

## Structure-Activity (Flotation) Relationship Modeling of Flotation of Sphalerite by N-Arylhydroxamic Acids

N. Ganapathy Subramanian<sup>1</sup>, R. Natarajan<sup>2\*</sup>, I. Nirdosh<sup>2</sup> and R. Jayalakshmi<sup>3</sup>

<sup>1</sup>Department of Mechanical Engineering, Mepco Schlenk Engineering College,  
Sivakasi, Tamil Nadu, India 626 005

<sup>2</sup>Department of Chemical Engineering, Lakehead University,  
Thunder Bay, Ontario, Canada P7B 5E1

<sup>3</sup>Department of Bioinformatics, Thanthai Hans Roever College,  
Perambalur, Tamil Nadu, India 621 212

\* Present Address: Natural Resources Research Institute, University of Minnesota Duluth,  
5013 Miller Trunk Highway, Duluth, MN 55811, USA

### Abstract

Molecular structure is known to play a vital role in the efficiency of chleating collectors in mineral flotation. Hence, flotation efficiencies of congeneric organic compounds used as mineral collectors are amenable to QSAR modeling. Sphalerite grade of the floats of a set of flotation tests conducted with a copper-zinc ore using a series of twenty seven N-arylhydroxamic acids of the generic structure  $Ar-N(OH)C(=O)-R$  ( $R = \text{aryl/alkyl/aaralkyl}$ ) were modeled using calculated molecular descriptors such as, topochemical, topostructural, quantum chemical, and geometrical parameters. In addition to these molecular descriptors, calculated physicochemical properties such as octanol-water partition coefficient ( $ClogP$ ), and organic carbon-water partition coefficient ( $\log K_{oc}$ ) were also used to build the regression models. The collectors were classified into C-aryl, C-alkyl, and C-aralkyl. Octanol-water partition coefficient ( $ClogP$ ) was found to give the best quadratic fit for C-aryl, and the combined set of C-aralkyl and C-alkyl. It was interesting to note that the data for individual sets namely, C-alkyl, and C-aralkyl gave linear fits with positive and negative slopes, respectively. This indicated that the points were distributed on the right hand and left hand sides of the parabola that fits the combined data set.

### INTRODUCTION

Biological activities as well as physicochemical and chemical properties of organic compounds depend on molecular structure. The quantitative-relation that bring out the dependence of a property or activity (P) on entire molecular structure or its sub-structural fragment, is called Quantitative Structure-Activity Relationship (QSAR).

$$P = a_1 \times S_1 + a_2 \times S_2 + \dots$$

Where  $a_1, a_2$  etc., are the coefficients and  $S_1, S_2$  are non-empirical structural property or an empirical property of the whole molecule or any of its sub-structural unit. QSAR approach is extensively used in medicinal chemistry and is a proven tool known to reduce the time taken in the selection of lead compounds and lead optimization. Development of QSAR models with good predictive abilities rely on the effective transformation of structural peculiarities of molecules into numerical quantities. Topological indices (TIs) are extensively used to encode structural details into numbers, i.e., they are the molecular metrics. There are over 300 topological indices and their definitions and calculations

can be obtained from the monograph by Devillers and Balaban (1999). In addition to TIs, quantum mechanical parameters such as energy of Highest Occupied Molecular Orbital ( $E_{HOMO}$ ), energy of the Lowest Unoccupied Molecular Orbital ( $E_{LUMO}$ ), and partial atomic charges are also used in forming QSAR models.

There is implicit dependence of efficiency of a chelating agent (organic compound) as flotation collector on the molecular architecture. Though the effect structure on the performance of a chemical is known, QSAR is the only approach that quantifies this intuitiveness. Natarajan et al. (1999) visualized that the extension of QSAR approach to model the efficiency of chelating collectors would eliminate the trial-and-error type exploratory tests. They successfully showed that the separation efficiencies of chelating collectors are amenable to QSAR modeling and topological indices, electronic parameters, physicochemical properties such as octanol-water coefficient ( $\log P$ ) and soil-water coefficient ( $\log K_{oc}$ ) and geometrical parameters that describe the molecular surface area and volume could be used to form the QSAR models. The success of the approach led to the emergence of a scientific tool for the selection of collectors.

Due to the large number of TIs and their mutual relatedness, Principal Component Analysis (PCA) was performed to extract the minimum number of orthogonal parameters. Principal Components (PC) extracted were used as independent parameters to form the regression equations. Results of these studies were recently reviewed by Natarajan et al. 2003 and regression equations were developed for the following collectors: substituted-cupferron in the flotation of a uranium ore (Natarajan et al. 2001, 2002a, 2002b), substituted-mercaptobenzothiazoles (MBT) in the flotation of lead mineral from an oxidized lead-zinc ore (Natarajan et al. 2002a, 2002b), *o*-Aminothiophenols (ATP) in the flotation of zinc mineral from an oxidized lead-zinc ore (Natarajan et al. 2002a, 2002b), Aylhydroxamic acids (*N*-aryl-*C*-aryl & *N*-aryl-*C*-alkyl combined) in the flotation of copper minerals from a copper-zinc sulfide ore (Natarajan and Nirdosh, 2003).

The important findings of the QSAR modeling are:

1. Topological indices or the Principal Components extracted from them could be used as independent parameters to describe the molecular architecture.
2. In the case of congeneric collectors, (collectors having the same chelating group), introduction of electronic and quantum mechanical parameters did not improve the predictive ability significantly.
3. Computed soil-water partition coefficient,  $\log K_{oc}$ , and octanol-water partition coefficient ( $\log P$ ) were found to be good computable property parameters that gave QSAR equations of good predictive ability.
4. The regression equation formed had good predictive ability and the differences between predicted and experimental separation efficiencies were less than the experimental errors.
5. When the chelating group was at the end of the hydrocarbon chain an increase in carbon chain increased the efficiency of the collector. However, increase in the chain length resulted in significant decrease in the solubility of the compound in water and this became a limiting factor in increasing the length of the hydrocarbon part. However, when the chelating group was in the middle of the hydrocarbon chain there was a maximum size beyond which increase in size of the alkyl chain or the alkyl group in an aromatic ring decreased the efficiency.
6. QSAR modeling had lead to the selection of a few compounds from a large list of possible structures with a given chelating group.

The selection of a small number of compounds from a virtual data base of compounds was illustrated by Natarajan and Nirdosh (2004). This method involves molecular-similarity based selection of compounds using the principal components extracted from 100 TIs.

This paper is an extension of the application of QSAR approach to model the flotation efficiency of *N*-arylhydroxamic acids used to float sphalerite in two stage flotation experiments (Natarajan and Nirdosh, 2006). Several calculated molecular descriptors, semi empirical quantum chemical parameters and physicochemical properties were used to model. The best fitting models and their significance are discussed.

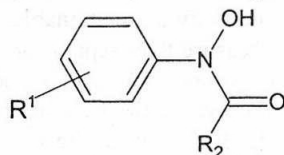
**Table 1: Symbols, Definitions and Classification of Structural Molecular Descriptors**

<b>Topostructural (TS)</b>	
$I_D^W$	Information index for the magnitudes of distances between all possible pairs of vertices of a graph
$\bar{I}_D^W$	Mean information index for the magnitude of distance
$W$	Wiener index = half-sum of the off-diagonal elements of the distance matrix of a graph
$\overline{IC}$	Information content of the distance matrix partitioned by frequency of occurrences of distance $h$
$M_1$	A Zagreb group parameter = sum of square of degree over all vertices
$M_2$	A Zagreb group parameter = sum of cross-product of degrees over all neighboring (connected) vertices
${}^h\chi$	Path connectivity index of order $h = 0-10$
${}^h\chi_C$	Cluster connectivity index of order $h = 3-6$
${}^h\chi_{PC}$	Path-cluster connectivity index of order $h = 4-6$
$P_h$	Number of paths of length $h = 0-10$
$J$	Balaban's $J$ index based on topological distance
<b>Topochemical (TC)</b>	
${}^h\chi^b$	Bond path connectivity index of order $h = 0-6$
${}^h\chi_C^b$	Bond cluster connectivity index of order $h = 3-6$
${}^h\chi_{PC}^b$	Bond path-cluster connectivity index of order $h = 4-6$
${}^h\chi^v$	Valence path connectivity index of order $h = 0-10$
${}^h\chi_C^v$	Valence cluster connectivity index of order $h = 3-6$
${}^h\chi_{PC}^v$	Valence path-cluster connectivity index of order $h = 4-6$
<b>Geometrical</b>	
$SAA$	Connolly Solvent Accessible Surface Area (angstrom <sup>2</sup> ) = The locus of the center of a solvent probe (sphere) as it is rolled over the van der Waal surface of a molecule
$MSA$	Connolly Molecular Surface Area (angstrom <sup>2</sup> ) = The contact surface created when a solvent probe (sphere) is rolled over van der Waal surface a molecule
$V_{ex}$	Connolly Solvent Excluded Volume (angstrom <sup>3</sup> ) = The volume contained within the molecular surface
$Oval$	Ovality = The ratio of the molecular surface area to the minimum surface area
<b>Quantum chemical</b>	
$E_{HOMO}$	Energy of the highest occupied molecular orbital
$E_{LUMO}$	Energy of the lowest unoccupied molecular orbital
$IP$	Molecular ionization energy = $E_{HOMO} - E_{LUMO}$
$\mu$	Dipole moment
<b>Calculated Physicochemical Property</b>	
$\log K_{oc}$	Soil-water partition coefficient
$ClogP$	Calculated waterOctanol partition coefficient

**MATERIALS AND METHOD**

Twenty seven *N*-arylhydroxamic acids were synthesized according to a reported procedure (Priyadarshini and Tandon, 1967) and were tested as collectors to float sphalerite from a copper-zinc ore from Canada. The details of the flotation tests and results were reported by Natarajan and Nirdosh (2006). Grades of the floats and the collector names are given in Table 1.

**Table 2: Structures of Collectors Tested and the Sphalerite Grade of the Floats**



#	R <sup>1</sup>	R <sup>2</sup>	Abbreviation	ClogP	Sp Grade%
<b>Type I. N-aryl-C-alkyl</b>					
1	H	n-C <sub>3</sub> H <sub>7</sub>	NBuPHA	2.721	23.82
2	4-C <sub>2</sub> H <sub>5</sub>	n-C <sub>3</sub> H <sub>7</sub>	NBuEPHA	3.769	22.09
3	H	C(CH <sub>3</sub> ) <sub>3</sub>	TMAPHA	2.900	24.76
4	H	CH <sub>2</sub> C(CH <sub>3</sub> ) <sub>3</sub>	tBuAPHA	3.519	30.15
5	2,6-diCH <sub>3</sub>	n- C <sub>3</sub> H <sub>7</sub>	NBuXHA	3.759	29.49
6	2,6-diCH <sub>3</sub>	C(CH <sub>3</sub> ) <sub>3</sub>	TMAXHA	3.938	24.66
7	2,6-diCH <sub>3</sub>	CH <sub>2</sub> C(CH <sub>3</sub> )	tBuAXHA	4.557	27.38
8	4-C <sub>2</sub> H <sub>5</sub>	C(CH <sub>3</sub> ) <sub>3</sub>	TMAEPHA	3.948	27.89
9	H	n-C <sub>5</sub> H <sub>11</sub>	NHexPHA	3.779	27.41
10	4-Br	n- C <sub>3</sub> H <sub>7</sub>	NBuBPHA	2.694	20.94
<b>Type II. N-Aryl-C-aryl</b>					
11	H	C <sub>6</sub> H <sub>5</sub>	NBPHA	3.694	22.15
12	H	4-CH <sub>3</sub> O-C <sub>6</sub> H <sub>4</sub>	NAnPHA	2.323	26.57
13	H	4NO <sub>2</sub> -C <sub>6</sub> H <sub>4</sub>	NNBPHA	1.318	23.29
14	4-F	C <sub>6</sub> H <sub>5</sub>	NBFPHA	3.087	19.76
15	2-OCH <sub>3</sub>	C <sub>6</sub> H <sub>5</sub>	NBoMPHA	3.316	21.54
16	4-F	4-CH <sub>3</sub> O-C <sub>6</sub> H <sub>4</sub>	NAnFPHA	1.716	22.54
17	4-C <sub>2</sub> H <sub>5</sub>	C <sub>6</sub> H <sub>5</sub>	NBEPHA	4.742	21.36
18	2,6-diCH <sub>3</sub>	C <sub>6</sub> H <sub>5</sub>	NBXHA	4.732	24.47
19	4-n-C <sub>3</sub> H <sub>7</sub>	C <sub>6</sub> H <sub>5</sub>	NBPPHA	5.271	18.47
20	4-i-C <sub>2</sub> H <sub>7</sub>	C <sub>6</sub> H <sub>5</sub>	NBiPPHA	5.141	18.99
21	4-t-C <sub>4</sub> H <sub>9</sub>	C <sub>6</sub> H <sub>5</sub>	NBtBuPHA	5.540	18.71
<b>Type III. N-Aryl-C-aralkyl</b>					
22	H	C <sub>6</sub> H <sub>5</sub> CH <sub>2</sub>	PANPHA	4.313	29.25
23	H	C <sub>6</sub> H <sub>5</sub> CH <sub>2</sub> CH <sub>2</sub>	HCNPHA	4.842	28.51
24	4-C <sub>2</sub> H <sub>5</sub>	C <sub>6</sub> H <sub>5</sub> CH <sub>2</sub>	PANEPHA	5.361	24.9
25	4-C <sub>2</sub> H <sub>5</sub>	C <sub>6</sub> H <sub>5</sub> CH <sub>2</sub> CH <sub>2</sub>	HCNEPHA	5.890	22.11
26	2,6-diCH <sub>3</sub>	C <sub>6</sub> H <sub>5</sub> CH <sub>2</sub>	PANXHA	5.880	25.69
27	2,6-diCH <sub>3</sub>	C <sub>6</sub> H <sub>5</sub> CH <sub>2</sub> CH <sub>2</sub>	HCNXHA	5.351	22.26

Several topological indices were calculated using an in-house software *INDCAL* (Natarajan et al, 2002c), geometrical parameters such as molecular surface area, solvent accessible surface area, solvent excluded surface area, and semi empirical quantum chemical parameters namely E<sub>HOMO</sub>, E<sub>LUMO</sub>, heat of formation, and dipole moment and octanol water partition coefficient ClogP were calculated using Chem 3D (CambridgeSoft 2004). Soil-water partition coefficient logK<sub>oc</sub> was

computed using Toolkit for estimating physicochemical properties of organic compounds (Reinhard et al. 1999). A list of calculated parameters and their brief definitions are given in Table 2. However, the numerical values of all these parameters are not given for the sake of brevity.

## RESULTS AND DISCUSSION

In the earlier attempts on QSAR modeling (Natarajan et al. 2003b) of performance of chelating collectors separation efficiency was used as the dependant parameter. Though collectors have selectivity between minerals, absolute specificity is unattainable. Hence, in a poly mineral dispersion grade of the float seemed to be a better measure than separation efficiency. Moreover, the data used to model here were from two-stage flotation experiments. In each of these tests a chalcopyrite rich float was obtained as a collectorless prefloat such that the collectors were adsorbed mostly on sphalerite. In the current situation grade of the sphalerite concentrate (float) was considered as a better measure of collector's efficiency. Sphalerite grades (%) of the floats are given in Table 1 along with the abbreviated names of the hydroxamic acids used in the study.

When the sphalerite grades obtained for all the 27 collectors were treated together only a scatter plot was obtained and none of the available parameters was found to model all the data pooled as a single set. The data were split into three sets viz., *C*-alkyl (set-I), *C*-aryl (set-II), and *C*-aralkyl (set-III) and curve fitting was repeated using the calculated parameters. Of all the parameters *ClogP* was found to fit well with all the three data sets, this parameter represent the hydrophobicity of the collectors. For set-I and set-II linear regression models were obtained but the most interesting aspect was that the slope of the linear equation for set-I had a positive slope ( 3.33)while that for set-III had a negative slope (-0.158). The linear regression equations are:

$$\text{For Set I: Sp.grade} = 14.86 + 3.33 * \text{Clog}P$$

$$n = 8, r^2 = 0.511; F = 6.28; \text{SEE} = 2.33$$

$$\text{For Set III: Sp.grade} = 9.29 - 0.158 * \text{Clog}P$$

$$n = 6, r^2 = 0.605; F = 6.14; \text{SEE} = 0.43$$

A phenomenon similar to this was observed in biological activities (Lein and Hansch, 1972) and when the complete structure space was included a parabolic relation was obtained. Thus, a parabolic relation might result in straight lines of positive and negative slopes depending on the distribution of the data points on the parabola. This was exactly matched our observation when we combined set-I and set-III, i.e., combined set of *C*-alkyl and *C*-arklyl. See Figure 1 for the parabola obtained for the combined data set of I and II (Set-I is in filled circles, and Set-III is in open circles). However, two of the data points (NBuEPA and TMAXHA) were to be dropped as outliers. The quadratic fit for the remaining data points of the combined sets of I and II is:

$$\text{Sp.grade} = 19.29 \text{Clog}P - 2.27 * (\text{Clog}P)^2 - 12.29$$

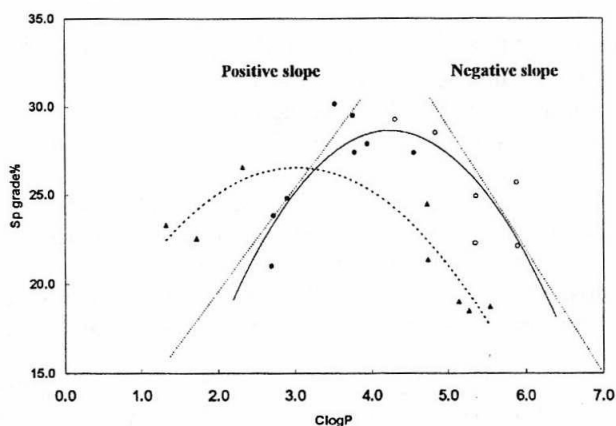
$$N = 14, r^2 = 0.647; F = 10.07; \text{SEE} = 1.92$$

Similar results had been obtained for a set of 17 similar compounds for the flotation of chalcopyrite (Natarajan, and Nirdosh, 2003). For the data set-II also *ClogP* was found to give the best quadratic fit and in this case the data seemed to be distributed to get a parabola (data points of Set II are marked in Figure 1 by filled triangles. In the case of Set-I also two outliers were observed and they were NBPHA and NBoMPHA. Even in the earlier study (Natarajan, and Nirdosh, 2003) on the flotation of chalcopyrite NBPHA was observed as an outlier, and the inability to scale the change in structure for the change of a hydrogen atom to a methyl group or a halogen was offered as a possible explanation. The other outlier NBoMPHA was the only *ortho* substituted compound used in the study. Steric factors and other secondary proximity effects normally make *ortho* substituted compound to behave differently from *meta* and *para* substituted compounds. The quadratic equation for set-II, *C*-aryl

compounds is:

$$\text{Sp. grade} = 8.53\text{Clog}P - 1.41 * (\text{Clog}P)^2 + 13.65$$

$$n = 8, r^2 = 0.795; F = 9.67; \text{SEE} = 1.59$$



**Fig. 1: Quadratic fit of sphalerite (Sp) grade with ClogP**

(Set-I, C-alkyl in filled circles; Set-II, C-aryl in filled triangles; Set-III, C-aralkyl in open circles)

It is to be noted that C-alkyl and C-aralkyl which have nearly same chemical behaviour could be grouped together while the C-aryl compounds that have two benzene rings directly attached to the chelate ring that would be formed by the collector with the metal ion formed a different group.

## CONCLUSION

It was quite a satisfying experience to have the same type of observation repeated in different studies, and it reinforces the hypothesis that flotation efficiencies of chelating collectors are amenable to QSAR modeling. Computed molecular descriptors or calculated physicochemical properties could be used for modeling is a very important observation in directing future research in the scientific design of collectors once the chelating group is decided based on the concepts of chelation chemistry such as stability constants. QSAR a powerful tool in pharmaceutical chemistry can become a scientific method of screening compounds from a virtual library (compounds that are yet to be synthesized may also be in the database). Considerable amount of money and time to be spent on exploratory tests can be saved. One might argue that the Molecular Recognition Mechanism (Pradip and Beena Rai, 2002, 2003 and references there in) is much better than QSAR approach. This is a computer intensive approach that involves optimization of molecular geometry of a collector using quantum mechanical calculations, surface modeling of the mineral, and the energy calculations of interactions between the collector and the surface. Screening a big database using this approach would involve thousands of hour of computation. Thus short listing a few compounds or structural analogs before going in for any optional computer intensive modeling appears to be a good strategy in developing new tailor made collectors that are more selective.

## ACKNOWLEDGEMENT

Financial assistance from Natural Sciences and Engineering Research Council of Canada (NSERC) is greatly acknowledged.

## REFERENCES

- [1] Chem3DPro 2005. Chem3D Ultra 8, CambridgeSoft Corporation, Cambridge, MA..

- [2] Devillers, J. and Balaban, A.T. (Editors) 1999. Topological Indices and Related Descriptors in QSAR and QSPR. Gordon and Breach, Amsterdam, Netherlands.
- [3] Lein, J.L., Hanssch, C., 1972. Structure-activity relationships in antifungal agents: A survey. In, Gould, R.F. Ed. Biological Correlations-The Hansch Approach, American Chemical Society, Washington, D.C., USA.
- [4] Natarajan, R., Nirdosh, I., Muthuswami, S.V., 1999. Application of topological indices to froth-flotation of a uranium ore. *Curr. Sci.* 77 (9), 1170– 1174.
- [5] Natarajan, R., Nirdosh, I., 2001. QSAR modeling of flotation collectors: Part 1. Application of valence connectivity indices to the flotation of a uranium ore using substituted-cupferrons. *Indian J. Chem.* 40(A), 130– 134.
- [6] Natarajan, R., Nirdosh, I., Venuvanalingam, P., Ramalingam, M., 2002a. Quantitative property – property relationship (QPPR) approach in predicting flotation efficiency of chelating agents as mineral collectors. *SAR QSAR Environ. Res.* 13 (5), 499– 508.
- [7] Natarajan, R., Nirdosh, I., Basak, S.C., Mills, D.R., 2002b. QSAR modelling of flotation collectors using principle components extracted from topological indices. *J. Chem. Inf. Comput. Sci.* 42, 1425–1430.
- [8] Natarajan, R., Nirdosh, I., Anbazhagan, T.M. and Murali, T.C, 2002c. Topological Index Calculator (INDCAL): A Computer Program to Calculate Topographical and Topostructural Indices, in Proceedings of the International seminar on Mineral Processing Technology, MPT 2002, 3-5 January 2002; Eds. S. Subramanian, K.A Natarajan, B.S. Rao and T.R.R. Rao, Vol. 1 . 301-306,
- [9] Natarajan, R., Nirdosh, I. 2003a. Application of topochemical, topostructural, physicochemical and geometrical parameters to model the flotation efficiencies of N-arylhydroxamic acids. *Int. J. Miner. Process.* 71, 113– 129.
- [10] Natarajan, R., Kamalakannan, P. and Nirdosh, I., 2003b. Applications of Topological Indices to Structure-Activity Relationship Modelling and Selection of Mineral Collectors.” *Indian J. Chem.*, 42A, 2003, pp. 1330-1346.
- [11] Natarajan, R., Nirdosh, I. 2004. Flotation of base metal ores with arylhydroxamic acids, Technical report submitted to Natural Sciences and Engineering Research Council of Canada (NSERC).
- [12] Natarajan, R., Nirdosh, I. 2006. New collectors for sphalerite flotation, *Int. J. Miner. Process* (in press)
- [13] Pradip, Beena Rai, B., 2002. Design of tailor-made surfactants for industrial applications using a molecular modeling approach. *Colloids and Surfaces A.* 205, 139– 148.
- [14] Pradip, Beena Rai 2003. Molecular modeling and rational design of flotation reagents *Int. J. Miner. Process.* 72, 95– 110
- [15] Priyadarshini, A., and Tandon, S.G. 1967. Preparation and properties of some N-arylhydroxamic acids, *J. Chem. Eng. Data.* 12, 143-144.
- [16] Reinhard, M. and Drefahl, A. 1999b. (Software by Baričič, P. and Mackov, M.) The Toolkit for Estimating Physicochemical Properties of Organic Compounds, Wiley, New York, U.S.A.