



MIT Open Access Articles

Unimolecular Reaction Pathways of a γ -Ketohydroperoxide from Combined Application of Automated Reaction Discovery Methods

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation	Grambow, Colin A. et al. "Unimolecular Reaction Pathways of a γ -Ketohydroperoxide from Combined Application of Automated Reaction Discovery Methods." <i>Journal of the American Chemical Society</i> 140, 3 (January 2018): 1035–1048 © 2017 American Chemical Society
As Published	http://dx.doi.org/10.1021/jacs.7b11009
Publisher	American Chemical Society (ACS)
Version	Author's final manuscript
Citable link	http://hdl.handle.net/1721.1/119819
Terms of Use	Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.

Unimolecular Reaction Pathways of a γ -Ketohydroperoxide from Combined Application of Automated Reaction Discovery Methods

Colin A. Grambow, Adeel Jamal, Yi-Pei Li, and William H. Green*

Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, Massachusetts 02139, United States

Judit Zádor

Combustion Research Facility, Sandia National Laboratories, 7011 East Ave., Livermore, California 94551, United States

Yury V. Suleimanov*

Computation-based Science and Technology Research Center, Cyprus Institute, 20 Kavafi Str., Nicosia 2121, Cyprus

Abstract

Ketohydroperoxides are important in liquid phase autoxidation and in gas phase partial oxidation and pre-ignition chemistry, but because of their low concentration, instability, and various analytical chemistry limitations, it has been challenging to experimentally determine their reactivity, and only a few pathways are known. In the present work, 75 elementary-step unimolecular reactions of the simplest γ -ketohydroperoxide, 3-hydroperoxypropanal, were discovered by a combination of density functional theory with several automated transition state search algorithms — the Berny algorithm coupled with the freezing string method (FSM), single- and double-ended growing string methods (SSM and GSM), the heuristic KinBot algorithm, and the single-component artificial force induced reaction method (SC-AFIR). The present joint approach significantly outperforms previous manual and automated transition state searches — 68 of the reactions of γ -ketohydroperoxide discovered here were previously unknown and completely unexpected. All methods found the lowest energy transition state, which corresponds to the first step of the Korcek mechanism, but each algorithm except for SC-AFIR detected several reactions not found by any of the other methods. We show that the low-barrier chemical reactions involve promising new chemistry that may be relevant in atmospheric and combustion systems. Our study highlights the complexity of chemical space exploration and the advantage of combined application of several approaches. Overall, the present work demonstrates both the power and the weaknesses of existing fully automated approaches for reaction discovery which suggest possible directions for further method development and assessment in order to enable reliable discovery of all important reactions of any specified reactant(s).

1. INTRODUCTION

For several decades, chemists have been trying to develop theoretical approaches to predict the reactivity of specific chemical compounds and to guide chemical discovery. Recent advances in electronic structure theory and high-performance computer technologies should make it possible to achieve this long-standing goal, and so achieve a much better understanding of systems where multiple reactions are occurring simultaneously. [1,2] Complex chemistry is common in combustion chemistry, [3] polymerization, [4] catalysis, [5,6,7] and environmental processes, [8] all of which are subject to continuous study due to their fundamental and industrial importance. A fundamental understanding of all chemical compounds and elementary reactions of a given chemical process can facilitate the design of more effective technologies. [9] In general, multiple competing reaction paths exist, which lead to a variety of products, especially if initial species are highly reactive (e.g., radicals, peroxides, and catalytic intermediates) or if a system is at high temperature. Historically, new reactions were discovered experimentally by the serendipitous detection of unexpected products. New reactions can also be discovered serendipitously on the computer. [10,11,12] Recently, Zimmerman has intentionally discovered several important new reactions using quantum chemistry. [13,14] However, existing search techniques are CPU-intensive and sole reliance on experiments to find new reactions is insufficient. With modern computational capabilities, it should become possible to discover all the important reaction pathways more reliably and more rapidly than is possible using experiments alone.

A potential energy surface (PES) is a multidimensional function of atomic coordinates that provides comprehensive information on all reaction paths. Its local minima correspond to reactants, intermediates, and products. These are generally easy to characterize due to simple chemical bonding rules (in most cases) making it relatively easy to predict their 3D geometries. Numerical optimization of these geometries is straightforward because the negative of the gradient along the PES always points downhill (i.e., it is a local minimization problem). [15] In fact, automated searches for local and global minima have already proved to be very successful. [16] A more challenging task is to detect and characterize first-order saddle points that connect local minima along minimum energy paths (MEPs), which are necessary to describe most transition states (TSs), the key regions of the PES to calculate reaction rate coefficients. Predicting geometries of TSs is more difficult. Numerical procedures for saddle points must step uphill in

one certain direction (the reaction coordinate, usually unknown *a priori*) and downhill in all other orthogonal directions — a more challenging task than finding a local minimum. [17] Typically, many low-energy saddle points exist in the vicinity of minima, which correspond to torsional rearrangements and lead to different conformations of a given minimum structure. Reactive TSs are mostly higher in energy and correspond to a change in bonding. Additionally, if one saddle point can be found that directly connects two structures in a reactive event, often many additional saddle points exist, several of them representing conformers of the TS. [18] Another concern is that the dimension of a PES ($3N-6$) increases with the number of atoms (N) in the system. Therefore, construction and subsequent global mapping of PESs have prohibitive computational costs even for reactive systems consisting of only 5–6 atoms. At present, most saddle points are found by human-guided exploration of reaction pathways. Human expectations and chemical intuition bias such an approach, usually limiting the search to expected reaction paths. The process is also slow and tedious. It is therefore highly desirable to develop efficient methods for automatically searching for unexpected TSs and corresponding reaction pathways on PESs.

Several effective methods for automatically searching reaction pathways with given reactant(s) have been proposed recently. Some of these methods are based on adding an external temperature/pressure control or artificial forces in the initial reactive system, which drive the reaction to occur in the direction of different products. These include, but are not limited to, metadynamics, [19] an *ab initio* nanoreactor [20], and the artificial-force-induced reaction (AFIR) method. [21] While the former two are based on molecular dynamics (MD) with special techniques to accelerate the evolution of the reactive system, the artificial force induced reaction (AFIR) strategy of the global reaction route mapping (GRRM) method utilizes specially designed minimization functions that are composed of the adiabatic PES and an artificial force term. In the case of a unimolecular initial reactant channel, this strategy, called single-component AFIR (SC-AFIR), utilizes local optimization procedures between two fragments to explore the reactions possible due to intramolecular pathways. Several methods which are conceptually similar to AFIR have also been proposed. These are the single-ended growing string method (SSM), [22] in which several nodes along coordinates (defined in terms of bonds, angles, and torsions) are added to the reactant(s) to drive the search towards a desired product, and the coordinate driving method (CD),

[23] which uses constrained electronic structure optimizations along a series of proposed reaction coordinates in order to detect feasible reaction pathways.

An alternative strategy is to use a two-step approach. During the first step, a set of possible product channels is generated using graphical (or combinatorial) rules based on the concept of the chemical bond [17,24] or using a heuristic generation of high-energy reactive complexes followed by their relaxation to minima. [25] During the second step, the algorithm attempts to connect them with the reactant channel using double-ended saddle point search methods, such as freezing string (FSM) [17,26] or growing string (GSM) methods. [24,27] Nudged elastic band methods could also be used. [28,29] It is also possible to heuristically generate guess structures for the TSs [30] which are further refined by conventional methods, such as the Berny method, [31,32,33] as is implemented in the KinBot program. [34] If one TS has been identified, programs such as MSTor can automatically search for its conformers, [35,36] and the rate constant can be calculated from the set of conformers, e.g., by multistructural transition state theory. [37]

In addition, it is worth noting that some methods based on machine learning algorithms have recently been proposed for the prediction of organic chemistry reactions. [38] However, the predictive capabilities of such approaches are implicitly limited by the range of reaction types contained in the training set used, and usually more explicitly limited by the use of specific reaction templates. Therefore, these methods are unlikely to discover a new and unexpected types of reactions.

While the aforementioned methods have been successfully applied to some organic (and organometallic) reactions for searching reaction pathways, [2,3,13,14,17,39] it is difficult to assess which methods are most effective, since the simulations were performed separately for different systems using different algorithms. There is a paucity of comparative studies and understanding of reaction discovery algorithm performance. In the present work, we aim to address this issue by performing a joint study of the unimolecular decomposition and isomerization of a γ -keto hydroperoxide (KHP), due to the importance of this class of molecules in autoxidation and low-temperature combustion chemistry [10,40,41,42]. We have already studied the chemistry of this KHP using the Berny method and FSM for which six unimolecular decomposition saddle

points were found [17]. Thus, this system also serves as a reference point for the methods that were available to us when we initiated this project and that we selected for the present calculations. This includes one single-ended (SSM) and two double-ended (FSM and GSM) methods, the heuristic KinBot algorithm, and SC-AFIR. In the following sections, we discuss the rich chemistry discovered using these methods and the advantages of such a joint approach. We also compare the performance and analyze the main pitfalls of each method. We hope that the results of this work will lead to the development of more cost-effective and reliable automated reaction discovery methods for general application in complex chemical systems.

2. COMPUTATIONAL PROCEDURE

In this section, we discuss the computational procedure of each method. Some of the more technical details are given in the Supporting Information (SI) file.

2.1 Combinatorial Search Using String Methods and Berny Optimization

The computational procedure for the automated identification of reaction pathways using string methods is nearly identical to the one previously proposed by us for the freezing string method (FSM). [17] In the present study, we considered only breaking and forming a maximum of three bonds. Reaction Mechanism Generator (RMG) thermodynamic libraries [43] and Benson's group additivity approach [44] were used to estimate the standard reaction enthalpies, ΔH_r^0 , for all the generated reactions. We did not attempt to discover reactions with an estimated ΔH_r^0 higher than 20 kcal/mol. Following this filtering step, 562 product structures remained. In order to verify whether the group additivity estimates were a sufficient proxy for the true reaction energies, more accurate estimates using density functional theory (DFT) were calculated for the set of filtered reactions. This analysis showed that group additivity was sufficient for the reaction filtering step (see Figure S1 in the SI file), although several reactions were included in the filtered set that would have been excluded based on the DFT criterion, and it is possible that some additional reactions would have been included; however, we chose to only consider the set of 562 products, as this already constitutes a very broad search space. Initial geometries of the reactant and product(s) were generated using the Open Babel program. [45] The reactant and product conformers were optimized using Berny optimization at the B3LYP/6-31+G* level of theory implemented in the Gaussian 09 program. [46] Searches for guess transition state (TS) structures were initiated using

the three string methods — FSM, GSM, and SSM. For this work, we used our own implementation of the FSM algorithm and interfaced it with Gaussian 09. For GSM and SSM, we used the program developed by one of the authors of those methods. [47] After successful generation of the string paths, the first reactive peak was selected as an initial guess TS structure (i.e., provided that at least one bond changed in the structure). The next step was the optimization of these guess structures using the Berny optimization algorithm as implemented in Gaussian 09. Finally, intrinsic reaction path (IRC) calculations were performed to verify whether the detected saddle point corresponded to the expected reactant and product paths.

2.2 Single-Component Artificial Force Induced Reaction Method

For SC-AFIR, [21,48] we used the M06-2X/3-21G level of theory [49] to enable inexpensive searches along the SC-AFIR paths. To be consistent with the other methods, we then re-optimized all critical points from the SC-AFIR search with B3LYP/6-31+G* using Berny optimization in Gaussian 09, including IRC calculations leading to products. It should be noted that only mechanistic studies were considered and gradient calls are not reported here for the SC-AFIR calculations. We set the γ -value of the SC-AFIR function to 400 kJ/mol to encompass a large search space, and considered all artificial forces to every fragment within KHP through intramolecular interactions.

2.3 Heuristic KinBot Approach

KinBot is a heuristic search program, which proposes sensible guesses for certain types of very broad reaction classes based on hard-won chemical knowledge. [34] For instance, during *internal* abstractions of atoms or groups, symbolically $A \cdots B \overset{\curvearrowright}{\curvearrowleft} C \rightarrow A \overset{\curvearrowright}{\curvearrowleft} B \cdots C$, the atom that is being placed from one connectivity to another (B) is typically half-way between the original A and final C atoms, with further prescriptions for bond angles, for instance depending on the relative positions of A and C in the molecule, or the type of the atoms involved. Creating these prescribed centers automatically and systematically essentially obviates human effort and eliminates human mistakes (hence the name, KinBot = “Kinetics roBot”), while still capitalizing on the knowledge we gathered for certain types of reactions. Note that KinBot was initially created and optimized for reactions involving C, H, and O atom containing radicals. Creating heuristics for radicals is a much simpler task than creating ones for closed shell molecules, because the sensible, i.e., low energy,

pathways in a radical decomposition or isomerization almost always involve the radical center. For closed shell molecules, such as the KHP in this study, it is harder to predict (and code) the preferred pathways.

In the KinBot case, we analyzed the synergetic effect of the joint application of several approaches. For this purpose, the corresponding calculations were performed in two steps. During the first run, similar to the other methods, we used KinBot's default reaction types and parameters without any prior knowledge about the KHP chemistry. In the second run, we extended the reaction types and included some variations to allow for more proposed structures than previously while taking into account the chemical reactions detected in the first run and by the other four methods (FSM, GSM, SSM, and SC-AFIR). Most significantly, in the first run, KinBot's constraints limited the search space primarily to the transfer of H atoms to other atoms, but in the second run KinBot considered transferring every atom type to every other one since many reactions of this type have been observed with the other methods. As a result of relaxing KinBot's criteria, the possibilities were significantly expanded resulting in the detection of significantly more channels. Because of the different nature of these calculations, we will distinguish the KinBot results from the first calculations (as a 'blind method assessment') and the second calculations (as an 'extended guided run') throughout the text as well as in the relevant figures and table. Note that a similar 'refining' approach could be implemented for the other methods. For instance, in the case of the string methods (FSM/GSM/SSM), inclusion of zwitterionic structures in the initial set of product channels as well as inclusion of channels with ΔH_r^0 higher than 20 kcal/mol would lead to the detection of more channels. However, in the present study, we limit ourselves only to illustrative refining of KinBot calculations due to the flexibility of the KinBot algorithm and the ease of implementing modifications in the corresponding code.

In Figure 1 we summarize the broad reaction types invoked by KinBot for KHP. It is important to note, however, that many times the intended reactions do not happen, but often the calculations converge rapidly to a slightly different, yet chemically significant saddle point. It is possible to increase the number of templates further to allow for even more complex rearrangements, but we did not extend beyond the ones in Figure 1.

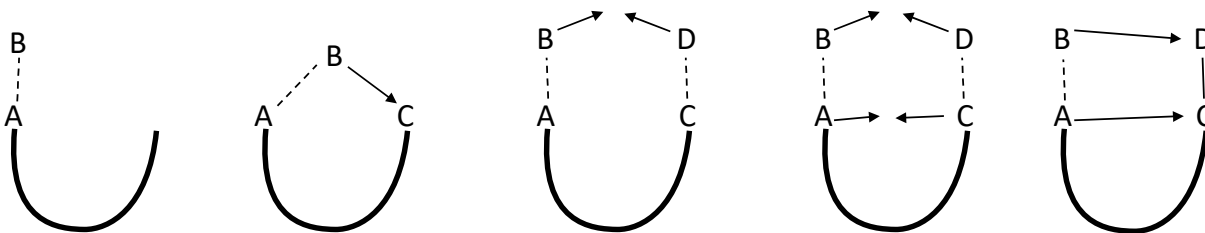


Figure 1. The schematics of the reactions invoked by KinBot automatically for KHP. The exact bond lengths, angles, and dihedrals depend on the type of the atoms in the active center and the length of the chain between A and C.

We used HF/STO-3G (first, generic run) and AM1 (second, expanded run) as cheap levels in this work, both of which have negligible computational costs, especially considering that in these constrained optimizations (all bond lengths are frozen) only a greatly reduced dimensionality gradient is needed for most steps. Once the desired conformation was achieved, we invoked the Berny algorithm to optimize the structure to a first-order saddle point at the B3LYP/6-31G* level of theory. Finally, for the successfully optimized structures we used IRC calculations to identify the reactant and the products, considering success if the reactant is the initial structure and the product is not. KinBot does not check whether the found saddle point is the intended one or not. KinBot also uses Gaussian 09 [46] to carry out both the constrained and the final optimizations.

Table 1. Summary of the automated detection of elementary chemical reaction steps using string methods (FSM, GSM, SSM), SC-AFIR, and KinBot.

Method	FSM			GSM			SSM			SC-AFIR			KinBot First run (Generic)			KinBot Second run (Expanded)		
Total number of channels tested (N_{channel})	562			562			562			74 ^l			200			443		
Total number of gradient calls (excluding IRC)	65420 ^a			756227 ^a			649589 ^a			N/A			8046 ^a (10093 ^b)			23458 ^a (25333 ^c)		
Total number of detected valid (i.e., reactive) saddle points (N_{detected})	153	27% ^d	-	215	38% ^d	-	371	66% ^d	-	17	23% ^d	-	83	42% ^d	-	162	37% ^d	-
• Unique saddle points	39	7% ^d	26% ^e	46	8% ^d	21% ^e	50	9% ^d	14% ^e	7	10% ^d	41% ^e	32	16% ^d	39% ^e	48	11% ^d	30% ^e
• Exact duplicates ^g	102	18% ^d	67% ^e	155	28% ^d	72% ^e	297	53% ^d	80% ^e	10	14% ^d	59% ^e	51	26% ^d	61% ^e	114	26% ^d	70% ^e
• Equivalent hydrogens ^h	12	2% ^d	8% ^e	14	2% ^d	6% ^e	24	4% ^d	6% ^e	0	0% ^d	0% ^e	0	0% ^d	0% ^e	0	0% ^d	0% ^e
• Intended reactions ⁱ	19	3% ^d	12% ^e	24	4% ^d	11% ^e	27	5% ^d	7% ^e	14	19% ^d	82% ^e	-	-	-	-	-	-
• Unintended reactions	134	24% ^d	88% ^e	191	34% ^d	89% ^e	344	61% ^d	93% ^e	3	4% ^d	18% ^e	-	-	-	-	-	-
Wrong reactant	85	15% ^d	-	39	7% ^d	-	23	4% ^d	-	4	5% ^d	-	13	6% ^d	-	22	5% ^d	-
Conformational saddle points	3	0.5% ^d	-	13	2% ^d	-	0	0% ^d	-	53	72% ^d	-	48	24% ^d	-	69	16% ^d	-
Crashed searches (N_{crashes})	321	57% ^d	-	295	52%	-	168	30% ^d	-	-	-	-	56 ^m	28% ^d	-	190 ^m	43% ^d	-
• Berny optimization failure	230	41% ^d	72% ^f	36	6% ^d	12% ^f	20	4% ^d	12% ^f	-	-	-	-	-	-	-	-	-
• Gradient failure in string method ^j	91	16% ^d	28% ^f	88	16% ^d	30% ^f	124	22% ^d	74% ^f	-	-	-	-	-	-	-	-	-
• Max number of iterations reached	0	0% ^d	0% ^f	171	30% ^d	58% ^f	18	3% ^d	11% ^f	-	-	-	-	-	-	-	-	-
• Dissociative channel ^k	0	0% ^d	0% ^f	6	0% ^d	0% ^f	6	1% ^d	4% ^f	-	-	-	-	-	-	-	-	-

^a B3LYP/6-31+G*;

^b Gradients in HF/STO-3G constrained minimization steps;

^c Gradients in AM1 constrained minimization steps;

^d $1/N_{\text{channel}} \times 100 \%$ (percentage of total number of channels);

^e $1/N_{\text{detected}} \times 100 \%$ (percentage of total valid saddle points);

^f $1/N_{\text{crashes}} \times 100 \%$ (percentage of total crashed searches);

^g Product and saddle point have the same adjacency matrix as a unique reaction product and saddle point;

^h Reaction is a duplicate because energetically equivalent hydrogens undergo the same reaction but with different atom indices;

ⁱ For double-ended FSM and GSM: reaction that connects reactant and product structure specified in the input, for single-ended SSM: reaction that contains the bond changes that correspond to the driving coordinates specified in the input; for SC-AFIR: reaction where the IRC calculation at the high level matches the products predicted by the low-level AFIR calculation; KinBot does not test this;

^j Umbrella term for any error in the electronic structure program during a single gradient calculation;

^k Combinatorially generated reactions do not include barrierless dissociation reactions so these channels are placed together with crashed searches;

^l For SC-AFIR, the total number of channels is the number of first-step channels as given by the low-level M06-2X/3-21G search;

^m KinBot crashed searches includes internal coordinate failures in the pre-optimization phase and any type of error in the Berny optimization.

3. COMPARATIVE ANALYSIS OF METHODS

Although the comparative analysis of the methods presented in this section focuses on rather technical aspects of the present calculations, we note that automated reaction discovery methodology is in an early stage and such an analysis can significantly improve the capability of discovering new and important chemistry.

3.1 Computational/Statistical Details

The statistical details of the automated search calculations are summarized in Table 1. This includes the total number of channels tested, total number of gradient calls (excluding IRC), total number of detected valid (i.e., reactive) saddle points, number of channels with wrong reactants, number of conformational saddle points, number of crashed searches, and the sources of failures. The statistics on the energy barriers and detected chemical reactions are discussed in the following subsections. Table 1 shows that SSM detected the highest number of chemical reactions corresponding to elementary steps of converting γ -ketohydroperoxide — 371 out of 562 channels resulted in detecting valid (i.e., reactive) saddle points. However, most of these channels correspond to duplicates (either exact duplicates or duplicates due to the same reaction with different but equivalent hydrogens) and only 50 saddle points are truly unique. Nevertheless, SSM provides the highest number of unique saddle points among all methods tested in the present work with GSM discovering 46 unique saddle points. However, Table 1 shows that the cost for such success is rather high — the computational expenses of these two growing string methods, estimated as the total number of gradient calls, were more than an order of magnitude higher than those of the FSM calculations. From this perspective, FSM represents an inexpensive alternative — 39 unique saddle points were detected and the total number of gradient calls was smaller by a factor of ten compared to GSM and SSM.

Rough estimates using all geometries of the detected saddle points in the KinBot study show that the Hessian matrix calculations are approximately five times more expensive than the corresponding gradient calculations at the high level of theory (B3LYP/6-31+G*), therefore, the initial construction of the Hessian adds only a negligible cost to its updates, which require only the gradients. The cost of the highly constrained Hessian calculations used to pre-optimize the structures is negligible at the HF/STO-3G (first run) or AM1 (second run) levels. This means that

the computational expense of KinBot (first run) was about eight times less than that of FSM, with the total number of saddle points detected by KinBot being slightly lower (32 in the first run). When KinBot was run with the extended set of rules, it was still about three times cheaper than FSM and detected more saddle points (48). Unfortunately, we were unable to extract the total number of gradient calls for the SC-AFIR calculations as we did not have access to the corresponding source code. The total number of unique saddle points detected by SC-AFIR was rather small (7).

For the three string methods, the ratio of the average number of gradient calls per successfully detected reaction is 308 (FSM):793 (GSM):1221 (SSM) (see the SI file). These numbers take into account both the string method and Berny optimization steps. This ratio is in line with previous estimates for string method calculations, [22,24,26,27] which indirectly shows that all three string methods provide equally good saddle point guess structures for subsequent Berny optimization in cases where the saddle point was eventually successfully detected. However, it is clearly different from the ratio for the number of gradient calls in the corresponding total calculations from Table 1 suggesting that our procedure of detecting saddle points can be optimized in the future. For instance, Table 1 shows that many string method calculations resulted in crashed searches. FSM calculations crashed in more than 50% of cases — 321 out of 562 channels finished with an error. Interestingly, the source of errors is different for all string methods. For FSM, the crashed searches were primarily due to failure of the Berny optimization step (~70%). Within the Berny optimization failures, almost 60% were due to reaching the maximum number of optimization steps (100), which signifies the inability to converge to a transition state structure in a reasonable amount of time. Approximately 20% of Berny optimization failures resulted from redundant internal coordinate errors in Gaussian 09. It is possible that such errors could be prevented by restarting the optimizations, but they typically are a result of chemically unreasonable geometries. Fewer than 20% of Berny optimization failures were due to convergence issues in the SCF calculations. Apparently, the maximum from the FSM path often provides a poor guess structure for further optimization [17] and often leads to crashes. Moreover, such guess structures can lead to saddle points which do not connect directly to the initial reactant — 85 wrong saddle points were detected in the FSM calculations. Note that GSM and SSM also detected a high number of saddle points for wrong reactant channels — 39 and 23, respectively. In such cases, GSM string

paths exhibit multiple elementary steps with apparently bad guess structures for further Berny optimization.

For GSM, a high number of crashes (295) was also observed, although most of them originated from reaching the maximum number of iterations inside the GSM algorithm (100 per channel), most probably due to the impossibility of constructing a growing string path in a limited number of iterations in cases where the string contains multiple elementary steps. It should be noted that such a path profile does not imply that a single-elementary-step profile does not exist for the same reaction, but GSM does not attempt to exclusively search for single-step pathways in its current implementation. Among the string methods, SSM exhibited the lowest amount of crashes, mainly from convergence failures encountered in the external quantum chemistry package Gaussian 09 during gradient calculations. While analyzing/fixing failures of external software packages is not included in the scope of the present study, the other sources of crashes can be considered as “essential” failures of the algorithms as they require many gradient calls and lead to undesired results. To summarize our comparison between the string methods, we point out that SSM provides the highest number of detected saddle points, although many of them are duplicates. It also exhibits the lowest number of crashes and undesired channels, i.e., wrong reactant(s) or conformational saddle points. However, it requires more gradient calls per successfully detected channel than the other string methods.

For SC-AFIR, we located 74 channels resulting from KHP as a reactant, 53 of which lead back to the reactant. The IRC scans and analysis of atom indices show that all of these channels are torsional conformation changes, i.e., they do not result from bond scission and rearrangements. The reason for such a result is likely due to the different nature of the SC-AFIR algorithm, which treats saddle points relatively equally, whereas the string methods more strongly bias the search space by directly specifying reactants and products to only obtain reactive saddle points. Similarly, KinBot automatically performs conformational rearrangements before directly searching for reactive saddle points, thus leading to a smaller percentage of conformational saddle points relative to the total number of channels. The level of theory used for the SC-AFIR calculations (M06-2X/3-21G), which may enable more accurate treatment of long-range dispersion interactions resulting from intramolecular hydrogen bonding interactions involving any of the three oxygen atoms, may

also somewhat influence the SC-AFIR search, [50] although further study would be required to validate this hypothesis. It is possible that dispersion interactions may only become significant with a larger basis set. After Berny optimization of saddle point structures at the B3LYP/6-31+G* level, 72% of the channels found were conformational saddle points (see Table 1). Yet, from the 17 saddle points corresponding to chemical reactions, 41% were unique. Meanwhile, 59% were exact duplicates (which includes different TS conformers), suggesting that SC-AFIR was relatively efficient at locating unique saddle points after the high level optimization. Additionally, it is possible that some saddle points that exist on the B3LYP/6-31+G* PES may not exist on the M06-2X/3-21G PES.

For KinBot, the percentage of crashed searches is comparable to the string methods (28% and 43% in the first and second runs, respectively) but the number of undesired saddle points (conformational or leading to wrong reactant(s)) is higher (61 and 91 in the first and second runs, respectively), originating mainly from conformational saddle points (48 and 69 in the first and second runs, respectively), i.e., points that do not correspond to a real chemical reaction event. KinBot is most efficient in that it has the highest percentage of detected unique saddle points relative to its total number of channels searched. KinBot crashed for two main reasons: (1) the created guess structure was too far from any saddle points, so the number of maximum iterations (100) was reached in the Berny optimization, or (2) the Cartesian geometry modifications prescribed by KinBot could not be completed in redundant coordinates in Gaussian. While the first mode of failure is clearly an intrinsic property of KinBot, i.e., not all saddle points can be predicted using templates, the second mode of failure can be prevented by taking smaller steps between consecutive geometry modification steps or imposing a looser convergence criterion for the intermediate step. We applied such techniques to minimize the second failure mode in the second run. In addition, in the second set of KinBot calculations finer IRC profiles were constructed leading to the detection of new channels. For example, a finer IRC profile may result in the detection of a biradical structure, which otherwise would not be detected if the IRC calculation steps past the shallow biradical well. It is interesting to note that although the total number of channels for the extended KinBot run was significantly larger than the total number of channels in the first run, 443 versus 200 (see Table 1), there were a significantly larger number of crashes, thus only resulting in 162 additional valid channels while yielding 16 additional unique channels.

Evidently, some of the additional channels that KinBot explored in the extended run were less chemically feasible and resulted in worse initial saddle point structures, which then ran into errors during the Beryn optimization step. This behavior is more similar to that observed with the string methods in that it renders KinBot less efficient at finding valid channels compared to the first run, but it significantly increases the number of unique saddle points found.

Clearly, all methods exhibit different behaviors, sources of error, and superfluous calculations.

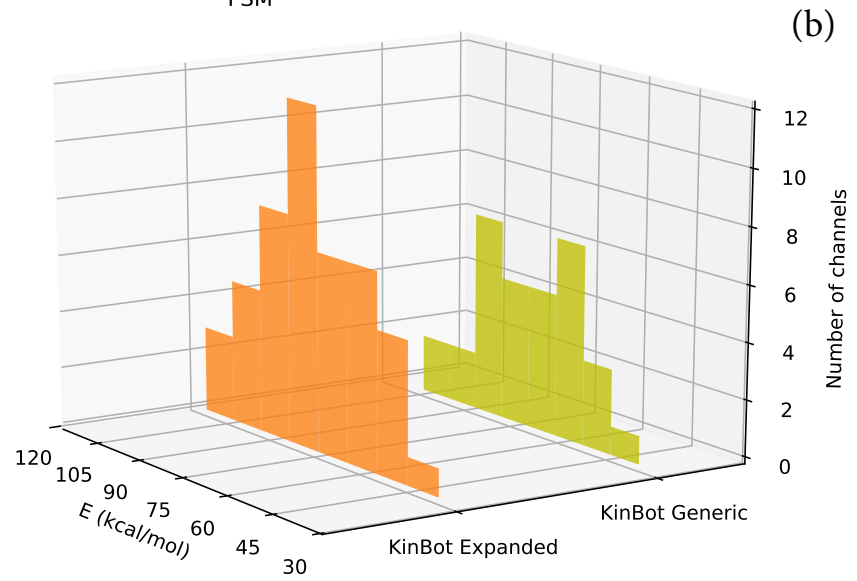
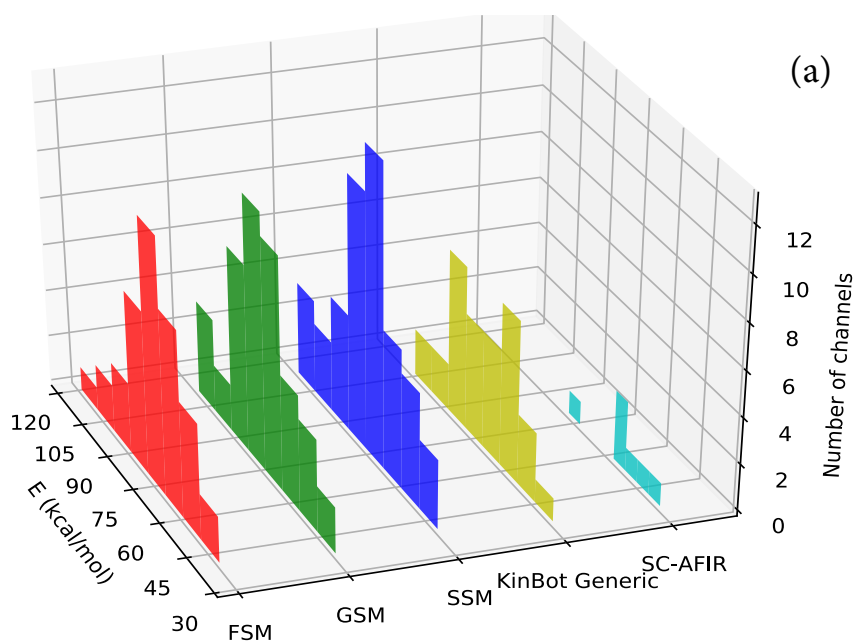


Figure 2. Histograms of the energy barriers (E_a) of the chemical reactions of unimolecular decomposition and isomerization of γ -ketohydroperoxide (with 8 equal-width bins containing the number of detected saddle points within each bin). (a) Reactions detected by FSM, GSM, SSM, SC-AFIR, and KinBot (first, generic run). (b) Reactions detected by KinBot during the first (generic) and second (expanded) runs.

3.2 Statistics on Energy Barriers

Figure 2a shows the histograms of the energy barriers of the reactions detected by FSM, GSM, SSM, SC-AFIR, and KinBot (first generic run). We are reporting energy differences relative to the reactant *excluding* the zero-point energy. Note that the various methods yield energy barriers in slightly different ways (see the SI file). Note, however, that the energy range of the conformers of the reactant was less than 5 kcal/mol for the set of reactant structures for all discovered reactions. All string methods have comparable energy barrier distributions and demonstrate a large peak in the intermediate range. The spread of energy barriers discovered by GSM is larger than that of FSM and KinBot, while SSM exhibits both the largest spread and finds the lowest barrier saddle points among the string methods. KinBot tends to avoid very high-barrier saddle points and is good at finding fairly low-barrier reactions. Figure 2b shows that this tendency remained in the extended run in which significantly more channels were detected — 48. However, as a result of extending the reaction rules such that all atom types are transferred to all other atom types, many more high-energy reactions were found in the second run, thus creating a histogram more similar to that of SSM. SC-AFIR also tends to find the lowest energy barriers — 6 out of 7 detected reactions have an energy barrier below 55 kcal/mol.

Figure 3 shows how the barrier heights obtained from the string methods compare to the barrier heights after Berny optimization of the saddle points. Clearly, FSM differs significantly from both GSM and SSM in that most of the FSM barriers are much greater than the corresponding optimized barrier heights. This again highlights the low quality of saddle point guess structures obtained from the FSM algorithm and explains the large number of crashes in the subsequent Berny optimization. Due to the exact saddle point search implemented in both growing string methods, the barrier heights of their guess geometries agree much better with the Berny optimized barrier heights for both GSM and SSM. Interestingly, there exist some GSM barriers that are lower in energy than

the corresponding Berny barriers. This stems from the fact that an exact saddle point search is only initiated for the highest energy structure in the GSM code if the string is composed of multiple elementary steps. If the first elementary step in the string does not contain the maximum energy point, then the saddle point of the first step is not optimized and may lie in a lower region on the PES. Compared to FSM, the GSM saddle point guess structures are of much higher quality, as indicated by energies closer to the Berny optimized energies, even though some of them have not yet undergone the exact saddle point search as part of the string method (as just discussed). This is due to the gradual convergence of the string to a minimum energy path with subsequent GSM iterations, which does not take place in FSM, in which all structures are frozen in place after initial optimization. For SSM, only single step paths exist due to the nature of the algorithm, such that all SSM barriers are larger than the Berny optimized barriers (within numerical accuracy). For both GSM and SSM, some points lie relatively high above the line of equality. These are mostly points where the predicted saddle points from GSM/SSM converged to a chemically different saddle point after Berny optimization. This is likely a result of the differences in the saddle point optimization algorithm implemented in GSM and SSM versus the Berny algorithm. It is also possible that some error is incurred due to the fact that GSM and SSM only construct an approximate Hessian, whereas the first step in the Berny optimization is the calculation of an analytical Hessian.

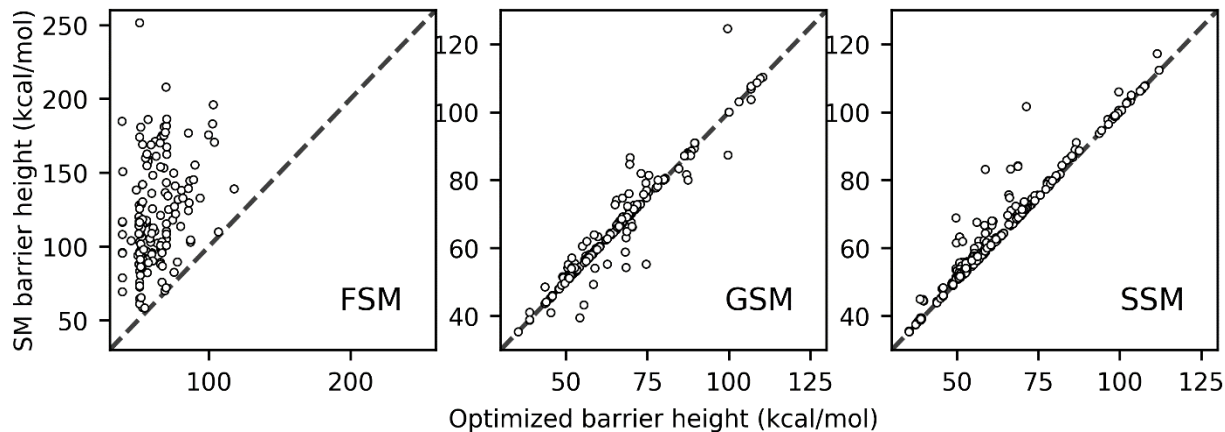


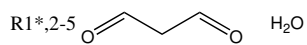
Figure 3. Comparison of the barrier heights from the three string methods (SM) to the optimized barrier heights after Berny optimization. The dashed line represents the line of equality.

4. ANALYSIS OF IDENTIFIED CHEMICAL REACTIONS

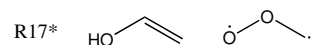
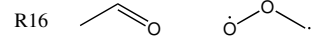
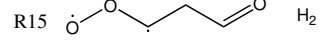
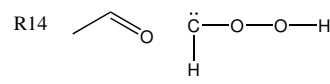
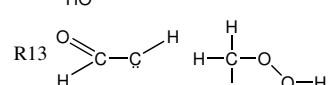
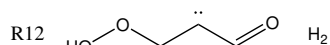
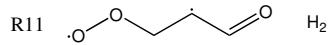
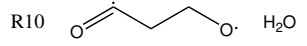
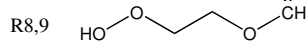
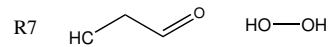
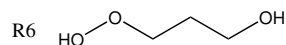
The previously outlined procedure found 75 distinct transition states originating from 3-hydroperoxypropanal, 68 of which were completely unexpected. Each search algorithm except for SC-AFIR discovered several TSs not found by any of the other methods. Figure 4 summarizes all identified products of the unimolecular one-step decomposition and isomerization of KHP obtained by all five methods (with two sets of results for KinBot). The mechanisms of the corresponding reactions, characteristic Cartesian coordinates, their energy barriers and differences in electronic energy of the reactant and products, as well as some details on the calculations, such as intended products and string path profiles, and the 3-D molecular structures of the saddle points found by KinBot are summarized in the SI file.

In total, 75 unique reactions were identified using all five methods. This significantly outperforms our previous FSM study [17] where only six chemical reactions were detected (reactions 4, 37, 47, 59, 69, and 72), which is partially due to the difficulties with the previous FSM calculations mentioned above as well as a result of the fact that the overwhelming majority of new chemical reactions have high-energy saddle points. In the previous study, such FSM profiles were discarded. Interestingly, in the present study, reactions 4 and 59 were detected by the other methods and not by FSM. Apparently, such discrepancy between the present and previous FSM calculations originates from the sensitivity of FSM to slight changes in starting geometries and the subsequently generated suboptimal saddle point guess structures. Another possible explanation is the different level of theory used previously (M06-2X/6-311++G*). [17] In addition, the reverse channel of reaction 17 was published in literature while we were preparing this manuscript. [51]

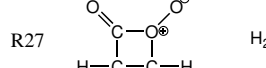
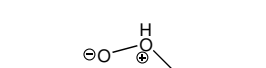
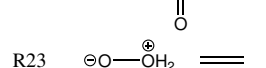
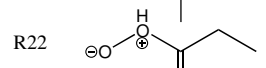
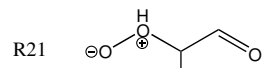
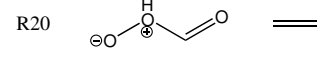
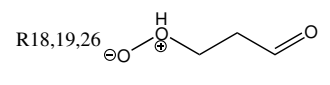
H₂O + malondialdehyde channels:



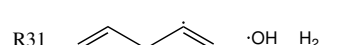
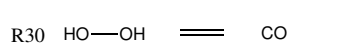
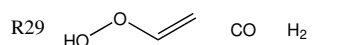
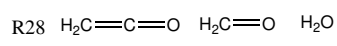
Biradical products including carbenes and the Criegee intermediates:



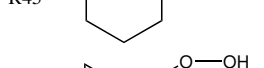
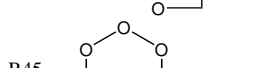
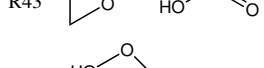
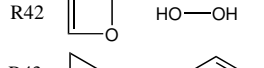
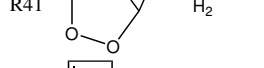
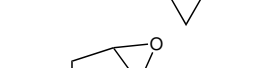
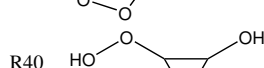
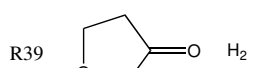
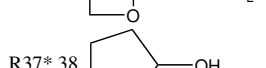
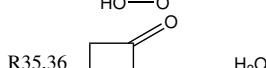
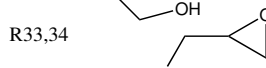
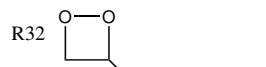
Zwitterionic structures:



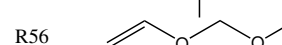
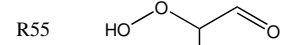
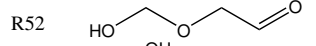
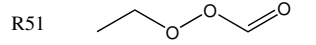
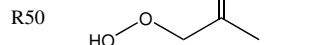
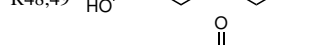
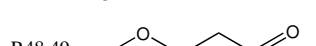
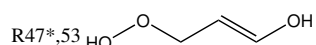
Channels with three products except zwitterionic structures:



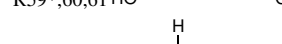
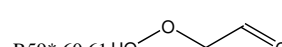
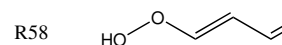
Channels with cyclic products:



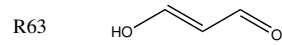
Stable (not radical or zwitterionic) unimolecular noncyclic channels:



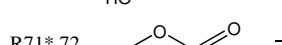
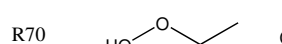
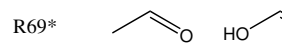
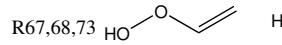
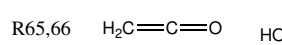
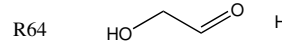
H₂ elimination channels:



Non-malondialdehyde H₂O elimination channel:



CH₂-CH₂ or CH₂-CHO bond breaking and forming two non-cyclic products:



HOOH elimination channels:

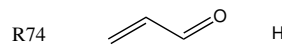


Figure 4. Automatically identified products of unimolecular reactions of γ -keto hydroperoxide (OOCCC=O) by using the FSM, GSM, SSM, KinBot (first and second runs), and SC-AFIR methods. Details on the mechanism of these reactions can be found in the Supporting Information file. *Reactions discovered previously (reactions 1, 37, 47, 59, 69, 71 in Ref. [17] and reaction 17 in Ref. [51]).

4.1. Comparison with the Existing Chemical Kinetic Database

The present results also significantly outperform the existing RMG kinetic database where only four of the detected reactions (reactions 37, 40, 47, and 56) are present (see Figure S2 in the SI file). In addition, RMG predicts 13 other chemical reactions which were not detected in the present study by any of the methods. Comparison between the present results and the RMG kinetic database suggests that the simple rules implemented in the RMG package may predict the correct bond breaking but they do not take into account additional internal H-transfers which stabilize the products. For instance, reaction 74 in Figure 4 has a similar RMG counterpart (RMG reaction 11 in the SI file) but involves an additional H-transfer between the products, which stabilizes and reduces the overall product energy. In addition, the RMG library does not take into account channels with unusually long bond distances in the transition state that are detected in the present study (reactions 3, 4, 34, 36, 62) (note, however, that the bond lengths are not long enough for these reactions to be called roaming ones). Thus, the present work also suggests improvements to the RMG rules that could be implemented in the future. However, it is important to note that all of the methods in this study are primarily looking for saddle points, whereas some important dissociation channels (e.g., H elimination from the carbonyl group, see RMG reaction 13 in the SI file) are barrierless (at least at the DFT level), therefore, the real number of reaction channels is always larger than the number of reactive saddle points for closed shell systems. For example, the Jalan et al. [10] paper features the O–OH homolytic scission as the lowest barrierless channel, with a 49.5 kcal/mol barrier at the CCSD(T)//M06-2X level. Our B3LYP calculations predict a ~58-88 kcal/mol range for the water elimination channels. Given the tightness of these channels, the homolytic scission will dominate over the found saddle points. Moreover, some of the other reactions predicted by RMG involve additional biradical products, which are likely formed by high barrier processes that would have been excluded in the initial thermodynamic filtering for FSM/GSM/SSM and were also not found from sub-optimal saddle point guesses that converged to unintended saddle points.

4.2 Types of Identified Chemical Reactions

The reactions presented in Figure 4 are divided in the following way: H₂O + malondialdehyde channels (reactions 1–5); formation of biradical products including carbenes and Criegee intermediates (reactions 6–17); products with zwitterionic structures including (reactions 18–27);

channels with three products except zwitterionic structures (reactions 28–31); channels with cyclic products (reactions 32–46); stable (not radical or zwitterionic) unimolecular noncyclic channels (reactions 47–57); H₂ elimination channels (reactions 58–62); another, non- malondialdehyde H₂O elimination channel (reaction 63); CH₂-CH₂ bond breaking and forming two non-cyclic products (reactions 64–68); CHO-CH₂ bond breaking channels (reactions 69–73); HOOH elimination channels (reactions 74 and 75). Note that the reactions are grouped into these somewhat loosely defined groups primarily for orientation (for instance, the Criegee intermediates in reaction 15-17 can also be represented as their zwitterionic resonant structures). Also note that some of the detected reactions lead to the same product(s). For instance, we detected five different routes for H₂O elimination forming malondialdehyde (reactions 1–5, see the SI file). We labeled these as separate reaction channels, because their mechanisms are distinct, which would, for instance, gain importance in an experimental study using isotopically labeled KHP. We emphasize that our choice of considering such saddle points as fundamentally different is somewhat subjective especially since no rigorous conformational search was done to find the lowest energy conformers of the saddle points for any of the methods tested in the present work; however, channels marked as separate do not appear to be simple conformers of any other channels. It should be noted that SSM, SC-AFIR, and KinBot are capable of initiating multiple searches which could end in the same product because the product structure is not explicitly prescribed initially, whereas FSM and GSM only allow for one search per product. This enables SSM, SC-AFIR, and KinBot to find multiple distinct transition states corresponding to the same product. However, the initial set of reactions for SSM is the same as that for FSM and GSM in this study so that the methods can be compared directly in terms of their ability to find the same reactions. It is possible that this explains why KinBot was able to find all five routes leading to malondialdehyde, whereas none of the other methods were able to find all of those channels (see SI).

The majority of the detected reactions proceed through breaking two bonds (52), while the number of reactions with three bonds being broken is more than a factor of two smaller (21). Interestingly, we also observed one reaction with only one bond breakage (reaction 19 detected by all the methods) and one with four bond breakages (reaction 63 detected by FSM, SSM and KinBot) even though none of these methods were targeted at reactions with four bond breakages. It is also worth noting that some of the detected product structures are highly unstable (e.g., the biradical structure

in reaction 10 and the zwitterionic trivalent O structure in reaction 20) and therefore are expected to undergo fast subsequent reactions. Note that for these cases the nature of the product is sometimes very sensitive to the conformer of the saddle point and how the IRC is carried out (step lengths, etc.). Some conformations stabilize the biradicals, while others lead to ring-closures or further decomposition. For instance, the products of reaction 10 could rapidly undergo beta scission to yield the products corresponding to reaction 28. These reactions, if important kinetically, require further dynamical investigations (such as the analysis of plateau regimes of the real-time correlation function responsible for describing elementary chemical reactions in order to rigorously separate chemical events and detect the corresponding product channels), which lies outside the scope of this work. Some of the high-energy structures found by KinBot were a result of the reaction type extension in the second KinBot run. In the extended search, KinBot attempted to transfer all atom types to all other ones, which led to the additional discovery of reactions 8 and 18.

4.3 Exclusiveness of Identified Chemical Reactions

Only four reactions were detected by all five methods — 1, 19, 37, and 74. If we exclude SC-AFIR, 18 reactions were detected by the remaining four methods (including FSM, GSM, SSM, and the first run of KinBot) — 1, 2, 14, 15, 16, 19, 28, 30, 37, 47, 48, 51, 58, 64, 69, 71, 73, and 74. Including the results of the second KinBot run in the comparison adds three more reactions — 8, 13, and 39. Still, the reproducibility of the four methods is low in terms of the total number of channels found for each method. Except SC-AFIR, each method was able to detect some reactions exclusive to that method, although the majority of such saddle points are high in energy. The string methods detect several such exclusive reactions (7, 20, 35, and 72 from FSM; 22, 42, 43, and 65 from GSM; 23, 24, 38, 44, 68, and 75 from SSM). During the first run, KinBot also detects a few of them (5, 12, 17, 45, 66), but during the second “expanded” iteration the total number of exclusive reactions has more than doubled (13, 25, 27, 34, 46, 61, 62 in addition to the original set).

All five methods were able to detect the channel with the lowest energy barrier (approximately 35 kcal/mol) — formation of 1,2-dioxolan-3-ol (reaction 37) which, as indicated in the Introduction, is the first step in the so-called Korcek mechanism. [10] However, among the string methods, only

GSM finds the Korcek reaction from the combinatorially intended product channel while FSM and SSM detect the corresponding saddle point from unintended channels. Interestingly, SSM was able to detect another enantiomer of 1,2-dioxolan-3-ol formed via the mechanism similar to the Korcek one (reaction 38), although its barrier is significantly higher (67.2 kcal/mol). KinBot detects the Korcek reaction as an internal 1,2 addition channel, a generic reaction from the first run. All methods except SC-AFIR were also able to detect the second lowest energy saddle point, for the reaction forming acetaldehyde and formic acid (reaction 69).

The analysis of the string path profiles for the detected reactions in the SI file shows that FSM often generates paths with multiple barriers (most noticeable reactions are 33, 37, 49, 62, 64, 70, and 72). In the previous study, [17] such profiles were discarded but the present results show that such cases can still lead to a discovery of the saddle points for single elementary steps (although typically with high energy barriers). Interestingly, in the previous study [17], the Korcek reaction (reaction 37) was detected by FSM from the intended channel while the present FSM calculations detected it from the Berny optimization of the maximum along the string path for another channel (see the SI file) again confirming the seemingly random nature of the method and high sensitivity to the choice of input parameters.

SSM finds more intended reactions, however many of them represent a transfer of equivalent hydrogen atoms — only 9 out of 27 SSM channels are unique (see Table 1 and the SI file). This does not outnumber the FSM and GSM statistics by a lot — 7 out of 19 and 11 out of 24, respectively. Only three reactions were detected by all three string methods from the intended channels (reactions 9, 47, and 51) while the majority of saddle points detected by the string methods originate from unintended channels. Whether to consider this as a drawback of the proposed approach is not obvious — many of the detected channels were unexpected and were not even included in the initial set of product channels (e.g., zwitterionic structures such as in reaction 19 where only one bond breaks and excessive valences are encountered, or reactions 48 and 49 where the products are isomorphic with the reactant molecule). However, our approach can easily fail if one has a specific reactant and product in mind. Because many saddle points were detected from unintended channels with string paths containing multiple elementary steps, it is expected that refining the initial set of channels to focus on more easily accessible single elementary steps

would lead to a higher percentage of intended reactions, although this would limit the breadth of discovered reactions.

4.3 Chemical Reactions with Low Energy Barriers and Potentially Important Chemistry

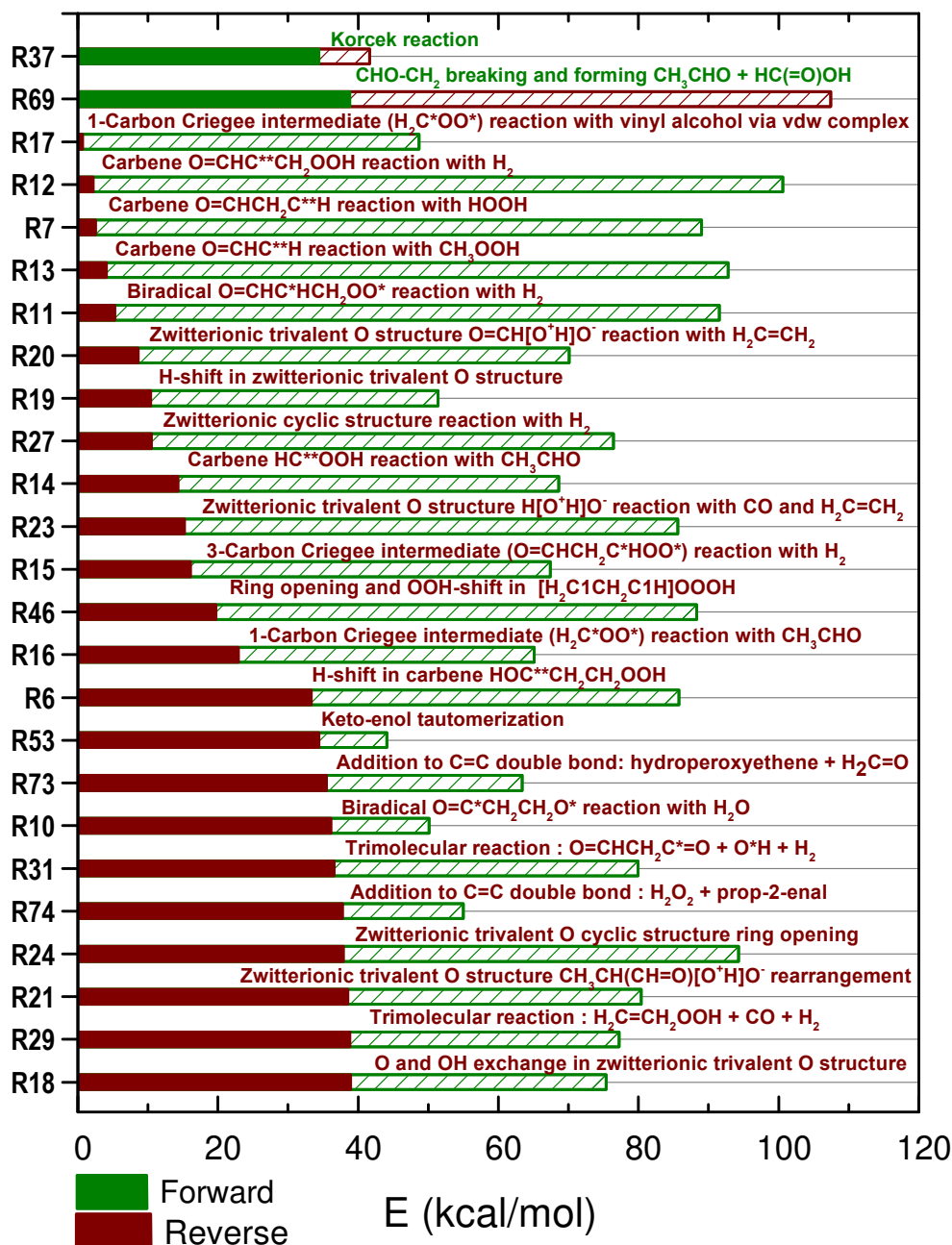


Figure 5. Energy barriers (E_a) of 25 chemical reactions of unimolecular decomposition and isomerization of γ -ketohydroperoxide for which E_a is below 40 kcal/mol (either for the forward or

reverse channel). (Note that bars representing the barrier heights are not stacked but overlapped for each reaction.)

25 chemical reactions with energy barriers below 40 kcal/mol (either forward or reverse channels) are summarized in Figure 5. Only two reactions (37 and 69) satisfy this criterion in the forward direction and both of them were detected in the previous FSM study [17] suggesting that there is a low probability of finding additional low energy channels (although not guaranteed). However, one can notice a significant number of previously unexpected reverse channels with low energy barriers, many of which involve biradical and zwitterionic structures.

Several of the low-barrier reverse reactions leading to the formation of KHP involve Criegee intermediates as the reactant (reactions 15-17). These radicals play an important role in atmospheric chemistry and are in the focus of current intense research [51-60]. Reaction 17 involves the reaction of the simplest Criegee intermediate (CI) with vinyl alcohol (VA), which leads to KHP in a practically barrierless reaction. Vinyl alcohol is a detectable tautomer of acetaldehyde in flames [61] and in the atmosphere [62,63]. Interestingly, the reaction of the same CI with acetaldehyde (reaction 16) has a fairly noticeable barrier of 23 kcal/mol when the product is KHP. According to Jalan et. al., [60] CI can also react with acetaldehyde in a barrierless manner, but such a reaction leads to a secondary ozonide. Another pair of low-barrier chemical reactions presented in Figure 4 is H₂ fission by a 3-carbon Criegee intermediate (reaction 15) with $E_a = 16$ kcal/mol and a reaction involving a similar biradical structure (reaction 11) which exhibits a much smaller barrier of 5 kcal/mol. However, reactions of CI with H₂ are likely not of relevance in the atmosphere due to the low concentration of hydrogen molecules [64].

The zwitterionic structures detected in reactions 18-27 contain hypervalent O-atoms, which made their discovery highly surprising. They were not included in the product set of the string methods used in the present study due to their unusual valency and were thus discovered serendipitously. However, because KinBot does not consider products to guide its search, the H-transfer to the inner O atom of the –OOH groups was an allowed step. It is notable that six out of ten detected reactions of these zwitterionic structures fell within our category of “low energy” barriers. Reaction 23 forms water oxide [65,66] and the other channels form ylides of varying complexity.

Although usually unstable in the gas phase, certain subgroups of zwitterions often appear in biochemical studies (under physiological conditions, a variety of biomolecules, such as amino acids, exist as zwitterions) and in organic synthesis. The present results particularly contribute to the chemical knowledge for zwitterionic structures with positively and negatively charged atoms in the O₂ fragment. This has potential implications on the chemistry of alkylperoxides — a class of species important in low-temperature autoignition and many atmospheric chemistry problems. For instance, it has been shown that water elimination from alkylperoxy radicals can involve zwitterionic states [11]. Apart from water oxide, the ylides discovered by the automated search have not been analyzed in much detail in literature. Schalley et al. conducted an extensive characterization of the PESs of methanol oxide and dimethyl ether oxide [67] and Vereecken recently discovered the relevance of alcohol and ether oxides in atmospheric chemistry systems involving a Criegee intermediate [68]. Alcohol oxides, such as those in reactions 18, 19, 21, 24, 25, and 26 are expected to undergo rapid tautomerization to the corresponding hydroperoxides, whereas ether oxides are more stable [67,68]. In particular, reaction 27 involves a cyclic ether oxide for which isomerization to an alkylperoxide is expected to be associated with a very significant energy barrier indicating that the species could be long-lived enough to participate in other reactions, although it is not clear in which chemical systems such a species would be present in significant amounts. To the best of our knowledge, carboxylic acid oxides, such as those in reactions 20 and 22, have not been reported in literature, and the low barrier of the reaction between formic acid oxide and ethylene indicates that these species could potentially be relevant if the tautomerization to the corresponding peroxy acid is slow enough.

Another interesting class of structures — carbenes — was detected in reactions 12, 7, 13, 14, 6 (sorted according to the height of the corresponding energy barriers in Figure 5). These reactions were not expected as the forward reactions are associated with large energy barriers. The version of group additivity used for the creation of the filtered data set mostly underestimates the endothermicity for such reactions compared to more accurate B3LYP energies, which led to their inclusion in the set of filtered reactions (see the SI file). Reaction 6 corresponds to an intramolecular O-H insertion (or internal H-shift), while the remaining four reactions are intermolecular insertions. Both are common types of carbene reactions; however, the context in which they were discovered is surprising. The observed ‘kinetic’ preference of carbene insertion

into an existing bond is in accordance with the common rule for carbene X-H insertion: reactions 12 (X = H) and 7 (X = O) have lower energy barriers than reactions 13 and 14 (X = C). The intramolecular carbene insertion reaction 8 (O-C) was not included in Figure 5 due to an energy barrier higher than 40 kcal/mol (see the SI file). Our results show that the carbene insertions occur in single elementary steps, which is possible because we are only considering the singlet surface for this study [69], whereas triplet carbenes would have to react in a sequence of two steps. Reaction 7 is notable due to the reaction with hydrogen peroxide, which is present in most combustion models and therefore might constitute a valid path to KHP. Even if the carbene is only present for short amount of times, its singlet state will result in high rate coefficients, and might make such a reaction important. Cyclopropene, or its isomers allene and propyne, is produced in a reaction between methylene and acetylene [70,71], and is relevant in combustion systems [72]. It is conceivable for oxygen atoms to add to the double bond in cyclopropene, which could then undergo further ring-opening to the carbene in reaction 7. Alternatively, it would be possible for oxygen to react directly with acetylene, the product of which could then react with methylene and also lead to the carbene. We note that such thoughts are purely hypothetical at this point, but they may warrant further investigation, which could show that species like the carbene in reaction 7 are important in combustion systems. Similarly, reaction 13 involves a carbene that might be formed from vinoxy radicals, which are also important in combustion systems [73], and thus may be a relevant reaction in such systems.

Although the low-barrier chemical reactions described above are potentially important, calculating rate coefficients for all of the reactions is not the goal of the present work. However, we decided to investigate one reaction (17) in more detail as it involves promising new chemistry that may be relevant in atmospheric systems. A characterization of KHP formation resulting from the reaction of CI with VA — reaction 17 — was published by Vereecken during the preparation of this manuscript [51]. In accordance with our findings, Vereecken found that the reaction proceeds through a van der Waals complex and a subsequent shallow barrier. Apart from the brief mention in Ref. [51], we are not aware of any other publications that have studied this reaction in greater depth, even though its small submerged barrier suggests a fast rate, which may render it important in atmospheric systems. In general, tautomerization of acetaldehyde to VA is associated with a high barrier in the gas phase, although it has been discovered that photo-tautomerization of

acetaldehyde is a viable process in the atmosphere, which can lead to up to 21% stable VA produced from excited acetaldehyde [63,74]. Due to the large energy of CI and VA relative to KHP, it should be possible to skip the KHP well at finite pressures and to obtain other bimolecular decomposition products of KHP directly. Likely candidates are the products obtained from the dissociation of the oxygen single bond in the hydroperoxy group of KHP, which yields two radicals (note that such products were not included in the search space for the automated TS search methods because the methods are not suitable for purely barrierless processes). To obtain quantitatively accurate results, we reoptimized the relevant geometries and calculated frequencies at the UM06-2X/aug-cc-pVTZ level of theory and subsequently calculated energies with the RCCSD(T)-F12a/cc-pVTZ-F12 method (geometries, energies, and frequencies are shown in the SI). The PES for the CI + VA network is shown in Figure 6. The Cantherm code, distributed as part of RMG [43], was used for all kinetic rate calculations. Further technical details elaborating upon the rate constant calculations are described in the SI. The rate constant for the reaction of CI and VA was calculated to be $3.7 \times 10^{-11} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$ at 298 K, which is faster or on the same order of other dominant reactions of the Criegee intermediate [75]. At atmospheric conditions, virtually all of the flux is towards the two radical products including OH and none of the product is obtained in the KHP form, which indicates that this reaction may be a significant source of OH and alkoxy radicals in the atmosphere and could help explain the underprediction of OH in current models compared to experiment [76,77]. However, it should be noted that the majority of Criegee intermediates available for reaction with VA in the atmosphere have larger substituent groups as a result of easier collisional stabilization after their production from larger primary ozonides [76]. Therefore, thermalization to KHP will be more likely for larger CI. *A priori*, it is conceivable for CI to add to the double bond in VA instead of forming KHP, which can lead to two separate ring structures depending on the orientation during the addition reaction (in fact, one of the structures is identical to the product of the first step of the Korcek mechanism). We calculated the rate coefficients for these reactions using the same methodology as described in the SI, which showed that they are not relevant at atmospheric conditions. Reaction with OH and O₂ constitutes the main reaction of VA in the atmosphere with a total rate constant of $6.8 \times 10^{-11} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$ at 298 K [78]. The concentration of stabilized CI in the atmosphere is associated with significant uncertainty and ranges from approximately 10^3 to $5.5 \times 10^4 \text{ molecule cm}^{-3}$ for global average concentrations and has peak concentrations of $10^5 \text{ molecule cm}^{-3}$ [75,79]. Global average OH

concentration can be estimated as 10^6 molecule cm^{-3} [78]. This suggests that the branching ratio between the reaction of CI and VA and the reaction of OH and VA is approximately 0.05 at peak CI conditions, although this estimate is quite variable based on uncertainties in the concentrations and in the calculated rates. As it is beyond the scope of the current study, we did not evaluate the effect of different substituents in CI, which are known to strongly affect rate constants [75]. Such an analysis may lead to more competitive rates of the CI + VA reaction. Enols other than vinyl alcohol may also play a role in atmospheric chemistry and could also undergo reaction with CI in a similar process. For example, acetylacetone exists predominantly in its enol form and is expected to be present in the atmosphere due to its widespread use in industry and ensuing release [80]. Such stable enols may be especially prone to reacting with Criegee intermediates, as photochemical activation is not required for their production, which may even render them important in nighttime conditions. Therefore, a subsequent study analyzing both the effect of substituents in CI and the reactivity towards stable enols would be of significance.

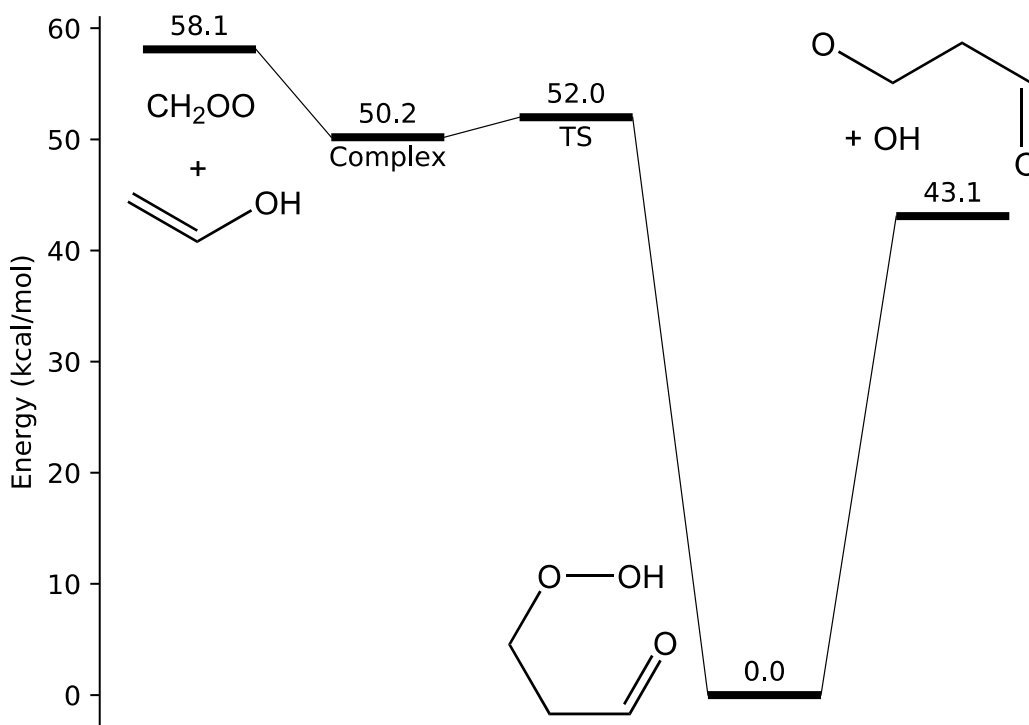


Figure 6. Potential energy surface of the Criegee + vinyl alcohol reaction at the RCCSD(T)-F12a/cc-pVTZ-F12//UM06-2X/aug-cc-pVTZ level of theory, which proceeds through a barrierless formation of a van der Waals complex followed by a shallow transition state to KHP,

which can dissociate to two radicals. Well-skipping directly to the radical products is the preferred channel at atmospheric pressures. All values include the zero-point energy.

5. CONCLUSIONS

Transition states for 68 previously unknown reactions of the simplest γ -ketohydroperoxide were discovered automatically by combining five different saddle-point-search algorithms. This significantly broadens the chemical knowledge for the present system and for all γ -ketohydroperoxides. The single-ended growing string method (SSM) found transition states leading to 50 distinct products, the most of all methods tested, while the single-component artificial force induced reaction method (SC-AFIR) detected the fewest (7). The remaining methods demonstrated intermediate results — transition states to 39 product channels were discovered by the freezing string method (FSM), 46 were discovered by the double-ended growing string method (GSM), and 32 were found by KinBot (generic run). The present FSM results outperform the previous ones [17], where only six reaction channels were detected using the same method, showing that the overall performance of the method is highly dependent on the input parameters and the internal details of the computational procedure. All of the methods were able to detect the reaction with the lowest energy barrier, which corresponds to the first step of the so-called Korcek mechanism [10]. However, among the string methods, only GSM detected the Korcek reaction when searching for this channel, while FSM and SSM found this saddle point serendipitously while searching for transition states corresponding to different products. All methods, except for SC-AFIR, detected several exclusive channels (i.e., reactions detected only by a given method). SC-AFIR found the highest percentage of conformational saddle points.

Analysis of the reverse reactions leading to the formation of γ -ketohydroperoxide reveals promising new chemistry with low energy barriers for reactions involving zwitterions, biradicals, including carbenes, and Criegee intermediates. Similar species are in the focus of current intense research and we show that several of the detected reactions may be relevant in atmospheric and combustion chemistry. In particular, we study the reaction of the simplest Criegee intermediate with vinyl alcohol in more detail, which leads to γ -ketohydroperoxide via a submerged barrier reaction with possible well-skipping to two radical products, by calculating the rate constant for this reaction at atmospheric conditions and comparing it to other possible decay processes of vinyl

alcohol. As a result of our analysis, directions for subsequent chemical rate studies have been formulated exhibiting a clear example of the benefits of applying automated transition state search algorithms for the discovery of new chemical reactions.

The present results highlight the complexity of chemical space exploration and the sensitivity to input parameters and level of theory of the methods assessed in the present work. They also demonstrate the power of joint application of several automated approaches for discovery of elementary chemical reactions since none of the above listed methods is greatly superior to the others and all methods discover transition states not detected by any other algorithm (except SC-AFIR). Moreover, chemical knowledge obtained using one or more methods can be used to improve the performance of another method. For example, during the second iteration of the KinBot calculations, selected due to its flexibility and the ease of implementing modifications, the reaction types were extended using the chemical knowledge obtained in the first run by all five methods resulting in significant improvement — transition states leading to 16 additional product channels were detected by the extended version of KinBot. At the same time, such results also show that it is impossible to determine how exhaustive an automated search was, as none of the methods are capable of finding all of the reactions and there is no guarantee that all reactions can be discovered by the joint approach.

Our analysis also shows various sources of crashed searches and excessive (useless or repetitive) computation for the methods selected for the present comparative study. We hope this information can be used to improve future applications of the methods. In the future, we plan to extend the present work to similar algorithms based on other representative methods for saddle point detection, such as the nudged elastic band method, [28,29] conjugate peak refinement, [81] the ridge method, [82] and others. We envision that similar method assessment for bimolecular chemical reactions would be a natural subsequent step for the community.

The results of this study indicate a vast number of unexpected elementary-step chemical reactions remain to be discovered, and that despite the weaknesses of existing transition state search algorithms, this reaction discovery process can be greatly accelerated by automated search of potential energy surfaces.

ASSOCIATED CONTENT

Supporting Information

Additional information on the computational procedure of the string methods combined with Berny optimization and KinBot approach. Comparison of group additivity and DFT reaction energies for the set of filtered reactions. Data for 75 detected unimolecular reactions of γ -ketohydroperoxide including pictures depicting the reaction mechanisms, string paths for the string methods, pictures with the geometry of the saddle points from the KinBot calculations (where available), as well as some descriptions (intended channels for the string methods, energy barriers, reaction energies and the number of gradient calls (excluding the IRC calculations)), and geometries of the saddle points. Energies are reported at the B3LYP/6-31+G* level of theory without ZPE correction. RMG kinetic database for the unimolecular decomposition of γ -ketohydroperoxide. Geometries, energies, and frequencies for the intermediates in the Criegee + vinyl alcohol reaction at the RCCSD(T)-F12a/cc-pVTZ-F12//UM06-2X/aug-cc-pVTZ level of theory. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

*Corresponding Author

whgreen@mit.edu; y.suleymanov@cyi.ac.cy

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

C.A.G., A.J., Y.-P.L., W.H.G., and Y.V.S. gratefully acknowledge financial support from the Air Force Office of Scientific Research, through Grant FA9550-16-1-0208. The authors thank Paul Zimmermann for useful discussions on the growing string method and permission to use his code. A.J. thanks the Academic Center for Computing and Media Studies (ACCMS) at Kyoto University for providing computational resources, and also acknowledges helpful discussions with Pedro J. Castro. A.J. would also like to thank Marcelo Delgado for programming support in providing some of the data used in this study. J.Z. was funded under DOE BES, the Division of Chemical Sciences, Geosciences, and Biosciences. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned

subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Y.V.S. also thanks the European Regional Development Fund and the Republic of Cyprus for support through the Research Promotion Foundation (Project Cy-Tera NEA ΥΠΟΔΟΜΗ/ΣΤΡΑΤΗ/0308/31).

REFERENCES

- [1] Martinez, T. J. *Acc. Chem. Res.* **2017**, *50*, 652.
- [2] Dewyer, A. L.; Zimmerman, P. M. *Org. Biomol. Chem.* **2017**, *15*, 501.
- [3] Klippenstein, S. J. *Proc. Comb. Inst.* **2017**, *36*, 77.
- [4] Vinu, R.; Broadbelt, L. J. *Annu. Rev. Chem. Biomol. Eng.* **2012**, *3*, 29.
- [5] Schlögl, R. *Angew. Chem. Int. Ed.* **2015**, *54*, 3465.
- [6] Poree, C.; Schoenebeck, F. *Acc. Chem. Res.* **2017**, *50*, 605.
- [7] Hammes-Schiffer, S. *Acc. Chem. Res.* **2017**, *50*, 561.
- [8] Vereecken, L.; Glowacki, D. R.; Pilling, M. J. *Chem. Rev.* **2015**, *115*, 4063.
- [9] Van de Vijver, R.; Devocht, B. R.; Van Geem, K. M.; Thybaut, J. W.; Marin, G. B. *Curr. Opin. Chem. Eng.* **2016**, *13*, 142.
- [10] Jalan, A.; Alecu, I. M.; Meana-Pañeda, R.; Aguilera-Iparraguirre, J.; Yang, K. R.; Merchant, S. S.; Truhlar, D. G.; Green, W. H. *J. Am. Chem. Soc.* **2013**, *135*, 11100.
- [11] Welz, O.; Klippenstein, S. J.; Harding, L. B.; Taatjes, C. A.; Zádor, J. *J. Phys. Chem. Lett.* **2013**, *4*, 350.
- [12] Quelch, G. E.; Gallo, M. M.; Schaefer, H. F. *J. Am. Chem. Soc.* **1992**, *114*, 8239.
- [13] Ludwig, J. R.; Zimmerman, P. M.; Gianino, J. B.; Schindler, C. S. *Nature*, **2016**, *533*, 374.
- [14] Khomutnyk, Y. Y.; Argüelles, A. J.; Winschel, G. A.; Sun, Z.; Zimmerman, P. M.; Nagorny

P. J. Am. Chem. Soc., **2016**, 138, 444.

[15] Magoon, G. R.; Green, W. H. *Comput. Chem. Eng.* **2013**, 52, 35.

[16] Bera, P. P.; Sattelmeyer, K. W.; Saunders, M.; Schaefer, H. F., III; Schleyer, P. v. R. *J. Phys. Chem. A* **2006**, 110, 4287.

[17] Suleimanov, Y. V.; Green, W. H. *J. Chem. Theory Comput.* **2015**, 11, 4248.

[18] Yu, T.; Zheng, J.; Truhlar, D. G. *J. Phys. Chem. A* **2012**, 116, 297.

[19] Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, 99, 12562.

[20] Wang, L.-P.; Titov, A.; McGibbon, R.; Liu, F.; Pande, V. S.; Martinez, T. J. *Nat. Chem.* **2014**, 6, 1044.

[21] Maeda, S.; Harabuchi, Y.; Takagi, M.; Taketsugu, T.; Morokuma, K. *Chem. Rec.* **2016**, 16, 2232.

[22] Zimmerman, P. M. *J. Comput. Chem.* **2015**, 36, 601.

[23] Yang, M.; Zou, J.; Wang, G.; Li, S. *J. Phys. Chem. A* **2017**, 121, 1351.

[24] Zimmerman, P. M. *J. Chem. Theory Comput.* **2013**, 9, 3043.

[25] Bergeler, M.; Simm, G. N.; Proppe, J.; Reiher, M. *J. Chem. Theory Comput.* **2015**, 11, 5712.

[26] Behn, A.; Zimmerman, P. M.; Bell, A. T.; Head-Gordon, M. *J. Chem. Phys.* **2011**, 135, 224108.

[27] Zimmerman, P. M. *J. Chem. Phys.* **2013**, 138, 184102.

- [28] Henkelman, G.; Uberuaga, B. P.; Jónsson, H. *J. Chem. Phys.* **2000**, 113, 9901.
- [29] Henkelman, G.; Jónsson, H. *J. Chem. Phys.* **2000**, 113, 9978.
- [30] Bhoorasingh, P. L.; West, R. H. *Phys. Chem. Chem. Phys.* **2015**, 17, 32173.
- [31] Schlegel, H. B. *J. Comput. Chem.* **1982**, 3, 214.
- [32] Peng, C.; Ayala, P. Y.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **1996**, 17, 49.
- [33] Schlegel, H. B. *Theor. Chem. Acc.* **1984**, 66, 333.
- [34] Zádor, J.; Najm, H. N. *Kinbot 1.0: A code for automatic PES exploration*; SAND2012-8095; 2013.
- [35] Zheng, J.; Mielke, S. L.; Clarkson, K. L.; Truhlar, D. G. *Comput. Phys. Commun.* **2012**, 183, 1803.
- [36] Zheng, J.; Meana-Pañeda, R.; Truhlar, D. G. *Comput. Phys. Commun.* **2013**, 184, 2032.
- [37] Yu, T.; Zheng, J.; Truhlar, D. G. *Chem. Sci.* **2011**, 2, 2199.
- [38] Wei, J. N.; Duvenaud, D.; Aspuru-Guzik, A. *ACS Cent. Sci.* **2016**, 2, 725.
- [39] Sameera, W. M. C.; Maeda, S.; Morokuma, K. *Acc. Chem. Res.* **2016**, 49, 763.
- [40] Jensen, R. K.; Korcek, S.; Mahoney, L. R.; Zinbo, M. *J. Am. Chem. Soc.* **1979**, 101, 7574.
- [41] Jensen, R. K.; Korcek, S.; Mahoney, L. R.; Zinbo, M. *J. Am. Chem. Soc.* **1981**, 103, 1742.

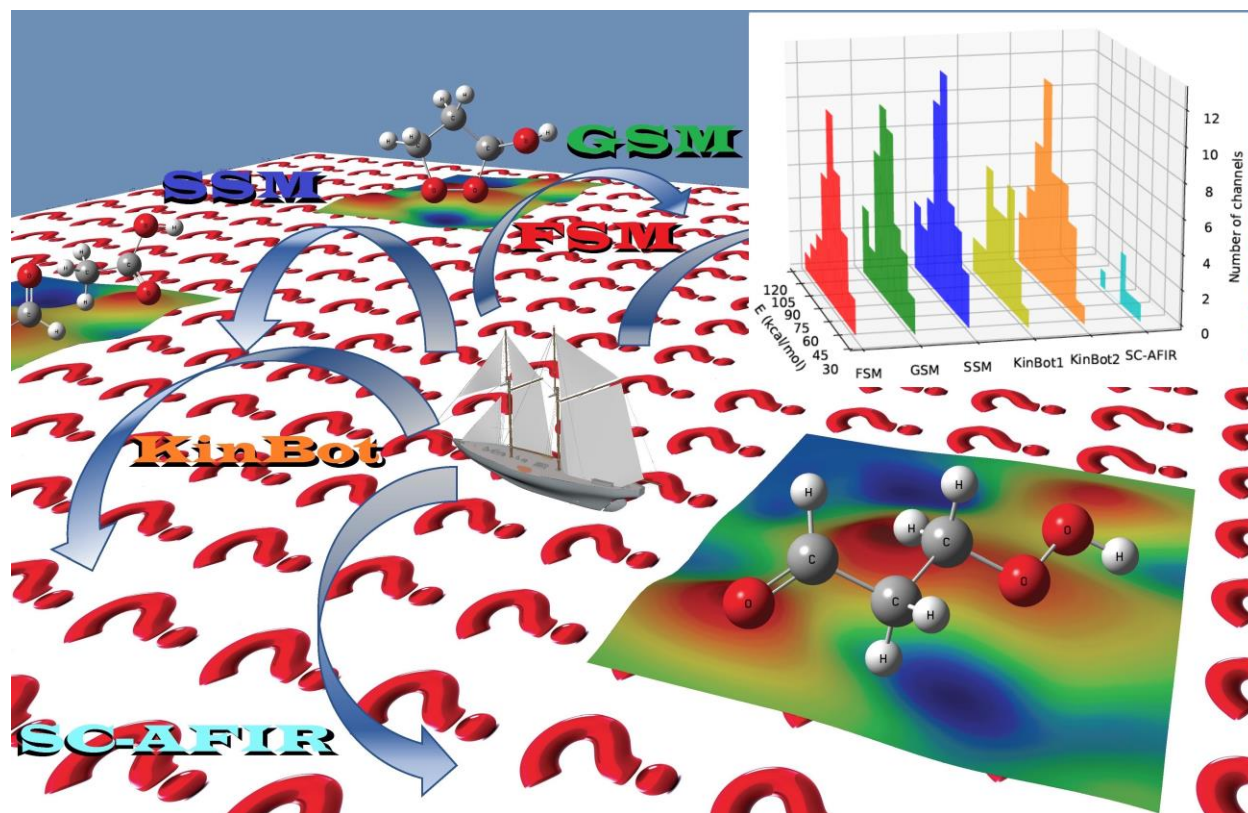
- [42] Zádor, J.; Taatjes, C. A.; Fernandes, R. X. *Prog. Energy Combust. Sci.* **2011**, *37*, 371.
- [43] Gao, C. W.; Allen, J. W.; Green, W. H.; West, R. H. *Comput. Phys. Commun.* **2016**, *203*, 212.
- [44] Benson, S. W. *Thermochemical Kinetics: Methods for the Estimation of Thermochemical Data and Rate Parameters*, 2nd ed.; Wiley: New York, 1976.
- [45] O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. *J. Cheminform.* **2011** *3*, 33.
- [46] Gaussian 09, Revision C.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.; 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, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; 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.
- [47] <https://github.com/ZimmermanGroup/molecularGSM> (accessed August 2017).
- [48] Maeda, S.; Taketsugu, T.; Morokuma, K. *J. Comput. Chem.* **2014**, *35*, 116.
- [49] Zhao, Y.; Truhlar, D. G. *Chem. Acc.* **2008**, *120*, 215.
- [50] Walker, M.; Harvey, A. J. A.; Sen, A.; Dessent, C. E. H. *J. Phys. Chem. A.* **2013**, *117*, 12590.

- [51] Vereecken, L. *Phys. Chem. Chem. Phys.* **2017**, *19*, 28630.
- [52] Welz, O.; Savee, J. D.; Osborn, D. L.; Vasu, S. S.; Percival, C. J.; Shallcross, D. E.; Taatjes, C. A. *Science* **2012**, *335*, 204.
- [53] Taatjes, C. A.; Welz, O.; Eskola, A. J.; Savee, J. D.; Scheer, A. M.; Shallcross, D. E.; Rotavera, B.; Lee, E. P. F.; Dyke, J. M.; Mok, D. K. W.; Osborn, D. L.; Percival, C. J. *Science* **2013**, *340*, 177.
- [54] Taatjes, C. A.; Welz, O.; Eskola, A. J.; Savee, J. D.; Osborn, D. L.; Lee, E. P. F.; Dyke, J. M.; Mok, D. W. K.; Shallcross, D. E.; Percival, C. J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 10391.
- [55] Tobias, H. J.; Ziemann, P. J. *J. Phys. Chem. A* **2001**, *105*, 6129.
- [56] Taatjes, C. A.; Meloni, G.; Selby, T. M.; Trevitt, A. J.; Osborn, D. L.; Percival, C. J.; Shallcross, D. E. *J. Am. Chem. Soc.* **2008**, *130*, 11883.
- [57] Andersen, A.; Carter, E. A. *J. Phys. Chem. A* **2003**, *107*, 9463.
- [58] Buras, Z. J.; Elsamra, R. M. I.; Green, W. H. *J. Phys. Chem. Lett.* **2014**, *5*, 2224.
- [59] Buras, Z. J.; Elsamra, R. M. I.; Jalan, A.; Middaugh, J. E.; Green, W. H. *J. Phys. Chem. A* **2014**, *118*, 1997.
- [60] Jalan, A.; Allen, J. W.; Green, W. H. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16841.
- [61] Taatjes, C. A.; Hansen, N.; McIlroy, A.; Miller, J. A.; Senosiain, J. P.; Klippenstein, S. J.; Qi, F.; Sheng, L. S.; Zhang, Y. W.; Cool, T. A.; Wang, J.; Westmoreland, P. R.; Law, M. E.; Kasper, T.; Kohse-Höinghaus, K. *Science* **2005**, *308*, 1887.

- [62] Heazlewood, B. R.; Maccarone, A. T.; Andrews, D. U.; Osborn, D. L.; Harding, L. B.; Klippenstein, S. J.; Jordan, M. J. T.; Kable, S. H. *Nat. Chem.* **2011**, *3*, 443.
- [63] Andrews, D. U.; Heazlewood, B. R.; Maccarone, A. T.; Conroy, T.; Payne, R. J.; Jordan, M. J. T.; Kable, S. H. *Science* **2012**, *337*, 1203.
- [64] Novelli, P. C.; Lang, P. M.; Masarie, K. A.; Hurst, D. F.; Myers, R.; Elkins, J. W. *J. Geophys. Res.* **1999**, *104*, 30427.
- [65] Schröder, D.; Schalley, C. A.; Goldberg, N.; Hrusak, J.; Schwarz, H. *Chem. Eur. J.* **1996**, *2*, 1235.
- [66] Bach, R. D.; Owensby, A. L.; Gonzalez, C.; Schlegel, H. B.; McDouall, J. J. W. *J. Am. Chem. Soc.* **1991**, *113*, 6001.
- [67] Schalley, C. A.; Harvey, J. N.; Schröder, D.; Schwarz, H. *J. Phys. Chem. A* **1998**, *102*, 1021.
- [68] Vereecken, L.; Rickard, A. R.; Newland, M. J.; Bloss, W. J. *Phys. Chem. Chem. Phys.* **2015**, *17*, 23847.
- [69] Bach, R. D.; Su, M.-D.; Aldabbagh, E.; Andrés, J. L.; Schlegel, H. B. *J. Am. Chem. Soc.* **1993**, *115*, 10237.
- [70] Yu, H.-G.; Muckerman, J. T. *J. Phys. Chem. A* **2005**, *109*, 1890.
- [71] Frankcombe, T. J.; Smith, S. C. *Faraday Discuss.* **2001**, *119*, 159.
- [72] Blanquart, G.; Pepiot-Desjardins, P.; Pitsch, H. *Combust. Flame* **2009**, *156*, 588.
- [73] Senosiain, J. P.; Klippenstein, S. J.; Miller, J. A. *J. Phys. Chem. A* **2006**, *110*, 5772.

- [74] Clubb, A. E.; Jordan, M. J. T.; Kable, S. H. *J. Phys. Chem. Lett.* **2012**, *3*, 3522.
- [75] Vereecken, L.; Novelli, A.; Taraborrelli, D. *Phys. Chem. Chem. Phys.* [Online early access]. DOI: 10.1039/C7CP05541B. Published Online: November 28, 2017. <http://pubs.rsc.org> (accessed Nov 28, 2017).
- [76] Vereecken, L.; Francisco, J. S. *Chem. Soc. Rev.* **2012**, *41*, 6259.
- [77] Stone, D.; Whalley, L. K.; Heard, D. E. *Chem. Soc. Rev.* **2012**, *41*, 6348.
- [78] So, S.; Wille, U.; da Silva, G. *Environ. Sci. Technol.* **2014**, *48*, 6694.
- [79] Novelli, A.; Hens, K.; Ernest, C. T.; Martinez, M.; Nölscher, A. C.; Sinha, V.; Paasonen, P.; Petäjä, T.; Sipilä, M.; Elste, T.; Plass-Dülmer, C.; Phillips, G. J.; Kubistin, D.; Williams, J.; Vereecken, L.; Lelieveld, J.; Harder, H. *Atmos. Chem. Phys.* **2017**, *17*, 7807.
- [80] Zhou, S.; Barnes, I.; Zhu, T.; Bejan, I.; Albu, M.; Benter, T. *Environ. Sci. Technol.* **2008**, *42*, 7905.
- [81] Fischer, S.; Karplus, M. *Chem. Phys. Lett.* **1992**, *194*, 252.
- [82] Ionova, I. V.; Carter, E. A. *J. Chem. Phys.* **1993**, *98*, 6377.

TOC Graphic



Unimolecular Reaction Pathways of a γ -Ketohydroperoxide from Combined Application of Automated Reaction Discovery Methods

SUPPORTING INFORMATION

Colin A. Grambow, Adeel Jamal, Yi-Pei Li, and William H. Green*

Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, Massachusetts 02139, United States

Judit Zádor

Combustion Research Facility, Sandia National Laboratories, 7011 East Ave., Livermore, California 94551, United States

Yury V. Suleimanov*

Computation-based Science and Technology Research Center, Cyprus Institute, 20 Kavafi Str., Nicosia 2121, Cyprus

*Email: whgreen@mit.edu; y.suleymanov@cyi.ac.cy

The Supporting Information file contains additional information on the computational procedures used for the characterization of reaction paths with string methods combined with Berny optimization and using the heuristic KinBot approach (Section S1). The same section also described the methodology for the calculation of rate constants in detail. Comparison of group additivity and DFT reaction energies for the set of filtered reactions is given in the next Section S2. It continues (Section S3) with the detailed description of 75 unimolecular reactions of γ -ketohydroperoxide detected by the freezing string method (FSM), growing string method (GSM), single-ended GSM (SSM), single-component artificial force induced reaction (SC-AFIR), and by the heuristic approach implemented in the KinBot program during the first (generic) and second (extended) calculations (see the main text for more detail). The reactions are grouped loosely into

some meaningful groups for orientation. For each reaction, the picture depicting the reaction mechanism, string paths from the FSM, GSM, SSM calculations, the picture with the geometry of the saddle points from the KinBot calculations (where available) as well as some description (intended channels for the string methods, energy barriers, reaction energies (calculated using final IRC geometries but separately if more than one fragment is radical) and the number of gradient calls (excluding the IRC calculations)) are included. Energies are reported at the B3LYP/6-31+G* level of theory without the ZPE correction. Note that small differences in the energy barriers and reaction energies occur because the located transition states connect two relevant isomers along all different reaction pathways and may be different for each method. For FSM and GSM, the intended channel is the reaction that connects reactant and product structures specified in the input; for single-ended SSM it is the reaction that contains the bond changes that correspond to the driving coordinates specified in the input; for SC-AFIR it is the reaction where the IRC calculation at the high level matches the products predicted by the low-level AFIR calculation. The geometries of the saddle points (selected for the lowest value of the energy barrier) for all 75 reactions are given below this information (in Angstrom). Next, the RMG kinetic database for the unimolecular decomposition of γ -keto hydroperoxide is shown (Section S4). The document concludes with the geometries, energies, and frequencies of all intermediates in the Criegee + vinyl alcohol reaction at the RCCSD(T)-F12a/cc-pVTZ-F12//UM06-2X/aug-cc-pVTZ level of theory (Section S5).

S1. COMPUTATIONAL PROCEDURE

S1.1 Combinatorial Search Using String Methods and Berny Optimization

In order to mathematically describe the original system and reactions that can occur (as well as the expected products), our approach utilized the Bond Electron (BE) matrix representation of species [S1,S2]. In the present study, we considered only breaking and forming a maximum of three bonds and only allowed products with permissible valences (e.g., maximum of four for carbon). In addition, the breaking of double (or higher order) bonds was not permitted. Products were only added to the set of new products if their connectivity was different from all other products and if they were not isomorphic with the reactant molecule (i.e., conformational changes were not counted as reactions). This procedure resulted in a set of 4324 possible product structures.

Initial geometries of the reactant and product(s) were generated by constructing linear connections based on rules corresponding to the hybridization of atoms and by constructing ring structures from templates as implemented in the Open Babel program [45]. The energies of generated structures were minimized using the MMFF94 force field also implemented in the program. When the product channels contained two or three fragments, they were translated and rotated in space in order to maximize the overlap with the initial reactant molecule while constrained by a minimum distance from each other. The dihedral angles of rotatable bonds were also modified to further maximize the overlap.

In the FSM case, we used the same default spacing parameters as in our previous study [17] (20 nodes, 6 perpendicular gradients per node). Our previous work was performed with the Q-Chem package [S3]. For this work, we rewrote the FSM algorithm, included more termination criteria checks and interfaced it with Gaussian 09. We also added one restarting option which we found to provide more stability to the FSM calculations — when the electronic structure calculation failed, the previous gradient was used for a given node. This was not done if the node had just been generated in the interpolation step of the FSM algorithm where a previous gradient calculation was not yet available. We also set the maximum number of nodes to 40 in order to terminate the FSM calculation when it becomes stuck and ends up oscillating back and forth between geometries.

In the GSM and SSM cases, we also used standard spacing parameters and settings (100 maximum iterations, exactly 11 nodes for GSM, a maximum of 30 nodes for SSM with a minimum node spacing of 1.0 Å) and, as indicated in the main text, used the program developed by one of the authors of these methods [47].

The first reactive peak along the string path was selected as an initial guess TS structure provided that at least one bond changed in the structure, which was determined using automatic identification of bonds as implemented in Open Babel [45].

While the GSM and SSM code already incorporates a Hessian-free exact saddle point search, we nevertheless re-optimized the structures with the Berny algorithm for consistency across all methods. This was required anyway for GSM paths where the first reactive barrier closest to the

reactant was not the maximum barrier in the entire GSM path, since the GSM code only completes an exact saddle point search for the highest energy structure.

S1.2 Heuristic KinBot Approach

When constructing the corresponding 3D structures for the reaction types in Figure 1 in the main text, moving the atoms in the reactive center at once to their desired position is not feasible in most cases, because it causes the other, spectator atoms to clash, which, in turn leads to the rapid fragmentation or chaotic rearrangement of the structure in the subsequent optimization steps. To achieve these generic, but prescribed steric structures in the reactive center of the molecule, KinBot manipulates some large-amplitude motions, typically involving rotations around the appropriate dihedral angles, in a stepwise fashion. In each step, a constrained optimization is carried out at a very cheap level of theory to lead the structure approximately along a MEP-like valley.

KinBot calculations used the Berny optimization with the CalcFC keyword of Gaussian.

S1.3 Rate Constant Calculation Methodology

Geometries and frequencies were calculated at the UM06-2X/aug-cc-pVTZ level of theory with an ultrafine integration grid and tight SCF convergence criteria. Energies were calculated using the RCCSD(T)-F12a/cc-pVTZ-F12 method. 1D hindered rotor scans were completed for all rotatable bonds to enable accurate calculation of relevant partition functions. Cantherm [43] was used for the rate constant calculations. Frequencies were scaled by a factor of 0.971, consistent with Ref. [S4]. For pressure-dependent calculations, the one-dimensional master equation was solved using the Reservoir State approximation employing an exponential down collisional energy-transfer model with nitrogen as the bath gas and using an average downward collisional energy transfer of 150 cm^{-1} [63] and an active *K*-rotor and inactive *J*-rotor. Lennard-Jones parameters were estimated using the RMG website (rmg.mit.edu). The rate of formation of the van der Waals complex was not assumed to affect the overall rate, the justification for which is shown in a more in-depth analysis in Ref. [S5]. The rate for the barrierless dissociation of KHP to the two radical products was estimated from an analogous reaction of ethyl hydroperoxide [S6].

S2. COMPARISON BETWEEN GROUP ADDITIVITY AND DFT

Figure S1 below compares group additivity and DFT (B3LYP/6-31+G*) reaction energies for the set of filtered reactions. For most reactions, the energies are very similar and lie along the line of equality. However, there are several reactions for which the DFT energy is much larger, which either correspond to carbenes, for which group additivity underestimates the energy because it assumes the triplet state (although the relatively low DFT level may also not be accurate for such species), or to fused bicyclic structures, for which group additivity severely underestimates the energy. Note that bicyclic corrections have now become available for RMG (which was the software used to estimate the group additivity values). There also exists a cluster of points for which the group additivity energy estimate is larger than DFT, but common features among these reactions are not apparent. It is likely that most reactions are either well represented by group additivity estimates or that group additivity underestimates the energy, which would lead to fewer reactions being included in the filtered set.

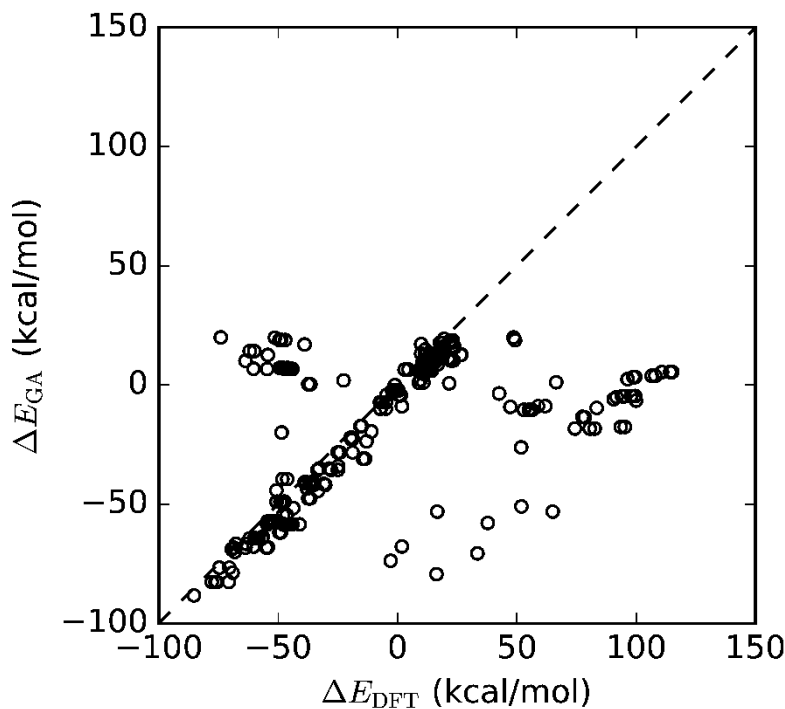


Figure S1. Comparison of group additivity and DFT reaction energies for the set of filtered reactions.

S3. UNIMOLECULAR REACTION PATHWAYS OF γ -KETOHYDROPEROXIDE

S3.1 Geometries of the reactant

The geometry of the reactant relative to which the energies are reported for KinBot (in Angs):

C	0.058961	0.149964	-0.044295
H	0.159216	0.260637	1.038365
O	1.393791	0.102392	-0.496141
C	-0.730849	-1.116942	-0.414583
H	-0.448119	1.030953	-0.448343
H	-0.963575	-1.095480	-1.480073
C	0.105016	-2.338178	-0.130474
H	-1.662572	-1.139229	0.154968
O	0.931415	-2.750953	-0.902996
H	-0.031952	-2.839205	0.846227
O	1.370385	-0.043900	-1.907537
H	1.548849	-0.994322	-2.009415

For FSM, we used the initially optimized reactant structure. It is slightly different for every FSM channel, because we use different starting conformers. For GSM and SSM, we use the first node along the string, which will also be a slightly different optimized reactant for each channel.

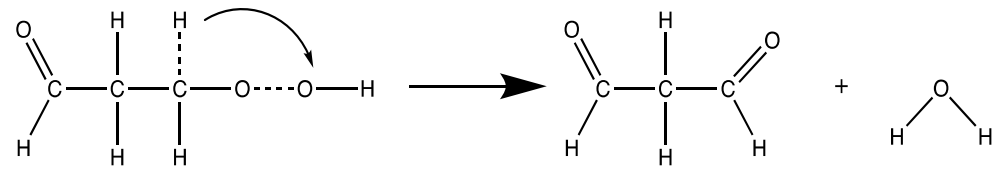
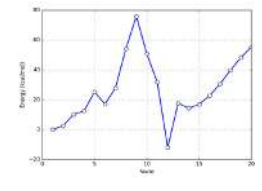
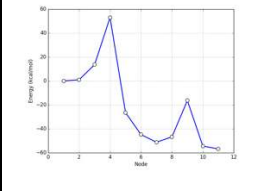
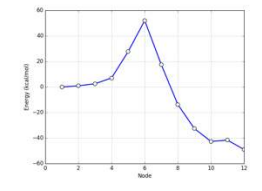
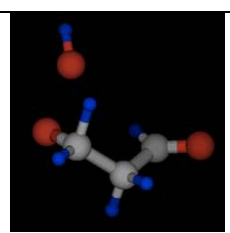
For SC-AFIR, we use the following geometry:

C	4.412143000	0.122278000	0.267626000
C	2.900718000	0.194683000	0.238405000
C	2.283613000	-0.632536000	1.370329000
H	4.841989000	-0.902697000	0.216882000
H	2.585301000	1.240627000	0.287643000
H	2.563831000	-0.230629000	-0.719189000
H	2.457727000	-0.165551000	2.348467000
H	2.687981000	-1.653518000	1.386840000
O	0.885700000	-0.684332000	1.095460000
O	0.294780000	-1.521513000	2.133243000

H -0.351658000 -0.897455000 2.510983000
 O 5.144192000 1.086353000 0.332597000

S3.2 Detected Chemical Reactions

- **H₂O + Malonaldehyde channels**

<p>R1</p> 				
1,2-H ₂ O elimination on the OOH group				
FSM path	GSM path	SSM path	SC-AFIR	KinBot 1 and 2
				
Intended channel				
[CH2]C=O.[CH][O].O	O=CCC(O)O	[CH2]C=O.[CH][O].O	O=CCC=O.O (intended)	N/A
Energy barrier, kcal/mol				
51.0	53.1	51.0	52.8	52.6
Reaction energy, kcal/mol				
-49.5	-49.5	-49.5	-49.5	-50.0
Number of gradient calls				
244	594	253		

TS geometry (in Angs):

O 0.66532400 -2.20938100 -1.05670400
 O 0.44413200 -0.29327500 -1.41417100
 C -0.47089000 -0.29085600 -0.47414800
 C -0.11652500 0.40104400 0.84707000
 C 0.03907200 1.89917800 0.63491500
 O -0.45812100 2.74230500 1.34574800
 H 1.17528100 -3.00130200 -1.35483100

H	-1.50500000	-0.16024200	-0.84224800
H	-0.42055600	-1.49654800	-0.24972700
H	-0.86861400	0.21312300	1.61968200
H	0.85474600	0.01522900	1.18708800
H	0.66114400	2.18072300	-0.24267800

1,3-H ₂ O elimination on the OOH group and H-shift			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
<chem>C1O[C@@H]2[C@H]1O2.O</chem>	<chem>C1O[C@@H]2[C@H]1O2.O</chem>	<chem>[O][C]CC[O].O (intended)</chem>	
Energy barrier, kcal/mol			
59.8	60.0	60.3	59.5
Reaction energy, kcal/mol			
-49.3	-49.5	-49.5	-50.0
Number of gradient calls			
271	566	489	

TS geometry (in Angs):

O	1.70521100	-1.05153200	-0.25277800
O	0.23394600	-0.86986100	-1.38576700
C	-0.75791400	-0.79572700	-0.43393300
C	-0.40412900	0.10660000	0.70657000
C	-0.24110200	1.53789600	0.48469200
O	-0.13034200	2.38031800	1.37000000

H	2.21326000	-0.46526600	-0.84276800
H	-1.54774400	-0.31854300	-1.06481100
H	-1.11645100	-1.78162500	-0.10875400
H	-0.79185700	-0.16731900	1.68749300
H	1.01462300	-0.42614400	0.42280200
H	-0.17750900	1.85119700	-0.58275400

1,2-H ₂ O elimination on the OOH group with H “roaming” around O in O-OH before connecting to another O atom in the OH group.			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
O[CH]CC([O])O		[CH2]C(O)O.[O][CH]	
Energy barrier, kcal/mol			
67.1		68.0	68.6
Reaction energy, kcal/mol			
-49.4		-49.5	-50.0
Number of gradient calls			
298		304	

TS geometry (in Angs):

O	-0.84618600	-1.43197500	1.89144700
O	-1.32786700	-0.28108500	0.49214800
C	-0.55377300	-0.56038900	-0.55172500
C	0.86272000	-0.07147500	-0.52441900
C	0.93862600	1.44472400	-0.43772700
O	-0.05358500	2.14040500	-0.33165700
H	-1.70325500	-1.49604300	2.34988300

H	-0.82452600	-1.41839400	-1.17103300
H	-1.23347900	0.66739800	-0.22660500
H	1.36089300	-0.47740300	0.37374800
H	1.43688300	-0.41862700	-1.39310300
H	1.94354000	1.90285800	-0.47096400

<p>R4</p>			
<p>1,2-H₂O elimination on the OOH group with H "roaming" around O of the CHO group before connecting to another O atom in the OH group.</p>			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
	OO[C@H]1C[C@@H]1O		
Energy barrier, kcal/mol			
	63.0		63.0
Reaction energy, kcal/mol			
	-49.4		-50.0
Number of gradient calls			
	748		

TS geometry (in Angs):

O	1.77931500	-0.48719100	-1.37872800
O	0.90923600	-1.50188000	-0.06847500
C	-0.26229800	-0.99539200	0.05094700
C	-0.47767600	0.07451800	1.11786300
C	-0.45918100	1.39388700	0.40327000
O	-0.32409100	1.36720200	-0.83112000
H	2.16328000	-1.19276800	-1.92973000
H	-0.17387400	0.18952200	-0.99693700

H -1.12914200 -1.54018700 -0.34774700
H -1.42952000 -0.01188100 1.66670000
H 0.32310200 0.07159500 1.87291800
H -0.55829400 2.37155800 0.88212300

1,2-H ₂ O elimination which starts as 1,1-H ₂ O ₂ elimination.			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
Energy barrier, kcal/mol			
			89.8
Reaction energy, kcal/mol			
			-50.0
Number of gradient calls			

TS geometry (in Angs):

C -0.369647 0.119291 0.106689
H -0.019041 -0.051363 1.132672
O 1.274209 0.356730 -0.556567
C -0.848660 -1.170198 -0.514362
H 0.522782 1.091402 -0.578915
H -0.945197 -1.087541 -1.597415
C 0.089056 -2.275435 -0.117935
H -1.845286 -1.333281 -0.089282
O 1.101310 -2.533910 -0.728677
H -0.118778 -2.796994 0.834665

O 1.373603 -0.065840 -1.901990
H 1.516217 -1.027123 -1.783180

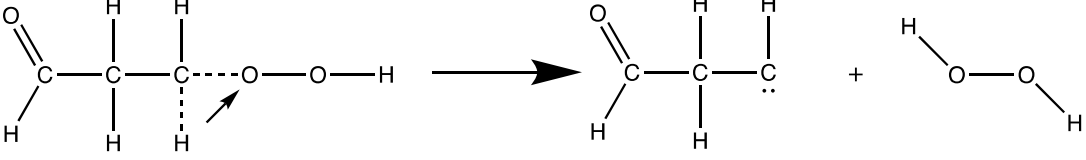
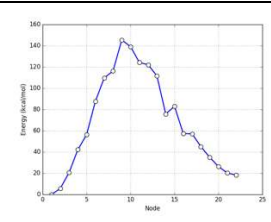
• **Biradical products (including carbenes)**

H-shift forming carbene			
FSM path	GSM path	SSM path	KinBot 2 only
Intended channel			
OOC/C=C/O	OOC/C=C/O		
Energy barrier, kcal/mol			
85.8	87.0		85.8
Reaction energy, kcal/mol			
52.5	55.1		52.0
Number of gradient calls			
211	485		

TS geometry (in Angs):

O 2.14184900 -1.21727900 -0.09977600
O 0.80328200 -1.76858700 -0.29888800
C -0.05092500 -0.67684000 -0.62365900
C -0.22748400 0.29651600 0.54734400
C -1.05893900 1.48285500 0.19394000
O -1.88088700 1.93880800 1.10909500
H 2.63252900 -1.72622100 -0.77005800
H 0.31988200 -0.15650500 -1.51501700
H -0.99484500 -1.17714000 -0.87009200
H -0.64147700 -0.19916200 1.43882400

H 0.76526000 0.69677500 0.80862100
H -1.80825700 2.50677000 0.07965800

R7			
H ₂ O ₂ elimination forming carbene			
FSM path	GSM path	SSM path	KinBot
			
			
Intended channel			
C=CC=O.OO			
Energy barrier, kcal/mol			
89.0			
Reaction energy, kcal/mol			
86.5			
Number of gradient calls			
316			

TS geometry (in Angs):

O 0.13397400 -2.01108500 2.30350900
O 0.82977600 -1.49386000 1.14522300
C -0.51993200 0.36514700 0.47810700
C -0.60356200 0.50043800 -0.98270100
C 0.46521800 1.40354300 -1.59129200
O 0.24506000 2.16250600 -2.50612500
H 0.84058600 -1.99163400 2.97425900
H -1.48848600 0.02769400 0.87854600
H 0.56376300 -0.45782900 1.14447400
H -0.34315900 -0.53591300 -1.30104700
H -1.59219800 0.70963400 -1.41675200

H 1.46895400 1.32135000 -1.12620700

O- and (CH)- exchange forming carbene			
FSM path	GSM path	SSM path	KinBot 2 only
Intended channel			
[CH]OCCOO (intended)	O1CCOOC1	[CH]OCCOO (intended)	
Energy barrier, kcal/mol			
117.7 (the same mechanism but higher- energy conformer)	108.6	111.5	112.0
Reaction energy, kcal/mol			
65.3	62.2	65.3	64.8
Number of gradient calls			
455	467	501	

TS geometry (in Angs):

O 1.71751300 -0.48057600 -1.29594500
O 1.16862700 -1.31356900 -0.23631500
C -0.23055700 -1.12802200 -0.20963200
C -0.67716900 0.06031900 0.60706000
C 0.22197300 1.57420900 0.06606900
O -1.00129800 1.50670100 -0.36711000
H 1.52492200 0.41937500 -0.90954900
H -0.63888700 -1.07727800 -1.22517700
H -0.62856900 -2.02651700 0.28687900
H -1.74956800 0.07728500 0.83216200

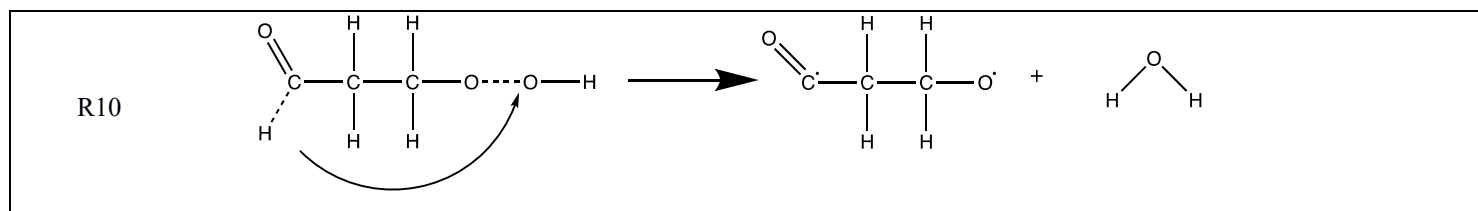
H -0.14996200 0.21513800 1.56297100
 H 0.31664900 2.35163500 0.87501300

C-C and C-O bonds breaking forming carbene			
FSM path	GSM path	SSM path	KinBot
Intended channel			
[CH]OCCOO (intended)	[CH]OCCOO (intended)	[CH]OCCOO (intended)	
Energy barrier, kcal/mol			
106.8	106.8	107.4	
Reaction energy, kcal/mol			
66.5	66.5	66.5	
Number of gradient calls			
142	469	608	

TS geometry (in Angs):

O -1.18223700 -2.06125300 -0.46150400
 O -0.80469400 -1.43233900 0.80229900
 C -0.65263500 0.66049700 0.06699800
 C 0.58421100 -0.14997800 0.29368000
 C 1.29349700 1.50094800 -0.32710300
 O 0.11408000 1.95580100 -0.44258100
 H -0.88408900 -2.97385100 -0.30998700
 H -1.29016500 0.35411200 -0.75722600
 H -1.20191800 0.98367900 0.94626200
 H 1.02755500 -0.23830500 1.27872600

H 0.93720500 -0.84292100 -0.46039900
H 2.05918100 2.24360000 -0.62917300



1,4- H₂O elimination forming biradical (highly unstable intermediate structure)

FSM path	GSM path	SSM path	SC-AFIR	KinBot
Intended channel				
O[CH]CC(=O)O.[H]	[CH2][C][O].C=O.O	OCCC(=O)O	OCCC(=O)O	
Energy barrier, kcal/mol				
63.4	63.7	63.7	50.1 (different conformer. Mechanism is the same)	
Reaction energy, kcal/mol*				
14.9	14.8	14.0	14.9	
Number of gradient calls				
254	691	1071		

* unstable fragment (kept frozen without optimization)

TS geometry (in Angs):

C 4.080394000 -0.009626000 0.521297000
C 3.051457000 -0.770442000 -0.397120000
C 2.063840000 -1.683323000 0.382926000
H 3.221641000 0.700551000 1.318292000
H 2.470733000 -0.009089000 -0.926986000
H 3.660499000 -1.343337000 -1.108732000
H 2.618271000 -2.348817000 1.068157000
H 1.656194000 -2.311574000 -0.442212000

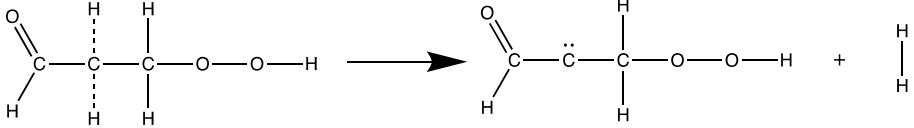
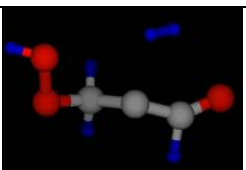
O	0.980705000	-1.091284000	0.930051000
O	2.120522000	0.371665000	1.721216000
H	1.826720000	0.813800000	2.548896000
O	5.231428000	-0.229621000	0.667761000

<p>R11</p>			
1,4- H ₂ elimination forming biradical			
FSM path	GSM path	SSM path	KinBot 2 only
Intended channel			
	<chem>O1C[C@H]2[C@@H](O1)O2.[H][H]</chem>		
Energy barrier, kcal/mol			
	109.6		91.5
Reaction energy, kcal/mol			
	87.1		86.2
Number of gradient calls			
	911		

TS geometry (in Angs):

C	0.228708	0.031315	0.249360
H	0.480921	-0.003899	1.311367
O	1.508215	-0.267462	-0.473602
C	-0.538257	-1.102720	-0.277116
H	-0.106978	1.020790	-0.060070
H	0.002378	-1.001107	-2.436036
C	-0.287601	-2.454116	0.203224
H	-1.374212	-0.950577	-0.949645

O -1.018397 -3.398822 -0.071840
H 0.594737 -2.592583 0.858631
O 1.688811 0.431993 -1.566015
H 0.552242 -0.487073 -2.582554

R12			
1,1-H ₂ elimination forming carbene			
FSM path	GSM path	SSM path	KinBot 1 and 2
			
Intended channel			
Energy barrier, kcal/mol			
			100.6
Reaction energy, kcal/mol			
			98.4
Number of gradient calls			

TS geometry (in Angs):

C 0.633637 -0.242973 -0.009680
H 1.061647 -0.253414 1.001434
O 1.628161 0.364456 -0.804427
C 0.317311 -1.628207 -0.372020
H -0.262857 0.390185 0.014234
H -0.869730 -1.367852 -1.813818
C -0.427603 -2.481553 0.451620
H -1.489726 -1.604926 -1.449101
O -1.528907 -2.980731 0.418611
H 0.336476 -2.788649 1.224157

O 1.062528 0.452941 -2.109127
H 1.269629 1.366461 -2.346181

CH ₃ OOH + HC-CHO channel			
FSM path	GSM path	SSM path	KinBot 2 only
Intended channel			
Energy barrier, kcal/mol			
			92.8
Reaction energy, kcal/mol			
			88.7
Number of gradient calls			

TS geometry (in Angs):

C -0.001432 0.607883 -0.007442
H -0.086969 0.835666 1.059928
O 1.362383 0.491690 -0.283967
C -0.259986 -1.978870 -1.129425
H -0.492475 1.364013 -0.629781
H -1.279120 -1.968216 -1.527421
C 0.011106 -2.839117 -0.044247
H -0.525318 -0.376743 -0.200709
O -0.000688 -3.997871 -0.493566
H 0.349670 -2.571142 0.968343
O 1.489160 0.298508 -1.713195

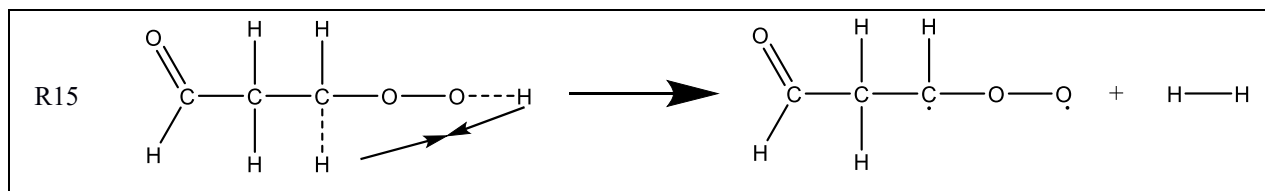
H 1.164235 -0.640063 -1.792814

C-C break and H-shift forming carbene peroxide			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
CC(=O)COO	COOC(=O)C	OO[C@H]1C[C@H]1O	
Energy barrier, kcal/mol			
71.0	75.2	69.4 (mechanism is the same for FSM/GMS/SSM. Different conformers)	70.3 68.6
Reaction energy, kcal/mol			
54.5	54.8	54.3	54.3
Number of gradient calls			
295	2104	1173	

TS geometry (in Angs):

O	1.41073000	0.49076700	1.77891800
O	0.34443600	-0.78969200	1.52093400
C	0.51087800	-1.28039700	0.38701800
C	-0.61304700	-0.12155600	-1.29068800
C	-0.80373100	1.13668900	-0.55927500
O	-1.83625200	1.41278700	0.03751000
H	1.17602900	0.66842000	2.71046500
H	1.48567900	-1.27353000	-0.10034900
H	-0.41809400	-1.29636700	-0.33904900

H 0.20770700 -0.16882300 -2.00503900
 H -1.53708800 -0.60550700 -1.61506100
 H 0.06095600 1.82733200 -0.53661200



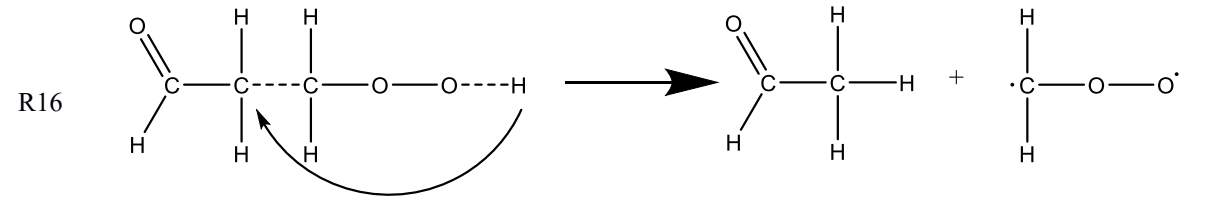
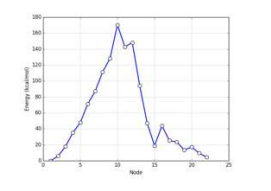
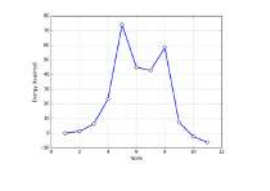
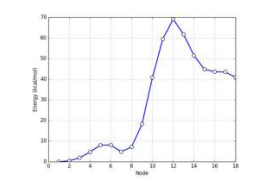
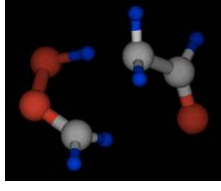
1,3-H₂ elimination from OOH forming 3-carbon Criegee

FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
O=C1CCOO1.[H][H]	O=C1CCOO1.[H][H]	COOC(=O)C	
Energy barrier, kcal/mol			
69.9	67.4	70.0	69.5
Reaction energy, kcal/mol			
52.1	52.1	51.3	51.6
Number of gradient calls			
349	1063	682	

TS geometry (in Angs):

O 1.05937500 0.76143000 1.70833600
 O 1.57381400 0.58342600 0.43551000
 C 0.60012100 0.50872800 -0.43830000
 C -0.60773100 -0.37006100 -0.25337300
 C -0.26441300 -1.81623800 -0.62321700
 O -0.81960300 -2.42522400 -1.50949400
 H 0.15068400 1.78815400 1.09568100
 H -0.09219900 1.90238600 0.25829500
 H 0.97496700 0.66432500 -1.45135800
 H -0.92157900 -0.34281300 0.79456300
 H -1.42209700 -0.03027700 -0.89906800

H 0.53219600 -2.28996400 -0.01521300

			
C-C break and H-shift forming 1-carbon Criegee			
FSM path	GSM path	SSM path	KinBot 1 and 2
			
Intended channel			
COOCC=O	CC(=O)COO	C[C@@H](C=O)OO	
Energy barrier, kcal/mol			
65.6	65.6	67.1	65.1
Reaction energy, kcal/mol			
42.7	42.7	42.7	42.2
Number of gradient calls			
277	1016	1021	

TS geometry (in Angs):

O	1.90214300	-0.68360100	-0.13521500
O	1.06412500	-1.74646100	-0.52981500
C	-0.05154200	-1.27662500	-0.96810800
C	-0.28295900	0.35192600	0.95991100
C	-0.52653100	1.54971600	0.16592900
O	-1.18525600	1.57162600	-0.87493000
H	1.05245000	-0.11636500	0.43000600
H	-0.12455300	-0.30888400	-1.45719000
H	-0.83158700	-2.02880500	-1.06153800
H	-1.14827200	-0.31124600	1.01810900
H	0.14548400	0.52156600	1.95363200
H	-0.01351800	2.47713600	0.49919300

R17			
<p>H-shift and C-C break forming vinyl alcohol + CH₂OO (Criegee)</p> <p>N.B.: The dissociation leads initially to a H-bonded complex of the two fragments. This weakly bound complex is 0.7 kcal/mol lower than the saddle point (0.8, if ZPE is included). The asymptote leading to the fully separated products from this weakly bound complex is barrierless.</p>			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
Energy barrier, kcal/mol			
			48.7
Reaction energy, kcal/mol			
			48.0(56.5)
Number of gradient calls			

TS geometry (in Angs):

C	0.358722	0.563940	0.008549
H	0.377703	0.669909	1.090287
O	1.461740	0.210036	-0.514582
C	-0.972273	-1.424904	-0.457538
H	-0.404829	1.022061	-0.609397
H	-1.029475	-0.910710	-1.411585
C	0.024167	-2.337497	-0.245812
H	-1.865075	-1.466177	0.157878
O	1.161826	-2.408433	-0.907359

H -0.034128 -3.072205 0.557002
 O 1.374104 -0.012345 -1.928637
 H 1.278085 -1.607938 -1.533101

• Products with two charged centers

O and OH exchange in the OOH group forming zwitterionic trivalent O structure			
FSM	GSM	SSM	KinBot 2 only
Intended channel			
<chem>C[C@H](C=O)OO</chem>	<chem>[CH]C[C@@H](OO)O</chem>	<chem>O=CCC(O)O</chem>	
Energy barrier, kcal/mol			
78.3	75.4	78.3	75.4
Reaction energy, kcal/mol			
41.0	36.5 (another conformer)	41.0	40.5
Number of gradient calls			
347	451	346	

TS geometry (in Angs):

O -1.23065800 -0.74660600 -1.19450000
 O -0.14993700 0.04103200 -1.93961800
 C 0.37352600 -0.87603700 -0.33096200
 C 1.01862800 0.35379800 0.25118500
 C 0.04379300 1.41216600 0.71981300
 O -1.16573100 1.28569300 0.75060800
 H -1.67128100 -0.02674700 -0.68785200

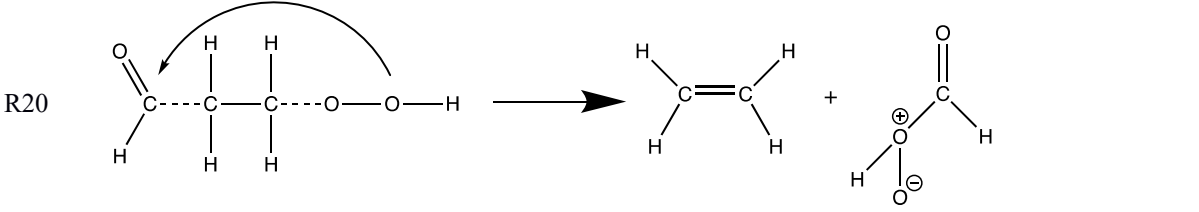
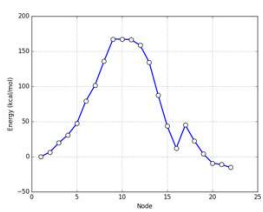
H -0.14055500 -1.45864200 0.44234900
 H 1.01118500 -1.50786400 -0.93701300
 H 1.59747600 0.05436600 1.14747200
 H 1.72234600 0.81983800 -0.44657900
 H 0.50906800 2.35530500 1.07031000

<p>R19</p>				
H-shift forming a zwitterionic trivalent O structure				
FSM path	GSM path	SSM path	SC-AFIR	KinBot 1 and 2
Intended channel				
C=CC=O.OO	C=CC=O.OO	O[C@@H](C=O)CO	OCCC=O.[O] (intended)	
Energy barrier, kcal/mol				
51.7	53.0	51.4	51.5	52.7
Reaction energy, kcal/mol				
41.0	41.0	41.1	41.0	40.6
Number of gradient calls				
252	514	1104		

TS geometry (in Angs):

O 1.96584300 -1.41267300 0.78840300
 O 0.55065300 -1.78572700 0.14384600
 C 0.04392500 -0.60558800 -0.53305100
 C -0.34732200 0.41919900 0.51685500
 C -0.89181000 1.69123700 -0.08118900
 O -1.01879000 1.88944400 -1.27301300
 H 1.47880200 -2.14713500 -0.23401700
 H 0.81007700 -0.22406400 -1.21129600

H -0.81476200 -0.96396200 -1.10948800
H -1.10161100 0.01011500 1.20556400
H 0.52713300 0.66014100 1.13761500
H -1.18528600 2.47185800 0.65219100

			
C-C and C-O bond breaking forming ethane and zwitterionic trivalent O structure (unstable intermediate structure)			
FSM path	GSM path	SSM path	KinBot
			
Intended channel			
CCOOC=O			
Energy barrier, kcal/mol			
70.1			
Reaction energy, kcal/mol*			
61.5			
Number of gradient calls			
264			

* unstable fragment (kept frozen without optimization)

TS geometry (in Angs):

O 0.29712300 0.79256800 -1.42087700
O 0.51048800 -0.59803800 -1.69022800
C -0.33015900 -1.41676500 0.19851800
C -0.43556500 -0.36252100 1.06886100
C 0.61335300 1.28410300 0.02858600
O 0.13537100 2.33812700 0.34025400
H -0.63962700 1.00709000 -1.62160600

H	-1.16478600	-1.73122400	-0.41780800
H	0.51615400	-2.09459700	0.22409900
H	-1.37704800	0.16829000	1.18504900
H	0.26797800	-0.25492200	1.89102800
H	1.60670000	0.86787500	0.21411400

R21			
H-shift and C-O break forming a zwitterionic trivalent O structure			
FSM	GSM	SSM	KinBot
Intended channel			
	<chem>O[C@@H](C=O)CO</chem>	<chem>C[C@H](C=O)OO</chem>	
Energy barrier, kcal/mol			
	80.5	80.4	
Reaction energy, kcal/mol			
	41.9	41.9	
Number of gradient calls			
	1093	678	

TS geometry (in Angs):

O	1.78829900	-0.42990500	-0.25706000
O	1.16216200	-1.45169900	-0.96276900
C	-0.77599600	-0.92856400	-0.15530600
C	-0.23158100	0.09924900	0.67500200
C	-0.16281300	1.50419100	0.24120500
O	-0.63263000	1.92202100	-0.80885200
H	1.93848100	0.26178700	-0.93348600

H -1.09094900 -0.69287000 -1.16455100
H -0.81150700 -1.95188700 0.19473800
H -1.62700200 -0.35197800 0.55178700
H 0.07264900 -0.15301600 1.68619400
H 0.37704000 2.17711800 0.93484200

H-shift and C-O break forming a zwitterionic trivalent O structure			
FSM path	GSM path	SSM path	KinBot
Intended channel			
	C[CH][C][O].OO		
Energy barrier, kcal/mol			
	100.0		
Reaction energy, kcal/mol			
	26.7		
Number of gradient calls			
	663		

TS geometry (in Angs):

O -0.04015800 0.66436100 0.12045500
O 0.30911800 -0.56159500 0.67736200
C -1.84208300 -1.38153700 0.89983200
C -2.30458000 0.00516000 1.34445400
C -2.09343100 0.48620000 -0.06414200
O -2.64412200 1.21156800 -0.83921700

H 0.25964900 0.61800300 -0.81049800
H -2.60665000 -1.97366300 0.39294300
H -1.14657000 -1.96367700 1.48900600
H -3.33636600 0.08407300 1.70184200
H -1.61638800 0.47344800 2.04926900
H -1.55007700 -0.67471600 -0.39847200

H-shift and C-C and C-O break forming CO + ethane + a zwitterionic trivalent O structure			
FSM path	GSM path	SSM path	KinBot
Intended channel			
		CCC(=O)OO	
Energy barrier, kcal/mol			
		85.6	
Reaction energy, kcal/mol			
		70.4	
Number of gradient calls			
		2353	

TS geometry (in Angs):

O 1.82421800 0.78758100 -0.21500500
O 1.60457600 -0.33338700 0.73285200
C -0.07834100 -1.27368700 0.22142600
C -1.15746800 -0.43015500 -0.02532100
C -0.53245100 1.69094100 0.13060400
O -1.45458000 2.36050500 0.32643500
H 1.93954300 0.36826200 -1.08851500

H 0.41946800 -1.77999300 -0.60156800
H 0.03183900 -1.76561100 1.18462800
H -1.51968800 -0.31143700 -1.04520300
H -1.89013300 -0.26698700 0.76000400
H 0.80930200 1.32983200 -0.12749900

C-O break forming a zwitterionic trivalent O cyclic structure			
FSM path	GSM path	SSM path	KinBot
Intended channel			
		<chem>OO[C@H]1CCO1</chem> (almost intended)	
Energy barrier, kcal/mol			
		94.3	
Reaction energy, kcal/mol			
		56.4	
Number of gradient calls			
		1600	

TS geometry (in Angs):

O -1.22842100 0.79662400 -0.48441500
O -1.12253900 -0.51289700 -1.06685200
C 0.72571000 -1.31354400 0.30581300
C 0.10936600 -0.39745200 1.34778800
C 0.45800000 0.82623100 0.54172300

O 1.32251500 0.46198500 -0.31924600
 H -1.01029000 1.37298900 -1.24368500
 H 0.10098100 -2.00736000 -0.23834700
 H 1.77712300 -1.56892200 0.38941700
 H -0.95661100 -0.55168000 1.51136300
 H 0.65398900 -0.42378100 2.30150800
 H 0.36054400 1.87850100 0.82677900

O-O break forming a zwitterionic trivalent O cyclic structure via triangle transition state			
FSM path	GSM path	SSM path	KinBot 2 only
Intended channel			
		OO[C@H]1CCO1 (almost intended)	
Energy barrier, kcal/mol			
			101.8
Reaction energy, kcal/mol			
			56.2
Number of gradient calls			

TS geometry (in Angs):

C -0.470352 0.401171 -0.261480
 H -0.378960 0.939199 0.698403
 O 0.717474 0.395901 -1.015481
 C -0.694921 -1.115714 -0.068698
 H -1.306756 0.870092 -0.811232

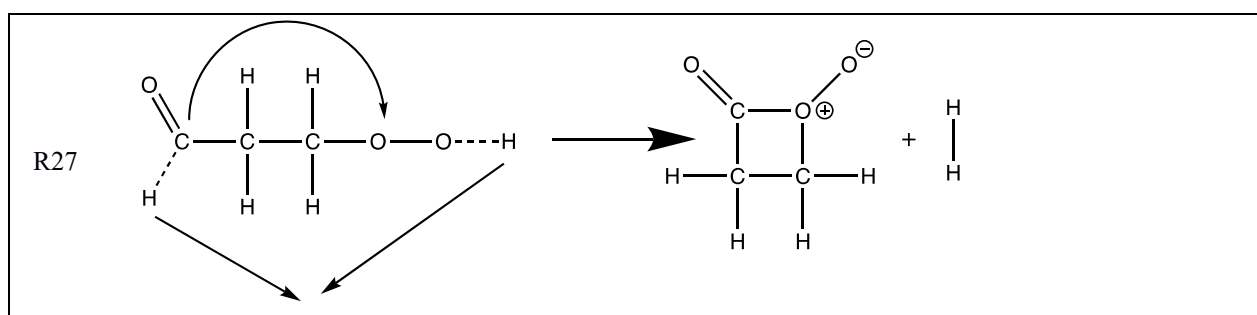
H -1.647831 -1.534049 -0.416080
 C 0.439391 -1.559688 -0.953909
 H -0.521495 -1.473733 0.951185
 O 1.017817 -3.004372 -0.607109
 H 0.414583 -1.548557 -2.032940
 O 1.726692 -1.641103 -0.302386
 H 2.434925 -1.503409 -0.974569

2 H-shifts forming a zwitterionic trivalent O cyclic structure			
FSM path	GSM path	SSM path	KinBot 2 only
Intended channel			
		COOCC=O	
Energy barrier, kcal/mol			
		78.7	78.4
Reaction energy, kcal/mol			
		36.5	36.1
Number of gradient calls			
		818	

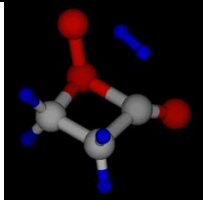
TS geometry (in Angs):

O 1.03237300 -1.11863500 -1.17241000
 O 1.44661300 -0.39117000 -0.02784000
 C 0.36456500 -0.55007500 0.98722800
 C -0.88549700 0.13345500 0.39988100

C -0.39153000 0.90610100 -0.84379000
 O -0.96882800 1.85011900 -1.35074100
 H 0.32238600 0.05759600 -1.53015000
 H 0.27985200 -1.62195400 1.15613100
 H 0.76328100 -0.04659800 1.87233200
 H -1.65226800 -0.58801500 0.09619400
 H -1.33747700 0.83082500 1.11109800
 H 1.01019500 0.67833000 -0.50386400



H₂ + zwitterionic cyclic compound

FSM path	GSM path	SSM path	KinBot 2 only
			
Intended channel			
Energy barrier, kcal/mol			
			76.4
Reaction energy, kcal/mol			
			65.9
Number of gradient calls			

TS geometry (in Angs):

C -0.220521 0.230761 -0.327886
 H 0.253071 0.603615 0.577730
 O 0.852971 -0.293291 -1.211845

C -0.946768 -1.117847 -0.199849
 H -0.747393 1.011533 -0.879130
 H -1.989679 -1.177651 -0.523633
 C 0.011013 -1.787376 -1.167089
 H -0.840679 -1.542214 0.801336
 O -0.031863 -2.521312 -2.082070
 H 1.358181 -2.294421 -0.059066
 O 2.146165 -0.208489 -0.63881
 H 1.886068 -1.677571 -0.083985

• Channels with three products (except R22 listed in the previous category)

1,4-H ₂ O elimination and C-C break forming H ₂ O + ketene + CH ₂ O			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
[CH2]CC(=O)[O].O	[CH2]CC(=O)[O].O	[CH2]CC(=O)[O].O	
Energy barrier, kcal/mol			
50.5	50.5	50.5	50.5
Reaction energy, kcal/mol			
-24.9	-24.9	-24.9	-24.2
Number of gradient calls			
268	892	2111	

TS geometry (in Angs):

O -1.40066900 1.98840500 0.98081000
 O -2.95815700 2.05749400 -0.29775100

C -2.52082400 0.95425000 -0.94163600
 C -1.01554900 0.91200100 -1.33042800
 C -0.06592100 0.45909100 -0.15892800
 O 0.48116800 -0.57828000 -0.01981900
 H -1.51084700 2.34753800 1.88955100
 H -2.82188400 0.00018000 -0.47368600
 H -2.99923600 1.00407000 -1.94668200
 H -0.82596700 0.20812400 -2.15153300
 H -0.70797900 1.92103600 -1.62161700
 H -0.33987000 1.42816600 0.76896900

<p>R29</p>				
1,3-H ₂ elimination and C-C break				
FSM path	GSM path	SSM path	SC-AFIR	KinBot 1 and 2
Intended channel				
<chem>OO[C]CC=O.[H][H]</chem>		<chem>[CH]CC(=O)OO.[H][H]</chem>	<chem>OOC=C.O=[C].[H][H]</chem> (intended)	
Energy barrier, kcal/mol				
75.2		78.2 (different conformer. Mechanism is the same)	77.2	75.5
Reaction energy, kcal/mol				
38.4		38.4	38.4	38.7
Number of gradient calls				
363		3031		

TS geometry (in Angs):

O	1.25345600	-2.17361700	-0.39108900
O	0.20804700	-1.54974300	0.43618200
C	-0.35837400	-0.54487800	-0.31242400
C	-0.98732300	0.47337900	0.41773900
C	0.19918400	2.00657500	0.39948300
O	-0.10326300	3.08566200	0.69802500
H	1.18116600	-3.08859800	-0.06174800
H	0.97131100	0.27867700	-0.80305600
H	-0.65977000	-0.83053200	-1.31796000
H	-1.86494000	0.93089400	-0.03473100
H	-1.04626700	0.35589900	1.49850100
H	1.20673800	1.05623400	-0.52894900

1,3-H ₂ O ₂ elimination and C-C break			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
[CH2]O[C][CH2].OO	[CH2]O[C][CH2].OO	CCC(=O)OO	
Energy barrier, kcal/mol			
69.7	69.7	70.3	69.9
Reaction energy, kcal/mol			
25.6	25.6	25.6	25.1
Number of gradient calls			
291	690	1274	

TS geometry (in Angs):

O	0.12756700	-0.31275300	-2.12540000
---	------------	-------------	-------------

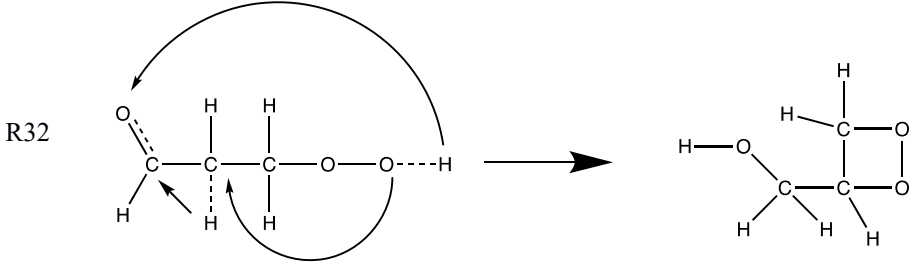
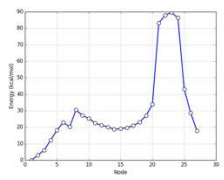
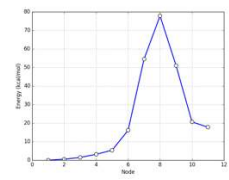
O	0.80393900	-0.92230800	-0.97979900
C	-0.73945900	-0.85394800	0.44702800
C	-0.64542200	0.35602800	1.14792000
C	0.90245400	1.20825700	0.37683900
O	1.19648900	2.25759100	0.79734600
H	0.67616400	-0.65872900	-2.85081600
H	-1.41928100	-0.97038800	-0.38928900
H	-0.33455700	-1.77174300	0.86058000
H	-1.37832800	1.13152400	0.93019500
H	-0.30754200	0.32501400	2.18314500
H	1.11796600	0.21144400	-0.39775800

R31			
1,1-H ₂ elimination and O-O break			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
	C=CC=O.OO	OO[C]CC=O.[H][H]	
Energy barrier, kcal/mol			
	88.4	78.2 (another conformer with CHO group closer to OOH)	79.9
Reaction energy, kcal/mol			
	42.7	42.7	43.3
Number of gradient calls			
	2143	3697	

TS geometry (in Angs):

O -1.14002500 1.52817600 0.68563200
 O -1.61694300 -0.17770500 0.64455800
 C -0.63088500 -0.88095000 0.39389300
 C 0.39610400 -0.69031800 -0.67567800
 C 1.09340300 0.68384800 -0.73561600
 O 1.41608800 1.17877400 -1.79058100
 H -1.99188800 1.93369400 0.43666900
 H 0.28815900 -0.90461400 1.66874500
 H -0.26065700 -1.57439600 1.38618700
 H 1.17736300 -1.45549400 -0.56879800
 H -0.09345900 -0.85424600 -1.64271400
 H 1.33509900 1.15777000 0.23280700

- Cycle formation channels (except two-center-charged R23, R24, R27 listed above)

			
2 H-shifts and 4-member cycle formation			
FSM path	GSM path	SSM path	KinBot
			
Intended channel			
OC[C@H]1COO1 (intended)	OC[C@H]1COO1 (intended)		
Energy barrier, kcal/mol			
77.8	77.8		
Reaction energy, kcal/mol			

17.7	17.7		
Number of gradient calls			
251	287		

TS geometry (in Angs):

O -0.13219700 -0.41684000 -1.73895500
O -1.34251600 -0.42284100 -0.91915400
C -0.94271000 -0.78706400 0.39824400
C 0.49864800 -0.26090300 0.61768600
C 0.81889900 1.09832500 0.40916700
O 0.45214100 1.73710700 -0.65331100
H 0.09306800 0.99708000 -1.34502600
H -1.68694700 -0.34159700 1.06796900
H -0.92583500 -1.87761300 0.53587200
H 1.03084300 -0.74243900 1.43529100
H 0.67296000 -0.61044200 -0.64113600
H 1.50085600 1.65298500 1.05689700

H-shift and 3-member cycle formation			
FSM path	GSM path	SSM path	KinBot
Intended channel			
OC/C=C\OO	O[C@H](C=O)CO	O[C@H]1CO1.C=O	
Energy barrier, kcal/mol			
85.3	87.5	86.6	
Reaction energy, kcal/mol			
25.1	21.8	21.8	

Number of gradient calls			
451	1186	2421	

TS geometry (in Angs):

O	2.26411400	-0.97612300	-1.09204100
O	1.33615300	-0.27588200	-0.21544500
C	0.05732400	-0.80670000	-0.51185000
C	-0.97634200	-0.12538300	0.28580100
C	-0.86652300	1.14398000	1.01038100
O	-1.91130100	1.93422000	0.85968200
H	2.75361600	-0.21658800	-1.45853500
H	0.03482000	-1.89167000	-0.31147500
H	-0.17390100	-0.68529400	-1.58828500
H	-1.99043500	-0.50905900	0.17240400
H	-0.63621200	0.77309600	2.07260600
H	0.10862600	1.63536400	0.77672000

H-shift and 3-member cycle formation (via roaming)			
FSM path	GSM path	SSM path	KinBot 2 only
Intended channel			
Energy barrier, kcal/mol			
			102.7
Reaction energy, kcal/mol			
			24.6

Number of gradient calls

TS geometry (in Angs):

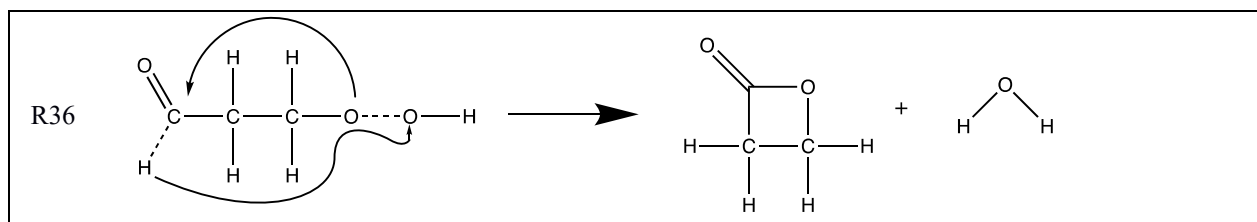
C	-0.129855	0.403451	-0.728224
H	-0.820703	0.817395	0.022929
O	1.163821	0.812525	-0.292492
C	-0.264591	-1.005578	-1.079376
H	-0.379535	0.903473	-1.695847
H	-1.328347	-1.203274	-1.300052
C	-0.167923	-2.863365	0.447033
H	-0.310982	-1.783473	0.750076
O	-1.111589	-3.611853	0.356224
H	0.883078	-3.176083	0.338420
O	2.112557	0.455853	-1.327163
H	2.084636	-0.523334	-1.285824

1,4-H ₂ O elimination on the OOH group and 4-member cycle formation (oxetane ketone)			
FSM path	GSM path	SSM path	KinBot
Intended channel			
OC1(O)CCO1			
Energy barrier, kcal/mol			
76.0			
Reaction energy, kcal/mol			

-56.5			
Number of gradient calls			
259			

TS geometry (in Angs):

O	1.17232400	0.43249800	-1.81606500
O	-0.61221700	0.24324300	-0.77878500
C	-0.48266000	-1.02571900	-0.06639500
C	-0.30190100	-0.34028900	1.26255700
C	-0.04292700	1.03241700	0.35083100
O	-0.45067100	2.16292700	0.64705100
H	1.37557700	0.49096900	-2.77478700
H	-1.41914800	-1.57834000	-0.18803600
H	0.36334700	-1.58614600	-0.46752000
H	-1.21909900	-0.18877300	1.83059400
H	0.54964000	-0.58021000	1.89715900
H	1.06772400	0.93741700	0.10338900



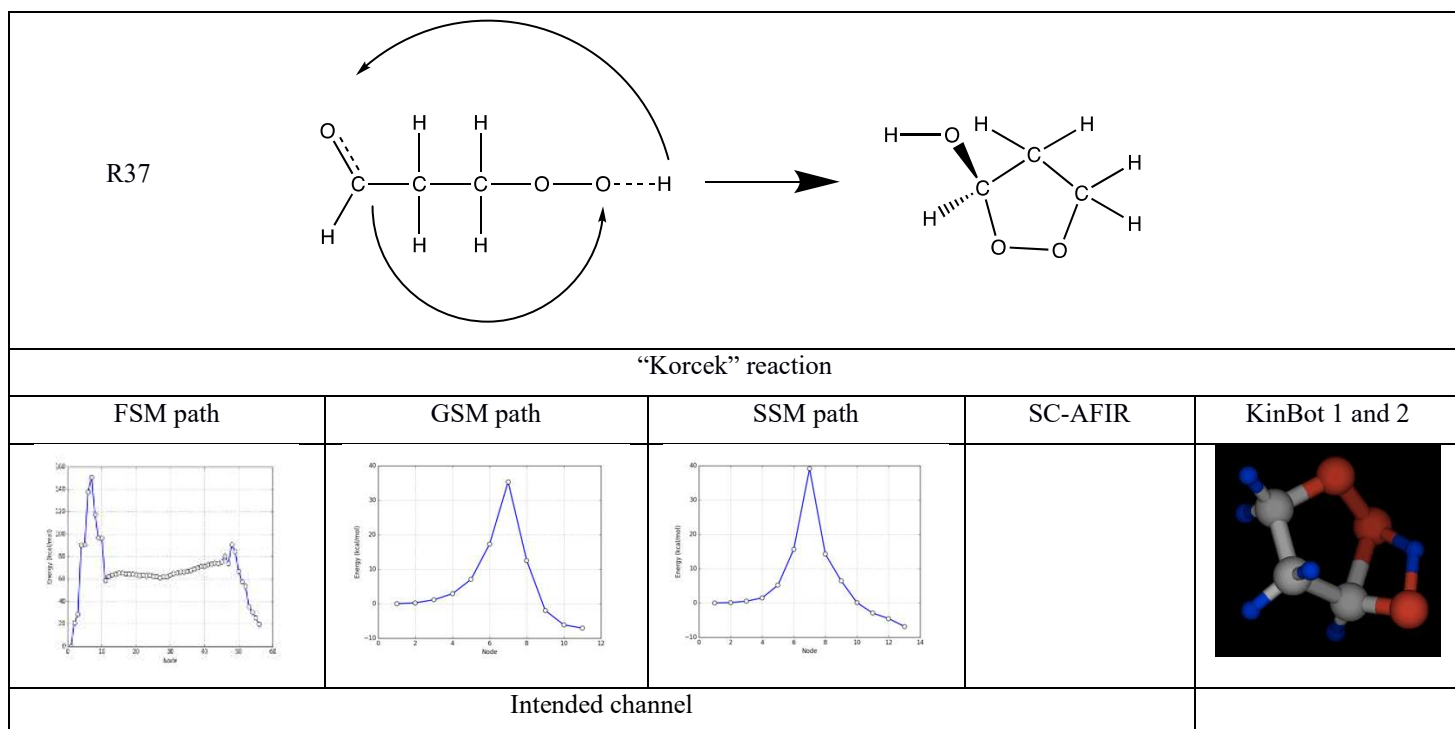
1,4-H₂O elimination on the OOH group and 4-member cycle formation (oxetane ketone) with H roaming around O in the OOH group

FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
	<chem>O[C@H]1C[C]1[O].O</chem>	<chem>O[C@H]1C[C]1[O].O</chem>	
Energy barrier, kcal/mol			
	58.1	58.1	58.3

Reaction energy, kcal/mol			
	-56.5	-56.5	-56.4
Number of gradient calls			
	676	843	

TS geometry (in Angs):

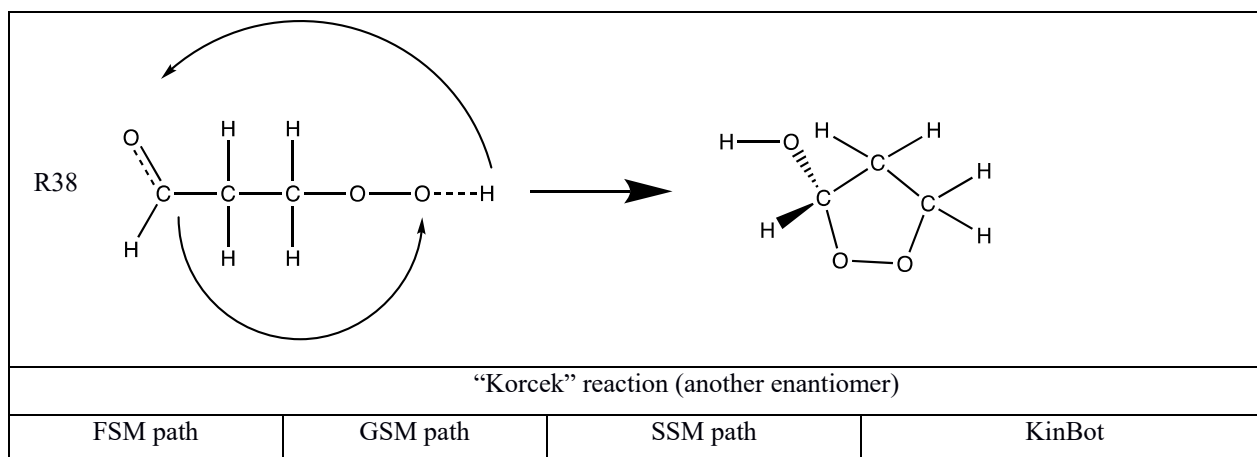
O -0.30975700 1.41109200 -1.72245700
O -1.10236100 0.23112200 -0.05724600
C -0.18321900 0.92005600 0.80804700
C -0.76860300 0.81849200 2.23272800
C -1.40538100 -0.52776900 2.27026600
O -1.81082300 -1.32797800 3.02172000
H -0.96120300 1.00872900 -2.32417100
H 0.81038100 0.46110100 0.75254000
H -0.12349100 1.94486600 0.45440200
H -0.05767600 0.94847000 3.06219200
H -1.59814000 1.53035900 2.35882400
H -1.30052500 -0.64754800 0.42439100

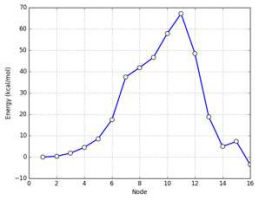


OO[C@H]1C[C@@H]1O	O[C@H]1CCOO1 (intended)	OCCC1OO1	O[C@H]1CCOO1 (intended)	
Energy barrier, kcal/mol				
39.1	35.3	39.1	34.4	34.9
Reaction energy, kcal/mol				
-7.2	-7.1	-7.2	-7.2	-7.7
Number of gradient calls				
352	193	405		

TS geometry (in Angs):

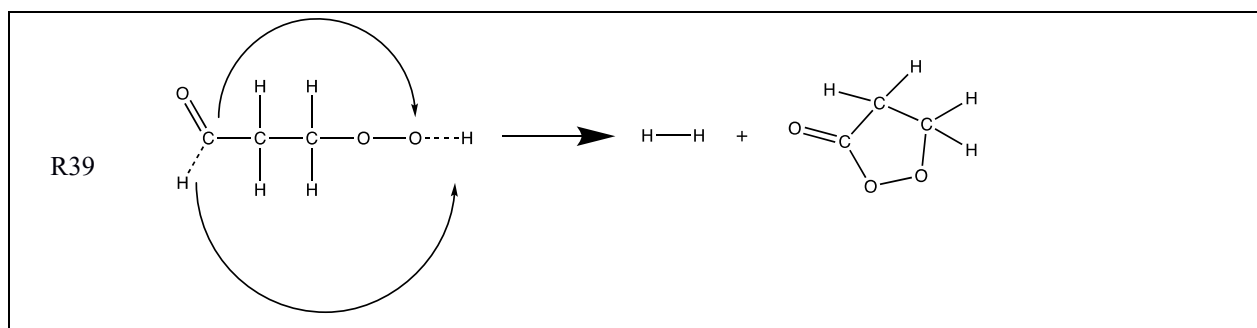
C	3.712092000	0.346330000	0.987643000
C	3.480240000	-0.896260000	0.149667000
C	2.272348000	-1.591798000	0.788212000
H	3.960070000	1.272747000	0.452118000
H	3.270062000	-0.627643000	-0.890913000
H	4.370779000	-1.532790000	0.188418000
H	2.541015000	-2.050364000	1.747904000
H	1.793492000	-2.331963000	0.141872000
O	1.281849000	-0.569796000	0.988649000
O	2.013975000	0.569284000	1.475565000
H	2.769190000	0.302458000	2.402226000
O	4.059015000	0.239550000	2.245395000

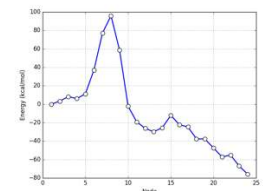
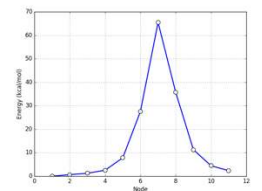
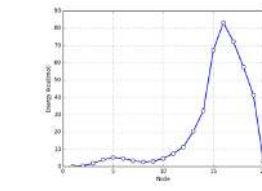
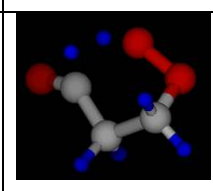


			
Intended channel			
		[CH2]COC(=O)O.[H]	
Energy barrier, kcal/mol			
		67.2	
Reaction energy, kcal/mol			
		-7.1	
Number of gradient calls			
		1361	

TS geometry (in Angs):

O -1.51775800 0.83989400 0.54714500
O -0.33578000 0.48981300 1.41214900
C -0.00899000 -0.84755700 1.03035300
C 0.26275300 -0.68597000 -0.47343100
C 0.47500000 0.80489700 -0.49167200
O 0.00410500 1.64757400 -1.35721600
H -0.92379600 1.38355000 -1.54862200
H 0.88864200 -1.12729100 1.59661600
H -0.83156600 -1.53328100 1.26099500
H 1.17679000 -1.18703400 -0.82833400
H -0.57765400 -0.97843500 -1.10360400
H 1.38330800 1.19223400 -0.03406500



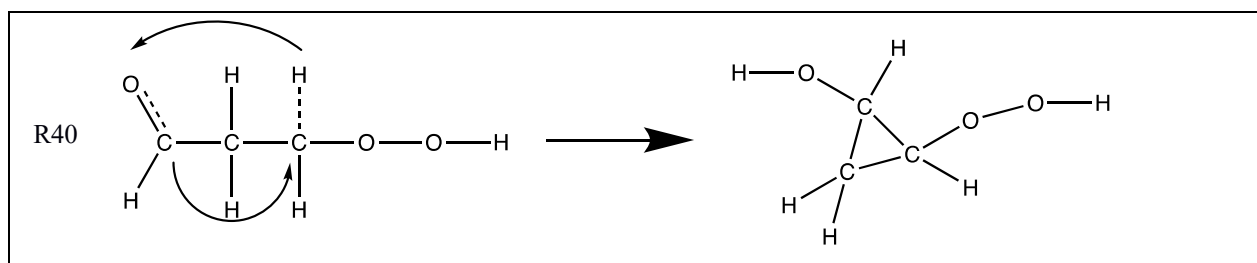
H ₂ elimination and 5-member cycle formation			
FSM path	GSM path	SSM path	KinBot 2 only
			
Intended channel			
OCCC(=O)O	O=C1CCOO1.[H][H] (intended)	O=C1CCOO1.[H][H] (intended)	
Energy barrier, kcal/mol			
66.8	65.4	66.4	65.9
Reaction energy, kcal/mol			
1.8	1.8	1.8	1.4
Number of gradient calls			
296	289	1757	

TS geometry (in Angs):

```

O  1.53722900  0.59753800 -0.76122300
O  1.40298200 -0.81191600 -0.47223000
C  0.00954400 -0.98938000 -0.25351500
C -0.39908900  0.06317900  0.78776700
C  0.20689300  1.36055800  0.30172000
O -0.23713100  2.37625500 -0.10762600
H  1.85800200  1.16243600  0.43642700
H -0.54738500 -0.85687800 -1.19000600
H -0.11197800 -2.01201300  0.11365500
H -1.48196600  0.21036200  0.84517700
H -0.00815600 -0.19743400  1.77332500
H  1.47167100  1.57089300  1.18210300

```



H-shift and 3-member ring formation

FSM path	GSM path	SSM path	KinBot 2 only
Intended channel			
<chem>OCC[C]OO</chem>		<chem>O[C]CCOO</chem>	
Energy barrier, kcal/mol			
93.8		96.3	96.1
Reaction energy, kcal/mol			
22.1		23.0	21.6
Number of gradient calls			
257		592	

TS geometry (in Angs):

O	1.46738200	-1.78245100	-0.49472400
O	0.42710500	-1.38654000	0.48351100
C	-0.34559500	-0.43195700	-0.12945400
C	-0.90282700	0.63540400	0.81102100
C	0.16176200	1.24948300	-0.02131200
O	-0.05823100	1.94438700	-1.15809200
H	1.48928900	-2.74230900	-0.32185700
H	0.23209500	0.59579600	-1.13768600
H	-1.01637200	-0.76597200	-0.92770900
H	-1.92458500	0.93751400	0.59098000
H	-0.68935600	0.47903500	1.86645300
H	1.15929700	1.26757500	0.43884600

R41			
1,4-H ₂ elimination and bridged cycle formation			
FSM path	GSM path	SSM path	KinBot
Intended channel			
	[CH]COOC=O.[H][H]	[CH]COOC=O.[H][H]	
Energy barrier, kcal/mol			
	106.8	106.8	
Reaction energy, kcal/mol			
	42.6	42.6	
Number of gradient calls			
	810	2173	

TS geometry (in Angs):

```

O  -0.45203200  -1.04012500  1.23484300
O  -0.08610500  -1.59299700  -0.08000800
C   0.44177600  -0.49794300  -0.86314800
C   0.30790000   0.70661800   0.01343000
C   0.55681800   0.25859700   1.41790800
O   1.81413300   0.08894700   1.43910700
H  -1.34659600   0.09261100   0.57149900
H   1.49953100  -0.68380000  -1.08358800
H  -0.12216900  -0.40122300  -1.79888800
H   0.55820100   1.70233400  -0.34949100
H  -1.24888200   0.79657800   0.07676300
H   0.00119200   0.71877900   2.25708100

```

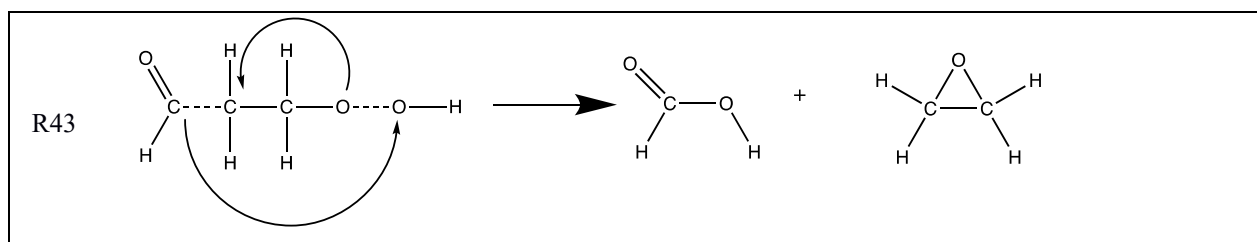
1,2-H ₂ O ₂ elimination and 4-member cycle formation			
FSM path	GSM path	SSM path	KinBot
Intended channel			
	<chem>OO[C@@H]1CCO1</chem>		
Energy barrier, kcal/mol			
	89.3		
Reaction energy, kcal/mol			
	47.8		
Number of gradient calls			
	805		

TS geometry (in Angs):

```

O  -0.81008000  -2.14330000  0.00422200
O  -0.66990700  -1.24924400  -1.12661400
C   0.95766400   0.09514800  0.10021300
C  -0.02075800   1.08203500  -0.52082900
C   0.40842900   2.01922900  0.51178000
O   1.30958400   1.33332700  1.12949300
H  -1.66816100  -2.56223000  -0.17694000
H   0.55351400  -0.67486300  0.74783800
H   1.87726800  -0.16601600  -0.41514900
H   0.16539400   1.37258300  -1.56096300
H  -1.03985000   0.62911100  -0.51427300
H   0.12702500   3.01322500  0.85884100

```



1,3 HCOOH elimination and 3-member cycle formation

FSM path	GSM path	SSM path	KinBot
Intended channel			
CCOC(=O)O			
Energy barrier, kcal/mol			
65.5			
Reaction energy, kcal/mol			
-35.8			
Number of gradient calls			
1961			

TS geometry (in Angs):

```

O   -1.69435200  -0.51505600  -0.44357900
O   -0.01157000  -0.05801100  -0.47967800
C    1.06750600  -1.10781900  -0.48080400
C    2.09227700  -0.07848000  -0.35546100
C    0.40348500   0.94199000   0.60683500
O    0.39560300   0.64503200   1.76869700
H   -1.89801300  -0.17482600  -1.33521800
H    0.91645100  -1.76411400   0.37564000
H    0.95078500  -1.63111900  -1.42957000
H    2.69051800  -0.00063800   0.54431100
H    2.44838100   0.41623800  -1.25326100
H    0.25404100   1.94163700   0.18342200

```

OH-shift and 4-member cycle formation			
FSM path	GSM path	SSM path	KinBot
Intended channel			
<chem>O[C@H]1COC1=O.[H][H]</chem>			
Energy barrier, kcal/mol			
			71.6
Reaction energy, kcal/mol			
			14.2
Number of gradient calls			
			2051

TS geometry (in Angs):

```

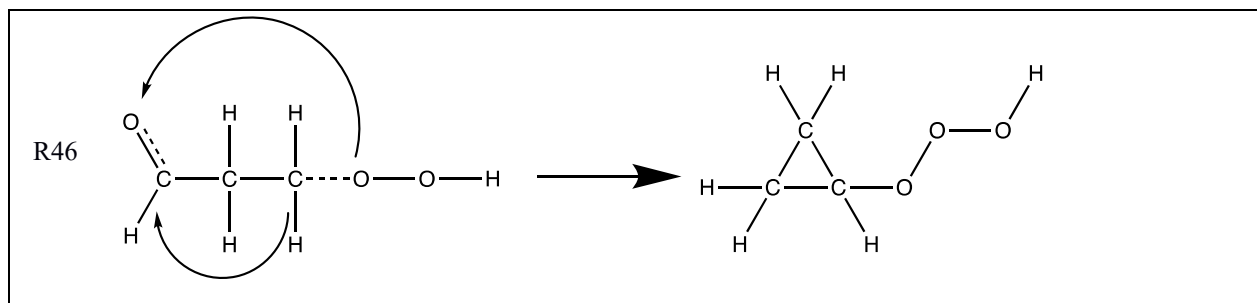
O  -1.80711300  0.59402700 -0.82124900
O   0.16610600 -0.17393000 -1.14736700
C   0.09055300 -1.22273100 -0.14053200
C   0.34760300 -0.23551900  1.00341300
C   0.63625500  0.80434200 -0.13488600
O   0.11233000  1.96374900 -0.23925800
H  -1.82836600  1.57193700 -0.81302500
H  -0.91075900 -1.66283300 -0.17185800
H   0.84771200 -1.98753700 -0.33955400
H  -0.54962600  0.04607100  1.55456300
H   1.17570800 -0.45608200  1.68108300
H   1.73825600  0.85000100 -0.34008900

```

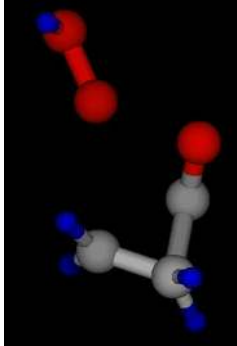
“inverse” Korcek with O-O-O in 6-member cycle			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
Energy barrier, kcal/mol			
			89.0
Reaction energy, kcal/mol			
			41.9
Number of gradient calls			

TS geometry (in Angs):

C	-0.141763	0.311301	-0.449185
H	-0.609904	1.103782	0.131929
O	1.274264	0.544007	-0.303378
C	-0.565745	-1.096897	-0.033814
H	-0.341789	0.450789	-1.513462
H	-1.618425	-1.204314	-0.309459
C	0.315799	-2.101183	-0.788655
H	-0.484690	-1.239447	1.046701
O	0.647364	-1.747316	-2.001619
H	0.009831	-3.163475	-0.715519
O	1.949690	-0.491538	-0.685206
H	1.295935	-2.139972	-0.172629



The OOH moves to the other O and the three carbons make a cycle.

FSM path	GSM path	SSM path	KinBot 2 only
			
Intended channel			
Energy barrier, kcal/mol			
			88.3
Reaction energy, kcal/mol			
			68.6
Number of gradient calls			

TS geometry (in Angs):

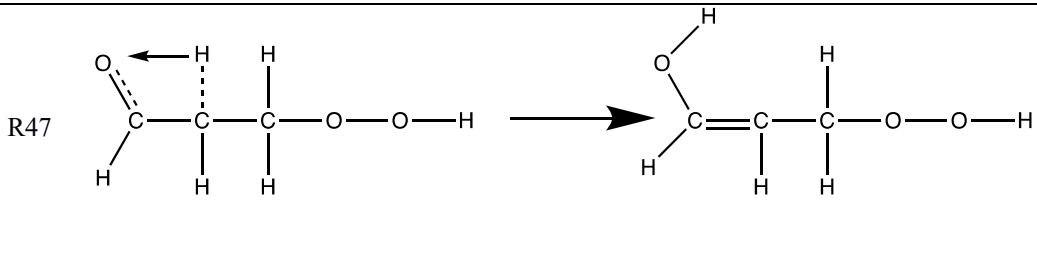
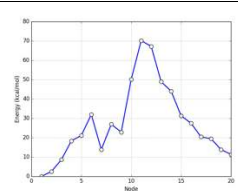
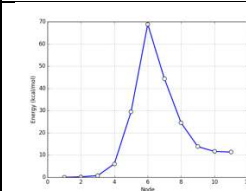
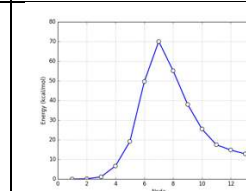
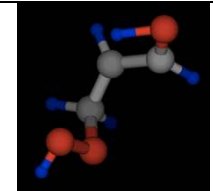
```

C  -0.798892  -0.138102  -0.416142
H  -0.007608   0.540503  -0.131251
O   1.693771  -0.664097  -1.290194
C  -1.234092  -1.191145   0.476254
H  -1.311552   0.053598  -1.351667
H  -2.233183  -1.600103   0.322791
C  -0.178582  -1.923102  -0.393565
H  -0.932483  -1.125898   1.519192

```

O 0.998249 -2.073242 0.004489
H -0.541570 -2.376278 -1.322364
O 2.895944 -0.426661 -1.917941
H 3.380563 0.150267 -1.293898

- **Unimolecular noncyclic channels with a stable product (non-radical or two-center-charged)**

			
enol-oxo tautomerization			
FSM path	GSM path	SSM path	KinBot 1 and 2
			
Intended channel			
OOC/C=C/O (intended)	OOC/C=C/O (intended)	OOC/C=C/O (intended)	
Energy barrier, kcal/mol			
68.8	68.8	70.0	70.6
Reaction energy, kcal/mol			
11.4	11.4	9.5	10.9
Number of gradient calls			
206	264	253	

TS geometry (in Angs):

O 1.96012300 0.00212900 -0.78602200
O 1.40329200 -1.22808900 -0.21817500
C -0.02910300 -1.08879600 -0.27922000
C -0.60042000 -0.01854500 0.60621800
C -0.68979000 1.35210600 0.23437300

O	-1.67929300	1.90820600	0.82990200
H	2.58310600	-0.38491400	-1.42643700
H	-0.32491500	-0.95568900	-1.32772100
H	-0.36375800	-2.07840700	0.05346900
H	-1.88006800	0.66375000	1.11651400
H	-0.29746800	-0.10878400	1.65694700
H	-0.08171700	1.93701900	-0.45986700

1,2-H exchange			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
OO[C@H]1CCO1	OOC/C=C\O	OOC/C=C/O	
Energy barrier, kcal/mol			
69.2	74.6	81.6	81.5
Reaction energy, kcal/mol			
-0.96	-0.94	0.15	-1.5
Number of gradient calls			
268	1319	735	

TS geometry (in Angs):

O	-2.06222900	-0.33882800	-0.25227500
O	-1.51112000	0.21938700	0.96834300
C	-0.29483900	-0.43558800	1.21158500
C	0.79112800	-0.14457100	0.24607400
C	0.71811700	0.52116700	-1.04364600
O	-0.32905200	0.99789900	-1.64503100

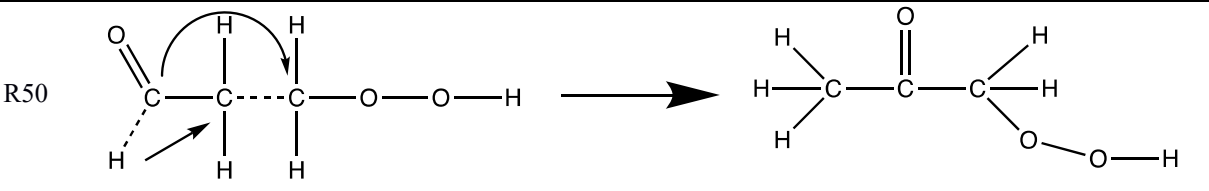
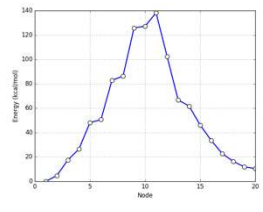
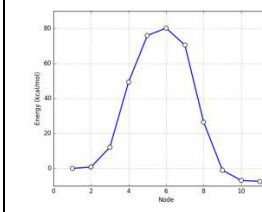
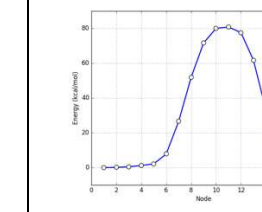
H	-1.60337700	0.26039000	-0.93184600
H	0.01955400	-0.11846800	2.22037500
H	-0.43382700	-1.53443200	1.26131300
H	1.79277400	-0.47100400	0.53311500
H	1.34013300	-0.22241800	-1.65323800
H	1.57271700	1.26644700	-0.91479400

1,2 (CH ₂) exchange			
FSM path	GSM path	SSM path	KinBot
Intended channel			
[CH]COO.C=O	OO[C@@H]1C[C@H]1O		
Energy barrier, kcal/mol			
90.2	89.5		
Reaction energy, kcal/mol			
1.06	1.06		
Number of gradient calls			
510	634		

TS geometry (in Angs):

O	0.72988400	-3.00928300	0.94330900
O	0.53110500	-1.66905400	0.58801900
C	0.20020600	0.13429800	-0.71636700
C	-0.41023300	0.37360400	0.53068700
C	-0.38866300	1.82082000	-0.53227400

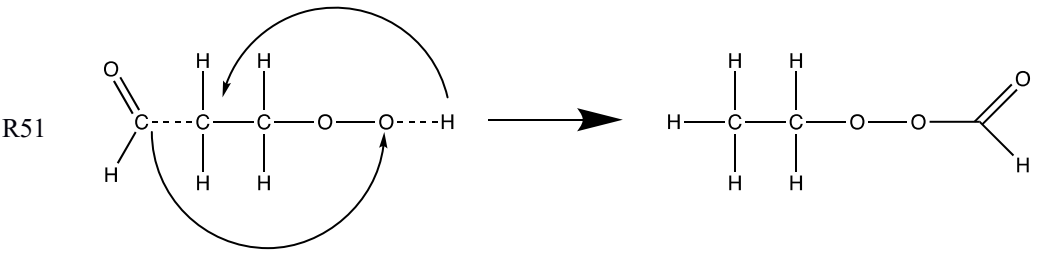
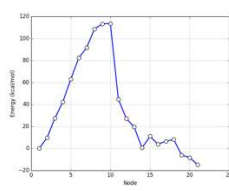
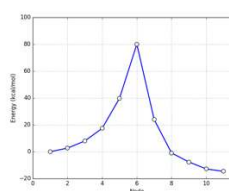
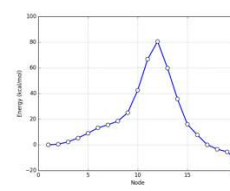
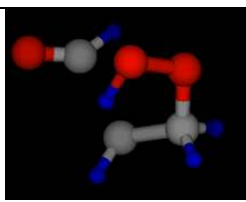
O -1.40313200 2.17227700 -1.09673900
H 1.60758100 -3.00371000 1.37104600
H 1.27596100 0.23068100 -0.82798700
H -0.36977200 -0.31052400 -1.52302400
H -1.45036400 0.11302300 0.68463700
H 0.19002700 0.65625900 1.39025500
H 0.43569600 2.49871600 -0.25872400

			
C-C break and H-shift forming 1-hydroperoxyacetone			
FSM path	GSM path	SSM path	KinBot
			
Intended channel			
OOCC(=C)O	CC(=O)COO (intended)	OOCC(=C)O	
Energy barrier, kcal/mol			
80.5	80.0	80.5	
Reaction energy, kcal/mol			
-6.3	-6.3	-6.3	
Number of gradient calls			
287	319	1826	

TS geometry (in Angs):

O 1.57047700 -0.40546000 -1.53401900
O 0.67350000 -1.27903900 -0.78152500
C -0.54743600 -0.64298300 -0.74586100
C -0.22267000 0.31344800 1.57242300
C -0.69357400 0.73424100 0.29697000
O -1.79454400 1.31590600 0.00203500

H 2.41770400 -0.85335700 -1.35192100
H -0.83171900 -0.19442500 -1.69735200
H -1.27074200 -1.37318900 -0.38714400
H -0.76487900 0.52951100 2.49134900
H 0.71202000 -0.23627400 1.63973900
H 0.25760600 1.15764100 -0.17183100

			
CHO-shift and H-shift forming ethoxy formate			
FSM path	GSM path	SSM path	KinBot 1 and 2
			
Intended channel			
CCOOC=O (intended)	CCOOC=O (intended)	CCOOC=O (intended)	
Energy barrier, kcal/mol			
80.0	80.0	80.0	78.9
Reaction energy, kcal/mol			
-15.3	-15.3	-15.3	-15.8
Number of gradient calls			
239	642	954	

TS geometry (in Angs):

O 0.33911700 0.16074400 -1.62967600
O 1.38000400 -0.30841400 -0.75520000
C 0.70763000 -1.08036600 0.30113700
C -0.56475400 -0.31984600 0.64552200
C -0.42532700 1.49245700 -0.38444100
O -1.39086700 2.10275600 -0.67825900

H	-0.52955800	0.04099900	-0.67131500
H	0.49806900	-2.08452000	-0.08505600
H	1.44336600	-1.13190100	1.10688200
H	-1.50582100	-0.87445400	0.59761300
H	-0.52060400	0.25940500	1.56782600
H	0.56872200	1.74311700	-0.01505000

O and CH ₂ group exchange forming hydroxymethyl acetate			
FSM path	GSM path	SSM path	KinBot
Intended channel			
CC(=O)OCO (intended)	[O]C[CH]C=O.O	CC(=O)OCO (intended)	
Energy barrier, kcal/mol			
66.8	67.2	66.8	
Reaction energy, kcal/mol			
-47.3	-47.6	-47.3	
Number of gradient calls			
331	792	807	

TS geometry (in Angs):

O	-1.04407300	-1.28529800	-2.14154600
O	0.31916800	-0.47428100	-1.03435000
C	-0.64163200	-0.78583800	-0.23961100
C	0.82657800	0.04785200	0.99212600
C	0.72817200	1.50091900	0.79408900
O	-0.20336200	2.16999100	1.21984400

H	-1.38979400	-0.47863300	-2.57230100
H	-1.43799800	-0.07242600	-0.00809700
H	-0.73060600	-1.80260000	0.15104400
H	0.25679200	-0.35437800	1.83234100
H	1.78149400	-0.43516600	0.80674500
H	1.53523400	1.96982600	0.19968600

2 H-shifts forming O ₂ CC=CO (internally catalyzed enol-oxo tautomerization)				
FSM path	GSM path	SSM path	SC-AFIR	KinBot
Intended channel				
	O=CCC=O.O	[CH]COO.C=O	O[CH][CH]COO (intended)	
Energy barrier, kcal/mol				
	45.8	45.8	44.1	
Reaction energy, kcal/mol				
	9.7	9.7	9.7	
Number of gradient calls				
	1406	361		

TS geometry (in Angs):

C	4.282000000	0.111585000	0.841924000
C	3.528901000	-0.886899000	0.183775000
C	2.562401000	-1.689724000	1.023165000
H	5.268164000	0.409608000	0.460942000
H	2.371780000	0.139712000	0.075888000
H	3.953648000	-1.358998000	-0.696461000

H	2.938228000	-1.882202000	2.034393000
H	2.271947000	-2.623951000	0.538192000
O	1.283401000	-0.946918000	1.205878000
O	1.594651000	0.439695000	0.891902000
H	2.434118000	0.690634000	1.593925000
O	3.808092000	0.780186000	1.821341000

R54			
H- and OH-shifts on the OOH group forming 3,3-dihydroxypropanal			
FSM path	GSM path	SSM path	KinBot
Intended channel			
	O=CCC(O)O (intended)	OO[C@H]1C[C@@H]1O	
Energy barrier, kcal/mol			
	63.6	63.6	
Reaction energy, kcal/mol			
	-62.0	-62.0	
Number of gradient calls			
	616	1282	

TS geometry (in Angs):

O	1.51296700	0.00098700	-0.71362900
O	0.46189500	-1.61739700	-0.50283400
C	-0.65126200	-1.02039500	-0.06828200
C	-0.25740500	-0.04281900	1.01553100
C	0.43777800	0.94620200	0.04227600
O	-0.37141700	1.31846600	-0.93846500

H 1.29620700 0.22046900 -1.64476000
H -1.05669900 -0.26418900 -0.87933700
H -1.48233800 -1.74898500 -0.05873000
H -1.09277000 0.47630300 1.49428100
H 0.41325200 -0.47755200 1.76169400
H 1.06047600 1.71857700 0.51041200

1,2-exchange of OOH and H groups			
FSM path	GSM path	SSM path	KinBot 2 only
Intended channel			
	<chem>C[C@H](C=O)OO</chem> (intended)	<chem>C[C@H](C=O)OO</chem> (intended)	
Energy barrier, kcal/mol			
	87.2	87.4	87.4
Reaction energy, kcal/mol			
	-1.6	-1.6	-1.9
Number of gradient calls			
	565	1800	

TS geometry (in Angs):

O 2.42025800 -0.68636400 -1.06301100
O 1.16001800 -0.56868100 -0.42623900
C -0.97917400 -0.79691900 -0.06290400
C -0.43812700 0.25699700 0.71378000

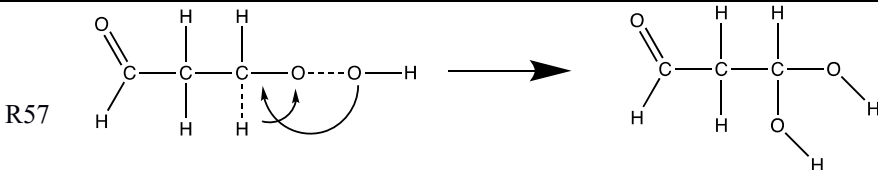
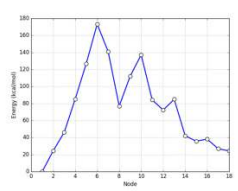
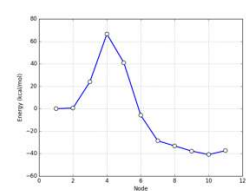
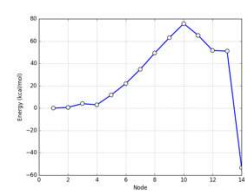
C -0.19230900 1.58477000 0.08042900
 O -0.82102300 1.98977900 -0.87941100
 H 2.91030400 -1.28279600 -0.46732500
 H -1.31562700 -0.58292600 -1.07020700
 H -1.09893200 -1.78824900 0.35777900
 H -1.73042600 0.02354000 0.73480000
 H -0.06125400 0.04518400 1.70955100
 H 0.60062400 2.18365100 0.56099300

C-C break forming C=COCOO			
FSM path	GSM path	SSM path	KinBot 2 only
Intended channel			
	[CH]COCOO	[CH]COCOO	
Energy barrier, kcal/mol			
	78.2	68.7 (another set of initial and final conformers but the same mechanism)	78.8
Reaction energy, kcal/mol			
	12.5	13.3	12.8
Number of gradient calls			
	388	570	

TS geometry (in Angs):

O -2.03772500 -0.43673200 -1.62707500

O -1.38794600 -1.36578100 -0.72206000
 C -1.11508800 -0.79558300 0.39411800
 C 1.33419900 0.44322200 1.01409100
 C 1.23519500 0.61962600 -0.34484100
 O 0.14778900 0.70117000 -1.03682600
 H -1.25049700 0.19212100 -1.70938800
 H -1.60420100 0.12852700 0.67243800
 H -0.55426400 -1.42163200 1.07869400
 H 0.50278600 0.68898600 1.67051800
 H 2.30643100 0.31606600 1.48095300
 H 2.16913100 0.71370700 -0.92288100

			
H-shift and O-O break			
FSM path	GSM path	SSM path	KinBot
			
Intended channel			
C=COOC=O.[H][H]	O[CH]C[C][O].O	[CH]CC=O.OO	
Energy barrier, kcal/mol			
65.8	65.8	65.9	
Reaction energy, kcal/mol			
-60.7	-60.7	-60.7	
Number of gradient calls			
290	440	765	

TS geometry (in Angs):

O 0.44952800 0.35825100 -1.85592800

O	-0.28086600	-1.29541900	-1.07837600
C	-0.73653000	-0.74323300	0.04869400
C	0.20198900	-0.21845300	1.05475200
C	0.55095500	1.28706700	0.70116400
O	0.16274200	2.19345800	1.39019300
H	0.17743500	0.48103500	-2.78381900
H	-1.76822700	-0.37614800	0.03800200
H	-0.78995600	-2.09418200	-0.37947800
H	-0.24531700	-0.20563800	2.05236900
H	1.13857100	-0.78542500	1.04161300
H	1.13965000	1.39867400	-0.22919800

- **H₂ elimination channels with a noncyclic stable product**

1,2-H ₂ elimination on the CH ₂ groups			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
OOC[C]C=O.[H][H]	OOC[C]C=O.[H][H]	OO[C]CC=O.[H][H]	
Energy barrier, kcal/mol			
99.7	99.7	107.6	103.5
Reaction energy, kcal/mol			
29.3	29.3	30.9	29.5
Number of gradient calls			
261	1504	1126	

TS geometry (in Angs):

O	-2.00304700	-0.97032400	1.51448600
O	-0.56542000	-0.63854000	1.37497900
C	-0.35302600	-0.22339300	0.09818800
C	1.03755400	0.08315500	-0.22374400
C	1.28558300	1.26671500	-0.98951800
O	0.41673200	1.99278900	-1.49017000
H	-2.16397600	-0.56673400	2.38824800
H	-0.85093600	-1.36160600	-0.60373700
H	-1.06824100	0.47272300	-0.34459800
H	1.82943000	-0.51237200	0.21535900
H	0.07758700	-1.05446300	-0.80809200
H	2.35774000	1.51203400	-1.13142000

R59			
1,2-H ₂ elimination on CHO and CH ₂ groups			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
	OOCC(=O)[CH].[H][H]	OOCC(=O)[CH].[H][H]	
Energy barrier, kcal/mol			
	84.7	84.7	86.1
Reaction energy, kcal/mol			
	34.1	34.1	34.7
Number of gradient calls			
	813	599	

TS geometry (in Angs):

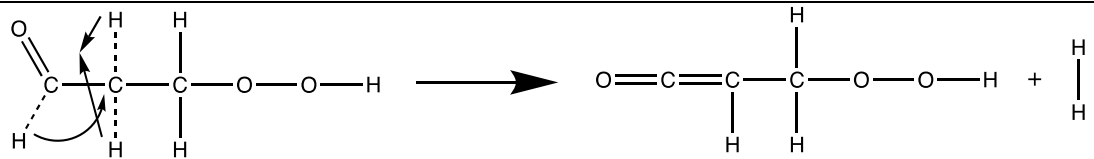
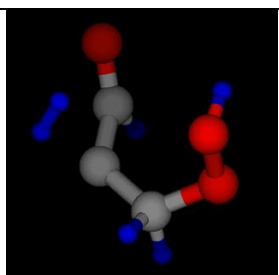
O	-2.19962100	-0.13799600	1.99089800
---	-------------	-------------	------------

O -1.11907900 -0.74774100 1.23198100
 C -0.61139100 0.27004300 0.36702700
 C 0.35206100 -0.39431200 -0.55296600
 C 1.36472300 0.31276300 -1.18843700
 O 1.77911600 1.48090300 -1.23712400
 H -1.84341100 -0.21160300 2.89541600
 H -1.44541000 0.71874700 -0.19625000
 H -0.12199900 1.07135000 0.93845500
 H 2.19292800 -0.61912400 -1.29466300
 H 0.24353700 -1.45512300 -0.75917200
 H 1.61487700 -0.41484000 -2.16906900

1,5-H ₂ elimination and H-shift			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
		<chem>O=C1COOC1.[H][H]</chem>	
Energy barrier, kcal/mol			
		82.2	81.8
Reaction energy, kcal/mol			
		33.6	33.1
Number of gradient calls			
		3506	

TS geometry (in Angs):

O 1.41407100 -0.83403400 1.10144900
 O 1.18755100 -1.01925100 -0.29051300
 C -0.16852900 -1.15446500 -0.50259100
 C -0.88097700 0.11837300 0.14971000
 C -0.49887900 1.30092800 -0.55315600
 O -0.45440800 2.04574600 -1.43223800
 H 1.47587700 0.96612700 0.97794500
 H -0.33910700 -1.23749100 -1.58024600
 H -0.60768600 -2.00436700 0.03693100
 H -1.97606500 0.07487800 0.21612300
 H -0.38833800 0.08973100 1.14144700
 H 1.23833600 1.67325000 0.73170000

<p>R61 </p>			
Same product as R58 and R59, but different mechanism.			
FSM path	GSM path	SSM path	KinBot 2 only
			
Intended channel			
Energy barrier, kcal/mol			
			98.4
Reaction energy, kcal/mol			
			32.6
Number of gradient calls			

TS geometry (in Angs):

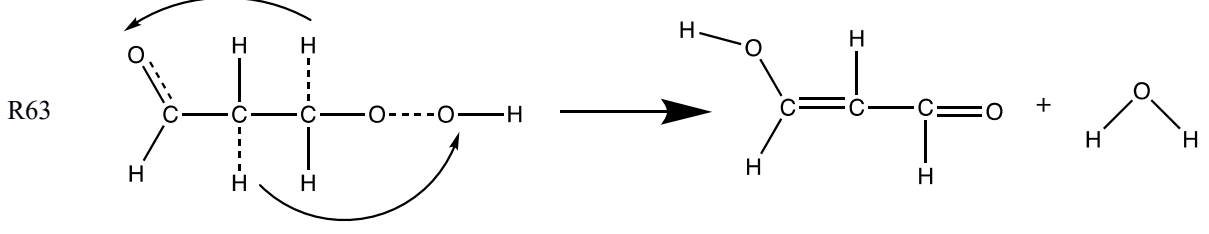
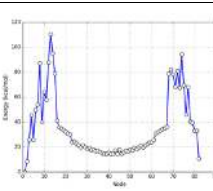
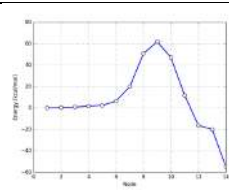
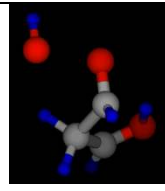
C 0.225538 0.128500 0.201557
 H 0.492485 0.405901 1.232979
 O 1.463657 0.262001 -0.511683
 C -0.446019 -1.174893 0.228555
 H -0.488635 0.888459 -0.143676
 H -1.431461 -1.313504 -1.730768
 C 0.127233 -2.414739 -0.030975
 H -1.761252 -1.186497 -1.057941
 O 0.065055 -3.330763 -0.835287
 H 0.724681 -2.515839 0.936372
 O 1.143999 0.149789 -1.926205
 H 1.615285 -0.672679 -2.157224

$\text{H}_2 + \text{HO}(\text{O}=\text{C})\text{-CH}_2\text{-CHO}$ (pyruvic acid), as the roaming OH will eventually bind with the carbon			
FSM path	GSM path	SSM path	KinBot 2 only
Intended channel			
Energy barrier, kcal/mol			79.9
Reaction energy, kcal/mol			-56.0
Number of gradient calls			

TS geometry (in Angs):

C 0.276855 -0.206259 0.073647
 H -0.295650 0.762861 0.643822
 O 1.414195 0.055813 -0.332266
 C -0.632526 -1.292717 -0.424934
 H -0.649579 1.043788 -0.151287
 H -0.637640 -1.273403 -1.525781
 C -0.112215 -2.666055 -0.004365
 H -1.654331 -1.167212 -0.057392
 O -0.846150 -3.574385 0.313014
 H 0.987268 -2.798289 -0.033067
 O 1.464848 0.069369 -2.112686
 H 2.415492 0.272228 -2.183003

• **Another H₂O-elimination channel (two stable noncyclic products)**

			
1,3-H ₂ O elimination and H-shift			
FSM path	GSM path	SSM path	KinBot 1 and 2
			
Intended channel			
OOC/C=C\O		OC/C=C/OO	
Energy barrier, kcal/mol			
61.6		61.6	60.3
Reaction energy, kcal/mol			
-54.6		-54.6	-55.1
Number of gradient calls			

442		1098	
-----	--	------	--

TS geometry (in Angs):

O	-0.16639500	-2.46265600	0.16503100
O	-1.18999500	-0.92579300	-0.34275300
C	-0.35066300	-0.17920100	-0.94812900
C	0.79209900	0.30517700	-0.05260300
C	0.31685800	1.47220500	0.70585100
O	-0.86733700	1.86323500	0.51805500
H	-0.83604700	-2.93951000	0.68645100
H	-0.12715900	-0.33082100	-2.01261600
H	-1.19652200	1.17855900	-0.27365400
H	0.95824400	-0.54851800	0.64801000
H	1.76393700	0.52225600	-0.52009800
H	0.90287900	2.04496800	1.42638200

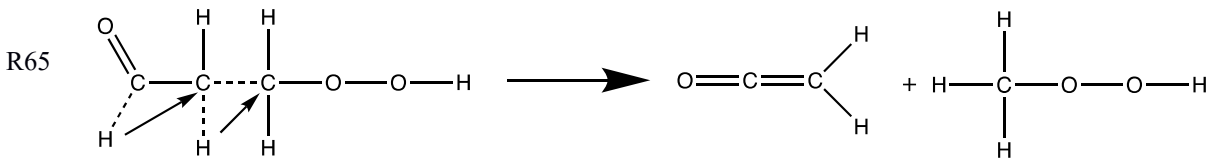
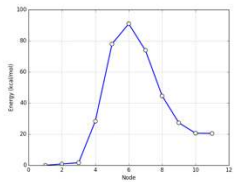
- **CH₂-CH₂ bond breaking and forming two stable non-cyclic products**

C-C and O-O break			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
OCC1OCO1	OCC=O.C=O (intended)	CO[C@@H](C=O)O	
Energy barrier, kcal/mol			
76.4	74.7	77.7	76.8

Reaction energy, kcal/mol			
-38.6	-38.6	-38.6	-36.0
Number of gradient calls			
279	661	620	

TS geometry (in Angs):

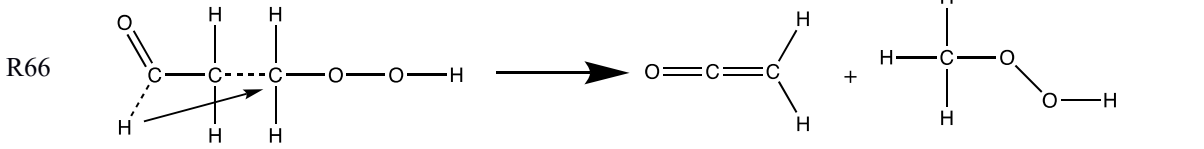
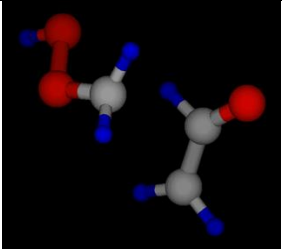
O	-0.18565800	-1.86864200	0.77649800
O	-0.41992100	-1.71959800	-1.12913200
C	-0.67520600	-0.49947400	-1.13610100
C	0.69702500	0.45606400	0.74081300
C	0.29440900	1.69931700	0.10708200
O	-0.66012100	2.38345100	0.46546900
H	-0.25338400	-2.74133600	1.22838800
H	-0.06541900	0.14670300	-1.78860000
H	-1.58965100	-0.09263900	-0.69097800
H	0.30625700	0.23727300	1.72698400
H	1.61818400	-0.03030600	0.43937400
H	0.93347700	2.02918200	-0.73980300

			
2 H-shifts and C-C break forming ethenone and hydroperoxymethane			
FSM path	GSM path	SSM path	KinBot
			
Intended channel			
COO.[CH2][C][O]			
Energy barrier, kcal/mol			
	89.6		

Reaction energy, kcal/mol			
	19.6		
Number of gradient calls			
	690		

TS geometry (in Angs):

O 1.16536500 -1.06137900 0.52143400
O -0.11848600 -1.39338500 -0.09796500
C -0.45494000 -0.30103600 -0.88899600
C -1.56913900 1.48496900 0.68226700
C -2.80140900 1.58015000 0.08033600
O -3.43578000 1.41759100 -0.94588100
H 1.52588000 -1.96066500 0.62940300
H 0.33452600 -0.02857000 -1.59889900
H -1.39421200 -0.55581000 -1.38270100
H -1.46406000 0.89915300 1.59839600
H -0.60737800 0.64827100 -0.25226600
H -3.20350300 2.22613500 0.96173800

			
1 H-shift and C-C break forming ethenone and hydroperoxymethane			
FSM path	GSM path	SSM path	KinBot 1 and 2
			
Intended channel			
Energy barrier, kcal/mol			
			80.4

Reaction energy, kcal/mol			19.1
Number of gradient calls			

TS geometry (in Angs):

C	0.256665	0.028804	0.105551
H	0.865139	-0.206511	0.975890
O	0.811702	0.722046	-0.885354
C	-1.226760	-2.038091	-1.039055
H	-0.772816	0.334522	0.238843
H	-0.983550	-1.475488	-1.934330
C	-0.256596	-2.190561	-0.053139
H	-2.131175	-2.642495	-1.037449
O	-0.209805	-2.750436	1.048887
H	0.716199	-1.685680	-0.430601
O	2.219147	0.357073	-0.964091
H	2.442418	0.772555	-1.819449

C-C break and OH shift (+rotation) forming formaldehyde + hydroperoxyethene			
FSM path	GSM path	SSM path	KinBot
Intended channel			
	<chem>OO[C@H]1CCO1</chem>	<chem>OO[C@H]1C[C@@H]1O</chem>	
Energy barrier, kcal/mol			
	80.3	80.3	

Reaction energy, kcal/mol			
	28.4	28.4	
Number of gradient calls			
	454	1106	

TS geometry (in Angs):

O	2.00971100	0.46370000	-1.50056300
O	0.56630700	-1.34534400	-0.53168800
C	-0.64615000	-1.07204000	-0.12913800
C	-0.79764200	0.35991600	0.83394900
C	0.37168800	1.14151400	0.41502900
O	0.43270400	1.65511500	-0.69807600
H	1.72044700	-0.47442100	-1.56527500
H	-1.39850600	-0.94142800	-0.92318900
H	-0.99363900	-1.76022400	0.67133700
H	-1.75601500	0.75666600	0.49160600
H	-0.80028200	0.12309600	1.89908100
H	1.26900800	1.10731300	1.05169400

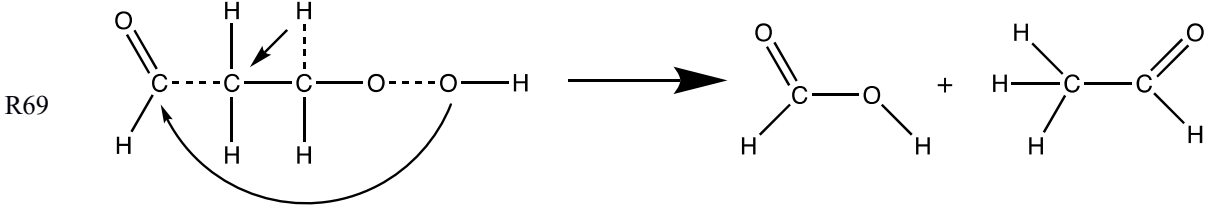
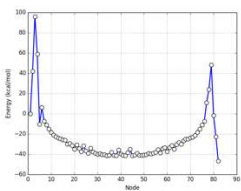
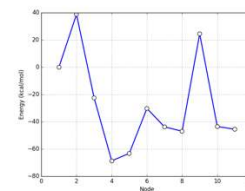
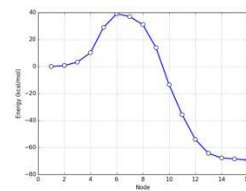
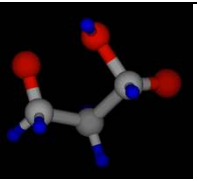
C-C break and OH shift forming formaldehyde + hydroperoxyethene			
FSM path	GSM path	SSM path	KinBot
Intended channel			
		OC/C=C\OO	
Energy barrier, kcal/mol			
		72.4	

Reaction energy, kcal/mol			
		27.5	
Number of gradient calls			
		1546	

TS geometry (in Angs):

O	-1.30614300	0.74821000	-0.52994400
O	-1.69158100	-0.90331400	0.69838000
C	-0.51331300	-0.79022300	1.17959900
C	0.81871300	-0.49610700	-0.28326900
C	1.00457000	0.84234800	-0.72514000
O	0.10661200	1.69436700	-1.01228800
H	-2.03486100	1.37354300	-0.74970400
H	-0.26366300	0.08085200	1.80143900
H	-0.02524500	-1.72802000	1.49169600
H	1.75335200	-1.01884100	-0.08099300
H	0.13635900	-1.04869300	-0.92744300
H	2.01519900	1.24587800	-0.86233500

- CHO-CH₂ breaking channels**

			
H and OH shifts and C-C break forming acetaldehyde + formic acid			
FSM path	GSM path	SSM path	KinBot 1 and 2
			
Intended channel			
OC=O.OC=C	OC=O.OC=C	OC=O.OC=C	

Energy barrier, kcal/mol			
38.8	38.8	38.8	48.8 (another conformer)
Reaction energy, kcal/mol			
-68.6	-68.6	-68.6	-69.1
Number of gradient calls			
517	1040	555	

TS geometry (in Angs):

O	0.97606200	0.45252600	-1.15818800
O	-0.15657400	-1.10243000	-1.31783200
C	-0.74908000	-1.05509700	-0.09481300
C	0.12073200	-0.23101800	0.89981100
C	0.35991800	1.09495600	0.07722200
O	-0.58523900	1.90211100	-0.10660300
H	0.54713900	0.88383400	-1.92824100
H	-1.76035000	-0.62407300	-0.12172200
H	-0.73725600	-2.08026700	0.32647600
H	-0.40749500	-0.01411400	1.83178700
H	1.06992400	-0.73614300	1.09983700
H	1.32219400	1.50969900	0.49225000

H-shift and C-C break forming CO and hydroperoxyethane			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
<chem>C[C@@H](C=O)OO</chem>		<chem>CC(=O)COO</chem>	
Energy barrier, kcal/mol			

86.6		86.3	83.8
Reaction energy, kcal/mol			
0.7		1.0	0.3
Number of gradient calls			
464		1940	

TS geometry (in Angs):

O 1.25327200 0.19911500 -1.22872300
O 1.37703600 -0.76061900 -0.16491100
C 0.04345900 -1.29560800 0.01302300
C -0.88056800 -0.07103000 0.04125700
C 0.30430100 1.70653200 0.56708400
O -0.20855700 2.57777700 1.12813300
H -0.01732100 0.44926800 -0.80039000
H -0.19158700 -1.96836500 -0.82057200
H 0.08329700 -1.85621200 0.95058300
H -1.75379000 -0.08771800 -0.62412000
H -1.24056000 0.13415200 1.05143800
H 1.16928400 1.38088800 -0.13264800

O-O and C-C break forming C ₂ H ₄ + peroxy formic acid			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
OOC=O.C=C (intended)	C=CC=O.OO	C=COC=O.O	

Energy barrier, kcal/mol			
56.6	58.5	56.6	56.6
Reaction energy, kcal/mol			
10.0	7.1	10.0	9.5
Number of gradient calls			
379	658	1539	

TS geometry (in Angs):

O	0.91831900	1.43817600	1.56923100
O	-0.04723500	0.66117400	0.81690500
C	-0.26766900	-1.24775200	0.46630300
C	-0.45913700	-0.92354700	-0.89412900
C	0.22566300	0.77745600	-0.78524000
O	-0.35926200	1.65968500	-1.38029200
H	0.34823200	2.14623900	1.92338200
H	0.68715300	-1.62214800	0.82474700
H	-1.09535600	-1.33595800	1.16140800
H	0.21426400	-1.36700600	-1.62569500
H	-1.48118100	-0.80221300	-1.24468400
H	1.31619500	0.61588800	-0.83194300

C-C and C-O break forming ethane and peroxyformic acid			
FSM path	GSM path	SSM path	KinBot
Intended channel			

[CH]OOCO			
Energy barrier, kcal/mol			
85.5			
Reaction energy, kcal/mol			
10.0			
Number of gradient calls			
300			

TS geometry (in Angs):

O	-0.09821800	1.35007000	1.45685000
O	-0.84091600	1.44532600	0.24586700
C	-0.72863100	-1.31658800	-0.31242300
C	0.51158000	-0.70681700	-0.08368800
C	0.71522200	0.64413900	-0.94116900
O	0.48137400	0.65254400	-2.12995300
H	-0.79345500	1.21126600	2.12792700
H	-0.88585800	-2.38508000	-0.17357800
H	-1.57365900	-0.71033600	-0.62417900
H	0.45905600	-0.16883700	0.94279200
H	1.37432600	-1.38473900	-0.06163500
H	1.37904700	1.36895700	-0.44689600

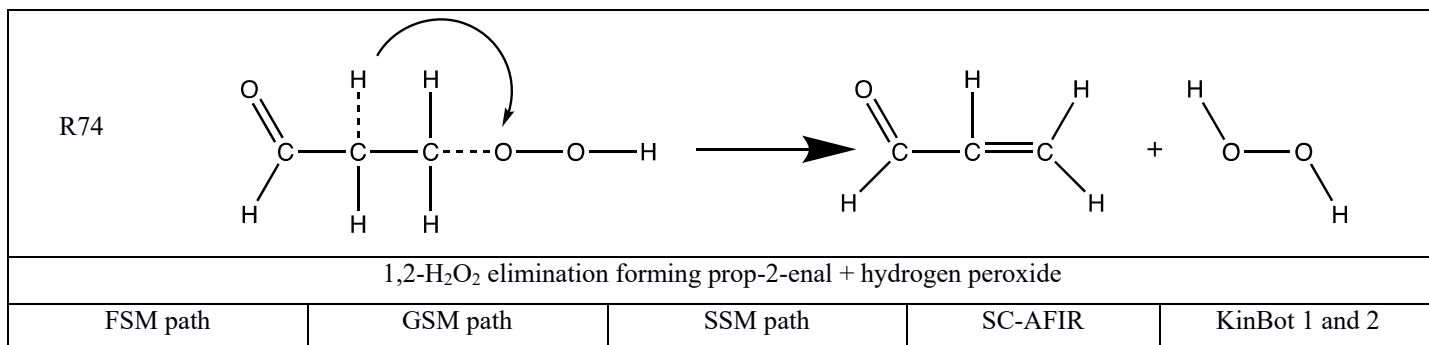
<p>R73</p>			
C-C break and H-shift forming hydroperoxyethene and formaldehyde			
FSM path	GSM path	SSM path	KinBot 1 and 2
Intended channel			
OCCC1OO1	OO[C@H]1CCO1	O[C@@H]1CO1.C=O	

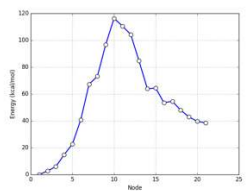
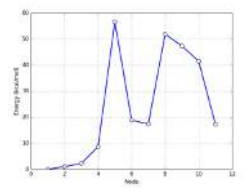
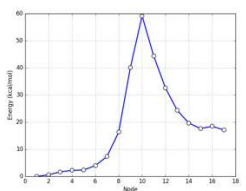
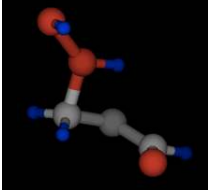
Energy barrier, kcal/mol			
69.9	72.8	66.4	63.4
Reaction energy, kcal/mol			
28.4	27.5	28.4	27.9
Number of gradient calls			
249	674	1655	

TS geometry (in Angs):

O 0.54413500 -2.41457000 -0.62505500
O -0.41150000 -1.50167200 0.00104100
C 0.24527600 -0.51863200 0.60408800
C -0.53635800 0.65822200 0.92215500
C -0.08042900 1.21551500 -0.55973600
O -0.93142200 1.16765300 -1.49148600
H 0.26755700 -2.31040700 -1.55773300
H 1.24530000 -0.73748700 0.96798600
H 0.86364500 0.43019100 -0.71877100
H -1.61721700 0.53067300 0.92913800
H -0.15942700 1.22498300 1.77393000
H 0.54138400 2.10722500 -0.34926200

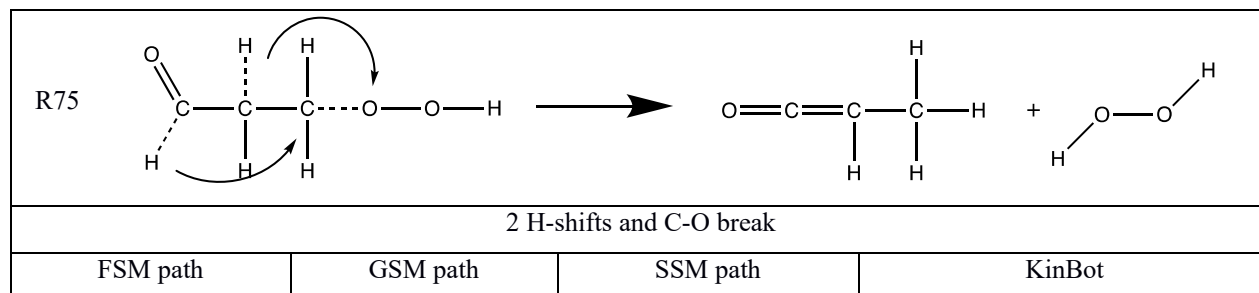
- **HOOH- elimination channels**

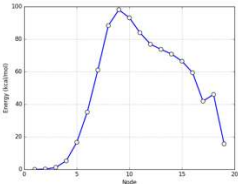


				
Intended channel				
OCO.[CH]C=O	C=CC=O.OO (intended)	C[C]C=O.OO	C=CC=O.OO (intended)	
Energy barrier, kcal/mol				
55.9	56.5	56.8	55.0	55.4
Reaction energy, kcal/mol				
17.2	17.2	17.2	17.2	16.7
Number of gradient calls				
257	846	1045		

TS geometry (in Angs):

C	3.992260000	0.209527000	-0.037963000
C	3.387928000	-0.715563000	0.916441000
C	2.775571000	-0.199067000	2.132546000
H	4.530916000	-0.269683000	-0.881338000
H	1.795729000	-0.629843000	0.514207000
H	3.644614000	-1.770183000	0.854156000
H	2.936527000	0.858594000	2.337080000
H	2.771100000	-0.828452000	3.019276000
O	1.345823000	-0.330946000	1.502049000
O	-0.027293000	0.213435000	2.077301000
H	0.024986000	1.125312000	1.729819000
O	3.922134000	1.433323000	0.028634000



			
Intended channel			
		CC(=O)COO	
Energy barrier, kcal/mol			
		97.6	
Reaction energy, kcal/mol			
		17.4	
Number of gradient calls			
		690	


TS geometry (in Angs):

```

O  1.59953200 -2.26092600  0.28383100
O  0.35591900 -1.67893400  0.65455100
C  0.37649100  0.19386500 -0.38385600
C -0.71338200  0.58580700  0.60972300
C -1.05410600  1.78562100 -0.20588900
O -1.91524000  2.39312900 -0.71494500
H  1.33271600 -2.93685500 -0.36534900
H  1.40883400  0.16257200 -0.05673700
H  0.09754400 -0.41414000 -1.23658000
H -1.54477000 -0.10933300  0.76065500
H -0.28101700  0.86160100  1.57593800
H  0.33554500  1.40789900 -0.91645700

```

Like G+ Log In



Documentation
Learn more about the RMG software

Resources
RMG related publications and presentations

Database
Browse the RMG database of chemical and kinetic parameters

Create Input File
Online form for making an RMG input file

Molecule Search
Draw a molecule from its adjlist and search or estimate its properties

Kinetics Search
Search or estimate kinetics of a chemical reaction

Solvation Search
Search or estimate solvation properties of solvent and a solute

Pressure Dependent Networks
CanTherm pdep kinetic calculations

Other RMG Tools
Additional tools to supplement RMG

RMG » Database » Kinetics » Search » Results

Kinetics Search Results

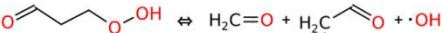




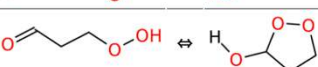
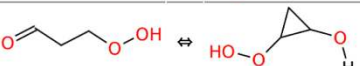

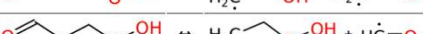
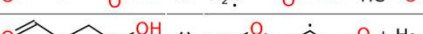







1.		3 results
2.		2 results
3.		2 results
4.		2 results
5.		2 results
6.		4 results
7.		2 results
8.		2 results
9.		2 results
10.		2 results
11.		2 results
12.		2 results
13.		2 results
14.		2 results
15.		2 results
16.		2 results
17.		2 results

Figure S2. RMG kinetic database for unimolecular decomposition of γ -keto hydroperoxide (screenshot from <http://rmg.mit.edu>).

RMG reaction 6 correlates with discovered R37; RMG reaction 7 \leftrightarrow R40; RMG reaction 16 \leftrightarrow R56; RMG reaction 17 \leftrightarrow R47.

S5. ELECTRONIC STRUCTURE CALCULATION RESULTS FOR CRIEGEE PLUS VINYL ALCOHOL REACTION

Geometries (Angstrom), frequencies (cm^{-1}), and zero-point energies (Hartree) are reported at the UM06-2X/aug-cc-pVTZ level of theory. Electronic energies at 0 K (Hartree) are reported at the RCCSD(T)-F12a/cc-pVTZ-F12 level of theory.

Criegee

O	1.106079000	0.016474000	0.150046000
O	2.406709000	0.003340000	-0.185680000
C	2.959890000	1.077042000	-0.461771000
H	2.375050000	1.989041000	-0.422889000
H	4.006582000	1.013554000	-0.725207000

E0: -189.404732250898

ZPE: 0.031764

Frequencies:	539.2210	696.2095	928.9247
	1021.3231	1261.9628	1434.2591
	1630.0476	3140.0660	3290.9550

Vinyl alcohol

O	0.989740000	-0.002242000	-0.159566000
C	0.418330000	0.208376000	1.052212000
C	1.050336000	0.407662000	2.198631000
H	1.947794000	0.012695000	-0.071350000
H	-0.660549000	0.195109000	0.976302000
H	0.484676000	0.564688000	3.102743000
H	2.131083000	0.418732000	2.261449000

E0: -153.643661125960

ZPE: 0.057085

Frequencies:	451.3134	496.7543	723.2478
	863.6811	967.9395	1018.2999
	1137.4540	1329.5169	1354.6649
	1448.8976	1731.1855	3177.3367

3219.0470 3280.9532 3857.2808

Van-der-Waals complex

C	-0.302052000	1.389830000	0.665536000
H	-0.128629000	1.552928000	1.722133000
O	0.649318000	0.927785000	0.027641000
C	-1.924784000	-0.961778000	0.148315000
H	-1.217095000	1.644016000	0.146638000
H	-1.915227000	-0.367696000	-0.756124000
C	-0.879972000	-1.728790000	0.473539000
H	-2.825827000	-1.013263000	0.739219000
O	0.295863000	-1.787192000	-0.154610000
H	-0.914666000	-2.398258000	1.325661000
O	0.416363000	0.663525000	-1.308087000
H	0.357661000	-1.079820000	-0.834913000

E0: -343.063865324791

ZPE: 0.091735

Frequencies:	103.5134	109.0123	151.1174
	196.8300	236.2835	270.7315
	517.5051	529.9932	691.3430
	745.4340	854.9784	882.8587
	893.8989	990.2308	1049.4531
	1066.6057	1211.7480	1250.0190
	1345.6598	1414.1271	1442.6099
	1462.6470	1659.3759	1695.9907
	3147.9528	3176.5866	3188.8587
	3279.4444	3292.2666	3410.0170

Transition state

C	0.386795000	0.539079000	0.011828000
H	0.400121000	0.622085000	1.092018000

O	1.460020000	0.184892000	-0.517467000
C	-0.989193000	-1.438496000	-0.484615000
H	-0.395885000	0.958729000	-0.603551000
H	-1.024453000	-0.916178000	-1.430321000
C	0.026075000	-2.296310000	-0.238158000
H	-1.873853000	-1.475441000	0.131891000
O	1.155103000	-2.363513000	-0.900372000
H	-0.018139000	-2.994336000	0.592551000
O	1.354806000	-0.014234000	-1.909830000
H	1.249169000	-1.580539000	-1.538269000

E0: -343.060933688771

ZPE: 0.091605

Frequencies:	-212.5452	144.7148	197.4182
	255.5758	308.7842	367.5488
	511.6534	541.3193	758.7919
	775.0936	871.7029	888.7673
	1011.4915	1018.4173	1054.6165
	1139.3328	1236.6571	1262.9379
	1340.7754	1425.8449	1443.8894
	1500.9733	1603.8491	1673.3029
	2792.2568	3153.9137	3172.7640
	3182.6412	3281.6593	3293.3838

Ketohydroperoxide

C	0.055564000	0.147660000	-0.043017000
H	0.137237000	0.252510000	1.039620000
O	1.396464000	0.099745000	-0.474440000
C	-0.728046000	-1.115652000	-0.427071000
H	-0.446563000	1.027165000	-0.450749000
H	-0.954080000	-1.092463000	-1.491060000
C	0.114079000	-2.327913000	-0.143298000

H -1.656796000 -1.145766000 0.142155000
O 0.888898000 -2.782441000 -0.944167000
H 0.032170000 -2.779047000 0.861769000
O 1.397580000 -0.047638000 -1.886851000
H 1.494056000 -1.010423000 -1.977189000

E0: -343.147262083172

ZPE: 0.095130

Frequencies:	89.0389	148.0662	224.6892
	237.0168	383.9257	516.5165
	531.5177	563.0697	829.9229
	895.1895	951.3767	1000.4442
	1027.1262	1124.8504	1159.2852
	1236.8712	1272.5640	1341.3229
	1393.1324	1426.7418	1460.9086
	1479.6474	1492.4210	1845.4482
	2969.5597	3065.9673	3095.0239
	3120.4663	3156.1356	3718.8558

Alkoxy radical

O 2.152667000 0.953333000 -0.973840000
C 2.583501000 0.027440000 -0.067856000
C 2.504973000 0.488214000 1.379162000
C 2.838235000 -0.601539000 2.356146000
O 3.076401000 -1.734476000 2.034772000
H 1.949316000 -0.862514000 -0.220624000
H 3.590066000 -0.323440000 -0.331784000
H 3.177054000 1.331814000 1.557741000
H 1.500134000 0.855502000 1.607297000
H 2.850079000 -0.308280000 3.422021000

E0: -267.3937207

ZPE: 0.075595

Frequencies:	61.0572	163.3842	202.7942
	399.9409	563.8936	682.2130
	715.3779	918.0278	963.1604
	1067.9424	1097.5702	1177.3395
	1265.4787	1314.8923	1366.9097
	1406.3194	1423.0652	1454.3203
	1851.7393	2958.5647	2960.8874
	3010.5793	3060.8281	3096.0159

Hydroxyl radical

O	0.000000000	0.000000000	0.000535000
H	0.000000000	0.000000000	0.972465000

E0: -75.67395032

ZPE: 0.008582

Frequency: 3767.2169

REFERENCES

[S1] Tarjan, R. E. Graph Algorithms in Chemical Computation. In *Algorithms for Chemical Computations*; Christoffersen, R. E., Ed.; ACS Symposium Series 46; American Chemical Society: Washington, DC, 1977.

[S2] Ugi, I.; Bauer, J.; Brandt, J.; Friedrich, J.; Gasteiger, J.; Jochum, C.; Schubert, W. *Angew. Chem., Int. Ed. Engl.* **1979**, *18*, 111.

[S3] Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T.B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X.; Ghosh, D.; Goldey, M.; Horn, P. R.; Jacobson, L. D.; Kaliman, I.; Khaliullin, R. Z.; Kuś, T.; Landau, A.; Liu, J.; Proynov, E. I.; Rhee, Y. M.; Richard, R. M.; Rohrdanz, M. A.; Steele, R. P.; Sundstrom, E. J.; Woodcock, H. L., III; Zimmerman, P. M.; Zuev, D.; Albrecht, B.; Alguire, E.; Austin, B.; Beran, G. J. O.; Bernard, Y. A.; Berquist, E.; Brandhorst, K.; Bravaya, K. B.; Brown, S. T.; Casanova, D.; Chang, C.-M.; Chen, Y.; Chien, S. H.; Closser, K. D.; Crittenden, D. L.; Diedenhofen, M.; DiStasio, R. A., Jr.; Do, H.; Dutoi, A. D.; Edgar, R. G.; Fatehi, S.; Fusti-Molnar, L.; Ghysels, A.; Golubeva-Zadorozhnaya, A.; Gomes, J.; Hanson-Heine, M. W. D.; Harbach, P. H. P.; Hauser, A. W.; Hohenstein, E. G.; Holden, Z. C.; Jagau, T.-C.; Ji, H.; Kaduk, B.; Khistyayev, K.; Kim, J.; Kim, J.; King, R. A.; Klunzinger, P.; Kosenkov, D.; Kowalczyk, T.; Krauter, C. M.; Lao, K. U.; Laurent, A. D.; Lawler, K. V.; Levchenko, S. V.; Lin, C. Y.; Liu, F.; Livshits, E.; Lochan, R. C.; Luenser, A.; Manohar, P.; Manzer, S. F.; Mao, S.-P.; Mardirossian, N.; Marenich, A. V.; Maurer, S. A.; Mayhall, N. J.; Neuscamman, E.; Oana, C. M.; Olivares-Amaya, R.; O'Neill, D. P.; Parkhill, J. A.; Perrine, T. M.; Peverati, R.; Prociuk, A.; Rehn, D. R.; Rosta, E.; Russ, N. J.; Sharada, S. M.; Sharma, S.; Small, D. W.; Sodt, A.; Stein, T.; Stück, D.; Su, Y.-C.; Thom, A. J. W.; Tsuchimochi, T.; Vanovschi, V.; Vogt, L.; Vydrov, O.; Wang, T.; Watson, M. A.; Wenzel, J.; White, A.; Williams, C. F.; Yang, J.; Yeganeh, S.; Yost, S. R.; You, Z.-Q.; Zhang, I. Y.; Zhang, X.; Zhao, Y.; Brooks, B. R.; Chan, G. K. L.; Chipman, D. M.; Cramer, C. J.; Goddard, W. A., III; Gordon, M. S.; Hehre, W. J.; Klamt, A.; Schaefer, H. F., III; Schmidt, M. W.; Sherrill, C. D.; Truhlar, D. G.; Warshel, A.; Xu, X.; Aspuru-Guzik, A.; Baer, R.; Bell, A. T.; Besley, N. A.; Chai, J.-D.; Dreuw, A.; Dunietz, B. D.; Furlani, T. R.; Gwaltney, S. R.; Hsu, C.-P.; Jung, Y.; Kong, J.; Lambrecht, D. S.; Liang, W.; Ochsenfeld, C.; Rassolov, V. A.; Slipchenko,

L. V.; Subotnik, J. E.; Van Voorhis, T.; Herbert, J. M.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Mol. Phys.* **2015**, *113*, 184.

[S4] Alecu, I. M.; Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2010**, *6*, 2872.

[S5] Elsamra, R. M. I.; Jalan, A.; Buras, Z. J.; Middaugh, J. E.; Green, W. H. *Int. J. Chem. Kinet.* **2016**, *48*, 474.

[S6] Chen, D.; Jin, H.; Wang, Z.; Zhang, L.; Qi, F. *J Phys. Chem. A* **2011**, *115*, 602.