

Topological models for the prediction of host: guest ratio of urea inclusion compounds

Seema Thakral · A. K. Madan

Received: 19 February 2009 / Accepted: 10 April 2009 / Published online: 2 May 2009
© Springer Science+Business Media B.V. 2009

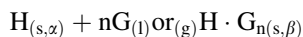
Abstract Correlation between composition (expressed as molal host: guest ratio) of urea inclusion compounds and molecular descriptors was studied. *Molecular connectivity index*, an adjacency-based topological descriptor, *Wiener's index*—a distance-based topological descriptor and *eccentric connectivity index*—an adjacency-cum-distance based topological descriptor were employed for the present investigations. A dataset comprising of 73 aliphatic compounds of diverse nature was selected for the present study. Values of all the three topological indices for all the compounds constituting the dataset were computed using an in-house computer program. All the possible uniparametric and biparametric mathematical models were attempted and internally cross-validated using leave one out (LOO) procedure. Among, the topological models investigated in the present study, uniparametric model based on molecular connectivity index yielded the best statistical parameters.

Keywords Eccentric connectivity index · Molal ratio · Molecular connectivity index · QSPR · Topological descriptors · Urea inclusion compounds · Wiener index

Introduction

Ever since the discovery of urea inclusion compounds, their chemistry has received much attention and a considerable research is invested in them. Many different families of guest molecules based on long alkane chains are known to form inclusion compounds with urea and X-ray diffraction studies confirmed that the host structure in these crystalline solids contain linear parallel tunnels [1, 2]. The urea inclusion compounds are of practical significance in purification, separation and stabilization technology [3]. Furthermore they are of use in basic scientific research from the point of view of studies of structural, energetic and dynamic properties of guest molecules [4]. They are the ideal representative systems for molecular isolation and separation as well as their crystalline environment is useful for the studies of molecular motions and interactions and chemical reactions [5].

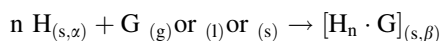
The process of formation of an inclusion compound may be depicted as follows:



where g, l and s represent gas, liquid and solid state respectively, α represents the nonporous phase of the pure host; H is the apophost (pure host; it might represent preformed cavities or it might be non-porous), and β is the phase of the host-guest compound with guest: host ratio n . When β phase decomposes upon heating, it may do so either in multiple step via formation of an intermediate γ -phase of composition $H \cdot G_{(n-x)_{(s,\gamma)}}$ or directly reverting to separate nonporous α -phase host and the guest [6]. However, with respect to urea inclusion compounds, which are a polymolecular kind of inclusion compound [7], the above equation may be modified as follows:

S. Thakral
PDM College of Pharmacy, Bahadurgarh, India

A. K. Madan (✉)
Faculty of Pharmaceutical Sciences, M.D. University, Rohtak
124-001, India
e-mail: madan_ak@yahoo.com



where α represents the tetragonal phase of urea while β represents the phase of hexagonal host–guest compound with host: guest ratio n . Composition of urea adducts are frequently expressed as molal ratio m [8] which is defined as $m = m_1/m_2$, where m_1 is the number of moles of urea containing m_2 moles of guest component or as the number of moles of urea required to enclose each mole of guest component. Molal fraction and molal ratio are also used frequently in experimentation involving theoretical considerations because it gives a measure of the relative proportion of moles of each constituent in a system [9].

In a structure like urea inclusion compounds, where one component occupies channels of indefinite length formed by the other, the molal ratio of the components depends on the length of the molecules of the endocytic component [10]. The knowledge of molal ratio m together with sufficient information regarding activity coefficients permits the prediction of the maximum yield and various other basic data for applications in which a single substance is precipitated [10]. In instances where the guest species is a drug moiety, an insight into host: guest molal ratio can be helpful in predicting expected proportion of drug loading in urea channels. A mathematical model was proposed by Redlich et al. to calculate molal ratio m (urea/reactant) for n -paraffin guests, which yielded values in good agreement (within the experimental errors of a few percent) with the values calculated from the results of X-ray analysis [8, 11]. However, the particular model was proposed to be applicable only for n -paraffins having number of carbon atoms (n) $6 < n < 17$. In the present paper, topological models have been developed for prediction of the molal host: guest ratio of a urea inclusion compound for linear/branched adductible molecules of diverse nature utilizing molecular topological descriptors.

Molecular topological descriptors represent an important advancement in automated computer treatment of chemical structures and QSPR through application of a mathematical technique, namely Graph Theory, to chemistry [12, 13]. In chemical graph theory, molecular structures are represented as hydrogen-suppressed graphs, commonly known as molecular graphs, in which the atoms are represented by vertices and the bonds by edges. The connections between atoms can be described by various types of topological matrices (e.g. distance or adjacency matrices), which can be mathematically manipulated so as to derive a single number, usually known as graph invariant or topological index. In consequence, the topological indices can be defined as two-dimensional descriptors that can be easily calculated from the molecular graphs, and do not depend on the way the graph is depicted or labeled.

They offer a simple way of measuring molecular branching, shape and size [14]. This graph description of molecules neglects information on bond lengths, bond angles, and torsion angles, but is able to encode in a numerical form the important atom connectivity information that determine a wide range of physical, chemical, and biological properties [15]. The paradigm of the use of topological indices and molecular descriptor in general in quantitative structure-activity relationship (QSAR) and quantitative structure-property relationship (QSPR) studies was defined as “it is desired to have as many molecular descriptors as possible at our disposition, but is preferred to include as few of them as possible in the QSPR and QSAR models to be developed” [16].

Recently, reasonably good results were obtained for prediction of adductability in urea for a dataset of diverse range of branched aliphatic and cyclic guest compounds using three topological descriptors, *molecular connectivity index*—an adjacency based topological descriptor, *Wiener's index*—a distance based topological descriptor, and *eccentric connectivity index*—an adjacency plus distance based topological descriptor [17, 18]. In the present study, the relationship of these three descriptors with host: guest molal ratio of urea inclusion compounds having aliphatic compounds as the guest, have been investigated.

Methodology

Calculations of topological indices

Molecular connectivity index (χ) is defined as the sum over all the edges (ij) as per following:

$$\chi = \sum_{\substack{i=1 \\ j=1}}^n (V_i V_j)^{-1/2}$$

where V_i and V_j are the degrees of adjacent vertices i and j and n is the number of vertices in a hydrogen depleted molecular structure [19]. The vertex degree or valency of atom i (V_i) is the number of bonds incident on the vertex i .

Wiener index (W) is defined as sum of the distances between all the pairs of vertices in a hydrogen-suppressed molecular graph

$$W = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n P_{ij}$$

where P_{ij} is the length of the path that contains the least number of edges between vertex i and vertex j in the graph G and n is the maximum possible number of i and j [20].

Eccentric connectivity index (ζ^c) is defined as the summation of the product of eccentricity and the degree of each vertex in the hydrogen depleted molecular graph having n vertices.

$$\zeta^c = \sum_{i=1}^n (E_i \times V_i)$$

where E_i is the eccentricity of the vertex i , V_i is the degree of vertex i and n is the number of the vertices in graph G . The eccentricity E_i of a vertex i in a graph G is the path length from vertex i to vertex j that is farthest from i ($E_i = \max d(ij); j \in G$); d_{ij} is the length of the shortest path between vertices i and j . The eccentric connectivity index takes into consideration the eccentricity as well as valency of the vertices in a hydrogen depleted graph [21].

Model design

A dataset comprising of 73 compounds of diverse nature with the reported composition of their urea inclusion complex as molal ratio m (urea: guest) was extracted from the literature [8, 22]. Values of χ , W and ζ^c for all the compounds in the dataset were computed using an in-house computer program. Subsequently, mathematical models were developed for prediction of molal host: guest ratio of urea inclusion compounds using these descriptors. The quality of models was expressed by fitted (descriptive) statistical parameters: the correlation coefficient r_{fit} , the standard error of estimate s_{fit} and F is the result of Fischer's test. In addition, the models were internally cross-validated by a leave-one-out procedure (LOO). Statistical parameters for the cross-validated models were denoted as r_{cv} and s_{cv} where cv stands for cross-validation procedure [23].

The standard error of estimate, s_{fit} and s_{cv} , were computed using the following expressions:

$$s_{\text{fit}} = \left\{ \left[\sum_{i=1}^N (P_i - P(\text{fit})_i^{\text{est}})^2 \right] / (N - I - 1) \right\}^{1/2}$$

$$s_{\text{cv}} = \left\{ \left[\sum_{i=1}^N (P_i - P(\text{cv})_i^{\text{est}})^2 \right] / (N - I - 1) \right\}^{1/2}$$

where P_i denotes experimental properties, $P(\text{fit})_i^{\text{est}}$ estimated properties based on the fit statistical procedure, $P(\text{cv})_i^{\text{est}}$ estimated properties based on the leave-one-out cross-validation procedure, N is the number of data points used in the model building and I is the number of descriptors contained in the model [23].

The following mathematical model was developed for the estimation of m using molecular connectivity index (χ):

$$m = 1.4132(\pm 0.0375)\chi + 0.42986(\pm 0.2455) \\ N = 73 \quad r_{\text{fit}} = 0.952 \quad s_{\text{fit}} = 0.762 \quad r_{\text{cv}} = 0.935 \quad (\text{Model I}) \\ s_{\text{cv}} = 0.891 \quad F = 1422.74$$

The following mathematical model was developed for the estimation of m using Wiener index (W):

$$m = 0.00496(\pm 0.00037)W + 6.4811(\pm 0.02876) \\ N = 73 \quad r_{\text{fit}} = 0.722 \\ s_{\text{fit}} = 1.842 \quad r_{\text{cv}} = 0.691 \quad s_{\text{cv}} = 2.317 \quad F = 183.93 \\ (\text{Model II})$$

The following mathematical model was developed for the estimation of m using eccentric connectivity index (ζ^c):

$$m = 0.0169(\pm 0.0006)\zeta^c + 5.2136(\pm 0.1957) \\ N = 73 \quad r_{\text{fit}} = 0.904 \quad s_{\text{fit}} = 1.085 \quad r_{\text{cv}} = 0.851 \\ s_{\text{cv}} = 2.156 \quad F = 665.84 \\ (\text{Model III})$$

Statistical analysis (Table 1) showed that pair of indices χ and ζ^c as well as W and ζ^c are appreciably correlated and indices χ and W are weakly correlated ($0.9 < r < 0.97$, appreciably correlated; $0.5 < r < 0.89$, weakly correlated) [23]. Hence a single possible biparametric model involving χ and W was investigated in the present study. Following topological model was developed for the estimation of m using χ and W .

$$m = 1.3116(\pm 0.0693)\chi + 0.00048(\pm 0.0003)W \\ + 0.797(\pm 0.334) \\ N = 73 \quad r_{\text{fit}} = 0.954 \quad s_{\text{fit}} = 0.752 \quad r_{\text{cv}} = 0.894 \\ s_{\text{cv}} = 0.898 \quad F = 732.85 \\ (\text{Model IV})$$

The results indicating index values and predicted m for models I–IV are compiled in Table 2.

Results and discussion

In the present study a set of three commonly used topological descriptors of diverse nature has been utilized for the development of models for the prediction of composition of urea inclusion compounds. Thorough study of the

Table 1 Intercorrelation matrix

Topological descriptors	χ	W	ζ^c
X	1	0.846	0.967
W		1	0.927
ζ^c			1

Table 2 Composition of urea inclusion compounds expressed as host: guest molal ratio m using topological models

S. no.	Category/compound name	χ	W	ζ^c	Predicted m using model				Reported m
					I	II	III	IV	
<i>Alkanes/alkenes</i>									
1	<i>n</i> -Nonane	4.41	120	96	6.66	7.08	6.84	6.64	7.4
2	<i>n</i> -Decane	4.91	165	122	7.37	7.3	7.28	7.32	8.1
3	1-Decene	4.91	165	122	7.37	7.3	7.28	7.32	7.8
4	1-Undecene	5.41	220	150	8.07	7.57	7.75	8	8.8
5	<i>n</i> -Heptane	3.414	552	54	5.25	9.22	6.13	5.54	6.1
6	<i>n</i> -Octane	3.914	964	74	5.96	11.26	6.46	6.4	6.73
7	1-Octene	3.914	964	74	5.96	11.26	6.46	6.4	6.3
8	2-Octene	3.914	964	74	5.96	11.26	6.46	6.4	6.5
9	Dodecane	5.91	286	182	8.78	7.9	8.29	8.69	9.3
10	Dodecene-1	5.91	286	182	8.78	7.9	8.29	8.69	9.3
11	Hexadecane	7.91	680	338	11.61	9.85	10.93	11.5	11.95
12	Hexadecene-1	7.91	680	338	11.61	9.85	10.93	11.5	12
13	1-Tetradecene	6.914	455	254	10.2	8.74	9.51	10.1	11.3
14	Octadecene	8.914	969	434	13.02	11.29	12.55	13	14.1
15	Undecane	5.41	220	150	8.07	7.57	7.75	8	8.7
16	Tetracosane	11.914	2300	794	17.26	17.89	18.63	17.5	17.8
17	Octacosane	13.914	3654	1094	20.09	24.6	23.7	20.8	21.8
<i>Haloalkanes</i>									
18	1-Bromodecane	5.414	220	150	8.08	7.57	7.75	8.01	8.9
19	2-Bromodecane	5.27	212	139	7.88	7.53	7.56	7.81	7.6
20	Lauryl chloride	6.414	364	216	9.49	8.29	8.86	9.39	10.6
21	Lauryl alcohol	6.414	364	216	9.49	8.29	8.86	9.39	9.5
22	Lauryl bromide	6.414	364	216	9.49	8.29	8.86	9.39	10.4
23	1-Chlorooctane	4.414	122	96	6.67	7.09	6.84	6.65	7.7
24	1-Chlorododecane	6.414	364	216	9.49	8.29	8.86	9.39	10
25	1-Chlorotetradecane	7.414	562	294	10.9	9.27	10.18	10.8	11.6
26	1-Bromooctane	4.414	122	96	6.67	7.09	6.84	6.65	7.2
27	1,6-Difluorohexane	3.914	84	74	5.96	6.9	6.46	5.97	5.6
28	1,4-Dichlorobutane	2.914	35	38	4.55	6.65	5.86	4.64	4.5
29	1,6-Dichlorohexane	3.914	84	74	5.96	6.9	6.46	5.97	6
30	1,6-Dibromohexane	3.914	84	74	5.96	6.9	6.46	5.97	5.3
<i>Alcohols</i>									
31	Octanol-2	4.27	114	87	6.46	7.05	6.68	6.45	6.4
32	Octanol	4.41	120	96	6.66	7.08	6.84	6.64	6.7
33	1-Octanol	4.41	120	96	6.66	7.08	6.84	6.64	6.85
34	1-Decanol	5.414	222	152	8.08	7.58	7.78	8.01	8.1
35	1-Dodecanol	6.414	364	216	9.49	8.29	8.86	9.39	9.25
36	1-Tetradecanol	7.414	562	294	10.9	9.27	10.18	10.8	10.7
37	1-Hexadecanol	8.414	816	384	12.32	10.53	11.7	12.2	12.3
38	6-Tridecanol	6.808	420	231	10.05	8.56	9.12	9.93	10.1
39	7-Tetradecanol	7.308	518	269	10.76	9.05	9.76	10.6	10.7
40	5-Decanol	5.308	200	133	7.93	7.47	7.46	7.86	7.8
<i>Ethers/ketones</i>									
41	Di- <i>n</i> -butyl ether	4.414	120	96	6.67	7.08	6.84	6.65	6.5
42	Methyl ethyl ketone	2.27	18	19	3.64	6.57	5.54	3.78	3.78
43	3-Pentanone	2.828	212	139	4.42	7.53	7.56	4.61	4.78

Table 2 continued

S. no.	Category/compound name	χ	W	ξ^c	Predicted m using model				Reported m
					I	II	III	IV	
44	4-Heptanone	3.828	75	61	5.84	6.85	6.24	5.86	6.03
45	1,4-Dioxane	3	27	36	4.67	6.62	5.82	4.75	2.15
<i>Carboxylic acids</i>									
46	Undecylenic acid	6.27	354	203	9.29	8.24	8.64	9.19	7.9
47	Caprylic acid	4.77	158	111	7.17	7.26	7.09	7.13	6.6
48	Myristic acid	7.77	667	321	11.41	9.79	10.64	11.3	11.1
49	Butanoic acid	2.77	32	31	4.34	6.64	5.74	4.45	3.98
50	Pentanoic acid	3.27	52	47	5.05	6.74	6.01	5.11	4.63
51	Oleic acid	9.77	1313	521	14.23	12.99	14.02	14.3	12.2
52	Hexanoic acid	3.77	89	65	5.76	6.92	6.31	5.79	5.5
53	Heptanoic acid	4.27	114	87	6.46	7.05	6.68	6.45	6
54	Nonanoic acid	5.27	212	139	7.88	7.53	7.56	7.81	7.62
55	Dodecanoic acid	6.77	444	239	10	8.68	9.25	9.89	10
56	Palmitic acid	8.77	954	415	12.82	11.21	12.23	12.8	12.8
57	Stearic acid	9.77	1313	521	14.23	12.99	14.02	14.3	14.2
58	Sebacic acid	6.626	433	224	9.79	8.63	9	9.7	10.9
<i>Esters</i>									
59	Methyl undecylenate	6.808	435	237	10.05	8.64	9.22	9.94	9.5
60	Methyl stearate	10.308	1506	577	14.99	13.95	14.97	15.1	14.8
61	Methyl oleate	10.308	1506	577	14.99	13.95	14.97	15.1	14.5
62	Methyl linoleate	10.308	1506	577	14.99	13.95	14.97	15.1	14.2
63	Methyl linolenate	10.308	1506	577	14.99	13.95	14.97	15.1	13.7
64	Methyl tetradecanoate	8.308	790	365	12.17	10.4	11.38	12.1	12.2
65	Methyl hexadecanoate	9.308	1110	465	13.58	11.99	13.07	13.6	13.4
66	Methyl butanoate	3.328	52	45	5.13	6.74	5.97	5.19	5.4
67	Methyl caprylate	5.328	226	137	7.96	7.6	7.53	7.9	8.3
68	Methyl laurate	7.328	538	277	10.78	9.15	9.9	10.7	10.5
69	Dimethyl adipate	5.722	254	152	8.51	7.74	7.78	8.43	7.3
70	Dimethyl sebacate	7.702	632	300	11.31	9.62	10.28	11.2	10.1
71	Diethyl sebacate	8.702	891	388	12.72	10.9	11.77	12.7	11.6
<i>Nitrogen containing compounds</i>									
72	1,6-Diaminohexane	3.914	84	74	5.96	6.9	6.46	5.97	5.5
73	Subnitrile	4.91	165	122	7.37	7.3	7.28	7.32	5.7

topological models developed as above indicate that it is possible to establish a reasonably accurate quantitative structure–property relationship to estimate molal ratio of urea inclusion compound based on a simple set of molecular topological descriptors. On the basis of value of correlation coefficient r for these models, molal ratio of urea inclusion compounds can be considered to be appreciably correlated with the two of the topological descriptors selected for the study i.e. molecular connectivity index (χ) and eccentric connectivity index (ξ^c) ($0.90 \leq r \leq 0.97$ appreciably correlated) while weakly correlated with descriptor W ($0.50 \leq r \leq 0.89$ weakly correlated) [23].

The biparametric model developed utilizing descriptors χ and W also appears to predict molal ratio of urea inclusion compounds with reasonable accuracy. Judging by the statistical parameters, Model I appears to be the best model, which explains almost 90.6% of the variance in the host: guest ratio for the series of compounds included in the study. Quite proximal values of r_{fit} and r_{cv} and values of s_{fit} and s_{cv} indicate the stability of the models. A plot of reported vs predicted m for the fit and cross-validated values for Model I–IV is shown in Fig. 1, 2, 3, 4.

It is generally accepted that in the adduct, urea molecules are connected into spirals by hydrogen bonds

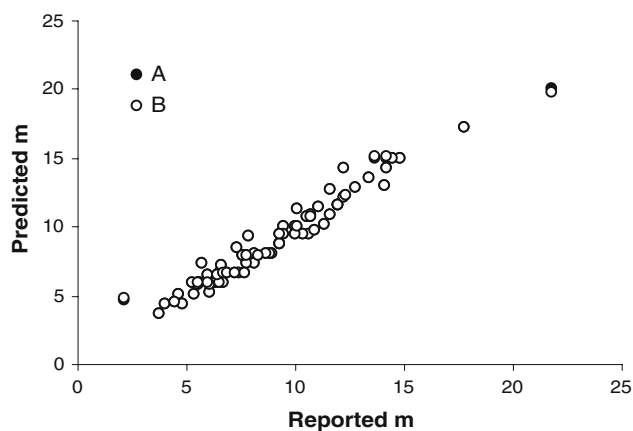


Fig. 1 Scatter plots between the reported and predicted host: guest molal ratio m of urea inclusion compounds for the fit A and cross-validated B [Model I](#)

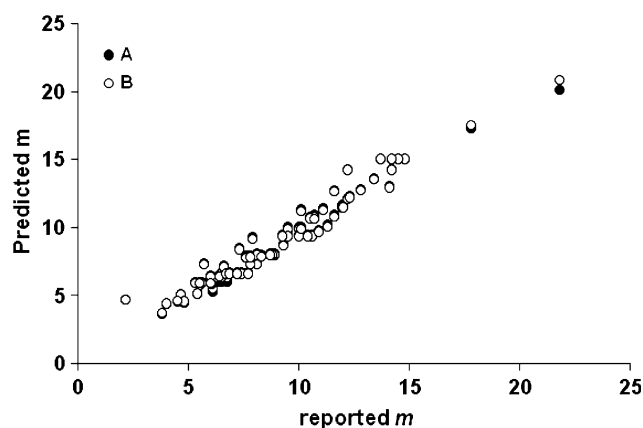


Fig. 4 Scatter plots between the reported and predicted host: guest molal ratio m of urea inclusion compounds for the fit A and cross-validated B [Model IV](#)

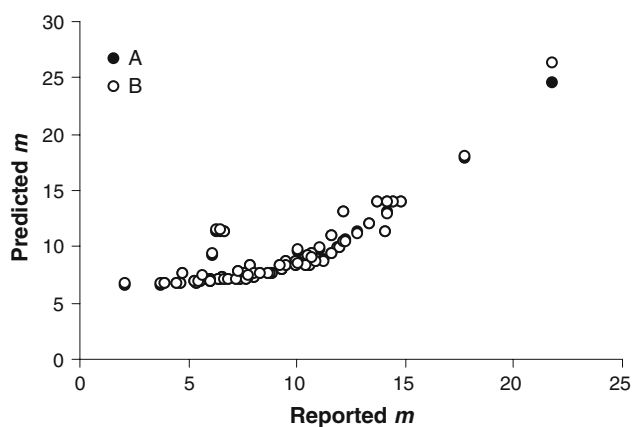


Fig. 2 Scatter plots between the reported and predicted host: guest molal ratio m of urea inclusion compounds for the fit A and cross-validated B [Model II](#)

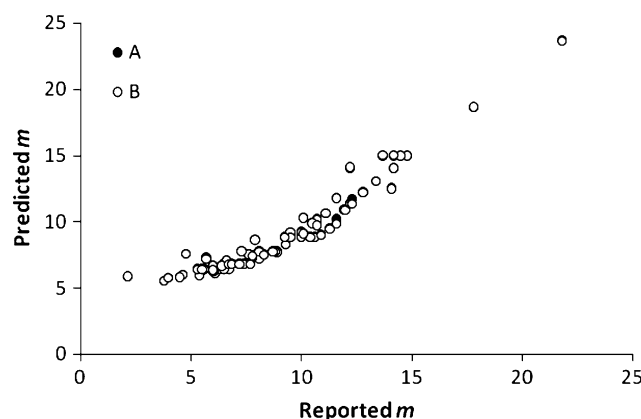


Fig. 3 Scatter plots between the reported and predicted host: guest molal ratio m of urea inclusion compounds for the fit A and cross-validated B [Model III](#)

between oxygen and amino group of adjacent urea molecules, resulting in a hexagonal channel into which the adducting molecule can fit [24]. The size of the channel limits the molecules that may adduct to those having cross-sectional dimensions equal to, or less than those of the channel in the urea adduct, but in certain cases, distortion of the urea lattice may occur and slightly larger molecules can adduct [25, 26]. These distortions in urea channels may lead to increase in the amount of urea required to enclose the adducting guest culminating to higher host: guest ratio.

Suitable topological models have already been developed earlier for the prediction of adductability in urea and for prediction of heat of decomposition of urea inclusion compound for a dataset of diverse range of branched aliphatic compounds using the above three topological descriptors, wherein, the models based on molecular connectivity index yielded the best statistical results [27]. Thus calculation of a topological index as simple as molecular connectivity index can lead to prediction regarding adductability of a guest compounds, prediction of the heat of decomposition as well as the composition of the resulting urea inclusion compound, giving an insight into the stability and kinetics of inclusion complex formation.

References

- Smith, A.E.: The crystal structure of urea-hydrocarbon complexes. *Acta Crystallogr.* **5**, 224 (1952). doi:[10.1107/S0365110X52000629](https://doi.org/10.1107/S0365110X52000629)
- Harris, K.D.M., Thomas, J.M.: Structural aspects of urea inclusion compounds and their investigation by X-ray diffraction: a general discussion. *J. Chem. Soc., Faraday Trans.* **86**, 2985–2996 (1990). doi:[10.1039/ft9908602985](https://doi.org/10.1039/ft9908602985)
- Hayes, D.: Urea inclusion compound formation. *Inform* **13**, 781–801 (2002)

- Harris, K.D.M.: Fundamental and applied aspects of urea and thiourea inclusion compounds. *Supramol. Chem.* **9**, 47–53 (2007). doi:[10.1080/10610270600977706](https://doi.org/10.1080/10610270600977706)
- Harris, K.D.M.: Meldola lecture: understanding properties of urea and thiourea inclusion compounds. *Chem. Soc. Rev.* **26**, 279–289 (1997). doi:[10.1039/cs9972600279](https://doi.org/10.1039/cs9972600279)
- Nassimbeni, L.R.: Inclusion compounds: selectivity, thermal stability and kinetics. In: Atwood, J.L., Steed, J.W. (eds.) *Encyclopedia of Supramolecular Chemistry*, vol. 2, pp. 696–703. Marcel Dekker, New York (2004)
- Frank, S.: Inclusion compounds. *J. Pharm. Sci.* **64**, 1585–1604 (1975)
- Redlich, O., Gable, C.M., Dunlop, A.K., Millar, R.W.: Addition compounds of urea and organic substances. *J. Am. Chem. Soc.* **72**, 4153–4160 (1950). doi:[10.1021/ja01165a087](https://doi.org/10.1021/ja01165a087)
- Sinko, P.J.: *Martin's Physical Pharmacy and Pharmaceutical Sciences*, p. 121. Lippincot Williams and Wilkins, Philadelphia (2005)
- Bhatnagar, V.M.: Clathrate compounds of urea and thiourea. *J. Struct. Chem.* **8**(3), 513–529 (1968). doi:[10.1007/BF00751656](https://doi.org/10.1007/BF00751656)
- Smith, A.E.: *J. Chem. Phys.* **18**, 150 (1950). doi:[10.1063/1.1747442](https://doi.org/10.1063/1.1747442)
- Gutman, I., Polansky, O.E.: *Mathematical Concepts in Organic Chemistry*. Springer-Verlag, Berlin (1987)
- Trinajstić, N.: *Chemical Graph Theory*. CRC Press, Boca Raton, FL (1992)
- van de Waterbeemd, H., Carter, R.E., Grassy, G., Kubinyi, H., Martin, Y.C., Tute, M.S., Willett, P.: *Annu. Rep. Med. Chem.* **33**, 397 (1998). doi:[10.1016/S0065-7743\(08\)61102-1](https://doi.org/10.1016/S0065-7743(08)61102-1)
- Estrada, E., Uriate, E.: Recent advances on the role of topological indices in drug discovery research. *Curr. Med. Chem.* **8**, 1573–1588 (2001)
- Estrada, E.: Novel strategies in the search of topological indices. In: Devillers, J., Balaban, A.T. (eds.) *Topological Indices and Related Descriptors in QSAR and QSPR*, pp. 403–453. Gordon and Breach Sci. Pub., Amsterdam (1999)
- Thakral, S., Madan, A.K.: Topological models for prediction of adductability of branched aliphatic compounds in urea. *J. Incl. Phenom. Macrocycl. Chem.* **56**, 405–412 (2006). doi:[10.1007/s10847-006-9123-0](https://doi.org/10.1007/s10847-006-9123-0)
- Thakral, S., Madan, A.K.: Topological models for prediction of adductability of substituted cyclic organic compounds in urea. *J. Incl. Phenom. Macrocycl. Chem.* **58**, 321–326 (2007). doi:[10.1007/s10847-006-9160-8](https://doi.org/10.1007/s10847-006-9160-8)
- Randic, M.: On characterization of molecular branching. *J. Am. Chem. Soc.* **97**, 6609–6615 (1975). doi:[10.1021/ja00856a001](https://doi.org/10.1021/ja00856a001)
- Wiener, H.: Correlation of heat of isomerization, and difference in heat of vaporization of isomers, among paraffin hydrocarbons. *J. Am. Chem. Soc.* **69**, 2636–2638 (1947). doi:[10.1021/ja01203a022](https://doi.org/10.1021/ja01203a022)
- Sharma, V., Goswami, R., Madan, A.K.: Eccentric connectivity index: a novel highly discriminating topological descriptor for structure - property and structure- activity studies. *J. Chem. Inf. Comput. Sci.* **37**, 273–282 (1997). doi:[10.1021/ci960049h](https://doi.org/10.1021/ci960049h)
- Schlenk, W.: Urea addition of aliphatic compounds. *Ann. Chem.* **565**, 204–240 (1949)
- Nikolic, S., Kovacevic, G., Milicevic, A., Trinajstić, N.: The Zagreb indices 30 years after. *Croat. Chem. Acta* **76**, 112–124 (2003)
- Takemoto, K., Sonoda, N.: Inclusion compounds of urea, thiourea and selenourea. In: Atwood, J.W., Davis, J.E.D., MacNicol, D.D. (eds.) *Inclusion compounds*, vol. 2, pp. 47–67. Academic Press, London (1984)
- Speight, J.G.: *Chemistry and Technique of Petroleum*, pp. 233–234. CRC Press, Boca Raton, FL (2006)
- Findlay, R.A.: Adductive crystallization. In: Schoen, H.M., Mcketta, J.J. (eds.) *Interscience Library of Chemical Engineering and Processing. New Chemical Engineering Separation Techniques*, vol. 1, pp. 257–318. Interscience Publishers, New York (1962)
- Thakral, S., Madan, A.K.: Topological models for prediction of heat of decomposition of urea inclusion compounds containing aliphatic endocycles. *J. Incl. Phenom. Macrocycl. Chem.* **60**, 187–192 (2008). doi:[10.1007/s10847-007-9345-9](https://doi.org/10.1007/s10847-007-9345-9)