

## Quantum algebraic–combinatoric study of the conformational properties of $n$ -alkanes. II

Gyula Tasi<sup>a,b</sup>, Fujio Mizukami<sup>a</sup>, József Csontos<sup>c</sup>,  
Werner Győrffy<sup>d</sup> and István Pálinkó<sup>e</sup>

<sup>a</sup> Department of Surface Chemistry, National Institute of Materials and Chemical Research, 1-1 Higashi, Tsukuba, Ibaraki 305-8565, Japan

<sup>b</sup> Department of Applied and Environmental Chemistry, József Attila University, Rerrich B. tér 1, H-6720 Szeged, Hungary

<sup>c</sup> Plant Biology Institute, Biological Research Center, Temesvári krt 62, H-6701 Szeged, Hungary

<sup>d</sup> Section of Risk Assessment, National Institute of Chemical Safety, József Fodor National Center of Public Health, POB 64, H-1966 Budapest, Hungary

<sup>e</sup> Department of Organic Chemistry, József Attila University, Dóm tér 8, H-6720 Szeged, Hungary

Received 1 March 2000; revised 12 June 2000

In a previous paper we solved a well-defined subcase of the whole conformational enumeration problem of  $n$ -alkanes. This paper offers a general solution. Most probably the rules determining the sequences of the conformers slightly depend on the force field applied. Nevertheless, the methods presented here can be applied in every case.

**KEY WORDS:**  $n$ -alkanes, conformers, theoretical chemical rules, algebraic and combinatoric results

### 1. Introduction

One hundred and twenty five years ago the concept of the tetrahedral carbon atom [1] of van't Hoff and Le Bell initiated the development of a new branch of chemistry, stereochemistry, with inherent three-dimensional viewpoint of molecules and their transformations [2].

One of the basic tasks of stereochemistry is to enumerate the structural, configurational and conformational isomers of molecules. Since its birth the set of aliphatic alkanes has been the most popular target for structural isomer enumeration studies [3]. In the 1930s Pólya established a general combinatorial theory [4], which was gradually modified [5] and extended [6]. Details concerning the current status of the enumeration of the structural isomers of aliphatic alkanes are to be found in recent papers [7].

Even for an  $n$ -alkane molecule due to its flexible carbon backbone the question of the number of possible conformers arises. Further important questions are: (i) what are the relative stabilities of the conformers and (ii) how can they transform to each other?

After all, the macroscopic physical and chemical properties of *n*-alkanes are statistical averages over the ensembles of the conformers.

The conformational enumeration problem of *n*-alkanes cannot be solved by means of abstract mathematics [8] only and the current advanced experimental techniques cannot give a decisive argument in this case [9] either. At present, the best solution seems to be the combined application of theoretical chemical (physical) and abstract mathematical tools.

By conventional quantum chemical treatment, the various conformers of a given molecule are related to the local minima of its conformational potential energy surface (PES) deduced from the approximate solution of the non-relativistic, time-independent Schrödinger equation within the Born–Oppenheimer approximation [10]. As far as the saddle points of the conformational PES are concerned, the most important first-order ones (transition states) connect the conformers to each other. Accordingly, the topology of the conformational PES [11] determines the conformational properties of a molecule.

In a previous paper [12], a well-defined subcase of the conformational enumeration problem of *n*-alkanes was solved. By chance the conformers with *t*,  $g^+$  and  $g^-$  torsional angles are the most important ones for *n*-alkanes from thermodynamic point of view. Our recent studies, however, have revealed that this does not hold true for *n*-silanes [13]. Therefore, the solution of the whole enumeration problem has some general relevance. We have derived several independent solutions: this paper presents the simplest ones.

## 2. Rules determining the sequences of the conformers of free *n*-alkanes

In several papers, the conformational properties of *n*-alkanes were studied by theoretical chemical methods at various levels of theory [14–16]. For the butane, pentane, hexane, and heptane molecules, full ab initio conformational analyses were performed at Hartree–Fock (HF) and second-order Moeller–Plesset (MP2) levels [15]. For the octane molecule, the HF level was the highest considered [15]. With the help of an effective one-electron method (SEOEM) parametrized on high-quality G2 data [17], all the conformers in the conformational space were determined up to undecane [16].

The following notation is used to designate the C–C–C torsional angles in the isolated *n*-alkane molecules: *t*:  $\sim 180^\circ$ ,  $g^+$ :  $\sim +60^\circ$ ,  $g^-$ :  $\sim -60^\circ$ ,  $x^+$ :  $\sim +95^\circ$  and  $x^-$ :  $\sim -95^\circ$ . From the quantum chemical calculation results, four rules were deduced for the sequences of the conformers [15]:

- (1) double sequences: (a)  $g^+g^-$  and  $g^-g^+$  as well as (b)  $x^+x^-$  and  $x^-x^+$  cannot occur;
- (2) double sequences:  $g^-$  should exist adjacent to  $x^+$  and  $g^+$  adjacent to  $x^-$ :  $\dots x^+g^- \dots, \dots g^-x^+ \dots, \dots x^-g^+ \dots$ , and  $\dots g^+x^- \dots$ ;
- (3) triple sequences: (a)  $g^+x^-g^+$  and  $g^-x^+g^-$  as well as (b)  $x^+g^+x^-$ ,  $x^-g^+x^+$ ,  $x^+g^-x^-$ , and  $x^-g^-x^+$  are not allowed;
- (4) quadruple sequences:  $x^+g^-g^-x^+$  and  $x^-g^+g^+x^-$  are forbidden.

As far as rule (4) is concerned, our previous calculations [15] revealed that the sequences of the heptane and octane molecules containing the quadruple sequences under consideration are not local minima on the conformational PES at the HF level regardless of the size of the basis set applied. By considering electron correlation effects at the MP2 level, the small basis set calculations afforded similar results [15]. However, it is noteworthy that the MM2 force field [18] seems to support only the first three rules [14,15]. Most likely, for other force fields the rules under consideration are a little bit different, but the technique presented here may be applied in every case.

At present, two models are worth examining: model 1, when all four rules apply and model 2 when only the first three rules are valid.

### 3. Conformers with $t$ , $g^+$ and $g^-$ torsional angles only

In our previous paper [12], for the number of the conformers with  $t$ ,  $g^+$  and  $g^-$  torsional angles, i.e. the conformers in the set  $B_1$ , the following equation was obtained:

$$C_n(B_1) = 3^n - \frac{2}{3} \sum_{i=1}^{n-1} \omega_i 3^{n-i}, \quad (1)$$

where  $n$  is the number of the free C–C–C torsional angles in the molecule;  $\omega_i$  are integers and a recurrence relation is valid for them:

$$\omega_{i+2} = 2\omega_{i+1} + \omega_i \quad \text{with } \omega_0 = 0 \text{ and } \omega_1 = 1. \quad (2)$$

By application of mathematical induction, however, equation (3) can be easily obtained from equations (1) and (2):

$$C_n(B_1) = \omega_{n+2} - \omega_{n+1} = \omega_{n+1} + \omega_n, \quad (3)$$

which leads to the following simple recurrence relation for the number of the conformers in  $B_1$ :

$$C_{n+2}(B_1) = 2C_{n+1}(B_1) + C_n(B_1) \quad \text{with } C_0(B_1) = 1 \text{ and } C_1(B_1) = 3. \quad (4)$$

This relation can be transformed into an explicit formula:

$$C_n(B_1) = \left(\frac{1}{2} + \frac{1}{\sqrt{2}}\right)(1 + \sqrt{2})^n + \left(\frac{1}{2} - \frac{1}{\sqrt{2}}\right)(1 - \sqrt{2})^n. \quad (5)$$

Similarly, recurrence relations can be deduced for the equivalence classes of the conformers in the set  $B_1$ :

$$\begin{aligned} k_n^\alpha &= 1, \\ k_{n+4}^\beta &= 2k_{n+2}^\beta + k_n^\beta + 2 \quad \text{with } k_0^\beta = 0, k_1^\beta = 1, k_2^\beta = 1, \text{ and } k_3^\beta = 4, \\ k_n^\delta &= \frac{1}{4}(C_n(B_1) + 2k_n^\beta - 1), \end{aligned} \quad (6)$$

where  $k_n^\alpha$ ,  $k_n^\beta$ , and  $k_n^\delta$  are the numbers of the equivalence classes with one, two, and four elements, respectively.

Relation (4) can also be derived with the help of graph theory. Some part of the mixed binary and ternary rooted family tree of the *n*-alkane conformers belonging to the set  $B_1$  is shown in figure 1. The vertices  $t$  and  $g$  have three and two children, respectively. It is clearly seen that the following recurrence relations hold for them:

$$\begin{aligned} t_n &= t_{n-1} + g_{n-1}, \\ g_n &= 2t_{n-1} + g_{n-1}, \end{aligned} \quad (7)$$

where  $t_n$  and  $g_n$  are the numbers of the vertices at level  $n$  with three and two children, respectively. The level number corresponds to the number of the free torsional angles. The number of the vertices at level  $n$  affords the number of the conformers of the *n*-alkane molecule with  $n$  free torsional angles in the set  $B_1$ :

$$\begin{aligned} C_n(B_1) &= t_n + g_n = t_{n-1} + g_{n-1} + 2t_{n-1} + g_{n-1} = 2(t_{n-1} + g_{n-1}) + t_{n-1} \\ &= 2C_{n-1}(B_1) + C_{n-2}(B_1). \end{aligned} \quad (8)$$

It is to be seen that the same result was obtained as above.

#### 4. Conformers in the set $B$ : solution of the whole conformational enumeration problem

The previously applied graph theoretical method can be used to enumerate all the conformers in the conformational space, i.e. in set  $B$  [12]. Figure 2 shows the mixed unary, ternary, quaternary and pentanary rooted family tree of the *n*-alkane conformers in the set  $B$  up to  $n = 3$ . An exhaustive detailed analysis of the whole family tree results in the following set of simultaneous recurrence relations for model 1, i.e. when all the four rules apply:

$$\begin{aligned} a_n &= a_{n-1} + b_{n-1} + d_{n-1}, \\ b_n &= 2a_{n-1} + b_{n-1} + b_{n-3} + c_{n-3} + c_{n-4}, \\ c_n &= 2a_{n-1} + b_{n-1} + b_{n-2} + b_{n-3} + 2c_{n-3} + c_{n-4}, \\ d_n &= b_{n-1} + b_{n-2} + c_{n-1} + 2c_{n-2} + c_{n-3}, \end{aligned} \quad (9)$$

where  $a_n$ ,  $b_n$ ,  $c_n$  and  $d_n$  are the numbers of the vertices at level  $n$  with five, four, one and three children, respectively. For  $n = 0$ ,  $a_0 = 1$ ,  $b_0 = c_0 = d_0 = 0$  and for  $n < 0$ ,  $a_n = b_n = c_n = d_n = 0$ . And finally, the number of the conformers of the *n*-alkane molecule with  $n$  free torsional angles can be obtained via equation (10):

$$C_n(B) = a_n + b_n + d_n. \quad (10)$$

It is worthwhile to note that the vertices with one child should be neglected in equation (10) because of rule (2). Scheme 1 presents a short program written in C to enumerate the conformers based on equations (9) and (10).

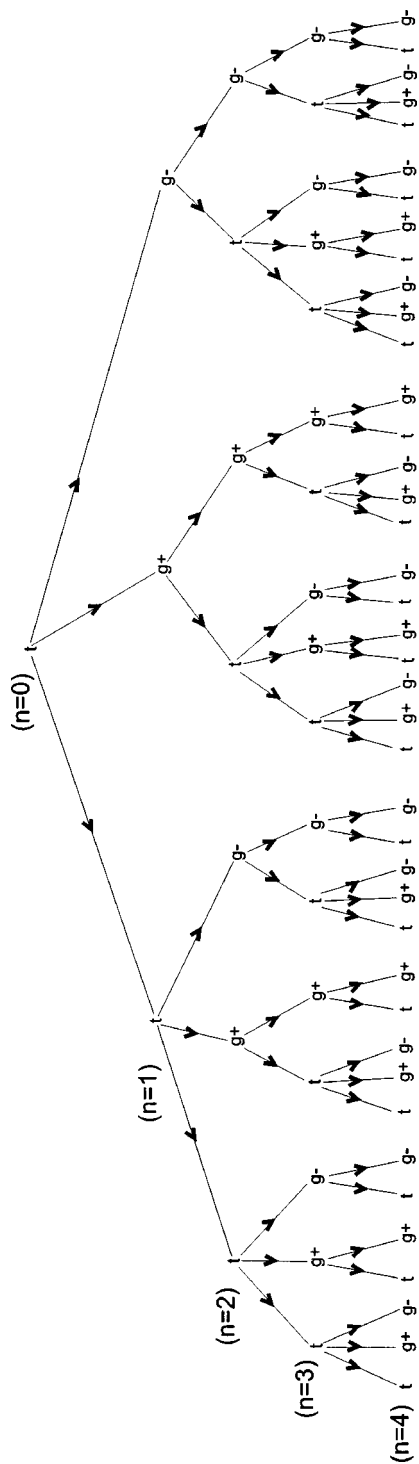


Figure 1.

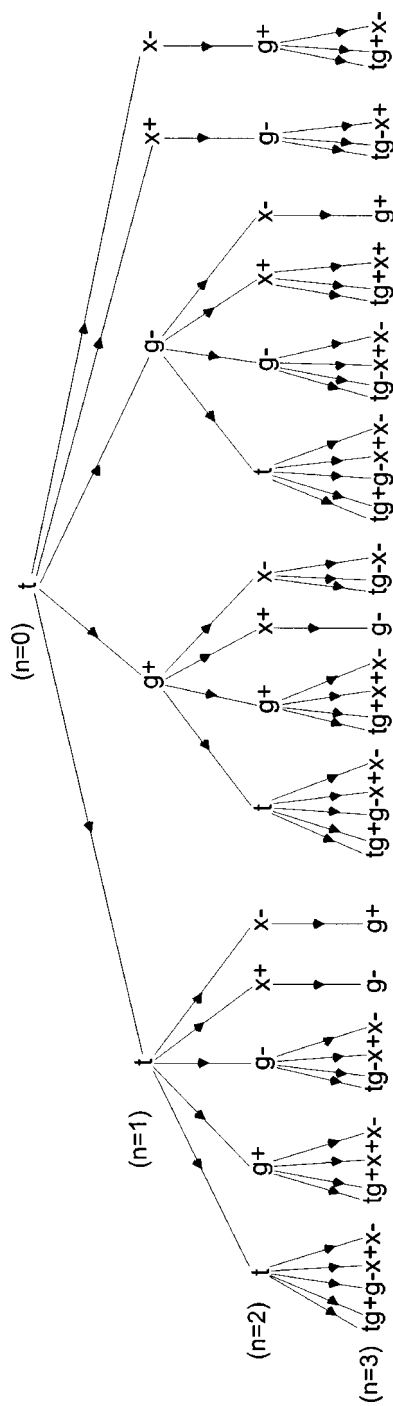


Figure 2.

```

# include <stdio.h>

# define MAXN 60 /* maximum dimension */
# define N 40 /* number of the C-C-C-C torsional angles considered */

main() {

    double a[MAXN], b[MAXN], c[MAXN], d[MAXN];
    int i, j;

    for (j = 1; j <= N; ++j) {
        a[0] = 1.0;
        b[0] = c[0] = d[0] = 0.0;
        for (i = 1; i <= j; ++i) {
            a[i] = a[i-1] + b[i-1] + d[i-1];
            b[i] = 2 * a[i-1] + b[i-1];
            c[i] = b[i];
            d[i] = b[i-1] + c[i-1];
            if (i > 2) {
                c[i] += b[i-2];
                d[i] += b[i-2] + 2 * c[i-2];
            }
            if (i > 3) {
                b[i] += c[i-3] + b[i-3];
                c[i] += 2 * c[i-3] + b[i-3];
                d[i] += c[i-3];
            }
            if (i > 4) {
                b[i] += c[i-4];
                c[i] += c[i-4];
            }
        }
        printf("%3d\t%40.0f\n", j, a[j] + b[j] + d[j]);
    }
}

```

Scheme 1. A C program to enumerate the conformers of *n*-alkanes according to model 1.

In the case of model 2, i.e. neglecting rule (4), the above set of simultaneous recurrence relations reads as follows:

$$\begin{aligned}
 a_n &= a_{n-1} + b_{n-1} + d_{n-1}, \\
 b_n &= 2a_{n-1} + b_{n-1} + b_{n-3} + c_{n-2} + c_{n-4}, \\
 c_n &= 2a_{n-1} + b_{n-1} + b_{n-2} + b_{n-3} + c_{n-3} + c_{n-4}, \\
 d_n &= b_{n-1} + b_{n-2} + c_{n-1} + c_{n-2} + c_{n-3}.
 \end{aligned}
 \tag{11}$$

Using equation (11), the number of the conformers can be obtained with the help of equation (10) in this case too. Table 1 lists the calculated values for the various models

Table 1  
 Number of the conformers of free *n*-alkanes by the various models. Model 0: rule (1a),  
 model 1: rules (1)–(4) and model 2: rules (1)–(3).

<i>n</i>	Model 0	Model 1	Model 2
1	3	3	3
2	7	11	11
3	17	35	35
4	41	107	109
5	99	339	347
6	239	1073	1105
7	577	3375	3507
8	1393	10633	11135
9	3363	33525	35371
10	8119	105651	112343
11	19601	332941	356797
12	47321	1049305	1133207
13	114243	3306957	3599143
14	275807	10421967	11431063
15	665857	32845327	36305649
16	1607521	103513709	115308721
17	3880899	326228241	366226729
18	9369319	1028123557	1163155753
19	22619537	3240180157	3694245273
20	54608393	10211580633	11733121945
21	131836323	32182277499	37265026883
22	318281039	101423965833	118355731271
23	768398401	319642412979	375904173873
24	1855077841	1007368140211	1193891891823
25	4478554083	3174768208785	3791864917001
26	10812186007	10005431759263	12043167097497
27	26102926097	31532590122723	38249747002215
28	63018038201	99376445091607	121483255514989
29	152139002499	313189554080745	385837359123841
30	367296043199	987031652156513	1225440222727991
31	886731088897	3110676807889763	3892064114500595
32	2140758220993	9803444683903509	12361405142762380
33	5168247530883	30896018328432150	39260488164678835
34	12477253282759	97370259059799520	124693423856546550
35	30122754096401	306866964169575350	396033026585146800
36	72722761475561	967105711825511600	1257822211431372000
37	175568277047523	3047879267084092000	3994911053787981000
38	423859315570607	9605535272029320000	12688052558331027000
39	1023286908188737	30272297481937875000	40297937940401080000
40	2470433131948081	95404573392538760000	127988420191571700000

up to  $n = 40$ . In the table, model 0 designates the results obtained via equation (4). All the calculation results were verified against those afforded by the ENUMERAT program [15]. They are precisely the same. It is noteworthy that model 1 and model 2





- [8] S.J. Cyvin, *J. Math. Chem.* 17 (1995) 291;  
J. Wang, S. Cao and Y. Li, *J. Math. Chem.* 20 (1996) 211.
- [9] G.T. Fraser, R.D. Suenram and C.L. Lugez, *J. Phys. Chem. A* 104 (2000) 1141;  
G.T. Fraser, L.-H. Xu, R.D. Suenram and C.L. Lugez, *J. Chem. Phys.* 112 (2000) 6209.
- [10] W.J. Hehre, L. Radom, P.v.R. Schleyer and J.A. Pople, *Ab Initio Molecular Orbital Theory* (Wiley, New York, 1986);  
A. Szabo and N.S. Ostlund, *Modern Quantum Chemistry* (Dover, New York, 1996);  
D.B. Cook, *Handbook of Computational Quantum Chemistry* (Oxford University Press, Oxford, 1998).
- [11] P.G. Mezey, *Theor. Chem. Acc.* 102 (1999) 279; *J. Math. Chem.* 14 (1993) 79.
- [12] G. Tasi and F. Mizukami, *J. Math. Chem.* 25 (1999) 55.
- [13] G. Tasi et al., under preparation.
- [14] H. Gotō, E. Ōsawa, M. Yamato, *Tetrahedron* 49 (1993) 387;  
S. Tsuzuki, L. Schäfer, H. Gotō, E.D. Jemmis, H. Hosoya, K. Siam, K. Tanabe and E. Ōsawa, *J. Am. Chem. Soc.* 113 (1991) 4665.
- [15] 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.
- [16] G. Tasi and F. Mizukami, *J. Chem. Inf. Comput. Sci.* 38 (1998) 632.
- [17] L.A. Curtiss, K. Raghavachari, G.W. Trucks and J. Pople, *J. Chem. Phys.* 94 (1991) 7221.
- [18] U. Burkert and N.L. Allinger, *Molecular Mechanics* (American Chemical Society, Washington, DC, 1982).
- [19] N.J.A. Sloane and S. Plouffe, *The Encyclopedia of Integer Sequences* (Academic Press, San Diego, CA, 1995); on-line version at URL: <http://www.research.att.com>.
- [20] J.M. Goodman, *J. Chem. Inf. Comput. Sci.* 37 (1997) 876.
- [21] N. Nair and J.M. Goodman, *J. Chem. Inf. Comput. Sci.* 38 (1998) 317.
- [22] I. Kolossváry and W.C. Guida, *J. Comp. Chem.* 20 (1999) 1671.