ACADEMIA ROMÂNĂ
Revue Roumaine de Chimie
http://web.icf.ro/rrch/

Rev. Roum. Chim.,
2011, *56*(4), 347-350

Dedicated to Professor Alexandru T. Balaban on the occasion of his 80th anniversary

PARTIAL LEAST SQUARES (PLS) ANALYSIS FOR HYDROPHILE-LIPOPHILE BALANCE (HLB) OF SOME NONIONIC SURFACTANTS. QSPR STUDY

Adrian BETERINGHE,[a,b,*] Ana Cristina RĂDUȚIU,[a] Marioara BEM,[a] Rodica Daniela BĂRĂȚOIU,[a] Elena-Nuşa HRISTEA[a] and Titus CONSTANTINESCU[a]

[a] Roumanian Academy, "Ilie Murgulescu" Institute of Physical Chemistry, Laboratory of Coordinative and Supramolecular Chemistry, Splaiul Independenței 202, Bucharest, 060021, Roumania
[b] Lumina – University of South-East Europe, Coletina Road 64B, Bucharest-021187, Roumania

Received September 15, 2010

HLB (hydrophile-lipophile balance) values for 14 commercial nonionic emulsifiers were correlated with molecular descriptors calculated using CODESSA program. Because these molecular descriptors are many and highly collinear, PLC regression was used to obtain a QSPR model with high predictive power (R^2=0.908). On this model made a mark the orthogonal Factors (latent variables) obtained by linear combinations of the original Descriptors.|

INTRODUCTION

Surfactants (surface-active agents) are most commonly used materials of the chemical industry that appear in various products such as the motor oils, pharmaceuticals, detergents, the drilling sludge used in prospecting for petroleum or flotation agents.[1] In last decades, surfactants have found applications in high-technology areas such as electronic printing, magnetic recording, biotechnology, viral research and micro-electronics.[1, 2]

In solution, surfactants have the ability to concentrate on the surface, making soluble materials that have low affinity to each other. Because of yhe phenomenon of accumulation on surfaces, surfactants, even in low concentrations, can modify the superficial tension of liquids that dissolve.[3, 4]

These properties are due to amphiphilic structure, represented by polar, weakly polar or in some cases ionisable group that ensure dissolution in water or strongly polar liquids together with a non-polar group that assure dissolution in hydrophobic environments[1, 5, 6] realizing the structures for the main type of surfactants: anionics[5, 7, 8], cationics[5, 9, 10], nonionics[5, 11] and amphoterics.[5, 12]

Of the four classes, nonionic surfactants have a number of physicochemical characteristics[11] that are different from those of ionic surfactants, recommending them for special applications.

Physicochemical properties including the superficial one of the surfactants is found in the HLB (hydrophile-lipophile balance) value, which describes in fact the structural contribution of amphiphilic structure on these properties. In the present paper we report the quantitative structure–property relationships (QSPR) for HLB values of 14 commercial nonionic emulsifiers involving PLS regression.

RESULTS AND DISCUSSION

HLB was originally defined by Griffin[13, 14] for surface-active agents with ethylene oxide moieties (eq. 1) as:

* Corresponding author: chemworks2003@yahoo.com

$$HLB = \frac{E+P}{5} \quad (1)$$

where E is the percentage weight of oxyethylene and P the percentage weight of polyhydric alcohol. Griffin's method for non-ionic surfactants as described in 1954 works[14] as follows (eq. 2):

$$HLB = 20\frac{Mh}{M} \quad (2)$$

where Mh is the molecular mass of the hydrophilic portion of the Molecule, and M is the molecular mass of the whole molecule.

In the esters, fatty acids or polyhydroxilic alcohols situation, HLB is calculated using the formula (eq. 3):[14]

$$HLB = 20\left(1-\frac{S}{A}\right) \quad (3)$$

where S is the saponification number of the ester and A is the acid number of the acid.

Lin and Marszall[15] found a relationship (eq. 4) between the partition coefficient and HLB:

$$\mathrm{HLB} = 0.36 \cdot \ln(1/\mathrm{K}p) + 7 \quad (4)$$

R.C. Little[16] reported the following relationship between the Hildebrand solubility parameter (δ) and HLB (eq. 5):

$$\mathrm{HLB} = 54\,((\delta - 8.2)/(\delta - 6)) \quad (5)$$

where δ is Hildebrand solubility parameter.

In the present paper we report the QSPR for HLB of 14 commercial nonionic emulsifiers involving PLS regression.

PLS is a new method developed by Herman Wold as a method for constructing predictive models when the explanatory variables are many and highly collinear[17, 18] and has been used as an alternative approach to the ordinary least squares regression in ill-conditioned linear regression models that arise in several disciplines such as chemistry, economics, medicine, psychology, and pharmaceutical science.[19]

In QSAR and QSPR techniques, PLS regression is used when the number of descriptors (independent variables) is much greater than the number of compounds (data points). In this case the classical solution does not lead to satisfactory results while the PLS regression lead to highly predictive models.[20-22]

Partial Least Squares regression is based on linear transition from a large number of original descriptors to a new variable space based on *small number of orthogonal factors (latent variables)*. Those factors are mutually independent (orthogonal) linear combinations of original descriptors. Unlike some similar approaches (e.g. principal component regression PCR), latent variables are chosen in such a way as to provide maximum correlation with dependent variable; thus, PLS model contains the smallest necessary number of factors.[21, 22] With increasing number of factors, PLS model converges to ordinary multiple linear regression model (if one exists). In addition, PLS approach allows one to detect relationship between activity and descriptors even if key descriptors have little contribution to the first few principal components. Basic algorithm of PLS method[20, 22] for the step of building k-th factor can be represented as follows:

$$w = X'y\frac{1}{\sqrt{norm(X'y)}};\, t = Xw;\, p = X't\frac{1}{norm(t)};\, q = y't\frac{1}{norm(t)}$$
$$X^{(k+1)} = X - tp';\, y^{(k+1)} = y - qt \quad (6)$$

where N – number of compounds, M – number of descriptors, $X[N,M]$ – descriptor matrix, $y[N]$ – activity vector, $w[M]$ – auxiliary weight vector, $t[N]$ – factor coefficient vector, $p[M]$ – loading vector, q – scalar coefficient of relationship between factor and activity; all vectors are columns, entities without index "($k+1$)" are for the current (k-th) factor.

In cases, where initial descriptor scales have missing values, the nonlinear partial least-squares (NIPALS) algorithm[23] for finding the principal components is preferred.

In this study, the dependent variables are the HLB values of 14 commercial nonionic emulsifiers[24] and independent variables (molecular descriptors) and orthogonal factors were calculated using the CODESSA program.[25] For this study were calculated an optimum number of 11 factors.

The optimal PLC regression with 4 orthogonal factors is:

$$\mathrm{HLB} = 0.114(\pm0.013)\boldsymbol{factor1} - 0.162(\pm0.054)\boldsymbol{factor4} + 0.223(\pm0.097)\boldsymbol{factor7} + 0.278(\pm0.118)\boldsymbol{factor8} + 10.80 \quad (7)$$

N = 14 $R^2 = 0.908$ SD = 1.603 F = 22.25 $R^2(\mathrm{CV}) = 0.771$

Table 1

Values in literature (lit.) for HLB [24], those calculated with eq. 7
(involving values of factors 1, 4, 7 and 8) in case of 14 commercial nonionic surfactant studies

Name[a]	Chemical Designation[b]	***Factor 1***	***Factor 4***	***Factor 7***	***Factor 8***	HLB (lit.)	HLB (eq. 7)	Resid.
Atlas G-3806	Polyoxyethylene cetyl ether	26.56	-4.991	-9.353	-0.070	10.3	12.54	-2.240
Brij 30	Polyoxyethylene lauryl ether	-23.94	6.272	1.539	4.722	9.5	8.702	0.798
Brij 72	Polyoxyethylene stearyl ether	-23.95	14.39	3.361	-1.963	4.9	5.933	-1.033
Brij 97	Polyoxyethylene oleyl ether	-0.764	-7.107	2.435	0.540	12.4	12.55	-0.150
PEG 400 ML	Polyoxyethylene glycol 400 monolaurate	-7.805	-12.54	1.667	-0.795	13.0	12.08	0.920
PEG 600 ML	Polyoxyethylene glycol 600 monolaurate	6.274	-15.71	0.213	-1.750	14.6	13.62	0.980
Span 20	Sorbitan monolaurate	-40.53	-0.334	-3.778	8.005	8.6	7.605	0.995
Span 40	Sorbitan monopalmitate	-32.87	2.303	-0.850	0.141	6.7	6.519	0.181
Span 60	Sorbitan monostearate	-33.12	4.255	-2.990	-8.836	4.7	3.201	1.499
Span 80	Sorbitan monooleate	-32.73	-2.522	1.192	-1.578	4.3	7.293	-2.993
Tween 20	Polyoxyethylene sorbitan monolaurate	31.43	-1.365	11.17	0.851	16.7	17.33	-0.630
Tween 40	Polyoxyethylene sorbitan monopalmitate	38.69	3.450	0.066	-0.807	15.6	14.45	1.150
Tween 60	Polyoxyethylene sorbitan monostearate	43.89	5.309	-2.585	2.645	14.9	15.11	-0.210
Tween 80	Polyoxyethylene sorbitan monooleate	48.84	8.585	-2.092	-1.103	15.0	14.22	0.780

[a] Commercial name; [b] Structural design

The HLB values in literature,[24] those calculated using eq. 7 and the values of **factors 1, 4, 7** and **8** in the situation of 14 commercial nonionic emulsifiers[24] are presented in Table 1.

Satisfactory results through (resid.) differences presented in Table 1 validate the type of equation 7 which helps to appreciate the HLB value considering the surfactant structure. Future studies will seek to expand such a strategy in the situation of other surfactant-type compounds.

CONCLUSIONS

The correlation of HLB values in the case of 14 commercial nonionic emulsifiers with values of molecular descriptors calculated using CODESSA[25] program led to a QSPR model with satisfactory statistical parameters appealing to the PLC regression. Such a strategy is that it may be involved in the future to predict the HLB values of some surfactants which present interest.

Acknoledgements: This article was done within the research programme "Coordinative and Supramolecular Chemistry" of the "Ilie Murgulescu" Institute of Physical Chemistry, finaced by the Roumanian Academy.

REFERENCES

1. M. J. Rosen, "Surfactants and interfacial phenomena", John Wiley & Sons, Hoboken, New Jersey, 3rd edition, 2004, p. 1.
2. A. M. Badawi, A. S. Makawi Mohamed, M. Z. Mohamed and M. M. Khowdairy, *J. Cancer Res. Therapeutics*, **2007**, *3*, 198-206.
3. M. J. Rosen and M. Dahanayake, "Industrial Utilization of Surfactants", AOCS, 2000.
4. D. T. Wasan, M. E. Ginn and D. O. Shah, editors, "Surfactants in Chemical/Process Engineering", New York: Marcel Dekker, 1988.
5. C. B. F. Young and K. N. Coons, "Surface Active Agents. Theoretical Aspects and Applications", Brooklyn, New York, Chemical Publishing Co., 1945.
6. D. Myers, "Surfactant Science and Technology", John Wiley & Sons, Hoboken, New Jersey, 3rd edition, 2006, p. 48, 66, 69, 74.
7. H. Stache, "Anionic surfactants: organic chemistry", New York: Marcel Dekker, 1996; J. Cross, "Anionic surfactants: analytical chemistry", New York: Marcel Dekker, 1998.
8. H. Waldhoff and R. Spilker, editors, " Handbook of Detergents, Part C: Analysis (Surfactant Science)", CRC Press, 2004.
9. E. Jungermann, "Cationic surfactants", New York: Marcel Dekker, 1990.
10. J. M. Richmond, "Cationic surfactants: organic chemistry", New York: Marcel Dekker, 1988; D. N. Rubingh and P. M. Holland, "Cationic surfactants: physical chemistry", New York: Marcel Dekker, 1991.
11. M. J. Schick, "Nonionic surfactants: physical chemistry", New York: Marcel Dekker, 1987; N. M. van Os, "Nonionic surfactants: organic chemistry", New York: Marcel Dekker, 1998.
12. E. G. Lomax, "Amphoteric surfactants", New York: Marcel Dekker, 1996.
13. W. C. Griffin, *J. Soc. Cosmetic Chemist*s, **1949**, *1*, 311-326.
14. W. C. Griffin, *J. Soc. Cosmetic Chemists*, **1954**, *5*, 249-256.
15. I. J. Lin and L. Marszall, *Progress in Colloid and Polymer Science*, **1978**, *63*, 99-104.
16. R. C. Little, *J. Colloid Interface Sci*., **1978**, *65*, 587-588.
17. H. Wold, "Estimation of principal components and related models by iterative least squares", in "Multivariate Analysis", P. R. Krishnaiah, editor, New York: Academic Press, 1966, pp. 391-420.
18. H. Wold, "Nonlinear Iterative Partial Least Squares (NIPALS) modeling: some current developments", in "Multivariate Analysis", P. R. Krishnaiah, editor, New York: Academic Press, 1973, pp. 383-407.
19. S. de Jong, *Chemometrics Intell. Lab. Syst.*, **1993**, *18*, 251-263.
20. H. Martens and T. Naes, "Multivariate Calibration", John Wiley & Sons, Chichester, 1989.
21. A. Höskuldsson, *J. Chemometrics.,* **1988,** *2*, 211-228.
22. I. V. Tetko, J. Gasteiger, R. Todeschini, A. Mauri, D. Livingstone, P. Ertl, V. A. Palyulin, E. V. Radchenko, N. S. Zefirov, A. S. Makarenko, V. Y. Tanchuk and V. V. Prokopenko, *J. Comput. Aid. Mol. Des.*, **2005**, *19*, 453-463.
23. P. Geladi and B. Kowalski, *Anal. Chim. Acta* **1986**, *185*, 1-17.
24. P. Becher, "Emulsions: theory and practice", Monography series no 135, Reinhold, New Jersey, USA, 1957.
25. A. R. Katritzky, S. Perumal, R. Petrukhin, and E. Kleinpeter, *J. Chem. Inf. Comp. Sci*. **2001**, *41*, 569-574; A. R. Katritrzky, V. S. Lobanov, and M. Karelson, CODESSA: A Reference Manual (Version 2.0), Gainesville, Florida, 1994.