Research Article

International Journal of Probiotics and Dietetics

3D-QSTR study of the aquatic toxicity of phenol derivatives to *Chlorella Vulgaris*

Assia Belhassan[1], Tahar Lakhlifi[1] and Mohammed Bouachrine[1,2*]

[1]*MCNS Laboratory, Faculty of Science, Moulay Ismail University, Meknes, Morocco*
[2]*EST Khenifra, Sultan MoulaySlimane University, Benimellal, Morocco*

[*]Corresponding author
Mohammed Bouachrine, EST Khenifra, Sultan MoulaySliman University, Benimellal, Morocco, E-mail: m.bouachrine@est-umi.ac.ma;
Tel. +212661416217

Submitted: 27 Aug 2020; **Accepted:** 03 Sept 2020; **Published:** 10 Sept 2020

Abstract
This study aims to develop quantitative structure-toxicity models to predict the toxicity of phenol derivatives. For that purpose, thetoxicity data of a total of 43substituted phenols to Chlorella Vulgaris (C. Vulgaris), were used to build and validate the three dimensional quantitative structure–toxicity relationships (3D-QSTR) models. The 3D-QSTR models were constructed using Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Indices Analysis (CoMSIA) methods. The created models displayed goodpredictivity, which can be used to predict the toxicity of other phenol derivatives on C. Vulgaris; the statistical results indicate that the predicted values are in good agreement with the experimental results for CoMFAand CoMSIAmodels; (r^2 = 0.760; Q^2 = 0.559;r^2_{test} = 0.761)and (r^2 = 0.778; Q^2 = 0.680; r^2_{test} = 0.534), respectively.

Keywords:
Chlorella vulgaris; Phenols; Three-dimensional quantitative structure toxicity relationship, Comparative Molecular Field Analysis; Comparative Molecular Similarity Indices Analysis.

Introduction
Phenols are aromatic chemical compounds bearing a hydroxyl function -OH. Derivatives carrying several hydroxyl functions are called polyphenols.Phenols and derivatives have important biological activities (biochemical defense against microbes and fungi in plants in particular);however, they are toxic to humans and other species and when they are abnormally disseminated in the environment, phenols are pollutants of air, soil or water. In the aquatic environment,these molecules arise from industrial, agricultural activities and natural compounds degradation [1]. Pharmaceuticals additives and food practices are also sources of phenoliccompounds [2].

The eukaryotic microalgae *Chlorella Vulgaris* is an algae of the genus Chlorella, present on Earth since the Precambrian period. Martinus Willem Beijerinck discovered this unicellular alga in 1890 as the first well-defined kernel microalgae.Chlorellaspeciesare also a promising biomass for biofuel production and have a high potential for replacing fossil fuels. In addition, it has been reported that *Chlorella* species can produce a high yield of bioenergy and have been grown on an industrial scale. Therefore, evaluation of data from toxicity tests using microalgae is an integral part of the environmental risk assessment [3,4]. Therefore, the REACH legislation (Registration, Evaluation,Authorization and Restriction of Chemicals) [5]requires the provision of ecotoxicological information obtained from algal growth inhibition tests for manufactured or imported compounds. With regard to substances without experimental data, REACH also encourages the use of non-experimental methods, including the Quantitative Structure Activity Activity (QSAR) approach, to predict the toxicity of untested chemicals [6].

The three-dimensional quantitative-activity/property-relationship (3D-QSAR/QSPR) method is one of the most widely computational methods used for predicting the activity/property of molecules[7,8]. With continued advancements, the QSAR/QSPR method has been remarkably successful in a variety of areas, such as medicinal chemistry, materials science and predictive toxicology [9,10]. Quantitative structure-toxicity relationship (QSTR) is becoming a gooddevice for predicted the toxicity of a chemical using computational methods. It should be remembered that a little work has been done on the 3D-QSTR study of the structure-toxicity correlation especially of phenol derivatives; so we have seen fit to look at this type of study in the case of our molecules.

In the present study, 43 substituted phenols were used to construct and validate the 3D-QSTR modelsusing the toxicity data for *C. Vulgaris*. The selected chemicals for toxicity include chlorine-, methyl-, and nitro-substituted phenols exhibited a wide range of algal toxicity from -0.60 to 2.34[11], which enables the construction of models with confidence.

Material and methods
2.1. Data set
In the present study, a series of 43 selected phenol derivatives with reported toxicity values (pT) were taken from literature[11], these molecules were considered to carry out the 3D-QSTR analysis, 35 molecules are selected to propose the quantitative model (training set), and 8 compounds have been selected randomly and have served to test the performance of the proposed model (test set) (Table 1).

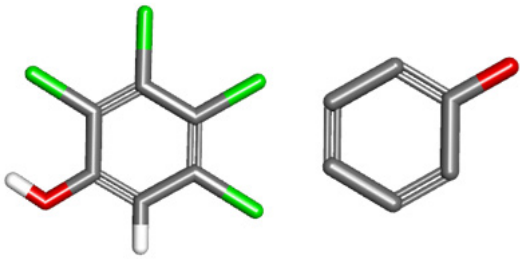
Table 1: Chemical name and toxicity values(*p*T) of studied compounds

N°	Name	pT	N°	Name	pT
1	Phenol	-0.60	23	Tetrachlorohydro-quinone	1.46
2	2-chlorophenol	0.17	24	Catechol	0.27
3*	3-chlorophenol	0.36	25*	4-chlorocatechol	1.13
4	4-chlorophenol	0.46	26*	3,5-dichlorocate-chol	1.95
5	2,3-dichlorophenol	1.09	27	Resorcinol	-0.49
6	2,4-dichlorophenol	1.24	28	4-chlororesorcinol	0.27
7	2,5-dichlorophenol	1.12	29	4,6-dichlororesor-cinol	1.02
8*	2,6-dichlorophenol	0.88	30	2-methylphenol	-0.08
9	3,4-dichlorophenol	1.47	31	2,3-dimethylphenol	0.39
10	3,5-dichlorophenol	1.67	32	2,4-dimethylphenol	0.44
11	2,3,4-trichlorophe-nol	1.64	33	2,5-dimethylphenol	0.33
12	2,3,5-trichlorophe-nol	1.86	34*	2,6-dimethylphenol	0.14
13	2,3,6-trichlorophe-nol	1.51	35*	3,4-dimethylphenol	0.58
14	2,4,5-trichlorophe-nol	1.67	36	3,5-dimethylphenol	0.51
15	2,4,6-trichlorophe-nol	1.53	37*	4-chloro-3-methyl-phenol	1.17
16	3,4,5-trichlorophe-nol	2.18	38	2-nitrophenol	1.12
17	2,3,4,5-tetrachloro-phenol	2.34	39	3-nitrophenol	0.70
18	2,3,4,6-tetrachloro-phenol	1.45	40	4-nitrophenol	1.23
19	2,3,5,6-tetrachloro-phenol	1.43	41	2,4-dinitrophenol	1.05
20	Pentachlorophenol	1.45	42	2,5-dinitrophenol	1.81
21	Hydroquinone	0.02	43*	5-methyl-2-nitro-phenol	1.15
22	Chlorohydroqui-none	1.13			

* Test set.

2.2. Minimization and alignment

Molecular structures sketched with sketch module in SYBYL [12,13] were minimized using Tripos force field [14] with the Gasteigere-Huckel charges [15]and conjugatedgradient convergence criteria of 0.01 kcal/mol (gradient method). Simulated annealing on the energy minimized structures was used with 20 cycles.In this study, the studied compounds were aligned on the common core (all compounds in test ant training set)by distil alignment technique available in SYBYL. Compound N°17, which was the most active compound (with highest toxicity (pT) value,was used as template (Fig.1).



Template (Molecule N° 17) Core

Figure. 1: 3D structure of the core and the templates (molecule N° 17).

2.3. CoMFA and CoMSIA studies

CoMFA studies were performed to analyze the specific contributions of steric and electrostatic effects. The calculation of these interactions waseffected using the Tripos force field with a distance-dependent dielectric constant at all interactions in a regularly spaced (2Å) grid taking a sp3 carbon atom as steric probe and a+1 charge as electrostatic probe. The cutoff was set to 30 kcal/mol [16].

CoMSIA method is an extension of CoMFA that uses, in addition to the steric and electrostatic interaction fields, a hydrophobic, a hydrogen bond acceptor and donor fields. The two techniques also differ in the way molecular interaction fields are implanted. They generally give comparable results, but CoMSIA models are often richer and easier to interpret[17,18].

2.4. Partial least squares analysis (PLS)

3D-QSTR models were generated using a series of 35compounds (training set). PLS method was used in this study to determine the optimal numbers of components using cross-validated coefficient Q (with leave-one-out (LOO) procedure). The final analysis (non-cross-validated analysis) was carried out using the optimum number of components obtained from the leave-one-out (LOO) cross-validation analysis to get correlation coefficient r. The best QSTR model was chosen on the basis of a combination of Q2 and r2[19,20].

2.5. Validation and predictive power of the model

The main objective of any QSTR study is to obtain a model with the highest predictive and generalization abilities. So, predictive power of the constructed models was examined using a test set of 8 molecules (Table 1)[13,21]. These compounds were aligned using the same methods described above, and then their toxicity values were predicted using the generated CoMFA and CoMSIA models.

Results and discussions

3.1. Molecular alignment

Alignment of training and test set compounds using distill module is shown in Fig. 2.

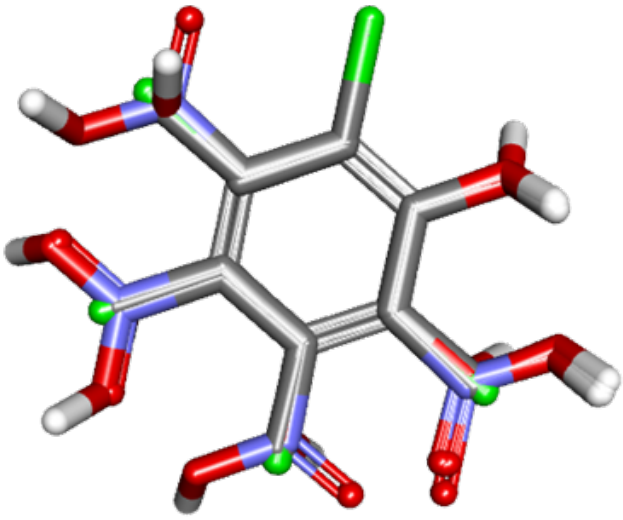
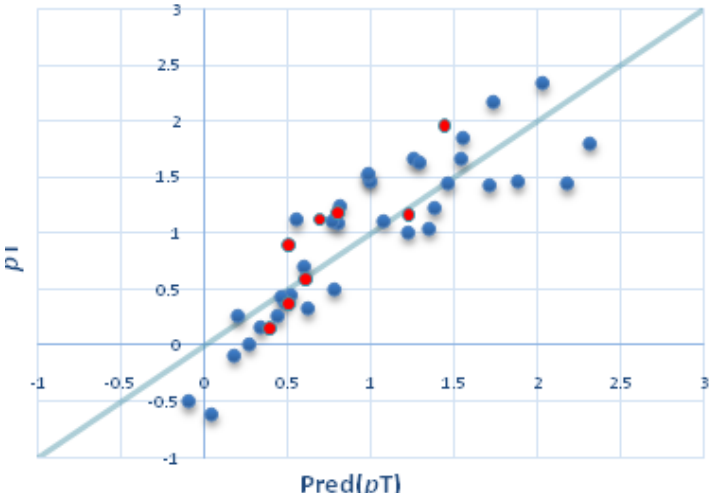


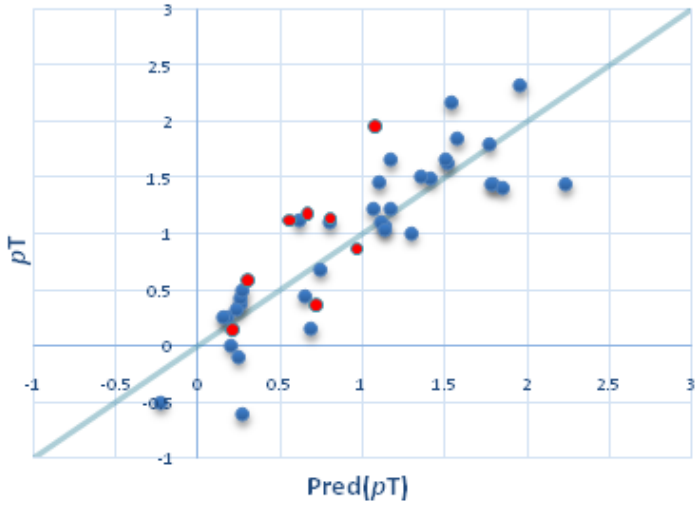
Figure.2: 3D-QSTR structure superposition and alignment of studied compounds (training and test set) using molecule N°17 as a template.

3.2. CoMFA and CoMSIA results

The 3D-QSTR models were generated from CoMFA and CoMSIA analyses. The correlations of predicted and observed toxicity values are illustrated in Fig. 3.



CoMFA analysis



CoMSIA analysis

Figure. 3: Correlations of observed and predicted pT derived from CoMFA and CoMSIA models (training set in blue; test set in red).

We use cross-validation as an internal test of the quality of the PLS models. True predictive power of a QSTR model is to test their ability to predict accurately the toxicity of compounds from an external test set (compounds which were not used for the model development), the toxicityof the remained set of 8 compounds are deduced from the quantitative model proposed with the 35 molecules (training set) by CoMFA and CoMSIA models (Fig. 3).The statistical parameters generated from CoMFA and CoMSIA analyses are listed in Table 2.

Table 2: PLS Statistics of CoMFA and CoMSIA models.

						Fractions				
Model	Q2	r2	Scv	N	r2test	Ster	Elec	Acc	Don	Hyd
CoMFA	0.559	0.760	0.368	2	0.761	0.552	0.448	-	-	-
CoMSIA	0.680	0.778	0.354	2	0.534	0.010	0.491	0.172	0.087	0.241

Q^2: Cross-validated determination coefficient; N: Optimum number of components obtained from cross-validated PLS analysis and same used in final non-cross-validated analysis; r^2: Non-cross-validated determination coefficient; Scv: Standard error of the estimate; r^2_{test}: External validation determination coefficient.

The obtained 3D-QSTR models gave good statistical results in terms of r2 values(r2 = 0.760 and r2 = 0.778 for CoMFA and CoMSIA models, respectively).The two approaches have good predictive capability gives good results (Q2 = 0.559 and Q2 = 0.680 for CoMFA and CoMSIA models, respectively). The models were able to establish a satisfactory relationship between the molecular descriptors and the toxicity of the studied compounds. The results obtained by 3D-QSTR analysesare sufficient to conclude the performance of the models; it's confirmed by the test done with the 8 compounds (r2test= 0.761 and r2test= 0.534 for CoMFA and CoMSIA models, respectively). The values of predicted and observed toxicity are presented in Table 3.

Table 3: Experimental and calculated toxicity (pT) of compounds in the training set and the test set for CoMFA and CoMSIA models.

N°	*p*T (Obs.)	*p*T (Calc.)				N°	*p*T (Obs.)	*p*T (Calc.)			
		CoMFA	Residu	CoMSIA	Residu			CoMFA	Residu	CoMSIA	Residu
1	-0.60	0.036	-0.636	0.272	-0.872	23	1.46	1.877	-0.417	1.789	-0.329
2	0.17	0.328	-0.158	0.686	-0.516	24	0.27	0.201	0.069	0.181	0.089
3*	0.36	0.509	-0.149	0.728	-0.368	25*	1.13	0.689	0.441	0.569	0.561
4	0.46	0.520	-0.060	0.652	-0.192	26*	1.95	1.440	0.510	1.090	0.860
5	1.09	0.801	0.289	1.137	-0.047	27	-0.49	-0.099	-0.391	-0.226	-0.264
6	1.24	0.807	0.433	1.068	0.172	28	0.27	0.431	-0.161	0.153	0.117
7	1.12	1.076	0.044	1.115	0.005	29	1.02	1.222	-0.202	1.293	-0.273
8*	0.88	0.509	0.371	0.964	-0.084	30	-0.08	0.171	-0.251	0.242	-0.322
9	1.47	0.993	0.477	1.106	0.364	31	0.39	0.476	-0.086	0.260	0.130
10	1.67	1.257	0.413	1.168	0.502	32	0.44	0.460	-0.020	0.260	0.180
11	1.64	1.290	0.350	1.523	0.117	33	0.33	0.612	-0.282	0.235	0.095
12	1.86	1.549	0.311	1.576	0.284	34*	0.14	0.401	-0.261	0.221	-0.081
13	1.51	0.979	0.531	1.417	0.093	35*	0.58	0.612	-0.032	0.308	0.272
14	1.67	1.542	0.128	1.508	0.162	36	0.51	0.776	-0.266	0.268	0.242
15	1.53	0.986	0.544	1.352	0.178	37*	1.17	0.809	0.361	0.665	0.505
16	2.18	1.736	0.444	1.547	0.633	38	1.12	0.765	0.355	0.801	0.319
17	2.34	2.024	0.316	1.959	0.381	39	0.70	0.596	0.104	0.745	-0.045
18	1.45	1.464	-0.014	1.800	-0.350	40	1.23	1.381	-0.151	1.173	0.057
19	1.43	1.714	-0.284	1.853	-0.423	41	1.05	1.351	-0.301	1.137	-0.087
20	1.45	2.182	-0.732	2.234	-0.784	42	1.81	2.313	-0.503	1.768	0.042
21	0.02	0.259	-0.239	0.194	-0.174	43*	1.15	1.224	-0.074	0.804	0.346
22	1.13	0.546	0.584	0.614	0.516						

* Test set.

CoMFA and CoMSIA contour plots were able to identify molecular fragments, functional groups and physicochemical properties strongly correlated with the toxicity of this series. CoMFA steric and electrostatic contours are shown in Fig. 4.

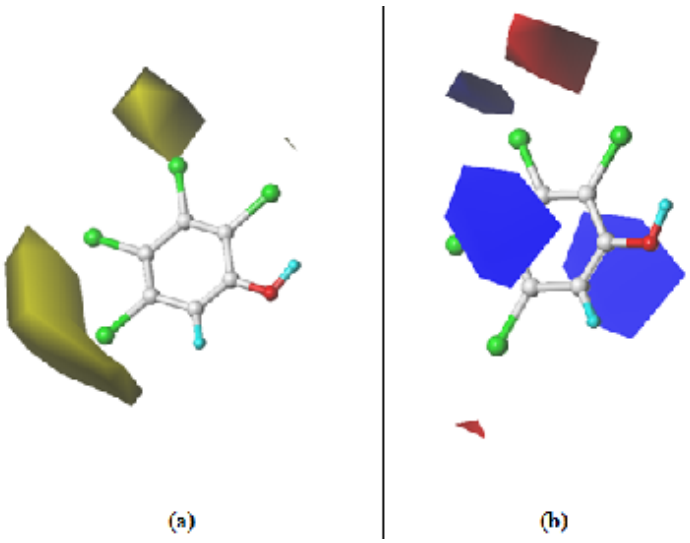


Fig.4: Std*coeff. contour maps of CoMFA analysis with 2 Å grid spacing in combination with compound N° 17. (a) Steric fields: green contours (80% contribution) indicate regions where bulky groups increase activity, while yellow contours (20% contribution) indicate regions where bulky groups decrease activity. (b) Electrostatic fields: blue contours (80% contribution) indicate regions where more positive/less negative electrostatic potential increase activity, while red contours (20% contribution) indicate regions where more negative/less positive electrostatic potential-increase activity.

The steric interaction is represented by green and yellow contours, while electrostatic interaction is denoted by red and blue contours (Fig. 4). The bigger yellow region observed around position 3, 4 and 5 (the carbon to which the initial –OH is bonded is counted as the first position) (Fig. 4a) suggesting that groups with steric tolerance are not required at this position, which means to decrease the toxicity.

CoMFA electrostatic contour plot is displayed in Fig. 4b. A blue contour indicate that substituents should be electron deficient and red color indicates that substituents should be electron rich[7]. The blue contour near position 3 (Fig. 4b) indicates that groups with positive charges may increase the activity. The electrostatic contour map displays a region of red contours neighbor to position 5, indicating that groups with negative charges are beneficial for activity in this area. Also we can comment that the blue contours above and below the plane of the ring suggests that lower electron density and hence lower negative electrostatic potential in the Pi cloud of the ring is generally associated with more activity.

For CoMSIA contour maps (Fig. 5), the steric and electrostatic field contours were constantly similar to the corresponding CoMFA contour maps[18]. Therefore, our following discussion will focus on the hydrophobic (Fig. 5(a)), hydrogen bond donor (Fig. 5(b)) and acceptor (Fig. 5(c)) fields.

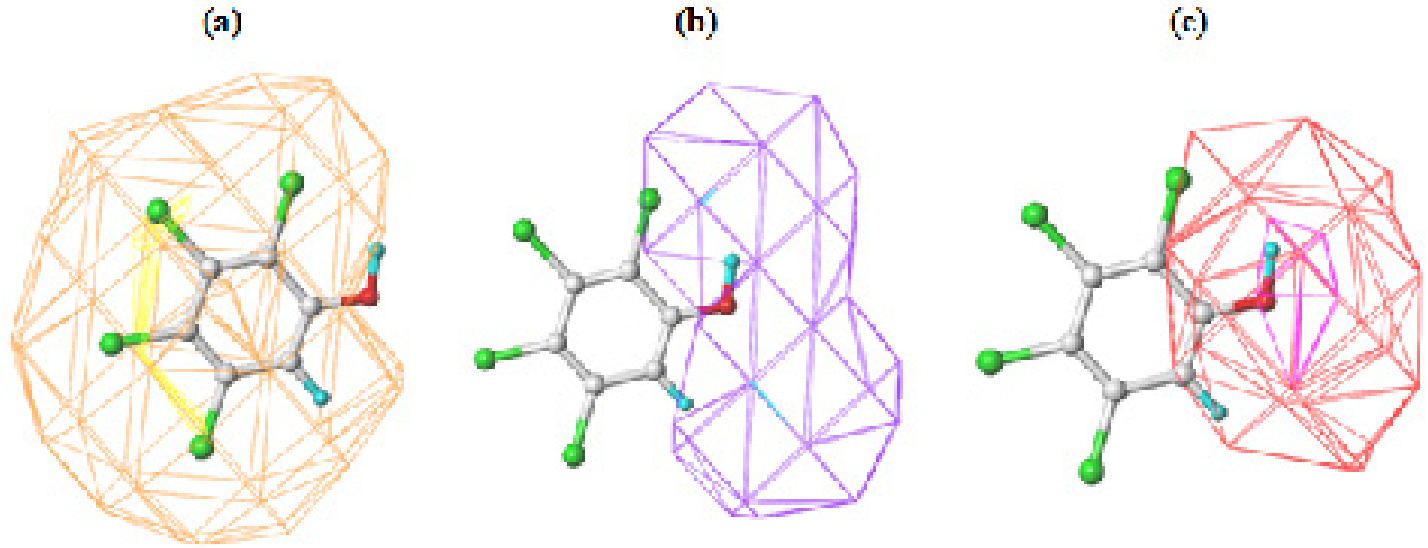


Fig.5: Std* coeff. contour maps of CoMSIA analysis with 2 Å grid spacing in combination with compound N°17. (a) Hydrophobic fields: yellow contours (80% contribution) indicate regions where hydrophobic substituents groups increase activity, while orange contours (20% contribution) indicate regions where hydrophobic substituents decrease activity. (b) Hydrogen bond donor field: cyan contours (80% contribution) regions where hydrogen bond donor groups increase activity, purple contours (20% contribution) regions where hydrogen bond donor groups decrease activity, (c) Hydrogen bond acceptor fields: magenta contours (80% contribution) indicate regions where hydrogen-bond acceptor groups favor activity, while red contours (20% contribution) indicate regions where hydrogen bond acceptor groups decrease activity.

A yellow contour near position 3, 4 and 5 means that hydrophobic substituent groups such as alkyl groups may increase the toxicity (Fig. 5(a)). A purple contour near position 1, 2 and 6means that groups with hydrogen bond donor decrease the toxicity (Fig. 5(b)). The magenta contour near position 1 indicates that groups with hydrogen bond acceptor groups increase the toxicity (Fig. 5(c)). While a red contour near to position 1, 2 and 6 means that groups with hydrogen bond acceptor groups decrease the activity (Fig. 5(c)).

All these findings may be used to design compounds with more or less toxicity values. As observed in CoMFA and CoMSIA maps, by adding suitable substituents.

Conclusion

In this study, 3D-QSTR analyses were used to predict the toxicity of a set of 43substituted phenols to *Chlorella Vulgaris*. In order to establish 3D-QSTR models, a Partial Least Squares (PLS) statistical method are used in deriving the Comparative Molecular Field Analysis (CoMFA) and Comparative Molecular Similarity Indices Analysis (CoMSIA). In conclusion, the established 3D-QSTR models gave good statistical results in terms of Q2,r2and rtest-2values. CoMFAand CoMSIAmodels exhibited high internal and external consistency as demonstrated by the several validation methods employed to assess their statistical quality.

Through this study; we have been able to develop models that will be used later to propose new structures of phenolic molecules with controlled toxicity. This work shows the interest of 3D-QSTR studies that allow the correlation between the experimental and theoretical results,to predict the toxicity of other molecules of the same series.

Acknowledgements

We are grateful to the "Association Marocaine des Chimistes-Théoriciens" (AMCT) for its pertinent help concerning the programs.

References

1. W. Duan, F. Meng, H. Cui, Y. Lin, G. Wang, J. Wu (2018)Ecotoxicity of phenol and cresols to aquatic organisms: A review, Ecotoxicology and Environmental Safety 157: 441–456. doi:10.1016/j.ecoenv.2018.03.089.
2. C. Selassie, R.P. Verma (2015) QSAR of toxicology of substituted phenols, Journal of Pesticide Science 40: 1–12.
3. E.R. Christensen, K.O. Kusk, N. Nyholm (2009) Dose–response regressions for algal growth and similar continuous endpoints: calculation of effective concentrations, Environmental Toxicology and Chemistry 28: 826–835.
4. M.D. Ertürk, M.T. Saçan (2013) Assessment and modeling of the novel toxicity data set of phenols to Chlorella vulgaris, Ecotoxicology and Environmental Safety 90: 61–68.
5. E. Commission, Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC, In: Official Journal, L 396/1, 30/12/2006. (2006).
6. T.I. Netzeva, M. Pavan, A.P. Worth (2018) Review of (quantitative) structure–activity relationships for acute aquatic toxicity, QSAR & Combinatorial Science 27: 77–90.
7. A. Belhassan, S. Chtita, T. Lakhlifi, M. Bouachrine (2018) 3D-QSAR Study of Biodegradability in Water for Aromatic Compounds, RHAZES: Green and Applied Chemistry 1: 21–30.
8. Correlation Structure-Property in the case of Odorant Molecules: QSPR study and Molecular Docking (Paperback) par Assia Belhassan, Mohammed Bouachrine, Tahar Lakhlifi: Noor Publishing, Germany 9786202359597 Paperback, Aufl. - Book Depository hard to find, (n.d.). https://www.abebooks.fr/Correlation-Structure-Property-case-Odorant-Molecules-QSPR/30230220509/bd (accessed March 17, 2019).
9. T. Wang, X. Yuan, M.-B. Wu, J.-P. Lin, L.-R. Yang (2017) The advancement of multidimensional QSAR for novel drug discovery-where are we headed?, Expert Opinion on Drug Discovery 12: 769–784.
10. A.S. de Souza, L.G. Ferreira, A.D. Andricopulo (2017) 2D and 3D QSAR Studies on a Series of Antichagasic Fenarimol Derivatives, International Journal of Quantitative Structure-Property Relationships (IJQSPR) 2: 44–63.
11. G. Tugcu, M.D. Ertürk, M.T. Saçan (2017)On the aquatic toxicity of substituted phenols to Chlorella vulgaris: QSTR with an extended novel data set and interspecies models, Journal of Hazardous Materials 339: 122–130.
12. Tripos Inc., St. Louis, MO, USA, SYBYL-X 2.0, (n.d.). http://www.tripois.com.
13. A. Aouidate, A. Ghaleb, M. Ghamali, S. Chtita, M. Choukrad, A. Sbai, M. Bouachrine, T. Lakhlifi (2017) Combined 3D-QSAR and molecular docking study on 7, 8-dialkyl-1, 3-diaminopyrrolo-[3, 2-f] Quinazoline series compounds to understand the binding mechanism of DHFR inhibitors, Journal of Molecular Structure 1139: 319–327.
14. M. Clark, R.D. Cramer, N. Van Opdenbosch (1989) Validation of the general purpose Tripos 5.2 force field, Journal of Computational Chemistry 10: 982–1012.
15. W.P. Purcell, J.A. Singer (1967)A brief review and table of semiempirical parameters used in the Hueckel molecular orbital method, Journal of Chemical and Engineering Data. 12: 235–246.
16. L. Ståhle, S. Wold, 6 multivariate data analysis and experimental design in biomedical research, in: Progress in Medicinal Chemistry, Elsevier, 1988: pp. 291–338.
17. B.L. Bush, R.B. Nachbar (1993) Sample-distance partial least squares: PLS optimized for many variables, with application to CoMFA, Journal of Computer-Aided Molecular Design. 7: 587–619.
18. H. Zaki, A. Belhassan, A. Aouidat, A. Taourati, T. Lakhlifi, M. Benlyas, M. Bouachrine (2018) In–silico Study of 3-(2-amino-6-phenylpyrimidin-4-yl)-Ncyclopropyl-1-methyl-1H-indole-2-carboxamide derivatives as antifungal agents: CoMFA, CoMSIA analyses, Molecular Docking and ADMET Properties, RHAZES: Green and Applied Chemistry 1: 64–76.
19. V. Srivastava, A. Kumar, B.N. Mishra, M.I. Siddiqi, CoMFA and CoMSIA (2008) 3D-QSAR analysis of DMDP derivatives as anti-cancer agents, Bioinformation. 2: 384.

20. A Ghaleb, A Aouidate, M Ghamali, A Sbai, MBouachrine, T Lakhlifi (2017) 3D-QSAR modeling and molecular docking studies on a series of 2, 5 disubstituted 1, 3, 4-oxadiazoles, Journal of Molecular Structure 1145: 278–284.
21. A. Belhassan, S. Chtita, T. Lakhlifi, M. Bouachrine (2017) QSPR study of the retention/release property of odorant molecules in water using statistical methods, Orbital: The Electronic Journal of Chemistry9: 234–247.