Reaction Kinetics, Mechanisms and Catalysis (2022) 135:1439–1455
https://doi.org/10.1007/s11144-022-02170-5

Reaction streams in overall gas-phase chemical reactions

Homayoon Rafatijo[1]

Received: 20 October 2021 / Accepted: 18 January 2022 / Published online: 21 March 2022

Abstract

The phase space data produced in molecular dynamics (MD) simulations of many-atom systems contain complete classical descriptions of the processes determined by the force field for the specific initial conditions. The usefulness of the simulation results comes from gleaning physically meaningful information from the mass of phase space points. This is especially difficult in the case of chemical reactions in bulk phases. Determining the major pathways from reactants to products directly from MD simulation results and without assuming a given mechanism is particularly challenging. We propose the *reaction stream* method based on the cluster representation for gases to identify, directly from MD simulations, the sequences of isolated reactions that are responsible for progress of the overall chemistry in complex reactions. The goal is to obviate the need for pre-convinced mechanisms to characterize the overall chemistry. To illustrate the applicability of the *reaction stream* method, we identify the dominant reaction pathways for the dissociation of H_2O_2 in MD simulations with the ReaxFF force field. The significance of rare reaction pathways to predict the progress of the evolving chemistry is also studied.

Keywords Reaction Stream · Reaction Clusters · Combustion Kinetics · Reaction Network

Introduction

Many overall reactions involve many species that undergo various kinds of reactions at various times over the course of the reaction. Thus, it is tempting to consider them to be complex processes; however, given that the driving forces in overall reactions can be expected to induce an orderly progression from reactants to products it is perhaps not appropriate to consider them complex if we wish to ultimately understand them better. The objective should be to discover the nature of the order within the

✉ Homayoon Rafatijo
rafatijoh@missouri.edu

[1] Department of Chemistry, University of Missouri-Columbia, Columbia, MO 65211-7600, USA

complicated overall chemistry. We present a method for determining the details of how overall chemical reactions in bulk gases proceed.

We begin with the idea that complex overall gas-phase reactions proceed via a stream comprising various kinds of elementary reactions; and we assume that they follow a definite pathway determined by the chemistry given a specific set of physical conditions. One can think of the stream of reactions as a channel, defined by the evolving sequential, branching reactions of mostly highly reactive radicals leading from reactants to products, that the overall reaction will always follow given the same initial and physical conditions. This reaction stream is defined by the chemical forces, for they cause atoms to trace the same classical trajectory, commonly referred to as the minimum energy path, given a specific set of initial conditions.

Determining a precise reaction stream experimentally is not feasible, but molecular dynamics (MD) is a practical way to simulate the evolving chemistry of an overall reaction in bulk gases, and thus illustrate the concept. With MD simulations, one would need algorithms to reveal significant information about how the overall chemistry proceeds, ideally without human guidance or preconceived notions of what occurs in the process. However, an algorithm requires some prior knowledge and assumptions about chemical processes being studied in order to target the events that are meaningful within the context of the problem of interest. In this respect, we establish our reaction stream method upon one general assumption: distinct isolated chemical events can be identified in reaction streams.

The idea of reaction stream is that there is for each overall reaction a pathway, like a central current of a stream, leading from the reactants to the products. There are at least a couple of ways we can characterize and represent the reaction stream. It can be considered to be a network of reduced size, illustrated by arborescence graphs, for example, that includes only dominant sequences of reactions that are responsible for conversion of reactants to products. Alternatively, the reaction stream can be described as a lower dimensional sub-manifold in the reaction phase space to which all solution trajectories are attracted after a short transient, that is, a slow invariant manifold. These two interpretations have been used to reduce reaction mechanisms in master equation models; however, to the best of our knowledge, they have not been used to describe the governing chemistry at the atomic level. Using either interpretation for our results would, we believe, lead to essentially the same picture of the chemistry and conclusions. At this point it is not clear how best to describe or represent reactions streams, so this work is an initial step in learning more about how the overall processes proceed. We have chosen to use the first of these two approaches.

We begin by defining an *atomic reaction sequence*, based on the cluster representation for gases, as sequence of isolated reactions that are responsible for conversion of reactant species to product species [1]. Previously, we have shown how the cluster representation can be used to identify reaction clusters, that is, isolated elementary reactions in simulations of bulk gases [2, 3]. Also, we have shown how one can compute activation energies for reaction clusters using a generalized Tolman activation energy concept [4]. Here we use atomic reaction sequences to characterize overall reactions, thus obtaining a description of reaction streams. Green et al. [5, 6] have recently presented a related method based on the idea of mapping [7] in

dynamical systems to report physically meaningful sequences of species in an overall reaction. They used MD simulation results for hydrogen combustion to periodically determine to what species each hydrogen and oxygen atom belongs. In Green's model, the determination of species to which an atom belongs were made discreetly at regular time intervals; thus the sequence of reactions was not determined. We pick where Green and his coworkers left off; our goal here is to explore the elementary reactions in an overall reaction, and to describe the process as a definable stream of elementary reactions. We illustrate the applicability of the reaction stream method using the unimolecular dissociation reaction $10H_2O_2 \rightarrow 10H_2O + 5O_2$.

Computational methods

Microcanonical (NVE) MD simulations were carried out using the 2016 version of the LAMMPS [8, 9] code with the ReaxFF [10, 11] force field parametrized for hydrogen combustion to study the unimolecular dissociation reaction $10H_2O_2 \rightarrow 10H_2O + 5O_2$. Eighteen trajectories were run in a periodic cell of dimensions $12.7 \times 12.7 \times 12.7$ Å^3 for density $\rho = 275.6$ kg m^{-3}.

The initial configurations for equilibration of the NVT ensemble were generated using PACKMOL [12]. Ten H_2O_2 molecules were positioned in the periodic cell such that all non-bonding atom–atom distances were at least 2.2 Å. The molecular intranuclear distances 0.970 Å and 1.461 Å, respectively, were chosen for the O–H and O–O bonds in each H_2O_2 molecule. The initial configurations were then adjusted by using the Polak-Ribiere [13] conjugate-gradient algorithm for energy minimization until either the energy or the force criterion 1×10^{-4} or 1×10^{-6} kcal mol^{-1} Å^{-1}, respectively, was satisfied. Atomic velocities were randomly selected from the Gaussian distribution for 2100 K and assigned to atoms such that the total linear momentum of the system was zero. The trajectories were propagated using the velocity-Verlet [14] algorithm with time step 0.1 fs. Equilibration of the system was ascertained by observation that the kinetic temperatures consistently fluctuated about 2100 K. The translational kinetic temperature reached this value in less than 1 ps, but the internal modes required up to 30 ps in some simulations. If a reaction occurred the equilibration simulation was abandoned. The final states of the equilibration simulations were used as the initial states for the production NVE simulations, which were run for 2 ns with time step 0.01 fs. The atomic coordinates and velocities were recorded every 0.5 fs.

We used a method based on the concept of *reaction clusters* applicable to MD simulation of bulk gases to identify isolated reactions. The individual clusters are conservative systems within the bulk gas. The reaction clusters are defined such that sequences of chemical change occurring in isolation are completed, with subsequent reactions of any of the participants being considered in the formation of a new reaction cluster. We considered two kinds of reactions: those that are initiated by a bond breaking but which can subsequently involve bond formations and those that involve only bond formation (radical recombination).

The bond distances were checked every 0.5 fs. The time at which a bond length exceeded the cutoff distance was taken as the time t_r that a reaction had begun.

The bond rupture cutoff distance for each bond was defined as the distance at which the bond energy is within 1 kcal mol^{-1} of the bond dissociation energy predicted by the force field. At t_r all the species within interaction ranges of the breaking bond were identified using the non-bonding atom–atom cutoff distances 3.0 Å, 1.8 Å, and 3.5 Å, respectively, for H···H, H···O, and O···O; which are based on ReaxFF. This defines the reaction cluster and the reaction molecularity.

The internuclear distances of the species in a reaction cluster were monitored beginning 210 fs before t_r. For bond-rupture reactions that were followed by a bond formation, we considered a bond to be formed when it had undergone four inner turning points. When a non-bonding atom–atom distance between two species in a reaction cluster exceeded the critical cutoff value the reaction cluster was considered to have dissociated and the products were identified.

A production run was begun with the atomic Cartesian coordinates of the final state from an equilibration simulation. Atoms, molecules, bonds, and radicals were assigned numeric IDs; and the atoms making up each species were tagged to its ID. The IDs of the atoms making up each bond in a species were tagged to the species ID. The integration was continued with periodic checks of all covalent bond lengths every 0.5 fs until a bond distance exceeded a bond cutoff distance, then IDs of the species and atoms and the composition of the reaction cluster were recorded. The algorithm returned plots of the internuclear distances as functions of time for the species making up the reaction cluster beginning 210 fs before the bond rupture, which were used to determine the reaction mechanism.

When a non-bonding atom–atom distance between any two species in a reaction cluster exceeded a critical cutoff value the reaction cluster was considered to have dissociated and the constituents were identified. When an individual reaction completed the chemical composition of the entire system was updated with all atoms and species were assigned new IDs, which were used in the identifications of the new clusters and next reaction.

The number of species present in the system as the overall reaction progresses was monitored for each simulation. We defined a "reduced" time for each reaction as the time at which the reaction occurred divided by the overall course of time for the trajectory in which the reaction was detected. The overall trajectory time was measured from the beginning of the simulation up until the time the final products reach the final equilibrium state. The final equilibrium state was ascertained when the kinetic internal temperatures fluctuated about the target thermodynamic temperature. The reduced time was used so that the results for different trajectories can be combined to compute the ensemble average behaviors.

We used the version of the *Nudged Elastic Band* (NEB) [15] method implemented in LAMMPS to determine the energetics precited by ReaxFF for the reactions involving radicals. A minimum of 36 replicas distributed equidistantly along the minimum energy path of the reaction was used to perform the NEB calculations. For each reaction, the initial and final replicas were selected from an actual simulation. Values of the force constants were iteratively changed within the interval 1.0×10^{-5} E_h a_0^{-2} to 1.0×10^{-1} E_h a_0^{-2} until the force criterion of 1.0×10^{-5} E_h a_0^{-1} was met for the saddle points. The NEB calculations did not converge for reactions involving $(OH)_3$ and symmetric exchange reactions. The

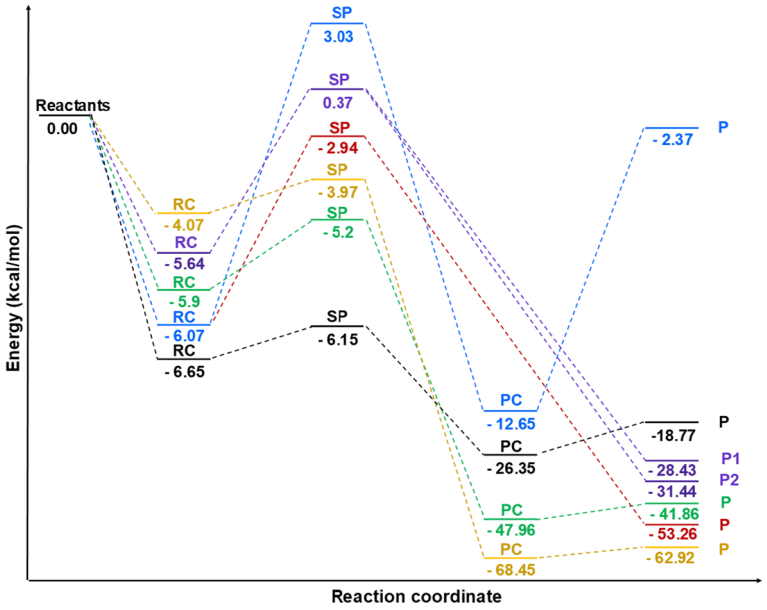

Fig. 1 ReaxFF predictions of the potential energy profiles of reactions (black) $H_2O_2+OH\rightarrow H_2O+HO_2$, (blue) $2OH\rightarrow H_2O+O$, (green) $2HO_2\rightarrow H_2O_2+O_2$, (purple) $HO_2+OH\rightarrow H_2O_3$, (yellow) $HO_2+OH\rightarrow H_2O+O_2$, and (red) $2OH\rightarrow H_2O_2$. RC, SP, PC, and P stand for reactant complex, saddle point, product complex, and product.

potential energy profiles for other reactions are shown in Fig. 1, which we show to define the model we have used, not to predict or represent the real chemistry. According to ReaxFF, reaction $HO_2+OH\rightarrow H_2O+O_2$ with classical energy barrier of 0.1 kcal mol^{-1} relative to the bottom of the RC well and $\Delta H(0\ \mathrm{K})$ of 62.92 kcal mol^{-1} is the most kinetically and thermodynamically favored reaction.

We note that ReaxFF predictions for the energetics of reactions involving radicals are inherently inaccurate. It is not parametrized to account for spin–spin interactions, which is indispensable to characterize the potential energy surface for such reactions. But nonetheless, for the purpose of this study we use the ReaxFF force field as a tool that, though not reliably accurate, provides the kind of complex chemistry we wish to characterize. For example, the barrier height predicted by ReaxFF for $2HO_2\rightarrow H_2O_2+O_2$ is 0.98 kcal mol^{-1} lower than that reported by Zhou et al. [16] who characterized the saddle point region of the triplet PES by using CASPT2/CBS/aug-cc-pVTZ level of the theory.

Results and discussion

While our purpose is to illustrate a method for characterizing reaction streams of complex reactions, we have chosen to use a relatively simple overall reaction. We have obtained the sequences of individual reactions in the reaction stream of the overall reaction $10H_2O_2\rightarrow 10H_2O+5O_2$ by using a reaction identification

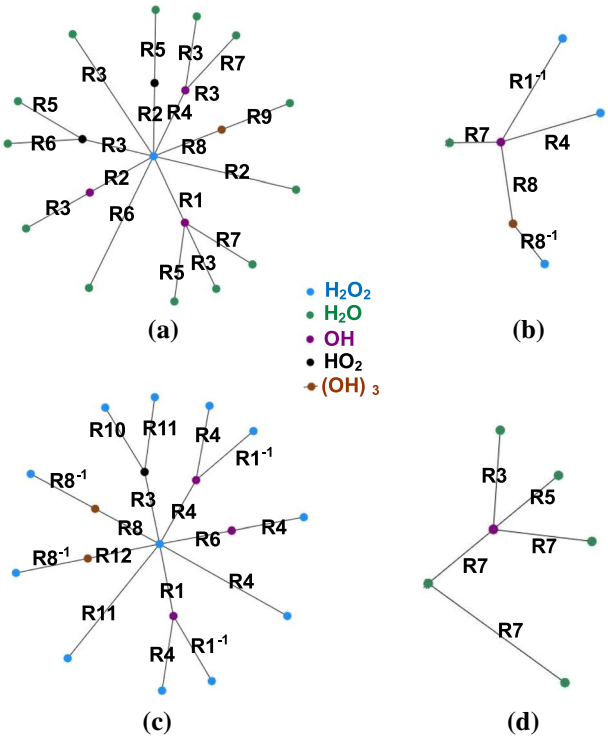

Fig. 2 Dominant atomic reaction sequences for hydrogen atoms in the overall reaction $10H_2O_2 \rightarrow 10H_2O + 5O_2$ that connect (**a**) a H_2O_2 molecule to a water molecule, (**b**) a water molecule to a H_2O_2 molecule, (**c**) a H_2O_2 molecule to another H_2O_2 molecule, and (**d**) a H_2O molecule to another H_2O molecule. Circular points represent different types of species to which a hydrogen atom belongs after a reaction. Lines represent reactions

algorithm. To determine the details of the chemistry in the overall reaction we need to determine the sequences of elementary reactions that leads from each reactant to each product. Thus, we defined four sets that contain all hydrogen atomic reaction sequences; specifically, the four sets contain atomic sequences that transmit hydrogen atoms from H_2O_2 to H_2O, from H_2O to H_2O_2, from H_2O_2 to another H_2O_2 ($HO'O'H' \rightarrow \cdots \rightarrow HOOH$), and from H_2O to another H_2O ($HO'H' \rightarrow \cdots \rightarrow HOH$). This was done by monitoring all hydrogen atoms in the system. We used arborescence graphs to depict these sets.

In the arborescence graphs of Fig. 2, the vertices represent different species distinguished by color, and a line between two vertices represents a reaction transmitting hydrogen atoms between two vertices. Depending on the set, the root vertex is either H_2O_2 or H_2O. Atomic reaction sequences for each set are ranked based on their frequency, which are given in Table 1. The frequency of an atomic reaction sequence in each set is defined as the ratio of the number of times that sequence occurred in the set and the total number of sequences in that

Table 1 The frequencies of most frequent atomic reaction sequences (ARS) for each set illustrated in Fig. 2

$H_2O_2 \rightarrow \cdots \rightarrow H_2O$ (63%)[a]		$H_2O \rightarrow \cdots \rightarrow H_2O_2$ (3%)[a]		HO′O′H′ → (23%)[a]		HO′H′ → ⋯ → HOH (11%)[a]	
ARS	Frequency[b] (%)	ARS	Frequency[b] (%)	ARS	Frequency[b] (%)	ARS	Frequency[b] (%)
R3	24	R7–R1^{-1}	59	R4	30	R7	56
R1–R7	14	R7–R4	26	R3–R10	13	R7-R7	27
R1–R3	13	R7–R8–R8^{-}	12	R1–R4	12	R7-R5	8
R6	6			R8–R8^{-1}	7	R7-R3	5
R3–R5	5			R11	7		
R4–R3	4			R6–R4	7		
R2	3			R4–R4	5		
R3–R6	3			R12-R8^{-1}	3		
R1–R5	3			R3-R11	2		
R2–R3	3			R1-R1^{-1}	2		
R8–R9	2			R4-R1^{-1}	2		
R4–R7	2						
R2–R5	1						

[a]This percentage represents the contribution of the set to the reaction stream, that is, ARSs in this set constitute this percentage of all hydrogen atomic reaction sequences

[b]The frequency for each ARS in the set was calculated as the ratio of number of times the sequence was detected and the total number of ARSs in that set multiplied by 100

set multiplied by 100. In drawing Fig. 2 and making Table 1, we excluded the atomic reaction sequences with frequency smaller than 1%. The reactions in the sequences listed in Table 1 were identified as the most frequent reactions.

We have determined the contribution of each set to the reaction stream. The two sets beginning with the reactant comprise large numbers of reaction sequences (~60), while those beginning with the product comprise relatively few (<8), as expected. Appendix I contains a complete of list of reactions identified in the simulations. We did not determine if these reactions occurred early, middle, or late in the overall reaction.

ReaxFF-based NVE simulations of ten H_2O_2 molecules at 2100 K predict two initial reactions[1]:

$$H_2O_2\left(+H_2O_2\right) \rightarrow 2OH \quad (12\%) \tag{R1}$$

and

$$2H_2O_2 \rightarrow H_2O + HO_2 + OH, \quad (5\%) \tag{R2}$$

with contributions of 67% and 33% to the initial radical production, respectively. The most frequent subsequent reactions are

$$2OH \rightarrow H_2O_2 \quad (4\%) \tag{R1^{-1}}$$

$$H_2O_2 + OH \rightarrow H_2O + HO_2 \quad (18\%) \tag{R3}$$

$$OH + HOO'H' \rightarrow O'H' + HOOH \quad (9\%) \tag{R4}$$

$$HO_2 + OH \rightarrow H_2O + O_2 \quad (5\%) \tag{R5}$$

$$H_2O_2 + HO_2 \rightarrow H_2O + O_2 + OH \quad (2\%) \tag{R6}$$

$$OH + HO'H' \rightarrow O'H' + HOH \quad (15\%) \tag{R7}$$

$$H_2O_2 + OH \rightarrow (OH)_3 \quad (3\%) \tag{R8}$$

$$(OH)_3 \rightarrow H_2O + HO_2 \quad (2\%) \tag{R9}$$

$$2HO_2 \rightarrow H_2O_2 + O_2 \quad (5\%) \tag{R10}$$

$$H'O'O'H + HO_2 \rightarrow H'O'O' + H_2O_2 \quad (1\%) \tag{R11}$$

[1] The percentage for each reaction was calculated as the ratio of the number of times the reaction was detected in all simulations and the total number of reactions identified multiplied by 100.

$$2H_2O_2 \rightarrow (OH)_3 + OH \quad (2\%) \tag{R12}$$

Fig. 2a illustrates the arborescence graphs for the atomic reaction sequences transmitting hydrogen atoms from H_2O_2 molecules to water molecules that constitute 1% or more of all the sequences in the $H_2O_2 \rightarrow \cdots \rightarrow H_2O$ set. The transmission of hydrogen atoms from reactants to products predominatly occurs via sequences of one or two reactions.The frequencies of occurrence of reaction sequences given in Table 1; these sequences cumulatively constitute 83% of the $H_2O_2 \rightarrow \cdots \rightarrow H_2O$ set. The most frequent reaction, which constitutes 24% of the $H_2O_2 \rightarrow \cdots \rightarrow H_2O$ set, is reaction R3.

"Rare" atomic reaction sequences that individually constitute less than 1% of all reaction sequences in the $H_2O_2 \rightarrow \cdots \rightarrow H_2O$ set cumulatively constitute 18% of the overall reaction stream. Table 2 provides a complete list of those sequences. We have excluded rare sequences in drawing the arborescence graph of Fig. 2a; however, including them would be necessary for a full description of the reaction stream. A more extensive sampling of the overall reaction could reveal more rare sequences not listed in Table 2.

The simulation results show that accurate characterization of transient behaviors, such as, ignition delay time, in an overall reaction is contingent on identifying both the initial reaction and rare sequences. The kinetic temperature of an isolated system steadily increases before a sudden jump, which signals an ignition event [17]. Fig. 3 shows the evolution of kinetic temperature from the beginning of the simulation until the time a final equilibrium state was reached for four trajectories. For three of the trajectories (shown in red, purple, and green) reaction R1 was the initial reaction, while for the trajectory shown in blue reaction R2 produced the first radical.

Progress of chemistry in the disequilibrium region of the thermodynamics phase space depends on reactions of radicals, and as such different initial reactions significantly alter the overall chemistry, as expected. Fig. 3 shows that trajectories in which R1 is the initial reaction reached the final equilibrium more rapidly than when R3 is the initial reaction. Differently put, trajectories in which the chemistry begins with reaction R1 have shorter ignition time than that begins with reaction R3. More important though, Fig. 3 shows that trajectories in which chemistry began by reaction R1—but differ in their atomic reaction sequences—qualitatively exhibit similar ignition behavior. Comparing sequences in the $H_2O_2 \rightarrow \cdots \rightarrow H_2O$ set for these trajectories shows that about fifty percent of atomic reaction sequences are the same; see Table 3. The set $H_2O_2 \rightarrow \cdots \rightarrow H_2O$ in each trajectory contains unique rare sequences that cumulatively amount to between 20 and 30% of the set. Therefore, excluding rare atomic reaction sequences renders parametrized kinetic models attempting to predict complex chemical processes inaccurate, although such models may be able to describe the chemistry "qualitatively."

Fig. 2a shows that there is in fact a pathway leading from reactants to products, which is controlled by dynamics within non-equilibrium thermodynamic conditions. However, the dynamics is predominantly chaotic. This is because the time required for completion of an atomic reaction sequence (that is, the time required for a hydrogen atom to transmit from a H_2O_2 molecule to a water molecule in the arborescence graph in Fig. 2a) varies from one hydrogen atom to another. For example, consider

sequences R1-R7 and R1-R3 that constitutes 14% and 13% of the $H_2O_2 \rightarrow \cdots \rightarrow H_2O$ set, respectively (see Table 1). Fig. 4 shows the distributions of time gaps between the two reactions making up these sequences. The average time gap between reactions R1 and R7 is 3.1 ps smaller than that between R1 and R3, indicating that completion of the sequence R1-R7 is more sensitive to the lifetime of the OH radicals produced by reaction R1. Furthermore, according to Fig. 4, in presence of equal number of H_2O_2 and H_2O molecules, a nascent OH radical produced by reaction R1 is twice more likely to participate in reaction R7 within 3 ps of its production than does it in reaction R3. Note that since the classical barrier height for reaction R3 is 0.5 kcal mol^{-1} relative to the bottom of the RC well (see Fig. 1), almost every bimolecular collision between OH radicals and H_2O_2 molecules has enough kinetic energy (~R*T*) under the thermodynamic condition of the system to surmount the barrier. Therefore, the difference in the barrier heights of reactions R3 and R7 is not the reason for the difference in the time gaps between their occuances in sequences R1-R7 and R1-R3.

Reverse sequences are responsible for transmitting hydrogen atoms in water molecules back to a H_2O_2 molecules; these are illustrated in Fig. 2b. There are three major reverse sequences for hydrogen atoms. The frequencies corresponding to the pathways illustrated in Fig. 2b are given in Table 1. Less than 4% of all hydrogen atomic sequences are those in which a hydrogen atom goes from a water molecule to a H_2O_2 molecule. This is to be expected because the potential energy profile for reactions involving reacting water molecules have significantly higher barrier heights than that involving reacting H_2O_2 molecules. For example, the barrier height to reaction R3 relative to the bottom of the well is 0.5 kcal mol^{-1} while that for the backward reaction is 26.2 kcal mol^{-1}; see Fig. 1.

Symmetric exchange reactions such as $OH + HOO'H' \rightarrow O'H' + HOOH$ (R4), $OH + HO'H' \rightarrow O'H' + HOH$ (R7), and $H'O'O'H + HO_2 \rightarrow H'O'O' + H_2O_2$ (R11) occur frequently (~25% of all reactions identified) as the overall reaction progresses. Symmetric exchange reactions result from collisions between hot molecules (H_2O_2 or H_2O) with kinetic energy higher than the average kinetic energy of the system and cold radicals (OH or HO_2) with kinetic energies lower than the average kinetic energy of the system. The outcomes of these reactions are energetically stabilized molecules (H_2O_2 or H_2O) without changing the total kinetic energy of the system. Since the total kinetic energy and the composition of system remains unaffected by symmetric exchange reactions, these reactions indicate the tendency of a chemical system to preserve its global thermodynamic state. In other words, by replacing hot H_2O_2 molecules with energetically stabilized H_2O_2 molecules, symmetric exchange reactions hamper the progress of the system toward completion. We note that the kinetic energies of product radicals of these reactions (OH or HO_2) are higher than that of their reactant radicals. However, since nascent radicals with energies lower than the average energy of the system can promote chemical reactions, the increase in the energy of these radicals has less significant effect on the progress of overall reaction than the decrease in the energy of H_2O_2 molecules produced by symmetric exchange reactions, such as, R4.

Table 2 Rare atomic reaction sequences for each set $H_2O_2 \rightarrow \cdots \rightarrow H_2O$

$H_2O_2 \rightarrow \cdots \rightarrow H_2O$	
ARS[a]	ARS
R22	R11–R6
R23	R12–R9
R34	R1–R8–R9
R12–R3	R3–R25^{-1}–R29
R3–R25	R30–R14^{-1}–R7
R4–R44	R4–R15–R44
R8–R26	R20–R20^{-1}–R5
R4–R26	R20–R48–R5
R4–R5	R12–R8^{-1}–R7
R30–R7	R12–R8–R8^{-1}–R7
R6–R7	R1–R14–R37–R3
R23–R7	R4–R37–R14^{-1}–R7
R3–R31	R1–R3–R20–R44^{-1}
R1–R28^{-1}	R20–R14–R14^{-1}–R3
R36–R28^{-1}	R3–R15–R15^{-1}–R7
R33–R32	R4–R15–R15^{-1}–R5
R23–R3	R3–R25–R17–R15^{-1}–R7
R12–R7	R2–R36^{-1}–R28^{-1}–R25^{-1}–R5
R3–R35	R8–R26–R15^{-1}–R25–R23
R20–R44	
R1–R45	

[a]ARS stands for atomic reaction sequences. All reactions are listed in Appendix I

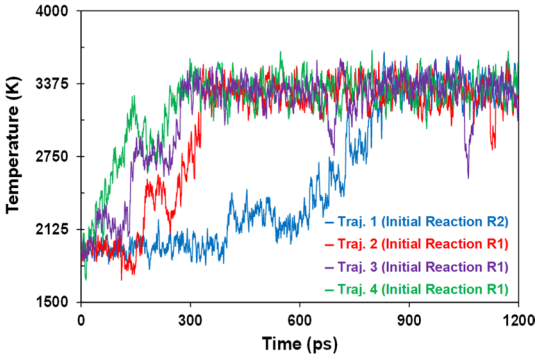

Fig. 3 The evolution of kinetic temperature from the beginning of simulation up until the time final equilibrium state was reached for four trajectories in the reaction overall reaction $10H_2O_2 \rightarrow 10H_2O + 5O_2$. The progress of overall chemistry toward final equilibrium state largely depends on the initial reaction. Two different initial reactions were identified: reactions R2 (blue) and R1 (red, purple, green). Atomic reaction sequences for set $H_2O_2 \rightarrow \cdots \rightarrow H_2O$ for red, purple, and green trajectories are given in Table 3, showing that the difference between trajectories regarding rare sequences does not qualitatively affect the progress of chemistry

Table 3 Atomic reaction sequences (ARS) in set $H_2O_2 \rightarrow \cdots \rightarrow H_2O$ for the red, green, and purple trajectories in Fig. 3

Red		Green		Purple	
ARS[c]	Frequency[b]	ARS	Frequency[b]	ARS	Frequency[b]
R3	35%	R3	21%	R3	24%
R1-R3	20%	R1-R7	25%	R1-R3	19%
R1-R7	10%	R4-R3	11%	R1-R7	9.5%
R6	10%	R1-R3	7%	R3-R5	9.5%
R3-R6	5%	R6	7%	R1-R5	4.7%
R3-R35	5%	R11-R6	4%	R34	9.5%
R12-R3	5%	R12-R9	4%	R20-R44	9.5%
R20-R48-R5	5%	R30-R7	4%	R1-R45	9.5%
R12-R8^{-1}-R7	5%	R4-R5	4%	R20-R44-R14^{-1}-R3	4.7%
		R3-R25^{-1}-R29	4%		
		R30-R14^{-1}R7	4%		
		R4-R15-R15^{-1}-R5	4%		
		R3-R15-R15^{-1}-R7	4%		

[a] Rare sequences are shown in color

We also identified atomic reaction sequences that do not change the composition of the system. Fig. 2c illustrates atomic reaction sequences that transmit a hydrogen atom in a H_2O_2 molecule to another H_2O_2 molecule. Similarly, Fig. 2d illustrates those that transmit a hydrogen atom in a H_2O molecule to another H_2O molecule. Comparing Fig. 2c and d, it follows that atomic reaction sequences in set $HO'O'H' \rightarrow \cdots \rightarrow$ HOOH have more variety than that of set $HO'H' \rightarrow \cdots \rightarrow$ HOH. The frequencies corresponding to the pathways illustrated in Fig. 2c and d are given in Table 1. These two sets of atomic reaction sequences constitute 34% of the reaction stream, indicating their importance in the evolution of the overall reaction.

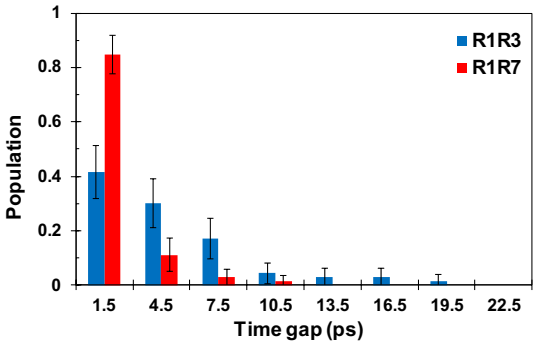

Fig. 4 The distributions of time gaps for two sequences of length two. Both sequences begin with the dissociation of a H_2O_2 molecule (R1) to produce OH. The results shown in blue are for $H_2O_2 + OH \rightarrow H_2O + HO_2$ (R3) and those in red for $OH + HO'H' \rightarrow O'H' + HOH$ (R7). These two sequences were observed equally frequently in the system; see Table 1. The average time gap between the two sequential reactions in R1-R7 sequence is 3.1 ps lower than that in R1-R3 sequence

The concept of reaction clusters in MD simulations can be further used to compute the rate of progress of the overall reaction. The objective here is to compute the rate of reaction directly from MD simulation without consideration of a particular reaction mechanism. We introduce the *reactants consumption parameter* (*RCP*) as a measure of the deviation of composition from the initial composition of the system as a function of time. For example, consider a unimolecular reaction $A \rightarrow$ Products with a the initial number of A, then the *RCP* as a function of time is

$$RCP(t) = \frac{A(t)}{a}, \tag{1}$$

where $A(t)$ is the number of species A at time t. At each instant of time, *RCP* has a value between zero and unity, where zero indicates that no reactant species is present in the system and unity indicates that only reactant species are present. A *reaction progress index* (*RPI*) at any instant of time t can be defined as unity minus the $RCP(t)$; that is,

$$RPI(t) = 1 - RCP(t) \tag{2}$$

Evolution of the *RPI* describes the overall rate of progress of the system toward its final equilibrium state.

For each trajectory, we plotted the curves of the total number of species and *RPI* during the course of the overall reaction against the reduced time. Frames (a) and (b) of Fig. 5 show, respectively, the evolutions of the average total number of species and the average *RPI* for 18 ensembles of $10H_2O_2 \rightarrow 10H_2O + 5O_2$ reactions. The data in frame (b) that shows *RPI*, was fitted (red curve) with a sigmoidal function, which is,

$$RPI(t) = d - \frac{c}{(1 + \exp(a \times t - b))}, \tag{3}$$

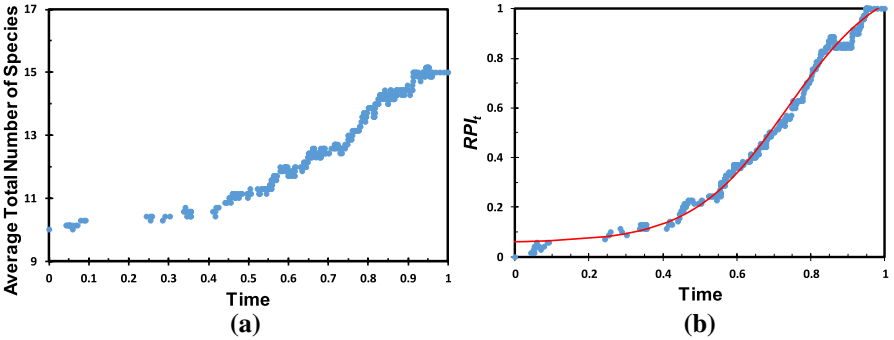

Fig. 5 The evolution of (**a**) the average total number of species and (**b**) the *reaction progress index*, *RPI*, both were averaged over 18 ensembles of $10H_2O_2 \rightarrow 10H_2O + 5O_2$ reactions. The horizontal axis is the reduced time obtained for each reaction by dividing the time at which the reaction occurred divided by the overall course of time for the trajectory to reach its final equilibrium. The *RPI* data was fitted (red curve) with the sigmoid expression, Eq. 3. Parameters of the fit are $d = 1.147$, $c = 1.090$, $a = 7.681$, and $b = 5.666$. The R^2 value for the fit is 0.99

Here fit parameters a, b, c, and d are given in the caption of Fig. 5. While we did not carry out MD simulations under different initial thermodynamic conditions, it is reasonable to expect that these parameters depend on initial conditions. Note that parameter a determines the rate of progress of the overall chemistry toward the final equilibrium, which is computed directly from MD simulations.

Conclusions

We used reaction cluster construct to develop the reaction stream method for simulations of overall complex reactions. We showed that the reaction stream method reveals the atomic-level details of and the dominant pathways for complex chemistry starting with reactants, passing through many intermediates, to the final products for various conditions (pressure, temperature, relative concentrations). The reaction stream method obviates the need for assumed mechanisms to predict the evolution of the chemistry in complex reactive systems. We illustrated the applicability of the method by identifying all sequences of isolated reactions transmitting hydrogen atoms from H_2O_2 molecules to water molecules for reaction $10H_2O_2 \rightarrow 10H_2O + 5O_2$ for a given thermodynamic condition.

The reaction stream method reveals information about the frequencies of reactions occurring in different stages (equilibrium, disequilibrium, and nonequilibrium) of a reactive system. As such, it guides experimentalists in the development of more efficient use of existing fuels by promoting sequences of reactions responsible for conversion of reactants into products. Of course, results given by the reaction stream method can only be considered to be authentic for a model system with the potential energy profile described by the reactive force field used.

Finally, with this study, we communicated that the evolution of the overall chemistry in a unimolecular dissociation is described by a sigmoidal function. We noted that sigmoid function is a natural solution and characteristic of any one-dimensional dynamical system, such as combustion.

Appendix I NVE reaction stream results for thermal dissociation of H_2O_2

We performed 20 simulations of 10 H_2O_2 molecules in a box of $12.7 \times 12.7 \times 12.7$ $Å^3$ with periodic boundary conditions for NVE ensemble using the ReaxFF force field. The reaction identification algorithm identified the folowing reactions.

R1	$H_2O_2 \rightarrow 2OH$
R2	$2H_2O_2 \rightarrow H_2O + HO_2 + OH$
R3	$H_2O_2 + OH \rightarrow H_2O + HO_2$
R4	$H_2O_2 + OH \rightarrow H_2O_2 + OH$
R5	$HO_2 + OH \rightarrow H_2O + O_2$
R6	$H_2O_2 + HO_2 \rightarrow H_2O + O_2 + OH$
R7	$H_2O + OH \rightarrow H_2O + OH$
R8	$OH + H_2O_2 \rightarrow (OH)_3$
R9	$(OH)_3 \rightarrow H_2O + HO_2$
R10	$2HO_2 \rightarrow H_2O + O_2$
R11	$HO_2 + H_2O_2 \rightarrow H_2O_2 + HO_2$
R12	$2H_2O_2 \rightarrow (OH)_3 + OH$
R13	$H_2O_2 + HO_2 \rightarrow H_3O_4$
R14	$O_2 + OH \rightarrow HO_3$
R15	$HO_2 + OH \rightarrow H_2O_3$
R16	$H + H_2O \rightarrow H_3O$
R17	$H + HO_3 \rightarrow H_2O_3``$
R18	$HO_2 + O_2 \rightarrow HO_2 + O_2$
R19	$HO_2 + OH \rightarrow HO_2 + OH$
R20	$H_2O_2 + HO_2 \rightarrow H_2O_3 + OH$
R21	$H + OH + H_2O \rightarrow 2H_2O$
R22	$H + 2H_2O_2 \rightarrow H_2O + (OH)_3$
R23	$H + H_2O_2 \rightarrow H_2O + OH$
R24	$H_2O_3 \rightarrow HO_2 + OH$
R25	$HO_2 \rightarrow H + O_2$
R26	$(OH)_3 + OH \rightarrow H_2O_3 + H_2O$
R27	$H_2 + OH \rightarrow H_2O + H$
R28	$H_2O + H \rightarrow H_2 + OH$
R29	$H + H_2O \rightarrow H + H_2O$
R30	$O_2 + H_2O_2 \rightarrow HO_3 + OH$

R31	$H_2O + HO_2 \rightarrow H + O_2 + H_2O$
R32	$H_3O_2 \rightarrow H_2O + OH$
R33	$2\ H_2O_2 \rightarrow H_3O_2 + HO_2$
R34	$O + H_2O_2 \rightarrow O_2 + H_2O$
R35	$HO_3 + HO_2 \rightarrow H_2O + 2O_2$
R36	$H_2O_2 + HO_2 \rightarrow H_2 + O_2 + HO_2$
R37	$HO_3 + OH \rightarrow HO_3 + OH$
R38	$HO_3 + H_2 \rightarrow (OH)_3$
R39	$H_2 + OH \rightarrow H_3O$
R40	$2HO_2 \rightarrow 2OH + O_2$
R41	$(OH)_3 + H_2O_2 \rightarrow H_2O + HO_2 + H_2O_2$
R42	$H + HO_2 \rightarrow 2OH$
R43	$O_3 \rightarrow O_2 + O$
R44	$H_2O_3 + OH \rightarrow H_2O + HO_3$
R45	$2OH \rightarrow H_2O + O$
R46	$H_2O + O \rightarrow (OH)_2$
R47	$(OH)_2 + O_2 \rightarrow O_3 + H_2O$
R48	$H_2O_3 + HO_2 \rightarrow HO_2 + H_2O_3$

Acknowledgements This work was supported by the U.S. Army Research Laboratory and the Army Research Office under grant number ARO-MURI W911NF-14-1-0359. The author expresses his deepest gratitude to Professor Donald L. Thompson for his unwavering intellectual and financial supports throughout this project. Professor Thompson's brilliant ideas sparked off a long-term undertaking that leads to the development of the reaction stream method. As such, the author dedicates this publication to Professor Donald L. Thompson.

Data availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Andersen HC (1977) Cluster methods in equilibrium statistical mechanics of fluids. Statistical mechanics: part a: equilibrium techniques. Plenum, New York
2. Rafatijo H, Thompson DL (2017) J Chem Phys 147:224111
3. Rafatijo H, Monge-Palacios M, Thompson DL (2019) J Phys Chem A 123:1131
4. Rafatijo H (2020). Mol. Phys. 118(17). https://doi.org/10.1080/00268976.2020.1749952
5. Alaghemandi M, Green JR (2016) Phys Chem Chem Phys 18:2810
6. Nicholson SB, Alaghemandi M, Green JR (2018) J Chem Phys 148:44102
7. Scott SK (1991) Chemical chaos. Oxford University Press Inc, New York
8. Aktulga HM, Fogarty JC, Pandit SA, Grama AY (2012) Parallel Comput 38:245
9. Plimpton S (1995) J Comp Phys 117:1
10. Senftle TP, Hong S, Islam MM, Kylasa SB, Zheng Y, Shin YK, Junkermeier C, Engel-Herbert R, Janik MJ, Aktulga HM, Versraelen T, Grama A, van Duin ACT (2016) NPJ Comput Mater 2:15011
11. Nielson KD, van Duin ACT, Oxgaard J, Deng WQ, Goddard WA III (2005) J Phys Chem A 109:493
12. Martinez L, Andrade R, Birgin EG, Martinez JM (2009) J Comput Chem 30:2157
13. Polak E, Ribiere G (1969) Rev Fran Informat Rech Oper 3:35–43
14. Verlet L (1967) Phys Rev 159:98–103

15. Nakano A (2008) Comp Phys Comm 178:280
16. Zhou DDY, Han K, Zhang P, Harding LB, Davis MJ, Skodje RT (2012) J Phys Chem A 116:2089
17. Alaghemandi M, Newcomb LB, Green JR (2017) J Phys Chem A 121:1686