

Prediction of transformation products during ozonation of micropollutant-containing waters: Kinetics and mechanisms

THÈSE N° 7177 (2016)

PRÉSENTÉE LE 16 SEPTEMBRE 2016

À LA FACULTÉ DE L'ENVIRONNEMENT NATUREL, ARCHITECTURAL ET CONSTRUIT
LABORATOIRE POUR LE TRAITEMENT ET LA QUALITÉ DE L'EAU
PROGRAMME DOCTORAL EN GÉNIE CIVIL ET ENVIRONNEMENT

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Minju LEE

acceptée sur proposition du jury:

Prof. A. Meibom, président du jury
Prof. U. von Gunten, Dr K. Fenner, directeurs de thèse
Prof. Y. Lee, rapporteur
Dr H. Lutze, rapporteur
Prof. S. Takahama, rapporteur



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2016

Acknowledgements

Thanks to ...

- ... Urs von Gunten for offering me this privileged opportunity to do a PhD in LTQE at EPFL. I thoroughly enjoyed my project and learned a lot under your guidance. Through all the years, Urs has remained responsive, supportive, patient, friendly, and humorous, for which I am deeply grateful. He is the epitome of an excellent PhD supervisor, and I could not have ever hoped to meet and know a more respectable person. Thank you Urs.
- ... Kathrin Fenner for her excellent co-supervision. I was privileged to have her as my cosupervisor. She was supportive and willing to share her insights with me. I could not have completed my PhD without her support. Thank you Kathrin.
- ... the president of the jury, Meibom Anders, and the examiners, Yunho Lee, Takahama Satoshi, and Holger Lutze, for accepting their role as jury members for my private defense, sparing their precious time especially in the summer vacation season, and for their insights and feedback on my thesis.
- ... Lorenz Blum and Emanuel Schmid who made it possible to develop a prediction platform. It was great to have you for our project, Lorenz, and I really appreciate your contributions and our intercity meetings between Lausanne and Zurich. Thank you so much Emanuel for making a smooth transition from Lorenz and helping me to finish my thesis.
- ... J. Samuel Arey, Silvio Canonica, Jakov Bolotin, Tony Merle, Daniel Rentsch, and Daniela Trologo for their invaluable input which greatly improved the quality of my PhD research.
- ... Swiss Federal Office for Environment and the “Regional Water Supply Beselland 21” project for financing this PhD project.
- ... former and present group members and guests in Urs von Gunten’s group at EPFL and Eawag. I have met many good people who helped me with so many things and kindly shared their moments with me. Thanks to you all, my PhD life was much more meaningful than just being isolated in research. In no particular order: Michèle Heeb, Florian Breider, Eva Rodriguez Franco, Yang Song, Margoux Voumard, Glen De Vera, Saskia Zimmermann-Steffens, Catherine Hoffman, Ina Kristina, Caroline Gachet, Alina Tominiak, Sébastien Allard, Jaedon Shin, Ioannis Katsoyiannis, Manuel Sanchez Polo, Fabian Soltermann, Matthias Rudolf von Rohr, Sabrina Bahn Müller, Paul Borer, Jan-nis Wenk, Hana Mestankova, Justine Criquet, Melissa Huguet, Clara Loi, Kangmin Chon, Frank Leresche, Tony Merle, Marc Bourgin, Boris Droz, Sung Eun Lim, Zhengqian Liu, Xiodan Zhao, Shaogan Liu, Peter Tentscher, Stephanie Sphar, Linda Oennby, Sebastian Stoll, Daniel Stalter, Elisabeth Salhi, Hans-Ueli Laubscher, Ursula Schoenenberger, and Claire Wedema.
- ... former and present LCE group members with whom we had a joint group seminar to encourage each other’s research and share ideas and also had occasional non-academic events like beer-tasting: Tamar Kohn, Qingxia Zhong, Simon Meister, Heather Bischel, Camille Wolf, Anna Carratalà Ripollès, Michael Mattle, Loïc Decrey, Therese Sigstam, and Florence Bonvin.

... friends and colleagues thanks to them I came to learn how to enjoy swiss life (e.g., beer, food, hike, bike, ski, pingpong, etc.) outside the office and the laboratory and have good moments together during my PhD in Switzerland: Lorenzo Gorla, Pierre Queloz, Flavio Finger, Antoine Petrelli, Sylvain Coutu, Benoît Cruzy, Kathrina Edmaier, Ana Clara Santos, Amin Niayifar, Mohsen Cheraghi, Abolfazl Rahaghi, Tristan Brauchli, Holly Jayne Oldroyd, Imants Kreituss, Heui-sook Weman, and Sang Hoon Chin.

... my wife Hyeryung Kim for her never-ending love and support, my parents Youngro Lee and Cheongsu Seo for their unconditional respect on all the decisions I have made, my brother and sister Seok Lee and Songmi Lee for cherishing their youngest brother, and my parents-in-law, Mun ho Kim and Kyungok Oh for their warm care for me and my wife.

Lausanne, 24 August, 2016

Minju Lee

Abstract

Ozonation, which is widely used for drinking water disinfection, has recently been applied to mitigate potentially harmful effects of micropollutants (e.g., pharmaceuticals, personal care products, pesticides, etc.) present in municipal wastewater effluents. Generally, ozonation is efficient for the abatement of biological effects caused by micropollutants. However, limited empirical information is available about the transformation products formed during ozonation of micropollutants due to analytical limitations and a large number of micropollutants present in wastewater effluents. In this thesis, a computer-based prediction platform for kinetics and mechanisms for the reactions of ozone with micropollutants was developed to provide information about (i) the reactivity of micropollutants with ozone expressed as second-order rate constants (k_{O_3} , $M^{-1}s^{-1}$) and (ii) potential transformation products formed from the reactions of ozone with micropollutants. Regarding (i), k_{O_3} for micropollutants were predictable using linear relationships between experimental k_{O_3} in log units for compounds of certain chemical groups (e.g., phenols, olefins, amines, etc.) and the corresponding molecular orbital energies (e.g., highest occupied molecular orbital (HOMO) or natural bond orbital (NBO)) obtained from quantum chemical computations (mostly $R^2 = 0.75 - 0.95$ for 14 compound groups consisting of 284 model compounds in total). Overall, the developed k_{O_3} prediction models could predict k_{O_3} on average within a factor of ~ 5 of an experimental k_{O_3} for model compounds used for the development of the k_{O_3} prediction models as well as tetrachlorobutadienes, which were externally validated. In contrast, poor k_{O_3} predictions (>10 fold) were observed for some model compounds excluded from the correlations as outliers as well as cetirizine, two pentachlorobutadiene congeners, and hexachlorobutadiene, which were used for external validation. (ii) A prediction tool for potential transformation products was developed based on numerous reaction pathways proposed in literature, which were encoded into 340 individual reaction rules using appropriate cheminformatics tools. The predicted pathways and the transformation products for some micropollutants (i.e., carbamazepine and tramadol) were shown to be consistent with experimental observations. However, in the future, both k_{O_3} and the pathway prediction modules need to be further validated with more compounds with experimental data and to be improved/updated accordingly. The developed prediction platform is expected to be useful for various groups of end-users in research and practice such as environmental engineers, chemists, or toxicologists. In addition, the treatability of 9 polychlorobutadienes, which are groundwater contaminants, with ozone, UV photolysis at 254nm, and their advanced oxidation processes (i.e., O_3/H_2O_2 and UV/H_2O_2) was investigated. The abatement efficiencies for polychlorobutadienes during ozonation or O_3/H_2O_2 in a natural groundwater could be well explained based on the experimental k_{O_3} and k_{OH} -values. UV treatment was shown to be effective for the abatement of polychlorobutadienes. However, the potential formation of photoisomers from UV irradiation of chlorobutadienes with either *E* or *Z* configurations needs to be taken into account because this isomerization will not necessarily lead to a loss of the biological effects of these compounds.

Keywords: ozonation, micropollutants, second-order rate constant, prediction, transformation products, reaction pathways, highest occupied molecular orbital, natural bond orbital, UV irradiation, polychlorobutadienes

Résumé

L'ozonation qui est utilisée pour la désinfection de l'eau potable depuis le début du 20^{ème} siècle a récemment été introduite dans les procédés de traitement des eaux usées dans le but de réduire les quantités de micropolluants déversées dans les milieux aquatiques. Cependant, en raison de limitations expérimentales et analytiques relativement peu d'informations sont disponibles en ce qui concerne la formation de produits de transformation lors que la réaction entre l'ozone et les micropolluants. De plus, l'identification expérimentale de potentiels produits de transformation formés lors de l'ozonation des micropolluants est une tâche extrêmement difficile en raison du grand nombre de substances à considérer. Dans cette thèse, un modèle informatique visant à prédire les réactions entre l'ozone et les micropolluants ainsi que leurs constantes cinétiques a été développé dans le but (i) de fournir des informations sur la réactivité des micropolluants avec l'ozone exprimée sous forme de constantes de vitesse de second ordre (k_{O_3} , $M^{-1}s^{-1}$) et de (ii) prédire la formation de produits de transformation. (i) La réactivité des micropolluants avec de l'ozone a pu être prédite en utilisant une relation linéaire entre le logarithme des constantes de vitesse (k_{O_3}) de différents composés chimiques (p.ex. phénol, oléfines, amines, etc.) obtenues expérimentalement et les énergies des orbitales moléculaires délocalisées (HOMO, *highest occupied molecular orbital*) et des orbitales localisées (NBO, *natural bond orbital*) obtenues à l'aide de calculs basés sur des théories de chimie quantique ($R^2 = 0.75 - 0.95$ pour 14 groupes de composés représentant un total de 284 substances modèles). Dans son ensemble, le modèle développé dans cette thèse permet de prédire les valeurs de k_{O_3} avec en moyenne une précision d'un facteur ~ 5 par rapport aux valeurs expérimentales des composés modèles utilisés pour le développement de ce modèle ainsi que des chlorobutadiènes utilisés pour la validation externe. A l'inverse, de mauvaises prédictions (>10) ont été obtenues pour certains composés modèles exclus de la corrélation comme la cetirizine, deux pentachlorobutadiènes et l'hexachlorobutadiène qui ont été utilisés pour la validation externe de ce modèle. (ii) Le modèle présenté ici a été développé sur la base de nombreuses voies de réaction proposées dans la littérature scientifique. Une vaste compilation de mécanismes réactionnels publiés dans la littérature a été effectuée et seulement ceux considérés comme probable ont été encodés sous la forme de 340 règles de réactivité en utilisant des outils informatiques appropriés. Les voies réactionnelles ainsi que les produits de transformation prévus par le modèle pour certains micropolluants (la carbamazépine et le tramadol) se sont révélés compatibles avec les observations expérimentales. Cependant, les modules de prédiction des constantes cinétiques k_{O_3} et des voies réactionnelles doivent être encore validés avec les données expérimentales d'autres classes de composés afin d'améliorer et de mettre à jour ce modèle. La plate-forme de prédiction développée dans cette thèse pourrait s'avérer utile pour différents groupes d'utilisateurs tels que les ingénieurs en environnement, les chimistes et les toxicologues. Outre le développement d'un modèle prédictif, la traitabilité par ozonation, photolyse UV et d'autres procédés d'oxydation avancée (O_3/H_2O_2 et UV/H_2O_2) de neuf polychlorobutadiènes, connus pour être des contaminants des eaux souterraines, a été évaluée. La réduction des concentrations observées pour les polychlorobutadiènes traités par ozonation ou O_3/H_2O_2 dans une eau souterraine naturelle peuvent être expliquées sur la base des valeurs expérimentales de k_{O_3} et k_{OH} . Le traitement par UV des polychlorobutadiènes s'est également montré particulièrement efficace. Cependant, la formation de photo-isomère *E* et *Z* ont été observés. Par conséquent ceci indique que l'abattement des chlorobutadiènes ne conduit pas nécessairement à une réduction de la toxicité. Ainsi la formation de photo-isomères doit être soigneusement prise en compte pour l'évaluation du traitement des polychlorobutadiènes par photolyse UV.

Mots clés: ozonation, micropolluants, constante cinétique de second ordre, mécanisme réactionnel, prédiction, produits de transformation, rayonnement UV, polychlorobutadiène, HOMO, NBO

Table of Contents

Acknowledgements	i
Abstract (English/Français)	iii
Table of Contents	v
Chapter 1. General introduction	1
1.1. Micropollutants and their presence in aquatic systems	2
1.2. Introduction of micropollutants into aquatic systems	2
1.3. Delayed awareness of the presence of polar micropollutants	3
1.4. Environmental concerns and mitigation strategies associated with micropollutants ...	5
1.5. Ozonation as a mitigation strategy for micropollutants in wastewaters and ensuing chal- lenges	7
1.6. Reactions of ozone with organic compounds in aqueous solution: Kinetics and mecha- nisms	9
1.6.1. Kinetics of the reactions of ozone with organic compounds	9
1.6.2. Mechanisms for the reactions of ozone with organic compounds.....	13
1.7. Thesis outline	17
References	19
Chapter 2. Development of prediction models for the reactivity of organic compounds with ozone in aqueous solution by quantum chemical computations: role of delocalized and localized molecular orbitals	25
Abstract	26
2.1. Introduction	27
2.2. Materials and methods	28
2.2.1. Data set	28
2.2.2. Computational methodology.....	28

2.2.3. Data treatment and statistical model evaluation	28
2.3 Results and discussion	29
2.3.1 Boundary conditions for the quantum chemical computations	29
2.3.2 Correlations of molecular orbital energies with second order rate constants	29
2.3.3 Comparison of molecular orbital (MO) models with QSAR models	39
2.3.4 Practical implications	41
References	44
Supporting information for chapter 2	47

Chapter 3. Abatement of polychoro-1,3-butadienes in aqueous solution by ozone, UV-photolysis, and advanced oxidation processes (O_3/H_2O_2 and UV/H_2O_2) 119

Abstract	120
3.1. Introduction	121
3.2. Materials and methods	122
3.2.1 Standards and reagents	122
3.2.2 Analytical methods	122
3.2.3 Determination of second-order rate constants for the reactions of CBDs with ozone (k_{O_3}) and hydroxyl radicals ($k_{\bullet OH}$)	122
3.2.4. Determination of photon fluence-based first-order rate constants (k'), molar absorption coefficients (ϵ), and quantum yields (Φ) for phototransformation of CBDs at UV 254 nm	122
3.2.5. Ozonation, direct UV photolysis (254nm), and advanced oxidation of the selected CBDs and micropollutants in a natural groundwater	123
3.2.6. Quantum chemical computations for the prediction of k_{O_3} of CBDs	123
3.2.7. Kinetic modeling	123
3.3. Results and discussion	124
3.3.1. Kinetics of the reactions of CBDs with ozone ($k_{O_3, \text{exp}}$)	124
3.3.2. Prediction of rate constants for the reactions of CBDs with ozone ($k_{O_3, \text{pred}}$)	125
3.3.3. Kinetics for the reaction of CBDs with hydroxyl radicals ($k_{\bullet OH}$)	126
3.3.4. Determination of fluence-based first-order rate constants (k'), molar absorption coefficients (ϵ), and quantum yields (Φ)	126

3.3.5. Abatement of CBDs and micropollutants and bromate formation in a natural groundwater during ozonation, direct UV photolysis, and the AOPs O ₃ /H ₂ O ₂ and UV/H ₂ O ₂	129
3.3.6. Practical implications	132
References	134
Supporting information for chapter 3	137
Chapter 4. Development of a computer-based prediction platform for the reaction of ozone with organic compounds in aqueous solution: Kinetics and mechanisms	155
Abstract	156
4.1. Introduction	157
4.2. Materials and methods	158
4.2.1. Chemoinformatics and quantum chemical computation tools	158
4.2.2. Development of a computer-based prediction platform	158
4.3. Results and discussion	162
4.3.1. Rate constant (k_{O_3}) prediction	162
4.3.2. Pathway enumeration	166
4.3.3. Practical implications	169
References	171
Supporting information for chapter 4	175
Chapter 5. General conclusions and perspectives	239

Curriculum vitae

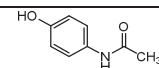
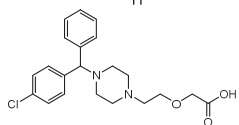
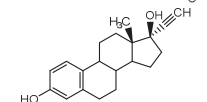
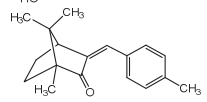
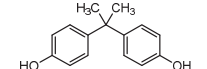
Chapter 1. General introduction

Chapter 1

1.1. Micropollutants and their presence in aquatic systems

Micropollutants or trace organic contaminants are collectively referred to as anthropogenic chemical substances as well as their (a)biotic transformation products present in aquatic environments at low concentrations (sub-ng/L – µg/L).¹ They include a variety of chemical compounds used for domestic, agricultural, and industrial purposes such as pharmaceuticals, personal care products, biocides, pesticides, additives, etc. (Table 1.1 shows some selected micropollutants). Numerous studies have reported the presence of micropollutants worldwide in various natural and technical aquatic systems such as lakes,²⁻⁴ rivers,⁴⁻⁸ groundwaters,⁸⁻¹¹ drinking water,^{12,13} wastewater,^{2,4,5,7,8,14-16} and even seawater.¹⁷ Bioaccumulation and harmful biological effects of micropollutants have been recognized for various aquatic organisms such as gemfibrozil in goldfish (*Carassius auratus*) with a reduced plasma testosterone¹⁸ and diclofenac in rainbow trout (*Oncorhynchus mykiss*) with a kidney alteration and the occurrence of an interstitial nephritis.^{19,20} As such, potential adverse impacts of micropollutants on the aquatic environment have been of concern.

Table 1.1. Selected micropollutants detected in the aquatic environment. The presented compounds were selected to represent a wide range of micropollutants in terms of the classes, applications, and structural characteristics relevant for their reactions with ozone (e.g., phenols, benzenes, amines, and olefins).

Class	Compound	Use	Chemical structure
Pharmaceuticals	Acetaminophen	Analgesic	
	Cetirizine	Antihistamine	
	17α-Ethinyl estradiol	Oral contraceptive	
Personal care products	Enzacamene	Sunscreen agent	
Industrial products	Bisphenol A	Plasticizer	

1.2. Introduction of micropollutants into aquatic systems

The introduction of micropollutants into the aquatic environment takes place through the water cycle shown in Figure 1.1. Once chemical substances are discharged after their use, wastewaters containing such substances (i.e., micropollutants) are collected through the urban drainage system and flow to wastewater treatment plants (WWTPs). As these plants were not designed to treat micropollutants, conventional municipal wastewater treatment processes inevitably discharge micropollutants in either their original form or as transformation products/metabolic products into the receiving water bodies. Con-

ventional wastewater treatment processes can serve as a partial barrier for some hydrophobic micropollutants such as tetracyclines and fluoroquinones by their sorption on activated sludge.^{21–24} Direct discharge without any treatment barrier can occur via leakage, runoffs from farms or livestock, or combined sewer overflows (CSOs). Once transported to a natural water system, micropollutants as any other water matrix component are circulated between different compartments through hydrological processes such as infiltration, evaporation, and precipitation while undergoing physical, chemical, and biological transformation processes. This eventually leads to the ubiquitous presence of micropollutants in aquatic environments.

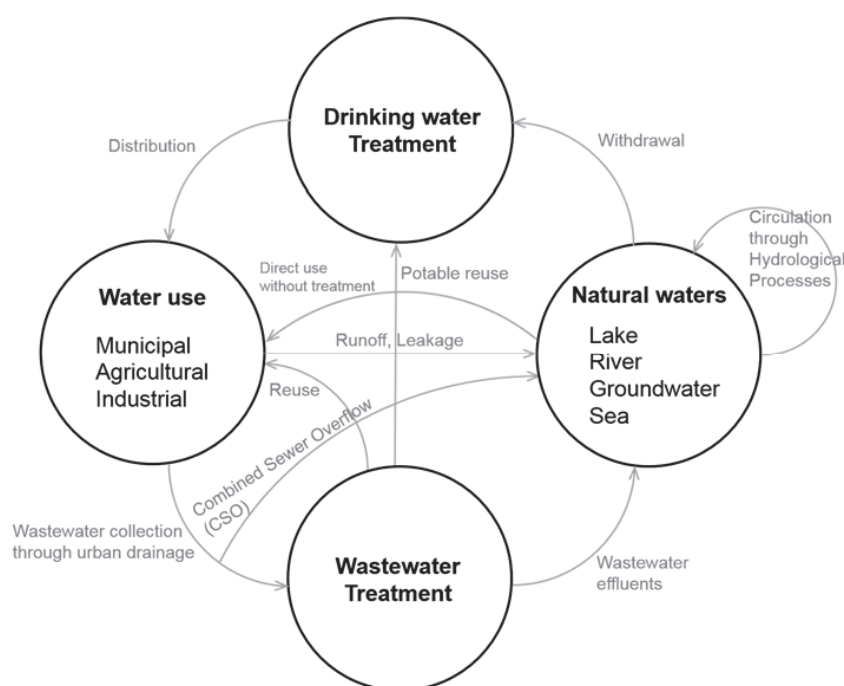


Figure 1.1. Urban water cycle comprising of four major elements: natural waters, drinking water and treatment, water use, and wastewater treatment.

1.3. Delayed awareness of the presence of polar micropollutants

Figure 1.2 shows a brief statistics about the number of research articles published between 1990 and July 2015 in selected academic journals in the field of aquatic environmental sciences, which were found by the keyword ‘pharmaceuticals’. Note that the keyword ‘pharmaceuticals’ only represents a fraction of the chemically diverse suite of micropollutants. As shown in Figure 2, no relevant research seems to have been conducted until the beginning of the 21st century. Since then, a significant increase of research papers has been observed up until 2015. A similar trend was also reported with a broader coverage in terms of keywords and academic journals.²⁵ Considering the fact that commercial pharma-

Chapter 1

ceuticals were already available more than half a century ago (e.g., peniciline was commercially available from around 1940) and a tremendous amount of drugs has been produced and consumed, comprehensive awareness of the ubiquitous occurrence of micropollutants has been apparently delayed.

This delay is mainly attributed to developments in analytical techniques, which allow to measure more polar compounds. The detection of micropollutants in the previous century was limited to rather non-polar volatile compounds, which could be analyzed by gas chromatography coupled with various detectors including mass spectrometry. Over the last several decades, various state-of-the-art analytical elements for polar organic compounds such as preconcentration, separation, detection, and data acquisition and interpretation have been improved for environmental samples. Especially, coupling of liquid chromatography with mass spectrometry with (online) solid phase extraction as a preconcentration step provides analytical capabilities for diverse polar compounds. Moreover, the price of such commercial analytical instruments has decreased, which made studies with state-of-the-art analytical equipment affordable (e.g., on-line solid phase extraction-gas or liquid chromatography-(high resolution) mass spectrometry) enabling more comprehensive investigations on micropollutants with various goals.

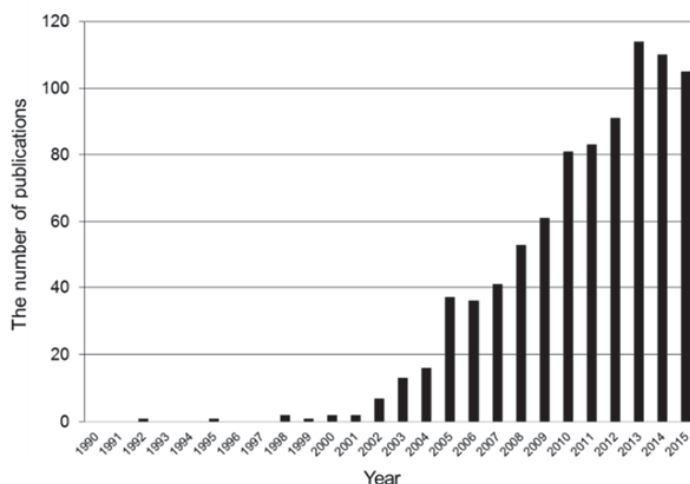


Figure 1.2. Number of publications found for with the keyword ‘pharmaceuticals’ in selected academic journals between 1990 and July 2015. The selected journals include Chemosphere (Elsevier), Environmental Science and Technology (American Chemical Society), Journal of Hazardous Materials (Elsevier), Science of the Total Environment (Elsevier), and Water Research (Elsevier) in alphabetical order.

1.4. Environmental concerns and mitigation strategies associated with micropollutants

Potential detrimental effects of micropollutants on both humans and aquatic organisms have attracted the attention of various stakeholders such as researchers, the public, policy makers, and governments. As for a human health risks, a human can be exposed to organic micropollutants via food ingestion (e.g., vegetables, crops, dairy products, meat or fish) or consumption of drinking water. The fraction of the

organic micropollutant exposure via drinking water is generally considered to be minor ($\leq 10\%$). Nonetheless, careful considerations are needed when direct potable reuse of wastewater effluents is practiced. Due to limited dilution, wastewater effluents to be reclaimed typically contain higher concentrations of micropollutants than receiving waters. Therefore, the state-of-the-art for direct water reuse includes reverse osmosis, which is an effective barrier against micropollutants.

In contrast, comparatively little information is available about detrimental effects of micropollutants on aquatic organisms except for a few prominent cases.^{26,27} The adverse effect of endocrine disrupting chemicals (EDCs), for example, is well recognized and established. Widespread feminization of wild-fish was reported in a Canadian lake²⁷ and in rivers in the UK.^{28,29} Natural steroidal estrogens, 17β -estradiol and estrone, and a synthetic steroid estrogen, 17α -ethinylestradiol, which are discharged from wastewater treatment plants were reported to be involved in sexual disruption in fishes.²⁸ Their low concentrations in the order of ng/L are known to be sufficient to manifest estrogenic effects on fish.^{27,30} There are various reasons for the elusiveness of ecotoxicological effects by most micropollutants, and some of the related major challenges are discussed in the following.

First, chemical analyses, even with the cutting-edge analytical techniques, only allow to detect, identify and quantify a limited suite of micropollutants in environmental samples. This can be linked to a conceptual iceberg of chemicals shown in Figure 1.3. The small tip of the iceberg above the water surface represents known or identified micropollutants, while the large chunk of the iceberg underneath is a plethora of unknown compounds. Therefore, elucidating a direct cause-effect relationship between an observed toxicity and micropollutants in a cocktail of numerous micropollutants is difficult to achieve. In other words, even if a toxicity is observed in an environmental sample, it would be difficult to pinpoint which chemicals lead to the effect, because the responsible micropollutants may not have been detected.

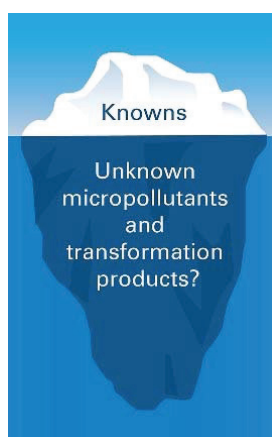


Figure 1.3. The distribution of known and unknown micropollutants and transformation/metabolic products in the aquatic environment can be conceptually represented by an iceberg model. From <http://www.australianwaterrecycling.com.au/projects/micropollutants-mixtures-and-transformation-products>

Chapter 1

Second, a number of *in vivo* and *in vitro* bioassays need to be simultaneously performed for obtaining comprehensive toxicological information since the effect of micropollutants on aquatic species may manifest in various modes of toxic actions (MOAs) (e.g., cytotoxicity, estrogenicity, mutagenicity, genotoxicity, etc.). This inevitably necessitates labor- and cost-intensive empirical investigations. For example, pharmaceuticals are designed for a specific biological target to achieve their therapeutic effect in humans or animals. Similar effects may occur to some extent for some lower aquatic animals. However, MOAs triggered by pharmaceuticals may be different from the intended target organisms for certain aquatic organisms.³¹ Therefore, various MOAs other than the expected MOA need to be explored using diverse *in vivo* and *in vitro* bioassays.³² For example, Escher *et al.* (2014) reported extensive research results obtained from the application of 103 *in vitro* bioassays to 10 water samples containing micropollutants.³³ Designing test batteries of bioassays, which are efficient and convenient as a routine monitoring tool for water utilities, is currently an important research area.^{34–38}

Third, little information is available on the effect of a long-term low level exposure to micropollutants. Micropollutants are present at a trace concentration to which aquatic organisms are continuously exposed for their whole lifespan. Thus, information about chronic toxic effects is equally or even more relevant than acute toxic effects. Nevertheless, the importance of chronic toxicity started to be recognized only about a decade ago,²⁶ wherefore, the corresponding information is currently still marginal.³¹

Having recognized the environmental concerns related to micropollutants as addressed above, various measures have been suggested and are under implementation or development. Mitigation of the aquatic pollution by micropollutants can be achieved by either source control or by enhancing end-of-pipe treatment at, e.g., municipal wastewater treatment plants. One of the source control strategies is a urine separation by which a substantial load of micropollutants into wastewaters can be avoided.³⁹ The nutrients such as nitrogen and phosphorus separated from urine can be used as fertilizers once it is properly stabilized.³⁹ Recognizing wastewater effluents as an unceasing input of micropollutants, the enhancement of the conventional wastewater treatment process with an advanced treatment technology can be considered as an end-of-pipe strategy. Viable technologies include ozonation, advanced oxidation processes (AOPs), activated carbon, membranes, etc. Briefly, ozonation is a chemical treatment process which utilizes ozone as a primary oxidant for oxidizing various contaminants.^{40,41} Moreover, hydroxyl radicals ($\cdot\text{OH}$), formed as secondary oxidants during ozonation, assist further oxidation.⁴² More details on ozone chemistry in the application of water and wastewater treatment are provided in section 1.6. AOPs are a chemical process using $\cdot\text{OH}$ as the main oxidant. They are generated by means of UV photolysis of H_2O_2 (UV/ H_2O_2), the reaction of O_3 with H_2O_2 ($\text{O}_3/\text{H}_2\text{O}_2$), the Fenton reaction ($\text{Fe(II)}/\text{H}_2\text{O}_2$), or combinations of the aforementioned processes.⁴³ The aforementioned oxidants only transform the chemical structures of micropollutants at practical oxidant dosages. Thus, it is an open question whether or not biological effects diminish with the decrease of a parent micropollutant and whether or not they are influenced by the formation of transformation products. In contrast, activated carbon is a physical treatment process which removes micropollutants via adsorption process.⁴³ In this case, there is no concern about transformation products in the treated wastewater. However, the disposal of the used activated carbon needs to be taken into account. Powered activated carbon (PAC) is removed via the sludge, which is incinerated in Switzerland.

1.5. Ozonation as a mitigation strategy for micropollutants in wastewaters and ensuing challenges

Ozonation is considered as a promising means to mitigate the discharge of micropollutants via wastewater effluents. It has been applied for disinfection in drinking water treatment since the early 20th century. The inactivation of microorganisms is thought to be mainly achieved by the oxidation of electron-rich nucleic acids by ozone.^{40,44} The same principle applies to the degradation of micropollutants because a number of micropollutants contain electron-rich moieties (see below section 1.6 for more details). The application of ozonation for mitigating micropollutants in wastewater effluents has been investigated in bench-,⁴⁵⁻⁴⁹ pilot-,^{45,49-53} and full-scale^{45,54-56} studies. These studies as a whole demonstrated that ozonation could induce significant degradation (>80%) for numerous micro-pollutants at reasonable specific ozone doses (0.4 ~ 1.0 g O₃/g DOC) (Figure 1.4).

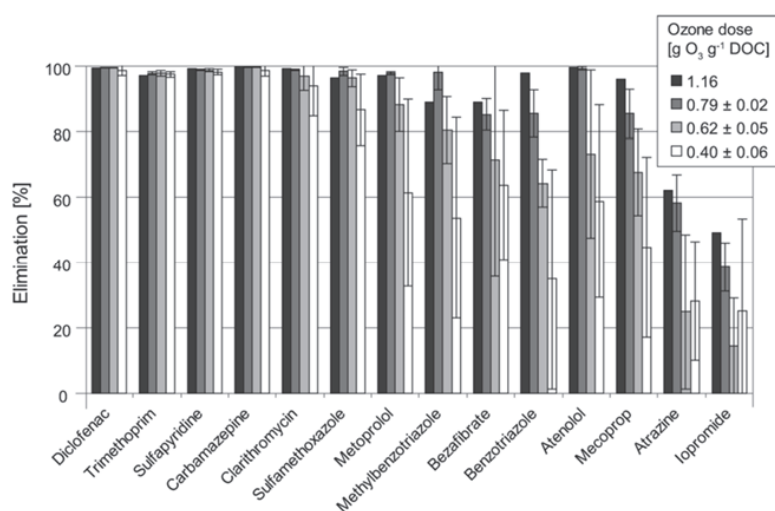


Figure 1.4. Effect of the specific ozone dose on the relative elimination of selected micropollutants in the ozone reactor of the Wüeri WWTP (Regensdorf, Switzerland). Reprinted with permission from Hollender et al., 2009.⁵⁴ Copyright (2009) American Chemical Society.

However, as ozone only induces a transformation of a chemical structure of micropollutants rather than a full mineralization, it may be possible that the resulting transformation products from ozonation are more toxic than their parent compounds. This indicates that ozone-transformation products have to be considered if an ozonation is applied to wastewater treatment for micropollutant abatement. To this end, along with the abatement of the parent micropollutants by ozone, many studies have been dedicated to shed light on (eco)toxicological aspects of ozone transformation products by various viable bioassays. Two differing approaches are mainly applied: (1) exposure-based methods and (2) effect-based methods.

Chapter 1

In the exposure-based methods, the disappearance of a micropollutant during ozonation is compared with a MOA of interest. For example, the estrogenicity of a solution containing 17α -ethinylestradiol (EE2) was assessed as a function of the ozone dose.^{57,58} The relative estrogenic activity decreased proportionally with the relative residual EE2 concentration. The ozonated solution as a composite sample contains not only the residual parent EE2 but also the transformation products. Therefore, it could be deduced that none of the oxidation products from the reaction of EE2 with ozone manifests estrogenicity. Various micropollutants were investigated in the same manner using various bioassays.^{59–62} The biological effects of many micropollutants were reported to be proportionally correlated to the residual concentration of a parent micropollutant. However, several cases with differing observations were also reported. Sulfoxides, which are ozone transformation products of penicillin G and cephalixin, were reported to retain antibacterial activities⁵⁹ although significantly lower than the parent compounds.⁶³ A temporary moderate increase in antiviral activity was reported in ozonated solutions containing oseltamivir acid, the active metabolite of Tamiflu[®], for about 50% transformation, which disappeared when samples were analyzed after 24h.⁶²

In the effect-based methods, (eco)toxicological assessments are carried out for ozonated wastewater effluents by selected bioassays without identification nor quantification of individual micropollutants. Many studies reported improvements as well as deterioration of the water quality related to biological effects upon ozonation.^{36,64–74} Often, an improvement of the water quality to the level prior to ozonation or better was reported after a biological post-treatment or activated carbon filtration.^{64,66,73}

Ecotoxicologically harmful transformation or oxidation products may have resulted from the reaction of ozone with either micropollutants, water matrix components such as dissolved organic matter (DOM), or combinations thereof. Oxidation by-products from ozonation of DOM are difficult to characterize because of the structural complexity of DOM, whereas chemical analysis for the transformation products can be achieved using viable state-of-the-art analytical techniques such as mass spectrometry and nuclear magnetic resonance. However, based on the fact that current analytical methods allow only a partial identification of the total pool of micropollutants (see above), this task is still formidable to achieve. The analyses of the huge number of environmentally relevant micropollutants including hitherto undetected compounds and their ozonation transformation products would be extremely challenging. In addition, an empirical toxicity evaluation of all individual transformation products is impossible.

Due to its importance in technical aquatic systems (section 1.4), ozone chemistry in aqueous phase has been studied over the last century. The elucidation of kinetics and mechanisms for the reactions of ozone with organic compounds was central. Often in literature, similar transformation products were observed for micropollutants possessing similar reaction centers (see the following section for more details). Therefore, predicting chemical structures of potential transformation products is possible for micropollutants for which a reference study is available, and the resulting information can be useful in many ways. First, the structural information of potential transformation products can assist environmental analytical chemists for a suspected-target screening. Second, it can serve as a toxicity screening battery when combined with *in silico* toxicity prediction tools. This could help to narrow down the list

of micropollutants, for which ozone transformation products are potentially toxic. Moreover, a predicted reactivity of a micropollutant with ozone (expressed by k_{O_3} , see below) can be used to evaluate the abatement efficiency of a parent micropollutant during ozonation.

Nevertheless, no attempt to develop a prediction system based on the accumulated knowledge in ozone chemistry has been made so far. In the subsequent section 1.6, the knowledge about kinetics and mechanisms for the reactions of organic compounds with ozone are briefly presented with a focus on developing a prediction system for the formation of transformation products during ozonation.

1.6. Reactions of ozone with organic compounds in aqueous solution: Kinetics and mechanisms

1.6.1. Kinetics of the reactions of ozone with organic compounds

Reactivity (k_{O_3}) The reactivity of a target micropollutant (A) with ozone is critical for ozonation processes and it is reflected by a second-order rate law with a second order rate constant (k_{O_3}) (Eq. 1.1).

$$-\frac{d[A]}{dt} = k_{O_3}[A][O_3] \quad (1.1)$$

The rate at which the concentration of A ($[A]$ in units of $M \cdot s^{-1}$) decreases is first-order with respect to the concentration of a chemical compound ($[A]$, M) and the concentration of ozone ($[O_3]$, M), and k_{O_3} is the second-order rate constant for the reaction of ozone with a chemical compound in units of $M^{-1}s^{-1}$. An integration of Eq. 1.1 over time yields

$$-\ln\left(\frac{[A]}{[A]_0}\right) = k_{O_3} \int [O_3] dt \quad (1.2)$$

$\int [O_3] dt$ is the ozone exposure in units of $M \cdot s$. Note that $\cdot OH$ are formed during ozonation (see section 1.6.2) and they can also contribute to the abatement of A. Hydroxyl radicals can be suppressed using a radical scavenger (e.g., *tert*-butanol) for k_{O_3} measurements when necessary.

Based on Eq. 1.2 with $\cdot OH$ suppressed by a radical scavenger, k_{O_3} -values have been determined and are available for several hundred organic and inorganic compounds using different experimental procedures.^{40,75-77} Among those, the following four reactive moieties are commonly found in many organic compounds and are of importance for ozonation: aromatic systems, olefins, amines, and organosulfur moieties. Figure 1.5 shows the range of experimental k_{O_3} for the respective moieties. Overall, the ranges of k_{O_3} -values for aromatic compounds, olefins, amines, and organosulfur compounds span over about 12, 6, 7, and 6 orders of magnitude, respectively, indicating a great variability in k_{O_3} even for one compound group.

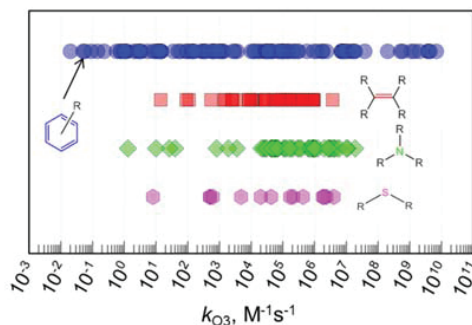


Figure 1.5. Range of experimental k_{O_3} for the reactions of ozone with aromatic compounds (n=112), olefins (n=45), amines (n=59), sulfur-containing compounds (n=14). All k_{O_3} -values were obtained from von Sonntag and von Gunten⁴⁰ and the individual compounds are not listed.

As ozone is an electrophile, which reacts preferentially with electron rich chemical moieties, the broad range of k_{O_3} is attributed to the electron density of these moieties in an organic compound. The electron density of a moiety can be significantly influenced by the electron-donating/withdrawing properties of substituents, the protonation state of a substituent and of the reactive moiety itself, and combinations thereof.

In Figure 1.6, selected compounds are presented to show the effect of substituents on k_{O_3} for olefins, aromatic compounds, and amines. While the presence of electron-donating substituents such as methyl groups leads to an increased k_{O_3} relative to the unsubstituted group, electron-withdrawing groups such as halogen atoms or nitro groups result in a decrease of k_{O_3} . This is shown for olefins with k_{O_3} -values decreasing consecutively from ethene to tetrachloroethene upon chlorine addition and the substituent effect is additive. Moreover, the protonation state of a reactive moiety or a substituent is of importance. k_{O_3} of the phenolate is about six orders of magnitude higher than that of phenol. This is attributable to the increased electron density on the aromatic ring by delocalization of the negative charge on the oxygen of the deprotonated hydroxy group. The lone-pair electrons of the amine nitrogen are the site of attack by ozone. Upon protonation, the amine nitrogen becomes non-reactive with ozone as the lone-pair electrons on the nitrogen are no longer available (e.g., $k_{O_3} = 1.9 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$ for dimethylamine and $k_{O_3} < 0.1 \text{ M}^{-1}\text{s}^{-1}$ for dimethylammonium,⁷⁶ respectively).

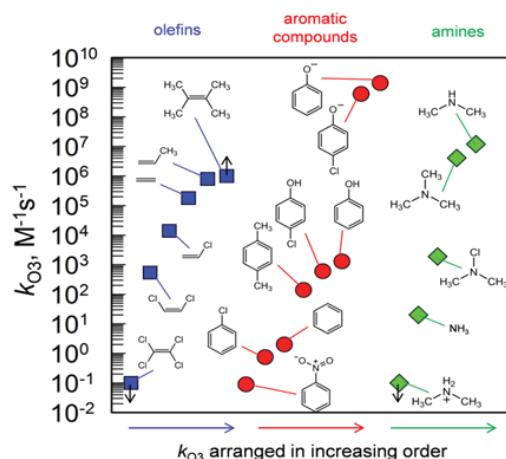


Figure 1.6. k_{O_3} for the reactions of selected olefins, aromatic compounds, and amines with ozone. k_{O_3} are sorted by increasing values on the abscissa. k_{O_3} -values were obtained from a compilation of experimental k_{O_3} -values.⁴⁰ The data points with arrows indicate that the reported values are only given as an approximate range.

The significant dependence of k_{O_3} on the protonation state of a substituent or a reactive moiety suggests that k_{O_3} can vary dramatically as a function of the pH. For a monoprotic compound (HA) as an example, the fractions of the compound (α_{HA}) and its conjugate base (α_{A^-}) at a certain pH can be calculated as follows (Eqs. 1.3 and 1.4).

$$\alpha_{HA} = \frac{[HA]}{[HA]_{total}} = \frac{[H^+]}{K_a + [H^+]} \quad (1.3)$$

$$\alpha_{A^-} = \frac{K_a}{K_a + [H^+]} = 1 - \alpha_{HA} \quad (1.4)$$

K_a is the dissociation constant for HA and $[H^+]$ is the concentration of protons.

Taking into account the presence of the multiple species for the reaction kinetics, Eq. 1.1 can be reformulated as follows:

$$-\frac{d[HA]_{total}}{dt} = k_{O_3,HA}[HA][O_3] + k_{O_3,A^-}[A^-][O_3] = k_{O_3,HA}\alpha_{HA}[HA]_{total}[O_3] + k_{O_3,(A^-)}\alpha_{A^-}[HA]_{total}[O_3] \quad (1.5)$$

$$-\frac{d[HA]}{dt} = k_{O_3,app}[HA]_{total}[O_3] \quad (1.6)$$

$$k_{O_3,app} = k_{O_3,HA}\alpha_{HA} + k_{O_3,(A^-)}\alpha_{A^-} \quad (1.7)$$

$k_{O_3,HA}$, k_{O_3,A^-} , and $k_{O_3,app}$ are the second-order rate constants for the reaction of ozone with HA and A^- , and the apparent second-order rate constant for the reaction of HA and A^- with ozone at the pH of interest, respectively, and $[HA]_{total}$ is the sum of the concentrations of HA and A^- .

Chapter 1

To date k_{O_3} -values have been mainly empirically determined for the compounds of interest. However, this approach requires time-consuming and labor- and resource-intensive laboratory work. Especially in the context of the large numbers of micropollutants that are already known to be present in aquatic environments and since increasing numbers of micropollutants are discovered every year, this approach seems ineffective.

Several methods for predicting k_{O_3} for organic compounds were previously developed and may aid to overcome the aforementioned problems. Quantitative structure-activity relationships (QSAR) were developed for aromatic compounds, olefins, and amines.⁷⁸ Therein, the sum of empirically derived substituent constants of substituents at the reaction center was correlated satisfactorily with the corresponding k_{O_3} -values ($R^2 > 0.8$, $n = 14 \sim 54$) (Figure 1.7a). Hammett constants denoted as σ , σ^+ , or σ^- were successfully correlated with k_{O_3} for aromatic compounds. Taft constants denoted as σ^* were used for aliphatic compounds such as olefins and amines. However, some compounds could not be included due to the unavailability of Hammett or Taft constants for certain substituents. Despite the fact that the performance of the QSAR model was successful, this aspect remains a limitation. A multi-linear regression (MLR)-based QSAR model was also developed employing three molecular descriptors such as the ionization potential, double bond equivalence, and weakly-polar component of solvent accessible surface area (WPSA) ($R^2 > 0.83$, $n = 27$).⁷⁹ In contrast to the QSAR model above, the developed MLR model is one unified model, which in principle predicts k_{O_3} for any chemical structure. However, as only 27 and 13 compounds were used for the model development and an external validation, respectively, the model performance remains to be evaluated with a larger data set. Moreover, the MLR approach seeks for a model with the best goodness-of-fit regardless of the mechanistic relevance of the parameters incorporated. This indicates that the multiple numbers of the molecular descriptors incorporated may or may not be mechanistically relevant and it is difficult to generate mechanistic insights with this model.

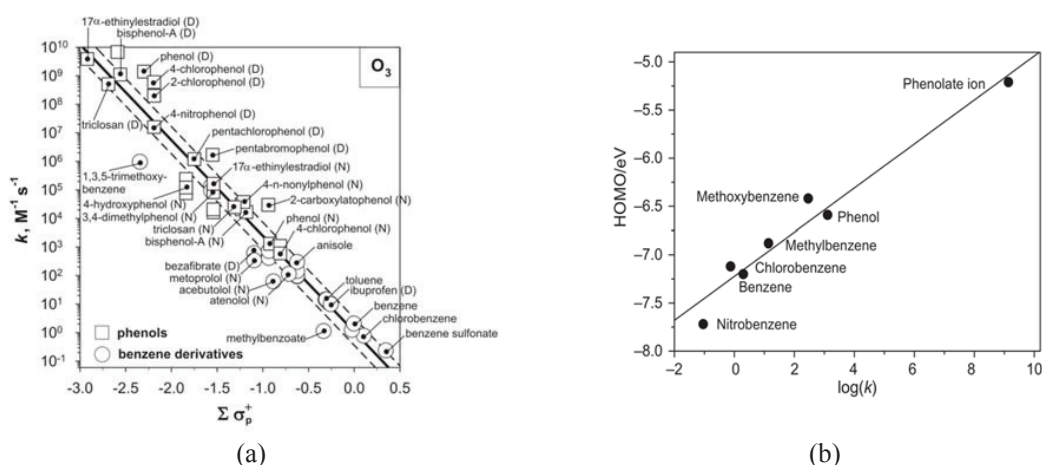


Figure 1.7. Currently available k_{O_3} prediction methods (aromatic compounds). (a) Correlation between $\log k_{O_3}$ for the reaction of ozone with benzene derivatives ($n=50$) vs. $\Sigma \sigma_p^+$. Reprinted from Lee and von Gunten., 2012.,⁷⁸ Copyright (2012), with permission from Elsevier. (b) Correlation of the energy of the HOMO with $\log k_{O_3}$ for the reaction of ozone with the seven selected benzene derivatives. Reprinted from Naumov and von Sonntag., 2010.⁸⁰ Copyright (2010), with permission from Taylor & Francis.

Another approach employs a quantum molecular descriptor, which can be obtained from quantum chemical computations. Amongst several quantum molecular descriptors investigated, the highest occupied molecular orbital (HOMO) energy of an aromatic compound was reported to correlate well with the corresponding k_{O_3} -values ($R^2=0.92$, $n=7$) (Figure 1.7b).⁸⁰ Moreover, the Gibbs free energy for the formation of an ozone adduct to an aromatic compound at the para position showed an excellent correlation ($R^2=0.99$, $n=8$).⁸⁰ As these quantum molecular descriptors are in principle attainable for any chemical compound, the applicability of this approach is more general compared to the QSAR model described above. Moreover, both descriptors are supported theoretically and mechanistically, thus they can be helpful for a better understanding of ozone reactions with organic compounds. For example, the success of the HOMO can be supported by the frontier molecular orbital (FMO) theory according to which the reaction of ozone with aromatic compounds is driven by the interaction between the HOMO of the aromatic compound and the LUMO of ozone.^{80,81} Since the LUMO of ozone is constant, the differing HOMO energies of aromatic compounds were correlated with the corresponding k_{O_3} -values. Despite the aforementioned advantages, the correlation of the quantum chemical approach is premature to be accepted because only a small number of aromatic compounds ($n=8-9$) was investigated. Therefore, a general application of the quantum chemical model for predicting k_{O_3} remains to be tested.

Based on the prediction models using the QSAR approach and the quantum chemical computations developed in this PhD thesis (see chapter 2), it is worthwhile to note that a predicted k_{O_3} by such models and an experimental k_{O_3} intrinsically differ in that the former and the latter are functional site-specific and species-specific, respectively. Note that in this discussion the acid-base speciation is ignored and the presence of multiple conformers was not considered for simplicity. For example, the chemical structure of cetirizine (Table 1.1) contains two amino groups and two benzene rings. Only one k_{O_3} is empirically measured, which represents the overall reactivity of the whole structure including these four moieties. In contrast, four k_{O_3} -values will be predicted by a quantum chemical computation. In principle, the sum of the four predicted k_{O_3} -values corresponds to the experimental k_{O_3} .

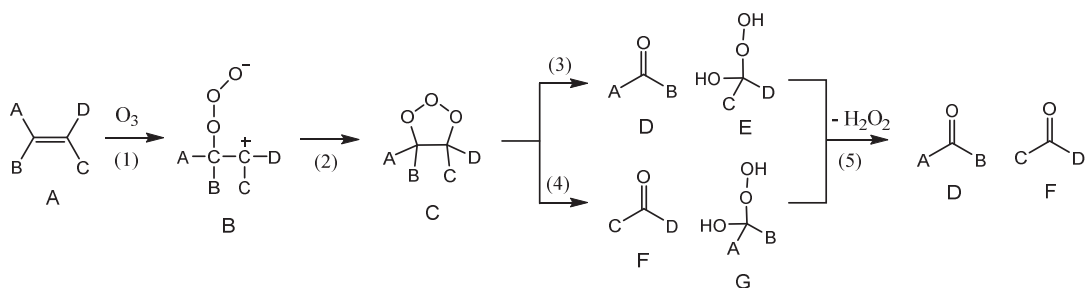
1.6.2. Mechanisms for the reactions of ozone with organic compounds

Over the last decades, numerous studies have been dedicated to the investigation of ozone transformation products of organic compounds containing aforementioned moieties. Based on these studies, various generalized reaction mechanisms have been proposed. In the following, a brief introduction to widely accepted reaction mechanisms for olefins, aromatic compounds, amines, and sulfur-containing compounds is provided. For detailed mechanisms, refer to von Sonntag and von Gunten (2012)⁴⁰ and the references therein.

Olefins. The Criegee mechanism, which was suggested by Rudolf Criegee⁸² and is named after him, is widely accepted for the reactions of olefins with ozone (Scheme 1.1).^{40,83} As shown in Scheme 1.1, ozone initially attacks at the double bond to form a cyclic ozonide intermediate (C). The ozonide subsequently undergoes a cleavage of the C-C σ bond to yield two primary products, a carbonyl compound and a hydroxyhydroperoxide. The latter product (E or G) releases hydrogen peroxide to eventually form a carbonyl compound. For asymmetrically substituted olefins, two differing reaction pathways from the

Chapter 1

ozonide to two primary products are possible (reactions (3) and (4)). The organic hydroxyhydroperoxides decompose to yield another carbonyl compound and hydrogen peroxide (reaction (5)). Due to H_2O_2 elimination, identical end products can be obtained for asymmetrically substituted olefins. It is well known that no radical species are formed during the Criegee mechanism.

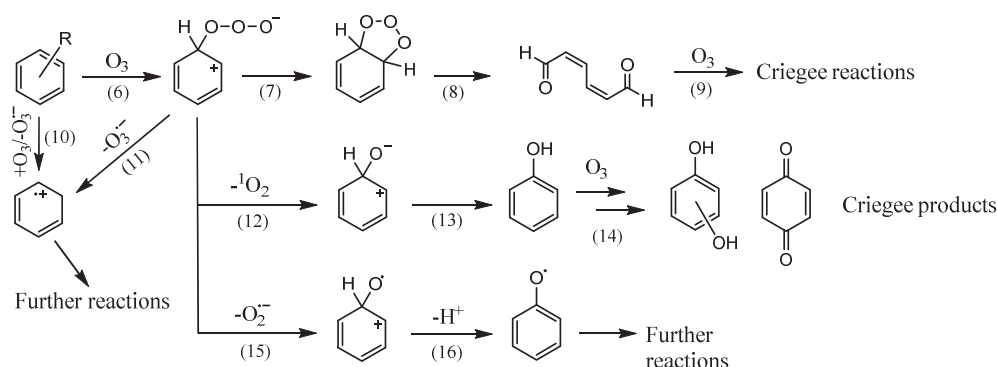


Scheme 1.1. Major reaction mechanisms for the reactions of ozone with olefins

Micropollutants containing an olefinic moiety such as carbamazepine,⁸⁴ cephalixin,⁶³ and progesterone⁸⁵ were reported to give rise to products via the Criegee mechanism. Moreover, several heteroaromatic compounds containing a C-C double bond such as thymine and thymidine were also reported to react by a Criegee-type mechanism.^{40,86} The resulting transformation products may undergo intra-molecular reactions, leading to the formation of the ensuing products.

Aromatic compounds. Contrary to olefins, which react predominantly via the Criegee mechanism, the reactions of aromatic compounds with ozone proceed via various reaction pathways, enumerating various oxidation products.⁴⁰ The elucidation of reaction mechanisms and oxidation products for ozonation of aromatic compounds is challenging for several reasons. Firstly, primary products often react by several orders of magnitude faster with ozone than the parent compound, which hinders a detection of the primary products. For example, while k_{O_3} for the reaction of benzene with ozone is only $2 \text{ M}^{-1}\text{s}^{-1}$,⁷⁵ phenol and phenolate, a hydroxylated product of benzene, have k_{O_3} of $1.3 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$ and $1.4 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$,⁷⁶ respectively. Moreover, olefinic products such as muconic acid have k_{O_3} in the order of $10^4 \text{ M}^{-1}\text{s}^{-1}$ in absence of electron-withdrawing substituents.⁴⁰ Consequently, primary products may not be present in detectable concentrations because they are further oxidized with ozone as soon as they are formed. Secondly, free radicals such as $\cdot\text{OH}$, which can be formed via several reaction pathways, can also play an important role. As the reaction of $\cdot\text{OH}$ with organic compounds is less selective, a number of minor products can be produced. Thirdly, the resulting oxidation products would differ in their physico-chemical properties (e.g., ionic vs non-ionic, chromophoric vs non-chromophoric, volatile vs non-volatile, etc). This necessitates the implementation of various analytical techniques. With all difficulties mentioned above, a poor mass balance between the oxidized parent compound and the identified products is inevitable, and a complete elucidation of reaction mechanisms is consequently difficult to achieve. Phenol^{87,88} and methoxylated benzenes⁸⁹ are good candidates for investigations of their oxidation products and reaction mechanisms because these compounds can kinetically compete with primary products

(e.g., phenolic products) toward ozone. Nonetheless, the mass balance is still not complete due to the complexity of the reactions involved as described above. Based on the finding of these studies, reaction mechanisms for aromatic compounds are briefly presented in Scheme 1.2.



Scheme 1.2. Major reaction mechanisms for the reactions of ozone with aromatic compounds

An ozone attack on an aromatic compound is considered to be initiated by the formation of an ozone adduct at the electron-rich aromatic carbon (e.g., carbons in ortho or para position for phenol) (reaction (6)). A cyclic ozonide formation can follow subsequently (reaction (7)), leading to a ring cleavage and the formation of olefinic products (reaction (8)). The resulting olefinic compounds can be further oxidized by ozone via the Criegee mechanism (reaction (9)), i.e., reactions (1)-(5)). An ozonide radical anion ($O_3^{\bullet-}$) can be formed via either direct electron transfer from the aromatic ring to ozone (reaction (10)) or the dissociation of the ozone adduct (reaction (11)) and its formation gives rise to $\cdot OH$ via reactions (17) and (18).⁹⁰



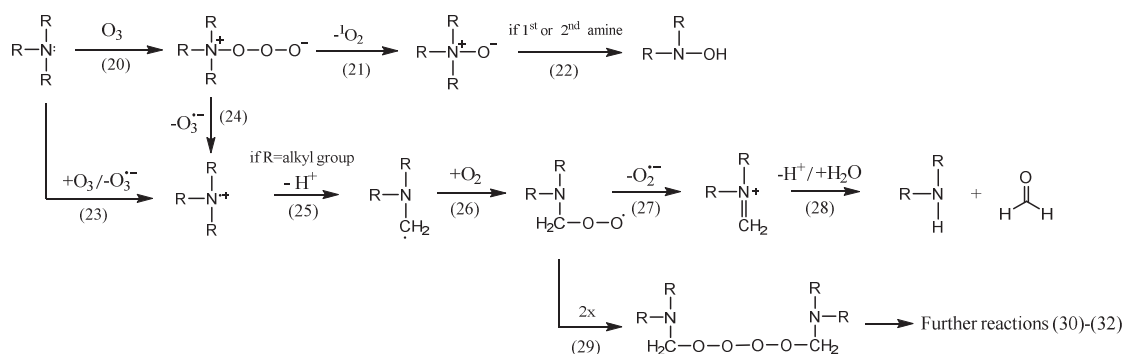
Hydroxylation via reactions (12) - (13) are well established based on the detection of hydroxylated products (i.e., phenolic compounds) as well as the detection of singlet oxygen (1O_2). Phenolic compounds are further oxidized by ozone to form hydroxylated products, quinones, and Criegee products (reaction (14)). Superoxide and phenoxyl radicals can be formed from the ozone adduct (reactions (15) and (16)), where the former also serves as a precursor to $\cdot OH$ via reaction (19), which entails an electron transfer between superoxide and ozone to form oxygen and the ozonide radical anion.



Many micropollutants containing an aromatic ring such as 17 α -ethinylestradiol,⁹¹ protonated propranolol,⁹² metoprolol,⁹³ bisphenol A,⁹⁴ estrone,⁹¹ benzotriazole,⁹⁵ and triclosan⁹⁶ have been reported to give rise to the common products resulting from reaction pathways for aromatic compounds (Scheme 1.2).

Chapter 1

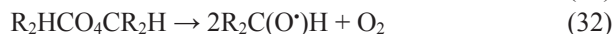
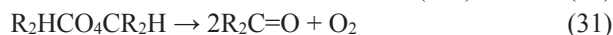
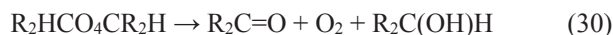
Amines. Ozone readily reacts with the lone-electron pair of the amine-nitrogen to form an ozone adduct (reaction (20) in Scheme 1.3), which can subsequently decompose into a *N*-oxide and singlet oxygen ($^1\text{O}_2$) (reaction (21)), or undergo a heterolytic cleavage of the nitrogen–oxygen bond to form an amine radical cation and an ozonide radical (reaction (24)). These two radical species can also be formed via a direct single electron transfer (reaction (23)).



Scheme 1.3. Major reaction mechanisms for the reactions of ozone with amines

For secondary amines, a *N*-oxide seems to rearrange into an isomeric hydroxylamine via reaction (22) which has been reported for metoprolol, propranolol, and piperidine all of which contain a secondary amine moiety.^{92,93,97} Moreover, quantum chemical computations for dimethylamine reactions (20), (21), and (22) are exergonic ($\Delta G^0 = -41 \text{ kJ mol}^{-1}$, -74 kJ mol^{-1} , and -20 kJ mol^{-1} , respectively) and support this mechanism.⁴⁰ Although there is no certain empirical evidence yet, it may be also possible that a primary amine gives rise to a hydroxylamino group following the rearrangement of the *N*-oxide.

Dealkylation takes place via reactions (20) and (23) – (28) involving the formation of peroxy radicals. Micropollutants containing tertiary amines such as clarithromycin,⁶⁰ roxithromycin,⁹⁸ and tramadol⁹⁹ were reported to yield both *N*-oxide as a major product and dealkylated amine products as a minor product. Moreover, peroxy radicals may undergo a bimolecular decay (reaction (29)) through which different ensuing pathways are known to be possible (reactions (30)-(33)).^{40,100}



Note that tertiary peroxy radicals are restricted to the reactions (32) and (33) as a α -C-H bond to the peroxy group is required for the reactions (30) and (31).

As a unique case, ozonation of hydrochlorothiazide gave rise to an exclusive conversion to chlorothiazide, which can be explained by reactions (20) and (23) – (27).¹⁰¹ In contrast to other cases described

above, reaction (25) gives rise to an imine, which is terminal for hydrochlorothiazide. Similar to aromatic compounds, precursors to hydroxyl radicals such as the ozonide radical and superoxide are formed in reactions (23), (24), and (27). This indicates that hydroxyl radicals may contribute to the transformation of a micropollutant containing amino groups.

Sulfur-containing compounds. An oxygen-transfer is known to be the major reaction mechanism for many sulfur-containing compounds based on the observations of nearly 100 % $^1\text{O}_2$ yields from ozonation of sulfides, disulfides, and sulfinic acids.¹⁰² The resulting product from the O-transfer is sulfoxide.⁴⁰ Penicillin G and cephalixin both containing a sulfide group were reported to yield the corresponding sulfoxide products as the major product from ozonation.⁶³ Sulfoxides are no longer reactive with ozone (e.g., $k_{\text{O}_3}=8 \text{ M}^{-1}\text{s}^{-1}$ for dimethyl sulfoxide¹⁰³). In contrast, an O-transfer seems not to be the only mechanism for thiols, as low $^1\text{O}_2$ yields (<50%) were observed for 1,4-dithiothreitol.¹⁰² Moreover, a pH-dependence on the $^1\text{O}_2$ yields was reported for 1,4-dithiothreitol (46% and 17% at pH 4.8 and 9, respectively).¹⁰²

1.7. Thesis outline

This doctoral thesis, which was funded by the Swiss Federal Office for the Environment (FOEN), was dedicated to the development of an *in silico* prediction system for transformation products formed during ozonation of micropollutants in aqueous solution. With the information about reaction kinetics and mechanisms available in the literature as a basis, which are briefly introduced in chapter 1, a computer-based prediction tool for reaction kinetics and reaction mechanisms was developed. The methodological developments for the prediction platform are presented in detail in chapters 2 and 4. Chapter 3 describes an experimentally-derived validation data set for a particular class of contaminants (i.e., olefins).

Chapter 1. The motivation for the research project of this PhD thesis is presented by giving a state-of-the-art compilation of the significance of micropollutants in the aquatic environment. Moreover, a brief introduction to the current knowledge of ozone chemistry is presented, which was the basis for the development of the computer-based prediction system.

Chapter 2. A prediction model for second-order rate constants (k_{O_3}) for the reactions of ozone with organic micropollutants was developed. Linear relationships between $\log k_{\text{O}_3}$ for compounds with a large variety of chemical moieties and the corresponding molecular orbital energies obtained from quantum chemical computations were developed. The prediction performances and the practical implications of the developed models are discussed.

Chapter 3. The performance of the k_{O_3} prediction model developed in chapter 2 was tested for nine polychlorobutadienes, which are problematic groundwater contaminants in the Hardwald in the canton of Baselland, Switzerland. Experimentally determined k_{O_3} -values for chlorobutadienes were compared with the predicted k_{O_3} . Besides ozonation, the treatability of the nine polychlorobutadienes by direct UV photolysis and the advanced oxidation processes ozone/hydrogen peroxide and UV/hydrogen peroxide was also investigated.

Chapter 1

Chapter 4. A computer-based prediction platform for the oxidation of organic compounds with ozone was developed as a standalone graphical user interface (GUI) application. Development details such as the structure of the prediction platform, the role and integration of individual elements, and the utilized cheminformatics tools are presented. Finally, the developed prediction platform is demonstrated for the reactions of ozone with some relevant micropollutants, for which experimental data are available for comparison. Based on these initial validation exercises, the opportunities/advantages and limitations of the developed prediction platform are highlighted.

Chapter 5. General conclusions and perspectives on the critical findings of this thesis are given. Practical implications, limitations, and further improvements for the currently developed prediction platform and key findings and discussions of abatement of polychlorobutadienes are presented.

References

- (1) Schwarzenbach, R. P.; Escher, B. I.; Fenner, K.; Hofstetter, T. B.; Johnson, C. A.; von Gunten, U.; Wehrli, B. The Challenge of Micropollutants in Aquatic Systems. *Science* (80-.). **2006**, *313*, 1072–1077.
- (2) Vanderford, B. J.; Snyder, S. A. Analysis of Pharmaceuticals in Water by Isotope Dilution Liquid Chromatography/Tandem Mass Spectrometry. *Environ. Sci. Technol.* **2006**, *40*, 7312–7320.
- (3) Bonvin, F.; Rutler, R.; Chèvre, N.; Halder, J.; Kohn, T. Spatial and Temporal Presence of a Wastewater-Derived Micropollutant Plume in Lake Geneva. *Environ. Sci. Technol.* **2011**, *45*, 4702–4709.
- (4) Kim, S. D.; Cho, J.; Kim, I. S.; Vanderford, B. J.; Snyder, S. A. Occurrence and removal of pharmaceuticals and endocrine disruptors in South Korean surface, drinking, and waste waters. *Water Res.* **2007**, *41*, 1013–1021.
- (5) Ashton, D.; Hilton, M.; Thomas, K. V. Investigating the environmental transport of human pharmaceuticals to streams in the United Kingdom. *Sci. Total Environ.* **2004**, *333*, 167–184.
- (6) Kolpin, D. W.; Furlong, E. T.; Meyer, M. T.; Thurman, E. M.; Zaugg, S. D.; Barber, L. B.; Buxton, H. T. Pharmaceuticals, Hormones, and Other Organic Wastewater Contaminants in U.S. Streams, 1999–2000: A National Reconnaissance. *Environ. Sci. Technol.* **2002**, *36*, 1202–1211.
- (7) Glassmeyer, S. T.; Furlong, E. T.; Kolpin, D. W.; Cahill, J. D.; Zaugg, S. D.; Werner, S. L.; Meyer, M. T.; Kryak, D. D. Transport of Chemical and Microbial Compounds from Known Wastewater Discharges: Potential for Use as Indicators of Human Fecal Contamination. *Environ. Sci. Technol.* **2005**, *39*, 5157–5169.
- (8) Ternes, T. A.; Hirsch, R. Occurrence and Behavior of X-ray Contrast Media in Sewage Facilities and the Aquatic Environment. *Environ. Sci. Technol.* **2000**, *34*, 2741–2748.
- (9) Barnes, K. K.; Kolpin, D. W.; Furlong, E. T.; Zaugg, S. D.; Meyer, M. T.; Barber, L. B. A national reconnaissance of pharmaceuticals and other organic wastewater contaminants in the United States — I) Groundwater. *Sci. Total Environ.* **2008**, *402*, 192–200.
- (10) Sacher, F.; Lange, F. T.; Brauch, H.-J.; Blankenhorn, I. Pharmaceuticals in groundwaters. *J. Chromatogr. A* **2001**, *938*, 199–210.
- (11) Lapworth, D. J.; Baran, N.; Stuart, M. E.; Ward, R. S. Emerging organic contaminants in groundwater: A review of sources, fate and occurrence. *Environ. Pollut.* **2012**, *163*, 287–303.
- (12) Benotti, M. J.; Trenholm, R. A.; Vanderford, B. J.; Holady, J. C.; Stanford, B. D.; Snyder, S. A. Pharmaceuticals and Endocrine Disrupting Compounds in U.S. Drinking Water. *Environ. Sci. Technol.* **2009**, *43*, 597–603.
- (13) Loraine, G. A.; Pettigrove, M. E. Seasonal Variations in Concentrations of Pharmaceuticals and Personal Care Products in Drinking Water and Reclaimed Wastewater in Southern California. *Environ. Sci. Technol.* **2006**, *40*, 687–695.
- (14) Lindqvist, N.; Tuhkanen, T.; Kronberg, L. Occurrence of acidic pharmaceuticals in raw and treated sewages and in receiving waters. *Water Res.* **2005**, *39*, 2219–2228.
- (15) Nakada, N.; Tanishima, T.; Shinohara, H.; Kiri, K.; Takada, H. Pharmaceutical chemicals and endocrine disruptors in municipal wastewater in Tokyo and their removal during activated sludge treatment. *Water Res.* **2006**, *40*, 3297–3303.
- (16) Stasinakis, A. S.; Gatidou, G.; Mamais, D.; Thomaidis, N. S.; Lekkas, T. D. Occurrence and fate of endocrine disruptors in Greek sewage treatment plants. *Water Res.* **2008**, *42*, 1796–1804.
- (17) Weigel, S.; Kuhlmann, J.; Hühnerfuss, H. Drugs and personal care products as ubiquitous pollutants: occurrence and distribution of clofibric acid, caffeine and DEET in the North Sea. *Sci. Total Environ.* **2002**, *295*, 131–141.
- (18) Mimeault, C.; Woodhouse, A. J.; Miao, X.-S.; Metcalfe, C. D.; Moon, T. W.; Trudeau, V. L. The human lipid regulator, gemfibrozil bioconcentrates and reduces testosterone in the goldfish, *Carassius auratus*. *Aquat. Toxicol.* **2005**, *73*, 44–54.
- (19) Schwaiger, J.; Ferling, H.; Mallow, U.; Wintermayr, H.; Negele, R. D. Toxic effects of the non-steroidal anti-inflammatory drug diclofenac. *Aquat. Toxicol.* **2004**, *68*, 141–150.

Chapter 1

- (20) Triebkorn, R.; Casper, H.; Heyd, A.; Eikemper, R.; Köhler, H.-R.; Schwaiger, J. Toxic effects of the non-steroidal anti-inflammatory drug diclofenac. *Aquat. Toxicol.* **2004**, *68*, 151–166.
- (21) Lindberg, R. H.; Wennberg, P.; Johansson, M. I.; Tysklind, M.; Andersson, B. A. V. Screening of Human Antibiotic Substances and Determination of Weekly Mass Flows in Five Sewage Treatment Plants in Sweden. *Environ. Sci. Technol.* **2005**, *39*, 3421–3429.
- (22) Golet, E. M.; Alder, A. C.; Giger, W. Environmental Exposure and Risk Assessment of Fluoroquinolone Antibacterial Agents in Wastewater and River Water of the Glatt Valley Watershed, Switzerland. *Environ. Sci. Technol.* **2002**, *36*, 3645–3651.
- (23) Kim, S.; Eichhorn, P.; Jensen, J. N.; Weber, A. S.; Aga, D. S. Removal of Antibiotics in Wastewater: Effect of Hydraulic and Solid Retention Times on the Fate of Tetracycline in the Activated Sludge Process. *Environ. Sci. Technol.* **2005**, *39*, 5816–5823.
- (24) Li, B.; Zhang, T. Biodegradation and Adsorption of Antibiotics in the Activated Sludge Process. *Environ. Sci. Technol.* **2010**, *44*, 3468–3473.
- (25) Hughes, S. R.; Kay, P.; Brown, L. E. Global Synthesis and Critical Evaluation of Pharmaceutical Data Sets Collected from River Systems. *Environ. Sci. Technol.* **2013**, *47*, 661–677.
- (26) Crane, M.; Watts, C.; Boucard, T. Chronic aquatic environmental risks from exposure to human pharmaceuticals. *Sci. Total Environ.* **2006**, *367*, 23–41.
- (27) Kidd, K. A.; Blanchfield, P. J.; Mills, K. H.; Palace, V. P.; Evans, R. E.; Lazorchak, J. M.; Flick, R. W. Collapse of a fish population after exposure to a synthetic estrogen. *Proc. Natl. Acad. Sci.* **2007**, *104*, 8897–8901.
- (28) Tyler, C. R.; Jobling, S. Roach, Sex, and Gender-Bending Chemicals: The Feminization of Wild Fish in English Rivers. *Bioscience* **2008**, *58*, 1051.
- (29) Jobling, S.; Nolan, M.; Tyler, C. R.; Brighty, G.; Sumpter, J. P. Widespread Sexual Disruption in Wild Fish. *Environ. Sci. Technol.* **1998**, *32*, 2498–2506.
- (30) Pickering, A. D.; Sumpter, J. P. Peer Reviewed: Comprehending endocrine disruptors in aquatic environments. *Environ. Sci. Technol.* **2003**, *37*, 331A – 336A.
- (31) Fent, K.; Weston, A. A.; Caminada, D. Ecotoxicology of human pharmaceuticals. *Aquat. Toxicol.* **2006**, *76*, 122–159.
- (32) Escher, B.; Leusch, F. *Bioanalytical tools in water quality assessment*; IWA publishing, 2012.
- (33) Escher, B. I.; Allinson, M.; Altenburger, R.; Bain, P. A.; Balaguer, P.; Busch, W.; Crago, J.; Denslow, N. D.; Dopp, E.; Hilscherova, K.; et al. Benchmarking Organic Micropollutants in Wastewater, Recycled Water and Drinking Water with In Vitro Bioassays. *Environ. Sci. Technol.* **2014**, *48*, 1940–1956.
- (34) Escher, B. I.; Bramaz, N.; Quayle, P.; Rutishauser, S.; Vermeirssen, E. L. M. Monitoring of the ecotoxicological hazard potential by polar organic micropollutants in sewage treatment plants and surface waters using a mode-of-action based test battery. *J. Environ. Monit.* **2008**, *10*, 622.
- (35) Macova, M.; Toze, S.; Hodggers, L.; Mueller, J. F.; Bartkow, M.; Escher, B. I. Bioanalytical tools for the evaluation of organic micropollutants during sewage treatment, water recycling and drinking water generation. *Water Res.* **2011**.
- (36) Escher, B. I.; Bramaz, N.; Ort, C. JEM Spotlight: Monitoring the treatment efficiency of a full scale ozonation on a sewage treatment plant with a mode-of-action based test battery. *J. Environ. Monit.* **2009**, *11*, 1836.
- (37) Tang, J. Y. M.; Glenn, E.; Thoen, H.; Escher, B. I. In vitro bioassay for reactive toxicity towards proteins implemented for water quality monitoring. *J. Environ. Monit.* **2012**, *14*, 1073.
- (38) Jia, A.; Escher, B. I.; Leusch, F. D. L.; Tang, J. Y. M.; Prochazka, E.; Dong, B.; Snyder, E. M.; Snyder, S. A. In vitro bioassays to evaluate complex chemical mixtures in recycled water. *Water Res.* **2015**, *80*, 1–11.
- (39) Maurer, M.; Pronk, W.; Larsen, T. A. Treatment processes for source-separated urine. *Water Res.* **2006**, *40*, 3151–3166.
- (40) von Sonntag, C.; von Gunten, U. *Chemistry of ozone in water and wastewater treatment: From basic principles to applications*; IWA publishing, 2012.
- (41) von Gunten, U. Ozonation of drinking water: part I. Oxidation kinetics and product formation. *Water Res.* **2003**, *37*, 1443–1467.
- (42) Hoigne, J.; Bader, H. Ozonation of Water: Role of Hydroxyl Radicals as Oxidizing

- Intermediates. *Science (80-.)*. **1975**, *190*, 782–784.
- (43) Crittenden, J. C.; Trussell, R. R.; Hand, D. W.; Howe, K. J.; Tchobanoglous, G. *MWH's Water Treatment: Principles and Design*; John Wiley & Sons, 2012.
- (44) von Gunten, U. Ozonation of drinking water: part II. Disinfection and by-product formation in presence of bromide, iodide or chlorine. *Water Res.* **2003**, *37*, 1469–1487.
- (45) Gerrity, D.; Gamage, S.; Jones, D.; Korshin, G. V.; Lee, Y.; Pisarenko, A.; Trenholm, R. A.; von Gunten, U.; Wert, E. C.; Snyder, S. A. Development of surrogate correlation models to predict trace organic contaminant oxidation and microbial inactivation during ozonation. *Water Res.* **2012**, *46*, 6257–6272.
- (46) Lee, Y.; Gerrity, D.; Lee, M.; Bogeat, A. E.; Salhi, E.; Gamage, S.; Trenholm, R. A.; Wert, E. C.; Snyder, S. A.; von Gunten, U. Prediction of micropollutant elimination during ozonation of municipal wastewater effluents: use of kinetic and water specific information. *Environ. Sci. Technol.* **2013**, *47*, 5872–5881.
- (47) Lee, Y.; Kovalova, L.; McArdell, C. S.; von Gunten, U. Prediction of micropollutant elimination during ozonation of a hospital wastewater effluent. *Water Res.* **2014**, *64*, 134–148.
- (48) Rosal, R.; Rodríguez, A.; Perdigón-Melón, J. A.; Petre, A.; García-Calvo, E.; Gómez, M. J.; Agüera, A.; Fernández-Alba, A. R. Occurrence of emerging pollutants in urban wastewater and their removal through biological treatment followed by ozonation. *Water Res.* **2010**, *44*, 578–588.
- (49) Wert, E. C.; Rosario-Ortiz, F. L.; Snyder, S. A. Effect of ozone exposure on the oxidation of trace organic contaminants in wastewater. *Water Res.* **2009**, *43*, 1005–1014.
- (50) Gerrity, D.; Gamage, S.; Holady, J. C.; Mawhinney, D. B.; Quiñones, O.; Trenholm, R. A.; Snyder, S. A. Pilot-scale evaluation of ozone and biological activated carbon for trace organic contaminant mitigation and disinfection. *Water Res.* **2011**, *45*, 2155–2165.
- (51) Huber, M. M.; Göbel, A.; Joss, A.; Hermann, N.; Löffler, D.; McArdell, C. S.; Ried, A.; Siegrist, H.; Ternes, T. A.; von Gunten, U. Oxidation of Pharmaceuticals during Ozonation of Municipal Wastewater Effluents: A Pilot Study. *Environ. Sci. Technol.* **2005**, *39*, 4290–4299.
- (52) Nakada, N.; Shinohara, H.; Murata, A.; Kiri, K.; Managaki, S.; Sato, N.; Takada, H. Removal of selected pharmaceuticals and personal care products (PPCPs) and endocrine-disrupting chemicals (EDCs) during sand filtration and ozonation at a municipal sewage treatment plant. *Water Res.* **2007**, *41*, 4373–4382.
- (53) Ternes, T. A.; Stüber, J.; Herrmann, N.; McDowell, D.; Ried, A.; Kampmann, M.; Teiser, B. Ozonation: a tool for removal of pharmaceuticals, contrast media and musk fragrances from wastewater? *Water Res.* **2003**, *37*, 1976–1982.
- (54) Hollender, J.; Zimmermann, S. G.; Koepke, S.; Krauss, M.; McArdell, C. S.; Ort, C.; Singer, H.; von Gunten, U.; Siegrist, H. Elimination of Organic Micropollutants in a Municipal Wastewater Treatment Plant Upgraded with a Full-Scale Post-Ozonation Followed by Sand Filtration. *Environ. Sci. Technol.* **2009**, *43*, 7862–7869.
- (55) Reungoat, J.; Escher, B. I. I.; Macova, M.; Argaud, F. X. X.; Gernjak, W.; Keller, J. Ozonation and biological activated carbon filtration of wastewater treatment plant effluents. *Water Res.* **2012**, *46*, 863–872.
- (56) Zimmermann, S. G.; Wittenwiler, M.; Hollender, J.; Krauss, M.; Ort, C.; Siegrist, H.; von Gunten, U. Kinetic assessment and modeling of an ozonation step for full-scale municipal wastewater treatment: Micropollutant oxidation, by-product formation and disinfection. *Water Res.* **2011**, *45*, 605–617.
- (57) Lee, Y.; Escher, B. I.; von Gunten, U. Efficient Removal of Estrogenic Activity during Oxidative Treatment of Waters Containing Steroid Estrogens. *Environ. Sci. Technol.* **2008**, *42*, 6333–6339.
- (58) Huber, M. M.; Ternes, T. A.; von Gunten, U. Removal of Estrogenic Activity and Formation of Oxidation Products during Ozonation of 17 α -Ethinylestradiol. *Environ. Sci. Technol.* **2004**, *38*, 5177–5186.
- (59) Dodd, M. C.; Kohler, H.-P. E.; von Gunten, U. Oxidation of Antibacterial Compounds by Ozone and Hydroxyl Radical: Elimination of Biological Activity during Aqueous Ozonation Processes. *Environ. Sci. Technol.* **2009**, *43*, 2498–2504.

Chapter 1

- (60) Lange, F.; Cornelissen, S.; Kubac, D.; Sein, M. M.; von Sonntag, J.; Hannich, C. B.; Golloch, A.; Heipieper, H. J.; Möder, M.; von Sonntag, C. Degradation of macrolide antibiotics by ozone: a mechanistic case study with clarithromycin. *Chemosphere* **2006**, *65*, 17–23.
- (61) Rutishauser, B. V.; Pesonen, M.; Escher, B. I.; Ackermann, G. E.; Aerni, H.-R.; Suter, M. J.-F.; Eggen, R. I. L. Comparative analysis of estrogenic activity in sewage treatment plant effluents involving three in vitro assays and chemical analysis of steroids. *Environ. Toxicol. Chem.* **2004**, *23*, 857.
- (62) Mestankova, H.; Schirmer, K.; Escher, B. I.; von Gunten, U.; Canonica, S. Removal of the antiviral agent oseltamivir and its biological activity by oxidative processes. *Environ. Pollut.* **2012**, *161*, 30–35.
- (63) Dodd, M. C.; Rentsch, D.; Singer, H. P.; Kohler, H.-P. E.; von Gunten, U. Transformation of β -Lactam Antibacterial Agents during Aqueous Ozonation: Reaction Pathways and Quantitative Bioassay of Biologically-Active Oxidation Products. *Environ. Sci. Technol.* **2010**, *44*, 5940–5948.
- (64) Magdeburg, A.; Stalter, D.; Oehlmann, J. Whole effluent toxicity assessment at a wastewater treatment plant upgraded with a full-scale post-ozonation using aquatic key species. *Chemosphere* **2012**, *88*, 1008–1014.
- (65) Petala, M.; Kokokiris, L.; Samaras, P.; Papadopoulos, A.; Zouboulis, A. Toxicological and ecotoxic impact of secondary and tertiary treated sewage effluents. *Water Res.* **2009**, *43*, 5063–5074.
- (66) Stalter, D.; Magdeburg, A.; Weil, M.; Knacker, T.; Oehlmann, J. Toxication or detoxication? In vivo toxicity assessment of ozonation as advanced wastewater treatment with the rainbow trout. *Water Res.* **2010**, *44*, 439–448.
- (67) Reungoat, J.; Macova, M.; Escher, B. I.; Carswell, S.; Mueller, J. F.; Keller, J. Removal of micropollutants and reduction of biological activity in a full scale reclamation plant using ozonation and activated carbon filtration. *Water Res.* **2010**, *44*, 625–637.
- (68) Stalter, D.; Magdeburg, A.; Wagner, M.; Oehlmann, J. Ozonation and activated carbon treatment of sewage effluents: Removal of endocrine activity and cytotoxicity. *Water Res.* **2011**, *45*, 1015–1024.
- (69) Macova, M.; Escher, B. I.; Reungoat, J.; Carswell, S.; Chue, K. L.; Keller, J.; Mueller, J. F. Monitoring the biological activity of micropollutants during advanced wastewater treatment with ozonation and activated carbon filtration. *Water Res.* **2010**, *44*, 477–492.
- (70) Maletz, S.; Floehr, T.; Beier, S.; Klümper, C.; Brouwer, A.; Behnisch, P.; Higley, E.; Giesy, J. P.; Hecker, M.; Gebhardt, W.; et al. In vitro characterization of the effectiveness of enhanced sewage treatment processes to eliminate endocrine activity of hospital effluents. *Water Res.* **2013**, *47*, 1545–1557.
- (71) Yan, Z.; Zhang, Y.; Yuan, H.; Tian, Z.; Yang, M. Fish larval deformity caused by aldehydes and unknown byproducts in ozonated effluents from municipal wastewater treatment systems. *Water Res.* **2014**, *66*, 423–429.
- (72) Cao, N.; Yang, M.; Zhang, Y.; Hu, J.; Ike, M.; Hirotsuji, J.; Matsui, H.; Inoue, D.; Sei, K. Evaluation of wastewater reclamation technologies based on in vitro and in vivo bioassays. *Sci. Total Environ.* **2009**, *407*, 1588–1597.
- (73) Magdeburg, A.; Stalter, D.; Schlüsener, M.; Ternes, T.; Oehlmann, J. Evaluating the efficiency of advanced wastewater treatment: Target analysis of organic contaminants and (geno-)toxicity assessment tell a different story. *Water Res.* **2014**, *50*, 35–47.
- (74) Stalter, D.; Magdeburg, A.; Oehlmann, J. Comparative toxicity assessment of ozone and activated carbon treated sewage effluents using an in vivo test battery. *Water Res.* **2010**, *44*, 2610–2620.
- (75) Hoigné, J.; Bader, H. Rate constants of reactions of ozone with organic and inorganic compounds in water—I. Non-dissociating organic compounds. *Water Res.* **1983**, *17*, 173–183.
- (76) Hoigné, J.; Bader, H. Rate constants of reactions of ozone with organic and inorganic compounds in water—II. Dissociating organic compounds. *Water Res.* **1983**, *17*, 185–194.
- (77) Yao, C. D.; Haag, W. Rate constants for direct reactions of ozone with several drinking water contaminants. *Water Res.* **1991**, *25*, 761–773.

- (78) Lee, Y.; von Gunten, U. Quantitative structure–activity relationships (QSARs) for the transformation of organic micropollutants during oxidative water treatment. *Water Res.* **2012**, *46*, 6177–6195.
- (79) Sudhakaran, S.; Amy, G. L. QSAR models for oxidation of organic micropollutants in water based on ozone and hydroxyl radical rate constants and their chemical classification. *Water Res.* **2013**, *47*, 1111–1122.
- (80) Naumov, S.; von Sonntag, C. Quantum Chemical Studies on the Formation of Ozone Adducts to Aromatic Compounds in Aqueous Solution. *Ozone Sci. Eng.* **2010**, *32*, 61–65.
- (81) Fukui, K. The Role of Frontier Orbitals in Chemical Reactions (Nobel Lecture). *Angew. Chemie Int. Ed. English* **1982**, *21*, 801–809.
- (82) Criegee, R. Mechanism of Ozonolysis. *Angew. Chemie Int. Ed. English* **1975**, *14*, 745–752.
- (83) Dowideit, P.; von Sonntag, C. Reaction of Ozone with Ethene and Its Methyl- and Chlorine-Substituted Derivatives in Aqueous Solution. *Environ. Sci. Technol.* **1998**, *32*, 1112–1119.
- (84) McDowell, D. C.; Huber, M. M.; Wagner, M.; von Gunten, U.; Ternes, T. A. Ozonation of Carbamazepine in Drinking Water: Identification and Kinetic Study of Major Oxidation Products. *Environ. Sci. Technol.* **2005**, *39*, 8014–8022.
- (85) Barron, E.; Deborde, M.; Rabouan, S.; Mazellier, P.; Legube, B. Kinetic and mechanistic investigations of progesterone reaction with ozone. *Water Res.* **2006**, *40*, 2181–2189.
- (86) Flyunt, R.; Theruvathu, J. A.; Leitzke, A.; von Sonntag, C. The reactions of thymine and thymidine with ozone. *J. Chem. Soc. Perkin Trans. 2* **2002**, 1572–1582.
- (87) Mvula, E.; Sonntag, C. von. Ozonolysis of phenols in aqueous solution. *Org. Biomol. Chem.* **2003**, *1*, 1749–1756.
- (88) Ramseier, M. K.; Gunten, U. von. Mechanisms of Phenol Ozonation—Kinetics of Formation of Primary and Secondary Reaction Products. *Ozone Sci. Eng.* **2009**, *31*, 201–215.
- (89) Mvula, E.; Naumov, S.; von Sonntag, C. Ozonolysis of Lignin Models in Aqueous Solution: Anisole, 1,2-Dimethoxybenzene, 1,4-Dimethoxybenzene, and 1,3,5-Trimethoxybenzene. *Environ. Sci. Technol.* **2009**, *43*, 6275–6282.
- (90) Merényi, G.; Lind, J.; Naumov, S.; Sonntag, C. von. Reaction of Ozone with Hydrogen Peroxide (Peroxone Process): A Revision of Current Mechanistic Concepts Based on Thermokinetic and Quantum-Chemical Considerations. *Environ. Sci. Technol.* **2010**, *44*, 3505–3507.
- (91) Pereira, R. de O.; de Alda, M. L.; Joglar, J.; Daniel, L. A.; Barceló, D. Identification of new ozonation disinfection byproducts of 17 β -estradiol and estrone in water. *Chemosphere* **2011**, *84*, 1535–1541.
- (92) Benner, J.; Ternes, T. A. Ozonation of Propranolol: Formation of Oxidation Products. *Environ. Sci. Technol.* **2009**, *43*, 5086–5093.
- (93) Benner, J.; Ternes, T. A. Ozonation of Metoprolol: Elucidation of Oxidation Pathways and Major Oxidation Products. *Environ. Sci. Technol.* **2009**, *43*, 5472–5480.
- (94) Deborde, M.; Rabouan, S.; Mazellier, P.; Duguet, J.-P.; Legube, B. Oxidation of bisphenol A by ozone in aqueous solution. *Water Res.* **2008**, *42*, 4299–4308.
- (95) Mawhinney, D. B.; Vanderford, B. J.; Snyder, S. A. Transformation of 1 H -Benzotriazole by Ozone in Aqueous Solution. *Environ. Sci. Technol.* **2012**, *46*, 7102–7111.
- (96) Chen, X.; Richard, J.; Liu, Y.; Dopp, E.; Tuerk, J.; Bester, K. Ozonation products of triclosan in advanced wastewater treatment. *Water Res.* **2012**, *46*, 2247–2256.
- (97) Tekle-Röttering, A.; Jewell, K. S.; Reisz, E.; Lutze, H. V.; Ternes, T. A.; Schmidt, W.; Schmidt, T. C. Ozonation of piperidine, piperazine and morpholine: Kinetics, stoichiometry, product formation and mechanistic considerations. *Water Res.* **2016**, *88*, 960–971.
- (98) Radjenović, J.; Godehardt, M.; Hein, A.; Farré, M.; Jekel, M.; Barceló, D. Evidencing Generation of Persistent Ozonation Products of Antibiotics Roxithromycin and Trimethoprim. *Environ. Sci. Technol.* **2009**, *43*, 6808–6815.
- (99) Zimmermann, S. G.; Schmukat, A.; Schulz, M.; Benner, J.; Gunten, U. von; Ternes, T. A. Kinetic and Mechanistic Investigations of the Oxidation of Tramadol by Ferrate and Ozone. *Environ. Sci. Technol.* **2012**, *46*, 876–884.
- (100) von Sonntag, C. *Free-radical-induced DNA damage and its repair*; Springer, 2006.
- (101) Borowska, E.; Bourgin, M.; Hollender, J.; Kienle, C.; McArdell, C. S.; von Gunten, U.

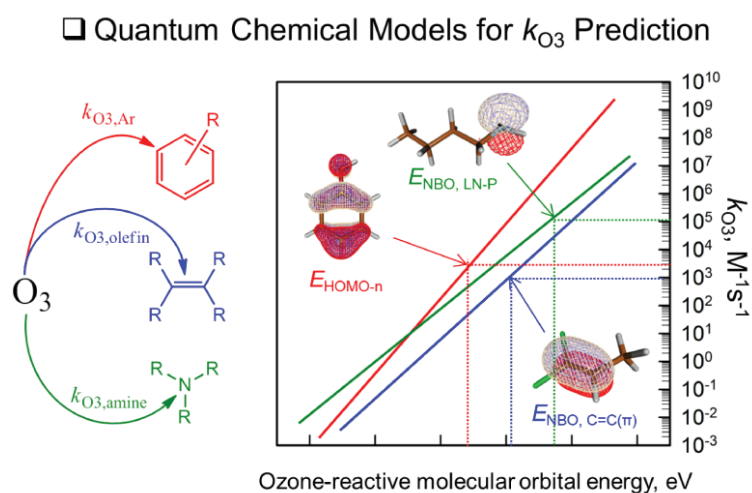
Chapter 1

Oxidation of cetirizine, fexofenadine and hydrochlorothiazide during ozonation: Kinetics and transformation products. *Water Res.* **2016**.

- (102) Muñoz, F.; Mvula, E.; Braslavsky, S. E.; von Sonntag, C. Singlet dioxygen formation in ozone reactions in aqueous solution. *J. Chem. Soc. Perkin Trans. 2* **2001**, 1109–1116.
- (103) Pryor, W. A.; Giamalva, D. H.; Church, D. F. Kinetics of ozonation. 2. Amino acids and model compounds in water and comparisons to rates in nonpolar solvents. *J. Am. Chem. Soc.* **1984**, *106*, 7094–7100.

Chapter 2.

Development of prediction models for the reactivity of organic compounds with ozone in aqueous solution by quantum chemical computations: Role of delocalized and localized molecular orbitals



Lee, M.; Zimmermann-Steffens, S. G.; Arey, J. S.; Fenner, K.; von Gunten, U.. *Environmental Science Technology* **2015**, 49, 9925–9935.

This chapter was prepared by Lee, M. (overall contribution >80%) with scientific advices and editorial comments by all coauthors, particularly under a joint supervision of Arey, J. S., Fenner, K., and von Gunten, U.

Abstract

Second-order rate constants (k_{O_3}) for the reaction of ozone with micropollutants are essential parameters to assess their elimination efficiency during ozonation in water and wastewater treatment. k_{O_3} prediction models were developed for aromatic compounds, olefins, and amines by quantum chemical molecular orbital computations employing *ab initio* Hartree-Fock (HF) and Density Functional Theory (B3LYP) methods. k_{O_3} for aromatic compounds correlated well with the energy of a delocalized molecular orbital first appearing on an aromatic ring, i.e., the highest occupied molecular orbital (HOMO), or the HOMO- n ($n \geq 0$) when the HOMO is not located on the aromatic ring (number (N) of compounds tested=112, correlation coefficient $R^2=0.82-1.00$). k_{O_3} for olefins and amines correlated well with the energy of a localized molecular orbital, i.e., the natural bond orbital (NBO) energy of the carbon-carbon π bond of olefins ($N=45$, $R^2=0.82-0.85$) and the NBO energy of the nitrogen lone-pair electrons of amines ($N=59$, $R^2=0.81-0.83$), respectively. Considering the performance of the k_{O_3} prediction model and computational costs, the HF/6-31G method is recommended for all aromatic groups and olefins investigated herein, whereas HF/MIDI!, HF/6-31G*, or HF/6-311++G** is recommended for amines. Based on their mean absolute errors, the above models could predict k_{O_3} within a factor of 4, on average, relative to the experimentally determined values. Overall, good correlations were also observed ($R^2=0.77-0.96$) between k_{O_3} predictions by quantum molecular orbital descriptors in this study and by Hammett (σ) and Taft (σ^*) constants from previously developed quantitative structure-activity relationship (QSAR) models. Hence, the quantum molecular orbital descriptors are an alternative to σ - and σ^* -values in QSAR applications and can also be utilized to estimate unknown σ - or σ^* -values.

2.1. Introduction

Ozonation is used widely in drinking water treatment for disinfection and oxidation of problematic substances (e.g., taste and odor compounds) or a combination thereof.¹⁻³ In recent years, ozonation has been successfully applied in advanced wastewater treatment for mitigation of the discharge of micropollutants such as pharmaceuticals, personal care products, and pesticides from municipal wastewater effluents to the aquatic environment.³⁻⁶ The benefits of ozone (O_3) are related to its electrophilic characteristics, selectively oxidizing electron-rich moieties such as activated aromatic compounds, olefins, free amines, and reduced sulfur-groups present in organic compounds.³ Moreover, substances without such electron-rich moieties can be oxidized to some extent by a secondary oxidant, the hydroxyl radical ($\bullet OH$), formed during ozonation.^{7,8} The elimination efficiency of a compound (P) during ozonation can be estimated by Eq. 2.1 provided that the two product terms on the right-hand side are known.¹

$$-\ln\left(\frac{[P]}{[P]_0}\right) = k_{O_3} \int [O_3] dt + k_{\bullet OH} \int [\bullet OH] dt \quad (2.1)$$

k_{O_3} and $k_{\bullet OH}$ are the second-order rate constants for the reaction of P with ozone and hydroxyl radicals, respectively. $\int [O_3] dt$ and $\int [\bullet OH] dt$ are the exposures to ozone and hydroxyl radical, respectively. The two oxidant exposure terms vary depending on the water matrix, thus necessitating an empirical determination. In contrast, the second-order rate constants (k_{O_3} and $k_{\bullet OH}$) are physico-chemical constants universally applicable for a given compound P regardless of the water matrix.

Although many rate constants are available for both oxidants (a few hundred values for k_{O_3} and a few thousand values for $k_{\bullet OH}$, respectively),^{3,7,9-14} rate constants for emerging compounds are often missing. Since the number of newly detected micropollutants is increasing rapidly,¹⁵ the experimental determination of rate constants becomes a time and cost factor for the assessment of ozonation efficiency.

Various approaches have been suggested for predicting k_{O_3} -values. A univariate quantitative structure-activity relationship (QSAR) model was reported in which the logarithm of k_{O_3} correlates well with Hammett constants (σ) for aromatic compounds (the number (N) of compounds=64, correlation coefficient (R^2)>0.85) and Taft constants (σ^*) for olefins ($N=48$, $R^2=0.86$) and amines ($N=54$, $R^2=0.86$).¹⁶ However, this approach is limited to compounds with substituents for which Hammett and Taft constants are available. In another study, a multivariate QSAR model ($N=27$, $R^2=0.83$) was developed in which several descriptors such as double bond equivalence and ionization potential were incorporated.¹⁷ Moreover, strong correlations were found for benzene derivatives using quantum molecular orbital descriptors such as the highest occupied molecular orbital energy (E_{HOMO}) ($N=7$, $R^2=0.92$) and the Gibbs free energy for adduct formation ($N=8$, $R^2=0.99$).¹⁸ The success of E_{HOMO} is supported by the frontier molecular orbital (FMO) theory,¹⁹ which explains the ozone reaction with substituted benzenes in terms of the interaction between the HOMO of substituted benzenes and the lowest unoccupied molecular orbital (LUMO) of ozone. The thermodynamic approach using the Gibbs free energy of adduct formation gave a better correlation than the FMO approach. However, it necessitates quantum chemical computations for ozone, the aromatic compound, and the intermediate

Chapter 2

ozone adduct. In contrast, the FMO theory only requires the computation of the E_{HOMO} of the aromatic compound, since the LUMO of ozone is constant.

In the present study, published k_{O_3} values of organic compounds such as aromatic compounds, olefins, amines, and sulfur(S)-containing compounds were correlated to quantum molecular orbital descriptors such as the energy of HOMO (E_{HOMO}) and natural bond orbital (NBO) (E_{NBO}), which feature a delocalized and localized molecular orbital, respectively, to establish a semi-empirical k_{O_3} prediction tool. Quantum chemical computations were performed with Hartree-Fock (HF) and density functional theory (DFT) methods to calculate the quantum molecular orbital descriptors. The developed molecular orbital models were compared to previously developed QSAR models.¹⁶

2.2. Materials and methods

2.2.1. Data set

All second-order rate constants (k_{O_3}) for the reaction of ozone with organic compounds in aqueous solution investigated in this study were obtained from literature, and these are summarized in the Supporting Information (SI) in Table S2.1 for aromatic compounds, Table S2.2 for aliphatic olefins, Table S2.3 for nucleic acid constituents and miscellaneous conjugated olefins, Table S2.4 for aliphatic amines, and Table S2.5 for S-containing compounds. pKa-values for compounds with acid-base speciation and Hammett and Taft constants¹⁶ are also included if available.

2.2.2. Computational methodology

Initial three-dimensional (3D) geometries of all electronic structures were obtained by the conformer module of ChemAxon MarvinSketch (Marvin 5.10.0, 2012, <http://www.chemaxon.com>). Quantum chemical computations were conducted by *ab initio* HF and B3LYP methods using the Gaussian 09 program.²⁰ The Pople basis sets (6-31G and 6-311++G**) ²¹⁻²⁵ were used for all atoms except for iodine for which a quasi-relativistic pseudo potential (LANL2DZ)²⁶ was used. In addition, 6-31G*²⁷ and MIDI!²⁸ were employed to further investigate the influence of polarization functions. Computations with 6-31G and 6-311++G** were found to produce the best correlations overall, and these basis sets are used for presentation of the main results. The results obtained with 6-31G* and MIDI! are discussed only briefly. All computations were consistently conducted at the same level of theory, unless stated otherwise. The integral equation formalism polarizable continuum model (IEF-PCM)²⁹ was used for all computations to take into account the solvation effect of water. NBO analysis was carried out using the NBO 3.1 program included in Gaussian 09.³⁰ A conformational search was performed for trimethoxybenzenes using Gabedit 2.4.8. A more detailed description of the methods used is given in Text S2.1 and the overall procedure for k_{O_3} prediction is summarized in Scheme S2.1.

2.2.3. Data treatment and statistical model evaluation

Molecular orbital descriptors such as the $E_{\text{HOMO}-n}$ ($n \geq 0$) or the E_{NBO} of organic compounds, obtained from the quantum chemical computations, were correlated with the logarithm of the corresponding

k_{O_3} -values to develop linear models by least squares regression: $\log(k_{O_3})=a \times (\text{descriptor variable})+y_0$, where a and y_0 are the slope and the y-intercept, respectively. Compounds for which predictions differed from measured k_{O_3} by more than 10-fold were excluded as outliers from the regressions. The performance of the models was then evaluated by R^2 , the mean unsigned error (MUE) in log units, and the root mean square error (RMSE) in log units. Whereas R^2 -values were determined with the outliers excluded, the MUE and the RMSE were determined with the set of compounds used for model development and also the outliers, so that the outliers are considered in the evaluation of the predictive power of the model.

2.3. Results and discussion

2.3.1. Boundary conditions for the quantum chemical computations

Prior to presenting the results, two caveats should be noted. First, the geometries of the investigated compounds, from which the orbital energy was obtained, were confirmed on a local minimum, not on a global minimum. Finding a global minimum of the potential energy surface of the selected compounds (>200) using quantum chemical computations is computationally expensive. Therefore, an initial geometry was obtained as the lowest energy conformer at the molecular mechanics level and was further optimized and confirmed at the stationary point with quantum mechanics. The model development using this geometry, which may or may not be on a global minimum, turned out to be satisfactorily successful for the compounds investigated herein, giving a good performance for predicting k_{O_3} . It may be deduced that an orbital energy corresponding to the ozone-reaction center seems to be mainly influenced by substituents of the target compound directly connected to the reaction center except for a few rotamers (for further discussions see below). For the selected compounds, the conformation of a reaction center and its adjacent molecular structure may be analogous for the local minimum and the global minimum. Thus, a similar orbital energy for a local and a global minimum generally could be obtained. Nonetheless, further improvements of the model performance may be possible by using a global minimum geometry, and future efforts could include the analysis of several conformers as an attempt to estimate the range of errors in predicting a k_{O_3} value.

Second, two different electronic structure methods, HF and DFT-B3LYP, were explored. The two methods differ in that the HF formalism derives a wave function of the system without taking dynamical electron correlation into account, whereas (among other differences) the Kohn-Sham (KS) DFT formalism derives an electron density including electron correlation.³¹ A physical interpretation is available for HF orbitals. According to Koopmans' theorem,³² the negative of the HOMO energy is an approximation of the ionization potential (IP). In contrast, KS orbitals at their first advent were considered as a pure mathematical expression.³³ However, many studies have reported their similarity with the HF orbitals,³⁴⁻³⁶ and KS orbital energies can additionally be used to approximate the IP.³⁷ Although a previous study³⁷ reports a strong correlation ($R^2 \sim 0.99$) between HF and DFT methods for E_{HOMO} -values for 11 substituted benzenes, an extensive comparative study of E_{HOMO} -values for a large number of aromatic compounds with diverse substituents ($N=112$ in the present study) has not been

performed yet. To our knowledge, moreover, the comparison between HF and DFT methods has not been yet investigated for olefins and amines for which the E_{NBO} is to be correlated with the corresponding k_{O_3} -values. With these two caveats addressed above in mind, we present the developed models.

2.3.2. Correlations of molecular orbital energies with second order rate constants

Aromatic compounds. The correlations between logarithms of k_{O_3} -values of aromatic compounds and the corresponding $E_{\text{HOMO}-n}$ ($n \geq 0$) are presented in Figures 2.1 and S2.1. The data were obtained from HF/6-31G computations and three other electronic structure methods, respectively. In certain cases with B3LYP computations, $E_{\text{HOMO}-n}$ is used instead of E_{HOMO} and this is explained further below. Overall, good correlations ($N=112$, $R^2=0.65-0.997$, $\text{MUE}=0.03-0.69$, and $\text{RMSE}=0.03-0.84$) were obtained upon separating the aromatic compounds into five different groups based on their chemical structures: (i) phenol derivatives, (ii) benzene derivatives, (iii) aniline derivatives, (iv) mono- and di-alkoxybenzenes, and (v) trimethoxybenzenes. Additionally, ozone attack on nitrogen-containing aromatic compounds such as azobenzene, pyridine and several triazines were investigated as group (vi) although they are not included into the model (see Text S2.5 for further details). Chemical structures of all aromatic compounds investigated are presented in Figure S2.2, and a summary of the prediction model equations and the corresponding statistical evaluation parameters from four different computation levels is presented in Table 2.1. Alternatively, the correlations for the aromatic compounds could be divided into two groups according to the formal charge (FC): one with the FC of 0, +1, and +2 and the other with the FC of -1 (Figure S2.3). However, the predictive power of this approach was worse than that of the structure-based models. Hence the FC-based model was not pursued further.

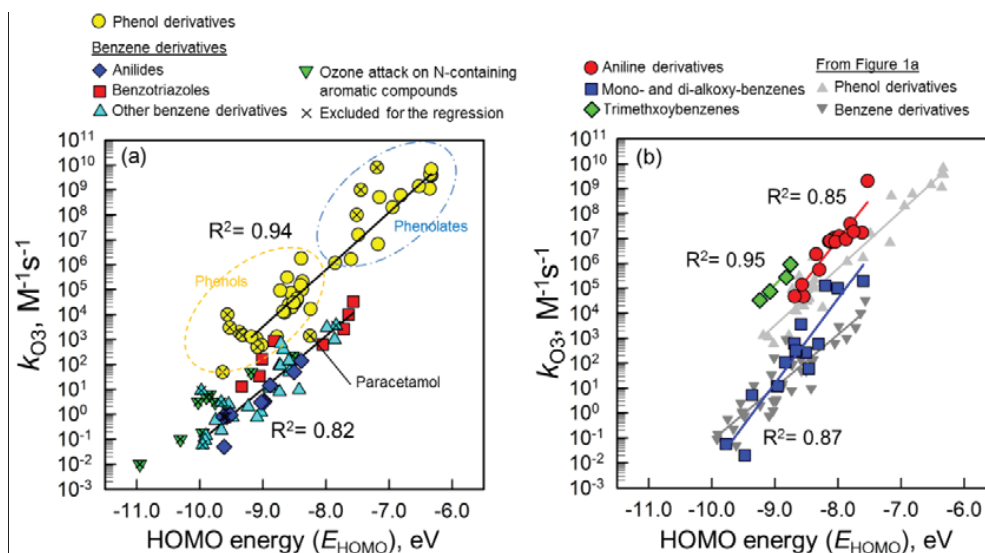


Figure 2.1. Correlations between second-order rate constants (k_{O_3}) for the reaction of aromatic compounds with ozone and the corresponding HOMO energies calculated with HF/6-31G: (a) phenol and benzene derivatives, (b) aniline derivatives, mono- and di-alkoxybenzenes, and trimethoxybenzenes. Crossed symbols indicates compounds excluded from the regression.

i) *Phenol derivatives*: As shown in Figure 2.1a, an excellent correlation ($R^2=0.94$) between $\log k_{O_3}$ and the corresponding E_{HOMO-n} was obtained with HF/6-31G computations for 35 phenolic compounds out of 45 in total (See Table S2.1 for a list of the excluded compounds). A distinctive separation between phenols and phenolates was observed (Figure 2.1a). The model performance remains stable with all electronic structure methods used ($R^2=0.94-0.95$ excluding outliers and $MUE=0.38-0.44$ and $RMSE=0.52-0.61$ both evaluated with 1 outlier (paracetamol) included) (Table 2.1 and Figure S2.1). The MUE value of 0.45 log units indicate that the predicted k_{O_3} values were on average within a factor of 2.8 ($=10^{0.45}$) of the measured k_{O_3} . The reaction with aromatic compounds is initiated by an ozone attack at the ring mainly forming an ozone adduct.³ Based on this mechanism, the HOMO should be located on the ring. Similar to the HOMO of phenol shown in Figure 2.2a, the HOMOs of all other phenol derivatives were located on the aromatic ring regardless of the computation methods (data not shown).

Ten phenolic compounds denoted by a diagonal cross in Figure 2.1a were excluded from the correlation for the following reasons. Nine compounds were excluded from this study because their rate constants were not reported accurately but only approximately, e.g., $k_{O_3} < 2 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$ for 2,3-dichlorophenol¹¹. However, these compounds were presented to show that their approximate values are nonetheless in relative proximity to the confident k_{O_3} values (Figure 2.1a). In addition, paracetamol was excluded from the regression due to increased uncertainty arising from its rotamers. An analysis of the rotamers of the selected aromatic compounds including paracetamol was conducted, for which a detailed discussion is given in Text S2.2. In brief, even with the same initial geometry, different electronic structure methods may lead to differing optimized geometries for compounds with substituents such as hydroxyl, amide, or alkoxy groups (i.e., phenols, anilides, and alkoxybenzenes). This may result in substantially different E_{HOMO} values between different rotamers, leading to inflated uncertainty in prediction of k_{O_3} . Thus, it is recommended to check the optimized geometry of a compound with such a substituent based on the rotational information provided in Text S2.2 prior to applying it to the model. However, this approach does not provide a fool-proof recommendation for the geometry of a compound.

ii) *Benzene derivatives*: A reasonably good correlation ($R^2=0.82$) was observed for 40 benzene derivatives out of 47 in total at the HF/6-31G level (7 compounds excluded due to uncertain k_{O_3} -values, see Table S2.1) (Figure 2.1a). Benzene derivatives are presented as three separate subgroups to show visually a structural diversity of the benzene derivative group: benzotriazoles ($N=8$), anilides ($N=11$), and other benzene derivatives ($N=28$) (Figures 2.1a and S2.1a-c). Models developed by the HF method overall performed better than the B3LYP method (HF: $R^2=0.81-0.82$, $MUE=0.47-0.48$, and $RMSE=0.61-0.62$; B3LYP: $R^2=0.65-0.74$, $MUE=0.57-0.69$, and $RMSE=0.71-0.84$, no outliers). Based on this, the HF method is recommended for k_{O_3} prediction for benzene derivatives.

Chapter 2

Table 2.1. Summary of the developed models for the selected organic compounds.

Compound group	Computation method	Model equation	N^a	R^{2bc}	MUE ^{d,f}	RMSE ^{e,f}
Aromatic compounds						
<i>Phenols</i>						
	HF/6-31G	$\log(k_{03}) = 2.29 \times (E_{\text{HOMO}}) + 24.15$	35	0.94	0.41	0.61
	HF/6-311++G**	$\log(k_{03}) = 2.69 \times (E_{\text{HOMO}}) + 27.37$	35	0.94	0.38	0.52
	B3LYP/6-31G	$\log(k_{03}) = 2.40 \times (E_{\text{HOMO}}) + 18.88$	35	0.95	0.41	0.58
	B3LYP/6-311++G**	$\log(k_{03}) = 2.98 \times (E_{\text{HOMO}}) + 23.08$	35	0.94	0.44	0.60
<i>Anilines</i>						
	HF/6-31G	$\log(k_{03}) = 3.18 \times (E_{\text{HOMO}}) + 32.43$	16	0.85	0.34	0.43
	HF/6-311++G**	$\log(k_{03}) = 3.41 \times (E_{\text{HOMO}}) + 34.60$	16	0.86	0.33	0.43
	B3LYP/6-31G	$\log(k_{03}) = 3.28 \times (E_{\text{HOMO-n}}) + 24.84$	16	0.82	0.38	0.48
	B3LYP/6-311++G**	$\log(k_{03}) = 3.48 \times (E_{\text{HOMO-n}}) + 27.19$	16	0.84	0.38	0.45
<i>Mono- and di-alkoxybenzenes</i>						
	HF/6-31G	$\log(k_{03}) = 3.40 \times (E_{\text{HOMO}}) + 31.77$	17	0.87	0.57	0.69
	HF/6-311++G**	$\log(k_{03}) = 4.17 \times (E_{\text{HOMO}}) + 38.22$	17	0.90	0.54	0.59
	B3LYP/6-31G	$\log(k_{03}) = 3.80 \times (E_{\text{HOMO}}) + 25.78$	17	0.80	0.63	0.84
	B3LYP/6-311++G**	$\log(k_{03}) = 4.65 \times (E_{\text{HOMO}}) + 31.81$	17	0.80	0.65	0.84
<i>Trimethoxybenzenes</i>						
	HF/6-31G	$\log(k_{03}) = 2.77 \times (E_{\text{HOMO}}) + 30.09$	4	0.95	0.11	0.12
	HF/6-311++G**	$\log(k_{03}) = 2.70 \times (E_{\text{HOMO}}) + 29.34$	4	0.98	0.07	0.07
	B3LYP/6-31G	$\log(k_{03}) = 2.34 \times (E_{\text{HOMO}}) + 20.10$	4	0.997	0.03	0.03
	B3LYP/6-311++G**	$\log(k_{03}) = 2.17 \times (E_{\text{HOMO}}) + 19.56$	4	0.99	0.06	0.07
<i>Benzene derivatives</i>						
	HF/6-31G	$\log(k_{03}) = 2.12 \times (E_{\text{HOMO}}) + 20.12$	40	0.82	0.47	0.61
	HF/6-311++G**	$\log(k_{03}) = 2.45 \times (E_{\text{HOMO}}) + 23.13$	40	0.81	0.48	0.62
	B3LYP/6-31G	$\log(k_{03}) = 2.20 \times (E_{\text{HOMO-n}}) + 15.58$	40	0.74	0.57	0.71
	B3LYP/6-311++G**	$\log(k_{03}) = 2.29 \times (E_{\text{HOMO-n}}) + 16.64$	40	0.65	0.69	0.84
<i>Olefins</i>						
	HF/6-31G	$\log(k_{03}) = 1.13 \times (E_{\text{NBO, C-C}\pi}) + 16.57$	45	0.84	0.51	0.68
	HF/6-311++G**	$\log(k_{03}) = 1.32 \times (E_{\text{NBO, C-C}\pi}) + 18.54$	45	0.82	0.53	0.69
	B3LYP/6-31G	$\log(k_{03}) = 1.37 \times (E_{\text{NBO, C-C}\pi}) + 15.14$	45	0.85	0.50	0.66
	B3LYP/6-311++G**	$\log(k_{03}) = 1.67 \times (E_{\text{NBO, C-C}\pi}) + 17.74$	45	0.84	0.51	0.66
<i>Aliphatic amines</i>						
	HF/6-31G	$\log(k_{03}) = 0.77 \times (E_{\text{NBO, LP-N}}) + 14.26$	59	0.76	0.75	1.01
	HF/6-311++G**	$\log(k_{03}) = 0.83 \times (E_{\text{NBO, LP-N}}) + 15.80$	59	0.83	0.59	0.86
	B3LYP/6-31G	$\log(k_{03}) = 0.65 \times (E_{\text{NBO, LP-N}}) + 10.02$	59	0.67	0.86	1.14
	B3LYP/6-311++G**	$\log(k_{03}) = 0.85 \times (E_{\text{NBO, LP-N}}) + 12.35$	59	0.81	0.63	0.90

^aThe number of compounds used for the model development excluding outliers, ^bCoefficient of determination, ^cEvaluated for the number of compounds (N in the fourth column) excluding outliers, ^dMean unsigned error, ^eRoot-mean-square error, ^fOutliers excluded for model development were included in calculating the MUE and the RMSE of the respective model, and the total number of the compounds for MUE and RSMIE determination was 36, 53, and 62 for phenols, olefins, and amines, respectively.

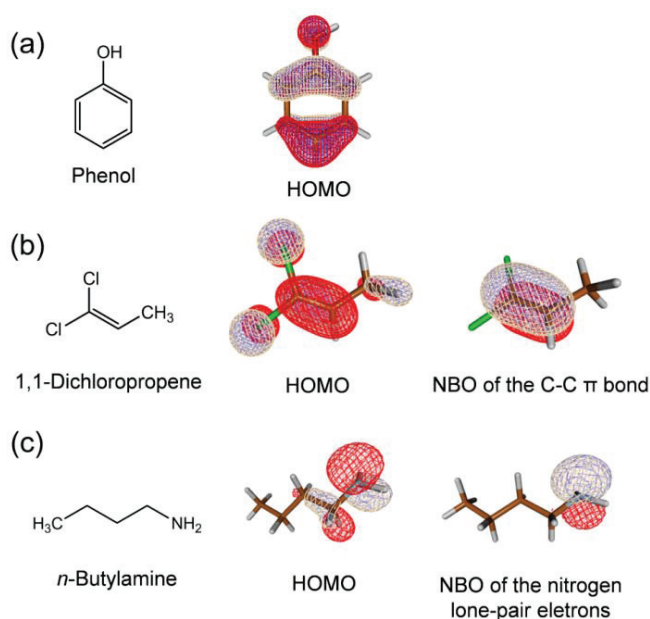


Figure 2.2. Visual representation of molecular orbitals: (a) the HOMO of phenol, (b) comparison between the HOMO and the natural bond orbital (NBO) of the C–C π bond of 1,1-dichloropropene, and (c) comparison between the HOMO and the NBO of the nitrogen lone-pair electrons of *n*-butylamine. Computation method: HF/6-31G. No noticeable difference in the shape and the location of the orbitals was observed between the different computation methods.

The HOMOs of all benzene derivatives were found on the ring for the HF method regardless of the basis set. However, the HOMOs of several compounds such as ibuprofen, benzoate ion, benzene sulfonate, benzaldehyde, and amidotrizoic acid were found on other parts of the molecules for the B3LYP method. Therefore, a HOMO-*n* located on a ring *n* levels lower in energy than the HOMO (*n*=0) was investigated for these compounds for B3LYP computations. For example, the HOMO of the benzoate is located on the carboxylate group and the first appearance of a ring-based molecular orbital was found for HOMO-3 (Figure S2.5). For this molecule, it seems that both HF and B3LYP yield the same molecular orbitals but with different energy orders (Figure S2.5). This phenomenon was observed previously for conjugated bases of various carboxylic acids.³⁸ As shown in Figure S2.6, replacing E_{HOMO} of these compounds by the corresponding ring-located $E_{\text{HOMO}-n}$ improved the B3LYP model performance significantly ($R^2=0.55$ with E_{HOMO} versus $R^2=0.74$ with $E_{\text{HOMO}-n}$). Note that despite this improvement, the performance obtained with B3LYP remained worse than of the HF method.

As described above, rotamers of anilides need to be taken into account. For all electronic structure methods used, the plane of the amide group of alachlor, amidotrizoic acid, iomeprol, iopamidol, iopromide, metolachlor, and propachlor were found to be perpendicular to the aromatic ring plane. The remaining phenylurea herbicides such as isoproturon, linuron, chlorotoluron, and diuron feature a

Chapter 2

geometry close to planar with the dihedral angle ranging from 0 to 38° (except isoproturon, 56°, when calculated at HF/6-311++G**).

iii) *Aniline derivatives*: 16 aniline derivatives were used for the model development and good correlations were obtained with all computation methods ($R^2=0.82-0.86$, MUE=0.33–0.38, and RMSE=0.43–0.48, no outliers) (Figures 2.1b and S2.2d-f). Two different geometries of anilines were investigated: i) planar anilines and ii) pyramidal anilines, and the planar geometry of anilines was selected as representative for the k_{O_3} prediction model since they showed better correlations than the pyramidal anilines. This can be supported by transition state theory: a planar aniline is considered to be closer to the structure of the transition state, thus more relevant to reaction rates than a pyramidal aniline (see Text S2.3 for further discussion).

iv) *Mono- and di-alkoxybenzenes*: Overall good correlations ($R^2=0.80-0.90$, MUE=0.54–0.65, and RMSE=0.59–0.84, no outliers) were obtained for 17 mono- and di-alkoxybenzenes with all computation methods (Figures 2.1b and S2.1d-f). The HF method performed better than the B3LYP method regardless of the basis set (HF: $R^2=0.87-0.90$, MUE=0.54–0.57, and RMSE=0.59–0.69; B3LYP: $R^2=0.80$, MUE=0.63–0.65, and RMSE=0.84). Among two groups of mono- and di-alkoxybenzene rotamers, the group mainly consisting of planar geometry rotamers was chosen as representative for the k_{O_3} prediction model over the group mainly consisting of out-of-plane geometry rotamers. A planar group not only yielded a better model performance than a non-planar one but also their electronic energies were confirmed to be lower (Text S2.2 for more details).

v) *Trimethoxybenzenes*: The trimethoxybenzene group consists of only four compounds: 1,3,5-trimethoxybenzene, 3,4,5-trimethoxytoluene, mono-, and di-protonated trimethoprim. The deprotonated trimethoprim was not included because its measured k_{O_3} seems to correspond to the ozone reaction with the 2,4-diaminopyrimidine (DAP) moiety³⁹ in which a carbon-carbon double bond is presumably attacked by ozone.⁴⁰ Thus, it was grouped with the miscellaneous olefins instead. Excellent correlations were observed for the trimethoxybenzene group with all computation methods ($R^2=0.95-0.997$, MUE=0.03–0.11, and RMSE=0.03–0.12, no outliers). For the selected trimethoxybenzenes all methyl groups are out of plane and they were chosen after analyzing various rotamers of trimethoxybenzenes. It should be noted that these selected geometries are not the lowest in energy among the investigated rotamers (see Text S2.4 for further details) and that the developed model for trimethoxybenzenes is provisional because only four compounds were used.

Olefins. The ozone reaction with olefins, as a 1,3-dipolar cycloaddition, is explained by the FMO theory in which the HOMO of an olefin (dipolarophile) interacts with the LUMO of ozone (dipole).^{41,42} However, inferior correlations ($R^2<0.4$) using the E_{HOMO} of 60 selected olefins with diverse substituents (Figure S2.13 for chemical structures) were observed with all computation methods (Figure S2.14 for the HF/6-31G level). The reaction of an olefin with ozone is driven by the π -like orbital interaction between the HOMO of the olefin and the LUMO of ozone (Scheme S2.2 in the SI), where the responsible molecular orbitals are situated on a C–C double bond and two terminal oxygen atoms, respectively.^{43,44} The HOMOs of all olefins investigated are not only located on a C–C double bond but also spread over substituents (with the exception of ethene). For instance, 1,1-

dichloropropene has 62%, 33%, and 6% of its HOMO spread over the C–C double bond, the two chlorines, and the methyl group, respectively (Figure 2.2b). Concern about the HOMO delocalizing over not only one site (i.e., C–C double bond) relevant to the reaction, but also other sites which are mechanistically irrelevant was previously raised for the cycloaddition reactivity. For these cases, the role of the delocalization of the HOMO has not been clearly resolved yet.⁴¹ Therefore, we sought to find mechanistically plausible molecular orbitals of olefins and correlate their orbital energies with the corresponding k_{O_3} . A natural bond orbital (NBO) analysis was performed using the NBO 3.1 program³⁰, which transforms the delocalized canonical molecular orbital into a localized bonding and lone-pair orbital, expressing a molecule by a formal Lewis structure. As shown in Figure 2.2b, the NBO analysis successfully resulted in a π -type orbital exclusively localized on the C–C double bond of 1,1-dichloropropene.

A correlation of the NBO energies of the C–C π bond of olefins ($E_{NBO, C-C(\pi)}$) with the corresponding k_{O_3} -values was tested for 13 olefins substituted by methyl groups and/or halogens. A near-double orbital occupancy (>1.9) was observed for 13 olefins (upper part Figure 2.3a), indicating a successful computation of the localized π -type orbitals. An excellent correlation ($R^2=0.94$) was observed for 10 olefins at the HF/6-31G level ($R^2=0.95$ – 0.96) (data are not shown for other computation levels, which had similar performance) (Figure 2.3a). In contrast, for the same olefins inferior model performance ($R^2=0.42$) was observed using E_{HOMO} with HF/6-31G (data not shown).

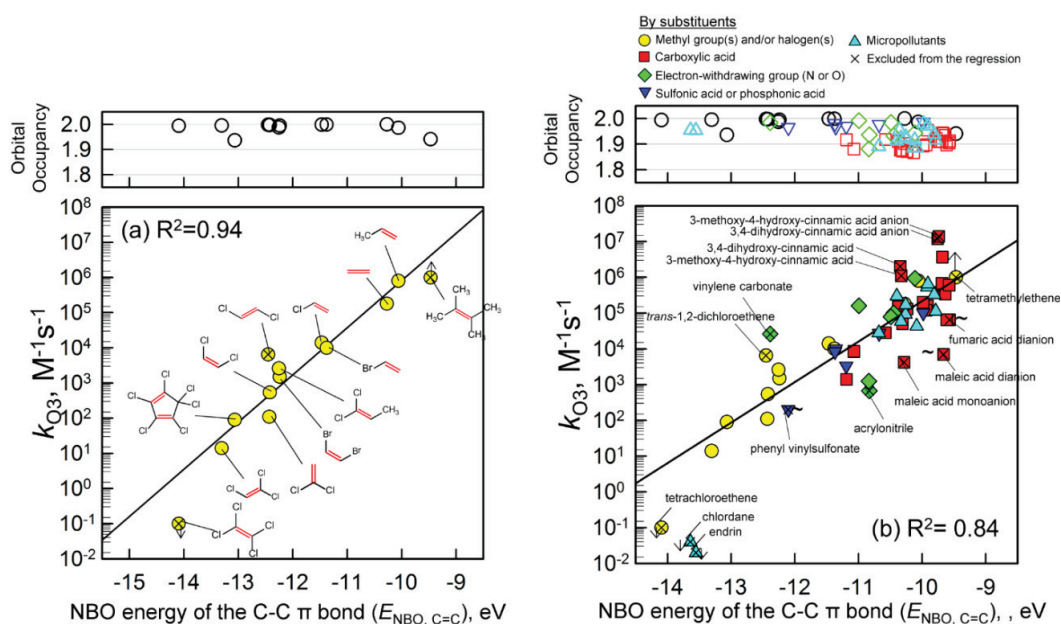


Figure 2.3. Correlations between second-order rate constants (k_{O_3}) for the reaction of selected olefins with ozone and the corresponding NBO energies of the C–C π bond ($E_{NBO, C-C(\pi)}$) calculated with HF/6-31G using the NBO 3.1 package. (a) olefins substituted by methyl group(s) and/or halogen(s); (b) all olefins investigated in this study. Crossed symbols: compounds excluded as outliers. Crossed symbols

Chapter 2

and arrow or a tilde: compounds excluded due to uncertain k_{O_3} . Upper Figures: orbital occupancies of the calculated NBOs of the olefins (full occupancy = 2).

This suggests that the NBO theory, which relates to the localized orbital of the C–C π bond of olefins, is more appropriate for interpreting the reaction of olefins with ozone than the widely accepted FMO theory, which uses the delocalized orbital.^{41,42} The HF orbital energy and the NBO energy differ in that the former is an eigenvalue obtained from diagonalizing the Fock operator and the latter is not. Whereas the HOMO of a molecule may be physically interpreted based on Koopmans' theorem (described above), the NBO remains more conceptual. Ozonation products of methyl- and chlorine-substituted olefins were formed in almost 100 % yield without involving radical species.^{3,45} This indicates that mechanistically an electron transfer from olefins to ozone is not dominant, which is consistent with the poor correlation obtained between k_{O_3} and E_{HOMO} . We propose that as ozone approaches olefins, the HOMO undergoes a localization of electron density on a C–C double bond to form an NBO. This then interacts with ozone leading to the formation of an intermediate ozonide. This interpretation is similar for amines (see below).

Three olefins were excluded from the model development, namely tetramethylethene and tetrachloroethene, which have uncertain experimental k_{O_3} -values, and *trans*-1,2-dichloroethene, which was a regression outlier. The k_{O_3} of *trans*-1,2-dichloroethene, excluded from the regression, was about one order of magnitude higher than k_{O_3} of *cis*-1,2-dichloroethene with nearly the same $E_{NBO, C-C(\pi)}$. *cis*-1,2-Dichloroethene is more stable than *trans*-1,2-dichloroethene, which is known as the *cis*-effect.^{46–48} However, the E_{NBO} apparently does not account for this effect.

We also found a correlation for 8 out of 13 investigated olefins above, for which k_{O_3} -values were determined in carbon tetrachloride (CCl₄) as a solvent (Figure S2.15). The $E_{NBO, C-C(\pi)}$ values were calculated with HF/6-31G employing the IEF-PCM solvation model parameterized for CCl₄ with a dielectric constant (ϵ) set to 2.23. An excellent correlation ($R^2=0.99$) was found with 7 olefins, with *trans*-1,2-dichloroethene excluded as a regression outlier. In particular, tetrachloroethene and tetramethylethene, which were excluded from regressions of aqueous k_{O_3} due to experimental uncertainty (above), were reported with more confident values in CCl₄ solvent^{49,50} and fit well to the regression line.

To test the applicability of the model using $E_{NBO, C-C(\pi)}$ for diverse olefins, another correlation of 60 olefins including the 13 olefins discussed above was investigated. The compounds were divided into five different groups, according to the substituents: (i) methyl group(s) and/or halogen(s), (ii) carboxylic acids, (iii) phosphonic and sulfonic acids, (iv) electron-withdrawing groups composed of oxygen and nitrogen, and (v) miscellaneous micropollutants. For the regressions, we left out 8 outliers and also 7 compounds having uncertain k_{O_3} (denoted by an arrow or a tilde), finding a good correlation ($R^2=0.84$) at the HF/6-31G level (Figure 2.3b). The model performance remained stable with respect to different computational methods ($R^2=0.82$ – 0.85 excluding outliers and MUE= 0.50 – 0.53 and RMSE= 0.66 – 0.69 both evaluated with 8 outliers included) (Table 2.1 and Figure S2.16). Among the 8 excluded outliers were those with an anionic carboxylate group (e.g., maleic acid and fumaric acid) and the ones with electron-withdrawing substituents such as nitrogen or oxygen (e.g., acrylonitrile and vinylene carbonate). Thus, the developed model may be inaccurate for olefins with those substitu-

uents. The bad performance of the correlation for 3,4-dihydroxycinnamic acids and 3-methoxy-4-hydroxycinnamic acids is unclear (see Text S2.6 for further discussions about the outliers).

The reported k_{O_3} values for olefins with two double bonds such as *trans,trans*-muconic acid, 1,4-benzoquinone, β -ionone, sorbic acid, microcystin-LR, *trans-cis*-2,6-nonadienal, and protonated tylosin is the overall k_{O_3} ($k_{O_3, \text{overall}}$) which is the summation of two site-specific k_{O_3} values. These compounds were integrated into the model development as follows. First, the $E_{\text{NBO, C-C}(\pi)}$ correlation model using one overall k_{O_3} was used to predict two site-specific k_{O_3} values ($k_{O_3, \text{pred}}$) for the two C-C double bonds. We then evaluated the fraction contribution of each $k_{O_3, \text{pred}}$ to the overall $k_{O_3, \text{pred}}$. The computed fractions associated with the higher $E_{\text{NBO, C-C}(\pi)}$ calculated averaged over four different computation methods were 0.63, 0.71, 0.75, 0.81, and 0.56 for sorbic acid (both neutral and anionic), β -ionone, *trans-cis*-2,6-nonadienal, microcystin-LR, and tylosin (cationic), respectively. A value of 0.5 is assigned for *trans,trans*-muconic acid and 1,4-benzoquinone because the two $E_{\text{NBO, C-C}(\pi)}$ are equivalent by symmetry. The k_{O_3} prediction model for olefins presented herein was developed using the corrected k_{O_3} values ($k_{O_3, \text{overall}} \times \text{a corresponding fraction}$) for such olefins with two double bonds.

Quantum chemical computations are demanding for macromolecules such as microcystin-LR (MW=995.17) and tylosin (MW=917.10). A structural approximation was conducted for these compounds by modeling the reaction center (i.e., carbon-carbon double bond) and neighboring sites only, and omitting the rest of the molecule, which is assumed to contribute negligibly to the NBO energy. Full molecular structures and their approximated structures were investigated using the 6-31G basis set (Figure S2.13), and similar $E_{\text{NBO, C-C}(\pi)}$ -values were obtained in both cases. Thus the approximated structures were used for the computations using the larger 6-311++G** basis set, and these calculated $E_{\text{NBO, C-C}(\pi)}$ -values fit well into the regression model. This shows that pared approximate structures can be substituted for macromolecules that are expensive to compute. This approach is analogous to approximating micropollutants using reactive substructures, which has been employed previously for the experimental determination of second-order rate constants and elucidation of reaction mechanisms.^{39,51}

Nucleic acid constituents, its derivatives, and miscellaneous olefins (conjugated olefins).

38 Nucleic acid constituents including adenine, cytosine, guanine, thymine, and uracil and their derivatives were analyzed here as a group of conjugated olefins, despite that they are heterocyclic aromatic compounds. These compounds all contain a C-C double bond, which is known to be the dominant reaction site of ozone attack.^{3,52} We also considered several olefinic compounds such as cyclohexenones, diaminopyrimidines, imidazoles, indoles, isoxazoles, quinolone, and indigotrisulfonic acid, which a conjugated C=C bond. These were categorized as miscellaneous olefins. The results and discussion for these compounds are given in the Text S2.7.

Aliphatic amines. Amines are attacked by ozone primarily at the nitrogen lone-pair (LP) electrons to form an ozone adduct, followed by various decay patterns.^{1,3} The nitrogen-occupying $E_{\text{HOMO-n}}$ of 65 amines (ammonia, 26 primary amines, 17 secondary amines, and 21 tertiary amines; see Figure S2.22 for the chemical structures) were found to correlate poorly with the corresponding k_{O_3} (HF/6-31G, $R^2=0.60$) (Figure S2.23). Similar to olefins, this might result from the delocalization of the HOMO(-n)

over the molecule. Thus, the orbital energy of the localized nitrogen lone-pair electrons ($E_{\text{NBO, LP-N}}$), which is mechanistically more relevant than the $E_{\text{HOMO-n}}$, was calculated. For example, whereas for *n*-butylamine 74% of the HOMO was located on the nitrogen, its orbital calculated by the NBO analysis was localized on the nitrogen with near-double occupancy (1.973 out of 2) (Figure 2.2c). Good correlations ($R^2=0.81\text{--}0.83$ excluding outliers and $\text{MUE}=0.59\text{--}0.63$ and $\text{RMSE}=0.86\text{--}0.90$ both evaluated with 4 outliers included) were observed using the $E_{\text{NBO, LP-N}}$ of the amines calculated with HF/6-311++G** and B3LYP/6-311++G** (Figures 2.4 and S2.24(c)).

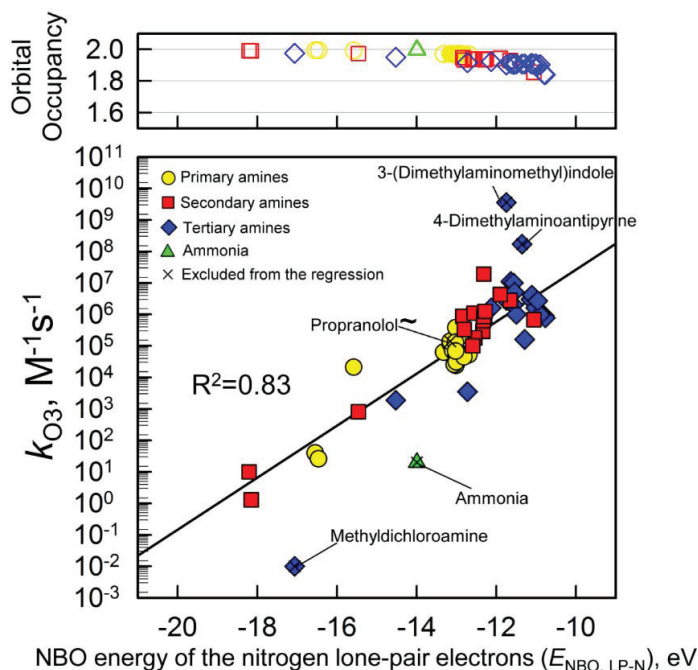


Figure 2.4. Correlations between second-order rate constants (k_{O_3}) for the reaction of the selected aliphatic amines with ozone and the corresponding NBO energies of the nitrogen lone-pair electrons ($E_{\text{NBO, LP-N}}$) calculated with HF/6-311++G** using the NBO 3.1 package. Crossed symbols: compounds excluded as outliers. Crossed symbol and a tilde: propranolol excluded due to uncertain k_{O_3} . Upper Figure: orbital occupancies of the calculated NBOs of the amines (full occupancy = 2).

A somewhat inferior performance was observed when using the 6-31G basis set for both methods ($R^2=0.67\text{--}0.76$ excluding outliers and $\text{MUE}=0.75\text{--}0.86$ and $\text{RMSE}=1.01\text{--}1.14$ both evaluated with 4 outliers included) (Figures S2.24a and b). This seems mainly attributable to the primary amines that fall below the regression line. Ammonia, 3-(dimethylaminomethyl)indole, and 4-dimethylaminoantipyrine were excluded as outliers, for unclear reasons, and methylchloroamine were excluded due to their uncertain k_{O_3} . For azithromycin, DABCO, and EDTA, each of which have two ozone-reactive nitrogens, $0.5 \times k_{\text{O}_3, \text{overall}}$ was used for the model development. A fraction of 0.5 was assigned because the two nitrogen atoms have nearly identical $E_{\text{NBO, LP-N}}$. Approximated structures of macromolecules such as azithromycin, tylosin, and roxythromycin were used for the compu-

tations using 6-311++G** (see Figure S2.22 for the approximated chemical structures), and the corresponding calculated $E_{\text{NBO, LP-N}}$ -values fit well to the regression model.

Sulfur-containing compounds. Only a few k_{O_3} -values are available to date for sulfur (S)-containing compounds.³ 14 neutral and ionic S-containing compounds comprising of thiols, sulfides, disulfides, a sulfoxide, and a sulfinic acid were investigated (Figure S2.25 for the chemical structures), and only a rather qualitative trend was observed between k_{O_3} and the corresponding $E_{\text{HOMO-n}}$ or the corresponding NBO energy of the sulfur lone-pair electron ($E_{\text{NBO, LP-S}}$) (Figures S2.26a-d) (Text S2.8 for further discussions).

2.3.3. Comparison of molecular orbital (MO) models with QSAR models

The MO regression models developed are compared with previously developed QSAR models¹⁶ for aromatic compounds, olefins, and aliphatic amines (Figures 2.5a-o and S28a-t). The performance (R^2) of the MO models calculated with the HF/6-311++G** method (Figures 2.5b, e, h, k, and n) was comparable to or better than the QSAR models (Figures 2.5a, d, g, j, and m) for all groups except the benzene derivatives, for which the QSAR model outperformed the MO model even after leaving out two outliers (1,3,5-trimethoxybenzene and ibuprofen). Reasonable to good correlations between the QSAR descriptors and the corresponding quantum molecular orbital descriptors were obtained ($R^2=0.77-0.91$) (Figures 2.5c, f, i, l, and o). Similar results were obtained using several different computational methods (Figure S2.28).

Several olefinic compounds such as diethyl vinylphosphonate, hexachlorocyclopentadiene, phenyl vinylsulfonate, phosphonic acid, and sorbic acid, which were successfully included in the NBO model (Figure 2.6a as an example at the HF/6-31G level), could not be included in the QSAR model¹⁶ because Taft (σ^*) constants of the substituents were unknown. The good correlations ($R^2=0.88-0.96$) from the insets in Figures S2.28n-p) between the σ^* constant and the $E_{\text{NBO, C-C}(\pi)}$ for olefins can be used to estimate unknown σ^* constants of a substituent, provided that the σ^* -values of the other substituents are known. For example, the total sum of the substituents of diethyl vinylphosphonate ($\Sigma\sigma^*$) can be estimated as 4.68 using the correlation in Figure 2.6b (green dotted line). Because the σ^* -value of three hydrogen substituents is 1.47 (0.49×3), the unknown σ^* -value of the diethyl vinylphosphonate group can be estimated as 3.21. In this manner, the unknown σ^* -values of substituents of diethyl vinylphosphonate (-PO(OEt)₂), phenyl vinylsulfonate (-SO₃-C₆H₅), phosphonic acid (-PO(OH)₂ and -PO(O)(O)²⁻), and sorbic acid (-CH=CH-CH₃) were estimated (see caption of Figure 2.6 for examples of the estimation of unknown σ^* -values at the HF/6-31G level). The estimated average σ^* -values from four different computation levels are given in Table S2.7. This approach is based on the assumption that the sum of the Taft constants of all substituents correlates strongly with the NBO energy of the corresponding substituted olefin. Analogous approaches have been proposed to estimate unknown Hammett and Taft substituent constants based on correlations with theoretical electronic structure descriptors (e.g., atomic charge and electrophilicity) or experimental parameters (e.g., p*K*_a).⁵³⁻⁵⁷

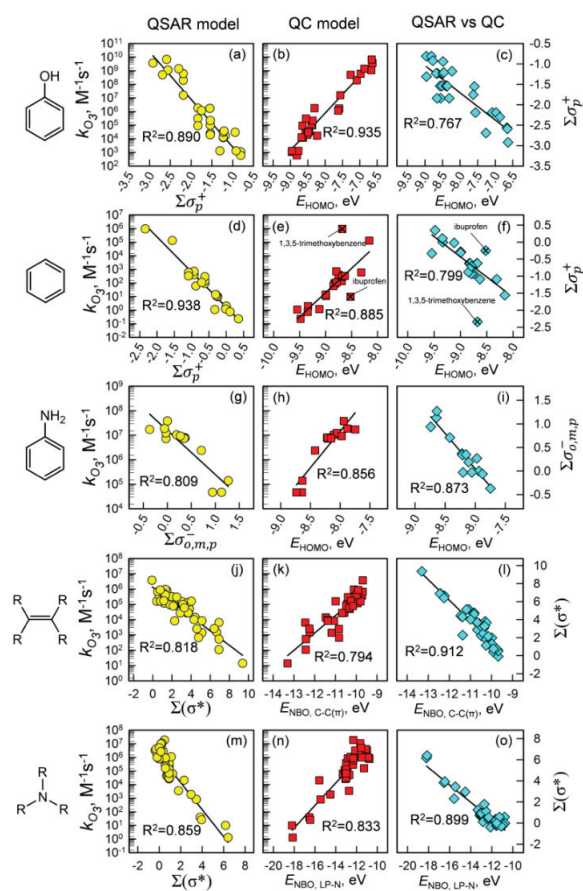


Figure 2.5. Comparison of quantum chemical molecular orbital (MO) computation models with QSAR models.¹⁶ (a-c) Phenol derivatives ($N=28$); (d-f) benzene derivatives ($N=21$); (g-i) aniline derivatives ($N=14$); (j-l) olefins ($N=40$); (m-o) amines ($N=48$). Left column: QSAR models¹⁶; middle column: MO models calculated using HF/6-311++G**; right column: correlations between QSAR descriptor and quantum molecular orbital descriptors.

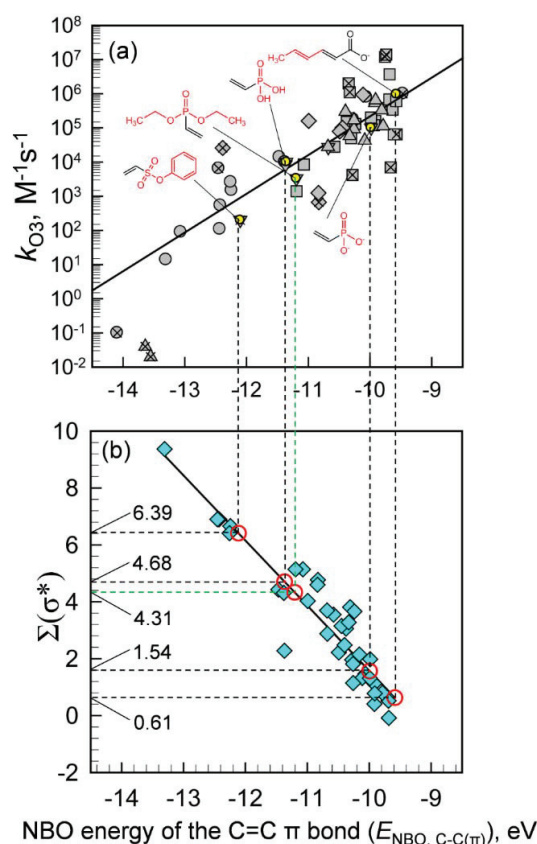


Figure 2.6. Estimation for unknown Taft constants (σ^*) of substituents of olefins using the correlation between the sums of Taft constants ($\Sigma\sigma^*$) and the NBO energies of the C-C π bond ($E_{\text{NBO, C-C}(\pi)}$) of olefins from HF/6-31G. (a) Correlation between the $E_{\text{NBO, C-C}(\pi)}$ of olefins and the corresponding k_{O_3} (from Figure 2.3b) and (b) correlation between $\Sigma\sigma^*$ and the $E_{\text{NBO, C-C}(\pi)}$ (from the inset in Figure S2.25n). Structures in red in Figure 2.6a are the substituents with unknown Taft constants. Examples of the estimation of a Taft constant at the B3LYP/6-31G level: $\sigma^*(-\text{PO}(\text{OEt})_2) = \Sigma\sigma^* - (3 \times \sigma^*(-\text{H})) = 4.68 - (3 \times 0.49) = \underline{3.21}$, $\sigma^*(-\text{SO}_3-\text{C}_6\text{H}_5) = \Sigma\sigma^* - (3 \times \sigma^*(-\text{H})) = 6.39 - (3 \times 0.49) = \underline{4.92}$, $\sigma^*(-\text{CH}=\text{CH}-\text{CH}_3) = \Sigma\sigma^* - (2 \times \sigma^*(-\text{H}) + \sigma^*(-\text{C}(\text{O})\text{O}^-)) = 0.61 - (2 \times 0.49 - 1.06) = \underline{0.69}$, $\sigma^*(-\text{PO}(\text{OH})_2) = \Sigma\sigma^* - (3 \times \sigma^*(-\text{H})) = 4.68 - (3 \times 0.49) = \underline{3.21}$, $\sigma^*(-\text{PO}(\text{O})(\text{O})^2) = \Sigma\sigma^* - (3 \times \sigma^*(-\text{H})) = 1.54 - (3 \times 0.49) = \underline{0.07}$. σ^* -values taken from reference.⁵⁹

2.3.4. Practical implications

k_{O_3} prediction models were successfully developed in this study using correlations between k_{O_3} and quantum molecular orbital descriptors: $E_{\text{HOMO-n}}$ for aromatic compounds, $E_{\text{NBO, C-C}(\pi)}$ for olefins, and $E_{\text{NBO, LP-N}}$ for amines. Based on the observed MUE values, the predicted k_{O_3} were on average within a factor of 3.7, 3.1, and 3.9 of the experimental values for aromatic compounds, olefins, and amines, respectively. The extent of elimination of a micropollutant during ozonation and the propagated uncertainty (%) as a function of k_{O_3} , assuming an error of a factor of 4 in the predicted k_{O_3} , are shown in

Figure 2.7 (see Text S2.9 for details of the computations). Three regions are defined depending on the degree of uncertainty in elimination efficiency: Region I (20–78% uncertainty), II (1–20%), and III (<1%). In brief, the extent of elimination can be estimated with low uncertainty (<20%) in the regions II and III, where the predicted k_{O_3} range is $<6 \text{ M}^{-1}\text{s}^{-1}$ and $>640 \text{ M}^{-1}\text{s}^{-1}$, whereas in the region I the elimination estimated using the predicted k_{O_3} between 6 to $640 \text{ M}^{-1}\text{s}^{-1}$ possesses high uncertainty (>20%) with the maximum of 78% at $74 \text{ M}^{-1}\text{s}^{-1}$ (see Text S2.9 for detailed explanation). It is noteworthy that for ozonation of micropollutants with k_{O_3} of $< 640 \text{ M}^{-1}\text{s}^{-1}$, which covers the regions I, II (left part in Figure 2.7), and III (left), the oxidation by hydroxyl radicals is predominant.^{5,58} Thus, the error in the region with k_{O_3} of $< 640 \text{ M}^{-1}\text{s}^{-1}$ may be minor for the assessment of the overall extent of elimination during ozonation.

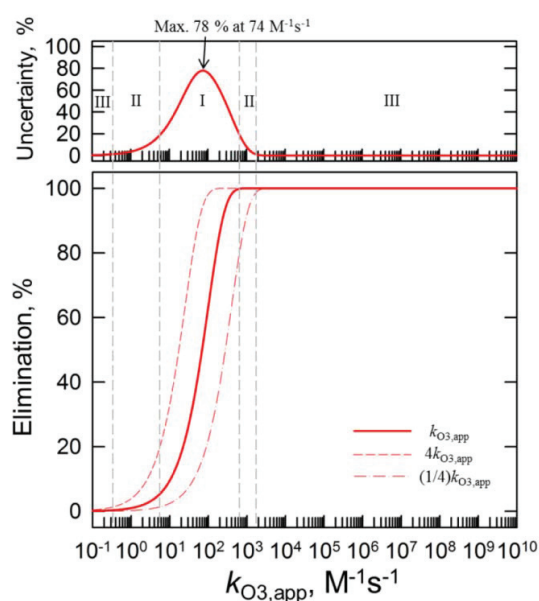


Figure 2.7. The predicted elimination (%) of a micropollutant (P) (bottom) as a function of the apparent second-order rate constant ($k_{O_3,app}$) for the reaction of ozone with P and the uncertainty (%) in the predicted elimination (top) due to the errors in predicted k_{O_3} values using the developed quantum chemical models. Solid lines: with $k_{O_3,app}$, dashed lines: with $1/4k_{O_3,app}$ and $4k_{O_3,app}$ based on the MUE of 0.57 log units for the models of aromatic compounds. Regions I, II, and III are defined by the range of the uncertainty being 20–78%, 1–20%, and <1%, respectively.

Based on the performance of the correlation models (Table 2.1) and the computational cost (HF/6-31G < B3LYP/6-31G < HF/6-311++G** < B3LYP/6-311++G**), the HF/6-31G method is recommended as the most efficient electronic structure method, among the four methods tested, for all compound groups except amines. Correlation results obtained with the 6-31G* and MIDI! basis sets were of similar or worse quality, on average, compared to those obtained with 6-31G and 6-311++G** (Figure S2.29). The success of the 6-31G basis set, which lacks polarization basis functions, was not foreseen. Polarization functions are known to offer generally improved performance for molecular

energies and geometries, molecular properties.³¹ Additionally, for most computational chemistry applications, the HF method is inferior to B3LYP and other competitive model chemistries, and the results of the present study should not be interpreted as a recommendation to use the HF method generally. The success of HF/6-31G for the particular applications in this study is ostensibly attributable to a fortuitous cancellation of errors.

However, for amines, the 6-31G basis set gave worse performance than 6-311++G**. To further investigate the role of the basis set, HF and B3LYP were additionally tested with the MIDI! and 6-31G* basis sets, both of which contain polarization functions. For this compound set, correlations using the HF/6-31G method ($R^2=0.77$) were inferior to those obtained with HF/6-311++G** ($R^2=0.83$), HF/MIDI! ($R^2=0.79$, slope=0.77, and y-intercept=15.23), or HF/6-31G* ($R^2=0.80$, slope=0.83, and y-intercept=15.57) (Figure S2.29f). The regression slope and the intercepts values are provided here in case basis sets less costly than 6-311++G** are desired.

Strong correlations between quantum molecular descriptors and QSAR descriptors such as Hammett and Taft constants indicate that (1) quantum molecular descriptors can be used for various QSAR/QSPR (Quantitative Structure/Property Relationship) models as surrogates for σ -values and (2) the developed quantum chemical models can be utilized to estimate unknown σ -values of substituents.

Acknowledgements

This study is dedicated to the memory of Prof. Jürg Hoigné whose pioneering studies on ozone chemistry were the basis of this publication. This study was funded by the Swiss Federal Office for the Environment (FOEN). The authors would like to thank Daniela Trogolo (EPFL) and Peter Rudolf Tenschler (EPFL/Eawag) for fruitful discussions.

References

- (1) von Gunten, U. Ozonation of drinking water: part I. Oxidation kinetics and product formation. *Water Res.* **2003**, *37*, 1443–1467.
- (2) von Gunten, U. Ozonation of drinking water: part II. Disinfection and by-product formation in presence of bromide, iodide or chlorine. *Water Res.* **2003**, *37*, 1469–1487.
- (3) von Sonntag, C.; von Gunten, U. *Chemistry of ozone in water and wastewater treatment: From basic principles to applications*; Iwa publishing, 2012.
- (4) Reungoat, J.; Escher, B. I.; Macova, M.; Argaud, F. X.; Gernjak, W.; Keller, J. Ozonation and biological activated carbon filtration of wastewater treatment plant effluents. *Water Res.* **2012**, *46*, 863–872.
- (5) Lee, Y.; Gerrity, D.; Lee, M.; Bogeat, A. E.; Salhi, E.; Gamage, S.; Trenholm, R. A.; Wert, E. C.; Snyder, S. A.; von Gunten, U. Prediction of micropollutant elimination during ozonation of municipal wastewater effluents: use of kinetic and water specific information. *Environ. Sci. Technol.* **2013**, *47*, 5872–5881.
- (6) Hollender, J.; Zimmermann, S. G.; Koepke, S.; Krauss, M.; McArdell, C. S.; Ort, C.; Singer, H.; von Gunten, U.; Siegrist, H. Elimination of organic micropollutants in a municipal wastewater treatment plant upgraded with a full-scale post-ozonation followed by sand filtration. *Environ. Sci. Technol.* **2009**, *43*, 7862–7869.
- (7) Hoigne, J.; Bader, H. Ozonation of Water: Role of Hydroxyl Radicals as Oxidizing Intermediates. *Science (80-)*. **1975**, *190*, 782–784.
- (8) Buffle, M.-O.; von Gunten, U. Phenols and Amine Induced HO • Generation During the Initial Phase of Natural Water Ozonation. *Environ. Sci. Technol.* **2006**, *40*, 3057–3063.
- (9) Pryor, W. A.; Giamalva, D. H.; Church, D. F. Kinetics of ozonation. 2. Amino acids and model compounds in water and comparisons to rates in nonpolar solvents. *J. Am. Chem. Soc.* **1984**, *106*, 7094–7100.
- (10) Hoigné, J.; Bader, H. Rate constants of reactions of ozone with organic and inorganic compounds in water—II. *Water Res.* **1983**, *17*, 185–194.
- (11) Hoigné, J.; Bader, H. Rate constants of reactions of ozone with organic and inorganic compounds in water—I. *Water Res.* **1983**, *17*, 173–183.
- (12) Haag, W. R.; Yao, C. C. D. Rate constants for reaction of hydroxyl radicals with several drinking water contaminants. *Environ. Sci. Technol.* **1992**, *26*, 1005–1013.
- (13) Buxton, G. V.; Greenstock, C. L.; Helman, W. P.; Ross, A. B. Critical review of rate constants for reactions of hydrated electrons, hydrogen atoms and hydroxyl radicals ($\cdot\text{OH}/\text{O}^-$ in aqueous solution). *J. Phys. Chem. Ref. data* **1988**, *17*, 513–886.
- (14) David Yao, C. ; Haag, W. R. Rate constants for direct reactions of ozone with several drinking water contaminants. *Water Res.* **1991**, *25*, 761–773.
- (15) Schymanski, E. L.; Singer, H. P.; Longrée, P.; Loos, M.; Ruff, M.; Stravs, M. A.; Ripollés Vidal, C.; Hollender, J. Strategies to characterize polar organic contamination in wastewater: exploring the capability of high resolution mass spectrometry. *Environ. Sci. Technol.* **2014**, *48*, 1811–1818.
- (16) Lee, Y.; von Gunten, U. Quantitative structure–activity relationships (QSARs) for the transformation of organic micropollutants during oxidative water treatment. *Water Res.* **2012**, *46*, 6177–6195.
- (17) Sudhakaran, S.; Amy, G. L. QSAR models for oxidation of organic micropollutants in water based on ozone and hydroxyl radical rate constants and their chemical classification. *Water Res.* **2013**, *47*, 1111–1122.
- (18) Naumov, S.; von Sonntag, C. Quantum Chemical Studies on the Formation of Ozone Adducts to Aromatic Compounds in Aqueous Solution. *Ozone Sci. Eng.* **2010**, *32*, 61–65.
- (19) Fukui, K. The Role of Frontier Orbitals in Chemical Reactions (Nobel Lecture). *Angew. Chemie Int. Ed. English* **1982**, *21*, 801–809.
- (20) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. Gaussian 09, Rev. C. 01, Gaussian, Inc., Wallingford CT **2009**.

- (21) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **1980**, *72*, 650.
- (22) Hehre, W. J. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257.
- (23) Frisch, M. J.; Pople, J. A.; Binkley, J. S. Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets. *J. Chem. Phys.* **1984**, *80*, 3265.
- (24) McLean, A. D.; Chandler, G. S. Contracted Gaussian basis sets for molecular calculations. I. Second row atoms, Z=11–18. *J. Chem. Phys.* **1980**, *72*, 5639.
- (25) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+G basis set for first-row elements, Li–F. *J. Comput. Chem.* **1983**, *4*, 294–301.
- (26) Wadt, W. R.; Hay, P. J. Ab initio effective core potentials for molecular calculations. Potentials for main group elements Na to Bi. *J. Chem. Phys.* **1985**, *82*, 284.
- (27) Hariharan, P. C.; Pople, J. A. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta* **1973**, *28*, 213–222.
- (28) Easton, R. E.; Giesen, D. J.; Welch, A.; Cramer, C. J.; Truhlar, D. G. The MIDI! basis set for quantum mechanical calculations of molecular geometries and partial charges. *Theoretica Chimica Acta*, 1996, *93*, 281–301.
- (29) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999–3093.
- (30) E. D. Glendening, A. E. Reed, J. E. Carpenter, F. W. NBO Version 3.1.
- (31) Cramer, C. J. *Essentials of computational chemistry: theories and models*; John Wiley & Sons, 2013.
- (32) Koopmans, T. Über die Zuordnung von Wellenfunktionen und Eigenwerten zu den Einzelnen Elektronen Eines Atoms. *Physica* **1934**, *1*, 104–113.
- (33) Parr, R. G.; Yang, W. *Density-functional theory of atoms and molecules*; Oxford university press, 1989; Vol. 16.
- (34) Zhao, Q.; Parr, R. Quantities $T_s[n]$ and $T_c[n]$ in density-functional theory. *Phys. Rev. A* **1992**, *46*, 2337–2343.
- (35) Zhao, Q.; Parr, R. G. Constrained-search method to determine electronic wave functions from electronic densities. *J. Chem. Phys.* **1993**, *98*, 543.
- (36) Baerends, E. J.; Gritsenko, O. V.; VanLeeuwen, R. Effective one-electron potential in the Kohn-Sham molecular orbital theory. *Chem. Appl. Density-Functional Theory* **1996**, *629*, 20–41.
- (37) Chong, D. P.; Gritsenko, O. V.; Baerends, E. J. Interpretation of the Kohn–Sham orbital energies as approximate vertical ionization potentials. *J. Chem. Phys.* **2002**, *116*, 1760.
- (38) Da Silva, R. R.; Ramalho, T. C.; Santos, J. M.; Figueroa-Villar, J. D. On the limits of highest-occupied molecular orbital driven reactions: the frontier effective-for-reaction molecular orbital concept. *J. Phys. Chem. A* **2006**, *110*, 1031–1040.
- (39) Dodd, M. C.; Buffle, M.-O.; von Gunten, U. Oxidation of Antibacterial Molecules by Aqueous Ozone: Moiety-Specific Reaction Kinetics and Application to Ozone-Based Wastewater Treatment. *Environ. Sci. Technol.* **2006**, *40*, 1969–1977.
- (40) Kuang, J.; Huang, J.; Wang, B.; Cao, Q.; Deng, S.; Yu, G. Ozonation of trimethoprim in aqueous solution: identification of reaction products and their toxicity. *Water Res.* **2013**, *47*, 2863–2872.
- (41) Sustmann, R. Orbital energy control of cycloaddition reactivity. *Pure Appl. Chem* **1974**, *40*, 569–593.
- (42) Houk, K. N.; Sims, J.; Watts, C. R.; Luskus, L. J. Origin of reactivity, regioselectivity, and periselectivity in 1,3-dipolar cycloadditions. *J. Am. Chem. Soc.* **1973**, *95*, 7301–7315.
- (43) McKee, M. L.; Rohlfing, C. M. An ab initio study of complexes between ethylene and ozone. *J. Am. Chem. Soc.* **1989**, *111*, 2497–2500.
- (44) Kuczkowski, R. L. The structure and mechanism of formation of ozonides. *Chem. Soc. Rev.* **1992**, *21*, 79.
- (45) Dowideit, P.; von Sonntag, C. Reaction of Ozone with Ethene and Its Methyl- and Chlorine-Substituted Derivatives in Aqueous Solution. *Environ. Sci. Technol.* **1998**, *32*, 1112–1119.

Chapter 2

- (46) Chaudhuri, R. K.; Hammond, J. R.; Freed, K. F.; Chattopadhyay, S.; Mahapatra, U. S. Reappraisal of cis effect in 1,2-dihaloethenes: an improved virtual orbital multireference approach. *J. Chem. Phys.* **2008**, *129*, 064101.
- (47) Yamamoto, T.; Kaneno, D.; Tomoda, S. The Origin of Cis Effect in 1,2-Dihaloethenes: The Quantitative Comparison of Electron Delocalizations and Steric Exchange Repulsions. *Bull. Chem. Soc. Jpn.* **2008**, *81*, 1415–1422.
- (48) Bingham, R. C. The stereochemical consequences of electron delocalization in extended π systems. An interpretation of the cis effect exhibited by 1,2-disubstituted ethylenes and related phenomena. *J. Am. Chem. Soc.* **1976**, *98*, 535–540.
- (49) Williamson, D. G.; Cvetanovic, R. J. Rates of reactions of ozone with chlorinated and conjugated olefins. *J. Am. Chem. Soc.* **1968**, *90*, 4248–4252.
- (50) Williamson, D. G.; Cvetanovic, R. J. Rates of ozone-olefin reactions in carbon tetrachloride solutions. *J. Am. Chem. Soc.* **1968**, *90*, 3668–3672.
- (51) Huber, M. M.; Ternes, T. A.; von Gunten, U. Removal of Estrogenic Activity and Formation of Oxidation Products during Ozonation of 17 α -Ethinylestradiol. *Environ. Sci. Technol.* **2004**, *38*, 5177–5186.
- (52) Prasse, C.; Wagner, M.; Schulz, R.; Ternes, T. A. Oxidation of the antiviral drug acyclovir and its biodegradation product carboxy-acyclovir with ozone: kinetics and identification of oxidation products. *Environ. Sci. Technol.* **2012**, *46*, 2169–2178.
- (53) McDaniel, D. H.; Brown, H. C. An Extended Table of Hammett Substituent Constants Based on the Ionization of Substituted Benzoic Acids. *J. Org. Chem.* **1958**, *23*, 420–427.
- (54) Galabov, B.; Ilieva, S.; Schaefer, H. F. An efficient computational approach for the evaluation of substituent constants. *J. Org. Chem.* **2006**, *71*, 6382–6387.
- (55) Korenaga, T.; Kadowaki, K.; Ema, T.; Sakai, T. Reestimation of the Taft's substituent constant of the pentafluorophenyl group. *J. Org. Chem.* **2004**, *69*, 7340–7343.
- (56) Domingo, L. R.; Pérez, P.; Contreras, R. Electronic contributions to the sigma(p) parameter of the Hammett equation. *J. Org. Chem.* **2003**, *68*, 6060–6062.
- (57) Emilia, M.; Silva, N. P. R. A.; Pombeiro, A. J. L.; da Silva, J. J. R. F.; Herrmann, R.; Deus, N.; Castilho, T. J.; Silva, M. F. C. G. Redox potential and substituent effects at ferrocene derivatives. Estimates of Hammett σ_p and Taft polar σ substituent constants. *J. Organomet. Chem.* **1991**, *421*, 75–90.
- (58) Buffle, M.-O.; Schumacher, J.; Salhi, E.; Jekel, M.; von Gunten, U. Measurement of the initial phase of ozone decomposition in water and wastewater by means of a continuous quench-flow system: application to disinfection and pharmaceutical oxidation. *Water Res.* **2006**, *40*, 1884–1894.
- (59) Hansch, C.; Leo, A.; Hoekman, D.; Heller, S. R. *Exploring Qsar*; American Chemical Society Washington, DC, 1995.

Supporting information for chapter 2.

Development of prediction models for the reactivity of organic compounds with ozone in aqueous solution by quantum chemical computations: Role of delocalized and localized molecular orbitals

including 9 texts, 2 schemes, 29 figures, and 8 tables.

Text S2.1. Computational methodology

Initial three-dimensional (3D) geometries of all chemical compounds were obtained by deriving a lowest energy conformer calculated by the conformer module of ChemAxon MarvinSketch (Marvin 5.10.0, 2012, <http://www.chemaxon.com>) employing the Dreiding force field.¹ These geometries were further fully optimized without symmetry constraints by *ab initio* Hartree Fock (HF) and the Density Functional Theory (DFT) method B3LYP^{2,3} using the Gaussian 09 D.01 program⁴. The “ultrafine” grid was employed in Gaussian for numerical integration of the density. The Pople 6-31G and 6-311++G** basis sets⁵⁻⁹ were used for all atoms except iodine for which a quasi-relativistic pseudopotential (LANL2DZ)¹⁰ was used. The 6-31G*¹¹ and MIDI!¹² basis sets were additionally used to investigate the effect of polarization functions. The geometry optimization computations were performed and the resulting optimized structures are confirmed to be the local minimum on the potential energy surface by frequency analysis with the same method used for geometry optimization. Single point calculations were then performed at the same level of computation, unless stated otherwise. In additional computations, the optimized geometry with 6-31G was used for computations with 6-311++G** to investigate further improvement. The effects of aqueous solvent were taken into account by implementing the integral equation formalism polarizable continuum model (IEF-PCM),¹³ for which the Universal Force Field (UFF) was used for the atomic radii, with a dielectric constant (ϵ) of 78.3 for water throughout all computations in this study, unless stated otherwise. NBO analysis was carried out using the NBO 3.1 program^{14,15} included in Gaussian 09 and the NBO 6.0 standalone package¹⁶, and the NBO energies by the NBO 3.1 were presented herein because the energies were nearly the same between two different versions. The conformers for four compounds belonging to the trimethoxybenzene group were searched using Gabedit 2.4.8¹⁷ based on a molecular dynamic simulation using molecular mechanics with the Amber 99 force field). Molden was used for checking 3D structures and the visualization of molecular orbitals and natural bond orbitals.¹⁸

Text S2.2. Rotamers of selected aromatic compounds

i) Paracetamol: For the initial geometry with a dihedral angle of 17° between the phenyl ring and the acetamido group, the geometry optimization at the HF/6-311++G** level resulted in a perpendicular paracetamol conformer with a dihedral angle of 88°. In contrast, planar or close-to-planar conformers (0.6–8°) were obtained from the three other computation methods. To check the sensitivity of the method employed, the geometry of paracetamol was additionally investigated by the other hybrid density functionals such as M06-2X¹⁹ and MPW1K²⁰ with the initial geometries of 0° and 90° dihedral angles. Similar to the B3LYP method, planar and close-to-planar structures (0.0–24°) were obtained. Thus, a near-planar structure seems to be more plausible than the perpendicular conformer for paracetamol. However, as shown in Figures 2.1a (main text) and S2.1, the E_{HOMO} of only the perpendicular paracetamol conformer at the HF/6-311++G** is located close to the E_{HOMO} of the phenolic compounds. Nevertheless, a perpendicular geometry was not considered representative of the correlation model, and paracetamol remained unresolved in this study for the following reason: the nitrogen of the acetamido group has a sp^2 hybridized nitrogen, thus the p orbital of the nitrogen lone-pair overlaps

with the adjacent aromatic π system, increasing the planarity of the system. The π conjugation may be influenced by an attack of ozone on the ring. However, it is unclear whether a somewhat distorted geometry or even a completely perpendicular geometry can be relevant for ozone reactions.

ii) Analysis of rotamers of 8 selected aromatic compounds: as observed in the case of paracetamol above, even with the same initial geometry, the optimization at the different levels of theory may result in differing geometries, leading to differing $E_{\text{HOMO}(-n)}$ and thereby an error in predicting k_{O_3} . Therefore, the effect of the rotation of the substituent on the E_{HOMO} was investigated for 8 aromatic compounds with different substituents such as phenol, nitrobenzene, benzoate ion, benzaldehyde, ethylbenzene, anisole, paracetamol, and linuron. The rotamers of the selected compounds were investigated by changing the dihedral angle (φ) from 0 to 90° or 360° by steps of 10, 15, or 30°, depending on the symmetry of the molecule, with both the HF and the B3LYP method using the 6-311++G** basis set. At each fixed value of dihedral angles, the remainder of the internal coordinates was optimized with the same electronic structure method, which is also known as a relaxed scan or minimum energy pathway.

In Figure S2.4, the profiles of the E_{HOMO} (ΔE_{HOMO} , eV) and the total electronic energy (ΔE_{tot} , kcal/mol) relative to the E_{HOMO} and the E_{tot} for the planar geometry ($\varphi = 0^\circ$), respectively, as a function of a dihedral angle (φ) are shown for both the HF/6-311++G** and B3LYP/6-311++G** methods. Note that the quantum chemical computations for these compounds were performed from 0° up to 90° and these results were used for the range between 90° and 180° by symmetry. Exceptionally, the full scan from 0° to 360° was performed for paracetamol and linuron because those are not symmetric.

For anisole, the E_{HOMO} decreases as φ increases up to 90°, where the lowest E_{HOMO} is found. In contrast, the E_{tot} increases as φ increases, thus the minimum was at 0°. This inversely proportional relationship between ΔE_{HOMO} and ΔE_{tot} as a function of φ was also observed for phenol and linuron. In contrast, a proportional relationship between ΔE_{HOMO} and ΔE_{tot} was observed for benzaldehyde, benzoate ion, and ethylbenzene. For nitrobenzene and paracetamol the proportional relationship between ΔE_{HOMO} and ΔE_{tot} depended on the electronic structure method used. For nitrobenzene a proportional relationship was found with the B3LYP method, whereas an inversely proportional relationship was found with the HF method. For paracetamol the HF and the B3LYP methods showed a proportional and inversely proportional relationship, respectively. As the main interest in this study is the analysis of the E_{HOMO} of rotamers, E_{tot} will not be discussed hereafter.

The maximum ΔE_{HOMO} was consistently found at φ of 90° for all the compounds investigated. A relatively small effect of the substituent rotation on E_{HOMO} was found for ethylbenzene (the maximum ΔE_{HOMO} : 0.03 eV and 0.04 eV for HF and B3LYP, respectively), nitrobenzene (0.03 eV and 0.17 eV), benzaldehyde (0.03 eV and 0.12 eV), and benzoate ion (0.22 eV and 0.18 eV). Among these compounds, the largest observation variation in ΔE_{HOMO} with substituent rotation is about 0.2 eV, which leads to a discrepancy of a factor of ~ 3 in the predicted k_{O_3} when using the corresponding model equation for the benzene derivatives in Table 2.1. However, a larger effect was obtained for phenol (0.40 eV and 0.48 eV), anisole (0.44 eV and 0.56 eV), paracetamol (0.70 eV and 0.60 eV), and linuron (0.70 eV and 0.60 eV). A larger uncertainty in the predicted k_{O_3} , from 10 up to ~ 400 times the ex-

pected value, can be obtained for these compounds. The hydroxyl group of the phenols was confirmed to be in the phenyl plane with all electronic structure methods, which comes as no surprise as it is commonly known to be planar. Exceptionally, 2,6-di-*tert*-butyl-4-methylphenol gave inconsistent geometries between the different electronic structure methods. Having performed geometry optimization of the two different initial geometries of 2,6-di-*tert*-butyl-4-methylphenol, i.e., planar and perpendicular, the B3LYP method consistently yielded the planar geometry conformer. While the HF/6-31G method favored only the perpendicular geometry, the HF/6-311++G** retained its initial geometry, i.e., a planar geometry remained as planar and a perpendicular geometry remained as perpendicular.

The perpendicular conformation of 2,6-di-*tert*-butyl-4-methylphenol can be explained by the steric effect of the neighboring substituents, two *tert*-butyl groups in 2 and 6 positions. However, the B3LYP methods and the HF method with 6-311++G** basis set suggests that the steric effect is not strong enough to favor a perpendicular geometry. Interestingly, the k_{O_3} values predicted for both conformations were within a factor of 3 of the experimental value, hence it is difficult to deduce which conformation may be representative based on the E_{HOMO} value alone. Discussions about rotamers of linuron and anisole are available in the main manuscript and below, respectively. As discussed above in section i), paracetamol remained unresolved in this study.

iii) Mono- and di-alkoxybenzenes: the geometry optimization was performed for two different initial geometries of 17 mono- and di-alkoxybenzenes: a planar geometry and a perpendicular geometry. For the HF computations regardless of the basis set, both initial geometries were retained following the geometry optimization for all alkoxybenzenes except for four compounds, 4-methoxy-1-naphthalenesulfonic acid, carbofuran, 2,4,6-trichloroanisole, and 2,4,6-tribromoanisole. Only the planar geometry was obtained for 4-methoxy-1-naphthalenesulfonic acid, that is, the perpendicular initial geometry was converted into a planar one after the geometry optimization. In contrast, the perpendicular geometry was consistently obtained for the three other compounds, i.e., carbofuran and two trihaloanisoles, regardless of whether the initial geometry was planar or perpendicular. The two perpendicular trihaloanisoles can be understood by the steric repulsion acting on the methoxy group by the neighboring halogens. Carbaryl, which has a carbamate group as carbofuran, was reported to have a perpendicular conformation.²¹ A correlation using the E_{HOMO} of the group of alkoxybenzenes dominated by a planar-type geometry (14 planar-type and 3 perpendicular-type) was compared with the correlation using the E_{HOMO} of the group dominated by a perpendicular-type geometry (1 planar-type and 16 perpendicular-type) at the HF/6-31G and HF/6-311++G** levels. Each respective model group is a mixture of both planar and perpendicular geometries because there were some compounds that retained only one conformation following the geometry optimization, regardless of their initial geometry. Those compounds are carbofuran, 2,4,6-trichloroanisole, and 2,4,6-tribromoanisole which are perpendicular-type, and 4-methoxy-1-naphthalenesulfonic acid, which is planar-type. See above for discussions. As shown in Figure S2.8, better correlations were observed with the planar-type dominated group with the HF method regardless of the basis set. This is plausible because the p orbital of the oxygen-lone pair electron conjugates with the π system of the aromatic ring. Moreover, the fact that planar rotamers were consistently lower in energy than perpendicular ones also supports this result.

In contrast, the B3LYP computations transformed the perpendicular geometries of many compounds to planar geometries after the optimization, while carbofuran, 2,4,6-trichloroanisole, and 2,4,6-tribromoanisole remained perpendicular. Thus, only the correlation using the E_{HOMO} of the planar-type dominated group could be made. In this regard, the B3LYP method is expected to give more plausible geometries than the HF method. More attention seems necessary for the HF computations because both planar and perpendicular minimum geometries of many alkoxybenzenes can be found. However, the performance of the correlation models with B3LYP turned out to be worse than the models by the HF method ($R^2=0.87$ and 0.90 by the HF/6-31G and HF/6-311++G** and $R^2=0.80$ and 0.80 by the B3LYP/6-31G and the B3LYP/6-311++G**, respectively).

Text S2.3. Analysis of two different geometries of anilines

Using planar initial geometries of anilines, for which the dihedral angle between the amino group and the phenyl plane is zero, the geometry optimization by both HF and B3LYP methods using the 6-31G basis set predicted a planar conformation for all 16 aniline derivatives. Frequency analysis confirmed that these planar anilines are local minima. Even with initial geometries of pyramidal anilines with a dihedral angle of 25° and 45° , all anilines reverted to the planar structures after the optimization at both methods using the 6-31G basis set. The dihedral angle of aniline is reported to be 42.4° in the gas phase²² and the substituted anilines investigated herein are also expected to be pyramidal in aqueous solution. Thus, the planar geometry does not seem to be the true equilibrium conformation of aniline. Planar configurations result from the absence of the d -polarization function (*) in the basis set.²³ As shown in Figures S2.7a and c, nevertheless, satisfactory correlations using the E_{HOMO} of the planar anilines were observed ($R^2=0.85$ by the HF/6-31G and 0.82 by the B3LYP/6-31G). In contrast, many planar anilines were found to be in a transition state (TS), with a single imaginary frequency, when using either HF or B3LYP with the 6-311++G** basis set. As the vibration of the imaginary frequency was at the amino group, changing the geometries from planar to pyramidal led to a local minimum eliminated the imaginary frequency. As described above, a pyramidal geometry is considered to be the true geometry of anilines as a reactant and in this respect the 6-311++G** basis set seems to yield more reasonable geometries than the 6-31 basis set. Therefore, pyramidal geometries were consistently used for computations using the 6-311++G** basis set. Compared to the computations using the 6-31G basis set above, however, somewhat worse correlations were obtained ($R^2=0.75$ by the HF/6-311++G** and 0.78 by the B3LYP/6-311++G**) (Figures S2.7b and d). The inferior performance is mainly attributed to the results for 3-nitroaniline/anionic sulfamethoxazole and 3-nitroaniline with the HF/6-311++G** and B3LYP/6-311++G** methods, respectively. Without these compounds, an R^2 -value of 0.83 was obtained with both methods. Since good correlations were obtained above with the planar anilines, another correlation was made with planar geometries calculated by the 6-311++G** basis set. Single point calculations were performed for the planar geometries optimized by the 6-31G basis set without further optimization. As shown in Figures S2.7b and d, the correlations of planar anilines outperformed the one of pyramidal anilines ($R^2=0.86$ for the planar geometries and $R^2=0.75$ for the pyramidal geometries with HF computations and $R^2=0.84$ for the planar geometries and $R^2=0.78$ for the pyramidal geometries with B3LYP computations). This mainly results from the fact

Supporting information for chapter 2

that 3-nitroaniline and anionic sulfamethoxazole, which were outlying when their geometry was optimized by the 6-311++G** basis set, are located closer to the regression line when optimized with 6-31G. Based on these observations, we propose to use a planar geometry of anilines as representative for the model. This decision can be loosely supported by transition state (TS) theory. The geometry of the TS is more relevant to reaction rates than the geometry of the reactant. As anilines are expected to be transformed from pyramidal to planar upon the addition of ozone to the aromatic ring, this may explain why the E_{HOMO} of planar geometries correlate better with the corresponding k_{O_3} . The detection of an imaginary frequency of planar anilines, using either HF or B3LYP with the 6-311++G** basis set, also supports this explanation.

Text S2.4. Analysis of the conformers of trimethoxybenzenes

Having three methoxy groups (-O-CH₃)₃, trimethoxybenzenes can have numerous rotamers depending on the planar and non-planar combinations of the three O-CH₃ bonds. For this group, experimental k_{O_3} values are only available for 1,3,5-trimethoxybenzene, 3,4,5-trimethoxytoluene, mono-, and di-protonated trimethoprim. A conformational search using Gabedit 2.4.8 returned 2, 4, 6, and 6 conformers for 1,3,5-trimethoxybenzene, 3,4,5-trimethoxytoluene, mono-, and di-protonated trimethoprim, respectively. The initial geometries obtained were further optimized by quantum chemical computations with four different electronic structure methods. The comparison of the geometries before and after geometry optimization in terms of planar and non-planar arrangements of the substituents is shown in Figure S2.9.

The conformers I and II of 1,3,5-trimethoxybenzene, which are planar, and the conformers I of 3,4,5-trimethoxytoluene, mono-, and di-protonated trimethoprim, which are non-planar, retained their planarity and non-planarity, respectively. In contrast, all remaining conformers, which have two or three in-plane methoxy groups, underwent a rotation of one or two methoxy groups from in-plane to out-of-plane (Figure S2.9). Eventually, no planar structure was obtained except 1,3,5-trimethoxybenzene following the geometry optimization with all computation methods. It was observed that a few conformers were similar in their arrangement of the substituents. For instance, the conformer VI of both mono- and di-protonated trimethoprim became similar to the conformer I thereof.

By putting the E_{HOMO} values of all conformers for the four compounds together, two separate groups resulted, each assigned a distinct correlation model (Figure S2.10). Note that Figures S2.10a, c, e, and g show the conformers of trimethoxybenzenes in comparison with the phenol and aniline derivatives (grey symbols), whereas Figures S2.10b, d, f, and h show the exclusive zoom-ins of the trimethoxybenzenes. One correlation, which is lower in E_{HOMO} energy than the other, consists of 3,4,5-trimethoxytoluene, mono-, and di-protonated trimethoprim with all three methoxy groups being out-of-plane, grouped together with the conformer I of 1,3,5-trimethoxybenzene. Another correlation consists of 3,4,5-trimethoxytoluene, mono-, and di-protonated trimethoprim with either one or two methoxy groups being in-plane, grouped together with the conformer II of 1,3,5-trimethoxybenzene. As shown in Figures S2.10a, c, e, and g, the latter correlation is located closer to the phenol and the aniline derivatives.

For the trimethoxybenzenes, the relative total electronic energies (ΔE) and the relative Gibbs free energies (ΔG) at 298.15 K of the conformers computed with respect to the lowest energy conformer are presented in Table S2.8 for the HF/6-311++G** and B3LYP/6-311++G** methods. The populations of the conformers calculated using the Boltzmann equilibrium is also presented, based on the computed Gibbs free energies. As shown in Table S2.8, the conformer I of 1,3,5-trimethoxybenzene, for which E_{HOMO} is lower than for conformer II, is more stable than conformer II, and more populated. In contrast, for trimethoxybenzenes with three methoxy groups substituted in 3, 4, and 5 positions, the opposite trend was observed, i.e., the higher its E_{HOMO} is, the more stable and populated the conformer is.

Among the two correlations shown in Figures S2.10b, d, f, and h, the group consisting of the trimethoxybenzenes with the lower E_{HOMO} was chosen as representative in this study since it correlated better than the group with the high E_{HOMO} ($R^2=0.95-0.99$ and $0.85-0.93$ for the ones with low E_{HOMO} and the ones with high E_{HOMO} , respectively). If the k_{O_3} prediction model is based on thermodynamically stable conformers, then 1,3,5-trimethoxybenzene does not fit into the model because its conformer I, which is more stable than conformer II based on ΔE and ΔG , has a lower E_{HOMO} as well (see Figure S2.10 and Table S2.8). Prior to applying trimethoxybenzenes to this model, therefore, it is recommended that the geometry of the trimethoxybenzene with the lowest E_{HOMO} among the investigated conformers be used. However, it should be noted that this decision is empirically based. Therefore, the suggested model is provisional because only four compounds were available for investigation of this group, and there is no other information available about the geometry of trimethoxybenzene relevant for ozone reactions.

Text S2.5. Ozone attack on nitrogen-containing aromatic compounds

trans-Azobenzene, 2-isopropyl-3-methoxypyrazine (IPMP), pyridine, and several triazines were separately investigated in addition to the five groups of aromatic compounds in this study (see Figure S2.2 for the chemical structures). Based on product studies, the reaction site of these aromatic compounds for the reaction with ozone seems to be the nitrogen atom instead of the ring.^{24,25} The products pyridine-*N*-oxide and azoxybenzene were mainly detected from ozonation of pyridine and azobenzene, respectively.^{24,25} Similarly, we expected IPMP to also be attacked at the nitrogen, although there has not yet been any experimental evidence.²⁵ In addition to three nitrogens in the heterocyclic ring, triazines such as atrazine, deethylatrazine, deethyldeisopropylatrazine, deisopropylatrazine, and simazine have two exocyclic nitrogens. Since 4-acetamino-2-chloro-6-isopropylamino-*s*-triazine, 2-chloro-4-ethylimino-6-isopropylamino-*s*-triazine, and deethylatrazine were reported to be the dominant products from ozonation of atrazine,²⁶ the exocyclic nitrogen for these compounds seems to be the site of ozone attack. Note that the product studies above were conducted in the presence of a hydroxyl radical scavenger. Thus, it was interpreted that the detected products exclusively resulted from the reaction with ozone.

As explained in the main manuscript, the selection of a molecular orbital for the model development in this study requires a mechanistic consideration, i.e., a molecular orbital should be entirely or par-

Supporting information for chapter 2

tially but not insignificantly located in the place where ozone attacks. Based on this condition, representative molecular orbitals for the ozone reaction were sought.

The molecular orbitals of *trans*-azobenzene, pyridine, IPMP, and atrazine, which are considered to be important, and their energy levels are presented in Figure S2.11. Moreover, the correlation of the $E_{\text{HOMO-n}}$ values of those molecular orbitals with their corresponding k_{O_3} values are shown in Figure S2.12 in comparison with the data points of phenol and benzene derivatives from Figures 2.1a and S2.1.

As shown in Figure S2.11, two molecular orbitals (MOs) located on the nitrogen were found to be relevant for the ozone reaction of *trans*-azobenzene. While the MO-A features π bonding spread over not only the aromatic ring but also the N=N double bond (denoted as the MO-A(π)), the MO-B shows appreciably non-bonding (nb) lone pairs of the nitrogen (denoted as the MO-B(nb)). Having considered that the nitrogen lone pairs of aliphatic amines react with ozone to form an ozone adduct which is followed by the loss of singlet oxygen to form the N-oxide,²⁵ the nitrogen lone-pairs of *trans*-azobenzene are also thought to be responsible for the formation of azoxybenzene during ozonation. Therefore, mechanistically the MO-B(nb) seems to be more appropriate for the representative molecular orbital than the MO-A(π). As shown in Figures S2.12a and b, however, the MO-B(nb), which was found to be the HOMO-4 for the HF computations, is located far from the data points of the benzene derivatives. Rather, the MO-A(π), which is the HOMO, turned out to be in proximity to the benzene derivatives. The use of the HOMO can be explained by the frontier molecular orbital theory but cannot be justified by the generally accepted reaction mechanism. It is unclear whether or not the success of the HOMO is fortuitous. In contrast to the HF computations, for the B3LYP method, the MO-A(π) and the MO-B(nb) only appeared between the HOMO and the HOMO-1, and a small energy gap between them was obtained (Figure S2.11). This near-degeneracy of these two MOs leads to both energies being located close to those of the benzene derivatives (Figures S2.12c and d).

A similar phenomenon was also observed for pyridine and IPMP (Figures S2.11 and S2.12). Interestingly, a good correlation was found within the group of the three compounds discussed above by using the mechanistically chosen non-bonding MOs for the HF computations ($R^2=0.94$ and 0.99 for the HF/6-31G and the HF/6-311++G** in Figures S2.12a and b, respectively).

Two near-degenerate MOs, which have the electron density on the exocyclic nitrogen, were found as the HOMO and the HOMO-1 for atrazine for the HF computations (Figure S2.11). Among two exocyclic nitrogens, the MO-A(nb) has an orbital located on the nitrogen attached to the isopropyl group, whereas the MO-B(nb) has a bigger orbital on the nitrogen with the ethyl group than the one with the isopropyl group. As discussed above, the major products from ozonation of atrazine seem to have resulted from the attack of the nitrogen with the ethyl group. Thus, the MO-B(nb) was chosen as the representative MO. Note that the MO-A(nb) would lead to almost the same result as the MO-B(nb) because those are near-degenerate. In contrast, the B3LYP computations resulted in the HOMO which has the orbital seemingly equally located on both exocyclic nitrogens. Thus, this HOMO was chosen and denoted as the MO-C(nb) in Figure S2.11. Similar molecular orbitals were obtained for the other triazines, and thus we selected the same orbitals as for atrazine. With HF computations, the selected

$E_{\text{HOMO}-n}$ values are rather close to the benzene derivatives, whereas for the B3LYP computations these values are very close to the benzene derivatives (Figure S2.12).

Overall, the k_{O_3} prediction model for the benzene derivatives may be used to estimate k_{O_3} for ozone-reactive nitrogen-containing compounds, by using the mechanistically chosen MOs from B3LYP computations. However, this approach is considered provisory, based on the limited number of data available for this group.

Text S2.6. Outliers from the k_{O_3} prediction model for olefins

Based on the fact that 8 olefins were excluded as outliers (see the main manuscript and Figure 2.3 for the list of outliers), it must be assumed that the developed model is likely to be inaccurate for certain olefins. While neutral maleic and fumaric acid fit well into the model, three conjugate bases were outliers, especially the maleate ion (Figures 2.3b and S2.16). This indicates that inferior k_{O_3} prediction for olefins with anionic carboxylic groups can be expected. The incorporation of diffuse functions denoted as ++ in the Pople basis set is considered important for anionic compounds because of the spread-out electron density over the molecule. However, the inclusion of diffuse functions did not improve the performance of the correlation model (Figures S2.16a and c).

As the solvation method can also influence the energy of ionic species, we also tried the CPCM solvation method for neutral and ionic maleic and fumaric acids calculated at the HF/6-311++G** level; however no improvement was observed (data not shown). Other implicit solvation methods (e.g., Conductor-like Screening Model (COSMO)), or an explicit approach using water molecules remains to be investigated.

While cinnamic acid, 4-methoxycinnamic acid, and 4-nitrocinnamic acid fit into the model, 3,4-dihydroxycinnamic acids and 3-methoxy-4-hydroxycinnamic acids (neutral and anionic) were outliers. These compounds are activated by a hydroxy and a methoxy group at the aromatic ring. However, the k_{O_3} -values of the aromatic rings predicted by the phenol model using the E_{HOMO} were lower than the measured k_{O_3} -values by about one order of magnitude. It is unclear what the reason for this discrepancy is.

For olefins with electron-withdrawing substituents including nitrogen or oxygen, vinylene carbonate and acrylonitrile were excluded as outliers. However, the errors of several other members such as 1,4-benzoquinone, vinyl acetate, and 2-acetaminoacrylate were as large as a factor of 5 to ~10. This indicates that the developed NBO model might not be as reliable for olefins with electron-withdrawing groups as for other olefins.

Text S2.7. Nucleic acid constituents and miscellaneous olefins (conjugated olefins)

38 neutral and ionic nucleic acid constituents were investigated: adenine and 7 derivatives, cytosine and 7 derivatives, guanosine and 7 derivatives, thymine and 4 derivatives, and uracil and 8 derivatives

Supporting information for chapter 2

(see Figure S2.17 for chemical structures). Prototropic tautomerism in aqueous solution was taken into account for the determination of their initial geometries. First, the amino and keto form over the imino and enol form were chosen for exocyclic nitrogen and oxygen of pyrimidine and purine bases, respectively, as it is commonly known that the amino and keto forms are predominant. Second, the protonation sites of the ring of neutral and ionic compounds were determined based on the literature, when there are multiple sites of nitrogen to be protonated (Table S2.3). Numbering conventions of purines and pyrimidines by the International Union of Pure and Applied Chemists (IUPAC) were used and are shown in Figure S2.17 as the representative for adenine, cytidine, guanosine, thymine, and uracil. Exceptionally, the weighted average of $E_{\text{HOMO-n}}$ and $E_{\text{NBO, C-C}(\pi)}$ of two tautomers was used for the neutral adenine and the anionic uracil because it was reported that two tautomers coexist in aqueous solution: neutral adenine (0.22:0.78, 7H:9H)²⁷ and anionic uracil (0.5:0.5, 1H:3H),²⁸ respectively, where the number in front of H is the protonation position on the ring. To our knowledge, there are no experimental observations reported for anionic 5-chlorouracil, anionic 6-methyluracil, and anionic thymine (5-methyluracil); for these compounds the observed ratio for anionic uracil (0.5:0.5, 1H:3H) was applied based on structural similarity.

As shown in Figure S2.18, good correlations ($R^2=0.83-0.99$) were found within the respective nucleic acid classes regardless of the computation method. However, only a small number of nucleic acid constituents were close to the regression line of the olefins. Note that 5-chlorouracil and cylin-drospormopsin were excluded as outliers from the uracil group and due to their uncertain k_{O_3} values, respectively. The noticeable difference between the aliphatic olefins and the nucleic acid constituents is the orbital occupancy of the NBO. The occupancy of the olefins lies between approximately 1.8 and 2.0, whereas that of the nucleic acid constituents is between 1.5 and 1.9, with thymines and uracils as the least delocalized (1.7-1.9) and with adenine as the most delocalized (1.5-1.7), respectively (see the upper graphs in Figure S2.18). Nucleic acid constituents consist of purine and pyrimidine, which are heterocyclic aromatic compounds. Their aromatic characteristics are supported by their low orbital occupancy (1.5 - 1.9) of the three double bonds in a ring, which is in a similar range as aromatic compounds investigated in this study (1.5 - 1.8) (data not shown). Therefore, we considered applying correlations with the $E_{\text{HOMO-n}}$, which was successful for aromatic compounds discussed above (Figure S2.19). The mechanistic assumption that ozone attacks the C=C bond of a heteroaromatic ring was assured by confirming that part of the selected HOMO-n is located on the C=C bond, even if the orbital is also delocalized over other parts of the molecule. Similar to the NBO energy, good correlations ($R^2=0.84-0.99$) were found using the $E_{\text{HOMO-n}}$ within the respective nucleic acid classes. Exceptionally, the model performance of the adenine group using the $E_{\text{HOMO-n}}$ was worse than the model using the $E_{\text{NBO, C-C}(\pi)}$ ($R^2=0.67-0.74$ by HOMO and $R^2=0.75-0.86$ by NBO) (Figure S2.19). It has to be noted that 5-chlorouracil, which was an outlier in the NBO model, can be included in the FMO model.

Miscellaneous olefins, which commonly feature a conjugated C=C bond, such as cyclohexenones, diaminopyrimidines, imidazoles, indoles, isoxazoles, quinolone, and indigotrisulfonic acid (Figure S2.17 for the chemical structures) were grouped separately from other olefins because they manifested a correlation of their own. While $E_{\text{NBO, C-C}(\pi)}$ did not produce a satisfactory correlation with k_{O_3} (Figures S2.20a and S21a-c), a good correlation ($R^2=0.77-0.91$) was observed using $E_{\text{HOMO-n}}$ for the miscellaneous olefins, excluding the diprotonated diaminopyrimidine as an outlier (Figures S2.20b and S21d-f). Note that thymines and uracils were included in this group as they were close to the miscellaneous olefins. Many compounds investigated in this section feature a heteroaromatic ring con-

taining a C=C bond. Thus, their grouping with aromatic compounds using $E_{\text{HOMO-n}}$ was considered to be possible, but no meaningful correlations with any developed model for aromatic compounds were observed (data not shown). A summary of the developed models for the nucleic acid constituents and the miscellaneous olefins is given in Table S6.

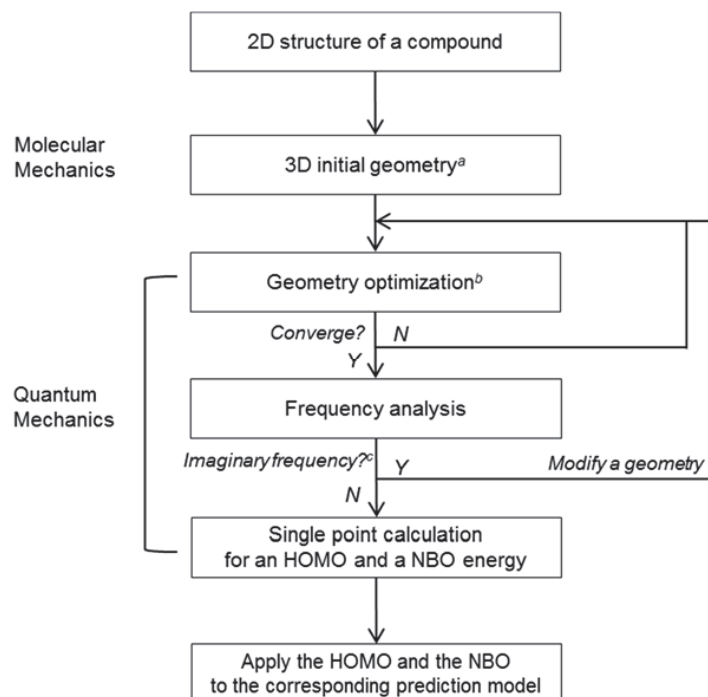
Text S2.8. Sulfur-containing compounds

14 neutral and ionic S-containing compounds comprising of thiols, sulfides, disulfides, a sulfoxide, and a sulfinic acid were investigated (Figure S2.25 for the chemical structures). Qualitatively, a trend was observed that $E_{\text{HOMO-n}}$ and the corresponding k_{O_3} increase similarly for different computation levels (Figures S2.26a-d). The results of $E_{\text{NBO, LP-S}}$ without dimethyl sulfoxide and methanesulfinate turned out to be similar to the one of $E_{\text{HOMO-n}}$ (Figures S2.26e-h). Note that dimethyl sulfoxide and methanesulfinate could not be included in the NBO model for the following reason. As shown in Figure S2.27 as an example, Aldicarb (sulfide) has two NBOs of the lone-pair electrons of sulfur, one (LP1) is the orbital resulting from mixing s and p atomic orbitals and another (LP2) purely comprises only of the p orbital. The latter one is much higher in energy, thus its $E_{\text{NBO, LP-S}}$ was used for the NBO model. While this is true for thiols, sulfides, and disulfides, dimethyl sulfoxide and methanesulfinate only have the LP1 orbital because the LP2 orbital is taken up to form a C-O π bond. Therefore, the E_{HOMO} seems to be the better descriptor than the $E_{\text{NBO, LP-S}}$ although both descriptors remain only qualitative.

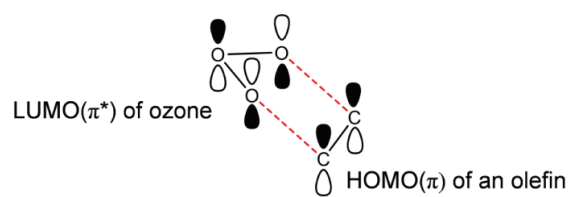
Text S2.9. Uncertainty (%) in estimating the elimination of micropollutants using the predicted k_{O_3}

Shown in Figure 2.7 is the extent of elimination of a micropollutant and the uncertainty (%) in predicting the elimination, plotted as a function of k_{O_3} , assuming an uncertainty of a factor of 4 in the predicted k_{O_3} . The % elimination was calculated using Eq. 2.1 in the main manuscript, $\left(1 - \frac{[P]}{[P]_0}\right) \times 100 = (1 - \exp(-k_{\text{O}_3, \text{app}} \int O_3 dt)) \times 100$, where the ozone exposure ($\int O_3 dt$) was set to 0.01 M s reported for wastewater effluents for an ozone dose of 1.0 g O₃/g DOC²⁹ and the hydroxyl radical oxidation is neglected in this case. The uncertainty (%) is derived by the difference between the eliminations obtained using $4k_{\text{O}_3, \text{app}}$ and $(1/4)k_{\text{O}_3, \text{app}}$.

As shown in Figure 2.7, three regions are defined. Region I has k_{O_3} between 6 M⁻¹s⁻¹ and 640 M⁻¹s⁻¹ with intermediate and high uncertainty (20–78%). The corresponding elimination range is between 6 and 99.8%. Region II has low uncertainty (1–20%) with a k_{O_3} range of 0.26–6 M⁻¹s⁻¹ and 640–1800 M⁻¹s⁻¹. The corresponding ranges of elimination are 0.27–5.9% and 99.8–100%, respectively. It is shown in this region that the error propagated by $4k_{\text{O}_3}$ or $(1/4)k_{\text{O}_3}$ becomes insignificant as full or no elimination is approached. Region III has less than 1 % uncertainty, which corresponds to a range of $k_{\text{O}_3} < 0.26 \text{ M}^{-1}\text{s}^{-1}$ and $> 1800 \text{ M}^{-1}\text{s}^{-1}$. There is no longer a noticeable error for either $4k_{\text{O}_3}$ or $(1/4)k_{\text{O}_3}$ because full or no elimination is achieved by $4k_{\text{O}_3}$ and $(1/4)k_{\text{O}_3}$, respectively.



Scheme S2.1. Procedure for k_{O_3} prediction using the developed *in silico* model. ^aA 3D initial geometry is obtained as the lowest energy conformer by the conformer module of Chemaxon MarvinSketch. ^bGroups with a geometry optimization requirement: mono- and di-alkoxybenzenes: an alkoxy group initially set to be planar followed by the optimization (Text S2.2); anilines: an amino group ought to be planar following the optimization (Text S2.3); trimethoxybenzenes: all methoxy groups initially set to be perpendicular to the aromatic ring followed by the optimization (Text S2.4). ^cNote that planar anilines may yield an imaginary frequency with the basis set containing polarization functions. HF/6-31G, which is the recommended method herein, is confirmed to yield no imaginary frequency for the selected planar anilines.



Scheme S2.2. The molecular orbital interaction between the HOMO of an olefin and the LUMO of ozone.

Supporting information for chapter 2

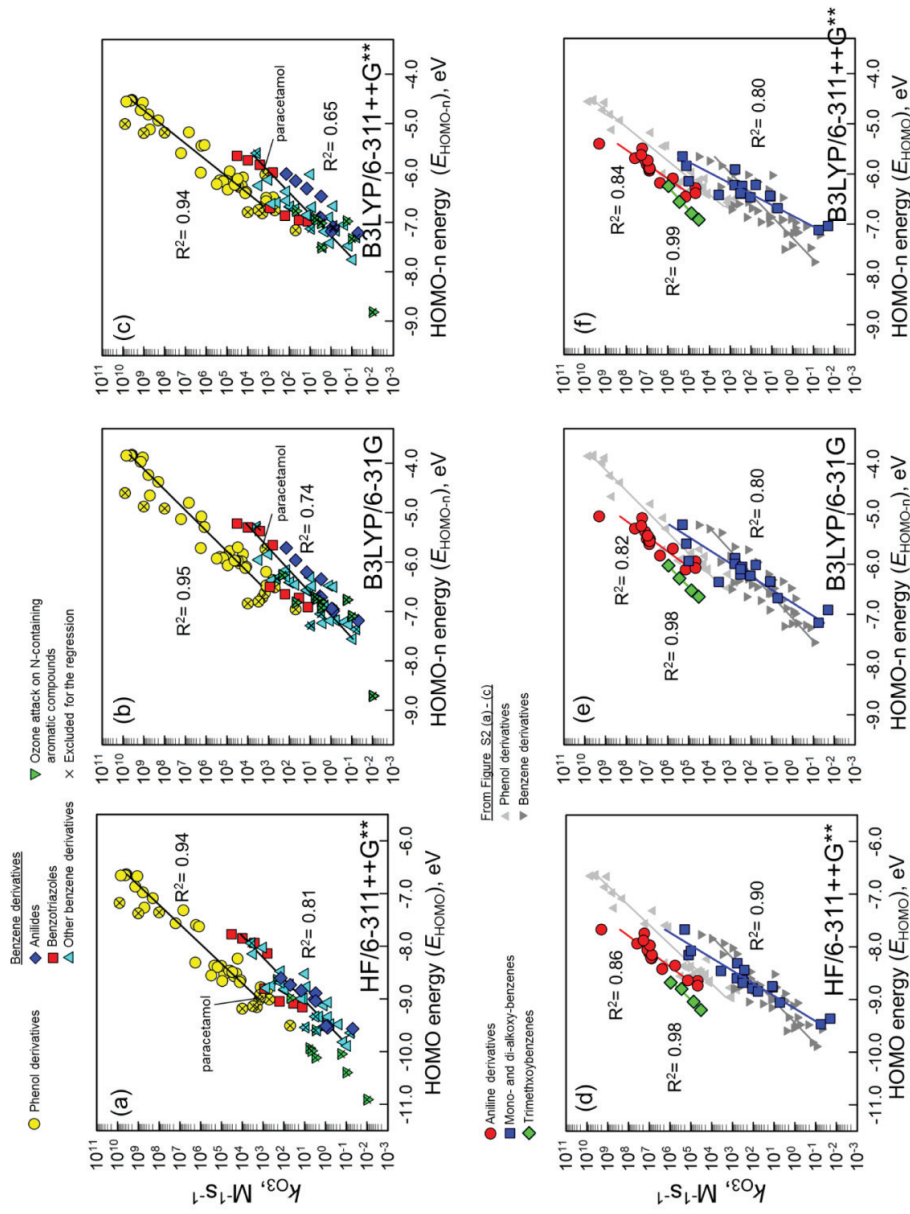


Figure S2.1. Correlations between the second-order rate constants (k_{O_3}) for the reaction of aromatic compounds with ozone and the corresponding HOMO-n energies (E_{HOMO-n}) ($n \geq 0$) with differing computation methods: (a) – (c) for phenol and benzene derivatives and (d) – (f) for anilines derivatives, mono- and di-alkoxybenzenes, and trimethoxybenzenes, respectively.

Phenol derivatives

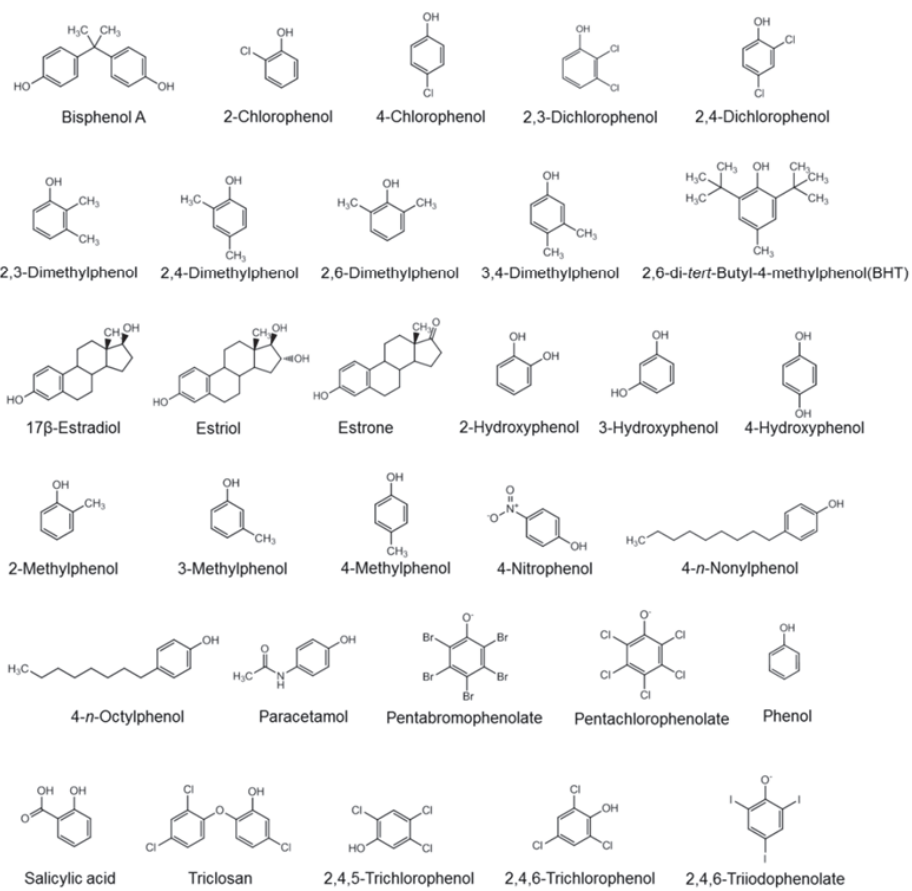
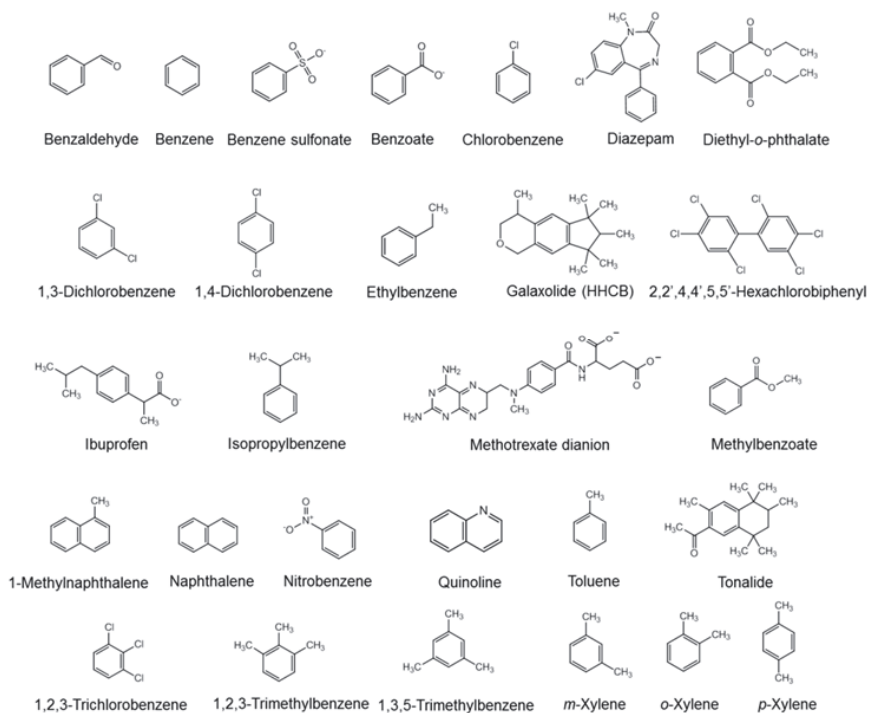


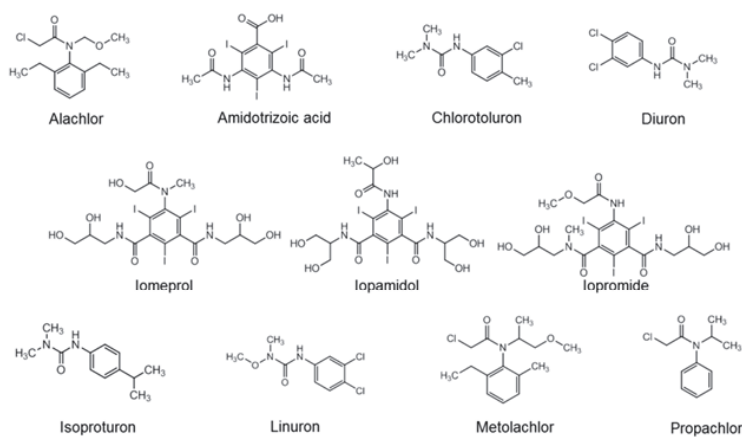
Figure S2.2. Chemical structures of the selected aromatic compounds.

Supporting information for chapter 2

Benzene derivatives



Benzene derivatives - Anilides



Benzene derivatives - Benzotriazoles

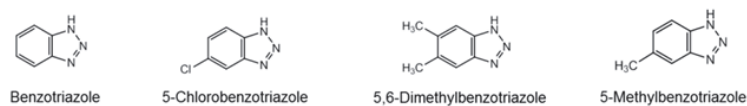
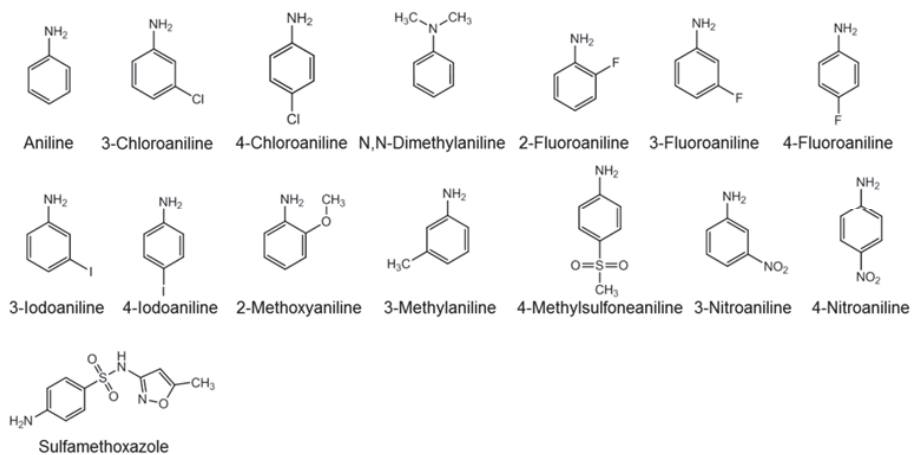
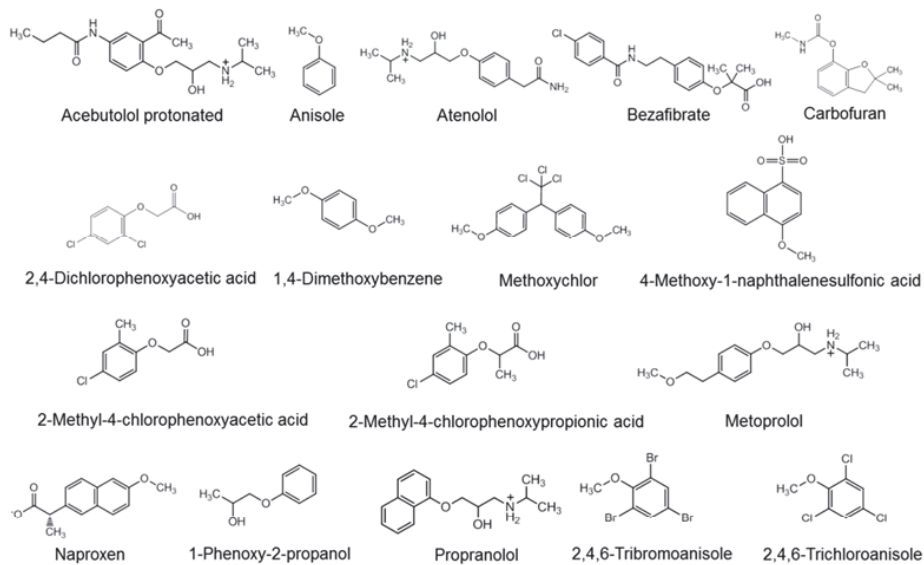


Figure S2.2. Chemical structures of the selected aromatic compounds (continued).

Aniline derivatives



Mono- and di-oxybenzenes



Trimethoxybenzenes

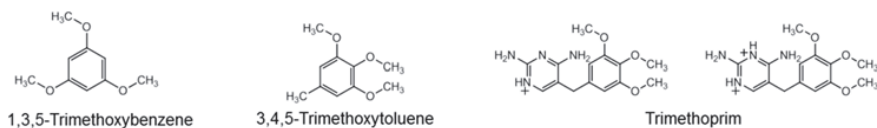


Figure S2.2. Chemical structures of the selected aromatic compounds (continued).

Ozone attack on nitrogen-containing aromatic compounds

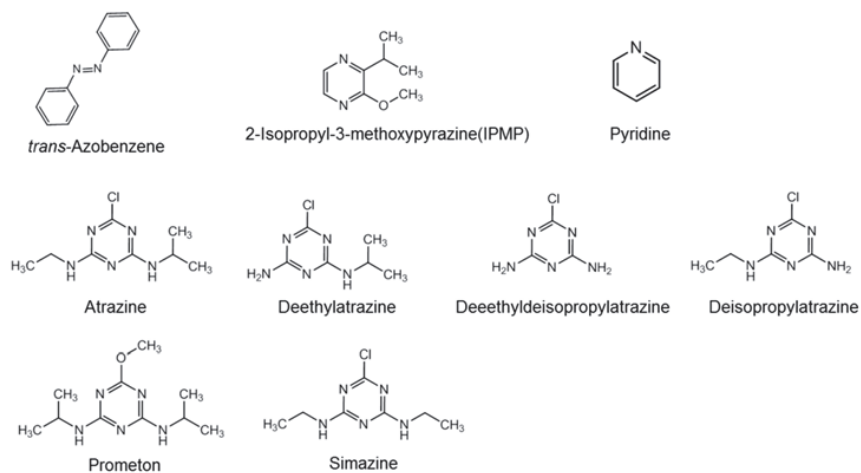


Figure S2.2. Chemical structures of the selected aromatic compounds (continued).

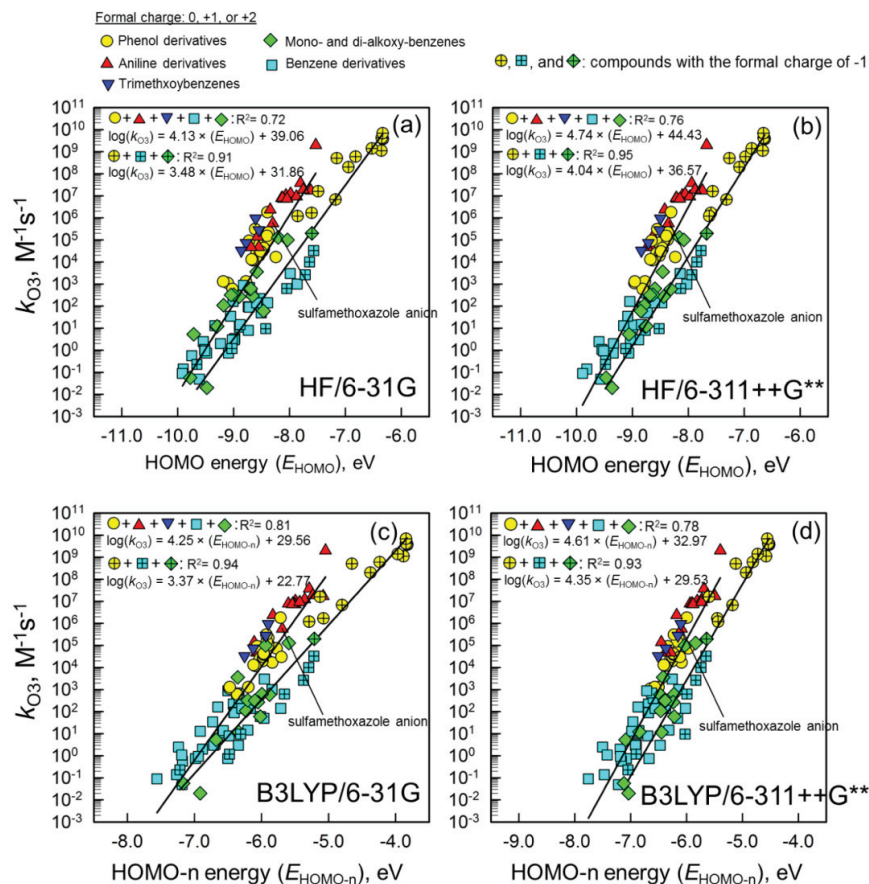


Figure S2.3. Comparison between the correlation of the model developed with aromatic compounds with the formal charge (FC) of 0, +1, or +2 and the correlations of the model developed with aromatic compounds with the FC of -1 at differing computation levels: (a) HF/6-31G, (b) HF/6-311++G**, (c) B3LYP/6-31G, and (d) B3LYP/6-311++G**. Sulfamethoxazole anion is not included in the anionic model as an exceptional case but in the model with the FC of 0, +1, or +2. The E_{HOMO} for the four trimethoxybenzenes which is higher in energy among respective conformers was chosen because their location is closer to the other aromatic compounds, leading to a better correlation for the model with the neutral and protonated aromatic compounds (see Text S2.4).

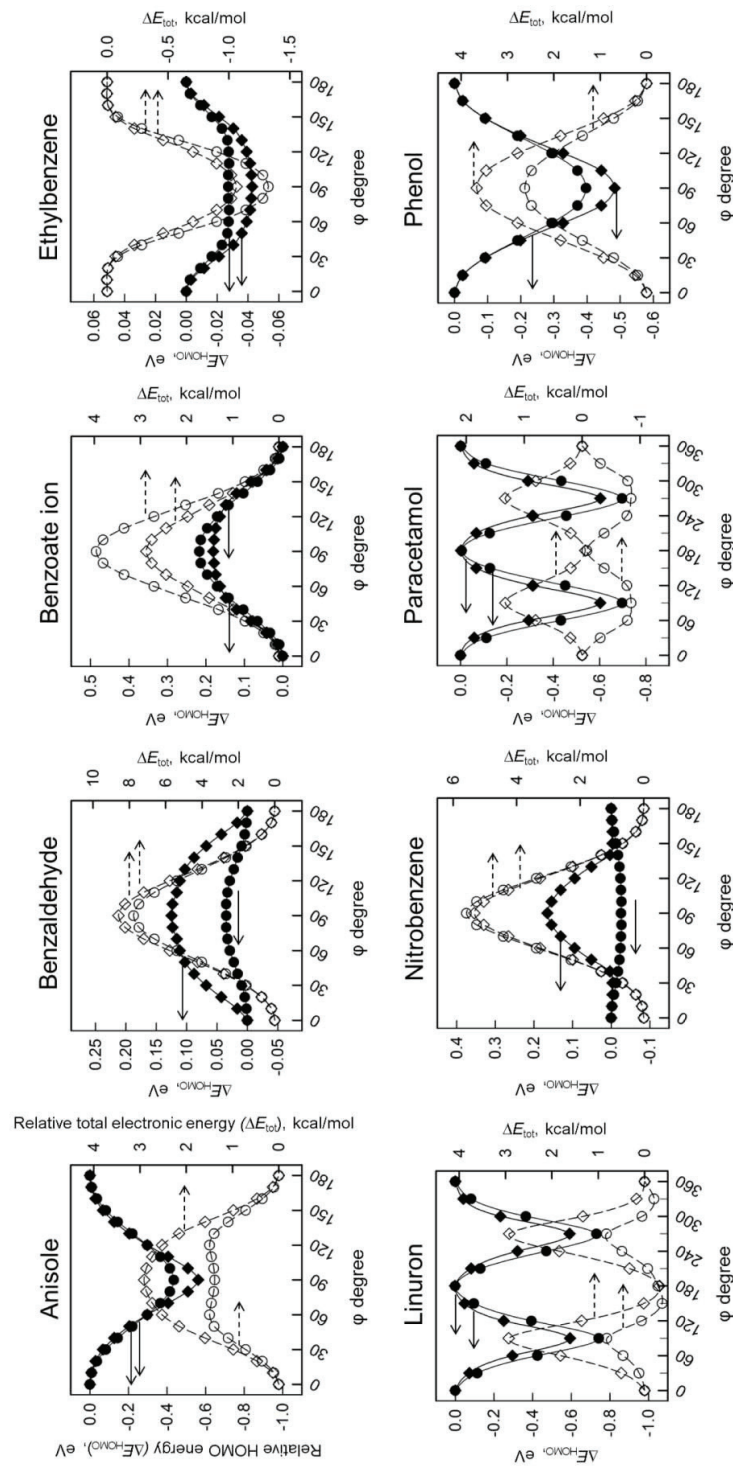


Figure S2.4. Analysis of the rotamers of 8 aromatic compounds: solid circles: relative HOMO energy (ΔE_{HOMO}) to the E_{HOMO} for the planar geometry with 0° of a dihedral angle (φ) for the HF/6-311++G** method, solid diamonds: ΔE_{HOMO} for the B3LYP/6-311++G** method, open circles: relative total electronic energy (ΔE_{tot}) to the E_{tot} for planar geometry with 0° of φ for the HF/6-311++G** method, and open diamonds: ΔE_{tot} for the B3LYP/6-311++G** method.

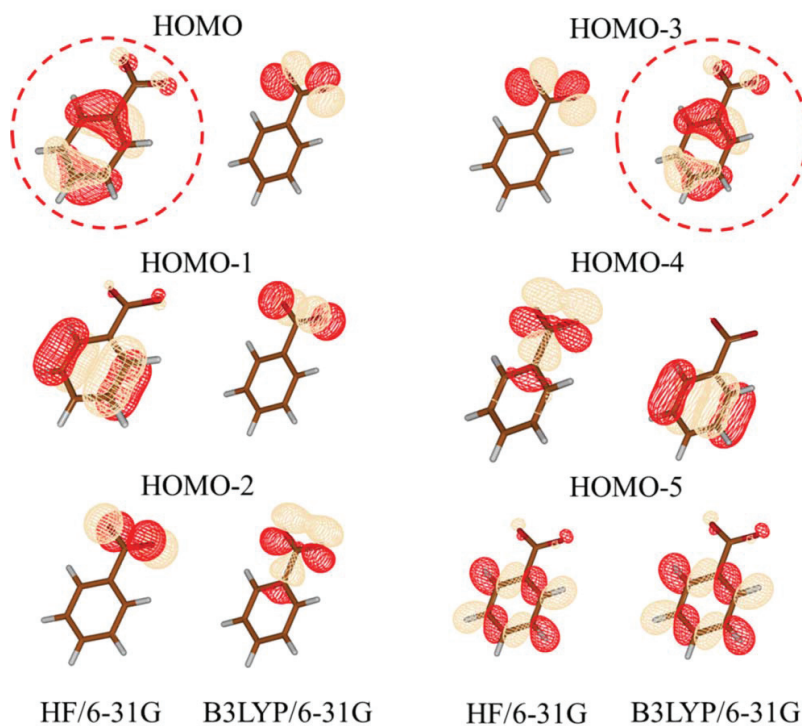


Figure S2.5. Comparison of HOMO to HOMO-5 for the benzoate ion for the HF/6-31G and the B3LYP/6-31G level. Molecular orbitals in the dashed circles (red) are the ones chosen for the model development.

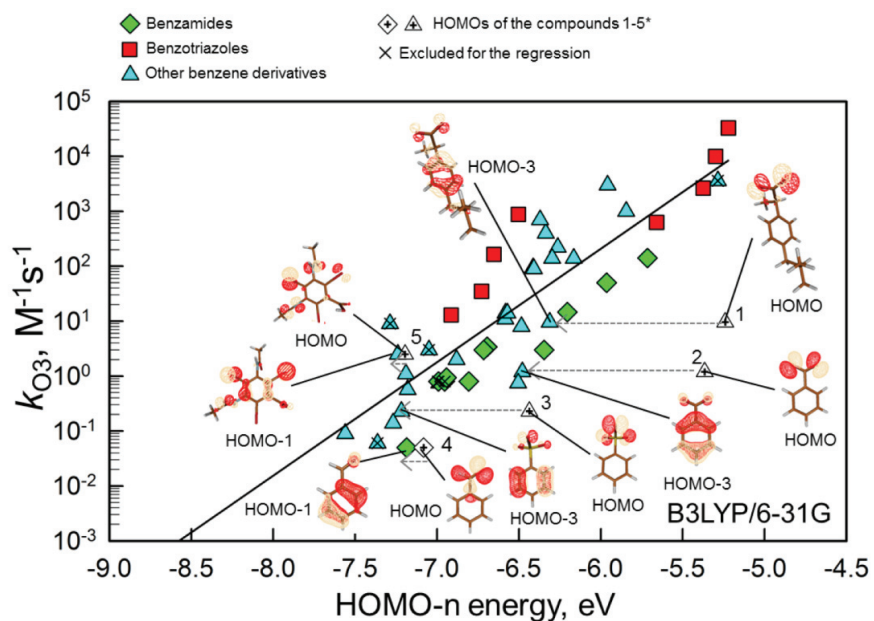


Figure S2.6. The effect of the replacement of HOMOs of aromatic compounds, which are not located on a ring, by HOMO-n values, which are instead located on the ring, on the correlation between the second-order rate constants (k_{O_3}) for the reaction of aromatic compounds with ozone and the corresponding HOMO-n energies. 1. ibuprofen, 2. benzoate ion, 3. benzene sulfonate ion, 4. benzaldehyde, 5. amidotrizoic acid. Computational method: B3LYP/6-31G.

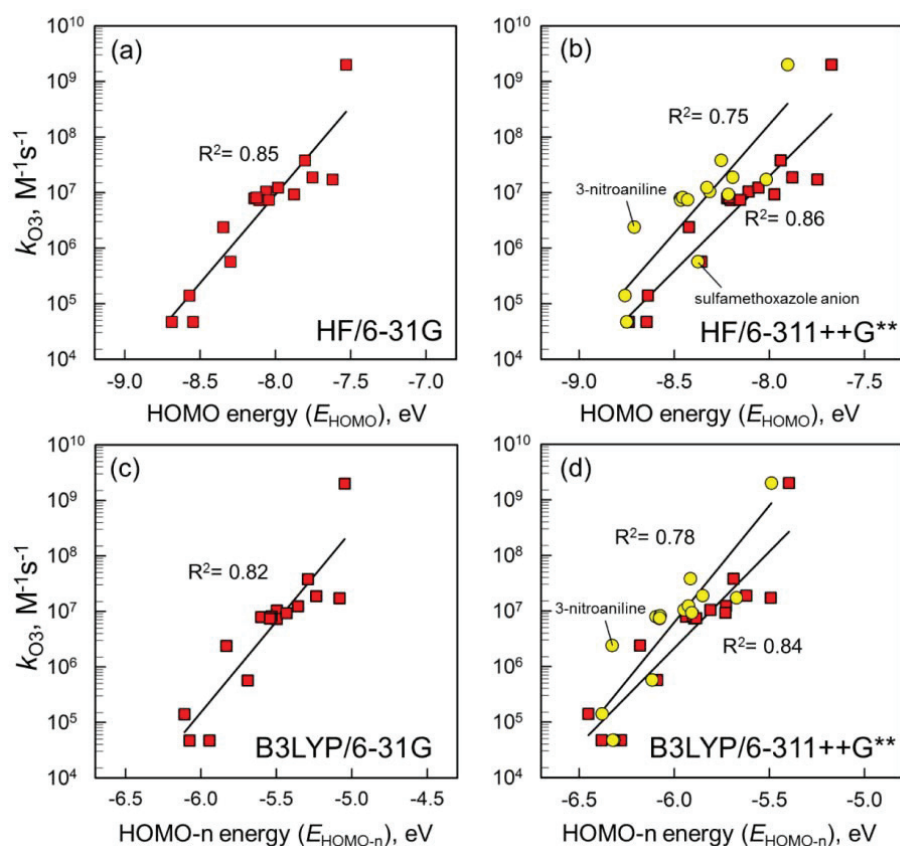


Figure S2.7. Comparison of HOMO and HOMO-n between planar-type anilines (red squares) and pyramidal-type anilines (yellow circles) for the correlation between second-order rate constants (k_{O_3}) for the reaction of anilines with ozone and the corresponding HOMO-n energies ($n \geq 0$) with the different computational methods: (a) HF/6-31G, (b) HF/6-311++G**, (c) B3LYP/6-31G, and (d) B3LYP/6-311++G**.

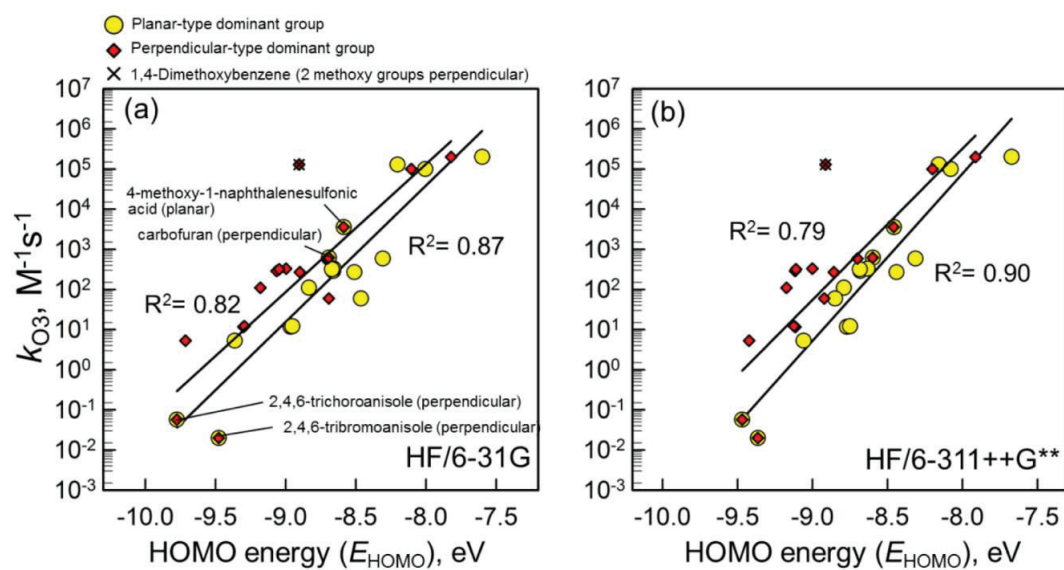


Figure S2.8. Comparison of correlations using the E_{HOMO} of the planar-type alkoxybenzenes (yellow circles) with the correlation using the E_{HOMO} of the perpendicular-type alkoxybenzenes (red diamonds): (a) HF/6-31G level and (b) HF/6-311++G** method.

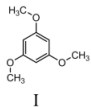
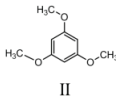
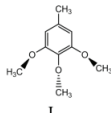
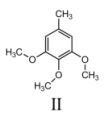
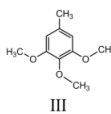
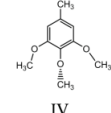
Initial geometry	Optimized geometry			
	HF/6-31G	HF/6-311++G**	B3LYP/6-31G	B3LYP/6-311++G**
<p><u>1,3,5-Trimethoxybenzene</u></p>  <p>I</p>  <p>II</p>	Same as the initial geometry			
<p><u>3,4,5-Trimethoxytoluene</u></p>  <p>I</p>  <p>II</p>  <p>III</p>  <p>IV</p>	Same as the initial geometry			

Figure S2.9. Initial and optimized structures of the selected rotamers of 1,3,5-trimethoxybenzene, 3,4,5-trimethoxytoluene, mono-, and di-protonated trimethoprim guided with wedges and dashes.

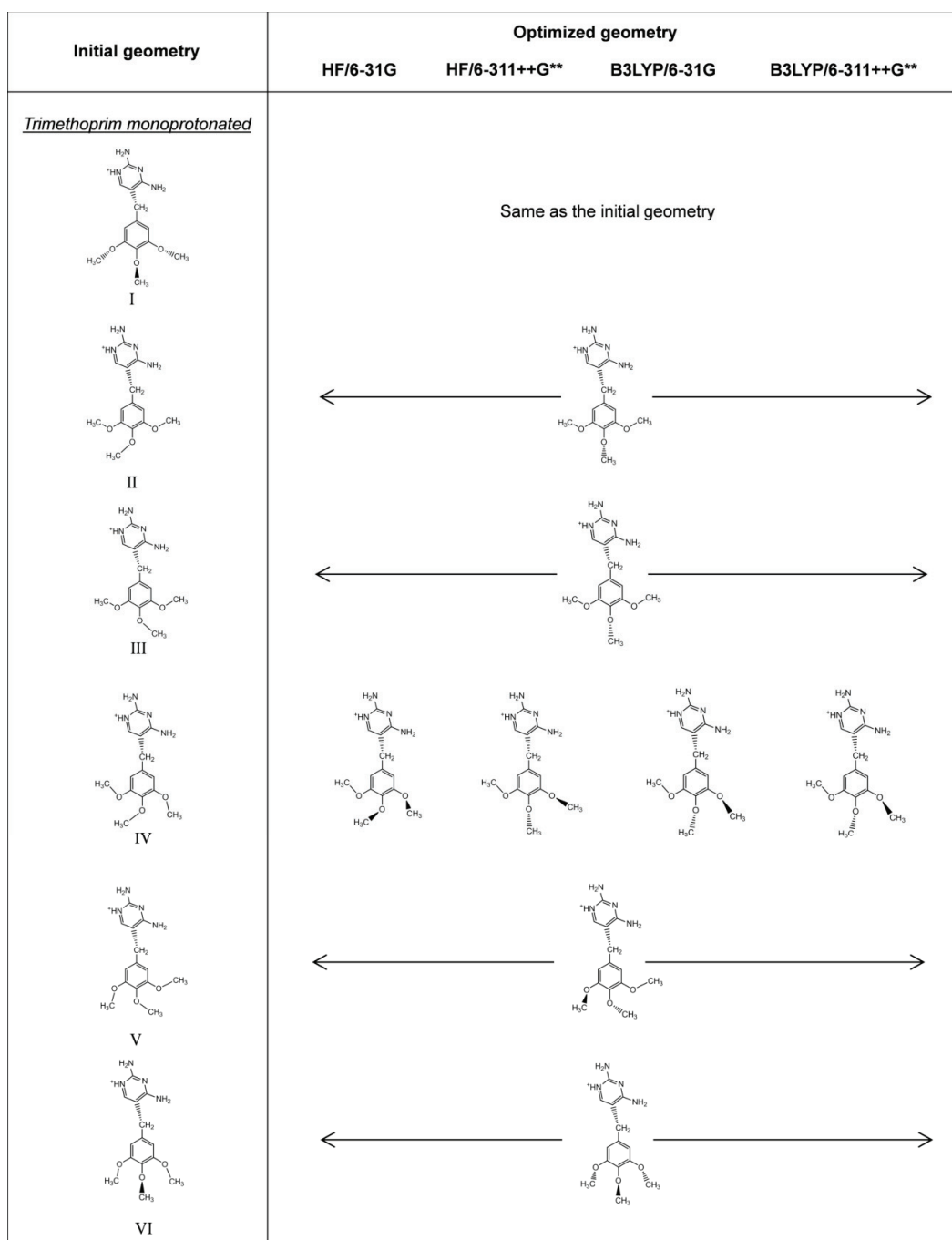


Figure S2.9. Initial and optimized structures of the selected rotamers of 1,3,5-trimethoxybenzene, 3,4,5-trimethoxytoluene, mono-, and di-protonated trimethoprim guided with wedges and dashes (continued).

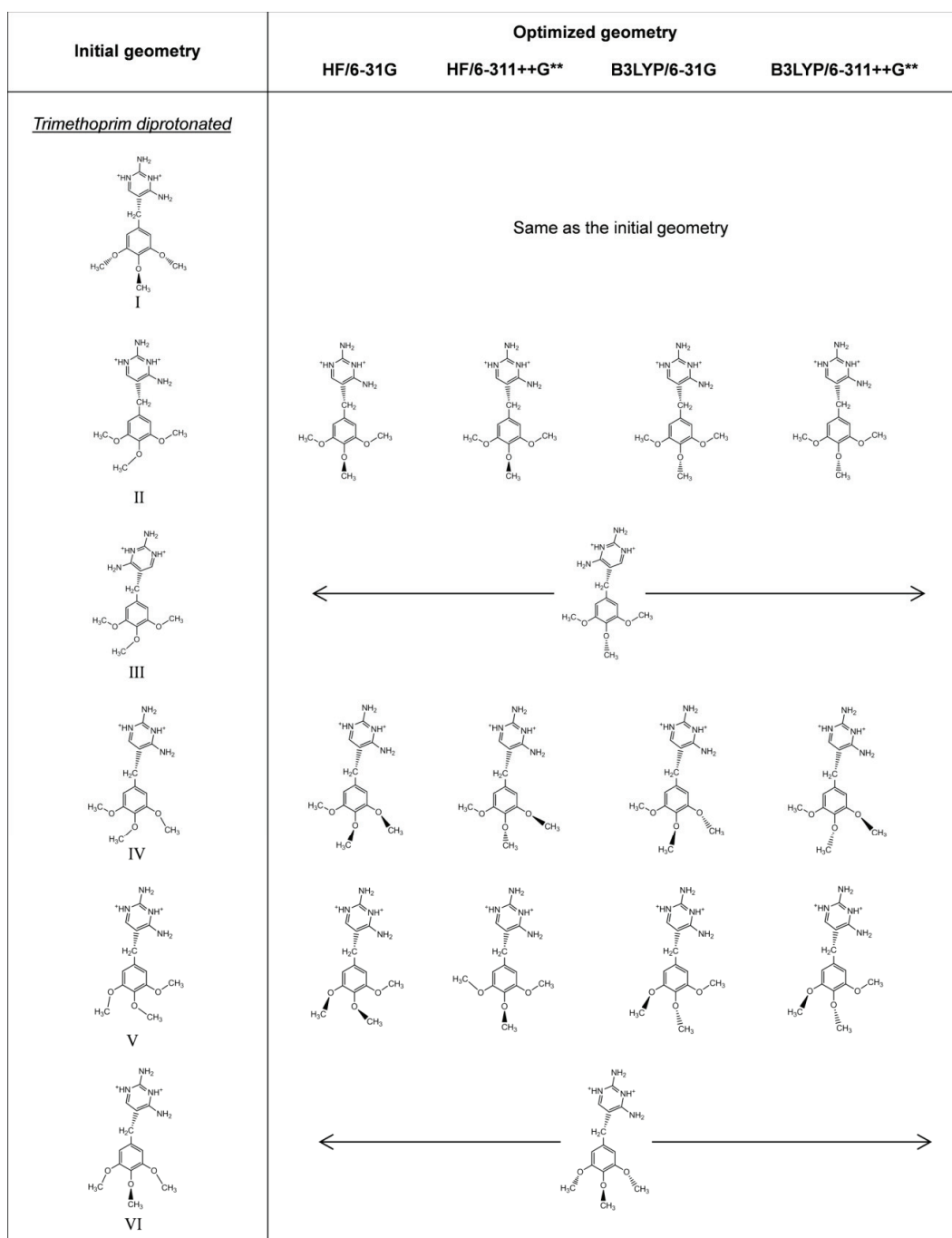


Figure S2.9. Initial and optimized structures of the selected conformers of 1,3,5-trimethoxybenzene, 3,4,5-trimethoxytoluene, mono-, and di-protonated trimethoprim guided with wedges and dashes (continued).

Supporting information for chapter 2

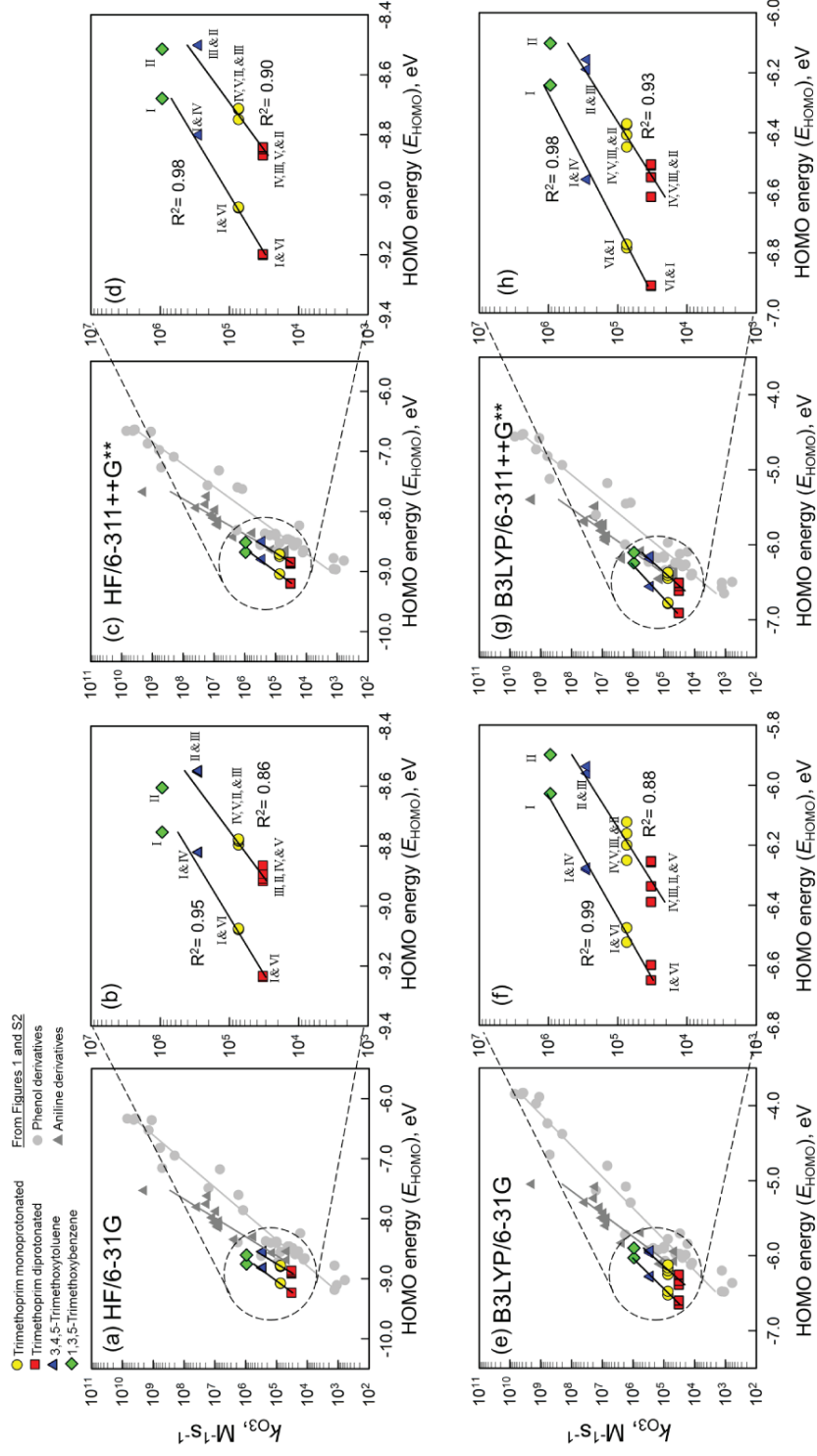


Figure S2.10. Correlations between the second-order rate constants (k_{03}) for the reaction of trimethoxybenzenes with ozone and the corresponding HOMO energies of their conformers calculated with different computational methods in comparison with the phenol derivatives and the anilines derivatives: (a) HF/6-31G, (b) the zoom-in of the trimethoxybenzenes for HF/6-311++G**, (c) HF/6-311++G**, (d) the zoom-in of the trimethoxybenzenes for HF/6-311++G**, (e) B3LYP/6-31G, (f) the zoom-in of the trimethoxybenzenes for B3LYP/6-31G, (g) B3LYP/6-311++G**, and (h) the zoom-in of the trimethoxybenzenes for B3LYP/6-311++G**. The roman numbers I to VI in Figures S2.10b, d, g, and h correspond to respective conformers shown in Figure S2.9.

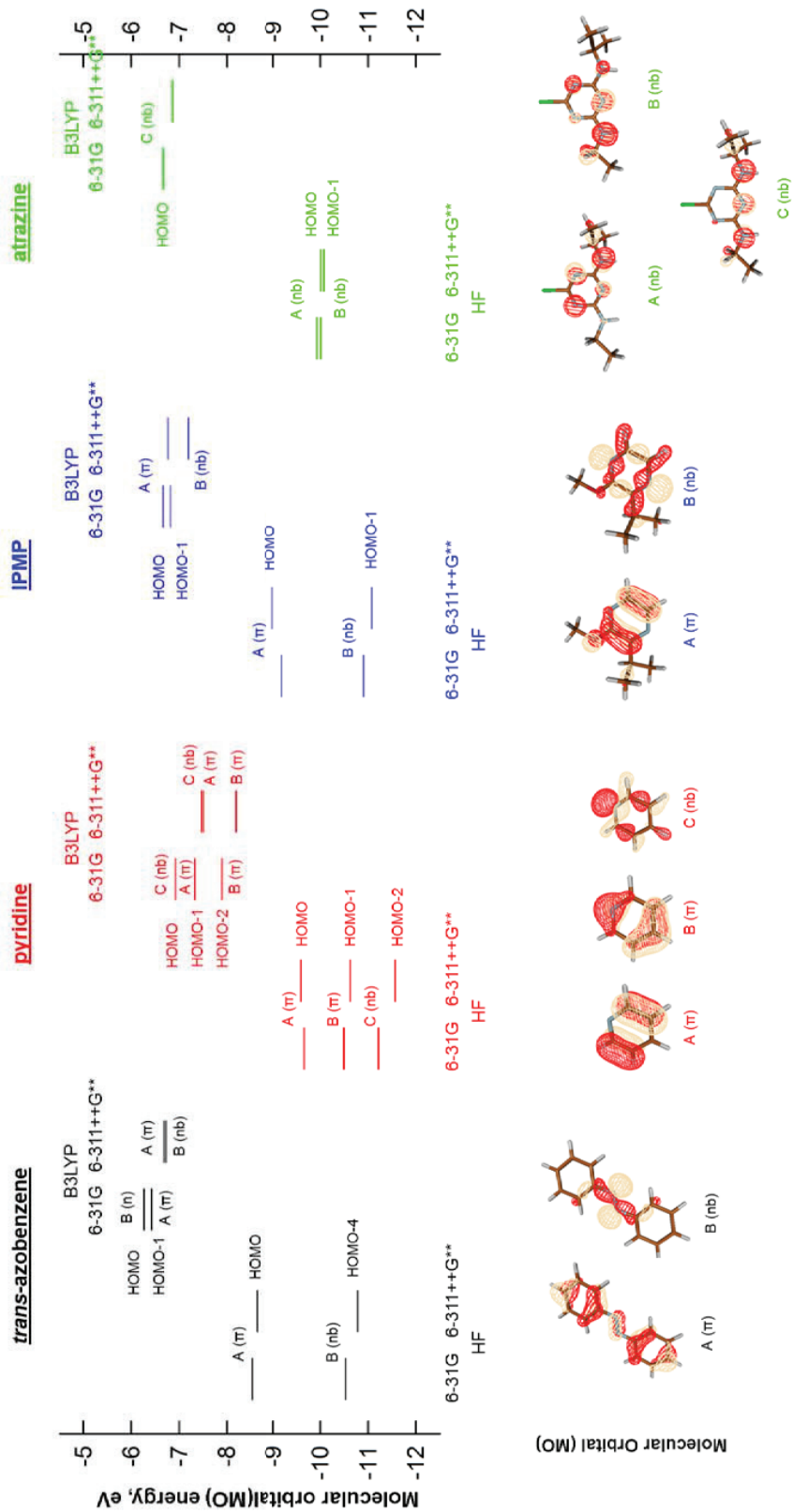


Figure S2.11. Selected molecular orbitals and the corresponding HOMO-n energy levels for ozone-attacked nitrogen-containing aromatic compounds with the different computational methods.

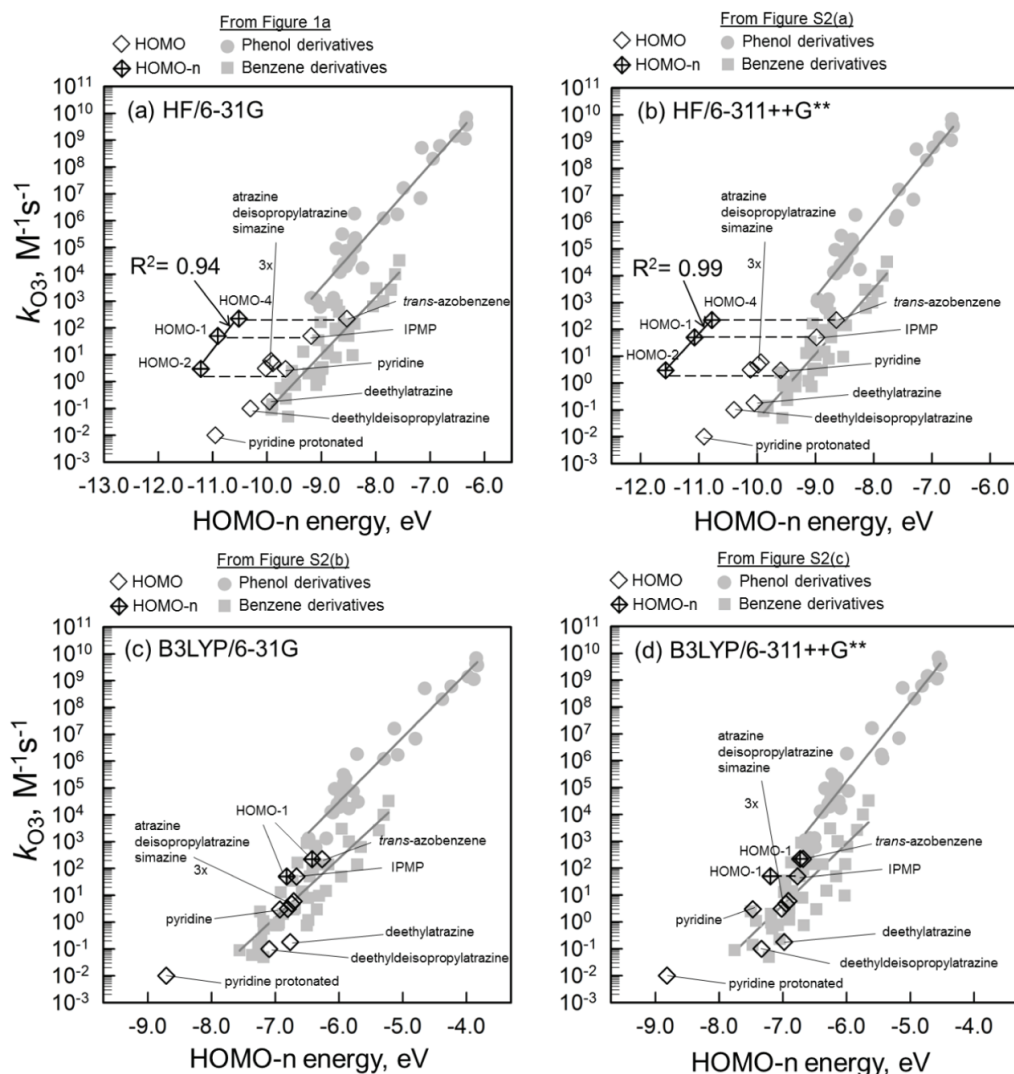
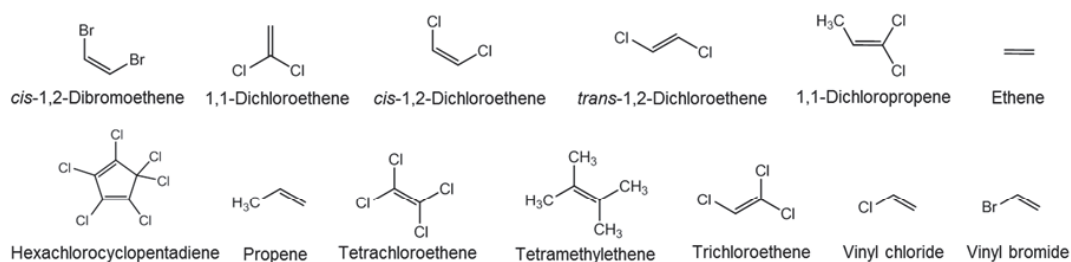


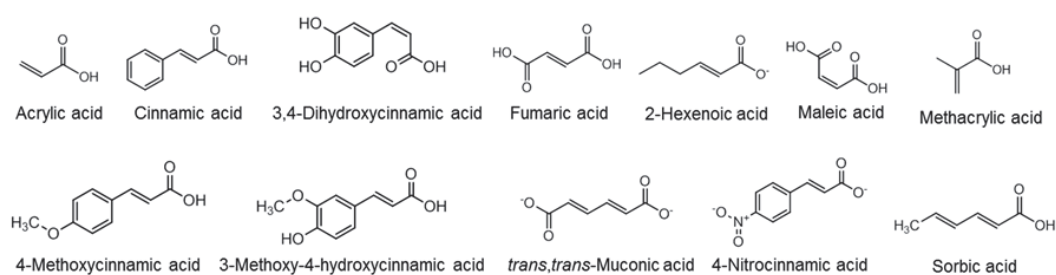
Figure S2.12. Correlations between the second-order rate constants (k_{O_3}) for the reaction of ozone-attacked nitrogen-containing aromatic compounds with ozone and the corresponding HOMO-n energies (E_{HOMO-n}) ($n \geq 0$) in comparison with phenol derivatives and benzene derivatives with the different computational methods: (a) HF/6-31G, (b) HF/6-311++G**, (c) B3LYP/6-31G, and (d) B3LYP/6-311++G**.

Olefins by substituents

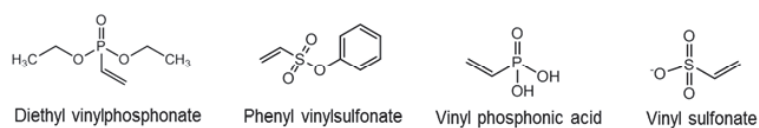
Methyl group or halogen



Carboxylic acid



Phosphonic acid or sulfonic acid



Electron-withdrawing group including oxygen or nitrogen

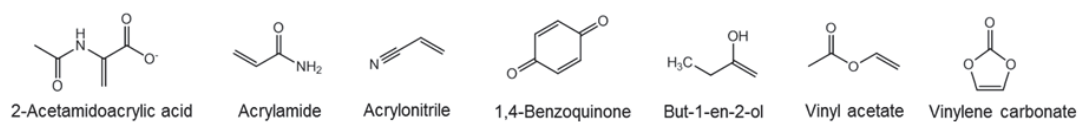


Figure S2.13. Chemical structures of the selected aliphatic olefins.

Micropollutant

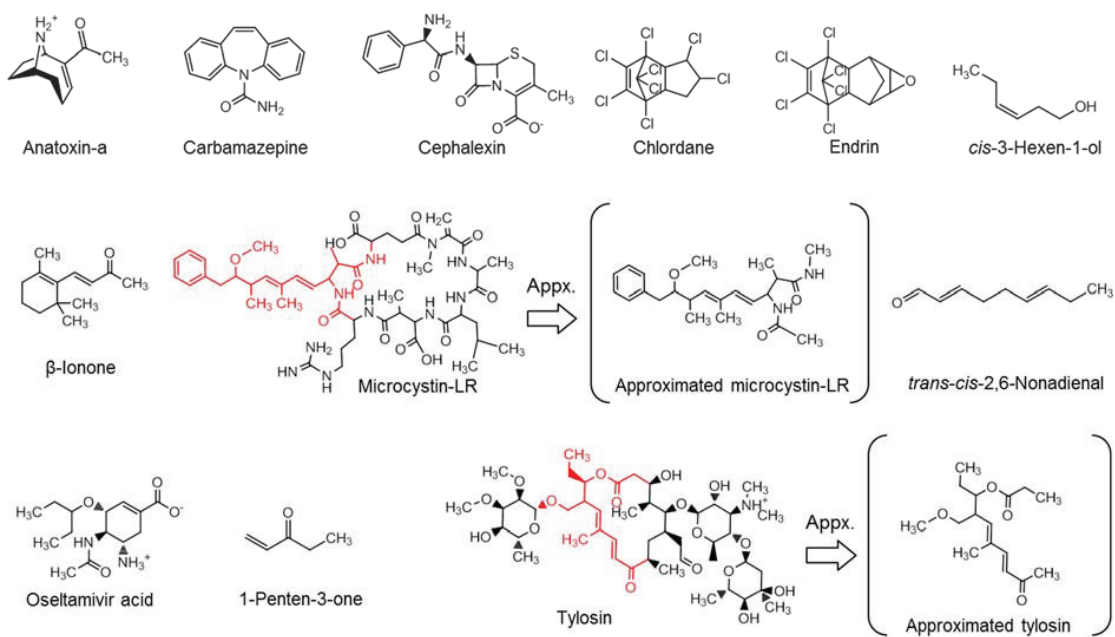


Figure S2.13. Chemical structures of the selected aliphatic olefins (continued).

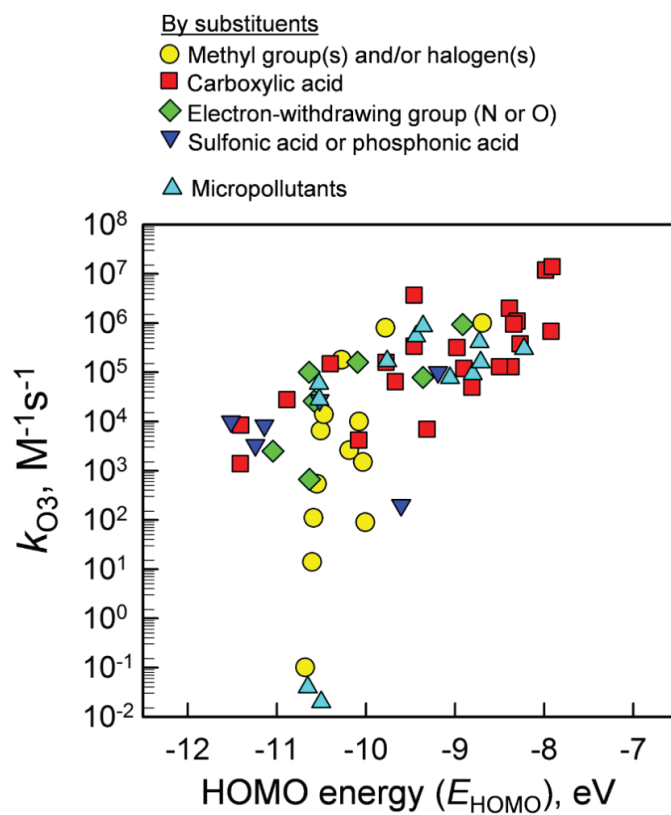


Figure S2.14. Correlations between the second-order rate constants (k_{O_3}) for the reaction of 60 olefins with ozone and the corresponding HOMO energies. Computational method: HF/6-31G. The data for other computational methods are not shown because no substantial difference was observed.

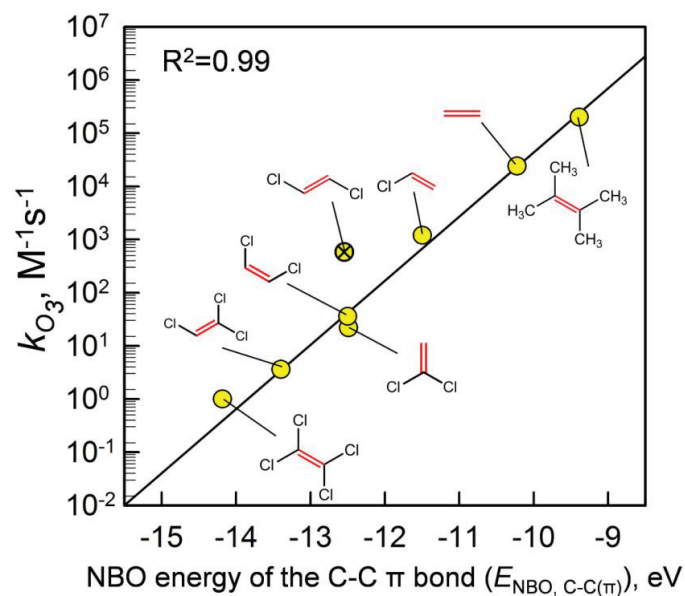


Figure S2.15. Correlations between the second-order rate constants (k_{O_3}) for the reaction of 8 olefins with ozone in carbon tetrachloride (CCl_4) and the corresponding NBO energies of the carbon-carbon π bond ($E_{NBO, C-C(\pi)}$) calculated with the HF/6-31G method using the NBO 3.1 package with the solvation model set to CCl_4 . Selected olefins (k_{O_3} , $M^{-1}s^{-1}$): 1,1-dichloroethene (1.1×10^2)³⁰, *cis*-1,2-dichloroethene (36)³⁰, *trans*-1,2-dichloroethene (5.9×10^2)³⁰, ethene (2.4×10^4)³¹, tetrachloroethene (1)³⁰, tetramethylethene (2.0×10^5)³¹, trichloroethene (3.6)³⁰, vinyl chloride (1.2×10^3)³⁰ (crossed symbol: excluded from the regression as an outlier).

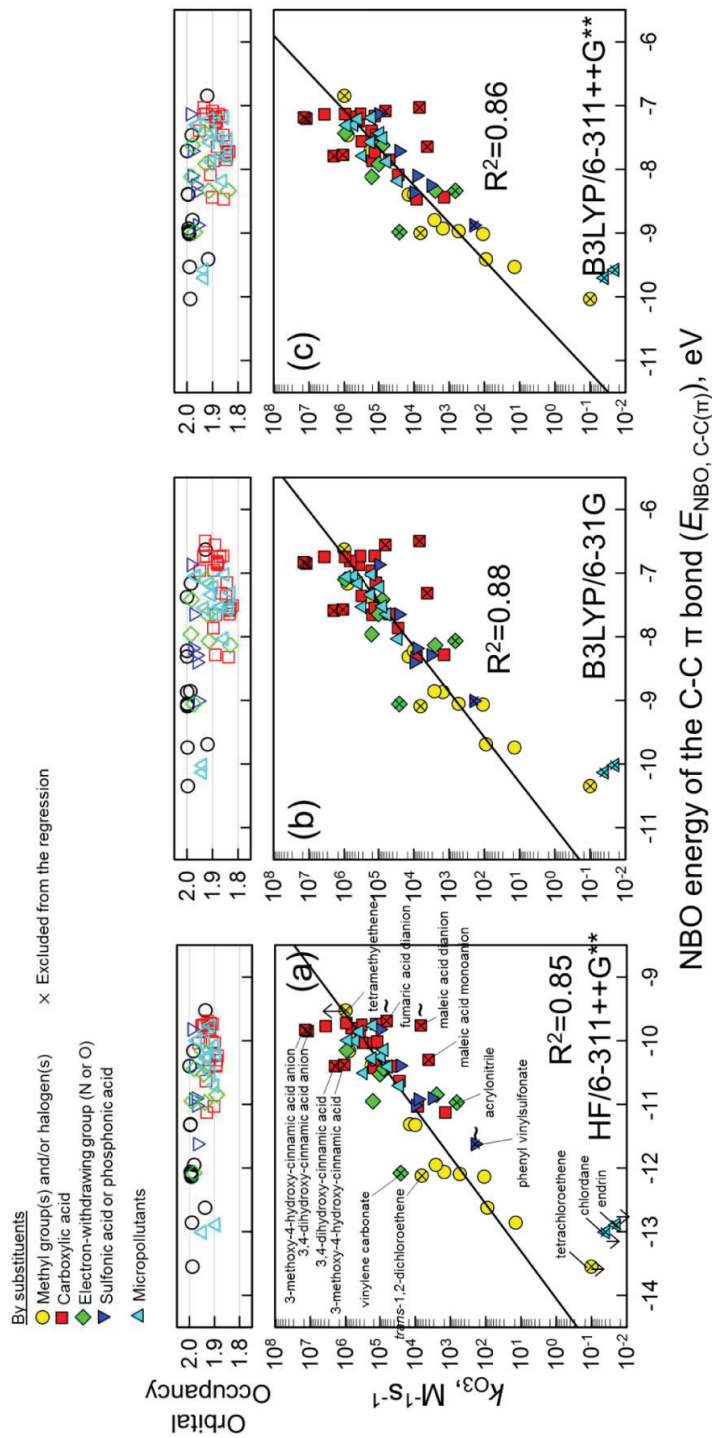
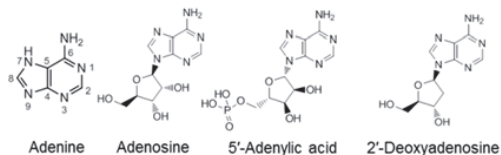


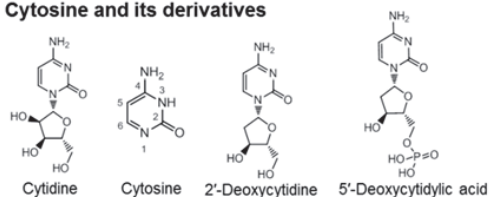
Figure S2.16. Correlations between the second-order rate constants (k_{O_3}) for the reaction of the selected olefins with ozone and the corresponding NBO energies of the carbon-carbon π bond ($E_{\text{NBO, C-C}(\pi)}$) calculated with the different computational methods using the NBO 3.1 package: (a) HF/6-311++G**, (b) B3LYP/6-31G, and (c) B3LYP/6-311++G**. The orbital occupancies (full occupancy = 2) of the NBO of the olefins are shown in the upper Figures.

Nucleic acid constituents and its derivatives

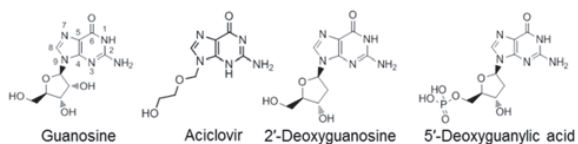
Adenine and its derivatives



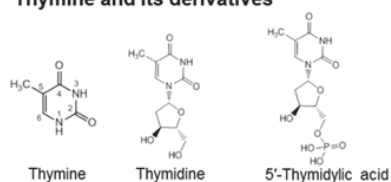
Cytosine and its derivatives



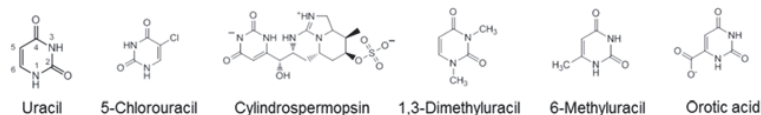
Guanosine and its derivatives



Thymine and its derivatives

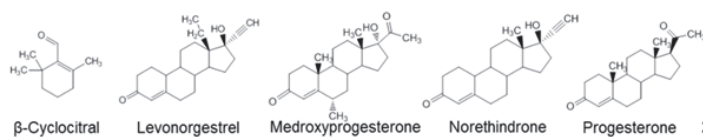


Uracil and its derivatives

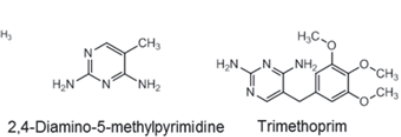


Miscellaneous olefins

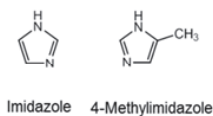
Cyclobenzenones



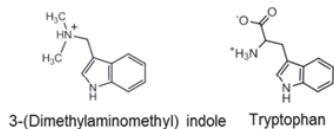
Diaminopyrimidines



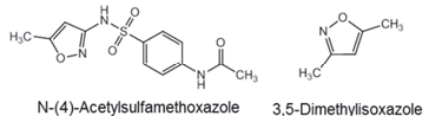
Imidazoles



Indoles



Isoxazoles



Quinolone

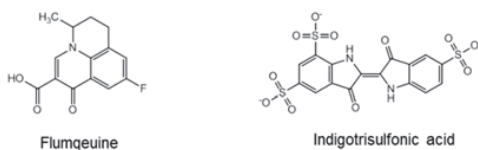


Figure S2.17. Chemical structures of the selected nucleic acid constituents, their derivatives, and miscellaneous olefins.

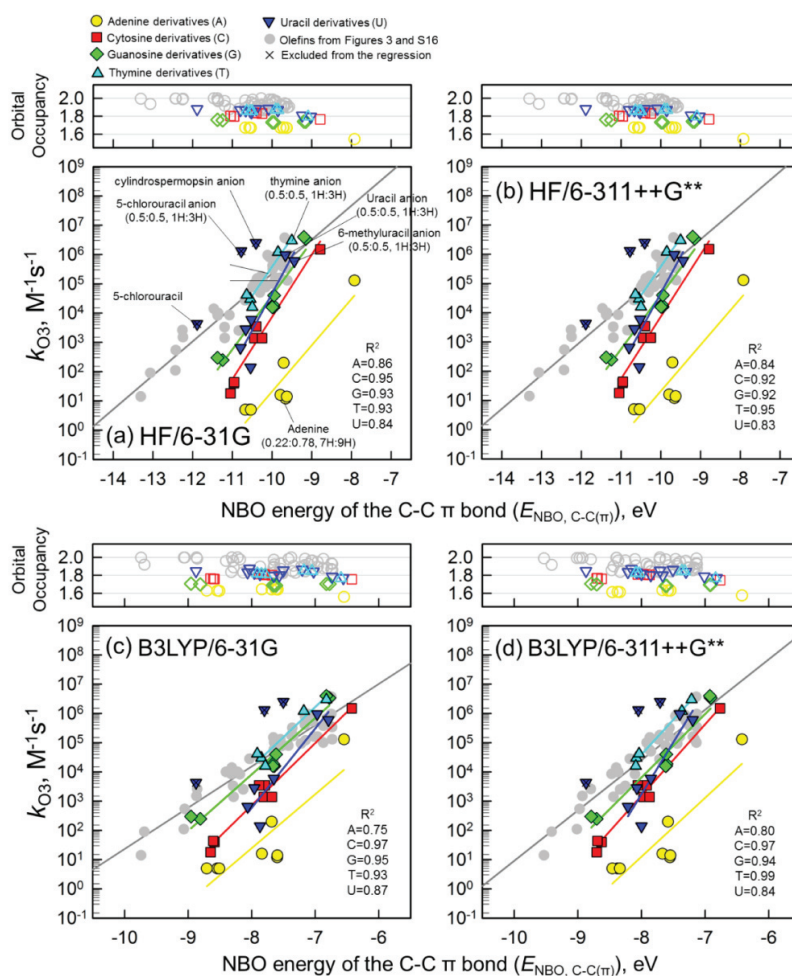


Figure S2.18. Correlations between the second-order rate constants (k_{O_3}) for the reaction of nucleic acid constituents with ozone and the corresponding NBO energies of the carbon-carbon π bond ($E_{NBO, C-C(\pi)}$) calculated at different levels of theory using the NBO 3.1 package in comparison with the data of the olefins (grey) from Figure 2.3 in the main manuscript: (a) HF/6-31G, (b) HF/6-311++G^{**}, (c) B3LYP/6-31G, and (d) B3LYP/6-311++G^{**}. The orbital occupancies (full occupancy = 2) of the NBO of the nucleic acid constituents are shown in the upper Figures.

Supporting information for chapter 2

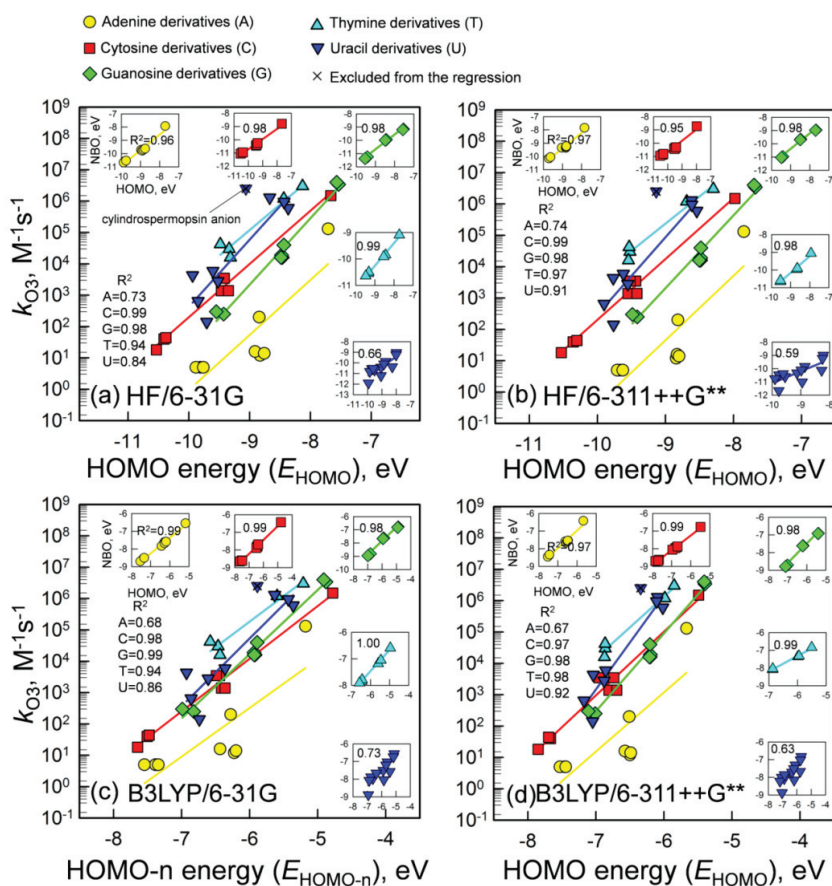


Figure S2.19. Correlations between the second-order rate constants (k_{O_3}) for the reaction of the selected nucleic acid constituents with ozone and the corresponding HOMO-n energies (E_{HOMO-n}) computed at different methods: (a) HF/6-31G, (b) HF/6-311++G**, (c) B3LYP/6-31G, and (d) B3LYP/6-311++G**. The insets are the correlations between the E_{NBO} , $C-C(\pi)$ and the E_{HOMO-n} of the nucleic acid constituents. Anionic cylindrospermopsin (crossed) was excluded as an outlier.

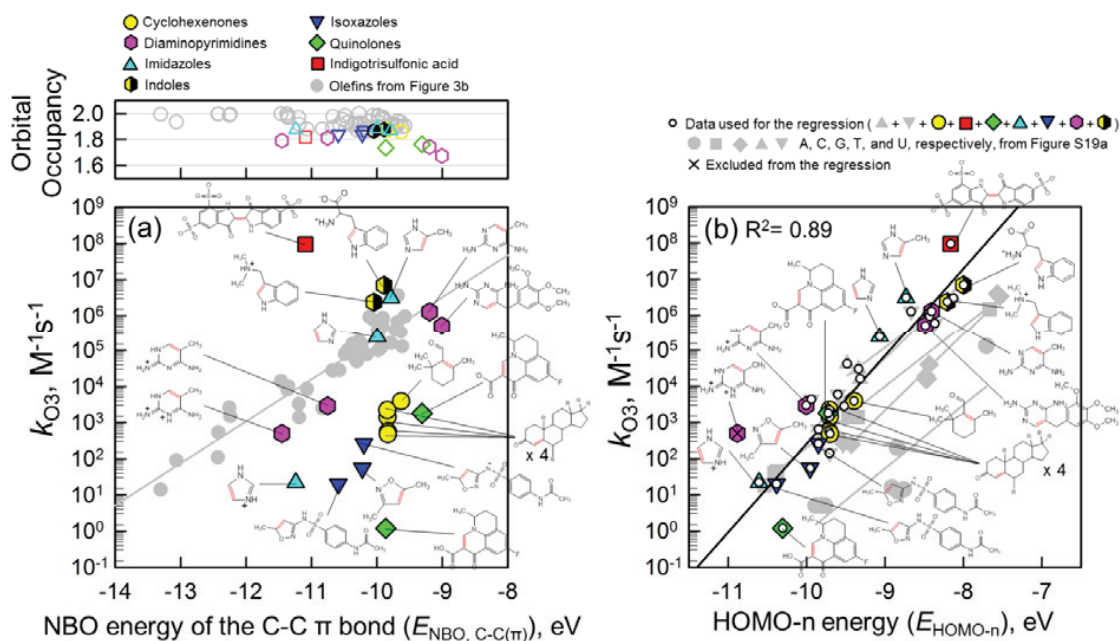


Figure S2.20. Correlations between the second-order rate constants (k_{O_3}) for the reaction of miscellaneous olefins with ozone and (a) the corresponding NBO energies of the carbon-carbon π bond ($E_{\text{NBO, C-C}(\pi)}$) computed with the HF/6-31G method using the NBO 3.1 package in comparison with the olefins from Figure 2.3b and (b) the corresponding HOMO-n energies ($E_{\text{HOMO-n}}$) calculated at the HF/6-31G level in comparison with the nucleic acid constituents from Figure S2.16a. The orbital occupancies (full occupancy = 2) of the miscellaneous olefins are shown in the upper Figure (a).

Supporting information for chapter 2

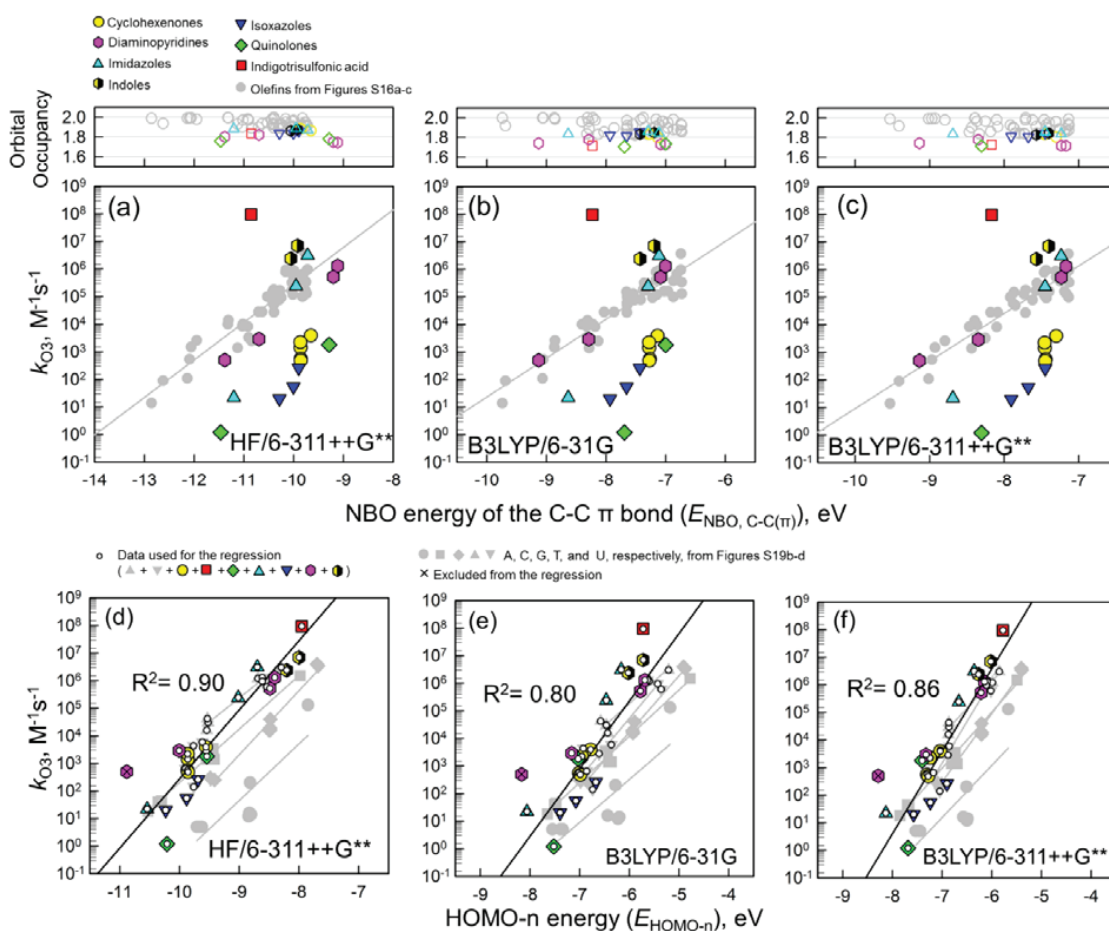


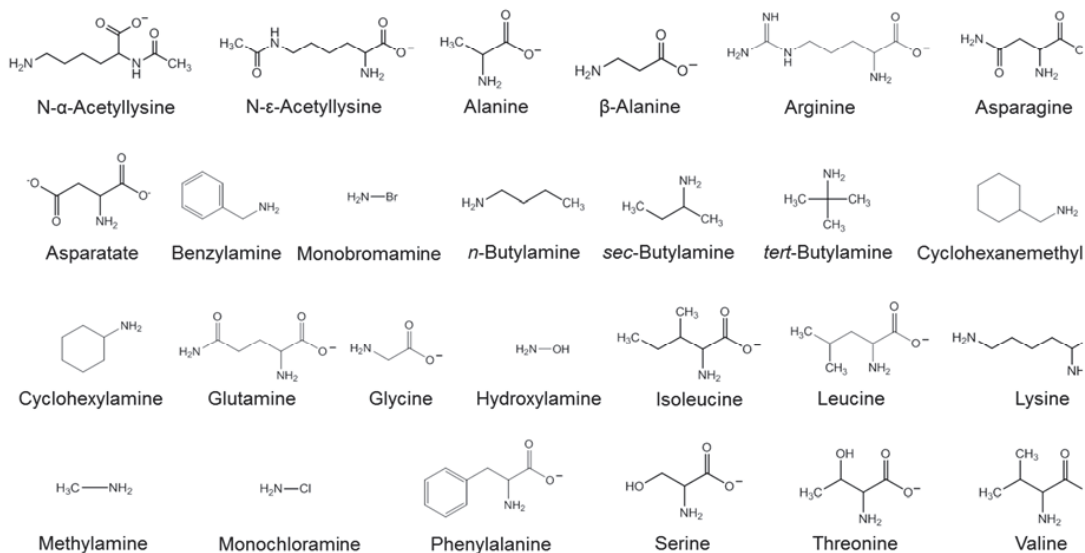
Figure S2.21. Correlations between the second-order rate constants (k_{O_3}) for the reaction of the selected miscellaneous olefins with ozone and (a-c) the corresponding NBO energies of the carbon-carbon π bond ($E_{NBO, C-C(\pi)}$) determined with different computational methods using the NBO 3.1 package in comparison with the olefins from Figure S2.14 and (d-f) the corresponding HOMO-n energies (E_{HOMO-n}), further compared with the ones of the nucleic acid constituents from Figure S2.17. The orbital occupancies (full occupancy = 2) of the miscellaneous olefins are shown in the upper Figures (a-c) in comparison with olefins from Figure S2.17.

Amines



Ammonia

Primary amines



Secondary amines

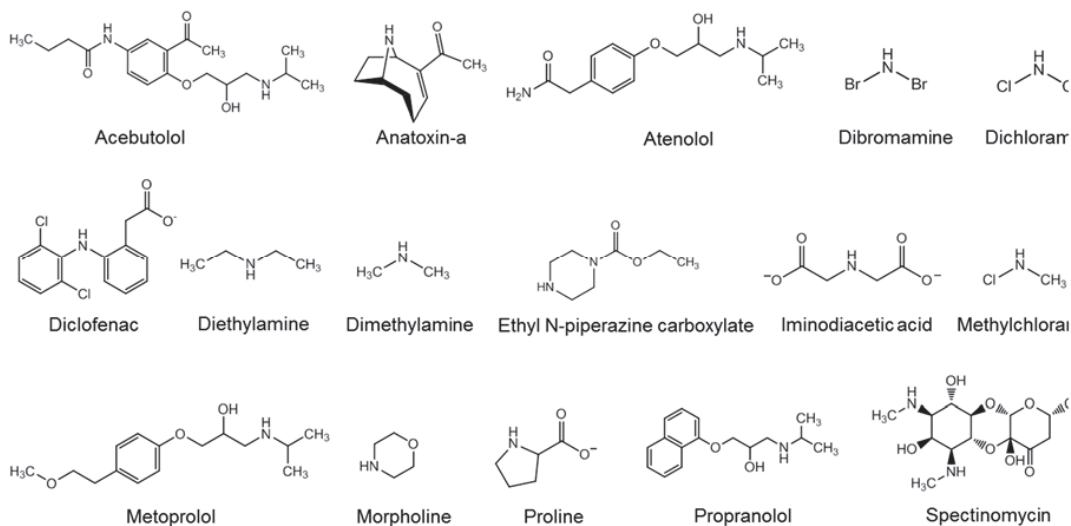


Figure S2.22. Chemical structures of the selected aliphatic amines.

Tertiary amines

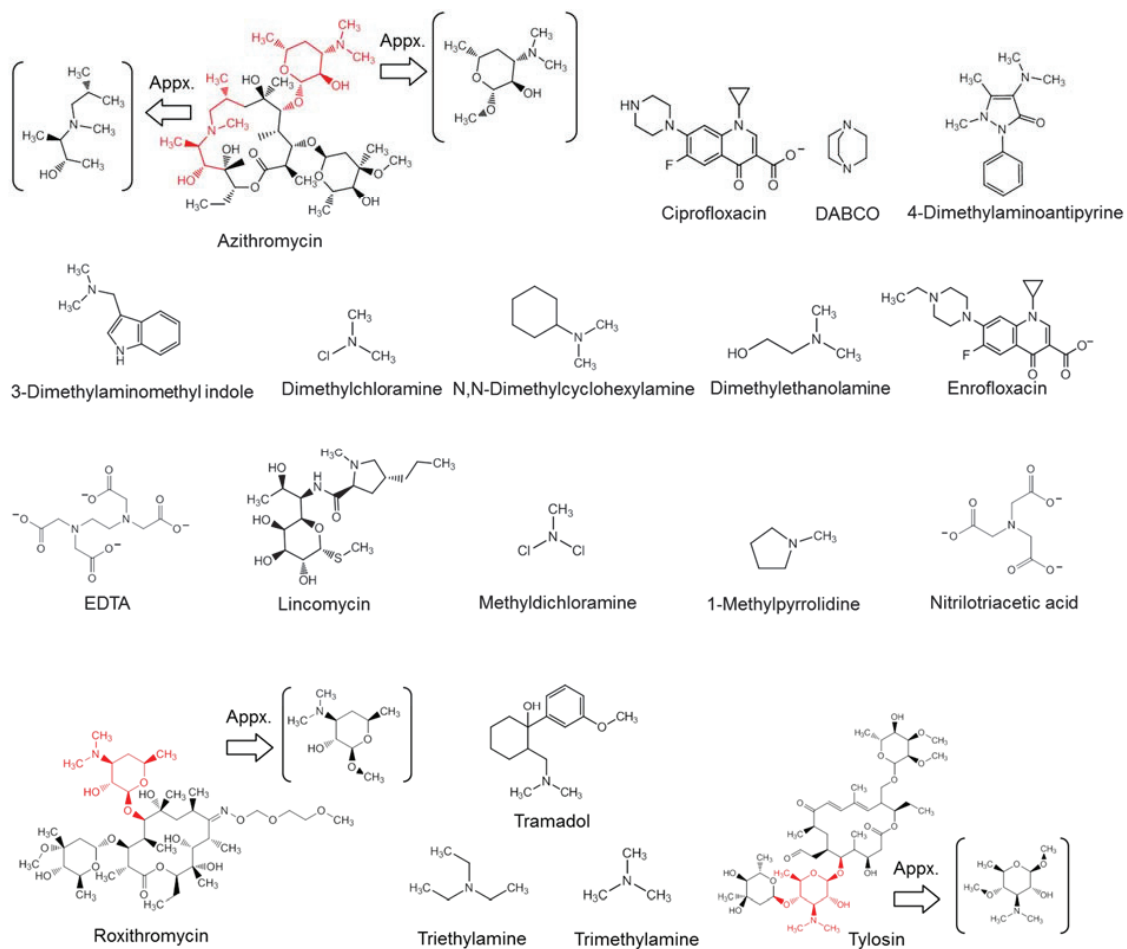


Figure S2.22. Chemical structures of the selected aliphatic amines (continued).

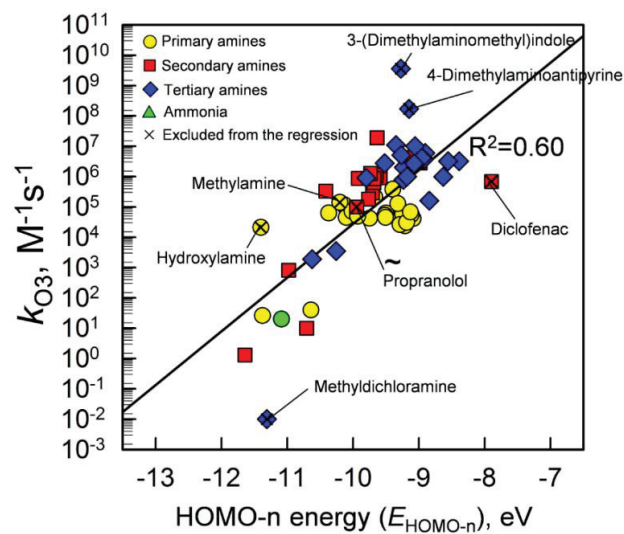


Figure S2.23. Correlation between the second-order rate constants (k_{O_3}) for the reaction of the selected aliphatic amines with ozone and the corresponding HOMO-n energies (E_{HOMO-n}) determined with the HF/6-31G method. Crossed symbols: excluded as outliers.

Supporting information for chapter 2

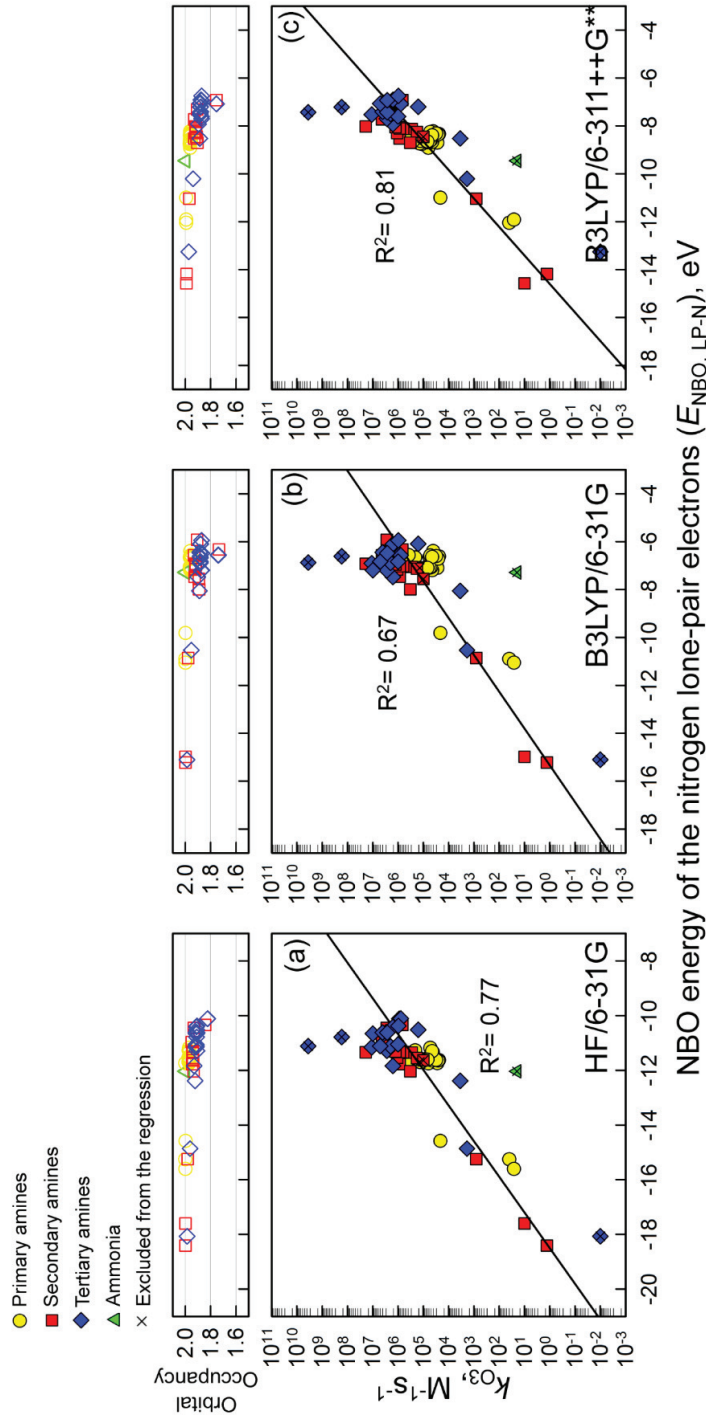
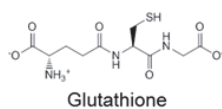
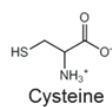
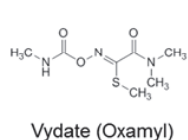
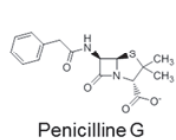
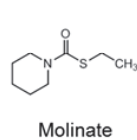
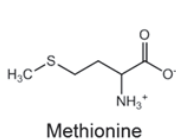
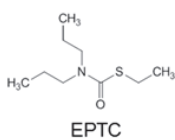
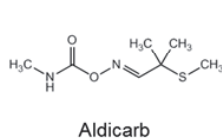


Figure S2.24. Correlations between the second-order rate constants (k_{O_3}) for the reaction of the selected aliphatic amines with ozone and the corresponding NBO energies of the nitrogen lone-pair electrons ($E_{NBO, LP-N}$) determined with different computational methods using the NBO 3.1 package: (a) HF/6-31G, (b) B3LYP/6-31G, and (c) B3LYP/6-311++G**. The orbital occupancies (full occupancy = 2) of the amines are shown in the upper figures (S24a-c).

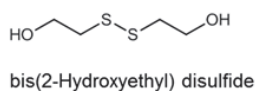
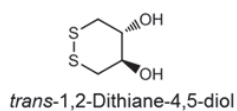
Thiols



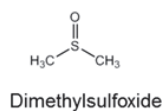
Sulfides



Disulfides



Sulfoxide



Sulfinic acid

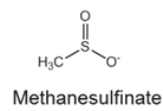


Figure S2.25. Chemical structures of the selected sulfur-containing compounds.

Supporting information for chapter 2

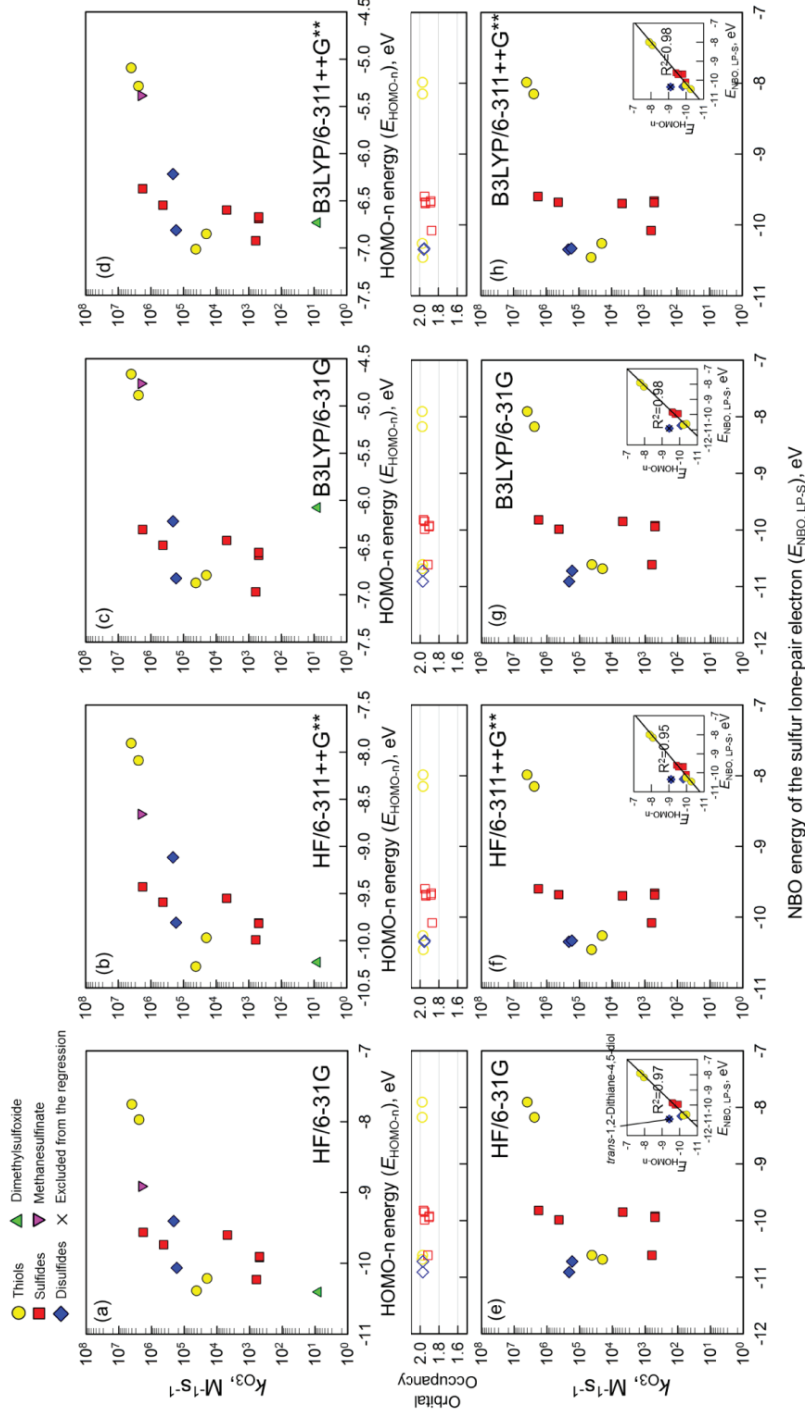
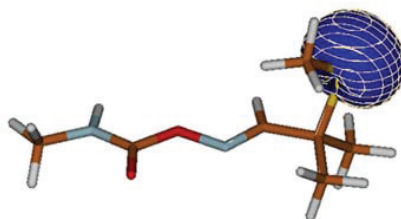


Figure S2.26. Correlations between the second-order rate constants (k_{O_3}) for the reaction of the selected sulfur-containing compounds with ozone and (a-d) the corresponding HOMO-n energies (E_{HOMO-n}) and (e-h) the corresponding NBO energy of the sulfur lone-pair electron ($E_{NBO, LP-S}$) determined with different computational methods: (a/e) HF/6-31G, (b/f) HF/6-311++G**, (c/g) B3LYP/6-31G, and (d/h) B3LYP/6-311++G**. The NBO 3.1 package was used for the computation of the $E_{NBO, LP-S}$. The orbital occupancies (full occupancy = 2) of the NBO are shown in the upper Figures e-h. The insets in Figures S2.23e-h show the correlations between the E_{HOMO-n} and the $E_{NBO, LP-S}$.

LP1- Mixed with *s* and *p* orbitals



LP2- pure *p* orbital

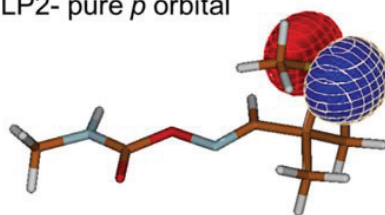


Figure S2.27. Natural bond orbitals of two lone-pair electrons (LP1 and LP2) of sulfur in Aldicarb calculated with the HF/6-31G method using the NBO 3.1 package.

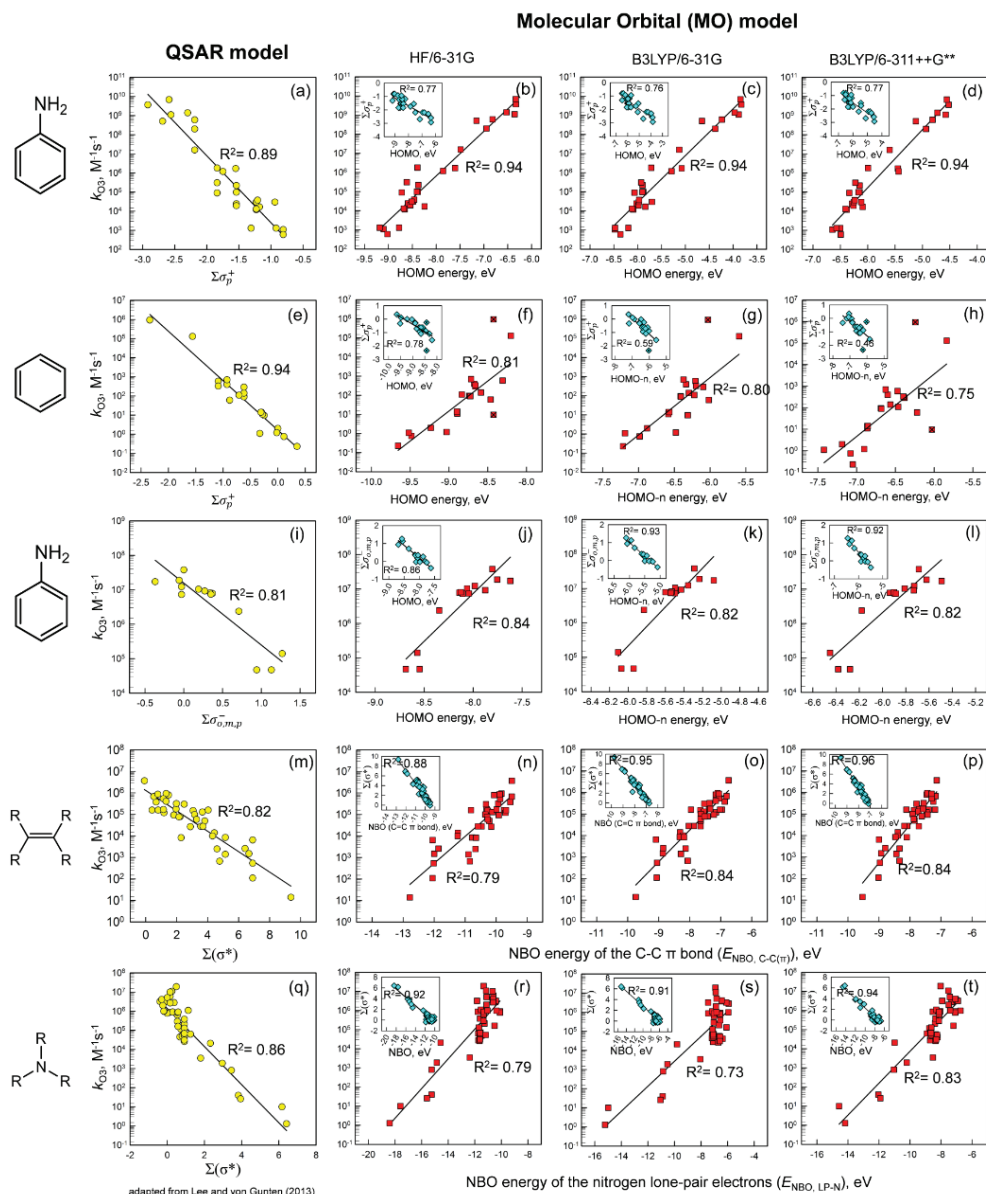
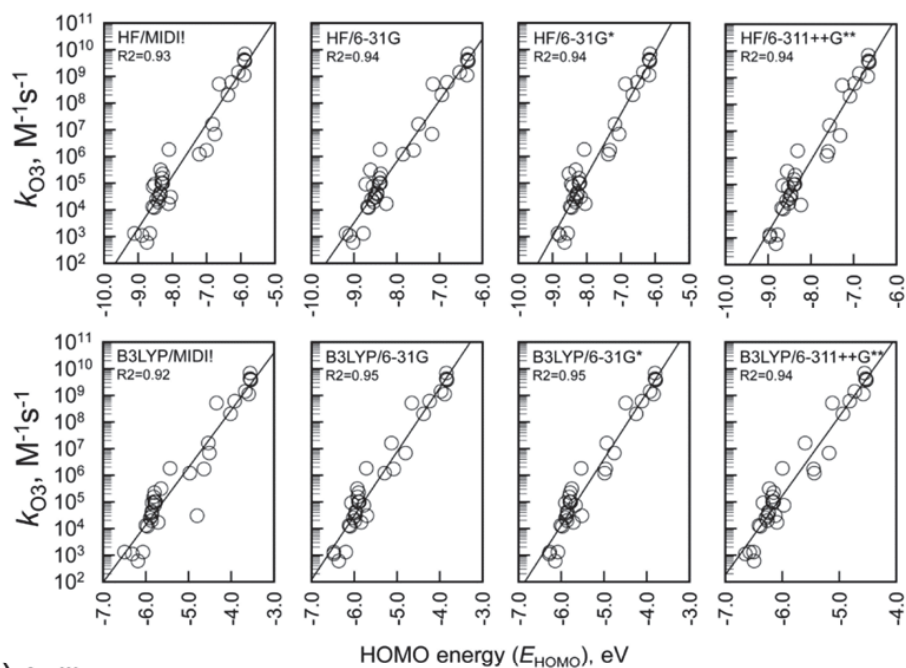
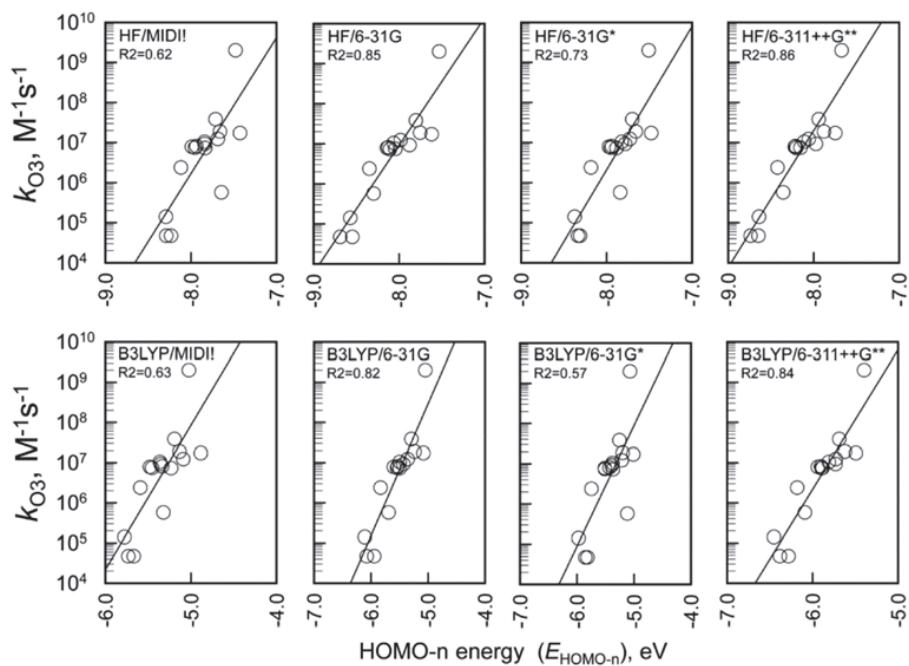


Figure S2.28. Comparison of quantum chemical molecular orbital (MO) models with QSAR models for (a-d) phenol derivatives, (e-h) benzene derivatives, (i-l) aniline derivatives, (m-p) olefins, and (q-t) amines. First column: QSAR models, second, third, and fourth columns: quantum chemical MO models with the HF/6-31G, B3LYP/6-31G, and B3LYP/6-311++G** methods, respectively. The insets in the second, third, and fourth column are the correlations between the quantum molecular descriptor and the corresponding QSAR descriptor.

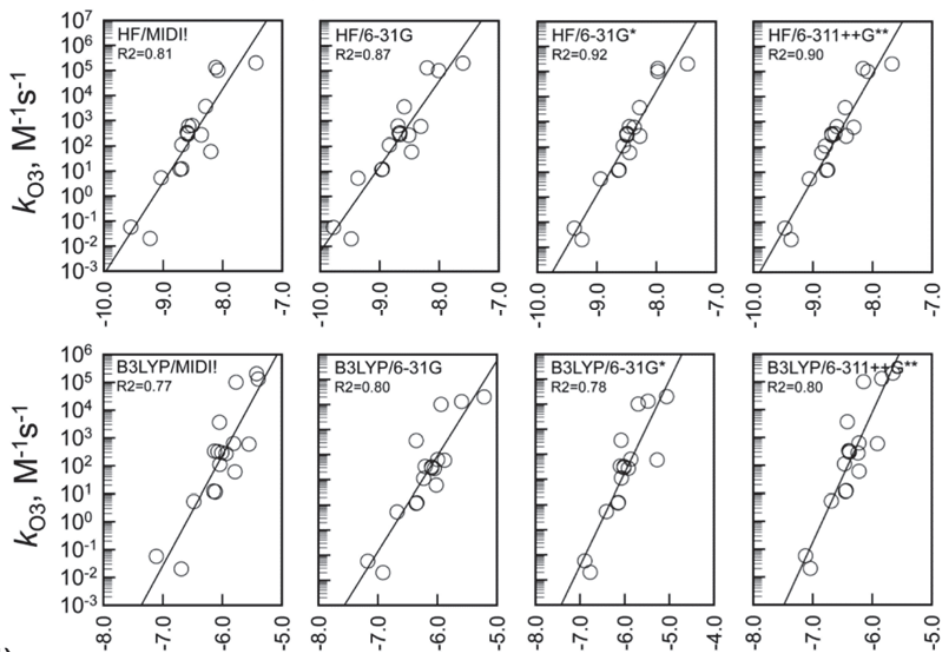
(a) Phenols



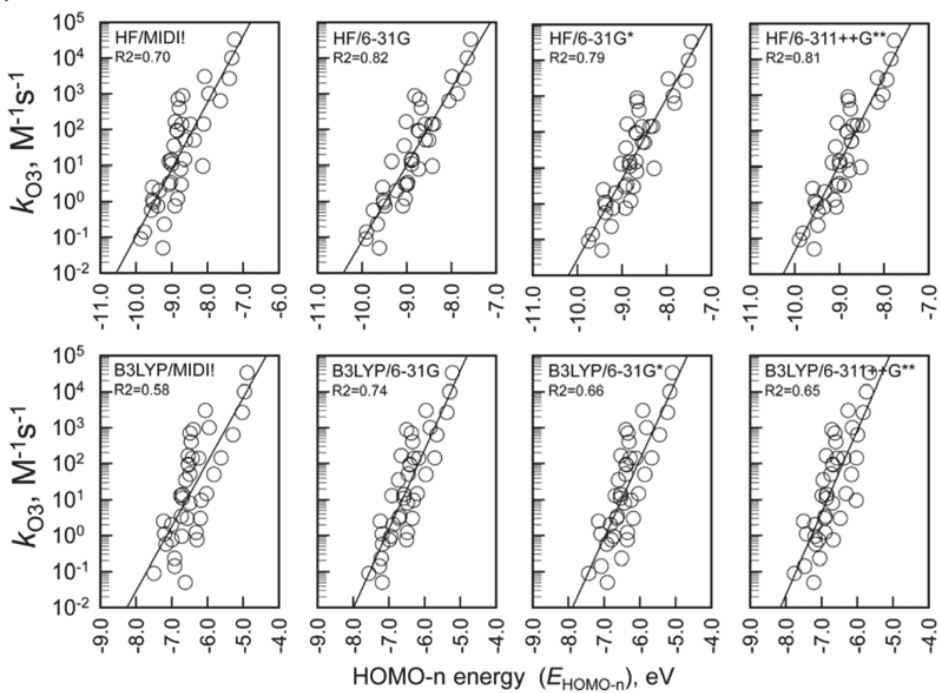
(b) Anilines



(c) Mono- and di-alkoxybenzenes



(d) Other benzene derivatives



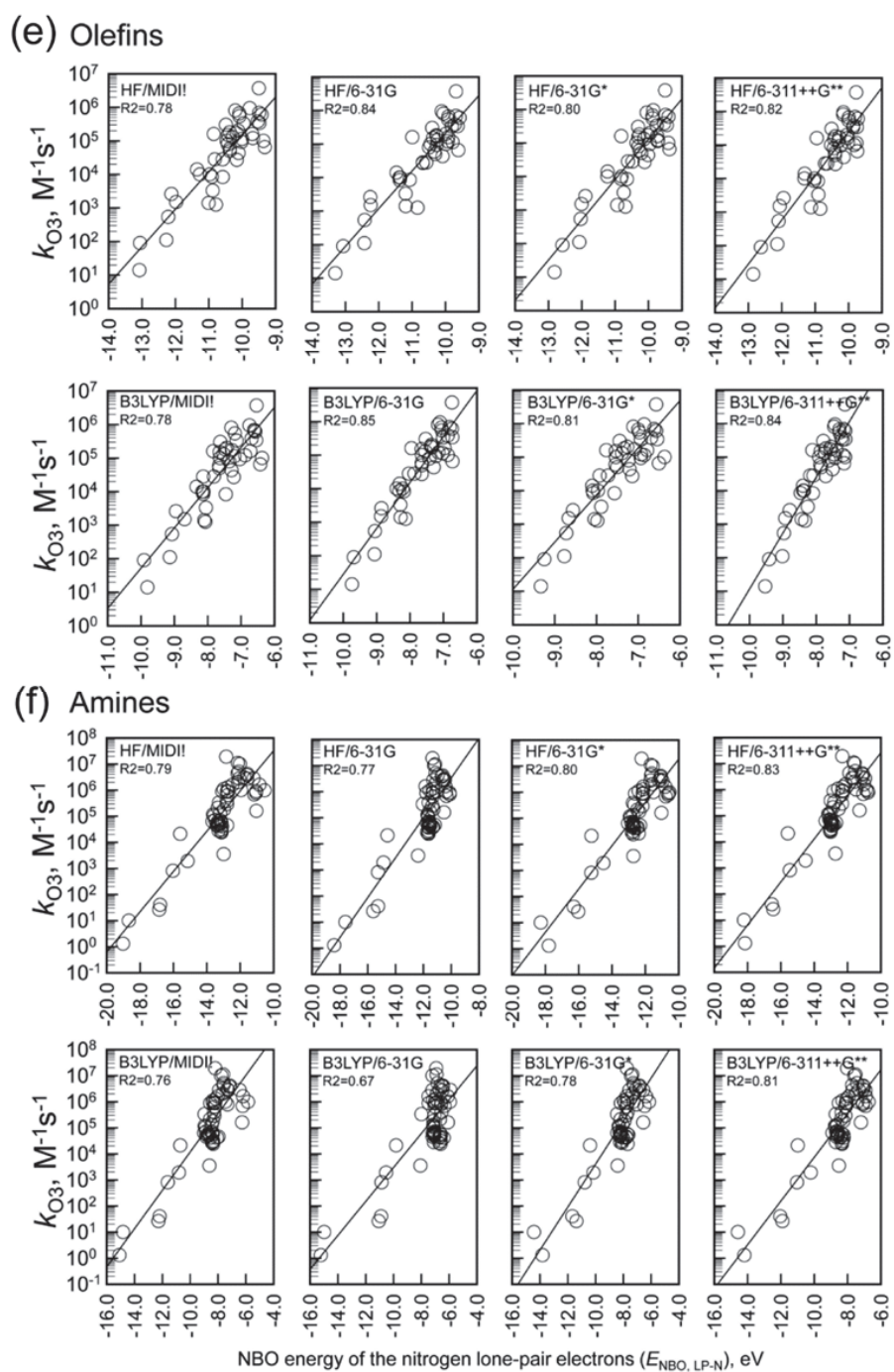


Figure S2.29. Comparison of various methods for the orbital energy computations for (a) phenols, (b) anilines, (c) mono- and dialkoxybenzenes, (d) other benzene derivatives, (e) olefins and (f) amines

Supporting information for chapter 2

Table S2.1. Second-order rate constants (k_{O_3}), dissociation constants (pK_a), and the sum of Hammett constants ($\Sigma\sigma_p^{+a}$) of the selected aromatic compounds.

Compound	pK_a	$k_{O_3}, M^{-1}s^{-1}$	$\Sigma\sigma_p^{+a}$
<i>Phenol derivatives</i>			
Bisphenol-A	9.6, 10.2 (32)	1.68×10^4 (33)	-1.18
Bisphenol-A monoanion		1.06×10^9 (33)	-2.56
2-Chlorophenol	8.3 (34)	1.10×10^3 (34)	-0.81
2-Chlorophenol anion		2.00×10^8 (34)	-2.19
4-Chlorophenol	9.2 (34)	6.00×10^2 (34)	-0.81
4-Chlorophenol anion		6.00×10^8 (34)	-2.19
2,3-Dichlorophenol*	7.7 (34)	$< 2.00 \times 10^3$ (34)	
2,4-Dichlorophenol*	7.8 (34)	$< 1.50 \times 10^3$ (34)	-2.08
2,4-Dichlorophenol anion*		$\sim 8.00 \times 10^9$ (34)	
2,3-Dimethylphenol		2.47×10^4 (35)	-1.54
2,4-Dimethylphenol		9.88×10^4 (35)	-1.54
2,6-Dimethylphenol		1.95×10^4 (35)	-1.54
3,4-Dimethylphenol		9.88×10^4 (35)	-1.54
2,6-di- <i>tert</i> -Butyl-4-methylphenol (BHT)		7.40×10^4 (36)	
17 β -Estradiol	10.4 (33)	2.21×10^5 (33)	-1.54
17 β -Estradiol anion		3.69×10^9 (33)	-2.92
Estriol	10.4 (33)	1.01×10^5 (33)	
Estriol anion		3.89×10^9 (33)	
Estrone	10.4 (33)	1.53×10^5 (33)	
Estrone anion		4.24×10^9 (33)	
2-Hydroxyphenol (catechol)	9.25, 13.0(37)	3.10×10^5 (35)	-1.84
3-Hydroxyphenol (resorcinol)	9.20, 10.9 (38)	9.10×10^4 (35)	-1.84
4-Hydroxyphenol (hydroquinone)		1.80×10^6 (35)	-1.84
2-Methylphenol (<i>o</i> -cresol)	10.2 (34)	1.20×10^4 (34)	-1.23
3-Methylphenol (<i>m</i> -cresol)	10.0 (34)	1.30×10^4 (34)	-1.23
4-Methylphenol (<i>p</i> -cresol)	10.2 (34)	3.00×10^4 (34)	-1.23
4-Nitrophenol*	7.2 (34)	$< 5.00 \times 10^1$ (34)	
4-Nitrophenol anion		1.60×10^7 (34)	-2.19
4- <i>n</i> -Nonylphenol	10.4 (39)	3.80×10^4 (33)	-1.21
4- <i>n</i> -Nonylphenol anion		6.83×10^9 (33)	-2.59
4- <i>n</i> -Octylphenol		4.30×10^4 (40)	
Paracetamol*	9.4 (41)	1.40×10^3 (42)	
Pentabromophenol anion		1.70×10^6 (43)	-1.55
Pentachlorophenol anion		1.20×10^6 (43)	-1.75
Phenol	9.9 (34)	1.30×10^3 (34)	-0.92
Phenol anion		1.40×10^9 (34)	-2.30
Salicylic acid*	13.4 (34)	$< 5.00 \times 10^2$ (34)	
Salicylic acid anion		3.00×10^4 (34)	-0.94
Triclosan	8.1 (44)	1.30×10^3 (45)	-1.31
Triclosan anion		5.10×10^8 (45)	-2.69
2,4,5-Trichlorophenol*	6.9 (34)	$< 3.00 \times 10^3$ (34)	
2,4,5-Trichlorophenol anion*		$> 1.00 \times 10^9$ (34)	
2,4,6-Trichlorophenol*	6.1 (34)	$< 1.00 \times 10^4$ (34)	
2,4,6-Trichlorophenol anion*		$> 1.00 \times 10^8$ (34)	
2,4,6-Triiodophenol anion	6.6 (46)	6.80×10^6 (43)	

^aThe sum of Hammett constants for the selected aromatic compounds reported in reference (47) *Excluded from the regression

Supporting information for chapter 2

Table S2.1. Second-order rate constants (k_{O3}), dissociation constants (pK_a), and the sum of Hammett constants ($\Sigma\sigma_p^+$) of the selected aromatic compounds (continued).

Compound	pK_a	$k_{O3}, M^{-1}s^{-1}$	$\Sigma\sigma_p^+{}^a$
<i>Benzene derivatives</i>			
Benzaldehyde		2.50 (48)	0.73
Benzene		2.00 (48)	0.00
Benzene sulfonate ion		0.23 (48)	0.35
Benzoate ion		1.20 (48)	-0.02
Chlorobenzene		0.75 (48)	0.11
Diazepam		0.75 (49)	
Diethyl- <i>o</i> -phthalate		0.14 (50)	
1,3-Dichlorobenzene		0.57 (50)	
1,4-Dichlorobenzene*		< 3.00 (48)	
Ethylbenzene		14 (48)	-0.30
Galaxolide (HHCB)		1.40×10^2 (51)	
2,2',4,4',5,5'-Hexachlorobiphenyl*		< 0.90 (50)	
Ibuprofen	4.2 (49)	9.6 (49)	-0.25
Isopropylbenzene		11 (48)	-0.28
Methotrexate dianion*	4.8, 5.2 (52)	$> 3.6 \times 10^3$ (53)	
Methylbenzoate		1.1 (48)	-0.33
2-Methylnaphthalene		1.00×10^3 (54)	
Naphthalene		3.00×10^3 (48)	
Nitrobenzene		0.09 (48)	0.79
Quinoline		51 (55)	
Toluene		14 (48)	-0.31
Tonalide		8 (51)	
1,2,3-Trichlorobenzene*		≤ 0.06 (50)	
1,2,3-Trimethylbenzene		4.00×10^2 (48)	-0.93
1,3,5-Trimethylbenzene		7.00×10^2 (48)	-0.93
<i>m</i> -Xylene		94 (48)	-0.62
<i>o</i> -Xylene		90 (48)	-0.62
<i>p</i> -Xylene		1.40×10^2 (48)	-0.62
<i>Anilides</i>			
Alachlor		3.4 (56)	
Amidotrizoic acid		0.05 (57)	
Chlorotoluron		50 (56)	
Diuron		14.7 (56)	
Iomeprol*		< 0.8 (49)	
Iopamidol*		< 0.8 (49)	
Iopromide*		< 0.8 (49)	
Isoproturon		1.41×10^2 (56)	
Linuron		3.0 (56)	
Metolachlor		3.0 (56)	
Propachlor		0.94 (56)	

^aThe sum of Hammett constants for the selected aromatic compounds reported in reference (47) *Excluded from the regression

Supporting information for chapter 2

Table S2.1. Second-order rate constants (k_{O_3}), dissociation constants (pK_a), and the sum Hammett constants ($\Sigma\sigma_p^{+a}$) of the selected aromatic compounds.

Compound	pK_a	$k_{O_3}, M^{-1}s^{-1}$	$\Sigma\sigma_p^{+a}$
<i>Benzotriazoles</i>			
Benzotriazole	1.6, 8.2 (58)	35 (58)	
Benzotriazole anion		2.65×10^3 (58)	
5-Chlorobenzotriazole	-0.04, 7.5 (58)	13 (58)	
5-Chlorobenzotriazole anion		6.30×10^2 (58)	
5,6-Dimethylbenzotriazole	1.97, 9.28 (58)	8.80×10^2 (58)	
5,6-Dimethylbenzotriazole anion		3.30×10^5 (58)	
5-Methylbenzotriazole	1.65, 8.5 (58)	1.64×10^2 (58)	
5-Methylbenzotriazole anion		1.00×10^4 (58)	
<i>Ozone-attacked nitrogen-containing aromatic compounds</i>			
Azobenzene*		2.20×10^2 (50)	
2-Isopropyl-3-methoxypyrazine *(IPMP)		50 (36)	
Pyridine*	5.2 (34)	3 (34)	
Pyridine protonated*		0.01 (34)	
<i>Triazines</i>			
Atrazine*	1.71 (59)	6 (26)	
Deethylatrazine*	1.65 (59)	0.18 (26)	
Deethyldeisopropylatrazine*		< 0.1 (26)	
Deisopropylatrazine*	1.58 (59)	3.1 (26)	
Simazine*		4.8 (50)	
<i>Aniline derivatives</i>			
Aniline	4.58 (60)	1.40×10^7 (61)	0
3-Chloroaniline	3.34 (60)	7.84×10^6 (61)	0.37
4-Chloroaniline	3.98 (60)	1.04×10^7 (61)	0.19
Dimethylaniline	5.21 (62)	2.00×10^9 (63)	
2-Fluoroaniline	3.20 (62)	7.28×10^6 (61)	-0.03
3-Fluoroaniline	3.59 (60)	8.12×10^6 (61)	0.35
4-Fluoroaniline	4.65 (60)	1.23×10^7 (61)	-0.03
3-Iodoaniline	3.61 (62)	7.42×10^6 (61)	0.35
4-Iodoaniline	3.78 (62)	9.24×10^6 (61)	0.27
2-Methoxyaniline	4.49 (62)	1.72×10^7 (61)	-0.37
3-Methylaniline	4.69 (60)	1.88×10^7 (61)	-0.06
4-Methylsulfone aniline	1.48 (62)	4.70×10^4 (61)	1.13
3-Nitroaniline	2.50 (60)	2.38×10^6 (61)	0.71
4-Nitroaniline	1.02 (60)	1.40×10^5 (61)	1.27
Sulfamethoxazole	1.7, 5.6 (64)	4.70×10^4 (65)	0.94
Sulfamethoxazole anion		5.70×10^5 (65)	

^aThe sum of Hammett constants for the selected aromatic compounds reported in reference (47) *Excluded from the regression

Table S2.1. Second-order rate constants (k_{O3}), dissociation constants (pK_a), and the sum of Hammett constants ($\Sigma\sigma_p^+$) of the selected aromatic compounds (continued).

Compound	pK_a	$k_{O3}, M^{-1}s^{-1}$	$\Sigma\sigma_p^+{}^a$
<i>Mono and di-methoxybenzenes</i>			
Acebutolol protonated		60 (66)	-0.88
Anisole		2.90×10^2 (48)	-0.62
Atenolol protonated	9.6 (66)	1.10×10^2 (66)	-0.71
Bezafibrate	3.6 (49)	5.90×10^2 (49)	-1.09
Carbofuran		6.20×10^2 (50)	
2,4-Dichlorophenoxyacetic acid		5.3 (67)	
1,4-Dimethoxybenzene		1.30×10^5 (68)	-1.56
Methoxychlor		2.70×10^2 (50)	
4-Methoxy-1-naphthalenesulfonic acid		3.60×10^3 (66)	
2-Methyl-4-chlorophenoxyacetic acid (MCPA)		11.7 (67)	
2-Methyl-4-chlorophenoxypropionic acid (Mecoprop)		12.2(67)	
Metoprolol protonated	9.7 (66)	3.30×10^2 (66)	-1.09
Naproxen anion	4.5(69)	2.00×10^5 (69)	
1-Phenoxy-2-propanol		3.20×10^2 (66)	
Propranolol protonated	9.5 (66)	1.00×10^5 (66)	
2,4,6-Tribromoanisole		0.02 (36)	
2,4,6-Trichloroanisole		0.057 (36)	
<i>Trimethoxybenzenes</i>			
Trimethoprim monoprotonated	3.2, 7.1 (65)	7.40×10^4 (65)	
Trimethoprim deprotonated		3.30×10^4 (65)	
1,3,5-Trimethoxybenzene		9.40×10^5 (68)	-2.34
3,4,5-Trimethoxytoluene		2.80×10^5 (65)	

^aThe sum of Hammett constants for selected aromatic compounds reported in reference (47)

Supporting information for chapter 2

Table S2.2. Second-order rate constants (k_{O_3}), dissociation constants (pK_a), and the sum of Taft constants ($\Sigma\sigma^*$) of the selected olefins.

Compound	pK_a	$k_{O_3}, M^{-1}s^{-1}$	$\Sigma\sigma^*$ ^a
<i>Grouped by substituents</i>			
<i>Methyl group and/or halogens</i>			
<i>cis</i> -1,2-Dibromoethene		1.50×10^3 (70)	6.66
1,1-Dichloroethene		1.10×10^2 (48)	6.90
<i>cis</i> -1,2-Dichloroethene		5.40×10^2 (71)	6.90
<i>trans</i> -1,2-Dichloroethene*		6.50×10^3 (71)	6.90
1,1-Dichloropropene		2.60×10^3 (71)	6.41
Ethene		1.80×10^5 (71)	1.96
Hexachlorocyclopentadiene		90 (50)	
Propene		8.00×10^5 (71)	1.47
Tetrachloroethene*		< 0.1 (48)	
Tetramethylethene*		> 10^6 (71)	
Trichloroethene		14 (71)	9.37
Vinyl chloride		1.40×10^4 (71)	4.43
Vinyl bromide		1.00×10^4 (70)	4.31
<i>Carboxylic acid</i>			
Acrylic acid	4.25 (72)	2.80×10^4 (72)	3.55
Acrylic acid anion		1.60×10^5 (72)	0.41
Cinnamic acid	4.27(73)	5.00×10^4 (74)	3.81
Cinnamic acid anion		3.80×10^5 (74)	0.67
3,4-Dihydroxycinnamic acid*	4.43, 8.69 (73)	2.00×10^6 (73)	
3,4-Dihydroxycinnamic acid anion*		1.20×10^7 (73)	
Fumaric acid	3.0, 4.4 (72)	8.50×10^3 (72)	5.14
Fumaric acid dianion*		$\sim 6.50 \times 10^4$ (72)	
2-Hexenoic acid	6.69(75)	3.40×10^5 (75)	
Maleic acid	1.1, 6.1(72)	1.40×10^3 (72)	5.14
Maleic acid anion*		4.20×10^3 (72)	
Maleic acid dianion*		$\sim 7.00 \times 10^3$ (72)	
Methacrylic acid	4.66 (72)	1.50×10^5 (72)	3.06
Methacrylic acid anion		3.70×10^6 (72)	-0.08
4-Methoxycinnamic acid		1.30×10^5 (74)	3.66
4-Methoxycinnamic acid anion		6.80×10^5 (74)	0.52
3-Methoxy-4-hydroxycinnamic acid*	4.69, 9.3 (73)	1.10×10^6 (73)	
3-Methoxy-4-hydroxycinnamic acid anion*		1.40×10^7 (73)	
<i>trans,trans</i> -Muconic acid dianion	2.70, 4.66 (76)	1.30×10^5 (77)	
4-Nitrocinnamic acid anion		1.20×10^5 (74)	1.18
Sorbic acid	4.76 (78)	3.20×10^5 (78)	1.98
Sorbic acid anion		9.60×10^5 (78)	

^aThe sum of Taft constants for the selected olefins reported in reference (47) *Excluded from the regression

Supporting information for chapter 2

Table S2.2. Second-order rate constants (k_{O_3}), dissociation constants (pK_a), and the sum of Taft constants ($\Sigma\sigma^*$) of the selected olefins (continued).

Compound	pK_a	$k_{O_3}, M^{-1}s^{-1}$	$\Sigma\sigma^*$ ^a
<i>Sulfonic acid or phosphonic acid</i>			
Diethyl vinylphosphonate		3.30×10^3 (70)	
Phenyl vinylsulfonate*		$\sim 2.00 \times 10^2$ (70)	
Vinyl phosphonic acid	2.2, 7.8 (70)	1.00×10^4 (70)	
Vinyl phosphonic acid monoanion		2.70×10^4 (70)	2.88
Vinyl phosphonic acid dianion		1.00×10^5 (70)	
Vinylsulfonate ion		8.30×10^3 (70)	2.28
<i>Electron withdrawing group including nitrogen or oxygen</i>			
2-Acetamidoacrylic acid anion		9.40×10^5 (78)	1.32
Acrylamide		1.00×10^5 (50)	3.15
Acrylonitrile*		6.70×10^2 (70)	4.77
1,4-Bezoquinone		2.50×10^3 (43)	4.60
But-1-en-2-ol		7.90×10^4 (71)	2.22
Vinyl acetate		1.60×10^5 (70)	4.03
Vinylene carbonate*		2.60×10^4 (70)	
<i>Miscellaneous micropollutants</i>			
Anatoxin	9.36 (78)	2.80×10^4 (78)	3.70
Carbamazepine		3.00×10^5 (49)	2.48
Cephalexin anion	2.5, 7.1(79)	9.30×10^4 (65)	1.81
Chlordane*		< 0.04 (50)	
Endrin*		< 0.02 (50)	
cis-3-Hexen-1-ol		5.40×10^5 (36)	1.13
β -Ionone		1.60×10^5 (36)	0.80
Microcystin-LR		4.10×10^5 (78)	1.31
trans-cis-2,6-Nonadienal		8.70×10^5 (36)	0.78
Oseltamivir acid	3.6, 8.9 (80)	1.70×10^5 (81)	1.15
1-Penten-3-one		5.90×10^4 (36)	3.28
Tylosin protonated	7.73(81)	7.70×10^4 (65)	2.15

^aThe sum of Taft constants for the selected olefins reported in reference (47) *Excluded from the regression

Supporting information for chapter 2

Table S2.3. Second-order rate constants (k_{O_3}) and dissociation constants (pK_a) of the selected nucleic acid constituents and miscellaneous olefins.

Compound	pK_a	$k_{O_3}, M^{-1}s^{-1}$
<i>Adenine and its derivatives</i>		
Adenine (7H- and 9H-) ^a	4.15, 9.8 (83)	12 (84)
Adenine anion		1.30×10^5 (84)
1H-9H-Adenine protonated (82, 83)		5 (84)
Adenosine	3.5 (83)	16 (84)
1H-Adenosine protonated (84)		5 (84)
5'-Adenylic acid	4.4 (83)	2.00×10^2 (88)
2'-Deoxyadenosine	3.8 (83)	14 (84)
1H-2'-Deoxyadenosine protonated		5 (84)
<i>Cytosine and its derivatives</i>		
Cytidine	4.15 (83)	3.50×10^3 (84)
Cytidine protonated		40 (84)
3H-Cytosine (86)	4.6, 12.2 (83)	1.40×10^3 (84)
Cytosine anion		1.50×10^6 (84)
Cytosine protonated		18 (84)
2'-Deoxycytidine	4.3 (83)	3.50×10^3 (84)
2'-Deoxycytidine protonated		44 (84)
5'-Deoxycytidylic acid	4.6 (83)	1.40×10^3 (88)
<i>Guanosine and its derivatives</i>		
1H-Aciclovir (87)	3.4, 9.6 (91)	1.80×10^4 (92)
Aciclovir anion		3.40×10^6 (92)
1H-7H-Aciclovir protonated		2.50×10^2 (92)
1H-2'-Deoxyguanosine	2.5, 9.2 (83)	1.90×10^4 (84)
1H-5'-Deoxyguanylic acid	2.9, 9.7 (83)	5.00×10^4 (88)
1H-Guanosine (90)	2.5, 9.2 (83)	1.60×10^4 (84)
Guanosine anion		4.00×10^6 (84)
1H-7H-Guanosine protonated (90)		$< 3.00 \times 10^2$ (84)
<i>Thymine and its derivatives</i>		
Thymidine	9.8 (83)	3.00×10^4 (84)
Thymidine anion		1.20×10^6 (84)
5'-Thymidylic acid	10.0 (83)	1.60×10^4 (88)
Thymine	9.9, 12.2 (83)	4.20×10^4 (84)
Thymine anion (1H- and 3H-) ^b		$\sim 3.00 \times 10^6$ (84)
<i>Uracil and its derivatives</i>		
5-Chlorouracil	8.0 (83)	4.30×10^3 (84)
5-Chlorouracil anion (1H- and 3H-) ^c		1.30×10^6 (84)

^a7H-Adenine and 9H-adenine are reported to coexist in aqueous solution with the ratio of 0.22 : 0.78²¹. ^b1H- and 3H-thymine (5-methyluracil) anions are assumed to coexist in aqueous solution with the ratio of 0.5:0.5 based on the structural similarity with uracil anion (see the annotation e for uracil below). ^c1H- and 3H-5-chlorouracil anions are assumed to coexist in aqueous solution with the ratio of 0.5:0.5 for the same reason for thymine anion.

Table S2.3. Second-order rate constants (k_{03}) and dissociation constants (pK_a) of the selected nucleic acid constituents and miscellaneous olefins (continued).

Compound	pK_a	$k_{03}, M^{-1}s^{-1}$
<i>Uracil and its derivatives (continued)</i>		
1H-Cylindrospermopsin anion*	8.8 (78)	$\sim 2.50 \times 10^6$ (78)
1,3-Dimethyluracil		2.80×10^3 (84)
6-Methyluracil	9.8 (83)	1.40×10^2 (84)
6-Methyluracil anion (1H- and 3H-) ^d		6.00×10^5 (84)
Uracil	9.5, >13 (83)	6.50×10^2 (84)
Uracil anion (1H- and 3H) ^e		9.20×10^5 (84)
Orotic acid	2.1, 9.45 (83)	5.90×10^3 (84)
<i>Miscellaneous olefins</i>		
<i>Cyclohexenones</i>		
β -Cyclocitral		3.90×10^3 (36)
Levonorgestrel		1.43×10^3 (94)
Medroxyprogesterone		5.58×10^2 (94)
Norethindrone		2.22×10^3 (94)
Progesterone		4.80×10^2 (95)
<i>Diaminopyrimidines</i>		
2,4-Diamino-5-methylpyrimidine	3.2, 7.1 (65)	1.30×10^6 (65)
2,4-Diamino-5-methylpyrimidine monoprotonated		2.90×10^3 (65)
2,4-Diamino-5-methylpyrimidine deprotonated*		5.00×10^2 (65)
Trimethoprim	3.2, 7.1 (96)	5.20×10^5 (65)
<i>Imidazoles</i>		
Imidazole	6.95 (75)	4.00×10^5 (34)
Imidazole protonated		22 (34)
4-Methylimidazole	7.52 (75)	3.10×10^6 (75)
<i>Indoles</i>		
3-(Dimethylaminomethyl) indole	10.0 (63)	2.40×10^6 (63)
Tryptophan	9.38 (75)	7.00×10^6 (75)
<i>Isoxazoles</i>		
N-(4)-Acetylsulfamethoxazole	5.5 (97)	20 (65)
N-(4)-Acetylsulfamethoxazole anion		2.60×10^2 (65)
3,5-Dimethylisoxazole		54 (65)
<i>Quinolones</i>		
Flumequine	6.5 (98)	1.2 (65)
Flumequine anion		1.80×10^3 (65)
Indigotrisulfonic acid		9.40×10^7 (68)

^d1H- and 3H-6-methyluracil anions are assumed to coexist in aqueous solution with the ratio of 0.5:0.5 for the same reason for thymine anion. ^e1H- and 3H-uracil anions are reported to coexist in aqueous solution with the ratio of 0.5:0.5²⁵.

*Excluded from the regression

Supporting information for chapter 2

Table S2.4. Second-order rate constants (k_{O_3}), dissociation constants (pK_a), and the sum of Hammett constants ($\Sigma\sigma^*$) of the selected amines.

Compound	pK_a	$k_{O_3}, M^{-1}s^{-1}$	$\Sigma\sigma^*$ ^a
Ammonia*	9.3(34)	20 (34)	
<i>Primary amines</i>			
N- α -Acetyllysine	10.53 (75)	1.03×10^5 (75)	
N- ϵ -Acetyllysine	9.46 (75)	2.40×10^4 (75)	
Alanine	9.87 (75)	2.80×10^4 (99)	0.92
β -Alanine	10.3 (34)	6.20×10^4 (34)	
Arginine	8.99 (75)	5.70×10^4 (75)	0.92
Asparagine	8.85 (75)	4.20×10^4 (75)	
Aspartate	9.82 (75)	4.10×10^4 (75)	0.92
Benzylamine	9.33 (75)	6.30×10^4 (75)	1.25
<i>n</i> -Butylamine	10.77 (75)	1.17×10^5 (75)	0.73
<i>sec</i> -Butylamine	10.63 (75)	5.20×10^4 (75)	0.79
<i>tert</i> -Butylamine	10.83 (75)	4.50×10^4 (75)	0.68
Cyclohexanemethylamine	10.3 (100)	7.10×10^4 (65)	0.91
Cyclohexylamine	10.6 (101)	4.90×10^4 (65)	0.80
Glutamine	9.28 (75)	2.60×10^4 (75)	0.92
Glycine	9.78 (75)	2.10×10^5 (75)	0.92
Hydroxylamine	6.0 (102)	2.10×10^4 (102)	2.32
Isoleucine	9.76 (75)	5.60×10^4 (75)	0.92
Leucine	9.74 (75)	5.30×10^4 (75)	0.92
Lysine	9.18 (75)	3.10×10^4 (75)	0.92
Methylamine	10.7 (34)	$< 1.40 \times 10^5$ (34)	
Monobromamine		40 (102)	3.82
Monochloramine		26 (102)	3.94
Phenylalanine	9.24 (75)	3.80×10^5 (75)	0.92
Serine	9.21 (75)	1.29×10^5 (75)	0.92
Threonine	9.10 (75)	4.50×10^4 (75)	
Valine	9.72 (75)	6.80×10^4 (75)	0.92
<i>Secondary amines</i>			
Acebutolol	9.2 (66)	2.90×10^5 (66)	0.55
Anatoxin-a	9.36 (103)	8.70×10^5 (78)	0.11
Atenolol	9.6 (66)	6.30×10^5 (66)	0.55
Dibromamine		10 (102)	6.17
Dichloramine		1.3 (102)	6.41
Diclofenac anion		6.80×10^5 (104)	
Diethylamine	10.5 (99)	9.10×10^5 (99)	0.29
Dimethylamine	11.0 (34)	1.90×10^7 (34)	0.49
Ethyl N-piperazine carboxylate	8.3 (105)	1.10×10^6 (65)	1.13
Iminodiacetic acid		2.80×10^6 (99)	0.37
Methylchloramine		8.10×10^2 (106)	3.45
Metoprolol	9.7 (66)	8.60×10^5 (66)	0.55
Morpholine	8.36 (107)	1.80×10^5 (107)	
Proline	10.6 (75)	4.30×10^6 (75)	
Propranolol*	9.5 (66)	$\sim 1.00 \times 10^5$ (66)	

^aThe sum of Taft constants for the selected amines reported in reference (47) *Excluded from the regression

Table S2.4. Second-order rate constants (k_{O3}), dissociation constants (pK_a), and the sum of Taft constants ($\Sigma\sigma^*$) of the selected amines (continued).

Compound	pK_a	$k_{O3}, M^{-1}s^{-1}$	$\Sigma\sigma^*$ ^a
<i>Secondary amines(continued)</i>			
Spectinomycin	6.80, 8.80 (108)	1.27×10^6 (109)	
Spectinomycin monoprotonated		3.30×10^5 (109)	
<i>Tertiary amines</i>			
Azithromycin	8.7, 9.5 (82)	6.00×10^6 (65)	0.15
Ciprofloxacin anion	6.2, 8.8 (65)	9.00×10^5 (65)	0.59
1,4-Diazabicyclo[2.2.2]octane (DABCO)	2.97, 8.82 (110)	3.20×10^6 (99)	-0.39
1,4-Diazabicyclo[2.2.2]octane monoprotonated		3.50×10^3 (99)	1.80
4-Dimethylaminoantipyrine*	5.1 (63)	1.70×10^8 (63)	
3-Dimethylaminomethyl indole*	10.0 (63)	3.60×10^9 (63)	
Dimethylchloramine		1.90×10^3 (106)	2.96
N,N-Dimethylcyclohexylamine	10.7 (101)	3.70×10^6 (65)	-0.18
Dimethylethanolamine	9.2 (111)	1.10×10^7 (63)	0.25
Enrofloxacin	6.1, 7.7 (112)	7.80×10^5 (65)	-0.00
Ethylenediaminetetraacetic acid (EDTA)	0.26, 0.96, 2.0, 2.7, 6.2, 10.2 (113)	3.20×10^6 (99)	-0.07
Ethylenediaminetetraacetic acid monoprotonated (HEDTA)		1.60×10^5 (99)	0.58
Lincomycin	7.79 (108)	2.76×10^6 (109)	
Methyldichloramine*		< 0.01 (106)	
1-Methylpyrrolidine	10.2 (100)	2.00×10^6 (65)	-0.24
Nitrioltriacetic acid (NTA)		9.80×10^5 (99)	-0.18
Roxithromycin	9.2 (108)	1.00×10^7 (65)	0.15
Tramadol	9.40 (114)	1.00×10^6 (115)	-0.19
Triethylamine	11.0 (99)	4.10×10^6 (99)	-0.30
Trimethylamine	11.01 (99)	5.10×10^6 (99)	0
Tylosine	7.7 (82)	2.70×10^6 (65)	0.15

^aThe sum of Taft constants for the selected olefins reported in reference (47) *Excluded from the regression

Supporting information for chapter 2

Table S2.5. Second-order rate constants (k_{O_3}) and dissociation constants (pK_a) of the selected sulfur-containing compounds.

Compound	pK_a	$k_{O_3}, M^{-1}s^{-1}$
<i>Thiols</i>		
Cysteine	8.14 (75)	4.20×10^4 (75)
Cysteine anion		2.40×10^6 (75)
Glutathione	8.75 (75)	4.00×10^6 (75)
Glutathione protonated		2.00×10^4 (75)
<i>Sulfides</i>		
Aldicarb		4.30×10^5 (50)
EPTC (S-ethyl-N,N-dipropylthiocarbamate)		5.00×10^2 (116)
Methionine		1.80×10^6 (75)
Molinate		5.00×10^2 (116)
Penicilline G		4.80×10^3 (65)
Vydate		6.20×10^2 (50)
<i>Disulfides</i>		
<i>trans</i> -1,2-Dithiane-4,5-diol		2.10×10^5 (117)
bis(2-Hydroxyethyl) disulfide		1.70×10^5 (117)
<i>Sulfoxide</i>		
Dimethyl sulfoxide		8.0 (75)
<i>Sulfinic acid</i>		
Methanesulfinate		2.00×10^6 (118)

Table S2.6. Summary of the developed models for the nucleic acid constituents and the miscellaneous olefins.

Group	Computation level	Model equation	N	R ²	MUE ^a	RMSE ^b
Nucleic acid constituents						
<i>Adenine and its derivatives</i>						
	HF/6-31G	$\log(k_{03}) = 1.76 \times (E_{\text{HOMO}}) + 17.54$	8	0.73	0.66	1.26
	HF/6-311++G**	$\log(k_{03}) = 2.05 \times (E_{\text{HOMO}}) + 20.13$	8	0.74	0.66	1.24
	B3LYP/6-31G	$\log(k_{03}) = 1.55 \times (E_{\text{HOMO}_n}) + 11.83$	8	0.68	0.69	1.32
	B3LYP/6-311++G**	$\log(k_{03}) = 1.89 \times (E_{\text{HOMO}}) + 14.40$	8	0.67	0.71	1.33
	HF/6-31G	$\log(k_{03}) = 1.58 \times (E_{\text{NBO, C-C(ri)}}) + 17.13$	8	0.86	0.50	1.00
	HF/6-311++G**	$\log(k_{03}) = 1.84 \times (E_{\text{NBO, C-C(ri)}}) + 18.89$	8	0.84	0.53	1.05
	B3LYP/6-31G	$\log(k_{03}) = 1.84 \times (E_{\text{NBO, C-C(ri)}}) + 16.14$	8	0.75	0.63	1.25
	B3LYP/6-311++G**	$\log(k_{03}) = 2.02 \times (E_{\text{NBO, C-C(ri)}}) + 17.23$	8	0.80	0.59	1.15
<i>Cytosine and its derivatives</i>						
	HF/6-31G	$\log(k_{03}) = 1.70 \times (E_{\text{HOMO}}) + 19.27$	8	0.99	0.14	0.23
	HF/6-311++G**	$\log(k_{03}) = 1.94 \times (E_{\text{HOMO}}) + 21.70$	8	0.99	0.12	0.19
	B3LYP/6-31G	$\log(k_{03}) = 1.68 \times (E_{\text{HOMO}_n}) + 14.16$	8	0.98	0.16	0.23
	B3LYP/6-311++G**	$\log(k_{03}) = 2.03 \times (E_{\text{HOMO}}) + 17.22$	8	0.97	0.18	0.32
	HF/6-31G	$\log(k_{03}) = 2.13 \times (E_{\text{NBO, C-C(ri)}}) + 25.18$	8	0.95	0.32	0.73
	HF/6-311++G**	$\log(k_{03}) = 2.15 \times (E_{\text{NBO, C-C(ri)}}) + 25.25$	8	0.92	0.37	1.12
	B3LYP/6-31G	$\log(k_{03}) = 2.13 \times (E_{\text{NBO, C-C(ri)}}) + 19.92$	8	0.97	0.19	0.29
	B3LYP/6-311++G**	$\log(k_{03}) = 2.44 \times (E_{\text{NBO, C-C(ri)}}) + 22.71$	8	0.97	0.20	0.29
<i>Guanine and its derivatives</i>						
	HF/6-31G	$\log(k_{03}) = 2.12 \times (E_{\text{HOMO}}) + 22.41$	8	0.98	0.17	0.28
	HF/6-311++G**	$\log(k_{03}) = 2.32 \times (E_{\text{HOMO}}) + 24.22$	8	0.98	0.20	0.31
	B3LYP/6-31G	$\log(k_{03}) = 2.05 \times (E_{\text{HOMO}}) + 16.52$	8	0.99	0.14	0.23
	B3LYP/6-311++G**	$\log(k_{03}) = 2.48 \times (E_{\text{HOMO}}) + 19.83$	8	0.98	0.18	0.26
	HF/6-31G	$\log(k_{03}) = 1.85 \times (E_{\text{NBO, C-C(ri)}}) + 23.06$	8	0.93	0.34	0.52
	HF/6-311++G**	$\log(k_{03}) = 1.91 \times (E_{\text{NBO, C-C(ri)}}) + 23.19$	8	0.92	0.39	0.56
	B3LYP/6-31G	$\log(k_{03}) = 1.93 \times (E_{\text{NBO, C-C(ri)}}) + 19.35$	8	0.95	0.28	0.44
	B3LYP/6-311++G**	$\log(k_{03}) = 2.17 \times (E_{\text{NBO, C-C(ri)}}) + 21.17$	8	0.94	0.33	0.49
<i>Thymine and its derivatives</i>						
	HF/6-31G	$\log(k_{03}) = 1.63 \times (E_{\text{HOMO}}) + 19.74$	5	0.94	0.17	0.34
	HF/6-311++G**	$\log(k_{03}) = 1.73 \times (E_{\text{HOMO}}) + 20.95$	5	0.97	0.14	0.29
	B3LYP/6-31G	$\log(k_{03}) = 1.63 \times (E_{\text{HOMO}}) + 15.04$	5	0.94	0.18	0.34
	B3LYP/6-311++G**	$\log(k_{03}) = 1.95 \times (E_{\text{HOMO}}) + 17.84$	5	0.98	0.11	0.21
	HF/6-31G	$\log(k_{03}) = 1.98 \times (E_{\text{NBO, C-C(ri)}}) + 25.35$	5	0.93	0.20	0.38
	HF/6-311++G**	$\log(k_{03}) = 1.93 \times (E_{\text{NBO, C-C(ri)}}) + 24.95$	5	0.95	0.17	0.47
	B3LYP/6-31G	$\log(k_{03}) = 2.06 \times (E_{\text{NBO, C-C(ri)}}) + 20.62$	5	0.93	0.20	0.39
	B3LYP/6-311++G**	$\log(k_{03}) = 2.38 \times (E_{\text{NBO, C-C(ri)}}) + 23.61$	5	0.99	0.08	0.15
<i>Uracil and its derivatives</i>						
	HF/6-31G	$\log(k_{03}) = 2.14 \times (E_{\text{HOMO}}) + 24.03$	8	0.84	0.45	0.76
	HF/6-311++G**	$\log(k_{03}) = 2.46 \times (E_{\text{HOMO}}) + 27.11$	8	0.91	0.34	0.55
	B3LYP/6-31G	$\log(k_{03}) = 2.16 \times (E_{\text{HOMO}_n}) + 17.68$	8	0.86	0.39	0.72
	B3LYP/6-311++G**	$\log(k_{03}) = 3.00 \times (E_{\text{HOMO}}) + 24.19$	8	0.92	0.32	0.54
	HF/6-31G	$\log(k_{03}) = 2.51 \times (E_{\text{NBO, C-C(ri)}}) + 29.80$	6	0.84	0.49	0.81
	HF/6-311++G**	$\log(k_{03}) = 2.61 \times (E_{\text{NBO, C-C(ri)}}) + 30.71$	6	0.83	0.49	0.83
	B3LYP/6-31G	$\log(k_{03}) = 2.71 \times (E_{\text{NBO, C-C(ri)}}) + 24.48$	6	0.87	0.42	0.73
	B3LYP/6-311++G**	$\log(k_{03}) = 3.55 \times (E_{\text{NBO, C-C(ri)}}) + 31.66$	6	0.84	0.47	0.83
Miscellaneous olefins^c						
	HF/6-311++G**	$\log(k_{03}) = 2.59 \times (E_{\text{HOMO}_n}) + 28.36$	32	0.89	0.51	0.67
	HF/6-31G	$\log(k_{03}) = 2.45 \times (E_{\text{HOMO}_n}) + 20.06$	32	0.90	0.47	0.65
	B3LYP/6-31G	$\log(k_{03}) = 2.45 \times (E_{\text{HOMO}_n}) + 20.06$	32	0.80	0.70	0.92
	B3LYP/6-311++G**	$\log(k_{03}) = 2.99 \times (E_{\text{HOMO}_n}) + 24.54$	32	0.86	0.57	0.75

^aMean unsigned error, ^bRoot-mean-square error, ^cThymine, uracil, and their derivatives included

Supporting information for chapter 2

Table S2.7. Estimation of unknown Taft constants (σ^*) for substituents of aliphatic olefins by the correlation between the $E_{\text{NBO}, \text{C}-\text{C}(\pi)}$ of the NBO model and σ^* of the QSAR model.

Compound	Known substituents ^a	Unknown substituent ^b	Estimated σ^{*c}
Diethyl vinylphosphonate	-H, -H, -H	-PO(OEt) ₂	2.83±0.38
Phenyl vinylsulfonate	-H, -H, -H	-SO ₃ -C ₆ H ₅	4.94±0.46
Sorbic acid deprotonated	-C(O)O ²⁻ , -H, -H	-CH=CH-CH ₃	0.48±0.22
Vinyl phosphonic acid	-H, -H, -H	-PO(OH) ₂	3.18±0.40
Vinyl phosphonic acid dianion	-H, -H, -H	-PO(O)(O) ²⁻	-0.67±0.55

^aSubstituents for which Taft constants (σ^*) are available in literature. ^bA substituent for which a σ^* is not available in literature. ^cThe average of the estimated σ^* -values from four different levels of computations with a standard deviation.

Table S2.8. The highest occupied molecular orbital energies (E_{HOMO}), the relative total energies (ΔE), the relative Gibbs free energies (ΔG), and the populations (P) of the conformers of the selected trimethoxybenzenes.

Conformer	HF/6-311++G**			B3LYP/6-311++G**		
	E_{HOMO} , eV	$\Delta E^{a,d}$	$\Delta G^{b,d}$	E_{HOMO} , eV	$\Delta E^{a,d}$	$\Delta G^{b,d}$
<i>1,3,5-Trimethoxybenzene</i>						
I	-8.68	0.0	0.0	-6.24	0.0	0.0
II	-8.51	3.25	1.64	-6.10	2.35	2.22
<i>3,4,5-Trimethoxytoluene</i>						
I	-8.80	4.79	1.80	-6.56	15.62	11.96
II	-8.50	0.05	0.0	-6.19	10.98	10.03
III	-8.50	0.0	0.80	-6.16	0.0	0.0
IV	-8.80	4.81	2.66	-6.56	15.69	12.70
<i>Trimethoprim monoprotonated</i>						
I	-9.04	4.21	7.95	-6.77	16.45	11.47
II	-8.71	0.44	3.66	-6.37	0.0	0.0
III	-8.71	0.0	0.0	-6.37	0.43	1.35
IV	-8.75	1.47	7.81	-6.45	7.22	5.34
V	-8.71	2.79	10.77	-6.41	7.86	6.19
VI	-9.04	4.40	8.95	-6.78	16.25	8.74
<i>Trimethoprim diprotonated</i>						
I	-9.20	4.53	10.71	-6.91	18.84	17.17
II	-8.84	0.0	0.0	-6.51	0.0	5.66
III	-8.85	0.62	11.65	-6.51	0.61	0.0
IV	-8.87	3.27	7.33	-6.61	8.19	13.78
V	-8.84	0.49	2.88	-6.55	9.29	6.63
VI	-9.20	4.57	11.27	-6.91	17.38	12.91

^aZero point vibrational energy (ZPVE) corrected total energy in kJ/mol. ^bGibbs free energy in kJ/mol at 298.15 K. ^cCalculated by the Boltzmann equilibrium based on the relative Gibbs free energies. ^dThe relative energy of conformers with respect to the lowest energy conformer.

References

- (1) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. DREIDING: a generic force field for molecular simulations. *J. Phys. Chem.* **1990**, *94*, 8897–8909.
- (2) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648.
- (3) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.
- (4) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenb, 2009. Gaussian 09, Revision D.01.
- (5) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **1980**, *72*, 650.
- (6) Hehre, W. J. Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257.
- (7) Frisch, M. J.; Pople, J. A.; Binkley, J. S. Self-consistent molecular orbital methods 25. Supplementary functions for Gaussian basis sets. *J. Chem. Phys.* **1984**, *80*, 3265.
- (8) McLean, A. D.; Chandler, G. S. Contracted Gaussian basis sets for molecular calculations. I. Second row atoms, Z=11–18. *J. Chem. Phys.* **1980**, *72*, 5639.
- (9) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+G basis set for first-row elements, Li-F. *J. Comput. Chem.* **1983**, *4*, 294–301.
- (10) Wadt, W. R.; Hay, P. J. Ab initio effective core potentials for molecular calculations. Potentials for main group elements Na to Bi. *J. Chem. Phys.* **1985**, *82*, 284.
- (11) Hariharan, P. C.; Pople, J. A. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta* **1973**, *28*, 213–222.
- (12) Easton, R. E.; Giesen, D. J.; Welch, A.; Cramer, C. J.; Truhlar, D. G. The MIDI! basis set for quantum mechanical calculations of molecular geometries and partial charges. *Theor. Chim. Acta*, **1996**, *93*, 281–301.
- (13) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999–3093.
- (14) E. D. Glendening, A. E. Reed, J. E. Carpenter, F. W. NBO Version 3.1.
- (15) Reed, A. E.; Curtiss, L. A.; Weinhold, F. Intermolecular interactions from a natural bond orbital, donor-acceptor viewpoint. *Chem. Rev.* 1988, *88*, 899–926.
- (16) Glendening, E. D.; Badenhoop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Landis, C. R.; Weinhold, F. NBO 6.0. *Univ. Wisconsin, Madison, WI*, **2013**.
- (17) Allouche, A.-R. Gabedit—a graphical user interface for computational chemistry softwares. *J. Comput. Chem.* **2011**, *32*, 174–182.
- (18) Schaftenaar, G.; Noordik, J. H. Molden: a pre-and post-processing program for molecular and electronic structures*. *J. Comput. Aided. Mol. Des.* **2000**, *14*, 123–134.
- (19) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other function. *Theor. Chem. Acc.* **2007**, *120*, 215–241.
- (20) Adamo, C.; Barone, V. Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The mPW and mPW1PW models. *J. Chem. Phys.* **1998**, *108*, 664.
- (21) Sandoz, C.; Lesca, P.; Narbonne, J. F.; Carpy, A. Molecular characteristics of carbaryl, a CYP1A1 gene inducer. *Arch. Biochem. Biophys.* **2000**, *373*, 275–280.

- (22) Roussy, G.; Nonat, A. Determination of the equilibrium molecular structure of inverting molecules by microwave spectroscopy: Application to aniline. *J. Mol. Spectrosc.* **1986**, *118*, 180–188.
- (23) Bock, C. W.; George, P.; Trachtman, M. A molecular orbital study of nitrogen inversion in aniline with extensive geometry optimization. *Theor. Chim. Acta* **1986**, *69*, 235–245.
- (24) Andreozzi, R.; Insola, A.; Caprio, V.; D'Amore, M. Ozonation of pyridine in aqueous solution: Mechanistic and kinetic aspects. *Water Res.* **1991**, *25*, 655–659.
- (25) von Sonntag, C.; von Gunten, U. *Chemistry of ozone in water and wastewater treatment: From basic principles to applications*; Iwa publishing, 2012.
- (26) Acero, J. L.; Stemmler, K.; von Gunten, U. Degradation Kinetics of Atrazine and Its Degradation Products with Ozone and OH Radicals: A Predictive Tool for Drinking Water Treatment. *Environ. Sci. Technol.* **2000**, *34*, 591–597.
- (27) Dreyfus, M.; Dodin, G.; Bensaude, O.; Dubois, J. E. Tautomerism of purines. I. N(7)H ↔ N(9)H equilibrium in adenine. *J. Am. Chem. Soc.* **1975**, *97*, 2369–2376.
- (28) Nakanishi, K.; Suzuki, N.; Yamazaki, F. Ultraviolet Spectra of N-Heterocyclic Systems. I. The Anions of Uracils. *Bull. Chem. Soc. Jpn.* **1961**, *34*, 53–57.
- (29) Lee, Y.; Gerrity, D.; Lee, M.; Bogeat, A. E.; Salhi, E.; Gamage, S.; Trenholm, R. A.; Wert, E. C.; Snyder, S. A.; von Gunten, U. Prediction of micropollutant elimination during ozonation of municipal wastewater effluents: use of kinetic and water specific information. *Environ. Sci. Technol.* **2013**, *47*, 5872–5881.
- (30) Williamson, D. G.; Cvetanovic, R. J. Rates of reactions of ozone with chlorinated and conjugated olefins. *J. Am. Chem. Soc.* **1968**, *90*, 4248–4252.
- (31) Williamson, D. G.; Cvetanovic, R. J. Rates of ozone-olefin reactions in carbon tetrachloride solutions. *J. Am. Chem. Soc.* **1968**, *90*, 3668–3672.
- (32) Kosky, P. G.; Silva, J. M.; Guggenheim, E. A. The aqueous phase in the interfacial synthesis of polycarbonates. Part 1. Ionic equilibria and experimental solubilities in the BPA-sodium hydroxide-water system. *Ind. Eng. Chem. Res.* **1991**, *30*, 462–467.
- (33) Deborde, M.; Rabouan, S.; Duguet, J.-P.; Legube, B. Kinetics of aqueous ozone-induced oxidation of some endocrine disruptors. *Environ. Sci. Technol.* **2005**, *39*, 6086–6092.
- (34) Hoigne, J.; Bader, H. Rate Constants of Reactions of Ozone with Organic and Inorganic Compounds in Water --II. Dissociating organic compounds. *Water Res.* **1983**, *17*, 185–194.
- (35) Gurol, M.; Nekouinaini, S. Kinetic behavior of ozone in aqueous solutions of substituted phenols. *Ind. Eng. Chem.* **1984**, *23*, 54–60.
- (36) Peter, A.; von Gunten, U. Oxidation kinetics of selected taste and odor compounds during ozonation of drinking water. *Environ. Sci. Technol.* **2007**, *41*, 626–631.
- (37) Schweigert, N.; Zehnder, A. J. B.; Eggen, R. I. L. Chemical properties of catechols and their molecular modes of toxic action in cells, from microorganisms to mammals. *Environ. Microbiol.* **2001**, *3*, 81–91.
- (38) Blanco, S. E.; Almandoz, M. C.; Ferretti, F. H. Determination of the overlapping pKa values of resorcinol using UV-visible spectroscopy and DFT methods. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **2005**, *61*, 93–102.
- (39) Maguire, R. J. Review of the persistence of nonylphenol and nonylphenol ethoxylates in aquatic environments. *Water Qual. Res. J. Canada* **1999**, *34*, 37–78.
- (40) Ning, B.; Graham, N.; Zhang, Y. Degradation of octylphenol and nonylphenol by ozone—Part I: Direct reaction. *Chemosphere* **2007**, *68*, 1163–1172.
- (41) Jones, O.; Vouvolis, N.; Lester, J. Aquatic environmental assessment of the top 25 English prescription pharmaceuticals. *Water Res.* **2002**, *36*, 5013–5022.
- (42) Andreozzi, R.; Caprio, V.; Marotta, R.; Vogna, D. Paracetamol oxidation from aqueous solutions by means of ozonation and H₂O₂/UV system. *Water Res.* **2003**, *37*, 993–1004.
- (43) Mvula, E.; Sonntag, C. von. Ozonolysis of phenols in aqueous solution. *Org. Biomol. Chem.* **2003**, *1*, 1749–1756.
- (44) Pillonel, L. Diploma thesis. *ETH Zurich* **1999**.
- (45) Suarez, S.; Dodd, M.; Omil, F.; von Gunten, U. Kinetics of triclosan oxidation by aqueous ozone and consequent loss of antibacterial activity: relevance to municipal wastewater ozonation. *Water Res.* **2007**, *41*, 2481–2490.

Supporting information for chapter 2

- (46) Perrin, D. D.; Dempsey, B.; Serjeant, E. P. *pKa prediction for organic acids and bases*; Chapman and Hall London; New York, 1981; Vol. 1.
- (47) Lee, Y.; von Gunten, U. Quantitative structure-activity relationships (QSARs) for the transformation of organic micropollutants during oxidative water treatment. *Water Res.* **2012**, *46*, 6177–6195.
- (48) Hoigné, J.; Bader, H. Rate Constants of Reactions of Ozone with Organic and Inorganic Compounds in Water --I. Non-dissociating compounds. *Water Res.* **1983**, *17*, 173–183.
- (49) Huber, M.; Canonica, S. Oxidation of pharmaceuticals during ozonation and advanced oxidation processes. *Environ. Sci. Technol.* **2003**, *37*, 1016–1024.
- (50) Yao, C. D.; Haag, W. Rate constants for direct reactions of ozone with several drinking water contaminants. *Water Res.* **1991**, *25*, 761–773.
- (51) Nöthe, T.; Hartmann, D.; von Sonntag, J.; von Sonntag, C.; Fahlenkamp, H. Elimination of the musk fragrances galaxolide and tonalide from wastewater by ozonation and concomitant stripping. *Water Sci. Technol.* **2007**, *55*, 287–292.
- (52) Reid, T.; Yuen, A.; Catolico, M.; Carlson, R. W. Impact of omeprazole on the plasma clearance of methotrexate. *Cancer Chemother. Pharmacol.* **1993**, *33*, 82–84.
- (53) Garcia-Ac, A.; Broséus, R.; Vincent, S. Oxidation kinetics of cyclophosphamide and methotrexate by ozone in drinking water. *Chemosphere* **2010**, *79*, 1056–1063.
- (54) Legube, B.; Sugimitsu, H.; Guyon, S.; Dore, M. Ozonation of Naphthalene in Aqueous Solution. *Water Res.* **1986**, *20*, 209–214.
- (55) Wang, X.; Huang, X.; Zuo, C.; Hu, H. Kinetics of quinoline degradation by O₃/UV in aqueous phase. *Chemosphere* **2004**, *55*, 733–741.
- (56) Laet, J. De; Maouala-Makata, P.; Dore, M. Constantes Cinétiques de Réaction de L'Ozone Moléculaire et des Radicaux Hydroxyles Sur Quelques Phenyl-Ureés et Acétamides. *Environ. Technol.* **1996**, *17*, 707–716.
- (57) Real, F.J., Benitez, F.J., Acero, J.L., Sagasti, J.P., C. F. Kinetics of the chemical oxidation of the pharmaceuticals primidone, ketoprofen, and diatrizoate in ultrapure and natural waters. *Ind. Eng. Chem. Res.* **2009**, *48*, 3380–3388.
- (58) Lutze, H.; Naumov, S.; Liu, A.; von Gunten, U.; von Sonntag, C.; Schmidt, T. S. Ozonation of benzotriazolones: rate constants and mechanistic aspects. *in preparation*.
- (59) Vermeulen, N. M. J.; Apostolides, Z.; Potgieter, D. J. J.; Nel, P. C.; Smit, N. S. H. Separation of atrazine and some of its degradation products by high-performance liquid chromatography. *J. Chromatogr. A* **1982**, *240*, 247–253.
- (60) Gross, K. C.; Seybold, P. G.; Peralta-Inga, Z.; Murray, J. S.; Politzer, P. Comparison of Quantum Chemical Parameters and Hammett Constants in Correlating pK_a Values of Substituted Anilines. *J. Org. Chem.* **2001**, *66*, 6919–6925.
- (61) Pierpoint, A. C.; Hapeman, C. J.; Torrents, A. Linear free energy study of ring-substituted aniline ozonation for developing treatment of aniline-based pesticide wastes. *J. Agric. Food Chem.* **2001**, *49*, 3827–3832.
- (62) Pankratov, A. N.; Uchaeva, I. M.; Doronin, S. Y.; Chernova, R. K. Correlations between the basicity and proton affinity of substituted anilines. *J. Struct. Chem.* **2001**, *42*, 739–746.
- (63) Lee, C.; Schmidt, C.; Yoon, J.; von Gunten, U. Oxidation of N-Nitrosodimethylamine (NDMA) Precursors with Ozone and Chlorine Dioxide: Kinetics and Effect on NDMA Formation Potential. *Environ. Sci. Technol.* **2007**, *41*, 2056–2063.
- (64) Lucida, H.; Parkin, J. E.; Sunderland, V. B. Kinetic study of the reaction of sulfamethoxazole and glucose under acidic conditions. *Int. J. Pharm.* **2000**, *202*, 47–62.
- (65) Dodd, M. C.; Buffle, M.-O.; von Gunten, U. Oxidation of Antibacterial Molecules by Aqueous Ozone: Moiety-Specific Reaction Kinetics and Application to Ozone-Based Wastewater Treatment. *Environ. Sci. Technol.* **2006**, *40*, 1969–1977.
- (66) Benner, J.; Salhi, E.; Ternes, T.; von Gunten, U. Ozonation of reverse osmosis concentrate: Kinetics and efficiency of beta blocker oxidation. *Water Res.* **2008**, *42*, 3003–3012.
- (67) Xiong, F.; Graham, N. J. D. Rate Constants for Herbicide Degradation by Ozone. *Ozone Sci. Eng.* **1992**, *14*, 283–301.
- (68) Muñoz, F.; von Sonntag, C. Determination of fast ozone reactions in aqueous solution by competition kinetics. *J. Chem. Soc. Perkin Trans. 2* **2000**, 661–664.

- (69) Huber, M. M.; Göbel, A.; Joss, A.; Hermann, N.; Löffler, D.; McArdell, C. S.; Ried, A.; Siegrist, H.; Ternes, T. A.; von Gunten, U. Oxidation of Pharmaceuticals during Ozonation of Municipal Wastewater Effluents: A Pilot Study. *Environ. Sci. Technol.* **2005**, *39*, 4290–4299.
- (70) Leitzke, A.; Flyunt, R.; Theruvathu, J. A.; von Sonntag, C. Ozonolysis of vinyl compounds, CH₂=CH-X, in aqueous solution—the chemistries of the ensuing formyl compounds and hydroperoxides. *Org. Biomol. Chem.* **2003**, *1*, 1012–1019.
- (71) Dowideit, P.; von Sonntag, C. Reaction of Ozone with Ethene and Its Methyl- and Chlorine-Substituted Derivatives in Aqueous Solution. *Environ. Sci. Technol.* **1998**, *32*, 1112–1119.
- (72) Leitzke, A.; Sonntag, C. von. Ozonolysis of Unsaturated Acids in Aqueous Solution: Acrylic, Methacrylic, Maleic, Fumaric and Muconic Acids. *Ozone Sci. Eng.* **2009**, *31*, 301–308.
- (73) Jans, U. Doctoral thesis. *ETH Zurich* **1996**.
- (74) Leitzke, A.; Reisz, E.; Flyunt, R.; von Sonntag, C. The reactions of ozone with cinnamic acids: formation and decay of 2-hydroperoxy-2-hydroxyacetic acid. *J. Chem. Soc. Perkin Trans. 2* **2001**, 793–797.
- (75) Pryor, W. A.; Giamalva, D. H.; Church, D. F. Kinetics of ozonation. 2. Amino acids and model compounds in water and comparisons to rates in nonpolar solvents. *J. Am. Chem. Soc.* **1984**, *106*, 7094–7100.
- (76) Mei, X.; Wolf, C. Neutral and Ionic Supramolecular Structures of Unsaturated Dicarboxylic Acids and Acridine: Significance of Molecular Geometry and Proton Transfer. *European J. Org. Chem.* **2004**, *2004*, 4340–4347.
- (77) Ramseier, M. K.; von Gunten, U. Mechanisms of Phenol Ozonation—Kinetics of Formation of Primary and Secondary Reaction Products. *Ozone Sci. Eng.* **2009**, *31*, 201–215.
- (78) Onstad, G. D.; Strauch, S.; Meriluoto, J.; Codd, G. A.; von Gunten, U. Selective Oxidation of Key Functional Groups in Cyanotoxins during Drinking Water Ozonation. *Environ. Sci. Technol.* **2007**, *41*, 4397–4404.
- (79) Takács-Novák, K.; Box, K. J.; Avdeef, A. Potentiometric pKa determination of water-insoluble compounds: validation study in methanol/water mixtures. *Int. J. Pharm.* **1997**, *151*, 235–248.
- (80) Straub, J. O. An environmental risk assessment for oseltamivir (Tamiflu) for sewage works and surface waters under seasonal-influenza- and pandemic-use conditions. *Ecotoxicol. Environ. Saf.* **2009**, *72*, 1625–1634.
- (81) Mestankova, H.; Schirmer, K.; Escher, B. I.; von Gunten, U.; Canonica, S. Removal of the antiviral agent oseltamivir and its biological activity by oxidative processes. *Environ. Pollut.* **2012**, *161*, 30–35.
- (82) McFarland, J. W.; Berger, C. M.; Froshauer, S. A.; Hayashi, S. F.; Hecker, S. J.; Jaynes, B. H.; Jefson, M. R.; Kamicker, B. J.; Lipinski, C. A.; Lundy, K. M.; et al. Quantitative structure-activity relationships among macrolide antibacterial agents: in vitro and in vivo potency against *Pasteurella multocida*. *J. Med. Chem.* **1997**, *40*, 1340–1346.
- (83) Richards, E. G.; Fasman, G. D. Handbook of Biochemistry and Molecular Biology: Nucleic Acids. by CD Fasman, CRC, Clevel. **1975**, *1*, 197.
- (84) Theruvathu, J. A.; Flyunt, R.; Aravindakumar, C. T.; von Sonntag, C. Rate constants of ozone reactions with DNA, its constituents and related compounds. *J. Chem. Soc. Perkin Trans. 2* **2001**, 269–274.
- (85) Major, D. T.; Laxer, A.; Fischer, B. Protonation Studies of Modified Adenine and Adenine Nucleotides by Theoretical Calculations and 15 N NMR. *J. Org. Chem.* **2002**, *67*, 790–802.
- (86) Turecek, F.; Chen, X. Protonated adenine: tautomers, solvated clusters, and dissociation mechanisms. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1713–1726.
- (87) Markowski, V.; Sullivan, G. R.; Roberts, J. D. Nitrogen-15 nuclear magnetic resonance spectroscopy of some nucleosides and nucleotides. *J. Am. Chem. Soc.* **1977**, *99*, 714–718.
- (88) Ishizaki, K.; Shinriki, N.; Ueda, T. Degradation of nucleic acids with ozone. V. Mechanism of action of ozone on deoxyribonucleoside 5'-monophosphates. *Chem. Pharm. Bull. (Tokyo)*. **1984**, *32*, 3601–3606.
- (89) Elguero, J. *Tautomerism of heterocycles*; Academic Press, 1976.
- (90) Kristl, A.; Mrhar, A.; Kozjek, F. The ionisation properties of acyclovir and deoxyacyclovir. *Int. J. Pharm.* **1993**, *99*, 79–82.

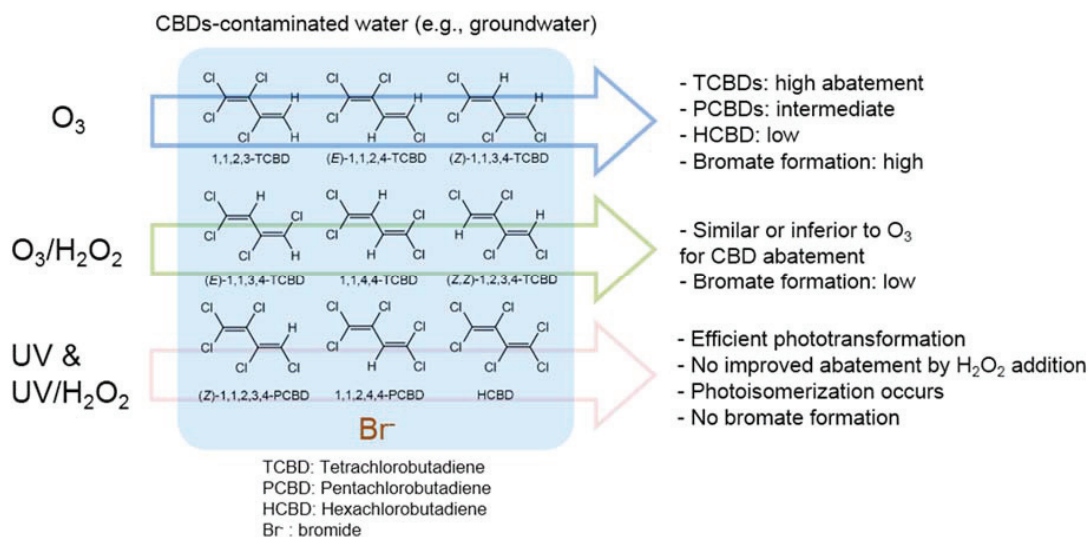
Supporting information for chapter 2

- (91) Verdolino, V.; Cammi, R.; Munk, B. H.; Schlegel, H. B. Calculation of pK_a values of nucleobases and the guanine oxidation products guanidinohydantoin and spiroiminodihydantoin using density functional theory and a polarizable continuum model. *J. Phys. Chem. B* **2008**, *112*, 16860–16873.
- (92) Prasse, C.; Wagner, M.; Schulz, R.; Ternes, T. A. Oxidation of the antiviral drug acyclovir and its biodegradation product carboxy-acyclovir with ozone: kinetics and identification of oxidation products. *Environ. Sci. Technol.* **2012**, *46*, 2169–2178.
- (93) Miles, H. T.; Howard, F. B.; Frazier, J. Tautomerism and Protonation of Guanosine. *Science*. **1963**, *142*, 1458–1463.
- (94) Broséus, R.; Vincent, S.; Aboulfadl, K.; Daneshvar, A.; Sauvé, S.; Barbeau, B.; Prévost, M. Ozone oxidation of pharmaceuticals, endocrine disruptors and pesticides during drinking water treatment. *Water Res.* **2009**, *43*, 4707–4717.
- (95) Barron, E.; Deborde, M.; Rabouan, S.; Mazellier, P.; Legube, B. Kinetic and mechanistic investigations of progesterone reaction with ozone. *Water Res.* **2006**, *40*, 2181–2189.
- (96) Roth, B.; Strelitz, J. Z. Protonation of 2,4-diaminopyrimidines. I. Dissociation constants and substituent effects. *J. Org. Chem.* **1969**, *34*, 821–836.
- (97) Rieder, J. Physiklisch-chemische und biologische Untersuchungen an Sulfonamiden. 1. Mitt. : Pharmakologisch interessante physikalisch-chemische Merkmale von 21 Sulfonamiden und 6 Sulfonamid-Metaboliten. *Arzneimittel-Forsch.* **1963**, *13*, 81.
- (98) Barbosa, J.; Barrón, D.; Cano, J.; Jiménez-Lozano, E.; Sanz-Nebot, V.; Toro, I. Evaluation of electrophoretic method versus chromatographic, potentiometric and absorptiometric methodologies for determining pK_a values of quinolones in hydroorganic mixtures. *J. Pharm. Biomed. Anal.* **2001**, *24*, 1087–1098.
- (99) Muñoz, F.; von Sonntag, C. The reactions of ozone with tertiary amines including the complexing agents nitrilotriacetic acid (NTA) and ethylenediaminetetraacetic acid (EDTA) in aqueous solution. *J. Chem. Soc. Perkin Trans. 2* **2000**, 2029–2033.
- (100) Weber, E. J.; Kenneke, J. F. SPARC (<http://www.epa.gov/athens/research/projects/sparc>). *US EPA, Natl. Expo. Res. Lab. Athens, GA*.
- (101) Christensen, J. J.; Izatt, R. M.; Wrathall, D. P.; Hansen, L. D. Thermodynamics of proton ionization in dilute aqueous solution. Part XI. pK , ΔH° , and ΔS° values for proton ionization from protonated amines at 25°. *J. Chem. Soc. A Inorganic, Phys. Theor.* **1969**, 1212.
- (102) Hoigné, J.; Bader, H.; Haag, W. ; Staehelin, J. Rate constants of reactions of ozone with organic and inorganic compounds in water—III. Inorganic compounds and radicals. *Water Res.* **1985**, *19*, 993–1004.
- (103) Koskinen, A. M. P.; Rapoport, H. Synthetic and conformational studies on anatoxin-a: a potent acetylcholine agonist. *J. Med. Chem.* **1985**, *28*, 1301–1309.
- (104) Sein, M. M.; Zedda, M.; Tuerk, J.; Schmidt, T. C.; Golloch, A.; Sonntag, C. von. Oxidation of Diclofenac with Ozone in Aqueous Solution. *Environ. Sci. Technol.* **2008**, *42*, 6656–6662.
- (105) Hall, H. K. Potentiometric Determination of the Base Strength of Amines in Non-protolytic Solvents. *J. Phys. Chem.* **1956**, *60*, 63–70.
- (106) Haag, W. R.; Hoigné, J. Ozonation of water containing chlorine or chloramines. Reaction products and kinetics. *Water Res.* **1983**, *17*, 1397–1402.
- (107) Tekle-Röttering, A.; Schmidt, W.; Schmidt, C.; von Sonntag, C. Kinetics and OH radical yield of the reaction of ozone with aniline and morpholine. *Gordon Conference Proceedings*. **2011**.
- (108) Qiang, Z.; Adams, C. Potentiometric determination of acid dissociation constants (pK_a) for human and veterinary antibiotics. *Water Res.* **2004**, *38*, 2874–2890.
- (109) Qiang, Z.; Adams, C.; Surampalli, R. Determination of Ozonation Rate Constants for Lincomycin and Spectinomycin. *Ozone Sci. Eng.* **2004**, *26*, 525–537.
- (110) Shi, Y.-J.; Humphrey, G.; Maligres, P. E.; Reamer, R. A.; Williams, J. M. Highly Regioselective DABCO-Catalyzed Nucleophilic Aromatic Substitution (S_NAr) Reaction of Methyl 2,6-Dichloronicotinate with Phenols. *Adv. Synth. Catal.* **2006**, *348*, 309–312.
- (111) Handbook, C. R. C. Handbook of Chemistry and Physics. *Lide, D. R., Ed* **1986**.
- (112) Barbosa, J.; Barrón, D.; Jiménez-Lozano, E.; Sanz-Nebot, V. Comparison between capillary electrophoresis, liquid chromatography, potentiometric and spectrophotometric techniques for

- evaluation of pKa values of zwitterionic drugs in acetonitrile–water mixtures. *Anal. Chim. Acta* **2001**, *437*, 309–321.
- (113) Serjeant, E. P.; Dempsey, B. *Ionisation constants of organic acids in aqueous solution*; Pergamon, 1979; Vol. 23.
- (114) Pospíšilová, M.; Polášek, M.; Jokl, V. Determination of tramadol in various dosage forms by capillary isotachopheresis. *J. Pharm. Biomed. Anal.* **1998**, *18*, 777–783.
- (115) Zimmermann, S. G.; Schmukat, A.; Schulz, M.; Benner, J.; von Gunten, U.; Ternes, T. A. Kinetic and mechanistic investigations of the oxidation of tramadol by ferrate and ozone. *Environ. Sci. Technol.* **2012**, *46*, 876–884.
- (116) Chen, W.; Wu, C.; Elovitz, M.; Linden, K.; Suffet, I. Reactions of thiocarbamate, triazine and urea herbicides, RDX and benzenes on EPA Contaminant Candidate List with ozone and with hydroxyl radicals. *Water Res.* **2008**, *42*, 137–144.
- (117) Muñoz, F.; Mvula, E.; Braslavsky, S. E.; von Sonntag, C. Singlet dioxygen formation in ozone reactions in aqueous solution. *J. Chem. Soc. Perkin Trans. 2* **2001**, 1109–1116.
- (118) Flyunt, R.; Makogon, O.; Schuchmann, M. N.; Asmus, K.-D.; von Sonntag, C. OH-Radical-induced oxidation of methanesulfinic acid. The reactions of the methanesulfonyl radical in the absence and presence of dioxygen. *J. Chem. Soc. Perkin Trans. 2* **2001**, 787–792.

Chapter 3.

Abatement of polychloro-1,3-butadienes in aqueous solution by ozone, UV-photolysis, and advanced oxidation processes (O_3/H_2O_2 and UV/H_2O_2)



Lee, M., Merle, T., Rentsch, D., Canonica, S., von Gunten, U., *in preparation*

This chapter was prepared by Lee, M. (overall contribution >80%) with scientific advices and editorial comments by all coauthors, particularly under a joint supervision of Canonica, S. and von Gunten, U.

Merle, T. (Eawag) performed the LC/MS/MS analysis for the quantification of micropollutants.

Rentsch, D. (EMPA) performed the NMR analysis for a 1,1,3,4-tetrachlorobutadiene isomer mixture and (Z)-1,1,2,3,4-pentachlorobutadiene.

Abstract

The abatement of 9 polychloro-1,3-butadienes (CBDs) in aqueous solution by ozone, UV-C(254nm) photolysis, and the corresponding advanced oxidation processes (AOPs) (i.e., O₃/H₂O₂ and UV/H₂O₂) was investigated. The following parameters were determined for 9 CBDs: Second-order rate constants for the reactions of CBDs with ozone (k_{O_3}) ($<0.1 - 7.9 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$) or with hydroxyl radicals (k_{OH}) ($0.9 \times 10^9 - 6.5 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$), photon fluence-based rate constants (k') ($210 - 2730 \text{ m}^2 \text{ einstein}^{-1}$), and quantum yields (Φ) ($0.03 - 1.23 \text{ mol einstein}^{-1}$). During ozonation of CBDs in a natural groundwater, appreciable abatements ($>50\%$ at specific ozone doses of $0.5 \text{ gO}_3/\text{gDOC}$ to $\sim 100\%$ at $\geq 1.0 \text{ gO}_3/\text{gDOC}$) were achieved for tetra-CBDs followed by (*Z*)-1,1,2,3,4-penta-CBD and hexa-CBD. This is consistent with the magnitude of the determined k_{O_3} and k_{OH} . The formation of bromate, a potentially carcinogenic ozonation by-product, could be significantly reduced by addition of H₂O₂. For a typical UV disinfection dose (400 J/m^2), various extents of phototransformations (10-90%) could be achieved. However, the efficient formation of photoisomers from CBDs with *E/Z* configuration must be taken into account because of their potential residual toxicity. Under UV-C(254nm) photolysis conditions, no significant effect of H₂O₂ addition on CBDs abatement was observed due to an efficient direct phototransformation of CBDs.

3.1. Introduction

Poly-chloro-1,3-butadiene (CBD) congeners (i.e., tetra-, penta-, hexa-CBD) have been detected over the last four decades in river water,¹⁻⁵ groundwater,⁶⁻⁸ and wastewater^{4,9} worldwide in the concentration range of tens of ng/L to a few µg/L. Among CBDs, hexa-chloro-1,3-butadiene (HCBD) has various industrial applications such as an intermediate in manufacturing rubber compounds, as a solvent, as a fumigant, etc.¹⁰ Moreover, HCBD is inadvertently produced as a by-product from the manufacturing of chlorinated hydrocarbons such as trichloroethene (TCE) and tetrachloroethene (PCE).¹⁰ Due to the potential hazardous impacts on humans and aquatic organisms,¹⁰⁻¹² HCBD has been subject to environmental guidelines worldwide: 0.6 µg/L for drinking water by World Health Organization¹¹ and an environmental quality standard value of 0.1 µg/L for inland surface water by the European Commission have been set.¹³⁻¹⁵ In contrast, little attention was drawn to other CBDs such as pentachloro-1,3-butadienes (PCBDs) and tetrachloro-1,3-butadienes (TCBDs).^{1,6-8} Such CBDs are known to be formed from reductive microbial dechlorination of HCBD under anaerobic conditions.^{16,17} Dechlorinated products are not necessarily less harmful, e.g., vinyl chloride, which is carcinogenic,¹⁸ is produced from the reductive dechlorination of PCE and TCE.¹⁹ Therefore, the ecotoxicological impact of other CBDs than HCBD should not be overlooked. A positive clastogenic activity was recently reported for 3 TCBDs and 2 PCBDs and a target value of 75 ng/L as the sum of TCBDs and PCBDs was proposed for drinking water using the threshold of toxicological concern (TTC) concept.⁸

Physical treatment technologies such as activated carbon and air stripping are commonly applied for eliminating highly chlorinated butadienes due to high hydrophobicity and volatility (e.g., $\log K_{ow}=4.78$ ²⁰ and Henry's Law Constant = 15.3×10^{-3} atm m³/mol at 25°C²¹ for HCBD). Alternatively, other treatment techniques such as ozonation, UV photolysis, or their combination with H₂O₂, i.e., O₃/H₂O₂ or UV/H₂O₂, also known as advanced oxidation processes (AOPs), can be potentially considered. AOPs are chemical treatment processes in which hydroxyl radicals (*OH) are produced *in situ*.²² However, the current lack of information on key parameters for such treatment options hampers the assessment of the abatement efficiency for CBDs, namely, second-order rate constants for the reaction of CBDs with ozone (k_{O_3}) or hydroxyl radicals (k_{OH}) for ozonation, O₃/H₂O₂, and UV/H₂O₂^{23,24} and a photon fluence-based pseudo first-order rate constant (k') for UV-based processes such as UV photolysis or UV/H₂O₂,^{25,26} respectively.

In the present chapter, k_{O_3} , k_{OH} , and k' for UV photolysis at 254 nm were experimentally determined for 9 CBDs, i.e., 6 TCBD isomers (1,1,2,3-, (*E*)-1,1,2,4-, (*Z*)-1,1,3,4-, (*E*)-1,1,3,4-, 1,1,4,4-, and (*Z,Z*)-1,2,3,4-), 2 PCBD isomers ((*Z*)-1,1,2,3,4- and 1,1,2,4,4-), and HCBD (Figure 3.1). Moreover, a previously developed quantum chemical model²⁷ was applied to predict k_{O_3} -values for 9 CBDs, which were then compared to the experimental k_{O_3} -values. Finally, the treatment efficacy for the selected CBDs by ozonation, direct UV photolysis, and the AOPs O₃/H₂O₂ and UV/H₂O₂ was evaluated in a natural groundwater matrix. For ozonation and O₃/H₂O₂, the abatement of other common micropollutants and bromate formation was also evaluated.

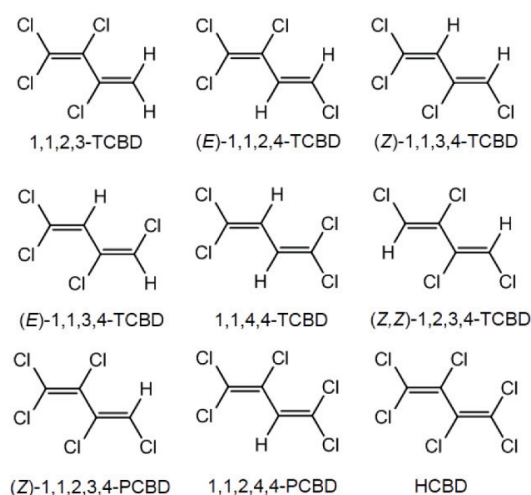


Figure 3.1. The selected 9 tetra-, penta-, and hexa-chlorobutadienes (TCBD, PCBD, and HCBD, respectively) for this study.

3.2. Materials and methods

3.2.1. Standards and reagents

All chemicals used are summarized in Text S3.1 in the Supporting Information.

3.2.2. Analytical methods

9 CBDs, *p*-chlorobenzoic acid (*p*CBA), atrazine, 5 micropollutants (benzotriazole, carbamazepine, diatrizoate, iopamidol, and tramadol), bromate, ozone, and H₂O₂ were quantified using appropriate analytical methods (Text S3.2).

3.2.3. Determination of second-order rate constants for the reactions of CBDs with ozone (k_{O_3}) and hydroxyl radicals (k_{OH})

k_{O_3} for CBDs was determined by simultaneously measuring the abatement of CBDs and the ozone decay using Eq. S1 (Text S3.3). Determination of k_{OH} for CBDs was performed by competition kinetics (Eqs. S3.2-S3.4) using *p*CBA as a competitor. $\cdot OH$ were produced by photolysis of H₂O₂. More details about analytical instruments and experimental procedures are provided in Text S3.3.

3.2.4. Determination of photon fluence-based first-order rate constants (k'), molar absorption coefficients (ϵ), and quantum yields (Φ) for phototransformation of CBDs at UV 254 nm

k' (m² einstein⁻¹) for 9 CBDs for UV irradiation at 254 nm were determined according to Eq. 3.1.

$$k' = \frac{k_{\text{CBD},254\text{nm}}}{E_p^0} = \frac{2.303\Phi_{\text{ATZ},254\text{nm}}\epsilon_{\text{ATZ},254\text{nm}}k_{\text{CBD},254\text{nm}}}{k_{\text{ATZ},254\text{nm}}} \quad (3.1)$$

$k_{\text{CBD},254\text{nm}}$ and $k_{\text{ATZ},254\text{nm}}$ (s^{-1}) are a time-based pseudo first-order rate constant (s^{-1}) for the phototransformation of a CBD and atrazine at 254 nm, respectively, E_p^0 is a photon fluence rate ($\text{einstein m}^{-2} \text{s}^{-1}$) derived based on actinometry using atrazine (ATZ),²⁶ and $\epsilon_{\text{ATZ},254\text{nm}}$ and $\Phi_{\text{ATZ},254\text{nm}}$ are $386 \text{ m}^2 \text{ mol}^{-1}$ and $0.046 \text{ mol einstein}^{-1}$ for atrazine at 254nm, respectively.²⁶ Direct UV photolysis to obtain $k_{\text{CBD},254\text{nm}}$ was measured using the same photoreactor used for the k_{OH} determination (Text S3.3), but employing a low pressure (LP) mercury lamp (model TNN 15/32, Heraeus Noblelight, Hanau, Germany) emitting monochromatic UV light at 254 nm and a quartz cooling jacket. Sample solutions containing individual CBDs (1.0–3.5 μM , pH 7, 5mM phosphate buffer) prepared in quartz tubes were subject to UV irradiation. 5 μM of an atrazine solution at pH 7 (5mM phosphate buffer) separately prepared was irradiated simultaneously to obtain the corresponding $k_{\text{ATZ},254\text{nm}}$. $\epsilon_{\text{CBD},254\text{nm}}$ for the 9 CBDs were determined based on the Beer-Lambert law by monitoring the absorbance at 254 nm using a quartz cuvette ($l = 10 \text{ cm}$) for various concentrations of CBDs (0.3 - 6 μM). $\Phi_{\text{CBD},254\text{nm}}$ was calculated according to Eq. 3.2.²⁵

$$\Phi_{\text{CBD},254\text{nm}} = k' / (2.303 \epsilon_{\text{CBD},254\text{nm}}) \quad (3.2)$$

3.2.5. Ozonation, direct UV photolysis (254nm), and advanced oxidation of the selected CBDs and micropollutants in a natural groundwater

A natural groundwater was sampled from the Hardwald site (Latitude 47°32' N and Longitude 7°37' E) in Switzerland (pH = 7.95 – 8.05, DOC = 0.5 mgC/L, alkalinity as $\text{HCO}_3^- = 2.84 \text{ mM}$, and bromide = 48 $\mu\text{g/L}$). For ozonation, the groundwater was spiked with a mixture of the selected CBDs ((*E*)-1,1,2,4-, (*E*)-1,1,3,4-, (*Z*)-1,1,3,4-, 1,1,4,4-, and (*Z,Z*)-1,2,3,4-TCBDs, (*Z*)-1,1,2,3,4-PCBD, and HCBD, 5 - 9 $\mu\text{g/L}$) and 5 micropollutants (2 $\mu\text{g/L}$ each). Varying ozone doses (0.25, 0.5, 1, 2, and 4 mg/L) were investigated in the absence and presence of H_2O_2 with ratios $\text{O}_3:\text{H}_2\text{O}_2$ (w/w) of 0.5, 1, and 2. UV photolysis for the selected CBDs was carried out in the absence and presence of H_2O_2 (0, 2.5, and 5.0 mg/L) up to a UV dose of 8600 J/m^2 . Light screening by groundwater matrix components, estimated as a light attenuation factor as described elsewhere,²⁸ was negligible (<2%). Therefore, no light screening correction for the UV dose was applied.

3.2.6. Quantum chemical computations for the prediction of k_{O_3} of CBDs

Quantum chemical computations were conducted using Gaussian 09 (Revision C.01).²⁹ *Ab initio* Hartree-Fock (HF) and the Density Functional Theory (DFT)-B3LYP method with the 6-311++G** basis set were used. The integral equation formalism polarizable continuum model (IEF-PCM)³⁰ was used for all the computations to account for the solvent effect applying a dielectric constant of 78.3 for water. Natural bond orbital (NBO) analyses were conducted using the NBO 3.1 program³¹ to obtain a NBO energy ($E_{\text{NBO}, \text{C-C}(\pi)}$) corresponding to the π orbital of a C-C double bond of the CBDs. More details about the model development and the application procedure are given elsewhere.²⁷

3.2.7. Kinetic modeling

Kinetic simulations using Kintecus 5.00³² were performed to derive the respective k' for photoisomerization and other phototransformation processes of (*Z*)-1,1,3,4- and (*E*)-1,1,3,4-TCBDs in a mixture (Text S3.5) and to derive a simulated abatement of CBDs based on derived k' -values (Table 3.2).

3.3. Results and discussion

3.3.1. Kinetics of the reactions of CBDs with ozone ($k_{O_3, \text{exp}}$)

The experimentally determined k_{O_3} ($k_{O_3, \text{exp}}$) are shown in Table 3.1. Overall, $k_{O_3, \text{exp}}$ -values for the 9 CBDs ranged over more than 4 orders of magnitude; TCBDs (1.6×10^2 – $7.8 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$), PCBDs (1 – $10 \text{ M}^{-1}\text{s}^{-1}$), and HCBd ($< 0.1 \text{ M}^{-1}\text{s}^{-1}$). Note that (*E*)-1,1,3,4-TCBD and (*Z*)-1,1,3,4-TCBD was ozonated as a mixture as received from the supplier. Nonetheless, the obtained $k_{O_3, \text{exp}}$ -values are considered to be valid because the ozone concentration is in excess (>15 fold) to 1,1,3,4-TCBDs ensuring that each isomer degrades independently of the other isomer. $k_{O_3, \text{exp}}$ for HCBd could not be determined because no appreciable degradation was observed under the experimental conditions. Thus, $k_{O_3, \text{exp}}$ for HCBd was estimated to be $< 0.1 \text{ M}^{-1}\text{s}^{-1}$ which is the reported $k_{O_3, \text{exp}}$ range for PCE.³³

Overall, with increasing chlorine substituents on CBDs from four to six, the $k_{O_3, \text{exp}}$ -values decrease. This suggests the electron-withdrawing effect of chlorine substituents on $k_{O_3, \text{exp}}$ for CBDs is additive, which is in agreement with previous observations for chlorinated ethenes.³⁴ The range of $k_{O_3, \text{exp}}$ for TCBD, PCBD, and HCBd is similar to the range of $k_{O_3, \text{exp}}$ for dichloroethenes (DCEs) (1.1×10^2 – $7.9 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$),³⁴ TCE ($14 \text{ M}^{-1}\text{s}^{-1}$),³⁴ and PCE ($< 0.1 \text{ M}^{-1}\text{s}^{-1}$),³³ respectively.

Table 3.1. Experimental k_{O_3} ($k_{O_3, \text{exp}}$) and predicted k_{O_3} ($k_{O_3, \text{pred}}$) or experimental $k_{\bullet\text{OH}}$ for the reactions of CBDs with ozone or hydroxyl radicals.

Compound	$k_{O_3, \text{exp}}^a$ (N^b)	$k_{O_3, \text{pred}}^c$		$k_{\bullet\text{OH}}^a$ (N^b), $\times 10^9 \text{ M}^{-1}\text{s}^{-1}$
		HF/6-311++G**	B3LYP/6-311++G**	
1,1,2,3-TCBD	$(3.0 \pm 0.6) \times 10^2$ (4)	1.4×10^3	1.6×10^3	3.4 ± 0.5 (4)
(<i>E</i>)-1,1,2,4-TCBD	$(7.8 \pm 1.0) \times 10^3$ (4)	2.2×10^3	2.5×10^3	6.0 ± 1.6 (6)
(<i>E</i>)-1,1,3,4-TCBD	$(1.1 \pm 0.2) \times 10^3$ (3)	3.4×10^2	5.0×10^2	5.3 ± 1.8 (3)
(<i>Z</i>)-1,1,3,4-TCBD	$(4.0 \pm 2.4) \times 10^2$ (3)	3.6×10^2	5.7×10^2	5.7 ± 3.2 (3)
1,1,4,4-TCBD	$(2.7 \pm 0.4) \times 10^2$ (3)	6.9×10^2	1.0×10^3	5.4 ± 0.8 (5)
(<i>Z,Z</i>)-1,2,3,4-TCBD	$(1.6 \pm 0.3) \times 10^2$ (3)	4.2×10^2	6.0×10^2	7.1 ± 1.5 (6)
(<i>Z</i>)-1,1,2,3,4-PCBD	0.8 ± 0.3 (4)	1.5×10^2	2.4×10^2	2.1 ± 0.8 (3)
1,1,2,4,4-PCBD	10.0 ± 2.3 (5)	1.5×10^2	1.8×10^2	3.9 ± 0.9 (5)
HCBd	< 0.1	25	47	0.9 ± 0.7 (3)

^aExperimental k -values ($\text{M}^{-1}\text{s}^{-1}$) for ozone and hydroxyl radicals given as a mean value ($\pm 95\%$ confidence interval determined by the Student's t test), ^bnumber of replicates, ^cpredicted k_{O_3} -value ($\text{M}^{-1}\text{s}^{-1}$) derived from Eq. S3.6 in SI. The experimental data used to determine k_{O_3} and $k_{\bullet\text{OH}}$ are presented in Figures S3.2 and S3.3.

Differences in $k_{O_3, \text{exp}}$ -values were noticeable between isomers with the same number of chlorine substituents: ~ 10 -fold or ~ 50 -fold for the two PCBD isomers or the six TCBDs isomers, respectively. This indicates that differing positions of the chlorine substituents influence electronic interactions of

CBDs with ozone. Although this variation is not well understood, the *cis* effect may explain the specific observation that $k_{O_3, \text{exp}}$ for *cis*- or (*Z*)-1,1,3,4-TCBD was lower than $k_{O_3, \text{exp}}$ for *trans*- or (*E*)-1,1,3,4-TCBD by a factor of ~ 2.4 . The *cis* effect is a phenomenon that the *cis*-isomer is more stable compared to the *trans*-isomer for 1,2-difluoroethene and 1,2-DCE^{35–37} and was also observed in k_{O_3} for 1,2-DCE: k_{O_3} for (*Z*)-1,2-DCE ($5.4 \times 10^2 \text{ M}^{-1}\text{s}^{-1}$) is ten times lower than k_{O_3} for (*E*)-1,2-DCE ($6.5 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$).³⁴

3.3.2. Prediction of rate constants for the reactions of CBDs with ozone ($k_{O_3, \text{pred}}$)

k_{O_3} -values for CBDs were predicted by taking into account (non)-planar CBD conformers using the previously developed quantum chemical model for olefins (Table 3.1).²⁷ Detailed procedures and discussion for deriving a predicted k_{O_3} are presented in Text S3.4. A k_{O_3} prediction procedure is briefly demonstrated for (*E*)-1,1,2,4-TCBD in Figure 3.2 where two conformers corresponding to minima in the potential energy curve (PEC) were found: The first conformer has a dihedral angle (i.e., the angle between the planes comprising the atoms $C_1-C_2-C_3$ and $C_2-C_3-C_4$, respectively) of 0° and the other 116° (Text S3.4 for more details). The planar *s-trans* global minimum conformer ($\varphi=0^\circ$) of (*E*)-1,1,2,4-TCBD reads -11.50 eV and -12.77 eV of $E_{\text{NBO}, C-C(\pi)}$ (open circles) for the two C=C bonds, respectively. Based on a linear regression equation ($\log k_{O_3} = 1.32 \times E_{\text{NBO}, C-C(\pi)} + 18.54$ for the HF/6-311++G** method),²⁷ $k_{O_3, \text{global}(C1=C2)} = 48 \text{ M}^{-1}\text{s}^{-1}$ and $k_{O_3, \text{global}(C3=C4)} = 2265 \text{ M}^{-1}\text{s}^{-1}$ were obtained, respectively, resulting in $2314 \text{ M}^{-1}\text{s}^{-1}$ for a species-specific second-order rate constant ($k_{O_3, \text{global}}$) for the global conformer based on Eq. S3.5. As expected, the C=C bond with a higher chlorine substitution had a lower k_{O_3} . In the same manner, $k_{O_3, \text{local}}$ of $1378 \text{ M}^{-1}\text{s}^{-1}$ was obtained for the *gauche* local minimum conformer ($\varphi=116^\circ$). Applying Eq. S3.6 with k_{O_3} -values for global and local conformers and their relative populations (p_i in Table S3.3), $k_{O_3, \text{pred}}$ for (*E*)-1,1,2,4-CBD was predicted to be $2.3 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$ with a prediction error of a factor of ~ 3 from $k_{O_3, \text{exp}} = 7.8 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$.

Similarly, $k_{O_3, \text{pred}}$ -values were obtained for all 9 CBDs for both the HF and the B3LYP methods (Table 3.1). The k_{O_3} prediction was satisfactory for the 6 TCBDs: 5 out of 6 TCBDs were predicted within a factor of 4 for both methods (Figure S3.6). In contrast, an overestimation occurred for the two PCBs and HCBd by more than an order of magnitude. This may be caused by steric hindrances by chlorine substituents for the orbital interaction between ozone and PCBd/HCBd, which are not taken into account in $E_{\text{NBO}, C-C(\pi)}$ calculations. Alternatively, the C=C bond is so electron-poor that the formation of a cyclic ozonide^{24,34} is no longer dominant. Therefore, the reactivity of PCBs and HCBd with ozone cannot be interpreted by only $E_{\text{NBO}, C-C(\pi)}$.

For CBDs with multiple conformers (Table S3.3), $k_{O_3, \text{pred}}$ -values in Table 3.1 derived by considering both global and local minimum conformers (i.e., $k_{O_3, \text{global}}$, $k_{O_3, \text{local}}$, and their relative populations) based on Eq. 3.6 are almost the same as $k_{O_3, \text{global}}$ in Table S3.4. This suggests that the contribution of a local minimum conformer (i.e., $k_{O_3, \text{local}}$) to $k_{O_3, \text{pred}}$ is insignificant, which is understood by the predominant population of a global minimum (Table S3.3). Moreover, the difference between $k_{O_3, \text{global}}$ and $k_{O_3, \text{local}}$ is less than a factor of 2 except for 1,1,4,4-TCBD (~ 3 and ~ 4 folds for HF and B3LYP method, respectively). Therefore, it seems that $k_{O_3, \text{global}}$ can be used as a good approximation. However, it should be noted that this does not justify the general representability of $k_{O_3, \text{global}}$ for other compounds.

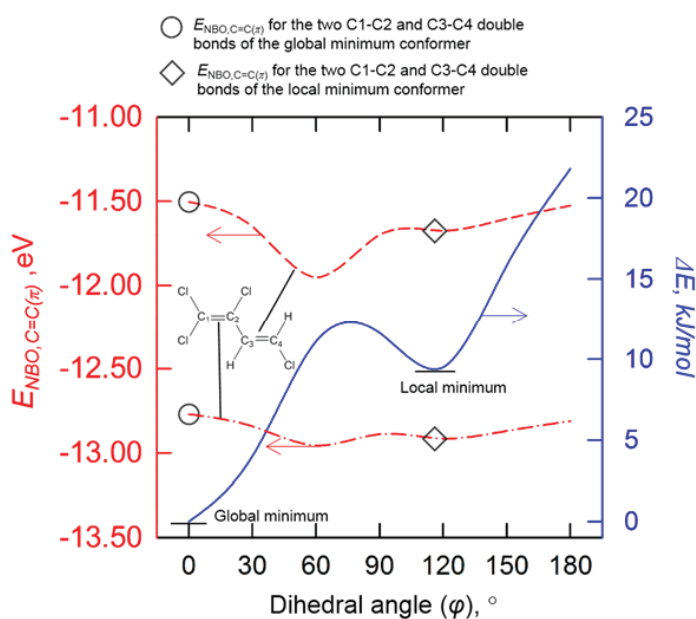


Figure 3.2. Profiles of the difference of a total electronic energy (blue solid line) (ΔE) of the (*E*)-1,1,2,4-TCBD conformers and the lowest energy of the (*E*)-1,1,2,4-TCBD conformer and the corresponding natural bond orbital energy ($E_{\text{NBO}, \text{C}=\text{C}(\pi)}$) (red dash-dot line and red dashed line for the $\text{C}_1\text{-C}_2$ and $\text{C}_3\text{-C}_4$ double bonds, respectively) of the π orbital of a C-C double bond of TCBD conformers as a function of the dihedral angle of the $\text{C}_2\text{-C}_3$ bond. All the energies were obtained at the HF/6-311++G** level.

3.3.3. Kinetics for the reaction of CBDs with hydroxyl radical (k_{OH})

Experimentally determined k_{OH} -values for 9 CBDs are presented in Table 3.1. The determined k_{OH} -values are higher than $5 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$ for all TCBDs except 1,1,2,3-TCBD ($3.4 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$). A lower reactivity with $\cdot\text{OH}$ was observed for the 2 PCBDs ($2.1 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$ and $3.9 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$), followed by the lowest reactivity for HCBD ($0.9 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$). Overall, a higher chlorine substitution of CBD leads to a lower reactivity with $\cdot\text{OH}$. Decreased k_{OH} with increasing chlorine substitution was also observed for chlorinated ethenes: $3.8 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$, $2.8 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$, and $2.0 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$ for *cis*-1,2-DCE, TCE, and PCE, respectively.³⁸

3.3.4. Determination of fluence-based first-order rate constants (k'), molar absorption coefficients (ϵ), and quantum yields (Φ)

Fluence-based first-order rate constants (k') and photoisomerization processes. The determined k' -values based on Eq. 1 ranged from $210 \pm 30 \text{ m}^2 \text{ einstein}^{-1}$ for 1,1,4,4-TCBD to $2730 \pm 440 \text{ m}^2 \text{ einstein}^{-1}$ for (*Z,Z*)-1,2,3,4-TCBD, respectively. Two k' -values were reported for (*E*)-1,1,3,4- and (*Z*)-1,1,3,4-TCBDs for separate photoisomerization and other phototransformation processes (see below).

A noticeable formation of photo-products was detected by GC/MS analyses for (*E*)-1,1,2,4-, (*Z*)- and (*E*)-1,1,3,4- as a mixture, (*Z,Z*)-1,2,3,4-TCBD and (*Z*)-1,1,2,3,4-PCBD (Figures 3.3 and S3.7). Because of their identical mass fragmentation patterns to those of the parent compounds (Figure S3.7), these products are considered to be the isomers of the parent CBDs, and are termed as photoisomers hereafter. No other confirmative examinations of their molecular structure than the mass fragmentation were used.

Only relative responses of the photoisomers are presented in Figure 3.3 for (*E*)-1,1,2,4- and (*Z,Z*)-1,2,3,4-TCBD and (*Z*)-1,1,2,3,4-PCBD because there were no commercially available standards. The relative response was calculated as the peak area of the parent CBD or its photoisomers divided by the initial peak area of the parent CBD (all peak areas corrected by normalization to the internal standard).

Photoisomerization of less than 40% relative to the parent CBD based on the relative response were observed for (*E*)-1,1,2,4-TCBD (Figure 3.3a) and (*Z*)-1,1,2,3,4-PCBD (Figure 3.3d), whereas a significant conversion of up to 85% was observed for (*Z,Z*)-1,2,3,4-TCBD (Figure 3.3b). For UV irradiation (254 nm) of the mixture of (*E*)- and (*Z*)-1,1,3,4-TCBD (Figure 3.3c), the (*E*)-1,1,3,4-TCBD concentration increased by ~10% compared to its initial concentration up to a UV photon fluence of 1 meinstein m^{-2} , while a significant degradation (>50%) of (*Z*)-1,1,3,4-TCBD occurred for the same UV dose. This indicates that a photoisomerization of the (*Z*)-1,1,3,4- to the (*E*)-1,1,3,4-TCBD occurs.

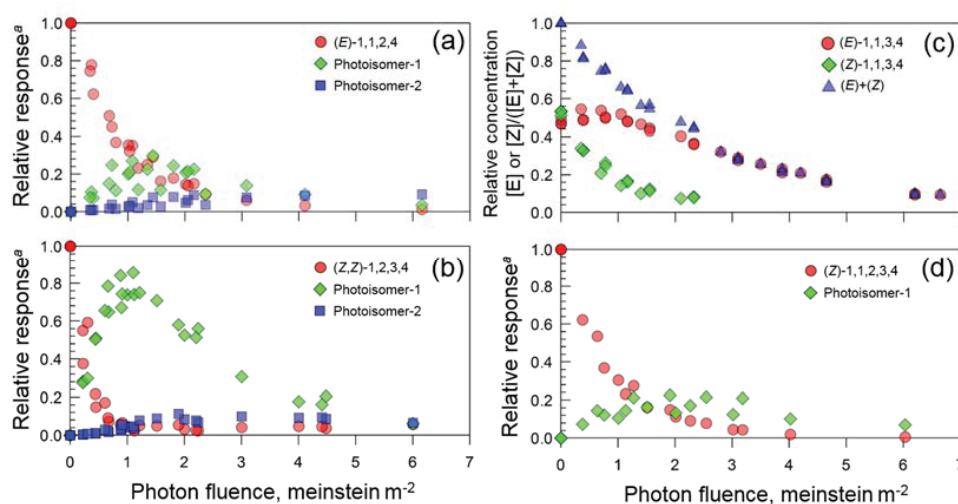
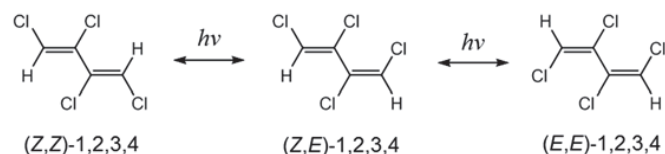


Figure 3.3. Evolution of (a) (*E*)-1,1,2,4-TCBD, (b) (*Z,Z*)-1,2,3,4-TCBD, (c) (*E*)-1,1,3,4- and (*Z*)-1,1,3,4-TCBD, and (d) (*Z*)-1,1,2,3,4-PCBD and their photoisomers as a function of the photon fluence (meinstein m^{-2}). ^aA relative response is the peak area of the parent CBDs or its photoisomers divided by the initial peak area of the parent CBD. All peak areas are normalized by the internal standard.

As rotation about a double bond is allowed in excited electronic states of organic compounds, all the CBDs with a *Z* or *E* configuration can undergo photoisomerization. Depending on the substitution

Chapter 3

pattern of the chlorine atoms, for TCBDs *Z* and *E* configuration pairs or (*Z,Z*), (*E,Z*) and (*E,E*) configuration triads are possible. In the case of (*Z,Z*)-1,2,3,4-TCBD, photoisomerization is proposed to occur following Scheme 3.1, which accounts for the formation of the two observed photoisomers (see Figure 3.3b). Thereby, the peaks with a higher relative response (photoisomer-1 in Figures 3.3b and S3.7) and a lower relative response (photoisomer-2) were assigned to be the (*Z,E*)-photoisomer and the (*E,E*)-photoisomer, respectively, assuming a stepwise conversions (Scheme 3.1).



Scheme 3.1 Proposed photoisomerization of (*Z,Z*)-1,2,3,4-TCBD to (*Z,E*)-1,2,3,4-TCBD and (*E,E*)-1,2,3,4-TCBD.

For (*E*)-1,1,2,4-TCBD two photoisomers were detected (Figure 3.3a and S3.7), although only one photoisomer is predicted for an *E-Z* photoisomerization reaction. The additional photoisomer could be the result of a photocyclization reaction. Indeed, the formation of cyclobutene and cyclopropene isomers has been shown to occur upon direct UV irradiation of 1,3-dienes.^{39,40} Only one photoisomer was observed for (*Z*)-1,1,2,3,4-PCBD, which was assigned to its *E* configuration.

$k_{\text{CBD},254\text{nm}}$ in Eq. 3.1 determined for photo-isomerizable CBDs takes into account both photoisomerization and other phototransformations. This propagates into $\Phi_{254\text{nm}}$ and k' reported in Table 3.2 as those are determined based on $k_{\text{CBD},254\text{nm}}$. Only for two 1,1,3,4-TCBD isomers, both photochemical processes could be differentiated by kinetic simulations because standards were available (see Text S3.5 and Figure S3.8 for more details), yielding the individual $\Phi_{254\text{nm}}$ - and k' -values in Table 3.2.

Table 3.2. $\epsilon_{254\text{nm}}$, $\Phi_{254\text{nm}}$, and k' determined for the 9 selected CBDs^a

Compound	$\epsilon_{254\text{nm}}^a$, $\text{m}^2 \text{mol}^{-1}$	$\Phi_{254\text{nm}}^b$, mol einstein ⁻¹	k'^a , $\text{m}^2 \text{einstein}^{-1}$
1,1,2,3-TCBD	620±80	0.22±0.04	310±40
(<i>E</i>)-1,1,2,4-TCBD	2490±140	0.14±0.04	820±370
(<i>E</i>)-1,1,3,4-TCBD	1580±140 ^c	0.05±0.03 ^d , 0.07±0.04 ^e	200±110 ^d , 270±150 ^e
(<i>Z</i>)-1,1,3,4-TCBD	1580±140 ^c	0.13±0.11 ^d , 0.27±0.06 ^e	1000±180 ^d , 460±400 ^e
1,1,4,4-TCBD	3210±100	0.03±0.004	210±30
(<i>Z,Z</i>)-1,2,3,4-TCBD	1620±140	0.73±0.13	2730±440
(<i>Z</i>)-1,1,2,3,4-PCBD	1110±470	0.32±0.26	810±560
1,1,2,4,4-PCBD	540±60	0.37±0.09	460±110
HCBD	230±60	0.95±0.28	510±110

^a95% confidence intervals determined by the student's *t* test are given for all the reported values. ^buncertainties estimated based on the error propagation through the multiplication of uncertainties in k' and $\epsilon_{254\text{nm}}$ for the respective CBDs. ^caverage $\epsilon_{254\text{nm}}$ for the isomer mixture of (*E*)-1,1,3,4-TCBD (43%) and (*Z*)-1,1,3,4-TCBD (57%), ^dfor photoisomerization, ^efor other phototransformations. The number of replicates for determination of k' for CBDs is 3 except for (*E*)-1,1,2,4- and (*Z,Z*)-1,2,3,4-TCBD for which the number is 4.

Molar absorption coefficients (ϵ). The determined $\epsilon_{254\text{nm}}$ for the 9 selected CBDs are summarized in Table 3.2. The $\epsilon_{254\text{nm}}$ -values for 1,1,2,3-, and (Z)-1,1,3,4-TCBD, 1,1,2,4,4-PCBD, and HCBd are in the order of $10^2 \text{ m}^2 \text{ mol}^{-1}$, whereas they are in the order of $10^3 \text{ m}^2 \text{ mol}^{-1}$ for (E)-1,1,2,4-, (E)-1,1,3,4-, 1,1,4,4-, and (Z,Z)-1,2,3,4-TCBD and (Z)-1,1,2,3,4-PCBD (Table 3.2 and Figure S3.5). Note that since there were no separate standards available for the individual isomers, an average $\epsilon_{254\text{nm}}$ was determined for the isomer mixture of (E)-1,1,3,4- and (Z)-1,1,3,4-TCBDs.

Quantum yields (Φ). The determined $\Phi_{254\text{nm}}$ values based on Eq. 3.2 are presented in Table 3.2. They vary between 0.03 and 0.95, where the highest and the lowest $\Phi_{254\text{nm}}$ were determined for HCBd and 1,1,4,4-TCBD, respectively. Note that $\Phi_{254\text{nm}}$ for (E)-1,1,3,4- and (Z)-1,1,3,4-TCBD were determined using an average $\epsilon_{254\text{nm}}$ of their isomer mixture.

3.3.5. Abatement of CBDs and micropollutants and bromate formation in a natural groundwater during ozonation, direct UV photolysis, and the AOPs $\text{O}_3/\text{H}_2\text{O}_2$ and UV/ H_2O_2

Ozonation and $\text{O}_3/\text{H}_2\text{O}_2$. Figure 3.4 shows the abatement of 7 selected CBDs and 3 micropollutants as well as the bromate formation during ozonation and the AOP $\text{O}_3/\text{H}_2\text{O}_2$ in Hardwald groundwater. For 0.5 mg O_3/L , a >95% abatement (based on the calibration range) was observed for all TCBDs, whereas 47% and 16% abatement was obtained for (Z)-1,1,2,3,4-PCBD and HCBd, respectively. Note that the data for 0.25 mg O_3/L without H_2O_2 cannot be presented due to an analytical failure. A gradual increase in degradation was observed for (Z)-1,1,2,3,4-PCBD by increasing the ozone dose to 4.0 mg O_3/L achieving >95% abatement, while an abatement of up to 40% was observed for HCBd.

Among the 5 micropollutants investigated, carbamazepine and tramadol are not presented in Figure 3.4 because a complete abatement was always observed due to high k_{O_3} -values of $3 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ ⁴¹ for carbamazepine and $3.8 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ ⁴² for tramadol at pH 8.0, respectively. Benzotriazole shows overall similar degradation efficiencies as TCBDs, which can be explained by the similar k_{O_3} -values ($\sim 1.0 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$ at pH 8.0 for benzotriazole²⁴). The degradation of iopamidol is comparable to (Z)-1,1,2,3,4-PCBD, which is consistent with rather similar k_{O_3} and k_{OH} ($k_{\text{O}_3} = <0.8 \text{ M}^{-1}\text{s}^{-1}$ ⁴³ and $k_{\text{OH}} = 3.34 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$ ⁴⁴ for iopamidol and $k_{\text{O}_3} = 1.0 \text{ M}^{-1}\text{s}^{-1}$ and $k_{\text{OH}} = 2.2 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$ for (Z)-1,1,2,3,4-PCBD, this study). A higher abatement efficiency was observed for diatrizoate compared to HCBd. While two k_{OH} -values are reported for diatrizoate, namely, $0.5 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$ ⁴⁵ and $0.9 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$ ⁴⁴, $0.9 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$ was determined for HCBd in this study (Table 3.1).

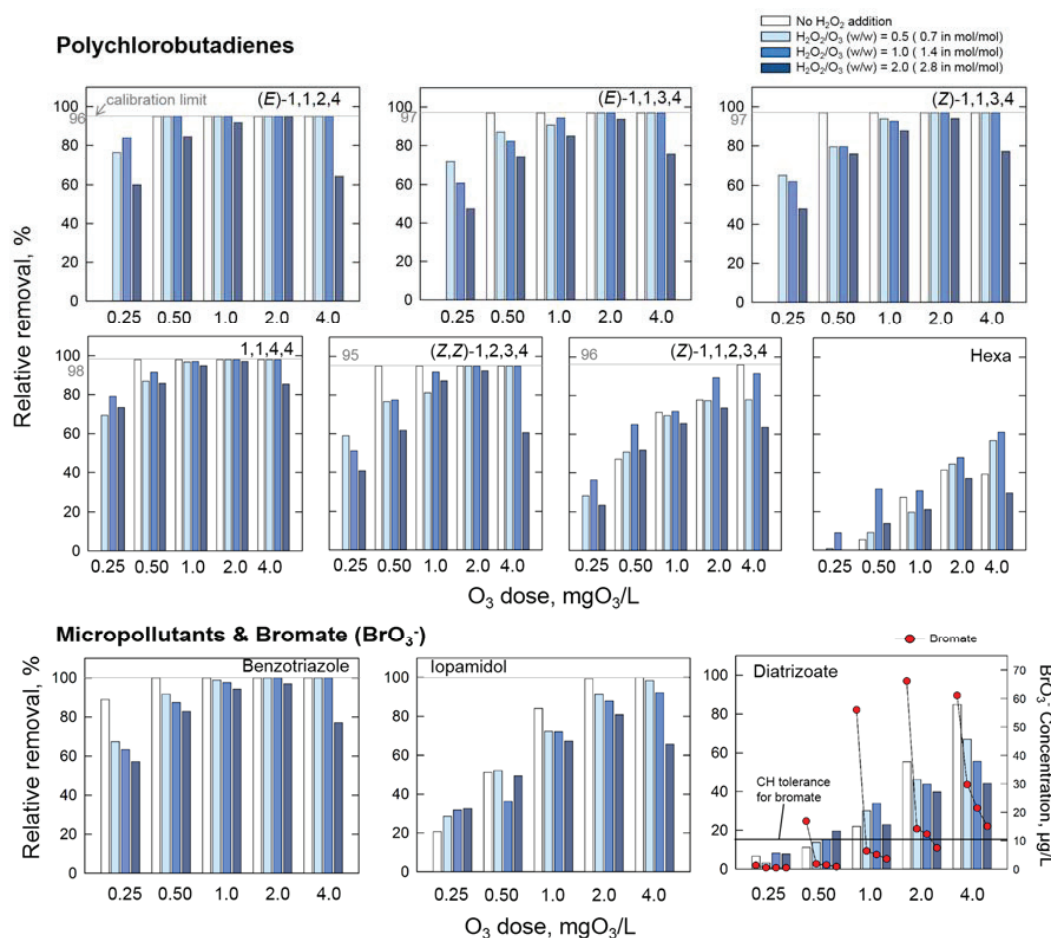


Figure 3.4. Abatement of CBDs and micropollutants and bromate formation during ozonation of Hardwald groundwater in the absence/presence of H_2O_2 . The data for CBDs at 0.25 mgO_3/L without H_2O_2 are not presented due to an analytical failure. $\text{pH} = 7.95 - 8.05$, temperature = 25°C , and bromide concentration = $48 \mu\text{g}/\text{L}$.

Overall H_2O_2 addition led to a lower abatement efficiencies of the investigated compounds (Figure 3.4). For ozone-reactive TCBDs and benzotriazole, it may be due to a decreased ozone exposure by a fast transformation of ozone to $\cdot\text{OH}$ by hydrogen peroxide (mainly by $\text{HO}_2\cdot$).²⁴ For ozone-resistant PCBD, HCBd, and iopamidol and diatrizoate at the highest H_2O_2 dose (i.e., $\text{H}_2\text{O}_2/\text{O}_3 = 2.0$ (w/w)) compared to the lower doses, it may be due to an increased $\cdot\text{OH}$ scavenging rate by added hydrogen peroxide (both H_2O_2 and $\text{HO}_2\cdot$) (e.g., about 8% increase of $\cdot\text{OH}$ scavenging rate (s^{-1}) was estimated for $\text{H}_2\text{O}_2/\text{O}_3 = 0.2$ (w/w) and $\text{O}_3 = 2.0 \text{ mg}/\text{L}$).

Bromate formation dramatically increased with increasing ozone doses from $1.3 \mu\text{g}/\text{L}$ for 0.25 mgO_3/L to $56 - 66 \mu\text{g}/\text{L}$ for the higher ozone doses (1.0 - 4.0 mgO_3/L) (Figure 3.4), which corresponds to 31-37 % molar conversion from bromide. Upon H_2O_2 addition, the bromate formation was significantly reduced to below $10 \mu\text{g}/\text{L}$, the drinking water standard in Switzerland,⁴⁶ EU,⁴⁷ and

USA,⁴⁸ up to 1.0 mgO₃/L. The bromate concentrations remained above 10 µg/L for the higher ozone doses except for 2.0 mgO₃/L with H₂O₂/O₃=2.0 (w/w). The reduced bromate formation is attributable to the decreased ozone exposure caused by H₂O₂ addition, as explained above for TCBDs and benzotriazole.

Direct UV photolysis and UV/H₂O₂. Direct UV photolysis for the 7 selected CBDs was compared with a simulated degradation based on the *k'*-values in Table 3.2 (Figure 3.5). Overall, the model could predict the % abatement up to 50 % relatively well for all CBDs (lines in Figure 3.5). For further degradation, relatively good predictions were observed for (*E*)-1,1,3,4- and 1,1,4,4-TCBD and HCBD. In contrast, poor predictions were obtained for (*E*)-1,1,2,4-, (*Z*)-1,1,3,4-, and (*Z,Z*)-1,2,3,4-TCBD and (*Z*)-1,1,2,3,4-PCBD, which is due to an appearance of a plateau without further degradation. Interestingly, the CBDs with a poor prediction coincided with photo-isomerizable compounds except for (*E*)-1,1,3,4-TCBD. The concentration even slightly increased for (*Z,Z*)-1,2,3,4-TCBD. Concurrently with the appearance of a plateau, the enhanced formation of photoisomers was observed in a groundwater compared to a buffered purified water. For example, the maximum relative response of the photoisomer-1 of (*Z,Z*)-1,2,3,4-TCBD was ~0.9 and ~1.6 in a buffered purified water (Figure 3.3a) and a groundwater (Figure 3.5), respectively. The interplay between CBDs and dissolved organic matter (DOM) present in a groundwater may be attributable to this phenomenon. Adsorption of CBDs to DOM could influence photoisomerization, inducing potentially a higher rotational barrier in the excited states. Moreover, the energy transfer between the species in the excited triplet states may play a role, especially when CBDs are adsorbed to DOMs. Further investigations to verify these hypotheses are necessary.

Although some enhancements (~10%) and inhibitions (~15%) were seemingly observed for (*E*)-1,1,3,4- and 1,1,4,4-TCBDs and (*Z*)-1,1,2,3,4-PCBD and HCBD (Figure 3.5), respectively, overall no significant effect on the degradation of the investigated CBDs seemed to manifest upon H₂O₂ addition. This may be mainly due to the efficient UV phototransformation of CBDs. For non-photoisomerizable CBDs such as 1,1,4,4-TCBD and HCBD, the matrix effect of the groundwater seems to play a minor role because their degradation could be relatively well predicted using parameters determined in purified water.

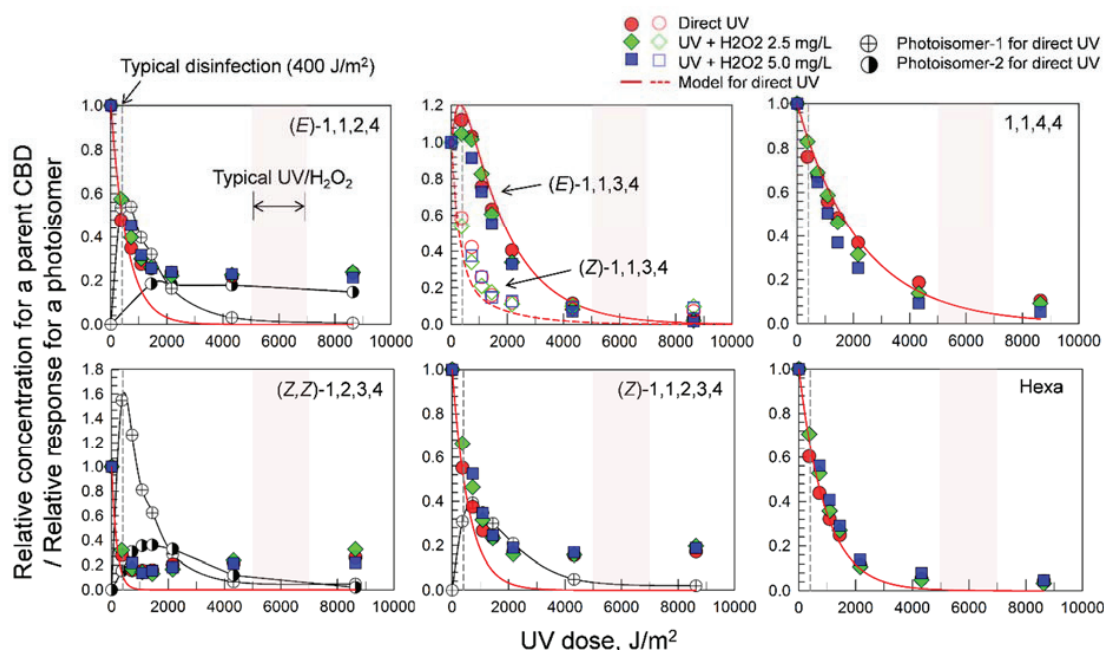


Figure 3.5. Abatement of CBDs during UV irradiation in absence/presence of H_2O_2 in Hardwald groundwater. A relative response is the ratio of the normalized peak area of the photoisomers by the internal standard to the normalized initial peak area of the parent CBD normalized by the internal standard. UV dose (J/m^2) was calculated by the equation: $\text{UV dose} = E_p^0 \times \text{irradiation time (seconds)} \times U$, where $U = 4.72 \times 10^5 \text{ J einstein}^{-1}$ at 254nm. $\text{pH} = 7.95 - 8.05$ and temperature = 25°C . Note the data for UV dose of about $360 \text{ J}/\text{m}^2$ for $\text{H}_2\text{O}_2 = 5.0 \text{ mg}/\text{L}$ are not presented due to an analytical failure.

3.3.6. Practical implications

During ozonation 50-100% of the TCBDs were abated at ozone doses of 0.25-1.0 mgO_3/L . The abatement efficiencies were smaller for PCBd and HCBd. Overall abatement efficiencies were consistent between CBDs and other selected micropollutants when evaluated based on the corresponding experimental k_{O_3} and k_{OH} . k_{O_3} predictions for TCBDs by quantum chemical models is an alternative to an empirical k_{O_3} determination. H_2O_2 addition is advantageous for minimizing the bromate formation during ozonation, however, it may lead to a decreased abatement of CBDs or micropollutants. Therefore, the ozone and H_2O_2 doses have to be carefully optimized to maximize pollutant abatement and to minimize bromate formation. For a typical UV dose for disinfection ($400 \text{ J}/\text{m}^2$), various degrees of phototransformation could be achieved for the investigated CBDs from 10 % for 1,1,4,4-TCBD to 90% for (Z,Z)-1,2,3,4-TCBD. However, the photoisomer formation from photoisomerizable CBDs and their biological effects must be taken into account for assessing the benefits of this treatment process. The beneficial effect of H_2O_2 addition on the UV-induced abatement of CBDs in the tested groundwater was observed to be minor (<10%).

Acknowledgements

This study was funded by the project “Regional Water Supply Baselland 21”. The authors thank Jakov Bolotin (Eawag), Elisabeth Salhi (Eawag), and Ursula Schönenberger (Eawag) for technical support and Frank Leresche (EPFL/Eawag) for fruitful discussions.

References

- (1) Burkhard, L. P.; Sheedy, B. R.; McCauley, D. J.; DeGraeve, G. M. Bioaccumulation factors for chlorinated benzenes, chlorinated butadienes and hexachloroethane. *Environ. Toxicol. Chem.* **1997**, *16*, 1677–1686.
- (2) Goldbach, R. W.; Van Genderen, H.; Leeuwangh, P. Hexachlorobutadiene residues in aquatic fauna from surface water fed by the river Rhine. *Sci. Total Environ.* **1976**, *6*, 31–40.
- (3) Laska, A. L.; Bartell, C. K.; Laseter, J. L. Distribution of hexachlorobenzene and hexachlorobutadiene in water, soil, and selected aquatic organisms along the lower Mississippi River, Louisiana. *Bull. Environ. Contam. Toxicol.* **1976**, *15*, 535–542.
- (4) Nikolaou, A. D.; Golfinopoulos, S. K.; Kostopoulou, M. N.; Kolokythas, G. A.; Lekkas, T. D. Determination of volatile organic compounds in surface waters and treated wastewater in Greece. *Water Res.* **2002**, *36*, 2883–2890.
- (5) G. Oliver, Barry, L.E. Kaiser, K. Chlorinated organics in nearshore waters and tributaries of the St. Clair River. *Water Poll. Res. J. Canada*, **1986**, *21*, 344–350.
- (6) Fattore, E.; Benfenati, E.; Fanelli, R. Analysis of Chlorinated 1,3-Butadienes by Solid-Phase Microextraction and Gas Chromatography-Mass Spectrometry. *J. Chromatogr. A* **1996**, *737*, 85–91.
- (7) Botta, D.; Dancelli, E.; Mantica, E. A Case History of Contamination by Polychloro-1,3-butadiene Congeners. *Environ. Sci. Technol.* **1996**, *30*, 453–462.
- (8) Brüsweiler, B. J. TTC-based risk assessment of tetrachlorobutadienes and pentachlorobutadienes—the in vitro genotoxic contaminants in ground and drinking water. *Regul. Toxicol. Pharmacol.* **2010**, *58*, 341–344.
- (9) Barco-Bonilla, N.; Plaza-Bolaños, P.; Fernández-Moreno, J. L.; Romero-González, R.; Frenich, A. G.; Vidal, J. L. M. Determination of 19 volatile organic compounds in wastewater effluents from different treatments by purge and trap followed by gas-chromatography coupled to mass spectrometry. *Anal. Bioanal. Chem.* **2011**, *400*, 3537–3546.
- (10) Agency for Toxic Substances and Disease Registry (ATSDR). *Toxicological Profile for Hexachlorobutadiene*; 1994.
- (11) *Hexachlorobutadiene in Drinking-water, Background document for development of WHO Guidelines for Drinking-water Quality, WHO/SDE/WSH/03.04/101*; 2004.
- (12) Stockholm Convention on Persistent Organic Pollutants.
- (13) *Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy, Off. J. Eur. Union, L 326, 2000.*
- (14) *Decision No 2455/2001/EC of the European Parliament and of the Council of 20 November 2001 establishing the list of priority substances in the field of water policy and amending Directive 2000/60/EC, Off. J. Eur. Union, L331, 2001.*
- (15) *Directive 2008/105/EC of the European Parliament and of the Council of 16 December 2008 on environmental quality standards in the field of water policy, amending and subsequently repealing Council Directives 82/176/EEC, 83/513/EEC, 84/156/EEC, 84/491/EEC, .*
- (16) Bosma, T. N.; Cottaar, F. H.; Posthumus, M. A.; Teunis, C. J.; van Veldhuizen, A.; Schraa, G.; Zehnder, A. J. Comparison of Reductive Dechlorination of Hexachloro-1,3-butadiene in Rhine Sediment and Model Systems with Hydroxocobalamin. *Environ. Sci. Technol.* **1994**, *28*, 1124–1128.
- (17) Booker, R. S.; Pavlostathis, S. G. Microbial reductive dechlorination of hexachloro-1,3-butadiene in a methanogenic enrichment culture. *Water Res.* **2000**, *34*, 4437–4445.
- (18) Atlanta, GA: U.S. Department of Health and Human Services, P. H. S. *Toxicological profile for Vinyl Chloride*.; 2006.
- (19) Vogel, T. M.; McCARTY, P. L. Biotransformation of tetrachloroethylene to trichloroethylene, dichloroethylene, vinyl chloride, and carbon dioxide under methanogenic conditions. *Appl. Environ. Microbiol.* **1985**, *49*, 1080–1083.
- (20) Banerjee, S.; Yalkowsky, S. H.; Valvani, S. C.; Valvani, C.; Valvani, S. C. Water Solubility and Octanol / Water Partition Coefficients of Organics . Limitations of the Solubility-Partition

- Coefficient Correlation. *Environ. Sci. Technol.* **1980**, *14*, 1227–1229.
- (21) Dewulf, J.; Van Langenhove, H.; Everaert, P. Determination of Henry's law coefficients by combination of the equilibrium partitioning in closed systems and solid-phase microextraction techniques. *J. Chromatogr. A* **1999**, *830*, 353–363.
 - (22) Glaze, W. H.; Kang, J.-W.; Chapin, D. H. The Chemistry of Water Treatment Processes Involving Ozone, Hydrogen Peroxide and Ultraviolet Radiation. *Ozone Sci. Eng.* **1987**, *9*, 335–352.
 - (23) von Gunten, U. Ozonation of drinking water: part I. Oxidation kinetics and product formation. *Water Res.* **2003**, *37*, 1443–1467.
 - (24) von Sonntag, C.; von Gunten, U. *Chemistry of ozone in water and wastewater treatment: From basic principles to applications*; IWA publishing, 2012.
 - (25) Bolton, J. R.; Stefan, M. I. Fundamental photochemical approach to the concepts of fluence (UV dose) and electrical energy efficiency in photochemical degradation reactions. *Res. Chem. Intermed.* **2002**, *28*, 857–870.
 - (26) Canonica, S.; Meunier, L.; von Gunten, U. Phototransformation of selected pharmaceuticals during UV treatment of drinking water. *Water Res.* **2008**, *42*, 121–128.
 - (27) Lee, M.; Zimmermann-Steffens, S. G.; Arey, J. S.; Fenner, K.; von Gunten, U. Development of Prediction Models for the Reactivity of Organic Compounds with Ozone in Aqueous Solution by Quantum Chemical Calculations: The Role of Delocalized and Localized Molecular Orbitals. *Environ. Sci. Technol.* **2015**, *49*, 9925–9935.
 - (28) Wenk, J.; von Gunten, U.; Canonica, S. Effect of Dissolved Organic Matter on the Transformation of Contaminants Induced by Excited Triplet States and the Hydroxyl Radical. *Environ. Sci. Technol.* **2011**, *45*, 1334–1340.
 - (29) Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H. Gaussian 09, revision C. 01; Gaussian, Inc. Wallingford, CT **2009**.
 - (30) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999–3093.
 - (31) E. D. Glendening, A. E. Reed, J. E. Carpenter, F. W. NBO Version 3.1.
 - (32) Ianni, J. C. Kintecus, Windows Version 5.00, 2014.
 - (33) Hoigné, J.; Bader, H. Rate constants of reactions of ozone with organic and inorganic compounds in water—I. Non-dissociating organic compounds. *Water Res.* **1983**, *17*, 173–183.
 - (34) Dowideit, P.; von Sonntag, C. Reaction of Ozone with Ethene and Its Methyl- and Chlorine-Substituted Derivatives in Aqueous Solution. *Environ. Sci. Technol.* **1998**, *32*, 1112–1119.
 - (35) Craig, N. C.; Brandon, D. W.; Stone, S. C.; Lafferty, W. J. Partial structure for trans-1,2-difluoroethylene from high-resolution infrared spectroscopy. *J. Phys. Chem.* **1992**, *96*, 1598–1605.
 - (36) Chaudhuri, R. K.; Hammond, J. R.; Freed, K. F.; Chattopadhyay, S.; Mahapatra, U. S. Reappraisal of cis effect in 1,2-dihaloethenes: an improved virtual orbital multireference approach. *J. Chem. Phys.* **2008**, *129*, 064101.
 - (37) Pitzer, K. S.; Hollenberg, J. L. cis- and trans-Dichloroethylenes. The Infrared Spectra from 130-400 Cm. ⁻¹ and the Thermodynamic Properties I. *J. Am. Chem. Soc.* **1954**, *76*, 1493–1496.
 - (38) Getoff, N. Radiation- and photoinduced degradation of pollutants in water. A comparative study. *Int. J. Radiat. Appl. Instrumentation. Part C. Radiat. Phys. Chem.* **1991**, *37*, 673–680.
 - (39) Boue, S.; Srinivasan, R. Differences in reactivity between excited states of cis- and trans-1,3-pentadiene. *J. Am. Chem. Soc.* **1970**, *92*, 3226–3227.
 - (40) Srinivasan, R. Kinetics of the Photochemical Dimerization of Olefins to Cyclobutane Derivatives. I. Intramolecular Addition. *J. Am. Chem. Soc.* **1962**, *84*, 4141–4145.
 - (41) Huber, M. M. M.; Canonica, S.; Park, G.-Y.; von Gunten, U. Oxidation of Pharmaceuticals during Ozonation and Advanced Oxidation Processes. *Environ. Sci. Technol.* **2003**, *37*, 1016–1024.
 - (42) Zimmermann, S. G.; Schmukat, A.; Schulz, M.; Benner, J.; von Gunten, U.; Ternes, T. A. Kinetic and Mechanistic Investigations of the Oxidation of Tramadol by Ferrate and Ozone.

Chapter 3

- Environ. Sci. Technol.* **2012**, *46*, 876–884.
- (43) Huber, M. M.; Göbel, A.; Joss, A.; Hermann, N.; Löffler, D.; McArdell, C. S.; Ried, A.; Siegrist, H.; Ternes, T. A.; von Gunten, U. Oxidation of Pharmaceuticals during Ozonation of Municipal Wastewater Effluents: A Pilot Study. *Environ. Sci. Technol.* **2005**, *39*, 4290–4299.
- (44) Jeong, J.; Jung, J.; Cooper, W. J.; Song, W. Degradation mechanisms and kinetic studies for the treatment of X-ray contrast media compounds by advanced oxidation/reduction processes. *Water Res.* **2010**, *44*, 4391–4398.
- (45) Real, FJ, Benitez, FJ, Acero, JL, Sagasti, JP, C. F. Kinetics of the chemical oxidation of the pharmaceuticals primidone, ketoprofen, and diatrizoate in ultrapure and natural waters. *Ind. Eng. Chem. Res.* **2009**, *48*, 3380–3388.
- (46) Swiss Federal Department of the Interior (2000). Ordinance on Contaminants in Food., Bern, Switzerland.
- (47) EU. Official Journal of the European Community L 330: Directive 98/83/EG, 1998.
- (48) USEPA. National Primary Drinking Water Regulations (NPDWRs). Federal Register, 1989, *54*, 27485–27541.

Supporting information for chapter 3.

Abatement of polychoro-1,3-butadienes in aqueous solution by ozone, UV-photolysis, and advanced oxidation processes (O₃/H₂O₂ and UV/H₂O₂)

including 5 texts, 4 tables, 1 scheme, and 8 figures.

Text S3.1. Standards and reagents

6 Tetrachlorobutadiene (TCBD) isomers (1,1,2,3-, (*E*)-1,1,2,4-, (*Z*)-1,1,3,4-, (*E*)-1,1,3,4-, 1,1,4,4-, and 1,2,3,4-) and 2 pentachlorobutadiene (PCBD) isomers ((*Z*)-1,1,2,3,4- and 1,1,2,4,4-) ($\geq 96\%$) were obtained from Neochema GmbH & Co. KG (Bodenheim, Germany), HCBd (98%), 4-chlorobenzoic acid (*p*CBA) (99%), *tert*-butanol (*t*-BuOH) ($\geq 99.7\%$), hydrogen peroxide ($\geq 35\%$), atrazine (98.8%), benzotriazole (99%), tramadol ($\geq 99\%$), carbamazepine (99%) and sodium sulfite ($\geq 98\%$) from Sigma Aldrich, and EPA 524 Internal Standard Mix containing fluorobenzene and 1,2-dichlorobenzene-*d*₄ from Supelco, respectively. Diatrizoic acid (92%) and iopamidol (99.6%) were purchased from Dr. Ehrenstorfer (Augsburg, Germany) and U.S. Pharmacopeia (MD, USA), respectively. Note that (*E*)- and (*Z*)-1,1,3,4-TCBD isomers were provided as a mixture, and their compositions were determined to be 43% and 57%, respectively, by ¹H and ¹³C nuclear magnetic resonance (NMR) analysis (Table S3.1 and Figure S3.1). The fractions of (*E*)-1,1,3,4- and (*Z*)-1,1,3,4-TCBD were estimated by comparing the areas of two proton signals (22.85 and 29.91 for the proton attached to the C2 of (*E*)-1,1,3,4-TCBD and for the proton attached to the C4 of (*Z*)-1,1,3,4-TCBD, respectively). The *Z* configuration of 1,1,2,3,4-PCBD was confirmed by a ¹H NMR analysis (Table S3.2). All aforementioned chemicals were used without further purification. Solutions were prepared in purified water produced by the Milli-Q (Millipore) nanopure system for all experiments except for the experiments conducted in a natural groundwater.

Text S3.2. Analytical methods

The 9 CBDs were quantified by a purge and trap/gas chromatography/mass spectrometry (PT/GC/MS) system. The GC (Thermo Scientific Trace GC Ultra) equipped with Rts-VMS[®] capillary column (60m length, 0.32 mm i.d., 1.8 μ m film thickness; Restek) was connected to a MS (Thermo Scientific DSQ II). The CBDs were extracted from aqueous samples by the PT (Tekmar Dohrmann 3100 Sample Concentrator), which are then introduced with a splitless injection into the GC-MS system using helium (99.999%) as a carrier gas. Both a total ion current (TIC) chromatogram and a selected ion monitoring (SIM) chromatogram were employed depending on the purpose of the measurements. Note that the highest oven temperature and the MS source temperature were set to 220°C and 250°C, respectively, and no thermal isomerization was observed when CBDs were individually analyzed. *p*CBA (competition kinetics) and atrazine (UV dosimetry) were quantified using a high-performance liquid chromatography (HPLC) (Dionex Ultimate 3000) at 240 nm with an eluent of methanol and 20 mM phosphate buffer (pH 2) (60:40, v:v) and at 240 nm with an eluent of water and acetonitrile (50:50, v:v), respectively. A UV spectrophotometer (Cary 100, Varian) was used for various purposes: the determination of the ozone stock solution (molar extinction coefficient (ϵ) = 3200 M⁻¹ cm⁻¹ at 258 nm)¹, the residual ozone concentration using the indigo method (ϵ = 20000 M⁻¹ cm⁻¹ at 600 nm)², the quantification of H₂O₂ stock solution (ϵ = 40 M⁻¹ cm⁻¹ at 240 nm)³, and the determination of ϵ for 9 CBDs at 254 nm. The 5 micropollutants (benzotriazole, carbamazepine, diatrizoate, iopamidol, and tramadol) were quantified by a fully automated online solid phase extraction/liquid chromatography/mass spectrometry (online SPE/LC/MS/MS). In a first step, 20 mL of sample was loaded with a solution of acidic nanopure water (0.1 % formic acid) on a 2-layer extraction cartridge made of 8 mg

OASIS HLB (15 μm , Waters) and 14 mg Isolute ENV+ (70 μm , Biotage, Uppsala, Sweden). Hence, the separation was performed at ambient temperature using an Atlantis T3 column (3.0 x 150 mm, 3 μm particle size) with a HPLC gradient containing acidic nanopure water (0.1% formic acid) and methanol. The detection was operated with a Thermo TSQ-Vantage equipped with an electrospray ionization probe (ESI). The method detection limits (MDL) were 1 ng L^{-1} for carbamazepine, 2 ng L^{-1} for tramadol and 5 ng L^{-1} for diatrizoate, iopamidol and benzotriazole. Bromate concentration was measured by capillary ion chromatography (Thermo Dionex ICS-4000) coupled with tandem mass spectrometry (Thermo TSQ-Vantage). The water samples were diluted 10 times prior to analysis and the MDL was 0.5 $\mu\text{g L}^{-1}$.

Text S3.3. Determination of second-order rate constants for the reactions of CBDs with ozone (k_{O_3}) and hydroxyl radicals ($k_{\cdot\text{OH}}$)

k_{O_3} for CBDs in aqueous solution were determined by simultaneously monitoring two terms in Eq. S3.1 during ozonation, i.e., the relative residual concentrations of CBDs at time, t , expressed as the natural logarithm and the corresponding ozone exposures, respectively.

$$-\ln \frac{[\text{CBD}]_t}{[\text{CBD}]_0} = k_{\text{O}_3} \int [\text{O}_3] dt \quad (\text{S3.1})$$

An ozone stock solution was prepared in purified water cooled in an ice bath by sparging ozone-containing oxygen produced from an ozone generator using pure oxygen gas (>99.999%). The ozone concentration range in the stock solution was between 1 and 1.3 mM. For the determination of k_{O_3} , an appropriate aliquot of the ozone stock solution, which renders always an excess of ozone to CBD, was transferred into the solution containing the individual CBDs (1.0 - 3.5 μM) at pH 2 in the presence of 5 mM *t*-BuOH as a scavenger for hydroxyl radicals. Samples from the ozonated CBD solution withdrawn at appropriate time intervals were quenched with sulfite. Thereafter, the CBDs were quantified with GC/MS and residual ozone was measured by the indigo method², respectively. The k_{O_3} -values were determined separately for individual CBDs. The experimental data used to determine k_{O_3} in Table 3.1 are presented in Figure S3.2.

$k_{\cdot\text{OH}}$ for CBDs were determined by UV-photolysis of H_2O_2 by competition kinetics with *para*-chlorobenzoic acid (*p*CBA) ($k_{\cdot\text{OH},p\text{CBA}} = 5 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$)⁴. A medium pressure (MP) mercury UV lamp housed in a Pyrex[®] cooling jacket with tap water as cooling water was used. The filtered MP-UV lamp, the main emission line responsible for H_2O_2 photolysis being centered at 313 nm, was placed in the center of a DEMA 125 merry-go-round photoreactor (MGRR) (Hans Mangels, Bornheim-Roisdorf, Germany) filled with purified water and constantly kept at 25°C. Details about the photoreactor are found elsewhere.^{5,6} The Pyrex[®] glass cooling jacket, which ensured that UV light with wavelength <290 nm was cut off, was chosen to minimize direct photolysis of the CBDs (Figure S3.4) while sustaining an acceptable rate of $\cdot\text{OH}$ production. A composite solution (pH 7, 5 mM phosphate buffer) of 20 mL containing CBD (1.0-3.5 μM), *p*CBA (0.2-2.0 μM) as a competitor, and H_2O_2 (10

Supporting information for chapter 3

mM) as $\cdot\text{OH}$ source in a quartz tube (ID: 15 mm and OD: 18 mm) was exposed to UV at 313 nm in the MGRR. Samples were collected at appropriate irradiation times for the quantification of the respective compounds. Despite the use of 313 nm to achieve exclusive degradation of CBDs by $\cdot\text{OH}$, a loss of the target compound still occurred to some extent by direct photolysis at 313nm and volatilization. Therefore, the photo-degradation of CBDs in the absence of H_2O_2 was also monitored to correct the loss. The net degradation of CBDs by $\cdot\text{OH}$ after the correction was used to determine $k_{\cdot\text{OH}}$. The correction was in the range of 0~13%. Consequently, $k_{\cdot\text{OH,CBD}}$ could be determined based on Eq. S3.4 which is obtained by combining Eqs. S3.2 and S3.3. The slope of a linear regression line obtained by $\ln([\text{CBD}]_t/[\text{CBD}]_0)/\ln([p\text{CBA}]_t/[p\text{CBA}]_0)$ is multiplied by $k_{\cdot\text{OH},p\text{CBA}}$ ($5 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$)⁴ to give rise to the corresponding $k_{\cdot\text{OH,CBD}}$.

$$-\ln \frac{[\text{CBD}]_t}{[\text{CBD}]_0} = k_{\cdot\text{OH,CBD}} \int [\cdot\text{OH}] dt \quad (\text{S3.2})$$

$$-\ln \frac{[p\text{CBA}]_t}{[p\text{CBA}]_0} = k_{\cdot\text{OH},p\text{CBA}} \int [\cdot\text{OH}] dt \quad (\text{S3.3})$$

$$k_{\cdot\text{OH,CBD}} = k_{\cdot\text{OH},p\text{CBA}} \ln \left(\frac{[\text{CBD}]_t}{[\text{CBD}]_0} \right) / \ln \left(\frac{[p\text{CBA}]_t}{[p\text{CBA}]_0} \right) \quad (\text{S3.4})$$

The \ln term is a relative concentration of respective compounds at the time (t) with respect to its initial concentration ($t=0$), $k_{\cdot\text{OH}}$ is a second-order rate constant for the reaction of $\cdot\text{OH}$ for the respective compounds in units of $\text{M}^{-1}\text{s}^{-1}$, and $\int [\cdot\text{OH}] dt$ is the exposure to $\cdot\text{OH}$ until t in units of $\text{M}\cdot\text{s}$. The experimental data used to determine $k_{\cdot\text{OH,CBD}}$ in Table 3.1 are presented in Figure S3.3.

Text S3.4. Prediction of second-order rate constants for the reactions of CBDs with ozone (k_{O_3}) using a previously developed quantum chemical prediction model.

The k_{O_3} prediction for CBDs by a quantum chemical model was conducted and is of special interest for the following two reasons:

(i) Model applicability to non-planar dienes: (near)-Planar dienes such as sorbic acid and microcystin-LR were successfully included in a previously developed quantum chemical model.⁷ Unlike those, CBDs manifest the high degree of a non-planarity driven by the steric repulsion between chlorine substituents (e.g., (*E,E*)-1,2,3,4-TCBD and HCBD were found to be in the *gauche* form with torsion angles of 75° and 83.6° , respectively.⁸) (Scheme S3.1). Therefore, the applicability of the model to non-planar dienes such as CBDs has to be tested.

(ii) Consideration of co-present CBD conformers: 1,3-butadiene, by virtue of facile rotation about a central C–C single bond, can exist in planar forms such as *s-trans* and *s-cis* and/or in a non-planar (i.e., *gauche*) form (Scheme S3.1). An unsubstituted 1,3-butadiene is known to predominantly exist in the *s-trans* form^{9–11} with the *gauche* form co-present^{12,13} in a fraction of approximately ~2%¹⁴ in the gas phase at room temperature.

In contrast to an experimental k_{O_3} , which yields an overall value of aggregate reactions of all present conformers with ozone, the quantum chemical model allows to predict k_{O_3} for a specific species under consideration. As conformers may have differing populations and reactivities with ozone, their role to a predicted k_{O_3} is of interest when CBD conformers are present. Note that the previously developed prediction model was using only one local minimum conformer without considering multiple conformers. Thus, the prediction performance when considering multiple conformers is unknown. As such, there is an interest to explore the sensitivity of the model on the treatment of multiple CBD conformers. With these two facets in mind, a k_{O_3} prediction for CBDs was performed.

A relaxed potential energy curve (PEC) scan with a dihedral angle comprising the atoms C₁-C₂-C₃-C₄ (φ) (Scheme S3.1) as a variable from the *s-trans* to the *s-cis* configuration was performed for nine CBDs to locate stationary points on the PEC for HF and B3LYP with the 6-311++G** basis set. Note that the *s-trans* configuration is adopted to be 0° as a convention throughout this study. In Figure 3.2 in the main manuscript, the profile of the relative total electronic energy of the (*E*)-1,1,2,4-TCBD rotamers is presented compared to the conformer with the lowest energy and the corresponding $E_{NBO, C-C(\pi)}$ of the two C=C bonds of (*E*)-1,1,2,4-TCBD computed at the HF/6-311++G**. Two minima were identified in the vicinity of 0° and 120° with the former being the lowest energy minimum, i.e., the global minimum. As the geometries obtained from the relaxed scan are approximate, two *s-trans* and *gauche* conformers were further optimized and confirmed to be the minimum structure by frequency analysis at the same level of theory. The dihedral angles of the two optimized minimum structures were 0° and 116°, respectively. With $\Delta G=7.63$ kJ mol⁻¹ calculated between the two minima, the calculation of the Boltzmann distribution at room temperature (298K) yielded the population (p) of two *s-trans* and *gauche* conformers to be 92% and 8%, respectively. Note that a frequency analysis for the optimized geometry was conducted to derive ΔG . The minimum structures for the 9 CBDs with their dihedral angles and populations obtained at HF/6-311++G** and B3LYP/6-311++G** level are summarized in Table S3.3. While the B3LYP method identified 6 CBDs with two minima among the 9 CBDs, the HF method could only identify 3 CBDs with two minima. For such CBDs, the population of the local minimum structures ranged from 1 – 43% over both methods. A geometry obtained by B3LYP is commonly known to be more accurate than HF,¹⁵ minimum structures obtained by B3LYP are considered more accurate. However, the HF method has not been excluded from further investigations because it was previously shown that the HF method does not necessarily yield a worse prediction performance, which is supposedly due to the fortuitous cancellation of errors in a relative correlation.⁷

For a CBD with two C-C double bonds, an overall species-specific $k_{O_3,i}$ can be derived according to Eq. S3.5.

$$k_{O_3,i} = k_{O_3, i(C1=C2)} + k_{O_3, i(C3=C4)} \quad (S3.5)$$

$k_{O_3, i(C1=C2)}$ and $k_{O_3, i(C3=C4)}$ are the site-specific second-order rate constant for the reaction of ozone with $C_1=C_2$ and $C_3=C_4$ of the i -th CBD conformer, respectively. $k_{O_3, i(C1=C2)}$ and $k_{O_3, i(C3=C4)}$ are derived from using the respective $E_{NBO, C-C(\pi)}$ of the two C=C bonds of the optimized CBDs for the linear regression equation from the previous study.⁷ The predicted $k_{O_3,i}$ for the 9 CBDs are presented in Table S3.4.

Taking the presence of multiple conformers into account for the k_{O_3} assessment, an overall second order rate constant for the reaction of ozone with all minimum conformers, denoted as $k_{O_3,pred.}$, is represented by Eq. S3.6.

$$k_{O_3,pred.} = \sum(p_i \times k_{O_3,i}) \quad \text{over all conformers } i \quad (S3.6)$$

p_i is the relative population of a i -th conformer derived from a Boltzmann distribution (Table S3.3) and $k_{O_3,i}$ is the species-specific second-order rate constant for the reaction of ozone with an i -th conformer from Eq. S3.5, respectively. The predicted $k_{O_3,pred.}$ for the 9 CBDs for both calculation methods are presented in Table 3.1.

Text S3.5. Estimation of time-based first-order rate constants (s^{-1}) for phototransformation and photoisomerization of (*E*)-1,1,3,4- and (*Z*)-1,1,3,4-TCBD during UV irradiation at 254 nm by kinetic simulation

A pseudo-first order rate constant for the phototransformation of a CBD, denoted as k , was obtained during direct UV photolysis, which was then used to determine Φ and k' in the main manuscript. However, for CBDs that undergo a photoisomerization, the derived k -value can contain errors because of a reversibility of the photoisomerization (Eqs. S3.7 and S3.8).

$$\frac{d[Z]}{dt} = -k_{Z \rightarrow P}[Z] - k_{Z \rightarrow E}[Z] + k_{E \rightarrow Z}[E] \quad (S3.7)$$

$$\frac{d[E]}{dt} = -k_{E \rightarrow P}[E] - k_{E \rightarrow Z}[E] + k_{Z \rightarrow E}[Z] \quad (S3.8)$$

$[Z]$ and $[E]$ are the concentrations (mol L^{-1}) of a CBD in the *Z* configuration or the *E* configuration, respectively. $k_{Z \rightarrow P}$, $k_{Z \rightarrow E}$, $k_{E \rightarrow P}$, and $k_{E \rightarrow Z}$ are the first-order rate constants (s^{-1}) for the corresponding processes of CBDs during UV irradiation at 254 nm. The subscripts before and after the arrow (\rightarrow) indicate the initial configuration (*E* or *Z*) of the CBD before photoisomerization and the final configuration of the CBD after the photoreaction, respectively. Unknown products are collectively denoted as *P* and the photoisomer is denoted as *E* or *Z*, respectively. Note that Eqs. S3.7 and S3.8 above are shown as examples for the CBDs which yield one photoisomer, such as (*E*)-1,1,3,4-TCBD, (*Z*)-1,1,3,4-TCBD, and (*Z*)-1,1,2,3,4-PCBD.

Based on the measured concentrations of two 1,1,3,4-TCBD isomers as a function of the UV irradiation time, the differential equations above (Eqs. S3.7 and S3.8) could be numerically solved by a kinetics simulation software (Kintecus 5.00¹⁶, fitting/optimization algorithm: Nelder & Mead,¹⁷ compar-

ison operator: standard least squares) to derive individual k -values (Figure S3.8). The average k -values from triplicates with 95% confidence interval were 0.061 ± 0.017 , 0.078 ± 0.035 , 0.009 ± 0.004 , and 0.038 ± 0.013 (s^{-1}) for $k_{Z\rightarrow P}$, $k_{Z\rightarrow E}$, $k_{E\rightarrow P}$, and $k_{E\rightarrow Z}$, respectively. Based on the k -value (0.0032 ± 0.00003 , s^{-1}) for atrazine determined from the atrazine actinometry, the corresponding k' -values for photoisomerization and other phototransformations are calculated according to Eq. 3.1 and presented in Table 3.2 in the main manuscript.

Supporting information for chapter 3

Table S3.1. ^1H and ^{13}C chemical shifts (δ) and coupling constants (J) of (*E*)-1,1,3,4-Tetrachlorobutadiene (TCBD) and (*Z*)-1,1,3,4-TCBD measured by ^1H NMR and ^{13}C NMR for a mixture of (*E*)-1,1,3,4-TCBD and (*Z*)-1,1,3,4-TCBD prepared in deuterated tetrahydrofuran (THF-*d*₈) at a concentration of 20 mg/mL.

C position ^a	This study				Reference study ¹⁸		$\Delta\delta$ (^{13}C), ppm
	δ (^1H), ppm	J (^1H , ^1H), Hz	δ (^{13}C), ppm	J (^{13}C , ^1H), Hz	δ (^{13}C), ppm	J (^{13}C , ^1H), Hz	
<i>(E)</i> -1,1,3,4-TCBD							
1			126.6	1.6, 0.5	126.0		0.6
2	6.78	1.40	123.3	170.1, 6.9 ^b	122.1	169.0, 6.9 ^b	1.2
3			129.3	1.4, 1.1	128.9		0.4
4	6.69	1.40	121.0	204.0, 4.2 ^c	120.2	201.9, 4.4 ^c	0.8
<i>(Z)</i> -1,1,3,4-TCBD							
1			126.2	1.3, 1.3	125.7		0.5
2	6.70	1.16	125.6	169.4, 3.4 ^b	124.7	168.7, 3.5 ^b	0.9
3			128.9	13.8 ^d	128.2	13.9 ^d	0.7
4	6.99	1.16	124.1	200.0, 5.9 ^c	123.2	199.1, 5.7 ^c	0.9

^aThe position of a terminal carbon with two chlorine substituents is defined as 1, ^b J (^{13}C -2, ^1H -4), ^c J (^{13}C -4, ^1H -2), ^d J (^{13}C -4, ^1H -4).

Table S3.2. Confirmation of the *Z* configuration for 1,1,2,3,4-PCBD by ^1H NMR and ^{13}C NMR for a sample solution of (*Z*)-1,1,2,3,4-PCBD prepared in deuterated acetone at a concentration of 15 mg/mL.

C position ^a	This study		Reference study ¹⁸		$\Delta\delta$ (^{13}C), ppm	ΔJ (^{13}C , ^1H), Hz
	δ (^{13}C), ppm	J (^{13}C , ^1H), Hz	δ (^{13}C), ppm	J (^{13}C , ^1H), Hz		
1	126.1	0.9	125.5	-	0.6	0.9
2	126.0	3.0	125.6	2.9	0.4	0.1
3	128.1	13.4	127.9	13.2	0.2	0.2
4	127.2	202.2	126.2	200.7	1.0	1.5

^aThe position of a terminal carbon with two chlorine substituents is defined as 1.

Table S3.3. Summary of minimum structures of CBDs

Compound	HF/6-311++G**			B3LYP/6-311++G**		
	φ^a	ΔG^b	Pop. ^c	φ^a	ΔG^b	Pop. ^c
1,1,2,3-TCBD	103.3	0	1.00	105.9	0	1.00
(<i>E</i>)-1,1,2,4-TCBD	0	0	0.92	0	0	0.96
	116.0	7.6	0.08	126.1	12.3	0.04
(<i>E</i>)-1,1,3,4-TCBD	105.6	0	1.00	38.5	0	0.58
				114.7	0.8	0.42
(<i>Z</i>)-1,1,3,4-TCBD	118.0	0	1.00	0	3.1	0.11
				128.5	0	0.89
1,1,4,4-TCBD	0	0	1.00	0	0	1.00
	106.2	16.6	0.00	119.0	24.0	0.00
(<i>Z,Z</i>)-1,2,3,4-TCBD	0	6.6	0.03	0	2.7	0.15
	117.8	0	0.97	122.0	0	0.85
(<i>Z</i>)-1,1,2,3,4-PCBD	101.7	0	1.00	104.5	0	1.00
1,1,2,4,4-PCBD ^d	100	0	1.00	58.1	0	0.57
				109.3	0.7	0.43
HCBD	92.2	0	1.00	94.4	0	1.00

^aDihedral angle (°), ^bRelative Gibbs free energy (kJ/mol) of the conformers to the lowest energy conformer

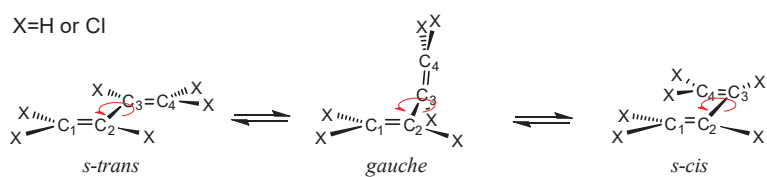
^cPopulation of the conformers calculated by the Boltzmann distribution equation using ΔG .

Table S3.4. Predicted second-order rate constant for the reactions of ozone with a global minimum CBD ($k_{O_3,global}$) and a local minimum CBD ($k_{O_3,local}$) for various CBDs

Compound	$k_{O_3,global}^a/k_{O_3,local}^b$	
	HF/6-311++G**	B3LYP/6-311++G**
1,1,2,3-TCBD	1.4×10^3	1.6×10^3
(<i>E</i>)-1,1,2,4-TCBD	$2.3 \times 10^3 / 1.4 \times 10^3$	$2.5 \times 10^3 / 1.6 \times 10^3$
(<i>E</i>)-1,1,3,4-TCBD	3.4×10^2	$4.5 \times 10^2 / 5.4 \times 10^2$
(<i>Z</i>)-1,1,3,4-TCBD	3.6×10^2	$5.2 \times 10^2 / 8.7 \times 10^2$
1,1,4,4-TCBD	$6.9 \times 10^2 / 2.5 \times 10^2$	$1.0 \times 10^3 / 2.6 \times 10^2$
(<i>Z,Z</i>)-1,2,3,4-TCBD	$5.2 \times 10^2 / 4.2 \times 10^2$	$5.9 \times 10^2 / 7.0 \times 10^2$
(<i>Z</i>)-1,1,2,3,4-PCBD	1.5×10^2	2.4×10^2
1,1,2,4,4-PCBD	1.5×10^2	$1.6 \times 10^2 / 2.4 \times 10^2$
HCBD	25	47

^aPredicted k_{O_3} -values ($M^{-1}s^{-1}$) calculated by only considering the global minimum conformer. ^bPredicted k_{O_3} -values ($M^{-1}s^{-1}$) calculated by only considering the local minimum conformer. Only one k_{O_3} for CBDs with only one conformer (Table S3.3) is presented (Table 3.1).

Supporting information for chapter 3



Scheme S3.1. Conformations of (chloro)-1,3-butadiene

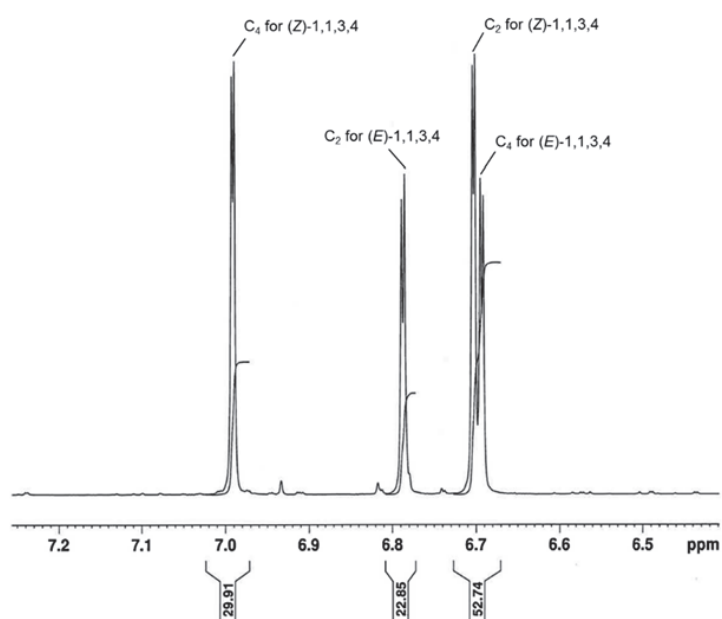


Figure S3.1. ^1H NMR of the mixture of (*E*)-1,1,3,4- and (*Z*)-1,1,3,4-TCBD prepared in $\text{THF-}d_8$ at a concentration of 20 mg/mL. The areas of two proton signals (29.91 for (*Z*)-1,1,3,4-TCBD and 22.85 for (*E*)-1,1,3,4-TCBD) were used to estimate an isomer composition of 57 and 43 % for (*Z*)-1,1,3,4-TCBD and for (*E*)-1,1,3,4-TCBD, respectively.

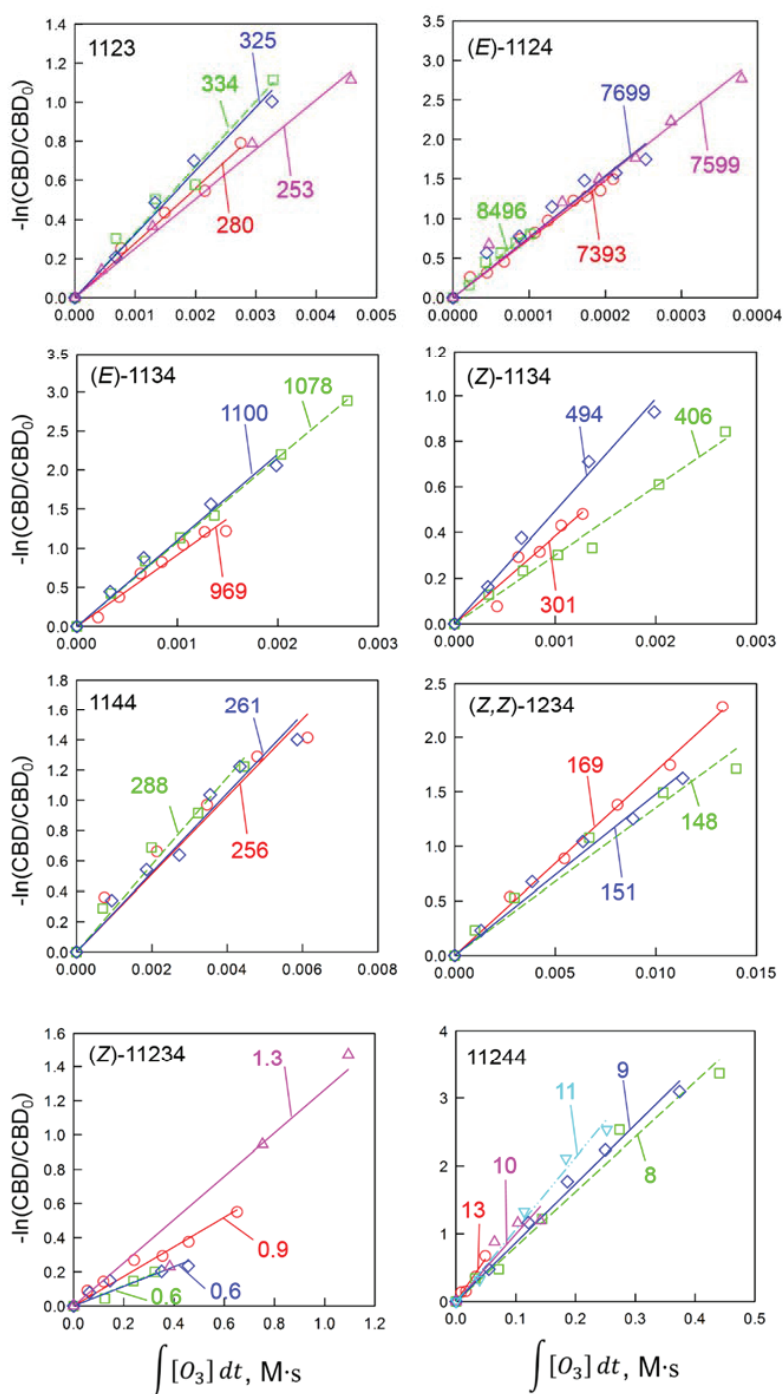


Figure S3.2. Second order rate constants (k_{O_3}) for the reaction of CBDs with ozone. Natural logarithm of the relative decrease of CBDs as a function of the ozone exposures obtained from ozonation experiments. y -intercept was set to 0 for the linear regression. Experimental conditions: $\text{pH} = 2$, $[\text{CBDs}]_0 = 1.0 - 3.5 \mu\text{M}$, and $[t\text{-BuOH}] = 5\text{mM}$. The number indicated to a slope of individual linear regression lines represents the determined k_{O_3} .

Supporting information for chapter 3

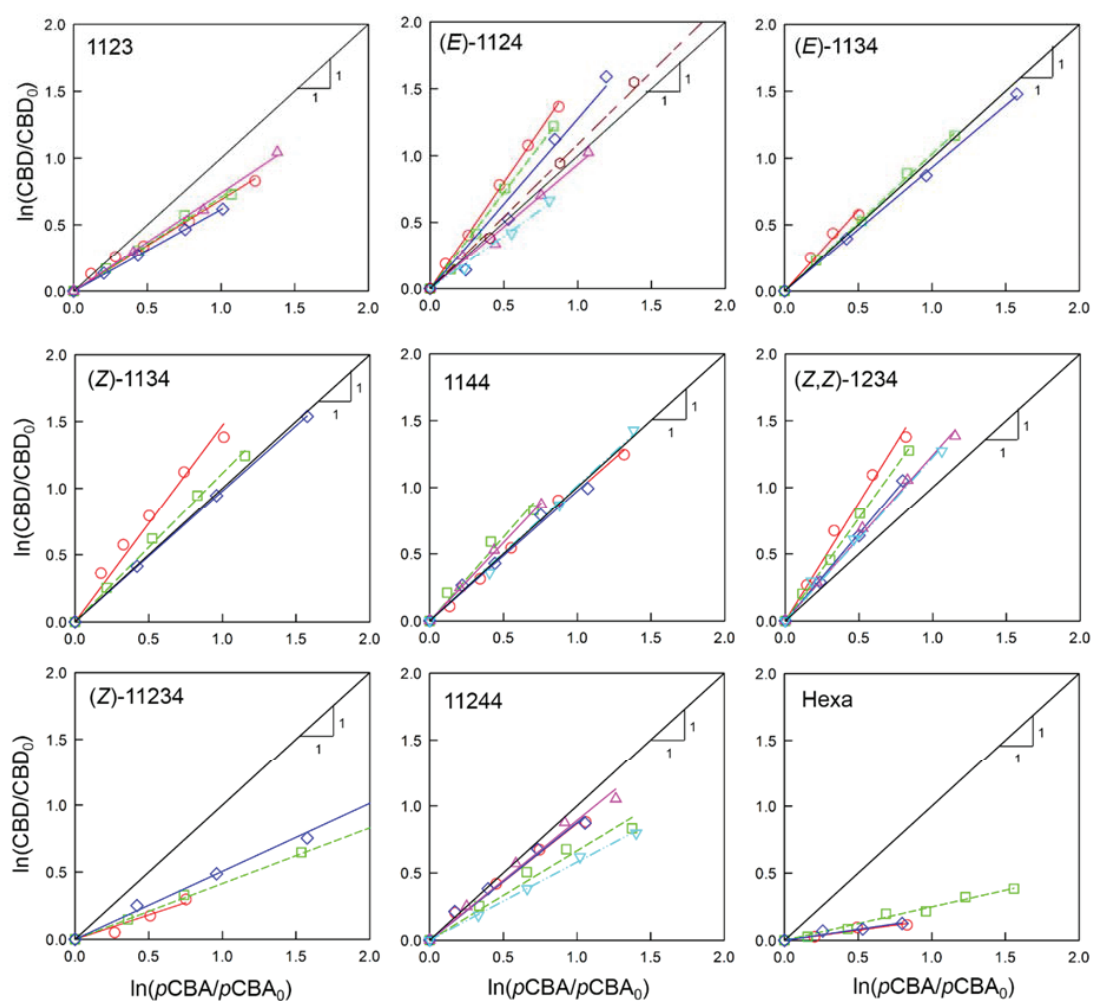


Figure S3.3. Competition kinetics to determine the second order rate constants ($k_{\text{OH,CBD}}$) for the reactions of CBDs with hydroxyl radicals. Natural logarithm of the relative residual concentration of CBDs versus the corresponding natural logarithm of the relative residual concentration of *p*CBA during UV irradiation by the MP-UV (313nm). y -intercept was set to 0 for the linear regression. Experimental conditions: pH 7 buffered by 5 mM phosphate, $[\text{CBDs}]_0 = 1.0 - 3.5 \mu\text{M}$, $[\text{pCBA}]_0 = 0.2 - 2.0 \mu\text{M}$, and $[\text{H}_2\text{O}_2]_0 = 10 \text{ mM}$. The slopes of the straight lines represent $k_{\text{OH,CBD}}/k_{\text{OH,pCBA}}$

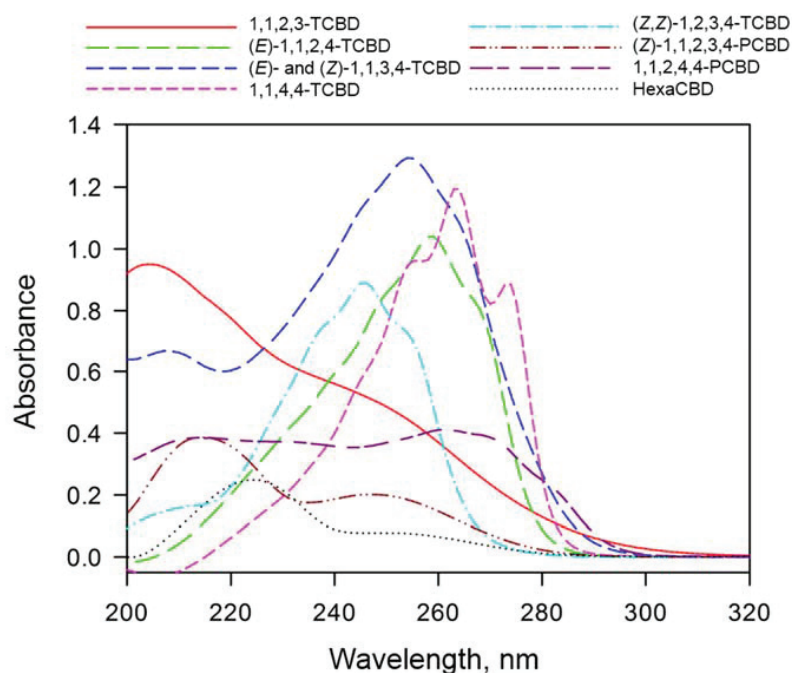


Figure S3.4. UV spectra between 200 and 320 nm for the selected chlorobutadienes (CBDs) (path-length = 10 cm). 1,1,2,3- (7.2 μM), (*E*)-1,1,2,4- (3.8 μM), 1,1,3,4- (8.2 μM as the total concentration of the two isomer mixture), 1,1,4,4- (2.9 μM), and (*Z,Z*)-1,2,3,4-TCBD (4.5 μM), (*Z*)-1,1,2,3,4- (7.2 μM) and 1,1,2,4,4-PCBD (7.1 μM), and HCBD (3.2 μM)

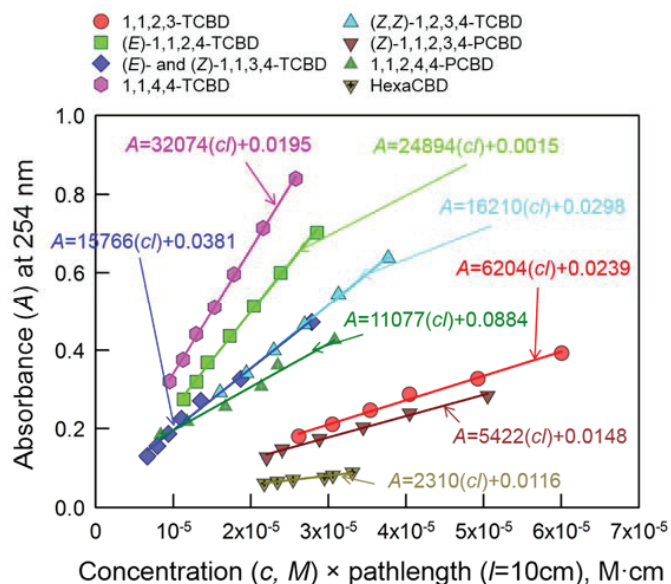


Figure S3.5. Determination of the molar absorption coefficients of the 9 selected CBDs (path length (*l*)=10 cm).

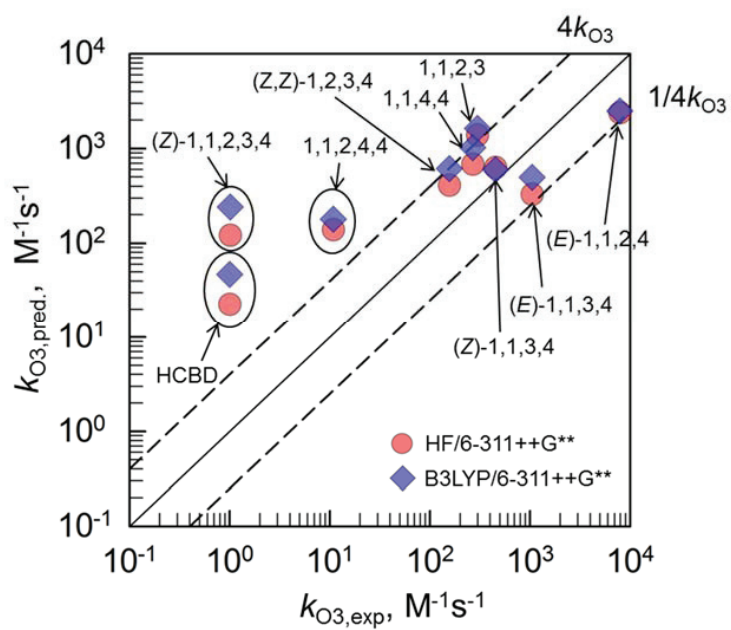


Figure S3.6. Predicted k_{O_3} ($k_{O_3,\text{pred.}}$) versus experimental k_{O_3} ($k_{O_3,\text{exp.}}$) for the nine selected CBDs. Both $k_{O_3,\text{exp.}}$ - and $k_{O_3,\text{pred.}}$ -values are from Table 3.1. The dotted line indicates an error of a factor of 4 of $k_{O_3,\text{pred.}}$.

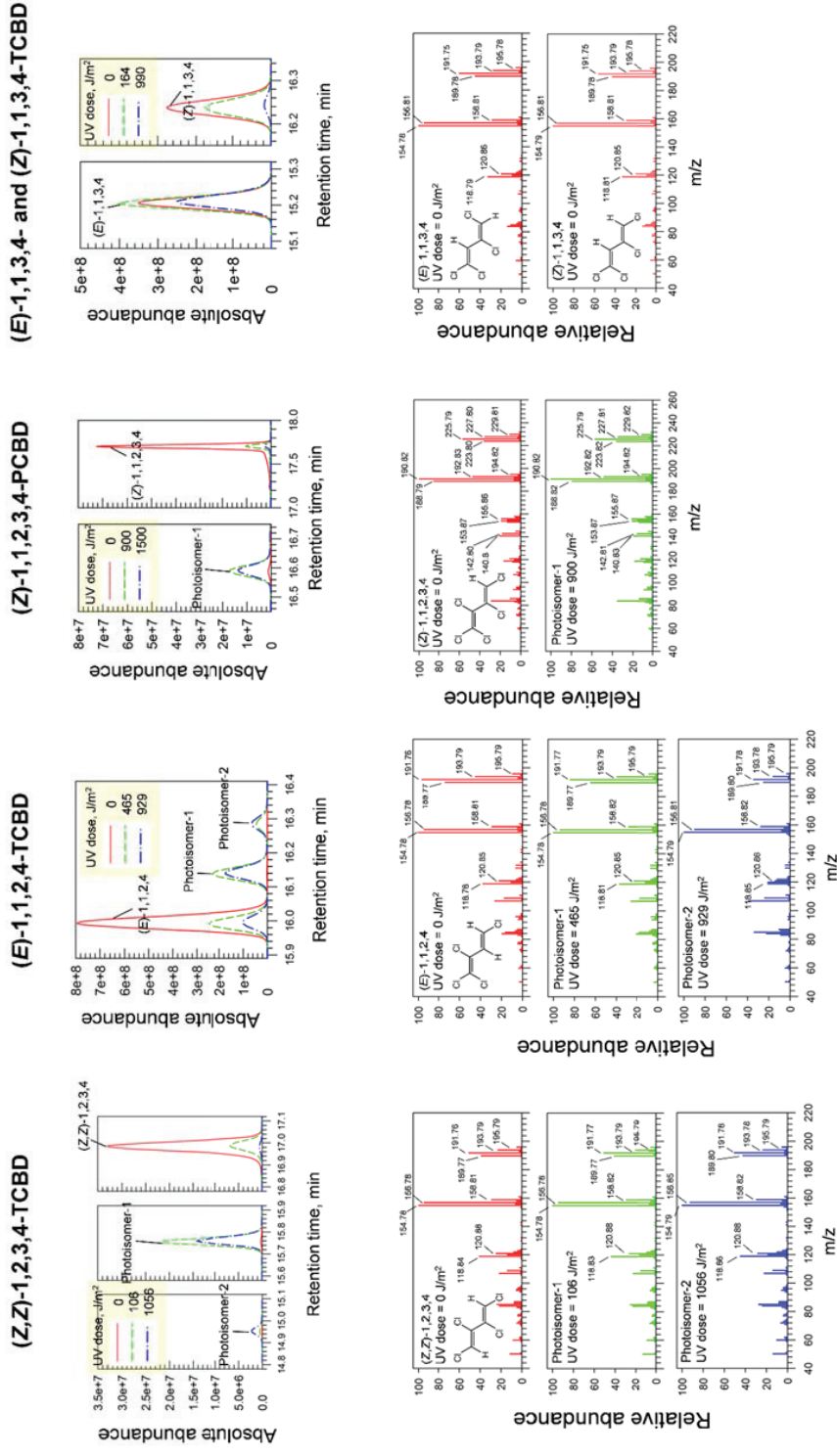


Figure S3.7. GC/MS chromatograms and mass spectra for photoisomerizable CBDs and their photoisomers. For the chromatograms, the peaks for parent and photoisomer CBDs are presented in an absolute abundance. This is only to visually guide their evolutions during UV irradiation. No quantitative information can be obtained because the responses from the three different samples were not corrected by an internal standard and no standards are available for photoisomers except for 1,1,3,4-TCBD.

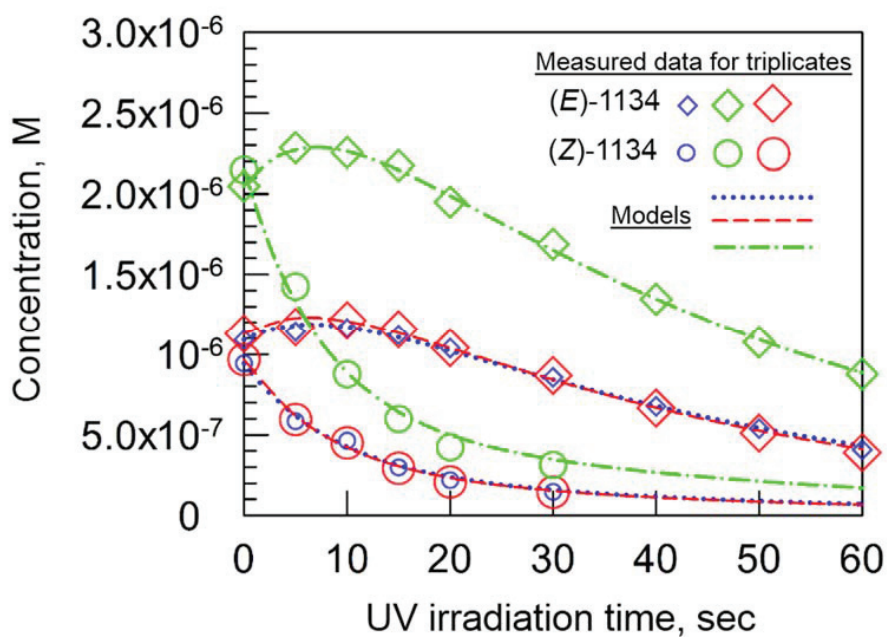


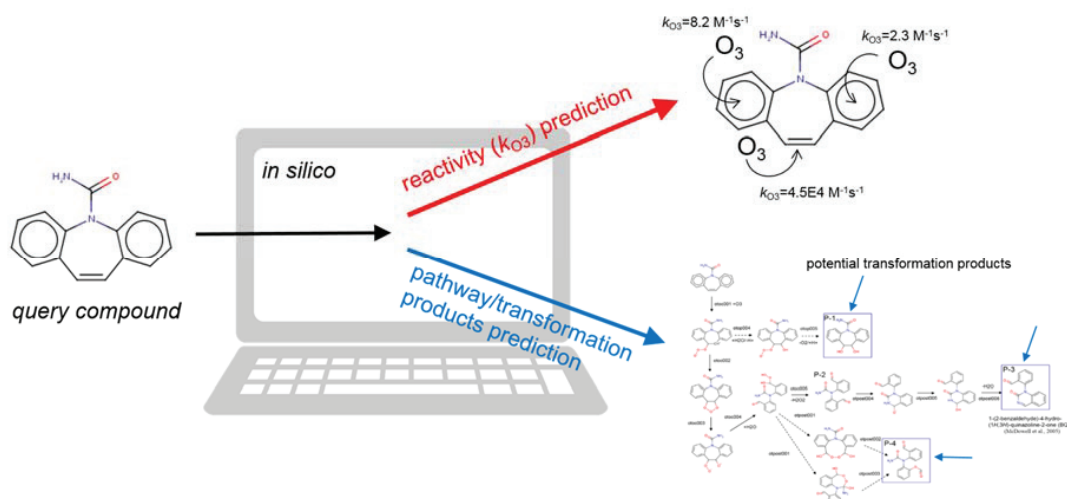
Figure S3.8. Experimental (symbols) and modeled (lines) evolution of a mixture of (*E*)-1,1,3,4- and (*Z*)-1,1,3,4-TCBDs (in triplicate) during UV irradiation at 254 nm. The modeled values (lines) were obtained based on Eqs. S3.7 and S3.8 by fitting of the experimental data.

References

- (1) von Sonntag, C.; von Gunten, U. *Chemistry of ozone in water and wastewater treatment: From basic principles to applications*; IWA publishing, 2012.
- (2) Bader, H.; Hoigné, J. Determination of ozone in water by the indigo method. *Water Res.* **1981**, *15*, 449–456.
- (3) Bader, H.; Sturzenegger, V.; Hoigné, J. Photometric method for the determination of low concentrations of hydrogen peroxide by the peroxidase catalyzed oxidation of N,N-diethyl-p-phenylenediamine (DPD). *Water Res.* **1988**, *22*, 1109–1115.
- (4) Neta, P., D. L. M. Pulse Radiolysis Studies XIII. Rate Constants for the Reaction of Hydroxyl Radicals with Aromatic Compounds in Aqueous Solutions, Advances in Chemistry Series. *Radiat. Chem.* **1968**, *81*, 222–230.
- (5) Wegelin, M.; Canonica, S.; Mechsner, K.; Fleischmann, T.; Pesaro, F.; Metzler, A. Solar water disinfection: Scope of the process and analysis of radiation experiments. *Aqua J. Water Supply Res. Technol.* **1994**, *43*, 154–169.
- (6) Canonica, S.; Jans, U.; Stemmler, K.; Hoigne, J. Transformation Kinetics of Phenols in Water: Photosensitization by Dissolved Natural Organic Material and Aromatic Ketones. *Environ. Sci. Technol.* **1995**, *29*, 1822–1831.
- (7) Lee, M.; Zimmermann-Steffens, S. G.; Arey, J. S.; Fenner, K.; von Gunten, U. Development of Prediction Models for the Reactivity of Organic Compounds with Ozone in Aqueous Solution by Quantum Chemical Calculations: The Role of Delocalized and Localized Molecular Orbitals. *Environ. Sci. Technol.* **2015**, *49*, 9925–9935.
- (8) Gundersen, G.; Nielsen, C. J.; Thomassen, H. G.; Becher, G. Vibrational spectra of trans,trans-1,2,3,4-tetrachloro-1,3-butadiene; and the molecular structure of trans,trans-1,2,3,4-tetrachloro-1,3-butadiene and of hexachloro-1,3-butadiene (a reinvestigation) determined by gas-phase electron diffraction. *J. Mol. Struct.* **1988**, *176*, 33–60.
- (9) Marais, D. J.; Sheppard, N.; Stoicheff, B. P. An investigation of the structure of butadiene by high resolution infra-red and raman spectroscopy. *Tetrahedron*, 1962, *17*, 163–169.
- (10) Almenningen, A.; Bastiansen, O.; Trætteberg, M.; Grönvall, A.; Zaar, B.; Diczfalusy, E. An Electron Diffraction Reinvestigation of the Molecular Structure of 1,3-Butadiene. *Acta Chemica Scandinavica*, 1958, *12*, 1221–1225.
- (11) Schomaker, V.; Pauling, L. The electron diffraction investigation of the structure of benzene, pyridine, pyrazine, butadiene-1,3, cyclopentadiene, furan, pyrrole, and thiophene. *J. Am. Chem. Soc.* **1939**, *61*, 1769–1780.
- (12) Furukawa, Y.; Takeuchi, H.; Harada, I.; Tasumi, M. Molecular Force Fields of s-trans-1,3-Butadiene and the Second Stable Conformer. *Bull. Chem. Soc. Jpn.* **1983**, *56*, 392–399.
- (13) Boopalachandran, P.; Craig, N. C.; Laane, J. Gas-Phase Raman Spectra of s-trans- and s-gauche-1,3-Butadiene and Their Deuterated Isotopologues. *J. Phys. Chem. A* **2011**, *116*, 271–281.
- (14) Craig, N. C.; Sams, R. L. An investigation of the rotamers of butadiene by high-resolution infrared spectroscopy. *J. Phys. Chem. A* **2008**, *112*, 12637–12646.
- (15) Cramer, C. J. *Essentials of computational chemistry: theories and models*; John Wiley & Sons, 2013.
- (16) Ianni, J. C. Kintecus, Windows Version 5.00, 2014.
- (17) Nelder, J. A.; Mead, R. A Simplex Method for Function Minimization. *Comput. J.* **1965**, *7*, 308–313.
- (18) Botta, D.; Mantica, E.; Castellani, L.; Dotelli, G.; Zetta, L. ¹H and ¹³C NMR study of some polychlorobuta-1,3-diene derivatives. *Magn. Reson. Chem.* **1998**, *36*, 885–891.

Chapter 4.

Development of a computer-based prediction platform for the reaction of ozone with organic compounds in aqueous solution: Kinetics and mechanisms



Lee, M., Blum, L. C., Schmid, E., Fenner, K., von Gunten, U., *in preparation*

This chapter was prepared by Lee, M. (overall contribution >70%) with scientific advices and editorial comments by all co-authors, particularly under a joint supervision of Fenner, K. and von Gunten, U.

Blum, L. C. and Schmid, E. developed the graphical-user-interface (GUI) application of a computer-based prediction platform.

Abstract

An ozonation of secondary wastewater effluents can reduce the discharge of parent micropollutants by transforming their chemical structures. Therefore, a better understanding of the formation of transformation products during ozonation is important. In this chapter, a computer-based prediction platform for kinetics and mechanisms for the reactions of ozone with organic compounds was developed to pursue *in silico* predictions for transformation products. In the developed prediction platform, the reaction kinetics expressed as second-order rate constants for the reactions of ozone with selected organic compounds (k_{O_3} , $M^{-1}s^{-1}$) can be predicted based on an adapted k_{O_3} prediction models from a previous study (Lee et al., 2015¹) (mostly $R^2 = 0.75-0.95$ with on average an error of about 6-fold for 14 compound classes with 284 model compounds). Ozone reaction mechanisms in literature have been reviewed and encoded into about 340 individual reaction rules using chemoinformatics tools, which can be generally applied to predict transformation products of a micropollutant. The demonstrations of the predictions for k_{O_3} and/or transformation products were overall consistent with experimental data for some micropollutants (e.g., carbamazepine, tramadol, and triclosan). However, the limitations of the current prediction platform were also identified, e.g., the k_{O_3} prediction for one tested micropollutant (cetirizine) was off by > 2 orders of magnitude, demonstrating that further improvements are needed. The developed prediction platform, made available as a stand-alone graphical user interface (GUI) application, is expected to provide valuable and critical information to various groups of end users such as environmental chemists, engineers, or toxicologists.

4.1. Introduction

Ozonation, which has been used for drinking water disinfection since the early 20th century,² has recently been considered as a viable tertiary treatment process for mitigating the release of various micropollutants (e.g., pharmaceuticals, personal care products, pesticides, etc.) from secondary wastewater effluents.²⁻⁶ However, as ozone mainly transforms the chemical structure of a micropollutant rather than mineralizing it, the formation of transformation products has been a concern related to e.g., the biological effects.^{2,7-11} To this end, numerous studies have been carried out with ecotoxicological test systems employing various *in vitro* and *in vivo* bioassays after ozonation. The findings generally support the beneficial effects of ozonation, which was demonstrated both for individual compounds with exposure-based essays¹¹⁻¹⁵ and for ozonated wastewaters (i.e., effect-based assays)¹⁶⁻²⁰. It is noteworthy, however, that in certain wastewaters elevated biological effects were observed after ozonation, which were later reduced in a subsequent biological treatment step.²¹⁻²³

The identification of ozone transformation products is a prerequisite for the empirical or theoretical assessment of their toxicity and biodegradability. Considering the ever increasing number of chemicals detected in natural and technical aquatic systems, an empirical elucidation of the transformation products for individual micropollutants is a formidable task to be achieved.

Over the last decades, knowledge on aqueous chemistry regarding reaction kinetics and mechanisms for the reactions of ozone with organic compounds, has greatly advanced.^{2,24,25} Several hundred second-order rate constants (k_{O_3}) for the reactions of ozone with organic compounds were empirically determined. Various classes of compounds such as aromatic compounds, olefins, amines, and organosulfur compounds were investigated. The range of k_{O_3} spans over >11 orders of magnitude from $<10^{-2} \text{ M}^{-1}\text{s}^{-1}$ – $10^9 \text{ M}^{-1}\text{s}^{-1}$.² With these experimental k_{O_3} -values as a basis, several k_{O_3} prediction tool such as a quantitative structure-activity relationship (QSAR) model²⁶ and a quantum chemical (QC) model¹ have been developed. Moreover, for numerous organic substances, reaction mechanisms were proposed based on the identified intermediates and final products.² Common intermediates and transformation products were often reported for differing organic compounds with similar ozone-reacting moieties: Criegee products for olefins (e.g., cephalexin,¹¹ progesterone²⁷), hydroxylation and ring cleavage for aromatic ring-containing compounds (e.g., phenol,^{28,29} bisphenol A,³⁰ and methoxylated benzenes³¹) and *N*-oxide formation and *N*-dealkylation for tertiary amino group-containing compounds (e.g., tramadol³² and clarithromycin³³), hydroxylamine for secondary amines (e.g., propranolol,³⁴ piperidine,³⁵ and morpholine³⁵), respectively. Therefore, it seems possible to deduce chemical structures of ozone transformation products of hitherto uninvestigated compounds if a reference study is available.

Based on available kinetic and mechanistic information for aqueous ozone reactions, a computer-based prediction platform for kinetics and pathways for reactions of organic compounds with ozone in aqueous solution has been developed in this chapter. It largely consists of two prediction modes: (i) reactivity (k_{O_3}) prediction and (ii) pathway prediction. In (i), k_{O_3} -values for chemical moieties of a query compound potentially reacting with ozone are provided. Either k_{O_3} is predicted by a prediction protocol that has been derived from a previously developed quantum chemical model¹ or an estimate

for k_{O_3} from a chemical analogue is suggested. In (ii), ozone reaction pathways proposed in relevant peer-reviewed research articles have been generalized into transformation rules to predict potential reaction pathways and transformation products. In this chapter, we present a detailed description of the development of the prediction platform and demonstrate its application for some examples.

4.2. Materials and methods

4.2.1. Chemoinformatics and quantum chemical computation tools

Various applications of the MarvinBeans/JChem package (Linux version 16.2.29.0, ChemAxon)³⁶ in Java have been used as follows. Marvin was used for drawing input molecules, converting them into 1D SMILES (simplified molecular-input line-entry system)³⁷ strings, and visualizing molecular structures. The pK_a calculator plugin was used for generating acid-base species, predicting acid dissociation constants, and producing a species distribution as a function of pH. Conformer plugin and Reactor were utilized for generating the 3D structure of a molecular structure from a 1D SMILES string for further quantum chemical computations and for enumerating reaction pathways, respectively. For quantum chemical computations, ORCA³⁸ was used for semi-empirical and *ab initio* quantum chemical computations. Natural bond orbital (NBO) analysis was performed by the NBO program 6.0.³⁹

4.2.2. Development of a computer-based prediction platform

A computer-based prediction platform for ozone reactions was developed based on the workflow given in Figure 4.1. To initiate a prediction, a *query compound* has to be submitted. A *query compound* is limited to organic compounds comprising of carbon and atoms such as hydrogen, nitrogen, oxygen, fluorine, phosphorus, sulfur, chlorine, bromine, and iodine. With the submission of a *query compound*, a user can choose between two prediction modes: (i) *direct prediction* and (ii) *pH-dependent prediction*. (i) The prediction is performed exclusively for the species specified as a chemical structure by the user, while in (ii) a comprehensive pH-dependent prediction for a query compound undergoing acid-base speciation is performed.

For a *direct prediction*, a *query compound* is subject to (a) a *reactive site search*, which elucidates chemical moieties in the *query compound* potentially reacting with ozone and presents them to the user. In a further step, the user can choose either (b) *rate constant (k_{O_3}) prediction* or (c) *pathway prediction*. In (b), k_{O_3} is predicted for the identified reactive sites and in (c) the reaction pathways for the reactive site selected by the user is enumerated. More information for (a) *reactive site search* and the two prediction modes is given below. Both (b) *rate constant (k_{O_3}) prediction* and (c) *pathway enumeration* can be operated independently or in combination, in which case (b) precedes (c). *rate constant (k_{O_3}) prediction* provides kinetic information (k_{O_3}) as to which reactive site reacts dominantly with ozone, and this is decisive for the pathway enumeration. For a pH-dependent prediction, (a) *reactive site search* and (d) *pH-dependent rate constant prediction* are comprehensively carried out for all the relevant acid-base species of the *query compound* for a specific pH or a pH range (see below for details). The pH-dependent prediction outputs provide information on the extent of the contribu-

tion of individual acid-base species to the overall reactivity at a certain pH. Based on the respective contributions of the individual species, the user can export the species of interest ((e) *species selection*) to the direct prediction train for further predictions (an arrow goes from (e) to (a) in Figure 4.1). Note that (a) and (b) were already implemented for the exported species from the pH-dependent prediction. Therefore, no additional processes for (a) and (b) are necessary. Further oxidation products can be predicted by feeding back the predicted transformation products as *query compounds*.

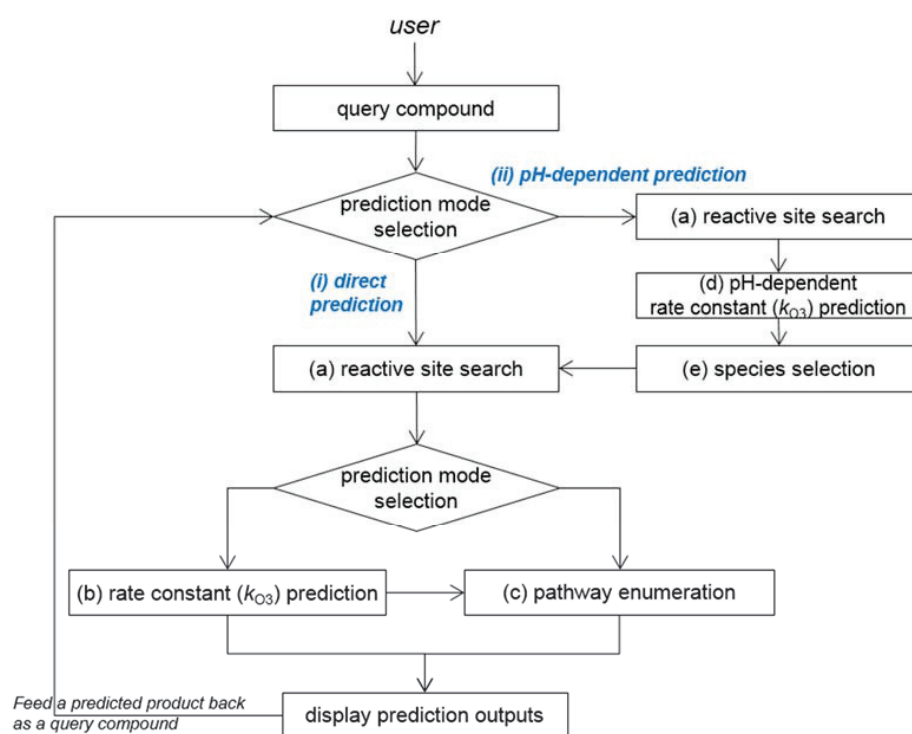


Figure 4.1 Workflow of the developed prediction platform to predict kinetics and transformation products for the reactions of ozone with organic compounds.

Reactive site search. A reactive site is a chemical moiety potentially reactive with ozone, for which a prediction can be implemented. An entire list of reactive sites currently defined (46 sites in total for five compound groups) is presented in the first column in Table S4.1 of the Supporting Information (SI) (Figure S4.1 for their chemical structures). The names of the reactive sites have been proposed not only to provide a chemical identification but also for appropriate/convenient assignments of an identified reactive site to the subsequent prediction models. As shown in Table S4.1, individual reactive sites are coupled with the corresponding k_{O_3} prediction model group and reaction pathway group in the third and the fourth columns, respectively. Details for definitions for individual reactive sites and their assignments are presented in Text S4.1. A SMARTS (SMiles ARbitrary Target Specification)⁴⁰ string is used for identifying substructural patterns (i.e., reactive sites) in a *query compound*.

Chapter 4

Two differing sets of chemical compounds, ~300 compounds used to develop k_{O_3} prediction models in this study and ~ 500 environmentally relevant micropollutants used as target analytes in the Department of Environmental Chemistry, Eawag (Switzerland), have been selected to train the *reactive site search* module. It has been manually checked that reactive sites for all the compounds are assigned as intended (data not shown).

Rate constant prediction. k_{O_3} for a reactive site denoted as ‘*prediction*’ in the 2nd column in Table S4.1 is to be predicted by the corresponding quantum chemical models (3rd column in Table S4.1) adapted from the originally proposed model.¹ For a reactive site with no k_{O_3} prediction model assigned, either an empirical k_{O_3} estimate (2nd column in Table S4.1) is used instead when there is no prediction model developed but a reference value is available (e.g., 200 M⁻¹s⁻¹ for ethynyl group determined for 1-ethynyl-1-cyclohexanol¹⁰), or k_{O_3} is considered unknown (denoted as *n.a.* in 2nd column in Table S4.1) when a site is potentially ozone-reactive but no information is available yet (e.g., thiophenol). k_{O_3} estimates are suggested based on the assumption that k_{O_3} is similar for structurally analogous reactive sites. Therefore, it should be noted that these k_{O_3} estimates would bear a large uncertainty, especially for a reactive site in a different environment than the reference case. These k_{O_3} estimates are summarized with the reference compounds in Table S4.2.

As mentioned above, ‘*prediction*’ in the second column in Table S4.1 indicates that quantum chemical computations will be conducted to predict k_{O_3} for the reactive site. This is achieved by applying the corresponding prediction model with the associated orbital energy (3rd column in Table S4.1) following the workflow shown in Figure S4.4. A k_{O_3} prediction consists of a ‘*speciation analysis*’ followed by a ‘*quantum chemical computation*’. Briefly, a speciation analysis is for a query compound undergoing acid-base speciation for which a specific pH or any pH range between 0 and 14 can be chosen. Compound-specific speciation parameters such as dissociation constants (pK_a) and a tautomeric fraction (f), if relevant, can be either predicted on the fly or provided by the user. Based on this, a species distribution between pH 0 and 14 for the query compound is derived. f indicates the fraction of individual tautomers (i.e., species with the same net charge). A more detailed explanation of the species analysis is given in Text S4.2.

Quantum chemical computations are performed to obtain an orbital energy (e.g., highest occupied molecular orbital energy (E_{HOMO}) or natural bond orbital energy (E_{NBO})) corresponding to a reactive site, which is then used in a linear prediction model to predict k_{O_3} for the reactive site. The computation protocol denoted as ‘HF/3-21G//MMFF94’, which was adapted from the original method in a previous study,¹ has been proposed to be universally applicable to all the compound groups in this study. HF/3-21G//MMFF94 implies that an orbital energy is derived from a single point calculation with the *ab initio* HF method using the Pople 3-21G basis set⁴¹⁻⁴⁵ for a geometry obtained using the molecular mechanics MMFF94 method.⁴⁶⁻⁵⁰ This protocol provides a similar prediction performance as the original model¹ while being much less computationally expensive. The methodology for development of the k_{O_3} prediction model (Text S4.3), the evaluation of the k_{O_3} prediction models with differing computational methods (Text S4.4), the detailed description and the computation costs of the computational methods chosen in this study for k_{O_3} predictions (Text S4.5), the hierarchy for the assignment of ambivalent aromatic compounds (Text S4.6), and the assignment of E_{HOMO-n} -values to the

corresponding reactive sites for a compound with multiple $E_{\text{HOMO-n}}$ -requiring reactive sites (Text S4.7) are presented in more detail in the corresponding texts in the SI, respectively. Combining both the predicted k_{O_3} and the speciation information for a *query compound*, k_{O_3} -values corresponding to a reactive site at a specific pH (or as a function of pH) can be predicted. A derivation of different types of predicted k_{O_3} such as (apparent) site-specific or species-specific k_{O_3} and their correspondence to experimentally measured k_{O_3} -values is described in detail in Text S4.8.

Pathway enumeration. Ozone reaction pathways proposed in peer-reviewed research articles have been comprehensively reviewed and compiled. The selected reaction pathways are discussed in detail in Text S4.9. Using Reactor of the JChem package (Chemaxon), all the selected reaction pathways (Schemes S4.1-S4.17) were segmented into individual unit reactions encoded as reaction rule files. Series of unit reaction rules have then been organized into pre-defined tree-like patterns to reproduce the structure of the reference pathways from literature. For instance, the reaction rule tree for phenols (Schemes S4.1-S4.3) is shown in Figure 4.2. It consists of three major branches, namely, “ortho”, “para”, and “radical”. In Figure 4.2, the sequence of reactions for the ortho position of phenols is shown, based on which the reaction rules were defined. The ortho branch is initiated with a reaction rule ‘otapho01’. When a *query compound* contains a phenol group, the rule ‘otapho01’ is triggered to produce an ozone adduct zwitterion at the ortho position. Note that although only phenol is shown in Figure 4.2 for simplicity, the reaction rules were encoded to be triggered for all phenol- and phenolate-containing compounds, respectively. The system continues to trigger subsequent reaction rules with the previous product until there is no match found. The ‘post-Criegee mechanism’ in Figure 4.2 is the pathways defined for the olefin group in Scheme S4.11 and is also applied to the resulting products formed from phenolic compounds via reactions (7)-(10) or (12)-(14). To date, about 340 unit reaction rules have been defined, based on the reaction pathways in Schemes S4.1-S4.17.

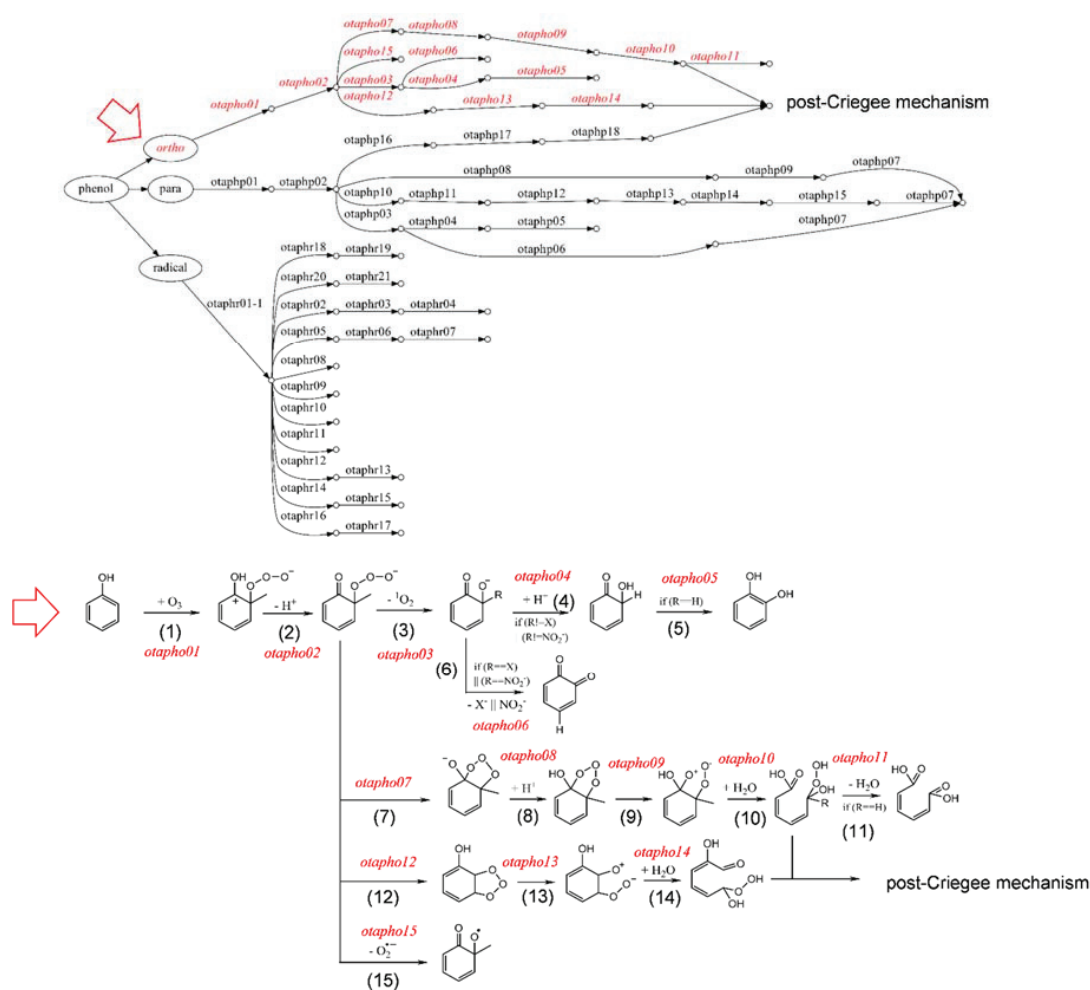


Figure 4.2 Reaction pathway enumeration (top) for a phenolic moiety and the reaction pathways (bottom) defined for the ortho position of phenol reacting with ozone. The corresponding pathways for “para” and “radical” are given in Schemes S4.2 and S4.3. The pathways for the post-Criegee mechanism, which deals with various ensuing reactions such as hydrolysis and fragmentations, are presented in Scheme S4.11. Note that further pathway predictions for muconic-type products is to be implemented by feeding such products back as an input structure.

4.3. Results and discussion

A standalone graphical user interface (GUI) has been developed for the prediction platform. Demonstrations for differing prediction modes, namely, direct k_{O_3} prediction, pH-dependent k_{O_3} prediction, and pathway enumeration, are presented below for selected micropollutants.

4.3.1. Rate constant (k_{O_3}) prediction

Direct k_{O_3} prediction. The main window of the developed prediction platform is shown in Figure 4.3. In the left panel named ‘Input’, the chemical structure for a *query compound* can be manually drawn

on a canvas or be imported as a 1D SMILES string. Carbamazepine is displayed as an example. The user can then choose between direct prediction (A) and pH-dependent prediction (B) (Figure 4.3). Direct prediction performs a reactive site search and displays the results on the right panel. For carbamazepine, two benzene rings and one olefin group were detected as reactive sites. Note that none of the two nitrogens in carbamazepine has been detected since amide nitrogens have a very low ozone reactivity² (see Text S4.1 for details). In the right panel, the user can choose to predict k_{O_3} (C), reaction pathways (D), or transformation products (E). For k_{O_3} prediction, no selection of the reactive site is necessary because relevant orbital energies are obtained all together from a single quantum chemical computation for the entire structure. In contrast, one needs to select a reactive site for pathway or product prediction as the reaction rules are applied to a specific reactive site (see below for more details).

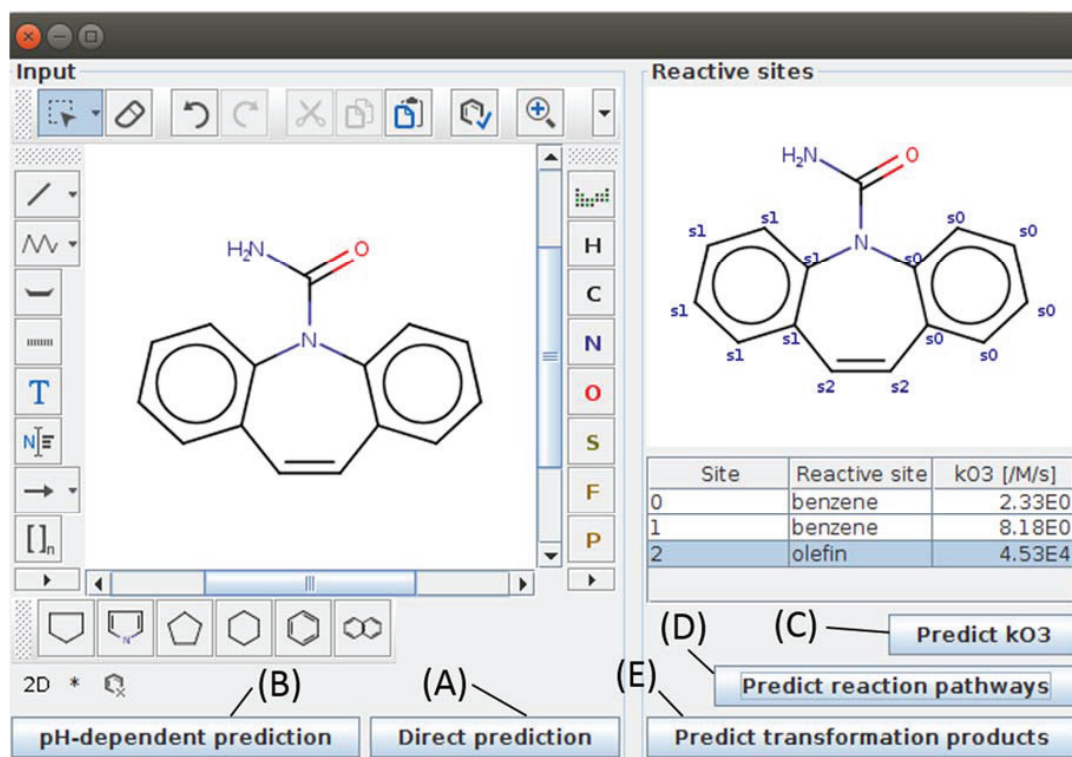


Figure 4.3. The main window for the developed kinetics and pathway prediction system. In the left panel of the main window, the user can choose between (A) direct prediction and (B) pH-dependent prediction. By selection of A, the right panel of the main window is activated and further predictions, (C), (D), or (E), can be conducted.

Upon choosing ‘Predict k_{O_3} ’, the quantum chemical computations are carried out following the workflow shown in Figure S4.4 without speciation analysis and the predicted k_{O_3} -values are displayed in the third column ($[k_{O_3} / M/s]$) after the computation. Together with the predicted k_{O_3} , k_{O_3} estimates are also presented for the assigned reactive sites. While $k_{O_3}=4.53 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ was predicted for the olefin, k_{O_3} of $2.3 \text{ M}^{-1}\text{s}^{-1}$ and $8.2 \text{ M}^{-1}\text{s}^{-1}$ was reported for both benzene rings. The prediction of two

differing k_{O_3} for the two benzene rings despite the same expected reactivity based on the symmetry of the molecule is attributed to the intrinsic ambiguity for the assignment of the responsible orbital energy (E_{HOMO}) to the benzene rings (see Texts S4.7 for more details). However, the difference in the predicted rate constants is small. Nonetheless, the k_{O_3} prediction indicates that the olefin moiety is the main site for ozone attack. Indeed, it was reported that the main oxidation products of carbamazepine resulted from an ozone attack at the olefin.⁸ Based on Eq. S4.1 in Text S4.8, the species-specific k_{O_3} is calculated as $4.5 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ (i.e., $4.5 \times 10^4 + 2.3 + 8.2$), which is lower than the experimental k_{O_3} ($\sim 3 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$,⁵¹) by a factor of 6.7. Compared to a mean unsigned error (MUE) of 0.57 for $\log k_{O_3}$ (corresponding to a factor of $10^{0.57} = 3.7$) for the general k_{O_3} prediction model for olefins (Table S4.3), the prediction for carbamazepine is rather at the bad end of the model predictions.

pH-dependent k_{O_3} prediction. Carbamazepine has no group that undergoes acid-base speciation in the environmentally relevant pH-range. Thus, its reactivity with ozone is barely influenced by pH. However, there are numerous organic compounds with changing acid-base speciations in the environmentally relevant pH range, which can significantly modify their reactivity. Therefore, a pH-dependent k_{O_3} prediction has been implemented in the prediction platform. Upon selecting *pH-dependent prediction* (B) for a *query compound* in the main window, a new window pops up (Figure 4.4). This window contains four panels, and to the right of them control tables are shown. The prediction output for tramadol is presented as an example. The upper-left pK_a panel (F) shows the moieties undergoing acid-base speciation for which the predicted pK_a -values are presented in the pK_a table (J). The bottom-left panel (G) shows the acid-base microspecies for the *query compound* and relevant information is shown in the microspecies table (K), with the tautomeric fraction (f) in the last column. Both pK_a and f can be manually replaced if empirical or theoretical values with a better accuracy are available.

By selecting ‘*Plot species distribution*’, a species distribution (H) for the query compound is generated based on the given pK_a -values and tautomeric fractions. The panel in the middle bottom (I) is a plot for apparent site-specific k_{O_3} ($k_{O_3,site(f)-app(pH)}$ in Eq. S4.2) for individual reactive sites for a user-defined pH range in Table (L). Apparent species-specific k_{O_3} ($k_{O_3,species(i)-app(pH)}$ in Eq. S4.3) for all microspecies can be calculated at a user-defined pH as in panel (M). In the last column, the contribution of $k_{O_3,species(i)-app(pH)}$ for individual microspecies (i) to the overall apparent k_{O_3} at the chosen pH is given. Upon selecting a microspecies and clicking ‘forward the selected species to main window’, the selected microspecies is exported to the main window where further k_{O_3} predictions and pathway predictions for the forwarded microspecies can be implemented.

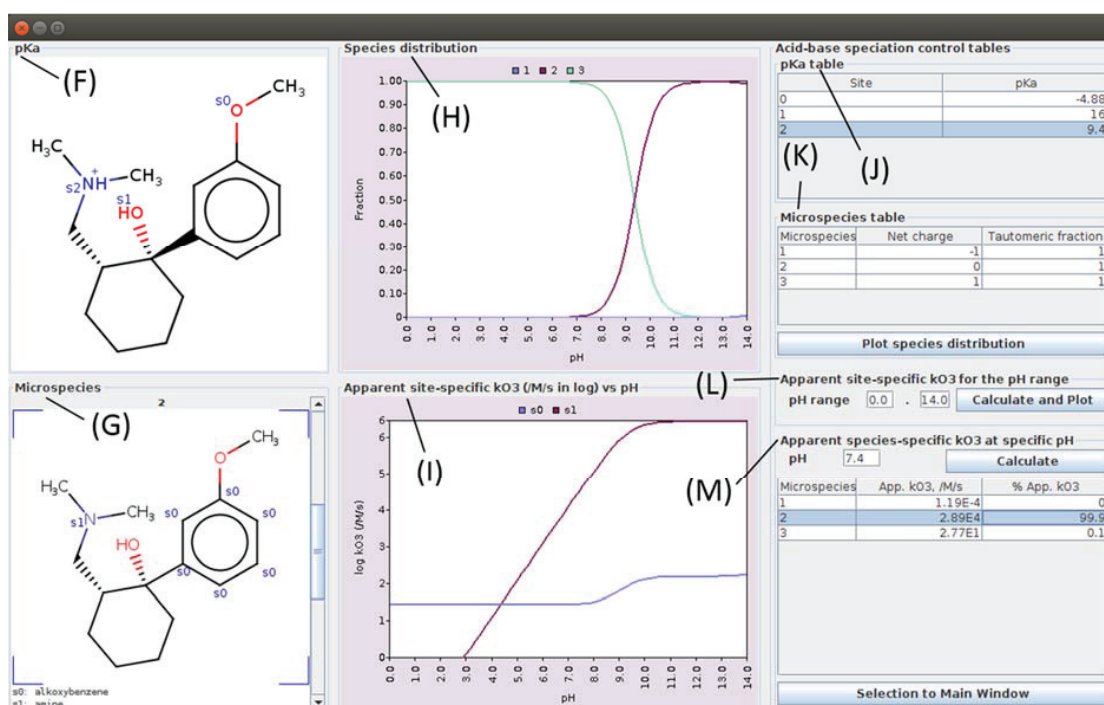


Figure 4.4. Window for the pH-dependent prediction of the kinetics of tramadol oxidation by ozone. Panel (F) shows the identified ionizable groups of the *query compound* for which pK_a -values are given in the table (J). Panel (G) shows individual species of the *query compound* and the identified reactive sites. Panel (H) is a species distribution plot created by choosing ‘Plot species distribution’ in the control table (K). The species distribution is drawn based on the presented pK_a (J) and tautomeric fraction (f) (K). The pH-dependent site-specific k_{O_3} prediction can be implemented for a user-defined pH-range (L), and is presented in panel (I). In table (M), apparent species-specific k_{O_3} are predicted at a specific user-defined pH and the user can export (by ‘Selection to Main Window’) the species of interest to the main window in Figure 4.3 for further predictions. While the pK_a -value of -4.9 for the methoxy group was used as predicted by Marvin (Chemaxon. Ltd), the predicted pK_a -values of 9.2 and 13.8 for respective amino group and alcohol group were replaced by 9.4⁵² (experimental) and 16 (typical pK_a range for an alcohol group), respectively.

As shown in Figures 4.5 a and b, good agreements were obtained between the predicted apparent k_{O_3} ($k_{O_3,app}$) and the experimental $k_{O_3,app}$ for tramadol within a factor of 1.2 – 5.8 and for triclosan within a factor of 1.3– 4.0, respectively. For tramadol, the predicted dominant sites of reaction with ozone are the tertiary amine above and the benzene ring below pH 4.4, respectively. This is consistent with the experimental data, which shows the appearance of a plateau at pH < 5 (Figure 4.5a). Interestingly, a pH-dependence was observed in the predicted k_{O_3} for the aromatic ring of tramadol. Two factors seem to be associated with this phenomenon. Firstly, the protonation of the amine led to a decreased E_{HOMO} for the aromatic ring by its electron withdrawing property and concomitantly to a decrease in k_{O_3} . Secondly, the conformation of the protonated tramadol in the gas phase is bent in a way that a hydrogen-bond is formed between the alcohol and the amine. Consequently, the inductive effect of the protonated nitrogen on the aromatic ring may be enhanced as it is closer to the aromatic ring. As described in Text S4.4. a gas-phase geometry is used for quantum chemical computations. A potential error in a predicted k_{O_3} associated with the use of a gas-phase geometry is addressed in detail for ce-

tirizine (Text S4.10). E_{HOMO} for the aromatic ring of a spread-out conformation without the hydrogen bond was calculated (data not shown) and observed to be higher than E_{HOMO} for the protonated tramadol in a bent conformation but still lower than E_{HOMO} for the deprotonated tramadol. This suggests that both inductive effect by the protonated nitrogen and the hydrogen bond in a gas-phase conformation are implicated in the observed pH-dependence of the predicted k_{O_3} for the aromatic ring.

For triclosan, it was predicted that the ozone-phenol reaction would dominate over the whole pH range. In contrast to those successful predictions, a poor prediction with a difference to the measurements of more than two orders of magnitude was observed for cetirizine. Efforts to improve the prediction for cetirizine were made. However, only minor improvements could be achieved by using an aqueous-phase relevant geometry for cetirizine. Thus, it can be concluded that the k_{O_3} prediction model developed for amino groups fails in the case of cetirizine (see Text S4.10 for more details).

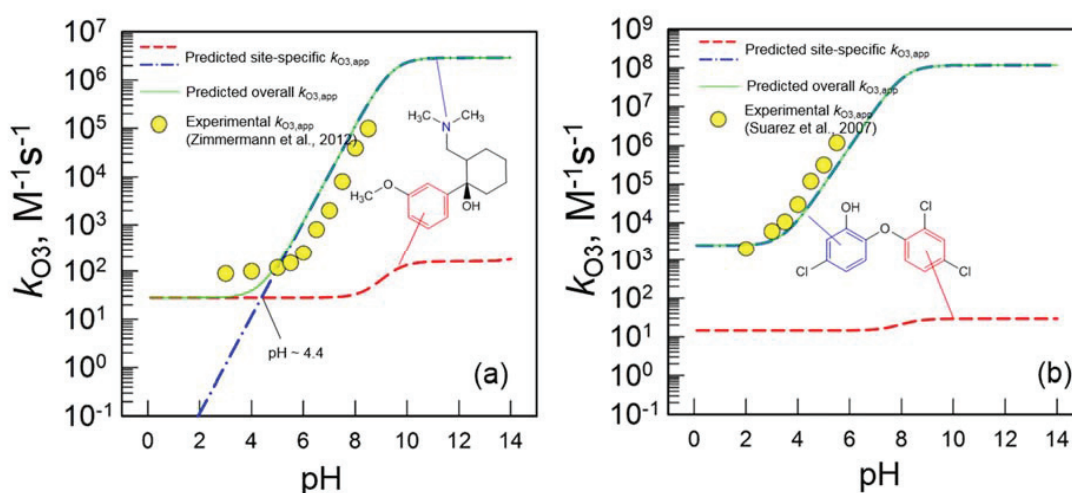


Figure 4.5. Comparison between predicted and experimental k_{O_3} as a function of pH for (a) tramadol and (b) triclosan. The k_{O_3} prediction outputs for both compounds are plotted and compared with the experimental data.

4.3.2. Pathway enumeration

By selection of a reactive site for the pathway prediction and choosing ‘*Predict reaction pathways*’ (D) in Figure 4.3, a pathway enumeration is carried out. From the k_{O_3} prediction above, it was identified that the olefin or the amine moiety is the dominant reactive site with ozone at pH 7 for carbamazepine or tramadol, respectively. The corresponding predicted pathways for carbamazepine and tramadol are presented in Figures 4.6 or 4.7, respectively. Note that raw pathway outputs were re-arranged for an appropriate representation in Figures 4.6 and 4.7 with no change in the contents except for the solid and dashed arrows (see below for details). As shown in Figure 4.6, carbamazepine underwent a typical Criegee mechanism (otoc001-otoc005) (Scheme S4.11), giving rise to a Criegee product (P-2) following H_2O_2 elimination. Additionally, a partial oxidation (otop004 and otop005) giving rise to product (P-1) has been proposed (see Text S4.9). Pathways with dashed arrows are considered less

likely, based on the reference information and authors' judgement. Note that a likelihood representation by solid and dashed arrows is not yet available in the current prediction platform and planned to be implemented in the near future. Prior to H₂O₂ elimination, a Bayer-Villiger type reaction (otpost001-otpost002) may occur to form an ester (P-4) (Text S4.9 and Scheme S4.17 for details), which is of minor importance (dashed arrows). Upon an intramolecular reaction (otpost004-otpost006), the Criegee product (P-2) transforms into P-3, 1-(2-benzaldehyde)-4-hydro-(1*H*,3*H*)-quinazoline-2-one (BQM), which was empirically found to be the major oxidation product of carbamazepine during ozonation.⁸ Note that the predicted BQM from ozonation of carbamazepine is really only a confirmation for the correct definition of the reaction rules because the intramolecular reactions (otpost004-otpost006) were defined based on the reference study for carbamazepine.⁸ To our knowledge, there is no other empirical data available to these intramolecular pathways. Therefore, its general applicability remains to be tested. For ozonation of tramadol (Figure 4.7), tramadol *N*-oxide (P-5), *N*-desmethyl-tramadol (P-6), and the *N*-deaminated-tramadol (P-10) were predicted to be the major products. This is consistent with the experimental observation with the exception that deamination was of minor importance in the experiments. Minor products (P-8, P-9, P-12, and P-13) formed via the decay of tetroxide intermediates were also considered (dashed arrows). In contrast to carbamazepine, the reaction rules applied to tramadol were defined from other reference compounds (Text S4.9) and found to be successfully applicable to tramadol. In that sense, the predicted tramadol pathway can be considered as a validation by external data. In addition to the entire pathway, transformation products, which are considered to be stable, thus potentially relevant to an empirical detection, are implemented by selecting '*Predict transformation products*' (E). For example, only the products in blue boxes in Figures S4.6 and S4.7 will be presented for carbamazepine or tramadol, respectively.

Chapter 4

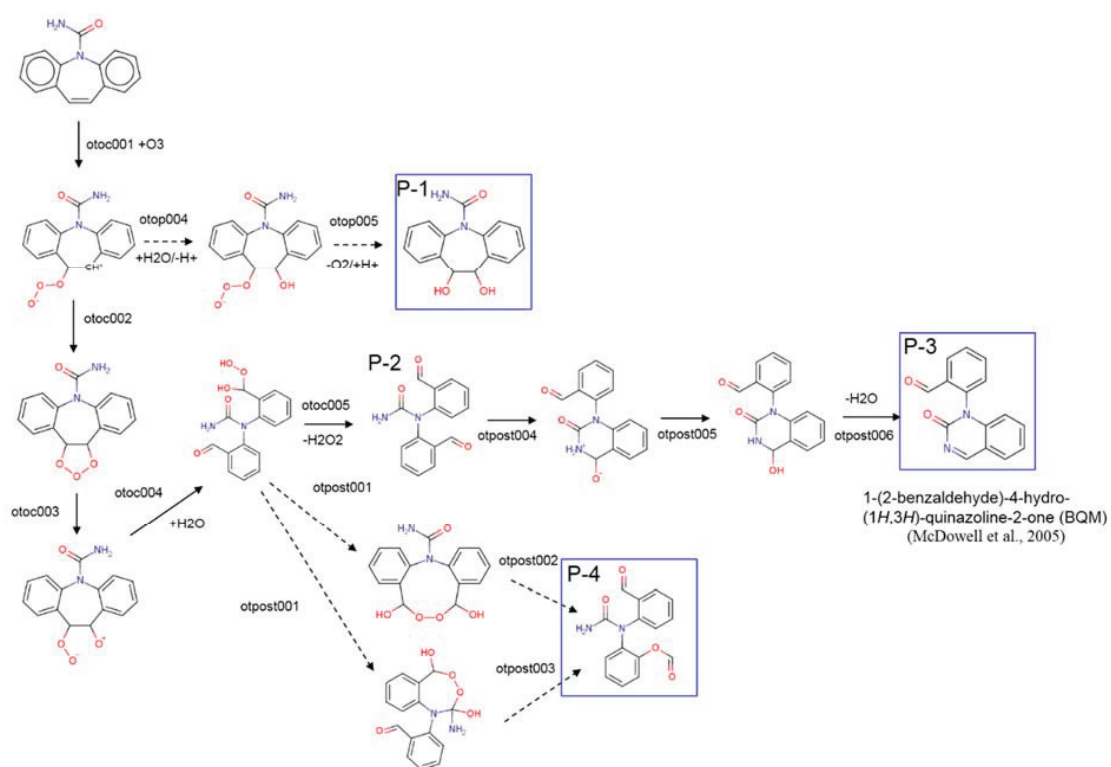


Figure 4.6. Predicted pathways for the reactions of carbamazepine with ozone. The dashed arrows indicate that the pathways are less likely. Predicted potential final products are presented in blue boxes. Product (P-3) was reported to be experimentally identified as the major transformation product.⁸

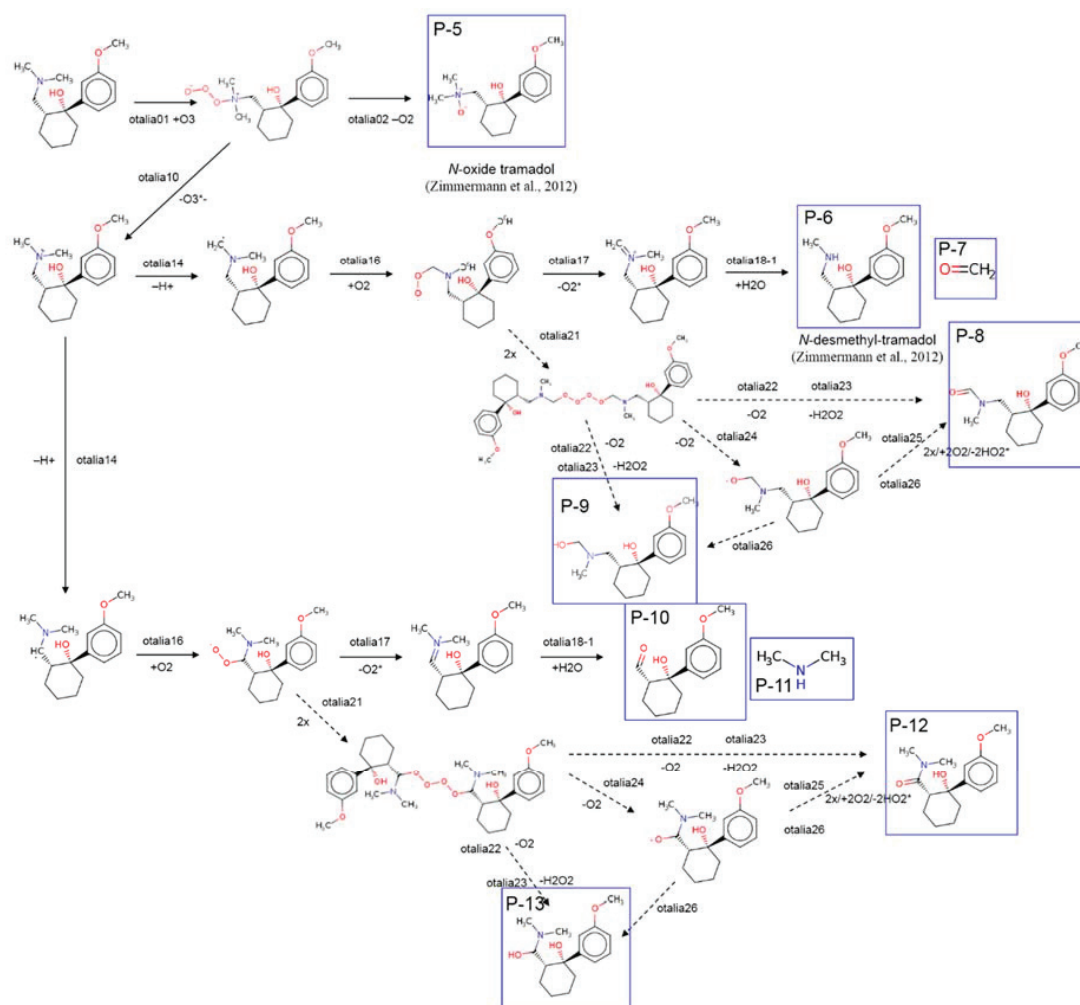


Figure 4.7. Predicted pathways for the reactions of tramadol with ozone. The dashed arrows indicate that the pathways are less likely. Predicted potential final transformation products are presented in blue boxes. The products P-5 and P-6 were reported to be the major experimental transformation products.

4.3.3. Practical implications

The developed computer-based prediction platform can serve various purposes. As demonstrated above, the predicted k_{O_3} -values can be used to identify a predominant site for ozone attack, which enables further pathway predictions. Moreover, k_{O_3} -values derived based on Eqs. S4.1 or S4.4 for compounds (not) undergoing acid-base speciation can be used for estimating the elimination efficiency during ozonation, for which theoretical/methodological details can be found elsewhere.^{1,2,4,25} Environmental engineers or treatment plant operators can utilize the predicted k_{O_3} for the optimization/design of an ozone treatment process for abating micropollutants. A predicted k_{O_3} -value can also be used to plan appropriate kinetic laboratory experiments. Pathway prediction can further be used as a tool for suspect screening, using high-resolution mass spectrometry, to detect and identify transformation products from the reaction of ozone with micropollutants. Moreover, the chemical structure of

Chapter 4

predicted transformation products can be screened for their toxicity using viable *in silico* toxicity prediction tools. Environmental chemists or toxicologists are considered to be a group of end-users who would benefit from these prediction outputs.

Besides possible applications of the currently developed prediction platform, major limitations and planned improvements have to be highlighted. As mentioned above, the current prediction tool does not give any information on reactions with hydroxyl radical. As k_{O_3} for a compound decreases, the contribution of hydroxyl radicals to its abatement and the formation of transformation product increases. Therefore, ozone transformation products predicted for compounds with low k_{O_3} , i.e., $<10^3 \text{ M}^{-1}\text{s}^{-1}$, may only be partially relevant for secondary wastewater effluents,⁴ because transformation products formed from reactions with hydroxyl radicals may be dominant. For detailed discussions see previous studies.^{1,2,4,25} As highlighted with the case study for cetirizine, a gas phase geometry may aggravate k_{O_3} prediction for some compounds. Therefore, further investigations on geometries relevant to the aqueous phase are necessary (Text S4.10). Hence, various chemical moieties need to be further investigated for a better understanding of the predictability for the developed k_{O_3} prediction models. For high-throughput screening purposes, a batch mode, which enables to process a large number of compounds, is currently under development.

Acknowledgements

This study was financially supported by the Swiss Federal Office for the Environment (FOEN). The authors would like to thank Peter Rudolf Tentscher (Eawag) for fruitful discussions. We thank Chemaxon Ltd. and the ORCA development team at the Max-Planck-Institute for Chemical Energy Conversion for providing academic licenses for MarvinBeans/JChem package and ORCA, respectively.

References

- (1) Lee, M.; Zimmermann-Steffens, S. G.; Arey, J. S.; Fenner, K.; von Gunten, U. Development of Prediction Models for the Reactivity of Organic Compounds with Ozone in Aqueous Solution by Quantum Chemical Calculations: The Role of Delocalized and Localized Molecular Orbitals. *Environ. Sci. Technol.* **2015**, *49*, 9925–9935.
- (2) von Sonntag, C.; von Gunten, U. *Chemistry of ozone in water and wastewater treatment: From basic principles to applications*; IWA publishing, 2012.
- (3) Lee, Y.; Kovalova, L.; McArdell, C. S.; von Gunten, U. Prediction of micropollutant elimination during ozonation of a hospital wastewater effluent. *Water Res.* **2014**, *64*, 134–148.
- (4) Lee, Y.; Gerrity, D.; Lee, M.; Bogeat, A. E.; Salhi, E.; Gamage, S.; Trenholm, R. A.; Wert, E. C.; Snyder, S. A.; von Gunten, U. Prediction of micropollutant elimination during ozonation of municipal wastewater effluents: use of kinetic and water specific information. *Environ. Sci. Technol.* **2013**, *47*, 5872–5881.
- (5) Huber, M. M.; Göbel, A.; Joss, A.; Hermann, N.; Löffler, D.; McArdell, C. S.; Ried, A.; Siegrist, H.; Ternes, T. A.; von Gunten, U. Oxidation of Pharmaceuticals during Ozonation of Municipal Wastewater Effluents: A Pilot Study. *Environ. Sci. Technol.* **2005**, *39*, 4290–4299.
- (6) Hollender, J.; Zimmermann, S. G.; Koepke, S.; Krauss, M.; McArdell, C. S.; Ort, C.; Singer, H.; von Gunten, U.; Siegrist, H. Elimination of Organic Micropollutants in a Municipal Wastewater Treatment Plant Upgraded with a Full-Scale Post-Ozonation Followed by Sand Filtration. *Environ. Sci. Technol.* **2009**, *43*, 7862–7869.
- (7) Sein, M. M.; Zedda, M.; Tuerk, J.; Schmidt, T. C.; Golloch, A.; von Sonntag, C. Oxidation of Diclofenac with Ozone in Aqueous Solution. *Environ. Sci. Technol.* **2008**, *42*, 6656–6662.
- (8) McDowell, D. C.; Huber, M. M.; Wagner, M.; von Gunten, U.; Ternes, T. A. Ozonation of Carbamazepine in Drinking Water: Identification and Kinetic Study of Major Oxidation Products. *Environ. Sci. Technol.* **2005**, *39*, 8014–8022.
- (9) Mawhinney, D. B.; Vanderford, B. J.; Snyder, S. A. Transformation of 1H-benzotriazole by ozone in aqueous solution. *Environ. Sci. Technol.* **2012**, *46*, 7102–7111.
- (10) Huber, M. M.; Ternes, T. A.; von Gunten, U. Removal of Estrogenic Activity and Formation of Oxidation Products during Ozonation of 17 α -Ethinylestradiol. *Environ. Sci. Technol.* **2004**, *38*, 5177–5186.
- (11) Dodd, M. C.; Rentsch, D.; Singer, H. P.; Kohler, H.-P. E.; von Gunten, U. Transformation of β -Lactam Antibacterial Agents during Aqueous Ozonation: Reaction Pathways and Quantitative Bioassay of Biologically-Active Oxidation Products. *Environ. Sci. Technol.* **2010**, *44*, 5940–5948.
- (12) Mestankova, H.; Schirmer, K.; Escher, B. I.; von Gunten, U.; Canonica, S. Removal of the antiviral agent oseltamivir and its biological activity by oxidative processes. *Environ. Pollut.* **2012**, *161*, 30–35.
- (13) Mestankova, H.; Escher, B.; Schirmer, K.; von Gunten, U.; Canonica, S. Evolution of algal toxicity during (photo)oxidative degradation of diuron. *Aquat. Toxicol.* **2011**, *101*, 466–473.
- (14) Lee, Y.; Escher, B. I.; von Gunten, U. Efficient Removal of Estrogenic Activity during Oxidative Treatment of Waters Containing Steroid Estrogens. *Environ. Sci. Technol.* **2008**, *42*, 6333–6339.
- (15) Mestankova, H.; Parker, A. M.; Bramaz, N.; Canonica, S.; Schirmer, K.; von Gunten, U.; Linden, K. G. Transformation of Contaminant Candidate List (CCL3) compounds during ozonation and advanced oxidation processes in drinking water: Assessment of biological effects. *Water Res.* **2016**, *93*, 110–120.
- (16) Reungoat, J.; Escher, B. I. I.; Macova, M.; Argaud, F. X. X.; Gernjak, W.; Keller, J. Ozonation and biological activated carbon filtration of wastewater treatment plant effluents. *Water Res.* **2012**, *46*, 863–872.
- (17) Margot, J.; Kienle, C.; Magnet, A.; Weil, M.; Rossi, L.; de Alencastro, L. F.; Abegglen, C.; Thonney, D.; Chèvre, N.; Schärer, M.; et al. Treatment of micropollutants in municipal wastewater: Ozone or powdered activated carbon? *Sci. Total Environ.* **2013**, *461-462*, 480–498.
- (18) Escher, B. I.; Bramaz, N.; Ort, C. JEM Spotlight: Monitoring the treatment efficiency of a full

- scale ozonation on a sewage treatment plant with a mode-of-action based test battery. *J. Environ. Monit.* **2009**, *11*, 1836.
- (19) Cao, N.; Yang, M.; Zhang, Y.; Hu, J.; Ike, M.; Hirotsuji, J.; Matsui, H.; Inoue, D.; Sei, K. Evaluation of wastewater reclamation technologies based on in vitro and in vivo bioassays. *Sci. Total Environ.* **2009**, *407*, 1588–1597.
- (20) Schindler Wildhaber, Y.; Mestankova, H.; Schärer, M.; Schirmer, K.; Salhi, E.; von Gunten, U. Novel test procedure to evaluate the treatability of wastewater with ozone. *Water Res.* **2015**, *75*, 324–335.
- (21) Stalter, D.; Magdeburg, A.; Weil, M.; Knacker, T.; Oehlmann, J. Toxication or detoxication? In vivo toxicity assessment of ozonation as advanced wastewater treatment with the rainbow trout. *Water Res.* **2010**, *44*, 439–448.
- (22) Stalter, D.; Magdeburg, A.; Oehlmann, J. Comparative toxicity assessment of ozone and activated carbon treated sewage effluents using an in vivo test battery. *Water Res.* **2010**, *44*, 2610–2620.
- (23) Magdeburg, A.; Stalter, D.; Oehlmann, J. Whole effluent toxicity assessment at a wastewater treatment plant upgraded with a full-scale post-ozonation using aquatic key species. *Chemosphere* **2012**, *88*, 1008–1014.
- (24) von Gunten, U. Ozonation of drinking water: part I. Oxidation kinetics and product formation. *Water Res.* **2003**, *37*, 1443–1467.
- (25) Lee, Y.; von Gunten, U. Advances in predicting organic contaminant abatement during ozonation of municipal wastewater effluent: reaction kinetics, transformation products, and changes of biological effects. *Environ. Sci. Water Res. Technol.* **2016**, *2*, 421–442.
- (26) Lee, Y.; von Gunten, U. Quantitative structure–activity relationships (QSARs) for the transformation of organic micropollutants during oxidative water treatment. *Water Res.* **2012**, *46*, 6177–6195.
- (27) Barron, E.; Deborde, M.; Rabouan, S.; Mazellier, P.; Legube, B. Kinetic and mechanistic investigations of progesterone reaction with ozone. *Water Res.* **2006**, *40*, 2181–2189.
- (28) Mvula, E.; von Sonntag, C. Ozonolysis of phenols in aqueous solution. *Org. Biomol. Chem.* **2003**, *1*, 1749–1756.
- (29) Ramseier, M. K.; von Gunten, U. Mechanisms of Phenol Ozonation—Kinetics of Formation of Primary and Secondary Reaction Products. *Ozone Sci. Eng.* **2009**, *31*, 201–215.
- (30) Deborde, M.; Rabouan, S.; Mazellier, P.; Duguet, J.-P.; Legube, B. Oxidation of bisphenol A by ozone in aqueous solution. *Water Res.* **2008**, *42*, 4299–4308.
- (31) Mvula, E.; Naumov, S.; von Sonntag, C. Ozonolysis of Lignin Models in Aqueous Solution: Anisole, 1,2-Dimethoxybenzene, 1,4-Dimethoxybenzene, and 1,3,5-Trimethoxybenzene. *Environ. Sci. Technol.* **2009**, *43*, 6275–6282.
- (32) Zimmermann, S. G.; Schmukat, A.; Schulz, M.; Benner, J.; von Gunten, U.; Ternes, T. A. Kinetic and Mechanistic Investigations of the Oxidation of Tramadol by Ferrate and Ozone. *Environ. Sci. Technol.* **2012**, *46*, 876–884.
- (33) Lange, F.; Cornelissen, S.; Kubac, D.; Sein, M. M.; von Sonntag, J.; Hannich, C. B.; Golloch, A.; Heipieper, H. J.; Möder, M.; von Sonntag, C. Degradation of macrolide antibiotics by ozone: a mechanistic case study with clarithromycin. *Chemosphere* **2006**, *65*, 17–23.
- (34) Benner, J.; Ternes, T. A. Ozonation of Propranolol: Formation of Oxidation Products. *Environ. Sci. Technol.* **2009**, *43*, 5086–5093.
- (35) Tekle-Röttering, A.; Jewell, K. S.; Reisz, E.; Lutze, H. V.; Ternes, T. A.; Schmidt, W.; Schmidt, T. C. Ozonation of piperidine, piperazine and morpholine: Kinetics, stoichiometry, product formation and mechanistic considerations. *Water Res.* **2016**, *88*, 960–971.
- (36) MarvinBeans/JChem suite (16.2.29.0), 2016, Chemaxon (<http://www.chemaxon.com>).
- (37) Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Model.* **1988**, *28*, 31–36.
- (38) Neese, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (39) Glendening, E. D.; Landis, C. R.; Weinhold, F. NBO 6.0: Natural bond orbital analysis program. *J. Comput. Chem.* **2013**, *34*, 1429–1437.

- (40) Daylight Chemical Information Systems, Inc., Laguna Niguel, CA, USA.
- (41) Binkley, J. S.; Pople, J. A.; Hehre, W. J. Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements. *J. Am. Chem. Soc.* **1980**, *102*, 939–947.
- (42) Gordon, M. S.; Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. Self-consistent molecular-orbital methods. 22. Small split-valence basis sets for second-row elements. *J. Am. Chem. Soc.* **1982**, *104*, 2797–2803.
- (43) Dobbs, K. D.; Hehre, W. J. Molecular orbital theory of the properties of inorganic and organometallic compounds 4. Extended basis sets for third-and fourth-row, main-group elements. *J. Comput. Chem.* **1986**, *7*, 359–378.
- (44) Dobbs, K. D.; Hehre, W. J. Molecular orbital theory of the properties of inorganic and organometallic compounds 5. Extended basis sets for first-row transition metals. *J. Comput. Chem.* **1987**, *8*, 861–879.
- (45) Dobbs, K. D.; Hehre, W. J. Molecular orbital theory of the properties of inorganic and organometallic compounds. 6. Extended basis sets for second-row transition metals. *J. Comput. Chem.* **1987**, *8*, 880–893.
- (46) Halgren, T. A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, *17*, 490–519.
- (47) Halgren, T. A. Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94. *J. Comput. Chem.* **1996**, *17*, 553–586.
- (48) Halgren, T. A. Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules. *J. Comput. Chem.* **1996**, *17*, 616–641.
- (49) Halgren, T. A. Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *J. Comput. Chem.* **1996**, *17*, 520–552.
- (50) Halgren, T. A.; Nachbar, R. B. Merck molecular force field. IV. conformational energies and geometries for MMFF94. *J. Comput. Chem.* **1996**, *17*, 587–615.
- (51) Huber, M. M. M.; Canonica, S.; Park, G.-Y.; von Gunten, U. Oxidation of Pharmaceuticals during Ozonation and Advanced Oxidation Processes. *Environ. Sci. Technol.* **2003**, *37*, 1016–1024.
- (52) Pospíšilová, M.; Poláček, M.; Jokl, V. Determination of tramadol in various dosage forms by capillary isotachopheresis. *J. Pharm. Biomed. Anal.* **1998**, *18*, 777–783.

Supporting information for chapter 4.

Development of a computer-based prediction platform for the reaction of ozone with organic compounds in aqueous solution: Kinetics and mechanisms

includes 10 texts, 8 tables, 13 figures, and 17 schemes.

Text S4.1 Definition of reactive sites and their connections to k_{O_3} prediction and pathway prediction

A reactive site is defined as a chemical moiety potentially reacting with ozone in a query compound and is therefore subject to a prediction in the current prediction platform. As shown in Table S4.1 and Figure S4.1 (chemical structures), various reactive sites for five differing functional groups such as aromatic compounds, olefin/ethynyl groups, amines, heteroaromatic compounds, and organosulfur compounds were defined. The nomenclature of the reactive sites was chosen such as to not only give chemically meaningful names but also for convenient assignments of a reactive site to the subsequent k_{O_3} prediction group (2nd and 3rd columns in Table S4.1) or reaction pathway group (4th column in Table S4.1).

Here, the defined reactive sites are individually discussed with respect to their assignment to a k_{O_3} prediction group and a reaction pathway group. An *italic* font was adopted as a convention for individual reactive sites hereafter. More details for the respective predictions are provided in Texts S4.4 and S4.9.

It should be noted that the currently defined reactive sites and the corresponding predictions were based on a limited number of currently available reference studies and does not cover all the chemical structures of micropollutants potentially found in natural and technical aquatic systems. Therefore, the predictions contain uncertainty and further empirical investigations are recommended for the verification of the predictions as well as for updates of the prediction platform.

Aromatic compounds Seven reactive sites are defined for the aromatic compounds, namely *phenol*, *3alkoxybenzene*, *aniline*, *cyclic_aniline*, *12alkoxybenzene*, *benzazole_dervs*, and *benzene*. *3alkoxybenzene* and *12alkoxybenzene* are abbreviations for trialkoxybenzene and mono- and dialkoxybenzene, respectively, and *dervs* in *benzazole_dervs* stands for derivatives. Five of those reactive sites, i.e., *phenol*, *aniline*, *3alkoxybenzene*, *12alkoxybenzene*, and *benzene*, were defined based on the k_{O_3} prediction models separately available for the respective compound groups (Text S4.4) as well as the distinctive reaction pathway groups (Text S4.9) (see below for *benzazole_dervs*). For an ambivalent case with multiple aromatic reactive sites in the same aromatic ring (e.g., phenol and aniline coexist in the same benzene ring), the assignment is carried out based on the proposed hierarchy (see Text S4.6 for details).

As shown in Figure S4.1 (below ‘Subclassification for [*aniline*]’), different subgroups were defined for *aniline*. The *aniline* group is defined exclusively for primary and tertiary anilines, where the tertiary anilines with the cyclic amine nitrogen (e.g., *N*-phenylpiperidine) are excluded from this definition (see below). In contrast, secondary anilines are divided into two sub-groups, namely, secondary anilines with (i) an acyclic nitrogen and (ii) a cyclic nitrogen. As for (i), the benzene ring is to be defined as any corresponding aromatic group depending on the hierarchy of the substituents and the acyclic nitrogen is assigned as an *amine*, respectively. For example, the amine nitrogen and the benzene ring of 4-(methoxy)-*N*-methylaniline are assigned as an *amine* and a *12alkoxybenzene*, respectively. This definition is based on the observation of ozonation of diclofenac¹ which bears an aniline with an acyclic secondary nitrogen. The nitrogen of diclofenac was reported to be a dominant reactive

site with ozone rather than the benzene ring, and this was also corroborated by the QSAR analysis for this compound using a Hammett coefficient.¹ Moreover, diclofenac has been successfully included in the currently developed k_{O_3} prediction model for amines using $E_{NBO,LP-N}$ rather than E_{HOMO} for the aniline group (data not shown). All this evidence suggests that the secondary aliphatic nitrogen seemingly behaves as an amine rather than an activated benzene ring. Therefore, the secondary aliphatic nitrogen of aniline was assigned as an *amine*, while its benzene ring was assigned to the corresponding aromatic group. Note that a prerequisite for these assignments as *aniline* and *amine* to be valid is that the nitrogen of the *aniline* and the *amine* should not be a part of any excluded moieties (e.g., amide) described in the ‘Amine’ section in Figure S4.1 (see ‘Amines’ section below for more details).

For (ii), secondary anilines with the cyclic nitrogen (e.g., tetrahydroquinoline) are defined as *cyclic_aniline* together with tertiary anilines with a cyclic nitrogen. As shown in Figure S4.1, any atom except hydrogen (expressed as ‘A’) can be in an aliphatic cyclic system along with the cyclic nitrogen. A k_{O_3} estimate of $450 \text{ M}^{-1}\text{s}^{-1}$ is assigned to *cyclic_aniline* as the average of experimental k_{O_3} -values for hydrochlorothiazide ($600 \text{ M}^{-1}\text{s}^{-1}$),² enrofloxacin ($330 \text{ M}^{-1}\text{s}^{-1}$ in diprotonated form),³ and ciprofloxacin ($400 \text{ M}^{-1}\text{s}^{-1}$ in diprotonated form).³ We assume that the benzene ring of these compounds is responsible for the observed k_{O_3} . The reason why the k_{O_3} estimate was given was because the experimental k_{O_3} of this class of compounds could not be predicted well with the currently available k_{O_3} prediction models with the E_{HOMO} for the benzene ring and the E_{NBO} for the cyclic amine (data not shown). Note that when the benzene ring of *cyclic_aniline* has a substituent corresponding to any aromatic compound group except *benzazole_dervs* or *benzene*, such an aromatic compound group will be detected over *cyclic_aniline*. For example, when the benzene ring of tetrahydroquinoline is substituted by a hydroxyl substituent, *phenol* is detected as the reactive site instead of the *cyclic_aniline*. This is based on the assumption that in presence of such a substituent, the k_{O_3} will be higher than the k_{O_3} estimate ($450 \text{ M}^{-1}\text{s}^{-1}$) due to the activation of the benzene ring by the substituent. Therefore, it has to be predicted by the corresponding k_{O_3} prediction model. There is no empirical information for the oxidation products from the reaction of ozone with a *cyclic amine* and it is currently assumed to be the same as for *aniline*. However, this remains to be verified empirically in future.

In addition to the reactive sites described above for the aromatic compound class, *benzazole_dervs* was additionally defined based on the experimental data for benzotriazoles. Differing benzotriazoles such as unsubstituted, 5-chloro, 5,6-dimethyl, and 5-methylbenzotriazole and their conjugate bases were included in the k_{O_3} prediction model for the benzene group (Text S4.9). However, an explicit definition of *benzazole_dervs* rather than including it in *benzene* was preferred for clarity but also to account for the additional reaction pathways associated with the heteroatom of a 5-membered ring (e.g., *N* or *S*-oxide formation in Scheme S4.10). It is noted that benzazoles feature a benzene ring fused with azoles at the 4- and 5-positions. As an azole is defined as a five-membered heteroaromatic ring containing a nitrogen and at least one other heteroatom (nitrogen, oxygen, or sulfur), it by definition does not include isoindole, isobenzofuran, and 2-benzothiophene shown in Figure S4.1 (top right chemical structure of *benzazole_dervs*). Since those were assumed to behave similar to benzazoles for ozone reactions, *benzazole_dervs* was defined instead of *benzazole* to include isoindole, isobenzofuran, and 2-benzothiophene.

Olefins/Ethynyl As shown in Figure S4.1, any carbon-carbon double bond (C=C) in an aliphatic system is defined as *olefin* with the exception of *olefin_conj*, which contains a conjugation with another functional group. The first exception included in *olefin_conj* is an olefin in a cyclic ring conjugated with an α -keto group, and the second is an olefin conjugated with both α -keto and α -amino groups at both carbons. β -cyclocitral, progesterone, medroxyprogesterone, norethindrone, and levonorgestrel were the basis for the former case and indigotrisulfonic acid was the reference compound for the latter case. All of these compounds were included in the k_{O_3} prediction model as miscellaneous olefins (denoted as *miscolefin* in Table S4.1) using E_{HOMO} (Text S4.4) rather than $E_{NBO,C=C}$, which is typically used for olefins.⁴ The detected products for progesterone during ozonation were according to the Criegee mechanism.⁵ Based on the reported pathways for progesterone, the olefin pathways are assumed to be generally applicable to *olefin_conj* members. ‘*Ethynyl*’ refers to a carbon-carbon triple bond. As there is no corresponding k_{O_3} prediction model developed, a k_{O_3} estimate of $200 \text{ M}^{-1}\text{s}^{-1}$ was assigned based on the second order rate constant for the ethynyl group in 17α -ethynylestradiol (Tables S4.1 and S4.2). A specific pathway reaction group named ethynyl was assigned.

Amines *Amine*, *hydrazine*, *N_sulfenamide*, *amine_conj*, *phenylazo*, *sulfonamide(1st)_prot*, *sulfonamide(1°)_deprot*, *sulfonamide(2°)_prot*, *sulfonamide(2°)_deprot*, and *sulfonamide(3°)* were defined as reactive amine sites (Table S4.1). For sulfonamides, 1°, 2°, and 3° stand for primary, secondary, and tertiary and “prot” or “deprot” stands for protonated or deprotonated, respectively. For all reactive sites to be valid as defined, the nitrogen should not be assigned to any of the excluded moieties in Figure S4.1, which lists amide, imine, nitrile, nitro, nitroso, nitroxide, aminoacrylaldehyde, and quaternary amines. The nitrogen lone-pair electrons in such excluded moieties are considered to be inactivated either by multiple bonds (e.g. imine), by conjugation (amide), by protonation (protonated amine), or by substitution (quaternary ammonium). As a consequence, their reactivities with ozone are very low; e.g., *N*-methylacetamide ($0.6 \text{ M}^{-1}\text{s}^{-1}$),⁶ tetranitromethane ($10 \text{ M}^{-1}\text{s}^{-1}$),⁷ *N*-nitrosodimethylamine ($<0.1 \text{ M}^{-1}\text{s}^{-1}$),⁸ protonated amines ($<0.1 \text{ M}^{-1}\text{s}^{-1}$).⁹ Therefore, those moieties were neglected in this study.

Amine is any primary, secondary, and tertiary amine, which does not belong to any other reactive sites in the amine group (e.g., *hydrazine* and *N-sulfenamide*). For *amine*, a reaction pathway group with the same name, i.e., amine, was assigned (see Text S4.9 for details). *Amine_conj* is defined as a nitrogen conjugated by either a C=N (amidine) or a C=S (thioamide) bond and no k_{O_3} prediction model is available for this group. The reaction pathway group for the amines was also assigned for *amine_conj* based on ozonation of cylindrospermopsin containing a guanidine moiety (Text S4.9 for details).¹⁰ Despite the fact that it is structurally categorized for amine, *phenylazo* was assigned to the *hetero_ar* group for both predictions. Azobenzene, which was the only reference compound for *phenylazo*,¹¹ was reported to be oxidized to azoxybenzene. This is consistent with other heteroaromatic compounds in which the nitrogen is a major reaction center with ozone to form an oxide product (Text-S4.9). Moreover, its $E_{NBO,LP-N}$ did not fit with other amines but its E_{HOMO} correlated well with the other *hetero_ar* group members (Text S4.4). Empirical k_{O_3} estimates were suggested for *sulfonamide(1°)_prot*, *sulfonamide(1°)_deprot*, and *sulfonamide(2°)_deprot*, based on experimental k_{O_3} for the reference compounds as there is no available prediction model (Tables S4.1 and S4.2). *Sulfonamide(2°)_deprot*

was assigned to the reaction pathway group for the amine on ozonation of hydrochlorothiazide (Text S4.9).²

It should be noted that there are several reactive sites for the amine group (e.g., *hydrazine* and *sulfonamide*(3°)) for which no empirical k_{O_3} estimate, no k_{O_3} prediction model group, and no reaction pathway group is assigned. In other words, for these classes of reactive sites no prediction is available in the current prediction platform. Nevertheless, these reactive sites are explicitly presented to provide information as to the limitations of the currently developed prediction system and also to highlight the needs for further investigations. This is the same for many reactive sites for the organosulfur compound group below.

Heteroaromatic compounds *Adenine* and *guanine* were treated separately because the corresponding k_{O_3} prediction models were developed separately (Text S4.4). Note that *guanine* refers not only to guanine but also to other purine bases with a keto group in the 6 position (e.g., hypoxanthine, xanthine, caffeine, and uric acid). A pathway prediction for *guanine* was developed based on ozonation of guanine, guanosine, and its derivatives, e.g., acyclovir.¹² Despite analogous chemical structures, the pathways for *guanine* were not extended for *adenine* in the current study (Text S4.9). Separate k_{O_3} prediction models were developed for *thymine*, *uracil*, and *cytosine* (Text S4.4) and the pathway group, uracil, is applied for *thymine*, *uracil*, and *cytosine* (Text S4.9).

Pyridine and *diazine* are defined as heteroaromatic compounds in which the aromatic nitrogen is not conjugated with the exocyclic nitrogen, oxygen (as a double bond), or sulfur (as a double bond). In contrast, the conjugated forms for pyridine and diazine are defined as *pyridine_conj* and *diazine_conj*, respectively. This classification was based on the observation that non-conjugated pyridine and diazine (2-isopropyl-3-methoxypyrazine) were successfully included in the k_{O_3} prediction model for the heteroaromatic compound group (hetero_ar) using E_{HOMO} , while a conjugated pyrimidine (2,4-diamino-5-methylpyrimidine) was included in the miscellaneous olefin group using E_{HOMO} (Text S4.4). In contrast to *pyridine* and *diazine*, which were defined separately from *pyridine_conj* and *diazine_conj*, no distinction by conjugation was made for *triazine*. Empirical k_{O_3} -values were only available for conjugated 1,3,5-triazines such as atrazine and simazine¹¹ and those were included in the k_{O_3} prediction model for the heteroaromatic compound group (hetero_ar) together with *pyridine* and *diazine* compounds (Text S4.4). It was assumed that other triazines such as 1,2,3-triazine or 1,2,4-triazine would behave similarly. For pyridine, diazine, and triazine, the same reaction pathway group, hetero_ar, was assigned (Text S4.9).

5hetero_ring was defined for 5-membered heteroaromatic compounds containing at least two neighboring aromatic carbons. Such compounds as imidazoles, indoles, isoxazoles were included in the miscellaneous olefin group for the k_{O_3} prediction (Text S4.4). The corresponding reaction pathways were assigned to olefins based on the detection of the Criegee products from ozonation of several *5hetero_ring* members such as tryptophan^{13,14} histidine,¹⁴ and vinylene carbonate.¹⁵

Quinolone and *quinolone_carboxylate* were defined based on the empirical observation for flumequine, enrofloxacin, and ciprofloxacin,³ and k_{O_3} estimates of $1.0 \text{ M}^{-1}\text{s}^{-1}$ and $2 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ were assigned to the two types of reactive sites (see below). Both neutral and anionic flumequine, which

Supporting information for chapter 4

bear the respective *quinolone* and *quinolone_carboxylate*, were included in the k_{O_3} prediction model for the miscellaneous olefin group in the previous study.⁴ However, the miscellaneous olefin group in the current prediction was not used to predict k_{O_3} for *quinolone* and *quinolone_carboxylate* because of the ambiguity of selecting an appropriate orbital for the quinolone moieties of flumequine, enrofloxacin, and ciprofloxacin (not addressed in the present study as it is highly hypothetical).

Quinolone is detected when the 4-pyridone of the quinolone moiety is not substituted by carboxylate (COO^-). For example, the quinolone moiety with a carboxylic group ($COOH$) as in enrofloxacin, ciprofloxacin, or flumequine is detected as *quinolone*. However, it is important to note that for enrofloxacin and ciprofloxacin *cyclic aniline* is also detected for the benzene ring of the quinolone because it has the cyclic nitrogen substituent in a piperazine group (see Text S4.1 for the definition of *cyclic aniline*). A k_{O_3} estimate of $1.0 \text{ M}^{-1}\text{s}^{-1}$ was assigned based on the observed k_{O_3} for the neutral flumequine with no nitrogen substituent.³ As shown in Figure S4.1, when the 4-pyridone of *quinolone* is substituted by carboxylate (COO^-), *quinolone_carboxylate* is detected instead. A k_{O_3} estimate of $2.0 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ was the average of the k_{O_3} -values for deprotonated flumequine ($1.8 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$),³ monoprotonated ciprofloxacin ($7.5 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$),³ and monoprotonated enrofloxacin ($4.6 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$),³ all the compounds of this group with COO^- . In analogy to *quinolone*, *cyclic aniline* will also be detected provided that the benzene ring of the quinolone has a cyclic nitrogen substituent. Note that three reference compounds used herein have a fluorine substituent at the phenyl ring. Although fluorine may have contributed to some extent to the observed k_{O_3} , *quinolone* and *quinolone_carboxylate* were defined without fluorine for a general application for quinolones. For reaction pathways, *quinolone* and *quinolone_carboxylate* are assigned to olefins. Two unfused carbons of 4-pyridone react according to the Criegee mechanism (see Text S4.9 below) based on the observed oxidation products of ciprofloxacin¹⁶.

Organosulfur compounds In contrast to all the reactive sites mentioned above, there is no k_{O_3} prediction model available for organosulfur compounds. Therefore, k_{O_3} estimates are suggested for some reactive sites. However, similar to the amines above, reactive sites (e.g., thioketone, thioamide, and thiophosphoramidate) with neither k_{O_3} estimates nor available pathways were defined in this study. These reactive sites are presented with the intention to stress the necessity for further investigations in future. As explained above for the amines, for a reactive site to be valid, sulfur should not be a part of sulfonamide and sulfonic acid, which are defined as the excluded moieties in Figure S4.1, because those moieties are known to be ozone-resistant.

Text S4.2 Speciation analysis

A speciation analysis is included for a *query compound* undergoing acid-base speciation, thus providing relevant acid-base species and their relative concentrations (p) as a function of pH between 0 and 14. This is a prerequisite for a pH-dependent k_{O_3} prediction, i.e., predicting an apparent k_{O_3} for a reactive site of a *query compound* at a specific pH. A workflow of the speciation analysis is shown in Figure S4.4. A built-in pK_a calculator plugin of Marvin (Chemaxon)¹⁷ generates all acid-base species and predicts associated pK_a -values. For a *query compound*, which has multiple species with the same

net charge (i.e., tautomers), a tautomeric fraction (f), which represents the fraction of each tautomer species, is additionally derived. Both pK_a and f calculation results are presented to the user and remain subject to confirmation or manual modification. As predicted values can be highly uncertain, it is recommended that the user seeks for experimental or theoretical values obtained from high level quantum chemical computations. Based on the confirmed or updated speciation parameters, p_i of a species (i) is recalculated and the corresponding species distribution as a function of pH is derived accordingly (e.g., panel (H) in Figure 4.4).

Text S4.3 Methodology for the development of k_{O_3} prediction models for organic compounds using quantum chemical computations and statistical evaluation of the model performance.

Quantum chemical models for predicting k_{O_3} for organic compounds of differing classes of organic compounds such as aromatic compounds, olefins, amines, etc. were previously developed.⁴ This approach was chosen as a k_{O_3} prediction method in this study not only because of its comprehensive applicability to various organic compounds but also because the k_{O_3} prediction procedure can be implemented with the help of existing quantum chemical computation softwares and other chemoinformatics applications. However, the previously proposed models⁴ were not directly used but modified and adapted for the following reasons. The previously proposed quantum chemical model followed a computational protocol using *ab initio* methods (Hartree Fock (HF) or Density Function Theory (DFT)) for all the computation steps such as geometry optimization, frequency analysis, and single point calculation. All the steps were carried out at the same level of theory.⁴ The prediction platform to be developed in this study was intended to be as responsive and interactive as possible. Although the prediction performance of the previous models was (highly) satisfactory and the protocol was straightforward, performing *ab initio* quantum chemical computations for all aforementioned steps would have hampered this goal due to extended computation times, especially for large molecules (e.g., molecular weight > 300). This is especially important given that the development of this prediction platform was targeted to be implemented by an end-user on a personal computer or as a web-based tool with minimal resources and maintenance. Moreover, as the prediction platform was intended to be available to the public, k_{O_3} predictions with license-free softwares were sought for. ORCA 3.0.3,¹⁸ which is a quantum chemical computation software available free of charge for academic users, was used instead of the commercial software Gaussian,¹⁹ which was applied previously.⁴ However, the commercial natural bond orbital (NBO) software 6.0²⁰ could not be replaced as there is no free substitute currently available. The details on modifications/adaptations of the original computation protocol to fulfill these requirements/specifications are given in Text S4.4.

General information about the establishment of k_{O_3} prediction models and statistical parameters used to evaluate its performance is provided here. A k_{O_3} prediction for organic compounds using quantum chemical computations exploits a linear relationship between the logarithms of empirical k_{O_3} -values for the organic compounds and the corresponding orbital energies. The corresponding linear correlation is established by a least-squares regression with $\log(k_{O_3}) = y_0 + aE$, where y_0 , a , and E are the y -intercept, the slope, and the corresponding orbital energy of a compound, respectively. It has previ-

ously been derived which type of orbital energy most strongly correlates for which class of organic compound (e.g., aromatic compounds correlate with the highest occupied molecular orbital energies (E_{HOMO})).⁴ Here, the same general relationships were used (Table S4.1) and the same set of training compounds were used unless stated otherwise. Statistical evaluations of the developed k_{O_3} prediction models were carried out based on R^2 , the mean unsigned error (MUE) in log units, and the root-mean square error (RMSE) in log units. The previous models were established by excluding outliers from the training compounds for which the predicted k_{O_3} differed by more than a factor of 10, and α , y_0 , and R^2 were reported accordingly. In contrast, those outliers were included in the reported MUEs and RMSEs to better demonstrate the predictive power of the models. In this study, the same compounds and the same outliers were used for the model development and evaluation for a consistent comparison of the model performances unless stated otherwise. A complete list of training compounds is given in Tables S4.4, S4.5, and S4.6, but their experimental k_{O_3} -values can be found elsewhere.⁴

Text S4.4 Evaluation of k_{O_3} prediction models developed with differing computational methods for the selected organic compound groups

The k_{O_3} prediction models were developed for the 14 model groups in the 3rd column in Table S4.1 with seven combinations of computational methods, HF/6-31G//PM3//MMFF94, HF/SV//PM3//MMFF94, HF/SV//PM3//Dreiding, HF/3-21G//PM3//MMFF94, HF/3-21G//PM3//Dreiding, HF/3-21G//MMFF94, and HF/3-21G//Dreiding. All the models were developed based on the same methodology for the model development (Text S4.3) and the complete list of model compounds as given in Tables S4.4, S4.5, and S4.6.

The computational methods presented above, divided by ‘//’, comprise of either two or three consecutive computations to be conducted in a reverse order from the right to the left. ‘MMFF94’ or ‘Dreiding’ indicate a molecular mechanics method to obtain an initial 3D geometry from 1D SMILES using either a Merck molecular force field (MMFF94) method^{21–25} or the Dreiding force field²⁶, respectively. Both methods were used as implemented in Marvin (Chemaxon). The 3D initial geometry of a *query compound* is then subjected to either semi-empirical PM3^{27,28} geometry optimization followed by *ab initio* Hartree-Fock (HF) single point calculations or directly to the HF single point calculations without geometry optimization with the semi-empirical PM3 method. The semi-empirical PM3 method was performed in gas phase. Three different basis sets (Pople 6-31G and 3-21G^{29–33} and Ahlrichs split valence³⁴ (SV)) were explored for single point calculations.

The single point calculations for all seven computational methods were performed by an *ab initio* HF method. A semi-empirical PM3 method showed no meaningful correlations ($R^2 < 0.4$) along with computation failures for many compounds (data not shown). The DFT was not considered as it was previously reported that for some compounds (e.g., benzoate, benzenesulfonate, and benzaldehyde) a molecular orbital located on the aromatic ring obtained by the DFT-B3LYP was not actually the HOMO but the HOMO- n ($n=1,2,3, \dots$), which is a molecular orbital n levels lower than the HOMO.⁴ In contrast, with the HF method, the HOMO consistently appeared on the aromatic ring for all the aromatic model compounds investigated. Because of this consistency, the HF method was preferred. Moreover,

it was reported that the prediction performance of the HF method was superior or comparable to the DFT-B3LYP method.⁴ Therefore, only the HF method was investigated in this study.

Although not presented in a computational method convention for brevity, auxiliary methods such as COSMO (conductor-like screening model), ECP (effective core potential), and NBO (natural bond orbital) were implemented for all single point calculations. For all single point calculations, the COSMO set to water (dielectric constant (ϵ) = 80.4 and refractive index = 1.33) was implemented to take the solvation effect into account. Los Alamos effective core potential (ECP)³⁵⁻³⁷ was applied for iodine for which the 3-21G basis set was unavailable. While E_{HOMO} is obtained from the implementation of the described method above, for compounds containing olefins or amines, the NBO analysis was additionally implemented by the NBO program 6.0²⁰ to derive the corresponding natural bond orbital (NBO) energies ($E_{\text{NBO,C=C}}$ or $E_{\text{NBO,LP-N}}$).

The performances of all the developed k_{O_3} prediction models were evaluated with statistical parameters such as R^2 , MUE, and RMSE (Text S4.3) and are presented in Figures S4.2 and S4.3. The average CPU time, which is the average of the individual CPU times measured for training compounds relative to the reference method (i.e., HF/6-31G//PM3//MMFF94), is additionally presented for individual k_{O_3} prediction model groups. The CPU times for 3D geometry generation was not included as it was completed in only a fraction of the CPU time required for single point calculations. The measurements were conducted on the same hardware (Intel® Core™2 Duo CPU at 3.00 GHz with 4 GB RAM) operated on Ubuntu Linux 14.04.4 LTS for all the computational methods.

Due to an extensive amount of data represented in Figures S4.2 and S4.3, all k_{O_3} prediction model groups are not individually discussed in detail. Rather, general discussions are given for specific model groups with noteworthy behavior such as mono- and di-alkoxybenzene, trialkoxybenzene, guanine, and aniline. Along with the discussions given in this text, a brief summary of the prediction model evaluations is additionally provided in Table S4.7.

For most model groups, the prediction performance evaluated by R^2 , MUE, and RMSE is consistently stable for differing methods, whereas the shortest CPU time (30% to 60% of the reference method) was required for the HF/3-21G method without the PM3 geometry optimization (i.e., HF/3-21G//MMFF94 and HF/3-21//Dreiding). As the MMFF was superior to the Dreiding force field for the mono- and di-alkoxybenzene group (see below), the HF/3-21G//MMFF method was chosen to be the default in this study.

Mono- and di-alkoxybenzene. Comparing the two molecular mechanics methods (i.e., MMFF94 and Dreiding force field), a worse performance was consistently observed with the Dreiding force field compared to the MMFF94 for the mono- and di-alkoxybenzene groups (Figure S4.2) regardless of the basis sets and the geometry optimization. As previously reported,⁴ rotamers can play an important role for alkoxybenzenes as E_{HOMO} significantly fluctuates upon rotation of an alkoxy group. A better correlation was found with alkoxybenzenes mainly consisting of planar alkoxybenzenes for which the alkyl group is aligned to be planar to the benzene ring.⁴ It turned out that initial geometries of more alkoxybenzenes were planar or nearly planar for the MMFF94 than for the Dreiding force field. For example, a planar and perpendicular geometry was consistently obtained for 1,4-dimethoxybenzene

Supporting information for chapter 4

with the MMFF94 and the Dreiding force field, respectively. The exclusion of 1,4-dimethoxybenzene from the correlation improved the performance of all methods with the Dreiding force field (e.g., R^2 increased by more than 0.1) (data not shown). Moreover, further improvements could be achieved by replacing E_{HOMO} of non-planar alkoxybenzenes with E_{HOMO} of their planar conformations, e.g., from $R^2=0.80$ to $R^2=0.84$ for the HF/SV method by using planar alkoxybenzenes (acebutolol, atenolol, 2,4-dichlorophenoxyacetic acid, and propranolol) (data not shown). Although automated geometric modifications could be pursued by programming an algorithm, it was not considered in this study to avoid additional complications. Overall, it was concluded that the MMFF94 has advantages over the Dreiding force field for the mono- and di-alkoxybenzene groups.

Trialkoxybenzenes. For the trialkoxybenzene group, poor correlations ($R^2 < 0.5$) were obtained for all the computational methods except for the HF/3-21G//MMFF94 method for which a partly meaningful correlation ($R^2=0.69$) was obtained. As shown in Figure S4.5a (green triangles), however, a negative slope of the regression line was obtained for HF/3-21G//MMFF94. Similar to the mono- and di-alkoxybenzene groups, this phenomenon is related to the rotamers of trialkoxybenzenes.⁴ However, no further investigation was conducted as automatic geometry modification was not planned to be developed in this study. Moreover, there are only four training compounds available for this model group, which makes it difficult to perform an in-depth evaluation. Nonetheless, the prediction performances evaluated with the MUE (0.26-0.51) and the RMSE (0.31-0.55) were comparable to those of other compound groups with good correlations. This is attributed to the fact that the calibration ranges for k_{O_3} between the minimum and the maximum is only a factor of 30 for trialkoxybenzenes, while other model groups span over several orders of magnitudes (e.g., 7 orders of magnitude for phenol). Therefore, its prediction may be acceptable as long as the derived orbital energy lays in the calibration range (e.g., E_{HOMO} of $-0.33649 \sim -0.32043$ hartree for the HF/3-21G//MMFF94), while it may completely fail when extrapolated.

Guanine. For guanine, an aggravated performance was observed for the method involving geometry optimization with a Dreiding force field-derived geometry, i.e., HF/SV//PM3//Dreiding ($R^2=0.62$) and HF/3-21G//PM3//Dreiding ($R^2=0.65$) compared to the other methods ($R^2=0.94-0.98$). This was because 5'-deoxyguanylic acid was an outlier in the Dreiding force field-based method. During the geometry optimization using the initial geometry obtained by the Dreiding force field, it was observed that an intramolecular hydrogen transfer from an amino group to a phosphate group occurred, which led to a significantly higher E_{HOMO} (data not shown). Because of this phenomenon, the MMFF94 method is recommended for guanines for the geometry optimization.

Aniline. It should be noted that in contrast to the other model groups for which k_{O_3} prediction models were developed with the same training compounds as used in the previous study,⁴ a k_{O_3} prediction model for anilines has been established in this study with different sets of anilines. This was because k_{O_3} -values for several anilines used for the previous model were updated with recently reported empirical k_{O_3} -values.³⁸ Compared to the previous study, moreover, k_{O_3} -values for anilines such as *p*-methylaniline, *p*-methoxyaniline, *p*-phenylenediamine, *o*-hydroxyaniline, *p*-hydroxyaniline, and *p*-aminobenzoic acid³⁸ were additionally included in the model. k_{O_3} for the unsubstituted aniline was reported to be $1.3 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$,³⁸ which is almost an order of magnitude lower than the k_{O_3} -value used

for the previous model, i.e., $1.4 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$.³⁹ We have adapted the currently reported k_{O_3} in this study as the empirical determination has been appropriately conducted using a well-established competitor (3-buten-2-ol) and at neutral to alkaline pH in which the ozone decomposition from the reaction with aminyl radicals can be reduced. Moreover, the determined value was theoretically corroborated by an excellent correlation between k_{O_3} for differing aromatic compounds including unsubstituted aniline and the corresponding Gibbs free energy (ΔG^0) for the formation of an ozone adduct in the para position.³⁸ As the previous k_{O_3} -values for many other anilines were reported relative to the k_{O_3} for unsubstituted aniline,³⁹ k_{O_3} -values for such anilines were corrected accordingly and used in the current model development. Moreover, a significant correction was made for *N,N*-dimethylaniline that while the previous k_{O_3} was $2.0 \times 10^9 \text{ M}^{-1}\text{s}^{-1}$,⁴⁰ the k_{O_3} that has been recently reported is $1.6 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$.³⁸ The k_{O_3} -values for 22 anilines used for the model development in this study are compiled in Table S4.8 where they are also compared to the previous values.

With 22 anilines and their up-to-date k_{O_3} values, overall inferior correlations ($R^2=0.43\text{-}0.64$) to the original model ($R^2=0.85$ for the HF/6-31G method)⁴ were obtained for the current model. Nonetheless, the prediction performance was better than with the original model, MUEs=0.24-0.30 and RMSEs=0.30-0.37 throughout all seven computational methods of the current model and MUE=0.34, RMSE=0.43 for the previous model, respectively. This is related to the fact that the current model was developed with the corrected k_{O_3} -values, which were in a narrower range than the original model. While the k_{O_3} range of the 22 anilines used for the current model only spans a factor of 50, it was ~ 4 orders of magnitudes for the previous model (Table S4.8).

It is worthwhile to note that the previous model for anilines has been developed using planar anilines which have two substituents on the amino group aligned to be planar to the phenyl ring because its correlation was better than that with pyramidal anilines (see reference⁴ for more details). However, a somewhat worse performance was obtained using planar anilines with the new set of 22 anilines (data not shown). Therefore, it is concluded that the previously observed better correlation with planar anilines was rather fortuitous and it is recommended to use the current model rather than the previous model.

Other heteroaromatic compounds (*hetero_ar*). Although correlation to the HOMO was shown to be somewhat successful, no prediction model was previously proposed for other heteroaromatic compounds because the model seemed to be rather premature to be established (see Text S2.5 in the Supporting Information for Chapter 2). In this study, nonetheless, a prediction model was developed to provide kinetic information for these chemical moieties.

Text S4.5 Computational protocol for k_{O_3} predictions

Based on the evaluation of the seven different combinations of computational methods in Text S4.4 above, the HF/3-21G//MMFF94 method was chosen as default in this study because it performed similarly or better compared to the other computational methods tested in this study, while the least CPU time was required (for more details see Text S4.4 above). The individual models are presented

Supporting information for chapter 4

in Figure S4.5 and their prediction performances as well as the parameters for linear regression equations are summarized in Table S4.3 in comparison with the original models. Note that this computational protocol does not carry out geometry optimization and frequency analysis. Nevertheless, the models' performances turned out to be acceptable although it was inferior to the original models (Table S4.3). This indicates that orbital energies obtained from molecular mechanics-level geometries of training compounds are similar to those of the optimized local minimum geometries from the original computational protocol involving geometry optimization and frequency analysis. Moreover, fortuitous cancellation of errors may have played a role as only relative orbital energies are relevant in the models while absolute orbital energies may still be biased. However, one should keep in mind that a prediction may fail for some compounds due to the use of an approximate geometry and further investigations are needed to be able to justify the general application of molecular mechanics-based geometries.

The CPU times required to complete single point calculations for the 294 individual compounds (10 outliers included) used for model development were measured and are plotted as a function of the number of atoms (atom count) in Figure S4.6. A good relationship between the CPU time and the atom count was obtained when fitted by a quadratic regression. While less than 700 seconds were required for most of the training compounds with less than about 60 atoms, CPU times in the order of several thousand seconds were needed to complete single point calculations for macromolecules such as azithromycin, roxithromycin, tylosin, and microcystin-LR with 124 – 145 atoms.

Text S4.6 Hierarchy for the assignment of an ambivalent aromatic compound

Compound groups such as olefins, amines, guanine, etc. are individually unique, therefore, the assignment of reactive sites can be conducted independently. However, reactive sites for aromatic compounds may overlap, e.g., aminophenol contains phenol and aniline in the same benzene ring. As such ambivalent compounds were not included in the model development, a hierarchy between aromatic compound models was established (arrow in top-left part of Figure S4.1), based on the relative positions of individual k_{O_3} prediction models in Figure S4.5a. As phenolate has the highest k_{O_3} -values, it is assigned to have the highest priority over all the other model groups. As shown in Figure S4.5a, a hierarchy between trialkoxybenzene, aniline, and phenol is ambiguous because most of the data points are in close proximity. It is assumed in this study that trialkoxybenzenes should be given a higher hierarchy than aniline, followed by phenol. Another approach to resolve this issue would be to choose the model group with the highest predicted k_{O_3} among the model groups. However, this approach has not been investigated yet. Therefore, further improvements/modifications can be expected in the future. Mono- and dialkoxybenzene and benzene are the lowest in the hierarchy, with the alkoxybenzene given higher hierarchy than the benzene as they have a higher slope (86.74 in Table S4.3) of the correlation than that of benzene (46.68 in Table S4.3) (Figure S4.5a).

Text S4.7 Assignment of E_{HOMO} to the corresponding reactive sites of a compound

As shown in Table S4.1, there are many reactive sites for which k_{O_3} -values are predicted using $E_{\text{HOMO-n}}$. There can be compounds with multiple reactive sites, which necessitate the assignment of multiple $E_{\text{HOMO-n}}$ -values to the corresponding reactive site. In this section, the strategy to achieve the correct assignment of $E_{\text{HOMO-n}}$ -values to individual reactive sites is discussed.

The workflow for the assignment is shown in Figure S4.7. Quantum chemical computation outputs, in which the molecular orbital coefficients for HOMO-n are printed out, are a prerequisite for this process (see below). The first step is to identify reactive sites of a query compound for which k_{O_3} is to be predicted using $E_{\text{HOMO-n}}$. $\text{SS}_{\text{MOcoef-n}(i)}$ for individual reactive sites are subsequently estimated. $\text{SS}_{\text{MOcoef-n}(i)}$, which is an empirical parameter introduced in this study, is the sum of squares of molecular orbital coefficients (c_i) of basis functions (ϕ_i) for the atoms of reactive sites i for HOMO-n. c_i is derived from the best wavefunction (ψ) of the molecular orbital (e.g., HOMO-n) obtained by the variational method where molecular orbitals are constructed by a linear combination of atomic orbitals (LCAO). Therefore, $\text{SS}_{\text{MOcoef}}$ can be used as an approximate index for the electron population of the HOMO-n on the atoms of a reactive site. In the following, threshold values of $\text{SS}_{\text{MOcoef}}$ assigned to all the relevant reactive sites are presented.

Six aromatic carbons in a benzene ring were considered as the atoms for aromatic compounds for which $\text{SS}_{\text{MOcoef-n}(i)}$ are calculated. For example, three aromatic compounds, i.e., benzene, phenol, and aniline, are presented in Figure S4.8a. $\text{SS}_{\text{MOcoef}}$ calculated for the part of the HOMO localized on the carbons of benzene, phenol, and aniline was 0.53, 0.49, and 0.43, respectively. Note that, as defined above, for phenol and aniline the contribution of an orbital to the exocyclic oxygen or nitrogen was not accounted for in the estimated $\text{SS}_{\text{MOcoef}}$. In this manner, $\text{SS}_{\text{MOcoef}}$ -values for all the aromatic compounds used to develop the k_{O_3} prediction models were calculated and were found to range from 0.24 for bisphenol A dianion to 0.53 for methylbenzoate. As shown in Figure S4.8a, the two phenolates of the bisphenol A dianion have almost identical $\text{SS}_{\text{MOcoef}}$ and their summation gives ~ 0.5 which is similar to the $\text{SS}_{\text{MOcoef}}$ of benzene and phenol. This indicates that the electron density of the HOMO for the bisphenol A dianion is delocalized equally over the two phenolates. As the bisphenol A dianion was successfully included in the k_{O_3} prediction model for the phenol group, the HOMO-n with a $\text{SS}_{\text{MOcoef}}$ of 0.25 should be considered to be acceptable for aromatic compounds. Therefore, a minimum $\text{SS}_{\text{MOcoef}}$ threshold for an aromatic ring was set to 0.22 with about 10% of a margin applied to 0.24 for the current prediction platform. It turned out that all the heteroaromatic model compounds have $\text{SS}_{\text{MOcoef}}$ -values higher than 0.22. Therefore, the same threshold $\text{SS}_{\text{MOcoef}}$ -value was set to reactive sites belonging to heteroaromatic compounds. Consequently, reactive sites with a minimum $\text{SS}_{\text{MOcoef}}$ threshold of 0.22 include *phenol*, *aniline*, *12alkoxybenzene*, *3alkoxybenzene*, *benzene*, *benzazole_dervs*, *adenine*, *guanine*, *cytosine*, *thymine*, *uracil*, *pyridine*, *diazine*, *triazine*, *pyridine_conj*, and *diazine_conj*.

For *5hetero_ring*, two adjacent carbons (e.g., imidazole) or three consecutive carbons (e.g., isoxazole) were used to estimate $\text{SS}_{\text{MOcoef}}$ (atoms with black circles Figure S8b are the carbons used to estimate $\text{SS}_{\text{MOcoef}}$). This is based on the observation that the ozone attack occurs not at the hetero atoms but at the two adjacent carbons (see *5hetero_ring* in Text S4.1). $\text{SS}_{\text{MOcoef}}$ -values of 0.32-0.35 were observed.

Supporting information for chapter 4

Therefore, the threshold SS_{MOcoef} was set to 0.3 for *5hetero_ring*. For *olefin_conj* which has two adjacent carbons (black circles in Figure S4.8c) as the estimation center atoms for the same reason as for *5hetero_ring*, a rather low threshold SS_{MOcoef} of 0.15 was set based on the indigotrisulfonic acid (Figure S4.8c).

Since the threshold values for SS_{MOcoef} were defined based on a limited number of model compounds, the assignment of E_{HOMO} may be wrong, leading to an error in a predicted k_{O_3} . Therefore, the assigned threshold values need to be modified when necessary based on new findings or identified improvements for a compound of interest. A visual examination of HOMO-n to check whether E_{HOMO} is reasonably assigned to the corresponding reactive site is also recommended. In addition, the current approach using SS_{MOcoef} has the intrinsic weakness mainly associated with the delocalization of HOMO-n, making it difficult to assign a responsible orbital to the specific reactive site. Therefore, the development of a better approach to resolve this rather ambiguous assignment is desired in future.

Text S4.8 Interpretation of predicted k_{O_3} -values for reactive sites of a query compound and their relationships with experimental k_{O_3} -values

Based on the computational protocol proposed in Text S4.5, k_{O_3} for identified reactive sites of a *query compound* will be obtained. For some reactive sites, for which a k_{O_3} prediction model is unavailable, empirical k_{O_3} estimates when available are assigned instead (Table S4.1). The obtained k_{O_3} -values can provide various useful informations. Note that k_{O_3} prediction in the current prediction platform is carried out for only one conformer, assuming that the investigated conformer is the dominant species for the reaction with ozone, thus representative of the overall reactivity. Therefore, the treatment of multiple conformers as shown in Text S4.10 for cetirizine needs to be conducted externally. For a non-ionizing *query compound* (i) with j reactive sites ($j=1,2,3,\dots,n$), the ‘rate constant prediction’ process provides the same number of predicted $k_{O_3,j}$ as the number of reactive sites. Comparison of $k_{O_3,j}$ between differing reactive sites allows to identify which reactive site dominantly reacts with ozone. The summation of $k_{O_3,j}$ for all the reactive sites (j) gives rise to a species (i)-specific k_{O_3} , i.e., $k_{O_3,i}$ as follows:

$$k_{O_3,i} = \sum k_{O_3,j} \text{ over all reactive sites } j \approx k_{O_3,exp} \quad (S4.1)$$

In principle, $k_{O_3,i}$ corresponds to an experimental k_{O_3} ($k_{O_3,exp}$) for a non-ionizing compound.

For an ionizing *query compound* having i acid-base microspecies, k_{O_3} predictions are performed for individual microspecies. Therefore, $i \times j$ k_{O_3} -values ($k_{O_3,i,j}$) are derived in total. Based on p_i obtained from the speciation analysis (see Text S4.2 above), an apparent reactive site(j)-specific k_{O_3} ($k_{O_3,site(j)-app(pH)}$) or an apparent species(i)-specific k_{O_3} ($k_{O_3,species(i)-app(pH)}$) at a specific pH can be derived by Eqs. S4.2 and S4.3 below.

$$k_{O_3,site(j)-app(pH)} = \sum k_{O_3,i,j} \times p_{i(pH)} \text{ over all microspecies } i \quad (S4.2)$$

$$k_{O_3,species(i)-app(pH)} = \sum k_{O_3,i,j} \times p_{i(pH)} \text{ over all reactive sites } j \quad (S4.3)$$

In the same manner as for a non-ionizing compound above, a comparison of $k_{\text{O}_3, \text{site}(j)\text{-app}(\text{pH})}$ between different reactive sites allows to identify which reactive site dominantly reacts with ozone at a specific pH and a comparison of $k_{\text{O}_3, \text{species}(i)\text{-app}(\text{pH})}$ between different species allows to identify which species dominantly reacts with ozone at a specific pH, respectively.

The summation of $k_{\text{O}_3, \text{site}(j)\text{-app}(\text{pH})}$ over all the reactive sites (j) or of $k_{\text{O}_3, \text{site}(i)\text{-app}(\text{pH})}$ over all the species (i) gives rise to an apparent k_{O_3} ($k_{\text{O}_3, \text{app}(\text{pH})}$) for the query compound at a specific pH, which in principle corresponds to an experimental apparent k_{O_3} ($k_{\text{O}_3, \text{exp-app}(\text{pH})}$) at a specific pH.

$$\begin{aligned} k_{\text{O}_3, \text{app}(\text{pH})} &= \sum k_{\text{O}_3, \text{site}(j)\text{-app}(\text{pH})} \text{ over all reactive sites } j \\ &= \sum k_{\text{O}_3, \text{species}(i)\text{-app}(\text{pH})} \text{ over all microspecies } i \approx k_{\text{O}_3, \text{exp-app}(\text{pH})} \end{aligned} \quad (\text{S4.4})$$

Text S4.9 Selected reaction pathways from literature for pathway predictions

As explained in the main manuscript, reaction rules for pathway predictions were defined based on published reaction pathways in literature. In this section, the selected reaction pathways are discussed in detail divided into different reaction pathway groups, which are shown in the 4th column of Table S4.1. In the given Schemes for reaction pathways, various logical operators such as \equiv (is equal to), \neq (is not equal to), $\&\&$ (and), and \parallel (or) are presented in ‘if ()’ statements when necessary to provide specificity for certain reaction pathways. Reaction pathways will be appropriately modified/updated in future when new information is available.

Phenol (Schemes S4.1-S4.3). The discussion of the mechanisms for the reaction of ozone with phenol is divided into three categories; *ortho*, *para*, and *radical*. Ozone reactions initiated by ozone attack at ortho and para positions are discussed in ‘*ortho*’ and ‘*para*’ groups, and those involving radicals generated by a one-electron transfer between phenol and ozone are dealt with in the ‘*radical*’ group.

Phenol-ortho (Scheme S4.1): The formation of hydroxylated products is the most commonly reported pathway in many studies for aromatic compounds in general.^{1,38,41-49} Hydroxylation of phenol at the ortho position gives rise to the formation of catechol provided that the ortho position is unsubstituted, i.e., substituted by hydrogen. The reaction mechanism was proposed as follows.⁴⁸ Initially, ozone attacks at the ortho position to form an ozone adduct (1 in Scheme S4.1) followed by a singlet oxygen release (3). It undergoes a rearrangement giving rise to a hydroxylated benzene provided that a substituent of the carbon to the ozone adduct was hydrogen (4-5). A substituent release (6) is expected when the substituent is a halide or a nitro group. This is based on the study of ozonation of *p*-chloro- or *p*-nitro-aniline,³⁸ in which such mechanisms were proposed to explain appreciable detections of chloride or nitrite and nitrate. Although an aryl carbocation was proposed therein following the substituent release without a further proposed pathway, it may be that the aryl carbocation can be stabilized by forming a carbonyl bond between oxygen of a hydroxyl group and the ipso carbon, which gives rise to 1,2-benzoquinone. Due to its unstability, 1,2-benzoquinone was not detected.⁴⁸ Various pathways of 1,2-benzoquinone seem to be possible; reaction with phenol to form a biphenyl dimer and the ensuing polymerization,⁵⁰ reaction with primary amines and the ensuing polymerization,⁵¹ and

Supporting information for chapter 4

reaction with thiols.⁵² In the context of water treatment, 1,2-benzoquinone would undergo reactions with various moieties including the aforementioned ones present in dissolved organic matters (DOM), thus its detection is highly unlikely. An ozone adduct can transform into a primary ozonide with either an ipso carbon (7) or a meta carbon (12), both of which follow a typical Criegee mechanism (7-14) leading to a ring cleavage (see the section for 'olefin' below for more details). Similar to the hydroxylation, a ring cleavage has been observed for many aromatic compounds.^{41,43,44,48,53-56} The resulting product undergoes a post-Criegee mechanism which gives rise to differing products depending on the substituents such as halide or carboxyl groups, amongst which hydrogen peroxide elimination is the most commonly observed (see below for details about the post-Criegee mechanism). Exceptionally, *cis,cis*-muconic acid formed from water elimination rather than hydrogen peroxide elimination was detected in ozonation of unsubstituted phenol (11).⁴⁸ An ozone adduct may undergo a homolytic cleavage of an oxygen-oxygen bond, giving rise to a superoxide ion and an oxyl radical species (15). Further reaction pathways of the oxyl radical are unknown and the superoxide ion is a precursor to hydroxyl radicals. An oxyl radical can in principle be formed from the dissociation of any ozone adduct. However, this reaction pathway for the oxyl radical is not explicitly presented hereafter for brevity.

Phenol-para (Scheme S4.2): The analogous reaction mechanisms such as hydroxylation, substituent release, and ring cleavage were proposed for the para position, resulting in the formation of the corresponding products. The formation of 1,4-benzoquinone was additionally proposed via reactions (23-30)

Phenol-radical (Scheme S4.3): A direct electron transfer from phenol to ozone gives rise to a phenol radical cation and an ozonide radical.⁴⁸ As the phenol radical cation has a pK_a of -2,⁵⁷ it quickly deprotonates to a phenoxyl radical in aqueous solution (34). An ozonide radical is a precursor to the hydroxyl radical.⁵⁸ The formation of a phenol radical cation can be also achieved via a dissociation of an ozone adduct.¹¹ As there is no mechanistic difference between the two routes in terms of concomitant reaction pathways, no distinction was made herein. Phenoxyl radicals can concurrently undergo various reactions. Superoxide ion described above can combine with the phenoxyl radical either in the para or in the ortho position, leading to the formation of the corresponding quinone (35-37 or 38-40). Quinones can also be formed either from disproportionation of the respective semiquinone radicals (41 or 43) or from two consecutive one-electron oxidations of hydroquinone (42) or catechol (44) by ozone. Alternatively, phenoxyl radicals may combine either at the ortho or at the para position to form various dimers (45-54). This is only important for higher concentrations.

Aniline (Schemes S4.4-S4.7). Reaction mechanisms for the reactions of ozone with aniline are discussed by dividing them into four categories; *ortho*, *para*, *nitrogen*, and *radical*.

Aniline-ortho (Scheme S4.4): The same reaction mechanisms such as hydroxylation, substituent release, and ring cleavage, which were applied to the phenol, are also considered for the reaction of aniline after attack of ozone at the ortho position. It is assumed herein that a benzoquinone imine is formed following deprotonation (60) of the product resulting from the substituent release. The 1,2-benzoquinone imine is prone to hydrolysis to yield 1,2-benzoquinone and the corresponding amine.

Similar to above, a ring opening takes place via the Criegee mechanism. In addition, picolinic acid was reported as one of the products resulting from an intramolecular condensation reaction (68-70) of the Criegee product.^{11,38}

Aniline-para (Scheme S4.5): The same reaction mechanisms such as hydroxylation, substituent release, and ring cleavage, which were applied for the ortho position of aniline (Scheme S4.4), are also considered for the reaction of aniline attacked by ozone at the para position.

Aniline-nitrogen (Scheme S4.6): An exocyclic nitrogen of aniline is available for ozone attack. Several previous studies reported formation of products involving oxidation of the nitrogen such as nitrosobenzene, nitrobenzene, and azobenzene.^{38,59,60} It is worthwhile to mention that it was reported that the yields of the products are too low to consider a exocyclic nitrogen as a major site of ozone attack.³⁸ An ozone adduct to the nitrogen is initially formed (81) which is followed by various pathways. In the case of a primary amino group, an ensuing hydrogen transfer gives rise to a hydrotrioxide which decays either via hydrogen peroxide elimination to form nitrosobenzene (83, 84) or via a peroxy radical (HO_2^\cdot) to form a nitroxyl radical. Despite the fact that a second order rate constant for the reaction between nitrosobenzene and azobenzene is very slow ($k=8.4\times 10^{-6} \text{ M}^{-1} \text{ s}^{-1}$)⁶¹ nitrosobenzene was reported to degrade over time. This indicates that there are hitherto unknown reaction pathways leading to the disappearance of nitrosobenzene. The nitroxyl radical can either disproportionate into nitrobenzene and aniline (87) or react with superoxide in the solvent cage to form nitrobenzene (89, 90) for which superoxide stems from reaction (88).³⁸ An oxygen transfer (91) is well-established for aliphatic amines in terms of the singlet oxygen and the ensuing products, i.e., *N*-oxide.^{2,62-64} For aromatic amines, however, 7% of the singlet oxygen yield was reported for *N,N*-dimethylaniline without the detection of a *N*-oxide product.⁶³ Therefore, this pathway as well as an ensuing hydrogen shift to form a hydroxyl amino group (92) remains to be tentative.

Aniline-radical (Scheme S4.7): An electron transfer between ozone and aniline (or the dissociation of an ozone adduct) leads to the formation of an aniline radical cation and an ozonide radical (93). Depending on substituents, primary aniline radical cations have a pK_a range between 4 – 12,⁴⁷ depending on which it deprotonates to an anilino radical provided that the amino group is primary or secondary (94). The resulting anilino radical can be further oxidized by ozone to form an ozone adduct radical which eventually leads to the formation of a 1,2-benzoquinone imine (95-99). 2-Haloaniline and 2-aminophenol may give rise to the same product via a cage reaction between the anilino radical and the ozonide radical (102)¹ or via another one-electron oxidation by ozone (103). These reaction pathways (95-103) apply the same for the para position. Alternatively, an anilino radical can be further oxidized by ozone to form a nitroxyl radical (104), which in turn participates in a radical disproportionation (87) or a nitrosation (89) as described above. There are other pathways for the aniline radical cation, which are in competition with a deprotonation of an amino group. For secondary or tertiary amino groups containing an alkyl group, a deprotonation can occur at the α -carbon to form a carbon-centered radical. The ensuing reactions are the same as those for aliphatic amines, details of which are presented below. Similarly, for anilines with a substituent such as an alkyl or an alkoxy group, a deprotonation occurs at the α -carbon of the substituent instead of the amino group (107 and 111), which via superoxide elimination gives rise to a dealkylated aniline and an aminophenol as well as their corresponding al-

Supporting information for chapter 4

dehyde products (108-110 and 112-114), respectively.³⁸ This mechanism was proposed to explain the formation of formaldehyde from ozonation of different anilines.³⁸ Moreover, an aniline radical cation can dimerize into a benzidine (105,106).⁶⁵ However, the formation of a benzidine was reported in a pulse radiolysis study in absence of dissolved oxygen as it was saturated by N₂O and an aniline radical cation is generated by another oxidant, i.e., N₃[•]. Indeed, it was not detected from ozonation of different anilines.³⁸ Therefore, this pathway is likely to be of minor importance.

Alkoxybenzene (Schemes S4.8-S4.9). Reaction mechanisms presented herein for alkoxybenzene are relevant to the reactive sites named *1*alkoxybenzene and *3*alkoxybenzene in Table S4.1.

Alkoxybenzene-ortho (Scheme S4.8): The same reaction mechanisms such as hydroxylation (116-118), substituent release (119), and ring cleavage (120-125), which were discussed above for phenol and aniline, are also considered. In addition to those, an alcohol elimination may occur.⁶⁶ For an alkoxybenzene with a hydrogen in ortho position, a proton transfer from the ortho position to the terminal oxygen of the ozone adduct rearomatizes while forming a very unstable hydrotrioxide (126), which subsequently decays into an oxyl radical and a peroxy radical (127). The two intermediates may undergo a cage reaction (128) provided that a substituent in para position with respect to the oxyl radical is an alkoxy group. Subsequently, alcohol and hydrogen peroxide elimination occur, giving rise to a 1,4-benzoquinone (129). Alternatively, an alkoxy substituent at the ipso position can be eliminated as alcohol, yielding 2-hydroxyphenol (130-135). A carbocation from reaction (115) may react with water to an intermediate with a hydroxyl group at the ipso position (130). This intermediate subsequently undergoes either an alcohol elimination (131) followed by an ensuing singlet oxygen elimination (132,133) or a singlet oxygen elimination (134) followed by an alcohol elimination (135), both of which give rise to 2-hydroxyphenol as a final product (135).

Alkoxybenzene-para (Scheme S4.9): The same reaction mechanisms as for alkoxybenzene-ortho such as hydroxylation (136-139), substituent release (140), and ring cleavage (141-143) are considered.

Alkoxybenzene-radical (Scheme S4.9): It was assumed that a superoxide elimination mechanism may occur, which initiates upon deprotonation of a exocyclic α -carbon of a substituent such as an alkyl or alkoxy group as described above for aniline-radical (107-114). This would lead to a change of an alkoxy group into a hydroxyl group and the formation of the corresponding aldehyde product.

Benzazole_dervs (Scheme S4.10): *1H*-1,2,3-triazole-4,5-dicarbaldehyde was reported as a major oxidation product during ozonation of *1H*-benzotriazole.⁵⁴ This can be formed via typical reaction pathways for aromatic compounds; ring cleavage of a benzene ring followed by further oxidation of the product by ozone via the Criegee mechanism. Therefore, the same reaction mechanisms such as hydroxylation, substituent release, and ring cleavage, which were discussed above, are also considered for both positions (6 and 7) of a benzene ring, while preventing an ozonide formation with a fused carbon, i.e., ozonide formation between C7 and C7a or C4 and C3a. When a nitrogen or a sulfur is present in the position 1 in a 5-membered ring, an oxygen transfer may additionally occur to form the corresponding oxide product. A subsequent proton shift may be possible for the resulting *N*-oxide product to form a hydroxylated product (148). Quantum chemical computations suggest that a *N*-oxide formation is thermodynamically less favored than formation of an ozone adduct on a benzene

ring (Lutze *et al.*, private communication). Radical reactions were not considered because only little information is available.

Benzene (No scheme provided): The same reaction mechanisms such as hydroxylation, substituent release, and ring cleavage, which were discussed above, are also considered for benzenes. It seems that general rules of regioselectivity by substituents (e.g., ortho-para directors or meta directors) in electrophilic aromatic substitutions can also be applicable to the reaction of ozone with substituted benzenes other than the aromatic compound classes previously described (i.e., phenol, aniline, and alkoxybenzene). However, it was assumed in the current prediction platform that ozone may attack any carbon of the benzene ring because their transformation products from the reaction with ozone would be of minor importance compared to other compound groups. This is because dominant transformation products will be formed from the reaction with hydroxyl radicals due to their low reactivity with ozone. Nonetheless, it is planned to update this aspect in future. Radical reactions were not considered.

Olefins (Scheme S4.11): As shown in Table S4.1, the reaction pathway group ‘olefin’ is applied to reactive sites such as *olefin*, *olefin_conj*, *quinolone*, and *quinolone_carboxylate*. The reaction pathways for olefins with ozone are well established by the so-called Criegee mechanism.⁶⁷ A carbon-carbon double bond (C=C) reacts with ozone to form a primary ozonide (151) via a primary zwitterion (150). It subsequently decays into a secondary zwitterion (152). The secondary zwitterion via hydrolysis decomposes into a α -hydroxyhydroperoxide and a carbonyl compound (153). Depending on substituents of these Criegee products, various ensuing pathways, which are named as a post-Criegee mechanism, are possible.^{68,69} A α -hydroxyhydroperoxide gives rise to a carbonyl compound and hydrogen peroxide provided that its two substituents are not halogen (154). When it is substituted by a halogen, a peracid and a halogen ion are generated (159). Depending on a remaining substituent of the peracid, various decays can follow (160-162). Moreover, decarboxylations can occur for the hydroxyhydroperoxide substituted by a carboxyl group (163, 164). The ensuing product is then subject to another decomposition depending on the presence of substituents such as a halogen and a carboxyl group (155-158). The secondary zwitterion formed in reaction (152) may also undergo decarboxylation provided that the carbon with cationic oxygen is substituted by a carboxyl group, giving rise to a ketone group (172). The additional presence of a carboxyl group to the opposite side leads the resulting product to decarboxylation yielding another ketone group (173). An identical product can be also formed via hydrolysis (174) in the absence of a carboxyl group. The final product is subject to further reactions (155-158). Products, which would result from the partial cleavage of a C=C bond, have been reported for a few olefins such as tetramethylethene, 1,1-dichloroethene, and trichloroethene.⁶⁸ In competition with a ring-closing reaction (151), a primary zwitterion may react with water (165) or undergo a chlorine shift (169). An oxygen is subsequently released (166, 167) and the halogen substituent when present is also released (167, 168). However, products from the partial cleavage were not detected for most olefinic compounds.⁶⁸⁻⁷³ Therefore, this mechanism is considered to be of minor importance.

Amine (Scheme S4.12): As shown in Table S4.1, the amine pathway group is assigned for four reactive sites; *amine*, *amine_conj*, *phenylazo* and *sulfonamide(2°)_deprot*. As reaction pathways for

amine_conj and *sulfonamide(2°)_deprot* belong to the pathways for *amine* under specific conditions, the overall reaction pathways for *amine* are first described (Scheme S4.12), followed by *amine_conj* and *sulfonamide(2°)_deprot*. As described above, the *phenylazo* is exceptionally defined as an amine for a structurally consistent categorization, even though, its pathways are implemented in *hetero_ar*. The reaction of amine with ozone is initiated by an electrophilic attack of ozone at the nitrogen lone-pair electrons to form an ozone adduct (175). A subsequent loss of oxygen to *N*-oxide (176) is commonly observed as a stable product for tertiary amines.^{2,16,62,64,74–76} The resulting *N*-oxide transforms into a hydroxyl group (177) for secondary amines^{2,41,53,77} or potentially primary amines. A proton transfer can occur for primary or secondary amines to form a hydrotrioxide (178), which subsequently decomposes into a nitroxyl radical and a peroxy radical (179). The nitroxyl radical can also be formed via a dissociation of the ozone adduct (180) or via oxygen transfer for an aminyl radical by ozone (187) formed from an electron-transfer-involving reaction pathway (see below for details). The nitroxyl radical and the aminyl radical can play a role to consume additional ozone via chain reactions (187 and 188).¹ This indicates a high ozone consumption compared to the degradation of an amine, therefore, supports a poor mass balance between ozone and amine frequently reported in relevant studies.^{1,2,77} The nitroxyl radical can be further oxidized by ozone via electron transfer leading to an oxoazanium formation (181)⁷⁷ or disproportionate into two different pairs of products (182, 183).^{11,78} Similar to aromatic compounds such as phenol and aniline described above, electron transfer-initiated reactions are of consideration for amine, which results in the formation of an amine radical cation. As described above, a secondary aniline nitrogen was defined as an amine for its k_{O_3} and pathway prediction. The resulting radical cation species can have common reaction pathways for the aniline radical cation starting from reaction (94) in Scheme S4.7 as it has the same reaction center. Depending on pH of an aqueous system and the pK_a value of an amine radical cation, it can deprotonate to an aminyl radical (186). A pK_a range of 5.3 – 6.8 was reported for the amine radical cations of dimethylamine, diethylamine, pyrrolidine, and piperidine.⁷⁹ As already described above, an oxygen transfer by ozone (187) transforms an aminyl radical to a nitroso radical followed by various ensuing pathways (181–183). In competition with deprotonation on the nitrogen, deprotonation can occur at the α -carbon to the nitrogen (189), yielding a carbon-centered radical species. Provided that the α -carbon is substituted by a carboxyl substituent, decarboxylation occurs, giving rise to the same carbon-centered radical. Most carbon-centered radicals are known to react with oxygen with a close to diffusion-controlled rate (k in the order of $10^9 \text{ M}^{-1}\text{s}^{-1}$) to form an α -amino-peroxy radical (191).⁸⁰ The α -amino-peroxy radical undergoes an unimolecular superoxide ion ($O_2^{\cdot-}$)-elimination to yield an iminium ion. It subsequently hydrolyzes into a carbonyl compound and an amine provided that a substituent of the amine is not a hydroxyl group.^{81,82} If a hydroxyl group is present, it is assumed to deprotonate to nitrene which then hydrolyzes into the same products as above.⁷⁸ Alternatively, peroxy radicals can undergo a bimolecular recombination to form a tetroxide intermediate for which there are several possible reaction pathways.⁸³ The tetroxide decomposes into a carbonyl product, a product with a hydroxyl group, and oxygen (197), which is also known as the Russell mechanism.⁸⁴ Another possible reaction gives rise to two carbonyl products and hydrogen peroxide, which is also known as the Bennett mechanism (198).⁸⁵ These two mechanisms necessitate the presence of hydrogen for the carbon bound to a tetraoxide, i.e., not possible for a tertiary peroxy radical. An appreciable amount of hydrogen peroxide was reported for ozonation of morpholine (28% per compound degraded) and piperazine (60%),

but not for piperidine (4.6%).⁷⁷ This indicates that the Bennett mechanism may have played an important role depending on the target compound. Alternatively, two oxyl radicals can be generated with oxygen (199), which undergo subsequent reactions (200, 201) to form the same products as above.^{83,86}

The pathways for *amine_conj* are only based on ozonation of cylindrospermopsin bearing a single-bonded tertiary nitrogen in a tricyclic guanidine,¹⁰ in which hydroxylation via oxygen transfer (176, 177 in Scheme S4.12) and dealkylation (175, 185, and 189-195 in Scheme S4.12) via electron transfer were reported. While the *N*-oxide formation is generally known to be the final product for tertiary amines as described above, it was proposed therein that an intermolecular hydrogen transfer takes place from a protonated secondary amine of the guanidine moiety to the oxide to form a hydroxyl group. However, as both *N*-oxide and hydroxylamine products have the same mass and identifications were conducted only with mass spectrometry, only *N*-oxide formation was tentatively considered in the current prediction. Moreover, it is assumed that a secondary nitrogen has the same fate.

The reaction pathways for *sulfonamide(2°)_deprot* are based on ozonation of hydrochlorothiazide,² in which a cyclic imine product (i.e., chlorothiazide) was reported to be a stable product, which can be formed via a partial dealkylation (175, 185, and 189-192 in Scheme S4.12). Although an imine product typically undergoes hydrolysis, the formation of an amidine seems to guarantee its stability by tautomerism. Therefore, the reaction rule is defined to cease for the case of amidine formation. Moreover, a protonation of the radical intermediate to form a radical cation was additionally proposed herein for *sulfonamide(2°)_deprot*. pK_a -values of aniline radical cations are several orders of magnitude higher than those of aniline (e.g., 4.63³⁸ and 7.05⁴⁷ for unsubstituted aniline and its radical cation and 5.08³⁸ and 8.50⁴⁷ for 4-methylaniline and its radical cation). Assuming that the same account can be also applied to secondary sulfonamides, the secondary sulfonamide radical species formed from oxidation of the sulfonamide nitrogen anion would readily be protonated due to an increased basicity, which then undergoes a radical shift to α -carbon (189). Another reaction mechanism leading to the same product, i.e., chlorothiazide, via a tautomeric rearrangement was suggested. However, this pathway was not defined herein because it is considered rather exceptional.

Guanine (Scheme S4.13) The reaction pathways in Scheme S4.13 are based on guanine without an amino group at C(2). It was assumed in this study that the pathways can be generally applied to various guanine derivatives. The reaction pathways for guanine were not extended for *adenine* despite the similar chemical structure because the assumption for 1,2-acyl shift seems to be premature (see below). The oxidation of guanine analogues by ozone was reported to be initiated by the formation of an ozonide at the 4,5-double bond (202).^{12,87} This is reminiscent of the Criegee mechanism. However, the formation of an epoxide intermediate via a subsequent singlet oxygen release (203) was proposed to take place for carboxy-acyclovir bearing a guanine moiety,¹² which is not a typical Criegee reaction. The detection of the epoxide was reported from ozonation of uric acid in microdroplets using online mass spectrometry,⁸⁷ and the yields of singlet oxygen per ozone were 40%, 37%, 16%, and 10% for 2'-deoxyguanosine, guanosine, 2'-deoxyadenosine, and adenosine, respectively. Further a 1,2-acyl shift (204) and the subsequent hydrolysis of the imidazole (205) give rise to the final product. An additional hydrolysis (206, 207) of the imine is expected for a guanine analogue in which an amino

Supporting information for chapter 4

group is absent at C(2). For guanine with an amino group at C(8), a hydrolytic product formation (207) is not expected.

However, it is unclear whether or not the final product is formed via an epoxide formation from ozonation of acyclovir.¹² The conventional ozonation pathways can also lead to a singlet oxygen release, the intermediate with the same mass as the epoxide, and the same final product. Applying the conventional ozone reaction pathways to guanine we proposed reactions (208)-(212) based on an ozone adduct formed at C(5) (208). Singlet oxygen release (209) subsequently follows, leaving an oxygen anion in the intermediate with the same mass as the epoxide. Via ensuing rearrangement (210) and hydrolysis (211) the same product is formed. Moreover, the conversion of the amino group into an imino group seems to be better explained by the presence of a carbocation (C4) (208) which leads to a imine bond shift from C(2)=N(3) to N(3)=C(4) (209). In this study, therefore, we chose to perform a pathway prediction for purines with the conventional ozonation pathways. It is noted that in the current study the pathway prediction for guanine is limited to the neutral form as little information is available for the ionic form.

Uracil (Scheme S4.14) The reaction pathways for the reaction of ozone with uracil are presented in Scheme S4.14, which is based on the reaction pathways proposed for thymine and 5-chlorouracil during ozonation.^{63,88} Those pathways are assumed to be also applicable to uracil and cytosine analogues as a ring opening via the Criegee mechanism is also likely to occur for cytosines (see below). The C4 or C5 of a pyrimidine is the initial site of attack by ozone to form an ozone adduct (213), and subsequently undergoes the typical Criegee mechanism (214-218), which was already presented for aromatic compounds and olefins above. Asymmetric cleavage of the ozonide (215 versus 217) takes places, giving rise to two differing zwitterions (A and B). The zwitterionic intermediate A undergoes a series of subsequent reactions (219)-(225), giving rise to a hydrantoin. Hydrantoin can also be formed via the reactions (226)-(231) from A. The same reactions (226)-(231) can be applied to B, giving rise to a hydrotoin. The intermediate (B) alternatively undergoes fragmentation reactions (232)-(237) initiated by the deprotonation of N(3) (233). A partial oxidation rather than a ring opening may occur for a pyrimidine with the anionic N3. The carbocation at C6 of the intermediate (C) is neutralized with the formation of an imine bond (N1=C6) (238). This intermediate undergoes a single oxygen release (239) followed by hydrolysis (241) to form a glycol product. Alternatively, a halogen ion can be released (242) when a halogen substituent is present at the carbon of the ozone adduct to form 6-hydroxypyrimidine-2,4,5-trione via the subsequent hydrolysis (243).

Hetero_ar (no Scheme provided) There is only one reaction pathway defined for non-conjugated *pyridine*, *diazine*, *triazine*, and *phenylazo*. It is a *N*-oxide formation (175, 176 in Scheme S4.12). Pyridine during ozonation was reported to yield almost 80% *N*-oxide per pyridine degraded.⁸⁹ Moreover, azoxybenzene was reported to be the major product from ozonation of azobenzene.¹¹ Based on these reports, it is assumed that *N*-oxide formation may be generally applicable for non-conjugated *pyridine*, *diazine*, and *triazine*, and *phenylazo*. For *triazine*, an aromatic nitrogen conjugated with an amino group was excluded from *N*-oxide formation. Dealkylation of the exocyclic secondary or tertiary amino group were reported for atrazine.⁹⁰ Therefore, a dealkylation mechanism was additionally defined

for triazine with a secondary or a tertiary amino substituent, which is the same pathway as the dealkylation of amine (175, 185, and 189-193 in Scheme S4.12).

Diaminopyrimidine (Scheme S4.15) Despite the fact that this reaction pathway group is defined for *diazine_conj*, by which diverse chemical structures are represented (Figure S4.1), only the reaction mechanism for 2,4-diaminopyrimidines was defined herein (Scheme S4.15) due to a limited information for other moieties.

Ozonation products of the 2,4-diaminopyrimidine were reported for trimethoprim,^{91,92} However, the presented pathways in Scheme S4.15 are not the same as the ones proposed therein.^{91,92} In the reference study,⁹² carbonylation of the carbon in the position 5 of a diaminopyrimidine moiety was proposed, giving rise to a product with the molecular weight of 324. However, the proposed product seems to have a MW of 322 rather than 324. Assuming that the MW 324 of the detected product is correct, we propose pathways giving rise to a product with the same mass, which can be explained based on the known ozone chemistry. In Scheme S4.15, the C(5) of a diaminopyrimidine is available for the attack by ozone to form a carbocation at the C(6) (244). Note that this pathway is also applicable to the C(6) but not shown for simplicity. Assuming that the carbocation is stable enough, it may react with water (245)⁶⁶ rather than undergoing other decay routes. Singlet oxygen can then be released from the ozone adduct with a subsequent proton addition (246). The resulting diol has the same mass (MW=324) of the detected product for trimethoprim. We additionally propose a halogen elimination in its presence to form a product with a carbonyl group (247 and 248).

Ethynyl (Scheme S4.16) An ethynyl group was reported to undergo reactions analogous to the Criegee mechanism for olefins.⁴⁶ As shown in Scheme S4.16, a primary ozonide is formed resulting in a C-C double bond (249). Further hydrolysis breaks down the C=C bond to form a carbonyl and a hydroxyhydroperoxide group for the respective carbons (250, 215). Hydrogen peroxide elimination gives rise to a glyoxal (252). Alternatively, hydrolysis may lead to the cleavage of the C-C bond, yielding two terminal carboxyl groups (253).

Sulfur (no Scheme provided) The oxygen transfer from ozone to sulfur via an ozone adduct is known to be predominant for most organosulfur compounds such as sulfides,^{63,93} sulfinic acids,^{63,94} and thiols.^{63,95,96} For sulfides, sulfoxide is the stable ozone oxidation product. From ozonation of glutathione, thiol was reported to be oxidized to sulfonic acids.⁹⁵ Although a stepwise oxidation from thiol to sulfonic acid via successive oxygen transfer is expected, detailed mechanisms have not been elucidated yet. Therefore, a direct oxidation pathway to sulfonate from thiol and thiolate was defined without intermediates.

Post_ozone reactions (Scheme S4.17) Oxidation products formed via various reaction pathways described above can undergo various post-ozone reactions such as intramolecular rearrangements, hydrolysis, and fragmentation. Such reaction pathways, which are generally applicable, are discussed herein. They are simultaneously applied to predicted final products of all the compound groups. First, post-Criegee reactions ((154)-(164) in Scheme S4.11), which mainly deal with hydrolysis and decarboxylation, can be generally applied. In addition, it was reported from ozonation of progesterone that an intramolecular Bayer-Villiger type reaction ((254)-(255)) may occur to a minor extent for an α -

hydroxyhydroperoxide product,⁵ which converts ketones to esters. The resulting product from the Bayer-Villiger type reaction for progesterone is an anhydride, which undergoes hydrolysis to the corresponding carboxylic acids (256). The hydrolysis of anhydrides was defined to be generally applicable, not only specific to the product from the Bayer-Villiger type reactions.

Imine formation between primary amine and ketone via reactions ((257)-(259)) was observed from ozonation of carbamazepine.⁷² As described above in the amine section (reaction (193) in Scheme S4.12), imine is prone to hydrolysis, therefore, a reverse reaction is possible, which depends on the stability of the imine. The imine product of carbamazepine during ozonation, 1-(2-benzaldehyde)-(1*H*,3*H*)-quinazoline-2,4-dione (BQD), was stabilized by forming an aromatic system. Moreover, amidine was discussed above to be stabilized by tautomerism. It is generally known that hydrolytic stability of imines are attributed to delocalization of the imine double bond with electron-withdrawing substituents, e.g., oxime and hydrazone.⁹⁷ Although there will be various other environments where imines are resistant to hydrolysis, only the aforementioned conditions are implemented.

Text S4.10. Case study: comparison between experimental k_{O_3} and predicted k_{O_3} for cetirizine

Figure S4.9a shows a comparison between the experimental k_{O_3} ² and the predicted k_{O_3} for cetirizine. The predicted k_{O_3} -values were obtained from the default method of this study (i.e., HF/3-21G//MMFF in Text S4.5) using a species distribution shown in Figure S10c. The acid-base speciation scheme for cetirizine is shown in Figure S4.10b where individual acid-base species are denoted as X_i ($i=1$ to 8). Note that the most basic $pK_{a,3}$ -value (Figure S4.10b) is reported to be ~ 8.0 in literature.^{98,99} However, a $pK_{a,3}$ of 7.0, which was suggested from the determination of k_{O_3} for cetirizine,² was instead chosen to have a consistent comparison between the measured k_{O_3} and the predicted k_{O_3} . As shown in Figure S4.9a, the predicted apparent k_{O_3} (dash-double-dot cyan line) was much higher than the experimental k_{O_3} (yellow circles) for the entire pH range. Noticeably, the error increases to more than an order of magnitude with decreasing pH below 7. The experimental k_{O_3} manifests the effect of amine protonation ($pK_{a,3}=7.0$), which decreases the reactivity of the amine nitrogen with ozone, while the predicted apparent k_{O_3} is uninfluenced by pH until pH 3. In Figures S4.9c and S4.9d, individual nitrogen-specific k_{O_3} -values of the species contributing to the predicted apparent nitrogen-specific k_{O_3} in Figure S4.9a are presented (black lines). The species X_2 , X_6 , and X_8 and the species X_3 , X_7 , and X_8 contribute to the predicted k_{O_3} of the nitrogens 1 and 2, respectively. The dominant species at pH > 8 and acidic to neutral pH (2-7) were X_8 (dotted black line) and X_7 (long-short dashed black line), respectively. As shown in Figure S10b, the species X_7 has the protonated nitrogen 1 and the free amine nitrogen 2 and it was expected that the protonation of an amine nitrogen in piperazine would have decreased the reactivity of the other nitrogen by withdrawing electron density from the free nitrogen to the protonated nitrogen. However, almost the same k_{O_3} -values were predicted for the nitrogen 2 of X_7 , and X_8 , $1.31 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$ for X_7 to $1.14 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$ for X_8 , respectively. As shown in Figure S4.11 (top row), it is observed that a moderate hydrogen bonding was formed between the hydrogen of the protonated amine and the carboxylate oxygen with a distance of 2.67 Å. Moreover, other species with a protonated amine such as X_2 , X_3 , and X_6 show stronger hydrogen bonding with shorter distances (1.57-2.25 Å)

(Figure S4.11). The ionic intramolecular interaction between the two atoms makes an overall geometry more stabilized. However, as the presented conformations were obtained by the MMFF94 method in the gas phase, they may not be relevant for aqueous solution where ionic charges can be locally stabilized in the polar environment. Therefore, initial geometries of cetirizine with spread-out conformations were optimized by the B3LYP/6-31G* method employing the integral equation formalism polarizable continuum model¹⁰⁰ (IEFPCM) as a solvation method. Note that this computation was performed using Gaussian.¹⁹ The optimized geometries are presented in Figure S4.11 (bottom row) in comparison with the geometries obtained by MMFF94 and straight conformations retained following the geometry optimization. When the total electronic energies are compared with the default method (HF/3-21G) for single point calculation (Figure S4.11), the optimized geometries in aqueous medium were more stable than the gas phase geometries obtained by the MMFF94 method. This supports that the MMFF94 conformations are less relevant to the experimental data obtained in aqueous solution. The predicted k_{O3} using the geometries relevant to the aqueous solution are presented in Figures S4.9b, S4.9c, and S4.9d. Note that the predicted k_{O3} -values for the benzene rings are neglected because of the low reactivity ($< 10^2 \text{ M}^{-1}\text{s}^{-1}$). With the spread-out conformers, a predicted apparent k_{O3} drop for nitrogen 2 was observed (Figure S4.9b), which is consistent with the amine protonation and the experimental data, however, the extent of the drop was still small compared to the experimental data.

In fact, all the prediction outputs presented in Figure S4.9 are obtained with the acid-base species of dextrocetirizine (*S* enantiomer). The default MMFF94 method for the 3D geometry generation consistently yielded dextrocetirizine (Figure S4.11 in which it can be recognized by a hydrogen bond to the chiral center carbon in comparison with the 2D chemical structure in Figure S4.10). For a consistent comparison between k_{O3} predictions with different conformers, therefore, dextrocetirizine was only investigated. However, the racemic mixture of levocetirizine (*R* enantiomer) and dextrocetirizine was used to determine experimental k_{O3} .² Thus, the presence of levocetirizine needs to be taken into account for the k_{O3} prediction to be consistent with the experimental conditions. Therefore, a k_{O3} prediction for levocetirizine was performed as well. Same as above for dextrocetirizine, initial geometries were prepared with spread-out conformations, which were then optimized by B3LYP/6-31G* with the solvation method employed by Gaussian.¹⁹ The optimized geometries were submitted to single point calculations with HF/3-21G and the corresponding k_{O3} were predicted accordingly. As shown in Figure S4.12a, the predicted k_{O3} for levocetirizine overall shows a similar pattern to the predicted k_{O3} for dextrocetirizine (Figure S4.9b) except that a shift of the dominant nitrogen from 1 to 2 at pH below 7 was observed for levocetirizine. The difference in the predicted apparent nitrogen-specific k_{O3} between levo- and dextro-cetirizines was within a factor of 5 at the maximum. The predicted apparent k_{O3} of two enantiomers were combined with an equal weighting factor (0.5) to derive the predicted apparent k_{O3} for the racemic mixture, which is a more appropriate representation of the experimental k_{O3} (Figure S4.12b). Even with the consideration of the enantiomers, the prediction fails still with an error of more than 2 orders of magnitude.

k_{O3} for unsubstituted piperazine bearing two secondary amine nitrogens was reported to be $2.6 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$.⁷⁷ In contrast, the k_{O3} prediction by the HF/3-21G/MMFF method yields $1.2 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ with an overestimation of almost two orders of magnitude. Based on the fact that the k_{O3} -values for secondary amines range typically in the order of $10^5 - 10^6 \text{ M}^{-1}\text{s}^{-1}$,¹¹ or even $10^7 \text{ M}^{-1}\text{s}^{-1}$ (e.g., $1.2 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$ for

Supporting information for chapter 4

dimethylamine⁹), the measured k_{O_3} of $2.6 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ for cetirizine is dramatically different from acyclic secondary amines. Interestingly, lowering the predicted k_{O_3} by the difference between the measured k_{O_3} and the predicted k_{O_3} for piperazine, i.e., two orders of magnitudes, led to a prediction in a close proximity with the experimental data (dotted line in Figure S4.12b). The overestimation by an order of magnitude was also observed for piperidine: experimental $k_{O_3} = 2.6 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$,⁷⁷ and the predicted $k_{O_3} = 1.2 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$. Based on this, it may be that the k_{O_3} prediction model may consistently overestimate k_{O_3} for cyclic amines. However, further investigations are needed to confirm whether this correction can be generally applicable to cyclic amines or if it was only coincidental to cetirizine.

Some aspects of the developed prediction platform, which may lead to an error in the predicted k_{O_3} , have been highlighted with the example of cetirizine. The MMFF94 method, a default method for 3D geometry generation in this study, may give rise to a geometry of a compound irrelevant to the aqueous phase, potentially leading to an error in predicting k_{O_3} . Therefore, the development of k_{O_3} prediction models by various conformer search tools generating an aqueous-phase relevant geometry and a model verification with external training compounds such as cetirizine need to be explored. Although it was evaluated for cetirizine that the overall improvement by using geometries relevant to aqueous phase was rather small, this aspect is still considered to be important because the effect of a protonation could be consistently represented in a predicted k_{O_3} obtained using aqueous-phase relevant geometries.

Moreover, the currently developed platform considers only a conformer for the k_{O_3} prediction. As shown for cetirizine, experiments may be conducted for a compound with isomerism, i.e., co-presence of multiple conformers. However, it is not planned to be implement this aspect due to associated complications for the development, e.g., predicting weighting factors (relative populations) of individual isomers involved unless otherwise provided by a user. However, one has to keep in mind that k_{O_3} predictions may be significantly influenced by conformers for some compounds, which are yet unknown.

Lastly, as it turned out above that the consideration of the geometry and enantiomers did not lead to a significant improvement, the poor k_{O_3} prediction model for cetirizine seems to stem from the failure of the prediction model itself. Further investigations on whether this failure occurs consistently for similar compounds as cetirizine (e.g., a compound containing a piperazine moiety or a cyclic amine in general) would be helpful to shed light on.

Although the main focus of this text was on the nitrogens of the piperazine moiety, it is worthwhile to discuss the predicted k_{O_3} for the two benzene rings. As shown in Figure S4.9a, k_{O_3} for the chlorine-substituted benzene ring is lower than k_{O_3} for unsubstituted benzene ring over the whole pH-range between 0 and 14 except for the pH range 4 – 6 where almost the same k_{O_3} was observed. Interestingly, an increase in k_{O_3} for the chlorine-substituted benzene was observed in the pH range 3 – 7 (Figure S4.9a). This is counter-intuitive because the protonation of the nitrogen 2 would lead to a decreased k_{O_3} of the benzene ring if ever manifested. Firstly, a gas-phase geometry with the hydrogen bonding may have played a role to to this phenomenon as described in the k_{O_3} for the amine above. Secondly, it may be also attributed to the assignment of E_{HOMO} based on the threshold $\text{SS}_{\text{MOcoef}}$ -value (0.22 for

aromatic compounds in general). Because of the $SS_{MO_{coef}}$ threshold of 0.22, several E_{HOMO-n} of micro-species for cetirizine could not be assigned although their calculated $SS_{MO_{coef}}$, i.e., 0.17-0.22, may be high enough to be meaningful, thus acceptable. As described above in Text S4.7, the ambiguity in assigning E_{HOMO} may lead to a difficulty in interpreting the predicted k_{O3} -values and needs to be resolved in future.

Supporting information for chapter 4

Table S4.1. A list of all reactive sites defined for the prediction platform and the associated k_{O_3} estimate, k_{O_3} prediction model group, and reaction pathway group. For structures see Figure S4.1.

Reactive site	k_{O_3} estimate ^{a,b} , M ⁻¹ s ⁻¹	k_{O_3} prediction model group (associated orbital energy)	reaction pathway group
Aromatic compounds			
phenol	<i>prediction</i>	phenol (E_{HOMO-n})	phenol
aniline	<i>prediction</i>	aniline (E_{HOMO-n})	aniline
cyclic_aniline	4.5×10^2	<i>n.a.</i>	aniline
12alkoxybenzene	<i>prediction</i>	mono- and di-alkoxybenzene (E_{HOMO-n})	alkoxybenzene
3alkoxybenzene	<i>prediction</i>	trialkoxymbenzene (E_{HOMO-n})	alkoxybenzene
benzene	<i>prediction</i>	benzene (E_{HOMO-n})	benzene
benzazole_dervs	<i>prediction</i>	benzene (E_{HOMO-n})	benzazole_dervs
Olefins/Ethynyl group			
olefin	<i>prediction</i>	olefin ($E_{NBO, C-C}$)	olefin
olefin_conj	<i>prediction</i>	miscolefin (E_{HOMO-n})	olefin
ethynyl	2×10^2	<i>n.a.</i>	ethynyl
Amines			
amine	<i>prediction</i>	amine ($E_{NBO, LP-N}$)	amine
hydrazine	<i>n.a.</i>	<i>n.a.</i>	<i>n.a.</i>
N_sulfenamide	<i>n.a.</i>	<i>n.a.</i>	amine
amine_conj	<i>n.a.</i>	<i>n.a.</i>	<i>n.a.</i>
phenylazo	<i>prediction</i>	hetero_ar (E_{HOMO-n})	hetero_ar
sulfonamide(1°)_prot	1	<i>n.a.</i>	<i>n.a.</i>
sulfonamide(1°)_deprot	7×10^3	<i>n.a.</i>	<i>n.a.</i>
sulfonamide(2°)_prot	<i>n.a.</i>	<i>n.a.</i>	<i>n.a.</i>
sulfonamide(2°)_deprot	3×10^5	<i>n.a.</i>	amine
sulfonamide(3°)	<i>n.a.</i>	<i>n.a.</i>	<i>n.a.</i>
Heteroaromatic compounds			
adenine	<i>prediction</i>	adenine (E_{HOMO-n})	<i>n.a.</i>
guanine	<i>prediction</i>	guanine (E_{HOMO-n})	guanine
cytosine	<i>prediction</i>	cytosine (E_{HOMO-n})	uracil
thymine	<i>prediction</i>	thymine (E_{HOMO-n})	uracil
uracil	<i>prediction</i>	uracil (E_{HOMO-n})	uracil
pyridine	<i>prediction</i>	hetero_ar (E_{HOMO-n})	hetero_ar
diazine	<i>prediction</i>	hetero_ar (E_{HOMO-n})	hetero_ar
triazine	<i>prediction</i>	hetero_ar (E_{HOMO-n})	hetero_ar
pyridine_conj	<i>prediction</i>	miscolefin (E_{HOMO-n})	<i>n.a.</i>
diazine_conj	<i>prediction</i>	miscolefin (E_{HOMO-n})	diaminopyrimidine
quinolone	1.0	<i>n.a.</i>	olefin
quinolone_carboxylate	2×10^4	<i>n.a.</i>	olefin
Shetero_ring	<i>prediction</i>	miscolefin (E_{HOMO-n})	olefin
Organosulfur compounds			
sulfide	8×10^5	<i>n.a.</i>	sulfur
disulfide	2×10^5	<i>n.a.</i>	sulfur
sulfide_6cyc	5×10^4	<i>n.a.</i>	sulfur
sulfide_5cyc	5×10^3	<i>n.a.</i>	sulfur
thiol_prot	3×10^4	<i>n.a.</i>	sulfur
thiol_deprot	3×10^6	<i>n.a.</i>	sulfur
thiophenol	<i>n.a.</i>	<i>n.a.</i>	<i>n.a.</i>
thiophosphoramidate	<i>n.a.</i>	<i>n.a.</i>	<i>n.a.</i>
S_sulfenamide	<i>n.a.</i>	<i>n.a.</i>	<i>n.a.</i>
thioamide	<i>n.a.</i>	<i>n.a.</i>	<i>n.a.</i>
thioketone	<i>n.a.</i>	<i>n.a.</i>	<i>n.a.</i>
sulfide_conj	6×10^2	<i>n.a.</i>	<i>n.a.</i>
sulfoxide	10	<i>n.a.</i>	<i>n.a.</i>
sulfinic_acid	2×10^6	<i>n.a.</i>	sulfur
organothiophosphate	<i>n.a.</i>	<i>n.a.</i>	<i>n.a.</i>

n.a.: not available. ^areference compounds, from which k_{O_3} estimates are derived, are presented in Table S4.2. ^ba reactive site, to which 'prediction' is assigned, is to be predicted by the corresponding k_{O_3} prediction model group in the 3rd column.

Table S4.2. k_{O_3} estimates for reactive sites for which no k_{O_3} prediction model is available and the corresponding reference compounds from which k_{O_3} estimates are derived.

Reactive site ^a	k_{O_3} estimate, $M^{-1}s^{-1}$	Reference compound (empirical k_{O_3} , $M^{-1}s^{-1}$)
Aromatic compounds		
cyclic_aniline	4.5×10^2	hychlorochlorothiazide (6×10^2), ² ciprofloxacin diprotonated (4×10^2), ³ enrofloxacin diprotonated (3.3×10^2) ³
Olefin/Ethynyl group		
ethynyl	2×10^2	1-ethynyl-1-cyclohexanol (2×10^2) ⁴⁶
Amine		
sulfonamide(1°)_prot	1	<i>N,N</i> -dimethylsulfamide(1.24 ^c) ¹⁰¹
sulfonamide(1°)_deprot	7×10^3	<i>N,N</i> -dimethylsulfamide(7×10^3 , ^d) ^{101,102}
sulfonamide(2°)_deprot	3×10^5	hydrochlorthiazide (3×10^5) ²
Heteroaromatic compounds		
quinolone	1	flumequine neutral (1.2) ³
quinolone_carboxylate	2×10^4	flumequine deprotonated (1.8×10^3), ³ ciprofloxacin monoprotonated (7.5×10^3), ³ enrofloxacin monoprotonated (4.6×10^4) ³
Organosulfur compounds		
thiol-prot	3×10^4	cysteine (4.2×10^4), ¹ glutathione (2.0×10^4) ¹
thiol-deprot	3×10^6	cysteine anion (2.4×10^6), ¹ glutathione anion(4.0×10^6) ¹
sulfide	8×10^5	aldicarb (4.3×10^5), ² methionine (2.4×10^6) ¹
sulfide_5cyc ^b	5×10^3	penicillin G (4.8×10^3) ³
sulfide_6cyc ^b	5×10^4	cephalexin ($5 - 10 \times 10^4$) ³
sulfide_conj	6×10^2	EPTC (500;estimated) ⁴ , molinate (500;estimated), ⁴ vydate (620) ²
disulfides ^c	2×10^5	<i>trans</i> -1,2-dithiane-4,5-diol(2.1×10^5), ⁵ bis(2-hydroxyethyl) disulfide (1.7×10^5) ⁵
sulfoxide	10	dimethylsulfoxide (8) ¹
sulfinic acid	2×10^6	methanesulfinic acid (2.0×10^6)

^asee Figure S4.1 for their chemical structures. ^bThe number indicates the number of the atoms of a cyclic ring. ^c k_{O_3} of 550 and $1.0 \times 10^3 M^{-1}s^{-1}$ was reported for cystine at pH 1 and 3.1, respectively. These values are significantly different from *trans*-1,2-dithiane-4,5-diol and bis(2-hydroxyethyl) disulfide and the reason for this is yet unclear. In this study, the representative k_{O_3} for disulfides was set excluding cystine. ^cApparent k_{O_3} at pH 5.¹⁰¹ ^dEstimated based on the reported apparent k_{O_3} measurements between pH 5 and 10^{101} and the reported pK_a .¹⁰²

Supporting information for chapter 4

Table S4.3. The currently developed k_{O_3} prediction models for various organic compound groups in comparison with the previously developed models

k_{O_3} prediction model group	Orbital energy	Present model ^a			Reference model ^b		
		slope(a) /y-intercept(y_0) ^c	N_{model}^d (N_{total}^e)	R^2 (MUE ^f , RMSE ^g)	N_{model}^d (N_{total}^e)	R^2 (MUE ^f , RMSE ^g)	
Aromatic compounds							
phenol	$E_{\text{HOMO-n}}$	55.43/22.19	35 (36)	0.94 (0.46, 0.67)	35 (36)	0.94 (0.42, 0.61)	
aniline ^h	$E_{\text{HOMO-n}}$	21.30/12.31	22 (22)	0.43 (0.26, 0.30)	16 (16)	0.85 (0.34, 0.43)	
mono- and dialkoxybenzenes	$E_{\text{HOMO-n}}$	86.74/30.55	17 (17)	0.74 (0.71, 0.96)	17 (17)	0.87 (0.57, 0.69)	
trialkoxybenzenes	$E_{\text{HOMO-n}}$	-95.32/-25.80	4 (4)	0.69 (0.26, 0.31)	4 (4)	0.95 (0.11, 0.12)	
benzenes	$E_{\text{HOMO-n}}$	46.68/16.71	39 (40)	0.75 (0.58, 0.78)	40 (40)	0.82 (0.47, 0.61)	
Olefin	$E_{\text{NBO,C=C}}$	31.26/16.78	45 (53)	0.79 (0.57, 0.76)	45 (53)	0.84 (0.51, 0.69)	
Amine	$E_{\text{NBO,L,P-N}}$	29.83/19.25	59 (59)	0.75 (0.61, 0.84)	59 (62)	0.83 (0.59, 0.86)	
Heteroaromatic compounds							
adenine	$E_{\text{HOMO-n}}$	49.79/18.12	8 (8)	0.76 (0.63, 0.70)	8 (8)	0.73 (0.66, 1.26)	
cytosine	$E_{\text{HOMO-n}}$	41.19/17.35	8 (8)	0.91 (0.35, 0.45)	8 (8)	0.98 (0.17, 0.28)	
guanine	$E_{\text{HOMO-n}}$	57.06/21.45	8 (8)	0.95 (0.31, 0.33)	8 (8)	0.99 (0.14, 0.23)	
thymine	$E_{\text{HOMO-n}}$	35.50/16.40	5 (5)	0.92 (0.22, 0.26)	5 (5)	0.94 (0.17, 0.34)	
uracil	$E_{\text{HOMO-n}}$	50.55/20.75	8 (8)	0.80 (0.52, 0.65)	8 (8)	0.84 (0.45, 0.76)	
other heteroaromatic compounds (hetero_ar)	$E_{\text{HOMO-n}}$	52.78/19.51	9 (9)	0.80 (0.48, 0.57)	<i>n.a</i>	<i>n.a</i>	
Miscellaneous (miscolefin)	olefins	$E_{\text{HOMO-n}}$	75.10/29.60	17 (17)	0.95 (0.43, 0.48)	32 (32)	0.89 (0.51, 0.67)

^a k_{O_3} prediction models developed with the HF/3-21G//MMFF94 method. See Texts S4.2 and S4.3 for the details. ^bPreviously developed models with HF/6-31G method used throughout all the computations for all k_{O_3} prediction model groups except for the amines for which HF/6-311++G** was used. See the reference^d for more details. ^capplied to a linear regression model to predict k_{O_3} in log units: $\log k_{O_3} = aE + y_0$. ^dthe number of model compounds used to develop a linear regression model. ^ethe number of the compounds including both model compounds and outliers used to calculate MUEs and RMSEs. ^fmean unsigned error. ^groot-mean-square error. ^hmean unsigned error. ^groot-mean-square error.

Table S4.4. The selected model compounds used for the development of a k_{O_3} prediction model for differing aromatic compound groups

No.	Phenol	Aniline ^a	Mono- and di-alkoxybenzene	Trialkoxybenzene	Benzene
1	bisphenol-A	aniline	acetutolol protonated	trimethoprim monoprotontaed	benzaldehyde
2	bisphenol-A dianion	3-chloro aniline	anisole	trimethoprim diprotontaed	benzene
3	2-chloro phenol	4-chloro aniline	atenolol protonated	1,3,5-trimethoxybenzene	benzene sulfonate
4	2-chloro phenolate	2-fluoro aniline	bezafibrate	3,4,5-trimethoxytoluene	benzoate
5	4-chloro phenol	3-fluoro aniline	carbofuran		chlorobenzene
6	4-chloro phenolate	4-fluoro aniline	2,4-dichlorophenoxyacetic acid		diazepam
7	2,3-dimethyl phenol	3-iodo aniline	1,4-dimethoxy benzene		diethyl- <i>o</i> -phthalate
8	2,4-dimethyl phenol	4-iodo aniline	methoxychlor		1,3-dichlorobenzene
9	2,6-dimethyl phenol	2-methoxy aniline	4-methoxy-1-naphthalenesulfonic acid		ethylbenzene
10	3,4-dimethyl phenol	3-methyl aniline	2-methyl-4-chlorophenoxyacetic acid		galaxolide
11	2,6-di- <i>n</i> -butyl-4-methylphenol	4-methylsulfone aniline	2-methyl-4-chlorophenoxypropionic acid		ibuprofen
12	17 β -estradiol	3-nitro aniline	metoprolol protonated		isopropylbenzene
13	17 β -estradiol anion	4-nitro aniline	naproxen anion		methylbenzoate
14	estriol	sulfamethoxazole	1-phenoxy-2-propanol		1-methylnaphthalene
15	estriol anion	sulfamethoxazole anion	propranolol protonated		naphthalene
16	estrone	4-methyl aniline	2,4,6-tribromoanisole		nitrobenzene
17	estrone anion	4-methoxy aniline	2,4,6-trichloroanisole		toluene
18	catechol	4-phenylenediamine			tonalide
19	resorcinol	2-hydroxy aniline			1,2,3-trimethylbenzene
20	hydroquinone	4-hydroxy aniline			1,3,5-trimethylbenzene
21	<i>o</i> -cresol	4-aminobenzoic acid			<i>m</i> -xylene
22	<i>m</i> -cresol	<i>N,N</i> -dimethylaniline			<i>o</i> -xylene
23	<i>p</i> -cresol				<i>p</i> -xylene
24	4-nitrophenolate				alachlor
25	4- <i>n</i> -nonylphenol				chlorotoluron
26	4- <i>n</i> -nonylphenolate				diuron
27	4-octylphenol				isoproturon
28	pentabromo phenolate				linuron
29	pentachloro phenolate				metolachlor
30	phenol				propachlor
31	phenolate				benzotriazole
32	2-carboxylate phenol anion				benzotriazole anion
33	triclosan				5-chlorobenzotriazole
34	triclosan anion				5-chlorobenzotriazole anion
35	2,4,6-triiodophenol anion				5,6-dimethylbenzotriazole
36	<i>paracetamol</i>				5,6-dimethylbenzotriazole anion
37					5-methylbenzotriazole
38					5-methylbenzotriazole anion
39					quinoline
40					<i>amidotrizoic acid</i>

Refers the SI in the original reference⁴ for experimental k_{O_3} and their references. Outliers, which are excluded from the model calibration but included in the prediction performance evaluation (i.e., MUE, RMSE), are given in *italic* ^aSee Table S4.8 for a list of the selected anilines and up-to-date k_{O_3} -values used for the model development.

Supporting information for chapter 4

Table S4.5. The selected model compounds used for the development of a k_{O_3} prediction model for olefins, amines, and heteroaromatic compounds

No.	Olefin	Amine	Heteroaromatic compound (hetero ar)
1	<i>cis</i> -1,2-dibromoethene	<i>N</i> - α -acetyllysine	azobenzene
2	1,1-dichloroethene	<i>N</i> - ϵ -acetyllysine	2-isopropyl-3-methoxypyrazine
3	<i>cis</i> -1,2-dichloroethene	alanine	pyridine
4	1,1-dichloropropene	β -alanine	pyridine protonated
5	ethene	arginine	atrazine
6	hexachlorocyclopentadiene	asparagine	deethylatrazine
7	propene	aspartate	deethyldeisopropylatrazine
8	trichloroethene	benzylamine	deisopropylatrazine
9	vinyl chloride	bromamine,	simazine
10	vinyl bromide	<i>n</i> -butylamine	
11	acrylic acid	<i>sec</i> -butylamine	
12	acrylic acid anion	<i>tert</i> -butylamine	
13	cinnamic acid	cyclohexanemethylamine	
14	cinnamic acid anion	cyclohexylamine	
15	fumaric acid	glutamine	
16	2-hexenoic acid anion	glycine	
17	maleic acid	hydroxylamine	
18	methacrylic acid	isoleucine	
19	methacrylic acid anion	leucine	
20	4-methoxy cinnamic acid	lysine	
21	4-methoxy cinnamic acid anion	monochloramine	
22	<i>trans-trans</i> -muconic acid dianion	phenylalanine	
23	4-nitro-cinnamic acid anion	serine	
24	sorbic acid	threonine	
25	sorbic acid anion	valine	
26	diethyl vinylphosphonate	acebutolol	
27	vinyl phosphonic acid	anatoxin	
28	vinyl phosphonic acid monoanion	atenolol	
29	vinyl phosphonic acid dianion	dibromamine,	
30	vinylsulfonate	dichloramine	
31	2-acetamidoacrylic acid	diclofenac	
32	acrylamide	diethylamine	
33	1,4-benzoquinone	dimethylamine	
34	but-1-en-2-ol	ethyl <i>N</i> -piperazinecarboxylate	
35	vinyl acetate	iminodiacetic acid	
36	anatoxin	methylchloramine	
37	cephalexin	metoprolol	
38	<i>cis</i> -3-hexen-1-ol	morpholine	
39	β -ionone	proline	
40	<i>trans-cis</i> -2,6-nonadienal	spectinomycin,	
41	oseltamivir acid	spectinomycin protonated	
42	1-penten-3-one	ciprofloxacin	
43	carbamazepine	1,4-diazabicyclo[2.2.2]octane	
44	tylosin protonated	DABCO monoprotonated	
45	MCLR	dimethylchloramine	
46	<i>trans</i> -1,2-dichloroethene	<i>N,N</i> -dimethylcyclohexylamine	
47	3,4-dihydroxycinnamic acid	dimethylethanolamine	
48	3,4-dihydroxycinnamic acid anion	enrofloxacin	
49	maleic acid anion	ethylenediaminetetraacetic acid	
50	3-methoxy-4-hydroxy cinnamic acid	EDTA mono protonated	
51	3-methoxy-4-hydroxy cinnamic acid anion	lincomycin	
52	acrylonitrile	1-methylpyrrolidine	
53	vinylene carbonate	nitrilotriacetic acid	
54		tramadol	
55		triethylamine	
56		trimethylamine	
57		azithromycin	
58		roxithromycin	
59		tylosine	
60		<i>ammonia</i>	

Refers the SI in the original reference⁴ for experimental k_{O_3} and their references. Outliers, which are excluded from the model calibration but included in the prediction performance evaluation (i.e., MUE, RMSE), are given in *italic*

Table S4.6. The selected model compounds used for the development of a k_{03} prediction model for adenine, guanine, cytosine, thymine, uracil, and miscellaneous olefins

No.	Adenine	Guanine	Cytosine	Thymine	Uracil	Miscellaneous olefins
1	adenine (9H)	aciclovir (1H)	cytidine	thymidine	5-chlorouracil	β -cyclocitral
2	adenine protonated (1,9H)	aciclovir anion	cytidine protonated	thymidine anion	5-chlorouracil anion (1H)	levonogestrel
3	adenine anion	aciclovir protonated (1,7H)	cytosine (1H)	5-thymidylic acid	1,3-dimethyluracil	medroxyprogesterone
4	adenosine	2'-deoxyguanosine (1H)	cytosine anion	thymine	6-methyluracil (1,3H)	norethindrone
5	adenosine protonated (1H)	5'-deoxyguanylic acid (1H)	cytosine protonated	thymine anion (1H)	6-methyluracil anion (1H)	progesterone
6	5'-adenylic acid	guanosine (1H)	2'-deoxycytidine		orotic acid anion	indigotrisulfonic acid
7	2'-deoxyadenosine	guanosine anion	2'-deoxycytidine protonated		uracil	imidazole
8	2'-deoxyadenosine protonated (1H)	guanosine protonated (1,7H)	5'-deoxyeyidylic acid		uracil anion (1H)	imidazole protonated
9						4-methylimidazole
10						3,5-dimethylisoxazole
11						2,4-diamino-5-methylpyrimidine
12						2,4-diamino-5-methylpyrimidine monoprotonated
13						trimethoprim deprotonated
14						tryptophan
15						3-(dimethylaminomethyl) indole protonated
16						N-(4)-acetylsulfamethoxazole deprotonated
17						N-(4)-acetylsulfamethoxazole protonated

Refers the SI in the original reference⁴ for experimental k_{03} and their references.

Supporting information for chapter 4

Table S4.7. A brief summary for the performance evaluation of the k_{03} prediction model developed with seven different computational methods.

No.	Computational method	3D geometry generation	Geometry Optimization	Single point calculation	Remark ^a
1	HF/6-31G//PM3/MMFF	MMFF	PM3	HF/6-31G	- Most expensive computational costs - Overall stable prediction performance
2	HF/SV//PM3/MMFF	MMFF	PM3	HF/SV	- 2 nd most expensive computational costs - Overall stable prediction performance - 2 nd most expensive computational costs
3	HF/SV//PM3/Dreiding	Dreiding	PM3	HF/SV	- Poor performance for mono- and di-alkoxybenzene, trialkoxybenzene, and guanine attributable to the Dreiding force field method
4	HF/3-21G//PM3/MMFF	MMFF	PM3	HF/3-21G	- Overall stable prediction performance
5	HF/3-21G//PM3/Dreiding	Dreiding	PM3	HF/3-21G	- Overall stable prediction performance - Poor performance for mono- and di-alkoxybenzene, trialkoxybenzene, and guanine attributable to the Dreiding force field-derived initial 3D geometry
6	HF/3-21G//MMFF ^b	MMFF	<i>n. i.</i>	HF/3-21G	- The least expensive computational costs - Overall stable prediction performance - Selected method for this study
7	HF/3-21G/Dreiding	Dreiding	<i>n. i.</i>	HF/3-21G	- The least expensive computational costs - Poor performance for mono- and di-alkoxybenzene and trialkoxybenzene attributable to the Dreiding force field-derived initial 3D geometry

n. i. not implemented. ^aremarks were based on the evaluation outputs presented in Figures S4.2 and S4.3 and detailed descriptions are given in Text S4.4. ^bdefault method chosen in this study for a k_{03} prediction.

Table S4.8. Second-order rate constants (k_{O_3}) of the selected anilines used for development of the previous model⁴ and the current model.

Compound	Previous $k_{O_3}^a$, $M^{-1}s^{-1}$	Current $k_{O_3}^b$, $M^{-1}s^{-1}$
Aniline	1.40×10^7 ⁽³⁹⁾	1.30×10^6 ⁽³⁸⁾
3-chloroaniline	7.84×10^6 ⁽³⁹⁾	7.28×10^5 ^{5,c}
4-chloroaniline	1.04×10^7 ⁽³⁹⁾	1.40×10^6 ⁽³⁸⁾
2-fluoroaniline	7.28×10^6 ⁽³⁹⁾	6.76×10^5 ^{5,c}
3-fluoroaniline	8.12×10^6 ⁽³⁹⁾	7.54×10^5 ^{5,c}
4-fluoroaniline	1.23×10^7 ⁽³⁹⁾	1.14×10^6 ^{6,c}
3-iodoaniline	7.42×10^6 ⁽³⁹⁾	6.89×10^5 ^{5,c}
4-iodoaniline	9.24×10^6 ⁽³⁹⁾	8.58×10^5 ^{5,c}
2-methoxyaniline	1.72×10^7 ⁽³⁹⁾	1.60×10^6 ^{6,c}
3-methylaniline	1.88×10^7 ⁽³⁹⁾	1.75×10^6 ^{6,c}
4-methylsulfone aniline	4.70×10^4 ⁽³⁾	4.70×10^4 ⁽³⁾
3-nitroaniline	2.38×10^6 ⁽³⁹⁾	2.21×10^5 ^{5,c}
4-nitroaniline	1.40×10^5 ⁽³⁹⁾	1.20×10^5 ⁽³⁸⁾
sulfamethoxazole	4.70×10^4 ⁽³⁾	4.70×10^7 ⁽³⁾
sulfamethoxazole anion	5.70×10^5 ⁽³⁾	5.70×10^5 ⁽³⁾
<i>N,N</i> -dimethylaniline	2.00×10^9 ⁽⁴⁰⁾	1.60×10^6 ⁽³⁸⁾
4-aminobenzoate	<i>n.a.</i>	8.50×10^5 ⁽³⁸⁾
4-methylaniline	<i>n.a.</i>	2.40×10^6 ⁽³⁸⁾
4-methoxyaniline	<i>n.a.</i>	1.10×10^6 ⁽³⁸⁾
2-hydroxyaniline	<i>n.a.</i>	3.70×10^5 ⁽³⁸⁾
4-hydroxyaniline	<i>n.a.</i>	7.80×10^5 ⁽³⁸⁾
<i>p</i> -phenylenediamine	<i>n.a.</i>	1.30×10^6 ⁽³⁸⁾

n.a. not available when the model was established. ^a k_{O_3} for anilines used for the development of the previous model⁴, ^b k_{O_3} for anilines used for the development of the current model, ^ccorrected k_{O_3} for anilines based on the relative difference of k_{O_3} for unsubstituted aniline between the previous (i.e., $1.40 \times 10^7 M^{-1}s^{-1}$)³⁹ and the recent (i.e., $1.30 \times 10^6 M^{-1}s^{-1}$)³⁸ measurement.

Supporting information for chapter 4

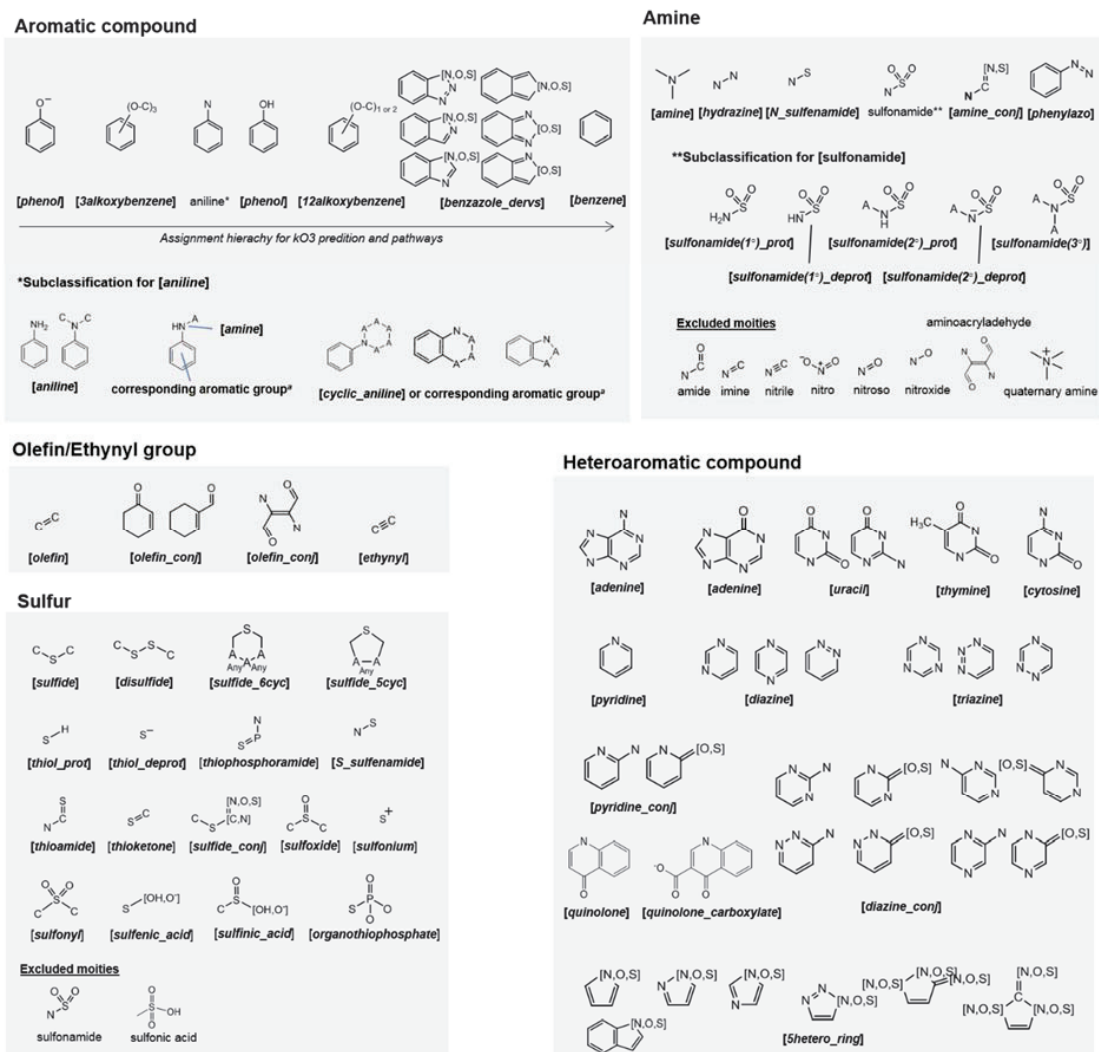


Figure S4.1. Chemical structures for defined reactive sites. Names in **bold** and *italic* within brackets are the reactive sites in the first column in Table S4.1. ‘A’ stands for any aliphatic atom except for hydrogen. A bond denoted as ‘any’ shown in *sulfide_6cyc* and *sulfide_5cyc* indicates any bond between single, double, aromatic, and triple bonds. ^aA corresponding aromatic compound group to the substituent is to be assigned to the benzene ring (see Text S4.1 for more details).

Aromatic compounds

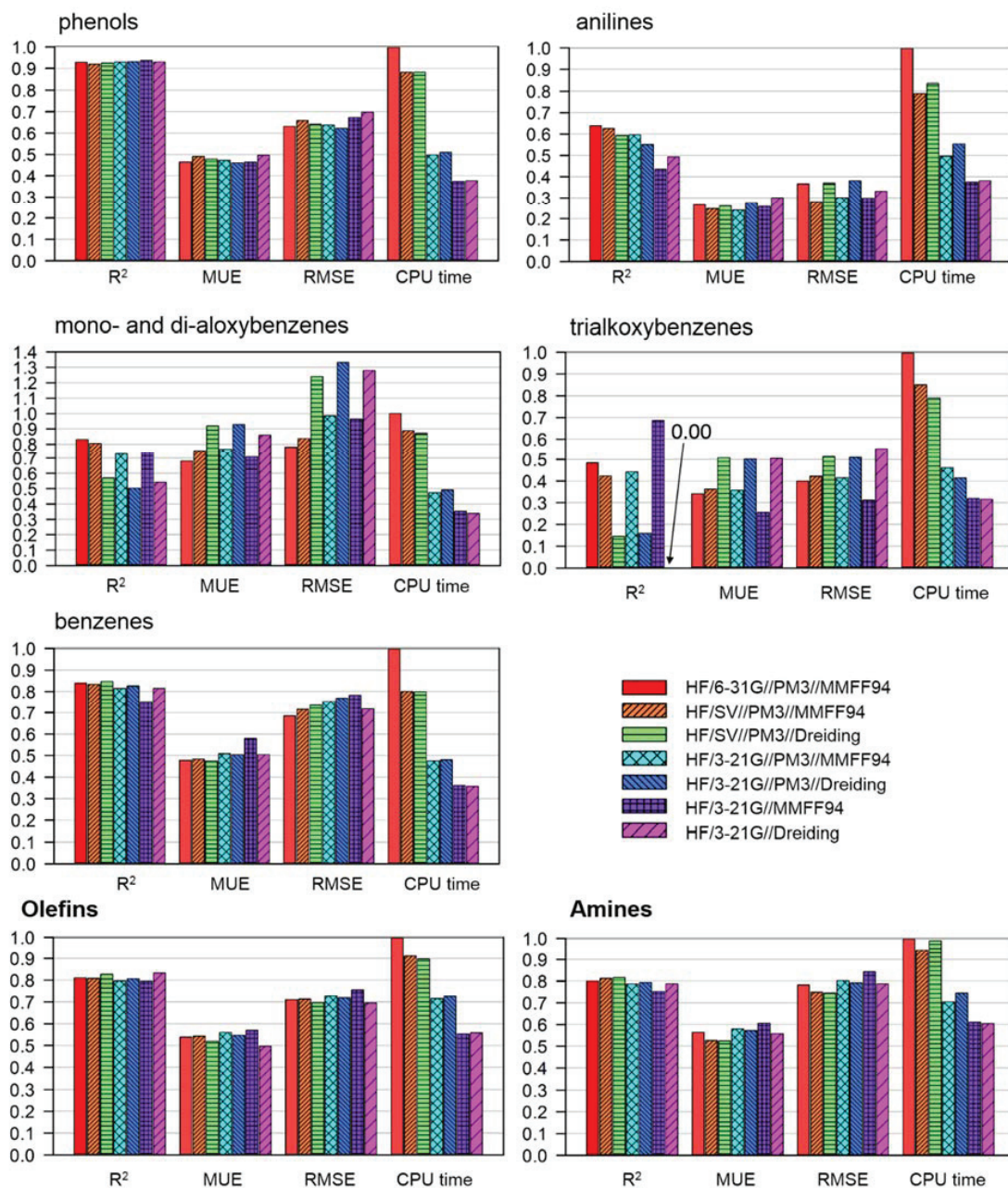


Figure S4.2. Comparison of the performances of the k_{O_3} prediction models for aromatic compound groups (phenols, anilines, mono- and di-alkoxybenzenes, trialkoxybenzenes, and benzenes), olefins, and amines developed with seven different computational methods. The relative CPU times were measured on the same hardware (Intel® Core™2 Duo CPU at 3.00 GHz with 4 GB RAM) operated on Ubuntu Linux 14.04.4 LTS. The numerical summary of R^2 , MUE, and RMSE for HF/3-21G//MMFF94 selected in this study for all prediction models are given in Table S4.3. The absolute CPU times are given in Figure S4.6 for the single point calculations of the selected method, i.e., HF/3-21G.

Heteroaromatic compounds

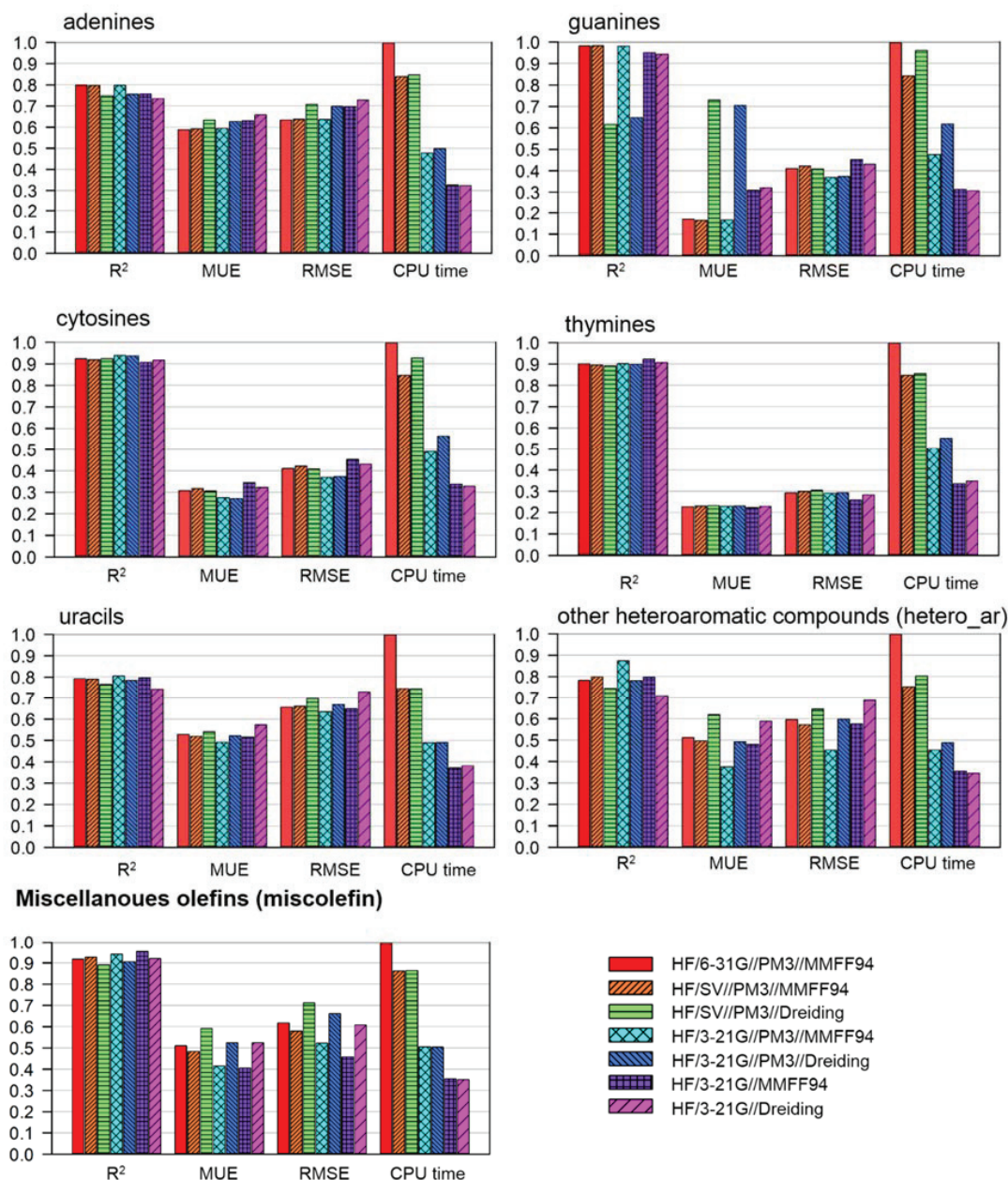


Figure S4.3. Comparison of the performances of the k_{O3} prediction models for adenines, guanines, cytosines, thymines, uracils, heteroaromatic compounds (hetero_ar), and miscellaneous olefins (miscolefin) developed with seven different computational methods. The relative CPU times were measured on the same hardware (Intel® Core™2 Duo CPU at 3.00 GHz with 4 GB RAM) operated on Ubuntu Linux 14.04.4 LTS. The numerical summary of R^2 , MUE, and RMSE for HF/3-21G//MMFF94 selected in this study for all prediction models are given in Table S4.3. The absolute CPU times are given in Figure S4.6 for the single point calculations of the selected method, i.e., HF/3-21G.

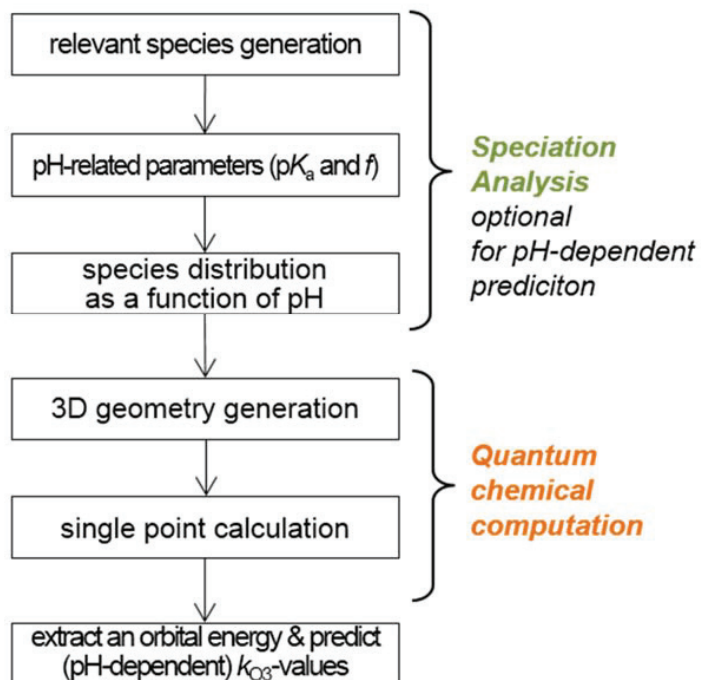
rate constant prediction

Figure S4.4. Workflow of the pH-dependent rate constant prediction

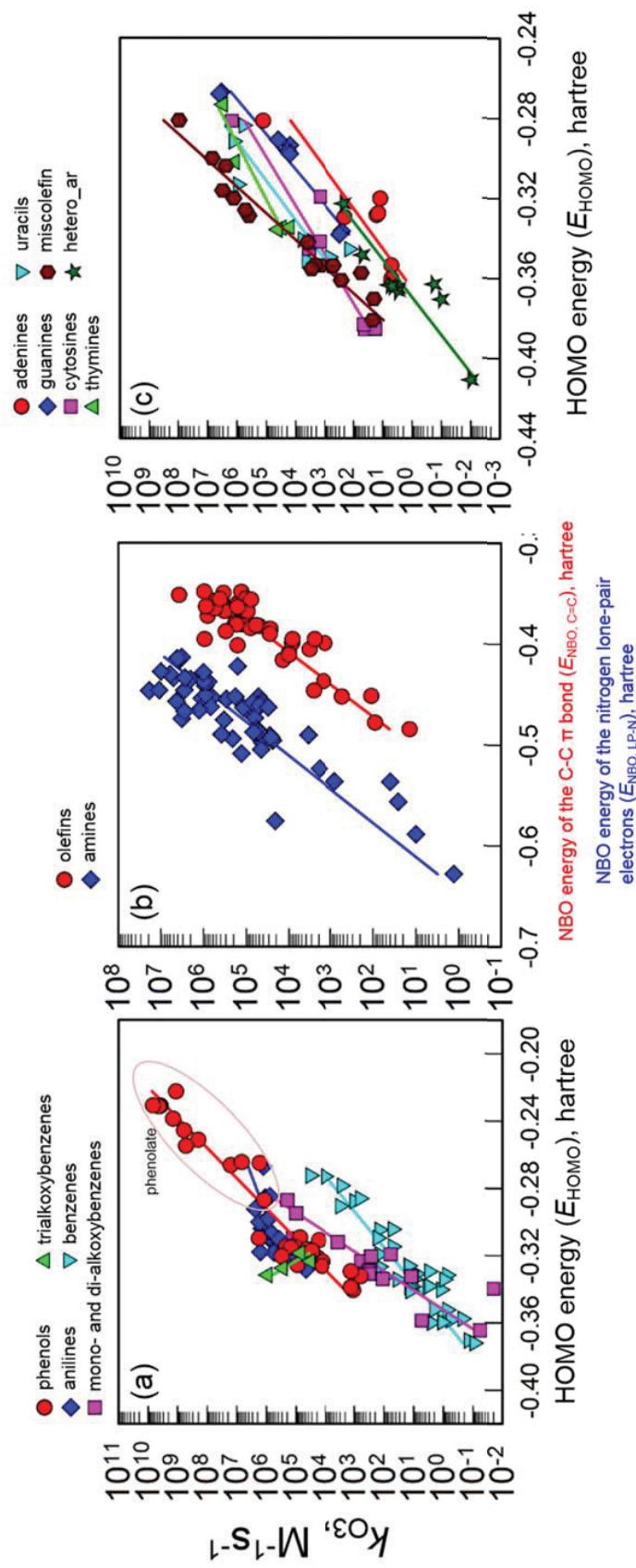


Figure S4.5. Correlations between experimental k_{O_3} for the reactions of ozone with (a) aromatic compound groups, (b) olefins, (b) amines, and (c) heteroaromatic compound groups with the corresponding orbital energies. All the orbital energies were obtained with the HF/3-21G//MMFF method. The statistical parameters for the prediction performance as well as slopes and y-intercept for the individual regression line equations are presented in Table S4.3.

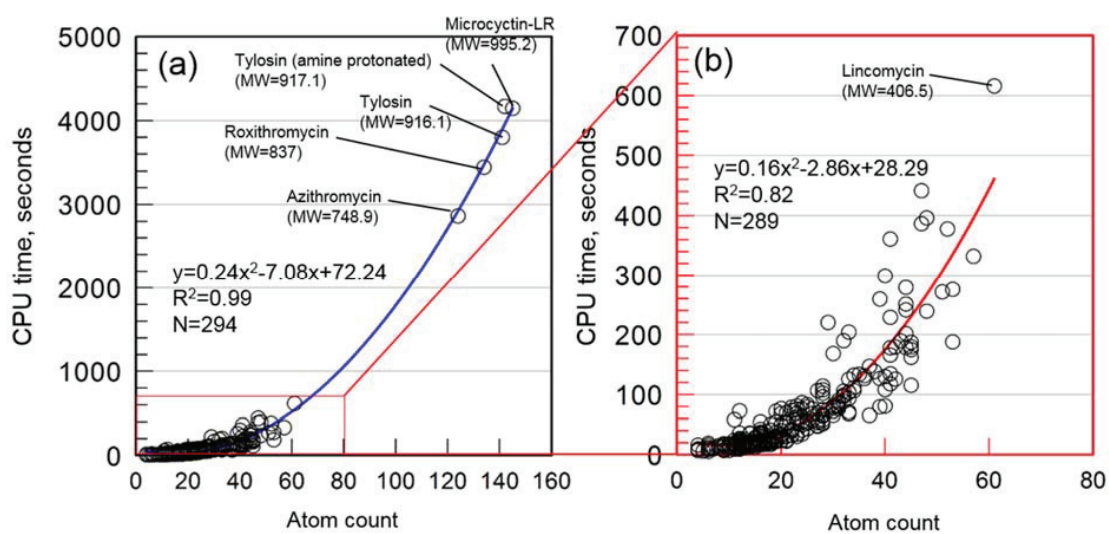


Figure S4.6. The CPU times in seconds for (a) 294 training compounds (10 outliers included, N_{total} in Table S4.3) and (b) 289 training compounds excluding 5 macromolecules from (a) as a function of atom count of the training compound. The CPU time was measured for single point calculations with the HF/3-21G. The CPU times were measured on the same hardware (Intel® Core™2 Duo CPU at 3.00 GHz with 4 GB RAM) operated on Ubuntu Linux 14.04.4 LTS.

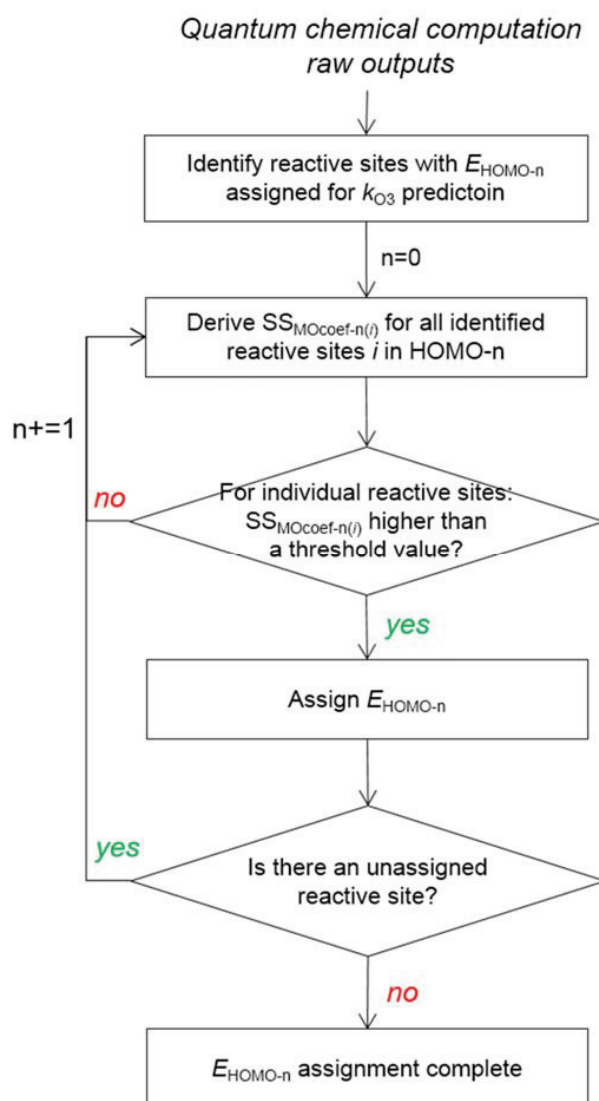


Figure S4.7. Workflow for the assignment of $E_{\text{HOMO}-n}$ to the corresponding reactive sites. $SS_{\text{MOcoef}-n(i)}$ is the sum of squares of molecular orbital coefficients of atoms of a reactive site (i) responsible for the reaction with ozone in HOMO- n .

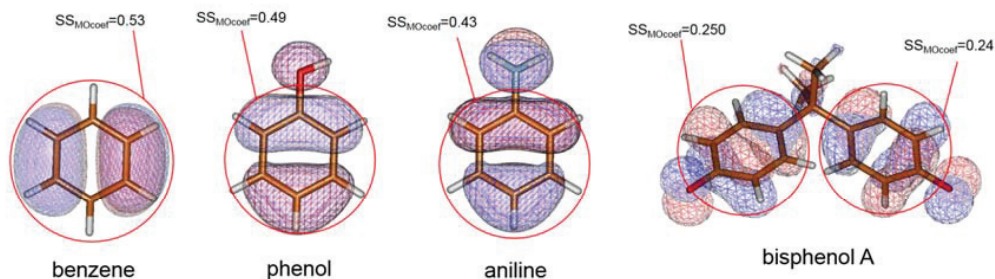
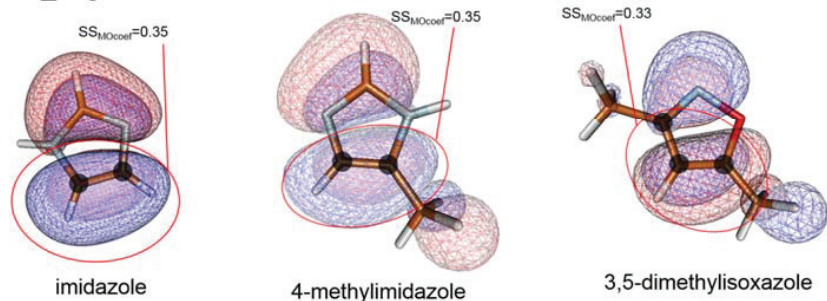
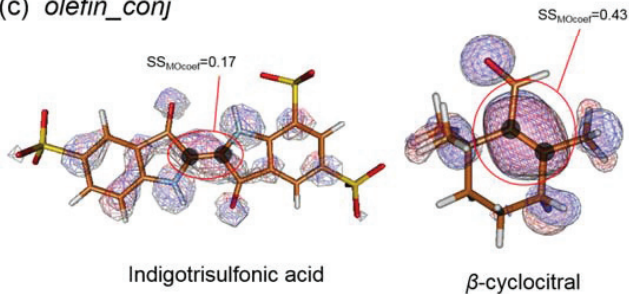
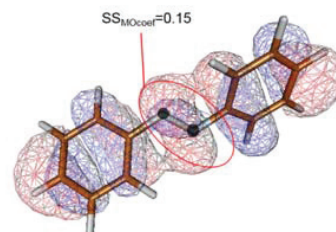
(a) aromatic compounds (*benzene*, *phenol*, *aniline*, and *bisphenol A*)(b) *5hetero_ring*(c) *olefin_conj*(d) *phenylazo*

Figure S4.8. HOMO and the estimated SS_{MOcoef} for the selected model compounds containing reactive sites such as (a) aromatic compounds (b) *5hetero_ring*, (c) *olefin_conj*, and (d) *phenylazo*. Computational method: HF/3-21G//MMFF94. The carbons with black circles are the atoms used to estimate SS_{MOcoef} for *5hetero_ring*, *olefin_conj*, and *phenylazo*.

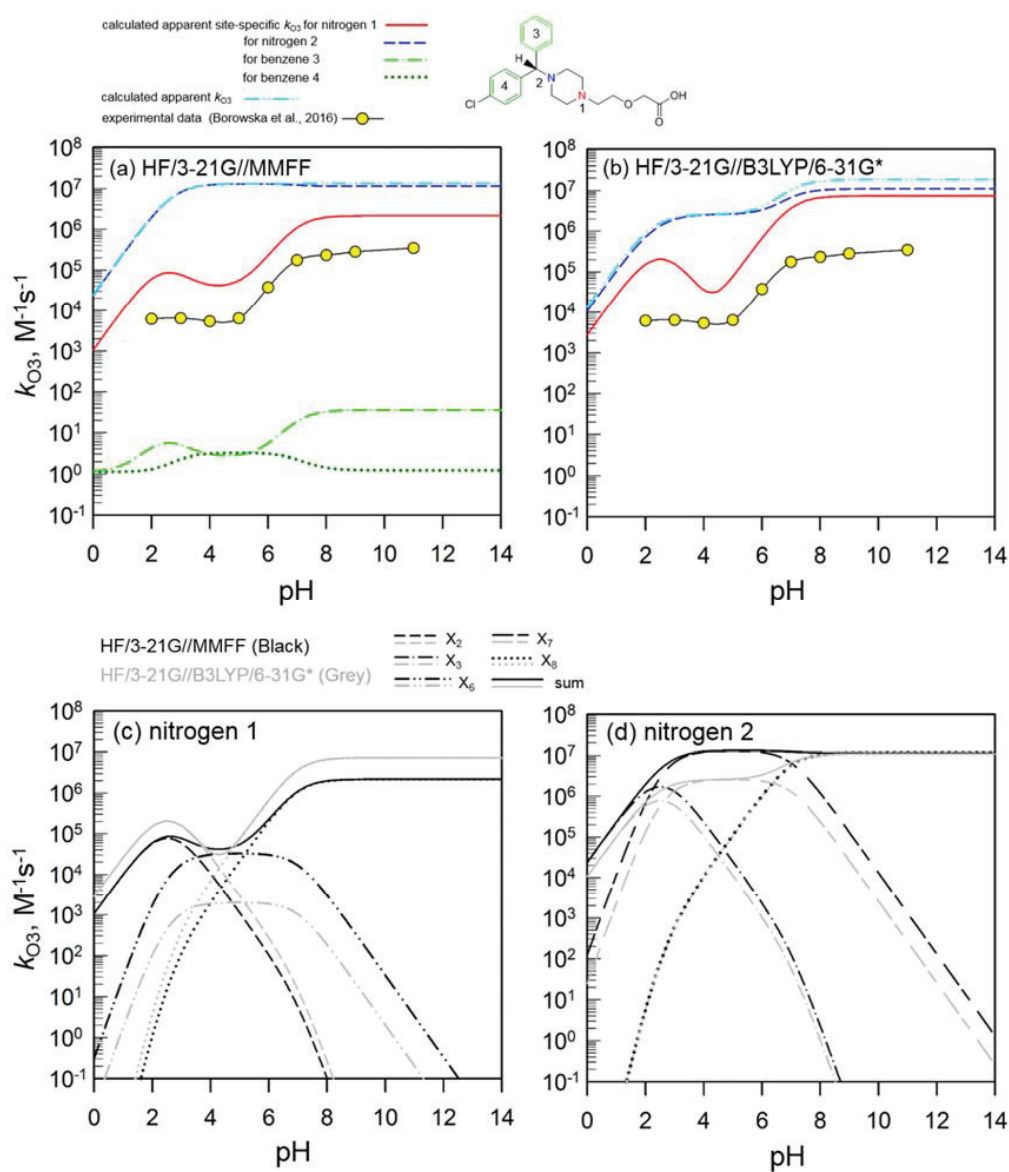
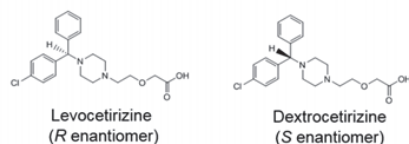
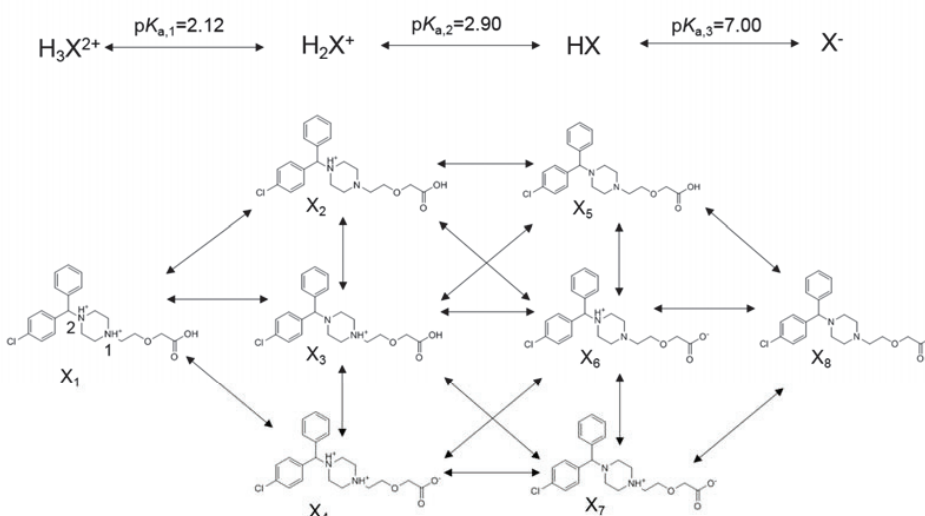


Figure S4.9. Apparent second order rate constants for the reaction of ozone with cetirizine as a function of the pH. (a) Comparison between predicted and experimental k_{O_3} -values, the prediction data were produced by the HF/3-21G//MMFF method. (b) Comparison between predicted and experimental k_{O_3} -values, the prediction data were produced by the HF/3-21G//B3LYP/6-31G*. (c) Comparison of the predicted k_{O_3} for nitrogen 1 of cetirizine for HF/3-21G//MMFF (black) and for HF/3-21G//B3LYP/6-31G* (grey). (d) Comparison of the predicted k_{O_3} for the nitrogen 2 of cetirizine for HF/3-21G//MMFF (black) and for HF/3-21G//B3LYP/6-31G* (grey). All the predicted apparent site-specific k_{O_3} were calculated based on Eq. S4.2 using the species distribution in Figure S4.10. The species X_i corresponds to the species shown in Figure S4.8(b). All the acid-base species of cetirizine used for the k_{O_3} prediction were dextrocetirizine (*S* enantiomer) (see Text S4.10 for details).

(a) cetirizine: a racemic mixture of two enantiomers



(b) acid-base speciation scheme for cetirizine



(c) species distribution for cetirizine as a function of pH

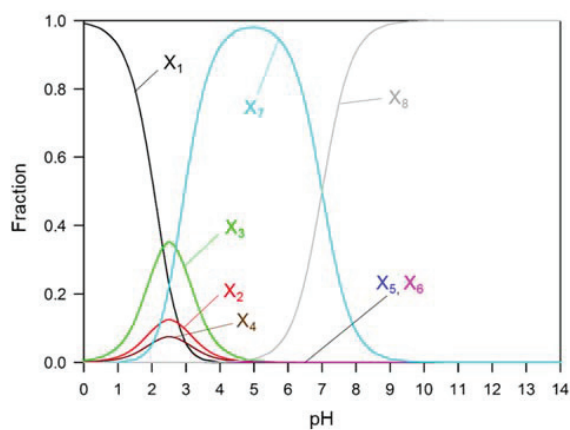


Figure S4.10. Relevant information about cetirizine for k_{O_3} prediction. (a) The chemical structures of a racemic mixture of cetirizine used for the determination of experimental k_{O_3} .² (b) Acid-base speciation scheme of cetirizine. $\text{p}K_{\text{a},1}$ and $\text{p}K_{\text{a},2}$ values were from ref⁹⁸ and $\text{p}K_{\text{a},3}$ were from ref.² (c) Species distribution for cetirizine as a function of pH between 0 and 14. Tautomeric fractions between X_2 , X_3 , and X_4 and X_5 , X_6 , and X_7 were 0.226, 0.638, and 0.136 and 0.000, 0.001, and 0.999, respectively, from the literature.⁹⁸

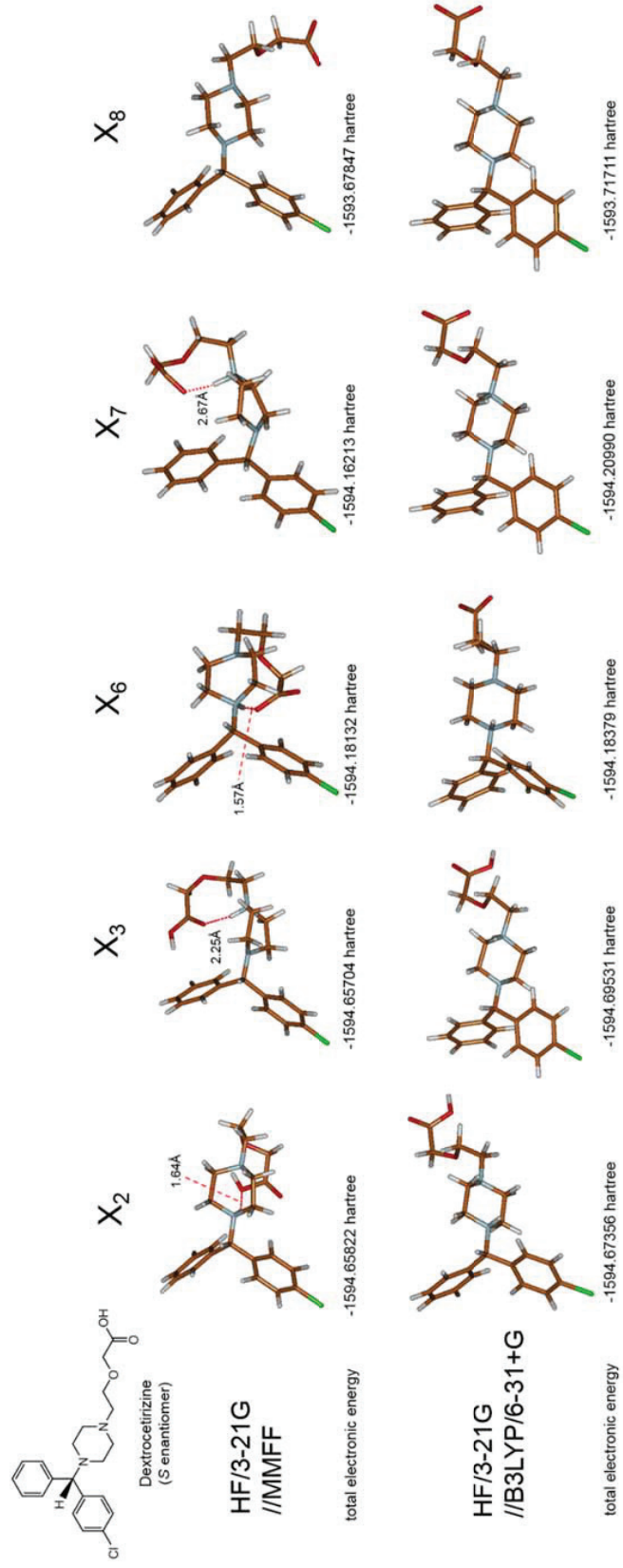


Figure S4.11. Comparison of geometries and total electronic energies of acid-base species of dextroctetrizine used for the k_{03} prediction presented in Figure S4.10a and S4.10b for the HF/3-21G//MMFF (top row) and HF/3-21G//B3LYP/6-31G* (bottom row), respectively.

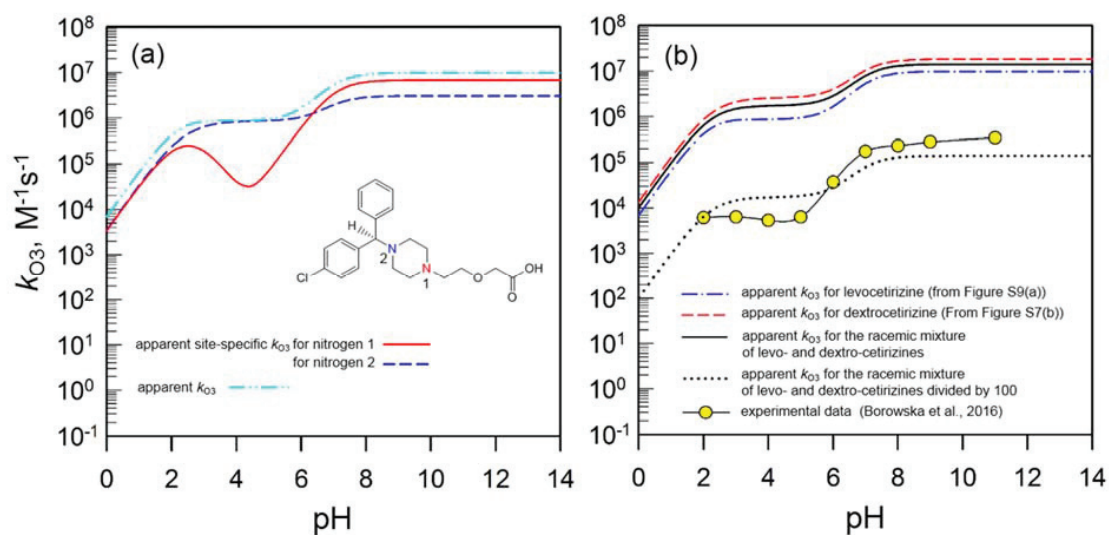


Figure S4.12. Apparent second-order rate constants (k_{O_3}) for the reaction of ozone with cetirizine as a function of pH. (a) predicted k_{O_3} obtained by the HF/3-21G//B3LYP/6-31G* method for levocetirizine (*R* enantiomer) and (b) comparison between the experimental k_{O_3} -values and the predicted k_{O_3} obtained by the HF/3-21G//B3LYP/6-31G* method for the racemic mixture of levocetirizine and dextrocetirizine. All the predicted apparent site-specific k_{O_3} were calculated based on Eq. S4.2 using the species distribution in Figure S4.10.

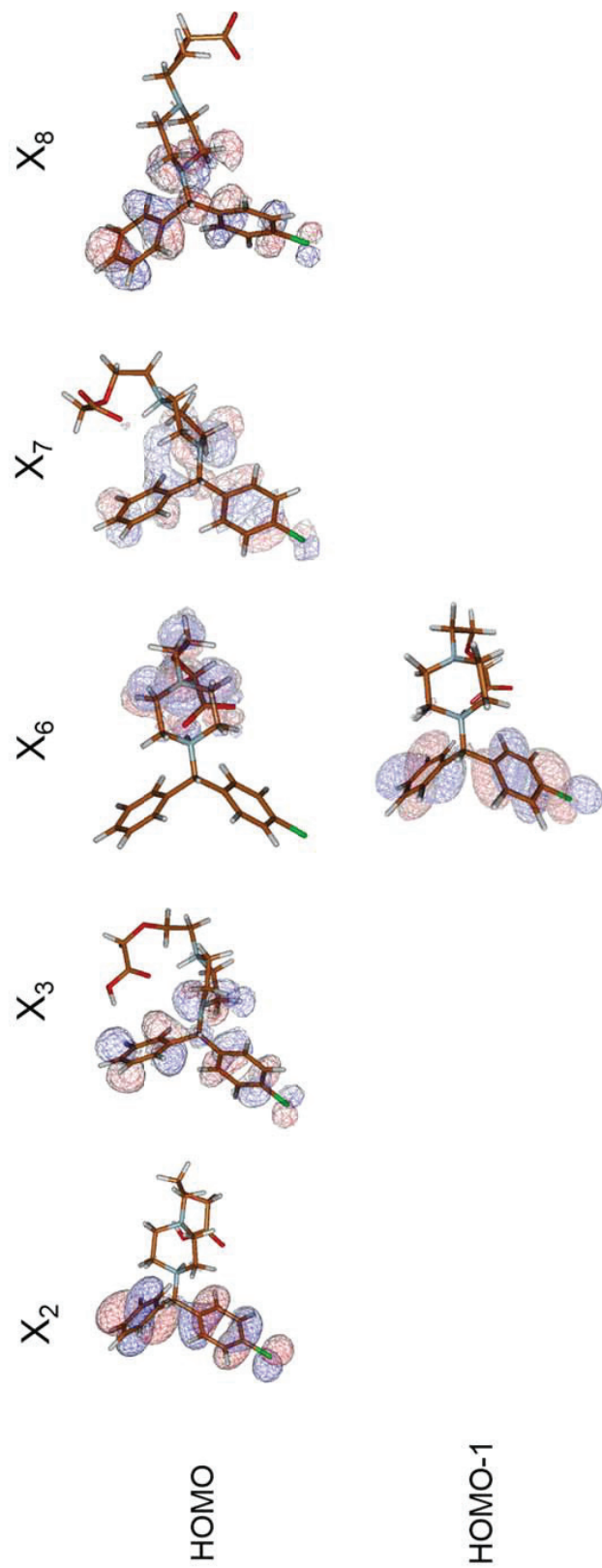
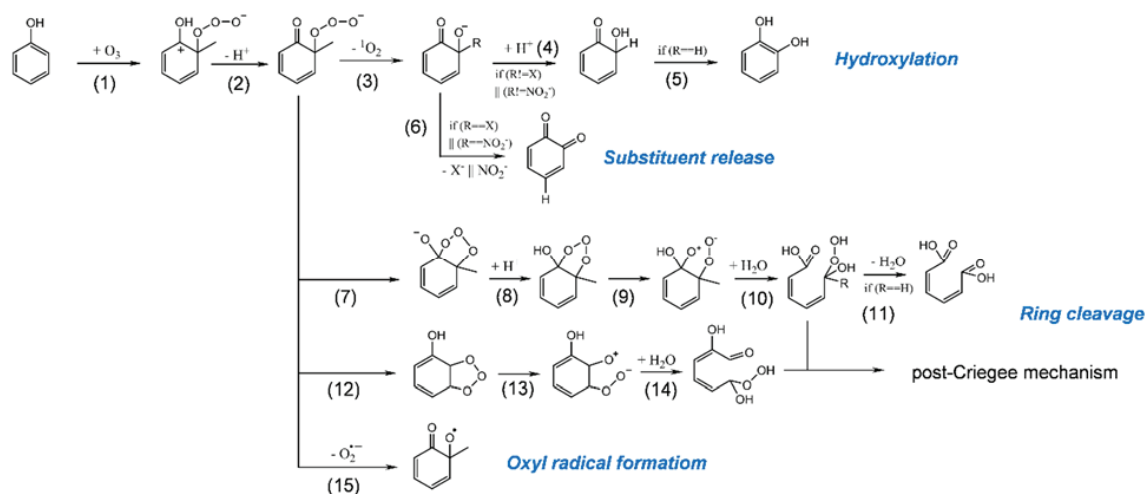
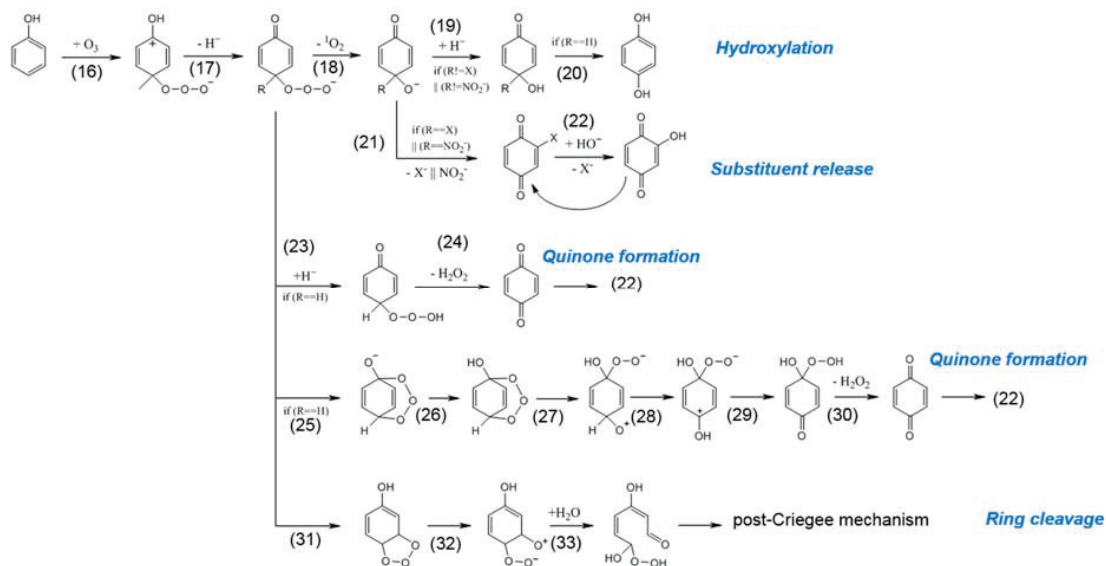


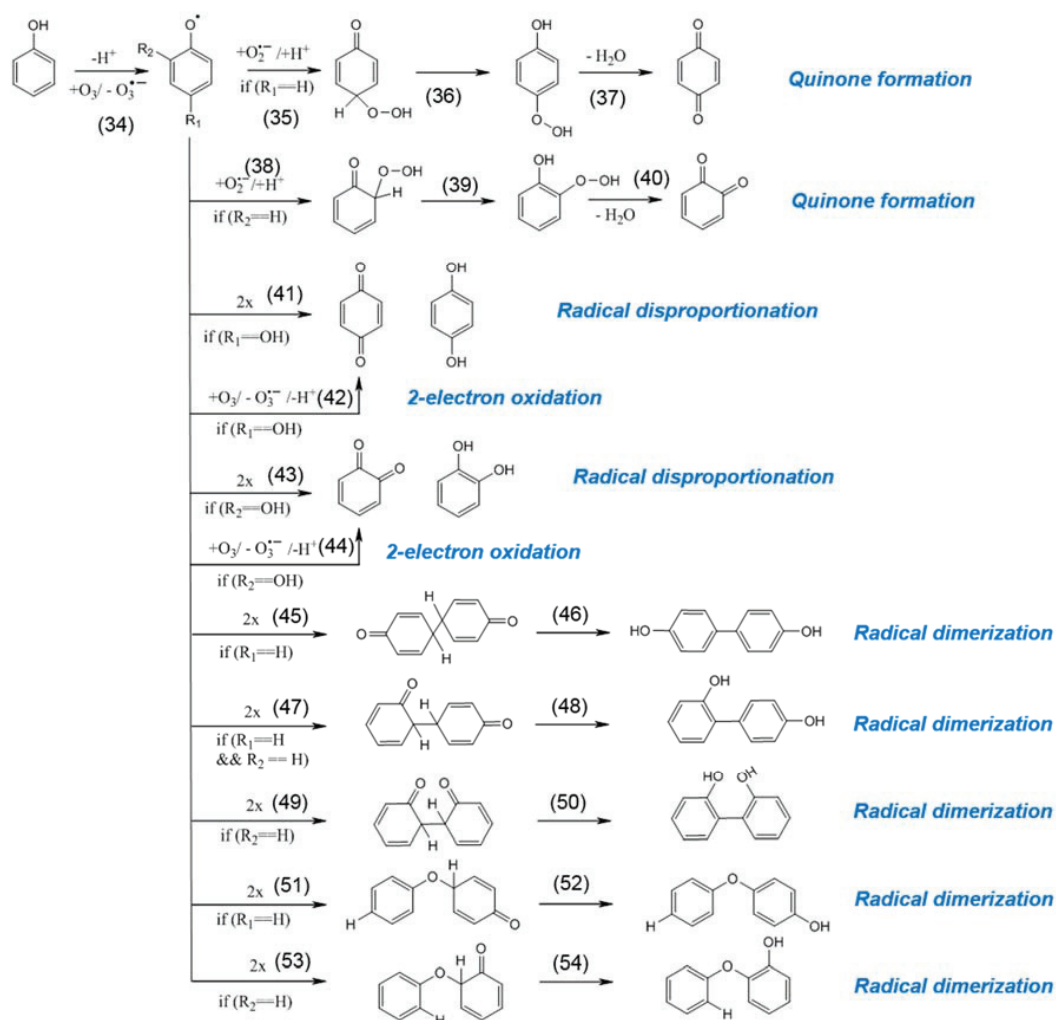
Figure S4.13. HOMO (and HOMO-1 for X₆) of the selected dextrocetirizine species obtained by the HF/3-21G//MMFF. Although not presented, a similar distribution of the HOMO was observed for other species, X₁, X₄, and X₅. The 2D chemical structure for dextrocetirizine is shown in Figure S4.10.



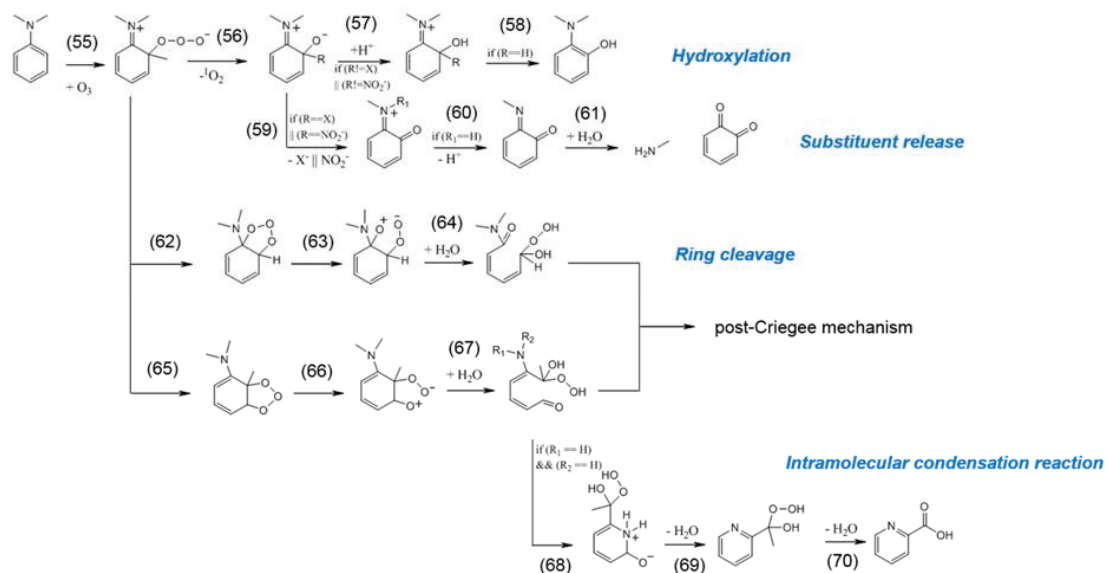
Scheme S4.1. Reaction pathways for the reactions of ozone with phenols at the ortho position adapted from various references (see Text S4.9) (logical operators such as ‘=’ (is equal to), ‘!=’ (is not equal to), and ‘||’ (or) are presented in ‘if ()’ statements). Implicit atoms can be hydrogen, any atom, or any organic moiety.



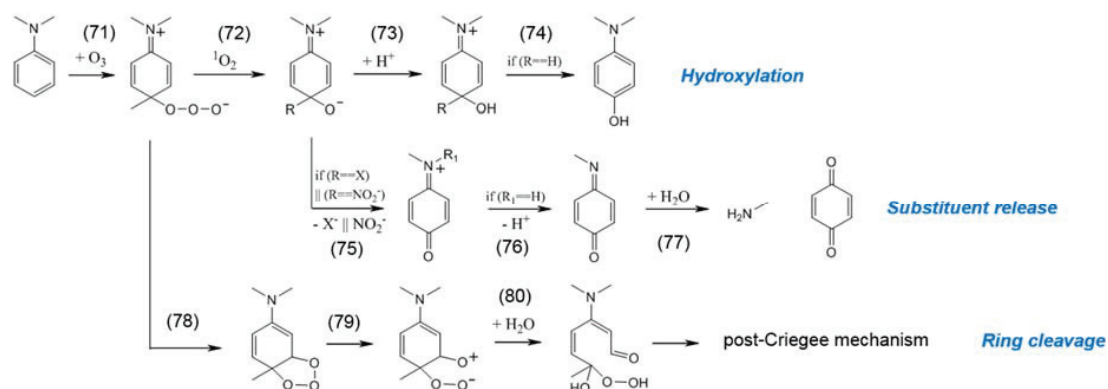
Scheme S4.2. Reaction pathways for the reactions of ozone with phenols at the para position adapted from various references (see Text S4.9) (logical operators such as ‘=’ (is equal to), ‘!=’ (is not equal to), and ‘||’ (or) are presented in ‘if ()’ statements). Implicit atoms can be hydrogen, any atom, or any organic moiety.



Scheme S4.3. Reaction pathways for the reactions of ozone with phenols via electron transfer adapted from various references (see Text S4.9) (logical operators such as ‘=’ (is equal to) and ‘&&’ (and) are presented in ‘if ()’ statements). Implicit atoms can be hydrogen, any atom, or any organic moiety.

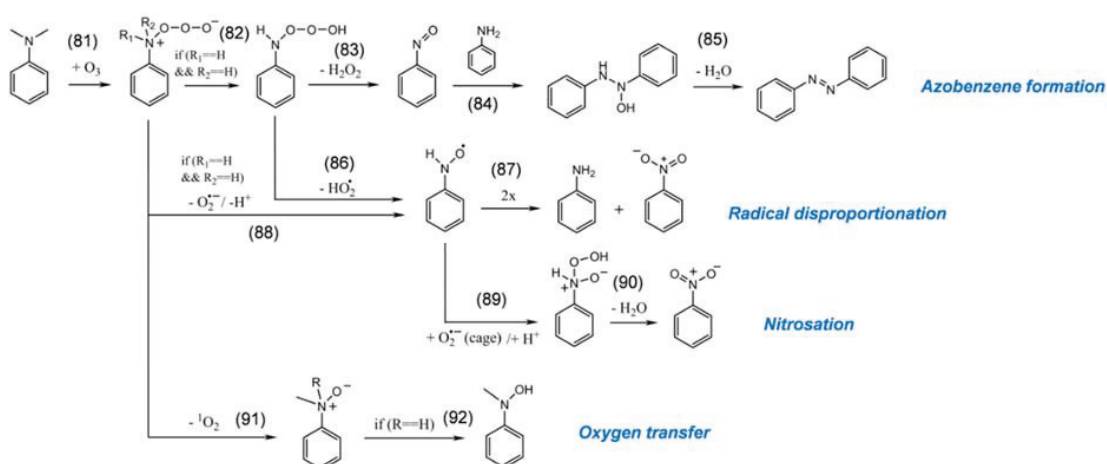


Scheme S4.4. Reaction pathways for the reactions of ozone with anilines at the ortho position adapted from various references (see Text S4.9) (logical operators such as ‘==’ (is equal to), ‘&&’ (and), and ‘||’ (or) are presented in ‘if ()’ statements). Implicit atoms can be hydrogen, any atom, or any organic moiety.

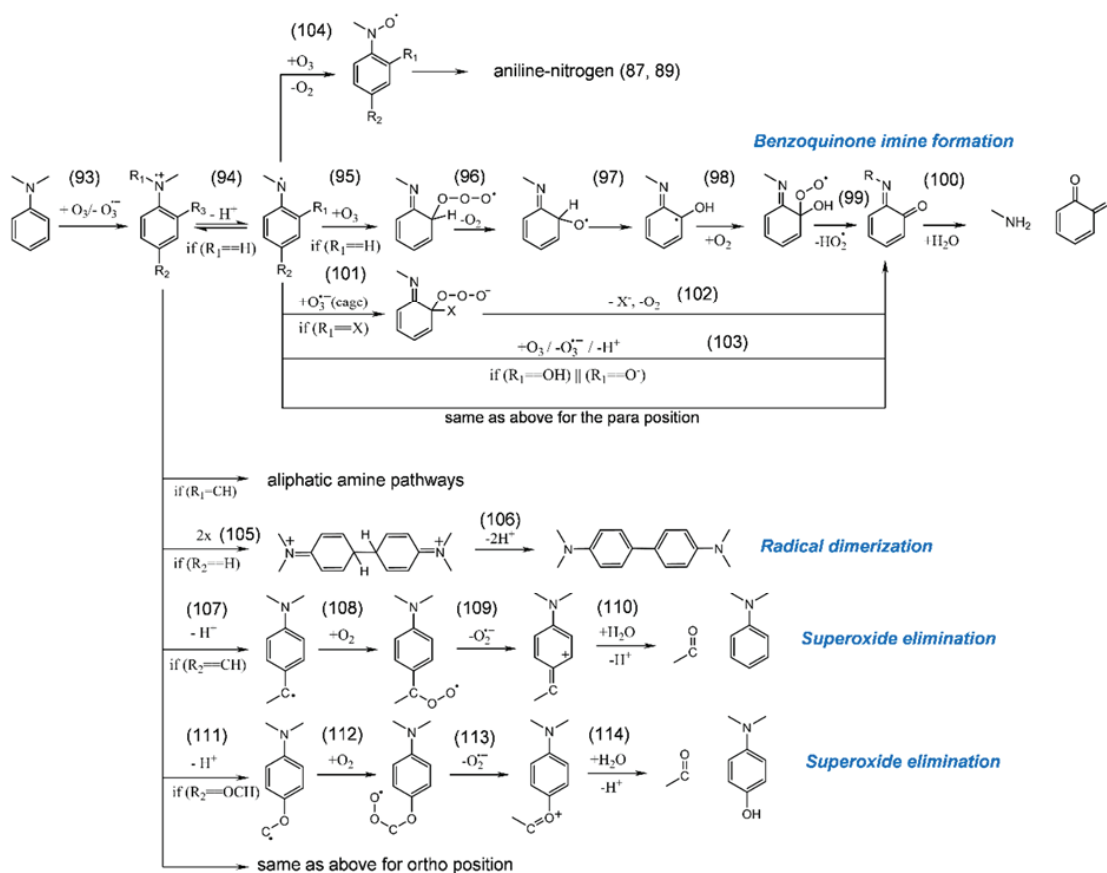


Scheme S4.5. Reaction pathways for the reactions of ozone with anilines at the para position adapted from various references (see Text S4.9) (logical operators such as ‘==’ (is equal to) and ‘||’ (or) are presented in ‘if ()’ statements). Implicit atoms can be hydrogen, any atom, or any organic moiety.

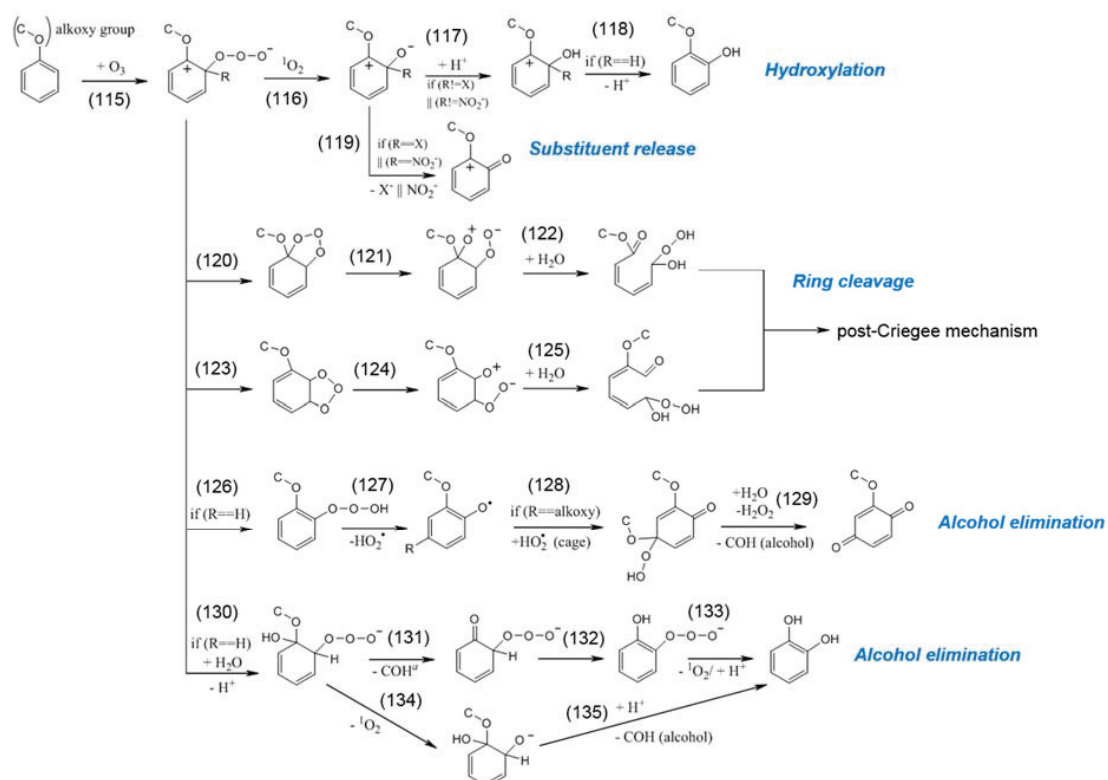
Supporting information for chapter 4



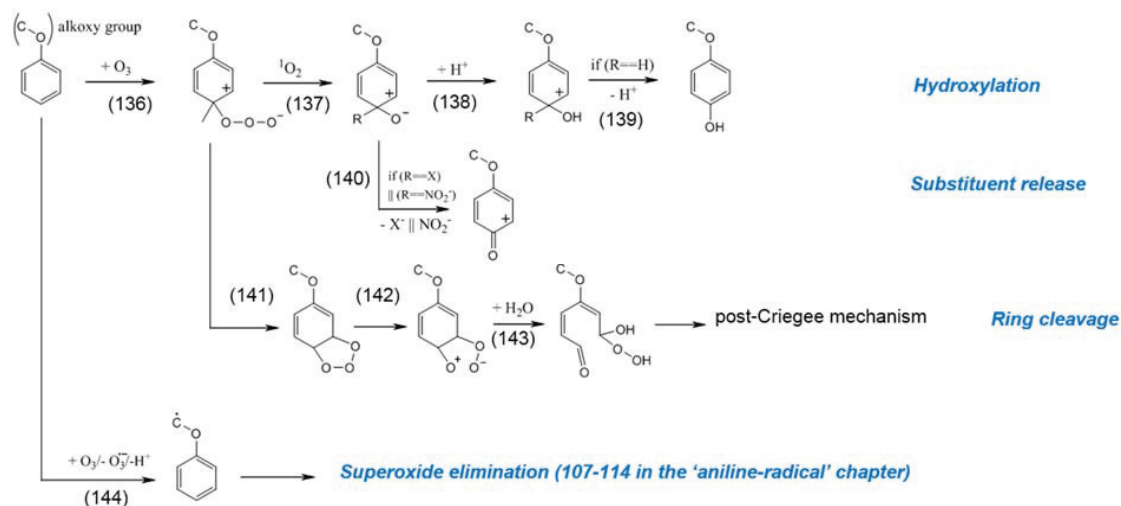
Scheme S4.6. Reaction pathways for the reactions of ozone with anilines at the exocyclic nitrogen adapted from various references (see Text S4.9) (logical operators such as ‘=’ (is equal to) and ‘&&’ (and) are presented in ‘if ()’ statements). Implicit atoms can be hydrogen, any atom, or any organic moiety.



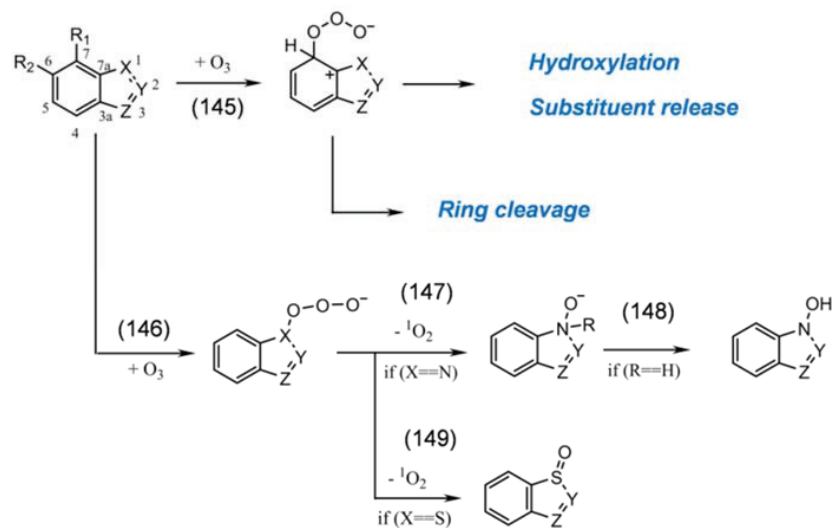
Scheme S4.7. Reaction pathways for the reactions of ozone with anilines via electron transfer adapted from various references (see Text S4.9) (logical operators such as ‘=’ (is equal to) and ‘||’ (or) are presented in ‘if ()’ statements). Implicit atoms can be hydrogen, any atom, or any organic moiety.



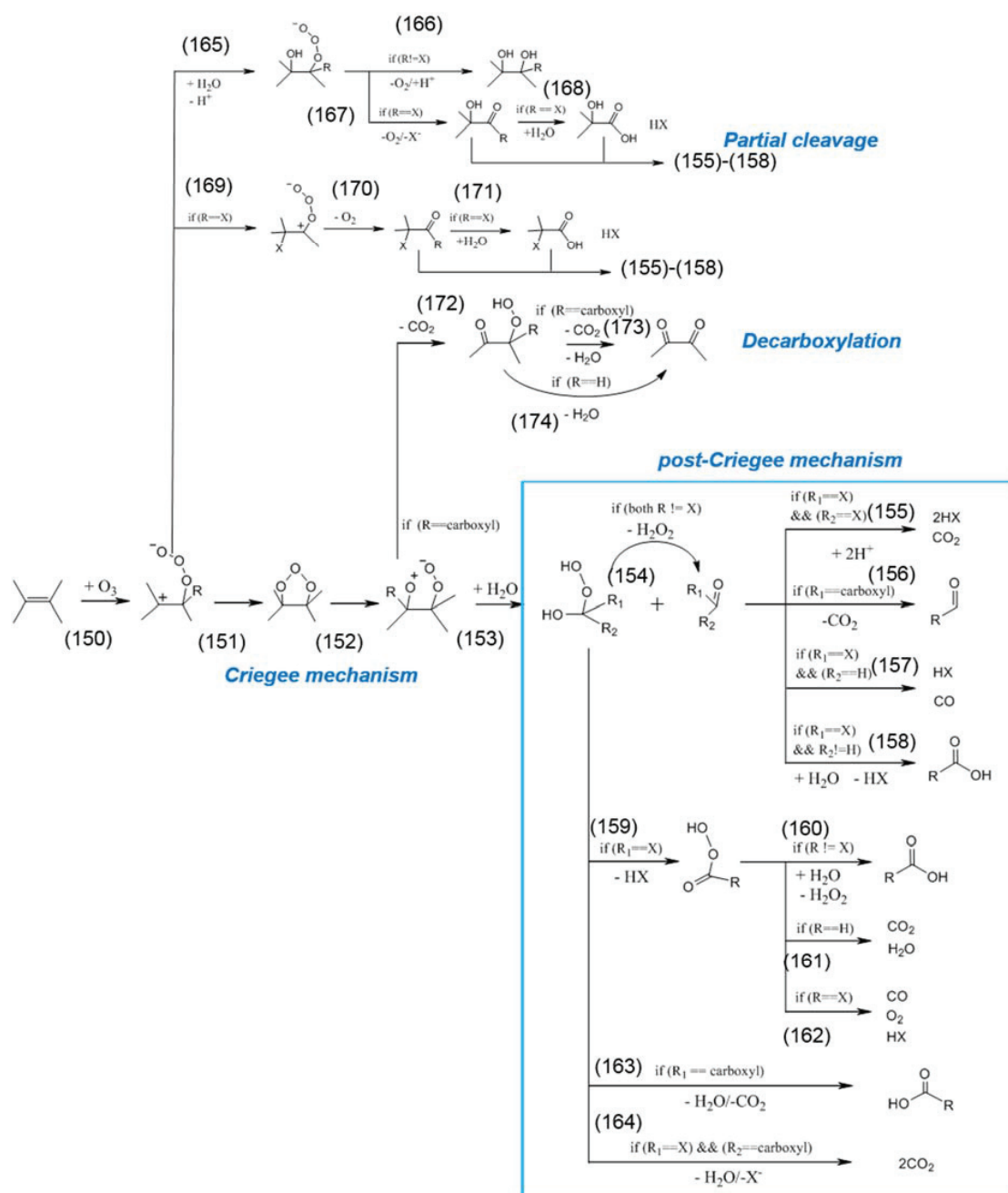
Scheme S4.8. Reaction pathways for the reactions of ozone with alkoxybenzenes at the ortho position adapted from various references (see Text S4.9) (logical operators such as ‘=’ (is equal to) and ‘||’ (or) are presented in ‘if ()’ statements). Implicit atoms can be hydrogen, any atom, or any organic moiety.



Scheme S4.9. Reaction pathways for the reactions of ozone with alkoxybenzenes at the para position or via electron transfer adapted from various references (see Text S4.9) (logical operators such as ‘=’ (is equal to) and ‘||’ (or) are presented in ‘if ()’ statements). Implicit atoms can be hydrogen, any atom, or any organic moiety.

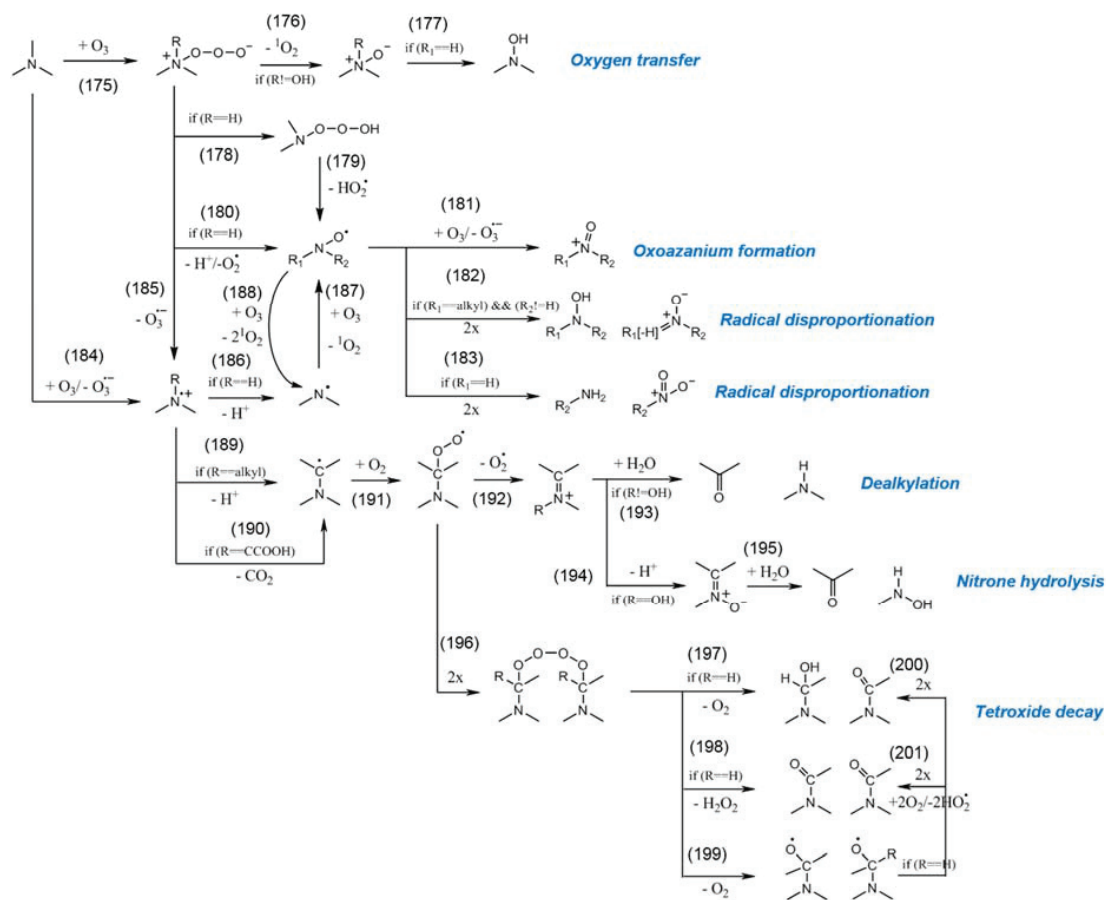


Scheme S4.10. Reaction pathways for the reactions of ozone with benzazoles adapted from various references (see Text S4.9) (a logical operator ‘==’ (is equal to) is presented in ‘if ()’ statements). Implicit atoms can be hydrogen, any atom, or any organic moiety.

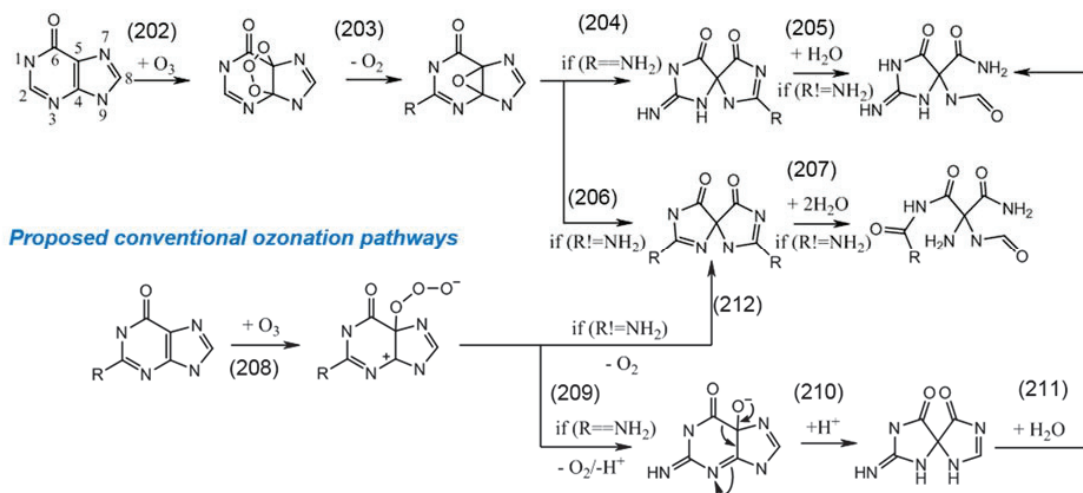


Scheme S4.11. Reaction pathways for the reactions of ozone with olefins adapted from various references (see Text S4.9) (logical operators such as ‘=’ (is equal to), ‘!=’ (is not equal to), ‘&&’ (and), and ‘||’ (or) are presented in ‘if ()’ statements). Implicit atoms can be hydrogen, any atom, or any organic moiety.

Supporting information for chapter 4

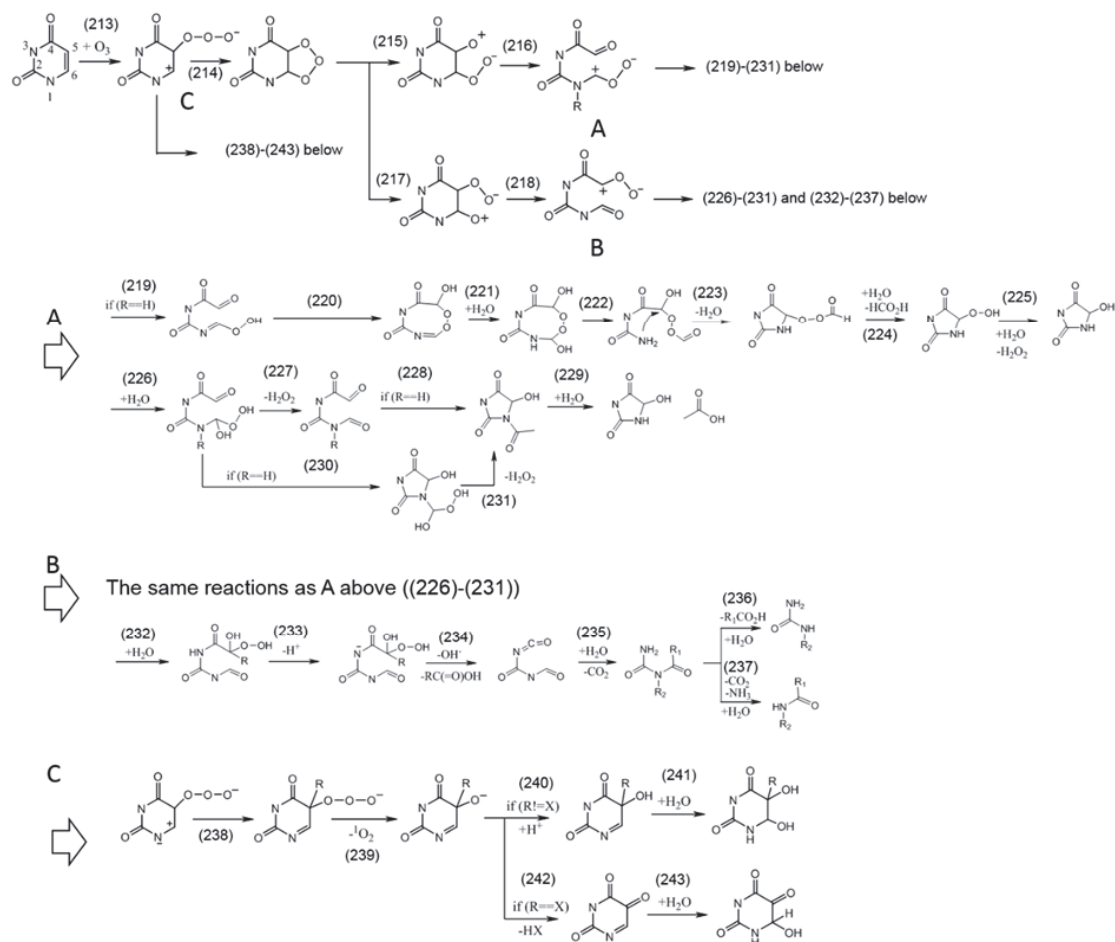


Scheme S4.12. Reaction pathways for the reactions of ozone with amines adapted from various references (see Text S4.9) (logical operators such as ‘==’ (is equal to), ‘!=’ (is not equal to), and ‘||’ (or) are presented in ‘if ()’ statements). Implicit atoms can be hydrogen, any atom, or any organic moiety.

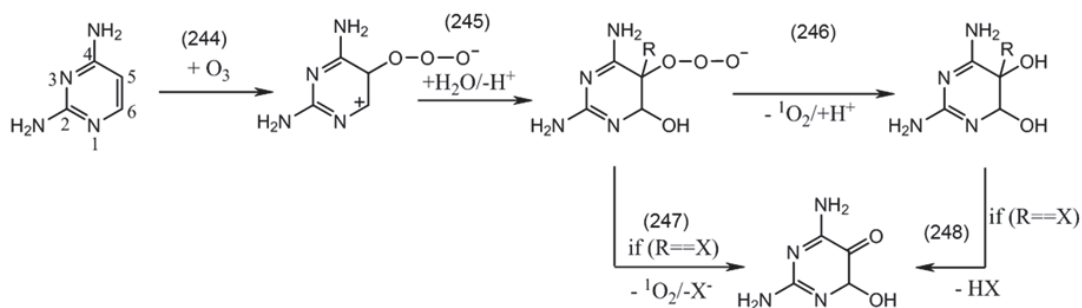


Scheme S4.13. Reaction pathways for the reactions of ozone with guanine analogues adapted from reference for reactions (202)-(207) and proposed herein for reactions (208)-(212) (logical operators such as '=' (is equal to) and '!=' (is not equal to) are presented in 'if ()' statements). Implicit atoms can be hydrogen, any atom, or any organic moiety.

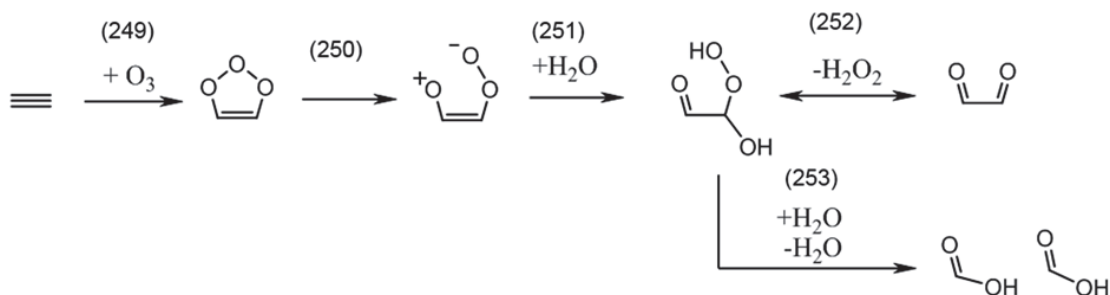
Supporting information for chapter 4



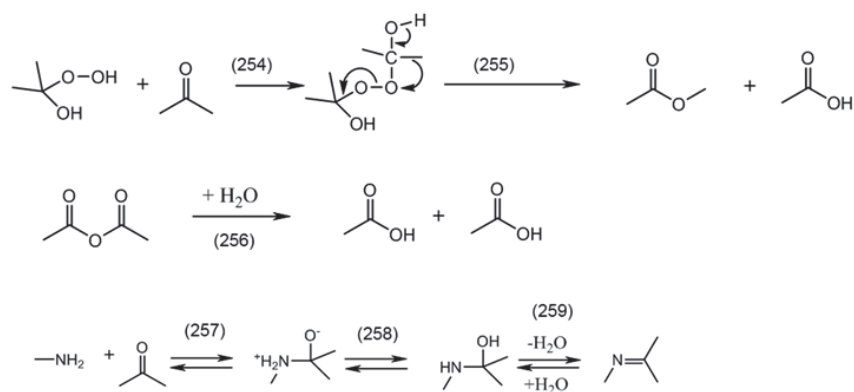
Scheme S4.14. Reaction pathways for the reaction of ozone with uracil adapted from the references^{63,88} (logical operators such as ‘==’ (is equal to) and ‘!=’ (is not equal to) are presented in ‘if ()’ statements). Implicit atoms can be hydrogen, any atom, or any organic moiety.



Scheme S4.15. Reaction pathways for the reactions of ozone with diaminopyrimidines proposed in the present study (a logical operator, ‘==’ (is equal to), is presented in a ‘if ()’ statement). Implicit atoms can be hydrogen, any atom, or any organic moiety.



Scheme S4.16. Reaction pathways for the reactions of ozone with an ethynyl group adapted from reference⁴⁶. Implicit atoms can be hydrogen, any atom, or any organic moiety.



Scheme S4.17. The selected post-ozonation reaction pathways adapted from various references (see Text S4.9). Implicit atoms can be hydrogen, any atom, or any organic moiety.

References

- (1) Sein, M. M.; Zedda, M.; Tuerk, J.; Schmidt, T. C.; Golloch, A.; von Sonntag, C. Oxidation of Diclofenac with Ozone in Aqueous Solution. *Environ. Sci. Technol.* **2008**, *42*, 6656–6662.
- (2) Borowska, E.; Bourgin, M.; Hollender, J.; Kienle, C.; McArdell, C. S.; von Gunten, U. Oxidation of cetirizine, fexofenadine and hydrochlorothiazide during ozonation: Kinetics and transformation products. *Water Res.* **2016**.
- (3) Dodd, M. C.; Buffle, M.-O.; von Gunten, U. Oxidation of Antibacterial Molecules by Aqueous Ozone: Moiety-Specific Reaction Kinetics and Application to Ozone-Based Wastewater Treatment. *Environ. Sci. Technol.* **2006**, *40*, 1969–1977.
- (4) Lee, M.; Zimmermann-Steffens, S. G.; Arey, J. S.; Fenner, K.; von Gunten, U. Development of Prediction Models for the Reactivity of Organic Compounds with Ozone in Aqueous Solution by Quantum Chemical Calculations: The Role of Delocalized and Localized Molecular Orbitals. *Environ. Sci. Technol.* **2015**, *49*, 9925–9935.
- (5) Barron, E.; Deborde, M.; Rabouan, S.; Mazellier, P.; Legube, B. Kinetic and mechanistic investigations of progesterone reaction with ozone. *Water Res.* **2006**, *40*, 2181–2189.
- (6) Pryor, W. A.; Giamalva, D. H.; Church, D. F. Kinetics of ozonation. 2. Amino acids and model compounds in water and comparisons to rates in nonpolar solvents. *J. Am. Chem. Soc.* **1984**, *106*, 7094–7100.
- (7) Flyunt, R.; Leitzke, A.; Mark, G.; Mvula, E.; Reisz, E.; Schick, R.; von Sonntag, C. Determination of OH, O₂⁻, and Hydroperoxide Yields in Ozone Reactions in Aqueous Solution. *J. Phys. Chem. B* **2003**, *107*, 7242–7253.
- (8) Lee, C.; Yoon, J.; von Gunten, U. Oxidative degradation of N-nitrosodimethylamine by conventional ozonation and the advanced oxidation process ozone/hydrogen peroxide. *Water Res.* **2007**, *41*, 581–590.
- (9) Hoigné, J.; Bader, H. Rate constants of reactions of ozone with organic and inorganic compounds in water—II. Dissociating organic compounds. *Water Res.* **1983**, *17*, 185–194.
- (10) Yan, S.; Jia, A.; Merel, S.; Snyder, S. A.; O'Shea, K. E.; Dionysiou, D. D.; Song, W. Ozonation of Cylindrospermopsin (Cyanotoxin): Degradation Mechanisms and Cytotoxicity Assessments. *Environ. Sci. Technol.* **2016**, acs.est.5b04540.
- (11) von Sonntag, C.; von Gunten, U. *Chemistry of ozone in water and wastewater treatment: From basic principles to applications*; IWA publishing, 2012.
- (12) Prasse, C.; Wagner, M.; Schulz, R.; Ternes, T. A. Oxidation of the antiviral drug acyclovir and its biodegradation product carboxy-acyclovir with ozone: kinetics and identification of oxidation products. *Environ. Sci. Technol.* **2012**, *46*, 2169–2178.
- (13) Pryor, W. A.; Uppu, R. M. A kinetic model for the competitive reactions of ozone with amino acid residues in proteins in reverse micelles. *J. Biol. Chem.* **1993**, *268*, 3120–3126.
- (14) Kotiaho, T.; Eberlin, M. N.; Vainiotalo, P.; Kostianinen, R. Electrospray mass and tandem mass spectrometry identification of ozone oxidation products of amino acids and small peptides. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 526–535.
- (15) Leitzke, A.; Flyunt, R.; Theruvathu, J. A.; von Sonntag, C. Ozonolysis of vinyl compounds, CH₂=CH-X, in aqueous solution—the chemistries of the ensuing formyl compounds and hydroperoxides. *Org. Biomol. Chem.* **2003**, *1*, 1012–1019.
- (16) DeWitte, B.; Dewulf, J.; Demeestere, K.; Van De Vyvere, V.; De Wispelaere, P.; Van Langenhove, H. Ozonation of Ciprofloxacin in Water: HRMS Identification of Reaction Products and Pathways. *Environ. Sci. Technol.* **2008**, *42*, 4889–4895.
- (17) MarvinBeans/JChem suite (16.2.29.0), 2016, Chemaxon (<http://www.chemaxon.com>).
- (18) Neese, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (19) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenb, 2009. Gaussian 09, Revision D.01.
- (20) Glendening, E. D.; Landis, C. R.; Weinhold, F. NBO 6.0: Natural bond orbital analysis

- program. *J. Comput. Chem.* **2013**, *34*, 1429–1437.
- (21) Halgren, T. A. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *J. Comput. Chem.* **1996**, *17*, 490–519.
 - (22) Halgren, T. A. Merck molecular force field. III. Molecular geometries and vibrational frequencies for MMFF94. *J. Comput. Chem.* **1996**, *17*, 553–586.
 - (23) Halgren, T. A. Merck molecular force field. V. Extension of MMFF94 using experimental data, additional computational data, and empirical rules. *J. Comput. Chem.* **1996**, *17*, 616–641.
 - (24) Halgren, T. A. Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions. *J. Comput. Chem.* **1996**, *17*, 520–552.
 - (25) Halgren, T. A.; Nachbar, R. B. Merck molecular force field. IV. conformational energies and geometries for MMFF94. *J. Comput. Chem.* **1996**, *17*, 587–615.
 - (26) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. DREIDING: a generic force field for molecular simulations. *J. Phys. Chem.* **1990**, *94*, 8897–8909.
 - (27) Stewart, J. J. P. Optimization of parameters for semiempirical methods II. Applications. *J. Comput. Chem.* **1989**, *10*, 221–264.
 - (28) Stewart, J. J. P. Optimization of parameters for semiempirical methods I. Method. *J. Comput. Chem.* **1989**, *10*, 209–220.
 - (29) Binkley, J. S.; Pople, J. A.; Hehre, W. J. Self-consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements. *J. Am. Chem. Soc.* **1980**, *102*, 939–947.
 - (30) Gordon, M. S.; Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. Self-consistent molecular-orbital methods. 22. Small split-valence basis sets for second-row elements. *J. Am. Chem. Soc.* **1982**, *104*, 2797–2803.
 - (31) Dobbs, K. D.; Hehre, W. J. Molecular orbital theory of the properties of inorganic and organometallic compounds 4. Extended basis sets for third-and fourth-row, main-group elements. *J. Comput. Chem.* **1986**, *7*, 359–378.
 - (32) Dobbs, K. D.; Hehre, W. J. Molecular orbital theory of the properties of inorganic and organometallic compounds 5. Extended basis sets for first-row transition metals. *J. Comput. Chem.* **1987**, *8*, 861–879.
 - (33) Dobbs, K. D.; Hehre, W. J. Molecular orbital theory of the properties of inorganic and organometallic compounds. 6. Extended basis sets for second-row transition metals. *J. Comput. Chem.* **1987**, *8*, 880–893.
 - (34) Schäfer, A.; Horn, H.; Ahlrichs, R. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.* **1992**, *97*, 2571.
 - (35) Hay, P. J.; Wadt, W. R. Ab initio effective core potentials for molecular calculations. Potentials for K to Au including the outermost core orbitals. *J. Chem. Phys.* **1985**, *82*, 299.
 - (36) Hay, P. J.; Wadt, W. R. Ab initio effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg. *J. Chem. Phys.* **1985**, *82*, 270.
 - (37) Wadt, W. R.; Hay, P. J. Ab initio effective core potentials for molecular calculations. Potentials for main group elements Na to Bi. *J. Chem. Phys.* **1985**, *82*, 284.
 - (38) Tekle-Röttering, A.; von Sonntag, C.; Reisz, E.; Eyser, C. vom; Lutze, H. V.; Türk, J.; Naumov, S.; Schmidt, W.; Schmidt, T. C. Ozonation of anilines: Kinetics, stoichiometry, product identification and elucidation of pathways. *Water Res.* **2016**, *98*, 147–159.
 - (39) Pierpoint, A. C.; Hapeman, C. J.; Torrents, A. Linear free energy study of ring-substituted aniline ozonation for developing treatment of aniline-based pesticide wastes. *J. Agric. Food Chem.* **2001**, *49*, 3827–3832.
 - (40) Lee, C.; Schmidt, C.; Yoon, J.; von Gunten, U. Oxidation of N-Nitrosodimethylamine (NDMA) Precursors with Ozone and Chlorine Dioxide: Kinetics and Effect on NDMA Formation Potential. *Environ. Sci. Technol.* **2007**, *41*, 2056–2063.
 - (41) Benner, J.; Ternes, T. A. Ozonation of Metoprolol: Elucidation of Oxidation Pathways and Major Oxidation Products. *Environ. Sci. Technol.* **2009**, *43*, 5472–5480.
 - (42) Coelho, A. D.; Sans, C.; Agüera, A.; Gómez, M. J.; Esplugas, S.; Dezotti, M. Effects of ozone pre-treatment on diclofenac: Intermediates, biodegradability and toxicity assessment. *Sci. Total Environ.* **2009**, *407*, 3572–3578.
 - (43) Dantas, R. F.; Canterino, M.; Marotta, R.; Sans, C.; Esplugas, S.; Andreozzi, R. Bezafibrate

- removal by means of ozonation: Primary intermediates, kinetics, and toxicity assessment. *Water Res.* **2007**, *41*, 2525–2532.
- (44) Deborde, M.; Rabouan, S.; Mazellier, P.; Duguet, J.-P.; Legube, B. Oxidation of bisphenol A by ozone in aqueous solution. *Water Res.* **2008**, *42*, 4299–4308.
- (45) Decoret, C.; Royer, J.; Legube, B.; Dore, M. Experimental and theoretical studies of the mechanism of the initial attack of ozone on some aromatics in aqueous medium. *Environ. Technol. Lett.* **1984**, *5*, 207–218.
- (46) Huber, M. M.; Ternes, T. A.; von Gunten, U. Removal of Estrogenic Activity and Formation of Oxidation Products during Ozonation of 17 α -Ethinylestradiol. *Environ. Sci. Technol.* **2004**, *38*, 5177–5186.
- (47) Jonsson, M.; Lind, J.; Eriksen, T. E.; Merenyi, G. Redox and Acidity Properties of 4-Substituted Aniline Radical Cations in Water. *J. Am. Chem. Soc.* **1994**, *116*, 1423–1427.
- (48) Mvula, E.; von Sonntag, C. Ozonolysis of phenols in aqueous solution. *Org. Biomol. Chem.* **2003**, *1*, 1749–1756.
- (49) Nogami, T.; Hishida, T.; Yamada, M.; Mikawa, H.; Shiota, Y. Formations and Reactions of o-Benzoquinone Mono- and Di-imines. I. *Bull. Chem. Soc. Jpn.* **1975**, *48*, 3709–3714.
- (50) Li, H.; Guo, A.; Wang, H. Mechanisms of oxidative browning of wine. *Food Chem.* **2008**, *108*, 1–13.
- (51) Davies, R.; Frahn, J. L. Addition of primary aliphatic amines to 1,2-benzoquinone. The absence of reaction between a secondary amide and 1,2-benzoquinone. *J. Chem. Soc. Perkin Trans. 1* **1977**, 2295.
- (52) Singleton, V. L.; Salgues, M.; Zaya, J.; Trousdale, E. Caftaric acid disappearance and conversion to products of enzymic oxidation in grape must and wine. *Am. J. Enol. Vitic.* **1985**, *36*, 50–56.
- (53) Benner, J.; Ternes, T. A. Ozonation of Propranolol: Formation of Oxidation Products. *Environ. Sci. Technol.* **2009**, *43*, 5086–5093.
- (54) Mawhinney, D. B.; Vanderford, B. J.; Snyder, S. A. Transformation of 1 H -Benzotriazole by Ozone in Aqueous Solution. *Environ. Sci. Technol.* **2012**, *46*, 7102–7111.
- (55) Ramseier, M. K.; von Gunten, U. Mechanisms of Phenol Ozonation—Kinetics of Formation of Primary and Secondary Reaction Products. *Ozone Sci. Eng.* **2009**, *31*, 201–215.
- (56) Sotelo, J. L.; Beltran, F. J.; Gonzalez, M. Ozonation of aqueous solutions of resorcinol and phloroglucinol. I. Stoichiometry and absorption kinetic regime. *Ind. Eng. Chem. Res.* **1990**, *29*, 2358–2367.
- (57) Dixon, W. T.; Murphy, D. Determination of the acidity constants of some phenol radical cations by means of electron spin resonance. *J. Chem. Soc. Faraday Trans. 2* **1976**, *72*, 1221.
- (58) Hoigne, J.; Bader, H. Ozonation of Water: Role of Hydroxyl Radicals as Oxidizing Intermediates. *Science (80-)*. **1975**, *190*, 782–784.
- (59) Sarasa, J. Study of the aromatic by-products formed from ozonation of anilines in aqueous solution. *Water Res.* **2002**, *36*, 3035–3044.
- (60) Turhan, K.; Uzman, S. The Degradation Products of Aniline in the Solutions with Ozone and Kinetic Investigations. *Ann. Chim.* **2007**, *97*, 1129–1138.
- (61) Dalmagro, J.; Yunes, R. A.; Simionatto, E. L. Mechanism of reaction of azobenzene formation from aniline and nitrosobenzene in basic conditions. General base catalysis by hydroxyide ion. *J. Phys. Org. Chem.* **1994**, *7*, 399–402.
- (62) Zimmermann, S. G.; Schmukat, A.; Schulz, M.; Benner, J.; von Gunten, U.; Ternes, T. A. Kinetic and Mechanistic Investigations of the Oxidation of Tramadol by Ferrate and Ozone. *Environ. Sci. Technol.* **2012**, *46*, 876–884.
- (63) Muñoz, F.; Mvula, E.; Braslavsky, S. E.; von Sonntag, C. Singlet dioxygen formation in ozone reactions in aqueous solution. *J. Chem. Soc. Perkin Trans. 2* **2001**, 1109–1116.
- (64) Lange, F.; Cornelissen, S.; Kubac, D.; Sein, M. M.; von Sonntag, J.; Hannich, C. B.; Golloch, A.; Heipieper, H. J.; Möder, M.; von Sonntag, C. Degradation of macrolide antibiotics by ozone: a mechanistic case study with clarithromycin. *Chemosphere* **2006**, *65*, 17–23.
- (65) Qin, L.; Tripathi, G. N. R.; Schüler, R. H. Radiation chemical studies of the oxidation of aniline in aqueous solution. *Zeitschrift für Naturforsch. A* **1985**, *40*, 1026–1039.

- (66) Mvula, E.; Naumov, S.; von Sonntag, C. Ozonolysis of Lignin Models in Aqueous Solution: Anisole, 1,2-Dimethoxybenzene, 1,4-Dimethoxybenzene, and 1,3,5-Trimethoxybenzene. *Environ. Sci. Technol.* **2009**, *43*, 6275–6282.
- (67) Criegee, R. Mechanism of Ozonolysis. *Angew. Chemie Int. Ed. English* **1975**, *14*, 745–752.
- (68) Dowideit, P.; von Sonntag, C. Reaction of Ozone with Ethene and Its Methyl- and Chlorine-Substituted Derivatives in Aqueous Solution. *Environ. Sci. Technol.* **1998**, *32*, 1112–1119.
- (69) Leitzke, A.; von Sonntag, C. Ozonolysis of Unsaturated Acids in Aqueous Solution: Acrylic, Methacrylic, Maleic, Fumaric and Muconic Acids. *Ozone Sci. Eng.* **2009**, *31*, 301–308.
- (70) Jans, U. Doctoral thesis. *ETH Zurich* **1996**.
- (71) Leitzke, A.; Reisz, E.; Flyunt, R.; von Sonntag, C. The reactions of ozone with cinnamic acids: formation and decay of 2-hydroperoxy-2-hydroxyacetic acid. *J. Chem. Soc. Perkin Trans. 2* **2001**, 793–797.
- (72) McDowell, D. C.; Huber, M. M.; Wagner, M.; von Gunten, U.; Ternes, T. A. Ozonation of Carbamazepine in Drinking Water: Identification and Kinetic Study of Major Oxidation Products. *Environ. Sci. Technol.* **2005**, *39*, 8014–8022.
- (73) Muñoz, F.; von Sonntag, C. Determination of fast ozone reactions in aqueous solution by competition kinetics. *J. Chem. Soc. Perkin Trans. 2* **2000**, 661–664.
- (74) Lester, Y.; Mamane, H.; Zucker, I.; Avisar, D. Treating wastewater from a pharmaceutical formulation facility by biological process and ozone. *Water Res.* **2013**, *47*, 4349–4356.
- (75) Tay, K. S.; Madehi, N. Ozonation of ofloxacin in water: By-products, degradation pathway and ecotoxicity assessment. *Sci. Total Environ.* **2015**, *520*, 23–31.
- (76) Witte, B. De; Langenhove, H. Van; Hemelsoet, K.; Demeestere, K.; Wispelaere, P. De; Van Speybroeck, V.; Dewulf, J. Levofloxacin ozonation in water: rate determining process parameters and reaction pathway elucidation. *Chemosphere* **2009**, *76*, 683–689.
- (77) Tekle-Röttering, A.; Jewell, K. S.; Reisz, E.; Lutze, H. V.; Ternes, T. A.; Schmidt, W.; Schmidt, T. C. Ozonation of piperidine, piperazine and morpholine: Kinetics, stoichiometry, product formation and mechanistic considerations. *Water Res.* **2016**, *88*, 960–971.
- (78) von Gunten, U. Ozonation of drinking water: part I. Oxidation kinetics and product formation. *Water Res.* **2003**, *37*, 1443–1467.
- (79) Jonsson, M.; Wayner, D. D. M.; Luszyk, J. Redox and Acidity Properties of Alkyl- and Arylamine Radical Cations and the Corresponding Aminyl Radicals. *J. Phys. Chem.* **1996**, *100*, 17539–17543.
- (80) Neta, P.; Huie, R. E.; Ross, A. B. Rate Constants for Reactions of Peroxyl Radicals in Fluid Solutions. *J. Phys. Chem. Ref. Data* **1990**, *19*, 413.
- (81) Das, S.; von Sonntag, C. The oxidation of trimethylamine by OH radicals in aqueous solution, as studied by pulse radiolysis, ESR, and product analysis. **1986**.
- (82) Das, S.; Schuchmann, M. N.; Schuchmann, H.-P.; Sonntag, C. Von. The production of the superoxide radical anion by the OH radical-induced oxidation of trimethylamine in oxygenated aqueous solution. The kinetics of the hydrolysis of (hydroxymethyl)dimethylamine. *Chem. Ber.* **1987**, *120*, 319–323.
- (83) von Sonntag, C.; Schuchmann, H.-P. The Elucidation of Peroxyl Radical Reactions in Aqueous Solution with the Help of Radiation-Chemical Methods. *Angew. Chemie Int. Ed. English* **1991**, *30*, 1229–1253.
- (84) Russell, G. A. Deuterium-isotope Effects in the Autoxidation of Aalkyl Hydrocarbons. Mechanism of the Interaction of Peroxy Radicals 1. *J. Am. Chem. Soc.* **1957**, *79*, 3871–3877.
- (85) Bennett, J. E.; Summers, R. Product Studies of the Mutual Termination Reactions of sec - Alkylperoxy Radicals: Evidence for Non-Cyclic Termination. *Can. J. Chem.* **1974**, *52*, 1377–1379.
- (86) von Sonntag, C. *Free-radical-induced DNA damage and its repair*; Springer, 2006.
- (87) Enami, S.; Hoffmann, M. R.; Colussi, A. J. Ozonolysis of Uric Acid at the Air/Water Interface. *J. Phys. Chem. B* **2008**, *112*, 4153–4156.
- (88) Flyunt, R.; Theruvathu, J. A.; Leitzke, A.; von Sonntag, C. The reactions of thymine and thymidine with ozone. *J. Chem. Soc. Perkin Trans. 2* **2002**, 1572–1582.
- (89) Andreozzi, R.; Insola, A.; Caprio, V.; D'Amore, M. G. M. Ozonation of pyridine in aqueous

Supporting information for chapter 4

- solution: Mechanistic and kinetic aspects. *Water Res.* **1991**, *25*, 655–659.
- (90) Acero, J. L.; Stemmler, K.; von Gunten, U. Degradation Kinetics of Atrazine and Its Degradation Products with Ozone and OH Radicals: A Predictive Tool for Drinking Water Treatment. *Environ. Sci. Technol.* **2000**, *34*, 591–597.
- (91) Radjenović, J.; Godehardt, M.; Hein, A.; Farré, M.; Jekel, M.; Barceló, D. Evidencing Generation of Persistent Ozonation Products of Antibiotics Roxithromycin and Trimethoprim. *Environ. Sci. Technol.* **2009**, *43*, 6808–6815.
- (92) Kuang, J.; Huang, J.; Wang, B.; Cao, Q.; Deng, S.; Yu, G. Ozonation of trimethoprim in aqueous solution: identification of reaction products and their toxicity. *Water Res.* **2013**, *47*, 2863–2872.
- (93) Dodd, M. C.; Rentsch, D.; Singer, H. P.; Kohler, H.-P. E.; von Gunten, U. Transformation of β -Lactam Antibacterial Agents during Aqueous Ozonation: Reaction Pathways and Quantitative Bioassay of Biologically-Active Oxidation Products. *Environ. Sci. Technol.* **2010**, *44*, 5940–5948.
- (94) Flyunt, R.; Makogon, O.; Schuchmann, M. N.; Asmus, K.-D.; von Sonntag, C. OH-Radical-induced oxidation of methanesulfinic acid. The reactions of the methanesulfonyl radical in the absence and presence of dioxygen. *J. Chem. Soc. Perkin Trans. 2* **2001**, 787–792.
- (95) Enami, S.; Hoffmann, M. R.; Colussi, A. J. Ozone Oxidizes Glutathione to a Sulfonic Acid. *Chem. Res. Toxicol.* **2009**, *22*, 35–40.
- (96) Enami, S.; Hoffmann, M. R.; Colussi, A. J. Simultaneous Detection of Cysteine Sulfenate, Sulfinate, and Sulfonate during Cysteine Interfacial Ozonolysis. *J. Phys. Chem. B* **2009**, *113*, 9356–9358.
- (97) Kalia, J.; Raines, R. T. Hydrolytic Stability of Hydrazones and Oximes. *Angew. Chemie Int. Ed.* **2008**, *47*, 7523–7526.
- (98) TAM, K. Y.; QUÉRÉ, L. Multiwavelength Spectrophotometric Resolution of the Micro-Equilibria of Cetirizine. *Anal. Sci.* **2001**, *17*, 1203–1208.
- (99) Pagliara, A.; Testa, B.; Carrupt, P.-A.; Jolliet, P.; Morin, C.; Morin, D.; Urien, S.; Tillement, J.-P.; Rihoux, J.-P. Molecular Properties and Pharmacokinetic Behavior of Cetirizine, a Zwitterionic H₁-Receptor Antagonist. *J. Med. Chem.* **1998**, *41*, 853–863.
- (100) Tomasi, J.; Mennucci, B.; Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **2005**, *105*, 2999–3093.
- (101) von Gunten, U.; Salhi, E.; Schmidt, C. K.; Arnold, W. A. Kinetics and Mechanisms of N - Nitrosodimethylamine Formation upon Ozonation of N,N-Dimethylsulfamide-Containing Waters: Bromide Catalysis. *Environ. Sci. Technol.* **2010**, *44*, 5762–5768.
- (102) Trogolo, D.; Mishra, B. K.; Heeb, M. B.; von Gunten, U.; Arey, J. S. Molecular Mechanism of NDMA Formation from N , N -Dimethylsulfamide During Ozonation: Quantum Chemical Insights into a Bromide-Catalyzed Pathway. *Environ. Sci. Technol.* **2015**, *49*, 4163–4175.

Chapter 5. General conclusions and perspectives

Chapter 5

In recent years, ozone applications were extended from drinking water disinfection to tertiary wastewater treatment to mitigate the micropollutant load from the discharge of secondary wastewater effluents to the aquatic environment. Empirical findings from diverse wastewater effluents worldwide showed efficient alleviation of the toxicological potential by ozonation, especially in combination with biological post-filtration. However, the application of ozonation to wastewater effluents leads to the formation of numerous transformation products from the oxidation of micropollutants by ozone and hydroxyl radicals and limited information is available on these compounds. To overcome this deficit, the goal of this PhD thesis was the development of a prediction platform for ozone transformation products based on currently available empirical information. In this first step of development, the reactions of hydroxyl radicals were excluded. The prediction platform for transformation products was developed as a graphical user interface (GUI) application. The developed prediction platform consists of two main prediction modules: (i) prediction of second-order rate constants (k_{O_3}) for the reaction of ozone with micropollutants and (ii) reaction pathways and transformation product formation.

(i) The prediction module for k_{O_3} is described in chapter 2 and it was demonstrated that k_{O_3} for a micropollutant can be predicted using a good correlation ($R^2=0.82-1.00$) between the experimental k_{O_3} (in log units) for compounds of a specific chemical group (e.g., phenols) and the corresponding orbital energies (e.g., highest occupied molecular orbital (HOMO) energy or natural bond orbital (NBO) energy) obtained from quantum chemical computations. The developed prediction model has useful applications. Firstly, predicted k_{O_3} -values for micropollutants can be used for estimating their elimination efficiencies during ozonation under different operating conditions (i.e., different ozone exposures), which enables an optimization of an ozonation process. Secondly, the dominant sites of reaction with ozone can be elucidated for a compound with multiple ozone-reactive moieties. This serves as critical information for the prediction module for reaction pathways and the formation of transformation products.

In the developed GUI application in chapter 4, the computational protocol for the k_{O_3} prediction was adapted from the protocol in chapter 2 to significantly reduce computational costs, while achieving the acceptable prediction performance. Consequently, k_{O_3} prediction can be implemented on a personal computer with a reasonable CPU time. The adapted k_{O_3} prediction models (mostly $R^2 = 0.75 - 0.95$) were developed for 14 compound classes consisting of 284 model compounds. The prediction performance of the adapted k_{O_3} prediction models was reasonable with a predicted k_{O_3} on average within a factor of ~ 5 of the experimental k_{O_3} for all the selected model compounds and the externally validated tetrachlorobutadienes. However, poor prediction was also observed for some model compounds, which were excluded from model developments, as well as for some micropollutants such as cetirizine, pentachlorobutadienes, and hexachlorobutadienes, which were used as the external validation set.

The use of a gas-phase geometry for quantum chemical computations was shown to contribute to an error for cetirizine in chapter 4. The employment of aqueous phase-relevant geometries may lead to a better prediction. Moreover, as already shown for aromatic compounds in which different linear relationships for k_{O_3} were established for different aromatic compound classes (e.g., phenols, anilines, alkoxybenzenes), the establishment of separate correlations for subclasses of a certain compound

group can be considered. For example, a separate correlation may be pursued for cyclic amines (e.g., piperazine, piperidine) from the amine group. However, the correlation must be supported by the appropriate number of model compounds with empirical k_{O_3} . Therefore, empirical determinations of k_{O_3} for additional model compounds may be necessary. Overall, the developed k_{O_3} prediction models will be further tested and improved in future with more compounds with various chemical moieties for a better understanding of their limitations and the options for improvements.

A k_{O_3} prediction model using other quantum molecular descriptors than orbital energies can also be considered. The Gibbs free energy for the formation of an ozone adduct or for the formation of an activated complex between ozone and a compound could be an option for aromatic compounds. However, this approach is more involved than the orbital energy approach, because it necessitates the computation for both a reactant and a product (i.e., an ozone adduct or an activated complex). It may also demand computations with a high level of theory, which would significantly increase the computational costs.

pK_a and tautomeric fractions (f) can also play a critical role for the accuracy of the pH-dependence of k_{O_3} predictions. By default, pK_a - and f -values are predicted by a built-in chemoinformatic tool. As predicted values are not necessarily reliable, it is critical to provide an empirical or theoretical input, which is as reliable as possible. However, such parameters may not be available for certain micropollutants, especially for compounds with complex acid-base speciations. This caveat should be kept in mind when performing a pH-dependent prediction for k_{O_3} .

(ii) The predictions for reaction pathways and transformation products are carried out according to pre-defined mechanisms of ozone reactions for individual moieties from the peer-reviewed literature. Although the demonstrations with a few micropollutants were shown to be successful, verifications with more micropollutants with empirical data and the subsequent updates and improvements are necessary in future.

The following improvements and perspectives for the developed prediction platform should be considered in future:

- As mentioned in chapter 4, there are still various chemical moieties for which k_{O_3} and/or pathway predictions are not implemented mainly due to the lack of empirical information (e.g., thiophenol and hydrazine). Therefore, further empirical studies need to focus on such chemical moieties.
- The scope of the current prediction platform is exclusive for ozone. However, for ozone-resistant micropollutants, hydroxyl radical, a secondary oxidant formed during ozonation, plays a dominant role for its abatement. Reactions with hydroxyl radical give rise to different transformation products than the reactions with ozone. Therefore, the extension of the prediction tool from ozone only to a combination of ozone and hydroxyl radical should be a priority for future developments. Moreover, the formation of well-known ozonation by-products (e.g., nitrosamines) and the role of water matrix components (e.g., Br^-) should also be included in future.
- The prediction platform is currently a standalone GUI application, which can be run on a personal computer. The development of an online prediction platform based on the developed GUI applica-

Chapter 5

tion, to which anyone has free access without any software/application installations, will be pursued in future. Financial and technical aspects such as software licenses and web-server managements need to be taken into account.

- The developed prediction platform is versatile and can also be implemented for other oxidants such as chlorine, chloramine, chlorine dioxide, ferrate, etc. However, this can be achieved only if there is sufficient empirical information (kinetics and mechanisms) to support the development of the respective transformation rules.
- The computer-based prediction platform for kinetics and mechanisms for the reactions of ozone with micropollutants developed in this PhD thesis can be beneficial to various groups of end users. Environmental engineers and treatment plant operators can benefit from predicted k_{O_3} for designing/optimizing ozonation processes. Predicted reaction pathways or chemical information about transformation products such as molecular mass or chemical structure can be useful for environmental chemists/toxicologists or analytical chemists for elucidation of ozone reactions, identification of transformation products by mass spectrometry, or (*in silico*) toxicity screening of transformation products.

In chapter 3, empirical investigations on abatement of nine polychlorobutadienes (CBDs) by ozonation, direct UV photolysis at 254 nm, and their H_2O_2 -based advanced oxidation processes (i.e., O_3/H_2O_2 and UV/H_2O_2) were conducted. The order of the determined k_{O_3} for CBDs and second-order rate constants for the reaction of CBDs with hydroxyl radicals (k_{OH}) was consistent with the extent of abatement efficiency for the selected CBDs and micropollutants. While such k -values are usually determined from experiments, it was shown in this study that the developed k_{O_3} prediction model in chapter 2 is an alternative to the empirical determination. Using the k_{O_3} prediction model developed for olefins, k_{O_3} for tetraCBDs could be predicted within a factor of ~ 4 of the experimental k_{O_3} . In experiments in a groundwater, the formation of bromate, which is a potentially carcinogenic by-product formed during ozonation, can be reduced by the H_2O_2 addition. As the H_2O_2 addition can lead to a decreased abatement efficiency of micropollutants, the O_3/H_2O_2 system needs to be optimized in terms of sufficient CBDs abatement while minimizing bromate formation. CBDs were shown to be labile to UV photolysis in agreement with the general properties of dienes. However, CBDs with *Z* or *E* configuration underwent photoisomerization to form the counter-configuration isomers, which may not lead to a loss of the biological effects. Therefore, photoisomerization of CBDs should also be taken into account for UV treatment of CBDs. H_2O_2 addition to UV photolysis (i.e., UV/H_2O_2 system) did not improve abatement efficiencies for CBDs, supposedly due to the efficient transformation of CBDs by direct UV photolysis.

Curriculum vitae

Minju Lee

minju.lee@epfl.ch

GR A1 475 Station 2
CH-1015 Lausanne
Switzerland
+41 78 611 2794 (mobile)

EDUCATION

École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Ph.D candidate in Environmental Engineering, School of Architecture, Civil and Environmental Engineering (ENAC), December 2011-present

Ph.D. thesis: Prediction of Transformation Products during Ozonation of Micropollutant-containing waters: Kinetics and Mechanisms

Ph.D. advisors: Prof. Urs von Gunten (EPFL/EAWAG)
PD. Dr. Kathrin Fenner (EAWAG/ETH Zurich)

Swiss Federal Institute of Technology Zurich (ETH Zurich), Zurich, Switzerland

Visiting student in Dept. of Environmental Science (D-UWIS) (Sep. 2010 – Dec. 2011)

as a Swiss Government Scholarship holder for 2010-2011

Chung-Ang University (CAU), Seoul, South Korea

M.S in Environmental Engineering (2008-2010)

B.S in Civil and Environmental Engineering (2000-2001 and 2005-2008)

*Korean military service (2002-2004)

PROFESSIONAL EXPERIENCES

Swiss Federal Institute of Aquatic Science and Technology (EAWAG), Dübendorf, Switzerland

Visiting researcher in Dept. of Water resources and Drinking water (Sep. 2010 – Dec. 2011)

as a Swiss Government Scholarship holder for 2010-2011

Korea Institute of Science and Technology (KIST), Seoul, South Korea

Research Scientist in Center for Environmental Technology Research (Jan. 2010 – Jun. 2010)

Chung-Ang University (CAU), Seoul, South Korea

Graduate Research Assistant in Dept. of Civil and Environmental Engineering (2008-2010)

LABORATORY SKILLS

UV-Vis spectrophotometry
Gas Chromatography (GC)
Ion Chromatography (IC)
High-Performance Liquid
Chromatography (HPLC)
Mass spectrometry (MS)

LANGUAGE SKILLS

Korean (Mother tongue)
English (Fluent)
French (Elementary)

SOFTWARE SKILLS

MS office suites (PowerPoint,
Excel, & MS Word)
Java (Programming language)
Gaussian and ORCA (Quantum
Chemical Computation software)
Kintecus (Chemical Kinetics
Simulation software)

PROFESSIONAL CERTIFICATES

Engineer in Civil Engineering certified in 2008 by Human Resources Development Service of Korea

PUBLICATIONS

M. Lee, T. Merle, D. Rentsch, S. Canonica, U. von Gunten, “Abatement of Polychloro-1,3-butadienes in Aqueous Solution by Ozone, UV-photolysis, and Advanced Oxidation Processes (O_3/H_2O_2 and UV/H_2O_2)”, *in preparation*.

M. Lee, L. C. Blum, E. Schmid, K. Fenner, U. von Gunten, “Development of a Computer-based Prediction Platform for the Reaction of Ozone with Organic Compounds in Aqueous Solution: Kinetics and Mechanisms”, *in preparation*.

D. Gerrity, Y. Lee, S. Gamage, **M. Lee**, A. Pisarenko, R. A. Trenholm, U. von Gunten, and S. A. Snyder, “Emerging investigators series: prediction of trace organic contaminant abatement with UV/H_2O_2 : development and validation of semi-empirical models for municipal wastewater effluents”, *Environmental Science: Water Research & Technology*, 2, 460-473 (2016).

Y. Lee, D. Gerrity, **M. Lee**, S. Gamage, A. Pisarenko, R. A. Trenholm, S. Canonica, S. A. Snyder, and U. von Gunten, “Organic Contaminant Abatement in Reclaimed Water by UV/H_2O_2 and a Combined Process consisting of O_3/H_2O_2 followed by UV/H_2O_2 : Prediction of abatement efficiency, energy consumption, and by-product formation”, *Environmental Science and Technology*, 50(7), 3809-3819 (2016).

M. Lee, S. G. Zimmermann-Steffens, J. Samuel Arey, K. Fenner, U. von Gunten, “Development of Prediction Models for the Reactivity of Organic Compounds with Ozone in Aqueous Solution by Quantum Chemical Calculations: The Role of Delocalized and Localized Molecular Orbitals”, *Environmental Science and Technology*, 49(16), 9925-9935(2015).

Y. Lee, D. Gerrity, **M. Lee**, A. E. Bogeat, E. Salhi, S. Gamage, R. A. Trenholm, E. C. Wert, S. A. Snyder, and U. von Gunten, “Prediction of Micropollutant Elimination during Ozonation of Municipal Wastewater Effluents: Use of Kinetic and Water Specific Information”, *Environmental Science and Technology*, 47(11), 5872-5881 (2013).

M. Lee, Y. Lee, F. Soltermann, and U. von Gunten, "Analysis of *N*-nitrosamines and other nitro(so) compounds in water by high-performance liquid chromatography with post-column UV photolysis/Griess reaction", *Water Research*, 47(14), 4893-4903 (2013).

F. Soltermann, M. Lee, S. Canonica, and U. von Gunten, "Enhanced *N*-nitrosamine formation in pool water by UV irradiation of chlorinated secondary amines in the presence of monochloramine", *Water Research*, 47(1), 79-90 (2013).

M. Lee and J. Oh, "The synergistic effect of hydrogen peroxide production and sonochemiluminescence under dual frequency ultrasound irradiation", *Ultrasonics Sonochemistry*, 18(3), 781-788 (2011).

M. Lee and J. Oh, "Sonolysis of aqueous mixtures of trichloroethylene and carbon tetrachloride", *Ultrasonics Sonochemistry*, 17(1), 207-212 (2010).

CONFERENCE PRESENTATIONS

"Prediction of transformation products during ozonation processes", Peak Course, Eawag, (2016), *oral presentation*

"Assessment of the reactivity of organic compounds with ozone in aqueous solution by quantum chemical calculations: Role of delocalized and localized molecular orbitals", 5th EuCheMS Chemistry Congress, Istanbul, Turkey (2014), *oral presentation*

"Computer-based predictions of transformation products", Internal workshop for evaluation of biological treatment processes after ozonation of wastewater from nutrient removal plants, EAWAG, Dübendorf, Switzerland (2014), *oral presentation*

"Prediction of rate constants for the reaction of ozone with micropollutants (k_{O_3}): Quantum chemical approach", 6th Late Summer Workshop: Micropollutants in the water cycle, Schloss Maurach, Lake Constance, Germany (2013), *oral presentation*

"Micropollutants elimination during ozonation: prediction of kinetics and reaction pathways", Green days, EPFL internal conference, Switzerland (2013), *oral presentation*

"Analysis of *N*-nitrosamines and other Nitro(so) compounds in Water by High-Performance Liquid Chromatography with Post-column UV Photolysis / Griess Reaction", 8th Micropol & Ecoharzard, Zurich, Switzerland (2013), *oral presentation*

TEACHING EXPERIENCES

Teaching assistant, "Environmental Chemistry" for Bachelor, EPFL, 2012-2016

Teaching assistant, "Water and Wastewater Treatment" for Master, EPFL, 2012-2015