Topological modeling of lipophilicity of some crown compounds

1SALMA KHAN, 1S. M.ALI and 2VINAY KUMAR DUBEY

1Research Scholar, Islamia Karimia College, Indore M.P. India
1Ex. Head, Department of Chemistry I. K. College Indore M.P. India
2Department of Chemistry, APS University, Rewa M.P. India

Email of Corresponding Author: salmazoyakhan786@gmail.com
http://dx.doi.org/10.22147/juc/130101

Acceptance Date 7th Nov., 2016, Online Publication Date 2nd January, 2017

Abstract

In the present paper a study on Benzo and di Benzo crown ethers has been done for the estimation of lipophilicity of these compounds. For this few topological parameters viz. Wiener and Balaban index along with MW gives best result. Proposed three-parametric model is capable of explaining the 97% of the variance.

Key words: Lipophilicity, log P, Topological modeling, Balaban index, Molecular Weight. QSAR

Introduction

Crown ethers are macro cyclic polyether that are characterized by having a different number of ethylene oxide unit, joined covalently in a macr cyclic ring so that in their simplest form they are cyclic oligomers of dioxane.1

It is well known that macro cyclic compounds have ability to surround or enclose many different cations.2 Consequently, interest in the complexes of macro cyclic compounds has increased in recent years. Complexes of these crown-ligands with alkali and alkaline earth cations have been reported extensively, however, their complexes with transition and rare-earth cations are very scare. Complexes of multidentate macro cyclic ligands, viz., crown ethers and cryptand e with lanthanide cations have, therefore, attracted considerable interest.2-3

Crown ether and related compounds become a central part of development in supra molecule (host/Guest) chemistry4-6 many different modification of crown ethers, such as changing ring size, the kinds of substituent’s and the types donor atoms, have been made to enhance their complexation properties.7

The partitioning of organic compounds acting as drugs between aqueous and lipophilic phase is of utmost importance for drug potency. Lipophilicity is an important endpoint used extensively in medicinal chemistry, drug design, pharmacy and environmental toxicity in predicting biological and hazardous effects of chemicals.8 Consequently, the development of new methods for the quantitative estimation of the lipophilicity of organic molecules is extremely important. It is well known and well established that the useful measure of lipophilicity/hydrophobicity of organic compounds acting as drugs is then passion coefficient between n-octanol and water9-10. No other physicochemical property has attracted as much interest in quantitative structure-activity relation (QSAR) studies as lipophilicity
Salma Khan, et. al.

(synonymously called hydrophobicity, any difference between both these terms is only a Semitic nemicity; the opposite of lipophilicity is hydrophobicity) generally expressed as log P.

Lipophilicity (Log P) is a physio-chemical property of principle importance in drug discovery and development.\(^1\) Log P has been modeled by Khadikar and Agrawal\(^12^-^14\) for many compounds using topological indices such as Randic connectivity, Kier & Hall valence connectivity, Wiener index, Balaban Index, Padmaker Ivan index\(^15^-^18\) and many others.

They have successfully demonstrated that using topological indices one can develop excellent models which can predict log P value of different compounds. They developed the models and tested them also.

**Methodology:**

In Quantitative structure activity relationship (QSAR) studies the lipophilicity expressed as log P has a major role. C. Hansch\(^19^-^20\) and coworkers have widely used lipophilicity (log P) in the field of chemical–biological interaction which can be modeled by using few independent parameters. We develop a mathematical model using multiple regression analysis\(^21\).

\[ Y = m_1 x_1 + m_2 x_2 + \ldots + m_i x_i + c \quad (1) \]

Where

\( Y = \) biological activity and \( x_1, x_2, \ldots, x_i = \) independent parameters, \( C = \) intercept.

It has been observed that since, the discovery of crown compound by Pederson they are widely used as membranes in medicinal chemistry. The membrane power of crown compounds is decided on the basis of \( \text{Log P} \). Such Log P value for the crowns were quantitative calculated by using lipophilicity increments \( \pi \) by Hansch. However, not much work in QSAR has been done in modeling lipophilicity (Log P) of crown compounds.

The crown compounds used in the present study are reported in Table 1. The log P values for these compounds have been taken from literature. These values are also reported in this table. Topological parameters for these compounds have been calculated by the Dragon software\(^22\) before that we used Chem. Sketch software\(^23\) to draw the molecular structures and then mole file were used for calculation of Topological indices depleting all the C-H bonds.

**Parameters Used:**

We have used following parameters for modeling log P of present set of crown ethers.

1. Wiener Index (W):  
2. Balaban Index (J)
3. Zero order connectivity index \( (\chi^0) \)
4. Molecular Weight (MW)

The calculation of these indices have been given in the literature\(^15\).

**Result and Discussion**

These compounds can be arranged in increasing order of lipophilicity (log P):

\[ 8 > 9 > 11 > 13 > 10 > 12 > 1 > 2 > 3 > 4 > 5 > 7 > 6 \]

This means that di Benzo,18-crown-6 has highest lipophilicity (log P) while Benzo 30-crown-10 has lowest value of lipophilicity (log P), this sequence do not give any structure-activity relationship. No one correlation is seen in activity and structures of the compounds. The structures of the compounds were drawn from ACD labs chem. Sketch soft ware.\(^23\) the calculated topological Parameters are reported in Table-2.

**Table-1 Compounds used in the present study and lipophilicity in terms of Log P**

| Comp.No. | Name of Comp.       | Logp   |
|----------|---------------------|--------|
| 1        | Bz,15-Crown-5       | 0.91   |
| 2        | Bz, 18-crown-6      | 0.58   |
| 3        | Bz,21-crown-7       | 0.57   |
| 4        | Bz,24-crown-8       | 0.45   |
| 5        | Bz,27-crown-9       | 0.23   |
| 6        | Bz,30-crown-30      | 0.03   |
| 7        | Bz,33-crown-11      | 0.09   |
| 8        | Dbz,18-crown-6      | 2.20   |
| 9        | Dbz,24-crown-8      | 2.11   |
| 10       | Dbz,27-crown-9      | 1.63   |
| 11       | Dbz,30-crown-10     | 1.80   |
| 12       | Dbz33-crown-11      | 1.45   |
| 13       | Dbz,21-crown-7      | 1.74   |

Bz-benzo; Dbz-di Benzo; Log P-lipophilicity;
Topological modeling of lipophilicity of some crown compounds.

Table 2: Lipophilicity (Log P) and calculated values of descriptors.

| Compound No. | Log P   | W    | J   | $\chi^0$ | MW   |
|--------------|---------|------|-----|---------|------|
| 1            | 0.91    | 744  | 1.74| 13.176  | 268.34|
| 2            | 0.58    | 1171 | 1.65| 15.297  | 312.4 |
| 3            | 0.57    | 1731 | 1.65| 17.418  | 400.52|
| 4            | 0.45    | 2458 | 1.615| 19.539 | 444.58|
| 5            | 0.23    | 3357 | 1.591| 21.661 | 488.64|
| 6            | 0.03    | 4465 | 1.567| 23.782 | 532.7 |
| 7            | 0.09    | 5784 | 1.55 | 25.903 | 580.74|
| 8            | 2.20    | 1797 | 1.438| 17.866 | 360.44|
| 9            | 2.11    | 3420 | 1.363| 22.108 | 429.62|
| 10           | 1.63    | 4947 | 1.341| 24.23 | 536.68|
| 11           | 1.80    | 5835 | 1.31 | 26.351 | 580.74|
| 12           | 1.45    | 7368 | 1.292| 28.472 | 634.68|
| 13           | 1.74    | 2499 | 1.41 | 19.987 | 404.5 |

We have tried variable selection method to get multiparametric correlations and the results are reported in Table 3.

The Balaban index (J) plays a dominating role in modeling of Log P of present set of compounds. Similarly MW also plays a dominating role in trimetric model using maximum R² method.

Table 3. Result of variable selection.

| Model Size | R-Squared | R-Squared Change | Variable Names |
|------------|-----------|------------------|----------------|
| 1          | 0.532353  | 0.532353         | J              |
| 2          | 0.970016  | 0.437663         | $\chi^0$, MW   |
| 3          | 0.974543  | 0.004527         | W, J, MW       |
| 4          | 0.974590  | 0.000047         | W, J, $\chi^0$, MW |

The above table shows that J, $\chi^0$, and MW are excellent parameters for modeling Log P of present set of compounds. However, the change in R² value is drastic from mono to bi-parametric correlation. We have also obtained correlation matrix which is reported in Table 4.

Table 4. The correlation matrix.

|       | J       | MW     | log P   |
|-------|---------|--------|---------|
| J     | 1.0000  | -0.7184| -0.7296 |
| MW    | -0.7184| 1.0000 | 0.0668  |
| log P | -0.7296| 0.0668 | 1.0000  |

Correlation matrix shows that J and MW are highly correlated. Similarly $\chi^0$ is also moderately correlated with log P, J and MW. log P shows strong correlation with MW.

The data was subjected to regression analysis and gave few statistically significant models which are reported in Table 5.

Mono-parametric Model:

Out of three mono-parametric models, one with J (Balaban index) gives maximum R² value which is 0.5324. The value of adjusted R² comes out to be 0.4898. The model is as under:

\[
\text{log } P = -3.7625 (\pm 1.0633)J + 6.7164 \quad (2)
\]

Here and hereafter, n is total no of compounds, R² is variance (square of correlation constant), R²_adj is adjusted R², Se is standard error of estimation, F is Fischer’s ratio and ‘Q’ is Pogliani’s quality factor which is the ratio of R/Se.

Bi-parametric Model:
When \( MW \) is added to mono-parametric model discussed above, a bi-parametric model is obtained. A drastic improvement in \( R^2 \) is observed. The \( R^2 \) value changes from 0.5323 to 0.9645. The adjusted \( R^2 (R^2_{Adj}) \) also changes from 0.4898 to 0.9573 which clearly indicates that the addition of \( MW \) is significant and it has its fair share in the model. The model is as given below:

\[
\log P = -7.2886(\pm 0.4435)J -0.0080(\pm 0.0007)MW +15.4801 (3)
\]

\[ n=13, Se=0.09197, R^2=0.9645, R^2_{Adj}=0.9573, F=135.654, Q=1.06783 \]

**Tri-parametric Model:**

Finally a tri-parametric model is obtained when Wiener index \( W \) is added to the above model. For this model a slight improvement in \( R^2 \) value is observed. The value changes from 0.9645 to 0.9745 and \( R^2_{Adj} \) also increases from 0.9573 to 0.9661. Hence, Addition of \( MW \) is significant and has fair share in the model. The model is discussed below:

\[
\log P = 0.0002(\pm 0.0001)W -7.4068(\pm 0.4006)J -0.0123(\pm 0.0024) MW+16.8537
\]

\[ n=13, Se=1.0963, R^2=0.9745, R^2_{Adj}=0.9661, F=114.846, Q=0.90045 \]

A perusal of equation 4 i.e. the model 6 and its comparison with equation 2 clearly shows that the \( Q \) value in the latter has decreased. Therefore, though the \( R^2 \) value of tri-parametric model is more than what we have in bi-parametric model, it cannot be accepted as best model. The cross validated \( R^2 \) value also confirms this finding, which has been discussed below.

Therefore the tri-parametric model is the best among all the discussed models for estimation of \( \log P \) value of present set of compounds.

To confirm above finding, \( \log P \) values have been estimated using best Tri-parametric model and calculated values are shown in Table 6.

**Table 6.** Observed and estimated \( \log P \) values of compounds using model no. 6

| Comp. no. | Observed \( \log P \) | Estimated \( \log P \) | Residual |
|-----------|----------------------|-----------------------|----------|
| 1         | 0.91                 | 0.801                 | 0.109    |
| 2         | 0.58                 | 0.749                 | -0.169   |
| 3         | 0.57                 | 0.576                 | -0.006   |
| 4         | 0.45                 | 0.436                 | 0.014    |
| 5         | 0.23                 | 0.249                 | -0.019   |
| 6         | 0.03                 | 0.103                 | -0.073   |
| 7         | 0.09                 | -0.052                | 0.142    |
| 8         | 2.2                  | 2.11                  | 0.09     |
| 9         | 2.11                 | 1.901                 | 0.209    |
| 10        | 1.63                 | 1.735                 | -0.105   |
| 11        | 1.8                  | 1.687                 | 0.113    |
| 12        | 1.45                 | 1.582                 | -0.132   |
| 13        | 1.74                 | 1.913                 | -0.173   |
These values are in good agreement with the observed activities. Such correlation is obtained by plotting a graph between observed and estimated log P which is demonstrated in Fig. 1. The predictive power of model comes out to be 0.974, which shows that model explain more than 97% variance of the data used in the present study.

The above bi-parametric model is also tested by evaluating cross validation parameters reported in Table-5. The lowest PRESS/SSY value 0.22519 for the bi-parametric model and higher $R^2_{cv}$ value 0.9334 confirms above findings. The PSE value0.1954 for this model is lowest and $S_{PRESS}$ comes out to be 0.2226, further verifies our result.

The cross validated $R^2_{cv}$ value for the threeparametric correlation is lower than the bi-parametric model hence that model cannot be accepted as best model.

![Fig.1 correlation between observed and estimate Log P using model 6.](image)

Table 7. Cross Validated parameters for the obtained models.

| Model No | Parameters Used | PRESS/SSY | $R^2_{cv}$ | $S_{PRESS}$ | PSE  |
|----------|-----------------|-----------|------------|-------------|------|
| 1.       | J               | 0.74027   | 0.3611     | 0.1811      | 0.19835 |
| 2.       | J,MW            | 0.22519   | 0.9334     | 0.2226      | 0.1954 |
| 3.       | W,J,MW          | 0.23594   | 0.9313     | 0.2383      | 0.1666 |

In order to test whether the model is free from the defect of collinearity, we have calculated VIF, $\lambda_i$, condition number ($k$), tolerance ($T$) for all independent parameters used in proposed models recorded in Table-8. It is clear bi-parametric model is free from defect of collinearity.

Table 8. Ridge regression parameters for the best models

| Model No | Parameters Used | VIF   | T     | $\lambda_i$ | k     |
|----------|-----------------|-------|-------|-------------|-------|
| 1.       | J               | 0.9901| 1.0000| 1.00000     | 1.00  |
| 2.       | J               | 2.0031| 0.4840| 1.71833     | 1.00  |
|          | MW              | 2.0031| 0.4840| 0.28166     | 6.10  |
| 3.       | W               | 16.2530| 0.0319| 2.59491     | 1.00  |
|          | J               | 2.0138| 0.4722| 0.38633     | 6.72  |
|          | MW              | 17.6608| 0.0352| 0.01875     | 135.36|

The VIF for the proposed model also rejects the tri-parametric model as some of the parameters exceed the limited values. The ridge traces are recorded in fig. 2 and fig.3.

![Fig. 2. Ridge Trace for Best bi parametric model no. 5](image)

![Fig. 3. VIF plot for bi parametric Model no. 5](image)
Conclusions

1. The result and discussion made above indicates that the balaban index (J) alone is not capable of modeling lipophilicity of the bz, dbz, crown compounds used in the present study.
2. The bi-parametric model comprising of J and MW is best for modeling the Log P value of present set of crown ethers.
3. The coefficient of J and Mw are negative suggesting that these parameters have retarding effect towards exhibition of Log P. Hence the new compounds with low value of J and MW will favor the activity Log P.

Acknowledgment

Authors are thankful to Prof. Vijay K. Agrawal, Head Department of Chemistry, APS University, Rewa for his kind help and support. The authors are also indebted to late. Professor P.V. Khadikar, a world known topologist, who inspired us for carrying out the present work when he was alive.

References

1. Gokel, G.W., Leevy, W.M. and Weber M.E., Crown Ethers: Sensors for Ions and Molecular Scaffolds for Materials and Biological Models, Chem. Rev., 104, 2723-2750 (2004).
2. C.J. Pedersen, J. Ame. Chem. Soc., 89, 7017 (1967).
3. S. Gurrieri, A. Seminara, G. Siracusa, A. Cassol; Thermochim. Acta. 11, 433 (1975).
4. Lehn, J.M. Supra molecular Chemistry: Scope and Perspectives Molecules – Supra molecules -Molecular Devices. Angew. Chem., Int.Ed. Engl., 27, 89-112 (1988).
5. Lehn, J. M. Supra molecular Chemistry: Concept and Perspectives; VC Publishers: Weinheim (1995).
6. Cram, D. J. The Design of Molecular Hosts, Guests, and Their A. Complexes. Angew. Chem., Int. Ed. Engl., 27, 1009-1020 (1988).
7. Marko Marjanovic, Marijeta Kralj, Fran Supek, Leo Frkanec, Ivo Piantanida, TomislavSmuc and Ljertka Tusek-Bozic. J. Med. Chem., 50, 1007-1018 (2007).
8. Agrawal, Singh, J., Khadikar, P.V. Bioorg. Med. Chem., 10, 3981 (2002).
9. Vijay K. Agrawal, Jyoti Singh, Madhu Gupta an V.K.; d Padmakar V. Khadikar QSAR Study on the Estimation of Lipophilicity Using Distance-Based Topological Indices: Dominating Role of Equalized Electro negativity, Bio.org. Med. Chem., 13, 2109-2120 (2005).
10. The Topology of Molecule and Its LipophilicityVijay K. Agrawal, Jyoti Singh, Bruno Louis and Padmakar V. Khadikar Current Computer-Aided Drug Design, (CCADD), 2, 369-403 (2006).
11. Ewelina R., Karolina P., and Krzysztof J., Acta Poloniae Pharmaceutical-drug Research, Vol.70 No 1 pp. 3-18 (2013).
12. Agrawal, V.K.; Singh, J.; Louis, B.; Joshi, A. and Khadikar P.V.; Current Computer-Aided drug Design, Vol. 2, No.4, 31-35 (2006).
13. Khadikar P.V.; Karmakar S., Gour, K.; Agrawal, V.K.; Singh, S., Oxi. Comm., 27, No.1, 1-11 (2004).
14. Khadikar P.V.; Mandloi, D., Bajaj, A.V.; Joshi, S., Bio org. Med. Chem. Lett., 13, 419 (2003).
15. Todeschini R., Consonni V.; Handbook of Molecular Descriptors Wiley-VCH. Weinhum (GER) in “Methods and Principles in Medicinal Chemistry, 11, 667 (2000).
16. Karelson M., Molecular Descriptors in QSAR/QSPR, J. Wiley and Sons, New York (NY), 430 (2000).
17. Trinajestic, N.; Chemical graph theory, 2nd edition, CRC Press, BocaRoton,Florida, chapter 10, 225-273 (1992).
18. Diudea, M.V.; (Ed.) QSAR/QSPR studies by molecular descriptors Babes-BolyaiUniversityClujRomonia, (2000).
19. Hansch, C.; Leo, A.; Hockman, D.; Eds. Exploring QSAR: Hydrophobic, Electronic and Steric Constant; ACS Professional reference book, Washington, DC, P. 110 (1995).
20. Hansch, C.; The QSAR paradigm in the design of less toxic molecules, Drug Me tab, Rev., 1984-1985, 15: 1279-1294.
21. Chatterjee, S.; Hadi, A.S.; Price, B.; “Regression analysis examples”3rd ed.; Wiley: NewYork, (2000).
22. DRAGON Software. (Version 5.0 – 2004).
23. Balaban A.T., Chem. Phys. Lett., 89, 399 (1982).
24. Chem. Sketch ACD / Labs Software (11.01 version) www.acdlabs.com
25. Pogliani, L.J.; Phys. Chem., 100, 18065–18077 (1996).