Extended Regression Modeling of the Toxicity of Phenol Derivatives to *Tetrahymena pyriformis* Using the Electronic-Structure Informatics Descriptor

Algafari Bakti Manggara^a and Manabu Sugimoto^{a,b,c,*}

^aGraduate School of Science and Technology, ^bFaculty of Advanced Science and Technology,
^cInstitute of Industrial Nanomaterials, Kumamoto University,
2-39-1 Kurokami, Chuo-ku, Kumamoto 860-8555, Japan

(Received: July 7, 2021; Accepted: August 4, 2021)

The extreme gradient boosting regression (XGBR) method was applied to 245 phenol derivatives in order to establish a regression model to predict their toxicity to *Tetrahymena pyriformis*. The modeling was done by using the set of “electronic-structure informatics” (ESI) descriptors recently suggested by the present authors. It is shown that the XGBR method is successful in predicting the toxicity of phenols in each class of five modes of action (MOA). A feature importance analysis showed that the different ESI descriptors were found to be important depending on the MOA. Through comparisons with the optimized descriptor set previously suggested by M. T. D. Cronin *et al.* (*Chemosphere.*, **49**, 1201-1221, (2002)), it is shown that the ESI descriptor set, which has been applied to different types of target variables, is of similar quality in regression modeling.

Key Words: electronic-structure informatics, machine learning, DFT, toxicity, mode of action

1. Introduction

In-silico prediction of toxicity is challenging [1] since its evaluation requires a lot of chemical and biological experiments. From the view of avoiding ethical issues, progress in such a computational technique to quantitatively predict the toxicity of compounds is demanded in order to reduce the number of animal applications [2].

The toxicity of a compound may arise from a variety of modes of action (MOA). This information is essentially indispensable for quantitative toxicity prediction since different factors would play roles in different MOA.

Therefore, it is desired to establish computational methods that facilitate reliable feature importance analysis for such prediction.

A lot of efforts have been made to develop such a method in chemoinformatics. For example, Schultz *et al.* extensively investigated the toxicity of phenol derivatives to ciliated protozoan such as *Tetrahymena pyriformis* [3]. They experimentally evaluated their toxicity by referring to the concentration in millimoles of a chemical for 50 % inhibition of growth (*IGC*₅₀). They applied the so-called QSAR (quantitative structure-activity relationship) method. Since the toxicity of chemicals arises from chemical reactions such as electrophilic reaction, Cronin *et al.* suggested to include electronic descriptors such as

* sugimoto@kumamoto-u.ac.jp

the energy level of the lowest unoccupied molecular orbital (LUMO) in addition to the original QSAR descriptors such as $\log P$ where P is a partition coefficient of a chemical into aqueous and octanol phases in a water/octanol bilayer system [4].

We have recently suggested a set of “electronic-structure informatics” (ESI) descriptors. The descriptor set has been successfully applied to predict biological activities such as toxicity and inhibitor activity of small molecules [5-7]. In the toxicity prediction of 33 alkylphenols to *Tetrahymena pyriformis* [5], vertical electron affinity (EA_v) and size parameters were found to be dominant among the ESI descriptors. This result implies that a molecular recognition process would play a critical role in MOA for their toxicity.

It is well known that phenol has various derivatives and exhibit toxicity via different MOAs [8]. Schultz *et al.* [3] suggested the existence of the five MOAs known as polar narcotic (PN), weak acid respiratory uncoupler (WARU), pro-electrophile (PE), soft electrophile (SE), and pro-redox cyler (PRC). Knowing dominant factors for each MOA would be important for deeper understanding of these MOA classes. For this purpose, it is desired to establish a machine-learning regression model with a flexible descriptor.

Nowadays a variety of machine learning models are available for evaluating the toxicity [9]. So far, the multiple linear regression (MLR) analysis has been extensively used for predicting the toxicity of phenol derivatives [4, 5, 10, 11]. The method has limitations in establishing general relation including nonlinearity between the target variable (toxicity) and descriptors. Therefore, in order to perform extended regression modeling, a machine learning method beyond MLR seems demanded for toxicity prediction for phenols.

In the present study, we try to establish a reliable regression model for predicting the toxicity of phenol derivatives to *Tetrahymena pyriformis*. Herein we apply the extreme gradient boosting regression (XGBR) method because many successful applications have been recently reported in the literature [12]. This method provides the biggest advantage even on the prediction speed and performance [13-15]. The method also has advantage for our purpose since it has capability of preforming feature importance analysis [14, 16]. We apply a simplified version of a set of the ESI descriptors in the XGBR regression modeling 245 phenol derivatives whose $pIGC_{50}$ were reported by Cronin *et al.* [4].

2. Calculation of the ESI Descriptors

The ESI descriptors used in the present paper are summarized in Table S1 in Supplementary Materials. They were numerically evaluated in electronic-structure calculations performed using the Gaussian 16 program

[17]. The density functional theory (DFT) method with the M06-2X functional and the 6-31G (d, p) basis set were applied.

In the descriptor evaluation procedure, we firstly optimized molecular structures, which was followed by the normal mode analysis. Then, the molecular properties, corresponding to the ESI descriptors, were calculated using the same computational method. In the set of the ESI descriptors, information on electronically excited states is included. This calculation was carried out using the time-dependent DFT (TDDFT) method where 30 low-lying excited states were calculated [5-7]. The solvation energy was evaluated using the polarizable continuum model (PCM). The Mathematica 11 software [18] was also used to determine the size parameters (B_1 , B_2 , and B_3 in Table S1). The details of the evaluation method of the ESI descriptors were explained in [5-7].

3. The Method of Machine Learning

Regression modeling was performed by using the XGBR method, which is one of the supervised machine learning method. The reference data for the XGB modeling was taken from Cronin *et al.* [4, 19] in which all the $pIGC_{50}$ values of 250 phenol derivatives are tabulated together with their assigned MOA classes. Among the 250 phenol compounds, we excluded 5 molecules in our reference set: 4 iodine-containing molecule and 4-methylcyanophenol for convenience in computation. In addition, we changed the assignment of the MOA class of tetrachlorocatechol (from PE to PRC) following the original MOA classification scheme suggested by Schultz *et al.* [3]. Thus, the reference dataset in this present study contains 169, 26, 27, 18, and 5 phenol derivatives in the PN, PE, SE, WARU, and PRC classes, respectively.

This choice of the XGBR method was due to our preliminary study on application of the multiple linear regression (MLR) method. In this implementation with the ESI descriptor, the R^2 values for the training and test sets for the 245 phenol derivatives were calculated to be 0.48 and 0.66, respectively. When the descriptor set in Cronin *et al.* [4] was used, the R^2 values were 0.64 and 0.60, respectively. In Ref. [4], the R^2 value for the comparable model given in eq. (8) therein was 0.65 for 200 phenol derivatives. Graphical comparison between the XGBR and MLR using the so-called y-y plot is made in Figure S1. It is evident that XGBR is much better than MLR.

In this work, the computer programs were developed in our laboratory by using Python 3.7.1 where the sci-kit learn package was applied for machine learning. The XGBR method requires tuning of the hyperparameters. Their optimization was repeatedly done for the 245 molecules and for each MOA class.

In the hyperparameter optimization, the Bayesian

optimization method was applied. The hyperparameters in XGBR, which are labeled as *eta*, *gamma*, *max_depth*, *min_child_weight*, *subsample*, and *colsample_bytree* [12, 13] were optimized in the following ranges: *eta* = [0,0.5]; *gamma* = [0,1]; *max_depth* = [1,5]; *min_child_weight* = [0,1]; *subsample* = [0,0.5]; and *colsample_bytree* = [0,0.5]. Their optimized values are tabulated in Table S4 in Supplementary Materials.

After the hyperparameter optimization, the regression model was established. In this step, we divided the dataset into the training and test sets with a ratio of 4:1. This data division was repeated 100 times in order to evaluate averaged performance through evaluation of the coefficient of determination (R^2) and to choose the best regression model. The best model was chosen as the one giving the largest R^2 value for the test set.

In order to know important descriptors describing characteristics of each MOA, we performed feature importance (*FI*) analysis. In order to select important descriptors, we took a stepwise approach: in one step, *FI* for all the descriptors were evaluated, and then the descriptor with the lowest score was removed. This step was repeated until all the remaining descriptors showed the *FI* value larger than 0.1.

4. Results and Discussion

We evaluated all the ESI descriptors for all the 245 molecules by using the DFT method. The calculated values are summarized in Table S2 and S3 in the Supplementary Materials. Although it would be desired to carefully investigate the accuracy of the calculated values, we herein compare only some descriptor values: the IP_v and EA_v of phenol were calculated to be 8.458 and -2.182 eV, which are very close to the experimental values of 8.470 and -2.190 eV, respectively [21, 22]. The calculated solvation energy of 4-tert-butylphenol was calculated to be 0.184 eV. This is comparable with its experimental value (0.140 eV) [23]. These results imply that the calculated value would be reasonable.

The obtained ESI descriptors were used to perform the regression modeling of XGBR for molecules belonging to all classes then followed in each MOA class. The results for the best model in the 100-times data divisions are shown in Figure 1. Table 1 summarizes the evaluated R^2 values for the training and test sets for the best model among the 100-times trials. For the former set, the R^2 values are 0.98 except for PRC, while, for the test set, they are larger than 0.76. This result indicates that the modeling would be satisfactory in its accuracy for a practical purpose.

When the result for all MOA classes is compared with the one for each MOA class, we see that the R^2 values for the PN, SE, and WARU classes are improved, and the value for PE class is nearly the same. Therefore, we may

conclude that the regression modeling on each MOA class is more preferable to that performed for the all-molecule set. This seems reasonable since dominant factors would be different depending on an MOA class.

Figure 1. Correlations between the predicted (horizontal axis) and experimental (vertical axis) $pIGC_{50}$ values. The former was calculated by using the XGBR method with the ESI descriptor set.

Table 1. The R^2 values for the best models in XGBR using the ESI and Cronin's descriptor sets.

MOA class	ESI Descriptor		Cronin's Descriptor	
	training set	test set	training set	test set
All classes	0.98	0.77	0.85	0.80
PN	0.99	0.85	0.92	0.92
PE	0.98	0.76	0.98	0.80
SE	0.99	0.80	0.99	0.82
WARU	0.99	0.98	0.99	0.92
PRC	0.85	—	0.93	—

It is noted here that the above-mentioned XGBR

results show the one obtained for the best regression model in each MOA class. For the averaged R^2 values, the quality of the modeling varies strongly depending on the data division. For the MOA classes with a smaller number of molecules such as PE, SE, and WARU, their R^2 values indicate that the modelings are disastrous: they are -0.33, -0.08, and -0.12, respectively, where the number of molecules included therein are 26, 27, and 18. The standard deviation (SD) for these MOA classes shows that the R^2 values appreciably fluctuate in the 100-times data division.

This is in sharp contrast to the PN class in which 169 molecules are included. In this case, the averaged R^2 value is 0.76, which is close to the one for the best model. The SD value (0.0597) is much smaller than the other MOA classes, indicating that the fluctuation of the R^2 value is much suppressed. This trend implies that the difference in the R^2 values between Tables 1 and 2 would be due to the number of molecules in the MOA classes. Namely, it seems reasonable to conclude that, without increasing the experimental data for the supervised machine learning, such an undesired characteristic in modeling would not be overcome. For practical purposes, the choice of the best model after many trials, which is 100-times trials in the present study, seems reasonable although the number of trials becomes arbitrary.

It is interesting to investigate characteristics of the ESI descriptor set applied in the present study. For comparison, we performed XGBR with Cronin's descriptor set. The results are summarized in Figure 2 and Table 1. Although the latter set consists of only 7 descriptors, the accuracy is generally comparable or even better than the ESI descriptor set. Their descriptor set include distribution coefficient ($\log D$), the LUMO level (represented as *LUMO* hereafter), *MW*, negatively-charged molecular surface area (P_{NEG}), sum of absolute charges on nitrogen and oxygen atoms in a molecule (*ABSQon*), largest positive charge on a hydrogen atom (*MaxHp*), and electrotopological state index for the hydroxy group (*SsOH*). These descriptors were selected among the 108 descriptors in [4]. The selection process would be nearly equivalent to the optimization of their descriptor set. Therefore, it can be reasonably concluded that the ESI descriptor set is nearly equivalent to the empirically optimized descriptor set by Cronin *et al.* in its regression modeling accuracy.

Obvious differences between the ESI and Cronin's descriptor sets in their derivation are that the former mainly includes energy-related quantities of a molecule, which corresponds to its overall (global) properties and essentially experimentally observable. In contrast, the latter descriptor set consists of both global ($\log D$, *LUMO*, *MW*, and P_{NEG}) and local characteristics (*ABSQon*, *MaxHp*, and *SsOH*). Interestingly, the local characteristics can never be measured experimentally and are relative values without a clear physical definition. Such a detailed comparison of characteristics among various descriptor

sets would be important to know and/or understand its appropriate choice.

Table 2. The averaged R^2 and standard deviation (SD) values for the training and test sets in XGBR with the ESI descriptor set.

MOA class	Training set		Test set	
	R^2	SD	R^2	SD
All classes	0.98	0.0021	0.52	0.1194
PN	1.00	0.0004	0.76	0.0597
PE	0.99	0.0056	-0.33	0.6948
SE	0.99	0.0023	-0.08	0.7555
WARU	1.00	0.0016	-0.12	2.2906
PRC	0.85	0.0436	—	—

Figure 2. Correlations between the predicted (horizontal axis) and experimental (vertical axis) $pIGC_{50}$ values. The former was calculated by using the XGBR method with Cronin's descriptor set [4].

In order to highlight dominant characteristics in each MOA class, we have performed *FI* analysis. This analysis was done for the two XGBR model using the ESI

and Cronin's descriptor sets. The results are shown in Figures 3 and 4, respectively. As was mentioned in Sec. 3, it was sequentially performed through systematic elimination of unimportant descriptors until all the descriptors would have *FI* larger than 0.1. The *FI* for all ESI descriptors in the first step of this procedure is shown in Figure S2 in the Supplementary Materials.

Figure 3. The *feature importance* score in XGBR for the ESI descriptor after the feature selection.

For the PN class, the important descriptors are *MW* and ΔE_{solv} . The importance of these descriptors implies that transport phenomena would be important in its MOA. Castillo-Garit, *et al.* [24] suggests that poisoning should occur through dissolving process into a cell membrane and accumulation in lipid tissue [25].

The result is essentially similar to Cronin's one [4]: $\log D$, which can be related to ΔE_{solv} in the ESI descriptor set, was shown to be one of the important descriptors. Even in the present XGBR modeling shown in Figure 4, $\log D$ is the most important descriptor. The secondary important descriptor in Figure 4(b) is molecular mass (weight) represented as *M* (*MW*). It is interesting that the top-two important descriptors are common in XGBR with the ESI and Cronin's descriptor set.

Figures 3 and 4 indicate that the important descriptors for each MOA class are different. Thus, regression modeling should be done separately after classification of the MOA classes. The important descriptors suggested

therein reflect their characteristic features and would be good measures for MOA classification of a given molecule.

Figure 4. The *feature importance* score in XGBR using Cronin's descriptor set.

Since similarly successful regression modeling has been done with the ESI and Cronin's descriptor sets, it seems reasonable to conclude that both sets have similar characteristics and should have similar implication even in the feature importance analysis. At present, mutual relationship between the two descriptor sets has not been known and will be investigated in future studies.

5. Summary and Conclusions

In this study, the XGB method with the ESI descriptor set has been applied to regression modeling for the *pIGC*₅₀ values of the 245 phenol derivatives [4] representing their influence on growth of *Tetrahymena pyriformis*. Through statistical analyses of the classification, the modeling has been shown successful, justifying the present computational approach. The accuracy of the modeling has been improved when it was applied to each MOA class. A comparison with MLR revealed that the nonlinear modeling is more effective to increase quantitative accuracy in the modeling.

The present XGB modeling with the ESI descriptor set has shown to yield results similar to the one obtained with the optimal descriptor set suggested by Cronin et al. [4]. This means that the ESI descriptor set is enough accurate to describe the toxicity of the phenol derivatives. It is also noted that the ESI descriptor has been successful in regression modeling for different types of biological activities [6, 7]. This implies that the ESI descriptor has general applicability in various QSAR and quantitative structure-property (QSPR) studies. This would be its unique characteristic and could be advantage over any kinds of descriptors.

The feature importance analysis has revealed that characteristic features in each MOA class are different. This suggests that, in general, regression modeling for toxicity should be done separately for each MOA. Therefore, for a given set of new molecules, it is recommended to perform machine-learning-based classification reflecting MOA before regression modeling for accurate and informative QSAR/QSPR studies.

Acknowledgments

We are grateful to Kazufumi Ohkawa, Takafumi Inoue, and Toshihiro Ideo for their assistance in a part of the descriptor calculation. This work was partially supported by Grant-in-Aid for Scientific Research on Innovative Areas (2601): π -System Figuration (26102015 to MS) from MEXT, Japan.

References and Notes

- [1] M. Nichkova, R. Galve, M. P. Marco, *Chem. Res. Toxicol.*, **15**, 1371-1379 (2002).
- [2] T. W. Schultz, *Chem. Res. Toxicol.*, **12**, 1262-1267 (1999).
- [3] T. W. Schultz, G. D. Sinks, M. T. D. Cronin, *Quant. Struct. Act. Relat. In Environ. Sci.*, **23**, 330-342 (1997).
- [4] M. T. D. Cronin, A. O. Aptula, J. C. Duffy, T. I. Netzeva, P. H. Rowe, I. V. Valkova, T. W. Schultz, *Chemosphere.*, **49**, 1201-1221 (2002).
- [5] M. Sugimoto, A. B. Manggara, K. Yoshida, T. Inoue, T. Ideo, *Mol. Inf.*, **39**, 1900121 (2020).
- [6] M. Sugimoto, T. Ideo, A. B. Manggara, K. Yoshida, T. Inoue, *J. Comput. Aid. Chem.*, **20**, 65-75 (2019).
- [7] T. Ideo, K. Yoshida, M. Sugimoto, *Chem. Lett.*, **50**, 849-852 (2021).
- [8] R. Garg, A. Kurup, C. Hansch, *Crit. Rev. Toxicol.*, **31(2)**, 223-245 (2001).
- [9] J. B. O. Mitchell, *WIREs Comput. Mol. Sci.*, **4**, 468-481 (2014).
- [10] D. X. Jiang, Y. Li, J. Li, G. X. Wang, *Int. J. Environ. Res.*, **5(4)**, 923-938 (2011).
- [11] K. Dieguez-Santana, H. Pham-The, P. J. Villegas-Aguilar, H. Le-Thi-Thu, J. A. Castillo-Garit, G. M. Casanola-Martin, *Chemosphere*, **165**, 434-441 (2016).
- [12] T. Chen, C. Guestin, in *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-794 (2016).
- [13] R. P. Sheridan, W. M. Wang, A. Liaw, J. Ma, E. M. Gifford, *J. Chem. Inf. Model.*, **56**, 2353-2360 (2016).
- [14] X. Ji, W. Tong, Z. Liu, T. Shi, *Front. in. Gen.*, **10**, 600 (2019).
- [15] Z. Wu, M. Zhu, Y. Kang, E. L. Leung, T. Lei, C. Leu, C. Shen, D. Jiang, Z. Wang, D. Cao, T. Hou, (In press), *Brief. Bioinf.*, (2020). doi:10.1093/bib/bbaa321.
- [16] D. Yu, Z. Liu, C. Su, Y. Han, X. Duan, R. Zhang, X. Liu, Y. Yang, S. Xu, *Thor. Can.*, **11**, 95-102 (2020).
- [17] M. J. Frisch et al., *Gaussian 16, Revision A.03*, Gaussian, Inc., Wallingford CT, (2016).
- [18] Mathematica Ver. 11.0, Wolfram Research Inc., Champaign, Illinois, (2016).
- [19] <http://qsar.db.org/repository/compounds/10967/30?property=MOA>
- [20] I. A. Miniakhmetov, S. A. Semenov, *Russ. J. Inorg. Chem.*, **56(4)**, 597-599 (2011).
- [21] D. R. Lide, *Handbook of Chemistry and Physics*, CRC Press Inc., USA, (1995).
- [22] G. H. Parsons, C. H. Rochester, A. Rostron, P. C. Sykes, *J. Chem. Soc. Perkin Trans.*, **2**, 136-138 (1972).
- [23] J. A. Castillo-Garit, G. M. Casanola-Martin, S. J. Barigye, H. Pham-The, F. Torrens, A. Torreblanca, *SAR QSAR Environ. Res.*, **28(9)**, 735-747 (2017).
- [24] G. Schuurmann, A. O. Aptula, R. Kuhne, R. U. Ebert, *Chem. Res. Toxicol.*, **16**, 974-987 (2003).

Supplementary Materials

Table S1. Descriptor set used in the present study. “S₀”, “T₁”, “C”, and “A” denote the spin-singlet ground, lowest spin-triplet, ionized (cationized), and anionized (electron-attached) states, respectively. **R_X** (X = T₁, C, A) indicates the geometry optimized in the X state. Otherwise, the values were evaluated at the S₀-state optimized geometry.

no.	descriptor	symbol
1	Similarity in UV/vis spectra (ρ_{UV})	$S(\rho_{UV})$
2	Similarity in density of electronically excited states (ρ_{ex})	$S(\rho_{ex})$
3	Similarity in density of molecular orbitals (ρ_{orb})	$S(\rho_{orb})$
4	Similarity in IR spectra (ρ_{IR})	$S(\rho_{IR})$
5	Similarity in density of vibrational states (ρ_{vib})	$S(\rho_{vib})$
6	Vertical ionization potential from S ₀	IP_v
7	Vertical electron affinity from S ₀	EA_v
8	Energy of vertical transition from S ₀ to T ₁	ΔST_v
9	Structural relaxation energy in C after vertical ionization	$\lambda_{S_0 \rightarrow C}$
10	Structural relaxation energy in S ₀ after vertical de-ionization at R_C	$\lambda_{C \rightarrow S_0}$
11	Structural relaxation energy in A after vertical anionization (electron attachment)	$\lambda_{S_0 \rightarrow A}$
12	Structural relaxation energy in S ₀ after vertical de-anionization at R_A	$\lambda_{A \rightarrow S_0}$
13	Structural relaxation energy in T ₁ after vertical transition from S ₀	$\lambda_{S_0 \rightarrow T_1}$
14	Structural relaxation energy in S ₀ after vertical transition from T ₁ at R_{T1}	$\lambda_{T_1 \rightarrow S_0}$
15	Dipole moment	μ
16	Solvation energy in aqueous solution at the optimized geometry in S ₀ in vacuum	ΔE_{solv}
17	Molecular volume	V_{mol}
18	Length of the longest edge of the cuboid box enclosing a molecule	B_1
19	Length of the second-longest edge of the cuboid box enclosing a molecule	B_2
20	Length of the shortest edge of the cuboid box enclosing a molecule	B_3
21	Molecular mass	M

Table S2. Summary of MOA and the values of $pIGC_{50}$, $S(\rho_{UV})$, $S(\rho_{ex})$, $S(\rho_{ab})$, $S(\rho_R)$, $S(\rho_{ib})$, IP_v , EA_v , ΔST_v , $\lambda_{80 \rightarrow C}$, and $\lambda_{C \rightarrow 80}$ evaluated for phenol derivatives (Compound 1-40)

compound			descriptors										
label	name	MOA	$pIGC_{50}$	$S(\rho_{UV})$	$S(\rho_{ex})$	$S(\rho_{ab})$	$S(\rho_R)$	$S(\rho_{ib})$	IP_v (eV)	$E A_v$ (eV)	ΔST_v (eV)	$\lambda_{80 \rightarrow C}$ (eV)	$\lambda_{C \rightarrow 80}$ (eV)
1	Nonyl-4-hydroxybenzoate	PN	2.630	16.290	48.080	65.077	19.397	40.459	8.581	-1.042	4.070	0.253	0.259
2	2-Ethylhexyl-4'-hydroxybenzoate	PN	2.510	16.155	54.271	66.347	23.900	42.995	8.582	-0.961	4.060	0.260	
3	4-Nonylphenol	PN	2.470	7.910	31.419	62.966	15.054	37.141	8.007	-2.042	4.527	0.250	0.257
4	2,3,5-Trichlorophenol	PN	2.370	67.778	71.858	82.265	66.320	66.547	9.021	-0.877	4.592	0.237	0.241
5	4-(4-Bromophenyl)phenol	PN	2.310	17.247	65.795	65.755	17.184	52.180	7.827	-0.860	4.124	0.268	0.251
6	4-tert-Octylphenol	PN	2.100	4.413	45.707	64.505	17.336	38.250	7.919	-1.869	4.487	0.260	
7	2,4,5-Trichlorophenol	PN	2.100	55.331	74.685	81.907	25.839	65.890	8.789	-0.866	4.404	0.253	0.257
8	2,4,6-Tribromophenol	PN	2.030	91.210	73.581	60.212	14.895	50.572	8.621	-0.705	4.332	0.228	0.233
9	4-Heptyloxyphenol	PN	2.030	7.663	36.390	61.651	11.242	38.794	7.603	-2.020	4.056	0.335	0.346
10	4-Chloro-2-isopropyl-5-methylphenol	PN	1.850	6.630	53.159	66.457	28.574	42.260	8.033	-1.738	4.568	0.251	0.252
11	2,6-Di-tert-butyl-4-methylphenol	PN	1.800	10.572	35.281	63.509	25.393	35.520	7.602	-1.958	4.287	0.260	0.259
12	4-Hexylresorcinol	PN	1.800	5.131	37.607	63.814	17.960	38.929	7.854	-2.229	4.606	0.255	0.260
13	4-Bromo-2,6-dichlorophenol	PN	1.780	82.089	75.718	75.526	23.448	54.764	8.722	-0.812	4.381	0.244	0.246
14	3,4-Dichlorophenol	PN	1.750	14.292	63.763	71.322	14.103	41.640	8.641	-1.274	4.473	0.262	0.265
15	3,5-Di-tert-butylphenol	PN	1.640	6.483	40.892	62.844	22.988	34.325	7.884	-1.999	4.727	0.239	0.242
16	3,5-Dibromosalicylaldehyde	PN	1.640	38.143	69.565	60.622	24.044	43.322	8.678	0.543	3.315	0.244	0.259
17	4-Hexyloxyphenol	PN	1.640	9.871	28.312	62.342	10.557	42.295	8.085	-1.958	4.428	0.903	0.257
18	3,5-Dichlorophenol	PN	1.570	21.708	66.314	70.487	36.615	52.075	8.958	-1.229	4.679	0.253	0.258
19	4-Cyclohexylphenol	PN	1.560	7.641	40.668	60.514	16.989	40.718	7.974	-2.006	4.526	0.243	0.245
20	3,5-Dichlorosalicylaldehyde	PN	1.550	34.721	63.675	73.640	23.349	39.060	8.814	0.503	3.320	0.275	0.289
21	Benzyl-4-hydroxybenzoate	PN	1.550	13.299	80.689	65.504	32.133	44.582	8.534	-0.943	4.055	0.218	0.253
22	Isoamyl-4-hydroxybenzoate	PN	1.480	16.765	49.000	66.806	22.347	42.792	8.589	-1.006	4.063	0.253	0.259
23	2-Hydroxy-4-methoxybenzophenone	PN	1.420	28.542	68.900	65.414	23.127	51.770	8.135	0.042	3.476	0.284	0.307
24	2,4,6-Trichlorophenol	PN	1.410	63.126	71.469	78.590	27.741	56.140	8.799	-0.840	4.383	0.262	0.263
25	2,4-Dibromophenol	PN	1.400	36.871	66.717	61.499	15.519	46.001	8.529	-1.127	4.401	0.235	0.237
26	4-Phenylphenol	PN	1.390	17.124	76.184	63.286	17.557	43.039	7.770	-1.251	4.138	0.266	0.248
27	n-Pentyloxyphenol	PN	1.360	7.381	35.413	62.362	11.776	41.125	7.622	-2.002	4.056	0.336	0.345
28	3-Phenylphenol	PN	1.350	43.856	79.792	61.097	13.672	37.461	8.085	-1.062	4.229	0.204	0.196
29	Butyl-4-hydroxybenzoate	PN	1.330	17.582	47.960	66.184	17.831	45.260	8.599	-1.011	4.066	0.251	0.256
30	2-Hydroxydiphenylmethane	PN	1.310	14.466	77.198	62.192	31.357	46.649	8.098	-1.686	4.447	0.229	0.264
31	5-Pentylresorcinol	PN	1.310	7.727	39.742	64.672	12.735	43.689	8.060	-2.153	4.758	0.271	0.279
32	2-Tert-butyl-4-methylphenol	PN	1.300	2.617	35.465	65.359	18.076	41.308	7.890	-2.026	4.569	0.259	0.253
33	4-Cyclopentylphenol	PN	1.290	6.403	39.880	62.771	19.194	38.914	7.958	-2.024	4.522	0.242	0.243
34	2-Tert-butylphenol	PN	1.290	1.982	34.089	63.254	30.968	45.850	8.168	-2.064	4.737	0.242	0.241
35	4-Bromo-6-chloro-2-cresol	PN	1.280	22.477	51.180	66.045	27.898	49.671	8.409	-1.290	4.489	0.244	0.244
36	2,3-Dichlorophenol	PN	1.280	19.346	52.487	73.224	45.486	47.723	8.824	-1.298	4.610	0.235	0.235
37	4-Bromo-3,5-dimethylphenol	PN	1.270	6.706	46.441	59.119	19.955	41.400	8.111	-1.803	4.654	0.244	0.246
38	2-Hydroxybenzophenone	PN	1.230	21.980	73.937	63.544	14.540	38.161	8.375	0.202	3.451	0.324	0.393
39	4-Tert-Pentylphenol	PN	1.230	6.841	37.025	63.758	15.545	40.649	7.999	-2.004	4.550	0.249	0.257
40	4-Chloro-3,5-dimethylphenol	PN	1.200	3.899	47.618	63.025	16.275	37.139	8.157	-1.836	4.640	0.262	0.266

Table S2. (Continued: Compound 41-83)

label	compound name	MOA	pIC_{50}	descriptors									
				$S(\rho_{IV})$	$S(\rho_{V})$	$S(\rho_{AB})$	$S(\rho_B)$	$S(\rho_{AB})$	IP_s (eV)	E_A (eV)	ΔST_s (eV)	λ_{SO-C} (eV)	λ_{C-SO} (eV)
41	4-Bromo-2,6-dimethylphenol	PN	1.170	5.606	41.710	61.931	25.684	39.939	8.049	-1.763	4.572	0.240	0.240
42	6-Tert-butyl-2,4-dimethylphenol	PN	1.160	5.847	33.680	64.645	24.870	36.227	7.727	-2.061	4.597	0.255	0.253
43	3-Bromophenol	PN	1.150	10.872	48.405	59.767	31.550	53.705	8.633	-1.599	4.403	0.231	0.235
44	2,5-Dichlorophenol	PN	1.130	11.654	48.104	73.165	30.450	58.722	8.832	-1.241	4.397	0.242	0.244
45	3-Chloro-4-fluorophenol	PN	1.130	4.561	32.380	64.572	13.864	40.821	8.643	-1.383	4.345	0.291	0.297
46	2-Phenylphenol	PN	1.090	36.791	79.657	61.907	28.442	43.254	8.024	-1.057	4.215	0.228	0.227
47	4-Chloro-3-ethylphenol	PN	1.080	2.182	52.832	64.923	11.754	38.362	8.251	-1.766	4.581	0.261	0.263
48	Benzyl-4-hydroxyphenyl ketone	PN	1.070	21.210	83.783	65.758	24.125	47.674	8.492	-0.589	3.882	0.142	0.148
49	2,4,6-Tribromoresorcinol	PN	1.060	92.382	85.223	57.226	26.047	34.779	8.414	-0.936	4.729	0.227	0.234
50	2,4-Dichlorophenol	PN	1.040	6.130	61.143	73.785	29.810	44.139	8.657	-1.228	4.430	0.262	0.262
51	4-Benzoyloxyphenol	PN	1.040	9.602	85.280	62.905	16.018	42.997	7.721	-1.631	4.116	0.413	0.392
52	4-Hydroxybenzophenone	PN	1.020	18.818	72.769	64.325	21.920	46.489	8.519	-0.202	3.834	0.285	0.342
53	3-Hydroxydiphenylamine	PN	1.010	47.094	85.314	65.189	17.594	50.407	7.102	-1.601	4.167	0.084	0.084
54	4-Sec-butylphenol	PN	0.980	7.678	34.489	62.646	20.320	44.203	8.026	-2.023	4.526	0.261	0.266
55	4,6-Dichlororesorcinol	PN	0.970	19.899	49.815	71.287	34.715	48.991	8.488	-1.361	4.450	0.260	0.265
56	2,2',4,4'-Tetrahydroxybenzophenone	PN	0.960	20.774	67.627	67.786	28.158	49.290	8.033	-0.207	3.509	0.321	0.377
57	3,4,5-Trimethylphenol	PN	0.930	3.548	42.974	59.168	18.634	43.263	7.826	-2.247	4.649	0.244	0.246
58	4-Tert-butylphenol	PN	0.910	6.697	36.829	63.300	16.165	40.201	8.034	-2.016	4.551	0.246	0.247
59	2,2'-Biphenol	PN	0.880	42.073	86.091	63.956	33.568	47.436	7.893	-1.138	4.227	0.321	0.310
60	3-Chlorophenol	PN	0.870	3.556	43.184	63.223	26.409	42.962	8.690	-1.679	4.425	0.244	0.245
61	4-Chloro-3-methylphenol	PN	0.800	2.374	52.663	62.408	12.176	43.045	8.314	-1.758	4.569	0.262	0.265
62	2-Isopropylphenol	PN	0.800	2.411	35.568	61.111	32.300	46.144	8.165	-2.128	4.730	0.239	0.239
63	2,6-Dichloro-4-fluorophenol	PN	0.800	17.489	48.235	71.857	26.104	55.562	8.803	-0.957	4.272	0.294	0.296
64	2-Methyl-3-nitrophenol	PN	0.780	34.125	60.477	65.230	31.353	37.893	8.963	0.066	3.781	0.269	0.252
65	3-Chloro-5-methoxyphenol	PN	0.760	22.245	46.036	63.315	30.865	55.089	8.381	-1.852	4.733	0.267	0.281
66	2-Methoxy-4-propenylphenol	PN	0.750	64.240	54.637	65.540	27.350	46.893	7.343	-1.457	3.602	0.281	0.260
67	3-Tert-butylphenol	PN	0.730	1.964	34.735	63.901	11.705	36.916	8.162	-2.082	4.731	0.236	0.238
68	2,6-Dichlorophenol	PN	0.730	21.057	47.884	70.179	40.621	42.937	8.804	-1.280	4.610	0.248	0.247
69	2-Hydroxy-4,5-dimethylacetophenone	PN	0.710	62.669	48.975	61.053	12.715	38.689	8.096	-0.365	3.426	0.266	0.280
70	4-Butoxyphenol	PN	0.700	9.710	29.241	60.366	11.126	40.095	8.100	-1.962	4.431	0.905	0.256
71	4-Chloro-2-methylphenol	PN	0.700	2.127	53.613	65.711	17.068	41.002	8.279	-1.752	4.535	0.258	0.259
72	2-Chloro-4,5-dimethylphenol	PN	0.690	6.532	35.320	64.514	39.527	42.286	8.176	-1.757	4.557	0.240	0.240
73	4-Bromophenol	PN	0.680	3.212	49.740	56.995	16.824	40.950	8.399	-1.622	4.494	0.242	0.243
74	4-Propylphenol	PN	0.640	7.383	31.248	59.601	19.821	43.411	8.055	-2.063	4.529	0.248	0.254
75	5-Bromovanillin	PN	0.620	56.134	74.523	61.999	22.793	51.170	8.439	-0.286	3.520	0.276	0.293
76	α,α -Trifluoro-4-cresol	PN	0.620	5.163	32.063	59.133	17.140	50.414	8.963	-1.605	4.791	0.272	0.277
77	Methyl-4-methoxysalicylate	PN	0.620	51.041	46.343	61.081	33.178	52.589	8.225	-0.886	3.905	0.287	0.309
78	3-Isopropylphenol	PN	0.610	2.212	35.193	60.218	26.952	43.071	8.192	-2.100	4.726	0.237	0.240
79	3-Methyl-2-nitrophenol	PN	0.610	33.937	47.784	62.766	39.390	48.994	8.884	-0.472	3.399	0.250	0.262
80	2-Bromo-4-methylphenol	PN	0.600	12.672	44.457	61.151	26.641	37.917	8.288	-1.603	4.307	0.240	0.240
81	2,4-Difluorophenol	PN	0.600	5.965	24.402	59.096	10.593	41.841	8.655	-1.745	4.456	0.318	0.320
82	5-Methyl-2-nitrophenol	PN	0.590	43.856	48.118	62.978	27.532	52.759	9.009	0.429	3.496	0.247	0.259
83	Ethyl-4-hydroxybenzoate	PN	0.570	18.704	49.874	64.379	26.428	41.691	8.618	-1.030	4.063	0.252	0.257

Table S2. (Continued: Compound 84-125)

compound				descriptors											
label	name	MOA	pIC ₅₀	S(ρ _{UV})	S(ρ _{OH})	S(ρ _{NH})	S(ρ _{OH})	S(ρ _{NH})	S(ρ _{OH})	IP _v (eV)	EA _v (eV)	ΔST _v (eV)	λ _{80-c} (eV)	λ _{c-80} (eV)	
84	4-Chlorophenol	PN	0.550	4.931	53.102	61.230	14.664	35.822	8.477	-1.668	4.492	0.263	0.265	0.265	
85	2-Hydroxy-4-methoxyacetophenone	PN	0.550	65.794	52.495	63.590	21.594	47.993	8.406	-0.454	3.697	0.260	0.276	0.276	
86	4-Cyanophenol	PN	0.520	5.157	45.580	64.300	21.301	35.291	8.975	-0.972	4.089	0.201	0.201	0.201	
87	4-Propyloxyphenol	PN	0.520	7.701	37.818	60.555	12.236	37.439	7.648	-2.037	4.205	0.335	0.344	0.344	
88	Ethyl-3-hydroxybenzoate	PN	0.480	55.172	48.179	64.886	21.702	40.327	8.581	-0.746	3.921	0.269	0.269	0.269	
89	2,6-Difluorophenol	PN	0.470	4.179	16.723	57.850	12.900	29.542	8.834	-2.099	4.966	0.294	0.295	0.295	
90	4-Isopropylphenol	PN	0.470	7.458	32.000	59.494	19.336	39.220	8.061	-2.041	4.531	0.244	0.245	0.245	
91	4-Allyl-2-methoxyphenol	PN	0.420	13.463	47.890	64.377	32.564	44.974	7.689	-2.130	4.495	0.305	0.321	0.321	
92	2-Hydroxybenzaldehyde	PN	0.420	70.306	45.120	58.246	9.913	29.186	8.730	-0.189	3.500	0.251	0.267	0.267	
93	2-Chloro-5-methylphenol	PN	0.390	4.837	31.360	66.110	39.803	43.903	8.463	-1.710	4.663	0.250	0.239	0.239	
94	3-Methoxysalicylaldehyde	PN	0.380	76.693	47.288	65.104	21.246	43.163	8.306	-0.600	3.717	0.317	0.329	0.329	
95	Salicylhydroxamic acid	PN	0.380	51.272	52.956	63.865	16.342	45.536	8.635	-0.447	3.809	0.473	0.851	0.851	
96	3-Fluorophenol	PN	0.380	3.476	31.586	55.925	10.080	46.439	8.719	-2.051	4.545	0.249	0.253	0.253	
97	2,3,5-Trimethylphenol	PN	0.360	7.388	35.655	59.077	23.736	40.710	7.946	-2.081	4.690	0.263	0.237	0.237	
98	5-Bromo-2-hydroxybenzyl alcohol	PN	0.340	7.035	51.366	60.728	15.734	49.877	8.141	-1.506	4.226	0.332	0.322	0.322	
99	2-Bromophenol	PN	0.330	4.541	44.864	57.416	21.448	37.918	8.595	-1.628	4.627	0.229	0.228	0.228	
100	2-Allylphenol	PN	0.330	8.254	47.859	62.108	23.010	48.108	8.182	-1.755	4.664	0.310	0.288	0.288	
101	2-Hydroxy-5-methylacetophenone	PN	0.310	71.072	49.521	60.859	15.213	40.234	8.215	-0.308	3.375	0.272	0.287	0.287	
102	2,3,6-Trimethylphenol	PN	0.280	6.617	28.544	59.172	32.672	36.043	7.936	-2.182	4.735	0.234	0.233	0.233	
103	2,4,6-Trimethylphenol	PN	0.280	3.142	28.951	59.607	26.567	32.894	7.774	-2.220	4.609	0.242	0.242	0.242	
104	4-Hydroxybenzaldehyde	PN	0.270	40.908	50.611	59.951	23.832	37.308	8.821	-0.644	3.802	0.227	0.232	0.232	
105	Salicylaldoxime	PN	0.250	67.033	43.423	60.672	14.444	43.952	8.195	-0.756	3.538	0.233	0.234	0.234	
106	3-Ethylphenol	PN	0.230	2.460	31.554	60.099	21.008	44.582	8.245	-2.080	4.702	0.248	0.244	0.244	
107	4-Ethylphenol	PN	0.200	6.467	31.119	61.774	21.541	38.484	8.065	-2.102	4.321	0.248	0.249	0.249	
108	4-Hydroxy-2-methylacetophenone	PN	0.190	43.223	45.665	67.237	20.312	41.936	8.485	-0.854	3.889	0.266	0.303	0.303	
109	2-Fluorophenol	PN	0.190	4.624	27.112	55.271	15.480	33.398	8.642	-2.062	4.484	0.265	0.267	0.267	
110	Salicylhydrazide	PN	0.180	46.173	51.776	61.821	8.941	39.878	8.463	-0.708	3.852	0.471	0.764	0.764	
111	2-Chlorophenol	PN	0.180	2.021	31.989	60.309	30.765	36.904	8.655	-1.705	4.656	0.240	0.239	0.239	
112	Syringaldehyde	PN	0.170	28.731	54.649	64.563	25.363	52.337	7.921	-0.575	3.314	0.298	0.311	0.311	
113	2-Ethylphenol	PN	0.160	2.831	33.249	60.697	27.089	45.984	8.172	-2.240	4.750	0.233	0.234	0.234	
114	4-Chlororesorcinol	PN	0.130	4.810	41.903	64.830	32.184	41.719	8.346	-1.827	4.559	0.271	0.277	0.277	
115	4-Hydroxypropiophenone	PN	0.120	35.678	51.776	64.698	25.453	42.254	8.622	-0.768	3.894	0.235	0.238	0.238	
116	2,3-Dimethylphenol	PN	0.120	2.153	32.261	54.246	16.889	42.841	8.116	-2.143	4.717	0.235	0.234	0.234	
117	3,4-Dimethylphenol	PN	0.120	3.552	37.657	56.321	20.055	41.336	7.969	-2.186	4.599	0.239	0.242	0.242	
118	3,5-Dimethylphenol	PN	0.110	2.909	33.176	58.142	14.291	35.353	8.139	-2.129	4.742	0.268	0.246	0.246	
119	3-Hydroxybenzaldehyde	PN	0.090	55.625	54.715	60.686	14.476	35.235	8.833	-0.354	3.747	0.250	0.254	0.254	
120	2-Hydroxyacetophenone	PN	0.080	71.339	50.001	61.341	13.362	36.995	8.513	-0.303	3.536	0.272	0.295	0.295	
121	2,5-Dimethylphenol	PN	0.080	2.683	28.493	61.195	27.532	36.966	8.083	-2.165	4.713	0.252	0.231	0.231	
122	Methyl-4-hydroxybenzoate	PN	0.080	19.585	47.136	62.632	24.314	46.942	8.648	-1.053	4.065	0.247	0.252	0.252	
123	2,4-Dimethylphenol	PN	0.070	2.489	32.793	61.195	10.978	31.955	7.940	-2.162	4.560	0.242	0.242	0.242	
124	4-Hydroxypropiophenone	PN	0.050	35.491	57.336	64.464	27.664	47.934	8.621	-0.725	3.867	0.236	0.243	0.243	
125	5-Fluoro-2-hydroxyacetophenone	PN	0.040	53.268	45.833	61.837	18.890	39.958	8.509	-0.062	3.283	0.321	0.341	0.341	

Table S2. (Continued: Compound 126-167)

label	compound		MOA	pIC_{50}	descriptors									
	name				$S(\rho_{IV})$	$S(\rho_{NV})$	$S(\rho_{OH})$	$S(\rho_{OR})$	$S(\rho_{OH})$	IP_s (eV)	E_A (eV)	ΔST_s (eV)	λ_{SO-C} (eV)	λ_{C-SO} (eV)
126	2-Cyanophenol		PN	0.030	27.395	44.074	63.687	13.909	35.454	9.047	-0.636	4.047	0.212	0.213
127	4-Fluorophenol		PN	0.020	7.157	30.535	54.630	10.533	37.859	8.459	-1.832	4.353	0.294	0.297
128	4-Ethoxyphenol		PN	0.010	7.580	37.600	58.993	12.529	43.158	7.671	-2.043	4.105	0.334	0.344
129	3-Ethoxy-4-hydroxybenzaldehyde		PN	0.010	21.387	54.555	64.889	35.663	48.811	8.250	-0.627	3.545	0.299	0.315
130	4-Hydroxy-3-methoxybenzonitrile		PN	-0.030	68.266	47.650	67.561	29.903	49.614	8.469	-0.949	3.824	0.258	0.268
131	3-Methoxy-4-hydroxybenzaldehyde		PN	-0.030	35.918	49.938	63.663	27.942	52.536	8.285	-0.623	3.516	0.296	0.311
132	Methyl-3-hydroxybenzoate		PN	-0.050	54.462	47.301	65.458	20.460	40.059	8.612	-0.766	3.924	0.264	0.265
133	3-Cyanophenol		PN	-0.060	2.035	29.635	57.878	16.769	41.132	9.066	-0.632	3.983	0.221	0.223
134	3-Methylphenol		PN	-0.060	36.835	45.995	66.466	33.624	47.650	8.288	-2.102	4.684	0.264	0.239
135	2-Hydroxyethylsalicylate		PN	-0.080	2.232	34.176	54.769	8.988	35.682	8.604	-0.456	3.782	0.450	0.461
136	3,5-Dimethoxyphenol		PN	-0.090	15.265	32.728	58.104	19.483	42.187	8.003	-2.571	4.631	0.289	0.304
137	4-Hydroxy-3-methoxyacetophenone		PN	-0.120	38.163	53.740	65.455	32.498	53.413	8.147	-0.750	3.597	0.300	0.313
138	4-Methoxyphenol		PN	-0.140	27.500	51.115	63.582	25.953	49.396	7.735	-2.046	4.107	0.325	0.332
139	3-Hydroxy-4-methoxybenzaldehyde		PN	-0.140	7.472	36.411	58.039	17.645	40.990	8.307	-0.619	3.647	0.297	0.308
140	3-Acetamidophenol		PN	-0.160	63.806	44.527	62.856	29.583	48.375	8.023	-1.673	4.296	0.295	0.306
141	4-Hydroxy-3-methoxyphenethylalcohol		PN	-0.180	7.001	29.875	59.995	17.562	37.726	7.739	-2.192	4.503	0.299	0.306
142	4-Methylphenol		PN	-0.180	2.563	43.238	63.213	34.698	51.770	8.129	-2.132	4.533	0.244	0.246
143	Phenol		PN	-0.210	5.181	30.386	52.852	14.652	30.569	8.458	-2.182	4.710	0.233	0.234
144	Ethyl-4-hydroxy-3-methoxyphenylacetate		PN	-0.230	16.650	48.647	64.526	29.789	49.513	7.815	-1.822	4.211	0.295	0.304
145	2-Hydroxybenzamide		PN	-0.240	57.919	47.208	60.453	16.338	45.246	8.418	-0.817	3.837	0.466	0.725
146	2-Methylphenol		PN	-0.290	2.244	31.570	57.700	26.288	36.084	8.249	-2.220	4.730	0.232	0.232
147	3-Ethoxy-4-methoxyphenol		PN	-0.300	36.552	48.380	62.456	21.535	45.678	7.326	-2.320	4.241	0.351	0.363
148	4-Acetylphenol		PN	-0.300	5.523	40.748	64.233	23.319	47.412	8.662	-0.790	3.889	0.235	0.238
149	3-Methoxyphenol		PN	-0.330	2.473	35.289	58.094	13.482	53.206	8.144	-2.251	4.728	0.283	0.293
150	2-Ethoxyphenol		PN	-0.360	1.793	30.480	60.707	26.008	41.691	7.869	-2.367	4.597	0.309	0.321
151	3-Hydroxyacetophenone		PN	-0.380	68.021	61.712	63.869	30.617	36.033	8.673	-0.518	3.830	0.252	0.254
152	5-Methylresorcinol		PN	-0.390	4.223	39.286	58.907	11.081	43.436	8.138	-2.270	4.775	0.279	0.279
153	4-(4-Hydroxyphenyl)-2-butanone		PN	-0.500	6.526	45.342	64.890	23.310	40.099	8.033	-1.876	4.434	0.410	0.611
154	2-Methoxyphenol		PN	-0.510	66.366	51.888	63.958	9.953	39.538	7.934	-2.367	4.600	0.297	0.307
155	Salicylic acid		PN	-0.510	1.820	30.374	56.634	27.939	48.736	8.821	-0.461	3.805	0.324	0.361
156	2,4,6-Tris(dimethylaminomethyl)phenol		PN	-0.520	23.807	41.988	59.410	12.103	46.025	7.503	-1.727	4.378	0.518	0.646
157	2,6-Dimethoxyphenol		PN	-0.600	36.789	27.024	61.019	16.515	50.575	7.578	-2.511	4.184	0.314	0.323
158	Resorcinol		PN	-0.650	1.817	37.934	56.100	8.199	44.865	8.258	-2.308	4.708	0.263	0.269
159	4-Hydroxy-3-methoxybenzyl alcohol		PN	-0.700	4.988	40.407	61.431	36.981	44.355	7.862	-2.124	4.248	0.463	0.383
160	4-Hydroxybenzamide		PN	-0.780	8.449	47.702	66.775	13.316	47.751	8.642	-1.269	4.159	0.278	0.333
161	3-Hydroxybenzoic acid		PN	-0.810	55.622	49.913	63.506	20.917	41.096	8.717	-0.724	3.900	0.266	0.266
162	4-Hydroxy-3-methoxybenzyl amine		PN	-0.820	3.266	40.733	62.908	8.048	31.654	7.623	-2.319	4.263	0.326	0.328
163	4-Hydroxyphenethyl alcohol		PN	-0.830	6.950	40.445	59.737	22.388	43.999	8.180	-2.013	4.559	0.254	0.256
164	2-Hydroxybenzylalcohol		PN	-0.950	2.318	37.377	58.504	17.424	43.855	8.172	-1.997	4.382	0.357	0.343
165	3-Hydroxy-4-methoxybenzyl alcohol		PN	-0.990	4.510	41.542	60.381	33.090	49.735	7.875	-2.118	4.507	0.357	0.368
166	4-hydroxybenzoic acid		PN	-1.020	23.128	43.466	60.773	30.559	49.558	8.767	-1.017	4.049	0.248	0.251
167	3-Hydroxybenzyl alcohol		PN	-1.040	1.701	41.336	59.648	11.395	50.086	8.245	-2.109	4.663	0.254	0.255

Table S2. (Continued: Compound 168-208)

label	compound		MOA	$p(GC_{50})$	descriptors									
	name				$S(\rho_{IV})$	$S(\rho_{SV})$	$S(\rho_{RH})$	$S(\rho_{IR})$	$S(\rho_{th})$	IP_r (eV)	$E A_{SV}$ (eV)	ΔST_r (eV)	$\lambda_{80 \rightarrow C}$ (eV)	$\lambda_{C \rightarrow 80}$ (eV)
168	1,3,5-Trihydroxybenzene		PN	-1.260	3.483	41.425	53.058	20.727	38.212	8.340	-2.700	4.727	0.320	0.302
169	4-Hydroxyphenylacetic acid		PN	-1.500	4.972	42.754	63.756	22.625	42.819	8.325	-1.848	4.584	0.346	0.272
170	Methoxyhydroquinone		PE	2.210	2.872	31.231	58.660	32.839	54.764	7.527	-2.278	4.192	0.338	0.344
171	3,5-Di-tert-butylcatechol		PE	2.110	16.859	45.862	66.284	34.375	36.904	7.555	-2.075	4.543	0.287	0.289
172	Phenylhydroquinone		PE	2.000	21.008	77.862	62.985	30.387	42.629	7.564	-1.093	3.791	0.267	0.272
173	Methylhydroquinone		PE	1.860	3.668	38.913	58.849	15.975	38.849	7.686	-2.110	4.164	0.307	0.311
174	Bromohydroquinone		PE	1.680	17.706	45.835	61.438	27.606	40.584	8.054	-1.503	4.048	0.307	0.308
175	4-Amino-2,3-dimethylphenol		PE	1.440	31.618	41.744	59.102	19.319	37.820	7.189	-2.170	4.043	0.518	0.342
176	2,3-Dimethylhydroquinone		PE	1.410	12.150	37.646	53.875	21.348	34.230	7.578	-2.039	4.150	0.307	0.309
177	Trimethylhydroquinone		PE	1.340	20.349	40.996	60.674	16.670	36.966	7.411	-2.078	4.184	0.304	0.306
178	4-Amino-2-cresol		PE	1.310	3.957	41.119	59.973	13.612	42.361	7.238	-2.247	4.009	0.480	0.327
179	Chlorohydroquinone		PE	1.260	5.812	35.967	64.404	23.689	40.515	8.098	-1.578	4.120	0.313	0.315
180	4-Chlorocatechol		PE	1.060	3.738	41.615	60.437	19.207	46.623	8.255	-1.849	4.457	0.291	0.296
181	2-Aminophenol		PE	0.940	3.327	41.674	59.424	12.173	41.092	8.181	-2.101	4.524	0.662	0.456
182	6-Amino-2,4-dimethylphenol		PE	0.890	4.750	42.391	60.349	10.316	36.050	7.711	-2.134	4.418	0.481	0.470
183	1,2,3-Trihydroxybenzene		PE	0.850	6.385	22.558	57.841	18.121	37.152	8.070	-2.613	4.430	0.325	0.329
184	2-Amino-4-chlorophenol		PE	0.780	3.152	61.402	64.829	8.915	46.521	8.233	-1.602	4.318	0.528	0.462
185	Catechol		PE	0.750	2.068	33.953	54.370	22.446	40.883	8.102	-2.339	4.583	0.286	0.290
186	Hydroquinone		PE	0.470	8.684	30.998	53.404	16.292	30.702	7.870	-2.041	4.100	0.313	0.316
187	5-Amino-2-methoxyphenol		PE	0.450	27.421	40.477	61.626	18.484	49.633	7.200	-2.431	4.161	0.482	0.359
188	1,2,4-Trihydroxybenzene		PE	0.440	3.849	39.969	57.665	19.081	42.031	7.733	-2.273	4.235	0.334	0.340
189	4-Methylcatechol		PE	0.370	2.510	34.730	59.416	14.844	41.170	7.901	-2.252	4.486	0.309	0.296
190	2-Amino-4-tert-butylphenol		PE	0.370	4.731	44.378	65.386	11.278	40.061	7.799	-1.949	4.387	0.553	0.456
191	3-Methylcatechol		PE	0.280	2.714	33.083	58.115	11.979	32.917	7.949	-2.358	4.349	0.284	0.287
192	2,4-Diaminophenol		PE	0.130	45.403	52.893	61.266	15.356	51.017	7.187	-2.280	3.966	0.519	0.393
193	4-Aminophenol		PE	-0.080	64.581	56.650	63.017	22.255	39.149	7.385	-2.182	3.927	0.487	0.325
194	3-Aminophenol		PE	-0.520	25.322	59.567	63.914	11.083	41.822	7.758	-2.466	4.360	0.379	0.221
195	4-Acetamidophenol		PE	-0.820	3.516	42.186	63.926	33.418	49.442	7.701	-1.950	4.236	0.353	0.356
196	4-Chloro-2-nitrophenol		SE	2.050	52.600	59.659	68.387	30.362	49.648	9.111	0.814	3.336	0.264	0.273
197	4-Bromo-2-nitrophenol		SE	1.870	48.063	70.849	63.416	25.974	50.530	9.014	0.831	3.330	0.243	0.250
198	3-Methyl-4-nitrophenol		SE	1.730	33.143	62.619	65.453	26.427	49.347	9.131	-0.049	3.871	0.278	0.305
199	4-Nitro-3-(trifluoromethyl)phenol		SE	1.650	44.550	63.116	62.181	44.136	45.919	9.598	0.314	3.866	0.288	0.302
200	4-Chloro-6-nitro-3-cresol		SE	1.640	48.769	67.051	67.536	30.058	49.942	8.965	0.728	3.378	0.266	0.274
201	2-Chloro-4-nitrophenol		SE	1.590	45.734	71.784	67.596	31.243	42.309	9.415	0.302	3.849	0.228	0.232
202	4-Nitrophenol		SE	1.420	27.791	60.093	60.339	31.785	44.689	9.303	-0.010	3.880	0.220	0.224
203	4-Fluoro-2-nitrophenol		SE	1.380	28.016	47.278	60.238	24.353	43.743	9.146	0.709	3.237	0.296	0.308
204	2,6-Dibromo-4-nitrophenol		SE	1.360	41.230	71.633	62.867	21.133	41.233	9.328	0.589	3.794	0.211	0.221
205	4-Nitrocatechol		SE	1.170	32.270	59.055	60.225	25.338	51.615	8.884	0.046	3.582	0.286	0.296
206	2-Amino-4-chloro-5-nitrophenol		SE	1.170	40.274	82.585	70.031	28.914	56.337	8.739	0.075	3.649	0.440	0.431
207	5-Fluoro-2-nitrophenol		SE	1.130	38.144	48.703	58.867	18.276	58.767	9.455	0.573	3.664	0.258	0.275
208	2-Fluoro-4-nitrophenol		SE	1.070	32.249	59.173	57.077	26.471	48.500	9.447	0.215	3.805	0.256	0.262

Table S2. (Continued: Compound 209-245)

label	compound name	MOA	pIC_{50}	descriptors									
				$S(\rho_{UV})$	$S(\rho_{Kv})$	$S(\rho_{ab})$	$S(\rho_B)$	$S(\rho_{ab})$	IP_v (eV)	EA_v (eV)	ΔST_v (eV)	λ_{SO-c} (eV)	λ_{C-SO} (eV)
209	3-Fluoro-4-nitrophenol	SE	0.940	35.799	67.124	61.947	21.833	48.831	9.544	0.077	3.994	0.297	0.317
210	4-Amino-2-nitrophenol	SE	0.880	49.936	53.971	62.029	24.278	50.676	8.022	0.405	2.718	0.464	0.300
211	2-Bromo-2-hydroxy-5'-nitroacetamide	SE	0.870	46.316	66.726	66.261	21.998	50.268	9.551	0.535	3.889	0.848	0.428
212	2-Chloromethyl-4-nitrophenol	SE	0.750	59.534	76.824	68.062	29.934	48.902	9.353	0.247	3.839	0.280	0.264
213	4-Methyl-3-nitrophenol	SE	0.740	43.721	57.860	67.146	22.255	42.842	8.876	0.176	3.598	0.257	0.261
214	2-Nitrophenol	SE	0.670	38.032	49.185	59.429	26.333	45.197	9.161	0.490	3.489	0.246	0.263
215	2-Nitroresorcinol	SE	0.660	36.230	37.735	61.646	13.835	42.677	8.750	0.832	3.018	0.246	0.254
216	4-Nitrosophenol	SE	0.650	7.500	47.365	58.386	34.164	42.096	8.310	0.178	0.886	0.560	0.837
217	4-Hydroxy-3-nitrobenzaldehyde	SE	0.610	19.753	72.877	61.898	28.233	54.934	9.480	0.889	3.567	0.236	0.253
218	4-Methyl-2-nitrophenol	SE	0.570	52.941	44.061	63.839	24.882	49.445	8.823	0.465	3.312	0.251	0.261
219	3-Nitrophenol	SE	0.510	48.421	57.194	61.460	31.826	44.595	9.211	0.218	3.712	0.246	0.248
220	2-Amino-4-nitrophenol	SE	0.470	28.606	52.954	64.180	25.992	55.491	8.999	-0.031	3.745	0.761	0.482
221	5-Hydroxy-2-nitrobenzaldehyde	SE	0.330	38.175	79.327	63.775	26.359	47.601	9.514	0.685	3.607	0.254	0.269
222	3-Hydroxy-4-nitrobenzaldehyde	SE	0.270	59.875	63.320	65.222	20.781	54.225	9.433	1.198	3.192	0.274	0.294
223	2,3,4,5-Tetrachlorophenol	WARU	2.710	100.000	100.000	100.000	100.000	100.000	8.899	-0.578	4.378	0.251	0.256
224	Pentabromophenol	WARU	2.660	4.469	30.591	51.423	25.612	49.993	8.682	0.019	4.318	0.201	0.206
225	3,4,5,6-Tetrabromo-2-cresol	WARU	2.570	25.530	47.139	57.070	39.106	45.589	8.518	-0.379	4.433	0.216	0.221
226	2,3,5,6-Tetrachlorophenol	WARU	2.220	96.450	90.237	85.401	70.160	68.234	9.062	-0.575	4.240	0.234	0.243
227	Pentachlorophenol	WARU	2.050	58.067	75.330	87.071	77.341	73.204	8.955	-0.296	4.320	0.239	0.245
228	2,4-Dichloro-6-nitrophenol	WARU	1.750	23.583	73.680	73.492	16.424	48.966	9.174	1.093	3.297	0.264	0.276
229	4,6-Dinitro-2-methylphenol	WARU	1.720	32.091	74.942	63.950	26.378	45.238	9.615	1.076	3.621	0.262	0.290
230	Pentafluorophenol	WARU	1.640	2.143	19.732	49.871	11.473	42.287	9.297	-1.647	4.989	0.353	0.363
231	4-Bromo-2-fluoro-6-nitrophenol	WARU	1.620	51.510	72.688	60.237	29.139	40.821	9.198	0.492	3.711	0.261	0.267
232	2,6-Dinitro-4-cresol	WARU	1.230	24.546	77.963	64.751	27.702	42.159	9.461	1.181	3.366	0.268	0.287
233	2,3,5,6-Tetrafluorophenol	WARU	1.170	2.432	19.617	54.537	21.732	45.844	9.337	-1.742	4.590	0.330	0.357
234	2,4-Dinitrophenol	WARU	1.080	31.336	74.323	59.736	32.034	54.173	9.912	1.132	3.700	0.242	0.273
235	2,5-Dinitrophenol	WARU	0.950	63.975	80.834	62.491	30.778	55.374	9.803	1.528	3.248	0.258	0.284
236	2,6-Dichloro-4-nitrophenol	WARU	0.630	69.764	87.143	70.121	27.924	48.898	9.485	0.570	3.814	0.236	0.239
237	2,6-Dinitrophenol	WARU	0.540	17.842	79.231	60.760	30.453	55.171	9.833	1.221	3.569	0.263	0.288
238	2,3-Dinitrophenol	WARU	0.460	29.274	78.501	67.908	22.111	50.678	9.805	1.197	3.516	0.287	0.316
239	3,4-Dinitrophenol	WARU	0.270	43.159	77.100	65.579	39.195	50.888	9.830	0.835	3.721	0.278	0.296
240	2,4,6-Trinitrophenol	WARU	-0.160	59.548	59.863	65.339	25.114	47.045	10.508	1.820	3.757	0.258	0.299
241	Tetrachlorohydroquinone	PRC	2.110	92.093	78.173	80.785	24.043	60.963	8.482	-0.516	3.942	0.287	0.292
242	Tetrafluorohydroquinone	PRC	1.840	3.199	32.894	52.747	14.000	49.188	8.700	-1.848	4.702	0.371	0.381
243	Tetrachlorocatechol	PRC	1.700	82.831	88.984	85.279	30.025	66.833	8.630	-0.816	4.338	0.265	0.271
244	Tetramethylhydroquinone	PRC	1.280	46.144	42.250	60.101	41.722	42.445	7.254	-2.012	4.136	0.302	0.305
245	Tetrabromocatechol	PRC	0.980	16.148	52.053	55.432	23.186	49.793	8.435	-0.348	4.308	0.232	0.239

Table S3. Summary of the values of $\lambda_{S0 \rightarrow A}$, $\lambda_{A \rightarrow S0}$, $\lambda_{S0 \rightarrow T1}$, $\lambda_{T1 \rightarrow S0}$, μ , ΔE_{sol} , V_{mol} , B_1 , B_2 , B_3 , and M evaluated for phenol derivatives (Compound 1–40)

label	compound name	descriptors										
		$\lambda_{S0 \rightarrow A}$ (eV)	$\lambda_{A \rightarrow S0}$ (eV)	$\lambda_{S0 \rightarrow T1}$ (eV)	$\lambda_{T1 \rightarrow S0}$ (eV)	μ (D)	ΔE_{sol} (eV)	V_{mol} (Å ³)	B_1 (Å)	B_2 (Å)	B_3 (Å)	M (amu)
1	Nonyl-4-hydroxybenzoate	0.371	0.368	0.477	0.477	3.594	0.924	1318.033	21.099	7.658	4.172	264.173
2	2-Ethylhexyl-4'-hydroxybenzoate	0.404	0.412	0.474	0.474	3.586	1.119	1216.672	16.805	7.651	6.049	250.157
3	4-Nonylphenol	0.256	0.432	0.242	0.242	4.285	1.378	1203.039	18.675	6.310	6.171	220.183
4	2,3,5-Trichlorophenol	1.343	4.029	1.071	1.071	3.521	1.004	682.566	8.857	7.990	3.496	195.925
5	4-(4-Bromophenyl)phenol	0.325	0.281	0.867	0.867	3.258	2.738	877.068	13.741	6.446	4.183	247.984
6	4-tert-Octylphenol	0.229	0.253	0.725	0.725	3.762	1.342	1068.414	11.006	7.856	6.574	206.167
7	2,4,5-Trichlorophenol	1.321	3.989	1.035	1.035	3.369	2.042	663.558	8.870	7.320	3.490	195.925
8	2,4,6-Tribromophenol	1.366	2.562	0.966	0.966	3.365	1.493	730.704	9.290	8.763	3.694	327.773
9	4-Heptyloxyphenol	0.253	0.258	0.453	0.453	3.603	0.749	1039.244	17.765	7.036	4.176	208.146
10	4-Chloro-2-isopropyl-5-methylphenol	0.358	0.534	0.277	0.277	4.291	2.354	838.123	10.500	7.767	5.438	184.065
11	2,6-Di-tert-butyl-4-methylphenol	0.243	0.245	1.003	1.003	3.285	1.867	1168.775	9.992	9.779	6.715	220.183
12	4-Hexylresorcinol	0.369	0.567	0.265	0.265	4.341	1.469	989.326	14.992	6.650	5.890	194.131
13	4-Bromo-2,6-dichlorophenol	1.332	3.877	0.901	0.901	3.480	1.401	705.468	9.079	8.471	3.637	239.874
14	3,4-Dichlorophenol	0.981	2.745	0.908	0.908	3.565	3.779	589.764	8.362	7.302	3.484	161.964
15	3,5-Di-tert-butylphenol	0.240	0.244	0.227	0.227	4.500	1.345	1077.093	9.984	9.389	6.715	206.167
16	3,5-Dibromosalicylaldehyde	0.325	0.319	0.919	0.919	2.396	3.430	704.722	9.340	8.659	3.691	277.858
17	4-Hexyloxyphenol	0.268	0.410	0.914	0.914	3.515	1.785	980.585	16.139	6.670	5.006	194.131
18	3,5-Dichlorophenol	0.832	2.633	0.239	0.239	4.440	2.357	593.331	8.827	7.903	3.495	161.964
19	4-Cyclohexylphenol	0.233	0.238	0.236	0.236	4.291	1.411	883.269	12.384	5.925	5.884	176.120
20	3,5-Dichlorosalicylaldehyde	0.329	0.323	0.933	0.933	2.388	3.356	682.782	8.849	8.555	3.490	189.959
21	Benzyl-4-hydroxybenzoate	0.445	0.414	0.472	0.472	3.583	1.236	994.844	13.539	6.991	6.831	228.079
22	Isoamyl-4-hydroxybenzoate	0.381	0.378	0.474	0.474	3.589	0.980	990.497	14.414	6.290	5.815	208.110
23	2-Hydroxy-4-methoxybenzophenone	0.385	0.382	0.859	0.859	2.617	3.963	983.318	14.093	7.571	4.663	228.079
24	2,4,6-Trichlorophenol	1.349	3.886	0.951	0.951	3.432	1.425	675.594	8.816	8.488	3.493	195.925
25	2,4-Dibromophenol	1.534	3.276	1.009	1.009	3.392	0.940	640.620	9.342	7.331	3.683	249.863
26	4-Phenylphenol	0.356	0.301	0.804	0.804	3.333	1.375	776.730	12.075	6.267	3.243	170.073
27	n-Pentyloxyphenol	0.265	0.415	0.453	0.453	3.603	0.659	904.075	14.134	6.849	5.179	180.115
28	3-Phenylphenol	0.373	0.314	0.795	0.795	3.434	1.340	774.321	11.612	7.000	4.322	170.073
29	Butyl-4-hydroxybenzoate	0.379	0.375	0.475	0.475	3.592	3.050	906.177	14.409	6.208	5.825	194.094
30	2-Hydroxydiphenylmethane	0.191	0.199	0.705	0.705	3.742	1.551	862.004	11.175	7.077	6.483	184.089
31	5-Pentylresorcinol	0.341	0.523	0.290	0.290	4.468	1.355	889.615	12.713	8.151	5.287	180.115
32	2-Tert-butyl-4-methylphenol	0.238	0.240	0.258	0.258	4.311	1.766	824.715	10.106	7.581	6.127	164.120
33	4-Cyclopentylphenol	0.266	0.412	0.733	0.733	3.789	1.362	773.063	11.209	5.932	5.917	162.104
34	2-Tert-butylphenol	0.238	0.240	0.248	0.248	4.490	1.391	765.281	7.573	7.346	6.716	150.104
35	4-Bromo-6-chloro-2-eresol	1.582	4.006	0.981	0.981	3.509	3.589	704.594	9.069	8.284	4.130	219.929
36	2,3-Dichlorophenol	1.562	4.101	1.087	1.087	3.523	1.324	590.520	7.987	7.773	3.491	161.964
37	4-Bromo-3,5-dimethylphenol	1.359	2.811	0.235	0.235	4.419	1.911	682.930	8.264	8.247	4.139	199.984
38	2-Hydroxybenzophenone	0.341	0.333	0.974	0.974	2.477	4.067	850.092	11.887	7.604	4.676	198.068
39	4-Tert-Pentylphenol	0.248	0.395	0.716	0.716	3.833	1.389	855.778	11.098	6.744	6.118	164.120
40	4-Chloro-3,5-dimethylphenol	0.345	0.545	0.238	0.238	4.402	1.924	669.234	8.260	8.197	4.154	156.034

Table S3. (Continued: Compound 41 -83)

label	compound name	descriptors										
		$\lambda_{\text{SP} \rightarrow \text{A}}$ (eV)	$\lambda_{\text{A} \rightarrow \text{SP}}$ (eV)	$\lambda_{\text{SP} \rightarrow \text{T}}$ (eV)	$\lambda_{\text{T} \rightarrow \text{SP}}$ (eV)	μ (D)	ΔE_{olv} (eV)	V_{mol} (Å ³)	B_1 (Å)	B_2 (Å)	B_3 (Å)	M (amu)
41	4-Bromo-2,6-dimethylphenol	1.205	2.517	1.207	3.365	2.913	0.183	691.097	8.967	8.248	4.159	199.984
42	6-Tert-butyl-2,4-dimethylphenol	0.243	0.245	0.264	4.333	1.691	0.144	890.640	7.959	7.649	3.395	178.136
43	3-Bromophenol	1.212	2.545	0.853	3.550	3.141	0.192	564.689	8.548	7.326	3.665	171.952
44	2,5-Dichlorophenol	1.509	4.045	0.896	3.501	1.471	0.127	600.163	7.913	5.429	3.756	161.964
45	3-Chloro-4-fluorophenol	0.234	0.241	0.855	3.489	2.242	0.200	532.795	8.294	7.288	3.482	145.993
46	2-Phenylphenol	0.338	0.293	0.785	3.431	1.734	0.173	789.903	11.653	6.942	4.591	170.073
47	4-Chloro-3-ethylphenol	0.267	0.273	0.243	4.338	1.851	0.192	660.018	9.188	7.993	4.124	156.034
48	Benzyl-4-hydroxyphenyl ketone	0.318	0.313	0.481	3.402	2.466	0.359	951.744	14.606	6.471	5.128	212.084
49	2,4,6-Tribromoresorcinol	1.542	2.768	1.379	3.350	2.826	0.189	782.491	9.337	8.765	3.695	343.768
50	2,4-Dichlorophenol	1.517	3.992	0.238	4.191	1.035	0.128	597.509	8.857	7.300	3.497	161.964
51	4-Benzylloxyphenol	1.111	2.575	0.472	3.644	2.266	0.281	915.002	14.243	6.963	4.166	200.084
52	4-Hydroxybenzophenone	0.266	0.256	0.501	3.333	2.225	0.311	899.018	12.840	7.517	4.564	198.068
53	3-Hydroxydiphenylamine	0.158	0.204	0.278	3.888	0.820	0.290	844.546	12.694	6.954	4.572	185.084
54	4-Sec-butylphenol	0.250	0.411	0.239	4.287	1.379	0.184	762.773	11.207	5.931	5.772	150.104
55	4,6-Dichlororesorcinol	1.621	4.028	1.041	3.409	2.227	0.162	618.468	8.905	7.341	3.486	177.959
56	2,2',4,4'-Tetrahydroxybenzophenone	0.354	0.353	0.817	2.692	3.472	0.513	959.840	13.656	7.348	5.350	246.053
57	3,4,5-Trimethylphenol	0.327	0.495	0.247	4.403	1.482	0.187	669.606	8.315	8.165	4.160	136.089
58	4-Tert-butylphenol	0.234	0.237	0.721	3.830	1.355	0.184	764.112	10.092	6.717	6.104	150.104
59	2,2'-Biphenol	0.296	0.281	0.707	3.521	3.152	0.313	802.403	11.626	7.284	4.835	186.068
60	3-Chlorophenol	1.063	2.968	0.782	3.643	3.197	0.191	509.195	8.292	7.302	3.485	128.003
61	4-Chloro-3-methylphenol	0.321	0.491	0.242	4.328	1.962	0.195	605.090	8.637	7.263	4.137	142.019
62	2-Isopropylphenol	0.286	0.438	0.266	4.464	1.749	0.169	683.115	9.429	7.064	5.392	136.089
63	2,6-Dichloro-4-fluorophenol	1.370	3.920	0.835	3.437	1.381	0.141	604.491	8.466	8.190	3.491	179.955
64	2-Methyl-3-nitrophenol	0.498	0.464	0.604	3.177	3.299	0.257	629.549	8.635	7.761	3.951	153.043
65	3-Chloro-5-methoxyphenol	1.223	2.912	1.021	3.712	4.513	0.229	650.664	9.260	8.285	4.148	158.013
66	2-Methoxy-4-propenylphenol	0.319	0.277	1.113	2.489	2.283	0.197	768.687	11.651	7.931	4.148	164.084
67	3-Tert-butylphenol	0.323	0.301	0.235	4.496	1.138	0.184	773.791	9.441	6.707	6.144	150.104
68	2,6-Dichlorophenol	1.560	3.982	1.070	3.540	2.100	0.154	599.269	8.482	7.705	3.479	161.964
69	2-Hydroxy-4,5-dimethylacetophenone	0.300	0.283	0.816	2.610	4.703	0.185	781.261	9.765	7.748	4.157	164.084
70	4-Butoxyphenol	0.258	0.261	0.770	3.661	1.823	0.229	823.766	13.606	6.668	4.790	166.099
71	4-Chloro-2-methylphenol	0.347	0.518	0.283	4.253	2.819	0.194	588.051	8.731	7.210	4.121	142.019
72	2-Chloro-4,5-dimethylphenol	1.749	4.041	0.849	3.708	1.640	0.130	680.552	8.763	7.483	4.153	156.034
73	4-Bromophenol	0.259	0.269	0.665	3.829	2.420	0.202	559.136	8.308	7.363	3.683	171.952
74	4-Propylphenol	0.255	0.436	0.241	4.288	1.375	0.184	709.411	11.124	5.926	5.446	136.089
75	5-Bromovanillin	0.286	0.288	0.451	3.069	3.059	0.300	733.607	9.936	8.790	4.168	229.958
76	α,α -Trifluoro-4-cresol	0.314	0.418	0.247	4.544	3.154	0.203	564.415	9.190	5.746	5.075	162.029
77	Methyl-4-methoxysalicylate	0.378	0.369	0.331	3.574	2.967	0.199	748.809	12.312	7.231	4.143	182.058
78	3-Isopropylphenol	0.331	0.465	0.259	4.467	1.597	0.185	699.927	9.402	6.116	5.983	136.089
79	3-Methyl-2-nitrophenol	0.466	0.440	1.303	2.096	3.424	0.165	610.653	8.320	7.957	4.080	153.043
80	2-Bromo-4-methylphenol	1.723	3.461	0.980	3.326	1.368	0.123	615.325	8.973	7.260	4.108	185.968
81	2,4-Difluorophenol	0.285	0.289	0.753	3.703	0.625	0.131	466.688	7.626	7.249	3.398	130.023
82	5-Methyl-2-nitrophenol	0.476	0.459	1.295	2.201	4.144	0.177	619.966	9.632	7.005	4.099	153.043
83	Ethyl-4-hydroxybenzoate	0.373	0.371	0.474	3.589	1.095	0.262	741.168	12.040	6.587	4.893	166.063

Table S3. (Continued: Compound 84-125)

compound		descriptors										
label	name	$\lambda_{80\rightarrow\Lambda}$ (eV)	$\lambda_{A\rightarrow S0}$ (eV)	$\lambda_{S0\rightarrow T1}$ (eV)	$\lambda_{T1\rightarrow S0}$ (eV)	μ (D)	ΔE_{calc} (eV)	$V_{\text{mol}}(\text{\AA}^3)$	B_1 (Å)	B_2 (Å)	B_3 (Å)	M (amu)
84	4-Chlorophenol	0.259	0.268	0.242	4.249	2.452	0.201	518.676	7.915	7.426	3.483	128.003
85	2-Hydroxy-4-methoxyacetophenone	0.335	0.326	0.884	2.813	2.564	0.236	722.049	11.554	7.214	4.171	166.063
86	4-Cyanophenol	0.240	0.242	0.509	3.580	5.049	0.321	528.877	8.596	7.516	3.396	119.037
87	4-Propyloxyphenol	0.259	0.265	0.547	3.658	1.796	0.227	728.137	12.498	6.807	4.156	152.084
88	Ethyl-3-hydroxybenzoate	0.385	0.381	0.470	3.451	1.055	0.240	731.326	11.440	6.610	5.395	166.063
89	2,6-Difluorophenol	0.627	1.236	1.289	3.677	1.994	0.151	454.575	7.820	7.243	3.396	130.023
90	4-Isopropylphenol	0.235	0.240	0.714	3.817	1.388	0.184	690.432	10.219	5.918	5.870	136.089
91	4-Allyl-2-methoxyphenol	0.431	0.542	0.888	3.607	2.802	0.181	771.246	9.719	8.430	5.201	164.084
92	2-Hydroxybenzaldehyde	0.300	0.293	0.885	2.614	4.059	0.165	529.515	8.478	7.354	3.399	122.037
93	2-Chloro-5-methylphenol	1.729	4.050	0.880	3.783	1.275	0.126	603.585	8.882	7.328	3.913	142.019
94	3-Methoxysalicylaldehyde	0.296	0.295	0.392	3.325	5.911	0.251	645.771	9.081	8.283	4.136	152.047
95	Salicylhydroxamic acid	0.503	0.428	0.673	3.136	5.203	0.249	614.062	9.698	7.280	3.675	153.043
96	3-Fluorophenol	0.277	0.280	0.525	4.019	2.705	0.187	460.086	7.817	7.235	3.398	112.032
97	2,3,5-Trimethylphenol	0.249	0.399	0.237	4.452	1.672	0.172	702.497	8.762	7.984	4.154	136.089
98	5-Bromo-2-hydroxybenzyl alcohol	0.369	0.464	0.555	3.672	3.824	0.268	664.363	10.160	7.171	4.463	201.963
99	2-Bromophenol	1.769	3.482	0.773	3.854	0.971	0.122	538.329	7.978	7.308	3.691	171.952
100	2-Allylphenol	1.377	2.221	0.281	4.384	1.875	0.158	636.525	9.957	6.534	5.217	134.073
101	2-Hydroxy-5-methylacetophenone	0.306	0.290	0.812	2.564	4.384	0.177	688.328	9.366	7.749	4.146	150.068
102	2,3,6-Trimethylphenol	0.265	0.345	0.241	4.495	1.165	0.160	693.485	8.922	7.934	4.163	136.089
103	2,4,6-Trimethylphenol	0.325	0.364	0.286	4.323	1.396	0.161	696.291	8.660	8.328	4.161	136.089
104	4-Hydroxybenzaldehyde	0.261	0.264	0.466	3.336	3.072	0.279	522.869	9.373	6.679	3.398	122.037
105	Salicylaldoxime	0.312	0.401	0.738	2.800	1.920	0.215	598.323	10.056	7.338	3.396	137.048
106	3-Ethylphenol	0.286	0.458	0.547	4.457	1.586	0.184	608.988	9.206	7.280	4.836	122.073
107	4-Ethylphenol	0.227	0.230	0.537	3.785	1.626	0.185	603.672	10.213	6.676	4.154	122.073
108	4-Hydroxy-2-methylacetophenone	0.321	0.305	0.577	3.313	2.251	0.288	672.871	8.612	8.512	4.492	150.068
109	2-Fluorophenol	0.274	0.275	0.555	3.928	2.608	0.133	468.638	7.743	7.110	3.394	112.032
110	Salicylhydrazide	0.464	0.423	0.731	3.121	2.574	0.265	620.997	9.880	7.259	3.771	152.059
111	2-Chlorophenol	1.752	4.044	0.238	4.418	0.949	0.125	524.096	7.767	7.321	3.477	128.003
112	Syringaldehyde	0.291	0.292	0.432	2.881	3.127	0.319	755.658	9.956	9.422	4.170	182.058
113	2-Ethylphenol	0.387	0.400	0.249	4.500	1.726	0.173	594.140	8.482	7.823	4.161	122.073
114	4-Chlororesorcinol	1.883	4.123	1.142	3.416	0.974	0.222	537.190	8.147	7.316	3.493	143.998
115	4-Hydroxypropionophenone	0.295	0.291	0.483	3.412	2.241	0.275	702.166	11.222	6.673	4.146	150.068
116	2,3-Dimethylphenol	0.241	0.399	0.237	4.481	0.823	0.174	606.902	7.977	7.667	4.157	122.073
117	3,4-Dimethylphenol	0.240	0.243	0.242	4.358	1.726	0.187	607.342	8.234	7.485	4.156	122.073
118	3,5-Dimethylphenol	0.320	0.336	0.269	4.473	1.413	0.185	625.577	8.381	8.250	4.146	122.073
119	3-Hydroxybenzaldehyde	0.279	0.279	0.472	3.275	1.770	0.251	533.507	8.255	7.527	3.399	122.037
120	2-Hydroxyacetophenone	0.311	0.294	0.886	2.650	1.743	0.176	596.051	9.115	7.386	4.139	136.052
121	2,5-Dimethylphenol	0.286	0.311	0.249	4.465	1.397	0.172	613.179	9.239	6.953	4.153	122.073
122	Methyl-4-hydroxybenzoate	0.356	0.354	0.474	3.591	1.178	0.272	639.601	10.897	6.491	4.128	152.047
123	2,4-Dimethylphenol	0.272	0.351	0.260	4.300	1.650	0.172	606.164	8.763	7.310	4.127	122.073
124	4-Hydroxypropionophenone	0.282	0.280	0.482	3.385	2.241	0.279	682.749	10.258	6.813	4.931	150.068
125	5-Fluoro-2-hydroxyacetophenone	0.326	0.309	0.838	2.445	2.958	0.173	627.291	9.128	7.367	4.134	154.043

Table S3. (Continued: Compound 126-167)

compound		descriptors										
label	name	$\lambda_{SO \rightarrow A}$ (eV)	$\lambda_{A \rightarrow SO}$ (eV)	$\lambda_{SO \rightarrow TI}$ (eV)	$\lambda_{TI \rightarrow SO}$ (eV)	μ (D)	ΔE_{olv} (eV)	V_{mol} (Å ³)	B_1 (Å)	B_2 (Å)	B_3 (Å)	M (amu)
126	2-Cyanophenol	0.244	0.244	0.444	3.603	3.316	0.204	528.357	8.344	7.342	3.398	119.037
127	4-Fluorophenol	0.247	0.252	0.481	3.872	1.968	0.190	468.169	7.475	7.201	3.395	112.032
128	4-Ethoxyphenol	0.258	0.264	0.447	3.657	0.629	0.229	647.027	11.105	6.662	4.163	138.068
129	3-Ethoxy-4-hydroxybenzaldehyde	0.280	0.280	0.435	3.110	3.722	0.237	749.440	10.566	8.188	4.161	166.063
130	4-Hydroxy-3-methoxybenzonitrile	0.260	0.261	0.480	3.344	4.203	0.289	641.294	9.806	7.654	4.142	149.048
131	3-Methoxy-4-hydroxybenzaldehyde	0.279	0.280	0.438	3.078	3.692	0.265	638.439	10.077	7.658	4.149	152.047
132	Methyl-3-hydroxybenzoate	0.368	0.364	0.470	3.454	1.019	0.250	653.456	10.343	7.341	4.093	152.047
133	3-Cyanophenol	0.244	0.246	0.509	3.474	1.658	0.186	538.945	8.333	7.252	4.036	108.058
134	3-Methylphenol	0.256	0.444	0.231	4.453	5.739	0.295	528.924	8.871	7.345	3.399	119.037
135	2-Hydroxyethylsalicylate	0.397	0.381	0.591	3.191	2.106	0.301	510.442	8.744	6.679	3.396	109.053
136	3,5-Dimethoxyphenol	0.521	1.009	0.757	3.874	0.441	0.260	671.776	10.343	8.172	4.166	154.063
137	4-Hydroxy-3-methoxyacetophenone	0.301	0.298	0.455	3.141	2.957	0.280	719.450	9.568	8.281	4.169	166.063
138	4-Methoxyphenol	0.257	0.263	0.448	3.659	2.228	0.283	661.660	10.154	7.169	4.153	152.047
139	3-Hydroxy-4-methoxybenzaldehyde	0.292	0.293	0.446	3.200	2.542	0.234	563.822	8.819	7.343	4.111	124.052
140	3-Acetamidophenol	0.233	0.213	0.549	3.748	1.855	0.387	736.926	9.922	9.122	3.907	228.997
141	4-Hydroxy-3-methoxyphenethylalcohol	0.511	0.640	0.901	3.603	1.340	0.183	526.815	8.938	6.519	4.070	108.058
142	4-Methylphenol	0.262	0.436	0.252	4.281	2.843	0.284	759.424	10.672	7.917	4.881	168.079
143	Phenol	0.236	0.242	0.239	4.470	1.369	0.185	442.183	7.326	7.244	3.398	94.042
144	Ethyl-4-hydroxy-3-methoxyphenylacetate	0.378	0.380	0.616	3.594	1.264	0.277	913.009	10.451	8.488	6.266	210.089
145	2-Hydroxybenzamide	0.449	0.404	0.782	3.054	4.927	0.285	572.797	9.012	7.331	3.508	137.048
146	2-Methylphenol	0.244	0.248	0.243	4.487	1.699	0.175	542.996	7.826	7.192	4.093	108.058
147	3-Ethoxy-4-methoxyphenol	0.548	0.643	0.462	3.780	2.390	0.284	616.079	10.220	6.602	4.118	136.052
148	4-Acetylphenol	0.288	0.285	0.483	3.406	1.300	0.299	741.512	10.280	8.916	4.175	168.079
149	3-Methoxyphenol	0.288	0.290	0.226	4.501	2.507	0.225	569.494	9.259	7.061	4.138	124.052
150	2-Ethoxyphenol	0.300	0.297	0.917	3.680	2.097	0.172	642.403	10.179	7.167	4.157	138.068
151	3-Hydroxyacetophenone	0.305	0.299	0.479	3.351	1.701	0.257	614.337	9.670	7.350	4.098	136.052
152	5-Methylresorcinol	0.364	0.565	0.952	3.823	2.017	0.280	568.580	8.291	7.876	3.988	124.052
153	4-(4-Hydroxyphenyl)-2-butanone	0.646	0.458	0.528	3.907	2.848	0.286	786.740	10.775	6.885	5.970	164.084
154	2-Methoxyphenol	0.297	0.295	0.928	3.673	6.498	0.288	551.673	8.773	7.173	3.464	138.032
155	Salicylic acid	0.365	0.350	0.738	3.067	0.789	0.175	561.058	8.332	7.709	4.142	124.052
156	2,4,6-Tris(dimethylaminomethyl)phenol	0.320	0.343	0.631	3.747	1.324	0.295	484.618	8.117	7.266	3.392	109.053
157	2,6-Dimethoxyphenol	0.283	0.288	0.559	3.624	3.275	0.260	677.377	10.610	7.535	4.139	154.063
158	Resorcinol	0.302	0.305	0.232	4.476	1.370	0.281	482.717	7.932	7.234	3.395	110.037
159	4-Hydroxy-3-methoxybenzyl alcohol	0.545	0.596	0.649	3.599	1.684	0.247	665.050	8.976	8.251	4.827	154.063
160	4-Hydroxybenzamide	0.466	0.409	0.559	3.600	4.696	0.402	582.345	10.076	6.646	3.728	137.048
161	3-Hydroxybenzoic acid	0.363	0.360	0.464	3.436	3.444	0.284	545.779	9.358	7.434	3.396	138.032
162	4-Hydroxy-3-methoxybenzyl amine	0.531	0.662	0.663	3.599	4.722	0.376	673.526	11.237	6.748	3.949	151.063
163	4-Hydroxyphenethyl alcohol	0.243	0.246	0.239	4.320	2.012	0.295	640.440	10.890	5.963	5.435	138.068
164	2-Hydroxybenzylalcohol	0.348	0.356	0.616	3.766	1.859	0.260	555.378	7.959	7.800	4.389	124.052
165	3-Hydroxy-4-methoxybenzyl alcohol	0.647	0.781	0.799	1.783	3.118	0.256	685.768	10.769	7.077	4.528	154.063
166	4-hydroxybenzoic acid	0.350	0.348	0.466	3.583	4.643	0.305	559.162	9.908	6.676	3.395	138.032
167	3-Hydroxybenzyl alcohol	0.446	0.421	0.835	3.828	0.401	0.293	544.162	9.228	7.165	4.100	124.052

Table S3. (Continued: Compound 168-208)

label	compound name	descriptors										
		$\lambda_{80 \rightarrow A}$ (eV)	$\lambda_{A \rightarrow 80}$ (eV)	$\lambda_{80 \rightarrow T1}$ (eV)	$\lambda_{T1 \rightarrow 80}$ (eV)	μ (D)	ΔE_{sub} (eV)	V_{mol} (Å ³)	B_1 (Å)	B_2 (Å)	B_3 (Å)	M (amu)
168	1,3,5-Trihydroxybenzene	0.633	1.212	0.607	4.120	2.748	0.365	512.322	8.143	7.910	3.399	126.032
169	4-Hydroxyphenylacetic acid	0.293	0.342	0.261	4.324	5.148	0.330	631.042	9.931	5.933	5.430	152.047
170	Methoxyhydroquinone	0.315	0.318	0.624	3.568	1.752	0.271	581.895	9.335	7.141	4.142	140.047
171	3,5-Di-tert-butylcatechol	0.477	0.557	1.001	3.542	2.861	0.197	1121.473	10.196	9.661	6.373	222.162
172	Phenylhydroquinone	0.360	0.315	0.572	3.219	0.253	0.336	794.457	11.539	8.066	4.456	186.068
173	Methylhydroquinone	0.262	0.267	0.532	3.632	3.042	0.276	553.571	8.248	7.168	4.089	124.052
174	Bromohydroquinone	1.640	3.445	0.822	3.225	1.618	0.227	561.897	8.971	7.125	3.670	187.947
175	4-Amino-2,3-dimethylphenol	0.283	0.396	0.520	3.524	1.934	0.273	634.593	8.715	7.839	4.094	137.084
176	2,3-Dimethylhydroquinone	0.273	0.432	0.494	3.656	3.183	0.265	636.248	8.584	7.715	4.112	138.068
177	Trimethylhydroquinone	0.311	0.448	0.505	3.679	2.500	0.255	725.851	8.952	8.263	4.166	152.084
178	4-Amino-2-cresol	0.359	0.415	0.507	3.502	2.237	0.287	571.564	8.747	7.222	4.106	123.068
179	Chlorohydroquinone	1.616	4.047	0.496	3.624	1.512	0.229	556.061	8.709	7.146	3.483	143.998
180	4-Chlorocatechol	0.487	0.663	0.824	3.633	3.972	0.246	537.994	8.618	7.319	3.491	143.998
181	2-Aminophenol	0.217	0.220	0.918	3.606	2.700	0.203	488.513	7.536	7.286	4.006	109.053
182	6-Amino-2,4-dimethylphenol	0.294	0.326	0.628	3.791	1.462	0.187	667.465	8.905	8.010	4.155	137.084
183	1,2,3-Trihydroxybenzene	0.519	1.031	0.817	3.613	2.132	0.260	509.852	7.952	7.632	3.398	126.032
184	2-Amino-4-chlorophenol	0.239	0.248	0.828	3.490	2.628	0.215	581.338	8.395	7.274	4.009	143.014
185	Catechol	0.299	0.297	0.943	3.640	2.553	0.229	478.372	7.639	7.282	3.396	110.037
186	Hydroquinone	0.260	0.266	0.461	3.639	0.007	0.287	477.327	7.628	7.332	3.400	110.037
187	5-Amino-2-methoxyphenol	0.505	0.660	0.674	3.487	2.188	0.282	622.926	10.001	7.095	4.129	139.063
188	1,2,4-Trihydroxybenzene	0.513	0.752	0.715	3.520	3.800	0.321	511.652	8.145	7.330	3.399	126.032
189	4-Methylcatechol	0.439	0.561	0.845	3.641	2.723	0.227	555.458	8.870	6.984	4.103	124.052
190	2-Amino-4-tert-butylphenol	0.212	0.213	0.617	3.770	1.381	0.201	835.860	10.138	6.551	6.534	165.115
191	3-Methylcatechol	0.454	0.599	0.556	3.792	2.881	0.218	577.239	7.881	7.805	4.106	124.052
192	2,4-Diaminophenol	0.465	0.423	0.581	3.385	2.391	0.340	550.796	8.575	7.339	3.624	124.064
193	4-Aminophenol	0.293	0.407	0.456	3.471	3.915	0.253	761.425	11.363	7.749	4.685	182.058
194	3-Aminophenol	0.382	0.529	0.564	3.796	1.081	0.263	1388.995	11.830	11.437	6.226	265.215
195	4-Acetamidophenol	0.330	0.578	0.477	3.759	1.294	0.256	702.720	10.633	7.538	4.823	153.079
196	4-Chloro-2-nitrophenol	0.479	0.464	1.244	2.092	5.448	0.167	609.120	9.006	7.561	3.481	172.988
197	4-Bromo-2-nitrophenol	0.478	0.464	1.234	2.096	5.433	0.167	642.507	9.505	7.483	3.679	216.937
198	3-Methyl-4-nitrophenol	0.390	0.366	0.483	3.388	5.067	0.295	618.467	8.658	7.954	4.072	153.043
199	4-Nitro-3-(trifluoromethyl)phenol	0.511	0.414	0.564	3.302	6.515	0.331	664.542	8.533	8.267	4.913	207.014
200	4-Chloro-6-nitro-3-cresol	0.469	0.459	1.261	2.116	6.059	0.165	690.376	9.585	7.355	4.123	187.004
201	2-Chloro-4-nitrophenol	0.395	0.382	0.479	3.370	5.054	0.223	613.761	9.475	7.362	3.471	172.988
202	4-Nitrophenol	0.393	0.378	0.480	3.399	5.188	0.306	543.330	9.221	6.686	3.395	139.027
203	4-Fluoro-2-nitrophenol	0.476	0.455	1.221	2.016	5.499	0.166	551.003	8.714	7.149	3.397	157.018
204	2,6-Dibromo-4-nitrophenol	0.391	0.379	0.465	3.329	3.746	0.228	735.836	9.228	9.140	3.693	294.848
205	4-Nitrocatechol	0.410	0.394	0.466	3.116	6.590	0.329	569.893	9.432	7.411	3.394	155.022
206	2-Amino-4-chloro-5-nitrophenol	0.621	0.424	0.650	2.999	7.280	0.395	675.574	9.412	7.562	3.943	187.999
207	5-Fluoro-2-nitrophenol	0.502	0.511	1.376	2.289	4.555	0.162	571.029	9.167	7.142	3.392	157.018
208	2-Fluoro-4-nitrophenol	0.402	0.386	0.495	3.310	5.019	0.236	571.187	9.158	7.132	3.391	157.018

Table S3. (Continued: Compound 209-245)

compound			descriptors										
label	name		$\lambda_{S0 \rightarrow A}$ (eV)	$\lambda_{A \rightarrow S0}$ (eV)	$\lambda_{S0 \rightarrow T1}$ (eV)	$\lambda_{T1 \rightarrow S0}$ (eV)	μ (D)	ΔE_{S0} (eV)	$V_{mol}(\text{\AA}^3)$	$B_1(\text{\AA})$	$B_2(\text{\AA})$	$B_3(\text{\AA})$	M (amu)
209	3-Fluoro-4-nitrophenol		0.383	0.371	0.487	3.506	6.025	0.334	561.573	9.311	7.168	3.396	157.018
210	4-Amino-2-nitrophenol		0.492	0.463	1.115	1.603	6.443	0.289	598.643	9.359	7.501	3.391	154.038
211	2-Bromo-2'-hydroxy-5'-nitroacetanilide		0.411	0.399	0.499	3.390	4.591	0.481	883.014	11.486	7.257	6.469	273.959
212	2-Chloromethyl-4-nitrophenol		0.412	0.409	0.477	3.362	5.157	0.274	707.397	10.214	7.077	5.008	187.004
213	4-Methyl-3-nitrophenol		0.426	0.382	0.496	3.102	3.120	0.264	615.271	8.743	7.900	4.076	153.043
214	2-Nitrophenol		0.473	0.458	1.311	2.178	5.827	0.171	546.717	8.689	7.146	3.396	139.027
215	2-Nitroresorcinol		0.485	0.450	0.929	2.088	2.320	0.174	570.768	8.359	7.982	3.393	155.022
216	4-Nitrosophenol		0.343	0.348	0.165	0.721	4.043	0.261	518.117	9.208	6.687	3.393	123.032
217	4-Hydroxy-3-nitrobenzaldehyde		0.587	0.855	1.310	2.257	3.873	0.243	631.929	8.822	8.669	3.391	167.022
218	4-Methyl-2-nitrophenol		0.464	0.447	1.242	2.070	6.044	0.172	616.327	9.309	7.522	4.114	153.043
219	3-Nitrophenol		0.413	0.394	0.501	3.211	5.800	0.273	543.936	8.637	7.537	3.396	139.027
220	2-Amino-4-nitrophenol		0.401	0.386	0.893	2.852	4.692	0.310	593.373	9.357	7.256	4.004	154.038
221	5-Hydroxy-2-nitrobenzaldehyde		0.479	0.374	0.501	3.106	5.710	0.333	633.204	8.391	8.336	3.754	167.022
222	3-Hydroxy-4-nitrobenzaldehyde		0.415	0.400	1.034	2.158	3.224	0.235	622.478	10.008	7.277	3.397	167.022
223	2,3,4,5-Tetrachlorophenol		0.883	2.250	1.065	3.314	2.075	0.129	733.995	8.929	8.507	3.496	229.886
224	Pentabromophenol		0.999	2.067	1.268	3.050	1.759	0.150	943.847	9.424	9.359	3.697	483.594
225	3,4,5,6-Tetrabromo-2-cresol		0.947	1.537	1.338	3.096	3.547	0.121	903.159	9.400	9.383	4.149	419.700
226	2,3,5,6-Tetrachlorophenol		1.220	3.015	0.828	3.411	1.408	0.143	758.996	8.862	7.967	3.499	229.886
227	Pentachlorophenol		0.767	2.162	1.122	3.198	1.923	0.142	797.004	9.148	8.894	3.499	263.847
228	2,4-Dichloro-6-nitrophenol		0.492	0.490	1.296	2.002	4.108	0.173	684.885	8.890	8.588	3.489	206.949
229	4,6-Dinitro-2-methylphenol		0.641	0.849	1.380	2.242	6.699	0.256	702.576	9.318	8.838	4.123	198.028
230	Pentafluorophenol		1.039	1.268	1.639	3.350	1.974	0.142	513.937	8.328	7.732	3.397	183.995
231	4-Bromo-2-fluoro-6-nitrophenol		0.418	0.378	0.460	3.251	1.708	0.275	645.558	9.504	7.990	3.675	234.928
232	2,6-Dinitro-4-cresol		0.554	0.848	1.250	2.116	4.622	0.311	716.212	9.431	8.839	4.140	198.028
233	2,3,5,6-Tetrafluorophenol		0.705	1.165	1.013	3.577	1.405	0.146	496.141	7.959	7.649	3.395	166.004
234	2,4-Dinitrophenol		0.618	0.867	1.364	2.336	6.075	0.264	630.979	9.198	8.325	3.393	184.012
235	2,5-Dinitrophenol		0.415	0.404	1.057	2.191	1.295	0.246	619.980	9.878	7.274	3.399	184.012
236	2,6-Dichloro-4-nitrophenol		0.398	0.386	0.479	3.335	2.993	0.232	688.813	9.197	8.534	3.493	206.949
237	2,6-Dinitrophenol		0.554	0.828	1.332	2.237	4.103	0.314	636.155	8.680	8.581	3.877	184.012
238	2,3-Dinitrophenol		0.508	0.458	1.343	2.173	5.522	0.292	633.593	8.155	8.141	4.605	184.012
239	3,4-Dinitrophenol		0.541	0.467	0.721	2.999	6.629	0.427	643.841	8.708	7.959	4.530	184.012
240	2,4,6-Trinitrophenol		0.654	0.834	1.382	2.375	2.591	0.352	665.918	10.367	7.626	3.920	151.063
241	Tetrachlorohydroquinone		1.133	2.948	0.686	3.255	2.695	0.200	764.892	8.850	8.616	3.496	245.881
242	Tetrafluorohydroquinone		0.394	0.399	0.278	4.424	2.790	0.211	527.529	8.327	7.726	3.396	181.999
243	Tetrachlorocatechol		1.586	4.083	0.955	3.382	3.490	0.199	758.208	9.116	8.926	3.494	245.881
244	Tetramethylhydroquinone		0.242	0.244	0.486	3.650	2.616	0.230	775.572	8.893	8.748	4.163	166.099
245	Tetrabromocatechol		1.235	2.368	1.282	3.026	3.151	0.199	849.709	9.415	9.209	3.691	421.679

Table S4. Summary of the optimized hyperparameters for all molecules and molecules in each MOA class

hyperparameter	range	optimal value					
		all MOA classes	PN	PE	SE	WARU	PRC
eta	[0,0.5]	0.35	0.50	0.50	0.00	0.00	0.44
gamma	[0,1]	0.00	0.00	0.00	0.00	0.00	0.00
max_depth	[1,5]	4.00	5.00	2.00	5.00	5.00	4.00
min_child_weight	[0,1]	0.00	1.00	1.00	0.00	0.00	0.00
subsample	[0,0.5]	0.50	0.50	0.50	0.50	0.50	0.39
colsample_bytree	[0,0.5]	0.40	0.50	0.50	0.00	0.50	0.39

Figure S1. Correlations between the predicted (horizontal axis) and experimental (vertical axis) $pIGC_{50}$ values for ESI descriptors (a-b) and Cronin's descriptor (c-d) in XGBR and MLR modeling's.

Figure S2. The *feature importance score* for the 21 ESI descriptors evaluated in the best XGBR model in the 100-times data separation.