CHEMIA

S T U D I A

UNIVERSITATIS BABEŞ–BOLYAI

CHEMIA

4

Volume dedicated to Professor Mircea V. Diudea at the 60th Anniversary

CUPRINS – CONTENT – SOMMAIRE – INHALT

Studia Universitatis Babes-Bolyai Chemia has been selected for coverage in Thomson Reuters products and custom information services. Beginning with V. 53 (1) 2008, this publication is indexed and abstracted in the following:

- Science Citation Index Expanded (also known as SciSearch®)
- Chemistry Citation Index®
- Journal Citation Reports/Science Edition

Professor Mircea V. Diudea is born in 1950, November, 11, in Silivas, a village of Transylvania, Romania, in a family of primary school teachers. He followed the high school "Nicolae Balcesc", Cluj and next the Faculty of Chemistry, University of Cluj (1969-1974). The PhD thesis, entitled "Phenothiazines and related structures" (performed in organic synthesis, under the guiding of two bright Chemists, Professors Valer Farcasan and Cornel Bodea) was defended in 1979, at Institute of Chemistry, Cluj. He worked six years (1974-1980) as a Chemist at "Terapia" Drug Factory, Cluj and the next seven years (1980-1987) as Researcher, at Chemical-Pharmaceutical Research Institute, Cluj. From 1987 was admitted at the Faculty of Chemistry and Chemical Engineering, of "Babes-Bolyai" University, Cluj, as Assistant Professor (1987-1990), next as Associate Professor, (1990-1996) and from 1996 as full Professor, at the Department of Organic Chemistry. His main courses in Chemistry: Organic Chemistry and Biologically Active Compounds.

In 1986 he established the TOPO GROUP CLUJ, and in 2007 founded the European Society of Mathematical Chemistry ESMC, of which first president is. These data include a period of 21 years when a new interdisciplinary science called Molecular Topology has been developed under his guidance and resulted in publication of more than 250 scientific articles (Hirsch index 22 (ISI) or 26 (Scopus), with more than 1200 citations in ISI journals) and 14 authored or edited books, in three directions:

1. **Molecular Topology** (basic theory, with the main results including matrices: Cluj, Schell, Combinatorial, matriceal operators; topological indices: Cluj, Cluj-Ilmenau,Cluj-Tehran, super-index Cluj-Niš, centrality, centric connectivity, etc.; algorithms: for inter and intra-molecular ordering, topological symmetry (equivalence classes of subgraphs), for enumerating the Kekulé valence structures).

2. **QSAR/QSPR** (correlating studies, with contributions in: data reduction procedures, clustering procedures based on similarity 2D and 3D, optimal regression procedures, modeling various phyisico-chemical properties and biological activities, algorithms for similarity 2D and 3D and Drug Design).

3. **Nanoscience** (basic theory, with contributions in: Design of nanostructures by operations on maps and nets, the Romanian "Capra" being the first pro-chiral basic operation, rules of stability of fullerenes, a modified Euler

theorem in multi-shell nanostructures, original counting polynomials: Omega, Pi, Theta, Cluj, the new diamond D_5 (hoped to be produced in the NanoLab of Carbon Allotropes, organized under his guidance), a Gallery of molecular art, etc.).

The scientific activity in his group is supported by 10 original software programs. As didactical activity, Mircea Diudea delivered courses of Molecular Topology, QSAR/QSPR and Fullerenes and Nanostructures, basically at Master and PhD levels (with 12 PhD theses defended so far).

Professor Mircea Diudea is member of International Academy of Mathematical Chemistry (2005) and member of Editorial Board of: *Croatica Chemica Acta*, *MATCH*, *Commun. Math. Comput. Chem.*, *Internet El. J. Molec. Design, Carpath. J. Math., Iran. J. Math. Chem.*, *Acta Univ. Cibin. and* Senior Editor, *Int. J. Chem. Model.*, NOVA Publishers, New York, USA. Also he is a referee at the following Scientific Journals:

(1) *Rev. Roum. Chim.*, (2) *Studia Univ. Babes-Bolyai*, (3) *Croat. Chem. Acta*, (4) *J. Chem. Inf. Comput. Sci.*, (5) *Chem. Phys. Lett.*, (6) *Int. Elect. J. Mol. Design*, (7) *New J. Chem.*, (8) *SAR/QSAR Env. Res.*, (9) *Bioorg. Med. Chem. Lett.*, (10) *MATCH Commun. Math. Comput. Chem.* (11) *Fullerenes, Nanotubes Carbon Nanostruct.*, (l2) Molecules, (13) *J. Am. Chem. Soc.* (14) *Romanian Chemical Quarterly Reviews,* (15) *Ars Combinatorica*, (16) *Arkivoc*, (17) *Utilitas Math*, (18) *Eur. J. Operational Res.*, (19) *Math. Comput. Model.*, (20) *J. Math. Chem*.

In the list of his stages of international collaboration, invited lectures, conferences, the following are included: Zelinsky Institute of Organic Chemistry, Russian Academy, Moscow, Russia, Rudger Bošković Institute, Zagreb, Croatia, Central Chemical Research Institute of Hungarian Academy, Budapest, Hungary, University of Bayreuth, Germany (as a DAAD fellowship), Technical University of Ilmenau, Germany, University of Bielefeld, Germany (as a second DAAD fellowship), University of Kiel, Germany, Forschungszentrum Karlsruhe, Institut für Nanotechnology, Germany, University of Karlsruhe, Germany, University of Erlangen, Germany, University of Exeter, UK, University of Hiroshima, Japan, University of Sendai, Japan, University of Tsukuba, Japan (five months), Catholic University Leuven, Belgium, University of Gent, Belgium, University of Sheffield, U.K., University of Miskolc, Hungary, National Institute of Chemistry, Ljubljana, Slovenia, University of Valencia, Spain, University of Milano-Biccoca, Italy, University of Tehran, Iran, Technical University, Isfahan, Iran, University of Warsaw, Poland, Collegium Budapest, Hungary, University of Kashan, Iran, University of Shiraz, Iran, Tarbiat Modares University, Tehran, Iran, State University, St. Petersburg, Russia, University of Ljubljana, Slovenia).

Books (authored):

1. M.V. Diudea, S. Todor, F. Igna, *Aquatic Toxicology*. DACIA, Cluj, 1986 (in Romanian), 320p.
2. M.V. Diudea, M. Pitea, M. Butan, *Fenothiazines and structurally related drugs*. DACIA, Cluj, 1992 (in Romanian), 278p.
3. M.V. Diudea, O. Ivanciuc, *Molecular Topology*, COMPREX, Cluj, 1995 (in Romanian), 320p.
4. M.V. Diudea, I. Gutman, L. Jäntschi, *Molecular Topology*, NOVA, New York, 2002, 329p.
5. M.V. Diudea, M. S. Florescu, and P. V. Khadikar, *Molecular Topology and Its Applications*, EFICON, Bucharest, 2006, 381 pp. (Eficon Press, Bucuresti, ISBN 978-973-87904-0-7)
6. M.V. Diudea, Cs. L. Nagy, *Periodic Nanostructures*, SPRINGER, 2007 (207p).
7. M.V. Diudea, Nanomolecules and Nanostructures – Polynomials and Indices, M C M, No. 10, University of Kragujevac, 2010 (472p).

Edited Books (Ed):

1. M.V. Diudea, *QSPR/QSAR Studies by Molecular Descriptors*, NOVA, New York, 2001, 438p.
2. M.V. Diudea, *Nanostructures, Novel Architecture*, NOVA, New York, 2005, 420pp.

Journal Issues (Guest Ed):

1. M.V. Diudea, O. Ivanciuc, *MATCH Commun. Math. Comput. Chem.,* **2001**, *44*, 428p.
2. M.V. Diudea, *Internet Electron. J. Mol. Des.*, **2002**, *1* (1-6), 427p.
3. M.V. Diudea, *J. Math. Chem.*, **2009**, 45.
4. M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2008**, 60.
5. M.V. Diudea, *Symmetry: Culture and Science,* **2008**, 19.

Book Chapters

1. M.V. Diudea and G. Katona, Molecular Topology of Dendrimers. In: Newkome, G.A. Ed., *Advan. Dendritic Macromol.*, **1999**, *4*, 135-201.
2. O.M. Minailiuc and M.V. Diudea, TI-MTD Model. Applications in Molecular Design. In: M.V. Diudea, Ed., *QSPR/QSAR Studies by Molecular Descriptors*. NOVA, New York, **2001**, pp. 363-388.
3. Cs. L. Nagy, M.V. Diudea, and T.S. Balaban, Coalescence of fullerenes. In: M.V. Diudea, Ed., *Nanostructures-Novel Architecture*, NOVA, New York, **2005**, 25-60.

4. M.V. Diudea, Cs. L. Nagy and A. Graovac, Periodic Finite Nanostructures. In: M.V. Diudea, Ed., *Nanostructures-Novel Architecture*, NOVA, New York, **2005**, 61-84.
5. M.V. Diudea, Covering Nanostructures, In: M.V. Diudea, Ed., *Nanostructures-Novel Architecture*, NOVA, New York, **2005**, 203-242.
6. M. Stefu, D. Butyka, M.V. Diudea, L. Jantschi, and B. Parv, Algorithms for basic operations on maps, In: M.V. Diudea, Ed., Nanostructures-Novel Architecture, NOVA, New York, **2005**, 243-267.
7. M. Stefu and M.V. Diudea, Distance Counting in Tubes and Tori: Wiener Index and Hosoya polynomial. In: M.V. Diudea, (Ed.), *Nanostructures–Novel Architecture*, Nova, New York, **2005**, 127-165.
8. Cs. L. Nagy and M.V. Diudea, Nanoporous carbon structures, In: M.V. Diudea, (Ed.), *Nanostructures–Novel Architecture*, Nova, New York, **2006**, 311-334.
9. M.V. Diudea, Covering nanostructures, in: A. Graovac, I. Gutman, and D. Vukicevic, Eds., Mathematical Methods for Students of Chemistry and Biology, Proceedings of the MMC 2007, Hum, Zagreb, **2009**.
10. M.V. Diudea, Nanostructure design - between science and art, in: M. Putz, Ed., NOVA, **2009**.
11. E. Vizitiu and M.V. Diudea, C_{60} Structural Relatives-An Omega-aided Topological Study, in: A. Graovac, A. Ottorino, Eds., **2010**.
12. M.V. Diudea, Counting polynomials in partial cubes, in: I. Gutman, Ed., New Topological Descriptors, MCM series, MATCH, **2010**.
13. A. Ilić, F. Gholami-Nezhaad, A. Vizitiu and M.V. Diudea, Counting polynomials and indices in Nanocones, in: I. Gutman, Ed*., New Topological Descriptors*, MCM series, MATCH, **2010**.
14. M.V. Diudea, M. Petitjean, Symmetry in multi tori, in: M.V. Diudea (Ed), *Simmetry, Culture, Sci*., **2008**, *19* (4), 285-305.

Representative articles:

1. M.V. Diudea, Cluj polynomials, *J. Math. Chem*., **2009**, *45*, 295 -308.
2. M.V. Diudea, Omega Polynomial, *Carpath. J. Math*., **2006**, *22*, 43-47.
3. M.V. Diudea, Hosoya polynomial in tori. *MATCH Commun*. *Math*. *Comput*. *Chem*., **2002**, *45*, 109-122.
4. M.V. Diudea, D. Vukicevic, Kekule structure count in corazulenic fullerenes, *J. Nanosci. Nanotechnol.,* **2007**, *7*, 1321–1328.
5. Sh. Xu, H. Zhang, M.V. Diudea, Hosoya polynomials of zig-zag open-ended nanotubes, *MATCH Commun*. *Math*. *Comput*. *Chem*., **2007**, 57(2), 443-456.
6. R.B. King and M.V. Diudea, The chirality of icosahedral fullerenes: a comparison of the tripling (leapfrog), quadrupling (chamfering), and septupling (capra) transformations, *J. Math. Chem.,* **2006**, *39*(3-4), 597-604.

7. O. Ursu, M.V. Diudea and Sh-i. Nakayama, 3D Molecular similarity: method and algorithms, *J. Comput. Chem. Jpn.*, **2006**, *5*, 39–46.

8. M.V. Diudea, Nanoporous carbon allotropes by septupling map operations, *J. Chem. Inf. Model.,* **2005**, *45,* 1002-1009.

9. M.V. Diudea, Corannulene and corazulene tiling of nanostructures, *Phys. Chem., Chem. Phys.,* **2005**, *7*, 3626-3633.

10. M.V. Diudea, Covering Forms in Nanostructures, *Forma* (Tokyo), **2004**, *19* (3), 131-163.

11. M.V. Diudea, Stability of tubulenes, *Phys. Chem., Chem. Phys.*, **2004**, *6*, 332-339.

12. M.V. Diudea and Cs. L. Nagy, Euler Formula in Multi-Shell Polyhedra, *MATCH Commun. Math. Comput. Chem.*, **2008**, *60*, 835-844.

13. M.V. Diudea, Phenylenic and naphthylenic tori. *Fullerenes, Nanotubes Carbon Nanostruct.*, **2002**, *10*, 273-292.

14. M.V. Diudea, M. Ştefu, P. E. John, and A. Graovac, Generalized operations on maps, *Croat. Chem. Acta,* **2006**, *79*, 355-362.

15. P.E. John, A.E. Vizitiu, S. Cigher and M.V. Diudea, CI Index in Tubular Nanostructures, *MATCH Commun. Math. Comput. Chem.*, **2007**, *57*(2), 479-484.

16. M. Lillington, P.W. Fowler and M.V. Diudea, Patterns of counter-rotating ring currents in two valence isomers of corazulene, *Polish J. Chem.*, **2007**, *81*, 653-662.

17. E. Lijnen, A. Ceulemans, M.V. Diudea and Cs. L. Nagy, Double toroids as model systems for carbon nanotube junctions: Through-bond currents, *J. Math. Chem.*, **2009**, *45*, 417-430.

18. M.V. Diudea, A.E. Vizitiu, Cs.L. Nagy, T. Beu, A. Bende and Dušanka Janežič, Circulene covered fullerenes, *J. Mol. Struct. THEOCHEM*, **2009**, *904*, 28-34.

19. M.V. Diudea, Counting polynomials in tori T(4,4)S[c,n], *Acta Chim. Slo.*, **2010**, *57*, 551-558.

20. M.V. Diudea and A. R. Ashrafi, Shell-polynomials and Cluj-Tehran index in tori T(4,4)[5,n], *Acta Chim. Slo.*, **2010**, *57*, 559-564.

DIAMOND D$_5$, A NOVEL ALLOTROPE OF CARBON

MIRCEA V. DIUDEA[*]

ABSTRACT. Design of a hypothetical crystal network, consisting mostly of pentagon rings and called diamond D$_5$, is presented. It is shown that the seed and repeat-units, as hydrogenated species, show good stability, compared with that of C$_{60}$ fullerene, as calculated at three levels of theory (PM3, HF/6-31G(d,p), B3LYP/6-31G). The topology of the network is described in terms of the net parameter.

Keywords: *diamond D$_5$; nano-dendrimer; multi-tori; crystal-like network.*

INTRODUCTION

The nano-era, a period starting, since 1985 with the discovery of C$_{60}$, is dominated by the carbon allotropes, studied for applications in nano-technology. Among the carbon structures, fullerenes (zero-dimensional), nanotubes (one dimensional), graphene (two dimensional), diamond and spongy nanostructures (three dimensional) were the most studied [1,2]. Inorganic compounds also attracted the attention of scientists. Recent articles in crystallography promoted the idea of topological description and classification of crystal structures [3-7].

Diamond D$_6$, the classical, beautiful and useful diamond has kept its leading interest among the carbon allotropes, even as the newer "nano" varieties [8-10]. Along with electronic properties, the mechanical characteristics appear of great importance, as the composites can overpass the resistance of steel or other metal alloys. A lot of efforts were done in the production and purification of "synthetic" diamonds, from detonation products [11-14].

Dendrimers are hyper-branched nano-structures, made by a large number of (one or more types) substructures called monomers, synthetically joined within a rigorously tailored architecture [15-17]. They can be functionalized at terminal branches, thus finding a broad pallet of applications in chemistry, medicine, etc [18,19].

Multi-tori MT are structures of high genera [1,2,20], consisting of more than one tubular ring. They are supposed to result by self-assembly of some repeat units (*i.e.*, monomers) which can be designed by opening of cages/ fullerenes or by appropriate map/net operations. Such structures can appear in spongy carbons or in zeolites [20]. Spongy carbons have been recently synthesized [21,22].

---

[*] *Faculty of Chemistry and Chemical Engineering, "Babes-Bolyai" University, 400028 Cluj-Napoca, Romania, mpop@chem.ubbcluj.ro; diudea@gmail.com*

There are rigid monomers that can assembly in dendrimers, but the growing process stops rather at the first generation. At a second generation, yet the endings of repeat units are not free any more, they fit to each other, thus forming either an infinite lattice, if the unit symmetry is octahedral or a spherical multi-torus, if the unit symmetry is tetrahedral. The last one is the case of structures previously discussed in refs [23,24].

RESULTS AND DISCUSSION ULTI-TORUS DESIGN AND STABILITY

A tetrapodal monomer M_1(Figure 1, left), designed by $Trs(P_4(T))$ sequence of map operations [25-28] and consisting of all pentagonal faces, can self-arrange to form a dendrimer M_5, at the first generation stage (Figure 1, right).

M₁   **M₅**

Figure 1. Tetrapodal unit designed by $Trs(P_4(T))$ and the corresponding dendrimer, at 1st generation stage

The "growing process" is designed occurring by identifying the trigonal faces of two opposite M_1 units; at the second generation, six pentagonal hyper-cycles are closed, as in molecule M_{17}, Figure 1.

M₁₇   **M₅₇**
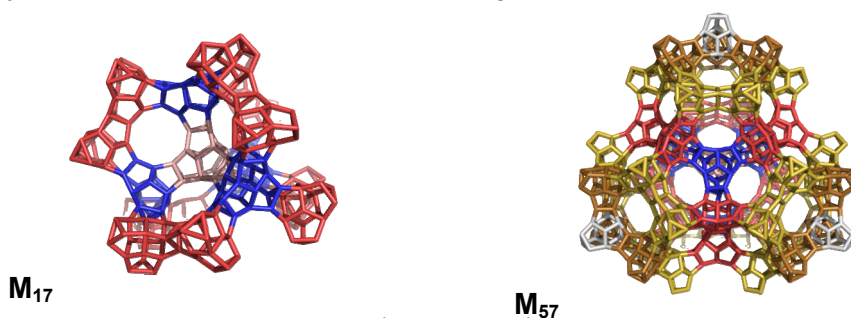
Figure 2. Dendrimer at 2nd (left) and 5th (right) generation stage; M_{57}=4S_MT; v=972; e=1770; f_5=684; g=58 (infinite structure); adding f_3=40, then g=38 (finite structure)

The process is imagined as a "dendrimer growing", and it is limited here at the fifth generation (Figure 2), when a tetrahedral array results: 4S_MT= M_{57}.

Multi-tori herein considered can be viewed either as infinite (*i.e.*, open) structures or as closed cages; then, it is not trivial to count the number of simple tori (*i.e.*, the genus *g*) in such complex structures.

The Euler's formula [29]: $v - e + f = 2(1-g)$, where *v*, *e* and *f* are the number of vertices/atoms, edges/bonds, and faces, respectively, is applicable only in case of single shell structures. In multi shell structures, we modified [30] the Euler formula as: $v - e + r - p(s-1) = 2(1-g)$, where *r* stands for the number of hard rings (*i.e.*, those rings which are nor the sum of some smaller rings), *p* is the number of smallest polyhedra filling the space of the considered structure while *s* is the number of shells. In case of an infinite structure, the external trigonal faces are not added to the total count of faces/rings. The calculated *g*-values are given in figures.

The number of tetrapodal monomers, added at each generation, up to the 5th one, realized as M_{57}, is: 1; 4; 12, 24, 12, 4. The connections in M_{57} are complex and to elucidate the large structures up to the fifth generation, the reduced graph drawing (Figure 3) was needed [24].
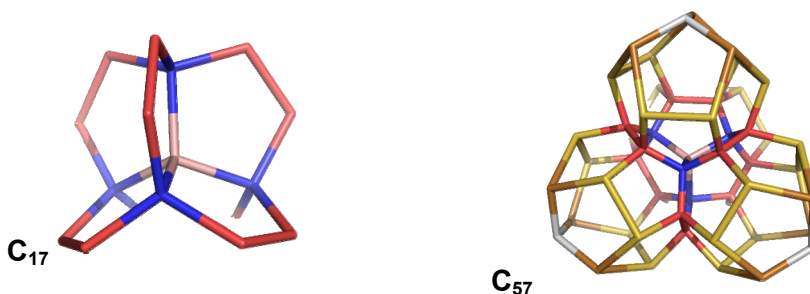
Figure 3. Reduced graphs at 2nd (left) and 5th (right) generation stage; C_{57}: v=57; e=94; r_5=42; g=0.5; $R(x)$=42x^5+82x^9+144x^{10}

The structure C_{17} (Figure 3, left) we call the "seed" of all the hereafter structures. The structure C_{57} (Figure 3, right) corresponds to 4S_MT, and is equivalent to 4 "condensed" dodecahedra, sharing a common point. By considering this common point as an internal shell *s*, the modified [30] Euler formula will give (for v=57; e=94; r=42; p=4 and s=2) a (non-integer) genus g=0.5. The ring polynomial R(*x*) is also given, at the bottom of Figure 3.

Energetic data, calculated at three levels of theory (Table 1) show a good stability of the structures shown in Figure 3 (as hydrogenated species), in comparing to C_{60}, the reference structure in Nano-science. Differences in HOMO-LUMO gap (in eV) in favor of the new (all sp^3 carbon atoms) structures, are just expected for hydrogenated (stable) species while the total energies are close comparable. The corresponding fullerene-like (containing both sp^3 and sp^2 carbon atoms) C_{57} is close to C_{60} as total energy while the gap vary, function of the considered approach: in PM3 and B3LYP the gap is lower for C_{57} than for C_{60} while in Hartree-Fock HF this parameter is in favor

of C_{57}. The above structures represent energetic minima, as shown by the simulated IR vibrations. All-together, these data reveal the proposed structures as pertinent candidates to the status of real molecules.

Table 1. Data on the structures in Figure 3 (HOMO-LUMO gap in eV, heat of formation Hf in kcal mol^{-1}, and total energy E_{tot} in a.u.) for the intermediate structures in building the C_{57} multi-cage, calculated at the PM3, HF/6-31G(d,p), B3LYP/6-31G(d,p) levels of theory; C_{60} is taken as reference structure.

Struct	Sym	PM3		HF		B3LYP	
		Gap	Hf	Gap	E_{tot}	Gap	E_{tot}
$C_{17}H_{12}$	T_d	10.53	131.66	12.99	-650.66	6.04	-654.92
$C_{57}H_{40}$	T_d	13.438	-49.11	14.270	-2181.99	7.365	-2170.67
C_{57}	D_{2d}	6.432	1497.92	7.574	-2156.98	1.888	-2196.27
C_{60}	I_h	6.596	810.82	7.420	-2271.83	2.761	-2286.17

DIAMOND D$_5$ NETWORK

A monomer C_{81}, derived from C_{57} and consisting of four closed C_{20} units and four open units, and its mirror image (Figure 4) is used to build the alternant network of the spongy diamond SD$_5$ (Figure 5).
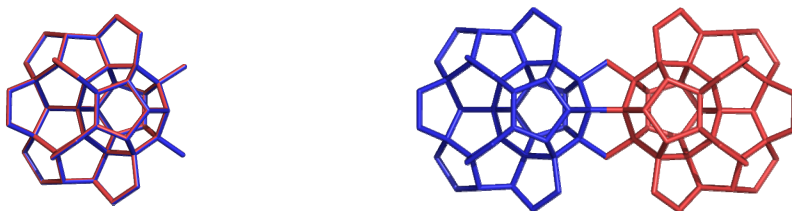
Figure 4. Monomer C_{81} unit (left-up), and its mirror image-pair (right-up); the repeating units of the spongy diamond SD$_5$ network.

The PM3 optimized structure SD$_5$(2,2,2)=$C_{600}H_{160}$ shows: heat of formation HF=4492.414 kcal/mol; heat of formation per heavy atoms HF/HA=7.487; Homo-Lumo-Gap=8.457 eV. In comparison, HF/HA (C_{60})=13.514 kcal/mol, thus justifying future studies (see also Table 1).
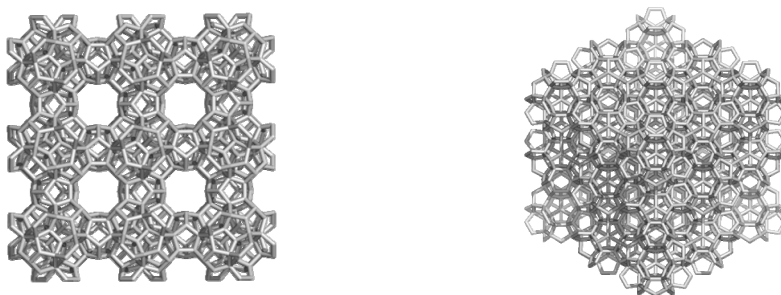
Figure 5. Spongy diamond SD$_5$ (3,3,3) network: top (left) and corner (right) views.

Formulas enabling the calculation of net parameter (vertices *v*, edges *e*, and the number of monomers *m* in a cubic domain), the ring polynomial $R(x)$, number of carbon atoms $C(sp^2)$ and $C(sp^3)$ and their ratios together with the limits of these quantities (when *k* trends to infinity) are presented in Table 2, function of *k* – the number of monomers along the edge of cubic domain. The network in Figure 5 is related to the P-type crystal proposed by Mackay for the Schwarzites [31].

Table 2. Topology of spongy diamond SD$_5$ (C$_{81}$) network

Formulas	
1	$v(\text{SD}_5) = 3k^2[27 + 23(k-1)] = 69k^3 + 12k^2$; $e(\text{SD}_5) = 130k^3$; $m(\text{SD}_5) = k^3$
2	$R(x) = ax^5 + bx^9 + cx^{10}$ $a = 6k^2[9 + 11(k-1)]$; $b = 6k^2[18 + 26(k-1)]$; $c = 6k^2[27 + 44(k-1)]$
3	$C(sp^3) = 53k^3 - 12k^2$
4	$C(sp^3\%) = (53k^3 - 12k^2) / (69k^3 + 12k^2)$ $\lim_{k\to\infty}(C(sp^3\%)) = 53/69$; 0.768116

Notice that there exist other diamond structures, either as real (Lonsdaleite, a rare stone of pure carbon discovered at Meteor Crater, Arizona, in 1967) or hypothetical [2,33] ones.

The structure C$_{17}$, has the skeleton of centrohexaquinane, and was synthesized (so far) as centrohexaindane [34], or C-trioxa-s-hexaquinane [35]

CONCLUSIONS

A new, yet hypothetically, carbon allotrope, called diamond D$_5$, was designed by using the structure C$_{81}$ as the repeating unit (for the spongy form SD$_5$). Diamond D$_5$ was theorized here for the first time in literature (even the origins of this completely new idea were presented in two previous articles [23,24]). The geometric and energetic arguments/proofs, summarily presented here, as well as the *dense* D$_5$ diamond network will be completed in further papers [36,37].

We expect the same mechanical, thermal, electrical, lubricating, catalyst support, biological, etc. properties as those found for the nano-diamond D$_6$.

ACKNOWLEDGMENTS

REFERENCES

1. M.V. Diudea, Ed., "*Nanostructures, novel architecture*", NOVA, **2005**.
2. M.V. Diudea and Cs.L. Nagy, "*Periodic Nanostructures*", Springer, **2007**.
3. L. Carlucci, G. Ciani and D. Proserpio, *Cryst. Eng. Comm*., **2003**, *5*, 269.
4. V.A. Blatov, L. Carlucci, G. Ciani and D. Proserpio, *Cryst. Eng. Comm*., **2004**, *6,* 377.
5. I.A. Baburin, V.A. Blatov, L. Carlucci, G. Ciani and D. Proserpio, *J. Solid State Chem*., **2005**, *178*, 2452.
6. O. Delgado-Friedrichs and M. O'Keeffe, *J. Solid State Chem*., **2005**, *178*, 2480.
7. V.A. Blatov, O. Delgado-Friedrichs, M. O'Keeffe, and D. Proserpio, *Acta Cryst.*, **2007**, *A63*, 418.
8. E. Ōsawa, *Diam. Rel. Mat*. **2007**, *16*, 2018.
9. E. Ōsawa*, Pure Appl. Chem*., **2008**, *80*, 1365.
10. V.N. Mochalin and Yu. Gogotsi, *J. Am. Chem. Soc.*, **2009**, *131*, 4594.
11. P.S. DeCarli, J.C. Jamieson, *Science*, **1961**, *133*, 1821.
12. A.E. Aleksenski, M.V. Baidakova, A.Ya. Vul', V.Yu. Davydov, Yu.A. Pevtsova, *Phys. Solid State*, **1997**, *39*, 1007.
13. A. Krüger, F. Kataoka, M. Ozawa, T. Fujino, Y. Suzuki, A.E. Aleksenskii, A.Ya. Vul', E. Ōsawa, *Carbon*, **2005**, *43*, 1722.
14. O.A. Williams, O. Douhéret, M. Daenen, K. Haenen, E. Ōsawa, M.M. Takahashi, *Chem. Phys. Lett*., **2007**, *445*, 255.
15. M.V. Diudea and G. Katona, in: Newkome, G.A. Ed., *Advan. Dendritic Macromol*. **1999**, *4*, 135.
16. G.R. Newcome, V.K. Gupta, G.R. Baker, Z.-Q.Yao, *J. Org. Chem.* **1985**, *50*, 2003.
17. D.A. Tomalia, *Aldrichimica Acta,* **1993**, *26*, 91.
18. M. Tang, C.T. Redeman, F.C. Szoka, Jr., *Bioconjugate Chem.,* **1996**, 7, 703.
19. B.F. Pan, D.X. Cui, P. Xu, T. Huang, Q. Li, R. He, and F. Gao, *J. Biomed. Pharmaceut. Eng*., **2007**, *1*, 13.
20. M.V. Diudea, *Nanomolecules and Nanostructures-Polynomials and Indices*, MCM, No. 10, Univ. Kragujevac, Serbia, **2010**.
21. G. Benedek, H. Vahedi-Tafreshi, E. Barborini, P. Piseri, P. Milani, C. Ducati, and J. Robertson, *Diamond Relat. Mater*., **2003**, *12*, 768.
22. E. Barborini, P. Piseri, P. Milani, G. Benedek, C. Ducati, and J. Robertson, *Appl. Phys. Lett*., **2002**, *81*, 3359.
23. M.V. Diudea, A. Ilić, *J. Comput. Theoret. Nanosci.,* **2011**, *8*, 000-000.
24. M.V. Diudea, Int. J. Chem. Model., **2010** (accepted).
25. M.V. Diudea, M. Ştefu, P.E. John, and A. Graovac, *Croat. Chem. Acta,* **2006**, *79*, 355.
26. M.V. Diudea, *J. Chem. Inf. Model.,* **2005**, *45,* 1002.
27. M.V. Diudea, *Forma* (Tokyo), **2004**, *19* (3), 131.
28. M. Stefu, M.V. Diudea and P.E. John, *Studia Univ. Babes-Bolyai Chemia*, **2005**, *50* (2), 165.

29. L. Euler, *Novi Comm. Acad. Scient. Imp. Petrop.*, **1758**, *4*, 109.
30. M.V. Diudea and C.L. Nagy, *MATCH Commun. Math. Comput. Chem.*, **2008**, *60*, 835.
31. A.L. Mackay and H. Terrones, *Nature*, **1991**, *352*, 762.
32. Zh. Chen, T. Heine, H. Jiao, A. Hirsch, W. Thiel, and P. von Rague Schleyer, *Chem. Eur. J.*, **2004**, *10*, 963.
33. T. Sunada, *Notices AMS*, **2008**, *55*, 208.
34. D. Kuck, *Pure Appl Chem.*, **2006**, *78*, 749.
35. L.A Paquette, M. Vazeux, *Tetrahedron Lett.*, **1981**, *22*, 291.
36. M.V. Diudea, Cs.L. Nagy, *Diam. Rel. Mat.*, **2010** (submitted).
37. M.V. Diudea, Cs.L. Nagy, *Carbon*, **2011** (submitted).

CORRELATING STUDY OF NEW MOLECULAR GRAPH DESCRIPTORS

MONICA L. POP[a], MIRCEA V. DIUDEA[a*]
AND ALEKSANDAR ILIĆ[b]

ABSTRACT. A correlating study, using as independent variables the topological indices newly developed within the design of the super-index Cluj-Niš *CJN*, is performed on the set of octane isomers. Among the modeled properties, the boiling point BP, entropy S and total surface area TSA, gave best scores, in mono- to three-variable regressions, with respect to our novel descriptors. The most important result was the monovariate description of octane boiling points, which is the best result so far published in literature.

Keywords: Octanes, topological indices, QSPR.

INTRODUCTION

Alkanes represent an interesting class of compounds as a starting point for the application of molecular modeling procedures. Many properties of alkanes vary function of molecular mass or branching, and alkanes can be described by using a single type of (carbon) atom. There are properties well accounted by a single molecular descriptor, e.g., octane number MON, entropy S, molar volume MV, molar refraction MR, etc. Other properties, such as boiling point BP, heat of vaporization HV, total surface area TSA, partition coefficient LogP, density DENS, critical temperature CT, critical pressure CP, and heat of formation DHF, are notable exceptions, being not well modeled by a single parameter [1,2].

The purpose of the present report is to evaluate the relative performances of new topological indices developed within the design of super-index Cluj-Niš *CJN* [3] in relating the hydrocarbon molecular structures to a set of physical properties.

LAYER MATRICES

Layer matrices have been proposed in connection with sequences of metrics: *DDS* (Distance Degree Sequence), *PDS* (Path Degree Sequence), and *WS* (Walk Sequence) [4-8]. They are built up on layer partitions in a graph.

---

[a] Faculty of Chemistry and Chemical Engineering, "Babes-Bolyai" University, 400028 Cluj-Napoca, Romania, mpop@chem.ubbcluj.ro; diudea@gmail.com
[b] Faculty of Sciences and Mathematics, University of Niš, Višegradska 33, 18000 Niš, Serbia, aleksandari@gmail.com

Let $G(v)_k$ be the kth layer of vertices v lying at distance k, in the partition $G(i)$:

$$G(v)_k = \{v \mid d_{i,v} = k\} \tag{1}$$

$$G(i) = \left\{ G(v)_k ; k \in [0,1,..,ecc_i] \right\} \tag{2}$$

with ecc_i being the *eccentricity* of i. The entries in the layer matrix (of a vertex property) **LM**, are defined as

$$\left[\mathbf{LM}\right]_{i,k} = \sum_{v\mid d_{i,v}=k} p_v \tag{3}$$

The layer matrix is a collection of the above defined entries:

$$\mathbf{LM}(G) = \left\{ [\mathbf{LM}]_{i,k} ; i \in V(G); k \in [0,1,..,d(G)] \right\} \tag{4}$$

with $d(G)$ standing for the diameter of the graph (*i.e.*, the largest distance in G

SHELL MATRICES

The entries in the shell matrix **ShM** (of a vertex pair property) are defined as follows [8]:

$$\left[\mathbf{ShM}\right]_{i,k} = \sum_{v\mid d_{i,v}=k} [\mathbf{M}]_{i,v} \tag{5}$$

The shell matrix is a collection of the above defined entries:

$$\mathbf{ShM}(G) = \left\{ [\mathbf{ShM}]_{i,k} ; i \in V(G); k \in [0,1,..,d(G)] \right\} \tag{6}$$

A shell matrix **ShM**(G) will partition the entries of a square matrix according to the vertex (distance) partitions in the graph. It represents a true decomposition of the property collected by the info square matrix according to the contributions brought by vertex pairs pertaining to shells located at distance k around each vertex.

The shell matrices were recently used as the basis in the calculation of super-index super index Cluj-Niš[3] *CJN* [3].

SHELL-DEGREE-DISTANCE POLYNOMIALS

The Cramer product of the diagonal matrix of vertex degrees **D** with the distance **DI** matrix provides the matrix of degree distances [9] denoted **DDI**.

$$\mathbf{D}(G) \times \mathbf{DI}(G) = \mathbf{DDI}(G) \tag{7}$$

The above Cramer product is equivalent (gives the same half sum of entries) with the pair-wise (Hadamard) product of the vectors "row sum" *RS* in the Adjacency **A** and Distance **DI** matrices, respectively.

$$RS(\mathbf{A}) \bullet RS(\mathbf{DI}) = RS(\mathbf{DDI}) \tag{8}$$

Next, by applying the Shell operator, we obtain the matrix **ShDDI**, of which column half sums are just the coefficients of the corresponding Shell-polynomial.

$$P(\mathbf{ShDDI}, x) = \sum_{k} p(G, k) \cdot x^k \tag{9}$$

An index, called *Cluj-Tehran CT*(**ShM**,*G*), with specified **M,** is defined as

$$CT(\mathbf{ShM}, G) = P'(\mathbf{ShM}, 1) + (1/2)P''(\mathbf{ShM}, 1) \tag{10}$$

where P' and P" refers to the polynomial firs and second derivative, respectively (here calculate dat x=1) [9].

ECCENTRIC CONNECTIVITY INDEX

Sharma *et al*. [10] introduced a distance-based molecular structure descriptor, the eccentric connectivity index, which is defined as:

$$\xi^c(G) = \sum_{v \in V} \deg(v) \cdot \mathrm{ecc}(v) \tag{11}$$

The eccentric connectivity index was successfully used for mathematical models of biological activities of diverse nature. The index ξ^c has been shown to give a high degree of predictability of pharmaceutical properties, and provide leads for the development of safe and potent anti-HIV compounds. The investigation of its mathematical properties started only recently, and has so far resulted in determining the extremal values and the extremal graphs, and also in a number of explicit formulas for the eccentric connectivity index of various products of graphs, several families of benzenoid graphs, zigzag and armchair hexagonal belts, nanotubes and nanotori [11].

MODELING OCTANE PROPERTIES

To test the correlating ability of the descriptors derived from the degree-distance matrices (**ShDDI**$_k$) and Shell-polynomials, we focused attention to the set of octanes, as one of the benchmark-sets [12,13] in correlating studies by using topological indices. Among several properties tested, three ones (listed in Table 1) provided best scores: boiling point BP, entropy S and total surface area TSA, in mono- to three- variable regressions, by our novel descriptors.
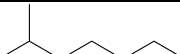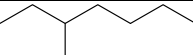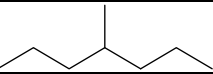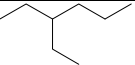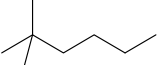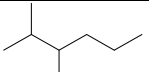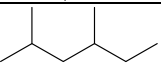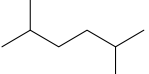
The octane topological descriptors are given in Table 2. They include the numbers derived from the ShM$_1$ (the basic degree-distance matrix, in the Shell-matrix format), the first and second derivatives, D$_1$ and DD$_1$, of the corresponding Shell-polynomial, al level *k*=1 (equivalent to A$_1$) and the same descriptors at *k*=2 (equivalent to A$_2$ – the exponent referring to the remote adjacency rank) and also Cluj-Tehran CT$_1$ index and the eccentric connectivity index, denoted ECC in Table 2. Statistics are presented in Table 3.
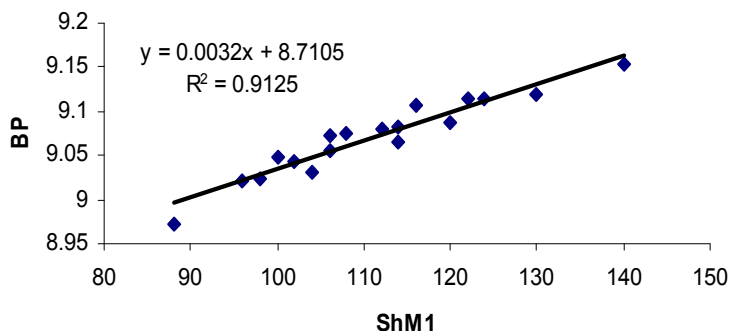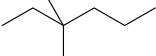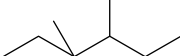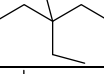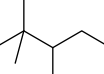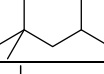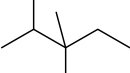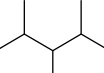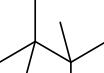
Figure 1. The plot of BP vs ShM_1 index

The best monovariate model of BP (by the index ShM_1) is plotted in Figure 1. Compare our results with the best results reported so far in literature (Table 3, entry 7) and remark the best result (R^2= 0.913) for the monovariate model of the octane boiling points. The other results in Table 3 (mono- or tri- variable models) can be considered as good (or acceptable) results.

Table 1. Octanes boiling point BP, entropy S and total surface area TSA values

	Molecule	BP	S	TSA
1		9.153	111.67	415.3
2		9.120	109.84	407.85
3		9.115	111.26	397.34
4		9.114	109.32	396.04
5		9.108	109.43	379.04
6		9.065	103.42	405.11
7		9.079	108.02	384.93
8		9.082	106.98	388.11
9		9.088	105.72	395.08

Molecule	BP	S	TSA
10	9.056	104.74	389.79
11	9.074	106.59	376.91
12	9.073	106.06	368.1
13	9.049	101.48	366.99
14	9.023	101.31	371.75
15	9.031	104.09	392.19
16	9.020	102.06	377.4
17	9.044	102.39	368.93
18	8.971	93.06	390.47

Table 2. Topological Indices of Octanes

Molecule	ShM_1	D_1	DD_1	CT_1	ShM_2	D_2	DD_2	ECC
1	140	532	1876	1470	258	972	3402	74
2	130	452	1406	1155	263	921	2900	65
3	124	404	1148	978	248	804	2270	63
4	122	388	1066	921	243	765	2070	61
5	116	340	808	744	230	668	1572	54
6	114	340	854	767	269	815	2088	56
7	112	320	746	693	244	704	1658	54
8	114	332	790	727	250	740	1796	54
9	120	376	998	875	266	856	2356	56
10	106	284	602	585	247	663	1390	52
11	108	292	620	602	234	636	1350	52
12	106	276	538	545	230	606	1202	43
13	100	244	432	460	232	568	996	41

23

Molecule	ShM_1	D_1	DD_1	CT_1	ShM_2	D_2	DD_2	ECC
14	98	236	414	443	250	624	1138	43
15	104	272	546	545	268	732	1552	45
16	96	224	370	409	244	588	1000	41
17	102	256	476	494	243	633	1236	43
18	88	184	246	307	262	592	876	34

Table 3. Statistics (R^2) of QSPR Study on Octanes

	Descriptors	BP	S	TSA
1	ShM_1	0.913	0.771	0.520
2	D_1	0.863		
3	CT_1	0.818	0.646	0.613
4	ECC	0.854	0.747	0.583
5	ShM2&ECC	0.938	0.873	0.902
6	ShM_1&D_1&DD_1	0.987	0.933	0.914
7	ShM_2&D_2&DD_2	0.991	0.924	0.870
8	Best in Octanes (monovariable)	(0.78)[13] (0.77)[14]	(0.92)[13] (0.93)[14] (0.950)[1]	(0.72)[13] (0.92)[14]

CONCLUSIONS

A correlating study, using as independent variables the topological indices newly developed within the design of the super-index Cluj-Niš *CJN*, was performed on the set of octane isomers. The modeled properties were: boiling point BP, entropy S and total surface area TSA, which gave best scores, among several properties, in mono- to three-variable regressions, with respect to our novel descriptors. The most important result was the monovariate description of octane boiling points, which is the best result so far published in literature.

ACKNOWLEDGMENTS

REFERENCES

1. D. Stevanović, A. Ilić, C. Onisor, and M.V. Diudea, *Acta Chim, Slov*., **2009**, *56*, 410.

2. M.V. Diudea (Ed.), *QSPR/QSAR Studies by Molecular Descriptors*, **NOVA**, New York, 2001.

3. M.V. Diudea, A. Ilić, K. Varmuza, M. Dehmer, *Complexity*, **2010** (in press).

4. M.V. Diudea, I. Gutman; L. Jäntschi, *Molecular Topology*, **NOVA**, New York, 2002.

5. M.V. Diudea, M.S. Florescu, and P.V. Khadikar, *Molecular Topology and Its Applications*, **EFICON**, Bucharest, 2006.

6. M.V. Diudea, *J. Chem. Inf. Comput. Sci*., **1994**, *34*, 1064.

7. M.V. Diudea, M.I. Topan, and A. Graovac, *J. Chem. Inf. Comput. Sci*., **1994**, *34*, 1072.

8. M.V. Diudea and O. Ursu, *Indian J. Chem*., **2003**, *42A*, 1283.

9. M.V. Diudea, Nanomolecules and Nanostructures - Polynomials and Indices, MCM, No. 10, Univ. Kragujevac, Serbia, 2010.

10. V. Sharma, R. Goswami, and A.K. Madan, *J. Chem. Inf. Comput. Sci.,* **1997**, *37*, 273.

11. A. Ilić, in: I. Gutman, B. Furtula, Novel Molecular Structure Descriptors – Theory and Applications II, MCM Vol. 9, University of Kragujevac, 2010.

12. Milano Chemometrics & QSAR Research Group, Molecular Descriptors Dataset (available at http://www.moleculardescriptors.eu/dataset/dataset.htm)

13. http://www.moleculardescriptors.eu/dataset/dataset.htm; **http://www.iamc-online.org/**.

14. D. Vukičević, *Bond Additive Modeling. Adriatic Indices – Overview of the Results,* in I. Gutman and B. Furtula (Eds.), Novel Molecular Structure Descriptors - Theory and Applications II, MCM, Kragujevac, 2010, p. 269.

EVALUATION OF THE ANTIOXIDANT CAPACITY OF A SERIES OF ACYL-HYDRAZONES BEARING 2-ARYL-THIAZOLE

**CRISTINA MARIANA MOLDOVAN[a], ALINA PÂRVU[b],
BRÎNDUŞA TIPERCIUC[a], OVIDIU ONIGA[a]**

ABSTRACT. The antioxidant properties of a series of acyl-hydrazones bearing 2-aryl-thiazole are explored. Some specific parameters were measured: total oxidant status (TOS), total antioxidant response (TAR) and oxidative stress index (OSI). The study indicates antioxidant properties for compounds 4-6, 10, 12, while acyl-hydrazones 1 and 2 developed a prooxidant effect.

Keywords: acyl-hydrazone, 2-aryl-thiazole, antioxidant activity

INTRODUCTION

Reactive oxygen species (ROS) are produced in metabolic and physiological processes, and harmful oxidative reactions may occur in organisms, which are removed via enzymatic and non-enzymatic antioxidative mechanisms. Under some conditions, the increase in oxidants and decrease in antioxidants cannot be prevented, and the oxidative/ antioxidative balance shifts toward the oxidative status. Consequently, oxidative stress, involved in over 100 disorders, occurs [1-3].

ROS are implicated in the pathophysiology of ageing and oxidative stress associated pathologies such as diabetes, neurodegenerative diseases, atherosclerosis and cardiovascular complications [4-5].

ROS are normally produced throughout oxygen metabolism and play a major role in physiological and pathological cell redox signalling. Oxidative stress appears in the context of non-equilibrium of overproduction of ROS of various cellular sources (the mitochondrial respiratory chain, nicotinamide adenine dinucleotide phosphate hydride oxidases (NADPHOXs or NOXs), xanthine oxidase, lipoxygenases, cytochromes P450, and other oxidases) and decreased cellular and plasma antioxidant defenses [6].

Hydroxyl group (OH) and its subsequent radicals are the most harmful ROS and they are mainly responsible for the oxidative injury of biomolecules. Alone hydrogen peroxide and superoxide molecules cannot directly oxidize

---
[a] *Iuliu Haţieganu University of Medicine and Pharmacy, Faculty of Pharmacy, Department of Pharmaceutical Chemistry, 41 Victor Babes Street, RO-400010 Cluj-Napoca, Romania, cris.moldovan@yahoo.com, cmoldovan@umfcluj.ro*
[b] *Iuliu Haţieganu University of Medicine and Pharmacy, Faculty of Medicine, Department of Physiology,Victor Babeş Street, Cluj-Napoca, Romania*

lipids, nucleic acids and sugars. These species can lead to oxidative injury in biomolecules indirectly by producing ·OH via Fenton reaction and/or iron-catalyzed Haber–Weiss reaction [7]. Oxidized molecules generally form new radicals leading to radical chain reactions or they are neutralized by antioxidants.

Antioxidant molecules prevent and/or inhibit these harmful reactions remaining very efficient in preventing the early atherosclerotic lesions and inflammatory events implicated in the evolution of the lesions toward [8].

Hydrazones display diverse biological and pharmaceutical activities, such as antimicrobial, antitumoral, antiinflammatory, antioxidant properties [9, 10]. The antioxidant activity may be due to their capacity of metal chelating. Under certain abnormal conditions, activated oxygen species release iron from the transport and storage proteins, and the resulting "free" iron (Fe^{2+}) promotes the formation of the devastatingly reactive toxic ·OH. Thus, chelating the iron, hydrazones may inhibit free-radical formation and the consequent free radical tissue damage [11].

Chromone derivatives possess a wide spectrum of biological activities, such as anti-inflammatory, antifungal, antimicrobial, antiviral, antitumour, mainly due to their well-recognised antioxidant properties, which stem from their ability to neutralise active forms of oxygen and to cut off free radical processes. This potential health benefit is ruled by strict structure-activity/ structure-property relationships, which, apart from determining their biological action, modulate their systemic distribution and bioavailability in sites of oxidation within the cell [12, 13].

Prompted by these reports, we tested the antioxidant capacity of 14 acyl-hydrazones bearing 2-aryl-thiazole scaffold. Some of these hydrazones have a chromone moiety in their structures, too.

RESULTS AND DISCUSSION

Our study investigated the effects of the acyl-hydrazones (Scheme 1) in an acute experimental inflammation, because of the close relationship between this process and ROS as endogenous mediators. The inflammation was induced by the i.m. injection of turpentine oil [14, 15]. The antioxidant effect of the tested compounds was assessed by evaluating some specific parameters: total antioxidant response (TAR), total oxidant status (TOS), and the index of oxidative stress (OSI) [OSI=(TOS/TAR)x100].

Determination of TAR in animal serum is done by a method that allows the simultaneous measurement of more molecules with an antioxidant potential, against the oxidants from serum. The method is based on the suppression of the obtention of dianisidil radicals from the oxidative process of orto-dianisidine, radicals colored in brown-yellow, by the antioxidant substances present in serum [16]. Therefore, a standardized solution of Fe^{+2}-o-dianisidine complex suffers a Fenton reaction with a standardized solution of H_2O_2, forming ·OH radicals. These radicals, in the presence of an acid, oxidize o-dianisidine to dianisidil radicals, which determine further oxidation reactions. The antioxidant

agents from the sample inhibit the oxidation reactions and determines the apparition of color. The method for evaluating the total antioxidant response is a colorimetric technique. The intensity of the color at the end of reactions is spectrophotometrically determined.

The method used to determine the total oxidant status is based on the oxidation of Fe^{+2}-o-dianisidine complex to Fe^{+3}, which forms a colored complex with xilenol-orange [17]. The intensity of this complex is colorimetrically determined and is in a direct relationship with the total quantity of oxidant molecules present in the sample.

The calculation of the index of oxidative stress is very useful for investigating and comparing the oxidant-antioxidant status of the tested compounds. Big values of OSI indicate oxidant properties, while small values suggest good antioxidant capacity.
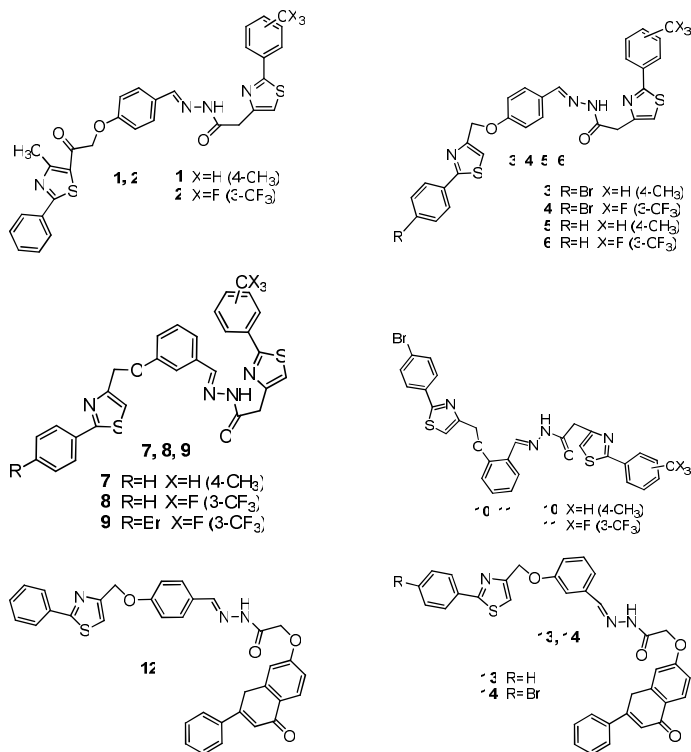
Scheme 1

The study was performed on adult male Wistar-Bratislava albino rats, divided in groups, which received food and water *ad libitum*. The effects of the compounds were compared with those from the inflammation group (I), and with those from the group treated with Meloxicam (M), as a reference NSAID with an antioxidant activity. A negative control group (C) of healthy rats without any

treatment was also used. All the tests were performed in triplicate and the average was taken as final reading.

The total antioxidant response, as a measure of each organism to protect himself against the oxidant agents, by releasing different physiological antioxidant substances and/or using different exogenous ones, showed a higher value than that of the total oxidant status for the control group. This resulted in a small TOS/TAR ratio (Tabel 1). On the other hand, for the inflammation group, OSI had a bigger value, as expected, because of the important increase of TOS, doubled by the decrease of TAR.

Table 1. Effects of the compounds on the oxidative stress

Compound	TOS	TAR	OSI (TOS/TAR)x100
C	1.52±0.61	2.332±0.033	0.06518
I	30.25±2.2	1.0922±0.0029	2.769639
M	19.99±2.5	1.1017±0.0079	1.814469
1	33.54±2.97	1.0969±0.0026	3.057708
2	31.04±3.78	1.097±0.004 *	2.829535
3	27.27±3.52 *	1.1018±0.0029 **	2.475041
4	25.92±2.64 *	1.104±0.0065 **	2.347826
5	26.66±2.83 *	1.0989±0.0014 **	2.426062
6	22.57±3.44 **	1.0952±0.001 *	2.060811
7	28.04±3.3	1.0983±0.0007 **	2.553037
8	10.98±1.36 **, ***	1.0882±0.002	1.009006
9	35.77±3.3	1.0963±0.006	3.262793
10	27.62±2.03 *	1.0978±0.0022 **	2.515941
11	40.18±1.83	1.103±0.003	3.642792
12	20.35±0.69 **	1.0992±0.0032	1.851346
13	9.82±1.1 **, ***	1.0855±0.0047	0.904652
14	45.9±2.46	1.0979±0.0023	4.180709

* $p < 0.05$, **$p < 0.001$ (comparing with inflammation group)
*** $p < 0.001$ (comparing with Meloxicam group)

The results registered for the tested acyl-hydrazones showed a significant decrease of TOS for the compounds **4**-**6**, **10**, **12** and also an increase of TAR, compared to the inflammation group ($p < 0.05$) (Figures 1 and 2). These values resulted in an important reduction of OSI compared to inflammation (Figure 3). The results obtained for these compounds reflect their antioxidant properties.

For compounds **8** and **13**, the reduction of TOS/TAR ratio was ascribed to the more pronounced decrease of TOS. As for compound **7**, the reduction was assigned to the more pronounced increase of TAR. These three hydrazone derivatives bearing 2-aryl-thiazole may be considered for their antioxidant capacity, too.
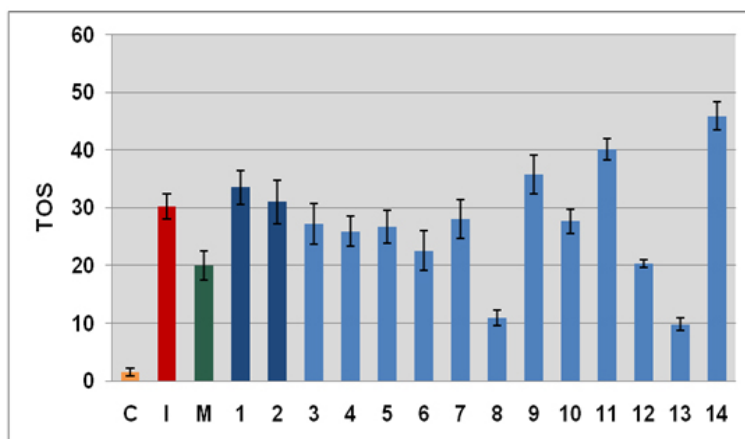
Figure 1. Influence of the compounds on the total oxidant status (TOS)
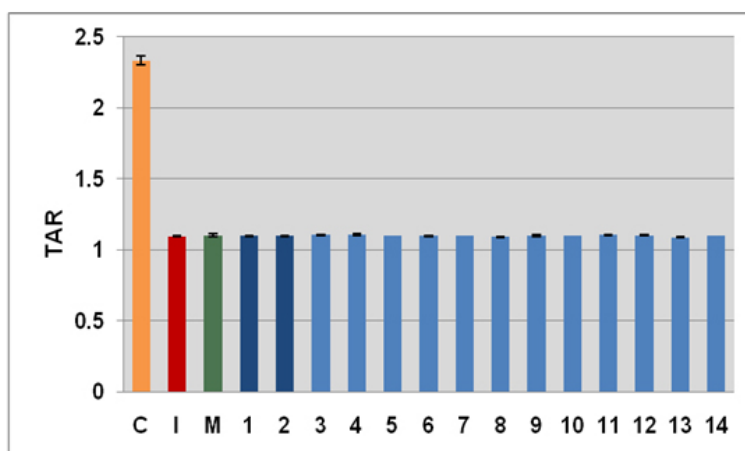
Figure 2. Influence of the compounds on the total antioxidant response (TAR)

Compounds **8** and **13** reduced TOS more powerful (p<<0.001) than meloxicam, the reference drug. This reduction of TOS determined the reduction of OSI, too. Values for TAR were superior to those from the group inflammation, for all the tested derivatives, excepting compounds **8** and **13**. For other compounds (**1**, **2**, **9, 11, 14**), both TOS and TAR increased, but in the case of TOS, the raise was more significant. This led to values of OSI bigger than for inflammation. Therefore, we could suspect a prooxidant effect of these compounds.
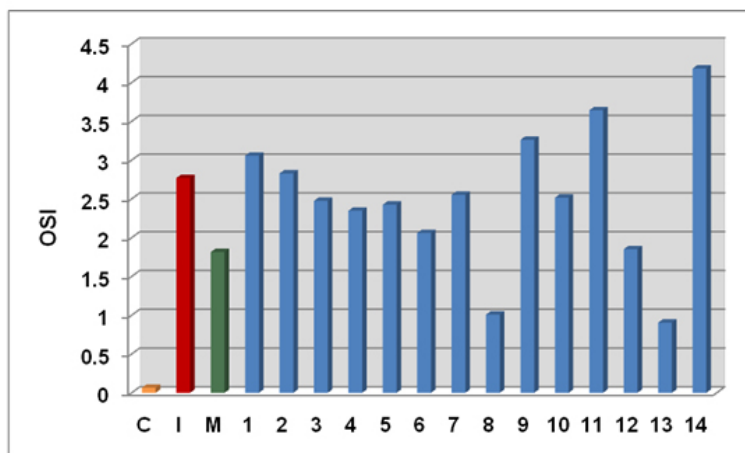
Figure 3. Variation of the index of oxidative stress (OSI) for the tested compounds

Analyzing the structural profile of the tested acyl-hydrazones, it could be observed that N'-(4-(2-(2-phenyl-4-methyl-thiazole-5-yl)-2-oxoethoxi)-benzyliden-aryl-hydrazides **1** and **2** presented a prooxidant activity. The substitution in para of the benzylidene fragment from position 2 of thiazole with brome resulted in the reduction of the antioxidant potential of the N'-(2), (3), 4-((2-phenyl-thiazole-4-yl)-methoxi)-benzyliden-hidrazides **3**-**14**, compared with the unsubstituted compounds. Two compounds of this series, **8** and **13**, demonstrated a more potent antioxidant capacity than meloxicam, the reference drug. For these two derivatives, a significant decrease of OSI was registered.

CONCLUSIONS

The study on the effect of acyl-hydrazones on the oxidative stress indicated antioxidant properties for compounds **4**-**6**, **10**, **12**. These derivatives determined a significant reduction of the total oxidant status and an increase of the total antioxidant response. For compounds **8** and **13**, the decrease of the index of oxidative stress resulted from the more pronounced reduction of TOS and more pronounced increase of TAR, in the case of compound **7**. On the other hand, acyl-hydrazones **1** and **2** developed a prooxidant effect.

SAR study showed that the substitution of phenyl from position 2 of thiazole with brome, in position 4, led to a reduction of the antioxidant capacity.

EXPERIMENTAL SECTION

The experiments were performed on adult male Wistar-Bratislava albino rats, weighing 200–250g. The animals were obtained from the Biobase of University of Medicine and Pharmacy Cluj-Napoca and housed at 25 ± 2 C^{o},

50 ± 5% relative humidity and 12 h light/dark cycle. They were distributed in groups of ten and had free access to water and food. All the experimental procedures and protocols used in this study were reviewed and approved by the Institutional Animal Ethical Committee (IAEC) of University of Medicine and Pharmacy Cluj-Napoca. Experiments were performed in triplicate and the average was taken as final reading.

For the group called Inflammation, each animal was injected i.m. with 0.6mL/100g (body weight) of turpentine oil, the pro-inflammatory substance. The same procedure and dose were used for the other groups, too. After that, a 3.2mg/kg dose, equivalent to 0.0091168mmol/kg of Meloxicam, the reference standard drug, was administered i.p. to the animals from the reference group. The test groups received the synthesized compounds in an equi-molar dose with Meloxicam, by the i.p. administration of its 1% carboxymethylcelullose suspension.

Determination of **TAR**: Serum (20 µl) is mixed with Fe^{+2}-o-dianisidine complex R_1 (obtained from the dissolution of $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ and of 3.17 g o-dianisidine in a KCl solution) (800 µl) and solution of R_2 H_2O_2 7.5 mM (40 µl). The intensity of the color obtained after 3-4 minutes after the mixing is spectrofotometrically determined at λ=444 nm. The blank is represented by 860 µl of R_1. Calibration is done using serial dillutions of a 1mM/l Trolox solution (pH 7.4). The results are expressed in mmol equivalent Trolox/l.

Determination of **TOS**: Serum (140 µl) is mixed with xilenol-orange solution R_1 (obtained from the dissolution of 0.114 g xilenol-orange and of 8.18 g NaCl in 900 ml of H_2SO_4 25 mM) (900 µl) and a solution of $Fe(NH_4)_2(SO_4)_2 6H_2O$ R_2 (obtained from the dissolution of 1.96 g of $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ and of 3.17 g of o-dianisidine chlorhydrate in 1000 ml H_2SO_4 25 mM) (44 µl). The intensity of the color obtained after 3-4 minutes after the mixing is spectrofotometrically determined at λ=560 nm. The blank is represented by 900 µl R_1 and 184 µl distilled water. Calibration is done using serial dillutions of a 200 µmol/l H_2O_2 solution. The results are expressed in µmol H_2O_2 equivalent /l.

Determination of **OSI**: OSI = (TOS/TAR) x 100.

The values are expressed as mean±S.D. for Inflammation group, Meloxicam group and the healthy population, separately. The comparisons of parameters were performed with Student's t-test. A p-value < 0.05 was accepted as significant. Data were analyzed using the SPSS for Windows computing program (Version 11.0).

REFERENCES

1. M. Harma, O. Erel. Oxidative stress in women with preeclampsia, *Am J Obstet Gynecol,* **2005**,*192(2)*, 656.
2. E. Yeni, M. Gulum, S. Selek, et al, *Int J Impot Res*, **2005**, *17(1)*,19.
3. M. Yanik, O. Erel, M. Kati, *Acta Neuropsychiatr*, **2004**, *16(4),* 200.
4. E.R. Stadtman, *Ann. N.Y. Acad. Sci.*, **2001**, *928*, 22.
5. A. Negre-Salvayre, C. Coatrieux, C. Ingueneau, R. Salvayre, *Br. J. Pharmacol.*, **2008**, *153*, 6.
6. H. Sies, *Am. J. Med.*, **1991**, *91*, 31S.
7. C.E. Thomas, L.A. Morehouse, S.D. Aust, *J Biol Chem*, **1995**, *260(6)*, 3275.
8. I.S. Young, J.V. Woodside, *J Clin Pathol*, **2001**, *54*, 176.
9. Y. Li, Z.Y. Yang, M.F. Wang, *Eur J Med Chem*, **2009**, *44*, 4585.
10. N. Belkheiri, B. Bouguerne, F. Bedos-Belval, H. Duran, C. Bernis, R. Salvayre, A. Nègre-Salvayre, M. Baltas, *Eur J Med Chem*, **2010**, *45*, 3019.
11. M. Horackova, P. Ponka, Z. Byczko, *Cardiovascular Research*, **2000**, *47 (3)*, 529.
12. N.F.L. Machado, M.P.M. Marques, *Current Bioactive Compounds*, **2010**, *6 (2)*, 76.
13. A. Gomes, O. Neuwirth, M. Freitas, D. Couto, D. Ribeiro, A.G. Figueiredo, et al, *Bioorg Med Chem*, **2009**, *17(20)*, 7218.
14. L. Pleşca-Manea, A.E. Pârvu, M. Pârvu, M. Tămaş, R. Buia, M. Puia, *Phytother Res*, **2002**, *16(4)*, 316.
15. B. Tiperciuc, A. Pârvu, M. Palage, O. Oniga, D. Ghiran, *Farmacia*, **1999**, *5*, 77.
16. O. Erel, *Clinical Biochemistry*, **2004**, *37*, 112.
17. O. Erel, *Clinical Biochemistry*, **2005**, *38*, 1103.

TO WHAT EXTENT THE NMR "MOBILE PROTONS" ARE RELEVANT FOR RESTRICTED ROTATIONAL STEREOCHEMISTRY PHENOMENA?

A CASE IN AMINO-*s*-TRIAZINE SERIES

OANA MOLDOVAN[a,b], PEDRO LAMEIRAS[c], ERIC HENON[c], FLAVIA POPA[a,b], AGATHE MARTINEZ[c], DOMINIQUE HARAKAT[c], CARMEN BÂTIU[a], YVAN RAMONDENC[b], MIRCEA DARABANTU[a*]

ABSTRACT. The use of the so-called NMR "mobile protons" in investigation of restricted rotational phenomena about partial double bonds, *i.e.* C^{sp2}(*s*-triazine)-N(exocyclic), is examined in the case of twelve highly elaborated amino-*s*-triazines.

Keywords: amino-s-triazines, NMR, restricted rotation, serinols

INTRODUCTION

The ^{1}H NMR assignment of the so-called "mobile (exchangeable, labile) protons" *XH* (X = N, O, S, etc.) is usually achieved by taking into account the crucial influence of the solvent (hydrogen bond donor or acceptor), heteroatom (X), temperature and molecular environment [1a]. Hydrogen bond acceptor solvents, *e.g.* [D6]DMSO, allow, by their chelatizing aptitude, detection of vicinal couplings $^3J_{H,H}$ in A_nX systems (n = 1, 2) of type >C*H*-N*H*- and >C*H*-O*H*. In contrast, in hydrogen bond acceptor solvents, *e.g.* CDCl₃, these "mobile protons" are observed much upfield and their broad shaped signals are somehow "classical", for example in the case of NH groups, due also to the quadrupolar moment of the isotope ^{14}N (I = 1) [1].

Higher temperatures increase the *XH* intra- or intermolecular mobility. For complex molecular environments, the correct significance of the "mobile protons" ^{1}H NMR location is still a challenging task [1b, 1c].

In the above context, the aim of this study is to present the synthesis and rotational stereochemistry about the C^{sp2}-N partial double bonds in some elaborated amino-*s*-triazines possessing a plethora of "mobile protons" together with their versatile role in evaluation of this dynamic behaviour.

[a] *"Babes-Bolyai" University, Department of Organic Chemistry, 11 Arany Jànos st., 400028 Cluj-Napoca, Romania darab@chem.ubbcluj*
[b] *University and INSA of Rouen, IRCOF – LCOFH, UMR 6014 CNRS COBRA, 76821 Mont Saint-Aignan Cedex, France*
[c] *University of Reims Champagne-Ardenne, ICMR - LIS, UMR 6229, BP 1039, 51687 Reims, France*

RESULTS AND DISCUSSION

Synthesis

The chemistry of the performed reactions is resumed in *Scheme 1*. Quantitative data are listed in *Table 1*.
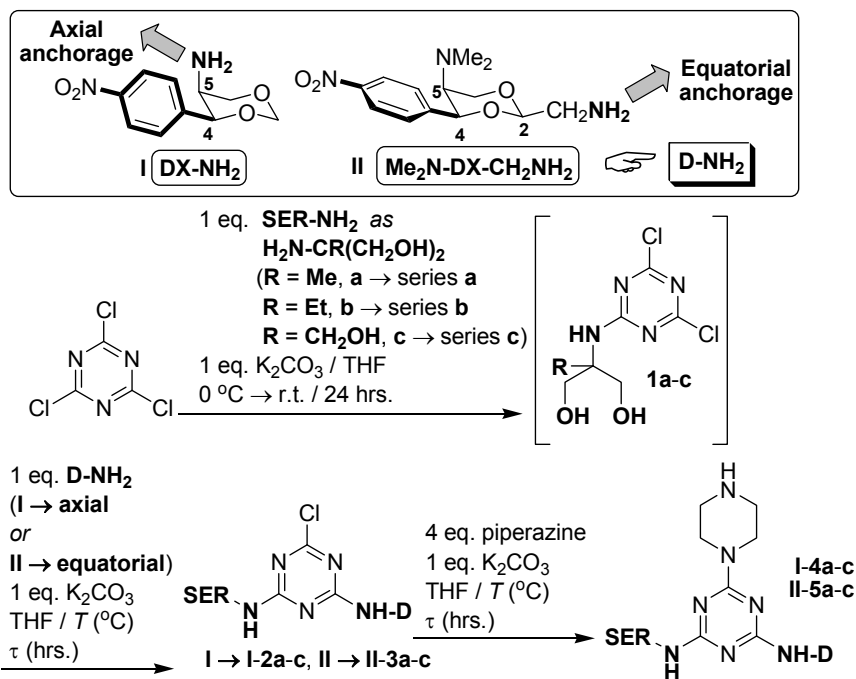
Scheme 1

Except our earlier results [2] on this topic, no similar approach was reported so far. All compounds are new ones.

Thus, by successive three highly selective aminations of cyanuric chloride with amino-nucleophiles of type *C*-substituted 2-aminopropane-1,3-diol (**SER-NH₂**, "*serinol*", **a**-**c**), we accessed the *N*-unsymmetrically substituted triamino-*s*-triazines ("*melamines*") **I**-**4a**-**c** and **II**-**5a**-**c**. They can be seen as novel building-blocks for further iterative synthesis. This account was performed in the presence of piperazine, a widely recognised dendritic linker [3].

The use of tandem two-type serinolic amino-nucleophiles (**a**-**b** and **I**, **II**) needs the comments below:

i) C-2-substituted serinols (**SER-NH₂**, **a** "*Methylserinol*", **b** "*Ethylserinol*" and **c** TRIS), were designed to play the role of an "*open-chain*" *N*-ligand in the target melamines. The first step-amination, carried out with **a-c,** occurred quantitatively in a very clean but slow evolution, due most likely to solvation effects diminishing their nucleophilicity.

36

Table 1. Reaction conditions and quantitative results in the synthesis
of compounds I-**2a**-**c**, II-**3a**-**c**, I-**4a**-**c** and II-**5a**-**c** (*Scheme 1*)

No.[a]	SER-NH (R)	D-NH (I *or* II)	T (°C) / τ (hrs.)	Yield (%)[b]
I-**2a**	Me	DX-NH (I)	reflux / 16	80
I-**2b**	Et		reflux / 22	66
I-**2c**	CH$_2$OH		reflux / 12	84
II-**3a**	Me	Me$_2$N-DX-CH$_2$NH (II)	-10 → r.t. / 24 reflux / 14	83
II-**3b**	Et		-10 → r.t. / 24 reflux / 16	42
II-**3c**	CH$_2$OH		-10 → r.t. / 24 reflux / 12	95
I-**4a**	Me	DX-NH (I)	r.t. /10[c]	80
I-**4b**	Et		r.t. / 24	84
I-**4c**	CH$_2$OH			86
II-**5a**	Me	Me$_2$N-DX-CH$_2$NH (II)		71
II-**5b**	Et			67
II-**5c**	CH$_2$OH			81

[a] The synthesis and stereochemistry of intermediates **1a**-**c** we reported elsewhere [2a];
[b] In case of compounds I-**2a**-**c** and II-**3a**-**c** as isolated global yields, after two-steps synthesis; in case of melamines I-**4a**-**c** and II-**5a**-**c** as isolated yields, after one-step synthesis.
[c] Time required by the slow addition, portionwise, of chlorodiamino-*s*-triazine to a four fold molar excess of piperazine (see **EXPERIMENTAL SECTION**).

        ii) *Mutatis-mutandis*, enantiopure amino-1,3-dioxanes **D-NH$_2$**, **I** and **II**, should be seen as "*closed-chain*" *N*-ligands. They were readily available by our "sulphuric acetalisation" methodology [2b,2d] applied to the corresponding (1*S*,2*S*)-2-amino-1-(4-nitrophenyl)propane-1,3-diols ("*Threo-p-nitrophenylserinols*"). **I** and **II** were used in the second step-amination. In the case of nucleophile **II**, precautions against side reactions, *i.e.* amination by *N*-demethylation, required milder conditions [2b, 4].
        iii) The selective attachment of the third nucleophile was accomplished based on our previously reported procedure [2a], consisting of the portionwise addition of chlorodiamino-*s*-triazines I-**2a**-**c** and II-**3a**-**c** (10 hrs. at room temperature) to a four fold molar amount of piperazine. Melamines I-**4a**-**c** and II-**5a**-**c** were purified by column chromatography on partially deactivated silica gel.

Rotational stereochemistry phenomena

Brief overview of our problem
        Starting from the well-known herbicide ATRAZINE®, 2-chloro-4-ethylamino-6-isopropylamino-*s*-triazine structural assignment [5], the elucidation of rotational diastereomers of *N*-substituted amino-*s*-triazines, in solution, is a quite difficult task [6]. They are issued from the lpN(exocyclic) → π (deficient *s*-triazine) delocalisation determining an increased order of bonds C(*s*-triazine)-

N(exocyclic). In addition to NMR and 2D-^{1}H,^{1}H NMR techniques [6], in investigating this phenomenon, computational methods [6a, 6g] including DFT approaches [2a, 2b] are also of interest. Usually, these studies focused on (un)symmetrically *N*-substituted melamines [6] by means of (VT) NMR at low temperature. In contrast, apart from ATRAZINE®, minor attention was paid to *N,N'*-substituted-2-chloro-4,6-diamino-*s*-triazines concerning their dynamic behaviour [5b, 6a-c], limited to *symmetric* 2-chloro-4,6-bis(*N,N*-dialkylamino) derivatives only.

Some introductory structural observations on our *unsymmetrically N*-substituted amino-*s*-triazines are mandatory.

The serinolic "*open-chain*" site, containing a variable number of geminal hydroxymethyl groups, is *a priori* seen as the most solvated region of the molecule while the amino-1,3-dioxanic "*closed-chain*" units, **I** and **II**, are anancomeric structures due to the overwhelmingly one-sided conformational equilibria, by the adoption of an equatorial position by the C-4'-*p*-nitrophenyl ring (*i.e.*, anancomerising group) [7]. However, their amino-anchorage to the *s*-triazine is essentially different, either *axial* (in **I**-**2a**-**c** and **I**-**4a**-**c**) or *equatorial* (in **II**-**3a**-**c** and **II**-**5a**-**c**).

As expected, the increased bond order of bonds C(*s*-triazine)-N(exocyclic) in our compounds creating restricted rotation, determined their NMR rather complicate appearance, at room temperature (*e.g.* a "sugar like" aspect). Indeed, depending on the π-deficiency of the *s*-triazine ring, higher in chlorodiamino-*s*-triazines than in melamines, and in an idealised topological model, four stereoisomers are possible (*Scheme 2*). Each of them can be generated by a single rotation / (frozen) equilibrium (a step-by-step interconversion).

Next, since the intimate rotational status of these four species was very different in the above two series, **I**-**2** and **II**-**3** *vs.* **I**-**4** and **II**-**5**, they will be discussed separately, in the decreasing order of *s*-triazine π-deficiency.

Rotational stereochemistry phenomena in chlorodiamino-s-triazines

As predicted, at *room temperature*, chlorodiamino-*s*-triazines **I**-**2a**-**c** and **II**-**3a**-**c** consisted of mixtures of four frozen rotamers (*Scheme 2*). Their abundance could be evaluated by means of ^{1}H NMR resonance of protons D-N*H* (α, β, γ) and SER-N*H* (α', β', γ'), the best separated, hence the single ones indicative for rotational behaviour. [D$_6$]DMSO was the only appropriate NMR solvent in all investigations (*Table 2*, *Table 3*).

In series **I**-**2**, the individual assignment of rotamers starts from the 2D-^{1}H,^{1}H-NOESY chart of compound **I**-**2a** (*Figure 1*, *Scheme 2*) disclosing dipolar interactions between the proton SER-N*H* (6.46 ppm, signal γ') and the *p*-nitrophenyl ring of the axially anchored D-NH moiety in rotamer **I**-**2a** (4-*a*-6-*s*), hence a "*trans*" relationship between the *N,N'*-ligands (*Table 2*).

Axial anchorage Content of blocked rotamers (%)[c]
of D *N*-ligand in

	a-a	4-s-6-a	4-a-6-s	s-s
I-2a	62.0	19.5	16.5	2.0
I-2b	56.5	19.5	16.0	8.0
I-2c	52.0	23.0	18.0	7.0

Equatorial anchorage
of D *N*-ligand in

	s-s	4-s-6-a	4-a-6-s	a-a
II-3a	62	19	19	-
II-3b	67	18	15	-
II-3c	57	21	22	-
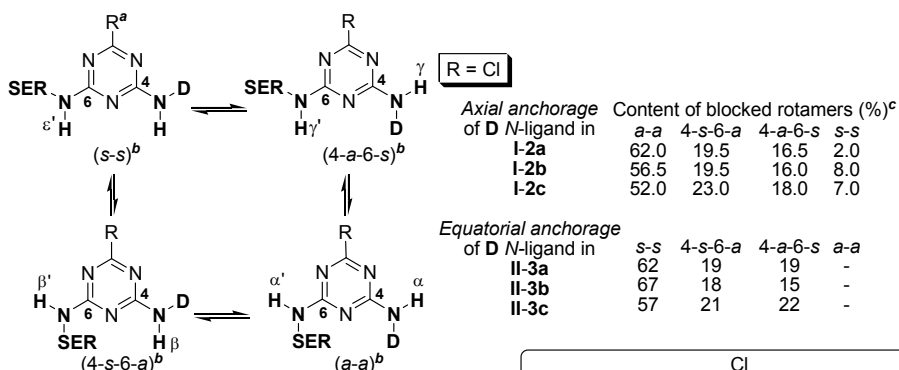
[a]R = Cl, in series **I-2**, **II-3**; piperazin-1-yl in series **I-4**, **II-5**
[b]Descriptors *s* (*syn*) and *a* (*anti*) are used with respect to
orientation of "*open-chain*" **SER**inolic and "*closed-chain*"
"**D**ioxanic *N*-ligands *vs*. **R**.
[c]Throughout percentages deduced from the ^{1}H NMR
spectra performed at room temperature in [D₆]DMSO
using signals of the *indicative protons* D-N*H* (α, β, γ)
and SER-N*H* (α', β', γ', ε') (see *Table 2*, *Table 3*).

I-2a (4-*a*-6-*s*)
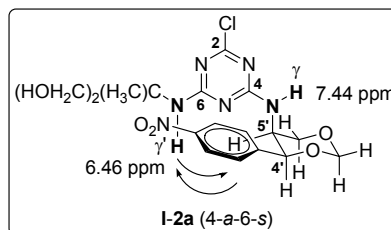
Scheme 2
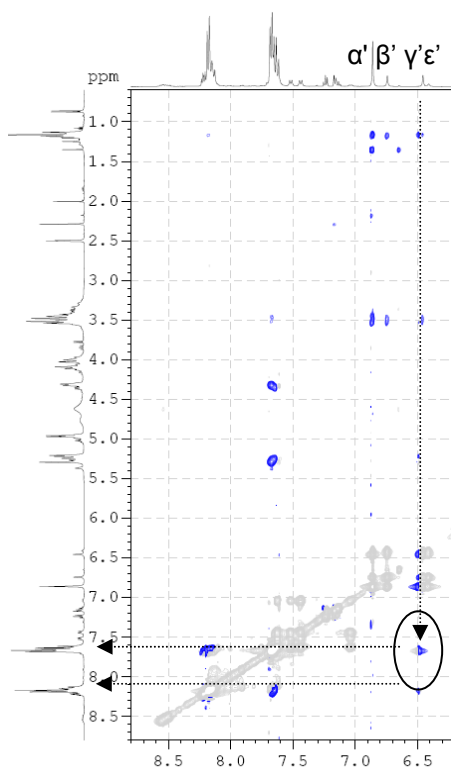
Figure 1. 2D-^{1}H,[1]H-NOESY chart of
compound **I-2a** (500 MHz, [D₆]DMSO)

The "*trans*" analogue **I-2a** (4-*s*-6-*a*)
was deduced logically, since its
incidence was comparable. The
major rotamer **I-2a** (*a-a*) was
preliminarily established by
considering the two closed SER-
N*H* δ values, signals α' and β', in
their *anti* local environment: 6.74
ppm in **I-2a** (4-*s*-6-*a*, signal β') and
6.86 ppm in **I-2a** (*a-a*), signal α'
(*Table 2*). If so, in the *syn* local
environments, rotamers **I-2a** (4-*a*-
6-*s*), **I-2a** (*s-s*), the δ values of
protons SER-N*H* are also very
related, 6.46 (signal γ') and 6.41
(signal ε') respectively. It is to note
that compound **I-2a** was the single
case that made possible the
rotamerism recognition based on
NOESY Experiment.

39

Table 2. Relevant ^{1}H NMR data of restricted rotation about
C(*s*-triazine)- N(exocyclic) bonds in compounds **I-2a**-**c**:
axial anchorage of the 1,3-dioxane N-ligand

No.	*T* (K)	Indicative protons	Discriminating δ_H (ppm) values and multiplicity $^3J_{H,H}$ (Hz)[a] in blocked rotamers			
			a-a	*4-s-6-a*	*4-a-6-s*	*s-s*
			α[b] α'	β β'	γ γ'	ε ε'
I-2a	298	D-N*H*	7.62 (d, 10.0)	7.52 (d, 9.5)	7.44 (d, 9.5)	-[c]
		SER-N*H*	6.86 (s)	6.74 (s)	6.46[d] (s)	6.41 (bs)
		OH	4.95, 4.75 (2×bs)			
	353		D-N*H* 7.01 (bs); SER-N*H* 6.51, 6.41 (2×bs); OH 4.50 (bs)			
I-2b	303	D-N*H*	7.62 (d, 9.0)	7.52 (d, 9.5)	7.46 (d, 10.0)	-
		SER-N*H*	6.77 (s)	6.64 (s)	6.34 (s)	6.33 (bs)
		OH	4.72-4.68 (m, 5.5), 4.56 (dd, 6.0), 4.52 (dd, 5.8), 4.47 (dd, 5.3)			
	353		D-N*H* 7.02 (bs); SER-N*H* 6.43, 6.31 (2×bs); OH 4.37 (bs)			
I-2c	303	D-N*H*	7.53, 7.52, 7.49[e]			
		SER-N*H*	6.57 (s)	6.50 (s)	6.27 (s)	6.21 (bs)
		OH	4.51 - 4.58 (m, 6.0)			
	353		D-N*H* 7.01 (bs); SER-N*H* 6.30, 6.24 (2×bs); OH 4.53, 4.36 (2×bs)			

Final rotational status of I-2a-c
*i) slow free rotation about bond C-4(s-triazine)-NH (**D** N-ligand)*
*ii) slow exchange about bond C-6(s-triazine)-NH (**SER** N-ligand)*

[a]As ^3J(ax-N*H-H*-5-e) in **D** N-ligand, ^3J(C*H₂*OH) in **SER** N-ligand, also supported by the 2D-^{1}H, ^{1}H-COSY Charts.
[b]Relevant peaks for (VT) ^{1}H NMR analysis (*Scheme 2, Figure 3*).
[c]Rotamers not found on the D-N*H* zone of the spectrum: the corresponding abundance was adopted from the SER-N*H* signal, ε'.
[d]Deduced from the 2D-^{1}H,[1]H-NOESY Experiment (*Figure 1*).
[e]Not assignable as overlapped signals.

Therefore, in order to validate this assignment for the entire series **I-2** and to predict the rotamerism occurrence in series **II-3** (*Table 3*), computational methods were applied to compounds **I-2a** (*axially* anchored) and **II-3a** (*equatorially* anchored) (*Table 4*).

Thus, by optimisation of rotational stereoisomers (*a-a*) and (*s-s*) of compounds **I-2a** and **II-3a** at B3LYP/6-311++G** level of theory and taking into account the effect of solvent (DMSO), we found out that:

i) In compounds **I-2a**-**c**, in agreement with ^{1}H NMR data, frozen stereoisomers **I-2a**-**c** (*a-a*) were indeed dominant while **I-2a**-**c** (*s-s*) should be the minor ones.

ii) In contrast, in series **II-3**, if the 1,3-dioxanic *N*-ligand was *equatorially* amino-linked to *s*-triazine, the less crowded rotamers **II-3a**-**c** (*s-s*) were, this time, the major species.

If so, regardless the type of anchorage of **D** *N*-ligand (*Table 2*, *Table 3*, *Table 4*) in each series, the most polar rotamer, hence the highest solvated, was dominant, displaying the most deshielded indicative protons SER-N*H* and D-N*H* as well. Surprisingly, in spite of opposite incidence of the major stereoisomer, **I-2** (*a-a*) *vs.* **II-3** (*s-s*), overall, the rotameric content was similar in the two series (*Scheme 2*).

Table 3. Relevant ^{1}H NMR data of restricted rotation about
C(*s*-triazine)-N(exocyclic) bonds in compounds **II-3a**-**c**:
equatorial anchorage of the 1,3-dioxane N-ligand

No.	T (K)	Indicative protons	Discriminating δ_H (ppm) values and multiplicity $^3J_{H,H}$ (Hz)^a in blocked rotamers			
			s-s	4-*s*-6-*a*	4-*a*-6-*s*	*a-a*
			α^b α'	β β'	γ γ'	ε^c ε'
II-3a	**298**	D-N*H*	8.03 (dd, 6.0)	7.91 (dd, 6.0)	7.89 (dd, 6.0)	-
		SER-N*H*	6.92 (s)	6.82 (s)	6.80 (s)	-
		OH	4.80 (dd, 5.8), 4.67 (dd, 5.8), 4.66 (dd, 6.0)			
	353		D-N*H* 7.48 (bs); SER-N*H* 6.58, 6.45 (2×bs); OH 4.52 (bs)			
II-3b	**298**	D-N*H*	8.05 (dd, 5.5)	7.91 (dd, 6.3)	7.88 (dd, 6.3)	-
		SER-N*H*	6.83 (s)	6.75 (s)	6.71 (s)	-
		OH	4.81 (dd, 5.5, 6.5), 4.76 (dd, 5.5), 4.64 – 4.60 (m, 5.8)			
	353		D-N*H* 7.50 (bs); SER-N*H* 6.47, 6.39 (2×bs); OH 4.53 (bs)			
II-3c	**303**	D-N*H*	8.01 (dd, 4.8)	7.94 (dd, 6.5)	7.92 (dd, 6.5)	-
		SER-N*H*	6.64 (s)	6.60 (s)	6.56 (s)	-
		OH	4.68 (dd, 5.5), 4.56 (dd, 5.5), 4.53 – 4.48 (m, 5.5)			
	353		D-N*H* 7.60 (bs); SER-N*H* 6.40, 6.31 (2×bs); OH 4.42 (bs)			
Final rotational status of II-3a-c						
*i) slow free rotation about bond C-4(s-triazine)-NH (**D** N-ligand)*						
*ii) slow exchange about bond C-6(s-triazine)-NH (**SER** N-ligand)*						
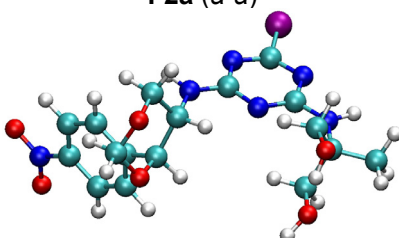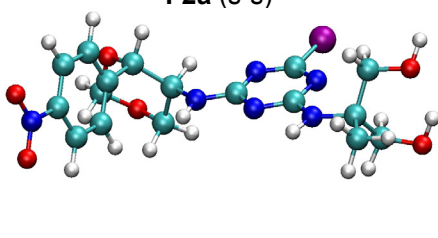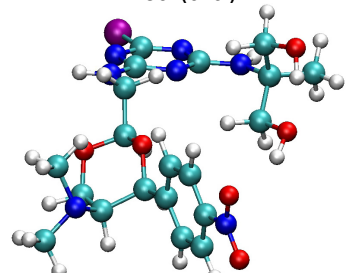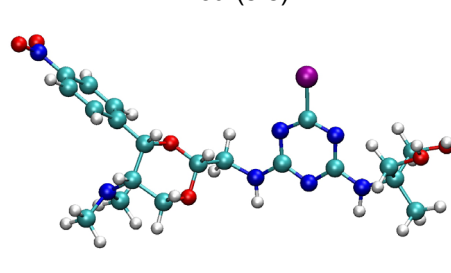
^{a}As ^3J(eq-C*H$_2$*-N*H*) in **D** *N*-ligand, ^3J(C*H$_2$OH*) in **SER** *N*-ligand, also supported by the 2D-^{1}H, ^{1}H-COSY Charts.
^{b}Relevant peaks for (VT) ^{1}H NMR analysis (*Scheme 2*).
^{c}Rotamers not found on the NH zone of the spectra.

^{1}H NMR data in *Tables 2* and *3* also deserved supplementary comments with respect to their accuracy (*Figure 2*).

The anticipated number of rotamers was not ^{1}H NMR observed in all cases, *e.g.* terms **II**-**3a-c** (*a-a*) were throughout absent (*Table 3*) as well as D-N*H* signals of **I-2a-c** (*s-s*) (*Table 2*). It was ^{13}C NMR spectra, which revealed, for almost all positions, distinct peaks for each rotameric environment (see **EXPERIMENTAL SECTION**).

Table 4. Bond orders (B.O.), dipole moments μ (D), ZPE corrected relative energies ΔH_{0K} (kJ/mol) and relative free energies ΔG_{298} (kJ/mol) of blocked rotamers (*a-a*) and (*s-s*) of compounds **I-2a** and **II-3a**.

Axial anchorage of the 1,3-dioxane N-ligand	
I-2a (*a-a*)	**I-2a** (*s-s*)

Equatorial anchorage of the 1,3-dioxane N-ligand	
II-**3a** (*a-a*)	**II**-**3a** (*s-s*)

B.O.[a]		D	ΔH_0 K	ΔG_{298}	B.O.		D	ΔH_0 K	ΔG_{298}
C(4)-N<	C(6)-N<				C(4)-N<	C(6)-N<			
B3LYP/6-311++G* / CPCM / DMSO[b]									
I-2a (*a-a*)					**I-2a** (*s-s*)				
1.23	1.25	**4.84**	0.00	**0.00**	1.22	1.24	**2.07**	2.93	**3.14**
II-3a (*a-a*)					**II-3a** (*s-s*)				
1.24	1.25	**7.81**	0.00	**1.38**	1.24	1.24	**11.63**	1.38	**0.00**

[a]Wiberg bond order calculated within the NBO (Natural Bonding Orbital) analysis.
[b]The effect of solvent took into account by using the implicit solvent method CPCM (Conductor-like Polarizable Continuum Model) implemented in Gaussian 09.

In both series **I**-**2** and **II**-**3**, protons N*H* were strongly chelatized by DMSO since clear ^3J coupling patterns through nitrogen, ax-N*H*-C*H*-5-*e* (*Table 2*) and eq–C*H*$_2$-N*H*- (*Table 3*) were detectable respectively and fully confirmed by the 2D-^{1}H,^{1}H-COSY Charts. Hence, they were not "mobile" at all as their lifetime of the spin state, τ_1 was greater than 0.1 sec. in compounds **I**-**2a**-**c** and greater than 0.15 sec. in **II**-**3a**-**c** [1a]. No broadening of the NH lines due to $^1J(^{14}$N-H) heterocoupling was observed [1c]. Particularly, the magnitude of vicinal coupling ax-N*H*-C*H*-5-*e* (*Table 2*, 9-10 Hz) we considered rather stereospecific [2b] for a preferred "s-*trans*-out" spatial arrangement between s-triazinyl and 1,3-dioxan-*c*-5-yl units with respect to the axial bond C-5-N (*Scheme 2*) [2b].[*]
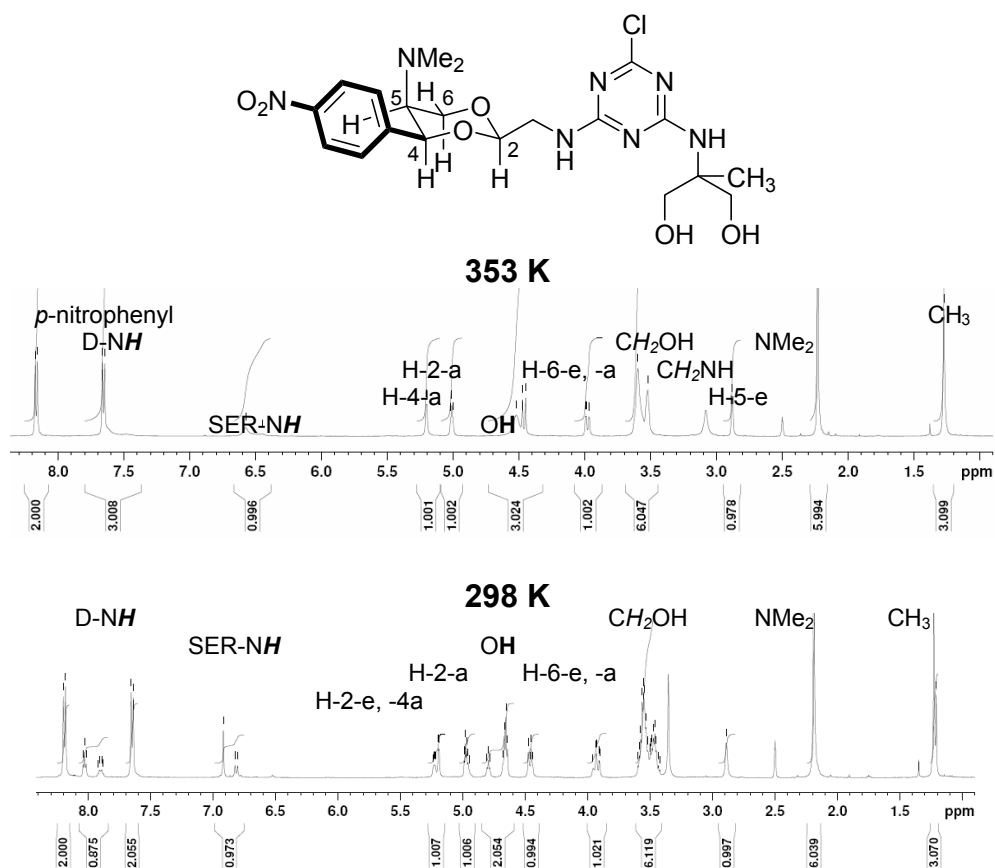
Figure 2. ^{1}H NMR evolution of compound **II-3a** (500 MHz time scale, [D$_6$]DMSO)

---

[*] Calculated as $\tau_1 > J^1$ (sec.) where *J* is the vicinal coupling constant >C*H*-N*H*-, ~ 9.5 Hz in **I**-**2a**-**c** (*Table 2*) and ~ 6.5 Hz in **II**-**3a**-**c** (*Table 3*).

The same was valid for hydroxyl protons who displayed typical (overlapped) dd signals in all rotamers, $^3J(CH_2\text{-}OH)$ ~ 5.5 Hz, except compound **I**-**2a** (*Table 2*). As one can observe, hydroxyl protons were less sensitive to the distinct rotameric ambiances, hence not useful for their quantitative evaluation.

On heating up to 80 °C, VT ^{1}H NMR spectra disclosed many successive coalescences, consistent with a complex dynamic evolution. Therefore, three hypotheses simplifying the problem, we had to adopt in the following:

i) By intercalation of the *s*-triazine ring, the two *N,N'*-ligands were "sufficiently remote" for a certain signal NH exhibited by each of them, D-N*H* (α, β, γ) or SER-N*H* (α', β', γ') be relevant for the rotational behaviour of the group to which this signal belongs, **D** or **SER**, only (*Figure 3*).

ii) Regardless series, **I**-**2** or **II**-**3**, completely de-blocking of a certain rotamer (*Scheme 2*) was mandatory to the equilibria involving, distinctly, its two *N*-ligands, **SER** and **D**. However, the final status of the entire molecule, as a wholly free rotating structure, required the validation of all fourth pathways.

iii) The consecutive nature of the four equilibria (*Scheme 2*) was disclosed as the succession in which the coalescences of NH signals ("indicative protons") appeared (*Figure 3*).

Accordingly, one can reach the preliminary conclusion that our *N,N'*-unsimmetrically substituted chlorodiamino-*s*-triazines could be but partially deblocked, with respect to **D** *N*-ligand about bonds C(4)(*s*-triazine)-N(exocyclic) only (*Figure 3*). This acquired *local rotational status* (*Table 2*, *Table 3*)
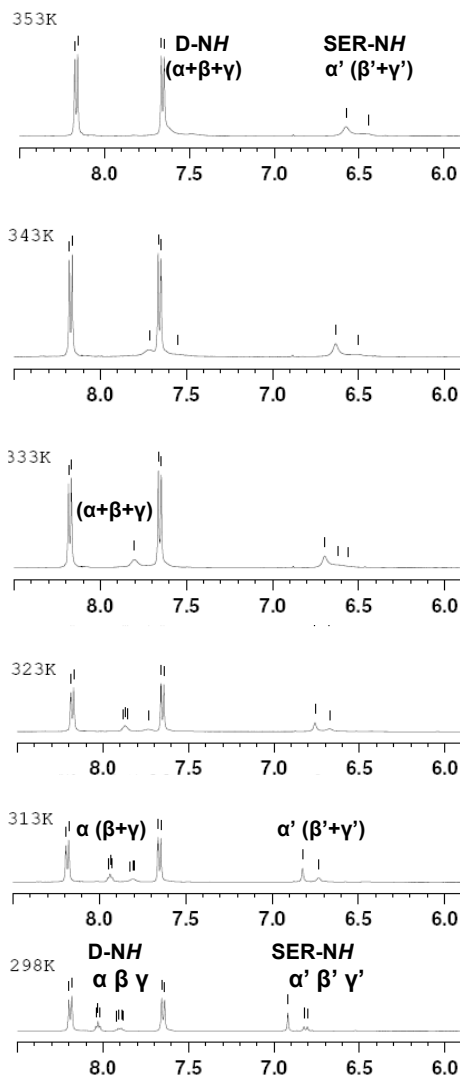
Figure 3. VT ^{1}H NMR spectra of compound **II**-**3a** on 500 MHz time scale ([D$_6$]DMSO)

we entitled "*slow free rotation*" (see the later discussion) while the **SER** *N*-ligand, more solvated at room temperature, displayed, at 80 ^{o}C, a typical *slow exchange status between unequally populated sites* [1a].

Furthermore, since computational data (*Table 4*) predicted the same bond order concerning C(*s*-triazine)-N(exocyclic) connections, it followed that the above delay in the dynamic behaviour of the two *N*-ligands was dictated mainly by solvation and not by electronic effects.

If so, we calculated the so-called "*temperature gradients*" $[\Delta\delta(NH)/\Delta T]\times10^3$ of protons NH for the major rotational diastereomer in series **I**-**2** and **II**-**3** (*Table 5*) [1b].

Table 5. Temperature gradients of protons NH of compounds **I**-**2a**-**c** and **II**-**3a**-**c**

Compd. (as major rotamer)		δ_H (ppm) (*T*, K)		$[\Delta\delta(NH)/\Delta T]\times10^3$ (ppb/K)[a]	
	T (K)	D-N*H*	SER-N*H*	D-N*H*	SER-N*H*
I-**2a** (*a*-*a*)	298	7.62	6.86	-11.1	-8.2
	353	7.01	6.41[b]		
I-**2b** (*a*-*a*)	303	7.62	6.77	-12.0	-6.8
	353	7.02	6.43		
I-**2c** (*a*-*a*)	303	-	6.57	-	-6.6
	353		6.24		
II-**3a** (*s*-*s*)	298	8.03	6.92	-10.0	-6.2
	353	7.48	6.58		
II-**3b** (*s*-*s*)	298	8.05	6.83	-10.0	-6.5
	353	7.50	6.47		
II-**3c** (*s*-*s*)	303	8.01	6.64	-8.2	-4.8
	353	7.60	6.40		

[a]Calculated as $[\delta(NH)_{353\,K} - \delta(NH)_{r.t.}] / (353 – T_{r.t.})$;
[b]At 353 K, in all cases, as signal displayed by the major SER-NH rotational site.

Thus, as recently Simanek observed in the case of elaborated amino-*s*-triazines [8], although temperature gradient is usually applied to peptides and proteins [1b], it is generally accepted and indicative that if this coefficient is more negative than -4 ppb/K in aqueous solution, the NH group was, initially, exposed to solvent and not involved in intramolecular hydrogen bonds. Conversely, a temperature gradient less negative than -4 ppb/K indicates the NH protons being, primarily, involved in intramolecular hydrogen bonding.

Our temperature gradients (*Table 5*) were consistent with the below assignments:

i) Although, at room temperature, D-N*H* protons were by far more chelated by the solvent, upon heating, they faster "escaped" from the solvent cage, in agreement with the *slow free rotating status* reached by these "*closed-chain*" *N*-ligands. The D-N*H*-solvation in our chlorodiamino-*s*-triazines was not dependent on the type of **D**-anchorage, *axial* or *equatorial*.

ii) In the "*open-chain*" SER-NH part, the NH protons less interacted with the solvent, presumably because of an intramolecular >NH…OH- partial association, clearly observed in the case of TRIS derivatives I-**2c** and II-**3c**.

To conclude, the late deblocking of the **SER** *N*-ligand was due to solvation of its OH protons and not to NH. Moreover, NH signals change progressively, from "amide type protons" (r.t.) to authentic "amine protons" upon heating up to 353 K.

In the end, we note the above "unsynchronised" **D** *vs.* **SER** evolution" to be completely different with respect to that of *symmetrically N,N'*-substituted chlorodiamino-*s*-triazines with the same *N*-ligands, previously reported by us. Thus, if the *N,N'*-identical ligands were serinols **a**-**c** (*Scheme 1*) a complete but slow rotational mobility (a single mediated structure) was observed at 353 K ($\Delta G^{\neq}$ = 68.10 – 69.22 kJ/mol). In contrast, if *N,N'*-identical ligands **I** or **II** were present (*Scheme 1*), a single mediated structure (slow rotation, $\Delta G \neq$ 71.20 kJ/mol) was found only in the case of the double equatorial **D** linkage of **II** [2b].

Rotational stereochemistry phenomena in melamines

Keeping in mind the above findings, melamines **I**-**4a**-**c** and **II**-**5a**-**c** were examined following the same algorithm (*Table 6*, *Table 7*, *Figure 4*).

By replacing the *s*-triazine C-2-chlorine atom with a bulky and strong releasing substituent, piperazine, the π-deficiency of *s*-triazine ring obviously decreased. However, *at room temperature*, essentially unlike from other simpler melamines [6], rotamerism was still existent. Moreover, the independent rotational evolution of the two *N*-ligands was once more revealed, since the number of broad singlets, D-N*H vs.* SER-N*H*, was not the same (*Table 6*)[**]. Therefore, the rotational situation of our melamines, at room temperature, we assigned as a *slow exchange between unequally populated sites* [1a]. Once again, since but in one case, compound **I**-**4c**, the 2D-^{1}H,^{1}H-COSY chart detected some [3]*J*(ax-N*H*-C*H*-5-e) coupling patterns, the D-N*H* and SER-N*H* lines width we associated to the above slow motion and not to a [1]*J*(^{14}N-H) heterocoupling.

Another interesting feature we observed concerning the "mobile protons", OH and NH of piperazine, Pip-N*H*. In the less hydroxylated compounds **I**-**4a**, **I**-**4b**, **II**-**5a** **II**-**5b**, these protons exhibited a unique broad singlet, suggesting a rapid intermolecular exchange defining a mediated environment, -CH$_2$O*H* ⇆ *H*N-Pip. Hence, an important intermolecular interaction of these melamines we had to suppose, *e.g.* a polymeric self-assembly. If so, in the case of TRIS based melamines **I**-**4c** and **I**-**5c**, it is very likely that the *internal* association of its three geminal hydroxymethyl groups prevailed against the alternative one, *external*, that with NH of piperazine.

[**] The number of these broad NH singlets is not indicative at all for the number of rotational sites because in the case of compound **I**-**4c** only the 2D-^{1}H,^{1}H-COSY chart clearly exhibited an ax-N*H*-C*H*-5-e [3]*J* coupling; it disclosed four rotamers displaying the two broad D-N*H* singlets.

Table 6. Relevant ^{1}H NMR data of restricted rotation about
C(*s*-triazine)-N(exocyclic) bonds in melamines **I**-**4a**-**c** and **II**-**5a**-**c**

No.	Discriminating δH (ppm) values and multiplicity 3JH,H (Hz)a				
	T (K)	D-NH	SER-NH	Pip-NH	OH
Axial anchorage of the 1,3-dioxane N-ligand					
I-4a	**298**	5.80, 5.70 (2×bs)	5.56 (bs)	4.77 (bs)	
	Tc	5.59 (313, bs)	-	-	
	363	5.49 (d, 9.5)	5.43 (s)	4.54 (bs)	
I-4b	**298**	5.81, 5.68 (2×bs)	5.55, 5.44 (2×bs)	4.74 (bs)	
	Tc	5.65 (313, bs)	5.44 (313, bs)	-	
	363	5.52 (d, 9.5)	5.35 (s)	4.54 (bs)	
I-4c	**293**	5.92, 5.81 (2×bs)	5.57, 5.49 (2×bs)	3.45 (bs)	2.63 (s)
	Tc	5.77 (313, bs)	5.48 (313, bs)	-	-
	353	5.61 (d, 9.5)	5.43 (s)	-	2.68 (s)
Equatorial anchorage of the 1,3-dioxane N-ligand					
II-5a	**298**	6.77, 6.60 (2×bs)	5.71, 5.60 (2×bs)	4.84 (bs)	
	Tc	6.49 (323, bs)	5.62 (323, bs)		
	353	6.27 (bdd as bt, 5.5)	5.55 (s)	4.66 (bs)	
II-5b	**298**	6.82, 6.58 (2×bs)	5.59 (bs)	4.77 (bs)	
	Tc	5.51 (323, bs)	-	-	
	353	6.29 (bdd as bt, 5.5)	5.46 (s)	4.68 (bs)	
II-5c	**303**	6.93, 6.80, 6.69 (3×bs)	5.62 (bs)	3.60 (bs)	4.81 (bs)
	Tc	6.60 (323, bs)	-	-	
	353	6.38 (bs)	5.53 (s)	-	4.65 (bs)

Final rotational status of I-4a-c and II-5a-c
Fast free rotation about bonds C-4(s-triazine)-NH (D ligand),
C-6(s-triazine)-NH (SER ligand) and C-2(s-triazine)-N (Pip. Ligand)

[a] As ^3J(ax-N*H*-C*H*-5-e) in **D** *N*-ligand, ^3J(C*H$_2$*O*H*) in **SER** *N*-ligand in series **I**-**4**; as ^3J(eq-C*H$_2$*-N*H*) in **D** *N*-ligand, ^3J(C*H$_2$*O*H*) in **SER** *N*-ligand in series **II**-**5**.

On heating up to 80-90 °C, all our melamines could be completely "activated" about all connections C(*s*-triazine)-N(exocyclic) reaching a *fast free rotation status* as a single mediated structure.

Table 7. Maximum temperature gradients of protons NH
of compounds I-**4a**-**c** and II-**5a**-**c**

Compd. T (K)		Relevant δ_H (ppm)			$[\Delta\delta(NH)/\Delta T]\times10^3$ (ppb / K)		
		D-N*H*	SER-N*H*	Pip-N*H*	D-N*H*	SER-N*H*	Pip-N*H*
I-**4a**	298	5.80	5.56 [a]	4.77	-4.8	-2.0	-3.5
	363	5.49	5.43	4.54			
I-**4b**	298	5.81	5.55	4.74	-4.5	-3.1	-3.1
	363	5.52	5.35	4.54			
I-**4c**	298	5.92	5.57	3.45	-5.6	-2.5	-
	353	5.61	5.43	-			
II-**5a**	298	6.77	5.71	4.84	-9.1	-2.9	-3.3
	353	6.27	5.55	4.66			
II-**5b**	298	6.82	5.59	4.77	-9.6	-2.4	-1.6
	353	6.29	5.46	4.68			
II-**5c**	303	6.93	5.62	-	-11.0	-1.8	-
	353	6.38	5.53	-			

[a]At r.t., in all cases, as signal displayed by the most deshielded broad singlet, D-N*H* or SER-N*H* proton respectively.

Their dynamic evolution can be subjected to the below comments:

i) If different slow exchanging rotameric sites were initially detected for each **D** and **SER** N-ligand, the coalescence was reached, however, almost simultaneously by the two fragments. Higher T_c values (+10 K) were observed for melamines II-**5a**-**c** against I-**4a**-**c** in agreement with the more solvated ground states of equatorially anchored derivatives (*Table 7*).

ii) The final *fast free rotation* status acquired by all melamines was nicely supported by the line-shape analysis of signals D-N*H* (sharp doublet or broad triplet) and SER-N*H* (sharp singlet) (*Figure 4*). Undoubtedly, one cannot assign this line-width as to belong to "exchangeable" protons. In contrast, OH and Pip-N*H* signals, as unique broad lines, were fully consistent with the mobility of these protons involved in a fast exchange with the solvent in a molecular unique mediated environment.

iii) Temperature gradients (*Table 7*) crucially discriminated between amino-1,3-dioxane anchorages, *axial* or *equatorial vs.* solvent, melamines II-**5c** being, at room temperature, by far more D-N*H*…O=SMe$_2$ solvated. That is, the slow turning *equatorial* amino-linkage to the *s*-triazine ring was sterically less crowded, facilitating the access of DMSO to this connection. In the SER-N*H* and Pip-N*H* counterparts, the $\Delta\delta$(N*H*)/ΔT values, much less negative than -4, confirmed our earlier assignment, namely the intermolecular CH$_2$O*H* ⇆ *H*-Pip interchange in the case of *methyl*- (**a**) and *ethylserinol* (**b**) based melamines.
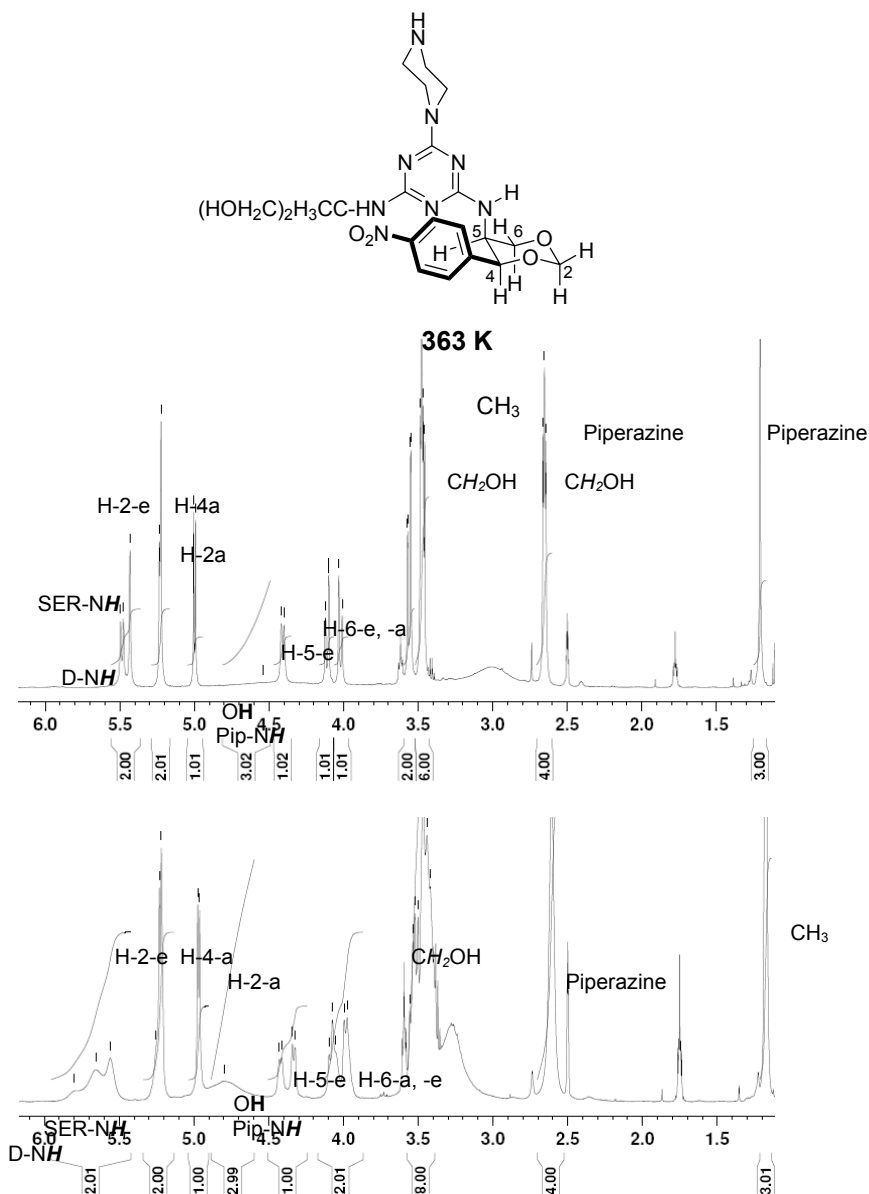
CONCLUSIONS

A rapid and efficient access to highly elaborated enantiopure melamines based on serinols was described. To the title report question,our answer consists of considering (i) primarily, the NH protons, involved in partial double

bonds, as very relevant for distribution of rotamers, dynamic behavior and sterically conditioned relationships with the solvent; (ii) the OH protons, indicative for solvation effects as intra- -OH…HO- or intermolecular -OH…HN< associations determining the stability of ground states.

EXPERIMENTAL SECTION

General. Melting points are uncorrected; they were carried out on ELECTROTHERMAL® instrument. Conventional NMR spectra were recorded on a Bruker® AM300 instrument operating at 300 and 75 MHz for ^{1}H and ^{13}C nuclei respectively. VT NMR experiments were performed on a Bruker® DMX500 instrument. All NMR spectra were measured in anhydrous commercially available deuteriated solvents. The ^{13}C NMR description of compounds exhibiting frozen rotamers at room temperature was made by considering them as one global structure. Multiple δ_C values for the same-labelled position means mixture of rotamers. Some specific abbreviations were used: bd (broad doublet) and bm (broad multiplet), *p*-NPh (*p*-nitrophenyl).TLC was performed by using aluminium sheets with silica gel 60 F_{254} (Merck®); flash column chromatography was conducted on Silica gel Si 60 (40–63 μm, Merck®). All visualisations were realised in UV at 254 nm. IR spectra were performed on a Perkin-Elmer® Paragom FT-IR spectrometer. Only relevant absorptions are listed [throughout in cm^{-1}: weak (w), medium (m) or (s) strong]. Microanalyses were performed on a Carlo Erba® CHNOS 1160 apparatus. Mass spectra (MS) were recorded on a Bruker® Esquire Instrument. Specific rotations were measured on a POLAMAT® Karl-Zeiss Jena instrument. Full characterisation and synthesis of compounds **I** and **II** we reported in detail elsewhere [2b].

*Typical procedure for the synthesis of compounds **I-2a-c** and **II-3a-c**. Preparation of compound **I-2c***

To anh. K_2CO_3 (1.512 g, 100%, 10.944 mmol) suspended in anh. THF (100 ml), solid 2-amino-2-(hydroxymethyl)propane-1,3-diol (TRIS, 1.325 g, 10.944 mmol) was added with vigorous stirring. The resulted suspension was cooled at -15 °C when cyanuric chloride (2.018 g, 10.944 mmol) as clear anh. THF (25 ml) solution was injected rapidly. The reaction mixture was let gently to reach room temperature and was kept as such for additional 24 hrs. with stirring. After this period, TLC monitoring indicated the intermediate 2,4-dichloro-6-{[1,3-dihydroxy-2-(hydroxymethyl)prop-2-yl]amino}-*s*-triazine **1c** as a single spot (eluent acetone : ligroin 2:1, R_f = 0.80). Freshly prepared (4*S*,5*S*)-5-amino-4-(4-nitrophenyl)-1,3-dioxane (**I**, **DX-NH$_2$**) (2.452 g, 100%, 10.944 mmol) and anh. K_2CO_3 (1.512 g, 100%, 10.944 mmol) were added and the reaction mixture was heated at reflux (65 °C) for 12 hrs. (TLC monitoring, eluent toluene : isopropanol 2:1, R_f = 0.80). When **I** and **1c** were detected in small traces only, the reaction mixture was cooled at room temperature.

Minerals were filtered off and well washed with anh. THF. The organic filtrate was evaporated under reduced pressure to dryness to provide 5.222 g crude product. This was purified by column chromatography on silica gel (eluent toluene : isopropanol 2:1) affording 4.110 g compound **I**-**2c** (84 % yield with respect to cyanuric chloride).

2-Chloro-6-{[1,3-dihydroxy-2-(methyl)prop-2-yl]amino}-4-{[(4S,5S)-4-(4-nitrophenyl)-1,3-dioxan-5-yl]amino}-s-triazine I-2a (80 %) yellowish crystalline powder, mp 107-118 °C (flash column chromatography, eluent toluene : isopropanol, 2:1 v/v). [Found: C 45.98, H 5.11, N 18.80; $C_{17}H_{21}ClN_6O_6$ (440.12) requires: C 46.32, H 4.80, N 19.06%]. R_f 0.86 (66% toluene/ isoporpanol). IR $\nu_{max.}$ (KBr) 3320 (s), 2946 (m), 2867 (m), 1581 (s), 1520 (s), 1411 (m), 1347 (s), 1242 (m), 1175 (s), 1094 (s), 1027 (s), 966 (m), 852 (m), 804 (s), 744 (m), 711 (m), 592 (w) cm^{-1}. ^{1}H NMR (500 MHz, $[D_6]$DMSO, 353 K)**:** δ_H 1.12 (3 H, s, CH_3), 3.48-3.57 (4 H, m, CH_2OH), 4.03 (1 H, d, $^2J_{H,H}$=11.0 Hz, H-6-a D-NH), 4.14 (1 H, d, $^2J_{H,H}$=11.5 Hz, H-6-e D-NH), 4.37 (1 H, d, $^3J_{H,H}$=9.0 Hz, H-5-e D-NH), 4.50 (2 H, bs, OH), 5.00 (1 H, d, $^2J_{H,H}$=6.0 Hz, H-2-a D-NH), 5.23 (1 H, d, $^2J_{H,H}$=6.0 Hz, H-2-e D-NH), 5.28 (1 H, s, H-4-a D-NH), 6.41, 6.51 (1 H, 2×bs SER-NH), 7.01 (1 H, bs D-NH), 7.66 (2 H, d, $^3J_{H,H}$=8.5 Hz, H-2, -6 p-NPh), 8.14 (2 H, d, $^3J_{H,H}$=7.0 Hz, H-3, -5 p-NPh) ppm; ^{13}C NMR (125 MHz, $[D_6]$DMSO, 298 K): δ_C 18.7, 18.8, 19.0, 19.4 (1 C, CH_3), 49.3, 49.5, 49.8 (1 C, C-5 D-NH), 58.8, 59.0 (1 C, C-2 SER-NH), 63.4, 63.5, 63.6, 63.7, 63.8 (2 C, CH_2-OH), 69.7, 70.1, 70.2, 70.5 (1 C, C-6 D-NH), 77.9, 78.2, 78.4, 78.8 (1 C, C-4 D-NH), 93.88, 93.94, 94.0, 94.2 (1 C, C-2 D-NH), 123.3, 123.5, 123.6, 123.8 (2 C, C-2, -6 p-NPh), 127.4, 127.5, 127.6, 127.9 (2 C, C-3, -5 p-NPh), 146.7, 146.76, 146.78 (1 C, C-1 p-NPh), 147.17, 147.21, 147.24 (1 C, C-4 p-NPh), 164.9, 165.1, 165.3, 165.4, 165.6 (2 C, C-4, -6 s-triazine), 167.76, 167.82, 167.9, 168.2 (1 C, C-2 s-triazine) ppm. MS (ESI+), m/z (rel. int. %) 463.1 [M+Na$^+$] (7.5), 443.1 [M$^+$+2] (38), 441.3 [M$^+$+H] (100), 315.1 (10). $[\alpha]_D^{25}$=-46 (0.5 % DMSO).

2-Chloro-6-{[1-hydroxy-2-(hydroxymethyl)but-2-yl]amino}-4-{[(4S,5S)-4-(4-nitrophenyl)-1,3-dioxan-5-yl]amino}-s-triazine I-2b (66 %) yellowish crystalline powder, mp 97-102 °C (flash column chromatography, eluent toluene : isopropanol, 2:1 v/v). [Found: C 47.35, H 5.25, N 18.79; $C_{18}H_{23}ClN_6O_6$ (454.14) requires: C 47.53, H 5.10, N 18.48%]. R_f 0.83 (66% toluene/i-PrOH). IR $\nu_{max.}$ (KBr) 3372 (s), 2972 (m), 2864 (m), 1587 (s), 1521 (s), 1414 (m), 1346 (s), 1242 (m), 1175 (s), 1095 (s), 1028 (s), 966 (m), 852 (m), 804 (m), 745 (w), 711 (m), 582 (w) cm^{-1}. ^{1}H NMR (500 MHz, $[D_6]$DMSO, 353 K): δ_H 0.72 (3 H, t, $^3J_{H,H}$=7.3 Hz, CH$_3$), 1.70 (2 H, bq, $^3J_{H,H}$=7.2 Hz, CH_2-CH$_3$), 3.50 (2 H, d, $^2J_{H,H}$=11.0 Hz, CH_2-OH), 3.59 (2 H, bd, $^2J_{H,H}$=8.5 Hz, CH_2OH), 4.02 (1 H, d, $^2J_{H,H}$=11.0 Hz, H-6-a D-NH), 4.14 (1 H, d, $^2J_{H,H}$=11.5 Hz, H-6-e D-NH), 4.37 (3 H, bs, H-5-e D-NH, OH), 5.00 (1 H, d, $^2J_{H,H}$=6.5 Hz, H-2-a D-NH), 5.23 (1 H, d, $^2J_{H,H}$=6.0 Hz, H-2-e D-NH), 5.28 (1 H, s, H-4-a D-NH), 6.31, 6.43 (1 H, 2×bs

SER-N*H*), 7.02 (1 H, bs D-N*H*), 7.65 (2 H, d, $^3J_{H,H}$=7.5 Hz, H-2, -6 *p*-NPh), 8.14 (2 H, bd, $^3J_{H,H}$=6.0 Hz, H-3, -5 *p*-NPh) ppm; ^{13}C NMR (125 MHz, [D$_6$]DMSO, 298 K): δ_C 7.8, 7.9 (1 C, CH$_3$), 22.1, 22.5, 23.0, 23.1 (1 C, *C*H$_2$-CH$_3$), 49.3, 49.37, 49.4, 49.9 (1 C, C-5 D-NH), 60.9 (1 C, C-2 SER-NH), 61.2, 61.3, 61.4, 61.5 (2 C, CH$_2$OH), 70.07, 70.1, 70.2, 70.5 (1 C, C-6 D-NH), 78.2, 78.4, 78.5, 78.8 (1 C, C-4 D-NH), 93.89, 93.94 (1 C, C-2 D-NH), 123.2, 123.3, 123.4, 123.6, (2 C, C-2, -6 *p*-NPh), 127.6, 127.7, 127.9, 128.0 (2 C, C-3, -5 *p*-NPh), 146.7, 146.8, (1 C, C-1 *p*-NPh), 147.1, 147.17, 147.22 (1 C, C-4 *p*-NPh), 164.7, 164.99, 165.04, 165.3, 165.4, 165.7 (2 C, C-4, -6 *s*-triazine), 167.8 (1 C, C-2 *s*-triazine) ppm. MS (ESI+), *m/z* (rel. int. %) 493.1 [M+K$^+$] (9), 477.1 [M+Na$^+$] (18), 457.1 [M$^+$+2] (35), 455.1 [M$^+$+H] (100), 438.2 (11), 437.2 (54), 419.2 (41). [α]$_D^{25}$=-34 (0.5 % DMSO).

2-Chloro-6-{[1,3-dihydroxy-2-(hydroxymethyl)prop-2-yl]amino}-4-{[(4*S*, 5*S*)-4-(4-nitrophenyl)-1,3-dioxan-5-yl]amino}-*s*-triazine I-2c (84 %) yellowish crystalline powder, mp 200-205 °C (flash column chromatography, eluent toluene : isopropanol, 2:1 v/v). [Found: C 45.01, H 4.39, N 18.59; C$_{17}$H$_{21}$ClN$_6$O$_7$ (456.12) requires: C 44.69, H 4.63, N 18.40%]. *R*$_f$ 0.80 (66% toluene/*i*-PrOH). IR $\nu_{max.}$ (KBr) 3369 (s), 2950 (m), 2865 (m), 1586 (s), 1519 (s), 1418 (m), 1387 (m), 1347 (s), 1243 (m), 1175 (s), 1096 (s), 1026 (s), 967 (m), 852 (w), 804 (m), 743 (m), 711 (m), 593 (w) cm^{-1}. ^{1}H NMR (500 MHz, [D$_6$]DMSO, 353 K): δ_H 3.66 (6 H, bs, C*H$_2$*OH), 4.05 (1 H, bd, $^2J_{H,H}$=9.5 Hz, H-6-a D-NH), 4.14 (1 H, d, $^2J_{H,H}$=11.5 Hz, H-6-e D-NH), 4.36 (1 H, bs, H-5-e D-NH), 4.36, 4.53 (3 H, 2×bs, OH), 5.00 (1 H, d, $^2J_{H,H}$=6.0 Hz, H-2-a D-NH), 5.23 (1 H, d, $^2J_{H,H}$=6.0 Hz, H-2-e D-NH), 5.27 (1 H, bs, H-4-a, D-NH), 6.24, 6.30 (1 H, 2×bs SER-N*H*), 7.01 (1 H, bs D-N*H*), 7.65 (2 H, d, $^3J_{H,H}$=8.0 Hz, H-2, -6 *p*-NPh), 8.14 (2 H, bd, $^3J_{H,H}$=7.0 Hz, H-3, -5 *p*-NPh) ppm; ^{13}C NMR (75 MHz, [D$_6$]DMSO, 298 K): δ_C 49.2, 49.3, 49.7 (1 C, C-5 D-NH), 59.4, 59.6, 60.1 (1 C, C-2 SER-NH), 62.1, 62.3, 62.4 (3 C, CH$_2$OH), 69.9, 70.3 (1 C, C-6 D-NH), 78.0, 78.1, 78.6 (1 C, C-4 D-NH), 93.8 (1 C, C-2 D-NH), 123.2, 123.4 (2 C, C-2, -6 *p*-NPh), 127.4, 127.6, 127.7 (2 C, C-3, -5 *p*-NPh), 146.6 (1 C, C-1 *p*-NPh), 147.1 (1 C, C-4 *p*-NPh), 164.9, 165.0, 165.2, 165.4 (2 C, C-4, -6 *s*-triazine), 167.6, 167.8 (1 C, C-2 *s*-triazine) ppm. MS (CI, isobutane) *m/z* (rel. int. %) 513 [M$^+$ +HC(CH$_3$)$_3$-2 H] (20), 495 [M+K$^+$] (9), 457 [M$^+$] (100), 421 (10), 225 (11),140(10). [α]$_D^{25}$=-36 (0.5 % DMSO).

2-Chloro-6-{[1,3-dihydroxy-2-(methyl)prop-2-yl]amino}-4-{[(2*R*,4*S*,5*S*)-5-(dimethylamino)-4-(4-nitrophenyl)-1,3-dioxan-2-yl]methylamino}-*s*-triazine II-3a (83 %) yellowish crystalline powder, mp 126-134 °C (flash column chromatography, eluent toluene : ethanol, 1:5 v/v). [Found: C 47.95, H 5.51, N 19.39; C$_{20}$H$_{28}$ClN$_7$O$_6$ (497.18) requires: C 48.24, H 5.67, N 19.69%]. *R*$_f$ 0.80 (17% toluene/EtOH). IR $\nu_{max.}$ (KBr) 3382 (s), 3276 (s), 2941 (m), 2878 (m), 1587 (s), 1521 (s), 1462 (m), 1412 (m), 1349 (s), 1153 (m), 1113 (m), 1057 (s),

52

852 (w), 804 (m), 753 (w), 709 (m), 571 (w) cm^{-1}. ^{1}H NMR (500 MHz, [D$_6$]DMSO, 353 K)**:** δ_H 1.27 (3 H, s, CH$_3$), 2.23 [6 H, s, N(CH$_3$)$_2$], 2.88 (1 H, dd as t, $^3J_{H,H}$=3.0 Hz, H-5-e D-NH), 3.52 (2 H, bs, C*H$_2$*-NH), 3.60 (4 H, bs, C*H$_2$*OH), 3.98 (1 H, dd, $^3J_{H,H}$=2.0 Hz, $^2J_{H,H}$=12 Hz, H-6-a D-NH), 4.46 (1 H, d, $^2J_{H,H}$=12.0 Hz, H-6-e D-NH), 4.52 (2 H, bs, OH), 5.01 (1 H, dd as t, $^3J_{H,H}$=4.3 Hz, H-2-a D-NH), 5.20 (1 H, d, $^3J_{H,H}$=2.0 Hz, H-4-a D-NH), 6.45, 6.58 (1 H, 2×bs SER-N*H*), 7.48 (1 H, bs, CH$_2$N*H*), 7.66 (2 H, d, $^3J_{H,H}$=8.5 Hz, H-2, -6 *p*-NPh), 8.17 (2 H, d, $^3J_{H,H}$=9.0 Hz, H-3, -5 *p*-NPh) ppm; ^{13}C NMR (125 MHz, [D$_6$]DMSO, 298 K): δ_C 18.8, 19.0 (1 C, CH$_3$), 43.8 [2 C, N(CH$_3$)$_2$], 44.3, 44,4, 44,8 (1 C, CH$_2$NH), 58.46, 58.51 (1 C, C-5 D-NH), 58.98, 59.04 (1 C, C-2 SER-NH), 63.5, 63.6, 63.6, 63.9 (2 C, CH$_2$OH), 64.4, 64.6 (1 C, C-6 D-NH), 80.05, 80.12, 80.3 (1 C, C-4 D-NH), 99.0, 99.2, 99.5 (1 C, C-2 D-NH), 123.3 (2 C, C-2, -6 *p*-NPh), 127.1 (2 C, C-3, -5 *p*-NPh), 146.7 (1 C, C-1 *p*-NPh), 148.8, 148.9 (1 C, C-4 *p*-NPh), 165.0, 165.3, 165.6, 165.85, 165.91 (2 C, C-4, -6 *s*-triazine), 167.88, 167.93, 168.4 (1 C, C-2 *s*-triazine) ppm. MS (ESI+), *m/z* (rel. int. %) 537.2 [M+K$^+$] (2), 520.1 [M+Na$^+$] (3.5), 500.2 [M$^+$+3], 498.1 [M$^+$+H] (100), 462.2 (7), 273.2 (7), 208.0 (11), 182.2 (17). [α]$_D^{25}$=+147 (0.5 % DMSO).

2-Chloro-6-{[1-hydroxy-2-(hydroxymethyl)but-2-yl]amino}-4-{[(2*R*,4*S*, 5*S*)-5-(dimethylamino)-4-(4-nitrophenyl)-1,3-dioxan-2-yl]methylamino}-*s*-triazine II-3b (42 %) yellowish crystalline powder, mp 110-115 °C (flash column chromatography, eluent diethyl ether : ethanol : water, 0.5:8:1 v/v/v). [Found: C 48.98, H 5.81, N 19.39; C$_{21}$H$_{30}$ClN$_7$O$_6$ (511.19) requires: C 49.27, H 5.91, N 19.15%]. R_f 0.73 (5% Et$_2$O/84% EtOH/H$_2$O). IR $\nu_{max.}$ (KBr) 3275 (s), 2971 (m), 2881 (m), 1591 (s), 1521 (s), 1464 (m), 1411 (s), 1349 (s), 1154 (m), 1117 (m), 1059 (s), 852 (w), 802 (m), 752 (w), 708 (m), 570 (w) cm^{-1}. ^{1}H NMR (500 MHz, [D$_6$]DMSO, 353 K): δ_H 0.779, 0.784 (3 H, 2×t, $^3J_{H,H}$=7.5 Hz, CH$_3$), 1.79 (2 H, q, $^3J_{H,H}$=7.3 Hz, C*H$_2$*CH$_3$), 2.27 [6 H, s, N(CH$_3$)$_2$], 2.97 (1 H, bs, H-5-e D-NH), 3.51-3.54 (2 H, m, C*H$_2$*-OH), 3.56 (2 H, dd as t, $^3J_{H,H}$=3.0 Hz, C*H$_2$*NH), 3.67 (2 H, 2×d as t, $^2J_{H,H}$=11.0, 12.5 Hz, C*H$_2$*OH), 4.00 (1 H, d, $^2J_{H,H}$=11.5 Hz, H-6-a D-NH), 4.49 (1 H, d, $^2J_{H,H}$=12.5 Hz, H-6-e D-NH), 4.53 (2 H, bs, OH), 5.01 (1 H, dd as t, $^3J_{H,H}$=4.5 Hz, H-2-a D-NH), 5.24 (1 H, bs, H-4-a D-NH), 6.39, 6.47 (1 H, 2×bs SER-N*H*), 7.50 (1 H, bs, CH$_2$N*H*), 7.67 (2 H, d, $^3J_{H,H}$=8.5 Hz, H-2, -6 *p*-NPh), 8.17 (2 H, d, $^3J_{H,H}$=9.0 Hz, H-3, -5 *p*-NPh) ppm; ^{13}C NMR (125 MHz, [D$_6$]DMSO, 298 K): δ_C 7.9, 7.96, 8.01, 8.1 (1 C, CH$_3$), 22.2, 22.3, 22.6, 23.0, 23.1 (1 C, CH$_2$-CH$_3$), 43.8 [2 C, N(CH$_3$)$_2$], 44.2, 44.4, 44.8 (1 C, CH$_2$NH), 58.6 (1 C, C-5 D-NH), 60.9, 61.0, 61.1, 61.2, (1 C, C-2 SER-NH), 61.46, 61.51, 61.65, 61.71 (2 C, CH$_2$OH), 64.4, 64.6 (1 C, C-6 D-NH), 80.0, 80.1, 80.3 (1 C, C-4 D-NH), 99.1, 99.3, 99.6 (1 C, C-2 D-NH), 123.4 (2 C, C-2, -6 *p*-NPh), 127.1 (2 C, C-3, -5 *p*-NPh), 146.8 (1 C, C-1 *p*-NPh), 148.6, 148.8 (1 C, C-4 *p*-NPh), 165.0, 165.2, 165.6, 165.8, 165.9, 166.0, (2 C, C-4, -6 *s*-triazine), 167.3, 167.9, 168.3 (1 C, C-2 *s*-triazine) ppm. MS (ESI+), *m/z* (rel. int. %) 515.4 [M$^+$+4] (8), 514.4 [M$^+$+3], 512.4 [M$^+$+H] (100), 498.4 (4). [α]$_D^{25}$=+128 (0.5 % DMSO).

2-Chloro-6-{[1,3-dihydroxy-2-(hydroxymethyl)prop-2-yl]amino}-4-{[(2*R*, 4*S*,5*S*)-5-(dimethylamino)-4-(4-nitrophenyl)-1,3-dioxan-2-yl]methylamino}-*s*-triazine II-3c (95 %) yellowish crystalline powder, mp 138-140 °C (flash column chromatography, eluent ligroin : acetone, 1:4 v/v). [Found: C 47.08, H 5.55, N 19.38; $C_{20}H_{28}ClN_7O_7$ (513.17) requires: C 46.74, H 5.49, N 19.08%]. R_f 0.80 (20% ligroin/acetone)**.** IR $\nu_{max.}$ (KBr) 3369 (s), 2945 (m), 2878 (m), 1583 (s), 1520 (s), 1412 (m), 1348 (s), 1299 (m), 1154 (m), 1113 (m), 1054 (s), 1014 (m), 852 (w), 804 (m), 710 (m), 597 (w) cm^{-1}. ^{1}H NMR (500 MHz, [D$_6$]DMSO, 353 K): δ_H 2.28 [6 H, s, N(CH$_3$)$_2$], 2.97 (1 H, bs, H-5-e D-NH), 3.53 (2 H, s, C*H$_2$*-NH), 3.72 (6 H, s, C*H$_2$*OH), 4.01 (1 H, d, $^2J_{H,H}$=11.5 Hz, H-6-a D-NH), 4.42 (3 H, bs, OH), 4.49 (1 H, d, $^2J_{H,H}$=12.5 Hz, H-6-e D-NH), 5.03 (1 H, bs, H-2-a D-NH), 5.24 (1 H, bs, H-4-a D-NH), 6.31, 6.40 (1 H, 2×bs SER-N*H*), 7.60 (1 H, bs CH$_2$N*H*), 7.68 (2 H, d, $^3J_{H,H}$=8.5 Hz, H-2, -6 *p*-NPh), 8.18 (2 H, d, $^3J_{H,H}$=8.5 Hz, H-3, -5 *p*-NPh) ppm; ^{13}C NMR (75 MHz, [D$_6$]DMSO, 298 K): δ_C 43.7 [2 C, N(CH$_3$)$_2$], 44.2, 44.6 (1 C, CH$_2$NH), 58.5 (1 C, C-5 D-NH), 59.6, 60.0, 60.2 (1 C, C-2 SER-NH), 62.4, 62.5 (3 C, CH$_2$OH), 64.4, 65.3, 67.4 (1 C, C-6 D-NH), 80.5 (1 C, C-4 D-NH), 99.2 (1 C, C-2 D-NH), 123.2 (2 C, C-2, -6 *p*-NPh), 126.9 (2 C, C-3, -5 *p*-NPh), 146.7 (2 C, C-1, -4 *p*-NPh), 165.2, 165.7 (2 C, C-4, -6 *s*-triazine), 167.7, 168.2 (1 C, C-2 *s*-triazine) ppm. MS (CI, isobutane), *m/z* (rel. int. %) 514 [M$^+$+H] (25) 278 (5), 178 (100), 140 (18), 116(11), 104 (21), 87 (18). [α]$_D^{25}$=+157 (0.5 % DMSO).

Typical procedure for the synthesis of compounds I-4a-c and II-5a-c. Preparation of compound I-4c

At room temperature and with vigorous stirring, anh. K$_2$CO$_3$ (0.604 g, 4.377 mmol) was suspended in a solution obtained by dissolving anh. piperazine (1.504 g, 17.508 mmol) in anh. THF (125 mL). To this suspension, chlorodiamino-*s*-triazine I-2c (2.000 g, 4.377 mmol) was added portionwise (5 equal portions, 0.400 g I-2c / portion, each 2 hours). After each addition and within 2 hours, TLC monitoring indicated the completion of reaction as follows: total consumption of I-2c (eluent toluene : isopropanol = 2:1 v/v, R_f = 0.8, visualisation in UV 254 nm) and formation of I-4c (eluent ethanol : aq. NH$_3$ 25% = 9:1 v/v, R_f = 0.76, double visualisation: UV 254 nm then I$_2$ bath). After addition, the reaction mixture was stirred at room temperature for 24 hrs. Minerals were filtered off and well-washed with anh. THF. The combined THF solution was evaporated under reduced pressure to yield 2.900 g crude material which was separated by column chromatography on silica gel (eluent ethanol : aq. NH$_3$ 25% = 9:1 v/v, R_f = 0.76, double visualisation: UV 254 nm then I$_2$ bath). The isolated I-4c, 2.267 g was taken with anh. THF (2 mL) then diethyl ether was added and the resulted fine yellow suspension was stirred at room temperature for 1 hr. After cooling at -20 °C for 12 hrs., filtering off, washing with could diethyl ether and drying 1.910 g pure I-4c were obtained (86% yield with respect to I-2c).

1-{6-{[1,3-Dihydroxy-2-(methyl)prop-2-yl]amino}-4-{[(4*S*,5*S*)-4-(4-nitro-phenyl)-1,3-dioxan-5-yl]amino}-*s*-triazin-2yl}-piperazine I-4a (80 %) yellowish crystalline powder, mp 123-133 °C (flash column chromatography, eluent ethanol : aq. NH_3 25% 9:1 v/v). [Found: C 51.55, H 5.80, N 23.03; $C_{21}H_{30}N_8O_6$ (490.23) requires: C 51.42, H 6.16, N 22.84%]. R_f 0.77 (90% ethanol/aq. NH_3 25%). IR ν_{max} (KBr) 3400 (s), 2922 (m), 2855 (s), 1548(s), 1501 (s), 1444 (s), 1346 (s), 1274 (m), 1174 (m), 1106 (m), 1056 (m), 1027 (m), 875 (w), 852 (w), 810 (m), 744 (w),711 (m), 583 (w) cm^{-1}. ^{1}H NMR (500 MHz, [D$_6$]DMSO, 363 K): δ_H 1.21 (3 H, s, CH$_3$), 2.65 (4 H, t, $^3J_{H,H}$=5.0 Hz, H-3, -5 Piperazine), 3.47 (2 H, d, $^2J_{H,H}$=10.5 Hz, C*H$_2$*OH), 3.48 (4 H, t, $^3J_{H,H}$=5.0 Hz, H-2, -6 Piperazine), 3.56 (2 H, d, $^2J_{H,H}$=10.5 Hz, C*H$_2$*OH), 4.02 (1 H, d, $^2J_{H,H}$=11.0 Hz, H-6-a D-NH), 4.11 (1 H, dd, $^3J_{H,H}$=1.5 Hz, $^2J_{H,H}$ = 11.5 Hz, H-6-e D-NH), 4.41 (1 H, d, $^3J_{H,H}$ = 9.0 Hz, H-5-e D-NH), 4.54 (3 H, bs, OH, Pip-N*H*), 5.00 (1 H, d, $^2J_{H,H}$=6.5 Hz, H-2-a D-NH), 5.225 (1 H, s, H-4-a D-NH), 5.230 (1 H,d, $^2J_{H,H}$=5.5 Hz, H-2-e D-NH), 5.43 (1 H, s SER-N*H*), 5.49 (1 H, d, $^3J_{H,H}$=9.5 Hz D-N*H*), 7.62 (2 H, d, $^3J_{H,H}$=9.0 Hz, H-2, -6 *p*-NPh), 8.11 (2 H, d, $^3J_{H,H}$=8.5 Hz, H-3, -5 *p*-NPh) ppm; ^{13}C NMR (125 MHz, [D$_6$]DMSO, 298 K): δ_C 19.2 (1 C, CH$_3$), 44.2, 44.3 (2 C, C-3, -5 Piperazine), 45.8, 45.9, 46.0 (2 C, C-2, -6 Piperazine), 48.8, 48.9 (1 C, C-5 D-NH), 57.9 (1 C, C-2 SER-NH), 64.6, 64.9 (2 C, CH$_2$OH), 70.5, 70.6, 70.8, 71.0 (1 C, C-6 D-NH), 78.5, 78.6, 78.9, 79.0 (1 C, C-4 D-NH), 93.9, 94.0, 94.1 (1 C, C-2 D-NH), 123.2, 123.5 (2 C, C-2, -6 *p*-NPh), 127.4, 127.6 (2 C, C-3, -5 *p*-NPh), 147.0, 147.1 (2 C, C-1, -4 *p*-NPh), 164.17, 164.22, 164.3, 164.5 (1 C, C-2 *s*-triazine), 165.3, 165,4 (2 C, C-4, -6 *s*-triazine) ppm. MS (ESI+), *m/z* (rel. int. %) 491.2 [M$^+$+H] (100), 403.2 (22), 208.0 (29). [α]$_D^{25}$=+28 (0.5 % DMSO).

1-{6-{[1-Hydroxy-2-(hydroxymethyl)but-2-yl]amino}-4-{[(4*S*,5*S*)-4-(4-ni- trophenyl)-1,3-dioxan-5-yl]amino}-*s*-triazin-2-yl}-piperazine I-4b (84 %) yellowish crystalline powder, mp 125-130 °C (flash column chromatography, eluent ethanol : aq. NH_3 25% 9:1 v/v). [Found: C 51.99, H 6.22, N 21.95; $C_{22}H_{32}N_8O_6$ (504.24) requires: C 52.37, H 6.39, N 22.21%]. R_f 0.66 (90% ethanol/aq. NH_3 25%). IR ν_{max} (KBr) 3401 (s), 2966 (m), 2856 (s), 1552 (s), 1500 (s), 1445 (s), 1346 (s), 1174 (m), 1106 (m), 1061 (m), 1026 (m), 873 (w), 852 (w), 809 (m), 744 (w), 710 (w), 583 (w) cm^{-1}. ^{1}H NMR (500 MHz, [D$_6$]DMSO, 363 K): δ_H 0.74 (3 H, t, $^3J_{H,H}$ = 7.3 Hz, CH$_3$), 1.73, 1.74 (2 H, 2×q, $^3J_{H,H}$=7.5 Hz, C*H$_2$*CH$_3$), 2.69 (4 H, t, $^3J_{H,H}$=5.0 Hz, H-3, -5 Piperazine), 3.49-3.51 (6 H, m, C*H$_2$*OH, H-2, -6 Piperazine), 3.56 (2 H, d, $^2J_{H,H}$=10.5 Hz, C*H$_2$*OH), 4.02 (1 H, d, $^2J_{H,H}$=11.5 Hz, H-6-a D-NH), 4.12 (1 H, dd, $^3J_{H,H}$=1.5 Hz, $^2J_{H,H}$ = 11.5 Hz, H-6-e D-NH), 4.41 (1 H, d, $^3J_{H,H}$ = 9.0 Hz, H-5-e D-NH), 4.54 (3 H, bs, OH, Pip-N*H*), 5.00 (1 H, d, $^2J_{H,H}$=6.5 Hz, H-2-a D-NH), 5.23 (1 H, d, $^2J_{H,H}$=5.5 Hz, H-2-e D-NH), 5.24 (1 H, s, H-4a D-NH), 5.35 (1 H, s SER-N*H*), 5.52 (1 H, d, $^3J_{H,H}$=9.5 Hz D-N*H*), 7.62 (2 H, d, $^3J_{H,H}$=8.0 Hz, H-2, -6 *p*-NPh), 8.10 (2 H, d, $^3J_{H,H}$=8.5 Hz, H-3, -5 *p*-NPh) ppm; ^{13}C NMR (125 MHz, [D$_6$]DMSO, 298 K): δ_C 8.1 (1 C, CH$_3$), 23.4, 23.46, 23.54 (1 C, C*H$_2$*CH$_3$), 43.5, 44.0 (2 C, C-3, -5 Piperazine),

45.45, 45.55, 45.64, 45.7 (2 C, C-2, -6 Piperazine), 48.8, 48,9 (1 C, C-5 D-NH), 60.2 (1 C, C-2 SER-NH), 62.5, 62.9 (2 C, CH_2OH), 70.5, 70,6 70.8, 70.9 (1 C, C-6 D-NH), 78.5, 78.95, 79.04 (1 C, C-4 D-NH), 93.9, 94.0 94.1 (1 C, C-2 D-NH), 123.2, 123.5 (2 C, C-2, -6 p-NPh), 127.3, 127.6 (2 C, C-3, -5 p-NPh), 147.0, 147.1 (2 C, C-1, -4 p-NPh), 164.5 (1 C, C-2 s-triazine), 165.4, 165.6 (2 C, C-4, -6 s-triazine) ppm. MS (ESI+), m/z (rel. int. %) 505.3 [M^++H] (100), 403.2 (25), 224.0 (12), 208.0 (37). $[\alpha]_D^{25}$=+42 (0.5 % DMSO).

1-{6-{[1,3-Dihydroxy-2-(hydroxymethyl)prop-2-yl]amino}-4-[(4S,5S)-4-(4-nitrophenyl)-1,3-dioxan-5-yl]amino}-s-triazin-2-yl}-iperazine I-4c (86 %) yellowish crystalline powder, mp 146-157 °C (flash column chromatography, eluent ethanol : aq. NH_3 25% 9:1 v/v). [Found: C 50.11, H 5.88, N 21.90; $C_{21}H_{30}N_8O_7$ (506.22) requires: C 49.80, H 5.97, N 22.12%]. R_f 0.71 (90% ethanol/aq. NH_3 25%). IR ν_{max} (KBr) 3392 (m), 2943 (m), 2856 (m), 1549 (s), 1504 (s), 1446 (m), 1346 (s), 1273 (m), 1174 (m), 1105 (m), 1025 (m), 872 (w), 852 (w), 809 (m), 744 (w), 711 (w), 584 (w) cm^{-1}. ^{1}H NMR (500 MHz, [D_6]DMSO, 353 K): δ_H 2.65 (3 H, t, $^3J_{H,H}$=4.8 Hz, H-3, -5 Piperazine), 2.68 (3 H, s, OH), 3.46 (4 H, t, $^3J_{H,H}$=5.0 Hz, H-2, -6 Piperazine), 3.62 (6 H, s, CH_2OH), 4.02 (1 H, d, $^2J_{H,H}$=11.0 Hz, H-6-a D-NH), 4.11 (1 H, $^2J_{H,H}$=11.0 Hz, H-6-e D-NH), 4.40 (1 H, bd, $^3J_{H,H}$ = 7.5 Hz, H-5-e D-NH), 5.00 (1 H, d, $^2J_{H,H}$=6.5 Hz, H-2-a D-NH), 5.22 (1 H, s, H-4-a D-NH), 5.23 (1 H, d, $^2J_{H,H}$=5.5 Hz, H-2-e D-NH), 5.43 (1 H, s SER-NH), 5.61 (1 H, d, $^3J_{H,H}$=9.5 Hz D-NH), 7.63 (2 H, d, $^3J_{H,H}$=9.0 Hz, H-2, -6 p-NPh), 8.12 (2 H, d, $^3J_{H,H}$=8.0 Hz, H-3, -5 p-NPh) ppm; ^{13}QC NMR (75 MHz, [D_6]DMSO, 298 K): δ_C 43.9 (2 C, C-3, -5 Piperazine), 45.4 (3 C, C-2, -6 Piperazine), 48.8 (1 C, C-5 D-NH), 61.0, 61.3 (4 C, C-2, CH_2OH, SER-NH), 70.5, 70.7 (1 C, C-6 D-NH), 78.4, 78,8 (1 C, C-4 D-NH), 93.9 (1 C, C-2 D-NH), 123,1, 123,3 (2 C, C-2, -6 p-NPh), 127.1, 127.4 (2 C, C-3, -5 p-NPh), 146.9 (2 C, C-1, -4 p-NPh), 164.2 (1 C, C-2 s-triazine), 165.0, 165.2, 165.4, 165.5 (2 C, C-4, -6 s-triazine) ppm. MS (DCI positive, 200 eV, isobutane), m/z (rel. int. %) 563 [M^++HC(CH$_3$)$_3$-2 H] (9), 507 [M^++H] (100), 489 (10), 477 (10), 404 (10),282 (5), 225 (10), 115 (8), 104 (20), 87 (75). $[\alpha]_D^{25}$=+24 (0.5 % DMSO).

1-{6-{[1,3-Dihydroxy-2-(methyl)prop-2-yl]amino}-4-{[(2R,4S,5S)-5-(dimethylamino)-4-(4-nitrophenyl)-1,3-dioxan-2-yl]methylamino}-s-triazin-2-yl}-piperazine II-5a (71 %) yellowish crystalline powder, mp 118-123 °C (flash column chromatography, eluent ethanol : aq. NH_3 25% 9:1 v/v). [Found: C 52.88, H 7.07, N 22.85; $C_{24}H_{37}N_9O_6$ (547.29) requires: C 52.64, H 6.81, N 23.02%]. R_f 0.57 (90% ethanol/aq. NH_3 25%). IR ν_{max} (KBr) 3295 (s), 2932 (s), 2860 (s), 1670 (m), 1548 (s), 1516 (s), 1444 (s), 1348 (s), 1296 (m), 1151 (m), 1113 (m), 1055 (m), 1016 (m), 852 (m), 809 (m), 710 (m), 668 (w), 572 (w) cm^{-1}. ^{1}H NMR (500 MHz, [D_6]DMSO, 353 K): δ_H 1.27 (3 H, s, CH$_3$), 2.23 [6 H, s, N(CH$_3$)$_2$], 2.69 (4 H, t, $^3J_{H,H}$=4.8 Hz, H-3, -5 Piperazine), 2.86 (1 H, dd as t, $^3J_{H,H}$=2.8 Hz, H-5-e D-NH), 3.496 (2 H, dd as t, $^3J_{H,H}$=4.5 Hz, CH_2NH), 3.504 (2 H, d, $^2J_{H,H}$=10.0 Hz, CH_2OH), 3.58 (4 H, t, $^3J_{H,H}$=5.3 Hz, H-2, -6 Piperazine), 3.60

56

(2 H, d, $^2J_{H,H}$=10.5 Hz, CH_2OH), 3.96 (1 H, dd, $^2J_{H,H}$=12.5 Hz, $^3J_{H,H}$=3.0 Hz, H-6-a D-NH), 4.46 (1 H, d, $^2J_{H,H}$=12.0 Hz H-6-e), 4.66 (3 H, bs, OH, Pip-NH), 4.99 (1 H, dd as t, $^3J_{H,H}$=4.8 Hz, H-2-a D-NH), 5.18 (1 H, d, $^3J_{H,H}$=3.5 Hz, H-4-a D-NH), 5.55 (1 H, s SER-NH), 6.27 (1 H, bdd as bt, $^3J_{H,H}$=5.5 Hz D-NH), 7.64 (2 H, d, $^3J_{H,H}$=8.5 Hz, H-2, -6 p-NPh), 8.16 (2 H, d, $^3J_{H,H}$=8.5 Hz, H-3, -5 p-NPh) ppm; ^{13}C NMR (125 MHz, [D$_6$]DMSO, 298 K): δ_C 19.3 (1 C, CH$_3$), 43.8 [2 C, N(CH$_3$)$_2$], 44.4 (2 C, C-3, -5 Piperazine), 45.87, 45.94, 46.03 (3 C, C-2, -6 Piperazine, CH$_2$NH), 58.0 (1 C, C-2 SER-NH), 58.5, 58.6 (1 C, C-5 D-NH), 64.4, 64.5, 64.6 (2 C, CH$_2$OH), 64.88, 64.92, 65.3 (1 C, C-6 D-NH), 80.1, 80.2 (1 C, C-4 D-NH), 99.8, 99.9 (1 C, C-2 D-NH), 123.19, 123.24 (2 C, C-2, -6 p-NPh), 127.0 (2 C, C-3, -5 p-NPh), 146.7 (1 C, C-1 p-NPh), 148.9 (1 C, C-4 p-NPh), 164.7 (1 C, C-2 s-triazine), 165.9 (2 C, C-4, -6 s-triazine) ppm. MS (ESI+), *m/z* (rel. Int. %) 548.3 [M$^+$+H] (100), 543.3 (8), 295.2 (10), 217.1 (15), 154.1 (10), 120.1 (3). [α]$_D^{25}$= +145 (0.5 % DMSO).

1-{6-{[1-Hydroxy-2-(hydroxymethyl)but-2-yl]amino}-4-{[(2*R*,4*S*,5*S*)-5-(dimethylamino)-4-(4-nitrophenyl)-1,3-dioxan-2-yl]methylamino}-*s*-tri-azin-2-yl}-piperazine II-5b (67 %) yellowish crystalline powder, mp 112-118 °C (flash column chromatography, eluent ethanol : aq. NH$_3$ 25% 9:1 v/v). [Found: C 53.55, H 7.17, N 22.22; C$_{25}$H$_{39}$N$_9$O$_6$ (561.30) requires: C 53.46, H 7.00, N 22.45%]. R_f 0.70 (90% ethanol/aq. NH$_3$ 25%). IR ν$_{max}$ (KBr) 3392 (s), 2939 (s), 2856 (s), 1553 (s), 1514 (s), 1446 (s), 1348 (s), 1298 (m), 1151 (m), 1113 (m), 1055 (s), 1011 (m), 852 (m), 710 (m), 573 (w) cm^{-1}. ^{1}H NMR (500 MHz, [D$_6$]DMSO, 353 K): δ_H 0.79, 0.80 (3 H, 2×t, $^3J_{H,H}$ = 7.5 Hz, CH$_3$), 1.80, 1.81 (2 H, 2×q, $^3J_{H,H}$=7.5 Hz, CH_2CH$_3$), 2.23 [6 H, s, N(CH$_3$)$_2$], 2.70 (4 H, t, $^3J_{H,H}$=4.8 Hz, H-3, -5 Piperazine), 2.86 (1 H, dd as t, $^3J_{H,H}$=2.3 Hz, H-5-e D-NH), 3.50 (2 H, dd as t, $^3J_{H,H}$=4.5 Hz, CH_2NH), 3.51-3.62 (8 H, m, CH_2OH, H-2, -6 Piperazine), 3.96 (1 H, dd, $^2J_{H,H}$=12.5 Hz, $^3J_{H,H}$=3.0 Hz, H-6-a D-NH), 4.46 (1 H, d, $^2J_{H,H}$=12.5 Hz, H-6-e D-NH), 4.68 (3 H, bs, OH, Pip-NH), 4.99 (1 H, dd as t, $^3J_{H,H}$ = 4.8 Hz, H-2-a D-NH), 5.18 (1 H, d, $^3J_{H,H}$=3.0 Hz, H-4-a D-NH), 5.46 (1 H, s SER-NH), 6.29 (1 H, bdd as bt, $^3J_{H,H}$=5.5 Hz D-NH), 7.64 (2 H, d, $^3J_{H,H}$=8.5 Hz, H-2, -6 p-NPh), 8.16 (2 H, d, $^3J_{H,H}$=9.0 Hz, H-3, -5 p-NPh) ppm; ^{13}C NMR (125 MHz, [D$_6$]DMSO, 298 K): δ_C 8.3 (1 C, CH$_3$), 23.5, 23.6 (1 C, CH_2CH$_3$), 43.8 [2 C, N(CH$_3$)$_2$], 44.3 (2 C, C-3, -5 Piperazine), 45.8, (3 C, C-2, -6 Piperazine, CH$_2$NH), 58.6 (1 C, C-5 D-NH), 60.4 (1 C, C-2 SER-NH), 62.8, 63.4 (1 C, C-6 D-NH), 64.5 (2 C, CH$_2$OH), 80.1, 80.3 (1 C, C-4 D-NH), 99.8, 100.1 (1 C, C-2 D-NH), 123.2 (2 C, C-2, -6 p-NPh), 127.0 (2 C, C-3, -5 p-NPh), 146.7 (1 C, C-1 p-NPh), 148.93, 148.95 (1 C, C-4 p-NPh), 164.7 (1 C, C-2 s-triazine), 165.9, 166.0 (2 C, C-4, -6 s-triazine) ppm. MS (ESI+), *m/z* (rel. Int. %) 562.3 [M$^+$+H] (100), 548.3 (10), 302.2 (15), 281.7 (25), 208.0 (13), 143.1 (7). [α]$_D^{25}$=+136 (0.5 % DMSO).

1-{6-{[1,3-Dihydroxy-2-(hydroxymethyl)prop-2-yl]amino}-4-{[(2*R*,4*S*,5*S*)-5-(dimethylamino)-4-(4-nitrophenyl)-1,3-dioxan-2-yl]methylamino}-*s*-triazin-2-yl}-piperazine II-5c (81 %) yellowish crystalline powder, mp 140-145 °C (flash

column chromatography, eluent ethanol : aq. NH_3 25% 9:1 v/v). [Found: C 50.98, H 6.77, N 22.55; $C_{24}H_{37}N_9O_7$ (563.28) requires: C 51.15, H 6.62, N 22.37%]. R_f 0.57 (90% ethanol/aq. NH_3 25%). IR ν_{max} (KBr) 3298 (s), 2940 (s), 2858 (s), 1551 (s), 1515 (s), 1446 (s), 1347 (s), 1297 (m), 1151 (m), 1112 (m), 1054 (m), 1015 (m), 852 (m), 809 (m), 710 (w), 578 (w) cm^{-1}. ^{1}H NMR (500 MHz, [D_6]DMSO, 353 K): δ_H 2.23 [6 H, s, $N(CH_3)_2$], 2.69 (4 H, t, $^3J_{H,H}$=4.8 Hz, H-3, -5 Piperazine), 2.86 (1 H, dd as t, $^3J_{H,H}$=2.8 Hz H-5-e.), 3.50 (2 H, dd as t, $^3J_{H,H}$=4.8 Hz, CH_2NH), 3.57 (4 H, t, $^3J_{H,H}$=4.3 Hz, H-2, -6 Piperazine), 3.66 (6 H, s, CH_2OH), 3.97 (1 H, dd, $^2J_{H,H}$=12.5 Hz, $^3J_{H,H}$=3.0 Hz, H-6-a D-NH), 4.46 (1 H, d, $^2J_{H,H}$=12.5 Hz H-6-e D-NH), 4.65 (3 H, bs, OH), 5.00 (1 H, dd as t, $^3J_{H,H}$=4.5 Hz, H-2-a D-NH), 5.19 (1 H, d, $^3J_{H,H}$=3.0 Hz, H-4-a D-NH), 5.53 (1 H, s SER-NH), 6.38 (1 H, bs D-NH), 7.64 (2 H, d, $^3J_{H,H}$=8.5 Hz, H-2, -6 p-NPh), 8.17 (2 H, d, $^3J_{H,H}$=8.5 Hz, H-3, -5 p-NPh) ppm; ^{13}C NMR (75 MHz, [D_6]DMSO, 298 K): δ_C 43.7 [2 C, $N(CH_3)_2$], 44.3 (2 C, C-3, -5 Piperazine), 45.8, 46.1, 46.3 (3 C, C-2, -6, Piperazine, CH_2NH), 58.4 (1 C, C-5 D-NH), 61.3, 61.4 (1 C, C-2 SER-NH), 64.3 (3 C, CH_2OH), 65.3 (1 C, C-6 D-NH), 80.0 (1 C, C-4 D-NH), 99.6 (1 C, C-2 D-NH), 123.1 (2 C, C-2, -6 p-NPh), 126.9 (2 C, C-3, -5 p-NPh), 146.5 (1 C, C-1 p-NPh), 148.8 (1 C, C-4 p-NPh), 164.4 (1 C, C-2 s-triazine), 165.8 (2 C, C-4, -6 s-triazine) ppm. MS (DCI positive, 200 eV, isobutane), m/z (rel. Int. %) 620 [M^++$HC(CH_3)_3$-2 H] (5), 564 [M^++1] (65), 449 (28), 380 (55), 300 (10), 282 (15), 221 (5), 178 (100), 165 (25), 148 (10), 104 (45), 87 (35). $[\alpha]_D^{25}$=+137 (0.5 % DMSO).

ACKNOWLEDGMENTS

REFERENCES

1. a) H. Friebolin, *Basic One- and Two Dimensional NMR Spectroscopy*, VCH Verlagsgesellschaft / VCH: Weinheim / New York, NY, **1991**, pp. 53, 93, 131, 132, *263–291;* b) H. Kessler, *Angew. Chem. Int. Ed. Engl*., **1982***, 21,* 512.c) R.M. Silverstein, F.X. Webster, D.J. Kiemle, *Identification spectrométrique de composes organiques*, 2-ième edition, De Boeck & Larcier s.a., **2007**, p 150-155.

2. a) M. Pintea, M. Fazekas, P. Lameiras, I. Cadis, C. Berghian, I. Silaghi-Dumitrescu, F. Popa, C. Bele, N. Plé, M. Darabantu, *Tetrahedron*, **2008**, *64*, 8851; b) M. Fazekas, M. Pintea, P. Lameiras, A. Lesur, C. Berghian, I. Silaghi-Dumitrescu, N. Plé, M. Darabantu, *Eur. J. Org. Chem.*, **2008**, *14*, 2473 ; c) M. Darabantu, L. Boulitreau, I. Olosutean, P. Lameiras, C. Batiu, *Studia Universitatis Babes-Bolyai, Chemia*, **2008**, LIII, *4*, 21; d) M. Darabantu, G. Plé, S. Mager, C. Puscas, E. Cotora, *Tetrahedron*, **1997**, *53*, 1909.
3. a) W. Zhang, D.T. Nowlan, L.M. Thomson, M.W. Lackowski, E.E. Simanek, *J. Am. Chem. Soc.*, **2001**, *123*, 8914; b) M.B. Steffensen, E.E. Simanek, *Angew. Chem. Int. Ed.*, **2004**, *43*, 5178; c) J. Lim, E.E. Simanek, *Molecular Pharmaceutics*, **2005**, *2*, 273; d) G.R. Newcome, J. Gross, C.N. Moorefield, B.D. Woosley, *Chem. Commun.*, **1997**, 515; e) A.P. Umali, E.E. Simanek, *Org. Lett.*, **2003**, *5*, 1245; f) H.T. Chen, M.F. Neerman, A.R. Parrish, E.E. Simanek, *J. Am. Chem. Soc.*, **2004**, *126*, 10044; g) L.–L. Lai, L.–Y. Wang, C.-H. Lee, C.Y. Lin, K.–L. Cheng, *Org. Lett.*, **2006**, *8*, 1541; h) E.J. Acosta, S.O. Gonzalez, E.E. Simanek, *J. Polym. Sci. Part A.*, **2005**, *43*, 168; i) W. Zhang, S.O. Gonzalez, E.E. Simanek, *Macromolecules*, **2002**, *35*, 9015; j) E. Hollink, E.E. Simanek, *Org. Lett.*, **2006**, *8*, 2293; k) M.F. Neerman, W. Zhang, A.R. Parrish, E.E. Simanek, *Int. J. Pharm.*, **2004**, *281*, 129; l) L.–L. Lai, C.–H. Lee, L.–Y. Wang, K.L. Cheng, H.–F. Hsu, *J. Org. Chem.,* **2008**, 73, 485.
4. a) R.C. Galesworthy, R.M. Show, B.C. Smith, *J. Chem. Soc.*, **1962**, 1507; b) R.M. Show, P.J. Ward, *J. Chem. Soc.*, **1965**, 5446; c) B.I. Stepanov, G.I. Migachev, E.F. Ilina, *Zh. Obshch. Khim.*, **1967**, *37*, 2140; d) R.M. Show, P.J. Ward, *J. Chem. Soc.*, **1968**, *12*, 1431; e) G.I. Migachev, E.F. Ilina, B.I. Stepanov, *Zh. Obsch. Khim.*, **1968**, *38*, 1370; f) R.M. Show, P.J. Ward, *J. Chem. Soc.*, **1969**, *6*, 596; g) G. Boncic-Caricic, Z.D. Tadic, M.D. Muskatirovic, *J. Serb. Chem. Soc.*, **1994**, *59*, 929; h) N.D. Reddy, A.J. Elias, N. Ashwani-Vij, *J. Chem. Research (S)*, **1998**, *9*, 504, *J. Chem. Research (M)* **1998**, 2033; i) B. Kolesinska, Z.J. Kaminski, *Tetrahedron*, **2009**, *65*, 3573.
5. a) T. Drakenberg, S. Forsen, *Chem. Commun.*, **1971**, *21*, 1404; b) S.S. Mirvish, P. Gannett, D.M. Babcook, D. Williamson, S.C. Chen, D.D. Weisenburger, *J. Agric. Food Chem.*, **1991**, *39*, 1205.
6. a) I. Willner, J. Rosengaus, Y. Eichen, *J. Phys. Org. Chem.*, **1993**, *6*, 29; b) A.R. Katritzky, I. Ghiviriga, D.C. Oniciu, A. Barkock, *J. Chem. Soc., Perkin Trans.*, *2* **1995**, *4*, 785; c) A.R. Katritzky, I. Ghiviriga, P.G. Steel, D.C. Oniciu, *J. Chem. Soc., Perkin Trans., 2* **1996**, *3*, 443; d) I. Ghiviriga, D.C. Oniciu, *Chem. Commun.*, **2002**, *22*, 2718; e) M. Amm, N. Platzer, J. Guilhem, J.P. Bouchet, J.P. Volland, *Magn. Reson. Chem.*, **1998**, *36*, 587; f) M. Amm, N. Platzer, J.P. Bouchet, J.P. Volland, *Magn. Reson. Chem.*, **2001**, *39*, 77; g) H.E. Birkett, R.K. Harris, P. Hodgkinson, K. Carr, M.H. Charlton, J.C. Cherryman, A.M. Chippendale, R.P. Glover, *Magn. Reson. Chem.*, **2000**, *38*, 504; h) H.E. Birkett, J.C. Cherryman, A.M. Chippendale, J.O.S. Evans, R.K. Harris, M. James, I.J. King, G.Mc. Pherson, *Magn. Reson. Chem.*, **2003**, *41*, 324.
7. a) E.L. Eliel, H.S. Wilen, *Stereochemistry of the Organic Compounds*, John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore **1994**, pp. 21, 465, 466, 503-507, 696, 697, 1191, 1200, 1206; b) M.J.O. Anteunis, D. Tavernier, F. Borremans, *Heterocycles*, **1976**, *4*, 293.
8. X.C. Moreno, E.E. Simanek, *Macromolecules*, **2008**, *41*, 4108.

THE RELATIONSHIP BETWEEN ENERGY CALCULATIONS AND BOILING POINTS OF N-ALKANES

LORENTZ JÄNTSCHI[a], SORANA D. BOLBOACĂ[a,b]

ABSTRACT. The relationship between energy calculations and boiling points was studied on a set of fourteen n-alkanes. The correlation analysis clearly showed that the best relationship is not linear. The regression analysis showed that a dose-response logistic function provided a very good agreement between the boiling points of alkanes and their heat of formation.

Keywords: regression, correlation, alkanes, boiling point, energy

INTRODUCTION

Boiling point, the temperature at which the vapor pressure of the liquid equals the environmental pressure surrounding the liquid [1], of organic compounds is an important property since it can provide information about other physical properties and structural characteristics [2]. Molecules with strong intermolecular forces are known to have higher boiling points [2].

The boiling point of alkanes, chemical structures with a C_nH_{2n+2} generic formula, increases with the chain length (number of carbon atoms).

The relationship between the boiling points of alkanes and other properties or descriptors have previously been studied using simple or multiple linear regression models [3-5] or non-linear models [6]. Since the boiling point of alkanes is determined by their molecular weight, this property shows a linear relationship with the size of the molecules [7]. Kozioł obtained, on a set of fourteen n-alkanes, a non-linear model with five descriptors having a determination coefficient of 0.9993 [6]. Moreover, simple exponential models estimated the critical temperature, pressure, and volume of alkanes as function of the normal boiling point and molecular weight [8].

The present study is aimed to carry out correlation and regression analyses in order to establish the relationships between the calculated energy and the boiling points of n-alkanes (an "easy to predict" property).

---

[a] *Technical University of Cluj-Napoca, 103-105 Muncii Bvd., RO-400641 Cluj-Napoca, Romania, lori@academicdirect.org*
[b] *"Iuliu Hațieganu" University of Medicine and Pharmacy Cluj-Napoca, 13 Emil Isac, RO-400023 Cluj-Napoca, Romania, sbolboaca@umfcluj.ro*

RESULTS AND DISCUSSION

The results of the correlation analyses are presented in Table 1. The dipole moment property was excluded from further analyses since the Pearson correlation coefficient was of -0.0391. The analysis of the obtained correlation coefficients revealed that Spearman and Gamma correlation coefficients had higher values compared to the Pearson correlation coefficients.

Table 1. Results of correlation analysis

X (Y= boiling point)	r (p)	ρ (p)	Γ (p)
heat-of-formation	0.9515958 ($1.67 \cdot 10^{-7}$)	1[*]	1[*]
scf-binding-energy	0.9499073 ($2.05 \cdot 10^{-7}$)	1[*]	1[*]
total-energy	0.9498675 ($2.06 \cdot 10^{-7}$)	1[*]	1[*]
scf-atom-energy	0.9498641 ($2.06 \cdot 10^{-7}$)	1[*]	1[*]
scf-electronic-energy	0.9060543 ($8.09 \cdot 10^{-6}$)	1[*]	1[*]
scf-core-energy	0.8992529 ($1.21 \cdot 10^{-5}$)	1[*]	1[*]
dipole-moment	-0.0391090 (0.8943)	0.0681 (0.8094)	0.0989 (0.9618)

Correlation coefficients: r = Person; ρ = Spearman; Γ = Gamma
[*] $p < 10^{-7}$;

The 0.9515958 value of the Pearson correlation coefficient revealed that the linear relationship with the heat of formation was able to explain almost 91% of boiling points variation of the studied n-alkanes, which is a good estimation. Since the Spearman correlation coefficient was equal to the Gamma correlation coefficient and both of them were higher than the Pearson correlation coefficient, the relationship between boiling points and energy calculations could be non-linear.

Non-linear regression analysis was carried out in order to identify the type of relationship between the boiling points of alkanes and energy calculations. The best performing models, in terms of determination coefficients, F-value and coefficient significance proved to be of the *dose-response logistic function* type. The top three models in terms of the above-presented criteria are shown in Table 2.

The analysis of the results in Table 2 revealed that the best performing model, able to explain the boiling points of alkanes (as estimator) used the heat of formation (as predictor, H_F) through a dose-response logistic function. As it can be observed, a four-variable equation was able to fully predict the variation of boiling points as function of the heat of formation. The smallest difference between the determination coefficient and the adjusted determination coefficient was obtained using the first equation (boiling point as function of the heat of formation). The smallest value of the standard error was of 0.33°C and provided by the first equation (boiling point as function of the heat of formation). Note that the highest t-values associated to the coefficients and the smallest values of the standard errors were obtained when the boiling points were investigated as function of the heat of formation.

Table 2. Regression analysis results

Type		r^2	r^2_{adj}	F	C	Value [95%CI]	StErr	t
Y	X			(FitStErr)				
DoseRspLgstc $\hat{Y} = a_0 + a_1/(1+(x/a_2)^{\wedge}a_3)$								
B_P	H_F	0.999997	0.999996	1090130	a_0	1142.31 [1111.59; 1173.03]	13.78	82.85
				(0.32797)	a_1	-1435.64 [-1470.43; -1400.85]	15.61	-91.94
					a_2	-191.47 [-200.82; -182.11]	4.20	-45.59
					a_3	0.7518 [0.7386; 0.7656]	0.01	121.71
B_P	T_E	0.999864	0.999823	24478	a_0	-324.89 [-367.34; -282.43]	19.06	-17.05
				(2.18849)	a_1	1836.08 [1332.98; 2339.17]	225.80	8.13
					a_2	-179833.96 [-305299; -54369]	56313	-3.19
					a_3	-0.6190 [-0.7225; -0.5155]	0.046	-13.32
B_P	SBE	0.999857	0.999814	23351	a_0	-359.58 [-416.26; -302.91]	25.44	-14.14
				(2.24065)	a_1	1925.18 [1315.38; 2534.99]	273.70	7.03
					a_2	-14657.09 [-26730; -258]	5418.9	-2.70
					a_3	-0.5950 [-0.7137; -0.4764]	0.0532	-11.15

DoseRspLgstc = dose-response logistic function;
B_P = boiling point; H_F = heat-of-formation; T_E = total-energy; SBE = scf-binding-energy;
r^2 = determination coefficient; r^2_{adj} = adjusted determination coefficient; F = F-value;
C = coefficient; 95%CI = 95% coefficient confidence interval; StErr = standard error;
t = t-value

The graphical representation of the best performing model (B_P^= (1142.31±30.72)-(1435.6±34.79)/(1+(H_F/(-191.47±9.35)))$^{(0.7518±0.0132)}$)) is presented in Figure 1.
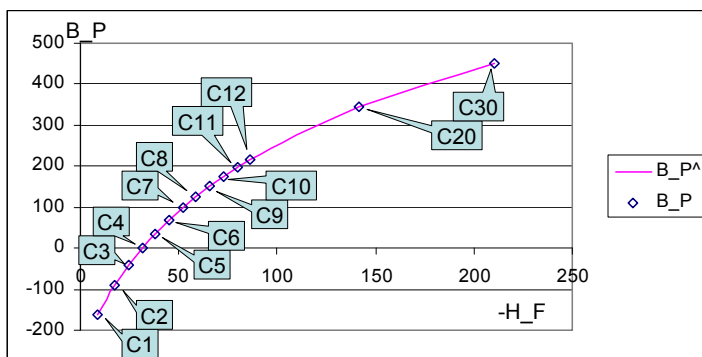
Figure 1. Boiling points of alkanes as heat of formation function

The analysis of Figure 1 revealed that the identified dose-response logistic function is the best one in estimating the relationship between the heat of formation and the boiling points of the studied n-alkanes. This statement is also supported by the value of the correlation coefficient associated to the model (see Table 2). A statistically significant linear relationship could also be identified between boiling points and the heat of formation, but this

relationship had lower performances compared to the best scoring dose-response logistic function (r^2 = 0.9062, F = 116, p = $1.6 \cdot 10^{-7}$, standard error of estimated = 52.44).

The estimated boiling points when the first equation was used (boiling point as function of the heat of formation), abbreviated as B_P^, and the measured boiling points, abbreviated as B_P, is graphically presented in Figure 2.
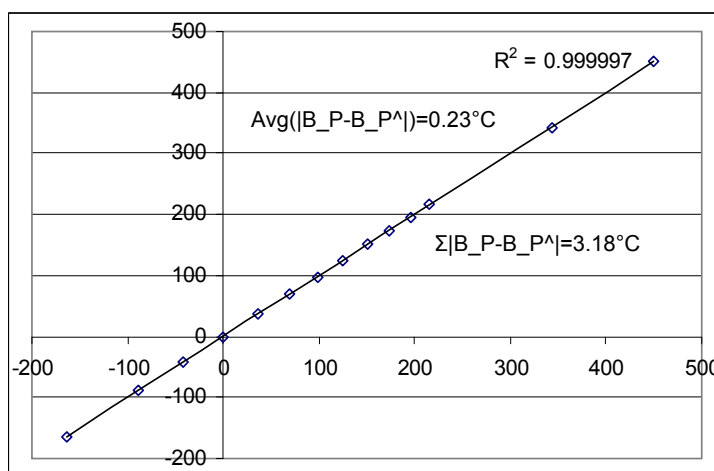
Figure 2. Estimated (horizontal) versus measured (vertical) boiling points using the dose-response logistic function

The validity and reliability of the best performing relationship obtained in the study on n-alkanes is supported by the smallest value of the absolute value of residuals (equal to 0.23°C) and by the sum of the absolute difference of residuals (equal to 3.18°C) (Figure 2). Moreover, the sum of residuals was 0.01°C while the squared sum of residuals was 1.08.

The objective of this research was met as soon as the best model able to estimate the boiling points of alkanes as functional dependence on energy calculations was identified. The value of the Person correlation coefficient, which proved to be smaller in comparison to the Spearman and Gamma correlation coefficients, determined the investigation of non-linear relationships even if the linear relationship was statistically significant. A dose response logistic function proved to better explain the boiling points as function of energy calculations for the studied n-alkanes when the molecules were prepared for analysis by applying the mm+ as molecular mechanics and the AM1 as semi-empirical method.

CONCLUSIONS

If ρ^2(Spearman), Γ^2(Gamma) $>r^2$(Pearson), the relationship between variables is not linear; non-linear relationships must always be checked. Thus, the best performing relationship between boiling points and the energy calculations of the investigated n-alkanes was expected not to be linear.

A functional dependence was identified between boiling points and the energy calculations of the investigated n-alkanes. This functional dependence proved to be a dose-response logistic function when mm+ molecular mechanics and AM1 semi-empirical methods were used to prepare the studied n-alkanes for analysis.

The following model was identified as the model with the highest performance:

B_P^ = $(1142.31\pm30.72)-(1435.6\pm34.79)/(1+(H\_F/(-191.47\pm9.35)))^{(0.7518\pm0.0132)}$,

where B_P^ is the estimated boiling point and H_F is the heat of formation. The validity of the model is supported by the small value of the standard error, the high F-value and the small p-value.

EXPERIMENTAL SECTION

Fourteen normal alkanes (C_1-C_{12}, C_{20}, C_{30}), chemical compounds consisting of carbon and hydrogen elements, were analyzed (see Table 3).

Table 3. Characteristics of alkanes: boiling point, dipole-moment, total-energy, atom-energy, binding-energy, core-energy, electronic-energy, and heat-of-formation

Name	Formula	B_P	D_M	T_E	SAE	SBE	SCE	SEE	H_F
Methane	CH_4	-164	$1.12\cdot10^{-6}$	-4225	-3837	-388	4619	-8844	-9
Ethane	C_2H_6	-89	$6.87\cdot10^{-7}$	-7821	-7149	-672	13638	-21459	-18
Propane	C_3H_8	-42	$4.28\cdot10^{-3}$	-11415	-10461	-954	26313	-37727	-24
Butane	C_4H_{10}	-0.5	$1.01\cdot10^{-7}$	-15008	-13773	-1236	41607	-56615	-31
Pentane	C_5H_{12}	36	$6.28\cdot10^{-3}$	-18602	-17084	-1518	59034	-77636	-38
Hexane	C_6H_{14}	69	$3.06\cdot10^{-7}$	-22196	-20396	-1800	78191	-100387	-45
Heptane	C_7H_{16}	98	$6.57\cdot10^{-3}$	-25790	-23708	-2082	98835	-124624	-52
Octane	C_8H_{18}	125	$1.52\cdot10^{-7}$	-29383	-27020	-2364	120757	-150141	-59
Nonane	C_9H_{20}	151	$6.65\cdot10^{-3}$	-32977	-30331	-2646	143819	-176796	-66
Decane	$C_{10}H_{22}$	174	$3.95\cdot10^{-7}$	-36571	-33643	-2928	167892	-204463	-73
Undecane	$C_{11}H_{24}$	196	$8.13\cdot10^{-3}$	-40165	-36955	-3210	192888	-233052	-80
Dodecane	$C_{12}H_{26}$	216	$1.35\cdot10^{-7}$	-43758	-40267	-3492	218724	-262482	-86
Eicosane	$C_{20}H_{42}$	343	$8.61\cdot10^{-7}$	-72508	-66760	-5748	449165	-521673	-142
Triacontane	$C_{30}H_{62}$	450	$1.59\cdot10^{-6}$	-108445	-99878	-8567	779447	-887893	-210

B_P = boiling point; D_M = dipole-moment; T_E = total-energy;
SAE = scf-atom-energy; SBE= scf-binding-energy; SCE = scf-core-energy;
SEE = scf-electronic-energy; H_F = heat-of-formation.

Eight properties of the above-mentioned alkanes were investigated: boiling point [°C] [9], total-energy (T_E) [kcal/mol], dipole-moment (D_M) [Debyes], scf-atom-energy (SAE) [kcal/mol], scf-binding-energy (SBE) [kcal/mol], scf-core-energy (SCE) [kcal/mol], scf-electronic-energy (SEE) [kcal/mol], and heat-of-formation (H_F) [kcal/mol]. Except for the boiling points, all the other properties were calculated with HyperChem v. 8.0 using the following criteria: optim-converged=true, molecular mechanics method: mm+ [10], and semi-empirical method: AM1 [11].

Correlation and regression analyses were carried out in order to meet the objective of the study. Pearson ("r") [12], Spearman ("ρ") [13] and Gamma ("Γ") [14] correlation coefficients were used to find the power and the sign of the relationship between boiling points and the investigated properties.

Regression analyses were carried out with the SlideWrite Plus software. The following possibilities of regression search were used:

- *Linear:* ▪ Linear Group; ▪ Exponential Group; ▪ Power Group; ▪ Polynomial Group.

- *Nonlinear:*
  - *Standard*: ▪ User-Defined (any function defined by the user); ▪ Exponential – $Y=a_0+a_1*\exp(-x/a_2)$; ▪ Power - $Y=a_0+a_1*x^{\wedge}a_2$.
  - *Transitional*: ▪ 1-Site Ligant – $Y=a_0*x/(a_1+x)$;
    ▪ Cumulative – $Y=a_0+a_1*0.5*(1+\mathrm{erf}((x-a_2)/\sqrt{(2)}*a_3))$;
    ▪ DoseRspLgstc - $Y=a_0+a_1/(1+(x/a_2)^{\wedge}a_3)$;
    ▪ Photosynthesis - $Y=a_0*a_1*x/(a_0+a_1*x)$;
    ▪ PH Activity – $Y=(a_0+a_1*10^{\wedge}(x-a_2))/(1+10^{\wedge}(x-a_2))$;
    ▪ Sigmoidal – $Y=a_0+a_1/(1+\exp(-(x-a_2)/a_3))$.
  - *Peak*: ▪ Erfc Peak, Gaussian – $Y=a_0+a_1*\exp(-0.5*((x-a_2/a_3)^2)$;
    ▪ Logistic Peak – $Y=a_0+a_1*4*(\exp(-(x-a_2)/a_3))/(1+\exp(-(x-a_2)/a_3))^2$;
    ▪ Log-Normal – $Y=a_0+a_1*\exp(-0.5*(\ln(x/a_2)/a_3)^2)$;
    ▪ Lorentzian – $Y=a_0+a_1/(1+((x-a_2)/a_3)^2)$.
  - *Waveform*: ▪ SineWave – $Y=a_0+a_1*\sin(2*pi*x/a_3+a_2)$;
    SineWaveSquared – $Y=a_0+a_1*(\sin(2*pi*x/a_3+a_2))^2$

- *User-Defined:* allows to define any equation with a maximum of 7 coefficients.

ACKNOWLEDGMENTS

REFERENCES

1. D.E. Goldberg, 3,000 Solved Problems in Chemistry (1st ed.). McGraw-Hill. ISBN 0-07-023684-4. Section 17.43, 1988, pp. 321.

2. M.R. Riazi, T.E. Daubert, *Hydrocarbon Processing,* **1980**, *59(3)*, 115.

3. E.S. Souza, C.A. Kuhnen, B.S. Junkes, R.A. Yunes, V.E.F. Heinzen, *Journal of Chemometrics,* **2010**, *24(3-4)*, 149.

4. D. Ciubotariu, M. Medeleanu, V. Vlaia, T. Olariu, C. Ciubotariu, D. Dragos, S. Corina, *Molecules,* **2004**, *9(12)*, 1053.

5. S.D. Bolboacă, L. Jäntschi, *International Journal of Pure and Applied Mathematics,* **2008**, *47(1)*, 23.

6. J. Kozioł, *Polish Journal of Chemistry,* **2009**, *83(12)*, 2173.

7. R.T. Morrison, R.N. Boyd. Organic Chemistry, 6th ed. (New Jersey: Prentice Hall, 1992.

8. F. Vejahati, M.B. Nikoo, S. Mokhatab, B.F. Towler, *Petroleum Science and Technology,* **2007**, 25(9), 1115.

9. C.E. Ophardt. *Virtual Chembook*, Elmhurst College, Elmhurst, IL, USA, 2003.

10. A. Hocquet, M. Langgard, *Journal of Molecular Modeling,* **1998**, *4*, 94.

11. M.J.S. Dewar, E.G. Zoebisch, E.F. Healy, J.J.P. Stewart, *Journal of the American Chemical Society,* **1985**, *107*, 3902.

12. K. Pearson, *Philosophical Magazine,* **1900**, *50*, 157.

13. C. Spearman, *American Journal of Psychology*, **1904**, *15*, 201.

14. L.A. Goodman, W.H. Kruskal, *Journal of the American Chemical Society,* **1963**, *58*, 310.

DIAGNOSTIC OF A QSPR MODEL: AQUEOUS SOLUBILITY OF DRUG-LIKE COMPOUNDS

SORANA D. BOLBOACĂ[a,b], LORENTZ JÄNTSCHI[b]

ABSTRACT. A diagnostic test for a qSPR (quantitative Structure-Property Relationship) model was carried out using a series of statistical indicators for correctly classifying compounds into actives and non-actives. A previously reported qSPR model, able to characterize the aqueous solubility of drug-like compounds, was used in this study. Eleven statistical indicators like those used in medical diagnostic tests were defined and applied on training, test and overall data sets. The associated 95% confidence interval under the binomial distribution assumption was also computed for each defined indicator in order to allow a correct interpretation. Similar results were obtained in the training and test sets with some exceptions. The prior probabilities of active and non-active compounds proved not to be significantly different in the training and test sets. However, the probability of classification as active compounds proved to be significantly smaller in the training set as compared to the test set (p = 0.0042). The total fraction of correctly classified compounds proved to be identical in the training and test sets as well as in the overall set. Nevertheless, the overall model and the model obtained in the test set show a higher ability to correctly assign the non-active compounds to the non-active class while the model obtained in the training set has a higher ability to correctly assign the active compounds to the active class.

Keywords: quantitative Structure-Property Relationships (qSPR), diagnostic parameters, 2×2 contingency table, solubility, drug-like compounds

INTRODUCTION

Quantitative structure-property relationships (qSPRs) procedures able to quantitatively correlate the chemical structure with a defined property [1], are widely used in drug design [2,3], drug classification [4,5] and screening [5,6].

A series of studies were drawn in order to establish the validation methods of a qSPR model [7,8], including the principle of parsimony, selection of the simplest model, cross-validation, Y scrambling and external predictability [9]. Various procedures for variable selection have been created [10-13] and statistical

---

[a] *"Iuliu Hațieganu" University of Medicine and Pharmacy Cluj-Napoca, 13 Emil Isac, RO-400023 Cluj-Napoca, Romania, sbolboaca@umfcluj.ro*

[b] *Technical University of Cluj-Napoca, 28 Memorandumului, RO-400114 Cluj-Napoca, Romania, lori@academicdirect.org*

analysis of molecular similarity matrices was developed in order to identify the best quantitative structure-activity relationships [14]. Reliability and accuracy have also been introduced for the validation of QSPR models [15,16]. The information criteria (Akaike's information criteria - AIC [17], corrected AIC [18], Schwarz (or Bayesian) Information Criterion – BIC, Amemiya Prediction Criterion – APC, and Hannan-Quinn Criterion - HQC) and Kubinyi's function [19, 20] are the parameters used to compare different qSPR/qSAR models [21-23].

The aim of this study was to carry out a diagnostic test on a qSPR (quantitative Structure-Property Relationships) model, by using a series of statistical indicators for correctly classifying compounds into actives and non-actives.

RESULTS AND DISCUSSION

Eleven statistical indicators were proposed as diagnostic parameters for qSPR models. The contingency tables used to calculate these parameters are presented in Table 1. The statistical indicators computed for the training, test and overall data sets are presented in Table 2 – 4.

Table 1. 2×2 contingency tables for the investigated qSPR model

Generic Table	Observed				Test Set	Observed		
Estimated	+	-	Σ		Estimated	+	-	Total
+	TP	FP			+	26	10	36
-	FN	TN			-	4	29	33
Σ			n		Total	30	39	69

Training Set	Observed				Overall	Observed		
Estimated	+	-	Total		Estimated	+	-	Total
+	28	7	35		+	54	22	76
-	12	48	60		-	11	77	88
Total	40	55	95		Total	65	99	164

+ = active class; - = non-active class;
Estimated = aqueous solubility estimated by Duchowitz's et al. qSPR model

The chi-squared test was applied on contingency tables in order to test the null hypotheses that the estimated class (active and non-active) is independent from the observed class (active and non-active). The value of the chi-squared statistics and associated significance level, presented at the bottom of Tables 2 - 4, supported the rejection of the null hypotheses that the estimated classification into active and non-active compounds is unrelated to the observed classification. These results sustain the ability of the qSPR model to classify compounds as actives and non-actives. The degree of association between the estimated and the observed classification of compounds proved to be a positive and moderate one, in all the investigated sets (training, test and overall set of studied compounds). The moderate association, expressed as the Φ contingency correlation coefficient, revealed that the reported qSPR [24] is not a perfect model.

70

Table 2. Statistical indicators for assessing the qSPR model: training set

Parameter (Abbreviation)	Value	95%CI
Concordance / Accuracy / Non-error Rate (CC/AC)	80.00	[71.07-87.02]
Error Rate (ER)	20.00	n.a.
Prior proportional probability of an active class	0.4211	[0.3254-0.5215]
Prior proportional probability of a non-active class	0.5789	n.a.
Sensitivity (Se)	70.00	[54.76-82.39]
False-negative rate (under-classification, FNR)	30.00	[17.61-45.24]
Specificity (Sp)	87.27	[76.39-93.96]
False-positive rate (over-classification, FPR)	12.73	[6.04-23.61]
Positive predictivity (PP)	80.00	[64.55-90.44]
Negative predictivity (NP)	80.00	[68.52-88.49]
Probability of classification		
- as active (PCA)	0.3684	[0.2766-0.4682]
- as non-active (PCIC)	0.6316	[0.5318-0.7234]
Probability of a wrong classification		
- as active compound (PWCA)	0.2000	[0.0956-0.3545]
- as non-active compound (PWCI)	0.2000	[0.1151-0.3148]
Odds Ratio (OR)	16.0000	[5.7090-45.0262]

95% CI = confidence interval at a significance level of 5%; n.a. = not available;
χ^2 = 30.2305 (p < 0.0001) (Chi-squared statistics); Contingency correlation coefficient Φ = 0.5641

Table 3. Statistical indicators for assessing the qSPR model: test set

Parameter (Abbreviation)	Value	95%CI
Concordance / Accuracy / Non-error Rate (CC/AC)	79.71	[69.04-87.79]
Error Rate (ER)	20.29	n.a.
Prior proportional probability of an active class	0.4348	[0.3225-0.5524]
Prior proportional probability of a non-active class	0.5652	n.a.
Sensitivity (Se)	86.67	[70.96-95.08]
False-negative rate (under-classification, FNR)	13.33	[4.92-29.04]
Specificity (Sp)	74.36	[59.21-85.91]
False-positive rate (over-classification, FPR)	25.64	[14.09-40.79]
Positive predictivity (PP)	72.22	[56.25-84.67]
Negative predictivity (NP)	87.88	[73.29-95.52]
Probability of classification		
- as active (PCA)	0.5217	[0.4050-0.6367]
- as non-active (PCIC)	0.4783	[0.3633-0.5950]
Probability of a wrong classification		
- as active compound (PWCA)	0.2778	[0.1533-0.4375]
- as non-active compound (PWCI)	0.1212	[0.0448-0.2671]
Odds Ratio (OR)	18.8500	[5.4919-64.5994]

95% CI = confidence interval at a significance level of 5%; n.a. = not available;
χ^2 = 22.9206 (p < 0.0001) (Chi-squared statistics); Contingency correlation coefficient Φ = 0.5764

The accuracy of the qSPR model proved to be almost 80% in all the investigated sets of compounds. The accuracy of the model in the training set proved not to be statistically different from the accuracy of the model in the test set (the confidence intervals overlap, see Tables 2 and 3). A similar interpretation is true when the values and associated confidence intervals of other statistical indicators are analyzed (see Tables 2 -4).

Table 4. Statistical indicators for assessing the qSPR model: overall set

Parameter (Abbreviation)	Value	95%CI
Concordance / Accuracy / Non-error Rate (CC/AC)	79.88	[73.22-85.43]
Error Rate (ER)	20.12	n.a.
Prior proportional probability of an active class	0.3963	[0.3238-0.4725]
Prior proportional probability of an non-active class	0.6037	n.a.
Sensitivity (Se)	83.08	[72.50-90.55]
False-negative rate (under-classification, FNR)	16.92	[9.45-27.50]
Specificity (Sp)	77.78	[68.82-85.05]
False-positive rate (over-classification, FPR)	22.22	[14.95-31.18]
Positive predictivity (PP)	71.05	[60.19-80.30]
Negative predictivity (NP)	87.50	[79.26-93.06]
Probability of classification		
- as active (PCA)	0.4634	[0.3883-0.5398]
- as non-active (PCIC)	0.5366	[0.4602-0.6117]
Probability of a wrong classification		
- as active compound (PWCA)	0.2895	[0.1970-0.3981]
- as non-active compound (PWCI)	0.1250	[0.0694-0.2074]
Odds Ratio (OR)	17.1818	[7.7989-38.1475]

95% CI = confidence interval at a significance level of 5%; n.a. = not available
χ^2 = 83.6385 (p < 0.0001) (Chi-squared statistics); Contingency correlation coefficient Φ = 0.5761

The Z test was applied in order to compare the statistical indicators expressed as probabilities obtained in training and test sets. The prior probabilities of active and non-active compounds proved not to be statistically different in training and test sets. The absence of statistically significant differences between prior probabilities of active and non-active compounds in training and test sets supports the correct assignment of compounds to the active/non-active sets. However, the probability of classification as active compounds proved to be statistically smaller in the training set compared to the test set (p=0.0042); thus, the classification model proved to perform better in terms of correct classification of active compounds when applied on test set.

The objective of this study is to propose a series of statistical indicators as diagnostic tools for the qSPR model. In achieving this, various aspects are considered:
- Analyzing the correct assignment of compounds to training and test sets: prior proportional probability of an active class & prior proportional probability of a non-active class
- Analyzing the correct classification of active and non-active compounds: all the other statistical indicators (see Table 2-4).

The proposed statistical indicators have to assess the qSPR model in training and test sets: as the indicators have similar performances in training and test sets, it could involve the model has similar classification abilities, thus being considered as a good model. The best model is the one with the highest possible accuracy and the smallest possible error rate. The best model is also the one with the highest sensitivity and specificity and the smallest false-negative and false-positive rates. In this respect, it can be observed that sensitivity is smaller than specificity in the training set while sensitivity is

higher than specificity in the test set (see Tables 2 and 3). In other words, the investigated qSPR model has a higher ability to correctly assign active compounds to the active class in the test set and a higher ability to correctly assign non-active compounds to the non-active class in the training set. An excellent classification model should also have the best possible positive and negative predictability values while the probability values of a wrong classification into active and non-active compounds should have the smallest possible values.

Similar statistical parameters are used to assess the performances of machine learning classification models: accuracy, recall (true positive rate, false positive rate, true negative rate, false negative rate, and precision) [25, 26]. These parameters are calculated based on the confusion matrix [27]. Note that the confusion matrix is the same as the generic contingency table presented in Table 1.

The present study is aimed to introduce a series of statistical indicators in order to diagnose a qSPR model. Useful information related to the assignment of compounds in the training and test sets could be obtained by using prior proportional probability of an active class & prior proportional probability of a non-active class. All the other proposed statistical indicators allow the characterization of a qSPR model in terms of total fraction of correctly classified compounds (accuracy), correct assignment to active or non-active class (sensitivity and specificity, false positive and false negative rates), etc. Statistical indicators were applied on a 2×2 confusion matrix but the same approach could also be applied on r×c confusion matrices when compounds are classified into more than two groups (e.g., non-active, active, and very active). The usefulness of this approach in diagnosing qSPR/qSAR models is currently investigated in our laboratory.

CONCLUSIONS

The total fraction of compounds correctly classified by the qSPR model proved to be identical in the training and test sets as well as in the overall set. However, the overall model and the model obtained in the test set showed a higher ability to correctly assign the non-active compounds to the negative class while the model obtained in the training set had a higher ability to correctly assign the active compounds to the active class.

EXPERIMENTAL SECTION

A previously reported qSPR model [28] able to characterize the aqueous solubility of drug-like compounds was herein used. The experimental aqueous solubility measured at 298K and expressed in mg/ml (values taken from Merck Index 13th [28]) was modeled using molecular descriptors [24].

The best model obtained in the training set (n=97) proved to be a model with 3 descriptors and the following characteristics [24]:

$R^2 = 0.871; S = 0.903$
$R^2_{loo} = 0.849; S_{loo} = 0.971$
$R^2_{val} = 0.848; S_{val} = 0.899$

where R^2 = determination coefficient; S = standard deviation of the model; R^2_{loo} = determination coefficient on leave one out analysis; S_{loo} = standard deviation on leave-one-out analysis; R^2_{val} = determination coefficient on validation set; S_{val} = standard deviation on validation set.

      A series of statistical indicators similar with those used in medical diagnostic tests [29, 30] were defined as diagnostic parameters for the qSPR model (Table 5).

      The experimental and estimated aqueous solubility of the studied compounds was transformed as dichotomial variables in order to calculate the defined statistical indicators (Table 5) using the following criteria: if experimental data ≥ 0, the compound was considered active, if experimental data < 0, the compound was considered non-active.

Table 5. Statistical indicators calculated on the 2×2 contingency table

Indicator (Abbreviation)	Formula	Definition
Accuracy / Non-error Rate (AC)	100*(TP+TN)/n	Total fraction of correctly classified compounds
Error Rate (ER)	100* (FP+FN)/n = 1-CC	Total fraction of misclassified compounds
Prior proportional probability of a class (PPP)	n_i/n	Fraction of compounds belonging to class *i*
Sensitivity (Se)	100*TP/(TP+FN)	Percentage of active compounds correctly assigned to the active class
False-negative rate (under-classification, FNR)	100*FN/(TP+FN) = 1-Se	Percentage of active compounds falsely assigned to the non-active class
Specificity (Sp)	100*TN/(TN+FP)	Percentage of non-active compounds correctly assigned to the non-active class
False-positive rate (over-classification, FPR)	100*FP/(FP+TN) = 1-Sp	Percentage of non-active compounds falsely assigned to the active class
Positive predictivity (PP)	100*TP/(TP+FP)	Percentage of compounds correctly assigned to the active class out of all compounds assigned to the active class
Negative predictivity (NP)	100*TN/(TN+FN)	Percentage of compounds correctly assigned to the non-active class out of all compounds assigned to the non-active class
Indicator (Abbreviation)	Formula	Definition
Probability of classification - as active (PCA)	(TP+FP)/n	- Probability to classify a compound as active (true positive & false

- as inactive (PCIC)	(FN+TN)/n	positive) - Probability to classify a compound as non-active (true negative & false negative)
Probability of a wrong classification - as active compound (PWCA) - as non-active compound (PWCI)	FP/(FP+TP) FN/(FN+TN)	Probability of a false positive classification Probability of a false negative classification
Odds Ratio (OR)	(TP*TN)/(FP*FN)	The odds of correct classification in the group of active compounds divided to the odds of an incorrect classification in the group of non-active compounds

The associated 95% confidence interval under the binomial distribution assumption [31] was also computed for the correct interpretation of the indicators [32].

ACKNOWLEDGMENTS

REFERENCES

1. L.P. Hammett, *Chemical reviews,* **1935**, *17*, 125.
2. I.M. Kapetanovic, *Chemico-Biological Interactions,* **2008**, *171(2)*, 165.
3. C.H. Andrade, K.F. Pasqualoto, E.I. Ferreira, A.J. Hopfinger, *Molecules,* **2010**, *15(5)*, 3281.
4. V. Potemkin, M. Grishina, *Drug Discov Today,* **2008**, *13(21-22)*, 952.
5. J. Li, P. Gramatica, *Journal of Chemical Information and Modeling,* **2010**, *50(5)*, 861.
6. M.C. Hutter, *Current Medicinal Chemistry,* **2009**, *16(2)*, 189.
7. M. Pavan, T.I. Netzeva, A.P. Worth, *SAR and QSAR in Environmental Research,* **2006**, *17(2)*, 147.
8. S. Wold, *Quantitaive Structure-Activity Relationship,* **1991**, *10*, 191.
9. R.D. Cramer III, J.D. Bunce, D.E. Patterson, I.E. Frank, *Quantitaive Structure-Activity Relationship,* **1988**, 7, 18; Erratum **1988**, 7, 91.
10. H. Kubinyi, *Quantitaive Structure-Activity Relationship,* **1994**, *13*, 285.
11. H. Kubinyi, *Quantitaive Structure-Activity Relationship,* **1994**, *13*, 393.
12. K. Héberger, *TrAC - Trends in Analytical Chemistry,* **2010**, *29(1)*, 101.
13. P. Ghosh, M.C. Bagchi, *Current Medicinal Chemistry,* **2009**, *16(30)*, 4032.
14. A.C. Good, S.J. Peterson, W.G. Richards, *Journal of Medicinal Chemistry,* **1993**, *36(20)*, 2929.

15. B.H. Su, M.Y. Shen, E.X. Esposito, A.J. Hopfinger, Y.J. Tseng, *Journal of Chemical Information and Modeling,* **2010**, *50(7)*, 1304.
16. L.G. Valerio Jr., *Toxicology and Applied Pharmacology,* **2009**, *241(3)*, 356.
17. H. Akaike, *Annals of the Institute of Statistical Mathematics,* **1969**, *21*, 243.
18. C.M. Hurvich, C. Tsai, *Biometrika,* **1989**, *76*, 297.
19. H. Kubinyi, *Quantitative Structure-Activity Relationships,* **1994**, *3*, 393.
20. H. Kubinyi, *Quantitative Structure-Activity Relationships,* **1994**, *13*, 285.
21. S.D. Bolboacă, L. Jäntschi, *TheScientificWorldJOURNAL*, **2009**, *9(10)*, 1148.
22. S.D. Bolboacă, M.M. Marta, C.E. Stoenoiu, L. Jäntschi, *Applied Medical Informatics*, **2009**, *25(3-4)*, 65.
23. S.D. Bolboacă, M.M. Marta, L. Jäntschi, *Folia Medica,* **2010**, *52(3)*, 37.
24. P.R. Duchowicz, A. Talevi, L.E. Bruno-Blanch, E.A. Castro, *Bioorganic & Medicinal Chemistry,* **2008**, 16(17), 7944.
25. A. Lombardo, A. Roncaglioni, E. Boriani, C. Milan, E. Benfenati, *Chemistry Central Journal,* **2010**, *4 Suppl 1*, S1.
26. N. Fjodorova, M. Vracko, M. Novic, A. Roncaglioni, E. Benfenati, *Chemistry Central Journal,* **2010**, Jul 29; 4 Suppl 1:S3.
27. M. Kubat, S. Matwin, Addressing the Curse of Imbalanced Training Sets: One Sided Selection. In *Proceedings of the Fourteenth International Conference on Machine Learning*, pages 179–186, Nashville, Tennesse. Morgan Kaufmann, **1997**.
28. The Merck Index An Encyclopedia of Chemicals, Drugs, and Biologicals; Merck & Co.: NJ, **2001**.
29. S.D. Bolboacă, L. Jäntschi *Electronic Journal of Biomedicine,* **2007**, *2*, 19-28.
30. S.D. Bolboacă, L. Jäntschi, A. Achimaş Cadariu, *Applied Medical Informatics,* **2004**, *14*, 27.
31. S.D. Bolboacă, L. Jäntschi, *International Journal of Pure and Applied Mathematics*, **2008**, *47(1)*, 1.
32. L. Jäntschi, S.D. Bolboacă, *TheScientificWorldJOURNAL*, **2010**, *10*, 865.

MODELING THE BIOLOGICAL ACTIVITY OF
2-ARYL-THIAZOLE DERIVATIVES

ERIKA TASNÁDI[a], CRISTINA MOLDOVAN[b]

ABSTRACT A QSAR study on a set of biologically active molecules belonging to the class of 2-aryl-thiazole, using topological indices, is reported. The purpose of the study is to find the best regression model for prediction of two biological activities: anti-oxidant and anti-inflammatory ones**.**

Keywords: QSAR, biological activity, prediction, regression analyses, correlation coefficient.

INTRODUCTION

Quantitative structure-activity relationship (QSAR) is the process by which chemical structures are quantitatively related with a well defined process, such as biological activity. The identification of the crucial factors involved in the relation structure-property is gained by the comparative analysis of a set of molecules. It is achieved with the help of topological descriptors and regression analysis, included in various algorithms. The topological characterization of the chemical structures allows their classification based on a similarity criterion.

The 14 molecules taken in study show anti-oxidant and anti-inflammatory activity and belong to the class of 2-aryl-thiazole derivatives. Their anti-inflammatory capacity was assessed by evaluating the acute phase bone marrow response, phagocytes' activity and NO synthesis (see below). The antioxidant effect of the tested compounds was assessed by evaluating: the total antioxidant response (TAR), the total oxidant status (TOS) and the index of oxidative stress (OSI) [OSI=(TOS/TAR)x100].

Phagocytic activity was assessed with the *in vitro* phagocytosis test by calculating two parameters: the phagocytic index (PI) (PI% = phagocytes with at least one phagocyted germ from 200 leukocytes counted) and the phagocytic activity (PA) (PA = number of germs phagocyted by 100 leukocytes) [3-8].

In acute inflammation there is a significant increase of NO synthesis due to the expression of iNOS (inducible nitric oxide synthesis). This will raise serum nitrates/nitrites concentration, as side metabolites of nitric oxide.

---

[a] *"Babes-Bolyai" University, Faculty of Chemistry and Chemical Engineering, Organic Chemistry Department*
[b] *UMF "Iuliu Hatieganu", Faculty of Pharmacy, Department of Pharmaceutical Chemistry*

In order for a molecule to have anti-oxidant effect TAR should raise or TOS should drop, and if both parameters drop, TAR should drop less, or if both rise, TAR should raise more. In order to have anti-inflammatory effect IF, AF and NO should drop.

The anti-inflammatory activity of the tested compounds was higher than that of Meloxicam, the drug taken as reference.

METHOD

The following procedure was used to find the best relationship between structures and the studied biological properties:

- structures are optimized to find a minimum-energy (stable) configuration (PM3, HYPER CHEM version 7.52);
- an index database is generated by using DRAGON 5.0 software and TOPOCLUJ software;
- an exhaustive search to find the best equations (i.e., with the correlation coefficient (R) higher than 0.90), by STATISTICA 6.0, software;

The molecules were designed by the aid of HYPER CHEM software. Geometry optimization was performed with the molecular mechanics method MM+, of the Polak-Ribiere conjugate gradient, at RMS lower than 0.009.

The topological indices were calculated by DRAGON (1630 indices) and TOPOCLUJ (962 indices) software. From these indices, the ones showing the best correlation coefficient in monovariate regression against the biological activity were selected out.

The statistical analysis was performed with STATISTICA software package, consisting in finding the best mono-, bi- and tri-variate regression equation, which can be further used to predict the biological activity of molecules belonging to the same class of those (2-aryl-thiazole derivatives) present in this study.

RESULTS AND DISCUSSION

Fourteen new 2-aryl-thiazole derivatives were synthesized by condensation between derivatives of 4-[2-(4-methyl-phenyl-thiazole-5-yl)-2-oxo-ethoxy]-benzaldehyde and 2-, 3- or 4-(2-aryl-thiazole-4-ylmethoxy)-benzaldehyde, and different carboxylic acid hydrazides.

For these new structures five parameters were calculated, further used in this study, the goal being to find the best regression equation between chemical structure and biological activity.

Table 1 presents the molecules from our set of study and the calculated parameters.

The meaning of the five parameters illustrated in Table 1, are: IF = phagocytic index (phagocytes with at least one phagocyted germ from 200 leukocytes counted); AF=phagocytic activity (number of germs phagocyted by 100 leukocytes); NO=nitric oxide (NO synthesis was evaluated measuring

nitrates/nitrites concentration); TAR=total antioxidant respons; TOS = total oxidant status. The first tree parameters are determined in case of an inflammatory process and the last two are used for testing antioxidant activity.

Table 1. 2-aryl-thiazole derivatives and their properties

Molecule	Formula	IF	AF	NO	TOS	TAR
5b	$C_{31}H_{26}N_4O_3S_2$	16.57 ±1.51	22.86 ±3.02	577.49 ±96.07	33.54 ±2.97	1.0969 ±0.0026
5c	$C_{31}H_{23}F_3N_4O_3S_2$	22.57 ±2.76	45.71 ±4.23	595.8 ±38.61	31.04 ±3.78	1.097 ±0.004
7b	$C_{29}H_{23}BrN_4O_2S_2$	34.57 ±5.13	55.43 ±3.41	1183.35 ±134.13	27.62 ±2.03	1.0978 ±0.0022
7c	$C_{29}H_{20}BrF_3N_4O_2S_2$	23.43 ±3.95	32.28 ±4.07	1121.64 ±123.65	40.18 ±1.83	1.103 ±0.003
7e	$C_{29}H_{20}BrF_3N_4O_2S_2$	31.14 ±4.74	25.14 ±4.88	558.92 ±72.49	35.77 ±3.3	1.0963 ±0.006
7f	$C_{34}H_{24}BrN_3O_5S$	15.86 ±3.29	16.57 ±2.15	580.35 ±72.28	45.9 ±2.46	1.0979 ±0.0023
7h	$C_{29}H_{23}BrN_4O_2S_2$	29.43 ±4.12	58.71 ±2.21	1031.16 ±146.91	27.27 ±3.52	1.1018 ±0.0029
7i	$C_{29}H_{20}BrF_3N_4O_2S_2$	31 ±3.6	73.28 ±4.5	1228.8 ±120.04	25.92 ±2.64	1.104 ±0.0065
7k	$C_{29}H_{24}N_4O_2S_2$	28.57 ±2.99	42.28 ±4.23	1192.44 ±49.95	28.04 ±3.3	1.0983 ±0.0007
7l	$C_{29}H_{21}F_3N_4O_2S_2$	16.57 ±1.51	28.28 ±1.38	888.45 ±155.09	10.98 ±1.36	1.0882 ±0.002
7m	$C_{34}H_{25}N_3O_5S$	30.86 ±2.54	47.43 ±2.22	662.32 ±142.95	9.82 ±1.1	1.0855 ±0.0047
7o	$C_{29}H_{24}N_4O_2S_2$	27.43 ±5.09	62.28 ±4.53	1152.59 ±79.89	26.66 ±2.83	1.0989 ±0.0014
7p	$C_{29}H_{21}F_3N_4O_2S_2$	16 ±3.51	19.14 ±2.54	1257.04 ±157.09	22.57 ±3.44	1.0952 ±0.001
7r	$C_{34}H_{25}N_3O_5S$	26.28 ±4.82	15.43 ±0.79	1158.25 ±91.61	20.35 ±0.69	1.0992 ±0.0032

We looked for the best regression equation in modeling all the five parameters; the best regression equations are listed in the following.

Property: IF.

Monovariate regression:

$$y = 27.02 + 0.849 \times Mor13u$$
$$R = 0.8489; \quad s = 1.14; \quad F = 30.94$$

Bivariate regression:

$$y = 21.70 - 0.31 \times MATS3m + 0.725 \times Mor13u$$
$$R = 0.8958; \quad s = 2.68; \quad F = 22.35$$

Trivariate regression:

$$y = 30.69 - 0.58 \times MATS3m - 1.7 \times Mor13v + 2.20 \times Mor13p$$
$$R = 0.9481; \quad s = 2.73; \quad F = 29.64$$

Property: NO.

Monovariate regression:

y = 6564.2-0.89×TIC2

R = 0.8940; s = 139.58;   F = 47.79

Bivariate regression:

y = -9333.75-0.6×IC3+0.475×X[Sh[CfMax[Charge]]]

R = 0.9389; s = 112.03;   F = 40.91

Trivariate regression:

y =-22353.63+0.225×ATS4p-0.68×IC3+0.628×X[Sh[CfMax[Charge]]]

R = 0.9427; s = 113.87;   F = 26.61

Property: AF.

Monovariate regression:

y = -14.33+0.77×R7p+

R = 0.7703; s = 12.25;   F = 17.51

Bivariate regression:

y = -4.31+0.658×H2u-0.67×PDS8[Sh[D3D]]

R = 0.8829; s = 9.42;   F = 19.44

Trivariate regression:

y = 101.94-0.55×E3u-0.54×PDS10[Sh[D3D]]+1.09×R7p+

R = 0.9180; s = 8.35;   F = 17.85

Property: TOS.

Monovariate regression:

y = -333.6+0.87×EEig10d

R = 0.8702;   s = 5.45;   F = 37.44

Bivariate regression:

y = -278.28+0.741×EEig10d-0.23×Mor10v

R = 0.8903;   s = 5.26; F = 21.02

Trivariate regression:

y =134.52+0.15×RDF040m-0.4×RDF135u-0.35×WkOp[SzMinSzMax U]

R = 0.9473;   s = 7.23; F = 6.03

Property: TAR.

Monovariate regression:

y = 1.22-0.88×GATS2v

R = 0.8818;   s = 0.003;   F = 41.93

Bivariate regression:

y = 0.96+0.562×EEig10d-0.49×Mor10v

R = 0.9355;   s = 0.002;   F = 38.59

Trivariate regression:

y = 0.96+0.58×EEig10d+0.566×Mor10v-1.00×Mor10v

R = 0.9550;   s = 0.002;   F = 34.54

Considering that a biological activity is a multi-conditional response, the models showed a clear correlation between activity and molecular structure, particularly in bi-and tri-variate equations. The study needs to be continued to enlarge the data set for a better statistical significance.

CONCLUSION

The class of 2-aryl-thiazoles is known for various biological activities, the anti-oxidant and anti-inflammatory included. The present article reported the modeling of these two bio-activities by using topological indices. Based on the regression models here presented we can predict the biological activity for molecules belonging to the same class and not included in the regression equation.

This theoretical study stand as a support for further experiments in finding molecules with desired anti-oxidant and anti-inflammatory activity.

ACKNOWLEDGEMENT

REFERENCES

1. G. Katona, G. Turcu, A.A. Kiss, O.M. Minailiuc, and M.V. Diudea, *Rev. Roumaine Chim.*, **2001**, *46*, 137.
2. O. Ursu, G. Katona, and M.V. Diudea, *Rev. Roum. Chim.*, **2003**, *48*(4), 321.
3. M.V. Diudea, (Ed.), *QSPR/QSAR* Studies by Molecular Descriptors, *NOVA*, New York, **2001**, 438.
4. R. Todeschini et al., Dragon software, http://www.talete.mi.it.
5. L. Pleşca-Manea, A.E. Pârvu, M. Pârvu, M. Tămaş, R. Buia, M. Puia, *Phytother Res*, **2002**, *16(4)*, 316.
6. B. Tiperciuc, A. Pârvu, M. Palage, O. Oniga, D. Ghiran, *Farmacia*, **1999**, *5*, 77.
7. S. Greenberg, S. C. Silverstein, Phagocytosis, Fundamental Immunology, third ed., *Raven Press*, New York, **1993**, 941.
8. M.A. Gougerot-Pocidalo, J. El Benna, C. Elbim, S. Chollet-Martin, Mc. Dang, *J Soc Biol*, **2002**, *196 (1)*, 37.
9. J. Rodenas, M.T. Mitjavila, T. Carbonell, *Free Radic Biol Med*, **1995**, *18 (5)*, 5869.
10. A. Hrabak, T. Bajor, I. Csuka, *Inflamm Res*, **2006**, *55(1)*, 23.
11. A. Hrabak, T. Bajor, I. Csuka, *Inflamm Res*, **2008**, *57(2)*, 75.
12. R. Stoika, N. Kashchak, M. Lutsik-Kordovsky, M. Boyko, M. Barska, A. Tsyrulnyk, *Med Sci Monit*, **2001**, *7(4)*, 652.
13. O. Erel, *Clinical Biochemistry*, **2004**, *37*, 112.
14. O. Erel, *Clinical Biochemistry*, 2005, *38*, 1103.

STRUCTURE BASED ALGORITHM FOR CLASSIFICATION OF MALEIMIDE DERIVATIVES ATP-COMPETITIVE INHIBITORS OF GSK-3

LILIANA M. PĂCUREANU[a], ALINA BORA[a], LUMINIȚA CRIȘAN[a], LUDOVIC KURUNCZI[b]

ABSTRACT. The structure based retrospective virtual screening algorithm employed the docking engine FRED (Fast Rigid Exhaustive Docking) to dock 74 inhibitors (4-aryl-3-anilino-maleimide derivatives) and 1778 decoy molecules into glycogen synthase kinase-3 β, GSK-3β, ATP-binding site (PDB code 1Q4L).

The input database of 74 ligands was prepared following the OpenEye protocol by adding tautomers and ionization states, generating conformers, and performing charge corrections with AM1BCC option from QUACPAC software. The protein preparation has been carried out with Chimera software by deleting water molecules (except water near Thr 138), adding hydrogen and charges (AM1BCC). The energy component values of the scoring functions were subsequently submitted to PLS-DA (Projections in Latent Structures, Discriminant Analysis). The final PLS-DA result contains only the essential energy factors that describe most accurately the interactions in the ATP binding site. The results obtained are of better quality than those obtained using the total scores provided by initial scoring functions in terms of AUC (Area Under Curve) 0.938 (chemgauss2 donor + screenscore rotatable bonds) with respect to 0.887 (chemgauss3). Moreover, the early enrichment of the PLS-DA term at 1% of the database is 13.514% while for Chemgauss 3 was only 8.108%.

Keywords: molecular docking, Projections in Latent Structures - Discriminant Analysis (PLS-DA), glycogen synthase kinase-3β (GSK-3β)

INTRODUCTION

The identification of selective inhibitors of protein kinases by virtual screening strategies withdraw much interest in the area of drug discovery by helping in terms of time and money the high throughput screening (HTS) experiments [1]. GSK-3 is a serine/threonine protein kinase, discovered as the enzyme that inactivates the glycogen synthase (GS), the rate limiting enzyme in glycogen synthesis [2]. Besides glycogen metabolism regulation [2,3], GSK-3 controls a large number of cellular processes such as microtubule stability [4], β-catenin degradation [5], protein translation [6], etc.

[a] *Institute of Chemistry of Romanian Academy, 24 Mihai Viteazul Bvd., RO-300223, Timisoara, Romania, pacureanu@acad-icht.tm.edu.ro*
[b] *University of Medicine and Pharmacy "Victor Babes", Faculty of Pharmacy, 2 Eftimie Murgu Avenue, RO-300041, Timisoara, Romania, dick@acad-icht.tm.edu.ro*

Maleimide derivatives have been identified as ATP competitive inhibitors of GSK-3α at Smithkline Beecham pharmaceutical company by means of a high throughput screening experiment [7]. GSK-3 inhibition by maleimide derivatives caused the acceleration of glycogen synthesis in the liver suggesting the utility of maleimide inhibitors for the treatment of diabetes [3]. Moreover, additional biological investigations demonstrated that maleimide derivatives prevent neuronal death through a mechanism that involve, interactions with tau and β-catenin [8].

Structural characteristics of GSK-3 inhibitors have been investigated by various techniques including QSAR, docking and ligand based virtual screening [9,10,11,12,13,14,15].

Our investigation is directed towards a structure-based methodology due to the availability of X-ray cocrystal GSK-3β - maleimide derivative [16]. The high identity (similarity) of human GSK-3α and β 83% (89%) overall and 91% (97%) of the catalytic domain [17] permitted us to use the X-ray stucture of GSK3 β to dock the maleimide inhibitors tested in GSK 3α [7]. The docking algorithm has to check that the chemical compounds make favorable interactions with the enzyme. Therefore, the set of inhibitors were mixed with a large number of inactives (decoys) in order to reproduce the real situation.

Scoring functions, as they have been constructed, display a series of shortcomings, especially high false positive rates. Consensus scoring has been introduced to counterweight for false positive rates of individual scoring functions. But the selection algorithm for the right, individual scoring functions represents the major challenge [18]. Jacobsson *et al.* [19] have used PLS-DA (Projections in Latent Structures - Discriminant Analysis) methodology to the total scores of individual scoring functions in order to improve the performance of individual scoring functions. In this paper we introduced the PLS-DA methodology [20] to the variables representing the components of individual scoring functions in order to get a new combination of terms that will rank more appropriately the actives with respect to inactives.

Dataset

In our study, a dataset of 74 derivatives of 3-anilino-4-arylmaleimide [7] (Figure 1) and their biological activity, measured as inhibitory activity IC_{50} (nM) evaluated against human GSK 3α, is considered. Our dataset is assembled/ mixed with a decoy set of 1778 molecules (CDK-2 decoys) downloaded from DUD (Directory of Useful Decoys) [21].
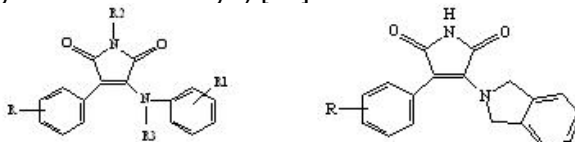
Figure 1. The template of maleimide derivatives (see ref 7):
R = H, 2-Cl, 2-OMe, 2-NO$_2$, 3-Cl, 3-OMe, 3-NO$_2$, 4-Cl, 4-OMe, 4-NO$_2$
R1 = H, 3-Cl, 3-OH, 4-OH, 3-Cl-4OH, 3,5-diCl-4-OH, 3-CO$_2$H, 4-Cl-3CO$_2$H, 4-SMe
R2, R3 = H, CH$_3$

The CDK-2 decoys were chosen on the basis of high similarity of aminoacid binding sites (85%) of GSK-3 and CDK-2 [16]. In the current study we assume the decoys are inactive, even not experimentally tested on GSK-3. Therefore, they probably can be active on this target [21]. The distribution of drug-like properties of actives and decoys are shown in Table 1.

Table 1. Drug-like properties of actives and decoys

	Molecular Weight	Rotatable Bond Number	Number of hydrogen bond donors	Number of hydrogen bond acceptors	MLOGP
Actives					
min	264.3	0	2	4	-3.569
max	575.68	12	5	12	2.773
Decoys					
min	298.37	1	0	4	-5.974
max	399.47	11	7	11	4.062

Protein preparation

The crystal structure of GSK-3 β (PDB entry: 1Q4L) in complex with inhibitor 2-chloro-5-[[4-(3-chlorophenyl)-2,5-dioxo-pyrrol-3-yl]amino]benzoic acid was downloaded from the PDB. The active site of the enzyme was prepared using Chimera package [22] deleting water molecules except water near Thr 138, that was kept as it mediates the hydrogen bonds to O^γ of Thr138 and $O^{\epsilon 2}$ of Gln185, [16] adding hydrogens and AM1BCC charges.

Assignment of ionization states and generation of tautomers

Database preparation before virtual screening analysis is important for the quality of the results. Kirchmaier demonstrated that tautomerism is essential for the classification of actives in virtual screening experiments [23]. The three-dimensional structures of 74 GSK-3α inhibitors were prepared using LigPrep 2.2 module of Maestro in the Schrödinger software [24]. For the ligands, the only reasonable tautomeric forms at pH=7.4±1.5 were selected.

Conformer generation

Conformer generation for ligands and decoys was performed with Omega version 2.-2.3.2 from OpenEye package [25]. Biologically active fragment conformations are available in Omega's library. The ligand is split into fragments and next reassembled according to energetic criteria and the conformations complying with the energy window and heavy atom root mean square (RMS) distance were saved. We used an increment-based methodology for energy window of "5.0, 6.0, 7.0" kcal/mol, and RMS distance of the heavy atom coordinates for conformer detection of "0.5, 0.4, 0.3" Å. The assignments

of appropriate atomic charges were carried out with QuacPac software [25], choosing AM1BCC option (AM1 bond charge correction). The resulting conformer enriched database of actives and decoys was used as input for docking.

Docking procedure

Docking investigation was carried out with FRED (Fast Rigid Exhaustive Docking) software version 2.2.5 (www.eyesopen.com) [25]. The docking procedure occurs in two steps: shape fitting and optimization. The ligand is placed into a 0.5Å resolution grid-box incorporating all active site atoms (including hydrogen atoms) using a smooth Gaussian potential [26]. To score the ligand in the docking procedure the binding site of GSK-3β was defined using the reference ligands and an addbox of 4Å around the ligand. The best docked pose per each ligand was saved and seven classical scoring functions including Chemscore (CS), Chemgauss-2 (CG2), Chemgauss-3 (CG3), Shapegauss (SG), Screenscore (SC), OEChemscore (OECS), and PLP were used.

PLS-DA analysis

In the present work, we attempted to implement a multivariate statistical method (PLS DA), with the values of scoring function components as descriptors, in order to classify the virtual screening results in active and inactive compounds [27]. PLS is a regression method that works with two matrices, X (e.g., chemical descriptors) and Y (e.g., biological responses), and has two objectives, namely to approximate well X and Y, and to model the relationship between them [28]. For PLS DA methodology two classes are defined: the actives (1) and the inactives (2) according to ligands and decoys.

The energetic component outputs of all scoring functions (see reference [25]) were submitted to the SIMCA P 9.0 package [29] to perform initially a PCA (Principal Component Analysis) analysis [30], followed by the PLS DA analysis.

RESULTS AND DISCUSSION

In the first step of PLS DA analysis, a PCA model for the whole X matrix (N=1852 rows/compounds, and K=32 columns/energetic terms) was performed and three principal components were obtained. These three principal components explain 47.7% of the information content of the X matrix and distinguish very well the actives (*in black*) from the inactives (*in red* - Figure 2).

The PLS DA models were further constructed starting from the same X matrix. In order to improve the PLS DA models, the coefficient sign and VIP >1 (variable influence on projection) were considered as significant. Based on these criteria, six out of thirty two energetic terms were selected: CG2 Donor (Chemgauss2 contributions from donors on the ligand interacting with acceptors on the protein), CG3 Steric (Chemgauss3 steric interactions), CS HB (Chemscore hydrogen bonds), SC RB (Screenscore rotatable bond), SC Ambig (Screenscore ambiggous interactions), and SC HB (Screenscore hydrogen bonds). For these

six energetic terms, all the possible combinations were made and the first significant combination (CG2 Donor + SC RB) was selected. The sum of these terms represents the PLS-DA equal weight "mixed" scoring function.

PLS_DAt[1]/t[2]/t[3]
Colored according to classes: decoys-red and active -black
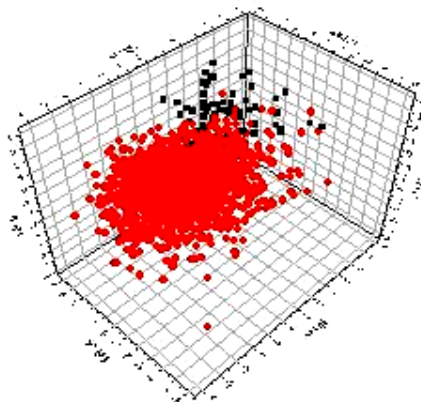
Figure 2. Classes of actives (in black) and inactives (in red)

In order to test the performance of the new "mixed" scoring function against classical scoring functions, the AUC and enrichment factors were compared. The results of ensemble AUC and enrichments are illustrated in Figure 3a and 3b.
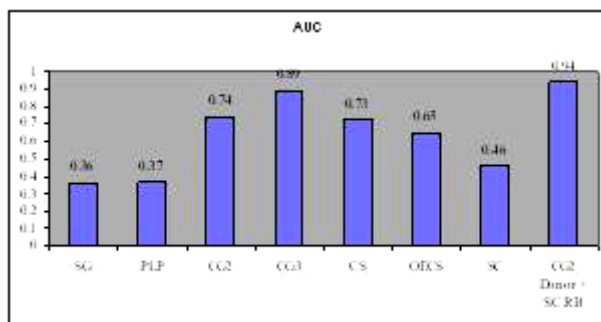
Figure 3.a) Bar chart showing AUC values obtained
with seven classical and the new „mixed" scoring functions

The AUC of 0.887 and enrichment factor of 8.108% at 1% of database show good performances of the classical CG3 at the beginning, but these results were surpassed by the corresponding values of the "mixed" components (CG2 donor + SC RB) scoring functions AUC (0.938 and enrichment factor 13.513 % at 1% of database for this combination).

87
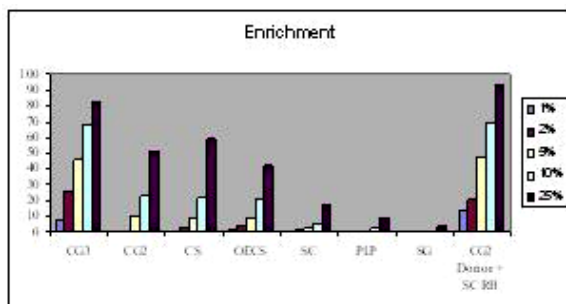
Figure 3.b) Enrichment performances at 1%, 2%, 5%, 10% and 25% of the database

Analyzing the classical CG2 and SC scoring functions, AUC is 0.735 and respectively 0.459 while the enrichment factor is 0.011% / 0.011% by the top 1% database and show low performances, but the donor + rotatable bond components (CG2 Donor + SC RB) seems to be significant in this combination.

The CG2 Donor energetic term into the "mixed" components scoring function measures the H-bond interaction energy between ligand and protein. The SC RB component is a penalty term proportional to the number of rotatable bonds in the ligand. SC RB is an important term in our situation since a number of compounds display a considerable number of flexible bonds in the decoys (up to 11) and ligands (up to 12).

In the top 2% - 25% of the database, the number of detected actives increases and the largest percentage (93.243%) was retrieved at 25% in the case of new "mixed" scoring functions.

CONCLUSIONS

Here we reported a promising workflow for structure-based virtual screening using rigid docking (FRED software) followed by PLS DA analysis. A new "mixed" scoring function was built. It collects the energy factors from different scoring functions that illustrate the particular interactions in the GSK3β site. In this way, the results here reported, are of better quality than those obtained by using every single scoring function available in the OpenEye package. The present study enabled us to indentify the optimal protocol for the highest enrichment of actives in the top 1% to 25% of the database for seven classical and one "mixed' scoring function. Therefore, in the following studies the algorithm for docking scoring aiming at ranking the actives versus decoys will be based on all possible combinations.

ACKNOWLEDGMENTS

REFERENCES

1. P. Cohen, *Nature Reviews Drug Discovery*, **2002**, *1*, 309.
2. P. Cohen, "Muscle Glycogen Synthase", Enzymes Academic Press, New York, **1986**, vol. XVII, 461-497.
3. M.P. Coghlan, A.A. Culbert, D.A.E. Cross, S.L. Corcoran, J.W. Yates, N.J. Pearce, O.L. Rausch, G.J. Murphy, P.S. Carter, R. Roxbee Cox, D. Mills, M.J. Brown, D. Haigh, R.W. Ward, D.G. Smith, K.J. Murray, A.D. Reith, J.C. Holder, *Chem. Biol.*, **2000**, *7*, 793.
4. M. Hong, D.C. Chen, P.S. Klein, V.M. Lee, *J. Biol. Chem*., **1997**, *272*, 25326.
5. M.J. Hart, R. de los Santos, I.N. Albert, B. Rubinfeld, P. Polakis, *Curr. Biol.*, **1998**, *8*, 573.
6. G.I. Welsh, C.M. Miller, A.J. Loughlin, N.T. Price, C. G. Proud, *FEBS Lett*., **1998**, *421*, 125.
7. D.G. Smitth, M. Buffet, A.E. Fenwick, D. Haigh, R. Ife, M. Saunders, B.P. Slingsby, R. Stacey, R.W. Ward, *Bioorg. Med. Chem. Lett.*, **2001**, *11*, 635.
8. D.A.E. Cross, A.A. Cubert, K.A. Chalmers, L. Facci, S.D. Skaper, A.D. Reith, *J. Neurochem*., **2001**, *77*, 94.
9. P. Polychronopoulos, P. Magiatis, A.L. Skaltsounis, V. Myrianthopoulos, E. Mikros, A. Tarricone, A. Musacchio, S.M. Roe, L. Pearl, M. Leost, P. Greengard, L. Meijer, *J. Med. Chem*., **2004**, *47,* 935.
10. F.X. Tavares, J.A. Boucheron, S.H. Dickerson, R.J. Griffin, F. Preugschat, S.A. Thomson, T.Y. Wang, H.Q. Zhou, *J. Med. Chem*., **2004**, *47*, 4716.
11. A. Peat, D. Garrido, J.A. Boucheron, S.L. Schweiker, S.H. Dickerson, J.R. Wilson, T.Y. Wang, S.A. Thomson, *Bioorg. Med. Chem. Lett*., **2004**, *14*, 2127.
12. M. Zeng, Y. Jiang, B. Zhang, Z. Kewen, N. Zhang, Q. Yu, *Bioorg. Med. Chem. Lett.*, **2005**, *15*, 395.
13. A.R. Katritzky, L.M. Pacureanu, D.A. Dobchev, D. Fara, P. Duchowitz, M. Karelson, *Bioorg. Med. Chem.,* **2006**, *14*, 4987.
14. N. Dessalev, P.V. Bharatam, *Eur. J. Med. Chem*., **2007**, *42*, 1014.
15. N. Dessalev, P.V. Bharatam, *Biophys. Chem*., **2007**, *128*, 165.
16. J.A. Bertrand, S. Thieffine, A. Vulpetti, C. Cristiani, B. Valsasina, S. Knapp, H.M. Kalisz, M. Flocco, *J. Mol. Biol.*, **2003**, *333*, 393.
17. A. Adnan, K.P. Hoeflich, J.R. Woodgett, *Chem. Rev*., **2001**, *101,* 2527.
18. M. Stahl, M. Rarey, *J. Med. Chem*., **2001**, 5, 375.
19. M. Jacobsson, P. Lidn, E. Stjernschantz, H. Bostrm, U. Norinder, *J. Med. Chem.*, **2003**, *46*, 5781.
20. S. Wold, E. Johansson, M. Cocchi, *ESCOM: Leiden*, **1993**, 523.
21. N. Huang, B.K. Shoichet, J.J. Irwin, *J. Med. Chem.*, **2006**, *49*, 6789. http://dud.docking.org./r2/
22. UCSF Chimera v 1.3 http://www.cgl.ucsf.edu/chimera
23. J. Kirchmaier, P. Markt, S. Distinto, G. Wolber, T. Langer, *J. Comput. Aided Mol. Des*., **2008**, *22*, 213.

24. LigPrep, version 2.2, Schrödinger, LLC, New York, NY, **2005** [http://www.schrodinger.com/].

25. OpenEye Scientific Software, Inc., 9 Bisbee Ct, Suite D Santa Fe, NM 87508 http://www.eyesopen.com/.

26. M.R. McGann, H.R. Almond, A. Nicholls, J.A. Grant, F.K. Brown, *Biopolymers*, **2003**, *68*, 76.

27. M. Jacobsson, P. Lidn, E. Stjernschantz, H. Bostrm, U. Norinder, *J. Med. Chem.,* ***2003, 46,*** 5781.

28. L. Eriksson, J. Gottfries, E. Johansson, S. Wold, *Chemometrics and Intelligent Laboratory Systems*, **2004**, *73*, 73.

29. SIMCA P, version 9.0; Umetrics AB: Umea, Sweden. http://www.umetrics.com.

30. M. Daszykowski, K. Kaczmarek, Y. Vander Heyden, B. Walczak, *Chemom. Intel. Lab. Syst.*, **2007**, *85*, 203.

KEKULÉ COUNT IN TUC$_4$C$_8$(R) NANOTUBES

A. R. ASHRAFI[a,*], P. NIKZAD[a], A. BEHMARAM[b], H. YOUSEFI-AZARI[b]

ABSTRACT. Counting Kekulé structures is a very difficult problem in chemical graph theory. Some recent techniques allowed to estimate the lower bound of this number in certain classes of graphs. In this note a formula for the number of Kekulé structures in TUC$_4$C$_8$(R) nanotube is given.

Keywords: TUC$_4$C$_8$(R) nanotube, Kekulé structure.

INTRODUCTION

Kekulé structures (perfect matchings in graph theory) in benzenoid hydrocarbons are discussed in the famous book of Cyvin and Gutman [1]. In physics, the enumeration of Kekulé structures is equivalent to the dimer problem of rectangle lattice graph in the plane [2]. The Kekulé count of nanostructures has become interesting subjects of research. Close formulas for the Kekulé count have been obtained in [3-6].

A graph G consist of a set of vertices V(G) and a set of edges E(G). In chemical graphs the vertices of the graph correspond to the atoms of the molecule and the edges represent the chemical bonds. The number of vertices and edges in a graph will be denoted by |V(G)| and |E(G)|, respectively.

A matching of a graph G is a set M of edges of G such that no two edges of M share an end-vertex; further a matching M of G is perfect if any vertex of G is incident with an edge of M. The concept of perfect matching in graphs coincides with the Kekulé structure in organic chemistry. In this paper we focus our attention on the number of Kekulé structures in TUC$_4$C$_8$(R) nanotube and a close formula is established, see [7-15] for details.

A C$_4$C$_8$ net is a trivalent decoration made by alternating rhombi C$_4$ and octagons C$_8$. It can cover either a cylinder or a torus. In some research papers, some topological indices of TUC$_4$C$_8$(R/S) nanotubes and TC$_4$C$_8$(R/S) nanotori have been investigated [16-22].

In this paper the TUC$_4$C$_8$(R)[p,q] = TU[p,q] nanotube is considered, where p and q are the number of octagons in each row and column, respectively. We explain the methods for computing the number of Kekulé structures in TU[p,q] and compute exact formula for the number of Kekulé structures in

---
[a] *Institute for Nanoscience and Nanotechnology, University of Kashan, Kashan 87317-51167, I. R. Iran*
[*] *Author to whom correspondence should be addressed. (E-mail: ashrafi@kashanu.ac.ir)*
[b] *School of Mathematics, Statistics and Computer Science, University of Tehran, Tehran, I. R. Iran*

some special case of TUC$_4$C$_8$(R) nanotubes, see Figure 1 (notice that the edges in the left side are affixed to the vertex in the right side of the figure to gain a tube in this way).
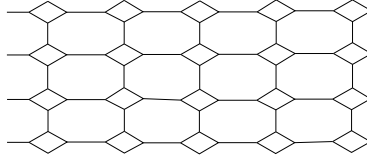
Figure 1. The chemical graph of TU[5,3].

MAIN RESULTS AND DISCUSSION

The aim of this section is to compute the number of Kekulé structures in TU[p,q] = TUC$_4$C$_8$(R)[p,q] nanotubes. The edges of rhombus in the molecular graph of $TU[p,q]$ are called the rhomboidal edges while the other edges are named octagonal.

LEMMA 1. Consider the molecular graph of TU[p,1] = TUC$_4$C$_8$(R)[p,1] and E is a Kekulé structure of TU[p,1] containing a horizontal edge, Figure 2. Then c, d $\notin$ E and a, b $\in$ E.
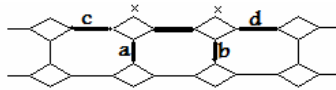
Figure 2. The molecular graph of TU[4,1].

PROOF. If E contains one of c or d then vertices shown by (x) could not be select in the matching, a contradiction. So, we must have the following figure for the matching:
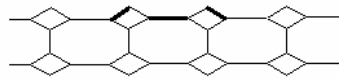
Figure 3. A part of a Kekulé structure without edges c and d.

By considering Figure 3 and the fact that in the upper selected rhomb, all the vertices must be covered, we have the following scheme for our Kekulé structure:
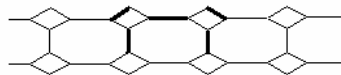
Figure 4. A part of a Kekulé structure containing a horizontal edge.

This completes our argument.   ▲

Corollary. There are exactly two Kekulé structure containing a given horizontal edge. These are as follows:

Figure 5. A part of two possible Kekulé structures containing a horizontal edge.

Theorem 1. Suppose K(p,1) denotes the number of Kekulé structures in a TU[p,1] nanotube. Then we have:

$$K(p,1) = \begin{cases} 2^{2p} + p(2^{2p-4} + 2^{2p-8} + ... + 2^4) + 4 & p \ is \ even \\ 2^{2p} + p(2^{2p-4} + 2^{2p-8} + ... + 2^2) & p \ is \ odd \end{cases}$$

Proof. We first note there are 2^{2p} Kekulé structures when we consider only the rhomboidal edges of TU = TU[p,1], see Figure 6.
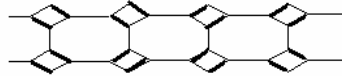
Figure 6. Kekulé structure containing rhomboidal edges.

Clearly, each of the rhomboidal edge can take part to a Kekulé structure in two schemes. Since the number of rhombi is 2p, we have 2^{2p} different choice for the number of Kekulé structures.

We now apply Lemma 1, to enumerate the Kekulé structures containing at least one non-rhomboidal edge.
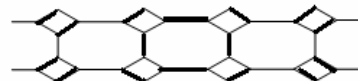
Figure 7. A Kekulé structure containing non-rhomboidal edges.

As it is shown in Figure 7, we have 2p − 4 rhombi, each of them belonging to two Kekulé structures and it is worth mentioning that this scheme can circulate in p situations. So, in this case we have p2^{2p-4} Kekulé structures. Figure 8 shows a Kekulé structure when two of the octagonal edges in a row take part in matching. By lemma 1, we know that no two incident edges in a row may belong to a Kekulé structure.
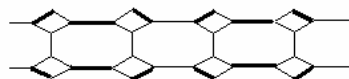
Figure 8. A Kekulé structure containing two of non-rhomboidal edges in a row.

At the end, we have a chain (Figure 9) that has two circulations for each of them. So we have 4 extra Kekulé structures, when p is even.
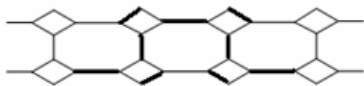
Figure 9. The extra Kekulé structures, when p is even.

This completes our proof.   ▲

Using a similar argument as Theorem 1, one can compute the number of Kekulé structures of TU[2,q].

Theorem 2. The number of Kekulé structure in TU[2,q] is 4×5^q.

Proof. To calculate the number of Kekulé structure in TU[2,q], we first find a recursive equation for the number of Kekulé structures and then solve it. Suppose A(q) denotes the set of all Kekulé structures of TU[2,q] and x_q is its size. From Figure 10, one can see that there are two types of Kekulé structures for TU[2,q] as follows: the first type Kekulé structures contain both e_1 and f_1; the second type Kekulé structures are those without e_1 and f_1.

Suppose L_1 and L_2 denote the number of Kekulé structures of the first and second types, respectively. Then from Figure 10, it can easily seen that $L_1 = 4x_{q-1}$. Suppose M is a Kekulé structure of the second type. Also, there are $4x_{q-2}$ Kekulé structures of the second type such that $e_2, f_2 \notin M$. Continue this argument, we can see that $x_q = 4[x_{q-1} + x_{q-2} + \ldots + x_1]$. To complete the proof, we must solve this recursive equation. To do this, notice that $x_{q-1} = 4[x_{q-2} + x_{q-3} + \ldots + x_1]$ and so $x_q - x_{q-1} = 4\,x_{q-1}$. Therefore, $x_q = 5 \times x_{q-1}$ which implies that $x_q = 5^{q-1} \times x_1$. An easy calculation shows that $x_1 = 20$ and so $x_q = 4 \times 5^q$.   ▲
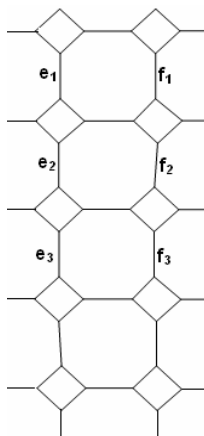
Figure 10. The Molecular Graph of TU[2,q].

From Theorems 1 and 2, we can find an upper and lower bounds for the number of Kekulé structures of TU[p,q] as follows:

Theorem 3. $(4 \times 5^q)^{p/2} \leq K(p,q) \leq (4 \times 5^q)^p$.

CONCLUSIONS

In this paper a simple method enabling to compute the Kekulé structures of TUC$_4$C$_8$(R) nanotubes with a small number of rows or columns was presented. By this method an upper and lower bound for this number is also calculated. It is possible to extend our method in view of obtaining better bounds.

ACKNOWLEDGMENTS

REFERENCES

1. S.J Cyvin, I. Gutman, Kekulé structures in benzenoid hydrocarbons, Lecture Notes in Chemistry, vol. 46, SSpringer, Berlin, **1988**.
2. P.W. Kasteleyn, *Physica*, **1961**, *27*, 1290.
3. J.G. Qian, F.G. Zhang, *J. Math. Chem*., **2005**, *38*, 233.
4. C.D. Lin, P. Tang, *J. Chem. Inform. Comput. Sci*., **2004**, *44*, 13.
5. S. Cigher and M.V. Diudea, *MATCH Commun. Math. Comput. Chem*., **2008**, *60*, 965.
6. S. Cigher and M.V. Diudea, *MATCH Commun. Math. Comput. Chem*., **2008**, *60*, 955.
5. M.V. Diudea, A. Graovac, *MATCH Commun. Math. Comput. Chem*., **2001**, *44*, 93.
6. M.V. Diudea, I. Silaghi-Dumitrescu, B. Parv, *MATCH Commun. Math. Comput. Chem*., **2001**, *44*, 117.
7. M.V. Diudea, P.E. John, *MATCH Commun. Math. Comput. Chem.,* **2001**, *44*, 103.
8. M.V. Diudea, *Bull. Chem. Soc. Jpn.,* **2002**, *75*, 487.
9. M.V. Diudea, *MATCH Commun. Math. Comput. Chem.,* **2002**, *45*, 109.
10. M.V. Diudea, D. Vukičević, J. Nanosci. *Nanotechnol*., **2007**, 7, 1321.
11. S. Cigher, D. Vukičević, M.V. Diudea, *J. Math. Chem.,* **2009**, *45*, 279.
12. S. Cigher and M.V. Diudea, *MATCH Commun. Math. Comput. Chem*., **2008**, *60*, 965.
13. S. Cigher and M.V. Diudea, *MATCH Commun. Math. Comput. Chem*., **2008**, *60*, 955.

14. A.R. Ashrafi, S. Yousefi, *MATCH Commun. Math. Comput. Chem.,* **2007**, *57*, 403.
15. S. Yousefi, A.R. Ashrafi, *J. Math. Chem.,* **2007**, *42*, 1031.
16. S. Yousefi, A.R. Ashrafi, *Current Nanoscience,* **2008**, *4*, 161.
17. S. Yousefi, A.R. Ashrafi, *MATCH Commun. Math. Comput. Chem*., **2006**, *56*, 169.
18. A.R. Ashrafi, S. Yousefi, *Nanoscale Res. Lett*., **2008**, *2*, 202.
19. A.R. Ashrafi, M. Faghani, S.M. Seyedaliakbar, *Digest Journal of Nanomaterials and Biostructures,* **2009**, *4*, 59.
20. A.R. Ashrafi, P. Nikzad, *Digest Journal of Nanomaterials and Biostructures*, **2009**, *4*, 383.

ON ESTRADA INDEX OF TWO CLASSES OF DENDRIMERS

GHOLAM HOSSEIN FATH-TABAR[a], ALI REZA ASHRAFI[a], ANTE GRAOVAC[b,c,d]

ABSTRACT. Suppose $G = (V, E)$ is a graph. The sequence $v_1 v_2 \ldots v_t (v_t = v_1)$ is called a closed walk with length $t - 1$ in G if v_i's are in $V(G)$ and $v_i v_{i+1} \in E(G)$. In this paper, the number of closed walks with length k, $C_w(G,k)$, for two classes of denderimers are computed.

Keywords: Dendrimer, closed walk.-graph spectrum, Estrada index.

INTRODUCTION

Dendrimers are polymeric macromolecules composed of multiple perfectly-branched monomers radially emanating from a central core, Figures 1, 2. The number of branching points increases upon moving from the dendrimer core to its surface and defines dendrimer generations. They are being investigated for possible uses in nanotechnology, gene therapy, and other fields [1-5].

In this paper, the word graph refers to a finite, undirected graph without loops and multiple edges. Suppose G is a graph. The vertices and edges of G are denoted by V(G) and E(G), respectively. A walk in G is an alternating sequence of graph vertices and edges such that any subsequent two edges are adjacent. A closed walk is a walk in which the first and the last vertices are the same. We encourage to the reader to consult papers [6-11] for background material, as well as basic computational techniques. Our notation is standard and mainly taken from the standard book of graph theory [12].

MAIN RESULTS AND DISCUSSION

Let $D_1[n]$ and $D_2[n]$ be the molecular graphs of the dendrimers depicted in Figures 1 and 2, respectively. In this section, some formulas are derived for the number of closed walks of length k, $C_w(G,k)$, where $1 \leq k \leq 10$ and G is one of the molecular graphs $D_1[n]$ and $D_2[n]$. For the sake of completeness, we mention here a well-known theorem in algebraic graph theory as follows:

[a] *Department of Mathematics, Statistics and Computer Science, Faculty of Science, University of Kashan, Kashan 87317-51167, I. R. Iran*
[b] *Faculty of Science, University of Split, Nikole Tesle 12, HR-21000, Split, Croatia*
[c] *NMR Center, Institute »R. Boskovic«, HR-10002 Zagreb, POB 180, Croatia*
[d] *IMC, University of Dubrovnik, Branitelja Dubrovnika 29, HR-20000 Dubrovnik, Croatia*

THEOREM 1. $C_w(D_1[n], 2k-1) = 0$ and $C_w(D_2[n], 2k-1) = 0$.

In the following theorems, the number of walks, of even length, are computed.

THEOREM 2. $C_w(D_1[n], 2) = 4 \times 3^{n+1} - 4$ and $C_w(D_2[n], 2) = 2 \times 3^{n+1} - 4$.

PROOF. Since for every graph G, $C_w(G, 2) = 2m$ we have $C_w(D_1[n], 2) = 4 \times 3^{n+1} - 4$ and $C_w(D_2[n], 2) = 2 \times 3^{n+1} - 4$.
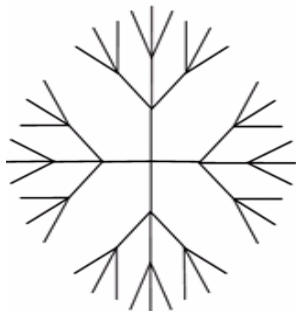
Figure 1. The Forth Generation of Dendrimer Molecule $D_1[4]$
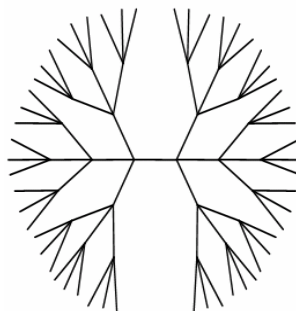
Figure 2. The Forth Generation of Dendrimer Molecule $D_2[4]$

THEOREM 3. $C_w(D_1[n], 4) = 48 \times 3^n + 24$ and $C_w(D_2[n], 4) = 24 \times 3^{n+1} - 66$.

PROOF. Every closed walk of length 4 in the dendrimer molecules $D_1[n]$ and $D_2[n]$ are constructed from one edge or a path of length 2. Therefore, we must count the following type of sequences:

a) $v_1v_2v_1v_2v_1$;
b) $v_1v_2v_3v_2v_1$;
c) $v_2v_1v_2v_3v_2$.

There are $4 \times 3^{n+1} - 4$ sequences of type (a) in $D_1[n]$, $2 \times 3^{n+1} - 4$ sequences of type (a) in $D_2[n]$, $24 \times 3^n + 12$ sequences of type (b) in $D_1[n]$ and $12 \times 3^n - 33$ sequences of type (b) in $D_2[n]$. So, there are $24 \times 3^n + 12$ sequences of type (c) in $D_1[n]$ and $12 \times 3^n - 33$ sequences of type (c) in $D_2[n]$. These facts imply that $C_w(D_1[n], 4) = 48 \times 3^{n+1} + 24$ and $C_w(D_2[n], 4) = 24 \times 3^{n+1} - 66$.

THEOREM 4. $C_w(D_1[n], 6) = 534 \times 3^n - 210$ and $C_w(D_2[n], 6) = 144 \times 3^n - 376$.

Proof. We apply a similar argument as in Theorem 1 to count the number of closed walk of length 6 in $D_1[n]$ and $D_2[n]$. Such walks constructed from an edge, a path of length 2, a path of length 3 or a star S_4. The number of closed walks of length 6 in $D_1[n]$ and $D_2[n]$ on an edge is $4 \times 3^{n+1} - 4$ and $2 \times 3^{n+1} - 4$, respectively. The number of closed walks of length 6 in $D_1[n]$ and $D_2[n]$ on a path with length 2 is $144 \times 3^n - 72$ and $72 \times 3^n - 216$, respectively and the number of closed walks of length 6 in $D_1[n]$ and $D_2[n]$ on a path with length 3 is $294 \times 3^n - 36$ and $18 \times 3^n - 108$, respectively. Finally, the number of closed walks

of length 6 in $D_1[n]$ and $D_2[n]$ on a star S_4 is $84 \times 3^n - 48$ and $48 \times 3^n - 48$, respectively. Therefore, by a simple calculation, one can see that $C_w(D_1[n], 6) = 534 \times 3^n - 210$ and $C_w(D_2[n], 6) = 144 \times 3^n - 376$.

THEOREM 5. Suppose k, k ≥ 8, is an even integer. Then

$$534 \times 3^n - 210 < C_w(D_1[n], k) < 2(4^k \cdot 3^{n+1} - 4^k), \quad (1)$$
$$144 \times 3^n - 376 < C_w(D_2[n], k) < (4^k \times 3^{n+1} - 4^k). \quad (2)$$

PROOF. For proving the left sides of inequalities (1) and (2), we note that $C_w(G, 2k) > C_w(G, 2(k-1))$, k > 3. Thus $C_w(G, 2k) > C_w(G, 6)$ and so $C_w(D_1[n], 2k) > C_w(D_1[n], 6) = 534 \times 3^n - 210$ and $C_w(D_1[n], 2k) > C_w(D_1[n], 6) = 144 \times 3^n - 376$. By an elementary fact in algebraic graph theory, the number of closed walks of length k connecting the i-th and j-th vertices of G is equal to the ij-th entry of A^k, where A denotes the adjacency matrix of G. Therefore, for an arbitrary eigenvalue λ, we have $|\lambda| \leq 4$. Thus, $C_w(D_2[n], k) < 2(4^k \times 3^{n+1} - 4^k)$ and $C_w(D_1[n], k) < (4^k \times 3^{n+1} - 4^k)$. A similar argument proves the same for $D_2[n]$. This completes the proof.

Using calculations given above, it is possible to evaluate the Estrada index of this class of dendrimers. To explain this topological index, we assume that G is a simple graph on n vertices. The adjacency matrix of G is the n × n matrix where the entry a_{ij} is 1 if vertex i is adjacent to vertex j, otherwise a_{ij} is 0. The eigenvalues of the adjacency matrix of G are said to be the eigenvalues of G and to form the spectrum of G [13]. A graph of order n has exactly n eigenvalues not necessarily distinct, but necessarily real-valued. We denote these eigenvalues by $\lambda_1, \lambda_2, \ldots, \lambda_n$. A graph-spectrum-based invariant, recently proposed by Estrada is defined as EE = EE(G) = $\sum_{i=1}^{n} e^{\lambda_i}$ [14-16]. We encourage the interested readers to consult papers [17,18] and references therein for more information on Estrada index and its computational techniques.

THEOREM 6. Consider the molecular graphs of dendrimers $D_1[n]$ and $D_2[n]$. Then there are constants c_i, $1 \leq i \leq 2 \times 3^{n+2} - 1$, and d_j, $1 \leq j \leq 3^{n+1} - 1$, such that $-4 \leq c_i, d_j \leq 4$ and the Estrada index of these graphs are computed as follows:

1) $EE(D_1[n]) = \frac{1769}{120} 3^n - \frac{55}{24} + \sum_{i=1}^{2 \times 3^{n+2}-1} \frac{\lambda_i^8 e^{c_i}}{8!}$,

2) $EE(D_2[n]) = \frac{31}{5} 3^n - \frac{1129}{180} + \sum_{i=1}^{3^{n+1}-1} \frac{\lambda_i^8 e^{d_i}}{8!}$.

CORROLAY 7. The Estrada index of dendrimers $D_1[n]$ and $D_2[n]$ are bounded above as follows:

1) $EE(D_1[n]) \leq \frac{1769}{120} 3^n - \frac{55}{24} + \frac{32e^4}{7!}(2 \times 3^{n+1} - 1)$,

2) $EE(D_2[n]) \leq \frac{31}{5} 3^n - \frac{1129}{180} + \frac{32e^4}{7!}(3^{n+1} - 1)$.

CONCLUSIONS

In this paper, a simple method enabling to compute the closed walks of dendrimers was presented. By our calculation it is possible to evaluate the Estrada index of these dendrimers. It is possible to extend our method in other classes of dendrimers.

ACKNOWLEDGEMENTS

REFERENCES

1. M.V. Diudea, G. Katona, *Molecular Topology of Dendrimers*, in: G.A. Newkome, Ed., *Advan. Dendritic Macromol*., **1999**, *4*, 135.
2. M.V. Diudea, (Ed.): *QSPR/QSAR Studies by Molecular Descriptors*, NOVA, New York, **2001**.
3. G.R. Newkome, C.N. Moorefield and F. Vögtle, *Dendrimers and Dendrons*, Wiley-VCH Verlag GmbH & Co. KGaA**, 2002**.
4. E. Buhleier, W. Wehner and F. Vogtle, *Synthesis,* **1978**, *2*, 155.
5. D.A. Tomalia, H. Baker, J. Dewald, M. Hall, G. Kallos, S. Martin, J. Roeck, J. Ryder and P. Smith, *Polym. J.*, **1985**, *17*, 117.
6. M.H. Khalifeh, H. Yousefi-Azari and A.R. Ashrafi*, Indian J. Chem*., **2008**, *47A*, 1503.
7. A.R. Ashrafi and M. Mirzargar, *Util. Math.,* **2008**, *77*, 249.
8. A. Karbasioun, A.R. Ashrafi and M.V. Diudea, *MATCH Commun. Math. Comput. Chem.,* **2010**, *63*, 239.
9. A.R. Ashrafi and H. Saati, *J. Comput. Theor. Nanosci*., **2008**, *5*, 681.
10. A. Karbasioun and A.R. Ashrafi, *Maced. J. Chem. Chem. Eng*., **2009**, *28*, 49.
11. A. Iranmanesh and N.A. Gholami, *Croat. Chem. Acta,* **2008**, *81*, 299.
12. F. Buckley and F. Harary, *Distance in Graphs*, Addison-Wesley, Reading, MA, **1990**.
13. D. Cvetković, M. Doob and H. Sachs, *Spectra of Graphs – Theory and Application*, third ed., Johann Ambrosius Barth Verlag, Heidelberg, Leipzig, **1995**.
14. E. Estrada, *Chem. Phys. Lett*., **2000**, *319*, 713.
15. E. Estrada, *Bioinformatics,* **2002**, *18*, 697.
16. E. Estrada, *Proteins*, **2004**, *54*, 727.
17. I. Gutman, S. Radenković, A. Graovac and D. Plavšić, *Chem. Phys. Lett*., **2007**, *446*, 233.
18. I. Gutman and A. Graovac, *Chem. Phys. Lett*., **2007**, *436*, 294.

COUNTING POLYNOMIALS OF A NEW INFINITE CLASS OF FULLERENES

MODJTABA GHORBANI[*]

ABSTRACT. Let m(G,c) be the number of strips of length c. The omega polynomial was defined by M. V. Diudea as $\Omega(G,x) = \sum_c m \cdot x^c$. One can obtain the Sadhana polynomial by replacing x^c with $x^{|E|-c}$ in omega polynomial. Then the Sadhana index will be the first derivative of Sd(G,x) evaluated at x = 1. In this paper, the Omega and Sadhana polynomials of a new infinite class of fullerenes is computed for the first time.

Keywords: Fullerene, Omega and Sadhana Polynomials, Sadhana Index.

INTRODUCTION

The discovery of C_{60} bucky-ball, which is a nanometer-scale hollow spherical structure, in 1985 by Kroto and Smalley, revealed a new allotrope of carbon element other than graphite, diamond and amorphous carbon [1,2]. Fullerenes are molecules in the form of cage-like polyhedra, consisting solely of carbon atoms and having pentagonal and hexagonal faces. In this paper, the [4,6] fullerenes C_{8n^2} with tetragonal and hexagonal faces are considered.

Let p, h, n and m be the number of tetragons, hexagons, carbon atoms and bonds between them, in a given fullerene F. Since each atom lies in exactly 3 faces and each edge lies in 2 faces, the number of atoms is n = (4p+6h)/3, the number of edges is m = (4p+6h)/2 = 3/2n and the number of faces is f = p + h. By the Euler's formula n − m + f = 2, one can deduce that (4p+6h)/3 – (4p+6h)/2 + p + h = 2, and therefore p = 6. This implies that such molecules, made entirely of n carbon atoms, have 6 tetragonal and (n/2 − 4) hexagonal faces.

Let G = (V, E) be a connected bipartite graph with the vertex set V = V(G) and the edge set E = E(G), without loops and multiple edges. The distance d(x,y) between x and y is defined as the length of a minimum path between x and y. Two edges e = ab and f = xy of G are called codistant, "e co f", if and only if d(a,x) = d(b,y) = k and d(a,y) = d(b,x) = k+1 or vice versa, for a non-negative integer k. It is easy to see that the relation "co" is reflexive and

---

[*] *Department of Mathematics, Faculty of Science, Shahid Rajaee Teacher Training University, Tehran, 16785 – 136, I R. Iran; mghorbani@srttu.edu*

symmetric but it is not necessary to be transitive. Set $C(e) = \{\, f \in E(G) \mid f \text{ co } e \,\}$. If the relation "co" is transitive on $C(e)$ then $C(e)$ is called an orthogonal cut "oc" of the graph G. The graph G is called a co-graph if and only if the edge set $E(G)$ is a union of disjoint orthogonal cuts. If any two consecutive edges of an edge-cut sequence are topologically parallel within the same face of the covering, such a sequence is called a quasi-orthogonal cut qoc strip. Three counting polynomials have been defined on the ground of qoc strips [3-7]:

$$\Omega(G, x) = \sum_c m \cdot x^c \tag{1}$$

$$\Theta(G, x) = \sum_c m \cdot c \cdot x^c \tag{2}$$

$$\Pi(G, x) = \sum_c m \cdot c \cdot x^{e-c} \tag{3}$$

$\Omega(G, x)$ and $\Theta(G, x)$ polynomials count equidistant edges in G while $\Pi(G, x)$, non-equidistant edges. In a counting polynomial, the first derivative (in x=1) defines the type of property which is counted; for the three polynomials they are:

$$\Omega'(G, 1) = \sum_c m \cdot c = e = \left| E(G) \right| \tag{4}$$

$$\Theta'(G, 1) = \sum_c m \cdot c^2 \tag{5}$$

$$\Pi'(G, 1) = \sum_c m \cdot c \cdot (e - c) \tag{6}$$

The Sadhana index Sd(G) for counting qoc strips in G was defined by Khadikar et al.[8,9] as $Sd(G) = \sum_c m(G, c)(|E(G)| - c)$. We now define the Sadhana polynomial of a graph G as $Sd(G, x) = \sum_c m x^{|E| - c}$. By definition of Omega polynomial, one can obtain the Sadhana polynomial by replacing x^c with $x^{|E| - c}$ in Omega polynomial. Then the Sadhana index will be the first derivative of Sd(G, x) evaluated at x = 1.

A topological index of a graph G is a numeric quantity related to G. The oldest topological index is the Wiener index, introduced by Harold Wiener. Padmakar Khadikar [10,11] defined the Padmakar–Ivan (PI) index as $PI(G) = \sum_{e=uv \in E(G)} [m_u(e|G) + m_v(e|G)]$, where $m_u(e|G)$ is the number of edges of G lying closer to u than to v and $m_v(e|G)$ is the number of edges of G lying closer to v than to u. Edges equidistant from both ends of the edge uv are not counted.

Ashrafi [12,13] introduced a vertex version of PI index, named the vertex PI index and abbreviated by PI_v. This new index is defined as $PI_v(G) = \sum_{e=uv \in E(G)} [n_u(e|G) + n_v(e|G)]$, where $n_u(e|G)$ is the number of vertices of G lying closer to u and $n_v(e|G)$ is the number of vertices of G lying closer to v. If G is bipartite then $n_u(e|G) + n_v(e|G) = n$ and so, $PI_v(G) = n\, |E(G)|$. Throughout this paper, our notation is standard and taken from the standard book of graph theory [14]. We encourage the reader to consult papers by Ashrafi et al and Ghorbani et al [15-23].

RESULTS AND DISCUSSION

The aim of this paper is to compute the counting polynomials of equidistant (Omega, Sadhana and Theta polynomials) of C_{8n^2} fullerenes with $8n^2$ carbon atoms and $12n^2$ bonds (the graph G, Figure 1, is n=2).
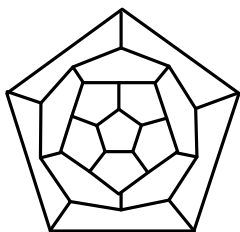
Figure 1. The Fullerene Graph C_{30}.
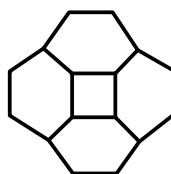
Figure 2. The Carbon Nanocone $CNC_4[1]$ with 16 vertices.

Example 1. Suppose C_{30} denotes the fullerene graph on 30 vertices, see Figure 1. Then $PI_v(C_{30}) = 1090$ and $\Omega(C_{30},x) = x^5 + 10x^2 + 20x$.

Example 2. Consider the carbon nanocones $G = CNC_4[1]$ with 16 vertices, Figure 2. Then $PI_v(G) = 320$ and $\Omega(G,x) = 2x^4 + 4x^3$.

Example 3. Suppose H is the graph of carbon nanocones $CNC_4[2]$ with 36 vertices, see Figure 3. Then $PI_v(H) = 1728$ and $\Omega(H,x) = 2x^6 + 4x^5 + 4x^4$.
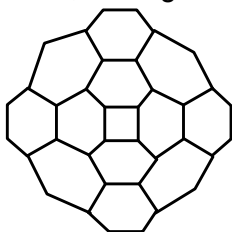
Figure 3. The Carbon Nanocone $CNC_4[2]$ with 36 vertices.
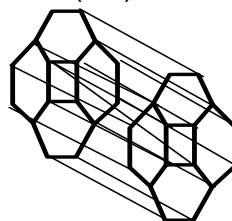
Figure 4. C_{32} obtained from two copies of $CNC_4[1]$.

Example 4. Consider the fullerene C_{32}, Figure 2. One can see that $PI_v(C_{32})=1536$ and $\Omega(C_{32},x) = 3x^8 + 4x^6$.

Lemma. Consider the fullerene graph C_{8n^2}. Then $PI_v(C_{8n^2})= 96n^4$.

Proof. Because the graph is bipartite, by above discussion we have:
$PI_v(G)= |E(G)||V(G) = 96n^4$.

Consider the fullerene graph C_{8n^2} (Figure 4). Its symmetry group is isomorphic to a non-Abelian group of order 96. The orders of elements of its symmetry group are 1, 2, 3, 4 and 6. The center of its symmetry group is isomorphic with the group C_2. In the Appendix one can see how its symmetry

group can be computed by GAP[31] software. We can draw the graph of C_{8n^2} by joining corresponding vertices of two copies of $CNC_4[n-1]$. For example C_{32} is obtained from two copies of $CNC_4[1]$ as follows:

Theorem. $\Omega(C_{8n^2}, x) = 3x^{4n} + 4(n-1)x^{3n}$.

Proof. By Figure 5, there are two distinct cases of qoc strips. We denote the corresponding edges by e_1, e_2, ..., e_{10}. By using Table 1 and Figure 5 the proof is completed.

Table 1. The number of co-distant edges of e_i, $1 \le i \le 5$.

No.	Number of co-distant edges	Type of Edges
3	4n	e_1
4(n-1)	3n	e_2

Corollary. $Sd(C_{8n^2}, x) = 3x^{12n^2-4n} + 4(n-1)x^{12n^2-3n}$.
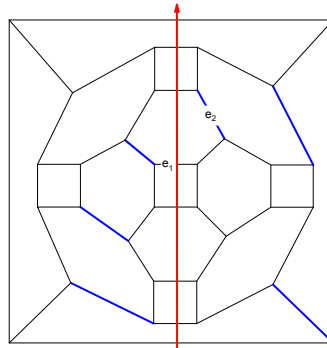
Figure 5. The graph of fullerene C_{8n^2} for n=2.

CONCLUSIONS

Fullerenes are molecules in the form of cage-like polyhedra, consisting solely of carbon atoms. In this paper, by constructing an infinite family of [4,6] fullerenes, we computed Omega and Sadhana polynomials of them for the first time.

Appendix(Symmetry Group of C_{32} Fullerene by GAP Software [31]
a:=(1,2)*(3,4)*(5,6)*(7,8)*(9,10)*(11,12)*(13,14)*(15,16)*(17,18)*(19,20)*(21,22)*(23,24)*(25,26)*(27,28)*(29,30)*(31,32);
b:=(1,3)*(2,4)*(5,7)*(6,8)*(9,11)*(10,12)*(13,15)*(14,16)*(17,19)*(18,20)*(21,23)*(22,24)*(25,27)*(26,28)*(29,31)*(30,32);
c:=(1,4)*(6,7)*(11,24)*(12,22)*(16,30)*(15,32)*(26,27)*(10,21)*(9,23)*(14,29)*(31,13)*(18,19);

```
        d:=(1,2,3,4)*(7,8,6,5)*(21,12,24,9)*(11,22,10,23)*(15,30,14,31)*(16,32,13,2
9)*(27,28,26,25)*(19,20,18,17);G:=Group(a,b,c,d);e:=Elements(G);Print("\n");Print("
e= ",Size(e),"\n");
        dd:=[ 1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23,
24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 64, 65,
66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88,
89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108,
109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125,
126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142,
143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159,
160, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200,
201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217,
218, 219, 226, 227, 228, 229, 230 ];w:=[];ww:=[];tt:=[];

        for i in  dd do
        ff:=Elements(SmallGroup(96,i));
        for j in ff do
        AddSet(w,Order(j));
        if w=[1,2,3,4,6] then AddSet(ww,i);fi;
        od;w:=[];
        od;
        for i in ww do
        if Size(NormalSubgroups(SmallGroup(96,i)))=12 then
        Add(tt,i);
        fi;
        od;
```

REFERENCES

1. H.W. Kroto, J.R. Heath, S.C. O'Brien, R.F. Curl, R.E. Smalley, *Nature*, **1985**, *318,* 162.
2. H.W. Kroto, J.E. Fichier, D.E. Cox, The Fullerene, Pergamon Press, New York, **1993**.
3. M.V. Diudea, S. Cigher, P.E. John, *MATCH Commun. Math. Comput.*, **2008**, *60,* 237.
4. P.E. John, A.E. Vizitiu, S. Cigher, M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2007**, *57,* 479.
5. M.V. Diudea, *Carpath. J. Math*., **2006**, *22,* 43.
6. M.V. Diudea, S. Cigher, A.E. Vizitiu, O. Ursu, P.E. John, *Croat. Chem. Acta*, **2006**, *79,* 445.
7. M.V. Diudea, S. Cigher, A.E. Vizitiu, M.S. Florescu and P.E. John, *J. Math. Chem*., **2009**, *45,* 316.
8. P.V. Khadikar, *Nat. Acad. Sci. Letters*, **2000**, *23*, 113.
9. P.E. John, P.V. Khadikar, J. Singh, *J. Math. Chem*., **2007**, *42*, 37.

10. P.V. Khadikar, S. Karmarkar and V.K. Agrawal, *J. Chem. Inf. Comput. Sci.*, **2001**, *41*, 934.

11. P.V. Khadikar, P.P. Kale, N.V. Deshpande, S. Karmarkar and V.K. Agrawal, *J. Chem. Inf. Comput. Sci.*, **2001**, *29*, 143.

12. A.R. Ashrafi, B. Manoochehrian, H. Yousefi-Azari, *Util. Math.*, **2006**, *71*, 97.

13. B. Manoochehrian, H. Yousefi-Azari, A.R. Ashrafi, *MATCH Commun. Math. Comput. Chem.*, **2007**, *57*, 653.

14. N. Trinajstic, Chemical Graph Theory, 2nd ed.; CRC Press: Boca Raton, FL. **1992**.

15. A.R. Ashrafi, M. Ghorbani and M. Jalali*, Indian J. Chem.*, **2008**, *47A(4)*, 538.

16. A.R. Ashrafi, M. Ghorbani and M. Jalali, *J. Theor. Comput. Chem.*, **2008**, *7(2),* 221.

17. A.R. Ashrafi, M. Jalali, M. Ghorbani and M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2008**, *60(3)*, 905.

18. A.R. Ashrafi, H. Saati and M. Ghorbani, *Digest Journal of Nanomaterials and Biostructures*, **2008**, *3(4)*, 227.

19. M. Ghorbani, and M. Jalali, *Digest Journal of Nanomaterials and Biostructures*, **2009**, *4(1)*, 177.

20. M. Ghorbani, and M. Jalali, *Digest Journal of Nanomaterials and Biostructures*, **2009**, *4(3)*, 403.

21. M. Ghorbani and M. Jalali, *MATCH Commun. Math. Comput. Chem.*, **2009**, *62*, 353.

22. A.R. Ashrafi and M. Ghorbani, *International Journal of Computer Mathematics*, (**2009**) Accepted.

23. M. Ghorbani, and M. Jalali, *Digest Journal of Nanomaterials and Biostructures*, **2009**, *4(2)*, 423.

24. A.R. Ashrafi, M. Ghorbani, M. Jalali, *Digest Journal of Nanomaterials and Biostructures*, **2008**, *3*, 245.

25. A.R. Ashrafi, M. Ghorbani, *Optoelectron. Adv. Mater. – Rapid Comm.*, **2009**, *3(6)*, 596.

26. A.R. Ashrafi, M. Ghorbani, M. Jalali, *Optoelectron. Adv. Mater. – Rapid Comm.*, **2009**, *3(8)*, 823.

27. M. Ghorbani, A.R. Ashrafi, M. Hemmasi, *Optoelectron. Adv. Mater. – Rapid Comm.*, **2009**, *3(12)*, 1306.

28. A.R. Ashrafi and M. Ghorbani, *Optoelectron. Adv. Mater. – Rapid Comm.*, **2009**, *3(3)*, 596.

29. M. Jalali and M. Ghorbani, *Studia Universitatis Babes-Bolyai, Chemia*, 2, **2009**, 2, 145.

30. M.A. Hosseinzadeh and M. Ghorbani, *J. Optoelectron. Adv. Mater*, **2009**, *11(11)*, 1671.

31. GAP, Groups, Algorithms and Programming, (Lehrstuhl De fur Mathematik, RWTH, Aachen), **1992**.

GEOMETRIC–ARITHMETIC INDEX: AN ALGEBRAIC APPROACH

HOSSEIN SHABANI[a], ALI REZA ASHRAFI[a*], IVAN GUTMAN[b]

ABSTRACT. The *Geometric-Arithmetic* (*GA*) index is a recently proposed topological index in mathematical chemistry. In this paper, a group theoretical method for computing the *GA* index of graphs is presented. We apply this method to some classes of dendrimers to calculate their *GA* index.

Keywords: geometric-arithmetic index, dendrimer

INTRODUCTION

A molecular graph is a simple graph such that its vertices correspond to the atoms and the edges to the bonds. Note that hydrogen atoms are often omitted. By IUPAC terminology, a topological index is a numerical value associated with a chemical constitution purporting for correlation of chemical structure with various physical properties, chemical reactivity or biological activity [1–3]. The name "topological index" was first used by Hosoya [4], in connection with his *Z* index, which he used for characterizing the topological nature of graphs.

A dendrimer is generally described as a macromolecule, which is built up from a starting atom, such as nitrogen, to which carbon and other elements are added by a repeating series of chemical reactions that produce a spherical branching structure. In a divergent synthesis of a dendrimer, one starts from the core (a multi connected atom or group of atoms) and growths out to the periphery. In each repeated step, a number of monomers are added to the actual structure, in a radial manner, resulting in quasi concentric shells, called generations. In a convergent synthesis, the periphery is first built up and next the branches (called dendrons) are connected to the core. The stepwise growth of a dendrimer follows a mathematical progression and its size is in the nanometer scale [5–7].

We now recall some algebraic notations that will be used throughout. Suppose *G* is a graph with vertex and edge sets *V*(*G*) and *E*(*G*), respectively. If *e* is an edge of *G*, connecting the vertices *u* and *v* then we write *e* = *uv*. For each vertex *a* and *b*, *d*(*a*,*b*) denotes the length of a minimal path connecting them. The eccentricity of a vertex *x*, $\varepsilon(x)$, is defined as the maximum of {*d*(*y*,*x*) | *y* ∈ *V*(*G*)}.

[a] *Department of Mathematics, Statistics and Computer Science, Faculty of Science, University of Kashan, Kashan 87317-51167, I. R. Iran*
[b] *Faculty of Science, University of Kragujevac, P. O. Box 60, 34000 Kragujevac, Serbia*

Following Vukičević and Furtula [8], the *GA* index of a molecular graph *G* is defined as $GA(G) = \sum_{uv \in E(G)} \frac{2\sqrt{deg(u)deg(v)}}{deg(u) + deg(v)}$, where *deg(u)* denotes the degree of vertex *u* in *G* and the sum is taken over all edges *e* = *uv* of *G*. We encourage the interested readers to consult the papers [9–16] for other variants of this new topological index and their mathematical properties, as well as the reviews [17,18].

In the present article, we continue our works on computing the topological indices of dendrimers [19–21]. Our notation is standard and mainly taken from the standard books of graph theory.

RESULTS AND DISCUSSION

The *GA* index of a molecular graph *G* is based on ratio of the geometric and arithmetic mean and can be computed by considering the automorphism group of *G*. One method to calculate this topological index efficiently is to use group theory and in particular the automorphism group of the graph [23–26]. An automorphism of a graph *G* is an isomorphism of *G* with itself and the set of all such mappings is denoted by *Aut*(*G*).

Figure 1. The All-Aromatic Dendrimer $DNS_1[1]$ and $DNS_1[3]$, respectively.

Figure 2. The Wang's Helicene-Based Dendrimers $DNS_2[2]$ and $DNS_2[3]$, respectively.

In mathematics, groups are often used to describe symmetries of objects. This is formalized by the notion of a group action: every element of the group "acts" like a bijective map (or "symmetry") on some set. To clarify this notion, we assume that *Γ* is a group and *X* is a set. *Γ* is said to act on *X*

when there is a map $\phi : \Gamma \times X \longrightarrow X$ such that all elements $x \in X$, (i) $\phi(e,x) = x$ where e is the identity element of Γ, and (ii) $\phi(g, \phi(h,x)) = \phi(gh,x)$ for all $g,h \in \Gamma$. In this case, Γ is called a transformation group; X is called a Γ -set, and ϕ is called the group action. For simplicity we define $gx = \phi(g,x)$.

In a group action, a group permutes the elements of X. The identity does nothing, while a composition of actions corresponds to the action of the composition. For a given X, the set $\{gx \mid g \in \Gamma\}$, where the group action moves x, is called the group orbit of x. The subgroup which fixes is the isotropy group of x.

Let H and K be two groups and K acts on a set X. The wreath product $H\sim K$ of these groups is defined as the set of all order pair $(f \ ; \ k)$, where $k \in K$ and $f : X \rightarrow H$ is a function such that $(f_1 \ ; \ k_1).(f_2 \ ; \ k_2) = (g \ ; \ k_1 k_2)$ and $g(i) = f_1(i)f_2(i^{k_2})$.

In the following simple lemma a formula for computing the *GA* index of a graph based on the action of $Aut(G)$ on $E(G)$ is obtained.
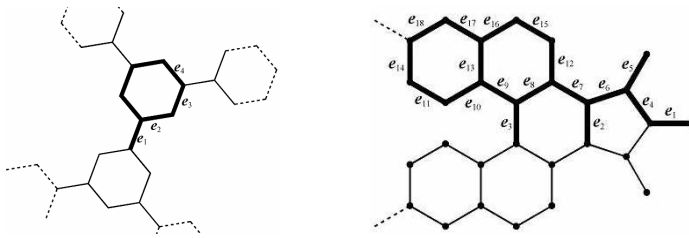
Figure 3. Some Elements of $E_{i,1}$, $E_{i,2}$, $E_{i,3}$, $E_{i,4}$ and .

Lemma. Consider the natural action of $Aut(G)$ on the set of edges containing orbits E_1, E_2, … , E_k. Then $GA(G) = \sum_{i=1}^{k} |E_i| \frac{2\sqrt{deg(u_i)deg(v_i)}}{deg(u_i) + deg(v_i)}$, where $u_i v_i$ is an edge of the i–th orbit. In particular, if the action is transitive and $e = uv$ is an edge of G then $GA(G) = |E(G)| \frac{2\sqrt{deg(u)deg(v)}}{deg(u) + deg(v)}$.

Proof. By definition, for each edge $e_1 = uv$ and $e_2 = xy$ in the same orbit O, there exists an automorphism f such that ($f(u)=x$ & $f(v)=y$) or ($f(u)=y$ & $f(v)=x$). Thus $\frac{2\sqrt{deg(u)deg(v)}}{deg(u) + deg(v)} = \frac{2\sqrt{deg(x)deg(y)}}{deg(x) + deg(y)}$. Since $E(G)$ is partitioned by orbits, $GA(G) = \sum_{i=1}^{k} \sum_{uv \in E_i} \frac{2\sqrt{deg(u)deg(v)}}{deg(u) + deg(v)} = \sum_{i=1}^{k} |E_i| \frac{2\sqrt{deg(u_i)deg(v_i)}}{deg(u_i) + deg(v_i)}$. This completes the proof.

109

We are now ready to calculate the *GA* index of dendrimers depicted in Figures 1 to 3. We have:

Theorem. The *GA* indices of dendrimers depicted in Figures 1 and 2 are as follows:

1. $GA(DNS_1[n]) = 9 \times 2^n + 24(2^n - 1)\sqrt{6}/5 - 3,$

2. $GA(DNS_2[n]) = 2^{n-1}(22 + 16\sqrt{6}/5 + 2\sqrt{3}) - (2^{n-1} - 1)(19 + 24\sqrt{6}/5 + 2\sqrt{3}) + (\sqrt{3} - 1).$

Proof. To compute the *GA* indices of these dendrimers, we first compute the number of their vertices and edges as follows:

$$|V(DNS_1[n])| = 18 \times 2^{n+1} - 12; \ |V(DNS_2[n])| = 27 \times 2^{n+1} - 1$$
$$|E(DNS_1[n])| = 21 \times 2^{n+1} - 15; \ |E(DNS_2[n])| = 33(2^{n+1} - 1)$$

Next we compute the automorphism group of $DNS_1[n]$. To do this, we assume that $T[n]$ is a graph obtained from $DNS_1[n]$ by contracting each hexagon to a vertex. The adjacencies of these vertices are same as the adjacencies of hexagons in $DNS_1[n]$. Choose the vertex x_0 of $T[n]$, associated to the central hexagon, as root. Label vertices of $T[n]$ adjacent to x_0 by 1, 2 and 3; the vertices with distance 2 from x_0 by 4, 5, 6, 7, 8, 9; the vertices with distance 3 from x_0 by 10, 11, 12, …, 21; … and vertices with distance n from x_0 by $3 \times (2^n - 1) + 1$, …, $3 \times (2^{n+1} - 1)$. Set $X = \{1, 2,…, 3 \times (2^{n+1} - 1)\}$. Then S_3 acts on $X = \{1, 2, …, 3 \times (2^n - 1)\}$ and the automorphism group of $DNS_1[n]$ is isomorphic to $Z_2 \sim S_3$, obtained from above action, see Figure 3. Suppose $Aut(DNS_1[n])$ acts on $E(DNS_1[n])$ and $E_{0,0}$, $E_{1,1}$, $E_{1,2}$, $E_{1,3}$, $E_{1,4}$, …, $E_{n,1}$, $E_{n,2}$, $E_{n,3}$, $E_{n,4}$ are orbits of this action. We also assume that H is the central hexagon and $E_{0,0}$ is the set of all edges of H. To obtain the edges $E_{i,1}$, $E_{i,2}$, $E_{i,3}$, $E_{i,4}$ we use the following algorithm:

1. $E_{i,1}$ is the set of all edges $e = uv$ such that $d(u,H) = 3i - 3$, $d(v,H) = 3i - 2$ and $\deg(u) = \deg(v) = 3$, where for each subset $Y \subseteq V(DNS_1[n])$, $d(u,Y) = Min\{d(u,b) \mid b \in Y\}$.
2. $E_{i,2}$ is the set of all edges $e = uv$ such that $d(u,H) = 3i - 2$ and $d(v,H) = 3i - 1$.
3. $E_{i,3}$ is the set of all edges $e = uv$ such that $d(u,H) = 3i - 1$ and $d(v,H) = 3i$.
4. $E_{i,4}$ is the set of all edges $e = uv$ such that $d(u,H) = 3i$, $d(v,H) = 3i + 1$, $deg(u) = 3$ and $deg(v) = 2$.

Obviously, for $DNS_1[n]$ if $e = uv \in E_{i,j}$ then

$$\frac{2\sqrt{deg(u)deg(v)}}{deg(u) + deg(v)} = \begin{cases} 1 & i = 1 \\ 2\sqrt{6}/5 & i = 2, 3, 4 \end{cases}.$$ Moreover, $|E_{i,1}| = 3 \times 2^{i-1}$

and $|E_{i,2}| = |E_{i,3}| = |E_{i,4}| = 6 \times 2^{i-1}$. This completes the proof of (1). To prove 2, it is enough to consider the action of the group $Aut(DNS_2[n])$

on $E(DNS_2[n])$ and use a similar method as given the case 1. Notice that in this case the automorphism group $Aut(DNS_2[n])$ is isomorphic to the wreath product $Z_2 \sim Z_2$, where Z_2 acts on the set $Z = \{1, 2, \ldots, 2^n-1\}$.

ACKNOWLEDGEMENT

REFERENCES

1. L.B. Kier and L.H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Research Studies Press, Latchworth, **1976**.
2. I. Gutman, N. Gavrilović, D. Banković, P.V. Khadikar, N.V. Deshpande and P.P. Kale, *J. Serb. Chem. Soc.*, **1994**, *59*, 519.
3. P.V. Khadikar, I. Lukovits, V.K. Agrawal, S. Shrivastava, M. Jaiswal, I. Gutman, S. Karmarkar and A. Shrivastava, *Indian J. Chem.*, **2003**, *42A,* 1436.
4. H. Hosoya, *Bull. Chem. Soc. Jpn.*, **1971**, *44*, 2332.
5. E. Buhleier, W. Wehner and F. Vogtle, *Synthesis*, **1978**, *2*, 155.
6. D.A. Tomalia, H. Baker, J. Dewald, M. Hall, G. Kallos, S. Martin, J. Roeck, J. Ryder and P. Smith, *Polym. J.*, **1985**, *17,* 117.
7. G.R. Newkome, C.N. Moorefield and F. Vogtle, *Dendrimers and Dendrons*, Wiley−VCH, Weinheim, **2002**.
8. D. Vukičević and B. Furtula, *J. Math. Chem.*, **2009**, *46*, 1369.
9. Y. Yuan, B. Zhou and N. Trinajstić, *J. Math. Chem.*, **2009**, *47*, 833.
10. K.C. Das, I. Gutman and B. Furtula, *MATCH Commun. Math. Comput. Chem.*, **2011**, *65*, 595.
11. H. Hua and S. Zhang, *MATCH Commun. Math. Comput. Chem.*, **2011**, *65*, 691.
12. Z. Tang and Y. Hou, *MATCH Commun. Math. Comput. Chem.*, **2011**, *65*, 705
13. G.H. Fath-Tabar, B. Furtula and I. Gutman: *J. Math. Chem.*, **2010**, *47*, 477.
14. M. Ghorbani and M. Jalali, *Dig. J. Nanomat. Bios.*, **2009**, *4*, 681.
15. M. Ghorbani, *Optoelectron. Adv. Mater.− Rapid Commun.*, **2010**, *4*, 261.
16. M. Gorbani and A. Zaeembashi, *Optoelectron. Adv. Mater.− Rapid Commun.*, **2010**, *4*, 1415.
17. B. Furtula and I. Gutman, in: I. Gutman and B. Furtula (Eds.), *Novel Molecular Structure Descriptors − Theory and Applications I*, Univ. Kragujevac, Kragujevac, **2010**, pp. 137.
18. K. Das, I. Gutman, B. Furtula, *MATCH Commun. Math. Comput. Chem.,* **2010**, *65*, 595.
19. A.R. Ashrafi and M. Mirzargar, *Indian J. Chem.*, **2008**, *47A*, 538.
20. A. Karbasioun and A.R. Ashrafi, *Maced. J. Chem. Chem. Eng.*, **2009**, *28*, 49.

21. A. Karbasioun, A.R. Ashrafi and M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2010**, *63*, 239.
22. A. Karbasioun, A.R. Ashrafi and M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2011**, *65*, 193.
23. S. Yousefi, H. Yousefi-Azari, A.R. Ashrafi and M.H. Khalifeh, *J. Sci. Univ. Tehran*, **2007**, *33*, 7.
24. M.R. Darafsheh, *Acta Appl. Math.*, **2010**, *110*, 1225.
25. M.R. Darafsheh and A. Darafsheh, *MATCH Commun. Math. Comput. Chem.*, **2006**, 56, 271.
26. M.R. Darafsheh, Y. Farjami and A.R. Ashrafi, *MATCH Commun. Math. Comput. Chem.*, **2005**, *54*, 53.

CLUJ AND RELATED POLYNOMIALS IN TORI

MIRCEA V. DIUDEA[a*]**, CSABA L. NAGY**[a]**,
PETRA ŽIGERT**[b]**, SANDI KLAVŽAR**[b,c]

ABSTRACT. Cluj polynomials are defined on the unsymmetric Cluj matrices or by a cutting procedure, as the counting polynomials which collect the vertex proximities in the graph. On these proximities, two Cluj polynomials CJS and CJP and the Plv polynomial can be defined. Formulas for these counting polynomials are derived in case of tori of several tessellation.

Keywords: Cluj polynomial, counting polynomial, torus

INTRODUCTION

Let $G=G(V,E)$ be a simple graph, with no loops and multiple edges and V(G), E(G) be its vertex and edge sets, respectively.

A graph G is a *partial cube* if it is embeddable in the *n*-cube Q_n, which is the regular graph whose vertices are all binary strings of length *n*, two strings being adjacent if they differ in exactly one position [1]. The distance function in the *n*-cube is the Hamming distance. A hypercube can also be expressed as the Cartesian product: $Q_n = W_{i=1}^{n} K_2$

For any edge $e=(u,v)$ of a connected graph G let n_{uv} denote the set of vertices lying closer to *u* than to *v*: $n_{uv} = \{w \in V(G) \mid d(w,u) < d(w,v)\}$. It follows that $n_{uv} = \{w \in V(G) \mid d(w,v) = d(w,u) + 1\}$. The sets (and subgraphs) induced by these vertices, n_{uv} and n_{vu}, are called *semicubes* of *G*; the semicubes are called *opposite semicubes* and are disjoint [2,3].

A graph G is bipartite if and only if, for any edge of *G*, the opposite semicubes define a partition of *G*: $n_{uv} + n_{vu} = v = |V(G)|$. These semicubes are just the vertex proximities (see above) of (the endpoints of) edge $e=(u,v)$, which *CJ* polynomial counts. In partial cubes, the semicubes can be estimated by an orthogonal edge-cutting procedure. The orthogonal cuts form a partition of the edges in *G*:

---

[a] *Universitatea Babeş-Bolyai, Facultatea de Chimie şi Inginerie Chimică,Str. Kogălniceanu, Nr. 1, RO-400084 Cluj-Napoca, Romania, diudea@chem.ubbcluj.ro*
[b] *Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška 160, 2000 Maribor, Slovenia*
[c] *Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia*

$$E(G) = c_1 \cup c_2 \cup \ldots \cup c_k, c_i \cap c_j = \varnothing, i \neq j \, .$$

To perform an orthogonal edge-cut, take a straight line segment, orthogonal to the edge e, and intersect e and all its parallel edges (in a plane graph). The set of these intersections is called an *orthogonal cut* $c_k(e)$ with respect to the edge e. An example is given in Table 1.

To any orthogonal cut c_k, two numbers are associated: first one is the *number of edges* $e_k = |c_k|$ intersected by the orthogonal segment while the second (in round brackets, in Figure 1) is v_k or the number of points lying to the left hand with respect to c_k.

Because in bipartite graphs the opposite semicubes define a partition of vertices, it is easily to identify the two semicubes: $n_{uv} = v_k$ and $n_{vu} = v - v_k$ or vice-versa.

The present study is focused on three counting polynomials of which coefficients can be calculated from the graph proximities/semicubes.

$CJS(x) = 3 \cdot 2 \cdot 3(x^5 + x^{121}) + 3 \cdot 2 \cdot 6(x^{16} + x^{110}) + 3 \cdot 2 \cdot 8(x^{31} + x^{95}) + 3 \cdot 2 \cdot 8(x^{47} + x^{79}) + 3 \cdot 1 \cdot 8(x^{63} + x^{63})$

$CJS'(1) = \mathbf{21924}; \ CJ_e \, S''(1) = 1762320$

$PI_v(x) = 3 \cdot 2 \cdot 3(x^5 + {}^{121}) + 3 \cdot 2 \cdot 6(x^{16} + {}^{110}) + 3 \cdot 2 \cdot 8(x^{31} + {}^{95}) + 3 \cdot 2 \cdot 8(x^{47} + {}^{79}) + 3 \cdot 1 \cdot 8(x^{63} + {}^{63})$

$PI_v'(1) = \mathbf{21924}; \ PI_v''(1) = 2740500$

$CJP(x) = 3 \cdot 2 \cdot 3(x^{5.121}) + 3 \cdot 2 \cdot 6(x^{16.110}) + 3 \cdot 2 \cdot 8(x^{31.95}) + 3 \cdot 2 \cdot 8(x^{47.79}) + 3 \cdot 1 \cdot 8(x^{63.63})$

$CJP'(1) = \mathbf{489090} = CJP(G) = SZ_v'(1)$
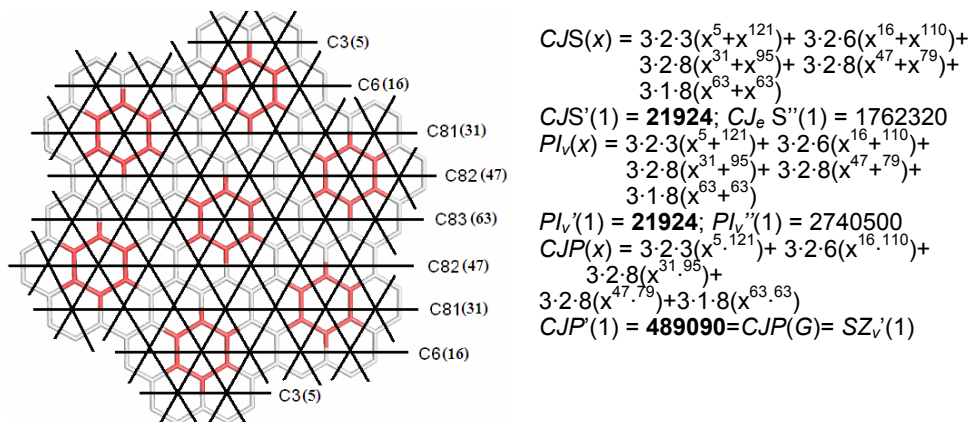
Figure 1. Cutting procedure in the calculus of several topological descriptors.

COUNTING POLINOMIALS OF PROXIMITY

According to the mathematical operations used in composing the edge contributions, these polynomials can be defined as [4]:

(i) *Summation*; the polynomial is called *Cluj-Sum* and is symbolized *CJS* -Diudea *et al* [5-9].

$$CJS(x) = \sum_e \left(x^{v_k} + x^{v - v_k} \right) \qquad (1)$$

(ii) *Pair-wise summation*; the polynomial is called PI_v (vertex-Padmakar-Ivan [10]) - Ashrafi *et al* [11-14].

$$PI_v(x) = \sum_e x^{v_k + (v - v_k)} \qquad (2)$$

(iii) *Pair-wise product*; the polynomial is called *Cluj-Product* (and symbolized *CJP*) [4,8,15-19] or also *Szeged* (and symbolized *SZ*) [12-14]:

$$CJP(x) = SZ(x) = \sum_e x^{v_k(v - v_k)} \qquad (3)$$

Their coefficients can be calculated from the graph proximities / semicubes as shown in Figure 1: the product of three numbers (in the front of brackets – right hand of Figure 1) has the meaning: (i) symmetry of G; (ii) occurrence of c_k (in the whole structure) and (iii) e_k.

The first derivative (in $x=1$) of a (graph) counting polynomial provides single numbers, often called topological indices.

Observe that the first derivative (in $x=1$) of the first two polynomials gives one and the same value (Figure 1), however, their second derivative is different and the following relations hold in any graph [4,7].

$$CJS'(1) = PI_v{}'(1); \ CJS''(1) \neq PI_v{}''(1) \qquad (4)$$

The number of terms, given by $P(1)$, is: $CJS(1)=2e$ while $PI_v(1)=e$ because, in the last case, the two endpoint contributions are pair-wise summed for any edge in a bipartite graph.

In bipartite graphs, the first derivative (in $x=1$) of $PI_v(x)$ takes the maximal value:

$$PI_v{}'(1) = e \cdot v = |E(G)| \cdot |V(G)| \qquad (5)$$

It can also be seen by considering the definition of the corresponding index, as written by Ilić [20]:

$$PI_v(G) = PI_v{}'(1) = \sum_{e=uv} n_{u,v} + n_{v,u} = |V| \cdot |E| - \sum_{e=uv} m_{u,v} \qquad (6)$$

where $n_{u,v}$, $n_{v,u}$ count the non-equidistant vertices with respect to the endpoints of the edge $e=(u,v)$ while $m(u,v)$ is the number of equidistant vertices *vs.* u and v. However, it is known that, in bipartite graphs, there are no equidistant vertices, so that the last term in (6) will disappear. The value of $PI_v(G)$ is thus maximal in bipartite graphs, among all graphs on the same number of vertices; the result of (5) can be used as a criterion for the "bipatity" of a graph [6].

The third polynomial is calculated as the pair-wise product; notice that Cluj-Product $CJP(x)$ is precisely the (vertex) Szeged polynomial $SZ_v(x)$, defined by Ashrafi *et al.* [12-14] This comes out from the relations between the basic Cluj (Diudea [16,21,22]) and Szeged (Gutman [22,23]) indices:

$$CJP'(1) = CJDI(G) = SZ(G) = SZ_v{}'(1) \qquad (7)$$

All the three polynomials (and their derived indices) do not count the equidistant vertices, an idea introduced in Chemical Graph Theory by Gutman [23].

CLUJ POLYNOMIAL IN (4,4), (6,3) AND ((4,8)3) COVERED TORI

In bipartite regular toroidal objects of (4,4), (6,3) and ((4,8)3) tessellation [26,27] (Figure 2) the Cluj and related polynomials (*i.e.*, polynomials counting non-equidistant vertices) and their indices show very simple forms, as given in Table 1. The formulas were obtained by cutting procedures similar to that presented in the introductory section. Note that the studied tori are non-twisted and (with some exceptions) all-even parity of the net parameters [*c,n*].
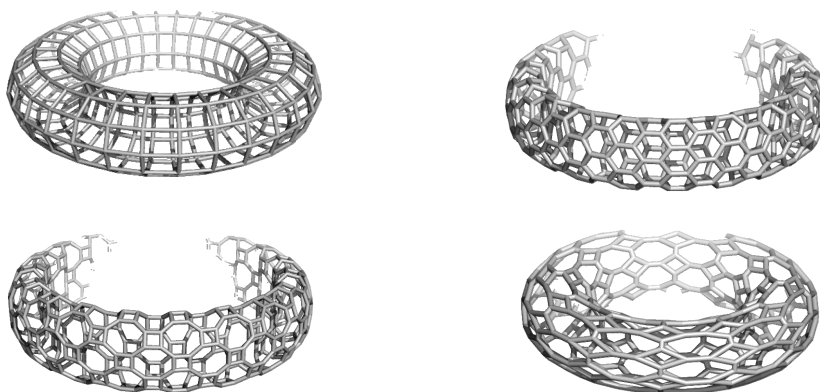
Figure 2. Tori of (4,4); (6,3) (top row) and
((4,8)3)S, ((4,8)3)R (bottom row) covering.

Table 1. Cluj counting polynomials and indices in regular toroidal structures.

$$CJS(x) = e(x^{v/2} + x^{v/2})$$

$$CJS'(1) = e(v/2 + v/2) = e \cdot v = 2(cn)^2$$

$$PI_v(x) = e(x^{v/2+v/2}) = e \cdot x^v$$

$$PI_v'(1) = e \cdot v = CJ_eS'(1)$$

$$CJP(x) = SZ(x) = e(x^{v/2 \cdot v/2})$$

$$CJP'(1) = e(v/2 \cdot v/2) = e(v/2)^2$$

$$= (1/2)v^3 = (1/2)(cn)^3$$

$$v = cn \,; \; e = (d/2)v$$

[c,n]	v	e	$PI_v(x)$	$CJS(x)$	CJS'(1)	SZ'(1)
(4,4); d=4						
10,10	100	200	$200x^{100}$	$400x^{50}$	20000	500000
12,14	168	336	$336x^{168}$	$672x^{84}$	56448	2370816
10,20	200	400	$400x^{200}$	$800x^{100}$	80000	4000000
10,50	500	1000	$1000x^{500}$	$2000x^{250}$	500000	62500000
(6,3); d=3						
H 8,8	64	96	$96x^{64}$	$192x^{32}$	6144	98304
H 8,10	80	120	$120x^{80}$	$240x^{40}$	9600	192000
V 8,26	208	312	$312x^{208}$	$624x^{104}$	64896	3374592
V 8,32	256	384	$384x^{256}$	$768x^{128}$	98304	6291456

[c,n]	v	e	$PI_v(x)$	$CJS(x)$	$CJS'(1)$	$SZ'(1)$
((4,8)3)S; d=3						
20,20 (m=1)	400	600	$600x^{400}$	$1200x^{200}$	240000	24000000
28,42 (m=1)	1176	1764	$1764x^{1176}$	$3528x^{588}$	2074464	609892416
5, 10 (m=8)[a]	400	600	$600x^{400}$	$1200x^{200}$	240000	24000000
7, 21 (m=8)[a]	1176	1764	$1764x^{1176}$	$3528x^{588}$	2074464	609892416
((4,8)3)R; d=3						
10,10 (m=4)	400	600	$600x^{400}$	$1200x^{200}$	240000	24000000
10,20 (m=4)	800	1200	$1200x^{800}$	$2400x^{400}$	960000	192000000
14,21 (m=4)[b]	1176	1764	$294x^{1176} +$ $1176x^{1162} +$ $294x^{1148}$	$1764x^{588} + 1764x^{574}$	2049768	595428792

[a] net design by Le_{22tt}; (m=8)

[b] net designed by Le; (m=4); in case c,n=odd, the graph is non-bipartite

CLUJ POLYNOMIAL IN NAPHTHYLENIC TORI

The naphthylenic pattern [28,29] is an analogue of phenylenic (6,4) pattern [30-32] it shows the ring sequence (6,6,4). Naphthylenic structures can be designed either by a cutting procedure (see above) or by using the map operation sequence: Le(Le(G)), applied on the square tessellation (4,4) embedded in the torus [28,29] We stress that Leapfrog Le operation performed on (4,4) results in the ((4,8)3)R tessellation, with the quadrilaterals disposed as rhombs R. A second iteration of Le operation will provide the naphthylenic pattern [28] eventually named H/VNPX, with the local signature: (4,6,8): (0,4,0);(1,3,2);(0,8,0), Figure 3, left.
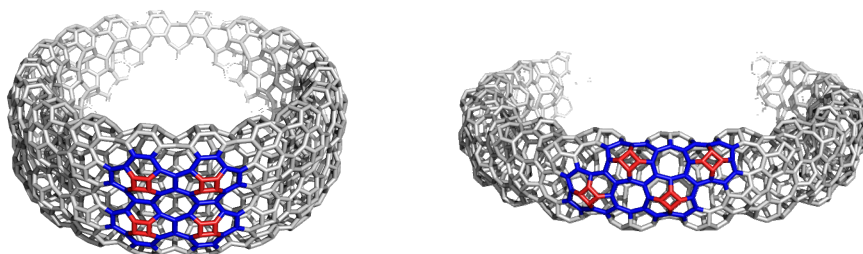
Figure 3. Isomeric Le(((4,8)3)R[6,18]) and Le(((4,8)3)S[12,36]) objects; v=1296; e=1944

Formulas to calculate Cluj and related polynomials, and derived indices as well, in toroidal structures designed by Le(T((4,8)3)R), are given in Table 2. Examples are given at the bottom of the table.

Table 2. Cluj counting polynomials and indices in $Le(Le(4,4))=Le(T((4,8)3)R)$ toroidal structures; c=even; Signature: (4,6,8): (0,4,0);(1,3,2);(0,8,0); m=12. (c=odd; non-Bipartite)

$$CJS(x) = (v/3)(x^{v/2+[a+18(c-4)/2]} + x^{v/2-[a+18(c-4)/2]}) +$$

$$(2v/3)(x^{v/2+[13+10(c-4)/2]} + x^{v/2-[13+10(c-4)/2]}) +$$

$$(v/2)(x^{v/2} + x^{v/2});$$

$$k=1; a=27; \quad k>1; a=31; \quad k=n/c$$

$$CJS'(1) = e \cdot v = (3/2) \cdot v^2 = 216k^2c^4; \quad k=1,2,...$$

$$PI_v(x) = e(x^{v/2+v/2}) = e \cdot x^v$$

$$PI_v'(1) = e \cdot v = CJ_eS'(1)$$

$$CJP(x) = (v/3)(x^{\{v/2+[a+18(c-4)/2]\}\{v/2-[a+18(c-4)/2]\}}) +$$

$$(2v/3)(x^{\{v/2+[13+10(c-4)/2]\}\{v/2-[13+10(c-4)/2]\}}) +$$

$$(v/2)(x^{(v/2)^2});$$

$$k=1; a=27; \quad k>1; a=31; \quad k=n/c$$

$$CJP'(1,k>1) = SZ'(1) = 4kc^2(162k^2c^4 - 131c^2 + 230c - 123)$$

$$CJP'(1,k=1) = SZ'(1) = 4c^2(162c^4 - 131c^2 + 302c - 179)$$

$$v = 12kc^2; \quad k=n/c = 1,2,...$$

$$e = (3/2)v$$

(c,n) v; e	CJS(x)	CJS'(1)	SZ'(1)
(4,8) 384; 576	$128x^{223} + 256x^{205} + 384x^{192} + 256x^{179} + 128x^{161}$	221184	21067392
(6,18) 1296; 1944	$432x^{697} + 864x^{671} + 1296x^{648} + 864x^{625} + 432x^{599}$	2519424	814799088
(8,8) 768; 1152	$256x^{447} + 512x^{417} + 768x^{384} + 512x^{351} + 256x^{321}$	884736	168295680
(10,40) 4800; 7200	$1600x^{2485} + 3200x^{2443} + 4800x^{2400} + 3200x^{2357} + 1600x^{2315}$	34560000	41454523200

When *Le* operation is applied to the ((4,8)3)S tessellation (with the quadrilaterals disposed as squares S) the resulted naphthylenic pattern will show the quadrilaterals disposed as rhombs (Figure 3, right).

As can be seen, the two series show the same tessellation signature (see above) and differ only in the embedding, thus resulting different classes of equivalence and corresponding polynomial terms. The first derivative *CJS*`(1) values are the same in isomeric structures (Tables 2 and 3, the first three rows, next last columns), as a consequence of the bipartity; also it can be considered as a case of (summation operation) degeneracy. In the opposite, the first derivative *SZ*`(1) shows different values (Tables 2 and 3, last columns), the multiplication operation involved being a stronger discriminating operation.

In series *Le*(TH((4,8)3)S), *c*=even, the case *c*=4*m* shows the smallest number of polynomial terms. In series *Le*(TH((4,8)3)R), *c*=even, there is no such a limitation; however, in case *c*=odd of this series, the graphs are non-bipartite and the polynomials show increased number of terms.

Table 3. Cluj counting polynomials and indices in *Le*(TH((4,8)3)S) toroidal structures, *c*=even. Signature: (4,6,8): (0,4,0);(1,3,2);(0,8,0); *m*=3; (*c*=odd; Bipartite)

$$CJS(x) = (v/6)(x^{v/2+[24+15(m-1)+12(k-1)(m+1)]} + x^{v/2-[24+15(m-1)+12(k-1)(m+1)]}) +$$

$$(v/3)(x^{v/2+[16+11(m-1)+12(k-1)(m+1)]} + x^{v/2-[16+11(m-1)+12(k-1)(m+1)]}) +$$

$$(v/6)(x^{v/2+[12+9(m-1)]} + x^{v/2-[12+9(m-1)]}) +$$

$$(v/3)(x^{v/2+[4+5(m-1)]} + x^{v/2-[4+5(m-1)]})$$

$$(v/2)(x^{v/2} + x^{v/2});$$

$$k = n/c; \quad m = (c-4)/4$$

$$CJS'(1) = e \cdot v = (3/2) \cdot v^2 = (27/2)k^2c^4; \quad k = 1, 2, ...$$

$$PI_v(x) = e(x^{v/2+v/2}) = e \cdot x^v$$

$$PI_v'(1) = e \cdot v = CJ_eS'(1)$$

$$CJP(x) = (v/6)(x^{\{v/2+[24+15(m-1)+12(k-1)(m+1)]\} \cdot \{v/2-[24+15(m-1)+12(k-1)(m+1)]\}}) +$$

$$(v/3)(x^{\{v/2+[16+11(m-1)+12(k-1)(m+1)]\} \cdot \{v/2-[16+11(m-1)+12(k-1)(m+1)]\}}) +$$

$$(v/6)(x^{\{v/2+[12+9(m-1)]\} \cdot \{v/2-[12+9(m-1)]\}}) +$$

$$(v/3)(x^{\{v/2+[4+5(m-1)]\} \cdot \{v/2-[4+5(m-1)]\}})$$

$$(v/2)(x^{(v/2) \cdot (v/2)});$$

$$k = n/c; \quad m = (c-4)/4$$

$$CJP'(1) = SZ'(1) = (kc^2/16)(162k^2c^4 - 216k^2c^2 - 12kc^2 - 71c^2 +$$

$$864kc + 480c - 1728)$$

$$v = 3kc^2; \quad k = n/c = 1, 2, ...$$

$$e = (3/2)v$$

(c,n) v; e	CJS(x)	CJS'(1)	SZ'(1)
(8,16) 384; 576	64x²⁴⁰ + 128x²³² + 64x²⁰⁴ + 128x¹⁹⁶ + 384x¹⁹² + 128x¹⁸⁸ + 64x¹⁸⁰ + 128x¹⁵² + 64x¹⁴⁴	221184	20870144
(12,36) 1296; 1944	216x⁷⁵⁹ + 432x⁷⁴⁷ + 216x⁶⁶⁹ + 432x⁶⁵⁷ + 1296x⁶⁴⁸ + 432x⁶³⁹ + 216x⁶²⁷ + 432x⁵⁴⁹ + 216x⁵³⁷	2519424	809267760
(16,16) 768; 1152	128x⁴³⁸ + 256x⁴²² + 128x⁴¹⁴ + 256x³⁹⁸ + 768x³⁸⁴ + 256x³⁷⁰ + 128x³⁵⁴ + 256x³⁴⁶ + 128x³³⁰	884736	168961024

CLUJ POLYNOMIAL IN TiO$_2$ TORI

After the discovery of carbon nanotubes many researchers addressed the question about the possible existence of nano-tubular forms of other elements and they tried to obtain inorganic nanostructures [33-35]. Among numerous oxide nanostructures, the titanium nanotubular materials are of high interest due to their chemical inertness, endurance, strong oxidizing power, large surface area, high photocatalytic activity, non-toxicity and low production cost. The application of TiO$_2$ nanotubes to photocatalysis, in solar cells, as nanoscale materials for lithium-ion batteries and as gas-sensing material was discussed in the literature [36-42]. The nanotubes were synthesized using various precursors [41-47], carbon nanotubes, porous alumina or polymer membranes as templates [40-56] fabrication by anodic oxidation of Ti [57-59], sol–gel technique [60-64] and sono-chemical synthesis [65]. Models of possible growth mechanisms of titanium nanotubes are discussed [48,49,64] but the details of the atomic structure of the nanotube walls and their stacking mode are unknown. TiO$_2$ nanotubes are semiconductors with a wide band gap and their stability increases with increasing of their diameters. The numerous studies for the use of nanotubular titania in technological applications require a lot of theoretical studies of stability and other properties of such structures. Theoretical studies on the stability and electronic characteristics of TiO$_2$ nanostructures were presented in ref. [66-68].

The titanium nanostructures on study below can be achieved by map operations: the sequence consists of *Du*[*Med*(*G*)], applied on polyhex tori or tubes (Figure 4).

Formulas for calculating Cluj and related polynomials, in toroidal TiO$_2$ structures, are given in Table 4.

Figure 4. TiO$_2$ covering embedded in the torus, designed by *Du*(*Med*(H(6,3)[*c*,*n*]))

Table 4. Cluj and Related Polynomials in TiO$_2$ Tori

$$CJS(x) = e(x^{e_{ka}} + x^{e_{kb}})$$
$$CJS'(1) = e(e_{ka} + e_{kb}) = e \cdot v = (1/2)e^2$$
$$PI_v(x) = e(x^{e_{ka}+e_{kb}}) = e \cdot x^v$$
$$PI_v'(1) = e(e_{ka} + e_{kb}) = e \cdot v = CJ_eS'(1)$$
$$CJP(x) = SZ(x) = e(x^{e_{ka} \cdot e_{kb}})$$
$$CJP'(1) = e(e_{ka} \cdot e_{kb}) = e(c/4)^2(3n-2)(3n+2)$$

$$v = (3/2)cn$$
$$e = 3cn$$
$$e_{ka} = e_1 k + (k-1)(c/2)$$
$$e_{kb} = e_{ka} + c$$
$$e_1 = c^2 - (c/2) \cdot (c/2 + 1)$$
$$k = n/c$$

Tori	*CJS*(x)	*CJS*`(1)	*CJP*(x)	*CJP*`(1)
H[10,10]	$300x^{70}+300x^{80}$	45000	$300x^{5600}$	1680000
H[10,20]	$600x^{145}+600x^{155}$	180000	$600x^{22475}$	13485000
H[10,30]	$900x^{220} + 900x^{230}$	405000	$900x^{50600}$	45540000
H[12,14]	$504x^{120} + 504x^{132}$	127008	$504x^{15840}$	7983360
V[8,10]	$240x^{56} + 240x^{64}$	28800	$240x^{3584}$	860160
V[10,20]	$600x^{145}+600x^{155}$	180000	$600x^{22475}$	13485000
V[10,30]	$900x^{220}+900x^{230}$	405000	$900x^{50600}$	45540000
V[10,50]	$1500x^{370}+1500x^{380}$	1125000	$1500x^{140600}$	210900000

ACKNOWLEDGMENTS

REFERENCES

1. F. Harary, *Graph theory*, Addison-Wesley, Reading, MA, 1969.
2. M.V. Diudea, S. Cigher, P.E. John, *MATCH Commun. Math. Comput. Chem.,* **2008**, *60*, 237-250.
3. M.V. Diudea, S. Klavžar, *Carpath. J. Math*., **2009**, accepted.
4. M.V. Diudea, Counting polynomials and related indices by edge cutting procedures, in: I. Gutman, Ed., *New topological descriptors*, MCM Series, *MATCH*, **2010**, accepted.
5. M.V. Diudea, Cluj polynomials. *J. Math. Chem*., **2009**, *45*, 295 -308.
6. M.V. Diudea, A.E. Vizitiu, D. Janežič, *J. Chem. Inf. Model*., **2007**, *47*, 864-874.
7. M.V. Diudea, A. Ilić, M. Ghorbani, A.R. Ashrafi, *Croat. Chem. Acta*, **2009**, accepted.
8. M.V. Diudea, N. Dorosti, A. Iranmanesh, *Carpath. J. Math*., **2009**, accepted.
9. A.E. Vizitiu, M.V. Diudea, *Studia Univ. Babes-Bolyai Chemia*, **2009**, *54*(1), 173-180.
10. P.V. Khadikar, *Nat. Acad. Sci. Lett*., **2000**, *23*, 113-118.
11. MH. Khalifeh, H. Yousefi-Azari, A.R. Ashrafi, *Discrete Appl. Math.* **2008**, *156*, 1780-1789.
12. M.H. Khalifeh, H. Yousefi-Azari, A.R. Ashrafi, *Linear Algebra Appl.*, **2008**, *429*, 2702-2709.
13. A.R. Ashrafi, M. Ghorbani, M. Jalali, *J. Theor. Comput. Chem.,* **2008**, *7*, 221-231.
14. T. Mansour, M. Schork, *Discr. Appl. Math*., **2009**, *157*, 1600-1606.
15. M.V. Diudea, *MATCH Commun. Math. Comput. Chem*., **1997**, *35*, 169-183.
16. M.V. Diudea, *J. Chem. Inf. Comput. Sci*., **1997**, *37*, 300-305.
17. M.V. Diudea, *Croat. Chem. Acta,* **1999**, *72*, 835-851.
18. M.V. Diudea, B. Parv, I. Gutman, *J. Chem. Inf. Comput. Sci*., **1997**, *37*, 1101-1108.
19. M.V. Diudea, Counting polynomials in partial cubes, in: I. Gutman, Ed., *New topological descriptors*, MCM Series, *MATCH*, 2010, accepted.
20. A. Ilić, *Appl. Math. Lett.*, **2009**, submitted.
21. M.V. Diudea, *Croat. Chem. Acta,* **1999**, *72*, 835-851.

22. M.V. Diudea, I. Gutman, L. Jäntschi, *Molecular Topology*, NOVA, New York, 2002.

23. I. Gutman, *Graph Theory Notes New York*, **1994**, *27*, 9-15.

24. A.E. Vizitiu, M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2008**, *60*, 927-933.

25. M.A. Alipour, A.R. Ashrafi, *J. Comput. Theoret. Nanosci.*, **2009**, *6*, 1204-1207.

26. M.V. Diudea, Ed., *Nanostructures, Novel Architecture*, Nova, New York, 2005.

27. M.V. Diudea, Cs.L. Nagy, *Periodic Nanostructures*, Springer, 2007.

28. M.V. Diudea, *Fullerenes, Nanotubes, Carbon Nanostruct.*, **2002**, *10*, 273-292

29. M.V. Diudea, *Phys.Chem.Chem.Phys.*, **2002**, *4*, 4740-4746.

30. B.C. Berris, G.H. Hovakeemian, Y.-H. Lai, H. Mestdagh, K.P.C. Vollhardt, *J. Am. Chem. Soc.*, **1985**, *107*, 5670-5687.

31. I. Gutman, *J. Chem. Soc. Faraday Trans.*, **1993**, *89*, 2413-2416.

32. I. Gutman, S. Klavžar, *ACH Models Chem.*, **1996**, *133*, 389-399.

33. R. Tenne, *Chem. Eur. J.*, **2002**, *8*, 5296-5304.

34. C.N.R. Rao, M. Nath, *Dalton Trans.*, **2003**, *1*, 1.

35. G.R. Patzke, F. Krumeich, R. Nesper, *Angew. Chem., Int. Ed.*, **2002**, *41*, 2447.

36. H. Imai, M. Matsuta, K. Shimizu, H. Hirashima, N. Negishi, *Solid State Ionics*, **2002**, *151*, 183-187.

37. M. Adachi, Y. Murata, I. Okada, S. Yoshikawa, *J. Electrochem. Soc.*, **2003**, *150*, G488.

38. Y. Zhou, L. Cao, F. Zhang, B. He, H. Li, *J. Electrochem. Soc.*, **2003**, *150*, A1246-1249.

39. O.K. Varghese, D. Gong, M. Paulose, K.G. Ong, C.A. Grimes, *Sens. Actuators B*, **2003**, *93*, 338-344.

40. O.K. Varghese, D. Gong, M. Paulose, K.G. Ong, E.C. Dickey, and C.A. Grimes, *Adv. Mater.*, **2003**, *15*, 624-627.

41. C.A. Grimes, K.G. Ong, O.K. Varghese, X. Yang, G. Mor, M. Paulose, E.C. Dickey, C. Ruan, M.V. Pishko, J.W. Kendig, A.J. Mason, *Sensors*, **2003**, *3*, 69-82.

42. G.K. Mor, M.A. Carvalho, O.K. Varghese, M.V. Pishko, C.A. Grimes, *J. Mater. Res.*, **2004**, *19*, 628-634.

43. T. Kasuga, M. Hiramatsu, A. Hoson, T. Sekino, K. Niihara, *Adv. Mater.*, **1999**, *11*, 1307-1311.

44. S. Zhang, J. Zhou, Z. Zhang, Z. Du, A.V. Vorontsov, Z. Jin, *Chin. Sci. Bull.*, **2000**, *45*, 1533.

45. G.H. Du, Q. Chen, R.C. Che, Z.Y. Yuan, L.-M. Peng, *Appl. Phys. Lett.*, **2001**, *79*, 3702.

46. D.-S. Seo, J.-K. Lee, H. Kim, *J. Cryst. Growth*, **2001**, *229*, 428.

47. C.-H. Lin, S.-H. Chien, J.-H. Chao, C.-Y. Sheu, Y.-C. Cheng, Y.-J. Huang, C.-H. Tsai, *Catal. Lett.*, **2002**, *80*, 153.

48. B.D. Yao, Y.F. Chan, X.Y. Zhang, W.F. Zhang, Z.Y. Yang, N. Wang, *Appl. Phys. Lett.*, **2003**, *82*, 281.

49. W. Wang, O.K. Varghese, M. Paulose, C.A. Grimes, *J. Mater. Res.*, **2004**, *19*, 417.

50. I. Sun, L. Gao, Q. Zhang, *J. Mater. Sci. Lett.*, **2003**, *22*, 339.

51. P. Hoyer, *Langmuir*, **1996**, *12*, 1411.

52. H. Imai, Y. Takei, K. Shimizu, M. Matsuda, H. Hirashima, *J. Mater. Chem.,* **1999**, *9*, 2971-2972.

53. S.M. Liu, L.M. Gan, L.H. Liu, W.D. Zhang, H.C. Zeng, *Chem. Mater.,* **2002**, *14*, 1391.

54. Y.-L. Shi, X.-G. Zhang, H.-L. Li, *Mater. Sci. Engin.* A, **2002**, *333*, 239.

55. X.H. Li, W.M. Liu, H.-L. Li, *Appl. Phys. A*, **2003**, *80*, 317.

56. T. Peng, H. Yang, G. Chang, K. Dai, K. Hirao, *Chem. Lett.,* **2004**, *33*, 336.

57. D. Gong, C.A. Grimes, O.K. Varghese, W. Hu, R.S. Singh, Z. Chen, E.C. Dickey, *J. Mater. Res.,* **2001**, *16*, 3331-3334.

58. O.K. Varghese, D. Gong, M. Paulose, C.A. Grimes, E.C. Dickey, *J. Mater. Res.,* **2003**, *18*, 156-165.

59. G.K. Mor, O.K. Varghese, M. Paulose, N. Mukherjee, C.A. Grimes, *J. Mater. Res.*, **2003**, *18*, 2588-2593.

60. B.B. Lakshmi, P.K. Dorhout, C.R. Martin, *Chem. Mater.,* **1997**, *9*, 857-872.

61. T. Kasuga, M. Hiramatsu, A. Hoson, T. Sekino, K. Niihara, *Langmuir,* **1998**, *14*, 3160-3163.

62. S. Kobayashi, K. Hanabusa, N. Hamasaki, M. Kimura, H. Shirai, *Chem. Mater.,* **2000**, *12*, 1523-1525.

63. M. Zhang, Y. Bando, K. Wada, *J. Mater. Sci. Lett.,* **2001**, *20*, 167.

64. Y.Q. Wang, G.Q. Hu, X.F. Duan, H.L. Sun, Q.K. Xue, *Chem. Phys. Lett.,* **2002**, *365*, 427-431.

65. Y. Zhu, H. Li, Y. Koltypin, Y.R. Hacohen, A. Gedanken, *Chem. Commun.,* **2003**, 2616.

66. V.V. Ivanovskaya, A.N. Enyashin, A.L. Ivanovskii, *Mendeleev Comm.,* **2003**, *13*, 5.

67. V.V. Ivanovskaya, A.N. Enyashin, A.L. Ivanovskii, *Russ. J. Inorg. Chem.*, **2004**, *49*, 2.

68. A.N. Enyashin, G. Seifert, *Phys. Stat. Sol.* **2005**, *242*, 7, 1361-1370.

WIENER INDEX OF MICELLE-LIKE CHIRAL DENDRIMERS

HASSAN YOUSEFI-AZARI[a,*], ALI REZA ASHRAFI[b,c],
MOHAMMAD HOSSEIN KHALIFEH[a]

ABSTRACT. A map taking graphs as arguments is called a graph invariant or topological index if it assigns equal values to isomorphic graphs. A dendrimer is an artificially manufactured or synthesized molecule built up from branched units called monomers. In this paper, the Wiener index of the micelle-like chiral dendrimers is computed.

Keywords: *Micelle-like chiral dendrimer, molecular graph, Wiener index.*

INTRODUCTION

The basic assumption for all molecules based hypothesis is that similar molecules have similar activities. This principle is also called Structure-Activity Relationship (SAR). Quantitative Structure Activity Relationship, QSAR, is the process by which a chemical structure is quantitatively correlated with a well defined process, such as biological activity or chemical reactivity.

In mathematical chemistry, molecules are often modeled by graphs named "molecular graphs". A molecular graph is a simple graph in which vertices are the atoms and edge are bonds between them. A topological index for a molecular graph G is a numerical value for correlation of chemical structure with various physical properties, chemical reactivity or biological activity [1]. The Wiener index [2] is the first topological index introduced by Harold Wiener. This index is defined as the sum of all topological distances between the pair vertices. In an exact phrase, if G is a graph and d(x,y) denotes the length of a minimal path connecting vertices x and y of G then $W(G) = \sum_{\{x,y\} \subseteq V(G)} d(x,y)$ will be the Wiener index of G.

Nano-biotechnology is a rapidly advancing area of scientific and technological opportunity that applies the tools and processes of nanofabrication to build devices for studying biosystems. Dendrimers are among the main objects of this new area of science. Here a dendrimer is a synthetic 3-dimensional

[a] *School of Mathematics, Statistics and Computer Science, University of Tehran, Tehran, I. R. Iran*
[b] *Department of Mathematics, Statistics and Computer Science, Faculty of Science, University of Kashan, Kashan 87317-51167, I. R. Iran*
[c] *Institute of Nanoscience and Nanotechnology, University of Kashan, Kashan 87317-51167, I. R. Iran*

macromolecule, prepared in a step-wise fashion from simple branched monomer units, the nature and functionality of which can be easily controlled and varied. The aim of this article is a mathematical study of this class of nano-materials. Cyclopropane and its derivatives are among the most intensely structurally studied molecules. Triangulanes are hydrocarbons consisting of spirofused cyclopropane rings. They are one of the most exotic groups of cyclopropane derivatives. Many of them show fascinating chemical, physical and sometimes biological properties [3].

Diudea and his co-workers [4-12] was the first scientist which considered the topological properties of nanostructures. After leading works of Diudea, some researchers from China, Croatia, Germany, India, Iran, Italy and UK continued these programs to compute distance-based topological indices of nanostructures [13-24].

MAIN RESULTS AND DISCUSSION

Consider the molecular graph of micelle-like chiral dendrimer G[2] depicted in Figure 1(c). We extend this molecular graph to the case that there exists a maximal chain of length n from the core to the end hexagon and denote its molecular graph by G[n]. The aim of this section is to compute the Wiener index of this class of dendrimers.

(a)

(b)

(c)

(d)
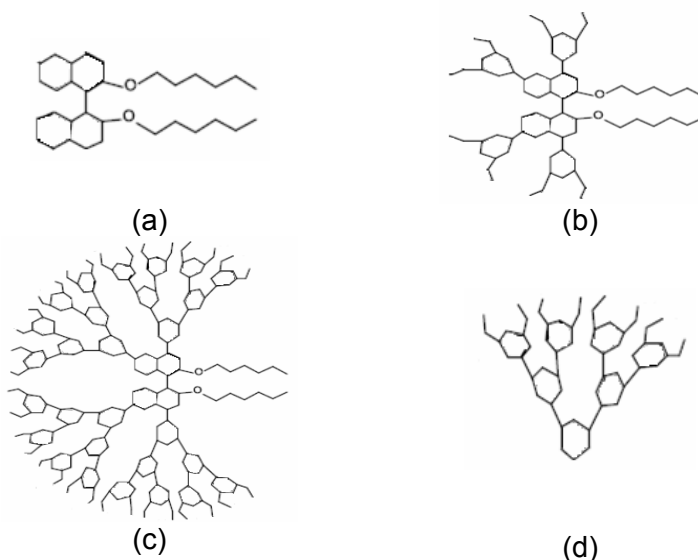
Figure 1. a) The Core of Micelle-Like Chiral Dendrimer G[n]; b) The Molecular Graph of G[0]; c) The Molecular Graph of G[2]; d) A Branch of G[2].

Let G be a simple molecular graph without directed and multiple edges and without loops, the vertex and edge-sets of which are represented by V(G) and E(G), respectively. A path of length n in G is a sequence of n + 1 vertices such that from each of its vertices there is an edge to the next vertex in the sequence. For two vertices x and y of G, d(x,y) denotes the length of a minimal path connecting x and y. A graph G is called connected, if there is a path connecting vertices x and y of G, for every x, y $\in$ V(G).

Suppose X is a set, X_i, $1 \le i \le m$, are subsets of X and F = $\{X_i\}_{1\le i\le m}$ is a family of subsets of X. If X_i's are non-empty, $X = \bigcup_{i=1}^{m} X_i$ and $X_i \cap X_j = \varnothing$, $i \ne j$, then F is called a partition of X. If G is not connected then G can be partitioned into some connected subgraphs, which is called component of G. Here a subgraph H of a graph G is a graph such that V(H) $\subseteq$ V(G) and E(H) $\subseteq$ E(G). A subgraph H of G is called *convex* if x, y $\in$ V(H) and P(x,y) is a shortest path connecting x and y in G then P is a subgraph of H.

Let's start by computing the number of vertices and edges of G[n]. From Figure 1(c), one can easily seen that this graph can be partitioned into four similar branches Figure 1(d) and a core depicted in Figure 1(a). Suppose a_n and b_n denote the number of edges and vertices in each branch of G[n], respectively. Then $a_n = 9 \times 2^{n+1} - 8$ and $b_n = 2^{n+4} - 6$. By Figure 1, one can see that $|V(G[n])| = 4b_n + 34 = 2^{n+6} + 10$ and $|E(G[n])| = 9 \times 2^{n+3} + 9$.

A graph G is called to satisfy the condition (*) if G is connected and there exists a partition $\{F_i\}_{1\le i\le k}$ for E(G) such that for each i, G – F_i has exactly two components, say $G_{i,1}$ and $G_{i,2}$, where they are convex subgraphs of G. The following theorem[25] is crucial in our calculations.

Theorem 1. If G satisfy the condition (*) then $W(G) = \sum_{i=1}^{k} |V(G_{i,1})| \times |V(G_{i,2})|$.

We are now ready to prove our main result. To do this, we first define the notion of parallelism in a graph. The edge $e_1 = xy$ said to be parallel with edge $e_2 = ab$, write $e_1 \| e_2$, if and only if D(x,ab) = D(y,ab), where D(x,ab) = *min*{d(x,a),d(x,b)} and D(y,ab) is defined similarly. In general this relation is not an equivalence relation; even it is not symmetric or transitive. But it is an equivalence relation in the edges of graph G[n] (by a few mathematical background one can see that this equivalence relation defines a partition on E(G[n]) each part being an equivalence class). The equivalence class of G[n] containing the edge e is denoted by [e]. So G[n] satisfies condition (*).

Theorem 2. The Wiener index of G[n] is computed as follows:

$$W(G[n]) = \left(\frac{33}{2}n + 61\right)2^{n+8} + 3\left(n + \frac{5}{8}\right)4^{n+6} + 1189.$$

Proof. Consider the parallelism relation "||" on the edges of G[n]. Since "||" is an equivalence relation on E(G), E(G) can be partitioned into equivalence classes. From Figure 1(c), there are two equivalence classes of size 3 and other classes have sizes 1 or 2. It is also clear that for each edge e $\in$ E(G[n]), G[n] – [e] has exactly two components where each of them is convex, thus we can use the Theorem 1. The hexagons nearest to the endpoints of G[n] are called the end hexagons of G[n].

Consider the subgraph A of G[n] depicted in Figure 2(a) is not an end hexagon. It is easy to see that $F_1 = \{e_7\}$, $F_2 = \{e_1, e_4\}$, $F_3 = \{e_3, e_6\}$ and $F_4 = \{e_2, e_5\}$ are the equivalence classes of A. The components of G[n] – F_1 have b_r and $b_r^c = |V(G[n])| - b_r$ vertices; the components of G[n] – F_4 have $b_r - 3$ and $(b_r - 3)^c$ vertices and the components of G[n] – F_2, G[n] – F_3 have exactly $b_{r-1} - 3$ and $(b_{r-1} - 3)^c$ vertices, where $1 \le r \le n$. One can see that for an arbitrary r, the number of hexagons in the (n – r)-th generation of G[n] is $4 \times 2^{n-r}$.

Next we consider an end hexagon, the subgraph B depicted in Figure 2(b). Then $H_1 = \{e_{11}\}$, $H_2 = \{e_7\}$, $H_3 = \{e_9\}$, $H_4 = \{e_8\}$, $H_5 = \{e_{10}\}$, $H_6 = \{e_2, e_6\}$, $H_7 = \{e_1, e_5\}$ and $H_8 = \{e_3, e_4\}$ are the equivalence classes of B. On the other hand, one of the component G[n] – H_1, G[n] – H_2, …, G[n] – H_8 have exactly 10, 2, 1, 2, 1, 5, 5 and 7 vertices, respectively. Also, one can see the number of end hexagons is 4×2^n.

The Subgraph A	(b) The Subgraph B
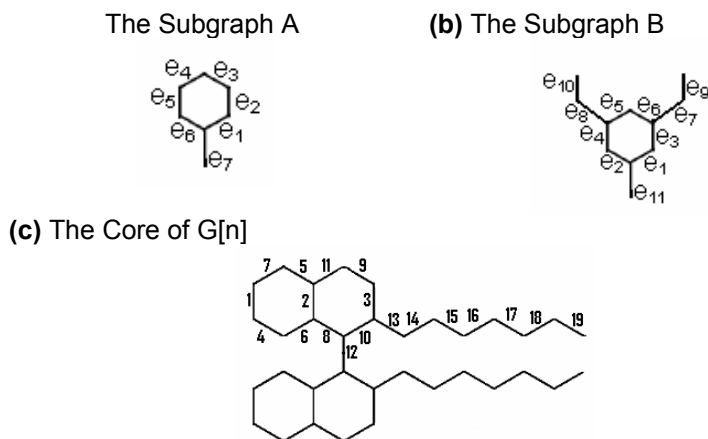
(c) The Core of G[n]

Figure 2. Fragments of the dendrimer G[n]

Finally, we consider the core of G[n], Figure 2(c). The equivalence classes of the core are $X_1 = \{1,2,3\}$, $X_2 = \{4,5\}$, $X_3 = \{6,7\}$, $X_4 = \{8,9\}$, $X_5 = \{10,11\}$, $X_6 = \{13\}$, $X_7 = \{14\}$, $X_8 = \{15\}$, $X_9 = \{16\}$, $X_{10} = \{17\}$, $X_{11} = \{18\}$, $X_{12} = \{19\}$ and

$X_{13} = \{12\}$. Again $G[n] - X_i$, $1 \le i \le 13$, are two component graphs, say $H_{i,1}$ and $H_{i,2}$. Define $a^* = a \times a^c$, a is integer, and $X_i^* = |V(X_{i,1})| \times |V(X_{i,2})|, 1 \le i \le 13$. Then we have the following equalities:

$$X_1^* = (2b_n + 5)^*, \ X_2^* = X_3^* = (b_n + 3)^*, \ X_4^* = (2b_n + 7)^*, \ X_5^* = (b_n + 10)^*,$$

$$X_{13}^* = \frac{|V(G[n])|^2}{4}, \ X_i^* = (13 - i)^*, \ 6 \le i \le 12.$$

Now, applying Theorem 1, we have:

$$W(G[n]) = \sum_{i=1}^{n-1} 4 \times 2^i \left(b_{n-i}^* + (b_{n-i} - 3)^* + 2(b_{n-i-1} + 3)^* \right)$$

$$+ 4 \times 2^n \left(2 \times (1^* + 2^* + 5^*) + 7^* + 10^* \right) + \left(\frac{|V(G[n])|}{2} \right)^* \qquad .$$

$$+ 2 \left[\sum_{i=1}^{7} i^* + (2b_n + 5)^* + (2b_n + 7)^* + 2(b_n + 3)^* + (b_n + 10)^* \right]$$

The proof is now complete by substituting the variables with those given above.

CONCLUSIONS

In this paper a simple method enabling to compute the Wiener index of dendrimers was presented. We apply this method on the molecular graph of a micelle-like chiral dendrimer to obtain an exact formula for the Wiener index of this class of dendrimers. Our method is efficient and can be applied on other classes of dendrimers.

ACLNOWLDGEMENT

REFERENCES

1. C. Hansch, L. Leo, Exploring QSAR Fundamentals and Applicability in Chemistry and Biology; American Chemical Society, Washington DC, **1996**.
2. H. Wiener, *J. Am. Chem. Soc.*, **1947**, *69*, 17.
3. G.R. Newkome, C.N. Moorefield, F. Vögtle, Dendrimers and Dendrons: Concepts, Syntheses, Applications, Wiley- VCH, Weinheim, **2001**.
4. M.V. Diudea, A.A. Kiss, E. Estrada, N. Guevara, *Croat. Chem. Acta*, **2000**, *73*, 367.
5. P.E. John, M.V. Diudea, *Croat. Chem. Acta*, **2004**, *77*, 127.
6. C.L. Nagy, M. Stefu, M.V. Diudea, A. Dress, A. Muller, *Croat. Chem. Acta,* **2004**, *77*, 457.

7. M.V. Diudea, M. Stefu, B. Parv, P.E. John, *Croat. Chem. Acta*, **2004**, *77*, 111.

8. A. Karbasioun, A.R. Ashrafi, M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2010**, *63*, 239.

9. M.V. Diudea, *MATCH Commun. Math. Comput. Chem.,* **2010**, *63*, 247.

10. N. Dorosti, A. Iranmanesh, M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2009**, *62*, 389.

11. M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **1995**, *32*, 71.

12. M.V. Diudea, A.E. Vizitiu, M. Mirzargar, A.R. Ashrafi*, Carpathian J. Math.*, **2010**, *26*, 59.

13. A. Karbasioun, A.R. Ashrafi, *Maced. J. Chem. Chem. Eng.,* **2009**, *28*, 49.

14. M.H. Khalifeh, H. Yousefi-Azari, A.R. Ashrafi, *Dig. J. Nanomater. Bios*., **2009**, *4*, 63.

15. M.H. Khalifeh, H. Yousefi-Azari, A.R. Ashrafi, *Indian J. Chem*., **2008**, *47A*, 1503.

16. A.R. Ashrafi, P. Nikzad, *Dig. J. Nanomater. Bios*., **2009**, *4*, 383.

17. S. Yousefi, A.R. Ashrafi, *J Math Chem*, **2007**, *42*, 1031.

18. A.R. Ashrafi, S. Yousefi, *MATCH Commun Math Comput Chem*, **2007**, *57*, 403.

19. S. Yousefi, A.R. Ashrafi, *Current Nanoscience*, **2008**, *4*, 161.

20. S. Yousefi, A.R. Ashrafi, *MATCH Commun. Math. Comput. Chem*., **2006**, *56*, 169.

21. M.H. Khalifeh, H. Yousefi-Azari, A.R. Ashrafi, *Current Nanosci*., **2010**, *6*, 155.

22. S. Yousefi, A.R. Ashrafi, *Current Nanosci*., **2008**, *4*, 161.

23. A.R. Ashrafi, M. Ghorbani, *Fullerenes, Nanotubes and Carbon Nanostructures*, **2010**, *18*, 198.

24. A. Heydari, B. Taeri, *MATCH Commun. Math. Comput. Chem*., **2007**, *57*, 665.

25. M.H. Khalifeh, H. Yousefi-Azari, A.R. Ashrafi, *Computers and Math. With Appl.,* **2010** doi:10.1016/j.camwa.2010.08.042.

26. M.H. Khalifeh, H. Yousefi-Azari, A.R. Ashrafi, *J. Comput. Appl. Math.*, submitted.

TUTTE POLYNOMIAL OF AN INFINITE CLASS
OF NANOSTAR DENDRIMERS

G. H. FATH-TABAR[*], F. GHOLAMI-NEZHAAD

ABSTRACT. Tutte polynomial T(G,x,y), is a precise topological description of an undirected graph G with two variables, which gives some information about the connectivity of the graph. Dendrimers are hyper-branched nano-structures with rigorously tailored architecture. In this paper, the formula for Tutte polynomial of an infinite nanostar dendrimer is derived.

Keywords: Dendrimer, Tutte polynomial.

INTRODUCTION

Dendrimers are hyper-branched macromolecules showing a rigorous, aesthetically appealing, architecture. They are synthesized, in a controlled manner, basically by two strategies: the *divergent* and *convergent* approaches. In the *divergent methods*, dendrimers are built up starting from a core out to the periphery. In each repeated step, a number of monomer units react with the end groups of the already existing periphery to add a new shell or generation. By each successive generation, the number of local coupling reactions increases. In the *convergent approach*, dendrimers are built from the periphery towards the central core. These rigorously tailored structures show, often at the fifth generation, a spherical shape, which resembles that of a globular protein. The size of dendrimers reaches the nanometer scale. The end groups can be functionalized, thus modifying their physico-chemical or biological properties [1]. Dendrimers have gained a wide range of applications in supramolecular chemistry, particularly in host-guest reactions and self-assembly processes [2-4]; their molecular graphs have been studied by the Mathematical-Chemistry tools [5-7].

Tutte polynomial T(G,*x,y*), is a precise topological description of an undirected graph G with two variables, which gives some information about the connectivity of the graph [8]. In order to define Tutte polynomial we need some notations.

---

[*] Department of Mathematics, Statistics and Computer Science, University of Kashan, Kashan 87317-51167, I. R. Iran

Suppose *G* is an undirected graph with the vertex set V(G) and the edge set E(G). Next, $e \in E(G)$ is neither a loop nor a bridge, then G-*e* is a graph in which the edge *e=uv* has been removed. The *edge contraction* G\e is obtained by linking the endpoints of edge *e=uv* together and making that edge as one vertex, (Figure 3). Then the Tutte polynomial of G is defined by the recurrence relation T(G) = T(G-*e*) + T(G\e). If *G* contains just *i* bridges and *j* loops, T(G,x,y) = $x^i y^j$. Also, T_G = 1 when *G* has no edges. By the above mathematical notations, we have:

$$T(G,x,y) = \begin{cases} 1 & \text{if } E(G) = \varnothing \\ x T(G\text{-}uv, x, y) & \text{if } uv \text{ is a bridge edge} \\ y T(G\text{-}e, x, y) & \text{if } e \text{ is a loop,} \\ T(G\text{-}uv, x, y) + T(G\backslash uv, x, y) & \text{otherwise} \end{cases}$$

In this paper, we derive the formula for computing the Tutte polynomial of the Nanostar Dendrimer Ns[n], Figure 1.
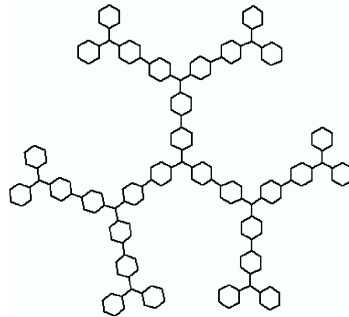
Figure 1. Nanostar dendrimer Ns[2].

MAIN RESULTS

Suppose *G* is an undirected graph with the vertex set V(G) and the edge set E(G). The vertices *v* and *u* of V(G), are in relation α (*v*α*u*), if there is a path in *G* connecting u and v. Each vertex is a path of length zero; so α is a reflexive relation. Moreover, we can easily prove that α is both symmetric and transitive. Thus α is an equivalence relation on *V(G)*, and its equivalence classes are called the *connected components* of G. Then the Tutte polynomial is defined as,

$$T(G, x, y) = \sum_{A \subseteq E} (x-1)^{c(A)-c(E)} (y-1)^{c(A)+|A|-|V|} ,$$

where, c(A) denotes the number of connected components of the graph of the vertex set V and the edge set A.
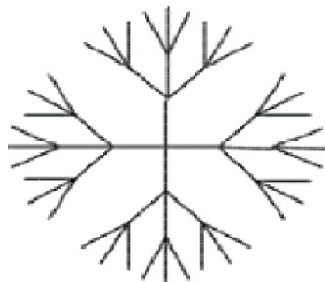
Figure 2. Denderimer D[2].

For example, let G be a tree with n vertices, then $T(G, x, y) = x^{n-1}$, because all the edges in a tree are bridges. The dendrimer D[n] in Figure 2 is a tree with $2\times3^{n+1}-1$ vertices, thus $T(D[n], x, y) = x^{2\times3^{n+1}-2}$.

The Figure 1 has been constructed by joining six Ns[0] units to the hexagons in the outer layers, as detailed in Figures 3 and 4.
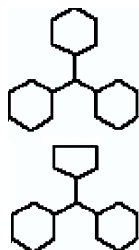
Figure 3. Ns[0] and Ns[0]-H_1+C_5.
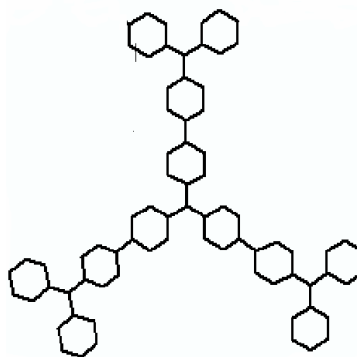
Figure 4. Ns[1].

Lemma 1. Let H be a hexagon. Then $T(D[H], x, y) = \left(\dfrac{x^6 - x}{x-1} + y\right)$.

Proof. By using the formula of Tutte polynomial, we have:

$$
\begin{aligned}
T(D[H], \ x, \ y) &= x^5 + T(D[C_5], \ x, \ y) \\
&= x^5 + x^4 + T(D[C_4], \ x, \ y) \\
&= x^5 + x^4 + x^3 + T(D[C_3], \ x, \ y) \\
&= \frac{x^6 - x}{x-1} + y.
\end{aligned}
$$

To compute the Tutte polynomial of Ns[n], we proceed inductively but at first, we compute T(Ns[0], x, y) in the following

Lemma 2. T(Ns[0], x, y) = $\left(\dfrac{x^6 - x}{x - 1} + y \right)^3 x^3$.

Proof. Suppose H_1, H_2 and H_3 are hexagons in Ns[0]; then

$$T(Ns[0], \ x, \ y) = x^5 T(Ns[0] - H_1, \ x, \ y) + T(Ns[0] - H_1 + C_5, \ x, \ y)$$
$$= x^5 T(Ns[0] - H_1, \ x, \ y) + x^4 T(Ns[0] - H_1, \ x, \ y) +$$
$$T(Ns[0] - H_1 + C_4, \ x, \ y)$$
$$= x^5 T(Ns[0] - H_1, \ x, \ y) + x^4 T(Ns[0] - H_1, \ x, \ y) +$$
$$x^3 T(Ns[0] - H_1, \ x, \ y) + T(Ns[0] - H_1 + C_3, \ x, \ y)$$
$$= \left(\frac{x^6 - x}{x - 1} + y \right) T(Ns[0] - H_1, x, y),$$

where $Ns[0] - H_1 + C_i$ is constructed from Ns[0] by removing H_1 and replacing C_i. As we did in the above,

$$T(Ns[0], \ x, \ y) = \left(\frac{x^6 - x}{x - 1} + y \right)^2 T(Ns[0] - H_1 - H_2, x, y).$$

Thus, T(Ns[0], x, y) = $\left(\dfrac{x^6 - x}{x - 1} + y \right)^3 T(Ns[0] - H_1 - H_2 - H_3, x, y).$ This

implies that

T(Ns[0], x, y) = $\left(\dfrac{x^6 - x}{x - 1} + y \right)^3 x^3$.

Lemma 3. T(Ns[1], x, y) = $x^{15} \left(\dfrac{x^6 - x}{x - 1} + y \right)^{12}$.

Proof. By a similar proof as Lemma 2, we can see that

$$T(Ns[1], x, y) = \left(\frac{x^6 - x}{x - 1} + y \right)^9 x^{12} T(Ns[0], x, y).$$

Thus, T(Ns[1], x, y) = $x^{15} \left(\dfrac{x^6 - x}{x - 1} + y \right)^{12}$.

Theorem 4. T(Ns[n], x, y) = $x^{2 \times 4^{n+1} + 7} \left(\dfrac{x^6 - x}{x - 1} + y \right)^{9 \times 2^n - 6}$.

Proof. Suppose b[n] and h[n] denote the number of bridges and hexagons of Ns[n], respectively. It is easy to see that $b[n] = 2 \times 4^{n+1} + 7$ and $h[n] = 9 \times 2^n - 6$. Thus $b[n] = b[n-1] + 6 \times 4^n$ bridges and $h[n] = h[n-1] + 9 \times 2^{n-1}$ hexagons. Now, by using the definition of Tutte polynomial for bridges and hexagons of Ns[n] - Ns[n-1], and lemma 2, we have

$$T(Ns[n], x, y) = x^{6 \times 4^n} \left(\frac{x^6 - x}{x - 1} + y \right)^{9 \times 2^{n-1}} T(Ns[n-1], x, y).$$

For solving this recursive sequence, we write

$$\prod_{m=2}^{n} \frac{T(Ns[m], x, y)}{T(Ns[m-1], x, y)} = \prod_{m=2}^{n} x^{6 \times 4^m} \left(\frac{x^6 - x}{x - 1} + y \right)^{9 \times 2^{m-1}}.$$

This implies that

$$T(Ns[n], x, y) = x^{2 \times 4^{n+1} - 8} \left(\frac{x^6 - x}{x - 1} + y \right)^{9 \times 2^n - 18} T(Ns[1], x, y).$$

Therefore by Lemma 3,

$$T(Ns[n], x, y) = x^{2 \times 4^{n+1} + 7} \left(\frac{x^6 - x}{x - 1} + y \right)^{9 \times 2^n - 6}.$$

This completes the proof.

REFERENCES

1. M.V. Diudea, G. Katona, *Molecular Topology of Dendrimers*, in: G.A. Newkome, Ed., *Advan. Dendritic Macromol.*, **1999**, *4*, 135.
2. G.R. Newkome, C.N. Moorefield and F. Vögtle, *Dendrimers and Dendrons*, Wiley-VCH Verlag GmbH & Co. KGaA, **2002**.
3. D.A. Tomalia, H. Baker, J. Dewald, M. Hall, G. Kallos, S. Martin, J. Roeck, J. Ryder and P. Smith, *Polym. J.*, **1985**, *17*, 117.
4. E. Buhleier, W. Wehner and F. Vogtle, *Synthesis,* **1978**, *2*, 155.
5. Z. Yarahmadi and G.H. Fath-Tabar, *MATCH Commun. Math. Comput. Chem.,* **2011**, *65*, 201.
6. M.V. Diudea, A.E. Vizitiu, F. Gholami-Nezhaad and A.R. Ashrafi, *MATCH Commun. Math. Comput. Chem.,* **2011**, *65*, 173.
7. A. Karbasioun, A.R. Ashrafi and M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2010**, *63*, 239.
8. H.C. Henry, *Aequationes Mathematicae*, **1969**, *3*, 211.

COMPUTING WIENER AND BALABAN INDICES OF DENDRIMERS BY AN ALGEBRAIC APPROACH

ALI REZA ASHRAFI[a*], HOSSEIN SHABANI[a], MIRCEA V. DIUDEA[b]

ABSTRACT. The Balaban index J is one of the oldest topological indices introduced by the Romanian chemist, A. T. Balaban. The chemical meaning of this topological index was investigated in several research papers. The aim of this paper is to introduce a new algebraic method for computing Wiener and Balaban indices of dendrimers.

Keywords: Dendrimer, Balaban index.

INTRODUCTION

A topological index for a graph G is a number invariant under the automorphism group of G. These numbers have been proposed for the characterization of chemical structures. The Wiener index, one of the oldest descriptors, was proposed by H. Wiener [1]. This topological index is defined as the sum of all distances in the hydrogen-depleted graph representing the skeleton of a molecule [2].

For a connected and simple (molecular) graph G, let $V(G)$ be a finite non-empty set of vertices/atoms and $E(G)$ the set of edges/bonds. The distance between the vertices x and y, $d(x,y)$, is defined as the length of a minimal path connecting x and y. The summation of all distances between a fixed vertex x and all other vertices of G, is denoted by $d(x)$.

The Balaban index is a topological index introduced by Balaban about 30 years ago [3,4]. It is defined as $J(G) = \dfrac{m}{\mu+1} \sum_{e=uv} [d(u)d(v)]^{-0.5}$, where $\mu = m - n + 1$ is called the cyclomatic number of G, with m being the number of edges and n the number of vertices of G. The Balaban index is one of the widely used topological indices for QSAR and QSPR studies, see [5−10] for details.

[a] *Institute for Nanoscience and Nanotechnology, University of Kashan, Kashan, I. R. Iran, Email: ashrafi@kashanu.ac.ir*
[b] *Faculty of Chemistry and Chemical Engineering, "Babes-Bolyai" University,400028 Cluj, Romania*

Two groups of problems for the topological indices associated to a graph can be distinguished. One is to ask the dependence of the index to the graph and the other is the calculation of these indices efficiently. For the Wiener index, the greatest progress in solving the above problems was reported for trees and hexagonal systems in [11-13]. Another method is to use the *Group Theory*, in particular the automorphism group of the graph under consideration [14].

Throughout this paper, our notations are standard and taken mainly from the standard books of graph theory as like as [15]. In this paper we continue our earlier works [16–22] on computing topological indices of dendrimers and derive a formula for the Balaban index of an infinite class of dendrimers. We encourage the reader to consult papers [23–27] for mathematical properties of the Balaban index, as well as basic computational techniques.
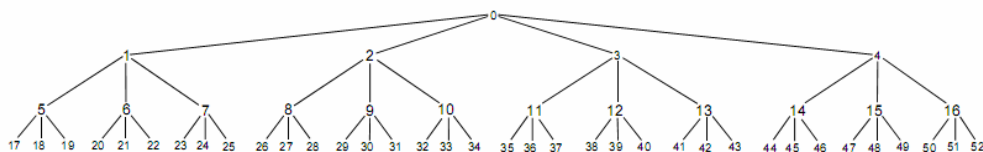
Figure 1. The molecular graph of a dendrimer with 52 vertices.

RESULTS AND DISCUSSION

In the recent years, some topological indices such as the Balaban index has attracted the interest of many chemists, mathematicians and computer scientists and has motivated a large number of research papers involving extremal properties and applications. In this section, we apply an algebraic procedure to obtain formula for computing the Balaban index of an infinite class of dendrimers, Figure 1. For this purpose we need some concepts.

Let G be a connected graph and let v be a vertex of G. The *eccentricity* $e(G)$ of v is the distance to a vertex farthest from v. So, $e(v) = \max\{d(u,v) : u \in V\}$. The *centre* of G we call all vertices with the minimum eccentricity.

Suppose H and K are two groups and K acts on the set Ω. The *wreath product* of $H \sim K$ is defined as the set of all ordered pairs (f, k) where $k \in K$ and $f : \Omega \to H$ is a function, such that $(f_1, k_1)(f_2, k_2) = (g, k_1 k_2)$ and $g(i) = f_1(i) f_2(i^{k_1})$. Observe if Ω, H and K are finite then $|H \sim K| = |H|^{|\Omega|}|K|$. Let's begin by making an isomorphic copy H_k of H for each $k \in K$. Now we can let K act on the right as an automorphism of direct product of all of

these H_k by defining $(a_k).g = a_{kg} \in H_{kg}$ where $g \in K$ and $a_k \in H_k$. So $H \sim K = \oplus_{k \in K} K \propto H$.

Proposition. In the graph G, if $Aut(G)$ acts on $V(G)$ and the orbits of this action are V_0, V_1, ... , V_k then $W(G) = \dfrac{1}{2} \sum_{i=1}^{k} |V_i| d(x_i)$ where $x_i \in V_i$. If $Aut(G)$ acts on $E(G)$ and the orbits of this action are E_1, E_2, ... , E_k then $J(G) = \dfrac{m}{\mu+1} \sum_{i=1}^{k} \dfrac{|E_i|}{\sqrt{d(x_{i-1})d(x_i)}}$ where $x_{i-1}x_i \in E_i$.

Proof. It is sufficient to show that if $\alpha \in Aut(G)$ then $d(u) = d(\alpha(u))$ that is evident.

Define D[k] as the dendrimer molecule depicted in Figure 2. We label the vertices of D[k] by 0, 1, ..., $2 \times (3^k - 1)$. If an edge ij $(i < j)$ is shown by j then the edges of D[k] can be labelled by 1, 2, ..., $2 \times (3^k - 1)$. So, the number of vertices and edges of D[k] are $1 + 2 \times (3^k - 1)$ and $2 \times (3^k - 1)$, respectively.
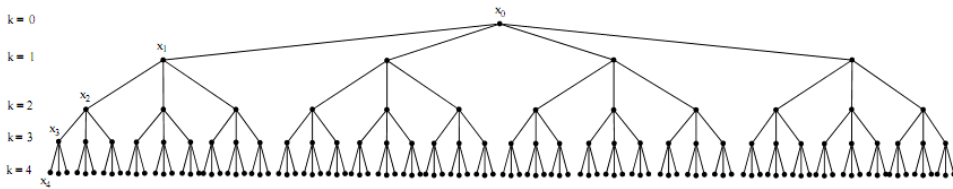
Figure 2. The Dendrimer Molecule D[4].

Theorem. The automorphism group of D[k] is isomorphic to the wreath product $S_3 \sim S_4$ where S_4 act on $\Omega = \{1,2,...,2 \times (3^k - 1)\}$.

Proof. Fix a vertex x_0 as root and assume that $\alpha \in Aut(D[k])$. Then for vertex v in level i, Figure 2, $\alpha(v)$ is also in level i, since v and $\alpha(v)$ have the same eccentricity. Consider the action of S_4 on $\{1,2,...,2 \times (3^k - 1)\}$. Therefore $Aut(D[k])$ is isomorphic to the wreath product of group S_3 via the permutation group S_4. ∎

Corollary 1. The orbits of $Aut(D[k])$ under its natural action on $V(D[k])$ are $V_0 = \{0\}$, $V_1 = \{1,2,3,4\}$, ... , $V_k = \{1 + 2 \times (3^{k-1} - 1),...,2 \times (3^k - 1)\}$.

Let $x_i \in V_i$, $0 \le i \le k$. In each orbit $d(v) = d(x_i)$ when v is a fixed element of V_i. Define:

$$\alpha(t,s) = (1+t).3^0 + (2+t).3^1 + (3+t).3^2 + \cdots + (s+t)3^{s-1}.$$

Then obviously $\alpha(t,s) = 1/4[1 - 2t + (2s + 2t - 1).3^s]$. Therefore,

$$d(x_i) = \alpha(i,k) + \alpha(0, k-i) + \sum_{j=0}^{i}[i - j + 2\alpha(i - j, k - j)]$$

To simplify above equation, we compute $d(x_i)$. We claim that

$$d(x_i) = 1 + (2k - \frac{5}{2}).3^k + (2 \times 3^k).i + (\frac{3^{k+1}}{2}).3^{-i}.$$

We now compute the Wiener and Balaban indices of D[k]. The Wiener index of a graph G is half of the summation of all $d(v)$ over all vertices of G. From the orbits of the action of $Aut(D[k])$ on $V(D[k])$, one can see that $W(G) = \frac{1}{2}\left(\sum_{i=1}^{k}[4 \times 3^{i-1}d(x_i)] + d(x_0)\right)$. So the Wiener index of D[k] is given by the following formula:

$$W(D[k]) = \frac{1}{4} \cdot \frac{1}{(32k^2 - 12k + 1)3^{2k} + (24k - 4)3^k + 4} \cdot (4 + 16k + (144k^2 + 16k)3^k$$

$$+ (416k^3 - 16k^2 - 40k + 5)3^{2k} + (384k^4 - 80k^3 - 44k^2 + 14k - 1)3^{3k})$$

Corollary 2. The orbits of the action of $Aut(D[k])$ on $E(D[k])$ are $E_1 = \{1,2,3,4\}$, ... , $E_k = \{1 + 2 \times (3^{k-1} - 1), \ldots, 2 \times (3^k - 1)\}$.

Since D[k] is a tree, $\mu(D[k]) = 0$ and next,

$$J(D[k]) = 2 \times (3^k - 1)\sum_{i=1}^{k}\frac{4 \times 3^{i-1}}{\sqrt{d(x_{i-1})d(x_i)}}$$

To simplify above equation, we first compute $d(x_{i-1})d(x_i)$. We have:

$$d(x_{i-1})d(x_i) = (1 - 7.3^k + 4k3^k + \frac{45}{4}.3^{2k} - 14k3^{2k} + 4k^2 3^{2k}) +$$

$$(4.3^k - 14.3^{2k} + 8k3^{2k})i + (4.3^{2k})i^2 + (6.3^k - 18.3^{2k} + 12k3^{2k})3^{-i}$$

$$+ (12.3^{2k})i3^{-i} + (\frac{27}{4}.3^{2k})3^{-2i}$$

Define: $f(i,k) = d(x_{i-1})d(x_i) = \alpha + \beta i + \lambda i^2 + \gamma 3^{-i} + \delta i 3^{-i} + \mu 3^{-2i}$, where,

$$\alpha = \alpha(k) = 1 - 7.3^k + 4k3^k + \frac{45}{4}.3^{2k} - 14k3^{2k} + 4k^2 3^{2k}$$

$$\beta = \beta(k) = 4.3^k - 14.3^{2k} + 8k3^{2k}$$

$$\lambda = \lambda(k) = 4.3^{2k}$$

$$\gamma = \gamma(k) = 6.3^k - 18.3^{2k} + 12k3^{2k}$$

$$\delta = \delta(k) = 12.3^{2k}$$

$$\mu = \mu(k) = \frac{27}{4}.3^{2k}$$

Therefore,

$$J(D[k]) = 8.(3^k - 1)\sum_{i=1}^{k} \frac{3^{i-1}}{\sqrt{f(i,k)}}$$

In the following table, the Balaban index $J(D[k])$ is computed, for some k.

Table 1. The Balaban index of D[k], k ≤ 20.

k	$J(D[k])$	k	$J(D[k])$	k	$J(D[k])$	k	$J(D[k])$
1	3.023715783	6	150.3527448	11	17983.21943	16	2.896950741×10^6
2	6.365606476	7	374.3679197	12	48967.03417	17	8.142315294×10^6
3	12.85128466	8	958.8910307	13	1.344850778×10^5	18	2.297691366×10^7
4	27.51789936	9	2509.007693	14	3.720218719×10^5	19	6.506856261×10^7
5	62.72145108	10	6673.758448	15	1.035416212×10^6	20	1.848495377×10^8

CONCLUSIONS

In this paper an algebraic method for computing Balaban index of dendrimers is presented. By this method the Wiener and Balaban indices of an infinite class of dendrimers are calculated. It is possible to extend our method for a general tree. These indices can be used in QSAR/QSPR studies.

REFERENCES

1. H. Wiener, *J. Am. Chem. Soc.*, **1947**, *69*, 17.
2. H. Hosoya, *Bull. Chem. Soc. Jpn.*, **1971**, *4*, 2332.
3. A.T. Balaban, *Chem. Phys. Lett.*, **1982**, *89*, 399.
4. A.T. Balaban, *Pure Appl. Chem.*, **1983**, *55*, 199.

5. A.T. Balaban, *MATCH Commun. Math. Comput. Chem*., **1986**, *21*, 115.

6. A.T. Balaban, P. Filip, *MATCH Commun. Math. Comput. Chem*., **1984**, *16*, 163.

7. A.T. Balaban, N. Ionescu-Pallas, T.S. Balaban, *MATCH Commun. Math. Comput. Chem.,* **1985**, *17*, 121.

8. A.T. Balaban, P.V. Khadikar, C.T. Supuran, A. Thakur and M. Thakur, *Bioorg. Med. Chem*., **2005**, *15*, 3966.

9. T.S. Balaban, A.T. Balaban and D. Bonchev, *J. Mol. Struct*. (*Theochem*), **2001**, *535*, 81.

10. J. Devillers, A.T. Balaban, "Topological Indices and Related Descriptors in QSAR and QSPR", Gordon & Breach, Amsterdam, **1999**.

11. A.A. Dobryanin, R. Entringer, I. Gutman, *Acta Appl. Math*., **2001**, *66*, 211.

12. A.A. Dobryanin, I. Gutman, S. Klavzar, P. Zigert, *Acta Appl. Math*., **2002**, *72*, 247.

13. A.A. Dobryanin, I. Gutmant, *Univ. Berogr. Publ. Elektrotehn. Fak. Ser. Mat*., **1997**, *8*, 106.

14. M.R. Darafsheh, *Acta. Appl. Math*., in press.

15. F. Buckley, F. Harary, "Distance in Graphs", Addison-Wesley, Reading, MA, **1990**.

16. A.R. Ashrafi, M. Mirzargar, *Util. Math*., **2008**, 77, 249.

17. A. Karbasioun, A.R. Ashrafi, *Macedonian J. Chem. Chem. Eng*., **2009**, *28*, 49.

18. A. Karbasioun, A.R. Ashrafi, M.V. Diudea, *MATCH Commun. Math. Comput. Chem*., **2010**, *63*, 239.

19. A.R. Ashrafi, A. Karbasioun, M.V. Diudea, *MATCH Commun. Math. Comput. Chem*., **2011**, *65*, 193..

20. A.R. Ashrafi, H. Saati, *J. Comput. Theoret. Nanosci*., **2008**, *5*, 681.

21. A.R. Ashrafi, M. Mirzargar, *Indian J. Chem*., **2008**, *47A*, 535.

22. M.V. Diudea, A.E. Vizitiu, M. Mirzargar, A.R. Ashrafi, *Carpath. J. Math*., **2010**, *26*, 59.

23. A. Iranmanesh, A.R. Ashrafi, *J. Comput. Theoret. Nanosci*., **2007**, *4*, 514.

24. H. Dong, X. Guo, *MATCH Commun. Math. Comput. Chem*., **2010**, *63*, 799.

25. M. Eliasi, B. Taeri, *J. Comput. Theor. Nanosci*., **2007**, *4*, 1174.

26. B. Zhou, N. Trinajstic, *Croat. Chem. Acta*, **2008**, *81*, 319.

27. B. Zhou, N. Trinajstic, *Croat. Chem. Acta*, **2009**, *82*, 537.

ON DIAMETER OF TUC$_4$C$_8$(S)[P,Q] LATTICE TITLE

MOHAMMAD A. IRANMANESH [a, b], A. ADAMZADEH [c]

ABSTRACT. In this paper we consider a TUC4C8(S)[p,q] nanotube lattice where $q = kp$ and we compute its diameter.

Keywords: TUC4C8 (S)[p,q] Nanotube lattice, Diameter of graph, Dual graph

INTRODUCTION

Let G be a molecular graph with $V(G)$ and $E(G)$ being the set of atoms/vertices and bonds/edges, respectively. The distance between vertices u and v of G is denoted by $d(u,v)$ and it is defined as the number of edges in a path with minimal length connecting the vertices u and v.

A topological index is a numerical quantity derived in an unambiguous manner from the structural graph of a molecule and it is a graph invariant.

The Wiener index of a graph represents the sum of all distances in the graph. Another index, the Padmakar-Ivan (PI) index, is defined as $PI(G) = \sum [n_{eu}(e|G) + n_{ev}(e|G)]$, where $(n_{eu}, e|G)$ is the number of edges of *G* lying closer to *u* than to *v*, and $(n_{eu}, e|G)$ is the number of edges of *G* lying closer to *v* than to *u* and summation goes over all edges of *G*. Also, the Szeged index of a graph G is defined as $Sz(G) = \sum_{e \in E(\Gamma)} n_1(e|G) \cdot n_2(e|G)$, where $n_i(e|G)$ is the number of elements in

$$N(e|G) = \{x \in V(G) \mid d(u,x) < d(v,x)\} \text{ and } e = \{u,v\} \in E(G).$$

For $G = TUC_4C_8(S)[p,q]$ nanotubes (Figure1) the Wiener W, Padmakar-Ivan PI and Szeged Sz indices are topological indices that have been computed in refs. [1-5].

One important application for graphs is to model computer networks or parallel processor connections. There are many properties of such networks that can be obtained by studying the characteristics of the graph models. For example, how do we send a message from one computer to another by using the least amount of intermediate nodes? This question is answered

[a] Corresponding Author
[b] Department of Mathematics, Yazd University, Yazd, P.O. Box 89195-741, I.R.IRAN, Iranmanesh@yazduni.ac.ir
[c] Islamic Azad University Branch of Najaf abad, adamzadeh@iaun.ac.ir

by finding a shortest path in the graph. We may also wish to know what is the largest number of communication links required for any two nodes to talk with each other; this is equivalent to find the diameter of the graph.
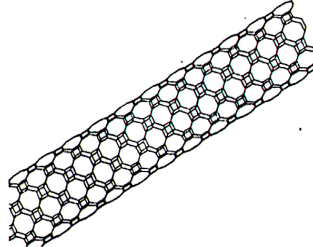
Figure 1: $G = TUC_4C_8(S)[p,q]$ Nanotube

Let $G = TUC_4C_8(S)[p,q]$ lattice (Figure 2) be a trivalent decoration made by alternating squares C_4 and octagons C_8. In [4] the diameter of a zig-zag polyhex lattice have been computed; in this paper the formula for the diameter of a $G = TUC_4C_8(S)[p,q]$ lattice will be derived.
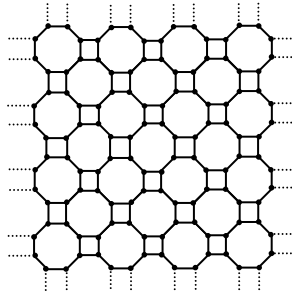
Figure 2: TUC_4C_8 lattice

RESULTS AND DISCUSSION

In this section we compute the diameter of the graph $G = TUC_4C_8(S)[p,q]$ where $q = kp$ and k, p are positive integers.

Definition 1. In $G = TUC_4C_8(S)[p,q]$ with $p = q$ any vertices of degree 2 and all vertices of degree 3 which are adjacent to a vertex of degree 2 is called boundary vertices (see Figure 3); a vertex which is not a boundary vertex is called an interior vertex. The set of all boundary vertices of G are denoted by $B(G)$.

Lemma 1. Let $d = d(G)$ be the diameter of G. Then for any $v - w$ path of length d, we have $\{v, w\} \in B(G)$.

The Proof is given by Lemma 2 [4].

Lemma 2. Suppose B_ℓ and B_β are the sets of left side and right side boundary vertices of G, respectively. If we sort the vertices of these sets from up to down such that $B_\ell = \{u_1, u_2, \ldots, u_n\}$ and $B_\beta = \{v_1, v_2, \ldots, v_n\}$, then $d(u_1, v_n) = d(G)$.

Proof. We consider the inner dual of $G = TUC_4C_8(S)[p,q]$, (Figure 3)
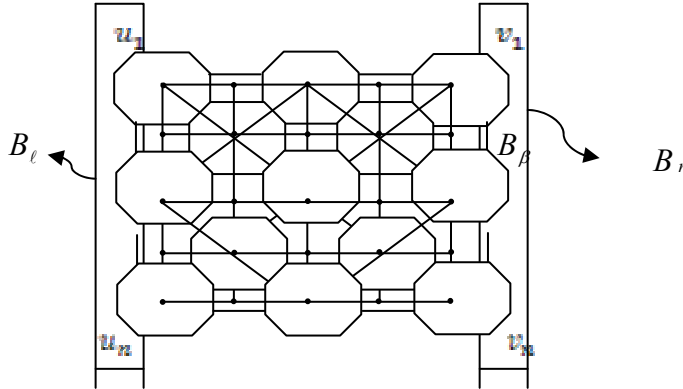
Figure 3: TUC₄C₈(S) lattice and its dual graph

Since

$$d(u_i, v_n) \le d(u_1, v_n),$$
$$i = 1, 2, \ldots, n$$

and

$$d(u_n, v_i) \le d(u_n, v_1).$$
$$i = 1, 2, \ldots, n$$

Then by symmetry of G we conclude that $d(G) = d(u_1, v_n) = d(u_n, v_1)$.

Theorem 1. Let $G = TUC_4C_8(S)[p,q]$, where $q = p$ and p, q are positive integers. Then we have $d(G) = 5p - 1$.

Proof. The proof is by induction on p. In case $p = 1$, G is an octagon with diameter 4 and the theorem is obviously true. Suppose that the theorem is true for the case $p = q = n - 1$ and consider a TUC₄C₈(S) [p,q] lattice of size $n \times n$. By Lemma 2, we have $d(G) = d(u_1, v_n)$. If we delete the last two rows and columns it is easy to see that $d(u_1, v_n) = d(u_1, v_{n-1}) + d(v_{n-1}, v_n)$. Hence,

$$d(G) = [5(n-1) - 2] + 6 = 5p - 1.$$

Theorem 2. Let, $G = TUC_4C_8(S)[p,q]$, where $q = kp$ and p, q are positive integers. Then we have $d(G) = (4k+1)p - 1$.

Proof. We prove the theorem by induction with respect to k. As *we* have seen for Theorem 1, the assertion is true for $k = 1$. Let $k > 1$ and suppose that the theorem is true for $q = kp$. Now consider a $G = TUC_4C_8(S)[p,q]$, lattice with $q = (k+1)p$. We may assume that the lattice contains $k+1$ blocks

145

$B_1, B_2, \ldots, B_{k+1}$ of $G = TUC_4C_8(S)[p, q]$ lattice each of size $(p \times p)$. Obviously it is enough to find the length of shortest path from vertex u_{11} to vertex $v_{(k+1)(p+1)}$.

By symmetry of $G = TUC_4C_8(S)[p, q]$ lattice, we conclude that $d(G) = d(u_{11}, v_{kp}) + d(v_{kp}, v_{(k+1)(p+1)}) = d(u_{11}, v_{kp}) + 4k + 2$.
Hence $d(G) = (4k + 1)p - 1$.

EXAMPLES

In this section, we give some examples in the following tables. The diameter calculations were done by the TOPO-CLUJ software package [6]. In Table 1 we consider some special cases where $k = 1$, while in Table 2, we consider cases for $k > 1$.

Table 1. Some cases of $d(G)$ with $k = 1$.

p	$d(G)$	$5p - 1$
2	9	$5(2) - 1 = 9$
3	14	$5(3) - 1 = 14$
4	19	$5(4) - 1 = 19$

Table 2. Some cases of $d(G)$ with $k > 1$.

k	p	q	$d(G)$ (distance matrix)	$p(4k + 1) - 1$ (Theorem 2)
2	3	6	26	3(8+1)-1=26
3	3	9	38	3(12+1)-1=38
3	4	12	51	4(12+1)-1=51
4	3	12	50	3(16+1)-1=50
4	4	16	67	4(16+1)-1=67

ACKNOWLEDGMENTS

REFERENCES

1. A. Ashrafi, A. Loghman, Padmakar-Ivan, *Journal of Computational and Theoretical Nanoscience*, **2006**, *3*, 378.
2. A. Heydari and B. Taeri, *MATCH Commun. Math. Comput. Chem.*, **2007**, *57*, 665.
3. A. Iranmanesh, Y. Pakravesh, *Ars. Comb.*, **2007**, 89.
4. M.A. Iranmanesh, A. Adamzadeh, *Journal of Computational and Theoretical Nanoscience*, **2008**, *5*, 1428.
5. M. Stefu, M.V. Diudea, *MATCH, Commun. Math. Comput. Chem.,* **2004**, *50*, 133.
6. O. Ursu, M.V. Diudea, TOPOCLUJ software program, Babes-Bolyai University, Cluj, **2005**; Available, on line at http://chem.ubbcluj.ro/~diudea.

146

EVALUATION OF THE AROMATIC CHARACTER OF λ^3-HETEROBENZENES USING THE MAGNETIC CRITERION

RALUCA O. POP[a], MIHAI MEDELEANU[b], MIRCEA MRACEC[c]

ABSTRACT. The aromaticity of the heterobenzenes containing 15-group elements, namely the six-membered homocycles $(\lambda^3\text{-}X)_6$ (X = CH, N, P, As) and the six-membered alternant heterocycles $(\lambda^3\text{-}X\text{-} \lambda^3\text{-}Y)_3$ (X, Y = CH, N, P, As) is investigated using magnetic and geometric criteria at B3LYP/6-31G (d,p) level.

The NMR isotropic shielding values of a molecule of H_2 placed at a distance of 2.5Å above the center of the heterobenzenes were computed using the GIAO method at 6-31G(d,p) level. The results are compared with the negative values of the absolute magnetic shielding, computed at the center of ring (NICS index), at B3LYP/6-31G (d,p) level. An angle-based index is used as geometric criterion for the evaluation of the aromatic character.

Keywords: heterobenzenes, aromaticity, shielding increments, angle-based index

INTRODUCTION

The benzene ring, with its delocalized 6 π-electrons, planar geometry, bond lengths equalization and ability of sustaining a diatropic ring current remains the key-molecule for the evaluation of the aromatic character [1]. There are also known hetero-analogues of benzene, compounds obtained by replacing one or many CH groups with isovalent atoms that show similar properties with the above-mentioned ones (an example is the pyridine). Hetero-analogues of benzene with dicoordinated trivalent atoms of 15 group (N, P, As) of type $(\lambda^3\text{-}X\text{-}\lambda^3\text{-}Y)_3$ (where X, Y = CH, N, P, As - Figure 1) have been investigated in order to evaluate their possible aromatic character. Our previous papers [2, 3] reported studies regarding the stability of homo- and heterobenzenes containing 15 group atoms, estimated by their heat of formation computations [2] and also an evaluation of the aromatic character of the heterobenzenes series by means of geometric, magnetic and chemical reactivity criteria [3].

a   *National Institute of R&D for Electrochemistry and Condensed Matter, Str. A. P. Podeanu, Nr. 144, RO-300569 Timisoara, Romania, ralucapop24@gmail.com*
b *University "Politehnica" Timisoara, Faculty of Industrial Chemistry and Environmental Engineering, P-ta Victoriei, nr. 2, RO-300006 Timisoara, Romania, mihai.medeleanu@chim.utt.ro*
c *Institute of Chemistry Timisoara of Roumanian Academy, Bd. Mihai Viteazu, Nr. 24, RO-300223 Timisoara, Romania, mracec@acad-icht.tm.edu.ro*

Concerning the use of the magnetic criteria to evaluate the aromaticity, one of the most used indices is NICS (Nucleus Independent Chemical Shift) [4], computed at the center of the molecule (NICS(0)) or at 1Å distance above the molecule's geometric center. Recent work of Martin et al. [5, 6] report the use of computed NMR shielding effects on a molecular probe (a diatomic hydrogen molecule, placed at a 2-2.5 Å above the geometric center of the ring) in order to evaluate the aromatic character of fused hydrocarbons. The same methodology was applied to the heterobenzenes of series $(\lambda^3\text{-X-}\lambda^3\text{-Y})_3$ (where X, Y = CH, N, P, As) and the shielding values of different heterobenzenes on the diatomic hydrogen probe (placed at 2.5 Å above the center of molecule) were compared with NICS(2.5) values. Moving the H_2 probe along x and y coordinates of the geometric center of the rings (step size 0.5 Å, three steps in each direction) leads to a complete evaluation of the shielding/deshielding effects.
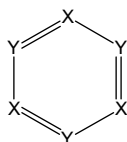
X, Y: CH, N, P, As

Figure 1. General formula of the heterobenzenes $(\lambda^3\text{-X-}\lambda^3\text{-Y})_3$

RESULTS AND DISCUSSION

In a previous paper [3] we reported the evaluation of the aromatic character of the same heterobenzenes series using bond length- and bond order-based indices. No concluding results were obtained, mostly due to the D_{3h} and D_{6h} symmetry of the compounds that strongly influences the values of bond length- and bond order-based indices. A possible method for avoiding such situation is the use of geometric angle-based indices.

Doerksen et al. [7, 8] mentioned two indices based on alternance of the angles, namely AVGNEXT and ADIFFPC and they assigned the maximum value (100) to benzene ring and the minimum value (0) to a six-membered cycle with alternant angles of 110° and 130°, respectively. For the heterobenzene series of this study, the minimum value will correspond to a six-membered cycle with alternant angles of 100° and 135°, respectively.

$$AVGNEXT = 100\left[1 - \frac{1}{35n}\sum_i^n |\theta_i - \theta_{i+1}|\right] \qquad (1)$$

$$ADIFFPC = 100\left[1 - \frac{1}{1837.5}\sum_i^n (\theta_{med} - \theta_i)^2\right] \qquad (2)$$

Table 1. Angle-based geometric indices, using geometries
optimized at *ab initio* level (HF/6-31G)

Heterobenzene	AVGNEXT	ADIFFPC
C_6H_6	100	100
$(CH-\lambda^3-N)_3$	71.4	90.1
$(CH-\lambda^3-P)_3$	22.1	38.7
$(CH-\lambda^3-As)_3$	17.1	31.3
$(\lambda^3-N)_6$	100	100
$(\lambda^3-N-\lambda^3-P)_3$	38.0	61.6
$(\lambda^3-N-\lambda^3-As)_3$	41.1	65.3
$(\lambda^3-P)_6$	100	100
$(\lambda^3-P-\lambda^3-As)_3$	97.1	99.9
$(\lambda^3-As)_6$	100	100

As it can be seen, the maximum value of both indices is obtained for the homocycles $(\lambda^3-X)_6$ but, even in this case, the angle-based geometric index is influenced by the symmetry of molecules.

NICS(2.5) values and shielding increments $\Delta\sigma$ were computed for the heterobenzenes series $(\lambda^3-X-\lambda^3-Y)_3$ (where X, Y = CH, N, P, As) at B3LYP/6-31G level (Table 2).

Table 2. NICS (2.5) (ppm) and shielding increments $\Delta\sigma$ (ppm)
computed at *ab initio* level (B3LYP/6-31G)

Heterobenzene	NICS 2.5 (ppm)	$\Delta\sigma$ (ppm)
C_6H_6	-0.2935	0.974
$(CH-\lambda^3-N)_3$	-2.7228	3.707
$(CH-\lambda^3-P)_3$	-3.6526	4.439
$(CH-\lambda^3-As)_3$	-3.8539	4.478
$(\lambda^3-N)_6$	-2.9665	4.353
$(\lambda^3-N-\lambda^3-P)_3$	-1.6690	2.647
$(\lambda^3-N-\lambda^3-As)_3$	-2.0210	2.362
$(\lambda^3-P)_6$	-5.7005	6.151
$(\lambda^3-P-\lambda^3-As)_3$	-6.7035	6.447
$(\lambda^3-As)_6$	-18.6155	7.203

A qualitative comparison of NICS(2.5) index and shielding increments $\Delta\sigma$ values shows a very similar trend. Higher values (even greater than for the benzene ring) are recorded for the heterobenzenes $(\lambda^3-P)_6$, $(\lambda^3-P-\lambda^3-As)_3$, $(\lambda^3-As)_6$ and predict a pronounced aromatic character. Our previous study [3] regarding the estimation of aromaticity of the heterobenzenes series using the magnetic (NICS(0) and NICS(1)) and reactivity-based criteria showed similar results; the highest NICS values belong to the heterobenzenes having the closest geometry to the benzene ring (equalization of bond lengths and bond angles), namely: $(\lambda^3-N)_6$, $(\lambda^3-P)_6$, $(\lambda^3-P-\lambda^3-As)_3$ and $(\lambda^3-As)_6$.

The same explanation seems to be appropriate for the values obtained for the shielding increments. The results presented in Table 1 show that all the heteroanalogues of benzene have positive shielding increments, revealing a possible aromatic character in all of the cases. The largest values of both NICS(2.5) index and $\Delta\sigma$ are obtained in case of heterobenzenes $(\lambda^3\text{-P})_6$, $(\lambda^3\text{-P-}\lambda^3\text{-As})_3$ and $(\lambda^3\text{-As})_6$.

Shielding increments have also been computed at three different steps in x and y directions (0.5 Å, 1.0 Å, 1.5 Å); the results are presented in Table 3 and Table 4.

Table 3. Shielding increments $\Delta\sigma$ (ppm) computed in three different points along the x axis (B3LYP/6-31G level)

Heterobenzene	$\Delta\sigma_{(x=0.5)}$ (ppm)	$\Delta\sigma_{(x=1.0)}$ (ppm)	$\Delta\sigma_{(x=1.5)}$ (ppm)
C_6H_6	0.974	0.938	0.896
$(CH\text{-}\lambda^3\text{-N})_3$	3.444	2.789	2.056
$(CH\text{-}\lambda^3\text{-P})_3$	4.147	3.385	2.463
$(CH\text{-}\lambda^3\text{-As})_3$	4.173	3.381	2.446
$(\lambda^3\text{-N})_6$	4.076	3.373	2.520
$(\lambda^3\text{-N-}\lambda^3\text{-P})_3$	2.513	2.162	1.678
$(\lambda^3\text{-N-}\lambda^3\text{-As})_3$	2.232	1.902	1.467
$(\lambda^3\text{-P})_6$	5.925	5.325	4.509
$(\lambda^3\text{-P-}\lambda^3\text{-As})_3$	6.215	5.600	4.778
$(\lambda^3\text{-As})_6$	6.690	6.315	5.453

Table 4. Shielding increments $\Delta\sigma$ (ppm) computed at three different points along the y axis (B3LYP/6-31G level)

Heterobenzene	$\Delta\sigma_{(y=0.5)}$ (ppm)	$\Delta\sigma_{(y=1.0)}$ (ppm)	$\Delta\sigma_{(y=1.5)}$ (ppm)
C_6H_6	1.009	0.936	0.898
$(CH\text{-}\lambda^3\text{-N})_3$	3.445	2.804	2.065
$(CH\text{-}\lambda^3\text{-P})_3$	4.158	3.456	2.573
$(CH\text{-}\lambda^3\text{-As})_3$	4.188	3.472	2.601
$(\lambda^3\text{-N})_6$	4.076	3.375	2.521
$(\lambda^3\text{-N-}\lambda^3\text{-P})_3$	2.505	2.119	1.579
$(\lambda^3\text{-N-}\lambda^3\text{-As})_3$	2.224	1.867	1.391
$(\lambda^3\text{-P})_6$	5.925	5.308	4.395
$(\lambda^3\text{-P-}\lambda^3\text{-As})_3$	6.219	5.591	4.670
$(\lambda^3\text{-As})_6$	6.961	6.300	5.323

The results presented in Table 2 and Table 3 show insignificant differences between the values of the shielding increments computed along x and y axes and this a possible consequence of the symmetry of the studied heterobenzenes.

CONCLUSIONS

Computations of AVGNEXT and ADIFFPC indices lead to maximum values for the benzene ring, as well as for the homocycles $(\lambda^3\text{-X})_6$. The results prove the difficulties that appear in quantifying the aromatic character of the heterobenzene series $(\lambda^3\text{-X-}\lambda^3\text{-Y})_3$ by using the geometric criterion. In the absence of any experimental data, it is difficult to make statements regarding the identical values obtained for the geometric indices in case of benzene, hexazine (known as unstable compound) and hexaphosphabenzene and hexaarsabenzene.

Using the magnetic criterion for evaluating the aromaticity, the results of NICS(2.5) index presented in Table 2 show negative values in case of all the heterobenzenes, thus predicting an aromatic character for all these compounds. As in the case of the angle-based geometric index, higher negative values of NICS index were obtained for the heteroanalogs of benzene $(\lambda^3\text{-P})_6$, $(\lambda^3\text{-P-}\lambda^3\text{-As})_3$, $(\lambda^3\text{-As})_6$. A similar situation is the one of the shielding increments $\Delta\sigma$, that also predict a strong aromatic character of the heterobenzenes $(\lambda^3\text{-P})_6$, $(\lambda^3\text{-P-}\lambda^3\text{-As})_3$, $(\lambda^3\text{-As})_6$ (Table 2). The very close values of the shielding increments computed in three different points along x and y axes also prove that the results may be a consequence of the geometry of the heterobenzenes series, especially of their symmetry.

COMPUTATIONAL SECTION

The heterobenzenes series have been first optimized at semi-empirical level (PM3 method) using MOPAC2009 program [9], then optimization at *ab initio* level (HF/6-31G) using Gaussian 03 package [10], has been performed. The geometries of all heteroanalogues of benzene are planar. NICS(2.5) indices have been computed at a 2.5 Å distance above the geometric center of each ring, using the GIAO method [g] at B3LYP/6-31G level. Computations of the shielding effects on the molecular probe of H_2 have also been performed at B3LYP/6-31G level, first at 2.5 Å above the geometric center of each ring and then at three different points along the x and y axis of the center of the molecule (0.5 Å, 1.0 Å and 1.5 Å in each direction). The shielding increments $(\Delta\sigma)$ were obtained by subtracting the shielding value corresponding to one of the H atoms of the H_2 probe alone (26.77 ppm) [5] from the shielding value of the closest H atom of the H_2 molecule relative to the heterobenzene plane.

REFERENCES

1. Balaban, A.T., Oniciu, D.C., Katritzky, A.R., *Chem. Rev.*, **2004**, *104*, 2777
2. Mracec, M., Pop, R.O., Medeleanu, M., Mracec, M., "Comparison of $(\lambda^3\text{-X-}\lambda^3\text{-Y})_3$ (X,Y = N, P, As, Sb, Bi) Aromatic Heterocycles Stability Through Quantum Chemical Methods", *ISNA13, 13th International Symposium on Novel Aromatic Compounds*, Luxembourg, July **2009**, 238

3. Pop, R.O., Medeleanu, M., Csunderlik, C., Mracec, M., *Rev. Chim.*, **2010**, *61*(5), 483
4. Schleyer, P.v.R.; Jiao, H., *Pure Appl.* Chem., **1996**, *68,* 209
5. Martin, N.H., Teague, M.P., Mills, K.H., *Symmetry*, **2010**, *2*, 418
6. Martin, N.H., Caldwell, B.W., Carlson, K.P., Teague, M.P., *Journal of Molecular Graphics and Modelling*, **2009**, *27*, 689
7. Doerksen, R.J., PhD Thesis „*Geometries, Polarizabilities and Aromaticity of Ring Molecules*", The University of New Brunswick, **1998**
8. El-Bakali Kassimi, N., Doerksen, R.J., Thakkar, A.J., *J. Phys. Chem.*, **1995**, *99* (34), 12790
9. *** MOPAC2009: James P. Stewart, Stewart Computational Chemistry, Version 9.069W; http://OpenMOPAC.net
10. Gaussian 03, Revision B.03, M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery, Jr., T. Vreven, K.N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Pittsburgh PA, **2003**

FUNCTIONALIZATION OF CARBON NANOTUBES

MELINDA E. FÜSTÖS*, ERIKA TASNÁDI*, GABRIEL KATONA*, MIRCEA V.DIUDEA*

ABSTRACT. The functionalization of single-walled carbon nanotubes (SWCNTs) is a timely topic in contemporary nanostructures literature. It is believed that modifications of SWCNTs properties could open the way towards real nanotechnology applications. In the present paper chemical functionalization of SWCNTs was performed to obtain first the carboxyl-functionalized species and then various synthetic approaches were investigated to obtain the target product (triethylene-glycol-functionalized SWCNTs), which can be used as a linker with medical purposes. The intermediate and final reaction products have been characterized by FT-IR spectroscopy, TEM analysis and micro-RAMAN spectroscopy.

Keywords: nanotechnology, carbon nanotubes, functionalization

INTRODUCTION

The concept of nanotechnology embraces applied science and technology. This field of study keeps developing day by day offering us information about the behavior of nanoparticles and their unique electrical, optical and magnetic properties [1]. Its practical use has many facets. This paper deals with the functionalization of SWCNT in order to increase their capacity of transporting therapeutic agents through cell membranes.

Carbon nanotubes can be classified in tree classes: single walled (SWCNT), double walled (DWCNT) and multiwalled (MWCNT) carbon nanotubes. They only consist of sp^2 hybridized carbon atoms (like the graphite), witch confer them a unique strength and toughness [2].

Depending on conditions, SWCNTs can form aggregates or they can exist as isolated tubes. Their ends can be opened or closed (the closing cap includes pentagons, also).

All SWCNTs can be represented by a pair of numbers, the so called chirality index (n, m) [3]. If n=m the tube is of armchair type, if one of the numbers is zero, then it is a zig-zag nanotube, and if n≠m≠0 the tube is chiral (Figure 1).

* *Faculty of Chemistry and Chemical Engineering, "Babes-Bolyai" University, Arany Janos Str. 11, 400084, Cluj, Romania*
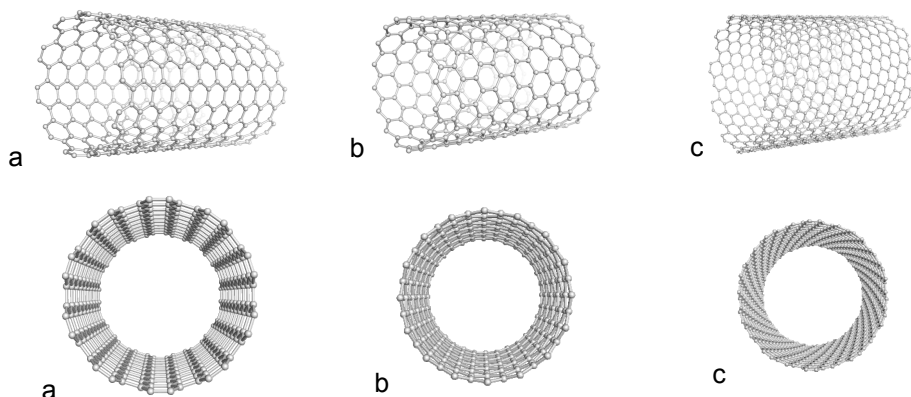
Figure 1. The tree types of SWCNTs (a – armchair; b – zig-zag; c - chiral)

Carbon nanotubes have incredible properties, such as hardness, thermal and electrical conductivity (all armchair nanotubes, 2/3 of the zig-zag type and 1/3 of the chiral ones are metallic, and the remaining ones are semi-conductors). Their tensile strength is 75 times higher than that of the steel, while their density is 6 times lower. Carbon nanotubes (CNT) are very light materials, bearing a density as low as 1.33-1.4 g/cm[3]. Another excellent physical property of CNT is their elasticity which helps them regain their original form after bending[4]. CNT are not miscible with any kind of solution; they only make suspensions. They can be synthesized in various ways, such as: chemical vapor deposition, arc discharge, laser ablation etc. [4-7].

RESULTS AND DISCUSSION

In the proposed synthesis the intermediates and the product were characterized by IR and microRaman spectroscopy, respectively and TEM microscopy as well. By using the IR spectroscopy allowed one to follow the reaction steps and verify the intermediates the reaction product, respectively.

The first intermediate, SWCNT-COOH, was characterization by IR analysis as shown in Figure 2. This spectrum proves the presence of the carbonyl group (-C=O) at 1655 cm^{-1}, the carboxyl group (–COOH) group can be observed at 1395 cm^{-1} and 3137 cm^{-1} and the carbon-oxygen bond (-C-O) shows a peak at 1066 cm^{-1}.

The IR spectrum of the second intermediate (SWCNT-COCl) (Figure 3) proves the presence of the group C=O of the chloride acid at 1705 cm^{-1} while the peak corresponding to the hydroxyl group from COOH (1395 cm^{-1} and 3137 cm^{-1}) disappeared.
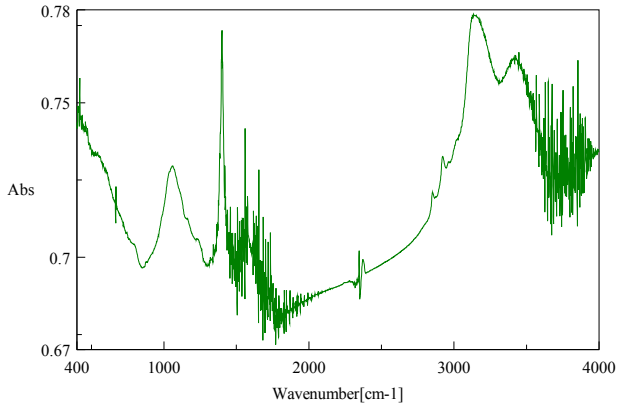
Figure 2. IR spectrum of SWCNT-COOH
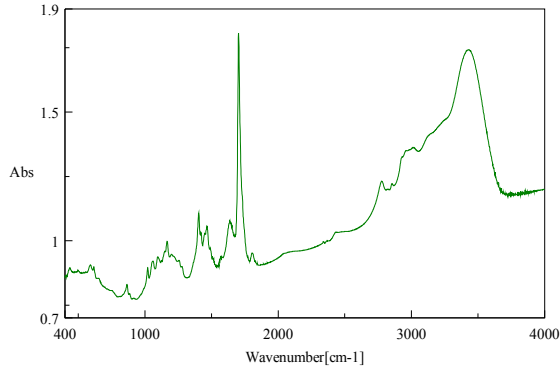
Figure 3. IR spectra of SWCNT-COCl

The IR spectrum of the final product is presented in Figure 4.
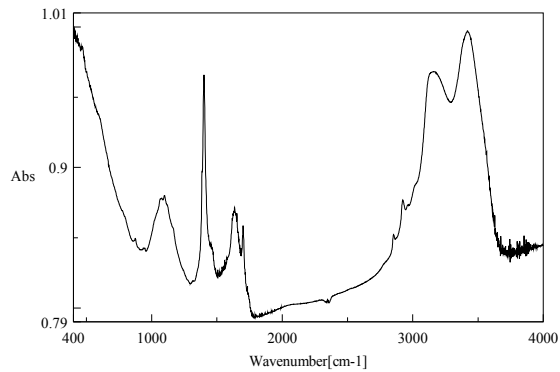
Figure 4. IR spectrum of SWNT-CO-O-$(CH_2)_2$-O-$(CH_2)_2$-O-$(CH_2)_2$-OH

155

The IR spectrum of the target product shows a peak at 1107 cm^{-1} which proves the presence of an ether bond (C-O-C), another peak at 1400 cm^{-1} of the –CH$_2$ group from the tryethylene-glycol. The peak at 1705 cm^{-1}, corresponding to the C=O bond from the chloride acid, (Figure 3) was missing but a peak at 3137 cm^{-1}, for the –OH group from the products chain end appeared.

The TEM microscopy images are shown in Figure 5: in comparison to the non-functionalized SWCNTs (5.a), the –COOH groups attached to the nanotubes are clearly seen in 6.b and 6.c (the final product).
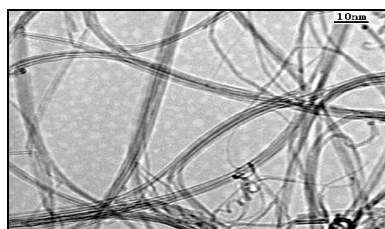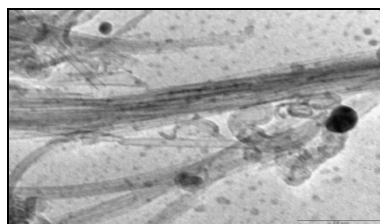
a – zoom of 120000x   b – zoom of 100000x
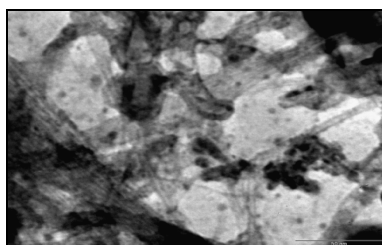
c – zoom of 100000x

Figure 5. a – SWCNT; b – SWCNT-COOH; c - SWNT-CO-[O-(CH$_2$)$_2$]$_3$-OH

The microRaman spectroscopy is a very sensitive tool, which allows one to observe fine structural modifications. From the spectrum represented in Figure 6 (a, b and c) we can see two important peaks: at 1585 cm^{-1} the so called G-band (a lower intensity band) and at 3186 cm^{-1} a higher intensity band. The G band for non-functionalized SWCNT was recorded at 1583 cm^{-1}, for SWCNT-COOH at 1582 cm^{-1} while in the case of final product at 1587 cm^{-1}.
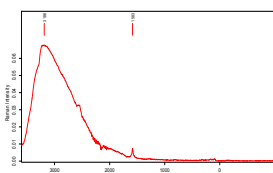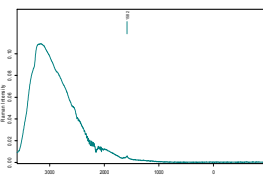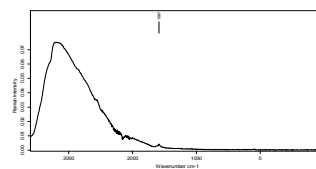
a.   b.   c.

Figure 6. a – SWCNT; b – SWCNT-COOH; c - SWNT-CO-[O-(CH$_2$)$_2$]$_3$-OH

Because in all three no any significant deviations were seen, especially at the G band, we can conclude that, during the functionalization processes, the carbon nanotubes did nof suffer any structural modifications.

CONCLUSIONS

Based on the IR, TEM and microRaman analysis, we proved the synthesis of the intermediates and the final product, that will be used in further biological studies, with the purpose of transporting therapeutic agents through cell membranes into the desired cells.

EXPERIMENTAL SECTION

The experimental part of this study was elaborated relying on the literature data available so far, the reaction parameters and reactive quantities being optimized according to available materials and resources. The experiment has multiple steps, represented in Figure 7.
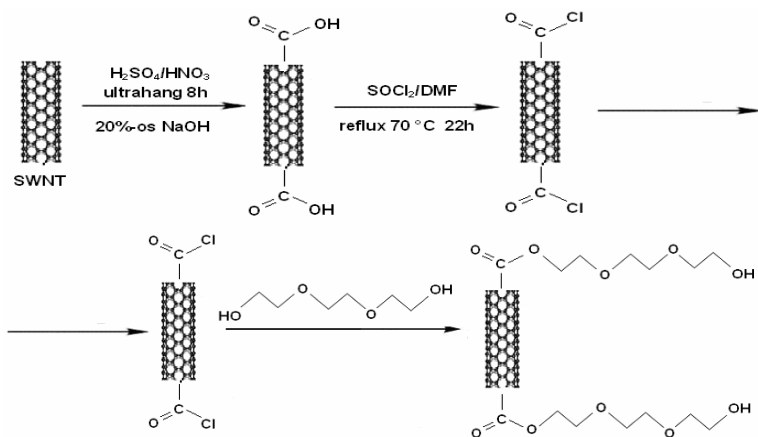
Figure 7. Experiment steps

The first step consists in the functionalization of SWCNTs by energic oxidation, with a mixture of H_2SO_4/HNO_3. The obtained carbonyl functionalized SWCNTs were then reacted with $SOCl_2$ resulting the acid chloride functionalized species. As a last reaction step, the above obtained functionalized SWCNTs were reacted with triethylene-glycol in order to obtain the desired product: triethylene-glycol-functionalized SWCNTs.

The importance of this study relies on the fact that these kind of functionalized SWCNTs have medical use, also.

The first step of the synthesis, as we mentioned above, was the synthesis of carboxyl functionalized SWCNTs and was performed according to C. Lynam et. al[6]. The procedure was the following: 20 mg of SWCNTs were suspended in a mixture of H_2SO_4/HNO_3 (ratio 3:1) in a well dried flask and left in an ultrasonic bath for 8 hours. After this reaction time the obtained mixture was diluted with bi-distilled water afterwards being neutralized with a 20% NaOH solution to pH 7. Later, the solution was filtered through PTFE membrane and the product dried.

The SWCNT-COOH intermediate, weighted 8 mg, was further suspended in a freshly distilled mixture of $SOCl_2/DMF$ (ratio 20:1) in a well dried flask[7]. This time the ultrasonic bath lasted for 20 minutes. Afterwards the mixture was refluxed in a stiring oil bath for 22 hours at 70°C. After cooling down, the mixture was distilled (to half of its volume), this way the remaining undesired $SOCl_2$ being removed. To make sure that all of the $SOCl_2$ was removed, the mixture was washed through with dioxane. The next step was rotavaporizing the mixture and than drying it for 48 hours under vacuum dryer.

The last step was the obtaining of the final product: triethylene-glycol-functionalized SWCNT s (SWCNT-CO-O-$(CH_2)_2$-O-$(CH_2)_2$-O-$(CH_2)_2$-OH)). First we attached to the SWCNT-COCl a six ringed 1,6-diol in order to position the hydroxyl group from the end of the chain further away from the carbon nanotubes. 4 mg SWCNT-COCl was suspended in a mixture of 40 µL triethylene-glycol and 1,5 mL 1,4-dioxane in a previously well dried flask. This was followed by sonication for 10 minutes and then refluxed for 52 hours at 110°C, under stirring on an oil bath. After cooling down, the mixture was filtered under vacuum and washed with 3x5 mL of THF (tetrahydrofuran) in order to remove the remaining triethylene-glycol, then dried for 24 hours in a vacuum dryer.

Instruments Used

The ultrasonic bath used for the dispersion of SWCNTs: TRANSSONIC 460/H, ELMA AUSTRIA, 100W, 40 kHz. To remove the remaining unwanted solvents we used the ROTAVAPOR P BÜCHI. The IR spectra of the intermediates was performed with a Fourier IR spectroscope (spectral range 7500-370 cm^{-1}, rezolution > 0.5 cm^{-1}, Michelson type interferometer, DLATGS detector). The analyzed samples were mixed with KBr and pressed into the form of a transparent tablete by hydraulic pressing.

For the electronmicroscopic illustration of the intermediates and of the final product, a transmission electron microscopy TEM was used (Hitachi Automatic TEM H7650, accelerating voltage 40-120 kV, zoom 200x-600000x).

For microRaman spectra, a FRA 106/S module was attached to the Fourier IR spectroscope, and to the module a microscope (Nikon ECLIPSE E400 – spectral range 3600-70 cm^{-1} for Stokes lines and 2000-100 cm^{-1} for the anti-Stokes lines, resolution > 1 cm^{-1}, Nd:YAG laser, ultrasensitive D418-T Ge detector) was linked with an optical wire.

Chemicals Used

The starting material was SWCNT synthesized by Chengu Organic Chemicals Co. Ltd, with the following characteristics: diameter 1-2 nm, lenght ~ 30 µm, purity > 90 tf%, special surface size (SSA) > 380 m^2/g, electrical conductivity > 10^2s/cm, production method CVD (chemical vapor deposition).

Reagents: 98% sulfuric acid (H_2SO_4, Mw = 98.08 g/mol, ρ_f = 1.84g/cm^{-3}), 70% nitric acid (HNO_3, Mw = 39.997g/mol, ρ_{sz} = 2.1g/cm^3), sodium hydroxyd (NaOH), tionil chloride ($SOCl_2$, Mw = 118.97g/mol, ρ_f = 1.638g/cm^3), N,N-dimethyl-formamide (DMF, Mw = 73.09g/mol, ρ_f = 0.944g/cm^3), tryethylene-glycol ($C_6H_{14}O_4$, Mw =150.17g/mol, ρ_f =1,1g/cm^3), tetrahyrofuran (THF, C_4H_8O, Mw = 72.11g/mol , ρ_f = 0.8892g/cm^3).

REFERENCES

1. H. Kuzmany*, A. Kukovecz, F. Simon, M. Holzweber, Ch. Kramberger, and T. Pichler, *Synthetic Metals*, **2004**, *141*, 113
2. S. Iijima, T. Ichihashi, and Y. Ando, *Nature,* **1992**, *356*, 776
3. E.B. Barrosa*, A. Joriob, G.G. Samsonidzef, R.B. Capazc*, A.G. Souza Filhoa, J. Mendes Filhoa, G. Dresselhause, and M.S. Dresselhausf, *Physics Reports*, **2006**, *431*, 261
4. J. Kürti, Szén nanocsövek, Fizikai Szemle, **2007**, *3*, 106
5. H. Dai, *Accounts of Chemical Research*, **2002**, *35*, 1035
6. C. Lynam and A.I. Minett, *Int. J. Nanotechnol*., **2008**, 5, 336
7. Y. Wang, Z. Iqbal, and S.V. Malhotra, *Chemical Physics Letters*, **2005**, *402*, 96

APPLICATION OF NUMERICAL METHODS IN THE TECHNOLOGY OF HYDROXYAPATITE

VALENTINA ROXANA DEJEU[a], SILVIA TOADER[b], BARABAS RÉKA[a], PAUL-ŞERBAN AGACHI[a]

ABSTRACT. Hydroxyapatite precipitation process involves the formation in the first phase of a solid-phase with amorphous structure (amorphous calcium phosphate), which in time turns into hydroxyapatite. Polynomial spline interpolation is a numerical method useful in mathematical modeling of this phase transformation process. By this method, experimental data are interpolated to obtain cubic spline polynomial function which can approximate reasonably well the experimental values of the degree of conversion at any pH between 8.5 ÷ 12. A very good agreement between the experimental and numerical results confirms the validity of the numerical procedure.

Keywords: *hydroxyapatite, phase transformation, numerical methods, cubic spline function*

INTRODUCTION

Numerical methods are used to determine approximate solutions of complex problems and use only simple arithmetic operations [1]. One of the simplest methods of approximation is interpolation and involves choosing a function data, which has a predetermined finite number of points (Δ): $x_0, x_1, x_2, ..., x_n$ chosen from its domain of definition. There is more than one class of interpolation functions, such as rational functions for rational interpolation, spline functions (polynomial or exponential) for spline interpolation, interpolation trigonometric functions for periodic functions [2]. The most suitable class of interpolation function is that where one can find an element closer to the function that interpolates [3]. This category includes cubic spline functions. Cubic spline function for function f and the above division (which are known values $f(x_i) = f_i$, $i = 1, ..., n$) satisfies the following three properties:

- It is a "segmental polynomial", meaning that each interval (x_{i-1}, x_i) is a polynomial $S_i(x)$ of degree 3

---

[a] *Universitatea Babeş-Bolyai, Facultatea de Chimie şi Inginerie Chimică, Str. Kogălniceanu, Nr. 1, RO-400084 Cluj-Napoca, Romania, vdejeu@chem.ubbcluj.ro*
[b] *Universitatea Tehnica Cluj-Napoca, Facultatea de Automatica si Calculatoare, Str. G. Baritiu nr. 26-28, RO-400027 Cluj-Napoca, Romania*

- two neighboring polynomials $S_i(x)$ and $S_{i+1}(x)$ have the following properties:

$$S_i(x_{i-1}) = f(x_{i-1}); \text{ for any } i = 1,...,n \tag{1}$$

$$S'_i(x_i) = S'_{i+1}(x_i), \text{ for any } i = 1,...,n-1 \tag{2}$$

General expresions for two adjacent cubic fuctions $S_i(x)$ and $S_{i+1}(x)$ are:

$$S_i(x) = \alpha \cdot a_i(x) + \beta \cdot b_i(x) + \gamma \cdot c_i(x) + \delta \cdot d_i(x) \tag{3}$$

$$S_{i+1}(x) = \beta \cdot a_{i+1}(x) + \gamma \cdot b_{i+1}(x) + \delta \cdot c_{i+1}(x) + \tau \cdot d_{i+1}(x) \tag{4}$$

where:

$$a_i(x) = \frac{(x_i - x)^2(x - x_{i-1})}{(x_i - x_{i-1})^2} \tag{5}$$

$$b_i(x) = -\frac{(x - x_{i-1})^2(x_i - x)}{(x_i - x_{i-1})^2} \tag{6}$$

$$c_i(x) = \frac{(x_i - x)^2[2(x - x_{i-1}) + (x_i - x_{i-1})]}{(x_i - x_{i-1})^3} \tag{7}$$

$$d_i(x) = \frac{(x - x_{i-1})^2[2(x_i - x) + (x_i - x_{i-1})]}{(x_i - x_{i-1})^3}. \tag{8}$$

Properties (1) and (2) become:

$$S_i(x_{i-1}) = \gamma = f(x_{i-1}); \quad S_i(x_i) = \delta = S_{i+1}(x_i) = f(x_i) \tag{9}$$

$$S'_i(x_i) = S'_{i+1}(x_i) = \beta \tag{10}$$

$$S'_i(x_{i-1}) = \alpha \tag{11}$$

To solve mathematical problems (scientific calculations), advanced software systems such as Matlab, Mathematica, or Mathcad are used [2,4,5].

It is generally acknowledged that in the crystallization of calcium phosphate first occurs the formation of a precursor phase (amorphous calcium phosphate) which is subsequently dissolved or restructured as the precipitation reaction occurs and turns into hydroxyapatite [6.7]. Transformation kinetics of amorphous calcium phosphate into hydroxyapatite, which can be described by a first order reaction law, is only a function of the solution pH at constant temperature [8.9]. Solution transformation depends on the conditions that regulate both amorphous calcium phosphate dissolution and formation of the first nuclei of hydroxyapatite [10]. In a recent study [11,12], the experimental results concerning the influence of pH and temperature on the transformation of amorphous calcium phosphate into hydroxyapatite have been presented.

From the kinetic data, the values of the rate constant and activation energy at pH 8.5, 9.1, 9.7, 10.2, 11.3, 12 were calculated. Based on the values obtained for activation energy, it has been established that the transformation of amorphous calcium phosphate into hydroxyapatite can be described by a combined macrokinetic mechanism: transfer-mass conversion. The mathematical model of the process was established and the constant values from the mathematical model equation were determined [11.12]. Simulations were made based on the proposed model and the results show a good agreement with the experimental data values, which confirms the validity of the model. Starting from these results, in the present work a method that can be used to determine quickly and easily the values for degree of conversion (η) in different experimental conditions is presented. A comparison between the numerical results and analytical results indicates that predictions obtained with the new technique are closer to the analytical solutions.

RESULTS AND DISCUSSION

The interpolation of experimental data in order to obtain spline interpolation function was performed with Mathcad 15. The experimental data presented in a previous communication [11.12] were used in the present study. Thus, for two temperatures $T_1^0 = 20\ ^0C$ and $T_2^0 = 50\ ^0C$, the spline $S(t)$ was determined, which is a function of one variable and approximates function η experimentally determined. Figures 1 and 2 show the spline function at various pH values:
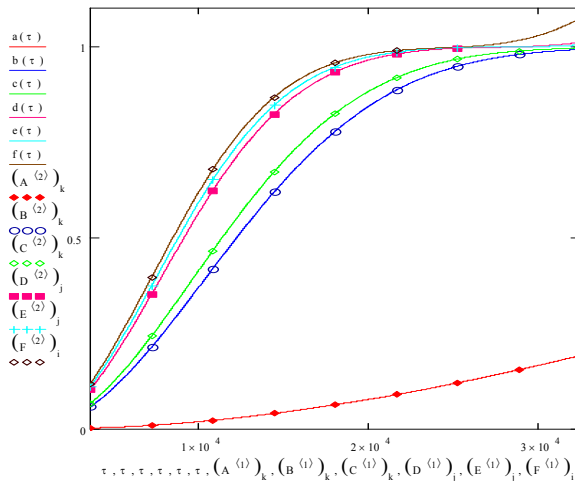
Figure 1. Graphic representation of the degree of transformation of amorphous phase in hydroxyapatite using spline function at 20 ^{0}C and 6 pH values.
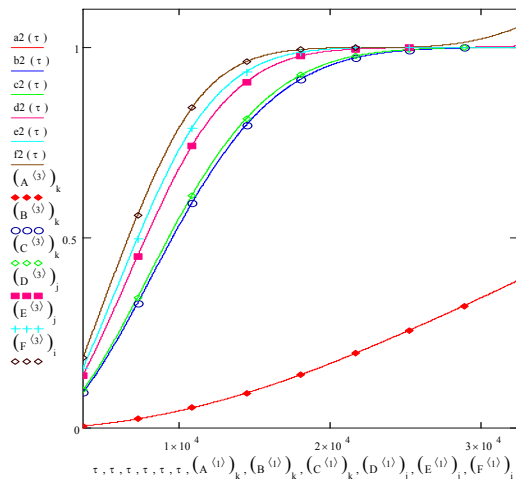
Figure 2. Graphic representation of the degree of transformation of amorphous phase in hydroxyapatite using spline function at 50 ^{0}C and 6 pH values

The analysis of the graphs obtained show that the interpolation spline nodes (t, T_1) have no large variations between nodes, so they model correctly the process of phase transformation.

Because function η varies with respect to time and pH, the interpolation of the function of two variables $\eta(t, pH)$ on domain containing (t, pH) experimentally determined at a fixed temperature T = 20 ^{0}C was made. The result is shown in Figure 3:
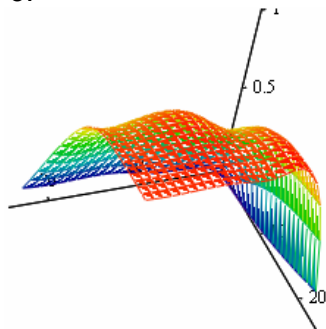
Figure 3. Function F (interpolation spline)

To determine the values of function η at fixed time and pH, the surface is divided with a plan that corresponds to ordinate pH point selected. The section curve of the plan with the spline surface $F(t, pH)$ is the same with the one obtained for one variable function η (in plan) at that temperature and pH.

164

So, the 3D cubic spline interpolation of function $\eta(t, pH)$ enables the determination of approximate values for function η for any other value from time and pH interpolation range at fixed temperature T.

CONCLUSIONS

In this paper we showed the possibility of using numerical methods in the technology of hydroxyapatite preparation. Compared with the mathematical model obtained in previously published works, this method is much simpler and allows the rapid determination of the degree of conversion values η, which in mathematical modeling of the process have a great importance. The advantage of mathematical formulation of the process is the reduction of the number of experiments, providing the ability to determine the values of function η at any time and in the pH range of interpolation.

Elaboration of experimental data obtained through spline functions is a new issue to the modeling domain when using the "black box" method. Polynomial spline interpolation method allows a rapid determination of conversion at any time and pH in the range of interpolation. It is necessary to continue the research in this field in order to determine the coefficients in polynomial equation of the spline functions. Therefore, using these functions in solving engineering problems, automation, management and optimization of processes would be more effective.

REFERENCES

1. Gh. Coman, "Analiza numerică", ed. Libris, Cluj-Napoca, **1995**.
2. S. Toader, Costin I., "Metode numerice", ed.Mediamira,Cluj-Napoca, **2009**.
3. I. Pavaloiu, "Rezolvarea ecuatiilor prin interpolare", ed. Dacia, Cluj-Napoca, **1981**.
4. M.G. Scheiber, D. Lixăndroiu, "Mathcad. Prezentare şi probleme rezolvate", Ed. Tehnica, Bucuresti, **1994**.
5. O. Cira, "Lectii de MathCad", ed Albastra, Cluj-Napoca, **2000**.
6. M. Johnsson, S.A. Nancollas, *Critical Reviews in Oral Biology & Medicine*, **1992**, *3*, 61.
7. S.V.J. Dorozhkin, *Material Science*, **2007**, *42*, 1061.
8. L. Bernard, M. Freche, J.L. Lacout, and B. Biscans, **1999**, *103(1)*, 19.
9. H.S. Liu, T.S. Chin, L.S. Lai, S.Y. Chiu, K.H. Chung, C.S. Chang, M.T. Lui, *Ceramics International,* **1997**, *23(1)*, 19.
10. A.L. Boskey, A.S. Posner, *The Journal of Physical Chemistry,* **1973**, *77*, 2313.
11. V.R. Dejeu, R. Barabás, Al. Pop, E.S. Bogya, P.-Ş. Agachi, *Studia Chemia*, **2009**, *3*, 61.
12. .V.R. Dejeu, R. Barabás, Al. Pop, E.S. Bogya, P.-Ş. Agachi, *Revista de chimie,* **2009**, *12*, 1251.

IMPLICATIONS OF SENSE/ANTISENSE
NUCLEIC-ACID CODONS ON AMINO-ACID COUNTS

VLADIMIR R. ROSENFELD*, DOUGLAS J. KLEIN*

ABSTRACT. We study the amino-acid content of protein sequence factors translated from *codonic palindromes* of nucleotide sequences, which have each half comprised from an integer number of codons. Alternatively, our study may be viewed to seek consequences if sense & antisense translations for proteins originate with the two (complementary) strands of RNA.

Under either of these presuppositions, we conclude: the total number of aspartic-acid, asparagine, tyrosine, and histidine residues produced equals the total number of isoleucine, methionine, and valine residues produced. Further, we find a suite of inequalities on amino acid counts. Our results provide a rigorous consequence to a relation considered by Zull *et al*. Further, a "parity rule" of Chargaff *et al*. gives some support for a sense/antisense presumption.

Keywords: nucleotide sequence, codonic palindrome, translation, parity rule, complementation

INTRODUCTION

Nucleotide sequences of DNA (desoxyribonucleic acid) and RNA (ribonucleic acid) are constructed from four types of nucleotides denoted by characters **A**, **C**, **G**, and either **T** (for DNA) or **U** (for RNA). DNA consists of two complementary strands, with these four characters matched into two complementary pairs: **A** & **T** and **C** & **G**.

Here, we investigate the consequences of protein translation from both sense & antisense directions along nucleotide sequences. This might [1, 2], sometimes, arise from oppositely oriented translation along strands from complementary DNA strands. Or it can arise from a single RNA strand which is a "codonic" palindrome. It is natural to interrelate amino acids as to whether they have inverted nucleotide codons, and, indeed, such has already been done by Zull & coworkers [3, 4]. From this interrelation (conveniently expressible as a "graph", of vertices representing amino acids, and edges representing the relation), consequences then are sought. Zull & Smith [3] questioned whether 3 portions of this graph correspond to 3 classes of amino acids manifesting

---

* *Mathematical Chemistry Group, Department of Marine Sciences, Texas A&M University at Galveston, 200 Seawolf Parkway, Galveston, TX 77553-1675, USA, rosenfev@tamug.edu, vladimir_rosenfeld@yahoo.com, kleind@tamug.edu*

different secondary protein structure (α-helix, β-sheet, or random). This could only be a statistical correlation, as many amino acids occur in 2 or 3 types of secondary protein structures (though with different frequencies), and, indeed, Zull & Smith found only a (very) weak correlation. We developed a different type of consequence which however is rigorous, under either the condition of sense/antisense translation of complementary RNA strands or translation from a "codonic" palindrome. We found equal net weights for the frequencies of occurrence of amino acids in two subclasses comprising one of the (bipartite) fragments of the codon inversion graph. Either Chargaff's proposal [5] of forward (sense) & reverse (antisense) translations nucleic acid sequences or Zull's idea of codonic palindromes leads to a general sense/antisense reading of individual codons.

Given a nucleotide sequence, a later disjoint sequence is termed an *inverted repeat* if it consists of the complements of the first sequence in reverse order. The initial sequence and the later inverse repeat are together termed ([6], p. 76) a *complementary palindrome* – elsewhere often termed simply a "palindrome". Sometimes, the direct sequence and its inverted repeat both consist of an integer number of codons. Such a pair of palindromic sequences (or subsequences) consisting exactly of an integer number of codons is called a *codonic palindrome*. We represent the situation when there are s codons in each of RNA sequence by

$$a_1 a_2 \cdot \cdot \cdot a_{3s-1} a_{3s} \cdot a_{3s}{}^* a_{3s-1}{}^* \cdot \cdot \cdot a_2{}^* a_1{}^*, \tag{1}$$

where a_i & $a_i{}^*$ are two complementary nucleotides (say, **C** & **G**) of the nucleotide alphabet $\mathcal{A}$ = {**A**, **C**, **G**, **U**}. Note: the "codonic" condition on this (complementary) palindrome means the direct & inverted sequences each comprise an integer number of codons.

A more general notion allows "concatenation" of different codonic fragments of a codonic palindrome. The codons of a codoinic palindrome may be moved around to be placed in different positions, still preserving all codons, just in a different order. We term such a reassemblage a *codonic palindromic conglomerate*.

This allows consecutive codonic-palindromic loops (such as occur with introns), and this also accounts for nested loops (*i. e.*, loops of smaller size inserted into contour sequences of loops of larger size), it allows even multiply nested loops. We may imagine: at the first hypothetic stage, starting from a single giant codonic palindrome, with direct sequence t and inverted repeat u, each of which are to be broken up into codon subsequences, say as $(t_1, t_2, \ldots, t_m)$ and $(u_s, u_{s-1}, \ldots, u_1)$, with possibly different numbers of different-lengthed subsequences t_i & u_j from t & u; and at the second step, putting these different subsequences back together in an arbitrary order. The superpalindrome need not be biologically realized but rather just the intermixed codonic palindromic conglomerates.

Granted these ideas, we develop some formalism in the next section, so as to identify notable consequences on the numbers of amino acids formed within different selected groups, under the assumption that the RNA is a codonic palindromic conglomerate. Most of the formal discussion is not needed to understand the final biological consequences, which come in Propositions 4, 5, & 6.

FORMAL RESULTS

To manipulate nucleotide sequences, one may use three commuting operators: α standing for complementation (as indicated by (*) in (1)) of nucleotides in a string; β for inversion of the string, and the composition $\gamma = \alpha\beta = \beta\alpha$. We can formally rewrite (1) using γ:

$$a_1a_2\cdot\cdot\cdot a_{3s-1}a_{3s}\cdot \gamma(a_1a_2\cdot\cdot\cdot a_{3s-1}a_{3s}). \qquad (2)$$

Let $\mathcal{B} = \{b_1, b_2, \ldots, b_{21}\}$ be the set of 21 amino acids (where the 21st amino acid terminologically corresponds to the triple of stop codons). For a nonempty subset $S \subseteq \mathcal{B}$ of amino acids, denote by $C(S)$ the set of all codons for the amino acids from S. And let $\gamma C(S)$ denote the result of action of the operator γ on each codon in $C(S)$.

We investigate the consequences of a pair of subsets S_1 and S_2 of amino acids, for which $\gamma C(S_1) = C(S_2)$, or equivalently $\gamma C(S_2) = C(S_1)$, since γ is idempotent (i. e., $\gamma^2 = 1$), as also are α and β.

Lemma 1. *Let T_1 & T_2 be two sets of amino acids such that $C(T1) = \gamma C(T2)$. Let $a = a_1a_2\cdot\cdot\cdot a_{3s-1}a_{3s}\cdot a_{3s}*a_{3s-1}*\cdot\cdot\cdot a_2*a_1*$ ($a_i, a_i* \in \mathcal{A}$; $1 \leq i \leq 3s \geq 3$) be a codonic palindrome. Moreover, let l_j (res. r_j) ($j = 1, 2$) be the total number of occurrences of codons belonging to $C(T_j)$ in*

$a_1a_2\cdot\cdot\cdot a_{3s-1}a_{3s}$ (res. $a_{3s}*a_{3s-1}*\cdot\cdot\cdot a_2*a_1*$). Then $l_1 = r_2$ & $r_1 = l_2$.

Proof. Since $a = t\gamma(t)$, with $t = a_1a_2\cdot\cdot\cdot a_{3s-1}a_{3s}$ & $\gamma(t) = a_{3s}*a_{3s-1}*\cdot\cdot\cdot a_2*a_1*$, the numbers of "direct" and inverted codons in a are equal. Also, by construction, $C(T_1)$ and $C(T_2)$ are sets of mutually inverted codons, whence we immediately arrive at the proof.

Note: for any codon t representing a respective amino acid b_i, the corresponding codon $u = \gamma(t)$ always represents a distinct amino acid b_j. Therefore, γ induces a binary relation on the set $\mathcal{B}$ of all amino acids which can be represented thereon by a simple graph Γ, where amino acids b_i & b_j are adjacent (linked by an edge) if there exist a codon t of the former and a codon u of the latter which are interchanged by γ ($u = \gamma(t)$ & $t = \gamma(u)$). Important here are the connected components (maximal connected subgraphs of Γ). We immediately use these considerations in the following

Lemma 2. *Let T_1 & T_2 be two sets of amino acids with $C(T_1) = \gamma C(T_2)$. Then, for any $U_1 \subseteq T_1$ corresponding to a connected component of Γ, either U_1 is entirely in T_2 or else entirely external to T_2 (i. e., either $U_1 \cap T_2 = U_1$ or $U_1 \cap T_2 = \varnothing$).*

Proof. Associate to the union $T_u = T_1 \cup T_2$ a graph H whose vertex set is T_u, and two vertices i & j are adjacent in H if there exist codons t_i & t_j such that $t_i = \gamma(t_j)$. Now, attach exactly one self-loop to every vertex of H to obtain a derivative graph $\hat{H}$ having the same connectivity components. Clearly, $\hat{H}$ is an equivalence relation on T_u where any pair of vertices i and j are equivalent iff these belong to one connected component. Indeed, three conditions are satisfied: (i) reflexivity, as guaranteed by 'self-connectivity' of every vertex having an attached self-loop; (ii) symmetry, since $t_i = \gamma(t_j) \Leftrightarrow t_j = \gamma(t_i)$; and (iii) transitivity, as follows from the connectivity within a component. Evidently, in our hypothesis, U_1 is a single equivalence class of vertices of T_u, while T_2 is the union of a number of equivalence classes of vertices thereof. Since two equivalence classes of objects either coincide or share no element, U_1 is either included as one such class in T_2 or intersects with no equivalence class of vertices comprising T_2. This completes the proof.

Corollary 2.1. *Let T_1 & T_2 be two sets with $T_1 \neq T_2$ and $C(T_1) = \gamma C(T_2)$. Then, if T_1 & T_2 are minimal, they are disjoint.*

Proof. This uses reasoning similar to Lemma 2. Namely, minimal sets T_1 and T_2 are both equivalence classes of $T_u = T_1 \cup T_2$. Since $T_1 \neq T_2$, we immediately arrive at the proof.

Corollary 2.2. *Let $a = a_1 a_2 \cdots a_{3s-1} a_{3s} \cdot a_{3s}{}^* a_{3s-1}{}^* \cdots a_2{}^* a_1{}^*$ ($a_i, a_i{}^* \in \mathcal{A}$; $1 \leq i \leq 3s \geq 3$) be a codonic palindrome. Moreover, let n_j ($j = 1, 2$) be the total number of occurrences of codons belonging to $C(T_j)$ in $a = a_1 a_2 \cdots a_{3s-1} a_{3s}$ (res. $a^* = a_{3s}{}^* a_{3s-1}{}^* \cdots a_2{}^* a_1{}^*$). Then $n_1 = n_2$.*

Proof. Note that $n_j = l_j + r_j$ ($j = 1, 2$). By virtue of the equalities $l_1 = r_2$ and $r_1 = l_2$ demonstrated in Lemma 1, the proof is immediate.

Proposition 3. *In a codonic palindromic conglomerate, there are equal amounts n_1 and n_2 of amino acids from respective minimal subsets T_1 and T_2, as in Corollary 2.1.*

Proof. The initial codonic superpalindrome has $n_1 = n_2$, by Corollary 2.2. But breaking up into codons and rearranging all the various codons does not change the numbers of the different codons, so that one still has $n_1 = n_2$.
Next, we frame these results more biologically.

AMINO-ACID COUNTS

The relation γ which acts on a nucleic acid string to reverse & complement it leads to a relation between amino acids: if an amino acid has codon $u = a_1 a_2 a_3$, then, it is related or linked to the amino acid with $\gamma(u) = u_3{}^* u_2{}^* u_1{}^*$.

This overall γ-relation is conveniently represented as a graph Γ where an edge occurs between the amino acids of codons u & $\gamma(u)$. Using the standard codons (*e. g.*, as in Ch. 13 of [7]), the graph Γ appears in Fig. 1 – also given by Zull *et al.* [4]. But now (following the results of our preceeding section) we seek a minimal pair of subsets S_1 & S_2 of amino acids for which $\gamma C(S_1) = C(S_2)$, and $\gamma C(S_2) = C(S_1)$, since $\gamma^2 = 1$. It turns out that in Γ there is a pair of such sets: $S_1 = \{$D, N, T, H$\}$ & $S_2 = \{$I, M, V$\}$, where D, N, T, H, I, M, V denote aspartic acid, asparagine, tyrosine, histidine, isoleucine, methionine, and valine, consecutively. One sees that our sets S_1 & S_2 are mutually interconnected while being completely disconnected from the remaining vertices. The corresponding codon sets are $C(S_1) = \{$**GAU**, **GAC**; **AAU**, **AAC**; **UAU**, **UAC**; **CAU**, **CAC**$\}$ and $C(S_2) = \{$**AUU**, **AUC**, **AUA**; **AUG**; **GUU**, **GUC**, **GUA**, **GUG**$\}$. Application of the operator γ to $C(S_1)$ gives $\gamma C(S_1) = \{$**AUC**, **GUC**, **AUU**, **GUU**, **AUA**, **GUA**, **AUG**, **GUG**$\}$, which is just $C(S_2)$. Hence, also $\gamma C(S_2) = C(S_1)$. This is the only pair of minimal distinct sets S_1 & S_2 of amino acids having the described property in Γ (to transform quantitatively into each other under the operator γ). The remnant set $S_3 = B \setminus S_1 \cup S_2$ of amino acids gives a minimal set $C(S_3)$ of codons closed under the action of γ.
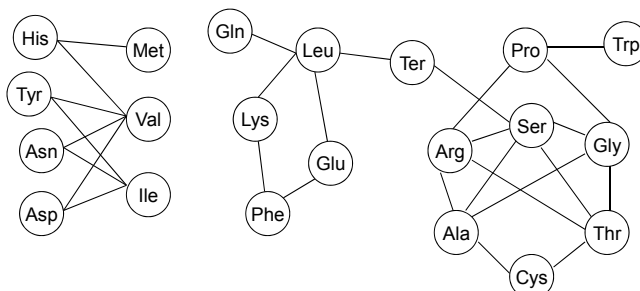
Figure 1: The graph Γ of γ-relations of amino acids; the left bipartite component displays the sets S_1 (the 4-site part: Hys, Tyr, Asn, Asp) and S_2 (the 3-site part: Met, Val, Ile), while the right component displays the set S_3 .

A codonic palindromic conglomerate merely conditions codons to occur in complementary pairs. Thence, allowing several codonic palindromes nested, or multiply nested, or interlinked in all kinds of ways. Instances of such objects can occur in introns. Recall that the mRNA of eukaryotes is obtained through *splicing* from pre-mRNA (precursor mRNA), which is similar to a portion of a strand of DNA. During splicing, relatively long factors called *introns* are removed from a pre-mRNA sequence. Most introns are 80 to 400 base pairs in size; though there also exist huge introns of length >10,000. While introns do not themselves participate in producing amino acids, it is of note that the intronic loops even of a very high degree are covered in the conditions of Proposition 3, where $n_1 = n_2$ is achieved. More explicitly for (T_1 & T_2 of Proposition 3 realized as) S_1 & S_2 in Fig. 1, with $\#_X$ being the number of amino acid moieties X produced, we arrive at a primary biological result:

Proposition 4. *In protein factors translated from codonic palindromic conglomerates, such as occur with various stem loops, numbers of amino-acid residues are related*

$$\#Asp + \#Asn + \#Tyr + \#His = \#Ile + \#Met + \#Val.$$

But, granted our codonic palindromic conglomerates, there are further (weaker) consequences, concerning inequalities on amino acid numbers. In particular, we have:

Proposition 5. *In protein factors translated from codonic palindromic conglomerates, there are inequalities on the numbers of different amino acids:*

$$\#Met \le \#His;$$
$$\#His \le \#Met + \#Val;$$
$$\#Ile \le \#Tyr + \#Asn + \#Asp;$$
$$\#Tyr + \#Asn + \#Asp \le \#Ile + \#Val;$$
$$\#Gln \le \#Leu;$$
$$\#Trp \le \#Pro;$$
$$\#Ter \le \#Leu + \#Ser;$$
$$\#Pro \le \#Trp + \#Arg + \#Gly;$$
$$\#Ser \le \#Ter + \#Arg + \#Gly + \#Ala + \#Thr$$
$$\#Cys \le \#Ala + \#Thr;$$
$$\#Ala + \#Thr \le \#S\ er + \#Arg + \#Gly + \#Cys;$$
$$\#Leu + \#Phe \le \#Gln + \#Lys + \#Glu + \#Ter;$$
$$\#Gln + \#Lys + \#Glu \le \#Phe + \#Leu;$$
$$\#Gln + \#Lys + \#Ter + \#Glu \le \#Leu + \#Phe + \#Ser;$$
$$\#Trp + \#Arg + \#Gly + \#Cys \le \#Pro + \#S\ er + \#Ala + \#Thr;$$
$$\#Pro + \#Ser + \#Cys \le \#Trp + \#Ter + \#Arg + \#Gln + \#Ala + \#Thr;$$

where the number #Ter of "stops" is conveniently identified to the number of different proteins.

Proof. Our proof begins with a transformation of G of Fig. 1 into a symmetric digraph Γ where each edge of G is converted into a pair of opposite directed arcs between the same two vertices (as originally connected by the replaced edge). We attach to every arc $i\,j$ of Γ a weight a_{ij} equal to the total multiplicity of all codons representing amino acid i which are transformed by the operator γ into codons of amino acid j. Next, we use a (common mathematical) definition that a subset I of vertices of G is *independent* if no two vertices of I are adjacent in G. Any independent subset I of amino acids (nontransformable one into another by γ) determines the set $J = N(I)$ of all amino acids adjacent to members of I. Evidently, the operator γ transforms all codons of amino acids from I into codons representing amino acids from J, but the converse is true if no two amino acids of J are adjacent in Γ (or Fig. 1). In general, there holds a (nonstrict) inequality interrelating the total numbers of codons of I transformed into codons of J and of codons of J itself, taking into account other possible transformations of codons of $N(I)$ – not into codons of I. Thus, we deduce

for the total numbers of codons in *I* and *J* that $|C(I)| \leq |C(J)|$. Hence, particular proofs for all cases considered in Proposition 3 follow, with different choices of independent *I* & neighbors $J = N(I)$ corresponding to the left & right sets of amino acids in each of these inequalities.

Presumably, these statements are most important when at least most of the RNA (or DNA) is comprised from codonic palindromic conglomerations. But, perhaps, most significantly Propositions 4 and 5 hold under the sense/antisense circumstance proposal in [1] & [2], [5] and explored in [3] & [4]. That is:

Proposition 6. *If in place of the condition of codonic palindromic conglomerates in Propositions 4 & 5, the protein factors are translated from RNA, obtained from both (sense & antisense) DNA strands, then the conclusions 4 & 5 still hold.*

Proof. The two corresponding RNA strands may be viewed as a single codonic palindrome, say each of the strands being separated from one another by a hypothetical "stop" codon. Thence, Propositions 4 & 5 apply.

DISCUSSION

Comparison of proportions of amino acids as indicated by Propositions 4 & 5 are perhaps of practical interest. Clearly, 4 & 5 are most relevant when all or at least a major part of the RNA (or DNA) is comprised from codonic palindromes – in as much as the various indicated amino acids may be coded for in different amounts by the portion of the nucleotide chain outside the codonic palindromes. Chargaff & coworkers' "parity rule" [5, 8–10] is in general a little weaker than the hypothesis of 4 & 5, but still is supportive of it, for some selected species. Most significantly, our results on amino-acid counts apply fully if the sense/antisense hypothesis of [1, 2] is met. As such, our Proposition 6 offers a strong test of the occurrence of sense/antisense translations – such as we imagine though not a general occurrence, could be the situation for selected species.

Further, note that a "parity rule" of Chargaff and coworkers [5] suggests that, in a wide class of single strands of DNA, the numbers of **A**&**T** nucleotides match as also do the numbers of **C**&**G** nucleotides. (This seems to occur [8] especially for eubacterial and chloroplast DNA.) That is, granted the satisfaction of this Chargaff's rule, single DNA strands have met (in our formal nomenclature) a first condition for the whole DNA molecule to be a codonic palindromic conglomerate. A strengthening of this rule to say that complementary nucleotides fully "condense" into complementary
codons would then imply our result for a single strand of DNA.

Again, our ideas are related to Zull and coworkers [3, 4], though they look at the possibility of the graphic structure of Fig. 1 to be statistically manifested in secondary protein structures, whereas what we focus on is what might be termed "0-ary" structure (of amino acid counts). A further point is that our results (of Propositions 4, 5, 6) are robust to certain rare complications involving the rare alternative translation of a "stop" to some

other rare amino acid – and this may be seen not to hurt any of the inequalities in 5. For instance, the "stop" codon **UGA** can in certain mitochondria code for tryptophan and for selenocystein in certain Archaea. Also, this occurs because [12] the second stop codon **UAG** can code for pyrrolysine in Archaea and bacteria.

Besides, the (standard) mode of forming RNA loops, another hypothetic possibility might be imagined to form "reverse loops" (*i. e.*, helixlike loops) interconnecting between a directed sequence and a second sequence of nucleotides which, though complemented from the first sequence, is not reversed in direction along the strand. If such is imagined: namely, to occur (as has indeed been entertained as a possibility by Chargaff *et al.* [10]), one could then inquire about the numbers of different amino acids which arise from two so-related sequences. That is, one would inquire about the interrelated amino acids, considering our complementation operator α as interrelating the two nucleotide sequences – conglomerated or not. Then, the same sort of results found in our formal section apply, with γ replaced by α, now with reference to the α-graph of Fig. 2.
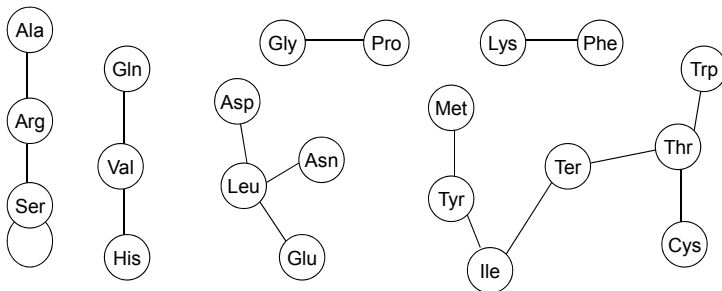
Figure 2: The graph of α-relations of amino acids.

With many bipartite components in this graph, this would evidently lead to a multiplicity of interrelations amongst numbers of various amino acids. For instance, this would imply that the amounts of glycine & and proline are the same (and also the amounts of lysine & phenylalanine) – seemingly, these equalities (and more) do not occur, so that the pairing between a direct sequence and a second sequence in a strand in the same direction, evidently, does not occur. The apparent reason must be that, *e. g.*, the pairing between **C** & **G** occurs only when the two nucleotides in making contact are oppositely oriented along a nucleotide chain, whence we might in fact distinguish the possibilities by $\overrightarrow{\mathbf{C}}\,\&\,\overrightarrow{\mathbf{G}}$ for nucleotides oriented in one "symparallel" direction along the chain, and $\overleftarrow{\mathbf{C}}\,\&\,\overleftarrow{\mathbf{G}}$ in the other direction along the chain – so that pairing occurs between antiparallel [10] $\overrightarrow{\mathbf{C}}\,\&\,\overleftarrow{\mathbf{G}}$ (or

174

between $\overrightarrow{\mathbf{G}} \& \overleftarrow{\mathbf{C}}$), but not between "symparallel" $\overrightarrow{\mathbf{C}} \& \overrightarrow{\mathbf{G}}$ (or $\overleftarrow{\mathbf{C}} \& \overleftarrow{\mathbf{G}}$). A similar comment applies for **A** & **U** (or **T**). That is, the conformational structure of each nucleotide is evidently different along the two different directions of a chain. Overall this evidently accounts for the fact that nucleotide sequences always form copies in antiparallel directions, rather than symparallel directions (with complementation). This, seemingly, is an evolutionarily selected (or) condition for faithful transcription.

CONCLUSION

Beyond the presumption of sense/antisense reading of codons, our exposition here arises from just very basic facts of molecular genetics. Under such (sense/antisense) conditions, we have found novel biological consequences enounced in Propositions 4, 5, and 6. Being rigorous consequences of these conditions, the amino-acid count relations may be used as tests for either Chargaff's sense/antisense hypothesis (in RNA) or for our codonic palindromic conglomerate condition (whence, then, Zull's hypothesis). That is, if our amino-acid conditions are not met, then this denies both Chargaff's and Zull's hypotheses. Finally, we may mention two other recent works [13, 14] which consider similar biological matters in a wider algebraic context.

ACKNOWLEDGMENTS

REFERENCES

1. L.B. Meckler, *Biofizika*, **1969**, *14*, 581.
2. R.G. Idlis, *Zhurnal Vses. Khim. Ob-va im. Mendeleeva*, **1980**, *25*, 431.
3. J.E. Zull and S.R. Smith, *Trends in Biochemical Sciences*, **1990**, *15*, 257.
4. J.E. Zull, R.C. Taylor, G.S. Michaels, and N.B. Rushforth, *Nucleic Acids Research*, **1994**, *22* (*16*), 3373.
5. R. Rudner, J.D. Karkas, and E. Chargaff, *Proc. Natl. Acad. Sci. USA*, **1968**, *60*, 921.
6. M.G. Sadovsky, J.A. Putintseva, and A.S. Shchepanovsky, *Theory Biosci.*, **2008**, *127*, 69.
7. J.D.D. Watson, "*Molecular Biology of the Gene*", 3rd ed., W. A. Benjamin Inc., New York, **1976**.
8. C. Nikolaou and Y. Almirantis, *Gene*, **2006**, *381*, 34.

9. D.R. Forsdyke and J.R. Mortimer, *Gene*, **2000**, *261*, 127.

10. E. Chargaff, *Nature,* **1965**, *206*, 145.

11. S. Castellano, A.M. Andrés, E. Bosch, M. Bayes, R. Guido, and A.G. Clark, *Mol. Biol. Evol.*, **2009**, *26* (*9*), 2031.

12. G. Srinivasan, C.M. James, and J. A. Krzycki, *Science*, **2002**, *296* (*5572*), 1409.

13. V.R. Rosenfeld, *MATCH Commun. Math. Comput. Chem.*, **2006**, *56 (2)*, 281.

14. V.R. Rosenfeld., *MATCH Commun. Math. Comput. Chem.*, **2007**, *57 (1)*, 134.

CYCLIC NUCLEOTIDE SEQUENCES
CODONICALLY INVARIANT UNDER FRAME SHIFTING

VLADIMIR R. ROSENFELD*, DOUGLAS J. KLEIN*

ABSTRACT. A shift of the frame in a polynucleotide sequence typically alters the codon content of the sequence. This provokes a question as to what sequence might be unaltered after shifting the frame. In fact, a linear sequence cannot exactly be so conserved – but there might be a possibility if it is a cyclic code subjected to a circular permutation, as we consider here. The solution is strikingly simple: A cyclic sequence of different nucleotides conserves a circular order of its codons under any shift of its frame if it has a length λ not divisible by 3 and is consecutively read κ times, or it is composed of κ repeated copies of a factor h of length λ, where κ is divisible by 3, while λ is not. For example, the sequence atcgatcgatcg has a factor atcg of length λ = 4 is repeated *κ* = 3 times. Translating this code without any shift gives isoleucine, aspartic acid, arginine, and serine, consecutively, or IDRS for short. The circular shift by 1 position results in SIDR, by 2 positions if produces RSID, and (here) at last, the circular shift by 3 positions gives DRSI. Apparently, all four translated codes of amino acids are the same relative to cyclic permutation. We conclude here discussing the *cyclically invariant* codes by noting that these can easily be enumerated using the famous Pólya's theorem.

Keywords: nucleotide sequence, codonic, frame shift, cyclically invariant, permutable

INTRODUCTION

Nucleotide sequences of DNA (desoxyribonucleic acid) and RNA (ribonucleic acid) are constructed from four types of nucleotides denoted by the characters A, C, G, and either T or U, where options T or U are used in cases of DNA or RNA, respectively. According to the complementarity of two strands in DNA, these four characters comprise two complementary pairs: A & T (or U) and C & G. See, *e. g.*, Ch. 13 in [1].

A cyclic shift of the frame in an RNA polynucleotide sequence, in general, alters the resulted sequenced codon content, by which we mean the net number of codons of each different type. This provokes questions as to whether there are codes unaltered after shifting the frame, and, if so, then what codes. Typically, a linear sequence of codons cannot exactly be

---
* *Mathematical Chemistry Group, Department of Marine Sciences, Texas A&M University at Galveston, 200 Seawolf Parkway, Galveston, TX 77553-1675, USA, rosenfev@tamug.edu, vladimir_rosenfeld@yahoo.com, kleind@tamug.edu*

so conserved – except possibly in certain circumstances. As announced in the title, cyclic sequences of nucleotides having this property do exist. Cutting such a cyclic sequence at an arbitrary position produces a linear factor f which then might be read as a sequence of codons. But with an alternative cut, a shift of the frame by one or two nucleotide positions can under suitable circumstances give the same codon content (*i. e.*, the same counts of acid type of codons -- and thence of each type of translated amino acid). Still one might imagine another scenario where a cyclic RNA is read without cutting, with the reading going round repeatedly – and this under different circumstances can again lead to codon conservation. That is, our considerations are connected with potential ways in which nature might create a kind of 'selfcorrecting code' for amino-acid content, or even codon sequences (up to cyclic permutation), such as to conserve the construction of proteins which are synthesized through translation of codons to amino acids. That is, regardless of the starting point, or reference frame choice for codon translation, the same result would be realized for a codonically invariant cyclic sequence. But also, instead of a cyclic RNA, one may also imagine a linear one having a similarly constructed, periodic factor whose frame shift produces the same circular shift of codons therein and thereby assures the same (apparently circular) order of a translated amino acid sequence.

RESULTS AND DISCUSSION

We begin with a selfevident statement:

Lemma 0. *A periodic cyclic sequence, of the length $\geq 3,$ obtained by repetition of just one nucleotide conserves a fixed codonic content which does not alter under any shift of frame.*

Note that both of DNA and RNA normally contain (long) stretches of mononucleotide repeats; besides the conservation of codon content, they may play an important role in base composition and genetic stability of a gene and gene functions, *etc.*. However, it is not yet properly understood -- how nature keeps a fixed-frame reading of general-type codons to reproduce many times the same polypeptide molecules, in organisms. Here, we apply some combinatorial reasoning to comprehend certain details of this complex natural phenomenon.

The first result of this paper is the following statement:

Lemma 1. *Let f be a cyclic sequence of nucleotides with a length $|f|$ not divisible by 3 and with not all nucleotides being the same. Then, there is conservation of a circular order of codons under any shift of frame if f is consecutively read κ times, where κ is divisible by 3.*

Proof: First, we address the case where the length of nucleotide sequence is not a multiple of 3, say $3k \pm 1$, with k being a positive integer. First, choose a cyclic sequence f of length $|f| = 3k + 1$. Starting from an arbitrary fixed point of the cycle, we can traverse $3k$ characters, or k complete codons, and have yet in reserve one spare nucleotide. Continuing cyclically, we utilize this remnant nucleotide as the first. Whence, codons in the second portion of k codons are all passed with the shift of nucleotides by one position to the left, with respect to the distribution into codons in the first $3k$-nucleotide string. Now, we have two remnant nucleotides, from the right "end" of which now constitute the first two nucleotides of the next codon. Making a third tour now of k more codons along the same sequence of nucleotides produces a sequence of codons which stops at the same point where it was originally begun. That is, we have overall traversed a sequence of $3k + 1$ complete codons where all the three possible shifts of the frame have been realized – meaning that the shifts have been made in a circular direction. Apparently, much the same holds true for a factor of length $3k - 1$. This completes the proof.

The next statement is related to the preceding one:

Lemma 2. *Let f be a cyclic nucleotide sequence obtained by the κ-fold repetition of a factor h of a length $|h|$, let the nucleotides not all be the same, and let the whole sequence be read just once. Then, there is conservation of a circular order of codons under any shift of frame if κ is a multiple of 3, while $|h|$ is not.*

Proof: First, take a sequence f which is the κ-fold repetition of a factor h of a length $|h|$ not divisible by 3 and with κ being a multiple of three, as in conditions of Lemma 1. Since the tour around such a κ-fold cycle is tantamount to the κ-fold rotation along a cycle of length h (obtained by cyclically closing a factor h), the application of Lemma 1 gives here the proof of the statement.

It is convenient to merge both lemmas to state the following:

Proposition 3. *Let f be cyclic sequence of nucleotides. Then, f conserves a circular order of its codons under any shift of its frame if (0) all the nucleotides are the same, (1) f has a length λ not divisible by 3 and is consecutively read κ times, with κ a multiple of 3, or (2) f is composed of κ repeated copies of a factor h of length λ, where κ is divisible by 3, while λ is not.*

As a case in point, consider the sequence ***atcgatcgatcg***; here, the factor ***atcg*** of length $\lambda = 4$ is repeated three times. Translating this code without any shift gives isoleucine, aspartic acid, arginine, and serine, consecutively, or **IDRS** for short. The circular shift by 1 position results in **SIDR**, by 2 positions produces **RSID**, and (here) at last, the circular shift by 3 positions gives **DRSI**.

Apparently, all the four translated codes of amino acids are the same relative to some circular permutation. Besides codon conservation, the circumstances of Proposition 3 lead to a further consequence:

Proposition 4. *Let f be a cyclic sequence of different nucleotides satisfying one of the conditions of Proposition 3. Then, the cyclic sequence q of amino acids so translated from f is conserved under any cyclic shift of f (with q defined only relative to circular order).*

  The Propositions 3 and 4 allow to conclude that a minimal linear factor g of a nucleotide sequence which guarantees to produce, upon translation, the respective factor of the amino acid 'with accuracy to a circular permutation' takes the form $g = acccb$, where c is a factor of length $|c| = \lambda$ not divisible by 3; and prefix a & suffix b factors of a total length $|a| + |b| = 2$ $(0 = a; b = 2)$ correspond to the last and first, consecutive nucleotides of *c*, respectively. The adjective "minimal" stands here to allow circular shifts by 1 or 2 positions. If a (res. b) is a longer factor of c and $|a| + |b| \geq \lambda,$ then g allows a (not minimal) number $|a| + |b|$ of circular permutations of the translated factor g. Accordingly, one or two 'sparse' nucleotides form codons with 2 or 1 external nucleotides, respectively. Codonic nucleotides of the factor g encode a (periodic) factor of the respective amino acid sequence containing a not necessary integer number of repeated translates of ccc. *Combinatorially, just this controls producing circular permutations in a protein domain.*

  One might also consider the enumeration of the types of sequences of our Lemma 2, say, with the enumeration at fixed κ & λ. That is, we seek the number of equivalence classes of cyclic sequences, where two such sequences are *equivalent* if one can be obtained from the other via a cyclic permutation (*i. e.*, a power of the permutation which cycles the members one unit along the sequence, with the last member permuted to the first). We let $\#_{\kappa, \lambda}$ be the number of such equivalence classes consisting of κ segments each of length $\lambda,$ with κ divisible by 3 and λ not. Then:

Proposition 5. *Let $\#_{\kappa, \lambda}$ be the number of equivalence classes of cyclic (nucleotide) sequences having κ segments of length λ. Then, $\#_{\kappa, \lambda} = \#_{1, \lambda}$.*

Proof: Each circular shift of an arbitrary (κ, λ)-sequence by one position is equivalent to asynchronous circular shift of every factor of length λ. Such a factor, if considered in a cyclic fashion, represents a $(1, \lambda)$-sequence, so that the number of distinct circular arrangements of nucleotides in both (fixed) (κ, λ)- and $(1, \lambda)$-sequences is the same. Since this is true for any (κ, λ)-sequence separately, it holds true for the entire set S of all circularly nonequivalent (κ, λ)-sequences with the set F of their representative λ-

factors (which are all distinct). With this one-to-one correspondence between the two sets, the proof is completed.

But now we note that $\#_{1,\lambda}$ is solved by Pólya's enumeration theory [2]. Indeed, a problem somewhat like this is a standard enumeration in many combinatorics texts: one ordinarily enumerates equivalence classes of beads on a necklace, with equivalence being determined by the dihedral group, rather than the cyclic group as here. The additional "reflective" permutations of the dihedral group are absent in our case, since our nucleotide "beads" have a direction (or orientation) along the sequence.

But further, we might clarify a point concerning sequences of types (κ, λ) and (κ', λ') with κ & κ' each divisible by 3 while λ & λ' not. In particular, it can turn out that some sequences can be of both types when there is a fixed total number of nucleotides $\#_{\kappa,\lambda} = \#_{\kappa',\lambda'} \equiv N$. In particular, if N has a maximum power $p > 1$ of 3 as a divisor, then $N \equiv 3^p \lambda''$ with λ'' not divisible by 3, and sequences of type (κ'', λ'') (with $\kappa'' = 3^p$, as both κ & κ' are divisible by κ''). Indeed, all the sequences of a type (κ, λ) are again counted in those of type (κ', λ') if κ is a divisor of κ' Thus, we might introduce the count $\#_{\kappa,\lambda}$ of cyclic sequences such that this includes no cyclic sequences of other types. By virtue of the Proposition 5, we can reduce this count to the enumeration of all cyclic $(1, \lambda)$-sequences that are a repetition of no block of length λ' being a divisor of λ. The calculation of $\#'_{\kappa,\lambda}$ first includes determining the numbers $\#_{1,\lambda'}$ for all distinct divisors λ' of λ; and, then, the general inclusion-exclusion procedure applies [2]. Now, one can in fact obtain those counts in terms of the classical Möbius functions [2]. In particular, the number-theoretic Möbius function $\mu(n)$ is defined as follows:

$$\mu(n) := \begin{cases} 0 & \text{if } n \text{ is not square-free;} \\ (-1)^k & \text{if } n = \text{product of } k \text{ distinct primes.} \end{cases} \tag{1}$$

Using the inclusion-exclusion principle, we state the following:

Proposition 6. *Let* $\#'_{\kappa,\lambda}$ *be the number of cyclic* (κ, λ)*-sequences such that includes no cyclic sequences of other types. Then,*

$$\#'_{\kappa,\lambda} = \sum_{d|\lambda} \mu(\lambda/d) \#_{(\lambda\kappa/d),d} = \sum_{d|\lambda} \mu(\lambda/d) \#_{1,d} \tag{2}$$

Where the d *summation is over divisors of* λ.

Proof: The first equality in (2) follows from the general inclusion-exclusion principle applying the number-theoretic Möbius function $\mu(n)$, as this given by (1). The second equality follows from the Proposition 5, which completes the overall proof.

Preceding experimental observations [3–9] of the last 30 years have unequivocally demonstrated the existence of naturally occurring cyclic permutations of the amino acid sequence of a protein. Our present Propositions 3 and 4 determine a sufficient combinatorial condition imposed on respective factors of a nucleotide sequence to guarantee the practical occurrence of this phenomenon.

Concluding, we also mention that in a wider context, which includes also an algebraic simulation of *alternative splicing*, two other cyclic invariances of nucleotide sequences were earlier considered by Propositions 1 and 2 in [10], which do not directly take into account the distribution of nucleotides into codons. Besides, in nature, there are cases of biologically tolerated shuffling of factors of a nucleotide sequence which conserves the inventory of translated amino acids, together with all multiplicities thereof [11]. In other words, there exist also noncircular permutations of a nucleotide sequence conserving the ratios of codonically encoded amino acids (and, maybe, the assortment of codons themselves, without equivalent replacements thereof), whereas a circular order in which they (would normally) follow may be altered. Here, "would" is also used to anticipate a possible perspective of gene engineering which might apply such a principle. Presumably, this may give a new impetus to further interdisciplinary studies of invariant permutable codes, including those which are not cyclically invariant.

ACKNOWLEDGMENTS

REFERENCES

1. J.D.D. Watson, "Molecular Biology of the Gene", 3rd ed., W.A. Benjamin Inc., New York, **1976**.

2. G.H. Hardy and E.M. Wright, An introduction to the Theory of Numbers, Oxford University Press, London, 1938; the 5th edition, 1979.

3. B.A. Cunningham, J.J., Hemperley T.P. Hopp, G.M. Edelman, *Proc. Natl. Acad. Sci. USA*, **1976**, *76*, 3215.

4. M. Hahn, K. Piotukh, R. Borriss, and U. Heinemann, *Proc. Natl. Acad. Sci. USA*, **1994,** *91* (*22*), 10417.

5. Y. Lindqvist and G. Schneider, *Curr. Opin. Struct. Biol., 1997*, *7* (*3*), 422.

6. J. Av, M. Hahn, K. Decanniere, K. Piotukh, R. Borriss, and U. Heinemann, *Proteins*, **1998**, *30* (*2*), 155.

7. S. Uliel, A. Fliess, A. Amir, and R. Unger*, Bioinformatics*, **1999**, *15* (*11*), 930.

8. S. Uliel, A. Fliess, and R. Unger, *Protein Engineering*, **2001**, *14* (*8*), 533.

9. J. Weiner 3rd, G. Thomas, and E. Bornberg-Bauer, *Bioinformatics,* **2005**, *21* (*7*), 932.

10. V.R. Rosenfeld, *MATCH Commun. Math. Comput. Chem.*, **2006**, *56* (*2*), 281.

11. E.A. Nalefski and J.J. Falke, *Protein Sci.,* **1996**, *5*, 2375.

COMPUTATION OF THE FIRST EDGE WIENER INDEX OF A COMPOSITION OF GRAPHS

MAHDIEH AZARI[a]*, ALI IRANMANESH[b], ABOLFAZL TEHRANIAN[a]

ABSTRACT The edge versions of Wiener index, based on distance between two edges in a connected graph G, were introduced by Iranmanesh et al. in 2009. In this paper, we find the first edge Wiener index of the composition of graphs.

Keywords: *Wiener index, finite graphs*

INTRODUCTION

Within this paper, we consider only simple, undirected, connected and finite graphs. A simple graph is a graph, without any loops or multiple bonds. Denote by $G = (V(G), E(G))$ a graph G with the set of vertices/ atoms $V(G)$ and the set of edges/bonds $E(G)$. For a (molecular) graph G, the degree of a vertex u is the number of edges incident to u and denoted by $\deg(u|G)$ and the distance between the vertices u and v of G, is denoted by $d(u, v|G)$ and it is defined as the number of edges in a shortest path, connecting u and v. In this paper, we denote by $[u, v]$, the edge connecting the vertices u, v of G.

A topological index is a real number related to the structural graph of a molecule. It dose not depend on the labeling or pictorial representation of a graph.

The ordinary (vertex) version of the Wiener index (or Wiener number) of G, is the sum of distances between all pairs of vertices of G, that is:

$$W(G) = W_v(G) = \sum_{\{u,v\} \subseteq V(G)} d(u, v|G).$$

This index was introduced by the Chemist, Harold Wiener [1], within the study of relations between the structure of organic compounds and their properties. This index is the first and most important topological index in Chemistry. So many interesting works have been done on it, in both Chemistry and Mathematics [2-13].

---

[a] *Department of Mathematics, Science and Research Branch, Islamic Azad University (IAU), P. O. Box: 14515-1775, Tehran, Iran, E-mail: azari@kau.ac.ir, tehranian@srbiau.ac.ir*
[b] *Department of Mathematics, Tarbiat Modares University, P. O. Box: 14115-137, Tehran, Iran, E-mail: iranmanesh@modares.ac.ir*

The Zagreb indices have been defined more than thirty years ago by Gutman and Trinajestic, [14].

Definition 1. [14] The first Zagreb index of G is defined as:

$$M_1(G) = \sum_{u \in V(G)} \deg(u|G)^2 .$$

The edge versions of Wiener index of G, which were based on the distance between all pairs of edges of G, were introduced by Iranmanesh et al. in 2009 [15]. We encourage the reader to consult [16-20], for computational techniques and mathematical properties of the edge Wiener indices. The first edge Wiener index of G, is defined as follows:

Definition 2. [15] The first edge Wiener index of G, is denoted by $W_{e_0}(G)$. That is:

$$W_{e_0}(G) = \sum_{\{e,f\} \subseteq E(G)} d_0(e,f|G) , \quad \text{where} \quad d_0(e,f|G) = \begin{cases} d_1(e,f|G)+1 & \text{if } e \neq f \\ 0 & \text{if } e = f \end{cases} \quad \text{and}$$

$d_1(e,f|G) = \min\{d(u,z|G), d(u,t|G), d(v,z|G), d(v,t|G)\}$, such that $e = [u,v]$, $f = [z,t]$. This index satisfies the relation $W_{e_0}(G) = W_v(L(G))$, where $L(G)$ is the line graph of G.

In this paper, we want to find the first edge Wiener index of the composition of graphs.

Recall definition of the composition of two graphs.

Definition 3. Let $G_1 = (V(G_1), E(G_1))$ and $G_2 = (V(G_2), E(G_2))$ be two connected graphs. We denote the composition of G_1 and G_2 by $G_1[G_2]$, that is a graph with the vertex set $V(G_1[G_2]) = V(G_1) \times V(G_2)$ and two vertices (u_1, u_2) and (v_1, v_2) of $G_1[G_2]$ are adjacent if and only if: $[u_1 = v_1 \text{ and } [u_2, v_2] \in E(G_2)]$ or $[u_1, v_1] \in E(G_1)$.

By definition of the composition, the distance between every pair of distinct vertices $u = (u_1, u_2)$ and $v = (v_1, v_2)$ of $G_1[G_2]$, is equal to

$$d(u,v|G_1[G_2]) = \begin{cases} d(u_1,v_1|G_1) & \text{if } u_1 \neq v_1 \\ 1 & \text{if } u_1 = v_1, \ [u_2,v_2] \in E(G_2) \\ 2 & \text{if } u_1 = v_1, \ v_2 \text{ is not adjacent to } u_2 \text{ in } G_2 \end{cases}$$

COMPUTATION OF THE FIRST EDGE WIENER INDEX OF THE COMPOSITION OF GRAPHS

Let $G_1 = (V(G_1), E(G_1))$ and $G_2 = (V(G_2), E(G_2))$ be two graphs. Consider the sets E_1 and E_2 as follows:

184

$$E_1 = \{[(u_1, u_2), (u_1, v_2)] \in E(G_1[G_2]) : u_1 \in V(G_1), \ [u_2, v_2] \in E(G_2)\}$$
$$E_2 = \{[(u_1, u_2), (v_1, v_2)] \in E(G_1[G_2]) : [u_1, v_1] \in E(G_1), \ u_2, v_2 \in V(G_2)\}$$

By definition of the composition, $E_1 \cup E_2 = E(G_1[G_2])$ and obviously,

$$E_1 \cap E_2 = \phi, \ |E_1| = |V(G_1)| \ |E(G_2)| \text{ and } |E_2| = |V(G_2)|^2 |E(G_1)|.$$

Set:

$$A = \{\{e, f\} \subseteq E(G_1[G_2]) : e \neq f, \ e, f \in E_1\}$$
$$B = \{\{e, f\} \subseteq E(G_1[G_2]) : e \neq f, \ e, f \in E_2\}$$
$$C = \{\{e, f\} \subseteq E(G_1[G_2]) : e \in E_1, \ f \in E_2\}$$

It is easy to see that each pair of the above sets is disjoint and the union of them is the set of all two element subsets of $E(G_1[G_2])$. Also we have:

$$|A| = \binom{|E_1|}{2} = \binom{|V(G_1)| \ |E(G_2)|}{2},$$

$$|B| = \binom{|E_2|}{2} = \binom{|V(G_2)|^2 |E(G_1)|}{2},$$

$$|C| = |E_1| \ |E_2| = |V(G_1)| \ |V(G_2)|^2 |E(G_1)| \ |E(G_2)|$$

Consider four subsets A_1, A_2, A_3 and A_4 of the set A as follows:

$$A_1 = \{\{e, f\} \in A : e = [(u_1, u_2), (u_1, v_2)], f = [(u_1, u_2), (u_1, z_2)], \ u_1 \in V(G_1),$$
$$u_2, v_2, z_2 \in V(G_2)\}$$

$$A_2 = \{\{e, f\} \in A : e = [(u_1, u_2), (u_1, v_2)], f = [(u_1, z_2), (u_1, t_2)], \ u_1 \in V(G_1),$$
$$u_2, v_2, z_2, t_2 \in V(G_2), \ \text{both } z_2 \text{ and } t_2 \text{ are adjacent neither}$$
$$\text{to } u_2 \text{ nor to } v_2 \text{ in } G_2\}$$

$$A_3 = \{\{e, f\} \in A : e = [(u_1, u_2), (u_1, v_2)], f = [(u_1, z_2), (u_1, t_2)], \ u_1 \in V(G_1),$$
$$u_2, v_2 \in V(G_2), \ z_2, t_2 \in V(G_2) - \{u_2, v_2\}\} - A_2$$

$$A_4 = \{\{e, f\} \in A : e = [(u_1, u_2), (u_1, v_2)], f = [(v_1, z_2), (v_1, t_2)], \ u_1, v_1 \in V(G_1),$$
$$v_1 \neq u_1, \ u_2, v_2, z_2, t_2 \in V(G_2)\}$$

It is clear that, every pair of the above sets is disjoint and $A = \bigcup_{i=1}^{4} A_i$.

In the next Proposition, we characterize $d_0(e, f | G_1[G_2])$ for all $\{e, f\} \in A$.

Proposition 1. Let $\{e, f\} \in A$.

(i) If $\{e,f\} \in A_1$, then $d_0(e,f|G_1[G_2]) = 1$

(ii) If $\{e,f\} \in A_2$, then $d_0(e,f|G_1[G_2]) = 3$

(iii) If $\{e,f\} \in A_3$, then $d_0(e,f|G_1[G_2]) = 2$

(iv) If $\{e,f\} \in A_4$, then $d_0(e,f|G_1[G_2]) = 1 + d(u_1,v_1|G_1)$,

where $e = [(u_1,u_2),(u_1,v_2)], f = [(v_1,z_2),(v_1,t_2)]$

Proof. (i) Let $\{e,f\} \in A_1$ and $e = [(u_1,u_2),(u_1,v_2)], f = [(u_1,u_2),(u_1,z_2)]$.
Due to distance between two vertices in $G_1[G_2]$ and by definition of $d_0(e,f)$, we have:

$$d_0(e,f|G_1[G_2]) = 1 + \min\{d((u_1,u_2),(u_1,u_2)|G_1[G_2]), d((u_1,u_2),(u_1,z_2)|G_1[G_2]),$$

$$d((u_1,v_2),(u_1,u_2)|G_1[G_2]), d((u_1,v_2),(u_1,z_2)|G_1[G_2])\} = 1 + \min\{0,1,1,d(v_2,z_2|G_2)\} = 1 + 0 = 1$$

(ii) Let $\{e,f\} \in A_2$ and $e = [(u_1,u_2),(u_1,v_2)], f = [(u_1,z_2),(u_1,t_2)]$.
By definition of the set A_2, z_2 is adjacent neither to u_2 nor to v_2 in G_2 and this is also true for t_2. Therefore,

$$d_0(e,f|G_1[G_2]) = 1 + \min\{d((u_1,u_2),(u_1,z_2)|G_1[G_2]), d((u_1,u_2),(u_1,t_2)|G_1[G_2]),$$

$$d((u_1,v_2),(u_1,z_2)|G_1[G_2]), d((u_1,v_2),(u_1,t_2)|G_1[G_2])\} = 1 + \min\{2,2,2,2\} = 3.$$

(iii) Let $\{e,f\} \in A_3$ and $e = [(u_1,u_2),(u_1,v_2)], f = [(u_1,z_2),(u_1,t_2)]$. By definition of the set A_3, $z_2 \notin \{u_2,v_2\}$, $t_2 \notin \{u_2,v_2\}$.
On the other hand $\{e,f\} \notin A_2$, so at least one of the following situations occurs:

$[u_2,z_2] \in E(G_2)$, $[u_2,t_2] \in E(G_2)$, $[v_2,z_2] \in E(G_2)$ or
$[v_2,t_2] \in E(G_2)$.

This means that, at least one of the distances $d((u_1,u_2),(u_1,z_2)|G_1[G_2])$,
$d((u_1,u_2),(u_1,t_2)|G_1[G_2])$, $d((u_1,v_2),(u_1,z_2)|G_1[G_2])$ or
$d((u_1,v_2),(u_1,t_2)|G_1[G_2])$ is equal to 1. Therefore,

$$d_0(e,f|G_1[G_2]) = 1 + \min\{d((u_1,u_2),(u_1,z_2)|G_1[G_2]), d((u_1,u_2),(u_1,t_2)|G_1[G_2]),$$

$$d((u_1,v_2),(u_1,z_2)|G_1[G_2]), d((u_1,v_2),(u_1,t_2)|G_1[G_2])\} = 1 + 1 = 2.$$

(iv) Let $\{e,f\} \in A_4$ and $e = [(u_1,u_2),(u_1,v_2)], f = [(v_1,z_2),(v_1,t_2)]$.
Thus $v_1 \neq u_1$ and

$$d_0(e,f|G_1[G_2]) = 1 + \min\{d((u_1,u_2),(v_1,z_2)|G_1[G_2]), d((u_1,u_2),(v_1,t_2)|G_1[G_2]),$$

$$d((u_1,v_2),(v_1,z_2)|G_1[G_2]), d((u_1,v_2),(v_1,t_2)|G_1[G_2])\} =$$

$1 + \min\{d(u_1, v_1|G_1), d(u_1, v_1|G_1), d(u_1, v_1|G_1), d(u_1, v_1|G_1)\} = 1 + d(u_1, v_1|G_1)$,

so the proof is completed

In follow, we define five subsets B_1, B_2, B_3, B_4 and B_5 of the set B.

$B_1 = \{\{e, f\} \in B : e = [(u_1, u_2), (v_1, v_2)], f = [(u_1, u_2), (v_1, z_2)], \ u_1, v_1 \in V(G_1),$
$\quad u_2, v_2, z_2 \in V(G_2)\}$

$B_2 = \{\{e, f\} \in B : e = [(u_1, u_2), (v_1, v_2)], f = [(u_1, z_2), (v_1, t_2)], \ u_1, v_1 \in V(G_1),$
$\quad u_2, v_2, z_2, t_2 \in V(G_2), \ z_2 \neq u_2, \ t_2 \neq v_2\}$

$B_3 = \{\{e, f\} \in B : e = [(u_1, u_2), (v_1, v_2)], f = [(u_1, u_2), (z_1, z_2)], \ u_1, v_1, z_1 \in V(G_1),$
$\quad u_2, v_2, z_2 \in V(G_2), \ z_1 \neq v_1\}$

$B_4 = \{\{e, f\} \in B : e = [(u_1, u_2), (v_1, v_2)], f = [(u_1, t_2), (z_1, z_2)], \ u_1, v_1, z_1 \in V(G_1),$
$\quad u_2, v_2, t_2, z_2 \in V(G_2), \ z_1 \neq v_1, \ t_2 \neq u_2\}$

$B_5 = \{\{e, f\} \in B : e = [(u_1, u_2), (v_1, v_2)], f = [(z_1, z_2), (t_1, t_2)], \ u_1, v_1 \in V(G_1),$
$\quad z_1, t_1 \in V(G_1) - \{u_1, v_1\}, \ u_2, v_2, z_2, t_2 \in V(G_2)\}$

It is clear that, each pair of the above sets is disjoint and $B = \bigcup_{i=1}^{5} B_i$.

The next Proposition, characterizes $d_0(e, f|G_1[G_2])$ for all $\{e, f\} \in B$.

Proposition 2. Let $\{e, f\} \in B$.

(i) If $\{e, f\} \in B_1$, then $d_0(e, f|G_1[G_2]) = 1$

(ii) If $\{e, f\} \in B_2$, then $d_0(e, f|G_1[G_2]) = 2$

(iii) If $\{e, f\} \in B_3$, then $d_0(e, f|G_1[G_2]) = d_0([u_1, v_1], [u_1, z_1]|G_1)$,

where $e = [(u_1, u_2), (v_1, v_2)], f = [(u_1, u_2), (z_1, z_2)]$

(iv) If $\{e, f\} \in B_4$, then $d_0(e, f|G_1[G_2]) = d_0([u_1, v_1], [u_1, z_1]|G_1) + 1$,

where $e = [(u_1, u_2), (v_1, v_2)], f = [(u_1, t_2), (z_1, z_2)]$

(v) If $\{e, f\} \in B_5$, then $d_0(e, f|G_1[G_2]) = d_0([u_1, v_1], [z_1, t_1]|G_1)$,

where $e = [(u_1, u_2), (v_1, v_2)], f = [(z_1, z_2), (t_1, t_2)]$

Proof. (i) Let $\{e, f\} \in B_1$ and $e = [(u_1, u_2), (v_1, v_2)], f = [(u_1, u_2), (v_1, z_2)]$.
Using the definition of $d_0(e, f)$, we have:

$d_0(e, f|G_1[G_2]) = 1 + \min\{d((u_1, u_2), (u_1, u_2)|G_1[G_2]), d((u_1, u_2), (v_1, z_2)|G_1[G_2]),$
$d((v_1, v_2), (u_1, u_2)|G_1[G_2]), d((v_1, v_2), (v_1, z_2)|G_1[G_2])\} =$
$1 + \min\{0, 1, 1, d((v_1, v_2), (v_1, z_2)|G_1[G_2])\} = 1 + 0 = 1$.

(ii) Let $\{e,f\} \in B_2$ and $e = [(u_1,u_2),(v_1,v_2)], f = [(u_1,z_2),(v_1,t_2)]$. By definition of B_2, $z_2 \neq u_2$, $t_2 \neq v_2$. So due to distance between two vertices in $G_1[G_2]$, the distances $d((u_1,u_2),(u_1,z_2)|G_1[G_2])$ and $d((v_1,v_2),(v_1,t_2)|G_1[G_2])$ are either 1 or 2. Therefore,

$$d_0(e,f|G_1[G_2]) = 1 + \min\{d((u_1,u_2),(u_1,z_2)|G_1[G_2]), d((u_1,u_2),(v_1,t_2)|G_1[G_2]),$$

$$d((v_1,v_2),(u_1,z_2)|G_1[G_2]), d((v_1,v_2),(v_1,t_2)|G_1[G_2])\} =$$

$$1 + \min\{d((u_1,u_2),(u_1,z_2)|G_1[G_2]),1,1, d((v_1,v_2),(v_1,t_2)|G_1[G_2]\} = 1 + 1 = 2$$

(iii) Let $\{e,f\} \in B_3$ and $e = [(u_1,u_2),(v_1,v_2)], f = [(u_1,u_2),(z_1,z_2)]$. By the definition of B_3 we have $z_1 \neq v_1$ and hence

$$d_0(e,f|G_1[G_2]) = 1 + \min\{d((u_1,u_2),(u_1,u_2)|G_1[G_2]), d((u_1,u_2),(z_1,z_2)|G_1[G_2]),$$

$$d((v_1,v_2),(u_1,u_2)|G_1[G_2]), d((v_1,v_2),(z_1,z_2)|G_1[G_2])\} =$$

$$1 + \min\{d(u_1,u_1|G_1), d(u_1,z_1|G_1), d(v_1,u_1|G_1), d(v_1,z_1|G_1)\} = d_0([u_1,v_1],[u_1,z_1]|G_1) \quad \textbf{(iv)}$$

Let $\{e,f\} \in B_4$ and $e = [(u_1,u_2),(v_1,v_2)], f = [(u_1,t_2),(z_1,z_2)]$. By definition of B_4, $z_1 \neq v_1$, $t_2 \neq u_2$. So $d(v_1,z_1|G_1) \geq 1$ and $d((u_1,u_2),(u_1,t_2)|G_1[G_2]) \geq 1$. Therefore

$$d_0(e,f|G_1[G_2]) = 1 + \min\{d((u_1,u_2),(u_1,t_2)|G_1[G_2]), d((u_1,u_2),(z_1,z_2)|G_1[G_2]),$$

$$d((v_1,v_2),(u_1,t_2)|G_1[G_2]), d((v_1,v_2),(z_1,z_2)|G_1[G_2])\} =$$

$$1 + \min\{d((u_1,u_2),(u_1,t_2)|G_1[G_2]),1,1, d(v_1,z_1|G_1)\} = 1 + 1 = d_0([u_1,v_1],[u_1,z_1]|G_1) + 1 \quad \textbf{(v)}$$

Let $\{e,f\} \in B_5$ and $e = [(u_1,u_2),(v_1,v_2)], f = [(z_1,z_2),(t_1,t_2)]$. By the definition of B_5, $z_1 \neq u_1$, $z_1 \neq v_1$, $t_1 \neq u_1$ and $t_1 \neq v_1$. So the edges $[u_1,v_1]$ and $[z_1,t_1]$ of G_1 are distinct. Therefore

$$d_0(e,f|G_1[G_2]) = 1 + \min\{d((u_1,u_2),(z_1,z_2)|G_1[G_2]), d((u_1,u_2),(t_1,t_2)|G_1[G_2]),$$

$$d((v_1,v_2),(z_1,z_2)|G_1[G_2]), d((v_1,v_2),(t_1,t_2)|G_1[G_2])\} =$$

$$1 + \min\{d(u_1,z_1|G_1), d(u_1,t_1|G_1), d(v_1,z_1|G_1), d(v_1,t_1|G_1)\} = d_0([u_1,v_1],[z_1,t_1]|G_1)$$

and the proof is completed.

Now, we consider three subsets C_1, C_2 and C_3 of the set C as follows:

$$C_1 = \{\{e,f\} \in C : e = [(u_1,u_2),(u_1,v_2)], f = [(u_1,u_2),(z_1,z_2)], \ u_1, z_1 \in V(G_1),$$
where $u_2, v_2, z_2 \in V(G_2)\}$

$$C_2 = \{\{e,f\} \in C : e = [(u_1,u_2),(u_1,v_2)], f = [(u_1,t_2),(z_1,z_2)], \ u_1, z_1 \in V(G_1),$$
where $u_2, v_2, t_2, z_2 \in V(G_2), \ t_2 \neq u_2, \ t_2 \neq v_2\}$

$$C_3 = \{\{e, f\} \in C : e = [(u_1, u_2), (u_1, t_2)], f = [(v_1, v_2), (z_1, z_2)], \ u_1, v_1, z_1 \in V(G_1),$$

where $u_2, t_2, v_2, z_2 \in V(G_2), \ v_1 \neq u_1, \ z_1 \neq u_1\}$

Clearly, every pair of the above sets is disjoint and $C = \bigcup_{i=1}^{3} C_i$.

In the following Proposition, we find $d_0(e, f | G_1[G_2])$ for all $\{e, f\} \in C$.

Proposition 3. Let $\{e, f\} \in C$.

    (i) If $\{e, f\} \in C_1$, then $d_0(e, f | G_1[G_2]) = 1$

    (ii) If $\{e, f\} \in C_2$, then $d_0(e, f | G_1[G_2]) = 2$

    (iii) If $\{e, f\} \in C_3$, then

$$d_0(e, f | G_1[G_2]) = 1 + \min\{d(u_1, v_1 | G_1), d(u_1, z_1 | G_1)\},$$

where $e = [(u_1, u_2), (u_1, t_2)], f = [(v_1, v_2), (z_1, z_2)]$

Proof. (i) Let $\{e, f\} \in C_1$ and $e = [(u_1, u_2), (u_1, v_2)], f = [(u_1, u_2), (z_1, z_2)]$.
By definition of $d_0(e, f)$, we have:

$$d_0(e, f | G_1[G_2]) = 1 + \min\{d((u_1, u_2), (u_1, u_2) | G_1[G_2]), d((u_1, u_2), (z_1, z_2) | G_1[G_2]),$$

$$d((u_1, v_2), (u_1, u_2) | G_1[G_2]), d((u_1, v_2), (z_1, z_2) | G_1[G_2])\} = 1 + \min\{0, 1, 1, 1\} = 1 + 0 = 1$$

    (ii) Let $\{e, f\} \in C_2$ and $e = [(u_1, u_2), (u_1, v_2)], f = [(u_1, t_2), (z_1, z_2)]$.
By definition of C_2 , $t_2 \neq u_2$, $t_2 \neq v_2$. Thus, due to the distance between
two vertices in $G_1[G_2]$, the distances $d((u_1, u_2), (u_1, t_2) | G_1[G_2])$ and
$d((u_1, v_2), (u_1, t_2) | G_1[G_2])$ are either 1 or 2. So

$$d_0(e, f | G_1[G_2]) = 1 + \min\{d((u_1, u_2), (u_1, t_2) | G_1[G_2]), d((u_1, u_2), (z_1, z_2) | G_1[G_2]),$$

$$d((u_1, v_2), (u_1, t_2) | G_1[G_2]), d((u_1, v_2), (z_1, z_2) | G_1[G_2])\} =$$

$$1 + \min\{d((u_1, u_2), (u_1, t_2) | G_1[G_2]), 1, d((u_1, v_2), (u_1, t_2) | G_1[G_2], 1\} = 1 + 1 = 2 .$$

    (iii) Let $\{e, f\} \in C_3$ and $e = [(u_1, u_2), (u_1, t_2)], f = [(v_1, v_2), (z_1, z_2)]$. By
definition of C_3, $v_1 \neq u_1$, $z_1 \neq u_1$. Therefore

$$d_0(e, f | G_1[G_2]) = 1 + \min\{d((u_1, u_2), (v_1, v_2) | G_1[G_2]), d((u_1, u_2), (z_1, z_2) | G_1[G_2]),$$

$$d((u_1, t_2), (v_1, v_2) | G_1[G_2]), d((u_1, t_2), (z_1, z_2) | G_1[G_2])\} =$$

$$1 + \min\{d(u_1, v_1 | G_1), d(u_1, z_1 | G_1), d(u_1, v_1 | G_1), d(u_1, z_1 | G_1)\} =$$

$$1 + \min\{d(u_1, v_1 | G_1), d(u_1, z_1 | G_1)\}, \text{ and the proof is completed}$$

Definition 4. Let $G = (V(G), E(G))$ be a graph.

(i) Let $u \in V(G)$. Set; $\Delta_u = \{z \in V(G) : [z, u] \in E(G)\}$. In fact, Δ_u is the set of all vertices of G, which are adjacent to u. Suppose that, δ_u is the number of all vertices of G, which are adjacent to u. Clearly, $\delta_u = |\Delta_u| = \deg(u|G)$.

(ii) For each pair of distinct vertices $u, v \in V(G)$, let $\delta_{(u,v)}$ be the number of all vertices of G, which are adjacent both to u and v. Obviously, $\delta_{(u,v)} = |\Delta_u \cap \Delta_v|$.

(iii) Let u, v and z be three vertices of G, which every pair of them is distinct. Assume that, $\delta_{(u,v,z)}$ denotes the number of all vertices of G which are adjacent to vertices u, v and z. It is easy to see that, $\delta_{(u,v,z)} = |\Delta_u \cap \Delta_v \cap \Delta_z|$.

(iv) Suppose that, u, v and z be three vertices of graph G, which every pair of them is distinct. Denote by $N_{(z,\tilde{u},\tilde{v})}$, the number of all vertices of G, which are adjacent to z, but neither to u nor to v. By the definition of $N_{(z,\tilde{u},\tilde{v})}$, we have:

$$N_{(z,\tilde{u},\tilde{v})} = |\Delta_z - (\Delta_u \cup \Delta_v)| = |\Delta_z| - |\Delta_z \cap (\Delta_u \cup \Delta_v)| = |\Delta_z| - |(\Delta_z \cap \Delta_u) \cup (\Delta_z \cap \Delta_v)| =$$
$$|\Delta_z| - (|\Delta_z \cap \Delta_u| + |\Delta_z \cap \Delta_v| - |\Delta_z \cap \Delta_u \cap \Delta_v|) = \delta_z - \delta_{(z,u)} - \delta_{(z,v)} + \delta_{(z,u,v)} .$$

Proposition 5.

$$\sum_{\{e,f\} \in A} d_0(e, f | G_1[G_2]) = |E(G_2)|^2 \left(\binom{|V(G_1)|+1}{2} + W(G_1) \right) - \frac{1}{4} |V(G_1)| (2M_1(G_2) - N(G_2)) ,$$

where, $N(G_2) = \displaystyle\sum_{[u_2, v_2] \in E(G_2)} \sum_{z_2 \in V(G_2) - (\Delta_{u_2} \cup \Delta_{v_2})} N_{(z_2, \tilde{u}_2, \tilde{v}_2)} .$

Proof. At first, we need to find $|A_2|$ and $|A_2 \cup A_3|$. It is easy to see that

$$|A_2| = \frac{1}{4} |V(G_1)| \sum_{[u_2, v_2] \in E(G_2)} \sum_{z_2 \in V(G_2) - (\Delta_{u_2} \cup \Delta_{v_2})} N_{(z_2, \tilde{u}_2, \tilde{v}_2)} = \frac{1}{4} |V(G_1)| N(G_2) ,$$

$$|A_2 \cup A_3| = \frac{1}{2} |V(G_1)| \sum_{[u_2, v_2] \in E(G_2)} (|E(G_2)| - (\delta_{u_2} + \delta_{v_2} - 1)) =$$

$$\frac{1}{2} |V(G_1)| \left(\sum_{[u_2, v_2] \in E(G_2)} |E(G_2)| - \sum_{[u_2, v_2] \in E(G_2)} (\delta_{u_2} + \delta_{v_2}) + \sum_{[u_2, v_2] \in E(G_2)} 1 \right) =$$

$$\frac{1}{2} |V(G_1)| (|E(G_2)|^2 + |E(G_2)| - M_1(G_2)) ,$$

Recall that, each pair of the sets $A_i\,(1 \le i \le 4)$ is disjoint and $A = \bigcup_{i=1}^{4} A_i$, then by

Proposition 1, we have:

$$\sum_{\{e,f\} \in A} d_0(e,f|G_1[G_2]) = \sum_{i=1}^{4} \sum_{\{e,f\} \in A_i} d_0(e,f|G_1[G_2]) = |A_1| + 3|A_2| + 2|A_3| +$$

$$\sum \{1 + d(u_1,v_1|G_1) : \{e,f\} \in A_4, e = [(u_1,u_2),(u_1,v_2)], f = [(v_1,z_2),(v_1,t_2)]\} =$$
$$|A_1| + 3|A_2| + 2|A_3| + |A_4| +$$

$$\sum \{d(u_1,v_1|G_1) : \{e,f\} \in A_4, e = [(u_1,u_2),(u_1,v_2)], f = [(v_1,z_2),(v_1,t_2)]\} =$$
$$\sum_{i=1}^{4} |A_i| + (|A_2| + |A_3|) + |A_2| + |E(G_2)|^2 \sum_{\{u_1,v_1\} \subseteq V(G_1)} d(u_1,v_1|G_1) =$$

$$\left| \bigcup_{i=1}^{4} A_i \right| + |A_2 \cup A_3| + |A_2| + |E(G_2)|^2 W(G_1) = |A| + |A_2 \cup A_3| + |A_2| + |E(G_2)|^2 W(G_1) =$$

$$\binom{|V(G_1)|\,|E(G_2)|}{2} + \frac{1}{2}|V(G_1)|\,(|E(G_2)|^2 + |E(G_2)| - M_1(G_2)) + \frac{1}{4}|V(G_1)|N(G_2) +$$

$$|E(G_2)|^2 W(G_1) =$$

$$\frac{1}{2}\left(|V(G_1)|^2 |E(G_2)|^2 - |V(G_1)|\,|E(G_2)| + |V(G_1)|\,|E(G_2)|^2 + |V(G_1)|\,|E(G_2)| \right) -$$

$$\frac{1}{2}|V(G_1)|M_1(G_2) + \frac{1}{4}|V(G_1)|N(G_2) + |E(G_2)|^2 W(G_1) =$$

$$|E(G_2)|^2 \left(\binom{|V(G_1)|+1}{2} + W(G_1) \right) - \frac{1}{4}|V(G_1)|\,(2M_1(G_2) - N(G_2))|$$

Proposition 6.

$$\sum_{\{e,f\} \in B} d_0(e,f|G_1[G_2]) = |V(G_2)|^2 \binom{|V(G_2)|}{2} M_1(G_1) + |V(G_2)|^4 W_{e_0}(G_1)$$

Proof. For the proof of this proposition, we need to obtain $|B_1|, |B_2|$ and $|B_4|$. It is easy to see that:

$$|B_1| = 2|E(G_1)| \, |V(G_2)| \binom{|V(G_2)|}{2}, \ |B_2| = 2|E(G_1)| \binom{|V(G_2)|}{2}^2,$$

$$|B_4| = |V(G_2)|^3 (|V(G_2)|-1) \sum_{u_1 \in V(G_1)} \binom{\delta_{u_1}}{2} = |V(G_2)|^2 \binom{|V(G_2)|}{2} (M_1(G_1) - 2|E(G_1)|)$$

Afterwards, we find $\displaystyle\sum_{\{e,f\} \in B_3 \cup B_4 \cup B_5} d_0(e, f | G_1[G_2])$. By Proposition 2.2, we have:

$$\sum_{\{e,f\} \in B_3} d_0(e, f | G_1[G_2]) =$$

$$\sum \{ d_0([u_1, v_1], [u_1, z_1] | G_1) : \{e, f\} \in B_3, e = [(u_1, u_2), (v_1, v_2)], f = [(u_1, u_2), (z_1, z_2)] \} =$$

$$|V(G_2)|^3 \sum_{u_1 \in V(G_1)} \sum_{\{[u_1, v_1], [u_1, z_1]\} \subseteq E(G_1)} d_0([u_1, v_1], [u_1, z_1] | G_1) =$$

$$\frac{1}{2} |V(G_2)|^3 \sum_{[u_1, v_1] \in E(G_1)} \sum_{\substack{z_1 \in \{u_1, v_1\}, \\ [z_1, t_1] \in E(G_1)}} d_0([u_1, v_1], [z_1, t_1] | G_1),$$

$$\sum_{\{e,f\} \in B_4} d_0(e, f | G_1[G_2]) =$$

$$\sum \{ d_0([u_1, v_1], [u_1, z_1] | G_1) + 1 : \{e, f\} \in B_4, e = [(u_1, u_2), (v_1, v_2)], f = [(u_1, t_2), (z_1, z_2)] \} =$$

$$(|V(G_2)|^4 - |V(G_2)|^3) \sum_{u_1 \in V(G_1)} \sum_{\{[u_1, v_1], [u_1, z_1]\} \subseteq E(G_1)} d_0([u_1, v_1], [u_1, z_1] | G_1) + |B_4| =$$

$$\frac{1}{2} (|V(G_2)|^4 - |V(G_2)|^3) \sum_{[u_1, v_1] \in E(G_1)} \sum_{\substack{z_1 \in \{u_1, v_1\}, \\ [z_1, t_1] \in E(G_1)}} d_0([u_1, v_1], [z_1, t_1] | G_1) + |B_4|,$$

$$\sum_{\{e,f\} \in B_5} d_0(e, f | G_1[G_2]) =$$

$$\sum \{ d_0([u_1, v_1], [z_1, t_1] | G_1) : \{e, f\} \in B_5, e = [(u_1, u_2), (v_1, v_2)], f = [(z_1, z_2), (t_1, t_2)] \} =$$

$$\frac{1}{2} |V(G_2)|^4 \sum_{[u_1, v_1] \in E(G_1)} \sum_{\substack{[z_1, t_1] \in E(G_1), \\ z_1, t_1 \notin \{u_1, v_1\}}} d_0([u_1, v_1], [z_1, t_1] | G_1).$$

Based on the above computations and since each pair of $B_i (1 \le i \le 5)$ is disjoint, we have:

$$\sum_{\{e,f\} \in B_3 \cup B_4 \cup B_5} d_0(e, f | G_1[G_2]) =$$

$$\sum_{i=3}^{5} \sum_{\{e,f\} \in B_i} d_0(e, f | G_1[G_2]) = \frac{1}{2} |V(G_2)|^3 \sum_{[u_1, v_1] \in E(G_1)} \sum_{\substack{z_1 \in \{u_1, v_1\}, \\ [z_1, t_1] \in E(G_1)}} d_0([u_1, v_1], [z_1, t_1] | G_1) +$$

$$\frac{1}{2}\left(|V(G_2)|^4 - |V(G_2)|^3\right) \sum_{[u_1,v_1]\in E(G_1)} \sum_{\substack{z_1\in\{u_1,v_1\}, \\ [z_1,t_1]\in E(G_1)}} d_0([u_1,v_1],[z_1,t_1]|G_1) + |B_4| +$$

$$\frac{1}{2}|V(G_2)|^4 \sum_{[u_1,v_1]\in E(G_1)} \sum_{\substack{[z_1,t_1]\in E(G_1), \\ z_1,t_1\notin\{u_1,v_1\}}} d_0([u_1,v_1],[z_1,t_1]|G_1) =$$

$$\frac{1}{2}|V(G_2)|^4 \sum_{[u_1,v_1]\in E(G_1)} \sum_{\substack{z_1\in\{u_1,v_1\}, \\ [z_1,t_1]\in E(G_1)}} d_0([u_1,v_1],[z_1,t_1]|G_1) + |B_4| +$$

$$\frac{1}{2}|V(G_2)|^4 \sum_{[u_1,v_1]\in E(G_1)} \sum_{\substack{[z_1,t_1]\in E(G_1), \\ z_1,t_1\notin\{u_1,v_1\}}} d_0([u_1,v_1],[z_1,t_1]|G_1) =$$

$$|B_4| + \frac{1}{2}|V(G_2)|^4(2W_{e_0}(G_1)) = |B_4| + |V(G_2)|^4 W_{e_0}(G_1).$$

Now, since $B = \bigcup_{i=1}^{5} B_i$, we have:

$$\sum_{\{e,f\}\in B} d_0(e,f|G_1[G_2]) = \qquad \sum_{\{e,f\}\in B_1} d_0(e,f|G_1[G_2]) + \qquad \sum_{\{e,f\}\in B_2} d_0(e,f|G_1[G_2]) +$$

$$\sum_{\{e,f\}\in B_3\cup B_4\cup B_5} d_0(e,f|G_1[G_2]) =$$

$$|B_1| + 2|B_2| + |B_4| + |V(G_2)|^4 W_{e_0}(G_1) = 2|E(G_1)|\binom{|V(G_2)|}{2}\left(|V(G_2)| + 2\binom{|V(G_2)|}{2} - |V(G_2)|^2\right) +$$

$$|V(G_2)|^2\binom{|V(G_2)|}{2}M_1(G_1) + |V(G_2)|^4 W_{e_0}(G_1) = |V(G_2)|^2\binom{|V(G_2)|}{2}M_1(G_1) + |V(G_2)|^4 W_{e_0}(G_1).$$

Proposition 7.
$$\sum_{\{e,f\}\in C} d_0(e,f|G_1[G_2]) =$$

$$|E(G_1)|\ |E(G_2)|\ |V(G_2)|\ (|V(G_1)|\ |V(G_2)| + 2|V(G_2)| - 4) + |E(G_2)|\ |V(G_2)|^2 Min(G_1)$$

where, $Min(G_1) = \sum_{u_1\in V(G_1)} \sum_{[v_1,z_1]\in E(G_1)} \min\{d(u_1,v_1|G_1), d(u_1,z_1|G_1)\}$

Proof. First, we find $|C_2|$ and $\sum_{\{e,f\}\in C_3} d_0(e,f|G_1[G_2])$. It is easy to see that:

$$|C_2| = |V(G_2)|\ (|V(G_2)| - 2)\ |E(G_2)| \sum_{u_1\in V(G_1)} \delta_{u_1} = 2|E(G_1)|\ |E(G_2)|\ |V(G_2)|\ (|V(G_2)| - 2)$$

and by Proposition 3, we have:

193

$$\sum_{\{e,f\}\in C_3} d_0(e,f|G_1[G_2]) =$$

$$\sum \{1 + \min\{d(u_1,v_1|G_1), d(u_1,z_1|G_1) : \{e,f\} \in C_3, e = [(u_1,u_2),(u_1,t_2)], f = [(v_1,v_2),(z_1,z_2)]\} =$$

$$|C_3| + |E(G_2)| |V(G_2)|^2 \sum_{u_1 \in V(G_1)} \sum_{\substack{[v_1,z_1] \in E(G_1) \\ v_1 \neq z_1, z \neq_1 u_1}} \min\{d(u_1,v_1|G_1), d(u_1,z_1|G_1)\} =$$

$$|C_3| + |E(G_2)| |V(G_2)|^2 \sum_{u_1 \in V(G_1)} \sum_{[v_1,z_1] \in E(G_1)} \min\{d(u_1,v_1|G_1), d(u_1,z_1|G_1)\} =$$

$$|C_3| + |E(G_2)| |V(G_2)|^2 Min(G_1).$$

Since each pair of the sets $C_i (1 \leq i \leq 3)$ is disjoint and $C = \bigcup_{i=1}^{3} C_i$, we have:

$$\sum_{\{e,f\}\in C} d_0(e,f|G_1[G_2]) = \sum_{\{e,f\}\in C_1 \cup C_2} d_0(e,f|G_1[G_2]) + \sum_{\{e,f\}\in C_3} d_0(e,f|G_1[G_2]) =$$

$$|C_1| + 2|C_2| + |C_3| + |E(G_2)| |V(G_2)|^2 Min(G_1) =$$

$$\sum_{i=1}^{3} |C_i| + |C_2| + |E(G_2)| |V(G_2)|^2 Min(G_1) =$$

$$\left|\bigcup_{i=1}^{3} C_i\right| + |C_2| + |E(G_2)| |V(G_2)|^2 Min(G_1) =$$

$$|C| + |C_2| + |E(G_2)| |V(G_2)|^2 Min(G_1) = |V(G_1)| |V(G_2)|^2 |E(G_1)| |E(G_2)| +$$

$$2|E(G_1)| |E(G_2)| |V(G_2)| (|V(G_2)| - 2) + |E(G_2)| |V(G_2)|^2 Min(G_1) =$$

$$|E(G_1)| |E(G_2)| |V(G_2)| (|V(G_1)| |V(G_2)| + 2|V(G_2)| - 4) + |E(G_2)| |V(G_2)|^2 Min(G_1)$$

Now, as the main purpose of this paper, we express the following theorem, which characterizes the first edge Wiener index of the composition of two graphs.

Theorem. Let $G_1 = (V(G_1), E(G_1))$ and $G_2 = (V(G_2), E(G_2))$ be two simple undirected connected finite graphs, then

$$W_{e_0}(G_1[G_2]) = |E(G_2)|^2 \binom{|V(G_1)|+1}{2} +$$

$$|E(G_1)| |E(G_2)| |V(G_2)| (|V(G_1)| |V(G_2)| + 2|V(G_2)| - 4) +$$

$$|E(G_2)|^2 W(G_1) + |V(G_2)|^4 W_{e_0}(G_1) + |V(G_2)|^2 \binom{|V(G_2)|}{2} M_1(G_1) +$$

$$|E(G_2)| |V(G_2)|^2 Min(G_1) -$$

$$\frac{1}{4}\big|V(G_1)\big|\,(2M_1(G_2)-N(G_2)),$$

where $Min(G_1) = \displaystyle\sum_{u_1 \in V(G_1)}\sum_{[v_1,z_1]\in E(G_1)} \min\{d(u_1,v_1|G_1),d(u_1,z_1|G_1)\}$ and

$$N(G_2) = \sum_{[u_2,v_2]\in E(G_2)}\sum_{z_2 \in V(G_2)-(\Delta_{u_2}\cup\Delta_{v_2})} N_{(z_2,\tilde{u}_2,\tilde{v}_2)}\,.$$

Proof. Recall that, each pair of the sets A, B and C is disjoint and union of them is the set of all two element subsets of $E(G_1[G_2])$. Now, using the definition of the first edge Wiener index, we obtain:

$$W_{e_0}(G_1[G_2]) = \sum_{\{e,f\}\subseteq E(G_1[G_2])}d_0(e,f|G_1[G_2]) = \sum_{\{e,f\}\in A}d_0(e,f|G_1[G_2])+$$

$\displaystyle\sum_{\{e,f\}\in B}d_0(e,f|G_1[G_2]) + \sum_{\{e,f\}\in C}d_0(e,f|G_1[G_2])$. Now, by the above Lemmas, the proof is completed.

REFERENCES

1. H. Wiener, *J. Am. Chem. Soc.*, **1947**, *69*, 17.
2. F. Buckley and F. Harary, *Distance in graphs*, Addison-Wesley, Redwood, CA, 1990.
3. A.A. Dobrynin, R. Entringer and I. Gutman, *Acta Appl. Math.*, **2001**, *66*, 211.
4. A.A. Dobrynin, I. Gutman, S. Klavzar and P. Zigert, *Acta Appl. Math.*, **2002**, *72*, 247.
5. A.A. Dobrynin and L.S. Mel'nikov, *MATCH Commun. Math. Comput. Chem.*, **2004**, *50*, 145.
6. A.A. Dobrynin and L.S. Mel'nikov, *MATCH Commun. Math. Comput. Chem.*, **2005**, *53*, 209.
7. I. Gutman, *J. Serb. Chem. Soc*, **2003**, *68*, 949.
8. I. Gutman and O.E. Polansky, *Mathematical concepts in organic chemistry*, Springer-Verlag, Berlin, Germany, **1986**.
9. A. Heydari and B. Taeri, *J. Comput. Theor. Nanosky*, **2007**, *4*, 158.
10. A. Iranmanesh, Y. Alizadeh and S. Mirzaie, *Fullerenes Nanotubes Carbon Nanostruct.,* **2009**, *17*, 560.
11. S. Klavzar and I. Gutman, *Disc. Appl. Math.*, **1997**, *80*, 73.
12. H.P. Schultz, *J. Chem. Inf. Comput. Sci.*, **1989**, *34*, 227.
13. G. Wagner, *Acta Appl. Math.*, **2006**, *91*, 119.
14. I. Gutman and N. Trinajstic, *Chem. Phys. Lett.*, **1972**, *17*, 535.
15. A. Iranmanesh, I. Gutman, O. Khormali and A. Mahmiani, *MATCH Commun. Math. Comput. Chem.*, **2009**, *61*, 663.

16. A. Iranmanesh and A.S. Kafrani, *MATCH Commun. Math. Comput. Chem.,* **2009**, *62*, 311.

17. A. Iranmanesh, A.S. Kafrani and O. Khormali, *Optoelectron. Adv. Mater - Rapid Commun.*, **2010**, *4*, 242.

18. A. Iranmanesh, O. Khormali and A. Ahmadi, *J. Comput. Theor. Nanosci.*, in press.

19. A. Mahmiani, O. Khormali and A. Iranmanesh, *Optoelectron. Adv. Mater - Rapid Commun.*, **2010**, *4*, 252.

20. A. Mahmiani, O. Khormali, A. Iranmanesh and A. Ahmadi, *Optoelectron. Adv. Mater - Rapid Commun.*, **2010**, *4*, 256.

21. B.E. Sagan, Y. Yeh and P. Zhang, *Inter. J. Quantum Chem.*, **1996**, *60*, 959.

22. M. Azari and A. Iranmanesh, *Ars Combinatoria*, in press.

ON OMEGA POLYNOMIAL OF ((4,7)3) NETWORK

MAHSA GHAZI[a], MODJTABA GHORBANI[a], KATALIN NAGY[b], MIRCEA V. DIUDEA[b]

ABSTRACT. The Omega polynomial $\Omega(x)$ was recently proposed by Diudea [*Carpath. J. Math.*, 2006, *22*, 43-47]. It is defined on the ground of "opposite edge strips" *ops*. The related polynomial: Sadhana $Sd(x)$ can also be calculated by ops counting. In this paper we compute these polynomials for the ((4,7)3) infinite network, designed by *Trs*(*Ca*(4,4)) sequence of map operations.

Keywords: polygonal structures, Omega and Sadhana polynomials

INTRODUCTION

A molecular graph is a simple graph such that its vertices correspond to the atoms and the edges to the covalent bonds. Note that hydrogen atoms are often omitted. Mathematical calculations are necessary in view of exploring important concepts in chemistry. Mathematical chemistry is a branch of theoretical chemistry enabling discussion and prediction of molecular structures or molecular properties, using methods of discrete mathematics, without referring to quantum mechanics. Chemical graph theory is an important tool in the study of molecular structures. This theory had an important impact in the development of chemical sciences.

Let $G(V,E)$ be a connected graph, with the vertex set $V(G)$ and edge set $E(G)$. Two edges $e = uv$ and $f = xy$ of G are called *codistant*, *e co f*, if they obey the following relation [1-3]:

$$d(v,x) = d(v,y) + 1 = d(u,x) + 1 = d(u,y) \tag{1}$$

Relation *co* is reflexive, that is, *e co e* holds for any edge *e* of G; it is also symmetric, if *e co f* then *f co e*. In general, relation *co* is not transitive, an example showing this fact is the complete bipartite graph $K_{2,n}$. If "*co*" is also transitive, thus an equivalence relation, then G is called a *co-graph* and the set of edges $C(e) := \{f \in E(G); f\ co\ e\}$ is called an *orthogonal cut oc* of G, $E(G)$ being the union of disjoint orthogonal cuts: $E(G) = C_1 \cup C_2 \cup ... \cup C_k$, $C_i \cap C_j = \varnothing, i \neq j$. Klavžar [4] has shown that relation *co* is a theta Djoković-Winkler relation [5,6].

[a] *Department of Mathematics, Faculty of Science, Shahid Rajaee, Teacher Training University, Tehran, 16785 – 136, I. R. Iran; mghorbani@srttu.edu*
[b] *Faculty of Chemistry and Chemical Engineering, "Babes-Bolyai" University, 400028 Cluj, Romania, diudea@gmail.com*

Let $e = uv$ and $f = xy$ be two edges of G which are *opposite* or topologically parallel and denote this relation by e *op* f. A set of opposite edges, within the same face/ring, eventually forming a strip of adjacent faces/rings, is called an *opposite* edge *strip*, *ops*, which is a quasi-ortogonal cut *qoc* (*i. e.*, the transitivity relation is not necessarily obeyed). Note that *co* relation is defined in the whole graph while *op* is defined only in a face/ring. The length of *ops* is maximal irrespective of the starting edge.

Let $m(G, s)$ be the number of *ops* strips of length s. The Omega polynomial is defined as [1]

$$\Omega(x) = \sum_s m(G,s) \cdot x^s \tag{2}$$

The first derivative (in $x=1$) equals the number of edges in the graph

$$\Omega'(1) = \sum_s m(G,s) \cdot s = e = |E(G)| \tag{3}$$

A topological index, called Cluj-Ilmenau [2], $CI=CI(G)$, was defined on Omega polynomial

$$CI(G) = \left\{ [\Omega'(1)]^2 - [\Omega'(1) + \Omega''(1)] \right\} \tag{4}$$

The Sadhana index $Sd(G)$ was defined by Khadikar *et al*. [7,8] as

$$Sd(G) = \sum_s m(G,s)(|E(G)|-s) \tag{5}$$

where $m(G,s)$ is the number of strips of length s. The Sadhana polynomial $Sd(G,x)$ was defined by Ashrafi *et al*. [9]

$$Sd(G,x) = \sum_s m(G,s) \cdot x^{|E(G)|-s} \tag{6}$$

Clearly, the Sadhana polynomial can be derived from the definition of Omega polynomial by replacing the exponent s by $|E(G)-s|$. Then the Sadhana index will be the first derivative of $Sd(x)$ evaluated at $x = 1$.

The aim of this study is to compute the Omega and Sadhana polynomials of the ((4,7)3) infinite network. This network can be seen as a modification of the graphene sheet [10-12].

RESULTS AND DISCUSSION

The design of ((4,7)3) network can be achieved by $Trs(Ca(4,4))$ sequence of map operations [13-16], where Ca is the pro-chiral "Capra" operation and Trs is the truncation operation, performed on selected atoms (those having the valence four); (4,4) is the Schläfli symbol [17] for the planar net made by squares and vertices of degree/valence four, which was taken as a ground for the map operations. Figure 1 illustrates the ((4,7)3) pattern. Since any net has its co-net, depending of the start/end view, the co-net of ((4,7)3) net (Figure 2) will also be considered.

Looking to these nets, one can see that there are $k^2+(k-1)^2$ squares in the net and $4k(k-1)$ in co-net, k being the number of repeat units. This implies there exactly exist $2(k^2+(k-1)^2)$ strips of length 2 in the net and $8k(k-1)$ in co-net and the others are of length 1. By definition of Omega polynomial, the formulas for the two polynomials and derived indices (Table 1) can be easily obtained. Some examples to prove the above formulas are collected in Table 2.

Figure 1. The 2-dimensional ((4,7)3) net (3×3 units) designed by the sequence of map operations *Trs*(*Ca*(4,4)): non-optimized (left) and optimized (right) structure.
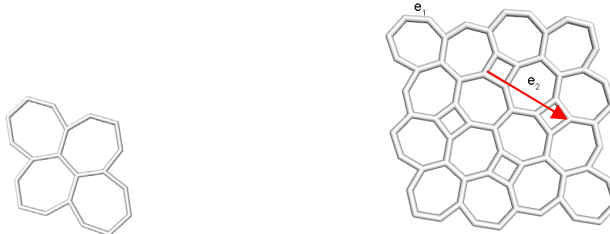
Figure 2. The ops strips of length s=1 and s=2 in the 2-dimensional co-net (the repeat unit (left) and 2×2 units (right) of ((4,7)3) pattern

Table 1. Omega and Sadhana polynomials in the ((4,7)3) modified graphene

Structure	Formulas
Net	$\Omega(x) = 2[k^2+(k-1)^2]\cdot x^2 + (10k^2+14k-4)\cdot x$
	$\Omega'(1) = 18x^2+6k = 6k(3k+1) = e(G)$
	$CI(G) = 2(162k^4+108k^3+5k^2+k-2)$
	$Sd(x) = 2[k^2+(k-1)^2]\cdot x^{18k^2+6k-2} + (10k^2+14k-4)\cdot x^{18k^2+6k-1}$
	$Sd'(1) = 6k(3k+1)(14k^2+10k-3) = e\cdot(14k^2+10k-3)$
	$Sd(G) = Sd'(1) = 252x^4+264x^3+6k^2-18k$
	$v(G) = 4k(5+3(k-1))$
Co-Net	$\Omega(x) = 4k(k-1)\cdot x^2 + (10k^2+14k-1)\cdot x$
	$\Omega'(1) = 18x^2+6k-1 = e(G)$
	$CI(G) = 2(162k^4+108k^3-13k^2-5k+1)$
	$Sd(x) = 4k(k-1)\cdot x^{18k^2+6k-3} + (10k^2+14k-1)\cdot x^{18k^2+6k-2}$
	$Sd'(1) = 2(7k^2+5k-1)(18k^2+6k-1) = 2(7k^2+5k-1)\cdot e$
	$Sd(G) = Sd'(1) = 252x^4+264x^3+10k^2-22k+2$
	$v(G) = 4k(5+3(k-1))$

Table 2. Examples for the formulas in Table 1.

k	Omega polynomial	v(G)	e(G)	CI(G)	Sd(G)
	Net				
2	$64X+10X^2$	64	84	6952	6132
3	$128X+26X^2$	132	180	32168	27540

k	Omega polynomial	$v(G)$	$e(G)$	$Cl(G)$	$Sd(G)$
4	$212X+50X^2$	224	312	96932	81432
5	$316X+82X^2$	340	480	229756	190560
6	$440X+122X^2$	480	684	466928	383724
7	$584X+170X^2$	644	924	852512	695772
	Co-Net				
2	$67X+8X^2$	64	83	6790	6142
3	$131X+24X^2$	132	179	31814	27566
4	$215X+48X^2$	224	311	96314	81482
5	$319X+80X^2$	340	479	228802	190642
6	$443X+120X^2$	480	683	465566	383846
7	$587X+168X^2$	644	923	850670	695942

CONCLUSIONS

Omega and Sadhana polynomials are useful theoretical tools in describing polygonal structures, such as the modified graphene of ((4,7)3) pattern. This modification can be acheved by using sequences of map operations.

Formulas to calculate the above polynomials and derived indices in an infinite ((4,7)3) lattice were given, along with some examples.

REFERENCES

1. M.V. Diudea, *Carpath. J. Math*., **2006,** *22*, 43.
2. P.E. John, A.E. Vizitiu, S. Cigher and M.V. Diudea, *MATCH Commun. Math. Comput. Chem*., **2007**, *57(2)*, 479.
3. A.R. Ashrafi, M. Jalali, M. Ghorbani and M.V. Diudea, *MATCH, Commun. Math. Comput. Chem*., **2008**, *60*, 905.
4. S. Klavžar, *MATCH Commun. Math. Comput. Chem*., **2008**, *59*, 217.
5. D.Ž. Djoković, *J. Combin. Theory Ser. B*, **1973**, *14*, 263.
6. P.M. Winkler, *Discrete Appl. Math*.,**1984**, *8*, 209.
7. P.V. Khadikar, V.K. Agrawal and S. Karmarkar, *Bioorg. Med. Chem*., **2002**, *10*, 3499.
8. P.V. Khadikar, S. Joshi, A.V. Bajaj and D. Mandloi, *Bioorg. Med. Chem. Lett*., **2004**, *14*, 1187.
9. A.R. Ashrafi, M. Ghorbani and M. Jalali, *Ind. J. Chem*., **2008**, *47A*, 535.
10. K.S. Novoselov and A.K. Geim, *Nat. Mater.,* **2007**, *6,* 183–191*.*
11. M.V. Diudea and A. Ilić, *Studia Univ. Babes-Bolyai Chemia*, 2009, *54(4),* 171.
12. M. Saheli, M. Neamati, K. Nagy and M.V. Diudea, *Studia Univ. Babes-Bolyai Chemia,* **2010**, *55 (1)*, 83.
13. M.V. Diudea, *Studia Univ. Babes-Bolyai Chemia*, **2003**, *48 (2)*, 3.
14. M.V. Diudea, M. Ştefu, P.E. John, and A. Graovac, *Croat. Chem. Acta,* **2006**, *79*, 355.
15. M.V. Diudea, *J. Chem. Inf. Model.,* **2005**, *45,* 1002.
16. M. Ştefu, M.V. Diudea and P.E. John, *Studia Univ. Babes-Bolyai Chemia*, **2005**, *50(2),* 165.
17. L. Schläfli, *Theorie der Vielfachen Kontinuität*, Denkschriften der Schweizerischen Naturforschenden Gesellschaft 38 (**1901**) (ed. J.H. Graf), 1{237 Gesammelte Mathematische Abhandlungen, Band 1, 167{387, Verlag BirkhÄauser, Basel **1950**.

OMEGA POLYNOMIAL IN TITANIUM OXIDE NANOTUBES

M. GHORBANI[a*], M.A. HOSSEINZADEH[b], M.V. DIUDEA[c]

ABSTRACT. A new counting polynomial, called Omega $\Omega(G,x)$, was recently proposed by Diudea. It is defined on the ground of "opposite edge strips" *ops*. Two related polynomials: Sadhana $Sd(G,x)$ and Theta $\Theta(G,x)$ polynomials can also be calculated by ops counting. Close formulas for calculating these three polynomials in infinite nano-structures resulted by embedding the titanium dioxide pattern in plane, cylinder and torus are derived. For the design of titanium dioxide pattern, a procedure based on a sequence of map operations is proposed.

Keywords: *Titanium oxide, Omega polynomial, Sadhana polynomial, Theta polynomial*

INTRODUCTION

Nano-era is a suitable name for the period started with the discovery of C_{60} fullerene and carbon nanotubes [1-3]. It opened a new gate for the science and technology at nanometer scale with wide implications in the human activities. After the discovery of carbon nanotubes, the question about the possible existence of nanotubular forms of other elements was addressed by scientists and they tried to obtain inorganic nanostructures [4-6]. Various oxides, sulfides, selenides, borates, silicates, etc of many metals show very ordered structures at the nano-scale. Many of these compounds form nanotubes, similar to those of carbon: MX2, M=Mo, W, Ta, In, Zn, Ti, Cd, X=O, S, Se, Te, CB_x, BN, etc. In the last years, oxides and other above mentioned inorganic substances found applications in the design of nanostructured functional materials as films, nanorods, porous systems, nanoclusters and nanocrystallites or as nanofibers [7-13].

Among these nanostructures, the titanium nanotubular materials, called "titania" by a generic name, are of high interest due to their chemical inertness, endurance, strong oxidizing power, large surface area, high photocatalytic activity, non-toxicity and low production cost. The applications of TiO_2 nanotubes include photocatalysis, solar cells systems, nanoscale materials for lithium-ion batteries, etc. The titanium oxide nanotubes were synthesized using various methods and precursors [14-20], carbon nanotubes, porous alumina or polymer

---

[a] *Institute of Nanoscience and Nanotechnology, University of Kashan, Kashan 87317−51167, I. R. Iran*
[b] *Department of Mathematical Science, Sharif University of Technology, Tehran, 11365-9415, I. R. Iran*
[c] *Faculty of Chemistry and Chemical Engineering, "Babes-Bolyai" University, 400028 Cluj, Romania*

membranes as templates [21-27], anodic oxidation of Ti [28-30], sol–gel technique [31-35] or sonochemical synthesis [36]. Models of possible growth mechanisms of titanium nanotubes, the atomic structure of the nanotube walls and their stacking mode are discussed [19,20,35]. TiO$_2$ nanotubes are semiconductors with a wide band gap and their stability increases with increasing of their diameters. The numerous studies on the production and technological applications of nanotubular titania also require theoretical studies on stability and other properties, the topological ones included [37-42].

DESIGN OF TITANIUM OXIDE LATTICE

A map *M* is a combinatorial representation of a (closed) surface. Several transformations or operations on maps are known and used for various purposes. We limit here to describe only those operations needed here to build the TiO$_2$ pattern. For other operations, the reader is invited to consult refs [43-48].

Medial Med is achieved by putting new vertices in the middle of the original edges. Join two vertices if the edges span an angle (and are consecutive within a rotation path around their common vertex in *M*). Medial is a 4-valent graph and *Med*(*M*) = *Med*(*Du*(*M*)).

Dualization of a map starts by locating a point in the center of each face. Next, two such points are joined if their corresponding faces share a common edge. It is the (Poincaré) *dual Du*(*M*). The vertices of *Du*(*M*) represent faces in *M* and *vice-versa.*

Figure 1 illustrates the sequence of map operations leading to the TiO$_2$ pattern: *Du*(*Med*(6,6)), the polyhex pattern being represented in Schläfli's symbols. Correspondingly, the TiO$_2$ pattern will be denoted as: (4(3,6)), squares of a bipartite lattice of 3 and 6 connected atoms, while the medial pattern: ((3,6)4). Clearly, the TiO$_2$ pattern can be done simply by putting a point in the centre of hexagons of the (6,6) pattern and join it alternately with the points on the contour. It is noteworthy that our sequence of operations is general, enabling transformation of the (6,6) pattern embedded on any surface and more over, it provides a rational procedure for related patterns, to be exploited in cage/cluster building.
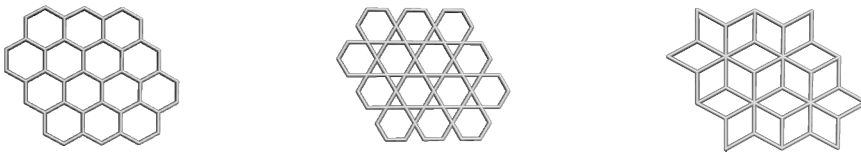
Figure 1. Way to TiO$_2$ lattice: (left) polyhex (6,6) pattern; (central) Med(6,6); (right) Du(Med(6,6))

OMEGA AND RELATED POLYNOMIALS

Let *G*(*V*,*E*) be a connected graph, with the vertex set *V*(*G*) and edge set *E*(*G*). Two edges *e* = *uv* and *f* = *xy* of *G* are called *codistant e co f* if they obey the following relation [49,50]:

$$d(v,x) = d(v,y)+1 = d(u,x)+1 = d(u,y) \tag{1}$$

Relation *co* is reflexive, that is, *e co e* holds for any edge *e* of *G*; it is also symmetric, if *e co f* then *f co e*. In general, relation *co* is not transitive, an example showing this fact is the complete bipartite graph $K_{2,n}$. If "*co*" is also transitive, thus an equivalence relation, then *G* is called a *co-graph* and the set of edges $C(e) := \{f \in E(G); f \; co \; e\}$ is called an *orthogonal cut oc* of *G*, *E(G)* being the union of disjoint orthogonal cuts: $E(G) = C_1 \cup C_2 \cup ... \cup C_k, \; C_i \cap C_j = \varnothing, i \neq j$. Klavžar [51] has shown that relation *co* is a theta Djoković-Winkler relation [52,53].

Let *e = uv* and *f = xy* be two edges of *G* which are *opposite* or *topologically parallel* and denote this relation by *e op f*. A set of opposite edges, within the same face/ring, eventually forming a strip of adjacent faces/rings, is called an *opposite* edge *strip ops*, which is a quasi-ortogonal cut *qoc* (*i.e.*, the transitivity relation is not necessarily obeyed). Note that *co* relation is defined in the whole graph while *op* is defined only in a face/ring. The length of *ops* is maximal irrespective of the starting edge.

Let *m(G,s)* be the number of *ops* strips of length *s*. The Omega polynomial is defined as [54]:

$$\Omega(G,x) = \sum_s m(G,s) \cdot x^s \tag{2}$$

The first derivative (in *x*=1) equals the number of edges in the graph

$$\Omega'(G,1) = \sum_s m(G,s) \cdot s = e = |E(G)| \tag{3}$$

A topological index, called Cluj-Ilmenau,[55] *CI=CI(G)*, was defined on Omega polynomial

$$CI(G) = \left\{ [\Omega'(G,1)]^2 - [\Omega'(G,1) + \Omega''(G,1)] \right\} \tag{4}$$

An example is given in Figure 2, which illustrates just the pattern of TiO$_2$ lattice.
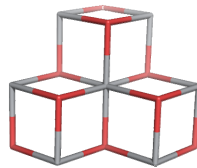
Figure 2. TiO$_2$ pattern; counting polynomial examples:

$$\Omega(G,x) = 3x^3 + 3x^5; \; \Omega'(G,1) = 24 = e(G); \; CI(G) = 474;$$

$$Sd(G,x) = 3x^{19} + 3x^{21}; \; Sd'(G,1) = 120 = Sd(G);$$

$$\Theta(G,x) = 9x^3 + 15x^5; \Theta'(G,1) = 27 + 75 = 102 = \Theta(G)$$

The Sadhana index *Sd(G)* was defined by Khadikar *et al.* [56,57] as

$$Sd(G) = \sum_s m(G,s)(|E(G)|-s) \tag{5}$$

where *m(G,s)* is the number of strips of length *s*. The Sadhana polynomial *Sd(G,x)* was defined by Ashrafi *et al.* [58] as

$$Sd(G,x) = \sum_s m(G,s) \cdot x^{|E(G)|-s} \qquad (6)$$

Clearly, the Sadhana polynomial can be derived from the definition of Omega polynomial by replacing the exponent s by $|E(G)-s|$. Then the Sadhana index will be the first derivative of $Sd(G, x)$ evaluated at $x=1$.

A third related polynomial is the Theta polynomial [59], defined in co-graphs as

$$\Theta(G,x) = \sum_s s \times m(G,s) \cdot x^s \qquad (7)$$

The aim of this study is to compute the Omega and its related counting polynomials in TiO_2 lattice, embedded in the plane but also in the cylinder and torus.

RESULTS AND DISCUSSION

We begin with the 2-dimensional graph, named K, (Figure 3). The various types of ops are drawn by arrows.
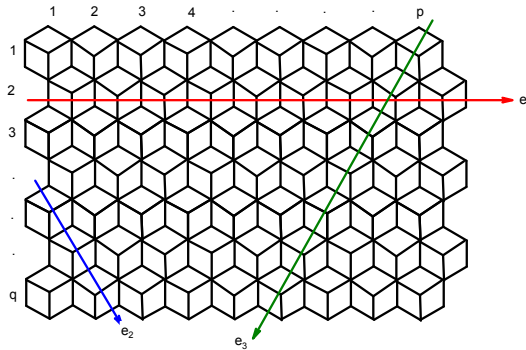
Figure 3. The ops strips of a 2-dimensional graph K of *Du*(*Med*(6,6)) TiO2 pattern.

By definition of Omega polynomial and Table 1 one can see that:

Table 1. The number of *ops* e_i, 1≤i≤6 in the graph K.

No.	Number of *ops*			Type of *ops*
q	2p+1			e_1
2 ⋮ 2	$\begin{cases} 3 \\ 5 \\ \vdots \\ 2\min\{2p,q\}-1 \end{cases}$			e_2
$\begin{cases} 2p-q+1 \\ q-2p+1 \end{cases}$	$\begin{cases} 2q+1 \\ 4p+1 \end{cases}$	$2p \ge q \ge 1$ $q \ge 2p.$		e_3

Now, we can derive the following formulas for the counting polynomials in the infinite 2-dimensional graph K:

204

$$\Omega(K,x) = \begin{cases} qx^{2p+1} + 2(x^3 + x^5 + ... + x^{2q-1}) + (2p-q+1)x^{2q+1} & 2p > q \geq 1 \\ qx^{2p+1} + 2(x^3 + x^5 + ... + x^{4p-1}) + (q-2p+1)x^{4p+1} & q \geq 2p \end{cases} \qquad (8)$$

$$Sd(K,x) = \begin{cases} qx^{|E(K)|-2p-1} + 2(x^{|E(K)|-3} + x^{|E(K)|-5} + ... + x^{|E(K)|-2q+1}) \\ + (2p-q+1)x^{E(K)|-2q-1} \qquad\qquad 2p > q \geq 1 \\ qx^{|E(K)|-2p-1} + 2(x^{|E(K)|-3} + x^{|E(K)|-5} + ... + x^{|E(K)|-4p+1}) \\ + (q-2p+1)x^{|E(K)|-4p-1} \qquad\qquad q \geq 2p \end{cases} \qquad (9)$$

$$\theta(K,x) = \begin{cases} a(x) + 2(3x^3 + ... + (2q-1)x^{2q-1}) + (2p-q+1)(2q+1)x^{2q+1} & 2p > q \geq 1 \\ a(x) + 2(3x^3 + ... + (4p-1)x^{4p-1}) + (q-2p+1)(4p+1)x^{4p+1} & q \geq 2p \end{cases} \qquad (10)$$

in which $a(x) = q(2p+1)x^{2p+1}$. Examples are given in Appendix.

We now consider the tubular structure G (Figure 4). Again the different cases of ops are drawn. One can see that $|S(e_1)| = 2p$ and $|S(e_2)| = 2q+1$. On the other hand, there are $q(e_1)$ and $2p(e_2)$ similar edges. This leads to the formulas

$$\Omega(G,x) = q \cdot x^{2p} + 2p \cdot x^{2q+1} \qquad (11)$$

$$Sd(G,x) = q \cdot x^{|E(G)|-2p} + 2p \cdot x^{|E(G)|-2q-1} \qquad (12)$$

$$\theta(G,x) = 2pq \cdot x^{2p} + 2p(2q+1) \cdot x^{2q+1} \qquad (13)$$

Figure 5 illustrates the case of a torus, denoted by H; it shows that there are two types of ops and their number is: $|S(e_1)| = 2p$, $|S(e_2)|=2pq$. On the other hand, there are $2q$ similar edges for each of e_1, e_2, respectively. With the above considerations we have the following formulas:

$$\Omega(H,x) = qx^{2p} + 2x^{2pq} \qquad (14)$$

$$Sd(H,x) = qx^{|E(H)|-2p} + 2x^{|E(H)|-2pq} \qquad (15)$$

$$\theta(H,x) = 2pqx^{2p} + 4pqx^{2pq} \qquad (16)$$

CONCLUSIONS

Nano-structured titania can be described, in topological terms by the aid of counting polynomials, such as Omega, Sadhana and Theta polynomials.

Close formulas for calculating these three polynomials in infinite nano-structures resulted by embedding the titanium dioxide pattern in plane, cylinder and torus are derived. A procedure based on a sequence of map operations is proposed for the design of titanium dioxide pattern.
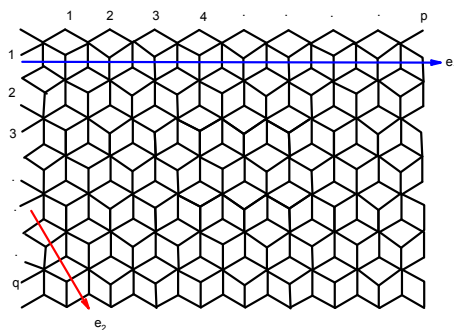
Figure 4. The ops strips of the nanotube G=TU[p,q].
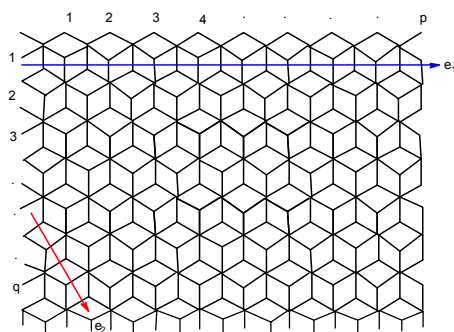
Figure 5. The ops strips of the nanotorus H=T[p,q].

REFERENCES

1. E. Osawa, Kagaku (Kyoto), **1970**, *25*, 854; C.A. 1971, *74*, 75698v.
2. H. Kroto, J.R. Heath, S.C. O'Brian, R.F. Curl, and R.E. Smalley, *Nature* (London), **1985**, *318*, 162.
3. W. Kraetschmer, L.D. Lamb, K. Fostiropoulos, and D.R. Huffman, *Nature* (London), **1990**, *347*, 354.
4. R. Tenne, *Chem. Eur. J.*, **2002**, *8*, 5296.
5. C.N.R. Rao, M. Nath, *Dalton Trans.,* **2003**, *1*, 1.
6. G.R. Patzke, F. Krumeich, R. Nesper, *Angew. Chem., Int. Ed.,* **2002**, *41*, 2447.
7. H. Imai, M. Matsuta, K. Shimizu, H. Hirashima, N. Negishi, *Sol. St. Ion.*, **2002**, *151*, 183.
8. M. Adachi, Y. Murata, I. Okada, S. Yoshikawa, *J. Electrochem. Soc.* **2003**, *150*, G488.
9. Y. Zhou, L. Cao, F. Zhang, B. He, H. Li, *J. Electrochem. Soc.,* **2003**, *150A*, 1246.
10. O.K. Varghese, D. Gong, M. Paulose, K.G. Ong, C.A. Grimes, *Sens. Actuators B*, **2003**, *93*, 338.

11. O.K. Varghese, D. Gong, M. Paulose, K.G. Ong, E.C. Dickey, and C.A. Grimes, *Adv. Mater.*, **2003**, *15*, 624.
12. C.A. Grimes, K.G. Ong, O.K. Varghese, X. Yang, G. Mor, M. Paulose, E.C. Dickey, C. Ruan, M.V. Pishko, J.W. Kendig, A.J. Mason, *Sensors*, **2003**, *3*, 69.
13. G.K. Mor, M.A. Carvalho, O.K. Varghese, M.V. Pishko, C.A. Grimes, *J. Mater. Res.*, **2004**, *19*, 628.
14. T. Kasuga, M. Hiramatsu, A. Hoson, T. Sekino, K. Niihara, *Adv. Mater.*, **1999**, *11*, 1307.
15. S. Zhang, J. Zhou, Z. Zhang, Z. Du, A.V. Vorontsov, Z. Jin, *Chin. Sci. Bull.,* **2000**, *45*, 1533.
16. G. H. Du, Q. Chen, R.C. Che, Z.Y. Yuan, L.-M. Peng, *Appl. Phys. Lett.,* **2001**, *79*, 3702.
17. D.-S. Seo, J.-K. Lee, H. Kim, *J. Cryst. Growth*, **2001**, *229*, 428.
18. C.-H. Lin, S.-H. Chien, J.-H. Chao, C.-Y. Sheu, Y.-C. Cheng, Y.-J. Huang, C.-H. Tsai, *Catal. Lett.,* **2002**, *80*, 153.
19. B.D. Yao, Y.F. Chan, X.Y. Zhang, W.F. Zhang, Z.Y. Yang, N. Wang, *Appl. Phys. Lett.,* **2003**, *82*, 281.
20. W. Wang, O.K. Varghese, M. Paulose, C.A. Grimes, *J. Mater. Res.,* **2004**, 19, 417.
21. J. Sun, L. Gao, Q. Zhang, *J. Mater. Sci. Lett.,* **2003**, *22*, 339.
22. P. Hoyer, *Langmuir,* **1996**, *12*, 1411
23. H. Imai, Y. Takei, K. Shimizu, M. Matsuda, H. Hirashima, *J. Mater. Chem.,* **1999**, *9*, 2971.
24. S.M. Liu, L.M. Gan, L.H. Liu, W.D. Zhang, H.C. Zeng, *Chem. Mater.,* **2002**, *14*, 1391.
25. Y.-L. Shi, X.-G. Zhang, H.-L. Li, *Mater. Sci. Engin.* A, **2002**, *333*, 239.
26. X.H. Li, W.M. Liu, H.L. Li, *Appl. Phys. A*, **2003**, *80*, 317.
27. T. Peng, H. Yang, G. Chang, K. Dai, K. Hirao, *Chem. Lett.,* **2004**, *33*, 336.
28. D. Gong, C.A. Grimes, O.K. Varghese, W. Hu, R.S. Singh, Z. Chen, E.C. Dickey, *J. Mater. Res.,* **2001**, *16*, 3331.
29. O.K. Varghese, D. Gong, M. Paulose, C. A. Grimes, E. C. Dickey, *J. Mater. Res.,* **2003**, *18*, 156.
30. G.K. Mor, O.K. Varghese, M. Paulose, N. Mukherjee, C.A. Grimes, *J. Mater. Res.*, **2003**, *18*, 2588.
31. B.B. Lakshmi, P.K. Dorhout, and C.R. Martin, *Chem. Mater.,* **1997**, *9*, 857.
32. T. Kasuga, M. Hiramatsu, A. Hoson, T. Sekino, K. Niihara, *Langmuir,* **1998**, *14*, 3160.
33. S. Kobayashi, K. Hanabusa, N. Hamasaki, M. Kimura, H. Shirai, *Chem. Mater.* **2000**, *12*, 1523.
34. M. Zhang, Y. Bando, K. Wada, *J. Mater. Sci. Lett.,* **2001**, *20*, 167.
35. Y.Q. Wang, G.Q. Hu, X.F. Duan, H.L. Sun, Q.K. Xue, *Chem. Phys. Lett.,* **2002**, *365*, 427.
36. Y. Zhu, H. Li, Y. Koltypin, Y.R. Hacohen, A. Gedanken, *Chem. Commun.,* **2003**, 2616.
37. V.V. Ivanovskaya, A.N. Enyashin, A.L. Ivanovskii, *Mendeleev Comm.,* **2003**, *13*, 5.
38. V.V. Ivanovskaya, A.N. Enyashin, A. L. Ivanovskii, *Russ. J. Inorg. Chem.,* **2004**, *49*, 1.
39. A.N. Enyashin, G. Seifert, *Phys. Stat. Sol.* **2005**, *242*, 7, 1361.
40. G.S. Zakharova, V.L. Volkov, V.V. Ivanovskaya, and A.L. Ivanovskii, *Russ. Chem. Rev.*, **2005**, *74*, 587.

41. A.E. Vizitiu, M.V. Diudea, *Studia Univ. Babes-Bolyai Chemia*, **2009**, *54*, 173.
42. M.V. Diudea and A. Ilić, *MATCH, Commun. Math. Comput. Chem.*, **2009** (submitted).
43. M.V. Diudea, P.E. John, A. Graovac, M. Primorac, and T. Pisanski, *Croat. Chem. Acta*, **2003**, *76*, 153.
44. M.V. Diudea, *Forma (Tokyo)*, **2004**, *19*, 131.
45. M.V. Diudea, M. Ştefu, P.E. John, and A. Graovac, *Croat. Chem. Acta*, **2006**, *79*, 355.
46. M. Ştefu, M.V. Diudea, and P.E. John, *Studia Univ. Babes-Bolyai Chemia*, **2005**, *50*, 165.
47. M.V. Diudea, *J. Chem. Inf. Model.,* **2005**, *45,* 1002.
48. M.V. Diudea and Cs.L. Nagy, *Periodic Nanostructures*, Springer, **2007**.
49. M.V. Diudea and A. Ilić, *Carpath. J. Math.*, **2009** (submitted).
50. A.R. Ashrafi, M. Jalali, M. Ghorbani and M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2008**, *60*, 905.
51. S. Klavžar, *MATCH Commun. Math. Comput. Chem.*, **2008**, *59*, 217.
52. D.Ž. Djoković, *J. Combin. Theory Ser. B*, **1973**, *14*, 263.
53. P.M. Winkler, *Discrete Appl. Math.*,**1984**, *8*, 209.
54. M.V. Diudea, *Carpath. J. Math.*, **2006**, *22*, 43.
55. P.E. John, A.E. Vizitiu, S. Cigher, and M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2007**, *57*, 479.
56. P.V. Khadikar, V.K. Agrawal and S. Karmarkar, *Bioorg. Med. Chem.*, **2002**, *10*, 3499.
57. P.V. Khadikar, S. Joshi, A.V. Bajaj and D. Mandloi, *Bioorg. Med.Chem. Lett.*, **2004**, *14*, 1187.
58. A.R. Ashrafi, M. Ghorbani and M. Jalali, *Ind. J. Chem.*, **2008**, *47A*, 535.
59. M.V. Diudea, S. Cigher and P.E. John, Omega and Related Counting Polynomials, *MATCH Commun. Math. Comput. Chem.*, **2008**, *60*, 237.

APPENDIX

Examples for calculating Omega polynomial.

1. Case of infinite 2-dimensional graph K.

We have the Omega polynomial: $qx^{2p+1} + 2(x^3 + x^5 + ... + x^{2\,q-1}) + (2p - q + 1)x^{2q+1}$

1.1. Case: $2p > q > p, 2\,|\,q,$
If q = 6, p = 5 then, the graph is:
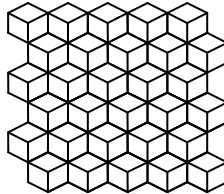
and
$\Omega(G,x) = 6x^{11} + 2(x^3 + x^5 + x^7 + x^9 + x^{11}) + 5x^{13}$

1.2. Case: $2p > q > p, 2 \nmid q,$
Now if p = 4, q = 7 then, the graph is:
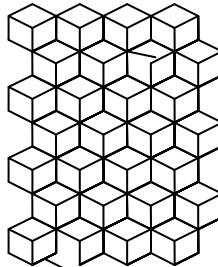
and
$\Omega(G,x) = 7x^9 + 2(x^3 + x^5 + x^7 + x^9 + x^{13}) + 2x^{15}$

We have also $qx^{2p+1} + 2(x^3 + x^5 + ... + x^{4p-1}) + (q - 2p + 1)x^{4p+1}$

1.3. Case: $q \geq 2p$.   If p = 4, q = 9 then, the graph is:
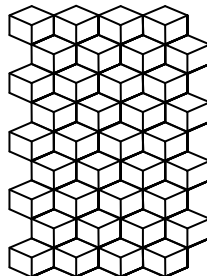
and
$\Omega(G,x) = 9x^9 + 2(x^3 + x^5 + x^7 + x^9 + x^{11} + x^{13} + x^{15}) + 2x^{17}$

209

2. Case of nanotubes G [p,q].

We have the Omega polynomial: $\Omega(G,x) = qx^{2p} + 2px^{2q+1}$
Now, if p = 5, q = 4 then, the graph is:
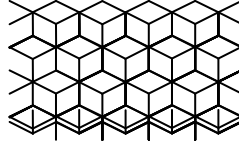
and

$$\Omega(G,x) = 4x^{10} + 10x^9$$

Or, if p = 6, q = 6 then, the graph is :
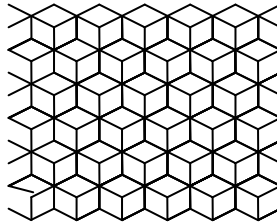
and

$$\Omega(G,x) = 6x^{12} + 12x^{13}$$

3. Case of nanotori H [p,q].

We have the Omega polynomial: $\Omega(H,x) = qx^{2p} + 2x^{2pq}$
Now, if p = 4, q = 5 then, the graph is:
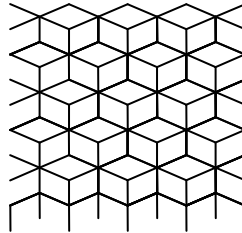
and
$$\Omega(H,x) = 5x^8 + 2x^{40}$$

OMEGA POLYNOMIAL IN P-TYPE SURFACE NETWORKS

MONICA STEFU[*], VIRGINIA BUCILA[*], M. V. DIUDEA[*]

ABSTRACT. Design of two crystal-like networks was achieved by embedding a zig-zag Z-unit and its corresponding armchair A-unit, of octahedral symmetry, in the P-type surface, by means of the original software Nano Studio. The hypothetical networks, thus obtained, were characterized in their topology by Omega counting polynomial.

Keywords: crystal-like networks, Omega polynomials, topology

INTRODUCTION

In the last two decades, novel carbon allotropes have been discovered and studied for applications in nano-technology. Among the carbon structures, fullerenes (zero-dimensional), nanotubes (one dimensional), graphene (two dimensional) and spongy carbon (three dimensional) were the most challenging [1,2]. Inorganic clusters, like zeolites, also attracted the attention of scientists. Recent articles in crystallography promoted the idea of topological description and classification of crystal structures [3-8].

The present study deals with two hypothetical crystal-like nano-carbon structures, of which topology is described in terms of Omega counting polynomial.

BACKGROUND ON OMEGA POLYNOMIAL

In a connected graph $G(V,E)$, with the vertex set $V(G)$ and edge set $E(G)$, two edges $e = uv$ and $f = xy$ of G are called *codistant e co f* if they obey the relation [9]:

$$d(v,x) = d(v,y)+1 = d(u,x)+1 = d(u,y) \qquad (1)$$

which is reflexive, that is, *e co e* holds for any edge *e* of *G*, and symmetric, if *e co f* then *f co e*. In general, relation *co* is not transitive; if "*co*" is also transitive, thus it is an equivalence relation, then *G* is called a *co-graph* and the set of edges $C(e) := \{f \in E(G); f\ co\ e\}$ is called an *orthogonal cut oc* of *G*, $E(G)$ being the union of disjoint orthogonal cuts: $E(G) = C_1 \cup C_2 \cup ... \cup C_k,\ C_i \cap C_j = \varnothing, i \neq j$.

Klavžar [10] has shown that relation *co* is a *theta* Djoković-Winkler relation [11,12].

* *Faculty of Chemistry and Chemical Engineering, "Babes-Bolyai" University, Arany Janos Str. 11, 400084, Cluj, Romania*

We say that edges *e* and *f* of a plane graph *G* are in relation *opposite, e op f,* if they are opposite edges of an inner face of *G*. Note that the relation *co* is defined in the whole graph while *op* is defined only in faces. Using the relation *op* we can partition the edge set of *G* into *opposite* edge *strips, ops.* An *ops* is a quasi-orthogonal cut *qoc*, since *ops* is not transitive.

Let *G* be a connected graph and $s_1, s_2, ..., s_k$ be the *ops* strips of *G*. Then the *ops* strips form a partition of *E(G)*. The length of *ops* is taken as maximum. It depends on the size of the maximum fold face/ring F_{max}/R_{max} considered, so that any result on Omega polynomial will have this specification.

Denote by *m(G,s)* the number of *ops* strips of length *s* and define the Omega polynomial as [13-15]:

$$\Omega(G, x) = \sum_s m(G, s) \cdot x^s \tag{2}$$

Its first derivative (in *x*=1) equals the number of edges in the graph:

$$\Omega'(G, 1) = \sum_s m(G, s) \cdot s = e = |E(G)| \tag{3}$$

On Omega polynomial, the Cluj-Ilmenau index [9], *CI=CI(G)*, was defined:

$$CI(G) = \left\{ [\Omega'(G, 1)]^2 - [\Omega'(G, 1) + \Omega''(G, 1)] \right\} \tag{4}$$

The Omega polynomial partitions the edge set of the molecular graph into opposite edge strips, by the length of the strips.

OMEGA POLYNOMIAL IN TWO P-SURFACE CRYSTAL NETWORKS

Design of two crystal-like networks was achieved by identifying the opposite open faces of a zig-zag Z-unit and its corresponding armchair A-unit (Figure 1), of octahedral symmetry and embedding them in the P-type surface, with the help of original software Nano Studio [16] .

Omega polynomials for the repeat units of the Z_P and A_P structures (Figure 2) herein discussed are listed in Table 1. The polynomials are calculated at R_{max}[8] as follows.

In the Z_P structure, the term at exponent 1 counts the edges in odd faces/rings that are not counted in even rings. The exponent 2 refers to isolated even rings while the exponent 4 represents strips of three even-membered faces/rings.

In the A_P structure, there are no odd faces so the polynomial has no terms at exponent 1. The exponent 6 represents strips of five even-membered faces/rings.

Figure 1. Units of the Z_P (left) and A_P (right) crystal-like structures

The polynomials are calculated on a cubic lattice of dimension (k,k,k), at R_{max}[8]; following similar considerations and analyzing the calculations made by our original Nano Studio [16] software, we derived the formulas, listed in Table2, and provided examples for some k-values, as well.

Figure 2. The Z_P (left) and A_P (right) crystal-like structures

Table 1. Topological data for the units of Z_P and A_P structures

Octahedral structure	Vertices	Edges	Faces f_8	Open Faces	Omega Polynomial R_{max}[8]	CI
Z_P	120	168	12	6	$48X^1 + 36X^2 + 12X^4$	27840
A_P	144	192	12	6	$12X^2 + 24X^4 + 12X^6$	36000

In the Z_P and A_P network structures, the term at exponent 8 represent the number of edge strips of length 8; these strips cross only f_8 when link 4 Z_P units, and cross faces f_8 and f_6 when link 4 A_P units respectively, so it is present starting with k=2. In case of Z_P net, the term at exponent 8 counts the large hollows, ordered as in zeolites, natural alumino-silicates, used as molecular sieves or in chemical catalysis.

Table 2. Omega polynomial in Z_P and A_P networks

Formulas for Z_P network
$\Omega(X, k, R_{max}[8]) = 48k^2 X^1 + 12k(4k^2 - 2k + 1)X^2 + 3k(5k^2 + 3k - 4)X^4 + 3k(k-1)^2 X^8$
$\Omega'(1) = 48k^2 + 2 \cdot 12k(4k^2 - 2k + 1) + 4 \cdot 3k(5K^2 + 3k - 4) + 8 \cdot 3k(k-1)^2 = 12k^2(15k - 1)$
$\Omega''(1) = 12k(37k^2 - 23k + 4)$
$CI(k) = 48k(675k^5 - 90k^4 + 3k^3 - 13k^2 + 6k - 1)$
Formulas for A_P network
$\Omega(X, k, R_{max}[8]) = 12kX^2 + 12k(k+1)X^4 + 3k(k-1)^2 X^8 + 12k^2 X^{5k+1} + 24k \sum_{i=2}^{k} X^{4(2i-1)}$
$\Omega'(1) = 12k^2(15k + 1)$
$\Omega''(1) = 4k(203k^3 + 33k^2 - 80k + 12)$
$CI(k) = 4k(8100k^5 + 1080k^4 - 167k^3 - 78k^2 + 77k - 12)$

k	Omega polynomial: examples	CI
	R_{max}[8], Z_P network	
1	$48X + 36X^2 + 12X^4$	27840
2	$192X + 312X^2 + 132X^4 + 6X^8$	1933728
3	$432X + 1116X^2 + 450X^4 + 36X^8$	22567104
4	$768X + 2736X^2 + 1056X^4 + 108X^8$	128288064
5	$1200X + 5460X^2 + 2040X^4 + 240X^8$	492768960
6	$1728X + 9576X^2 + 3492X^4 + 450X^8$	1478124000
	R_{max}[8] A_P network	

Table 2-continuation

1	$12X^2+24X^4+12X^6$	36000
2	$24X^2+72X^4+6X^8+48X^{11}+48X^{12}$	2199792
3	$36X^2+144X^4+36X^8+72X^{12}+108X^{16}+72X^{20}$	24609456
4	$48X^2+240X^4+108X^8+96X^{12}+96X^{20}+192X^{21}+96X^{28}$	136947840
5	$60X^2+360X^4+240X^8+120X^{12}+120X^{20}+300X^{26}+120X^{28}+120X^{36}$	519300960
6	$72X^2+504X^4+450X^8+144X^{12}+144X^{20}+144X^{28}+432X^{31}+144X^{36}+144X^{44}$	1544324400
7	$84X^2+672X^4+756X^8+168X^{12}+168X^{20}+168X^{28}+756X^{36}+168X^{44}+168X^{52}$	3882737712

Formulas for the number of atoms in the two networks are given in Table 3.

Table 3. Number of atoms $v = |V(G)|$

Z_P network structures $v_k = 120 \cdot k^3$	A_P network structures $v_k = 144 \cdot k^3 - 24 \cdot k^2(k-1)$				
k	1	2	3	4	5	6
v for Z_P	120	960	3240	7680	15000	25920
v for A_P	144	1056	3456	8064	15600	26784

ACKNOWLEDGMENTS

REFERENCES

1. M.V. Diudea, *Ed.*, "Nanostructures, novel architecture", NOVA, **2005**.
2. M.V. Diudea and Cs. L. Nagy, "Periodic Nanostructures", Springer, **2007**.
3. L. Carlucci, G. Ciani, and D. Proserpio, *Coord. Chem. Rev.*, **2003**, *246*, 247.
4. L. Carlucci, G. Ciani, and D. Proserpio, *Cryst. Eng. Comm.*, **2003**, *5*, 269.
5. V.A. Blatov, L. Carlucci, G. Ciani, and D. Proserpio, *Cryst. Eng. Comm.*, **2004**, *6*, 377.
6. I.A. Baburin, V.A. Blatov, L. Carlucci, G. Ciani, and D. Proserpio, *J. Solid State Chem.*, **2005**, *178*, 2452.
7. O. Delgado-Friedrichs and M. O'Keeffe, *J. Solid State Chem.*, **2005**, *178*, 2480.
8. V.A. Blatov, O. Delgado-Friedrichs, M. O'Keeffe, and D. Proserpio, *Acta Cryst.*, **2007**, *A63*, 418.
9. P.E. John, A.E. Vizitiu, S. Cigher, and M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2007**, *57*, 479.
10. S. Klavžar, *MATCH Commun. Math. Comput. Chem.*, **2008**, *59*, 217.
11. D.Ž. Djoković, *J. Combin. Theory Ser. B*, **1973**, *14*, 263.
12. P.M. Winkler, *Discrete Appl. Math.*, **1984**, *8*, 209.
13. M.V. Diudea, *Carpath. J. Math.*, **2006**, *22*, 43.
14. M.V. Diudea, S. Cigher, P.E. John, *MATCH Commun. Math. Comput. Chem.*, **2008**, *60*, 237.
15. M.V. Diudea, S. Cigher, A.E. Vizitiu, M.S. Florescu, and P.E. John, *J. Math. Chem.*, **2009**, *45*, 316.
16. Cs.L. Nagy, M.V. Diudea, Nano Studio software, *Babes-Bolyai University*, **2009**.

OMEGA POLYNOMIAL IN OCT-P$_4$TRS NETWORK

MAHBOUBEH SAHELI[a], RALUCA O. POP[b], MONICA L. POP[c], MIRCEA V. DIUDEA[c*]

ABSTRACT. Design of a hypothetical crystal network, by using $Trs(P_4(M))$ sequence of map operations, is presented. It is shown that the octahedral monomer is the most stable, among the similar structures designed from the Platonic solids, as hydrogenated species, and all these have a moderate stability, between adamantane and C_{60} fullerene, as calculated at the PM3 level of theory. The topology of the network is described in terms of Omega polynomial, function of the net parameters. Close formulas for this polynomial and examples are tabulated.

Keywords: Omega polynomial; crystal-like network.

INTRODUCTION

In the last two decades, several new carbon allotropes have been discovered and studied for applications in nano-technology. Among the carbon structures, fullerenes (zero-dimensional), nanotubes (one dimensional), graphene (two dimensional) and spongy nanostructures (three dimensional) were the most studied [1,2]. Inorganic compounds also attracted the attention of scientists. Recent articles in crystallography promoted the idea of topological description and classification of crystal structures [3-7].

The present study deals with a hypothetical crystal-like nano-carbon structure, designed by a sequence of map operations [8-11], of which topology is described in terms of Omega polynomial.

OMEGA POLYNOMIAL

Let $G(V,E)$ be a graph, with $V(G)$ and $E(G)$ being the sets of vertices/ atoms and edges/bonds, respectively. Two edges e and f of a plane graph G are in relation *opposite, e op f,* if they are opposite edges of an inner face of G. Relation *op* will partition the edges set of G into *opposite* edge *strips ops*, as follows. (i) Any two subsequent edges of an *ops* are in *op* relation; (ii) Any three subsequent edges of such a strip belong to adjacent faces; (iii) In a plane graph, the inner dual of an *ops* is a path, an open or a closed one (however,

---
[a] *Departmant of Mathematics, Faculty of Science, University of Kashan, Kashan 87317-51167, I.R. Iran, mmsaheli@yahoo.com*
[b] *National Institute of R&D for Electrochemistry and Condensed Matter, 300569 Timisoara, Romania, ralucapop24@gmail.com*
[c] *Faculty of Chemistry and Chemical Engineering, "Babes-Bolyai" University, 400028 Cluj, Romania, mpop@chem.ubbcluj.ro; diudea@gmail.com*

in 3D networks, the ring/face interchanging will provide *ops* which are no more paths); (iv) The *ops* is taken as maximum possible, irrespective of the starting edge [12-18].

The Ω-polynomial [12] is defined on the ground of opposite edge strips *ops* $S(G)=S_1, S_2, ..., S_k$ in the graph. Denoting by *m* the number of *ops* of length *s*=|*S*|, then we can write

$$\Omega(x) = \sum_s m \cdot x^s \qquad (1)$$

The first derivative (in *x*=1) can be taken as a graph invariant or a topological index; in this case, it equals the number of edges in the graph.

$$\Omega'(1) = \sum_s m \cdot s = e = |E(G)| \qquad (2)$$

On Omega polynomial, the Cluj-Ilmenau index [13], *CI*=*CI*(*G*), was defined:

$$CI(G) = \left\{ [\Omega'(G,1)]^2 - [\Omega'(G,1) + \Omega''(G,1)] \right\} \qquad (3)$$

The first derivative (in *x*=1) can be taken as a graph invariant or a topological index:

$$\Omega'(1) = \sum_s m \cdot s = |E(G)| \qquad (4)$$

In tree graphs, the Omega polynomial simply counts the non-opposite edges, being included in the term of exponent *c*=1.

LATTICE BUILDING AND MONOMER STABILITY

The lattice was constructed by using the unit designed with the net operation sequence *Trs*(P_4(M)), where M=Oct (Octahedron). More about map/net operations, the reader can find in refs. [8-11].

The net (Figure 1) was built up by identifying the identical (quadrilateral) faces of the unit structure The crystal-like structure shows oriented hollows, as those encountered in zeolites, natural alumino-silicates widely used in synthetic chemistry as catalysts.

The unit involved in these constructions, namely *Trs*(P_4(M)), M=Oct, as a hydrogenated structure, shows moderate stability as given by their heat of formation HF, total energy TE and HOMO-LUMO Gap HLGAP, calculated at the PM3 level of theory (Table 1).

Figure 1: Network *Trs*(P_4(M)); [2,2,2]; M=Octahedron, in two different views.

For example, the total energy per heavy atoms of the structures in Table 1 are between the values of adamantane (-3305.19 kcal/mol), which is the most related small structure (see Figure 2, left, in red) and C_{60} (-2722.45

kcal/mol), the standard molecule in nanostructures. The same is true about the HOMO-LUMO gap. Calculations by using a density functional-based tight binding method combined with the self-consistent charge technique (SCC-DFTB) on hydrogenated units of diamond and a diamond-like network[19] have shown the same ordering of stability as given by PM3 approach; thus, our results reported here can be considered as pertinent ones.

Table 1. Quantum Chemistry PM3 data for some units designed by $Trs(P_4(M))$: Heat of Formation HF, Total energy TE and HOMO-LUMO Gap HLGAP

M	N-heavy atoms	HF (kcal/mol)	HF/N heavy	TE (kcal/mol)	TE/N heavy	HLGAP (eV)	Sym.
Ico	110	1216.81	11.06	-328026	-2982.05	11.79	I_h
Oct	44	448.67	10.19	-131248	-2982.92	12.17	o_h
T	22	308.48	14.022	-65540	-2979.09	11.99	T_d

OMEGA POLYNOMIAL IN $Trs(P_4(M))$ Network

The Omega polynomial (calculated at $R_{max}[4]$) for the investigated network is as follows:

$$\Omega(G,x) = 24\,a(a+1)x + 12(a(a+1)+(a-2))x^2 + 24(a(a-2)+1)x^3 \qquad (5)$$
$$+ 4(a-1)^3 x^6 + 3(a-1)x^{(2a)^2}$$

$$\Omega'(G,1) = |E(G)| = 36a^2(a+1) \qquad (6)$$

$$CI(G) = 1296\,a^6 + 2544\,a^5 + 1344\,a^4 - 144\,a^3 + 144\,a^2 - 120\,a + 24 \qquad (7)$$

The above formulas can be verified with the examples listed in Table 2. Calculations were performed by our Nano Studio[20] software program.
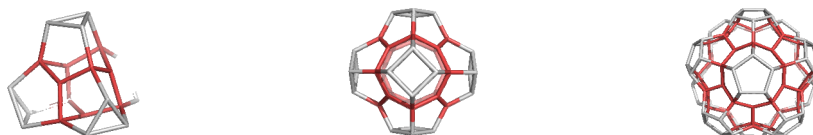
Figure 2. Platonic structures transformed by $Trs(P_4(M))$ sequence of map operations: M=Tetrahedron T (left); M=Octahedron Oct (central) and M=Icosahedron Ico (right). The red color is only to show the related substructures**.**

Table 2. Examples of Omega polynomial and CI calculation

a	Omega polynomial	CI
1	$48x + 12x^2$	5088
2	$144x + 72x^2 + 24x^3 + 4x^6 + 3x^{16}$	185064
3	$288x + 156x^2 + 96x^3 + 32x^6 + 6x^{36}$	1668912
4	$480x + 264x^2 + 216x^3 + 108x^6 + 9x^{64}$	8250168
5	$720x + 396x^2 + 384x^3 + 256x^6 + 12x^{100}$	29025024
6	$1008x + 552x^2 + 600x^3 + 500x^6 + 15x^{144}$	81963528

217

CONCLUSIONS

A hypothetical crystal network was built up by using a repeat unit designed by $Trs(P_4(M))$ sequence of map operations. It was shown that the octahedral monomer (i.e., the repeat unit of this network) is the most stable (as hydrogenated species), among the similar structures derived from the Platonic solids, and all these have a moderate stability, between adamantane and C_{60} fullerene, as calculated at the PM3 level of theory. The topology of the network was described in terms of Omega polynomial, function of the net parameters. Close formulas for this polynomial and examples were tabulated. Omega polynomial description proved to be a simple and efficient method in topological characterization of new designed nano-structures.

ACKNOWLEDGMENTS

REFERENCES

1. M.V. Diudea, Ed., "*Nanostructures, novel architecture*", NOVA, **2005**.
2. M.V. Diudea and Cs.L. Nagy, "*Periodic Nanostructures*", Springer, **2007**.
3. L. Carlucci, G. Ciani and D. Proserpio, *Cryst. Eng. Comm.*, **2003**, *5*, 269.
4. V.A. Blatov, L. Carlucci, G. Ciani and D. Proserpio, *Cryst. Eng. Comm.*, **2004**, *6,* 377.
5. I.A. Baburin, V.A. Blatov, L. Carlucci, G. Ciani and D. Proserpio, *J. Solid State Chem.*, **2005**, *178*, 2452.
6. O. Delgado-Friedrichs and M. O'Keeffe, *J. Solid State Chem.*, **2005**, *178*, 2480.
7. V.A. Blatov, O. Delgado-Friedrichs, M. O'Keeffe, and D. Proserpio, *Acta Cryst.*, **2007**, *A63*, 418.
8. M.V. Diudea, M. Ştefu, P.E. John, and A. Graovac, *Croat. Chem. Acta,* **2006**, *79*, 355.
9. M.V. Diudea, *J. Chem. Inf. Model.*, **2005**, *45,* 1002.
10. M.V. Diudea, *Forma* (Tokyo), **2004**, *19 (3)*, 131.
11. M. Stefu, M.V. Diudea and P.E. John, *Studia Univ. Babes-Bolyai Chemia*, **2005**, *50*, 2, 165.
12. M.V. Diudea, *Carpath. J. Math.,* **2006**, *22*, 43.
13. P.E. John, A.E., Vizitiu, S. Cigher, M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2007**, *57*, 479.
14. M.V. Diudea, S. Cigher, P.E. John, *MATCH Commun. Math. Comput. Chem.*, **2008**, *60*, 237.
15. M.V. Diudea, S. Cigher, A.E. Vizitiu, M.S. Florescu, P.E. John, *J. Math. Chem.*, **2009**, *45*, 316.
16. M. Saheli, M. Neamati, K. Nagy and M.V. Diudea, *Studia Univ. Babes-Bolyai Chemia*, **2010**, *55 (1)*, 83.
17. M.V. Diudea, *Acta Chim. Slov.*, **2010**, 57, 551.
18. M.V. Diudea, S. Klavžar, *Acta Chim. Slov.*, **2010**, *57*, 565.
19. M.V. Diudea, A. Bende and D. Janežič, *Fullerenes, Nanotubes, Carbon Nanostruct.*, **2010**, *18 (3)*, 236.
20. Cs.L. Nagy, M.V. Diudea, *Nano Studio software*, Babes-Bolyai Univ., **2**009.

OMEGA AND SADHANA POLYNOMIALS IN P-TYPE SURFACE NETWORKS

FARZANEH GHOLAMI-NEZHAAD[a], MIRCEA V. DIUDEA[b*]

ABSTRACT. Design of a hypothetical carbon crystal lattice, embedded in the P-type surface, was performed by identifying two opposite open faces of a unit, of octahedral symmetry, by the aid of Nano Studio software. The topology of the net and its co-net, thus obtained, was characterized by Omega and Sadhana counting polynomials.

Keywords: Omega polynomial, Sadhana polynomial, P-type surface networks

INTRODUCTION

Among the carbon allotropes, discovered in the nano-era, fullerenes (zero-dimensional), nanotubes (one dimensional), graphene (two dimensional) and spongy carbon (three dimensional) were the most challenging [1,2]. Inorganic compounds including oxides, sulfides, selenides, borates, silicates, etc. of many metals, also found applications as nano-structured functional materials [3-12].

Zeolites are natural or synthetic alumino-silicates with an open three-dimensional crystal structure. Zeolites are micro-porous solids known as "molecular sieves." The term molecular sieve refers to the property of these materials to selectively sort molecules, based primarily on a size exclusion process. This is due to a regular structure of pores, of molecular dimensions, forming channels [13-17].

The rigorous and often aesthetically appealing architecture of crystal networks attracted the interest of scientists in a broad area, from crystallographers, to chemists and mathematicians.

The present study deals with a hypothetical carbon crystal-like nanostructure, of which topology is described in terms of Omega and Sadhana counting polynomial.

NETWORK DESIGN

The hypothetical carbon crystal network herein discussed was built up by identifying two opposite open faces of a unit (Figure 1, left), of octahedral symmetry, by the aid of Nano Studio software [18], also enabling their embedding in the P-type surface [1,2], belonging to the space group $P_n 3 m$.

---

[a] *Faculty of Mathematics, University of Kashan, I. R. Iran*
[b] *Faculty of Chemistry and Chemical Engineering, Babes-Bolyai University, 400028 Cluj, Romania,*
  diudea@gmail.com

As any net has its co-net, this was identified to the structure presented in Figure 1, right. Indeed, when constructing the two infinite networks (Figure 2), a perfect superposition (Figure 2, central) can be evidenced: in fact is one and the same infinite network, differences appearing only at the boundaries. Thus, the topological characterization will be done on cubic (k,k,k) domains, separately, for the net and its co-net (see below).

v=98; e=120; f_6=8; f_8=12   v=144; e=192; f_6=20; f_8=24

Figure 1. Units of the net (left) and co-net (right)
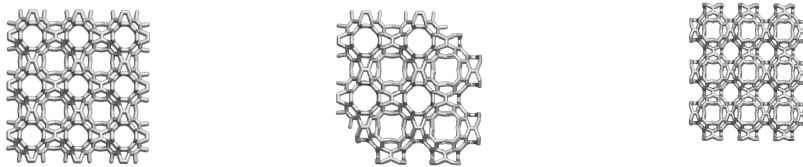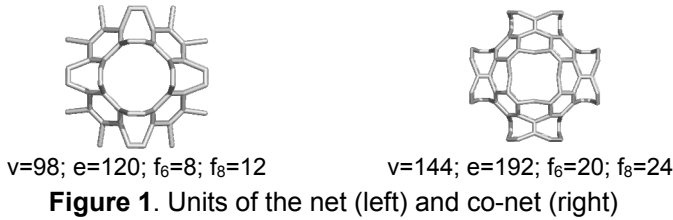
Figure 2. The net (3,3,3- left), superimposed net&co-net (2,2,2-central) and co-net (3,3,3- right) in a cubic (k,k,k) domain.

COUNTING POLYNOMIALS

A counting polynomial [19] is a representation of a graph $G(V,E)$, with the exponent k showing the extent of partitions $p(G)$, $\cup p(G) = P(G)$ of a graph property $P(G)$ while the coefficient $p(k)$ are related to the number of partitions of extent k.

$$P(x) = \sum_k p(k) \cdot x^k \qquad (1)$$

Let G be a connected graph, with the vertex set $V(G)$ and edge set $E(G)$. Two edges $e=(u,v)$ and $f=(x,y)$ of G are called *codistant* (briefly: *e co f*) if the notation can be selected such that [20]:

$$d(v,x) = d(v,y) + 1 = d(u,x) + 1 = d(u,y) \qquad (2)$$

where d is the usual shortest-path distance function. The above relation *co* is reflexive (*e co e*) and symmetric (*e co f*) for any edge e of G but in general is not transitive.

A graph is called a *co-graph* if the relation *co* is also transitive and thus an equivalence relation.

Let $C(e) := \{f \in E(G); f \; co \; e\}$ be the set of edges in G that are codistant to $e \in E(G)$. The set $C(e)$ can be obtained by an orthogonal edge-cutting procedure: take a straight line segment, orthogonal to the edge e, and intersect it and all other edges (of a polygonal plane graph) parallel to e. The set of these intersections is called an *orthogonal cut* (*oc* for short) of G, with respect to e.

If *G* is a *co-graph* then its orthogonal cuts $C_1, C_2, ..., C_k$ form a partition of *E(G)*: $E(G) = C_1 \cup C_2 \cup ... \cup C_k, \ C_i \cap C_j = \varnothing, i \neq j$.

A subgraph *H⊆ G* is called *isometric*, if $d_H(u,v) = d_G(u,v)$, for any $(u,v) \in H$; it is *convex* if any shortest path in *G* between vertices of *H* belongs to *H*. The relation *co* is related to ~ (Djoković [21]) and Θ (Winkler [22]) relations [23,24].

Two edges *e* and *f* of a plane graph *G* are in relation *opposite, e op f,* if they are opposite edges of an inner face of *G*. Then *e co f* holds by the assumption that faces are isometric. The relation *co* is defined in the whole graph while *op* is defined only in faces/rings. Note that John *et al*. [20] implicitly used the "*op*" relation in defining the Cluj-Ilmenau index *CI*.

Relation *op* will partition the edges set of *G* into *opposite* edge *strips ops*, as follows. (i) Any two subsequent edges of an *ops* are in *op* relation; (ii) Any three subsequent edges of such a strip belong to adjacent faces; (iii) In a plane graph, the inner dual of an *ops* is a path, an open or a closed one (however, in 3D networks, the ring/face interchanging will provide *ops* which are no more paths); (iv) The *ops* is taken as maximum possible, irrespective of the starting edge. The choice about the maximum size of face/ring, and the face/ring mode counting, will decide the length of the strip.

Also note that *ops* are *qoc* (quasi orthogonal cuts), meaning the transitivity relation is, in general, not obeyed.

The Omega polynomial [25-27] $\Omega(x)$ is defined on the ground of opposite edge strips *ops* $S_1, S_2, ..., S_k$ in the graph. Denoting by *m*, the number of *ops* of cardinality/length *s=|S|*, then we can write

$$\Omega(x) = \sum_s m \cdot x^s \qquad (3)$$

On *ops*, another polynomial, called Sadhana *Sd(x)* is defined [28,29]:

$$Sd(x) = \sum_s m \cdot x^{|E(G)|-s} \qquad (4)$$

The first derivative (in *x*=1) can be taken as a graph invariant or a topological index (e.g., *Sd'*(1) is the Sadhana index, defined by Khadikar *et al*. [30]):

$$\Omega'(1) = \sum_s m \cdot s = |E(G)| \qquad (5)$$

$$Sd'(1) = \sum_s m \cdot (|E(G)|-s) \qquad (6)$$

An index, called Cluj-Ilmenau [20], *CI(G)*, was defined on $\Omega(x)$:

$$CI(G) = \left\{ [\Omega'(1)]^2 - [\Omega'(1) + \Omega''(1)] \right\} \qquad (7)$$

In tree graphs, the Omega polynomial simply counts the non-opposite edges, being included in the term of exponent *s*=1.

POLYNOMIALS IN THE P-TYPE SURFACE NETWORKS

Omega and Sadhana polynomials are herein calculated at R_{max}[6]. Formulas for the two infinite networks are listed in Tables 1 and 2, with examples at the bottom of these tables.

In the discussed network, one can see that the coefficient $a(X^1)$ gives the number of octagons, by counting the edges not enumerated in the even faces. Next, $a(X^2)/3$ provides the number of hexagons while $a(X^4)/4$ counts the number of tubular necks (each bearing four anthracene units) joining the nodes of the net. In the co-net, the most informative is $a(X^4)/12$, giving the total number of the nodes while $(a(X^4)/12)^{1/3}=k$, the co-net parameter.

Table 1. Omega and Sadhana polynomials in the net

Formulas
$\Omega(X,R_{max}[6])=k^2(72+12(k-1))X+24k^3X^2+12k^2(k-1)X^4$
$\qquad =12k^2(k+5)X+24k^3X^2+12k^2(k-1)X^4$
$\Omega'(1)=12k^2(9k+1)$; $\Omega''(1)=48k^2(4k-3)$
$CI(G)=12k^2(972k^4+216k^3+12k^2-25k+11)$
$Sd(X,R_{max}[6])=k^2(72+12(k-1))X^{e-1}+24k^3X^{e-2}+12k^2(k-1)X^{e-4}$
$\qquad =12k^2(k+5)X^{12k^2(9k+1)-1}+24k^3X^{12k^2(9k+1)-2}+12k^2(k-1)X^{12k^2(9k+1)-4}$
$Sd'(1)=12k^2(9k+1)(48k^3+48k^2-1)=e(48k^3+48k^2-1)$

k	Omega polynomial: examples	$e(G)$	$CI(G)$
1	$72X^1+24X^2$	120	14232
2	$336X^1+192X^2+48X^4$	912	829872
3	$864X^1+648X^2+216X^4$	3024	9137664
4	$1728X^1+1536X^2+576X^4$	7104	50449728
	Sadhana polynomial: examples		$Sd'(1)$
1	$72X^{119}+24X^{118}$		11400
2	$336X^{911}+192X^{910}+48X^{908}$		524400
3	$864X^{3023}+648X^{3022}+216X^{3020}$		5222448
4	$1728X^{7103}+1536X^{7102}+576X^{7100}$		27272256

The number of atoms in the cubic domains (k,k,k) of the two lattices can be calculated by the formulas given in Table 3; some examples are available.

Table 2. Omega and Sadhana polynomials in co-net

Formulas
$\Omega(X,R_{max}[6])=k^2(96+12(k-1))X+24k^3X^2+12k^3X^4$
$\qquad =12k^2(k+7)X+24k^3X^2+12k^3X^4$
$\Omega'(1)=12k^2(9k+7)$; $\Omega''(1)=192k^3$
$CI(G)=12k^2(972k^4+1512k^3+588k^2-25k-7)$
$Sd(X,R_{max}[6])=k^2(96+12(k-1))X^{e-1}+24k^3X^{e-2}+12k^3X^{e-4}$
$\qquad =12k^2(k+7)X^{12k^2(9k+7)-1}+24k^3X^{12k^2(9k+7)-2}+12k^3X^{12k^2(9k+7)-4}$
$Sd'(1)=12k^2(9k+7)(48k^3+84k^2-1)=e(48k^3+84k^2-1)$

k	Omega polynomial: examples	e(G)	CI(G)
1	$96X^1+24X^2+12X^4$	192	36480
2	$432X^1+192X^2+96X^4$	1200	1437264
3	$1080X^1+648X^2+324X^4$	3672	13474728
4	$2112X^1+1536X^2+768X^4$	8256	68140992
	Sadhana polynomial: examples		Sd'(1)
1	$96X^{191}+24X^{190}+12X^{188}$		25152
2	$432X^{1199}+192X^{1198}+96X^{1196}$		862800
3	$1080X^{3671}+648X^{3670}+324X^{3668}$		7531272
4	$2112X^{8255}+1536X^{8254}+768X^{8252}$		36450240

Table 3. Number of atoms $v = |V(G)|$

Net			co-Net		
$v_k = 24 \cdot k^2(4 + 3(k-1)) = 24 \cdot k^2(3k+1)$			$v_k = k^2(144 + 72(k-1)) = 72k^2(k+1)$		
k	1	2	3	4
v for net	96	672	2160	4992
v for co-net	144	864	2592	5760

CONCLUSIONS

In this paper, the design of a hypothetical carbon crystal lattice, embedded in the P-type surface, achieved by identifying two opposite open faces of a unit, of octahedral symmetry, by the aid of Nano Studio software, was presented. The topology of the net and its co-net, thus obtained, was characterized by Omega and Sadhana counting polynomials. The ops strips proved to be informative about the construction of these infinite carbon nanostructures.

REFERENCES

1. M.V. Diudea and Cs.L. Nagy, "Periodic Nanostructures", Springer, **2007**.
2. M.V. Diudea, Ed., "Nanostructures, Novel Architecture", NOVA, N.Y., **2005**.
3. G.R. Patzke, F. Krumeich and R. Nesper, *Angew. Chem. Int. Ed.*, **2002**, *41*, 2447.
4. C.N.R. Rao and M. Nath, *Dalton Trans.*, **2003**, *1*, 1.
5. R. Tenne, *Chem. Eur. J.*, **2002**, *8*, 5296.
6. H. Imai, M. Matsuta, K. Shimizu, H. Hirashima and N. Negishi, *Solid State Ionics*, **2002**, *151*, 183.
7. M. Adachi, Y. Murata, I. Okada and S. Yoshikawa, *J. Electrochem. Soc.*, **2003**, *150*, G488.
8. Y. Zhou, L. Cao, F. Zhang, B. He and H. Li, *J. Electrochem. Soc.*, **2003**, *150*, A1246.
9. O.K. Varghese, D. Gong, M. Paulose, K.G. Ong and C.A. Grimes, *Sens. Actuators B*, **2003**, *93*, 338.
10. O.K. Varghese, D. Gong, M. Paulose, K.G. Ong, E.C. Dickey and C.A. Grimes, *Adv. Mater.*, **2003**, *15*, 624.
11. G.K. Mor, M.A. Carvalho, O.K. Varghese, M.V. Pishko and C.A. Grimes, *J. Mater. Res.*, **2004**, *19*, 628.

12. C.A. Grimes, K.G. Ong, O.K. Varghese, X. Yang, G. Mor, M. Paulose, E.C. Dickey, C. Ruan, M.V. Pishko, J.W. Kendig and A.J. Mason, *Sensors*, **2003**, *3*, 69.

13. R.W. Thompson and A. Dyer, *Zeolites*, **1985**, *5*, 292.

14. Zh. Liu, T. Ohsuna, O. Terasaki, M.A. Camblor, M.-J. Diaz- Cabañas and K. Hiraga, *J. Am. Chem. Soc.*, **2001**, *123*, 5370.

15. Zh. Yang, Y. Xia and R. Mokaya, *J. Am. Chem. Soc.*, **2007**, *129*, 1673.

16. G.O. Brunner, *Zeolites*, **1993**, *13*, *88.*

17. E.H. Ellison, *J. Phys. Chem.*, **2006**, *110*, 11406.

18. Cs.L. Nagy and M.V. Diudea, *Nano Studio*, "Babes-Bolyai" Univ., **2009**.

19. M.V. Diudea, "Nanomolecules and Nanostructures - Polynomials and Indices", MCM, No. 10, Univ. Kragujevac, Serbia, 2010.

20. P.E. John, A.E. Vizitiu, S. Cigher, and M.V. Diudea, *MATCH Commun. Math. Comput. Chem.,* **2007**, *57*, 479.

21. D.Ž. Djoković, *J. Combin. Theory Ser. B*, **1973**, *14*, 263.

22. P.M. Winkler, *Discrete Appl. Math.*, **1984**, *8*, 209.

23. S. Klavžar, *MATCH Commun. Math. Comput. Chem.*, **2008**, *59*, 217.

24. M.V. Diudea, S. Klavžar, *Acta Chem. Sloven.,* **2010**, *57*, 565.

25. M.V. Diudea, *Carpath. J. Math.*, **2006**, *22*, 43.

26. M.V. Diudea, S. Cigher, P.E. John, *MATCH Commun. Math. Comput. Chem.*, **2008**, *60*, 237.

27. M.V. Diudea, S. Cigher, A.E. Vizitiu, M.S. Florescu, and P.E. John*, J. Math. Chem.,* **2009**, *45*, 316.

28. A.R. Ashrafi, M. Ghorbani and M. Jalali, *Indian J. Chem.*, **2008**, *47A*, 535.

29. A.R. Ashrafi and M. Mirzargar, *Indian J. Chem.*, **2008**, *47A*, 538.

30. Khadikar, P.V.; Agrawal, V.K.; Karmarkar, S., *Bioorg. Med. Chem.*, **2002**, *10*, 3499.

OMEGA POLYNOMIAL FOR NANOSTRUCTURES
DESIGNED BY (P$_4$)^{k}Le OPERATIONS

MAHBOUBEH SAHELI[a], MIRCEA V. DIUDEA[b,*]

ABSTRACT. New cages are designed by repeating P$_4$ map operation and finalized by Le operation. The energy of some small non-classical fullerenes, tessellated according to above sequences of map operations was evaluated at the level of semiempirical method PM3. The topology of the networks is described in terms of Omega counting polynomial. Close formulas for this polynomial and the Cluj-Ilmenau index derived from it, as well as formulas to calculate the net parameters, are given.

Keywords: Counting polynomial, CI index, non-classical fullerenes.

INTRODUCTION

It is well established that covering/tessellation of fullerenes (nanostructures, in general) dictates the stability and reactivity of these molecules [1-3]. Covering and its modifications enables understanding of chemical reactions (their regioselectivity) occurring in nanostructures, particularly in carbon allotropes. In this respect, TOPO GROUP Cluj has developed some software programs [1], based on either well-known or original map operations [4-7]. A map M is a discretized (closed) surface [1].

We recall here the only two operations used in designing the proposed tessellation of the cages derived from the Platonic solids: tetrahedron T, octahedron Oct, Cube C, dodecahedron Do and icosahedron Ico.

Polygonal P$_4$ mapping is achieved by adding a new vertex in the center of each face and one point on the boundary edges; next, connecting the central point with one vertex on each edge, results in quadrilaterals covering [1,6].

Leapfrog Le is a composite operation, firstly described by Eberhard (1891) [8] and next by Fowler [9] and Diudea [6], that can be achieved as follows: add a point in the center of each face and join it with all the corners of a face, next truncate this point together with the edges incident in it (Figure 1). The original face will appear twisted by π/s, (s being the folding of the original face) and surrounded by polygons of 2d0 folding, where d0 is the degree of the parent vertices (in a regular graph).

---

[a] *Departmant of Mathematics, Faculty of Science, University of Kashan, Kashan 87317-51167, I.R. Iran, mmsaheli@yahoo.com*

[b] *Faculty of Chemistry and Chemical Engineering, "Babes-Bolyai" University, 400028 Cluj, Romania, diudea@gmail.com*
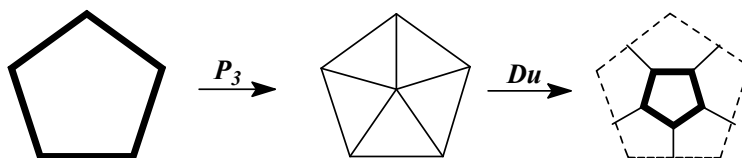
Figure 1. The Leapfrog *Le* operation on a pentagonal face

If the parent cage is a d_0 regular graph, the number of vertices in *Le*(M) is d_0 times larger than in the original map M, irrespective of the tessellation type. Note that in *Le*(M) the vertex degree is *always* 3, as a consequence of the involved triangulation P_3. In other words, the dual Du of a triangulation is a *cubic net* [2]. It is also true that truncation always provides a trivalent map. The leapfrog operation can be used to insulate the parent faces by surrounding (most often hexagonal) polygons.

CAGE BUILDING

Cages are built up, starting from the Platonic solids, by repeating the P_4 operation and finalized by Le operation; the sequence [10] of operations can be written as $Le(P_4(M))^k$. Due to dual pairs: Tetrahedron-Tetrahedron, T-T; Cube-Octahedron, C-Oct and Dodecahedron-Icosahedron, Do-Ico, there will be only three series of transformed cages (Figures 2 to 4, non-optimized). One can see that the central face/ring (in red) is twice the folding of parent face; similarly, the corner face (in blue) is twice the degree d_0 of parent vertices. These faces are distanced to each other by squares and octagons. The counting of faces/rings will be given below by the Ring polynomial [1].

Figure 2. $Le(P_4(T))^2$); v=192 ; 3D-vue (left) and orthoscopic vue (right)
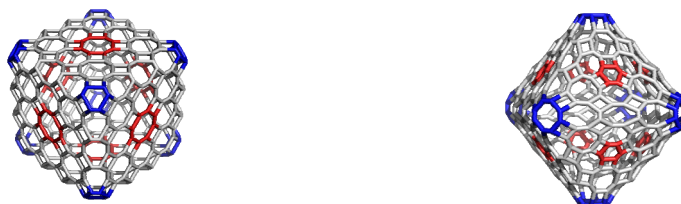
Figure 3. 3D-vue of $Le(P_4(C))^2$); v=192 ; (left) and $Le(P_4(Oct))^2$); v=192 ; (right)

Figure 4. 3D-vue of $Le(P_4(Do))^2$); v=960 ; (left) and $Le(P_4(Ico))^2$); v=960 ; (right)

ENERGETIC STABILITY

The calculations reported here were done at PM3 level of theory and serve only as arguments for the topological description of the interesting cages built up by $Le((P_4(M))^k)$ sequence of operations. Data, for the smallest representatives of this series (Figure 5) are listed in Table 1; for comparison, data for C_{60}, are also given.

Le(P₄(T))   Le(P₄(C))   Le(P₄(Do))
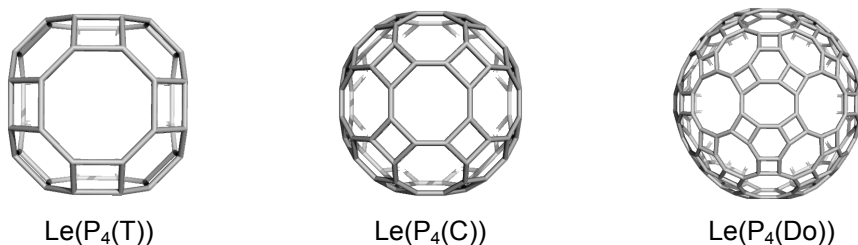
Figure 5. The smallest cages built up by $Le((P_4(M))^k)$.

It can be seen that the proposed cages show a moderate stability (by the values of heat of formation per number of atoms HF/N and HOMO-LUMO gap HLGAP), lower than that of C_{60}, the reference structure in nanoscience [1].

Regarding aromaticity, even C_{60} shows a low value of the geometry-based HOMA (harmonic oscillator model of aromaticity) index [11-13]; the new cages appear as anti-aromatic and this result is in agreement with the massive presence in structure of anti-aromatic faces f_4 and f_8, along with some aromatic f_6 and f_{10} (cf. Hückel theory) [14-16].

Table 1. Data for structures built up by $(P_4)^kLe$ and C_{60}; heat of formation per number of atoms HF/N; HOMO-LUMO gap HLGAP; point group symmetry Sym

	Name	N atoms	HF/N (kcal/mol)	HLGAP (eV)	Sym.	HOMA	POAV1
1	Le(P₄(T))	48	24.386	5.948	O_h	-0.871	9.457
2	Le(P₄(C))	96	20.633	5.917	O_h	-0.868	4.831
3	Le(P₄(Do))	240	19.597	6.047	I_h	-0.879	2.067
4	C_{60}	60	13.514	6.596	I_h	0.169	8.257

The last column in Table 1 refers to the strain of cage covering, in terms of Haddon's theory [17-19]. Clearly, the larger cage is the most relaxed structure and this is supported by the lowest value of HF/N.

Computations were done by MOPAC2009 software package [20]. Calculations at a higher level of quantum chemistry are in progress in our lab.

OMEGA POLYNOMIAL

In a connected graph $G(V,E)$, with the vertex set $V(G)$ and edge set $E(G)$, two edges $e = uv$ and $f = xy$ of G are called *codistant e co f* if they obey the relation [21]:

$$d(v,x) = d(v,y)+1 = d(u,x)+1 = d(u,y) \tag{1}$$

which is reflexive, that is, *e co e* holds for any edge *e* of *G*, and symmetric, if *e co f* then *f co e*. In general, relation *co* is not transitive; if "*co*" is also transitive, thus it is an equivalence relation, then *G* is called a *co-graph* and the set of edges $C(e) := \{f \in E(G); f\ co\ e\}$ is called an *orthogonal cut oc* of *G*, $E(G)$ being the union of disjoint orthogonal cuts: $E(G) = C_1 \cup C_2 \cup ... \cup C_k,\ C_i \cap C_j = \varnothing, i \neq j$. Klavžar [22] has shown that relation *co* is a theta Djoković-Winkler relation [23,24].

We say that edges *e* and *f* of a plane graph *G* are in relation *opposite, e op f,* if they are opposite edges of an inner face of *G*. Note that the relation *co* is defined in the whole graph while *op* is defined only in faces. Using the relation *op* we can partition the edge set of *G* into *opposite* edge *strips, ops.* An *ops* is a quasi-orthogonal cut *qoc*, since *ops* is not transitive.

Let *G* be a connected graph and $S_1, S_2, ..., S_k$ be the *ops* strips of *G*. Then the *ops* strips form a partition of $E(G)$. The length of *ops* is taken as maximum. It depends on the size of the maximum fold face/ring F_{max}/R_{max} considered, so that any result on Omega polynomial will have this specification.

Denote by $m(G,s)$ the number of *ops* of length *s* and define the Omega polynomial as [25-33]:

$$\Omega(G,x) = \sum_s m(G,s) \cdot x^s \tag{2}$$

Its first derivative (in $x=1$) equals the number of edges in the graph:

$$\Omega'(G,1) = \sum_s m(G,s) \cdot s = e = |E(G)| \tag{3}$$

On Omega polynomial, the Cluj-Ilmenau index [21], $CI=CI(G)$, was defined:

$$CI(G) = \left\{ [\Omega'(G,1)]^2 - [\Omega'(G,1) + \Omega''(G,1)] \right\} \tag{4}$$

RESULTS AND DISCUSSION

Cage parameters

Since the starting cages of this study are the graphs of Platonic solids, let's present the net parameters of these structures in Table 2, as $|p_0|$ parameters, p being vertices v (of degree d), edges e and faces f (of various folding s). By applying the sequence of operations $Le(P_4(M))^k$, the transformed maps will show all the vertex degree d=3. Formulas for the value of the other

228

parameters are given in Table 3. Observe, in the dual pair, the face of parent becomes the vertex of transform and this interchanging operates also on the corresponding parameters: s_0f_0 becomes d_0v_0, while the number of edges remains unchanged.

Table 2. Platonic solid graph parameters

Graph	Vertices $|v_0|$	Degree d_0	Edges $|e_0|$	Ring folding s_0	Faces $|f_0|$
T	4	3	6	3	4
C	8	3	12	4	6
Oct	6	4	12	3	8
Do	20	3	30	5	12
Ico	12	5	30	3	20

Table 3. Transforms of the Platonic solid graphs by $Le(P_4(M))^k)$

M	Vertices $|v_0|$	Edges $|e_0|$	Faces $|f_0|$
T	12×4^k	18×4^k	$6 \times 4^k + 2$
C	24×4^k	36×4^k	$12 \times 4^k + 2$
Do	60×4^k	90×4^k	$30 \times 4^k + 2$
Formula	$|v| = 4^k \times s_0 f_0$	$|e| = 3 \times 4^k \times e_0$	$|f| = 4^k \times e_0 + 2$
Oct	24×4^k	36×4^k	$12 \times 4^k + 2$
Ico	60×4^k	90×4^k	$30 \times 4^k + 2$
Formula	$|v| = 4^k \times d_0 v_0$	$|e| = 3 \times 4^k \times e_0$	$|f| = 4^k \times e_0 + 2$

Ring polynomial

The ring polynomial for the graphs originating in trivalent Platonics is as follows:

$$R(Le((P_4(\mathrm{T}))^k), x) = \left(3 \times 4^a\right) x^4 + 8x^6 + \left(3 \times 4^a - 6\right) x^8 \tag{5}$$

$$R(Le((P_4(\mathrm{C}))^k), x) = \left(6 \times 4^a\right) x^4 + 8x^6 + \left(6 \times 4^a - 6\right) x^8 \tag{6}$$

$$R(Le((P_4(\mathrm{Do}))^k), x) = \left(15 \times 4^a\right) x^4 + 20\, x^6 + \left(15 \times 4^a - 30\right) x^8 + 12\, x^{10} \tag{7}$$

Generalizing, for the graphs transformed from the trivalent Platonics, the formula for ring polynomial is of the form:

$$R(Le((P_4(G(d_0 : 3)))^k), x) = \left(s_0 f_0 \times 2^{2(k-1)}\right) x^4 + v_0 x^6 + \\ \left(s_0 f_0 \times 2^{k-1}(2^{k-1} - 1) + e_0(2^k - 1)\right) x^8 + f_0\, x^{2s_0} \tag{8}$$

Now, considering the relation between the dual pairs, for the trigonal Platonics we have:

$$R(Le((P_4(G(f_0 : 3)))^k), x) = \left(d_0 v_0 \times 2^{2(k-1)}\right) x^4 + f_0 x^6 + \\ \left(d_0 v_0 \times 2^{k-1}(2^{k-1} - 1) + e_0(2^k - 1)\right) x^8 + v_0\, x^{2d_0} \tag{9}$$

Omega Polynomial

The Omega polynomial (calculated at $R_{max}[8]$) for the graphs transformed from the trivalent Platonics is as follows:

$$\Omega(Le((P_4(\mathrm{T}))^k), x) = \left(3(2^k - 1) + 6\right) x^{2^{k+2}} + 4(2^{k-1} - 1)\, x^{3 \times 2^k} \quad (10)$$

$$\Omega(Le((P_4(\mathrm{C}))^k), x) = \left(4(2^k - 1) + 6\right) x^{3 \times 2^{k+1}} + \left(6(2^{k-1} - 1) + 3\right) x^{2^{k+2}} \ (11)$$

$$\Omega(Le((P_4(\mathrm{Do}))^k), x, R[8]) = 6(2^k - 1)\, x^{5 \times 2^{k+1}} + 12(2^{k-1} - 1)\, x^{5 \times 2^k} + 15\, x^{2^{k+3}} \ (12)$$

Generalizing, we have:

$$\Omega(Le((P_4(\mathrm{G}))^k), x) = \left(f_0(2^{k-1} - 1) + 3\frac{1 + (-1)^{s_0}}{2} \right) x^{s_0 \times 2^k} + $$

$$\left(s_0 + \left\lfloor \frac{s_0 + 1}{6} \right\rfloor \right)(2^k - 1)\, x^{\left(s_0 + \left\lfloor \frac{s_0+1}{6} \right\rfloor - 1\right) \times 2^{k+1}} + \frac{e_0}{\left\lceil \frac{s_0}{3} \right\rceil}\, x^{(s_0 - 1) \times 2^{k+1}} \quad (13)$$

And for *CI* we have:

$$CI\,(Le((P_4(\mathrm{T}))^k)) = 324 \cdot 4^{2k} - 6 \cdot 4^k (11 \cdot 2^k - 1) - 18 \cdot 4^k \quad (14)$$

$$CI\,(Le((P_4(\mathrm{C}))^k)) = 1296 \cdot 4^{2k} - 12 \cdot 4^k (16 \cdot 2^k - 1) - 36 \cdot 4^k \quad (15)$$

$$CI\,(Le((P_4(\mathrm{Do}))^k)) = 8100 \cdot 4^{2k} - 30 \cdot 4^k (25 \cdot 2^k - 1) - 90 \cdot 4^k \quad (16)$$

The Omega polynomial, calculated at $R_{max}=10$, in case M=Do, is as follows.

$$\Omega(Le((P_4(\mathrm{Do}))^k), x, R[10]) = 6(k^2 - p - 2) \cdot x^{5(k^2 - p)} + 15 \cdot x^{2^k + 3} + $$

$$6(k^2 - p - 1) \cdot x^{10(k^2 - p)} \quad (17)$$

$$\Omega'(Le((P_4(\mathrm{Do}))^k), 1, R[10]) = 120p - 180k^2 p + 120 \cdot 2^k - 120k^2 + 90k^4 + 90p^2 \ (18)$$

$$CI(Le((P_4(\mathrm{Do}))^k), R[10]) = 2250k^4 p - 1800k^2 p + 900k^4 - 750k^6 + $$

$$900p^2 + 750p^3 - 2250k^2 p^2 - 960 \cdot 2^{2k} + (\Omega'(Le((P_4(\mathrm{Do}))^k), 1, R[10]))^2 \quad (19)$$

Table 4 lists some examples for the formulas derived within this paper. Computations were made by Nano Studio software [34].

Table 4. Examples for the herein derived formulas

$Le((P_4(M))^k)$ M ; k ; R_{max}	V	Omega polynomial	CI	Ring polynomial
T; k=3 ; R[8]	768	$12x^{24}+27x^{32}$	1292544	$192x^4+8x^6+210x^8$
C; k=3 ; R[8]	1536	$21x^{32}+34x^{48}$	5208576	$384x^4+8x^6+402x^8$
Do; k=3;R[8]	3840	$30x^8+30x^{24}+36x^{40}+42x^{80}$	32832000	$960x^4+20x^6+930x^8+12x^{10}$
Do; k=3;R[10]		$36x^{40}+15x^{64}+42x^{80}$	32789760	
T; k=4; R[8]	3072	$28x^{48}+51x^{64}$	20960256	$768x^4+8x^6+786x^8$
C; k=4; R[8]	6144	$45x^{64}+66x^{96}$	84142080	$1536x^4+8x^6+1554x^8$
Do; k=4; R[8]	15360	$30x^{16}+30x^{48}+84x^{80}+90x^{160}$	527923200	$3840x^4+20x^6+3810x^8+12x^{10}$
Do; k=4; R[10]		$84x^{80}+15x^{128}+90x^{160}$	527754240	

CONCLUSIONS

In this article, new cages designed by $Le((P_4(M))^k)$ sequence of map operations are reported. The energy of some small non-classical fullerenes, tessellated according to the above map operations was evaluated at the level of semiempirical method PM3; it was shown that these non-classical fullerenes have a moderate stability, less than the reference C60 fullerene, a result pertinent for a tessellation with massive anti-aromatic faces R_4 and R_8. The topology of the networks was described in terms of Omega counting polynomial. Close formulas for this polynomial and the Cluj-Ilmenau index, as well as formulas to calculate the net parameters, were given.

ACKNOWLEDGEMENTS

REFERENCES

1. M.V. Diudea, "Nanomolecules and Nanostructures - Polynomials and Indices, MCM", No. 10, Univ. Kragujevac, Serbia, **2010**.
2. M.V. Diudea, Cs.L. Nagy, "Periodic Nanostructures", Springer, **2007**.
3. M.V. Diudea, (Ed.), "Nanostructures, Novel Architecture", NOVA, New York, **2005**.
4. M.V. Diudea, M. Ştefu, P.E. John, and A. Graovac, *Croat. Chem. Acta,* **2006**, *79*, 355.
5. M.V. Diudea, *J. Chem. Inf. Model.,* **2005**, *45,* 1002.
6. M.V. Diudea, *Forma* (Tokyo), **2004**, *19* (3), 131.
7. M. Stefu, M.V. Diudea and P.E. John, *Studia Univ. Babes-Bolyai Chemia*, **2005**, *50*, 2, 165.
8. V. Eberhard, *Zur Morphologie der Polyeder*, Leipzig,Teubner, **1891**.
9. P.W. Fowler, *Chem. Phys. Lett.*, **1986**, *131*, 444.
10. M.V. Diudea, *Phys. Chem., Chem. Phys.,* **2005**, 7, 3626.
11. T.M. Krygowski and A. Ciesielski, *J. Chem. Inf. Comput. Sci.*, **1995**, *35*, 203.
12. T.M. Krygowski and A. Ciesielski, *J. Chem. Inf. Comput. Sci.*, **1995**, *35*, 1001.
13. T.M. Krygowski and M. Cyranski, *Tetrahedron*, **1996**, *52*, 10255.
14. E. Hückel, *Z. Physik*, **1931**, *70*, 204;
15. E. Hückel, *Z. Physik,* **1931**, *72*, 310;
16. E. Hückel, Grundzüge der Theorie ungesättigter und *aromatischer Verbindungen*, Verlag Chemie: Berlin, 1938.
17. R.C. Haddon, *J. Am. Chem. Soc.*, **1987**, *109*, 1676.
18. R.C. Haddon, *J. Am. Chem. Soc.*, **1990**, *112*, 3385.
19. R.C. Haddon, S.-Y. Chow, *J. Am. Chem. Soc.*, **1998**, *120*, 10494.
20. Stewart, J.J.P., MOPAC2009. 2008, Stewart Computational Chemistry: Colorado Springs, CO, USA.

21. P.E. John, A.E. Vizitiu, S. Cigher, M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2007**, *57*, 479.
22. S. Klavžar, *MATCH Commun. Math. Comput. Chem.*, **2008**, *59*, 217.
23. D.Ž. Djoković, *J. Combin. Theory Ser. B*, **1973**, *14*, 263.
24. P.M. Winkler, *Discrete Appl. Math.*, **1984**, *8*, 209.
25. M.V. Diudea, *Carpath. J. Math.*, **2006**, *22*, 43.
26. A.R. Ashrafi, M. Jalali, M. Ghorbani and M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2008**, *60*, 905.
27. M.V. Diudea, S. Cigher, P.E. John, *MATCH Commun. Math. Comput. Chem.*, **2008**, *60*, 237.
28. A.E. Vizitiu, S. Cigher, M.V. Diudea, M.S. Florescu, *MATCH Commun. Math. Comput. Chem.*, **2007**,*57* (2) 457.
29. M.V. Diudea, S. Cigher, A.E. Vizitiu, O. Ursu and P.E. John, *Croat. Chem. Acta*, **2006**, *79*, 445.
30. M.V. Diudea, S. Cigher, A.E. Vizitiu, M.S. Florescu, P.E. John, *J. Math. Chem.*, **2009**, *45*, 316.
31. M. Saheli, M. Neamati, K. Nagy and M.V. Diudea, *Studia Univ. Babes-Bolyai Chemia*, **2010**, *55* (1), 83.
32. M.V. Diudea, *Acta Chim. Slov.*, **2010**, 57, 551.
33. M.V. Diudea, S. Klavžar, *Acta Chim. Slov.,* **2010**, *57*, 565.
34. Cs.L. Nagy, M.V. Diudea, *Nano Studio software*, Babes-Bolyai University, **2009**.

THE OMEGA POLYNOMIAL OF THE CORCOR DOMAIN OF GRAPHENE

MAHBOUBEH SAHELI[a], ALI REZA ASHRAFI[a*], MIRCEA V. DIUDEA[b]

ABSTRACT. An opposite edge strip *ops* with respect to a given edge of a graph is the smallest subset of edges closed under taking opposite edges on faces. The Omega polynomial is a counting polynomial whose k-th coefficient is the number m(G,k) of *ops* containing k-edges. In this paper an exact formula for the Omega polynomial of the molecular graph of a new type of graphene named CorCor is given. As a consequence, the PI index of this nanostructure is computed.

Keywords: Omega polynomial, CorCor

INTRODUCTION

Throughout this paper, a graph means a simple connected graph. Suppose G is a graph and u, v are vertices of G. The distance d(u,v) is defined as the length of a shortest path connecting u and v in G. A graph can be described by a connection table, a sequence of numbers, a matrix, a polynomial or a derived unique number which is called a topological index. When we describe a graph by a polynomial as $P(G,x) = \sum_k m(G,k)x^k$, then we must find algorithms to compute the coefficients m(G,k), for each k, see [1-3].

Suppose G is a connected bipartite graph, with the vertex set V(G) and edge set E(G). Two edges e = uv and f = xy of G are called co-distant (briefly: e co f) if d(v,x) = d(v,y) + 1 = d(u,x) + 1 = d(u,y). It is far from true that the relation "co" is equivalence relation, but it is reflexive and symmetric.

Let C(e) = { f $\in$ E(G) | f co e} denote the set of edges in G, co-distant to the edge e $\in$ E(G). If relation "co" is an equivalence relation then G is called a co-graph. Consequently, C(e) is called an orthogonal cut *oc* of G and E(G) is the union of disjoint orthogonal cuts. If two consecutive edges of an edge-cut sequence are opposite, or "topologically parallel" within the same face/ring of the covering, such a sequence is called an opposite edge strip *ops* which is a quasi-orthogonal cut *qoc* strip. This means that the transitivity relation of the "co" relation is not necessarily obeyed. Any *oc* strip is an *op* strip but the reverse is not always true.

Let m(G,k) denote the multiplicity of a qoc strip of length k. For the sake of simplicity, we define m = m(G,k) and e = |E(G)|. A counting polynomial can be defined in simple bipartite graphs as $\Omega(G,x) = \sum_e mx^k$, named Omega

---
[a] *Department of Mathematics, University of Kashan, Kashan 87317-51167, I. R. Iran, Email: ashrafi@kashanu.ac.ir*
[b] *Faculty of Chemistry and Chemical Engineering, "Babes-Bolyai" University, 400028 Cluj, Romania*

polynomial of G. This polynomial was introduced by one of the present authors (MVD) [4]. Recently, some researchers computed the Omega and related polynomials for some types of nanostructures [5-10].

In this paper, we continue our earlier works on the problem of computing Omega polynomials of nanostructures. We focus on a new type of nanostructures named CorCor, a domain of the graphene – a 2-dimensional carbon network, consisting of a single layer of carbon atoms, and compute its Omega polynomial, Figure 1. Our notation is standard and mainly taken from the standard books of graph theory.

Main Results and Discussion

In this section, the Omega polynomial of G[n] = CorCor[n] with n layers (Figure 1) is computed. At first, we notice that the molecular graph of G[n] has exactly $42n^2 - 24n + 6$ vertices and $63n^2 - 45n + 12$ edges. The molecular graph G[n] is constructed from $6n - 3$ rows of hexagons. For example, the graph G[3] has exactly 15 rows of hexagons and the number of hexagons in each row is according to the following sequence:

2, 5, 9, 10, 11, 12, 12, **11**, 12, 12, 11, 10, 9, 5, 2

The $(3n - 1)^{th}$ row of G[n] is called the central row of G[n]. This row has exactly $2\left(3\left\lceil\dfrac{n}{3}\right\rceil + 2(n - \left\lceil\dfrac{n}{3}\right\rceil)\right) - 3 = 4n + 2\left\lceil\dfrac{n}{3}\right\rceil - 3$ hexagons, where for a real number x, $\lceil x \rceil$ denotes the smallest integer greater or equal to x. The central hexagon of G[n] is surrounded by six hexagons. If we replace each hexagon by a vertex and connect such vertices according to the adjacency of hexagons, then we will find a new hexagon containing the central hexagon of G[n]. Next consider the adjacency relationship between the hexagons of the second layer of G[n] and construct a new hexagon containing the last one and so on, see Figure 1. The hexagons constructed from this algorithm are called the big hexagons. By our algorithm, the hexagons of G[n] are partitioned into the following two classes of hexagons:
   a) The hexagons crossing the edges of big hexagons, i.e. those depicted by thick line.
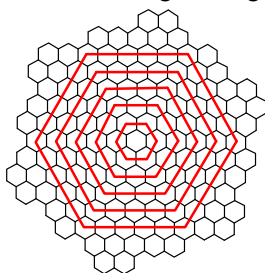   b) The hexagons outside the big hexagon.

Figure 1. The Molecular Graph of CorCor[3].

One can see that the number of rows inside and outside big hexagons are equal to $\left(6\left\lceil\dfrac{n}{3}\right\rceil + 4(n - \left\lceil\dfrac{n}{3}\right\rceil)\right) - 4 = 4n + 2\left\lceil\dfrac{n}{3}\right\rceil - 4$ and

$6n - 4 - (4n + 2\left\lceil\dfrac{n}{3}\right\rceil - 4) = 2n - 2\left\lceil\dfrac{n}{3}\right\rceil$, respectively. From Figure 1, one can see that the molecular graph of CorCor[n] can be partitioned into six equal parts with the same number of hexagons. If we consider one half of this graph then three cases of these six parts must be considered. Define three matrices A, A' and A'' as follows:

- A is an $\left(n - \left\lceil\dfrac{n}{3}\right\rceil\right) \times \left(2n - 2 + \left\lceil\dfrac{n}{3}\right\rceil\right)$ matrix with 0 & 1 entries.

The entries corresponding to the hexagons of CorCor[n] are equal to 1, and other entries are zero, see Figure 2. As an example, the matrix A_6 is as follows:

$$A_6 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

- Suppose $A = [a_{ij}]$, $A' = [b_{ij}]$ is an $\left(2n - 2 + \left\lceil\dfrac{n}{3}\right\rceil\right) \times \left(n - \left\lceil\dfrac{n}{3}\right\rceil\right)$ matrix defined by

$$b_{ij} = \begin{cases} a_{\left(n - \left\lceil\frac{n}{3}\right\rceil - j + 1\right)(i+j-1)} & i + j \leq 2n - 1 + \left\lceil\dfrac{n}{3}\right\rceil \\ a_{\left(n - \left\lceil\frac{n}{3}\right\rceil - j + 1\right)\left(i+j-2n+1-\left\lceil\frac{n}{3}\right\rceil\right)} & i + j > 2n - 1 + \left\lceil\dfrac{n}{3}\right\rceil \end{cases}$$

$A'' = [c_{ij}]$ is an $\left(2n - 2 + \left\lceil\dfrac{n}{3}\right\rceil\right) \times \left(n - \left\lceil\dfrac{n}{3}\right\rceil\right)$ matrix defined by $c_{ij} = a_{j\left(2n-1+\left\lceil\frac{n}{3}\right\rceil - i\right)}$.

It is easy to see that the number of hexagons in the central row of G[n] is $2\left(3\left\lceil\dfrac{n}{3}\right\rceil + 2\left(n - \left\lceil\dfrac{n}{3}\right\rceil\right)\right) - 3 = 4n + 2\left\lceil\dfrac{n}{3}\right\rceil - 3$. Suppose S_i' and S_i'' denote the summation of all entries in the ith row of the matrices A' and A'', respectively. Then

$S_i' = \sum_{j=0}^{n-1-\lceil n/3\rceil} a_{(n-j-\lceil n/3\rceil)(i+j)}$ and $S_i'' = \sum_{j=1}^{n-\lceil n/3\rceil} a_{j\left(2n-i-1+\lceil n/3\rceil\right)}$
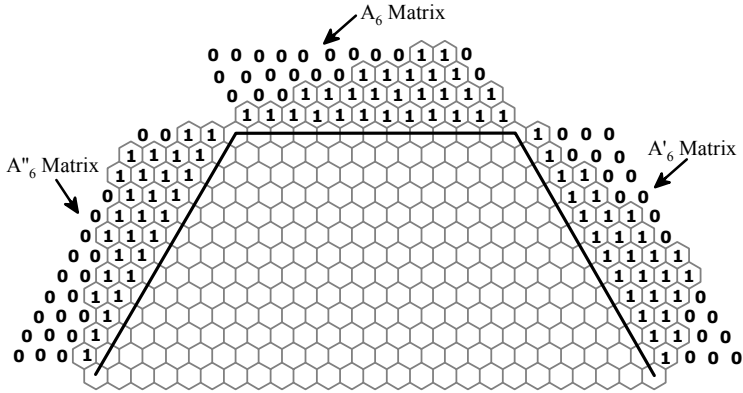
Figure 2. Construction of the Matrices A_6, A_6' and A_6''.

and so,

$$\Omega(G[n], x) = 6\sum_{j=1}^{n-\lceil n/3 \rceil} x^{3j-1+\lceil j/2 \rceil}$$

$$+ 6\sum_{j=1}^{2n-2+\lceil n/3 \rceil} x^{1+S_j'+S_j''+(2n+j-2+\lceil n/3 \rceil)}$$

$$+ 3x^{4n-2+2\lceil n/3 \rceil}$$

Thus for computing the omega polynomial of G[n], it is enough to compute S_i' and S_i''. By a simple calculations, one can see that $\Omega(G[1]) = 6x^3 + 3x^4$ and $\Omega(G[2]) = 6x^3 + 6x^6 + 15x^8$. So, we can assume that n ≥ 3. Our main proof consider three cases that n ≡ 0 (mod 3), n ≡ 1 (mod 3) and n ≡ 2 (mod 3).

We first assume that n ≡ 0 (mod 3). In this case the number of rows in the big hexagons is 7n/3 − 2. By definition of A_n, if *1 ≤ j ≤ 4n/3 − 2* then we have $S_j' = \lceil j/2 \rceil$. If *4n/3 − 1 ≤ j ≤ 7n/3 − 2* then we can define $j = 4n/3 - 2 + k$, where *1 ≤ k ≤ n*. Thus,

 ➢ $S_i' = 2n/3 - (2k - 2)$, where k ≡ 1 or 2 (mod 3),

 ➢ $S_i' = 2n/3 - (2k - 1)$, where k ≡ 0 (mod 3).

To compute S_j'', we consider four cases that *1 ≤ j ≤ n/3 − 1, j = n/3, j = n/3 + 1* and *n/3 + 2 ≤ j ≤ 7n/3 − 2*. In the first case $S_j'' = 2j$, and for the second and third cases we have $S_j'' = 2n/3$. For the last case, we assume that j = n/3 + k + 1, 1 ≤ k ≤ 2n − 3. Then $S_j'' = n - n/3 - \lceil k/3 \rceil = 2n/3 - \lceil k/3 \rceil$. To compute the omega polynomial, we define the following polynomials:

1) $\varepsilon_1 = \sum_{j=1}^{n/3-1} x^{3j-1+7n/3+\lceil j/2 \rceil}$,

2) $\varepsilon_2 = x^{-1+10n/3+\lceil n/6 \rceil} + x^{1+10n/3+\lceil (n-3)/6 \rceil}$,

3) $\varepsilon_3 = \sum\limits_{j=n/3+2}^{4n/3-2} x^{3n-1+j+\lceil j/2 \rceil - \lceil (3j-n-3)/9 \rceil}$,

4) $\varepsilon_4 = \sum\limits_{j=4n/3-1}^{7n/3-2} x^{j+7n/3+S_j'+S_j''}$.

Therefore, $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_4 = \sum\limits_{j=1}^{2n-2+\lceil n/3 \rceil} x^{1+S_j'+S_j''+(2n+j-2+\lceil n/3 \rceil)}$.

To simplify these quantities, two cases that n is odd or even are considered. If n is even then

$$\Omega(G,x) = [6/(1-x^7)] \times \left(\begin{array}{l} x^3+x^6+x^{\frac{7n}{2}}\left(2x^1+x^4+x^5\right) - \\ x^{\frac{14n}{3}}\left(\frac{1}{2}x^{-2}+x^{-1}+2x^1+x^3+x^4+\frac{1}{2}x^5+\frac{n}{3}(-x^{-2}-2x^{-1}+x^5+2x^6)\right) \end{array} \right)$$

if n is odd, then

$$\Omega(G,x) = [6/(1-x^7)] \times \left(\begin{array}{l} x^3+x^6+x^{\frac{7n}{2}}\left(x^{\frac{1}{2}}+x^{\frac{3}{2}}+2x^{\frac{9}{2}}\right) - \\ x^{\frac{14n}{3}}\left(\frac{1}{2}x^{-2}+x^{-1}+2x^1+x^3+x^4+\frac{1}{2}x^5+\frac{n}{3}(-x^{-2}-2x^{-1}+x^5+2x^6)\right) \end{array} \right)$$

Using a similar argument as above, if $n \equiv 1 \ (mod \ 3)$ then for even n,

$$\Omega(G,x) = [6/(1-x^7)] \times \left(\begin{array}{l} x^3+x^6+x^{\frac{7n}{2}}\left(2x+x^4+x^5\right) - \\ x^{\frac{14n}{3}}\left(\frac{2}{3}x^{\frac{-5}{3}}+\frac{5}{6}x^{\frac{-2}{3}}+x^{\frac{1}{3}}+x^{\frac{4}{3}}+2x^{\frac{10}{3}}+\frac{1}{3}x^{\frac{16}{3}}+\frac{1}{6}x^{\frac{19}{3}}+\frac{n}{3}(x^{\frac{19}{3}}+2x^{\frac{16}{3}}-2x^{\frac{-5}{3}}-x^{\frac{-2}{3}})\right) \end{array} \right)$$

and for odd n,

$$\Omega(G,x) = [6/(1-x^7)] \times \left(\begin{array}{l} -x^3-x^6+x^{\frac{7n}{2}}\left(x^{\frac{1}{2}}+x^{\frac{3}{2}}+2x^{\frac{9}{2}}\right) - \\ x^{\frac{14n}{3}}\left(\frac{2}{3}x^{\frac{-5}{3}}+\frac{5}{6}x^{\frac{-2}{3}}+x^{\frac{1}{3}}+x^{\frac{4}{3}}+2x^{\frac{10}{3}}+\frac{1}{3}x^{\frac{16}{3}}+\frac{1}{6}x^{\frac{19}{3}}+\frac{n}{3}(-x^{\frac{19}{3}}-2x^{\frac{16}{3}}+2x^{\frac{-5}{3}}+x^{\frac{-2}{3}})\right) \end{array} \right)$$

Finally, if $n \equiv 2 \ (mod \ 3)$ then for even n,

$$\Omega(G,x) = [6/x^4(1-x^7)] \times \left(\begin{array}{l} x^7+x^{10}+x^{\frac{7n}{2}}\left(2x^5+x^8+x^9\right) - \\ x^{\frac{14n}{3}}\left(\frac{3}{2}x^{\frac{8}{3}}+x^{\frac{14}{3}}+x^{\frac{17}{3}}+x^{\frac{20}{3}}+x^{\frac{23}{3}}+\frac{1}{2}x^{\frac{29}{3}}+n(-x^{\frac{8}{3}}+x^{\frac{29}{3}})\right) \end{array} \right),$$

and for odd n,

$$\Omega(G,x) = [6/x^4(1-x^7)] \times \left(\begin{array}{l} x^7 + x^{10} + x^{\frac{7n}{2}}\left(x^{\frac{9}{2}} + x^{\frac{11}{2}} + 2x^{\frac{17}{2}}\right) - \\[2mm] x^{\frac{14n}{3}}\left(\frac{3}{2}x^{\frac{8}{3}} + x^{\frac{14}{3}} + x^{\frac{17}{3}} + x^{\frac{20}{3}} + x^{\frac{23}{3}} + \frac{1}{2}x^{\frac{29}{3}} + n(-x^{\frac{8}{3}} + x^{\frac{29}{3}})\right) \end{array} \right)$$

It is now possible to simplify our calculations as follows:

$$\Omega(G,x) = [6/x^4(1-x^7)] \begin{cases} x^4\left(x^3 + x^6 + x^{7/2}(2x + x^4 + x^5) - x^{14/3} \times R_0(x)\right) & n \equiv 0 \ (\mathrm{mod}\ 6) \\[1mm] x^4\left(x^3 + x^6 + x^{7/2}(2x^{1/2} + x^{3/2} + 2x^{9/2}) - x^{14/3} \times R_1(x)\right) & n \equiv 1 \ (\mathrm{mod}\ 6) \\[1mm] x^7 + x^8 + x^9 + x^{10} + x^{12} - x^{14/3} \times R_2(x) & n \equiv 2 \ (\mathrm{mod}\ 6) \\[1mm] x^4\left(x^3 + x^6 + x^{7/2}(2x^{1/2} + x^{3/2} + 2x^{9/2}) - x^{14/3} \times R_0(x)\right) & n \equiv 3 \ (\mathrm{mod}\ 6) \\[1mm] x^4\left(x^3 + x^6 + x^{7/2}(2x + x^4 + x^5) - x^{14/3} \times R_1(x)\right) & n \equiv 4 \ (\mathrm{mod}\ 6) \\[1mm] x^7 + 2x^{17/2} + x^{10} + x^{23/2} + x^{25/2} - x^{14/3} \times R_2(x) & n \equiv 5 \ (\mathrm{mod}\ 6) \end{cases}$$

where,

$$R_0(x) = 1/2x^{-2} + x^{-1} + 2x + x^3 + x^4 + 1/2x^5 + n/3\left(-x^{-2} - 2x^{-1} + x^5 + 2x^6\right),$$

$$R_1(x) = 2/3x^{-5/3} + 5/6x^{-2/3} + x^{1/3} + x^{4/3} + 2x^{10/3} + 1/3x^{16/3} + 1/6x^{19/3}$$
$$+ (n/3)\left(-2x^{-5/3} + x^{-2/3} + 2x^{16/3} + x^{19/3}\right),$$

$$R_2(x) = (3/2 - n)x^{8/3} + x^{14/3} + x^{17/3} + x^{20/3} + x^{23/3} + (n + 1/2)x^{29/3}.$$

Since G[n] is a partial cube by a result of Klavzar [11],

$$PI(G[n]) = CI(G[n]) = \left[\Omega'(G[n], x)^2 - \Omega'(G[n], x) - \Omega''(G[n], x)\right]_{x=1}$$

We now apply above calculations to compute the PI index of G[n]. We have:

$$PI(G[n]) = \begin{cases} 3969n^4 - \frac{53333}{9}n^3 + \frac{22769}{6}n^2 - 1196n + 162 & n \equiv 0 \ (\mathrm{mod}\ 6) \\[2mm] 3969n^4 - \frac{53333}{9}n^3 + \frac{22769}{6}n^2 - 1196n + \frac{2809}{18} & n \equiv 1 \ (\mathrm{mod}\ 6) \\[2mm] 3969n^4 - \frac{53333}{9}n^3 + \frac{11458}{3}n^2 - \frac{3743}{3}n + \frac{1504}{9} & n \equiv 2 \ (\mathrm{mod}\ 6) \\[2mm] 3969n^4 - \frac{53333}{9}n^3 + \frac{22769}{6}n^2 - 1196n + \frac{321}{2} & n \equiv 3 \ (\mathrm{mod}\ 6) \\[2mm] 3969n^4 - \frac{53333}{9}n^3 + \frac{22769}{6}n^2 - 1196n + \frac{1418}{9} & n \equiv 4 \ (\mathrm{mod}\ 6) \\[2mm] 3969n^4 - \frac{53333}{9}n^3 + \frac{22769}{6}n^2 - \frac{3584}{3}n + \frac{2837}{18} & n \equiv 5 \ (\mathrm{mod}\ 6) \end{cases}$$

REFERENCES

1. M.V. Diudea and Cs. L. Nagy, "Periodic Nanostructures", Springer, Berlin, **2007**.
2. A.R. Ashrafi, B. Manoochehrian and H. Yousefi-Azari, *Utilitas Mathematica*, **2006**, *71*, 97.
3. A.R. Ashrafi, B. Manoochehrian and H. Yousefi-Azari, *Bull. Iranian Math. Soc.*, **2007**, *33* (1), 37.
4. M.V. Diudea, *Carpathian J. Math.*, **2006**, *22*, 43.
5. A.E. Vizitiu, S. Cigher, M.V. Diudea and M. S. Florescu, *MATCH Commun. Math. Comput. Chem.,* **2007**, *57*, 457.
6. M.V. Diudea, A.E. Vizitiu, F. Gholaminezhad and A.R. Ashrafi, *MATCH Commun. Math. Comput. Chem.,* **2008**, *60*, 945.
7. M.V. Diudea, *J. Math. Chem.*, **2009**, *45*, 309.
8. M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2008**, *60*, 935.
9. R. Ashrafi, M. Jalali, M. Ghorbani and M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2008**, *60*, 905.
10. R. Ashrafi, M. Ghorbani and M. Jalali, *Indian J. Chem.*, **2008**, *47A*, 538.
11. S. Klavzar, *MATCH Commun. Math. Comput. Chem.,* **2008**, *59*, 217.

OMEGA POLYNOMIAL IN CRYSTAL-LIKE NETWORKS

MAHBOUBEH SAHELI[a], MODJTABA GHORBANI[a], MONICA L. POP[b], MIRCEA V. DIUDEA[b*]

ABSTRACT. Omega polynomial $\Omega(G,x)$, defined by Diudea in *Carpath. J. Math*., 2006, *22*, 43-47, counts topologically parallel edges eventually forming a strip of adjacent faces/rings, in a graph *G=G*(*V,E*). The first and second derivatives, in *x*=1, of Omega polynomial enables the evaluation of the Cluj-Ilmenau *CI* index. Analytical close formulas for the calculation of this polynomial in two hypothetical crystal-like lattices are derived.

Keywords: Omega polynomial, crystal networks

INTRODUCTION

Design of polyhedral units, forming crystal-like lattices, is of interest in crystallography as many metallic oxides or more complex salts have found application in chemical catalysis. Various applied mathematical studies have been performed, in an effort to give new, more appropriate characterization of the world of crystals. Recent articles in crystallography promoted the idea of topological description and classification of crystal structures.[1-8] They present data on real but also hypothetical lattices designed by computer.

The geometry and polyhedral tiling is function of the experimental conditions and can be designed by dedicated software programs. Such a program, called Cage Versatile CV-NET, was developed at TOPO Group Cluj, Romania. It works by net operations, as a theoretical support.

Three basic net/map operations Leapfrog *Le*, Quadrupling *Q* and Capra *Ca*, associated or not with the more simple Medial *Med* operation, are most often used to transform small polyhedral objects (basically, the Platonic solids) into more complex units. These transforms preserve the symmetry of the parent net.[9-11]

The article is devoted to the study of two new double periodic crystal-like network, by using a topological description in terms of the Omega counting polynomial.

OMEGA POLYNOMIAL

A counting polynomial is a representation of a graph *G*(*V,E*), with the exponent *k* showing the extent of partitions *p*(*G*), $\cup p(G) = P(G)$ of a graph property *P*(*G*) while the coefficient $p(k)$ are related to the number of partitions of extent *k*.

[a] *Department of Mathematics, Faculty of Science, Shahid Rajaee Teacher Training University, Tehran, 16785 – 136, I. R. Iran; mghorbani@srttu.edu*
[b] *Faculty of Chemistry and Chemical Engineering, Babes-Bolyai University, 400028 Cluj, Romania, diudea@gmail.com*

$$P(x) = \sum_{k} p(k) \cdot x^k \tag{1}$$

Let G be a connected graph, with the vertex set $V(G)$ and edge set $E(G)$. Two edges $e=(u,v)$ and $f=(x,y)$ of G are called *codistant* (briefly: e *co* f) if the notation can be selected such that[12]

$$d(v,x) = d(v,y) + 1 = d(u,x) + 1 = d(u,y) \tag{2}$$

where d is the usual shortest-path distance function. The above relation *co* is reflexive (e *co* e) and symmetric (e *co* f) for any edge e of G but in general is not transitive.

A graph is called a *co-graph* if the relation *co* is also transitive and thus an equivalence relation.

Let $C(e) := \{f \in E(G); f\ co\ e\}$ be the set of edges in G that are codistant to $e \in E(G)$. The set $C(e)$ can be obtained by an orthogonal edge-cutting procedure: take a straight line segment, orthogonal to the edge e, and intersect it and all other edges (of a polygonal plane graph) parallel to e. The set of these intersections is called an *orthogonal cut* (*oc* for short) of G, with respect to e.

If G is a *co-graph* then its orthogonal cuts $C_1, C_2, ..., C_k$ form a partition of $E(G)$: $E(G) = C_1 \cup C_2 \cup ... \cup C_k$, $C_i \cap C_j = \varnothing, i \neq j$.

A subgraph $H \subseteq G$ is called *isometric*, if $d_H(u,v) = d_G(u,v)$, for any $(u,v) \in H$; it is *convex* if any shortest path in G between vertices of H belongs to H. The relation *co* is related to ~ (Djoković[13]) and Θ (Winkler[14]) relations.[15]

Two edges e and f of a plane graph G are in relation *opposite, e op f,* if they are opposite edges of an inner face of G. Then e *co* f holds by the assumption that faces are isometric. The relation *co* is defined in the whole graph while *op* is defined only in faces/rings.

Relation *op* will partition the edges set of G into *opposite* edge *strips ops*, as follows. (i) Any two subsequent edges of an *ops* are in *op* relation; (ii) Any three subsequent edges of such a strip belong to adjacent faces; (iii) In a plane graph, the inner dual of an *ops* is a path, an open or a closed one (however, in 3D networks, the ring/face interchanging will provide *ops* which are no more paths); (iv) The *ops* is taken as maximum possible, irrespective of the starting edge. The choice about the maximum size of face/ring, and the face/ring mode counting, will decide the length of the strip.

Also note that *ops* are *qoc* (quasi orthogonal cuts), meaning the transitivity relation is, in general, not obeyed.

The Omega polynomial[16,17] $\Omega(x)$ is defined on the ground of opposite edge strips *ops* $S_1, S_2, ..., S_k$ in the graph. Denoting by m, the number of *ops* of cardinality/length $s=|S|$, then we can write

$$\Omega(x) = \sum_{s} m \cdot x^s \tag{3}$$

The first derivative (in *x*=1) can be taken as a graph invariant or a topological index:

$$\Omega'(1) = \sum_s m \cdot s = |E(G)| \qquad (4)$$

An index, called Cluj-Ilmenau,[12] *CI(G)*, was defined on $\Omega(x)$:

$$CI(G) = \left\{ [\Omega'(1)]^2 - [\Omega'(1) + \Omega''(1)] \right\} \qquad (5)$$

In tree graphs, the Omega polynomial simply counts the non-opposite edges, being included in the term of exponent *s*=1.

Main Results

The nets herein discussed were built up by combinations of map operations.

Net A. The unit of this net is an isomer of cuboctahedron (which is the medial of Cube and Octahedron). The net is constructed by identifying some squares so that the net appears as "translated" on the Z-axis, each time one row (Figure 1)

111a   111b

333a   333b
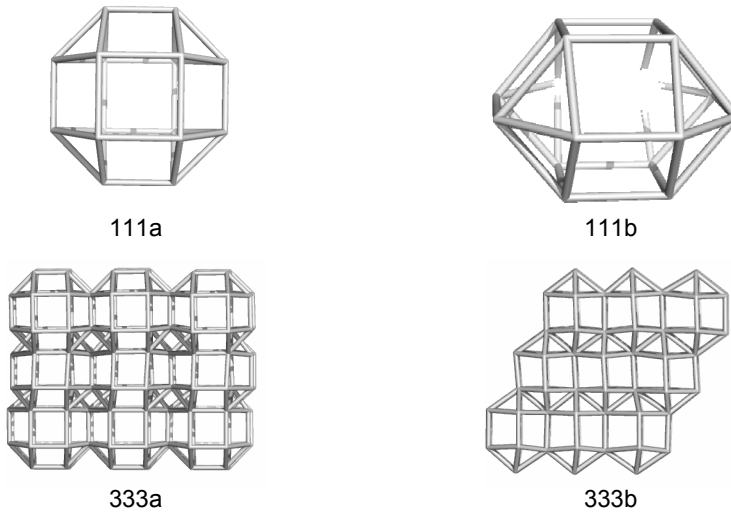
Figure 1. Net A; unit 111 (top) and 333 (bottom)

The computed data for the Omega polynomial of this net were rationalized as in the formulas presented below and Table 1.

$$a = 1 \implies \Omega(G,x) = 4x^1 + 8x^2 + 2x^6$$

$$\Omega(G,x) = 4a(2a-1)x^1 + (2a^3 + 7a^2 + 3a - 4)x^2 + ax^{2a(a+2)} + ax^{3a(a+1)} + (a-1)x^{4(a-1)(3a+2)} \qquad (6)$$

$$|E(G)| = \Omega'(G,1) = 21a^3 + 13a^2 - 2a \qquad (7)$$

$$CI(G) = 441a^6 + 389a^5 + 291a^4 - 5a^3 - 272a^2 - 8a + 80 \qquad (8)$$

Table 1. Omega polynomial and CI index of the Net A: Examples

a	Omega Polynomial	CI
1	$4x^1 + 8x^2 + 2x^6$	916
2	$24x^1 + 46x^2 + 2x^{16} + 2x^{18} + 1x^{32}$	44264
3	$60x^1 + 122x^2 + 3x^{30} + 3x^{36} + 2x^{88}$	437060
4	$112x^1 + 248x^2 + 4x^{48} + 4x^{60} + 3x^{168}$	2274544
5	$180x^1 + 436x^2 + 5x^{70} + 5x^{90} + 4x^{272}$	8280740
6	$264x^1 + 698x^2 + 6x^{96} + 6x^{126} + 5x^{400}$	23966456
7	$364x^1 + 1046x^2 + 7x^{126} + 7x^{168} + 6x^{552}$	59104804
8	$480x^1 + 1492x^2 + 8x^{160} + 8x^{216} + 7x^{728}$	129524240

Net B. The unit of this net is as for the case A but the edges sharing triangles were deleted. Moreover, the net is constructed not translated (Figure 2). Note, these networks and only double periodic, as can be seen from bottom rows of figures.

The computed data for the Omega polynomial of this net were rationalized as in the formulas presented below and Table 2.

$$a = 1 \quad \Rightarrow \quad \Omega(G, x) = 2x^6 + 2x^8$$

$$\Omega(G, x) = 4 \sum_{i=1}^{a-1} x^{(10+4(a-2))\,i} + 2x^{2a(2a+1)} + 1x^{16a^3} \tag{9}$$

$$|E(G)| = \Omega'(G,1) = 24a^3 + 4a^2 = 4a^2(6a+1) \tag{10}$$

$$|V(G)| = 8a^2(a+1) \tag{11}$$

$$\Omega''(G,1) = 256a^6 + \frac{64}{3}a^5 + \frac{64}{3}a^4 - 8a^3 + \frac{20}{3}a^2 + \frac{8}{3}a \tag{12}$$

$$CI(G) = 320a^6 + \frac{512}{3}a^5 - \frac{16}{3}a^4 - 16a^3 + \frac{32}{3}a^2 - \frac{8}{3}a =$$
$$8a\left(40a^5 + \frac{64}{3}a^4 - \frac{2}{3}a^3 - 2a^2 - \frac{4}{3}a - \frac{1}{3}\right) \tag{13}$$

Table 2. Omega polynomial and CI index of the Net B: Examples

a	Omega Polynomial	CI
1	$2x^6 + 2x^8$	584
2	$4x^{10} + 2x^{20} + 1x^{128}$	25680
3	$4x^{14} + 4x^{28} + 2x^{42} + 1x^{432}$	273784
4	$4x^{18} + 4x^{36} + 4x^{54} + 2x^{72} + 1x^{1024}$	1482912
5	$4x^{22} + 4x^{44} + 4x^{66} + 4x^{88} + 2x^{110} + 1x^{2000}$	5527720
6	$4x^{26} + 4x^{52} + 4x^{78} + 4x^{104} + 4x^{130} + 2x^{156} + 1x^{3456}$	16246256
7	$4x^{30} + 4x^{60} + 4x^{90} + 4x^{120} + 4x^{150} + 4x^{180} + 2x^{210} + 1x^{5488}$	40497240
8	$4x^{34} + 4x^{68} + 4x^{102} + 4x^{136} + 4x^{170} + 4x^{204} + 4x^{238} + 2x^{272} + 1x^{8192}$	89447744
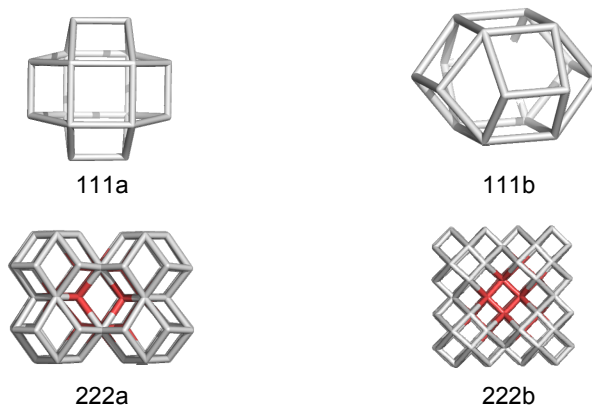
Figure 2. Net B; unit 111 (top) and 222 (bottom)

CONCLUSIONS

Omega polynomial can be used in topological description of polyhedral crystal networks.

REFERENCES

1. L. Carlucci, G. Ciani and D. Proserpio, *Coord. Chem. Rev*., **2003**, *246*, 247.
2. L. Carlucci, G. Ciani and D. Proserpio, *Cryst. Eng. Comm*., **2003**, *5*, 269.
3. V.A. Blatov, L. Carlucci, G. Ciani and D. Proserpio, *Cryst. Eng. Comm*., **2004**, *6*, 377.
4. I. Baburin, V.A. Blatov, L. Carlucci, G. Ciani and D. Proserpio, *J. Solid State Chem*., 2005, 178, 2452.
5. O. Delgado-Friedrichs and M. O'Keeffe, *J. Solid State Chem*., **2005**, *178*, 2480.
6. V.A. Blatov, O. Delgado-Friedrichs, M. O'Keeffe, and D. Proserpio, *Acta Cryst*., **2007**, A63, 418.
7. M.V. Diudea and A. Ilić, *MATCH, Commun. Math. Comput. Chem.,* **2011**, *65*, 153.
8. A.E. Vizitiu and M.V. Diudea, *Studia Univ. Babes-Bolyai Chemia*, **2009**, *54* (1), 173.
9. M.V. Diudea and P.E. John, *Commun*. *Math*. *Comput. Chem*. (*MATCH*)*,* **2001**, *44*, 103.

10. M.V. Diudea, P.E. John, A. Graovac, M. Primorac, and T. Pisanski, *Croat. Chem. Acta*, **2003**, *76*, 153.

11. M.V. Diudea, *J. Chem. Inf. Model.,* **2005**, *45,* 1002.

12. P.E. John, A.E. Vizitiu, S. Cigher, and M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2007**, *57*, 479.

13. D.Ž. Djoković, *J. Combin. Theory Ser. B*, **1973**, *14*, 263.

14. P.M. Winkler, *Discrete Appl. Math.*,**1984**, *8*, 209.

15. S. Klavžar, *MATCH Commun. Math. Comput. Chem.*, **2008**, *59*, 217.

16. M.V. Diudea, *Carpath. J. Math.*, **2006**, *22*, 43.

17. M.V. Diudea, *Nanomolecules and Nanostructures - Polynomials and Indices*, MCM, No. 10**,** Univ. Kragujevac, Serbia, **2010**.

CLUJ CJ POLYNOMIAL AND INDICES IN A DENDRITIC MOLECULAR GRAPH

MIRCEA V. DIUDEA[a], NASTARAN DOROSTI[b], ALI IRANMANESH[b,*]

ABSTRACT. The Cluj polynomials $CJ_e(x)$ and indices are calculable by either summation $CJ_eS(x)$ or multiplication $CJ_eP(x)$ of the sets of non-equidistant vertices related to the endpoints of any edge $e=(u,v)$ in the graph. A third polynomial, the (vertex) $PIv(x)$, is related to CJ_eS. In this paper, a procedure based on orthogonal cuts is used to derive the three above polynomials and indices in the molecular graph of a dendrimer.

Keywords: dendrimer, molecular graph, Cluj polynomial, Cluj index

INTRODUCTION

Cluj matrices and indices have been proposed by Diudea twelve years ago. A Cluj fragment [1-4] $CJ_{i,j,p}$ collects vertices v lying closer to i than to j, the endpoints of a path $p(i,j)$. Such a fragment collects the *vertex proximities* of i against any vertex j, joined by the path p, with the distances measured in the subgraph $D_{(G-p)}$:

$$CJ_{i,j,p} = \left\{ v \middle| v \in V(G); D_{(G-p)}(i,v) < D_{(G-p)}(j,v) \right\} \qquad (1)$$

In trees, $CJ_{i,j,p}$ denotes sets of (connected) vertices v joined with j by paths p going through i. The path $p(i,j)$ is characterized by a single endpoint, which is sufficient to calculate the unsymmetric matrix UCJ.

In graphs containing rings, the choice of the appropriate path is quite difficult, thus that path which provides the fragment of maximum cardinality is considered:

$$[\mathbf{UCJ}]_{i,j} = \max_{p} \left| CJ_{i,j,p} \right| \qquad (2)$$

When path p belongs to the set of distances DI(G), the suffix DI is added to the name of matrix, as in UCJDI. When path p belongs to the set of detours DE(G), the suffix is DE. When the matrix symbol is not followed by a suffix, it is implicitly DI. The Cluj matrices are defined in any graph and, except for some symmetric graphs, are unsymmetric and can be symmetrized by the Hadamard multiplication with their transposes[5]

$$SM_p = UM \bullet (UM)^T \qquad (3)$$

---

[a] *Faculty of Chemistry and Chemical Engineering, Babes-Bolyai University, Arany Janos 11, 400028 Cluj, Romania*
[b] *Department of Mathematics, Tarbiat Modares University. 14115-137 Tehran, IRAN,*
* *iranmanesh@modares.ac.ir*

If the matrices calculated on edges (*i.e.*, on adjacent vertex pairs) are required, the matrices calculated on paths must be multiplied by the adjacency matrix **A** (which has the non-diagonal entries of 1 if the vertices are joined by an edge and, otherwise, zero)

$$SM_e = SM_p \bullet A \qquad (4)$$

The basic properties and applications of the above matrices and derived descriptors have been presented elsewhere [6-11]. Notice that the Cluj indices, previously used in correlating studies published by TOPO GROUP Cluj, were calculated on the symmetric matrices, thus involving a multiplicative operation. Also, the symbol CJ (Cluj) is used here for the previously denoted CF (Cluj fragmental) matrices and indices.

Our interest is here related to the unsymmetric matrix defined on distances and calculated on edges **UCJ_e**

$$UCJ_e = UCJ_p \bullet A \qquad (5)$$

which provides the coefficients of the Cluj polynomials [12,13] (see below).

CLUJ POLYNOMIALS

A counting polynomial can be written in a general form as:

$$P(x) = \sum_k m(k) \cdot x^k \qquad (6)$$

It counts a graphical property, partitioned in *m* sets of extent *k*, of which re-composition will return the global property. As anticipated in introduction, the Cluj polynomials count the vertex proximity of the both ends of an edge *e*=(*u*,*v*) in *G*; there are Cluj-edge polynomials, marked by a subscript *e* (edge), to be distinguished to the Cluj-path polynomials (marked by a subscript *p*), defined on the concept of distance DI or detour DE in the graph [2,5].

The coefficients *m*(*k*) of eq. (6) can be calculated from the entries of unsymmetric Cluj matrices, as provided by the TOPOCLUJ software program [14] or other simple routines. In bipartite graphs, a simpler procedure enabling the estimation of polynomial coefficients is based on orthogonal edge-cutting. The theoretical background of the edge-cutting procedure is as follows.

A graph *G* is a *partial cube* if it is embeddable in the *n*-cube Q_n, which is the regular graph whose vertices are all binary strings of length *n*, two strings being adjacent if they differ in exactly one position.[15] The distance function in the *n*-cube is the Hamming distance. A hypercube can also be expressed as the Cartesian product: $Q_n = W_{i=1}^n K_2$.

For any edge *e*=(*u*,*v*) of a connected graph *G* let n_{uv} denote the set of vertices lying closer to *u* than to *v*: $n_{uv} = \{w \in V(G) \mid d(w,u) < d(w,v)\}$. It follows that $n_{uv} = \{w \in V(G) \mid d(w,v) = d(w,u) + 1\}$. The sets (and subgraphs) induced by these vertices, n_{uv} and n_{vu}, are called *semicubes* of *G*; the semicubes are called *opposite semicubes* and are disjoint [16,17].

A graph G is bipartite if and only if, for any edge of G, the opposite semicubes define a partition of G: $n_{uv} + n_{vu} = v = |V(G)|$. These semicubes are just the vertex proximities (see above) of (the endpoints of) edge $e=(u,v)$, which the Cluj polynomials count.

In bipartite graphs, the opposite semicubes can be estimated by an orthogonal edge-cutting procedure, as shown in Figure 1. The set of edges intersected by an orthogonal line is called an (orthogonal) cut Cn and consists of (topologically) parallel edges; the associate number counts the intersections with the orthogonal line. In the right hand part of Figure 1, there are three numbers in the front of brackets, with the meaning: (i) symmetry; (ii) occurrence (in the whole structure) and (iii) n, the number of edges cut-off by an ortogonal line. The product of the above three numbers will give the coefficients of the Cluj polynomials. The exponents in each bracket represent the number of points lying to the left and to the right of the corresponding ortogonal line segment. A similar procedure has been used by Gutman and Klavžar to calculate the Szeged index of polyhex graphs [18].
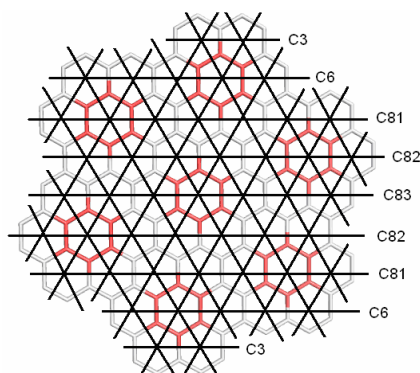
$$CJ_e\,S(x) = 3{\cdot}2{\cdot}3(x^5+x^{121})+ 3{\cdot}2{\cdot}6(x^{16}+x^{110})+$$
$$3{\cdot}2{\cdot}8(x^{31}+x^{95})+ 3{\cdot}2{\cdot}8(x^{47}+x^{79})+$$
$$3{\cdot}1{\cdot}8(x^{63}+x^{63})$$
$$CJ_e\,S'(1) = 21924;\; CJ_e\,S''(1) = 1762320$$

$$PI_v(x) = 3{\cdot}2{\cdot}3(x^{5+121})+ 3{\cdot}2{\cdot}6(x^{16+110})+$$
$$3{\cdot}2{\cdot}8(x^{31+95})+ 3{\cdot}2{\cdot}8(x^{47+79})+$$
$$3{\cdot}1{\cdot}8(x^{63+63})$$
$$PI_v'(1) = 21924;\; PI_v''(1) = 2740500$$

$$CJ_e\,P(x) = 3{\cdot}2{\cdot}3(x^{5.121})+ 3{\cdot}2{\cdot}6(x^{16.110})+$$
$$3{\cdot}2{\cdot}8(x^{31.95})+ 3{\cdot}2{\cdot}8(x^{47.79})+$$
$$3{\cdot}1{\cdot}8(x^{63.63})$$
$$CJ_e\,P'(1) = 489090$$

Figure 1. Edge-cutting procedure in the calculus of CJ polynomials of a bipartite graph

Three different counting polynomials can be defined on the vertex proximities/semicubes in bipartite graphs, which differ by the operation used in re-composing the edge contributions:

(i) *Summation*, and the polynomial is called *Cluj-Sum* (Diudea *et al.* [12,13,19,20]) and symbolized CJ_eS:

$$CJ_eS(x) = \sum_e \left(x^{n_e} + x^{v-n_e} \right) \tag{7}$$

(ii) *Pair-wise summation*, with the polynomial called (vertex) Padmakar-Ivan [21,22] (Ashrafi [23-26]) and symbolized PI_v:

$$PI_v(x) = \sum_e x^{n_e+(v-n_e)} \tag{8}$$

(iii) *Product*, while the polynomial is called *Cluj-Prod* and symbolized CJ_eP:

$$CJ_eP(x) = \sum_e x^{n_e(v-n_e)} \tag{9}$$

Because the opposite semicubes define a partition of vertices in a bipartite graph, it is easily to identify the two semicubes in the above formulas: $n_{uv}=n_e$ and $n_{vu}=v-n_e$, or vice-versa.

The first derivative (in $x=1$) of a (graph) counting polynomial provides single numers, often called topological indices.

It is not difficult to see that the first derivative (in $x=1$) of the first two polynomials gives one and the same value, however, their second derivative is different (see Figure 1) and the following relations hold in any graph [20]:

$$CJ_eS'(1) = PI_v'(1) \; ; \; CJ_eS''(1) \neq PI_v''(1) \tag{10}$$

The number of terms, $CJ_e(1)=2e$, is twice the number given by $PI_v(1)$ because, in the last case, the endpoint contributions are summed together for any edge in G (see (7) and (8)).

Clearly, the third polynomial is more different; notice that Cluj-Prod $CJ_eP(x)$ is precisely the (vertex) Szeged polynomial $Sz_v(x)$, defined by Ashrafi *et al*. [24-26] This comes out from the relations between the basic Cluj (Diudea [2,5]) and Szeged (Gutman [5,27]) indices:

$$CJ_eP'(1) = CJ_eDI(G) = Sz(G) = Sz_v'(1) \tag{11}$$

Recall the definition of the vertex PI_v index:

$$PI_v(G) = PI_v'(1) = \sum_{e=uv} n_{u,v} + n_{v,u} = |V| \cdot |E| - \sum_{e=uv} m_{u,v} \tag{12}$$

where $n_{u,v}$, $n_{v,u}$ count the non-equidistant vertices *vs*. the endpoints of $e=(u,v)$ while $m(u,v)$ is the number of vertices lying at equal distance from the vertices u and v. All the discussed polynomials and indices do not count the equidistant vertices, an idea introduced in Chemical Graph Theory by Gutman. In bipartite graphs, since there are no equidistant vertices vs any edge, the last term in (12) will disappear. The value of $PI_v(G)$ is thus maximal in bipartite graphs, among all graphs on the same number of vertices; the result of (12) can be used as a criterion for checking the "bipatity" of a graph.

APPLICATION

The three above polynomials and their indices are calculated on a dendritic molecular graph, a (bipartite) periodic structure with the repeat unit $v_0=8$ atoms, taken here both as the root and branching nodes in the design of the dendron (Figure 2, see also refs. [27-30]).

Formulas collect the contributions of the *Root*, the internal (*Int*) and external (*Ext*) parts of the structure but close formulas to calculate first derivative (in $x=1$) of polynomials were derived for the whole molecular graph. Formulas for calculating the number of vertices, in the whole wedge or in local ones, and the number of edges are also given. Examples, at the bottom of Tables 1 and 2, will enable the reader to verify the presented formulas.
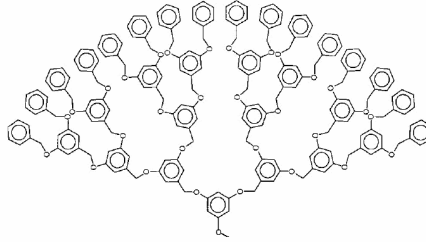
Figure 2. A dendritic wedge, of generation $r=4$; $v=248$; $e=278$.

Table 1. Formulas for counting CJ_eS and PI_v polynomials in a dendritic D wedge graph

$$CJ_eS(D,x) = CJ_eS(Root) + CJ_eS(Int) + CJ_eS(Ext)$$

$$CJ_eS(Root) = (x^1 + x^{v-1}) + (x^2 + x^{v-2}) + 1 \cdot 2 \cdot (x^5 + x^{v-5}) + 2 \cdot 2 \cdot (x^{v/2-1} + x^{v/2+1})$$

$$CJ_eS(Int) = \sum_{d=1}^{r-1} \{2^{r-d} \cdot 2 \cdot 2 \cdot [x^{v_d+3} + x^{v-(v_d+3)}] + 2^{r-(d+1)} \cdot 2 \cdot 2 \cdot [x^{v_d+1-5} + x^{v-(v_d+1-5)}] +$$

$$2^{r-(d+1)} \cdot 2 \cdot 1 \cdot \{[x^{v_d+1-2} + x^{v-(v_d+1-2)}] + [x^{v_d+1-1} + x^{v-(v_d+1-1)}] + [x^{v_d+1} + x^{v-v_d+1}]\}\}$$

$$CJ_eS(Ext) = 2^r \cdot 3 \cdot 2 \cdot (x^3 + x^{v-3}) + 2^r \cdot 1 \cdot 1 \cdot \{[x^{v_0-2} + x^{v-(v_0-2)}] +$$

$$[x^{v_0-1} + x^{v-(v_0-1)}] + [x^{v_0} + x^{v-v_0}]\}$$

$$CJ_e'S(1) = CJ_eS(D) = v \cdot (8 + \sum_{d=1}^{r-1} 18 \cdot 2^{r-(d+1)} + 9 \cdot 2^r) = v \cdot (18 \cdot 2^r - 10) = v \cdot e$$

$$v = v(D,r) = 2^3(2^{r+1}-1); \quad v_d = 2^3(2^d - 1); d = 1,2,.. \quad e(D) = 18 \cdot 2^r - 10$$

Example:
$v(r=3)=120$; $e(r=3)=134$; $v(r=4)=248$; $e(r=4)=278$

$$PI_v(x) = e \cdot x^v = (18 \cdot 2^r - 10) \cdot x^{2^3(2^{r+1}-1)}; \quad PI_v'(1) = v \cdot e$$

Example:
$CJ_eS(x,r=3)=(1x^1+1x^{119})+(1x^2+1x^{118})+(48x^3+48x^{117})+(2x^5+2x^{115})+(8x^6+8x^{114})+(8x^7+8x^{113})$
$+(8x^8+8x^{112})+(16x^{11}+16x^{109})+(8x^{19}+8x^{101})+(4x^{22}+4x^{98})+(4x^{23}+4x^{97})+(4x^{24}+4x^{96})+(8x^{27}+8x^{93})$
$+(4x^{51}+4x^{69})+(2x^{54}+2x^{66})+ (2x^{55}+2x^{65})+(2x^{56}+2x^{64})+ (4x^{59}+4x^{61})$
$CJ_eS'(1,r=3)=16080$; $CJ_eS'(1,r=4)=68944$.

Table 2. Formulas for counting CJ_eP polynomial in a dendritic D wedge graph

$CJ_eP(D,x) = CJ_eP(Root) + CJ_eP(Int) + CJ_eP(Ext)$
$CJ_eP(G) = \sum_e \left(x^{n_e(v-n_e)} \right)$
$CJ_eP(Root) = x^{1(v-1)} + x^{2(v-2)} + 1 \cdot 2 \cdot [x^{5(v-5)}] + 2 \cdot 2 \cdot [x^{(v/2-1)(v/2+1)}]$
$CJ_eP(Int) = \sum_{d=1}^{r-1} \{2^{r-d} \cdot 2 \cdot 2 \cdot [x^{(v_d+3)(v-(v_d+3))}] + 2^{r-(d+1)} \cdot 2 \cdot 2 \cdot [x^{(v_d+1-5)(v-(v_d+1-5))}] +$ $2^{r-(d+1)} \cdot 2 \cdot 1 \cdot \{[x^{(v_d+1-2)(v-(v_d+1-2))}] + [x^{(v_d+1-1)(v-(v_d+1-1))}] + [x^{(v_d+1)(v-v_d+1)}]\}\}$
$CJ_eP(Ext) = 2^r \cdot 3 \cdot 2 \cdot (x^{3(v-3)}) + 2^r \cdot 1 \cdot 1 \cdot \{[x^{(v_0-2)(v-(v_0-2))}] +$ $[x^{(v_0-1)(v-(v_0-1))}] + [x^{v_0(v-v_0)}]\}$

$$CJ_e'P(1) = CJ_eP(D) = 3626 \cdot 2^r + 256 \cdot 2^{2r} - 3872 \cdot 4^r +$$
$$+ 1792 \cdot 4^r \cdot r + 1120 \cdot 2^r \cdot r + 99$$

Example:

$CJ_eP(x, r=3) = x^{119} + x^{236} + 48x^{351} + 2x^{575} + 8x^{684} + 8x^{791} + 8x^{896} + 16x^{1199} + 8x^{1919} + 4x^{2156} + 4x^{2231} + 4x^{2304} + 8x^{2511} + 4x^{3519} + 2x^{3564} + 2x^{3575} + 2x^{3584} + 4x^{3599}$

$CJ_eP'(1, r=3) = 168627$; $CJ_eP'(1, r=4) = 1039107$.

CONCLUSIONS

Two Cluj polynomials $CJ_e(x)$ and indices, defined on vertex proximities/ semicubes, are calculable by either summation $CJ_eS(x)$ or multiplication $CJ_eP(x)$ of the sets of non-equidistant vertices related to the endpoints of any edge $e = (u, v)$ in the graph. A third polynomial, the (vertex) $PI_v(x)$, was shown to be related to the CJ_eS. A procedure based on orthogonal cuts, enabled us to derive the three above polynomials and indices in the molecular graph of a dendrimer. The procedure is applicable only in bipartite graphs.

ACKNOWLEDGEMENTS

REFERENCES

1. M.V. Diudea, *MATCH Commun. Math. Comput. Chem.,* **1997**, *35*, 169.
2. M.V. Diudea, *J. Chem. Inf. Comput. Sci*., **1997**, *37*, 300.
3. M.V. Diudea, *Croat. Chem. Acta,* **1999**, 72, 835.
4. M.V. Diudea, B. Parv, I. Gutman, *J. Chem. Inf. Comput. Sci*., **1997**, *37*, 1101.
5. M.V. Diudea, I. Gutman, L. Jäntschi, *Molecular Topology*, NOVA, New York, **2002**.
6. D. Opris, M. V. Diudea, *SAR/QSAR Environ. Res*., **2001**, *12*, 159.
7. L. Jäntschi, G. Katona, M.V. Diudea, *MATCH Commun. Math. Comput. Chem*., **2000**, *41*, 151.
8. G. Katona, G. Turcu, A.A. Kiss, O.M. Minailiuc, M.V. Diudea, *Rev. Roumaine Chim*., **2001**, *46*, 137.
9. M. Ardelean, G. Katona, I. Hopartean, M.V. Diudea, *Studia Univ. Babes-Bolyai Chemia*, **2001**, *45*, 81.
10. A.A. Kiss, G. Turcu, M.V. Diudea, *Studia Univ. Babes-Bolyai Chemia*, **2001**, *45*, 99.
11. G. Katona, M.V. Diudea, *Studia Univ. Babes-Bolyai Chemia*, **2003**, *48*, 41.
12. M.V. Diudea, A.E. Vizitiu, D. Janežič, *J. Chem. Inf. Model*., **2007**, *47*, 864.
13. M.V. Diudea, *J. Math. Chem*., **2009**, *45*, 295.

14. O. Ursu, M.V. Diudea, TOPOCLUJ software program, Babes-Bolyai University, Cluj, **2005**.
15. F. Harary, *Graph Theory*, Addison-Wesley, Reading, MA, 1969.
16. M.V. Diudea, S. Cigher, P.E. John, *MATCH Commun. Math. Comput. Chem.*, **2008**, *60*, 237.
17. M.V. Diudea, S. Klavžar, *Acta Chem. Sloven.*, **2010**, *57*, 565.
18. I. Gutman, S. Klavžar, *J. Chem. Inf. Comput. Sci.*, **1995**, *35*, 1011.
19. A.E. Vizitiu, M.V. Diudea, *Studia Univ. Babes-Bolyai Chemia*, **2009**, *54* (1), 173.
20. M.V. Diudea, A. Ilic, M. Ghorbani, A. R. Ashrafi, *Croat. Chem. Acta*, **2010**, *83*, 283.
21. P.V. Khadikar, S. Karmarkar, V.K. Agrawal, *Natl. Acad. Sci. Lett.*, **2000**, *23*, 165.
22. P.V. Khadikar, S.A. Karmarkar, *J. Chem. Inf. Comput. Sci.*, **2001**, *41*, 934.
23. M.H. Khalifeh, H. Yousefi-Azari, A.R. Ashrafi, *Discrete Appl. Math.,* **2008**, *156*, 1780.
24. M.H. Khalifeh, H. Yousefi-Azari, A.R. Ashrafi, *Linear Algebra Appl.*, **2008**, *429*, 2702.
25. A.R. Ashrafi, M. Ghorbani, M. Jalali, *J. Theor. Comput. Chem.*, **2008**, *7*, 221.
26. T. Mansour, M. Schork, *Discrete Appl. Math.*, **2009**, *157*, 1600.
27. I. Gutman, *Graph Theory Notes of New York*, **1994**, *27*, 9.
28. N. Dorosti, A. Iranmanesh, M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2009**, *62*, 389.
29. S. Hecht, J.M.J. Frechet, *Angew. Chem. Int. Ed.*, **2001**, *40*, 74.
30. M.V. Diudea, G. Katona, in: G.A. Newkome, Ed., *Advan. Dendritic Macromol.*, **1999**, *4*, 135.
31. M.V. Diudea, *Nanomolecules and Nanostructures - Polynomials and Indices*, MCM, No. 10, Univ. Kragujevac, Serbia, **2010**.

THE WIENER INDEX OF CARBON NANOJUNCTIONS

ALI REZA ASHRAFI[a], ASEFEH KARBASIOUN[b], MIRCEA V. DIUDEA[c]

ABSTRACT. Let G be a molecular graph. The Wiener index of G is defined as the sum of all distances between vertices of G. In this paper a method, which is useful to calculate the Wiener index of nanojunctions, is presented. We apply our method on the molecular graph of a carbon nanojunction $Le_{1,1}(Op(Q_{20}(T)))\_TU(3,3)$ and its Wiener index is given.

Keywords: Nanojunction, molecular graph, Wiener index.

INTRODUCTION

A molecular graph is a simple graph such that its vertices correspond to the atoms and the edges to the bonds. Note that hydrogen atoms are often omitted. By IUPAC terminology, a topological index is a numerical value associated with a chemical constitution purporting for correlation of chemical structure with various physical properties, chemical reactivity or biological activity [1−3]. This concept was first proposed by Hosoya [4] for characterizing the topological nature of a graph. Such graph invariants are usually related to the distance function d(-,-). To explain, we assume that G is a molecular graph with vertex set V(G) and edge set E(G). The mapping d(-,-): $V(G) \times V(G) \longrightarrow V(G)$ in which d(x,y) is the length of a minimum path connecting x and y, will be called "*distance function*" on G.

Recently, this part of Mathematical Chemistry was named "Metric Graph Theory". The first topological index of this type was proposed in 1947 by the chemist Harold Wiener [5]. It is defined as the sum of all distances between vertices of the graph under consideration. Suppose G is a graph with the vertex set V(G) = $\{v_1, v_2, \ldots, v_n\}$. The distance matrix of G is defined as D(G) = $[d_{ij}]$, where $d_{ij} = d(v_i, v_j)$. It is easy to see that the Wiener index of G is the half sum of entries of this matrix.

Recently many researchers were interested in the problem of computing topological indices of nanostructures. There are more than 200 published papers after 2000, but a few of them devoted to the Wiener index. On the other hand, there are not many methods to compute the Wiener index of molecular graphs and most of them are related to bipartite or planar graphs.

[a] *Department of Mathematics, Faculty of Science, University of Kashan, Kashan 87317-51167, I. R. IRAN*
[b] *Department of Mathematics, Faculty of Science, Islamic Azad University, Majlesi Branch, Majlesi, 8631656451, I. R. IRAN*
[c] *Babeş-Bolyai University, Faculty of Chemistry and Chemical Engineering, Arany Janos str 11, 400028, Cluj-Napoca, ROMANIA*

Since the molecular graphs of nanostructures are usually non-planar and most of them are not bipartite, every author applied a method designed for his/her problem.

In some research papers [7−11] one of present authors (MVD) applied some computer programs to compute the Wiener index of nanotubes and nanotori. In this method, we must decompose the problem in some cases and then prove that the Wiener index in each case is a polynomial of a given order. Finally, we compute the Wiener index in some case and find the coefficients of our polynomials. There is also a numerical method given in [12] for estimating the Wiener index.

In some papers [13−19], the authors presented a matrix method for computing Wiener index of nanotubes and nanotori. This method is appropriate for high symmetry objects and it is not general. The most general methods for computing Wiener index of nanostructures are those given in [20−22]. These methods are useful for graphs constructible by a few numbers of subgraphs. The aim of this paper is to apply the new method on the carbon nanojunction $Le_{1.1}(Op(Q_{2.0}(T)))\_ TU(3,3)$ and to compute its Wiener index.

RESULT AND DISCUSSION

Throughout this paper G[n] denotes the molecular graph of carbon nano- junction that show by $Le_{1.1}(Op(Q_{2.0}(T)))\_TU(3,3)$, Figure 1. At first, we introduce two notions. Suppose G and H are graphs such that $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. Then we call H to be a subgraph of G. H is called isometric if for each vertex x, $y \in V(H)$, $d_H(x,y) = d_G(x,y)$. In Figures 2−5, four isometric subgraphs of G[n] are depicted. Define n to be the number of rows in each arm tube (Figure 1, n=3). Then by a simple calculation, one can show that $|V(G)| = 48(n + 1)$.

To compute the Wiener index of $Le_{1.1}(Op(Q_{2.0}(T)))−TU(3,3)$, we first calculate the Wiener matrices of these subgraphs. Suppose S_1, .., S_4 are defined as follows:

- S_1 is the summation of distances between the vertices of core, Figure 2.
- S_2 is the summation of distances between vertices of a tube and the vertices of the core, Figure 3.
- S_3 is the summation of distances between two vertices of a tube, Figure 4.
- S_4 is the summation of distances between vertices of two different arm tubes, Figure 5.

We notice that the core has exactly 48 vertices and so its distance matrix is 48×48. By using HyperChem [23] and TopoCluj [24], one can see that $S_1 = 5664$. We consider the isometric subgraphs K, L and M depicted in Figures 3 to 5. To compute S_2, we consider the Figure 3. Suppose C denotes the subgraph core and D_i, $1 \le i \le n$, are the set of vertices in the ith row of a tube in G[n]. By TopoCluj, we calculate that the summation of

256

distances between vertices of the core and the set D_1 is 3480. In what follows, we obtain a recursive formula for computing S_2.

- The summation of distances between vertices of the core and the set D_1 is 3480,
- The summation of distances between vertices of the core and the set $D_1 \cup D_2$ is 3480 + 12 × 384,
- The summation of distances between vertices of the core and the set $D_1 \cup D_2 \cup D_3$ is 3480 + 12 × 384 + 12 × (384 + 96),
- The summation of distances between vertices of the core and the set $D_1 \cup D_2 \cup D_3 \cup D_4$ is 3480 + 12 × 384 + 12 × (384 + 96) + 12 × (384 + 2 × 96),
- The summation of distances between vertices of the core and the set $D_1 \cup ... \cup D_n$ is 3480 + 12 × 384 (n − 1) + 12 × 96 ×

$$\left[\frac{1}{2}\right](n-1)^2 - \frac{1}{2}n + \frac{1}{2}.$$

Therefore, $S_2 = -1128 + 4608n + 576(n-1)2 - 576n + 576$. Notice that for computing the Wiener index, we should compute $4S_2$.

We now calculate the quantity S_3. To do this, we assume that R_iR_j denote the summation of distances between vertices of D_i and D_j in subgraph L, Figure 4. For computing S_3 it is enough to compute R_iR_j, for $1 \leq i, j \leq n$. In Table 1, the occurrence of R_iR_j in S_3 is computed.

Table 1. The Number of R_iR_j in Computing S_3.

# Rows	The Number of R_iR_j
1	R_1R_1
2	$2R_1R_1 + R_1R_2$
3	$3R_1R_1 + 2R_1R_2 + R_1R_3$
⋮	⋮
N	$216n + 528(n-1) + \sum_{i=1}^{n-2}\left[792 + 288(i-1)\right]\left[n-(i+1)\right]$

From Table 1, one can compute S_3 as follows:

$$S_3 = 216n + 528(n-1) + \sum_{i=1}^{n-2}\left[792 + 288(i-1)\right]\left[n-(i+1)\right]$$
$$= 360n(n-1) + 84n + 132 - 252(n-1)^2 + 144n(n-1)^2 - 96(n-1)^3$$

Notice that in computing the Wiener index of G[n], we should consider $4S_3$, Figure 1.

To compute S_4, we assume that D_i and E_i, $1 \leq i \leq n$, denote the set of vertices in the ith row of two different arm tubes in G[n]. Using a similar argument as above, we assume that R_iS_j denote the summation of distances between vertices of D_i and E_j, $1 \leq i, j \leq n$. For computing S_4 it is enough to compute R_iS_j, for $1 \leq i, j \leq n$. In Table 2, the occurrence of R_iS_j in S_4 is computed.

Table 2. The Number of R_iS_j in Computing S_4.

# Rows	The Number of R_iR_j
1	1224
2	1224+2(1224+288)+(1224+2.288)
3	1224+2(1224+288)+3(1224+2.288)+2(1224+3.288)+(1224+4.288)
$\vdots$	$\vdots$
n	$\sum_{i=1}^{n} i(1224 + (i-1)288) + \sum_{i=1}^{n} (i-1)(1224 + (2n-i)288)$

Therefore,

$$S_4 = \sum_{i=1}^{n} i(1224 + (i-1)288) + \sum_{i=1}^{n} (i-1)(1224 + (2n-i)288)$$
$$= 1224(n+1)^2 - 1872n - 1224 - 864n(n+1) + 288n(n+1)^2.$$
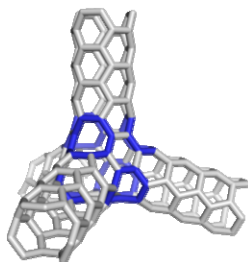
Figure 1. The Molecular Graph of $Le_{1.1}(Op(Q_{2.0}(T)))\_TU(3,3)$; n=3.
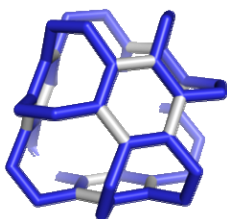
Figure 2. The Core.

Figure 3. The Subgraph K.

Figure 4. The Subgraph L.

Figure 5. The Subgraph M.

Finally, we notice that in computing the Wiener index of G[n], we should consider $\binom{4}{2}$ S$_4$ (Figure 1), thus 6S$_4$. We are now ready to state our main result.

Theorem. The Wiener index of the molecular graph of nanojunction G[n] is computed as follows:

$$W(G[n]) = 1920n^3 + 8352n^2 + 11856n + 5664.$$

Proof. By above calculations $W(G[n]) = S_1 + 4S_2 + 4S_3 + 6S_4$. Thus, a simple calculation will prove the result.
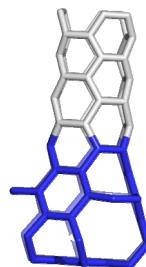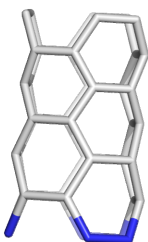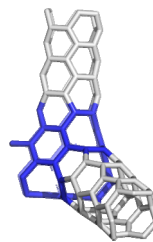
CONCLUSIONS

In this paper the Wiener index of a carbon nanojunction is computed for the first time. To the best of our knowledge it is the first paper considering the Wiener index of such nanostructures into account. A powerful method for this calculation is presented which is extendable to other nanojunctions.

REFERENCES

1. L.B. Kier and L.H. Hall, *Molecular Connectivity in Chemistry and Drug Research*, Research Studies Press, Latchworth, **1976**.
2. I. Gutman, N. Gaurilovic, D. Nankovic, P.V. Khadikar, N.V. Deshpande and P.P. Kale, *J. Serb. Chem. Soc.*, **1994**, *59*, 519.
3. P.V. Khadikar, I. Lukovits, V.K. Agrawal, S. Shrivastava, M. Jaiswal, I. Gutman, S. Karmarkar and A. Shrivastava, *Indian J. Chem.*, **2003**, *42A*, 1436.
4. H. Hosoya, *Bull. Chem. Soc. Jpn.,* **1971**, *44*, 2332.
5. H. Wiener, *J. Am. Chem. Soc.*, **1947**, *69*, 17.
6. M.V. Diudea, I. Silaghi-Dumitrescu, *MATCH Commun. Math. Comput. Chem.* **2001**, *44*, 117.
7. M.V. Diudea, P. E. John, *MATCH Commun. Math. Comput Chem.*, **2001**, *44*, 103.
8. M.V. Diudea, *Bull. Chem. Soc. Jpn.*, **2002**, *75*, 487.
9. M.V. Diudea, *MATCH Commun. Math. Comput. Chem.*, **2002**, *45*, 109.
10. P.E. John, M.V. Diudea, *Croat. Chem. Acta*, **2004**, *77*, 127.
11. M.V. Diudea, M. Stefu, B. Parv, P.E. John, *Croat. Chem. Acta*, **2004**, *77*, 111.
12. M.A. Alipour, A. R. Ashrafi, *J. Comput. Theor. Nanosci.*, **2009**, *6*, 1204.
13. S. Yousefi, A.R. Ashrafi, *MATCH Commun. Math. Comput. Chem.*, **2006**, *56*, 169.
14. S. Yousefi, A.R. Ashrafi**,** *Curr. Nanosci.*, **2008**, *4*, 161.
15. A.R. Ashrafi, S. Yousefi, *MATCH Commun. Math. Comput. Chem.*, **2007**, *57*, 403.

16. S. Yousefi, A. R. Ashrafi, *J. Math. Chem*., **2007**, *42*, 1031.

17. A.R. Ashrafi, S. Yousefi, *Nanoscale Res. Lett*., **2007**, *2*, 202.

18. S. Yousefi, H. Yousefi-Azari, A.R. Ashrafi, M.H. Khalifeh, *J. Sci. Univ. Tehran*, **2008**, *33* (3), 7.

19. S. Yousefi, A.R. Ashrafi, *Studia Univ. Babes-Bolyai*, *Chemia*, **2008**, *53* (4) 111.

20. A. Karbasioun, A.R. Ashrafi, *Macedonian J. Chem. Chem. Eng*., **2009**, *28* (1) 49.

21. A. Karbasioun, A.R. Ashrafi, M.V. Diudea, *MATCH Commun. Math. Comput. Chem*., **2010**, *63*, 239.

22. M.H. Khalifeh, H. Yousefi-Azari, A.R. Ashrafi, *Indian J. Chem*., **2008**, *47A*, 1503.

23. HyperChem package Release 7.5 for Windows, Hypercube Inc., 1115 NW 4th Street, Gainesville, Florida 32601, USA, **2002**.

24. M.V. Diudea, O. Ursu, Cs.L. Nagy, *TOPOCLUJ*, Babes-Bolyai University, Cluj, **2002**.

OBTAINING PYRAZINE-2,3-DICARBOXYLIC ACID THROUGH ELECTROCHEMICAL OXIDATION OF QUINOXALINE ON NICKEL ELECTRODE

POPA IULIANA[a*], DRAGOŞ ANA[a], VLĂTĂNESCU NANDINA[a], MARIUS DOBRESCU[a], ŢĂRANU BOGDAN[a]

ABSTRACT. The purpose of this paper is to show the studies made on the pyrazine-2,3-dicarboxylic acid (PDCA) synthesis process by quinoxaline chemical oxidation on the nickel electrode with electrochemically regenerated potassium permanganate ($KMnO_4$). It was followed the investigation of electrode reaction through cyclic voltammetry and the making of an efficient electrolyser for PDCA synthesis. Anodic regeneration of Mn^{7+} on the nickel electrode is possible. This process is favoured by KOH, Mn^{7+} (Mn^{6+} implicitly) and quinoxaline concentrations increase as well as temperature increase. Current and substance efficiencies of 80% and 85%, respectively, were achieved.

Keywords: quinoxaline, pyrazine-2,3-dicarboxylic acid, potassium permanganate, cyclic voltammetry, electrolysis, nickel electrode.

INTRODUCTION

Medical statistics show that tuberculosis is once again on the verge of becoming a threat. This is why any method for synthesizing drugs known to have antituberculosis effects must be carefully evaluated and investigated [1].

In this context pyrazine synthesis in the most advantageous conditions is of the outmost importance. The raw stock for the production of pyrazinamide is dipotassium-pyrazine-2,3-dicarboxylic acid (K_2PDCA), which can be synthesized through chemical oxidation of quinoxaline (Q) [2-4] with potassium permanganate in alkaline medium [5-7]:
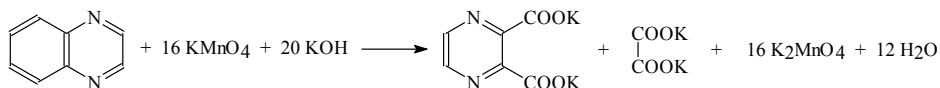
The chemical oxidation involves a very high consumption of potassium permanganate, Q: $KMnO_4$ = 1:16M (kg/kg) [8, 9]. By contrast, the original electrochemical process for PDCA synthesis proposed by us ensures considerable higher efficiencies. This paper focuses on how these efficiencies can be obtained using the perforated nickel plate electrode.

[a] *National Institute of Research-Development for Electrochemistry and Condensed Matter Timişoara, Romania 300569, Dr. A.P Podeanu, 144, Pho: 0256-222.119, Fax: 0256-201.382, *e-mail: iuliana.popa@incemc.ro*

The chemical reaction taking place in the electrochemical process is similar to that of the classical chemical process, but potassium permanganate is continuously regenerated due to the electro-oxidation of potassium manganate generated during the process. This leads to appreciable decrease of potassium permanganate consumption, the ratio of reactants being higher: Q: $KMnO_4$ = 1 – 3: 1kg/kg.

Previous studies have shown that Mn^{7+} regeneration on platinum electrode is possible both in the absence [10] and presence of quinoxaline [11]. The price of an electrolyser equipped with such an electrode is very high and finding a cheaper material for manufacturing of the anode, while maintaining the platinum performance, constitutes a strong issue for the process at hand.

This paper shows the results obtained through cyclic voltammetry in the study of the Mn^{6+}/Mn^{7+} couple behaviour on the nickel electrode as well as the manufacturing of the laboratory electrolyser made with perforated nickel plate electrode for PDCA synthesis using electrochemically regenerated potassium permanganate as chemical reagent.

RESULTS AND DISCUSSION

The Mn^{7+}/Mn^{6+} redox couple behaviour in alkaline medium was studied through cyclic voltammetry. The curves obtained using the nickel anode in 4M KOH solution in the presence of manganese ions at various concentrations, are shown in figure 1.
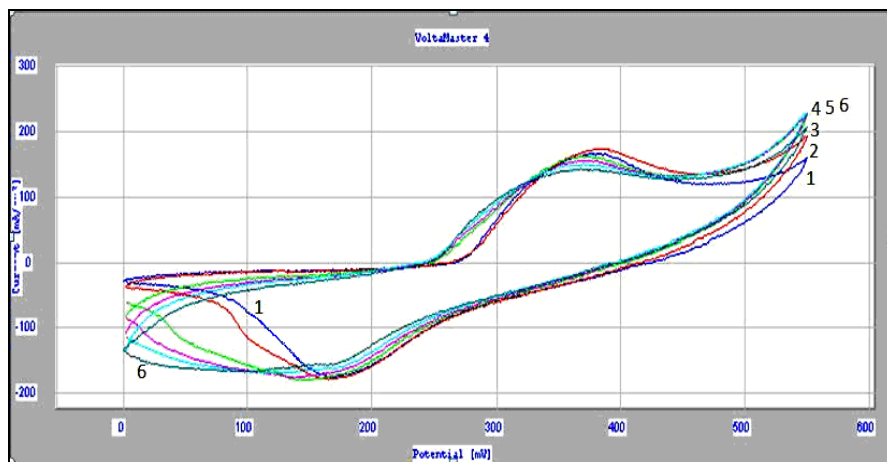
Figure 1. Cyclic voltammograms at different concentration (M) of Mn^{6+}: 0(1); $0.4 \cdot 10^{-3}$(2); $2 \cdot 10^{-3}$(3); $4 \cdot 10^{-3}$(4); $8 \cdot 10^{-3}$(5); $16 \cdot 10^{-3}$(6); [KOH] = 4M; 25ºC; v = 100 mV/s.

Cycle 1 (blue) – generated in the absence of Mn^{6+} ions – shows an anodic peak at ~ 0.38V and a cathodic peak at ~ 0.16V. The presence of the two peaks is due to the reversible process:

262

$$Ni^{2+} + e \leftrightarrow Ni^{3+}$$
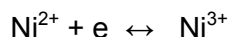
When metallic Ni is sunk in a NaOH solution, it gets covered with a $Ni(OH)_2$ monomolecular layer. During anodic polarization, the Ni^{2+} thus formed is converted in Ni^{3+} (NiOOH). The process is reversible and during cathodic polarization $Ni(OH)_2$ is obtained once again.

Increasing Mn^{6+} concentration (cycles 2-6) leads to a depolarization of the oxygen release and at the same time there is an observed decrease and slight displacement of the anodic peak towards more negative potentials. Another tendency towards more negative potentials is observed at the cathodic peak, starting at 0.16V. Besides this cathodic peak – present due to a reduction in Ni^{3+} – at the 0.100V potential a wave appears and increases. At an increase in Mn^{6+} ion concentration, the wave tends significantly towards more negative potentials and current intensity increases. The wave seems to appear as a result of the reduction in Mn^{7+} ions formed during the anodic process.

It's possible that Mn^{6+} oxidation on the nickel electrode takes place at the same time with oxygen release.

During the process the color of the electrolyte solution turns from green to violet.

An increase in temperature from 20 to 45ºC determines the increase of the peak currents. The anodic peak currents as functions of temperature and supporting electrolyte concentration for $4.10^{-3}M$ K_2MnO_4 are shown in figure 2.
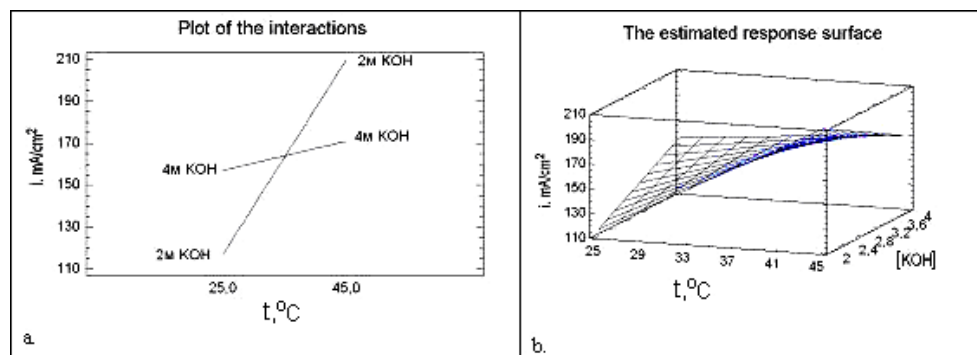
Figure 2. Anodic current density variation with temperature and supporting electrolyte concentration for $4.10^{-3}M$ K_2MnO_4.

The regression equation is:

$$i = -138 + 8.6 \cdot t + 69.375 \, [KOH] - 1.975 \cdot t \, [KOH]$$

The cyclic voltammograms corresponding to different quinoxaline concentrations, obtained at 45ºC in 2M KOH solution with $6 \cdot 10^{-2}M$ $KMnO_4$, are shown in figure 3. Cycle 1, obtained in the absence of Q, the peak pair due to the $Ni^{2+} + e^- \leftrightarrow Ni^{3+}$ balance can be observed. The presence of

quinoxaline in the electrolyte solution leads to the disappearance of the Ni^{2+} oxidation peak and to the appearance of a new anodic peak – at 0.30V. The intensity of this peak increases with the Q concentration. The peak potential moves slowly towards more positive values as the Q concentration increases. Two peaks appear on the cathodic branch. The cathodic peak present at 0.15V – it's associated with Ni^{3+} reduction – increases with the Q concentration and moves towards more negative potentials. Next to this peak there is another cathodic peak, present at a potential of 0.07V. This peak also increases with Q concentration and also tents towards more negative potentials. The peak appears in the same area where Mn^{7+} ions reduction takes place.

It seems that Q oxidation can be achieved in two ways: direct oxidation on the electrode (at ~ 0.3V) and mediated oxidation with electrochemically regenerated Mn^{7+}. The later takes place simultaneously with oxygen generation.
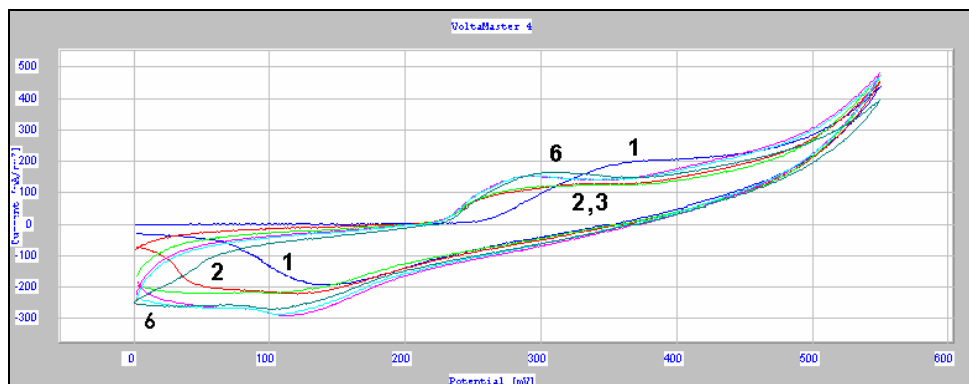
Figure 3. Cyclic voltammograms obtained for different [Q]: 0(1); $0.4 \cdot 10^{-3}$(2); $2 \cdot 10^{-3}$(3) ; $4 \cdot 10^{-3}$(4) ; $8 \cdot 10^{-3}$(5) and $16 \cdot 10^{-3}$(6) M ; t = 45°C; 2M KOH solution, [$KMnO_4$] = $6 \cdot 10^{-2}$M, v = 100 mV/s.

Cyclic voltammetry studies performed at 25°C and 45°C respectively, in KOH solution containing Mn^{6+} ions, on the nickel electrode show that:

- there are several processes taking place on nickel electrode:

$$Ni^{2+} \quad \leftrightarrow \quad Ni^{3+} + e^-$$
$$Mn^{7+} + e^- \quad \leftrightarrow \quad Mn^{6+}$$
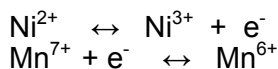
- the addition of Q in the electrolyte solution leads to disappearance of the Ni^{2+} anodic oxidation peak and to the appearance of a new anodic peak, at 0.30V. This peak is associated with Q oxidation. The cathodic peaks don't change when temperature is increased.

- the increase in KOH concentration, temperature, Mn^{6+} and Q concentrations favours the anodic regeneration process of Mn^{7+}.

264

The experimental results obtained on the nickel plate anode syntheses are shown in table 2, where: Q_{el} – electricity quantity; U_{med} - cell tension; m_{Qi} – initial quantity of quinoxaline; m_{Qf} – final quantity of quinoxaline; Conv.- conversion of quinoxaline and m_{K2PDCA} - K_2PDCA quantity.

Table 2. The experimental results for the PDCA synthesis on the nickel electrode with electrochemically regenerated $KMnO_4$.

I [A]	i [A/m²]	Q_{el} [C]	U_{med} [V]	T [°C]	m_{Qi} [g]	m_{Qf} [g]	Conv [%]	m_{K2PDC} [g]	η_s [%]	η_c [%]	C. En. KWh/Kg
1.8	3.5	20000	3.5	45	2	0.3	85	2.5	66.6	79.10	7.78
1.8	3.5	40000	3.5	45	2	0	100	3.2	85.2	50.62	12.15
0.9	1.7	30000	2.1	45	2	0	100	2.8	74.6	59.06	6.25
2.7	5.3	40000	3.8	45	2	0	100	2.5	66.6	39.55	16.89

- KOH concentration – 23 %
- quinoxaline concentration – 1,4 – 2,8 %
- Mn^{7+} concentration – 1,4 %.

The best results for the current yield η_c are achieved at a current density of 3.5 A/dm². Lower current densities lead to a higher current efficiency and a lower cell tension U_{med} – thus to a desirable lower specific energy consumption C.En. = 6.25 KWh/Kg. On the other hand there is a high increase in reaction time and thus a decrease in electrolyser productivity. At higher current densities (5.3 A/dm²) the substance η_s and current efficiencies are acceptable, but energy consumption increases significantly. However, in this case, the electrolyser productivity is higher.

CONCLUSIONS

From our studies (we did not identify any similar data in the scientific literature) the Mn^{7+} regeneration takes place on the nickel electrode even at low current densities. Current efficiencies of ~ 80% have been achieved at ~ 85% conversions and substance efficiencies of ~ 85% have been achieved at 100% conversions and a current efficiency of ~ 50%. Nickel constitutes a very good material for manufacturing the anode of a $KMnO_4$ regeneration electrolyser used for quinoxaline oxidation.

EXPERIMENTAL SECTION

Electrochemical cell - Cyclic voltammetry method

For the cyclic voltammetry studies we used a glass electrolysis cell (figure 4) equipped with a heating/cooling jacket and with three electrodes: the working electrode made from a nickel wire (0,008 cm²), the platinum counter electrode (1 cm²) and the SCE reference electrode. A PGZ 301 Dynamic-

EIS Voltammetry potentiostat with VoltaMaster 4 software manufactured by Radiometer Copenhagen was also used in these studies. All electrochemical potentials mentioned in this paper are related to the SCE electrode unless otherwise specified.
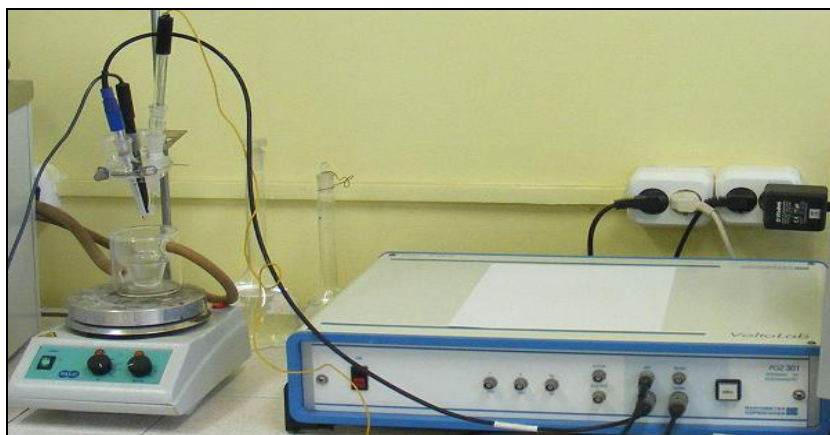
Figure 4. Electrochemical installation.

Electrolyte solution: 2 - 4M KOH (electrolyte support), K_2MnO_4 $0.4 \cdot 10^{-3}$ - $16 \cdot 10^{-3}$ M; $KMnO_4$ $2 \cdot 10^{-2}$ - $6 \cdot 10^{-2}$ M, quinoxaline $1.18 \cdot 10^{-2}$ – $3.62 \cdot 10^{-2}$ M. We used two temperatures: 25 and $45^{\circ}C$. The quinoxaline was from Merck, $KMnO_4$ from Riedel-de Haen and KOH, from Lach-Ner.

The method for synthesizing potassium manganate is as follows: an alkaline aqueous solution of 8N KOH containing 10g of potassium permanganate was heated at a temperature of 120°C. After the color changed from violet (Mn^{7+}) to intense green (Mn^{6+}) the supersaturated solution of Mn^{6+} was obtained. K_2MnO_4 crystals were filtered from this solution on a *S4* frit, washed with $CHCl_3$, dried and weighed, and then directly dissolved in 8N KOH solutions (25 ml measuring flask) and used in cyclic voltammetry tests.

Laboratory electrolyser

The perforated plate electrolyser had an electrolyser with a volume of ~ 100 ml. The perforated plate cathode and anode are shown in figures 5. The nickel anode underwent nitric acid pickling before each synthesis and between two syntheses it was washed with a mixture of sulphuric and oxalic acids.

The general characteristics of the electrolyser and the working conditions are the following:

- Anodic surface, cm^2 S_A = 0.51
- Cathodic surface, cm^2 S_C = 0.034
- Sa/Sc – ratio - 15
- Electrolyte volume, ml – 90
- Current density, mA/cm^2 – 1.7 – 5.3
- Working temperature, ^oC – 45
- Total volume of the electrolyser, ml – 150
- anodic material – nickel perforated plate
- cathodic material – stainless steel
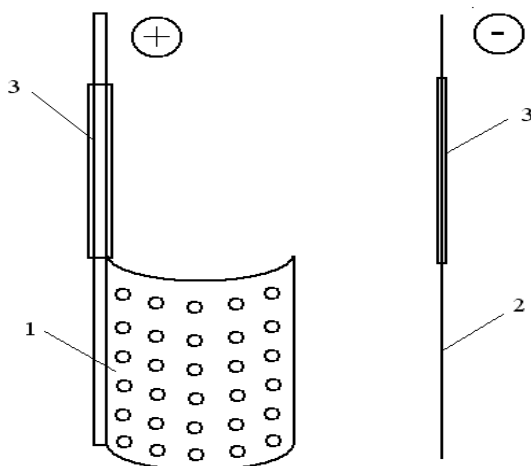
Figure 5. Components of the electrolyser used for the preliminary results of $KMnO_4$ regeneration. 1 - anode; 2 – cathode; 3 – insulating tube.

ACKNOWLEDGMENTS

REFERENCES

1. W. Wang, J.W. Lown, *J. Med. Chem*., **1992**, *35(15)*, 2890.

2. G. Reuben, G. Jones, K.C. McLaughlin, *Org. Syn.*, *Coll. Vol. 4*, **1963**, 824.

3. D. Arndt, *"*Manganese Compounds as Oxidizing Agents in Organic Chemistry"*, Open Court Publishing Co.: La Salle, IL, **1981**, 254.

4. J. March, "Advanced Organic Chemistry – Reactions, Mechanisms and Structure", 3rd Edition, *Wiley*, **1984**, 1048.

5. C.A. Obafemi, W. Pfleiderer*, Helv. Chim. Acta*, **1994**, *72,* 1549.

6. H. Lund, *J. Electrochem. Soc.*, **2002**, *149 (4),* S21.

7. S.A. Kotharkar; D.B. Shinde, *J. Iran. Chem. Soc.*, **2006**, *Vol.3, No 3,* 267.

8. C.A. Obafemi, W. Pfleiderer, *Helv. Chim. Acta,* **2004**, *77 (6),* 1549.

9. G. Reuben, G. Jones, K.C. McLaughlin, *Org. Syn.*, *Coll.* **1950**, *Vol. 30*, 86.

10. I. Popa, A. Dragos, I. Taranu, M. Stefanuţ, C. Vaszilcsin, D. Buzatu, *Bull. Politehnica, Ser. Chem. Environ. Eng.,* **2007**, *Vol. 52 (66), 1-2,* 99.

11. A. Dragoş, I. Popa, I. Taranu, "*New Trends and Strategies in the Chemistry of Advanced Materials with Relevance in Biological Systems, Technique and Environmental Protection",* Timişoara, 8 - 9 November **2007**, 28.

SCHULTZ, MODIFIED SCHULTZ AND SZEGED INDICES OF A FAMILIY OF FULLERENES

ALI IRANMANESH[a,b], YASER ALIZADEH, SAMANE MIRZAIE

ABSTRACT. Let G be a simple connected graph. Schultz and modified Schultz indices are defined as:

$$S(G) = \sum_{\{u,v\} \subseteq V(G)} (\delta_u + \delta_v) d(u,v); \; MS(G) = \sum_{\{u,v\} \subseteq V(G)} (\delta_u \delta_v) d(u,v),$$ where δ_u is

the degree of vertex u and $d(u,v)$ is the distance between u and v. Let e be an edge of a graph G connecting the vertices u and v. Define two sets $N_1(e|G)$ and $N_2(e|G)$ as follows:

$N_1(e|G) = \{x \in V(G) | d(x,u) < d(x,v)\}$ and

$N_2(e|G) = \{x \in V(G) | d(x,v) < d(x,u)\}$. The number of elements of

$N_1(e|G)$ and $N_2(e|G)$ are denoted by $n_1(e|G)$ and $n_2(e|G)$ respectively. Szeged index of G is defined as:

$$Sz(G) = \sum_{e \in E(G)} n_1(e|G).n_2(e|G).$$

In this paper we give a GAP program for computing the Schultz, the Modified Schultz and the Szeged indices of a simple connected graph. Also we compute and formulate these indices for a family of fullerenes by the software GAP and MAPLE.

Keywords: *Schultz index, Modified Schultz index, Szeged index, $C_{12(n-1)}$ fullerenes.*

INTRODUCTION

A topological index is a numerical quantity that is mathematically derived in a direct and unambiguous manner from the structural graph of a molecule. Let G be a simple connected graph, the vertex and edge sets of G being denoted by *V(G)* and *E(G),* respectively. The distance between two vertices *u* and *v* of G is denoted by *d(u,v)* and it is defined as the number of edges in a shortest path connecting *u* and *v*. Diameter of G is denoted by *d*. Distance is an important concept in graph theory and it has applications in computer science, chemistry, and a variety of other fields. Topological indices based on the distances in graph, like Wiener index [1], are widely used for

---

[a] *Department of Mathematics, Faculty of Mathematical Sciences, Tarbiat Modares University, P.O. Box: 14115-137, Tehran, Iran. E-mail: iranmanesh@modares.ac.ir*
[b] *Corresponding author Email: iranmanesh@modares.ac.ir*

establishing relationships between the structure of a molecular graph and its physicochemical properties.

In connection with certain investigations in mathematical chemistry, Schultz [2] considered a graph invariant that he called "molecular topological index" and denoted by MTI. It is defined as:

$$MTI(G) = \sum_{i=1}^{n}\sum_{j=1}^{n}(A_{ij} + D_{ij})\delta_i$$

where δ_i is the degree of vertex i in G and A_{ij} and D_{ij} are elements of the adjacency matrix and distance matrix of G respectively.

The essential part of MTI is the *Schultz index* $S(G)$ [3]:

$$S(G) = \sum_{\{u,v\}\subseteq V(G)}(\delta_u + \delta_v)d(u,v)$$

where δ_u is degree of vertex u and $d(u,v)$ denote the distance between vertices u and v.

Klavzar and Gutman in [4] defined a modified Schultz index as:

$$MS(G) = \sum_{\{u,v\}\subseteq V(G)}(\delta_u\delta_v)d(u,v)$$

Let e be an edge of a graph G connecting the vertices u and v. Define two sets $N_1(e|G)$ and $N_2(e|G)$ as follows:

$$N_1(e|G) = \{x \in V(G)|d(x,u) < d(x,v)\} \text{ and}$$

$$N_2(e|G) = \{x \in V(G)|d(x,v) < d(x,u)\}.$$

The number of elements of $N_1(e|G)$ and $N_2(e|G)$ are denoted by $n_1(e|G)$ and $n_2(e|G)$ respectively.

The Szeged index *Sz* was introduced by Gutman [5] and is defined as:

$$Sz(G) = \sum_{e \in E(G)} n_1(e|G).n_2(e|G).$$

Schultz, Modified Schultz and Szeged indices of the following nanotubes and fullerenes are computed: C_{60} fullerene [6], $HAC_5 C_7$ [p, q] [7], TUC_4C_8[p,q] [8,9], VC_5C_7[p,q] nanotube [10], $HAC_5C_6C_7$[p,q] [11], dendrimer nanostars [12], HC_5C_7[r, p] [13], zigzag nanotube [14].

In this paper, we give a GAP program for computing the Schultz, Modified Schultz and Szeged indices of $C_{12(n-1)}$ fullerenes.

RESULTS AND DISCUSSION

The Schultz, modified Schultz and Szeged indices are topological indices based on distances in a graph. To obtain these indices, it needs to compute the degree of the vertices and the distance between the vertices. The

set of vertices having their distance to the vertex u equal to t is denoted by $D_{u,t}$ and the set of vertices adjacent to vertex u is denoted by $N(u)$.

Let $e=uv$ be an edge connecting the vertices u and v, then we have the following result:

$$V(G) = \bigcup_{t \geq 0}^{d} D_t(u), \quad \forall u \in V(G)$$

$$S(G) = \sum_{\{u,v\} \subseteq V(G)} (\delta_u + \delta_v) d(u,v)$$

$$= \sum_t \sum_{u \in V(G)} \sum_{v \in D_{u,t}} (\delta_u + \delta_v) t$$

$$MS(G) = \sum_{\{u,v\} \subseteq V(G)} (\delta_u \delta_v) d(u,v)$$

$$= \sum_t \sum_{u \in V(G)} \sum_{v \in D_{u,t}} (\delta_u \delta_v) t$$

$$(D_t(u) \setminus D_t(v)) \subseteq (D_{t-1}(v) \bigcup D_{t+1}(v)), \ t \geq 1.$$

$$(D_t(u) \bigcap D_{t-1}(v)) \subseteq N_2(e|G) \quad and \quad D_t(u) \bigcap D_{t+1}(v) \subseteq N_1(e|G), t \geq 1$$

$$(D_1(u) \bigcup \{u\}) \setminus (D_1(v) \bigcup \{v\}) \subseteq N_1(e|G) \quad and \quad (D_1(v) \bigcup \{v\}) \setminus (D_1(u) \bigcup \{u\}) \subseteq N_2(e|G).$$

By using the following relations, we can determine the sets $D_{u,t}$.

$$D_{u,1} = N(u),$$

$$D_{u,t+1} = \bigcup_{j \in D_{u,t}} (N(j) \setminus (D_{u,t} \bigcup D_{u,t-1})), \ t \geq 1.$$

According to the above relations, by determining the sets $D_{u,t}$, we can compute the Schultz, the modified Schultz and the Szeged indices of a graph.

The fullerene is a hollow, pure carbon molecule in which the atoms lie at the vertices of a polyhedron with 12 pentagonal faces and any number (other than one) of hexagonal faces. The fullerenes discovered in 1985 by researchers at Rice University, are a family of carbon allotropes named after Buckminster Fuller. Spherical fullerenes are sometimes called buckyballs. A family of Fullerene is $C_{12(n-1)}$ (n denote the number of layers) Figure 1.
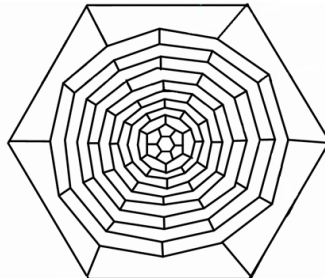
Figure 1. $C_{12(n-1)}$ fullerene (n=11).

By using the following GAP program, we can compute the Shultz, the Modified Schulttz and the Szeged indices of a graph. Input of the program is the set of adjacent vertices.

```
D:=[];  deg:=[];  e:=[];
for i in [1..n] do  D[i]:=[]; D[i][1]:=N[i]; deg[i]:=Size(N[i]);
u:=Union(u,D[i][1]);   r:=1; t:=1; u:=[i];
while r<>0 do   D[i][t+1]:=[];
for j in D[i][t] do
for m in Difference (N[j],u) do
AddSet(D[i][t+1],m);
od;
od;
u:=Union(u,D[i][t+1]);
  if D[i][t+1]=[] then r:=0;
fi;
t:=t+1;
od;
od;
A:=[]; T:=[];S:=0;  MS:=0;  sz:=0;
for i in [1..n] do
for t in [1..Size(D[i])] do
for j in D[i][t] do   S:=S+(deg[i]+deg[j])*t;
MS:=MS+(deg[i]*deg[j])*t);
od;   od;od;
for i in [1..n-1] doN1:=[];
for j in Difference(N[i],T) do   N2:=[];
N1[j]:=Union(Difference(N[i],Union([j],N[j])),[i]);
N2[i]:=Union(Difference(N[j],Union([i],N[i])),[j]);
for t in [2..Size(D[i])-1] do
for x in Difference(D[i][t],Union(D[j][t],[j])) do
if not x in D[j][t-1] then    AddSet(N1[j],x);
elif x  in D[j][t-1] then    AddSet(N2[i],x);
fi;
od;
od;
sz:=sz+Size(N1[j])*Size(N2[i]);
od;
Add(T,i);
od;
 S:=S/2; #(The value of S is equal to the Schultz index of the graph)
 MS:=MS/2; #(The value of MS is equal to the Modified Schultz index of the graph)
sz; #(The value of sz is equal to szeged index of the graph
```

In Table 1, the Shultz, the Modified Schulttz and the Szeged indices of $C_{12(n-1)}$ fullerene for some n are computed by the above GAP software.

Table 1. The Shultz, the Modified Schulttz and the Szeged indices of $C_{12(n-1)}$ fullerene.

n	Number of vertices	Schultz index	Modified Schultz index	Szeged index
5	48	28728	43092	29508
6	60	51084	76626	59616
7	72	82224	123336	104052
8	84	123768	185652	166236
9	96	177408	266112	247248
10	108	244872	367308	349332
11	120	327888	491832	474912
12	132	428184	642276	626460
13	144	547488	821232	806532
20	228	2059992	3089988	3083100
24	276	3608568	5412852	5410428
30	348	7158312	10737468	10741740
36	420	12512088	18768132	18779100
41	480	18620928	27931392	27947940
57	672	50845824	76268736	76303140
69	816	90884448	136326672	136374468
76	900	121862808	182794212	182849820
82	972	153447624	230171436	230233740
95	1128	239665392	359498088	359574900
100	1188	279928152	419892228	419974620

In the following, the formulas of these indices are obtained by the software Maple.

The Schultz index of $C_{12(n-1)}$ fullerene is :

$$S(C_{12(n-1)}) = 288n^3 - 864n^2 + 5832n - 15048 \quad n \geq 8$$

The Modified Schultz of $C_{12(n-1)}$ fullerene is:

$$M\,S(C_{12(n-1)}) = 432n^3 - 1296n^2 + 8748n - 22572 \quad n \geq 8,$$

The Szeged indices of $C_{12(n-1)}$ fullerene is:

$$Sz(C_{12(n-1)}) = 432n^3 - 1296n^2 + 9864n - 51780 \quad n \geq 12.$$

CONCLUSIONS

In this paper, a GAP program for computing the Schultz and Modified Schultz indices of a simple connected graph is presented. Input of the program is the set of adjacent vertices of the graph. The formulas for these indices in $C_{12(n-1)}$ fullerenes are derived by Maple software and examples are computed by the software GAP.

REFERENCES

1. H. Wiener, *J. Am. Chem. Soc*., **1947**, *69*, 17.
2. H.P. Schultz, *J. Chem. Inf. Comput. Sci*., **1989**, *34*, 227.
3. H.P. Schultz, T.P. Schultz, *J. Chem. Inf. Comput. Sci.,* **1993**, *33*, 240.
4. S. Klavzar and I. Gutman, *Disc. Appl. Math*., **1997**, *80*, 73.
5. P.P. Khadikar, N.V. Deshpande, P.P. Kale, A.A. Dobrynin, I. Gutman, G. Domotor, *J. Chem. Inf. Comput. Sci.,* **1997**, *35*, 545.
6. Y. Alizadeh, A. Iranmanesh and S. Mirzaie, *Dig. J. Nanomater. Biostruct.,* **2009**, *4*, 7.
7. A.Iranmanesh, Y. Alizadeh, *Am. J. Appl. Sci.,* **2008**, *5,* 1754.
8. A. Heydari, B. Taeri, *J. Comp. Theor. NanoSci.,* **2007**, *4*, 158.
9. A. Heydari, B. Taeri, *MATCH Commun. Math. Comput. Chem.,* **2007**, *57*, 665.
10. A. Iranmanesh, Y. Alizadeh, *Int. J. Mol. Sci.,* **2008**, *9*, 131.
11. A. Iranmanesh, Y. Alizadeh, *Dig. J. Nanomater. Biostruct.,* **2009**, *4*, 67.
12. A. Iranmanesh, N. Gholami, *Croat. Chem. Acta,* **2008**, *81*, 299.
13. A. Iranmanesh, A. Mahmiani, and Y. Pakravesh, *Szeged index of $HC_5C_7[r, p]$ Nanotubes. ArsCombinatorics,* **89** (**2008**), 309.
14. E. Eliasi and B. Taeri, *MATCH Commun. Math. Comput. Chem*., **2006**, *56*, 383.

# EXPRESSION OF PHENYLALANINE AMMONIA-LYASES IN *ESCHERICHIA COLI* STRAINS

KLAUDIA KOVÁCS[a, b], ANDRÁS HOLCZINGER[c], BEÁTA VÉRTESSY[b,c], LÁSZLÓ POPPE[a*]

ABSTRACT. Phenylalanine ammonia-lyase (PAL, EC 4.3.1.5) catalyzes the non-oxidative conversion of L-phenylalanine into (*E*)-cinnamate. PAL can be applied in organic synthesis, and can be considered also for enzyme supplementation cure for genetic disorder phenylketonuria. The aim of this study was to find optimal expression parameters of several previously cloned PAL's (bacterial, plant and a chimera) in pBAD vectors for further functional characterization. Investigation of the expression level of PAL's in *E. coli* hosts with SDS PAGE analysis as well as activity assay of the recombinant PAL enzymes was performed.

Keywords: phenylalanine ammonia-lyase, heterologous protein expression, biocatalysis, phenylketonuria

INTRODUCTION

Phenylalanine ammonia-lyase (PAL) catalyzes the non-oxidative conversion of L-phenylalanine into (*E*)-cinnamate (*Figure 1*) [1].
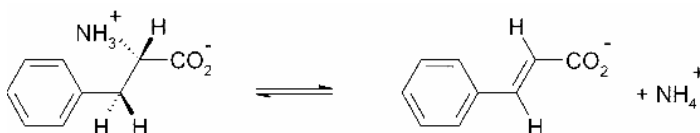
Figure 1. Non-oxidative deamination of phenylalanine

PAL is an important enzyme in both plant development and pathogen defense. In all plants PAL is encoded by a multi-gene family, ranging in copy number from four in *Arabidopsis* to a dozen or more copies in some higher plants [2]. The PAL participates in five metabolic pathways: tyrosine, phenylalanine and nitrogen metabolism, phenylpropanoid and alkaloid biosynthesis. Because of its key role between the primary and secondary metabolism in plants, PAL is a potential target for herbicides.

[a] *Budapest University of Technology and Economics, Department of Organic Chemistry and Technology, and Research Group for Alkaloid Chemistry of HAS, Műegyetem rkp 3, H-1111, Budapest, Hungary, poppe@mail.bme.hu*
[b] *Institute of Enzymology of the Hungarian Academy of Sciences, Karolina út 29, H-1113, Budapest, Hungary*
[c] *Budapest University of Technology and Economics, Department of Applied Biotechnology and Food Science, Műegyetem rkp 3, H-1111, Budapest, Hungary*

By reversing the natural direction of the PAL reaction in the presence of high concentration of ammonia, optically pure L-phenylalanine, which is the precursor molecule of the artificial sweetener aspartame (L-phenylalanyl-L-aspartyl methyl ester) can be produced. Similarly, starting from various (hetero)arylacrylates, further enantiopure L-phenylalanine analogues, such as L-piridil/pirimidil-alanines can be also prepared. Since neither cofactor recycling nor other additives are needed in these asymmetric syntheses, they are potentially interesting as industrial processes as well [3a-e].

In addition to its application in synthetic chemistry, PAL can be applied also in human medicine as treatment to avoid the effects (mental retardation, neurotoxic effects) of the most common congenital metabolic disease, phenylketonuria (PKU). Daily oral administration of micro-encapsulated PAL to PKU rats decreased the systemic toxic phenylalanine level by 75 ± 8% in 7 days (P < 0.001) [4]. Because of its many scopes, PAL is an extensively studied enzyme.

Although PAL is an ubiquitous higher-plant enzyme, it has only been encountered in a few bacteria, where cinnamic acid is involved in biosynthesis of several specific bacterial products, such as oligosaccharide antibiotics [5a-c].

The major difference between eukaryotic and prokaryotic PAL's is a ca. 150-residue long C-terminal extension of the eukaryotic PAL's (per monomeric unit) which is not present in the prokaryotic PAL's (*Figure 2*) [6]. In plant and
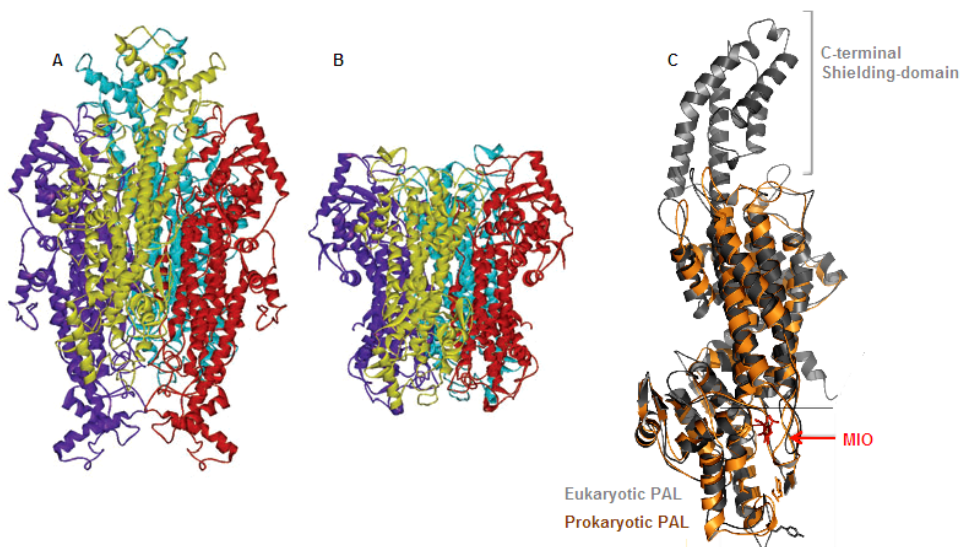
Figure 2. Tetrameric structures of PALs

(A) Homology model of *P. crispum* PAL [7]; (B) Homology model of *P. luminescens* PAL [6];
(C) Difference between eukaryotic and prokaryotic PAL monomers

fungal PAL's,the additional C-terminal domain forms an arch over the active site and has been proposed to function as a shielding domain by restricting substrate entry and product egress. Alternatively, this domain may influence the conformation of an active-site lid loop and thereby affect the stability and catalytic activity of the holoenzyme. Molecular dynamics studies confirmed the hypothesis that the C-terminal extension decreases the lifetime of eukaryotic PAL by destabilization, which might be important for the rapid responses in the regulation of phenylpropanoid biosynthesis [6].

One of the PAL enzymes expressed in this study was an artificial chimera [L. Poppe, A. Holczinger, *unpublished*] composed of the C-terminal domain of a bacterial PAL from *Photorhabdus luminescens* and of the catalytic N-terminal domain of the plant PAL from *Petroselinum crispum*. The main goal with the expression and thermostability investigations of this chimera PAL is to prove the hypothesis on the destabilizing role of the C-terminal extension of the eukaryotic PAL's.

RESULTS AND DISCUSSION

Expressions of *Petroselinum crispum* PAL and the *Photorhabdus luminescens* PAL were used as references. The catalytic N-terminal segment of the *Photorhabdus luminescens* PAL was expressed also, to compare the activity of this segment to native enzyme. In addition to the *Photorhabdus luminescens* PAL, an other PAL (hereafter HA1) gene of a bacterium growing at relatively high temperature was expressed as well.

To express the PAL genes from different organisms (bacteria, plant and a chimera), pBAD-24 and pBAD-HisB vector systems were investigated in two *E. coli* strains (Rosetta (DE3) and TOP 10) as hosts (*Table 1.*).

Table 1. The expressed genes and the applied vectors

Gene	Vector	Main objectives
Petroselinum crispum PAL (*Pc*PAL); ~2150 bp	pT7-7 (2470 bp)	Eukaryotic reference; optimized system
N-terminal segment of *Photorhabdus luminescens* PAL (PhN); ~1420 bp	pBAD-24 (4542 bp)	Study the catalytic activity of the N-terminal segment.
Chimera (CHI): N-terminal of *Pc*PAL and C-terminal of *Photorhabdus luminescens* PAL; ~1650 bp	pBAD-24 (4542 bp)	Activity and thermostability of the artificial enzyme
Photorhabdus luminescens PAL (Phl6); ~ 1600 bp	pBAD-24 (4542 bp)	Bacterial reference; known, characterized bacterial PAL
PAL from a thermophilic bacterium (HA1); ~ 1700 bp	pBAD-HisB (4092 bp.)	Activity and thermostability of the PAL from a thermophile.

The pBAD vectors containing the PAL genes with relatively weak promoter were used for expression work. The advantage of the the *AraC*- pBAD expression system is that in the presence of L-arabinose the expression from the promoter is turned on, while in the absence of L-arabinose very low level of transcription from pBAD promoter can occur. The uninduced level is repressed

even further by growth in the presence of glucose. By varying the concentration of L-arabinose, protein expression levels can be optimized to ensure maximum expression of soluble protein. In addition, the tight regulation of pBAD by AraC is really effective to minimize the leakiness of the promoter.

For cloning and transformation, a *rec*A, *end*A, *ara*BADC- and *ara*EFGH+ TOP10 strain was used, which is capable of transporting L-arabinose, but not metabolizing it, and Rosetta (DE3), which doesn't have the *ara*BADC- mutation to prevent arabinose degradation. To the latter strain, the same amount of inductor was re-added at 4 h after the first induction to maintain full induction.

Our aim was to optimize the conditions of the expression with the pBAD vector constructions in the two *E. coli* hosts to achieve expression levels of the PAL enzymes which are satifactory for further biochemical and biocatalytic characterizations. The effects of the expression conditions (temperature, expression time, inductor concentration) on the protein expression level and the activity and thermostability of the enzymes have been studied. Thus, the temperature of expression was varied between 18 ºC to 37 ºC in 3-5 ºC steps, the time of expression was changed between 8-22 h in 2-4 h steps. The inductor concentration was increased gradually from 0,002 to 0,02 %.

After cell disruption by sonication, the PAL activity of the crude extract was determined by measurement PAL by monitoring the formation of (*E*)-cinnamate at 290 nm ($\varepsilon_{290}= 10^4$ L M^{-1} cm^{-1} [3e]). At 290 nm, absorbance of aromatic amino acid residues of the proteins, nucleic acids and denaturation of proteins can influence the measurement, therefore the absorbance values were always corrected with the blind values from determinationd with same amount of substrate-free buffer and supernatant without substrate. The PAL content of the crude extract was confirmed also by SDS-PAGE investigation of samples from various fractions (supernatant, pellet etc.).

Increasing the expression temperature and shortening the expression time had favorable effect on expression levels of all bacterial expressions. The highest expression levels were found at the maximum concentration (0.02 %) of the arabinose inductor. In all cases, expressions with TOP 10 strain (*Figure 3a*) resulted higher level expression than with Rosetta (DE3) strain (*Figure 3b*). Unfortunately, no expression was detected with the chimera PAL in the pBAD systems in our hands (*Figure 3a*).

The optimal parameters of the expressions were determined by considering the SDS-PAGE and activity (U_s: Units/L of crude extract) results together, compared to the expression level and the activity of the *Petroselinum crispum* PAL expression as reference (*Table* 2.). Although the bacterial PAL's were expressed at low level according to the SDS-PAGE, the activites of the expressed bacterial PAL's in the crude extracts (3.3−7.5 U L^{-1}) were comparable to the PAL activity of the crude extract of *Pc*PAL (8.5 U L^{-1}) (*Table 2*). As the purification method for *Pc*PAL at this level of expression is well established [3e,7], expression of the bacterial PAL's at comparable levels followed by further purification steps will be enough to investigate the thermostability (*Figure 3a*) and biotransformation properties of the prokariotic PAL's.
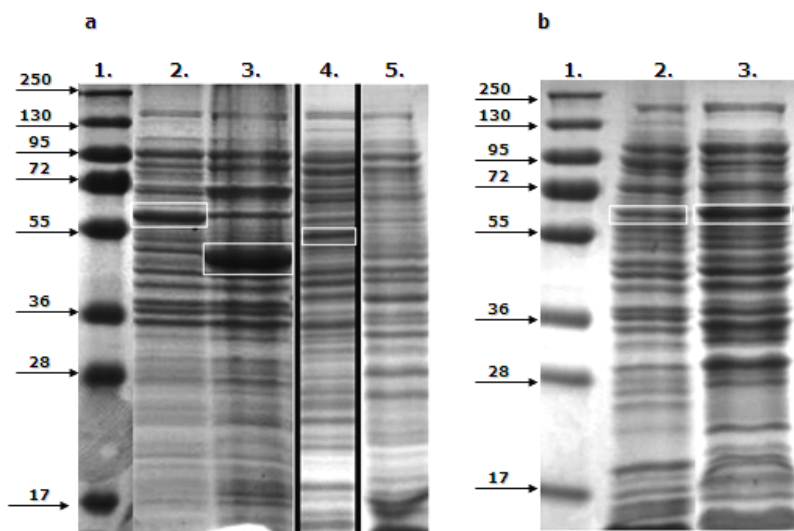
Figure 3. SDS-PAGE on PAL expressions in E. coli TOP 10 and Rosetta strains

(a) SDS-PAGE of the PAL expression in *E. coli* TOP 10 [*Lane 1*: molecular mass markers (250, 130, 95, 72, 55, 36, 28,17 kDa); *Lane 2*: supernatant after cell lysis of the expression of HA1 (58,4 kDa); *Lane 3*: supernatant after cell lysis of the expression of phN; (51,17 kDa); *Lane 4*: supernatant after cell lysis of the expression of Phl6; (57,57 kDa); *Lane 5*: supernatant after cell lysis of the expression of Chi (~58 kDa)];

(b) SDS-PAGE: the different between the expression of the two E. coli strains [*Lane 1:* molecular mass markers (250, 130, 95, 72, 55, 36, 28,17 kDa); *Lane 2:* expression of HA1 (58,4 kDa) in Rosetta (DE3) strain; *Lane 3:* expression of HA1 (58,4 kDa) in TOP 10 strain].

Table 2. The optimal expression conditions of PALs in pBAD vectors and activity of the crude extracts

Expressed enzyme	Inductor	Temperature (°C)	Time (h)	U_s (U/L)
P. crispum PAL [8]	1 mM IPTG	20	20	8.5
P. luminescens PAL	0,2 % arabinose	30	12	7.5
N-terminal segment of *P. luminescens* PAL	0,2 % arabinose	25	18	6.2
HA1 bacterial PAL	0,2 % arabinose	25	16	3.3

To achieve higher PAL yield, coexpression with pREP4-groESL chaperon plasmid was investigated as well (*Figure 4.*).

Analysis of the expression level of the bacterial PAL's with this coexpression system proved to be difficult, because the size of groEL chaperon (~65 kDa) is similar to the bacterial enzymes (~60 kDa). Thus, appraise the level of PAL expression by SDS-PAGE failed (*Figure 4.*). Because the level of chaperon expression seems to exceed the level of PAL expression, construction of a vector system carrying the PAL and chaperon genes under the contol of the same promoter is considered.
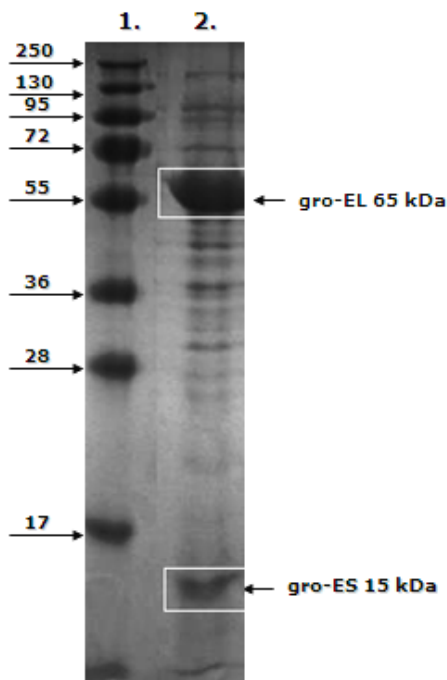
Figure 4. The gro-EL chaperone coexpression with bacterial PAL's (~60 kDa)

Lane 1: molecular mass markers (250, 130, 95, 72, 55, 36, 28, 17 kDa);
Lane 2: supernatant after cell lysis of the reported expression of HA1 (58,4 kDa) in *E. coli* TOP 10
    (coexpression of the pREP4-groESL chaperon plasmid)

CONCLUSIONS

In this study, expression of several bacterial phenylalanine ammonia-lyases in *Eschericia coli* TOP 10 using pBAD vactor was investigated. The activity assay of the expressed bacterial PAL's using the supernatant of the *E. coli* lysate indicated catalytic activity and thus the presence of active soluble enzymes. The expression levels of the bacterial PAL's were adequate for further investigations. The highest protein expression level by SDS-PAGE was found in the expression of *N*-terminal segment of *Photorhabdus luminescens* PAL, whereas the highest crude supernatant activity was achieved with the native *Photorhabdus luminescens* PAL. Due to the non-synchronized expression and overproduction of the chaperon protein in the pBAD-PAL / pREP4-groESL chaperon coexpression system, cloning of the PAL genes without and with groESL chaperon into pET vector is considered.

EXPERIMENTAL SECTION

Transformation and expression of PAL genes

E. coli TOP 10 and Rosetta (DE3) strains were transformed with different plasmids containing PAL genes (see *Table 1*.) and grown in 50 mL low salt LB broth/medium with 100 µg/mL ampicillin [and 35 µg/mL chlorampehnicol only for Rosetta (DE3) strain] overnight 37°C with shaking (220 rpm) until OD600= ~1.5. From the resulting culture, 800 µL volume was added to 50 mL LB (containing 100 µg/mL ampicillin and 35 µg/mL chlorampehnicol). The resulting cultures were grown at 37°C with vigorous shaking until OD600 = ~0.6. At this OD, the temperature was changed to the induction temperature (in a range of 18 ºC − 37 ºC; 4 ºC steps for screening) and the cells were induced with arabinose (in a range of 0,02-0,2 %; 0.05 % steps for screening). At 4 h after the first induction, the same amount of inductor was added to Rosetta cell cultures. After addition of the inducer, the cultures were further cultivated for 12 h and harvested by centrifugation. All subsequent operations were carried out at 4 ºC.

Cell disruption and activity measurement

The pellet was resuspended in 5 mL of lysis buffer (150 mM NaCl, 50 mM TRIS pH 8.0, 10 mM BME Protease inhibitor cocktail, 2 mM PMSF, 5 mM BA) and sonicated in ice bath at amplitude 40 % and pulsation 60 % using a Bandelin Sonopuls HD 2070 instrument. The sonication was performed until the viscosity of the suspension was significantly lowered. After centrifugation (30 min at 5000 x g), PAL activity was determined in the crude extract by monitoring the formation of (*E*)-cinnamate at 290 nm (ε_{290} =10^4 L M^{-1} cm^{-1}). The assay contained performed at room temperature by addition of 100 µL of the supernatant from the crude extract to 1000 µL of 0.1 M Tris-HCl, pH 8.8, containing 20 mM L-phenylalanine and recording the absorbance at 290 nm for 10 min. The PAL content of the crude extract, supernatant and pellet were analyzed by SDS-PAGE as well.

REFERENCES

1. H. Ritter, G.E. Schulz, *Plant Cell*, **2004**, *16*(12)*, 3426.*

2. A. Chang, M.H. Lim, S.W. Lee, E.J. Robb, R.N. Nazar, *J. Biol. Chem.*, **2008**, *283*(48), 33591.

3. (a) S. Yamada, K. Nabe, N. Izuo, K. Nakamichi, I. Chibata, *Appl. Environ. Microbiol.*, **1981**, *42*, 773. (b) C.T. Evans, K. Hanna, C. Payne, D. Conrad, M. Misawa, *Enz. Microb. Tech.,***1987**, *9*, 417. (c) M. Yanaka, D. Ura, A. Takahashi, N. Fukuhara, *Japan Patent,* **1994**, 06-113 870. (d) A. Gloge, B. Langer, L. Poppe, J. Rétey, *Arch. Biochem. Biophys.,***1998**, *359*, 1. (e) A. Gloge, J. Zon, Á. Kővári, L. Poppe, J. Rétey, *Chem. - Eur. J.*, **2000**, *6*, 3386.

4. L. Bourget, T.M. Chang, *FEBS Lett.*, **1985**, *180*, 5.

5. (a) A.V. Emes, L.C. Vining, *Can. J. Biochem.,* **1970**, *48*(5), 613. (b) L. Xiang, B.S. Moore, *J. Bact.*, **2005**, *187*(12), 4286. (c) J.S. Williams, M. Thomas, D.J. Clarke, *Microbiol.*, **2005**, *151*, 2543.

6. S. Pilbák, A. Tomin, J. Rétey, L. Poppe, *FEBS J.*, **2006**, *273*(6), 1004.

7. D. Röther, L. Poppe, G. Morlock, S. Viergutz, J. Rétey, *Eur. J. Biochem.*, **2002**, *269*(12), 3065.

8. M. Baedeker, G.E. Schulz, *FEBS Lett.*, **1999**, *457*, 57.

NOVEL SOLID SUPPORTS FOR LIPASES IN SOL-GEL IMMOBILIZATION SYSTEMS

DIÁNA WEISER[a], ANNA TOMIN[a], LÁSZLÓ POPPE[a,*]

ABSTRACT. Sol-gel encapsulation of lipases proved to be a particularly easy and effective way to enhance the mechanical and catalytic properties of biocatalysts. The sol-gel encapsulated enzymes usually retain their selectivity whereas their heat stability or specific activity may be significantly improved. The aim of our work was to improve the immobilization of lipases in supported sol-gel systems. First, the binding properties of lipase AK on various solid supports were studied. Next, the immobilization properties of the best adsorbent-lipase combinations were tested in sol-gel encapsulation using tetraethoxy-silane/octyltriethoxy-silane/phenyltriethoxy-silane 1/0.7/0.3 silane precursor system.

Keywords: lipase, biocatalysis, adsorption, supported sol-gel immobilization

INTRODUCTION

Lipases (EC 3.1.1.3) are extensively utilized biocatalysts in organic chemistry [1,2]. Lipases are essential in the digestion, transport and processing of lipids (e.g. triglycerides, fats, oils) in most, if not all, living organisms. However, lipases are also being exploited as inexpensive and easy-to-use biocatalysts in more modern applications [3,4,5]. For instance, lipases are used in applications such as baking and laundry detergents and even as biocatalysts in alternative energy strategies to convert vegetable oil into biofuel [6]. Lipases are flexible biocatalysts which can catalyze a wide range of regio- and enantioselective reactions such as hydrolysis, esterifications, transesterifications, aminolysis and ammoniolysis [1,7,8]. These reactions usually proceed with high regio- and/or enantioselectivity, therefore lipases became indispensable biocatalysts in various biotransformations. Development of efficient/economical biotransformations often requires robust technologies for immobilization of biomolecules or microorganisms. Immobilization of enzymes can enhance their activity, thermal and operational stability, and reusability which is important for industrial applications [9,10]. Among many available immobilization methods, including adsorption, covalent attachment to solid supports and entrapment within polymers [9,10], entrapment of enzymes in inorganic/organic hybrid polymer matrices has received a lot of attention in recent years and has provided new possibilities in the field of material science [11,12]. We report here the binding properties of lipases on various solid carriers and further

[a] *Budapest University of Technology and Economics, Department of Organic Chemistry and Technology, and Research Group for Alkaloid Chemistry of Hungarian Academy of Sciences, Műegyetem rkp 3, H-1111, Budapest, Hungary, * poppe@mail.bme.hu*

immobilization of the adsorbed lipases in hydrophobic sol-gel materials, which results in the formation of highly active, stable and reusable heterogeneous biocatalysts.

RESULTS AND DISCUSSION

Adsorption of the lipases at the large specific surface area of porous supports can avoid the aggregation of proteins and thus can result in an increased activity of the biocatalysts. In our study the adsorption behavior of four different lipases – from *Pseudomonas fluorescence* (lipase AK), from *Burkholderia cepacia* (lipase PS), from *Candida cylindracea* (lipase CcL) and BUTE-3 [13] – on various solid supports were investigated. Ten kinds of carriers with different porosity were examined. Most of them were different types of silica gel, but Celite 545 and Filtracel-950 were also studied.

Our recent study indicated that better sol-gel immobilization results can be achieved with lipases adsorbed previously on solid support than with simultaneous addition of the lipase and the supporting material to the silane precursor system [14]. Therefore, our further aim was to use the best supporting materials showing the most pronounced enhancement in the enzyme activity in combined sol-gel encapsulation as well. To test the biocatalytic properties of the resulting biocatalysts, the kinetic resolution of 1-phenylethanol with vinyl acetate in hexan/THF 2:1 was used as model reaction. The immobilization efficiency was characterized by several parameters such as specific biocatalyst activity (U_b), specific enzyme activity (U_e), enantiomer selectivity (E) and conversion (c).
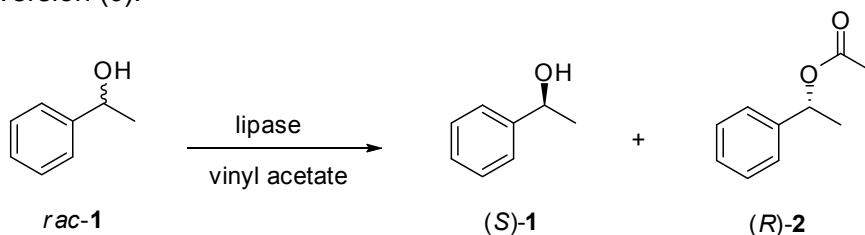
Figure 1. Lipase-catalyzed kinetic resolution of racemic 1-phenylethanol

Analysis of the above parameters for the free and adsorbed lipases indicated that in many cases the biocatalytic properties of enzymes adsorbed on solid support were superior compared to the native lipase (Table 1.).

Because the enantiomeric excess (*ee*) of the forming acetate (*R*)-**2** alone is not characteristic for the selectivity, the degree of enantiomer selectivity was characterized by the *E* value calculated from the conversion (c) and enantiomeric excess of the forming acetate (*R*)-**2** ($ee_{(R)\text{-}2}$) [15]. The activity yield [Y_A (%)] can be calculated from the effective specific activity of the immobilized biocatalyst ($U_{e,imm\text{-}LAK}$) compared to the effective specific activity of the free Lipase AK ($U_{e,LAK}$) [14].

Table 1. Behavior of free and adsorbed lipases in kinetic resolution of racemic 1-phenylethanol rac-**1** with vinyl acetate.

Lipase	Adsorbent[a]	c[b] %	E[b]	U_e (μmol min^{-1} g^{-1})	Y_A[c] %
Lipase AK[d]	-	49	>100	10	100
Lipase AK	Filtracel-950	22	>200	82	818
Lipase AK	Davisil 150	18	>100	69	687
Lipase AK	Grace 915	22	72	85	848
Lipase AK	Grace 920	30	>200	113	1124
Lipase AK	PQ 300	22	>200	85	847
Lipase PS[d]	-	29	»200	14	14
Lipase PS	Filtracel-950	48	>200	104	725
CcL[d]	-	7.6	3	2.6	100
CcL	Geduran Si 60	0.9	2,9	3.3	127
BUTE-3[d]	-	50	»200	17	100
BUTE-3	Grace 920	7.9	62.1	30	174

[a] The enzyme / adsorbent mass ratio was 1 / 10;
[b] Results after 4 h reaction time. The enantiomer selectivity (E) was calculated from c and $ee_{(R)-2}$ [15] and $ee_{(S)-1}$ and $ee_{(R)-2}$ [16] simultaneously. Due to sensitivity to experimental errors, E values calculated in the 100-200 range are reported as >100, values in the 200-500 range are reported as >200 and values calculated above 500 are given as »200;
[c] Activity yield; [d] Free lipase without immobilization

By entrapment the lipases adsorbed on a large surface in a sol-gel matrix the diffusional limitations can be decreased leading to immobilized lipases with enhanced catalytic properties. The study of adsorption of the lipase AK on four silica supports revealed that the activity yield (Y_A) could be increased more than 800 % compared to specific activity of the free enzyme. The reason of this large effect can be the large surface area of the silica supports of pore diameters between 60 and 150 Å. This pore size range allows the adsorption of the enzyme inside the pores and thus formation of a thin layer on the total surface of the catalyst with practically no diffusion limitations for the substrate to reach and the product to leave from the catalyst. The same effect can be observed with lipase PS using Filtracel-950. On the other hand, the activity yields (Y_A) for lipase CcL and BUTE-3 were not enhanced significantly by the adsorptive immobilization.

As the activity yield (Y_A) enhancement by adsorption was most pronounced with lipase AK, further investigation was performed with this enzyme using the four most effective solid supports in a combined adsorption / sol-gel immobilization process (Table 2.). In this study, the preadsorption of Lipase AK on silica-gels was followed by sol-gel immobilization using the tetraethoxy-silane/octyltriethoxy-silane/phenyltriethoxy-silane precursors in 1/0.7/0.3 molar ratio which was found optimal in our previous study with Celite as solid support [17]. The best results were achieved with the combined sol-gel entrapment using the Grace 915 support, because the activity yield (Y_A) enhancement (Table 2) was as high as for the simple adsorption without sol-gel (Table 1).

Besides the reusability of the sol-gel immobilized lipases, the formation of sol-gel polymer matrix can increase significantly the stability of the biocatalysts. The long-term stability test of entrapped lipases indicated that full activity of the

sol-gel biocatalysts was retained after storage in refrigerator or at room temperature (1 day, 1 week, 1 month tests) [18]. The full activity was also maintained for our novel sol-gel lipases after 1 month storage at room temperature.

Table 2. Behavior of sol-gel immobilized supported Lipase AK in kinetic resolution of racemic 1-phenylethanol rac-**1** with vinyl acetate.

Support[a]	c[b] %	E[b]	U_e (μmol min^{-1} g^{-1})	Y_A[c] %
-	49	>100	10	100
Grace 920	3,9	38	16	156
Grace 915	16	34	87	868
Davisil 150	8.2	75	40	400
PQ 300	5.1	50	22	218

[a] Lipase AK adsorbed on solid support (Table 1) was added to the tetraethoxy-silane/octyltriethoxy-silane/phenyltriethoxy-silane 1/0.7/0.3 silane precursor system during sol-gel matrix formation;
[b] Results after 4 h reaction time. The enantiomer selectivity (E) was calculated from c and $ee_{(R)\text{-}\mathbf{2}}$ [15] and $ee_{(S)\text{-}\mathbf{1}}$ and $ee_{(R)\text{-}\mathbf{2}}$ [16]. simultaneously. Due to sensitivity to experimental errors, E values calculated in the 100-200 range are reported as >100;
[c] Activity yield

CONCLUSION

The activity and stability of lipases had been increased significantly by applying commercially available solid silica supports of wide pores. The adsorptive or the combined sol-gel immobilization did not influence the selectivities of lipases, whereas the robust combined adsorption / sol-gel encapsulation resulted in biocatalysts which are reusable and thus applicable in various synthetic processes.

EXPERIMENTAL SECTION

Chemicals and enzymes

Lipase AK (lipase from *Pseudomonas fluorescens*), Lipase PS (lipase from *Burkholderia cepacia*), CcL (lipase from *Candida cylindracea*), and Davisil 150 were obtained from Sigma-Aldrich. 2-Propanol (IPA), vinylacetate and sodium fluoride (NaF) were products of Aldrich. 1- Phenylethanol, 2-heptanol, polyethylenglycol 1000 (PEG), Celite® 545, tetraethoxy-silane and phenyltriethoxy-silane were obtained from Fluka. Octyltriethoxy-silane were obtained from Alfa Aesar. Grace 920, Grace 915 were obtained from Grace. PQ 300 was obtained from PQ Corporation. Filtracel EFC-950 was a product of Rettenmaier and Söhne GMBH. Geduran® Si 60 was obtained from Merck KGaA. The lipase BUTE-3 was obtained as described earlier [13].

Preadsorption of lipases on solid adsorbent

The lipase powder (50 mg) was suspended in TRIS-HCl buffer (0.1 M, pH 7.5, 25 mL) at room temperature. The solid support (500 mg) was added to the solution. After and stirring (at 800 rpm for 15 min), the resulting suspension was kept at 4°C for 24 h. After filtration, the residual solid was washed with buffer (25 mL), dried at room temperature at air overnight and finally dried in vacuum exicator for 3 h.

Immobilization of lipases in sol-gel systems

The solution of TRIS-HCl buffer (0.1 M, pH 7.5, 390 μL), polyethylene glycol (PEG, 4% w/v, 200 μL), aqueous sodium fluoride (NaF, 1M, 100 μL) and 2-propanol (IPA, 200 μL) were shaken at room temperature for 10 minutes in a glass of 20 ml vial. Then the silane precursors [tetraethoxy-silane, octyltriethoxy-silane, phenyltriethoxy-silane precursors in 1/0.7/0.3 molar ratio; total 780 μmol] were added to the aqueous solution and the resulting two-phase emulsion was shaken for further 5 minutes until gelation. The lipase powder (22.7 mg free or 250 mg preadsorbed lipase) was added to the gel at intensive shaking. To complete the polymerization, the mixture was shaken for 12 h at room temperature. The forming fine, white powder was washed by 2-propanol (7 ml), distilled water (5 ml), 2-propanol (5 ml) and n-hexane (5 ml). The immobilized biocatalysts were dried in a vacuum exicator for 5 h then stored in air at room temperature. The immobilization efficiency was calculated on the basis of the enzyme supplied to the immobilization process.

Esterification assay

To a solution of racemic 1-phenylethanol (rac-**1**, 50 mg, mmol) in hexane/THF 2/1 (1 mL) and vinyl acetate (100 μL), biocatalyst (50 mg) was added and the mixture was shaken in a sealed glass vial at 1000 rpm at room temperature. For GC analyses, samples were taken directly from the reaction mixture (sample size: 10 μL, diluted with CH_2Cl_2 to 100 μL) at 2,4,8 and 24 h. Data on conversion and enantiomeric selectivity of the process with various enzymes are presented in Tables 1 and 2.

Gas chromatographic analysis of the products

The products of the kinetic resolutions with the various lipase biocatalysts [(R)-**2** and (S)-**1**] were analyzed by gas chromatography on Acme 6100, equipped with flame ionization detector and Hydrodex β-6TBDM [30 m × 0.25 mm × 0.25 μm film of heptakis-(2,3-di-O-methyl-6-O-t-butyldimethyl-silyl)-β-cyclodextrin] column. The oven temperature, injector and detector temperatures were 135, 250 and 250 °C, respectively. Hydrogen was used as carrier gas at constant flow (1.8 mL/min).

REFERENCES

1. U.T. Bornschauer, R.J. Kazlauskas, "Hydrolases in Organic Synthesis: Regio- and Stereoselective Biotransformations", Wiley-VCH, Weinheim-New York, **2006**.
2. R. Gupta, N. Gupta, P. Rathi, *Appl. Microb. Biotech.,* **2004**, *64*, 763.
3. R.D. Schmid, R. Verger, *Angew. Chem. Int. Ed.,* **1998**, *37*, 1608.
4. K.E. Jaeger, M.T. Reetz, *Tr. Biotech.,* **1998**, *16*, 396.
5. K.E. Jaeger, T. Eggert, *Curr. Opin. Chem. Biol.,* **2002**, *13*, 390.
6. M.G. Kulkarni, A.K. Dalai, *Ind. Eng. Chem. Res.,* **2006**, *45*, 2901.
7. M.T. Reetz, *Curr. Opin. Chem. Biol.,* **2002**, *6*, 145.
8. A. Liese, K. Seelbvach, C. Wandrey, „Industrial Biotransformations", 2nd Edition Wiley-VCH, Weinheim, **2006**.
9. R.A. Sheldon, *Adv. Synth. Catal.,* **2007**, *349*, 1289.
10. C. Mateo, J.M. Palomo, G. Fernandez-Lorente, J.M. Guisan, R. Fernandez-Lafuente, *Enzym. Microb. Tech.,* **2007**, *40*,1451.
11. J. Kim, J.W. Grate, P. Wang, *Chem. Eng. Sci.,* **2006**, *61*, 1017.
12. D. Avnir, T. Coradin, O. Lev, J. Livage, *J. Mat. Chem.,* **2006**, *16*, 1013.
13. V. Bódai, R. Peredi, J. Bálint, G. Egri, L. Novák, G. Szakács, L. Poppe, *Adv. Synth. Catal.,* **2003**, *345*, 811.
14. A. Tomin, D. Weiser, Z. Bata, L. Corici, F. Péter, L. Poppe, *Studia Univ.Babes-Bolyai Chemia*, **2009**, *54* (Sp2), 77.
15. C.S. Chen, Y. Fujimoto, G. Girdaukas, C.J. Sih, *J. Am. Chem. Soc.,* **1982**, *104*, 7294.
16. J.L.L. Rakels, A.J.J. Straathof, J.J. Heijnen, *Enzym. Microb. Tech.,* **1993**, *15*, 1051.
17. A. Tomin, D. Weiser, G. Hellner, Z. Bata, L. Corici, F. Péter, B. Koczka, L. Poppe, *Proc. Biochem.*, **2010**, *46*, 52.
18. T. Reetz, A. Zonta, *Angew. Chem. Int. Ed.,* **1995**, *34*, 301.

SYNTHESIS AND LIPASE CATALYSED KINETIC RESOLUTION OF RACEMIC AMINES

PÉTER FALUS[a], ZOLTÁN BOROS[a], GÁBOR HORNYÁNSZKY[a], JÓZSEF NAGY[a], LÁSZLÓ ÜRGE[b], FERENC DARVAS[c], LÁSZLÓ POPPE[a]*

ABSTRACT. Feasibility of production of amines from ketones employing metal and/or metal-catalysts in one-pot and one-step reductive amination (modified Leuckart- and Leuckart-Wallach- reaction) and lipase catalysed kinetic resolution of racemic amines in batch and continuous-flow reactor were investigated. In kinetic resolutions the effect of the solvent, the acetylating agent and the lipase itself was examined.

Keywords: *reductive amination, metal-catalysis, lipase, continuous-flow reactor, kinetic resolution*

INTRODUCTION

Enantiomerically pure chiral amines are valuable building blocks of quite a number of drugs [1], pesticides [2] and colour pigments [3]. Considerable amount of drugs are amines or amine derivatives.

Biotechnology and biocatalysis are increasingly employed to produce optically active intermediates of drugs [4]. Hydrolases can be efficiently used for synthetic biotransformations due to its favourable characteristics [5, 6, 7, 8]. Hydrolases can catalyze several related reactions such as hydrolysis, condensations, alcoholysis and aminolysis. Lipases are proved to be highly versatile biocatalyst in stereoselective biotransformations such as kinetic resolutions [9], deracemisations and dynamic kinetic resolutions [10]. Enantioselective enzymatic reactions are typically carried out in batch mode [5, 9, 11, 12], however a few studies state these are feasible in continuous-flow system [13, 14, 15, 16].

RESULTS AND DISCUSSION

In our previous work, one-step reductive amination of various ketones was examined [17]. It was found that in those cases when the carbonyl-group is at α-position from an aromatic or heteroaromatic ring Zn dust promoted reactions gave the corresponding amines without notable side reactions, whereas in the case of aliphatic and cycloaliphatic ketones 10% Pd/C-catalysis was suitable.

[a] *Budapest University of Technology and Economics, Department of Organic Chemistry and Technology, and Research Group for Alkaloid Chemistry of HAS, Műegyetem rkp 3, H-1111, Budapest, Hungary, * poppe@mail.bme.hu*
[b] *ThalesNano Inc., Graphisoft Park, Záhony u. 7., H-1031 Budapest, Hungary*
[c] *Department of Cellular Biology and Pharmacology, Herbert Weinheim College of Medicine, Florida International University, 11200 S.W. 8th Street, Miami, FL 33199, USA*

The amines **2a-d** for the enzymatic studies were prepared by the reductive amination of the corresponding ketone **1a-d** by our novel method [17] (Scheme 1). Aliphatic ketones **1a,b** were treated with ammonium formate in methanol at 40ºC until the disappearance of the starting ketone. The reaction was catalysed by 10% Pd/C and was performed in continuous-flow reactor. Conversely, Zn dust proved to be an effective catalyst for the transformation of carbonyl groups at benzylic sidechain position of aromatic system **1c,d** at the reflux temperature of methanol in batch reaction (Scheme 1). The following yields were achieved: 53% for *rac*-**2a**, 41% for *rac*-**2b**, 71% for *rac*-**2c** and 71% for *rac*-**2d**.
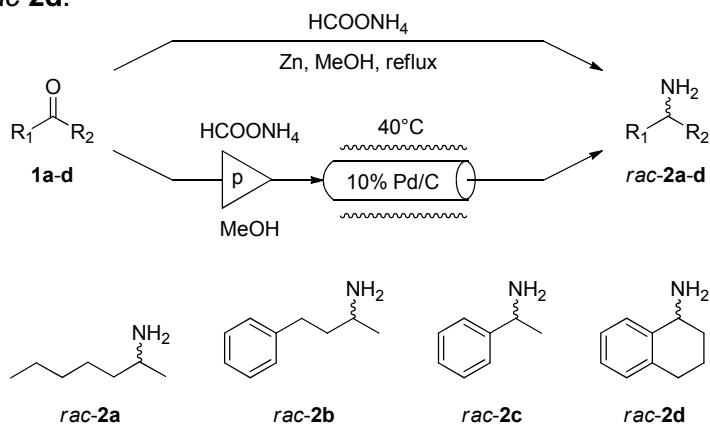
Scheme 1

Herein, we intended to study the effects of solvent and nature of the biocatalyst on the lipase catalysed kinetic resolution of four racemic amines **2a-d**. For selecting the proper catalyst for the continuous-flow mode kinetic resolution various lipases (immobilized and non-immobilized) were screened in batch mode. The two examined CalB (*Candida antarctica* Lipase B) enzymes – immobilized on polymeric carriers by two different methods – resulted in formation of (*R*)-*N*-acetamides in high enantiomeric excesses (*ee*) with moderate to good conversions (*c*). Our in house made BUTE 3 (F-4) [18] biocatalyst also gave acceptable enantiomeric excess and conversion in certain cases.

Efficiency of the biocatalytic reactions are highly influenced by the milieu. Thus, different solvents (toluene, trifluorotoluene, *tert*-butyl methyl ether, diisopropyl ether, ethyl acetate, hexane, tetrahydrofuran, hexane-tetrahydrofuran 2:1) and acetylating agents (ethyl acetate, isopropenyl acetate, ethylene glycol diacetate) were examined next. The best results (*ee* and *c*) were reached with the use of ethyl acetate as acetylating agent in toluene (representative results in toluene are in Table 1). It was also noticed that the use of ether-like solvents gave high enantiomeric excesses, however with lower conversions (data not shown). Our study also revealed, that the nature of immobilization of lipase B from *Candida antarctica* (Novozym 435 vs. CalB T2-150) had remarkable effects on the activity and selectivity of the enzyme (Table 1).

Table 1. Kinetic resolution of racemic amines **2a-d** with ethyl acetate in toluene

Substrate	Enzyme	Time [h]	c [%][a]	$ee_{(R)-3a}$ [%][a]	$E_{c,ee(P)}$ [%][b]
2a	Novozym 435	8	25.4	98.0	>100
2a	CalB T2-150	8	11.5	97.9	>100
2a	Novozym 435	24	37.0	96.0	86
2a	CalB T2-150	24	19.7	97.1	85
2b	Novozym 435	8	39.8	97.4	>100
2b	CalB T2-150	8	28.9	98.7	>200
2b	BUTE 3 (F-4)	8	18.7	97.7	>100
2b	Novozym 435	24	52.7	96.4	-[c]
2b	CalB T2-150	24	43.5	96.8	>100
2b	BUTE 3 (F-4)	24	28.0	98.4	>100
2c	Novozym 435	8	12.0	95.5	50
2c	CalB T2-150	8	3.2	85.3	13
2c	Novozym 435	24	25.5	94.2	46
2c	CalB T2-150	24	8.1	88.3	17
2d	CalB T2-150	8	3.3	94.3	35
2d	Novozym 435	8	16.4	96.4	66
2d	CalB T2-150	24	9.1	92.2	27
2d	Novozym 435	24	32.1	95.0	61

[a] Determined by chiral phase GC;
[b] Calculated by using c and $ee_{(R)-3a}$ [19]. Due to sensitivity to experimental errors, E values calculated in the 100–250 range are reported as >100, values above 250 are given as >200;
[c] Enantiomeric selectivity can not be calculated properly above 50% conversion [19].

Finally, kinetic resolutions of racemic amines **2a-d** were performed at preparative scale in X-Cube reactor operating in continuous-flow mode. The solution of the corresponding amine **2a-d** in toluene-ethyl acetate 9:1 was pumped through a Novozym 435 filled CatCart column thermostated to 30ºC at a flow rate of 0.5 mL min⁻¹ (Scheme 2).
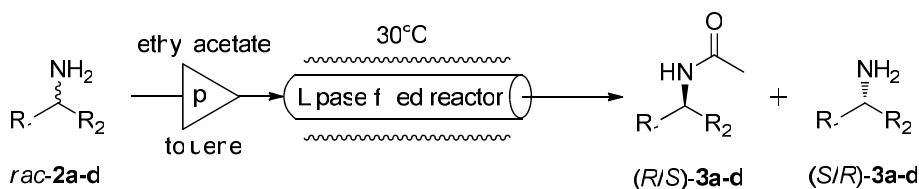
Scheme 2

After the stationary state reached (~8 times dead volume of the column, 40 min), feeding the reactor with amine **2a-d** was continued for 6 h. The *N*-acetamide **3a-d** was isolated from the collected homogenous solutions (Table 2). In all cases, higher enantiomeric excesses and conversions were obtained in continuous-flow reactor than in the corresponding batch reaction. The main advantage of the continuous-flow system with lipase filled-columns is the recyclability and the efficient and reproducible use of the catalyst. This was demonstrated by repeating the 6 h long reactions of amines **2a-d** in a second series of 6 h long reactions using the same Novozym 435-filled column with the same results.

Table 2. Kinetic resolution of racemic amines **2a-d** by Novozyme 435
in continuous-flow reactor

Substrate	c [%][a]	$ee_{(R)\text{-}3}$ [%][b]	$E_{c,ee(P)}$ [%][c]
2a	43.3	97.9	>200
2b	47.2	98.8	>200
2c	48.1	98.7	>200
2d	45.7	99.3	»200

[a] Determined after removal of the solvent from the resulting mixture;
[b] Determined by chiral phase GC;
[c] Calculated by using c and $ee_{(R)\text{-}3}$ [19]. Due to sensitivity to experimental errors, E values calculated in the 100–250 range are reported as >100, values in the 250–500 range are reported as >200 and values calculated above 500 are given as »200

CONCLUSIONS

In this study, solvent and catalyst effects on the kinetic resolutions of racemic amines **2a-d** (synthesized from the corresponding ketones **1a-d**) in batch and in continuous flow mode were investigated. It was found that immobilized forms of lipase B from (Novozym 435 and CalB T-2 150) are the most suitable biocatalysts in acylations with ethyl acetate in toluene with the examined substrates **2a-d**. Comparison of the batch and continuous mode reactions indicated that the continuous-flow process was superior to the corresponding batch reaction in cases of all the four (R)-acetamides (R)-**3a-d**.

EXPERIMENTAL SECTION

Analytical methods

The NMR spectra were recorded in DMSO on a Bruker DRX-500 spectrometer and are reported in ppm on the δ scale. Infrared spectra were recorded on a Bruker ALPHA FT-IR spectrometer. TLC was carried out on Kieselgel 60F254 (Merck) sheets. Spots were visualized under UV light (Vilber Lourmat VL-6.LC, 254 nm and 365 nm) or by treatment with 5% ethanolic phosphomolybdic acid solution and heating of the dried plates. GC analyses were carried out on Younglin ACME 6100 or Agilent 4890D instruments equipped with FID detector and Hydrodex-β-TBDAc column (50 m × 0.25 mm × 0.25 μm film with heptakis-(2,3-di-*O*-acetyl-6-*O*-*t*-butyldimethylsilyl)-β-cyclodextrin; Macherey&Nagel) or Hydrodex-β-6TBDM column (25 m × 0.25 mm × 0.25 μm film with heptakis-(2,3-di-*O*-methyl-6-*O*-*t*-butyldimethylsilyl)-β-cyclodextrine; Macherey&Nagel) using H_2 carrier gas (injector: 250°C, detector: 250°C, head pressure: 10 psi, 50:1 split ratio). The continuous flow reactions were performed by X-Cube™ laboratory flow reactor (X-Cube™ - trademark of ThalesNano, Inc.; Ser. No.: 002/2006) equipped with 10% Pd/C [THS 01111] or Novozym 435 [THS 01724] filled CatCart™ columns (CatCart™ – registered trademark of ThalesNano Inc.: cat. no.: THS X1175; stainless steel (INOX 316L); inner diameter: 4 mm; total length: 70 mm; packed length: 65 mm; inner volume: 0.816 mL).

Chemicals and enzymes

Heptan-2-one **1a**; 4-phenylbutan-2-one **1b**; acetophenone **1c**; 3,4-dihydronaphthalen-1(2*H*)-one **1d**; ammonium formate and all further chemicals and solvents were of analytical grade or higher were products of Sigma, Aldrich, Fluka, Alfa Aesar or Merck. 10% Pd/C [THS 01111] and Novozym 435 [THS 01724] filled CatCart™ columns were products of ThalesNano, Inc. Lipase AK, Lipase PS, Lipase AYS were obtained from Amano Europe. Novozym 435 was purchased from Novozymes, Denmark. *Mucor miehei* lipase, Lipase PPL, Lipase CcL, *Candida rugosa* were products of Sigma. Lipase AY, Lipase M were obtained from Amano Pharmaceutical. Lipozyme TL IM was purchased from Novo Nordisk A/S. CalB T2-150, CalA T2-150, CRL T2-150, IMMAULI T2-150 were products of Chiral Vision. Lipobond-Lipase PS was a kind gift of Iris Biotech GmbH. The sol-gel immobilzed lipases (sol-gel lipase AK: 338b and 251b; sol-gel lipase PS) [20], lipases from thermophilic fungi (BUTE 3 (T-2), BUTE (F-4)) [18] and lipases from solid-state fermentations (SSF9, SSF23) [21] were prepared in our laboratory.

General procedure of the continuous-flow reductive amination of ketones (*1a,b*) [17]

The solution of the corresponding ketone (5 mg mL^{-1} of **1a,b**: 0.044 mmol mL^{-1} of **1a**, 0.034 mmol mL^{-1} of **1b**) and 6 equiv. ammonium formate (16.65 mg mL^{-1} for **1a**, 12.86 mg mL^{-1} for **1b**) in methanol was pumped through the 10% Pd/C filled column thermostated to 40°C at a flow rate of 0.2 mL min^{-1} without choking (no measurable back-pressure). After the stationary state reached (approximately 8x the whole volume of the system) the mixture was pumped through the column for 6 hours. After a run the columns were routinely washed with methanol (0.5 mL min^{-1}, 20 min). The collected reaction mixture was concentrated under reduced pressure. The residue was treated with conc. HCl solution (3 mL) and water (20 mL) and extracted with diethyl ether (2x15 mL). The aqueous phase was treated with ammonia solution until pH=10 and extracted with dichloromethane (4x20 mL). The organic phase was treated with brine, dried over sodium sulfate and the solvent was distilled off from the resulting solution by rotary evaporation to give the amines **2a-d**.

racemic Heptan-2-amine *rac*-**2a**: pale yellow liquid; ^{1}H NMR (300 MHz, DMSO-d$_6$): 0.86 (3H, t, J=7.0 Hz, CH$_3$); 0.93 (3H, d, J=6.2 Hz, CH$_3$); 1.11-1.36 (8H, m, 4×CH$_2$); 2.93 (brs, NH$_2$+H$_2$O); 2.71 (1H, m, J=6.3 Hz, CHN); ^{13}C NMR (75 MHz, DMSO-d$_6$): 14.47 (CH$_3$); 22.70 (CH$_2$); 24.49 (CH$_3$); 26.08 (CH$_2$); 32.10 (CH$_2$); 40.29 (CH$_2$); 46.87 (CH); IR (cm^{-1}): 3325, 3284, 2956, 2927, 2858, 1564, 1456, 1362, 1294, 1166, 1097, 855, 814, 725

racemic 4-Phenylbutan-amine *rac*-**2b**: colourless liquid; ^{1}H NMR (300 MHz, DMSO-d$_6$, δ ppm): 1.00 (3H, d, J=6.3 Hz, CH$_3$); 1.43-1.58 (4H, brs+m, NH$_2$+CH$_2$); 2.50-2.70 (2H, m, CH$_2$); 2.74 (1H, q, J=6.3 Hz, CHN); 7.12-7.29 (5H, m, 5×CH); ^{13}C NMR (75 MHz, DMSO-d$_6$, δ ppm): 24.09 (CH$_3$); 32.10 (CH$_2$); 41.74 (CH$_2$); 45.92 (CH); 125.42 (CH); 128.14 (2×CH); 128.16 (2×CH); 142.51 (C); IR (cm^{-1}): 3412, 3396, 2956, 2925, 2858, 1561, 1458, 1363, 1302, 1163, 886, 854, 815, 723, 460

General procedure for the reduction of ketones 1c,d to amines 2c,d in one-step batch synthesis [17]

A mixture of the ketone **1c,d** (10 mmol: 1.46 g of **1c**, 1.20 g of **1d**), ammonium formate (60 mmol, 3.78 g) and Zn powder (30 mmol, 1.96 g) in methanol (30 mL) was stirred under reflux until disappearance of the ketone (monitored by TLC). The reaction mixture was strained through Celite®, and the solvent was removed under vacuum. The residue was treated with conc. HCl solution (4 mL) and water (30 mL) and extracted with diethyl ether (2x20 mL). The aqueous phase was treated with ammonia solution until pH=10 and extracted with dichloromethane (4x25 mL). The organic phase was treated with brine, dried over sodium sulfate and the solvent was distilled off from the resulting solution by rotary evaporation.

racemic 1-Phenylethanamine *rac*-**2c**: colourless liquid; ^{1}H NMR (300 MHz, DMSO-d_6, δ ppm): 1.24 (3H, d, J=6.6 Hz, CH$_3$); 3.02 (brs, NH$_2$+H$_2$O); 3.97 (1H, q, J=6.6 Hz, CHN); 7.18 (1H, m, CH); 7.29 (2H, m, 2×CH); 7.36 (2H, m, 2×CH); ^{13}C NMR (75 MHz, DMSO-d_6, δ ppm): 26.69 (CH$_3$); 51.16 (CH); 126.25 (2×CH); 126.64 (CH); 128.57 (CH); 149.26 (C); IR (cm^{-1}): 3362, 3301, 3061, 3026, 2965, 2924, 2868, 1604, 1579, 1492, 1475, 1450, 1364, 1024, 859, 762, 698, 591, 537

racemic 1,2,3,4-Tetrahydronaphthalen-1-amine *rac*-**2d**: yellow liquid; ^{1}H NMR (300 MHz, DMSO-d_6): 1.47-1.72 (2H, m, CH$_2$); 1.78-1.96 (2H, m, CH$_2$); 2.57-2.78 (2H, m, CH$_2$); 3.02 (brs, NH$_2$+H$_2$O); 3.79 (1H, t, J=6.2 Hz, CHN); 6.98-7.16 (3H, m, 3×CH); 7.40-7.48 (1H, m, CH); ^{13}C NMR (75 MHz, DMSO-d_6): 19.97 (CH$_2$); 29.67 (CH$_2$); 33.76 (CH$_2$); 49.41 (CH); 125.96 (CH); 126.42 (CH); 128.61 (CH); 128.89 (CH); 136.73 (C); 142.39 (C); IR (cm^{-1}): 3412, 3381, 3057, 3015, 2924, 2858, 1578, 1488, 1447, 1371, 1337, 1155, 884, 854, 761, 733, 584, 433

Enantiomer selective acetylation of racemic amines 2a-d in shake vials

To a solution of the racemic amine **2a-d** (20 mg) in toluene-ethyl acetate 9:1 mixture (1 mL), the enzyme (20 mg) was added in a sealed amber glass vial and the resulting mixture was shaken (1000 rpm) at 30°C for 24 h. The reactions were analyzed by GC and TLC after 1, 2, 4, 8 and 24 hours.

Enantiomer selective acetylation of racemic amines 3a-d in continuous mode

The solution of racemic amine **2a-d** (10 mg mL^{-1}) in toluene-ethyl acetate 9:1 mixture was pumped through three serially connected Novozym 435-filled CatCart columns operated at 30ºC and the product was collected at a flow rate of 0.5 mL min^{-1} without choking. After 40 min (which was necessary to reach the stationary state of the bioreactor) the reaction mixture was collected for 6 h. The solvent was removed from the reaction mixture by vacuum rotary evaporation. The residue was dissolved in dichloromethane (100 mL), treated with 5% HCl solution (10 mL) and the aqueous phase was extracted with dichloromethane (4x3 mL). The combined organic phases were dried over Na$_2$SO$_4$ and concentrated under vacuum. The aqueous phase was alkalized with ammonia solution to pH=10 and extracted with dichloromethane (3x10 mL). The organic phase was dried over Na$_2$SO$_4$ and concentrated under vacuum.

(R)-N-(Heptan-2-yl)acetamide (R)-**3a**: yellow oil; ^{1}H NMR (300 MHz, DMSO-d$_6$, δ ppm): 0.85 (3H, t, J=7.0 Hz, CH$_3$); 0.99 (3H, d, J=6.6 Hz, CH$_3$); 1.15-1.40 (8H, m, 4×CH$_2$); 1.76 (3H, s, CH$_3$); 3.70 (1H, m, CHN); 7.62 (1H, br d, J=8.1 Hz, NH); ^{13}C NMR (75 MHz, DMSO-d$_6$, δ ppm): 14.38 (CH$_3$); 21.22 (CH$_3$); 22.52 (CH$_2$); 23.14 (CH$_3$); 25.86 (CH$_2$); 31.72 (CH$_2$); 36.57 (CH$_2$); 44.54 (CH); 168.76 (CO); IR (cm^{-1}): 3275, 3079, 2959, 2928, 2858, 1638, 1550, 1453, 1371, 1293, 1157, 974, 726, 608

(R)-N-(4-Phenylbutan-2-yl)acetamide (R)-**3b**: pale yellow liquid; ^{1}H NMR (300 MHz, DMSO-d$_6$): 1.04 (3H, d, J=6.6 Hz, CH$_3$); 1.57-1.71 (2H, m, CH$_2$); 1.81 (3H, s, CH$_3$); 2.48-2.61 (2H, m, CH$_2$); 3.75 (1H, m, CHN); 7.11-7.21 (3H, m, 3×CH); 7.22-7.31 (2H, m, 2×CH); 7.74 (1H, br d, J=8.1 Hz, NH); ^{13}C NMR (75 MHz, DMSO-d$_6$): 21.24 (CH$_3$); 23.24 (CH$_3$); 32.42 (CH$_2$); 38.44 (CH$_2$); 44.42 (CH); 126.09 (CH); 128.83 (4×CH); 142.36 (C); 168.93 (CO); IR (cm^{-1}): 3271, 3084, 3065, 2967, 2929, 1637, 1546, 1495, 1453, 1371, 1292, 1144, 967, 746, 698, 609

(R)-N-(1-Phenylethyl)acetamide (R)-**3c**: tawny crystals; ^{1}H NMR (300 MHz, DMSO-d$_6$): 1.33 (3H, d, J=7.2 Hz, CH$_3$); 1.84 (3H, s, CH$_3$); 4.90 (1H, m, CHN); 7.14-7.26 (1H, m, CH); 7.27-7.36 (4H, m, 4×CH); 8.30 (1H, br d, J=7.9 Hz, NH); ^{13}C NMR (75 MHz, DMSO-d$_6$): 22.47 (CH$_3$); 22.64 (CH$_3$); 47.78 (CH); 125.93 (2×CH); 126.64 (CH); 128.28 (2×CH); 144.87 (C); 168.30 (CO); IR (cm^{-1}): 3266, 3071, 3023, 2980, 2929, 1643, 1555, 1451, 1375, 1286, 1278, 1216, 1136, 1070, 1027, 763, 702, 620, 533, 501

(R)-N-(1,2,3,4-Tetrahydronaphthalen-1-yl)acetamide (R)-**3d**: brown crystals; ^{1}H NMR (300 MHz, DMSO-d$_6$): 1.58-1.78 (2H, m, CH$_2$); 1.79-1.95 (5H, m, CH$_3$+CH$_2$); 2.63-2.82 (2H, m, CH$_2$); 4.90-5.03 (1H, m, CHN); 7.04-7.20 (4H, m, 4×CH); 8.22 (1H, br d, J=8.5 Hz, NH); ^{13}C NMR (75 MHz, DMSO-d$_6$): 20.47 (CH$_2$); 23.20 (CH$_3$); 29.25 (CH$_2$); 30.48 (CH$_2$); 46.86 (CH); 126.28 (CH); 127.22 (CH); 128.67 (CH); 129.28 (CH); 137.45 (C); 138.07 (C); 169.05 (CO); IR (cm^{-1}; KBr): 3240, 3062, 2928, 2854, 1633, 1544, 1445, 1371, 1283, 1095, 1038, 965, 764, 739, 610, 536, 446

ACKNOWLEDGEMENTS

REFERENCES

1. T. Henkel, R.M. Brunne, H. Müller, F. Reichel, *Angew. Chem. Int. Ed.,* **1999**, *38*, 643.
2. T.A. Unger, Ed., „Pesticide Synthesis Handbook" Noyes Publications, Park Ridge, **1996**.
3. W. Herbst, K. Hunger, Eds., „Industrial Organic Pigments: Production, Properties, Applications (3rd Completely Revised Edition)" Wiley-VCH, Weinheim, **2004**.

4. A.S. Bommarius, B.R. Riebel, "Biocatalysis", Wiley-VCH, Weinheim, **2004**.

5. U.T. Bornschauer, R.J. Kazlauskas, "Hydrolases in Organic Synthesis: Regio- and Stereoselective Biotransformations", Wiley-VCH, Weinheim, **2006**.

6. R.J. Kazlauskas, A.N.E. Weissfloch, A.T. Rappaport, L.A. Cuccia, *J. Org. Chem.*, **1991**, 56, 2656.

7. K. Burgess, L.D. Jennings, *J. Am. Chem. Soc.*, **1991**, *113*, 6129.

8. F. Theil, *Catal. Today*, **1994**, *22*, 517.

9. A. Ghanem, H.Y. Aboul-Enein, *Chirality*, **2005**, *17*, 1.

10. N.J. Turner, *Curr. Opin. Chem. Biol.*, **2004**, *8*, 114.

11. K. Faber, "Biotransformations in Organic Chemistry", Springer, Berlin, **2004**.

12. L. Poppe, L. Novák, "Selective Biocatalysis: A Synthetic Approach", VCH, Weinheim-New York, **1992**.

13. A. Liese, K. Seelbach, C. Wandrey, "Industrial Biotransformations", Wiley-VCH: Weinheim-New York, **2006**.

14. R.N. Patel, A. Banerjee, L.J. Szarka, *J. Am. Oil Chem. Soc.*, **1996***, 73,* 1363.

15. J.C. Chen, S.W. Tsai, *Biotech. Prog.*, **2000**, *16*, 986.

16. Z. Ujang, W.H. Husain, M.C. Seng, A.H.A. Rashid, *Proc. Biochem.*, **2003**, *38*, 1483.

17. P. Falus, Z. Boros, G. Hornyánszky, J. Nagy, F. Darvas, L. Ürge, L. Poppe, *Tetrahedron Lett.*, **2010**, accepted.

18. V. Bódai, R. Peredi, J. Bálint, G. Egri, L. Novák, G. Szakács, L. Poppe, *Adv. Synth. Catal.,* **2003**, *345*, 811.

19. C.S. Chen, Y. Fujimoto, G. Girdaukas, C.J. Sih, *J. Am. Chem. Soc.*, **1982**, *104*, 7294.

20. A. Tomin, D. Weiser, G. Hellner, Z. Bata, L. Corici, F. Péter, B. Koczka, L. Poppe, *Proc. Biochem.*, **2010**, doi:10.1016/j.procbio.2010.07.021

21. V. Nagy, E.R. Tőke, L. Chee Keong, G. Szatzker, D. Ibrahim, I. Che-Omar, G. Szakács, L. Poppe, *J. Mol. Catal. B: Enzym.*, **2006**, *39*, 141.

RAPID AND SIMPLE ANALYSIS OF ALLICIN IN
ALLIUM SPECIES BY LC-CIS-MS/MS

LAURIAN VLASE[a,*], MARCEL PÂRVU[b], ANCA TOIU[a], ELENA ALINA PÂRVU[c], SIMONA CODRUŢA COBZAC[d], MIHAI PUŞCAŞ[e]

ABSTRACT. A simple liquid chromatography-coordination ion spray-mass spectrometry method (LC-CIS-MS/MS) for analysis of allicin in *Allium* species extracts has been developed. Allicin was chromatographic separated under isocratic conditions using a mobile phase of 1 mM ammonium acetate. A silver nitrate solution was post-column added to enhance the allicin detection by formation an ionized coordinated complex. The overall time of one analysis was 1 min. The detection of allicin was performed in multiple reaction monitoring mode using an ion trap mass spectrometer with electrospray positive ionization. The linearity domain was established between 18 and 864 µg/mL. Inter-day accuracy and precision were less than 11% and 2.2%, respectively.

Keywords: allicin, Allium extracts, liquid chromatography, coordination mass spectrometry

INTRODUCTION

Allium species have been used for food and medicine for thousands of years, especially *Allium sativum* (garlic) and *Allium cepa* (onion), and recently interest in other species has been increasing [1,2]. Garlic is considered as a medicinal plant and especially one of the best disease-preventive foods against some forms of cancer and cardiovascular disorders. Its beneficial widespread effect on health is attributed to sulphur-containing compounds, and particularly to thiosulfinates [3]. When garlic is cut or crushed, the enzyme alliinase is released from its compartment and transforms S-allyl-L-cysteine sulfoxide (alliin) into diallyl thiosulfinate (allicin, **ALC**, Fig.1), the characteristic compound of garlic flavor. The hypocholesterolaemic activity of garlic has been attributed to diallyl disulphide, a decomposition product of ALC [4]. Ajoene (a secondary degradation product of alliin) inhibits platelet aggregation by altering the platelet membrane via an interaction with sulphydryl groups [4]. Antimicrobial activity is well documented for garlic, and antifungal activity is more effective

[a] *University of Medicine and Pharmacy "Iuliu Haţieganu", Faculty of Pharmacy, Emil Isac 13, RO-400023, Cluj-Napoca, Romania, * vlaselaur@yahoo.com; corresponding author*
[b] *"Babeş-Bolyai" University, Faculty of Biology and Geology, 42 Republicii Street, RO-400015 Cluj-Napoca, Romania*
[c] *University of Medicine and Pharmacy "Iuliu Haţieganu", Faculty of Medicine, Emil Isac 13, RO-400023, Cluj-Napoca, Romania*
[d] *"Babeş-Bolyai" University, Faculty of Chemistry and Chemical Engineering, Arany Janos 11, RO-400028, Cluj-Napoca, Romania*
[e] *"Babeş-Bolyai" University, "A. Borza" Botanical Garden, 42 Republicii Street, RO-400015 Cluj-Napoca, Romani*

than nystatin (ALC is the main active component by inhibition of lipid synthesis). In vitro antiviral activity was attributed to ALC and its derivatives, and alliin has anti-hepatotoxic activity in vitro and in vivo [4,5].

Several methods for qualitative or quantitative determination of ALC have been reported. Mainly, high performance liquid chromatography methods with UV or MS detection were described [6-10]. However, the main drawback of the previously reported methods for analysis of ALC are the long analysis time – about 10-30 minutes per sample and the poor specificity in case of UV detection (Table 1).

The analysis by liquid chromatography-coordination ion spray-mass spectrometry method (LC-CIS-MS/MS) has been used as an alternative technique instead of classical LC/MS, especially when a compound does not readily ionize in the ion source of mass spectrometer due low proton affinity. In the LC-CIS-MS/MS, the molecule is ionized by attachment of a metallic ion with coordination capabilities (usually silver) instead of a proton. The site of coordination is usually a carbon-carbon double bond [11] or sulfur atom [12].
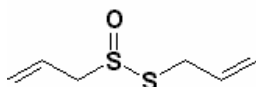
Figure 1. Molecular structure of ALC

The aim of present study is the development of a rapid and specific LC-CIS-MS/MS method for ALC quantification from *Allium* extracts. In comparison with previously published HPLC methods (Table 1), the proposed method is rapid and specific.

RESULTS AND DISCUSSION

Although ALC absorb UV light at 220 nm and in the literature there are reported analytical methods using this type of detection [6-8, 10], frequent interferences may appear at this wavelength because of lack of selectivity, leading to measurement errors.

Regarding mass spectrometry analysis, because ALC does not have ionizable chemical groups (either acidic or basic) it cannot be readily ionized under the influence of the pH of the mobile phase, so it cannot be detected „as it is" by LC/MS-electrospray with good sensitivity.

The sulphur-containing compounds have the ability to form adduct complexes with some transitional metals. Because of the metallic ion, such complex has an electric charge and it can be analyzed by mass spectrometry with electrospray ionization. For ALC and some of its derivatives, their complexes with silver were already reported, but only for qualitative analysis [12]. In order to obtain the selectivity in quantitative measurement of ALC by LC/MS we used for quantification the adduct complex formed by ALC and the silver ion. An aqueous silver nitrate 1mM solution at a flow rate of 10 µL/min was post-column added to effluent.

298

Table 1. Analytical characteristics of several reported HPLC or LC/MS methods for the determination of ALC in vegetal extracts

Extraction	Mobile phase	Detection	Run time (min)	Observations	Ref.
Extraction at room temperature	Sodium dihydrogenphosphate+ heptanesulfonic acid / acetonitrile, gradient	HPLC-UV, 208 nm	30	Other ALC derivatives also analysed	**6**
Ultrasonic extraction/ centrifugation	Water-methanol, isocratic	HPLC-UV, 254 nm	20	-	**7**
Turboextraction, liquid-liquid extraction	Water-methanol, isocratic	HPLC-UV, 254 nm after post-column photochemical derivatisation	10	-	**8**
Supercritical fluid extraction	Water-acetonitrile gradient	LC-APCI-MS	25	Other ALC derivatives also analysed	**9**
Vortex-sonication	Water-methanol, isocratic	HPLC-UV, 220 nm	15	-	**10**

The pseudo-molecular mass spectra (MS[1], no fragmentation applied) obtained for ALC-silver complex (Fig. 2) shows two main ions with m/z 449 and 451, corresponding to a molecular formula $[2*ALC+H_2O+Ag]^+$, in which the silver ion is surrounded by two ALC molecules and one molecule of water. The two major ions from the spectra (m/z 449 and 451) are the adducts formed by the two silver isotopes. The mass spectrum of ALC-silver complex (MS[2], fragmentation applied) is shown in Fig. 3 and the proposed main fragmentation pathways in Fig. 4.
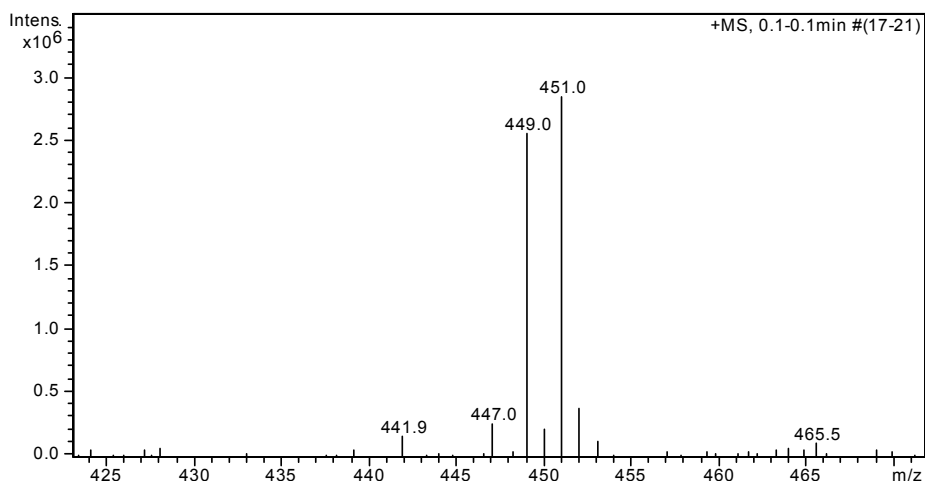
Figure 2. Pseudo-molecular (MS[1]) non-reactive ion on mass spectra of ALC-silver complex
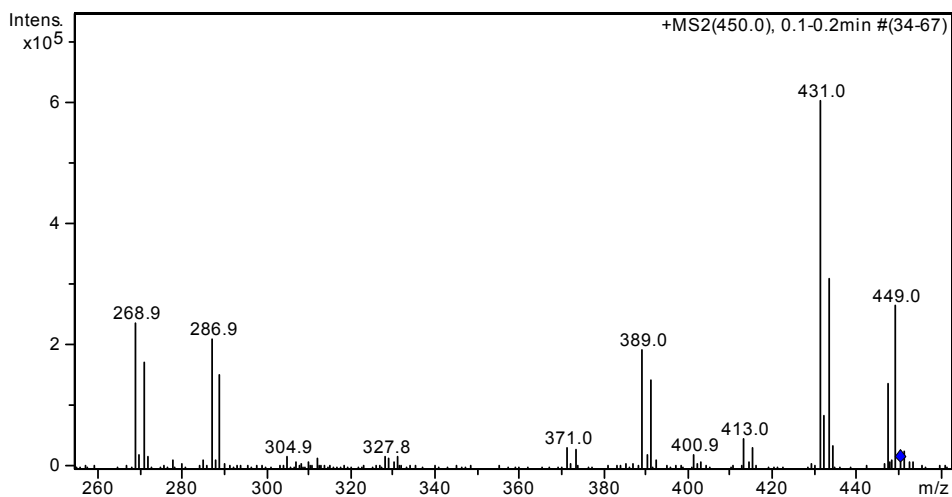
Figure 3. MS/MS mass spectra of ALC-silver complex
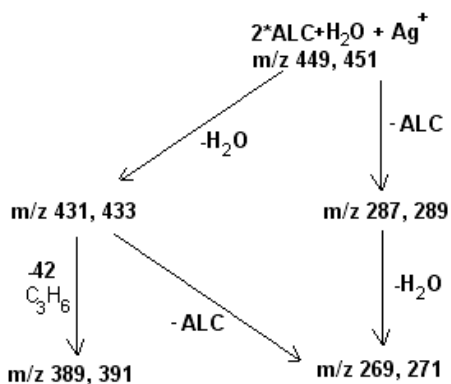
Figure 4. The proposed fragmentation pathways of ALC-silver complex

The peak of ALC was observed at RT = 0.9 min (Fig. 5).

The calibration curves showed linear response over the entire range of concentrations used in the assay procedure. The calibration curve for ALC was in the concentration range 18.0-864.0 µg/mL, using 7 calibration levels, n = 3 days, with a coefficient of correlation greater than 0.999. The residuals had no tendency of variation with concentration and were between ±13.1% values. The bias and precision of calibration curves are presented in Table. 2.

The developed analytical method was applied for analysis of ALC in five *Allium* species extracts. The found concentrations are presented in Table 3. The extracts prepared by heating at 60 ºC ("C" extracts) are richer in ALC content than "R" extracts, proving that extraction at higher temperatures favors the transformation of alliin to ALC.
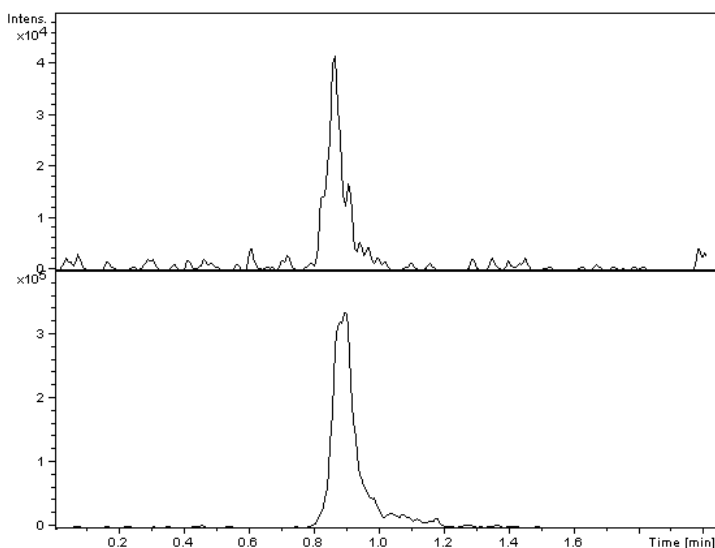
Figure 5. Typical chromatogram of the ALC standard (upper image)
and *A. obliquum* extract (lower image)

Table 2. The bias and precision for ALC determination method

Cnominal µg/mL	Calibration 1		Calibration 2		Calibration 3		Inter-day precision %	Inter-day bias %
	Cfound µg/mL	Bias %	Cfound µg/mL	Bias %	Cfound µg/mL	Bias %		
18.0	18.7	4.0	18.6	3.3	18.8	4.4	0.56	3.91
36.0	33.0	-8.2	32.8	-8.9	32.3	-10.2	1.09	-9.11
72.0	75.0	4.1	77.5	7.7	78.4	9.0	2.34	6.92
144.0	160.8	11.6	162.9	13.1	155.9	8.3	2.24	11.00
288.0	266.4	-7.5	267.0	-7.3	267.3	-7.2	0.17	-7.33
576.0	576.5	0.1	563.9	-2.1	570.2	-1.0	1.11	-1.00
864.0	874.0	1.2	871.9	0.9	869.7	0.7	0.25	0.91
Slope	5630.80		5576.19		5690.72			
Interrcept	44976.69		34772.45		48727.85			
r	0.99931		0.99915		0.99935			

Table 3. The ALC content found in various *Allium* extracts

Allium species	ALC (mg/ml in "R" extracts)	ALC (mg/100 g vegetal product, "R" extracts)	ALC (mg/ml in "C" extracts)	ALC (mg/100 g vegetal product, "C" extracts)
A. obliquum	2.819	272.806	5.579	426.629
A. senescens subsp. *montanum*	0.919	82.474	5.880	435.931
A. schoenoprasum	3.968	320.001	3.410	947.222
A. fistulosum	0.122	5.275	4.481	896.225
A. ursinum (leaves)	0.028	1.965	-	-
A. ursinum (flowers)	1.946	175.614	-	-

301

CONCLUSIONS

This is the first reported quantitative analytical method for determination of ALC by using LC-CIS-MS/MS. The developed method is fast (1 minute run-time), selective and provides high-throughput in analysis of ALC from *Allium* species extracts.

EXPERIMENTAL SECTION

Reagents

Reference standard of Allicin (ALC) was purchased from "Allicin International", Great Britain. Methanol, ammonium acetate and silver nitrate were Merck products (Merck KgaA, Darmstadt, Germany). Distilled, deionised water was produced by a Direct Q-5 Millipore (Millipore SA, Molsheim, France) water system.

Standard solutions

A stock solution of ALC with concentration of 4 mg/mL was prepared by dissolving appropriate quantity of reference substance in 10 mL metanol. The working solution was obtained by diluting a specific volume of stock solution with water. This solution was used to prepare 7 calibration solutions with concentration range between 18-864 µg/mL.

Chromatographic and mass spectrometry systems and conditions

The HPLC system was an 1100 series model (Agilent Technologies) consisted of a binary pump, an in-line degasser, an autosampler, a column thermostat, and an Ion Trap SL mass spectrometer detector (Brucker Daltonics GmbH, Germany). Chromatograms were processed using QuantAnalysis software. The detection of ALC was in MS/MS mode using electrospray positive ionisation (ESI+). The monitored ion transition was m/z (449 + 451) > m/z (269 + 271 + 287 + 289) Chromatographic separation was performed at 40°C on a Synergy Polar 100 mm x 2 mm i.d., 4 µm column (Phenomenex, Torrance, SUA), protected by an in-line filter.

Mobile phase

The mobile phase consisted in 100% ammonium acetate, 1mM in water, isocratic elution, flow 0.6 mL/min. A silver nitrate solution 1mM in water was post column added, with a flow of 10 µL/min.

Sample preparation

Fresh stems and leaves of *Allium obliquum*, *A. senescens* subsp. *montanum*, *A. schoenoprasum*, *A. fistulosum*, leaves of *A. ursinum* and flowers of *A. ursinum* were crushed and then extracted with ethanol 70% by repercolation method at room temperature ("R" extracts). In order to observe the influence of the temperature on the allicin extraction it were also performed extractions by heating the mixtures at 60°C on water bath for 30 min ("C" extracts). All extracts were filtered and adjusted up to the final volume. All plants were identified and voucher specimen (CL 659564, CL 659563, CL 659561, CL 659761

and CL 659750) was deposited at the Herbarium of "A. Borza" Botanical Garden, "Babes-Bolyai" University of Cluj-Napoca. Before analysis, the vegetal extracts were diluted 10 folds with distilled water, then 1 µL was injected in the chromatographic system.

Method validation

In three different days, a calibration curve was run. The linearity of the peak area against standard concentration was verified between 18-864 µg/mL ALC by applying least-squares linear regression. The applied calibration model was y = a*x+b, 1/y weight, where y is peak area and x, concentration. Distribution of the residuals (% difference of the back-calculated concentration from the nominal concentration) was investigated. The calibration model was accepted if the residuals were within ±20% at the lower limit of quantification and within ±15% at all other calibration levels and at least 2/3 of the standards meet this criterion [13-15]. The limit of quantification was established as the lowest calibration standard with an accuracy and precision less than 20%.

ACKNOWLEDGMENTS

REFERENCES

1. C. Nencini, F. Cavallo, A. Capasso, G.G. Franchi, G. Giorgio, L. Micheli, *Phytother. Res.*, **2007**, *21(9)*, 874.

2. D. Stajner, J. Canadanovic-Brunet, A. Pavlovic, *Phytother. Res.*, **2004**, *18*, 522.

3. H. Amagase, B.L. Petesch, H. Matsuura, S. Kasuga, Y. Itakura, *J. Nutr.,* **2001**, *131*, 955S.

4. C.A. Newall, A.A. Anderson, J.D. Phillipson, "Herbal Medicines. A guide for Health-care Professionals", The Pharmaceutical Press, London, **1996**.

5. M. Parvu, A.E. Parvu, O. Rosca-Casian, L. Vlase, G. Groza, *J. Med. Plant. Res.,* **2010**, *4(2)*, 138.

6. I. Arnault, J.P. Christides, N. Mandon, T. Haffner, R. Kahane, J. Auger, *J. Chromatogr. A*, **2003**, *991*, 69.

7. K. Baghalian, S.A. Ziai, M.R. Naghavi, H.N. Badi, A. Khalighi, *Scientia Horticulturae*, **2005**, *103*, 155.

8. P. Bocchini, C. Andalo, R. Pozzi, G.C. Galletti, A. Antonelli, *Anal. Chim. Acta*, **2001**, *441*, 37.

9. E.M. Calvey, K.D. White, J.E. Matusik, D. Sha, E. Block, *Phytochem.*, **1998**, *38(1)*, 248.

10. M. De Diego, M. Avello, S. Mennickent, M. Fernandez, P. Fernandez, *J. Sep. Sci.*, **2007**, *30*, 2703.

11. A. Medvedovici, K. Lazou, A. d'Oosterlinck, Y.N. Zhao, P. Sandra, *J. Sep. Sci.*, **2002**, *25*, 173.

12. E. Block, "Garlic and Other Alliums: The Lore and the Science". The Royal Society of Chemistry, Cambridge, **2009**.

13. D. Mihu, L. Vlase, S. Imre, C.M. Mihu, M. Achim, D.L. Muntean, *Studia Univ. Babeş-Bolyai, Chemia*, **2009**, *54(3)*, 151.

14. M. Achim, D. Muntean, L. Vlase, I. Bâldea, D. Mihu, S.E. Leucuţa, *Studia Univ. Babeş-Bolyai, Chemia*, **2009**, *54(3)*, 7.

15. D.S. Popa, L. Vlase, S.E. Leucuţa, F. Loghin, *Farmacia*, **2009**, *57(3)*, 301.

RAPID LC/MS^3 METHOD FOR DETERMINATION OF MEMANTINE IN PIG PLASMA

LAURIAN VLASE[a], DANA MUNTEAN[a], MARCELA ACHIM[a]

ABSTRACT. A simple and sensitive liquid chromatography coupled with tandem mass spectrometry (LC/MS^3) method for the quantification of memantine in pig plasma was developed and validated. The separation was performed on a Zorbax SB-C18 column under isocratic conditions using a mobile phase of 55:45 (v/v) methanol and 0.1% (v/v) formic acid in water at 45 ºC with a flow rate of 1 mL/min. The detection of memantine was performed in multiple reaction monitoring mode using an ion trap mass spectrometer with electrospray positive ionisation, operating in MS^3 mode. The pig plasma samples (0.2 mL) were deproteinized with 6% perchloric acid in water and aliquots of 10 µL from supernatants obtained after centrifugation were directly injected into the chromatographic system. The method shows a good linearity (r > 0.995), precision (CV < 12.2%) and accuracy (bias < 12.1%) over the range of 4.86-486 ng/mL plasma. The lower limit of quantification (LLOQ) was 4.86 ng/mL and the recovery was between 92.5-100.8%. The developed and validated method is simple, rapid and specific for the determination of memantine in pig plasma and was successfully applied to a pharmacokinetic study of intravenously administered memantine in pigs.

Keywords: memantine, LC/MS^3, human plasma, pig plasma

INTRODUCTION

Memantine, 3,5-dimethyladamantan-1-amine (Figure 1) is a N-methyl-D-aspartate (NMDA) receptor antagonist with neuroprotective effect and is currently used for treating of patients with vascular dementia, Alzheimer disease, hemorrhagic stroke and neuropathic pain [**1-3**].

The process of neuronal degradation in cases of deprived oxygen supply (ischemia), as in cardiac arrest, is closely connected with the activity of the NMDA receptors [**1**]. Despite recent acquisition regarding neuronal vulnerability to hypoxia, at the moment the only clinically available mean of neuroprotection is the therapeutic hypothermia, but this method has limited applicability [**4**]. The similar mechanisms of neuronal injury in Alzheimer disease and in cerebral ischemia suggested the idea of potentially neuroprotective effects of intravenous memantine administered during global cerebral ischemia due experimentally induced cardiac arrest in pigs. In order to evaluate the pharmacological effect of memantine in induced cerebral ischemia and for evaluation of its pharmacokinetics, a fast and reliable analytical method for determination of memantine in pig plasma was needed.

[a] *University of Medicine and Pharmacy "Iuliu Haţieganu", Faculty of Pharmacy, Emil Isac 13, RO-400023, Cluj-Napoca, Romania, * vlaselaur@yahoo.com*
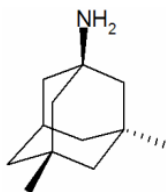
Figure 1. Chemical structure of memantine

Several methods involving gas-chromatography (GC) coupled with mass spectrometry (MS) [5] and high-performance liquid-chromatography (LC) with mass spectrometry (MS) [6-10] detection have been reported for determination of memantine in biological matrix.

The GC method requires derivatisation of memantine after liquid-liquid extraction (LLE) in order to transform it in a volatile molecule [5]. However, both extraction and derivatization are time-consuming steps, increasing the cost of assay and can affect the recovery. The LC/MS or LC/MS/MS methods offer considerable advantages by their powerful performances: speed, selectivity, sensitivity and robustness. However, the sample preparation procedure by solid phase extraction (SPE) [6] or LLE [7-9] may complicate the analysis in terms of speed and recovery.

The aim of this work was to develop and validate a new simple, specific and efficient LC/MS^3 assay for the quantification of memantine in pig plasma for application in a pharmacokinetic study.

RESULTS AND DISCUSSION

Sample preparation

In LC/MS assays the sensitivity depends on MS detection mode, but the method involved in sample preparation may also influence the chromatographic background level and can generate matrix suppression effects. Usually an extraction step of analyte from matrix prior to analysis (SPE or LLE) has two main advantages: sample purification and sample pre-concentration. As stated before, the extraction step (either SPE or LLE) is laborious, time consuming and usually needs an internal standard to compensate the extraction variability. The protein precipitation (PP), as sample processing method is desirable when one need a high-throughput analysis, and low sample-to sample extraction variability. However, the two main advantages of SPE or LLE extraction mentioned before become drawbacks in case of PP: first- the sample is not really purified so matrix interferences or high background noise may appear; and second - the sample is physically diluted during precipitation, lowering the method sensitivity. Thus, the working parameters in developing an analytical method are related to the performance needed: the sample preparation time and costs, the method speed and sensitivity.

Although after oral administration of memantine in humans, its maximum plasma levels are about 15-25 ng/ml [6-9], after intravenous administration (IV) in pigs the maximum plasma levels of memantine are much higher, about

400 ng/mL (pilot study, unpublished data). In that case the calibration curve range was adapted to the concentration range of the samples to be analyzed, in our case 4.86-486 ng/ml. Because the current assay does not require a high sensitivity as in human pharmacokinetic studies, the PP extraction method becomes an attracting alternative to SPE or LLE due the high speed and the high reproducibility of the extraction.

Table 1. Analytical characteristics of reported LC/MS and LC/MS/MS methods for the determination of memantine in biological matrix

Matrix	Pre-treatment/ extraction[a]	Mobile phase constituents[b]	Detection mode[c]	LOQ[d] (ng/mL)	Run time (min)	Reference
Human plasma	SPE	ACN - ammonium acetate buffer	ESI-MS/MS, SRM (m/z 180→163)	0.2	2	**6**
Human plasma	LLE	MeOH- formic acid in water	APCI-MS SIM, m/z 180	0.2 L	6	**7**
Human plasma	LLE	MeOH– formic acid in water	ESI-MS/MS, SRM (m/z 180→163)	0.1	2	**8**
Human plasma	LLE	MeOH– formic acid in water	ESI-MS/MS, SRM (m/z 180→107)	0.1	4	**9**
Rat plasma	Cloud point extraction	MeOH– formic acid in water	APCI-MS SIM, m/z 180	2	6	**10**

[a] SPE, solid-phase extraction; LLE, liquid–liquid extraction;
[b] ACN, acetonitrile; MeOH, methanol;
[c] ESI, electrospray ionisation; APCI, atmospheric pressure chemical ionisation; SIM, selected ion monitoring; SRM, selected reaction monitoring;
[d] LOQ, limit of quantification

In our method we analysed volumes of only 0.2 mL plasma by PP with 7% perchloric acid (0.1 mL) and direct injection into the chromatographic system from supernatant after centrifugation. We obtained a sensitivity corresponding to our needs (LLOQ of 4.86 ng/mL) and absolute recoveries between 92.5-100.8%, this being the first analytical method using the PP as plasma preparation procedure.

LC-MS assay

The analyte detection was optimized in several trials to achieve maximum sensitivity and specificity. The memantine is ionized in ESI source by proton addition, giving a pseudo-molecular ion with m/z 180. After fragmentation, the protonated memantine looses an ammonia molecule (MW 17 amu) and is converted to ion with m/z 163 (Figure 2). However, although in MS[2] mode the method specificity is increased in comparison with MS[1] mode, the M-17 transition is not specific. The reason is that fragmentation of ions occurs in fact in a narrow window of about 2 amu around the selected mass, so in case of

memantine the entire range of m/z 180±2 ions are prone to collision-induced dissociation processes. In that case all the ions in that range having a loss of 16 amu (methane), 17 amu (ammonia) and 18 amu (water) may interfere the analysis.

In order to obtain the needed specificity of analysis and a maximum signal-to-noise ratio (S/N) of analyte, we used the capability of Ion Trap mass spectrometer to do multiple stages isolation-fragmentation processes, that means MS^n analysis. This feature is specific to Ion Trap MS analyzers, other MS systems (single quadrupole, triple quadrupole, time of flight) don't have that capability. Thus, in a MS^3 stage the ion with m/z 163 obtained in the MS^2 stage was further fragmented and the obtained mass spectra was recorded (Figure 2). The obtained ions with m/z 107, 121 and 135 are specific to memantine and were used for quantification.
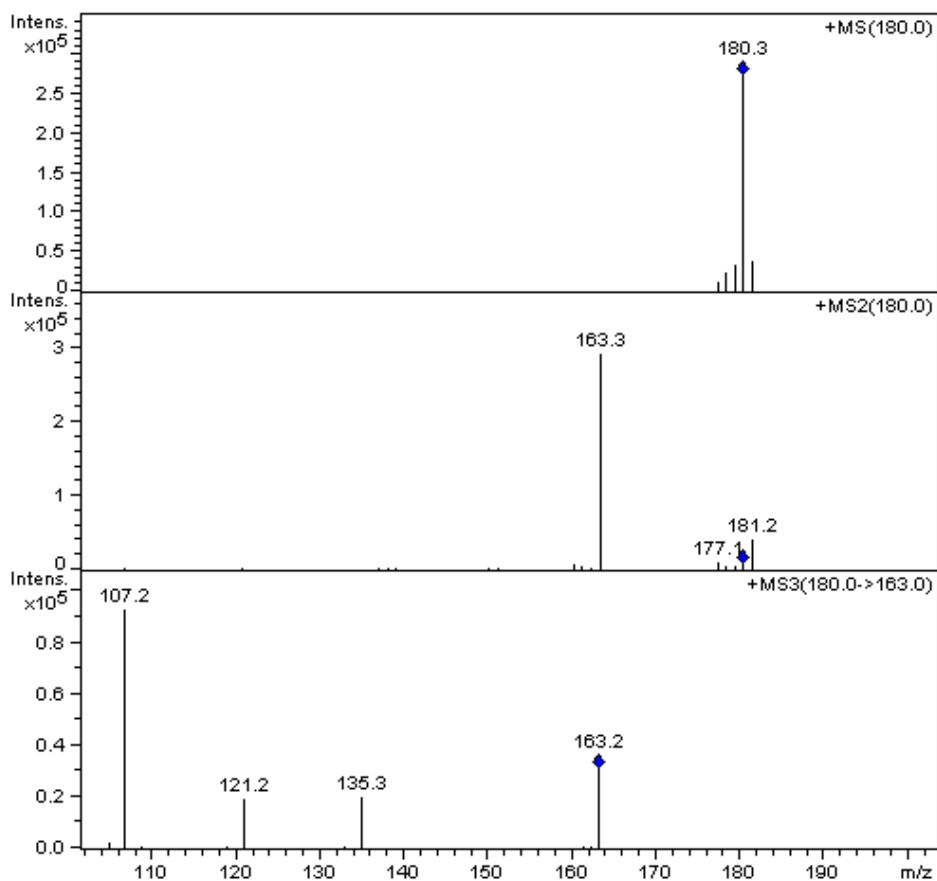
Figure 2. MS spectra of memantine – MS^1 spectra (upper image), MS^2 spectra (middle image) and MS^3 spectra (lower image)

Comparative chromatograms of memantine detected in MS[1], MS[2] and MS[3] mode are presented in Figure 3. By using MS[3] detection mode, due to high specificity, the overall method sensitivity is increased about 6 times in comparison with MS[1] detection.

The detection of memantine was carried out in multiple reaction monitoring (MRM). The extracted ion chromatogram (EIC) of m/z (107, 121, 135) from m/z 180 was analyzed. In the selected chromatographic conditions the retention time of memantine was 1.95 min.
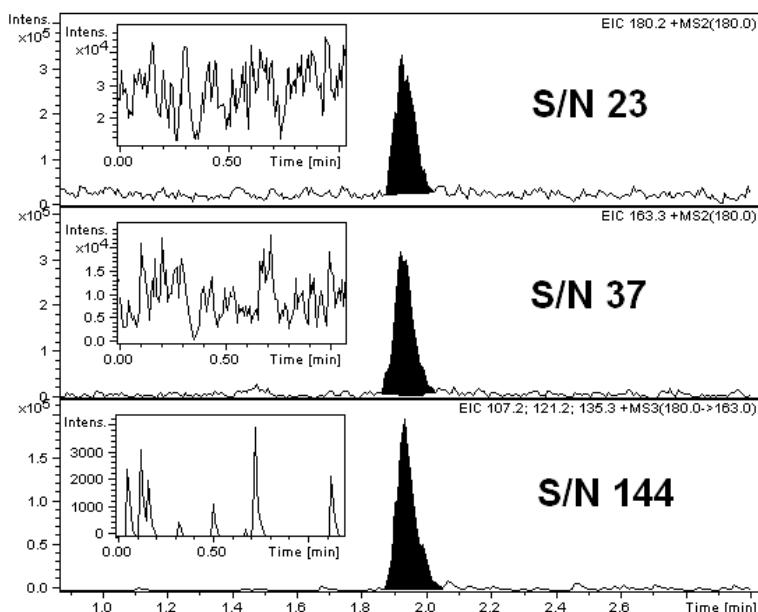
Figure 3. Chromatograms of memantine, same concentration, different MS detection modes: MS[1] (upper image), MS[2] (middle image) and MS[3] (lower image). In insert, the typical background noise pattern for each case, in absolute scale (ion abundance)

Assay validation

The method was validated in accordance with international regulations [**11,12**]. A representative chromatogram of pig plasma spiked with memantine at LLOQ is shown in Figure 4. No interfering peaks from the endogenous plasma components were observed at the retention time of memantine.

The calibration curves were linear over the concentration range of 4.86 – 486 ng/mL in pig plasma, with a correlation coefficient greater than 0.995. The LLOQ was 4.86 ng/mL. The values obtained for intra-day and inter-day precision and accuracy during the validation are shown in Tables 2 and 3, respectively. All values for accuracy and precision were within guidelines recommended limits (<15%) [**11,12**]. The absolute recovery values were between 92.5-100.8%.
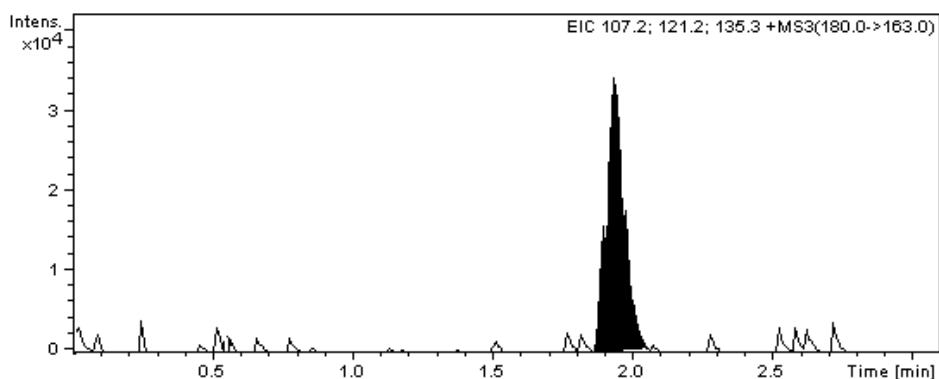
Figure 4. Representative chromatogram of pig plasma sample spiked with memantine at lower limit of quantification (4.86 ng/ml) (retention time – 1.95 min)

Table 2. Intra-day precision, accuracy and recovery (n = 5) for memantine

$c_{nominal}$ ng/mL	Mean c_{found} ng/mL (± S.D.)	CV %	Bias %	Recovery % (± S.D.)
4.86	4.73±0.58	12.2	-2.5	97.1±8.3
14.57	16.32±0.49	3.0	12.1	100.0±2.6
97.10	92.06±3.49	3.8	-5.2	96.2±3.3
194.20	178.85±6.73	3.8	-7.9	92.5±2.9

Table 3. Inter-day precision, accuracy and recovery (n = 5) for memantine

$c_{nominal}$ ng/mL	Mean c_{found} ng/mL (± S.D.)	CV %	Bias %	Recovery % (± S.D.)
4.86	4.94±0.37	7.5	1.7	100.0±5.7
14.57	16.09±0.66	4.1	10.5	100.8±6.5
97.10	96.05±3.52	3.7	-1.1	99.8±7.3
194.20	180.09±11.42	6.3	-7.3	98.1±7.7

Method application

The validated method for determination of memantine in pig plasma was successfully applied in a pharmacokinetic study of intravenously administered memantine in pigs.

CONCLUSION

Our developed LC/MS[3] assay is simple, rapid, specific, accurate and not expensive. This is the first published analytical method for analysis of memantine in biological matrix using protein precipitation as plasma processing method. This new fast and specific method was successfully applied in a pharmacokinetic study of intravenously administered memantine in pigs.

EXPERIMENTAL SECTION

Reagents

Memantine was reference standard from USP (Rockville, MD, USA). Methanol of gradient grade for liquid chromatography, formic acid and 70% perchloric acid of analytical-reagent grade were purchased from Merck KGaA (Darmstadt, Germany). Bidistilled, deionised water *pro injections* was purchased from Infusion Solution Laboratory of University of Medicine and Pharmacy Cluj-Napoca (Romania). The pig blank plasma was from drug-untreated pigs.

Apparatus

The following apparatus were used: 204 Sigma Centrifuge (Osterode am Harz, Germany); Analytical Plus and Precision Standard Balance (Mettler-Toledo, Switzerland); Vortex Genie 2 mixer (Scientific Industries, New York, USA); Ultrasonic bath Elma Transsonic 700/H (Singen, Germany). The HPLC system used was an 1100 series Agilent Technologies model (Darmstadt, Germany) consisting of one G1312A binary pump, an in-line G1379A degasser, an G1329A autosampler, a G1316A column oven and an Agilent Ion Trap Detector 1100 SL.

Chromatographic and spectrometric conditions

Chromatographic separation was performed on a Zorbax SB-C18 (100 mm x 3.0 mm i.d., 3.5 μm) column (Agilent Technologies) under isocratic conditions using a mobile phase of a 55:45 (v/v) mixture of methanol and 0.1% (v/v) formic acid in water at 45 ºC with a flow rate of 1 mL/min. In order to maintain the ESI source clean, the column effluent was diverted to waste for the first 1.5 minutes after injection. The detection of memantine was performed in multiple reaction monitoring (MRM) mode using an ion trap mass spectrometer with an electrospray ion (ESI) source, positive ionisation (capillary 3500 V, nebulizer 60 psi (nitrogen), dry gas nitrogen at 12 L/min, dry gas temperature 350ºC). The extracted ion chromatogram (EIC) of m/z (107, 121, 135) from m/z 163 from m/z 180 was analyzed (MS[3] mode).

Standard solutions

A stock solution of memantine (0.971 mg/mL) was prepared by dissolving an appropriate quantity of memantine in methanol. Two working solutions (19.42 µg/mL and 486 ng/mL, respectively) were prepared by appropriate dilution in drug-free pig plasma. These solutions were used to prepare eight plasma calibration standards with the concentrations between 4.86 and 486 ng/mL, respectively. Quality control (QC) samples of 14.6 ng/mL (lower), 97.1 ng/mL (medium) and 194.2 ng/mL (higher) were prepared by adding the appropriate volumes of working solutions to drug-free pig plasma.

Sample preparation

Standards and plasma samples (0.2 mL) were deproteinized with a 7% perchloric acid aqueous solution (0.1 mL). After vortex-mixture (10 s) and centrifugation (3 min at 12000 rpm), the supernatants (0.15 mL) were transferred in autosampler vials and 10 µL were injected into the HPLC system.

Method validation
The specificity of the method was evaluated by comparing the chromatograms obtained from the plasma samples containing memantine with those obtained from plasma blank samples.

The concentration of memantine was determined automatically by the instrument data system using peak areas and the external standard method. The calibration curve model was determined by the least squares analysis: y = b + ax, weighted (1/y) linear regression, where y - peak area of the analyte and x - concentration of the analyte (ng/mL).

The intra-day precision (expressed as coefficient of variation, CV %) and accuracy (expressed as relative difference between obtained and theoretical concentration, bias %) were determined by analysis of five different samples (n = 5) from each QC standards (at lower, medium and higher levels) on the same day. The inter-day precision and accuracy were determined by analysis on five different days (n = 5) of one sample from each QC standards (at lower, medium and higher levels). The lower limit of quantification (LLOQ) was established as the lowest calibration standard with an accuracy and precision less than 20%. The relative recoveries (at LLOQ, lower, medium and higher levels) were measured by comparing the response of the spiked plasma with the response of standards in solvent with the same concentration of memantine as the plasma (n = 5) [**11-15**].

ACKNOWLEDGMENTS

REFERENCES

1. R.G. Geocadin, M.A. Koenig, R.D. Stevens, M.A. Peberdy, *Crit Care Clin*, **2006**, *22*, 619.
2. S.A. Lipton, *Nature Rev*, **2006**, *5*, 160.
3. S. Varanese, J. Howard, A. Di Rocco, *Mov Disord*, **2010**, *25(4)*, 508.
4. R.S. Green, D.W. Howes, *Can Med Assoc J*, **2007**, *176*, 759.
5. H.J. Leis, G. Fauler, W. Windischhofer, *J Mass Spectrom*, **2002**, *37*, 477.
6. S.K. Dubey, A. Patni, A. Khuroo, N.R. Thudi, S. Reyar, A. Kumar, et al., *E-J Chem*, **2009**, *6(4)*, 1063.
7. M.Y. Liu, S.N. Meng, H.Z. Wu, S. Wang, M.J. Wei, *Clin Ther*, **2008**, *30(4)*, 641.
8. A.A. Almeida, D.R. Camposa, G. Bernasconi, S. Calafatti, F.A.P. Barros, M.N. Eberlin, et al., *J Chrom B*, **2007**, *848*, 311.
9. R.N. Pan, T.Y. Chian, B.P. Kuo, L.H. Pao, *Chromatogr.*, **2009**, *70(5-6)*, 783.
10. W. Liu, K. Bi, X. Liu, J. Zhao, X. Chen, *Chromatogr.*, **2009**, *69(9-10)*, 837.
11. Guidance for Industry, Bioanalytical Method Validation. U.S. Department of Health and Human Services, Food and Drug Administration. Federal Register, **2001**, *66*.
12. Guidance on the Investigation of Bioavailability and Bioequivalence. The European Agency for the Evaluation of Medicinal Products, Committee for Proprietary Medicinal Products, **2001**, CPMP/EWP/QWP/1401/98.
13. M. Achim, D. Muntean, L. Vlase, I. Bâldea, D. Mihu, S. E. Leucuţa, *Studia Univ. Babeş-Bolyai, Chemia*, **2009**, *54(3)*, 7.
14. D. S. Popa, L. Vlase, S. E. Leucuţa, F. Loghin, *Farmacia*, **2009**, *57(3)*, 301.
15. D. Mihu, L. Vlase, S. Imre, C.M. Mihu, M. Achim, D.L. Muntean, *Studia Univ. Babeş-Bolyai, Chemia*, **2009**, *54(3)*, 151.

CISPLATIN EFFECT ON HEMOGLOBIN AND MYOGLOBIN AUTOOXIDATION

CRISTINA BISCHIN[a], VICENTIU TACIUC[a], RADU SILAGHI-DUMITRESCU*[a]

ABSTRACT. It has previously been shown using mass-spectrometry that therapeutically-useful platinum-based compounds of the cisplatin family bind to a range of proteins, including hemoglobin and cytochrome *c*. On the other hand, currently available data suggests that the medically-relevant chemical properties of hemoglobin are in no significant way affected by Pt-based drugs, within the concentration ranges attainable during treatment of a patient. Here, cisplatin is shown to modulate two physiological parameters of hemoglobin and myoglobin - autooxidation and oxygen dissociation rate, as measured by UV-vis aspectra. The extent, to which these changes in reactivity impact the side-effects and even the therapeutic mechanisms of these drugs, appears to deserve further attention.

Keywords: hemoglobin, cytochrome c, cisplatin, autooxidation

INTRODUCTION

Cisplatin and related compounds are known to exert much of their useful therapeutic effects via binding to DNA.[1] The need for more effective drugs as well as the wide range of side-effects (nausea, progressive peripheral sensory neuropathy, fatigue, vomiting, alopecia, hematological suppression, renal damage) have, for several decades now, fuelled interest into understanding the complex mechanisms of interaction of cisplatin and related compounds with various biomolecules.[2,3] One notable observation in this respect has been that platinum can bind to a range of proteins, as demonstrated by elemental analyses, chromatography and mass spectrometry.[3,4] With such methods, cisplatin–Hb complexes were shown to be formed using clinically relevant concentrations of cisplatin and Hb. The interaction of oxaliplatin and carboplatin with hemoglobin was also studied with nanoelectrospray ionization quadrupole time-of-flight mass spectrometry (nanoESI-QTOF-MS) and size-exclusion high performance liquid chromatography/inductively coupled plasma mass spectrometry (HPLC/ICPMS) [5], showing similar results to those obtained with cisplatin; heme release was a noted side effect of platinum binding. Cisplatin-derived platinum was also found bound to cytochrome, with a preference for the iron ligand Met 65.[6]

Hb is present in high concentrations in blood and is particularly sensitive to changes in redox status, to the extent that under stress conditions such as physical effort or certain pathological conditions it engages in toxic reactions

---

[a] *Department of Chemistry and Physics, Babes-Bolyai University, 11 Arany Janos street, Cluj-Napoca 400028, Romania,*corresponding author; email: * rsilaghi@chem.ubbcluj.ro; Tel.: +402645903833; FAX: +40264590818*

with oxidative stress agents- primarily peroxide – yielding free radicals and highly-oxidizing states at the iron (ferryl, Compound II).[7-10] As such, it may be expected that Hb might be sensitive to the stress imposed by cisplatin in patients; indeed, in a preliminary report we have shown that the autooxidation rate of hemoglobin is affected by cisplatin and related compounds.[11] Cytochrome *c* not only serves as key component of the electron transport chain, but is also involved, most likely via redox reactions once again linked to peroxides, in the apoptosis process and as such would appear as a sensitive target for exogenous compounds such as cisplatin.[12-14] Here, data is shown establishing a link between previously-demonstrated binding of platinum to hemoglobin and myoglobin, and their reactivity towards dioxygen with possible relevance to the *in vivo* reactions of platinum-based drugs

RESULTS AND DISCUSSION

Table 1 illustrates that, in agreement with a previous preliminary report,[11] cisplatin enhances the rate of autooxidation of hemoglobin, as measured under two different conditions – at 37°C in pH 7.4 buffer or at room temperature under acceleration by guanidinium hydrochloride. The latter procedure is interpreted to effectively probe the stability of the protein: binding of platinum to the protein surface (previously proven by others using mass spectrometry [5]) is expected to affect the local mobility of the protein side-chains and hence the stability of the tertiary structure. As seen in Table 1, guanidine accelerates autooxidation; this is proposed to be due to partial denaturation of the protein induced by guanidine, which would result in opening the heme site towards solvent. Indeed, solvent accessibility and polarity at the metal site is a key parameter dictating autooxidation parameters in dioxygen carriers.[15] Under these conditions, it was expected that in the presence of guanidine the Pt-treated Hb would show slightly increased autooxidation – as confirmed by the data in Table 1. One may note, from a methodological point of view, that our approach of measuring globin autooxidation rate in the presence of a controlled amount of guanidine offers the advantage of measuring this rate somewhat faster than traditionally done at 37°C in phosphate buffer saline. While the absolute values obtained for autooxidation in the presence of guanidine at room temperature (shown in Table 1) lack direct physiological relevance, they do allow for a faster way of assessing the effect of exogenous factors on the autooxidation rate and on stability in general - such as exemplified here by the cisplatin case. Indeed, while traditional measurements of autooxidation rates in dioxygen-carrying proteins, done at 37°C and in PBS, may take several hours, the guanidine procedure requires no incubation and can be controlled, by choosing the appropriate guanidine concentrations, to proceed in only a few minutes.

Table 2 illustrates that cisplatin affects the rate of dioxygen dissociation from the hemoglobin ferrous heme (k_{off}). The 14-26% decreases brought by cisplatin compared to native Hb may be interpreted to suggest slightly reduced mobility of protein side-chains controlling dioxygen liberation. With myoglobin

314

the k_{off} rate was affected somewhat similar to hemoglobin (data not shown). Figure 1 illustrates UV-vis spectra collected during data collection for k_{off} measurements.

Table 1. The effect of cisplatin on Hb and Mb autooxidation rate as induced by temperature or by 2.5 M guanidine. Percentage increases or decreases relative to values seen for the respective proteins in the absence of cisplatin are shown.

	Hb (%)	Mb (%)
100 µM Pt	+ 2 (± 1)	+6 (± 1)
300 µM Pt	+ 34 ($\pm$ 1)	+11 (± 6)
30 µM Pt, guanidine	+48 ($\pm$ 13)	
100 µM Pt, guanidine	+13 ($\pm$ 5)	
300 µM Pt, guanidine	+33 ($\pm$ 1)	
500 µM Pt, guanidine	+36 ($\pm$ 1)	

Table 2. Effect of cisplatin on the dissociation rate of dioxygen from oxy-Hb. Percentage increases or decreases relative to values seen for Hb in the absence of cisplatin are shown.

	k_{off}
Hb, 200 µM Pt, inc	-23%$\pm$0.1
Hb, 200 µM Pt, non-inc	-26%$\pm$0.1
Hb, 400 µM Pt, inc 24 h (mM)	-14%+0.1
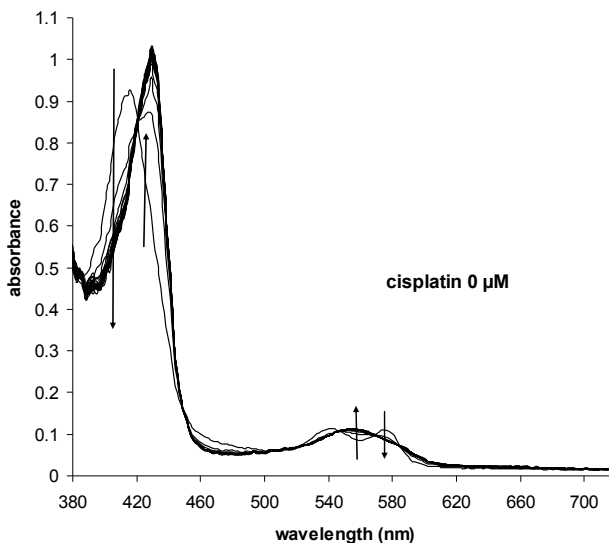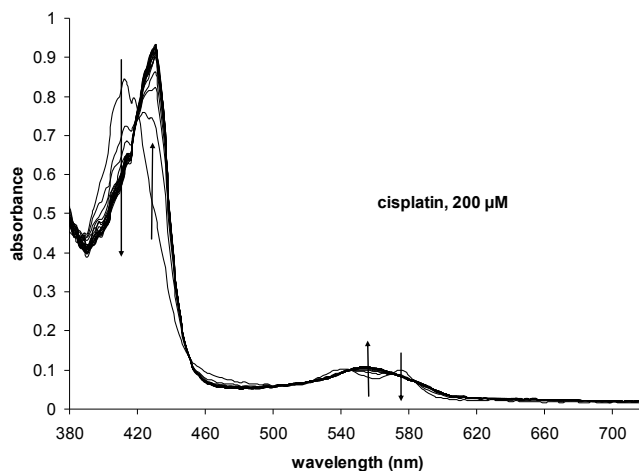
Figure 1. UV-vis spectra collected during a stopped-flow experiment where oxy-hemoglobin (14 µM) was mixed with dithionite (20 mM) in order to monitor the transition from the oxy to the deoxy state, in the presence or absence of 200 µM cisplatin. Spectra were collected every 13 ms for the first 2 seconds after mixing; arrows indicate the directions in which absorbance changes at various wavelengths.

CONCLUSIONS

Cisplatin affects the autooxidation rate and the oxygen dissociation rate in hemoglobin, more so than in myoglobin. The rate of dissociation of molecular oxygen from ferrous hemoglobin is, by contrast, slightly lowered by cisplatin. Together, these findings suggest that cisplatin is likely to limit the dioxygen-delivering abilities of hemoglobin. This is proposed to originate in a decrease in flexibility of the protein, as a result of platinum binding to the surface aminoacids. It remains to be established to what extent these findings have a clinically-relevant outcome.

METHODS

Bovine hemoglobin was purified from bovine blood as previously described.[16] Thus, the blood, freshly drawn on citrate, was centrifuged 15 minutes at 5000 rpm to separate the red blood cells, which were then washed three times with 5 mM phosphate pH 7.4 + 150 mM NaCl. Hemoglobin concentrations in text are given per heme rather than per tetramer. The met forms of the hemoglobins were prepared by ferricyanide treatment as previously described.[17-19] Where needed, guanidinium hydrochloride was added to reaction mixtures from 6 M stock solutions.

UV-vis spectra were recorded on Agilent 8453 (Agilent, Inc.) and Cary 50 (Varian, Inc) instruments. Stopped-flow spectra were collected on a Biologic SFM-300 system equipped with 3 syringes and capable of sequential mixing,

with a high-speed diode array detector. Stopped-flow data were analyzed within the SPECFIT32 software package (BioLogic Science Instruments, France) using Singular Value Decomposition (SVD) and global multiexponential fitting of the SVD treated data, with the spectra fitted to simple kinetic models using Levenberg-Marquardt or Simplex algorithms.

Autooxidation rates were measured at 37° C in PBS. Optionally, prior to measurements the samples were incubated at 4ºC overnight, with physiological serum (control sample) or with cisplatin dissolved in physiological serum at concentrations indicated in the Table 1. Alternatively, autooxidation measurements were measured at room temperature in the presence of 2.5 M guanidine, in which case pre-incubation with cisplatin did not affect the result.

ACKNOWLEDGMENTS

REFERENCES

1. S.E. Sherman,S.J. Lippard, *Chem. Rev.*, **1987**, *87*, 1153
2. J. Peng, R. Mandal,M. Sawyer, X. F. Li, *Clin. Chem.*, **2005**, *51*, 2274
3. R. Mandal, C. Teixeira, X. F. Li, *The Analyst*, **2003**, 629
4. A. Casini, A. Guerri,C. Gabbiani, L. Messori, *J. Inorg. Biochem.*, **2008**, *102*, 995
5. R. Mandal, M.B. Sawyer, X.F. Li, *Rapid Commun. Mass Spectrom.*, **2006**, *20*, 2533
6. T. Zhao, F.L. King, *J. Am. Soc. Mass. Spectrom.*, **2009**, *20*, 1141
7. R. Silaghi-Dumitrescu, B.J. Reeder, P. Nicholls, C.E. Cooper, M.T. Wilson, *Biochem. J.*, **2007**, *403*, 391
8. N.B. Vollaard, J.P. Shearman, C.E. Cooper, *Sports Med.*, **2005**, *35*, 1045
9. N.B. Vollaard, B.J. Reeder, J.P. Shearman, P. Menu, M.T. Wilson, C.E. Cooper, *Free Radic. Biol. Med.*, **2005**, *39*, 1216
10. B.J. Reeder, D.A. Svistunenko, C.E. Cooper, M.T. Wilson, *Antioxidant & Redox Signaling*, **2004**, *6*, 954
11. V. Taciuc, C. Bischin, R. Silaghi-Dumitrescu "A novel mechanism for platinum-based drugs: cisplatin and related compounds as pro-oxidants in blood", Silaghi-Dumitrescu, R. and Garban, G., Ed.; Cluj University Press: Cluj-Napoca, Romania, **2009**, pp 130
12. V.E. Kagan, G.G. Borisenko, Y.Y. Tyurina, V.A. Tyurin, J. Jiang, A.I. Potapovich, V. Kini, A.A. Amoscato, Y. Fujii, *Free Radic. Biol. Med.*, **2004**, *37*, 1963

13. V.E. Kagan, V.A. Tyurin, J. Jiang, Y.Y. Tyurina, V. B. Ritov, A.A. Amoscato, A.N. Osipov, N.A. Belikova, A.A. Kapralov, V. Kini, Vlasova, II, Q. Zhao, M. Zou, P. Di, D.A. Svistunenko, I.V. Kurnikov, G.G. Borisenko, *Nat. Chem. Biol.*, **2005**, *1*, 223

14. V.E. Kagan, H.A. Bayir, N.A. Belikova, O. Kapralov, Y.Y. Tyurina, V.A. Tyurin, J. Jiang, D.A. Stoyanovsky, P. Wipf, P.M. Kochanek, J.S. Greenberger, B. Pitt, A.A. Shvedova, G. Borisenko, *Free Radic. Biol. Med.*, **2009**, *46*, 1439

15. D.M. Kurtz, Jr., *Ess. Biochem.*, **1999**, 55

16. E. Antonini, M. Brunori, "Hemoglobin and Myoglobin in their Reaction with Ligands", North-Holland, Amsterdam, 1971.

17. B.J. Reeder, D.A. Svistunenko, M.A. Sharpe, M.T. Wilson, *Biochem.*, **2002**, *41*, 367

18. J. Dunne, D.A. Svistunenko, A.I. Alayash, M.T. Wilson, C.E. Cooper, *Adv. Exp. Med. Biol.*, **1999**, *471*, 9

19. J. Dunne, A. Caron, P. Menu, A.I. Alayash, P.W. Buehler, M.T. Wilson, R. Silaghi-Dumitrescu, B. Faivre, C.E. Cooper, *Biochem. J.*, **2006**, *399*, 513

CENTRIC CONNECTIVITY INDEX

MIRCEA V. DIUDEA[*]

ABSTRACT. Relative centricity *RC* values of vertices/atoms are calculated within the Distance and Cluj-Distance criteria. The vertex *RC* distribution in a molecular graph provides atom equivalence classes, useful in interpretation of NMR spectra. Timed by vertex valences, *RC* provides a new index, called Centric Connectivity *CC*, which can be useful in topological characterization of graphs and in QSAR/QSPR studies.

Keywords: *graph theory, Cluj index, relative centricity, centric connectivity index*

INTRODUCTION

Let $G = (V, E)$ be a connected graph, with no multiple bonds and loops. *V* is the set of vertices and *E* is the set of edges in *G*; $v = |V(G)|$ and $e = |E(G)|$ are their cardinalities.

A *walk w* is an alternating string of vertices and edges: $w_{1,n} = (v_1, e_1, v_2, e_2, ..., v_{n-1}, e_m, v_n)$, with the property that any subsequent pair of vertices represent an edge: $(v_{i-1}, v_i) \in E(G)$. Revisiting of vertices and edges is allowed [1-6].

The *length* of a walk, $l(w_{1,n}) = |E(w_{1,n})|$ equals the number of its traversed edges. In the above relation $E(w_{1,n})$ is the edge set of the walk $w_{1,n}$. The walk is *closed* if $v_1 = v_n$ and is *open* otherwise [3,5].

A *path p* is a walk having all its vertices and edges distinct: $v_i \neq v_j$, $(v_{i-1}, v_i) \neq (v_{j-1}, v_j)$ for any $1 \leq i < j \leq n$. As a consequence, revisiting of vertices and edges, as well as branching, is prohibited. The *length* of a path is $l(p_{1,n}) = |E(p_{1,n})| = |V(p_{1,n})| - 1$, with $V(p_{1,n})$ being the vertex set of the path $p_{1,n}$. A closed path is a *cycle* (i.e., *circuit*).

A path is *Hamiltonian* if all the vertices in *G* are visited at most once: $n = |V(G)|$. If such a path is closed, then it is a *Hamiltonian circuit.*

The *distance* d_{ij} is the length of a *shortest* path joining vertices v_i and v_j: $d_{ij} = min \; l(p_{ij})$; otherwise $d_{ij} = \infty$. The *set of all distances* (i.e., geodesics) in *G* is denoted by **DI**(*G*).

The *detour* δ_{ij} is the length of a *longest* path between vertices v_i and v_j: $\delta_{ij} = max \; l(p_{ij})$; otherwise $\delta_{ij} = \infty$. The *set of all detours* in *G* is denoted by **DE**(*G*).

The square arrays that collect the distances and detours, in *G* are called the *Distance* **DI** and *Detour* **DE** matrix**,** respectively [3,5]:

$$\left[\mathbf{DI}(G) \right]_{i,j} = \begin{cases} \min l(p_{i,j}), \text{if } \; i \neq j \\ 0 \text{ if } \; i = j \end{cases} \tag{1}$$

---

[*] *Faculty of Chemistry and Chemical Engineering, Babes-Bolyai University, 3400 Cluj, Romania*

$$| \mathbf{DE}(G) |_{i,j} = \begin{cases} \max l(p_{i,j}), if \;\; i \neq j \\ 0 \quad if \;\; i = j \end{cases} \tag{2}$$

In words, these matrices collect the number of edges separating the vertices i and j on the shortest and longest path $p_{i,j}$, respectively. The half sum of entries in the Distance and Detour matrices provide the well-known Wiener index W [7] and its analogue, the detour number w [8,9].

The *Cluj fragments* are sets of vertices obeying the relation [3,5,10-13]:

$$CJ_{i,j,p} = \left\{ v \middle| v \in V(G); D_{(G-p)}(i,v) < D_{(G-p)}(j,v) \right\} \tag{3}$$

The entries in the Cluj matrix **UCJ** are taken, by definition, as the maximum cardinality among all such fragments:

$$[\mathbf{UCJ}]_{i,j} = \max_p \left| CJ_{i,j,p} \right| \tag{4}$$

It is because, in graphs containing rings, more than one path can join the pair (i, j), thus resulting more than one fragment related to i (with respect to j and path p).

The Cluj matrix is defined by using either *distances* or *detour*s [14]: when path p belongs to the set of distances DI(G), the suffix DI is added to the name of matrix, as in UCJDI. When path p belongs to the set of detours DE(G), the suffix is DE.

Two graphs are called *isomorphic*, G $\approx$ G', if there exists a mapping $f : V \rightarrow V'$ that preserves adjacency (*i.e.*, if $(i,j) \in$ E(G), then $(f(i), f(j)) \in$ E'(G')). The function f provides a one-to-one correspondence between the vertices of the two sets. The isomorphism of G with itself is called an automorphism. It is demonstrated that all the automorphisms of G form a group, *Aut*(G) [3,5].

The symmetry of a graph is often called a topological symmetry; it is defined in terms of *connectivity*, as a constitutive principle of molecules and expresses equivalence relationships among elements of the graph: vertices, bonds, faces or larger subgraphs. The topological symmetry does not fully determine molecular geometry and it does not need to be the same as (*i.e.*, isomorphic to) the molecular point group symmetry. However, it represents the maximal symmetry which the geometrical realization of a given topological structure may posses [15-17].

Given a graph $G=(V,E)$ and a group Aut(G), two vertices, $i, j \in$ V are called *equivalent* if there is a group element, aut(v_i)$\in$Aut(G), such that j aut(v_i) i. The set of all vertices j (obeying the *equivalence relation*) is called the i's class of equivalence. Two vertices i and j, showing the same vertex invariant $In_i = In_j$ belong to the same *invariant class IC*. The process of vertex partitioning in *IC*-s leads to m classes, with $v_1, v_2,...v_m$ vertices in each class. Note that invariant-based partitioning may differ from the orbits of automorphism since no vertex invariant is known so far to discriminate two non-equivalent vertices in any graph [3,5].

In the chemical field, the isomorphism search could answer to the question if two molecular graphs represent or not one and the same chemical compound. Two isomorphic graphs will show the same topological indices, so that they cannot be distinguished by topological descriptors.

CENTRIC CONNECTIVITY CC INDEX

In studies on the centrality/centricity of graphs, Bonchev *et al*. [18,19] have proposed the distance-based criteria *1D-3D* as follows:

1D*: minimum vertex eccentricity: min ecc_i

2D*: minimum vertex distance sum*: min DIS_i

3D*: minimum number of occurrence of the largest distance*:

min **[LM, ShM]**$_{i,j\,\text{max}}$

When applied hierarchically, the above criteria lead to the center(s) of a graph.

Our older centrality index C(**LM**, **ShM**) is a function also giving the graph center(s), used alone or within the MOLORD algorithm [20]. In the above, **LM**, **ShM** denote the layer matrix and the shell matrix (of a given square info-matrix M), defined as follows [21-23].

The entries in the layer matrix (of vertex property) **LM**, are

$$\left[\mathbf{LM}\right]_{i,k} = \sum_{v\mid d_{i,v}=k} p_v \tag{5}$$

Layer matrix is a collection of the above defined entries:

$$\mathbf{LM}(G) = \left\{ [\mathbf{LM}]_{i,k}\,; i \in V(G); k \in [0,1,..,d(G)] \right\} \tag{6}$$

with $d(G)$ being the diameter of the graph (*i.e*., the largest distance in G). Any atomic/vertex property can be considered as p_i. More over, any square matrix **M** can be taken as *info matrix*, *i.e*., the matrix supplying local/vertex properties as row sum *RS*, column sum *CS*. The zero column is just the column of vertex properties $\left[\mathbf{LM}\right]_{i,0} = p_i$. When the vertex property is 1 (*i.e*., the counting property), the **LM** matrix will be **LC** (the Layer matrix of Counting).

Define the entries in the shell matrix **ShM** (of pair vertex property) as [23]

$$\left[\mathbf{ShM}\right]_{i,k} = \sum_{v\mid d_{i,v}=k} [\mathbf{M}]_{i,v} \tag{7}$$

The shell matrix is a collection of the above defined entries:

$$\mathbf{ShM}(G) = \left\{ [\mathbf{ShM}]_{i,k}\,; i \in V(G); k \in [0,1,..,d(G)] \right\} \tag{8}$$

A shell matrix **ShM**(G), will partition the entries of the square matrix according to the vertex (distance) partitions in the graph. It represents a true decomposition of the property collected by the info square matrix according to the contributions brought by pair vertices pertaining to shells located at distance k around each vertex. The zero column entries $\left[\mathbf{ShM}\right]_{i,0}$ are just the diagonal entries in the info matrix.

In this paper, the distance-based functions, expressing the topology related to the center of the graph, are as follows:

$$EP(i) = \sum_k P(i)_k \cdot k^n; k = 1,2,...d(G); n = 1,2,.. \tag{9}$$

$$P(i)_k = [\mathbf{LM}, \mathbf{ShM}]_{i,k} \qquad (10)$$

$$C(i) = (EP(i))^{-1} \qquad (11)$$

$$RC(i) = C(i) / C(i)_{max} = EP(i)_{min} / EP(i) \qquad (12)$$

$$RC(G) = \sum_i RC(i) \qquad (13)$$

$$CC(i) = RC(i) \cdot d(i) \qquad (14)$$

$$CC(G) = \sum_i CC(i) \qquad (15)$$

The distance-extension of the property $P(i)$ (collected in **LM** or **ShM**, (10)) is made by a variable power function, depending of the info matrix M, to ensure the separation of the resulting values, of which meaning is that of an eccentric property $EP(i)$ (9). There is a clear difference between $EP(i)$ and the eccentricity $\varepsilon(i)$ (counting the largest topological distance from i to any other vertex in G), used in the construction of "Eccentric Connectivity index" [24]. The vertex centricity $C(i)$ (11) is then calculated in the sense of the Bonchev's 1D-3D criteria, by virtue of the involved **LM, ShM** matrices.

The relative centricity (or centrality) $RC(i)$ (12) accounts for the deviation to the maximum centrality, equaling 1 in case of vertices being centers of the graph. The global value RC(G) (13) is useful in characterizing the distribution of the centricity function (11), particularly when normalized by the number of vertices of G.

Finally, the centric connectivity CC index (14,15) is hoped to be useful in QSAR/QSPR studies, their values being of the same order of magnitude as the number of vertices/atoms in the molecular graph. Relation (14) can be generalized by changing $d(i)$ by the "remote" degree [5,25] or by degrees of "extended connectivity" [26-30].

Tables 1 and 2 exemplify the above formulas for the molecular graphs illustrated in Figure 1. The sum in the EP(i) column gives twice the Wiener index [7]. Note that G_2 (Table 1) is a self-centered graph [31], of which all vertices are centers of the graph, as ranked by the RC(i) column. Also note that G_2 is a full Hamiltonian detour graph [14]; this means that all its detours are Hamiltonian path, visiting once all the vertices of the graph.

The invariant classes of equivalence IC-s are given at the bottom of tables, by their population Pop (no. of vertices in each class). IC-s are important in NMR spectra interpretation.
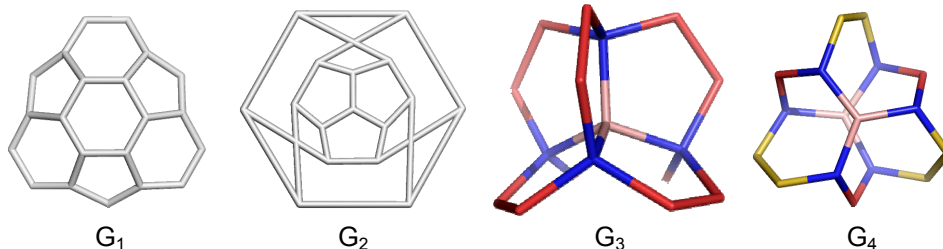
Figure 1. G_1 (v=21); G_2 (v=16); G_3 (v=17); G_4 (v=17)

Table 1. Vertex eccentric property EP(i), relative centricity RC(i) and centric connectivity CC(i) values, calculated on **LM=LC** matrix (k^1), unless otherwise specified; no. of "invariant classes" IC-s, by their population Pop (no. of vertices) for G_1 and G_2.

#	G_1				G_2 (self-centered graph)			
	EP(i)	RC(i)	CC(i)	d(i)	EP(i)	RC(i)	CC(i)	CC(i); (**ShCJDI**; k^2)
1	70	0.7714	1.5429	2	33	1	3	2.8385
2	70	0.7714	1.5429	2	33	1	3	2.9611
3	70	0.7714	1.5429	2	33	1	3	2.9358
4	54	1	3	3	33	1	3	2.9358
5	64	0.8438	2.5313	3	33	1	3	3
6	74	0.7297	1.4595	2	33	1	3	2.9611
7	64	0.8438	2.5313	3	33	1	3	2.9358
8	74	0.7297	1.4595	2	33	1	3	3
9	64	0.8438	2.5313	3	33	1	3	2.9611
10	54	1	3	3	33	1	3	2.9358
11	64	0.8438	2.5313	3	33	1	3	3
12	54	1	3	3	33	1	3	3
13	64	0.8438	2.5313	3	33	1	3	2.9358
14	74	0.7297	1.4595	2	33	1	3	2.9358
15	64	0.8438	2.5313	3	33	1	3	3
16	74	0.7297	1.4595	2	33	1	3	3
17	54	1	3	3	528=2W	16	48=CC	47.3364=CC
18	54	1	3	3				
19	74	0.7297	1.4595	2				
20	54	1	3	3				
21	74	0.7297	1.4595	2				
Sum	1362=2W	17.7552	46.5728=CC					
Pop			3,(6)$_3$					1,3,(6)$_2$

Table 2. Vertex eccentric property EP(i), relative centricity RC(i) and centric connectivity CC(i) values, calculated on **ShM=ShUCJDI** matrix (k^3); no. of "invariant classes" *IC*-s, by their population Pop (no. of vertices) for G_3 and G_4;

#	G_3				G_4			
	EP(i)	RC(i)	CC(i)	d(i)	EP(i)	RC(i)	CC(i)	d(i)
1	2478	0.5585	1.1170	2	2483	0.7962	2.3886	3
2	2478	0.5585	1.1170	2	2483	0.7962	2.3886	3
3	2044	0.6771	2.7084	4	2483	0.7962	2.3886	3
4	1384	1	4	4	1977	1	3	3
5	2044	0.6771	2.7084	4	3322	0.5951	1.1902	2
6	2478	0.5585	1.1170	2	2982	0.6630	1.3260	2
7	2478	0.5585	1.1170	2	2483	0.7962	2.3886	3
8	2044	0.6771	2.7084	4	2982	0.6630	1.3260	2
9	2478	0.5585	1.1170	2	2483	0.7962	2.3886	3
10	2478	0.5585	1.1170	2	3322	0.5951	1.1902	2
11	2044	0.6771	2.7084	4	2483	0.7962	2.3886	3
12	2478	0.5585	1.1170	2	1977	1	3	3
13	2478	0.5585	1.1170	2	3322	0.5951	1.1902	2
14	2478	0.5585	1.1170	2	2982	0.6630	1.3260	2
15	2478	0.5585	1.1170	2	3322	0.5951	1.1902	2
16	2478	0.5585	1.1170	2	3322	0.5951	1.1902	2
17	2478	0.5585	1.1170	2	3322	0.5951	1.1902	2
Sum	39296	10.4106	28.2380		47730	12.3370	31.4512	
Pop			1,4,12				2,3,(6)$_2$	

323

CONCLUSIONS

The relative centricity RC of vertices in a molecular graph were calculated within the Distance and Cluj-Distance criteria. The vertex RC distribution in a molecular graph gives information on the equivalence classes (as vertex invariant classes) of atoms, useful in interpretation of NMR spectra. Timed by vertex valences, RC provides an index, called Centric Connectivity CC. This index represents a new descriptor, which can be useful in topological characterization of graphs and in QSAR/QSPR studies. By definition, there is a clear difference between the Centric Connectivity index and the older Eccentric Connectivity index or its newer versions [32-34].

REFERENCES

1. N. Trinajstić, *Chemical Graph Theory,* CRC Press: Boca Raton, FL, **1983**.
2. F. Harary, *Graph Theory*, Addison-Wesley, Reading, M.A., **1969**.
3. M.V. Diudea, I. Gutman, L. Janschi, *Molecular Topology*, Nova Science, Huntington, N.Y., **2001**.
4. M.V. Diudea, O. Ivanciuc, *Molecular Topology*, Comprex, Cluj, **1995** (in Romanian).
5. M.V. Diudea, *Nanomolecules and Nanostructures, Polynomials and Indices*, MCM, No. 10, Univ. Kragujevac and Fac. Sci. Kragujevac, Serbia, **2010**.
6. M.V. Diudea, M.S. Florescu, P.V. Khadikar, *Molecular Topology and Its Applications*, EFICON, Bucharest, **2006**.
7. H. Wiener, *J. Am. Chem. Soc.,* **1947**, *69*, 17.
8. I. Lukovits, *Croat. Chem. Acta,* **1996**, *69*, 873.
9. I. Lukovits and M. Razinger, *J. Chem. Inf. Comput. Sci.*, **1997**, *37*, 283.
10. M.V. Diudea, *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 300.
11. M.V. Diudea, *MATCH, Commun. Math. Comput. Chem.*, **1997**, *35*, 169.
12. M.V. Diudea, I. Gutman, *Croat. Chem. Acta,* **1998**, *71*, 21.
13. M.V. Diudea, *Croat. Chem. Acta,* **1999**, *72*, 835.
14. M.V. Diudea, B. Parv, I. Gutman, *J. Chem. Inf. Comput. Sci.,* **1997**, *37*, 1101.
15. C.Y. Hu, L. Xu, *Anal. Chim. Acta,* **1994**, *295*, 127.
16. G.S. Ezra, *Symmetry Properties of Molecules*, Lecture Notes in Chemistry 28, Springer, **1982**.
17. M. Razinger, K. Balasubramanian, M.E. Munk, *J. Chem. Inf. Comput. Sci.*, **1993**, *33*, 197.
18. D. Bonchev, A.T. Balaban, M. Randić, *Int. J. Quantum Chem.*, **1981**, *19*, 61.
19. D. Bonchev, O. Mekenyan, A.T. Balaban, *J. Chem. Inf. Comput. Sci.*, **1989**, *29*, 91.
20. M.V. Diudea, D. Horvath, D. Bonchev, *Croat. Chem. Acta*, **1995**, *68*, 131.
21. M.V. Diudea, *J. Chem. Inf. Comput. Sci.,* **1994**, *34*, 1064.
22. M.V. Diudea, M. Topan, A. Graovac, *J. Chem. Inf. Comput. Sci.*, **1994**, *34*, 1071.
23. M.V. Diudea, O. Ursu, *Indian J. Chem.*, 42*A*, **2003**, 1283.
24. V. Sharma, R. Goswami, A.K. Madan, *J. Chem. Inf. Comput. Sci.*, **1997**, *37*, 273.
25. P.E. John, M.V. Diudea, *Carpath. J. Math.*, **2004**, *20* (2), 235.
26. A.T. Balaban, O. Mekenyan, D. Bonchev, *J. Comput. Chem.*, **1985**, *6*, 538.
27. A.T. Balaban, O. Mekenyan, D. Bonchev, *J. Comput. Chem.,* **1985**, *6*, 552.
28. O. Mekenyan, A.T. Balaban, D. Bonchev, *J. Magn. Reson.,* **1985**, *63*, 1.
29. A.T. Balaban, I. Moțoc, D. Bonchev, O. Mekenyan, *Top. Curr. Chem.*, **1993**, *114*, 21.
30. H Morgan, *J. Chem. Doc.*, **1965**, *5*, 107.
31. F. Buckley, *Ann. New York Acad. Sci.*, **1989**, *576*, 71.
32. A. Ilić, I. Gutman, *MATCH Commun. Math. Comput. Chem.*, **2011**, *65*, 731.
33. A. Ilić, in: I. Gutman, B. Furtula, "Novel Molecular Structure Descriptors -Theory and Applications II", MCM 9, University of Kragujevac, **2010**.
34. G. Yu, L. Feng, A. Ilić, *J. Math. Anal. Appl.*, **2011**, *375*, 934.