Quantum algebraic–combinatoric study of the conformational properties of n-alkanes. I

Gyula Tasi [a,b] and Fujio Mizukami [a]

[a] *Department of Surface Chemistry, National Institute of Materials and Chemical Research,*
1-1 Higashi, Tsukuba, Ibaraki 305-8565, Japan
[b] *Applied Chemistry Department, József Attila University, Rerrich B. tér 1, H-6720 Szeged, Hungary*

Based on quantum chemical calculation results, four rules were previously derived for the numbers and the sequences of the conformers of free n-alkane molecules. This paper builds up first an algebra to handle the conformational problem of n-alkanes. Partitioning the set of all sequences, the whole problem is then subdivided into three independent subcases. With the help of an equivalence relation, the sequences can be classified. According to the quantum chemical rules, certain equivalence classes do not represent conformers. A well-defined subcase of the whole problem is solved.

1. Introduction

A large number of papers have dealt with the conformational behaviour of n-alkanes. As regards the applied experimental techniques, microwave [13], infrared [12] and electron diffraction [2,7] investigations should be mentioned. For theoretical studies, algebraic and combinatoric [1,3–6,9,21], molecular mechanics [10,11] and quantum chemical [8,14–20,22] methods have been used.

In the algebraic and combinatoric studies, it is generally assumed that any isomer of n-alkanes can be embedded into the crystal lattice of diamond [1,3–6,9,21]. The embedded isomers (alkane systems, alkanoids), therefore, cannot have different C–C–C–C torsional angles from those occurring in the diamond lattice.

In several recent papers we analysed the conformational properties of n-alkanes by quantum chemical methods at various levels of theory [15–19]. With the help of an effective one-electron method (SEOEM) [16], all the existing conformers in the conformational space were determined up to undecane [19]. For the butane, pentane, hexane, heptane and octane molecules, calculations were also performed at Hartree–Fock (HF/6-31G*) and second-order Moeller–Plesset (MP2/6-311G**) levels [17,19]. Based on the results obtained for butane, pentane, hexane and heptane, four rules were derived, which exactly reproduced the quantum chemical calculations up to undecane [19]. The rules show that the most important factors governing the conformational behaviour of n-alkanes are the non-bonded repulsive–attractive (van der Waals)

interactions between the hydrogen atoms attached to the carbon atoms at positions 1,4, 1,5, 1,6 and 1,7.

A FORTRAN program (ENUMERAT) was written to determine the number and the sequences of the existing conformers of any n-alkane molecule [19]. With the help of similarity calculations, the structurally unique (non-isomorphic) conformers can be identified. The program generates first a list of 5^n items representing the available sequences. To obtain the number and the sequences of the existing conformers, the forbidden ones should then be eliminated according to the quantum chemical rules. This procedure, however, is neither neat nor fast enough.

The question now arises of whether it is possible to derive a simple explicit formula for the numbers of the conformers with algebraic and combinatoric tools. The present paper analyses this problem and affords a solution for a well-defined subcase of the whole problem. In the discussion to follow, methane, ethane and propane are excluded from the set of n-alkanes, because they have no free C–C–C–C torsional angle.

2.   Rules determining the sequences of the conformers

According to the quantum chemical calculations, there are two kinds of gauche C–C–C–C torsional angle in the isolated n-alkane molecules: one is around $\pm 60°$ and the other around $\pm 95°$ [19]. Let us introduce the following symbols for the various torsional angles: t, $\sim 180°$; g^+, $\sim +60°$; g^-, $\sim -60°$; x^+, $\sim +95°$; and x^-, $\sim -95°$. Since the diamond lattice contains t, g^+ and g^- torsional angles, the alkanoids can be used for modelling only a subset of the existing conformers.

From the quantum chemical calculation results, the following rules can be deduced for the sequences of the conformers [19]:

(1a)   the double sequences g^+g^- and g^-g^+ as well as

(1b) x^+x^- and x^-x^+ cannot occur,

(2)   the triple sequences $g^+x^-g^+$, $g^-x^+g^-$, $x^+g^+x^-$, $x^-g^+x^+$, $x^+g^-x^-$ and $x^-g^-x^+$ are likewise not allowed,

(3)   the quadruple sequences $x^+g^-g^-x^+$ and $x^-g^+g^+x^-$ are also forbidden and, finally,

(4) g^- should exist beside x^+ and, similarly, g^+ beside x^-: $\ldots x^+g^- \ldots$ or $\ldots g^-x^+ \ldots$ and $\ldots x^-g^+ \ldots$ or $\ldots g^+x^- \ldots$.

It is noteworthy that Bartell and coworkers recognised the validity of rule (1a) without performing any quantum chemical calculation [2,7]. They studied the conformational behaviour of butane, pentane, hexane, heptane and hexadecane in the gas phase by the electron diffraction method. To simulate the experimental data, t, g^+ and g^- C–C–C–C torsional angles were considered for the conformers and the sequences according to rule (1a) were omitted. For butane, pentane, hexane and heptane,

3, 7, 17 and 41 conformers were therefore considered, respectively [2]. A comparison of the experimental and simulation results led to the conclusion that further conformers might exist in traces in the gas phase. According to the quantum chemical studies, those conformers contain extended torsional angles x^+ and x^-.

3. Construction of an algebra for conformational studies of n-alkanes

Let us construct a set from the possible C–C–C–C torsional angles:

$$A = \left\{ t, g^+, g^-, x^+, x^- \right\}. \tag{1}$$

The sequences of the conformers of n-alkanes can be regarded as elements of the various Cartesian powers of set A. Let B denote the nth Cartesian power of A:

$$B = A^n = \left\{ (a_1, a_2, \ldots, a_n) \mid a_i \in A, \ 1 \leqslant i \leqslant n \right\}. \tag{2}$$

The order of B, i.e., the number of elements in set B is

$$|B| = |A|^n = 5^n. \tag{3}$$

For a particular n-alkane, n is the number of free C–C–C–C torsional angles in the molecule.

We can subdivide the conformational problem into different subcases making the following proper subsets of A:

$$A_1 = \left\{ t, g^+, g^- \right\}, \qquad A_2 = \left\{ t, x^+, x^- \right\}. \tag{4}$$

Let us make the nth Cartesian powers B_1 and B_2 of sets A_1 and A_2, respectively:

$$
\begin{aligned}
B_1 &= A_1^n = \left\{ (a_1, a_2, \ldots, a_n) \mid a_i \in A_1, \ 1 \leqslant i \leqslant n \right\}, \\
|B_1| &= |A_1|^n = 3^n, \\[4pt]
B_2 &= A_2^n = \left\{ (a_1, a_2, \ldots, a_n) \mid a_i \in A_2, \ 1 \leqslant i \leqslant n \right\}, \\
|B_2| &= |A_2|^n = 3^n.
\end{aligned} \tag{5}
$$

The intersection of sets B_1 and B_2 is just the "all-trans" element:

$$B_1 \cap B_2 = (t, t, \ldots, t). \tag{6}$$

Let B_0 denote the set that consists of the "all-trans" element only. The sets B_0, B_1 and B_2 are proper subsets of B. Let us make the following sets from them:

$$B_1' = B_1 - B_0, \qquad B_2' = B_2 - B_0, \qquad B_3 = B - \left(B_1' \cup B_2' \right), \qquad B_3' = B_3 - B_0. \tag{7}$$

It can be seen that the sets B_0, B_1', B_2' and B_3' are disjoint in pairs and their union is the set B. With this partition of B, the conformational problem can be subdivided into three subcases:

(i) According to rules (1)–(4), the whole set B_2' can be discarded because its elements do not represent existing conformers.

(ii) The set B_1' can be examined separately from the set B_3'. To determine the elements representing conformers in B_1', only rule (1a) should be considered.

(iii) The sequences corresponding to the elements of set B_3' contain at least one "x" and one "g" element. In this case, all the rules should be applied to determine the elements representing conformers.

Let us define two unary operations ($\phi_i, \phi_m : B \to B$) on the set B:

$$\varphi_i\big[(a_1, a_2, \ldots, a_n)\big] = (a_n, a_{n-1}, \ldots, a_1),$$

$$\varphi_m\big[(a_1, a_2, \ldots, a_n)\big] = (-a_1, -a_2, \ldots, -a_n). \tag{8}$$

To assure the closure of B under ϕ_m, the following relations are to be required:

$$-t = t, \qquad -g^+ = g^-, \qquad -g^- = g^+, \qquad -x^+ = x^-, \qquad -x^- = x^+. \tag{9}$$

From the definition of the operations ϕ_i and ϕ_m, it follows that

$$\begin{aligned}
\varphi_i\big[\varphi_i(\alpha)\big] &= \alpha, \\
\varphi_m\big[\varphi_m(\alpha)\big] &= \alpha, \\
\varphi_i\big[\varphi_m(\alpha)\big] &= \varphi_m\big[\varphi_i(\alpha)\big],
\end{aligned} \tag{10}$$

where $\alpha \in B$. We can readily generate a matrix or a permutation representation for this algebra; however, there is no need to use it in practice.

Let α and β be two elements of B and generate the following elements from them:

$$\begin{aligned}
\alpha_i &= \varphi_i(\alpha), & \beta_i &= \varphi_i(\beta), \\
\alpha_m &= \varphi_m(\alpha), & \beta_m &= \varphi_m(\beta), \\
\alpha_k &= \varphi_m\big[\varphi_i(\alpha)\big], & \beta_k &= \varphi_m\big[\varphi_i(\beta)\big].
\end{aligned} \tag{11}$$

Let ρ_α and ρ_β denote the following sets:

$$\rho_\alpha = \{\alpha, \alpha_i, \alpha_m, \alpha_k\}, \qquad \rho_\beta = \{\beta, \beta_i, \beta_m, \beta_k\}. \tag{12}$$

It can easily be shown that the sets ρ_α and ρ_β are the same or disjoint:

$$\rho_\alpha = \rho_\beta \quad \text{or} \quad \rho_\alpha \cap \rho_\beta = \emptyset. \tag{13}$$

Generating all the distinct ρs, a partition of B can be obtained. The sets ρ are therefore equivalence classes of B. In fact, with the help of the two unary operations under consideration an equivalence relation is defined on set B. In a similar manner, the sets B_1', B_2' and B_3' can be partitioned. The set B_0' is an equivalence class of B itself:

$$\rho_{(t,t,\ldots,t)} = \big\{(t, t, \ldots, t)\big\} = B_0. \tag{14}$$

The equivalence classes ρ contain one, two or four elements depending on the symmetry properties of the sequences:

(i) $|\rho_\alpha| = 1$ if $\alpha = \alpha_i = \alpha_m = \alpha_k$. In this case, the sequence has C_{2v} symmetry if n is even or C_{2h} if n is odd;

(ii) $|\rho_\alpha| = 2$ if $\alpha = \alpha_i \neq \alpha_m$ or $\alpha \neq \alpha_i = \alpha_m$. In the first case, the sequences have C_2 symmetry, while in the second one, C_s if n is even or C_i if n is odd;

(iii) $|\rho_\alpha| = 4$ if α, α_i, α_m and α_k are all distinct. In this case, the sequences have no symmetry (C_1).

The elements belonging to the same equivalence class represent structurally identical sequences. They are said to be isomorphic with each other. The number of structurally unique sequences is therefore equal to the number of equivalence classes. According to rules (1)–(4), certain equivalence classes of B do not represent conformers. Eliminating all of them, one can obtain the structurally unique (non-isomorphic) conformers of the molecule under consideration. In general, the non-isomorphic conformers are named spectral isomers in the literature because the isomorphic conformers cannot be distinguished by spectroscopic methods. The equivalence classes with C_1 and C_2 symmetry consist of chiral sequences (conformers).

4. Conformers in the set B_1

Let us determine first the equivalence classes of B_1. For the number of equivalence classes k in B_1, the following decomposition is valid:

$$k = k_1 + k_2 + k_4, \tag{15}$$

where k_1, k_2 and k_4 are the numbers of the equivalence classes with one, two and four elements, respectively. According to the symmetry properties of the sequences, the number k_2 is a sum of two members:

$$k_2 = \begin{cases} k_2(C_2) + k_2(C_s) & (n \text{ is even}), \\ k_2(C_2) + k_2(C_i) & (n \text{ is odd}), \end{cases} \tag{16}$$

where, for instance, $k_2(C_2)$ is the number of equivalence classes with C_2 symmetry. With elementary combinatorical skills, the following equations can be derived for the numbers of equivalence classes:

$$n \text{ is even}: \quad k_2(C_2) = \left(3^{n/2} - 1\right)/2,$$
$$k_2(C_s) = k_2(C_2),$$
$$k_2 = 3^{n/2} - 1, \tag{17}$$
$$k_4 = \frac{1}{4}\left(1 + 3^n - 2 \times 3^{n/2}\right),$$
$$k = \frac{1}{4}\left(1 + 3^n + 2 \times 3^{n/2}\right);$$

$$n \text{ is odd}: \quad k_2(\mathrm{C}_2) = \left(3^{(n+1)/2} - 1\right)/2,$$

$$k_2(\mathrm{C_i}) = \left(3^{(n-1)/2} - 1\right)/2,$$

$$k_2 = 2 \times 3^{(n-1)/2} - 1, \tag{18}$$

$$k_4 = \frac{1}{4}\left(1 + 3^n - 4 \times 3^{(n-1)/2}\right),$$

$$k = \frac{1}{4}\left(1 + 3^n + 4 \times 3^{(n-1)/2}\right).$$

Table 1 presents data up to $n = 20$. The numbers in table 1 are well-known from other combinatorical and graph theoretical studies [1,3–6,9,21].

Let us now consider rule (1a). To count the elements (sequences), which are forbidden according to rule (1a), a special procedure is needed. Figure 1 presents our counting procedure for the octane molecule ($n = 5$). It is to be seen that the number of forbidden elements in set B_1 is 144, therefore, the number of conformers is 99. For butane, pentane, hexane and heptane, the numbers of conformers in set B_1 are 3, 7, 17 and 41, respectively. It is worthwhile to note that these numbers were also obtained by Bartell and Kohl [2]. By application of mathematical induction, an equation can be deduced for the numbers of forbidden elements, N_f, in

Table 1
Equivalence classes of set B_1.

n	3^n	k_1	k_2	k_4	k
1	3	1	1	0	2
2	9	1	2	1	4
3	27	1	5	4	10
4	81	1	8	16	25
5	243	1	17	52	70
6	729	1	26	169	196
7	2187	1	53	520	574
8	6561	1	80	1600	1681
9	19683	1	161	4840	5002
10	59049	1	242	14641	14884
11	177147	1	485	44044	44530
12	531441	1	728	132496	133225
13	1594323	1	1457	397852	399310
14	4782969	1	2186	1194649	1196836
15	14348907	1	4373	3585040	3589414
16	43046721	1	6560	10758400	10764961
17	129140163	1	13121	32278480	32291602
18	387420489	1	19682	96845281	96864964
19	1162261467	1	39365	290545684	290585050
20	3486784401	1	59048	871666576	871725625

| g+ | g- | | | | | $3^3 = 27$ |

| g- | g+ | | | | | $3^3 = 27$ |

| | g+ | g- | | | | $3^3 - 3^2 = 18$ |

no g-

| | g- | g+ | | | | $3^3 - 3^2 = 18$ |

no g+

| | | g+ | g- | | | $3^3 - 3^2 - 3^1 = 15$ |

no g-

no g+g-
no g-g+

| | | g- | g+ | | | $3^3 - 3^2 - 3^1 = 15$ |

no g+

no g+g-
no g-g+

| | | | g+ | g- | | $3^3 - 3^2 - 3\,3^1 + 3\,3^0 = 12$ |

no g-

no g+g-
no g-g+

no g+g-
no g-g+

| | | | g- | g+ | | $3^3 - 3^2 - 3\,3^1 + 3\,3^0 = 12$ |

no g+

no g+g-
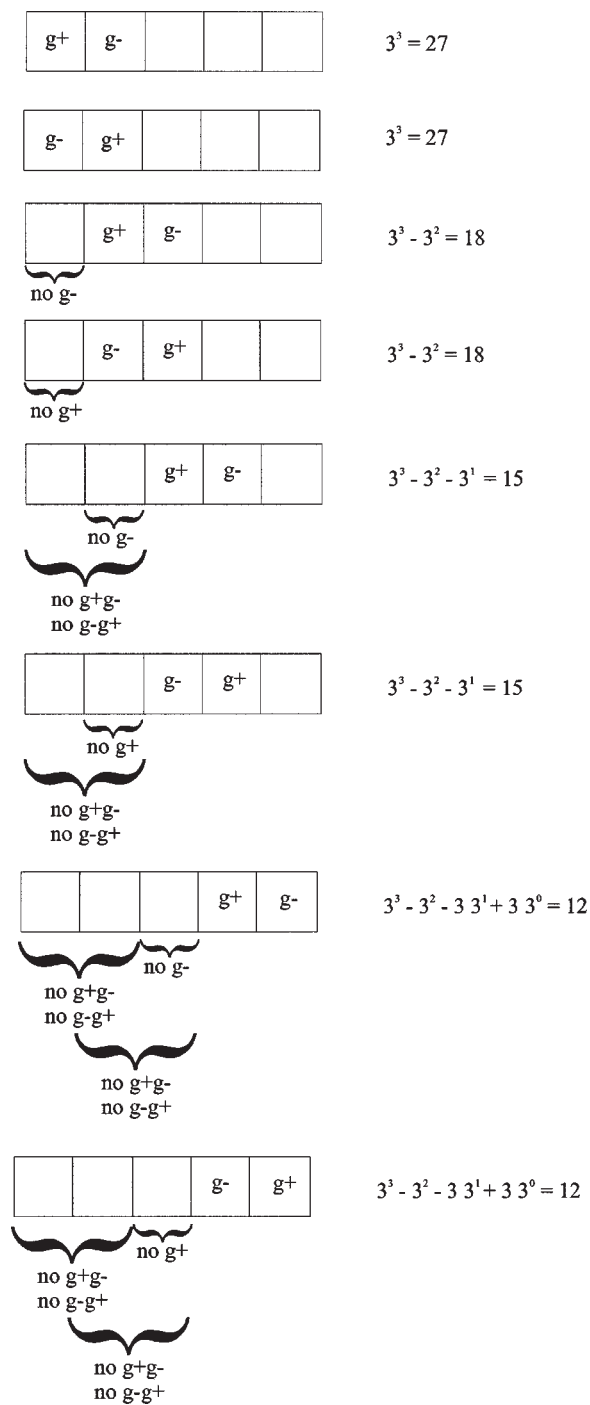no g-g+

no g+g-
no g-g+

Figure 1. Counting procedure to determine the forbidden sequences for octane ($n = 5$) in set B_1 according to rule (1a).

set B_1:

$$N_f = \frac{2}{3} \sum_{i=1}^{n-1} \omega_i 3^{n-i}, \tag{19}$$

where ωs are integers and the following recurrence relation is valid for them:

$$\omega_{i+2} = 2\omega_{i+1} + \omega_i \quad \text{with } \omega_1 = 1 \text{ and } \omega_2 = 2. \tag{20}$$

According to equation (19), the number of elements representing conformers in B_1 is

$$N_c = 3^n - \frac{2}{3} \sum_{i=1}^{n-1} \omega_i 3^{n-i} \tag{21}$$

or, transforming the recurrence relation (20) of ωs into explicit formula,

$$N_c = 3^n - \frac{1}{3} \sum_{i=1}^{n-1} \left[\frac{(1+\sqrt{2})^i}{\sqrt{2}} + \frac{(1-\sqrt{2})^{i+1}}{2-\sqrt{2}} \right] 3^{n-i}. \tag{22}$$

To determine the number of equivalence classes, which consist of elements representing conformers in B_1, the following expressions can be used:

$$n \text{ is even}: \quad k_2(C_2) = \frac{1}{2}\left(3^{n/2} - \frac{2}{3} \sum_{i=1}^{(n/2)-1} \omega_i 3^{(n/2)-i} - 1 \right),$$

$$k_2(C_s) = \frac{1}{2}\left(3^{(n/2)-1} - \frac{2}{3^2} \sum_{i=1}^{(n/2)-2} \omega_i 3^{(n/2)-i} - 1 \right); \tag{23}$$

$$n \text{ is odd}: \quad k_2(C_2) = \frac{1}{2}\left(3^{(n+1)/2} - \frac{2}{3} \sum_{i=1}^{(n-1)/2} \omega_i 3^{(n+1)/2-i} - 1 \right),$$

$$k_2(C_i) = \frac{1}{2}\left(3^{(n-1)/2} - \frac{2}{3} \sum_{i=1}^{(n-3)/2} \omega_i 3^{(n-1)/2-i} - 1 \right). \tag{24}$$

The number of equivalence classes with C_1 symmetry is

$$k_4 = \frac{1}{4}(N_c - 2k_2 - 1). \tag{25}$$

Table 2 presents data for the conformers in B_1 up to $n = 20$. The last column of table 2 gives the fractions r of the conformers in B_1.

Bartell and Fitzwater studied the conformational behaviour of hexadecane ($n = 13$) by electron diffraction method [7]. They wrote in their paper: "... A molecule of n-hexadecane has 1225 internuclear distances in a given conformation and can exist in well over a million conformations ...". In fact, it can be seen from table 2 that the

Table 2
Equivalence classes of the conformers in set B_1.

n	N_c	k_1	k_2	k_4	k	r
1	3	1	1	0	2	1.0000
2	7	1	1	1	3	0.7778
3	17	1	4	2	7	0.6296
4	41	1	4	8	13	0.5062
5	99	1	11	19	31	0.4074
6	239	1	11	54	66	0.3278
7	577	1	28	130	159	0.2638
8	1393	1	28	334	363	0.2123
9	3363	1	69	806	876	0.1709
10	8119	1	69	1995	2065	0.1375
11	19601	1	168	4816	4985	0.1106
12	47321	1	168	11746	11915	0.0890
13	114243	1	407	28357	28765	0.0717
14	275807	1	407	68748	69156	0.0577
15	665857	1	984	165972	166957	0.0464
16	1607521	1	984	401388	402373	0.0373
17	3880899	1	2377	969036	971414	0.0301
18	9369319	1	2377	2341141	2343519	0.0242
19	22619537	1	5740	5652014	5657755	0.0195
20	54608393	1	5740	13649228	13654969	0.0157

number of conformers according to their model is much less than one million: "only" 114,243.

Acknowledgements

References

[1] A.T. Balaban, Rev. Roum. Chim. 21 (1976) 1049.
[2] L.S. Bartell and D.A. Kohl, J. Chem. Phys. 39 (1963) 3097.
[3] J. Brunvoll, B.N. Cyvin, E. Brendsdal and S.J. Cyvin, Comput. Chem. 19 (1995) 379.
[4] S.J. Cyvin, J. Math. Chem. 17 (1995) 291.
[5] S.J. Cyvin, J. Brunvoll, B.N. Cyvin and W. Lüttke, Z. Naturforsch. 50A (1995) 857.
[6] S.J. Cyvin, J. Wang, J. Brunvoll, S. Cao, Y. Li, B.N. Cyvin and Y. Wang, J. Mol. Struct. 413–414 (1997) 227.
[7] S. Fitzwater and L.S. Bartell, J. Am. Chem. Soc. 98 (1976) 8338.
[8] R.F. Frey, M. Cao, S.Q. Newton and L. Schäfer, J. Mol. Struct. (Theochem) 285 (1993) 99.
[9] E. Funck, Z. Elektrochem. 62 (1958) 901.
[10] J.M. Goodman, J. Chem. Inf. Comput. Sci. 37 (1997) 876.

[11] H. Goto, E. Osawa and M. Yamato, Tetrahedron 49 (1993) 387.

[12] W.A. Herrebout, B.J. van der Veken, A. Wang and J.R. Durig, J. Phys. Chem. 99 (1995) 578.

[13] W. Hütner, W. Majer and H. Kästle, Mol. Phys. 67 (1989) 131.

[14] G.D. Smith and R.L. Jaffe, J. Phys. Chem. 100 (1996) 18718.

[15] G. Tasi and F. Mizukami, J. Chem. Inf. Comput. Sci. 38 (1998) 313.

[16] G. Tasi and F. Mizukami, J. Chem. Inf. Comput. Sci. 38 (1998) 632.

[17] G. Tasi and F. Mizukami, under preparation.

[18] G. Tasi, F. Mizukami and I. Pálinkó, J. Mol. Struct. (Theochem) 401 (1997) 21.

[19] G. Tasi, F. Mizukami, I. Pálinkó, J. Csontos, W. Győrffy, P. Nair, K. Maeda, M. Toba, S. Niwa, Y. Kiyozumi and I. Kiricsi, J. Phys. Chem. A 102 (1998) 7698.

[20] S. Tsuzuki, L. Schäfer, H. Goto, E.D. Jemmis, H. Hosoya, K. Siam, K. Tanabe and E. Osawa, J. Am. Chem. Soc. 113 (1991) 4665.

[21] J. Wang, S. Cao and Y. Li, J. Math. Chem. 20 (1996) 211.

[22] K.B. Wiberg and M.A. Murco, J. Am. Chem. Soc. 110 (1988) 8029.