JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

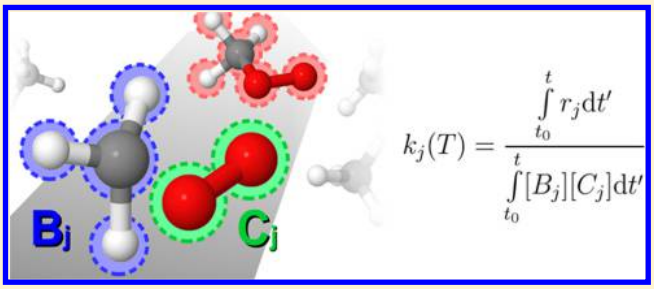
Automated Discovery of Reaction Pathways, Rate Constants, and Transition States Using Reactive Molecular Dynamics Simulations

Malte Döntgen,[†,§] Marie-Dominique Przybylski-Freund,[†] Leif C. Kröger,[†] Wassja A. Kopp,[†] Ahmed E. Ismail,[‡,§] and Kai Leonhard*[,†,§]

[†]Chair of Technical Thermodynamics, [‡]Aachener Verfahrenstechnik, Faculty of Mechanical Engineering, [§]AICES Graduate School, RWTH Aachen University, 52062 Aachen, Germany

Supporting Information

ABSTRACT: We provide a methodology for deducing quantitative reaction models from reactive molecular dynamics simulations by identifying, quantifying, and evaluating elementary reactions of classical trajectories. Simulations of the inception stage of methane oxidation are used to demonstrate our methodology. The agreement of pathways and rates with available literature data reveals the potential of reactive molecular dynamics studies for developing quantitative reaction models.



1. INTRODUCTION

Quantitative models of complex chemical reaction networks are widely used in chemistry for analysis, interpolation, and prediction. Mechanisms have always been essential for the analysis of complex chemical networks in chemical engineering,[1] polymerization,[2] and combustion,[3] e.g., to predict product yields[4] or pollutant formation.[5] Although mechanisms were originally used to elucidate fundamental chemical processes, such as enzyme−substrate interactions[6] or the effects of pressure,[7] nowadays, mechanisms can model networks of hundreds of species and thousands of reactions. However, accurate modeling is both experimentally and computationally challenging. Recently, reactive molecular dynamics (MD) has emerged as a key tool to understand these complex networks.[8] We present a methodology to derive both a network of reactions and its kinetic parameters from reactive MD simulations.

Challenges in reaction model development comprise experimental and computational issues. Rate constants can be estimated from Evans−Polanyi relationships,[9] Benson rules,[10] or group additivity[11] only for compounds chemically similar to well-studied compounds. On the other hand, experiments often cannot be applied to study reaction rate constants if the reaction cannot be isolated. Training the model to experiments measuring global quantities of the reaction network requires the assumption of a reaction scheme if one wants to determine the rate constants of the reactions. Alternatively, if one wants to build a mechanism using an automatic generator,[12] knowledge about rate constants of the reactions is required to evaluate their importance. Incompleteness of the assumed reaction scheme and faulty rate constants may lead to amplified errors.[13]

Rappoport et al. used a heuristic approach for mechanistic studies that does not require knowledge of the reaction kinetics.[14] Transformation rules were developed to create new chemical structures based on an initial structure. This allows new reaction pathways to be discovered independent of their dynamics as well as reactions that would be observed rarely during reactive MD simulations.

Zheng and Pfaendtner proposed the use of chemical trajectory simulations to gain insight into the elementary reactions of a mechanism.[15] The authors used Car−Parrinello MD (CPMD) accelerated with metadynamics to investigate rare reaction events of methanol oxidation chemistry. Changes in atomic connectivities were computed from the full connectivity matrix of a single methanol molecule interacting with a single oxygen molecule.

A very recent approach to discover reaction paths and species without prior knowledge is the *ab initio* nanoreactor,[8] which uses energies and forces calculated by Born−Oppenheimer MD (BOMD) to simulate the fate of specified reactants. To observe reaction events in systems that may require seconds in reality within the BOMD-acccessible timespan of hundreds of picoseconds, Langevin dynamics are performed at 2000 K in a spherical box whose volume fluctuates to increase reaction frequencies. Because of this trick, however, the dynamics of the system cannot be quantitatively related to the real dynamics. The computational resources needed are also substantial, on the order of 10^5 CPU and GPU hours for the pyrolysis of 39 acetylene molecules or a 228-atom simulation of the Urey−Miller experiment. Software tools to identify species and reactions and to remove high-frequency oscillations detected as spurious reactions have been developed using graph-theoretical and machine-learning techniques. The trajectories have been used to obtain so-called minimum energy paths (MEPs) by a string method. The MEP connects reactants, transition states,

Received: March 2, 2015
Published: April 29, 2015

DOI: 10.1021/acs.jctc.5b00201
J. Chem. Theory Comput. 2015, 11, 2517−2524

and products and therefore provides insight into the chemistry taking place. The MEPs have been determined at the B3LYP/6-31+G(d,p) level to provide barriers that are more accurate than those obtained during the simulation from Hartree−Fock (HF) with the 3-21G basis set.

Empirical reactive force fields provide an alternative to quantum mechanics (QM). They are at least 4 orders of magnitude more efficient computationally and are potentially more accurate than HF since they can be trained by more accurate DFT methods. Therefore, simulation times on the order of nanoseconds can be achieved without acceleration techniques. This allows reaction rates to be deduced directly from the trajectories because we keep the original dynamics of the system.

The potential of reactive force fields to study kinetics beyond the short simulation times of BOMD has been recognized in several studies. MD simulations using the ReaxFF[16] reactive force field have gained increasing attention in material technology[17] and biophysics[18] and have been used in combustion technology, fuel pyrolysis,[19−23] and oxidation[24,25] to extract reaction pathways. Elementary reactions extracted through manual observation have been used to compile parts of the underlying mechanisms of reactions.[24−26] Reaction mechanisms, reaction and activation energies, and rate constants have all been obtained from ReaxFF simulations of combustion and pyrolysis.[20,25,27,28] Although ReaxFF shows significant deviations for some reaction properties,[16] the simulation results were found to be in good agreement with available experimental data in most cases.[19−22,24] As can be seen from these studies, reactive MD is a promising tool to perform *in silico* experiments.[29]

In the present study, we show how ReaxFF simulations can be used to determine both the reaction mechanism and its rate constants. This allows us to determine reaction rates for a large mechanism from a single reactive MD trajectory rather than focusing on a particular reaction. A graph-theoretical analysis of the connectivities and their changes is used to automatically detect reactions. Furthermore, we assess the accuracy of the obtained rates and validate the molecular structures of transition states extracted directly from trajectories. From the trajectory in the vicinity of the transition state, we obtain starting points for a transition-state geometry optimization using the nudged-elastic band (NEB) approach.[30−32] From this, additional *ab initio* calculations can be performed to further increase the accuracy of the rate predictions.

2. THEORY

In previous reactive MD studies, reaction rate determinations either required individual computations[8] or were performed only for the first reaction of a simulation.[28] We derive here a formulation that can determine the rate constants of all reactions in a single simulation (see Supporting Information, Section S1 for validation for first-order reactions).

The production of a species D during simulation can be expressed as the integral of the corresponding reaction rate

$$[D]_{\mathrm{prod},i/j} = \int_{t_0}^{t} r_{i/j}\,\mathrm{d}t' \tag{1}$$

with the reaction rate $r_{i/j}$ being the average frequency of uni- and bimolecular reaction events, respectively. Here, $r_{i/j}$ is not the net rate of reaction but the rate of forward reactions leading to the production of D.

The uni- and bimolecular pathways that contribute to production of D give the following general macroscopic concentration balance

$$\frac{\mathrm{d}[D]_{\mathrm{prod}}}{\mathrm{d}t} = \sum_i k_i(T)[A_i] + \sum_j k_j(T)[B_j][C_j] \tag{2}$$

where A_i is a unimolecular source species reacting to D with rate constant $k_i(T)$ and B_j and C_j are bimolecular source species reacting to D with rate constant $k_j(T)$. All species' concentrations are time-dependent. The time-dependent concentration profile of D is obtained by integrating eq 2 over time, starting at t_0

$$[D]_{\mathrm{prod}} = \sum_i k_i(T)\int_{t_0}^{t}[A_i]\,\mathrm{d}t' + \sum_j k_j(T)\int_{t_0}^{t}[B_j][C_j]\,\mathrm{d}t' \tag{3}$$

The integrated concentration balance for D_{prod} includes information on the production of D due to single reaction channels. The uni- and bimolecular rate constants are computed by splitting the total production D_{prod} into the production from single channels $D_{\mathrm{prod},i/j}$

$$[D]_{\mathrm{prod},i} = k_i(T)\int_{t_0}^{t}[A_i]\,\mathrm{d}t',\qquad [D]_{\mathrm{prod},j} = k_j(T)\int_{t_0}^{t}[B_j][C_j]\,\mathrm{d}t' \tag{4}$$

$D_{\mathrm{prod},i/j}$ is not the total concentration of D but only that part of the concentration of D that is produced by channel i/j. In contrast to most experimental studies, this quantity is easily accessible in MD simulations.

While the rate constants $k_{i/j}(T)$ are unknown, the concentration profiles of A_i, B_j, C_j, and $D_{\mathrm{prod},i/j}$ can be easily computed by evaluating the reaction events. Combining eqs 1 and 4 yields the rate formulation

$$k_i(T) = \frac{\int_{t_0}^{t} r_i\,\mathrm{d}t'}{\int_{t_0}^{t}[A_i]\,\mathrm{d}t'} \tag{5}$$

$$k_j(T) = \frac{\int_{t_0}^{t} r_j\,\mathrm{d}t'}{\int_{t_0}^{t}[B_j][C_j]\,\mathrm{d}t'} \tag{6}$$

In case a reaction involves multiple reactants of the same species, their concentration needs to be adjusted to account for the smaller number of possible reaction partners. In the rate formulation in eq 6, the concentration $[C_j]$ would change to $([B_j] - 1/V)$ in case $[B_j]$ reacts with $[B_j]$.

3. METHODOLOGY

The rate formulation in eq 5 is based on knowledge of species concentrations and production of species via single reaction channels. In the present work, this information is extracted by quantitatively evaluating connectivity changes during reactive MD simulations using methane oxidation as an example.

3.1. Simulation Details. The reactions during the inception stage of hydrocarbon oxidation were studied by simulating a stoichiometic mixture of methane and oxygen at a

DOI: 10.1021/acs.jctc.5b00201
J. Chem. Theory Comput. 2015, 11, 2517−2524

density of 1125 mol m^{-3} and temperatures from 2000 to 2400 K. These conditions correspond to ideal gas pressures from 185 to 222 atm. We evaluated the effect of lower pressures and claim that the rate constants compared to literature data in Section 4.2 represent the high-pressure limit (cf. Supporting Information, Section S6). The methane−oxygen mixtures were simulated with the LAMMPS software package,[33] using the ReaxFF force field for hydrocarbon oxidation of Chenoweth et al.[16,34] This force field shows reasonable performance for methane pyrolysis[20] and oxidation.[25,34] The reactive force field trajectory simulations did not include quantum effects, such as tunneling. Nevertheless, zero-point corrections are partly included due to the choice of the force field training set.[16]

To study the gas phase chemistry of methane oxidation, we filled a cubic box with 30 methane molecules (CH_4) and 60 oxygen molecules (O_2) using the Packmol software package.[35] The box volume was calculated according to the initial number of molecules and the desired density, using the ideal gas law. Afterward, atomic velocities were assigned to be consistent with the Maxwell−Boltzmann distribution at the desired temperature.

For each temperature, 10 unique initial velocity distributions were assigned for the same initial positions. Each configuration was simulated for 5 ns, using a time step of 0.1 fs and a Nosé−Hoover thermostat[36,37] with a damping constant of 10 fs.[19] The atomic configurations and connectivities were written to text files every 1.0 ps and 20 fs (frame), respectively.

3.2. Equilibration of Molecular Modes. During the initial phase of the simulations, we observed thermal imbalance between molecular modes, resulting from the lack of vibrational potential energy of the energy-minimized molecules. The redistribution from overheated to vibrational modes, in turn, required collisions with other molecules, whose collision frequencies and energy transfer efficiencies determined the speed of thermalization. To check for thermalization, the distribution of thermal energy among the translational, rotational, and vibrational modes of the individual molecules was computed by projecting the atomic velocities on the center-of-mass motion, on the motion induced by angular velocity, and on the residual motion, respectively. The equilibration of molecular modes was monitored for all simulations (cf. Supporting Information, Figures S2−S4). On the basis of this data, we decided to omit the first nanosecond of the reactive MD simulations to ensure that any artificial disturbance was removed.

3.3. Detection and Determination of Reaction Pathways. Species concentrations (cf. eq 4) and integrals of reaction rates (cf. eq 1) were computed based on reaction events. In the following, we present a methodology for extracting reaction events from reactive MD trajectories.

From the atomic positions of each frame of the trajectory, the atomic connectivities were determined as a function of interatomic distance. This function must be robust against vibrational fluctuations of interatomic distances at elevated temperatures to avoid misinterpreting these periodic motions as dissociation and association. Therefore, we used the nonlinear bond-order relation proposed by van Duin et al.,[16] which is readily computed in ReaxFF simulations (cf. Supporting Information, Section S2). The chemical composition of each frame was determined using the graph-theoretical breadth-first search (BFS) algorithm[38] for extracting molecules from atomic connectivity. Each molecule was identified as a group of atoms separated from the remaining system after applying bond-order filters (cf. Supporting Information, Section S2).

Similar to recent approaches by Zheng and Pfaendtner[15] and Wang et al.,[8] we compared the atomic connectivities of consecutive frames to detect any change. A single connectivity change contained information about either dissociation, association, or isomerization reactions. In case multiple connectivity changes between the same frames involved the same molecule, we considered these events as parts of the same elementary reaction. For this procedure, the time between successive frames needs to be smaller than the time between consecutive elementary reactions to avoid merging these reactions into a single elementary reaction.

From the connectivity changes involved in an elementary reaction, the corresponding reaction equation, which included all participating molecules, was formulated. The graph representations used to describe these molecules were converted to string representations using the graph canonization and conversion algorithms implemented in the Open Babel software package.[39] We used the widely accepted canonical SMILES representation[40] for molecules to simplify the comparison of species.

Finally, the elementary reaction equations from all trajectories of all temperature conditions were compiled in a single reaction mechanism. This mechanism was reduced to important and competing pathways by omitting any pathway with an accumulated net flux of fewer than five reaction events. Furthermore, species were grouped according to their number of oxygen atoms. The 2D visualization of molecules and the graph layout was done using the depiction algorithms of the Open Babel software package[39] and the graph layout algorithms of the yEd software package,[41] respectively.

The above-described methodology was implemented in an in-house Python software package named **Chem**ical **Tra**jectory Anal**yzer** (ChemTraYzer), which is available online (cf. Supporting Information).

4. RESULTS AND DISCUSSION

We present the reaction pathways extracted from methane oxidation simulations and validate several important rate constants by comparing to literature data and quantum mechanical results.

4.1. Reaction Network of Methane Oxidation Simulations. The pathways extracted from the methane oxidation simulations at different temperatures were merged into a single mechanism. This network of reactions included 54 species and 257 elementary reactions leading to carbon dioxide and various C_2 species (cf. Supporting Information, Section S5). The mechanism presented in Figure 1 was generated by suppressing pathways with an accumulated net flux of less than five reaction events. Furthermore, pathways with a branching ratio below 5% were ignored, leading to a reduced mechanism of 18 species and 39 reactions, excluding C_0 species and their reactions.

The detailed mechanism that was extracted from reactive trajectories between 2000 and 2400 K included 22 of 34 relevant species and 46 of 168 relevant reactions of the widely accepted GRI-Mech 3.0 model,[42] referred to as GRI3 below. Furthermore, the remaining 32 species and 211 reactions detected in the present work are not considered in the GRI3 model, e.g., formic acid and part of its chemistry. The reduced mechanism in Figure 1 includes 17 of 34 relevant species and 10 of 168 relevant reactions; therefore, it includes 29 reactions that are not considered in the GRI3 model.

DOI: 10.1021/acs.jctc.5b00201
J. Chem. Theory Comput. 2015, 11, 2517−2524

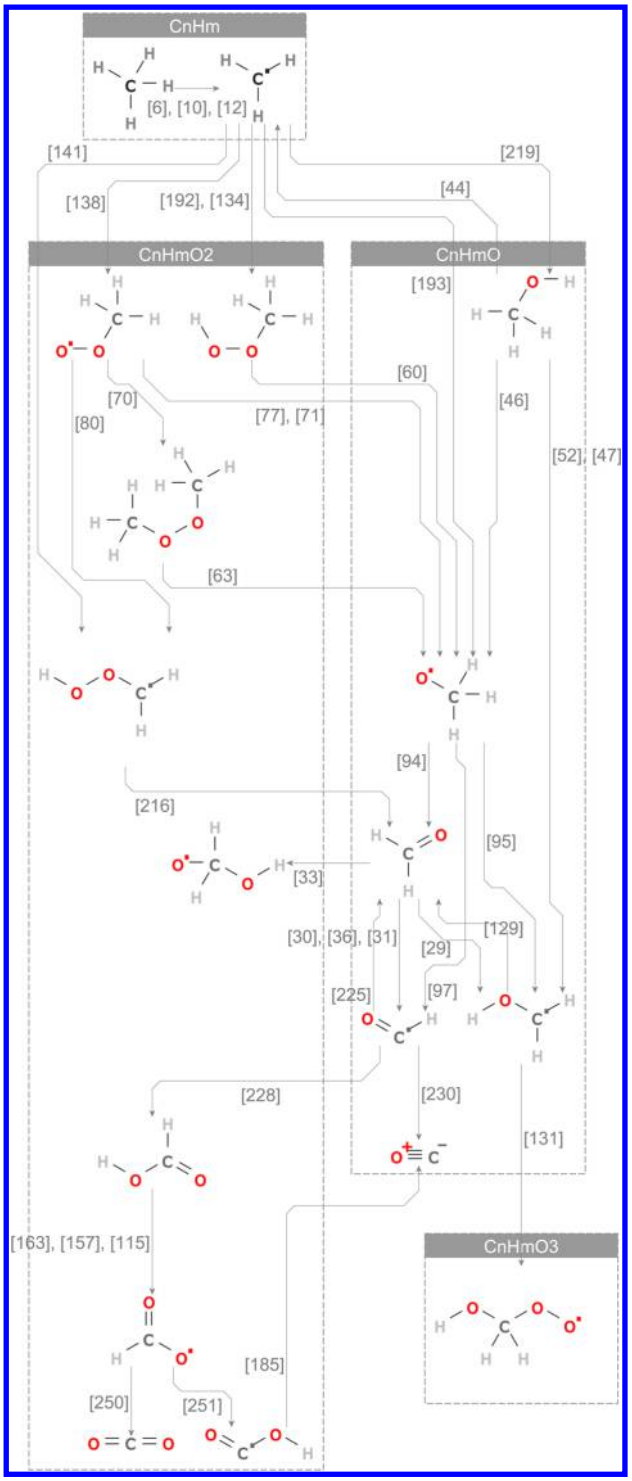


Figure 1. Important reactive MD pathways of the first 5 ns of methane oxidation. Species are grouped according to their count of oxygen atoms, and species without carbon are hidden. Each edge is labeled with the reaction IDs of all reactions connecting the species.

As expected, the combustion of methane was initialized by forming methyl radicals via hydrogen abstraction. Subsequent reactions of methyl radicals with molecular oxygen or hydrogenated oxygen species led to methoxy radicals, which, in turn, was stabilized to formaldehyde via hydrogen loss. While the GRI3 model includes methanol formation and subsequent reactions toward formaldehyde, it neglects the alternative pathways via methyldioxidanyl and methyl-hydroperoxide, which Reid et al.[43] include, emphasizing their significance. In the mechanism shown in Figure 1, formaldehyde can react to CO only by forming formyl radicals. These radicals, in turn, formed stable formic acid via association with hydroxyl radicals, the kinetics of which were studied recently by de Souza Machado and Bauerfeldt.[44] Subsequent hydrogen loss led to the formation of CO_2 and thus to complete oxidation. The formation of carbon dioxide via formic acid is not considered in the GRI3 model, but it is the dominant pathway in our MD simulations. Competing with the formation of CO_2, several C_2 species formed during the simulations. Species such as ethylene formed as expected from the GRI3 model, whereas others, such as dimethyl ether, were not predicted by that model.

Since the presented mechanism was deduced from high-temperature reactive MD investigations, pathways of the GRI3 model that dominate the low-temperature regime were not expected to be observed during simulations. Furthermore, alternative pathways observed during simulations need to be validated via QM computations before they can be used to update existing models.

4.2. Reaction Rates of Methane Oxidation Simulations. Rate constants of three frequently observed hydrogen-loss reactions of methane were compared to literature rates. Table 1 summarizes the reaction event counts in the 5 ns methane oxidation simulations at various temperatures. For each temperature, the reaction event counts were averaged over the 10 replicates.

Table 1. Averaged Reaction Event Counts[a]

reaction	temperature T/K		
	2000	2200	2400
$CH_4 \rightarrow CH_3^\bullet + H^\bullet$	0.1	0.6	2.0
$CH_4 + O_2 \rightarrow CH_3^\bullet + HO_2^\bullet$	0.6	1.1	1.7
$CH_4 + {}^\bullet OH \rightarrow CH_3^\bullet + H_2O$	0.5	2.2	4.4

[a]Reaction event counts for frequently observed hydrogen-loss reactions of methane in the 5 ns replica simulations. Event counts were averaged over 10 replicates for each temperature.

The authenticity of reaction pathways of classical trajectories can be proved by validating the respective transition state with high-level QM. This approach was used to validate the hydrogen-loss reactions (cf. Table 1), the transition states of which were extracted from reactive MD simulations using the NEB and the climbing image (CI) method.[30−32] The resulting ReaxFF geometries for reactants, transition states, and products were used as input for CCSD(T)/aug-cc-pVTZ//B2PLYPD3/6-311++G(d,p) calculations, performed with revision D.01 of the Gaussian09 software package.[45] Electronic and thermochemical energies were used to compute the energetic barriers and the rigid-rotor harmonic oscillator free energy barriers, respectively. Furthermore, the ReaxFF free energy barriers $\Delta G^{\ddagger}$ for the investigated reactions were deduced from the rate constants extracted from reactive MD simulations at different temperatures assuming transition-state theory

$$\Delta G^{\ddagger}(T) = -RT \ln\left(\frac{h \cdot k^*(T)}{k_B T}\right) \tag{7}$$

where R is the gas constant, T is the temperature, h is Planck's constant, k_B is Boltzmann's constant, and $k^*(T)$ is the first-order rate constant. The rate constants were arithmetically averaged over all 10 replicates.

DOI: 10.1021/acs.jctc.5b00201
J. Chem. Theory Comput. 2015, 11, 2517−2524

Since the theoretical background of eq 7 excludes recrossing of the reaction barrier,[46] we neglected any reaction event associated with fast back-and-forth reactions when computing $k(T)$. For that purpose, the history of reaction products was traced by storing reaction events for a short period of time (cf. Supporting Information, Section S3 for details).

4.2.1. Unimolecular Rate Constants. The unimolecular hydrogen-loss reaction of methane was frequently observed (cf. Table 1), despite its rather low rate.[47−50] These slow reactions could be observed due to the initially large amount of CH_4 molecules.

In Figure 2, the unimolecular reactive MD rate constant of the $CH_4 \rightarrow CH_3^\bullet + H^\bullet$ dissociation reaction is shown in comparison to theoretical[47,48] and experimental data[49,50] which was partly used in the GRI3 model.[42] The presented rate constants agree to within an order of magnitude with the literature data for the temperature regime from 2000 to 2400 K. While at high temperatures the agreement is better and the error bars of the reactive MD rate constant are small, at lower temperatures, the statistical basis is insufficient to make reliable statements about the accuracy of ReaxFF. The slightly overestimated rate constants of CH_4 dissociation are consistent with the underestimated heat release of CH_4 formation.[16] In contrast, the dissociation limit computed via ReaxFF was in very good agreement with QM results (cf. Supporting Information, Table S3).

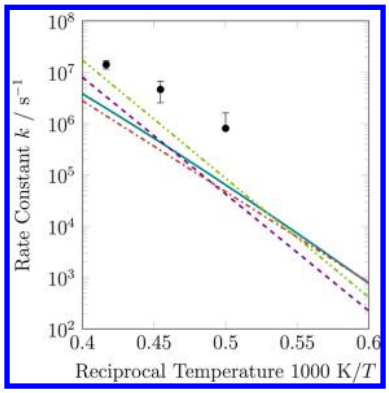


Figure 2. $CH_4 \rightarrow CH_3^\bullet + H^\bullet$ rate. Comparison of literature data of Cobos and Troe (solid blue/green line),[47] Takahashi et al. (dashed purple line),[48] Koike et al. (dashed-dotted red/orange line),[49] and Sutherland et al. (dashed-dotted light green line)[50] to the present work (●).

4.2.2. Bimolecular Rate Constants. A statistically reasonable number of hydrogen abstraction reactions from CH_4 by O_2 were observed (cf. Table 1), despite the rather low reactivity of O_2. Hydrogen abstraction reactions from CH_4 by hydroxyl radicals ($^\bullet$OH) were observed frequently, despite the low concentration of this radical, due to its high reactivity. This radical was produced by H_xO_y decomposition reactions and by some of the reactions shown in the mechanism in Figure 1.

In Figures 3 and 4, the bimolecular reactive MD rate constants of the abstraction reactions from CH_4 by O_2 and $^\bullet$OH are shown in comparison to theoretical[51,53,54] and experimental data.[43,55] Theoretical studies were based on different levels of transition state theory,[53,54] except for the work of Shaw,[51] which was based on estimations. Experimental studies were based on either shock tube experiments[55] or flame studies.[43] For reaction with O_2, reference data was available from Baulch et al.[52]

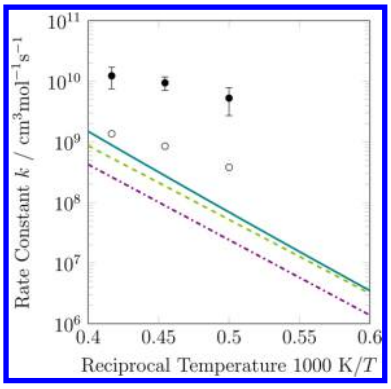


Figure 3. $CH_4 + O_2 \rightarrow CH_3^\bullet + HO_2^\bullet$ rate. Comparison of literature data of Shaw (solid blue/green line),[51] Reid et al. (dashed light green line),[43] and Baulch et al. (dashed-dotted purple line)[52] to the present work (●) and the present work with corrected barrier (○).

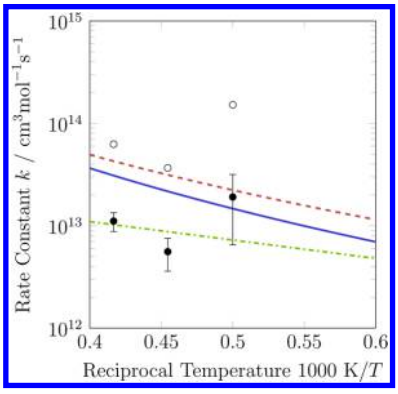


Figure 4. $CH_4 + {}^\bullet OH \rightarrow CH_3^\bullet + H_2O$ rate. Comparison of literature data of Melissas and Truhlar (solid purple line),[53] Schwartz et al. (dashed red/orange line),[54] and Srinivasan et al. (dashed-dotted light green line)[55] to the present work (●) and the present work with corrected barrier (○).

While the rate constants for hydrogen abstraction from CH_4 by $^\bullet$OH were in good agreement with literature data, abstraction by O_2 was overestimated by more than an order of magnitude. Comparing ReaxFF and QM barriers revealed deviations for both bimolecular reactions (cf. Supporting Information, Table S3). For abstraction by O_2, the ReaxFF barrier was too low, whereas the ReaxFF barrier was too high for abstraction by $^\bullet$OH. The observed deviation from the QM reaction barrier of abstraction by O_2 was in agreement with the deviation from the literature rate constants. For abstraction by $^\bullet$OH instead, the ReaxFF rates were already in reasonable agreement with literature data, although the averaged rate constant at 2000 K was larger than expected. This nonlinear behavior is expected to vanish as the number of available data points increases.

Furthermore, the deviation between the ReaxFF and QM free energy barriers was especially pronounced for abstraction by $^\bullet$OH (cf. Supporting Information, Table S4). Despite the rather large deviation, the rate constants agreed reasonably with the literature data of Melissas and Truhlar.[53] Since the authors used a variational transition state formulation, the good agreement between the ReaxFF and the literature data is assumed to be due to variational effects of hydrogen abstraction from CH_4 via $^\bullet$OH. These effects are readily implemented when using reactive MD simulations for studying kinetics.

DOI: 10.1021/acs.jctc.5b00201
J. Chem. Theory Comput. 2015, 11, 2517−2524

The reaction barriers at CCSD(T) level of theory were used to correct the bimolecular rate constants according to the following formulation

$$k_{corr} = k_{ReaxFF}\frac{\exp(-\Delta E^{\ddagger}_{CCSD(T)}/RT)}{\exp(-\Delta E^{\ddagger}_{ReaxFF}/RT)} \quad (8)$$

where $\Delta E_i^{\ddagger}$ is the energetic barrier obtained with method i, which is either CCSD(T) or ReaxFF (cf. Supporting Information, Table S3). Assuming that rate constants can be expressed by an Arrhenius law, eq 8 replaces the ReaxFF barrier with the CCSD(T) barrier. The corrected rate constants are given by the small circles (○) in Figures 3 and 4. This correction shifted the abstraction by •OH to larger rate constants and improved the agreement with literature data for abstraction by O_2. Nevertheless, when considering the strong multireference character of the $CH_4 + O_2$ transition state (T1 diagnostics of ca. 0.07), the CCSD(T) barrier used to update the respective rate constants gives only a rough estimate of the QM barrier. The complex electronic structure of this particular reaction would require multireference computations to obtain a precise barrier.

4.3. Applicability and Comparison. The presented rate formulation can be applied if the thermal energy distribution of the reactive MD simulations is not artificially perturbed. The methodology presented for analyzing reactive trajectories required a maximum time of 20 fs or less between consecutive frames to avoid merging of consecutive elementary reactions. Furthermore, the time period used to filter recrossing events (e.g., 2 ps) requires careful testing when investigating new systems.

The rate constants deduced from the presented reactive trajectories were in the high-pressure limit due to the high rate of collisions and the use of a thermostat for removing nonthermal effects. Unimolecular reactions with a rate of approximately 8×10^5 s^{-1} or higher and bimolecular reactions with a rate of roughly 5×10^9 cm^3 mol^{-1} s^{-1} or higher were observed. For the discovery of less frequent events, i.e., rare events, proper acceleration technique need to be incorporated, such as metadynamics[56] or the recent virtual piston approach.[57]

Reactive MD simulations performed with a state-of-the-art empirical force field gave insight into mechanisms on an atomistic level in the present study. Therefore, we expect the rate estimates to be at least as reliable as conventional rules for estimation.[9,11,58] Nevertheless, the small-sized systems investigated in the present work yield large statistical errors for rarely observed reactions. Furthermore, using quantum mechanical trajectory simulations, such as BOMD[8] or CPMD,[15] will most likely increase the accuracy of rate predictions. The presented methodology is equally applicable for postprocessing BOMD/CPMD trajectories.

QM trajectory simulations are typically limited to small systems of molecule pairs[15] at the DFT level or small molecule clusters[8] at the HF level. In the present work, the computationally fast force field computations allow for increasing the system size and therefore allow for reactions of up to 54 different species in this case. In comparison to the *ab initio* nanoreactor,[8] using a reactive force field reduced the computational effort required for investigation by almost 4 orders of magnitude (cf. Supporting Information, Section S8).

Complementary to reactive MD investigations for mechanism studies, the approach by Rappoport et al.[14] makes use of mathematical transformation rules for generating reaction events. Since these rules are not limited by the reaction kinetics of the investigated system, these rules are numerically more efficient for discovering new chemistry. Nevertheless, being able to discover new chemistry crucially depends on the transformation rules used. Inadequate or incomplete rules may lead to missing species. Instead, reactive MD aims to reproduce the physics of the system and thus is assumed to allow for all possible chemical reactions. Furthermore, Rappoport et al.[14] provided a framework for estimating the reaction kinetics between parts of a mechanism by approximating the transition states between low-energy states by high-energy intermediates. The present work, hence, aims to provide rate estimates for all elementary reactions.

5. CONCLUSIONS

We have presented a new approach to obtain quantitative models for complex chemical reaction networks based on reactive MD simulations. Compared to previous MD mechanism studies, our approach allows us to identify, quantify, and evaluate a multitude of reactions from a single simulation, including reactions of short-lived reaction intermediates. This automated procedure was used to generate qualitative and quantitative descriptions of complex chemical processes, as shown for methane oxidation.

Although an empirical force field was used, the rate constants deduced from the reactive MD simulations were in reasonable agreement with widely accepted literature data. Furthermore, using reactive trajectories allowed us to compute rate constants for barrierless reactions, which are difficult to handle in theoretical kinetic studies. Therefore, rate constants derived by means of the presented formulation may be used as good first approximations. Analyzing the sensitivity of a process to these estimates yields the most important reactions, the stationary points of which could be re-evaluated with QM subsequently. This would not only allow energetics to be updated but also for quantum effects to be included, such as tunneling.

Compared to the recently developed *ab initio* nanoreactor concept,[8] the computational effort of the present study is nearly 4 orders of magnitude lower. In addition, the unperturbed kinetics of the presented reactive MD simulations allowed us to extract rate constants, which cannot be done with the nanoreactor. Replica MD simulations can be extended to study various mechanisms, such as the combustion-relevant reaction network of methane oxidation at temperatures between 2000 and 2400 K at engine-relevant pressures. Future work includes implementing appropriate acceleration techniques to allow investigation of low-temperature reaction networks. Furthermore, updating kinetics with QM data will be possible by automatically extracting molecular structures of reactants and transition states from trajectories using NEB/CI methods.

The present study underlines the potential of reactive MD for developing and optimizing reaction mechanisms and models. We believe the presented methodology will increase the importance of kinetic studies with reactive MD for developing quantitative reaction models.

■ ASSOCIATED CONTENT

Supporting Information

Validation of the rate formulation (Section S1), limitations of the presented reaction detection methodology (Section S2), more detailed information on the recrossing filter (Section S3), information on thermalization of the presented simulations

DOI: 10.1021/acs.jctc.5b00201
J. Chem. Theory Comput. 2015, 11, 2517−2524

(Section S4) and more detailed results (Sections S5 and S6), results of quantum mechanical computations (Section S7) and the performance of the presented methodology (Section S8). The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.5b00201. The **Chem**ical **Tra**jectory Anal**yzer** (ChemTraYzer) software package for postprocessing and analyzing of ReaxFF trajectories can be downloaded from the following website: http://www.ltt.rwth-aachen.de/reaction_engineering.

AUTHOR INFORMATION

Corresponding Author

*E-mail: kai.leonhard@ltt.rwth-aachen.de.

Funding

This work was performed as part of the Cluster of Excellence "Tailor-Made Fuels from Biomass" (EXC 236), which is funded by the Excellence Initiative by the German federal and state governments to promote science and research at German universities. Financial support from the Deutsche Forschungsgemeinschaft (German Research Association) through grant GSC 111 is gratefully acknowledged.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Yang, G.; Pidko, E. A.; Hensen, E. J. M. *J. Catal.* **2012**, *295*, 122−132.
(2) Semsarilar, M.; Perrier, S. *Nat. Chem.* **2010**, *2*, 811−820.
(3) Harper, M. R.; van Geem, K. M.; Pyl, S. P.; Marin, G. B.; Green, W. H. *Combust. Flame* **2011**, *158*, 16−41.
(4) Balthasar, M.; Mauss, F.; Knobel, A.; Kraft, M. *Combust. Flame* **2002**, *128*, 395−409.
(5) Contino, F.; Masurier, J.-B.; Foucher, F.; Lucchini, T.; D'Errico, G.; Dagaut, P. *Fuel* **2014**, *137*, 179−184.
(6) Michaelis, L.; Menten, M. L. *Biochem Z.* **1913**, *49*, 333−369.
(7) Lindemann, F. A.; Arrhenius, S.; Langmuir, I.; Dhar, N. R.; Perrin, J.; Lewis, W. C. M. *Trans. Faraday Soc.* **1922**, *17*, 598−606.
(8) Wang, L.-P.; Titov, A.; McGibbon, R.; Liu, F.; Pande, V. S.; Martínez, T. J. *Nat. Chem.* **2014**, *6*, 1044−1048.
(9) Mok, M. H.; Polanyi, J. C. *J. Chem. Phys.* **1969**, *51*, 1451−1469.
(10) Benson, S. W.; Cruickshank, F. R.; Golden, D. M.; Haugen, G. R.; O'Neal, H. E.; Rodgers, A. S.; Shaw, R.; Walsh, R. *Chem. Rev.* **1969**, *69*, 279−324.
(11) Sumathi, R.; Green, W. H. *Phys. Chem. Chem. Phys.* **2003**, *5*, 3402−3417.
(12) Green, W. H.; Allen, J. W.; Buesser, B. A.; Ashcraft, R. W.; Beran, G. J.; Class, C. A.; Gao, C.; Goldsmith, C. F.; Harper, M. R.; Jalan, A.; Keceli, M.; Magoon, G. R.; Matheu, D. M.; Merchant, S. S.; Mo, J. D.; Petway, S.; Raman, S.; Sharma, S.; Song, J.; Suleymanov, Y.; van Geem, K. M.; Wen, J.; West, R. H.; Wong, A.; Wong, H.-W.; Yelvington, P. E.; Yee, N.; Yu, J. *Reaction Mechanism Generator*, v4.0.1; MIT: Cambridge, MA, 2014; http://rmg.sourceforge.net/.
(13) Barner-Kowollik, C.; Coote, M. L.; Davis, T. P.; Radom, L.; Vana, P. *J. Polym. Sci., Part A* **2003**, *41*, 2828−2832.
(14) Rappoport, D.; Galvin, C. J.; Zubarev, D. Y.; Aspuru-Guzik, A. *J. Chem. Theory Comput.* **2014**, *10*, 897−907.
(15) Zheng, S.; Pfaendtner, J. *J. Phys. Chem. C* **2014**, *118*, 10764−10770.
(16) van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A., III *J. Phys. Chem. A* **2001**, *105*, 9396−9409.
(17) Yue, D.-C.; Ma, T.-B.; Hu, Y.-Z.; Yeon, J.; van Duin, A. C. T.; Wang, H.; Luo, J. *J. Phys. Chem. C* **2013**, *117*, 25604−25614.
(18) Abolfath, R. M.; van Duin, A. C. T.; Brabec, T. *J. Phys. Chem. A* **2011**, *115*, 11045−11049.
(19) Chenoweth, K.; van Duin, A. C. T.; Dasgupta, S.; Goddard, W. A., III *J. Phys. Chem. A* **2009**, *113*, 1740−1746.
(20) Lümmen, N. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7883−7893.
(21) Liu, L.; Bai, C.; Sun, H.; Goddard, W. A., III *J. Phys. Chem. A* **2011**, *115*, 4941−4950.
(22) Ding, J.; Zhang, L.; Zhang, Y.; Han, K.-L. *J. Phys. Chem. A* **2013**, *117*, 3266−3278.
(23) Qi, T.; Bauschlicher, C. W., Jr.; Lawson, J. W.; Desai, T. G. *J. Phys. Chem. A* **2013**, *117*, 11115−11125.
(24) Cheng, X.-M.; Wang, Q.-D.; Li, J.-Q.; Wang, J.-B.; Li, X.-Y. *J. Phys. Chem. A* **2012**, *116*, 9811−9818.
(25) He, Z.; Li, X.-B.; Liu, L.-M.; Zhu, W. *Fuel* **2014**, *124*, 85−90.
(26) Yan, G.; Zhang, Z.; Yan, K. *Mol. Phys.* **2013**, *111*, 147−156.
(27) Furman, D.; Kosloff, R.; Dubnikova, F.; Zybin, S. V.; Goddard, W. G., III; Rom, N.; Hirshberg, B.; Zeiri, Y. *J. Am. Chem. Soc.* **2014**, *136*, 4192−4200.
(28) Joshi, K. L.; Raman, S.; van Duin, A. C. T. *J. Phys. Chem. Lett.* **2013**, *4*, 3792−3797.
(29) Lucas, K. *Molecular Models for Fluids*; Cambridge University Press: Cambridge, 2007; pp 127−143.
(30) Henkelman, G.; Jónsson, H. *J. Chem. Phys.* **2000**, *113*, 9978.
(31) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. *J. Chem. Phys.* **2000**, *113*, 9901.
(32) Nakano, A. *Comput. Phys. Commun.* **2008**, *178*, 280−289.
(33) Plimpton, S. J. *J. Comput. Phys.* **1995**, *117*, 1−19.
(34) Chenoweth, K.; van Duin, A. C. T.; Goddard, W. A., III *J. Phys. Chem. A* **2008**, *112*, 1040−1053.
(35) Martiínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. *J. Comput. Chem.* **2009**, *30*, 2157−2164.
(36) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511−519.
(37) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695−1697.
(38) Cormen, T. H.; Leiserson, C. E.; Rivest, R. L.; Stein, C. *Introduction to Algorithms*; MIT Press: Cambridge, MA, 2001; pp 531−539.
(39) O'Bolye, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeerch, T.; Hutchison, G. R. *J. Cheminf.* **2011**, *3*, 33.
(40) Weininger, D. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31−36.
(41) *yEd Graph Editor*, 3.12.2; yWorks GmbH: Tübingen, Germany, 2000−2015; http://www.yworks.com/de/products_yed_about.html.
(42) Smith, G. P.; Golden, D. M.; Frenklach, M.; Moriarty, N. W.; Eiteneer, B.; Goldenberg, M.; Bowman, C. T.; Hanson, R. K.; Song, S.; Gardiner, W. C., Jr.; Lissianski, V. V.; Qin, Z. *GRI-MECH* 3.0, 2015; http://www.me.berkeley.edu/gri-mech/version30/text30.html.
(43) Reid, I. A. B.; Robinson, C.; Smith, D. B. *Symp. Int. Combust.* **1985**, *20*, 1833−1843.
(44) de Souza Machado, G.; Bauerfeldt, G. F. *Theoretical investigation of formaldehyde reactions initiated by OH radicals*, 7th European Combustion Meeting, Budapest, Hungary, March 30−April 02, 2015.
(45) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision A.01; Gaussian, Inc.: Wallingford, CT, 2009.
(46) Truhlar, D. G.; Garrett, B. C. *Annu. Rev. Phys. Chem.* **1984**, *35*, 159−189.
(47) Cobos, C. J.; Troe, J. *Z. Phys. Chem.* **1992**, *176*, 161−171.
(48) Takahashi, J.; Momose, T.; Shida, T. *Bull. Chem. Soc. Jpn.* **1994**, *67*, 74−85.
(49) Koike, T.; Kudo, M.; Yamada, H. *Int. J. Chem. Kinet.* **2000**, *32*, 1−6.

DOI: 10.1021/acs.jctc.5b00201
J. Chem. Theory Comput. 2015, 11, 2517−2524

(50) Sutherland, J. W.; Su, M.-C.; Michael, J. V. *Int. J. Chem. Kinet.* **2001**, *33*, 669−684.
(51) Shaw, R. *J. Phys. Chem. Ref. Data* **1978**, *7*, 1179.
(52) Baulch, D.; Cobos, C.; Cox, R.; Esser, C.; Frank, P.; Just, T.; Kerr, J. A.; Pilling, M. J.; Troe, J.; Walker, R. W.; Warnatz, J. J. *Phys. Chem. Ref. Data* **1992**, *21*, 411−429.
(53) Melissas, V. S.; Truhlar, D. G. *J. Chem. Phys.* **1993**, *99*, 1013−1027.
(54) Schwartz, M.; Marshall, P.; Berry, R. J.; Ehlers, C. J.; Petersson, G. A. *J. Phys. Chem. A* **1998**, *102*, 10074−10081.
(55) Srinivasan, N. K.; Su, M. C.; Sutherland, J. W.; Michael, J. V. *J. Phys. Chem. A* **2005**, *109*, 1857−1863.
(56) Laio, A.; Gervasio, F. L. *Rep. Prog. Phys.* **2008**, *71*, 126601.
(57) Goldman, N. *J. Nat. Chem.* **2014**, *6*, 1033−1034.
(58) Benson, S. W. *Thermochemical Kinetics: Methods for the Estimation of Thermochemical Data and Rate Parameters*; John Wiley & Sons: New York, 1976; p 24ff.

DOI: 10.1021/acs.jctc.5b00201
J. Chem. Theory Comput. 2015, 11, 2517−2524