MODELING THE DEGRADATION OF ORGANIC POLLUTANTS AND THE SOLUBILITY OF DRUG-LIKE MOLECULES

by

Evrim Arslan

Integrated B.S. and M.S., Teaching Chemistry, Boğaziçi University, 2013

M.S., Chemistry, Boğaziçi University, 2016

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Chemistry

Boğaziçi University

2023

Dedicated to my family

ACKNOWLEDGEMENTS

I would like to state my deeply gratefulness to my supervisors Prof. Viktorya Aviyente to give me the opportunity to work in her research group. I would like to express my thanks to her for her support, scientific guidance through out these projects and endless attention. I gained lots of experience during these years.

Also, I would like to extend my thanks to my second supervisor Prof. İlknur Doğan for her kindness, scientific support and advices regarding my Projects.

I wish to express my appreciation to my committee members Prof. Nilsun İnce, Prof. Ayşe Neren Ökte and Prof. Fethiye Aylin Sungur and Prof. Nurcan Tüzün.

Moreover, I would like to express my great thanks Başak Fındık, Deniz Akgül, İpek Munar, Sesil Agopçan Çınar and Pınar Haşlak for their supports, friendships and helpfulness throughout these years. I would also thank all my friends and the members of the faculty in Chemistry Department.

Finally, my deepest thanks go to my whole family for their endless love, support and encouragement.

ABSTRACT

MODELING THE DEGRADATION OF ORGANIC POLLUTANTS AND THE SOLUBILITY OF DRUG-LIKE MOLECULES

In this doctoral thesis, three different subjects were investıgated with hybrid computational methods. The first part dealing with environmental chemistry covers the removal of pharmaceutical raw materials and personal care products (PPCPs) that pollute freshwater. In this section, the formation of possible intermediates detected during the removal processes of salicylic acid, methylparaben and caffeine by advanced oxidation techniques is modeled by the density functional theory. It is important to understand the removal mechanisms of these chemicals, as they have the potential to destroy the entire aquatic ecosystems. The second part of the thesis focuses on the determination of the potential of drug-like molecules to penetrate the cells by determining the hydrophilic/hydrophobic properties of the organic molecules that make up the drug raw material. In this section, inferences were made about the calculation methods of the dissociation constants (pK_a) of 11 protein kinase inhibitor molecules with the potential to become drugs. In addition, in the continuation of the study, the water/octanol dissociation constants (LogP) of 22 molecules were calculated and their dissociation abilities in water and organic solvent were determined. In the third part of the thesis, the method that can best reflect the dissociation constants of tiyazol-2-imin derivatives dissolved in organic solvent was determined. These molecules have been synthesized in the Chemistry Department of Bogazici University.

ÖZET

ORGANİK KİRLETİCİLERİN BOZUNMASININ VE İLAÇ BENZERİ MOLEKÜLLERİN ÇÖZÜNÜRLÜĞÜNÜN MODELLENMESİ

Bu doktora tez çalışmasında üç farklı konu hibrid hesapsal yöntemlerle çalışılmıştır. Çevre kimyası ile ilgili olan ilk kısım tatlı suları kirleten ilaç ham maddeleri ve kişisel bakım ürünlerinin (PPCP'ler) giderilmesini kapsamaktadır. Bu bölümde salisilik asit, metil paraben ve kafeinin ileri oksidasyon teknikleri ile giderilme süreçleri boyunca tespit edilen olası ara ürünlerin oluşması yoğunluk teorisi yöntemi ile modellenmiştir. Bahsi geçen kimyasalların sudaki tüm ekosistemleri oratadan kaldırma potansiyelleri olduğu için giderilme mekanizmalarının anlaşılması önem taşımaktadır. Ara ürünlerin oluşma olasılıkları tırmandıkları Gibbs eşik enerjileri saptanarak, bu yapıların Gibbs serbest aktivasyon enerjileri hesaplanarak değerlendirilmiştir. Tezin iki sayısal olarak saptanarak hücrelere girme potansiyellerinin tespit edilmesine yöneliktir. Bu kısımda 11 adet ilaç olma potansiyeli olan protein kinaz inhibitörü molekülün ayrışma sabitlerinin (pK_a) hesaplama yöntemleri hakkında çıkarımlarda bulunulmuştur. Bunun yanısıra çalışmanın devamında 22 adet molekülün su/oktanol ayrışma sabitleri (LogP) hesaplanarak suda ve organik çözücüde ayrışma yetileri tespit edilmiştir. Tezin üçüncü bölümünde, FEF-Kimya Bölümünde sentezlenmiş olan farmasötik özelliklere sahip olacağı düşünülen organik çözücüde eriyen tiyazol-2-imin türevlerinin ayrışma sabitlerini en iyi yansıtabilecek yöntem tespit edilmiştir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

ϵ	Dielectric Constant
$\Delta G^{\ddagger}$	Gibbs free energy of activation
ΔG_{rxn}	Gibbs free energy of reaction
ΔG_{solv}	Solvation Energy of a Reaction

LIST OF ACRONYMS/ABBREVIATIONS

AOP	Advanced Oxidation Process
CAF	Caffeine
CPCM	Conductor-Like Polarizable Continuum Model
DFT	Density Functional Theory
ECOSAR	Ecological Structure Activity Relationship
IEF-PCM	Integral Equation Formalism Polarizable Continuum Model
kHz	Kilohertz
LCMS	Liquid Chromatography–Mass Spectrometry
LogP	Water/octanol Partition Coefficient
MP	Methylparaben
NPA	Natural Population Analysis
QSAR	Quantitative Structure-Activity Relationship
PCP	Personal Care Product
pK_a	Acid-Dissociation Constant
PPCP	Pharmaceutical and Personal Care Product
SA	Salicylic Acid
SMD	Solvation Model Based on Density
TC	Thermodynamic Cycle
TOC	Total Organic Carbon

1. INTRODUCTION

1.1. Pharmaceuticals and Personal Care Products

Personal care products (PCPs) are used primarily to enhance daily living, whereas pharmaceuticals are described as prescription, over-the-counter, and veterinary therapeutic medications meant to prevent or treat human and animal disorders. The unintended presence of pharmaceuticals and personal care products (PPCPs) in various aquatic environment (Figure 1.1) compartments (such as water, sediments, and biota) at amounts capable of harming aquatic creatures has come to more people's attention in recent years [1]. Due to PPCPs' extensive and growing usage in human and veterinary medicine and the ongoing leakage of these substances into the environment, this has grown to be a serious issue [2].

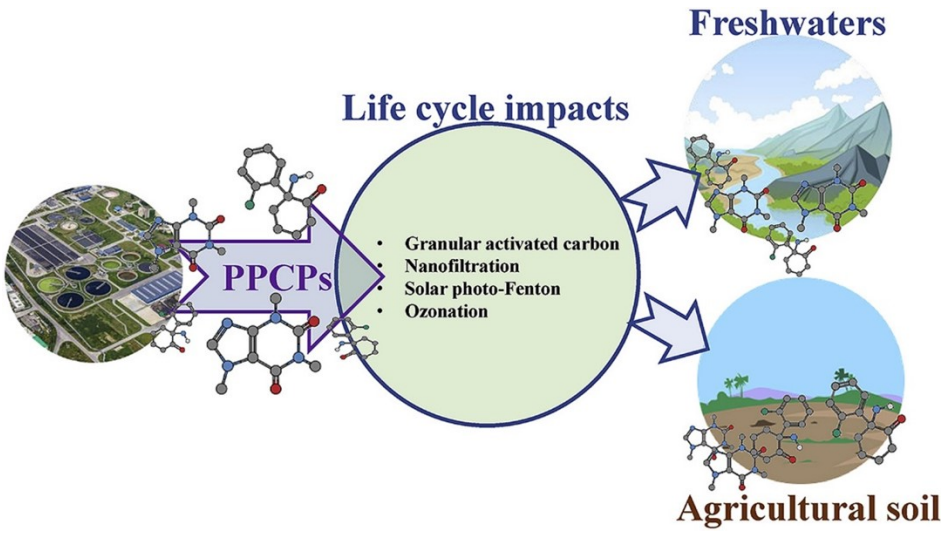


Figure 1.1. Representation of PPCPs presence in the environment, fate, transport, and accumulation in plants and animals [3].

Pharmaceuticals are substances used to treat medical conditions. Some examples of pharmaceuticals include antihyperlipidemic drugs, stimulants (caffeine), analgesics (ibuprofen, diclofenac, paracetamol, ketoprofen, naproxen, salicylic acid), psy-

chiatric drugs (carbamazepine, primidone), antibiotics (sulfamethoxazole, chloramphenicol, trimothoprin, ciprofloxacin), lipid regulators (gemfibrozil, bezafibrate, propranolol, atenolol, metoprolol), and antipyretics. PCPs include synthetic fragrances (nitropolycyclic musks), antimicrobial compounds (triclosan), UV blockers (methylbenzylidene camphor), antioxidants and preservatives (phenols and p-hydroxybenzoic acid [parabens]), and insect repellents (N,N-diethyl-m-toluamide [DEET]). Hormones (estrone E1, estradiol E2, ethynlestradiol EE2) are also frequently detected in the environment [4-8].

1.2. Advanced Oxidation Processes

Advanced oxidation processes (AOPs) are mainly used for the degradation of contaminants from the water systems. The processes are mainly based on the creation of highly reactive radicalic species that can destroy the contaminants in water systems. There are many different ways to generate radicalic species that are generally OH radicals since they are very reactive and easy to generate. Hydroxyl radicals are produced with the help of one or more primary oxidants (e.g. ozone, hydrogen peroxide, oxygen) and/or energy sources (e.g. ultraviolet light) or catalysts (e.g. titanium dioxide). Precise, pre-programmed dosages, sequences and combinations of these reagents are applied in order to obtain a maximum OH yield. In general, when applied in properly tuned conditions, AOPs can reduce the concentration of contaminants from several-hundreds ppm to less than 5 ppb and therefore significantly bring the quantity of total organic carbon (TOC) down [3]. AOPs are mainly based on three steps: formation of highly reactive radicalic species (like OH radicals), attacking OH radicals to the target molecules to breakdown to their fragments and subsequent attacking until the ultimate mineralizations. The mechanism of OH radical generation depends on the way of AOP technique that is used.

Ultraviolet irradiation (UV) and hydrogen peroxide (H_2O_2) can be used for generation of two OH radicals via homolytic bond cleavage O-O [10, 11]. Hypochlorous acid (HOCl) and UV can also be used generating OH and Cl radicals [12]. Moreover,

there is ozone (O_3) based AOPs (Eq. 1.1)

$$\begin{aligned} &O_3 + HO^- \rightarrow HO_2^- + O_2 \\ &O_3 + HO_2^- \rightarrow HO_2. + O_3 \\ &O_3^-. + H^+ \rightarrow O_3. \\ &O_3. \rightarrow .OH + O_2. \end{aligned} \tag{1.1}$$

Also, there are AOPs using Fenton reagent [13] and photocatalytic oxidation with titaniumdioxide (TiO_2) [14] (Eq. 1.2)

$$\begin{aligned} &TiO_2 + UV \rightarrow e^- + h^+ \\ &Ti(IV) + H_2O \rightleftharpoons Ti(IV) - H_2O \\ &Ti(IV) + H_2O + h^+ \rightleftarrows Ti(IV) - .HO + H^+. \end{aligned} \tag{1.2}$$

1.3. Octanol-Water Partition Coefficient (LogP)

The most popular method of quantifying a compound's lipophilicity is through its octanol-water partition ratio, which is the ratio of a solute's concentration in an octanolic phase that is water-saturated to that of the same solute in an aqueous phase that is octanol-saturated (Figure 1.2). Lipophilicity is frequently considered as one of the most crucial physicochemical characteristics to screen lead compounds in the early stages of the drug development process. It is also widely employed in the sectors of environmental risk assessment and agriculture. The octanol-water partition ratio, among other things, is frequently used as a descriptor in quantitative structure-activity/property interactions as a result of the association between lipophilicity and biological activity. In fact, it can be found in thousands of equations that deal with how biological systems behave [15].

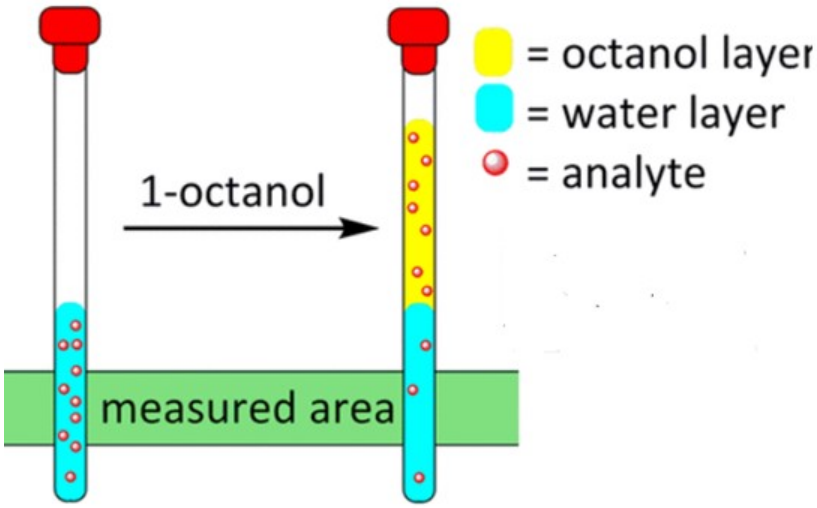


Figure 1.2. Schematic description of LogP [16].

The bioavailability of a medicine and its entry into the target are crucial factors in a rational drug design. To have an impact, a medicine needs to cross a cell membrane barrier. The pH of the environment can vary depending on where the target site is and how the drug is administered; for example, the pH of the kidneys is 4.2, that of plasma is 7.4, that of the stomach is 2.0, etc. To forecast a drug's effects on a target molecule, one must take into account the drug's affinity and capacity to penetrate a lipophilic environment at various pH levels [17]. A target-specific pharmacological treatment requires a solid understanding of drug dispersion. While hydrophobic medications with high LogP values are distributed to the hydrophobic portions of the body, such as lipid bilayers of cells, hydrophilic drugs with low LogP values are typically concentrated in the hydrophilic sections of the organism, such as blood serum. LogP coefficient is utilized to properly estimate a drug's pharmokinetic qualities for this same reason as mentioned in details in Section 5.2.

1.4. Dissociation Constants of Drug-Like Molecules

The pharmaceutical and chemical industries have long been aware of the impact of the acid-base dissociation constant, pK_a, on the biopharmaceutical properties of

medicines and chemicals. Knowing the dissociation constant in each case assists in understanding the ionic form a molecule will adopt over a variety of pH values because the majority of medications are weak acids and/or bases. This is crucial in physiological systems as the ionization state will impact how quickly a substance diffuses through barriers and membranes like the blood-brain barrier (BBB). Lipophilicity, solubility, protein binding, and permeability are all directly impacted by a drug's pK_a, and these factors in turn have an impact on pharmacokinetic (PK) properties like absorption, distribution, metabolism, and excretion (ADME)[18–21]. Since there is a known relationship between pK_a and PK, measuring pK_a levels is necessary for regulatory compliance. The determination of the pK_a benefits formulation processes for maximizing drug distribution. Given the significance of this parameter to the pharmaceutical industry [22], it stands to reason that being able to estimate or measure [23] the pK_a and understanding their distribution will be extremely beneficial. Given the vast number of compounds that might be taken into account for screening purposes, this is especially crucial (e.g. combinatorial libraries, third party compound collections). In terms of the percentage of ionizables and the distribution of different chemical classes, these groups of compounds should ideally be reflective of drug-like substances as a whole.

2. OBJECTIVE AND SCOPE

In this doctoral thesis research, three different topics are evaluated using hybrid computational methods.

The first part is related to environmental chemistry based on the pharmaceutical materials and personal care products (PPCPs) that pollute freshwater systems. In this section, the formation of possible intermediate products detected during the decontamination process of salicylic acid, methylparaben, and caffeine by advanced oxidation techniques are modeled by using hybrid computational methods. Since these chemicals can remove entire ecosystems from water, it is important to understand the removal mechanisms. The probability of formation of intermediates are assessed by determining the activation barriers at which they slowed down and also by calculating the Gibbs free energy.

The second part of the thesis aims to numerically determine the hydrophilicity/hydrophobicity of the organic molecules that make up the drug substance and their potential to enter cells. This section presents 11 protein kinase inhibitors with potential drug potential. For this purpose, the water/octanol partition coefficient (LogP) of 22 molecules are calculated to determine the degradation capacities in water and organic solvents.

In the third part of the thesis, the method that best reflects the dissociation constants of thiazole-2-imine derivatives dissolved in an organic solvent, is presented. These molecules have the potential of possessing pharmaceutical properties. For this reason, experimentally determined pK_a values of known molecules were found in the literature. Next, a method determination using a variety of computational methods was conducted, and a computational method was chosen. Using this computational approach, the pK_a values of thiazole-2-imine derivatives, whose pK_a values were unknown from experiments, were computed.

3. HYDROXYL RADICAL-MEDIATED DEGRADATION OF SALICYLIC ACID, METHYLPARABEN AND CAFFEINE: A COMPUTATIONAL APPROACH TO ASSESS THE REACTION MECHANISMS

This chapter is published as: Arslan, E., Hekimoglu, B. S., Cinar, S. A., Ince, N., Aviyente, V., "Hydroxyl radical-mediated degradation of salicylic acid and methyl paraben: a computational approach to assess the reaction mechanisms", Environmental Science and Pollution Research, Vol. 26 (32), 33125-33134, 2019.

3.1. Introduction

In this chapter, the degradation of some of chemicals in fresh water systems have been investigated.

Pharmaceuticals and personal care products (PPCPs) are extensively used in daily life via their presence in human and veterinary medicine, cosmetics, and numerous food products. Inevitably, they are readily discharged to sewage treatment facilities, which they usually bypass and find their way to the aquatic environment with the effluent of the treatment system [24, 25]. The presence of PPCPs in water is a global environmental concern due to their recognition with aquatic toxicity and potential threat to human health [26]. Advanced oxidation processes (AOPs), which are capable of in situ •OH generation, have lately received considerable attention in water treatment, owing to the high reactivity and non-selectivity of the OH radical, and the sludge-free character of the process [27]. The most common methods of AOPs are UV/H_2O_2, ozonation, electromagnetic radiation, TiO_2-photocatalysis, γ-irradiation, and ultrasonication [28]. In all AOPs, the attack of •OH to organic molecules is initiated either by H-abstraction or direct addition, both of which lead to the formation of lower molecular weight and less toxic intermediates. Ultrasonic irradiation of wa-

ter leads to high-energy chemistry via the formation of extreme temperatures and pressures upon the implosion of acoustic cavitation bubbles containing dissolved gases and water vapor [29]. Temperatures and pressures during the implosion process are as high as 5000 K and 2000 atm, at which water molecules and the entrapped gases undergo thermal fragmentation to yield a variety of radical species. Under these conditions, intense mechanical, chemical, and thermal effects are observed. Hence, the use of ultrasound in AOPs is based on the formation of OH radicals upon pyrolytic decomposition of water molecules during the collapse of cavity bubbles. It has been established that if sonication is applied at low frequencies, mechanical effects of ultrasound dominate, whereas at high frequencies chemical effects are more important. Although the full mechanism has not been so far elucidated, it is accepted that low frequencies (20–80 kHz) preferentially lead to mechanical and physical effects, while high frequencies (150–2000 kHz) favor the production of more HO radicals, leading to more intense chemical effects [30]. The attractiveness of sonochemistry in AOPs arise from the fact that no chemicals are required to produce radical species, and no waste sludge is produced as a byproduct of the process. Nevertheless, the challenge remains, as the energetic and financial competitiveness of the method to the existing processes performed in silent conditions are still under investigation [31].

This chapter includes three PPCPs commonly found in fresh water systems: Salicylic acid (SA)-methylparaben (MP)-caffeine (CAF).

Salicylic acid (SA) and its derivatives constitute a group of PPCPs that are commonly used in analgesics and antipyretics, and are available in a wide variety of formulations. Overdosing of the compound has been reported as toxic [32] leading to symptoms such as headache and nausea [33]. SA is frequently detected in water and wastewater systems at relatively high concentrations, indicating the necessity to eliminate it before the water is reclaimed and transferred to distribution systems [34]. Published studies related to sonochemical degradation of SA have shown that the compound is highly reactive with OH radicals to readily yield dihydroxybenzoic acid, followed by the production of smaller aromatic compounds [35]. Note also that salicylic

acid is commonly used as a model compound to monitor the formation of hydroxyl radicals in advanced oxidation or other chemical reactions [36].

Parabens are another group of PPCPs that are widely used in toothpaste, cosmetics, textiles, foodstuff, and beverages, owing to their preservative and antibacterial properties [24]. It has been reported that annually, 8000 tons of parabens are produced and a large fraction of them are introduced into the aquatic system via excretion and waste discharge [37]. Among all, methyl paraben (MP) is the most highly consumed paraben and is classified under "emerging contaminants" not only because of its aquatic toxicity, but also due to its potential contribution to the incidence of breast cancer [38].

Over the past years, many varieties of AOPs (ozonation, photocatalysis, electrochemical oxidation UV-irradiation [28]) have been extensively used for the degradation of MP, based on the power and non-selectivity of •OH in oxidizing organic molecules. Many of these studies, however, have focused mainly on setting the optimum experimental conditions, disregarding the reaction mechanism [39–41]. There are some studies reporting the attack of the •OH to both the benzene ring and the alkyl chain, based on the identified intermediates. However, the specific degradation route for MP is still unclear. In a study by Gao *et al.*, all possible degradation pathways of MP were elucidated by quantum chemical calculations and the toxicity of all intermediates was estimated [40]. However, to the best of our knowledge, there are no computational studies reporting the intermediates formed by bond rupture mechanisms. The literature on sonochemical decomposition of MP at high-frequency ultrasound is limited to a few studies focusing on the optimization of the operational parameters and integration of the method with clay minerals [42] and investigation of the driving mechanisms as hydroxylation and hydrolysis [43]. The latter research has also reported complete mineralization of the compound as the evidence of the high efficiency of sonochemistry compared with all other AOPs. The purpose of this study was to investigate the degradability of SA and MP by high-frequency ultrasound to elucidate the reaction kinetics and to propose reaction mechanisms that support the experimental results. Additionally, the study covers the prediction of the aquatic toxicity of the parent and

oxidation byproducts at three test levels to test the environmental safety of the ultrasonic method as an AOP.

The reaction of SA and MP with the medium (•OH, H_2O) was considered by using the basic principles of organic chemistry such as •OH addition, H abstraction, water elimination, bond rupture, and aromatization with a goal of rationalizing the formation of the experimentally observed byproducts.

3.1.1. Methodology

The proposed mechanisms were modeled using the density functional theory (DFT), the B3LYP functional and the 6–31 + G(d) basis sets [44–46]. A conformational search around single bonds was performed for the structures corresponding to local minima, the energetics discussed in the results are based on the global minima. Transition state structures were characterized by a single imaginary frequency. All the quantum chemical calculations were performed using Gaussian 09 software package [47]. The reactions were conducted in water (ϵ = 78.36) environment using the integral equation formalism polarizable continuum model (IEF-PCM) at 298K. The intrinsic reaction coordinate (IRC) calculations were held to prove that each transition state is connected to the corresponding reactants (RC) and products (PC). The acute toxicities of SA, MP, and their oxidation byproducts to green algae, daphnia, and fish were estimated in terms of maximum effective (EC_{50}) and lethal (LC_{50}) concentration using the Ecological Structure Activity Relationships software (ECOSAR 2014). Utilizing computational Structure Activity Relationships (SARs), the program calculates a chemical's acute (short-term) and chronic (long-term or delayed) toxicity to aquatic species, including fish, aquatic invertebrates, and aquatic plants. The program's features include forecasting the toxicity of novel or untested industrial chemicals, grouping structurally comparable organic molecules with experimental effect levels that are associated with physical and chemical attributes are combined, programming of a categorization system to determine which class is best representative for novel or unproven substances and updating aquatic toxicity continuously depending on experi-

mental results provided or collected from both public and private sources. For caffeine calculations, reactions were modeled using a quantum chemical computational method (Density Functional Theory) incorporated in the Gaussian 16 software [48]. B3LYP functional, and the 6-311G (d, p) basis sets were used for the conformational search of each stationary point [45, 46], [49, 50]. The triple zeta basis set is used trace the possible interactions of the proton with the nitrogen atom in the protonated caffeine during the process.

Transition state structures were characterized by a single imaginary frequency, and reactions were modeled in water ($\epsilon = 78.36$) at 298 K using the Integral Equation Formalism Polarizable Continuum Model [51]. The intrinsic reaction coordinates (IRC) were calculated to verify the relation of each transition state with the corresponding reactants and products. The oxidation reactions were modeled based on the attack of the OH radical to the reactant, while those occurring in water without AOP's (as verified by the peaks depicted in LCMS chromatograms at t = 0) were modeled assuming water as the attacking reactant. The energy barrier for the formation of each byproduct was referred to as activation Gibbs free energy $\Delta G^{\ddagger}$, and found by subtracting the Gibbs from free energy of the transition state from that of the corresponding reactants for both homogenous and heterogenous processes. In heterogenous processes, Quantum Espresso software was used to calculate the binding energies of the by-products on a TiO_2 surface. Quantum Espresso is a set of tools for modeling materials and performing first-principles computations on molecules' electronic structures. It is based on plane wave basis sets, pseudo potentials, and density-functional theory (both norm-conserving and ultra soft). The Quantum Espresso project has started in 2002, when three packages for DFT simulations were merged. These packages were PWscf (code for self-consistent field (SCF) solutions of Kohn–Sham equations), CP (a code for MD simulations of the Car–Parrinello type), and FPMD (a code similar to CP, but with a different and partially overlapping set of functionalities) [52]. It can be used for structural optimization, transition states, minimum energy path, ground state calculations, generation of pseudo potentials, surface calculations [53, 54].

In this study, PBE functional was used for ground state calculations. Hubbard U (1) (correction factor) was taken as 1.0. TiO_2 anatase surface with 001 facets was used for calculations.

3.2. Results and Discussion

3.2.1. Experimental

Normalized concentrations of SA and MP during 30-min sonolysis at 572 kHz are plotted in Figure 3.1. The fitted curves ($R^2 = 0.98$) show that the rate of reaction in both cases was pseudo first order, with rate constants of 0.067 and 0.091 min^{-1} for SA and MP, respectively. The inset at the top of the figure shows the degree of C-mineralization after 30-min sonolysis and that after a 24-h silent maintenance in the dark. It was interesting to find that the mineralization process continued for a long time after the ultrasonic generator was shut down, as the indication of "latent cavitation", i.e., the effect of stable reactive oxygen species (ROS) such as H_2O2 that form upon the violent collapse of acoustic cavitation bubbles.

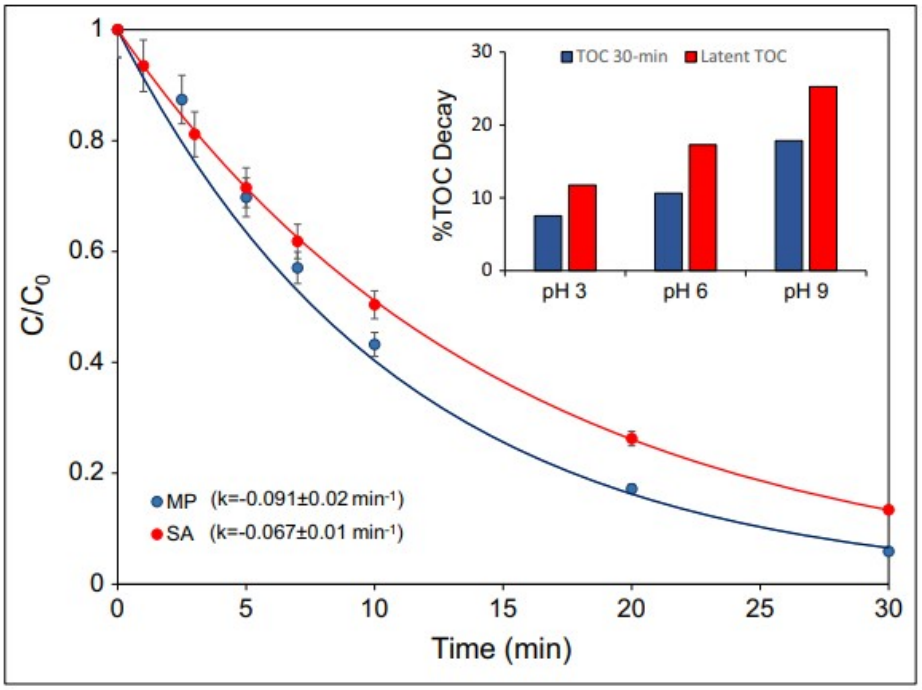


Figure 3.1. Normalized plots of concentration-time showing the rate of sonochemical degradation of SA and MP (C^0 = 10 mg/L) at pH 3.0. The bar chart at the right bottom shows the corresponding C-mineralization.

LC/MS/MS spectra of 10-min sonicated samples of SA and MP are presented in Figure 3.2, respectively. For SA, four major peaks at molecular weights of 62, 108, 110, and 154 g/mol with intensities of 2.44, 1.53, 1.36, and 1.83 were detected, respectively. For MP, six major peaks at molecular weights of 85, 108, 110, 134, 138, and 150 g mol^{-1} and intensities of 1.42, 0.30, 0.15, 0.50, 0.42, and 0.15 were also detected, respectively. The results for SA are consistent with those of Scheck and Frimmel, who identified benzoquinone (MW = 108) and dihydroxybenzoic acid (MW = 154) as the byproducts of UV/peroxide/O_2 process [55]. The latter was also reported by Chang *et al.*, together with catechol (MW = 110) as the byproducts of SA oxidation in a Fenton process [56]. In another study by Guinea et al., using electrochemical oxidation techniques, dihydroxybenzoic acid, maleic acid (MW = 116), fumaric acid (MW = 116), and malic acid (MW = 134) were identified as the byproducts of SA degradation (Guinea *et al.* 2008). A photocatalytic study of MP by Lin et al. with TiO2 has shown that the intermediate products were dihydroxybenzene (MW = 110), 4-hydroxybenzoic

acid (MW = 138), tartaric acid (MW = 150), and benzoquinone (MW = 108), which are consistent with our findings [57]. Another study done by Gmurek *et al.* focused on the degradation of hazardous water contaminants: methyl-, ethyl-, propyl-, butyl-, and benzylparaben using ultraviolet C lamps in the absence and presence of H2O2 [58]. From the byproduct analysis, it was concluded that the main degradation products were p-hydroxybenzoic acid (MW = 138); 2,4 dihydroxybenzoic acid (MW = 154); and 3,4 dihydroxybenzoic acid (MW = 154). These experimental results are consistent with the computational results.

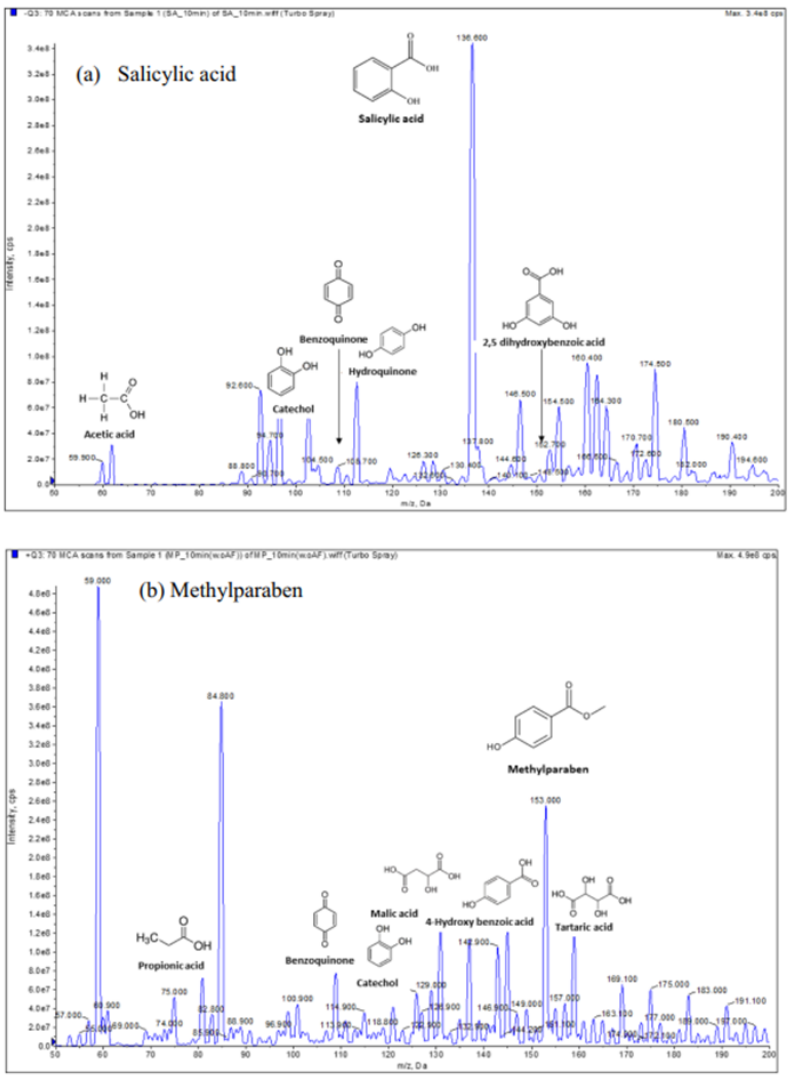


Figure 3.2. LC/MS/MS spectra of 10-min sonicated samples of SA (a) and MP (b) and structures of their identified byproducts.

3.2.2. Salicylic Acid

The energetics for the degradation mechanism of SA can be monitored step by step in Figure 3.3 and Figure 3.4. The transition state geometries are displayed in Appendix Section A as Figure A.1.. One can also find the toxicity results in Table A.1. and A.2.

SA is primarily attacked by •OH at the ortho/para positions of the OH since these are favorable positions for the ortho/para director OH group and the meta director carbonyl group which are already present on the ring. The ortho and para positions to •OH are attacked with very low barriers of 4.3 kcal/mol and 6.3 kcal/mol; the reaction is exothermic by 6.0 kcal/mol (SA6) and 3.2 kcal/mol (SA1) respectively. It is very likely that SA1 is attacked barrierless by another •OH to form the neutral product SA2 with very high exothermicity $\Delta G^{\ddagger}$ = -60.1 kcal/mol). At this stage, water elimination (TS-(SA2-SA3)) is very costly ($\Delta G^{\ddagger}$ = 41.8 kcal/mol) to yield the stable product SA3 (ΔG_{rxn} = -95.9 kcal/mol), which is detected experimentally (MW = 154). SA8, the isomer of SA3 is formed via similar steps. We suggest the other isolated intermediate SA5 (MW=110) to be formed in a stepwise manner, first a hydrogen shift from position 6 to position 5 forming SA4 and then decarboxylation. The driving force for this last reaction would be the high stability of the final product SA5 (ΔG_{rxn} = -110.5 kcal/mol). Overall, the stabilities of the intermediates SA3 (MW = 154) and 1,4-benzenediol/hydroxyquinone (SA5, MW = 110) trigger their formation. Note that SA8 yields 1,2-benzenediol/ catechol (SA10, MW=110) similarly along exothermic reactions. The formation of SA10 (MW=110) follows the same pattern as SA5.

The formation of 1,4-benzoquinone (MW=108) can be visualized in Figure 3.4 as a continuation of Figure 3.3. In this mechanism 1,4-benzenediol (SA5) gains easily 2 •OH groups sequentially to form the very stable species SA12 (-47.3 kcal/mol). SA12 yields SA14, namely 1,4-benzoquinone (MW = 108) after two consecutive H_2O eliminations. Further reactions to form the intermediate with MW = 126 involves bond rupture (TS-(SA14-SA15)) which is highly energetic (79.9 kcal/mol) to yield the

product SA15 (3.5 kcal/mol). SA15 undergoes another bond breakage with a barrier of 78.6 kcal/mol to yield SA16 (-6.2 kcal/mol) that abstracts H_2O to yield SA17 (MW = 108), (4.6 kcal/mol). SA17 can undergoes bond rupture to form SA18 (MW = 60) which was detected experimentally (-5.9 kcal/mol).

Overall, reaction profiles yielding the experimentally detected intermediates (SA18 (MW=60), SA14 (MW = 108), SA5 and SA10 (MW = 110), SA3 and SA8 (MW = 154)) have been suggested. Although Gibbs free energies of activation yielding the desired intermediates are high, the products are stable, the reactions are thermodynamically driven.

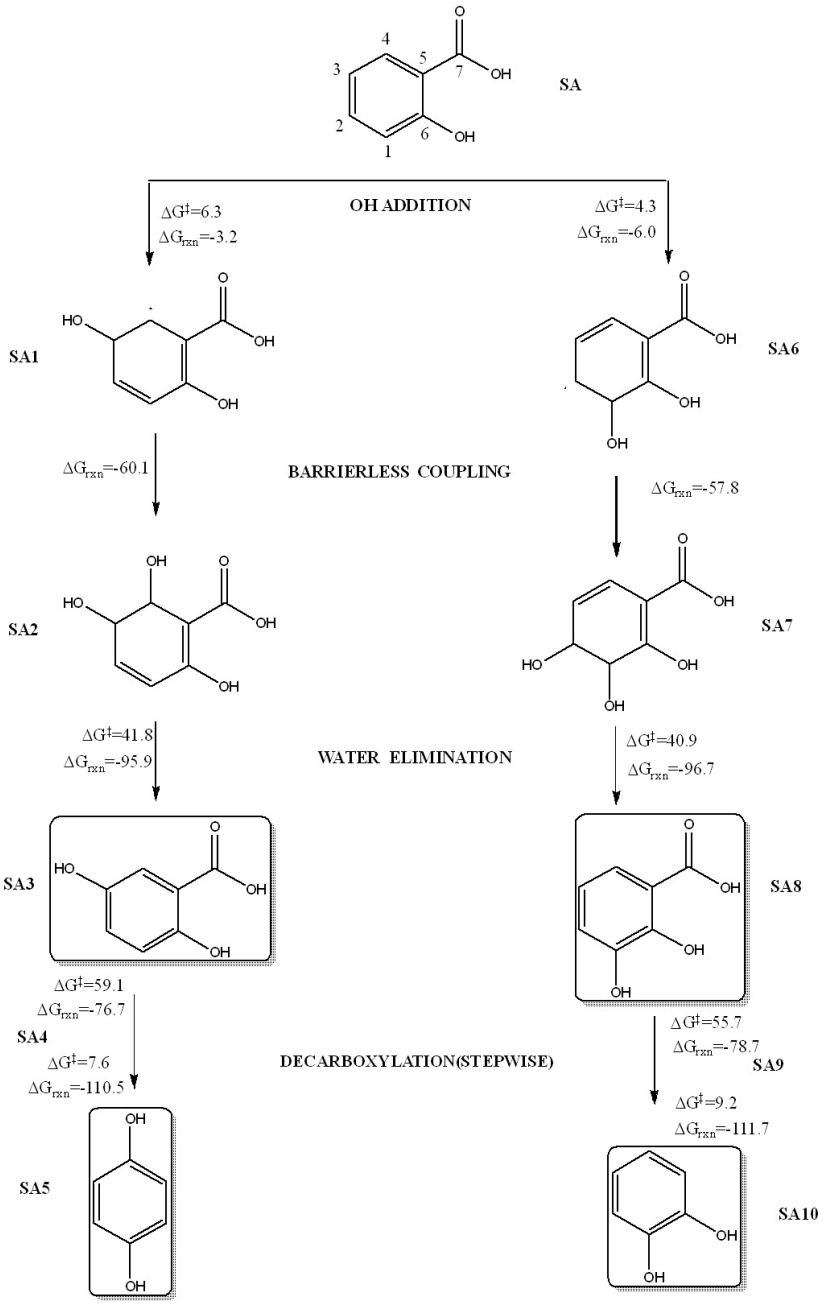


Figure 3.3. Gibbs free energies (kcal/mol) for the degradation pathway of SA into SA5 and SA10 (B3LYP/6-31+G*/IEF-PCM/water).

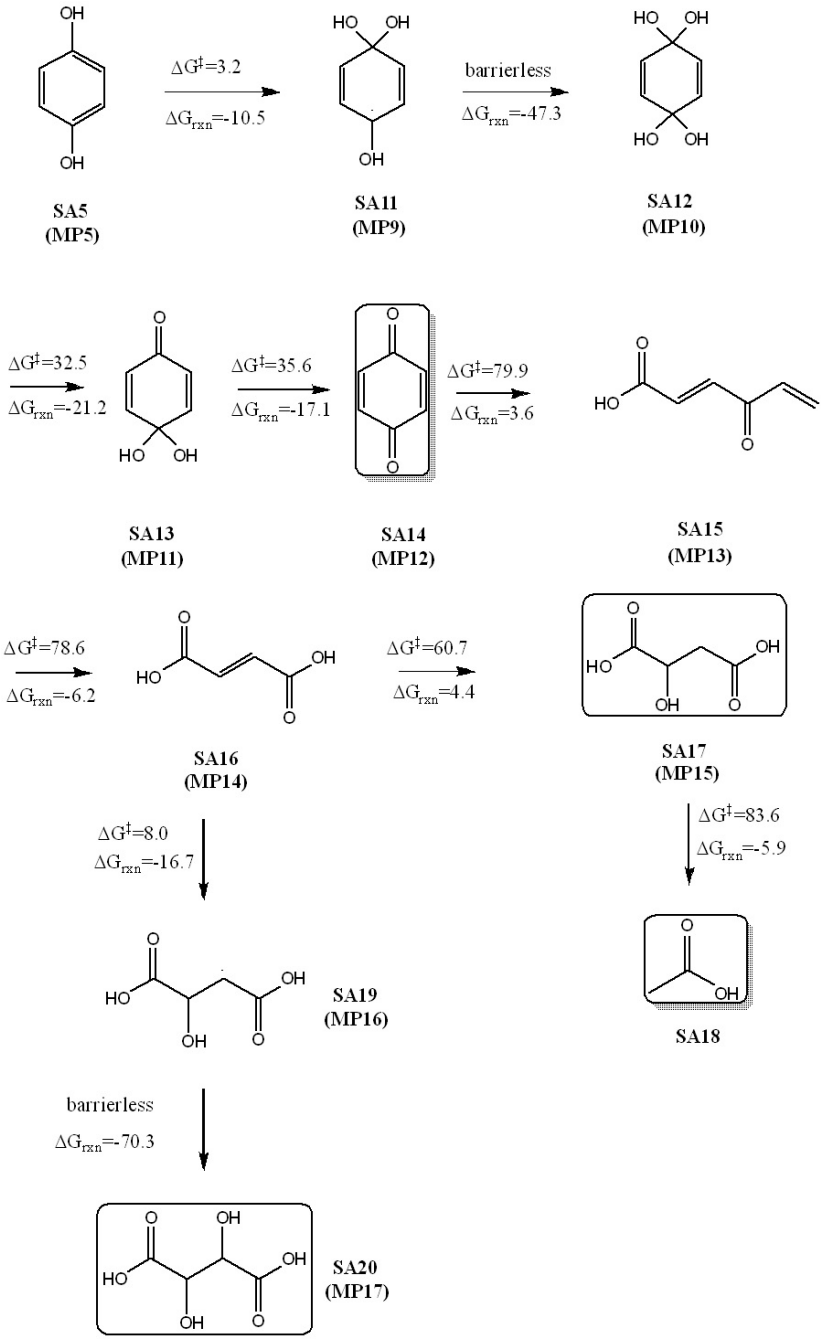


Figure 3.4. Gibbs free energies (kcal/mol) for the degradation pathway of SA5 into SA20 (B3LYP/6-31+G*/IEF-PCM/water).

3.2.3. Methyl Paraben

Figure 3.4 and Figure 3.5 show the proposed mechanisms in this work for the degradation of MP. All the transition state geometries of MP are displayed in Figure A.2. Moreover, one can see the toxicity results in Table A.1 and A.2. The energetics have been calculated based on the stoichiometrically balanced reactions. In water, MP (MW=150) may undergo hydrolysis yielding MP1 with a barrier 59.8 kcal/mol and the reaction is endothermic by 2.0 kcal/mol. Note that the latter intermediate MP1 (MW = 138) was also detected experimentally (Figure 3.2 (b)). The reaction is followed by decarboxylation to obtain MP2 (-12.1 kcal/mol). Then, MP2 (MW = 94) is attacked by •OH, with addition to the ortho and para positions yielding MP3 (MW = 112) and MP6 (MW = 112) with low barriers and high exothermicities ($\Delta G^{\ddagger}$= 4.8 kcal/mol and 5.6 kcal/mol; ΔG_{rxn}= -16.6 kcal/mol and -17.9 kcal/mol, respectively). It is very probable that MP3 and MP6 are attacked barrierless by another •OH to form the neutral species MP4 (MW = 128, -65.7 kcal/mol) and MP7 (MW = 128, -69.2 kcal/mol). At this stage, elimination of water (TS-(MP4-MP5) and TS-(MP7-MP8)) is costly (34.2 kcal/mol and 40.8 kcal/mol, respectively) to yield exothermically MP5 (MW=110) and MP8 (MW=110); the latter stable intermediates have also been detected experimentally. Note that MP5 and MP8 are identical to SA5 and SA10 respectively, thus the reactions that follow the degradation byproducts of MP5 and MP8 will be similar to the ones of SA5 and SA10 and have already been discussed in the previous section. The mechanism proposed above illustrates that •OH-addition is the governing route in the degradation of MP, as also reported in the literature [59] for shorter alkyl chain parabens. A study in the literature has investigated the effect of alkyl chain length on the degradation mechanism of parabens and found that •OH-addition was the more dominant route for the degradation of shorter parabens, such as MP [60]. The same authors have also demonstrated the dominance of •OH-addition over H-abstraction for the degradation of diclofenac [61]. As mentioned in Section 3.1, the intermediate products of oxidation detected by Lin et al. [57] for the degradation of MP are consistent with our results, i.e. dihydroxybenzene (MW = 110), 4-hydroxybenzoic acid (MW = 138), tartaric acid (MW = 150) and benzoquinone (MW

= 108) (Figure 3.5).

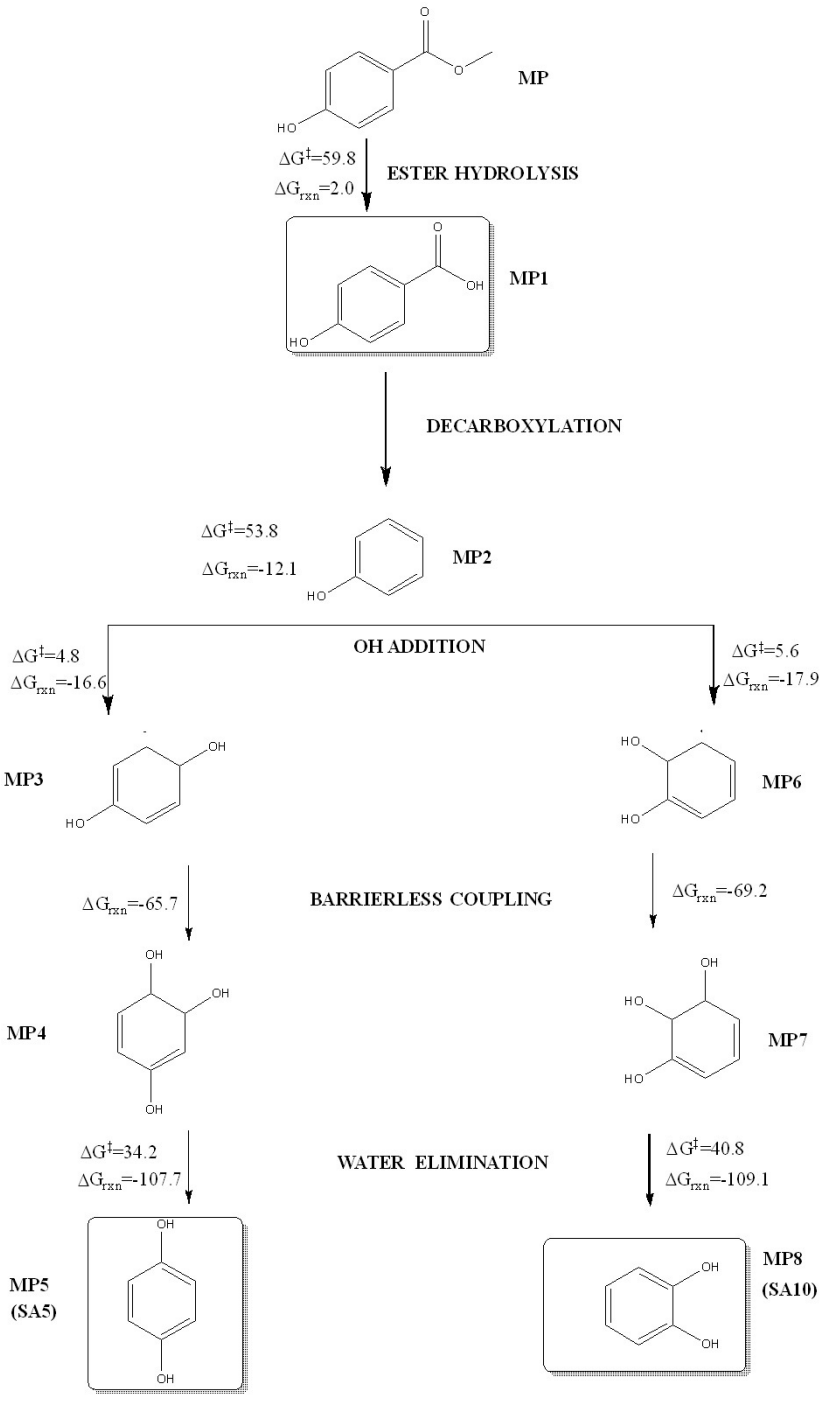


Figure 3.5. Gibbs free energies (kcal/mol) for the degradation pathway of MP into MP5 and MP8 (B3LYP/6-31+G*/IEF-PCM/water).

3.2.4. Caffeine

This study is about the degradation of caffeine that is detected in large amounts in fresh water systems. There are two parts in this study: homogeneous (no catalyst) and hetereogeneous (with catalyst (TiO_2)).

Caffeine is a common compound of PPCP's and an ingredient of coffee, tea, chocolate, cocoa and beverages; while it is widely used in the production of prescription medicines and personal care products [62]. Upon consumption, the compound and/or its metabolic byproducts are readily discharged (via wash water and excretion) into sewage treatment facilities with low biodegradability. Hence, caffeine is not only detected frequently in effluents of wastewater treatment operations, but also in soil, groundwater and fresh water systems [63]. Lately, the presence of caffeine in the water environment has been a serious environmental and public concern, due to its toxic effects to aquatic organisms, and potential health threat to humans as hyperactivity and cardiovascular diseases [64]. Consequently, a lot of research has lately focused on finding effective processes that control the discharge and/or the presence of caffeine in water supply systems. In this study, many AOPs were used as mentioned in Section 1.2.

3.2.4.1. Homogenous Degradation of Caffeine. Firstly, homogenous degradation (US/UV and US/UV-H_2O_2) of caffeine will be discussed in acidic medium (pH=4).

LCMS analyses of the reactor samples collected at t = 0, 30 and 60- min during single and hybrid processes revealed numerous fragments and byproducts, most of which were detected at retention times of 7.80 and 8.33 min [65]. Moreover, all samples of time zero showed not only a major peak (m/z 194) corresponding to the mother compound (caffeine), but also a variety of others (e.g. m/z 60, 83, 114, 123, 165, 187, 195, 210, 236), indicating fragmentation of the molecule at acidic pH, and/or reaction with water molecules (Figure 3.6, Figure 3.7, Figure 3.8 and Figure 3.9).

The first mechanism (a) shows the reaction of protonated caffeine with water leading to the rupture of C1-C14 bond to yield an intermediate P1 (MW=213) with a high energy barrier ($\Delta G^{\ddagger}$=42.4 kcal/mol, ΔG_{rxn}=16.6 kcal/mol). The reaction of P1 with another molecule of water yields two intermediates with MWs of 98 and 117 (ΔG_{rxn}=1.5 kcal/mol). The former is in protonated form (MW=99) and undergoes demethylation to produce another P1 ($\Delta G^{\ddagger}$=38.2 kcal/mol, ΔG_{rxn}=6.5 kcal/mol) with a high energy barrier, as also reported by Dalmazio et al, for demethylation of caffeine in protonated water [66] The second mechanism (b) shows the formation of P2 (MW=125) via the reaction of caffeine with water to yield an intermediate (MW=140) through bond rupture, followed by demethylation ($\Delta G^{\ddagger}$ =54.5 kcal/mol, ΔG_{rxn}=32.9 kcal/mol). In general, high energy barriers of bond rupture and demethylation reactions are well known and reported in the literature [67].

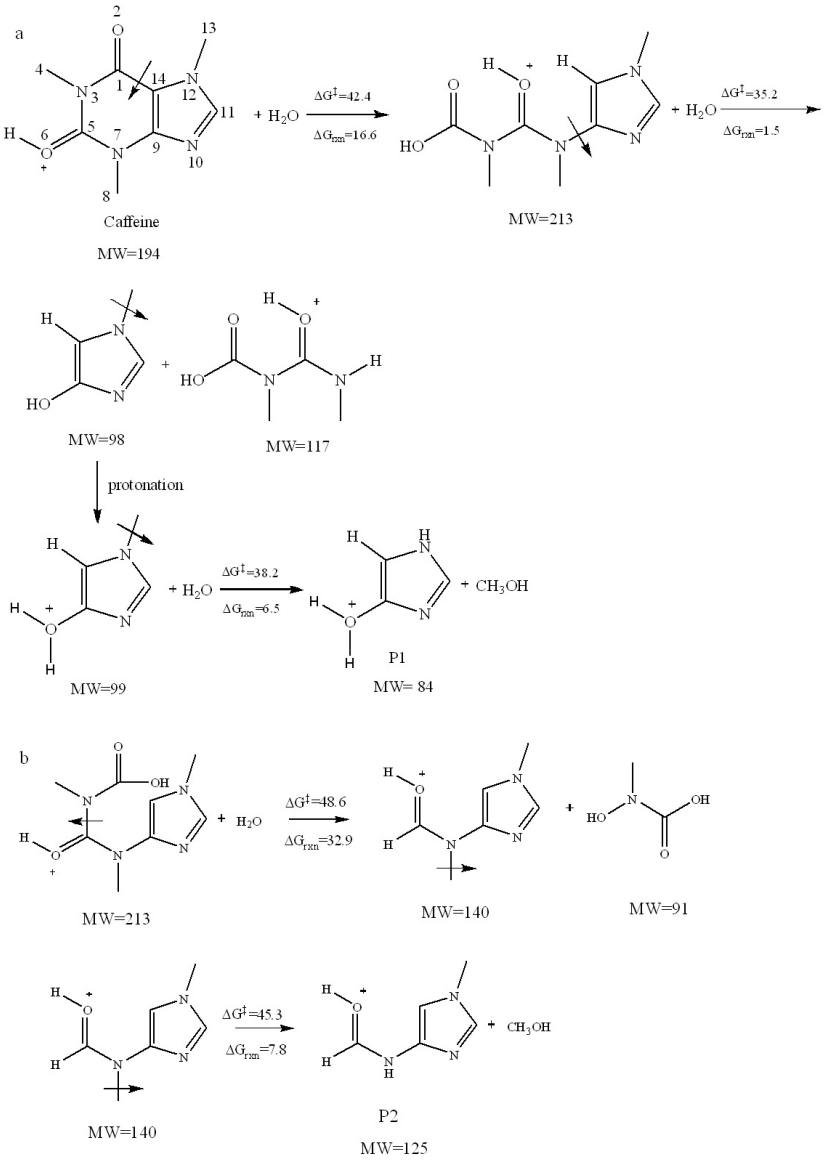


Figure 3.6. The reaction pathways and Gibbs free energies (kcal/mol) of caffeine decomposition in acidic water (pH 4), leading to the formation of P1 (a) and P2 (b) (B3LYP/6-311G(d,p)/SMD/water).

The most probable reaction mechanisms for the degradation of caffeine in neutral medium by •OH-mediated oxidation are presented in Figure 3.7. The first one (a) shows the formation of P3 (MW=221) and P4 (MW=228), which were both the experimentally byproducts.Accordingly, caffeine is first attacked by an OH radical at the bridge-side, leading to the formation of P3 ($\Delta G^{\ddagger}$=3.1 kcal/mol, ΔG_{rxn}=-3.3 kcal/mol),

which most likely reacts spontaneously with another •OH via electron-coupling to yield exothermically the product P4 (ΔG_{rxn}=-51.7 kcal/mol). On the other hand, •OH may also attack the relatively positive carbonyl group of caffeine at C1 to produce P32 with a relatively low barrier ($\Delta G^{\ddagger}$=23.7 kcal/mol, ΔG_{rxn}=7.5 kcal/mol), which reacts readily with a very low barrier (2.1 kcal/mol) to produce an intermediate with the same MW as P3, referred in the scheme as P33. The compound is also very reactive and is attacked spontaneously by an •OH via electron coupling (ΔG_{rxn}=-95.7 kcal/mol) to yield the enolic structure $P4_2$, which is readily converted to $P4_3$ via keto-enol tautomerism ($\Delta G^{\ddagger}$=23.4 kcal/mol, ΔG_{rxn}=-7.3 kcal/mol). The reaction of P43 with •OH gives an intermediate with MW 245, which undergoes bond rupture through a highly exothermic reaction (ΔG_{rxn}=-36.4 kcal/mol) to produce two intermediates with MWs of 120 and 126. The latter is a radical, and reacts barrierlessly with •OH to produce exothermically two intermediates with MWs of 119 and 143 (ΔG_{rxn}=-41.7 kcal/mol). The former (MW 119) is demethylated via reaction with a water molecule, leading to the formation of P5 (MW=105), which was one of the most commonly detected byproducts, as it was found in all process effluents. Finally, P5 undergoes reductive decomposition and is converted to P6 (MW=86), which was also detected experimentally. Reductive cleavage is a likely reaction in all sample solutions, due to the presence of excess air bubbles and nitrogen species, as NOx.

In general, the reaction mechanisms proposed in Figure 3.6 are in good agreement with the corresponding LCMS analyses of the samples. The intermediate product P1 was found to form with a reasonable barrier by the reaction of caffeine with water. Nevertheless, it was detected in all process samples at high intensity, indicating the resistance of the compound to oxidation, most likely due to the unique cyclic structure. The formation of P3 and P4 via reaction with •OH was much easier, as they proceeded with lower energy barriers. Moreover, these products were detected in all samples except those at t=0, and at lower intensities than that of P1, signifying their ease of reaction with •OH. More significantly, the intensity of both was reduced in hybrid operations, by which the abundance of OH radicals was increased via additional routes and/or sources of generation.

Homogenous degradation of caffeine has been discussed in neutral medium (Figure 3.7). Firstly, OH radical attacks to the caffeine giving a radical (MW=211) and this radical captures H radical from the medium barrierlessly. Therefore, formation of 212 molecule is easy. However, it can return the 194 by elimination of a water molecule. It is consistent with the experimental data since there is no 212 molecule in 30 minutes reaction but it was formed in high concentration in 60 minutes reaction. Moreover, byproduct with MW=183 increased from 30 minutes to 60 minutes in heterogenous part and it was consistent with heterogeneous process (Figure 3.8. 2. Pathway reaction) since caffeine molecule can release -CH_3 from three different places and the activation barrier is not very high (34.4 kcal/mol). Then, the intermediate (MW=183) turned into 165 by giving another -CH_3 group with reactiın barrier 33.2 kcal/mol and then 165 molecule was converted to 149 again the same procedure (-CH_3 breakage). Actually, 149 molecule should appear in LC/MS analyses because generating 170 from 149 has high activation barriers (Figure 3.8, path 4). It is inconsistent with our proposed mechanism but this reaction pathway is proposed pathway. There can be many plausible reaction pathways. Therefore, there may be some conflicts about the data. After generating 170, 59 and 83 molecules were formed that were also experimentally detected compounds.

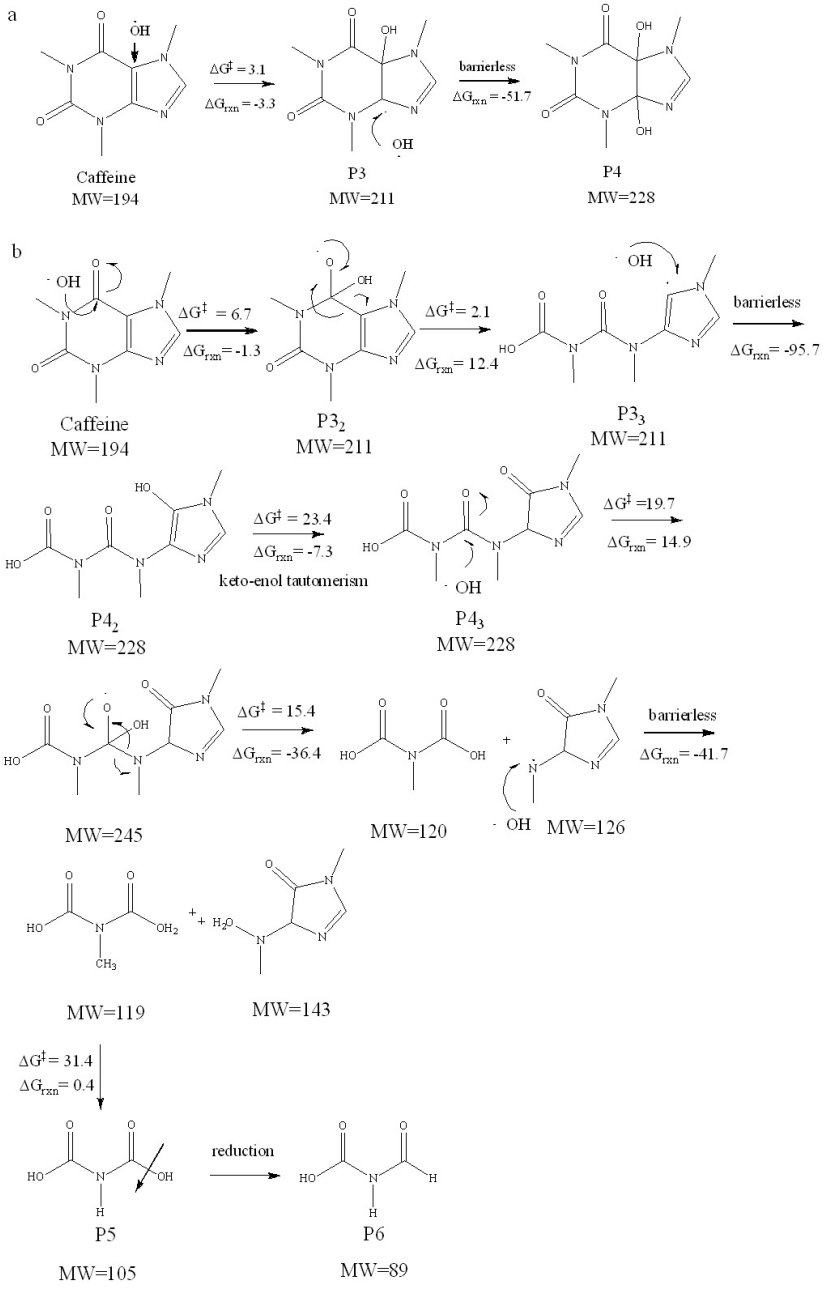


Figure 3.7. •OH-mediated homogeneous reaction pathways and Gibbs free energies (kcal/mol) for the oxidative degradation of caffeine at neutral medium to produce P3, P4 (a), $P3_2$, $P3_3$, $P4_2$, P5 and P6 (b) (B3LYP/6-311G(d,p)/SMD/water).

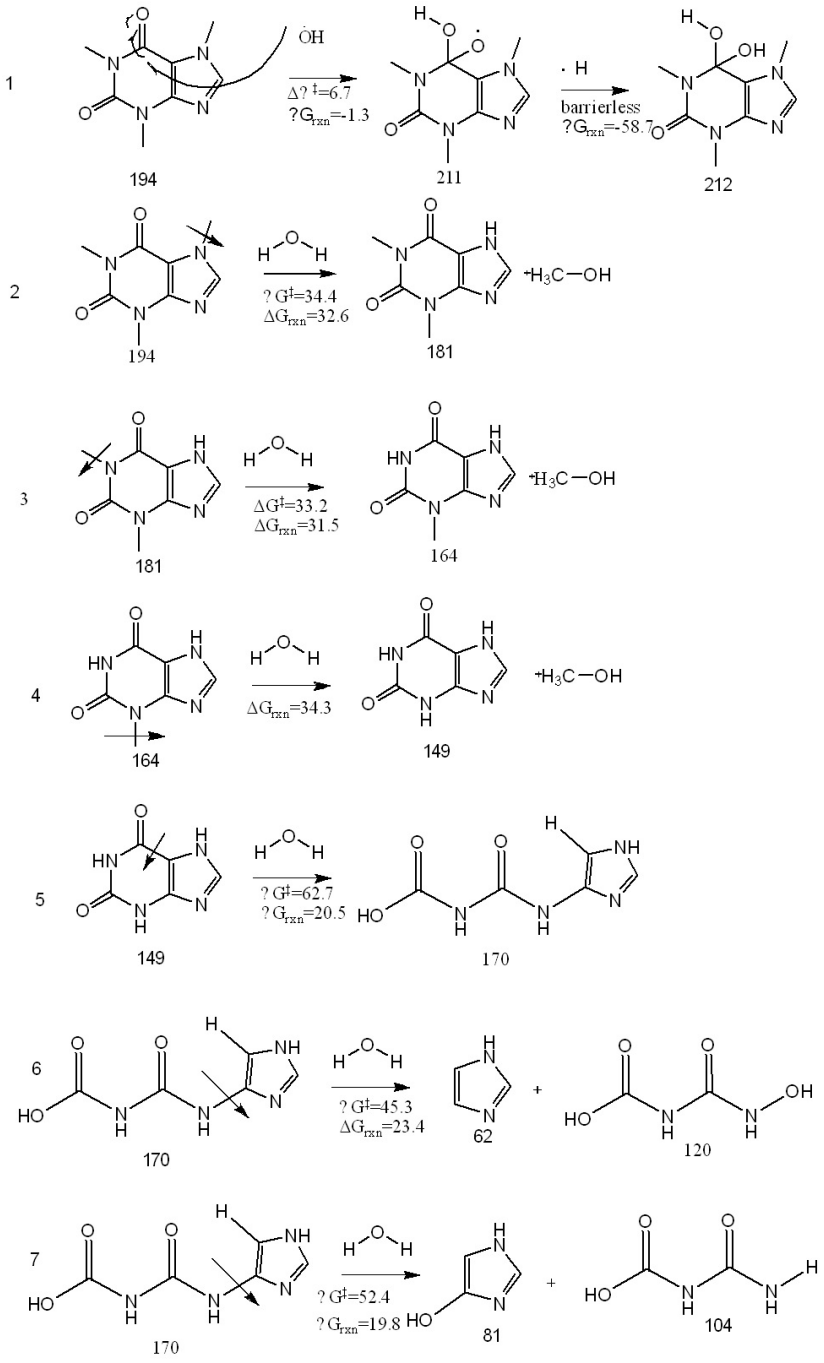


Figure 3.8. H_2O-mediated degradation pathway of caffeine in neutral homogenous medium (B3LYP/6-311G(d,p)/SMD/water).

3.2.4.2. Heterogeneous Degradation of Caffeine on TiO_2 surface.. The study described herein has shown that the emerging water contaminant caffeine is easily decomposed by

singly applied ultrasonic irradiation (577 kHz) and UV-H_2O_2 photolysis at optimized conditions, but with insufficient mineralization. The efficiency of the degradation reactions was significantly improved via hybrid applications, particularly when operated at a heterogeneous mode using commercial TiO_2 (P25) as the catalyst. It was found that under equivalent conditions, the percentage of TOC decay increased from 21.7 by single photolysis (UV-H_2O_2) to 35.4 by US-assisted photolysis. The result was attributed to the formation of excess •OH and H_2O_2 (via cavitation collapse), and enhanced mixing of the solution, that facilitated the transmission of UV-light. Similarly, mineralization increased from 60% by TiO_2-photocatalysis to 66% by US-assisted photocatalysis, owing mostly to the mechanical actions of ultrasound for cleaning and disintegrating of solid particles, thus increasing the number of fresh adsorption sites. Assessment of all hybrid processes for their synergy in TOC decay showed that the best was UV-H_2O_2/TiO_2 with a synergy index of 65 (as opposed to an index of 0.25 in US/UV hybrid). It was also found that further hybridization by adding ultrasound to the above process drastically reduced the synergy to 2.8, due to the presence of excess H_2O_2 that scavenged OH radicals, and the formation of cavity clouds and high density particle layers that inhibited the transmission of UV light to solution. The reaction mechanisms proposed using the computational technique described in the text were consistent with the experimental data, as many of the byproducts predicted were detected by LCMS analysis of the samples.

Table 3.1. Experimental results of degradation of caffeine (Binding energies (kcal/mol) are calculated by using Quantum Espresso) Retention times: 8.3-7.9.

MW of byproducts	Reaction time-Homogeneous		Reaction time-Heterogeneous		Binding energies
	30 min	60 min	30 min	60 min	
	Intensities (x10^3)				
210	——	25-60	——	60-80	-279.8
194 (CAF)	130	——	——	——	-245.8
183/187	130-170	210-220	40-60	70-80	-222.8
165	40-48	——	20	20-30	-204.9
100	70-78	90-100	100-110	110-120	-184.2
83	220-148	200-125	295-130	180-180	-183.2
59	240-330	240-390	310-480	330-410	-108.7

In Figure 3.9, binding of by-products are shown to the TiO_2 surface. The distance between atoms of byproducts and surface in consistent with the binding energies (Table 3.1). As the moleculer weight of product increases, the distance decreases meaning that molecule binds to the surface stronger and degrade easier.

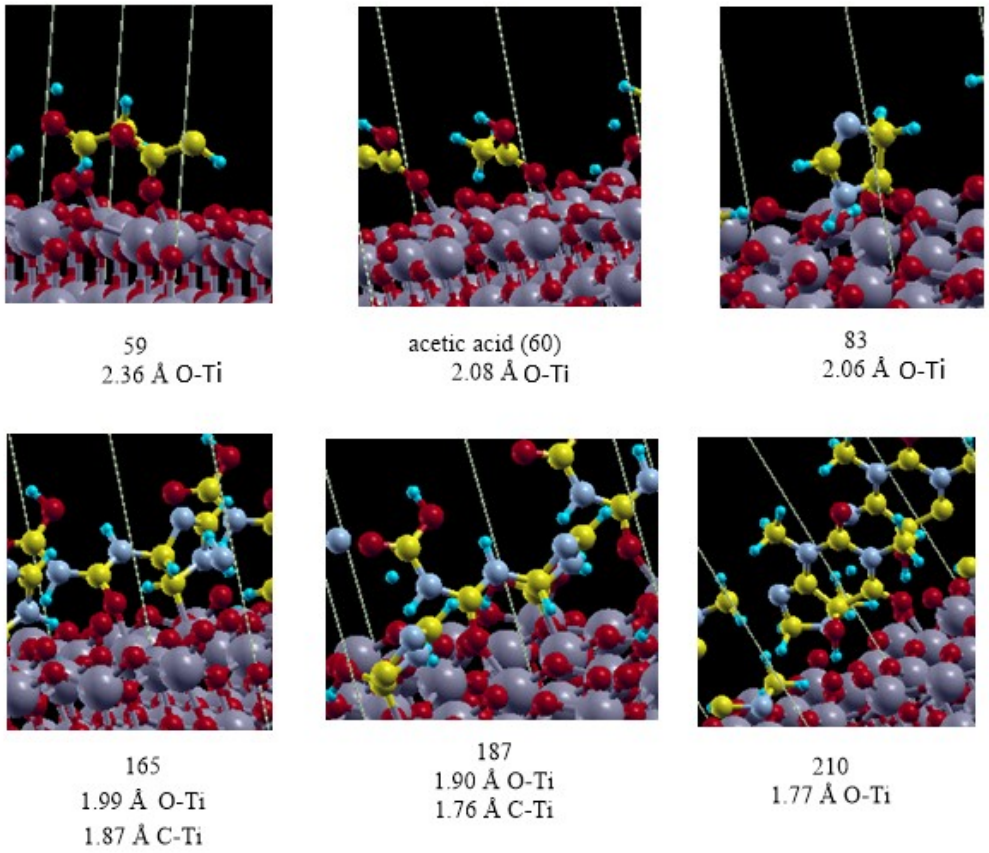


Figure 3.9. 3D images of experimentally identified by-products of caffeine binding to the surface (TiO_2).

As a conclusion, caffeine is one of mostly detected contaminant in fresh water systems. Therefore, it is very important to decontaminate it from the environment since it can cause hearth attack, anxiety or arrhythmia and thousands of tons of caffeine interfere seas, lakes and finally full ecosystems affected. According to some researches [68], caffeine is generally first detected compound in wastewater, surface water, ground water, drinking water, sediment, soil, sludge, plant and animals. Therefore, incomplete mineralization of caffeine can affect all the systems totally.

4. STATISTICAL ASSESSMENT OF THE MODELING OF PROTEINS AND LIGANDS BLIND CHALLENGES

This chapter is published as:

- Arslan, E., Findik, B.K., Aviyente, V., Journal of Computer Aided Molecular Design, 34, 463-470, 2020. (SAMPL6)

- Fındık, B.K., Haslak, Z.P., Arslan, E., Aviyente, V., Journal of Computer Aided Molecular Design, 35, 841-851, 2021. (SAMPL7).

4.1. Introduction

In this chapter, some important properties of drug-like molecules such as pK_a and LogP have been evaluated. The Statistical Assessment of the Modeling of Proteins and Ligands (SAMPL) challenges are widely known, the blind challenges are intended to advance the computational tools in rational drug design [69–72]. Over the last decade, the challenges have became attractive for computer-aided drug design (CADD) around the world since they have been basically based on the calculations of several molecular properties of drug-like small molecules. Up to now, many SAMPL challenges were organized in order to calculate some important chemical properties of drug-like molecules. In SAMPL0 challenge, two different methods of solvent-free energy estimation were performed in the prospective test: an implicit solvent approach based on the Poisson-Boltzmann equation and an explicit solvent approach based on alchemical-free energy calculations. The next challenge was SAMPL1 challenge. The purpose of this challenge was to predict the solvation-free energies of 63 different molecules that couldn't be calculated accurately with the current methods. SAMPL2 challenge represented the third annual blind prediction of transfer energies and was included in the challenge for the first time in the tautomer ratios. More than 60 sets of predictions were submitted and

each participant also tried to estimate the error in their predictions. SAMPL3 challenge was related with the binding affinities of host-guest systems. Second-generation mining minima method (M2) was used to compute the binding affinities of target host-guest systems. The predictions weren't close to the experimental values. Therefore, the force-field was changed and finally concluded that the binding energies were highly sensitive to the force-field used. Another challenge, SAMPL4, was about the binding energies of host-guest systems, HIV integrase inhibitors and hydration free energies of small molecules. In SAMPL5 challenge, participants tried to predict distribution coefficients and host-guest interactions of target molecules [69–75]. The SAMPL6 that consists of two parts [76]. The purpose of part I was the blind prediction of pK_a and part II was the prediction of LogP coefficients of drug-like molecules. The participation to the SAMPL7 challenge is based on the estimation of LogP coefficients and pK_a values of small drug-like sulfonamide derivatives.

4.2. LogP Calculations

In physical sciences, the ratio of a neutral solute in a mixture of two immiscible systems at equilibrium is known as the partition coefficient (P) [77, 78]. The hydrophobicity (lipophilicity) and hydrophilicity features of a drug are described as 1-octanol /water partition coefficient. Therefore this physicochemical property is a simple measure that has been used as a lipophilicity indicator for many years [79]. Partition coefficients for drugs are present in the biochemical and pharmaceutical literature in the course of structure-activity studies conducted over the past decade. Many other values have come to light from the references in the paper of Hansch and Leo [77]; these values were fed into a computer-based "keyed-retrieval" database. This pioneering work carried out by Hansch and Leo in 1971 has contributed to the usage of LogP, also refered as log Kow, as a general description of cell permeability in the quantitative structure-activity relationship (QSAR) methods. In pharmaceutical science, both phases usually are solvents, one is water and the other is hydrophobic, such as 1-octanol [80]. Therefore, the LogP coefficient indicates the hydrophobic (water-fearing) or hydrophilic (water-loving) nature of a molecule.

In a rational drug design, bioavailability of a drug and its entrance to the target are very important issues. The drug must pass through a cell membrane barrier in order to show its effects. Depending on the location of the target site and the way of the administration of the drug, the pH of the environment may change; the pH of kidneys is 4.2, the one of plasma is 7.4 that of stomach is 2.0 etc. Therefore, the affinity of a drug molecule and its ability to enter in a lipophilic environment under different pH's has to be considered to predict its effects on the target molecule. A good knowledge of the drug distribution is necessary to have a target-specific drug treatment. The hydrophilic drugs, having low LogP values, are generally localized in the hydrophilic parts of the organism such as blood serum, whereas hydrophobic drugs with high LogP values are distributed to the hydrophobic part of the organism such as lipid bilayers of cells. For the very reason, LogP coefficient is used to predict pharmokinetic properties of a drug accurately.

$$LogP = \frac{(\triangle G_{solv,water} - \triangle G_{solv,octanol})}{RT(ln10)}, \tag{4.1}$$

The absolute Gibbs energies of solvation calculated in diffferent phases is useful to estimate the partition coefficient. Since the LogP coefficient is proportional to the transfer of free energy between two phases (water and n-octanol) it can be related to solvation free energies [81], as shown in Eq. 4.1 where ΔG values are the Gibbs free energies of the solvated molecules in the relevant solvent.

There are a variety of experimental tools to predict LogP partition coefficient in pharmaceutical industry, such as the shake-flask [81], slow stir, reverse phase high performance liquid chromatography [82, 83] and voltammetry. However, these methods are expensive and time consuming , because they all require the actual chemical compounds and their synthesis take a long time. Accordingly, computational methods are very useful to accelerate the prediction process in drug discovery industry.

4.2.1. SAMPL6

In this study, we have considered the SAMPL6 part II challenge including the forecasting of the LogP of 11 protein kinase inhibitor-like small molecules shown in Figure 4.1. In line with this purpose, we have used quantum-chemical calculations to evaluate the LogP values of these 11 molecules.

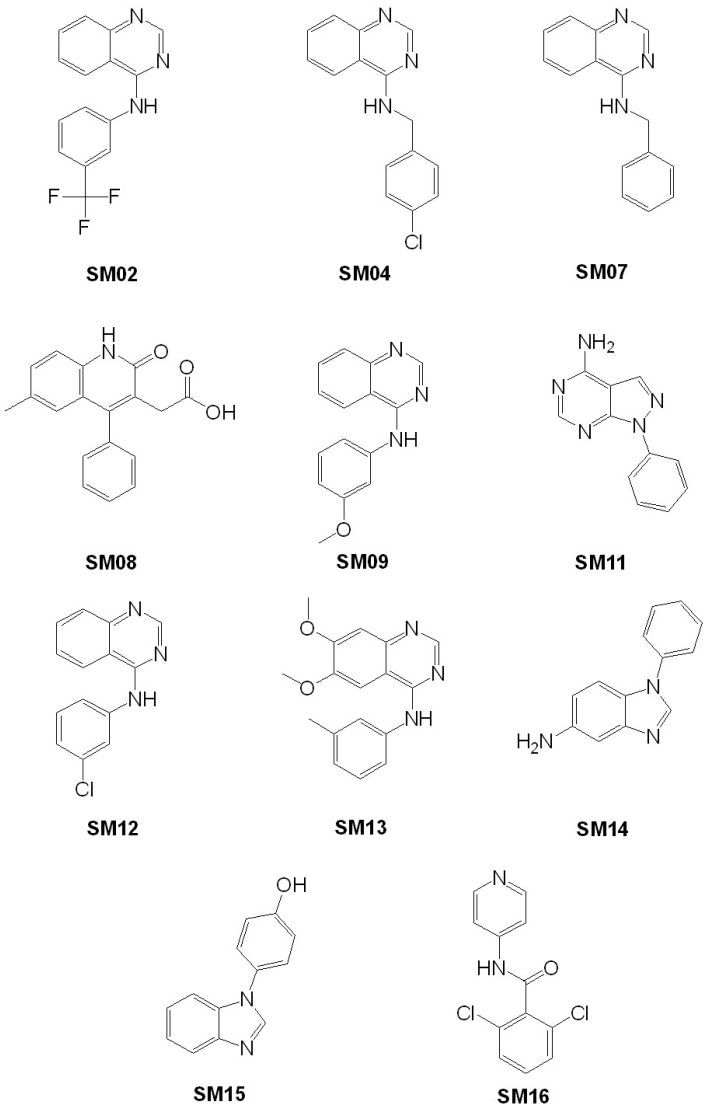


Figure 4.1. The molecules of the SAMPL6 challenge.

4.2.1.1. Methods. All the quantum chemical calculations were performed using the Gaussian 09 software package [84]. Basically, the initial compounds declared for the challenge are modeled to predict the partition coefficient by using quantum mechanical approaches taking into account the solvent environment both implicitly and explicitly. Density functional theory (DFT) is used by employing three hybrid density functionals in conjunction with two basis sets (6-31G(d) [85–88]and 6-311+G(d,p) [89, 90] in order to unravel the most appropriate methodology to reach the optimal estimation. Three functionals were chosen based on the Jacob's ladder of DFT [91], since the energy calculations become more accurate while the functional forms get more complex. For this reason, a hybrid functional (B3LYP [92, 93]), a hybrid meta-GGA (meta-generalized gradient approximation) functional (M06-2X [94]) and a range separated hybrid functional (ωB97x-D [50]) have been chosen. To the best of our knowledge, the previous studies based on the solvation free energy calculations include hybrid DFT functionals however, a specific functional or a class of functionals to calculate accurately the solvation free energy, has not been asserted yet [95–98].

A conformational search around single bonds was performed by using the semi-empirical PM3 method for the structures corresponding to local minima. The geometries corresponding to the local minima were separately optimized in each environment (gas phase, n-octanol and water), so that their effects on geometry were included. The most stable molecular geometry in each environment is relocated and the energetics reported are mainly based on the global minima. The Solvation Model based on the Density (SMD) at 298K [99] was used to take into account the n-octanol and water environments. The $G_{solution}$ and G_{gas} energies are the Gibbs energies calculated for the global minima in the relevant solvents and in gas phase, respectively. The "fine" integration grid, default integration grid of Gaussian 09, was used for all calculations.

LogP values were calculated according to Eq. 4.1, where the partition coefficient was calculated from the difference in the transfer free energies in n-octanol and in water. Therefore, Gsolv,octanol/water was obtained from the standard reaction free energies, by the difference of the Gibbs free energies in gas and in the relevant solvent.

For $\Delta G_{solv,octanol/water}$ calculation, Eq. 4.2 was used to evaluate the ΔG_{solv} for different media.

$$\Delta G_{solv} = G_{soln} - G_{gas}, \tag{4.2}$$

A charge distribution analysis was performed in order to determine the most suitable heteroatom to be attacked by water by using Charge Method 5, an extention of Hirshfeld population analysis [100], this method is chosen because of its less susceptibility to the basis set size. An explicit water molecule was placed in the vicinity of the heteroatoms where it was stabilized via hydrogen bonding. All the negatively charges heteroatoms have been considered for hydrogen bonding and the candidate structures have been re-optimized by using the B3LYP/6-311+G** methodology, in implicit water environment and gas. Structures corresponding to global minima (having positive vibrational frequencies) were used to calculate the solvation free energies in water. Eq. 4.3 and Eq. 4.4 were utilized for evaluating ΔG_{solv},water energies of drug-like molecules, the subscript w indicates the implicit solvation water phase.

$$\triangle G_{solv,water} = G_{(M+water),w} - G_{M,w} - G_{water,w}, \tag{4.3}$$

$$\triangle G_{solv,water} = G_{(M+water),w} - G_{Mexp,w} - G_{Wexp,w}. \tag{4.4}$$

In Eq. 4.3, G"(M+water),w" represents the total Gibbs free energy of the drug-like molecule interacting with one explicit water, GM,w is the Gibbs free energy of the molecule, G_{water} is the Gibbs free energy of one single water molecule.

On the other hand, the Eq. 4.4 calculates the $\Delta G_{solv,\ water}$, by using a different approach. In this equation, $G_{Mexp,w}$ is the Gibbs free energy of the drug-like molecule evaluated by using single point calculations on the drug in the geometry of the optimized complex, $G_{Wexp,w}$ represents the Gibbs free energy of water evaluated by using single point calculations on water in the geometry of the optimized complex; calculations for Eq. 4.3 and Eq. 4.4 are carried out in an implicit water environment.

$$\triangle G_{gas} = G_{(M+water)} - G_{Mexp} - G_{Wexp} \tag{4.5}$$

Similarly, ΔG_{gas} value is evaluated by means of Eq. 4.5, where $G_{M+water}$ is the Gibbs free energy of the drug-like molecule interacting with one explicit water, in gas. G_{Mexp} is the single-point Gibbs free energy of the drug molecule in the geometry of the complex. G_{Wexp} represents the single point calculated Gibbs free energy of the water molecule in the geometry of the complex, calculations for Eq. 4.5 are carried out gas.

4.2.1.2. Results. Our study has been carried out in two parts: the first part includes the estimation of LogP values by taking into account the solvent implicitly. In the second part we report the results with one explicit water molecule. Implicit solvation (sometimes called continuum solvation) is a technique for describing the solvent as a continuous medium, The approach is often used to estimate the free energy of solvent-solute interactions in structural and chemical processes, such as folding or conformation of proteins, DNA, RNA, and polysaccharides, combining biological macromolecules with ligands, or transporting drugs through biological membranes [101]. The molecules of interest in this challenge possess generally a quinazoline structure, the latter is a water-soluble molecule having low LogP value such as 1.00 [102]. However, due to the presence of different side groups the molecules considered have different LogP values (generally higher than 1.00). These molecules have hydrophobic parts such as benzene rings that increase the LogP values.

The SMD solvation model was applied to 11 different molecules with B3LYP, M06-2X and ωB97x-D as functionals and the 6-31G*, 6-311+G** as basis sets in order to test the methodology used for further calculations. Geometry optimizations and systematic conformer searches were carried out for all molecules in the gas phase, in water and in octanol (Figure B1-B3, Table B1-B5). Solvation free energies obtained in water and n-octanol and then LogP values were calculated and reported in Table 4.1. Figure 4.2 displays linear correlations of the experimental and calculated LogP values by using three different functionals with the 6-31G* and the 6-311+G** basis set. According to the R^2 values, the B3LYP /6-311+G** methodology was chosen for further calculations based on the highest R^2 value. Table 4.2 shows the error analysis of the calculated LogP values with six different methodologies. Root mean square errors

(RMSE), mean absolute errors (MAE) and mean signed errors (MSE) were analyzed to choose the methodology for further calculations. The MAE and MSE errors calculated with B3LYP/6-311+G** (RMSE = 1.29, MAE = 1.17, MSE = -1.17) were the smallest among the MAE and MSE calculated with the other methodologies: these findings have confirmed the usage of B3LYP/6-311+G** for further calculations.

Table 4.1. LogP values of molecules of interest with different functionals and basis sets.

		B3LYP		M06-2X		wB97X-D	
Molecules	**6-31G***	**6-311 +G****	**6-31G***	**6-311 +G****	**6-31G***	**6-311 +G****	**Exp. [23]**
SM02	7.86	3.12	2.73	2.19	3.04	2.97	4.09
SM04	7.43	3.25	2.52	2.48	2.44	2.33	3.98
SM07	6.07	2.52	1.92	2.69	2.69	2.60	3.21
SM08	6.03	1.38	1.54	2.14	1.43	1.36	3.10
SM09	7.37	0.69	4.04	0.69	4.44	0.73	3.03
SM11	-7.31	0.45	3.26	0.45	3.24	0.38	2.10
SM12	7.06	2.58	7.09	2.58	7.08	2.32	3.83
SM13	5.12	1.65	6.30	1.65	4.80	1.20	2.92
SM14	5.24	1.76	5.08	1.46	4.97	1.98	1.95
SM15	5.33	2.05	1.70	1.70	1.58	1.73	3.07
SM16	5.39	1.61	0.85	0.97	1.87	1.83	2.62

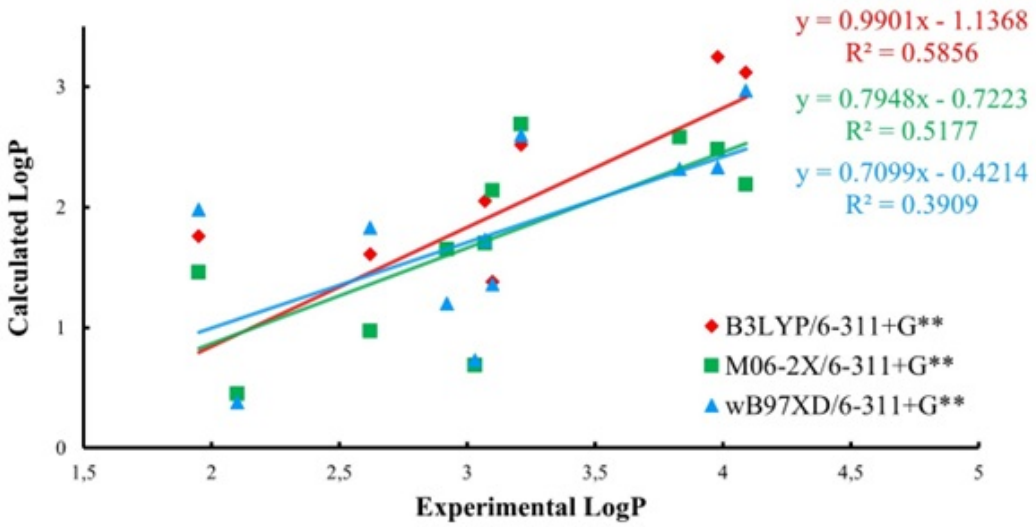


Figure 4.2. Linear correlations of experimental and calculated LogP values by using the B3LYP (red), M06-2X (green) and ωB97x-D (blue) with 6-311+G(d,p) basis set with SMD solvation model.

Table 4.2. Error analysis (Root Mean Square Error (RMSE), Mean Average Error (MAE) and Mean Signed Error (MSE) for the calculated LogP values with six different methodologies.

Method	RMSE	MAE	MSE
B3LYP/6-31G*	4.18	1.66	1.97
M06-2X/6-31G*	1.15	1.89	0.28
ωB97x-D/6-31G*	1.80	1.61	0.33
B3LYP/6-311+G**	1.29	1.17	-1.17
M06-2X/6-311+G**	1.45	1.36	-14.90
ωB97x-D/6-311+G**	1.46	1.34	-14.47

It is well known that hydrogen bond donors are less polar than hydrogen bond acceptors, the identification of the hydrogen bond type can be of interest in determining the affinity of the drug molecule [103]. The most stable molecular geometries of the system including the drug and water molecules are illustrated in Figure 4.3. In most

of the cases, the nitrogen atoms on the bridge between quinazoline and benzene rings are found to be the most negative and H-bond with water is expected to stabilize the molecule while water is in this location. However, the water molecule tends to bind to the nitrogen atom of the quinazoline ring itself probably to avoid steric encumbrance. The water molecule in most cases lies in the same plane as the quinazoline ring, allowing favorable interactions between the oxygen of water and the H on the quinazoline ring as well. Note that the halogen atoms have small negative charges except for SM12,where the charge of chlorine atom is calculated as -0.83 au. However the water doesn't tend to bind to Cl, this may be due to the size of chlorine as opposed to nitrogen.

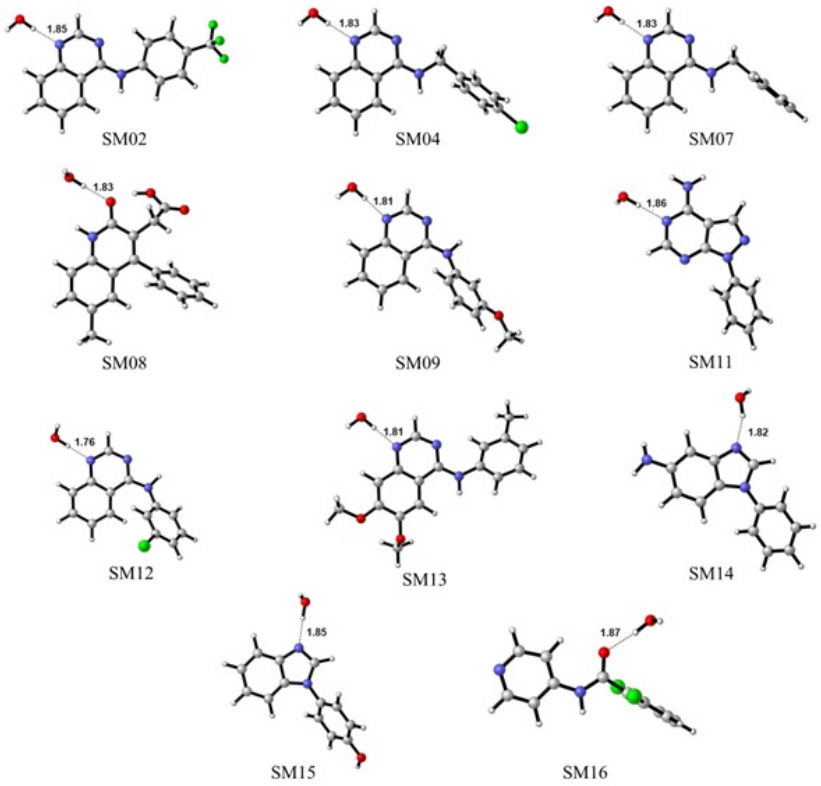


Figure 4.3. Explicitly solvated systems includingone water molecule. (B3LYP/6-311+G**/SMD/water) (Dashed lines represent hydrogen bonding distances in Angstroms.) (Color Key: N:Blue, O:Red, Cl: Green, F:Green, C:Gray).

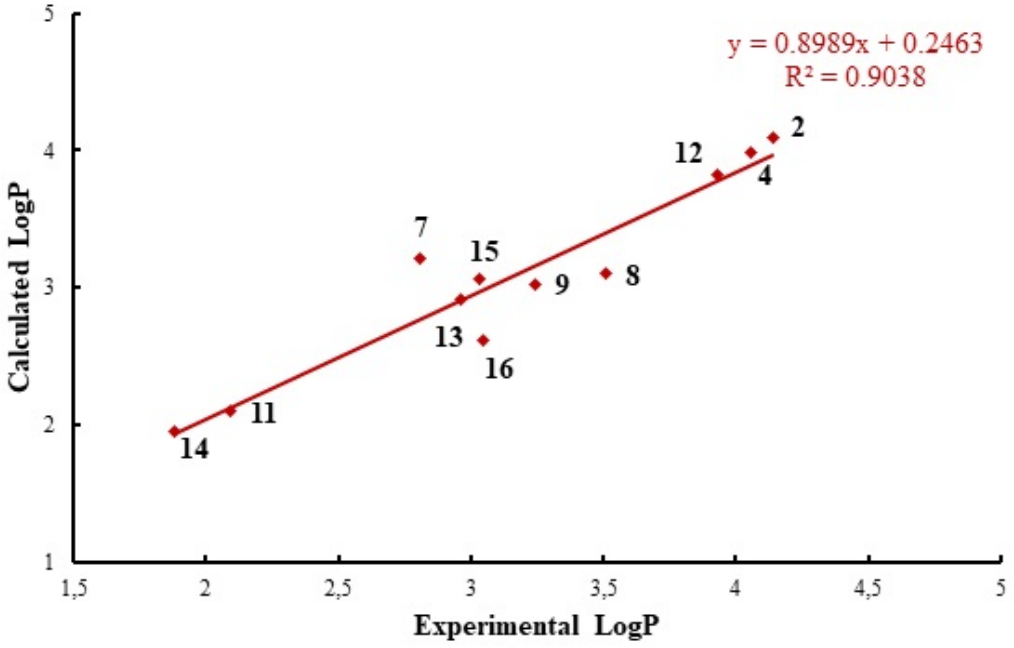


Figure 4.4. Correlation between experimental and calculated LogP values by using B3LYP/6-311+G**/SMD/water.

We have calculated the solvation free energies by substracting Eq. 4.5 from Eq. 4.3, and in Figure 4.4, we have established a correlation between calculated LogP values (R^2= 0.90). However, the ΔG_{solv} values for SM14 and SM15 which possess an imizadole ring were not obtained by using Eq. 4.3, another approach was proposed and their solvation energies were calculated by substracting Eq. 4.5 from Eq. 4.4. Note that, Eq. 4.3 takes into account single point calculated Gibbs free energies of the molecule and the water, while Eq. 4.3 directly uses the Gibbs free energies of the optimized species.

In this study, we were able to predict the exact location of a single water molecule for species having amino quinazoline structures (SM02, SM04, SM07, SM09, SM12 and SM13). Even though for these molecules the charges on nitrogens are more or less similar (Table B1-B5, Figure B1-B3). Water sticks preferentially on the quinazoline ring nitrogen, far away from the bridge. Structures with a water molecule in the vicinity

of the other nitrogens have been considered, it turned out that the ring nitrogen away from the bridge is better stabilized than the others. SM11 is unique in that it has a 1-H-pyrazolo [3,4-d] pyrimidine backbone. The water still interacts with the nitrogen located on the pyrimidine ring away from the substituent. SM08 and SM16 have amidic functional groups, even though the carbonyl oxygen and the amidic nitrogen have comparable charges, the carbonyl oxygen is richer in free electrons and keeps the water molecule. SM14 and SM15 have imidazole rings where water binds preferentially to the available ring nitrogen (sp2) where delocalization of electrons is favored.

Overall, quinazoline having species (SM02, SM 04, SM07, SM09, SM12 and SM13) and SM11 have the water next to the core ring N away from the substituents, in imidazole containing molecules (SM14, SM15) the water is next to the available core ring N, in compounds with amidic groups (SM08, SM16) the water molcule lies next to carbonyl oxygen. The calculated interaction distances are similar, they are around 1.80 Å.

4.2.1.3. Conclusions. This study utilizes quantum chemical methods for the prediction of the octanol/water partition coefficient for a variety of drug-like molecules as a part of SAMPL6 challenge. Gibbs free energies have been calculated implicitly and explicitly with one water molecule. In the first part, three hybrid functionals (B3LYP, M06-2X and ωB97x-D) have been examined with two different basis sets (6-31G* and 6-311+G**) The results have demonstrated that the B3LYP/6-311+G** methodology reproduces better the experimental results. Then, the interaction of the drug molecule with a single water molecule was examined, this methodology -with some exceptions- yields LogP values closer to the experimental ones. This study besides drawing attention to the importance of the inclusion of hydrogen-bonding effects, it correlates nicely experimental and computational LogP values. Drug-like molecules bearing functional groups similar to the ones utilized in this study –quinazoline ring- can be subjected to this relatively simple methodology to have a priori logical LogP values.

4.2.2. SAMPL7

In SAMPL7, there were also pK_a calculations of small drug-like molecules. SAMPL7 includes two parts: LogP calculations and pK_a calculations. LogP calculations were mentioned in SAMPL6 challenge. For pK_a calculations there are many methods that can be used by using quantum chemical tools. The physicochemical features of a drug molecule, such as solubility, partition coefficient, hydrogen bonding ability, degree of ionization, and protein binding affinity, are directly related to its therapeutic action. Among these properties, ionization degree (pK_a) plays a significant role for the design of smart drug delivery systems. High degree of ionization of a drug ligand is a pre-requisite for good water solubility, therefore high hydrophilicity, which is required for drug-receptor interactions. Whereas, the non-ionized form of the ligand improves the lipophilicity of a drug, and therefore, it helps the drug molecule to cross the cell membrane. Thus, the penetration of a drug ligand through the cell membranes is mainly governed by the pK_a of the drug molecule and the membrane environment. On the other hand, since the reactivity and stability of a molecule are highly dependent on its pK_a, the strength of acidity or basicity of a molecule in any solvent determines the mechanisms of chemical reactions involving synthesis, catalysis, oxidation-reduction and decomposition. There are many ways to calculate pK_a; direct method (Eq. 4.6, Eq. 4.7, Eq. 4.8, Eq. 4.9), thermodynamic cycles (TC1(Figure 4.10), TC2 (Figure 4.11), isodesmic calculations (Figure 4.12). In Equation 4.6, the general monoprotic acid (HA) dissociation equilibrium reaction and the deprotonation products (H+, A-) are shown. The pK_a of the given equilibrium is the negative logarithm of the equilibrium constant (K_a), as shown in Eq. 4.7. The standard free energy difference between the reactants and products of the dissociation reaction given in Eq. 4.6 is related to K_a and pK_a as shown in Eq. 4.8 and Eq. 4.9.

$$HA \leftrightarrows H^+ + A^-, \tag{4.6}$$

$$K_a = \frac{[H^+][A^-]}{[HA]}, \tag{4.7}$$

$$\triangle G_{aq} = -RTlnK_a, \tag{4.8}$$

$$pK_a = \frac{\triangle G_{aq}}{RTln10}. \tag{4.9}$$

Accurate and fast pK_a predictions by computational methods are crucial for more efficient drug-design. Quantum chemical approaches have been extensively used for the prediction of pK_a's of the ligands [104–115]. Most of the methods differ in the calculation of the dissociation free energies ($\Delta \mathrm{G_{aq}}$) of a given deprotonation reaction (Eq. 4.8). The most straightforward theoretical approach is based on the direct calculation of the Gibbs free energies of the equilibrium reaction of the acid in solution (Eq. 4.6). This method, so called direct method, faces many errors about determining free energies of these three chemical species in solution. significant errors arising from calculation of the free energy of H^+. $\Delta \mathrm{G_{solv}}(\mathrm{H^+})$ values used in the direct method vary between -265.9 and -270.3 kcal/mol and this brings large uncertainties to the predicted values [116]. In the thermodynamic cycle approaches, TC1 (Figure 4.5) and TC2 (Figure 4.6), the free energy of solvation, desolvation and deprotonation of the acid species HA is considered, where $\Delta \mathrm{G_{aq}}$ values can be calculated by:

$$\triangle G_{aq} = \triangle G_{gas} + \triangle\triangle G_{solv}, \tag{4.10}$$

ΔG_{gas} and ($\Delta\Delta G_{\mathrm{solv}}$) terms are calculated by ab into or Density Functional Theory (DFT) methods by adapting the expressions to the thermodynamic cycle constructed. According to the TC1 (Figure 4.5), the most commonly implemented thermodynamic cycle, these terms are:

$$\triangle G_{gas} = G_{gas}(H^+) + G_{gas} - (A^-)G_{gas}(HA), \tag{4.11}$$

$$\triangle\triangle G_{solv} = (H^+) + \triangle G_{solv}(A^-) - \triangle G_{solv}(HA). \tag{4.12}$$

The large uncertainties coming from the $\Delta \mathrm{G_{solv}H^+}$) values in TC1 can be prevented by substituting the proton H^+ by H_2O/H_3O^+ pair in TC2 (Figure 3.9), in which the ΔG_{gas} and $\Delta\Delta G_{\mathrm{solv}}$ are calculated as given in Eq. 4.13 and Eq. 4.14:

$$\triangle G_{gas} = G_{gas}(H_3O^+) + G_{gas}(A^-) - G_{gas}(HA) - G_{gas}(H_2O), \tag{4.13}$$

$$\triangle\triangle G_{solv} = \triangle G_{solv}(H_3O^+) + \triangle G_{solv}(A^-) - \triangle G_{solv}(HA) - \triangle G_{solv}(H_2O). \tag{4.14}$$

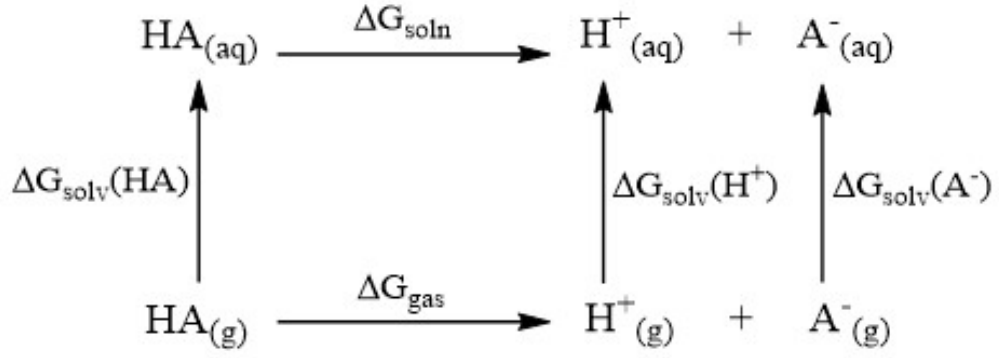


Figure 4.5. Thermodynamic cycle 1 (TC1) where the acid HA is dissociated in its conjugate base A^- and a proton H+. $\Delta G_{soln}$and ΔG_{gas} are the free energies of deprotonation in solution and gas phase, whereas ΔG_{solv} is the free energy of solvation.

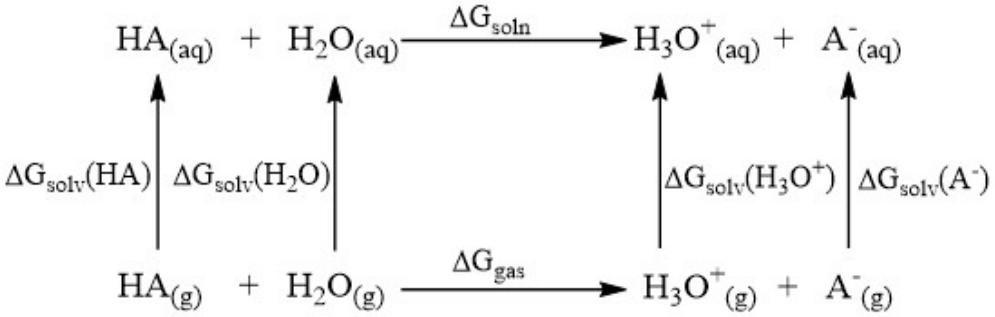


Figure 4.6. Thermodynamic cycle 2 (TC2) where the acid HA donates its proton H^+ to the water molecule to yield its conjugate base A^- and hydronium cation H_3O^+. ΔG_{soln} and ΔG_{gas} are the free energies of deprotonation in solution and gas phase, whereas ΔG_{solv} is the free energy of solvation.

Note that in these reactions the H_2O is the base and signifies a water cluster, whereas $H^+_{(aq)}$ and $H_3O^+_{(aq)}$ represent the hydrated proton product [117]. Even though

the former is claimed to be simpler and truer than the latter, both models are widely used for pK_a calculations. The isodesmic reaction scheme for pK_a calculations, is based on a reaction between an acid (HA) and the conjugate base of a reference molecule (B^-) with an experimentally known pK_a as shown in Figure 3.10, whose equilibrium constant is described in Eq. 4.10. In the isodesmic reactions the types of bonds that are made to form the products are the same as those which are broken in the reactant. The pK_a of acid HA relative to pK_a of reference molecule HB is calculated by Eq. 4.11. Since in the isodesmic method, only the free energies of the reactants and products in solution are calculated, the errors due the gas phase calculations are cancelled and as a result more accurate pK_a's are obtained. The protonation/deprotonation process occurs in similar environments in molecules A and B

$$\begin{aligned} &K_{eq} = \tfrac{[A^-][HB]}{[HB][B^-]}, \\ &pK_a(HA) = \tfrac{\triangle G_{soln}}{2.303RT} + pK_a(HB), \\ &\triangle G_{soln} = G_{soln}(HB) + G_{soln}(A^-) - G_{soln}(HA)G_{soln}(B^-), \\ &HA_{(aq)} + B^-_{aq} \xrightarrow{\triangle G_{soln}} A^-_{aq} + HB_{(aq)}. \end{aligned} \tag{4.15}$$

In SAMPL7, six data set groups have been proposed based on their backbone structures: N-acylsulfonamides in Group I, oxetane, thiethane, S+O- and SO_2derivatives of sulfonamides in Group II and heterocyclic sulfonamides in Groups III and IV as shown in Figure 4.8.

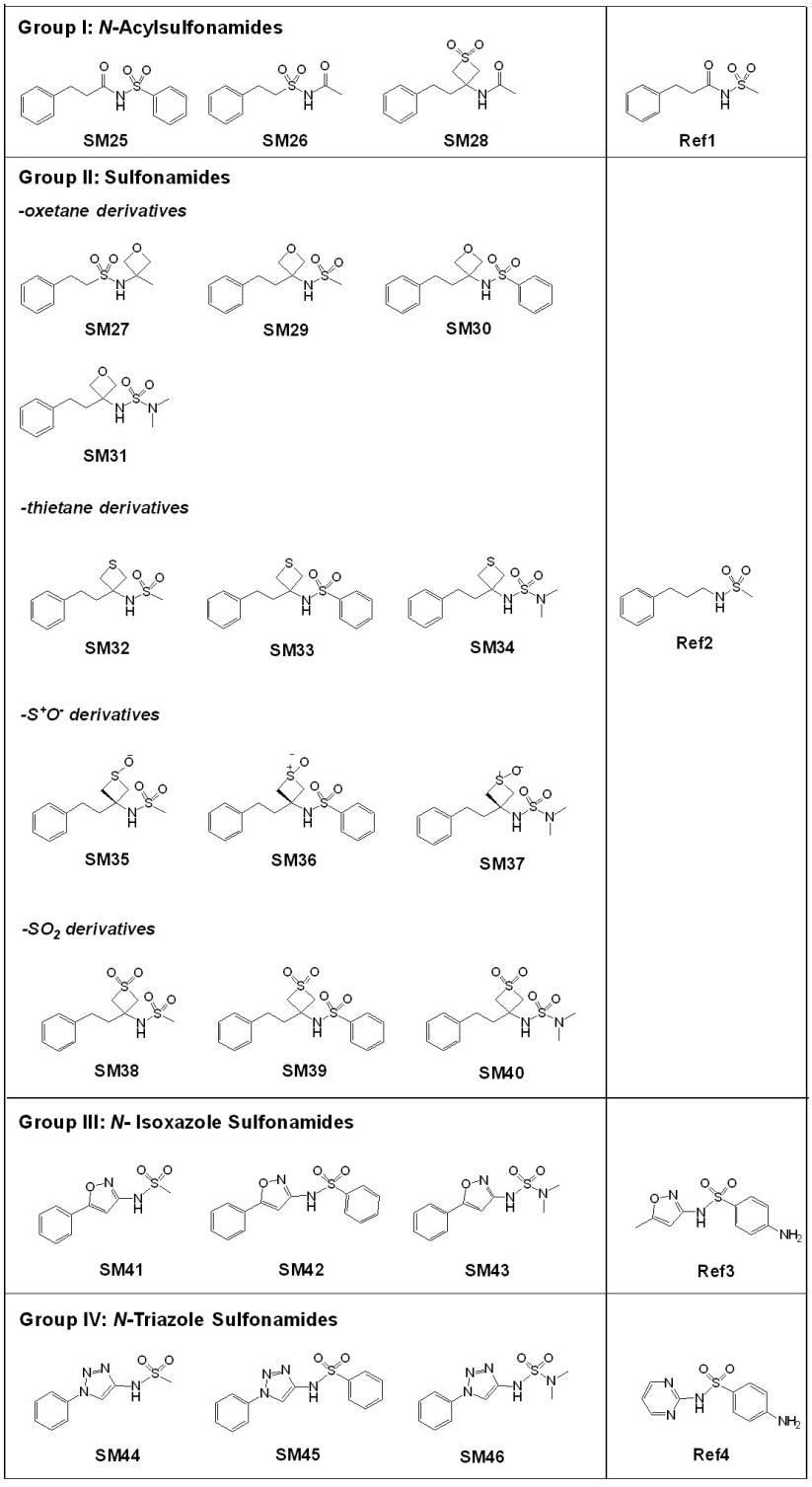


Figure 4.7. 2D images of the molecules and their classifications.

4.2.2.1. Computational Methodology. A systematic conformational search was conducted for 22 molecules using the semi-empirical PM3 method [118] and then the geometries obtained were further optimized with B3LYP-D3/6–311+G(d,p) level of theory [49], [119, 120] in the gas phase using the Gaussian 16 software [53]. The vibrational frequency analysis was implemented to ensure that no imaginary frequencies were present. The gas phase optimized geometries were further optimized at the B3LYP-D3/6-311+G(d,p)[121] level of theory in each solvent environment (n-octanol and water) by employing the SMD solvation model [122]. The lowest 3 energy conformers were re-optimized using the M06 Minnesota family functionals M062X [123] and M06-L [124] in combination with Pople basis sets 6-31G(d) [125, 126] and 6-311+G(d,p) [118] and split valence Karlsruhe basis set def2-SVP [127]. Minnesota functionals were shown to perform very well for the prediction of LogP values of small drug-like molecules [128]. In the previous SAMPL challenge [129] it was reported that the M06-L functional performs well for the lipophilicity estimations, nevertheless the sensitivity of the Minnesota functionals is highly dependent on the data types. Therefore we extended the choices of the functionals and chose pure Minnesota functionals, M11-L [130] and MN15-L [131], due to their relatively high prediction performance in thermochemistry [132]. Energy refinement calculations were done with def2-TZVP [86] basis set by using the corresponding functional, as well.

Unlike our previous study [129], where we used one explicit water molecule in close proximity to the most negative heteroatom, in this study the solvation effects on the geometries have been included with the implicit solute electron density (SMD) solvation model since the molecules in SAMPL7 challenge have more than one heteroatom and inclusion of a finite number of explicit water is not expected to mimic the hydrophilic environment.

4.2.2.2. Results for LogP Calculations. In this part, we performed a systematic investigation of the transfer free energies in n-octanol and water at different level of theories. Also, the statistical error assessment (RMSE, MAD and MD) for the calculated LogP values with respect to the experimental data is collected in Table 4.4. The deviations

allowed us to choose the appropriate functional and the basis set for the given data set. Note that regarding LogP calculations, the relative energetic distribution as well as the 3D geometries of the few best conformations is not the same with all the methodologies employed. Having access to the experimental data enabled us to choose the most adequate methodology based on its small RSME value, in this way the geometries of the best conformers in water and in octanol were identified. The best gas conformes are shown in Figure B.6.

The results show that the geometry optimization of the molecules with the local functional (M06-L) with 0% HF exchange gives better approximations when compared to global hybrid functional (M06-2X) and other local functionals (M11-L and MN15-L), and def2-SVP basis set is found to perform the best based on the low RMSE value along with low computational cost. LogP predictions with M06-L/def2-SVP has a relatively low RMSE (1.70) with reasonable mean deviation value (-1.54), whereas M06-2X/def2-SVP level of theory has RMSE value of 2.09 with mean deviation of -1.71.

Table 4.3. Experimental and Calculated LogP values with different level of theories together with their RMSE, MAD and MD values (A: M06-2X/6-31G(d) B: M06-2X/6-311+G(d,p) C: M06-2X/def2-SVP D: M06-L/6-31G(d) E: M06-L/6-311+G(d,p) F: M06-L/def2-SVP G: M06-L/def2-TZVP//M06-L/def2-SVP H: M11-L/def2-SVP I: M11-L/def2-TZVP//M11-L/def2-SVP J: MN15-L/def2-SVP K: MN15-L/def2-TZVP//MN15-L/def2-SVP).

IDs	Exp	A	B	C	D	E	F	G	H	I	J	K
SM25	2.67	0.55	1.5	-0.29	0.55	1.27	1.03	1.32	0.44	0.79	0.5	0.77
SM26	1.04	-0.76	-1.49	0.02	-0.76	-0.18	-0.52	0.53	-0.26	0.08	-1.4	-1.13
SM27	1.56	0.15	-0.11	0.26	0.15	1.09	0.57	2.28	1.01	1.23	2.26	1.66
SM28	1.18	-1.2	-2.03	-0.98	-1.2	-0.71	-0.45	0.33	-0.52	-0.29	0.02	0.22
SM29	1.61	-0.38	-0.89	-0.3	-0.38	-0.32	-0.19	1.33	0.37	0.95	-0.02	0.36
SM30	2.76	0.79	-0.18	0.87	0.79	1.06	1.12	2.64	1.24	2.28	0.83	1.19
SM31	1.96	-0.34	-0.19	0.13	-0.34	0.17	-0.53	1.6	-0.72	-0.47	0.38	0.91
SM32	2.44	1.45	0.62	1.16	1.45	1.44	1.44	3.8	2.04	2.79	1.77	2.5
SM33	2.96	2.75	2.62	2.43	3.53	2.77	2.14	2.18	3.04	2.65	2.75	2.38
SM34	2.83	1.62	1.9	1.82	1.71	2.03	2.08	2.22	2.34	2.37	1.91	1.92
SM35	0.88	-1.75	-1.98	-1.41	-1.43	-1.52	-1.48	-1.07	-1.55	-1.02	-1.43	-0.97
SM36	0.76	0.74	0	0.41	0.03	-0.18	0.4	2.36	0.94	1.38	1.02	1.38
SM37	1.45	-0.24	-0.67	-0.46	0.12	-0.23	0.52	0.77	0.48	0.79	0.12	0.31
SM38	1.03	-2.7	-2.86	0.85	-1.41	-2.1	-1.54	-0.88	-2.8	-1.87	-1.64	-1.15
SM39	1.89	2.27	-1.27	0.64	0.18	-0.42	-0.35	0.28	-0.89	-0.95	-0.63	-0.13
SM40	1.83	2.29	-0.97	1.67	-0.16	-0.96	-1.35	0.38	-0.41	0.09	-0.63	-0.21
SM41	0.58	-0.93	-0.96	-0.92	-0.6	-0.8	0.37	0.54	-0.36	-0.22	-1.18	-1.17
SM42	1.76	0.84	-0.4	0.78	1.11	0.64	0.94	1.07	1	1.11	1.05	1.14
SM43	0.85	-0.83	-0.63	0.02	-0.49	-0.49	-0.26	-0.27	-0.07	-0.07	-0.14	-0.22
SM44	1.16	-2.36	-2.36	-3.57	-1.78	-1.79	-1.62	-1.18	-1.69	-1.43	-1.94	-1.78
SM45	2.55	-2.28	-0.85	-1.52	-0.66	-0.28	-0.47	0.23	-0.31	-0.1	-0.41	-0.24
SM46	1.72	-2.3	-1.65	-1.97	-0.92	-2.67	-1.31	-0.48	-1.05	-1.22	-1.23	-1.27
RMSE		2.62	2.05	2.09	1.91	1.89	1.70	1.33	1.85	1.59	1.78	1.47
MAD		1.89	1.80	1.70	1.77	1.66	1.54	1.13	1.56	0.95	1.55	1.14
MD		-1.82	-1.80	-1.71	-1.72	-1.59	-1.54	-0.79	-1.03	-0.75	-1.51	-0.68

Furthermore, the performances of the local functionals to predict LogP values were evaluated by using graphical representations for the correlation plots of the calculated and experimental LogP values. As illustrated in Figure 4.9., the dashed lines represent the acceptable deviations (overestimation or underestimation of the computational method) while the solid line represents the 1:1 correlation. In a general scheme, the single point calculations reasonably improve the correlation and the RMSE values become slightly narrow, nevertheless a perfect relationship between experiment and calculations could not be established. After the energy refinement calculations with def2-TZVP basis set, the RMSE values were calculated as 1.33, 1.68 and 1.79 for M06-L, M11-L and MN15-L, respectively.

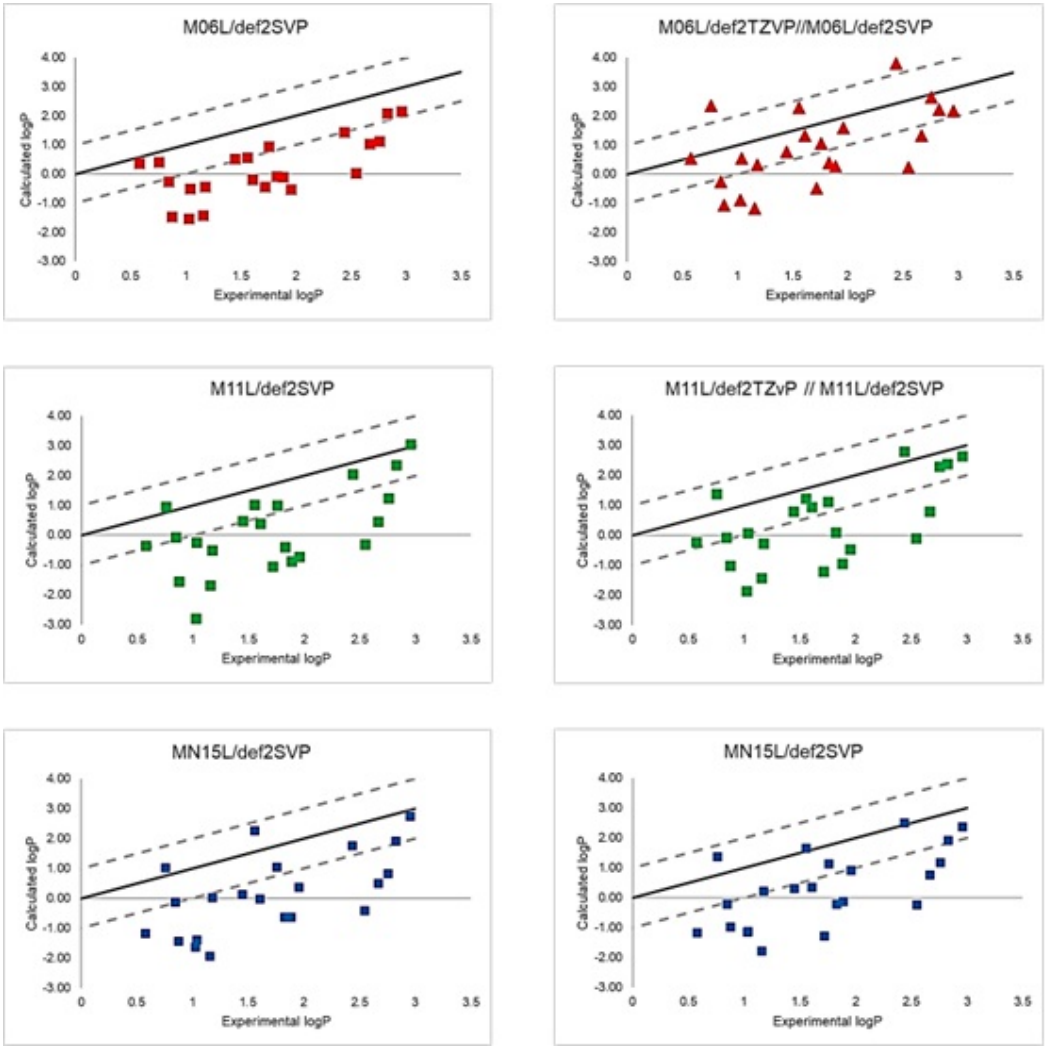


Figure 4.8. Linear correlation plots of pure Minnesota functionals M06-L, M11-L and MN15-L (red, green and blue respectively). The results obtained from geometry optimizations (with def2-SVP basis set) are reported as squares (left), while results obtained from single point energy refinement calculations (with def2-TZVP basis set) are represented as triangles (right).

We grouped the investigated molecules by taking into account their backbone similarities and the results were analyzed based on this classification. Figure 4.10 represents the correlation plot of the experimental and the predicted LogP values calculated at M06-L/def2-TZVP// M06-L/def2-SVP level of theory.

Group I molecules with N-acylsulfonamide functionality are represented as green

diagonals. Our approach underestimates the LogP value of SM25, while SM26 and SM28 are located in the correlation region. Moreover, the methodology performs differently for the molecules in Group II; the ones which contain SO group in their rings. They are represented as the solid blue triangles, and among them SM36 is overestimated, SM35 is underestimated while SM37 is in the acceptable region. The lilac cross corresponds to the molecules which contain the thioethane ring, and two of them (SM33 and SM34) are on the ideal correlation line and the other (SM32) is in the acceptable region, and hence, these molecules are well correlated with the experiment. Compounds with the oxetane rings (red plus), SM27, SM29, SM30 and SM31, are all found in the acceptable region. Group III sulfonamides with the isoxazole rings, namely SM41, SM42 and SM43 (represented with solid yellow spheres), have an acceptable correlation with the experimental data, especially SM41 lies on the ideal line. Conversely, Group IV molecules, SM44, SM45 and SM46, which have the triazole heterocycle, are out of the acceptable area. Nevertheless, the methodology and the classification used in this study suffers from the prediction of the partition coefficients of molecules SM38, SM39 and SM40, having the SO2 group (orange square), and N-triazole sulfonamide backboned heterocyclic rings, SM44, SM45 and SM46. A sound conclusion on this issue would only be possible by extending the data set.

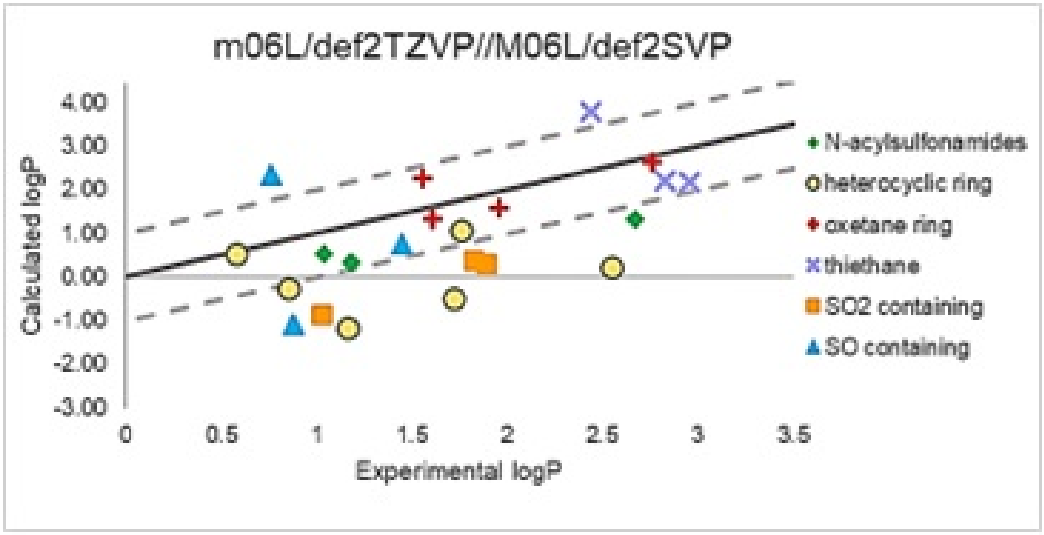


Figure 4.9. Correlation plot of the calculated (M06-L/def2-TZVP//M06-L/def2-SVP) and experimental LogP values of 22 molecules classified as described in the data set section.

4.2.2.3. Conclusion for LogP Calculations.. Eleven different methodologies were used in this study, M06-L and M06-2X functionals with 6-31G(d), 6-311+G(d,p) and def2-SVP, as well as M11-L and MN15-L functionals along with the def2-SVP basis set. Single point calculations with the corresponding functionals and the def2-TZVP basis set were performed on the optimized geometries. The M06-L functional was found to predict LogP values more accurately with relatively low RMSE value and low computational cost for the data set included in this study. The prediction performances of M06-L, M11-L and MN15-L were improved by extending the basis set (def2-TZVP), overall M06-L functional outperformed the others by yielding the lowest RMSE values. However, this method failed in the prediction of the experimental partition coefficients of sulfonamides with SO_2containing rings and N-triazole functionalized heterocyclic rings; a generalized conclusion can only be reached by extending the data set and using other recent DFT functionals as well.

4.2.2.4. Conclusion for pK_a Calculations. For the second part of the study pK_a values of the molecules were calculated. Isodesmic reaction scheme were applied to the 22 drug-like medium sized molecules in order to predict their pK_a values by using the neutral and protonated states of 4 diferent sulfonamide derivatives as reference species (Figure 4.9). Calculated pK_a's, deviations from the experimental pK_a's ($\triangle pK_a$), mean absolute error (MAE), mean deviation (MD) and root mean square error (RMSE) are presented in Table 4.4. MAE and RMSE were calculated to be 1.54 and 1.90 with M052X/6–31+G**, 0.82 and 1.08 with M062X/6–311+G (2df,2p)//M052X/6–31+G(d,p) level pointing out that M062X/6–311+G(2df,2p)//M052X/6–31+G(d,p) performs better than M052X/6–31+G** (SM28 and SM33 are discarded from statistical analysis since their exact experimental pK_a's are not reported). Predicted pK_a's of 15 over 20 molecules are within the range of ± 1 as compared to the experimental value: 7 are within the$\leq$0.5 pK_a confidence interval, 8 are within the 0.5 and $\leq$ 1.0 pK_a confidence and 5 predictions show a deviation of $\succ$ 1.0 pK_a units which indicate that even though the isodesmic method has a strong predictive power for these molecules, Ref2 and Ref4 may not be adequate reference molecules for SM38, SM39 and SM44, SM45, SM46 respectively.

Table 4.4. Calculated absolute macroscopic pK_a and experimental data.

	Experimental	M05-2X/6-31 +G(d,p)		M06-2X/6-311 +G(2df,2p)// M05-2X/6-31 +G(d,p)	
Molecule ID	pK_a	pK_a (calc)	$\lvert\triangle pK_a\rvert$	pK_a (calc)	$\lvert\triangle pK_a\rvert$
SM25	4.49±0.04	5.50	1.01	5.23	0.74
SM26	4.91±0.01	7.15	2.24	5.58	0.67
SM27	10.45±0.01	9.76	0.69	9.93	0.52
SM28	>12.00	14.41		13.72	

Table 4.4. Calculated absolute macroscopic pK_a and experimental data. (cont.)

	Experimental	M05-2X/6-31 +G(d,p)		M06-2X/6-311 +G(2df,2p)// M05-2X/6-31 +G(d,p)	
Molecule ID	pK_a	pK_a (calc)	$\lvert\triangle pK_a\rvert$	pK_a (calc)	$\lvert\triangle pK_a\rvert$
SM29	10.05±0.01	8.59	1.46	9.61	0.44
SM30	10.29±0.12	8.97	1.32	9.33	0.96
SM31	11.02±0.01	12.63	1.61	10.18	0.84
SM32	10.45±0.02	9.59	0.86	10.76	0.31
SM33	$\succ$12.00	8.85		10.82	
SM34	11.93±0.04	10.58	1.35	11.91	0.02
SM35	9.87±0.01	10.09	0.22	10.43	0.56
SM36	9.80±0.06	9.36	0.44	9.79	0.01
SM37	10.33±0.02	9.97	0.36	10.42	0.09
SM38	9.44±0.02	5.71	3.73	8.06	1.38
SM39	10.22±0.15	5.37	4.85	7.42	2.80
SM40	9.58±0.01	7.39	2.19	9.51	0.07
SM41	5.22±0.01	4.83	0.39	4.55	0.67
SM42	6.62±0.02	4.11	2.51	6.03	0.59
SM43	5.62±0.02	4.12	1.50	5.94	0.32
SM44	6.34±0.01	4.79	1.55	4.56	1.78
SM45	5.93±0.05	4.82	1.11	4.32	1.61
SM46	6.42±0.01	7.86	1.44	8.22	1.80
MAE			1.54		0.85
MD			-0.89		-0.36
MAD			1.54		0.81
RMSE			1.90		1.08

Group I compounds SM25 and SM26 possess a 3-phenyl-propanamide or acetamide scafolds attached to sulfonamide groups with phenyl or phenylethyl substituents. SM28 differs from the other compounds in the position of sulfonyl group which is not located on the backbone chain. On the other hand, the reference species Ref1 has a 3-phenylpropanamide group and a methyl group bonded to sulfonamide scafold. The calculated pK_a's of SM25 and SM26 are within ± 1 pK_a units with respect to their experimental pK_a's. We were not able to calculate ΔpK_a for SM28 since its experimental pK_a is reported to be larger than 12.00 and the predicted pK_a is 13.72 which can be accepted within the confidence interval. In Group II, there are 13 structurally resembling molecules which have either an oxetane or a thietane substituted 3-phenylpropyl or ethyl scafold attached to the sulfonamide nitrogen and methyl, phenyl, dimethylamine or 3-phenylpropyl groups bonded to the sulfonyl sulfur. Ref2 has a non-substituted 3-phenylpropyl scafold attached to methanesulfonamide. Among these 13 compounds, only SM38 and SM39's pK_a's are underestimated. These molecules have thietane 1,1-dioxide groups on their phenylpropyl scafolds and methyl and phenyl substituents on the sulfonamide scafolds. In order to check if Ref2 is a good reference species for SM38 and SM39, SM40, which only difers in the sulfonamide substituent from SM38 and SM39, was used as the reference molecule instead of Ref2. The respective pK_a values are calculated as 10.39 and 9.77 with ΔpK_a's of 0.95 and 0.45. Therefore, SM40 is a more adequate reference species for SM38 and SM39. This may be attributed to the fact that different functional groups present on the molecules may be the source of different charge distributions or solute–solvent interactions. Group I compounds SM25 and SM26 possess a 3-phenylpropanamide or acetamide scaffolds attached to sulfonamide groups with phenyl or phenylethyl substituents. SM28 differs from the other compounds in the position of sulfonyl group which is not located on the backbone chain.

On the other hand, the reference species Ref1 has a 3-phenylpropanamide group and a methyl group bonded to sulfonamide scaffold. The calculated pK_a's of SM25 and SM26 are within ± 1 pK_a units with respect to their experimental pK_a's. We were not able to calculate ΔpK_a for SM28 since its experimental pK_a is reported to be larger

than 12.00 and the predicted pK_a is 13.72 which can be accepted within the confidence interval.

In Group II, there are 13 structurally resembling molecules which have either an oxetane or a thietane substituted 3-phenylpropyl or ethyl scaffold attached to the sulfonamide nitrogen and methyl, phenyl, dimethylamine or 3-phenylpropyl groups bonded to the sulfonyl sulfur. Ref2 has a non-substituted 3-phenylpropyl scaffold attached to methanesulfonamide. Among these 13 compounds, only SM38 and SM39's pK_a's are underestimated. These molecules have thietane 1,1-dioxide groups on their phenylpropyl scaffolds and methyl and phenyl substituents on the sulfonamide scaffolds. In order to check if Ref2 is a good reference species for SM38 and SM39, SM40, which only differs in the sulfonamide substituent from SM38 and SM39, was used as the reference molecule instead of Ref2. The respective pK_a values are calculated as 10.39 and 9.77 with ΔpK_a's of 0.95 and 0.45. Therefore, SM40 is a more adequate reference species for SM38 and SM39. This may be attributed to the fact that different functional groups present on the molecules may be the source of different charge distributions or solute-solvent interactions.

Group III compounds have a N-(5-phenylisoxazole) scaffold attached to the sulfonamide group with either a methyl, phenyl or dimethylamine substituent and the reference molecule Ref3 has a N-(5-methylisoxazole) and 4-aminophenyl scaffold attached to the sulfonamide group. Both ethyl and phenyl substituents on the isoxazole ring and methyl, phenyl, dimethylamine, 4-aminophenyl substituents have similar electron donating effects; this makes Ref3 a suitable choice as a reference species for SM41-43 compounds. As there are two protonation sites in Group III compounds, the protonation on the amide nitrogen of the chain resulted in pK_a values of 2.19, 3.04 and 5.45 (data not shown) which correspond to more than 3 units of deviation for SM41 and SM42 whereas the protonation on the N of isoxazole ring resulted in pK_a values of 4.55, 6.03 and 5.94 for SM41, SM42 and SM43 respectively. These predicted pK_a's indicate that the dominant protonation sites in SM41 and SM42 are the N atoms on the isoxazole ring, whereas, for SM43 N atoms on the ring and the chain may contribute

equally to the absolute pK_a of this molecule.

Group IV consists of N-(1-phenyl-1H-1,2,3-triazol-4-yl)-R-sulfonamide derivatives where R is methyl, phenyl and dimethylamine. The reference molecule (Ref4) has a pyrimidine ring and 4-aminophenyl scaffolds attached to the sulfonamide fragment. SM44, SM45 and SM46 have 3 different protonation sites, however, due to the absence of triazole ring in the reference species, we could only protonate the chain amide nitrogen. The calculated pK_a values are 4.56, 4.32 and 8.22 for SM44, SM45 and SM46 respectively and the experimental values are under/over estimated by more than one unit. This could indicate that either the amide nitrogen which lies in the chain is not the primary protonation site or Ref4 is not a suitable reference molecule for this group of compounds. Thus, to clarify this issue, we used SM46 molecule as reference for SM44 and SM45 (Figure 4.11). The protonation on the chain N resulted in pK_a values of 2.84 and 2.46 for SM44 and SM45 respectively, which deviate by 3.50 and 3.47 units from the experimental pK_a values. On the other hand, protonation of the first nitrogen atom on the ring results in pK_a values of 5.14 and 3.60, with ΔpK_a's of 1.20 and 2.33 for SM44 and SM45 respectively, and protonation of the second nitrogen atom results in pK_a values of 7.06 and 6.19, with ΔpK_a's of 0.72 and 0.26 for SM44 and SM45 respectively. These results point out that SM46 is a more suitable reference molecule for the calculation of pK_a's of SM44 and SM45 by the isodesmic reaction scheme, and the primary protonation site is the second nitrogen atom of the triazole ring.

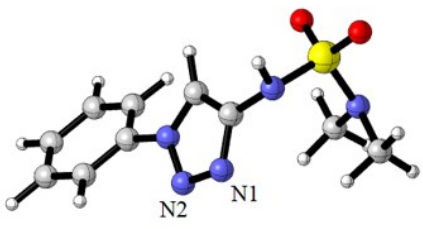


Figure 4.10. 3D representation of SM46 with the N atoms numbered as N1 and N2.

4.2.2.5. Conclusion for pK_a Calculations. Isodesmic reaction scheme was applied for the absolute pK_a calculations of SAMPL7 molecules. We classified the molecules in the data set in 4 groups based on their structural similarities. For each group we determined a reference molecule by taking into account that the reference species and the molecule in question should contain similar functional groups, similar solute-solvent interactions and similar charge distributions.The pK_a's were calculated with the M06-2X/6-311+G(2df,2p)//M05-2X/6-31+G(d,p) level of theory employing the SMD solvation model. Predicted pK_a's of 15 over 20 molecules were in the range of ± 1 unit as compared to the experimental value which points out that the isodesmic method can be applied successfully for the pK_a estimation of N-acylsulfonamide-like compounds. The outliers were further investigated in order to reveal whether the reference species used were suitable for these compounds or not. When we used SM40 as reference for SM38 and SM39 instead of Ref2, more accurate pK_a values were obtained. This may be attributed to the different charge distributions or solute-solvent interactions due to the different functional groups present on the molecules. On the other hand, SM44, SM45 and SM46 have 2 protonation sites on the triazole ring and 1 protonation site on the chain, while Ref4 contains the pyrimidine ring. Due to the absence of triazole ring in the reference species, only chain amide nitrogen was protonated in these compounds. The large deviations observed in these molecules were attributed to the structural differences between the molecule in question and the reference species, since when SM46 was used as reference, the primary protonation site was found to be the second nitrogen atom of the triazole ring.

5. QUANTUM-MECHANICAL PREDICTION OF DISSOCIATION CONSTANTS FOR THIAZOL-2-IMINE DERIVATIVES

5.1. Introduction

The physicochemical features of a drug molecule, such as solubility, partition coefficient, hydrogen bonding ability, degree of ionization, and protein binding affinity, are directly related to its therapeutic action. Among these properties, ionization degree (pK_a) plays a significant role for the design of smart drug delivery systems. High degree of ionization of a drug ligand is a pre-requisite for good water solubility, therefore high hydrophilicity, which is required for drug-receptor interactions. Whereas, the non-ionized form of the ligand improves the lipophilicity of a drug, and therefore, it helps the drug molecule to cross the cell membrane. Thus, the penetration of a drug ligand through the cell membranes is mainly governed by the pK_a of the drug molecule and the membrane environment. On the other hand, since the reactivity and stability of a molecule are highly dependent on its pK_a, the strength of acidity or basicity of a molecule in any solvent determines the mechanisms of chemical reactions involving synthesis, catalysis, oxidation-reduction and decomposition. pK_a calculations are mentioned in Section 4.2.2. In this study TC1, TC2 and isodesmic calculations were used.

With significant biological activities, thiazol-2-imine derivatives are commonly used in organic, synthetic, pharmaceutical and medicinal chemistry areas due to their anti-inflamatory, analgesic, antibacterial [133], antifungal [134], antiviral [135, 136] as well as kinase inhibition activities [137]. These molecules possess chiral centers, thus, synthesis of thiazol-2-imine derivatives as single enantiomers is of great interest. Moreover, these thiazol-2-imine derivatives that are present as single enantiomers can be used in asymmetric chemistry such as catalyst in converting olefins to primary amines [138], asymmetric hydrogenations [138] and generating chiral aminoni-

triles.[139] Recently, Tuncel and Dogan have synthesized single enantiomer thiazol-2-imines (5RR, 6RR, 7RR, 8RR) by water elimination reactions from the corresponding 2-iminothiazolidine-4-ones (1RR, 2RR, 3RR, 4RR) (Figure 5.1) [140]. The synthesized molecules are insoluble in water, their pK_a's can be determined in acetonitrile. Due to its relatively lower polarity compared to water, acetonitrile is a widely used solvent in pharmaceutical industry since it can mimic the membrane environment. As an aprotic solvent with pK_a of 25, MeCN is a poor H-bond acceptor and a weak base, thus, when it is used as a solvent, it increases the pK_a values of the solutes. Compounds 1-8RR, in addition to being drug-like molecules, are expected to be strong chiral organic bases as well because of the amidine conjugation in their structures (Figure 5.2). Since they are anticipated to be stronger bases than their amine precursors, their pK_a values are worth determining in different media.

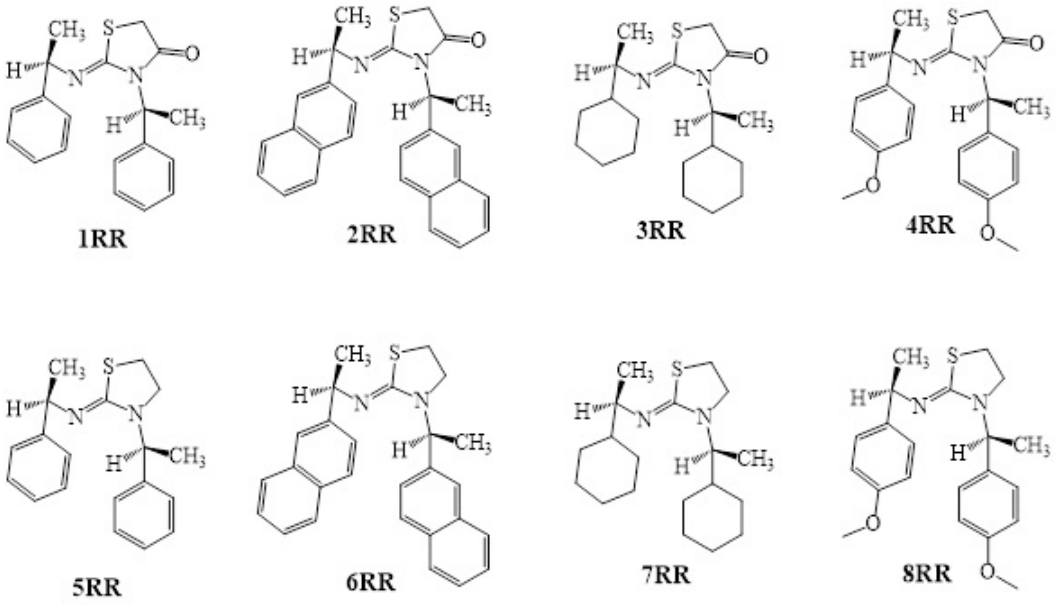


Figure 5.1. 2D representations of the single enantiomer of thiazol-2-imines synthesized by Tuncel and Dogan [140].

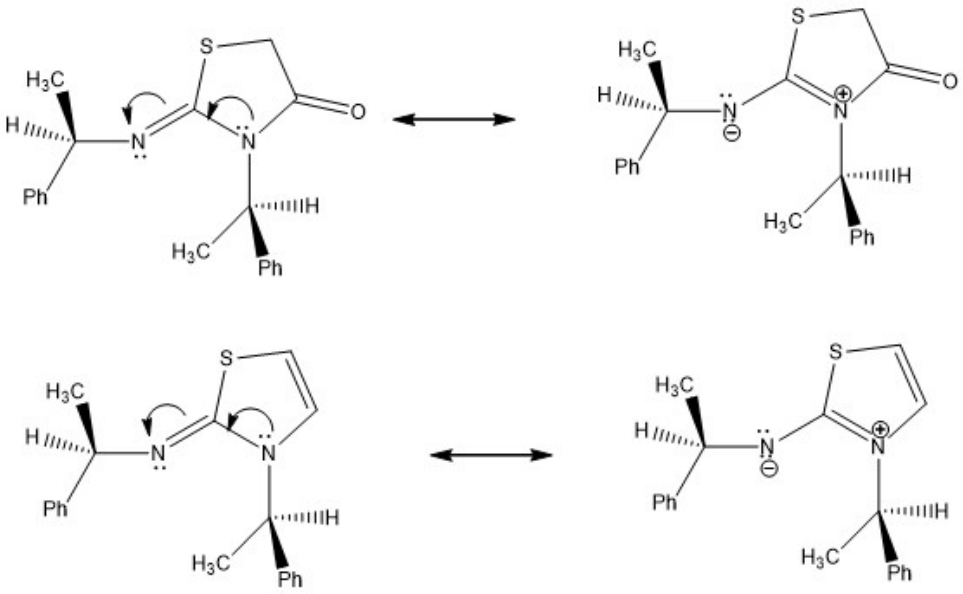


Figure 5.2. The amidine conjugation in thiazol-2-imine compounds.

In this study, the acetonitrile and water pK_a's of thiazol-2-imine derivatives will be predicted by employing the isodesmic reaction scheme, based on the previously reported success of the method [141–146]. With this purpose, by using several quantum mechanical approaches, we propose a very accurate protocol. In the first part, the methodology to be employed throughout the study is identified from a benchmark study, where several DFT functionals and solvent models are examined together with TC1 and TC2 schemes, for reproducing the experimentally known water pK_a of the reference molecule to be used in the isodesmic reaction. In the second and third parts, the proposed methodology is validated on two sets of test molecules whose water and acetonitrile pK_a's are experimentally known. In the fourth part, $pK_{a(\mathrm{MeCN})}$ of the reference molecule to be used in isodesmic reaction is calculated by the established methodology and finally, $pK_{a(\mathrm{MeCN})}$'s of thiazol-2-imines are calculated by using the isodesmic reaction scheme. The protocol proposed in this study can be used with confidence to predict the water and acetonitrile pK_a's of thiazol-2-imine derivatives.

5.2. Computational Methodology

A systematic conformational search was conducted for all the molecules employing the semi-emprical PM3 method [147] by using the SPARTAN14 software [148]. Free rotations around single bonds were taken into account and all the geometries corresponding to stationary points have been re-optimized by using the Gaussian16 program suite [58]. For the benchmarking study of the DFT method, combinations of a hybrid-GGA exchange-correlation functional (B3LYP[149]), a hybrid meta-GGA functional (M062X[150]) and a hybrid-GGA exchange correlation functional ωB97x-D[151] with 6-31+G*, 6-31G**, 6-31+G**, 6-31++G**, and 6-311++G** basis sets were tested. The universal solvent model (SMD)[152] and the conductor-like polarizable continuum model (CPCM) [153] were employed to mimic the aqueous solvent environment. The partial atomic charges were derived from Natural Population Analysis (NPA) [154], Charge Model 5 (CM5) [155] and Hirshfeld Analysis [156]. RMSE (root mean square error), MAD (mean absolute deviation), and MD (mean deviation) were calculated to compare the efficiency of our predictions.

5.3. Results And Discussion

Since thiazol-2-imines synthesized by Tuncel and Dogan in Figure 5.1 are water insoluble molecules, their pK_a's are predicted in acetonitrile environment, by employing the isodesmic reaction scheme. With this respect, several quantum chemical approaches were utilized. In the first place, in order to propose a methodology that describes the system best, the experimentally known water pK_a of the reference molecule to be used in isodesmic reaction was estimated employing the TC1 and TC2 schemes with a set of different DFT functionals and solvent models. In the second part of the study, the proposed methodology was validated on a set of 2-(phenylimino)imidazolidine derivatives and then in the third part the verification of the methodology in acetonitrile environment was performed for a set of nitrogen-containing heterocyclic compounds. The successful applications of the proposed methodology allowed us to predict the $pK_{a(\mathrm{MeCN})}$ of the reference molecule, 2-imino-thiazolidinone. And finally, with the pre-

dicted acetonitrile pK_a of the reference molecule, the $pK_{a(\mathrm{MeCN})}$'s of thiazol-2-imines in Figure 5.1 were calculated by employing the isodesmic reaction scheme.

5.3.1. Identification of the Methodology

As the reference molecule to be used in isodesmic reaction scheme for the prediction of $pK_{a(\mathrm{MeCN})}$'s of thiazol-2-imines, structurally resembling 2-imino-thiazolidinone (Figure 5.3) with the experimental water pK_a of 11.70 [157] is used. Since the experimental $pK_{a(\mathrm{MeCN})}$ of the reference molecule is not known, a 3-step protocol is followed to predict its $pK_{a(\mathrm{MeCN})}$. As well as the calculation scheme to be applied, the accuracy of the prediction is directly related to the quality of the QM method and the solvation model employed to describe the system's electronic nature and solute-solvent interactions. Therefore, the computational methodology to be used is identified by employing TC1 and TC2 schemes with a set of different DFT functionals, basis sets and solvation models for the estimation of the water pK_a of the reference molecule. The reference molecule has two conformers: hydrogen on 2-imine N being either up (towards S) or down (towards N3). 'Hydrogen up' conformation, which was calculated to be 0.76 kcal/mol more stable than the 'down' conformation, is further considered in our calculations.

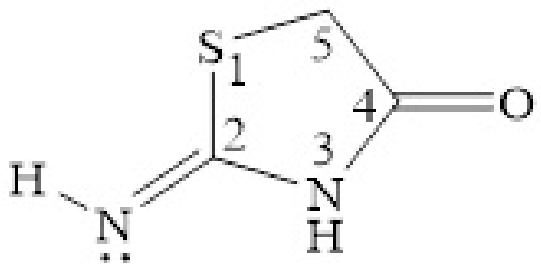


Figure 5.3. 2D representation of the reference molecule of 2-imino-thiazolidinone.

As there are two protonation sites in the reference molecule (N3 and 2-imino N), NPA, CM5 and Hirschfeld charge analysis were performed for the identification of the dominant protonation/deprotonation site. The computed charges with all methodolo-

gies were found to give the same trend with 2-imino N having a higher electron density $q_{N(NPA)}$ = -0.628 e compared to N3 $q_{(NPA),N3}$ = -0.521 e. Thus, the reported partial atomic charges of neutral and protonated (on 2-imino N) 2-imino-thiazolidinone are based on NPA calculations in the rest of the manuscript. First, TC1 scheme was employed with several levels of theory in order to predict the water pK_a of the reference molecule. Two $\Delta G_{solv(}H_+)$ values (-265.9 kcal/mol and -270.3 kcal/mol) were used for each trial, and the results presented in Table 5.1 show that TC1 scheme predicts much lower pK_a's than that of the experimental value which indicates the presence of a systematic error. TC1 seems to fail in the prediction of 2-imino-thiazolidinone pK_a with a very high ΔK_a (the difference between the predicted and experimental pK_a) range ($2.56 \leq |\Delta pK_a| \leq 10.43$) irrespective of the level of theory used. When $\Delta G_{solv(}H_+)$ is taken as -270.3 kcal/mol, the calculated 2-imino-thiazolidinone pK_a values (Calculated pK_{a2}) with respect to different levels of theory are observed to deviate more from the experimental value compared to calculations performed with $\triangle G_{solv}(H^+)$ = -265.9 kcal/mol (Calculated pK_{a1}).

Table 5.1. Benchmark study for the prediction of water pK_a of the reference molecule (pK_a(exp) = 11.70) by employing TC1 scheme.

Level of Theory	**Calculated** pK_a^1	$\triangle pK_a^1$	**Calculated** pK_a^1	$\triangle pK_a^2$
B3LYP/6-31+G* CPCM	2.13	-9.57	0.27	-10.43
B3LYP/6-31+G CPCM**	2.45	-9.25	0.59	-10.11
B3LYP/6-31G CPCM**	7.68	-4.02	5.82	-5.88
B3LYP/6-31++G CPCM**	3.65	-8.05	1.79	-9.91

Table 5.1. Benchmark study for the prediction of water pK_a. (cont.)

Level of Theory	**Calculated** pK_a^1	$\triangle pK_a^1$	**Calculated** pK_a^1	$\triangle pK_a^2$
B3LYP/6-31++G//B3LYP/6-31+G* CPCM**	3.63	-8.07	1.77	-9.93
B3LYP/6-31++G//B3LYP6-31G** CPCM**	3.62	-8.08	1.76	-9.94
B3LYP/6-31+G* SMD	6.55	-5.15	4.68	-7.02
B3LYP/6-31+G SMD**	6.60	-5.10	4.74	-6.96
B3LYP/6-31G SMD**	5.57	-5.13	3.83	-7.87
B3LYP/6-31++G SMD**	6.67	-5.03	4.40	-7.30
B3LYP/6-31++G//B3LYP/6-31+G* SMD**	6.02	-5.68	3.90	-7.80
B3LYP/6-31++G//B3LYP6-31G** SMD**	6.14	-5.56	4.48	-7.22
wB97XD/6-31+G* CPCM	3.75	-7.95	1.89	-9.81
wB97XD/6-31+G CPCM**	5.62	-6.08	4.70	-7.00

Table 5.1. Benchmark study for the prediction of water pK_a. (cont.)

Level of Theory	**Calculated** pK_a^1	$\triangle pK_a^1$	**Calculated** pK_a^1	$\triangle pK_a^2$
wB97XD/6-31G CPCM**	8.49	-3.21	6.63	-5.07
wB97XD/6-31++G CPCM**	5.15	-6.55	3.29	-8.41
wB97XD/6-31++G//wB97XD/6-31+G* CPCM**	5.14	-6.56	3.28	-8.42
wB97XD/6-31++G//wB97XD/6-31G** CPCM**	5.15	-6.55	2.29	-9.41
wB97XD/6-31+G* SMD	7.59	-4.11	6.02	-5.68
wB97XD/6-31+G SMD**	8.00	-3.70	6.14	-5.56
wB97XD/6-31G SMD**	7.43	-4.27	5.49	-6.21
wB97XD/6-31++G SMD**	8.45	-3.25	6.72	-4.98
wB97XD/6-31++G//wB97XD/6-31+G* SMD**	6.43	-5.27	4.58	-7.12

Table 5.1. Benchmark study for the prediction of water pK_a. (cont.)

Level of Theory	**Calculated** pK_a^1	$\triangle pK_a^1$	**Calculated** pK_a^1	$\triangle pK_a^2$
wB97XD/6-31++G//wB97XD/6-31G** SMD**	7.00	-4.70	5.92	-5.78
M062X/6-31+G* CPCM	5.55	-6.15	4.68	-7.02
M062X/6-31+G CPCM**	5.69	-6.01	4.23	-7.47
M062X/6-31G CPCM**	8.01	-3.69	6.75	-4.95
M062X/6-31++G CPCM**	5.59	-6.11	4.02	-7.68
M062X/6-31++G//M062X/6-31+G* CPCM**	5.52	-6.18	3.69	-8.01
M062X/6-31++G//M062X/6-31G** CPCM**	5.01	-6.69	3.53	-8.17
M062X/6-31+G* SMD	4.59	-7.11	2.73	-8.97
M062X/6-31+G SMD**	4.70	-7.00	2.84	-8.86
M062X/6-31G SMD**	9.14	-2.56	7.28	-4.42
M062X/6-31++G SMD**	6.04	-5.66	4.18	-7.52

Table 5.1. Benchmark study for the prediction of water pK_a. (cont.)

Level of Theory	**Calculated** pK_a^1	$\triangle pK_a^1$	**Calculated** pK_a^1	$\triangle pK_a^2$
M062X/6-31++G//M062X/6-31+G* SMD**	6.01	-5.69	4.15	-7.55
M062X/6-31++G//M062X/6-31G** SMD**	5.95	-5.75	4.09	-7.61
Calculated $pK_a^1 = \triangle G_{solv}(H+)$ value is taken as -265.9 kcal/mol				
Calculated $pK_a^2 = \triangle G_{solv}(H+)$ value is taken as -270.3 kcal/mol				

As an attempt to prevent the large uncertainties arising from the $\Delta G_{solv}(H^+)$ values in TC1, the proton H^+ was substituted by H_2O/H_3O^+ pair in TC2, and the calculated water pK_a values of 2-imino-thiazolidinone with different levels of theory are presented in Table 5.2. The differences between the predicted and experimental pK_a (ΔpK_a) have been calculated for each level of theory and represented in Figure 5.4. In TC2 scheme, lower $\triangle pK_a$ range ($0.77 \leq |\Delta pK_a| \leq 6.06$) is observed compared to TC1 scheme. As seen from Figure 5.4, with the CPCM solvation model, the M062X functional behaves better than the others except for the 6-31G** basis set; on the other hand with the SMD solvation model the 6-31G** basis set where polarization function yields the smallest ΔpK_a value.

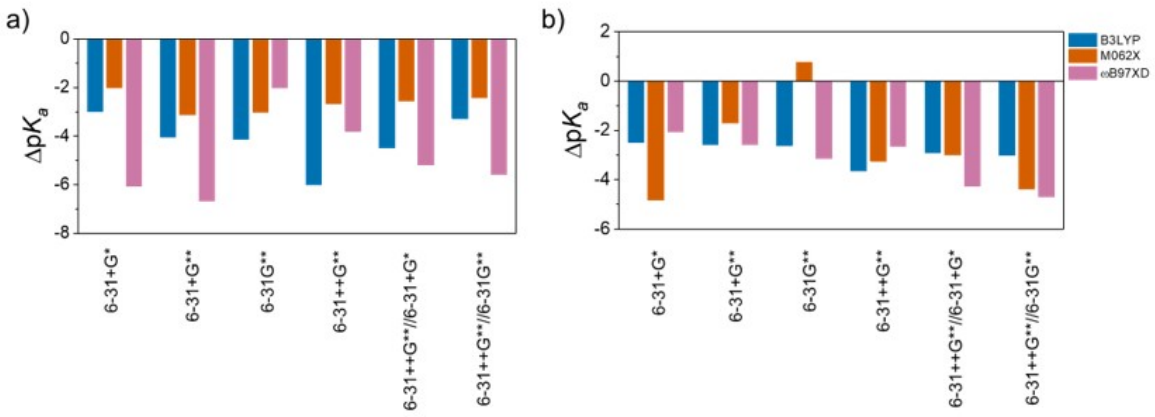


Figure 5.4. Differences between the predicted and experimental $pK_a(\Delta pK_a)$ for three different DFT functionals and six different basis sets considered in this study. Geometry optimizations were performed using the (a) CPCM model and (b) SMD model.

when the M062X/6-31G** level of theory is used in the presence of SMD solvation model, the calculated pK_a of 2-imino-thiazolidinone by employing TC2 scheme is 12.47, for which the $|\Delta pK_a|$ is 0.77 unit. Therefore, in order to predict the $pK_{a,\,(MeCN)}$ of the reference molecule, further verification of the established methodology is performed on a set of test molecules.

Table 5.2. Benchmark study for the prediction of water pK_a of the reference molecule ($pK_a(\text{exp}) = 11.70$) by employing TC2 scheme.

Level of Theory	**Calculated** pK_a	ΔpK_a
B3LYP/6-31+G* **CPCM**	8.17	-3.00
B3LYP/6-31+G** **CPCM**	7.65	-4.05
B3LYP/6-31G** **CPCM**	7.55	-4.15

Table 5.2. Benchmark study for the prediction of water pK_a. (cont.)

Level of Theory	**Calculated** pK_a	ΔpK_a
B3LYP/6-31++G** **CPCM**	5.69	-6.01
B3LYP/6-31++G//B3LYP/6-31+G*** **CPCM**	7.20	-4.50
B3LYP/6-31++G//B3LYP/6-31G**** **CPCM**	8.42	-3.28
B3LYP/6-31+G* **SMD**	9.21	-2.49
B3LYP/6-31+G** **SMD**	9.12	-2.58
B3LYP/6-31G** **SMD**	9.08	-2.62
B3LYP/6-31++G** **SMD**	8.05	-3.65
B3LYP/6-31++G//B3LYP/6-31+G*** **SMD**	7.80	-2.90
B3LYP/6-31++G//B3LYP/6-31G**** **SMD**	8.69	-3.01
wB97XD/6-31+G* **CPCM**	5.64	-6.06
wB97XD/6-31+G** **CPCM**	5.01	-6.69
wB97XD/6-31G** **CPCM**	9.68	-2.02
wB97XD/6-31++G** **CPCM**	7.90	-3.80
wB97XD/6-31++G//wB97XD/6-31+G*** **CPCM**	6.50	-5.20

Table 5.2. Benchmark study for the prediction of water pK_a. (cont.)

Level of Theory	**Calculated** pK_a	ΔpK_a
wB97XD/6-31+G** **SMD**	9.12	-2.58
wB97XD/6-31++G//wB97XD/6-31G**** **CPCM**	6.12	-5.58
wB97XD/6-31+G* **SMD**	9.64	-2.06
wB97XD/6-31G** **SMD**	8.55	-3.15
wB97XD/6-31++G** **SMD**	9.05	-2.65
wB97XD/6-31++G//wB97XD/6-31+G*** **SMD**	7.43	-4.27
wB97XD/6-31++G//wB97XD/6-31G**** **SMD**	7.01	-4.69
M062X/6-31+G* **CPCM**	9.68	-2.02
M062X/6-31+G** **CPCM**	8.56	-3.14
M062X/6-31G** **CPCM**	7.98	-3.02
M062X/6-31++G** **CPCM**	9.02	-2.68
M062X/6-31++G//M062X/6-31+G*** **CPCM**	9.13	-2.57
M062X/6-31++G//M062X/6-31G**** **CPCM**	9.27	-2.43
M062X/6-31+G* **SMD**	6.87	-4.83

Table 5.2. Benchmark study for the prediction of water pK_a. (cont.)

Level of Theory	**Calculated** pK_a	ΔpK_a
M062X/6-31+G** **SMD**	10.00	-1.70
M062X/6-31G** **SMD**	12.47	0.77
M062X/6-31++G** **SMD**	8.45	-3.25
M062X/6-31++G//M062X/6-31+G*** **SMD**	8.70	-3.00
M062X/6-31++G//M062X/6-31G**** **SMD**	7.32	-4.38

5.3.2. Validation of the Methodology for Water Environment

For the validation of the methodology established, a set of molecules with experimentally known water pK_a were collected from the literature. The selected ten 2-(phenylimino)-imidazolidine derivatives share common scaffolds with our target molecules, thiazol-2-imines, as displayed in Figure 5.5 and Figure C.1, and the pK_a's of these molecules vary between 9.08-10.78 [158]. The methodology proposed in the previous section is applied to the molecules in Figure 5.5 in order to predict their pK_a's in water. Initially, NPA charges on the N atoms were calculated for all the structures and the most electron rich N atom was selected as the first protonation site. For each molecule, the electron density around imine N was observed to be the highest and that their pK_a's are directly related to the electron density around N atom. For example, ph6 has the highest pK_a value (10.78) and the computed NPA charge on imine N for ph6 is the lowest (q_N = -0.781 e), whereas ph10 has the lowest pK_a (9.08) with the highest

NPA charge (q_N = -0.362 e) among the ph molecules. The presence of electron withdrawing groups such as carbonyl moities in the structures of ph8, ph9 and ph10 make the imine N more electron deficient compared to ph1-7, and as a result, less basic. To estimate the water pK_a's of the 10 molecules displayed in Figure 5.5, the TC2 scheme with M062X/6-31G** level of theory in conjunction with the SMD solvation model using water as solvent was applied by protonating the imine N.

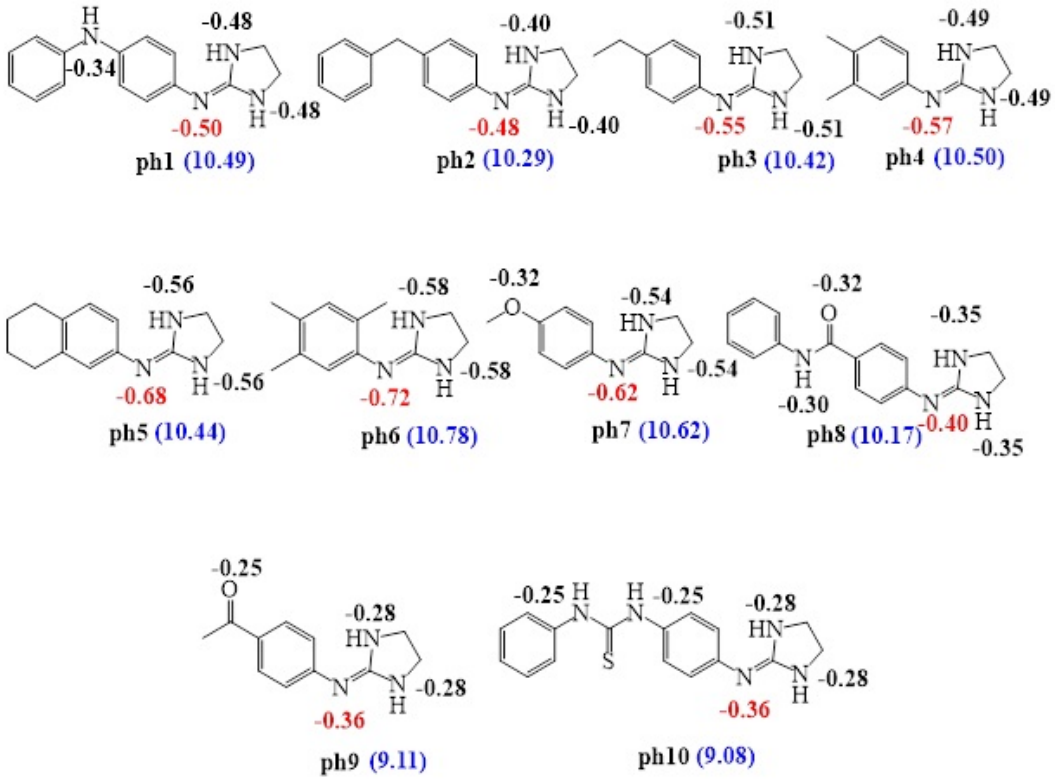


Figure 5.5. 2D representations of 2-(phenylimino)imidazolidine derivatives with computed NPA charges on the heteroatoms (the most negative one is colored with red). (M062X/6-31G**/SMD/water, pop=npa) (Experimental pK_a's are given in blue).

The results presented in Table C1 show that the predicted pK_a's of 2-(phenylimino) imidazolidine derivatives are very close to their experimentally determined values with a strong correlation (R^2=0.96, Figure 5.6). This impacts the success of the proposed methodology with high accuracy of MAD and $\triangle pK_a$ values (MAD = 0.12 and 0.05 $\leq$ $|\Delta pK_a|$ $\leq$ 0.21).

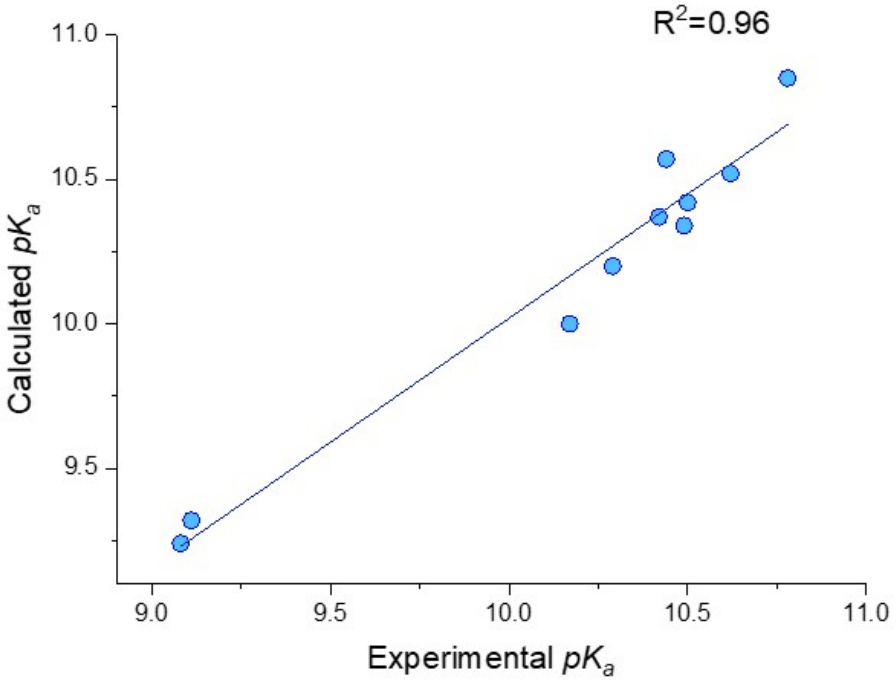


Figure 5.6. Linear regression of experimental vs calculated pK_a values of 2-(phenylimino)- imidazolidine derivatives (M062X/ 6-31G**/ SMD/ water).

5.3.3. Validation of the Methodology for Acetonitrile Environment

Following the validation of the proposed methodology in water environment, the applicability of the methodology was tested for a set of molecules for which there are experimentally determined $pK_{a(MeCN)}$'s. The selected 17 molecules are nitrogen-containing heterocycles as displayed in Figure 5.7 and Figure C.2. Since some of the molecules possess more than one N atom in their structures, calculated NPA atomic charges on the N atoms allowed us to determine the most electron rich N to identify the first protonation site. For the molecules which possess a $-NH_2$ functional group (3, 5, 6, 7, 8, 9,11, 12, 14) the highest electron density is observed to be located on the N atom of the amino substituent. Then, imine N and N on $-N(CH3)_2$ substituent were calculated to have more electron density in molecules 2 and 4, respectively. The

lack of amidine conjugation in 10, 13, 15, 16 and 17 make these molecules less basic compared to 1, in consistency with the calculated partial charges on N atoms. TC2 scheme was employed with M062X/6-31G** level of theory in conjunction with the SMD solvation model using acetonitrile as solvent for the prediction of $pK_{a(MeCN)}$'s of nitrogen-containing heterocycles.

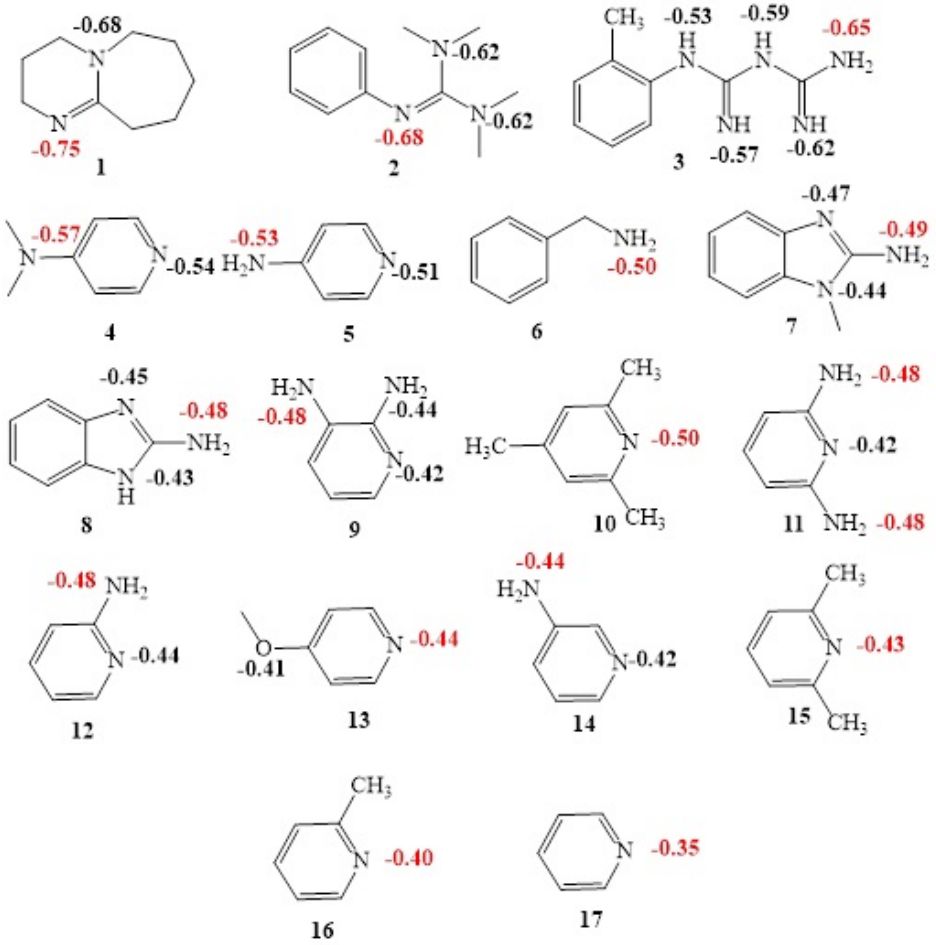


Figure 5.7. 2D representations of selected nitrogen-containing heterocycles with NPA charges on the heteroatoms (the most negative one is colored with red) (M062X/6-31G**/SMD/MeCN, pop=npa).

The experimental and predicted pK_a's of N-containing aromatic compounds are presented in Table 5.3. Maximum deviation of the predicted $pK_{a(MeCN)}$'s from experimental $pK_{a(MeCN)}$'s is found to be 0.77 unit and a very good agreement between calculated and experimental values were obtained with R^2=0.98 (Figure C.3). More-

over, small MAD value calculated (0.33) demonstrates the applicability of the proposed methodology to the acetonitrile environment. In addition to the $pK_{a(MeCN)}$ calculations, the water pK_a's of the molecules in Figure 5.7 are calculated with the proposed methodology and the results are presented in Table 5.3. Among the seventeen selected molecules, experimental $pK_{a(water)}$'s were collected from the literature for fourteen of them. A very good correlation is obtained between the calculated and the experimental pK_a's with a maximum deviation of 0.53 units. MeCN is a polar aprotic solvent with a polarized C≡N triple bond. Unlike water, it does not donate hydrogens to anions and therefore, when used as a solvent, the ionization of the molecules in MeCN is more difficult compared to water, which in turn results in higher pK_a's, and thus stronger basicity's in MeCN environment. The pK_a for these molecules in MeCN ranges between 24.77 and 12.56 [159], whereas in water it is between 13.50 and 5.23 [160].

Table 5.3. Calculated and eperimental pK_a's of nitrogen containing aromatic compounds (M062X/6-31G**/SMD/MeCN/water).

ID	**Experimental pK_a (MeCN)**	**Calculated pK_a (MeCN)**	$\triangle pK_a$ **(MeCN)**	**Experimental pK_a (water)**	**Calculated pK_a (water)**	$\triangle pK_a$ **(water)**
1	24.77	24.13	-0.64	13.50	13.02	-0.48
2	19.83	20.60	0.77	NA	11.56	-
3	19.36	19.43	0.07	NA	11.30	-
4	18.11	17.74	-0.37	9.60	9.07	-0.53
5	17.78	17.40	-0.38	9.17	9.02	-0.15
6	16.44	16.70	0.26	9.34	9.06	-0.28
7	16.02	16.11	0.09	NA	8.76	-
8	15.96	15.87	-0.09	NA	8.45	-
9	14.32	15.03	-0.71	6.78	6.43	-0.35
10	15.08	14.77	-0.31	7.43	7.70	0.27
11	13.94	14.56	0.12	6.13	5.85	-0.28
12	13.88	14.26	0.38	6.82	6.60	-0.22
13	14.02	14.04	0.02	6.58	7.00	0.42
14	14.31	13.96	-0.35	6.04	5.64	-0.40
15	13.77	13.92	0.15	6.72	6.32	-0.40
16	12.98	13.11	0.13	5.81	5.19	-0.38
17	12.56	12.33	-0.23	5.23	5.43	0.20
RMSE			0.39			0.27
MAD			0.33			0.21
MD			-0.05			-0.08

5.3.4. Prediction of pK_a(MeCN)'s of Thiazol-2-imines

In this study, the goal is to predict the pK_a's of thiazol-2-imines synthesized by Tuncel and Dogan[161] in acetonitrile environment. An accurate protocol is suggested by a two-step methodology validation. In the final step of the study, $pK_{a,(MeCN}$'s of thiazol-2-imines were calculated by employing the isodesmic reaction scheme at M062X/6-31G** level of theory in conjunction with the SMD solvation model using acetonitrile as solvent. Figure 5.8 illustrates an isodesmic reaction where the protonated 1RR is the acidic species HA^+ and 2-imino-thiazolidinone is the reference species B_{Ref}.

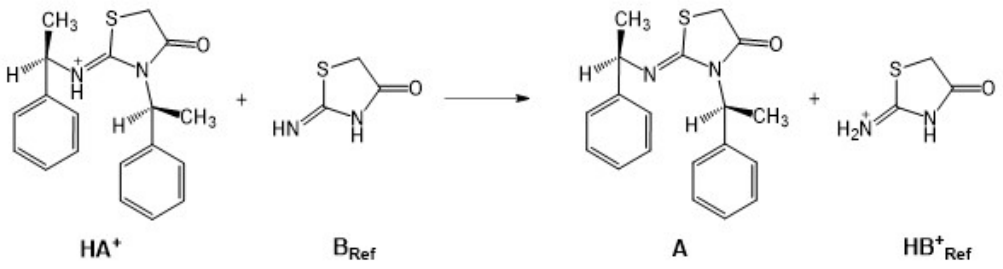


Figure 5.8. Isodesmic reaction between an acid species (HA^+) and a reference species (B_{Ref}).

To the best of our knowledge, the $pK_{a,(MeCN}$ of the reference molecule 2-imino-thiazolidinone is not known experimentally, its $pK_{a,(MeCN}$ is first predicted by employing the established methodology. The calculated $pK_{a,(MeCN}$ of 2-imino-thiazolidinone using acetonitrile as solvent is 19.41 (TC2/M062X/6-31G**/SMD), and this prediction will further be used in the next step to propose $pK_{a,(MeCN}$'s of thiazol-2-imine derivatives by employing the isodesmic reaction scheme.

The computed charges indicate that the protonation dominantly occurs on the imine-N as presented in Table 5.4 and the predicted pK_a's of thiazol-2-imines are given in Figure 5.9. We observe that 1RR, 2RR, 3RR and 4RR have relatively lower pK_a values compared to 5RR, 6RR, 7RR and 8RR. Withdrawal of the electrons from the

exocylic imine N by the carbonyl substituents on thiazole ring in 1RR, 2RR, 3RR and 4RR results in less electron density around N atom than imine N atom of 5RR, 6RR, 7RR and 8RR (q_N = -0.582 e for 1RR; q_N = -0.745 e for 5RR) (Table 5.4). The lower electron density on protonation site makes these molecules less available for proton abstraction. Thus, 1RR, 2RR, 3RR and 4RR were predicted to be less basic compared to 5RR, 6RR, 7RR and 8RR. For the molecules of our interest, the pK_a range in MeCN is between 17.16-10.59 and 7.42-1.77 in water.

Table 5.4. NPA[1], CM5[2] and Hirshfeld[3] charges of nitrogen atoms of thiazol-2-imines considered in this study (M062X/6-31G**/SMD/MeCN).

ID	**Imine-nitrogen**	**Ring nitrogen**
1RR	-0.582[1]	-0.336[1]
	-0.616[2]	-0.445[2]
	-0.602[3]	-0.403[3]
2RR	-0.594[1]	-0.345[1]
	-0.626[2]	-0.455[2]
	-0.612[3]	-0.413[3]
3RR	-0.612[1]	-0.409[1]
	-0.647[2]	-0.489[2]
	-0.625[3]	-0.447[3]
4RR	-0.623[1]	-0.421[1]
	-0.659[2]	-0.435[2]
	-0.646[3]	-0.428[3]
5RR	-0.745[1]	-0.471[1]
	-0.802[2]	-0.455[2]
	-0.7673	-0.4633
6RR	-0.704[1]	-0.436[1]
	-0.759[2]	-0.448[2]
	-0.738[3]	-0.442[3]
7RR	-0.824[1]	-0.503[1]
	-0.865[2]	-0.532[2]
	-0.847[3]	-0.517[3]
8RR	-0.805[1]	-0.487[1]
	-0.849[2]	-0.545[2]
	-0.836[3]	-0.503[3]

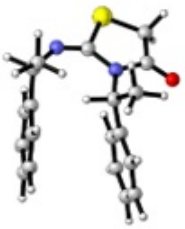
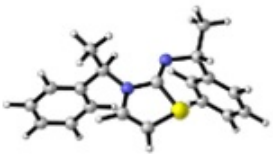
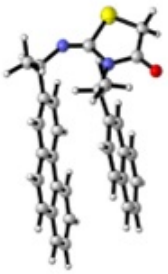
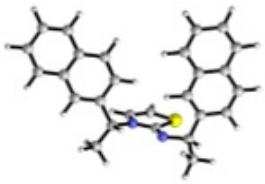
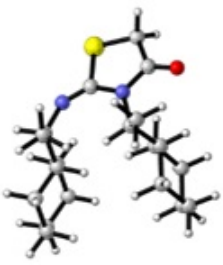
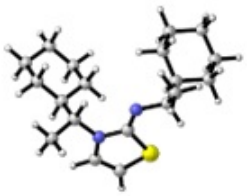
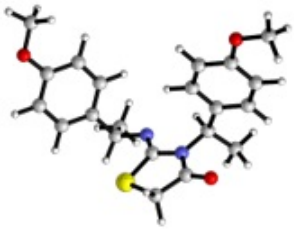
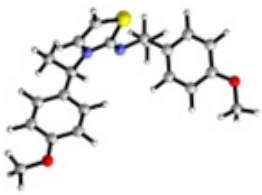
ID	Predicted pK_a	ID	Predicted pK_a
 **1RR**	10.65[1] 1.84[2]	 **5RR**	16.40[1] 6.73[2]
 **2RR**	10.59[1] 1.77[2]	 **6RR**	14.57[1] 4.35[2]
 **3RR**	11.08[1] 2.05[2]	 **7RR**	17.16[1] 6.52[2]
 **4RR**	11.57[1] 2.35[2]	 **8RR**	16.57[1] 7.42[2]

Figure 5.9. 3D representations of thiazol-2-imine derivatives and their predicted pK_a in MeCN[1] and water[2].

5.3.5. Conclusions

The pK_a of a drug molecule and the membrane environment are the key factors controlling how well a drug penetrates through the cell membranes. In this study, an elaborate protocol to evaluate the pK_a's of the single enantiomers of thiazol-2-imines is proposed. Different functionals and basis set combinations were tested to evaluate the $pK_{a,\,(water)}$ of a reference molecule resembling the molecules synthesized by Dogan et al., 2-imino-thiazolidinone, to be employed in isodesmic reaction scheme. M062X/6-31G** level of theory and SMD solvation model are found to be adequate for the prediction of $pK_{a,\,(water)}$ of the reference molecule. In order to ensure that the constructed methodology describes the nitrogen-containing heterocyclic compounds well, a two-step validation procedure is followed. First, $pK_{a,\,(water)}$ of 2-(phenylimino)imidazolidine derivatives were calculated with the proposed methodology and the predicted values were found to be very close to the experimental $pK_{a,\,(water)}$'s. Being interested in pK_a calculations in a less polar medium, it was necessary to test the methodology for compounds similar to the species of interest whose pK_a values in acetonitrile are experimentally known. Thus, the applicability of the method was verified on a set of nitrogen-containing heterocyclic compounds with experimentally known $pK_{a,\,(MeCN)}$. With the justified methodology, the experimentally unknown $pK_{a,\,(MeCN)}$ of the reference molecule, 2-imino-thiazolidinone, was calculated and finally, isodesmic reactions between the reference molecule and thiazol-2imine derivatives synthesized by Dogan et al. were proposed to evaluate the $pK_{a,\,(MeCN)}$'s of the latter. The predicted pK_a values of these drug-like molecules will enable the evaluation of their lipophilicity/hydophilicity at a given pH.

6. CONCLUDING REMARKS

In the first part of the dissertation, the effects of pharmaceuticals and personal care products (PPCPs) on the environment are disscussed since they can affect in a bad way all ecosystems. There are many PPCPs in the lakes, seas and oceans and automatically the organisms in the water systems and humanbeing are affected from these chemicals. Therefore, they should be eliminated from the environment. Advanced oxidation processes (AOPs) are one of the most popular technique to eliminate chemicals from the water and AOPs are based on generating a radical that attacks the hazardous material and turn them into enviromentally safe molecules. Generally, OH radicals are generated with different ways like UV/US, Fenton processes. OH radicals are non-selective and attack the molecules and degrade them easily. In this thesis, salicylic acid, methyl paraben and caffeine are chosen as PPCPs because they are mostly detected in waste-water systems. The degradation mechanisms and their byproducts are investigated. The toxicity of the byproducts to environment also have been evaluated.

In the second part of the study, Statistical Assesment of of the Modeling of Protein and Ligands (SAMPL) challenges are evaluated. SAMPL challenges are blind challenges. By using computational tools, crucial features of a drug-like molecules like pK_a and LogP have been estimated and then compared with the experimental data. This part includes; SAMPL6, SAMPL7 and SAMPL8 challenges. By using hybrid methods, the properties of molecules such as LogP and pK_a have been evaluted.

In the third part of the study, the pK_a values of thiazol-2-imine derivaties in MeCN are determined by using computational tools with isodesmic calculations. These molecules are important since they have some crucial biological activities, they have anti-fungal and anti-bacterial properties. This part reflects a successful achievement of various methodologies in order to calculate the dissociation constants of drug-like molecules.

REFERENCES

1. Boxall, A.B.A., M.A Rudd, B.W Brooks, D.J. Caldwell, K. Choi, S. Hickmann, E. Innes, K. Ostapyk, J.P. Staveley, T. Verslycke, G.T. Ankley, K.F. Beazley, S.E. Belanger, J.P. Berninger, P. Carriquiriborde, A. Coors, P.C. Deleo, S.D. Dyer, J.F. Ericson, F. Gagné, J.P. Giesy, T. Gouin, L. Hallstrom, M.V. Karlsson, D.G.J. Larsson, J.M. Lazorchak, F. Mastrocco, A. McLaughlin, M.E. McMaster, R.D. Meyerhoff, R. Moore, J.L. Parrott, J.R. Snape, R. Murray-Smith, M.R. Servos, P.K. Sibley, J.O. Straub, N.D. Szabo, E. Topp, G.R. Tetreault, V.L. Trudeau and G.V.D. Kraak, "Brogan & Partners Pharmaceuticals and Personal Care Products in the Environment: What Are the Big Questions? Review Pharmaceuticals and Personal Care Products in the Environment: What Are the Big Questions?", *Source: Environmental Health Perspectives*, Vol. 120, No. 9, pp. 1221–1229, 2012.

2. Nikolaou, A., S. Meric and D. Fatta, "Occurrence Patterns of Pharmaceuticals in Water and Wastewater Environments", *In Analytical and Bioanalytical Chemistry*, Vol. 387, No. 4, pp. 1225–1234, 2007.

3. Lian, P., R.C. Johnston, J.M. Parks and J.C. Smith, "Quantum Chemical Calculation of p Kas of Environmentally Relevant Functional Groups: Carboxylic Acids, Amines, and Thiols in Aqueous Solution", *Journal of Physical Chemistry A*, Vol. 122, No. 17, pp. 4366–4374, 2018.

4. Ebele, A.J., M. Abou-Elwafa Abdallah and S. Harrad, "Pharmaceuticals and Personal Care Products (Ppcps) in the Freshwater Aquatic Environment", *Emerging Contaminants*, Vol. 3, No. 1. KeAi Communications Co., pp. 1–16, 2017.

5. Dębska, J., A. Kot-Wasik and J. Namieśnik, "Fate and Analysis of Pharmaceutical Residues in the Aquatic Environment", *Critical Reviews in Analytical Chemistry*, Vol. 34, No. 1, pp. 51–67, 2004.

6. Rice, S.L. and S. Mitra, "Microwave-Assisted Solvent Extraction of Solid Matrices and Subsequent Detection of Pharmaceuticals and Personal Care Products (PPCPS) Using Gas Chromatography-Mass Spectrometry", *Analytica Chimica Acta*, Vol. 589, No. 1, pp. 125–132, 2007.

7. Roberts, P.H. and K.V. Thomas, "The Occurrence of Selected Pharmaceuticals in Wastewater Effluent and Surface Waters of the Lower Tyne Catchment", *Science of the Total Environment*, Vol. 356, No. 1-3, pp. 143–153, 2006.

8. Hernando, M.D., M. Mezcua, A.R. Fernández-Alba, and D. Barceló, "Environmental Risk Assessment of Pharmaceutical Residues in Wastewater Effluents, Surface Waters And Sediments", *In Talanta*, Vol. 69, No. 2, pp. 334–342, 2006.

9. Glaze, W.H., J.W. Kang and D.H. Chapin, "The Chemistry of Water Treatment Processes Involving Ozone, Hydrogen Peroxide and Ultraviolet Radiation", *Ozone Science Engineering*, Vol. 9, No. 4, pp. 335-352, 1987.

10. Lee, B.C.Y., F.Y. Lim, W.H. Loh, S.L. Ong and J. Hu, "Emerging Contaminants: An Overview of Recent Trends for Their Treatment and Management Using Light-Driven Processes", *Water (Switzerland)*, Vol. 13, No. 17, 2021.

11. Audenaert, W.T.M., Y. Vermeersch, S.W.H. van Hulle, P. Dejans, A. Dumoulin and I. Nopens, "Application of A Mechanistic UV/Hydrogen Peroxide Model at Full-Scale: Sensitivity Analysis, Calibration and Performance Evaluation", *Chemical Engineering Journal*, Vol. 171, No. 1, pp. 113–126, 2011.

12. Brillas, E., E. Mur, R. Sauleda, L. Sànchez, J. Peral, X. Domènech, J. Casado, "Aniline Mineralization By AOP's: Anodic Oxidation, Photocatalysis, Electro-Fenton and Photoelectro-Fenton Processes", *Applied Catalysis B: Environmental*, Vol. 16, No. 1, pp. 31-42, 1998.

13. Cai, Q.Q., L. Jothinathan, S.H. Deng, S L. Ong, H.Y. Ng and J.Y. Hu, "Fenton-

and Ozone-Based AOP Processes for Industrial Effluent Treatment", *In Advanced Oxidation Processes for Effluent Treatment Plants*, pp. 199–254, 2021.

14. Thompson, T.L. and J.T. Yates, "Surface Science Studies of The Photoactivation of TIO2 - New Photochemical Processes", *Chemical Reviews*, Vol. 106, No. 10. pp. 4428–4453, 2006.

15. Hodges, G., C. Eadsforth, B. Bossuyt, A. Bouvy, M.H. Enrici, M. Geurts, M. Kotthoff, E. Michie, D. Miller, J. Müller, G. Oetter, J. Roberts, D. Schowanek, P. Sun and J. Venzmer, "A Comparison of Log K Ow (N-Octanol–Water Partition Coefficient) Values for Non-Ionic, Anionic, Cationic and Amphoteric Surfactants Determined Using Predictions and Experimental Methods", *Environmental Sciences Europe*, Vol. 31, No. 1, 2019.

16. Cumming, H. and C. Rücker, "Octanol-Water Partition Coefficient Measurement by a Simple 1H NMR Method", *American Chemical Society*, Vol. 2, No. 9, pp. 6244–6249, 2017.

17. Bannan, C.C., G. Calabró, D.Y. Kyu and D.L. Mobley, "Calculating Partition Coefficients of Small Molecules in Octanol/Water and Cyclohexane/Water", *Academic Journals for Chemical Engineering*, Vol. 12, No. 8, pp. 4015–4024, 2016.

18. Mitani, G.M., I. Steinberg, E.J. Lien, E.C. Harrison, and U. Elkayam, "The Pharmacokinetics of Antiarrhythmic Agents in Pregnancy and Lactation", *National Library of Medicine*, Vol. 12, No. 4, pp. 253-291, 1987.

19. Avdeef, A., "Physicochemical Profiling (Solubility, Permeability and Charge State)", Current Topics in Medicinal Chemistry, 2004.

20. Kerns, E.H. and L. Di, "Physicochemical Profiling: Overview of The Screens1", *Drug Discovery Today: Technologies*, Vol. 1, No. 4. pp. 343–348, 2004.

21. Moor, M.J., S.H. Steiner, G. Jachertz and M.H. Bickel, "Adipose Tissue Distribu-

tion and Chemical Structure of Basic Lipophilic Drugs: Desipramine", *N-Acetyl Desipramine, and Halo Per Idol*, Vol. 70, No. 2, pp. 121-124, 1992.

22. Zhou, C., Y. Jin, J. R. Kenseth, M. Stella, K.R. Wehmeyer, and W.R. Heineman, "Rapid pKa Estimation Using Vacuum-Assisted Multiplexed Capillary Electrophoresis (VAMCE) With Ultraviolet Detection", *Journal of Pharmaceutical Sciences*, Vol. 94, No. 3, pp. 576–589, 2005.

23. Wan, H. and J. Ulander, "High-Throughput Pka Screening and Prediction Amenable for ADME Profiling", *Expert Opinion on Drug Metabolism and Toxicology*, Vol. 2, No. 1. pp. 139–155, 2006.

24. Savun-Hekimoğlu, B. and N.H. Ince, "Decomposition of PPCPS By Ultrasound-Assisted Advanced Fenton Reaction: A Case Study with Salicylic Acid", *Ultrason Sonochem*, Vol. 39, pp. 243–249, 2017.

25. Kasprzyk-Hordern, B., R. M. Dinsdale and A.J. Guwy, "The Occurrence of Pharmaceuticals, Personal Care Products, Endocrine Disruptors and Illicit Drugs in Surface Water in South Wales, UK", *Water Research*, Vol. 42, No. 13, pp. 3498–3518, 2008.

26. Brodin, T., J. Fick, M. Jonsson and J. Klaminder, "Dilute Concentrations of A Psychiatric Drug Alter Behavior of Fish From Natural Populations", PubMed records, Vol. 339, No. 6121, pp. 814–815, 2013.

27. Gültekin, I. and N.H. Ince, "Synthetic Endocrine Disruptors in The Environment and Water Remediation by Advanced Oxidation Processes", *Journal of Environmental Management*, Vol. 85, No. 4. pp. 816–832, 2007.

28. Ince, N.H., "Ultrasound-Assisted Advanced Oxidation Processes for Water Decontamination", *Ultrason Sonochem*, Vol. 40, pp. 97–103, 2018.

29. Suslick, K.S., "Variation of Liquid Temperature at Bubble Wall near the Sonolu-

minescence Threshold", *Journal of the Physical Society of Japan*, Vol. 65, No. 9, pp. 2830-2840, 1986.

30. Mason, T.J., J.P. Lorimer and D.M. Bates, "Quantifying Sonochemistry: Casting Some Light On A 'Black Art'", *Ultrasonics*, Vol. 30, No. 1, pp. 40-42, 1992.

31. Chatel, G., L. Novikova and S. Petit, "How Efficiently Combine Sonochemistry and Clay Science?", *Applied Clay Science*, Vol. 119. pp. 193–201, 2016.

32. Nakada, N., H. Shinohara, A. Murata, K. Kiri, S. Managaki, N. Sato, H. Takada, "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 Research*, Vol. 41, No. 19, pp. 4373–4382, 2007.

33. Chaniotakis, N.A., S.B. Park and M.E. Meyerhoff, "Salicylate-Selective Membrane Electrode Based on Tin(1V) Tetraphenylporphyrin", *American Chemical Society*, Vol. 61, No. 6, pp. 566–570, 1989.

34. Stamatis, N.K. and I.K. Konstantinou, "Occurrence and Removal oof Emerging Pharmaceutical, Personal Care Compounds and Caffeine Tracer In Municipal Sewage Treatment Plant In Western Greece", *Journal of Environmental Science and Health*, Vol. 48, No. 9, pp. 800–813, 2013.

35. Amin, A., S. Chauhan, M. Dare and A.K. Bansal, "Degradation of Parabens by Pseudomonas Beteli and Burkholderia Latens", *European Journal of Pharmaceutics and Biopharmaceutics*, Vol. 75, No. 2, pp. 206-212, 2010.

36. Karnik, B.S., S.H. Davies, M.J. Baumann and S.J. Masten, "Use of Salicylic Acid as A Model Compound to Investigate Hydroxyl Radical Reaction in an Ozonation-Membrane Filtration Hybrid Process", *Environmental Engineering Science*, Vol. 24, No. 6, pp. 852–860, 2007.

37. Ramaswamy, B.R., J.W. Kim, T. Isobe, K.H. Chang, A. Amano, T.W. Miller, F.P. Siringan, S. Tanabe, "Determination of Preservative And Antimicrobial Compounds in Fish From Manila Bay, Philippines Using Ultra High Performance Liquid Chromatography Tandem Mass Spectrometry, and Assessment of Human Dietary Exposure", *The Journal of Hazardous Materials*, Vol. 192, No. 3, pp. 1739-1745, 2011.

38. Márquez-Sillero, I., E. Aguilera-Herrador, S. Cárdenas, and M. Valcárcel, "Determination of Parabens in Cosmetic Products Using Multi-Walled Carbon Nanotubes as Solid Phase Extraction Sorbent and Corona-Charged Aerosol Detection System", *Journal of Chromatography A*, Vol. 1217, No. 1, pp. 1-6, 2010.

39. Verenitch, S.S., C.J. Lowe and A. Mazumder, "Determination of Acidic Drugs and Caffeine in Municipal Wastewaters and Receiving Waters by Gas Chromatography-Ion Trap Tandem Mass Spectrometry", J*ournal of Chromatography A*, Vol. 1116, No. 1–2, pp. 193–203, 2006.

40. Frontistis, Z., M. Antonopoulou, I. Konstantinou and D. Mantzavinos, "Degradation of Ethyl Paraben by Heat-Activated Persulfate Oxidation: Statistical Evaluation of Operating Factors and Transformation Pathways", *Environmental Science and Pollution Research*, Vol. 24, No. 2, pp. 1073–1084, 2017.

41. Cleuvers, M., "Mixture Toxicity of the Anti-Inflammatory Drugs Diclofenac, Ibuprofen, Naproxen, and Acetylsalicylic Acid", *Ecotoxicology and Environmental Safety*, Vol. 59, No. 3, pp. 309–315, 2004.

42. Savun-Hekimoğlu, B. and N.H. Ince, "Sonochemical and Sonocatalytic Destruction of Methylparaben Using Raw, Modified and SDS-Intercalated Particles of a Natural Clay Mineral", *Ultrason Sonochem*, Vol. 54, pp. 233–240, 2019.

43. Sasi, S., M.P. Rayaroth, D. Devadasan, U.K. Aravind and C.T. Aravindakumar, "Influence of Inorganic Ions and Selected Emerging Contaminants on the Degra-

dation of Methylparaben: A Sonochemical Approach", *The Journal of Hazardous Materials*, Vol. 300, pp. 202–209, 2015.

44. Lee, C., E. Yang, and R.G. Parr, "Development of the Colic-Salvetti Correlation-Energy Formula Into a Functional of the Electron Density", Vol. 37, pp. 785–790, 1988.

45. Becke, A.D., "A New Mixing of Hartree-Fock and Local Density-Functional Theories", *The Journal of Chemical Physics*, Vol. 98, No. 2, pp. 1372–1377, 1993.

46. Becke, A.D., "Density-Functional Thermochemistry, III. The Role of Exact Exchange", *The Journal of Chemical Physics*, Vol. 98, No. 7, pp. 5648–5652, 1993.

47. Frisch, M.J., G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, G.A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B.G. Janesko, R. Gomperts, B. Mennucci, H.P. Hratchian, J.V. Ortiz, A.F. Izmaylov, J.L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V.G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J.A. Montgomery, Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J.M. Millam, M. Klene, C. Adamo, R. Cammi, J.W. Ochterski, R.L. Martin, K. Morokuma, O. Farkas, J.B. Foresman and D.J. Fox, "Gaussian 09, Revision A.02", Gaussian Also Includes, Wallingford CT, 2016.

48. Frisch, M.J., G.W., Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, G.A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A.V. Marenich, J. Bloino, B.G. Janesko, R. Gomperts, B. Mennucci, H.P. Hratchian, J.V. Ortiz, A.F. Izmaylov, J.L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D.

Ranasinghe, V.G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J.A. Montgomery, J.E. Jr. Peralta, F. Ogliaro, M.J. Bearpark, J.J. Heyd, E.N. Brothers, K.N. Kudin, V.N. Staroverov, T.A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A.P. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, J.M. Millam, M. Klene, C. Adamo, R. Cammi, J.W. Ochterski, R.L. Martin, K. Morokuma, O. Farkas, J.B. Foresman, D.J. Fox, "Gaussian 16, Revision C.01", Gaussian Also Includes, Wallingford CT, 2016.

49. Becke, A.D., "Density-Functional Thermochemistry. III. The Role of Exact Exchange", *The Journal of Chemical Physics*, Vol. 98, No. 7, pp. 5648–5652, 1993.

50. Becke, A.D., "Density-Functional Thermochemistry. III. The Role of Exact Exchange" *The Journal of Chemical Physics*, Vol. 98, No. 7, pp. 5648–5652, 1993.

51. Giannozzi, P., S.Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo1, A.D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A.P. Seitsonen, A. Smogunov, P. Umari and R.M Wentzcovitch, "Quantum ESPRESSO Toward the Exascale'", *Journal of Chemical Physics*, Vol. 152, No. 15, 2020.

52. Giannozzi, P., O, Andreussi, T. Brumme, O. Bunau, M. Buongiorno Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. Dal Corso, S. de Gironcoli, P. Delugas, R.A. DiStasio Jr, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H-Y. Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N.L. Nguyen, H-V. Nguyen, A. Otero-de-la-Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A.P. Seitsonen, A. Smogunov, I. Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu and S.

Barn, “Advanced capabilities for materials modelling with Quantum ESPRESSO”, *Journal of Physics Condensed Matter*, Vol. 29, No. 46, 2017.

53. Giannozzi, P., S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni, I. Dabo, A.D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L.M. Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A.P. Seitsonen, A. Smogunov, P. Umari, R.M. Wentzcovitch, “Quantum Espresso: A Modular and Open-Source Software Project for Quantum Simulations of Materials”, *Journal of Physics Condensed Matter*, Vol. 21, No. 39, 2009.

54. Scheck, C.K. and F. H. Frimmel, “Degradation of Phenol and Salicylic Acid by Ultraviolet Radiation/Hydrogen Peroxide/Oxygen”, *Water Research*, Vol. 29, No. 10, pp. 2346-2352, 1995.

55. Chang, C.Y., Y.H. Hsieh, K.Y. Cheng, L.L. Hsieh, T.C. Cheng and K.S. Yao, “Effect of pH on Fenton Process Using Estimation of Hydroxyl Radical With Salicylic Acid as Trapping Reagent”, *Water Science and Technology*, Vol. 58, No. 4, pp. 873–879, 2008.

56. Lin, Y., C. Ferronato, N. Deng, F. Wu and J.M. Chovelon, “Photocatalytic Degradation of Methylparaben by TiO2: Multivariable Experimental Design and Mechanism”, *Applied Catalysis B*, Vol. 88, No. 1–2, pp. 32–41, 2009.

57. Gmurek, M., A.F. Rossi, R.C. Martins, R.M. Quinta-Ferreira, and S. Ledakowicz, “Photodegradation of Single and Mixture of Parabens - Kinetic, by-Products Identification and Cost-Efficiency Analysis”, *Chemical Engineering Journal*, Vol. 276, pp. 303–314, 2015.

58. Agopcan C.S., A. Ziylan-Yavaş, S. Catak, N.H. Ince, and V. Aviyente, “Hydroxyl Radical-Mediated Degradation of Diclofenac Revisited: A Computational Ap-

proach to Assessment of Reaction Mechanisms and by-Products", *Environmental Science and Pollution Research*, Vol. 24, No. 22, pp. 18458–18469, 2017.

59. Gao, Y., Y. Ji, G. Li, and T. An, "Theoretical Investigation on The Kinetics and Mechanisms of Hydroxyl Radical-Induced Transformation of Parabens and Its Consequences for Toxicity: Influence of Alkyl-Chain Length", *Water Research*, Vol. 91, pp. 77–85, 2016.

60. Agopcan Cinar, S., A. Ziylan-Yavaş, S. Catak, N. H. Ince and V. Aviyente, "Hydroxyl Radical-Mediated Degradation of Diclofenac Revisited: A Computational Approach to Assessment of Reaction Mechanisms and by-Products", *Environmental Science and Pollution Research*, Vol. 24, No. 22, pp. 18458–18469, 2017.

61. Bruton, T., A. Alboloushi, B. de La Garza, B.O. Kim, and R.U. Halden, "Fate Of Caffeine In The Environment And Ecotoxicological Considerations", *In ACS Symposium Series*, Vol. 1048, pp. 257–273, 2010.

62. Załeska-Radziwiłł, M., K. Affek, and J. Rybak, "Ecotoxicity of Chosen Pharmaceuticals in Relation to Micro-Organisms-Risk Assessment", *Desalination Water Treat*, Vol. 52, No. 19–21, pp. 3908–3917, 2014.

63. Gardinali, P.R. and X. Zhao, "Trace Determination of Caffeine in Surface Water Samples by Liquid Chromatography-Atmospheric Pressure Chemical Ionization-Mass Spectrometry (LC-APCI-MS)", *Environment International*, Vol. 28, No. 6, pp. 521-580, 2003.

64. Ziylan-Yavas, A., N.H. Ince, E. Ozon, E. Arslan, V. Aviyente, B. Savun-Hekimoğlu, A. Erdincler, "Oxidative Decomposition and Mineralization of Caffeine by Advanced Oxidation Processes: The Effect of Hybridization", *Ultrason Sonochem*, Vol. 76, 2021.

65. Dalmázio, I., L.S. Santos, R.P. Lopes, M.N. Eberlin, and R. Augusti, "Advanced

Oxidation of Caffeine in Water: On-Line and Real-Time Monitoring by Electrospray Ionization Mass Spectrometry", *Environmental Science Technology*, Vol. 39, No. 16, pp. 5982–5988, 2005.

66. Choi, Y., D. Lee, S. Hong, S. Khan, B. Darya, J.-Y. Lee, J. Chung and S.-H. Cho, "Investigation of the Synergistic Effect of Sonolysis and Photocatalysis oof Titanium Dioxide for Organic Dye Degradation", *Catalysts*, Vol. 10, No. 5, 2020.

67. Reyes, N.J.D.G., F.K.F. Geronimo, K.A. v. Yano, H.B. Guerra, and L.H. Kim, "Pharmaceutical and Personal Care Products in Different Matrices: Occurrence, Pathways, and Treatment Processes", *Water (Switzerland)*, Vol. 13, No. 9, 2021.

68. Geballe, M.T., A.G. Skillman, A. Nicholls, J.P. Guthrie, and P.J. Taylor, "The SAMPL2 Blind Prediction Challenge: Introduction and Overview", *Journal of Computer-Aided Molecular Design*, Vol. 24, No. 4, pp. 259–279, 2010.

69. Muddana, H.S., A.T. Fenley, D.L. Mobley, and M.K. Gilson, "The SAMPL4 Host-Guest Blind Prediction Challenge: An Overview", *Journal of Computer-Aided Molecular Design*, Vol. 28, No. 4. pp. 305–317, 2014.

70. Muddana, H.S., C.D. Varnado, C.W. Bielawski, A.R. Urbach, L. Isaacs, M.T. Geballe and M.K. Gilson, "Blind prediction of host-guest binding affinities: A new SAMPL3 challenge", *In Journal of Computer-Aided Molecular Design*, Vol. 26, No. 5, pp. 475–487, 2012.

71. Skillman, A.G., "SAMPL3: Blinded Prediction of Host-Guest Binding Affinities, Hydration Free Energies, and Trypsin Inhibitors", *Journal of Computer-Aided Molecular Design*, Vol. 26, No. 5, pp. 473–474, 2012.

72. Wu, S.C., C.S. Tan, and M.H. Huang, "Strong Facet Effects on Interfacial Charge Transfer Revealed through the Examination of Photocatalytic Activities of Various Cu2O–ZnO Heterostructures", *Advanced Functional Materials*, Vol. 27, No. 9,

2017.

73. Muddana, H.S. and M.K. Gilson, "Prediction of SAMPL3 Host-Guest Binding Affinities: Evaluating the Accuracy of Generalized Force-Fields", *In Journal of Computer-Aided Molecular Design*, Vol. 26, No. 5, pp. 517–525, 2012.

74. Rustenburg, A.S., J. Dancer, B. Lin, J.A. Feng, D.F. Ortwine, D.L. Mobley and J.D. Chodera, "Measuring Experimental Cyclohexane-Water Distribution Coefficients for the SAMPL5 Challenge", *The Journal of Computer-Aided Molecular Design*, Vol. 30, No. 11, pp. 945–958, 2016.

75. Bannan, C.C., D.L. Mobley, and A.G. Skillman, "SAMPL6 Challenge Results from Pka Predictions Based on a General Gaussian Process Model", *The Journal of Computer-Aided Molecular Design*, Vol. 32, No. 10, pp. 1165–1177, 2018.

76. Leo, A., C. Hansch and D. Elkins, "Chemical Reviews Partition Coefficients and Their Uses", *Partition Coefficients And Their Uses*, Vol. 71, No. 6, pp. 525–616, 1971.

77. Işık, M., D. Levorse, D.L. Mobley, T. Rhodes, and J.D. Chodera, "Octanol–Water Partition Coefficient Measurements for the SAMPL6 Blind Prediction Challenge", *The Journal of Computer-Aided Molecular Design*, Vol. 34, No. 4, pp. 405–420, 2020.

78. Tehrany, E.A., F. Fournier, and S. Desobry, "Simple Method to Calculate Octanol-Water Partition Coefficient of Organic Compounds", *Journal of Food Engineering*, Vol. 64, No. 3, pp. 315–320, 2004.

79. Sangster, J.M., "Octanolwater-Partition-Coefficients-Fundamentals-and-Physical-Ch-1997", Physical Chemistry, 1997.

80. Bannan, C.C., G. Calabró, D.Y. Kyu, and D.L. Mobley, "Calculating Partition Coefficients of Small Molecules in Octanol/Water and Cyclohexane/Water", *Journal*

of Chemical Theory and Computation, Vol. 12, No. 8, pp. 4015–4024, 2016.

81. Xiang, Q., G. Shan, W. Wu, H. Jin, and L. Zhu, "Measuring Log Kow Coefficients of Neutral Species of Perfluoroalkyl Carboxylic Acids Using Reversed-Phase High-Performance Liquid Chromatography", *Environmental Pollution*, Vol. 242, pp. 1283–1290, 2018.

82. Port, A. M. Bordas, R. Enrech, R. Pascual, M. Rosés, C. Ràfols, X. Subirats, E. Bosch, "Critical Comparison of Shake-Flask, Potentiometric and Chromatographic Methods for Lipophilicity Evaluation (Log Po/W) of Neutral, Acidic, Basic, Amphoteric, and Zwitterionic Drugs", *European Journal of Pharmaceutical Sciences*, Vol. 122, pp. 331–340, 2018.

83. Nicholls, A., D.L. Mobley, J.P. Guthrie, J.D. Chodera, C.I. Bayly, M.D. Cooper and V.S. Pande, "Predicting Small-Molecule Solvation Free Energies: An Informal Blind Test for Computational Chemistry", *Journal of Medicinal Chemistry*, Vol. 51, No. 4, pp. 769–779, 2008.

84. Ditchfield, R., W.J. Hehre, and J.A. Pople, "Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules", *The Journal of Chemical Physics*, Vol. 54, No. 2, pp. 720–723, 1971.

85. Hariharan, P.C. and J.A. Pople, "Accuracy of Ahn Equilibrium Geometries by Single Determinant Molecular Orbital Theory", *Molecular Physics*, Vol. 27, No. 1, pp. 209–214, 1974.

86. Hariharan, P.C. and J.A. Pople, "The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies", Springer-Verlag, 1973.

87. Hehre, W.J., K. Ditchfield and J.A. Pople, "Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use In Molecular

Orbital Studies of Organic Molecules", *The Journal of Chemical Physics*, Vol. 56, No. 5, pp. 2257–2261, 1972.

88. Krishnan, R., J.S. Binkley, R. Seeger, and J.A. Pople, "Self-Consistent Molecular Orbital Methods. XX. A Basis Set for Correlated Wave Functions", *The Journal of Chemical Physics*, Vol. 72, No. 1, pp. 650–654, 1980.

89. McLean, A.D. and G.S. Chandler, "Contracted Gaussian Basis Sets for Molecular Calculations. I. Second Row Atoms, Z=11-18", *The Journal of Chemical Physics*, Vol. 72, No. 10, pp. 5639–5648, 1980.

90. Perdew, J.P., "Jacob's Ladder of Density Functional Approximations for the Exchange-Correlation Energy", *AIP Publishing*, pp. 1–20, 2003.

91. Kelly, C.P., C.J. Cramer and D.G. Truhlar, "Aqueous Solvation Free Energies of Ions And Ion-Water Clusters Based on an Accurate Value for the Absolute Aqueous Solvation Free Energy of the Proton", *Journal of Physical Chemistry B*, Vol. 110, No. 32, pp. 16066–16081, 2006.

92. Zhao, Y. and D.G. Truhlar, "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 Fur M06-Class Functionals and 12 Other Functionals", *Theoretical Chemistry Accounts*, Vol. 120, No. 1–3, pp. 215–241, 2008.

93. Lee, C., E. Yang and R.G. Parr, "Development of the Colic-Salvetti Correlation-Energy Formula Into A Functional of the Electron Density", *Physical Review B*, Vol. 37, pp.785-795, 1988.

94. Takano, Y. and K.N. Houk, "Benchmarking The Conductor-Like Polarizable Continuum Model (CPCM) For Aqueous Solvation Free Energies of Neutral And Ionic Organic Molecules", *Journal of Chemical Theory and Computation*, Vol. 1, No. 1,

pp. 70–77, 2005.

95. Rayne, S. and K. Forest, "Accuracy of Computational Solvation Free Energies for Neutral and Ionic Compounds: Dependence on Level of Theory and Solvent Model", Nature Precedings, 2010.

96. Bryantsev, V.S., M.S. Diallo, A.C. T. van Duin, and W.A. Goddard, "Evaluation of B3LYP, X3LYP, and M06-Class density Functionals for Predicting the Binding Energies of Neutral, Protonated, and Deprotonated Water Clusters", *Journal of Chemical Theory and Computation*, Vol. 5, No. 4, pp. 1016–1026, 2009.

97. Da Chai, J. and M. Head-Gordon, "Long-Range Corrected Hybrid Density Functionals with Damped Atom-Atom Dispersion Corrections", *Physical Chemistry Chemical Physics*, Vol. 10, No. 44, pp. 6615–6620, 2008.

98. Marenich, A.V., C.J. Cramer, and D.G. Truhlar, "Universal Solvation Model Based on Solute Electron Density and on A Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions", *Journal of Physical Chemistry B*, Vol. 113, No. 18, pp. 6378–6396, 2009.

99. Marenich, A.V., S.V Jerome, C.J. Cramer, and D.G. Truhlar, "Charge Model 5: An Extension of Hirshfeld Population Analysis for the Accurate Description of Molecular Interactions in Gaseous and Condensed Phases", *Journal of Chemical Theory and Computation*, Vol. 8, No. 2, pp. 527–541, 2012.

100. Skyner, R.E., J.L. McDonagh, C.R. Groom, T. Van Mourik, and J.B.O. Mitchell, "A Review of Methods for the Calculation of Solution Free Energies and the Modelling of Systems in Solution", *Physical Chemistry Chemical Physics*, Vol. 17, No. 9, pp. 6174–6191, 2015.

101. Skyner, R.E., J.L. McDonagh, C.R. Groom, T. van Mourik, and J.B.O. Mitchell, "A Review of Methods for the Calculation of Solution Free Energies and the Mod-

elling of Systems iin Solution", *Physical Chemistry Chemical Physics*, Vol. 17, No. 9, pp. 6174–6191, 2015.

102. Borges, N.M., P.W. Kenny, C.A. Montanari, I.M. Prokopczyk, J.F.R. Ribeiro, J.R. Rocha and G.R. Sartori, "The Influence of Hydrogen Bonding on Partition Coefficients", *The Journal of Computer-Aided Molecular Design*, Vol. 31, No. 2, pp. 16,3–181, 2017.

103. Ahn, S., M. Hong, M. Sundararajan, D.H. Ess and M.H. Baik, "Design and Optimization of Catalysts Based on Mechanistic Insights Derived from Quantum Chemical Reaction Modeling", *Chemical Reviews*, Vol. 119, No. 11. pp. 6509–6560, 2019.

104. Grimme, S., F. Bohle, A. Hansen, P. Pracht, S. Spicher and M. Stahn, "Efficient Quantum Chemical Calculation of Structure Ensembles and Free Energies for Non-rigid Molecules", *Journal of Physical Chemistry A*, Vol. 125, No. 19 pp. 4039–4054, 2021.

105. Jensen, J.H., C.J. Swain, and L. Olsen, "Prediction of Pka Values For Druglike Molecules Using Semiempirical Quantum Chemical Methods", *Journal of Physical Chemistry A*, Vol. 121, No. 3, pp. 699–707, 2017.

106. Selwa, E., I.M. Kenney, O. Beckstein, and B.I. Iorga, "SAMPL6: Calculation of Macroscopic Pk A Values From Ab Initio Quantum Mechanical Free Energies", *The Journal of Computer-Aided Molecular Design*, Vol. 32, No. 10, pp. 1203–1216, 2018.

107. Pracht, P. and S. Grimme, "Efficient Quantum-Chemical Calculations of Acid Dissociation Constants from Free-Energy Relationships", *Journal of Physical Chemistry A*, Vol. 125, No. 25, pp. 5681–5692, 2021.

108. Caine, B.A., M. Bronzato, and P.L.A. Popelier, "Experiment Stands Corrected:

Accurate Prediction of The Aqueous P: K A Values of Sulfonamide Drugs Using Equilibrium Bond Lengths", *Chemical Science*, Vol. 10, No. 25, pp. 6368–6381, 2019.

109. Popelier, A., P.L., A.P. Harding, and D.C. Wedge, "pKa Prediction From 'Quantum Chemical Topology' Descriptors", *Journal of Chemical Information and Modeling*, Vol. 49, No. 8, pp. 1914–1924, 2009.

110. Guillaume, M., E. Botek, B. Champagne, F. Castet, and L. Ducasse, "Substituent Effects on the Electronic Structure and Pka Benzoic Acid", *In International Journal of Quantum Chemistry*, Vol. 90, No. 4–5, pp. 1396–1403, 2002.

111. Gangarapu, S., A.T.M. Marcelis, and H. Zuilhof, "Accurate pKa Calculation of the Conjugate Acids of Alkanolamines, Alkaloids and Nucleotide Bases by Quantum Chemical Methods", *Wiley Online Library*, Vol. 14, No. 5, pp. 990–995, 2013.

112. Sastre, S., R. Casasnovas, F. Muñoz, and J. Frau, "Isodesmic Reaction for pKa Calculations of Common Organic Molecules", *Theoretical Chemistry Accounts*, Vol. 132, No. 2, pp. 1–8, 2013.

113. Gross, K.C. and P.G. Seybold, "Substituent Effects on the Physical Properties and pK a of Phenol", *Quantum Chemistry*, Vol. 85, No. 4-5, pp. 569-579, 2001.

114. Gross, K.C., P.G. Seybold, and C.M. Hadad, "Comparison of Different Atomic Charge Schemes for Predicting pKa Variations in Substituted Anilines and Phenols", *International Journal of Quantum Chemistry*, Vol. 90, No. 1, pp. 445–458, 2002.

115. Dutra, F.R., C.D.S. Silva, and R. Custodio, "On the Accuracy of the Direct Method to Calculate P Kafrom Electronic Structure Calculations", *Journal of Physical Chemistry A*, Vol. 125, No. 1, pp. 65–73, 2021.

116. Silverstein, T.P., "The Solvated Proton is NOT H3O+!", *Journal of Chemical*

Education, Vol. 88, No. 7, pp. 875, 2011.

117. Stewart, J.J.P. and F.J. Seiler, "Optimization of Parameters for Semiempirical Methods I. Method", *Journal of Computational Chemistry*, Vol. 10, pp. 209-220, 1989.

118. Krishnan, R., J.S. Binkley, R. Seeger, and J.A. Pople, "Self-Consistent Molecular Orbital Methods. XX. A Basis Set for Correlated Wave Functions", *The Journal of Chemical Physics*, Vol. 72, No. 1, pp. 650–654, 1980.

119. Lee, C. E. Yang, and R.G. Parr, "Development of the Colic-Salvetti Correlation-Energy Formula into A Functional of the Electron Density", *Physical Review B*, Vol. 37, pp. 785-796, 1988.

120. Grimme, S., J. Antony, S. Ehrlich, and H. Krieg, "A Consistent and Accurate AB Initio Parametrization of Density Functional Dispersion Correction (DFT-D) For The 94 Elements H-Pu", *Journal of Chemical Physics*, Vol. 132, No. 15, 2010.

121. Marenich, A.V., C.J. Cramer, and D.G. Truhlar, "Universal Solvation Model Based on Solute Electron Density and on A Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions", *Journal of Physical Chemistry B*, Vol. 113, No. 18, pp. 6378–6396, 2009.

122. Hehre, W.J., K. Ditchfield, and J.A. Pople, "Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules", *The Journal of Chemical Physics*, Vol. 56, No. 5, pp. 2257–2261, 1972.

123. Zhao, Y. and D.G. Truhlar, "A New Local Density Functional For Main-Group Thermochemistry, Transition Metal Bonding, Thermochemical Kinetics, and Non-covalent Interactions", *Journal of Chemical Physics*, Vol. 125, No. 19, 2006.

124. Ditchfield, R., W.J. Hehre, and J.A. Pople, "Self-Consistent Molecular-Orbital

Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules", *The Journal of Chemical Physics*, Vol. 54, No. 2, pp. 720–723, 1971.

125. Zhao, Y., N.E. Schultz, and D.G. Truhlar, "Design of Density Functionals By Combining the Method of Constraint Satisfaction with Parametrization for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions", *Journal of Chemical Theory and Computation*, Vol. 2, No. 2, pp. 364–382, 2006.

126. Hariharan, P.C. and J. A. Pople, "Accuracy of AHn Equilibrium Geometries by Single Determinant Molecular Orbital Theory", *Molecular Physics*, Vol. 27, No. 1, pp. 209–214, 1974.

127. Weigend, F. and R. Ahlrichs, "Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy", *Physical Chemistry Chemical Physics*, Vol. 7, No. 18, pp. 3297–3305, 2005.

128. Arslan, E., B.K. Findik, and V. Aviyente, "A Blind SAMPL6 Challenge: Insight into the Octanol-Water Partition Coefficients of Drug-Like Molecules Via A DFT Approach", *The Journal of Computer-Aided Molecular Design*, Vol. 34, No. 4, pp. 463–470, 2020.

129. Guan, X., Y. Ma, H. Li, Y. Yusran, M. Xue, Q. Fang, Y. Yan, V. Valtchev and S. Qiu, "Fast, Ambient Temperature and Pressure Ionothermal Synthesis of Three-Dimensional Covalent Organic Frameworks", *Journal of the American Chemical Society*, Vol. 140, No. 13, pp. 4494–4498, 2018.

130. Peverati, R. and D.G. Truhlar, "M11-L: A Local Density Functional that Provides Improved Accuracy for Electronic Structure Calculations in Chemistry and Physics", *Journal of Physical Chemistry Letters*, Vol. 3, No. 1, pp. 117–124, 2012.

131. Yu, H.S., X. He, and D.G. Truhlar, "MN15-L: A New Local Exchange-Correlation Functional for Kohn-Sham Density Functional Theory with Broad Accuracy for Atoms, Molecules, and Solids", *Journal of Chemical Theory and Computation*, Vol. 12, No. 3, pp. 1280–1293, 2016.

132. Sondhi, S.M., N.Singh, A.M. Lahoti, K. Baja, A. Kumar, O. Lozach, L. Meijer, "Synthesis of Acridinyl-Thiazolino Derivatives and Their Evaluation for Anti-Inflammatory, Analgesic and Kinase Inhibition Activities", *Bioorganic and Medicinal Chemistry*, Vol. 13, No. 13, pp. 4291–4299, 2005.

133. Bae, S., H.G. Hahn, K.D. Nam, and H. Mah, "Solid-Phase Synthesis of Fungitoxic 2-Imino-1,3-Thiazolines", *Journal of Combinatorial Chemistry*, Vol. 7, No. 1, pp. 7–9, 2005.

134. Barreca M.L., A. Chimirri, L.D. Luca, A.M. Monforte, P. Monforte, A. Rao, M. Zappalà, J. Balzarini, E.D. Clercq, C. Pannecouque, M. Witvrouw, "Discovery of 2,3-Diaryl-1,3-thiazolidin-4-ones as Potent Anti-HIV-1 Agents", *Bioorganic and Medicinal Chemistry Letters*, Vol. 11, No. 13, 2001, pp. 1793-1796, 2001.

135. Sondhi, S.M., N. Singh, A.M. Lahoti, K. Bajaj, A. Kumar, O. Lozach, L. Meijer, "Synthesis of Acridinyl-Thiazolino Derivatives and Their Evaluation for Anti-Inflammatory, Analgesic and Kinase Inhibition Activities", *Bioorganic and Medicinal Chemistry*, Vol. 13, No. 13, pp. 4291–4299, 2005.

136. Chen, F., Z. Ding, Y. He, J. Qin, T. Wang, and Q.H. Fan, "Asymmetric Hydrogenation of N-alkyl and N-Aryl Ketoimines Using Chiral Cationic Ru(Diamine) Complexes As Catalysts: The Counter anion and Solvent Effects, And Substrate Scope", *Tetrahedron*, Vol. 68, No. 26, pp. 5248–5257, 2012.

137. Juaristi, E., J.L. García Ruano, M.C. García, A.L. Navarro, F. Tato and A.M.M. Castro, "Issue In Honor Of Prof Reactions Of Enantiopure β-Ketimino Sulfoxides With Et 2 Alcn. Scope and Limitations In Asymmetric Synthesis of α-

Aminonitriles", ARKIVOC, 2005.

138. Tuncel, S.T. and I. Dogan, "Synthesis of Thiazol-2-Imines from the Reduction of Single Enantiomer 2-Imino-Thiazolidin-4-Ones Followed by A Spontaneous Water Elimination", *Chirality*, Vol. 32, No. 6, pp. 866–875, 2020.

139. Sastre, S., R. Casasnovas, F. Muñoz, and J. Frau, "Isodesmic Reaction for Accurate Theoretical P: K A Calculations of Amino Acids and Peptides", *Physical Chemistry Chemical Physics*, Vol. 18, No. 16, pp. 11202–11212, 2016.

140. Biswas, A.K., R. Lo, and B. Ganguly, "Is the Isodesmic Reaction Approach a Better Model for Accurate Calculation of pKa of Organic Superbases? A Computational Study", *Synlett*, Vol. 24, No. 19, pp. 2519–2524, 2013.

141. Sastre, S., R. Casasnovas, F. Muñoz, and J. Frau, "Isodesmic Reaction for pKa Calculations of Common Organic Molecules", *Theoretical Chemistry Accounts*, Vol. 132, No. 2, pp. 1–8, 2013.

142. Casasnovas, R., J. Ortega-Castro, J. Frau, J. Donoso, and F. Muñoz, "Theoretical pKa Calculations with Continuum Model Solvents, Alternative Protocols to Thermodynamic Cycles", *International Journal of Quantum Chemistry*, Vol. 114, No. 20. pp. 1350–1363, 2014.

143. Fındık, B.K., Z.P. Haslak, E. Arslan, and V. Aviyente, "SAMPL7 Blind Challenge: Quantum–Mechanical Prediction of Partition Coefficients and Acid Dissociation Constants for Small Drug-Like Molecules", *Journal of Computer-Aided Molecular Design*, Vol. 35, No. 7, pp. 841–851, 2021.

144. Casasnovas, R., D. Fernández, J. Ortega-Castro, J. Frau, J. Donoso, and F. Muñoz, "Avoiding Gas-Phase Calculations in Theoretical pKa Predictions", *Theoretical Chemistry Accounts*, Vol. 130, No. 1, pp. 1–13, 2011.

145. Stewart, J.J.P. and F.J. Seiler, "Optimization of parameters for semiempirical

methods II. Applications", *Computational Chemistry*, Vol.10, No. 2, pp. 221-264, 1989.

146. Krylov, A.I. and P.M.W. Gill, "Q-Chem: An engine for innovation", *Wiley Interdisciplinary Reviews: Computational Molecular Science*, Vol. 3, No. 3, pp. 317–326, 2013.

147. Stephens, P.J., F.J. Devlin, C.F. Chabalowski, and M.J. Frisch, "Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields", *The Journal of Physical Chemistry A*, Vol. 98, No. 45, pp.11623–11627, 1994.

148. Zhao, Y., N.E. Schultz, and D.G. Truhlar, "Design of Density Functionals by Combining the Method of Constraint Satisfaction with Parametrization for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions", *Journal of Chemical Theory and Computation*, Vol. 2, No. 2, pp. 364–382, 2006.

149. Chai J.dD. and M. Head-Gordon, "Long-Range Corrected Hybrid Density Functionals With Damped Atom-Atom Dispersion Corrections", *Physical Chemistry Chemical Physics*, Vol. 10, No. 44, pp. 6615–6620, 2008.

150. Barone, V. and M. Cossi, "Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model", *The Journal of Physical Chemistry A*, Vol. 102, No. 11, pp. 1995–2001, 1998.

151. Marenich, A.V., C.J. Cramer, and D.G. Truhlar, "Universal Solvation Model Based on Solute Electron Density and on A Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions", *Journal of Physical Chemistry B*, Vol. 113, No. 18, pp. 6378–6396, 2009.

152. Reed, A.E., R.B. Weinstock, and F. Weinhold, "Natural Population Analysis", *The Journal of Chemical Physics*, Vol. 83, No. 2, pp. 735–746, 1985.

153. Marenich, A.V., S.V. Jerome, C.J. Cramer, and D.G. Truhlar, "Charge Model 5: An Extension of Hirshfeld Population Analysis for The Accurate Description of Molecular Interactions In Gaseous and Condensed Phases", *Journal of Chemical Theory and Computation*, Vol. 8, No. 2, pp. 527–541, 2012.

154. Hirshfeld, F.L., "Theoretica Chimica Acta Bonded-Atom Fragments for Describing Molecular Charge Densities", *Springer*, Vol. 44, pp.129–138, 1977.

155. Ginak, A.I., K.A. V'yunov, V.V. Barmina, and E.G. Sochilin, "Original Article Submitted", *Soviet Physics-Doklady*, Vol. 7, pp.173–174, 1971.

156. Caine, B.A., C. Dardonville, and P.L.A. Popelier, "Prediction of Aqueous P K A Values for Guanidine-Containing Compounds Using Ab Initio Gas-Phase Equilibrium Bond Lengths", ACS Omega, Vol. 3, No. 4, pp. 3835–3850, 2018.

157. Kaljurand, I., T. Rodima, I. Leito, I. A. Koppel, and R. Schwesinger, "Self-Consistent Spectrophotometric Basicity Scale in Acetonitrile Covering the Range between Pyridine and DBU", *Journal of Organic Chemistry*, Vol. 65, No. 19, pp. 6202–6208, 2000.

158. Tuncel, S.T. and I. Dogan, "Synthesis of Thiazol-2-Imines from the Reduction of Single Enantiomer 2-Imino-Thiazolidin-4-Ones Followed by A Spontaneous Water Elimination", *Chirality*, Vol. 32, No. 6, pp. 866–875, 2020.

159. Tay, K.S. and N. Madehi, "Ozonation of Ofloxacin in Water: By-Products, Degradation Pathway and Ecotoxicity Assessment", *Science of the Total Environment*, Vol. 520, pp. 23–31, 2015.

APPENDIX A: HYDROXYL RADICAL-MEDIATED DEGRADATION OF SALICYLIC ACID, METHYLPARABEN AND CAFFEINE: A COMPUTATIONAL APPROACH TO ASSESS THE REACTION MECHANISMS

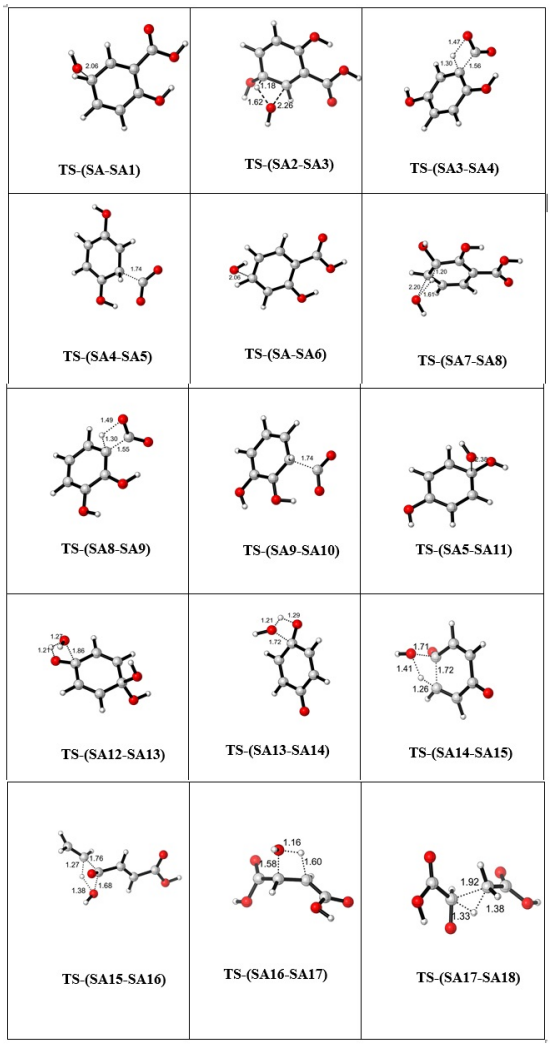


Figure A.1. 3D images of transition state geometries of SA.

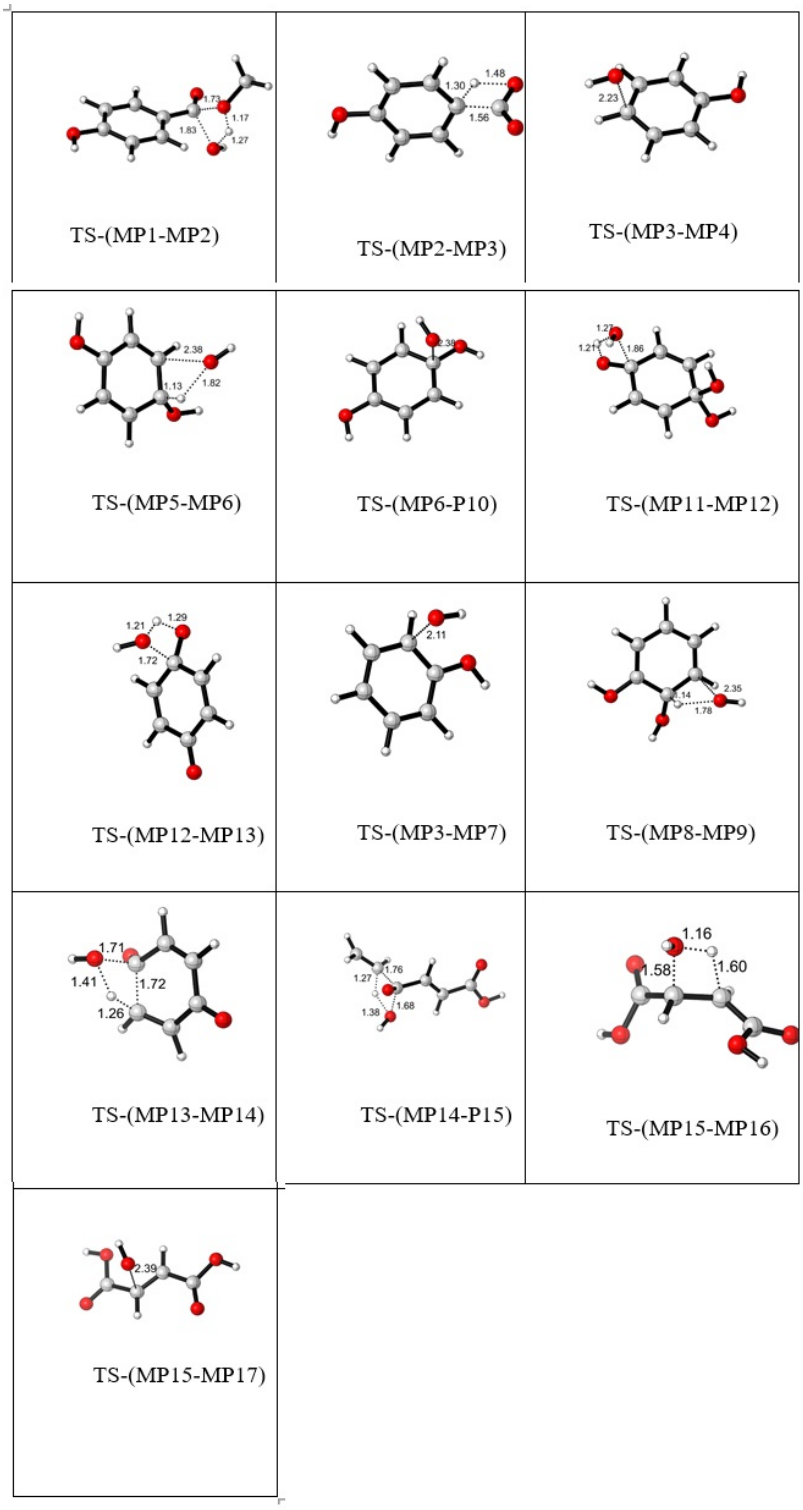


Figure A.2. 3D images of transition state geometries of MP.

Table A.1. Toxicity values of SA and MP with their degradation by-products to aquatic organisms (mg/L) using ECOSAR.

	Green algae EC_{50}	Daphnid LC_{50}	Fish LC_{50}
SA1	11.4	61.8	128
SA3	66.1	$1.34x10^3$	341
SA4	38.8	981	114
SA6	4.90	256	22.2
SA8	3.55	80.0	22.0
SA9	62.4	$2.89x10^3$	263
SA11	4.90	256	22.2
SA13	1.02	27.7	8.67
SA14	$2.30x10^3$	$5.86x10^3$	$4.66x10^3$
SA15	0.047	0.738	0.095
SA16	$3.50x10^3$	$7.29x10^3$	$7.94x10^3$
SA17	$8.96x10^3$	$2.53x10^4$	$5.33x10^4$
SA18	$1.65x10^5$	$9.01x10^5$	$2.23x10^6$
SA19	$4.40x10^3$	$1.23x10^4$	$2.58x10^4$
MP1	18.1	42.5	20.4
MP2	42.3	162	491
MP3	2.40	9.64	27.7
MP5	2.45	56.0	15.5
MP6	4.90	256	22.2
MP8	2.55	57.9	16.0
MP9	4.90	256	22.2
MP11	1.02	27.7	8.67
MP12	$2.30x10^3$	$5.86x10^3$	$4.66x10^3$
MP13	0.047	0.738	0.095
MP14	$3.50x10^3$	$7.29x10^3$	$7.49x10^3$
MP15	$8.96x10^3$	$2.53x10^4$	$5.33x10^4$
MP16	$1.65x10^5$	$9.01x10^{35}$	$2.23x10^6$
MP18	$6.24x10^4$	$2.63x10^{53}$	$6.12x10^5$

Table A.2. Toxicity classification according to the Globally Harmonized System (GHS) [162].

Toxicity range (mg/L)	Class
$LC_{50}/EC_{50} \leq 1$	Very toxic
$1 \prec LC_{50}/EC_{50} \leq 10$	Toxic
$10 \prec LC_{50}/EC_{50} \leq 100$	Harmful
$LC_{50}/EC_{50} \succ 100$	Not harmful

APPENDIX B: SAMPL BLIND CHALLENGES

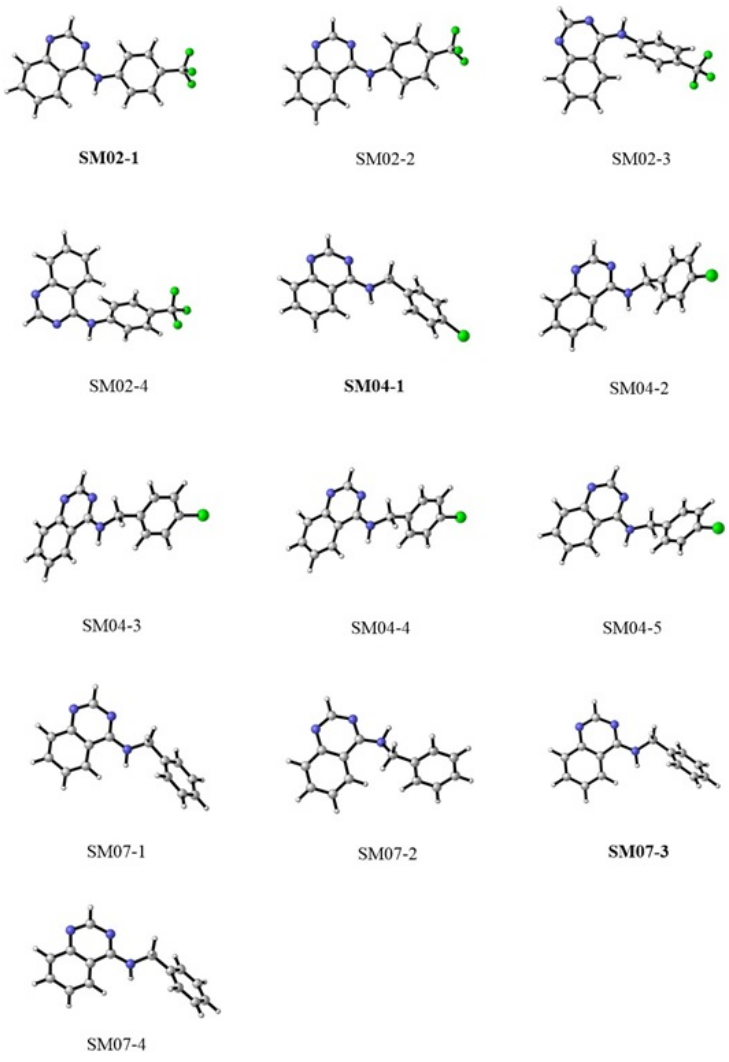


Figure B.1. Conformers of SM02, SM04 and SM07 (B3LYP/ 6-311+G**/ SMD/ water) (the most stable conformers are in bold) (Color Key: N:Blue, O:Red, Cl: Green, F:Green, C:Gray).

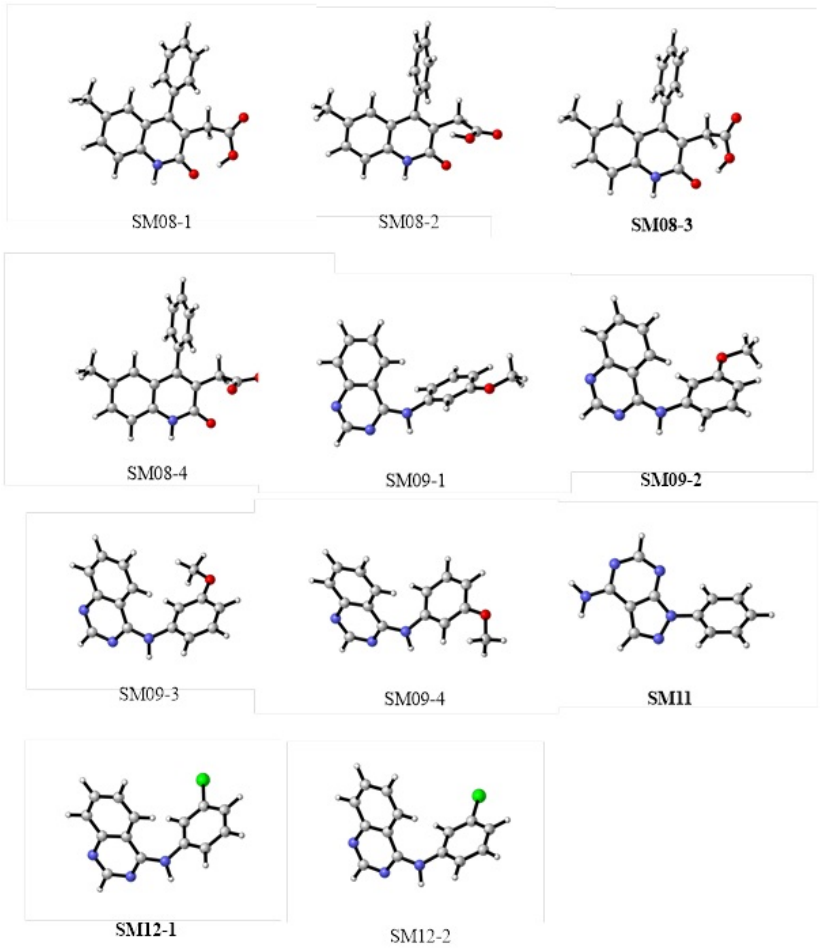


Figure B.2. Conformers of SM08, SM09, SM11 and SM12 (B3LYP/ 6-311+ G** /SMD/ water) (the most stable conformers are in bold) (Color Key: N:Blue, O:Red, Cl: Green, F:Green, C:Gray).

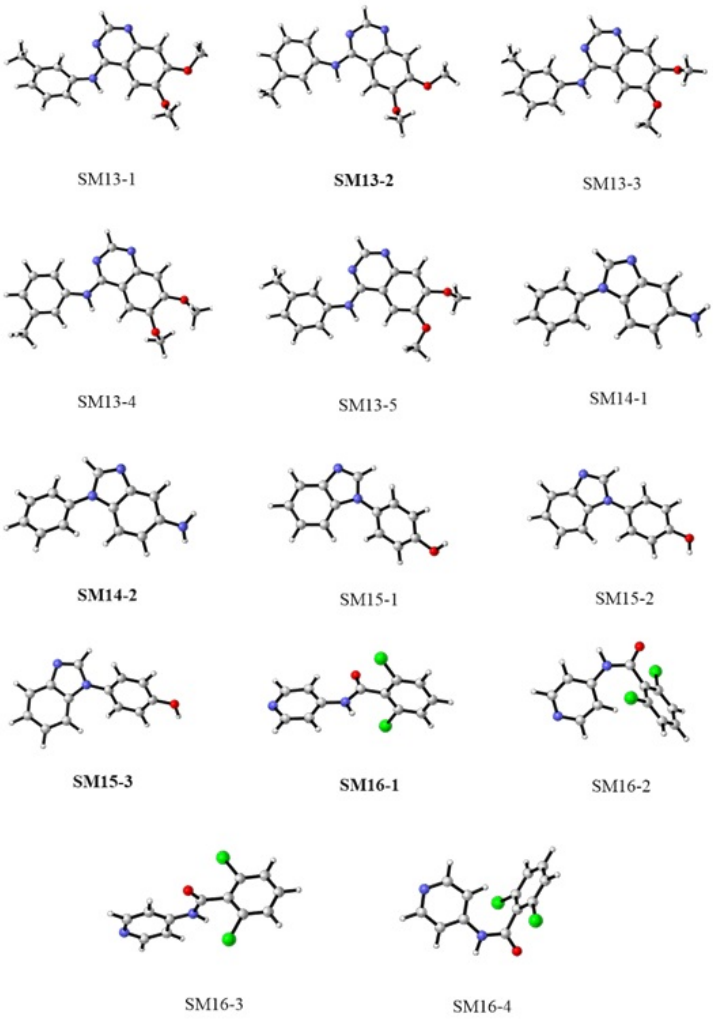


Figure B.3. Conformers of SM13, SM14, SM15 and SM16 (B3LYP/ 6-311+G**/ SMD/ water) (the most stable conformers are in bold) (Color Key: N:Blue, O:Red, Cl: Green, F:Green, C:Gray).

Table B.1. Relative Gibbs free energies of all the conformers of the molecules of interests (the most stable conformers are given in bold) (B3LYP/ 6-311+G**/ SMD/ water).

Molecules	**Relative Energies**	**Molecules**	**Relative Energies**
SM02-1	**0.00**	**SM11**	**0.00**
SM02-2	0.01	**SM12-1**	**0.00**
SM02-3	3.38	SM12-2	0.12
SM02-4	3.38	SM13-1	0.46
SM04-1	**0.00**	**SM13-2**	**0.00**
SM04-2	0.43	SM13-3	1.7
SM04-3	0.46	SM13-4	2.61
SM04-4	0.43	SM13-5	2.54
SM04-5	0.44	SM14-1	0.04
SM07-1	0.01	**SM14-2**	0.00
SM07-2	7.18	SM15-1	0.16
SM07-3	**0.00**	SM15-2	0.01
SM07-4	0.01	**SM15-3**	**0.00**
SM08-1	0.01	**SM16-1**	**0.00**
SM08-2	1.66	SM16-2	5.14
SM08-3	**0.00**	SM16-3	0.21
SM08-4	1.70	SM16-4	5.15
SM09-1	0.53		
SM09-2	**0.00**		
SM09-3	0.43		
SM09-4	0.42		

Table B.2. Relative Gibbs free energies (kcal/mol) of all conformers of the molecules of interests (the most stable conformers are bold) (B3LYP/ 6-311+G**/ SMD/ octanol).

Molecules	**Relative Energies**	**Molecules**	**Relative Energies**
SM02-1	0.01	**SM11**	**0.00**
SM02-2	**0.00**	**SM12-1**	**0.00**
SM02-3	3.94	SM12-2	0.01
SM02-4	3.94	**SM13-1**	**0.00**
SM04-1	**0.00**	SM13-2	0.02
SM04-2	0.75	SM13-3	1.7
SM04-3	0.65	SM13-4	2.75
SM04-4	0.37	SM13-5	1.90
SM04-5	1.11	**SM14-1**	**0.00**
SM07-1	0.00	SM14-2	0.07
SM07-2	5.66	**SM15-1**	**0.00**
SM07-3	**0.00**	SM15-2	0.49
SM07-4	0.12	SM15-3	0.49
SM08-1	0.01	SM16-1	0.11
SM08-2	3.08	SM16-2	5.16
SM08-3	**0.00**	**SM16-3**	**0.00**
SM08-4	2.27	SM16-4	5.15
SM09-1	0.36		
SM09-2	**0.00**		
SM09-3	0.19		
SM09-4	0.20		

Table B.3. Relative Gibbs free energies of all conformers of the molecules of interests (the most stable conformers are bold) (B3LYP/ 6-311+G**/ gas).

Molecules	**Relative Energies**	**Molecules**	**Relative Energies**
SM02-1	0.01	**SM11**	**0.00**
SM02-2	**0.00**	**SM12-1**	**0.00**
SM02-3	3.81	SM12-2	0.01
SM02-4	3.81	**SM13-1**	**0.00**
SM04-1	0.31	SM13-2	0.32
SM04-2	**0.00**	SM13-3	1.62
SM04-3	0.27	SM13-4	2.74
SM04-4	0.14	SM13-5	1.48
SM04-5	0.06	SM14-1	0.03
SM07-1	0.01	**SM14-2**	**0.00**
SM07-2	6.39	**SM15-1**	**0.00**
SM07-3	0.01	SM15-2	0.01
SM07-4	**0.00**	SM15-3	0.01
SM08-1	0.01	**SM16-1**	**0.00**
SM08-2	0.01	SM16-2	4.08
SM08-3	**0.00**	SM16-3	0.09
SM08-4	**0.00**	SM16-4	4.00
SM09-1	0.22		
SM09-2	0.28		
SM09-3	**0.00**		
SM09-4	0.17		

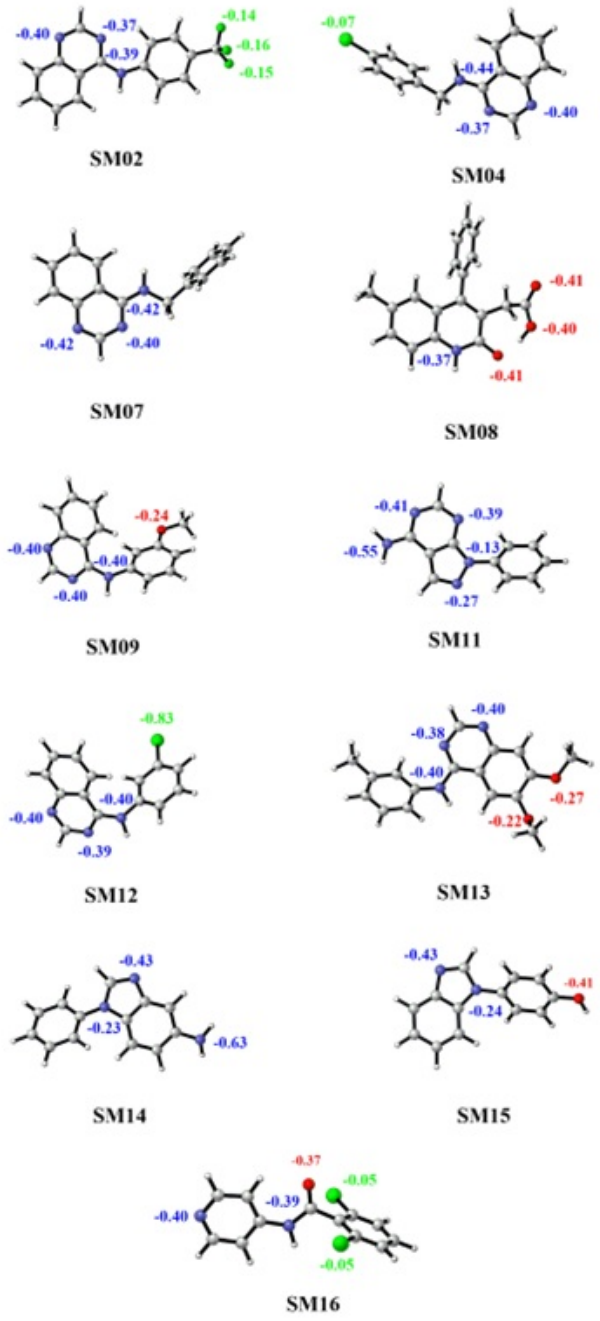


Figure B.4. 3D geometries and the calculated charge values (CM5) on the heteroatoms of the most stable conformer of each molecule. (Atomic charges are in au.) (Color Key: N:Blue, O:Red, Cl: Green, F:Green, C:Gray) (B3LYP/ 6-311+G**/ SMD/ water).

Table B.4. Calculated and experimental LogP values
(B3LYP/6-311+G**/SMD/water, with one explicit water molecule).

Molecules	**Calculated**	**Experimental**
SM02	4.14	4.09
SM04	4.06	3.98
SM07	2.81	3.21
SM08	3.24	3.1
SM09	3.51	3.03
SM11	2.09	2.1
SM12	3.93	3.83
SM13	2.96	2.92
SM14	1.88	1.95
SM15	3.03	3.07
SM16	3.05	2.62

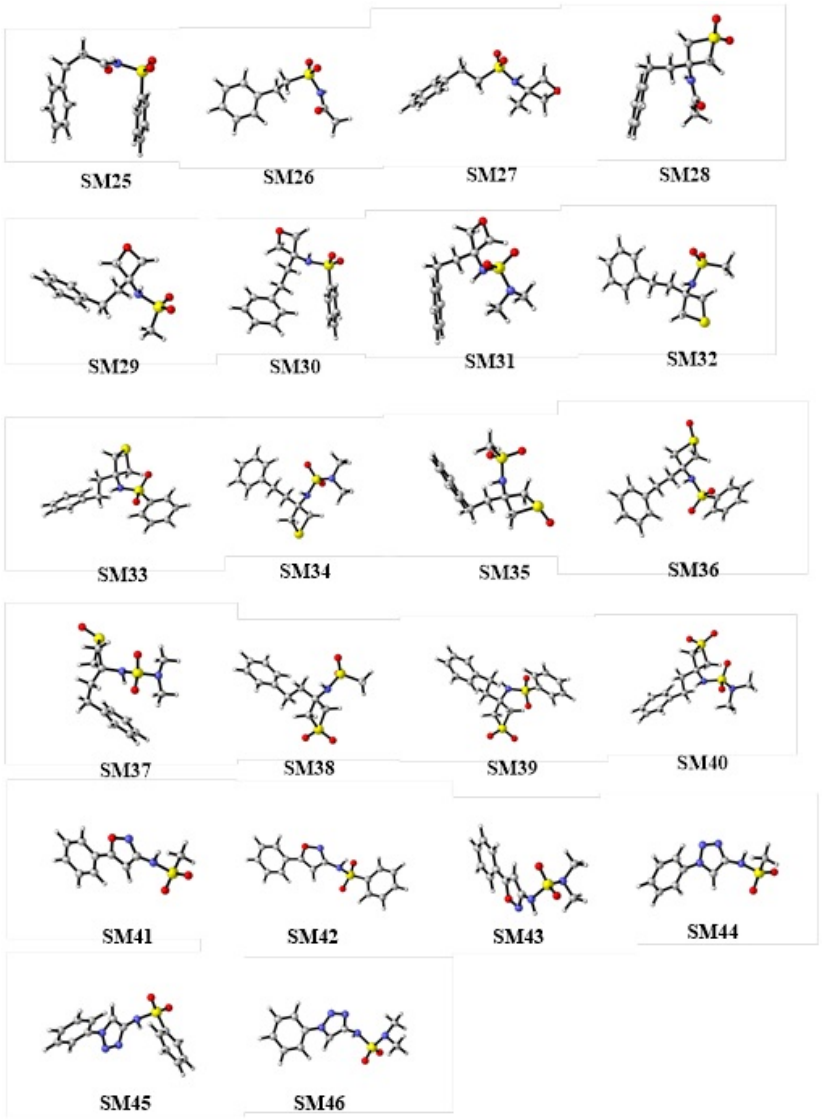


Figure B.5. Global minima structures of N-acylsulfonamides and sulfonamides derivatives (B3LYP- D3/ 6-311+ G(d,p)/ gas).

APPENDIX C: QUANTUM-MECHANICAL PREDICTION OF DISSOCIATION CONSTANTS FOR THIAZOL-2-IMINE DERIVATIVES

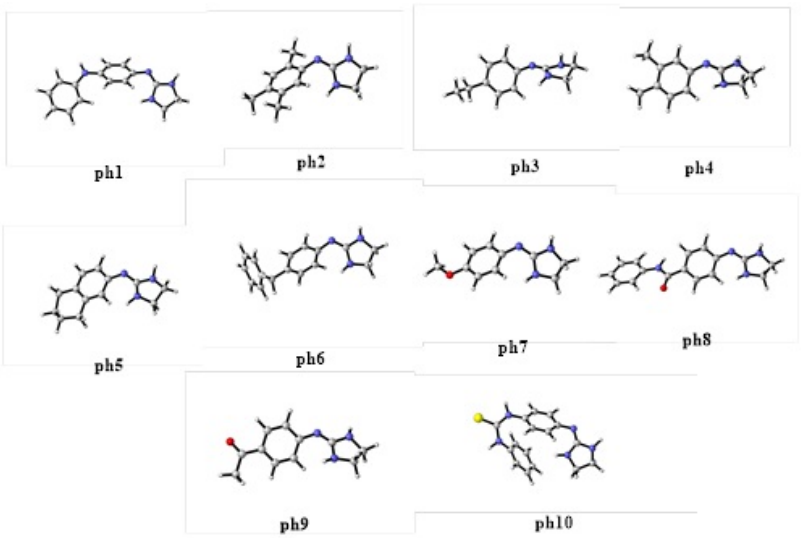


Figure C.1. 3D representations of 2- (phenylimino) imidazolidine derivatives (M062X/ 6-31G**/ SMD/ water).

Table C.1. Calculated and experimental water pK_a's of 2-(phenylimino)imidazolidine derivatives (M062X/ 6-31G**/ SMD/ water).

ID	**Experimantal** pK_a	**Calculated** pK_a	$\triangle pK_a$
ph1	10.49	10.34	-0.15
ph2	10.29	10.20	-0.09
ph3	10.42	10.37	-0.05
ph4	10.50	10.42	-0.08
ph5	10.44	10.57	0.13
ph6	10.78	10.85	0.07
ph7	10.62	10.52	-0.10
ph8	10.17	10.00	-0.17
ph9	9.11	9.32	0.21
ph10	9.08	9.24	0.16
RMSE			0.13
MAD			0.12
MD			-0.01

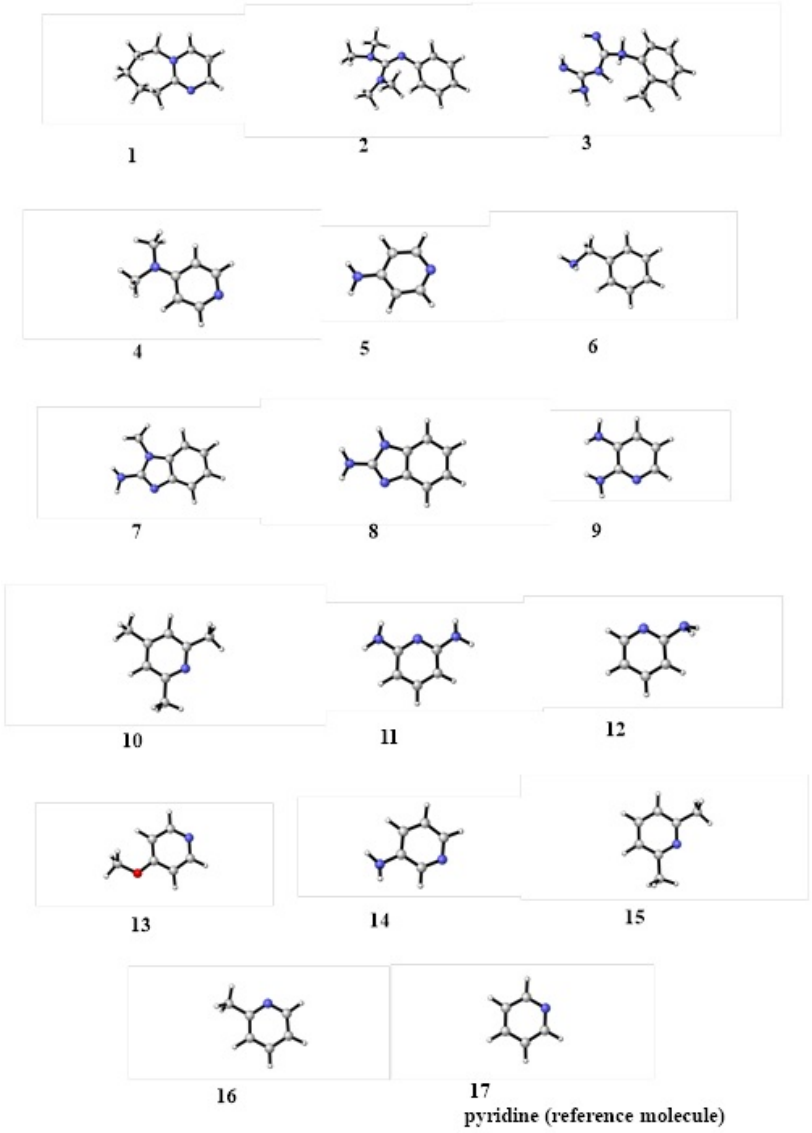


Figure C.2. 3D representations of nitrogen containing small aromatic compounds (M062X/ 6-31G**/ SMD/ MeCN).

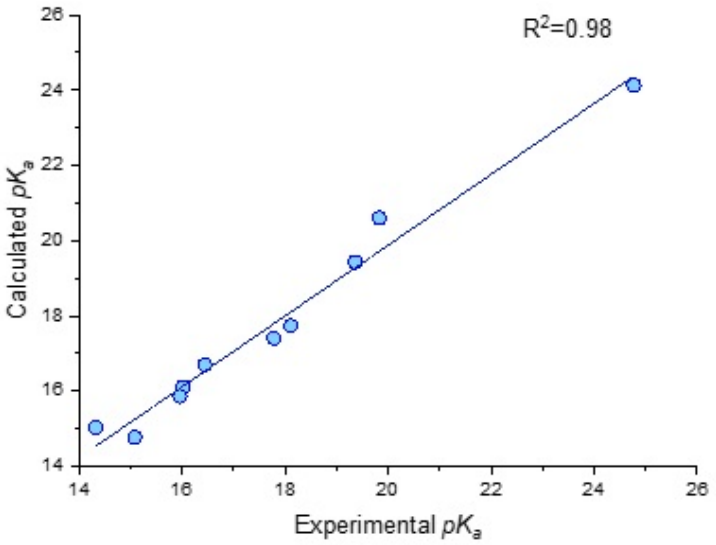


Figure C.3. Linear regression of experimental vs calculated pK_a values of nitrogen containing small aromatic compounds (M062X/ 6-31G**/ SMD/ MeCN).

APPENDIX D: COPYRIGHT LICENSES

This Agreement between Bogazici University – Evrim Arslan (''You'') and Elsevier consists of your license details and the terms anad conditions provided by Elsevier and Copyright Clearance Center.

License Number	5458650958918
License date	Dec 30, 2022
License Content Publisher	Elsevier
License Content Publication	Journal of Environmental Management
License Content Title	Life cycle environmental impacts of advanced wastewater treatment techniques for removal of pharmaceutical and personal care products (PPCPs)
Licensed Content Author	Raphael Ricardo Zepon Tarpani, Adisa Azapagic
Licensed Content Date	Jun 1, 2018
Licensed Content Volume	215
Licensed Content Issue	n/a
Licensed Content Date	15
End Page	272
Type of Use	reuse in a thesis/dissertation
Portions	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both electronic and printing
Are you the author of this Elsevier article?	No
Will you be translating?	Yes, including English rights
Number of languages	1
Title	MODELING THE DEGRADATION OF ORGANIC POLLUTANTS AND THE SOLUBILITY OF DRUG-LIKE MOLECULES
Institution name	Bogazici University

Figure D.1. Permission from [3] Elsevier, Copyright (2018)..

This Agreement between Bogazici University – Evrim Arslan (''You'') and Elsevier consists of your license details and the terms anad conditions provided by Elsevier and Copyright Clearance Center.

License Number	5460990738433
License date	Jan 02, 2023
License Content Publisher	Elsevier
License Content Publication	Elsevier Books
License Content Title	Advanced Oxidation Processes for Effluent Treatment Plants
Licensed Content Author	Q.Q.Chai,L. Jothinathan,S.H. Deng,S.L. Ong,H.Y.Ng, J.H.Yu
Licensed Content Date	Jun 1, 2021
Licensed Content Pages	56
Start Page	199
End Page	254
Type of Use	reuse in a thesis/dissertation
Portions	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both electronic and printing
Are you the author of this Elsevier article?	No
Will you be translating?	Yes, including English rights
Number of languages	1
Title	MODELING THE DEGRADATION OF ORGANIC POLLUTANTS AND THE SOLUBILITY OF DRUG-LIKE MOLECULES
Institution name	Bogazici University

Figure D.2. Permission from [13] Elsevier, Copyright (2021).