

AutoMeKin2021: An open-source program for automated reaction discovery

Emilio Martínez-Núñez¹  | George L. Barnes² | David R. Glowacki³ |
Sabine Kopec⁴ | Daniel Peláez⁴  | Aurelio Rodríguez⁵ |
Roberto Rodríguez-Fernández¹ | Robin J. Shannon³ | James J. P. Stewart⁶ |
Pablo G. Tahoces⁷ | Saulo A. Vazquez¹

¹Department of Physical Chemistry, University of Santiago de Compostela, Santiago de Compostela, Spain

²Department of Chemistry and Biochemistry, Siena College, Loudonville, New York, USA

³Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol, UK

⁴Institut de Sciences Moléculaires d'Orsay, UMR 8214, Université Paris-Sud - Université Paris-Saclay, Orsay, France

⁵Galicia Supercomputing Center (CESGA), Santiago de Compostela, Spain

⁶Stewart Computational Chemistry, Colorado Springs, Colorado, USA

⁷Department of Electronics and Computer Science, University of Santiago de Compostela, Santiago de Compostela, Spain

Correspondence

Emilio Martínez-Núñez, Department of Physical Chemistry, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain.

Email: emilio.nunez@usc.es

Funding information

Ministerio de Ciencia e Innovación, Grant/Award Number: PID2019-107307RB-I00; National Science Foundation, Grant/Award Number: 1763652

Abstract

AutoMeKin2021 is an updated version of tsscads2018, a program for the automated discovery of reaction mechanisms (*J. Comput. Chem.* **2018**, *39*, 1922). This release features a number of new capabilities: rare-event molecular dynamics simulations to enhance reaction discovery, extension of the original search algorithm to study van der Waals complexes, use of chemical knowledge, a new search algorithm based on bond-order time series analysis, statistics of the chemical reaction networks, a web application to submit jobs, and other features. The source code, manual, installation instructions and the website link are available at: <https://rxnkin.usc.es/index.php/AutoMeKin>

KEYWORDS

graph theory, kinetics, MD simulations, reaction mechanisms

1 | INTRODUCTION

Over the last several years, computational chemistry has witnessed a surge in the development of methods for reaction mechanism discovery.^{1–65} Many of these methods predict complex reaction networks in an automated manner, where the search of reactions is usually more thorough than the traditional “by hand” approach.

Our group is actively involved in this endeavor, and a few years ago we presented a new automated method called Transition State Search using Chemical Dynamics Simulations (TSSCDS).^{44,45} Our algorithm relied on a molecular dynamics (MD)-based exploration of configurational space, followed by a post-processing analysis to locate promising transition state (TS) candidates from the MD snapshots.^{43–45} While other methods also use the ability of MD simulations to discover reaction mechanisms, the distinctive feature of our approach is the focus on

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *Journal of Computational Chemistry* published by Wiley Periodicals LLC.

finding saddle-point structures. Up until now, locating TSs from MD simulations has been difficult, but the procedure described here has proven to be very effective and useful in predicting unexpected mechanisms. Our approach has been applied in combustion chemistry,^{66,67} cycloaddition reactions,⁶⁸ photodissociations,^{69–71} organometallic catalysis,⁴³ radiation damage of biological systems,⁷² simulation of mass spectrometry experiments,^{73,74} and other applications.^{75–77}

The first version of the computer program implementing our approach was released 3 years ago under the name *tsscds2018*.⁴⁶ This approach, along with the algorithms described in this paper, has now been combined and implemented in the software package named *AutoMeKin*,⁷⁸ which stands for *Automated Mechanisms and Kinetics*.

AutoMeKin2021 includes the following new features: (a) the rare-event acceleration method boxed molecular dynamics in energy space (BXDE)^{76,79}; (b) a generalization of a graph theory-based algorithm to locate TS structures for the study of non-covalent interactions⁸⁰; (c) a chemical knowledge-based method for reaction discovery; (d) a new TS search algorithm based on a bond-order time series analysis⁸¹; (e) a statistical analysis of the chemical reaction networks using the Python library *NetworkX*⁸²; (f) a web application for online job submission; as well as other features.

After a brief introduction of the original method, a description will be given of the new methods incorporated into *AutoMeKin2021*, as well as some test cases and sample input files. Its new capabilities, as well as some proposed future improvements, will be summarized in the conclusions.

2 | METHODS

AutoMeKin's main components are:

1. Short-time reactive MD simulations
2. Post-processing analysis of the MD simulations
3. Kinetics simulations

To run the MD simulations, a sizeable amount of vibrational energy is adaptatively placed in each vibrational mode, to trigger reactive events. Also, as described below, a new rare-event acceleration technique is available, which allows for an efficient sampling of reactive events by imposing a bias on the potential energy.

In the analysis of the MD trajectories, some concepts from Graph Theory are useful, including the Adjacency and Laplacian matrices and the SPRINT coordinates.⁴⁵ These are used in locating suitable TS guesses and in constructing the chemical reaction network. Specifically, the MD snapshots are screened to find TS candidates associated with a reactive event. This is accomplished by transforming a 3D molecular geometry into a graph, which is defined by its adjacency matrix \mathbf{A} , whose elements, a_{ij} , are given by⁴⁵:

$$a_{ij} = \begin{cases} 1 & \text{if } \delta_{ij} < 1 \\ 0 & \text{otherwise} \end{cases}, \text{ with } \delta_{ij} = \frac{r_{ij}}{r_{ij}^{\text{ref}}}, \quad (1)$$

where r_{ij} and r_{ij}^{ref} are the interatomic and reference distances, respectively, of each pair, ij , of atoms. Reference distances are determined from the sum of the covalent radii of atoms i and j .

A reactive event is then deemed to occur when for any atom j ⁴⁵:

$$\max(\delta_{jk}) > \min(\delta_{jl}), \quad (2)$$

where index k runs over the set of atoms that are covalently bonded to j (neighbors), and index l runs over the remaining (non-neighbor) atoms. In other words, the criterion of Equation (2) is met when the nearest atom to j is a non-neighbor. Since more than two bonds can be part of the reaction coordinate in a given TS, reactive events occurring within an adaptive time window of ~ 10 – 20 fs are merged.⁴⁵ The resulting structures are first subjected to a partial relaxation, with the atoms involved in the reactive event kept frozen, and then optimized to a TS (saddle point of index one). This search algorithm is named *bbfs*, which stands for bond breaking/formation search.

For the sake of efficiency, the trajectories are integrated with either MOPAC2016⁸³ or Entos Qcore⁸⁴ at a semiempirical quantum-mechanical (SQM) level, while the stationary points are re-optimized with Gaussian09⁸⁵ or Entos Qcore⁸⁴ using a higher level of electronic structure theory. More details about the method can be found in the original papers.^{44,45}

Table 1 shows a summary of the most important tools that have been implemented in the last version of *AutoMeKin*. They will be described in the next sections.

2.1 | Rare-event acceleration method BXDE

Standard MD simulations are typically biased toward the entropically favored reaction pathways. *AutoMeKin*'s standard MD module employs initial conditions with substantial amounts of vibrational energy to accelerate the incidence of reactive events.

An alternative way of accelerating reactive events, called BXDE⁷⁹ has recently been proposed. BXDE belongs to the family of BXD methods,^{86–89} which introduce reflective barriers in the phase space of an MD trajectory along a particular (collective) variable. The boundaries are employed to push the dynamics along the collective variable into regions of phase space which would rarely be sampled in an unbiased trajectory.

In BXDE, the bias is introduced into the potential energy, rather than in any particular collective variable of the system. The different chemical reaction channels are sampled by gradually scanning through potential energy “boxes” or energetic “windows.” The BXDE simulation module in *AutoMeKin* utilizes the Atomistic Simulation Environment (ASE) package, and MOPAC2016 or Entos Qcore is interfaced via the ASE calculator class.⁹⁰

Figure 1 shows the section of an input file where a BXDE calculation is requested. Each line in the input file consists of a *keyword value* pair. BXDE is one of the different sampling methods employed in *AutoMeKin* for finding stationary states in a potential energy surface, with other sampling alternatives such as: MD, MD-micro, external, Chemical Knowledge (ChemKnow), association, and van der Waals (vdW). Although some of the options are described in this work, the reader is referred to the documentation for more details.⁷⁸

Method ^a	Features	Dependencies	Ref
BXDE	Accelerated MD simulation	ASE	76,79
vdW	Sampling vdW structures	ASE	80
ChemKnow	Graph transformations and NEB	ASE NetworkX	This work
bots	Reactive event search algorithm	-	81
Reaction network properties	Graph-Theory-based statistics	NetworkX	76
Web application	Online submission of jobs	-	This work

Abbreviations: ASE, Atomistic Simulation Environment; BXDE, boxed molecular dynamics in energy space; vdW, van der Waals.

^aName of the tool/method.

TABLE 1 Main tools available in AutoMeKin2021

```
--Method--
sampling      BXDE
ntraj         1
fs            5000
fric          0.5
post_proc     bbfs 20 1
temp          1000
```

FIGURE 1 Section of an input file for a boxed molecular dynamics in energy space (BXDE) calculation

The number of trajectories and simulation time are specified through the keywords *fs* and *ntraj*, respectively. BXDE employs a Langevin thermostat whose friction coefficient in ps^{-1} (keyword *fric*) and temperature in K (keyword *temp*) must be entered as well.

An example of the use of BXDE combined with AutoMeKin's TS search algorithms is the recent study of the ozonolysis of α -pinene.⁷⁶ This reaction is known to follow the "Criegee mechanism" of alkene ozonolysis (see Figure 2), and was previously studied using ab initio methods.⁹¹ Despite the low level of electronic structure (PM7) employed to run BXDE and to optimize the stationary points,⁷⁶ not only was the new approach capable of predicting the major pathways (shown in black in Figure 2), but a significant number of new intermediates and pathways were also predicted. The figure also shows (in red) some of the most important pathways that were overlooked in the previous ab initio study and found in the BXDE sampling.⁷⁶ A full account of the new mechanisms predicted by BXDE is detailed elsewhere.⁷⁶

2.2 | Non-covalent interactions (van der Waals)

The original search algorithm relies on the adjacency matrix of Equation (1), where the reference distance is determined from the covalent radii of the atoms. Consequently, a complex where two molecules are held together by intermolecular interactions would not be regarded as a single entity, but as two separated fragments.

To expand the scope of the method, it was recently suggested that matrix **A** should be recast in a block structure that accounts for a system made up of molecules B and C⁸⁰:

$$\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{BC} \\ \mathbf{BC} & \mathbf{C} \end{pmatrix} \quad (3)$$

where the diagonal blocks **B** and **C** refer to the (covalent) connectivity within B and C, respectively, whereas the off-diagonal **BC** block corresponds to the non-covalent, that is, vdW, interacting system B–C. The matrix elements for **B** and **C** are evaluated according to Equation (1), with the reference distances determined from the covalent radii. In contrast, the matrix elements of the **BC** block are calculated using Equation (1) but with the reference distances determined from van der Waals radii.⁹²

In this new *ansatz*, non-covalent interactions in B–C are treated on the same footing as covalent ones within any of the fragments, thus permitting the detection of TSs connecting vdW structures. The method can be easily extended to more than two interacting molecules, and has been recently applied to study the Ar–benzene, N₂–benzene, (H₂O)_n–benzene ($n = 1$ –3), and (NO₂–benzene)⁺ systems.⁸⁰

To prevent the inherent bias of standard high-energy MD simulations toward dissociation of the complex, the BXDE sampling option is automatically used when a vdW calculation is called for. Figure 3 shows an example input file for a vdW calculation on the pyrene + NO₂ system.^{93,94}

In this example, a total simulation time of 2 ps is employed. The electronic structure level of theory employed in this example is GFN1-xTB⁹⁵ (*xtb*) using the Entos Qcore program,⁹⁶ which is requested using the keyword *LowLevel* followed by the computer program (*qcore*) and the method (*xtb*).

To generate the starting structures for the dynamics the keywords *rotate* and *Nassoc* are used. The former has four values:

FIGURE 2 Major reaction mechanisms of the α -pinene ozonolysis, featuring the new reaction pathways found by boxed molecular dynamics in energy space (BXDE)⁷⁶ (red)

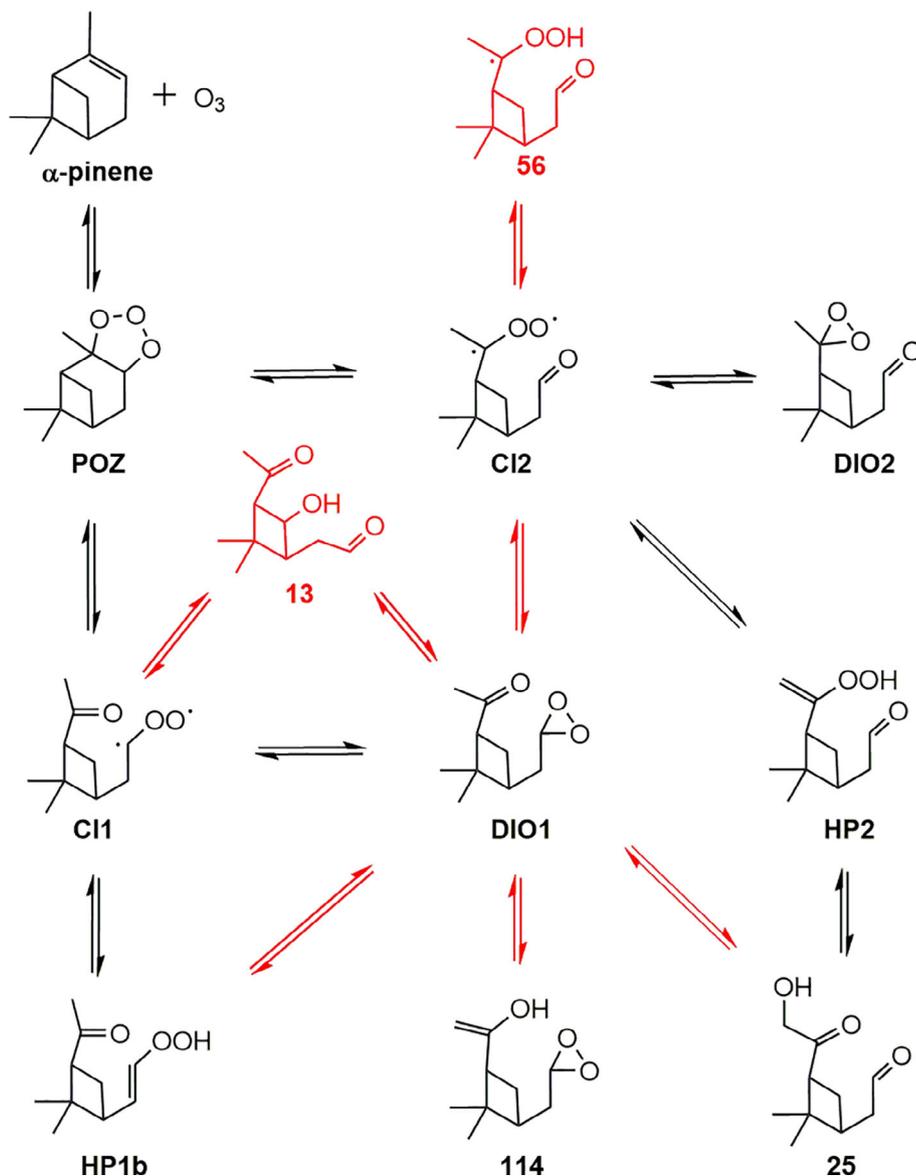


FIGURE 3 Two sections of an input file for a vdW calculation. The complete input file is in the Supporting Information

```
--General--
Molecule      pyrene-N02
fragmentA      pyrene
fragmentB      N02
LowLevel       qcore xtb

--Method--
sampling        vdW
rotate         com com 4.0 1.5
Nassoc         10
ntraj          1
fs             2000
```

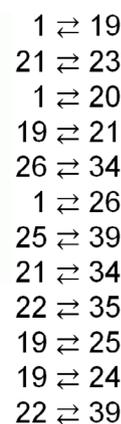
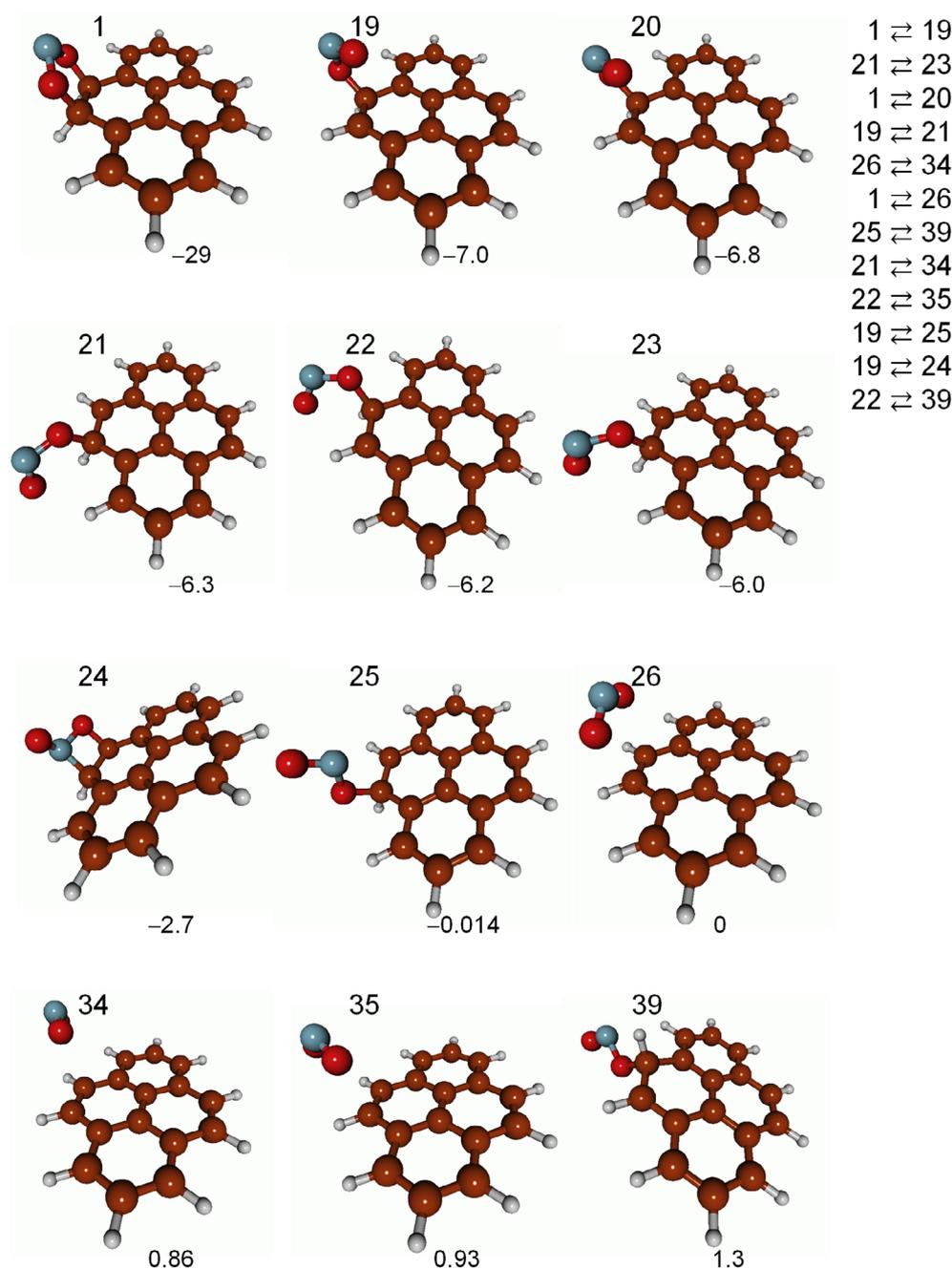


FIGURE 4 Example of a simple reaction network (using only three iterations of the vdW workflow; see text) obtained for the pyrene + NO₂ system featuring some covalent and non-covalent interactions. Numbers in the top left of each minimum energy structure are the labels, and the numbers at the bottom are their relative energies in kcal/mol

rotate pivotA pivotB r_pivot r_min

llcalcs.sh vdW.dat 100 3 48,

where *pivotA* and *pivotB* refer to the pivot points for the random rotations of fragments A and B, respectively (center of mass, *com*, for both). Then, *r_pivot* is the fixed distance between the pivot points (4.0 Å), and *r_min* is the minimum distance between any pair of atoms of different fragments (1.5 Å).

The keyword *Nassoc* is used to select the number of initial structures generated. The randomly generated structures are subjected to optimization using *xtb*. The global minimum from the set of optimized structures is then used as the starting point for the MD simulations.

In this example, only three iterations of AutoMeKin's workflow are employed:

where *llcalcs.sh* is AutoMeKin's script to run all low-level calculations,⁴⁶ *vdW.dat* is the input file (summarized in Figure 3), 100 is the number of BXDE trajectories per iteration, 3 is the number of iterations, and 48 is the number of concurrent simulations in a multithreading CPU.

This calculation results in a total of 112 minima and 115 TSs for the NO₂ + pyrene system. Ignoring the unconnected minima, the reaction network and its corresponding minimum energy structures are displayed in Figure 4.

In sum, the more general definition of the adjacency matrix of Equation (3) permits exploring both covalently and non-covalently

bound structures, as seen in Figure 4. In particular, nine of the 12 structures of Figure 4 present covalent bonds between NO₂ and pyrene, while three correspond to vdW structures. The geometries of all the structures, including those not connected to the network of Figure 4 (the vast majority), are collected in the Supporting Information.

2.3 | Chemical Knowledge

MD-based methods have shown great success exploring reaction mechanisms in small-medium sized systems. However, for larger ones, a thorough search of the full chemical reaction space with such approaches soon becomes intractable. One way to circumvent this is to focus only on those reactions which are relevant at the conditions of interest. With MD-based methods, this could be achieved simply by imposing upper limits to the energies of the TSs (an option available in AutoMeKin). We propose in this section an alternative method for this. The new approach, called ChemKnow, allows to impose a greater number of constraints in the search rendering the whole process more efficient than MD-based methods.

The potential gain in efficiency of ChemKnow comes from:

- (i) exploring only those reactions in which there is an interest,
- (ii) imposing limits for the minimum and maximum number of neighbors (valencies) of an atom, and (iii) restricting the maximum number of bonds that can break and form in each step.

The method also benefits from working in graph space, where reactions are simply graph transformations. This has been shown to be a practical way to explore reactive events by Habershon and co-workers.^{26,27,97–100}

The workflow of ChemKnow is detailed as follows:

1. The reactive sites (*active atoms*) of the system are selected, along with the maximum number of bonds that can break (n_b) and form (n_f) per elementary step, the allowed minimum (min_val) and maximum valencies (max_val) of each atom, and the maximum energy ($emax$) of the system.
2. Beginning with a given minimum energy geometry turned into its graph analogue, ChemKnow generates all possible graph transformations that comply with the constraints of the previous step. Three additional restrictions are imposed on top of those specified by the user: (1) reactions where the closest distance between the (linear) paths followed by the atoms in their rearrangement is lower than a threshold value are ignored, (2) bond formations between atoms that are at a distance greater than a certain value $startd$ are not allowed; and, (3) only those bonds whose bond order is lower than 1.5 can be broken.
3. The newly generated graphs are converted back into 3D structures using constrained Langevin dynamics, with external forces applied to the active atoms. The adjacency matrix is monitored along the trajectory, and the constraints are lifted once the product graph is obtained. At this point, the final geometry is

optimized, and its connectivity is checked once more to make sure that the desired product is obtained. As a second check, its energy must be lower than $emax$ to retain the newly generated geometry.

4. A path connecting the initial and final geometries is constructed using the nudged elastic band (NEB) method, and the highest point along the path is subjected to TS optimization.
5. Successive iterations of AutoMeKin start from a new (connected) graph, and steps 1–4 are repeated until no new graphs are found.

To avoid sampling equivalent paths multiple times in step 2, a descriptor for each unique TS is book-kept in a Python dictionary. The chosen descriptor is the list of eigenvalues of a TS adjacency matrix:

$$\mathbf{A}_{TS} = \frac{1}{2}(\mathbf{A}_R + \mathbf{A}_P) \quad (4)$$

where \mathbf{A}_R and \mathbf{A}_P are the reactant and product adjacency matrices, respectively, with the atomic numbers filling the diagonals. This descriptor has the property of being invariant with respect to permutations of like atoms, thus avoiding sampling equivalent paths more than once.

Additionally, steps 3 and 4 of the above pipeline are carried out using ASE's ExternalForces and AutoNEB classes, respectively,⁹⁰ with all the graph analysis and transformations performed using NetworkX.⁸²

By way of example, ChemKnow was employed to study the fragmentation channels of formic acid (FA) using the following constraints:

1. All atoms are *active*.
2. $n_f \leq 2$ and $n_b \leq 2$, where transformations with both $n_f = 2$ and $n_b = 2$ are discarded.
3. $min_val = [1,1,1]$ and $max_val = [1,4,2]$ for the list of atoms = [H,C,O], respectively.
4. $startd = 2.75 \text{ \AA}$
5. $emax = 150 \text{ kcal/mol}$.

When the initial structure is *cis*-FA, a total of 77 distinct (n_f, n_b) combinations are found, which breaks down into 3 (0,1); 3 (0,2); 5 (1,0); 15 (1,1); 15 (1,2); 9 (2,0); and 27 (2,1) combinations. This number excludes the cleavage of bonds with bond orders greater than 1.5. Additionally, when the other constraints were imposed, the number of combinations (paths) that start in *cis*-FA became 8.

Overall, this approach only needs four iterations of the workflow to reach convergence (no further minima found) for FA, which affords a total of 8 TSs and 4 minima at the PM7 SQM level after exploring a total of 24 paths. The CO₂:CO branching ratio obtained at 150 kcal/mol of excitation energy is 3:97. In comparison, an MD sampling with 200 trajectories leads to 11 TSs and 7 minima and a CO₂:CO ratio of 2:98.

A test was also done on vinyl cyanide (VC),⁷¹ to compare the performance of ChemKnow versus an MD-based sampling. Chemknow's

constraints are similar to those employed for FA, including now (2,2) combinations of (n_f, n_b) for the graph transformations, and min_val and max_val for N are 1 and 4, respectively. In this example, ChemKnow needs to sample 850 paths to obtain 59 TSs and 31 minima versus 2000 trajectories employed by the MD module, which affords 64 TSs and 27 minima.

On the other hand, if the set of active atoms and range of valences are reduced, ChemKnow probably outperforms MD-based methods in terms of efficiency. However, the decision to employ this method should also rely on its efficacy in finding the relevant structures for the system under study.

It was noted in passing that, while the MD-based methods have been heavily tested, ChemKnow needs further assessment and perhaps an optimization of steps 3 and 4, which are the major components of the TS search algorithm.

2.4 | New TS search algorithm (bots)

Our previous version of AutoMeKin included only one algorithm for the detection of reactive events. The algorithm (named *bbfs*, see above) is based on monitoring geometries along the MD trajectories and therefore does not encode information on the subtle changes that bond orders experience in some chemical reactions. It is therefore desirable to implement new search algorithms based on bond orders rather than interatomic distances to identify possible flaws in *bbfs*.

Wang and co-workers⁸¹ have recently shown that, by doing a time series analysis of the trajectories based on pairwise bond orders, reactive events involving formation and/or breaking of chemical bonds can be readily detected. The basis of this approach is that reactive events correspond to peaks on the first time derivative of the bond order time series. Its fundamental assumption is that bond orders do not change slowly in a chemical reaction. This new method is called *bots*, which stands for bond-order time series, and this is the workflow⁸¹:

1. A low-pass filter is applied to remove the fast fluctuations from the time series using a cutoff frequency ω .
2. The first time derivative of the smoothed time series is obtained using the central difference formula.
3. A threshold μ is applied to the first-order derivatives to select the peaks (those above $+\mu$ and below $-\mu$).
4. As in *bbfs*, peaks within an adaptive time window⁴⁵ are merged and regarded as multi-bond reactive events.

Tests showed that this algorithm works best with BXDE because high-energy MD simulations do not typically give rise to high-

frequency fluctuations in the bond-order time series. Therefore, the use of *bots* is restricted to BXDE-based methods.

Figure 5 shows part of an input file to run a BXDE sampling, followed by *bots* analysis of the trajectories. The keyword *post_proc* is employed to select *bbfs* (the default) or *bots*. In the latter case, two parameters are required: ω , in cm^{-1} , and μ , given as a multiple of σ . By way of example, the simple test case FA is employed to run 100 BXDE trajectories, which were then analyzed using *bots* with the parameters in Figure 5. The resulting structures and kinetics are very similar to the ones obtained using the standard *bbfs* method.

Figure 6 shows the variation of three bond orders and their time derivatives for a reactive BXDE simulation leading to $\text{H}_2\text{O} + \text{CO}$. In

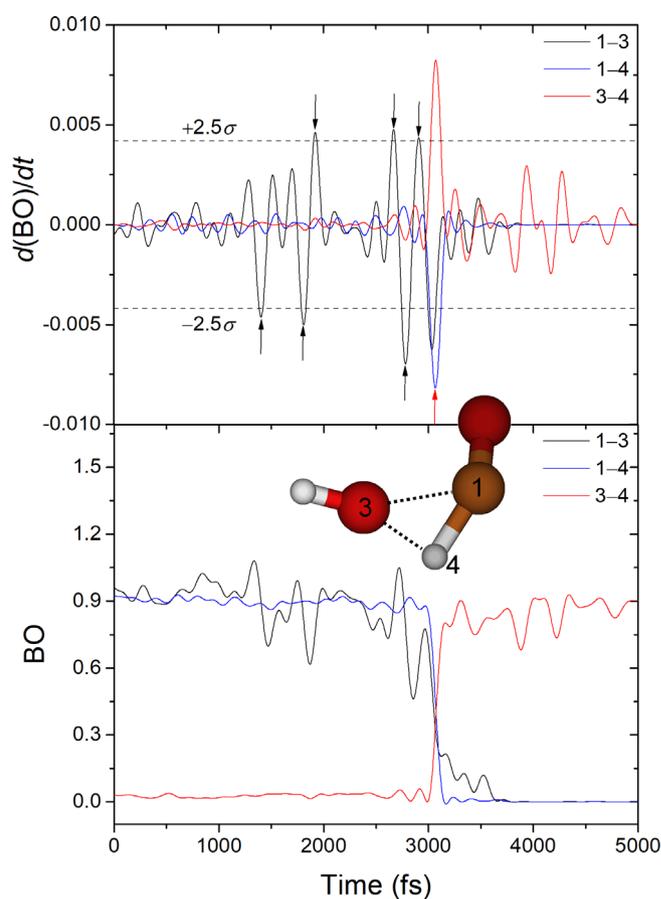


FIGURE 6 Bond orders (bottom) and their time derivative (top) for a reactive boxed molecular dynamics in energy space (BXDE) trajectory starting from *cis*-FA. The reactive event occurs at 3.1 ps and is successfully detected by *bots* (red arrow in the top panel). Black arrows correspond to false positives

```
--Method--
sampling      BXDE
ntraj         1
post_proc     bots 200 2.5
```

FIGURE 5 One section of an input file that employs *bots* TS search algorithm. The complete input file is in the Supporting Information

the figure, the three bonds correspond to those that change from reactant to products.

The use of *bots* requires fine-tuning two parameters (ω and μ) for the system under study. These values should be selected to find as many peaks (reactive events) as possible, while minimizing the number of false positives. Figure 6 shows that the only reactive event is successfully detected at ~ 3.1 ps, that is, the breaking of the 1–3 and 1–4 bonds and formation of the 3–4 bond. However, this comes at the cost of finding six false positives, at 1.4, 1.8, 1.9, 2.7, 2.8, and 2.9 ps. Although false positives also occur in *bbfs*, the major disadvantage of *bots* is its dependence on two parameters that strongly affect its performance.

2.5 | Properties of the reaction network

Complex reaction mechanisms can be represented as networks of reactions and studied by Graph Theory. A node of the network corresponds to either an intermediate or to a reaction product. Edges represent pathways connecting two nodes. Studying the structure of such reaction networks can be very useful to identify the number of steps between reactants and products or to spot the presence of important (highly connected) intermediates (or hubs) in the network.

In this section, statistical properties of the networks are presented seeking to answer two questions: (1) Do these networks exhibit “small-world” behavior? (2) are they scale-free? Small-world networks are defined as networks that present a short path length, even between distant nodes. In scale-free networks there exists a hierarchy of nodes, that is, a small number of them (called hubs) are highly connected and a large number of them present only a few connections.

In AutoMeKin2021, properties of the reaction networks are analyzed using the NetworkX Python library.¹⁰¹ Two types of networks are constructed in the example shown here. In the “all-states” network, every single intermediate constitutes a node, while in the “coarse-grained” one a family of conformers is lumped together to form each node. Additionally, in both networks, edges have weights representing the number of pathways connecting a pair of nodes. Finally, self-loops are avoided by removing paths connecting permutation-inversion isomers of the same node.

Common properties of a network can be studied in this new version of the program: the average shortest path length, the average clustering coefficient, the transitivity, and the assortativity. Some example systems that have been described using similar approaches are the network of organic chemistry,¹⁰² the network involved in the ozonolysis of α -pinene,⁷⁶ and a network of small clusters.¹⁰³ Below, we give a brief description of the properties provided in AutoMeKin's output.

The shortest path is the one connecting a pair of nodes through the least number of edges.¹⁰⁴ The clustering coefficient indicates the degree to which the neighbors of a node are also neighbors of each other, and an average clustering coefficient can be calculated for the network.¹⁰⁵ In turn, the transitivity is proportional to the ratio of the number of triangles over the number of triads in the network. For

any three nodes in the network, a triangle is formed when the three possible pairs of nodes are connected, while in a triad only two pairs are connected.

The assortativity is a measure of the tendency of nodes to have connections with nodes of a similar degree and can be measured through a coefficient¹⁰⁶ that varies from -1 to 1 . Values close to 1 indicate that nodes have a preference to connect with nodes of a similar degree, which is called assortative mixing, while values close to -1 indicate the opposite and is called disassortative mixing.

Figure 7 shows part of the coarse-grained network constructed from the results obtained with AutoMeKin for the decomposition of protonated uracil, which was chosen as an example. Table 2 shows the properties of the “all-states” and “coarse-grained” networks obtained from the same results.

In general, the networks of chemical reactions are sparsely connected⁷⁶ with a low density of edges (1%–2% in this example). An important feature of any network is whether they present small-world behavior, that is, when pairs of nodes are connected through a small number of edges. This property can be assessed by comparing the transitivity values and the average shortest path length with those of random networks. In this case, average shortest path lengths (5.49 and 3.89) are considerably shorter than those for the corresponding random networks, and the transivities are 1.7 times greater. These results point out a clear “small-world” behavior. Similar results of other chemical reaction networks can be found in the literature.^{76,102,103}

Clustering coefficients provide the proportion of interlinking between neighbors of a given node. The so-called scale-free networks are characterized by an enhanced clustering compared with a random network, just like the networks in this study (see Table 2).

Finally, the negative values of the assortativity indicate disassortativity mixing. That is, nodes of different degree tend to be connected. Disassortative mixing has also been observed in the ozonolysis of α -pinene,⁷⁶ in the network of organic chemistry,¹⁰² and in biological and technological networks.¹⁰⁷

The detailed reaction networks corresponding to the fragmentation of protonated uracil can be found in the Supporting Information.

2.6 | Web application

The web application is available at <https://rxnkin.usc.es/amk/> and works on most widely used web browsers. It is for demonstration and test purposes only. Therefore, reaction mechanisms and kinetics results are predicted at the PM7 SQM level of theory and the maximum number of atoms is limited to 15.

Figure 8 shows three screenshots of the most relevant sections of the web application. Briefly, users first need to register using a valid email account. Once this is done they can request a “New Job,” whose details need to be specified (Figure 8B): geometry of the system, charge, and the temperature or energy of the kinetics simulations.

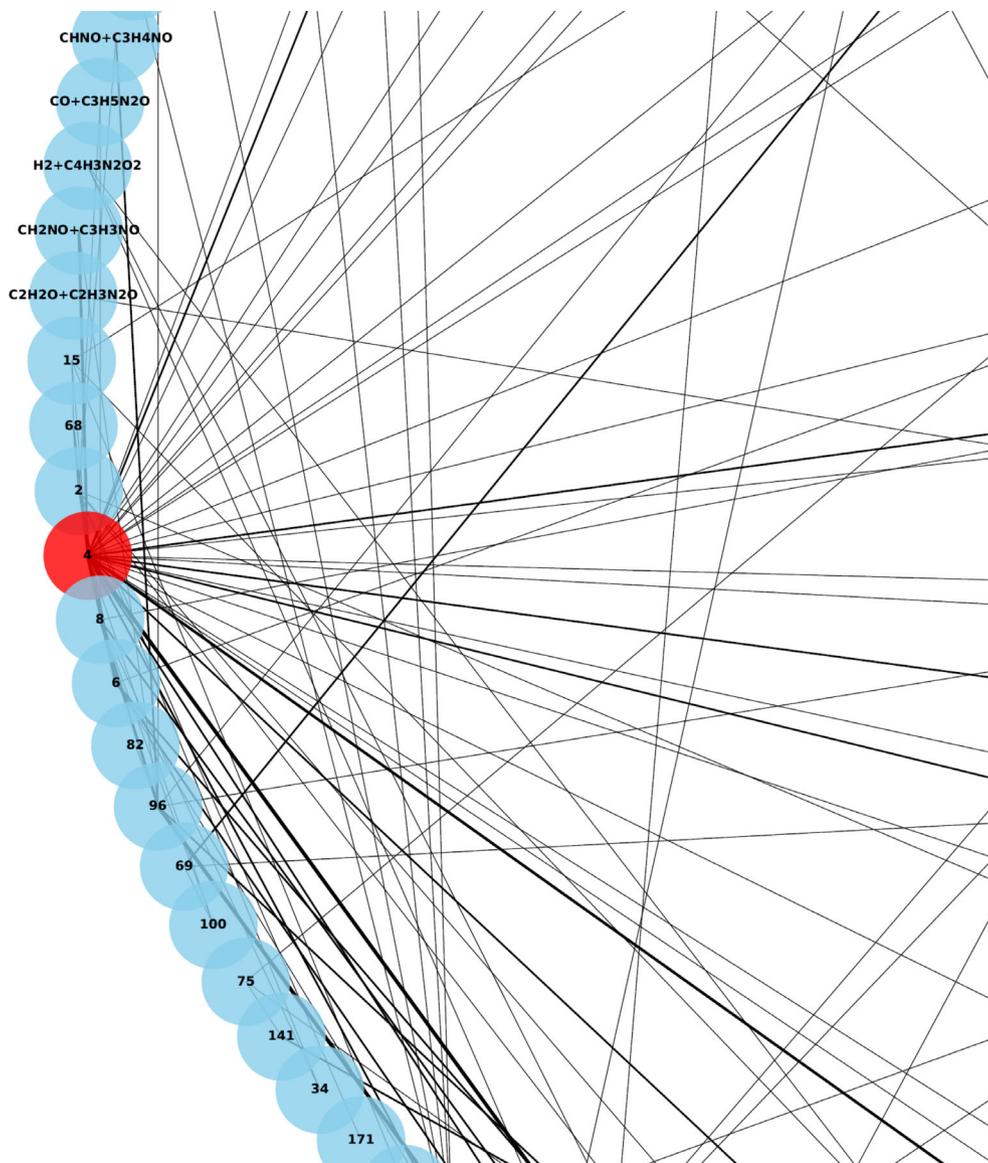


FIGURE 7 Part of the complex reaction network (in circular layout) involved in the fragmentation of protonated uracil. Nodes represent families of conformers (coarse-grained network), and the width of the edges is a measure of the number of paths between a pair of nodes. The red node corresponds to the starting structure

TABLE 2 Properties of the reaction networks^a

	All-states ^b	Coarse-grained ^c
Nodes	208	116
Edges	244	136
Density of edges (%) ^d	1.1	2.0
Average shortest path length	5.49 (0.5)	3.89 (0.4)
Average clustering coefficient	0.026 (3.5)	0.070 (5.0)
Transitivity	0.019 (1.7)	0.033 (1.7)
Assortativity	-0.024	-0.18

^aNumbers in parenthesis give the ratio of the value of the property over the corresponding value of a random (Erdős-Rényi) network with the same number of nodes and edges.

^bEvery structure is a node in the network.

^cFamilies of conformers form a node of the network.

^dPercentage of edges with respect to the maximum number of edges between the nodes of the network.

To input the geometries, a JSmol viewer integrated in our web interface was employed.¹⁰⁸ Once all parameters are specified, users can submit their jobs by clicking the “submit your job” button. The number of jobs is not limited, but users should try to limit the number of jobs they submit at once so that other users can also use the server at the same time.

The status of the jobs is shown on a different page (Figure 8C). Depending on the size of the system and workload, the execution time can vary substantially, and users can log out. Upon job completion, they will receive a notification in their email account.

Finished jobs appear with a “Completed” status in Figure 8C and users will be able to download a brief summary of the results in PDF format (“Report”), as well as a tarball file with detailed data (“Data”).

The web interface was built following HTML5 recommendations.¹⁰⁹ The Bootstrap framework¹¹⁰ is included to develop a responsive design. The Apache HTTP server, the MariaDB server, and PHP

FIGURE 8 Different screenshots of the web application featuring: (A) the front page, (B) the area employed to set up your calculation, and (C) the job queue

(A) **Automated Reaction Mechanisms and Kinetics**

Auto MeKin

MD simulations
Graph Theory
Kinetic Monte Carlo

Mechanism Structures Kinetics

AutoMeKin is an automated reaction discovery tool. This web server allows you to unravel the **reaction mechanisms** (with **intermediates** and **transition states**) of your system, just providing an equilibrium structure (reactant). As a bonus, you will also obtain **kinetics** results at the desired temperature or energy. Information on how to use this server can be found in **Tutorial**. Additionally, you can install and run AutoMeKin locally in your own machine, and instructions on how to do so are given in the **Wiki**.

Tutorial Wiki

(B) **Geometry:**

methanol

NCl(small molecules) Search

XYZ Geometry:

```
6
Model[1]: Jmol 14.29.28 2018-10-31 18:38
C 0.73700 -0.01470 0.00000
O -0.68960 0.06770 0.00000
H 1.07010 -0.54870 0.89000
H 1.07010 -0.54870 -0.89000
H 1.15900 0.99020 0.00000
H -1.13330 -0.79160 0.00000
```

Parameters:

Charge: 0

Kinetics: Temperature(K)

Value: 300

cancel submit your job

* Mandatory fields

Description:

your job description here

(C)

Inicio / Menu

Id	Description	Status	Report	Data	Delete
899	methanol	✓	📄	📄	🗑️

✓ Completed ⚙️ Running ⏸️ Waiting ❌ Error 🗑️ Deleted

New Job

are used to build the backend. A batch system written in Perl and C deals with the execution of each job, balancing the system workload, updating the status of the jobs, moving the results to the Apache download area and notifying the users upon job completion.

2.7 | Other improvements

This new version also includes the possibility of employing Entos Qcore⁸⁴ for the electronic structure calculations.

TS #	DE(kcal/mol)	Reaction path information			
====	=====	=====	=====	=====	=====
98	60.0	MIN	12	---->	PR5: CHNO + C3H4NO+
143	72.0	MIN	224	---->	PR6: C4H3N2O+ + H2O
144	72.2	MIN	50	---->	PR26: C4H4N2O2+ + H
300	98.4	MIN	10	---->	PR46: C3H5N2O+ + CO
369	110.5	PR6: C4H3N2O+ + H2O	<--->	PR38: C2HN2O+ + C2H2 + H2O	
396	114.0	MIN	6	---->	PR40: CH2NO+ + C3H3NO
421	118.7	MIN	216	---->	PR29: C4H3N2O2+ + H2
450	123.9	MIN	467	---->	PR22: CH2NO+ + CO + C2H3N
497	140.4	MIN	432	---->	PR12: CH4NO+ + C3HNO
504	143.4	PR8: C4H3N2O+ + H2O	<--->	PR52: C4H2N2O + H3O+	

FIGURE 9 Pathways involving different fragments for the high-level network of protonated uracil. A comprehensive list of all pathways is given in the Supporting Information

An example input file (FA_qcore.dat) employing this option for both low-level and high-level calculations can be found in the examples folder. The corresponding test can be run using:

```
run_test.sh -tests = FA_qcore
```

Briefly, these calculations can be requested through the keywords *LowLevel* and *HighLevel*. An example of a low-level calculation using Entos Qcore is also given above for the pyrene + NO₂ system. For the high-level calculations, the syntax is:

```
HighLevel qcore qcore_template
```

where *qcore_template* is the name of a file that contains the instructions to carry out the Entos Qcore high-level calculations:

```
dft (
xc=PBE
ao='6-31G*'
)
```

Since IRC calculations are not available in Entos Qcore, a damped velocity Verlet algorithm¹¹¹ is utilized to follow the reaction pathways.

A useful feature to study the decomposition of ions is the assignment of charges (and multiplicities) of the resulting fragments. This option is only available for high-level calculations using Gaussian.⁸⁵ Charges and multiplicities are assigned using the keyword *pop = (mk,nb)*. This keyword is added to a single point calculation for the geometry of the last point of an IRC leading to fragmentation.

As an example, Figure 9 shows part of the high-level pathways obtained for the decomposition of protonated uracil at the B3LYP/6-31+G(d,p) level of theory. The figure displays only a reduced part of the pathways (those involving different fragments) with the positive charges assigned to the corresponding fragments.

Since the reaction detection algorithms focus on bond formation/breakage, there is an additional tool to scan dihedral angles and find TSs for interconverting conformers. Torsions around bonds with bond orders greater than 2.0 and/or those belonging to rings are excluded.

Finally, an auto-installer script is now available, which eases the burden of installing third-party packages. The script installs singularity¹¹² and downloads the latest container image from sylabs (<https://sylabs.io/>). An instance of the container is started using a sandbox image deployed under \$(TMPDIR-/tmp) folder. The container comes with all AutoMeKin's tools installed in \$AMK.

3 | CONCLUSIONS

Presented here is the open-source software package AutoMeKin, for automated reaction discovery. AutoMeKin is an updated version of tsscds2018 featuring several new tools: rare-event MD simulations, a search algorithm to study van der Waals complexes, a chemical-knowledge based search procedure, a reactive-event detection method based on bond orders, statistics of the chemical reaction networks, and a web application to submit online jobs.

AutoMeKin is actively developed, and the most relevant functionalities that will be incorporated in the future include (but are not limited to):

1. An interface with M3C¹¹³ to study fragmentation of vibrationally excited molecules including barrierless mechanisms.
2. A deep-learning correction to SQM barrier heights to boost the performance and efficiency of the calculations.¹¹⁴
3. An interface with Pilgrim,¹¹⁵ a code to calculate thermal rate constants of chemical reactions including variational and tunneling effects.

ACKNOWLEDGMENTS

This work was partially supported by the Ministerio de Ciencia e Innovación (Grant # PID2019-107307RB-I00). G. L. B. gratefully

acknowledges support from the National Science Foundation under grant No. 1763652.

ORCID

Emilio Martínez-Núñez  <https://orcid.org/0000-0001-6221-4977>

Daniel Peláez  <https://orcid.org/0000-0003-3924-7804>

REFERENCES

- [1] H. B. Schlegel, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 790.
- [2] H. L. Davis, D. J. Wales, R. S. Berry, *J. Chem. Phys.* **1990**, *92*, 4308.
- [3] J. Q. Sun, K. Ruedenberg, *J. Chem. Phys.* **1993**, *98*, 9707.
- [4] C. J. Tsai, K. D. Jordan, *J. Phys. Chem.* **1993**, *97*, 11227.
- [5] Y. Abashkin, N. Russo, *J. Chem. Phys.* **1994**, *100*, 4477.
- [6] K. Bondensgard, F. Jensen, *J. Chem. Phys.* **1996**, *104*, 8025.
- [7] J. P. K. Doye, D. J. Wales, *Z. Phys. D* **1997**, *40*, 194.
- [8] W. Quapp, M. Hirsch, O. Imig, D. Heidrich, *J. Comput. Chem.* **1998**, *19*, 1087.
- [9] M. Černohorský, S. Kettou, J. Koča, *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 705.
- [10] K. M. Westerberg, C. A. Floudas, *J. Chem. Phys.* **1999**, *110*, 9259.
- [11] D. J. Wales, J. P. Doye, M. A. Miller, P. N. Mortenson, T. R. Walsh, *Adv. Chem. Phys.* **2000**, *115*, 1.
- [12] K. K. Irikura, R. D. Johnson, *J. Phys. Chem. A* **2000**, *104*, 2191.
- [13] E. M. Müller, A. D. Meijere, H. Grubmüller, *J. Chem. Phys.* **2002**, *116*, 897.
- [14] M. Dallos, H. Lischka, E. Ventura Do Monte, M. Hirsch, W. Quapp, *J. Comput. Chem.* **2002**, *23*, 576.
- [15] J. Baker, K. Wolinski, *J. Comput. Chem.* **2011**, *32*, 43.
- [16] P. M. Zimmerman, *J. Comput. Chem.* **2013**, *34*, 1385.
- [17] P. M. Zimmerman, *J. Chem. Phys.* **2013**, *138*, 184102.
- [18] P. M. Zimmerman, *J. Chem. Theory Comput.* **2013**, *9*, 3043.
- [19] P. M. Zimmerman, *J. Comput. Chem.* **2015**, *36*, 601.
- [20] P. M. Zimmerman, *Mol. Simul.* **2015**, *41*, 43.
- [21] M. Jafari, P. M. Zimmerman, *J. Comput. Chem.* **2017**, *38*, 645.
- [22] A. L. Dewyer, P. M. Zimmerman, *Org. Biomol. Chem.* **2017**, *15*, 501.
- [23] D. Rappoport, C. J. Galvin, D. Y. Zubarev, A. Aspuru-Guzik, *J. Chem. Theory Comput.* **2014**, *10*, 897.
- [24] B. Schaefer, S. Mohr, M. Amsler, S. Goedecker, *J. Chem. Phys.* **2014**, *140*, 214102.
- [25] D. J. Wales, *J. Chem. Phys.* **2015**, *142*, 130901.
- [26] S. Habershon, *J. Chem. Phys.* **2015**, *143*, 094106.
- [27] S. Habershon, *J. Chem. Theory Comput.* **2016**, *12*, 1786.
- [28] X.-J. Zhang, Z.-P. Liu, *Phys. Chem. Chem. Phys.* **2015**, *17*, 2757.
- [29] L.-P. Wang, R. T. McGibbon, V. S. Pande, T. J. Martinez, *J. Chem. Theory Comput.* **2016**, *12*, 638.
- [30] L.-P. Wang, A. Titov, R. McGibbon, F. Liu, V. S. Pande, T. J. Martinez, *Nat. Chem.* **2014**, *6*, 1044.
- [31] M. Yang, J. Zou, G. Wang, S. Li, *J. Phys. Chem. A* **2017**, *121*, 1351.
- [32] L. D. Jacobson, A. D. Bochevarov, M. A. Watson, T. F. Hughes, D. Rinaldo, S. Ehrlich, T. B. Steinbrecher, S. Vaitheeswaran, D. M. Philipp, M. D. Halls, R. A. Friesner, *J. Chem. Theory Comput.* **2017**, *13*, 5780.
- [33] K. Ohno, S. Maeda, *Chem. Phys. Lett.* **2004**, *384*, 277.
- [34] S. Maeda, K. Ohno, *J. Phys. Chem. A* **2005**, *109*, 5742.
- [35] K. Ohno, S. Maeda, *J. Phys. Chem. A* **2006**, *110*, 8933.
- [36] K. Ohno, S. Maeda, *Phys. Scr.* **2008**, *78*, 058122.
- [37] S. Maeda, K. Morokuma, *J. Chem. Phys.* **2010**, *132*, 241102.
- [38] S. Maeda, K. Morokuma, *J. Chem. Theory Comput.* **2011**, *7*, 2335.
- [39] S. Maeda, K. Ohno, K. Morokuma, *Phys. Chem. Chem. Phys.* **2013**, *15*, 3683.
- [40] S. Maeda, T. Taketsugu, K. Morokuma, *J. Comput. Chem.* **2014**, *35*, 166.
- [41] S. Maeda, Y. Harabuchi, M. Takagi, T. Taketsugu, K. Morokuma, *Chem. Rec.* **2016**, *16*, 2232.
- [42] S. Maeda, Y. Harabuchi, M. Takagi, K. Saita, K. Suzuki, T. Ichino, Y. Sumiya, K. Sugiyama, Y. Ono, *J. Comput. Chem.* **2017**, *39*, 233.
- [43] J. A. Varela, S. A. Vazquez, E. Martinez-Nunez, *Chem. Sci.* **2017**, *8*, 3843.
- [44] E. Martínez-Núñez, *Phys. Chem. Chem. Phys.* **2015**, *17*, 14912.
- [45] E. Martínez-Núñez, *J. Comput. Chem.* **2015**, *36*, 222.
- [46] A. Rodríguez, R. Rodríguez-Fernández, A. S. Vázquez, L. G. Barnes, J. P. Stewart, E. Martínez-Núñez, *J. Comput. Chem.* **2018**, *39*, 1922.
- [47] S. A. Vazquez, X. L. Otero, E. Martinez-Nunez, *Molecules* **2018**, *23*, 3156.
- [48] L. J. Broadbelt, S. M. Stark, M. T. Klein, *Ind. Eng. Chem. Res.* **1994**, *33*, 790.
- [49] D. M. Matheu, A. M. Dean, J. M. Grenda, W. H. Green, *J. Phys. Chem. A* **2003**, *107*, 8552.
- [50] C. W. Gao, J. W. Allen, W. H. Green, R. H. West, *Comput. Phys. Commun.* **2016**, *203*, 212.
- [51] P. L. Bhoorasingh, R. H. West, *Phys. Chem. Chem. Phys.* **2015**, *17*, 32173.
- [52] P. L. Bhoorasingh, B. L. Slakman, F. Seyedzadeh Khanshan, J. Y. Cain, R. H. West, *J. Phys. Chem. A* **2017**, *121*, 6896.
- [53] Y. V. Suleimanov, W. H. Green, *J. Chem. Theory Comput.* **2015**, *11*, 4248.
- [54] M. Bergeler, G. N. Simm, J. Proppe, M. Reiher, *J. Chem. Theory Comput.* **2015**, *11*, 5712.
- [55] J. Proppe, T. Husch, G. N. Simm, M. Reiher, *Faraday Discuss.* **2016**, *195*, 497.
- [56] G. N. Simm, M. Reiher, *J. Chem. Theor. Comput.* **2017**, *13*, 6108.
- [57] G. N. Simm, M. Reiher, *J. Chem. Theor. Comput.* **2018**, *14*, 5238.
- [58] G. N. Simm, A. C. Vaucher, M. Reiher, *J. Phys. Chem. A* **2019**, *123*, 385.
- [59] A. L. Dewyer, A. J. Argüelles, P. M. Zimmerman, *WIREs Comput. Mol. Sci.* **2018**, *8*, e1354. <https://doi.org/10.1002/wcms.1354>
- [60] Y. Kim, J. W. Kim, Z. Kim, W. Y. Kim, *Chem. Sci.* **2018**, *9*, 825.
- [61] J. W. Kim, Y. Kim, K. Y. Baek, K. Lee, W. Y. Kim, *J. Phys. Chem. A* **2019**, *123*, 4796.
- [62] M. Kawano, S. Koido, T. Nakatomi, Y. Watabe, T. Takayanagi, *Comput. Theor. Chem.* **2019**, *1155*, 31.
- [63] C. A. Grambow, A. Jamal, Y.-P. Li, W. H. Green, J. Zádor, Y. V. Suleimanov, *J. Am. Chem. Soc.* **2018**, *140*, 1035.
- [64] R. Van de Vijver, J. Zádor, *Comput. Phys. Commun.* **2020**, *248*, 106947.
- [65] T. A. Young, J. J. Silcock, A. J. Sterling, F. Duarte, *Angew. Chem., Int. Ed.* **2021**, *60*, 4266.
- [66] D. Ferro-Costas, E. Martínez-Núñez, J. Rodríguez-Otero, E. Cabaleiro-Lago, C. M. Estévez, B. Fernández, A. Fernández-Ramos, S. A. Vázquez, *J. Phys. Chem. A* **2018**, *122*, 4790.
- [67] Y. Fenard, A. Gil, G. Vanhove, H.-H. Carstensen, K. M. Van Geem, P. R. Westmoreland, O. Herbinet, F. Battin-Leclerc, *Combust. Flame* **2018**, *191*, 252.
- [68] A. Esteban, I. Izquierdo, N. García, M. J. Sexmero, N. M. Garrido, I. S. Marcos, F. Sanz, P. G. Jambriña, P. Ortega, D. Diez, *Tetrahedron* **2020**, *76*, 130764.
- [69] M. J. Wilhelm, E. Martínez-Núñez, J. González-Vázquez, S. A. Vázquez, J. M. Smith, H.-L. Dai, *ApJ* **2017**, *849*, 15.
- [70] R. Perez-Soto, S. A. Vazquez, E. Martinez-Nunez, *Phys. Chem. Chem. Phys.* **2016**, *18*, 5019.
- [71] S. A. Vazquez, E. Martinez-Nunez, *Phys. Chem. Chem. Phys.* **2015**, *17*, 6948.
- [72] F. F. da Silva, T. Cunha, A. Rebelo, A. Gil, M. J. Calhorda, G. García, O. Ingólfsson, P. Limão-Vieira, *J. Phys. Chem. A* **2021**, *125*, 2324.
- [73] E. Rossich Molina, J.-Y. Salpin, R. Spezia, E. Martinez-Nunez, *Phys. Chem. Chem. Phys.* **2016**, *18*, 14980.

- [74] V. Macaluso, D. Scuderì, M. E. Crestoni, S. Fornarini, D. Corinti, E. Dalloz, E. Martínez-Núñez, W. L. Hase, R. Spezia, *J. Phys. Chem. A* **2019**, *123*, 3685.
- [75] L. Song, C. Zhang, C.-G. Sun, B.-C. Hu, X.-H. Ju, *Monatsh. Chem.* **2021**, *152*, 421.
- [76] R. A. Jara-Toro, G. A. Pino, D. R. Glowacki, R. J. Shannon, E. Martínez-Núñez, *ChemSystemsChem* **2020**, *2*, e1900024. <https://doi.org/10.1002/syst.201900024>
- [77] E. Grajales-González, M. Monge-Palacios, S. M. Sarathy, *J. CO2 Util.* **2021**, *49*, 101554.
- [78] AutoMeKin's wiki, <https://rxnkin.usc.es/index.php/AutoMeKin> (2020). [accessed on 07 May 2017].
- [79] R. J. Shannon, S. Amabilino, M. O'Connor, D. V. Shalashilin, D. R. Glowacki, *J. Chem. Theor. Comput.* **2018**, *14*, 4541.
- [80] S. Kopec, E. Martínez-Núñez, J. Soto, D. Peláez, *Int. J. Quantum Chem.* **2019**, *119*, e26008.
- [81] M. Hutchings, J. Liu, Y. Qiu, C. Song, L.-P. Wang, *J. Chem. Theory Comput.* **2020**, *16*, 1606.
- [82] A. A. Hagberg, D. A. Shult, P. J. Swart, 7th Python in Science Conference (SciPy2008), Pasadena, CA **2008**, pp. 11–15.
- [83] J. J. P. Stewart *Stewart Computational Chemistry: Colorado Springs, CO*. <http://OpenMOPAC.net> (2016). [accessed on 07 May 2017].
- [84] M. Frederick, M. Thomas, B. Peter, D. Feizhi, D. Thomas, B.-R. Fidel, B. Alexander, B. Callum, L. Sebastian, M. Rocco, M. Kaito, S. Casper, T. Takashi, W. Matthew, W. Timothy, W. Zack, *ChemRxiv* **2019**. <https://doi.org/10.26434/chemrxiv.7762646.v2>.
- [85] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. N. Petersson, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, *Gaussian09*, Gaussian Inc., Wallingford, CT **2009**.
- [86] D. R. Glowacki, E. Paci, D. V. Shalashilin, *J. Chem. Theory Comput.* **2011**, *7*, 1244.
- [87] J. Booth, S. Vazquez, E. Martínez-Núñez, A. Marks, J. Rodgers, D. R. Glowacki, D. V. Shalashilin, *Philos. Trans. R Soc. A* **2014**, *372*, 20130384.
- [88] M. O'Connor, E. Paci, S. McIntosh-Smith, D. R. Glowacki, *Faraday Discuss.* **2016**, *195*, 395.
- [89] E. Martínez-Núñez, D. V. Shalashilin, *J. Chem. Theory Comput.* **2006**, *2*, 912.
- [90] A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dulak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng, K. W. Jacobsen, *J. Phys. Condens. Matter* **2017**, *29*, 273002.
- [91] D. Zhang, R. Zhang, *J. Chem. Phys.* **2005**, *122*, 114308.
- [92] M. Mantina, A. C. Chamberlin, R. Valero, C. J. Cramer, D. G. Truhlar, *J. Phys. Chem. A* **2009**, *113*, 5806.
- [93] R.-L. Panadés-Barrueta, *Ph.D. Thesis*, Université de Lille **2020**.
- [94] R.-L. Panadés-Barrueta, E. Martínez-Núñez, D. Peláez, in preparation.
- [95] S. Grimme, C. Bannwarth, P. Shushkov, *J. Chem. Theor. Comput.* **2017**, *13*, 1989.
- [96] F. Manby, T. Miller, P. Bygrave, F. Ding, T. Dresselhaus, F. Batista-Romero, A. Buccheri, C. Bungey, S. Lee, R. Meli, K. Miyamoto, C. Steinmann, T. Tsuchiya, M. Welborn, T. Wiles & Z. Williams *ChemRxiv Preprint*, <https://doi.org/10.26434/chemrxiv.7762646v2> (2019). [accessed on 07 May 2017].
- [97] C. Robertson, R. Hyland, A. J. D. Lacey, S. Havens, S. Habershon, *J. Chem. Theor. Comput.* **2021**, *17*, 2307.
- [98] I. Ismail, H. B. V. A. Stuttard-Fowler, C. Ochan Ashok, C. Robertson, S. Habershon, *J. Phys. Chem. A* **2019**, *123*, 3407.
- [99] C. Robertson, S. Habershon, *Catal. Sci. Technol.* **2019**, *9*, 6357.
- [100] C. Robertson, I. Ismail, S. Habershon, *ChemSystemsChem* **2020**, *2*, e1900047.
- [101] A. A. Hagberg, D. A. Shult, P. J. Swart, in *Proc. 7th Python in Sci. Conf. (SciPy2008)* (Eds: G. Varoquaux, T. Vaught, J. Millman), Pasadena, CA **2008**, p. 11. <https://networkx.org/documentation/networkx-2.1/citing.html>.
- [102] P.-M. Jacob, A. Lapkin, *React. Chem. Eng.* **2018**, *3*, 102.
- [103] J. W. R. Morgan, D. Mehta, D. J. Wales, *Phys. Chem. Chem. Phys.* **2017**, *19*, 25498.
- [104] E. Dijkstra, *Numer. Math.* **1959**, *1*, 83.
- [105] J. Saramäki, M. Kivelä, J.-P. Onnela, K. Kaski, J. Kertész, *Phys. Rev. E* **2007**, *75*, 027105.
- [106] M. E. J. Newman, *Phys. Rev. E* **2003**, *67*, 026126.
- [107] L. D. F. Costa, F. A. Rodrigues, G. Travieso, P. R. Villas Boas, *Adv. Phys.* **2007**, *56*, 167.
- [108] JSmol, http://wiki.jmol.org/index.php/Jmol_JavaScript_Object. [accessed on 07 May 2017].
- [109] HTML5, <https://www.w3.org/TR/html5/>. [accessed on 07 May 2017].
- [110] Bootstrap, <https://getbootstrap.com/>. [accessed on 07 May 2017].
- [111] H. P. Hratchian, H. B. Schlegel, *J. Phys. Chem. A* **2002**, *106*, 165.
- [112] G. M. Kurtzer, V. Sochat, M. W. Bauer, *PLoS One* **2017**, *12*, e0177459.
- [113] N. F. Aguirre, S. D-Tendero, P.-A. Hervieux, M. Alcamí, F. Martín, *J. Chem. Theor. Comput.* **2017**, *13*, 992.
- [114] C. A. Grambow, L. Pattanaik, W. H. Green, *J. Phys. Chem. Lett.* **2020**, *11*, 2992.
- [115] D. Ferro-Costas, D. G. Truhlar, A. Fernández-Ramos, *Comput. Phys. Commun.* **2020**, *256*, 107457.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: E. Martínez-Núñez, G. L. Barnes, D. R. Glowacki, S. Kopec, D. Peláez, A. Rodríguez, R. Rodríguez-Fernández, R. J. Shannon, J. J. P. Stewart, P. G. Tahoces, S. A. Vazquez, *J. Comput. Chem.* **2021**, *42*(28), 2036. <https://doi.org/10.1002/jcc.26734>