Chemical Physics 450-451 (2015) 21–31

Contents lists available at ScienceDirect

Chemical Physics

journal homepage: www.elsevier.com/locate/chemphys

A flexible transition state searching method for atmospheric reaction systems

Xiao-Xiao Lin [a], Yi-Rong Liu [a], Teng Huang [a], Jiao Chen [a], Shuai Jiang [a], Wei Huang [a,b,*]

[a] Laboratory of Atmospheric Physico-Chemistry, Anhui Institute of Optics & Fine Mechanics, Chinese Academy of Sciences, Hefei, Anhui 230031, China
[b] School of Environmental Science & Optoelectronic Technology, University of Science and Technology of China, Hefei, Anhui 230026, China

ARTICLE INFO

Article history:
Received 16 October 2014
In final form 3 February 2015
Available online 11 February 2015

Keywords:
Monte Carlo
Atmosphere
VOCs
Bimolecular reactions
OH radical
Aerosol

ABSTRACT

The precise and rapid exploration of transition states (TSs) is a major challenge when studying atmospheric reactions due to their complexity. In this work, a Monte Carlo Transition State Search Method (MCTSSM), which integrates Monte Carlo sampling technique with transition state optimization methods using an efficient computer script, has been developed for transition state searches. The efficiency and the potential application in atmospheric reactions of this method have been demonstrated by three types of test suits related to the reactions of atmospheric volatile organic compounds (VOCs): (1) OH addition, (2) OH hydrogen-abstraction, and (3) the other reactive group (e.g. Cl, O_3, NO_3), especially for the reaction of β-pinene-sCI (stabilized Criegee Intermediates) with water. It was shown that the application of this method with effective restricted parameters has greatly simplified the time-consuming and tedious manual search procedure for transition state (TS) of the bimolecular reaction systems.

1. Introduction

Volatile organic compounds (VOCs) are essential components of tropospheric chemistry and play a critical role in atmospheric chemistry [1–4]. Goldstein and Galbally estimated that there were 10,000 to 100,000 different organic compounds in the atmosphere [3]. Each VOC can produce a series of oxidized products through a number of atmospheric degradation processes. The products formed in the subsequent chemistry processes can contribute to secondary organic aerosol (SOA) formation and growth [5–7] and will thus directly impact air quality and the global climate [8–10], as well as human health and the environment [11–13]. The reactions of VOCs are primarily initiated by OH, O_3, NO_3, and Cl oxidants [14]. Among of them, the hydroxyl radical (OH) is particularly important because it controls the removal of most atmospheric pollutants, such as carbon monoxide (CO) and VOCs [15,16]. These reactions are often accompanied by a series of subsequent oxidation steps, such as alkyl, alkoxy, alkylperoxy and Criegee intermediates reactions, which make the chemical transformation mechanisms very intricate [1]. It is usually difficult to obtain a clear picture of various reaction channels simply from experimental studies and generally requires theoretical explanation or predication.

To explore reaction processes, where the rate-determining step is particularly important, the elusive transition state structures must be found. A transition state (TS) which can be determined by searching for the first-order saddle points on potential energy surfaces (PES), is the state corresponding to the highest energy along a particular reaction coordinate and is essential for analyzing the reaction path and rate. Usually, the determination of a TS is performed via manual searching, which begins from guessing a possible structure based on the reactant and product configurations. The performance of manual searching typically requires one should be familiar with the reaction mechanism and with good chemical intuition. However, this is not true in most cases, especially for those unknown chemical reactions. In addition, chemical reactions may contain multiple transition state structures and a variety of intermediates exist on the reaction pathways due to the complexity of the reaction systems. Even researchers specialized in chemical reactions may have difficulty in finding out all transition states. To determine the correct transition states is usually a tedious, time-consuming, and error-prone process. Therefore, a systematic and effective TS search method is urgently needed for the automatic exploration of transition state structures.

In this work, we developed a simple and efficient method called Monte Carlo Transition State Search Method (MCTSSM) which combines the Monte Carlo sampling technique and the TS optimization methods to determine the transition states of complex atmospheric reactions. Similar to our successful Basin-Hopping (BH) method with the Monte Carlo sampling technique and a local

* Corresponding author at: School of Environmental Science & Optoelectronic Technology, University of Science and Technology of China, Hefei, Anhui 230026, China.
E-mail address: huangwei6@ustc.edu.cn (W. Huang).

http://dx.doi.org/10.1016/j.chemphys.2015.02.002
0301-0104/

optimization method to obtain global minimum structures of atomic and molecular clusters [17–25], the MCTSSM method started from a Monte Carlo randomly generated initial guess based on the reactant configurations, and the initial guess was then optimized via local TS optimization methods implemented in quantum chemistry package like Gaussian [26]. MCTSSM can be combined with Gaussian 09 [26] or DMol^3 [27,28]. In principle, any TS optimization algorithm implemented in Gaussian or DMol^3 can be employed in combination with the MCTSSM method. In the present work, we chose the default Berny optimization algorithm [29] implemented in Gaussian 09 for TS optimization. To save computational cost, Hartree–Fock (HF) theory and Density Functional Theory (B3LYP) with small basis set were chosen for initial TS optimization. Then, all the TSs found from MCTSSM using HF or B3LYP were re-optimized using the higher-level theory recommended in the relevant literature for a particular system.

Many different strategies for finding TS have been suggested [29–43]. The MCTSSM presented here differs from previous efforts in two ways: (1) The Monte Carlo sampling technique and TS optimization method (which is implemented in widely used quantum chemistry package like Gaussian) are integrated via a computer script. This manner is flexible and allows for changes to the script, suggesting it can be easily improved in the future. (2) The initial guesses are generated via Monte Carlo random sampling with simple parameters, and this gives the guarantee to find enough TSs. The MCTSSM strategies on the combination of Monte Carlo sampling technique and TS optimization method are summarized in Section 2. Then examples of applications of the method are discussed in Section 3. The reactions of OH radicals with the general atmospheric VOCs are primarily considered to be the test suites. Other attack group reactions, such as O_3, NO_3, Cl and H_2O, are also studied. Additionally, the reaction of β-pinene-sCI (stabilized Criegee Intermediates) with water is chosen as a new system for testing.

2. Methods

Our successful Basin-Hopping method has been widely used on structure exploration of atomic clusters and molecular clusters [17–25]. Here we present a new method applying a Monte Carlo sampling technique and TS optimization method for automatically searching the TS for resolving complex atmospheric reactions. To start TSs searching, only the reactant structures are needed as initial inputs for the Monte Carlo random sampling; *a priori* knowledge of the products is not required. Fig. 1 shows the flowchart of the method. The method for finding TSs can be summarized as follows:

(1) Input the known reactants structure (Cartesian coordinates of reactant A and reactant B) and some simple parameters. (Sampling parameter R represents an initial distance between atoms, M represents the number of TS searches.)
(2) Start from the given reactants structure, the initial random guess is generated automatically via Monte Carlo sampling.
(3) Invoke Gaussian 09 (or other quantum chemistry package) to optimize the initial guess with the key words "opt = (ts, calcfc, maxcyc = 500) freq" in our script and calculate the frequency with low-cost electronic structure methods such as HF or DFT (method and basis set are set by the users in step (1)).
(4) Read the output frequencies. If the list of output frequencies has only one imaginary frequency (IF), then the optimized geometries with sampling parameter $R = R'$ are saved, and $R = R'$ replaces $R = R_0$ as the initial distance in the next cycle. If not, sampling continues with parameter $R = R_0$ until the structure with only one IF is found.

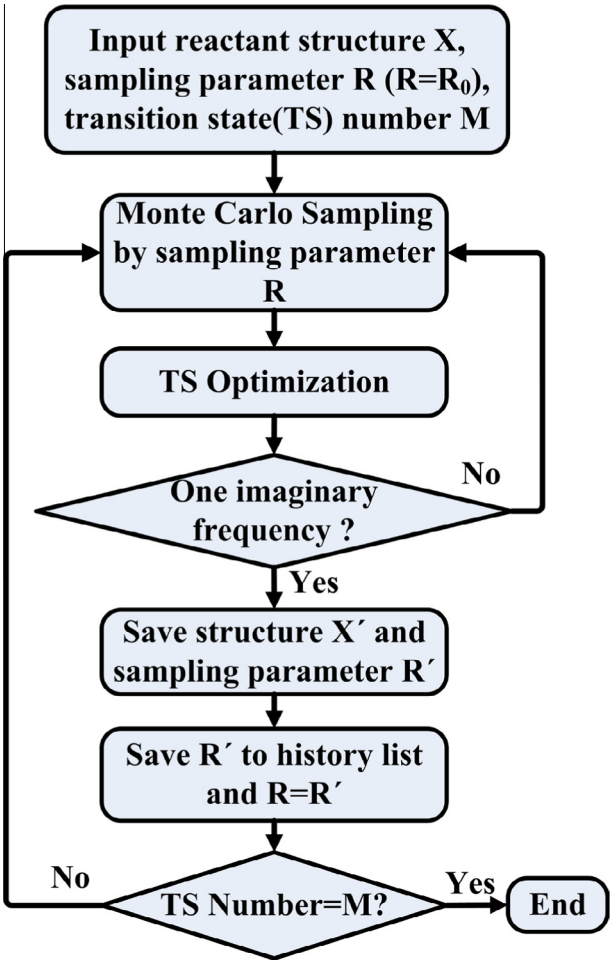


Fig. 1. Flowchart of the Monte Carlo Transition State Search Method (MCTSSM).

(5) Terminate the search. The TS searching number M is set to terminate the search. M defined by the user is based on the size of the system.

Sampling parameters are very important. The efficiency of the MCTSSM depends heavily upon the effective sampling. The initial guess generated via Monte Carlo sampling is completely random, thus structures taken from Monte Carlo (MC) sampling are often deviate from true TSs. As a result, the self-consistent field (SCF) or TS optimization calculation often fails to converge, which adds additional cost for finding the true TSs. Therefore, we introduced some reasonable empirical parameters to constrain traditional MC sampling. Such a sampling is not completely random and we call it "restricted" or "constrained" MC sampling. "Restricted" MC sampling can generate "higher quality" initial guesses which probably are located in the limited zone near the true TS.

Different from traditional Basin-Hopping (BH) method, which allows each atom to move independently, here we consider that the reaction system consists of two molecules, that is reactant A (such as OH radical) and reactant B (such as dimethyl ether). Generally, we designated reactant A (such as OH radical) as attacking molecule which attacked towards reactant B and rotate randomly around B (such as dimethyl ether), but the relative positions of atoms in A are fixed, and it is the same for molecule B. So the MCTSSM moves each molecule as an entirety, and that is to say, the relative positions of atoms in the same molecule are kept fixed. Sampling parameter R represents an initial distance between atoms which has been defined in step (1). In principle, there are

$N(N-1)/2$ distances R_{ij} between atom i and atom j, where N is the number of atoms, while not all the $N(N-1)/2$ R_{ij} are considered. For a specific reaction system, we can also designated atom i from reactant A as attacking atom and atom j from reactant B as attacked atom, which makes the initial guesses close to the zone near the true TS more easily. For example, in the H atom abstraction reaction of 1-Butene + OH, the attack atom is O from OH radical and the attacked atom is H from 1-Butene molecule. Separations between atom i and atom j that are too larger or too small will result in a failure of the optimization and require additional computation cost. Typically, R_0 is an initial approximate value that is often set between 1.0 and 2.0 Å. For example, if the R_0 is set to 1.2, then the MCTSSM started random sampling around the distance of 1.2 Å, such as 1.19, 1.18 or 1.21. After TS optimization, if the program found the structure with only one IF. Then the optimized structural parameter R' (such as $R' = 1.23$) will be saved and $R = R'$ replaces $R = R_0$ as the initial distance in the next cycle. The rough R_0 is adjusted to the appropriate value. In other words, the sampling parameter R is adaptive. However, it will take much more time for MCTSSM to get the appropriate value and the efficiency of searching reduced greatly if the R_0 deviate too far from the optimal value. For not missing some important TSs, we often perform MCTSSM for several runs with different R_0. Using this strategy, our method could find TSs more easily and quickly. The TS searching number M is defined to terminate the search. For example, if $M = 10$, the program does not stop running until it finds 10 TSs. However, if the value of M is large enough, the program will repeatedly find the same TSs, and if M is set to only 1, many TSs will be lost. To solve this issue, we can perform MCTSSM for several runs until no new TS structures are found. In this study, three searches have been performed and the value of M ranges from small to large (Table 3). The total number of the different TSs found is obtained from three searches which guarantees not missing some key TSs.

MCTSSM can be combined with Gaussian 09 and invokes TS optimization algorithm implemented in Gaussian 09 in step (3). To find the transition states within a short period of time, low-cost electronic structure methods were selected for initial TS searches. Higher-level calculations usually give higher accuracy but come at a significantly higher computational cost. Thus, for our TS search method, HF and B3LYP levels of theory were chosen for TS initial optimization and frequency computation. TSs obtained from HF and DFT methods with small basis set are rough structures and further higher-level structural optimization should be followed. To verify the TSs found by MCTSSM, all the TS structures were re-optimized using the higher-level theory recommended in the relevant literatures. Cartesian coordinates of all the TS structures obtained in the initial search from MCTSSM and the re-optimized TS structures at a higher level can be found in Supplementary material.

3. Results and discussions

The following three types of test suites consisting of 13 reactions were chosen for verifying the utility of MCTSSM, as summarized in Tables 1 and 2. The size of the test system ranges from 7 to 30 atoms. Basis sets ranging from 3-21G to 6-311++G(d,p) are used along with either wave function theory (HF) or density functional theory (B3LYP). In each case, simple input parameters were used (i.e., one structure of reactant A + reactant B, sampling parameters, and a TS optimization method), and a wide variety of TSs were rapidly found. To confirm the structures obtained from MCTSSM searches, all the TSs were re-optimized using the level of theory suggested in the relevant references (Tables 1 and 2). Cartesian coordinates of all the TSs obtained in the initial search from MCTSSM and the re-optimized TSs at a higher level afterward, as well as the vibrational frequencies of them can be found in Supplementary material. The number of Monte Carlo samplings, the number of TSs found, and the average time are listed in Table 3. As shown in Table 3, over 50% initial guesses generated by restricted Monte Carlo sampling lead to successful TSs for most reactions. The number of TSs found is determined by the parameter M (TS searching number) set by the users. However, among the structures found, some of them are identical or wrong TSs by checking the imaginary frequency vibrational modes and intrinsic reaction coordinate (IRC) calculations. The number of unique TS structures is given in parenthesis (Table 3). For example, in the case of the

Table 1
Test suite of OH radical reactions used for MCTSSM.

OH-initiated reaction	Number of atoms	Initial search		Further TS optimization		Ref.
		Basis set	Theory	Basis set	Theory	
DME H atom abstraction	9	6-31G	UHF	6-311++G(d,p)	M05-2X	50
Skew-1-butene H atom abstraction	14	6-31+G(d)	UHF	6-311G(d,p)	BH&HLYP	52
Syn-1-butene H atom abstraction	14	6-31++G(d,p)	UHF	6-311G(d,p)	BH&HLYP	52
Pinonaldehyde H atom abstraction	30	6-31G	UHF	6-31G*	MP2	60
Halogen substituted propene OH addition	11	6-31G	UHF	cc-pVTZ	BH&HLYP	63
β-pinene OH addition	28	6-31G	UHF	6-31G(d)	MP2	62

Table 2
Test suite of other attack group reactions used for MCTSSM.

Other attack group reactions	No. of atoms	Initial search		Further TS optimization		Ref.
		Basis set	Theory	Basis set	Theory	
$CH_3CN + Cl \rightarrow CH_2CN + HCl$	7	3-21G	UHF	6-311G(d,p)	BH&	
HLYP	7					
$H_2S + NO_3 \rightarrow HNO_3 + SH$	7	6-31G	UHF	6-311++G(3df,3pd)	B3LYP	71
$H_2S + NO_3 \rightarrow H_2SO + NO_2$	7	6-311++G(d,p)	UB3LYP	6-311++G(3df,3pd)	B3LYP	71
Ethylene + $O_3 \rightarrow$ POZ	9	6-31G	RHF	6-311G(2d,d,p)	B3LYP	72
Isoprene + $O_3 \rightarrow$ POZ	16	6-31G	RHF	6-311G(2d,d,p)	B3LYP	72
β-pinene-CI1 + $H_2O \rightarrow C_8H_{14}C(OH)OOH$	28	6-31 + G(d)	RHF	6-311 + G(2d,p)	M06-2X	This work
β-pinene-CI2 + $H_2O \rightarrow C_8H_{14}C(OH)OOH$	28	6-31 + G(d)	RHF	6-311 + G(2d,p)	M06-2X	This work

Table 3
Performance of MCTSSM integrates Monte Carlo sampling technique with TS optimization algorithm.

Reaction	1		2		3		Average time (min) [c]
	NO. of samplings[a]	NO. of TS found[b]	NO. of samplings[a]	NO. of TS found[b]	NO. of samplings[a]	NO. of TS found[b]	
DME H atom abstraction	4(75%)	3(3)	8(75%)	6(3)	16(56%)	9(3)	5
Skew-1-butene H atom abstraction	16(50%)	8(8)	21(76%)	16(8)	37(65%)	24(8)	7
Syn-1-butene H atom abstraction	10(80%)	8(8)	23(70%)	16(10)	30(80%)	24(10)	26
Pinonaldehyde H atom abstraction	22(73%)	16(13)	38(60%)	23(16)	42(71%)	30(16)	60
Halogen substituted propene OH addition	6(50%)	3(2)	8(75%)	6(2)	12(75%)	9(2)	8
β-pinene OH addition	4(100%)	4(2)	10(80%)	8(4)	16(75%)	12(4)	41
CH_3CN + Cl H atom abstraction	2(100%)	2(1)	4(100%)	4(1)	8(75%)	6(1)	1
$H_2S + NO_3$ H atom abstraction	4(50%)	2(1)	5(80%)	4(1)	10(60%)	6(1)	5
$H_2S + NO_3$ O atom abstraction	4(50%)	2(1)	6(67%)	4(1)	11(73%)	8(1)	120
Ethylene ozonolysis	4(100%)	4(1)	9(89%)	8(1)	14(86%)	12(1)	2
Isoprene ozonolysis	6(67%)	4(1)	13(62%)	8(2)	19(63%)	12(2)	11
β-pinene-CI1 water addition	7(57%)	4(2)	14(57%)	8(4)	18(67%)	12(4)	36
β-pinene-CI2 water addition	8(50%)	4(2)	13(62%)	8(4)	16(75%)	12(4)	38

[a] Initial guesses (%) of the total number of samplings converge to the TSs is in parentheses.
[b] The number of unique TS structures is in parenthesis.
[c] The average time of MCTSSM spends in finding every TS structure. Computations were performed using 2 core.

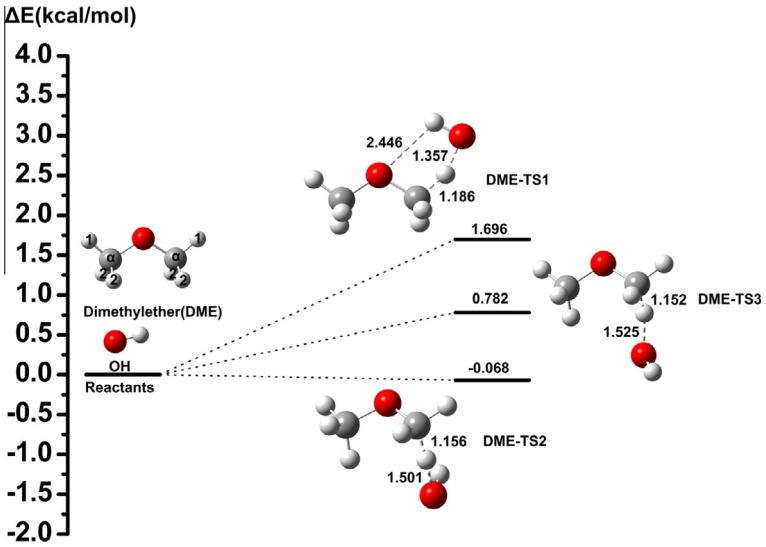


Fig. 2. Energy profile of the potential energy surface of the OH hydrogen abstractions from dimethyl ether (DME) obtained at the M05-2X/6-311++G(d,p) level. The energies of transition states TS1–TS3 are related to the energies of reactants at 0 K including ZPE correction. Initial TSs search were performed at the HF/6-31G level. For convenience, H atoms, C atoms, and O atoms are color coded in white, grey, and red, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

reaction of dimethyl ether (DME) H atom abstraction, upon performing the MCTSSM for the second time, 8 initial guesses were generated and 6 (3 unique) TS structures were found. Table 3 also indicates that the calculation time cost increases with number of atoms and is relevant to the choice of HF or DFT levels of theory. For instance, the average calculation time of MCTSSM spends on the $H_2S + NO_3$ O atom abstraction reaction at the UB3LYP/6-311++G(d,p) level is more than that of $H_2S + NO_3$ H atom abstraction reaction at the UHF/6-31G level with the same atoms. Various reactions will be discussed in detail below.

3.1. H atom abstraction reactions

A test suite that consists of 3 OH hydrogen abstraction reactions was chosen in this study, and the corresponding theoretical methods and basis sets are summarized in Table 1. For the systems of the H-abstraction reactions, we designate the attack atom O from OH radical and the attacked atom H from VOCs. The R_0 is often set between 1.1 and 1.3 Å for the H-abstraction reactions. Various transition state structures from different reaction sites of parent molecular hydrogen were obtained using the MCTSSM. Moreover, new TS structures were found (that had not been found in previous research).

3.1.1. Dimethyl ether (DME)

Ethers are used worldwide and considerable amounts of ethers are released into the atmosphere. Dimethyl ether (DME) is one of the most studied ethers that reacts with OH radicals [44,45]. There are 3 different TSs corresponding to H atom abstractions of dimethyl ether found in the three runs of MCTSSM (Table 3). The structures of TSs found are depicted in Fig. 2. From Fig. 2, one can easily observe that there are two types of H atoms (1-H and 2-H), which are

Table 4
Reaction barriers relative to isolated reactants electronic energies with zero-point correction (ZPE) included (in kcal mol^{-1}) at different levels of theory.

Levels energy kcal mol^{-1}	MP2/6-311++G(d,p)	B3LYP/6-311++G(d,p)	M05-2X/6-311++G(d,p)	M06-2X/6-311++G(d,p)
DME-TS1	5.916	−1.650	1.696	1.517
DME-TS2	4.478	−5.282	−0.068	−0.134
DME-TS3	5.064	−4.948	0.782	0.582

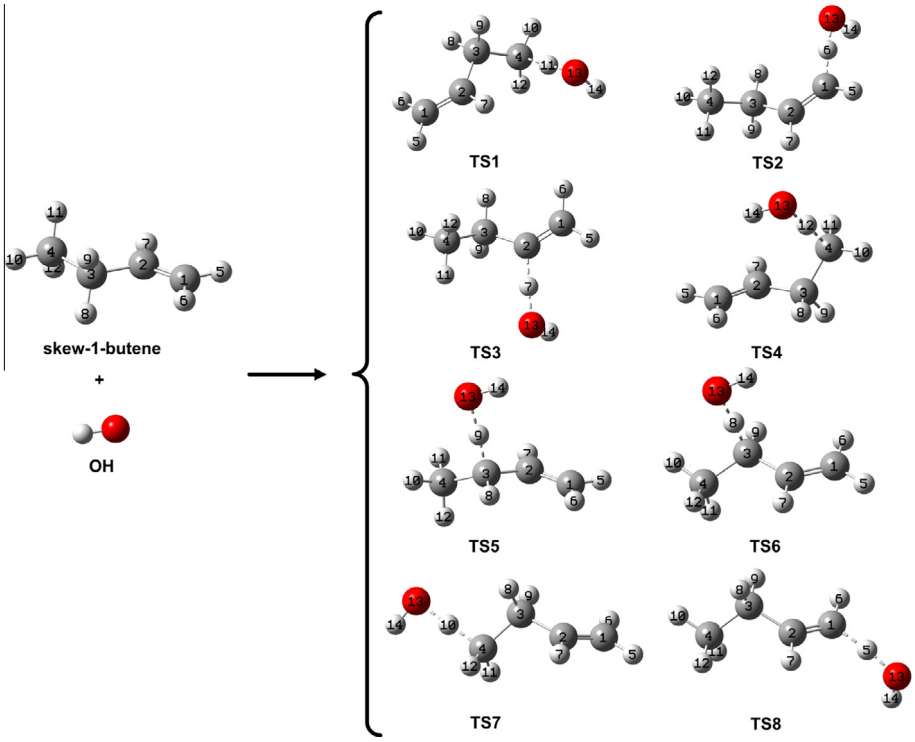


Fig. 3. Transition state structures corresponding to OH hydrogen abstractions from skew-1-butene got using MCTSSM at the HF/6-31+G(d) level. For convenience, H atoms, C atoms, and O atom are color coded in white, grey, and red, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

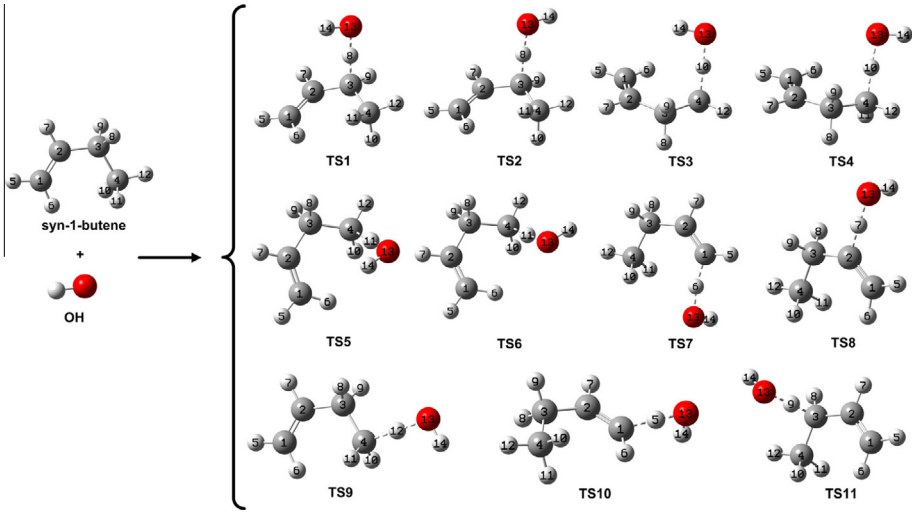


Fig. 4. Transition state structures corresponding to OH hydrogen abstractions from syn-1-butene got using MCTSSM at the HF/6-31++G(d,p) level. For convenience, H atoms, C atoms, and O atom are color coded in white, grey, and red, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

categorized by their positions. The earliest calculation by Good and Francisco [46] identified a single transition state for this H atom abstraction. Subsequent calculations [47–49] recognized two transition states with the reaction occurring with out-of-plane and in-plane geometries. Atadinc et al. [48] and El-Nahas et al. [49] were able to differentiate between two different out-of-plane abstraction

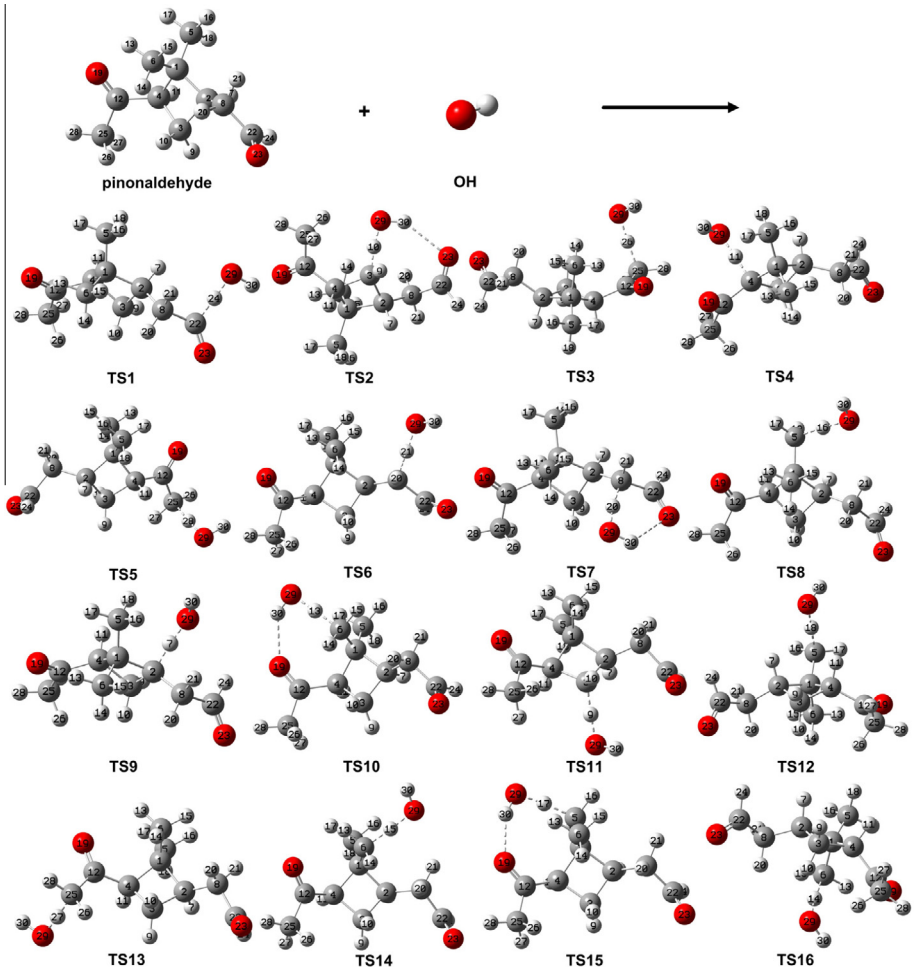


Fig. 5. Transition state structures corresponding to OH hydrogen abstractions from pinonaldehyde ($C_{10}H_{16}O_2$) calculated using MCTSSM at the HF/6-31G level. For convenience, H atoms, C atoms, and O atoms are color coded in white, grey, and red, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

pathways corresponding to DME-TS2 and DME-TS3 (Fig. 2). It can be seen that, just for the relative simple reaction system, researchers could hardly find out all transition states easily. It is very inspiring that our code can find all of the three TSs when running only once (see Table 3). The energy of a TS structure has a considerable impact on rate constants; if a TS structure is missed, it may lead to a large discrepancy between the calculated rates constants and the measured rates constants. The transition states of DME + OH searched using MCTSSM were further optimized at the M05-2X/6-311++G(d,p) level. The geometries of DME-TS1 and DME-TS2 are in agreement with those reported in the literature [50]. We also labeled out the electronic energies with the zero-point energy (ZPE) correction of DME TSs. The energies of DME-TS1, DME-TS2, and DME-TS3 relative to the energies of reactants are 1.696 kcal/mol, −0.068 kcal/mol, and 0.782 kcal/mol, respectively.

The three transition states identified for OH + DME reported by Carr et al. [51] are connected by internal rotation and are not independent. DME-TS1 and DME-TS3 are conformers of the lowest-energy saddle point DME-TS2 [51]. The aim of our method is to find the transition states as many as possible including different conformers of the transition state, and the code is proved useful and effective. To verify the geometry of DME-TS3, the TSs were also optimized at the MP2/6-311++G(d,p), B3LYP/6-311++G(d,p), M05-2X/6-311++G(d,p) and M06-2X/6-311++G(d,p) theory levels. The single-point energies with zero-point correction (ZPE) included of TSs at the various levels are listed in Table 4. It can be seen that DME-TS3 is closer in energy to DME-TS2 than that of DME-TS1 at various theoretical levels (MP2, B3LYP etc.). This is due to that the minor difference between DME-TS2 and DME-TS3 caused by the orientation of the OH group.

3.1.2. 1-butene

Butene is the smallest alkene that possesses isomers. For 1-butene, there are two rotational conformers in skew (C_1) and syn (C_s) forms, respectively. We selected skew-1-butene and syn-1-butene for testing; the results are shown in Figs. 3 and 4, respectively. Because each conformer contains eight different hydrogens (allylic, vinylic and alkyl hydrogens) according to their different positions, the transition state structures that were searched using MCTSSM were found to include eight TS structures (at least) for each conformer. Interestingly, when we increased the transition state searching number (M), the program found more structures regarding the hydrogen atom H-10 and H-11 in the reaction of syn-1-butene with OH. TS3 and TS5 of syn-1-butene with OH were not found in the first search maybe M was not sufficiently large. The TS1 and TS2 in the reaction of syn-1-butene with OH displayed in Fig. 4 are consistent with the prediction from the theoretical studies by Sun and Law [52]. As they discussed, TS1 is an OH rotamer of TS2. Similarly, TS3 and TS5 are OH rotamers of TS4 and TS6, respectively, which were not mentioned previously.

3.1.3. Pinonaldehyde

α-pinene is one of the most abundant terpenes emitted into the troposphere, and pinonaldehyde ($C_{10}H_{16}O_2$) is the main oxidation

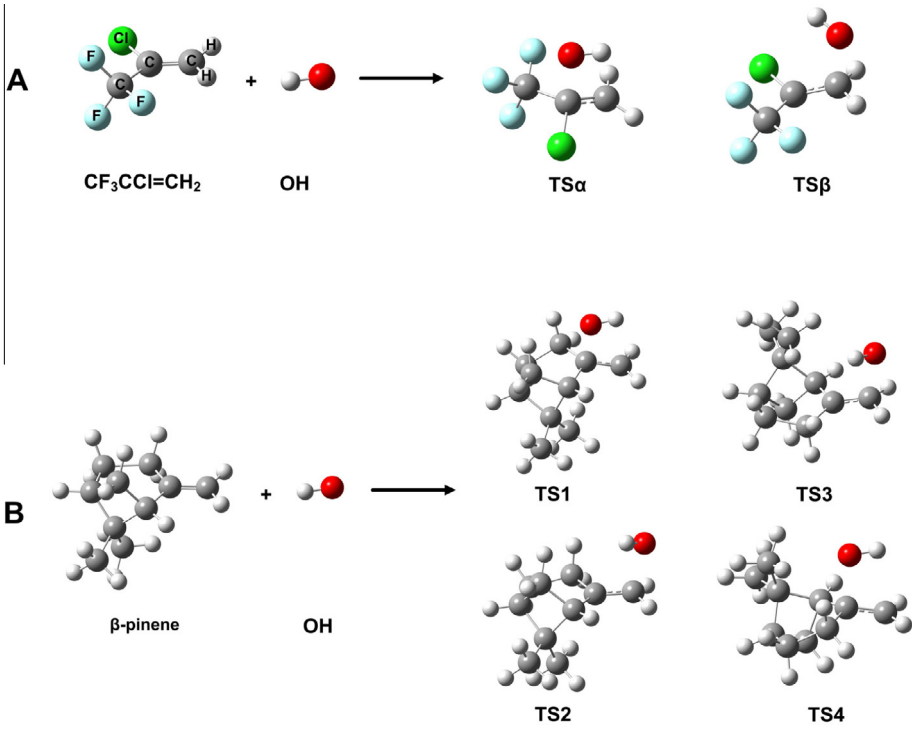


Fig. 6. Transition state structures corresponding to OH addition to $CF_3CCl{=}CH_2$ and β-pinene double bond calculated using MCTSSM at the HF/6-31G level. For convenience, H atoms, C atoms, Cl atom, F atoms and O atom are color coded in white, grey, green, blue, and red, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

product of α-pinene [53,54]. Pinonaldehyde has a several-hour lifetime and will react with OH and NO_3 radicals [55,56]. There are already several experimental and theoretical investigations on the reactions of pinonaldehyde with OH radicals [57–60]. We chose the pinonaldehyde ($C_{10}H_{16}O_2$) H atom abstraction reaction by the OH radical for testing. In the first run of MCTSSM, the number of initial guesses by Monte Carlo sampling was 22, and 73% initial guesses converged successfully with one imaginary frequency. However, there were some duplicate structures and 13 different TSs were found (Table 3). In the second run, the 3 more different TSs were found. In the third run, 16 different TSs were found as the same as the second run. All the 16 different TSs are presented in Fig. 5. Due to the large size of the reaction system which contains 30 atoms, it spent much more time to find the exact TSs than the small reactions (Table 3). All the TSs searched at HF/6-31G level were further optimized at MP2/6-31G*.

3.2. OH radical addition reactions

OH radical addition is a typical and important reaction in atmospheric chemistry, and most transition states in such reactions are formed with an OH attack on the double bond. A series of reaction systems, from simple to complex, have been chosen for test. We designate the attack atom O from OH radical and the attacked atom C from VOCs. The R_0 is often set between 1.6 and 1.8 Å for the OH radical addition reactions. The transition states obtained using MCTSSM are in good agreement with previous theoretical studies for the clear addition sites [61–64]. Particularly important is that

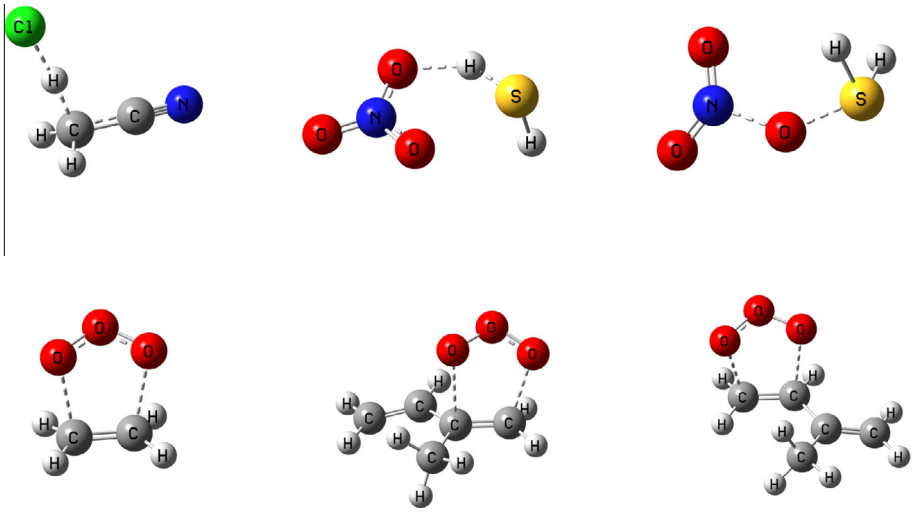


Fig. 7. Transition state structures corresponding to Cl, NO_3, H_2S and O_3 attack on acetonitrile, NO_3, H_2S and alkenes calculated using MCTSSM at the HF/3-21G, HF/6-31G, B3LYP/6-311++G(d,p), and HF/6-31G levels of theory, respectively.

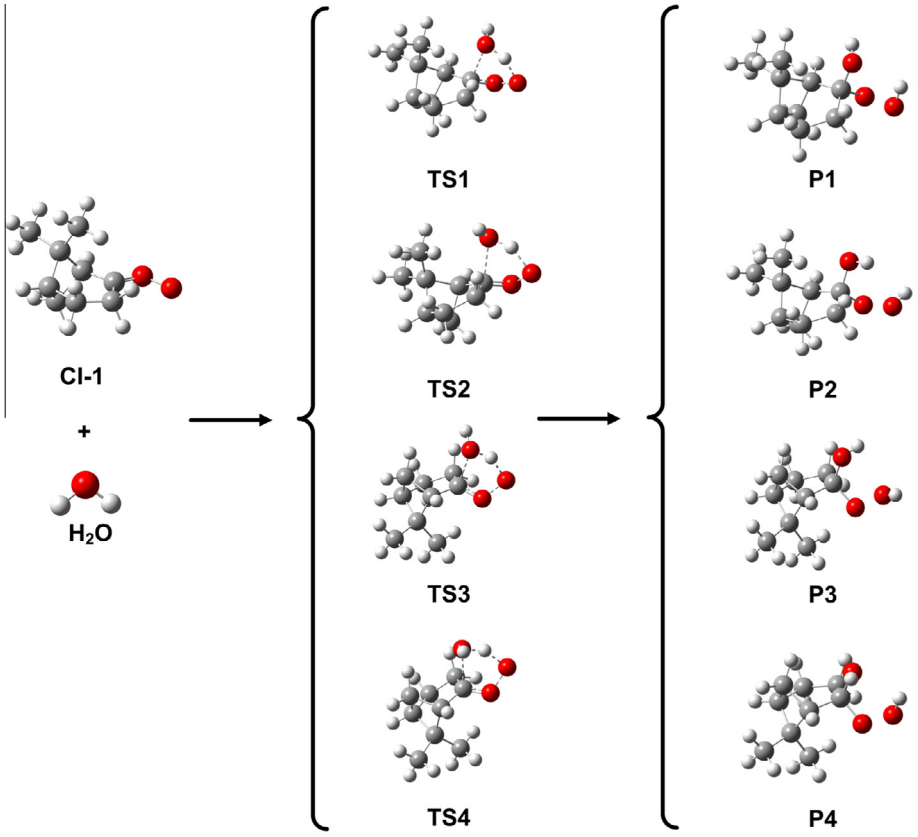


Fig. 8. Structure optimizations of reactants, transition states, and products were calculated at the M06-2X/6-311+G(2d,p) level. Initial TSs searching of β-pinene sCI1+H_2O were performed at the HF/6-31+G(d) level. For convenience, H atoms, C atoms, and O atoms are color coded in white, grey, and red, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the desired transition state structures were found within a very short time, and the method only requires reactant structures as prerequisites.

3.2.1. Halogen-substituted propene

Hydrochlorofluorocarbons (HCFCs) have small but non-negligible ozone depletion potentials because they contain chlorine [65]. However, due to the efficient scavenging of HCFCs by OH radicals in the troposphere, the ability of HCFCs to destroy stratospheric ozone is negligible [63]. OH radicals react with unsaturated HCFCs primarily via OH radical addition to the C=C double bond. The reaction of $CF_3CCl{=}CH_2$ with OH was chosen for testing, and the corresponding transition structures of OH radical addition to halogen-substituted propene are shown in Fig. 6(A). The geometries of TSs were found with low-level theoretical methods and a small basis set. TS structures were further optimized at the BH&HLYP/cc-pVTZ level of theory, and the results are in agreement with the literature [63].

3.2.2. β-pinene

Biogenic VOCs play a dominant role in the chemistry of the lower troposphere and atmospheric boundary layer [66]. As one of most abundant terpenes in the troposphere, β-pinene contributes significantly to the origin of tropospheric ozone and fine organic aerosols [61,62]. The degradation of pinenes is initiated primarily by OH radical addition to the C=C double bond [61]. Fig. 6(B) shows the geometries of TS structures corresponding to four pathways. These transition states obtained at the HF/6-31G level using MCTSSM were further optimized at the UMP2/6-31G(d) level (Table 1). It is shown that the optimized TS structures are in agreement with the literature [62].

3.3. Other attack group reactions

To further test MCTSSM, we chose other common attack groups in the earth's atmosphere such as chlorine, NO_3, O_3, and H_2O [67–72]. We did not specify the attack atom and the attacked atom for other attack group reactions. The optimal R_0 value is unknown for us and we only guess an initial rough value. All of the TSs found when performing MCTSSM for several runs with different R_0. The corresponding reactions, theoretical methods and basis sets are given in Table 2. It is noteworthy that the reactions of β-pinene stabilized Criegee Intermediates (β-pinene-sCI) + H_2O, to the best of our knowledge, have not been studied. For the simple reaction of chlorine with acetonitrile, it takes only an average of 1 min to find the transition state at the HF/3-21G. We also applied the HF/3-21G level search for an NO_3 hydrogen abstraction reaction with H_2S, and MCTSSM ran three times for repeat searching but could not find the transition states. Thus, we changed the basis sets from 3-21G to 6-31G and found the desired TSs. For the H abstraction from H_2S by O, we attempted to find the corresponding transition states using the same Hartree–Fock (HF) method. However, it was difficult to find the desired transition states. B3LYP-DFT method was chosen instead of HF which has been used before, and was proved to be effective in finding the transition state. Therefore, for different reaction systems, one needs to select the appropriate method and basis set. Usually, low-level theoretical methods and a small basis set for the initial search are suggested in order to complete the transition state structures search in the shortest amount of time. The transition state structures of other attack group reactions such as NO_3, H_2S, and O_3 are shown in Fig. 7.

Criegee intermediates (carbonyl oxides) are key intermediates in the ozonolysis of alkenes in the atmosphere. There have been many studies on the bimolecular reactions of stabilized CI (sCI)

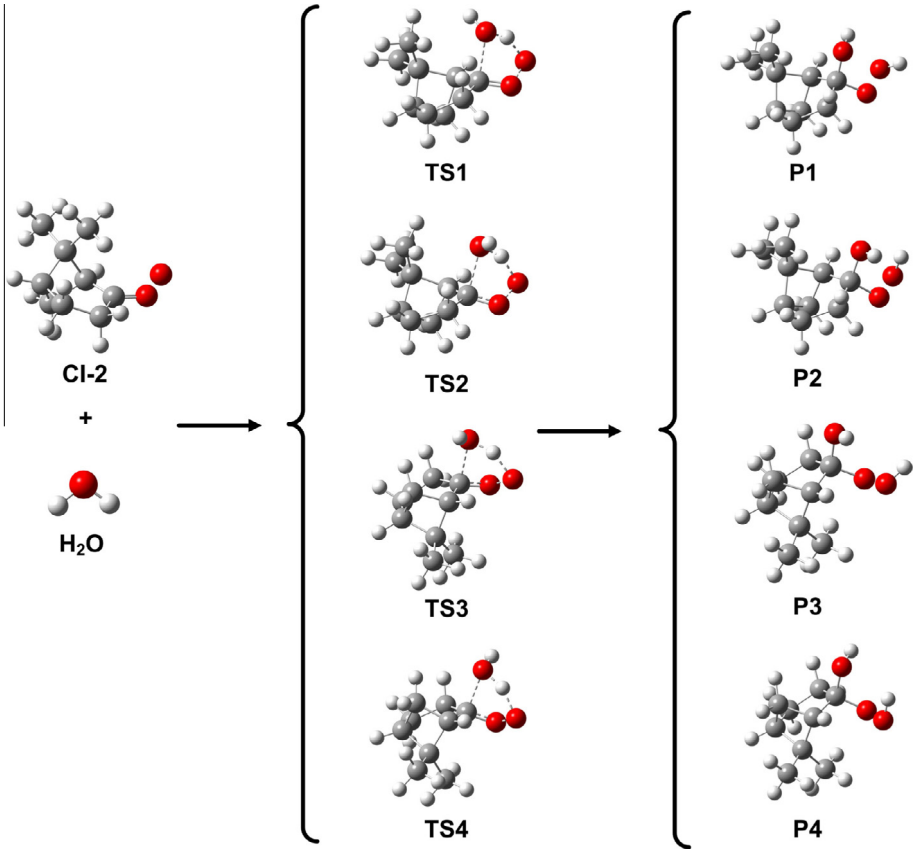
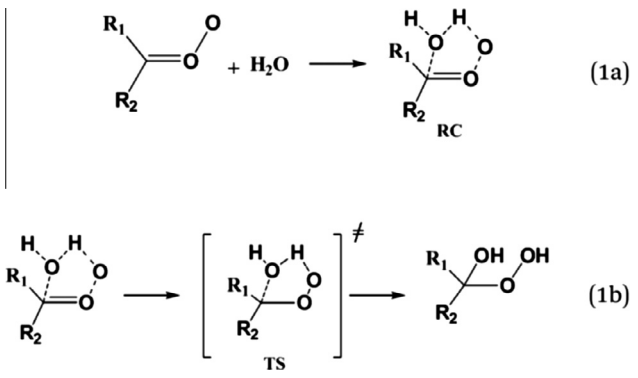


Fig. 9. Structure optimizations of reactants, transition states, and products were calculated at the M06-2X/6-311+G(2d,p) level. Initial TSs searching of β-pinene sCI2+H_2O were performed at the HF/6-31+G(d) level. For convenience, H atoms, C atoms, and O atoms are color coded in white, grey, and red, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

with NO_2, SO_2, and CO [73–76], and most are small sCI (one to four carbon atoms). Small CI such as CH_2OO will not be stabilized to a significant extent at atmospheric pressure, while larger CI with more carbons can be stabilized [1]. The reaction of sCI with water is often considered a major loss process for an atmospheric sCI in most instances. β-pinene-sCI are relatively heavily substituted Criegee intermediates and stabilize for approximately 40% at 1 atm. pressure [77]. Therefore, to learn the fate of β-pinene-sCI in the atmosphere can be very important.

It is well known that when sCI reacts with water, α-hydroxy hydroperoxides are primarily generated [74,78,79]. The reactions of the stabilized Criegee intermediate and H_2O which produce α-hydroxy hydroperoxides proceed as shown in reactions (1a) and (1b) [79]:



For β-pinene-sCI, there are two isomers, *anti*-sCI and *syn*-sCI. To improve the searching efficiency for the complex reactions, we designated O atom from water as the attack atom and C atom from Criegee intermediate as the attacked atom. The MCTSSM found 4 transition states for each isomer. Further structure optimizations of the transition states were calculated using the Density Functional Theory (M06-2X) method with the 6-311+G(2d,p) basis set. The corresponding TS structures are shown in Figs. 8 and 9. To confirm that the transition states connect the hydroxy-hydroperoxide channel, intrinsic reaction coordinate (IRC) calculations were performed at the M06-2X/6-311++G(2d,p) level, and the products were used to verify the correct transition states that connect the hydroxy-hydroperoxide reaction path. The corresponding hydroxy-hydroperoxide products calculated at the M06-2X/6-311++G(2d,p) level are also listed in Figs. 8 and 9. They were obtained by optimizing the final structures from IRC calculations at the products side.

4. Conclusions

In this work, an efficient method called MCTSSM for transition states searching by combing the Monte Carlo sampling technique with a transition state optimization method has been presented. It provides a new approach to automatically find TSs in an efficient manner via a computer script and with simple input parameters not heavily reliant on human intuition. The introduction of some effective constraints such as R_0 can greatly increase the efficiency of the MCTSSM for specific reactions such as OH addition reactions. The ability for the TS searching is demonstrated in three types of test suits related to the reactions of atmospheric volatile organic compounds (VOCs): (1) OH addition, (2) OH hydrogen-abstraction,

and (3) the other reactive group (e.g. Cl, O_3, NO_3). As shown in the present study, MCTSSM is very powerful in the systematic TS search for atmospheric reactions. The method could not only find the structures of TSs previously reported, but find some new TSs within a short period of time. The potential application in atmospheric reactions of this method has been further demonstrated by studying the reaction of β-pinene-sCI (stabilized Criegee Intermediates) with water, which has not been reported.

The MCTSSM can be easily combined with many quantum chemistry packages such as the Gaussian, DMol[3], and so on. In this work, Berny optimization algorithm implemented in Gaussian 09 has been invoked for TS optimization. In the future, some other TS optimization algorithms implemented in quantum chemistry packages will be employed in our method. At present, MCTSSM focuses on atmospheric bimolecular reactions. Unimolecular reactions such as isomerization or elimination, and the three-body reactions involvement of the third body such as H_2O and HCOOH as a proton transfer catalyst are not considered and will be implemented soon.

Conflict of interest

There is no conflict of interest.

Acknowledgments

The study was supported by grants from the National Natural Science Foundation of China (21133008), and "Interdisciplinary and Cooperative Team" of CAS. Acknowledgement is also made to the "Thousand Youth Talents Plan". Part of the computation was performed at the Supercomputing Center of the Chinese Academy of Sciences and Supercomputing Center of USTC.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.chemphys.2015.02.002.

References

[1] L. Vereecken, J.S. Francisco, Chem. Soc. Rev. 41 (2012) 6259.
[2] J.D. Gouw, J.L. Jimenez, Environ. Sci. Technol. 43 (2009) 7614.
[3] A.H. Goldstein, I.E. Galbally, Environ. Sci. Technol. 41 (2007) 1514.
[4] G.P. Ayers, S.A. Penkett, R.W. Gillett, B. Bandy, I.E. Galbally, C.P. Meyer, C.M. Elsworth, S.T. Bentley, B.W. Forgan, Nature 360 (1992) 446.
[5] M. Hallquist, J.C. Wenger, U. Baltensperger, Y. Rudich, D. Simpson, M. Claeys, J. Dommen, N.M. Donahue, C. George, A.H. Goldstein, J.F. Hamilton, H. Herrmann, T. Hoffmann, Y. Iinuma, M. Jang, M.E. Jenkin, J.L. Jimenez, A. Kiendler-Scharr, W. Maenhaut, G. McFiggans, T.F. Mentel, A. Monod, A.S.H. Prevot, J.H. Seinfeld, J.D. Surratt, R. Szmigielski, J. Wildt, Atmos. Chem. Phys. 9 (2009) 5155.
[6] M.O. Andreae, P.J. Crutzen, Science 276 (1997) 1052.
[7] M. Kalberer, J. Yu, D.R. Cocker, R.C. Flagan, J.H. Seinfeld, Environ. Sci. Technol. 34 (2000) 4894.
[8] M. Kulmala, T. Suni, K.E.J. Lehtinen, M.D. Maso, M. Boy, A. Reissell, Ü. Rannik, P. Aalto, P. Keronen, H. Hakola, J. Bäck, T. Hoffmann, T. Vesala, P. Hari, Atmos. Chem. Phys. 4 (2004) 557.
[9] P.J. Adams, J.H. Seinfeld, J. Geophys. Res. 107 (2002) 4370.
[10] G. Oberdorster, E. Oberdorster, J. Oberdorster, Health Perspect. 113 (2005) 823.
[11] R. Derwent, Ø. Hov, J. Geophys. Res: Atmos. 93 (1988) 5185.
[12] A.M. Hough, J. Geophys. Res: Atmos. 93 (1988) 3789.
[13] D. Schuetzle, D. Cronn, A.L. Crittenden, R.J. Charlson, Environ. Sci. Technol. 9 (1975) 838.
[14] R. Atkinson, J. Arey, Chem. Rev. 103 (2003) 4605.
[15] A. Hofzumahaus, F. Rohrer, K. Lu, B. Bohn, T. Brauers, C.C. Chang, H. Fuchs, F. Holland, K. Kita, Y. Kondo, X. Li, S. Lou, M. Shao, L. Zeng, A. Wahner, Y. Zhang, Science 324 (2009) 1702.
[16] R. Atkinson, W.P.L. Carter, Chem. Rev. 84 (1984) 437.
[17] W. Huang, R. Pal, L.M. Wang, X.C. Zeng, L.S. Wang, J. Chem. Phys. 132 (2010) 054305.
[18] W. Huang, A.P. Sergeeva, H.J. Zhai, B.B. Averkiev, L.S. Wang, A.I. Boldyrev, Nat. Chem. 2 (2010) 202.
[19] S. Jiang, T. Huang, Y.R. Liu, K.M. Xu, Y. Zhang, Y.Z. Lv, W. Huang, Phys. Chem. Chem. Phys. 16 (2014) 19241.
[20] H. Wen, Y.R. Liu, T. Huang, K.M. Xu, W.J. Zhang, W. Huang, L.S. Wang, J. Chem. Phys. 138 (2013) 174303.
[21] L.L. Yan, Y.R. Liu, T. Huang, S. Jiang, H. Wen, Y.B. Gai, W.J. Zhang, W. Huang, J. Chem. Phys. 139 (2013) 244312.
[22] S. Jiang, Y.R. Liu, T. Huang, H. Wen, K.M. Xu, W.X. Zhao, W.J. Zhang, W. Huang, J. Comput. Chem. 35 (2014) 159.
[23] K.M. Xu, T. Huang, H. Wen, Y.R. Liu, Y.B. Gai, W.J. Zhang, W. Huang, RSC Adv. 3 (2013) 24492.
[24] Y.R. Liu, H. Wen, T. Huang, X.X. Lin, Y.-B. Gai, C.J. Hu, W.J. Zhang, W. Huang, J. Phys. Chem. A 118 (2014) 508.
[25] Y.P. Zhu, Y.R. Liu, T. Huang, S. Jiang, K.M. Xu, H. Wen, W.J. Zhang, W. Huang, J. Phys. Chem. A 118 (2014) 7959.
[26] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery, T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, H.P. Hratchian, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K .Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q.C., A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez, J.A. Pople, Gaussian 09W, Revision A.02; Gaussian; Inc.: Wallingford, CT, 2009.
[27] B. Delley, J. Chem. Phys. 92 (1990) 508.
[28] B. Delley, J. Chem. Phys. 113 (2000) 7756.
[29] H.B. Schlegel, J. Comput. Chem. 3 (1982) 214.
[30] A. Heyden, A.T. Bell, F.J. Keil, J. Chem. Phys. 123 (2005) 224101.
[31] S.K. Burger, W. Yang, J. Chem. Phys. 124 (2006) 054109.
[32] A. Goodrow, A.T. Bell, M. Head-Gordon, J. Chem. Phys. 130 (2009) 244108.
[33] A. Behn, P.M. Zimmerman, A.T. Bell, M. Head-Gordon, J. Chem. Phys. 135 (2011) 224108.
[34] D. Mei, L. Xu, G. Henkelman, J. Catal. 258 (2008) 44.
[35] J.M. Del Campo, A.M. Koster, J. Chem. Phys. 129 (2008) 024107.
[36] L. Xu, D. Mei, G. Henkelman, J. Chem. Phys. 131 (2009) 244520.
[37] C. Shang, Z.P. Liu, J. Chem. Theory Comput. 8 (2012) 2215.
[38] H.-F. Wang, Z.-P. Liu, J. Am. Chem. Soc. 130 (2008) 10996.
[39] C. Grebner, L.P. Pason, B. Engels, J. Comput. Chem. 34 (2013) 1810.
[40] P.M. Zimmerman, J. Comput. Chem. 34 (2013) 1385.
[41] S. Maeda, K. Ohno, K. Morokuma, Phys. Chem. Chem. Phys. 15 (2013) 3683.
[42] K.K. Irikura, R.D. Johnson, J. Phys. Chem. A 104 (1999) 2191.
[43] K.M. Westerberg, C.A. Floudas, J. Chem. Phys. 110 (1999) 9259.
[44] J.Y. Wu, J.Y. Liu, Z.S. Li, C.C. Sun, J. Chem. Phys. 118 (2003) 10986.
[45] S. Urata, A. Takada, T. Uchimaru, A.K. Chandra, Chem. Phys. Lett. 368 (2003) 215.
[46] D.A. Good, J.S. Francisco, J. Phys. Chem. A 104 (2000) 1171.
[47] A. Bottoni, P.D. Casa, G. Poggi, J. Mol. Struct. THEOCHEM 542 (2001) 123.
[48] F. Atadinc, C. Selcuki, L. Sari, V. Aviyente, Phys. Chem. Chem. Phys. 4 (2002) 1797.
[49] A.M. El-Nahas, T. Uchimaru, M. Sugie, K. Tokuhashi, A. Sekiya, J. Mol. Struct. THEOCHEM 722 (2005) 9.
[50] C. Zavala-Oseguera, J.R. Alvarez-Idaboy, G. Merino, A. Galano, J. Phys. Chem. A 113 (2009) 13913.
[51] S.A. Carr, T.J. Still, M.A. Blitz, A.J. Eskola, M.J. Pilling, P.W. Seakins, R.J. Shannon, B. Wang, S.H. Robertson, J. Phys. Chem. A 117 (2013) 11142.
[52] H. Sun, C.K. Law, J. Phys. Chem. A 114 (2010) 12088.
[53] A. Guenther, C.N. Hewitt, D. Erickson, R. Fall, C. Geron, T. Graedel, P. Harley, L. Klinger, M. Lerdau, W.A. McKay, T. Pierce, B. Scholes, R. Steinbrecher, R. Tallamraju, J. Taylor, P. Zimmerman, J. Geophys. Res. 100 (1995) 8873.
[54] I.G. Kavouras, N. Mihalopoulos, E.G. Stephanou, Geophys. Res. Lett. 26 (1999) 55.
[55] A.C.M. Glasius, N.R. Jensen, J. Hjorth, C.J. Nielsen, Int. J. Chem. Kinet. 29 (1997) 527.
[56] M. Hallquist, Environ. Sci. Technol. 31 (1997) 3166.
[57] P.J. Lewis, K.A. Bennett, J.N. Harvey, Phys. Chem. Chem. Phys. 7 (2005) 1643.
[58] D. Grosjean, E.L. Williams, J.H. Seinfeld, Environ. Sci. Technol. 26 (1992) 1526.
[59] B. Nozière, S. Spittler, L. Ruppert, I. Barnes, K.H. Becker, M. Pons, K. Wirtz, Int. J. Chem. Kinet. 31 (1999) 291.
[60] M.R. Dash, B. Rajakumar, J. Phys. Chem. A 116 (2012) 5856.
[61] J. Fan, J. Zhao, R. Zhang, Chem. Phys. Lett. 411 (2005) 1.
[62] V.M. Ramírez-Ramírez, J. Peiró-García, I. Nebot-Gil, Chem. Phys. Lett. 391 (2004) 152.
[63] D.L. Thomsen, S.A. Jørgensen, Chem. Phys. Lett. 481 (2009) 29.
[64] A. Wisthaler, N.R. Jensen, R. Winterhalter, W. Lindinger, J. Hjorth, Atmos. Environ. 35 (2001) 6181.
[65] T.J. Wallington, W.F. Schneider, D.R. Worsnop, O.J. Nielsen, J. Sehested, W.J. Debruyn, J.A. Shorter, Environ. Sci. Technol. 28 (1994) 320A.
[66] M. Trainer, E.J. Williams, D.D. Parrish, M.P. Buhr, E.J. Allwine, H.H. Westberg, F.C. Fehsenfeld, S.C. Liu, Nature 329 (1987) 705.
[67] A.D. Keil, P.B. Shepson, J. Geophys. Res. 111 (2006) D17303.

[68] J.D. Fuentes, M. Lerdau, R. Atkinson, D. Baldocchi, J.W. Bottenheim, P. Ciccioli, B. Lamb, C. Geron, L. Gu, A. Guenther, T.D. Sharkey, W. Stockwell, Bull. Am. Meteorol. Soc. 81 (2000) 1537.
[69] S.C. Liu, M. McFarland, D. Kley, O. Zafiriou, B. Huebert, J. Geophys. Res. 88 (1983) 1360.
[70] Q.S. Li, C.Y. Wang, J. Phys. Chem. A 106 (2002) 8883.
[71] M.C.L. Scaldaferri, A.S. Pimentel, Chem. Phys. Lett. 470 (2009) 203.
[72] K.T. Kuwata, L.C. Valin, A.D. Converse, J. Phys. Chem. A 109 (2005) 10710.
[73] O. Horie, P. Neeb, G.K. Moortgat, Int. J. Chem. Kinet. 29 (1997) 461.
[74] J.M. Anglada, J. Gonzalez, M.A. Torrent-Sucarrat, Phys. Chem. Chem. Phys. 13 (2011) 13034.
[75] L. Vereecken, H. Harder, A. Novelli, Phys. Chem. Chem. Phys. 14 (2012) 14682.
[76] B. Long, X.F. Tan, Z.W. Long, Y.B. Wang, D.S. Ren, W.J. Zhang, J. Phys. Chem. A 115 (2011) 6559.
[77] T.L. Nguyen, J. Peeters, L. Vereecken, Phys. Chem. Chem. Phys. 11 (2009) 5643.
[78] P. Aplincourt, J.M. Anglada, J. Phys. Chem. A 107 (2003) 5812.
[79] X.X. Lin, Y.R. Liu, T. Huang, K.M. Xu, Y. Zhang, S. Jiang, Y.B. Gai, W.J. Zhang, W. Huang, RSC Adv. 4 (2014) 28490.