African Journal of Biotechnology Vol. 10(81), pp. 18863-18868, 16 December, 2011
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB11.2766
ISSN 1684–5315 ©2011 Academic Journals

Full Length Research Paper

Prediction of anticancer activity of aliphatic nitrosoureas using quantum chemical quantitative structure activity relation (QSAR) methods

Noorbatcha, I. A.[1]*, Hamzah, F.[1], Salleh, H. M.[1] and Idid, S. Z.[2]

[1]BioProcess and Molecular Engineering Research Unit (BPMERU), Department of Biotechnology Engineering, Faculty of Engineering, International Islamic University Malaysia (IIUM), P.O. Box 10, 50728 Kuala Lumpur, Malaysia.
[2]Department of Biomedical Science, Faculty of Science, International Islamic University Malaysia, Bandar Indera Mahota, 25200 Kuantan, Pahang, Malaysia.

Accepted 7 November, 2011

Design and development of new anticancer drugs with low toxicity is a very challenging task and computer aided methods are being increasingly used to solve this problem. In this study, we investigated the anticancer activity of aliphatic nitrosoureas using quantum chemical quantitative structure activity relation (QSAR) approach. In this method, the physic-chemical properties, known as descriptors, necessary for predicting quantitative structure activity relations was obtained from semi empirical quantum chemical methods. We used Recife Model 1 to optimize the structure of the molecules and to calculate the quantum chemical descriptors, while heuristic and best multilinear regression methods were applied to obtain the best correlation. Two data sets containing aliphatic nitrosoureas and chloroethyl substituted nitrosoureas were used in the present calculations. The QSAR equations obtained here can be used to design new anticancer drugs prior to resorting to experimental activity studies.

Key words: Quantitative structure activity relationship (QSAR), best multi linear regression (BMLR), quantum chemical method, Recife Model 1 (RM1).

INTRODUCTION

Nitrosoureas play an important role in the treatment of cancer (Gnewuch and Sosnovsky, 1997). Even though some nitrosoureas have been used in chemotherapy for treating leukemia, most of them are still not effective against solid tumor (Hansch, 1979). Hence a systematic study of the correlation between biological activity and the structure of nitrosoureas will lead to the design and synthesis of more bioactive nitrosoureas. In the conventional quantitative structure relation (QSAR) analysis, as initially proposed by Hansch and Fujita (Hanch et al., 1980), the correlation between the bioactivity and some physicochemical property of the chemical compound is represented by a mathematical equation. This QSAR allows the scientist to target new compounds with improved activity and the number of novel compounds that have to be synthesized. Thus, numbers of experimental tests needed for testing new compounds are reduced drastically. Additionally, if a newly discovered analogue does not fit the proposed QSAR equation, it explains that some other feature is important and provides a hint for further development (Patrick, 2005). In this work, we applied quantum chemical methods to calculate the various electronic and structural parameters and related these parameters to the bioactivity of the nitrosoureas.

MATERIALS AND METHODS

A series of nitrosoureas with known activity against L1210 leukemia cell lines were selected from the literature (Yaujihara et al., 1981; Johnston et al., 1963; Schmid et al., 1986; Schabel et al., 1963). The general structure of nitrosoureas is shown in Figure 1.

The biological activities of these compounds have been express-

*Corresponding author. E-mail: ibrahiman@iium.edu.my.

ed as percentage of increase in life span (%ILS). On the basis of the structural similarity, the selected nitorosureas are grouped into two sets, as given in Tables 1 and 2. The three dimensional (3D) structures of these compounds were drawn using ACD/ ChemSketch software and their initial geometry is optimized using molecular mechanics methods. The optimized 3D structures are then used for self-consistent field (SCF) calculation using MOPAC 6.0 (Stewart and Frank, 1990) using RM1 semi empirical quantum chemical method (Rocha et al., 2006). The optimized geometries are checked with the known structural information available. The electronic structural data produced by this quantum chemical approach is used by CODESSA program (Katrizky et al., 1995) to calculate conventional, topological, geometrical, electrostatic, quantum-chemical and thermodynamic descriptors. These descriptors are further analyzed for linear dependence.

Statistical methods were applied to develop the best QSAR model and to select or reduce the number of descriptors. The Heuristic method and the best multi linear regression (BMLR) method (Katrizky et al., 1995) are used to obtain the QSAR equation.

Heuristic method is used to identify descriptors with bad or missing values, insignificant and highly inter correlated descriptors. All descriptors are checked to ensure that values of each descriptor are available for each structure and that there is a variation in these values. Descriptors for which values are not available for every structure in the data in question are discarded. Descriptors having a constant value for all structures in the data set are also discarded. This information will reduce the number of descriptors involved to develop the best QSAR model. The statistical parameters such as regression or correlation coefficient (*R*), standard deviation (*s*) and *F* values are calculated for the QSARs obtained.

After developing the QSAR model or the correlation equation, it is necessary to analyze the model. There are several methods that can be used to analyze and improve the model. If there is a big number of descriptor, may be one or two of them are inter-correlated. The correlations were improved and new descriptors were constructed if the QSAR models are not fulfilling the satisfaction. The best QSAR model is the correlation model with the optimum values of statistical criteria which are highest values of regression correlation coefficient R^2, the cross validated R_{cv}^2 and the *F* value. Several criteria can be used to choose the best QSAR model. When adding additional descriptors to improve the quality of the QSAR, the model corresponding to the break point is considered to be the optimum model. The break point is the point where the statistical improvement in the R^2 value after the addition of new descriptor to the QSAR equation, drops below 0.02 after this point (Katrizky et al., 2004).

RESULTS AND DISCUSSION

QSAR-data set 1

The best nonlinear four descriptor for aliphatic nitrosoureas set was selected, and the fitted parameters with their statistical errors are given in Table 3. The QSAR with these parameters can be written as:

$$\log\left(\frac{1}{\%ILS}\right) = 63.5(ft) - 1.21(fi) - 0.575(^2X) + 0.214(^1\kappa) - 0.864 \quad (1)$$

It can be seen that Equation (1) describes the role of molecular topology and electrostatic interactions on the anti-cancer activity of nitorosoureas. In this equation, fractional negatively charged surface area of type 3

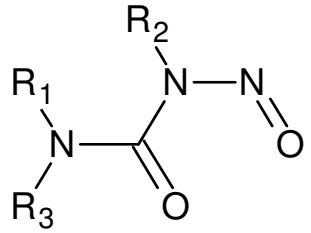

Figure 1. Molecular structure of nitrosoureas, where R_1, R_2 and R_3 are different substituent groups including hydrogen atoms.

(FNSA3) is calculated as the ratio of PNSA3 and TMSA, where PNSA3 refers to the partial negative surface area weighted by atomic charge and TMSA refer to the total molecular surface area (Stanton, 1990; Stanton et al., 1992). Topographic electronic index for all bonded and non-bonded pairs of atoms T_1^E (Osmialowski et al., 1986) is calculated as:

$$T_1^E = \sum_{(i<j)}^{N} \frac{|q_i - q_j|}{r_{ij}^2}$$

Where, q_i is a partial charge on the *i*-th atom and r_{ij} is a distance between *i*-th and *j*-th atoms.
The average information content of order 2, $^K IC$, (Bonchev, 1983) is defined as:

$$^K IC = \sum_i \frac{n_i}{n} \log_2 \frac{n_i}{n}$$

Where, n_i is the number of atoms in the i-th class and *n* is the total number of atoms in the molecule. Randic index and Kier shape index were calculated using the formula given by Bingham et al. (1975) and Kier (1990).

It is observed that Kier shape index and the cross product of FNSA-3 Fractional PNSA (PNSA-3/TMSA) and topological electronic index contribute positively to the obtained model while, the rest contribute negatively to the QSAR model. It shows that increasing the information content and Randic index, as well as decreasing the FNSA-3 topological electronic index and Kier shape index will increase the value of %ILS. A comparison of calculated and predicted values of %ILS using Equation 1 is shown in Figure 2.

QSAR-data set 2

The parameters for the best regression obtained for chloroethyl substituted nitrosoureas set are shown in Table 4. The QSAR for the chloroethyl substituted nitrosoureas set is given as follows:

$$\log\left(\frac{1}{\%ILS}\right) = 3.59(S_k^r) - 7.77(E_{tot}) + 4.37(E_R) + 45.2 \quad (2)$$

Table 1. %ILS values for the selected aliphatic nitrosoureas of the type R^1-NH-CO-N(NO)-R^2.

S/N	R^1	R^2	%ILS[#]
1	H	CH_3	109
2	H	$ClCH_2CH_2$	131
3	H	CH_3CH_2	34
4	H	$CH_3CH_2\,CH_2$	30
5	H	CH_2=CH CH_2	32
6	H	$CH_3(CH_2)_3$	29
7	CH_3	CH_3	61
8	$ClCH_2CH_2$	CH_3	46
9	$ClCH_2CH_2$	$ClCH_2CH_2$	184
10	FCH_2CH_2	$ClCH_2CH_2$	500
11	$(CH_3)_3\,CCH_2$	$ClCH_2CH_2$	275
12	$CH_3CH_2OCOCH_2$	$ClCH_2CH_2$	26
13	$NCCH_2CH_2$	$NCCH_2CH_2$	53
14	$NCCH_2CH_2$	$ClCH_2CH_2$	40
15	CH_3COCH_2	$ClCH_2CH_2$	186
16	$(CH_3)_2N$	$ClCH_2CH_2$	129
17	CH_3NCHO	$ClCH_2CH_2$	226
18	CH_3NCOCH_3	$ClCH_2CH_2$	229

[#]ILS values taken from Johnston et al., 1963, Schmid et al., 1986 and Schabel et al., 1963.

Table 2. %ILS values for the selected chloroethyl substituted nitrosoureas of the type R^1-R^3N-CO-N(NO)-CH_2CH_2CL.

S/N	R^1	R^3	%ILS[#]
1	CH_3	CH_3	249
2	CH_3	$CH_3(CH_2)_3$	421
3	CH_2=$CHCH_2$	CH_2=CH CH_2	103
4	H	$CH_2CH_2COCH_2CH_2$	623
5	CH_3	$HOCH_2CH_2$	140
6	CH_3	$HOCH_2(CHOH)CH_2$	433
7	CH_3	$HOCH_2(CHOH)_4CH_2$	84
8	$CH_3CH_2C\,H_2$	$HOCH_2(CHOH)CH_2$	231
9	$CH_3(CH_2)_3$	$HOCH_2(CHOH)CH_2$	278
10	$CH_3(CH_2)_3$	$HOCH_2(CHOH)_4CH_2$	143
11	$HOCH_2CH_2$	$HOCH_2CH_2$	688

[#]ILS values taken from Taujihara et al., 1981.

Table 3. The parameters for best nonlinear four descriptor QSAR for aliphatic nitrosoureas set (R^2 = 0.9642, F = 87.59, S^2 = 0.0077, R^2_{CV} = 0.9178).

S/N	X	±ΔX	*t*-test	Descriptor
0	-8.64e-01	1.37e-01	-6.3249	Intercept
1	6.35e+01	4.63e+00	13.7136	FNSA3 Fractional PNSA (PNSA3/TMSA) *x* Topological electronic index (*ft*)
2	-1.21e+00	1.23e-01	-9.8519	FNSA3 Fractional PNSA (PNSA3/TMSA) x Information content (*fi*)

Table 3. Contd.

3	-5.75e-01	8.05e-02	-7.1429	Randic index (order 2) (^2X)
4	2.14e-01	4.38e-02	4.8879	Kier shape index order 1)($^1\kappa$)

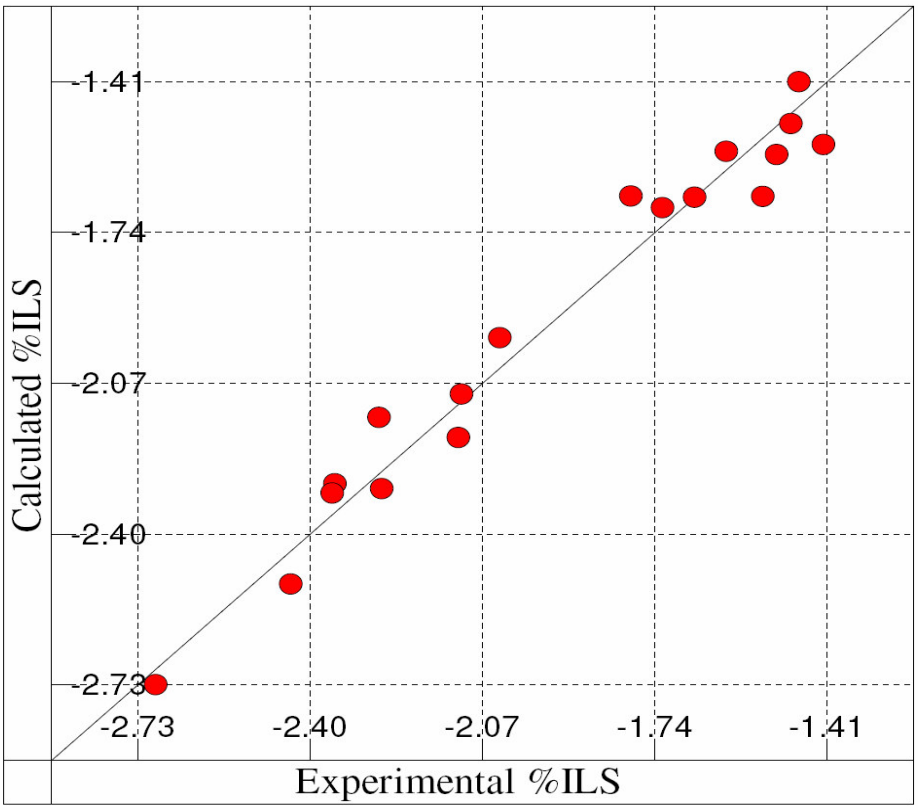

Figure 2. Calculated versus experimental activity according to the best nonlinear four parameters correlation Equation (1) for aliphatic nitrosoureas set.

Table 4. The best three descriptor QSAR for chloroethyl substituted nitrosoureas set-2 (R^2 = 0.9637, F = 61.93, S^2 = 0.0050, R^2_{CV} = 0.9178).

S/N	X	±ΔX	*t*-test	Descriptor
0	4.52e+01	6.18e+00	7.3174	Intercept
1	3.59e+00	8.47e-01	4.2395	ZX Shadow/ZX Rectangle (S_k^r)
2	-7.77e+00	6.97e-01	-11.1474	Min total interaction for a C-H bond (E_{tot})
3	4.37e+00	9.59e-01	4.5606	Min resonance energy for a C-H bond (E_R)

Figure 3 shows the correlation between calculated activity and experimental activity according to the best nonlinear four parameters correlation in Equation (2) for chloroethyl substituted nitrosoureas set.

The descriptors involved in QSAR for chloroethyl substituted nitrosoureas in Equation (2), are ZX shadow/ZX rectangle (S_k^r), minimum total interaction for a C-H bond (E_{tot}) and minimum resonance energy for a C-H bond

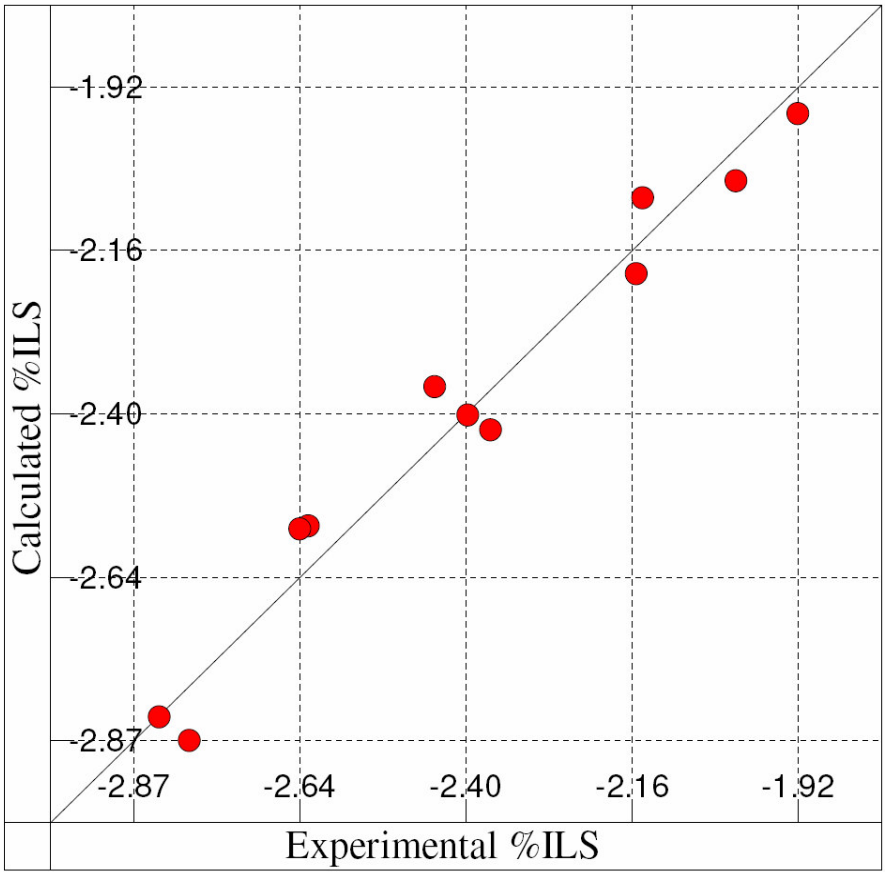

Figure 3. Calculated versus experimental activity according to the best three parameters correlation for chloroethyl substituted nitrosoureas.

(E_R). The descriptor S_k^r, is the area of the shadow of the molecule on the XZ plane divided by the area of the maximum rectangle covered by this area (Rohrbaugh and Jurs, 1987). Hence, this descriptor explains the influence of the size and shape of the molecule on the anti-cancer activity of these nitrosoureas. The minimum total interaction energy for a C-H bond (E_{tot}) and minimum resonance energy for a C-H bond (E_R) have related conformation changes in the molecules (Strouf, 1986). Hence, the QSAR Equation (2) provides a measure of the influence of the changes in the size, shape and conformation on the anticancer activity of the nitrosoureas investigated here.

Conclusion

The quantum chemical approach together with statistical analysis was used to establish quantitative relationship between structure and bioactivity. The topological, shape, size and conformation dependent descriptors are found to play a significant role in influencing the anti-cancer activities of the nitrosoureas. These QSARs can be used to predict the %ILS activity of new compounds. The implementation of quantum chemical method in QSAR approaches in this project circumvents experimental measurement of physicochemical properties required in traditional QSAR analysis, thereby reducing the efforts necessary to develop new bioactive nitrosoureas.

ACKNOWLEDEGMENT

This research work was funded by Fundamental Research Grant Scheme (FRGS) from Ministry of Higher Education, Malaysia.

REFERENCES

Bingham RC, Dewar MJS, Lo DH (1975). Ground states of molecules. XXV. MINDO/3. Improved version of the MINDO semiempirical SCF-MO method. J. Am. Chem. Soc. 97: 1285-1293

Bonchev D (1983) Information Theoretic Indices for Characterization of Chemical Structure, Wiley-Interscience, New York.

Gnewuch CT, Sosnovsky G (1997). A critical appraisal of the evolution of N-Nitrosoureas as anticancer drugs. Chem. Rev. 97: 829-1013.

Hansch C (1979). QSAR in cancer chemotherapy. Farmaco Sci. 34(1): 89-104.

Hansch C, Leo A, Schmidt C, Jow PY, Montgomery JY (1980). Antitumor structure activity relations. Nitrosoureas vs. L1210 leukemia. Med. Chem. 23(10): 1095-101.

Johnston TP, Mccaleb GS, Montgomery JA (1963). The Synthesis of Antineoplastic Agents. XXXII. N-Nitrosoureas. Int. J. Med. Chem. 6: 669-681.
Katritzky AR, Karelson M, Lobanov VS, Dennington R, Keith, T (1995). CODESSA v2.60, Semichem: 7204 Summit, Shawnee, Kansas, USA.
Katritzky AR, Fara DC, Yang H,. Karelson M, Suzuki T, Solovev VP, Varnek A (2004). Quantitative structure-property relationship modeling of β-Cyclodextrin complexation free energies. J. Chem. Int. Comput. Sci. 44(2): 529-541.
Kier LB (1990) In: Computational Chemical Graph Theory Rouvray, D.H. (editor), Nova Science Publishers, New York, pp. 151-174.
Osmialowski K, Halkiewicz J, Kaliszan R (1986). Quantum chemical parameters in correlation analysis of gas liquid chromatographic retention indices of amines. II. Topological electronic index. J. Chromatogr. 361: 63.
Patrick GL (2005). An introduction to medicinal chemistry, Oxford University Press, USA.
Rocha GB, Freire RO, Simas AM, Stewart JJP (2006). RM1: A reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I. J. Comput. Chem. 27: 1101-1111.
Rohrbaugh RH, Jurs PC (1987). Molecular shape and the prediction of high-performance liquid chromatographic retention indexes of polycyclic aromatic hydrocarbons. Anal. Chim. Acta. 199: 99.
Schabel FM Jr, Johnston TP, Mccaleb GS, Montgomery JA, Laster WR, Skipper HE (1963). Experimental evaluation of potential anticancer agents VIII. Effects of certain nitrosoureas on intracerebral L1210 leukemia. Cancer Res. 23: 725-33.
Schmid JR, Fiebig HH, Eisenbrand G, Löhr GW (1986). L1210 leukemia i. v. implanted as a model for testing short-chain nitrosourea analogs, J. Cancer Res. Clin. Oncol. 111: 31-34.
Stanton DT, Jurs PC (1990). Development and Use of Charged Partial Surface Area Structural Descriptors in Computer Assissted Quantitative Structure Property Relationship Studies. Anal. Chem. 62: 2323-2329.
Stanton DT, Egolf LM, Jurs PC, Hicks MG (1992). Computer-Assisted Prediction of Normal Boiling Points of Pyrans and Pyrroles. J. Chem. Int. Comput. Sci. 32: 306-316.
Stewart, JJP, Frank J (1990). MOPAC 6.0, Seiler Research Laboratory, United States Air Force Academy, Colorado, USA.
Strouf O (1986). Chemical Pattern Recognition, Wiley, New York
Taujihara K, Ozski M, Morikawa T, Arai Y (1981). A New Class of Nitrosoureas. I. Synthesis and Antitumor Activity of 1-(2-Chloroethyl)-3, 3-disubstituted-1-nitrosoureas having a Hydroxyl Group at the β-Position of the Substituents, Chem. Pharm. Bull. 29: 2509-15