RUHR-UNIVERSITÄT BOCHUM
FAKULTÄT FÜR CHEMIE UND BIOCHEMIE
LEHRSTUHL FÜR THEORETISCHE CHEMIE

RUB

Atomistic Multiscale Modeling of Char Combustion Reactions

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften

von

Özlem Yönder

Acknowledgements

I would like to thank the German Research Foundation (DFG) for funding the SFB/Transregio 129 Oxyflame. The interdisciplinary environment in Oxyflame made me see the potential of this project to make a real contribution to the development of furnaces, which kept me motivated. I would like to thank all of my colleagues who are part of Oxyflame for creating this environment.

I would like to express my deepest appreciation to my supervisor Prof. Christof Hättig. He was always attuned to my progress as well as the progress of the project. I have learned a lot from his invaluable insights, extensive knowledge and experience in theoretical chemistry. He has also always encouraged us to keep a work–life balance and provided a peaceful working environment.

I am also extremely grateful to my second supervisor Prof. Rochus Schmid, who always came up with creative ideas. His expertise in relational databases has been very enriching for this project and I have learned a lot on the way. I'd like to recognize the help that I received from his former PhD student Dr. Johannes Dürholt, who always made time to answer my questions on coding in Python and using pyDAL. I would also like to thank my collaborator Vanessa Angenent. I am also grateful to Dr. Gunnar Schmitz who has joined the project last year with his enthusiasm and efficiency. Collaborating with him has been a pleasure.

I would like to extend my sincere gratitude to Prof. Kai Leonhard, Felix Schmalz, and Lukas Krep for making ChemTraYzer code accessible to us and for providing a deeper understanding of the code.

I would also like to thank my colleagues from the Quantum Chemistry research group, especially my office–mates before the Covid-19 pandemic, Arjeta Rushiti and Matthias Rückert. I also very much appreciate the support of my friends from the Theoretical Chemistry Department, especially Esma Birsen Boydaş, Dr. Abhishek Khedkar, and Mihkel Ugandi, who have always given me useful feedback before an important oral

presentation.
The completion of my dissertation would not have been as easy without my dear friend and former colleague Dr. Marius Sebastian Frank. He has always kept the computer clusters running smoothly, answered my questions about coupled cluster calculations and made me join the "Pair Natural Orbitals Fan Club". He has supported me in every aspect of life and stood by my side during the whole process.
I am also very thankful to all my friends who have been part of my life during the process: to Dorine Chézière Théière who nurtured my inspiration through her art projects; to Gemma Maskell who has always been there for me with delicious food and heart-to-hearts and lately with her stories about the mysterious world of nano impact electrochemistry; to Zeynep Kapısız who has been a good friend to me and brought the memories of home to Bochum; to Helena Düwel who has brought joy and mindfulness to my life; to my acroyoga family, especially Nadine Naujoks, Aslı Yılmaz, and Jan Sessing, who always lifted my mood; to my bouldering and climbing partners, especially Robert Krieger, David Paul Maharam, Daniel Pohl, and Charlotte Herz, who have helped me push my limits, and through this, inspired me to finish the challenging tasks during my PhD project; and to all of my friends whom I could not mention here with their names.
I cannot express my thanks to my family, who always did their best to support me in every aspect of my life. Thank you my sister Çiğdem Yönder and my cousin Ayberk Özkırlı, I am so blessed to have you in my life and I know that I will never be alone in this world. I am so thankful to "Özkırlı kooperatifi", who always made my day start with a sunny "günaydın".
Last but not least, I would also like to thank myself. Living alone and working alone, the Covid-19 pandemic hit me hard because I love being with people and find energy through connection, but I showed up for myself to give a pep talk and sought ways to change my perspective and improve the situation.

Contents

I. Introduction 1

A. Motivation and research questions 1

B. Multiscale and multimethod modeling of char burnout: A look at the kinetic models for combustion modeling 7

C. Multiscale and multimethod modeling of char burnout: The role of computational chemistry 10

1. Quantum chemistry 15

2. Molecular mechanics 22

3. Statistical mechanics 24

4. The transition state theory 28

5. Composite methods: First multimethod approaches 31

II. CO formation through reactions of polyaromatic hydrocarbons with atomic oxygen 33

III. Can Small Polyaromatics Describe Their Larger Counterparts for Local Reactions? A Computational Study on the H-Abstraction Reaction by an H-Atom from Polyaromatics 45

IV. Computational Chemistry as a Tool for Predictive Combustion Modelling 59

V. An automatized workflow from molecular dynamic simulation to quantum chemical methods to identify elementary reactions and compute reaction constants 75

VI. Conclusions 128

I. Introduction

A. Motivation and research questions

Decreasing the anthropogenic greenhouse gas emissions caused by burning of fossil fuels, deforestation, land use changes, livestock production, fertilization, waste management, and industrial processes, is a big challenge of our age. Our world is already approximately 1°C warmer than the averages from the pre-industrial period. Some serious adverse effects of anthropogenic global warming include the increase in the number of extreme events such as heatwaves, droughts, extreme precipitation; the ice sheets decreasing in mass; the rising sea levels; the ocean acidification; increased duration and intensity of the wildfires due to surface dryness[1], which do not only damage the ecosystems but also contribute to carbon dioxide (CO_2) emissions; adverse effects on biodiversity, such as sea turtles who are at the risk of extinction due to their temperature dependent sex[2,3], and on ecosystems, such as thawing permafrost, or waning of the Gulf Stream which sustains the ecosystems of the Atlantic. There are also unforeseen effects of global warming. For example, a recent study has found that the plants are getting less and less efficient in capturing CO_2 from the atmosphere due to a decrease in the other ingredients of photosynthesis, such as nitrogen, phosphorus, or moisture[4].

Increasing global temperatures, will only intensify these adverse effects. In 2015, the Paris Agreement, a legally binding international treaty on climate change, has been signed by 196 parties with a goal to limit global warming below 2°C and preferably by 1.5°C by 2050 compared to the pre-industrial levels. The greenhouse gas (GHG) emissions from 2017 are shown in Figure 1. CO_2 makes up 74% of the GHG emissions. Among the CO_2 emissions, 86% originates from energy consumption. Indeed, the 1.5°C goal requires a fully decarbonized power sector[6]. Anthropogenic carbon dioxide removal (CDR) comprises mainly of agriculture, forestry and other land use (AFOLU), e.g., through reforestration and ecosystem restoration; and bioenergy with

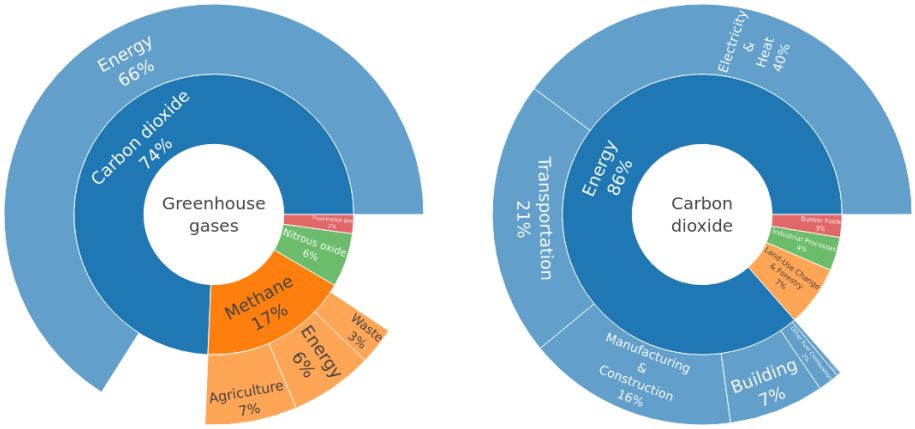


FIG. 1. The distribution of greenhouse gas emission (left) and CO_2 emissions (right) from different sectors in 2017. Data retrieved from Climate Analysis Indicators Tool (CAIT).[5]

carbon capture utilization and storage (BECCUS)[6]. The increase in the wild-fires and the decrease in the CO_2 fertilization effect (vide supra), might make the former less efficient. On the other hand, there are uncertainties whether or not BECCUS is the technology of the future[7]. However, despite the opposing public opinion and lack of political interest[8–10], the climate change mitigation pathways indicate[6] a reliance on BECCUS, particularly those with high overshoot of 1.5°C. Therefore, in order to overcome the technological readiness barrier of its deployment[9], the research and scaling should be done rather now than too late.

Any technology which relies on resources will face the risk of unsustainability. Even the publicly and politically accepted green energy systems, such as photovoltaics, wind-turbines, and batteries for electric-cars, are reliant on metal supply. This increasing demand risks to shift one non-renewable energy source to another one; in other words, instead of depending on fossil fuels, being dependent on the finite metal and mineral reserves[11]. Therefore, it is important to assure that the mining can be done sustainably. The doubts against biomass availability are similar. Bio-energy can become a "threat to energy security" if biomass comes along with long transportation costs, and hence, added CO_2 emissions,[7] or if it relies on unsustainable sources such as palm-

oil, which causes destruction of forests and peatlands[8]. Instead, the local marginal lands should be used to grow bio-energy crop mixtures which don't require nitrogen fertilizers or agrochemicals, and in fact, potentially enhance the biodiversity of these lands[12]. Another critical point is the carbon capture utilization and storage itself. These include, in terms of storage, the risk of leakage and the large capacity demand and, in terms of utilization, the availability of hydrogen that comes from low-carbon sources[10].

In terms of energy conversion, biomass can be used through different conversion pathways[13]. First, it can go through thermo-chemical conversion. For example, it can be gasified in presence of steam and used as syn-gas. Alternatively, it can be pyrolysed by increasing temperature at the absence of oxygen and converted to bio-oil, bio-char and bio-gas. Second, it can go through bio-chemical conversion to produce fuel such as bio-ethanol through fermentation. Finally, it can also be used as solid fuel through physical processing such as grinding, drying, pelletization, or briquetting.

The project Oxyflame CRC/TRR 129 is concerned with the application of BECCUS on power generation using oxy-fuel combustion technology with direct combustion of plant-based solid biofuels. Therefore, the project also aims developing a comprehensive predictive model of oxy-fuel combustion from microscopic level up to the scale of industrial burners and boilers. The main advantage of oxy-fuel combustion over post-combustion technology is that there is no need to use solvents for separating the CO_2 and N_2 from the effluent gas, because in the oxy-fuel combustion the N_2 is already separated before the combustion, e.g., through cryogenic air separation. During direct combustion, biomass goes through drying, pyrolysis (devolatilisation), char burnout and gas-phase reactions. The char burnout step is the rate limiting step for the total fuel conversion, and therefore, determines the design of the furnace and the boiler. Thus, it is important to understand this step. During char burnout the reacting gases diffuse to the particle surface and through the particle pores, the reacting gases are adsorbed on the surface, the chemical reactions take place, and the products desorb

and diffuse out of the pores and through the boundary layer[14]. At low temperatures reaction kinetics determine the rate of char burnout, whereas at high temperatures it is the diffusion[14]. Therefore, it is important to build predictive models for both the diffusion and the reaction kinetics which are applicable to solid fuels with different chemical compositions and mesoscopic structures, such as those with different pore size distributions.

In order to build a char burnout model, the structure of the char should also be understood. A comprehensive model of biochar is the dynamic molecular model of biochar, which explains the differences in the structure of biomass as the charring intensity, i.e., the pyrolysis temperature, increases[15]. Accordingly, at the temperature range below 200°C dehydration takes place and the structure of cellulose, lignin, and hemicellulose is not broken, whereas by increasing temperature the small volatiles such as anhydro sugars, pyrans, and furans start to be dissociated. At the range $\sim 350 - 450$°C the long-range ordering is lost and small, heat-resistant aliphatic and aromatic structural units form a randomly disordered carbon-phase called amorphous carbon. Above this temperature, turbostratic crystallites begin to form, and above 700°C there is no more amorphous carbon and the biomass is fully transformed into turbostratic crystallites and pore space.

The scientific work of this dissertation comprises the development of tools and workflows to be used for atomistic modeling of char burnout as a part of the A7 project within Oxyflame CRC/TRR 129. The investigation of the effects of heteroatoms, which may change the reaction mechanisms, is beyond the scope of this dissertation and only models composed of C, H, and O atoms will be considered.

The two significant aspects of atomistic modeling of char burnout is to describe the reaction kinetics as well as the weak intermolecular interactions which determine the sorption and diffusion properties. The complex structure of biochar and the different time scales of char burnout poses a multiscale problem, which requires a multimethod approach to describe both of these processes. In terms of the scale of computational

chemistry methods, the quantum mechanical methods resolve the electronic structure for small molecules whereas with molecular mechanics much larger length and time scales are accessible. Therefore, different computational chemistry methods are used in the studies presented in this dissertation. These methods are described briefly in section I C. The focus of this dissertation is the quantum chemical aspects of atomistic modeling of char burnout kinetics.

Concerning the description of the reaction kinetics, the main objective is to build detailed kinetic models. To achieve this, the classes of reactions should be identified and the prediction of the reaction and activation energies of these reactions should be formulated, such that they can be used without the requirement of computationally expensive quantum chemical calculations. Hence, the *first research question* of this dissertation is "how does the size and the different active sites of polyaromatics affect the reaction/activation energies and the rate constants of the elementary reaction steps?" The polyaromatics have been chosen as representatives of char fragments. For small molecules, such as alkanes, alkenes, alkynes, benzene, toluene, etc., detailed combustion kinetic models already exist[16]. It is expected that combustion reactions start at the edges or at the defects of aromatic structural units or of the turbostratic crystallites. Therefore, it is important to study polyaromatics which are accessible with accurate quantum chemical methods and can describe the properties of the edges and defects of their larger counterparts, so that rules can be formulated to estimate kinetic data of important elementary reactions for these larger polyaromatic structural units. The sections II and III tackle this specific research question by investigating, respectively, the reactions of polyaromatics with an O atom which leads to CO release and the H abstraction reaction from polyaromatics with an H atom.

The next research question is related with the exploration of relevant elementary reaction steps. Manual investigation of combustion reactions of polyaromatics has been widely conducted[17–24]. However, such an approach is limited to chemical intuition. Therefore, automated exploration schemes should be used to reduce the bias, as well

as decreasing the amount of time required to build the detailed kinetic mechanisms. Among different approaches, reactive molecular dynamics simulations have been used widely for this purpose. For example, semi-empirical density functional theory with tight binding, specifically GFN2-xTB, has been used to simulate combustion of cyclohexane by five dioxygen molecules conserving a total spin quantum number of S=5 along the entire trajectory[25]. A more generalizable approach, which also allows simulations of combustion of more realistic char structural models due its applicability to larger scales, is the use of a reactive force field such as ReaxFF[26].

To date ReaxFF has not only been used to simulate combustion of char models[27], but has also been combined with analysis and post-processing programs to automatically explore the elementary steps of combustion reactions, for example, to build detailed kinetic models of methane combustion[28,29]. Therefore, the *second research question* of this dissertation is "how applicable are the existing tools for the analysis of ReaxFF trajectories of large polyaromatics in order to explore the elementary reaction steps of combustion and, if not easily applicable, what kind of a workflow is suitable for this purpose?" By answering this question, we establish the first step of building detailed reaction networks of char combustion. Section IV presents the work done to answer this question using a modified version of the ReaxFF trajectory analysis and post-processing program ChemTraYzer[28,29]. Later, we identified the points which may be improved to work more effectively for larger systems, such as not relying on Simplified Molecular Input Line Entry Specification (SMILES) representations for several steps of this analysis, and wrote a new program which follows similar ideas as ChemTraYzer while also providing a flexibility to work with larger systems. Section V presents the results of this new program and evaluates its success for identifying potential elementary reactions of char burnout.

In the following parts of this section, the approaches to char burnout modeling will be explained and background information on the methods used for the atomistic multi-scale modeling within this project will be given.

B. Multiscale and multimethod modeling of char burnout: A look at the kinetic models for combustion modeling

First mechanistic studies on combustion included combustion of light hydrocarbon fuels such as methane, ethane, ethylene, or acetylene[30]. In order to solve the differential equations describing the conversion of the fuel into energy (heat/pressure) and combustion products in a reactor or flame, using the reaction rate constants and thermodynamic variables of the detailed kinetic models, software modules have been developed. One of the first such software frameworks to solve gas-phase chemical kinetics problems has been CHEMKIN, originally developed at the Sandia National Laboratories in 1980[31]. The pivotal GRI-Mech mechanism for methane combustion has been unraveled by the Gas Research Institute in 1990s and evaluates the rate coefficient expressions in CHEMKIN format and the thermodynamic variables as NASA polynomials[32]. Afterwards many other modeling frameworks and kinetic mechanisms utilised the same format.

Liquid fuel combustion is a more complex process than gas fuel combustion because it is a multiphase problem. That is, when the liquid fuel is injected as a liquid jet, it breaks up to form droplets before evaporating[33]. Additionally, the liquid fuels such as gasoline or diesel are complex mixtures of thousands of hydrocarbon compounds. Fortunately, generally they can be simplified as a mixture of a few species. For example, gasoline can be treated as a blend of ten compounds composed of n-alkanes, iso-alkanes, mono- and dicycloalkanes, mono- and diaromatics, and naphtho-aromatics[34]; or similarly diesel can be simplified as a mixture of n-heptane, toluene and 1-hexene[35]. As the size of the molecules in the fuel increases, the number of elementary reaction steps increases. Hence, the detailed reaction mechanisms of these fuels get very large which increases the computational cost. Therefore, for these systems, the detailed kinetic mechanisms are generally reduced before they are applied in a more complex computational fluid dynamics (CFD) simulation such as for turbulent flames, using a framework such as

CHEMKIN[31], Cantera[36], FlameMaster[37], OpenSmoke++[38]. These frameworks also allow users to simulate combustion for ideal reactors, i.e., batch reactor, continuously stirred reactor (CSTR), plug flow reactor (PFR), and shock-tubes, and for laminar flames, which can be used to validate the detailed and reduced kinetic mechanisms.

Similar to liquid fuel, also solid fuel combustion is a multiphase problem. On top of that, the particles can generally not be simplified as spheres[39] and are porous. Hence, other physical phenomena such as pore diffusion, surface and pore evolution, or fragmentation become important. Additionally, there are catalytic effects due to the presence of inorganic matter within the solid fuels, i.e., within biomass[40] or coal. As a consequence, to describe the kinetic mechanisms of solid fuel combustion, different approaches have been developed.

A well-known model for solid fuel pyrolysis has been developed in 1989, the Chemical Percolation Devolatalization (CPD) model[41]. It uses percolation theory to break the weak chemical bonds (labile bridges), which then results in the release of light gases, formation of heavier molecular fragments (which forms char), or formation of tar and metaplast. During the solid fuel combustion, pyrolysis is followed by char burnout and gas phase reactions. To predict the conversion during char burnout and ultimate fly ash carbon content, the Carbon Burnout Kinetic Model (CBK) has been developed in 1998. It integrates a global char combustion mechanism[42] (i.e., $C + O_2 \longrightarrow CO \,/\, CO_2$) with char deactivation submodels[43]. Particularly, this model took into account several chemical and physical phenomena such as thermal char deactivation, swelling, diameter/density changes during oxidation, and ash inhibition in the late stages of combustion[43]. In the following years, in order to describe the changing kinetic behaviour at the high temperature regime ($T > 110\,K$), a three step semi-global mechanism has been implemented instead and enhanced by considering two types of active sites, a and b, one of which with a higher activation energy[44]. It is called extended Carbon Burnout Kinetic (CBK/E) model and includes the reactions 1-3. In this model, the gasification reactions are not described. Three years later, the Car-

bon Burnout Kinetics–gasification (CBK/G) model has been developed[45]. This model describes gasification in a five step mechanism (reactions 4-8).

Carbon burnout reactions:

$$\mathrm{C} + \mathrm{O_2} \xrightarrow{1} 2\,\mathrm{C(O)} \qquad \textit{on type a,b sites}$$

$$\mathrm{C(O)} + \mathrm{O_2} \xrightarrow{2} \mathrm{CO_2/CO} + \mathrm{C(O)}$$

$$\mathrm{C(O)} \xrightarrow{3} \mathrm{CO/CO_2} \quad \textit{on type a',b' sites}$$

Gasification reactions:

$$\mathrm{C} + \mathrm{CO_2} \xrightarrow{4} \mathrm{C(O)} + \mathrm{CO}$$

$$\mathrm{C(O)} \xrightarrow{5} \mathrm{CO}$$

$$\mathrm{C} + \mathrm{H_2O} \xrightarrow{6} \mathrm{C(O)} + \mathrm{H_2}$$

$$\mathrm{C(O)} \xrightarrow{7} \mathrm{CO}$$

$$\mathrm{C} + 2\,\mathrm{H_2} \xrightarrow{8} \mathrm{CH_4}$$

Recently, the CPD, CBK/E and CBK/G models are combined under a comprehensive char conversion code (CCK)[46] and extended to the oxy-fuel combustion atmospheres via CCK/oxy model[47,48]. The strength of these models is that they include most of the relevant physical-chemical phenomena involved for carbon conversion, i.e., thermal annealing, film diffusion, pore diffusion, particle swelling[47]. However, the reaction kinetics data still need to be fitted to the experimental data for each new type of solid fuel.

In order to build a predictive kinetic model which can describe any solid fuel with less experimental fitting of the kinetic parameters, a kinetic model which takes into account structural features of the solid fuel is needed. Another approach to model solid fuel combustion addresses this aspect and uses a similar methodology as explained for liquid fuel combustion. It has been developed at Politecnico di Milano (PoliMi) and is called CRECK modeling. In 1999 a nuclear magnetic resonance (NMR) study has shown that the structural parameters can be estimated from ultimate and proximate analysis

of coal, which are more easily obtainable than performing NMR measurements, with a remark that there might be exceptions[49]. Based on this information, the CRECK model makes use of proximate composition and ultimate elemental analysis to form a mix of model fuels which represents coal[50], biomass[51], or algae[52]. The CRECK model also uses the CHEMKIN format, and therefore, is compatible with many of the available modeling frameworks. In order to use it in a more complex CFD simulation, the detailed kinetic mechanisms have been simplified with mechanism reductions[53,54]. Alternatively, supervised machine learning has been trained with the output of the CRECK model and shown to be predictive with a cheaper computational cost which allows large eddy simulations to model turbulent flames[55]. Unfortunately, this model has a limited description of char structure and does not describe pore evolution.

Neither the CCK nor the CRECK conversion kinetic models take into account important physico-chemical processes such as tar evolution, S and N release, or catalytic effects. The goal of CRC/TRR 129 OxyFlame is to develop an accurate simulation framework (OxySim-129) which tackles all of the relevant processes for solid fuel combustion[56]. Such a model will provide more realistic predictions, and hence, more reliable estimates for the design of burners in oxy-fuel power plants.

C. Multiscale and multimethod modeling of char burnout: The role of computational chemistry

In combustion modeling, the mathematical models used in computational fluid dynamics require accurate thermochemical data, such as heat capacity, enthalpy, equilibrium constants, and chemical reaction rates. Computational chemistry helps providing such accurate thermochemical data, especially for the species which are not easily measurable through experiments such as transition states or reaction intermediates with short lifetimes. Such accurate data can be obtained for very small molecules using composite methods, which can be considered as first multi-method approaches in computational

chemistry, such as the Gaussian-n methods (G1, G2, G3, G4) or the HEAT protocol explained in the section I C 5. However, although the prediction of accurate thermochemical data remains as a goal, the multimethod computational approach needed in this project originates from the fact that solid fuel combustion is spatially and temporally a multiscale process.

Considering the industrial scale combustion of solid fuels, there is the scale of the furnace, the scale of the solid fuel particles, and the scale of the pores, i.e., macro-, meso-, micro-pores, etc., within these particles. The particles can further be resolved at the atomistic scale. There are also different time scales, such as the diffusion rates through different sized pores or the rates of different elementary reaction steps. Therefore, to describe this multiscale problem, a multiscale model is required, which can only be achieved with a multimethod approach. Using computational chemistry tools, it is possible to build molecular structures and study their physico-chemical properties. Computational chemistry methods, based on quantum mechanics and molecular mechanics, resolve the atomistic scales to predict the reaction mechanisms, reaction rates, and adsorption and diffusion processes.

Figure 2 shows the time and length scales of the different computational chemistry methods used in this project in order to resolve combustion processes at the atomistic scale. This hierarchy forms the basis of the multiscale multimethod aproach followed in this project. The more approximations are introduced to a method, the less computationally expensive it becomes, making larger systems and time scales accessible. In general, the quantum mechanical methods (see section I C 1) resolve the electronic structure through approximate solutions to electronic Schrödinger equation of many electron systems, whereas the reactive force field ReaxFF (see section I C 2) uses parametric functions of nuclear coordinates to obtain the energy of the system and propagate the reacting molecules in space and time. The semiempirical methods such as density functional tight binding (DFTB) can resolve the electronic structure at a lower computational cost compared with density functional theory (DFT) (also

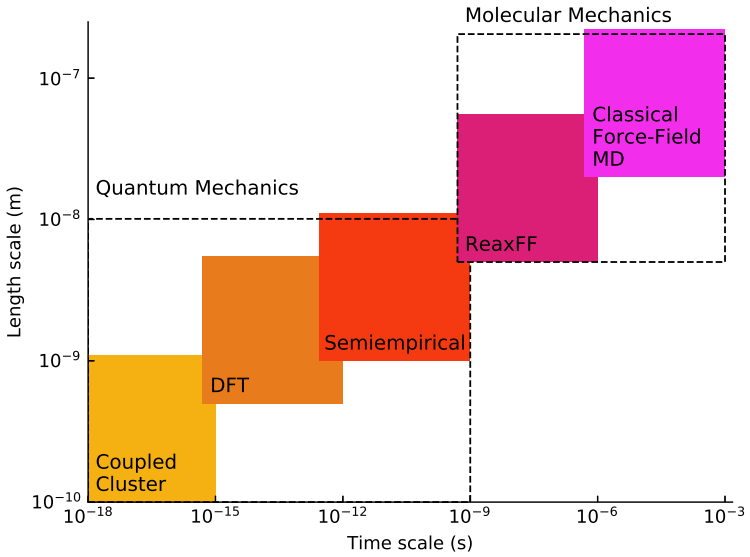


FIG. 2. The computational chemistry methods used for combustion modeling and the scales they cover.

see section I C 1), and hence, builds a bridge between ReaxFF and DFT. The coupled cluster (CC) theory provides very accurate energies when static correlation is small.

The scales connecting different methods should be linked based on their individual strengths. Figure 3 shows an outline of the computational scheme followed in this dissertation. ReaxFF can be used to identify the reaction sites on char structural models and to discover the elementary reaction steps of its combustion. For the calculation of the reaction rate constants of elementary reaction steps (see section I C 4) and the thermochemical properties of the molecules, accurate electronic energies and partition functions (see section I C 3) are needed.

In order to obtain accurate electronic energies, the coupled cluster singles and doubles with a perturbative triple excitation correction, CCSD(T), method, which is considered as the gold-standard in quantum chemistry[57], can be used. However, even with the local pair natural orbital based CCSD(T) implementation, systems with at most

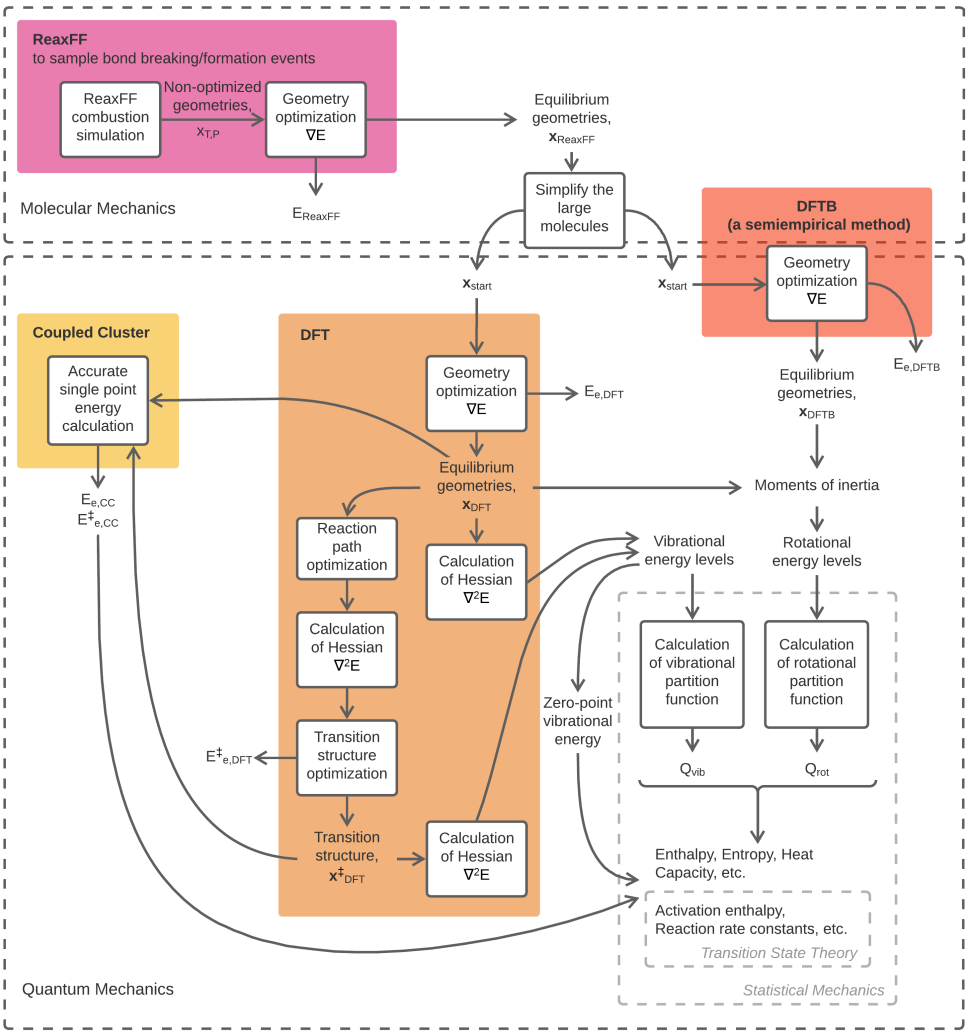


FIG. 3. Outline of the approximate computational scheme followed for the development of methods and models for char combustion reactions in this dissertation.

around 30 heavy atoms have been studied routinely[58]. Fortunately, most of the time, the Bell-Evans-Polanyi principle can be used to estimate activation energies given the reaction energies and the reaction energies can be estimated by generalizations made based on the chemical nature of the elementary reaction step. For example, in section III we have shown that for reactions which have a local character such as H-abstraction with an H-atom, the reaction energies are dominated by the local en-

vironment of the reactive site; hence, reaction energies of small model polyaromatics provide direct estimates for their larger counterparts. As another example, in section II we have shown that for reaction steps such as CO desorption or O atom addition to a polyaromatic, the reaction energies are affected by rearrangement of π-electrons; therefore, the stabilization energies calculated with Hückel Molecular Orbital theory (HMOT) can be correlated with the calculated reaction energies. As a consequence, the reaction energies of larger systems can be estimated using these correlations. In section III we have also demonstrated a multimethod approach to link coupled cluster theory, DFT, and DFTB to obtain reaction energies.

In addition to allowing extrapolations to larger length and time scales, multi-method or composite approaches are also used to obtain, for a given reaction, the thermochemical data with a good cost/accuracy balance in all steps. Accurate electronic energies require sufficiently accurate structures for the stationary points. That's why it is more desirable to use a method which resolves electronic structure when the geometries of stationary points are needed. Among the quantum chemistry methods, the practicable ones are the DFT and the semiempirical methods such as DFTB, whereas the methods such as CCSD(T) are only used to obtain accurate electronic energies due to their high computational cost. Both DFT and DFTB provide good equilibrium geometries.

Additionally, the Hessian matrix (force constants) is needed for the optimization of transition structures (saddle points) and to confirm the nature (curvature) of stationary points found in geometry optimizations. It is also needed to calculate vibrational energy contributions and vibrational partition functions within the harmonic approximation. The calculation of the gradient or the Hessian matrix would not be practicable at the CCSD(T) level except for molecules with only a few heavy atoms due to high computational cost. Therefore, the method of choice for the calculation of the Hessian matrix is in general DFT.

The rest of this subsection explains the concepts of the theories and methods used in computational chemistry relevant to this project and explains more in detail the

reasons for the particular method selections.

1. Quantum chemistry

The quantum chemical methods are based on the solution of the Schrödinger equation; therefore, they can resolve the electronic structure of the molecular systems. The electrons cannot be properly treated by classical mechanics because of wave–particle duality which results in the Heisenberg's uncertainty principle

$$\Delta p \Delta x \geq \frac{1}{2}\hbar \tag{1}$$

where $\hbar$ is the Planck constant, p is the momentum, and x is the position of the particle. An electron has a very small mass ($m_e \approx 9.109 \times 10^{-31} kg$) which results in a very small Δp; therefore, the uncertainty in their position gets large. For that reason, the position of electrons can only be described as a probability density function ($|\psi|^2$), where ψ is the wavefunction. In order to find ψ, the Schrödinger equation[a] should be solved. The first postulate of quantum mechanics states that to any well-defined observable in physics, such as linear momentum, angular momentum, energy, etc., there corresponds an operator such that measurement of the observable yields values which are eigenvalues of that operator. The *time-independent* Schrödinger equation is basically this postulate written for the energy operator, or Hamiltonian,

$$\hat{H}\psi = E\psi \tag{2}$$

where E is the total energy of the system and the Hamiltonian operator is defined as $\hat{H} = \hat{T} + \hat{V}$, where $\hat{T} = \hat{p}^2/2m$ is the kinetic energy operator, $\hat{p} = -i\hbar\nabla$ is the momentum operator, and $\hat{V}$ is the potential energy operator.

[a] The Schrödinger equation is valid only with the assumption that the electrons can be treated non-relativistically; that is, the velocity of the electrons is assumed to be small relative to the speed of light, $v \ll c$, so that the mass of the particle, $m = \frac{m_0}{\sqrt{1-v^2/c^2}}$, is almost constant. Otherwise, the relativistic Dirac equation should be solved instead.

A molecule is composed of nuclei and electrons. If there is no external field, the potential energy operator of a molecule is time-independent and is only a function of the position of the nuclei ($\mathbf{R}$) and electrons ($\mathbf{r}$), i.e., $\hat{V}(\mathbf{R},\mathbf{r},t)$=$\hat{V}(\mathbf{R},\mathbf{r})$. Therefore, the Hamiltonian is time-independent and the time dependence of ψ can be separated as a phase factor, i.e., $\psi(\mathbf{R},\mathbf{r},t) = \psi(\mathbf{R},\mathbf{r})e^{-iEt}$. Hence, the time-independent Schrödinger equation[b] can be used to obtain energy levels of the molecule, which is

$$\hat{H}_{tot}\psi_{tot}(\mathbf{R},\mathbf{r}) = E_{tot}\psi_{tot}(\mathbf{R},\mathbf{r}) \tag{3}$$

where the total Hamiltonian is defined as

$$\hat{H}_{tot} = \underbrace{-\sum_A \frac{\hbar^2}{2m_A}\nabla_A^2}_{\hat{T}_N} \underbrace{-\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2}_{\hat{T}_e} \underbrace{+\sum_{A<B} \frac{Z_A Z_B e^2}{4\pi\varepsilon_0 R_{AB}}}_{\hat{V}_{NN}} \underbrace{-\sum_{iA} \frac{Z_A e^2}{4\pi\varepsilon_0 r_{i,A}}}_{\hat{V}_{eN}} \underbrace{+\sum_{i<j} \frac{e^2}{4\pi\varepsilon_0 r_{ij}}}_{\hat{V}_{ee}} \tag{4}$$

where the indices A and B run over the nuclei, and i and j run over the electrons, $\hat{T}_N$ is the kinetic energy operator of the nuclei, $\hat{T}_e$ is the kinetic energy operator of the electrons, $\hat{V}_{NN}$, $\hat{V}_{eN}$ and $\hat{V}_{NN}$ are the potential energy operators due to Coulomb interactions, m_A and m_B are the masses of the nuclei A and B respectively, ε_0 is the electric constant, e is the elementary positive charge, Z is the charge number. In quantum chemistry programs the Hamiltonian operator is handled in atomic units instead, so that the equation is free of physical constants:

$$\hat{H}_{tot} = -\sum_A \frac{\nabla_A^2}{2m_A} - \sum_i \frac{\nabla_i^2}{2} + \sum_{A<B} \frac{Z_A Z_B}{R_{AB}} - \sum_{iA} \frac{Z_A}{r_{i,A}} + \sum_{i<j} \frac{1}{r_{ij}} \tag{5}$$

Away from conical intersections, it can be assumed that the nuclei are stationary relative to electrons. In other words, from the perspective of nuclei, the electrons immediately adjust to the changes of the nuclear coordinates. This approximation is

[b] The time development of the wavefunction $\psi(\mathbf{r},t)$ is described with the time-dependent Schrödinger equation $\hat{H}\psi(\mathbf{r},t) = i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t)$. When there is a time-dependent perturbation to the system such as an oscillating electric field, the potential field is time dependent; therefore the time-dependent Schrödinger equation should be solved instead, which is not an eigenvalue problem anymore. This is needed, for example, to describe nonlinear optical properties, such as hyperpolarizabilities.

called adiabatic or Born-Oppenheimer approximation[59]. Therefore, the motion of the electrons can be separated from the motion of the nuclei. Hence, the electronic wave function can be expressed as parametric functions of the nuclear positions

$$\psi(\mathbf{R}, \mathbf{r}) = \psi_N(\mathbf{R})\psi_e(\mathbf{r}; \mathbf{R}). \tag{6}$$

Then, the electronic Schrödinger equation is separated as

$$\hat{H}_e\psi_e(\mathbf{r}; \mathbf{R}) = E_e(\mathbf{R})\psi_e(\mathbf{r}; \mathbf{R}) \tag{7}$$

where the electronic Hamiltonian is

$$\hat{H}_e = \hat{T}_e + \hat{V}_{eN} + \hat{V}_{ee} + \hat{V}_{NN} \tag{8}$$

and the nuclear Hamiltonian is separated as

$$\hat{H}_N = \hat{T}_N + E_e(\mathbf{R}). \tag{9}$$

The analytical solution to the equation 7 is only possible for one-electron systems such as the H atom or $H_2{}^+$ because of the Coulomb interaction between the electrons.
The motion of the electrons is correlated to each other. Firstly, the electrons are fermions; therefore, according to the Pauli exclusion principle two or more electrons cannot have identical quantum states in a quantum system, i.e., the many-electron wavefunction should be anti-symmetric with respect to the exchange of two electrons. This results in Fermi correlation. The simplest form of many electron wavefunction which fulfills Pauli exclusion principle is a Slater determinant, which is anti-symmetric with respect to permutation of two electrons.
The Hartree Fock (HF) method determines the optimal Slater determinant ψ^{HF} expanded in a basis set of functions, such as atom centered functions, by the variational principle using the Self-Consistent Field (SCF) procedure. The orthonormality condition ($\langle\psi_i|\psi_j\rangle = \delta_{ij}$ where δ_{ij} is the Kronecker delta) for the orbitals can be enforced by Lagrange multipliers λ_{ij}:

$$\mathcal{L} = E_{\mathrm{HF}} - \sum_i \lambda_{ij} \left(\langle\psi_i|\psi_j\rangle - \delta_{ij}\right) \tag{10}$$

where $E_{\mathrm{HF}} = \frac{\langle \psi^{\mathrm{HF}} | \hat{H}_e | \psi^{\mathrm{HF}} \rangle}{\langle \psi^{\mathrm{HF}} | \psi^{\mathrm{HF}} \rangle} = \sum_i \langle \psi_i | - \frac{\nabla_i^2}{2} - \sum_A \frac{Z_A}{r_{i,A}} | \psi_i \rangle + \frac{1}{2} \sum_{i<j} \langle \psi_i \psi_j | \frac{1}{r_{ij}} | \psi_i \psi_j \rangle$ is the expectation value of the energy of ψ^{HF}. Applying the variational principle ($\delta \mathcal{L} = 0$) leads to the Hartree Fock equation:

$$\hat{F} \, |\psi_i\rangle = \epsilon_{ij} \, |\psi_j\rangle \, . \tag{11}$$

The Fock operator $\hat{F}$ is defined as

$$\hat{F} = \sum_i \Big(\hat{h}(i) + \hat{j}(i) - \hat{k}(i) \Big) \tag{12}$$

where $\hat{h}(i) = -\frac{\nabla_i^2}{2} - \sum_A \frac{Z_A}{r_{i,A}}$ evaluates the one-electron integrals. The *Coulomb operator* $\hat{j}(i)$ and the *exchange operator* $\hat{k}(i)$ (together also called *mean field operator*) account for the two electron interactions in a mean field of other electrons. In practice *canonical molecular orbitals* which diagonilize the Fock matrix are usually used. The canonical Hartree Fock equation reduces to an eigenvalue equation:

$$\hat{F} \, |\psi_i\rangle = \epsilon_i \, |\psi_i\rangle \tag{13}$$

where ϵ_i are the energies of the canonical orbitals.

The other type of electron correlation is the Coulomb correlation due to the Coulombic interaction between the electrons. Following Löwdin[60], the *electron correlation* E_{corr} is defined as

$$E_{corr} = E_e - E_{\mathrm{HF}} \tag{14}$$

where E_{HF} is the Hartree Fock energy. Therefore, the post-HF wavefunction methods are also called electron correlation methods. E_{corr} actually refers only to the Coulomb correlation. The Coulomb correlation can further be divided into *static correlation*, which is due to near degeneracies of the Slater determinants, and *dynamic correlation*, which is due to the instantaneous repulsion of electrons trying to occupy the same region in space.

The dynamic correlation can be taken into account by methods which use a trial wavefunction which uses more than one Slater determinant. Configuration Interaction

(CI) is one of the oldest electron correlation methods which is like Hartree-Fock a variational method. The CI wavefunction ansatz is linear combination of determinants with expansion coefficients c_i:

$$\psi^{\mathrm{CI}} = \sum_{i=0} c_i \psi^{CSF} \tag{15}$$

where the configuration state functions ψ^{CSF} are excited singly, doubly, ... for $i = 1, 2, ...$ relative to the configuration of ψ^{HF}. This method is impracticable without truncation except for systems with few electrons. The truncated CI method is not size consistent and the accuracy decreases with increasing number of electrons (i.e., not size extensive). Quadratic configuration interaction (QCI)[61] corrects the CI equations to solve the size consistency problem; however, it cannot be extended to hierarchy that would converge systematically to full CI. The coupled cluster (CC) method[62,63] uses the wavefunction ansatz

$$\psi^{\mathrm{CC}} = e^{\hat{\mathcal{T}}} \psi^{\mathrm{HF}} \tag{16}$$

where $\hat{\mathcal{T}} = \sum_\mu t_\mu \hat{\tau}$ is a linear combination of excitation operators $\hat{\tau}$ which excite electrons from one orbital to another. Even if the operator $\hat{\mathcal{T}}$ is truncated, the CC method is size consistent and size extensive.

For bond breaking events or molecules with near-degenerate electronic states, static correlation also becomes important. Multireference methods such as complete active-space second-order perturbation theory (CASPT2) or multireference CI (MRCI) can be used to also account for the static correlation for these cases.

For complete active-space calculations, the wavefunction is contructed as a full CI wavefunction on a predefined orbital subspace, which is the *active space*. However, the selection of the active space is a non-trivial task which requires expertise. In the recent years, there have been developments for the automatic selection of the active space, such as the autoCAS code[64]. However, autoCAS has still not been fully automated for reactions, which would require consistent selection of the active space along the reaction coordinate[64]. Additionally, these calculations come with a high computational

cost. Therefore, in this dissertation we leave out the multireference calculations and use CCSD(T) to obtain accurate electronic energies which covers most of the dynamic correlation.

In order to decrease the cost of calculation with the wavefunction based methods, a further approximation can be made. When canonical orbitals are used, the correlation between all of the electrons are considered. The local correlation methods exploit that dynamic electron correlation is a short ranged effect decaying with the distance of the interacting electron as r^{-6}. Pair natural orbitals (PNO)[58,65], cluster-in-molecules[66], divide expand and consolidate (DEC)[67] are some examples. In section III, we have used PNOs to calculate energies of larger polyaromatics, the largest being coronene ($C_{24}H_{12}$).

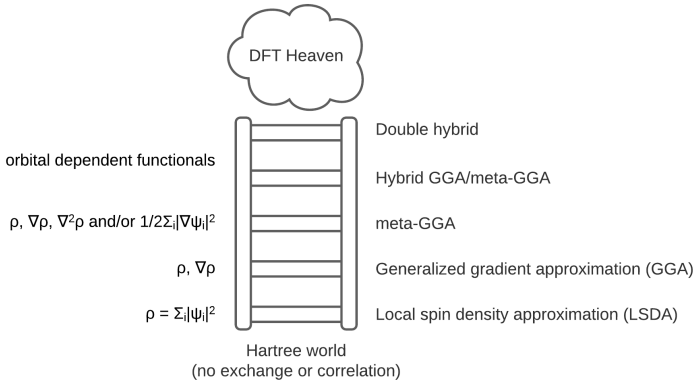


FIG. 4. Perdew's Jacob's ladder[68,69] composed of five rungs with increasingly sophisticated models for the exchange-correlation functional of the DFT.

Apart from the wavefunction based methods, Hohenberg and Kohn[70] proved that the electronic energy can completely be described by the electron density ($\rho = \sum_i |\psi_i|^2$). This is the basis of the density functional theory (DFT). However, with this formulation the exact exchange-correlation functional is not known; therefore, it is approximated by ab-initio or semiempirical functionals. Even if the DFT functionals do not systematically improve the description of the correlation energy, Perdew[68] presented a

DFT hierarchy called Jacob's Ladder (see Figure 4). Within this dissertation, we have used the meta-GGA functional TPSS as the initial step and then refined the calculations with its hybrid version TPSSh. The use of the hybrid meta-GGA functional is a balanced choice considering the trade-off between accuracy and computational cost.

Another set of methods are the semi-empirical quantum chemistry methods. One of the first is the Hückel molecular orbital theory proposed in 1931[71]. It treats only the π-electrons and generates MOs from the connectivity matrix of a molecule. This method can even be applied by pen and paper and provides chemical insights on the stability of the molecules. In 1963 Hoffmann developed the extended Hückel method[72], which treats only the valence electrons and parameterizes the Fock matrix elements with atomic orbital energies. Nowadays, this method is generally used to generate start orbitals for further QM calculations.

More recent semi-empirical quantum chemistry methods aim to build a bridge between density functional theory and force field methods by reducing the computational costs to be able to handle several thousands of atoms but still explicitly calculate the electronic structure. These methods reduce the computational cost of the evaluation of the two-electron integrals in the Fock matrix by either neglecting some of them or employing approximations to them. A common approximation is the zero-density-overlap (ZDO) approximation which ignores the overlap of basis functions if the basis functions are not on the same atom; hence, the three- and four-center two-electron integrals are neglected and these approximations are compensated using parameterized expressions for the remaining integrals[73]. For example, with increasing level of approximation the Neglect of Diatomic Differential Overlap (NDDO), Intermediate Neglect of Differential Overlap (INDO), Complete Neglect of Differential Overlap (CNDO) aprroximations are based on the ZDO approximation[73,74].

Conceptually different semi-empirical methods are the density functional tight binding (DFTB) methods[75]. They employ approximations to the Hamiltonian matrix elements by analytical functions. GFN-xTB is one of them which is parameterized globally and

consistently works through almost all elements of the periodic table[76]. This method has been parameterized for geometry optimizations, computations of vibrational frequencies, and non-covalent interaction energies. The GFN2-xTB variant improves the description of the non-covalent interactions and conformational energies[77].

2. Molecular mechanics

If the Born-Oppenheimer approximation is applicable for the system under study[c], the solution of the equation 7 provides an adiabatic potential energy surface, on which the motion of the nuclei can be obtained either approximately by Newtonian mechanics or by more accurately by quantum mechanics. However, most of the time the nuclei are heavy enough to be treated as classical particles[d]. Classical molecular dynamics reveal the simultaneous motion of the atoms and electrons based on Newtonian's equations of motion making use of analytical functions of nuclear coordinates, .e., a force field (FF), under these two assumptions.

The energy of a classical force field, which assumes that the electronic structure remains essentially unchanged, i.e., no bonds are broken or formed, is generally defined as

$$E_{FF} = E_{str} + E_{bend} + E_{tors} + E_{vdw} + E_{el} + E_{cross} \tag{17}$$

where the energy contributions are due to, respectively, the stretching of bonds between atoms, the bending of an angle, the torsional rotation around a bond, the van der Waals interactions, the electrostatic interaction, and the coupling term between E_{str}, E_{bend}, and E_{tors}[73]. The parameters can be obtained experimentally, from quantum mechanical calculations, or from a combination of both. In practical applications, once the initial conditions such as temperature and pressure are chosen, the initial

[c] The Born-Oppenheimer approximation breaks down for some systems. For example, coupling between electrons and nuclei may form a conical intersection between PESs[78], which is a non-adiabatic process important in photochemistry.

[d] There are exceptions to this assumption. For example, the motion of hydrogen atom, which is prone to tunneling due to its relatively small mass, can be described more accurately by quantum mechanics.

velocities and positions of the molecules in a simulation box can be created based on a statistical distribution. Then, the Newtonian equations of motion are applied to propagate in time.

Molecular mechanics can also be used to study reactions. In general, a reaction can either be due to a conformational change without bond breaking or due to bond breaking/formation. Classical MD simulations may be used to study conformational changes, such as of a protein[79]. However, because the bond stretching term E_{str} is generally evaluated as harmonic oscillator in the classical MD, the bonds cannot be broken. When bond breaking is important two approaches can be followed: Refining electronic structure during the simulation or moving from classical to reactive force fields. When the reactive sites are not large, hybrid quantum mechanics/molecular mechanics (QM/MM) approaches can be used. For that reason, they are generally applied in mechanistic studies of enzymatic reactions[80]. Ab initio molecular dynamics (AIMD) simulations, in which the nuclei propagate according to Newtonian mechanics while the forces acting on them are calculated on–the–fly quantum mechanically, can also be used to study chemical reactions, especially combined with acceleration schemes. However, these are still limited to small systems of around 100 heavy atoms, for example, to model polymerization of acetylene into larger molecules[81]. A similar approach can be applied with semi-empirical quantum chemistry methods with simple reactions where reasonable guesses for the reaction coordinates are known beforehand. DFTB/MD simulations has been applied to a system of around 200 carbons to study fullerene formation during benzene combustion[82]. Another such example is the prototypical study of oxidation of a single cyclohexane with 5 oxygen molecules using a GFN2-xTB nanoreactor[25]. The total computation time of this calculation was 12 min on a quad-core laptop computer[25]. That is, the method can be applied to larger systems as well. However, at the size- and time- scales and number of reactions that is needed for the current project, its use would have had to be restricted to a few cases.

Apart from these methods which refine the electronic structure by a quantum mechan-

ical treatment of the system, the bond breaking can also be modeled by not assuming harmonic stretching. An intuitive approach could be to use an anharmonic potential as a function of interatomic distances for bond stretches, such as the Morse potential[83]. However, for long stretches the restoring force becomes so small that distorted structures display slow convergence to an equilibrium bond length[73]. Additionally, for a reactive force field other concepts such as connectivity, pre-defined hybridization and atom types, etc. need to be generalized. Therefore, another approach is more common, which is to describe the stretching energy as a function of bond order instead of bond length, one of first is the Brenner potential[84]. Nowadays, the reactive force field ReaxFF is being used widely because apart from calculating bond energies based on bond orders, it also includes van der Waals and Coulomb potentials for non-bonded interactions and they have shown accuracy similar to semi-empirical methods with a much lower computational cost[26]. As long as an appropriate parameterization exists, ReaxFF can be applied to various systems consisting of thousands of atoms. Relevant to char combustion are the C/N/B/S/O/H parameterization which, for example, has been used to investigate heat transfer from the heated soot particles to the surrounding gas[85], or to simulate combustion of an Illionis No. 6 coal char model[27], and the C/H/O parameterization which, for example, has been used to study methane/O_2, o-xylene/O_2, propene/O_2, and benzene/O_2 mixtures[86]. In sections IV and V we have used the latter in order to generate a reaction pool for combustion of polyaromatics as small char fragment models, which are later refined with quantum mechanical methods.

3. Statistical mechanics

Statistical mechanics build a connection between the quantum mechanical energy levels ϵ_i of a system and its thermodynamic functions through partition functions. The partition function of an ensemble in which the number of molecules N, the volume V, and

the temperature T are fixed (canonical ensemble), the partition function $Q(N, V, T)$ is defined as

$$Q(N, V, T) = \sum_i e^{-\epsilon_i(N,V)/k_B T}. \tag{18}$$

The Boltzmann factor, $e^{-\epsilon_i(N,V)/k_B T}$, approximates the fraction of particles in a large system, where k_B is the Boltzmann's constant. The partition function of a system with N non-interacting identical particles can be written in terms of the partition functions of the individual components (i.e., individual atomic or molecular partition functions)

$$Q(N, V, T) = \frac{[q(V, T)^N]}{N!} \tag{19}$$

and the partition function of each individual components can approximately be separated in terms of the degrees of freedom the system has. For a mono-atomic gas, these are the translational q_{trans}, electronic q_{el}, and nuclear q_{nucl} degrees of freedom. The nuclear state generally does not change in chemical processes. Therefore, the nuclear contribution is generally neglected, i.e., $q_{nucl} = 1$. Then, the partition function of a mono-atomic gas is

$$q(V, T) = q_{trans} q_{el}. \tag{20}$$

The translational motion can be considered as the particle in a box problem. The Hamiltonian $\hat{H}$ and the energy eigenvalues of a particle in a 1-D box is given in Table I. Similarly, for a 3-D box, the energy levels of the particle become

$$\epsilon_{n_x,n_y,n_z} = \frac{\pi^2 \hbar^2}{2ma^2}(n_x^2 + n_y^2 + n_z^2), \quad n_x, n_y, n_z = 1, 2, 3, ... \tag{21}$$

Then, the translational partition function can be written combining equations 19 and 21 and replacing the summation by an integral because the successive terms of ϵ_{n_x,n_y,n_z} differ so little:

$$\begin{aligned} q_{trans} &= \sum_{n_x,n_y,n_z=1}^{\infty} e^{-\epsilon_{n_x,n_y,n_z}/k_B T} = \left(\sum_{n=1}^{\infty} e^{-\pi^2 \hbar^2 n^2/2ma^2 k_B T} \right)^3 \\ &\approx \left(\int_0^{\infty} e^{-\pi^2 \hbar^2 n^2/2ma^2 k_B T} dn \right)^3 = \left(\frac{m k_B T}{2\pi \hbar^2} \right)^{3/2} V \end{aligned} \tag{22}$$

where $V = a^3$ is the volume. Similarly, the electronic partition function can be written as

$$q_{el} = \sum_i g_i e^{-\epsilon_i / k_B T} \tag{23}$$

where g_i is the degeneracy of the ith electronic level.

TABLE I. The solutions to the Schrödinger equation of ideal quantum mechanical systems.

	Particle in a 1-D box [a]	Harmonic Oscillator [b]	2D-Rigid Rotor [c]
$\hat{H}$	$-\frac{\hbar}{2m}\frac{\partial^2}{\partial x^2}$	$-\frac{\hbar}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2}kx^2$	$-\frac{\hbar}{2I}\left[\frac{1}{sin\theta}\frac{\partial}{\partial\theta}\left(sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]$
Energy levels	$\frac{\pi^2\hbar^2 n^2}{2ma^2}, \quad n = 1, 2, 3, ...$	$\left(n + \frac{1}{2}\right)\hbar\omega, \quad n = 0, 1, 2, ...$	$\frac{J(J+1)\hbar^2}{2I}, \quad J = 0, 1, 2, ...$

[a] a is the width of the box.

[b] k is the spring constant and $\omega = (k/m)^{1/2}$.

[c] I is the moment of inertia and the Laplacian term ∇^2 of $\hat{H}$ is given in spherical coordinates.

For a polyatomic gas there are also rotational and vibrational degrees of freedoms:

$$q(V, T) = q_{trans} q_{rot} q_{vib} q_{el}. \tag{24}$$

These can be approximated as rigid rotor and harmonic oscillator (RRHO) whose Hamiltonian and energy eigenvalues are given also in Table I. For a diatomic, there is only one harmonic oscillator and the vibrational partition function can be found exactly by summation as a geometric series:

$$q_{vib} = \sum_{n=0}^{\infty} e^{-\frac{(n+1/2)\hbar\omega}{k_B T}} = e^{-\frac{\hbar\omega}{2k_B T}} \sum_{n=0}^{\infty} e^{-\frac{\hbar\omega n}{k_B T}} = \frac{e^{-\frac{\hbar\omega}{2k_B T}}}{1 - e^{-\frac{\hbar\omega}{k_B T}}}. \tag{25}$$

For a polyatomic gas, there are for linear molecules $\alpha = 3N - 5$ and for non-linear molecules $\alpha = 3N - 6$ normal modes of vibrations. If these vibrations are assumed to be independent harmonic oscillators, the vibrational partition function of a polyatomic molecule becomes

$$q_{vib} = \prod_{i=1}^{\alpha} \frac{e^{-\frac{\hbar\omega_i}{2k_B T}}}{1 - e^{-\frac{\hbar\omega_i}{k_B T}}} \tag{26}$$

For the rotational degrees of freedom, solution to the rigid rotor approximation is used. The summation at the rotational function can be approximated as an integral at high enough temperatures. For a diatomic, the degeneracy of the states is $2J+1$, then the partition function becomes

$$q_{rot} = \sum_{J=0}^{\infty}(2J+1)e^{-\frac{\hbar^2}{2Ik_BT}J(J+1)} \approx \int_0^{\infty} e^{-\frac{\hbar^2}{2Ik_BT}J(J+1)}d[J(J+1)] = \frac{2Ik_BT}{\hbar^2} = \frac{T}{\Theta_r} \quad (27)$$

where Θ_r is called the characteristic temperature of rotation. For a polyatomic molecule, symmetry number σ is introduced and the degeneracy of the states is $(2J+1)^2$. Then, the generalized rotational partition function for an asymmetric top is

$$q_{rot} = \frac{\pi^{1/2}}{\sigma}\left(\frac{T^3}{\Theta_A\Theta_B\Theta_C}\right)^{1/2} \quad (28)$$

where Θ_A, Θ_B, Θ_C are the rotational temperatures which correspond to the three principle moments of inertia I_A, I_B, I_C.

Once the partition function of a molecule is constructed, the thermodynamic variables can be evaluated. For example, in the canonical ensemble, the thermodynamic connection to entropy S is made by the formula[87]

$$S = k_B \ln Q + k_B T\left(\frac{\partial \ln Q}{\partial T}\right)_{N,V}. \quad (29)$$

The RRHO approximation is most of the time reasonable; however, the harmonic oscillator approximation becomes especially erroneous for low lying torsional modes. This can easily be seen from the mathematical expression of the contribution of the normal mode vibrations to the entropy S_{vib}, combining the equations 26 and 29:

$$S_{vib} = \sum_{i=1}^{\alpha} k_B T\left[\frac{\hbar\omega_i}{k_B\left(e^{-\frac{\hbar\omega_i}{k_BT}}-1\right)} - \ln\left(1-e^{-\frac{\hbar\omega_i}{k_BT}}\right)\right]. \quad (30)$$

As ω_i approaches zero, the logaritmic term in S_{vib} approaches infinity. In order to correct this error, the normal modes which correspond to hindered rotors can be

identified[88] and corrected, e.g. as 1-D hindered rotors[89–91]. A more simple, although more approximate, solution has been proposed and applied by Grimme to treat the low-lying vibrational modes by a free-rotor approximation[92]. In particular, they replace the contribution of the low-lying modes to the entropy by a rotational entropy using a weighting function of ω. This approximation is called quasi-rigid rotor harmonic oscillator (qRRHO) approximation. In the calculation of the rate constants presented in sections III and V, we have used this approach by applying the weighting factor to the partition functions instead of entropy. We have chosen to apply qRRHO approximation, because the explicit treatment of hindered-rotor motions is not as straightforward. At the reactive sites or around the large polyaromatics there might be more than a single hindered rotor. For example, the uncoupled hindered rotor treatment has been shown to work well for n-alkanes due to error cancellation[91]. Additionally, for example, for very shallow minima, there might still be low-lying vibrations which are not due to internal rotors. The qRRHO approximation basically cuts the error from all of the low-lying frequencies, which might even be an advantage for these cases.

4. The transition state theory

The development of thermodynamics, collision theory, and statistical mechanics led to the development of the *transition state theory* (TST) by Eyring and Polanyi in 1935[93,94]. This theory still provides an intuitive and efficient way to calculate reaction rate constants. The conceptualization of the theory is based on the potential energy surfaces.

The small region where the systems entering from the educts side pass to the products and those from the products side pass to the educts is the *dividing surface* as represented in Figure 5. Systems within this region are the *transition states*. The transition structure lies at a saddle point of index 1 on the potential energy surface of the re-

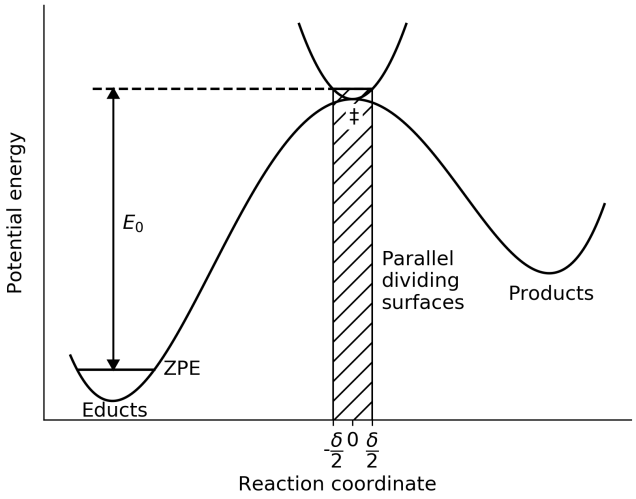


FIG. 5. Reaction coordinate profile. The saddle point at the transition state region defines the reaction path and the coordinates perpendicular to the reaction path.

acting system, i.e., the force constant matrix (Hessian) has one negative eigenvalue. The eigenvector of the negative eigenvalue indicates the direction of the reaction coordinate. Therefore, TST interprets the transition structure as a structure which has all the properties of a stable compound except in the normal mode of the reaction direction. TST assumes that this mode can be treated as a translational degree of freedom[93,95]. Based on this assumption, the partition function of the transition state per unit volume $Q^{\ddagger}_{tot}$ can be separated as

$$Q^{\ddagger}_{tot} = Q_s Q^{\ddagger} \tag{31}$$

where $Q_s = (2\pi\mu_s k_B T)^{1/2}\delta/h$ is the partition function for the translation along the reaction coordinate where k_B is the Boltzmann's constant, T the temperature, δ the separation distance between the two parallel dividing surfaces as shown in Figure 5, μ_s the reduced mass for the motion through the dividing surface, and $Q^{\ddagger}$ the partition function for all other degrees of freedom of the transition state. Another major assumption of the TST is that the educts are in quasi-equilibrium with the transition

state. That is, it is assumed that the systems which cross the transition state region in the direction of the products cannot recross in the direction of educts. Similarly, all products entering the transition state region must pass to the educts. For a reaction

$$\mathrm{A} + \mathrm{B} \longrightarrow \mathrm{X}^{\ddagger} \longrightarrow \mathrm{C}$$

if follows from statistical mechanics that the equilibrium constant for the quasi-equilibrium is approximately

$$K^{\ddagger} = \frac{[X^{\ddagger}]}{[A][B]} = \frac{Q^{\ddagger}_{tot}}{Q_A Q_B} e^{-E_0/k_B T} \tag{32}$$

where Q_A and Q_B are the partition functions of the educts per unit volume, E_0 is the difference of the zero-point energy level of the transition state from that of the educts, $[X^{\ddagger}]$, $[A]$ and $[B]$ are the concentrations of the transition states and the educts. According to the Maxwell-Boltzmann laws, the average velocity of crossing the barrier is $(2k_B T/\mu_s)^{1/2}$. Therefore, the net rate of reaction dN/dt can be written as[96]

$$\frac{dN}{dt} = \frac{[X^{\ddagger}]}{2} \left(\frac{2k_B T}{\mu_s} \right)^{1/2} \frac{1}{\delta}. \tag{33}$$

Combining the equations 31, 32, and 35, the rate of the reaction can be written as

$$\frac{dN}{dt} = \frac{k_B T}{h} \frac{Q^{\ddagger}}{Q_A Q_B} e^{-E_0/k_B T} [A][B]. \tag{34}$$

Hence, the rate constant k for the forward reaction is

$$k = \frac{k_B T}{h} \frac{Q^{\ddagger}}{Q_A Q_B} e^{-E_0/k_B T}. \tag{35}$$

Reaction rates can also be calculated following different approaches such as from molecular dynamics given a potential energy surface for the reaction. However, the accuracy of this approach depends highly on the accuracy of the approximate potential energy surface. Additionally, if tunneling is important, the above approximation becomes inaccurate. For such cases, the quantum mechanical transition state theory can be applied on small systems up to a few atoms[97], such as F+CH_4 reaction[98]. Alternatively, as a more approximate solution, the rate constant can simply be multiplied

by a correction such as the Eckart tunneling correction[99] or the approximate formula derived by Wigner for small degrees of tunneling[100]. In sections III and III we have used the classical transition state theory to calculate rate constants of elementary reactions by using the high temperature limit of the Eckart tunneling correction[101]. Above 400 K this approximation gives similar values as Eckart tunneling correction for $H+H_2$ system[101]. As char combustion takes place at high temperatures, this approximation should be reasonable.

5. Composite methods: First multimethod approaches

Composite methods can be considered as the first multimethod approaches in computational chemistry. They are used to obtain accurate thermodynamic properties at a modest computational cost but with predictive reliability. The first composite method has been developed by Pople and coworkers in late 1980s, and is denoted by Gaussian-1 theory or G1 method. It is based on the Møller-Plesset perturbation theory to obtain the geometries, Hartree-Fock theory to calculate the zero point vibrational energy corrections, includes corrections originating from the use of diffuse sp basis functions and d and f polarization functions on non-hydrogen atoms. To overcome deficiencies of Møller-Plesset perturbation theory, it uses quadratic configuration interaction with single and double excitations and a perturbative correction for triple excitations, QCISD(T), to calculate an additional correlation energy correction, and finally adds a higher order correlation defined by the number of bound unpaired valence electrons[102]. The following introduced Gn methods make improvements on that, decreasing the mean absolute deviation from experiment from 2 kcal/mol to 1 kcal/mol[103]. These methods have been used to generate thermochemical databases such as the Burcat thermochemical database which includes 300 species for combustion which is based on the G3//B3LYP method[104]. An alternative family of methods which achieve similar accuracy with a lower computational cost are the Complete Basis Set methods (CBS-

n) developed by Petersson and coworkers, which employ nonlinear pair natural orbital extrapolations to the complete basis set limit[105,106]. Composite methods which achieve a very high accuracy less than 1 kcal/mol include the Weizmann theories (Wn) and the high accuracy extrapolated ab initio thermochemistry (HEAT) protocol. Wn theory is based on coupled cluster with a full treatment of single and double excitations (CCSD) in spdf and spdfg basis sets and the CCSD with a perturbative triple excitation correction [CCSD(T)] in spd and spdf basis sets. With W1 theory, systems as large as benzene can be treated and it achieves 0.30 kcal/mol mean absolute error, whereas W2 theory uses larger basis sets and no empirical parameters to achieve 0.23 kcal/mol mean absolute error and can treat up to three heavy atoms[107]. Apart from geometry optimizations, total energy, and zero-point energy at the CCSD(T) level of theory, the HEAT theory achieves a very high accuracy by including a diagonal Born-Oppenheimer correction to account for mass dependency of the electronic structure, a correction for coupling of spin and orbital angular momentum, and a correction for scalar relativistic effects[108]. Though the composite methods follow a multimethod approach, they aim to achieve high chemical accuracy and are applicable to only small systems. Therefore, in this project they are not used; however, mentioned here as they can be considered as the first multimethod approaches in computational chemistry.

II. CO formation through reactions of polyaromatic hydrocarbons with atomic oxygen

This study includes partially the work done during my master studies, under the supervision of Prof. Dr. Christof Hättig. The O atom addition and H abstraction reaction, as well as the correlations between the stabilization energy and the reaction energies for the O atom addition step has been found during the master's project. The idea of analysing stabilization energies from the Hückel Molecular Orbital Theory belongs to Prof. Dr. Christof Hättig. During the PhD study, the subsequent CO-elimination mechanisms are analysed.

The following article is reprinted with permission from "Özlem Yönder, Christof Hättig, Density functional theory study of CO formation through reactions of polycyclic aromatic hydrocarbons with atomic oxygen ($O(^3P)$), Fuel, Volume 241, 2019, Pages 851-860. DOI: 10.1016/j.fuel.2018.12.047. " Copyright 2019 Elsevier.

The raw data of this publication can be reached under DOI: 10.5281/zenodo.4129481.

Fuel 241 (2019) 851–860

Contents lists available at ScienceDirect

Fuel

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

Full Length Article

Density functional theory study of CO formation through reactions of polycyclic aromatic hydrocarbons with atomic oxygen ($O(^3P)$)

Özlem Yönder*, Christof Hättig

Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, D-44780 Bochum, Germany

ARTICLE INFO

Keywords:
Polycyclic aromatic hydrocarbons
PAH
DFT
$O(^3P)$
CO formation
Oxy-fuel
Char
Combustion

ABSTRACT

To provide a qualitative analysis on how the reaction energies and mechanisms of initial steps of char combustion depend on a) the size of polycyclic aromatic hydrocarbons (PAHs) and b) the reaction sites, we investigate the initial steps of the individual reactions of thirteen PAHs with atomic oxygen with density functional theory. Specifically, we have studied three elementary steps: i) $O(^3P)$ addition, ii) subsequent oxyradical formation through H abstraction, and iii) CO elimination from the oxyradicals. We find that the CO elimination from the oxyradicals follows either a three, two, or single step mechanism, depending on the geometrical strain at the reaction site. In accordance, we observe that as the geometrical strain at the reaction site increases, the activation energy of the first step of the CO elimination from the oxyradicals increases. Important findings that are relevant for a multiscale modeling of the oxidation of large PAHs, e.g. during char burnout, are that the activation energies for these reactions can be related to the reaction energies through Bell-Evans-Polanyi (BEP) fits, and that for the O addition and the last step of the three or two step CO elimination mechanisms, i.e. the CO elimination from carbonyl radicals, the reaction energies are correlated with the stabilization energies calculated from the Hückel molecular orbital theory.

1. Introduction

The interest in oxy-fuel combustion in coal-fired power generation has been increasing during the last years due to its potential to reduce NO_x emissions and to capture CO_2 more easily as compared with conventional air combustion [1]. To optimize the combustion, an understanding of the reactions within an oxy-fuel atmosphere is crucial and, indeed, there is a body of studies focusing on the macroscopic and microscopic levels. However, atomistic studies of the reaction mechanisms are still rather uncommon.

Char includes amorphous and crystalline contents [2–4]. The amorphous content is composed of small, heat-resistant aliphatic and hetero(aromatic) elements such as phenols, quinones, pyrols, furans, and small polyaromatic units; whereas the crystalline part is composed of turbostratically stacked graphene layers [2]. Since both the amorphous and the crystalline content include aromatic structures, PAHs are important model systems to study the combustion of char. Additionally, the understanding of the oxidation of PAHs provides also important information on the oxidation of soot.

In the current contribution we studied initial steps of the reaction of atomic oxygen with PAHs of different sizes and at different reaction sites. Fig. 1 illustrates the different molecular regions of PAHs studied in this work. In IR spectroscopy of PAHs, the bending C–H vibrations of the molecular regions containing one (solo), two (duo), three (trio), or four (quartet) adjacent hydrogens show different frequency ranges [5]. The range is 900–860 cm^{-1} for the solo region, 860–800 cm^{-1} for the duo, 810–750 cm^{-1} for the trio and 770–735 cm^{-1} for the quartet regions [5]. The quartet and trio regions can further be differentiated by their relative location to the adjacent fused ring. The first carbon atom next to the neighboring ring is referred as inner (or α) and the second one as outer (or β) position. Another classification of positions as bay and fjord regions originates from NMR spectroscopy. The bay and fjord regions are sterically crowded; therefore, they show downfield shifts in NMR which are more pronounced for the fjord regions than the bay regions [5].

An important reaction which initiates the decomposition of polycyclic aromatic structures (C_xH_y) during combustion is the elimination of CO from oxyradicals $C_xH_{y-1}O\cdot$. A pathway leading to the formation of oxyradicals can be the addition of an O_2 molecule to a PAH radical, followed by the cleavage of the O–O bond [6]:

$$C_xH_{y-1} + O_2 \rightarrow C_xH_{y-1}O_2 \rightarrow C_xH_{y-1}\,O + O(^3P). \quad (1)$$

A second oxyradical formation pathway is the addition of $O(^3P)$ to PAHs forming a diradical and the subsequent H abstraction,

* Corresponding author.
E-mail addresses: oezlem.yoender@theochem.ruhr-uni-bochum.de (Ö. Yönder), christof.haettig@theochem.ruhr-uni-bochum.de (C. Hättig).

https://doi.org/10.1016/j.fuel.2018.12.047
Received 2 May 2018; Received in revised form 4 October 2018; Accepted 5 December 2018
0016-2361/ © 2018 Elsevier Ltd. All rights reserved.

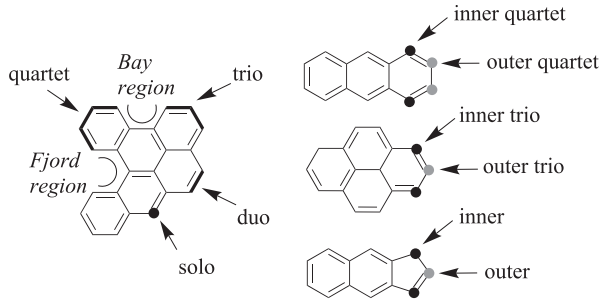


Fig. 1. Molecular regions of PAHs as classified in this work.

$$C_xH_y + O \rightarrow C_xH_y O \rightarrow C_xH_{y-1} O + H, \quad (2)$$

which can occur on the triplet potential energy surface. Following the formation of the oxyradical, CO can be eliminated as:

$$C_xH_{y-1} O \rightarrow C_{x-1}H_{y-1} + CO. \quad (3)$$

Atomic oxygen O(^{3}P) can be formed in a high temperature environment together with other radicals such as H, OH, or HO_2 during the initiation process (the so-called radical pool formation) [7,8]. For example, Brezinsky suggested that phenyl radicals can be formed from benzene through abstraction of H [8]:

$$C_6H_6 \rightarrow C_6H_5 + H \quad (4)$$

For toluene proposed initiation processes are the thermal elimination of the low strength benzylic H:

$$C_6H_5CH_3 \rightarrow C_6H_5CH_2 + H, \quad (5)$$

or the abstraction of hydrogen by O_2 [8]:

$$C_6H_5CH_3 + O_2 \rightarrow C_6H_5CH_2 + HO_2. \quad (6)$$

The branching reaction of H with O_2 leads to the formation of the radicals OH and O [8]:

$$H + O_2 \rightarrow OH + O. \quad (7)$$

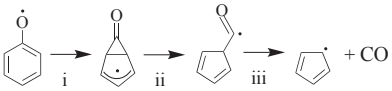
Another route for the formation of O(^{3}P) is the above mentioned cleavage of the O-O bond in peroxy (RO_2·) radicals.

Previous theoretical studies established that the most important pathway for the reaction of C_6H_6 with O(^{3}P) starts with an electrophilic addition of O(^{3}P) to benzene in the triplet state [9,10]. The calculated activation barrier of this reaction is 20.5 kJ/mol at the CBS-QB3 level [10] and 23.8 kJ/mol at the G3B3 level [9,10] of theory. Taatjes et al. observed, however, that in a flow reactor above 700 K, phenoxy (C_6H_5O) + H and cyclopentadiene (C_5H_6) + CO appear as products [9].

The diradical formed by the electrophilic addition of O(^{3}P) to benzene can pass through an intersystem crossing by a nonradiative transition to the singlet state followed by the almost barrierless isomerization to benzene oxide with 230 kJ/mol energy release as calculated by CASSCF [9]. The singlet benzene oxide can rearrange to 2,4-cyclohexadienone with a barrier of 160 kJ/mol which can further isomerize to phenol or cyclopentadiene + CO with energy barriers of, respectively, 213 and 251 kJ/mol [9,10].

Nguyen et al. studied the formation of the cyclopentadienyl radical (C_5H_5) on the triplet surface from the diradical formed upon O(^{3}P) addition to benzene through formyl cyclopentadiene (C_5H_5CHO) as intermediate [10]. Because of the high activation barrier of 135 kJ/mol (obtained from CBS-QB3) and lack of experimental evidence for C_5H_5 + CHO as products, they concluded that this route is not important for benzene [10].

Another reaction route that leads to the cyclopentadienyl radical is the decomposition of oxyradicals. An early theoretical study by Olivella et al. using CASPT2/6–31G(d,p) and a later study by Lui et al. using CASSCF/6–31G* established that the elimination of CO from the phenoxy radical follows a three step mechanism: i) Formation of a bicyclic radical intermediate, followed by ii) a ring opening of its cyclopropanone moiety to a carbonyl radical, and finally, iii) CO elimination from the carbonyl yielding the cyclopentadienyl radical [11,12]:



(8)

Concerning larger PAHs, You et al. studied the decomposition of pentacene oxyradicals at the DFT B3LYP/6–311G(d,p) level [13]. On the outer quartet site the reaction follows a similar mechanism as for the phenoxy radical with activation barriers of 195, 42.3, and 10.5 kJ/mol for the three elementary reaction steps. On the inner quartet site, the barriers were found to be 238, 6.7, and 10.5 kJ/mol [13]. For the solo sites they could not find a similar reaction path [13]. Instead they found pathways with different intermediates which lead to ring openings with barriers around 440 kJ/mol [13]. Liu et al. did a similar study at the DFT B3LYP/6–311 + G(d,p) level for the decomposition of benz[a]anthracene oxyradicals at the outer quartet, bay-inner quartet, solo, and bay-solo positions and of H-cyclopenta[a]anthracene oxyradical as an example for the decomposition of a five-membered ring [14]. The energy barriers for the three steps were respectively 197, 74.9, and 10.8 kJ/mol for the outer quartet site, and 217, 16.3, and 14.2 kJ/mol for the bay-inner quartet site. In contrast to the studies on the pentacene oxyradicals, they were able to find at a solo site a reaction path similar to that of benzene but occuring in a single step instead of in three steps. The energy barrier for the single step elimination at the solo site was calculated as 348 kJ/mol and at the bay-solo site as 335 kJ/mol [14].

In the current contribution we report a study on the reactions of PAHs with O(^{3}P) which lead to the formation of the oxyradicals (reaction (2)) and the subsequent elimination of CO (reaction (3)). Existing studies on these reactions covered only a few of the different possible site types mentioned above. As it is not feasible to treat all possible reaction sites for a sufficiently representative set of molecules (cmp. Fig. 2) with accurate wavefunction methods like CASPT2 and as we are aiming in the current study rather for the identification and understanding of trends transferable to larger PAHs we employ for the current study density functional theory.

2. Computational methods

For the optimizations of initial structures and reaction pathways we used DFT with the meta-GGA functional TPSS [15] and the SVP basis set [16]. The final geometries and electronic energies were obtained with the TPSSh hybrid exchange correlation functional [17] and the TZVP basis set [18] including the D3 dispersion correction [19]. The dispersion correction is included because for large molecules the London interactions have a significant influence on the potential energy surfaces. The zero-point vibrational energy (ZVPE) corrections to the electronic energies were computed at the same electronic structure level (TPSSh-D3/TZVP) within the harmonic approximation. We have used the scaling factor of 0.9874 optimized in Ref. [20] for TPSSh/6-31 + G(d,p) to fit the experimental ZPVE.

All the calculations were performed with the TURBOMOLE program package [21]. For all of the calculations, the resolution of the identity (RI-J) approximation has been used for the Coulomb energy with the optimized auxiliary basis sets from Ref. [22].

The reaction and activation energies are calculated as

$$\Delta E = \sum E^{el}_{products} - \sum E^{el}_{reactants} + \Delta ZPVE, \quad (9)$$

and

$$\Delta E^{\ddagger} = E^{el}_{TS} - \sum E^{el}_{reactants} + \Delta ZPVE. \quad (10)$$

The Gibbs free energies are calculated using the *freeh* program of the

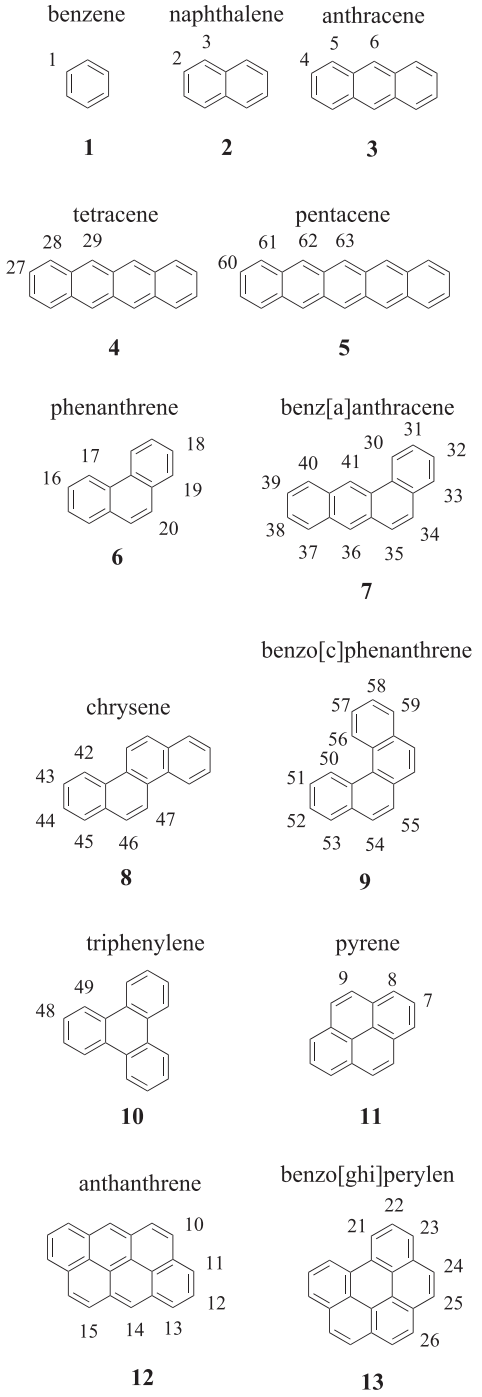


Fig. 2. The set of PAHs. The numbers represent studied reaction sites. Phenanthrene, benz[a]anthracene, chrysene, triphenylene, and benzo[ghi] perylen are the PAHs with bay region(s). Benzo[c]phenanthrene has a fjord region.

TURBOMOLE package. This code calculates the vibrational partition function within the harmonic oscillator approximation and the rotational partition function within the rigid rotor model. Based on the ideal gas assumption for the O and H atoms the internal energy contribution due translations is assumed to be 3R/2, where R is the gas constant, and the PV term of enthalpy is calculated as RT, where P is the pressure, V the molar volume and T the temperature.

Table 1
Comparison of the TPSSh-D3/TZVP energies with the literature. In the nomenclature SX-Y, X = 1,2,3,4,5,6 refers to the reaction step/intermediate with S1: PAH, S2: triplet intermediate after O(^{3}P) addition, S3: oxyradical, S4: radical with cylopropane moeity, S5: carbonyl radical, and S6: final radical with five-membered C-ring. The index Y refers to the reaction site (cmp. Fig. 2).

	$\Delta E^{\ddagger}$ (kJ/mol)		ΔE (kJ/mol)	
	TPSSh-D3	Literature	TPSSh-D3	Literature
S1-1 + O → S2-1	−4.3	17.2[a]	−63.7	−60.7[a]
S2-1 → S3-1 + H	41.1	43.9 [a]	16.9	−5.4[a]
S3-1 → S4-1	199.0	216.3[b]	184.3	192.5[b]
S4-1 → S5-1	29.6	34.7[b]	−35.3	−65.3[b]
S5-1 → S6-1	12.0	23.0[b]	−30.4	−33.5[b]
S3-2 → S4-2	185.9	187.0[c], 197.5[d]	169.9	170.3[c], 183.7[d]
S4-2R → S5-2R	27.9	32.2[c], 29.3[d]	−52.0	−76.6[c], −64.4[d]
S5-2R → S6-2R	13.4	22.2[c], 15.5[d]	−40.9	−50.6[c], −59.4[d]
S4-2L → S5-2L	66.8	82.8[c], 68.2[d]	33.3	10.5[c], 23.8[d]
S5-2L → S6-2L	3.1	18.8[c], 2.1[d]	−126.2	−137.7[c], −147.7[d]
S3-3 → S4-3	241.5	241.8[c], 257.7[d]	239.7	241.4[c], 258.2[d]
S4-3 → S5-3	0.0	2.1[c], 0.4[d]	−105.6	−130.1[c], −120.9[d]
S5-3 → S6-3	13.3	22.2[c], 15.5[d]	−41.0	−50.6[c], −59.0[d]
S3-9 → S5-9	252.7	252.7[e]	146.4	120.9[e]
S5-9 → S6-9	18.4	24.3[e]	−27.5	−42.3[e]

[a] CBS-QB3 energies of the B3LYP/6–311G(d,p) geometry optimized structures from Nguyen et al. [10].
[b] CASPT2/6–31G(d,p) + ZPVE results from Olivella et al. [11].
[c] G3(MP2,CC) energies of the B3LYP/6–311G(d,p) optimized geometries [6].
[d] B3LYP/6–311G(d,p) + ZPVE results from Zhou et al. [6].
[e] G3(MP2,CC) results from Singh et al. [23].

3. Accuracy of the results

In order to evaluate the accuracy of the TPSSh-D3/TZVP results, Table 1 compares the reaction energies for benzene **1**, naphthalene **2** for site 2 and 3, and pyrene **11** for site 9 with the computational reference data from the literature. The TPSSh-D3/TZVP results are in general less exothermic for the opening of the three-membered rings and the formation of the carbonyl radicals, but the deviations exceed only in two cases ± 26 kJ/mol. This is approximately the accuracy obtained at a DFT level for such systems if one takes into account the uncertainties from the transition state optimizations done with different functionals and basis sets.

4. Reaction mechanism and geometrical effects

4.1. Reaction mechanism

The first reaction step that we studied is the addition of O(^{3}P) to the PAHs which is exothermic. In a subsequent step a hydrogen atom is eliminated and the corresponding oxyradical is formed. Among the PAHs studied (Fig. 2), neither the O(^{3}P) addition nor the H atom abstraction are geometrically constrained. Therefore, all the PAHs studied in this work follow for these steps analogous pathways as benzene (Eq. (2)).

The reaction pathway for CO elimination is well-established for the phenoxy radical (Eq. (9)) but only few and partially contradictory results have been published for larger PAHs. For the solo sites of pentacene it has in Ref. [13], based on the result of a potential energy scan at the B3LYP/6-311(d,p) level, been concluded that reaction pathways similar to that of the phenoxy radical are unlikely. In the current study we were able to find for the solo sites of pentacene the analogous reaction pathway for a single step mechanism. For the solo site 62 of pentacene **5** we obtained the reaction energy as 73.0 kJ/mol and the activation energy as 327 kJ/mol, and for the solo site 63 as 70.9 and

368 kJ/mol, respectively. Both activation energies for the solo sites of pentacene are significantly smaller than the 440 kJ/mol calculated by You et al. [13] for the ring opening at the B3LYP/6-311G(d,p) level.

In another study using B3LYP/6-311 + G(d,p) Liu et al. has found that at the solo site of the benz[a]anthracene (**7**) oxyradical, which is geometrically similar to the solo sites of pentacene, CO eliminates along a similar pathway as for phenoxy radicals with the only difference that the elimination happens in a single step [14]. They reported the activation energies of the CO elimination from the solo (site 36) and bay-solo sites (site 41) of the benz[a]anthracene oxyradical as 348 kJ/mol and 335 kJ/mol respectively [14]. We obtained the activation energies of these sites at the TPSSh-D3/TZVP level as 304 kJ/mol and 312 kJ/mol, respectively. For the bay-solo site (41) the current result is close enough to the value reported by Liu et al. to assume that the two calculations follow a similar reaction pathway; however, for the solo site (36) it deviates significantly. Therefore, we re-optimized for these sites the structures of the oxyradical and transition states (T36) with the B3LYP functional (using the VWN(III) functional) and the 6-311 + G(d,p) basis set. At this level we obtained for the current pathway activation energies of 318 kJ/mol and 332 kJ/mol. At the bay-solo site 41, the geometry of the transition state structure are the same as the one reported by Liu et al.; however, for the solo site 36, as compared with those reported by Liu et al. [14], the transition state structure has the shorter C–C bond length of the cyclopropanone-like moeity at the opposite side and is lower in energy by approximately 30 kJ/mol. A similar difference in the activation energies depending on from which side of the cyclopropanone moeity the CO eliminates can also be seen in Fig. 4. The difference in the activation energies for the S4-12 → S5-12L and S4-12 → S5-12R reactions is 22.5 kJ/mol, which is of the same order of magnitude as the difference between the transition state found for the CO elimination at the solo site 36 in this study and that reported by Liu et al., [14]. That is, the transition states found in this study are comparable to the ones reported in the literature.

Analysis of the reactions at sixty-three sites (Fig. 2) indicates that, in general, the CO elimination from PAH oxyradicals can follow either a three, two, or single step mechanism depending on the geometrical strain at the reaction site. Fig. 3 shows the reaction pathways and Fig. 4 the corresponding intermediates along the reaction pathways for different sites of anthanthrene — an example that illustrate the relation between the site types and the three, two, and single step CO elimination mechanisms. During the O(^{3}P) addition, spin density is transferred from the oxygen atom to the ring system and a C–O σ bond is formed. During the H abstraction, the σ C–H bond of the triplet intermediate structure S2 breaks and the oxygen forms a π bond which is conjugated with the π system of the aromatic carbon rings. The doublet oxyradical can then undergo the CO elimination reaction. As can be seen from these figures, the three step CO elimination mechanism is analogous to that of the phenoxy radical sketched in Eq. (8). For this mechanism the second step, which is the ring opening of the cyclopropanone moeity to a carbonyl radical, possesses the overall highest barrier along the reaction pathway. Therefore, this step is the rate limiting step for the three step CO elimination mechanism. For the two step mechanism the structures with a strained three-membered carbon ring, which are the intermediates formed in the first step of the three step mechanism (S4 structure), are not local minima along the reaction pathway. However, the transition states of the first step of the two step mechanism (T35 structures) have a similar structure with a three-membered carbon as the intermediates S4 of the three step mechanism. The CO elimination from carbonyl radicals for the two and three step mechanisms are analogous and comparable. In the single step mechanism none of the two intermediates (S4 and S5 structures) are formed. But again, the transition state structure T36-14 is similar to the S4 structure of the three-step mechanism.

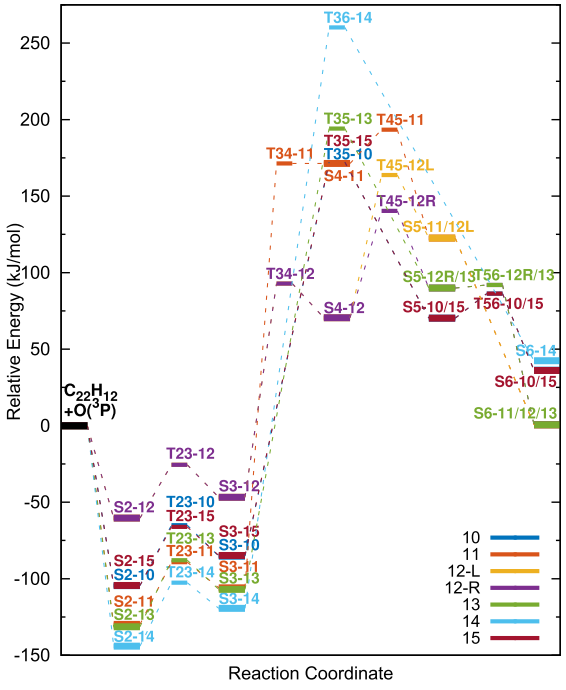


Fig. 3. Reaction profiles for anthanthrene. The corresponding structures are shown in Fig. 4. Site 14 follows a single step CO elimination mechanism, the sites 10, 13, and 15 follow a two step, and sites 11 and 12 follow a three step CO elimination mechanism.

4.2. First step of CO elimination and geometrical effects

The geometrical strain can be related to the site type (quartet, trio, duo, or solo) and the number of neighboring rings. Among the sample set of PAHs, the quartet sites exist in the acenes (**2–5**) and phenacenes (**6**, **8**) as well as in benz[a]anthracene (**7**), benzo[c]phenanthrene (**9**), and triphenylene (**10**). Trio sites are found in the peri-fused PAHs (**11–13**), and duo sites in the phenacenes (**6**,**9**), benz[a]anthracene (**7**), benzo[c]phenanthrene (**9**), and peri-fused PAHs (**11–13**). In our sample set, solo sites exist in the acenes larger than naphthalene (**3–5**) and in anthanthrene (**12**). In addition, the quartet and trio sites have different reactivities when they are located at an inner (α carbon) or outer (β carbon) positions (see Fig. 2), – the former are geometrically more constrained and require higher activation energies for reactions that involve a ring deformation.

Table 2 shows the ranges of activation energies of the first step of the CO elimination sorted according to the site type and the number of neighboring rings. For the three step CO elimination mechanism from the oxyradicals, this corresponds to the isomerization of the oxy- to cyclopropanone-like radicals. Among the reaction sites in the test set which follow the three step CO elimination mechanism, the most reactive ones are the outer trio sites with activation energies ranging from 140 to 162 kJ/mol. They are followed by the outer quartet sites with activation energies ranging from 170 to 194 kJ/mol. The activation energies of the inner quartet sites ($189 < \Delta E^{\ddagger}$ (kJ/mol) < 242) are higher than for the outer quartet sites. The activation energies for duo sites with two neighboring rings ($182 < \Delta E^{\ddagger}$ (kJ/mol) < 220) are in between those of two types of quartet sites.

The sites follow the two or the single step CO elimination mechanism when the geometrical strain raises the energy for the cyclopropanone structure so much that it becomes instead of an intermediate a transition state. For the duo sites with three neighboring rings, CO elimination occurs in two steps. The range of activation energies for the first reaction step is for these sites with 240–257 kJ/mol similar to that for the inner trio sites which follow the three step mechanism. On the other hand, the inner trio sites still follow a three step mechanism with

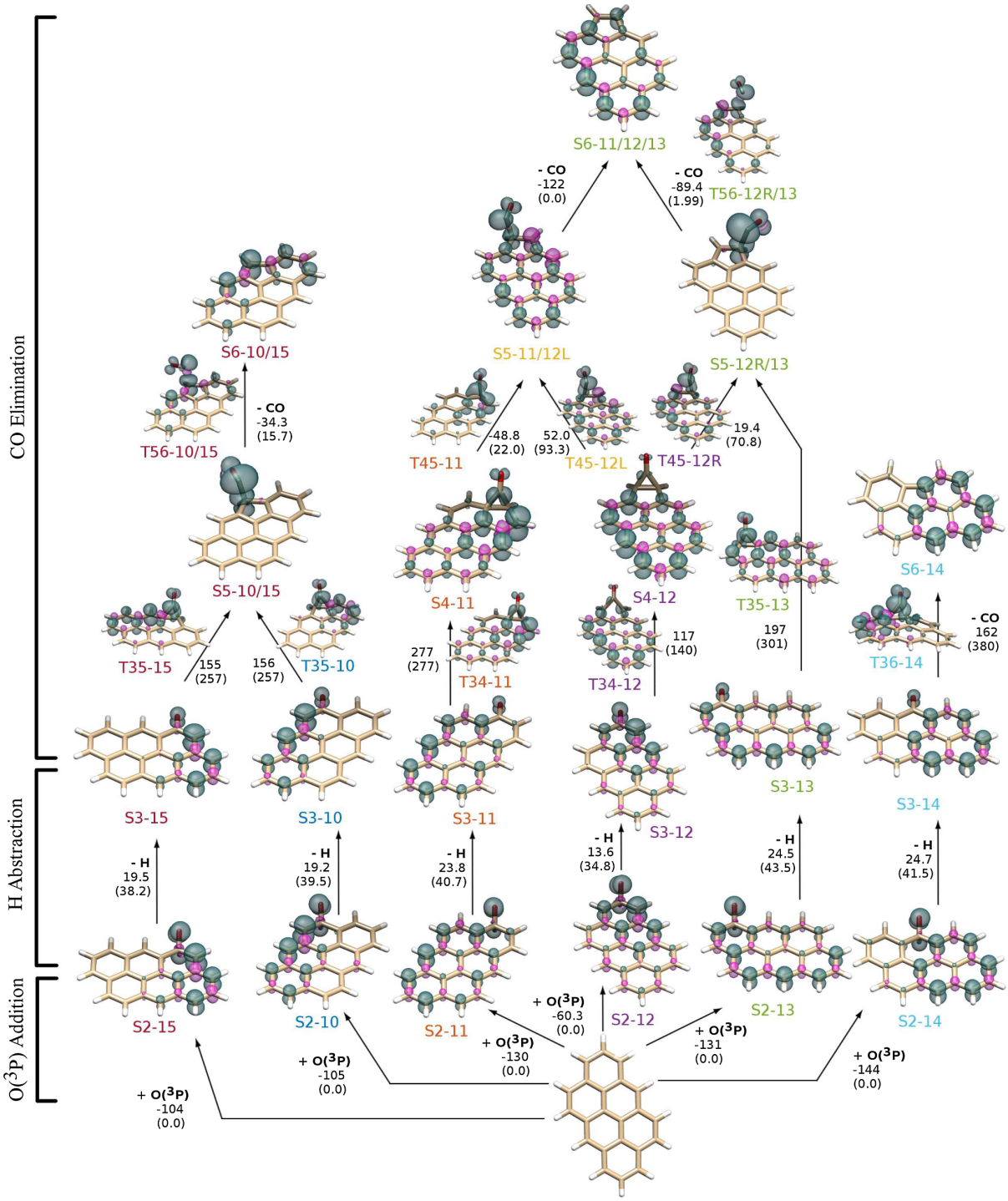


Fig. 4. The intermediates along the reaction pathways of anthanthrene as examples for the three different kinds of pathways. Reacting species are shown with their spin densities. Reaction and activation energies are given in kJ/mol, activation energies are given inside parantheses. The O(^{3}P) addition step occurs without a barrier for the sites on anthanthrene.

the exception of site 13 in anthanthrene **12**. The activation energies of the trio sites are in the range 236–277 kJ/mol, which is higher than that for the duo sites with three neighboring rings although the trio sites are geometrically less constraint. Among the inner trio sites, site 21 at the bay region of benzo[ghi]perylen **13** has the lowest activation energy due to its flexibility. In contrast to this, the exceptional inner trio site 13 of anthanthrene **12** has a very high activation energy of 301 kJ/mol. The reaction at this site occurs in a two step mechanism. In general, the activation energies for the CO elimination are for the solo sites the highest among all the sites studied in this work. The highest activation energy of 380 kJ/mol belongs to the solo site 14 in anthanthrene **12** which has four neighboring rings.

This analysis of the activation energies for the first step of the CO elimination for different types of sites indicates that as the geometric

Table 2
Activation energies of the first step of CO elimination ($\Delta E^{\ddagger} = E_{T3X} - E_{S3} + \Delta ZPVE$ where X = 4,5,6 for three, two, and single step mechanisms respectively) in kJ/mol. The range of activation energy for each site type indicates that geometrical strain at the reaction site determines the CO elimination mechanism and activation energy.

CO abstraction mechanism	Site type		$\Delta E^{\ddagger}$ range (kJ/mol)
Three step	Outer trio		140–162
	Outer quartet		170–194
	Duo & 2 neighbors		184–220
	Inner quartet		189–242
	Inner trio (except anthanthrene)		236–277
Two step	Duo & 3 neighbors		240–257
	Inner trio @anthanthrene		301
Single step	Solo & 2 neighbors		304–368
	Solo & 4 neighbors (@ anthanthrene)		380

constrains at the reaction site increases, the activation energy of this step also increases. However, we could not find a quantitative measure for only the geometrical strain, i.e. without including electronic effects from the aromatic π system, and therefore we could also not find a quantitative relation of it to the activation energy of the first step of the CO elimination from oxyradicals.

5. Electronic effects

5.1. Hückel molecular orbital theory analysis

The O(^{3}P) addition to PAHs, the subsequent H abstraction which leads to the oxyradicals, and the CO elimination from carbonyl radicals are expected to be dominated by electronic effects from the π-system. To better understand the differences and trends in the reaction energies for these steps we made an analysis based on the Hückel molecular orbital theory (HMOT). In HMOT only the π-electrons are considered and thus the energies obtained from it can be used as a measure of the resonance energy, i.e. the electronic stabilization due to delocalization of the electrons in the aromatic ring system. To evaluate the effect of rearrangements in the π-electron system on the variation of the reaction energies, we have calculated resonance energies for π system analogues

Table 3
The σ-framework of the structures used for the reaction of benzene in the HMOT analysis. The PAH analogues are used for the HMOT analysis of the influence of the π-electron delocalization on the reaction energies.

	Reactant	Product
O(^{3}P) Addition		
H Abstraction		
CO Elimination		

of the actual reaction intermediates shown in Table 3. These were chosen such that they have the same atoms and the same number of electrons in the π-system as the actual reaction intermediates but are planar and describable by HMOT. The contribution of the π-system to the reaction energy has been calculated for each reaction and site by taking the difference in the total resonance energies of the analogues for the reactant and the product.

Fig. 5 shows correlation plots where the reaction energies for the O(^{3}P) addition, the H abstraction, and the CO elimination from carbonyl radicals are compared with the contribution from the resonance energy obtained from HMOT. According to this analysis, the correlation between the actual reaction energies and the contribution from change in resonance energy is only weak for the H abstraction, but for the O addition and CO elimination there is a strong correlation. These results indicate that the contribution from the π-orbital rearrangements to the reaction energies play a major role for the O addition and CO elimination; on the other hand, for the H abstraction other effects are equally important.

5.2. Site specific analysis: O(^{3}P) addition

The strong correlation between the changes in the resonance energy and the reaction energies for the O(^{3}P) addition to PAHs indicates that mainly electronic effects from the aromatic π-system are responsible for the differences in the reactivity of different sites. In order to further analyze the relation between the reaction energies, and the particular site types and size of the PAHs, Fig. 6 shows the reaction energies at each studied reaction site.

For the linear PAHs **2–5** the reaction becomes in general more exothermic with increasing size of the PAH. Similarly, when we compare the sites of pyrene **11** (−50.8, −110, −98.0 kJ/mol) with analogous trio and duo sites of anthanthrene **12** (−60.3, −130/−131, −105/−104 kJ/mol), the energy release is higher for the larger PAH anthanthrene **12**. The increasing exothermicity of the O(^{3}P) addition to PAHs with increasing size of the aromatic system can be explained by the increasing delocalization of the unpaired electron in the product.

Another observation is that for the linear molecules (**2–5**) the reaction at the inner quartet site is more exothermic than at the outer quartet site, and that for this elementary step the solo sites on the inner ring are the most reactive ones. There are similarities between these patterns and the reactivity patterns described for the Diels–Alders reactions. For this class of reactions it has also been found that the reactivity of acenes increases with increasing size of the PAH and that the central ring is the most reactive one [24].

Lastly, when the PAH is branched the O(^{3}P) addition becomes in many cases less exothermic, e.g. when we compare analogous sites in anthracene **3** (−93.2, −108 kJ/mol) and phenanthrene **6** (−74.8/−69.3, −86.4/−88.6 kJ/mol) or in anthanthrene **12** (−105/−104, −130/−131, −60.3 kJ/mol) and benzo[ghi]perylen **13** (−102/−96.4/−102, −109/−107, −58.5 kJ/mol).

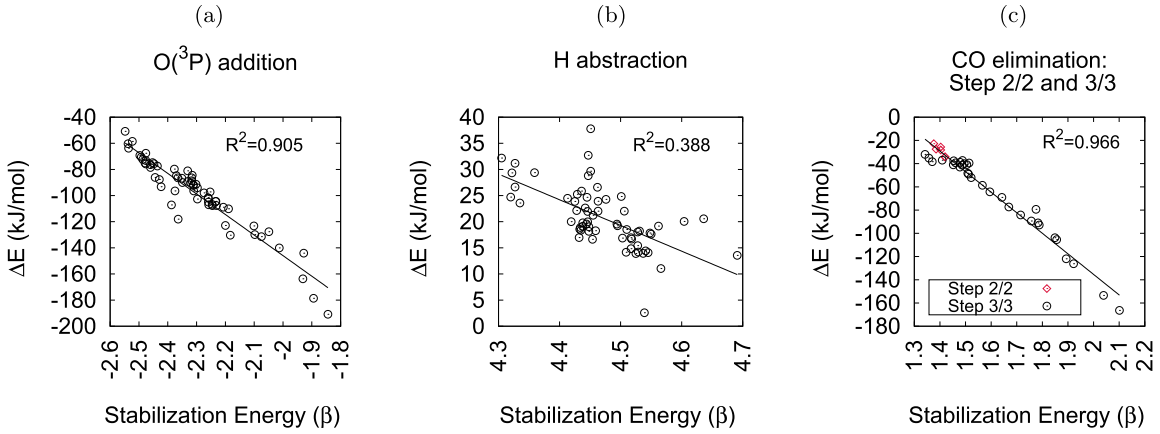


Fig. 5. Reaction energies calculated by DFT plotted versus the change in the resonance energies obtained from HMOT for (a) the O(^{3}P) addition to PAHs ($\Delta E/kJ{\cdot}mol^{-1} = -157\ E_{stabilization}/\beta - 461$, s= 8 kJ·mol^{-1}, $R^2 = 0.905$), (b) the subsequent H abstraction to from the oxyradical ($\Delta E/kJ{\cdot}mol^{-1} = -49.2\ E_{stabilization}/\beta + 241$, s= 5kJ·mol^{-1}, $R^2 = 0.388$), and (c) CO elimination from carbonyl radicals ($\Delta E/kJ{\cdot}mol^{-1} = -177\ E_{stabilization}/\beta + 219$, s= 7 kJ·mol^{-1}, $R^2 = 0.966$). Benzene is excluded. β is negative as it represents the (stabilizing) resonance energy for an isolated π bond. R^2 is the coefficient of determination and s the standard deviation of the fit.

5.3. Site specific analysis: The CO elimination from carbonyl radicals

The reaction energies of the CO elimination from carbonyl radicals at different sites are shown in Fig. 7. For the linear acenes this elementary step does not occur at the outer site of five membered rings when the initial PAH is larger than anthracene even though before the isomerization the site was a quartet site. The reason is the lower stability of the intermediate carbonyl radical at the outer quartet site compared to the inner quartet site. This stability difference increases with the size of the acenes, which is an electronic effect from the π-sytem: For naphthalene **2**, anthracene **3**, tetracene **4**, and pentacene **5** the difference in the resonance energies between the carbonyl radicals at the inner and outer sites of the five membered ring obtained from HMOT are respectively 0.4, 0.6, 0.6, and 0.7 β. As a consequence of this, only the more stable carbonyl radicals at the inner sites are formed in the preceding step of the CO elimination.

Additionally, for the CO abstraction from the carbonyl radicals we find the trend that the reaction becomes for analogous sites more exothermic with increasing size of the PAHs. Finally, if the pathway exists, the CO at the outer position abstracts significantly more exothermic than the one at the inner position, which is again a consequence of the different stabilities of the two isomeric structures for the carbonyl radicals.

6. Relation of the reaction and activation energies

Since the localization of the reaction pathways and the transition states is computationally much more demanding than the optimization of the structures of the reactants and products, it is desirable to find a correlation between the reaction and activation energies that can be used to extrapolate the results from the sample set to other molecules. Therefore, we analyzed the degree of correlation for each reaction step by means of Bell-Evans Polanyi (BEP) [25,26] plots which are shown in Fig. 8. For all of the reaction steps there is at least a general trend of a BEP correlation. However, we also note that some mechanism are sampled by our test set more than the others. For example, the set of sites corresponding to PAHs consists of thirty-four quartet, thirteen duo, eight trio, and seven solo sites; or among the sixty-three sites there are only seven sites at bay regions and only one site at a fjord region.

For the O(^{3}P) addition, the more exothermic the reaction is, the lower is the activation energy $\Delta E^{\ddagger}$. The negative activation energies are due to the way reaction and activation energies are calculated according to Eq. (10) relative to the fully separated, i.e. isolated, reactants. When calculated with respect to the shallow van der Waals (vdW) minima, the activation energies are positive. For example, if it is calculated from Eq. (10), for site number 1 at benzene (**1**) $\Delta E^{\ddagger}$ is −4.34 kJ/mol. On the other hand, when calculated with respect to the intermediate vdW minimum of the reactants, it is 7.51 kJ/mol. Similarly, the activation energies $\Delta E^{\ddagger}$ calculated according to Eq. (10) relative to the isolated reactants and relative to the vdW minima are, respectively, −11.5 and 7.39 kJ/mol for site 2, −10.5 and 8.02 kJ/mol for site 32, −22.2 and 1.93 kJ/mol for site 35, and, finally, −11.7 and 6.8 kJ/mol for site 43. The BEP plot of the O(^{3}P) addition has a slope less than 0.5 indicating a reactant-like early transition state. As the reaction gets exothermic, the activation energy becomes zero. Additionally, the more exothermic the reaction is, the earlier the transition state is, as evidenced by the larger C–O distance and in accordance with the Hammond's postulate [27] as can be seen from Fig. 9a.

The H abstraction has a transition state structure similar to the products, which results in a BEP fit with a slope larger than 0.5 (Fig. 8b). The outlier at $\Delta E^{\ddagger}$ = 63.0 kJ/mol and ΔE = 37.8 kJ/mol corresponds to the site 56 of benzo[c]phenanthrene **9** which is a site at the fjord region. The reason that the activation energy is in this case higher than the BEP fit is probably that the repulsive close H···H and the attractive H···O contacts in the fjord region lead to an additional stabilization of the non-planar structure of intermediate S2–56 relative to the more planar structure of the oxyradical as illustrated in Fig. 10. Similar as for the O(^{3}P) addition and in accordance with Hammond's postulate, the more endothermic the H abstraction is, the later the transition state and the larger the C–H distance in the transition state is, as can be seen from Fig. 9b.

Finally, we see from the BEP plot for the CO elimination in Fig. 8c that the first step of the three step mechanism shows a very good correlation. Also, the second step of the three step CO elimination mechanism shows a very good BEP correlation (see Fig. 8d). The last step of the two and the three step CO elimination mechanisms resembles each other. Therefore, they are plotted at the same BEP plot in Fig. 8e as step 2/2 and step 3/3 respectively. Fig. 9c shows the C–C(O) distance of the transition state structure for this reaction step. We observe again in accordance with Hammond's postulate the general trend that the less exothermic the reaction is, the later the transition state is which in this case means the larger the C–C(O) bond length is. As the sample of PAHs studied in this work contains only few sites where the CO elimination follows the two or the single step mechanism, we refrained from

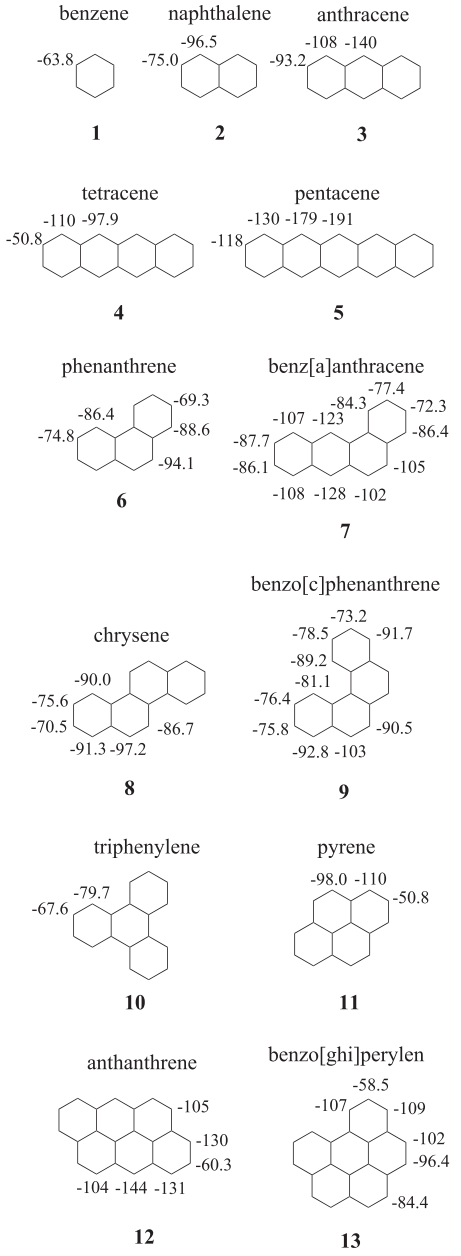


Fig. 6. Reaction energies for the O(^{3}P) addition to PAHs for each studied site in kJ/mol. For clarity only the σ-bond network of the carbons is shown.

analyzing them further. Since for them the activation barriers for the CO elimination are considerably higher than for the sites where the elimination follows the three step mechanism, they are not expected to contribute significantly to the overall reaction rates for PAH oxidation.

The above analysis was done based on the reaction energies. Fig. S1 in the Supplementary Information shows BEP plots based on the Gibbs free energies for three of the reaction steps calculated using the vibrational, rotational, and translational contributions to the enthalpy and the entropy at 1200 K and 0.1 MPa. For the H abstraction and the CO elimination from carbonyl radicals, the BEP correlation based on Gibbs free energies is weak and there is only a general trend. Nonetheless, for the first and second step of the three step CO elimination mechanism, which are the rate determining steps, the BEP plots based on the Gibbs free energies still show a good correlation.

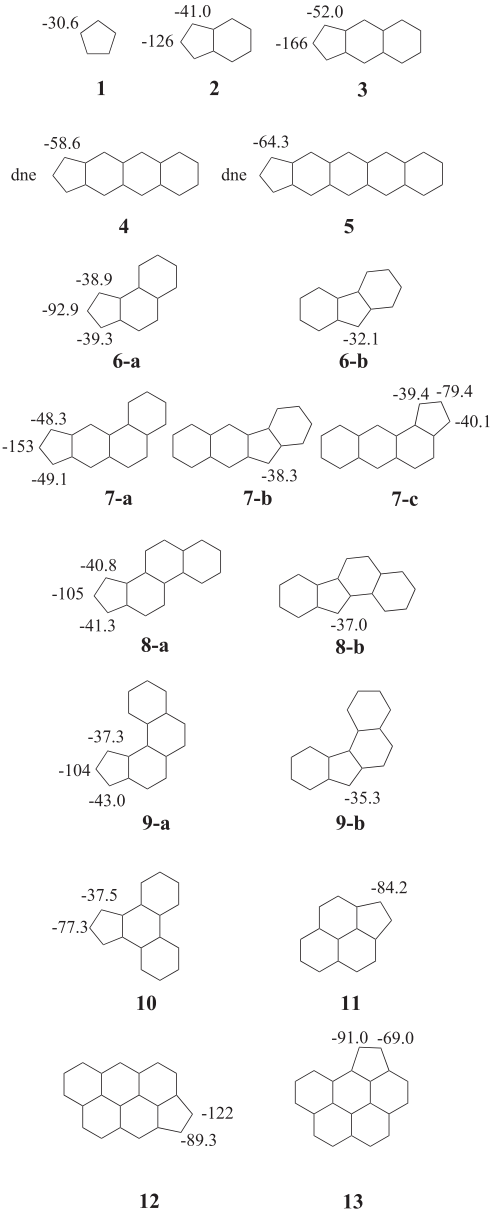


Fig. 7. Reaction energies of the CO elimination from carbonyl radicals in kJ/mol. Only the σ-bonding network of carbons are shown for the final polyaromatic structures. Some of the shown structures (**7-b**, **8-b**, **9-b**) and some other polyaromatics with a five-membered ring polyaromatics may be formed through single step CO elimination mechanism. However, as in such a case the reaction energy is dependent on the geometrical strain and CO eliminates in a single step, they are irrelevant in this discussion.

7. Summary and conclusions

For a qualitative analysis we have studied the reaction and activation energies of the reactions of PAHs with O(^{3}P) using the TPSSh density functional, Grimme's D3-dispersion correction, and the TZVP basis set based on a test set which samples different molecular regions and sizes of PAHs. Specifically we have studied i) the O(^{3}P) addition, ii) the subsequent H abstraction leading to an oxyradical intermediate, and iii) the CO elimination from the oxyradicals. We found that depending on the geometrical strain at the reaction site, the CO elimination takes place in either three steps, two steps, or a single step. Based on a HMOT analysis we could document that the changes in the resonance energy of

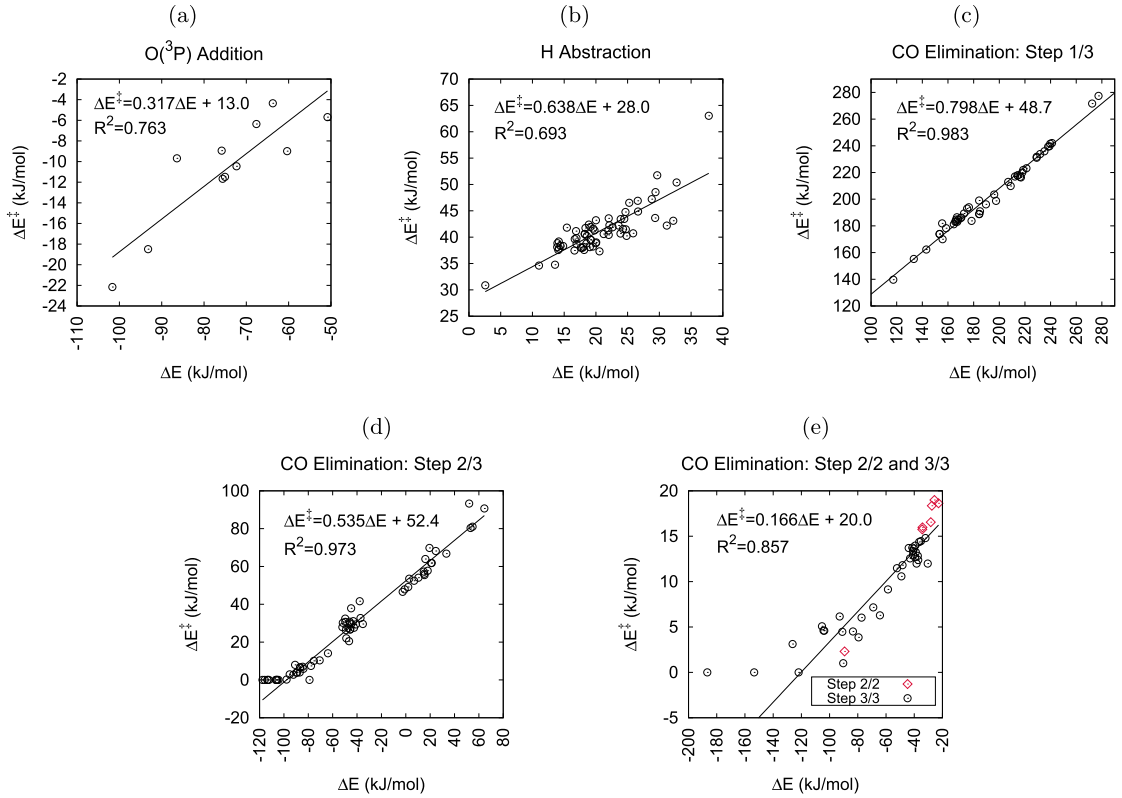


Fig. 8. BEP plots. (a) O(^{3}P) addition to PAHs, s = 2.7 kJ/mol, (b) following H abstraction to form oxyradical. (s = 2.5 kJ/mol) (c) First step of the three step CO elimination mechanism from the oxyradicals, s = 3.8 kJ/mol. (d) Second step of the three step CO elimination mechanism, s = 4.1 kJ/mol. (The points for which the activation energy is zero are excluded in the fit.) (e) CO elimination from carbonyl radicals, s = 1.9 kJ/mol. (The points for which the activation energy is zero are excluded in the fit.) The constant terms in the correlation equations has kJ/mol as units.

the aromatic π system are responsible for the different reactivities of the different reaction sites for the O(^{3}P) addition and the release of CO from carbonyl radicals.

The BEP correlation found for all the reactions studied in this work allows us to extrapolate the activation energies for other PAHs based on the reaction energies within ≈ ±4 kJ/mol for the rate limiting step of the three step CO elimination mechanism. This will facilitate the prediction of the activation energies for the studied reactions for larger systems, and most likely such BEP correlations can be transferred to other reactions that occur during the combustion of PAHs and the char burnout.

8. Associated content

The Supplementary Information includes tables of electronic energies calculated at TPSSh-D3/TZVP level, the zero point energy corrections, the contributions of vibrational, rotational, and translational partition functions to entropy and enthalpy at 1200 k and 0.1 MPa. Additionally, the BEP plots based on Gibbs free energy differences for some of the reaction steps are given.

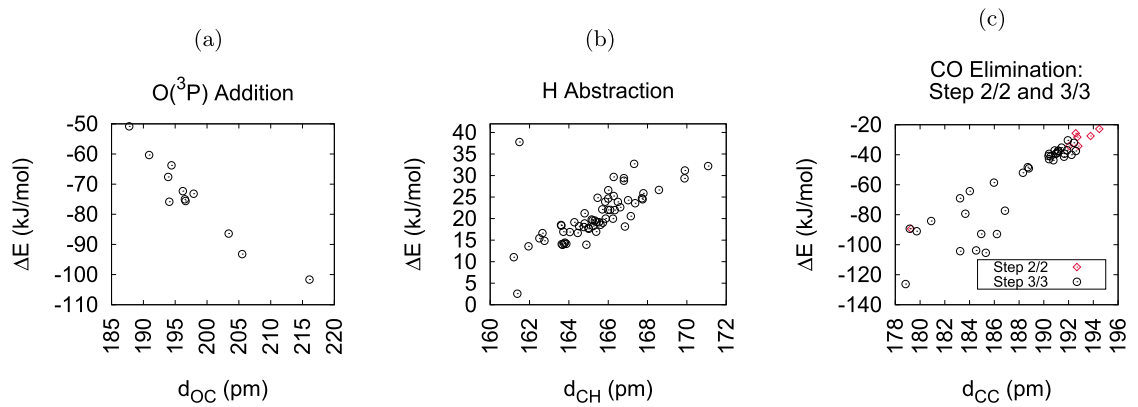


Fig. 9. Hammond postulate. (a) Reaction energy for the O(^{3}P) addition plotted versus the C–O bond length at the transition state. (b) Reaction energy for the H abstraction plotted versus the C–H bond length of the transition state structure. (c) Reaction energy for the CO elimination from carbonyl radicals plotted versus the bond length of the C–C bond that is broken for the transition state structure.

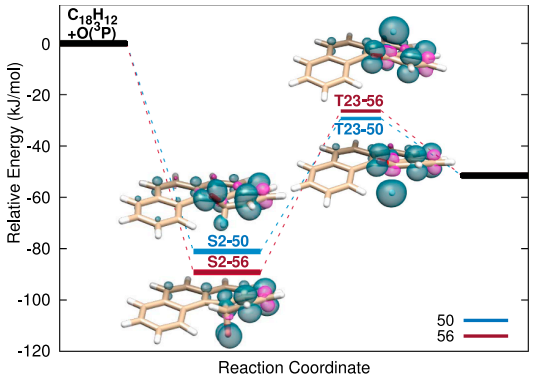


Fig. 10. Reaction path of oxyradical formation on the sites 50 and 56 at the fjord region of benzo[c]phenanthrene. The structures are shown with their spin densities.

Acknowledgment

This work was financially supported by the German Research Foundation (DFG) within the framework of the SFB/Transregio 129 Oxyflame.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, athttps://doi.org/10.1016/j.fuel.2018.12.047.

References

[1] Buhre B, Elliott L, Sheng C, Gupta R, Wall T. Oxy-fuel combustion technology for coal-fired power generation. Prog Energy Combust Sci 2005;31:283–307. https://doi.org/10.1016/j.pecs.2005.07.001.
[2] Keiluweit M, Nico PS, Johnson MG, Kleber M. Dynamic molecular structure of plant biomass-derived black carbon (biochar). Environ Sci Technol 2010;44(4):1247–53. pMID: 20099810. doi: https://doi.org/10.1021/es9031419.
[3] The combustion rates of coal chars: A review, Symposium (International) on Combustion 1982;19(1):1045–65, nineteenth Symposium (International) on Combustion. doi: https://doi.org/10.1016/S0082-0784(82)80281-6.
[4] Char structural ordering during pyrolysis and combustion and its influence on char reactivity, Fuel 2002;81(9):1215–25. doi: https://doi.org/10.1016/S0016-2361(02)00035-2.
[5] Harvey R. Polycyclic aromatic hydrocarbons: chemistry and carcinogenicity, cambridge monographs on cancer research. Cambridge University Press; 1991.
[6] Zhou C-W, Kislov VV, Mebel AM. Reaction mechanism of naphthyl radicals with molecular oxygen. 1. Theoretical study of the potential energy surface. J Phys Chem A 2012;116(6):1571–85. pMID: 22239650. doi: https://doi.org/10.1021/jp2119313.
[7] Emdee JL, Brezinsky K, Glassman I. A kinetic model for the oxidation of toluene near 1200 K. J Phys Chem 1992;96(5):2151–61. https://doi.org/10.1021/j100184a025.
[8] Brezinsky K. The high-temperature oxidation of aromatic hydrocarbons. Prog Energy Combust Sci 1986;12(1):1–24.
[9] Taatjes CA, Osborn DL, Selby TM, Meloni G, Trevitt AJ, Epifanovsky E, et al. Products of the benzene + O(3P) reaction. J Phys Chem A 2010;114(9):3355–70. pMID: 20099861. doi: https://doi.org/10.1021/jp9114145.
[10] Nguyen TL, Peeters J, Vereecken L. Theoretical reinvestigation of the O(3P) + C6H6 reaction: quantum chemical and statistical rate calculations. J Phys Chem A 2007;111(19):3836–49. pMID: 17253662. doi: https://doi.org/10.1021/jp0660886.
[11] Olivella S, Sole A, Garcia-Raso A. Ab initio calculations of the potential surface for the thermal decomposition of the phenoxyl radical. J Phys Chem 1995;99(26):10549–56. https://doi.org/10.1021/j100026a018.
[12] Liu R, Morokuma K, Mebel AM, Lin MC. Ab initio study of the mechanism for the thermal decomposition of the phenoxy radical. J Phys Chem 1996;100(22):9314–22. https://doi.org/10.1021/jp953566w.
[13] You X, Zubarev DY, Lester WA, Frenklach M. Thermal decomposition of pentacene oxyradicals. J Phys Chem A 2011;115(49):14184–90. pMID: 22054037. doi: https://doi.org/10.1021/jp208974b.
[14] Liu P, Lin H, Yang Y, Shao C, Gu C, Huang Z. New insights into thermal decomposition of polycyclic aromatic hydrocarbon oxyradicals. J Phys Chem A 2014;118(48):11337–45. pMID: 25386793. doi: https://doi.org/10.1021/jp510498j.
[15] Tao J, Perdew JP, Staroverov VN, Scuseria GE. Climbing the density functional ladder: nonempirical meta generalized gradient approximation designed for molecules and solids. Phys Rev Lett 2003;91:146401 https://doi.org/10.1103/PhysRevLett. 91.146401.
[16] Schäfer A, Horn H, Ahlrichs R. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. J Chem Phys 1992;97(4):2571–7.
[17] Staroverov VN, Scuseria GE, Tao J, Perdew JP. Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes. J Chem Phys 2003;119(23):12129–37. https://doi.org/10.1063/1.1626543.
[18] Schäfer A, Huber C, Ahlrichs R. Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. J Chem Phys 1994;100(8):5829–35.
[19] Grimme S, Antony J, Ehrlich S, Krieg H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. J Chem Phys 2010;132(15):154104 https://doi.org/10.1063/1.3382344.
[20] Csonka GI, Ruzsinszky A, Perdew JP. Estimation, computation, and experimental correction of molecular zero-point vibrational energies. J Phys Chem A 2005;109(30):6779–89. pMID: 16834032. doi: https://doi.org/10.1021/jp0519464.
[21] TURBOMOLE V7.1 2016 and V7.2 2017, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.
[22] Eichkorn K, Weigend F, Treutler O, Ahlrichs R. Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. Theoret Chem Acc 1997;97(1):119–24. https://doi.org/10.1007/s002140050244.
[23] Singh RI, Mebel AM, Frenklach M. Oxidation of graphene-edge six- and five-member rings by molecular oxygen. J Phys Chem A 2015;119(28):7528–47. pMID: 25894330. doi: https://doi.org/10.1021/acs.jpca.5b00868.
[24] Schleyer PVR, Manoharan M, Jiao H, Stahl F. The acenes: Is there a relationship between aromatic stabilization and reactivity? Organic Lett 2001;3(23):3643–6. pMID: 11700102. doi: https://doi.org/10.1021/ol016553b.
[25] The theory of reactions involving proton transfers, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 154 (882) (1936) 414–429. doi: https://doi.org/10.1098/rspa.1936.0060.
[26] Evans MG, Polanyi M. Inertia and driving force of chemical reactions. Trans Faraday Soc 1938;34:11–24. https://doi.org/10.1039/TF9383400011.
[27] Hammond GS. A correlation of reaction rates. J Am Chem Soc 1955;77(2):334–8. https://doi.org/10.1021/ja01607a027.

III. Can Small Polyaromatics Describe Their Larger Counterparts for Local Reactions? A Computational Study on the H-Abstraction Reaction by an H-Atom from Polyaromatics

The following article investigates H-abstraction reaction with an H-atom. This study links DFT and coupled cluster methods to approximate kinetic data for large systems via a composite scheme. Additionally, the predictive success of GFN-xTB methods and ReaxFF is investigated through comparisons to this data.
In this article, my contribution includes the selection of the test set, DFT , UHF-CCSD(T), ROHF-CCSD(T) calculations, calculation of the rate constants, and partially the ROHF-PNO-CCSD calculations. The analysis of these calculations are also done by me, together with discussions with Prof. Christof Hättig. The idea of using an average (T) correction and separation of the energy contributions came from Prof. Hättig. The GFN-xTB and ReaxFF calculations and their analysis are done by Dr. Gunnar Schmitz.

The following article is reprinted with permission from "Özlem Yönder, Gunnar Schmitz, Christof Hättig, Rochus Schmid, Paulo Debiagi, Christian Hasse, Andrea Locaspi, and Tiziano Faravelli, Can Small Polyaromatics Describe Their Larger Counterparts for Local Reactions? A Computational Study on the H-Abstraction Reaction by an H-Atom from Polyaromatics. The Journal of Physical Chemistry A, 124(46), 2020, Pages 9626-9637. DOI: 10.1021/acs.jpca.0c07133." Copyright 2020 American Chemical Society.

The raw data of this publication can be reached under DOI: 10.5281/zenodo.4122033.

THE JOURNAL OF PHYSICAL CHEMISTRY A

pubs.acs.org/JPCA Article

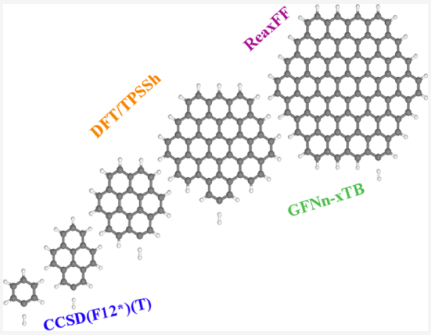
Can Small Polyaromatics Describe Their Larger Counterparts for Local Reactions? A Computational Study on the H-Abstraction Reaction by an H-Atom from Polyaromatics

Özlem Yönder,* Gunnar Schmitz,* Christof Hättig, Rochus Schmid, Paulo Debiagi, Christian Hasse, Andrea Locaspi, and Tiziano Faravelli

Cite This: J. Phys. Chem. A 2020, 124, 9626–9637 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Hydrogen abstraction is one of the crucial initial key steps in the combustion of polycyclic aromatic hydrocarbons. For an accurate theoretical prediction of heterogeneous combustion processes, larger systems need to be treated as compared to pure gas phase reactions. We address here the question on how transferable activation and reaction energies computed for small molecular models are to larger polyaromatics. The approximate transferability of energy contributions is a key assumption for multiscale modeling approaches. To identify efficient levels of accuracy, we start with accurate coupled-cluster and density functional theory (DFT) calculations for different sizes of polyaromatics. More approximate methods as the reactive force-field ReaxFF and the extended semi-empirical tight binding (xTB) methods are then benchmarked against these data sets in terms of reaction energies and equilibrium geometries. Furthermore, we analyze the role of bond-breaking and relaxation energies, vibrational contributions, and post-Hartree-Fock correlation corrections on the reaction, and for the activation energies, we analyze the validity of the Bell-Evans-Polanyi and Hammond principles. First, we find good transferability for this process and that the predictivity of small models at high theoretical levels is way superior than any approximate method can deliver. Second, ReaxFF can serve as a qualitative exploration method, whereas GFN2-xTB in combination with GFN1-xTB appears as a favorable tool to bridge between DFT and ReaxFF so that we propose a multimethod scheme with employing ReaxFF, GFN1/GFN2-xTB, DFT, and coupled cluster to cope effectively with such a complex reactive system.



INTRODUCTION

An accurate modeling of heterogeneous reactions of polyaromatic hydrocarbons with gas phase radicals is important for modeling and understanding combustion chemistry as well as soot formation. In particular, biochars are a diverse class of compounds due to the variations in the feedstock and due to the varying charring intensity applied in the production, which both affect its composition and structure.[1−4] Therefore, the effects of the structure and size of the aromatic compounds on the elementary reaction steps of biochar combustion is an important aspect for modeling its combustion.

A lot is known about reactions for small molecules,[5] but when it comes to the combustion of solid fuels such as biochar, accurate calculations on the full heterogeneous reaction systems or even on large model compounds are no longer possible. Their description is a complex multiscale problem. It requires on the small scale accurate first-principle quantum chemistry methods to obtain accurate thermodynamic data for elementary reaction steps. On the (super-)nanometer scale, force-field or coarse-grained force-field methods[6] have to be applied to generate realistic structure models, and in between, efficient tools are needed to select, out of the myriads of, in principle, thinkable reaction steps, those that could be relevant. In the context of heterogeneous combustion, where covalent bond breaking needs to be described, reactive force fields, for example, ReaxFF,[7] developed by van Duin et al., are often used to deal with the mere size of the system, however, oftentimes relying solely only on this approximate method.[8−10]

The limitation of accurate quantum chemistry methods to small systems raises the question whether, or to which extent, an analogy between gas phase reactions and heterogeneous reactions exists.[11] Detailed or semi-detailed mechanisms as commonly employed in computational fluid dynamics simulations involving heterogeneous gas/solid reactions adopt rate constants derived from analogous gas-phase reactions[12] or accounting for proper corrections proposed about 30 years ago

Received: August 4, 2020
Revised: September 17, 2020
Published: November 4, 2020



Downloaded via BOCHUM LIBRARIES on April 9, 2021 at 12:22:44 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

ACS Publications © 2020 American Chemical Society

https://dx.doi.org/10.1021/acs.jpca.0c07133
J. Phys. Chem. A 2020, 124, 9626–9637

by Frenklach and Wang.[13] These assumptions require specific validations. Some preliminary results show that the activation energies of H-abstractions on graphene edges by different radicals can be related to those of the analogous reactions in the gas phase.[11] For a multiscale modeling of biochar combustion, important questions are thus how transferable are results for reaction and activation enthalpies on small polycyclic aromatic hydrocarbons (PAHs) to larger systems and to which extent can results that are obtained with cost-efficient force-field and semi-empirical methods be combined with accurate electronic structure calculations. With this study, we aim to contribute to this question with the specific example of the hydrogen abstraction reactions of PAHs. This reaction has the importance as an initiation step for combustion and especially for soot formation through the hydrogen abstraction C_2H_2 addition (HACA) mechanism.[14]

In a recent study on the H-abstraction reactions on armchair and zigzag edges of graphene, it has been shown that the interactions between the layers do not affect the transition state and reaction barriers.[11] Therefore, we do not describe here the effect of the different numbers of layers on this reaction. In the current study, we use a test set of PAHs ranging from benzene to $C_{96}H_{24}$ that, apart from varying in size, exhibit a variety of reaction sites that represent different situations such as zigzag, armchair, or free edges.[15] As quantum chemistry electronic structure methods, we employ density functional theory (DFT) with a modern meta-GGA hybrid functional. For the smaller PAHs, these are combined with single point coupled-cluster (CC) calculations. Apart from the costly CC calculations which are only possible for small- and medium-sized molecules, a bottleneck for larger PAHs are the searches for the transitions states. Here, the questions are (1) how transferable are differences between DFT and CC energies from the small to the larger PAHS and (2) how well the activation energies correlate with the reactions energies (Bell−Evans−Polanyi principle),[16,17] that is, how accurate can activation energies be estimated based on it?

Individual local reaction can be targeted directly with static DFT and coupled-cluster calculations. A route to complex sequences of reactions, as they occur during the combustion of char, which is tractable for first-principles calculations is to preselect the most important elementary steps based on reactive molecular dynamics. This requires cost-efficient semi-empirical electronic structures or reactive force-field methods. To investigate their performance in the context of combustion, we do a similar assessment for the extended semi-empirical tight-binding methods xTB (GFN0-xTB, GFN1-xTB, and GFN2-xTB)[18,19] of Grimme and co-workers as well as for the reactive force-field ReaxFF.[7] Again, the aim is to compare these more approximate methods with the more accurate DFT and CC results. We compare reaction energies and geometrical similarity of stationary points on the potential energy surface (PES) such as minima and transition states. For modeling the combustion of solid fuels, it is vital to have cost-efficient methods which qualitatively and semi-quantitatively reproduce the correct PES. In this case, those methods can be used, as pointed out above, for prescreening or, for less local reactions, combined in hybrid approaches, where larger parts are optimized using ReaxFF or xTB and smaller active sites are optimized using DFT. Furthermore, only these methods scale efficiently enough to be applied in large scale molecular dynamic simulations to get an understanding on the atomistic scale of the processes involved in combustion.

COMPUTATIONAL DETAILS

For the current study, we have selected a test set consisting of 13 PAHs which range in size from benzene to $C_{96}H_{24}$, as shown in Figure 1. These PAHs cover a variety of reaction sites such as

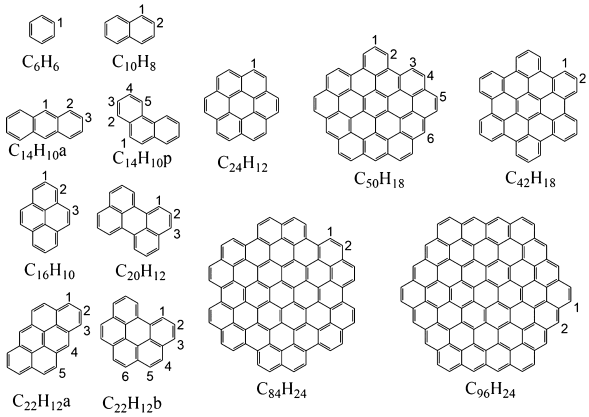


Figure 1. Selected reaction sites.

sterically unhindered sites at zig-zag, free edges (e.g., in $C_{96}H_{24}$), and sterically hindered bay region sites at armchair edges (e.g., the sites 1 in $C_{20}H_{12}$, $C_{42}H_{18}$, or $C_{84}H_{24}$). For details of the nomenclature for the sites, see Supporting Information.

All geometries have been optimized at the DFT level (vide infra). At the same level, we have also calculated the analytical second derivative of the PES with the harmonic approximation, from which the vibrational contributions are evaluated. For a few cases, we carried out explicitly correlated coupled cluster, CCSD(F12*)(T), and for a larger subset local pair-natural orbital coupled cluster, PNO-CCSD, calculations.

In this study, the performance of the DFT functional is important for the calculation of zero point vibrational energies (ZPVEs) and the partition functions, as well as for observing the general trends in the activation and reaction energies. Therefore, we have used the TPSSh functional. In a previous study, it was stated that for the H abstraction reaction with the H atom and methyl radical, the TPSSh functional has a mean absolute error of up to 38 kJ/mol as compared with G3(MP2)-RAD results; therefore, the PBE0 functional was recommended.[20] On the other hand, based on the C3/99 test set, the TPSSh functional performs as good as the most common hybrid functionals, such as PBE0 or B3LYP, for bond lengths, angles, bond dissociation energies, and the harmonic vibrational frequencies, and it outperforms them for the calculation of standard enthalpies of formation.[21]

The DFT calculations have been performed using TURBOMOLE's m5 integration grid[22] (except for reactions of $C_{84}H_{24}$ and $C_{96}H_{24}$ for which m3 grid is used), TZVP basis set,[23] and Grimme's D3 dispersion correction[24] and were performed with the dscf, grad, and aoforce programs of the TURBOMOLE package.[25] The zero-point vibrational energy (ZVPE) corrections to the electronic energies were computed at the same electronic structure level (TPSSh-D3/TZVP) within the harmonic approximation. We have used the scaling factor of 0.9874 for TPSSh/6-31+G(d,p) to fit the experimental ZPVE.[26]

All CC calculations employed the frozen core approximation for the $1s^2$ electrons on carbon. To obtain it for the smaller molecules, results close to the basis set limit, we used the

https://dx.doi.org/10.1021/acs.jpca.0c07133
J. Phys. Chem. A 2020, 124, 9626−9637

Table 1. Deviations in the C−H Bond Distance (in pm) and C−C−C Angle (in deg) for the Semi-empirical and FF Methods Compared to the TPSSh-D3/TZVP Reference Structure[a]

	PAHs				H-abstracted			
	GFN0-xTB	GFN1-xTB	GFN2-xTB	ReaxFF	GFN0-xTB	GFN1-xTB	GFN2-xTB	ReaxFF
				C−H Distance				
|MAX|	0.220	0.130	0.510	1.400	0.240	0.220	0.590	1.480
MEAN	−0.039	0.006	0.438	0.757	−0.046	0.006	0.432	0.657
RMSD	0.069	0.056	0.439	0.818	0.081	0.069	0.433	0.768
				C−C−C angle				
|MAX|	1.49	1.54	1.57	3.49	5.22	4.95	4.96	28.35
MEAN	0.04	0.07	0.06	−0.08	0.01	0.03	0.02	−0.48
RMSD	0.28	0.32	0.29	0.86	0.66	0.41	0.39	4.46

[a]It is differentiated between the native PAHs structures and the H abstracted radical structures.

explicitly correlated CCSD(F12*) approach[27] with the aug-cc-pVTZ orbital,[28] corresponding optimized complementary auxiliary,[29] and density-fitting basis[30] sets. The reference Hartree−Fock determinants are calculated using the dscf program,[31] the CC singles, doubles, and perturbative triples, CCSD(F12*)(T), and calculations for the reactions of benzene and naphthalene are performed using the ccsdf12 program.[25] For the PAHs up to $C_{24}H_{12}$, we did local pair natural orbital (PNO) based ROHF-CCSD, also known as UR-CCSD, calculations in the cc-pVTZ basis using a local development version of the pnoccsd program.[32−34] For anthracene and phenanthrene, we did in addition PNO-UHF-CCSD(F12*)(T) calculations to investigate the trend of coupled-cluster methods based on UHF reference determinants with the size of PAHs.

The ReaxFF calculations were performed using the reimplementation of ReaxFF[35] available in the LAMMPS Molecular Dynamics Simulator[36] and using the "CHO" parameterization for combustion, as published in ref 37. The xTB calculations were done with version 6.3.1 of the xTB code.[18,19,38] If not otherwise stated, we used in all calculations the standard settings in xTB. For the transition state optimization with xTB, we interfaced it to the TURBOMOLE reaction pathfinder statpt[39] and the transition state optimizer statpt.

The reaction rate constants have been calculated using transition state theory (TST). The partition functions are calculated based on a modification to the quasi rigid rotor and harmonic oscillator approximation.[40] The modification is that for the low lying modes (ω < 300 cm^{-1}), the rotational and harmonic oscillator approximations are combined for partition functions instead of the entropy. In other words, for the low lying modes, the Head-Gordon damping function ($W(\omega)$) is applied to its contribution q to the partition function as

$$q = q_{vib}^{W(\omega)} q_{rot}^{1-W(\omega)} \tag{1}$$

instead of to the entropy as in the original approach[40] given by

$$S = W(\omega)S_{vib} + [1 - W(\omega)]S_{rot} \tag{2}$$

Using the ideal gas law as the equation of state, the site-specific forward reaction rate constants without using symmetry numbers are calculated at 1 bar based on the TST according to

$$k_{forward}(T) = \kappa \frac{RT}{P}\frac{k_B T}{h}\frac{Q_{TS}}{Q_{PAH}Q_H}\exp\left(-\frac{E_{el,act}}{RT}\right) \tag{3}$$

where R is the gas constant, P the pressure, k_B the Boltzmann's constant, T the temperature, h Planck's constant, and Q_{TS}, Q_{PAH}, and Q_H are the partition functions which include translational, vibrational, and rotational contributions, respectively. The vibrational partition function is defined relative to the bottom of the PES, and $E_{el,act}$ is the electronic activation energy (without the ZPVE). κ is the transmission coefficient used to account for tunneling. It is calculated using the high temperature limit of the Eckart tunneling correction, which asymptotically approaches to Wigner tunneling correction[41]

$$\kappa = 1 - \frac{1}{24}\left(\frac{h\nu}{k_B T}\right)^2\left(1 + \frac{RT}{E_{el,act}}\right) \tag{4}$$

In eq 4, ν is the imaginary frequency at the transition state.

RESULTS

Geometrical Accuracy of Semi-empirical Methods and Reactive Force Fields. DFT methods give (usually) accurate geometries and, with hybrid functionals, reasonable reaction and activation energies but are, although computationally much less demanding than coupled-clusters methods, still computationally costly. If one is interested in large scale atomistic simulations or molecular dynamic simulations on long time scales, there is no way around more approximate but numerically more efficient methods. Here, we want to answer the question if the extended tight binding models xTB and the reactive force-field ReaxFF can be used as a bridge from DFT to larger scales. In a first step, we compare how well equilibrium geometries on the xTB and ReaxFF level agree with the DFT results and later judge on how accurate reaction energies are reproduced.

For the first purpose, we reoptimized the PAH structures and the H-abstracted structures on this level and compared the C−H bond lengths and the C−C−C angles in the aromatic rings. The reoptimization will result in smaller deviations than an optimization from other starting guesses, but this is intended. We want to assess how much the semi-empirical methods need relaxation in the found DFT minima.

The results are given in Table 1. The trend for the bond lengths is quite interesting: GFN0-xTB and GFN1-xTB reproduce the DFT equilibrium C−H bonds quite well, and the deviations increase for GFN2-xTB. This, however, does not translate into a better agreement in reaction energies for GFN0 and GFN1 (vide infra). For GFN2-xTB, we observe a RMSD in the C-H bond length of 0.43 pm for the H-abstracted species and 0.44 pm for the PAHs. This is a very consistent error for different molecular species (radical and no radical). For ReaxFF, we see a slightly more unbalanced description. For the PAHs, a RMSD of 0.082 pm is observed, while the C-H distances for the H-abstracted species deviate by an RMSD of 0.768 pm. When looking at the C−C−C angles, we see for the xTB methods a

https://dx.doi.org/10.1021/acs.jpca.0c07133
J. Phys. Chem. A 2020, 124, 9626−9637

quite good agreement, but the deviations are typically larger for the H-abstracted species. This especially holds for ReaxFF, where a RMSD of 4.43° is observed. For the radical structure, ReaxFF favors straightened C−C−C angles, which distorts the ring structure, which is presumably due to the fact that in ReaxFF, there is no concept of a radical, and a two coordinate carbon naturally tends toward a linear geometry (sp-hybridization).

It turned out to be difficult to optimize the transition states using the xTB methods and ReaxFF. For the latter case, we observed a rather non-smooth PES with large changes of the gradients in the region of the barrier, explaining the short timesteps necessary for energy-conserving MD. We located the barrier by a nudged elastic band calculation for the benzene case but refrained from an investigation of larger PAHs. Using the xTB methods, it was more straightforward to optimize transition states, but only GFN1-xTB showed a good agreement with the TPSSh-D3/TZVP-optimized structures, and since our ultimate goal is to bridge the scale from DFT calculations to larger scales, we propose to use GFN1-xTB for geometry optimizations and GFN2-xTB for energies. The results for geometrical deviations of the GFN1-xTB-optimized transition states are listed in Table 2. Especially, for the C−H bond distance (for the abstracted H atom), the deviations are larger compared to our results for the PAHs, but we see a reasonable agreement.

Table 2. Deviations in the C−H Bond Distance (in pm) and C−C−C Angle (in deg) for GFN1-xTB Compared to the TPSSh-D3/TZVP Reference Structure for the Transition State

	C−H	C−C−C
|MAX|	42.48	4.06
MEAN	−3.69	−1.26
RMSD	7.25	1.50

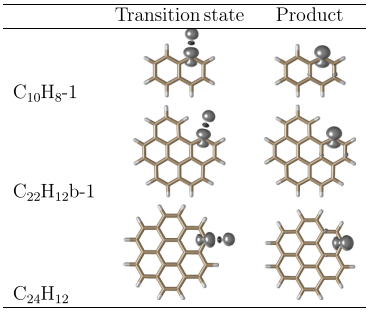
Trends from DFT. To study the abstraction of a hydrogen from PAHs, we considered the following reactions with atomic hydrogen

$$C_nH_m + H \rightarrow C_nH_{m-1} + H_2 \quad (5)$$

The reaction does not disturb the resonance in the π-system and should thus cause only local changes in the geometry and the electronic structure. As shown for some examples in Table 3, the spin density at the transition states and for the H-abstracted radicals looks very alike for any given system in this study.

Due to the local electronic character of the reaction, the electronic reaction and activation energies for this reaction are expected to be very similar and independent of the size of the polyaromatic. In accordance with this, the electronic activation energies differ only by 3.2 kJ/mol. The electronic activation energy is on average 35.6 kJ/mol, while the lowest one is 34.8 kJ/mol for $C_{22}H_{12}$a-2 and the largest one is 37.9 kJ/mol for $C_{84}H_{24}$-2. Even if the electronic activation energies are similar, the reaction energies at the sterically hindered sites ($C_{14}H_{10}$p-5, $C_{20}H_{12}$-1, $C_{22}H_{12}$b-1, $C_{42}H_{18}$-1, $C_{50}H_{18}$-2, $C_{50}H_{18}$-3, and $C_{84}H_{24}$-1) are lower than those at the sterically unhindered sites, with an average electronic reaction energy of 14.0 kJ/mol, whereas the average of those which are not sterically hindered is 21.2 kJ/mol. This can be explained by the unfavorable repulsion between the neighboring H atoms in the bay region sites, which is removed by the H-abstraction.

Table 3. Spin Densities of the Transition States and the H-Abstracted Product Radicals for Some Example Reaction Sites Calculated with DFT/TPSSh-D3/TZVP



The two determining contributions to the electronic reaction energies—for which one can expect a different sensitivity to the electronic structure method and the model compound size—are (1) the changes in the electronic structure upon removal of the H atom, that is, unrelaxed C−H bond breaking and H−H bond formation, and (2) the relaxation of the structure of the H-abstracted PAH during the geometry optimization, which is exothermic. To measure these two contributions, we calculated in addition to the reacting species also the energies of the H-removed educt structures ($E_{\text{H-removed PAH}}$) and defined the mentioned two contributions as

$$E_{C-H} = E_{\text{el,H-removed PAH}} + E_{\text{el},H_2} - E_{\text{el,PAH}} - E_{\text{el,H}} \quad (6)$$

and

$$E_{\text{relax}} = E_{\text{el,geoopt. H-abstracted PAH}} - E_{\text{el,H-removed PAH}} \quad (7)$$

which sum up to the electronic reaction energy for reaction 5

$$E_{\text{el,rxn}} = E_{C-H} + E_{\text{relax}} \quad (8)$$

Figure 2 shows the contributions to the reaction energy for different sites in a stack plot. E_{C-H} depends on the local environment of the site where the hydrogen is removed, which is strongly supported by the fact that this contribution is on average 7.0 kJ/mol less endothermic for the bay region reaction sites than those which are not sterically hindered. Additionally, it can be seen that E_{relax} does not vary much between different sites or with the size of the polyaromatics. This is also the case for the zero-point vibrational contribution to the reaction energy, which for our test set is −6.9 kJ/mol with a standard deviation of ±0.8 kJ/mol and a maximum deviation of ±2 kJ/mol. Therefore, for the H abstraction reaction, variations in the reaction energies are mainly due to unrelaxed bond breaking and formation energies, which are dependent on the local environment of the reacting site but are also independent of the size of the polyaromatic.

The reaction and activation energies of the sites without steric hindrance follow the Bell−Evans−Polanyi principle with an R^2 value of 0.9, and for the bay region sites, the correlation is less good with an R^2 value of 0.8, (Figure 3). The fact that the sterically hindered bay region sites do not fit into the BEP correlation found for the sites without steric hindrance has also been observed in previous studies.[42] Figure 4 shows that there is a correlation between the C−H distance at the transition states and the reaction energies (Hammond's postulate). Here, for the bay region sites, there is no correlation because the range of reaction energies is too small (2 kJ/mol) to observe this.

https://dx.doi.org/10.1021/acs.jpca.0c07133
J. Phys. Chem. A 2020, 124, 9626−9637

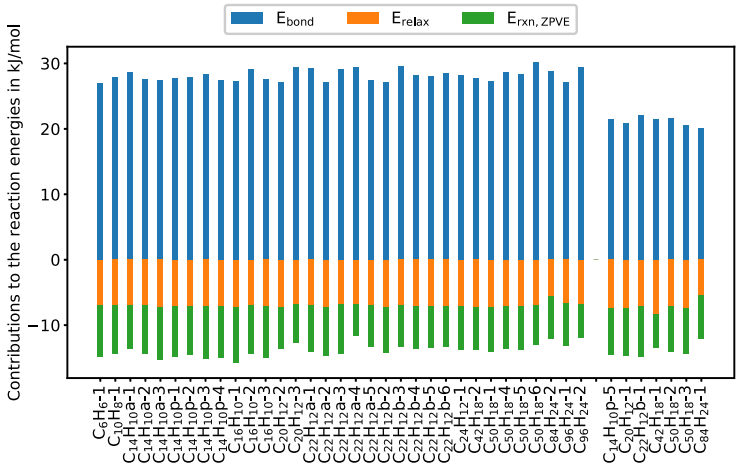


Figure 2. Stack plot of the contributions to the electronic reaction energy at the DFT/TPSSh-D3 level. The ZPV energies are stacked on top of the relaxation energies. Note that the last seven sites are the bay region sites.

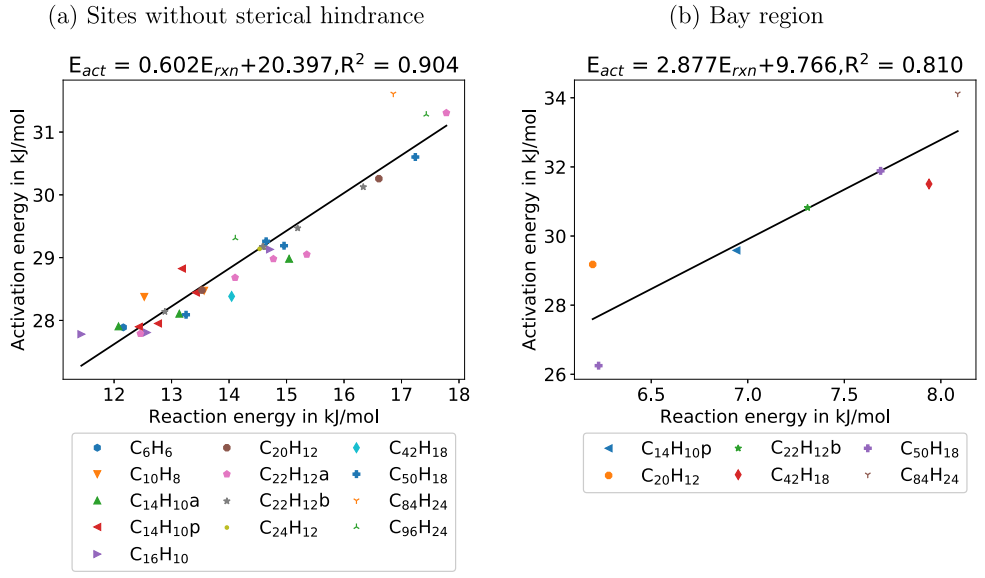


Figure 3. Relation between the reaction and activation energies (E_{rxn} and E_{act}) based on the TPSSh-D3/TZVP results.

Assessment of Semi-empirical and Reactive Force-Fields. Next we turn to a comparison of the xTB and ReaxFF reaction energies to our more accurate results, which will be presented in the subsequent sections. First, we will do a comparison to our TPSSh-D3/TZVP results and a composite scheme, based on PNO-CCSD and an extrapolation estimate for larger PAHs, which will be discussed in the sections "Coupled-Cluster Energy Corrections" and "Extrapolation to Larger PAHs". We solely concentrate on the electronic contribution to the reaction energy to allow for a better comparison. First, we will discuss the examples given in Table 4. In these cases, we see that GFN0-xTB is qualitatively not correct by a large magnitude, and GFN1-xTB starts to be at least qualitatively correct but underestimates the reaction energy nearly by a factor of 10. However, GFN2-xTB on the other hand shows a good agreement with our CC results but overestimates the reaction energy compared to the DFT results by a factor of two. Table 5 shows the average deviations of the xTB methods to the TPSSh-D3/TZVP results and the composite scheme. As for the already discussed example, GFN0-xTB is qualitatively not correct, and also, on average, GFN1-xTB is qualitatively not correct, but this is the result of a few outliers. For GFN2-xTB, the reaction energies are always qualitatively correct, but the average deviation is around 26 kJ/mol to the DFT reference and 6 kJ/mol to the composite scheme. Typically, the GFN2-xTB reactions energies are a factor of 2−2.5 larger than the DFT results and only slightly above the more accurate results given by the composite scheme. This is a positive finding since it allows us to use GFN2-xTB for prescreening of reactions and to use DFT or CC only for more accurate calculation for the reactions kept after the pre-screening. Although in our study GFN2-xTB seems to be more accurate than TPSSh-D3, we do not claim that this holds in general, and although the reaction energies are in better agreement with our most accurate reference, there were some problems with calculating activation energies on the GFN2-xTB level of theory.

https://dx.doi.org/10.1021/acs.jpca.0c07133
J. Phys. Chem. A 2020, 124, 9626−9637

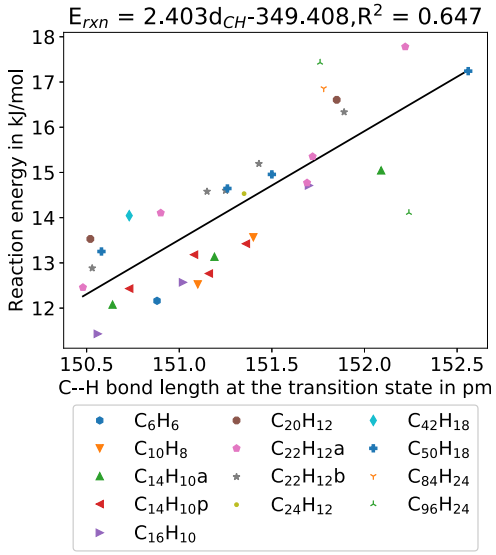


Figure 4. Reaction energies as a function of the C−H bond length of the transition state based on the TPSSh-D3/TZVP results for the sites without sterical hindrance.

In Table 4, we also give results for ReaxFF reaction energies. They are always qualitatively correct but overestimate the reaction energy on average by a factor of 4−5 with respect to the DFT reference. Compared to the composite scheme, the reaction energies improve a bit but are still 2.5−3.5 times larger than the reference values. Furthermore, we see that the reaction energy for ReaxFF is very site-dependent, which originates from ReaxFF favoring a sp^2 hybridization. Due to steric hindrance, this relaxation is not possible in all cases. The sites $C_{10}H_8$-1 and $C_{10}H_8$-2 are a good example since their reaction energies differ more than a factor of two. For $C_{10}H_8$-1, we can measure a C−C−C angle of 147.5°, while for $C_{10}H_8$-2, we observe an angle of 151.6°. Therefore, the product could in the second case relax more to its preferred state, which gives a lower reaction energy. One can conclude that ReaxFF should be used for aromatic ring-like systems in case of radicals with caution. Since the functional form for ReaxFF is not that flexible it is questionable if a different parameterization can cure this behaviour.

Besides the total electronic reaction energy, we were interested in the contribution, which stems from relaxation. This answers the question how well relative changes in a particular configuration are transferable from the semi-empirical methods to DFT. Table 6 shows the statistics over the deviations to the DFT reference. One has to note that the statistics look much better than that for the reaction energies, but still the hierarchy GFN0-xTB < GFN1-xTB < GFN2-xTB is retained. GFN1-xTB has a slightly better RMSD, but its outliers are smaller. The xTB methods seem to be accurate enough to describe the configurational changes. The relaxation in the ReaxFF optimized structures is approximately 10 times larger than that in the DFT reference, which results in large deviations with an RMSD around 75 kJ/mol, but this finding is not completely unexpected. When analyzing the deviations in the geometries, we already found that ReaxFF tends to straighten the C−C−C angle, which results in a large rearrangement and thus to a large relaxation energy.

In accordance with the overestimated reaction energies, also the reaction barrier for C_6H_6-1 is with 89 kJ/mol overestimated by a factor of about 3 for ReaxFF. In contrast to the electronic structure methods (xTB and DFT), the reactive force field also results in a non-linear arrangement of the C−H−H group at the barrier with an angle of 147°. For the assessment of the xTB reaction barriers, we used the GFN1-xTB geometries in combination with GFN1-xTB and GFN2-xTB energies to calculate the barrier heights. We compare the electronic contribution and the ZPVE contribution. A statistic for the comparison to the TPSSh-D3/TZVP results can be found in Table 7. The xTB methods underestimate the electronic contribution to the activation energy compared to the reference. GFN1-xTB predicts on average 5.9 times lower activation energies, while GFN2-xTB predicts on average 1.5 times lower barriers. For the contribution of the ZPVE, GFN1-xTB and GFN2-xTB have comparable errors but with different trends. GFN1-xTB underestimates the energies, while GFN2-xTB overestimates them. Excluding $C_{10}H_8$-1 as outlier, GFN2-xTB gives ZPEs which are on average 1.6 times larger than the TPSSh-D3/TZVP ones, while GFN1-xTB gives 0.4 times smaller ones.

Based on our gathered data, we can conclude that a combination of GFN1-xTB and GFN2-xTB is a good candidate to bridge the presented CC and DFT calculations to larger scales for the energetics GFN2-xTB is close to our CC and DFT reference, while GFN1-xTB performs better for geometry optimizations. The failure of GFN0-xTB is unfortunate since it is roughly 10 times more efficient than GFN2-xTB which would greatly improve the simulation time in MD studies, but in the case of H abstraction, GFN0-xTB overestimates the stability of molecular hydrogen. When rescaling the energy of atomic hydrogen to mimic the H_2 binding energy of CCSD(F12*), we obtain reaction energies slightly above those for GFN1-xTB. For individual reactions, such a correction term is feasible but not when used in MD simulations. GFN1-xTB and GFN2-xTB are more accurate than ReaxFF for our purposes, but since ReaxFF is still at least by a factor of 10 more efficient than GFN0-xTB, it can be a viable building block for a multiscale approach.

Table 4. Comparison of the Reaction Energies (in kJ/mol) on the Semi-empirical Level to the TPSSh-D3/TZVP, PNO-CCSD/cc-pVTZ, and ROHF-CCSD(F12*)(T)/aug-cc-pVTZ Results

	$E_{el,rxn}$						
site	GFN0-xTB	GFN1-xTB	GFN2-xTB	ReaxFF	PNO-CCSD	CCSD(F12*)(T)	TPSSh
$C_{10}H_8$-1	−189.2	4.5	49.3	107.6	47.3	43.6	20.94
$C_{10}H_8$-2	−188.0	4.9	49.7	49.0	46.8	42.7	27.60
$C_{14}H_{10}$a-1	−191.3	4.0	48.6	146.4	42.8		16.44
$C_{14}H_{10}$a-2	−189.8	4.2	48.8	105.2	41.1		15.35
$C_{14}H_{10}$a-3	−188.5	4.8	49.6	63.2	40.4		14.92
C_6H_6-1	−186.1	5.7	50.5	30.5	39.9	42.7	12.20

https://dx.doi.org/10.1021/acs.jpca.0c07133
J. Phys. Chem. A 2020, 124, 9626−9637

Table 5. Statistics of the Deviations in the Electronic Contribution of Reaction Energies (in kJ/mol) of the Semi-empirical Methods and ReaxFF Compared to Either the TPSSh-D3/TZVP Reference Values or Our Most Accurate Reaction Energies Based on the Composite Scheme, Presented in the Subsequent Sections[a]

	TPSSh-D3 reference				composite reference			
	GFN0	GFN1	GFN2	ReaxFF	GFN0	GFN1	GFN2	ReaxFF
|MAX|	220.94	27.55	30.09	154.55	239.80	45.04	9.71	117.68
MEAN	−213.36	−18.18	26.08	79.98	−228.71	−35.16	4.46	51.74
RMSD	213.38	18.50	26.33	87.03	230.73	38.73	5.88	60.07

[a]We only give the xTB parameterization GFN*n* and drop "xTB".

Table 6. Error Statistics for the Relaxation Contribution to the Reaction Energy (in kJ/mol) for the Different Semi-empirical Methods and ReaxFF Where the TPSSh-D3/TZVP Results Serve as the Reference[a]

	GFN0-xTB	GFN1-xTB	GFN2-xTB	ReaxFF
		Non-Bay Area		
|MAX|	6.56	5.56	3.87	135.78
MEAN	3.67	1.11	1.17	−74.47
RMSD	3.88	1.82	1.81	77.49
		Bay Area		
|MAX|	4.29	0.99	1.12	76.94
MEAN	4.00	0.73	0.96	−74.61
RMSD	4.01	0.77	0.96	74.63

[a]We distinguish between the bay area and the non bay area.

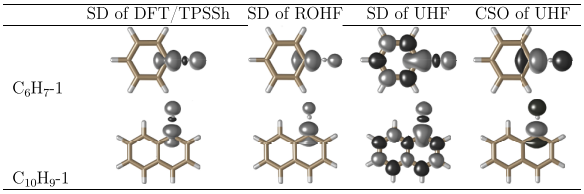
Table 7. Statistics for the Deviations in Activation Energies (in kJ/mol) for the xTB Methods Compared to the TPSSh-D3/TZVP Reference

	GFN1-xTB	GFN2-xTB
	Electronic Contribution	
MAX	35.09	46.43
MEAN	−29.61	−17.34
RMSD	29.70	18.23
	ZPVE Contribution	
MAX	8.13	13.12
MEAN	4.36	−4.37
RMSD	4.42	5.38

Coupled-Cluster Energy Corrections. For correlated wavefunction calculations, apart from the method, two important questions are the basis set convergence and, for radicals, the treatment of the spin. Single-reference coupled-cluster calculations can for radicals be based on unrestricted (UHF) or restricted open-shell (ROHF) Hartree−Fock reference determinants. As mentioned above, it was found at the DFT level that at the transition state and in the product, the spin density of the radical is localized close to the reaction site, essentially in a singly occupied σ-orbital. For the H-abstracted PAH radicals and the transitions states, we observe for unrestricted Hartree−Fock a strong spin-polarization of the orbitals, which also affects the π-system and leads to a delocalization of the spin-density over a large part of the PAH (see Table 8).

We performed UHF-based calculations at the UHF-CCSD(F12*)(T)/aug-cc-pVTZ level for benzene and naphthalene and with PNO-UHF-CCSD(F12*)(T)/aug-cc-pVTZ for phenanthrene. The results are listed in Table 9. In contrast to what one would expect based on the results from DFT, the UHF-coupled-cluster singles-and-doubles with perturbative triples [CCSD(T)] calculations predict that the activation and reaction energies would increase with the size of the aromatic ring system, for the reaction energy, by ≈16−17 kJ/mol from C_6H_6 to $C_{10}H_8$ and ≈14−23 kJ/mol from $C_{10}H_8$ to $C_{14}H_{10}$. In particular, the perturbative (T) triples correction is for the UHF-based calculations sizeable and grows with the system size.

Table 8. Spin Densities (SD) of the Transition State Structures for Example Reaction Sites Calculated with DFT/TPSSh-D3, ROHF, and UHF[a]



[a]The last column shows the corresponding spin orbital (CSO) of the UHF wavefunction with occupation closest to 1.

Table 9. UHF-CCSD(F12*)(T)/aug-cc-pVTZ Reaction Energies

	$E_{el,act}$	$E_{el,rxn}$	$(T)_{act}$	$(T)_{rxn}$
C_6H_6	81.7	51.8	−1.9	3.1
$C_{10}H_8$-1	96.4	67.8	1.2	6.7
$C_{10}H_8$-2	90.3	67.0	−3.6	6.5
$C_{14}H_{10}$p-1	109.9	80.5	3.3	8.2
$C_{14}H_{10}$p-5	110.3	75.0	3.1	8.7
$C_{14}H_{10}$a-1	116.6	87.7	9.4	14.4
$C_{14}H_{10}$a-2	117.1	87.3	9.4	14.0
$C_{14}H_{10}$a-3	116.7	86.5	9.3	13.9

The counter-intuitive behaviour, however, is an artifact of the spin-polarization and the unphysical delocalization of the spin density over the aromatic ring systems in the UHF reference. With ROHF reference determinants, the spin density is qualitatively similar to that obtained with DFT/TPSSh, as shown in Table 8. The ROHF-CCSD(F12*)(T) results for reaction energies of benzene and naphthalene are shown in Table 10.

Similar as at the DFT level and in agreement with results from Carissan and Klopper[20] at the G3(MP2)-RAD level, which also employs ROHF-CCSD(T) as part of the computational protocol, these results do not increase with the size of the aromatic ring system. Also, the (T) triples correction, which is the most severe computational bottleneck in the coupled-cluster calculations, varies very little between benzene and the two reaction sites at naphthalene. Explicitly correlated CCSD(F12*)/aug-cc-pVTZ results are usually close to the basis set limit, comparable to conventional CCSD/cc-pV5Z results.[27,43] For the reaction energies of H-abstraction from PAHs, both the basis set limits extrapolated from cc-pVDZ and cc-pVTZ result[44]

https://dx.doi.org/10.1021/acs.jpca.0c07133
J. Phys. Chem. A 2020, 124, 9626−9637

Table 10. ROHF-CCSD(T) Reaction Energies and the Contribution of the Triples Correction, (T), to Them in kJ/mol[a]

method	site	$E_{el,act}$	$E_{el,rxn}$	E_{act}	E_{rxn}	$(T)_{act}$	$(T)_{rxn}$
CCSD(T)/cc-pVDZ	C_6H_6	76.5	46.9	69.5	39.0	−3.3	0.7
	$C_{10}H_8$-1	76.8	48.1	70.2	40.7	−3.7	0.7
	$C_{10}H_8$-2	76.6	47.6	69.9	39.6	−3.4	0.7
CCSD(T)/cc-pVTZ	C_6H_6	72.4	41.5	65.4	33.6	−3.6	1.5
	$C_{10}H_8$-1	72.5	42.8	65.9	35.4	−4.0	1.6
	$C_{10}H_8$-2	72.4	42.1	65.7	34.2	−3.7	1.5
extrapolated[b]	C_6H_6	70.9	41.3	63.9	33.5	−3.7	1.9
	$C_{10}H_8$-1	71.1	42.8	64.5	35.4	−4.1	1.9
	$C_{10}H_8$-2	70.9	42.0	64.2	34.1	−3.8	1.9
CCSD(F12*)(T)/aug-cc-pVTZ	C_6H_6	71.8	42.1	64.8	34.3	−3.7	1.6
	$C_{10}H_8$-1	71.9	43.6	65.3	36.2	−4.1	1.6
	$C_{10}H_8$-2	71.8	42.7	65.1	34.8	−3.8	1.6

[a] $E_{act} = E_{el,act} + E_{act,ZPVE}$, $E_{rxn} = E_{el,rxn} + E_{rxn,ZPVE}$. [b] Extrapolated from the cc-pVDZ and cc-pVTZ basis sets.

and the raw cc-pVTZ numbers are close to the CCSD(F12*)-(T)/aug-cc-pVTZ values. For the larger PAHs, we thus continued with PNO-CCSD/cc-pVTZ and added to them the average result for the triples correction from the CCSD(F12*)-(T)/aug-cc-pVTZ calculations, 1.6 kJ/mol for reaction and −3.9 kJ/mol for the activation energies, and the vibrational correction from DFT/TPSSh/TZVP. The results are listed in Table 11. For the benchmark cases, benzene and naphthalene, the results agree within 1 kJ/mol with the CCSD(F12*)(T)/aug-cc-pVTZ reference values.

For the benzene, naphthalene, and phenanthrene accurate results are available in the literature (cmp. Table 12), including an experimental result for the activation energy for benzene.[45] They agree with the results from the current work within ±5 kJ/mol for the reaction and with ±3 kJ/mol for the activation energies, with the exception of the G2M(rcc,MP2) results from ref 46, which are somewhat higher.

Compared to the accurate coupled-cluster results, the DFT/TPSSh/TZVP results significantly underestimate the reaction energies and the activation energies. Closer inspection reveals that for the DFT/TPSSh/TZVP, results differ mainly by a constant offset, for the reaction energies of −20.5 kJ/mol and for the activation energies of −41.1 kJ/mol based on the average over the differences to those given in Table 11.

Extrapolation to Larger PAHs. The analysis of the DFT results showed that most of the site-specific differences between the reaction energies originate from the C−H bond breaking energy. Furthermore, if we distinguish between hindered bay and unhindered sites, the activation energies exhibit an approximately linear relation with the reaction energies (Bell−Evans−Polanyi principle). This allows us to go one step further than above and estimate the reaction and activation energies directly from the C−H bond breaking energies as

$$E_{rxn} = E_{C-H} + E_{relax} + E_{rxn,ZPVE} + E_{rxn,CCSD\text{-}corr} + E_{rxn,(T)}$$

$$E_{act} = E_{BEP} + f_{BEP}(E_{C-H} + E_{relax} + E_{rxn,ZPVE}) + E_{act,CCSD\text{-}corr} + E_{act,(T)}$$

with $E_{relax} = -7.1 \pm 0.4$ kJ/mol, $E_{rxn,ZPVE} = -6.9 \pm 0.8$ kJ/mol, $E_{rxn,CCSD\text{-}corr} = 20.5 \pm 0.3$ kJ/mol, $E_{rxn,(T)} = 1.6$ kJ/mol for the reaction and $E_{act,CCSD\text{-}corr} = 41.1 \pm 0.2$ kJ/mol and $E_{act,(T)} = -3.9$ kJ/mol for the activation energies. $E_{BEP} = 20.397$ kJ/mol and $f_{BEP} = 0.602$ for unhindered and $E_{BEP} = 9.766$ kJ/mol and $f_{BEP} = 2.877$ for bay region sites.

Table 11. PNO-CCSD/cc-pVTZ Results for Hydrogen Abstraction Energies in kJ/mol[a]

site	$E_{el,rxn,CCSD}$	E_{rxn}	$E_{el,act,CCSD}$	E_{act}
C_6H_6	39.9	33.7	75.8	64.9
$C_{10}H_8$-1	41.2	35.5	76.3	65.8
$C_{10}H_8$-2	40.5	34.2	75.9	65.3
$C_{14}H_{10}$a-1	42.8	37.7	77.1	66.8
$C_{14}H_{10}$a-2	41.1	35.2	76.1	65.4
$C_{14}H_{10}$a-3	40.4	33.9	75.9	64.9
$C_{14}H_{10}$p-1	41.2	35.0	76.2	65.4
$C_{14}H_{10}$p-2	41.1	35.3	76.3	65.7
$C_{14}H_{10}$p-3	41.6	35.1	76.7	65.8
$C_{14}H_{10}$p-4	40.5	34.3	75.8	64.8
$C_{14}H_{10}$p-5	34.8	29.3	76.7	67.2
$C_{16}H_{10}$-1	40.7	33.8	75.9	64.9
$C_{16}H_{10}$-2	42.4	36.5	76.9	66.1
$C_{16}H_{10}$-3	40.9	34.6	76.0	65.0
$C_{20}H_{12}$-1	34.0	28.3	77.0	66.6
$C_{20}H_{12}$-2	40.8	36.0	76.0	65.8
$C_{20}H_{12}$-3	42.9	38.6	77.2	67.2
$C_{22}H_{12}$a-1	42.5	37.1	76.9	65.8
$C_{22}H_{12}$a-2	40.6	34.8	75.8	64.9
$C_{22}H_{12}$a-3	42.5	36.6	76.8	65.8
$C_{22}H_{12}$a-4	43.5	40.2	77.4	68.7
$C_{22}H_{12}$a-5	40.8	36.0	75.7	65.7
$C_{22}H_{12}$b-1	35.4	29.3	77.2	68.1
$C_{22}H_{12}$b-2	40.6	35.2	75.8	65.3
$C_{22}H_{12}$b-3	43.0	38.2	77.3	67.0
$C_{22}H_{12}$b-4	41.6	36.8	76.4	66.4
$C_{22}H_{12}$b-5	41.5	36.7	76.4	66.4
$C_{22}H_{12}$b-6	42.1	37.4	76.7	66.7
$C_{24}H_{12}$	41.9	36.9	76.5	66.4

[a] $E_{rxn} = E_{el,rxn,CCSD} + E_{el,rxn,(T)} + E_{rxn,ZPVE}$ with $E_{el,rxn,(T)} = 1.6$ kJ/mol. E_{act} has been calculated correspondingly with $E_{el,act,(T)} = -3.9$ kJ/mol.

For applications in reduced kinetic models that require pretabulated data, calculations for a specific PAH can be replaced by using average values for the reaction site type such as solo sites on zigzag or bay sites on armchair edges. In Table 14, we collected the energy ranges for the seven different site types considered in the current study. With the exception of the K-region duo sites (also known as free edge sites), where the range of the reaction energies 4.3 kJ/mol, the ranges are for all other sites smaller than 1 kcal/mol. This means that the errors

https://dx.doi.org/10.1021/acs.jpca.0c07133
J. Phys. Chem. A 2020, 124, 9626−9637

Table 12. Hydrogen Abstraction Energies by H Atoms from the Literature in kJ/mol

site	method	E_{act}	E_{rxn}
C_6H_6	CCSD(T)-F12a/cc-pVTZ	66.1[47]	33.4[47]
	G2M(rcc,MP2)	83.3[46]	46.4[46]
	CCSD(T)-F12/aug-cc-pVTZ	66.5[48]	35.6[48]
	G3(MP2,CC)	71.1[48]	36.8[48]
	experimental	66.9[45]	36 ± 3[46]
$C_{10}H_8$-1	G3(MP2)-RAD	70[20]	38[20]
	CCSD(T)-F12a/cc-pVTZ	66.5[47]	35.6[47]
$C_{10}H_8$-2	G3(MP2)-RAD	70[20]	37[20]
$C_{14}H_{10}$p-1	G3(MP2)-RAD	70[20]	38[20]
$C_{14}H_{10}$p-2	G3(MP2)-RAD	70[20]	38[20]
$C_{14}H_{10}$p-3	G3(MP2)-RAD	70[20]	38[20]
$C_{14}H_{10}$p-4	G3(MP2)-RAD	63[20]	37[20]
$C_{14}H_{10}$p-5	G3(MP2)-RAD	70[20]	32[20]

[a]Experimental study in a shock tube with laser-schlieren technique over 1900−2400 K and 0.2−1.0 atm.

introduced by taking the average value for the site type are within the error bars of the quantum chemical calculations.

Reaction Rate Constants. The reaction rate constants are calculated using equation 3 based on the energies given in Tables 11 and 13. Table 15 lists the parameters fit to the modified Arrhenius equation for the average over different site types, as classified in Table 14. All reported rate constants are "per reaction site", that is, we excluded the symmetry numbers that account for the summation over symmetry equivalent reaction sites, for example, in benzene or naphthalene, to allow for direct comparability between the different PAHs. The individual rate constants can be found in Supporting Information. As already observed for the activation energies, the armchair bay region sites show the largest variations between the different cases in the test set; therefore, the uncertainty in the rate constant parameters given for the armchair bay region site is higher than the others.

Table 13. Reaction and Activation Energies for Larger Molecules Estimated as $E_{est} = E_{el,DFT} + E_{ZPVE} + E_{CCSD\text{-}corr} + E_{(T)}$ with $E_{rxn,CCSD\text{-}corr}$ = 20.5 kJ/mol, $E_{rxn,(T)}$ = 1.6 kJ/mol, $E_{act,CCSD\text{-}corr}$ = 41.1 kJ/mol, and $E_{act,(T)}$ = −3.9 kJ/mol

site	$E_{rxn,est}$	$E_{act,est}$
$C_{42}H_{18}$-1	30.0	68.7
$C_{42}H_{18}$-2	36.1	65.6
$C_{50}H_{18}$-1	35.3	65.3
$C_{50}H_{18}$-2	29.8	69.1
$C_{50}H_{18}$-3	28.3	63.4
$C_{50}H_{18}$-4	37.0	66.4
$C_{50}H_{18}$-5	36.7	66.4
$C_{50}H_{18}$-6	39.3	67.8
$C_{84}H_{24}$-1	30.2	71.3
$C_{84}H_{24}$-2	38.9	68.8
$C_{96}H_{24}$-1	36.2	66.5
$C_{96}H_{24}$-2	39.5	68.4

Figure 5a shows the average rate constants given in Table 15. In general, the rate constants do not vary significantly with the different sizes and sites of the molecules. The average activation energies (given in Table 14) are ordered as zigzag edge solo sites and armchair bay region sites ≫ K-region duo, inner trio > outer trio, and quartet sites. As can be seen in Figure 5a, the reaction rate constants follow a comparable trend. The armchair bay region and zigzag edge solo sites have the lowest reaction rate constants (56−61 and 41−58% lower compared to free edge outer quartet site respectively). Similarly, the K-region duo sites have around 24−33% lower rate constants compared with the free edge outer quartet sites. Figure 5b shows the calculated rate constants for benzene and naphthalene and compares them with the literature. The calculated reaction rate constant for benzene agree well with the experimental study[45] at the low temperature range. In general, earlier calculations with composite methods, for example, by Semenikhin *et al.*,[48] Kislov *et al.*,[49] and Mebel *et al.*,[46] predict lower rate constants than the ones calculated in this study. The rate constants of the armchair bay region sites are lower than that of the zigzag solo sites (by 9−25%), which is in accordance with the findings of the study by Semenikhin et al. based on G3(MP2,CC) calculations for chrysene and benzo-[*a*]pyrene.[48]

SUMMARY AND CONCLUSIONS

Hydrogen abstraction from large polyaromatic systems, as present in (bio)chars or soot, is a crucial first reaction step in heterogeneous combustion mechanism, and the determination of accurate rate parameters is pivotal for its modeling. Because of the systems size, often a single approximate method (e.g., ReaxFF) is used. We showed, however, that multilevel and multimethod approaches are needed to tackle this difficult problem accurately. With a combination of DFT, PNO-CCSD as a local correlation method and CCSD(F12*)(T) as highly accurate method, we extended in the current work the available data on reaction and activation energies and reaction rates from medium-sized systems to large PAHs up to $C_{96}H_{24}$. This has been possible since we could show for small and medium-sized PAHs that the energy corrections for connected triples excitations (T) and also the difference between PNO-CCSD

Table 14. Range of the Reaction Energies (in kJ/mol Including the ZPVE) for Different Types of Reaction Sites, Based on the Results Reported in Tables 11 and 13

site type	sites	E_{rxn}	E_{act}	$E_{act,average}$
zigzag edge solo	$C_{14}H_{10}$a-1, $C_{22}H_{12}$a-4, $C_{50}H_{18}$-6, $C_{96}H_{24}$-2	37.7−40.2	66.8−68.7	67.9 ± 0.8
K-region duo	$C_{14}H_{10}$p-1, $C_{16}H_{10}$-3, $C_{22}H_{12}$a-5, $C_{22}H_{12}$b-4, $C_{22}H_{12}$b-5, $C_{22}H_{12}$b-6, $C_{24}H_{12}$-1, $C_{50}H_{18}$-4, $C_{50}H_{18}$-5, $C_{84}H_{24}$-2, $C_{96}H_{24}$-1	34.7−38.9	65.0−68.8	66.4 ± 1.0
inner trio	C16H12-2, $C_{20}H_{12}$-3, $C_{22}H_{12}$a-1, $C_{22}H_{12}$a-3, $C_{22}H_{12}$b-3	36.6−38.6	65.8−67.2	66.4 ± 0.7
outer trio	$C_{16}H_{10}$-1, $C_{20}H_{12}$-2, $C_{22}H_{12}$a-2, $C_{22}H_{12}$b-2, $C_{42}H_{18}$-2, $C_{50}H_{18}$-1	33.8−36.1	64.9−65.8	65.3 ± 0.3
free edge inner quartet	$C_{10}H_8$-1, $C_{14}H_{10}$a-2, $C_{14}H_{10}$p-2	35.2−35.5	65.4−65.8	65.6 ± 0.2
free edge outer quartet	C_6H_6-1, $C_{10}H_8$-2, $C_{14}H_{10}$a-3, $C_{14}H_{10}$p-3, $C_{14}H_{10}$p-4	33.7−35.1	64.8−65.8	65.1 ± 0.4
armchair bay region	$C_{14}H_{10}$p-5, $C_{20}H_{12}$-1, $C_{22}H_{12}$b-1, $C_{42}H_{18}$-1, $C_{50}H_{18}$-2, $C_{50}H_{18}$-3, $C_{84}H_{24}$-1	28.3−30.2	63.4−71.3	67.8 ± 2.4

https://dx.doi.org/10.1021/acs.jpca.0c07133
J. Phys. Chem. A 2020, 124, 9626−9637

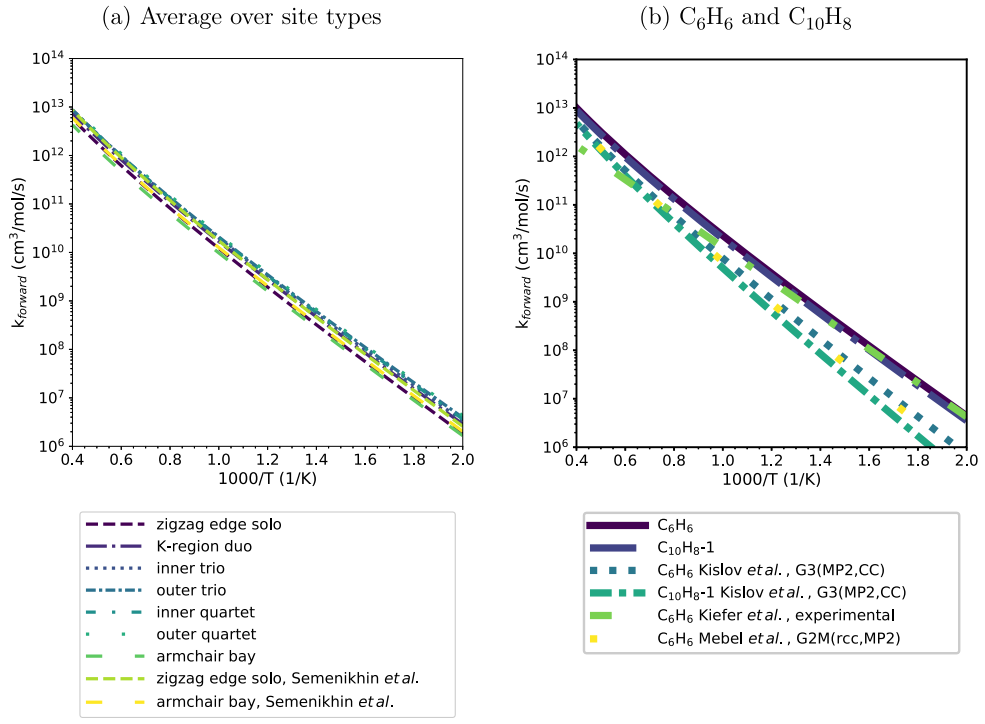


Figure 5. Semi-logarithmic plots of the site specific forward reaction rate constants as a function of inverse temperature for the temperature range 500−2500 K and comparison with the literature. Note that the experimental data for Kiefer *et al.*[45] is based on measurements at a temperature range 1800−2500 K; however, it is shown for the full temperature range. For comparison, the reference rate constants[45,46,49] are divided by 6 for benzene and by 4 for naphthalene so that the per site rate constants can directly be compared.

Table 15. Non-Linear Least Square Fit of the Calculated Rates Constants (k) for the Studied Sites to a Modified Arrhenius Equation, $k = AT^n \exp(-E/T)$, Where T and E Are in Units of K and k is in $cm^3\ mol^{-1}\ s^{-1}$

site type	A	N	E
zigzag edge solo	1.25×10^8	1.76	7585
K-region duo	1.30×10^8	1.78	7437
inner trio	1.82×10^8	1.76	7469
outer trio	1.27×10^8	1.80	7333
free edge inner quartet	1.17×10^8	1.81	7361
free edge outer quartet	1.20×10^8	1.82	7319
armchair bay region	9.62×10^7	1.75	7468

and DFT/TPSSh are transferable between different PAHs and reaction sites. Thus, we can answer the central question, raised in the title, whether small polyaromatics can be used to describe their larger counterparts for local reactions clearly with "yes". In other words, it is at the electronic structure level much more import to use sufficiently accurate methods for smaller models than to approximate the complete system. The effects of steric repulsion and geometric strain, which require at least somewhat larger models, can be described on lower levels with more approximate methods.

In this spirit, it is worth to note that most of the differences in the reaction energies between different sites enter actually via the C−H bond breaking energy at a fixed geometry for the PAH, the computationally cheapest contribution at the DFT level. The activation energies, on the other hand, can for this reaction estimated reliably from the reaction energy, as shown above at the DFT level.

Besides pushing the limits for accurate reference data, we assessed the extended semi-empirical tight binding methods (xTB) and a popular reactive force-field ReaxFF as tools for modeling combustion reactions. The cheapest xTB methods GFN0-xTB are not suitable for modeling combustion since it fails to qualitatively describe H-abstraction by a large order of magnitude. However, GFN2-xTB shows a good agreement with the DFT results in terms of reaction energies and equilibrium geometries, indicating that it could be used in future to bridge between DFT and ReaxFF. Since the latter is numerically the most efficient method, it is ideally suited for an exploration of unknown potential surfaces; however, it strongly overestimates reaction energies and barriers in the studied context, which makes these processes less often observed in an MD simulation, leading to longer simulation times.

Overall we demonstrate that a combination of judiciously chosen methods and exploiting transferability of energetic contributions in a multiscale and multimethod fashion is the ultimate choice for modeling such complex heterogeneous combustion processes, whereas single method strategies will most likely fail.

This work intends to give another contribution toward a better fundamental understanding of the gas/solid heterogeneous reactions. The further confirmation of the possibility of accurately estimating rate constants based on analogous reactions in the gas phase will allow the development of detailed mechanisms to increase the predictability of the actual macromolecular or lumped models.

For the specific reaction class studied in this investigation, the H-abstraction from PAHs with H atoms, we find that the rate

https://dx.doi.org/10.1021/acs.jpca.0c07133
J. Phys. Chem. A 2020, 124, 9626−9637

constants for the sterically unhindered sites can be very well approximated by site type-specific averages.[50] Only for the sterically hindered bay region sites, as found on armchair edges, we observe a significant variation of the rate constants between the cases within the test set. The steric repulsions at these sites induce some non-planarity and geometric strains,[42] partially released when the C−H bond is broken, that depend on the flexibility of the environment.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpca.0c07133.

Study on PNO selection thresholds, non-linear least square fits for the calculated rate constants for the studied sites using a modified Arrhenius equation, additional data for the relation between reaction and activation energy for TPSSh-D3/TZVP, and raw data from our TPSSh-D3/TZVP, xTB and ReaxFF calculations (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Özlem Yönder − *Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, D-44780 Bochum, Germany;* orcid.org/0000-0002-1464-2018; Email: oezlem.yoender@ruhr-uni-bochum.de

Gunnar Schmitz − *Computational Materials Chemistry Group, Ruhr-Universität Bochum, 44780 Bochum, Germany;* Email: Gunnar.Schmitz@rub.de

Authors

Christof Hättig − *Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, D-44780 Bochum, Germany;* orcid.org/0000-0002-5752-2710

Rochus Schmid − *Computational Materials Chemistry Group, Ruhr-Universität Bochum, 44780 Bochum, Germany;* orcid.org/0000-0002-1933-3644

Paulo Debiagi − *Simulation of Reactive Thermo-Fluid Systems (STFS), TU Darmstadt, 64827 Darmstadt, Germany*

Christian Hasse − *Simulation of Reactive Thermo-Fluid Systems (STFS), TU Darmstadt, 64827 Darmstadt, Germany*

Andrea Locaspi − *Department of Chemistry, Materials and Chemical Engineering "G. Natta", Politecnico di Milano, 20133 Milano, Italy*

Tiziano Faravelli − *Department of Chemistry, Materials and Chemical Engineering "G. Natta", Politecnico di Milano, 20133 Milano, Italy*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpca.0c07133

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank the German Research Foundation (DFG) for funding this research within the framework of the SFB/Transregio 129 "Oxyflame" (project number 215035359).

■ ABBREVIATIONS

DFT, density functional theory; CCSD(T), Coupled-cluster singles-and-doubles with perturbative triples; ReaxFF, reactive force field; xTB, extended semi-empirical tight-binding

■ REFERENCES

(1) Paris, O.; Zollfrank, C.; Zickler, G. A. Decomposition and carbonisation of wood biopolymers—a microstructural study of softwood pyrolysis. *Carbon* **2005**, *43*, 53−66.

(2) Amonette, J. E.; Joseph, S. *Biochar for Environmental Management: Science and Technology*; Earthscan, 2009; Chapter 3, pp 34−52.

(3) Keiluweit, M.; Nico, P. S.; Johnson, M. G.; Kleber, M. Dynamic Molecular Structure of Plant Biomass-Derived Black Carbon (Biochar). *Environ. Sci. Technol.* **2010**, *44*, 1247−1253.

(4) Cao, X.; Pignatello, J. J.; Li, Y.; Lattao, C.; Chappell, M. A.; Chen, N.; Miller, L. F.; Mao, J. Characterization of Wood Chars Produced at Different Temperatures Using Advanced Solid-State 13C NMR Spectroscopic Techniques. *Energy Fuels* **2012**, *26*, 5983−5991.

(5) Simmie, J. M. Detailed chemical kinetic models for the combustion of hydrocarbon fuels. *Prog. Energy Combust. Sci.* **2003**, *29*, 599−634.

(6) Kmiecik, S.; Gront, D.; Kolinski, M.; Wieteska, L.; Dawid, A. E.; Kolinski, A. Coarse-Grained Protein Models and Their Applications. *Chem. Rev.* **2016**, *116*, 7898−7936.

(7) van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A. ReaxFF: A Reactive Force Field for Hydrocarbons. *J. Phys. Chem. A* **2001**, *105*, 9396−9409.

(8) Salmon, E.; van Duin, A. C. T.; Lorant, F.; Marquaire, P.-M.; Goddard, W. A. Early maturation processes in coal. Part 2: Reactive dynamics simulations using the ReaxFF reactive force field on Morwell Brown coal structures. *Org. Geochem.* **2009**, *40*, 1195−1209.

(9) Castro-Marcano, F.; Kamat, A. M.; Russo, M. F.; van Duin, A. C. T.; Mathews, J. P. Combustion of an Illinois No. 6 coal char simulated using an atomistic char representation and the ReaxFF reactive force field. *Combust. Flame* **2012**, *159*, 1272−1285.

(10) Zheng, M.; Li, X.; Liu, J.; Wang, Z.; Gong, X.; Guo, L.; Song, W. Pyrolysis of Liulin Coal Simulated by GPU-Based ReaxFF MD with Cheminformatics Analysis. *Energy Fuels* **2013**, *28*, 522−534.

(11) Ding, Z.-B.; Di Marco, E.; Pelucchi, M.; Faravelli, T.; Maestri, M. First-principles assessment of the analogy between gas-phase and gas-solid H-abstraction reactions at graphene edges. *Chem. Eng. J.* **2019**, *377*, 119691. , ISCRE 25 Special Issue: Bridging Science and Technology

(12) Lacroix, R.; Fournet, R.; Ziegler-Devin, I.; Marquaire, P.-M. Kinetic modeling of surface reactions involved in CVI of pyrocarbon obtained by propane pyrolysis. *Carbon* **2010**, *48*, 132−144.

(13) Frenklach, M.; Wang, H. Detailed surface and gas-phase chemical kinetics of diamond deposition. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1991**, *43*, 1520−1545.

(14) Wang, H.; Frenklach, M. Calculations of Rate Coefficients for the Chemically Activated Reactions of Acetylene with Vinylic and Aromatic Radicals. *J. Phys. Chem.* **1994**, *98*, 11465−11489.

(15) Adkins, E. M.; Miller, J. H. Towards a taxonomy of topology for polynuclear aromatic hydrocarbons: linking electronic and molecular structure. *Phys. Chem. Chem. Phys.* **2017**, *19*, 28458−28469.

(16) Bell, R. P.; Hinshelwood, C. N. The theory of reactions involving proton transfers. *Proc. R. Soc. London, Ser. A* **1936**, *154*, 414−429.

(17) Evans, M. G.; Polanyi, M. Inertia and driving force of chemical reactions. *Trans. Faraday Soc.* **1938**, *34*, 11−24.

(18) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652−1671.

(19) Pracht, P.; Caldeweyher, E.; Ehlert, S.; Grimme, S. *A Robust Non-Self-Consistent Tight-Binding Quantum Chemistry Method for large Molecules*, 2019 DOI: 10.26434/chemrxiv.8326202.v1.

(20) Carissan, Y.; Klopper, W. Hydrogen abstraction from biphenyl, acenaphthylene, naphthalene and phenanthrene by atomic hydrogen and methyl radical: DFT and G3(MP2)-RAD data. *J. Mol. Struct.: THEOCHEM* **2010**, *940*, 115−118.

(21) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. Comparative assessment of a new nonempirical density functional:

Molecules and hydrogen-bonded complexes. *J. Chem. Phys.* **2003**, *119*, 12129−12137.
(22) Treutler, O.; Ahlrichs, R. Efficient molecular numerical integration schemes. *J. Chem. Phys.* **1995**, *102*, 346−354.
(23) Schäfer, A.; Huber, C.; Ahlrichs, R. Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr. *J. Chem. Phys.* **1994**, *100*, 5829−5835.
(24) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
(25) TURBOMOLE V7.4 2019, A Development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989−2007, TURBOMOLE GmbH, since 2007; available from https://www.turbomole.org.
(26) Csonka, G. I.; Ruzsinszky, A.; Perdew, J. P. Estimation, Computation, and Experimental Correction of Molecular Zero-Point Vibrational Energies. *J. Phys. Chem. A* **2005**, *109*, 6779−6789.
(27) Hättig, C.; Tew, D. P.; Köhn, A. Communications: Accurate and efficient approximations to explicitly correlated coupled-cluster singles and doubles, CCSD-F12. *J. Chem. Phys.* **2010**, *132*, 231102.
(28) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron-affinities of the 1st-row atoms revisited - systematic basis-sets and wave-functions. *J. Chem. Phys.* **1992**, *96*, 6796−6806.
(29) Yousaf, K. E.; Peterson, K. A. *Chem. Phys. Lett.* **2009**, *476*, 303.
(30) Weigend, F.; Köhn, A.; Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **2001**, *116*, 3175−3183.
(31) Häser, M.; Ahlrichs, R. Improvements on the direct SCF method. *J. Comput. Chem.* **1989**, *10*, 104−111.
(32) Schmitz, G.; Helmich, B.; Hättig, C. A scaling PNO-MP2 method using a hybrid OSV-PNO approach with an iterative direct generation of OSVs. *Mol. Phys.* **2013**, *111*, 2463−2476.
(33) Schmitz, G.; Hättig, C.; Tew, D. P. Explicitly correlated PNO-MP2 and PNO-CCSD and their application to the S66 set and large molecular systems. *Phys. Chem. Chem. Phys.* **2014**, *16*, 22167−22178.
(34) Development version of TURBOMOLE V7.5 2020, A Development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989−2007, TURBOMOLE GmbH, since 2007; available from https://www.turbomole.org.
(35) Aktulga, H. M.; Fogarty, J. C.; Pandit, S. A.; Grama, A. Y. Parallel Reactive Molecular Dynamics: Numerical Methods and Algorithmic Techniques. *Parallel Comput.* **2012**, *38*, 245−259.
(36) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.
(37) Chenoweth, K.; van Duin, A. C. T.; Goddard, W. A. ReaxFF Reactive Force Field for Molecular Dynamics Simulations of Hydrocarbon Oxidation. *J. Phys. Chem. A* **2008**, *112*, 1040−1053.
(38) Extended semi-empirical tight binding model xTB v6.3.1, available from https://github.com/grimme-lab/xtb 2020.
(39) Plessow, P. Reaction Path Optimization without NEB Springs or Interpolation Algorithms. *J. Chem. Theory Comput.* **2013**, *9*, 1305−1310.
(40) Grimme, S. Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory. *Chem.—Eur. J.* **2012**, *18*, 9955−9964.
(41) Shavitt, I. A Calculation of the Rates of the Ortho-Para Conversions and Isotope Exchanges in Hydrogen. *J. Chem. Phys.* **1959**, *31*, 1359−1367.
(42) Hemelsoet, K.; Van Speybroeck, V.; Moran, D.; Marin, G. B.; Radom, L.; Waroquier, M. Thermochemistry and Kinetics of Hydrogen Abstraction by Methyl Radical from Polycyclic Aromatic Hydrocarbons. *J. Phys. Chem. A* **2006**, *110*, 13624−13631.
(43) Tew, D. P.; Klopper, W.; Neiss, C.; Hättig, C. Quintuple-ζ quality coupled-cluster correlation energies with triple-ζ basis sets. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1921−1930.
(44) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-set convergence in correlated calculations on Ne, N2, and H2O. *Chem. Phys. Lett.* **1998**, *286*, 243−252.
(45) Kiefer, J. H.; Mizerka, L. J.; Patel, M. R.; Wei, H. C. A shock tube investigation of major pathways in the high-temperature pyrolysis of benzene. *J. Phys. Chem.* **1985**, *89*, 2013−2019.
(46) Mebel, A. M.; Lin, M. C.; Yu, T.; Morokuma, K. Theoretical Study of Potential Energy Surface and Thermal Rate Constants for the C6H5 + H2 and C_6H_6 + H Reactions. *J. Phys. Chem. A* **1997**, *101*, 3189−3196.
(47) Hou, D.; You, X. Reaction kinetics of hydrogen abstraction from polycyclic aromatic hydrocarbons by H atoms. *Phys. Chem. Chem. Phys.* **2017**, *19*, 30772−30780.
(48) Semenikhin, A. S.; Savchenkova, A. S.; Chechet, I. V.; Matveev, S. G.; Liu, Z.; Frenklach, M.; Mebel, A. M. Rate constants for H abstraction from benzo(a)pyrene and chrysene: a theoretical study. *Phys. Chem. Chem. Phys.* **2017**, *19*, 25401−25413.
(49) Kislov, V. V.; Islamova, N. I.; Kolker, A. M.; Lin, S. H.; Mebel, A. M. Hydrogen Abstraction Acetylene Addition and Diels-Alder Mechanisms of PAH Formation: A Detailed Study Using First Principles Calculations. *J. Chem. Theory Comput.* **2005**, *1*, 908−924.
(50) Violi, A.; Truong, T. N.; Sarofim, A. F. Kinetics of Hydrogen Abstraction Reactions from Polycyclic Aromatic Hydrocarbons by H Atoms. *J. Phys. Chem. A* **2004**, *108*, 4846−4852.

https://dx.doi.org/10.1021/acs.jpca.0c07133
J. Phys. Chem. A 2020, 124, 9626−9637

IV. Computational Chemistry as a Tool for Predictive Combustion Modelling

This work presents modifications to the original ChemTraYzer and evaluates the performance of this version to explore phenanthrene combustion reaction.
The coding and testing was done mainly by me through discussions with my supervisor Prof. Christof Hättig as well as with extensive discussions and explanations from Prof. Kai Leonhard, Felix Schmalz, and Lukas Krep. Except from the inputting of initial structures from a coordinate file, the additions to the code were made by me. I have also done the simulation and analysis of the trajectories mentioned here. The discussions around the results were finalized by discussions with Prof. Hättig.

Supplementary material can be found under DOI: 10.5281/zenodo.4675515.

Computational Chemistry as a Tool for Predictive Combustion Modelling

Özlem Yönder*, Christof Hättig
Chair of Theoretical Chemistry
Quantum Chemistry Group
Ruhr-University Bochum, 44780 Bochum, Germany

Vanessa Angenent, Rochus Schmid
Chair of Inorganic Chemistry 2
Computational Materials Chemistry Group
Ruhr-University Bochum, 44780 Bochum, Germany

Kai Leonhard, Felix Schmalz
Chair of Technical Thermodynamics
RWTH Aachen University, 52062 Aachen, Germany

Highlights

- The ChemTraYzer 2.0 program suite, which had previously been tested for the automated generation of reaction mechanisms for methane combustion, has been extended and tested for the combustion of phenanthrene.
- The program suite was able to locate most of the transition states with activation barriers larger than 30 kJ/mol.
- The extension of ChemTraYzer to PAHs opens a new perspective for multiscale simulation of char burn out starting from atomistic calculations on fragment models for which the results can then be lumped into a kinetic model for computational fluid dynamics simulations of char in boilers.

Abstract

The kinetic models used in computational fluid dynamics simulations of solid fuel boilers generally use parameters that are fitted to experimental data, which do not represent individual well-defined processes going on at the atomistic scale. This hampers extrapolations to new fuels or reaction conditions. Here, we test the performance of a modified version of the ChemTraYzer 2.0 program suite, which can be a useful tool for the automated creation of reaction networks of char combustion. The program suite first runs molecular dynamics simulations with the reactive force field ReaxFF, and then refines the detected reaction pathways with quantum mechanical calculations. ChemTraYzer 2.0 had previously been tested for the combustion of methane; however, it has not been tested for larger carbon materials relevant to char burnout reactions. Here, after explaining the modifications that are made to the existing ChemTraYzer program for polycyclic aromatic hydrocarbons, we investigate the combustion of phenanthrene to test the performance of the modified workflow. With the modifications the program suite is able to locate the transition states for most reactions with non-negligible activation barriers. This opens up a route towards multiscale simulations of char combustion that use in computational fluid dynamics simulations kinetic models that are obtained from electronic structure calculations by lumping the results for an automatically generated reaction network.

* Corresponding author, e-mail: oezlem.yoender@rub.de

Keywords

computational chemistry, atomistic modelling, automatized reaction kinetics, oxyfuel combustion capture, combustion chemistry

1. Introduction

Carbon capture and storage technologies are seen as intermediate solutions to the increasing CO_2 emissions until renewables become our main source of energy. It has also been stated as a technological option for reducing the CO_2 emissions to the atmosphere in the Intergovernmental Panel on Climate Change Report on CCS technologies[1]. The two carbon capture technologies which are feasible for retrofitting existing power plants are the post-combustion and the oxy-fuel combustion capture technologies[2].

In the *post-combustion capture*, the combustion conditions are the same as in the conventional power plants and due to the high concentration of nitrogen (N_2) in the flue gas from the boiler, the CO_2 is separated from the downstream of the boiler by a chemical CO_2 separation unit. The separation of CO_2 can be performed via absorption using solvents such as aqueous solutions of monoethanolamine (MEA)[3], [4], or alternatively, via adsorption using metal—organic frameworks[4]–[6], or via membranes[4], [7]. The *oxy-fuel combustion capture* technology, on the other hand, operates by eliminating the N_2 from air before the combustion process via an air separation unit. As a result, the flue gas is mainly composed of CO_2 and water vapour (H_2O) which is then separated by condensation. In other words, the aim of oxy-fuel combustion is to concentrate the CO_2 obtained and make it "easy to isolate after combustion"[8]. In oxy-fuel combustion, the flue gas from the boiler, after being used for power generation, is recycled to the boiler in order to control the combustion temperature[8]. As a result, the combustion chemistry is altered[9], [10]. Therefore, it is especially important for oxy-fuel combustion to understand the detailed chemistry and kinetics of the combustion of solid fuels. A molecular understanding of solid fuel combustion requires development of a kinetic model for volatile release and char burnout. Here, we target char burnout as a starting point.

One of the widely used models for the simulation of char conversion is the Carbon Burnout Kinetic Model (CBK)[11] family which includes the Carbon Burnout Kinetics-Extended Model (CBK/E), Carbon Burnout Kinetics-Gasification Model (CBK/G), and the recent Char Conversion Kinetics Model (CCK)[12] which combines the 3-step char oxidation reactions of the CBK/E model with the 5-step gasification reaction model of CBK/G. The char conversion mechanism in CCK includes the following eight steps where the indices α, β, and γ denote different types of char surface oxides:

Oxidation

$$2C + O_2 \rightarrow C(O)_\alpha + CO$$

$$C + C(O)_\alpha + O_2 \rightarrow CO_2 + C(O)_\alpha$$

$$C(O)_\alpha \rightarrow CO$$

Gasificati on

$$CO_2 + C \leftrightarrow C(O)_\beta + CO$$

$$C(O)_\beta \rightarrow CO$$

$$C + H_2O \leftrightarrow C(O)_\gamma + H_2$$

$$C(O)_\gamma \rightarrow CO$$

$$C + 2H_2 \rightarrow CH_4$$

The CCK model is able to match the experimental data for different O_2 conditions for a given coal with a single set of kinetic parameters[12]. However, it is still not able to predict the

kinetics for any solid fuel, as the elementary reactions are represented by an eight step mechanism. Building a predictive kinetic model of char burnout requires understanding of the reaction mechanisms at a molecular level and their relations with the functional groups and the structural features of char.

For small hydrocarbons including linear and cyclic alkanes, alkenes, alkynes, as well as small aromatics detailed kinetic models exist[13]. However, building a kinetic model for char burnout is cumbersome due to its relatively complex structure which is composed of amorphous carbon, turbostratic carbon crystallites, and pore space[10], [14], [15]. A predictive kinetic model of char burnout requires a large database. This database should include reactions of different functional groups which exists within char as well as the reactions at different sites of aromatic units. The functional groups within char include epoxy, C-OH hydroxyl, C-O-C ether, C=O carbonyl, (H)O-C=O carboxyl, and lactone groups and for a given solid fuel their abundance depends on the charring atmosphere[10]. The different sites of polyaromatic units can be classified as bay and fjord regions, or solo, duo, trio, and quartet regions[16]. A database that represents these features will be helpful to detect the patterns of reactions and search for similar products and educts to understand the combustion chemistry. Analysing the combustion reactions manually wouldn't be feasible; therefore, an automated exploration of reaction networks as well as their kinetic values would be of help.

In 2015 Döntgen et al.[17] automated the *simulation* of reactions using reactive molecular dynamics and the simultaneous *processing* of the trajectories. After the batched processing and simulation of the trajectories the quantum mechanical (QM) jobs are submitted and *analysed*, for example, to calculate the rate constants[18]. This scheme, called Chemical Trajectory Analyzer (ChemTraYzer), has been implemented in a Python code and shown to successfully perform for the methane combustion[18]. Here, we aim to test the performance of ChemTraYzer for larger and more complex molecules. Therefore, we modified the *simulation* class of the ChemTraYzer software. Additionally, we have interfaced the program with the TURBOMOLE quantum chemistry package[19], [20]. With this study, we analyse the performance of the *simulation* and *processing* part of the ChemTraYzer software for the combustion of polycyclic aromatic hydrocarbons (PAHs) as char models with the example of phenanthrene. After specifying in the next section the computation details, we present and discuss the results for phenanthrene combustion reactions as observed in two reactive molecular dynamics simulations. Of particular interest is the performance of the ChemTraYzer workflow to locate the minima and transition states. Finally, we will discuss some necessary modifications to the ChemTraYzer workflow for the combustion of PAHs.

2. Methods

The molecular dynamics (MD) simulations have been performed with the ReaxFF reactive force field for molecular dynamics simulations of hydrocarbon oxidation [21] by running LAMMPS[22] via the interface in the ChemTraYzer suite[17], [18]. The simulations have been performed at 2000 K (using a Nosé-Hoover thermostat[23], [24] with a damping constant of 100 fs) and 600 mol/m^3 density (at approximately 100 bar) with a composition of 53% O_2, 20% CO_2, 7% phenanthrene, 7% H_2O, 7% OH, and 7% O with 7 phenanthrene molecules and in total 100 molecules in the simulation box. The simulation length for the simulations were 0.5 ns using a time-step of 0.1 fs.

The quantum mechanical calculations have been performed using the TURBOMOLE quantum chemistry package[19]. For the Density Functional Theory (DFT) calculations, first the meta-GGA functional TPSS[25], including the D3 dispersion correction[26], and the SVP basis set[27] have been used to pre-optimize the stationary points. Later, the geometry optimizations and the calculation of the second derivatives of the potential energy surface to obtain zero point vibrational frequencies have been repeated with the TPSSh[28] hybrid

functional and TZVP basis set[29] to improve the accuracy. The scaling factor of 0.9874 optimized for TPSSh/6-31+G(d,p) is used for the zero point vibrational energies[30]. For all of the DFT calculations, the resolution of identity (RI-J) approximation[31] is used. For the cases in which the direct TS optimization failed, the reaction paths are investigated with a double-ended chain-of-states calculation between the educt and product structures as implemented in the woelfling[32] program of the TURBOMOLE package.

The details of the ChemTraYzer software can be found elsewhere [17], [18]. We use the workflow of ChemTraYzer as described in Ref. [17] which we interfaced to TURBOMOLE and modified for the current study to avoid a few problems that appeared for reactions with polycyclic aromatic hydrocarbons (PAHs), vide infra.

3. Improved workflow

For reactions on larger systems it is important to avoid as much as possible redundant calculations on the same structure. ChemTraYzer already did this partially by storing already investigated structures in a database and checking for any initial molecular structure based on the canonical Simplified Molecular Input Line Entry System (SMILES)[33], [34] representations if it is already available in the database. To reduce the number of QM calculations for larger systems even further, we included an additional similarity check[35], based on the root mean square deviation between the internal coordinates of two structures. This check is applied to any batch of (equilibrium) structures before they are passed to the workflow for the optimization at the QM level.

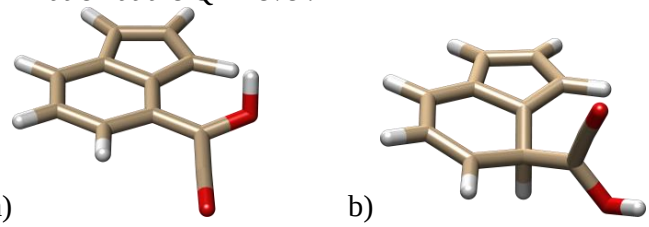


Figure 1 a) The original molecule $C_{11}H_7O_2$ with odd number of electrons. The canonical SMILES created from this molecule is O=C=C([C@H]1C=CC=C2C1=CC=C2)O. b) When the molecule is re-created from the canonical SMILES of the original molecule, there is an incorrectly added extra-hydrogen atom ($C_{11}H_8O_2$). This error also causes Open Babel to assign the spin multiplicity as 1, which is not possible due to the odd number of electrons

In the original implementation of ChemTraYzer the SMILES representations were used as input for the ChemTraYzer program and used to create the simulation box (using Packmol [36]). When an initial molecular structure does not already exist in the database, ChemTraYzer creates a 3D structure from the SMILES representation using Open Babel[37][38]. However, for large molecules the conversion from the SMILES representation to 3D structures does not always create the same molecule from which the SMILES representation was originally created, as for the molecule shown in Figure 1. Therefore, we have now modified the code such that 3D structures can be provided in the MDL Molfile format with title lines that include the canonical SMILES. ChemTraYzer then uses the 3D structure information from this file instead of creating it from the SMILES representation using Open Babel.

Another modification concerns the treatment of spin, which we adapted such that the spin multiplicities of equilibrium structures are no longer determined from Open Babel. The reason for this is that Open Babel assumes for molecules with more than one radical center always the high spin case[38] and thus sometimes results in an excited state multiplicity. It can also not cope with radical centers at atoms involved in a ring closure[38] as in Figure 1. In such cases, it will always result in the wrong spin multiplicity as it cannot comprehend the

correct system. Therefore, instead of using Open Babel, we determine the spin multiplicities of ReaxFF equilibrium structures by testing spin multiplicities up to quartet (that is, it tests doublet and quartet for molecules with an odd number of electrons, singlet and triplet for the molecules with an even number of electrons). Then, the occupation which results in the lowest energy is assigned to the molecule for further calculations (i.e. geometry optimizations and calculation of Hessians).

The assignment of spin multiplicities to transition state structures are not independent of the educts and products. ChemTraYzer was considering for the determination of the spin multiplicity of the transition states only the educts and determining the spin multiplicity using Open Babel for the educt complex. For example, in a hypothetical reaction where there are two doublet educts, which can couple to a singlet or triplet state. Open Babel would determine for the educt complex in this case triplet multiplicity (i.e. high spin case). However, if the product is a molecule whose ground state is singlet, assigning triplet to the transition state will lead to wrong transition states in the QM calculations. Instead, if one would like to search for the transition state of the ground state singlet product, one should perform spin-unrestricted DFT calculation with equal numbers of α and β spin electrons. That is, the original workflow would wrongly connect a triplet transition state to a singlet ground state. On the other hand, as will be shown in the following section, assuming that all equilibrium structures will be in their electronic ground states sometimes leads to reactions connected with spin-forbidden transitions.

To interface ChemTraYzer with TURBOMOLE, for the automated submission of QM jobs, a TURBOMOLE input is written using the TMOLE script[39] format. The user needs to specify the method as an input to the ChemTraYzer program (e.g., "tpss/SVP" for the calculations performed here). In the DFT calculations, the resolution of identity (RI-J) approximation[31] is used. The generated TMOLE scripts perform only single point calculations. Then, a separate Python module (named CTYturbomole) which follows a similar workflow as the original ChemTraYzer[18] is called to conduct the geometry optimizations, the calculation of the second derivatives, and for successfully optimized transition state structures calculation of the intrinsic reaction coordinate (IRC) pathways (either within ChemTraYzer or in the QM job submission scripts). This module can either be called within the simulation module or separately run in a queuing system.

The workflow for searching and optimizing the transition states starts from the structures where in the ReaxFF run a change in a bond order was detected. The first step at the QM level is a pre-optimization of the structures during which the internal coordinates of a set of active atoms are kept fixed. After this QM pre-optimization, the constraints on the internal coordinates are removed. Then, the Hessian matrix is computed and the number of negative eigenvalues is checked. After the calculation of the Hessian matrix, if there are no negative eigenvalues, the program does not perform any further calculations. For the cases where there are one or more negative eigenvalues, the trust-region image minimization (TRIM) algorithm[40] is used to optimize the transition states. For this, a reduced trust radius of 3×10^{-2} a.u. is used instead of the default value of TURBOMOLE which is 3×10^{-1} a.u., because fixing the internal coordinates of the active atoms to those from the ReaxFF trajectory during the QM pre-optimization might result in sizeable negative eigenvalues of the Hessian matrix and large gradients. After the transition state geometry optimization, the Hessian matrix is calculated again. For those cases where it has only one negative eigenvalue, the IRC pathways are calculated and the structures of the end points are optimized to find the corresponding educt and product for comparison with the educts and products determined originally by ReaxFF to confirm the transition state.

4. Results and Discussion

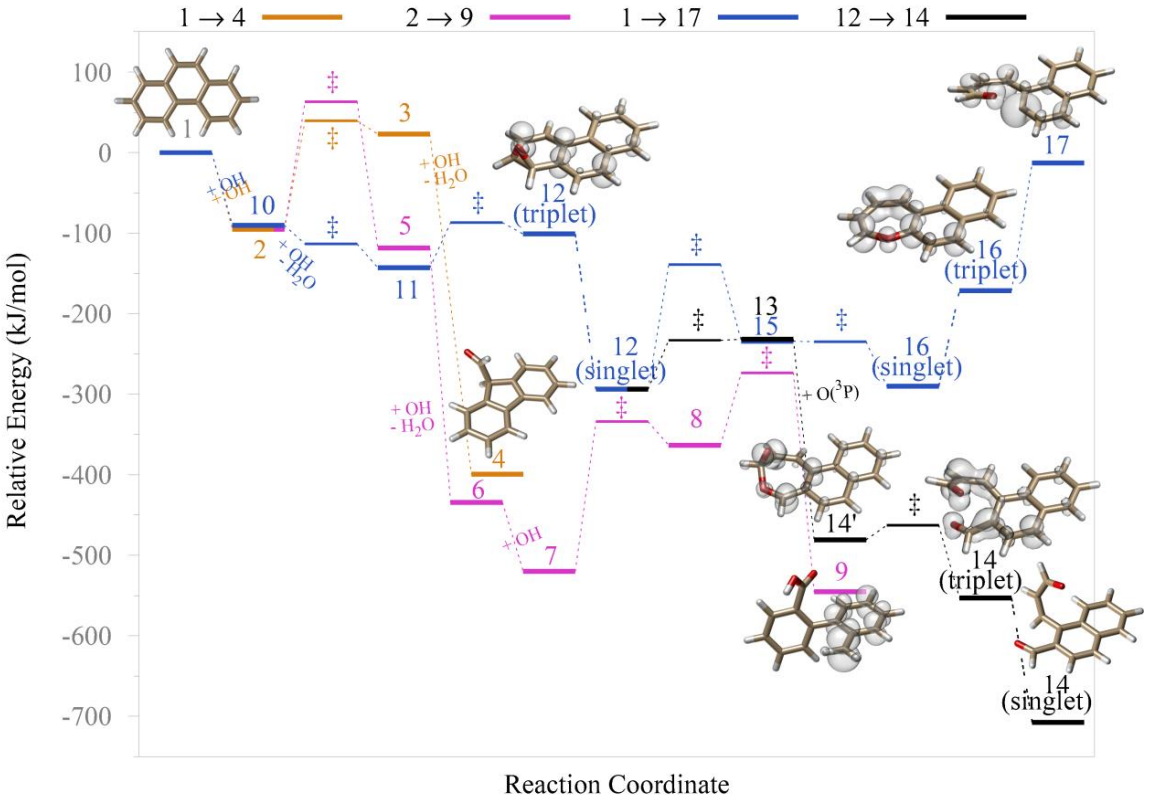


Figure 2 Reactions observed in two of the ReaxFF simulations. The structures of phenanthrene (1), the end products (4, 9, singlet 14, and 17) as well as some of the intermediate structures (triplet 12, triplet 16, 14', and triplet 14) are shown together with their spin densities. The symbol ‡ indicates transition states. The species connected with dashing style (… …) result from relaxation after spin-forbidden transitions (11→12, 13→14, and 16→17).

Figure 2 shows the reaction pathways for different reactions observed in the first steps of two trajectories. Once small molecules are formed and released, their detailed combustion mechanisms are already available in the literature, as e.g. for formic acid[41] or formaldehyde[42]. Therefore, the combustion kinetics of small molecules are here less interesting. In this section a detailed analysis of the reactions from the two simulations (shown in Figure 2) is presented to discuss the performance of the ChemTraYzer workflow for exploring the first steps of phenanthrene combustion.

4.1. Trajectory 1

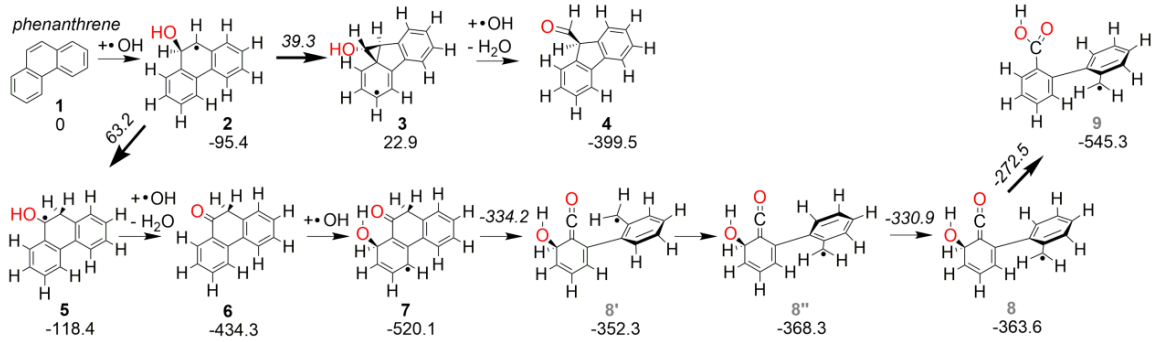


Figure 3: Some reactions initiated at the duo site of phenanthrene. The relative energies are given in kJ/mol and calculated at the DFT/TPSSh-D3/TZVP level. If given above an arrow, energies refer to the transition state of the reaction; if no transition state energy is given, the reaction is (approximately) barrierless. The bold arrows indicate the transition states for which the ChemTraYzer workflow was successful, regular arrows the reactions for which it failed or there were no significant barriers. Investigation of these reactions by a double-ended reaction path optimization method showed in most of these cases no sizeable maximum along the optimized reaction pathway at the DFT/TPSS-D3/SVP level of theory. For the molecules whose ID are shown in grey, the structures were not found with the ChemTraYzer workflow for reasons explained in the text.

4.1.1. Minima

Figure 3 shows reactions initiated at the duo site of phenanthrene. The main problem observed during the optimization of the minima was due to the fact that the initial structures were taken from the time-frames of the ReaxFF simulation where the transition state guess structures are taken from, i.e. the time-frames where a bond order of a broken or newly formed bond is close to 0.5. In ChemTraYzer, the molecules are then differentiated based on the bond connectivity between the atoms in the simulation box that belong to the molecule or the molecules involved in the bond order change at the presumed transition state. For unimolecular rearrangement reactions this approach generates only a single start structure which leads thus only to either the educt or the product of the reaction. This happens, for example, for the reactions 2→3 and 7→8', which produce one product molecule from one educt molecule. To be more precise, for the reaction 2→3, only the structure 3 was found and for the reaction 7→8' only the structure 7 was found. The structure 2 was detected only because it is the product of the non-unimolecular reaction of 1→2. In order to obtain the structure 8', the structure as taken from the ReaxFF trajectory has been optimized at the QM level.

4.1.2. Saddle points

The workflow described above failed to find a transition state for the OH radical addition to phenanthrene (i.e., reaction 1→2), because there were no imaginary frequencies for the QM pre-optimized transition state (TS) start structure. The approximate reaction pathway for this reaction was then computed with the double ended reaction path optimization method implemented in the woelfling[32] module of TURBOMOLE. The approximate reaction pathway was calculated between the van der Waals minimum of the educts and the product using the TPSS-D3/SVP level of theory. No sizeable maximum in the energy profile along the optimized pathway has been observed which implies a practically barrierless reaction.

The OH initiated oxidation of phenanthrene follows either formation of a cyclopropane-like moiety (reaction 2→3) or 9,10-shift of the hydrogen from the CHOH group to the neighbouring carbon (reaction 2→5, see Figure 3). The optimizations of the transition states were for both reactions successful with the ChemTraYzer workflow and their barriers are 134.7 and 158.7 kJ/mol, respectively.

Structure 3 further reacts with another OH radical and forms a water molecule and the product 4 which has an aldehyde group. As described above, we performed in these cases, an unrestricted DFT calculation with equal number of α and β spins. The QM pre-optimized transition state of the reaction 3→4 has three imaginary frequencies; however, none of the corresponding modes are in the direction of the reaction. Also in this case, the computed reaction pathway did not show any sizeable barrier (practically barrierless reaction), which explains why the transition state search failed. Structure 5 goes through a similar OH addition reaction leading to the formation of water and the product 6 with a keto group, where both products are in their singlet ground states. The QM pre-optimized structure for the transition state of reaction 5→6 had two imaginary frequencies. Again none associated with a mode in the direction of the studied reaction and also the computed reaction pathway did not show any sizeable barrier.

For the next addition of an OH radical (reaction 6→7) the ChemTraYzer workflow did not locate a transition state, even though after the QM pre-optimization there was a meaningful imaginary frequency. Similarly, as before, a check with the chain-of-states calculation for the reaction pathway showed no sizeable maximum, implying a practically barrierless reaction.

For the subsequent ring-opening reaction (reaction 7→8'), the imaginary frequency after the QM pre-optimization was associated with a mode oriented along the presumed reaction coordinate. However, the subsequent transition-state optimization with the TRIM algorithm without constraints on the coordinates failed and ended up at a structure which resembled the product. Figure 4 compares the transition state located after a reaction pathway optimization and the ChemTraYzer QM pre-optimized structure for the reaction 7→8', in both cases at the TPSS-D3/SVP level. The latter is more planar than the final optimized transition state and the C—C bond length is by 0.34 Å shorter. This implies that even though the single mode with an imaginary frequency shows a meaningful direction, the start structure is still too far away from the transition state structure to converge to the transition state. The energy of the transition state is in this case, however, also only ca. 29 kJ/mol above that of the product and also structure-wise the transition state is close to the product.

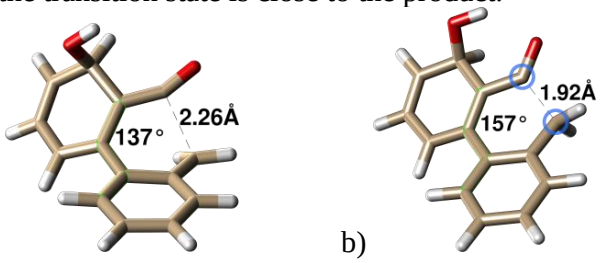


Figure 4: The TS structures for the reaction 7→ 8' at the TPSS-D3/SVP level of theory. a) The optimized structure of the transition state obtained from optimization of a start structure from a woelfling calculation. b) The QM pre-optimized structure of the transition state from ChemTraYzer. The active atoms are marked with circles.

The conformational changes from structure 8' to 8" and 8 are in the workflow not treated as chemical reactions, since they do not alter the bond connectivity. The intermediate minima and barriers which connect structure 8' with 8 have been found with reaction pathway calculations. For the rotation 8'→8" using TPSS-D3/SVP there is a barrier of 1.8 kJ/mol

without zero-point energy correction. On the TPSSh-D3/TZVP potential surface, this transition state does not have any negative eigenvalue. Since this rotation (8'→8") has a very small barrier, we did not search for the transition state of the rotation 8'→8" on the TPSSh-D3/TZVP surface. For the second rotation of the oxidized ring (8"→8), there is a barrier of 37.4 kJ/mol. For the formation of the carboxylic acid group (reaction 8→9), which has a barrier of 91.1 kJ/mol, the transition state was successfully located with the ChemTraYzer workflow.

4.2. Trajectory 2

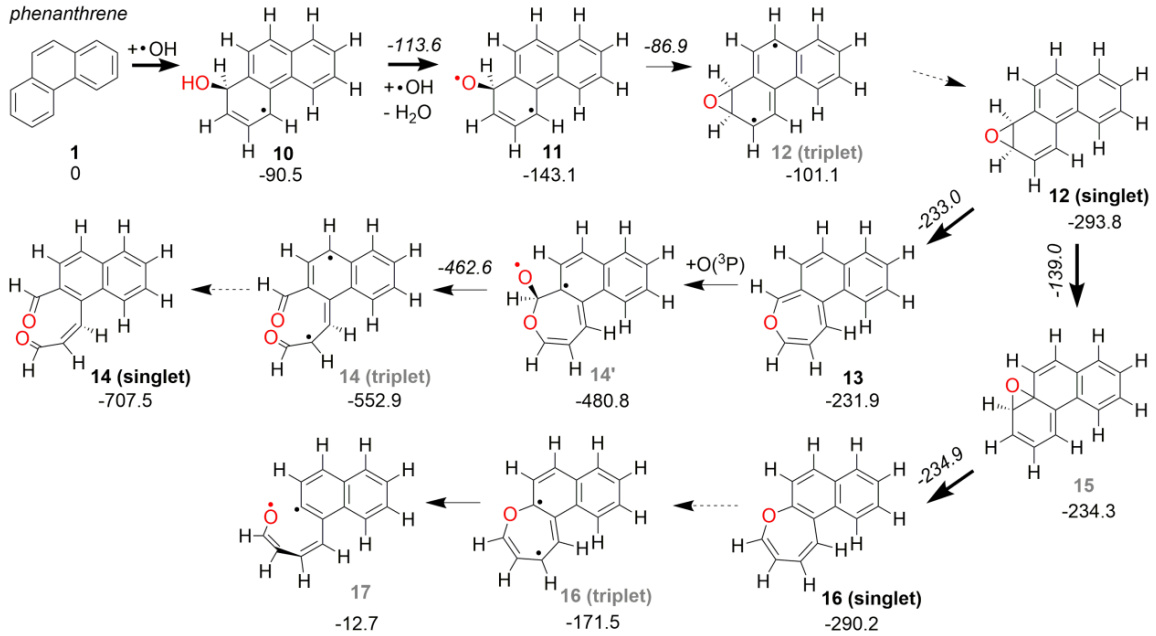


Figure 5: Some reactions initiated at the quartet site of phenanthrene. The relative energies are given in kJ/mol and calculated using the TPSSh-D3/TZVP method. The bold arrows indicate the transition states for which the ChemTraYzer workflow was successful. The dashed arrows indicate transitions between a singlet and a triplet state.

4.2.1. Minima

Figure 5 shows reactions of phenanthrene at the quartet region. The species 15 and 17 were not found by the automated workflow due to their connection to unimolecular reactions, i.e. the same problem as already explained for the first trajectory. During the ReaxFF pre-optimization of the educt and product for the reaction 12→15, both of the start structures converged to a structure with the C—C bond length of 1.87Å, which is in between the QM C—C bond length of 1.56Å for the species 15 and 2.33Å for the species 16. As the barrier for the opening of the three-membered ring is very low, the ReaxFF pre-optimized structure for the minima of the reaction 12→15 converged during the QM geometry optimization to the species 16. Also for the reaction 15→16, the educt and product structures of converged to the species 16 during the QM geometry optimization starting from the ReaxFF pre-optimized structure with the C—C bond length of 2.41Å. Therefore, the species 15 has later been detected from intrinsic reaction coordinate (IRC) calculations starting from the transition states of reaction 15→16 and 12→15 as the end point of the IRC calculation. We optimized the species 17 by taking the structure from the trajectory, because the ReaxFF pre-optimization has converged for the reaction 16→17 as well to the species 16.

Apart from the species 15 and 17, also 14' was not found with the ChemTraYzer workflow, because both the product of the oxygen atom addition (13→14') and both of the educt and product of the ring opening reaction (14'→14) converged to the end product (14) at the ReaxFF pre-optimization stage. The structure 14' has been later found in the IRC calculations starting from the transition state of the reaction 14'→14 on the triplet potential energy surface.

4.2.2. Saddle points

The ChemTraYzer workflow was successful to find the transition states for the OH addition reaction 1→10, for the subsequent reaction with another OH radical which leads to the abstraction of a H atom from the alcohol group (reaction 10→11), for the arene oxide–oxepin reactions (12→ 13 and 15→16), and for the diffusion of the oxygen around the carbon edge (12→15).

The energy of the transition state structure of the reaction 10→11 lies between the energies of the educts and products, because in this case the transition state refers to a small maximum along the reaction coordinate between the van der Waals minima for the educt and product complexes while the product and educt energies in Figure 5 refer to the separated non-interacting molecules. Additionally, the energies of the transition states for the reactions 12→ 13 and 15→16 are slightly lower than that for the product in case of 12→13 and lower than that for the educt in case of 15→16 (see Figure 5). This is due to the zero-point vibrational energy (ZPE) corrections that have been added to the electronic energies. Without the ZPE corrections, the transition state structures are higher in energy than those of the educts and products.

The reaction 11→12 as well as the reactions 13→14 and 16→17 require transitions between a triplet and a singlet state. They appear in the workflow because the reactive force field is not aware of electronic states and spin, but they are also part of the real reaction network. The singlet-triplet transitions can either be induced by spin-orbit coupling or collision induced spin-exchange with radicals present in the reaction medium.

The reaction 11→12 has a transition state on the triplet surface. In this case, since the ground state product spin multiplicity has been assigned as singlet whereas the educt was a triplet (species 11), the ChemTraYzer workflow assigned an unrestricted calculation with singlet spin multiplicity to the transition state search. However, after the QM pre-optimization, the Hessian did not have any imaginary frequencies. Later, the triplet transition state has been found via a woelfling calculation. Accordingly, the barrier for the reaction 11→12 (triplet) is 56.2 kJ/mol.

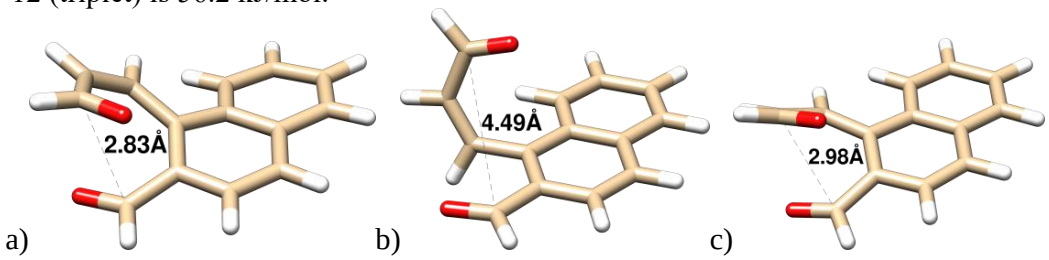


Figure 6 The structures for species 14: a) ReaxFF pre-optimized structure b) QM optimized structure with singlet spin multiplicity using TPSS-D3/SVP. c) QM optimized structure with triplet spin multiplicity using TPSS-D3/SVP.

Another reaction which must be investigated on potential energy surfaces for different spin quantum numbers is the reaction 13→14. This reaction was detected by the processing program of ChemTraYzer as happening in two steps: first, the oxygen atom adds to the molecule (13→14'), then the C—C bond breaks (14'→14). For neither of the reactions the

automated workflow located a transition state. On the singlet surface the optimized reaction path (13→14) has practically no barrier. In the triplet case, the ring opening step has a barrier of 18.2 kJ/mol, whereas the oxygen atom addition is practically barrierless. That is, the reason that the ChemTraYzer workflow could not locate a transition state is due to the practically barrierless reaction step. On the other hand, for the ring opening, because both the product and educt converged to the singlet species 14, the transition state spin multiplicity had been assigned as singlet, which cannot describe the triplet reaction 14'→14. As shown in Figure 6, the triplet product is also closer to the ReaxFF pre-optimized structure for this species. The energetically lower ground state can then be reached by spin-orbit or collision-induced intersystem crossing.

For the unimolecular reaction 16→17 the product structure 17 is not a minimum on the singlet potential energy surface. Therefore, this reaction will only take place if the educt 16 is promoted to the triplet state. It is, however, overall very endothermic and thus unlikely to play a role in the actual reaction network. For this reaction, both the educt and product has converged at the ReaxFF pre-optimization to species 16. As the ground state of the species 16 is a singlet, the ChemTraYzer workflow carried out the transition state calculations singlet multiplicity, which was not correct for this reaction.

5. Summary and Conclusions

The ChemTraYzer program suite has been interfaced with the TURBOMOLE quantum chemistry program package and some improvements have been made concerning i) the input of the structures, ii) the assignment of the spin multiplicities, and iii) avoiding the submission of QM jobs for the same structures. The performance of the modified version has been discussed at the example of two trajectories for the combustion of phenanthrene to demonstrate its potential as a tool for the automated generation of reaction networks for the oxidation of polycyclic aromatic hydrocarbons.

The results demonstrate that it is a valuable tool to identify possible reaction steps and, with some modifications, the workflow was already able to locate the transition states for reactions with barriers above ca. 30 kJ/mol. A remaining problem with the current modified ChemTraYzer implementation is the way product and educt species are identified. They should not be taken from the time step at which a reaction is detected. Instead, the educt and product structures should be taken from time frames with bond orders close to those of the presumed educts and products. The uncertainty caused by the failed transition state optimizations, e.g., due to a lacking imaginary frequency (e.g. reactions 1→2 or 11→12) or a failed optimization despite the existence of a meaningful imaginary frequency (e.g., reaction 5→6 or 7→8'), could be solved with reaction pathway calculations with a chain-of-states method between the educts and products. That is, whenever the transition state optimization fails, the energy profile of the reaction pathway can be used to identify if a reaction is practically barrierless (i.e. when no maximum is found within the resolution of the chain-of-states calculation), or if a maximum is found to obtain a better start guess to for the transition state optimization. This can be achieved by optimizing first the structures of the educts and of the products (which is a van der Waals minimum in the case of bimolecular reactions) using the positions of the corresponding atoms from the corresponding time frames which are close to, respectively, the educts and the products.

Acknowledgements

This project is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 215035359 – SFB/TRR 129 "Oxyflame".

References

[1] P. Freund, “IPCC Special Report on Carbon dioxide Capture and Storage,” 2018.

[2] M. B. Toftegaard, J. Brix, P. A. Jensen, P. Glarborg, and A. D. Jensen, “Oxy-fuel combustion of solid fuels,” *Prog. Energy Combust. Sci.*, vol. 36, no. 5, pp. 581–625, 2010.

[3] G. T. Rochelle, “Amine Scrubbing for CO_2 Capture,” *Science (80-.).*, vol. 325, no. September 2009, pp. 1652–1654, 2014.

[4] A. S. Bhown and B. C. Freeman, “Analysis and status of post-combustion carbon dioxide capture technologies,” *Environ. Sci. Technol.*, 2011.

[5] K. Sumida *et al.*, “Carbon Dioxide Capture in Metal–Organic Frameworks,” *Chem. Rev.*, 2012.

[6] J. A. Mason, K. Sumida, Z. R. Herm, R. Krishna, and J. R. Long, “Evaluating metal-organic frameworks for post-combustion carbon dioxide capture via temperature swing adsorption,” *Energy Environ. Sci.*, vol. 4, no. 8, pp. 3030–3040, 2011.

[7] P. Luis, T. Van Gerven, and B. Van Der Bruggen, “Recent developments in membrane-based technologies for CO_2 capture,” *Progress in Energy and Combustion Science*. 2012.

[8] M. Anheden, J. Yan, and G. De Smedt, “Denitrogenation (or oxyfuel concepts),” *Oil Gas Sci. Technol.*, 2005.

[9] L. Chen, S. Z. Yong, and A. F. Ghoniem, “Oxy-fuel combustion of pulverized coal: Characterization, fundamentals, stabilization and CFD modeling,” *Prog. Energy Combust. Sci.*, vol. 38, no. 2, pp. 156–214, 2012.

[10] O. Senneca *et al.*, “Assessment of combustion rates of coal chars for oxy-combustion applications,” *Fuel*, vol. 238, pp. 173–185, 2019.

[11] R. Hurt, J. K. Sun, and M. Lunden, “A kinetic model of carbon burnout in pulverized coal combustion,” *Combust. Flame*, vol. 113, no. 1–2, pp. 181–197, 1998.

[12] T. Holland and T. H. Fletcher, “Comprehensive Model of Single Particle Pulverized Coal Combustion Extended to Oxy-Coal Conditions,” *Energy and Fuels*, vol. 31, no. 3, pp. 2722–2739, 2017.

[13] J. M. Simmie, “Detailed chemical kinetic models for the combustion of hydrocarbon fuels,” *Prog. Energy Combust. Sci.*, vol. 29, no. 6, pp. 599–634, 2003.

[14] L. Lu, C. Kong, V. Sahajwalla, and D. Harris, “Char structural ordering during pyrolysis and combustion and its influence on char reactivity,” *Fuel*, vol. 81, no. 9, pp. 1215–1225, 2002.

[15] M. Keiluweit, P. S. Nico, M. Johnson, and M. Kleber, “Dynamic molecular structure of plant biomass-derived black carbon (biochar),” *Environ. Sci. Technol.*, vol. 44, no. 4, pp. 1247–1253, 2010.

[16] R. G. Harvey, *Polycyclic aromatic hydrocarbons: Chemistry and carcinogenicity*. Cambridge University Press, 1991.

[17] M. Döntgen, M. D. Przybylski-Freund, L. C. Kröger, W. A. Kopp, A. E. Ismail, and K. Leonhard, “Automated Discovery of Reaction Pathways, Rate Constants, and Transition States Using Reactive Molecular Dynamics Simulations,” *J. Chem. Theory Comput.*, vol. 11, no. 6, pp. 2517–2524, 2015.

[18] M. Döntgen, F. Schmalz, W. A. Kopp, L. C. Kröger, and K. Leonhard, “Automated Chemical Kinetic Modeling via Hybrid Reactive Molecular Dynamics and Quantum Chemistry Simulations,” *J. Chem. Inf. Model.*, vol. 58, no. 7, pp. 1343–1355, 2018.

[19] R. Ahlrichs *et al.*, “TURBOMOLE V7.3 2019, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.,” *Turbomole*. .

[20] F. Furche, R. Ahlrichs, C. Hättig, W. Klopper, M. Sierka, and F. Weigend,

"Turbomole," *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, vol. 4, no. 2, pp. 91–100, 2014.
[21] K. Chenoweth, A. C. T. Van Duin, and W. A. Goddard, "ReaxFF reactive force field for molecular dynamics simulations of hydrocarbon oxidation," *J. Phys. Chem. A*, vol. 112, no. 5, pp. 1040–1053, 2008.
[22] S. Plimpton, A. Thompson, S. Moore, A. Kohlmeyer, and R. Berger, "LAMMPS Documentation," 2019. [Online]. Available: https://lammps.sandia.gov/doc/Manual.html.
[23] S. Nosé, "A unified formulation of the constant temperature molecular dynamics methods," *J. Chem. Phys.*, vol. 81, no. 1, pp. 511–519, 1984.
[24] W. G. Hoover, "Canonical dynamics: Equilibrium phase-space distributions," *Phys. Rev. A*, vol. 31, no. 3, pp. 1695–1697, 1985.
[25] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, "Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids," *Phys. Rev. Lett.*, vol. 91, no. 14, pp. 146401-1,146401-4, 2003.
[26] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu," *J. Chem. Phys.*, 2010.
[27] A. Schäfer, H. Horn, and R. Ahlrichs, "Fully optimized contracted Gaussian basis sets for atoms Li to Kr," *J. Chem. Phys.*, vol. 97, no. 4, pp. 2571–2577, 1992.
[28] V. N. Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew, "Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes," *J. Chem. Phys.*, vol. 119, no. 23, pp. 12129–12137, 2003.
[29] A. Schäfer, C. Huber, and R. Ahlrichs, "Fully optimized contracted Gaussian basis sets of triple zeta valence quality for atoms Li to Kr," *J. Chem. Phys.*, vol. 100, no. 8, pp. 5829–5835, 1994.
[30] G. I. Csonka, A. Ruzsinszky, and J. P. Perdew, "Estimation, Computation, and Experimental Correction of Molecular Zero-Point Vibrational Energies," *J. Phys. Chem. A*, 2005.
[31] K. Eichkorn, F. Weigend, O. Treutler, and R. Ahlrichs, "Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials," *Theor. Chem. Acc.*, 1997.
[32] P. Plessow, "Reaction path optimization without NEB springs or interpolation algorithms," *J. Chem. Theory Comput.*, vol. 9, no. 3, pp. 1305–1310, 2013.
[33] D. Weininger, "SMILES, a Chemical Language and Information System: 1: Introduction to Methodology and Encoding Rules," *J. Chem. Inf. Comput. Sci.*, vol. 28, no. 1, pp. 31–36, 1988.
[34] D. Weininger, A. Weininger, and J. L. Weininger, "SMILES. 2. Algorithm for Generation of Unique SMILES Notation," *J. Chem. Inf. Comput. Sci.*, vol. 29, no. 2, pp. 97–101, 1989.
[35] B. Helmich and M. Sierka, "Similarity recognition of molecular structures by optimal atomic matching and rotational superposition," *J. Comput. Chem.*, vol. 33, no. 2, pp. 134–140, 2012.
[36] L. Martínez, R. Andrade, E. G. Birgin, and J. M. Martínez, "Packmol: A package for building initial configurations for molecular dynamics simulations," *J. Comput. Chem.*, vol. 30, no. 13, pp. 2157–2164, 2009.
[37] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, and G. R. Hutchison, "Open Babel: An Open chemical toolbox," *J. Cheminform.*, vol. 3, no. 33, 2011.
[38] "Open Babel v2.3.1 documentation: Radicals and SMILES extensions." [Online].

Available: https://openbabel.org/docs/dev/Features/Radicals.html. [Accessed: 11-Dec-2019].
[39] R. Ahlrichs *et al.*, "Chapter 3.3 Running Turbomole using the script TMOLE," in *Turbomole User's Manual Version 6.0*, 2009, pp. 38–44.
[40] T. Helgaker, "Transition-state optimizations by trust-region image minimization," *Chem. Phys. Lett.*, vol. 182, no. 5, pp. 503–510, 1991.
[41] P. Marshall and P. Glarborg, "Ab initio and kinetic modeling studies of formic acid oxidation," *Proc. Combust. Inst.*, vol. 35, no. 1, pp. 153–160, 2015.
[42] V. Dias, C. Duynslaegher, F. Contino, J. Vandooren, and H. Jeanmart, "Experimental and modeling study of formaldehyde combustion in flames," *Combust. Flame*, vol. 159, no. 5, pp. 1814–1820, 2012.

Supplementary Materials

Supplementary materials include the xyz coordinates of the species, their single point energies, zero point energies, expectation values of the S^2 operator, as well as the energy profiles of the reactions which are assigned as practically barrierless together with the corresponding xyz files named as 'path-*.xyz'.

V. An automatized workflow from molecular dynamic simulation to quantum chemical methods to identify elementary reactions and compute reaction constants

In this work, we present a new ReaxFF trajectory analysis program, which is suitable for using with large molecular systems. The refinement of the detected bond breaking/formation events via density functional theory sheds light on the strengths and weaknesses of the ReaxFF.

In this work, I have contributed to the design of the database structure and analysis of the trajectories only during brainstorming. The coding and creation of the database, identification of reaction events, or filtering of unique reactions, etc. were done by Prof. Rochus Schmid and Dr. Gunnar Schmitz. I have only contributed to debugging of this part by communicating the problems to Dr. Schmitz as they arise, and he did the debugging.

My major contribution to this work is that I have written the python module for the automated quantum chemistry calculations. In this process, I have actively solved the problems as they emerge either independently or through discussions with my supervisor Prof. Christof Hättig. For example, the idea of using Z-matrices to create the reacting complexes came from Prof. Hättig. The analysis of the trajectory presented in section "Quantum Chemical Calculations" is also mainly done by me and of course finalized through discussions with Prof. Hättig, Prof. Schmid, and Dr. Schmitz.

An automatized workflow from molecular dynamic simulation to quantum chemical methods to identify elementary reactions and compute reaction constants

Gunnar Schmitz*, Özlem Yönder†, Bastian Schnieder‡
Rochus Schmid§, Christof Hättig¶

May 7, 2021

Abstract

We present an automatized workflow which, starting from molecular dynamics simulations, identifies reaction events, filters them and prepares for accurate quantum chemical calculation with e.g. Density Functional Theory (DFT) or Coupled Cluster methods. The capabilities of the automatized workflow are demonstrated at the example of simulations for the combustion of some polycyclic aromatic hydrocarbons (PAHs). It is shown how key elementary reaction candidates are filtered out of a much larger set of redundant reactions and refined further. The molecular species in question are optimized using DFT and reactions energies, barriers heights and reaction rates are calculated. The setup is general enough to include at this stage configurational sampling, which can be exploited in the future. Using the introduced machinery we investigate how the observed reaction types depend on the gas atmosphere used in the molecular dynamics simulation. For the re-optimization on the DFT level we show how the additional information needed to switch from reactive force-field to electronic structure calculations can be filled in and study how well ReaxFF and DFT agree with each other and shine light on the perspective of using more accurate semi-emiprical methods in the MD simulation.

Keywords: Reaction finder, ReaxFF, DFT, automatized workflow, MD simulations. ■

*Corresponding author: gunnar.schmitz@rub.de, Computational Materials Chemistry group, Ruhr-Universität Bochum, D-44780 Bochum, Germany

†Corresponding author: oezlem.yoender@rub.de, Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, D-44780 Bochum, Germany

‡Computational Materials Chemistry group, Ruhr-Universität Bochum, D-44780 Bochum, Germany

§Computational Materials Chemistry group, Ruhr-Universität Bochum, D-44780 Bochum, Germany

¶Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, D-44780 Bochum, Germany

1 INTRODUCTION

Quantum chemical methods are nowadays important tools to model chemical reactions and supply researchers with additional data that is not easily accessible by experimental techniques. Due to the algorithmic progress over the last decades methods like density functional theory (DFT) or even some coupled cluster (CC) methods[1–4] can be routinely applied to large molecules with up to 100 atoms without creating a computational bottleneck. However, in large heterogeneous systems chemical reactions are often very complex in nature and require more than one or two reaction steps. In those cases selecting the elementary reaction steps that are quantum chemically studied only based on chemical or human intuition becomes problematic since the chemical reaction space is too huge to be sampled by man power alone. Thus, automatic schemes and tools are necessary to propose possible reactions which can be closer investigated by DFT or CC methods.

There are different routes to such a scheme. One is defining reaction families. The MIT reaction mechanism generator (RMG) is such an example, which has been used to generate kinetic models for the combustion of several fuels such as *n*-butanol[5], diisopropyl ketone[6], and JP-10 jet fuel[7]. RMG uses molecular graphs to convert the reactants to the products by the available reactions from its reaction library[8,9]. The thermodynamic properties are then calculated by means of Benson's additivity rules, semi-empirical or ab initio quantum chemistry methods[8,9]. Such kinetic model generators require a pre-knowledge on the type of reactions which would occur. On the other hand, for arbitrary systems, being restricted to reaction families may hinder the inclusion of new chemical reactions valid for that system and lead to biased results[10]. Other routes, which do not require this knowledge, employ the exploration of potential energy surfaces (PES). These can either be seen as an alternative or complementary to the reaction family based approaches. Chemoton is an example to such a reaction network discovery tool developed by Reiher and co-workers, which makes use of conformer generators to explore PESs while minimizing the exhaustivity of the possible reactions by introducing descriptors based on physical quantities such as the electron density[10]. Additional to this explorative scheme, there are also other automated first-principle approaches using different aspects or targeting a different scope. Having products and educts fixed,

Habershon[11] could for example demonstrate how to automatically identify different reaction pathways by introducing a Hamiltonian, which describes a reaction string connecting reactants and products, and performing molecular dynamics (MD) sampling of this Hamiltonian. The reactants and products in their approach are connected via graph theory.

Another approach to explore PESs is to start from reactive molecular dynamics simulations based on cheap force-field or semi empirical methods in order to pre-sample possible relevant reactions. In this work we propose such a workflow with the special interest in modelling the combustion of polycyclic aromatic hydrocarbons (PAHs) under oxyfuel conditions. Our goal for this initial publication is not to demonstrate it for an exhaustive exploration of the combustion reactions, but to introduce the work flow and to explore its potential for identifying important reaction families within the limitations of the reactive force field, ReaxFF[12,13]. This route, as many others, requires reduction of the number of reaction events that are detected on the MD trajectory. We achieve this by comparisons of molecular graphs for the involved molecular species and thereby identifying unique reactions.

For the automated framework our ideas are similar to those presented by Döntgen et al. and Kröger et al. in Refs. 14–16 which inspired the current work. We avoid, however, to analyse trajectories in batches and keep track of the configurations of the molecular species throughout our workflow and and avoid the reduction of chemical structure information to SMILES strings. Like in the work or the latter authors we use ReaxFF for the MD simulation due to its computational efficiency. However, for our specific interest ReaxFF has some weaknesses like a certain inertness of O_2. This makes it mandatory to perform the MD at rather high temperatures and under high pressure to trigger reaction events. This approach is acceptable since the MD simulation in this setup is considered as a sampling technique in the reaction space rather than a simulation approach for the combustion process. The actual reactions are then closer investigated by subsequent electronic structure calculations. This approach is not new. It is used by ChemTrayzer[14–16] and is, for example, also pursued by Wang et al.[17] in their chemical nano reactor, where they additionally apply periodically a force acting on all atoms directed to the center of the simulation box. Other ways to increase the rate of reaction events is to add artificial attractive interatomic forces[18–20]. All these approaches allow to identify possible reaction events, but this opens two new challenges

which we want to address in this work. First, the number of reaction events can become quite large and it is important to filter only those which are of interest for the main research objective. Second, the molecular structures extracted from the MD trajectories are due to the high temperature usually quite deformed and far from ground state equilibria, which makes subsequent optimization, at, for example, the DFT level, difficult, especially for the transition states. Here we present an automated scheme to handle those cases and comment on challenging aspects. The species can be (re-)optimized on various level of theory (LoT) to obtain reaction pathways and accurate reaction energies and activation barriers.

The paper is organized as follows: First we outline the basic theory of our automatized workflow. This is followed by a detailed overview of the computational settings used in this study. Afterwards the results are presented and discussed. For this purpose the reaction event screening is closer investigated. On a qualitative level it will be studied how the atmosphere applied in the simulation affects the observed reaction types and, more importantly, if common reaction types important for combustion can be identified. In the last part we present some exemplary DFT results and discuss some challenges that arise from using the reactive force field ReaxFF as precursor method for the MD simulations.

2 COMPUTATIONAL METHODS

2.1 The workflow

Our integrated work flow starts with MD simulations with the trajectories stored in HDF5 files[21]. The first central part of the workflow, which connects the MD simulations with the subsequent quantum chemical calculations, is the program *findR* (vide infra) which analyses the trajectories frame by frame and stores identified reaction events in a relational SQL database. The reaction events found are then filtered to arrive at a non-redundant set of reactions. As similarity measure we use different approaches based on graph theory, but the setup is general enough to use also other kinds of similarity measures. During this process it can also be ensured that not only the identified reactions are non-redundant, but that also the molecular species are stored in a non-redundant fashion. As next step a reaction

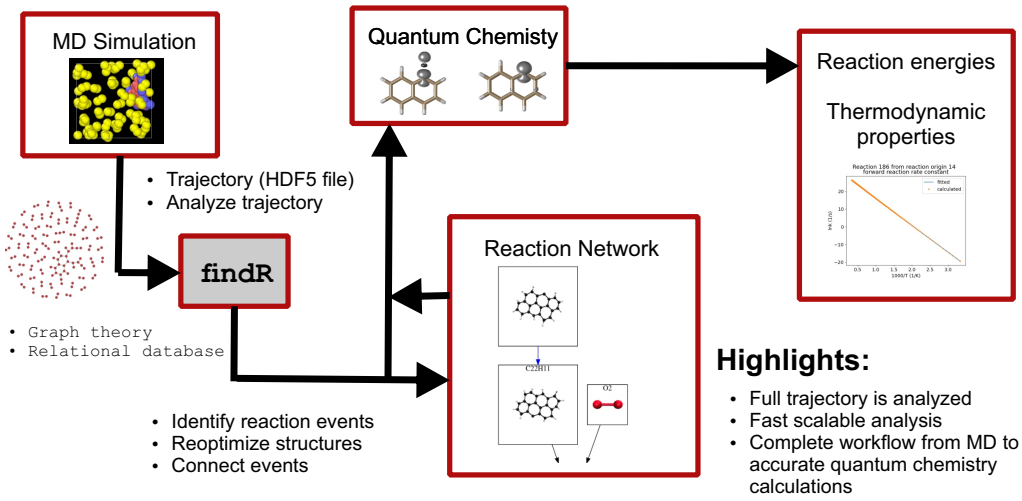


Figure 1: Illustration of the computational workflow employing *findR* for analyzing MD trajectories and the subsequent steps involving quantum chemistry calculations for determining reaction rates and thermodynamic parameters.

network can be created which connects the different reaction events for visual analysis and to ensure that no reactions were missed. In the Supporting Information we give two examples for this visual analysis as HTML file generated by our program. Furthermore, and key to our purpose, the found reactions can be investigated closer with (higher-level) quantum chemical methods. The educts, products, and transition states are optimized, which allows to calculate barriers and reaction rates.

2.2 Database layout

Before going into detail how the reaction database is populated and which algorithms are in play to avoid redundancies and to automate the quantum chemical calculations we want to give a simplified overview of the database layout. A simplified schematic overview of the tables in the database is depicted in Fig. 2. Arrows indicate which tables relate to each other and the arrow directions indicates the direction of the reference. The central table is `reac2spec` which links an abstract `species` table to the `reactions` table. The `species` table is general for a molecular species e.g. water and several optimized species (`opt_species`)

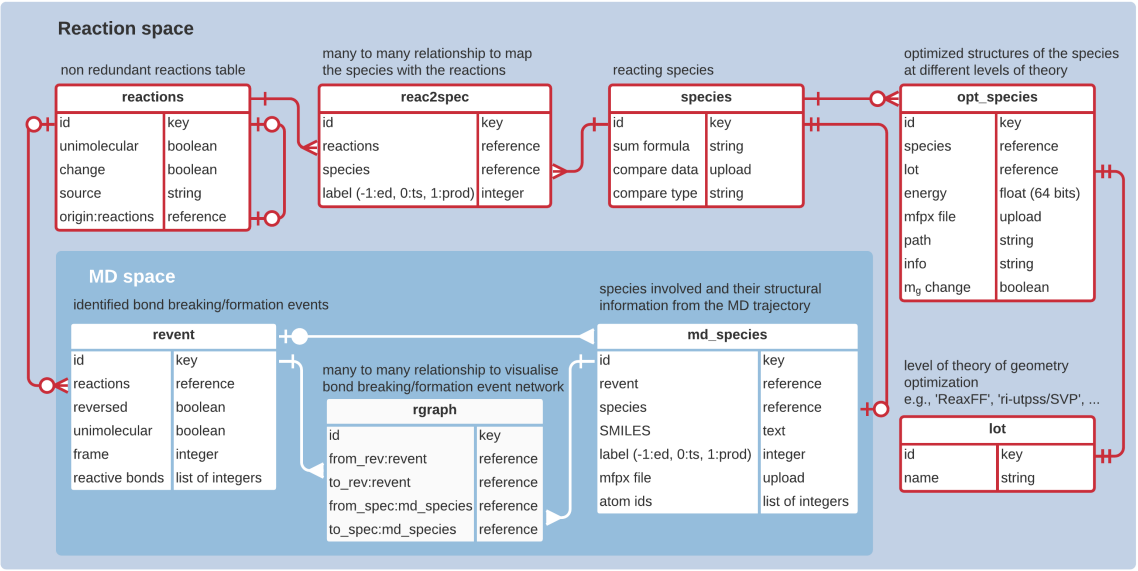


Figure 2: Simplified layout of the used database to store reactions and molecular species.

can point to it. It does not contain structural information, but is just a global label for a species and the idea is to keep this table non-redundant. Geometries can be retrieved on various LoT from the `opt_species` table, where they are stored as `mfpx` file, which is a self-developed file format containing the coordinates and the connectivity. However, the `species` table contains an abstract object which is the basis of comparisons. Currently we only use molecular graphs (vide infra), but other similarity measures can be implemented and stored in the database. The `reactions` table stores all reactions and allows for self referencing in the case a studied reaction can spawn new reactions e.g. due to configurational sampling or different possible spin cases in the involved educts and products. We refer to the `species`, `reac2spec`, `reactions`, `LoT` and `opt_species` table as reaction space and the `md_species` and `revent` table as MD space. The latter is extracted from the MD simulation and the tables belonging to the reaction space are obtained from post-processing steps. The first post-processing step is the filtering step described in section 2.5 which reduces the reaction events stored in `revent` to a non-redundant set.

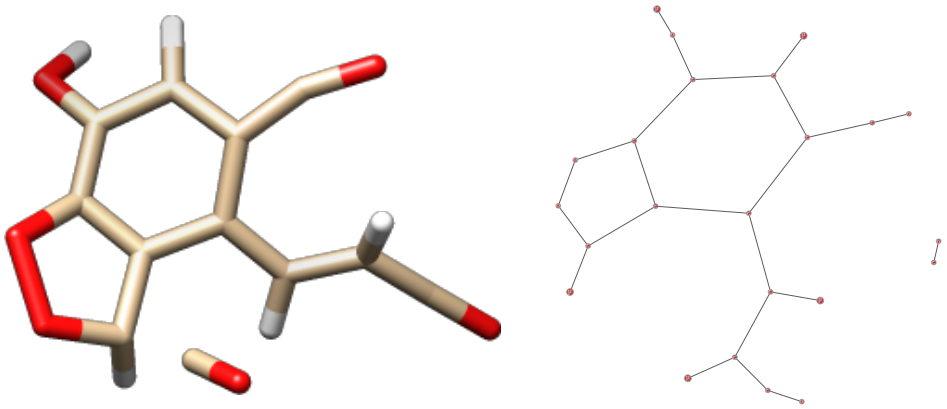


Figure 3: Illustration of how a molecular structure is mapped onto a graph, where the vertices are the atom types and the bonds are incorporated as edges. The depicted case shows a reactive complex involving two molecular species.

2.3 Molecular graph representation

Basis of all comparisons of molecular species in the trajectories is graph theory. A so-called molecular graph m_g is constructed for each molecular species, where the vertices are given by the atom types and the bonds are represented using edges. Figure 3 demonstrates how a molecular structure is converted into a graph. The connectivity for the graph construction can be obtained by several means. In the current work it is based on the ReaxFF bond order, but other schemes could be applied as well. Just looking at the visualization of the graph, it looks like a 2-dimensional projection of the molecular structure, but the graph representation has some advantages. There are efficient available algorithms[22,23] to identify if two graphs are isomorphic. Graph isomorphism ensures the same connectivity, but does not rely on the atom order or the relative orientation of functional groups. This makes it a very robust comparison measure. A drawback is that for many reactions the configuration of the involved molecular species matters. However, a configurational sampling can always be added additionally after the initial filtering. From this step onward new reactions can be added anytime in the reaction database and reference to the reaction where they originate from.

2.4 Identifying reaction events

To identify reaction events the frames in the MD trajectory are compared with each other on the basis of the molecular species. Since we are interested into the combustion of polycyclic aromatic hydrocarbons (PAHs) and not in the (known) gas phase chemistry of small molecules, not all species are tracked, but only those that fulfill certain criteria, for our examples only those which have at least 6 atoms and are composed of at least 2 carbon atoms. Starting from the first frame it is checked from frame to frame if the graph for the tracked species changed. For efficiency reasons the trajectory is not searched frame by frame, but the frame index is incremented by a user defined step size (default 200). If a change in a molecular system is detected, the search direction is reversed using a smaller step size (default 10) which will be reduced down to 1 to locate the frames, where the reaction event took place. There are two levels of comparison: At level 1 it will be checked if the species are composed of the same atoms. If the composition does not change it will be checked on level 2 if the bonding changed. As mentioned above, bonds are determined based on the ReaxFF bond order, where per default a bond is classified by having a bond order larger 0.5. Thus, changes on level 2 classify unimolecular reactions while changes on level 1 indicate a reaction with 2 or more reaction partners. If a reaction event was found three frames, representing the educt(s), product(s), and the transition state, are stored for later analysis. During the frame comparison always two frames are compared with each other. If a reaction event was located the frame in which the bond order of the reactive bond is closer to 0.5 is chosen as transition state frame and this frame is stored as frame for the reaction event. The frames for educt(s) and product(s) are chosen relatively (-1,+1) to the transition state frame. The stored frames, however, do not contain minimum energy structures, which makes it later necessary to optimize the species for which different LoTs can be used. Following the outlined recipes all reaction events in the MD are registered, but the number of reaction events can easily become quite large which makes it mandatory to use filtering mechanisms, like those described in the next section.

2.5 Filtering non-redundant molecular species and reaction events

One strength of the *findR* scheme is that the whole trajectory is analyzed and that one is not in danger of overlooking reaction steps. However, this strength also creates some challenges. Many reaction events occur during an MD simulation and some, if not most, of them are redundant or oscillatory motions where bond orders dropped only for a few frames below the threshold. Therefore, it is pivotal to develop a scheme to filter only the unique reactions. Along the same lines it is pivotal to store only a non-redundant set of molecular species. We will first describe how to filter the molecular species. For this purpose the full molecular graph is used, but additionally its stereo centers are considered in order to differentiate stereoisomers. For the detection of the latter we use the algorithm implemented in OpenBabel[24] to identify chiral centers. The structures in the database can later be optimized at various levels. As outlined in Fig. 2 our database is flexible enough to allow for different comparison measures besides the molecular graph. However, if the used comparison type differs two molecular species are automatically regarded as dissimilar.

To differentiate a reaction event r from its unique reaction class we use in the following a bar ($\bar{r}$). The list of reactions $\mathbf{r}$ contains all observed and possible redundant reactions while $\bar{\mathbf{r}}$ is a set of unique reaction events. Algorithm 1 outlines the filtering that is applied. After identifying all elementary reactions $\mathbf{r}$ we loop over them and compare the molecular species $\mathbf{m}^{(r)}$ i.e. educts, products and the transition state participating in the found reaction r with the analogue molecular species $\mathbf{m}^{(\bar{r})}$ in the set of already known unique reaction events $\bar{\mathbf{r}}$. The comparison is made on the basis of the molecular graphs constructed as outlined in sections 2.3; however, a few fast checks based on for example the sum formula are implemented to short cut graph comparisons if the atomic compositions differ.

During the comparison it is tracked if the molecular species belong to the educt or product side to classify the reaction r as forward or backward reaction. If there is no match in $\bar{\mathbf{r}}$ the reaction r is added as new unique reaction. The first reaction of its kind is always classified as forward reaction.

For the comparison we employ two schemes. In the first one, the comparison is made using the same molecular graph as used in the initial analysis to identify reaction events,

but in the second scheme we only consider the molecular graph based on certain atom types. Thus, a reaction event is only found if the connectivity to such an atom type changed. For a C-graph this means that we only consider C-C bonding. Reactions like H abstraction or the addition of O_2 are not regarded as reactions. For our intended application in char combustion a sane choice is to restrict the atom types to carbon (if the model system does not contain potential hetero-atoms, e.g., nitrogen or sulfur). This essentially gives two layers of filtering. We start from an initial list of all reactions r, but then filter only the unique reactions $\bar{r}$ based on the molecular graph. However, also this shortened list still contain reactions in which the molecular C-graph does not change. If we also consider the change in the C-graph we arrive at the final stage of filtering $\bar{r}^{(C)}$.

The number of reaction events $n_{\bar{r}}^{(C)}$ is a measure for the decomposition of the carbon structure. In Section 3.2 we investigate how this number depends on conditions like the atmosphere. During the analysis we need to keep track of the reactive species for a reaction. They will be stored in lists $\mathcal{L}_{m,t}$, where m refers to the molecular graph and t to the type (educt, product, or transition state).

Algorithm 1 Filtering unique reactions from redundant list

Read in all already found unique reactions $\bar{\mathbf{r}}$

Loop over all reactions in $\mathbf{r}$

for $r = r_1, ..., r_n$ **do**

Get set of molecular species $\mathbf{m}^{(r)}$ in r

for $m^{(r)} \in \mathbf{m}^{(r)}$ **do**

Identify type t as educt, product or transition state

Create molecular graph $m_g^{(r)}$

Store $m_g^{(r)}$ and t in list $\mathcal{L}_{m,t}$

end for

for $\bar{r} = \bar{r}_1, ..., \bar{r}_m$ **do**

for $m^{(\bar{r})} \in \mathbf{m}^{(\bar{r})}$ **do**

Identify type τ as educt, product or transition state

Create molecular graph $m_g^{(\bar{r})}$

Store $m_g^{(\bar{r})}$ and τ in list $\mathcal{L}_{\bar{m},\tau}$

end for

if $\mathcal{L}_{m,t} = \mathcal{L}_{\bar{m},\tau}$ **then**

$\rightarrow r \in \bar{\mathbf{r}}$

Exit Loop

end if

end for

if $r \notin \bar{\mathbf{r}}$ **then**

Add r to $\bar{\mathbf{r}}$

end if

end for

2.6 Refining the reaction and activation energies using density functional theory

A second central part of the automated workflow is a new self-developed module which automates the quantum chemistry calculations with the TURBOMOLE program package[25]. More specifically, this module is used to get initial equilibrium and transition state structures from the database and refine them to obtain accurate thermodynamic and kinetic data. In the current work, the ReaxFF optimized structures are used as initial structures and only the non-redundant (unique) reactions are considered. For the refinement of the molecular structures, density functional theory (DFT) is used.

In a first step, the ReaxFF optimized initial structures of the educts, products, and transition states are refined using a computationally efficient low-cost DFT functional (avoiding exact exchange) and a small basis set, TPSS-D3/SVP in this study. The data for reactions refined at the DFT level contains inherently different information than available from the ReaxFF pre-optimization. Therefore, each reaction at the DFT level is registered as a new reaction to the database together with the information about the reaction origin, i.e., the reaction ID of the reference MD ReaxFF reaction. Together with each new reaction, the successfully located transition state structures and the equilibrium structures connected to them or for the practically barrierless reactions just the equilibrium structures of the educts and products are registered in the optimized species table (`opt_species`) of the database. Afterwards, the data about these species can be retrieved from the database, re-optimized using a more accurate functional and a larger basis set, e.g., TPSSh-D3/TZVP, and be registered in the optimized species table with the new level of theory. The general workflow of the refinement of reaction and activation energies using an inexpensive DFT functional and small basis set is shown in algorithm 3.

Since the MD with ReaxFF provides only molecular geometries and energies, the quantum chemical electronic structure calculations need additional information to specify the electronic wavefunction. These are the charges (or number of electrons) and the z component of the electronic spin, which sets a lower bound for the electron spin and the number of unpaired electrons. Therefore, to determine the occupation of the molecular orbitals,

we have used Fermi smearing as implemented in TURBOMOLE[26] with unrestricted DFT orbitals as indicated in algorithm 2. In the current study it is assumed that all molecular species are charge neutral. For the equilibrium structures, an initial occupation for Fermi smearing is assigned based on the number of electrons in the system: for those with even number of electrons as (n_α-n_β) = 2, $m_s^{start} = 1$, and for those with odd number of electrons as (n_α-n_β) = 1, $m_s^{start} = 1/2$, where n_α and n_β are the number of electrons with, respectively, α and β spin and $m_s = (n_\alpha - n_\beta)/2$ is the secondary spin quantum number. The initial orbitals are generated using an Extended Hückel Theory (EHT) guess. After the Fermi smearing fractional occupation numbers are rounded to the closest integer values, resulting in a new value m_s^{Fermi} for the secondary spin quantum number. To assure that the electronic state with the lowest energy is found, the DFT energy is then also evaluated with the secondary spin quantum number increased and, if possible, also decreased by 1 from the value found by the Fermi smearing ($m_s^{Fermi} \pm 1$). The spin orbital occupation that resulted in the lowest energy (m_s^{Emin}) is then kept for the subsequent calculations.

Algorithm 2 Determine orbital occupations of equilibrium structures

if odd number of electrons **then**

$m_s^{start} = 1/2$

else if even number of electrons **then**

$m_s^{start} = 1$

end if

Create start orbitals from extended Hückel Theory $\rightarrow m_s^{start}$

Minimize the Kohn-Sham energy and orbital occupations with Fermi smearing $\rightarrow m_s^{Fermi}$

Turn off Fermi smearing

Minimize the Kohn-Sham energy of the species by modifying occupations $m_s^{Fermi} \pm 1$

Choose the one with lowest energy $\{m_s^{Fermi} - 1, m_s^{Fermi}, m_s^{Fermi} + 1\} \rightarrow m_s^{minE}$

Optimize the geometry of the species with $m_s^{minE} \rightarrow x^{QM,ref}$

Skip reaction

if $m_g^{(\mathrm{ReaxFF},ref)} = m_g^{(QM,ref)}$ **then**

return $m_s^{final} = m_s^{minE}$

else

Optimize the geometry of the species with $m_s^{minE} + 1 \rightarrow x^{QM,ref}$

if $m_g^{(\mathrm{ReaxFF},ref)} = m_g^{(QM,ref)}$ **then**

return $m_s^{final} = m_s^{minE} + 1$

else

Skip reaction

end if

end if

Then a geometry optimization is performed. To avoid that geometry optimizations for equilibrium structures get trapped on saddle points because a start structure had a higher symmetry than the searched-for minimum on the PES, a random noise with a range of $\pm 0.05\,\text{Å}$ is added to the coordinates of the ReaxFF optimized structures. For the transition state structures, the noise is not applied to the atoms which belong to the reacting bonds (see below). If the molecular graph changes, additional geometry optimizations are done with the original start geometry with $m_s^{Emin} + 1$. If the graph still changes, the reaction is

skipped because this means that this reactions has no counterpart at the DFT level. Also reactions where the initial educt or product structures dissociate into two or more molecules during the geometry optimization are skipped, because then it is not trivial to automatically track the actual reaction.

To ensure that the occupations of the spinorbitals of the educts, the transition state, and the products of a reaction are compatible also in non-unimolecular cases, we define for the educts and for the products sets of possible m_s value, m_s^{educts} and $m_s^{products}$, respectively, that can be reached by coupling of the individual spins, as shown in algorithm 3. If the two sets contain common m_s values which would result in a spin-conserving reaction, then only these values are considered. In the workflow for the location of the transition state on the DFT PES we have then the possibility to either refine the ReaxFF pre-optimized transition state, i.e. the structure which is taken from the MD trajectory where the bond order was 0.5 and then geometry optimized by constraining the bonds involved in the reaction using ReaxFF ($x^{\mathrm{ReaxFF},BO=0.5}$), with a geometry optimization for saddle points or, if this fails, to use a double ended reaction pathway optimization (*vide infra*). If the sets of possible values for the secondary spin quantum numbers of the educts and products do not contain a common element ($\mathbf{m}_s^{educts} \cap \mathbf{m}_s^{products} = \{\}$), only the transition state optimization workflow based on the ReaxFF start structure is carried out, trying in this case all possible values ($\mathbf{m}_s^{educts} \cup \mathbf{m}_s^{products}$). The educts and products of the reaction on the DFT PES are then determined starting from the DFT transition state as described below. For elementary reactions which involve more than two molecules the generation of a reactive complex (*vide infra*) is a complex problem. Such elementary reactions are, however, very rare. This happens at most for 1-3 reactions in an analyzed trajectory. Therefore, these cases are not automated and the program skips these reactions in this initial proof of concept study.

Algorithm 3 Automation of density functional theory calculations

Read in all unique reactions $\bar{\mathbf{r}}$ from the database

for $r = \bar{r}_1, ..., \bar{r}_n$ **do**

Read the ReaxFF geometry optimized reference structures $\mathbf{x}^{\mathrm{ReaxFF}}$

Add random noise $\delta x_{max} = 0.1\,\text{Å}$ to $x^{\mathrm{ReaxFF}} \in \mathbf{x}^{\mathrm{ReaxFF}}$

for $x^{\mathrm{ReaxFF}} \in (\mathbf{x}^{\mathrm{ReaxFF},educts} \cap \mathbf{x}^{\mathrm{ReaxFF},products})$ **do**

Determine the orbital occupations and optimize the geometries:

$x^{\mathrm{ReaxFF}} \rightarrow x^{QM,ref}, m_s^{QM,ref}$ (Algo. 2)

end for

if $n_{educts} \leq 2$ **and** $n_{products} \leq 2$ **then**

Couple the spins $\begin{cases} \mathbf{m}_s^{educts} = \left\{(m_s^{ed_1} + m_s^{ed_2}), (m_s^{ed_1} + m_s^{ed_2}) - 1, ..., |m_s^{ed_1} - m_s^{ed_2}|\right\} \\ \mathbf{m}_s^{products} = \left\{(m_s^{prod_1} + m_s^{prod_2}), (m_s^{prod_1} + m_s^{prod_2}) - 1, ..., |m_s^{prod_1} - m_s^{prod_2}|\right\} \end{cases}$

if $\mathbf{m}_s^{educts} \cap \mathbf{m}_s^{products} = \{\}$ **then**

for $m_s \in (\mathbf{m}_s^{educts} \cup \mathbf{m}_s^{products})$ **do**

Pre-optimize the TS structure $x^{\mathrm{ReaxFF},BO=0.5} \rightarrow x^{\dagger,QM,ref,m_s}$

Optimize TS structure and IRC end points $x^{\dagger,QM,ref,m_s} \rightarrow x^{\dagger,m_s}, x^{IRC,m_s}$ (Algo. 4)

end for

else

for $m_s \in (\mathbf{m}_{s,educts} \cap \mathbf{m}_{s,products})$ **do**

Pre-optimize the TS structure $x^{\mathrm{ReaxFF},BO=0.5} \rightarrow x^{\dagger,QM,ref,m_s}$

Optimize TS structure and IRC end points $x^{\dagger,QM,ref,m_s} \rightarrow x^{\dagger,m_s}, x^{IRC,m_s}$ (Algo. 4)

if $m_g^{(IRC)} \neq m_g^{(QM,ref)}$ **then**

Calculate approximate reaction path using woelfling

Select the peaks of the energy profile as TS start structures $x^{\dagger,QM,ref,m_s}$

Optimize TS structure and IRC end points $x^{\dagger,QM,ref,m_s} \rightarrow x^{\dagger,m_s}, x^{IRC,m_s}$ (Algo. 4)

end if

end for

end if

end if

end for

The module uses two alternative algorithms to locate a transition state. The first one is partially adopted from ChemTraYzer[14–16] and creates a start structure for the transition state optimization by constrained minimization of the ReaxFF pre-optimized 'transition state' structure ($x^{\mathrm{ReaxFF},BO=0.5} \rightarrow x^{\dagger,QM,ref}$). For this, the internal coordinates of the atoms which belong to the "reacting" bonds as identified by *findR*, i.e., the bonds for which the bond order changes (see Section 2.4), are fixed during this initial energy minimization. This resulting structure will be referred to as DFT pre-optimized TS structure. The Hessian of this structure is calculated using the *aoforce* module of TURBOMOLE[27,28]. Then, the constraints on the structure are removed for the optimization of the TS structure, i.e. the saddle point, with the trust-region image minimization (TRIM) algorithm[29] as implemented in the *statpt* program of TURBOMOLE. Then, the Hessian is calculated for the optimized saddle point. If this confirms that a saddle point is found successfully, i.e., that the geometry optimization converged to a stationary point with only one imaginary frequency ($N_{\nu_{im}}(x^{\dagger}) = 1$), an intrinsic reaction coordinate (IRC)[30] calculation is used to test if the saddle point connects the expected educts and products ($x^{QM,ref}$). It is checked if there are multiple molecules at the optimized end points, which should be the case for non-unimolecular reactions, and if this is the case the molecules are separated using the topology of the molecular graph[31]. The separated geometries are optimized following the same procedure as described in algorithm 2 to determine the occupation numbers for the α and β spin orbitals of the individual fragments. Afterwards, the molecular graphs of the optimized IRC end points ($\mathfrak{m}_g^{IRC}$) are compared to those of the reference educts and products ($\mathfrak{m}_g^{QM,ref}$). It should be noted that the educts and products are not necessarily the same as those detected during the *findR* step. If the geometry-optimized IRC end points (x^{IRC}) differ from the expected educts and products ($\mathfrak{m}_g^{IRC}$), the equilibrium structures found from the IRC end points (x^{IRC}) are assigned instead of the former as its educts and products.

If the stationary point found by the saddle point optimization has no imaginary frequencies or if it is connected to other educts or products than expected, an algorithm to locate the right transition state starting with a double-ended reaction path optimization between the DFT-optimized educt and product structures ($x^{QM,ref}$) is employed in those cases where this is possible (see above). For this we use the *woelfling* module[32] of TURBOMOLE. Then,

the maxima on the PES along the reaction pathways are chosen as new start structures for transition state optimizations.

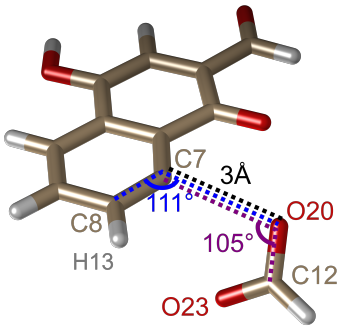


Figure 4: Sketch of the Z-matrix coordinates that determine the position and orientation of the formate radical relative to the PAH structure. The three torsion angles are not shown.

For bimolecular reactions this requires the combination of the structures of two reactants, stored separately in the database, to a reactive complex to get the start or end structures for the reaction path optimization. This is achieved by combining the Z-matrices of the two reacting species by making use of the Python module chemcoord[33]. In order to position the second molecule with respect to the first one, the missing internal coordinates related to the first three atoms of the second molecule have to be defined. Figure 4 shows an example of a reaction where a formate radical is released by breaking a C-O bond. Accordingly, atom O20 (cmp. Fig. 4) is set to be the reactive atom in the formate species and located at a distance of 3 Å from atom C7 in the remaining PAH fragment. For all internal coordinates of the separate fragments the values are taken from the respective optimized geometries. (The Z-matrices of the two products, the ReaxFF pre-optimized transition state structure, and the product complex are given in the Supporting Information.) The relative orientation of the fragments is determined as follows: For atom O20, the bond angle ($\angle$(O20C7C8)) and the dihedral angle (τ(O20C7C8H13)) are set to the values in the ReaxFF pre-optimized transition state structure. Similarly, for C12 the angle ($\angle$(C12O20C8)) and the dihedral (τ(C12O20C8H13)), and for O23 the dihedral (τ(O23O20C12H13)) are set to the values in the ReaxFF pre-optimized transition state structure. After the structure for the reactive complex has been generated, the orbitals and occupations of the DFT pre-optimized TS

structure are used to start the reaction path optimization.

For unimolecular reactions, before carrying out the reaction path optimization, the educt and product structures obtained at the DFT level ($x^{QM,ref}$) are compared via a similarity recognition program described in Ref. 34 which uses atomic matching and rotational superposition. If educt and product are similar, which can occur due the oscillatory motions at the high temperaturs in the MD trajectory (see above), the reaction path optimization is skipped. This almost never happens for bimolecular reactions as the reacting species are optimized separately.

Algorithm 4 Determine transition state (TS) and connect it with educts and products

Calculate the Hessian matrix of the TS start structure $\rightarrow \frac{\partial^2 E}{\partial (x^{\dagger,QM,ref})^2}$
if $N_{\nu_{im}}(x^{\dagger,QM,ref}) \geq 1$ **then**
  Trust region image minimization $x^{\dagger,QM,ref} \rightarrow x^{\dagger}$
  if Geometry optimization is converged **and** $N_{\nu_{im}}(x^{\dagger}) == 1$ **then**
    Optimize the end points from IRC $\rightarrow x^{IRC}$
    Calculate their Hessian matrices $\rightarrow \frac{\partial^2 E}{\partial (x^{IRC})^2}$
    if $N_{\nu_{im}}(x^{IRC}) == 0$ **then**
      Compare $m_g^{(IRC)}$ with $m_g^{(QM,ref)}$
      Register new reaction and optimized species $(x^{\dagger}, x^{IRC})$ to the database
    end if
  end if
end if

2.7 Calculation of reaction energies and rate constants

After the initial refinement of the reactions with the computationally non-demanding density functional, the optimizations of the geometries and calculations of the Hessian matrix with a more accurate density functional and basis set combination is straightforwardly automated using the lower level DFT data as start guesses. The additional refinement is especially important as the results of the DFT calculations like the molecular structures and harmonic vibrational frequencies enter the calculation of thermodynamic properties and rate constants

(*vide infra*). Since the accuracy of DFT energies is often not sufficient for the requirements on reaction energies and rate constants it is usually necessary to get more accurate electronic energies from wavefunction based methods as, for example, coupled cluster theory.

The reaction rate constants are calculated using transition state theory (TST). The partition functions needed for this are calculated using a modified version of the quasi-rigid rotor and harmonic oscillator (qRRHO) approximation[35,36]. Using the ideal gas law as the equation of state, the site specific reaction rate constants (excluding the symmetry numbers) can be calculated at constant pressure within TST according to

$$k(T) = \kappa \frac{k_B T}{h} \frac{Q_{\mathrm{TS}}/V}{\prod_{i=educts}(Q_{\mathrm{i}}/V)} \exp\left(-\frac{\Delta E_{el,act}}{RT}\right) \tag{1}$$

where R is the gas constant, k_B Boltzmann's constant, V the molar volume, T the temperature, h Planck's constant, and Q is the partition function which includes translational, vibrational, and rotational contributions. The vibrational partition function (q_{vib}) is here defined relative to the bottom of the potential energy surface as

$$q_{vib} = \prod_i \frac{\exp(-h\nu_i/2k_BT)}{1-\exp(-h\nu_i/k_BT)}, \tag{2}$$

where ν_i are the vibrational frequencies. κ in Eq. 1 is the transmission coefficient used to account for tunneling. It is calculated using the high temperature limit of the Eckart tunneling correction:[37]

$$\kappa = 1 - \frac{1}{24}\left(\frac{h\nu_{im}}{k_BT}\right)^2 \left(1 + \frac{RT}{\Delta E_{el,act}}\right), \tag{3}$$

where ν_{im} is the imaginary frequency and $\Delta E_{el,act}$ is the electronic activation energy without zero point vibrational correction. For a compact representation of the results, the parameters A, n, and E for the modified Arrhenius equation,

$$k(T) = AT^n \exp\left(-\frac{E}{T}\right) \tag{4}$$

where T and E are in units of K, are fitted to the rate constants calculated via Eq. 1 with a non-linear least square fitting for the temperature range of 300–3000 K.

3 SHOWCASE STUDY

In the following we want to demonstrate the presented automated workflow on the basis of studying the combustion of PAHs under oxyfuel conditions. First we summarize the used computational settings used in our study in Section 3.1. In Section 3.2 the introduced filtering scheme is assessed and how the composition of the gas phase affects the number of reactions. Afterwards follows a qualitative analysis of found reaction types in Section3.4 and in Section 3.5 the quantitative assessment of reactions employing the automated refinement via DFT calculations.

3.1 Computational details

The MD simulations were done with our in-house python interface to the LAMMPS Molecular Dynamics Simulator[38] employing the reimplementation of the reactive Force-Field ReaxFF[39]. For the ReaxFF calculations the "CHO" parameterization for combustion as published in Ref. 13 was used. All MD simulations reported in this study use the following settings if not otherwise stated: Periodic boundary conditions in a cubic simulation box were applied and depending on the case the size varied from 20 – 25 Å. The simulations were run for 500 ps with a time step of 0.1 fs at 2000 K using a Nosé–Hoover thermostat. Each simulation was repeated 5 times using different initial random velocities.

For the manipulation of graphs the graph-tools library[31] is applied and the SQLite[40] through database abstraction layer of web2py[41] is used to manage the relational database to store reaction events, reaction lists as well as the molecular species.

All quantum chemical calculations were performed with the Turbomole program package[25] version 7.5[26]. For the DFT calculations the ridft module was used. Basis sets applied in the DFT calculation include the TZVP[42] and SVP[43] basis sets together with the corresponding auxiliary basis sets for the resolution of the identity approximation from Ref. 44. Furthermore, Grimme's D3 dispersion correction[45] was included. As functionals TPSS[46] and TPSSh[47] were chosen. The TPSSh/TZVP zero point vibrational frequencies have been scaled by the factor of 0.9874 that was optimized for TPSSh/6-31+G(d,p) to fit the experimental zero point vibrational energies (ZPVE)[48].

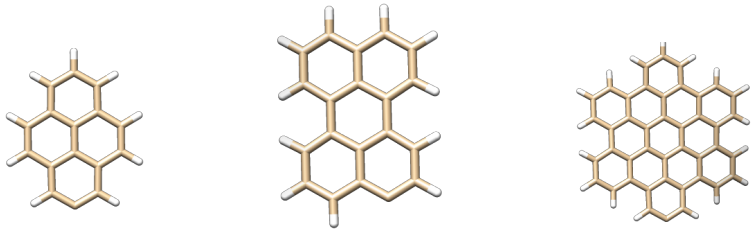


Figure 5: Systems (without O_2 and OH atmosphere) studied by ReaxFF MD simulation and analysis by *findR*.

In addition explicitly correlated coupled cluster singles doubles and perturbative triples, CCSD(F12*)(T*) calculations were carried with the ccsdf12 module[49] based on preceding restricted open-shell Hartree-Fock (ROFH) calculations with the dscf module[50]. The cc-pVDZ-F12 orbital[51], auxiliary[52,53], and complementary auxiliary basis sets[54] are used with the recommended exponent[51] γ=0.9 for the correlation factor $\exp(-\gamma r_{12})$.

3.2 Reaction screening

For the initial exploration we used the three systems shown in Fig. 5 as examples. We run simulations in an atmosphere enriched with OH radicals (40 O_2 molecules and 10 OH radicals) as well as in a pure O_2 atmosphere to investigate if these conditions lead to different reaction types. In all cases we started from a PAH, where already one H atom was abstracted to enhance reactivity. The H abstraction process was not studied in detail here since we already investigated it in a recent publication[35].

First we want to address the question how many unique reactions according to the defined uniqueness measure can be identified in the MD simulations and how this will increase performance in the subsequent analysis on the basis of for example quantum chemical calculations. Table 1 gives an overview of the number of non-redundant and total reactions found in the MD simulations in the OH enriched atmosphere, which were repeated each 5 times with different initial velocities. We further differentiate in the filtered reactions if the molecular C graph changed upon reaction.

In all cases a large reduction in the number of reactions is observed. Especially for the larger systems savings above 90% are possible. This translates directly to less computational

costs for the post-processing with more accurate quantum chemical methods, e.g. DFT, and during this stage reactions can again be added to the database if it turns out that for a specific example it is important to account for different conformations.

Table 1: Statistics of filtered reactions in OH enriched atmosphere. Given are the found unique reaction $n_{\bar{r}}$ as well as all found reactions n_r. Additionally the number of reactions is given in which the molecular C-graph changed upon reaction $n_{\bar{r}}^{(C)}$ and thus an important step in the combustion process happened. The screening ratio "Screen.", defined as the fraction of reactions that are skipped, given in %.

	(a) $C_{16}H_9$					(b) $C_{20}H_{11}$					(c) $C_{42}H_{17}$				
run	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
n_r	268	1348	853	1265	696	1061	775	2204	483	912	5943	7974	2458	159	1405
$n_{\bar{r}}$	43	143	169	134	147	71	209	260	91	554	195	281	100	48	72
Screen.	84	89	80	89	79	93	73	88	81	39	97	96	96	70	95
$n_{\bar{r}}^{(C)}$	10	43	49	44	49	13	84	83	46	63	51	95	22	13	17

3.3 Influence of radical enriched atmosphere

Next, we address the question how a radical enriched atmosphere can trigger reaction events in the sense that more reactions occur. Unfortunately performing ReaxFF simulations at experimentally realistic conditions is computationally too demanding and previous studies employed too high temperatures which leads to a bias towards fragmentation instead of oxidation reactions. To assess how the radical atmosphere results in more reactions also MD simulations in a pure O_2 atmosphere were performed and are compared to the previous ones. The statistics over the found reactions are listed in Table 2. While the average number of all reactions increased, the number of unique reactions and in particular unique reactions including a change in the C-graph decreased in most cases. The on average larger number of reaction events can be explained by O_2 attaching to the PAHs and detaching again in a subsequent step. These non-reactive collisions of O_2 have also been observed in hydrocarbon oxidation simulations by Chenoweth et al[13]. Furthermore, one can make a few more general

observations: In those cases where many unique reaction events occur some radical species in the atmosphere were created. If no gas phase radicals are involved, the reactions occur mainly in close proximity to the initial radical site at the PAH. At least at the ReaxFF level of theory one can conclude that the formation of radical species in the gas phase is an important step in the combustion process.

Table 2: Statistics of filtered reactions in pure O_2. Given are the found unique reactions $n_{\bar{r}}$ as well as all found reactions n_r. Additionally the number of reactions is given in which the molecular C-graph changed upon reaction $n_{\bar{r}}^{(C)}$ and thus an important step in the combustion process happened. The screening ratio Screen. which represents the skipped reactions is given in %.

	(a) $C_{16}H_9$					(b) $C_{20}H_{11}$					(c) $C_{42}H_{17}$				
run	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
n_r	1643	2494	1166	809	978	3829	3733	50	498	3933	601	3942	1842	1311	1673
$n_{\bar{r}}$	103	32	11	34	14	11	8	9	41	7	15	7	63	77	17
Screen.	94	99	99	96	99	99.9	99.9	82	92	99.9	98	99	97	94	99.9
$n_{\bar{r}}^{(C)}$	38	10	3	8	4	2	2	20	1		3	0	20	28	8

3.4 Identification of common elementary reactions

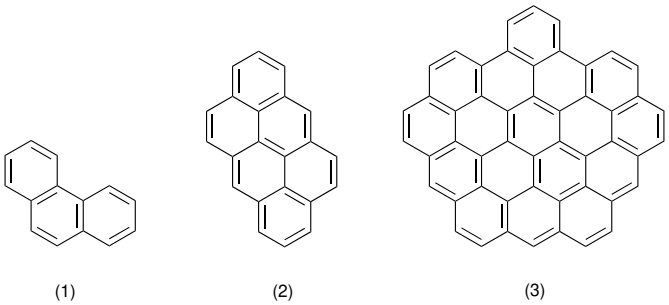


Figure 6: Model systems for the molecular dynamic simulations to identify key elementary reactions in the combustion process.

Table 3: Composition of the atmosphere in the molecular dynamic simulation as well as the studied system given in Fig. 6. For each system we repeated the simulation 5 times with different initial random velocities. Later in the text these simulations are referred to as run 1, run 2 etc.

Molecular Dynamic Simulation	System	hydroxyl radicals	oxygen molecules	water molecules
MD 1	(1)	10	40	-
MD 2	(1)	15	60	-
MD 3	(1)	10	40	10
MD 4	(2)	12	40	-
MD 5	(3)	12	40	-

A key feature of *findR* is the identification of elementary reactions. Our current focus lies on the atomistic modelling of combustion reactions of char models. Therefore, it is essential that *findR* can identify common key reaction mechanisms. To assess this, we created an additional test set of systems depicted in Fig. 6 where we again varied the composition of the atmosphere. An overview can be found in Table 3. Each MD was done 5 times using different initial velocities. For this purpose we generated reaction networks for each of the simulations and inspected them by hand to identify common reaction steps. This type of analysis only allows a qualitative description. In Section 3.5 we will present a more elaborate way to analysis reactions by means of post-processing with DFT.

Before discussing some common reaction steps that were found in the MD simulations, we want to discuss a few more general observations: For larger PAHs more reaction events are observed, which can be trivially explained by the larger number of reaction sites. Furthermore, we observed that the PAHs are decomposed starting form the exterior ends. Using different pathways, H atoms are abstracted, and radical species attach to the PAHs and initiate rearrangements. In these rearrangements some common reaction steps are observed like the formation of five-membered rings as well as the opening of 6-rings. In the following we will list a few typical reactions found in the simulations. We give the examples as found in the MD, but note that they represent general reactions that can occur on outer rings of PAHS or graphene-like areas in char.

Shrinkage of a 6-membered ring usually happens via the formation an additional bond that divides it into a 5- and a 3-membered ring. In subsequent steps the latter is opened and a fragment, e.g. CO_2 or formic acid, is released. Examples to these are shown in Figure 7

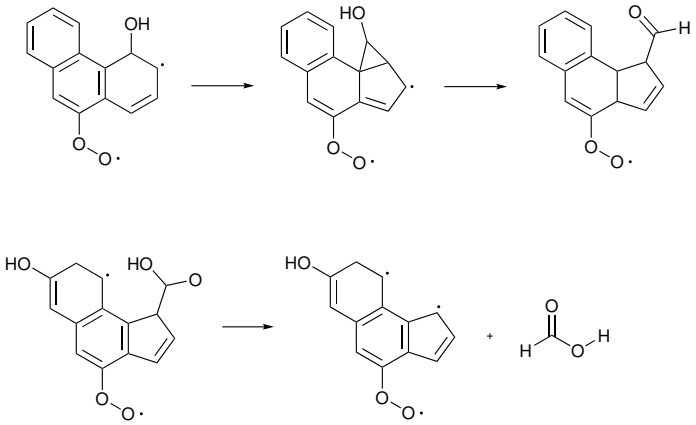


Figure 7: Top: Formation of a 5-membered ring via ring shrinkage. Bottom: Release of formic acid from a 5-membered ring.

Ring extension usually proceeds via insertion of O into the ring structure after a preceding formation of an epoxide as shown in Figure 8.

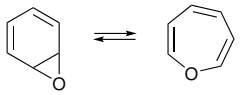


Figure 8: Benzene-oxide/oxepin tautomers.

The opening of a 6-ring occurs due to homolysis after oxidation of two neighboring carbon positions to a bicarbonyl such as represented in Figure 9.

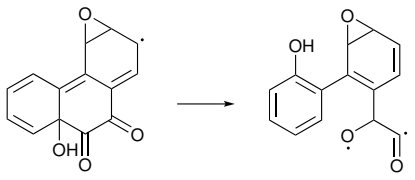


Figure 9: An example ring opening reaction.

We observed several times that formic acid is a favoured leaving group in a OH radical enriched atmosphere while in atmospheres with lower OH radical concentration CO and CO_2 are favoured leaving groups as depicted in Figure 10.

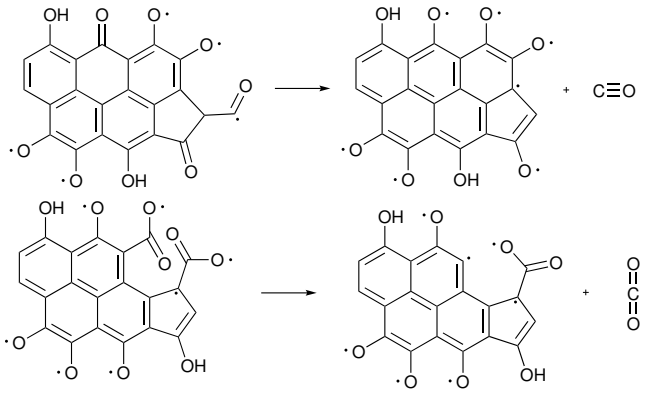


Figure 10: The release of CO and CO_2.

3.5 Quantum Chemical Calculations

Table 4: Educt, transition state, and product structures optimized at the TPSSh-D3/TZVP level for some reactions detected in the run 3 trajectory of MD3. Spin densities obtained from ROHF calculations are shown as isosurface plots. In the last two columns we give the CCSD(F12*)(T*) results for activation and reaction energies including the ZPVE corrections from TPSSh-D3/ZZVP, $\Delta E = \Delta E_{\mathrm{el,CCSD(F12*)(T*)/cc-pVDZ-F12}} + \Delta E_{ZVPE,\mathrm{TPSSh-D3/TZVP}}$

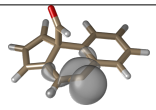
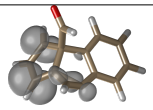
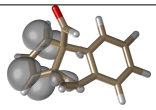
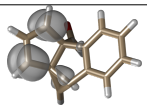
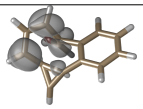
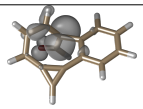
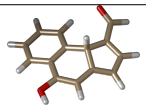
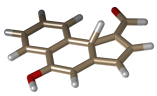
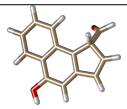
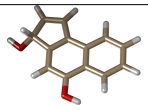
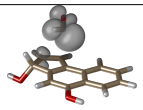
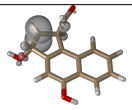
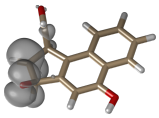
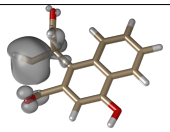
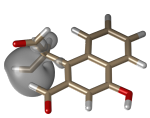
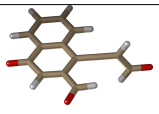
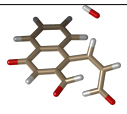
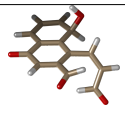
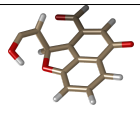
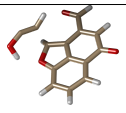
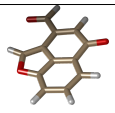
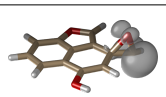
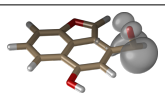
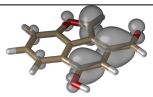
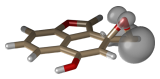
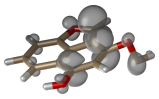
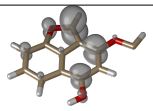
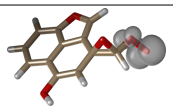
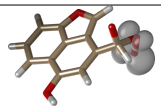
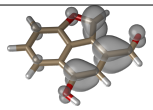
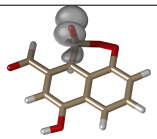
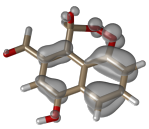
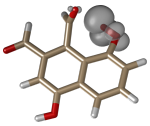
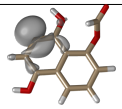
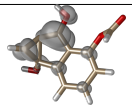
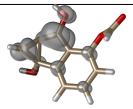
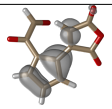
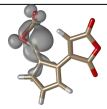
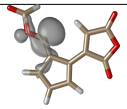
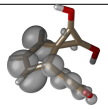
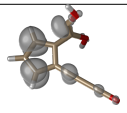
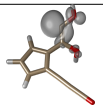
	Educt(s)	Transition State	Product(s)	ΔE_{act} kJ/mol	ΔE_{rxn} kJ/mol
1	 m_s=1/2	 m_s=1/2	 m_s=1/2	122.3	116.3
2	 m_s=1	 m_s=1	 m_s=1	82.5	16.9
3	 m_s=0	 m_s=0	 m_s=0	77.6	-54.0
4	 m_s=0 CHO m_s=1/2	 m_s=1/2	 m_s=1/2	20.1	-77.8
5	 m_s=1	 m_s=1	 m_s=1	80.4	62.5
6	 m_s=0 H_2O m_s=0	 m_s=0	 m_s=0	163.1	64.7
7	 m_s=0	 m_s=0	 m_s=0 ethynol m_s=0	7.4	-86.1

Table 5: The species geometry optimized at TPSSh-D3/TZVP level as obtained from the run 3 trajectory of MD3. The ROHF spin densities are shown. $\Delta E = \Delta E_{\mathrm{el,CCSD(F12*)(T*)/cc-pVDZ-F12}} + ZVPE_{\mathrm{TPSSh-D3/TZVP}}$ (cont.)

	Educt(s)	Transition State	Product(s)	ΔE_{act} kJ/mol	ΔE_{rxn} kJ/mol
8	 m_s=1/2	 m_s=1/2	 m_s=1/2	12.9	-258.1
9	 m_s=1/2	 m_s=1/2	 m_s=1/2	140.7	47.6
10	 m_s=1/2	 m_s=1/2	 m_s=-1/2 O_2 m_s=1	46.8	-328.2
11	 m_s=1/2	 m_s=1/2	 m_s=1/2	41.6	1.1
12	 m_s=1/2	 m_s=1/2	 m_s=1/2	202.5	201.5
13	 m_s=1/2	 m_s=1/2	 m_s=1/2	191.8	111.4
14	 m_s=1/2	 m_s=1/2	 m_s=1/2	52.4	-29.5

So far we discussed a more qualitative assessment of the reaction events found by *findR*. In this section we want to focus on the automated refinement with quantum chemical methods, which allows a more quantiative assessment.

Due to the approximate nature of ReaxFF as mentioned before it is not always straight forward to refine the found reactions on the DFT level. In the following we want to discuss some of these challenges on the basis of a particular trajectory we analyzed. For this purpose the run 3 trajectory from MD 3 (cmp. Table 3) was selected as an example since the used PAHs are sill relatively small and allow even Coupled Cluster calculations. Furthermore, this trajectory contains some examples for the certain types of challenges we faced and want to address in this work. Tables 4–5 summarize for this trajectory the reactions that were identified and successfully refined with the automated workflow until there was only one ring left. In Tables 4–5, reactions in which the molecular graph does not change, such as CHO or OH internal rotations, and practically barrierless reactions, such as many cases of O_2 addition or elimination, are not shown. The assignment of the reaction directions is in general arbitrary (as explained in Section 2.6). Only if the molecular graphs of the IRC end points (x^{IRC}) match those of the educts and products ($x^{QM,ref}$) that were obtained by refinement from the corresponding ReaxFF structures the reaction directions are defined in accordance with the MD trajectory. The activation and reaction energies given in Tables 4–5 include the zero point vibrational energy (ZPVE) corrections at the TPSSh-D3/TZVP level of theory and the electronic energies from the ROHF-CCSD(F12*)(T*)/cc-pVDZ-F12 calculations. The rate constants of each elementary reaction steps shown in these tables are calculated with as given in equation 1. The calculated rate constants are fitted to the modified Arrhenius equation (equation 4) and the fitting parameters can be found in the Supporting Information.

One of the challenges for the automatic refinement can be present in unimolecular reactions. For the given example among 73 unique unimolecular reactions, the educt and product structures extracted from the trajectory converged for 53 of them during the ReaxFF geometry optimization to equilibrium structures with the same molecular graph, indicating an intermediate bond breaking or formation reverted after a few time steps rather than a true reaction, which could be detected during the refinement. The reason why the educt and product structures from ReaxFF are the same is that these structures are, as explained in section 2.4, extracted directly from the trajectories of the MD simulations that are run at 2000 K where the molecular structure undergoes large deformations that, even though they

lead to changes in the ReaxFF bond order, are often not real reactions in the sense of transitions from one potential well over a saddle point to another well on the PES. That is, as the educt and product are not located through an intrinsic reaction coordinate search from an optimized transition state, it is not guaranteed that they are close to two equilibrium structures, i.e. minima on the PES which are connected by a transition state. Nevertheless, the automated refinement module tries for theses cases to locate a saddle point on the DFT PES starting from the pre-optimized ReaxFF transition state structure and find from the saddle point on the DFT PES the educt and product structures that are connected to the located transition state. The reactions 1, 2, 12, and 14 are such examples, where geometry optimizations of the structures taken from the MD trajectory ended up at the ReaxFF level in the same equilibrium structures, but where the ReaxFF pre-optimized transition state structures were successfully used as start guesses for the optimization of a transition state at the DFT level. As it can be seen in Table 4, they correspond to either C-C bond formation or breaking. For all four of these reactions the products are, however, structurally rather close to the transition state, which partially explains why on the ReaxFF PES no minima were found that correspond to the product structures on the DFT PES. For those reactions where the transition state guess from the ReaxFF structure was not a good initial structure, the reaction path calculation either located a conformational change such as rotation of a -OH group or it wasn't carried out because the educt and the product is the same (see Section 2.6). The automated refinement workflow located for 24 of the 73 unique unimolecular reaction events a saddle point on the DFT PES. 14 of them belong to reactions with different molecular graphs for the educt and product.

Other difficulties arise when the MD trajectory deviates qualitatively from (minimum energy) reaction pathways on the DFT PES. For example, the educt of reaction 1 in Table 4 is on the MD trajectory formed through several elementary reaction steps where the automated workflow has been challenged. The first elementary step has been on the MD trajectory a simultaneous addition of OH and O_2. As the automated refinement with DFT has not been implemented for intermolecular reactions, this step has been skipped in the proof-of-concept study. The following step, H abstraction from the OH group with another OH radical, results in an oxy radical (R-O·) which is stable only with three unpaired electrons (and a water

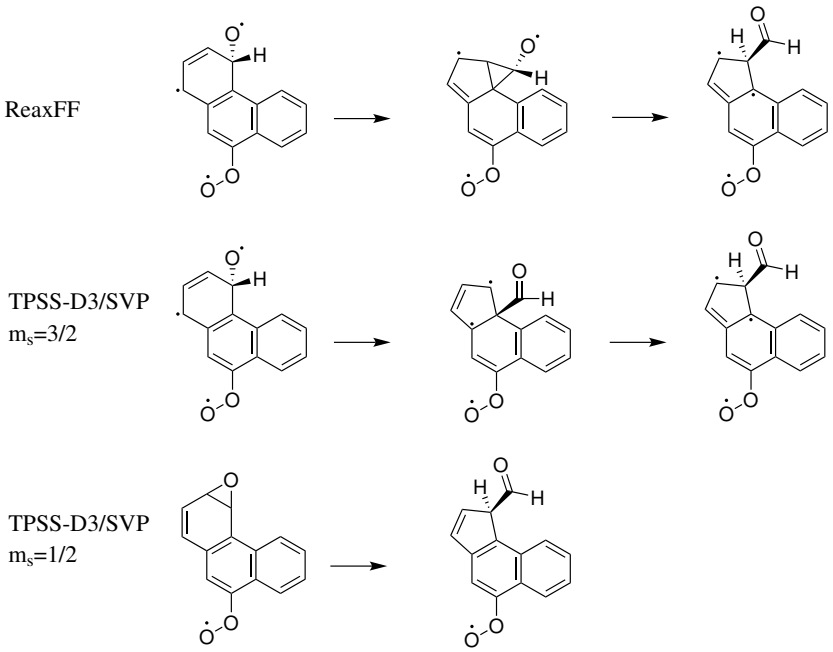


Figure 11: The elementary steps of the formation of the formyl radical following the formation of an oxy-radical or an epoxide.

molecule). On the PES for one unpaired electron, the equilibrium structure converges to an epoxide. Starting from the oxy radical or the epoxide, the following elementary steps of the formyl group formation are as shown in Figure 11. A problem for the H abstraction from the R-OH group by a hydroxy radical is the inability of ReaxFF to correctly describe radicals. As shown in Figure 12, the ReaxFF optimized educt has a structure that is intermediate between a water molecule bonded to an oxy radical and a hydroxyl radical bonded to an OH functional group, because 111 pm is obviously too short to be an intermolecular bond. Indeed when optimized at the DFT level, a water molecule separates from the rest. In the future one can potentially also avoid these difficulties in the first place by using MD simulations based on methods like the extended tight binding methods[55,56] (xTB), which should more closely resemble the DFT PES. For the outlined structure in Figure 12 we also optimized the molecular species obtained from the ReaxFF MD simulation on the GFN0-xTB, GFN1-xTB and GFN2-xTB level. All structures optimized on the xTB level show a better agreement with the DFT optimized ones than ReaxFF. This does especially hold for GFN1-xTB and

GFN2-xTB.

It can occur that ReaxFF optimized species exhibit unusual bonding, which is then resolved upon optimization on the DFT level. If this optimization leads to fragmentation of the species the reaction is currently flagged for inspection, but in the future it is planned to identify the reactive center(s) and register new reactions in the database. To keep track of which reactions originate from each other our database layout has the `origin` cell which can relate different reactions.

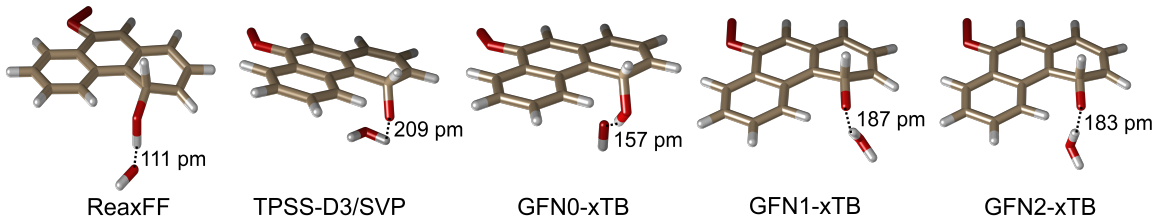


Figure 12: Comparison of the ReaxFF, DFT ($m_s = 3/2$), and xTB optimized educt structures for the detected dehydration reaction.

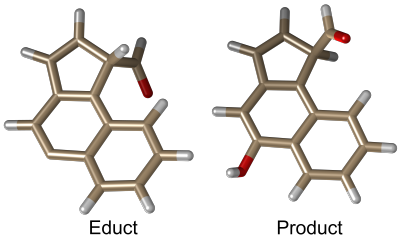


Figure 13: The educt and product structures of the detected OH addition, as geometry optimized with ReaxFF.

During the ReaxFF simulations, functional groups connected to the ring systems by single bonds such as alcohol (-OH) or aldehyde (-CHO) groups rotate or stereoisomerize in some occasions. The latter results in educt and product structures for a detected reaction where these functional groups are flipped relative to the plane of the aromatic rings as shown in Figure 13. The reason for the stereoisomerism not being detected by *findR* is that, as explained in section 2.4, only the composition and the bonding of the atoms (based on the molecular graphs) are used to detect the reaction events. Due to the two simultaneous reactions

happening at the same time and the large structural change during the stereoisomerisation, the default number of iteration steps used during the woelfling calculations (maximum of 40 iterations) as well as number of intermediates might not be enough to optimize a good approximate reaction path. Even so, the peaks along the approximate reaction path calculated from woelfling may be good initial points and converge to a transition state structure upon trust region image minimization. Reaction 3, migration of an H atom to a neighboring carbon, was located this way. The detected reaction event from the MD trajectory was addition of an OH radical, whose detected educt and product structures are shown in Figure 13, however, the formyl group flips along the trajectory very shortly after the OH is chemisorbed. Also reaction 4, addition of the formyl radical, was found in a similar way during the automated DFT refinement. The reaction event detected for the MD trajectory was again the addition of an OH radical. Thus, the quantum chemical refinement can in those cases resolve the issues related with *findR* not differentiating stereoismers. In the future a configurational sampling step for each involved reaction species can be included. The setup of the database is general enough to register new reactions at any step and to relate them with each other. Both of the mentioned reaction (3 and 4) might be an elementary step of the isometrization reaction. However, further investigation should be made. These reactions can easily be queried from the database for further analysis by selecting the reactions in which the molecular graph of the educt and product species are different than those of the start structures (here as reference to the ReaxFF optimized structures).

Sometimes, even if there is no simultaneous stereoisomerization happening, another reaction can be located where the initial educts and products ($x^{QM,ref}$) do not match the IRC end points (x^{IRC}). For example, reaction 6 is found during the reaction path optimization of a OH addition: when the OH radical approaches the molecule it passes through a structure close to the shown transition state structure and corresponds to a maximum on the energy profile along the reaction path. Another such example is reaction 11, which was located starting from the pre-optimized ReaxFF transition state structure. Compared with the non-optimized structures taken directly from the MD trajectory the molecular graphs of the educts and products match, however, upon geometry optimization of these structures with ReaxFF the educt and product converge to the same structure. Therefore, as compared

with the reference level of theory, i.e., using ReaxFF optimized structures as reference, the molecular graphs of the IRC end points do not match the initial ones.

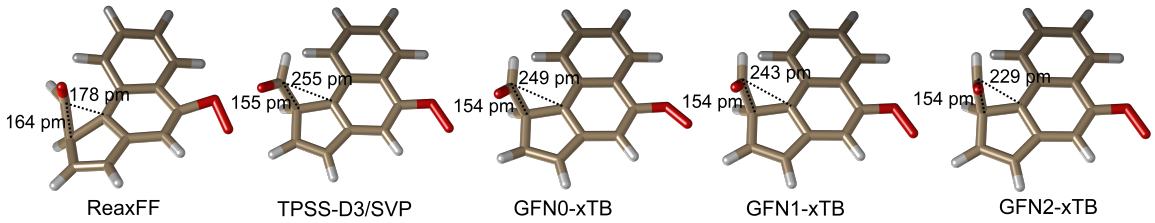


Figure 14: Comparison of the ReaxFF, DFT ($m_s = 1/2$), and xTB optimized structures.

Another obstacle for the DFT refinement of the reaction events detected in the ReaxFF MD trajectories is that some of the ReaxFF optimized equilibrium structures are too far from minima on the DFT PES with the same molecular graph or where corresponding minima do not at all exist on the DFT PES. An example is shown in Figure 14, which is at the ReaxFF LoT a stable structure with 5- and 3-membered rings. As mentioned above in the text and in Figure 11, in the MD simulation the formation of the formyl group occurs through ring shrinkage of a 6-membered ring to 5- and 3-membered rings. However, neither on the $m_s = 3/2$ or $m_s = 1/2$ DFT PESs such a 3-membered ring exists as an equilibrium structure. On the DFT PES such a structure is instead rather close to a transition state. The depicted structure in Figure 14 has also been geometry optimized with xTB methods starting from the MD structure. For this species as well the structures optimized on the xTB level show a better agreement with the DFT optimized ones than ReaxFF. In a previous study by Frenklach et al.[57] it has been reported that for 4-phenanthrol a similar shrinkage of a 6-membered ring through expulsion of -COH group happens in a single step. It is known that the reaction mechanisms at different reactive sites depend on the rigidity of the molecule at the reactive site[58] and ReaxFF might not be able to capture these subtle differences. Therefore, use of MD simulations based on xTB methods might be a solution to such cases. As at the DFT PES surface this reaction happens at a single step at this reactive site, it is not possible to find an equilibrium structure at the DFT surface whose graph would not change. Hence, in such cases elementary reaction steps on the ReaxFF PES do not have a direct counterpart on the DFT PES. For these cases the equilibrium structures found for

the educts and products on the DFT PES have to be added as new species to the database and the saddle point(s) connecting them can be located starting with a reaction pathway optimization with these species as start and end points.

Another example to the cases where the ReaxFF optimized equilibrium structures are too far from minima on the DFT PES is shown on the left of the Figure 15. The program *findR* has detected O_2 addition to the shown ReaxFF optimized structure, which has the same molecular graph as that of the educt of the reactions 8 and 9. For this detected reaction, in the ReaxFF optimized educt structure, the epoxide ring is almost parallel to the plane of the aromatic rings. The unrelaxed scan of the CCCH torsional angle (Figure 15) shows that this ReaxFF structure is far from the equilibrium structure at the DFT level which has the same molecular graph. For the DFT equilibrium structures where the CCCH torsional angle is around 160 ° and 340 °, the epoxide group is almost perpendicular to the aromatic ring plane. Because of this mismatch between the ReaxFF and DFT equilibrium structures, a geometry optimization at the DFT level starting from the ReaxFF structure ends up in another equilibrium structure where the epoxide ring has opened to a formyl group. On the other hand, the transformation of the formyl group to the epoxide is highly endothermic (reverse reaction 8). As can be seen in the following reactions, the ether was converted back to the formyl group. Thus, for modelling the kinetics of combustion processes, the potential, but unstable, epoxide intermediates that are not found by the automated DFT refinement procedure do not play a significant role.

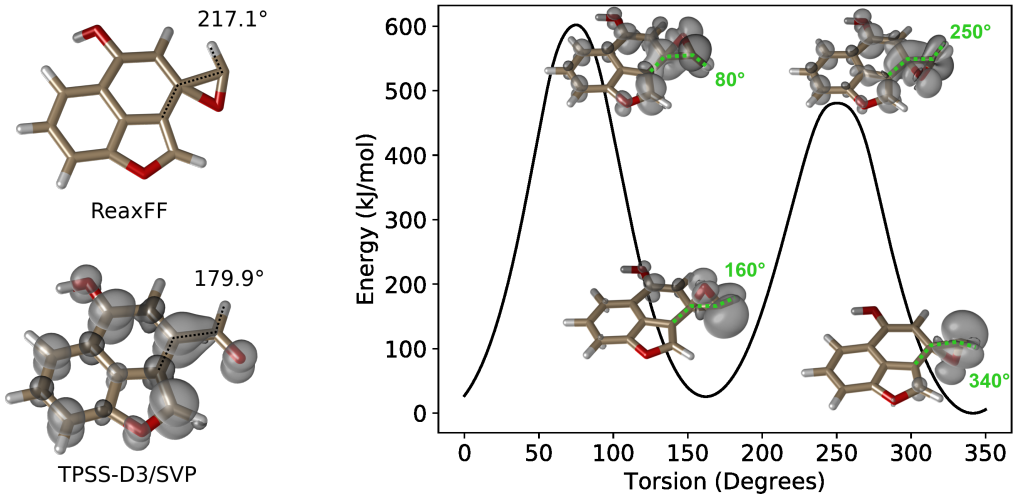


Figure 15: Left: Comparison of the ReaxFF and DFT optimized structures for an educt where the ReaxFF equilibrium structure was far from the minimum which describes the equilibrium structure with the same molecular graph on the DFT PES. Right: The unrelaxed scan of the CCCH torsional angle of the educt of the reactions 8 and 9 using TPSS-D3/SVP.

Reaction 5 is successfully located through the ReaxFF pre-optimized transition state. With reaction 7 the ethynol group (HCCOH) leaves the molecule. This reaction as well is successfully located starting from the ReaxFF pre-optimized transition state. The reaction 10, O_2 elimination, is another reaction where the IRC end points match the reference educt and the product. The transition state of this reaction is found through a start guess from the reaction path optimization. In reaction 13 the ketone group is converted to an ether group. Also for this case the automated refinement procedure could locate the transition state on the DFT PES starting directly from the pre-optimized ReaxFF transition state structure.

4 CONCLUSIONS

We presented an automatic workflow which uses molecular dynamic simulations to sample possible elementary reactions for which afterwards refined with quantum chemical methods,

in the current study DFT and coupled cluster, to obtain accurate thermodynamic and kinetic parameters. At the example of combustion reactions for polycyclic aromatic hydrocarbons, the capabilities of the workflow are demonstrated. The implemented automatic filtering technique reduces the number of reactions events that need to be refined at with quantum chemical calculations dramatically by 70 to 90% and allows to restrict the latter to non-redundant reactions, which dominate the combustion process and involve C-C bond breaking. To focus on combustion the simulation was performed at high temperature (2000K) and the combustion parameterization for ReaxFF was employed.

To enhance the rate of observed reaction events we started with radical species and an radical enriched atmosphere. The oxygen and radical concentration in the system influences the outcome of the observed reaction. For low hydroxyl radical concentrations, we observed that carbon monoxide and dioxide are the favoured products. With increasing hydroxyl radical concentration more formic acid-like species are formed. However, in all cases the steps for the decomposition of the carbon structure remained the same. The carbon rings are opened from the exterior side or functional groups are formed via oxygen and/or OH addition and later leave the system. At the enhanced temperatures of the MD simulations, we did not observe that for example epoxide groups form on the aromatic ring system itself. They have only found as short-living reaction intermediates at outer positions. The current simulation suggest that the PAHs decompose starting from the exterior parts. In the future it will be interesting to study if this trend remains if hetero atoms or defects are included in the aromatic ring structure.

Besides these qualitative findings automated DFT calculations were carried out for a subset of the reactions. In this process the transition states and the connecting educts and products are optimized. In general, exploring chemical space via ReaxFF simulations is a fast and useful way to create a reaction pool. However, though more expensive, this study shows that more accurate methods for the MD simulations should also be considered.

The usage of ReaxFF MD simulations for sampling possible reaction events causes some challenges for a subsequent refinement with electronic structure calculations since the PESs for those two methods can differ to a large extent. We observed for example O—H—O bond sequences with bond lengths that are far from equilibrium structures on the more

accurate PESs, or ReaxFF equilibrium structures which either do not exist as minima on the DFT PES or are far away from their DFT counterparts. Furthermore, there have been detected reactions on the ReaxFF level, where the automated refinement could not verify these reactions on the DFT PES. However, in some of those cases the ReaxFF transition states could still be used to locate transition states on the DFT PES for other reactions, which then could be added to the reaction database.

We started with ReaxFF MD as precursor to sample reactions because of its high computational efficiency and broad usage in the literature[59–62] for modelling char combustion. In future work we plan to generalize the implementation of the automated trajectory analyzer and work flow for MD simulations with semi-empirical electronic structure methods like GFN2-xTB. Because these methods already use an electronic wavefunction, they will already provide information about the secondary spin quantum numbers which reduces the amount of information that needs be filled in during the refinement of the reactions at the DFT level. Semi-empirical MD simulations will also potentially diminish the discrepancies between the structures and reactions found in the MD trajectories and those existing on the DFT PESs and will allow to study the influence of heteroatoms for which no suitable parameterization for ReaxFF exists. Furthermore, a future addition to our workflow is could be the incorporation of configurational sampling techniques for the molecular species involved in the reaction.

ACKNOWLEDGMENTS

We thank Kai Leonhard, Felix Schmalz, and Lukas Krep for helpful discussions and other authors for giving access to their code ChemTraYzer. We also thank to the German Research Foundation (DFG) for funding this research within the framework of the SFB/Transregio 129 “Oxyflame” (project number 215035359).

References

1. C. Riplinger, B. Sandhoefer, A. Hansen, and F. Neese, The Journal of Chemical Physics **139**, 134101 (2013).

2. C. Riplinger, P. Pinski, U. Becker, E. F. Valeev, and F. Neese, The Journal of Chemical Physics **144**, 024109 (2016).

3. G. Schmitz and C. Hättig, The Journal of Chemical Physics **145**, 234107 (2016).

4. A. Hansen, C. Bannwarth, S. Grimme, P. Petrović, C. Werlé, and J.-P. Djukic, ChemistryOpen **3**, 177 (2014), ISSN 2191-1363.

5. M. R. Harper, K. M. Van Geem, S. P. Pyl, G. B. Marin, and W. H. Green, Combustion and Flame **158**, 16 (2011), ISSN 0010-2180.

6. J. W. Allen, A. M. Scheer, C. W. Gao, S. S. Merchant, S. S. Vasu, O. Welz, J. D. Savee, D. L. Osborn, C. Lee, S. Vranckx, et al., Combustion and Flame **161**, 711 (2014), ISSN 0010-2180.

7. G. R. Magoon, J. Aguilera-Iparraguirre, W. H. Green, J. J. Lutz, P. Piecuch, H.-W. Wong, and O. O. Oluwole, International Journal of Chemical Kinetics **44**, 179 (2012).

8. G. R. Magoon and W. H. Green, Computers & Chemical Engineering **52**, 35 (2013), ISSN 0098-1354.

9. C. W. Gao, J. W. Allen, W. H. Green, and R. H. West, Computer Physics Communications **203**, 212 (2016), ISSN 0010-4655.

10. G. N. Simm and M. Reiher, Journal of Chemical Theory and Computation **13**, 6108 (2017).

11. S. Habershon, The Journal of Chemical Physics **143**, 094106 (2015).

12. A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard, The Journal of Physical Chemistry A **105**, 9396 (2001).

13. K. Chenoweth, A. C. T. van Duin, and W. A. Goddard, The Journal of Physical Chemistry A **112**, 1040 (2008).

14. M. Döntgen, M.-D. Przybylski-Freund, L. C. Kröger, W. A. Kopp, A. E. Ismail, and K. Leonhard, Journal of Chemical Theory and Computation **11**, 2517 (2015).

15. M. Döntgen, F. Schmalz, W. A. Kopp, L. C. Kröger, and K. Leonhard, Journal of Chemical Information and Modeling **58**, 1343 (2018).

16. L. C. Kröger, W. A. Kopp, M. Döntgen, and K. Leonhard, Journal of Chemical Theory and Computation **13**, 3955 (2017).

17. L.-P. Wang, A. Titov, R. McGibbon, F. Liu, V. S. Pande, and T. J. Martínez, Nature Chemistry **6**, 1044 (2014), ISSN 1755-4349.

18. S. Maeda and K. Morokuma, Journal of Chemical Theory and Computation **7**, 2335 (2011).

19. W. M. C. Sameera, S. Maeda, and K. Morokuma, Accounts of Chemical Research **49**, 763 (2016).

20. C. W. Lee, B. L. H. Taylor, G. P. Petrova, A. Patel, K. Morokuma, K. N. Houk, and B. M. Stoltz, Journal of the American Chemical Society **141**, 6995 (2019).

21. The HDF Group, *Hierarchical data format version 5* (2000-2010), URL `http://www.hdfgroup.org/HDF5`.

22. L. P. Cordella, P. Foggia, C. Sansone, and M. Vento, in *In: 3rd IAPR-TC15 Workshop on Graph-based Representations in Pattern Recognition, Cuen* (2001), pp. 149–159.

23. L. P. Cordella, P. Foggia, C. Sansone, and M. Vento, IEEE Transactions on Pattern Analysis and Machine Intelligence **26**, 1367 (2004).

24. N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, and G. R. Hutchison, Journal of Cheminformatics **3**, 33 (2011), ISSN 1758-2946.

25. S. G. Balasubramani, G. P. Chen, S. Coriani, M. Diedenhofen, M. S. Frank, Y. J. Franzke, F. Furche, R. Grotjahn, M. E. Harding, C. Hättig, et al., JOURNAL OF CHEMICAL PHYSICS **152** (2020).

26. *TURBOMOLE V7.5 2020, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from* `https://www.turbomole.org`.

27. P. Deglmann, F. Furche, and R. Ahlrichs, Chemical Physics Letters **362**, 511 (2002), ISSN 0009-2614.

28. P. Deglmann and F. Furche, The Journal of Chemical Physics **117**, 9535 (2002).

29. T. Helgaker, Chemical Physics Letters **182**, 503 (1991), ISSN 0009-2614.

30. K. Fukui, Acc. Chem. Res. **14**, 363 (1981).

31. T. P. Peixoto, figshare (2014), URL `http://figshare.com/articles/graph_tool/1164194`.

32. P. Plessow, Journal of Chemical Theory and Computation **9**, 1305 (2013).

33. O. Weser, Master's thesis, Georg August Universität Göttingen (2017).

34. B. Helmich and M. Sierka, Journal of Computational Chemistry **33**, 134 (2012).

35. O. Yönder, G. Schmitz, C. Hättig, R. Schmid, P. Debiagi, C. Hasse, A. Locaspi, and T. Faravelli, The Journal of Physical Chemistry A **124**, 9626 (2020).

36. S. Grimme, Chemistry – A European Journal **18**, 9955 (2012).

37. I. Shavitt, The Journal of Chemical Physics **31**, 1359 (1959).

38. S. Plimpton, Journal of Computational Physics **117**, 1 (1995), ISSN 0021-9991.

39. H. M. Aktulga, J. C. Fogarty, S. A. Pandit, and A. Y. Grama, Parallel Comput. **38**, 245–259 (2012), ISSN 0167-8191.

40. R. D. Hipp, *SQLite* (2020), URL `https://www.sqlite.org/index.html`.

41. M. Di Pierro, *Chapter 6: The database abstraction layer.*

42. A. Schäfer, C. Huber, and R. Ahlrichs, The Journal of Chemical Physics **100**, 5829 (1994).

43. A. Schäffer, H. Horn, and R. Ahlrichs, J. Chem. Phys. **97**, 2571 (1992).

44. K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs, Chem. Phys. Lett. **242**, 652 (1995).

45. S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, The Journal of Chemical Physics **132**, 154104 (2010).

46. J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, Phys. Rev. Lett. **91**, 146401 (2003).

47. V. N. Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew, The Journal of Chemical Physics **119**, 12129 (2003).

48. G. I. Csonka, A. Ruzsinszky, and J. P. Perdew, The Journal of Physical Chemistry A **109**, 6779 (2005), pMID: 16834032.

49. C. Hättig, D. P. Tew, and A. Köhn, The Journal of Chemical Physics **132**, 231102 (2010).

50. M. Häser and R. Ahlrichs, Journal of Computational Chemistry **10**, 104 (1989).

51. K. A. Peterson, T. B. Adler, and H.-J. Werner, The Journal of Chemical Physics **128**, 084102 (2008).

52. C. Hättig, Phys. Chem. Chem. Phys. **7**, 59 (2005).

53. F. Weigend, A. Köhn, and C. Hättig, The Journal of Chemical Physics **116**, 3175 (2002).

54. K. E. Yousaf and K. A. Peterson, The Journal of Chemical Physics **129**, 184108 (2008).

55. C. Bannwarth, S. Ehlert, and S. Grimme, Journal of Chemical Theory and Computation **15**, 1652 (2019).

56. P. Pracht, E. Caldeweyher, S. Ehlert, and S. Grimme (2019), URL `doi.org/10.26434/chemrxiv.8326202.v1`.

57. D. E. Edwards, D. Y. Zubarev, W. A. Lester, and M. Frenklach, The Journal of Physical Chemistry A **118**, 8606 (2014).

58. Özlem Yönder and C. Hättig, Fuel **241**, 851 (2019), ISSN 0016-2361.

59. Y. Qiu, W. Zhong, Y. Shao, and A. Yu, Powder Technology **361**, 337 (2020), ISSN 0032-5910.

60. C. Chen and X. Jiang, Phys. Chem. Chem. Phys. **22**, 20829 (2020).

61. F. Castro-Marcano, A. M. Kamat, M. F. Russo, A. C. van Duin, and J. P. Mathews, Combustion and Flame **159**, 1272 (2012).

62. M. Zheng, X. Li, J. Liu, Z. Wang, X. Gong, L. Guo, and W. Song, Energy & Fuels **28**, 522 (2013).

Supporting Information to "An automatized workflow from molecular dynamic simulation to quantum chemical methods to identify elementary reactions and compute reaction constants"

Gunnar Schmitz,[1, a)] Özlem Yönder,[2, b)] Bastian Schnieder,[1] Rochus Schmid,[1] and Christof Hättig[2]

[1)] *Computational Materials Chemistry group, Ruhr-Universität Bochum, D-44780 Bochum, Germany*

[2)] *Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, D-44780 Bochum, Germany*

(Dated: 7 May 2021)

[a)] Corresponding author: gunnar.schmitz@rub.de

[b)] Corresponding author: oezlem.yoender@rub.de

MODIFIED ARRHENIUS EQUATION FIT PARAMETERS

TABLE I: The non-linear least square fit of the calculated rates constants (k) for the studied sites to a modified Arrhenius equation, $k = AT^n \exp\left(-\frac{E}{T}\right)$, where T and E are in units of K.

reaction	lnA	n	E	units of k	reaction	lnA	n	E	units of k
1, forward	28.30	0.392	14985.6	s^{-1}	8, forward	29.80	0.086	1745.1	s^{-1}
1, reverse	29.67	0.070	888.3	s^{-1}	8, reverse	27.99	0.207	32641.7	s^{-1}
2, forward	29.09	0.159	10090.4	s^{-1}	9, forward	27.50	0.420	16973.8	s^{-1}
2, reverse	27.27	0.302	7920.7	s^{-1}	9, reverse	26.46	0.383	10898.0	s^{-1}
3, forward	26.01	0.611	9213.8	s^{-1}	10, forward	30.03	0.159	6006.2	s^{-1}
3, reverse	24.30	0.730	15554.0	s^{-1}	10, reverse	8.74	1.674	44330.0	$cm^3\ mol^{-1}s^{-1}$
4, forward	8.87	1.965	1787.3	$cm^3\ mol^{-1}s^{-1}$	11, forward	30.85	0.013	5285.4	s^{-1}
4, reverse	29.24	0.324	12159.0	s^{-1}	11, reverse	26.50	0.303	4765.6	s^{-1}
5, forward	30.42	0.166	1017.4	s^{-1}	12, forward	30.07	0.082	24575.9	s^{-1}
5, reverse	24.76	0.457	2074.5	s^{-1}	12, reverse	28.56	0.141	249.8	s^{-1}
6, forward	5.02	2.653	18589.9	$cm^3\ mol^{-1}s^{-1}$	13, forward	28.32	0.230	23243.5	s^{-1}
6, reverse	31.00	0.374	12413.3	s^{-1}	13, reverse	24.82	0.573	9740.0	s^{-1}
7, forward	30.81	-0.022	1186.4	s^{-1}	14, forward	26.58	0.303	6124.7	s^{-1}
7, reverse	5.97	1.934	10326.0	$cm^3\ mol^{-1}s^{-1}$	14, reverse	28.30	0.264	9917.6	s^{-1}

CONSTRUCTING Z-MATRIX OF A REACTIVE COMPLEX FOR BIMOLECULAR REACTIONS

This is an example for the construction of the Z-matrix of the product complex, given in Table V, of the formate radical elimination reaction. The Z-matrix of the reactive complex is constructed using the Z-matrix of the product 1 given in Table II, and combining it with the Z-matrix of the product 2 given in Table III. In order to fix the position of the product 2 3 Å away from the product 1, the Z-matrix of the ReaxFF pre-optimized transition state structure given in Table IV is used.

TABLE II: The Z-matrix of product 1 in chemcoord format.

atom index	atom	b	bond	a	angle	d	dihedral
2	C	origin	0.874688	e_z	91.452418	e_x	-152.022256
1	C	2	1.446514	e_z	81.549821	e_x	-53.769336
10	C	2	1.417747	1	122.032205	e_x	179.235964
3	C	2	1.436697	1	118.127669	10	-179.979322
6	C	1	1.401427	2	121.207736	10	179.945733
12	O	1	1.352896	2	116.322769	10	-0.034151
9	C	10	1.395587	2	120.74058	1	-179.970576
15	H	10	1.09304	2	118.804695	1	0.021696
4	C	3	1.491326	2	121.71541	1	-0.002394
7	C	3	1.389128	2	115.76535	1	179.994412
5	C	4	1.471034	3	115.786513	2	0.068066
18	H	6	1.098885	1	120.792635	2	-179.977663
16	H	12	0.976181	1	109.302615	2	179.99772
8	C	7	1.379784	3	126.481693	2	-0.009963
14	H	9	1.0957	10	119.915137	2	-179.996684
19	O	4	1.242963	3	121.798676	2	-179.880362
11	C	5	1.480193	4	119.622825	3	179.956272
13	H	8	1.095655	7	122.439193	3	179.96875
20	O	11	1.227004	5	123.03236	4	179.832039
17	H	11	1.115326	5	114.763792	4	-0.137868

TABLE III: The Z-matrix of product 2 in chemcoord format.

atom index	atom	b	bond	a	angle	d	dihedral
1	O	origin	1.229252	e_z	87.14591	e_x	148.748899
3	C	1	1.260991	e_z	40.968279	e_x	-26.326719
2	O	1	2.107755	3	33.327708	e_x	-2.485743
4	H	1	2.090702	3	26.392835	2	-179.996804

TABLE IV: The Z-matrix of the ReaxFF pre-optimized transition state structure.

atom index	atom	b	bond	a	angle	d	dihedral
2	C	origin	1.113894	e_z	85.358874	e_x	-100.258273
1	C	2	1.437909	e_z	119.322369	e_x	138.666961
10	C	2	1.437902	1	123.41761	e_x	156.218645
3	C	2	1.466253	1	117.196311	10	179.053916
6	C	1	1.397245	2	124.158224	10	-160.188653
24	O	1	1.423223	2	113.246744	10	15.378376
9	C	10	1.394124	2	120.263136	1	-179.663367
15	H	10	1.073874	2	119.673104	1	3.216709
4	C	3	1.526831	2	117.760838	1	-15.709834
7	C	3	1.374774	2	114.796079	1	168.096178
5	C	4	1.484354	3	117.166578	2	-0.05818
19	H	6	1.075975	1	121.23611	2	171.272191
16	H	24	0.981769	1	112.808756	2	110.018302
8	C	7	1.379033	3	127.033629	2	16.663697
14	H	9	1.075189	10	120.422978	2	-174.614094
21	O	4	1.305194	3	125.807478	2	-172.292007
11	C	5	1.468667	4	119.778294	3	-163.235937
13	H	8	1.082761	7	121.629074	3	172.813687
22	O	11	1.305042	5	120.482091	4	178.866012
18	H	11	1.083692	5	118.990579	4	2.538083
20	O	7	1.65943	8	111.136323	13	-5.266817
12	C	20	1.377391	8	90.240594	13	-47.932694
23	O	20	2.353672	12	27.699258	13	35.024443
17	H	20	2.138229	12	26.533316	23	179.371312

TABLE V: The Z-matrix of the product complex.

atom index	atom	b	bond	a	angle	d	dihedral
2	C	origin	0.874688	e_z	91.452418	e_x	-152.022256
1	C	2	1.446514	e_z	81.549821	e_x	-53.769336
10	C	2	1.417747	1	122.032205	e_x	179.235964
3	C	2	1.436697	1	118.127669	10	-179.979322
6	C	1	1.401427	2	121.207736	10	179.945733
24	O	1	1.352896	2	116.322769	10	-0.034151
9	C	10	1.395587	2	120.74058	1	-179.970576
15	H	10	1.093040	2	118.804695	1	0.021696
4	C	3	1.491326	2	121.71541	1	-0.002394
7	C	3	1.389128	2	115.76535	1	179.994412
5	C	4	1.471034	3	115.786513	2	0.068066
19	H	6	1.098885	1	120.792635	2	-179.977663
16	H	24	0.976181	1	109.302615	2	179.99772
8	C	7	1.379784	3	126.481693	2	-0.009963
14	H	9	1.095700	10	119.915137	2	-179.996684
21	O	4	1.242963	3	121.798676	2	-179.880362
11	C	5	1.480193	4	119.622825	3	179.956272
13	H	8	1.095655	7	122.439193	3	179.96875
22	O	11	1.227004	5	123.03236	4	179.832039
18	H	11	1.115326	5	114.763792	4	-0.137868
20	O	7	3.000000	8	111.136323	13	-5.266817
12	C	20	1.260991	8	90.240594	13	-47.932694
23	O	20	2.107755	12	33.327708	13	35.024443
17	H	20	2.090702	12	26.392835	23	-179.996804

THE STATISTICS OF FILTERED REACTIONS

TABLE VI: Statistics of filtered reactions in pure O_2. Given are the found unique reactions $n_{\bar{r}}$ as well as all found reactions n_r. Additionally the number of reactions is given in which the molecular C-graph changed upon reaction $n_{\bar{r}}^{(C)}$ and thus an important step in the combustion process happened. The screening ratio Screen. which represents the skipped reactions is given in %.

	(a) MD 1					(b) MD 2					(c) MD 3				
run	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
n_r	167	1144	490	324	378	597	1675	714	2625	346	1623	660	709	280	638
$n_{\bar{r}}$	50	135	73	103	91	193	170	142	161	130	140	197	182	98	173
Screen.	70	88	85	68	76	68	90	80	94	62	91	70	74	65	73
$n_{\bar{r}}^{(C)}$	19	52	25	32	26	67	44	42	62	31	44	65	66	36	50

(d) MD 4					(e) MD 5				
1	2	3	4	5	1	2	3	4	5
25675	1908	2390	1107	1682	2276	11971	29	2353	633
349	219	299	313	182	164	55	18	202	64
99	89	87	72	89	93	99.5	38	91	90
100	76	114	133	52	72	16	3	72	11

VI. Conclusions

The research reported in the previous chapters aimed to develop a multimethod modeling scheme for atomistic multiscale modeling of char burnout reactions as part of Oxyflame CRC/TRR 129. The model systems were restricted to only C, H, and O atoms. In order to build such a scheme, polyaromatic hydrocarbons (PAHs) have been chosen as models of char fragments. The effect of differences in the size and sites of PAHs has been investigated for several elementary reaction steps. It has been shown through these studies that, based on a set of PAHs which are accessible with accurate quantum chemistry methods, the thermochemical and kinetic properties of larger PAHs can be estimated. This is an important step for building detailed char burnout kinetic models at more realistic scales. Another aspect of this research was to establish a protocol to explore potential elementary reaction steps by means of automated analysis of ReaxFF trajectories. We developed such an automated analysis tool which shares similar ideas as ChemTraYzer, but has been generalized such that it is also applicable to large molecular systems. In benchmark studies, it has been shown that ReaxFF reveals reasonable bond formation/breaking events and that with DFT methods these events can be refined to locate transition structures and the connecting minima.

In order to investigate the effect of the size and sites of PAHs on elementary reaction steps, we have studied in section II, a CO formation pathway with an O-atom, and in section III H-abstraction reaction with an H-atom. Understanding these effects allows to extrapolate the reaction and activation energies to larger systems. For all of the elementary reactions studied in these sections, it has been shown that the Bell-Evans-Polanyi principle can be used to estimate the activation energies based on reaction energies. Therefore, it is useful to also find ways to estimate reaction energies for these elementary reaction steps without the need for optimization of the transition structures.

The CO formation and release pathways studied in section II happen through i) addition of an O(^{3}P), ii) abstraction of H from the oxyradical, and iii) CO elimination. It has been shown that the last step occurs either in three, two, or a single step mechanism depending on the geometrical strain at the reaction site. Although, a quantitative measure of the geometrical strain is difficult to obtain, there was enough evidence that similar reactive sites show similar mechanisms and reaction and activation energies. This implies that a sub-system of a larger polyaromatic which can capture the local environment can be used for estimating reaction mechanisms and energetics. The reaction energies of the elementary reaction steps of O(^{3}P) addition and CO elimination from the carbonyl intermediate correlated well with the stabilization energies calculated from Hückel molecular orbital theory, because the differences in the reaction energies are mainly determined by the rearrangements in the π-electron system. The H-abstraction reaction via an H-atom from PAHs studied in section III is a local reaction which does not cause a rearrangement in the π-electron system. For this reaction, we have observed that the differences in the activation and reaction energies are determined by the changes in the local environment of the reaction site upon H-removal (C-H bond breaking energy) because the relaxation energies of the H-abstracted structure is almost constant among different reaction sites. This suggests that using a fixed relaxation energy for all of the sites and only calculating the C-H bond breaking energies would approximate the reaction energies. Additionally, for this reaction, we have observed that the contribution of the triples correction (T) to the reaction energy using the ROHF-CCSD(T) method and the difference between the DFT and CCSD reaction energies are almost constant. Therefore, the reaction energies close to CCSD(T) accuracy could be achieved using a fixed CCSD and (T) correction to the estimated DFT energies.

In the future, as an alternative to the rules derived for the thermochemical parameters, i.e., the use of Bell-Evans-Polanyi principle, correlations between stabilization energies and reaction energies, etc., machine learning methods could also be used to achieve

the same outcome. This would also allow to decrease the manual workload needed to derive these rules.

Another aspect of the study in Section III was to link the ReaxFF, DFTB, and DFT methods. It has been observed that the ReaxFF geometries sometimes are not reliable. On the other hand, GFN1-xTB provides transition state and equilibrium geometries close to the DFT geometries, while GFN2-xTB provides more accurate energies which are closer to the coupled cluster energies than GFN1-xTB does. Additionally, compared with DFT TPSSh-D3/TZVP, among the zero point vibrational energies calculated with the xTB methods, GFN1-xTB deviates 0.4 times whereas GFN2-xTB deviates by 1.6 times. These also indicate that GFN1-xTB may also be suitable for estimation of thermochemical properties at a lower computational cost than DFT. The general ability of xTB methods to predict acceptable zero-point-vibrational energies and vibrational partition functions can be investigated in further studies.

Another remark which came out of the study of H-abstraction is that even though CCSD method is known to eliminate error in energies originating from spin contamination of UHF reference[109], we have observed that this might not be the case for the reaction intermediates of polyaromatics. Therefore, for the calculation of accurate energies with CCSD(T) of these systems, we recommend using ROHF reference determinant, especially for an automated scheme.

As ReaxFF is more efficient than xTB methods for the exploration of combustion reactions, ReaxFF remains the method of choice. As explained in section IV, the automated ReaxFF trajectory analyzer ChemTraYzer has been modified and its applicability for the analysis of phenanthrene combustion has been tested. The first modification was that the generation of structures and the assignment of the α- and β-spin occupations by OpenBabel using simplified molecular-input line-entry system (SMILES) format has been abandoned. Additionally, the redundant calculations of the reacting species are avoided through structural comparisons. This work showed that the approach of ChemTraYzer to generate start structures for transition states,

i.e., through preliminary constrained geometry optimization with fixed coordinates for the atoms which are involved in bond breaking/formation in the ReaxFF simulation, is especially effective for reactions with barriers higher than $\approx 30\,\mathrm{kJ/mol}$. We have concluded that for those cases where the ReaxFF pre-optimized transition structures were not close enough to the saddle point, a double-ended reaction path calculation could help generating better start structures. For this purpose, processing the trajectory information at once instead of in batches becomes advantageous, because it adds the flexibility to go back and forth through the trajectory without losing information between the batches. Additionally, a relational database structure makes the abstraction of data easier. Therefore, as described in section V, we wrote a new program which is more suitable for studying combustion of larger model systems but still uses some of the ideas from ChemTraYzer.

The new program described in section V analyses the trajectory to search for bond breaking/formation events by means of graph theory. For this task, I contributed to the automation of the refinement of these events by DFT and coupled cluster methods. The potential educt(s) and product(s) are chosen relative to (+1,-1) from the frame where bond order of the breaking/forming bond is 0.5. The non-redundant reactions can be filtered out. As a subsequent step, these reactions can be refined using density functional theory. Whenever a ReaxFF pre-optimized transition structure does not provide a good start guess, the corresponding reaction path is automatically optimized with a double ended method in order to generate another start guess for the transition state. We have shown that this new approach can refine the detected potential reaction events for large polyaromatics. Another future potential investigation that can be made is related to the unimolecular bond breaking/formation events. For some unimolecular reactions upon geometry optimization with ReaxFF, the potential educt and product structures end up in the same minima. These events seem like ring breathings at the ReaxFF trajectory due to high temperatures. However, this might need further investigation because at the DFT PES, corresponding reactions

have sometimes been located in the automated refining scheme. Finally, to better understand the type of reactions that ReaxFF might not be able to capture, a future research might include comparison of the types of reactions observed in xTB and in ReaxFF simulations on a small polyaromatic.

A potential roadmap for the future steps of building a detailed char kinetic model within this project could include: 1) Generation and analysis of elementary reaction steps using the methodology explained in section V, 2) building a connected reaction network, 3) test validity of the kinetic model by means of experiments. The last step can be done by implementing the developed detailed mechanisms into the CRECK modeling scheme. When this proof of concept is achieved, the effects of heteroatoms, nitrogen and sulfur release, etc. should also be studied and integrated in the detailed kinetic models of char burnout.

References

[1]S. W. Running. Is Global Warming Causing More, Larger Wildfires? Science. **313**(5789), 927 (2006).

[2]J. Braun McNeill, L. Avens, A. Goodman Hall, L. R. Goshe, C. A. Harms and D. W. Owens. Female-Bias in a Long-Term Study of a Species with Temperature-Dependent Sex Determination: Monitoring Sex Ratios for Climate Change Research. PLOS ONE. **11**(8), 1 (2016).

[3]B. Geffroy and C. Wedekind. Effects of global warming on sex ratios in fishes. Journal of Fish Biology. **97**(3), 596 (2020).

[4]S. Wang, Y. Zhang, W. Ju, J. M. Chen, P. Ciais, A. Cescatti, J. Sardans, I. A. Janssens, M. Wu, J. A. Berry, E. Campbell, M. Fernández-Martínez, R. Alkama, S. Sitch, P. Friedlingstein, W. K. Smith, W. Yuan, W. He, D. Lombardozzi, M. Kautz, D. Zhu, S. Lienert, E. Kato, B. Poulter, T. G. M. Sanders, I. Krüger, R. Wang, N. Zeng, H. Tian, N. Vuichard, A. K. Jain, A. Wiltshire, V. Haverd, D. S. Goll, and J. Peñuelas. Recent global decline of CO_2 fertilization effects on vegetation photosynthesis. Science. **370**(6522), 1295 (2020).

[5]Climate Analysis Indicators Tool. Historical CO_2 emissions by sector, 2020. Data retrieved from data source of Climate Watch, `https://www.climatewatchdata.org/ghg-emissions`.

[6]M. Allen, O. Dube, W. Solecki, F. Aragón-Durand, W. Cramer, S. Humphreys, M. Kainuma, J. Kala, N. Mahowald, Y. Mulugetta, R. Perez, M. Wairiu, and K. Zickfeld. Global warming of 1.5° C. An IPCC special report on the impacts of global warming of 1.5° C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty, 2018.

[7]M. Fajardy and N. Mac Dowell. The energy return on investment of BECCS: is

BECCS a threat to energy security? Energy Environ. Sci. **11**, 1581 (2018).

[8]A. Ernsting and O. Munnion. Last ditch climate option or wishful thinking? Bioenergy with carbon capture and storage, 2016.

[9]M. Fridahl and M. Lehtveer. Bioenergy with carbon capture and storage (BECCS): Global potential, investment preferences, and deployment barriers. Energy Research Social Science. **42**, 155 (2018).

[10]M. Bui, C. S. Adjiman, A. Bardow, E. J. Anthony, A. Boston, S. Brown, P. S. Fennell, S. Fuss, A. Galindo, L. A. Hackett, J. P. Hallett, H. J. Herzog, G. Jackson, J. Kemper, S. Krevor, G. C. Maitland, M. Matuszewski, I. S. Metcalfe, C. Petit, G. Puxty, J. Reimer, D. M. Reiner, E. S. Rubin, S. A. Scott, N. Shah, B. Smit, J. P. M. Trusler, P. Webley, J. Wilcox, and N. Mac Dowell. Carbon capture and storage (CCS): the way forward. Energy Environ. Sci. **11**, 1062 (2018).

[11]O. Vidal, B. Goffé and N. Arndt. Metals for a low-carbon society. Nature Geoscience. **6**(11), 894 (2013).

[12]M. Von Cossel, M. Wagner, J. Lask, E. Magenau, A. Bauerle, V. Von Cossel, K. Warrach-Sagi, B. Elbersen, I. Staritsky, M. Van Eupen, Y. Iqbal, N. D. Jablonowski, S. Happe, A. L. Fernando, D. Scordia, S. L. Cosentino, V. Wulfmeyer, I. Lewandowski, and B. Winkler. Prospects of Bioenergy Cropping Systems for A More Social-Ecologically Sound Bioeconomy. Agronomy. **9**, 605 (2019).

[13]A. Sharma, V. Pareek and D. Zhang. Biomass pyrolysis—A review of modelling, process parameters and catalytic studies. Renewable and Sustainable Energy Reviews. **50**, 1081 (2015).

[14]U. Kleinhans, S. Halama and H. Spliethoff. The role of gasification reactions during pulverized solid fuel combustion: A detailed char combustion model based on measurements of char structure and kinetics for coal and pre-treated biomass. Combustion and Flame. **184**, 117 (2017).

[15]M. Keiluweit, P. S. Nico, M. G. Johnson and M. Kleber. Dynamic Molecular Structure of Plant Biomass-Derived Black Carbon (Biochar). Environmental Sci-

ence & Technology. **44**(4), 1247 (2010). PMID: 20099810.

[16]J. M. Simmie. Detailed chemical kinetic models for the combustion of hydrocarbon fuels. Progress in Energy and Combustion Science. **29**(6), 599 (2003).

[17]C. A. Taatjes, D. L. Osborn, T. M. Selby, G. Meloni, A. J. Trevitt, E. Epifanovsky, A. I. Krylov, B. Sirjean, E. Dames, and H. Wang. Products of the Benzene + O(3P) Reaction. The Journal of Physical Chemistry A. **114**(9), 3355 (2010). PMID: 20099861.

[18]A. Shiroudi, M. S. Deleuze and S. Canneaux. Theoretical study of the oxidation mechanisms of naphthalene initiated by hydroxyl radicals: the O2 addition reaction pathways. Phys. Chem. Chem. Phys. **17**, 13719 (2015).

[19]Z. Zhang, L. Lin and L. Wang. Atmospheric oxidation mechanism of naphthalene initiated by OH radical. A theoretical study. Phys. Chem. Chem. Phys. **14**, 2645 (2012).

[20]D. E. Edwards, D. Y. Zubarev, W. A. Lester and M. Frenklach. Pathways to Soot Oxidation: Reaction of OH with Phenanthrene Radicals. The Journal of Physical Chemistry A. **118**(37), 8606 (2014). PMID: 24761798.

[21]N. Zhao, Q. Zhang and W. Wang. Atmospheric oxidation of phenanthrene initiated by OH radicals in the presence of O2 and NOx — A theoretical study. Science of The Total Environment. **563-564**, 1008 (2016).

[22]A. Raj, S. Y. Yang, D. Cha, R. Tayouo and S. H. Chung. Structural effects on the oxidation of soot particles by O2: Experimental and theoretical study. Combustion and Flame. **160**(9), 1812 (2013).

[23]J. Dang, X. Shi, J. Hu, J. Chen, Q. Zhang and W. Wang. Mechanistic and kinetic studies on OH-initiated atmospheric oxidation degradation of benzo[α]pyrene in the presence of O2 and NOx. Chemosphere. **119**, 387 (2015).

[24]R. I. Singh, A. M. Mebel and M. Frenklach. Oxidation of Graphene-Edge Six- and Five-Member Rings by Molecular Oxygen. The Journal of Physical Chemistry A. **119**(28), 7528 (2015). PMID: 25894330.

[25]S. Grimme. Exploration of Chemical Compound, Conformer, and Reaction Space with Meta-Dynamics Simulations Based on Tight-Binding Quantum Chemical Calculations. Journal of Chemical Theory and Computation. **15**(5), 2847 (2019). PMID: 30943025.

[26]A. C. T. van Duin, S. Dasgupta, F. Lorant and W. A. Goddard. ReaxFF: A Reactive Force Field for Hydrocarbons. The Journal of Physical Chemistry A. **105**(41), 9396 (2001).

[27]F. Castro-Marcano, A. M. Kamat, M. F. Russo, A. C. van Duin and J. P. Mathews. Combustion of an Illinois No. 6 coal char simulated using an atomistic char representation and the ReaxFF reactive force field. Combustion and Flame. **159**(3), 1272 (2012).

[28]M. Döntgen, M.-D. Przybylski-Freund, L. C. Kröger, W. A. Kopp, A. E. Ismail and K. Leonhard. Automated Discovery of Reaction Pathways, Rate Constants, and Transition States Using Reactive Molecular Dynamics Simulations. Journal of Chemical Theory and Computation. **11**(6), 2517 (2015). PMID: 26575551.

[29]M. Döntgen, F. Schmalz, W. A. Kopp, L. C. Kröger and K. Leonhard. Automated Chemical Kinetic Modeling via Hybrid Reactive Molecular Dynamics and Quantum Chemistry Simulations. Journal of Chemical Information and Modeling. **58**(7), 1343 (2018). PMID: 29898359.

[30]J. A. Miller, R. J. Kee and C. K. Westbrook. Chemical Kinetics and Combustion Modeling. Annual Review of Physical Chemistry. **41**(1), 345 (1990).

[31]R. J. Kee, J. A. Miller and T. H. Jefferson. CHEMKIN: a general-purpose, problem-independent, transportable, FORTRAN chemical kinetics code package. Technical report, United States, 1980. SAND–80-8003.

[32]M. Frenklach, H. Wang, C.-L. Yu, M. Goldenberg, C. Bowman, R. Hanson, D. Davidson, E. Chang, G. Smith, D. Golden, W. Gardiner, and V. Lissianski.

[33]J. E. Dec. A conceptual model of di diesel combustion based on laser-sheet imaging*. in International Congress Exposition. SAE International, feb 1997.

[34]C. J. Mueller, W. J. Cannella, T. J. Bruno, B. Bunting, H. D. Dettman, J. A. Franz, M. L. Huber, M. Natarajan, W. J. Pitz, M. A. Ratcliff, and K. Wright. Methodology for Formulating Diesel Surrogate Fuels with Accurate Compositional, Ignition-Quality, and Volatility Characteristics. Energy & Fuels. **26**(6), 3284 (2012).

[35]J. Luo, M. Yao, H. Liu and B. Yang. Experimental and numerical study on suitable diesel fuel surrogates in low temperature combustion conditions. Fuel. **97**, 621 (2012).

[36]D. G. Goodwin, R. L. Speth, H. K. Moffat and B. W. Weber. Cantera: An object-oriented software toolkit for chemical kinetics, thermodynamics, and transport processes. `https://www.cantera.org`, 2021. Version 2.5.1.

[37]H. Pitsch and M. Bollig. Flamemaster: A computer code for homogeneous combustion and one-dimensional laminar flame calculations. `https://www.itv.rwth-aachen.de/downloads/flamemaster/`, 2021.

[38]A. Cuoci, A. Frassoldati, T. Faravelli and E. Ranzi. OpenSMOKE++: An object-oriented framework for the numerical modeling of reactive systems with detailed kinetic mechanisms. Computer Physics Communications. **192**, 237 (2015).

[39]T. Gronarz, M. Schnell, C. Siewert, L. Schneiders, W. Schröder and R. Kneer. Comparison of scattering behaviour for spherical and non-spherical particles in pulverized coal combustion. International Journal of Thermal Sciences. **111**, 116 (2017).

[40]S. V. Vassilev, D. Baxter, L. K. Andersen and C. G. Vassileva. An overview of the chemical composition of biomass. Fuel. **89**(5), 913 (2010).

[41]D. M. Grant, R. J. Pugmire, T. H. Fletcher and A. R. Kerstein. Chemical model of coal devolatilization using percolation lattice statistics. Energy & Fuels. **3**(2), 175 (1989).

[42]R. E. Mitchell, R. H. Hurt, L. L. Baxter and D. R. Hardesty. Compilation of Sandia coal char combustion data and kinetic analyses. Milestone report. (1992).

[43]R. Hurt, J.-K. Sun and M. Lunden. A Kinetic Model of Carbon Burnout in

Pulverized Coal Combustion. Combustion and Flame. **113**(1), 181 (1998).

[44]R. H. Hurt and J. M. Calo. Semi-global intrinsic kinetics for char combustion modeling. Combustion and Flame. **125**(3), 1138 (2001).

[45]G.-S. Liu and S. Niksa. Coal conversion submodels for design applications at elevated pressures. Part II. Char gasification. Progress in Energy and Combustion Science. **30**(6), 679 (2004).

[46]R. C. Shurtz and T. H. Fletcher. Coal Char-CO_2 Gasification Measurements and Modeling in a Pressurized Flat-Flame Burner. Energy & Fuels. **27**(6), 3022 (2013).

[47]T. Holland and T. H. Fletcher. Comprehensive Model of Single Particle Pulverized Coal Combustion Extended to Oxy-Coal Conditions. Energy & Fuels. **31**(3), 2722 (2017).

[48]T. Holland and T. H. Fletcher. Global Sensitivity Analysis for a Comprehensive Char Conversion Model in Oxy-fuel Conditions. Energy & Fuels. **30**(11), 9339 (2016).

[49]D. Genetti, T. H. Fletcher and R. J. Pugmire. Development and Application of a Correlation of ^{13}C NMR Chemical Structural Analyses of Coal Based on Elemental Composition and Volatile Matter Content. Energy & Fuels. **13**(1), 60 (1999).

[50]S. Sommariva, T. Maffei, G. Migliavacca, T. Faravelli and E. Ranzi. A predictive multi-step kinetic model of coal devolatilization. Fuel. **89**(2), 318 (2010).

[51]P. E. A. Debiagi, C. Pecchi, G. Gentile, A. Frassoldati, A. Cuoci, T. Faravelli and E. Ranzi. Extractives Extend the Applicability of Multistep Kinetic Scheme of Biomass Pyrolysis. Energy & Fuels. **29**(10), 6544 (2015).

[52]P. E. A. Debiagi, M. Trinchera, A. Frassoldati, T. Faravelli, R. Vinu and E. Ranzi. Algae characterization and multistep pyrolysis mechanism. Journal of Analytical and Applied Pyrolysis. **128**, 423 (2017).

[53]A. Stagni, A. Cuoci, A. Frassoldati, T. Faravelli and E. Ranzi. Lumping and Reduction of Detailed Kinetic Schemes: an Effective Coupling. Industrial & Engineering Chemistry Research. **53**(22), 9004 (2014).

[54]E. Ranzi, A. Frassoldati, A. Stagni, M. Pelucchi, A. Cuoci and T. Faravelli. Reduced Kinetic Schemes of Complex Reaction Systems: Fossil and Biomass-Derived Transportation Fuels. International Journal of Chemical Kinetics. **46**(9), 512 (2014).

[55]P. Debiagi, H. Nicolai, W. Han, J. Janicka and C. Hasse. Machine learning for predictive coal combustion CFD simulations—From detailed kinetics to HDMR Reduced-Order models. Fuel. **274**, 117720 (2020).

[56]CRC/TRR-129: Research program.

[57]J. F. Stanton. Why CCSD(T) works: a different perspective. Chemical Physics Letters. **281**(1), 130 (1997).

[58]G. Schmitz and C. Hättig. Perturbative triples correction for local pair natural orbital based explicitly correlated CCSD(F12*) using Laplace transformation techniques. The Journal of Chemical Physics. **145**(23), 234107 (2016).

[59]M. Born and R. Oppenheimer. Zur Quantentheorie der Molekeln. Annalen der Physik. **389**(20), 457 (1927).

[60]P.-O. Löwdin. Quantum Theory of Many-Particle Systems. III. Extension of the Hartree-Fock Scheme to Include Degenerate Systems and Correlation Effects. Phys. Rev. **97**, 1509 (1955).

[61]J. A. Pople, M. Head-Gordon and K. Raghavachari. Quadratic configuration interaction. A general technique for determining electron correlation energies. The Journal of Chemical Physics. **87**(10), 5968 (1987).

[62]J. Čížek. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. The Journal of Chemical Physics. **45**(11), 4256 (1966).

[63]O. Sinanoğlu. Many-Electron Theory of Atoms and Molecules. I. Shells, Electron Pairs vs Many-Electron Correlations. The Journal of Chemical Physics. **36**(3), 706 (1962).

[64]C. J. Stein and M. Reiher. autoCAS: A Program for Fully Automated Multiconfigurational Calculations. Journal of Computational Chemistry. **40**(25),

2216 (2019).

[65]F. Neese, F. Wennmohs and A. Hansen. Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method. The Journal of Chemical Physics. **130**(11), 114108 (2009).

[66]W. Li, Z. Ni and S. Li. Cluster-in-molecule local correlation method for post-Hartree–Fock calculations of large systems. Molecular Physics. **114**(9), 1447 (2016).

[67]J. J. Eriksen, P. Baudin, P. Ettenhuber, K. Kristensen, T. Kjærgaard and P. Jørgensen. Linear-Scaling Coupled Cluster with Perturbative Triple Excitations: The Divide–Expand–Consolidate CCSD(T) Model. Journal of Chemical Theory and Computation. **11**(7), 2984 (2015). PMID: 26575735.

[68]J. P. Perdew and K. Schmidt. Jacob's ladder of density functional approximations for the exchange-correlation energy. AIP Conference Proceedings. **577**(1), 1 (2001).

[69]N. Mardirossian and M. Head-Gordon. Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals. Molecular Physics. **115**(19), 2315 (2017).

[70]P. Hohenberg and W. Kohn. Inhomogeneous Electron Gas. Phys. Rev. **136**, B864 (1964).

[71]E. Hückel. Quantentheoretische Beiträge zum Benzolproblem. Zeitschrift für Physik. **70**(3), 204 (1931).

[72]R. Hoffmann. An Extended Hückel Theory. I. Hydrocarbons. The Journal of Chemical Physics. **39**(6), 1397 (1963).

[73]F. Jensen. Introduction to Computational Chemistry. Wiley, 2006.

[74]J. A. Pople and D. L. Beveridge. Approximate Molecular Orbital Theory. McGraw-Hill Book Company, 1970.

[75]D. Porezag, T. Frauenheim, T. Köhler, G. Seifert and R. Kaschner. Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon. Phys. Rev. B. **51**, 12947 (1995).

[76]S. Grimme, C. Bannwarth and P. Shushkov. A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements (Z = 1–86). Journal of Chemical Theory and Computation. **13**(5), 1989 (2017). PMID: 28418654.

[77]C. Bannwarth, S. Ehlert and S. Grimme. GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. Journal of Chemical Theory and Computation. **15**(3), 1652 (2019). PMID: 30741547.

[78]G. A. Worth and L. S. Cederbaum. Beyond Born-Oppenheimer: Molecular Dynamics Through a Conical Intersection. Annual Review of Physical Chemistry. **55**(1), 127 (2004). PMID: 15117250.

[79]D. Zhao, L. Li, D. He and J. Zhou. Molecular dynamics simulations of conformation changes of HIV-1 regulatory protein on graphene. Applied Surface Science. **377**, 324 (2016).

[80]H. M. Senn and W. Thiel. QM/MM studies of enzymes. Current Opinion in Chemical Biology. **11**(2), 182 (2007). Bioinorganic chemistry / Biocatalysis and biotransformation.

[81]L.-P. Wang, A. Titov, R. McGibbon, F. Liu, V. S. Pande and T. J. Martínez. Discovering chemistry with an ab initio nanoreactor. Nature Chemistry. **6**(12), 1044 (2014).

[82]B. Saha, S. Shindo, S. Irle and K. Morokuma. Quantum Chemical Molecular Dynamics Simulations of Dynamic Fullerene Self-Assembly in Benzene Combustion. ACS Nano. **3**(8), 2241 (2009). PMID: 19702322.

[83]P. M. Morse. Diatomic Molecules According to the Wave Mechanics. II. Vibrational Levels. Phys. Rev. **34**, 57 (1929).

[84]D. W. Brenner. Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films. Phys. Rev. B. **42**, 9458 (1990).

[85]A. M. Kamat, A. C. T. van Duin and A. Yakovlev. Molecular Dynamics Simulations of Laser-Induced Incandescence of Soot Using an Extended ReaxFF Reactive Force Field. The Journal of Physical Chemistry A. **114**(48), 12561 (2010). PMID: 21067165.

[86]K. Chenoweth, A. C. T. van Duin and W. A. Goddard. ReaxFF Reactive Force Field for Molecular Dynamics Simulations of Hydrocarbon Oxidation. The Journal of Physical Chemistry A. **112**(5), 1040 (2008). PMID: 18197648.

[87]D. A. McQuarrie. Statistical Mechanics. Harper Row, 1976.

[88]P. Y. Ayala and H. B. Schlegel. Identification and treatment of internal rotation in normal mode vibrational analysis. The Journal of Chemical Physics. **108**(6), 2314 (1998).

[89]J. Pfaendtner, X. Yu and L. J. Broadbelt. The 1-D hindered rotor approximation. Theoretical Chemistry Accounts. **118**(5), 881 (2007).

[90]V. Van Speybroeck, D. Van Neck, M. Waroquier, S. Wauters, M. Saeys and G. B. Marin. Ab Initio Study of Radical Addition Reactions: Addition of a Primary Ethylbenzene Radical to Ethene (I). The Journal of Physical Chemistry A. **104**(46), 10939 (2000).

[91]V. V. Speybroeck, P. Vansteenkiste, D. V. Neck and M. Waroquier. Why does the uncoupled hindered rotor model work well for the thermodynamics of n-alkanes? Chemical Physics Letters. **402**(4), 479 (2005).

[92]S. Grimme. Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory. Chemistry – A European Journal. **18**(32), 9955 (2012).

[93]H. Eyring. The Activated Complex and the Absolute Rate of Chemical Reactions. Chemical Reviews. **17**(1), 65 (1935).

[94]M. G. Evans and M. Polanyi. Some applications of the transition state method to the calculation of reaction velocities, especially in solution. Trans. Faraday Soc. **31**, 875 (1935).

[95]H. Eyring. The Activated Complex in Chemical Reactions. The Journal of Chemical

Physics. **3**(2), 107 (1935).

[96]J. I. Steinfeld, J. S. Francisco and W. L. Hase. Chemical Kinetics and Dynamics. Prentice Hall, 1989.

[97]U. Manthe. Accurate calculations of reaction rates: predictive theory based on a rigorous quantum transition state concept. Molecular Physics. **109**(11), 1415 (2011).

[98]G. Czakó, B. C. Shepler, B. J. Braams and J. M. Bowman. Accurate ab initio potential energy surface, dynamics, and thermochemistry of the $F+CH_4 \longrightarrow HF+CH_3$ reaction. The Journal of Chemical Physics. **130**(8), 084301 (2009).

[99]C. Eckart. The Penetration of a Potential Barrier by Electrons. Phys. Rev. **35**, 1303 (1930).

[100]E. Wigner. On the Quantum Correction For Thermodynamic Equilibrium. Phys. Rev. **40**, 749 (1932).

[101]I. Shavitt. A Calculation of the Rates of the Ortho-Para Conversions and Isotope Exchanges in Hydrogen. The Journal of Chemical Physics. **31**(5), 1359 (1959).

[102]J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari and L. A. Curtiss. Gaussian-1 theory: A general procedure for prediction of molecular energies. The Journal of Chemical Physics. **90**(10), 5622 (1989).

[103]L. A. Curtiss, P. C. Redfern and K. Raghavachari. Gn theory. WIREs Computational Molecular Science. **1**(5), 810 (2011).

[104]A. Burcat, B. Ruscic, Chemistry and T. I. I. of Tech. Third millenium ideal gas and condensed phase thermochemical database for combustion (with update from active thermochemical tables). (2005).

[105]J. W. Ochterski, G. A. Petersson and J. A. Montgomery. A complete basis set model chemistry. V. Extensions to six or more heavy atoms. The Journal of Chemical Physics. **104**(7), 2598 (1996).

[106]J. A. Montgomery, M. J. Frisch, J. W. Ochterski and G. A. Petersson. A complete

basis set model chemistry. VII. Use of the minimum population localization method. The Journal of Chemical Physics. **112**(15), 6532 (2000).

[107] J. M. L. Martin and G. de Oliveira. Towards standard methods for benchmark quality ab initio thermochemistry—W1 and W2 theory. The Journal of Chemical Physics. **111**(5), 1843 (1999).

[108] A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez and J. F. Stanton. HEAT: High accuracy extrapolated ab initio thermochemistry. The Journal of Chemical Physics. **121**(23), 11599 (2004).

[109] J. F. Stanton. On the extent of spin contamination in open-shell coupled-cluster wave functions. The Journal of Chemical Physics. **101**(1), 371 (1994).