

Sveučilište u Zagrebu

FAKULTET KEMIJSKOG INŽENJERSTVA I TEHNOLOGIJE

Anamarija Pulitika

INTERAKCIJE MIKROPLASTIKE I ORGANSKIH ONEČIŠĆIVALA U VODI

DOKTORSKI RAD

Zagreb, 2025



University of Zagreb

FACULTY OF CHEMICAL ENGINEERING AND TECHNOLOGY

Anamarija Pulitika

INTERACTIONS OF MICROPLASTICS AND ORGANIC POLLUTANTS IN WATER

DOCTORAL THESIS

Supervisors:

prof. Ana Lončarić Božić, Ph.D.

Panaghiotis Karamanis, Ph.D.

Zagreb, 2025



Sveučilište u Zagrebu

FAKULTET KEMIJSKOG INŽENJERSTVA I TEHNOLOGIJE

Anamarija Pulitika

INTERAKCIJE MIKROPLASTIKE I ORGANSKIH ONEČIŠĆIVALA U VODI

DOKTORSKI RAD

Mentori:

prof. dr. sc. Ana Lončarić Božić

dr. sc. Panaghiotis Karamanis

Zagreb, 2025

Bibliographic data:

- UDK:
- Scientific area: Technical sciences
- Scientific field: Chemical engineering
- Scientific branch: Environmental protection in chemical engineering
- Institution: University of Zagreb, Faculty of Chemical Engineering and Technology,
 Department of Polymer Engineering and Organic Chemical Technology
- Supervisors: prof. Ana Lončarić Božić, Ph.D.

Panaghiotis Karamanis, Ph.D.

- Number of pages:
- Number of figures:
- Number of tables:
- Number of literature references:
- Defence date:
- Defence jury:

- Dissertation is stored in:

The topic of the dissertation was approved at the 257th session of the Faculty Council of the Faculty of Chemical Engineering and Technology, University of Zagreb on July 11th 2022, and approved at the 1st session of the Senate of the University of Zagreb on October 25th 2022.



This research is supported by Croatian Science Foundation: Young Researchers' Career Development Project – Training New Doctoral Students; Microplastics in water; fate and behavior and removal (Project IP-2020-02-6033; **ReMiCRO**)

A part of the research was performed using the computational resources provided by computer cluster Isabella based in SRCE – University of Zagreb, University Computing Centre and computational resources provided by "Direction du Numérique" of the "Université de Pau et des Pays de l'Adour".

Abstract

Nowadays, microplastics (MP), plastic particles smaller than 5 mm, are a hot topic in environmental science field due to their abundance and widespread in the environment and food chain, including human organism. The greatest concern is potential harmful and toxic effect MP can have on both the environment and human health. Additionally, MP can adsorb and transport other pollutants, possibly increasing overall toxicity compared to the toxicity of isolated MP and pollutants. Hence, it is important to understand the adsorption process to properly assess the potential risk of MP. Experimental results show that adsorption is mainly governed by weak hydrophobic, electrostatic, van der Waals and π - π interactions, and hydrogen bonds. However, due to system complexity and the many influencing factors, the results often lack consistency and reproducibility, which makes it impossible to draw clear conclusions and build a predictive model for MP adsorption. The computational chemistry studies on the adsorption of MP are still at the early stages, but they are believed to provide further insights into the adsorption mechanism.

In this thesis, polyethylene terephthalate (PET) MP was chosen as one of the most frequently found types of MP in the environment. Adsorption was studied using quantum mechanical (QM) and molecular dynamics (MD) methods. QM study revealed the interactions are weak and local, mainly dependent on the pollutants ability to form multiple contacts with MP. Further quantitative structure activity/property relationship (QSA/PR) modelling showed the key pollutants' properties governing the adsorption can be correlated to the size of pollutants and number of proton donor/acceptor functional groups. On the other hand, MD study showed that, when the explicit water molecules are included, the adsorption is mainly controlled by the pollutant's relative affinity for water and MP. Additionally, the affinity also depends on the structure of MP, as more fragmented model showed stronger capacity compared to unfragmented MP model. MD results showed similar trend as the experimental results. However, the difference in calculated interaction energies is could not be quantitatively correlated with the difference in adsorption capacities. Overall, computational studies showed a great potential as a tool to help to understand adsorption mechanism, although more work is needed to further develop the model's accuracy.

Keywords: adsorption, DFT, MD, microplastic, polyethylene terephthalate, QSA/PR

Sažetak

Mikroplastika (MP) su čestice polimera manje od 5 mm koje su danas vrlo popularna tema u istraživanju okoliša zbog velike količine i rasprostranjenosti u okolišu i hranidbenom lancu, pa čak i u ljudskom organizmu. MP može imati štetne i toksične učinke na okoliš i ljudsko zdravlje. Osim toga, MP može adsorbirati i prenositi druga štetna onečišćivala što može dovesti do povećanja njihove ukupne toksičnosti u odnosu na toksičnost same MP i onečišćivala. Upravo iz tog razloga je važno detaljno proučiti proces adsorpcije kako bi se moglo preciznije procijeniti potencijalne rizike povezane s MP. Dosadašnja eksperimentalna istraživanja su pokazala da adsorpciju organskih onečišćivala na MP kontroliraju slabe interakcije poput hidrofobnih, elektrostatskih, van der Waalsovih, π - π interakcija i vodikovih veza. S obzirom na složenost sustava i veliki broj čimbenika koji utječu na adsorpciju, rezultati iz literature su često nedosljedni te nisu reproducibilni. To otežava donošenje zaključaka i izradu pouzdanih predikcijskih modela za adsorpciju organskih onečišćivala na MP. Metode računalne kemije omogućuju dublje razumijevanje mehanizma adsorpcije te se sve više počinju koristiti u ovom području istraživanja.

U okviru ovog rada, odabran je poli(etilen-tereftalat) (PET), kao jedan od najčešćih tipova MP u okolišu. Za istraživanje korištene su kvantno-mehaničke (QM) i metode molekularne dinamike (MD). QM proračuni pokazali su da su interakcije između MP i onečišćivača slabe, lokalne interakcije te da adsorpcija ovisi o broju interakcija koje onečišćivalo može uspostaviti s MP. Modeliranjem odnosa strukture i aktivnosti/svojstva (QSA/PR) utvrđeno je da su ključna svojstva onečišćivala koja kontroliraju adsorpciju, povezana s veličinom molekule i brojem proton donora/akceptora. MD simulacije koje uključuju eksplicitne molekule vode su pokazale da adsorpcija značajno ovisi o razlici između afiniteta onečišćivala za molekule vode i za MP. Također su pokazale da struktura MP ima značajnu ulogu s obzirom na to da je fragmentirani model PET-a pokazao veću sposobnost adsorpcije od nefragmentiranog. Rezultati MD simulacija u skladu su s eksperimentalnim opažanjima, iako razlike u izračunatim energijama interakcije nisu bile u izravnoj kvantitativnoj korelaciji s razlikama u adsorpcijskom kapacitetu. Zaključno, računalne metode pokazuju velik potencijal u istraživanju adsorpcijskih mehanizama MP, no potrebna su dodatna istraživanja i razvoj modela kako bi se povećala njihova točnost i primjenjivost.

Ključne riječi: adsorpcija, DFT, MD, mikroplastika, poli(etilen-tereftalat), QSA/PR

Contents

1.	Introduc	ction	1		
2. Literature review					
2	2.1. The	e world of (micro)plastic	3		
	2.1.1.	Microplastic as contaminants of emerging concern	6		
2	2.2. Ads	sorption of organic pollutants on microplastic	8		
	2.2.1.	Factors influencing microplastics adsorption capacity	9		
	2.2.2.	Experimental approach to study the adsorption	13		
	2.2.3.	Computational approach to study the adsorption	17		
2	2.3. The	eoretical background	19		
	2.3.1.	Basics of quantum chemistry	19		
	2.3.2.	Density Functional Theory (DFT)	25		
	2.3.3.	Quantum chemistry calculations	30		
	2.3.4.	Molecular dynamic simulations	34		
	2.3.5.	QSA/PR mathematical modelling	46		
3.	Methods	S	55		
	8.1. Cor	mputational details	55		
	3.1.1.	DFT calculations	55		
	3.1.2.	QSA/PR formulation	55		
	3.1.3.	Molecular dynamic simulations	56		
	3.2. Exp	perimental details	56		
	3.2.1.	Materials	56		
	3.2.2.	Microplastic preparation	56		
	3.2.3.	Adsorption experiment	57		
4.	Results a	and discussion	59		

4.1. At	o initio study on the role of functional groups in PET MP-pollutants interact	ctions59
4.1.1.	Calculation of weak interactions	61
4.1.2.	Method performance	62
4.1.3.	PET model adsorption sites and size effects	67
4.1.4.	Molecular interactions	70
4.1.5.	Visual study of weak interactions	73
4.1.6.	The effect of folded particle model	79
4.1.7.	Synergistic and ageing effect	80
4.2. QS	SA/PR models of pollutants' functional groups and PET MP interaction end	ergies84
4.2.1.	Preparing dataset for QSA/PR modelling	
4.2.2.	Developing QSA/PR models	
4.2.3.	Model descriptors	94
4.3. Ad	dsorption of pollutants on PET MP in water	
4.3.1.	Force field	
4.3.2.	MD simulation of pesticides in water	
4.3.3.	Construction of PET models for MD simulations	111
4.3.4.	MD simulation of pesticides' adsorption on model PET	
4.3.5.	Adsorption capacities	117
5. Conclu	sion	119
Appendices		
Appendix	A	
Appendix	В	
Appendix	C	151
List of Fig	gures	
List of Ta	bles	
List of Ab	breviations	

1. Introduction

Microplastics (MP) are synthetic polymer particles smaller than 5 mm in diameter^[1]. They can be intentionally produced as microbeads (primary MP) or they can be derived from the fragmentation of larger plastic waste (secondary MP). MP have a low degradation rate and are very abundant, persistent, and widespread in the environment. By now, they have been found in every environmental compartment,^[2–4] including the most distant places as the north and south poles^[5,6]. The most found types of MP in the environment are polyethylene (PE), polypropylene (PP), polyethylene terephthalate (PET), polystyrene (PS) and polyvinyl chloride (PVC). Owing to their small size, MPs can easily be inhaled, ingested, or even absorbed by tissue or roots and thus enter the food chain and further propagate.^[7–10] Once in the organism, they can cause a series of negative effects including intestinal damage, liver necrosis and death.^[11–13] MPs have already been found in human blood, lungs and placenta^[14–16] proving they can penetrate human tissue. Due to the potential risks they may pose to both environment and human health, MP are now considered a new class of contaminants of emerging concern (CECs)^[4] and a lot of research is aimed to better understand the fate and behaviour of MPs in the environment.

Another concerning factor is MP ability to adsorb other contaminants from the environment and to transport them within the ecosystem.^[17,18] Co-contamination of MP with organic pollutants might increase the overall toxicity compared to the toxicity of isolated MP and pollutants.^[19–21] Therefore, to properly evaluate the potential risk of MP in the environment it is important to understand the adsorption process. Currently, a lot of studies aim to understand the complex mechanism of MP adsorption. So far, they have shown the adsorption of organic contaminants by MP is physical and governed by weak intermolecular forces.^[22] However, there are many factors controlling adsorption behaviour making it hard to differentiate the individual effects to properly describe adsorption mechanism. Additionally, there is no consistent way to design the experiment and interpret the results which leads to a lot of variations in results and draws to different conclusion. In order to better clarify the adsorption on molecular level, computational chemistry are more often used. Although still at its early stages, studies incorporating quantum chemistry and molecular dynamics simulations showed promising results in helping to understand the adsorption mechanism of MP. For that reason, it is important to expend the research of MP's adsorption in the field of computational chemistry. Therefore, this thesis aimed to integrate computational chemistry approaches, including quantum mechanics and molecular dynamics, into the study of MP adsorption mechanism. The focus of quantum mechanics study was to address the nature of interaction between MP and organic pollutants, and to understand the role of specific functional groups governing the sorption mechanism. The study also seeks to correlate the computational approach with experimental results to improve the interpretation of adsorption behaviour. Additionally, to develop cost-effective predictive methodologies to reduce the need for extensive experimental work.

2. Literature review

2.1. The world of (micro)plastic

Since its discovery, plastic has effectively replaced many materials due to its remarkable properties, including low density, inertness, impermeability, corrosion resistance, and excellent electrical and thermal insulation. Above all, plastic's ease of production and low cost have contributed significantly to its widespread use. Mass production of plastic began in 1950 and has continued ever since. In 2020, the demand for plastic in Europe alone reached 49.1 million tons, increasing to 58.7 million tons in 2022, which represents only 15% of the global plastic demand.^[23] The most common types of polymers used in plastic are high-density polyethylene (HDPE), low-density polyethylene (LDPE), polypropylene (PP), polyvinyl chloride (PVC), polystyrene (PS), and polyethylene terephthalate (PET) whose monomeric units are shown in Figure 2.1. The exponential growth of plastic production since 1950 to 2022 and the demand for these raw plastic materials by polymer type in world is illustrate in Figure 2.2, according to market research and statistics from Plastics Europe.^[23]



Figure 2.1. Structures of monomers of the most produced polymers.

The exponential growth in plastic production is naturally followed by the exponential growth in plastic waste. Plastics are generally characterised by low degradability, which can last for centuries. This characteristic poses an ecological problem as it leads to an accumulation of

plastic waste in the environment. Although plastic is very resistant, the environmental stressors, including chemical, biological and physical stressors can lead to the deterioration of structural properties and increase the fragility and brittleness of the material. Finally leading to the fragmentation of plastic to smaller particles of micro size called microplastics.



Figure 2.2. The exponential growth of the plastic production and the global demand for the raw plastic material according to the type of polymer in 2022, according to the market research from Plastic Europe.^[23]

Microplastics (MP) are defined as synthetic polymer particles with a diameter of less than 5 millimetres. Although the lower limit of microplastic size is not clearly defined, particles smaller than 1000 nm are commonly referred to as nanoplastics (NP).^[24] Sources of MP can be divided to primary MP and secondary MP. Primary MP are those already produced in the form of microparticles to be used in personal care products, especially in exfoliants, detergents, as a fertiliser additive, for medical devices, as a blasting agent, in adhesives, as a raw material to produce other plastic products and others. In 2019, the European Chemicals Agency (ECHA) presented a proposal to limit the production of primary MP and in September 2023, the European Commission adopted the restrictions of primary MP intentionally added to products under the European Union chemical legislation REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals).^[25] he mayor source of MP in the environment are secondary MP.

Secondary MP result from the fragmentation of larger plastics and represent a serious environmental problem due to the number and variety of possible sources. Fragmentation can result from mechanical damage during the manufacture and use of plastic products or from the effects of the environment on plastic waste.^[26] As MP are small and light, they can be easily transported by wind, rain, rivers, oceans, etc. The largest amount of plastic particles found consisted of PE, PP, PS, PET and PVC. Traces of MP have been discovered even in the most remote places in the world, from the top of the Himalayas^[27] and the Pic du Midi^[28] in the Pyrenees to the depths of the Mariana Trench.^[29] Materić et al.^[5] even detected MP in samples of Antarctic ice in the south to Greenland ice in the north, taken at a depth of 14 metres.

To date, MP has been detected in all components of the environment, including air, soil, freshwater and oceans.^[30] The main sources of airborne microplastics are synthetic fibres, car tyres and urban dust.^[30] The concentration of MP in outdoor air is generally lower than the concentration of MP indoors, which could be due to the release of plastic particles from indoor furniture, textiles, packaging materials or other indoor sources, as well as a slower removal rate. However, a systematic review by Eberhard et al.^[31] shows that outdoor exposure is highly dependent on the location. By far the highest concentrations of MP are therefore found on roadsides and in urban areas with high traffic volumes. Significant amounts of MP are found in soil, particularly in agricultural soils, where the main sources are plastics residues from mulching, municipal waste, plastic-coated fertilisers and sewage sludge.^[32] Once in the soil, MP can contaminate food, increasing the likelihood of human exposure. It can also come into contact with the pesticides used, which can alter their toxicity and impact on soil and biota. The sources of microplastics in freshwater and oceans are numerous, from direct fragmentation of waste to runoff from landfills, the air, wastewater treatment plants and more. The concentration of microplastics in freshwater and oceans depends on the region, human activities, wind and ocean currents.^[30,33,34] MP have been detected in tap water and bottled water, showing that MP can be transferred from natural sources despite drinking water treatment plants. A review by Gambino et al.^[35] shows that the concentration of MP in bottled water is higher than in tap water, suggesting that there may be additional contamination from the packaging process.

2.1.1. Microplastic as contaminants of emerging concern

There is growing evidence of the presence of microplastics and their harmful effects on the environment, but also on organisms and humans. As concern about microplastics grows, so does the number of studies on the fate, behaviour and toxicity of microplastics in the environment (Figure 2.3).^[36]



Figure 2.3. Number of published papers per year based on a search of keywords microplastic* OR nanoplastic* AND adsorption in Scopus database.

Most research focusses on microplastics in inland waters and oceans. Fish, crabs, aquatic mammals and other aquatic organisms can ingest MP.^[37–40] MP have also been detected in various terrestrial animals, both wild,^[41–43] and domestic animals,^[7,44,45] proving that MP are present in the food chain and affect the entire food web.^[46] In addition, traces of MP have also been found in fruits and vegetables that we eat, where it most likely enters the through direct absorption from the water and soil via the roots.^[9,47–52] Humans are also exposed to MP, which raises the question of its potential health risk. MP can enter the human organism either by ingestion or inhalation. Several studies have already detected MP in lung tissue, especially PE, PP and PET MP.^[15,53] MP are found in human faeces, as the majority of ingested MP are excreted from the organism via the gastrointestinal tract.^[54,55] Nevertheless, a study by Leslie et al.^[14] showed the presence of MP

traces in human blood. PET was the most common type of MP found in blood samples, followed by PS and PE in a size range between 700 nm and 500 μ m. The maximum concentrations were 2.4 μ g mL⁻¹ for PET, 4.8 μ g mL⁻¹ for PS and 7.1 μ g mL⁻¹ for PE. This study proves that MP can be transported through the bloodstream in the organism and reach other organs, with the highest MP concentration found in the colon and liver.^[56] A study by Ragusa et al.^[16] was the first to detect traces of MP in the placenta. Since then, further research has shown the presence of different types of MP in the placenta^[50,57] and in breast milk.^[58] These recent discoveries have raised serious concerns about the effects of MP on foetal development and infant exposure to MP.

The presence of MP in the organism can potentially have a toxic effect. Studies on freshwater zebrafish (lat. *Danio rerio*) have shown that the presence of microplastics can lead to intestinal and liver damage, oxidative stress and disruption of reproductive function, which can ultimately lead to death.^[11,13] MP also have cytotoxic effect. Rubio et al.^[59] studied the effect of PS MP on various human haematopoietic cell lines and showed that PS MP can induce intracellular oxidative stress and DNA damage. The cytotoxicity of MP varies depending on the size, type and dose of MP, type of cells, exposure time and presence of co-contaminants.^[60] Studies to date are limited to *in vivo* studies with model organisms and *in vitro* studies with different cell types. However, there is no evidence of a direct effect of MP on human health,^[61] but due to the increasing evidence of the presence of plastic in humans and its potentially harmful effects on the environment and organisms, MP is now considered a contaminant of emerging concern.

Another factor of concern is the co-contamination of MP with other environmental pollutants, microorganisms or heavy metals.^[62] MP may contain additives that are intentionally added during the production of the plastic through the refining process. The additives can be transported together with MP and released into the environment, leading to a toxic effect.^[63,64] In addition to the additives, MP can also adsorb other environmental pollutants such as heavy metals^[65] and organic pollutants^[66] on its surface. Studies show that the combined effect of MP with additives or adsorbed pollutants can lead to an increase^[19,21] or decrease in their overall toxicity.^[62] The co-contamination of MP with additives and environmental pollutants depends on the mechanism of their adsorption on the MP surface and desorption. To better understand the joint effect of MP and additives or adsorbed pollutants and to assess the potential risk of MP, it is therefore crucial to understand the adsorption mechanism. The state of the art on the adsorption mechanism of MP and organic pollutants is presented in the following Section 2.2.

2.2. Adsorption of organic pollutants on microplastic

Previous studies have shown that the adsorption of organic pollutants on microplastics is a weak physical sorption controlled by weak intermolecular interactions, mainly hydrophobic interactions, hydrogen bonding, electrostatic attraction and repulsion forces, van der Waals forces and π - π interactions (Figure 2.4).^[22,67] However, the extent of each interaction and its contribution to the mechanism of adsorption is not yet known, so further research is needed in this area.

Hydrophobic interactions are one of the most important mechanisms influencing the adsorption of hydrophobic organic pollutants on MPs in water. The hydrophobicity of an organic molecule is usually expressed as an octanol/water partition coefficient (K_{ow}). A higher value means a higher hydrophobicity. The adsorption capacity of MP for hydrophobic pollutants showed a positive correlation with the log K_{ow} value in several studies.^[68,69] In a study by Li et al.,^[70] hydrophobic antibiotics had a higher affinity for PP, PS, PE and PVC MP, but not for polyamide (PA) MP. PA MP is the only one of the materials tested that has an amide group, and its higher affinity for more polar antibiotics is attributed to its ability to form hydrogen bonds.

Hydrogen bonds are weak interactions in which a hydrogen atom is located between two electronegative atoms. Several other studies have shown the importance of hydrogen bonding, as PA MP has a significantly higher adsorption capacity for polar antibiotics such as amoxicillin, tetracycline, ciprofloxacin,^[70] sulfamethoxazole,^[71] and sulfamethazine^[72] compared to other types of plastics.

Electrostatic interactions are interactions between two charged molecules. Under natural environmental conditions, the surface of MP generally carries a net negative charge.^[70,73–75] Therefore, the repulsive electrostatic forces can act between MP and negatively charged organic pollutants and the attractive forces can contribute to the adsorption of positively charged organic pollutants. The adsorption of charged pollutants depends on the pH value of the water. In the adsorption of negatively charged triclosan on PVC, the adsorption capacity of PVC decreased drastically with increasing pH value.^[76] The same phenomenon was observed in the adsorption of chlorophenol on PET MP, proving that the process was controlled by electrostatic repulsion.^[77] On the other hand, pH value had no significant effect on the adsorption of neutral diethyl and dibutyl phthalate on MP, as no electrostatic interactions were involved.^[78]

Van der Waals forces are weak nonspecific intermolecular forces. Xu et al.^[73] showed that the adsorption of hydrophilic sulfamethoxazole on PE MPs was linear, albeit they were both negatively charged under the experimental conditions. Neither hydrophobic nor electrostatic interactions could explain the sorption which is likely governed by van der Waals interactions. Other studies also confirm that van der Waals forces may play an important role in the sorption mechanism.^[72,79]

 π - π interactions are a special type of dispersion force that occur between unsaturated benzene rings or polycyclic molecules. The higher adsorption of pollutants such as diethyl and dibutyl phthalate on PS MP compared to other types of MP is associated with their ability to form π - π interactions.^[68,78]



Figure 2.4. Interactions of MP and organic pollutants in water.

2.2.1. Factors influencing microplastics adsorption capacity

One of the major reasons why there is still a lot of questions concerning adsorption mechanism of microplastic, albeit the abundance of current research, is the complexity of the system and the large number of factors that can affect adsorption mechanism.^[80,81] The mechanism depends on the physicochemical properties of both microplastic and organic pollutants but also on environmental factors.

As mentioned before, hydrophobicity of pollutants is very important factor for the adsorption of organic pollutants onto microplastics. Li et at. proved the hydrophobic interactions are the main mechanism for the adsorption on PE, PP and PS.^[82] However, hydrophilic compound can also be adsorbed by other types of interactions. The adsorption of ionizable compound depends on pH value of the solution. Positively charged pollutants will be more readily adsorbed than negatively charged as the surface of MP is negative under the natural environmental conditions. ^[73,74] The ability to form hydrogen bond will favour the adsorption onto hydrogen donor and acceptor groups on the surface of microplastic.^[70–72]

The adsorption capacity of MP, its behaviour in the environment and the potential threat is largely dependent on the type of polymer in MP and its chemical and structural properties.^[26] The most commonly used types of plastics in adsorption studies are PE, PS, PVC, PP and PET.^[80] Some of their properties are listed in Table 2.1. Polymers are long-chain molecules consisting of many repeating organic units (monomers) connected by covalent bonds.^[83] The chains in the structure are held together by weak interactions, such as hydrogen bonds, van der Waals interactions, hydrophilic or hydrophobic interactions.

Polymers as PE, PP and PET have semi-crystalline structure. Semi-crystalline structure means their structure contains crystalline regions, where segments of chains are properly arranged, emersed in amorphous region, where chains are oriented randomly. When polymers are heated, they can transform from solid (glassy) state to more flexible rubbery state. The temperature of this transformation is called glass transition temperature (T_g) .^[84] PE and PP have low glass transition temperature, so they are soft and flexible at room temperature. In contrast, PET, PVC, and PS with high glass transition temperature are rigid. The state of the polymer might affect adsorption capacity as polymers in a rubbery state are more flexible.^[85] However, crystallinity and rubbery state are not essential factors.^[70] Size of the MP and the porosity of surface are very important factors.^[86] Smaller particles and particles with more developed pore structure have larger specific surface area (SSA) and, therefore, more adsorption sites available. Ma et al.^[76] showed the adsorption capacity of smaller PVC MP was 12.7 mg g^{-1} , while the adsorption capacity of larger PVC MP was 8.98 mg g^{-1} for triclosan. Nonpolar polymers such as PE, PP and PS are strongly hydrophobic, and they showed larger adsorption capacity for hydrophobic 17β-estradiol than polar PVC. On the contrast, adsorption capacity of highly polar PA for 17β -estradiol was the highest which is contributed to the formation of hydrogen bonds.^[87]

Name	Abb.	$T_{\rm g}$ / °C	$T_{\rm m}$ / °C	ho / g cm ⁻¹	χ/%	polarity
High-density polyethylene	HDPE	-110	130	0.96	60-90	nonpolar
Low-density polyethylene	LDPE	-100	110	0.92	30-50	nonpolar
Polypropylene	PP	-25	170	0.90	30-50	nonpolar
Polystyrene	PS	95	240	1.04	0	nonpolar
Polyvinyl chloride	PVC	81	227	1.39	0	polar
Polyethylene terephthalate	PET	70	240	1.33	10-30	polar

Table 2.1. Physicochemical properties of the most used types of plastics. T_g is the glass transition temperature, T_m is the temperature of melting, ρ is the average density and χ is the percent crystallinity.^[88]

MPs are very persistent in the environment because they are highly resistant to the degradation process. However, various external factors can still cause ageing of MPs by changing their physicochemical properties and consequently adsorption behaviour. The most important process of MPs degradation is photooxidation.^[89] Absorption of UV light causes chain scission and formation of free radicals, following the reaction with oxygen. The reaction is autocatalytic and can further self-propagate leading to the formation of oxygen containing functional groups and linking of chains.^[90] General mechanism of polymer photooxidation is shown in Figure 2.5. Carbonyl groups can act as chromophores through Norrish-type 1 and Norrish-type 2 mechanism. Norrish type 1 leads to the chain scission and the formation of carbon monoxide, while Norrish type 2 reaction leads to the formation of vinyl groups and aldehydes (Figure 2.6).^[91,92]

Initiation:	Polymer $\rightarrow R \cdot + R \cdot$
Propagation:	$R + O_2 \rightarrow ROO$
	$ROO + RH \rightarrow ROOH + R \cdot$
Chain branching:	$ROOH \rightarrow RO \cdot + OH \cdot$
	$RH+OH \cdot \rightarrow R \cdot + H_2O$
	$RO \cdot \rightarrow Chain-scission$
Termination:	$ROO \cdot + ROO \cdot \rightarrow Cross-linking$
	$ROO \cdot + R \cdot \rightarrow Cross-linking$
	$R \cdot + R \cdot \rightarrow Cross-linking$

Figure 2.5.General mechanism of photooxidation.



Figure 2.6. Mechanism of polymer photooxidation: a) chain-scission and propagation; b) Norrish type I; c) Norrish type II mechanism.^[93]

The surface of aged MP is more fragile with more visible pores and cracks^[94] and larger SSA with more adsorption sites available.^[95,96] FTIR spectra of PP, HDPE and LDPE showed wider peaks in the regions 3300-3500 cm⁻¹ and 1610-1800 cm⁻¹ after UV irradiation which corresponds to the formation of hydroxyl carbonyl and vinyl groups.^[97] Wu et al.^[98] showed a linear correlation of oxygen/carbon surface ratio (O/C ratio) and the duration of UV irradiation of PP MPs. More oxygen containing groups make MP more hydrophilic which favours the adsorption of hydrophilic contaminants.^[99] Oxidation of PET, PA and PP MPs increased their hydrophilicity and the adsorption capacity for sulfamethoxazole by 187.8%, 64.6% and 198% respectively.^[100] Adsorption of antibiotics amoxicillin and chlortetracycline on the surface of tire wear particles and PP MPs was also higher after ageing.^[101] Physicochemical changes caused by ageing affect the mechanism of adsorption. In the study by Yao et al.,^[102] the main interactions controlling adsorption of pharmaceuticals and personal care products on pristine PP MPs were hydrophobic interactions, while hydrogen bonds and electrostatic interactions controlled the adsorption on chemically oxidized PP. As ageing impacts the adsorption mechanism it also influences the potential risk of MP in the environment. Therefore, it is important to further investigate adsorption behaviour of aged MP.^[103]

The adsorption mechanism of MPs for organic pollutant depends on the type of environment and various environmental factors.^[82,104] Li et al.^[70] showed the adsorbed amount of five antibiotics on five types of MPs was higher in freshwater than in seawater. Lorca et al.^[105] also concluded the adsorption was more favoured in freshwater. Higher concentration of ions might compete for the sorption site and thus reduce the adsorption capacity.^[99] On the other hand, increased salinity might increase the sorption capacity by lowering the solubility of organic pollutants ("salting out")^[76,98] thus increasing the adsorption capacity in seawater. Zhang et al.^[95] found out the adsorption amount of oxytetracycline was higher in the presence of CaCl₂ than of NaCl or Na₂SO₄ which is probably due to the ability of Ca²⁺ ions to form ternary complex between oxytetracycline and MPs and enhance adsorption. Dissolved organic matter (DOM) present in natural water might also have a significant effect on the adsorption. Mo et al.^[106] showed the presence of DOM has a negative effect on the adsorption of carbofuran on PE which can be attributed to the higher affinity of DOM to pesticides. On the other hand, presence of humic acid promoted the adsorption of oxytetracycline on PS MP probably by acting as a bridge.^[95] The change of pH value can affect dissociation of organic pollutants.^[76,107] Consequently, adsorption capacity of dissociable pollutants will be lower at higher pH value due to electrostatic repulsions between negatively charged MP and pollutants. The change of pH value does not affect the adsorption of non-ionizable compounds.^[78]

2.2.2. Experimental approach to study the adsorption

Adsorption is a surface process in which a molecule of solute or gas (adsorbate) is transferred from a liquid to a solid surface (adsorbent). The adsorption can be chemical (chemisorption) or physical (physisorption). Chemisorption leads to the formation of new chemical bonds between the adsorbent and the adsorbate. In physisorption, the adsorbate and the adsorbent interact without bonding.^[108] The amount of adsorbate adsorbed in a solution is calculated as follows:

$$q_{\rm e} = \frac{c_0 - c_{\rm e}}{m} \times V \tag{2.1}$$

where $q_e \text{ [mg g}^{-1]}$ is the equilibrium adsorption capacity, $c_0 \text{ [mg dm}^{-3]}$ is the initial concentration of the adsorbate in a solution, $c_e \text{ [mg dm}^{-3]}$ is the final concentration of the adsorbate in a solution, m [mg] is the mass of the adsorbent, and V [dm³] is the volume of a solution. The adsorption behaviour of microplastics and organic pollutants is generally studied experimentally by fitting the data of the adsorption capacity to the kinetic or isothermal models. The way in which the data fit the different models provides information about the possible nature and mechanism of adsorption.

Adsorption kinetics describe the rate of an adsorption process by analysing the change in the amount of adsorbate adsorbed by the adsorbent over time. If the amount of adsorbate adsorbed is constant over time, the adsorption process has reached a state of equilibrium. The study of adsorption kinetics is the first step in determining the time required for the process to reach equilibrium under the experimental conditions. In a study by Raznajatovo et al.,^[69] adsorption to PE reached equilibrium after approximately 96 hours for sertraline, 24 hours for propranolol, and 3 hours for sulfamethoxazole. However, in a study by Xu et al.,^[73] 24 hours were required for the adsorption of sulfamethoxazole on PE MP to reach equilibrium. The different results show the influence of the experimental conditions on the sorption kinetics and the importance of preliminary kinetic investigations for estimating the adsorption equilibrium time. The study of adsorption kinetics can give us an insight into the efficiency of adsorption, the adsorption rate and the rate-limiting step. In most studies, the data are only fitted to pseudo-first order (PFO) and pseudo-second order (PSO) kinetic models derived from the following equation:

$$v = \frac{dq_{\rm t}}{dt} = k_1 (q_{\rm e} - q_{\rm t})^n \tag{2.2}$$

where $v \text{ [mg g}^{-1} \text{ s}^{-1}\text{]}$ is the adsorption rate, $q_e \text{ [mg g}^{-1}\text{]}$ is the equilibrium adsorption capacity, q_t [mg g $^{-1}$] is the amount of adsorbate adsorbed by the unit mass of adsorbent in a time t [s], n is the order of the mechanism, and $k \text{ [(g mg}^{-1)n s}^{-1}\text{]}$ is the rate constant. In general, the experimental data are transformed and fitted by linear regression methods. However, caution is needed as this approach might favour the PFO model over the PSO model, which is not the case when a non-linear approach is used.^[109] Moreover, J. Simonin^[110] has pointed out that in most studies, kinetic data near equilibrium or at equilibrium are included in the models, which favours the linear PSO model over the linear PFO model and might lead to incorrect conclusions. In this sense, Wang et Wang^[85] showed that the adsorption of pyrene on PE, PS and PVC MPs before reaching equilibrium agrees well with the PFO model. However, the results fitted PSO better when the data from equilibrium were also fitted. A good fit to PSO is sometimes incorrectly associated with chemisorption,^[96,111,112] such as in a study by Hu et al.^[113] in which the adsorption of 17 β -estradiol on MP is labelled as chemical due to the better fit of the data to PSO than to PFO. However, it is worth noting that PSO alone cannot indicate chemisorption without additional experimental evidence, such as from FTIR and XPS measurements.^[76,99] PFO and PSO are used to determine the

adsorption rate, but do not provide information about the mechanism of adsorption. To understand the mechanism and determine the rate-controlling step, the intraparticle diffusion (IP) model is often used. In most studies, the IP model suggests that adsorption on MP is controlled by both intraparticle and film diffusion processes.^[72,76,77,95] In the study by Wang et al.,^[114] the adsorption mechanism of tetracycline on virgin PE MP was also controlled by both intraparticle and film diffusion. However, when PE MP was exposed to environmental factors and a biofilm developed on its surface, the mechanism changed, and film diffusion was the only rate-limiting step.

The adsorption isotherm describes the relationship between the adsorption capacity and the concentration of the adsorbate in a state of equilibrium at a constant temperature and a constant pH value (q_e vs. c_0). Isotherms are used to predict the amount of adsorbate that can be adsorbed onto an adsorbent and the adsorption equilibrium constant. They can also help to elucidate the mechanism of adsorption. There are many isotherm models that have been formulated either theoretically or empirically to fit the experimental adsorption data. The Langmuir, Freundlich and linear isotherms are the most commonly used models to fit the adsorption in a single layer on a homogeneous surface. It assumes that all adsorption sites are uniform and the adsorbates are evenly distributed between them. It also assumes that there are no interactions between two adsorbates, so that adsorption can only take place on the surface of the adsorbent in a monolayer.^[116] The Langmuir adsorption is as follows:

$$q_{\rm e} = \frac{Q_{\rm max} K_{\rm L} c_{\rm e}}{1 + K_{\rm L} c_{\rm e}} \tag{2.3}$$

where $q_e \text{[mg g^{-1}]}$ is the equilibrium sorption capacity, $Q_{\text{max}} \text{[mg g^{-1}]}$ is the maximum adsorption capacity, K_{L} [dm³ mg⁻¹] is the Langmuir adsorption constant and c_e [mg dm⁻³] is the equilibrium concentration of the adsorbate. Freundlich adsorption model^[117] does not assume an ideal surface of the adsorbent. The surface is heterogeneous, i.e. the adsorption sites on the surface can have different affinities to the adsorbates. Since interactions between two adsorbates are possible, the model is not limited to single-layer adsorption and can also be used to describe multi-layer adsorption. Freundlich adsorption equation is represented as follows:

$$q_{\rm e} = K_{\rm F} c_{\rm e}^{1/n} \tag{2.4}$$

where $K_{\rm F}$ [dm³ mg⁻¹]^{*n*} is the adsorption constant, and 1/*n* is the adsorption intensity. The adsorption intensity, 1/*n*, represents the non-linearity factor. It can take values between 0 and 1.

Lower 1/n value indicates a more heterogeneous surface. Values above 1 indicate unfavourable sorption. The linear model is similar to the Freundlich model, in which the adsorption intensity is equal to 1. In the linear adsorption model, the adsorption capacity increases proportionally with the increasing concentration of the adsorbate.^[118] The linear isotherm equation is expressed as follows:

$$q_{\rm e} = K_{\rm d} c_{\rm e} \tag{2.5}$$

where K_d [dm³ mg⁻¹] is the linear adsorption constant. This model can describe the adsorption on an uniform surface at sufficiently low concentration when the adsorbates cannot feel one another.^[119] The linearity of adsorption varies amongst different types of polymers and pollutants.^[79] The adsorption isotherm of sulfamethazine on PA, PE, PVC, PS and PET MP showed a good fit to the linear model.^[72] The linear model also described well the adsorption of amoxicillin and sulfadiazine on the same five types of MPs, while the nonlinear Freundlich model was a better fit for the adsorption of tetracycline, ciprofloxacin and trimethoprim.^[70] Wang et al.^[75] compared the adsorption isotherms of tetracycline on virgin and aged PE MPs and showed that the Freundlich model was the best to describe the adsorption by virgin and air-exposed MPs, while the Langmuir model was better at describing the adsorption by water and soil-exposed MP. Dubinin-Radushkevich and Temkin^[120,121] isotherms are also used in studies on the microplastic adsorption capacity. The Dubinin-Radushkevich model is used to explain adsorption by pore-filling mechanism.^[122–124] It can estimate the mean free energy and distinguish the type of adsorption (physisorption or chemisorption). Li et al.^[125] used Dubinin-Radushkevich to calculate the mean adsorption energy for diclofenac and PS MPs. The adsorption energy obtained was 241 kJ mol⁻¹ which was attributed to chemisorption. However, this study neglects the fact that the Dubinin-Radushkevich model was developed for gas/solid interaction and ignores the effect of solvent, chemical species of the solutes, surface charge and other factors affecting solute/solid adsorption. Therefore, the results may not be accurate.^[123] The computation study by Cortés-Arriagada et al.^[126] suggests that the high mean adsorption energy is due to several pairwise interactions and is incorrectly associated with chemisorption.

As there is no standardised way to study the adsorption of organic pollutants on MP, the experimental results are still subject to high variations. For example, the K_d values of sulfamethoxazole adsorption on PE in three different studies were 700 dm³ kg⁻¹,^[69] 592 dm³ kg⁻¹,^[73] and 30 dm³ kg⁻¹.^[71] High variation in results, the different experimental designs, the use of

different models from study to study and the different interpretations of results make it difficult, almost impossible, to compare the results and draw clear conclusion.

2.2.3. Computational approach to study the adsorption

Theoretical research into the adsorption behaviour of microplastics is still at an early stage. Computational research can help to clarify the experimental results and gain a deeper understanding of the adsorption process at the molecular level, which is why research in the field of computational chemistry must be expanded. Density Functional Theory (DFT) methods are now increasingly used in combination with experimental results to understand the electronic properties of organic pollutants and to find the correlation with the results obtained or to directly calculate the interaction energies with MP and correlate them with the adsorption capacity. However, the methodology and the way in which the initial structures for the calculations are set up are still questionable. The combined experimental and DFT studies mostly represent the MP only as an oligomeric chain.^[106,127-131] Although some studies show that the adsorption capacities and the calculated interaction energies are correlated in this type of approach, the research of Cortes-Arriagada et al.^[132,133] showed that this approach strongly neglects the importance of the dispersion interaction in favour of the electrostatic interactions. Their research also showed that the organic pollutants can form multiple interactions with neighbouring parts of the MP, all of which contribute to the total interaction energy.^[126,134,135] In this sense, the nano- or microparticle model of MP should be used instead of a simple oligomeric chain.

The dynamics of adsorption can be studied by molecular dynamics (MD). Similar to DFT studies, MD is now mostly used in combination with experimental results to correlate the interaction energies with the adsorption capacity.^[136–139] However, MD simulations can contribute much more to the understanding of the mechanism of MP adsorption and the effects of the environment.^[140] The study by Feng et al.^[141] showed that the adsorption of benzo[α]pyrene is higher for PS-MP due to its ability to form sandwich-like π -stack structures. It was also shown that humic acid, which occurs naturally in water, can adsorb to the surface of MP and interact with benzo[α]pyrene. This prevents the direct interaction between MP and benzo[α]pyrene. Liu et al.^[142] showed that the adsorption capacity of MP in seawater is greater for aromatic carbonates than in pure water, which is consistent with the experimental results. The simulation showed that this is probably due to the more porous structure of MP in seawater than in pure water, where the pores

can serve as additional adsorption sites. Furthermore, adsorption in seawater can be additionally enhanced by cations, which could act as a bridge between MP and aromatic carbonates. MD has proven to be a very promising tool to study adsorption on MP and although the number of studies using it is still very small, it is increasing.

The number of organic pollutants whose adsorption on microplastics can be analysed is very large. Therefore, it is very useful to develop efficient mathematical models that are able to predict the adsorption of new, untested pollutants as QSA/PR (Quantitative Structure Activity/Property Relationship) models. QSA/PR modelling correlates the structural properties of molecules with their activity or property. To do this, the data obtained from experimental studies or calculations must be consistent, which is difficult to achieve as there are no standardised methods for collecting and validating the results. Nevertheless, some attempts have been made. The most common models use the linear free energy relationship (LFER) or the linear solvation energy relationship (LSER) to predict the equilibrium distribution coefficient ($\log K_d$) for the adsorption of pollutants on MP.^[82,143-146] Other adsorption-related parameters can also be modelled, such as the research by Zhang et al.^[147] who successfully developed models to predict the maximum adsorption capacity and adsorption affinity of organic compounds on PE, PET, PS and PVC MP. In order to develop further models that can predict the adsorption of a variety of organic pollutants and incorporate the effects of the types of MP and environmental effect, much more experimental data is needed, which once again underlines the importance of developing standardised methods. On the other hand, the tools of computational chemistry can be used to collect data on molecular interactions that would be difficult and time-consuming to obtain through experimental research alone. With this in mind, some QSA/PR models have been successfully developed that correlate the calculated interaction energies between compounds with the structural properties.^[148,149] Further details on the theoretical background of computational chemistry tools and modelling methods are described in section 2.3.

2.3. Theoretical background

2.3.1. Basics of quantum chemistry

The fundamental postulate of quantum mechanics states that every chemical system is completely represented by its wave function $\Psi(r, t)$ and that observables are represented by appropriate operators that return the value of a corresponding observable by acting on a system wave function.^[150] The wave function $\Psi(r, t)$ is the solution of the time-dependent Schrödinger equation:

$$\widehat{H}\Psi(r,t) = i\hbar \frac{\partial \Psi(r,t)}{\partial t}$$
(2.6)

where \hat{H} is the Hamiltonian operator and \hbar is Planck's constant. The product of a wave function with its complex conjugate $|\Psi(r,t)|^2$ is the probability density for finding the particle at position r, at time t. For a system with constant potential energy, the wave function can be separated to position dependent variable $\Psi(r)$, and time-dependent variable $\varphi(t)$. In this case, the probability density for finding a particle at position r does not depend on time $(|\Psi(r,t)|^2 = |\Psi(r)|^2)$, nor does any other expected value of observables. Therefore, the system is in its stationary state.^[151]

2.3.1.1. Time independent Schrödinger equation

Quantum chemistry methods use the principles and equations of quantum mechanics to find an approximate solution of non-relativistic, time-independent Schrödinger equation (stationary state):

$$\widehat{H}\Psi(r) = E\Psi(r) \tag{2.7}$$

where \hat{H} is Hamiltonian operator for a molecular system in the absence of magnetic or electric field, $\Psi(r)$ is a time-independent wave function and *E* is the total energy of the system.^[152] The Hamiltonian is the operator of total energy, which typically considers five contributions:

$$\widehat{H} = \widehat{T}_e + \widehat{T}_n + \widehat{V}_{ne} + \widehat{V}_{ee} + \widehat{V}_{nn}$$
(2.8)

the kinetic energy of electrons (\hat{T}_e) , the kinetic energy of nuclei (\hat{T}_n) , the potential energy of interactions between nuclei and electrons (\hat{V}_{ne}) , the potential energy of electron-electron repulsion (\hat{V}_{ee}) and nuclei-nuclei repulsion (\hat{V}_{nn}) . The upper equation is expressed as:

$$\widehat{\mathbf{H}} = -\sum_{i} \frac{\hbar^2}{2m_{\rm e}} \nabla_i^2 - \sum_{k} \frac{\hbar^2}{2m_{\rm k}} \nabla_k^2 - \sum_{i} \sum_{k} \frac{e^2 Z_k}{r_{ik}} + \sum_{i < j} \frac{e^2}{r_{ij}} \sum_{k < l} \frac{e^2 Z_k Z_l}{r_{kl}}$$
(2.9)

where *i* and *j* run over electrons, *k* and *l* over all nuclei, m_e is the mass of electron, m_k is the mass of nuclei, *Z* is the atomic number, *r* is the distance between two particles, \hbar is Planck's constant, and ∇^2 is the Laplacian.

The repulsion and attraction terms in Hamiltonian suggest that the particles do not move independently, i.e. their motion is correlated. For that reason, it is not possible to separate the variables and derive an analytical solution of Schrödinger equation for a system that contains more than one particle. Therefore, the Schrödinger equation of a many-body system is solved using approximative and numerical methods.

2.3.1.2. Born-Oppenheimer Approximation

The most important approximation in quantum mechanics is Born-Oppenheimer approximation that separates the motion of nuclei from the motion of electrons. It relies on a fact that one proton weight about 1800 times more than an electron, so the kinetic energy of a nuclei is negligible compared to the kinetic energy of an electron.^[152] Therefore, in scope of the Born-Oppenheimer approximation, the position of nuclei is fixed, and nuclei-nuclei repulsion term becomes a simple constant for a given geometry. The electronic time-independent Schrödinger equation than can be written as:

$$(\widehat{\mathbf{H}}_{el} + \widehat{\mathbf{V}}_{nn})\Psi_{el}(\mathbf{q}_i; \mathbf{q}_k) = E_{el}\Psi_{el}(\mathbf{q}_i; \mathbf{q}_k)$$
(2.10)

where q_i are electronic coordinates, q_k are nuclear coordinates parameters, E_{el} is electronic energy, and \hat{H}_{el} is the sum of the electrons' kinetic energies and potential energies of nuclei-electron attractions and electron-electron repulsions, as:

$$\widehat{H}_{el} = \widehat{T}_e + \widehat{V}_{ne} + \widehat{V}_{ee}$$
(2.11)

The Born-Oppenheimer approximation is one of the fundamental approximations in computational chemistry that significantly simplifies the Schrödinger equation and introduces the concept of potential energy surface (PES).

PES is a surface determined by electronic energy (E_{el}) over all possible nuclear coordinates (geometries). In computational chemistry, one of the main goals is to determine stable geometries of a system and transitions states that connect them. This can be done by analysis of PES and its stationary points, i.e. minima and maxima on potential energy surface.^[153] Stable geometries are represented by minima of energy. Small changes in geometry of minima will increase potential energy of a system. The lowest energy path connecting two local minima is called reaction coordinate and the maximum on reaction coordinate is a transition state between these two minima.

2.3.1.3. Variational principle

The Born-Oppenheimer approximation allows us to analytically solve Schrödinger equation for systems containing one electron, such as hydrogen atom and helium cation. For system with more than one electron, the approximative numerical methods are needed. One of those methods is variational method based on the variational principle. The variational principle states that the energy calculated as the expectation value of the Hamiltonian operator (\hat{H}) from any trial wavefunction (Ψ_{Trial}) must be greater than or equal to the true ground-state energy (E_0), i.e.:

$$\frac{\int \Psi_{\text{Trial}}^* \widehat{H} \Psi_{\text{Trial}}}{\int \Psi_{\text{Trial}}^* \Psi_{\text{Trial}}} \ge E_0 \tag{2.12}$$

This principle enables comparison of trial wave functions by their associated energies following the rule: the lower the better.^[152]

2.3.1.4. Basis set

The variational principle allows the use of any function that seems reasonable as trial wave function. Generally, the trial wave function is presented as linear combination of arbitrary functions:

$$\Psi_{\rm Trial} = \sum_{i} c_i \phi_i \tag{2.13}$$

where ϕ_i are arbitrary functions, and c_i are coefficients chosen in a way to minimise the energy. The set of mathematical functions ϕ_i used to construct the wave function is called basis set. In theory, the best wave function in a limit of chosen method is achieved by using an infinite basis set. However, in practice, this is impossible as computational cost increases with the number of basis functions used. In the light of computational efficiency, it is desirable to have a smaller number of basis set functions that are chemically reasonable, i.e. have a small amplitude in a region where the probability density is low, and larger in a region where the probability density is high, and that it can calculate the wave function in a computationally efficient way.^[154]

Slater type orbitals (STOs) are very attractive type of basis functions as they closely resemble hydrogenic atomic orbitals whose radial wave function has a cusp at the nucleus and exponential decay at larger distances from nucleus. Unfortunately, their use is quite limited as they require the use of numerical integration which becomes quite compilated for a system of any significant size. Gaussian type orbitals (GTOs) are another type of basis functions that can analytically solve high order integrals and therefore are more computationally efficient than STOs.

However, they do not have a cusp at r=0 and they decay too quickly at larger distances from nucleus. To overcome these limitations of both STOs and GTOs, most of the basis set functions are built as a linear combination of GTOs fit to reproduce the shape of STOs. A basis function defined as linear combination of GTOs is called contracted GTO (cGTO) and each GTO is called primitive function.^[155]

The size of a basis set is a very important feature. Basis set with smaller number of basis function are computationally more efficient, but the resulting wave function is less accurate compared to the wave function produced with more basis functions. The minimal basis, also called single- ζ basis set defines one basis function for each orbital. The basis set with two functionals for each orbital is called double- ζ basis set, with three functions triple- ζ , and so on. One way of adding basis functions is by decontracting basis set, that is by constructing more basis functions using the same number of primitive GTOs. In general, valence orbitals are significantly affected by chemical bonding as opposed to core orbitals. For that reason, it is more important to have more flexibility in valence basis functions than core functions. This led to the development of split-valence basis sets in which valence orbitals are represented by more basis functions while core orbitals are represented only by a single basis function.

Basis set that contains only functions cantered on the atoms lacks the flexibility to efficiently describe the molecular orbitals of a system with multiple nuclei. This flexibility is almost always increased by adding polarization function, that is by adding function that correspond to the orbitals with one quantum number higher angular momentum than the valance orbital. Thus, p-functions are added to polarize s-functions for hydrogen and helium, d-functions to polarize p-functions, and f-functions to polarize d-functions in heavy atoms. Lossely bound electrons that are found in anions and excited states, are on the average at larger distance from nuclei than core or bonding electrons. In that case diffuse functions are added to basis set to increase the electron density at larger distance from nuclei. Diffuse functions are Gaussian functions with smaller α value which causes slower descent of $e^{-\alpha r^2}$ with the distance r from the nuclei.^[156]

2.3.1.5. Hartree-Fock method

The Schrödinger equation for a system with more than one electron cannot be solved analytically because of the electron-electron repulsion term(s) are unknown. One way to approximate these terms is to separate the motion of electrons and solve a set of one electron wave function where electron move in an average potential from all the other electrons.^[152] This approximation is a basis

of Hartree-Fock (HF) method that is the simplest ab initio method that laid foundations for all others, more advanced methods. In HF method, the wave function on N electrons is represented as one Slater determinant (SD):

$$\Phi_{\rm SD} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(1) & \chi_2(1) & \dots & \chi_N(1) \\ \chi_1(2) & \chi_2(2) & \dots & \chi_N(2) \\ \vdots & \vdots & \ddots & \dots \\ \chi_1(N) & \chi_2(N) & \dots & \chi_N(N) \end{vmatrix}$$
(2.14)

where χ is a spin-orbital and $1/\sqrt{N!}$ is a normalization factor. Spin-orbital is a product of spatial orbital and electron spin eigenfunction chosen to be orthonormal for computational convenience:

$$\langle \chi_i | \chi_j \rangle = \delta_{ij} \tag{2.15}$$

Each row in SD represents one electron, while each column represents a unique spin-orbital where the electron could be found. In SD there are no more than one electron in each spin-orbital which follows the Pauli exclusion principle that forbids having two electrons in a system with all the same quantum numbers. SD also ensures the wave function is antisymmetric as the switching of two rows or columns will change the sign of the determinant. The one electron Hamiltonian in HF method, called Fock operator is defined as:

$$f_{i} = -\frac{1}{2}\nabla_{i}^{2} - \sum_{k}^{\text{nuclei}} \frac{Z_{k}}{r_{ik}} + V_{i}^{\text{HF}}\{j\}$$
(2.16)

where the first two terms are kinetic energy of the electron and potential energy of the electronnucleus attraction. $V_i^{\text{HF}}{j}$ is the Hartree-Fock potential that is the average repulsive potential that one electron feels from all the other electrons occupying orbitals {j}. $V_i^{\text{HF}}{j}$ can be computed as:

$$V_i^{\rm HF}\{j\} = \sum_{j \neq i} \frac{\int \rho_j}{r_{ij}} dr$$
(2.17)

where ρ_j is the probability density related to electron *j* which is calculated as $\rho_j = |\chi_i|^2$. Therefore, to calculate the spin-orbital of electron *j*, one must know the probability density of electron *j* that is the calculated using that exact spin-orbital. This circular problem is bypassed by iterative self consistent field (SCF) method. In SCF, the initial density matrix is guessed and used to solve HF equations to derive the new density matrix. The new density matrix is then used for the next iteration and so on until the self-consistency is reached, i.e. until the difference between two density matrices is less than the set threshold value.

HF derived energy of a system ($E_{\rm HF}$) is necessarily always larger (less negative) than the exact ground state energy (E_0) because of the variational principle.^[157] The difference between the exact energy of the ground state and the energy computed by HF method is called correlation energy ($E_c^{\rm HF}$) and is calculated as:

$$E_{\rm c}^{\rm HF} = E_0 - E_{\rm HF} \tag{2.18}$$

 E_c is a negative quantity that is caused by instantaneous repulsion of the electrons which is neglected in HF method as it treats repulsion only as average potential field. Although the error of HF might seem small, the electron correlation is very important quantity for the exact treatment of chemical systems. Considering the helium atom, the correlation energy is about 26 kcal mol⁻¹.^[157] As the system increase in size, so does the correlation energy. Therefore, new methods are developed and continue developing to include the correlation effect and improve the results of HF. The methods can be divided to semiempirical methods developed by fitting to empirical data and post-HF methods that are purely theoretical. These post-HF methods include configuration interaction (CI) methods, Møller-Plesset perturbation (MP*n*) methods and coupled cluster (CC) methods.^[157] In theory, ab initio post-HF methods can produce the exact solution if all possible excited states are considered with the infinite basis set. In practice this is not possible as the calculations become too expensive. Therefore, the post-HF ab initio methods with large basis set are usually used only for a small molecular system and the choice of computational method is a compromise between the computational efficiency and method accuracy.

2.3.2. Density Functional Theory (DFT)

The *N*-electron wave function depends on 4*N* coordinates (3*N* spatial and 1*N* spin coordinates). Thus, the methods based on wave function can become computationally very expensive for a system containing many atoms and electrons, which are of common interest in biological, chemical, and physical systems.^[158] This complexity of a wave function was a motivation for the development of density functional theory (DFT) that uses electron density ($\rho(r)$), i.e. probability density, as a variable to determine many-electron system's properties. Unlike wave function, electron density depends only on three spatial coordinates, regardless of the system size. Also, it has a physical meaning as it is a quantity that can be experimentally measured (observable). Moreover, electron density integrated over all space gives the total number of electrons:

$$\int \rho(r)dr = N \tag{2.19}$$

The local maxima of an electron density correspond to the positions of nuclei as the nuclei as there are effectively positive point charges that attract negative electrons. At these positions, the gradient of the density is discontinuous which results in cusp. The electron density at these positions can also give information about nuclear charges as:

$$\frac{\partial \bar{\rho}(r_{\rm A})}{\partial r_{\rm A}}\bigg|_{r_{\rm A}=0} = -2Z_{\rm A}\rho(r_{\rm A})$$
(2.20)

where $\bar{\rho}$ is a spherically averaged electron density, Z_A is nuclear charge of nucleus A, and r_A is the radial distance from A.^[158]

DFT is based on two theorems proved by Hohenberg and Kohn in 1964.^[159] The first theorem is an existence theorem that states that the energy of the ground electronic state is a unique functional of the electron density, i.e., the electron density determines the external potential (to an additive constant), which in turn determines the ground state wave function. The second one establishes the variational principle, that is the energy of a system calculated using trial electron density is higher or equal to the ground state energy of that system. These two theorems lead to a conclusion that there is a universal energy functional $E[\rho(r)]$ that can be used to obtain the energy of a system and, by energy minimisation, the exact ground state energy. The energy functional contains three terms and can be expressed as:

$$E[\rho(\mathbf{r})] = T[\rho(\mathbf{r})] + V_{\rm ne}[\rho(\mathbf{r})] + V_{\rm ee}[\rho(\mathbf{r})]$$
(2.21)

where *T* is kinetic energy of electrons, V_{ne} is potential energy of nuclear-electron attraction and V_{ee} is potential energy of electron-electron repulsion. *T* and V_{ee} are impossible to calculate due to the electron-electron repulsion in the correct Hamiltonian.

2.3.2.1. Kohn-Sham model

In 1965, Kohn and Sham (KS) proposed a key model that allowed the practical use of DFT in computational chemistry. They introduced a fictious system of non-interacting electrons that have, in their ground state, the same electron density as some real system where electrons do interact.^[160] As electron density determines the position of nuclei, the geometry of fictious and real system is the same. The energy of KS system is formulated as:

$$E[\rho(\mathbf{r})] = T_{\rm ni}[\rho(\mathbf{r})] + V_{\rm ne}[\rho(\mathbf{r})] + V_{\rm ee}[\rho(\mathbf{r})] + E_{\rm xc}[\rho(\mathbf{r})]$$
(2.22)

where T_{ni} is kinetic energy of the non-interaction electrons, V_{ne} is potential energy of nuclearelectron attraction, V_{ee} is classical electron-electron repulsion, and finally E_{xc} is exchange correlation term that holds all the corrections to the potential energy from the non-classical electron-electron repulsion and correction to kinetic energy that arise from the electron-electron interactions.^[161] The ground-state density can be calculated as:

$$\rho(r) = \sum_{i}^{N} |\chi_{i}(r)|^{2}$$
(2.23)

where *N* is the number of electrons and χ_i is the Kohn-Sham orbital of *i*-th electron represented as linear combination of basis set functions. The χ_i orbitals are solution on *N* single particle equations:

$$h_i^{\rm KS}\chi_i = \varepsilon_i\chi_i \tag{2.24}$$

where h_i^{KS} is one electron Kohn-Sham operator defined as:

$$h_{i}^{\rm KS} = -\frac{1}{2}\nabla_{i}^{2} - \sum_{k}^{\rm nuclei} \frac{Z_{k}}{|r_{i} - r_{k}|} + \int \frac{\rho(r')}{|r_{i} - r'|} dr' + V_{\rm xc}$$
(2.25)

$$V_{\rm xc}(r) = \frac{\delta E_{\rm xc}}{\delta \rho} \tag{2.26}$$

where i runs over all electrons and k over all nuclei. The KS orbitals form the Slater-determinant that is an eigenfunction of the sum of the Kohn-Sham operators:

$$\sum_{i}^{N} h_{i}^{\mathrm{KS}} |\chi_{1}\chi_{2} \dots \chi_{N}\rangle = \sum_{i}^{N} \varepsilon_{i} |\chi_{1}\chi_{2} \dots \chi_{N}\rangle$$
(2.27)
The form of KS equations is similar to HF equations. The solution of ground-state density is found using an iterative SCF procedure and minimizing the system energy.^[161] The key difference between HF and DFT is that the DFT is exact provided we know the exact functionals. In practice, however, the exchange-correlation functional is unknown thus different approximations are used to find a solution for ground-state density. It is worth noting that the computer cost of DFT methods formally scales as N^3 where N is the number of basis functions used, while HF method scales as N^4 . The convergence respect to basis set tends to be more rapid for DFT methods compared to methods based on molecular orbitals. DFT methods generally preform much better than HF and are less expensive than post-HF methods.^[161]

2.3.2.2. Exchange-correlation functionals

The quality of the DFT results depends on the quality of approximation used for exchangecorrelation functional. While as in methods based on wave functions, there is a clear path to the exact solution, in DFT there is no systematic way of improving the results. Therefore, many different functionals have been proposed in literature and the choice to use one of them depend mainly on the chemical system that needs to be solved.^[161] The functionals are generally expressed as:

$$E_{\rm xc}[\rho(r)] = \int \rho(r) \varepsilon_{\rm xc}[\rho(r)] dr \qquad (2.28)$$

where ε_{xc} is the energy density per particle density. This term is always treated as a sum of individual exchange and correlation contributions:

$$\varepsilon_{\rm xc} = \varepsilon_{\rm x} + \varepsilon_{\rm c}$$
 (2.29)

The functionals are generally divided according to their complexity to several categories.

Local density approximation (LDA) functionals use the simplest approximations in which the energy density ε_{xc} on position *r* depends solely on the value of density ρ at the same position. In practice, this is true only for functional derived from the analysis of the uniform electron gas. The exchange for the uniform electrons gas is known to be:^[162]

$$E_{\rm x}^{\rm LDA}[\rho] = -\frac{3}{4} \left(\frac{3}{\pi}\right)^{\frac{1}{3}} \int \rho^{\frac{4}{3}} dr$$
(2.30)

It is not possible to analytically derive the correlation functional for uniform gas. Instead, the results of Monte Carlo simulations that computed the uniform gas to high accuracy were used to determine

the correlation energy. Local spin density approximation (LSDA) additionally includes the local spin-polarization.

Generalized gradient approximation (GGA) functionals consider electron density and its gradient at position r. Most of the GGA functionals are constructed by adding correction to LDA functional in a way:

$$\varepsilon_{\rm xc}^{\rm GGA}[\rho(r)] = \varepsilon_{\rm xc}^{\rm LDA}[\rho(r)] + \Delta \varepsilon_{\rm xc}^{\rm GGA} \left[\frac{|\nabla \rho(r)|}{\rho^{\frac{4}{3}}(r)} \right]$$
(2.31)

GGA functionals are developed both as non-empirical and empirical functionals. Parameter free functionals use only theoretical conditions to determine new parameters in a general form. While as empirical functional are derived by fitting the parameters to experimental data or accurately calculated atomic and molecular properties. The most commonly used GGA functionals are Perdew, Burke and Ernzerhof (PBE),^[163] and Becke, Lee, Parr and Yang (BLYP).^[164]

Next in line are meta-GGA functionals that consider second order gradients of the density, and the Kohn-Sham kinetic energy density defined as:

$$\tau(r) = \sum_{i}^{\text{occupied}} \frac{1}{2} |\nabla \chi_i(r)|^2$$
(2.32)

meta-GGA are typically more accurate than GGA with a comparable cost. The most used meta-GGA functional is TPSS constructed by Tao, Perdew and co-workers.^[165]

All of the above functionals are referred to as "pure functional" as they are derived only from DFT theory. Hybrid functionals, on the other hand, mix a portion of HF exact exchange with GGA or LDA functionals. This allows the inclusion of the effects of static correlation which consequently leads to more accurate results. The parameters for hybrid functionals are determined by fitting to a set of molecular data. The most favoured functional of this kind is B3LYP as it showed a reasonably good performance over a wide range of chemical systems. B3LYP is defined as:

$$E_{\rm xc}^{\rm B3LYP} = (1-a)E_{\rm x}^{\rm LSDA} + aE_{\rm x}^{\rm HF} + b\Delta E_{\rm x}^{\rm B} + (1-c)E_{\rm c}^{\rm LSDA} + cE_{\rm c}^{\rm LYP}$$
(2.33)

where ΔE_x^B are GGA corrections to the LDA exchange energy and *a*, *b*, *c* are parameters that in this method equal 0.20, 0.72, and 0.81 respectively.^[166] Hybrid functionals are more costly than pure functionals due to the exact HF exchange that depends not only on density but also on density matrix.^[167]

Double hybrid functionals include virtual orbitals to treat non-local correlation energy by second order Møller-Plesset perturbation theory (MP2).^[168] Double-hybrid DFT is one of the most accurate quantum chemical method but are also computationally very expensive.^[169]



Hartree hell

Figure 2.7. Jacob's ladder of DFT functionals.

2.3.2.3. Dispersion energy correction

DFT methods generally lack the ability to account for inter or intramolecular dispersion (London) forces.^[162] Dispersion forces are attractive forces that derive from electron correlation at long distances. Local functionals are not able to properly describe dispersion as they rely entirely on local density, and adding non-local HF exchange does not help as it cannot account for electron correlation. Consequently, DFT methods are often inaccurate for system where these interactions play a significant role, as for an example, biological systems, molecular crystals, and adsorption on surface. New methods and functionals able to capture a portion of dispersion interactions and improve results are developed.^[170] However, the methods based entirely on first principle are computationally too demanding. On the other hand, semiempirical dispersion correction methods, proposed by Grimme et al. and called DFT-D, are able to largely improve the results where dispersion interactions play a significant roles without adding a notable computational cost.^[171] In DFT-D, empirical correction is added a posteriori to the energy calculated by KS DFT:

$$E_{\rm DFT-D} = E_{\rm DFT} + E_{\rm disp} \tag{2.34}$$

where E_{disp} is an empirical dispersion correction given as:

$$E_{\rm disp} = -s_6 \sum_{j>i}^{N} \frac{C_6^{ij}}{R_{ij}^6} f_{\rm dmp}(R_{ij})$$
(2.35)

N is the number of atoms in the molecule, s_6 is a global scaling factor that depends on the functional, C_6^{ij} are dispersion coefficients for atom pair ij, R_{ij} is a distance between atom pair ij, and f_{dmp} is a damping function that tends to 1 at large interatomic distance, and to 0 at small distances.^[172] The difference between DFT-D methods generally differ by the approach used to calculate dispersion coefficient.^[173,174]

2.3.3. Quantum chemistry calculations

The simplest kind of quantum chemistry calculation is a single point (SP) energy calculation. SP calculates the wave function of a chemical system at a specified geometric structure. This kind of calculation are often performed using more accurate and costly methods on systems whose geometries were previously optimized at lower level of theory.

Geometry optimization calculation locates the stationary point on potential energy surface, that is, the point where gradient of the energy is zero. The geometry optimisation starts with the input structure and continues along the potential energy surface. At each point, the energy and the gradient of the energy are calculated, i.e. the forces that determine the direction and slope on the PES in which the energy decreases. The optimisation algorithm then determines how to change the system coordinates in order to reduce the energy of the system. Most optimisation algorithms today also calculate the second derivative of the energy, also called the Hessian, which gives the curvature of the surface at that point to helps determine the next step. The process of changing the geometry and calculating the energy and forces is repeated until convergence, that is, until the forces are close to zero and the changes in geometry and energy are below a specified threshold. Geometry optimisation is generally used to find stable geometries, that is, minima on the PES. It should be noted that the optimisation algorithm always leads towards lower energies, so it finds a local minimum on the PES that is not necessarily the most stable structure, i.e. a global minimum. To find another local or global minimum, a new input geometry is needed.

Frequency calculations determine normal-mode vibrational frequencies that can be used to predict infrared spectrum and calculate zero-point energy and other thermodynamical properties of interest, such as enthalpy and entropy of a system at given temperature. Frequency analysis often follows geometry optimization to determine the nature of stationary point as it calculates the Hessian matrix that describes the curvature of PES. If an element in Hessian matrix is negative, the corresponding frequency is called imaginary frequency. Stationary point that has no imaginary frequencies is a true minimum on PES. If there is only one imaginary frequency, the stationary point is a saddle point that might correspond to transition state, and if it has *n* imaginary frequencies it is a saddle point od n^{th} order.



Figure 2.8. Flow diagram of general single point calculation and geometry optimization process.

2.3.3.1. Including the effect of solvent

The calculations in gas phase are much simpler than the ones in solvent. However, they are often inadequate to describe a molecule in solution as the solvent can affect its energy and geometry as well as other properties. The molecules of solvent can also directly affect the adsorption mechanism.^[175] As experimental studies suggest, the adsorption of organic pollutants to hydrophobic microplastic is often largely controlled by hydrophobic interaction.^[68,69] These types of interactions are directly related to the presence of water molecules in the system and can be

attributed to their tendency to interact with other water molecules rather than adsorbate and adsorbent molecules.^[176] Thus, the calculations in gas phase cannot account for the hydrophobic interactions. Other types of interactions are also affected by the presence of water or some other solvent in a studied system. For example, electrostatic interactions are much less important for the system placed in solvent with high dielectric constant, such as water, than for the system in a gas phase.

Generally, the effect of solvent is included in calculations by two different approaches: explicit (atomistic), or implicit (continuum) approach. In atomistic approach, the molecules of solvent are added around the solute to simulate the solvation shells, while in continuum approach, the solvent is treated as a continuum dielectric featuring specific dielectric properties. The clear advantage of atomistic approach over continuum, is the ability to describe a specific interaction between a solute and solvent molecule, such as hydrogen bonds.^[177] However, the addition of solvent molecules is computationally very demanding. Nowadays, QM/MM approaches are developed to include a larger quantity of explicit solvent molecules whereas a small number of solvent molecules closest to the solute are treated at quantum mechanics (QM) level of theory, while the bulk solvent is treated by molecular mechanics (MM) classical force field.

Implicit treatment of solvent is included in self-consistent reaction field (SCRF) method. In this method, the solvent is described as a continuum of uniform dielectric constant ε_r . The solute is placed in a molecular-shaped cavity within the solvent. Then, the solute is able to polarize the dielectric which creates an apparent charge density on the surface of the cavity that back-polarizes the solute.^[177] There are various SCRF methods available that differ by the way they define the cavity and dielectric constant.^[178] Some of the most commonly used are PCM,^[179] CPCPM,^[180] COSMO^[181] and SMD.^[182] The combination of implicit and explicit approach is also often used in a way that few solvent molecules are treated explicitly while the rest of the solvent is treated as continuum. These combined models are referred to as cluster or supermolecule models.

2.3.3.2. Basis set superposition error (BSSE)

The interaction energy of two monomers A and B ($E_{int}(AB)$) is calculated as the difference between the energy of a dimer (E(AB)) and energies of separated monomers (E(A) and E(B)):

$$E_{\rm int}(AB) = E(AB) - E(A) - E(B)$$
 (2.36)

One of the well-known problems that occur when calculating energy of a dimer using DFT methods is basis set superposition error (BSSE). In a calculation of a dimer, an individual monomer can

utilize some of the basis functions of other monomer and vice versa, which stabilizes the dimer and lowers its energy. This stabilization of a dimer leads to an artificial overestimation of interaction energy between monomers which is especially problematic in calculations of weakly bound systems.^[183] BSSE is a consequence of using a finite basis set to calculate energies of monomers and dimer. Thus, it can be reduced by using more basis functions, and completely eliminated by using infinite basis set, which is practically impossible due to the increasing cost of calculation.

A very popular method to deal with the BSSE is counterpoise-correction (CP) method proposed by Boys and Bernardi.^[183] In CP method, the energy of a monomer is calculated using the basis set of a whole dimer. That is, the energy of monomer A is calculated in a dimer basis set in a way that the electrons and nuclear charges of monomer B are neglected and vice versa. The atoms without nuclear charge and electrons are usually called ghost atoms, and their basis function are called ghost functions.^[184] The interaction energy can then be rewritten as:

$$E_{\text{int}}^{\text{CP}}(\text{AB}) = E_{\text{AB}}(\text{AB}) - E_{\text{AB}}(\text{A}) - E_{\text{AB}}(\text{B})$$
(2.37)

where subscript AB means that the whole dimer basis set is used. The stabilization of one monomer by extra basis functions of another monomer is estimated as:

$$E_{\text{BSSE}}(A) = E_{AB}(A) - E_A(A)$$
(2.38)

$$E_{\rm BSSE}(B) = E_{\rm AB}(B) - E_{\rm B}(B)$$
 (2.39)

and the counterpoise corrected energy for a dimer is expressed as:

$$E_{AB}^{CP}(AB) = E_{AB}(AB) - E_{BSSE}(A) - E_{BSSE}(B)$$
(2.40)

For interaction calculations in continuum solvent, there is still no widely used method to deal with BSSE. In some studies, the energies of dimer and monomers in continuum solvent are calculated using their own basis set, while the correction to the monomer energies due to BSSE are calculated in gas phase.^[185] In other studies, the interaction is calculated using a standard CP method, but the monomers are calculated in a dimeric cavities.^[186,187]

2.3.4. Molecular dynamic simulations

The quantum chemical methods described so far provide a static image of the molecular system in which the positions of the nuclei are fixed. Molecular dynamics (MD) simulations, on the other hand, enable the movement of atoms and time-evolution of a chemical system. This provides insights into various dynamic processes such as conformational changes, solvation processes, ligand binding and adsorption or absorption mechanisms as well as the calculation of the equilibrium and transport properties of a system.^[188–190]

In MD, each atom is characterised by the specific position and momentum coordinates. The system of N atoms therefore defines a 6N-dimensional space called phase space.^[191] The time evolution of a system (trajectory) is calculated according to the laws of classical mechanics by solving Newton's equations of motion:

$$f_i = m_i \frac{\partial^2 q_i}{\partial t^2} = m_i \frac{\partial v_i}{\partial t} = -\frac{\partial U(q)}{\partial q_i}$$
(2.41)

where f_i is the force acting on atom i, m_i is the mass of atom i, q_i is a 3D vector representing the position of atom i in space, t is the time, v_i is the velocity and U(q) is the potential energy defined by the classical force field. In MD, the properties of a system are calculated by time averaging a simulation that is long enough to achieve convergence. Convergence is achieved when enough phase points are sampled and when a further increase in time does not change the results.^[192] MD assumes the validity of the ergodic hypothesis, which states that all accessible microstates are equally likely over a long period of time. In other words, the average value of a system over all possible quantum states is equal to its time average.^[193] MD methods can be divided into classical and ab initio (AIMD) methods. In classical MD, the interatomic forces are calculated from a predefined force field, whereas in AIMD the forces are obtained from QM calculations. Although much higher accuracy can be achieved with AIMD, the computational cost of such calculations is considerably higher, so that only small systems can be treated.^[194] Therefore, it is only used when high-precision simulations are required or to treat effects that can hardly be described with the force field method, such as bond formation and breakage or polarisation and charge transfer effects.

2.3.4.1. Force field

A classical mechanic force field (FF) is a set of equations and associated parameters chosen to model the potential energy of a system from which the interatomic forces are calculated. The force field methods are based on the assumption that the potential energy can be written as the sum of different potentials with a simple physical interpretation.^[195] In most FFs, the total potential energy U_{FF} is composed as follows:

$$U_{\rm FF}(q) = U_{\rm bond}(d) + U_{\rm angle}(\theta) + U_{\rm dihedral}(\phi) + U_{\rm improper}(\psi) + U_{\rm ele}(r) + U_{\rm vdW}(r)$$
(2.42)

where U_{bond} , U_{angle} , U_{dihedral} and U_{improper} are bonded terms that refer to the stretching and compressing of bonds, the bending of angles, torsion (proper dihedrals) and out-of-plane bending (improper dihedrals), respectively. U_{ele} and U_{vdW} are nonbonded terms related to electrostatic and dispersion (van der Waals) interactions.



Figure 2.9. Graphic representation of the main terms defining the potential energy in classical force field, including potential energy of bond stretching $(U_{bond}(d))$, bond-angle bending $(U_{angle}(\theta))$, dihedral-angle torsion $(U_{dihedral}(\phi))$, improper dihedral-angle bending $(U_{improper}(\psi))$, and electrostatic $(U_{ele}(r))$ and van der Waals interactions $(U_{vdW}(r))$.

The bond stretching term can be defined by harmonic potential as following:

$$U_{\text{bond}}(d) = \sum_{i} \frac{1}{2} k_{d,i} (d_i - d_{i0})$$
(2.43)

where $k_{d,i}$ is the force constant, d_i is the bond length, and d_{i0} is the reference bond length for the bond *i*. This term is also often defined by the Morse potential as follows:

$$U_{\text{bond}}(d) = \sum_{i} D_{e,i} \left(1 - e^{-a_i(d_i - d_{i0})} \right)^2$$
(2.44)

where $D_{e,i}$ is the depth of the energy minimum and a_i is related to the force constant. The potential energy of angle bending is often defined as follows:

$$U_{\text{angle}}(\theta) = \sum_{i} \frac{1}{2} k_{\theta,i} (\theta_i - \theta_{i0})^2$$
(2.45)

where $k_{\theta,i}$ is the bending force constant, and θ_i and θ_{i0} are the bond angle and the reference bond angle respectively for angle *i*. The accuracy of the defined bond stretching and angle bending potentials is often improved by including additional terms in the Taylor expansion. The proper dihedral potential energy can be expressed as follows:

$$U_{\text{dihedral}}(\phi) = \sum_{i} k_{\phi,i} (1 + \cos(n_i \phi_i - \phi_{i0}))$$
(2.46)

where $k_{\phi,i}$ is related to the height of the potential energy barrier, ϕ_i is the torsion angle, ϕ_{i0} is the torsion angle that determines the position of a minimum, and n_i is the number of minima. The torsional energy can also be calculated as follows:

$$U_{\text{dihedral}}(\phi) = \sum_{i} \sum_{n=0}^{N} C_{n,i}(\cos(\phi_i))^n$$
(2.47)

where $C_{n,i}$ is a constant and n is a number of terms, normally between 4 and 6. Improper dihedrals potential is intended to preserve the planarity or chirality of particular groups. It provides a potential penalty when atoms in a planar molecule move out of plane or when chiral molecules turn into their mirror image. This potential energy is calculated in the same way as the potential energy for the proper dihedral. The non-bonded potential energy terms are calculated for atoms separated by three or more bonds and for atoms belonging to different molecules. The simplest form for calculating the non-bonded van der Waals potential energy is the Lennard-Jones 12–6 potential, which is expressed as follows:

$$U_{\rm vdW}(r) = \sum_{i,j} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$$
(2.48)

where ε_{ij} and σ_{ij} are constants, and r_{ij} is the distance between atoms *i* and *j*. The electrostatic interaction between two atoms is commonly represented as the Coulomb interaction of two point charges:

$$U_{\rm ele}(r) = \sum_{i,j} \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$
(2.49)

where q_i and q_j are the partial charges of atoms *i* and *j*, ϵ_0 is the dielectric constant and r_{ij} is the distance between atoms *i* and *j*. Different force fields may use different equations to calculate the

potential energy terms and different parameters obtained by fitting to the experimental data or by QM calculations of smaller set of molecules. The accuracy of classical MD simulations largely depend on the FF used. The most widely used forcefields today are the ones belonging to AMBER, CHARM, GROMOS and OPLS groups. These FFs are still actively developed, tested and improved, and, by now, provide parameters for a very large subset of molecules.^[196] Although very similar in their functional form, these FFs differ by the parametrization strategies. Therefore, it's important to use only one force field throughout the MD simulation and ensure that the same settings used during parameterization are applied.^[197]

For the simulation of systems in a water environment, it is essential to use explicit water molecules to account for the specific water-solute interactions. Since the number of water molecules surrounding the system is generally very high, simplified water models are used to reduce the computational cost of a simulation. In these models, bond lengths and bond angles are fixed, interatomic forces are represented by the Lennard-Jones potential and a water dipole moment is represented by two or more point charges. The most commonly used simplified water models in MD are SPC/E (extended simple point charge) and TIP*n*P (transferable intermolecular potential). SPC/E^[198] and TIP3P^[199] are three-sided models in which three point charges are centred on each atom, but they differ in parameters.^[200] TIP4P^[199] and TIP5P^[201] are also often used models from the same family as TIP3P, but with four and five point charges respectively. These models are more complex and therefore increase computational cost, but provide a more accurate description of water.^[202,203]

The simulation of a system in a bulk phase is usually carried out under periodic boundary conditions (PBC). In PBC, the system of interest is placed in a box of finite size surrounded by its copies translated in all directions. Thus, if a molecule exits the box on one side, it enters on the other side of a box. This approach eliminates the border effect that occurs in a finite system and allows the use of a smaller, computationally more efficient system to study the effect in a bulk. When using the PBC to simulate the finite system, it is important to choose the size of the box large enough so that the molecules do not interact with their own translated copies.^[191] For this purpose, it is necessary to define a cutoff distance r_{cutoff} above which all non-bonded (van der Waals and electrostatic) interactions can be neglected. Van der Waals interactions are proportional to r^{-6} , so that they quickly decay to zero at large distances. At about 10 to 12 Å, which is a typical r_{cutoff} , van der Waals interactions are close to zero, and truncation of the potential leads to only a small

error.^[204] The discontinuity of the potential can be adjusted by a switching function that makes the potential smoothly go to zero, or by shifting the function by a constant value so that it becomes zero at the cutoff distance.^[205] The range of the electrostatic interaction is much larger as it decays proportionally to r^{-1} . In general, r_{cutoff} is set to the same value as for the van der Waals interactions. The simplest truncation of the electrostatic potential at this distance leads to a significant error. One of the alternative solutions is to use methods based on Edwald summation, which split the Coulomb interaction into a short-range term calculated in real space and a long-range term calculated in reciprocal space.^[197] The most commonly used method today for considering a long-range electrostatic interaction is particle-mesh Ewald method (PME).^[197] Part of the calculation of non-bonding interactions is the evaluation of all interatomic distances which formally scales as N^2 , where *N* is the number of atoms. An easy way to speed up calculated that is typically updated every 10-20 steps of the calculations.^[206]



Figure 2.10. Illustration of periodic boundary conditions (PBC).

2.3.4.2. Integration algorithms

The equations of motion are solved using numerical algorithms. The most commonly used algorithms are Verlet-like algorithms due to their simplicity and stability. The Verlet algorithm is based on the Taylor expansion of a position coordinate by the time t.^[192] The positions of the atoms after the time step Δt are calculated only from the forces at time t:

$$q(t + \Delta t) = 2q(t) - q(t - \Delta t) + \frac{f(t)}{m} \Delta t^{2}$$
(2.50)

The velocities are not needed for the calculation of the next step, but they can be derived as follows:

$$v(t) = \frac{q(t + \Delta t) - q(t - \Delta t)}{2\Delta t}$$
(2.51)

The Leap-Frog algorithm is equivalent to Verlet.^[192] However, this algorithm uses the forces f(t) to update the velocities for half a time step $(t + 1/2 \Delta t)$ which are then used to update the positions at time $t + \Delta t$ by the following equations:

$$v\left(t + \frac{1}{2}\Delta t\right) = v\left(t - \frac{1}{2}\Delta t\right) + \frac{\Delta t}{m}f(t)$$
(2.52)

$$q(t + \Delta t) = q(t) + \nu \left(t + \frac{1}{2\Delta t} \right) \Delta t$$
(2.53)

In this way, the velocities and positions are not calculated simultaneously. Consequently, the kinetic and potential energy are not calculated simultaneously, so that the total energy of the system is not defined. Another popular alternative to the Verlet algorithm, in which positions and velocities are calculated simultaneously, is the Velocity-Verlet algorithm. In this algorithm, the new positions are calculated from the forces and velocities and the velocities are then updated in a way from the forces:

$$q(t + \Delta t) = q(t) + v(t)\Delta t + \frac{f(t)}{2m}\Delta t^2$$
(2.54)

$$v(t + \Delta t) = v(t) + \frac{\Delta t}{2m} [f(t) + f(t + \Delta t)]$$
(2.55)

In the limit of infinitely small time steps, the numerical integration algorithm will generate the exact trajectory. However, each time step requires a calculation of the forces. It therefore makes sense to use longer time steps in order to capture a trajectory over a longer period of time. In practice, the time step should be shorter than the fastest movement in the molecule. In a molecule with hydrogen atoms, the vibration of the bond between the hydrogen and the heavy atom is about 10^{-14} s. Therefore, the time step should not be greater than 0.1 fs. This means that at least 10^7 steps

are required for 10 ns simulation.^[192] Alternative algorithms have been proposed and implemented to reduce the computational cost for long simulations by increasing the time step. The higher order algorithms incorporate the higher order derivatives of the position to achieve higher accuracy at the given time and better predict the position after a longer time step. Other algorithms, such as SHAKE, SETTLE and LINCS, use a restricted dynamic approach by freezing the fast bonding motions such as those between hydrogen and heavy atoms.^[207–210]

2.3.4.3. Molecular dynamics ensembles

A trajectory produced by solving Newton's equations of motions probes the microcanonical (*NVE*) ensemble in which the number of atoms N, the volume of a system V and the energy E are held constant. However, MD simulations are usually performed in other types of ensembles, such as canonical (*NVT*) or isothermal-isobaric (*NPT*) ensembles, in which the temperature T and/or the pressure P are kept constant instead of the total energy of a system E.^[211] Changing the ensemble type in MD simulations can be very useful for various reasons. For example, to better match the chemical experiments that are typically performed at constant temperature or constant pressure, or to study the effects of temperature on the dynamic properties of the system. Switching from microcanonical to other types of ensembles is possible using algorithms that modify the equations of motion. Algorithms that control the temperature during the simulation are called thermostats and those that control the pressure are called barostats.

The temperature of a system is directly proportional to the kinetic energy. Keeping this value constant throughout the simulation would therefore mean that the kinetic energy remains constant, which is not desirable. Instead, the thermostat ensures that the average value of a system corresponds to the desired value, but allows the temperature and kinetic energy to fluctuate during the simulations. This fluctuation can be achieved by coupling a system to a large heat bath with the desired temperature T_0 .^[212] In most cases, simulations are performed in an isothermal-isobaric *NPT* ensemble controlled by both thermostats and barostats.^[213] Similar to thermostats, barostats control the average pressure of a system while allowing for pressure fluctuations during the simulation. This chapter introduces the most widely used thermostats, including Berensted and canonical sampling thermostats based on velocity rescaling, Anderson thermostats based on a stochastic approach, and Nosé-Hoover based on extended equations of motion. These are followed by the two most widely used barostats, Berensted and Parrinello-Rahman barostat.

The easiest way to control the temperature of a system is to rescale the velocity of a system using the equation:

$$v_{\rm new} = v_{\rm old}\lambda \tag{2.56}$$

where v_{new} is a new velocity of a system, which calculated by rescaling the velocity v_{old} after each time step by the scaling factor λ . In the strong coupling method, the factor λ is calculated as $\sqrt{T_0/T}$ where *T* is the current temperature of a system and T_0 desired temperature. In this approach the velocities change quite drastically, and the fluctuation of temperature is not allowed. The Berensted algorithm, also called the weak coupling algorithm, allows the control of the strength of the coupling between a system and a heat bath.^[214] In Berendsen thermostat, the rate of temperature change is proportional to the difference between the desired temperature T_0 and the actual temperature *T*:

$$\frac{dT}{dt} = \frac{1}{\tau} (T_0 - T)$$
(2.57)

where τ is the coupling time, which determines how tight the bath and a system are coupled and consequently determines the time scale for reaching the desired temperature. The scaling factor λ is calculated as:

$$\lambda = \sqrt{1 - \frac{\Delta t}{\tau} \left(1 - \frac{T_0}{T}\right)}$$
(2.58)

When the coupling parameter τ is high, the coupling is very weak and the scaling factor approaches unity, i.e. the microcanonical ensemble. The lower the τ values, the stronger the coupling and the more significant the temperature exchange between the bath and the system. And if the value of τ is equal to Δt , the scaling factor is $\sqrt{T_0/T}$, just like in the simplest velocity rescaling algorithm. The Berensted thermostat does not allow the kinetic energy to fluctuate and does not generate a proper canonical ensemble. The kinetic energy is not properly redistributed to the system, which invalidates the equipartition theorem and leads to an artifact called the flying ice cube effect.^[215] Considering this fact, the Berensted thermostat should not be used for MD simulations and is mostly only discussed for historical reasons, although it can still be used for fast equilibration of a system to the desired temperature.

An alternative algorithm based on velocity rescaling is the canonical sampling through velocity rescaling (CSVR) thermostat proposed by Bussi et al.^[216] The CSVR thermostat is similar

to the Berenested algorithm, but has an additional stochastic term that ensures the correct canonical ensemble. The scaling factor λ in CSVR is calculated as follows:

$$\lambda = \sqrt{\frac{K_0}{K}}$$
(2.59)

where *K* is the kinetic energy of a system and K_0 is the desired kinetic energy obtained by a stochastic method from the probability density function:

$$\bar{P}(K_0)dK_0 \propto K_0^{N_f/(2-1)} e^{-\frac{K_0}{k_b T_0}} dK_0$$
(2.60)

where N_f is the number of degrees of freedom, k_b is the Boltzman constant and T_0 is the desired temperature. In this way, the CSVR thermostat enables the correct distribution of kinetic energy and does not exhibit an ice-cube artefact.^[215]

The Andersen thermostat is based on the stochastic collision method.^[217] In this algorithm, the randomly selected particles in the system occasionally collide with the heat bath, resulting in a transfer of energy between them. Between these stochastic collisions, the system evolves according to the normal Newtonian equations of motion in the *NVE* ensemble. The strength of the coupling is determined by the frequency of the occasional collisions. If the collisions are uncorrelated, the distribution of the time intervals between two collisions follows the Poisson distribution:

$$P(t;\nu) = \nu e^{-\nu t} \tag{2.61}$$

where P(t; v)dt is the probability that the next collision will occur in the time interval [t, t + dt]. The Anderson thermostat can provide good results or time-independent properties. However, it should not be used to study the dynamic properties of a system.^[212]

One way to perform deterministic MD at constant temperature is to use the Nosé-Hoover thermostat, which uses an extended Lagrangian form to solve the equation of motion. The Lagrangian and Hamiltonian formulas are alternative expressions of classical mechanics that lead to the same solution as Newton's equation. These alternative formulations are used by more advantageous techniques to solve the equations of motion in a more computationally effective way than the Newtonian formulation. The Lagrangian \mathcal{L} is defined as the difference between kinetic and potential energy, which can be written as follows:

$$\mathcal{L}(q, \dot{q}) = \sum_{i=1}^{N} \frac{1}{2} m_i \dot{q}_i^2 - U(q)$$
(2.62)

where m_i and q_i are the mass and the coordinates of atom *i* and *U* is the potential energy.^[218] The equations of motion can be solved in terms of coordinates *q* and momentum *p* using the Hamiltonian formulation where the Hamiltonian \mathcal{H} is defined as follows:

$$\mathcal{H}(q,p) = p\dot{q} - \mathcal{L}(q,\dot{q},t) \tag{2.63}$$

In the Nosé-Hoover thermostat, the system of N particles with the coordinates q_i' , the masses m_i' , momenta p_i' and the potential energy U(q') is extended by an artificial dynamic variable s, which is associated with an effective mass Q.^[218] The equations of motion are solved for virtual variables q, p, t that are related to the real variables q', p', t' as follows:

$$q_i' = q_i \tag{2.64}$$

$$p_i' = \frac{p_i}{s} \tag{2.65}$$

$$\Delta t' = \frac{\Delta t}{s} \tag{2.66}$$

The form of extended Lagrangian in the Nosé-Hoover thermostat is:

$$\mathcal{L}_{\text{Nose}}(q,\dot{q}) = \sum_{i=1}^{N} \frac{1}{2} m_i s^2 \dot{q}_i^2 - U(q) + \frac{1}{2} Q s^2 - L k_b T \ln(s)$$
(2.67)

where L is the number of degrees of freedom whose value is chosen so that the canonical distribution is satisfied at equilibrium, k_b is the Boltzmann constant, T is the temperature and $\ln(s)$ is an additional term needed to achieve a canonical ensemble. The Nosé-Hoover thermostat solves the equations of motion in the Hamiltonian formalism. If p_i and p_s are expressen in terms of q_i and s as follows:

$$p_i = \frac{\partial \mathcal{L}_{\text{Nosé}}}{\partial \dot{q}_i} = m_i s^2 \dot{q}_i \tag{2.68}$$

$$p_s = \frac{\partial \mathcal{L}_{\text{Nosé}}}{\partial \dot{s}} = Q \dot{s} \tag{2.69}$$

then the Hamiltonian of the extended system can be written as:

$$\mathcal{H}_{\text{Nosé}}(p,q) = \sum_{i=1}^{N} \frac{p_i^2}{2m_i s^2} + U(q) + \frac{p_s^2}{2Q} - Lk_b T \ln(s)$$
(2.70)

and the equations of motions are defined as:

$$\frac{dq_i}{dt} = \frac{\partial \mathcal{H}_{\text{Nosé}}}{\partial p_i} = \frac{p_i}{m_i s^2}$$
(2.71)

$$\frac{dp_i}{dt} = \frac{\partial \mathcal{H}_{\text{Nosé}}}{\partial q_i} = -\frac{\partial U(q)}{\partial q_i}$$
(2.72)

$$\frac{ds}{dt} = \frac{\partial \mathcal{H}_{\text{Nosé}}}{\partial p_s} = \frac{p_s}{Q}$$
(2.73)

$$\frac{dp_s}{dt} = -\frac{\partial \mathcal{H}_{\text{Nosé}}}{\partial s} = \frac{\sum \frac{p_i^2}{m_i s^2} - Lk_b T}{s}$$
(2.74)

Nosé-Hoover yields a correct *NVT* and can be used to study the dynamic properties of a system as it is fully deterministic, i.e. it does not introduce any randomization of velocities that can change the dynamics of a system. However, it is not suitable for simulating a system that is too far away from the equilibrium, as it can lead to temperature oscillations.^[212] Therefore, the common practice in MD simulations is to equilibrate the system with another thermostat, e.g. the Berensted thermometer, before performing MD simulations with the Nosé-Hoover thermostat.

Barostats keep the average pressure at the desired value by controlling the volume of a simulation box.^[211] The Berensted barostat works in a similar way to the Berensted thermometer.^[214] It scales the volume of the simulation box the towards the desired pressure P_0 at each time step of the simulation according to:

$$\frac{dP}{dt} = \frac{P_0 - P}{\tau_P} \tag{2.75}$$

where P matrix is the current pressure, P₀ matrix is the desired pressure, and τ_P is the pressure time constant. The coordinates are scaled by a scaling matrix μ calculated as follows:

$$\mu_{ij} = \delta_{ij} - \frac{\Delta t}{3\tau_P} \beta_{ij} \left(P_{0ij} - P_{ij} \right)$$
(2.76)

where β is the isothermal compressibility of the system. In general, the value of β is unknown but it is sufficient to use a rough estimate. Most liquids have the same value of β and usually the value of water is used that is equal to 4.6×10^{-5} bar⁻¹. This method does not allow the volume fluctuation and therefore does not generate the correct *NPT* ensemble.

The other most commonly used barostat is the Parrinello-Rahman barostat, which is based on the extension of Hamilton's equations of motion, similar to the Nosé-Hoover thermostat.^[213,219] This barostat makes it possible to vary both the volume and the shape of a simulation box according to the equation of motion:

$$\frac{db^2}{dt^2} = VW^{-1}b'^{-1}(P - P_0)$$
(2.77)

where b is the matrix of the box vectors, V is the box volume and W is the matrix parameter that determines the strengths of the coupling which can be calculated from the approximate isothermal compressibility β and the box size. Just like the Nosé-Hoover thermostat, the Parinello-Rahman barostat also suffers from large volume oscillations when the initial system is far from equilibrium, which can cause the simulation to crash. Therefore, before performing an *NPT* simulation with a Parinello-Rahman barostat, the system should first be equilibrated to the desired pressure using a different type of barostat, e.g. a Berenested barostat.

2.3.5. QSA/PR mathematical modelling

The Quantitative Structure Activity/Property Relationship (QSA/PR) methodology is a very valuable tool for the mathematical description of the relationship between the activity (QSAR) or another property (QSPR) and the structural features of the molecule. The first QSAR model was developed by Corwin Hanchin 1961.^[220] Since then, QSA/PR methodology has evolved from simple regression models to models capable of analysing very large numbers of molecules with different structures using machine learning techniques. QSA/PR models are based on the principle that two molecules with similar structures show comparable behaviour under similar environmental conditions and the differences in their behaviour can be explained by the variations in their structure. Thus, if one knows the activity or other property of a large number of molecules, one can predict the activity or other property of a structurally similar molecule without the need for experimental data.^[221] QSA/PR models have been widely used in research, especially in drug development and in assessing the harmfulness of substances, materials or nanomaterials to human health and the environment. QSA/PR models are nowadays also used as guidelines for the development of regulatory measures, especially for the assessment and authorisation of the use of certain chemicals.^[222]

QSA/PR models are developed using a statistical regression analysis that estimates the relationship between dependent (Y) and independent (X) variables. In QSA/PR models, dependent or response variable (Y) is any type of molecular response derived from experiments or calculations that depends on the structural properties of the molecule. The response variables are often referred to as endpoints. The model can only be as good as its endpoints, so these must be both accurate and precise.^[223] The independent variables (X) are the structural properties of the molecule in numerical form, known as descriptors. The number of endpoints should be as high as possible in order to find the hidden relationship between the response variables and the structural features of the corresponding molecules. The "rule of thumb" in QSA/PR modelling is that the number of endpoints should be at least five times greater than the number of descriptors used in a model. It is also desirable to have a wide range of endpoint values to cover a variety of molecules so that the resulting model has a good range of applicability and is able to reliably predict the responses of new, untested molecules. To evaluate the predictive power of a model, the data set is usually split into two parts:

- training set used to develop a model,
- test set used to test the predictivity of a model.

The values of the responses in the test set should lie within the range of the response values of the training set in order to avoid extrapolation of the data when estimating the predictive power of a model. After the regression analysis, it is important to test the resulting model for its goodness-of-fit, robustness and predictive power using statistical validation methods. In addition to the statistical prediction of data, some models can also be used to draw conclusions about the underlying mechanism that results from the description of the most influential descriptors of the model.



Figure 2.11. A general scheme of QSA/PR model development.

2.3.5.1. Multiple linear regression analysis

Many different multivariate regression analyses are used in QSA/PR modelling, e.g., Multiple Linear Regression, Principal Component Regression, Partial Least Squares, Artificial Neural Nets and other.^[224] Multiple Linear Regression using ordinary least squares is still the most used method today.

Multiple linear regression (MLR) assumes the linear relationship between two or more independent variables (*x*) and dependent variable (*Y*).^[225] The general MLR model can be written as:

$$Y = a + b_1 x_1 + b_2 x_2 + \dots + b_k x_k + \epsilon$$
(2.78)

where ϵ is a random error with a normal distribution and mean that equals 0 ($\epsilon \sim N(0, \sigma^2)$). *a*, *b*₁, *b*₂, ..., *b*_k are regression parameters, and *x*_i, *i* = 1,2, ..., *k*, are independent variables, descriptors. The values of the regression parameters are most often determined by the ordinary least square (OLS) method that minimizes the square of the residuals, that is, it finds the minima of the following function:

$$f(\hat{a}, \hat{b}_1, \dots, \hat{b}_k) = \sum_{j=1}^n \left[y_j - \left(a + b_1 x_{1j} + b_2 x_{2j} + \dots + b_k x_{kj} \right) \right]^2$$
(2.79)

In functional minima, the partial derivation for all parameters equals 0. The further expansion leads to the set of normal functions:

The result of solving the set of normal functions are the values of \hat{a} , \hat{b}_1 , \hat{b}_2 , ..., \hat{b}_k which correspond to the estimated values of a, b_1 , b_2 , ..., b_k .

2.3.5.2. Fitting criteria

The goodness-of-fit of the QSA/PR model is assessed by calculating the statistical fit parameters. In the following equations, n stands for the total number of endpoints, p for the total number of descriptors (independent variables), and \bar{x} and \bar{y} for the mean values of variables x and y, i.e.:

$$\bar{x}_k = \sum_{j=1}^n \frac{x_{kj}}{n}$$
 $\bar{y} = \sum_{j=1}^n \frac{y_j}{n}$ (2.81)

The residuals (*e*) are the difference between the values of the response variables *y* and the ones estimated by the model \hat{y} for the given variables *x*, i.e.:

$$p_i = y_i - \hat{y}_i \tag{2.82}$$

The standard error of estimate (*s*) is calculated as:

$$s = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n - p - 1}}$$
(2.83)

Correlation coefficient (R) measures the linear correlation between two sets of data as:

$$R = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}}$$
(2.84)

The values of the correlation coefficient range from -1 to 1. Values closer to -1 or 1 indicate a strong negative or positive linear correlation, while as the values close to 0 indicate that there is no linear correlation between the data sets. The squared value of the correlation coefficient, the so-called coefficient of determination (R^2), is usually used to assess the goodness of fit. R^2 indicates the ratio of the variation in *y* values that can be predicted by a linear model and is calculated as follows:

$$R^2 = \frac{MSS}{TSS} = 1 - \frac{RSS}{TSS}$$
(2.85)

where *MSS* is the model sum of squares, *RSS* is the sum of the squares of the residuals, and *TSS* is the total sum of the squares of the difference between the response and the means, calculated as:

$$MSS = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$$
(2.86)

$$RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$
(2.87)

$$TSS = \sum_{i=1}^{n} (y_i - \bar{y})^2$$
(2.88)

Although adding more dependent variables (descriptors) to the model leads to higher R^2 values, the models with a large number of descriptors can often lack the ability to reliably predict the responses for the similar data set. To avoid overfitting, the adjusted coefficient of determination (R_{adj}^2) can be used as it also considers the number of descriptors in the model. The R_{adj}^2 is calculated as follows:

$$R_{\rm adj}^2 = 1 - \frac{n-1}{n - (p+1)} \times \frac{RSS}{TSS}$$
(2.89)

The Fisher statistics (F) evaluates the statistical significance of the regression model. The F-value is the ratio between the variation that can be explained by the model for a given data set and the variation that cannot be explained by the model, and is calculated as:

$$F = \frac{MSS}{RSS} \times \frac{p}{n - p - 1}$$
(2.90)

The statistical parameters explained above are used to select a model that describes the relationship between response variables and descriptors well. Internal and external validation methods are used to further assess the robustness and predictive power of the model.

2.3.5.3. Model validation

The robustness and stability of the model can be tested by methods of internal validations.^[226] Leave-One-Out (LOO) and Leave-Many-Out (LMO) are cross-validation methods that place one or more molecule from a training set to a test set and then calculates the new QSAP/R model. LOO omits one molecule from the training set and builds a new model with n - 1 endpoints. The method repats this process each time omitting a different molecule. For each new model, the response of the omitted molecule is predicted, \hat{y}_i . Finally, the coefficient of cross validation (Q^2) is calculated as:

$$Q^2 = 1 - \frac{PRESS_{\rm CV}}{TSS} \tag{2.91}$$

Where *TSS* is the total variation in *y* values and *PRESS*_{CV} is the variation in predicted values \hat{y} of the molecules placed in the test set, calculated as:

$$PRESS_{\rm CV} = \sum_{i}^{n} \left(y_i - \hat{y}_{\frac{i}{t}} \right)^2 \tag{2.92}$$

 Q^2 depends on the number of descriptors. If descriptors that are significant to the model are added, the values of Q^2 will increase. Otherwise, it will decrease.^[223] LMO method of cross-validation places a random set of molecules in a test set and builds a QSA/PR model with the remaining molecules in a training set. LMO can leave out a various number of molecules from the training set and therefore it is possible to have more iteration in comparison with LOO method. The maximum number of molecules LMO can leave out is 50% of the total number of molecules. If the model is stable and robust, the LMO method will produce a set of similar QSA/PR models. *Y*-scrambling is the internal validation method used to test whether the obtained regression model is just a result from the random correlation between variables or is there a real linear relationship between them. In this method, the values of response variables *y* are scrambled and randomly assigned to other molecules. The new QSA/PR models are then developed using a scrambled data. The process of scrambling and developing a model is repeated a few hundred times. The R^2 or Q^2 values of all the models are compared. If the original model is not a result of random correlation, the new models with mixed data should have a significantly lower values of R^2 and Q^2 .

The external validation is used to test a predictive power of a model. The responses of the test molecules are calculated by including the values of their descriptors X in the model equation. The model is then validated by calculating different statistical parameters. There are many different criteria to validate a predictive power of a model. Usually, the above-mentioned equations for fitting criteria are also used for external validation with the difference that they use values of responses y_i and predicted values \hat{y}_i of a test set, and not a training set. The concordance correlation coefficient *CCC* is a proposed method for the external validation as it is more precautionary in accepting QSA/PR models. *CCC* is calculated as follows:

$$CCC = \frac{2\sum_{i=1}^{n} (y_i - \bar{y}) (\hat{y}_i - \bar{\hat{y}})}{\sum_{i=1}^{n} (y_i - \bar{y})^2 + \sum_{i=1}^{n} (\hat{y}_i - \bar{\hat{y}})^2 + n(\bar{y} - \bar{\hat{y}})^2}$$
(2.93)

Where y_i is the response of a molecule from the test set, \bar{y} is the mean value of the responses from the test set, \hat{y}_i is the predicted value of the response of molecule from the test set, and \bar{y} is the mean of predicted values of the responses from the test set.^[223]

2.3.5.4. Molecular descriptors

Descriptors are chemical characteristics of a molecule in numerical form, which are used as independent variables in QSA/PR. Molecular descriptors can be experimentally measured physicochemical properties of the molecule (e.g. octanol-water partition coefficient, polarizability, molar refractivity) or theoretically derived from the structural representation of molecules.^[227] With the increasing popularity and applicability of the QSA/PR methodology, the need for new, well-defined descriptors capable of capturing some specific chemical information from the molecular structure is also growing. Several thousand descriptors have been proposed so far.^[228] Various software packages can be used to quickly retrieve the molecular descriptors from the

provided molecular structure, e.g. Dragon software can be used to calculate about 5000 descriptors.^[229] According to the standardised rules, all descriptors must:^[230]

- be invariant to atom labelling and numbering,
- be invariant to the molecule roto-translation,
- be defined by an unambiguous algorithm,
- have a well-defined applicability on molecular structures.

This means that a descriptor must be well-defined, should not depend on manipulations that do not affect the molecular structure, and that it should be known to which class of molecules it can be applied. Furthermore, it is desirable that a descriptor is simple and easy to interpret, possibly related to an experimental property but not derived from experiments, and able to account for small variations in molecules, as in isomers.

Molecular descriptors can be categorised into 0D, 1D, 2D, 3D and 4D descriptors according to different representations of a molecular structure.^[230] 0D and 1D descriptors are constitutional descriptors that do not contain information about the topology of molecules, such as the type and number of atoms, the number of bonds and the molecular weight of a molecule. While the 0D representation is the simplest chemical formula, the 1D representation contains a list of the structural fragments of a molecule. 2D descriptors or topological descriptors are derived from a 2D representation of a molecule, i.e. a molecular graph. 2D descriptors consider the connectivity of the atoms in a molecule. 3D descriptors or geometry descriptors take into account the arrangement of atoms in a three-dimensional space. The 3D structure of a molecule should be optimised before calculating 3D descriptors. 4D descriptors additionally take into account the electron distribution and the interactions of the molecule with the space around the molecule or the multiple conformational states of a molecule.^[231]

To create a valid QSA/PR model, it is important to select only a small set of descriptors that are relevant to the modelled activity or property. The selection process is quite difficult and various computational methods have been developed to speed up and improve the selection, with the genetic algorithm (GA) being the most popular method.^[232]

2.3.5.5. Genetic algorithm (GA)

The genetic algorithm (GA) is an optimisation method used to select significant descriptors.^[233] The GA mimics the principles of evolutionary biology by representing a subset of variables as a binary string called a chromosome. The value of a particular position on a chromosome (gene) can either be 1 if the variable is included in the model, or 0 if the variable is not included in the model. Each chromosome represents a model with its own set of variables. A set of chromosomes forms a population.^[234] The basic steps of the GA methodology are as follows:^[235]

- 1. generation of the initial population,
- 2. crossover operation,
- 3. mutation operation,
- 4. comparison operation.

In the first step, the models are created with randomly selected variables. The models with the highest fitness values are included in the elite population. Subsequently, well-fitting models are selected as "parent" models and the crossover operations are performed to create new models. The probability of being selected as a parent model is inversely proportional to the fitness value of a model. During mutation, all individuals are randomly changed. All new models and the elite population are compared with each other, and the models with the highest fitness score are included in an elite population. The GA steps are repeated until convergence is reached, i.e. until the overall fitness of the elite population can no longer be improved.

3. Methods

3.1. Computational details

3.1.1. DFT calculations

All calculations were performed with Gaussian 16 software for solving the electronic structure equations.^[236] Geometry optimizations and interaction energy computations have been performed with M06-L combined with def2-SVP basis set. Vibrational frequencies were calculated to ensure that the studied systems are true minima of the corresponding potential energy surfaces. Grimme's 3D empirical dispersion and density fitting were used in all computations.

Analysis of intermolecular weak interactions was performed in IGMPlot-3.08 program^[237] using Independent gradient model based on Hirshfeld partition of molecular density (IGMH)^[238,239] and Visual Molecular Dynamics (VMD) program^[240] to generate colour filled isosurface.

3.1.2. QSA/PR formulation

The molecular descriptors were calculated by Dragon 6.0. software^[241] that uses empirical and semi-empirical methods to calculate the structural features of chosen molecules. The models were developed using QSARINS 2.2.4 software; the calculated energies (E_{int}) were set as the response variables and the Dragon descriptors were used as independent variables. Genetic algorithm (GA) method^[233] was used to select the most relevant descriptors and the parameters were set as following: 100 random models, generation size of 500 interaction and the mutation probability specified as 20%. Multiple Linear Regression Analysis (MLRA) method was used to develop models with 1-5 variables.

Models were selected and validated by comparing the common statistical parameters such as: coefficient of determination (R^2), the leave-one-out cross-validation coefficient (Q^2_{LOO}), *F*-ratio between the variances of observed and calculated property (*F*), probability value for calculated *F* (*p*), standard error (*s*), and standard error of the predictive residue of sum of squares (S_{press}). The validation of models for each response variable was also performed by Leave Many Out (LMO) and "*Y*-scrambling" test. Williams plots were used to study the applicability domain (AD) and determine the possible response variable (*Y*) and structural (*X*) outliers.^[223]

3.1.3. Molecular dynamic simulations

All molecular dynamics simulations were carried out using GROMACS 2022.1. with the OPLS-AA force field. Both assembly and adsorption simulations were conducted in the NPT ensemble using the leapfrog algorithm. The pressure was maintained at 1 bar using the isobaric Parrinello-Rahman barostat with a coupling constant of 2.0 ps, while the temperature was set to 300 K and regulated with the velocity-rescale algorithm, applying a coupling constant of 0.1 ps. Periodic boundary conditions (PBC) were applied in all three dimensions, and the systems were solvated using the SPC water model. A cutoff value of 1.2 nm was used for both van der Waals and electrostatic interactions, with the latter accounted for using the particle-mesh Ewald (PME) summation method. Bonds involving hydrogen atoms were constrained using the LINCS algorithm, which allowed the use of an increased time step of 2 fs. The simulations were run for 10 ns, and snapshots of the systems were rendered using VMD.

3.2. Experimental details

3.2.1. Materials

Polyethylene terephthalate (PET) granules were obtained from Sigma Aldrich (Saint Louis, MO, USA). Pesticides alachlor (ALC; C14H20ClNO2, 99.8%), diuron (DIU; C9H10Cl2N2O, 99.6%) and isoproturon (IPT; C12H18N2O, 99.8%) were purchased from Sigma-Aldrich (Burlington, MA, USA).

3.2.2. Microplastic preparation

PET MP were prepared from PET granules. PET granules were melted and thermally aged at a temperature of 240 °C for 24 h in a drying oven (Memmert UN55, Germany). Pristine granules and thermally aged PET samples were ground in a cryomill (CryoMill, Retsch, Germany) and then sieved by a vibrate sieve (AS 200 basic, Retsch, Germany) to isolate the fraction sized 100–200 μ m that was used in further experiments.

3.2.3. Microplastic characterisation

Scanning electron microscopy (SEM) was used to analyse the surface morphology of pristine and aged PET MP using a Vega III scanning electron microscope (SEM) (Tescan, Brno, Czech Republic). The specific surface area of pristine and aged PET MP was determined by the multiple BET method (Micromeritics, Gemini) using nitrogen gas as the adsorbate.

X-ray photoelectron spectroscopy (XPS) was performed on both pristine and aged PET MP samples with and without adsorbed organic pollutants to evaluate the effect of thermal aging on the PET MP surface and its interaction with pollutants. The samples were mounted on adhesive graphite tapes and analysed without applying a conductive coating via vapor phase deposition. Measurements were carried out using a PHI VersaProbe III (Version AD) system (PHI, Chanhassen, MN, USA) equipped with a hemispherical analyser and a monochromatic Al K α X-ray source. Survey spectra were acquired at a pass energy of 224 eV with a step size of 0.8 eV. Data acquisition was performed using ESCApe 1.4 software, and spectral fitting of the C 1s level was carried out using CasaXPS 2.3.15 software.

3.2.4. Adsorption experiment

Solutions of IPT, ALC and DIU (50 μ mol dm⁻³) were prepared in ultra-pure water ($\rho = 18 \text{ M}\Omega \text{ cm}$, EMD Millipore, Burlington, MA, USA). A 1 ml aliquot of adsorbate solution was added in an Eppendorf microtube containing 100 mg of prepared PET MP. The microtubes were then placed in an OLS Aqua Pro orbital shaker (Grant, Royston, UK) and incubated at 25 °C with a shaking speed of 120 rpm for one week. All experiments were conducted in triplicate. Control samples contained only the adsorbate solution without PET MP.

After one week, the samples were filtered through 0.45 µm PTFE filter (Chromafil, Macherey-Nagel, Düren, Germany). The remaining adsorbate concentration in the filtered solution was analysed using high-performance liquid chromatography (HPLC, LC20, Shimadzu, Japan) coupled with UV/DAD detector (SPD-M20A, Shimadzu, Japan).

4. Results and discussion

4.1. Ab initio study on the role of functional groups in PET MP-pollutants interactions

In this part of the thesis, the study focused on understanding how different functional groups, commonly present in water pollutants, such as pesticides, pharmaceuticals and personal care products influence adsorption interactions with microplatic. The study aimed at analysing the interactions on molecular level employing the tools of computational chemistry, specifically DFT methods, to provide deeper insights into the nature of these interactions.

PET MP was selected due to its widespread presence in the environment and its significant impact on both ecological systems and human health.^[14,15] The molecular structure of PET was chosen to enable the study of specific interactions such as hydrogen bonds or van der Waals interactions across multiple adsorption sites, representing the non-homogeneous surface characteristics of real-world PET MP. The organic molecules examined in this study were selected through a careful analysis of the structures of common pharmaceuticals, pesticides, and personal care products, as depicted in Figure 4.1. This group of potential water pollutants includes compounds containing aromatic rings with various functional groups. Hydroxyl groups are present in pollutants such as 17β -estradiol, amoxicillin, nonylphenol, benzylparaben, butylparaben, and bisphenol A. Pollutants containing carboxyl groups, among others. Amide functional groups are part of oxytetracycline and carbamazepine, while trimethoprim and procainamide contain amine groups. Chlorine is a component of diuron, triclosan, and triclocarban. Dicamba, a pesticide, contains a benzene ring with carboxyl and methoxy groups along with two chlorines. Atrazine and simazine, common pesticides, are also included due to their triazine structures.



Figure 4.1. Structures of common organic pollutants in water.

This group of water contaminants shares several functional groups expected to play a significant role in their interactions with the surface of microplastic particles, particularly those with pristine or aged forms featuring oxygen-containing groups such as hydroxyls, carbonyls, and carboxyls.

These functional groups are represented by the simplified molecules shown in Figure 4.2. For computational efficiency and to preserve the essential chemical interactions, these reference molecules were considered as model water pollutants.



Figure 4.2. Organic pollutant molecular models.

4.1.1. Calculation of weak interactions

The influence of different functional groups was analysed by comparing the interaction energies between chosen water pollutant models and model PET. The interaction energies are determined using the following equation:

$$E_{\rm int} = E_{\rm mPET/pollutant} - E_{\rm mPET} - E_{\rm pollutant}$$
(4.1)

Where $E_{mPET/pollutant}$ represents the total energy of the model PET/pollutant complex, E_{mPET} the energy of model PET unit, and $E_{pollutant}$ the energy of the pollutant unit. It is important to note that the stronger interactions are indicated by more negative E_{int} . The interactions between PET MP and organic molecules are already described in the literature as weak physical interactions. As outlined in Section 2.3.3.2., one of the challenges in calculating weak interaction energies is the basis set superposition error (BSSE). BSSE is a mathematical error that arises due to the use of a finite basis set in calculations. To address this issue, the classical counterpoise correction (CP) method was applied. The CP method evaluates the energy of each unit within the complex using the basis set of the entire complex, where the atoms of the other unit are represented as "ghost atoms". These ghost atoms are characterized by the absence of nuclear charges and electrons, while the geometry of the individual units remains unchanged from that in the complex.

4.1.2. Method performance

Prior to modelling adsorption interactions and calculating weak intermolecular interactions, it is essential to evaluate the performance of DFT functionals, considering both their accuracy and computational cost. DFT functionals are designed with specific applications in mind, making the choice of the most suitable functional highly dependent on the studied system.^[167] A well-known limitation of DFT is its inability to account for dispersion interactions,^[162] which are crucial in systems where weak interactions dominate. Dispersion forces arise from the correlated fluctuations of electron density between atoms or molecules, exhibiting non-local character as they depend on the electron correlation over long distances.^[242] In contrast, most DFT methods primarily rely on local or semi-local approximations, making them insufficient for accurately capturing such interactions. These limitations can lead to significant errors in systems where weak interactions dominate.^[243]

As known from the literature, the adsorption of organic pollutants to microplastic is controlled by weak interactions.^[22,67] Consequently, addressing these limitations is essential for accurately modelling such systems. This requires a special care when selecting appropriate DFT functional by assessing the performance against more accurate, reference method. In this work, the performance of DFT functional was tested with respect to second order Møller-Plesset perturbation theory (MP2) method, which is commonly used to calculate non-covalent interactions with a reasonably accurate results for small systems.^[244,245] Several functional were considered, including B3LYP and Minnesota functionals (M05, M05-2X, M06, M06-2X, M06-HF and M06-L) from Truhlar group as a possible candidates. B3LYP was considered as one of the most popular and widely applied hybrid functional which combines Hartree-Fock exchange with gradient-corrected correlation,^[166,246–248] and is also found used in calculations of weak interactions.^[132] Minnesota functional are specially developed to describe dispersion interactions. M06-L is a non-hybrid functional from Minnesota family of DFT functionals developed to treat non-covalent interactions, thermochemistry, and thermochemical kinetics more accurately. One of the notable strengths of M06-L is its robust performance across a diverse range of molecular systems and properties, apart of the fact that is considerably less computationally expensive than conventional hybrid functionals. Moreover, local functionals allow the use of density fitting for basis set to improve the calculations of Coulomb forces.^[249–251] This all makes M06-L functional a very promising candidate for further calculations. Also we considered M05,^[252] known for its accurate treatment
of non-covalent interactions, M05-2X,^[253] the range-separated hybrid exchange-correlation functional and M06^[249] functional optimized for main group thermochemistry, thermochemical kinetics, and non-covalent interactions. Finally, we considered M06-2X,^[249] a global hybrid functional with increased Hartree-Fock exchange, and M06-HF.^[254] All DFT calculations were performed using Grimme's D3 empirical dispersion correction^[255] which improves the results of DFT calculations for weakly bound systems without additional cost.^[256]

To test the method performance, the interactions were calculated between the terephthalic acid and ten systems of small organic pollutant molecules chosen previously in two different configurations represented in Figure 4.3. In the first configuration (TA I), the functional group of the pollutants interacts with the carboxyl group of the terephthalic acid, and in the second configuration (TA II), the aromatic ring of the pollutants is located above the aromatic ring of the terephthalic acid.

Terephthalic acid and the model pollutants



Figure 4.3. Structural representation of terephthalic acid (TA) and the model pollutants in two configurations. The blue sphere represents the position of model pollutant's functional group.

MP2 method with smaller basis set generally underestimates non-covalent interactions, and the use of a larger basis set typically improves the results for hydrogen bonded systems, but can lead to overbinding of systems that interact mainly by dispersion forces.^[245,257] For that reason, we used three basis sets def2-SVP, def2-TZVP and def2-TZVPP to calculate energy using MP2 method. The interaction energies and basis set superposition errors (BSSE) are shown in Figure 4.4. The difference between the interaction energies calculated using def2-SVP and def2-TZVP averaged 4.5 kcal mol⁻¹ for TA II, and 1.3 kcal mol⁻¹ for TA I. Further expanding the basis set to def2-TZVPP increased interaction energies only by 0.3 kcal mol⁻¹ on average which is not a significant difference. In contrast, the difference in the energies calculated by M06-L method with def2-SVP and def2-TZVP averaged 0.3 kcal mol⁻¹, as DFT methods tend to converge faster than MP2 with respect to basis set. The same trend is observed for BSSE. Therefore, we used MP2





Figure 4.4. a) Interaction energies and b) basis set superposition error calculated for the interactions of the model organic pollutants and terephthalic acid in two configurations (TA (upper) I and TA II (lower)) using MP2 method.

The interaction energies calculated by MP2/def2-TZVP method and DFT method with def2-SVP basis set are shown in Figure 4.5 **a**), and the mean average errors (MAE) and root mean square errors (RMSE) of DFT methods compared to MP2 are shown in Figure 4.5 **b**). For configuration TA II, the interaction energies calculated by DFT methods were generally lower than those obtained using MP2/def2-TZVP. However, the variations among the functionals in this configuration were not substantial enough to determine the most suitable functional based on these values alone. The differences between functionals became more evident in configuration TA I. In this case, the functional groups of the model organic pollutants interact with the carboxylic group, forming stronger interactions such as hydrogen bonds. Overall, all the functionals tend to overestimate interaction energies, with this overestimation being particularly evident for systems with stronger adsorption energies. This effect is most pronounced for the M06-HF functional, which significantly overestimates interactions where hydrogen bonds can be formed.^[260] More detailed information about the difference between interaction energies for each system are shown in Figure A 1 and Figure A 2.

Finally, considering both configurations, M06-L demonstrated the most similar results to MP2, making it a strong candidate for further analysis. While some other functionals, such as M05, M06, M06-2X also showed reasonably good performance, the differences between these methods and MP2 were not significant enough to decisively recommend one over the others. However, M06-L stood out not only for its reliable accuracy in reproducing MP2-calculated interaction energies but also for being a pure functional, which makes it computationally less demanding compared to hybrid functionals. This computational efficiency, combined with its ability to balance accuracy and cost, was a critical factor in its selection. Given these considerations, all subsequent calculations were performed using M06-L/def2-SVP method with density fitting and D3 empirical dispersion.



Figure 4.5. Method performance of the selected DFT/def2-SVP methods against MP2/def2-TZVP method: a) interaction energies and b) mean average errors (MAE) and root mean square errors (RMSE) of interaction energies calculated by DFT/def2-SVP compared to the energies calculated by MP2/def2-TZVP method. The interactions were calculated between terephthalic acid and the model organic pollutants in two different configurations (TAI (upper) and TAII (lower)).

4.1.3. PET model adsorption sites and size effects

Polyethylene terephthalate (PET) is composed of elongated polymer chains that are held together through noncovalent intermolecular forces.^[261] A study conducted by Cortes-Arriagada et al.^[132,135] point out that the linear model commonly used for molecular modelling of MP has significant limitations. Specifically, it was shown that this model can underestimate the role of dispersion forces in favour of electrostatic forces. Based on these findings, the authors recommended adopting a folded nanoparticle model as a more accurate representation. Therefore, in the present work, PET MP was modelled in a folded form of nanoparticle.

The construction of PET model was initiated by optimizing a starting configuration consisting of two dimers of terephthalic acid. Following each optimization cycle, a monomeric unit was appended to each of the existing chains. The iterative optimization led to several models with 4 to 12 monomeric units. To obtain an estimation of the strength of the interactive forces between two chains, the interaction energies were calculated and presented as a function of the "number of PET units". The results depicted in Figure 4.6 suggest that the interaction energies converge to a value of about -9 kcal mol⁻¹ per unit at the M06-L level of theory.



Figure 4.6. Interaction energies per monomer of two PET chains.

Considering the size and shape of PET models, as well as the number of atoms for the reasons of computational efficiency, a model built of PET consisting of 2 chains with 3 monomers, with 138 atoms, and 620 electrons in total ($C_{60}H_{52}O_{26}$) was selected for further calculations. This model is hereafter referred to as mPET. The surface of mPET was analyzed based on its molecular electrostatic surface potentials (MEP), as shown in Figure 1.9, to identify potential adsorption sites

representing various parts of the PET molecule. Distinct regions of excess negative and positive charges were observed, which are expected to serve as adsorption sites for specific pollutants through electrostatic interactions. Specifically for the PET model under consideration, four potential adsorption sites of interest were identified.

Adsorption site I corresponds to the carboxyl group at the end of the PET chain. This site is less common in real polymers due to the limited number of terminal units compared to the overall surface area. However, the ageing of microplastic in nature may lead to chain breakage and the formation of new oxygen-containing functional groups, increasing the number of sites similar to adsorption site I, which could alter the adsorption capacity for pollutants. Adsorption site II is a benzene ring fused in a PET chain chosen to study the role of π - π interactions. Finally, adsorption sites III and IV represent the surface of PET, far from terminal units. These four adsorption sites are considered to be among the most common on the surfaces of micro-sized particles and are likely to play a key role in determining adsorption behaviour.



Figure 4.7. Structure and molecular electrostatic potential (MEP) of mPET. Red parts represent the excess of negative charge on the surface, and blue excess of positive charge. Adsorption sites I, II, III and IV are circled with orange lines.

To validate the size of the model and assess the effect of model size on pollutant interactions, the interaction energies of phenol, a representative model pollutant, were calculated and compared for PET models of two different sizes. Specifically, the previously selected mPET model (comprising 2 chains with 3 monomers) was compared with a second model twice as large, consisting of 2 chains with 6 monomers and approximately 2 nm in length. The interaction energies between phenol and the two models were calculated for adsorption sites I and II. The adsorption energy of phenol at adsorption site I of the mPET model was found to be -21.30 kcal mol⁻¹, while for the larger model, it was -21.27 kcal mol⁻¹. Similarly, the interaction energies for adsorption site II were -18.63 kcal mol⁻¹ for mPET and -18.71 kcal mol⁻¹ for the larger model as shown in Figure 4.8. These differences in interaction energies are minimal, indicating that the size of the model does not significantly affect the interactions. Therefore, the mPET model is considered valid for studying these interactions, as the results suggest that the interaction energies are largely independent of the model size. The Cartesian coordinates of the optimized final geometry are listed in Table A 6.



Figure 4.8. The effect of model PET size on the interaction energy of phenol and a) adsorption sites I and b) adsorption site II calculated by M06-L/def2-SVP method.

4.1.4. Molecular interactions

With the carefully chosen mPET structure and its identified adsorption sites of interest, the next step was to calculate their interactions with model organic pollutants. Additionally, the affinity of water molecules for the adsorption sites was examined to compare their interactions with those of the pollutants. These interactions were evaluated both in a vacuum and in implicit water using IEFPCM solvation model to examine the influence of the polarization field on the interaction energy.

The results are shown in Figure 4.9. Upon a detailed analysis of the data corresponding to the adsorption site I (Figure 4.9 a)) which refers to the terminal carboxyl group, it becomes evident that one could classify the considered model pollutants into three groups based on their interaction energies with adsorption site I. The first group consists of phenylacetic acid, benzoic acid, and benzamide, which are capable of forming two strong hydrogen bonds with the terminal carboxyl group. Their interaction energies with adsorption site I are about -31 kcal mol⁻¹, which is much stronger than the interactions between adsorption site I and the water molecule. Therefore, it is reasonable to speculate that these model pollutants could, to a certain extent, adsorb onto PET in an aqueous environment. Aniline, phenol, anisole, triazine, and chlorotriazine can be placed in the second group because their interaction energies with adsorption site I are approximately -20 kcal mol⁻¹. These energies are comparable to the interaction energy of a water molecule and adsorption site I, indicating a weaker affinity for the PET surface in competition with water. This similarity suggests that adsorption of these pollutants onto PET in an aqueous environment may be less favoured compared to the first group. Benzene and chlorobenzene are in the last group, since their interaction energies with adsorption site I are only about -11 kcal mol⁻¹, which is about two times weaker than the interaction of the water molecule. When comparing the interaction energies calculated in vacuum with those obtained using the PCM solvation model, a general weakening of the interaction energies is observed. For the first group, the interaction energies are reduced by approximately 5 kcal mol⁻¹, for the second group by around 2 kcal mol⁻¹, and for the third group by about 1.5 kcal mol⁻¹. Notably, the interaction energy of water remains unchanged under the PCM model. Despite these differences, the overall trend in interaction strengths among the three groups remains consistent.

The results for adsorption site II are shown in the Figure 4.9 b). This site is related to a benzene ring with a hydroxyl terminal group in its vicinity. The classification of model organic

pollutants to categories is not so evident as for adsorption site I. The average energies of phenylacetic acid, benzoic acid and benzamide, that exhibit the strongest adsorption, are about $-20 \text{ kcal mol}^{-1}$, which is only about 3 kcal mol⁻¹ stronger than the average interaction energies of aniline, phenol, and anisole. In contrast, triazine and chlorotriazine exhibit even weaker interaction energies, averaging around $-12 \text{ kcal mol}^{-1}$ which is comparable to the interaction energy of the water molecule. For benzene and chlorobenzene, the average interaction energies with adsorption site II are around $-10 \text{ kcal mol}^{-1}$.

Turning attention to adsorption sites III and IV, which are located on two separate parts of the mPET surface "far" from the terminal groups. At these sites, hydrogen bonding can occur exclusively between the proton donor groups present in model pollutants and the oxygen atoms of ester groups in the mPET structure. Interaction energies of water molecules and the latter adsorption sites III and IV are about –9 kcal mol⁻¹ for III and –11 kcal mol⁻¹ respectively as illustrated in Figure 4.9 c) and d). A consistent trend is evident across both sites. That is, model pollutants carrying proton donor groups, such as phenylacetic acid, benzoic acid, benzamide, aniline and phenol, demonstrate average interaction energies that are approximately 6 kcal mol⁻¹ and 4 kcal mol⁻¹ stronger than that of the water molecule at the adsorption sites III and IV, respectively. On the other hand, model pollutants lacking a proton donor group exhibit interaction energies that are either similar to or weaker than the interactions between the water molecule and these adsorption sites. In PCM, interaction energies change for 1.6 kcal mol⁻¹ in average, the overall trend remains unaltered.

This analysis shows the critical role of hydrogen bonding capabilities in determining adsorption behaviour at different sites on the mPET surface. It also highlights the relative importance of proton donor groups for enhancing interactions. These findings provide valuable insights into the factors governing pollutant adsorption on PET surfaces.



Figure 4.9. Interaction energies of the model pollutants and mPET on a) adsorption site I, b) adsorption site II, c) adsorption site III, and d) adsorption site IV calculated in vacuo and in implicit water as a solvent using IEFPCM model. Dashed lines represent the interaction energies of water molecules.

4.1.5. Visual study of weak interactions

Further analysis can be conducted through the visual study of weak interactions, allowing for the graphical representation of these interactions within chemical systems. This approach reveals the primary regions where interactions occur and distinguishes between different interaction types. These methods are extensively employed in studies of weak interactions due to their convenience and great universality. The non-covalent interaction (NCI) method proposed by Johnson et al. is based on reduced density gradient (RDG).^[262] The basis of this method is briefly outlined here, as it serves as a foundation for other methods of this kind. RDG is a dimensionless quantity defined as follows:

$$RDG(r) = \frac{1}{2(3\pi^2)^{1/3}} \times \frac{|\nabla\rho(r)|}{\rho(r)^{4/3}}$$
(4.2)

where $\rho(r)$ is electron density and $\nabla \rho(r)$ gradient of electron density. In regions far from the molecule, where the density decays to 0, RDG will have a very large positive value. On the opposite, in the regions of covalent bonds and weak interactions, RDG(r) values will be very small, approaching zero. Weak interactions are formed at a low density values between two monomers. Upon dimer formation from two distinct monomers, RDG(r) value quickly changes from very large value to near zero value. This can be seen as spikes in plots RDG(r) versus $\rho(r)$. ^[262] To distinguish interaction types, the NCI method utilizes the sign of the second derivative of the Laplacian of the electron density $(sign(\lambda_2))$. Bonding interactions are characterized by a negative $sign(\lambda_2)$ while nonbonding interactions are identified by a positive sign(λ_2). More specifically, the Laplacian of the density $(\nabla^2 \rho)$ is often decomposed into three components along the principal axes of maximal variation, represented by three eigenvalues ($\lambda_1 \leq \lambda_2 \leq \lambda_3$). In bonded interatomic regions, one eigenvalue is positive, and the other two are negative. For covalent interactions, the dominant negative contributions result in an overall negative Laplacian, whereas for weak non-covalent interactions, the dominant positive contributions lead to a positive Laplacian. Additionally, bonding weak interactions are associated with a negative λ_2 , while nonbonding are identified by a positive λ_2 in the interatomic region. This characteristic allows for the differentiation of interaction types.^[262] In weak interaction visualization analyses, it is common to employ a colour scheme to represent various interaction types, as illustrated in Figure 4.10. Plots of RDG(r) versus $\operatorname{sign}(\lambda_2)\rho(r)$ for water and water dimer are shown in Figure 4.11. where it can clearly be seen the presence the "spike" in the plot that arises from the formation of weak attractive interactions between dimers.



Figure 4.10. Colouring scheme for representation of weak interactions.



Figure 4.11. Plots of the reduced density gradient (RDG) versus the electron density multiplied by the sign of the second Hessian eigenvalue sign(λ_2) ρ for water and water dimer.

Another widely used method for visualizing non-covalent interactions is the Independent Gradient Model (IGM) method, initially proposed by Lefebvre et al.^[263] The original version of IGM is based on promolecular density, requiring only the system's geometry to visualize interactions. This approach makes the method highly efficient and computationally attractive. Compared to the RDG method, IGM method offers the advantage of distinguishing between intra-fragment and inter-fragment interactions individually. In this method, interactions are using δg function as compared to RDG function used in NCI method. δg function is defined as:

$$\delta g(r) = g^{\text{IGM}}(r) - g(r) \tag{4.3}$$

where $g^{\text{IGM}}(r)$ is IGM type of density gradient defined as the sum of absolute values of atom in free-states, and g(r) is the absolute sum of values of density gradients in their free state. That is, $g^{\text{IGM}}(r)$ and g(r) are defined as follows:

$$g(r) = \left| \sum_{i} \nabla \rho_{i}^{\text{free}}(r) \right|$$
(4.4)

$$g^{\rm IGM}(r) = \sum_{i} \left| \nabla \rho_i^{\rm free}(r) \right| \tag{4.5}$$

In similar way, $\delta g^{\text{inter}}(r)$ function was defined only for intermolecular interaction in a way that:

$$\delta g^{\text{inter}}(r) = g^{\text{IGM,inter}}(r) - g^{\text{inter}}(r)$$
(4.6)

where $g^{\text{IGM,inter}}(r)$ and $g^{\text{inter}}(r)$ are defined as follows:

$$g^{\text{inter}}(r) = \left| \sum_{A} \sum_{i \in A} \nabla \rho_i^{\text{free}}(r) \right|$$
(4.7)

$$g^{\text{IGM, inter}}(r) = \sum_{A} \left| \sum_{i \in A} \nabla \rho_i^{\text{free}}(r) \right|$$
(4.8)

where A loops over all fragments, and *i* loops over all atoms in the fragment. And finally, δg^{intra} is defined as:

$$\delta g^{\text{intra}}(r) = \delta g(r) - \delta g^{\text{inter}}(r)$$
(4.9)

In IGM analysis, sign $(\lambda_2)\rho$ function is mapped on δg^{intra} and δg^{inter} isosurface to reveal the regions and type of weak interactions. As already stated, IGM analysis is based only on density of atoms in their free state. Tian Lu and coworkers proposed a new version of IGM based on Hirshfeld partition of molecular density (IGMH).^[238,239] The density from atom *i* is this method is calculated as:

$$\rho_i^{\text{Hirsh}}(r) = \rho(r)w_i(r) \tag{4.10}$$

where $\rho(r)$ is the actual density calculated by quantum mechanics or obtained experimentally, and $w_i(r)$ is Hirshfeld weighting function of the atom *i* defined as:

$$w_i(r) = \frac{\rho_i^{\text{free}}(r)}{\sum_j \rho_j^{\text{free}}(r)}$$
(4.11)

IGMH offers several advantages over IGM. Most notably, it is defined using the actual electron density, providing a more rigorous physical foundation. In many cases, the isosurface of δg calculated by IGM can appear bulgy, and the associated colouring scheme may lead to incorrect conclusions. A detailed comparison of these two methods is provided in the original IGMH analysis paper. Overall, IGMH is generally recommended over IGM whenever possible. However, it is important to note that while IGMH relies on electron density calculations, IGM only depends on

the system's geometry. This characteristic makes IGM a more appealing option for very large systems where computational calculation are more demanding. However, for the systems in hands, IGMH analysis was used to further analyse the interactions between model organic pollutants and mPET.

The interaction energies of the model pollutants and mPET did not show the same trend for all four adsorption sites. In general, the energies were consistent with the strength and number of hydrogen-bonds the model pollutants could form. For instance, the interaction energies of benzamide, phenylacetic acid, and benzene were highly dependent on the adsorption site, while those of benzene and chlorobenzene showed less dependency. To illustrate this, scatter plots of weak interactions obtained from the IGMH analysis for benzamide and benzene with the four adsorption sites of mPET are presented in Figure 4.12. The scatter plot of benzamide interactions with the adsorption site I shows the large contribution of hydrogen bonding. However, that contribution is reduced on the adsorption sites II, III, and IV. This is consistent with the numerical results as the difference between the interaction energies of benzamide with adsorption site I, where the interactions were the strongest, and adsorption site III, where the interactions were the weakest, was about 16 kcal mol⁻¹. In contrast, the scatter plots for benzene interactions are similar across all adsorption sites, as only van der Waals interactions are involved. This also suggests that $\pi - \pi$ interactions, which could possibly form between the benzene ring of adsorption site II and the model pollutant, do not play an important role for the adsorption on mPET, as proposed by some experimental studies.^[264,265] Additional scatter plots of weak interactions for all model pollutants with water molecules and the four adsorption sites are provided in the supporting materials (Figure A 3 - Figure A 5).



Figure 4.12. Scatter plots of IGMH analysis of weak interactions between benzamide and benzene, and the four adsorption sites of mPET.

To check the qualitative validity of the results obtained for selected fractions of organic pollutants and to assess whether the observed trends hold true for more complex molecules, the interaction of a widely known organic pollutant, diclofenac, with mPET was investigated. Diclofenac was chosen due to its relevance as a common environmental contaminant and its structural complexity, which includes a larger, more flexible molecular framework compared to simpler model pollutants. The goal of this analysis was to compare the binding behavior of diclofenac with mPET to that of phenylacetic acid (PAA), a simpler structural fragment, to better understand how molecular size and flexibility influence adsorption interactions.

The results presented in Figure 4.13 point out that for site I, the interaction energy for mPET with diclofenac is comparable to that of PAA. In contrast, at adsorption site II, the binding affinity of diclofenac with mPET markedly surpasses that of PAA, suggesting a notably stronger interaction at this particular site. This pronounced affinity is consistently observed across sites III and IV. The IGMH analysis, shown in the same figure, further substantiates these findings by revealing a complex interplay of attractive forces, van der Waals interactions, and steric repulsions at these sites. Specifically, sites II, III, and IV demonstrate a convergence of multiple non-covalent interactions that are integral to the stabilization of the mPET diclofenac complexes. The results of the IGMH analysis, paired with the interaction energy profiles, indicate that the strength of the binding should be linked to the structural attributes of the pollutants, namely, the size and flexibility of the molecules. Diclofenac, with its larger and more flexible molecular structure, appears to

establish a more extensive network of interactions compared to its PAA fragment. This comprehensive network is likely facilitated by diclofenac's ability to adapt to the topography of the adsorption sites on mPET, forming multiple noncovalent bonds that reinforce the stability of the complex.





Figure 4.13. (top) Interaction energies of mPET and phenylacetic acid compared to interaction energies of mPET and diclofenac calculated by M06-L/def2-SVP method. Yellow/blue columns correspond to interaction energies between mPET and (phenylacetic acid)/(diclofenac). (bottom) Independent gradient model (IGMH) analysis scatter plots of weak interactions between diclofenac and mPET at all four adsorption sites. In the usual colouring scheme, the blue colour represents attractive interactions, the green colour represents van der Waals interactions, and the red colour represent steric effects.

4.1.6. The effect of folded particle model

Finally, it is meaningful to demonstrate, at this point, the difference between folded mPET models and oligomeric chains. This was revealed by comparing the interaction energies of the model pollutants with terephthalic acid (TA) in two configurations, corresponding to adsorption sites I and II, as shown in Figure 4.3. All the details of those interactions are shown in Figure A 1 and Figure A 3 in Appendix. As seen in the results depicted in Figure 4.14, the difference between the interaction energies of the model pollutants with mPET and TA was negative for all model pollutants, suggesting that they bind stronger on the surface of mPET than on TA.



Figure 4.14. Difference between interaction energies of pollutants with mPET and terephthalic acid on adsorption sites I and II.

This can be easily understood from the visual representations of the intermolecular interactions illustrated in Figure 4.15. As it is observed, the model pollutants can form multiple non-covalent interactions or hydrogen bonds with the neighbouring parts of mPET. This suggests that the adsorption mechanism of pollutants on micro and nano PET is different from the interaction with PET chains, which is consistent with other studies^[132,133] and should be considered in future studies. The same trend is observed for the rest of the systems considered in this work as shown in Figure A 6 – Figure A 9 in Appendix.



Figure 4.15. Schematic representations of the interactions between adsorption site I of mPET and a) benzoic acid; b) benzamide; c) aniline; d) benzene, and between phenol and mPET on e) adsorption site I; f) adsorption site II; g) adsorption site III; h) adsorption site IV analysed by independent gradient model (IGMH) and visualised on 0.01 a.u. isosurface. In the usual colouring scheme, the blue colour represents attractive interactions, the green colour represents van der Waals interactions, and the red colour represent steric effects.

4.1.7. Synergistic and ageing effect

It was interesting to go one step further and analyse how the adsorption of other pollutants on one site on mPET can affect the adsorption on another site. To assess this synergistic affect, the structure of mPET was first optimized with four phenols adsorbed on distant adsorption sites. The optimization process was then repeated by gradually removing phenols and re-optimizing the structure at each step, resulting in a series of optimized structures containing varying numbers of phenols. The interaction energies between each phenol and mPET were calculated for configurations with one, two, and three additional phenols adsorbed on the surface of mPET. The results are shown in Figure 4.16. Upon analysing the obtained data, it becomes evident that the local interactions play a dominant role in determining the binding behaviour. Specifically, the interaction energy for a particular phenol remained largely unaffected by the presence of additional phenols on the mPET surface. This suggests that each phenol's interaction with mPET is primarily determined by its immediate surroundings on the mPET surface, rather than distant phenol molecules. From a quantum chemistry standpoint, this observation indicates that the forces acting at the individual adsorption sites, such as localized electron density shifts via polarization and specific orbital interactions, are crucial in dictating the interaction energy.



Figure 4.16. Interaction energies of phenol and mPET with no other phenols adsorbed and with multiple phenols adsorbed on different parts of mPET surface. Highlighted phenol molecules correspond to ones whose interaction energies are calculated.

Microplastics in the environment are susceptible to photooxidative ageing. Ageing of MP can lead to morphological changes on the surface of MP and the formation of new oxygen containing functional groups such as hydroxyls, carboxyls, and ketones, which consequently can affect adsorption capacity. Experimental studies from the literature have shown that the aged PET MP generally has a higher adsorption capacity for organic pollutants than the pristine PET MP.^[100] In this thesis, aged PET is considered as a modified version of pristine mPET, with additional oxygen-containing groups introduced on each monomer. mPET–OH contains two additional hydroxyl groups on each benzene ring, mPET–COOH has an additional carboxyl group on the carbon atom of the ethylene group, and mPET=O includes an oxygen atom bonded through a double bond to the same carbon atom.

To understand the effects of ageing on the adsorption interactions, interaction energies of phenol with pristine and aged mPETs were calculated at all four adsorption sites. The molecular structures of the modified monomers and the results of interaction energy calculation are shown in Figure 4.17. Some minor variations in the interaction energies of phenol with pristine and aged mPETs were observed, which can be attributed to the structural changes and repositioning of PET chains and phenol molecules during the optimization process. However, no significant differences

in the interaction energies are noted. Therefore, it can be assumed that the formation of oxygencontaining functional groups does not influence the interactions between organic pollutants and PET microplastics.



Figure 4.17. Interaction energies of phenol with pristine (mPET) and aged (mPET–OH, mPET–COOH, and mPET=O) PET model at M06-L level of theory and def2-SVP basis set.

Ageing of PET MP could affect the adsorption capacity by providing more adsorption sites, since the aged particles generally have a larger specific surface area. Additionally, surface oxidation could lead to the formation of more adsorption sites able to form stronger interactions with organic pollutants through hydrogen bonding. As seen from our model, adsorption site I that contains – COOH terminal group showed the strongest adsorption interactions with pollutants containing hydrogen bond donor and acceptor groups compared to other model adsorption sites.

Finally, the interaction energies of twelve phenols adsorbed on the surface of mPET– COOH were calculated. The structure is shown in Figure 4.18. The interaction energies of all twelve phenols ranged from -11.7 to -22.7 kcal mol⁻¹, with an average interaction energy of -17.2kcal mol⁻¹. This range of energies signifies a stable interaction across all twelve phenolic species with the mPET–COOH surface, suggesting that even small nanoparticles of aged mPET possess the capacity to adsorb a considerable number of organic pollutants effectively.



Figure 4.18. The structure of mPET–COOH with twelve phenols adsorbed optimized in gas phase at M06-L level of theory and def2-SVP basis set.

4.2. QSA/PR models of pollutants' functional groups and PET MP interaction energies

The results obtained from DFT provided a model of PET MP that is considered appropriate and offered valuable insights into the nature of its interactions with small model organic pollutants. As the next step in this thesis, the research was expanded to include a broader range of organic pollutants with diverse functional groups. This was done to develop mathematical models capable of predicting interaction energies between new, uncalculated functional groups and the PET model. More importantly, it aimed to provide a deeper understanding of the functional group properties that govern adsorption on mPET. To achieve this, the QSA/PR modelling method was employed. This approach correlates structural properties of molecules, represented by descriptors, with a response variable, which in this case was the interaction energy between mPET and the model organic pollutants.

The set of model organic pollutants (hereafter referred to as m-OPs) was expanded to include 30 molecules, whose structures are shown in Figure 4.19. In QSA/PR modelling, a larger set of molecules is desirable to identify correlations between structural properties and the response variable. The chosen set is considered sufficiently large, as it includes a wide range of functional groups commonly found in organic pollutants, such as pesticides, pharmaceuticals and personal care products. This diversity enables the development of models with up to five variables, respecting the "rule of thumb" in QSPR modelling, which recommends maintaining a ratio greater than 1:5 between the number of variables (descriptors) and the number of molecules in the training set. It is assumed that using up to five variables is sufficient to single out the most significant structural properties of functional groups that guide adsorption.



Figure 4.19. Structures of 30 compounds representing common organic pollutants (m-OPs) to study the adsorption interactions with model PET surface.

4.2.1. Preparing dataset for QSA/PR modelling

The first step involved preparing a dataset of response variables (E_{int}) for the adsorption of all 30 m-OPs onto the four adsorption sites identified on mPET. The calculations were carried out using the M06-L/def2-SVP method with density fitting and Grimme's D3 empirical dispersion, as previously described. The optimized geometries of m-OPs adsorbed on adsorption sites I, II, III and IV on mPET surface are shown in Figure A 11 – Figure A 16 in Appendix, while the calculated interaction energies (E_{int}) are listed in Table 4.1. The interaction energies (E_{int}) on the adsorption site I are in range from -32.69 kcal mol⁻¹ (calculated for benzamide adsorption) to -9.89 kcal mol⁻¹ ¹ (calculated for toluene adsorption) with an average interaction energy of -19.42 kcal mol⁻¹ across all 30 m-OPs. The average interaction energies for the adsorption on sites II, III, and IV are -16.08 kcal mol⁻¹, -13.51 kcal mol⁻¹ and -12.43 kcal mol⁻¹ respectively. m-OPs that exhibited stronger adsorption at the sites (more negative interaction energies), such as benzamide (BAM), phenylsulfonic acid (PhSO₂OH), phenylsulfinic acid (PhSOOH), phenylphosphonic acid (PhPO(OH)₂), and benzenesulfonamide (BSAM), were capable of forming hydrogen bonds with the PET MPs, stabilizing the adsorption process. On the other hand, m-OPs incapable of forming hydrogen bonds, such as toluene (PhMe), benzene (Bz), and styrene (Sty), displayed the weakest adsorption affinities toward the PET MP model, as indicated by their higher interaction energy values. For instance, the lowest interaction energy for PhMe was -9.89 kcal mol⁻¹, whereas for catechol (o-Ph(OH)₂) was -25.03 kcal mol⁻¹. These results align with earlier findings that m-OPs capable of forming hydrogen bonds exhibit stronger interactions with the mPET surface. In contrast, m-OPs lacking proton donor or acceptor groups demonstrate weaker interaction energies across all four adsorption sites on the mPET surface.

The set of independent variables (descriptors) representing the structural properties of m-OPs were calculated and filtered to exclude highly correlated pairs. In total, 2526 descriptors were used to develop predictive models of interactions energies in dependence on the structural features of the m-OPs. Before modelling, the descriptor matrix was normalized by dividing the values of descriptors by the highest absolute value in the training set for each descriptor. The normalization of matrix allows us to study the contribution of each descriptor to the target response in selected model.

Compound	Abbreviation	$E_{\rm int}/~{\rm kcal~mol^{-1}}$			
		Ι	II	III	IV
Benzene	Bz	-10.20	-9.44	-7.06	-8.28
Toluene	PhMe	-9.89	-7.98	-9.01	-9.18
Styrene	Sty	-10.34	-13.00	<u>-11.28</u>	<u>-10.08</u>
Phenylacetylene	PhA	<u>-12.29</u>	-11.90	-11.18	-9.27
Chlorobenzene	PhCl	-11.71	-10.91	-10.75	<u>-8.88</u>
Bromobenzene	PhBr	-11.87	<u>-11.08</u>	-11.48	-8.97
Fluorobenzene	PhF	-12.10	-10.40	<u>-9.37</u>	-8.75
Phenol	PhOH	<u>-21.30</u>	-18.63	-16.59	-15.37
Anisole	PhOMe	-19.21	<u>-16.09</u>	-9.32	-11.53
Benzaldehyde	PhCHO	-20.34	-15.41	-10.63	-9.59
Acetophenone	AcPhO	-21.24	<u>-18.19</u>	-11.49	<u>-13.64</u>
Benzoic acid	BA	-29.43	-20.78	-17.58	-14.17
Phenylacetic acid	PAA	-29.89	-17.25	<u>-16.20</u>	-13.98
Aniline	PhNH ₂	<u>-20.18</u>	<u> </u>	-12.95	-12.79
N-Methylaniline	NMA	-17.28	-16.75	-11.81	-10.23
N,N-Dimethylaniline	DMA	-13.24	-17.05	-13.83	-11.22
Benzonitrile	PhCN	-19.47	-15.69	-13.60	-9.17
Nitrobenzene	PhNO ₂	-18.26	-15.58	-9.68	-12.31
Benzamide	BAM	-32.69	-22.49	<u> </u>	-17.78
Thiophenol	PhSH	<u>-12.99</u>	-15.35	<u>-12.80</u>	-12.25
Phenylsulfenic acid	PhSOH	-19.41	-18.92	-13.71	-14.99
Phenylsulfinic acid	PhSOOH	-30.18	-17.59	-19.43	-18.33
Phenylsulfonic acid	PhSO ₂ OH	-32.04	-25.21	-25.97	-18.49
Benzenesulfonamide	BSAM	-22.99	-22.14	-15.74	-20.58
Phenylphosphonic acid	PhPO(OH) ₂	-29.62	-23.72	-26.66	-19.26
o-xylene	o-PhMe ₂	-11.57	-10.58	-11.29	-9.19
Catechol	o-Ph(OH) ₂	-25.03	-22.44	-17.70	<u>-15.69</u>
o-Phenyldiamine	$o-Ph(NH_2)_2$	-18.95	-18.71	-14.70	-12.39
Triazine	Tr	<u>-18.96</u>	-11.88	-7.79	-7.79
Chlorotriazine	TrCl	-20.05	-11.97	-9.35	-8.65

Table 4.1. Interactions energies between 30 model organic pollutants (m-OPs) and model PET MPs surface on four adsorption sites (I, II, III and IV) used as a response variable for QSPR modelling (**bold underlined** values represent those used for the test sets, while others were part of training sets).

4.2.2. Developing QSA/PR models

Calculated interactions energies were further used as response variables to develop four QSA/PR models, each representing the adsorption on one adsorption site. For each model, m-OPs were first split to training (25 molecules) and test set (5 molecules) as indicated in Table 4.1. The training set was used to construct the QSA/PR models, while the test set was used to validate their predictive power. The splitting was carefully performed to ensure that the response variable values in the test set were evenly distributed and within the range of those in the training set. In accordance with the "rule of thumb," a set of QSA/PR models containing between one and five variables (descriptors) were developed for each adsorption site.

The model with the appropriate number of variables was then selected by comparing the R^2 values of models with 1 to 5 variables for each adsorption site. The plots of the coefficients of determination R^2 versus the number of variables are shown in Figure 4.20. As can be seen, R^2 values for the training set increase with the addition of variables, as each new variable contribute to better description of chosen data set. However, this trend does not hold for the R^2 values of the test set. Models with too many variables tend to be overfitted, reducing their predictive accuracy for the dataset. Therefore, to address this, the models with the highest R^2 values of the test set were selected to describe the interaction energies of our m-OPs and PET MPs. Hence, the 3-variable models were chosen for adsorption sites I, III, and IV, while the 4-variable model was determined to be the best for adsorption site II. The statistical parameters of chosen models are shown in Table A 7 in Appendix.



Figure 4.20. The values of coefficients of determination (R2) of the training set and test of the QSPR models with 1-5 variables developed to predict the interaction energies of m-OPs with adsorption sites I, II, III and IV of PET MP.

The selected models were further validated by methods of internal validation, including Leave Many Out (LMO) cross-validation and *Y*-scrambling. LMO studies the robustness of the model by excluding a random sample of molecules and recalculating the model, after which it tests the performance by predicting the excluded molecules. The scatter graphs obtained by LMO technique are shown in Figure 4.21. For all four models, the Q^2_{LMO} values are mainly close to Q^2_{LOO} value and not widely disperse. The stability seen in the LMO cross-validation demonstrates that the models remain consistent when tested with different subsets of data. The results of *Y*-scrambling test are shown Figure 4.22. R^2 and Q^2 values of all selected models are significantly greater that R^2 and Q^2 values of the models after *Y*-scrambling is performed. The clear difference in performance between the original models and the *Y*-scrambled models confirms that the

observed correlations are not results from random correlation. Overall, these results confirm that the developed models are accurate and useful for predicting the adsorption behaviour of m-OPs on PET MPs.



Figure 4.21. Scatter plot of Leave Many Out (LMO) technique performed to show the robustness of the selected QSPR models for the prediction of interaction energies between m-OPs and PET MP on adsorption sites I, II, III and IV.



Figure 4.22. Scatter plot of Y-scrambling validation method for the selected QSPR models for the prediction of interaction energies between m-OPs and PET MP on adsorption sites I, II, III and IV.

The fitting criteria values and the values of internal and external validation criteria for the selected models are summarized Table A 8 in Appendix. The scatter plots of QSPR-predicted vs. DFT-calculated interaction energies using the selected models are shown in Figure 4.23. All the models accurately represent the given data, as evidenced by their high R^2 values. The data points or data clusters are closely aligned with the diagonal line in all cases, further supporting the models' accuracy. To assess the applicability domain of the selected models and identify potential outliers, Williams plots were used, as shown in Figure 4.24. A Williams plot evaluates the reliability of a model's predictions by examining both structural (X) and response variable (Y) outliers. The plot is constructed using the residuals (differences between predicted and actual values) and leverage (HAT) values, where leverage measures the influence of a molecule's structural properties on the

model. Molecules with leverage values exceeding a critical threshold (h*h^*h*) are considered structural outliers (*X*-outliers), while those with residuals outside $\pm 3.0 \sigma$ are classified as response variable outliers (*Y*-outliers). For all four models, no *Y*-outliers were identified, as all residuals fall within the $\pm 3.0 \sigma$ range. However, in the Williams plot for Model IV, which predicts interaction energies between organic pollutants and adsorption site IV on the PET MP, the HAT value for catechol (*o*-Ph(OH)₂) exceeds the critical threshold (*h**), classifying it as an *X*-outlier. Catechol was part of the test set, but the HAT values for all other molecules remained below the limit.

Overall, these results indicate that the selected models are accurate and reliable for predicting the interaction energies of organic pollutants on PET MPs, with only one structural outlier identified across all models.



Figure 4.23. Calculated interaction energies and interaction energies predicted by QSPR Model I, Model II, Model III and Model IV build using training set (25 molecules) for the prediction of interaction energies of m-OPs with the adsorption sites I, II, III and IV on PET MP respectively.



Figure 4.24. Williams plots of QSPR Model I, Model II, Model III and Model IV for the prediction of interaction energies of m-OPs with the adsorption sites I, II, III and IV on PET MP respectively.

4.2.3. Model descriptors

The coefficients of descriptors used to build our selected Model I, Model II, Model III and Model IV with normalized training set that describe the interaction energies between m-OPs and PET MP on adsorption sites I, II, III and IV, respectively, are provided in Table 4.2 along with the corresponding statistical data. These descriptors were selected for their significance, as evidenced by their p-values, all of which are below the upper significance threshold of 0.05. This indicates that each descriptor contributes meaningfully to the model's predictive ability and is not included by chance.

Table 4.2. Descriptive statistical data of the coefficients in in QSPR Model I, Model II, Model III and Model IV for the prediction of interaction energies of m-OPs with the adsorption sites I, II, III and IV respectively build using training sets (25 molecules) with normalized descriptors

Model	Variable	Coefficient	Standard deviation	(+/-) 95% Confidence interval	<i>t</i> –ratio	<i>p</i> -value
I	Intercept	-10.78	0.90	1.88	-11.92	0.00
	RDF050m	-8.44	1.94	4.04	-4.35	< 0.01
	F01[C–O]	-4.29	1.51	3.14	-2.84	< 0.01
	SAacc	-12.98	1.98	4.11	-6.56	0.00
П	Intercept	11.79	3.70	7.72	3.19	< 0.01
	SM15_AEA(ed)	-39.90	4.36	9.09	-9.15	0.00
	Mor22s	-3.28	0.71	1.48	-4.61	< 0.01
	HATS4u	5.03	1.63	3.41	3.08	< 0.01
	MLOGP2	8.22	0.92	1.92	8.93	0.00
Ш	Intercept	-6.61	1.55	3.22	-4.26	< 0.01
	TDB04v	-5.40	2.22	4.61	-2.44	< 0.01
	CATS2D_03_DL	-6.81	1.00	2.08	-6.81	0.00
	F02[O–O]	-8.81	1.50	3.11	-5.89	0.00
IV	Intercept	-8.34	0.76	1.58	-10.98	0.00
	RDF050u	-8.03	1.02	2.11	-7.90	0.00
	MLOGP2	3.28	1.10	2.29	2.98	< 0.01
	SAdon	-5.76	1.06	2.20	-5.46	0.00

In addition to their significance, the descriptors were evaluated for multicollinearity. The cross-correlation coefficients between all pairs of descriptors in each model were calculated, and none exceeded the threshold of 0.7 (R_{ij} >0.7) as shown in the cross-correlation matrix in Figure 4.25. This confirms that the descriptors are not strongly correlated with one another and ensures that the models exhibit true linear behaviour.



Figure 4.25. Cross-correlation matrix of descriptor variables in QSPR Model I, Model II, Model III and Model IV for the prediction of interaction energies of m-OPs with the adsorption sites I, II, III and IV respectively.

The descriptors used to build these models are listed in Table 4.3. The used descriptors fall into several categories: 2D atom pairs, molecular properties, 3D-MoRSE, RDF descriptors, 3D-autocorrelation, CATS2D, edge adjacency indices, and GETAWAY. 2D atom pairs are very simple and straightforward descriptors that represent the number of times the exact atom pair occur at a certain topological distance. Molecular properties descriptors are derived from literature models. Others are calculated by more complex scheme. 3D-MoRSE (Molecular Representation of Structures based on Electronic diffraction) descriptors are 3D descriptors derived from scattering functions in electron diffraction studies.^[266] There is a high range of 3D-MoRSE descriptors that are usually denoted as Morsw, where *s* denotes the number of signal and takes values from 1 to 32, and *w* denotes a weighting scheme. 3D-MoRSE descriptors can be unweighted (*u*) or weighted by either mass (*m*), van der Waals volume (*v*), Sanderson electronegativity (*e*), polarizability (*p*), ionization potential (*i*) or I-state (*s*). These descriptors have proved they are very valuable for QSA/PR. However, they are hardly interpretable.^[267] RDF descriptors are based on the radial

distribution function (RDF).^[266] RDF can be interpreted as the probability distribution of finding an atom in a spherical volume of radius r.^[268] RDF descriptors are denoted as RDFsw, where s represents the distance in 0.1 Å and it takes up values from 10 to 155 in five-unit steps, and w stands for the weighting scheme same as for 3D-MoRSE descriptors. 3D-autocorrelation descriptors calculate the Euclidian distance between atom pairs within a molecule up to a cut-off distance.^[269] These descriptors are denoted as TDB*sw*, where *s* stands for oath length (lag) and *w* stands for weighting scheme. 3D-autocorrelation descriptors can be unweighted (u) or weighted by mass (m), van der Waals volume (v), Sanderson electronegativity (e), polarizability (p), ionization potential (i), I-state (s) or covalent radius (r). CATS2D descriptors are topological pharmacophore descriptors based on auto- and cross-correlation of pharmachophoric atom types.^[270] CATS2D descriptors recognize 5 pharmacophoric types of atoms: H-bond donor (D), H-bond acceptor (A), positively charged (P), negatively charged (N), and lipophilic (L) atom. Every atom in the molecule can be assign to none, one or two pharmacophoric types. CATS2D descriptors count the number of atom pairs in molecule with a defined pharmacophoric types at a specific topological distance that is in the range from 0 to 9. Edge adjacency indices are topological descriptors derived from edge adjacency matrix that encode the connectivity between graph edges.^[271] GETAWAY descriptors (GEometry, Topology and Atom-Weights Assembly) correlate geometric information from the leverage matrix (also known as the Molecular Influence Matrix (MIM)), topological information from the molecular graph, and physicochemical information using various weighting schemes such as mass (m), van der Waals volume (v), Sanderson electronegativity (e), polarizability (p), ionization potential (i) and I-state (s).^[272] GETAWAY descriptors are categorized into two groups: H-GETAWAY, which are calculated from MIM, and R-GATEWAY, which are calculated from the influence distance matrix that combines the elements of MIM with the elements of the geometry matrix.

Model	Descriptor name	escriptor name Descriptor definition		
Ι	RDF050m	Radial Distribution Function -050 / weighted by mass	RDF descriptors	
	F01[C-O]	Frequency of C – O at topological distance 1	2D Atom Pairs	
	SAacc	Surface area of acceptor atoms from P_VSA-like descriptors	Molecular properties	
Ш	SM15_AEA(ed)	Spectral moment of order 15 from augmented edge adjacency matrix weighted by edge degree	Edge adjacency indices	
	Mor22s	Signal 22 / weighted by I-state	3D-MoRSE descriptors	
	HATS4u	Leverage-weighted autocorrelation of lag 4 / unweighted		
	MLOGP2	Squared Moriguchi octanol-water partition coefficient $(\log P^2)$	Molecular properties	
III	TDB04v	3D Topological distance based descriptors – lag 4 weighted by van der Waals volume	3D autocorrelations	
	CATS2D_03_DL	CATS2D Donor-Lipophilic at lag 03	Pharmacophore descriptors	
	F02[O-O]	Frequency of O – O at topological distance 2	2D Atom Pairs	
IV	RDF050u	Radial Distribution Function – 050 / unweighted	RDF descriptors	
	MLOGP2	LOGP2 Squared Moriguchi octanol-water partition coefficient $(\log P^2)$		
	SAdon	Surface area of donor atoms from P_VSA-like descriptors	Molecular properties	

Table 4.3. Definitions of descriptors used in QSPR Model I, Model II, Model III and Model IV for the prediction of interaction energies of m-OPs with the adsorption sites I, II, III and IV respectively.

To further refine the analysis, the descriptors selected from the initial models developed with the training set of 25 molecules were used to construct new models based on the entire dataset of 30 m-OPs. This approach allows for a more comprehensive evaluation of the models by incorporating all available data, which can improve the accuracy and generalizability of the predictions. The performance of the newly developed models, using normalized descriptors, is presented in Figure 4.26. The corresponding plots of residuals, illustrating the differences between predicted and calculated interaction energies, are shown in Figure 4.27. Descriptive statistical data for the updated coefficients in these models is provided in Table 4.4. the corresponding models' equations are presented below by equations (4.12) to (4.15):

$$E_{\text{int}}(I) = -7.190 \text{RDF050m} - 4.155 \text{F01}[\text{C-O}] - 13.603 \text{SAacc} - 11.285$$
(4.12)

$$E_{int}(II) = -39.471 \text{SM15}_\text{AEA(ed)} - 3.261 \text{Mor22s} + 4.839 \text{HATS4u} + 8.064 \text{MLOGP2} + 11.572$$
(4.13)

$$E_{\text{int}}(\text{III}) = -5.811 \text{TDB04v} - 6.773 \text{CATS2D}_{03}\text{DL} - 8.446 \text{F02[O-O]} - 6.562$$
(4.14)

$$E_{\text{int}}(\text{IV}) = -8.152 \text{RDF050u} + 2.690 \text{MLOGP2} - 5.593 \text{SAdon} - 8.296$$
(4.15)


Figure 4.26. Calculated interaction energies and corresponding models' equations of QSPR Model I, Model II, Model III and Model IV that predict interaction energies of m-OPs with the adsorption sites I, II, III and IV on PET MP respectively when applied on the entire set of data (30 molecules) with normalized descriptors.



Figure 4.27. Plot of residuals for QSPR Model I, Model II, Model III and Model IV that predict interaction energies of m-OPs with the adsorption sites I, II, III and IV on PET MP respectively when applied on the entire set of data (30 molecules) with normalized descriptors.

Model	Variable	Coefficient Standard deviation		(+/-) 95% Confidence interval	<i>t</i> –ratio	<i>p</i> –value
	Intercept	-11.29	0.78	1.61	-14.40	0.00
T	RDF050m	-7.19	1.73	3.55	-4.16	< 0.01
•	F01[C-O]	-4.15	1.44	2.96	-2.89	< 0.01
	SAacc	-13.60	1.80	3.69	-7.57	0.00
	Intercept	11.57	3.45	7.11	3.35	< 0.01
	SM15_AEA(ed)	-39.47	3.92	8.07	-10.08	0.00
II	Mor22s	-3.26	0.64	1.33	-5.06	0.00
	HATS4u	4.84	1.28	2.63	3.79	< 0.01
	MLOGP2	8.06	0.79	1.64	10.15	0.00
	Intercept	-6.56	1.41	2.90	-4.66	< 0.01
Ш	TDB04v	-5.81	2.03	4.17	-2.87	< 0.01
	CATS2D_03_DL	-6.77	0.94	1.93	-7.23	0.00
	F02[O-O]	-8.45	1.40	2.88	-6.04	0.00
	Intercept	-8.30	0.76	1.56	-10.95	0.00
IV	RDF050u	-8.15	0.93	1.91	-8.77	0.00
1.4	MLOGP2	2.69	1.00	2.05	2.70	< 0.01
	SAdon	-5.59	0.89	1.82	-6.32	0.00

Table 4.4. Descriptive statistical data of the coefficients in in QSPR Model I, Model II, Model III and Model IV for the prediction of interaction energies of m-OPs with the adsorption sites I, II, III and IV respectively when applied on the entire set of data with normalized descriptors (30 molecules).

It should be noted that a more negative interaction energy (E_{int}) corresponds to a stronger adsorption. Therefore, the descriptors with a negative coefficient have a positive effect on the adsorption capacity. In the above models, only **MLOGP2** (Model II and Model IV) and **HATS4u** (Model II) have a negative effect on adsorption capacities, while all other descriptors have a positive effect.

The descriptors included in Model I that describe the interaction energies between m-OPs and PET-MP at adsorption site I are **RDF050m**, **F01[C-O]** and **SAacc**. As can be seen from equation (4.12), all three included descriptors have negative coefficients, meaning that all three synergistically contribute to the chosen response, i.e. the higher their values, the higher is adsorption. Due to the fact that descriptors are normalized, the coefficients allow us to estimate the extent of their contribution. Hence, the highest contribution is obtained by the **SAacc** descriptor, while other two contribute approx. 50% (**RDF050m**) and 25% (**F01[C-O]**) of its extent. The **SAacc** descriptor represents the surface area of the acceptor atoms.^[273] The organic molecules characterized by higher values of the **SAacc** descriptor are more susceptible to accept electrons and

form hydrogen bonds. The values of the descriptors used in Model I for all 30 m-OPs selected are listed in Table A 9 in Appendix. The m-OPs that adsorb most strongly (with the lowest interaction energy) at adsorption site I have very high SAacc values. The m-OPs with oxygen-containing functional groups such as PhSO₂OH, PhSOOH, BSAM and PhPO(OH)₂, in which the oxygen is bound to sulphur and phosphorus, and BAM and BA, in which the oxygen is bound to carbon, have the highest SAacc values. PhMe, Bz, Sty, o-PhMe₂, PhCl, PhBr, PhA, PhSH have an SAaac value of 0 and have the lowest interaction energies with adsorption site I. The descriptor F01[C-O] indicates the number of oxygen atoms directly bonded to the carbon atom. m-OPs with an oxygencontaining functional group can form hydrogen bonds with the carboxyl group at adsorption site I. However, for some m-OPs such as PhSO₂OH, PhSOOH, and PhPO(OH)₂, the value of this descriptor is equal to 0, as their oxygen is not bonded directly to the carbon, but to a sulphur, phosphorus or nitrogen atom. Since they can also form hydrogen bonds, as the hydrogen is bonded to a strongly electronegative atom, this contributes significantly to their interaction energies. Hence, their strong ability to form hydrogen bonds between O atoms and PET MPs surface is compensated with the highest SAaac values, as mentioned above. **RDF050m** is the probability distribution for finding an atom in a spherical volume with a radius of 5 Å, weighted by mass. The descriptor RDF050m has a positive effect on adsorption, presumably due to the fact that the larger molecules (recognized over weighting scheme m, i.e. molecular mass) can interact with the neighbouring parts of PET MPs, which can further stabilize the complex and facilitate the adsorption process.

Model II (eq. (4.13)) describes the relationship between the structural properties of selected m-OPs and their interaction energies with adsorption site II of PET MPs. In this case the 4-variable model which includes **SM15_AEA(ed)**, **Mor22s**, **HATS4u** and **MLOGP2** was determined to be the best-fitting model whose values are listed in Table A 10. First two mentioned descriptors have negative coefficients in eq. (4.13), providing synergistic effects to the chosen response, while later two possess positive confidents, thus antagonistically contributing to the predicted adsorption at site II. **SM15_AEA(ed)** makes by far the largest contribution, while the other three descriptors play a minor, but not negligible role in the adsorption prediction with contributions of approximately 20% for **MLOGP2**, 12% for **HATS4u**, 9% for **Mor22s** in relation to **SM15_AEA(ed)**. **SM15_AEA(ed)** is a spectral moment of order 15 from augmented edge adjacency matrix weighted by edge degree. The edge adjacency matrix is a matrix derived from

the molecular graph. The matrix element is equal to 1 if the edges are adjacent and 0 if they are not. In an augmented edge adjacency matrix, the main diagonal contains non-zero elements, which in our case are weighted by edge degree.^[274] SM15_AEA(ed) is a trace of the corresponding matrix raised to the power of 15. Molecules that contain more atoms and have a more branched structure have a higher value of the SM15 AEA(ed) descriptor than smaller linear or cyclic molecules. The larger m-OPs in terms of molecular weight are usually more branched and have a higher number of electronegative oxygen and/or nitrogen atoms. The m-OPs with higher values of SM15_AEA(ed) have more degrees of rotational freedom and their functional groups can interact with the nearby surface of PET MP, including the terminal hydroxyl group, which is located near adsorption site II and forms hydrogen bonds. Therefore, the high contribution of this descriptor in the model shows the importance of hydrogen bonding for the stabilization of the system. Following in terms of contribution to the selected response, specifically adsorption at site II, is MLOGP2, which represents a squared value of logP calculated by Moriguchi method.^[275,276] The high MLOGP2 value indicate very lipophilic or very hydrophilic molecules. Although these two properties are polar opposites, since logP values for hydrophilic molecules are negative, while for lipophilic are positive, by squaring their values within MLOGP2 descriptor, both values become positive. In our case, all the molecules are lipophilic, therefore the positive coefficient of MLOGP2 indicates the negative effect of lipophilicity on the adsorption on PET MP. HATS4u, as next contributing descriptor, does not have such straightforward correlation with ability to make hydrogen bonding as previously elaborated descriptors. Hence, HATS4u pertain to GETAWAY descriptors, whose interpretation is much more complicated and requires the consideration of three elements: weighting scheme, leverages, and the Dirac delta function. As HATS4u represents a leverage-weighted autocorrelation of lag 4 / unweighted, the weighting scheme is irrelevant. The descriptor with the fourth largest contribution in eq (4.13) is Mor22s, representing the signal 22 of 3D-MoRSE weighted by I-state. As mentioned above, 3D-MoRSE descriptors are very useful and are often used as the main or one of the contributing descriptors of various models. However, their meaning is also hardly interpretable, without a straightforward correlation with the adsorption effect in our case. However, its weighting scheme I-state, i.e. electrotopological state atom (I-state) indices, is developed to better indicate the important topological features and molecular fragments mediating a particular response, combining electronic and topological characteristics of atoms/molecules.^[277,278] Hence, it is has been found as important structural feature to reflected adsorption of selected m-OPs via hydrogen bonding to PET MPs surface in our case.

In the Model III (eq. 4.14), **CATS2D_03_DL**, **F02[O-O]** and **TDB04v** descriptors are highly correlated with the end-point, i.e. (E_{int}) values, at site III. Their values for all 30 m-OPs are listed in Table A 11. **CATS2D_03_DL** is the number of hydrogen bond donor atom-lipophilic atom pair that are separated by three bonds. This descriptor directly reflects importance of hydrogen bonding in adsorption of m-OPs onto PET MPs; higher **CATS2D_03_DL** value results in stronger adsorption of m-OPs. **F02[O-O]** represents the number of oxygen-oxygen atom pair separated by two bonds. As explained above, m-OPs possessing O atom showed higher susceptibility to hydrogen bonding to PET MPs, while double O atoms in the molecule even increase such a probability. The **TDB04v** descriptor is another descriptor that refers to the size of the molecule. This descriptor calculates the Euclidean distances between atoms with a topological distance of 4 (separated by 4 bonds). In this case, the descriptor is also weighted by the van der Waals radius. Therefore, more linear molecules with larger surface areas have a higher value of the **TDB04v** descriptor than small molecules with curved or cyclic structures. The coefficient of this descriptor is negative, which means that the interaction energies between PET MPs and m-OPs are stronger for the m-OPs with a larger surface area.

Model IV describes the relationship between the structural features of m-OPs and their interaction energies with adsorption site IV at PET MP surface. The values of descriptors used in model IV are listed in Table A 12. The descriptor with the highest contribution in eq. (4.15) is **RDF050u**. The same descriptor has been shown as contributing one in Model II, but there was its analogue weighted by molecular mass (**RDF050m**). It is assumed that RDF050u contributes by indicating that m-OPs with higher values possess larger surfaces. Consequently, these m-OPs are able to engage in a greater number of weak interactions with the surface of PET MPs compared to smaller m-OPs. **SAdon** falls into the same category of P_VSA like descriptors as **SAacc** (shown to be the most contributing in Model I (eq (4.12)), while it represents surface of electron donor atoms. **MLOGP2** negatively affects adsorption, the same as in in Model II (eq. (4.13)). The molecules with a higher **MLOGP2** value (the one with more lipophilic character) will have a higher interaction energy with the adsorption site IV which has a negative effect on adsorption capacity.

In general, the main descriptors used in the above models to correlate structural features of m-OPs and their ability to adsorb at PET MPs surface, represented in our case as interaction energies between m-OPs and PET MPs surface, refer either to their ability to form hydrogen bonds or to the size of the m-OPs. The contribution of the descriptors differs from model to model, i.e. from adsorption site to adsorption site. For adsorption site I, the most important descriptor, **SAacc**, whose contribution is over 54%, is associated with the ability to form hydrogen bonds between the functional groups and the terminal carboxyl group of PET MPs. For adsorption site II, the most important descriptor is **SM15_AEA(ed)**, which contributes over 70%. This descriptor is more related to the size and flexibility of the functional group. An examination of the structure of the m-OPs (Figure 4.19) reveals that the larger, more flexible functional groups contain atoms capable of forming hydrogen bonds. Furthermore, an analysis of the optimized geometries of the m-OP/PET systems (Figure A 11 – Figure A 16) demonstrates that these functional groups can extend toward the terminal hydroxyl group of PET near adsorption site II, enabling the formation of hydrogen bonds. The models representing adsorption at sites III and IV do not have as dominant descriptors as model I and model II. Since these sites do not have a terminal hydroxyl or carboxyl group, adsorption is not only dominated by the ability to form a stronger bond but also depends on the size of the m-OPs and the ability to have many interactions, all of which contribute to the total interaction energy.

The findings presented are of significant relevance to our understanding of the environmental fate and behaviour of PET micro- and nano-plastics. Ageing of PET leads to the formation of new carboxyl groups, i.e. by surface oxidation^[279] of the macromolecule. As was demonstrated by the findings presented herein, these sites facilitate adsorption of m-OPs, especially those able to form hydrogen bonds such as PhSO₂OH, PhSOOH, and PhPO(OH)₂ used in this study. While hydrophobic interactions are commonly considered to be the most relevant^[280,281] in the context of the environmental vectoring properties of the micro- and nano-plastics, this study underscores the importance of hydrogen bonding. In addition, it is important to note that hydrogen bonding is a much stronger type of a weak interaction than hydrophobic interactions. It is therefore more than likely that aged PET, but also other polymer particles that form carboxyl groups under the oxidative conditions of environmental ageing, exhibit enhanced adsorption of favourable aromatic moieties

4.3. Adsorption of pollutants on PET MP in water

Up to this point, this thesis has focused on the different functional groups and the nature of their interactions with PET MP. However, in the real environmental conditions, the adsorption of organic pollutants is influenced by numerous of factors, many of which cannot be captured by the so far computational approach. As previously discussed, the size of the adsorbate may play a role, as larger molecules can establish more interactions with MP. At the same time, they can also interact with the surrounding water molecules, meaning that the adsorption capacity of MP in the water for a certain adsorbate depends also on its affinity for water molecules. Keeping that in mind, the next step in this thesis involved the transition from small model organic molecules in vacuo to real water pollutants in the water surrounding.

One of the tools to computationally study the adsorption in more realistic environmental systems is molecular dynamics (MD).^[140,282–284] The use of classical force field (FF) in MD allows the simulation of larger and more complex systems that include the explicit molecules of solvent while keeping the reasonable accuracy-cost ratio. Unlike ab initio calculations done so far, MD simulation captures the movement of atoms and calculates system's properties by time-averaging, which is more similar to real experiments. Currently, MD simulations are widely accepted for studying adsorption interactions, particularly in biological systems, where well-developed FFs provide reliable.^[285-287] However, applying MD to other systems is more challenging due to the lack of well-parameterized FFs. The use of MD to study MP adsorption is still in its early stages. A comprehensive 2024 review by Townstead^[140] points out the lack of consensus among studies using MD for this purpose, particularly regarding the choice of force field, timescale, and initial system setup. Despite this limitation, there is a growing trend in using MD simulations to study adsorption on MP as it is a valuable tool for adsorption studies, cost-effective, and could reduce the need for experimental approaches. MD studies use purely theoretical or combined approach with experimental data where the adsorption capacity is generally correlated with the calculates interaction energies.^[72,141,288,289] The interaction energy between two or more molecules is defined as the contribution to the total energy of the system caused by interactions between atoms or molecules.^[290] In ab initio calculations, this is simply calculated as the difference between the total energy of the system and the individual energy of the molecules. However, in classical MD, the interactions between molecules are generally calculated form the distances between

atoms/molecules up until defined cutoff distance.^[192,291] The total interaction energy is the sum of non-bonding energy terms that is, electrostatic (Coulombic) term and dispersion interactions term usually represented as Lennard-Jones (LJ) potential. The absolute values of these quantities might not have a true physical meaning; however, they can still provide valuable insights into the dominant effects governing the system.

4.3.1. Force field

As previously defined in Section 2.3.4.1, force fields consist of a set of equations and associated parameters used to model a system's potential energy. While most force fields provide well-established parameters for proteins and nucleic acids, making them widely applicable to the simulation of biochemical systems, the parametrization of other types of molecules presents significant challenges. This process can be both time-consuming and challenging, as the accuracy of a force field is highly dependent on the quality of its parameters. Fortunately, several software tools have been developed to assist and to provide parameters for molecular dynamics simulations of a broad range of molecules.^[292–295] There is no single dominant FF used to study adsorption of MP. Some of the popular choices are COMPASS,^[296] COMPAS II,^[288,289] CHARMM,^[297] GROMOS,^[141] and OPLS-AA.^[142,298] However, there is no clear explanation for selecting one over another, as the current state of art lacks the validation of FF for these types of studies. In this thesis, OPLS-AA force field was chosen, as it is one of the often used FF that is still actively developing but keeping in mind the need for future validation of generated results.

Ther FF parameters for PET were obtained from the work of Sangkhawasi et al.^[299] who used DL_FIELD 4.1. software to generate OPLS-AA parameters. This software analyses the molecular structure and determines the type of atoms for a given FF scheme.^[300] DL_FIELD can produce force fields applicable to a wide range of models, including polymers, as demonstrated by Yong's work, where it was employed to generate force fields for simulating polyethylene and polystyrene nanoparticle models.^[301] In this study, the PET molecule is represented as a trimer consisting of three segments: the Head (carboxyl terminal), the Body (repeating unit), and the Tail (hydroxyl terminal), as depicted in Figure 4.28.



Figure 4.28. Structure oh PET: Head (blue), Body (yellow; repeating unit) and Tail (red) and their atom labels in generated OPLS-AA force field.

For the missing angles, i.e. unspecified in the provided parameters set: CA-CA-CA, CT-CT-OH, HAE-CT-OH, CT-OH-HO, and HAE-CT-HAE, parameters from corresponding angles in amino acids were applied. For unspecified dihedrals, including HO-OH-CO-CA, OES-CT-CT-OES, and OES-CT-CT-OH, parameters were obtained using LibParGen. LibParGen (Parameter Generator for Organic Ligands) is another software designed to provide OPLS-AA parameters for neutral organic molecules, and particularly useful to generate FF of small organic ligands.^[292,302,303] Since periodic boundary conditions (PBC) are applied during MD simulations, it is necessary to ensure that the total charge of the system equals zero. Therefore, the partial charges of the atoms were slightly adjusted to ensure that the total charge of the system is neutral. In the same manner, the parameters of the OPLS-AA force field for pesticides were obtained from LibParGen. Finally, all the assigned OPLS-AA parameters and atom charges can be found in Table A 13 – Table A 16 in Appendix.

4.3.2. MD simulation of pesticides in water

Five pesticides, known to be common organic pollutants in aquatic environment were chosen for MD simualtion, i.e. simazine (SMZ), atrazine (ATZ), diuron (DIU) alachlor (ALC) and isoproturon (IPT).^[304] Their molecular structures are shown in Figure 4.29.





To gain a better understanding of pollutants' affinity toward water molecules, the MD simulations were performed on systems containing one pollutant placed in the cubic box dimension $9 \text{ nm} \times 9 \text{ nm} \times 9 \text{ nm}$ solvated using SPC model of water. Before MD simulation, the systems were minimised and equilibrated to the pressure of 1 bar and the temperature of 300 K in NVT and NPT ensemble. The results of energy minimisation and equilibration of simazine are shown in Figure 4.30, while the results for other pesticides are shown in Figure A 17. in Appendix. After this, the MD simulation was run for 10 ns as described in Materials and Methods section (Section 3.1.3.).



Figure 4.30. Energy minimization and equilibration of simazine in water in NVT and NPT ensemble.

From the resulting MD trajectory, the short-range Lennard-Jones (LJ) and short-range Coulombic interaction energies were calculated between non-bonded atoms of the pollutant molecules and surrounding water molecules and the results are presented in Figure 4.31. The LJ potential corresponds to van der Waals interactions, which primarily govern interactions between uncharged species. In contrast, Coulombic energy arises from attractive or repulsive electrostatic forces between partially charged particles. More negative values of both interaction energies

indicate stronger interactions with water molecules. Therefore, the total interaction energy, that is the sum of two contributions, provides insights into the affinity of each pollutant for water. Charged molecules or those with polar functional groups tend to have dominant Coulombic interactions, whereas nonpolar molecules interact predominantly by LJ interactions. In this case, simazine and atrazine exhibit the most negative Coulombic interaction energies, but their LJ contributions are relatively weak, suggesting that electrostatic interactions predominantly govern their interactions with water molecules, likely due to the polar functional groups in their structures. For diuron and isoproturon, Coulombic interactions are less negative, while in the case of alachlor, the LJ and Coulombic contributions are nearly equal. This suggests that among these molecules, alachlor interacts with water more through van der Waals forces than electrostatic interactions. The balance between LJ and Coulombic interactions can be useful to predict molecule's solubility, tendency to aggregate, and potential adsorption onto microplastics.



Figure 4.31. Interaction energies of organic pollutants (pesticides) with water calculated from MD trajectories.

4.3.3. Construction of PET models for MD simulations

There are different way of setting up the initial system for MP adsorption found in the literature. The MP can be represented as a single chain with high degree of polymerisation that folds due to the intramolecular interactions,^[289] or as an aggregate of smaller oligomeric chains.^[72,141,142,297] In some studies, the aggregation of oligomeric chains occurs during the MD simulation of a system that includes water and pollutant molecules, which can become encapsulated in the process.^[142,297] In others, the MP aggregate is preassembled, with water and pollutant molecules added later.^[72,141] The later approach seems more in line with the events in realistic systems and therefore is used in this work.

To proceed to adsorption studies, first the two distinct PET model were prepared. Each model comprises 150 monomeric units. The first model was designed as a single, continuous 150mer chain, representing a more intact, unfragmented PET structure. In contrast, the second model was designed as an aggregate of thirty 5-mer chains, mimicking a more fragmented PET surface, which could arise from natural aging, degradation, or environmental processes. These two representations have the same number of monomeric units, but different number of terminal carboxyl and hydroxyl groups. While the single-chain model contains only one carboxyl terminal group and one hydroxyl terminal group, the aggregate model features 30 carboxyl and 30 hydroxyl terminal groups. This structural difference affects the availability of functional groups on the surface of PET MP that can potentially influence the adsorption mechanism of organic pollutants.



Figure 4.32. Two models of PET microplastic used for MD simulation of pesticides adsorption to PET microplastic. The structure on the left is designed as one folded 150-mer chain, and the one on the right as an aggregate of 30 5-mer chains that represent fragmented PET microplastic.

The spherical shape of the 150-mer chain was achieved through molecular dynamics (MD) simulation in the NVT ensemble. The system was initially simulated at 500 K, a temperature above the melting point of PET, for 100 ps to promote chain flexibility and rearrangement. Subsequently, the temperature was gradually reduced to 300 K, following the same parameters outlined in the Methods section (Section 3.1.3.). The aggregate model was created by placing thirty 5-mer chains into a cubic simulation box. The dimensions of the box were optimized to 7.5 nm × 7.5 nm × 7.5 nm in all three directions. A larger box size did not result in aggregation, while a smaller box caused the chains to connect with their periodic images due to insufficient spacing. The system was equilibrated in the NVT ensemble at successive temperatures of 500 K, 450 K, 400 K, and finally 300 K. Both structures (the single 150-mer chain and the aggregate of thirty 5-mer chains) were equilibrated and finalized into spherical shapes with an approximate diameter of 4 nm (Figure 4.32).

4.3.4. MD simulation of pesticides' adsorption on model PET

To simulate the adsorption and calculate the interaction energies between pollutants and the two models of PET, each PET model was placed in a box dimensions $12 \text{ nm} \times 12 \text{ nm} \times 12 \text{ nm}$ with one pollutant molecule in the vicinity. The systems were minimised and equilibrated before MD simulation was run. The interaction energies were calculated as the sum of non-bonded terms, i.e. Coulombic and Lennard-Jones potential. The results of the interaction energies between pollutants and both models of PET (aggregate that represent fragmented PET MP and one oligomeric chain that represent pristine PET MP) are presented in Figure 4.33. Note, that the more negative interaction energy is attributed to the stronger interactions.

The total interaction energy for the adsorption on PET model I is in range from -23.9 to -15.0 kcal mol⁻¹, and for the adsorption on PET model II from -26.0 to -18.6 kcal mol⁻¹. In both cases, the strongest interactions (most negative energy) occurs with alachlor, while the weakest with atrazine. These values align with previously reported interaction energies from DFT studies, where the adsorption energy of diclofenac on PET varied from -33 to -22 kcal mol⁻¹ depending on the adsorption site. However, it is important to note that those DFT values were calculated in a system without explicit water molecules, which could otherwise compete for interactions with both diclofenac and the PET surface. Feng et al.^[141] studied the adsorption of benzo[a]pyrene on five types of MPs (PET, PS, PVC, PP, and PE) using MD simulations. In their study, the MP particle

was modelled as a surface onto which benzo[a]pyrene could adsorb. Their findings indicated that the strongest interactions occurred with PS, followed by PVC, PET, PE, and PP. The interaction energy for PET was approximately -25 kcal mol⁻¹, which is of the same order of magnitude as the results obtained in this work. Interestingly, they also showed the adsorption on plastic surface is almost the twice stronger than the adsorption on single polymer chain. This difference is attributed to the greater number of hydrophobic interactions that can be realised with the surface compared to a single chain, a finding that also aligns with previous conclusions from DFT studies. Guo et al.^[72] studied the sorption of sulfamethazine on different MP in a vacuum layer, neglecting the effect of water. Their results show higher values for the adsorption, specifically for PET, the interaction energy with sulfamethazine is -40.87 kcal mol⁻¹. In a similar study from Liu et al.,^[288] the interaction energy of PET MP with sulfamonomethoxine calculated in a system in vacuum is -125 kcal mol⁻¹ which is one order of magnitude larger than results from the studies where explicit water molecules are included. However, when comparing their results for adsorption on PE, the results are much more consistent. The interaction energy of PE with benzo[a]pyrene, sulfamethazine, and sulfamonomethoxine was reported as -21 kcal mol⁻¹, -27 kcal mol⁻¹, and -34kcal mol⁻¹, respectively. This significant difference in adsorption results for PET between studies that include explicit water molecules and those in vacuum can be attributed to PET's molecular structure. PET contains oxygen atoms capable of forming hydrogen bonds and stronger electrostatic interactions with pollutants. However, these interactions can be disrupted in the presence of explicit water molecules. In contrast, PE lacks hydrogen-bonding sites and charged atoms, so the inclusion of explicit water molecules in the system has a much smaller effect on adsorption.

Further evidence for this can be found by examining the interaction energy terms. The results in Figure 4.33 show that for all the pesticides, electrostatic (Coulombic) interactions were significantly weaker when compared to van der Waals (LJ) interactions, which indicates dispersion forces are the dominant factor in pesticides-PET MP interactions in water environment. These results are in contrast with the previous ones from DFT calculations in continuum model that does not take into account explicit water molecules and emphasises the importance of hydrogen bonds in the adsorption mechanism on PET MP. This suggest DFT calculations in continuum model, while as can give the information about the nature of interactions between two interacting systems,

cannot accurately predict the adsorption capacity as it neglects the large influence of surrounding environmental factors, in this case, interactions with explicit water molecules.



Figure 4.33. Interaction energies between pesticides and PET model I, and PET model II (fragmented model) calculated from MD trajectories.

Finally, a comparison can be made between the adsorption of pesticides on PET model I and PET model II. It is evident that PET model II exhibits stronger interactions with all pesticides compared to PET model I. The difference in total interaction energy results from an increase in both Coulombic and LJ interactions. The stronger Coulombic interactions can be attributed to the higher number of terminal carboxyl and hydroxyl groups in PET model II, which provide additional hydrogen bonding and electrostatic interaction sites for pesticides. Despite the increase in Coulombic interactions, LJ forces remain the dominant contribution, suggesting that hydrophobic interactions still play a crucial role in adsorption.

The equilibrium structures of atrazine and simazine adsorption are shown in Figure 4.34 and the equilibrium structures of diuron, alachlor and isoproturon adsorption are shown in Figure 4.35. From the careful examination of figures, it can be observed that the pesticides can penetrate more deeply into the surface of PET model II compared to PET model I. This allows for increased number of interaction contacts that can be realised between PET model and pesticide, resulting in stronger adsorption. This can be linked to the greater flexibility of PET chains and a less compact structure, which is more adaptable and more easily modified, thereby stabilizing the adsorption process. A similar effect was observed in a study from Liu et al.^[142] where MD was used to

simulate the adsorption of aromatic hydrocarbons on PE and PS microplastic water. Their findings showed that hydrocarbons continuously modified the structure of the microplastic during sorption, further supporting the influence of polymer flexibility on adsorption mechanism.

Another important factor to be considered when running a simulation is the time-scale.^[140] The processes involving penetration of molecules and interaction with flexible chains of polymer in water may need more time to reach equilibrium state than process involving adsorption on more rigid plastic surface or in vacuum. As demonstrated in a study by Liu et al.,^[142] the interaction energies reached equilibrium after 1 to 2 ns, while the simulation was run for 50 ns. In a study from Feng et al.,^[141] the equilibrium was achieved even later, and the simulation was run for 100 ns. However, in studies from Guo et al. and Liu et al. in vacuum, the simulations were run only for 0.5 ns.^[72,288]



Figure 4.34. Structures of alachlor adsorption on PET model I and PET model II.



Figure 4.35. Structures of diuron, alachlor and isoproturon adsorption on PET model I and PET model II.

4.3.5. Adsorption capacities

The adsorption capacity of MP reported in the literature varies depending on the study. This is caused by the numerous of factors affect the MP adsorption capacity and different methodologies used to study the adsorption. The study by Miranda et al.^[305] reports significantly lower MP capacities for organic pollutants compared to similar research. Both pristine and aged PET MP did not show any significant adsorption affinity for most organic pollutants, including diuron and alachlor. On the other hand, a particularly high affinity was observed for pentachlorophenol, especially after the aging treatment. These findings align with those of Godoy et al.,^[306] who reported low adsorption affinity of PET for organic molecules such as diuron and atrazine but a high affinity for phenol. However, both studies lack an explanation for the increased adsorption capacity of pentachlorophenol or phenol. Overall, research suggests that the aging process enhances the sorption capacity of MP. Titov et al.^[307] investigated the sorption of micropollutants present in European river water on PET, PVC, and PS. Their results indicated the aging state of MP particles is the most critical factor influencing sorption, followed by polymer type.

Considering the variation between reported results, this study aimed to estimate the difference in the PET MP's adsorption capacity for various pesticides. To achieve this, the adsorption capacity of a very high relative concentration of PET MP was measured. Details of PET MP characterisation can be found in Appendix B. The experimental adsorption capacity of both pristine and thermally aged PET MP was measured for diuron, alachlor and isoproturon. Simazine and atrazine were excluded from the analysis due to their participation during the experiments. The results, presented in Table 4.5, indicate that thermally aged PET MP exhibit a higher adsorption capacity for all tested pesticides. Adsorption increased by 132%, 60%, and 92% for diuron, alachlor, and isoproturon, respectively, compared to pristine PET MP demonstrating that aging enhances the adsorption capacity of microplastics.

The adsorption affinity of both pristine and thermally aged PET was the highest for alachlor, followed by diuron, with the lowest observed for isoproturon. When these results are compared with the previously discussed results from the MD simulations, it is obvious that the average interaction energies of the pesticides adsorption followed the same trend for the adsorption on PET model I, showing alachlor exhibit the strongest interactions with PET, followed by diuron and isoproturon. However, for PET model II, the MD simulation results did not align with experimental findings, as they indicated the lowest adsorption for diuron instead of isoproturon.

Dollutont	Pristine	PET MP	Thermally aged PET MP			
Fonutant	q / mg g $^{-1}$	$\sigma \ / \ mg \ g^{-1}$	q / mg g $^{-1}$	$\sigma \ / \ mg \ g^{-1}$		
Diuron	0.0094	0.0004	0.0218	0.0012		
Alachlor	0.0250	0.0038	0.0399	0.0013		
Isoproturon	0.0037	0.0004	0.0071	0.0004		

Table 4.5. Equilibrium adsorption capacity (*q*) and standard deviation (σ) for the adsorption of diuron, isoproturon and alachlor (initial concentration 50 µmol dm⁻³, 1 mL) on pristine and thermally aged PET MP (100–200 µm; 100 mg) at the temperature of 25 °C and rotation speed of 120 rpm.

A combined MD and experimental study from Leng et al.^[308] on the adsorption of 17- β estradiol on MP, demonstrated a correlation between interaction energies from MD simulations and experimental adsorption capacities. In that case, adsorption capacity was highest for PE (q = 0.642mg g⁻¹; $E_{int} = -26.06 \text{ kcal mol}^{-1}$), followed by PP ($q = 0.545 \text{ mg g}^{-1}$; $E_{int} = -25.19 \text{ kcal mol}^{-1}$), and PS ($q = 0.415 \text{ mg g}^{-1}$; $E_{\text{int}} = -23.79 \text{ kcal mol}^{-1}$). Similarly, a study by Guo et al.^[72] confirmed agreement between adsorption capacities and interaction energies from MD simulations. However, this alignment is not universal. For example, research by Chen et al.^[139] on the adsorption of three antibiotics and by Li et al.^[136] on the adsorption of three pesticides on PE MP found that trends in adsorption capacities did not strongly align with the trend observed in calculated interaction energies. Looking back at the results from this study, MD simulations further indicated that adsorption is stronger on PET model II, which represents fragmented, aged PET MP, compared to PET model I, representing unfragmented, pristine MP which agrees with the experimental results. Although, MD results alone cannot quantitatively explain the increase in adsorption capacities after ageing. Despite this, interaction energies calculated by MD can still reasonably explain the trends observed in adsorption capacities. Difference in results might come from various sources, including error in experimental data, and calculated interaction energy.

MD has shown its potential as a valuable tool for understanding adsorption mechanisms. To better capture MP adsorption phenomena, advancing simulation methodologies are encouraged, including testing of different force fields and refining the representation of PET surfaces based on experimental findings. Developing more realistic PET models, incorporating surface heterogeneity, polymer flexibility, and degradation effects, could improve the agreement between simulations and experimental adsorption capacities. Additionally, longer simulation timescales, improved solvent models, and the inclusion of other effects, as the competitive adsorbents and ions, should be explored to gain a better understanding of pollutant adsorption onto MP in environment.

5. Conclusion

This research integrated computational chemistry methods into the study of MP adsorption mechanisms to reveal the nature of interactions between organic pollutants and PET MP. Additionally, simulation approaches were developed to better understand experimentally obtained results and to enable the prediction of adsorption behaviour for other pollutants.

DFT methods were used to analyse the specific interactions between PET MP and ten functional groups commonly found in organic pollutants in water. These functional groups were represented by small molecules containing either a single benzene or pyridine ring (m-OP). To ensure both computational efficiency and accuracy, the performance of multiple DFT functionals was evaluated against the MP2 method. Based on this comparison, the M06-L functional was selected, as it produced results most comparable to MP2 while being a pure functional, making it computationally less demanding than hybrid functionals. The DFT study confirmed that the adsorption interactions between organic pollutants and PET MP are weak physical interactions. The strongest interactions were observed between m-OP molecules containing proton donor and proton acceptor groups, specifically, phenylacetic acid, benzoic acid, and benzamide, and adsorption site I of the PET model, which featured a terminal carboxyl group. The differences in interaction energies among m-OP molecules were primarily attributed to the number of hydrogen bonds they could form with PET. In most cases, the interaction energy depended on the adsorption site of PET, with the exception of benzene and chlorobenzene, which exhibited the weakest adsorption across all adsorption sites. This weaker adsorption was attributed to the absence of partially charged groups, and the interactions only through van der Waals forces. making these molecules reliant solely on weak van der Waals interactions.

The study was further expanded to 30 m-OPs to capture more variety in functional groups and to be able to create mathematical models that correlate calculate interaction energies with structural properties of m-OPs using QSA/PR. The results for added functional groups were in accordance with previously reported ones. The interaction energies were the strongest for benzamide, benzoic acid, phenylacetic acid, phenylphosphonic acid, phenylsulfonic acid, and phenylsulfinic acid, all of which can form multiple hydrogen bonds with PET MP. For m-COPs whose functional group cannot form hydrogen bonds, such as benzene, toluene and styrene, the interaction with PET MPs was similar for all four adsorption sites. Four QSA/PR models were developed correlating the interaction energies between m-COPs and PET MPs at four adsorption sites. Descriptors of m-OP used in models provide further insight into their key structural features for adsorption to PET MPs. For adsorption site I (terminal carboxyl group), the most important descriptor was **SAacc**, which refers to the surface area of the electron acceptor atoms, which can be interpreted as the ability to form hydrogen bonds. For adsorption site II (terminal hydroxyl group), the most important descriptor is **SM15_AEA(ed)**, which refers to the size and flexibility of the functional groups. Considering the position of the terminal hydroxyl group at adsorption site II and the final optimised geometries of the binary PET MPs and m-COP systems, this descriptor is important due to the ability of larger and more flexible functional groups to interact with the terminal hydroxyl group of PET MPs through hydrogen bonding. Adsorption at sites III and IV (surface of PET) is not much facilitated by strong hydrogen bonds, but rather by the size of the functional group, which is able to form many weaker interactions with the surrounding surface of PET MPs.

In the next phase, the study transitioned from small organic molecules in a continuum model to real pollutants in an aqueous environment. The MD simulations aimed to clarify the dynamics of adsorptions of atrazine, simazine, diuron, alachlor and isoproturon on PET MP model including the explicit water molecules that could potentially compete for adsorption sites on both PET MP and pollutants. Two distinct PET models were constructed: one representing an unfragmented microplastic (PET model I) and the other simulating aged and fragmented PET MP (PET model II). The adsorption results for the selected pesticides revealed that van der Waals interactions contributed more significantly to the total interaction energy than electrostatic interactions. Additionally, adsorption onto the fragmented PET model II was stronger for all pesticides compared to PET model I. This increased adsorption was attributed to the greater flexibility and less compact structure of model II, which allowed for more structural modifications and a higher number of interaction contacts with the pesticides, thereby enhancing the total interaction energy. MD simulations with PET model I showed the strongest interactions for alachlor, followed by diuron and isoproturon, which aligns with the trend in adsorption capacities for pristine PET. However, in simulations with PET model II, the interaction energy followed a different order: alachlor, isoproturon, and diuron. This order did not match experimentally obtained results for adsorption on thermally aged PET MP, which indicated that alachlor had the highest adsorption capacity, followed by diuron and then isoproturon. Nevertheless, adsorption on PET model II was consistently higher than on PET model I for all tested pesticides. This trend agreed with experimental findings, which demonstrated that pesticide adsorption capacity increased after the aging treatment of PET MP. However, the exact increase in adsorption capacity could not be quantitatively explained through MD simulations. Overall, MD simulations proved to be a valuable tool for understanding adsorption mechanisms, highlighting their potential for further development in this field to better capture the adsorption behaviour of MP in water environment.

Appendices



Appendix A

Figure A 1. Differences in interaction energies between the model pollutants and terephthalic acid in configuration TA I calculated by MP2/def2-TZVP method and various DFT methods with def2-SVP basis set.



Figure A 2. Differences in interaction energies between the model pollutants and terephthalic acid in configuration TA II calculated by MP2/def2-TZVP method and various DFT methods with def2-SVP basis set.

	$E_{\rm int}({\rm TA~I})$ / kcal mol ⁻¹											
Pollutant				def2-TZVP	def2-TZVPP	def2-TZVP						
	B3LYP	M05	M05–2X	M06	M06–2X	M06–HF	M06-L	MP2	MP2	MP2	M06–L	
Phenylacetic acid	-24.28	-21.72	-24.45	-21.98	-23.38	-32.70	-21.22	-16.93	-18.48	-19.04	-21.66	
Benzoic acid	-24.44	-21.80	-24.66	-22.09	-23.67	-32.93	-21.36	-17.05	-18.61	-19.17	-21.86	
Benzamide	-22.07	-20.06	-21.31	-20.29	-20.83	-25.99	-19.54	-16.04	-17.57	-18.09	-20.15	
Aniline	-13.75	-12.19	-13.29	-12.61	-13.40	-16.49	-11.70	-10.38	-11.86	-12.07	-11.91	
Phenol	-12.71	-11.85	-11.93	-12.06	-11.87	-14.39	-11.38	-8.69	-10.19	-10.72	-11.80	
Anisole	-10.63	-9.92	-9.79	-9.46	-9.67	-11.69	-8.55	-7.54	-8.14	-8.45	-8.45	
Triazine	-13.20	-12.22	-12.31	-11.98	-12.32	-14.81	-11.08	-9.41	-10.25	-10.52	-11.12	
Chlorotriazine	-12.83	-11.86	-11.67	-11.57	-11.70	-13.70	-10.61	-9.02	-9.90	-10.18	-10.73	
Benzene	-5.01	-4.80	-5.28	-5.04	-5.16	-6.33	-4.23	-2.75	-4.99	-5.20	-4.62	
Chlorobenzene	-6.20	-6.03	-5.48	-5.33	-5.43	-5.97	-4.54	-3.71	-4.38	-4.69	-4.89	

Table A 1. Interaction energies of TA and organic pollutants in a system where the functional group of pollutant interacts with the carboxyl group of terephthalic acid (TA I) calculated by MP2/def2-TZVP method and various DFT methods with def2-SVP basis set.

	$E_{\rm int}({\rm TA~II}) \ / \ {\rm kcal~mol^{-1}}$											
Pollutant				def2-TZVP	def2-TZVPP	def2-TZVP						
	B3LYP	M05	M05–2X	M06	M06–2X	M06–HF	M06–L	MP2	MP2	MP2	M06–L	
Phenylacetic acid	-6.35	-6.07	-6.23	-7.57	-6.99	-7.25	-6.88	-3.45	-8.01	-8.32	-6.53	
Benzoic acid	-10.89	-10.55	-10.69	-12.61	-11.87	-12.33	-11.20	-5.62	-11.04	-11.34	-11.06	
Benzamide	-13.87	-12.82	-13.54	-15.10	-14.31	-17.76	-13.61	-9.15	-13.72	-14.12	-13.51	
Aniline	-6.56	-5.92	-6.33	-7.81	-7.06	-7.45	-7.18	-3.16	-7.97	-8.31	-6.96	
Phenol	-7.76	-7.42	-7.52	-8.77	-8.52	-9.02	-7.71	-4.46	-9.03	-9.32	-7.38	
Anisole	-7.34	-6.81	-7.35	-9.20	-8.28	-8.86	-8.86	-4.14	-9.18	-9.56	-8.82	
Triazine	-5.39	-5.24	-5.22	-6.07	-5.82	-5.96	-4.82	-2.48	-5.93	-6.05	-4.52	
Chlorotriazine	-5.87	-5.48	-5.51	-6.71	-6.48	-6.95	-5.98	-2.92	-7.27	-7.40	-5.66	
Benzene	-4.23	-4.30	-4.45	-5.65	-4.95	-5.30	-4.95	-2.22	-6.01	-6.25	-4.56	
Chlorobenzene	-5.50	-5.20	-5.35	-6.65	-6.09	-6.55	-6.03	-2.89	-7.59	-7.82	-5.75	

Table A 2. Interaction energies of TA and organic pollutants in a system where the aromatic ring of pollutant is placed on top of the aromatic ring of terephthalic acid (TA II) calculated by MP2/def2-TZVP method and various DFT methods with def2-SVP basis set.

	$\Delta E_{\rm int}({ m TA~I})$ / kcal mol ⁻¹											
Pollutant					def2-TZVPP	def2–TZVP						
	B3LYP	M05	M05–2X	M06	M06–2X	M06–HF	M06–L	MP2	MP2	M06-L		
Phenylacetic acid	5.80	3.24	5.97	3.50	4.90	14.22	2.74	-1.55	0.56	3.18		
Benzoic acid	5.83	3.19	6.05	3.48	5.06	14.32	2.75	-1.56	0.56	3.25		
Benzamide	4.50	2.49	3.74	2.72	3.26	8.42	1.97	-1.53	0.52	2.58		
Aniline	1.89	0.33	1.43	0.75	1.54	4.63	-0.16	-1.48	0.21	0.05		
Phenol	2.52	1.66	1.74	1.87	1.68	4.20	1.19	-1.50	0.53	1.61		
Anisole	2.49	1.78	1.65	1.32	1.53	3.55	0.41	-0.60	0.31	0.31		
Triazine	2.95	1.97	2.06	1.73	2.07	4.56	0.83	-0.84	0.27	0.87		
Chlorotriazine	2.93	1.96	1.77	1.67	1.80	3.80	0.71	-0.88	0.28	0.83		
Benzene	0.02	-0.19	0.29	0.05	0.17	1.34	-0.76	-2.24	0.21	-0.37		
Chlorobenzene	1.82	1.65	1.10	0.95	1.05	1.59	0.16	-0.67	0.31	0.51		
MAE	3.08	1.85	2.58	1.80	2.31	6.06	1.17	1.29	0.38	1.36		
RMSE	12.46	4.34	10.27	4.42	7.64	56.94	2.24	1.89	0.16	3.19		

Table A 3. The difference between MP2/ def2-TZVP and DFT/ def2-SVP interaction energies shown in Table A 1. calculated by MP2/def2-TZVP method and various DFT methods with def2-SVP basis set.

	$\Delta E_{\rm int}({ m TA~II})$ / kcal mol ⁻¹										
Pollutant			def2-TZVPP	def2-TZVP							
	B3LYP	M05	M05–2X	M06	M06–2X	M06–HF	M06–L	MP2	MP2	M06–L	
Phenylacetic acid	-1.66	-1.94	-1.78	-0.44	-1.02	-0.76	-1.13	-4.56	0.31	-1.48	
Benzoic acid	-0.15	-0.49	-0.35	1.57	0.83	1.29	0.16	-5.42	0.30	0.02	
Benzamide	0.15	-0.90	-0.18	1.38	0.59	4.04	-0.11	-4.57	0.40	-0.21	
Aniline	-1.41	-2.05	-1.64	-0.16	-0.91	-0.52	-0.79	-4.81	0.34	-1.01	
Phenol	-1.27	-1.61	-1.51	-0.26	-0.51	-0.01	-1.32	-4.57	0.29	-1.65	
Anisole	-1.84	-2.37	-1.83	0.02	-0.90	-0.32	-0.32	-5.04	0.38	-0.36	
Triazine	-0.54	-0.69	-0.71	0.14	-0.11	0.03	-1.11	-3.45	0.12	-1.41	
Chlorotriazine	-1.40	-1.79	-1.76	-0.56	-0.79	-0.32	-1.29	-4.35	0.13	-1.61	
Benzene	-1.78	-1.71	-1.56	-0.36	-1.06	-0.71	-1.06	-3.79	0.24	-1.45	
Chlorobenzene	-2.09	-2.39	-2.24	-0.94	-1.50	-1.04	-1.56	-4.70	0.23	-1.84	
MAE	1.23	1.59	1.36	0.58	0.82	0.90	0.89	4.53	0.27	1.10	
RMSE	1.96	2.95	2.27	0.60	0.80	2.06	1.02	20.78	0.08	1.62	

Table A 4. The difference between MP2/ def2-TZVP and DFT/ def2-SVP interaction energies shown in Table S2.

Dollutont		BSSE (TA I) / %	6	BSSE (TA II) / %			
Pollutalit	def2-SVP	def2-TZVP	def2-TZVPP	def2-SVP	def2-TZVP	def2-TZVPP	
Phenylacetic acid	36.5%	8.3%	9.1%	71.3%	30.8%	28.0%	
Benzoic acid	32.3%	8.2%	8.9%	66.6%	28.1%	25.9%	
Benzamide	27.6%	8.0%	8.6%	52.3%	21.2%	19.7%	
Aniline	36.5%	13.8%	13.1%	74.7%	31.9%	28.7%	
Phenol	43.3%	17.0%	15.4%	67.5%	29.3%	26.3%	
Anisole	38.2%	12.6%	12.7%	70.2%	29.8%	27.0%	
Triazine	32.3%	10.0%	10.2%	71.6%	29.0%	27.4%	
Chlorotriazine	33.6%	10.6%	10.9%	73.2%	29.2%	27.8%	
Benzene	58.1%	22.2%	19.4%	74.6%	30.8%	27.5%	
Chlorobenzene	33.6%	16.9%	16.5%	73.3%	29.5%	27.2%	

Table A 5. The dependence of basis set superposition error (BSSE) for the interaction of organic pollutants and TA calculated by MP2/def2-TZVP method and various DFT methods with def2-SVP basis set.

×	Coordinates		A 4	Coordinates			A 4	Coordinates			
Z	x	у	z	Atom	x	у	z	Atom	x	у	z
0	9.250543	1.004598	-2.661174	С	-1.301652	5.019308	-3.088444	Н	1.535956	-6.776552	1.609066
С	7.95902	0.613098	-3.046418	С	-1.383518	5.223525	-1.595832	Н	5.260703	1.097983	-3.968006
Н	9.38349	1.915264	-2.939021	Н	-0.335517	5.368942	-3.480048	Н	4.720395	2.664414	-4.629344
С	7.707333	-0.724468	-2.403834	Н	-2.109915	5.578427	-3.577012	0	3.354876	1.090193	-4.820685
Н	7.858289	0.508727	-4.146994	0	-0.191367	4.782865	-0.936725	0	-1.009021	-4.587159	1.762126
Н	7.184141	1.326086	-2.708818	Н	-2.257635	4.705549	-1.174204	Н	0.547921	-5.034181	3.065652
0	6.461389	-1.224704	-2.894748	Н	-1.461678	6.296269	-1.376996	Н	-0.654664	-6.323301	2.762483
Н	8.51379	-1.433247	-2.645101	С	-0.159886	3.506886	-0.508571	Н	3.698924	0.83184	-5.681931
Н	7.663747	-0.634547	-1.308174	С	1.153952	3.134035	0.065716	С	-0.622814	-3.292064	1.722307
С	6.156218	-2.478377	-2.511146	0	-1.107267	2.756215	-0.584892	С	-1.629686	-2.426999	1.064387
С	4.840663	-2.929066	-3.034543	С	1.266629	1.88985	0.699968	0	0.425995	-2.891612	2.175776
0	6.878187	-3.15374	-1.815919	С	2.270845	3.977128	-0.013968	С	-1.210818	-1.178038	0.583291
С	4.436347	-4.23424	-2.73201	С	2.471066	1.49522	1.267363	С	-2.975803	-2.800795	0.949365
С	4.003666	-2.105849	-3.799452	Н	0.38364	1.249419	0.754303	Н	-0.153705	-0.910802	0.678632
С	3.205983	-4.706718	-3.170907	Н	2.179662	4.938767	-0.520608	С	-2.124026	-0.306567	0.00352
Н	5.110693	-4.864404	-2.148191	С	3.481145	3.57725	0.537355	С	-3.899522	-1.911851	0.412946
С	2.777107	-2.581662	-4.242038	С	3.586993	2.337586	1.184042	Н	-3.29501	-3.780673	1.307722
Н	4.301302	-1.084003	-4.036082	Н	2.562723	0.538615	1.786623	Н	-1.812884	0.672519	-0.372981
С	2.353591	-3.878404	-3.915045	Н	4.359822	4.220553	0.470313	С	-3.479868	-0.660394	-0.056411
Н	2.913922	-5.736998	-2.95638	С	4.870675	1.8614	1.758867	Н	-4.953983	-2.183304	0.351067
Н	2.125107	-1.951242	-4.851217	0	5.034712	0.775603	2.262927	С	-4.436362	0.351495	-0.574282
С	1.054271	-4.350877	-4.476547	0	5.852784	2.775564	1.657684	0	-5.712276	0.054301	-0.25895
0	0.463487	-5.429771	-3.886343	Н	6.65013	2.354903	2.013837	0	-4.123471	1.354762	-1.173878
0	0.568835	-3.875771	-5.4689	Н	1.947494	-0.708242	-1.544036	С	-6.699606	1.037211	-0.563042
С	0.270652	-5.487928	-2.480503	С	2.601389	-1.30334	-0.907741	С	-6.421	2.407679	0.058766
С	-0.999222	-4.774572	-2.106604	С	2.172946	-2.520668	-0.403477	Н	-6.801026	1.145314	-1.654253
Н	0.207039	-6.551587	-2.210521	С	3.881944	-0.819277	-0.604518	Н	-7.628122	0.601701	-0.175654
Н	1.113282	-5.047512	-1.926864	Н	1.174431	-2.890002	-0.649782	0	-5.6467	2.212636	1.230143
0	-0.788038	-3.403902	-2.423762	С	3.013729	-3.258216	0.4444	Н	-7.372365	2.894339	0.322896
Н	-1.203781	-4.875543	-1.026112	С	4.720833	-1.555907	0.238367	Н	-5.875786	3.071998	-0.622826
Н	-1.870031	-5.154164	-2.659618	С	4.379181	0.447448	-1.194657	С	-4.508727	2.93829	1.360667
С	-1.866725	-2.655482	-2.754185	С	2.573205	-4.520744	1.077803	С	-3.526288	2.23251	2.219633
С	-1.469972	-1.272032	-3.104834	С	4.283108	-2.763988	0.767483	0	-4.31046	3.983225	0.792316
0	-2.999238	-3.072148	-2.720699	Н	5.71155	-1.16626	0.4768	С	-3.791997	0.967268	2.763442
С	-0.126755	-0.877595	-3.151466	0	5.471676	0.929251	-0.978337	С	-2.251212	2.795191	2.35919
С	-2.477221	-0.323676	-3.306497	0	3.467252	0.995854	-2.013005	С	-2.774027	0.24063	3.367112
Н	0.653545	-1.626736	-3.016401	0	1.446321	-5.019515	0.494746	Н	-4.787212	0.53402	2.653335
С	0.207384	0.457187	-3.334296	0	3.120425	-5.067352	2.003333	Н	-2.070566	3.78717	1.940706
С	-2.151349	1.015289	-3.470122	Н	4.917296	-3.344154	1.43986	С	-1.23338	2.067457	2.957209
Н	-3.518789	-0.650482	-3.27497	С	3.840061	2.154298	-2.76398	С	-1.478732	0.769315	3.425189
Н	1.253825	0.772085	-3.359812	С	0.797957	-6.075581	1.192851	Н	-2.96204	-0.76663	3.74259
С	-0.808925	1.415325	-3.459496	Н	2.917481	2.738591	-2.876682	Н	-0.219674	2.466478	3.035073
Н	-2.936486	1.765529	-3.548585	Н	4.591357	2.733184	-2.207863	С	-0.31522	-0.046726	3.85282
С	-0.425717	2.846064	-3.443626	С	4.375202	1.749571	-4.10973	0	0.835339	0.310597	3.751579
0	-1.505687	3.653088	-3.438068	С	-0.067666	-5.521243	2.297301	0	-0.668109	-1.256784	4.318921
0	0.712185	3.256682	-3.377139	Н	0.189533	-6.5975	0.441797	Н	0.156945	-1.760858	4.410312

Table A 6. X. Y. Z coordinates of o	ptimized geometry	v of polveth	vlene terephthalate	e model (mPET).
, , ,		J - F - J		



Figure A 3. Independent gradient model (IGMH) analysis scatter plots of weak interactions between phenylacetic acid, benzoic acid, benzamide, aniline and phenol, and mPET at all four adsorption sites. In the usual colouring scheme, the blue colour represents attractive interactions, the green colour represents van der Waals interactions, and the red colour represents non-bonding interactions.



Figure A 4. Independent gradient model (IGMH) analysis scatter plots of weak interactions between anisole, triazine, chlorotriazine, benzene and chlorobenzene, and mPET at all four adsorption sites. In the usual colouring scheme, the blue colour represents attractive interactions, the green colour represents van der Waals interactions, and the red colour represents non-bonding interactions.



Figure A 5. Independent gradient model (IGMH) analysis scatter plots of weak interactions between a water molecule and mPET at all four adsorption sites. In the usual colouring scheme, the blue colour represents attractive interactions, the green colour represents van der Waals interactions, and the red colour represents non-bonding interactions.



Figure A 6. Schematic representations of the interactions between the model pollutants and mPET at adsorption site I analysed by independent gradient model (IGMH) and visualised on 0.01 a.u. isosurface. In the usual colouring scheme, the blue colour represents attractive interactions, the green colour represents van der Waals interactions, and the red colour represents non-bonding interactions.



Figure A 7. Schematic representations of the interactions between the model pollutants and mPET at adsorption site II analysed by independent gradient model (IGMH) and visualised on 0.01 a.u. isosurface. In the usual colouring scheme, the blue colour represents attractive interactions, the green colour represents van der Waals interactions, and the red colour represents non-bonding interactions.



Figure A 8. Schematic representations of the interactions between the model pollutants and mPET at adsorption site III analysed by independent gradient model (IGMH) and visualised on 0.01 a.u. isosurface. In the usual colouring scheme, the blue colour represents attractive interactions, the green colour represents van der Waals interactions, and the red colour represents non-bonding interactions.



Figure A 9. Schematic representations of the interactions between the model pollutants and mPET at adsorption site IV analysed by independent gradient model (IGMH) and visualised on 0.01 a.u. isosurface. In the usual colouring scheme, the blue colour represents attractive interactions, the green colour represents van der Waals interactions, and the red colour represents non-bonding interactions.


Figure A 10. Schematic representations of the interactions between the water molecule and mPET at four adsorption sites analysed by independent gradient model (IGMH) and visualised on 0.01 a.u. isosurface. In the usual colouring scheme, the blue colour represents attractive interactions, the green colour represents van der Waals interactions, and the red colour represents non-bonding interactions.



Figure A 11. Structures of PET model with benzene, toluene, styrene, phenylacetylene and chlorobenzene adsorbed on four adsorption sites. The structures were optimized by M06-L/def2-SVP/svpfit method.



Figure A 12. Structures of PET model with bromobenzene, fluorobenzene, phenol, anisole. and benzaldehyde adsorbed on four adsorption sites. The structures were optimized by M06-L/def2-SVP/svpfit method.



Figure A 13. Structures of PET model with acetophenone, benzoic acid, phenylacetic acid, aniline, and *N*-methylaniline adsorbed on four adsorption sites. The structures were optimized by M06-L/def2-SVP/svpfit method.



Figure A 14. Structures of PET model with *N*,*N*-dimethylaniline, benzonitrile, nitrobenzene, benzamide, and thiophenol adsorbed on four adsorption sites. The structures were optimized by M06-L/def2-SVP/svpfit method.



Figure A 15. Structures of PET model with phenylsulfenic acid, phenylsulfinic acid, phenylsulfonic acid, benzenesulfonamide, and phenylphosphonic acid adsorbed on four adsorption sites. The structures were optimized by M06-L/def2-SVP/svpfit method.



Figure A 16. Structures of PET model with *o*-xylene, catechol, *o*-phenyldiamine, triazine, and chlorotriazine adsorbed on four adsorption sites. The structures were optimized by M06-L/def2-SVP/svpfit method.

			Trainin	Г	Test set $(n = 5)$	i)			
Model	R^2	Q^2	F	р	S	$S_{\rm PRESS}$	R^2	F	S
Ι	0.88	0.84	53.25	< 0.0001	2.73	3.21	0.81	17.48	3.35
II	0.93	0.89	65.95	< 0.0001	1.37	1.71	0.99	274.35	1.07
III	0.91	0.87	73.38	< 0.0001	1.57	1.94	0.89	31.27	1.92
IV	0.91	0.86	68.92	< 0.0001	1.27	1.57	0.94	60.38	1.92

Table A 7. Statistical evaluation of QSPR models for predicting interaction energies between m-OPs and the adsorption sites I, II, III and IV of PET MP derived for training set (25 molecules) and applied on test set (5 molecules).

Table A 8. The fitting criteria values and the values of internal and external validation criteria for the selected models QSPR Model I, Model II, Model III and Model IV for the interaction energies between m-OPs and the adsorption sites I, II, III and IV of PET MP respectively.

	Critoria	Model					
	Criteria	Ι	Π	III	IV		
	R^2	0.88	0.93	0.91	0.91		
	$R^2_{ m adj}$	0.86	0.92	0.89	0.89		
50	$R^2 - R^2_{ m adj}$	0.02	0.01	0.01	0.01		
ittir	LOF	10.81	3.20	3.81	2.36		
Ц	RMSE _{tr}	2.50	1.22	1.48	1.17		
	MAE _{tr}	1.95	0.97	1.15	0.94		
	CCC _{tr}	0.94	0.96	0.95	0.95		
	$Q^2_{ m loo}$	0.83	0.89	0.86	0.86		
_	$R^2 - Q^2_{ m loo}$	0.05	0.04	0.05	0.05		
tion	RMSE _{cv}	3.00	1.53	1.83	1.44		
lida	MAE _{cv}	2.35	1.22	1.41	1.15		
val	PRESS _{cv}	225.51	58.17	83.97	51.72		
mal	CCC_{cv}	0.91	0.94	0.92	0.93		
Inter	$Q^2_{ m LMO}$	0.81	0.88	0.82	0.85		
_	$R^2 Y_{ m scr}$	0.12	0.16	0.13	0.13		
	$Q^2 Y_{\rm scr}$	-0.26	-0.33	-0.27	-0.25		
	RMSE _{ext}	2.60	0.92	1.20	1.48		
atio	MAE _{ext}	2.06	0.84	0.79	1.27		
Exte alid	PRESS _{ext}	33.79	4.24	7.21	11.01		
E E	R^2_{ext}	0.94	0.99	0.92	0.94		

Table A 9. Normalized molecular descriptors of 30 molecules calculated by Dragon 6.0. and used to build
Model I that predicts the interaction energies between m-OPs and the adsorption site I of PET MP. The
molecules used for test set are underlined and typed in bold.

Compound	Abbrevation	$E_{\rm int}(I)$ / kcal mol ⁻¹	RDF050m	F01[C–O]	SAacc
Benzene	Bz	-10.20	0.01	0.00	0.00
Toluene	PhMe	-9.89	0.01	0.00	0.00
Styrene	Sty	-10.34	0.24	0.00	0.00
Phenylacetylene	PhA	<u>-12.29</u>	0.23	0.00	0.00
Chlorobenzene	PhCl	-11.71	0.03	0.00	0.00
Bromobenzene	PhBr	-11.87	0.35	0.00	0.00
Fluorobenzene	PhF	-12.10	0.00	0.00	0.16
Phenol	PhOH	<u>-21.30</u>	0.00	0.50	0.45
Anisole	PhOMe	-19.21	0.38	1.00	0.12
Benzaldehide	PhCHO	-20.34	0.35	0.50	0.27
Acetophenone	AcPHO	-21.24	0.76	0.50	0.27
Benzoic acid	BA	-29.43	0.78	1.00	0.72
Phenylacetic acid	PAA	-29.89	0.53	1.00	0.72
<u>Aniline</u>	$PhNH_2$	<u>-20.18</u>	0.00	0.00	0.35
N–Methylaniline	NMA	-17.28	0.40	0.00	0.54
<i>N</i> , <i>N</i> –Dimethylaniline	DMA	-13.24	0.18	0.00	0.19
Benzonitrile	PhCN	-19.47	0.07	0.00	0.33
Nitrobenzene	PhNO ₂	-18.26	0.47	0.00	0.03
Benzamide	BAM	-32.69	0.66	0.50	0.61
Thiophenol	PhSH	<u>-12.99</u>	0.08	0.00	0.00
Phenylsulfenic acid	PhSOH	-19.41	0.26	0.00	0.41
Phenylsulfinic acid	PhSOOH	-30.18	0.91	0.00	0.67
Phenylsulfonic acid	PhSO ₂ OH	-32.04	1.00	0.00	0.88
Benzenesulfonamide	BSAM	-22.99	0.92	0.00	0.80
Phenylphosphonic acid	PhPO(OH) ₂	-29.62	0.71	0.00	1.00
<i>o</i> -xylene	o–PhMe ₂	-11.57	0.01	0.00	0.00
Catechol	o-Ph(OH) ₂	-25.03	0.00	1.00	0.90
o-Phenyldiamine	$o-Ph(NH_2)_2$	-18.95	0.00	0.00	0.70
<u>Triazine</u>	Tr	<u>-18.96</u>	0.00	0.00	0.53
Chlorotriazine	TrCl	-20.05	0.00	0.00	0.53
Normalization f	factor:		3.26	2.00	94.53

Compound	Abbrevation	$E_{\rm int}({\rm II})$ / kcal mol ⁻¹	SM15_AEA(ed)	Mor22s	HATS4u	MLOGP2
Benzene	Bz	-9.44	0.71	0.09	0.46	0.59
Toluene	PhMe	-7.98	0.83	0.05	1.00	0.79
Styrene	Sty	-13.00	0.85	-0.02	0.35	0.94
Phenylacetylene	PhA	-11.90	0.85	0.03	0.38	0.94
Chlorobenzene	PhCl	-10.91	0.83	0.12	0.47	0.96
Bromobenzene	PhBr	<u>-11.08</u>	0.83	-0.10	0.48	1.07
Fluorobenzene	PhF	-10.40	0.83	-0.41	0.47	0.84
Phenol	PhOH	-18.63	0.83	0.28	0.49	0.26
<u>Anisole</u>	PhOMe	<u>-16.09</u>	0.85	-0.35	0.43	0.40
Benzaldehide	PhCHO	-15.41	0.85	-0.08	0.49	0.36
Acetophenone	AcPHO	<u>-18.19</u>	0.91	0.20	0.43	0.51
Benzoic acid	BA	-20.78	0.91	0.83	0.40	0.33
Phenylacetic acid	PAA	-17.25	0.86	0.14	0.81	0.36
<u>Aniline</u>	PhNH ₂	<u>-15.23</u>	0.83	0.22	1.05	0.26
<i>N</i> –Methylaniline	NMA	-16.75	0.91	0.29	0.49	0.41
N,N–Dimethylaniline	DMA	-17.05	0.85	-0.08	0.44	0.40
Benzonitrile	PhCN	-15.69	0.85	-0.54	0.50	0.36
Nitrobenzene	PhNO ₂	-15.58	0.91	-0.12	0.98	0.56
Benzamide	BAM	-22.49	0.91	0.64	0.69	0.19
Thiophenol	PhSH	-15.35	0.83	0.05	0.49	0.74
Phenylsulfenic acid	PhSOH	-18.92	0.85	0.24	0.65	0.26
Phenylsulfinic acid	PhSOOH	-17.59	0.91	-0.37	0.85	0.12
Phenylsulfonic acid	PhSO ₂ OH	-25.21	1.00	0.07	0.77	0.08
Benzenesulfonamide	BSAM	-22.14	1.00	-0.49	0.72	0.02
Phenylphosphonic acid	PhPO(OH) ₂	-23.72	1.00	0.01	0.80	0.10
<i>o</i> –xylene	o-PhMe ₂	-10.58	0.90	-0.52	0.79	1.00
<u>Catechol</u>	o-Ph(OH) ₂	-22.44	0.90	0.27	0.49	0.09
o-Phenyldiamine	$o-Ph(NH_2)_2$	-18.71	0.90	0.43	0.90	0.09
Triazine	Tr	-11.88	0.71	-0.47	0.62	0.01
Chlorotriazine	TrCl	-11.97	0.83	-1.00	0.66	0.13
Normalization f	actor:		29.43	1.37	1.09	8.65

Table A 10. Normalized molecular descriptors of 30 molecules calculated by Dragon 6.0. and used to build Model II that predicts the interaction energies between m-OPs and the adsorption site II of PET MP. The molecules used for test set are underlined and typed in bold.

Compound	Abbrevation	$E_{\rm int}({\rm III}) / {\rm kcal \ mol^{-1}}$	TDB04v	CATS2D_03_DL	F02[O-O]
Benzene	Bz	-7.06	0.36	0.00	0.00
Toluene	PhMe	-9.01	0.60	0.00	0.00
<u>Styrene</u>	Sty	<u>-11.28</u>	0.77	0.00	0.00
Phenylacetylene	PhA	-11.18	0.88	0.00	0.00
Chlorobenzene	PhCl	-10.75	0.63	0.00	0.00
Bromobenzene	PhBr	-11.48	0.70	0.00	0.00
<u>Fluorobenzene</u>	PhF	<u>-9.37</u>	0.47	0.00	0.00
Phenol	PhOH	-16.59	0.51	0.50	0.00
Anisole	PhOMe	-9.32	0.67	0.00	0.00
Benzaldehide	PhCHO	-10.63	0.76	0.00	0.00
Acetophenone	AcPHO	-11.49	0.82	0.00	0.00
Benzoic acid	BA	-17.58	0.86	0.50	0.33
Phenylacetic acid	PAA	<u>-16.20</u>	0.81	0.25	0.33
Aniline	PhNH ₂	-12.95	0.54	0.50	0.00
<i>N</i> –Methylaniline	NMA	-11.81	0.86	0.00	0.33
N,N–Dimethylaniline	DMA	-13.83	0.67	0.50	0.00
Benzonitrile	PhCN	-13.60	0.82	0.00	0.00
Nitrobenzene	PhNO ₂	-9.68	0.73	0.00	0.00
<u>Benzamide</u>	BAM	<u>-16.27</u>	0.81	0.50	0.00
Thiophenol	PhSH	<u>-12.80</u>	0.67	0.00	0.00
Phenylsulfenic acid	PhSOH	-13.71	0.80	0.50	0.00
Phenylsulfinic acid	PhSOOH	-19.43	0.92	0.50	0.33
Phenylsulfonic acid	PhSO ₂ OH	-25.97	1.00	0.50	1.00
Benzenesulfonamide	BSAM	-15.74	0.95	0.50	0.33
Phenylphosphonic acid	PhPO(OH) ₂	-26.66	0.95	1.00	1.00
<i>o</i> -xylene	o-PhMe ₂	-11.29	0.74	0.00	0.00
Catechol	o-Ph(OH) ₂	-17.70	0.64	1.00	0.00
o-Phenyldiamine	$o-Ph(NH_2)_2$	-14.70	0.66	1.00	0.00
Triazine	Tr	-7.79	0.25	0.00	0.00
Chlorotriazine	TrCl	-9.35	0.67	0.00	0.00
Normalization fa	actor:		0.34	4.00	3.00

Table A 11. Normalized molecular descriptors of 30 molecules calculated by Dragon 6.0. and used to build Model III that predicts the interaction energies between m-OPs and the adsorption site III of PET MP. The molecules used for test set are underlined and typed in bold.

Compound	Abbrevation	$E_{\rm int}({\rm IV})$ / kcal mol ⁻¹	RDF050u	MLOGP2	SAdon
Benzene	Bz	-8.28	0.38	0.55	0.00
Toluene	PhMe	-9.18	0.32	0.73	0.00
<u>Styrene</u>	Sty	<u>-10.08</u>	0.40	0.88	0.00
Phenylacetylene	PhA	-9.27	0.35	0.88	0.00
Chlorobenzene	PhCl	<u>-8.88</u>	0.30	0.89	0.00
Bromobenzene	PhBr	-8.97	0.53	1.00	0.00
Fluorobenzene	PhF	-8.75	0.26	0.79	0.00
Phenol	PhOH	-15.37	0.25	0.24	0.57
Anisole	PhOMe	-11.53	0.79	0.37	0.00
Benzaldehide	PhCHO	-9.59	0.48	0.34	0.00
Acetophenone	AcPHO	<u>-13.64</u>	0.67	0.48	0.00
Benzoic acid	BA	-14.17	0.59	0.31	0.57
Phenylacetic acid	PAA	-13.98	0.47	0.34	0.57
Aniline	PhNH ₂	-12.79	0.24	0.24	0.44
<i>N</i> –Methylaniline	NMA	-10.23	0.39	0.38	0.00
N,N–Dimethylaniline	DMA	-11.22	0.43	0.37	0.24
Benzonitrile	PhCN	-9.17	0.28	0.34	0.00
Nitrobenzene	PhNO ₂	-12.31	0.68	0.52	0.00
Benzamide	BAM	-17.78	0.73	0.18	0.44
Thiophenol	PhSH	<u>-12.25</u>	0.47	0.69	0.00
Phenylsulfenic acid	PhSOH	-14.99	0.57	0.24	0.53
Phenylsulfinic acid	PhSOOH	-18.33	0.80	0.11	0.53
Phenylsulfonic acid	PhSO ₂ OH	-18.49	1.00	0.07	0.53
Benzenesulfonamide	BSAM	-20.58	0.92	0.02	0.42
Phenylphosphonic acid	PhPO(OH) ₂	-19.26	0.80	0.09	1.00
<i>o</i> -xylene	o-PhMe ₂	-9.19	0.50	0.93	0.00
Catechol	o-Ph(OH) ₂	<u>-15.69</u>	0.13	0.09	1.15
o-Phenyldiamine	$o-Ph(NH_2)_2$	-12.39	0.12	0.09	0.89
Triazine	Tr	-7.79	0.00	0.01	0.00
Chlorotriazine	TrCl	-8.65	0.00	0.12	0.00
Normalization f	actor:		7.41	9.27	74.26

Table A 12. Normalized molecular descriptors of 30 molecules calculated by Dragon 6.0. and used to build Model IV that predicts the interaction energies between m-OPs and the adsorption site IV of PET MP. The molecules used for test set are underlined and typed in bold.

Atom type	<i>m</i> (amu)	charge	p-type	σ (nm)	ε (kJ/mol)
НО	1.008	0	А	0.050	0.126
OH	15.999	0	А	0.300	0.711
Ο	15.999	0	А	0.296	0.879
OES	15.999	0	А	0.290	0.586
CO	12.012	0	А	0.375	0.276
CA	12.012	0	А	0.355	0.293
СТ	12.012	0	А	0.350	0.276
HA	1.008	0	А	0.242	0.126
HAE	1.008	0	А	0.242	0.126

Table A 13. The assigned OPLS-AA Lennard-Jones parameters and atomic mass.

Table A 14. The assigned OPLS-AA bond parameters.

Aton	n type	Function	$d(\mathbf{nm})$	$k (k I/(mol nm^2))$	
i	j	j		x (x3/(1101 1111))	
НО	OH	1	0.0945	442400	
OH	CO	1	0.1364	360000	
CO	Ο	1	0.1229	456000	
CO	CA	1	0.1490	320000	
CA	CA	1	0.1400	375200	
CA	HA	1	0.1080	293600	
CO	OES	1	0.1327	171200	
OES	CT	1	0.1410	256000	
CT	CT	1	0.1529	214400	
CT	HAE	1	0.1090	272000	
СТ	OH	1	0.1410	267776	

	Atom type	1	Function	θ (°)	$k (kJ/(mol/rad^2))$
l	J	K			
OH	CO	0	1	123.32	836.8
OH	CO	CA	1	111.88	711.28
0	CO	CA	1	120.4	669.44
CO	CA	CA	1	120	711.28
CA	CA	HA	1	120	292.88
CA	CO	OES	1	111.4	677.808
0	CO	OES	1	123.4	694.544
CO	OES	CT	1	116.9	694.544
OES	CT	HAE	1	109.5	292.88
CT	CT	HAE	1	110.7	313.8
OES	CT	CT	1	109.5	418.4
HO	OH	CO	1	113	292.88
CA	CA	CA	1	120	527.184
CT	CT	OH	1	109.5	418.4
HAE	СТ	OH	1	109.5	292.88
CT	OH	HO	1	108.5	460.24
HAE	CT	HAE	1	107.8	276.144

Table A 15. The assigned OPLS-AA angle parameters.

Table A 16. The assigned OPLS-AA dihedral parameters.

	Atom	n type		Eurotion	co	c_1	c_2	C ₃	c_4	c_5
i	j	k	l	Function	(kJ/mol)	(kJ/mol)	(kJ/mol)	(kJ/mol)	(kJ/mol)	(kJ/mol)
HO	OH	CO	0	3	33.616	-3.688	-29.928	0.000	0.000	0.000
CO	CA	CA	HA	3	29.000	0.000	-29.000	0.000	0.000	0.000
CO	CA	CA	CA	3	29.000	0.000	-29.000	0.000	0.000	0.000
CA	CA	CA	CA	3	29.000	0.000	-29.000	0.000	0.000	0.000
CA	CA	CA	HA	3	29.000	0.000	-29.000	0.000	0.000	0.000
CA	CA	CO	0	3	8.400	0.000	-8.400	0.000	0.000	0.000
CA	CA	CO	OES	3	7.900	0.000	-7.900	0.000	0.000	0.000
CA	CO	OES	CT	3	28.386	-2.398	-25.988	0.000	0.000	0.000
0	CO	OES	CT	3	24.000	0.000	-24.000	0.000	0.000	0.000
CO	OES	CT	CT	3	-4.326	2.446	3.760	-1.880	0.000	0.000
CO	OES	CT	HAE	3	0.322	0.966	0.000	-1.288	0.000	0.000
OES	CT	CT	HAE	3	0.936	2.808	0.000	-3.744	0.000	0.000
HAE	CT	CT	HAE	3	0.600	1.800	0.000	-2.400	0.000	0.000
CT	CT	OH	HO	3	-6.386	-9.030	15.416	0.000	0.000	0.000
HAE	CT	CT	OH	3	0.600	1.800	0.000	-2.400	0.000	0.000
HAE	CT	OH	HO	3	0.778	2.334	0.000	-3.112	0.000	0.000
HO	OH	CO	CA	3	29.288	-8.368	-20.920	0.000	0.000	0.000
OES	CT	CT	OES	3	-1.151	1.151	0.000	0.000	0.000	0.000
OES	CT	CT	OH	3	-1.151	1.151	0.000	0.000	0.000	0.000
OH	CO	CA	CA	3	8.786	0.000	-8.786	0.000	0.000	0.000



Figure A 17. Energy minimization and equilibration of a) simazine, b) atrazine, c) diuron, d) alachlor and e) isoproturon molecule in water in NVT and NPT ensemble.

Appendix B



Figure A 18. SEM images of pristine (left) and thermally aged (right) PET MP.

Appendix C

Listing A 1. Python code used to extract quantitative results from Gaussian output file (.out) after optimization or single point calculation.

```
1. #Extract HF, RMSD, Dipole, HOMO, LUMO from Gaussian out files (.out) in csv file
 2.
 3. #input your path to the folder containing output files
 4. pathS = "/path_to_foder_containing_your_output_files/"
 5.
 6. import subprocess
7.
 8. def sysstout(cmd):
9.
        cnt=0;rc=1
10.
        process = subprocess.Popen(
11.
                    cmd, stdout=subprocess.PIPE, stderr=subprocess.PIPE,
12.
                    shell=True,universal_newlines=True
13.
14.
        process.wait()
15.
        rc=process.returncode
16.
        return process.communicate()
17.
18. import numpy as np
19.
20. def find_indices(list_to_check, item_to_find):
21.
        array = np.array(list_to_check)
22.
        indices = np.where(array == item_to_find)[0]
23.
        return list(indices)
24.
25. lSL=sysstout("ls "+pathS)[0].split()
26.
27. outSL = [lineS for lineS in ISL if ".out" in lineS and "slurm" not in lineS ]
28. resultSLL = []
29. resultSL = []
30.
31. for out in outSL:
32.
        fo= open(pathS + out)
33.
        do = fo.readlines()
34.
        fo.close
        d = "".join(do).replace("\n", "").replace(" ", "").split("\\")
35.
        nSL = [lineS for lineS in d if "HF=" in lineS or "RMSD=" in lineS
36.
                    or "Dipole=" in lineS ][-3:]
37.
        numSL = [lineS[lineS.index("=") + 1:] for lineS in nSL]
38.
39.
        dipoleL = [float(result) for result in numSL[2].split(",")]
        dipoleS = str((dipoleL[0]**2+dipoleL[1]**2+dipoleL[2]**2)**0.5 / 0.393430307)
40.
        oSL = [lineS for lineS in do if "Alpha occ. eigenvalues" in lineS][-1]
41.
42.
        HOMO=oSL[oSL.index("\n")-8:oSL.index("\n")]
43
        a = find_indices(do,
                                          Population analysis using the SCF Density.\n")
44.
        bo=do[a[-1]:]
45.
        oSL = [lineS for lineS in bo if "Alpha virt. eigenvalues" in lineS][0]
        LUMO=oSL[oSL.index("--")+2:oSL.index("--")+13]
46.
47.
        resultSL = [out, numSL[0], numSL[1], dipoleS, HOMO, LUMO, "\n"]
48.
        resultSLL += [", ".join(resultSL)]
49.
50. result = "name, HF, RMSD, Dipole, HOMO, LUMO \n"
51. resultS=" ".join(resultSLL)
52.
53. f = open(pathS + "HF RMSD DIPOLE.txt", "wt")
54. f.write(result)
55. f.write(resultS)
56. f.close
```

Listing A 2. Python code used to extract quantitative results from Gaussian output file (.out) after frequency calculation.

```
1. #Extract results from frequency calculation from Gaussian out files (.out) in csv file
 2. #ZPC, DE, DH, DG, Eo, E, H, G, E(thermal), Cv, S, Polarizability
 3.
 4. #input Path to the folder containing output files
 5. pathS = "/path to foder containing your output files/"
 6.
7. import subprocess
 8.
9. def sysstout(cmd):
10.
        cnt=0;rc=1
        process = subprocess.Popen(
11.
                    cmd, stdout=subprocess.PIPE, stderr=subprocess.PIPE,
12.
                    shell=True,universal_newlines=True
13.
14.
                    )
15.
        process.wait()
16.
        rc=process.returncode
17.
        return process.communicate()
18.
19. lSL=sysstout("ls "+pathS)[0].split()
20. outSL = [lineS for lineS in ISL if ".out" in lineS and "slurm" not in lineS ]
21. resultSLL = []
22. resultSL = []
23.
24. for out in outSL:
25.
        fo= open(pathS + out)
26.
        do = fo.readlines()
27.
        fo.close
        ZPCSL = [lineS for lineS in do if "Zero-point correction" in lineS][:]
28.
29.
        ZPCnumSL = [lineS[lineS.index("=") + 1:-19] for lineS in ZPCSL]
        DESL = [lineS for lineS in do if "Thermal correction to Energy" in lineS][:]
30.
31.
        DEnumSL = [lineS[lineS.index("=") + 1:-1] for lineS in DESL]
        DHSL = [lineS for lineS in do if "Thermal correction to Enthalpy" in lineS][:]
32.
        DHnumSL = [lineS[lineS.index("=") + 1:-1] for lineS in DHSL]
33.
34.
        DGSL = [lineS for lineS in do if "Thermal correction to Gibbs Free Energy" in lineS][:]
35.
        DGnumSL = [lineS[lineS.index("=") + 1:-1] for lineS in DGSL]
        EoSL = [lineS for lineS in do if "Sum of electronic and zero-point Energies" in lineS][:]
36.
37.
        EonumSL = [lineS[lineS.index("=") + 1:-1] for lineS in EoSL]
        ESL = [lineS for lineS in do if "Sum of electronic and thermal Energies" in lineS][:]
38.
39.
        EnumSL = [lineS[lineS.index("=") + 1:-1] for lineS in ESL]
        HSL = [lineS for lineS in do if "Sum of electronic and thermal Enthalpies" in lineS][:]
40.
        HnumSL = [lineS[lineS.index("=") + 1:-1] for lineS in HSL]
41.
        GSL = [lineS for lineS in do if "Sum of electronic and thermal Free Energies" in lineS][:]
42.
43.
        GnumSL = [lineS[lineS.index("=") + 1:-1] for lineS in GSL]
        PSL = [lineS for lineS in do if "Isotropic polarizability" in lineS][:][0]
44.
45.
        P = PSL[45:-9]
46.
        a=do[do.index("
                                             E (Thermal)
                                                                     CV
                                                                                        S\n")+2]
47.
        E = a[23:32]
48.
        CV=a[35:53]
49
        S=a[53:-1]
50
51.
        resultSL = [out, ZPCnumSL[0], DEnumSL[0], DHnumSL[0], DGnumSL[0],
52.
                    EonumSL[0], EnumSL[0], HnumSL[0], GnumSL[0], E, CV, S, P ]
        resultSLL += [", ".join(resultSL) + "\n"]
53.
54
55. result = "name, ZPC/Ha, DE/Ha, DH/Ha, DG/G, Eo/Ha, E/Ha, H/Ha, G/Ha, E(thermal)/kcal mol-1, " \
56.
             "Cv/cal mol-1 K-1, S/cal mol-1, Polarizability \n"
57. resultS=" ".join(resultSLL)
58. f = open(pathS + "freq.txt", "wt")
59. f.write(result)
60. f.write(resultS)
61. f.close
```

Listing A 3. Python code to create two .gjf files from one containing the geometry of the system with 2 fragments. The code creates 2 more files that contain only geometry of a single fragment where the second one is represented as ghost molecule. Used to calculate interaction energies using CPC method.

```
1. #creates input files to calculate interaction energy of 2 fragments using cpc
 2. #results in 3 new input files used to calculate energies of each fragment, and of a whole system
 з.
 4. #input your path to the folder containing output files
 5. pathS = "/path_to_foder_containing_your_input_file/"
 6.
7. #define the multiplicity and the charge of the system
8. mab = "2 1"
9.
10. #define the multiplicity and the charge for each fragment
11. ma = "2 1"
12. mb = "0 1"
13.
14. a = ma + " \n"
15. b = mb + " \setminus n"
16. m = mab + "" + ma + "" + mb
17.
18. import subprocess
19.
20. def sysstout(cmd):
21.
        cnt=0;rc=1
22.
        process = subprocess.Popen(
23.
                    cmd, stdout=subprocess.PIPE,
                    stderr=subprocess.PIPE,
24.
25.
                    shell=True, universal_newlines=True
26.
                    )
27.
        process.wait()
28.
        rc=process.returncode
29.
        return process.communicate()
30.
31. lSL=sysstout("ls "+pathS)[0].split()
32. filenameS = lSL[0].replace(".gjf",
                                        '")
33.
34. fo = open(pathS + 1SL[0])
35. inpL = fo.readlines()
36. fo.close()
37.
38. e abL = inpL
39. e_aL=[ a if m in lineS else lineS[:lineS.index("(")]+"-Bq"+lineS[lineS.index("("):]
         if "Fragment=2)" in lineS else lineS for lineS in inpL ]
40.
41. e_bL=[ b if m in lineS else lineS[:lineS.index("(")]+"-Bq"+lineS[lineS.index("("):]
          if "Fragment=1)" in lineS else lineS for lineS in inpL
42.
43.
44. inputSLL = [e_abL, e_aL, e_bL]
45. fileextSL = ["_Eab_AB", "_Ea_ABC", "_Eb_ABC", "_Ec_ABC"]
46. inputSL=[]
47.
48. for i in range(len(inputSLL)):
        inputSLL[i][2] = "%chk = " + filenameS + fileextSL[i] + ".chk\n"
49.
50.
51. for i in range(len(inputSLL)):
52.
        inputSL.append("".join(inputSLL[i]))
53.
54. for i in range(len(inputSL)):
55.
        fo = open(pathS + filenameS + fileextSL[i] + ".gjf", "wt")
56.
        fo.write(inputSL[i])
57.
        fo.close()
58.
59. args = "rm", pathS + 1SL[0]
60. subprocess.call('%s %s ' % args, shell=True)
```

Listing A 4. Python code to create 6 .gjf files from one containing the geometry of the system with 3 fragments. The code creates 6 more files of which 3 contain only geometry of a single fragment where the other two are represented as ghost molecules and 3 contain the geometries of 2 fragments where the third one is represented as ghost molecule. Used to calculate interaction energies of systems with three fragments using CPC method. Modification of code from Listing A 3.

```
1. #creates input files to calculate interaction energy between 3 fragments using cpc.
 2. #reusult in 7 input files in total.
 3.
 4. #input your path to the folder containing output files
 5. pathS = "/path_to_foder_containing_your_input_file/"
 6.
7. #define the multiplicity and the charge of the system
 8. mabc = "2 1"
9.
10. #define the multiplicity and the charge for each fragment
11. ma = "0 1"
12. mb = "0 1"
13. mc = "2 1"
14.
15. #define the multiplicity and the charge for combinations of fragments
16. mab = "0 1"
17. mac = "2 1"
18. mbc = "2 1"
19.
20. a = ma + " \setminus n"
21. b = mb + " \setminus n"
22. c = mc + " \setminus n"
23. ab = mab + " " + ma + " " + mb + " \n"
24. ac = mac + " " + ma + " " + mc + " \n"
25. bc = mbc + " " + mb + " " + mc + " n"
26. m = mabc + " " + ma + " " + mb + " " + mc
27.
28. import subprocess
29.
30. def sysstout(cmd):
        cnt=0;rc=1
31.
32.
        process = subprocess.Popen(
33.
                    cmd, stdout=subprocess.PIPE,
34.
                    stderr=subprocess.PIPE,
                    shell=True, universal_newlines=True
35.
36.
                    )
37.
        process.wait()
38.
        rc=process.returncode
39.
        return process.communicate()
40.
41. lSL=sysstout("ls "+pathS)[0].split()
42. filenameS = lSL[0].replace(".gjf",
                                        "")
43.
44. fo = open(pathS + 1SL[0])
45. inpL = fo.readlines()
46. fo.close()
47.
48. e_abcL = inpL
49. e_aL=[ a if m in lineS else lineS[:lineS.index("(")]+"-Bq"+lineS[lineS.index("("):]
50.
         if "Fragment=2)" in lineS or "Fragment=3" in lineS else lineS for lineS in inpL
51. e_bL=[ b if m in lineS else lineS[:lineS.index("(")]+"-Bq"+lineS[lineS.index("("):]
          if "Fragment=1)" in lineS or "Fragment=3" in lineS else lineS for lineS in inpL ]
52.
53. e_cL=[ c if m in lineS else lineS[:lineS.index("(")]+"-Bq"+lineS[lineS.index("("):]
54.
          if "Fragment=1)" in lineS or "Fragment=2" in lineS else lineS for lineS in inpL ]
55.
56. e_abL=[ ab if m in lineS else lineS[:lineS.index("(")]+"-Bq"+lineS[lineS.index("("):]
          if "Fragment=3" in lineS else lineS for lineS in inpL ]
57.
58. e_acL=[ ac if m in lineS else lineS[:lineS.index("(")]+"-Bq"+lineS[lineS.index("("):]
```

```
59.
          if "Fragment=2" in lineS else lineS for lineS in inpL ]
60. e_bcL=[ bc if m in lineS else lineS[:lineS.index("(")]+"-Bq"+lineS[lineS.index("("):]
          if "Fragment=1" in lineS else lineS for lineS in inpL ]
61.
62.
63. inputSLL = [e_abcL,e_abL, e_acL, e_bcL,e_aL, e_bL, e_cL ]
64. fileextSL = ["_Eabc_ABC", "_Eab_ABC", "_Eac_ABC", "_Ebc_ABC", "_Ea_ABC", "_Eb_ABC", "_Ec_ABC"]
65. inputSL=[]
66.
67. for i in range(len(inputSLL)):
         inputSLL[i][2] = "%chk = " + filenameS + fileextSL[i] + ".chk\n"
68.
69.
70. for i in range(len(inputSLL)):
71. inputSL.append("".join(inputSLL[i]))
72.
73. for i in range(len(inputSL)):
74.
         fo = open(pathS + filenameS + fileextSL[i] + ".gjf", "wt")
         fo.write(inputSL[i])
75.
76.
         fo.close()
77.
78. args = "rm", pathS + 1SL[0]
79. subprocess.call('%s %s ' % args, shell=True)
```

Listing A 5. Python code used to extract energies and calculate counterpoise corrected interaction energy from three Gaussian output (.out) files that calculate the energy of a whole system and single fragments where the other fragment is represented as ghost atoms (the output files after the calculations of input files (.gjf) created by code presented in Listing A 3).

```
1. #Extracts interaction energies calculated using cpc
2.
 3. #input your path to the directory that contains output files
 4. pathS = "/path_to_foder_containing_your_output_files/"
 5.
 6. import subprocess
 7.
8. def sysstout(cmd):
9.
        cnt=0;rc=1
10.
        process = subprocess.Popen(
11.
                    cmd, stdout=subprocess.PIPE,
12.
                    stderr=subprocess.PIPE,
                    shell=True, universal_newlines=True)
13.
14.
        process.wait()
15.
        rc=process.returncode
16.
        return process.communicate()
17.
18. lSL=sysstout("ls "+pathS)[0].split()
19. outSL = [lineS for lineS in ISL if ".out" in lineS and "slurm" not in lineS]
20. SCFeL = []
21.
22. for out in outSL:
23.
        fo = open(pathS + out)
        d = fo.readlines()
24.
25.
        fo.close
        elineS = [lineS for lineS in d if "SCF Done" in lineS][0]
26.
27.
        SCFe = float(elineS.split()[4])
28.
        SCFeL+=[SCFe]
29.
30. corr = (SCFeL[1]-SCFeL[0]-SCFeL[2])*627.5
31. Eab =SCFeL[1]
32. Ea =SCFeL[0]
33. Eb =SCFeL[2]
34.
35. resultL = [corr, Eab, Ea, Eb]
36. resultSL = [str(result) for result in resultL]
37. resultS = outSL[0][:-10] +","+ ", ".join(resultSL) + "\n"
38.
39. nresultS = "file name, CP_corrected_energy/kcal mol-1, Eab/Ha, Ea/Ha, Eb/Ha\n"
40.
41. f = open(pathS + "cpc2_reduced_results.txt", "wt")
42. f.write(nresultS)
43. f.write(resultS)
44. f.close
```

Listing A 6. Python code used to extract energies and calculate counterpoise corrected interaction energy from multiple directories that contain three Gaussian output (.out) files that calculate the energy of a whole system and single fragments where the other fragment is represented as ghost atoms (the output files after the calculations of input files (.gjf) created by code presented in Listing A 3). Modification of code from Listing A 5.

```
1. #Extracts interaction energies calculated using cpc
 2. #extracts from multiple directories
 3.
 4. #input your path to the directory that contains directories with output files
 5. pathS = "/path_to_directory/"
 6.
7. import subprocess
8.
9. def sysstout(cmd):
10.
        cnt=0;rc=1
11.
        process = subprocess.Popen(
12.
           cmd, stdout=subprocess.PIPE, stderr=subprocess.PIPE,shell=True, universal_newlines=True)
13.
        process.wait()
14.
        rc=process.returncode
15.
        return process.communicate()
16.
17. dirSL=sysstout("ls "+pathS)[0].split()
18.
19. resultSLL = []
20.
21. for dir in dirSL:
        ISL = sysstout("ls "+pathS + dir)[0].split()
22.
23.
        outSL = [lineS for lineS in ISL if ".out" in lineS and "slurm" not in lineS]
        SCFeL = []
24.
25.
        for out in outSL:
26.
            fo = open(pathS+ dir +"/"+ out)
            d = fo.readlines()
27.
28.
            fo.close
29.
            elineS = [lineS for lineS in d if "SCF Done" in lineS][0]
30.
            SCFe = float(elineS.split()[4])
31.
            SCFeL+=[SCFe]
32.
        corr = (SCFeL[1]-SCFeL[0]-SCFeL[2])*627.5
33.
        Eab =SCFeL[1]
34.
        Ea =SCFeL[0]
35.
        Eb =SCFeL[2]
36.
        resultL = [corr, Eab, Ea, Eb]
37.
        resultSL = [str(result) for result in resultL]
        resultSLL += [outSL[0][:-10] +","+ ", ".join(resultSL) + "\n"]
38.
39.
40. resultS = "".join(resultSLL)
41. nresultS = "file name, CP corrected energy/kcal mol-1, Eab/Ha, Ea/Ha, Eb/Ha\n"
42.
43. f = open(pathS + "cpc2_reduced_results.txt", "wt")
44. f.write(nresultS)
45.
46. f.write(resultS)
47. f.close()
```

List of Figures

Figure 2.1. Structures of monomers of the most produced polymers
Figure 2.2. The exponential growth of the plastic production and the global demand for the raw
plastic material according to the type of polymer in 2022, according to the market
research from Plastic Europe. ^[23] 4
Figure 2.3. Number of published papers per year based on a search of keywords microplastic* OR
nanoplastic* AND adsorption in Scopus database6
Figure 2.4. Interactions of MP and organic pollutants in water9
Figure 2.5.General mechanism of photooxidation11
Figure 2.6. Mechanism of polyethylene photooxidation: a) chain-scission and propagation; b)
Norrish type I; c) Norrish type II mechanism. ^[93] 12
Figure 2.7. Jacob's ladder of DFT functionals
Figure 2.8. Flow diagram of general single point calculation and geometry optimization process.
Figure 2.9. Graphic representation of the main terms defining the potential energy in classical force
field, including potential energy of bond stretching $(Ubond(d))$, bond-angle bending
$(Uangle(\theta))$, dihedral-angle torsion $(Udihedral(\phi))$, improper dihedral-angle bending
(Uimproper ψ), and electrostatic (Uele(r)) and van der Waals interactions
(<i>U</i> vdW(<i>r</i>))35
Figure 2.10. Illustration of periodic boundary conditions (PBC)
Figure 2.11. A general scheme of QSA/PR model development47
Figure 4.1. Structures of common organic pollutants in water
Figure 4.2. Organic pollutant molecular models61
Figure 4.3. Structural representation of terephthalic acid (TA) and the model pollutants in two
configurations. The blue sphere represents the position of model pollutant's functional
group63
Figure 4.4. a) Interaction energies and b) basis set superposition error calculated for the interactions
of the model organic pollutants and terephthalic acid in two configurations (TA (upper)
I and TA II (lower)) using MP2 method

Figure 4.15. Schematic representations of the interactions between adsorption site I of mPET and
a) benzoic acid; b) benzamide; c) aniline; d) benzene, and between phenol and mPET
on e) adsorption site I; f) adsorption site II; g) adsorption site III; h) adsorption site IV
analysed by independent gradient model (IGMH) and visualised on 0.01 a.u.
isosurface. In the usual colouring scheme, the blue colour represents attractive
interactions, the green colour represents van der Waals interactions, and the red colour
represent steric effects80
Figure 4.16. Interaction energies of phenol and mPET with no other phenols adsorbed and with
multiple phenols adsorbed on different parts of mPET surface. Highlighted phenol
molecules correspond to ones whose interaction energies are calculated
Figure 4.17. Interaction energies of phenol with pristine (mPET) and aged (mPET-OH, mPET-
COOH, and mPET=O) PET model at M06-L level of theory and def2-SVP basis set.
Figure 4.18. The structure of mPET–COOH with twelve phenols adsorbed optimized in gas phase
at M06-L level of theory and def2-SVP basis set
Figure 4.19. Structures of 30 compounds representing common organic pollutants (m-OPs) to study
the adsorption interactions with model PET surface
Figure 4.20. The values of coefficients of determination (R2) of the training set and test of the
QSPR models with 1-5 variables developed to predict the interaction energies of m-
OPs with adsorption sites I, II, III and IV of PET MP89
Figure 4.21. Scatter plot of Leave Many Out (LMO) technique performed to show the robustness
of the selected QSPR models for the prediction of interaction energies between m-OPs
and PET MP on adsorption sites I, II, III and IV.
Figure 4.22. Scatter plot of Y-scrambling validation method for the selected QSPR models for the
prediction of interaction energies between m-OPs and PET MP on adsorption sites I,
II, III and IV91
Figure 4.23. Calculated interaction energies and interaction energies predicted by QSPR Model I,
Model II, Model III and Model IV build using training set (25 molecules) for the
prediction of interaction energies of m-OPs with the adsorption sites I, II, III and IV on
PET MP respectively

Figure 4.24. Williams plots of QSPR Model I, Model II, Model III and Model IV for the prediction
of interaction energies of m-OPs with the adsorption sites I, II, III and IV on PET MP
respectively93
Figure 4.25. Cross-correlation matrix of descriptor variables in QSPR Model I, Model II, Model
III and Model IV for the prediction of interaction energies of m-OPs with the adsorption
sites I, II, III and IV respectively95
Figure 4.26. Calculated interaction energies and corresponding models' equations of QSPR Model
I, Model II, Model III and Model IV that predict interaction energies of m-OPs with
the adsorption sites I, II, III and IV on PET MP respectively when applied on the entire
set of data (30 molecules) with normalized descriptors
Figure 4.27. Plot of residuals for QSPR Model I, Model II, Model III and Model IV that predict
interaction energies of m-OPs with the adsorption sites I, II, III and IV on PET MP
respectively when applied on the entire set of data (30 molecules) with normalized
descriptors
Figure 4.28. Structure oh PET: Head (blue), Body (yellow; repeating unit) and Tail (red) and their
atom labels in generated OPLS-AA force field108
Figure 4.29. Molecular structure of pesticides simazine (SMZ), atrazine (ATZ), diuron (DIU)
alachlor (ALC) and isoproturon (IPT)108
Figure 4.30. Energy minimization and equilibration of simazine in water in NVT and NPT
ensemble109
Figure 4.31. Interaction energies of organic pollutants (pesticides) with water calculated from MD
trajectories110
Figure 4.32. Two models of PET microplastic used for MD simulation of pesticides adsorption to
PET microplastic. The structure on the left is designed as one folded 150-mer chain,
and the one on the right as an aggregate of 30 5-mer chains that represent fragmented
PET microplastic111
Figure 4.33. Interaction energies between pesticides and PET model I, and PET model II
(fragmented model) calculated from MD trajectories114
Figure 4.34. Structures of alachlor adsorption on PET model I and PET model II
Figure 4.35. Structures of diuron, alachlor and isoproturon adsorption on PET model I and PET
model II

- Figure A 7. Schematic representations of the interactions between the model pollutants and mPET at adsorption site II analysed by independent gradient model (IGMH) and visualised on 0.01 a.u. isosurface. In the usual colouring scheme, the blue colour represents

attractive interactions, the green colour represents van der Waals interactions, and the
red colour represents non-bonding interactions
Figure A 8. Schematic representations of the interactions between the model pollutants and mPET
at adsorption site III analysed by independent gradient model (IGMH) and visualised
on 0.01 a.u. isosurface. In the usual colouring scheme, the blue colour represents
attractive interactions, the green colour represents van der Waals interactions, and the
red colour represents non-bonding interactions
Figure A 9. Schematic representations of the interactions between the model pollutants and mPET
at adsorption site IV analysed by independent gradient model (IGMH) and visualised
on 0.01 a.u. isosurface. In the usual colouring scheme, the blue colour represents
attractive interactions, the green colour represents van der Waals interactions, and the
red colour represents non-bonding interactions
Figure A 10. Schematic representations of the interactions between the water molecule and mPET
at four adsorption sites analysed by independent gradient model (IGMH) and visualised
on 0.01 a.u. isosurface. In the usual colouring scheme, the blue colour represents
attractive interactions, the green colour represents van der Waals interactions, and the
red colour represents non-bonding interactions
Figure A 11. Structures of PET model with benzene, toluene, styrene, phenylacetylene and
chlorobenzene adsorbed on four adsorption sites. The structures were optimized by
M06-L/def2-SVP/svpfit method136
Figure A 12. Structures of PET model with bromobenzene, fluorobenzene, phenol, anisole. and
benzaldehyde adsorbed on four adsorption sites. The structures were optimized by
M06-L/def2-SVP/svpfit method137
Figure A 13. Structures of PET model with acetophenone, benzoic acid, phenylacetic acid, aniline,
and N-methylaniline adsorbed on four adsorption sites. The structures were optimized
by M06-L/def2-SVP/svpfit method138
Figure A 14. Structures of PET model with N,N-dimethylaniline, benzonitrile, nitrobenzene,
benzamide, and thiophenol adsorbed on four adsorption sites. The structures were
optimized by M06-L/def2-SVP/svpfit method139

- Figure A 15. Structures of PET model with phenylsulfenic acid, phenylsulfinic acid, phenylsulfonic acid, benzenesulfonamide, and phenylphosphonic acid adsorbed on four adsorption sites. The structures were optimized by M06-L/def2-SVP/svpfit method......140

Figure A 18. SEM images of pristine (left) and thermally aged (right) PET MP......150

List of Tables

Table 2.1. Physicochemical properties of the most used types of plastics. Tg is the glass transition
temperature, Tm is the temperature of melting, ρ is the average density and χ is the
percent crystallinity. ^[88] 11
Table 4.1. Interactions energies between 30 model organic pollutants (m-OPs) and model PET MPs
surface on four adsorption sites (I, II, III and IV) used as a response variable for QSPR
modelling (bold underlined values represent those used for the test sets, while others
were part of training sets)
Table 4.2. Descriptive statistical data of the coefficients in in QSPR Model I, Model II, Model III
and Model IV for the prediction of interaction energies of m-OPs with the adsorption
sites I, II, III and IV respectively build using training sets (25 molecules) with
normalized descriptors94
Table 4.3. Definitions of descriptors used in QSPR Model I, Model II, Model III and Model IV for
the prediction of interaction energies of m-OPs with the adsorption sites I, II, III and
IV respectively
Table 4.4. Descriptive statistical data of the coefficients in in QSPR Model I, Model II, Model III
and Model IV for the prediction of interaction energies of m-OPs with the adsorption
sites I, II, III and IV respectively when applied on the entire set of data with normalized
descriptors (30 molecules)101
Table 4.5. Equilibrium adsorption capacity (q) for the adsorption of diuron, isoproturon and
alachlor (initial concentration 50 µmol dm ⁻³ , 1 mL) on pristine and thermally aged PET
MP (100–200 $\mu m;$ 100 mg) at the temperature of 25 °C and rotation speed of 120 rpm.

- Table A 1. Interaction energies of TA and organic pollutants in a system where the functional groupof pollutant interacts with the carboxyl group of terephthalic acid (TA I) calculated byMP2/def2-TZVP method and various DFT methods with def2-SVP basis set.........125
- Table A 2. Interaction energies of TA and organic pollutants in a system where the aromatic ring of pollutant is placed on top of the aromatic ring of terephthalic acid (TA II) calculated by MP2/def2-TZVP method and various DFT methods with def2-SVP basis set....126

Table A 3. The difference between MP2/ def2-TZVP and DFT/ def2-SVP interaction energies
shown in Table A 1. calculated by MP2/def2-TZVP method and various DFT methods
with def2-SVP basis set127
Table A 4. The difference between MP2/ def2-TZVP and DFT/ def2-SVP interaction energies
shown in Table S2128
Table A 5. The dependence of basis set superposition error (BSSE) for the interaction of organic
pollutants and TA calculated by MP2/def2-TZVP method and various DFT methods
with def2-SVP basis set129
Table A 6. X, Y, Z coordinates of optimized geometry of polyethylene terephthalate model
(mPET)
Table A 7. Statistical evaluation of QSPR models for predicting interaction energies between m-
OPs and the adsorption sites I, II, III and IV of PET MP derived for training set (25
molecules) and applied on test set (5 molecules)142
Table A 8. The fitting criteria values and the values of internal and external validation criteria for
the selected models QSPR Model I, Model II, Model III and Model IV for the
interaction energies between m-OPs and the adsorption sites I, II, III and IV of PET
MP respectively
Table A 9. Normalized molecular descriptors of 30 molecules calculated by Dragon 6.0. and used
to build Model I that predicts the interaction energies between m-OPs and the
adsorption site I of PET MP. The molecules used for test set are underlined and typed
in bold143
Table A 10. Normalized molecular descriptors of 30 molecules calculated by Dragon 6.0. and used
to build Model II that predicts the interaction energies between m-OPs and the
adsorption site II of PET MP. The molecules used for test set are underlined and typed
in bold144
Table A 11. Normalized molecular descriptors of 30 molecules calculated by Dragon 6.0. and used
to build Model III that predicts the interaction energies between m-OPs and the
adsorption site III of PET MP. The molecules used for test set are underlined and typed
in bold145

Table A 12. Normalized molecular descriptors of 30 molecules calculated by Dragon 6.0. and us	sed
to build Model IV that predicts the interaction energies between m-OPs and t	the
adsorption site IV of PET MP. The molecules used for test set are underlined and typ	ed
in bold14	46
Table A 13. The assigned OPLS-AA Lennard-Jones parameters and atomic mass.	47
Table A 14. The assigned OPLS-AA bond parameters. 14	47
Table A 15. The assigned OPLS-AA angle parameters. 1	48
Table A 16. The assigned OPLS-AA dihedral parameters. 14	48

List of Abbreviations

AcPhO	Acetophenone
AD	Applicability domain
AIMD	Ab initio molecular dynamics
ALC	Alachlor
ATZ	Atrazine
BA	Benzoic acid
BAM	Benzamide
BLYP	Becke, Lee, Parr and Yang
BSAM	Benzenesulfonamide
BSSE	Basis set superposition error
Bz	Benzene
CC	Coupled cluster
CECs	Contaminants of emerging concern
CI	Configuration interaction
СР	Counterpoise-correction
CSVR	Canonical sampling through velocity rescaling
DFT	Density functional theory
DIU	Diuron
DMA	<i>n</i> , <i>n</i> -dimethylaniline
DOM	Dissolved organic matter
ECHA	European chemicals agency
FF	Force field
FTIR	Fourier transform infrared spectroscopy
GA	Genetic algorithm
GATEWAY	Geometry, topology and atom-weights assembly
GGA	Generalized gradient approximation
GTO	Gaussian type orbital
HDPE	High-density polyethylene
HF	Hartree-Fock
HPLC	High-performance liquid chromatography
IGM	Independent gradient model
IGMH	Independent gradient model based on Hirshfeld partition
IP	Intraparticle diffusion
IPT	Isoproturon
KS	Kohn-Sham

LDA	Local density approximation
LDPE	Low-density polyethylene
LFER	Linear free energy relationship
LMO	Leave-many-out
LOO	Leave-one-out
LSDA	Local spin density approximation
LSER	Linear solvation energy relationship
LJ	Lennard-jones
MAE	Mean average error
MD	Molecular dynamic
MEP	Molecular electrostatic surface potentials
MIM	Molecular influence matrix
MLR	Multiple linear regression
MLRA	Multiple linear regression analysis
MM	Molecular mechanics
m-OP	Model organic pollutant
MP	Microplastic
MP ₂	Second order Møller-Plesset perturbation theory
mPET	Model polyethylene terephthalate
NCI	Non-covalent interaction
NMA	N-methylaniline
NP	Nanoplastic
OLS	Ordinary least square
$o-Ph(NH_2)_2$	o-phenyldiamine
o-Ph(OH) ₂	Catechol
o-PhMe ₂	o-xylene
PA	Polyamide
PAA	Phenylacetic acid
PBC	Periodic boundary conditions
PBE	Perdew, Burke and Ernzerhof
PE	Polyethylene
PES	Potential energy surface
PET	Polyethylene terephthalate
PFO	Pseudo-first order
PhA	Phenylacetylene
PhBr	Bromobenzene
PhCHO	Benzaldehyde
PhCl	Chlorobenzene

PhCN	Benzonitrile
PhF	Fluorobenzene
PhMe	Toluene
PhNH ₂	Aniline
PhNO ₂	Nitrobenzene
PhOH	Phenol
PhOMe	Anisole
PhPO(OH) ₂	Phenylphosphonic acid
PhSH	Thiophenol
PhSO ₂ OH	Phenylsulfonic acid
PhSOH	Phenylsulfenic acid
PhSOOH	Phenylsulfinic acid
PME	Particle-mesh Ewald method
PP	Polypropylene
PS	Polystyrene
PSO	Pseudo-second order
PVC	Polyvinyl chloride
QM	Quantum mechanic
QSA/PR	Quantitative structure activity/property relationship
RDG	Reduced density gradient
REACH	Registration, evaluation, authorisation and restriction of chemicals
RMSE	Root mean square errors
SCF	Self-consistent field
SCRF	Self-consistent reaction field
SD	Slater determinant
SMZ	Simazine
SP	Single point
SPC/E	Extended simple point charge
SSA	Specific surface area
STO	Slater type orbital
Sty	Styrene
ТА	Terephthalic acid
TIPnP	Transferable intermolecular potential
Tr	Triazine
TrCl	Chlorotriazine
VMD	Visual molecular dynamics
XPS	X-ray photoelectron spectroscopy
References

- N. B. Hartmann, T. Hüffer, R. C. Thompson, M. Hassellöv, A. Verschoor, A. E. Daugaard, S. Rist, T. Karlsson, N. Brennholt, M. Cole, M. P. Herrling, M. C. Hess, N. P. Ivleva, A. L. Lusher, M. Wagner, Are We Speaking the Same Language? Recommendations for a Definition and Categorization Framework for Plastic Debris, Environ. Sci. Technol. 53 (2019) 1039–1047.
- [2] C. Wang, J. Zhao, B. Xing, Environmental source, fate, and toxicity of microplastics, J. Hazard. Mater. 407 (2021) 124357.
- [3] J. Li, H. Liu, J. Paul Chen, Microplastics in freshwater systems: A review on occurrence, environmental effects, and methods for microplastics detection, Water Res. 137 (2018) 362– 374.
- [4] M. Cole, P. Lindeque, C. Halsband, T. S. Galloway, Microplastics as contaminants in the marine environment: A review, Mar. Pollut. Bull. 62 (2011) 2588–2597.
- [5] D. Materić, H. A. Kjær, P. Vallelonga, J. L. Tison, T. Röckmann, R. Holzinger, Nanoplastics measurements in Northern and Southern polar ice, Environ. Res. 208 (2022) 112741.
- [6] D. Materić, M. Peacock, J. Dean, M. Futter, T. Maximov, F. Moldan, T. Rockmann, R. Holzinger, Presence of nanoplastics in rural and remote surface waters, Environ. Res. Lett. 17 (2022) 054036.
- [7] A. Omidi, H. Naeemipoor, M. Hossein, Plastic Debris in the Digestive Tract of Sheep and Goats : An Increasing Environmental Contamination in Birjand , Iran, Bull Env. Contam Toxicol 88 (2012) 691–694.
- [8] R. Wang, H. Mou, X. Lin, H. Zhu, B. Li, J. Wang, M. Junaid, J. Wang, Microplastics in Mollusks: Research Progress, Current Contamination Status, Analysis Approaches, and Future Perspectives, Front. Mar. Sci. 8 (2021) 1–12.
- [9] G. Oliveri Conti, M. Ferrante, M. Banni, C. Favara, I. Nicolosi, A. Cristaldi, M. Fiore, P. Zuccarello, Micro- and nano-plastics in edible fruit and vegetables. The first diet risks assessment for the general population, Environ. Res. 187 (2020) 109677.
- [10] B. R. Baechler, C. D. Stienbarger, D. A. Horn, J. Joseph, A. R. Taylor, E. F. Granek, S. M. Brander, Microplastic occurrence and effects in commercially harvested North American finfish and shellfish: Current knowledge and future directions, Limnol. Oceanogr. Lett. 5

(2020) 113-136.

- [11] Y. Lu, Y. Zhang, Y. Deng, W. Jiang, Y. Zhao, J. Geng, L. Ding, H. Ren, Uptake and Accumulation of Polystyrene Microplastics in Zebrafish (Danio rerio) and Toxic Effects in Liver, Environ. Sci. Technol. 50 (2016) 4054–4060.
- [12] Q. Chen, M. Gundlach, S. Yang, J. Jiang, M. Velki, D. Yin, H. Hollert, Quantitative investigation of the mechanisms of microplastics and nanoplastics toward zebrafish larvae locomotor activity, Sci. Total Environ. 584–585 (2017) 1022–1031.
- [13] L. Lei, S. Wu, S. Lu, M. Liu, Y. Song, Z. Fu, H. Shi, K. M. Raley-Susman, D. He, Microplastic particles cause intestinal damage and other adverse effects in zebrafish Danio rerio and nematode Caenorhabditis elegans, Sci. Total Environ. 619–620 (2018) 1–8.
- [14] H. A. Leslie, M. J. M. van Velzen, S. H. Brandsma, A. D. Vethaak, J. J. Garcia-Vallejo, M. H. Lamoree, Discovery and quantification of plastic particle pollution in human blood, Environ. Int. 163 (2022) 107199.
- [15] L. C. Jenner, J. M. Rotchell, R. T. Bennett, M. Cowen, V. Tentzeris, L. R. Sadofsky, Detection of microplastics in human lung tissue using µFTIR spectroscopy, Sci. Total Environ. 831 (2022) 154907.
- [16] A. Ragusa, A. Svelato, C. Santacroce, P. Catalano, V. Notarstefano, O. Carnevali, F. Papa, M. C. A. Rongioletti, F. Baiocco, S. Draghi, E. D'Amore, D. Rinaldo, M. Matta, E. Giorgini, Plasticenta: First evidence of microplastics in human placenta, Environ. Int. 146 (2021) 106274.
- [17] X. Guo, J. Wang, The chemical behaviors of microplastics in marine environment: A review, Mar. Pollut. Bull. 142 (2019) 1–14.
- [18] C. Campanale, C. Massarelli, I. Savino, V. Locaputo, V. F. Uricchio, A detailed review study on potential effects of microplastics and additives of concern on human health, Int. J. Environ. Res. Public Health 17 (2020) 1212.
- [19] A. E. Rubin, I. Zucker, Interactions of microplastics and organic compounds in aquatic environments: A case study of augmented joint toxicity, Chemosphere 289 (2022) 133212.
- [20] S. Liu, J. Wang, J. Zhu, J. Wang, H. Wang, X. Zhan, The joint toxicity of polyethylene microplastic and phenanthrene to wheat seedlings, Chemosphere 282 (2021) 130967.
- [21] R. Zhou, G. Lu, Z. Yan, R. Jiang, X. Bao, P. Lu, A review of the influences of microplastics on toxicity and transgenerational effects of pharmaceutical and personal care products in

aquatic environment, Sci. Total Environ. 732 (2020) 139222.

- [22] O. D. Agboola, N. U. Benson, Physisorption and Chemisorption Mechanisms Influencing Micro (Nano) Plastics-Organic Chemical Contaminants Interactions: A Review, Front. Environ. Sci. 9 (2021) 1–27.
- [23] URL: https://plasticseurope.org/knowledge-hub/plastics-the-facts-2021.
- [24] J. Gigault, A. ter Halle, M. Baudrimont, P. Y. Pascal, F. Gauffre, T. L. Phi, H. El Hadri, B. Grassl, S. Reynaud, Current opinion: What is a nanoplastic?, Environ. Pollut. 235 (2018) 1030–1034.
- [25] URL: https://echa.europa.eu/hr/hot-topics/microplastics.
- [26] A. L. Andrady, The plastic in microplastics: A review, Mar. Pollut. Bull. 119 (2017) 12–22.
- [27] I. E. Napper, B. F. R. Davies, H. Clifford, S. Elvin, H. J. Koldewey, P. A. Mayewski, K. R. Miner, M. Potocki, A. C. Elmore, A. P. Gajurel, R. C. Thompson, Reaching New Heights in Plastic Pollution—Preliminary Findings of Microplastics on Mount Everest, One Earth 3 (2020) 621–630.
- [28] S. Allen, D. Allen, F. Baladima, V. R. Phoenix, J. L. Thomas, G. Le Roux, J. E. Sonke, Evidence of free tropospheric and long-range transport of microplastic at Pic du Midi Observatory, Nat. Commun. 12 (2021) 7242.
- [29] X. Peng, M. Chen, S. Chen, S. Dasgupta, H. Xu, K. Ta, M. Du, J. Li, Z. Guo, S. Bai, Microplastics contaminate the deepest part of the world's ocean, Geochemical Perspect. Lett. 9 (2018) 1–5.
- [30] C. Wang, J. Zhao, B. Xing, Environmental source, fate, and toxicity of microplastics, J. Hazard. Mater. 407 (2021) 124357.
- [31] T. Eberhard, G. Casillas, G. M. Zarus, D. B. Barr, Systematic review of microplastics and nanoplastics in indoor and outdoor air: identifying a framework and data needs for quantifying human inhalation exposures, J. Expo. Sci. Environ. Epidemiol. 34 (2024) 185– 196.
- [32] M. Sajjad, Q. Huang, S. Khan, M. A. Khan, Y. Liu, J. Wang, F. Lian, Q. Wang, G. Guo, Microplastics in the soil environment: A critical review, Environ. Technol. Innov. 27 (2022) 102408.
- [33] S. Upadhyay, P. K. Sharma, K. Dogra, P. Bhattacharya, M. Kumar, V. Tripathi, R. Karmakar, Microplastics in freshwater: Unveiling sources, fate, and removal strategies,

Groundw. Sustain. Dev. 26 (2024) 101185.

- [34] A. Cera, G. Cesarini, M. Scalici, Microplastics in freshwater: What is the news from the world?, Diversity 12 (2020) 276.
- [35] I. Gambino, F. Bagordo, T. Grassi, A. Panico, A. De Donno, Occurrence of Microplastics in Tap and Bottled Water: Current Knowledge, Int. J. Environ. Res. Public Heal. Rev. 19 (2022) 5283.
- [36] T. Jenkins, B. D. Persaud, W. Cowger, K. Szigeti, D. G. Roche, E. Clary, S. Slowinski, B. Lei, A. Abeynayaka, E. S. Nyadjro, T. Maes, L. Thornton Hampton, M. Bergmann, J. Aherne, S. A. Mason, J. F. Honek, F. Rezanezhad, A. L. Lusher, A. M. Booth, R. D. L. Smith, P. Van Cappellen, Current State of Microplastic Pollution Research Data: Trends in Availability and Sources of Open Data, Front. Environ. Sci. 10 (2022) 1–10.
- [37] G. B. Merrill, L. Hermabessiere, C. M. Rochman, D. P. Nowacek, Microplastics in marine mammal blubber, melon, & other tissues: Evidence of translocation, Environ. Pollut. 335 (2023) 122252.
- [38] L. G. A. Barboza, A. Dick Vethaak, B. R. B. O. Lavorante, A. K. Lundebye, L. Guilhermino, Marine microplastic debris: An emerging issue for food security, food safety and human health, Mar. Pollut. Bull. 133 (2018) 336–348.
- [39] M. Lehtiniemi, S. Hartikainen, P. Näkki, J. Engström-Öst, A. Koistinen, O. Setälä, Size matters more than shape: Ingestion of primary and secondary microplastics by small predators, Food Webs 17 (2018) e00097.
- [40] R. J. E. Vroom, A. A. Koelmans, E. Besseling, C. Halsband, Aging of microplastics promotes their ingestion by marine zooplankton, Environ. Pollut. 231 (2017) 987–996.
- [41] H. S. Charlton-Howard, A. L. Bond, J. Rivers-Auty, J. L. Lavers, 'Plasticosis': Characterising macro- and microplastic-associated fibrosis in seabird tissues, J. Hazard. Mater. 450 (2023) 131090.
- [42] X. Zheng, X. Wu, Q. Zheng, B. X. Mai, R. Qiu, Transfer of Microplastics in Terrestrial and Aquatic Food Webs: The Impact of E-Waste Debris and Ecological Traits, Environ. Sci. Technol. 57 (2023) 1300–1308.
- [43] I. G. Hallanger, A. Ask, E. Fuglei, Occurrence of ingested human litter in winter arctic foxes (Vulpes lagopus) from Svalbard, Norway, Environ. Pollut. 303 (2022) 1–5.
- [44] F. Bahrani, A. Mohammadi, S. Dobaradaran, G. E. De-la-Torre, H. Arfaeinia, B.

Ramavandi, R. Saeedi, A. Tekle-Röttering, Occurrence of microplastics in edible tissues of livestock (cow and sheep), Environ. Sci. Pollut. Res. 31 (2024) 22145–22157.

- [45] H. Li, Z. Yang, F. Jiang, L. Li, Y. Li, M. Zhang, Z. Qi, R. Ma, Y. Zhang, J. Fang, X. Chen, Y. Geng, Z. Cao, G. Pan, L. Yan, W. Sun, Detection of microplastics in domestic and fetal pigs' lung tissue in natural environment: A preliminary study, Environ. Res. 216 (2023) 114623.
- [46] E. Jeong, J. Y. Lee, M. Redwan, Animal exposure to microplastics and health effects: A review, Emerg. Contam. 10 (2024) 100369.
- [47] W. Gong, W. Zhang, M. Jiang, S. Li, G. Liang, Q. Bu, L. Xu, H. Zhu, A. Lu, Speciesdependent response of food crops to polystyrene nanoplastics and microplastics, Sci. Total Environ. 796 (2021) 148750.
- [48] N. N. Lazăr, M. Călmuc, Ștefania A. Milea, P. L. Georgescu, C. Iticescu, Micro and nano plastics in fruits and vegetables: A review, Heliyon 10 (2024) 28291.
- [49] L. E. Tympa, K. Katsara, P. N. Moschou, G. Kenanakis, V. M. Papadakis, Do microplastics enter our food chain via root vegetables? A raman based spectroscopic study on raphanus sativus, Materials (Basel). 14 (2021) 1–11.
- [50] L. Zhu, J. Zhu, R. Zuo, Q. Xu, Y. Qian, L. AN, Identification of microplastics in human placenta using laser direct infrared spectroscopy, Sci. Total Environ. 856 (2023) 159060.
- [51] Y. Dong, M. Gao, W. Qiu, Z. Song, Uptake of microplastics by carrots in presence of As (III): Combined toxic effects, J. Hazard. Mater. 411 (2021) 125055.
- [52] L. Giorgetti, C. Spanò, S. Muccifora, S. Bottega, F. Barbieri, L. Bellani, M. Ruffini Castiglione, Exploring the interaction between polystyrene nanoplastics and Allium cepa during germination: Internalization in root cells, induction of toxicity and oxidative stress, Plant Physiol. Biochem. 149 (2020) 170–177.
- [53] L. F. Amato-Lourenço, R. Carvalho-Oliveira, G. R. Júnior, L. dos Santos Galvão, R. A. Ando, T. Mauad, Presence of airborne microplastics in human lung tissue, J. Hazard. Mater. 416 (2021) 126124.
- [54] J. Zhang, L. Wang, L. Trasande, K. Kannan, Occurrence of Polyethylene Terephthalate and Polycarbonate Microplastics in Infant and Adult Feces, Environ. Sci. Technol. Lett. 8 (2021) 989–994.
- [55] P. Schwabl, S. Koppel, P. Konigshofer, T. Bucsics, M. Trauner, T. Reiberger, B. Liebmann,

Detection of various microplastics in human stool: A prospective case series, Ann. Intern. Med. 171 (2019) 453–457.

- [56] G. Kutralam-Muniasamy, V. C. Shruti, F. Pérez-Guevara, P. D. Roy, Microplastic diagnostics in humans: "The 3Ps" Progress, problems, and prospects, Sci. Total Environ. 856 (2023) 159164.
- [57] F. Amereh, N. Amjadi, A. Mohseni-Bandpei, S. Isazadeh, Y. Mehrabi, A. Eslami, Z. Naeiji,
 M. Rafiee, Placental plastics in young women from general population correlate with
 reduced foetal growth in IUGR pregnancies, Environ. Pollut. 314 (2022) 120174.
- [58] A. Ragusa, V. Notarstefano, A. Svelato, A. Belloni, G. Gioacchini, C. Blondeel, E. Zucchelli, C. De Luca, S. D'avino, A. Gulotta, O. Carnevali, E. Giorgini, Raman Microspectroscopy Detection and Characterisation of Microplastics in Human Breastmilk, Polymers (Basel). 14 (2022) 1–14.
- [59] L. Rubio, I. Barguilla, J. Domenech, R. Marcos, A. Hernández, Biological effects, including oxidative stress and genotoxic damage, of polystyrene nanoparticles in different human hematopoietic cell lines, J. Hazard. Mater. 398 (2020) 122900.
- [60] A. Banerjee, W. L. Shelver, Micro- and nanoplastic induced cellular toxicity in mammals: A review, Sci. Total Environ. 755 (2021) 142518.
- [61] B. Zhao, P. Rehati, Z. Yang, Z. Cai, C. Guo, Y. Li, The potential toxicity of microplastics on human health, Sci. Total Environ. 912 (2024) 168946.
- [62] N. Rafa, B. Ahmed, F. Zohora, J. Bakya, S. Ahmed, S. F. Ahmed, M. Mofijur, A. A. Chowdhury, F. Almomani, Microplastics as carriers of toxic pollutants: Source, transport, and toxicological effects, Environ. Pollut. 343 (2024) 123190.
- [63] B. Cunningham, B. Harper, S. Brander, S. Harper, Toxicity of micro and nano tire particles and leachate for model freshwater organisms, J. Hazard. Mater. 429 (2022) 128319.
- [64] M. Abboudi, A. Odeh, K. Aljoumaa, Carbonyl compound leaching from polyethylene terephthalate into bottled water under sunlight exposure, Toxicol. Environ. Chem. 98 (2016) 167–178.
- [65] V. Godoy, G. Blázquez, M. Calero, L. Quesada, M. A. Martín-Lara, The potential of microplastics as carriers of metals, Environ. Pollut. 255 (2019) 113363.
- [66] F. Wang, C. S. Wong, D. Chen, X. Lu, F. Wang, E. Y. Zeng, Interaction of toxic chemicals with microplastics: A critical review, Water Res. 139 (2018) 208–219.

- [67] S. H. Joo, Y. Liang, M. Kim, J. Byun, H. Choi, Microplastics with adsorbed contaminants: Mechanisms and Treatment, Environ. Challenges 3 (2021) 100042.
- [68] L. Zhao, L. Rong, J. Xu, J. Lian, L. Wang, H. Sun, Sorption of five organic compounds by polar and nonpolar microplastics, Chemosphere 257 (2020) 127206.
- [69] R. M. Razanajatovo, J. Ding, S. Zhang, H. Jiang, H. Zou, Sorption and desorption of selected pharmaceuticals by polyethylene microplastics, Mar. Pollut. Bull. 136 (2018) 516–523.
- [70] J. Li, K. Zhang, H. Zhang, Adsorption of antibiotics on microplastics, Environ. Pollut. 237 (2018) 460–467.
- [71] X. Guo, C. Chen, J. Wang, Sorption of sulfamethoxazole onto six types of microplastics, Chemosphere 228 (2019) 300–308.
- [72] X. Guo, Y. Liu, J. Wang, Sorption of sulfamethazine onto different types of microplastics: A combined experimental and molecular dynamics simulation study, Mar. Pollut. Bull. 145 (2019) 547–554.
- [73] B. Xu, F. Liu, P. C. Brookes, J. Xu, The sorption kinetics and isotherms of sulfamethoxazole with polyethylene microplastics, Mar. Pollut. Bull. 131 (2018) 191–196.
- [74] Y. Dong, M. Gao, Z. Song, W. Qiu, As(III) adsorption onto different-sized polystyrene microplastic particles and its mechanism, Chemosphere 239 (2020) 124792.
- [75] Y. Wang, X. Wang, Y. Li, J. Li, Y. Liu, S. Xia, J. Zhao, Effects of exposure of polyethylene microplastics to air, water and soil on their adsorption behaviors for copper and tetracycline, Chem. Eng. J. 404 (2021) 126412.
- [76] J. Ma, J. Zhao, Z. Zhu, L. Li, F. Yu, Effect of microplastic size on the adsorption behavior and mechanism of triclosan on polyvinyl chloride, Environ. Pollut. 254 (2019) 113104.
- [77] Z. Liu, Q. Qin, Z. Hu, L. Yan, U. I. Ieong, Y. Xu, Adsorption of chlorophenols on polyethylene terephthalate microplastics from aqueous environments: Kinetics, mechanisms and influencing factors, Environ. Pollut. 265 (2020) 114926.
- [78] F. Liu, G. Liu, Z. Zhu, S. Wang, F. Zhao, Interactions between microplastics and phthalate esters as affected by microplastics characteristics and solution chemistry, Chemosphere 214 (2019) 688–694.
- [79] T. Hüffer, T. Hofmann, Sorption of non-polar organic compounds by micro-sized plastic particles in aqueous solution, Environ. Pollut. 214 (2016) 194–201.
- [80] E. Costigan, A. Collins, M. D. Hatinoglu, K. Bhagat, J. MacRae, F. Perreault, O. Apul,

Adsorption of organic pollutants by microplastics: Overview of a dissonant literature, J. Hazard. Mater. Adv. 6 (2022) 100091.

- [81] F. Yu, C. Yang, Z. Zhu, X. Bai, J. Ma, Adsorption behavior of organic pollutants and metals on micro/nanoplastics in the aquatic environment, Sci. Total Environ. 694 (2019) 133643.
- [82] M. Li, H. Yu, Y. Wang, J. Li, G. Ma, X. Wei, QSPR models for predicting the adsorption capacity for microplastics of polyethylene, polypropylene and polystyrene, Sci. Rep. 10 (2020) 1–11.
- [83] R. C. Thompson, S. H. Swan, C. J. Moore, F. S. Vom Saal, Our plastic age, Philos. Trans.R. Soc. B Biol. Sci. 364 (2009) 1973–1976.
- [84] G. Strobl, The Physics of Polymers: Concepts for Understanding Their Structures and BehaviorSpringer Berlin Heidelberg, New York, 1997.
- [85] W. Wang, J. Wang, Comparative evaluation of sorption kinetics and isotherms of pyrene onto microplastics, Chemosphere 193 (2018) 567–573.
- [86] X. Guo, X. Wang, X. Zhou, X. Kong, S. Tao, B. Xing, Sorption of four hydrophobic organic compounds by three chemically distinct polymers: Role of chemical and physical composition, Environ. Sci. Technol. 46 (2012) 7252–7259.
- [87] X. Liu, J. Xu, Y. Zhao, H. Shi, C. H. Huang, Hydrophobic sorption behaviors of 17B-Estradiol on environmental microplastics, Chemosphere 226 (2019) 726–735.
- [88] J. P. Greene, Microstructures of Polymers, in Plast. Des. Libr. (Ed.: J.P.B.T.-A.P. and C. Greene), William Andrew Publishing, 2021, pp. 27–37.
- [89] H. Luo, C. Liu, D. He, J. Xu, J. Sun, J. Li, X. Pan, Environmental behaviors of microplastics in aquatic systems: A systematic review on degradation, adsorption, toxicity and biofilm under aging conditions, J. Hazard. Mater. 423 (2022) 126915.
- [90] J. F. Rabek, Polymer Photodegradation: Mechanisms and Experimental Methods, in Polym. Photodegradation, Chapman & Hall, London, 1995, pp. 67–103.
- [91] T. Sang, C. J. Wallis, G. Hill, G. J. P. Britovsek, Polyethylene terephthalate degradation under natural and accelerated weathering conditions, Eur. Polym. J. 136 (2020) 109873.
- [92] A. K. Rodriguez, B. Mansoor, G. Ayoub, X. Colin, A. A. Benzerga, Effect of UV-aging on the mechanical and fracture behavior of low density polyethylene, Polym. Degrad. Stab. 180 (2020) 109185.
- [93] I. Grigoriadou, K. M. Paraskevopoulos, K. Chrissafis, E. Pavlidou, T. G. Stamkopoulos, D.

Bikiaris, Effect of different nanoparticles on HDPE UV stability, Polym. Degrad. Stab. 96 (2011) 151–163.

- [94] J. W. Scott, A. Turner, A. F. Prada, L. Zhao, Heterogeneous weathering of polypropylene in the marine environment, Sci. Total Environ. 812 (2021) 152308.
- [95] H. Zhang, J. Wang, B. Zhou, Y. Zhou, Z. Dai, Q. Zhou, P. Chriestie, Y. Luo, Enhanced adsorption of oxytetracycline to weathered microplastic polystyrene: Kinetics, isotherms and influencing factors, Environ. Pollut. 243 (2018) 1550–1557.
- [96] X. Fan, Y. Zou, N. Geng, J. Liu, J. Hou, D. Li, C. Yang, Y. Li, Investigation on the adsorption and desorption behaviors of antibiotics by degradable MPs with or without UV ageing process, J. Hazard. Mater. 401 (2021) 123363.
- [97] N. M. Ainali, D. N. Bikiaris, D. A. Lambropoulou, Aging effects on low- and high-density polyethylene, polypropylene and polystyrene under UV irradiation: An insight into decomposition mechanism by Py-GC/MS for microplastic analysis, J. Anal. Appl. Pyrolysis 158 (2021) 105207.
- [98] X. Wu, P. Liu, H. Huang, S. Gao, Adsorption of triclosan onto different aged polypropylene microplastics: Critical effect of cations, Sci. Total Environ. 717 (2020) 137033.
- [99] G. Liu, Z. Zhu, Y. Yang, Y. Sun, F. Yu, J. Ma, Sorption behavior and mechanism of hydrophilic organic chemicals to virgin and aged microplastics in freshwater and seawater, Environ. Pollut. 246 (2019) 26–33.
- [100] F. Kong, X. Xu, Y. Xue, Y. Gao, L. Zhang, L. Wang, S. Jiang, Q. Zhang, Investigation of the Adsorption of Sulfamethoxazole by Degradable Microplastics Artificially Aged by Chemical Oxidation, Arch. Environ. Contam. Toxicol. 81 (2021) 155–165.
- [101] X. Fan, R. Gan, J. Liu, Y. Xie, D. Xu, Y. Xiang, J. Su, Z. Teng, J. Hou, Adsorption and desorption behaviors of antibiotics by tire wear particles and polyethylene microplastics with or without aging processes, Sci. Total Environ. 771 (2021) 145451.
- [102] J. Yao, J. Wen, H. Li, Y. Yang, Surface functional groups determine adsorption of pharmaceuticals and personal care products on polypropylene microplastics, J. Hazard. Mater. 423 (2022) 127131.
- [103] C. Y. Shih, Y. H. Wang, Y. J. Chen, H. A. Chen, A. Y. C. Lin, Enhanced sorption of the UV filter 4-methylbenzylidene camphor on aged PET microplastics from both experimental and theoretical perspectives, RSC Adv. 11 (2021) 32494–32504.

- [104] X. Guo, J. Wang, Sorption of antibiotics onto aged microplastics in freshwater and seawater, Mar. Pollut. Bull. 149 (2019) 110511.
- [105] M. Llorca, G. Schirinzi, M. Martínez, D. Barceló, M. Farré, Adsorption of perfluoroalkyl substances on microplastics under environmental conditions, Environ. Pollut. 235 (2018) 680–691.
- [106] Q. Mo, X. Yang, J. Wang, H. Xu, W. Li, Q. Fan, S. Gao, W. Yang, C. Gao, D. Liao, Y. Li,
 Y. Zhang, Adsorption mechanism of two pesticides on polyethylene and polypropylene microplastics: DFT calculations and particle size effects, Environ. Pollut. 291 (2021) 118120.
- [107] S. Seidensticker, P. Grathwohl, J. Lamprecht, C. Zarfl, A combined experimental and modeling study to evaluate pH-dependent sorption of polar and non-polar compounds to polyethylene and polystyrene microplastics, Environ. Sci. Eur. 30 (2018) 1–12.
- [108] X. Xin, B. Chen, M. Yang, S. Gao, H. Wang, W. Gu, X. Li, B. Zhang, A critical review on the interaction of polymer particles and co-existing contaminants: Adsorption mechanism, exposure factors, effects on plankton species, J. Hazard. Mater. 445 (2023) 130463.
- [109] E. C. Lima, F. Sher, A. Guleria, M. R. Saeb, I. Anastopoulos, H. N. Tran, A. Hosseini-Bandegharaei, Is one performing the treatment data of adsorption kinetics correctly?, J. Environ. Chem. Eng. 9 (2021) 104813.
- [110] J. P. Simonin, On the comparison of pseudo-first order and pseudo-second order rate laws in the modeling of adsorption kinetics, Chem. Eng. J. 300 (2016) 254–263.
- [111] A. Kayan, Inorganic-organic hybrid materials and their adsorbent properties, Adv. Compos. Hybrid Mater. 2 (2019) 34–45.
- [112] X. Sheng, J. Wang, W. Zhang, Q. Zuo, The potential for PE microplastics to affect the removal of carbamazepine medical pollutants from aqueous environments by multiwalled carbon nanotubes, Toxics 9 (2021) 139.
- [113] B. Hu, Y. Li, L. Jiang, X. Chen, L. Wang, S. An, F. Zhang, Influence of microplastics occurrence on the adsorption of 17β-estradiol in soil, J. Hazard. Mater. 400 (2020) 123325.
- [114] Y. Wang, X. Wang, Y. Li, J. Li, F. Wang, S. Xia, J. Zhao, Biofilm alters tetracycline and copper adsorption behaviors onto polyethylene microplastics, Chem. Eng. J. 392 (2020) 123808.
- [115] I. Langmuir, The adsorption of gases on plane surface of glass, mica and platinum, J. Am.

Chem. Soc. 40 (1918) 1361–1403.

- [116] K. Y. Foo, B. H. Hameed, Insights into the modeling of adsorption isotherm systems, Chem.Eng. J. 156 (2010) 2–10.
- [117] H. Freundlich, Über die Adsorption in Lösungen, Zeitschrift für Phys. Chemie 57 (1907) 385–470.
- [118] J. Wang, X. Guo, Adsorption isotherm models: Classification, physical meaning, application and solving method, Chemosphere 258 (2020) 127279.
- [119] A. Khan, J. E. Szulejko, P. Samaddar, K.-H. Kim, W. Eom, S. B. Ambade, T. H. Han, The effect of diverse metal oxides in graphene composites on the adsorption isotherm of gaseous benzene, Environ. Res. 172 (2019) 367–374.
- [120] I. M. Temkin, Adsorption equilibrium and the kinetics of processes on nonhomogeneous surfaces and in the interaction between adsorbed molecules, Zh. Fiz. Chim. 15 (1941) 296– 332.
- [121] J. Kapelewska, U. Klekotka, E. Żadziłko, J. Karpińska, Simultaneous sorption behaviors of UV filters on the virgin and aged micro-high-density polyethylene under environmental conditions, Sci. Total Environ. 789 (2021) 147979.
- [122] M. M. Dubinin, V. A. Astakhov, Development of concepts of the volume filling of micropores in the adsorption of gases and vapors by microporous adsorbents -Communication 4. Differential heats and entropies of adsorption, Bull. Acad. Sci. USSR Div. Chem. Sci. 20 (1971) 3–7.
- [123] Q. Hu, Z. Zhang, Application of Dubinin–Radushkevich isotherm model at the solid/solution interface: A theoretical analysis, J. Mol. Liq. 277 (2019) 646–648.
- [124] V. J. Inglezakis, Solubility-normalized Dubinin-Astakhov adsorption isotherm for ionexchange systems, Microporous Mesoporous Mater. 103 (2007) 72–81.
- [125] J. Li, X. Huang, Z. Hou, T. Ding, Sorption of diclofenac by polystyrene microplastics: Kinetics, isotherms and particle size effects, Chemosphere 290 (2022) 133311.
- [126] D. Cortés-Arriagada, S. Miranda-Rojas, M. Belén, D. E. Ortega, V. B. Alarcón-Palacio, The interaction mechanism of polystyrene microplastics with pharmaceuticals and personal care products, Sci. Total Environ. 861 (2022) 160632.
- [127] R. C. Lajmanovich, A. M. Attademo, G. Lener, A. P. Cuzziol Boccioni, P. M. Peltzer, C. S. Martinuzzi, L. D. Demonte, M. R. Repetti, Glyphosate and glufosinate ammonium,

herbicides commonly used on genetically modified crops, and their interaction with microplastics: Ecotoxicity in anuran tadpoles, Sci. Total Environ. 804 (2022) 150177.

- [128] S. Tang, S. Ma, T. Zhang, X. Liu, N. Pervez, C. Cao, Y. Zhao, Sorption of bisphenol A onto microplastics and associated environmental risks in comparison to engineered carbonous materials and natural media, Gondwana Res. 117 (2023) 295–306.
- [129] T. Wan, S. Lu, W. Cheng, J. Ren, M. Wang, B. Hu, Z. Jia, Y. Li, Y. Sun, A spectroscopic and theoretical investigation of interaction mechanisms of tetracycline and polystyrene nanospheres under different conditions, Environ. Pollut. 249 (2019) 398–405.
- [130] J. Wu, Q. Ye, P. Wu, S. Xu, Y. Liu, Z. Ahmed, S. Rehman, N. Zhu, Heteroaggregation of nanoplastics with oppositely charged minerals in aquatic environment: Experimental and theoretical calculation study, Chem. Eng. J. 428 (2022) 131191.
- [131] H. Yu, B. Yang, M. G. Waigi, F. Peng, Z. Li, X. Hu, The effects of functional groups on the sorption of naphthalene on microplastics, Chemosphere 261 (2020) 127592.
- [132] D. Cortés-Arriagada, Elucidating the co-transport of bisphenol A with polyethylene terephthalate (PET) nanoplastics: A theoretical study of the adsorption mechanism, Environ. Pollut. 270 (2021) 116192.
- [133] D. E. Ortega, D. Cortés-Arriagada, Atmospheric microplastics and nanoplastics as vectors of primary air pollutants - A theoretical study on the polyethylene terephthalate (PET) case, Environ. Pollut. 318 (2023) 120860.
- [134] D. Cortés-Arriagada, D. E. Ortega, Interaction mechanism of triclosan on pristine microplastics, Sci. Total Environ. 891 (2023) 164470.
- [135] D. Cortés-Arriagada, D. E. Ortega, S. Miranda-Rojas, Mechanistic insights into the adsorption of endocrine disruptors onto polystyrene microplastics in water, Environ. Pollut. 319 (2023) 121017.
- [136] H. Li, F. Wang, J. Li, S. Deng, S. Zhang, Adsorption of three pesticides on polyethylene microplastics in aqueous solutions: Kinetics, isotherms, thermodynamics, and molecular dynamics simulation, Chemosphere 264 (2021) 128556.
- [137] Y. Leng, W. Wang, H. Cai, F. Chang, W. Xiong, J. Wang, Sorption kinetics, isotherms and molecular dynamics simulation of 17β-estradiol onto microplastics, Sci. Total Environ. 858 (2023) 159803.
- [138] M. A. Dias, P. R. Batista, L. C. Ducati, C. C. Montagner, Insights into sorption and

molecular transport of atrazine, testosterone, and progesterone onto polyamide microplastics in different aquatic matrices, Chemosphere 318 (2023) 137949.

- [139] Y. Chen, J. Li, F. Wang, H. Yang, L. Liu, Adsorption of tetracyclines onto polyethylene microplastics: A combined study of experiment and molecular dynamics simulation, Chemosphere 265 (2021) 129133.
- [140] P. A. Townsend, Adsorption in Action: Molecular Dynamics as a Tool to Study Adsorption at the Surface of Fine Plastic Particles in Aquatic Environments, ACS Omega 9 (2024) 5142–5156.
- [141] H. Feng, Y. Liu, Y. Xu, S. Li, X. Liu, Y. Dai, J. Zhao, T. Yue, Benzo[a]pyrene and heavy metal ion adsorption on nanoplastics regulated by humic acid: Cooperation/competition mechanisms revealed by molecular dynamics simulations, J. Hazard. Mater. 424 (2022) 127431.
- [142] W. Liu, H. Tang, B. Yang, C. Li, Y. Chen, T. Huang, Molecular level insight into the different interaction intensity between microplastics and aromatic hydrocarbon in pure water and seawater, Sci. Total Environ. 862 (2023) 160786.
- [143] D. Hatinoglu, A. Adan, F. Perreault, I. Imamoglu, O. G. Apul, Linear solvation energy relationships for adsorption of aromatic organic compounds by microplastics, Chem. Eng. Sci. 282 (2023) 119233.
- [144] M. D. Hatinoglu, F. Perreault, O. G. Apul, Modified linear solvation energy relationships for adsorption of perfluorocarboxylic acids by polystyrene microplastics, Sci. Total Environ. 860 (2023) 160524.
- [145] C. W. Cho, Y. Zhao, Y. S. Yun, QSAR modelling for predicting adsorption of neutral, cationic, and anionic pharmaceuticals and other neutral compounds to microalgae Chlorella vulgaris in aquatic environment, Water Res. 151 (2019) 288–295.
- [146] B. Gui, X. Xu, S. Zhang, Y. Wang, C. Li, D. Zhang, L. Su, Y. Zhao, Prediction of organic compounds adsorbed by polyethylene and chlorinated polyethylene microplastics in freshwater using QSAR, Environ. Res. 197 (2021) 111001.
- [147] Y. Zhang, H. Mao, Q. Ma, Z. Chen, H. Wang, A. Xu, Y. Zhang, A QSAR prediction model for adsorption of organic contaminants on microplastics: Dubinin-Astakhov plus linear solvation energy relationships, Sci. Total Environ. 930 (2024) 172801.
- [148] A. U. Khan, G. M. Porta, M. Riva, A. Guadagnini, In-silico mechanistic analysis of

adsorption of Iodinated Contrast Media agents on graphene surface, Ecotoxicol. Environ. Saf. 280 (2024) 116506.

- [149] Y. Wang, J. Chen, X. Wei, A. J. Hernandez Maldonado, Z. Chen, Unveiling Adsorption Mechanisms of Organic Pollutants onto Carbon Nanomaterials by Density Functional Theory Computations and Linear Free Energy Relationship Modeling, Environ. Sci. Technol. 51 (2017) 11820–11828.
- [150] D. J. Griffiths, D. F. Schroeter, The wave function, in Introd. to Quantum Mech., Cambridge University Press, Cambridge, 2018, pp. 15–41.
- [151] D. J. Griffiths, D. F. Schroeter, Time-independent Schrödinger Equation, in Introd. to Quantum Mech., Cambridge University Press, Cambridge, 2018, pp. 42–117.
- [152] C. J. Cramer, Foundations of Molecular Orbital Theory, in Essentials Comput. Chem. Theor. Model., John Wiley & Sons, Ltd, New York, 2004, pp. 105–130.
- [153] E. G. Lewars, The Concept of the Potential Energy Surface, in Comput. Chem. Introd. to Theory Appl. Mol. Quantum Mech., Springer Netherlands, Dordrecht, 2011, pp. 9–43.
- [154] C. J. Cramer, Ab Initio Implementations of Hartree–Fock Molecular Orbital Theory, in Essentials Comput. Chem. Theor. Model., John Wiley & Sons, Ltd, New York, 2004, pp. 165–202.
- [155] I. S. Ulusoy, A. K. Wilson, Slater and Gaussian basis functions and computation of molecular integrals, in Dev. Phys. Theor. Chem. (Eds.: S.M. Blinder, J.E.B.T.-M.P. in T.C. House), Elsevier, 2019, pp. 31–61.
- [156] E. G. Lewars, Ab initio Calculations, in Comput. Chem. Introd. to Theory Appl. Mol. Quantum Mech., Springer Netherlands, Dordrecht, 2011, pp. 175–390.
- [157] C. J. Cramer, Including Electron Correlation in Molecular Orbital Theory, in Essentials Comput. Chem. Theor. Model., John Wiley & Sons, Ltd, New York, 2004, pp. 203–248.
- [158] W. Koch, M. C. Holthausen, The Electron Density as the Basic Variable: Early Attempts, in A Chem. Guid. to Density Funct. Theory, John Wiley & Sons, Ltd, 2001, pp. 29–32.
- [159] W. Koch, M. C. Holthausen, The Hohenberg-Kohn Theorems, in A Chem. Guid. to Density Funct. Theory, John Wiley & Sons, Ltd, 2001, pp. 33–40.
- [160] A. E. Mattsson, P. A. Schultz, M. P. Desjarlais, T. E. Matsson, K. Leung, Designing meaningful density functional theory calculations in materials science — a primer, Model. Simul. Mater. Sci. Eng. 13 (2005) R1–R31.

- [161] C. J. Cramer, Density Functional Theory, in Essentials Comput. Chem. Theor. Model., John Wiley & Sons, Ltd, New York, 2004, pp. 249–304.
- [162] A. J. Cohen, P. Mori-Sánchez, W. Yang, Challenges for density functional theory, Chem. Rev. 112 (2012) 289–320.
- [163] J. P. Perdew, K. Burke, M. Ernzerhof, Generalized Gradient Approximation Made Simple, Phys. Rev. Lett. 77 (1996) 3865–3868.
- [164] B. Miehlich, A. Savin, H. Stoll, H. Preuss, Results obtained with the correlation energy density functionals of becke and Lee, Yang and Parr, Chem. Phys. Lett. 157 (1989) 200– 206.
- [165] J. Tao, J. P. Perdew, A. Ruzsinszky, G. E. Scuseria, G. I. Csonka, V. N. Staroverov, Metageneralized gradient approximation: non-empirical construction and performance of a density functional, Philos. Mag. 87 (2007) 1071–1084.
- [166] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields, J. Phys. Chem. 98 (1994) 11623–11627.
- [167] D. Rappoport, N. R. M. Crawford, F. Furche, K. Burke, Approximate Density Functionals: Which Should I Choose?, in Encycl. Inorg. Chem., John Wiley & Sons, Ltd, 2009.
- [168] S. Grimme, Semiempirical hybrid density functional with perturbative second-order correlation, J. Chem. Phys. 124 (2006) 034108.
- [169] S. Grimme, F. Neese, Double-hybrid density functional theory for excited electronic states of molecules, J. Chem. Phys. 127 (2007) 154116.
- [170] L. A. Burns, Á. Vázquez-Mayagoitia, B. G. Sumpter, C. D. Sherrill, Density-functional approaches to noncovalent interactions: A comparison of dispersion corrections (DFT-D), exchange-hole dipole moment (XDM) theory, and specialized functionals, J. Chem. Phys. 134 (2011) 084107.
- [171] W. Reckien, F. Janetzko, M. F. Peintinger, T. Bredow, Implementation of empirical dispersion corrections to density functional theory for periodic systems, J. Comput. Chem. 33 (2012) 2023–2031.
- [172] S. Grimme, Accurate description of van der Waals complexes by density functional theory including empirical corrections, J. Comput. Chem. 25 (2004) 1463–1473.
- [173] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio

parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. 132 (2010) 154104.

- [174] E. Caldeweyher, C. Bannwarth, S. Grimme, Extension of the D3 dispersion coefficient model, J. Chem. Phys. 147 (2017) 034112.
- [175] N. S. Gould, S. Li, H. J. Cho, H. Landfield, S. Caratzoulas, D. Vlachos, P. Bai, B. Xu, Understanding solvent effects on adsorption and protonation in porous catalysts, Nat. Commun. 11 (2020) 1–13.
- [176] Q. Sun, The Hydrophobic Effects: Our Current Understanding, Molecules 27 (2022) 7009.
- [177] T. Giovannini, C. Cappelli, Continuum vs. atomistic approaches to computational spectroscopy of solvated systems, Chem. Commun. 59 (2023) 5644–5660.
- [178] J. Tomasi, B. Mennucci, R. Cammi, Quantum mechanical continuum solvation models, Chem. Rev. 105 (2005) 2999–3093.
- [179] B. Mennucci, Polarizable continuum model, Wiley Interdiscip. Rev. Comput. Mol. Sci. 2 (2012) 386–404.
- [180] G. Scalmani, M. J. Frisch, Continuous surface charge polarizable continuum models of solvation. I. General formalism., J. Chem. Phys. 132 (2010) 114110.
- [181] V. Barone, M. Cossi, Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model, J. Phys. Chem. A 102 (1998) 1995–2001.
- [182] A. V Marenich, C. J. Cramer, 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, J. Phys. Chem. B 113 (2009) 6378–6396.
- [183] M. L. Senent, S. Wilson, Intramolecular basis set superposition errors, Int. J. Quantum Chem. 82 (2001) 282–292.
- [184] C. D. Sherrill, Counterpoise Correction and Basis Set Superposition Error, Self-published (2010) 1–5.
- [185] R. Benda, T. Vezin, B. Lebental, Prediction of the interaction strength of an urea-based probe toward ions in water by means of Density Functional Theory/Polarizable Continuum Model calculations, Int. J. Quantum Chem. 122 (2022) 1–21.
- [186] A. Zawada, R. W. Góra, M. M. Mikołajczyk, W. Bartkowiak, On the Calculations of interaction energies and induced electric properties within the polarizable continuum model, J. Phys. Chem. A 116 (2012) 4409–4416.

- [187] Q. Wang, M. D. Newton, Structure, energetics, and electronic coupling in the (TCNE2) -•
 Encounter complex in solution: A polarizable continuum stud, J. Phys. Chem. B 112 (2008) 568–576.
- [188] M. Sri Harish, P. K. Patra, Temperature and its control in molecular dynamics simulations, Mol. Simul. 47 (2021) 701–729.
- [189] J. Choi, A Review of the Mechanical Design of Materials Based on Molecular Dynamics Simulations, Multiscale Sci. Eng. 5 (2023) 86–103.
- [190] S. A. Hollingsworth, R. O. Dror, Molecular Dynamics Simulation for All, Neuron 99 (2018) 1129–1143.
- [191] C. J. Cramer, Simulations of Molecular Ensembles, in Essentials Comput. Chem. Theor. Model., John Wiley & Sons, Ltd, New York, 2004, pp. 105–130.
- [192] D. Frenkel, B. Smit, Chapter 4 Molecular Dynamics Simulations, in Underst. Mol. Simul. (Second Ed. (Eds.: D. Frenkel, B. Smit), Academic Press, San Diego, 2002, pp. 63–107.
- [193] D. Frenkel, B. Smit, Chapter 2 Statistical Mechanics, in (Eds.: D. Frenkel, B.B.T.-U.M.S. (Second E. Smit), Academic Press, San Diego, 2002, pp. 9–22.
- [194] M. E. Tuckerman, Ab initio molecular dynamics : basic concepts , current, J. Phys. Condens. Matter 14 (2002) R1297–R1355.
- [195] L. Monticelli, D. P. Tieleman, Force fields for classical molecular dynamics., Methods Mol. Biol. 924 (2013) 197–213.
- [196] O. Guvench, A. D. MacKerell, Comparison of Protein Force Fields for Molecular Dynamics Simulations, in Mol. Model. Proteins (Ed.: A. Kukol), Humana Press, Totowa, NJ, 2008, pp. 63–88.
- [197] S. Riniker, Fixed-Charge Atomistic Force Fields for Molecular Dynamics Simulations in the Condensed Phase: An Overview, J. Chem. Inf. Model. 58 (2018) 565–578.
- [198] H. J. C. Berendsen, J. R. Grigera, T. P. Straatsma, The missing term in effective pair potentials, J. Phys. Chem. 91 (1987) 6269–6271.
- [199] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926–935.
- [200] P. Mark, L. Nilsson, Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K, J. Phys. Chem. A 105 (2001) 9954–9960.
- [201] M. W. Mahoney, W. L. Jorgensen, A five-site model for liquid water and the reproduction

of the density anomaly by rigid, nonpolarizable potential functions, J. Chem. Phys. 112 (2000) 8910–8922.

- [202] P. Florová, P. Sklenovský, P. Banáš, M. Otyepka, Explicit water models affect the specific solvation and dynamics of unfolded peptides while the conformational behavior and flexibility of folded peptides remain intact, J. Chem. Theory Comput. 6 (2010) 3569–3579.
- [203] D. R. Nutt, J. C. Smith, Molecular dynamics simulations of proteins: Can the explicit water model be varied?, J. Chem. Theory Comput. 3 (2007) 1550–1560.
- [204] C. J. Cramer, Molecular Mechanics, in Essentials Comput. Chem. Theor. Model., John Wiley & Sons, Ltd, New York, 2004, pp. 17–68.
- [205] D. Frenkel, B. Smit, Chapter 3 Monte Carlo Simulations, in Underst. Mol. Simul. (Second Ed. (Eds.: D. Frenkel, B. Smit), Academic Press, San Diego, 2002, pp. 23–61.
- [206] M. P. Allen, D. J. Tildesley, Some tricks of the trade, in Comput. Simul. Liq., Oxford University Press, Oxford, 2017.
- [207] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, LINCS: A linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
- [208] S. Miyamoto, P. A. Kollman, Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models, J. Comput. Chem. 13 (1992) 952–962.
- [209] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, J. Comput. Phys. 23 (1977) 327–341.
- [210] H. C. Andersen, Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations, J. Comput. Phys. 52 (1983) 24–34.
- [211] D. Frenkel, B. Smit, Chapter 6 Molecular Dynamics in Various Ensembles, in Underst. Mol. Simul. (Eds.: D. Frenkel, B. Smit), Academic Press, San Diego, 2002, pp. 139–163.
- [212] P. H. Hünenberger, Thermostat algorithms for molecular dynamics simulations, Adv. Polym. Sci. 173 (2005) 105–147.
- [213] Q. Ke, X. Gong, S. Liao, C. Duan, L. Li, Effects of thermostats/barostats on physical properties of liquids by molecular dynamics simulations, J. Mol. Liq. 365 (2022) 120116.
- [214] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690.
- [215] E. Braun, S. M. Moosavi, B. Smit, Anomalous Effects of Velocity Rescaling Algorithms:

The Flying Ice Cube Effect Revisited, J. Chem. Theory Comput. 14 (2018) 5262–5272.

- [216] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (2007) 014101.
- [217] H. C. Andersen, Molecular dynamics simulations at constant pressure and/or temperature, J. Chem. Phys. 72 (1980) 2384–2393.
- [218] D. Frenkel, B. Smit, Eds., Appendix A Equations of Motion from the Lagrangian or Hamiltonian, in Underst. Mol. Simul. (Second Ed., Academic Press, San Diego, 2002, pp. 481–494.
- [219] M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, J. Appl. Phys. 52 (1981) 7182–7190.
- [220] Y. C. Martin, Hansch analysis 50 years on, Wiley Interdiscip. Rev. Comput. Mol. Sci. 2 (2012) 435–442.
- [221] H. Kubinyi, QSAR and 3D QSAR in drug design Part 2: applications and problems, Drug Discov. Today 2 (1997) 538–546.
- [222] A. Cherkasov, E. N. Muratov, D. Fourches, A. Varnek, I. I. Baskin, M. Cronin, J. Dearden,
 P. Gramatica, Y. C. Martin, R. Todeschini, V. Consonni, V. E. Kuz'Min, R. Cramer, R.
 Benigni, C. Yang, J. Rathman, L. Terfloth, J. Gasteiger, A. Richard, A. Tropsha, QSAR
 modeling: Where have you been? Where are you going to?, J. Med. Chem. 57 (2014) 4977–5010.
- [223] P. Gramatica, On the Development and Validation of QSAR Models, in Comput. Toxicol.(Eds.: B. Reisfeld, A. N. Mayeno) Vol. 930, Humana Press, Totowa NJ, 2013, pp. 499–526.
- [224] C. Yap, H. Li, Z. Ji, Y. Chen, Regression Methods for Developing QSAR and QSPR Models to Predict Compounds of Specific Pharmacodynamic, Pharmacokinetic and Toxicological Properties, Mini-Reviews Med. Chem. 7 (2007) 1097–1107.
- [225] S. M. Ross, Chapter 9 Regression, in Introd. to Probab. Stat. Eng. Sci. (Ed.: S.M. Ross), Academic Press, Boston, 2014, pp. 357–444.
- [226] P. Gramatica, Principles of QSAR models validation: Internal and external, QSAR Comb. Sci. 26 (2007) 694–701.
- [227] J. Leszczynski, A. Kaczmarek-Kedziera, T. Puzyn, M. G. Papadopoulos, H. Reis, M. K. Shukla, Handbook of computational chemistry, in Handb. Comput. Chem. , 2017, pp. 1– 2381.

- [228] R. Todeschini, V. Consonni, Molecular Descriptors for ChemoinformaticsVol. 2Wiley-VCH, Weinheim, 2010.
- [229] Danishuddin, A. U. Khan, Descriptors and their selection methods in QSAR analysis: paradigm for drug design, Drug Discov. Today 21 (2016) 1291–1302.
- [230] A. Mauri, V. Consonni, R. Todeschini, Molecular Descriptors, in Handb. Comput. Chem.
 (Eds.: J. Leszczynski, A. Kaczmarek-Kedziera, T. Puzyn, M. G. Papadopoulos, H. Reis, M. K. Shukla), Springer International Publishing, Cham, 2017, pp. 2065–2093.
- [231] F. Grisoni, D. Ballabio, R. Todeschini, V. Consonni, Molecular Descriptors for Structure--Activity Applications: A Hands-On Approach, in Comput. Toxicol. Methods Protoc. (Ed.: O. Nicolotti), Springer New York, New York, NY, 2018, pp. 3–53.
- [232] M. Shahlaei, Descriptor Selection Methods in Quantitative Structure–Activity Relationship Studies: A Review Study, Chem. Rev. 113 (2013) 8093–8103.
- [233] N. Sukumar, G. Prabhu, P. Saha, Applications of Genetic Algorithms in QSAR/QSPR Modeling, in Appl. Metaheuristics Process Eng. (Eds.: J. Valadi, P. Siarry) Vol. 9783319065, Springer, Cham, 2014, pp. 1–444.
- [234] S. Katoch, S. S. Chauhan, V. Kumar, A Review on Genetic Algorithm: Past, Present, and FutureVol. 80Multimedia Tools And Applications, 2021.
- [235] T. J. Hou, J. M. Wang, X. J. Xu, Applications of genetic algorithms on the structure-activity correlation study of a group of non-nucleoside HIV-1 inhibitors, Chemom. Intell. Lab. Syst. 45 (1999) 303–310.
- [236] M. J. Frisch, 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, Williams, 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. 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, (2019).

- [237] C. Lefebvre, J. Klein, H. Khartabil, J. C. Boisson, E. Hénon, IGMPlot: A program to identify, characterize, and quantify molecular interactions, J. Comput. Chem. 44 (2023) 1750–1766.
- [238] C. Lefebvre, H. Khartabil, J.-C. Boisson, J. Contreras-García, J.-P. Piquemal, E. Hénon, The Independent Gradient Model: A New Approach for Probing Strong and Weak Interactions in Molecules from Wave Function Calculations., Chemphyschem 19 (2018) 724–735.
- [239] T. Lu, Q. Chen, Independent gradient model based on Hirshfeld partition: A new method for visual study of interactions in chemical systems, J. Comput. Chem. 43 (2022) 539–555.
- [240] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics., J. Mol. Graph. 14 (1996) 27-28,33-38.
- [241] A. Mauri, V. Consonni, M. Pavan, R. Todeschini, DRAGON software: An easy approach to molecular descriptor calculations, Match 56 (2006) 237–248.
- [242] J. Ángyán, J. Dobson, G. Jansen, T. Gould, London Dispersion Forces in Molecules, Solids and Nano-Structures: An Introduction to Physical Models and Computational MethodsThe Royal Society Of Chemistry, 2020.
- [243] S. Grimme, Density functional theory with London dispersion corrections, WIREs Comput. Mol. Sci. 1 (2011) 211–228.
- [244] T. J. Daas, E. Fabiano, F. Della Sala, P. Gori-Giorgi, S. Vuckovic, Noncovalent Interactions from Models for the Møller–Plesset Adiabatic Connection, J. Phys. Chem. Lett. 12 (2021) 4867–4875.
- [245] K. E. Riley, J. A. Platts, J. Řezáč, P. Hobza, J. G. Hill, Assessment of the performance of MP2 and MP2 variants for the treatment of noncovalent interactions, J. Phys. Chem. A 116 (2012) 4159–4169.
- [246] C. Lee, W. Yang, R. G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density., Phys. Rev. B. Condens. Matter 37 (1988) 785– 789.
- [247] A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange, J. Chem. Phys. 98 (1993) 5648–5652.
- [248] S. H. Vosko, L. Wilk, M. Nusair, Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis, Can. J. Phys. 58 (1980) 1200–

1211.

- [249] Y. Zhao, 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 four M06-class functionals and 12 other function, Theor. Chem. Acc. 120 (2008) 215–241.
- [250] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, Auxiliary basis sets to approximate Coulomb potentials, Chem. Phys. Lett. 240 (1995) 283–290.
- [251] K. Eichkorn, F. Weigend, O. Treutler, R. Ahlrichs, Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials, Theor. Chem. Acc. 97 (1997) 119–124.
- [252] Y. Zhao, N. E. Schultz, D. G. Truhlar, Exchange-correlation functional with broad accuracy for metallic and nonmetallic compounds, kinetics, and noncovalent interactions, J. Chem. Phys. 123 (2005) 161103.
- [253] Y. Zhao, N. E. Schultz, D. G. Truhlar, Design of Density Functionals by Combining the Method of Constraint Satisfaction with Parametrization for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions, J. Chem. Theory Comput. 2 (2006) 364–382.
- [254] Y. Zhao, D. G. Truhlar, Comparative DFT study of van der Waals complexes: Rare-gas dimers, alkaline-earth dimers, zinc dimer and zinc-rare-gas dimers, J. Phys. Chem. A 110 (2006) 5121–5129.
- [255] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. 132 (2010) 154104.
- [256] L. Goerigk, H. Kruse, S. Grimme, Benchmarking density functional methods against the S66 and S66x8 datasets for non-covalent interactions, ChemPhysChem 12 (2011) 3421– 3433.
- [257] S. Vuckovic, E. Fabiano, P. Gori-Giorgi, K. Burke, MAP: An MP2 Accuracy Predictor for Weak Interactions from Adiabatic Connection Theory, J. Chem. Theory Comput. 16 (2020) 4141–4149.
- [258] F. Weigend, Accurate Coulomb-fitting basis sets for H to Rn, Phys. Chem. Chem. Phys. 8 (2006) 1057–1065.

- [259] F. Weigend, 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, Phys. Chem. Chem. Phys. 7 (2005) 3297–3305.
- [260] M. Walker, A. J. A. Harvey, A. Sen, C. E. H. Dessent, Performance of M06, M06-2X, and M06-HF density functionals for conformationally flexible anionic clusters: M06 functionals perform better than B3LYP for a model system with dispersion and ionic hydrogen-bonding interactions, J. Phys. Chem. A 117 (2013) 12590–12600.
- [261] S. Koltzenburg, M. Maskos, O. Nuyken, Polymers as Materials, in Polym. Chem., Springer Berlin Heidelberg, Berlin, Heidelberg, 2017, pp. 141–162.
- [262] E. R. Johnson, S. Keinan, P. Mori-Sánchez, J. Contreras-García, A. J. Cohen, W. Yang, Revealing noncovalent interactions, J. Am. Chem. Soc. 132 (2010) 6498–6506.
- [263] C. Lefebvre, G. Rubez, H. Khartabil, J. C. Boisson, J. Contreras-García, E. Hénon, Accurately extracting the signature of intermolecular interactions present in the NCI plot of the reduced density gradient: Versus electron density, Phys. Chem. Chem. Phys. 19 (2017) 17928–17936.
- [264] M. Munoz, D. Ortiz, J. Nieto-Sandoval, Z. M. de Pedro, J. A. Casas, Adsorption of micropollutants onto realistic microplastics: Role of microplastic nature, size, age, and NOM fouling, Chemosphere 283 (2021) 131085.
- [265] J. Lin, D. Yan, J. Fu, Y. Chen, H. Ou, Ultraviolet-C and vacuum ultraviolet inducing surface degradation of microplastics, Water Res. 186 (2020) 116360.
- [266] M. C. Hemmer, V. Steinhauer, J. Gasteiger, Deriving the 3D structure of organic molecules from their infrared spectra, Vib. Spectrosc. 19 (1999) 151–164.
- [267] O. Devinyak, D. Havrylyuk, R. Lesyk, 3D-MoRSE descriptors explained, J. Mol. Graph. Model. 54 (2014) 194–203.
- [268] M. P. González, Z. Gándara, Y. Fall, G. Gómez, Radial Distribution Function descriptors for predicting affinity for vitamin D receptor, Eur. J. Med. Chem. 43 (2008) 1360–1365.
- [269] C. T. Klein, D. Kaiser, G. Ecker, Topological Distance Based 3D Descriptors for Use in QSAR and Diversity Analysis, J. Chem. Inf. Comput. Sci. 44 (2004) 200–209.
- [270] J. Dreher, J. Scheiber, N. Stiefl, K. Baumann, XMaP An Interpretable Alignment-Free Four-Dimensional Quantitative Structure-Activity Relationship Technique Based on Molecular Surface Properties and Conformer Ensembles, J. Chem. Inf. Model. 58 (2018)

165-181.

- [271] E. Estrada, A. Ramírez, Edge Adjacency Relationships and Molecular Topographic Descriptors. Definition and QSAR Applications, J. Chem. Inf. Comput. Sci. 36 (1996) 837– 843.
- [272] V. Consonni, R. Todeschini, M. Pavan, Structure/response correlations and similarity/diversity analysis by GETAWAY descriptors. 1. Theory of the novel 3D molecular descriptors, J. Chem. Inf. Comput. Sci. 42 (2002) 682–692.
- [273] A. Gajewicz-Skretna, M. Gromelski, E. Wyrzykowska, A. Furuhama, H. Yamamoto, N. Suzuki, Aquatic toxicity (Pre)screening strategy for structurally diverse chemicals: global or local classification tree models?, Ecotoxicol. Environ. Saf. 208 (2021) 111738.
- [274] D. Janežič, A. Miličević, S. Nikolić, N. Trinajstić, Graph-Theoretical Matrices in Chemistry2015.
- [275] I. Moriguchi, S. Hirono, I. Nakagome, H. Hirano, Comparison of Reliability of log P Values for Drugs Calculated by Several Methods, Chem. Pharm. Bull. 42 (1994) 976–978.
- [276] I. Moriguchi, S. Hirono, Q. Liu, I. Nakagome, Y. Matsushita, Simple Method of Calculating Octanol/Water Partition Coefficient, Chem. Pharm. Bull. 40 (1992) 127–130.
- [277] K. Roy, I. Mitra, Electrotopological State Atom (E-State) Index in Drug Design, QSAR, Property Prediction and Toxicity Assessment, Curr. Comput. Aided-Drug Des. 8 (2012) 135–158.
- [278] M. Cvetnic, D. Juretic Perisic, M. Kovacic, S. Ukic, T. Bolanca, B. Rasulev, H. Kusic, A. Loncaric Bozic, Toxicity of aromatic pollutants and photooxidative intermediates in water: A QSAR study, Ecotoxicol. Environ. Saf. 169 (2019) 918–927.
- [279] J. Papac Zjačić, S. Tonković, A. Pulitika, Z. Katančić, M. Kovačić, H. Kušić, Z. Hrnjak Murgić, A. Lončarić Božić, Effect of Aging on Physicochemical Properties and Size Distribution of PET Microplastic: Influence on Adsorption of Diclofenac and Toxicity Assessment, Toxics 11 (2023) 615.
- [280] H. Lee, W. J. Shim, J. H. Kwon, Sorption capacity of plastic debris for hydrophobic organic chemicals, Sci. Total Environ. 470–471 (2014) 1545–1552.
- [281] B. Gui, X. Xu, S. Zhang, Y. Wang, C. Li, D. Zhang, L. Su, Y. Zhao, Prediction of organic compounds adsorbed by polyethylene and chlorinated polyethylene microplastics in freshwater using QSAR, Environ. Res. 197 (2021) 111001.

- [282] J. Chen, Z. Gu, J. M. Perez-Aguilar, Y. Luo, K. Tian, Y. Luo, Molecular dynamics simulations reveal efficient heavy metal ion removal by two-dimensional Cu-THQ metalorganic framework membrane, Sci. Rep. 15 (2025) 1–11.
- [283] B. Barzegar, F. Feyzi, Designing cellulose based biochars for CO2 separation using molecular simulations, Sci. Rep. 15 (2025) 1643.
- [284] M. Tognia, M. Hou, X. Dong, L. Li, H. Wang, T. Wang, Integrated molecular and computational fluid dynamics study of CO2 mass transport mechanisms in carbon molecular sieve membranes, Sep. Purif. Technol. 363 (2025) 132068.
- [285] T. Kameda, A. Awazu, Y. Togashi, Molecular dynamics analysis of biomolecular systems including nucleic acids, Biophys. physicobiology 19 (2022) 1–18.
- [286] H. Zhang, J. Zheng, C. Lin, S. Yuan, Molecular dynamics study on adsorption and desorption of lysozyme above polymer antifouling membranes, Colloids Surfaces A Physicochem. Eng. Asp. 649 (2022) 129466.
- [287] J. W. Carter, A. S. Tascini, J. M. Seddon, F. Bresme, Molecular Dynamics Computer Simulations of Biological Systems, in Comput. Tools Chem. Biol., The Royal Society Of Chemistry, 2017.
- [288] C. Liu, Z. Zhao, J. Sui, H. Ma, L. Zhu, H. Jiang, R. Zhou, S. Wang, Y. Dai, The Sword of Damocles: Microplastics and the molecular dynamics of sulfamonomethoxine revealed, Ecotoxicol. Environ. Saf. 285 (2024).
- [289] G. Zhao, W. Li, C. Xu, Q. Qin, W. Fan, X. Li, D. Zhao, Adsorption mechanism of cefradine on three microplastics: A combined molecular dynamics simulation and density functional theory calculation study, Sci. Total Environ. 951 (2024) 175690.
- [290] P. Hobza, Chapter 20 Potential Energy and Free Energy Surfaces of Floppy Systems. Ab initio Calculations and Molecular Dynamics Simulations, in Mol. Dyn. (Eds.: P.B. Balbuena, J.M. Seminario) Vol. 7, Elsevier, 1999, pp. 829–857.
- [291] M. P. Allen, D. J. Tildesley, Introduction, in Comput. Simul. Liq., Oxford University Press, Oxford, 2017.
- [292] L. S. Dodda, I. C. De Vaca, J. Tirado-Rives, W. L. Jorgensen, LigParGen web server: An automatic OPLS-AA parameter generator for organic ligands, Nucleic Acids Res. 45 (2017) W331–W336.
- [293] J. D. Yesselman, D. J. Price, J. L. Knight, C. L. 3rd Brooks, MATCH: an atom-typing toolset

for molecular mechanics force fields., J. Comput. Chem. 33 (2012) 189–202.

- [294] J. Wang, W. Wang, P. Kollman, D. Case, ANTECHAMBER: an accessory software package for molecular mechanical calculations, J. Chem. Inf. Comput. Sci. - JCISD 222 (2000).
- [295] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. D. J. Mackerell, CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields., J. Comput. Chem. 31 (2010) 671–690.
- [296] S. Shang, Y. Liu, M. Liu, Y. Bai, X. Wang, B. Wu, J. Chen, J. Dong, Y. Liu, Studying the adsorption mechanisms of nanoplastics on covalent organic frameworks via molecular dynamics simulations, J. Hazard. Mater. 421 (2022).
- [297] P. Gupta, A. Mahapatra, B. Manna, A. Suman, S. S. Ray, N. Singhal, R. K. Singh, Sorption of PFOS onto polystyrene microplastics potentiates synergistic toxic effects during zebrafish embryogenesis and neurodevelopment, Chemosphere 366 (2024) 143462.
- [298] O. Hollóczki, S. Gehrke, Nanoplastics can change the secondary structure of proteins, Sci. Rep. 9 (2019) 1–7.
- [299] M. Sangkhawasi, T. Remsungnen, A. S. Vangnai, R. P. Poo-Arporn, T. Rungrotmongkol, All-Atom Molecular Dynamics Simulations on a Single Chain of PET and PEV Polymers, Polymers (Basel). 14 (2022) 1–11.
- [300] C. W. Yong, Descriptions and Implementations of DL-F Notation: A Natural Chemical Expression System of Atom Types for Molecular Simulations, J. Chem. Inf. Model. 56 (2016) 1405–1409.
- [301] C. W. Yong, Study of interactions between polymer nanoparticles and cell membranes at atomistic levels, Philos. Trans. R. Soc. B Biol. Sci. 370 (2015) 1–7.
- [302] W. L. Jorgensen, J. Tirado-Rives, Potential energy functions for atomic-level simulations of water and organic and biomolecular systems, Proc. Natl. Acad. Sci. U. S. A. 102 (2005) 6665–6670.
- [303] L. S. Dodda, J. Z. Vilseck, J. Tirado-Rives, W. L. Jorgensen, 1.14*CM1A-LBCC: Localized Bond-Charge Corrected CM1A Charges for Condensed-Phase Simulations, J. Phys. Chem. B 121 (2017) 3864–3870.
- [304] K. E. Murray, S. M. Thomas, A. A. Bodour, Prioritizing research for trace pollutants and

emerging contaminants in the freshwater environment, Environ. Pollut. 158 (2010) 3462–3471.

- [305] M. N. Miranda, A. R. Lado Ribeiro, A. M. T. Silva, M. F. R. Pereira, Can aged microplastics be transport vectors for organic micropollutants? – Sorption and phytotoxicity tests, Sci. Total Environ. 850 (2022).
- [306] V. Godoy, M. A. Martín-Lara, M. Calero, G. Blázquez, The relevance of interaction of chemicals/pollutants and microplastic samples as route for transporting contaminants, Process Saf. Environ. Prot. 138 (2020) 312–323.
- [307] I. Titov, J. Semerád, J. Boháčková, H. Beneš, T. Cajthaml, Microplastics meet micropollutants in a central european river stream: Adsorption of pollutants to microplastics under environmentally relevant conditions, Environ. Pollut. 360 (2024).
- [308] Y. Leng, W. Wang, H. Cai, F. Chang, W. Xiong, J. Wang, Sorption kinetics, isotherms and molecular dynamics simulation of 17β-estradiol onto microplastics, Sci. Total Environ. 858 (2023) 159803.

Curriculum vitae