The effect of the functional groups of organic contaminants on their adsorption to PET microplastics: a combined DFT and QSPR approach

Anamarija Pulitika, Panaghiotis Karamanis, Marin Kovačić, Ana Loncaric Bozic, Hrvoje Kusic

The effect of organic contaminants functional groups on their adsorption to PET microplastic; a combined DFT and QSPR approach

Anamarija Pulitika[a,b], Panaghiotis Karamanis[b], Marin Kovačić[a], Ana Loncaric Bozic[a], Hrvoje Kušic[a]*

[a] University of Zagreb Faculty of Chemical Engineering and Technology, 10000 Zagreb, Croatia

[b] E2S UPPA, CNRS, IPREM, Université de Pau et des Pays de l'Adour, 64053 Pau, France

*Corresponding author:

Phone: +385 1 4597 160

Email: hkusic@fkit.unizg.hr

Abstract

The adsorption of contaminants of emerging concern (CECs) onto microplastics (MPs) can alter their environmental behavior and toxicity. Thus, understanding adsorption and desorption processes is essential for accurate risk assessment. Models based on adsorption capacities lack robustness due to experimental variability. To address this, we applied density functional theory (DFT) and developed quantitative structure property relationship (QSPR) models to study interactions between polyethylene terephthalate (PET) MPs and model organic pollutants (m-COPs). Four QSPR models were developed by correlating interaction energies (E_{int}) at distinct PET adsorption sites with molecular descriptors of m-COPs. Model performance was confirmed through internal and external validation. Key descriptors reflect structural features influencing adsorption, particularly functional group size, flexibility, and the ability to form hydrogen bonds. Adsorption mechanisms varied depending on the adsorption site. At the sites close to terminal hydroxyl or carboxyl groups adsorption is mainly controlled by the ability of m-COPs to form hydrogen bonds. At other sites, adsorption was driven by molecular size and the ability to form multiple interactions all of which contribute to the total interaction energy.

Keywords: microplastic, adsorption, organic pollutants, DFT, QSPR

1. Introduction

Microplastics (MPs) are polymer particles smaller than 5 millimeters in diameter that are either produced intentionally or are the result of the fragmentation of larger plastics (Hartmann et al., 2019). Just like larger plastics, MPs are not degradable and therefore accumulate in the environment. A considerable amount of MPs is now found in all environmental compartments worldwide (Cole et al., 2011; Li et al., 2018A; Wang et al., 2021), from where they can easily enter organisms through ingestion, inhalation or direct tissue absorption (Urli et al., 2023). Consequently, traces of MPs are found in fruits, vegetables (Oliveri Conti et al., 2020), fish (Baechler et al., 2020), livestock (Omidi et al., 2012), and humans (Ragusa et al., 2021; Jenner et al., 2022; Leslie et al., 2022), raising the question of their toxicity and effect on health. Indeed, MPs have been shown to have different effects on organisms (Kogel et al., 2020). The studies on the toxic effect of MPs on zebrafish (*Danio rerio*) have shown that the presence of MPs in the organism can lead to intestinal and liver damage, oxidative stress, disruption of the reproductive system and ultimately death (Lu et al., 2016; Lei et al., 2018).

In addition, MPs can interact with inorganic and organic compounds (Godoy et al., 2019; Li et al, 2021; Papac, Zjačić et al., 2023), including contaminants of emerging concern (CECs) that involve cyanotoxins, flame retardants, organohalogens and perfluorinated compounds, personal care products, pesticides, plasticizers, and pharmaceuticals (Kovacic et al., 2020). Studies on MPs collected from the environment have also showed that they can adsorb various contaminants on their surface (Camacho et al., 2019; Capriotti et al., 2021; Elseblani et al., 2023). Therefore, MPs can contribute to the migration of contaminants through the environment and within an organism. Adsorption and desorption are dynamic processes that can alter the behaviour

of MPs and the overall toxicity of MPs and vectored contaminants (Bhagat et al., 2021; Rubin and Zucker, 2022).

For that reason, to properly assess the risk of MPs, it is crucial to understand the mechanism of adsorption/desorption and types of interactions between MPs and different types of contaminants. Previous studies have shown that the adsorption of organic contaminants on MPs results from the interplay of hydrophobic interactions, van der Waals forces, electrostatic interactions, π-π interactions and hydrogen bonding (Xia et al., 2023). The adsorption capacity and the extent of these interactions are determined by the type and physico-chemical properties of MPs and contaminants, as well as the surrounding matrix (Fu et al., 2021; Costigan et al., 2022). For example, MPs with a larger specific surface area generally have more adsorption sites and therefore a greater adsorption capacity (Li et al., 2018B; Ma et al., 2019). Polyamide (PA) MPs showed a higher adsorption capacity for pharmaceuticals compared to other types of MPs which is attributed to the formation of hydrogen bonds (Guo et al., 2019A; Guo et al., 2019B). Some studies suggest the adsorption of contaminants on MPs is higher in freshwater than in seawater (Li et al., 2018; Llorca et al., 2018), presumably due to the competition of ions and contaminants for the adsorption sites. On the other hand, a high ion concentration could increase adsorption by reducing the solubility of contaminants (Ma et al., 2019; Wu et al., 2020). Hence, quantitative experimental results are highly dependent on methodology, which makes the comparison of results from study to study quite challenging.

Quantitative Structure Property Relationship (QSPR) is a method used to correlate the quantitative results of adsorption studies and structural features of contaminants to develop models able to predict the adsorption of similar set of contaminants and potentially provide insight into the underlying adsorption mechanism (Ling et al., 2019; Tomic et al., 2022). Previous QSPR

studies on MPs adsorption indicate that lipophilicity, ionization potential, molecular size and branching were the main structural properties of contaminants that influenced the adsorption capacities on polyethylene (PE), polypropylene (PP) and polyvinyl chloride (PVC) MPs (Zhu et al., 2022), while π-electron conjugation influenced adsorption on polystyrene (PS) MPs (Li et al., 2020). These studies also suggest different adsorption mechanisms depending on the type and pretreatment of MPs as well as the type of environment (Li et al., 2020; Gui et al., 2021; Yao et al., 2022; Zhu et al., 2022; Hatinoglu et al., 2023). Therefore, the application of these models is limited to the specific MPs and environmental conditions used to obtain data for model development.

For this reason, in our work, we used computational chemistry combined with QSPR modelling to single out key structural features of common organic pollutants (m-COPs) that affect their adsorption on polyethylene terephthalate (PET) MPs. Specifically, we used density functional theory (DFT) methods to calculate the interaction energies between PET MPs and m-COPs, representing the response variable for QSPR modelling. The interaction energy is related to the number and strength of specific interactions (Cortés-Arriagada, 2021; Cortés-Arriagada et al., 2022; Cortés-Arriagada et al., 2023) and can therefore be used to provide insights into the mechanism (Yu et al., 2020; Mo et al., 2021). PET MPs was chosen because it is one of the most abundant plastics in the environment, with a significantly higher adsorption capacity compared to PE and PP (Munoz et al., 2021). Our goal is to develop a simple models able to explain the effects of different functional groups on adsorption, since m-COPs functional groups are often found in CECs determining their chemical properties and fate in the environment (Juretic et al., 2015).

2. Computational details

2.1. Interaction energy calculations

Gaussian 16 software (Frisch et al., 2019) was used to perform all the calculations in the framework of density functional theory. The method of choice was M06-L/def2-SVP. M06-L (Zhao and Truhlar, 2006) is a non-hybrid functional chosen for its ability to treat non-covalent interactions at a computationally lower cost compared to the hybrid functionals. Additionally, Grimme's D3 empirical dispersion (Goerigk et al., 2011) was used to account for the dispersive forces and density fitting of basis set that can accelerate the calculation of non-hybrid functionals without the significant loss of the accuracy.

Our PET model contains two trimers of ethylene terephthalate with a molecular formula $C_{60}H_{52}O_{26}$, optimized in the shape of a nanosphere. This structure was chosen as the previous studies (Cortés-Arriagada, 2021; Ortega and Cortés-Arriagada, 2023) showed that the adsorption onto MPs is controlled by many interactions that can be formed between pollutants and folded nanoparticle model, as opposed to the models that use only monomers or single oligomeric chains to represent MPs. Furthermore, our recent work (Pulitika et al., 2024) showed that the interactions between pollutants and MPs are local and that the further increase in the size of the model had no effect on the interaction energies.

Since the surface of PET MPs is heterogeneous, we assume it contains multiple adsorption sites that differ in their ability to adsorb organic pollutants. In our previous work (Pulitika et al., 2024), we identified four adsorption sites of interest, as showed in **Figure 1**.

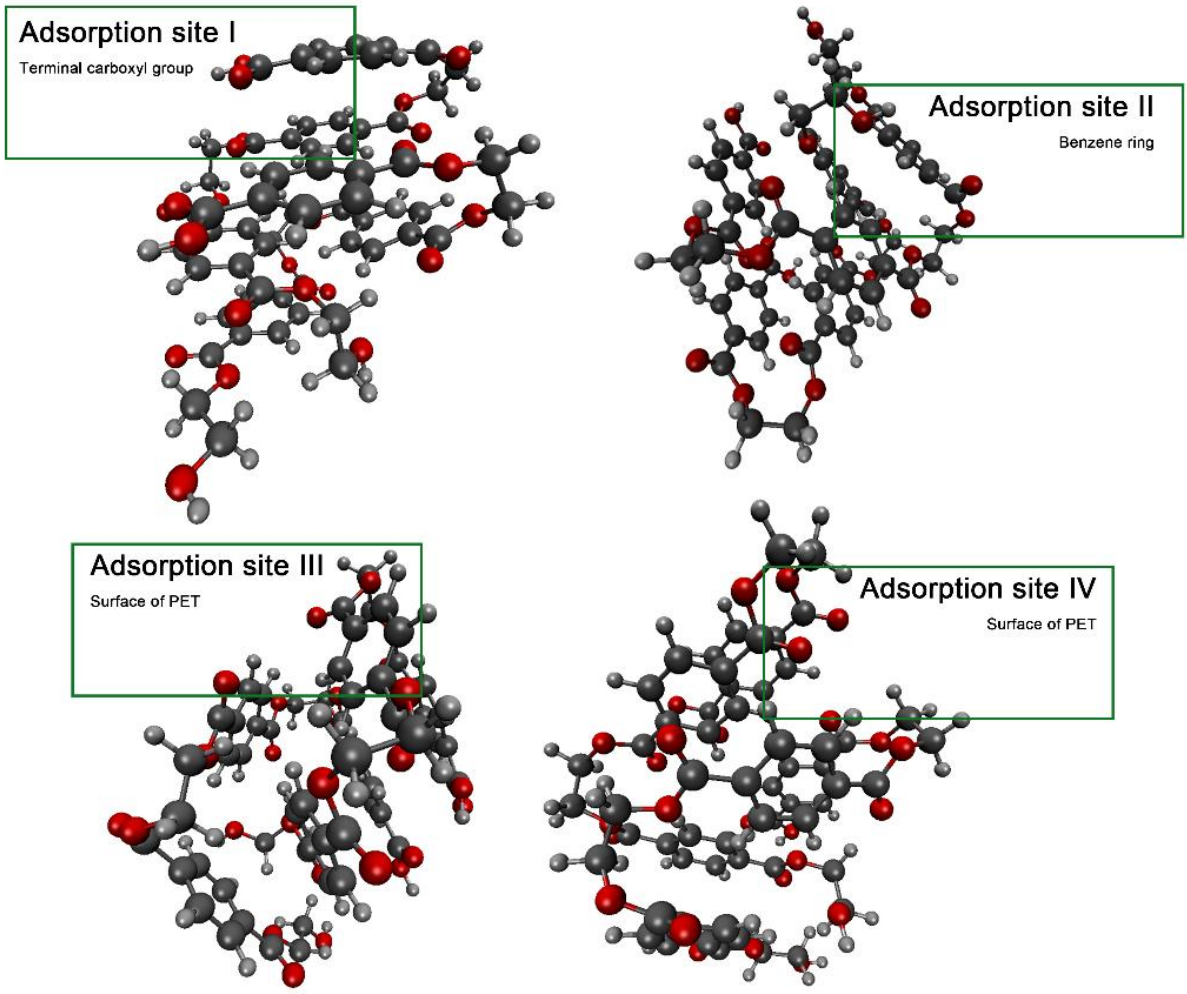

Figure 1. The structure of model PET microplastic with four adsorption sites.

The adsorption sites were chosen to represent different adsorption sites to account for the heterogeneity and complexity of the real polymer. Adsorption site I is the terminal carboxyl group, which can form strong hydrogen bonds, adsorption site II is the benzene ring, which can contribute to the interactions by forming π-π bonds. Adsorption site III and adsorption site IV are located on the surface of PET MPs and are presumably the most common types of adsorption sites in a realistic PET MPs.

The interactions were studied between our PET MPs model surface and a series of 30 m-COPs structured as single-benzene compounds containing different types of functional groups as

substituents as well as pyridine analogues. Their structures are showed in **Figure 2**. They were chosen because of their key features and functional groups found in complex organic pollutants, frequently abounded in environment (e.g. pesticides, pharmaceuticals, personal care products, organics dyes, etc), as well as for their simplicity, which is necessary for the efficiency of the calculations. All the binary systems containing PET MPs model surface and one of chosen m-COPs adsorbed to each of the adsorption sites were optimized and the vibrations frequencies were calculated as to avoid any imaginary frequencies and to ensure the system is in its true minima of energy. The interaction energies (E_{int}) between PET MPs and m-COPs were calculated as the difference between the total energy of the system ($E_{\text{PET/m-COP}}$) and the energies of only PET MPs unit (E_{PET}) and only m-COP unit ($E_{\text{m-COP}}$):

$$E_{\text{int}} = E_{\text{PET/m-COP}} - E_{\text{PET}} - E_{\text{m-COP}}$$

The more negative interaction energy corresponds to a stronger adsorption. The interaction energies were calculated using the counterpoise correction (CP) method (Gray et al., 2022) to account for the basis set superposition error (BSSE). BSSE is a mathematical error that arises from the way in which the energies of individual molecules and dimer are calculated, i.e. in a calculation of a dimer, an individual monomers can utilise some of the basis functions from one another which leads to the artificial stabilization of a complex and overestimation of interaction energy. In CP method the the energies of the dimer and the individual units are calculated using the same basis set which avoids BSSE. The interaction energies are presented for all four sites in **Table 1** and were further used as the response variables in QSPR modelling.

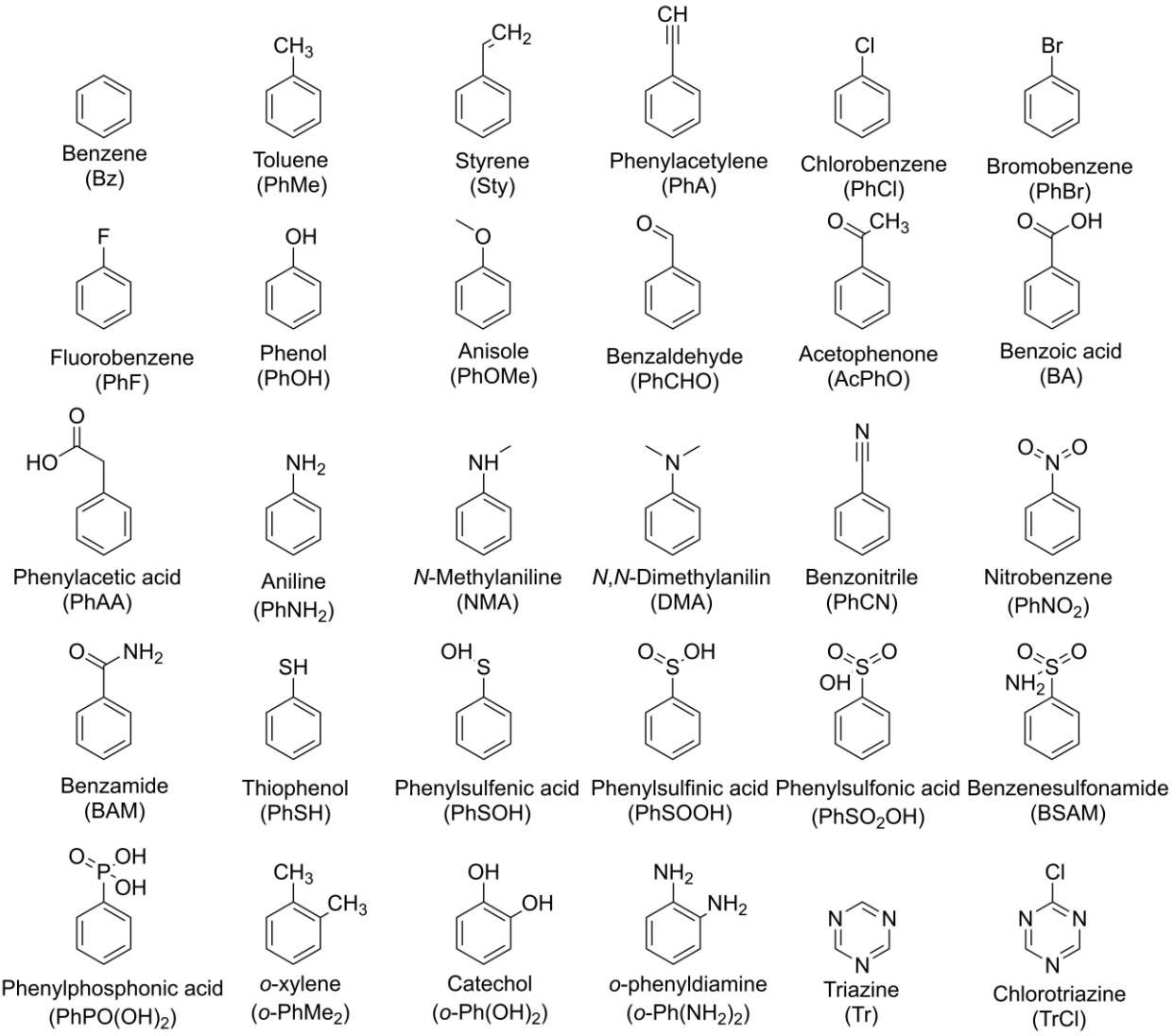

Figure 2. Structures of 30 compounds representing common organic pollutants (m-COPs) to study the adsorption interactions with model PET surface.

2.2. QSPR formulation

The structures of all 30 m-COPs were optimized by M06-L/def2-SVP method in Gaussian 16. The molecular descriptors were calculated by Dragon 6.0. software (Mauri et al., 2006) that uses empirical and semi-empirical methods to calculate the structural features of chosen m-COPs.

In total, 2526 descriptors were used to develop predictive models of interactions energies in dependence on the structural features of the m-COPs.

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. For each model, 25 m-COPs were selected for training set, while rest five compounds were placed into test set. The criteria for splitting into training and test sets were standard and rather simple: similar intervals of targeted values used in QSPR modeling were kept in both sets, meaning that . none of sets could contain only upper or bottom values of targeted responses. Hence, all 30 m-COPs were ordered by energy (from highest to lowest) and every 5^{th} compound (+/- one) was arbitrarily selected and transferred to test set, while remaining 25 m-COPs were part of training set. It should be noted that only the models up to 5 variables (i.e. descriptors) were considered, respecting "the rule of thumb" in QSPR modelling; the ratio of more than 1:5 between the number of variables (descriptors) and the number of molecules in training set is not desirable.

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.

Genetic algorithm (GA) method (Sukumar et al., 2014) 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. During GA and MLRA modeling actions, the filtering rules were employed to discard models with either highly correlated descriptors (Rij > 0.7) or those with the inadequate significance of either model of model term (p_M

or $p_T \geq 0.05$); QUIK rule was applied prior modeling, CI rule was activated after models were built.

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 (Gramatica, 2013).

Four models were chosen as the best-fitting ones, each predicting the interaction energies (E_{int}) of m-COPs (**Fig. 2**) at one adsorption site (**Fig. 1**). The models' coefficients were recalculated using the entire set of data (all 30 molecules) with normalized descriptors. The final Model I predicts the interaction energies with adsorption site I on PET MPs, Model II with the adsorption site II, Model III and Model IV, with adsorption sites III and IV respectively.

3. Results and Discussion

3.1 Interaction energies

The optimized geometries of m-COPs adsorbed on adsorption sites I, II, III and IV on PET MP model surface are showed in Figure S1-S6, while the interaction energies (E_{int}) calculated by M06-L/def2-SVP method with Grimme's D3 empirical dispersion for mentioned binary systems (i.e. PET MPs model surface and 30 m-COPs on four adsorption sites) are listed in **Table 1**. The interaction energies (E_{int}) on the adsorption site I are in range from –32.69 kcal mol^{-1} (calculated in the case of benzamide, BAM, adsorption) to –9.89 kcal mol^{-1} (calculated in the case of toluene,

PhMe, adsorption). The average interaction energy for all 30 m-COPs is –19.42 kcal mol^{-1}. 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. The m-COPs that adsorb strong to the adsorption sites (have more negative interaction energies), for example benzamide (BAM), phenylsulfonic acid ($PhSO_2OH$), phenylsulfinic acid (PhSOOH), phenylphosphonic acid ($PhPO(OH)_2$), and benzenesulfonamide (BSAM), have the ability to form hydrogen bonds with the PET MPs which can stabilize the adsorption. On the other hand, m-COPs incapable of forming hydrogen bonds, namely toluene (PhMe), benzene (Bz) and styrene (Sty), demonstrated the weakest adsorption affinity towards the model PET MP, as evidenced by the relatively high interaction energy values. Hydrogen bonds can be formed between the functional groups of the m-COPs and the carboxyl group of the adsorption site I, hydroxyl group found in the vicinity of the adsorption site II on our model PET MPs surface, and oxygen atoms of the ester group on the surface of PET MPs.

Table 1. Interactions energies between 30 m-COPs and model PET MPs surface of on four chosen 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).

Compound	Abbreviation	E_{int} / kcal mol^{-1}			
		I	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	**–11.28**	**–10.08**
Phenylacetylene	PhA	**–12.29**	–11.90	–11.18	–9.27
Chlorobenzene	PhCl	–11.71	–10.91	–10.75	**–8.88**
Bromobenzene	PhBr	–11.87	**–11.08**	–11.48	–8.97
Fluorobenzene	PhF	–12.10	–10.40	**–9.37**	–8.75
Phenol	PhOH	**–21.30**	–18.63	–16.59	–15.37
Anisole	PhOMe	–19.21	**–16.09**	–9.32	–11.53
Benzaldehyde	PhCHO	–20.34	–15.41	–10.63	–9.59
Acetophenone	AcPhO	–21.24	**–18.19**	–11.49	**–13.64**
Benzoic acid	BA	–29.43	–20.78	–17.58	–14.17
Phenylacetic acid	PhAA	–29.89	–17.25	**–16.20**	–13.98

Aniline	$PhNH_2$	**–20.18**	**–15.23**	–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	**–16.27**	–17.78
Thiophenol	PhSH	**–12.99**	–15.35	**–12.80**	**–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	$PhSO_2OH$	–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	**–15.69**
o-Phenyldiamine	*o*-$Ph(NH_2)_2$	–18.95	–18.71	–14.70	–12.39
Triazine	Tr	**–18.96**	–11.88	–7.79	–7.79
Chlorotriazine	TrCl	–20.05	–11.97	–9.35	–8.65

3.2. Validation of QSPR models

To further elucidate the effect of m-COP's structural features on the adsorption to PET MPs, the calculated interaction energies (E_{int}) were used as the responses to develop QSPR models. For each adsorption site, the m-COPs were split into training (25 molecules) and test set (5 molecules) as indicated in **Table 1** and QSPR modelling was performed to obtain a set of models with 1 to 5 variables (descriptors).

In **Figure 3** the coefficients of the determinations of models for the training and test set in accordance with the number of variables used in the model for each adsorption site are showed. As can be seen, R^2 values of the training set increase as the number of variables in the models also increases, since added variables contribute to better description of chosen data set. However, this does not apply for R^2 values of the test set, due to the fact that the models with too many variables are overfitted, providing lower predictivity of responses for chosen data set. Therefore, we have chosen as the best models for each adsorption site on the basis of cross-validation taking into account R^2 values calculated for both training and test sets, as well as values of other statistical

parameters listed in **Table S1** of supporting materials. Hence, in the case of adsorption sites I, III and IV the 3-variable models were chosen as the best, and for adsorption site II, the 4-variable model was chosen.

The statistical parameters of chosen models are showed in **Table S1** of supporting materials. The selected models were further validated by Leave Many Out technique (LMO) that 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 showed in **Figure S7**. For all four models, the Q^2_{LMO} values are mainly close to Q^2_{LOO} value and not widely disperse which indicates that our selected models are stable. The results of "*Y*-scrambling" test are showed in **Figure S8**. 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 which proves that models are valid, and do not result from the chance correlation. The fitting criteria values and the values of internal and external validation criteria for the selected models are listed in **Table S2** in the supporting materials.

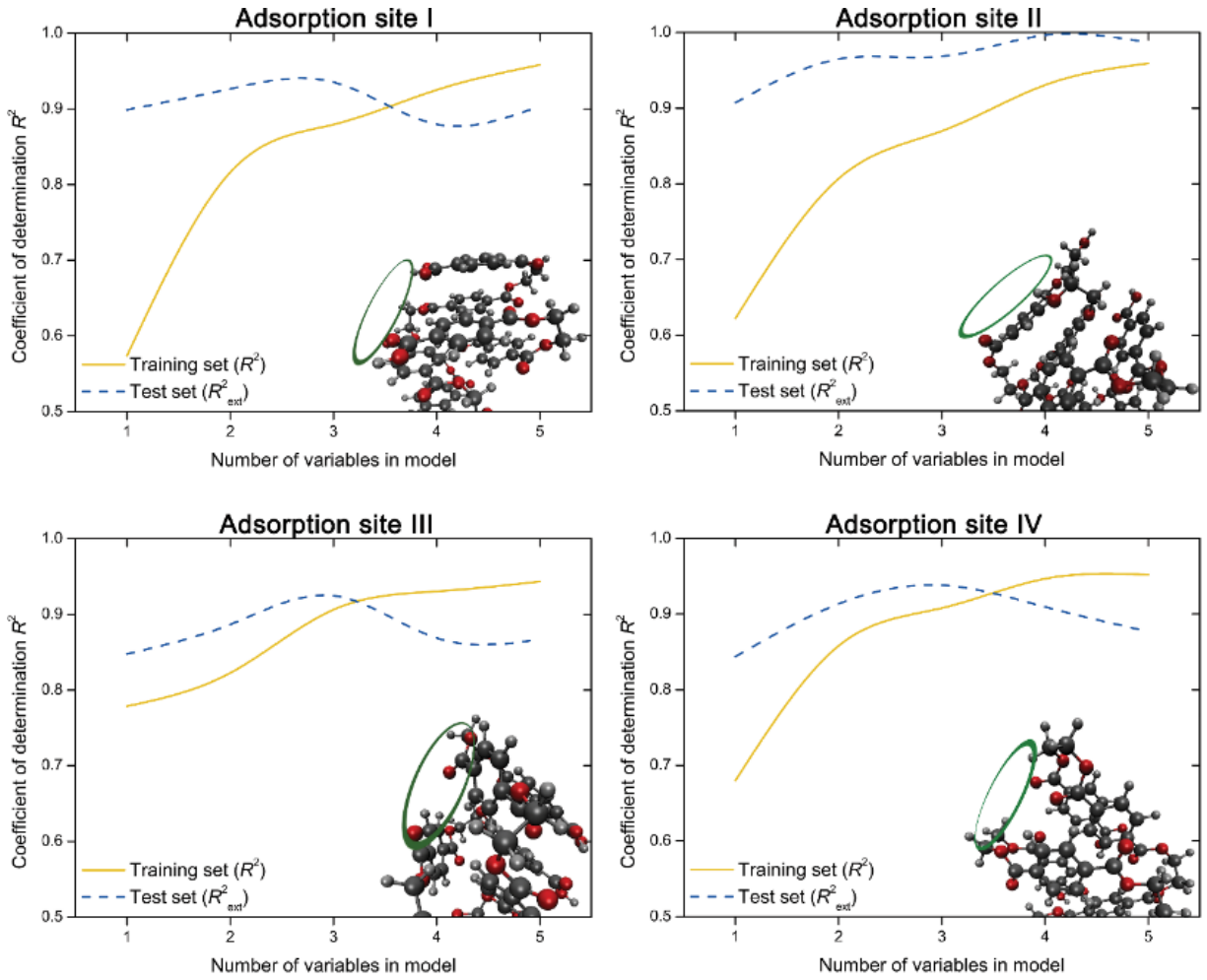

Figure 3. The values of coefficients of determination (R^2) of the training set and test of the QSPR models with 1-5 variables developed to predict the interaction energies of m-COPs with adsorption sites I, II, III and IV of PET MP.

The scatter plots of QSPR predicted vs. DFT calculated interaction energies employing selected models are showed in **Figure S9**. All the models represent given data very accurately which is also seen from the corresponding R^2 values; points and/or point clusters are located very close to the diagonal line in all cases. In **Figure S10** we show the Williams plots that study the applicability domain of the selected models and point out to response variable (Y) or structure (X)

outliers. In most cases there are no *Y*-outliers as all residuals fall inside ±3.0 σ limits; the only exception is BSAM compound in model for adsorption site I, which standard residual value slightly exceeds -3.0 σ limit. Williams plot of Model IV, that predicts the interaction energy between organic pollutants and adsorption site IV on the PET MP, shows the HAT value of catechol (*o*-Ph$(OH)_2$), which was placed in test set, is higher than the limit value (h^*) and is considered as *X*-outlier. However, HAT values of all the other molecules are below the limit. Overall, we conclude that are models can be considered as accurate.

3.3. 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-COPs and PET MP on adsorption sites I, II, III and IV, respectively, are listed in **Table S3**, along with pertaining statistical data. All the descriptors included are significant as their *p*-value is lower than the upper significance limit of 0.05. All descriptor pairs used in a specific model have a cross-correlation coefficients lower than the limit of 0.7 (R_{ij}>0.7) as can be seen in cross-correlation matrix presented in **Figure 4**, i.e., the models are linear.

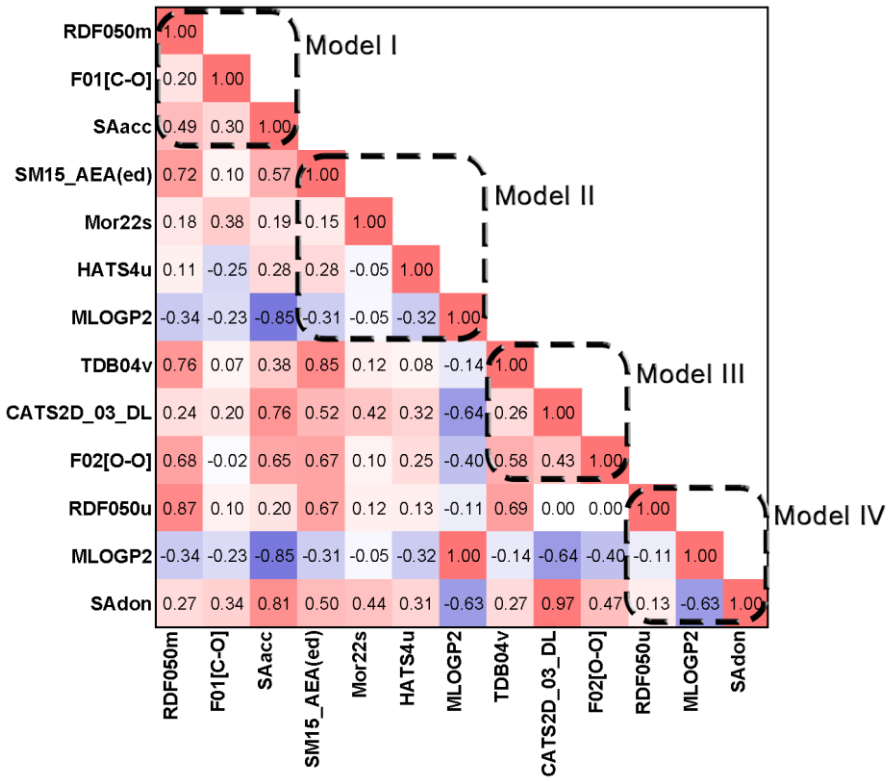

Figure 4. Cross-correlation matrix of descriptor variables in QSPR Model I, Model II, Model III and Model IV for the prediction of interaction energies of m-COPs with the adsorption sites I, II, III and IV respectively.

The descriptors used to build our models are listed in **Table 2**. 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 (Juretic et al., 2015). There is a high range of

3D-MoRSE descriptors that are usually denoted as Mor*sw*, 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 (Devinyak et al., 2014). RDF descriptors are based on the radial distribution function (RDF). RDF can be interpreted as the probability distribution of finding an atom in a spherical volume of radius *r* (Gonzalez et al., 2008). RDF descriptors are denoted as RDF*sw*, 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 (Sliwoski et al., 2015). 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 (Dreher et al., 2018). 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 (Tomic et al., 2022). 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*) (Zapadka et al., 2022). 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.

Table 2. Definitions of descriptors used in QSPR Model I, Model II, Model III and Model IV for the prediction of interaction energies of m-COPs with the adsorption sites I, II, III and IV respectively.

Model	Descriptor name	Descriptor definition	Descriptor type
I	**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
II	**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	GETAWAY descriptors
	MLOGP2	Squared Moriguchi octanol-water partition coefficient (logP^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	Squared Moriguchi octanol-water partition coefficient (logP^2)	Molecular properties
	SAdon	Surface area of donor atoms from P_VSA-like descriptors	Molecular properties

The performance of the selected models when applied on the entire set of data (30 m-COPs) with normalized descriptors is showed in **Figure 5,** the plot of residuals is showed in **Figure S11** and descriptive statistical data of the new coefficients is given in **Table 3**. The corresponding models' equations, together with their statistical parameters are presented below by equations (1) to (4):

- Model I:

$$E_{int}(\mathrm{I}) = -7.190\ (\pm 3.554)\mathbf{RDF050m} - 4.155\ (\pm 2.958)\mathbf{F01[C\text{-}O]} - 13.603\ (\pm 3.695)\mathbf{SAacc} - 11.285\ (\pm 1.612) \quad (1)$$

(n = 30; R^2 = 0.872; s = 2.670; F = 58.876; p < 0.0001; Q^2 = 0.826; S_{Press} = 3.108; S_{DEP} = 2.943)

- Model II:

$$E_{int}(\mathrm{II}) = -39.471\ (\pm 8.067)\ \mathbf{SM15\_AEA(ed)} - 3.261\ (\pm 1.327)\ \mathbf{Mor22s} + 4.839\ (\pm 2.634)\mathbf{HATS4u} + 8.064(\pm 1.636)\ \mathbf{MLOGP2} + 11.572\ (\pm 7.107) \quad (2)$$

(n = 30; R^2 = 0.929; s = 1.295; F = 83.296; p < 0.0001; Q^2 = 0.899; S_{Press} = 1.557; S_{DEP} = 1.445)

- Model III:

$$E_{int}(\mathrm{III}) = -5.811\ (\pm 4.168)\mathbf{TDB04v} - 6.773\ (\pm 1.926)\ \mathbf{CATS2D\_03\_DL} - 8.446\ (\pm 2.875)\ \mathbf{F02[O\text{-}O]} - 6.562\ (\pm 2.896) \quad (3)$$

(n = 30; R^2 = 0.903; s = 1.539; F = 80.418; p < 0.0001; Q^2 = 0.861; S_{Press} = 1.837; S_{DEP} = 1.739)

- Model IV:

$$E_{int}(\mathrm{IV}) = -8.152\ (\pm 1.911)\ \mathbf{RDF050u} + 2.690\ (\pm 2.051)\ \mathbf{MLOGP2} - 5.593\ (\pm 1.821)\ \mathbf{SAdon} - 8.296\ (\pm 1.558) \quad (4)$$

(n = 30; R^2 = 0.893; s = 1.280; F = 72.702; p < 0.0001; Q^2 = 0.859; S_{Press} = 1.472; S_{DEP} = 1.394)

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.

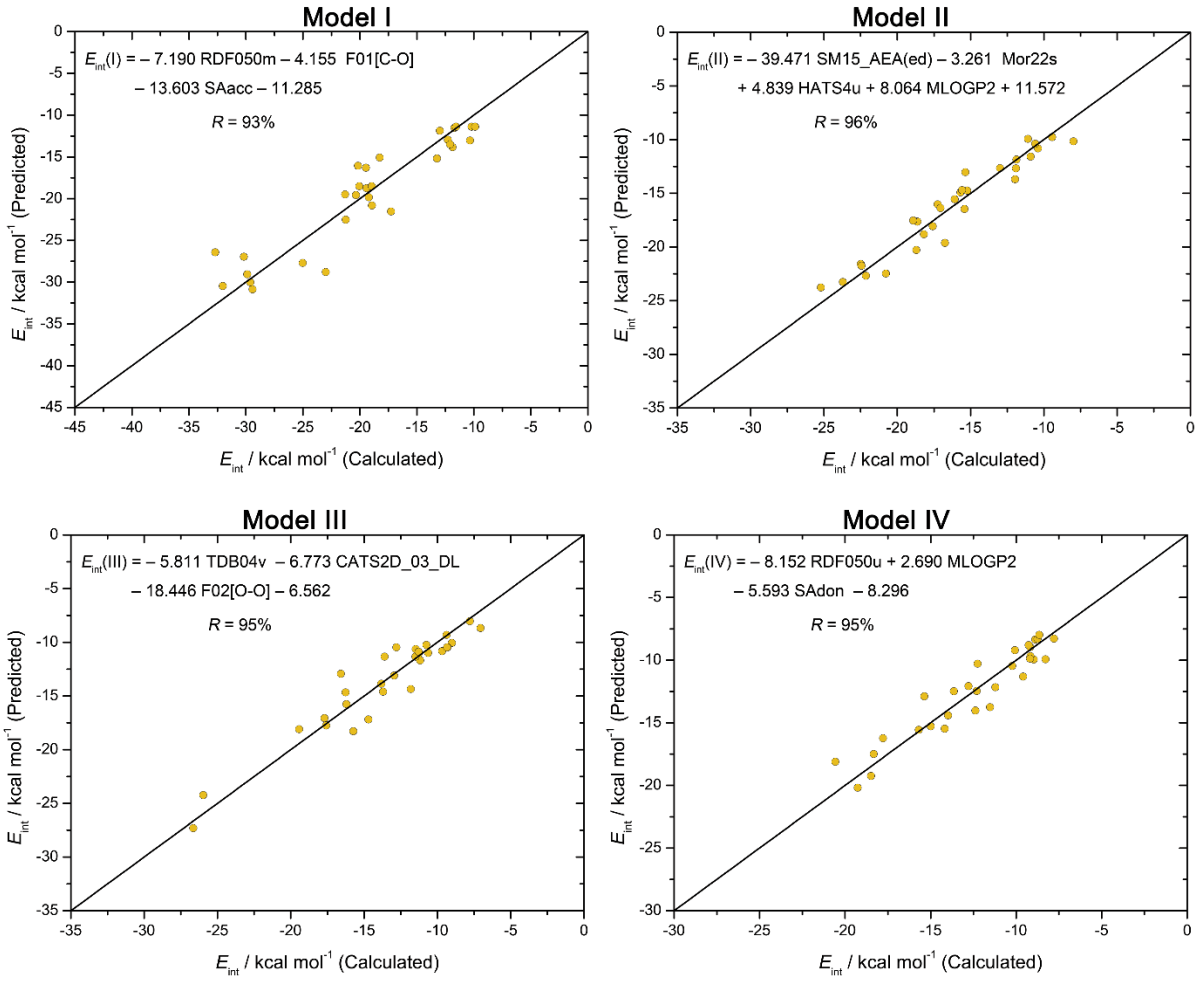

Figure 5. Calculated interaction energies and corresponding models' equations of QSPR Model I, Model II, Model III and Model IV that predict interaction energies of m-COPs with the adsorption sites I, II, III and IV on PET MP respectively when applied on the entire set of data (30 molecules) with normalized descriptors.

Table 3. 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-COPs with the adsorption sites I, II, III and IV respectively when applied on the entire set of data with normalized descriptors (30 molecules).

Model	Variable	Coefficient	Standard deviation	(+/-) 95% Confidence interval	*t*-ratio	*p*-value
I	Intercept	-11.29	0.78	1.61	-14.40	0.00
	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
II	Intercept	11.57	3.45	7.11	3.35	<0.01
	SM15_AEA(ed)	-39.47	3.92	8.07	-10.08	0.00
	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
III	Intercept	-6.56	1.41	2.90	-4.66	<0.01
	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
IV	Intercept	-8.30	0.76	1.56	-10.95	0.00
	RDF050u	-8.15	0.93	1.91	-8.77	0.00
	MLOGP2	2.69	1.00	2.05	2.70	<0.01
	SAdon	-5.59	0.89	1.82	-6.32	0.00

The descriptors included in Model I that describe the interaction energies between m-COPs and PET-MP at adsorption site I are **RDF050m**, **F01[C-O]** and **SAacc**. As can be seen from equation (1), 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 (Gajewicz-Skretna et al., 2021). 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-COPs selected are listed in Table S4 of the supporting material. The m-COPs that adsorb most strongly (with the lowest interaction energy) at adsorption site I have very high **SAacc** values. The m-COPs with oxygen-containing functional groups such as $PhSO_2OH$, 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*-$PhMe_2$, 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-COPs with an oxygen-containing functional group can form hydrogen bonds with the carboxyl group at adsorption site I. However, for some m-COPs such as $PhSO_2OH$, 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 neighboring parts of PET MPs, which can further stabilize the complex and facilitate the adsorption process.

Model II (eq. (2)) describes the relationship between the structural properties of selected m-COPs 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 S5. First two mentioned descriptors have negative coefficients in eq. (2), 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 (Janežič et al., 2015). **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-COPs in terms of molecular weight are usually more branched, and have a higher number of electronegative oxygen and/or nitrogen atoms. The m-COPs 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 (Lipinski et al., 2012). 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, which has been shown as essential for effective adsorption on PET surface (Pulitika et al., 2024), 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 (2) 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 (Roy and Mitra, 2012). Hence, it is has been found as important structural feature to reflected adsorption of selected m-COPs via hydrogen bonding to PET MPs surface in our case.

In the Model III (eq. 3), **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 c-COPs are listed in Table S6. **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-COPs onto PET MPs; higher **CATS2D_03_DL** value results in stronger adsorption of m-COPs. **F02[O-O]** represents the number of oxygen-oxygen atom pair separated by two bonds. As explained above, m-COPs 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-COPs are stronger for the m-COPs with a larger surface area.

Model IV describes the relationship between the structural features of m-COPs and their interaction energies with adsorption site IV at PET MP surface. The values of descriptors used in model IV are listed in Table S7. The descriptor with the highest contribution in eq. (4) is **RDF050u**. The same descriptor has been shown as contributing one in Model II, but there was its analogue weighted by molecular mass (**RDF050m**). We assume RDF050u contributes in a way that m-COPs characterized by a higher value have larger surface, therefore those m-COPs are able to interact by more weak interactions with the surface of PET MP than smaller m-COPs. **SAdon** falls into the same category of P_VSA like descriptors as **SAacc** (shown to be the most contributing in Model I, eq (1)), while it represents surface of electron donor atoms. **MLOGP2** negatively affects adsorption, the same as in in Model II (eq. (2)). 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-COPs and their ability to adsorb at PET MPs surface, represented in our case as interaction energies between m-COPs and PET MPs surface, refer either to their ability to form hydrogen bonds or to the size of the m-COPs. 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. When we examine the structure of the m-COPs (**Fig. 1**), we observed that the larger, more flexible functional groups have atoms that can form hydrogen bonds, while the optimized geometries of the m-COP/PET systems (**Fig S1 - Fig. S6**) were examined, it was evident that these groups are able to "reach" the terminal hydroxyl group of PET near adsorption site II and therefore form 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-COPs and the ability to have many interactions, all of which contribute to the total interaction energy. These results are consistent with our previous work where we examine intermolecular weak interactions using DFT methods (Pulitika et al., 2024).

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 of the macromolecule (Papac Zjačić et al., 2023). As was demonstrated by the findings presented herein, these sites facilitate adsorption

of m-COPs, especially those able to form hydrogen bonds such as $PhSO_2OH$, PhSOOH, and $PhPO(OH)_2$ used in this study. While hydrophobic interactions are commonly considered to be the most relevant (Lee et al., 2014; Gui et al., 2021) 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 favorable aromatic moieties.

4. Conclusion

In this work, we studied the adsorption of PET MPs and 30 single benzene ring compounds and pyridine analogues, i.e. m-COPs, containing various functional groups commonly present in CECs. The quantitative results of the interaction energies between PET MPs model surface and 30 m-COP were used as end-points for the QSPR modelling. Our results showed that the adsorption of m-COPs differs depending on the adsorption site at the PET MPs surface. The interaction energies between PET MPs and m-COPs were strongest for adsorption site I, especially for benzamide, benzoic acid, phenylacetic acid, phenylphosphonic acid, phenylsulfonic acid, and phenylsulfinic acid, all of which can form multiple hydrogen bonds with terminal carboxyl group on PET MPs. 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 QSPR models developed by correlating the interaction energies between m-COPs and PET MPs at all four adsorption sites and descriptors of m-COPs provide further insight into the key structural

features of m-COPs for adsorption to PET MPs. For adsorption site I, 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, 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, we concluded that 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 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. These findings indicate that ageing of PET, and other MPs as well, in the environment could lead to an increased affinity towards adsorption of certain m-COPs due to the formation of hydrogen bonds with the carboxyl groups on the surface of aged MPs.

Acknowledgements

We would like to acknowledge the financial support by the Croatian Science Foundation under the project *Microplastic in water; fate and behaviour and removal*, ***ReMiCRO*** (IP-2020-02-6033). 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. We also acknowledge the "Direction du Numérique" of the "Université de Pau et des Pays de l'Adour" for the computing facilities provided. Part of this work was granted access to the HPC resources of (CCRT/CINES/IDRIS) under the allocations 2022-AD010807031R1, 2023- AD010807031R2

made by GENCI (Grand Equipment National de Calcul Intensif). We also gratefully acknowledge Prof. Paola Gramatica for enabling a free license of software QSARins (version 2.2.4).

References

Baechler, B.R., Stienbarger, C.D., Horn, D.A., Joseph, J., Taylor, A.R., Granek, E.F., Brander, S.M. 2020. Microplastic occurrence and effects in commercially harvested North American finfish and shellfish: Current knowledge and future directions. Limnol. Oceanogr. Lett. 5, 113–136.

Bhagat, J., Nishimura, N., Shimada, Y., 2021. Toxicological interactions of microplastics/nanoplastics and environmental contaminants: Current knowledge and future perspectives. J. Hazard. Mater. 405, 123913.

Camacho, M., Herrera, A., Gómez, M., Acosta-Dacal, A., Martínez, I., Henríquez-Hernández, L.A., Luzardo, O.P. 2019. Microplastic debris in beaches of Tenerife (Canary Islands, Spain). Sci. Total Environ. 662, 22–31.

Capriotti, M., Cocci, P., Bracchetti, L., Cottone, E., Scandiffio, R., Caprioli, G., Sagratini, G., Mosconi, G., Bovolin, P., Palermo, F.A. 2021. Microplastics and their associated organic pollutants from the coastal waters of the central Adriatic Sea (Italy): Investigation of adipogenic effects in vitro. Chemosphere 263, 128090.

Cole, M., Lindeque, P., Halsband, C., Galloway, T.S. 2011. Microplastics as contaminants in the marine environment: A review. Mar. Pollut. Bull. 62, 2588–2597.

Cortés-Arriagada, D. 2021. Elucidating the co-transport of bisphenol A with polyethylene terephthalate (PET) nanoplastics: A theoretical study of the adsorption mechanism. Environ. Pollut. 270, 116192.

Cortés-Arriagada, D., Miranda-Rojas, S., Belén, M., Ortega, D.E., Alarcón-Palacio, V.B. 2023. The interaction mechanism of polystyrene microplastics with pharmaceuticals and personal care products. Sci. Total Environ. 861, 160632.

Cortés-Arriagada, D., Ortega, D.E., Miranda-Rojas, S. 2023. Mechanistic insights into the adsorption of endocrine disruptors onto polystyrene microplastics in water. Environ. Pollut. 319, 121017.

Costigan, E., Collins, A., Hatinoglu, M. D., Bhagat, K., MacRae, J., Perreault, F., Apul, O. 2022. Adsorption of organic pollutants by microplastics: overview of a dissonant literature. J. Hazard. Mater. Adv. 6, 100091.

Devinyak, O., Havrylyuk, D., Lesyk, R. 2014. 3D-MoRSE descriptors explained. *J*. Mol. Graph. Model. 54, 194–203.

Dreher, J., Scheiber, J., Stiefl, N., Baumann, K. 2018. 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, 165–181.

Elseblani, R., Cobo-Golpe, M., Godin, S., Jimenez-Lamana, J., Fakhri, M., Rodríguez, I., Szpunar, J. 2023. Study of metal and organic contaminants transported by microplastics in the Lebanese coastal environment using ICP MS, GC-MS, and LC-MS. Sci. Total Environ. 887, 164111.

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

J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery, J. A., Jr., Peralta, J.E., Ogliaro, F., Bearpark, M.J., Heyd, J.J., Brothers, E.N., Kudin, K.N., Staroverov, V.N., Keith, T.A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A.P., Burant, J.C., Iyengar, S.S., Tomasi, J., Cossi, M., Millam, J.M., Klene, M., Adamo, C., Cammi, R., Ochterski, J.W., Martin, R.L., Morokuma, K., Farkas, O., Foresman, J.B., Fox, D.J. 2019. Gaussian 16, Revision C.01.

Fu, L., Li, J., Wang, G., Luan, Y., Dai, W. 2021. Adsorption behavior of organic pollutants on microplastics. Ecotoxicol. Environ. Saf. 217, 112207.

Gajewicz-Skretna, A., Gromelski, M., Wyrzykowska, E., Furuhama, A., Yamamoto, H., Suzuki, N. 2021. Aquatic toxicity (Pre)screening strategy for structurally diverse chemicals: global or local classification tree models? Ecotoxicol. Environ. Saf. 208, 111738.

Godoy, V., Blázquez, G., Calero, M., Quesada, L., Martín-Lara, M.A. 2019. The potential of microplastics as carriers of metals. Environ. Pollut. 255, 113363.

Goerigk, L., Kruse, H., Grimme, S. 2011. Benchmarking density functional methods against the S66 and S66x8 datasets for non-covalent interactions. ChemPhysChem 12, 3421–3433.

González, M.P., Gándara, Z., Fall, Y., Gómez, G. 2008. Radial Distribution Function descriptors for predicting affinity for vitamin D receptor. Eur. J. Med. Chem. 43, 1360–1365.

Gramatica, P. In: Computational Toxicology. Humana Press, (Eds.: Reisfeld, B., Mayeno, A. N.), Totowa NJ, 2013, pp. 499–526.

Gray, M., Bowling, P. E., Herbert, J. M. 2022. Systematic Evaluation of Counterpoise Correction in Density Functional Theory. J. Chem. Theory Comput. 18 (11), 6742-6756.

Gui, B., Xu, X., Zhang, S., Wang, Y., Li, C., Zhang, D., Su, L., Zhao, Y. 2021. Prediction of

organic compounds adsorbed by polyethylene and chlorinated polyethylene microplastics in freshwater using QSAR. Environ. Res. 197, 111001.

Guo, X., Chen, C., Wang, J. 2019A. Sorption of sulfamethoxazole onto six types of microplastics. Chemosphere 228, 300–308.

Guo, X., Liu, Y., Wang, J. 2019B. Sorption of sulfamethazine onto different types of microplastics: A combined experimental and molecular dynamics simulation study. Mar. Pollut. Bull. 145, 547–554.

Hartmann, N.B., Hüffer, T., Thompson, R.C., Hassellöv, M., Verschoor, A., Daugaard, A.E., Rist, S., Karlsson, T., Brennholt, N., Cole, M., Herrling, M.P., Hess, M.C., Ivleva, N.P., Lusher, A.L., Wagner, M., 2019. Are we speaking the same language? recommendations for a definition and categorization framework for plastic debris. Environ. Sci. Technol. 53, 1039–1047.

Hatinoglu, D., Adan, A., Perreault, F., Imamoglu, I., Apul, O.G. 2023. Linear solvation energy relationships for adsorption of aromatic organic compounds by microplastics. Chem. Eng. Sci. 282, 119233.

Janežič, D., Miličević, A., Nikolić, S., Trinajstić, N. Graph-Theoretical Matrices in Chemistry. CRC Press, Boca Raton, USA, 2015.

Jenner, L.C., Rotchell, J.M., Bennett, R.T., Cowen, M., Tentzeris, V., Sadofsky, L.R. 2022. Detection of microplastics in human lung tissue using μFTIR spectroscopy. Sci. Total Environ. 831, 154907.

Juretic, D., Kusic, H., Dionysiou, D.D., Rasulev, B., Peternel, I., Loncaric Bozic, A. 2015. Prediction of key structural features responsible for aromaticity of single-benzene ring pollutants and their photooxidative intermediates. Chem. Eng. J. 276, 261–273.

Kovacic, M., Kusic, H., Loncaric Bozic, A., Dionysiou, D.D., In: Encyclopedia of Water: Science, Technology, and Society. Part IV Water Technology: Water Treatment and Supply. Wiley, (Ed.: Maurice, P.) 2020. pp. 1925-1940.

Kögel, T., Bjorøy, Ø., Toto, B., Bienfait, A.M., Sanden, M. 2020. Micro- and nanoplastic toxicity on aquatic life: determining factors. Sci. Total Environ. 709, 136050.

Lee, H., Shim, W.J., Kwon, J.H. 2014. Sorption capacity of plastic debris for hydrophobic organic chemicals. Sci. Total Environ. 470–471, 1545–1552.

Lei, L., Wu, S., Lu, S., Liu, M., Song, Y., Fu, Z., Shi, H., Raley-Susman, K.M., He, D. 2018. Microplastic particles cause intestinal damage and other adverse effects in zebrafish *Danio rerio* and nematode *Caenorhabditis elegans*. Sci. Total Environ. 619–620, 1–8.

Leslie, H.A., van Velzen, M.J.M., Brandsma, S.H., Vethaak, A.D., Garcia-Vallejo, J.J., Lamoree, M.H. 2022. Discovery and quantification of plastic particle pollution in human blood. Environ. Int. 163, 107199.

Li, H., Wang, F., Li, J., Deng, S., Zhang, S. 2021. Adsorption of three pesticides on polyethylene microplastics in aqueous solutions: Kinetics, isotherms, thermodynamics, and molecular dynamics simulation. Chemosphere 264, 128556.

Li, J., Liu, H., Chen, J.P. 2018A. Microplastics in freshwater systems: A review on occurrence, environmental effects, and methods for microplastics detection. Water Res. 137, 362–374.

Li, J., Zhang, K., Zhang, H., 2018B. Adsorption of antibiotics on microplastics. Environ. Pollut. 237, 460–467.

Li, M., Yu, H., Wang, Y., Li, J., Ma, G., Wei, X. 2020. QSPR models for predicting the adsorption capacity for microplastics of polyethylene, polypropylene and polystyrene. Sci. Rep. 10, 1–11.

Ling, Y., Klemes, M.J., Steinschneider, S., Dichtel, W.R., Helbling, D.E. 2019. QSARs to predict adsorption affinity of organic micropollutants for activated carbon and β-cyclodextrin polymer adsorbents. Water Res. 154, 217–226.

Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. 2012. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Deliv. Rev. 64, 4-17.

Llorca, M., Schirinzi, G., Martínez, M., Barceló, D., Farré, M. 2018. Adsorption of perfluoroalkyl substances on microplastics under environmental conditions. Environ. Pollut. 235, 680–691.

Lu, Y., Zhang, Y., Deng, Y., Jiang, W., Zhao, Y., Geng, J., Ding, L., Ren, H. 2016. Uptake and accumulation of polystyrene microplastics in zebrafish (*Danio rerio*) and toxic effects in liver. Environ. Sci. Technol. 50, 4054–4060.

Ma, J., Zhao, J., Zhu, Z., Li, L., Yu, F. 2019. Effect of microplastic size on the adsorption behavior and mechanism of triclosan on polyvinyl chloride. Environ. Pollut. 254, 113104.

Mauri, A., Consonni, V., Pavan, M., Todeschini, R. 2006. DRAGON software: an easy approach to molecular descriptor calculations. Match 56, 237–248.

Mo, Q., Yang, X., Wang, J., Xu, H., Li, W., Fan, Q., Gao, S., Yang, W., Gao, C., Liao, D., Li, Y., Zhang, Y. 2021. Adsorption mechanism of two pesticides on polyethylene and polypropylene microplastics: DFT calculations and particle size effects. Environ. Pollut. 291, 118120.

Munoz, M., Ortiz, D., Nieto-Sandoval, J., M. de Pedro, Z., Casas, J.A. 2021. Adsorption of micropollutants onto realistic microplastics: Role of microplastic nature, size, age, and NOM fouling. Chemosphere 283, 131085.

Oliveri Conti, G., Ferrante, M., Banni, M., Favara, C., Nicolosi, I., Cristaldi, A., Fiore, M., Zuccarello, P. 2020. Micro- and nano-plastics in edible fruit and vegetables. The first diet risks assessment for the general population. Environ. Res. 187, 109677.

Omidi, A., Naeemipoor, H., Hossein, M., 2012. Plastic debris in the digestive tract of sheep and goats: an increasing environmental contamination in Birjand, Iran. Bull. Env. Contam. Toxicol. 88, 691–694.

Ortega, D.E., Cortés-Arriagada, D. 2023. Atmospheric microplastics and nanoplastics as vectors of primary air pollutants - A theoretical study on the polyethylene terephthalate (PET) case. Environ. Pollut. 318, 120860.

Papac Zjačić, J., Tonković, S., Pulitika, A., Katančić, Z., Kovačić, M., Kušić, H., Hrnjak Murgić, Z., Lončarić Božić, A. 2023. Effect of aging on physicochemical properties and size distribution of PET microplastic; influence on adsorption of diclofenac and toxicity assessment. Toxics 11, 615.

Pulitika, A., Karamanis, P., Kovačić, M., Lončarić Božić, A., Kušić, H. An atomic-level perspective on the interactions between organic oollutants and PET particles: a comprehensive ab-initio and DFT investigation. ChemPhysChem 25, e202300854.

Ragusa, A., Svelato, A., Santacroce, C., Catalano, P., Notarstefano, V., Carnevali, O., Papa, F., Rongioletti, M.C.A., Baiocco, F., Draghi, S., D'Amore, E., Rinaldo, D., Matta, M., Giorgini, E. 2021. Plasticenta: First evidence of microplastics in human placenta. Environ. Int. 146, 106274.

Roy, K., Mitra, I., 2012. Electrotopological state atom (E-State) index in drug design, QSAR, property prediction and toxicity assessment. Curr. Comput. Aided-Drug Des. 8, 135–158.

Rubin, A.E., Zucker, I. 2022. Interactions of microplastics and organic compounds in aquatic

environments: A case study of augmented joint toxicity. Chemosphere 289, 133212.

Sliwoski, G., Mendenhall, J., Meiler, J. 2015. Autocorrelation descriptor improvements for QSAR: 2DA_sign and 3DA_sign. J. Comput. Aided Mol. Des. 30(3), 209–217

Sukumar, N., Prabhu, G., Saha, P. In: Applied Metaheuristics Process Enginerring. Springer (Eds.: Valadi, J., Siarry, P.), Cham, 2014, pp. 1–444.

Tomic, A., Cvetnic, M., Kovacic, M., Kusic, H., Karamanis, P., Bozic, A.L. 2022. Structural featrues promoting adsorption of contaminants of emerging concern onto TiO_2; experimental and computational approaches. Environ. Sci. Pollut. Res. 29, 87628–87644.

Urli, S., Corte Pause, F., Crociati, M., Baufeld, A., Monaci, M., Stradaioli, G. 2023. Impact of microplastics and nanoplastics on livestock health: an emerging risk for reproductive efficiency. Animals 13, 1–16.

Wang, C., Zhao, J., Xing, B., 2021. Environmental source, fate, and toxicity of microplastics. J. Hazard. Mater. 407, 124357.

Wu, X., Liu, P., Huang, H., Gao, S. 2020. Adsorption of triclosan onto different aged polypropylene microplastics: Critical effect of cations. Sci. Total Environ. 717, 137033.

Xia, Y., Niu, S., Yu, J. 2023. Microplastics as vectors of organic pollutants in aquatic environment: A review on mechanisms, numerical models, and influencing factors. Sci. Total Environ. 887, 164008.

Yao, J., Wen, J., Li, H., Yang, Y. 2022. Surface functional groups determine adsorption of pharmaceuticals and personal care products on polypropylene microplastics. J. Hazard. Mater. 423, 127131.

Yu, H., Yang, B., Waigi, M.G., Peng, F., Li, Z., Hu, X. 2020. The effects of functional groups on the sorption of naphthalene on microplastics. Chemosphere 261, 127592.

Zapadka, M., Dekowski, P., Kupcewicz, B. 2022. HATS5m as an example of GETAWAY molecular descriptor in assessing the similarity/diversity of the structural features of 4-thiazolidinone. Int. J. Mol. Sci. 23, 6576.

Zhao, Y., Truhlar, D.G. 2006. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. J. Chem. Phys. 125, 194101.

Zhu, T., Tao, C., Cheng, H., Cong, H. 2022. Versatile in silico modelling of microplastics adsorption capacity in aqueous environment based on molecular descriptor and machine learning. Sci. Total Environ. 846, 157455.