

# Numbers of Monosubstituted Alkanes as Stereoisomers

Shinsaku FUJITA

Department of Chemistry and Materials Technology, Kyoto Institute of Technology  
Matsugasaki, Sakyo-ku, Kyoto, 606-8585 Japan  
e-mail: fujitas@chem.kit.ac.jp

(Received: September 28, 2006; Accepted for publication: December 5, 2006; Published on Web: February 28, 2007)

Monosubstituted alkanes as stereoisomers, not as constitutional isomers, are regarded as planted three-dimensional (3D) trees, which are defined as a 3D extension of planted trees (graphs). They are thus recognized as 3D-objects (planted promolecules) and enumerated by Fujita's proligand method (Fujita S (2005) Theor. Chem. Acc. 113:73–79, 113:80–86, Fujita S (2006) Theor. Chem. Acc. 115:37–53). By starting from three kinds of sphericity indices, i.e.,  $a_d$  for homospheric cycles,  $c_d$  for enantiospheric cycles, and  $b_d$  for hemispheric cycles, cycle indices with chirality fittingness (CI-CFs) are obtained to enumerate planted 3D-trees or equivalently monosubstituted alkanes as stereoisomers. Functional equations  $a(x)$ ,  $c(x^2)$ , and  $b(x)$  for recursive calculations are derived from the CI-CFs and programmed in three ways by means of the Maple programming language. The three recursive procedures for calculating the numbers of planted 3D-trees are executed to give identical results, which are collected up to 100 carbon content in a tabular form. The results are compared with the enumeration of planted trees (as graphs).

**Keywords:** Monosubstituted alkane, Stereoisomer, Enumeration, Sphericity

## 1 Introduction

Combinatorial enumeration of rooted trees as models of monosubstituted alkanes has been initiated by a mathematician Cayley [1, 2]. Henze and Blair [3] obtained the numbers of aliphatic alcohols of a given carbon content by means of recursive equations. Later, a more systematic method based on Pólya's theorem [4, 5] was applied to the evaluation of the numbers of rooted trees, which was differently formulated as planted trees.

Although Pólya's theorem [4, 5] has been widely applied to chemical combinatorics [6–11], enumerated objects in most cases are “graphs”, not as “three-dimensional objects”. Chemically speaking, they are constitutional (structural) isomers, but not stereoisomers. As pointed out in our recent reports [12–14], Pólya's theorem lacks the concept of sphericities [15, 16] so as to be incapable of solving stereochemical problems. Although Robinson *et al.* [17] have reported the enumeration of monosubstituted alkanes by modifying Pólya's cycle indices (CIs), their treatment also lacks the concept of sphericities.

Without the sphericity concept, in general, any methods are incapable of treating fundamental stereochemical problems such as *meso*-compounds and pseudoasym-

metry. Because the stereochemical problems have been already solved in the last quarter of the 19th century in a descriptive or non-mathematical fashion [18–21], the present status of chemical combinatorics based on Pólya's theorem does not properly reflect the advance of stereochemistry before and during the 20th century.

On the basis of the sphericity concept, we have developed a general method named *the proligand method* for enumerating stereoisomers [12–14]. To show the versatility of the proligand method, we have recently applied it to the enumeration of monosubstituted alkanes (planted 3D-trees) as 3D-objects [22].

The sphericity concept was originally proposed in the form of “the sphericities of orbits”, which provides us with various tools for combinatorics in connection with Fujita's USCI (unit-subduced-cycle-index) approach [16]. Then, it was modified into the concept “the sphericities of orbits based on cyclic groups”, which gives a basis to other types of combinatorial tools [23–25]. Further, it was transformed into the concept “the sphericities of cycles”, which gives the proligand method [12–14].

As found by the short history described in the preceding paragraph, the original formulation [12–14] was too mathematical to assure chemists of the traceability of

the proligand method. However, there is a short cut to become practised in the usage of the proligand method without acquainting such mathematical formulation. The short cut is to study suitable examples which disclose the details of calculation procedures. We can safely say that understanding mathematics is usually different from implementing mathematics as a program for concrete calculation. In other words, “The proof of the pudding is in the eating. You can’t eat mathematics, but you *can* digest it.”, as Pólya phrased it [10](page 68). Note that the word “eat” corresponds to “understand” and the word “digest” corresponds to “implement”.

As clarified in the preceding paragraphs, the aim of the present paper is to report the details of calculation procedures for obtaining the numbers of monosubstituted alkanes as 3D-objects (stereoisomers), but not as graphs, where the procedures are implemented by the Maple programming language [26].

## 2 Enumeration of Monosubstituted Alkanes as Stereoisomers

### 2.1 Planted Promolecules and Their Nesting

A monosubstituted alkane is regarded as a planted 3D-tree of degree 4, where a principal vertex (a central carbon atom) is connected with a root (the mono-substituent) and three substituents (alkyl ligands or hydrogen atoms). Each one of the three substituents is called a *proligand*, which is regarded as being a hypothetical structureless object with chirality/achirality. By considering such proligands, the original monosubstituted alkane (or the planted 3D-tree) is regarded as a *planted promolecule*, which is a 3D-skeleton (a tetrahedral skeleton) having the root (●) and the tree proligands on its vertices.

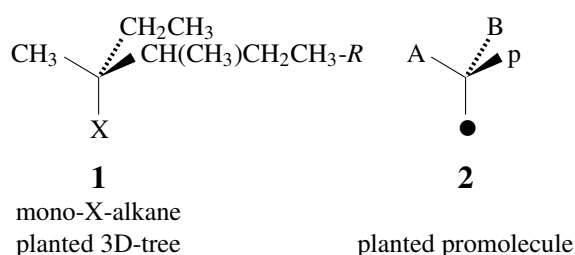


Figure 1. Monosubstituted alkane as a planted 3D-tree and a planted promolecule

For example, let us examine a 3,4-dimethyl-3-X-hexane (**1**), which is selected from a set of four diastereomers. The mono-X-substituted alkane (**1**) is regarded as a planted 3D-tree with a root X. When the three substituents are replaced by proligands, i.e., A = CH<sub>3</sub>, B =

CH<sub>2</sub>CH<sub>3</sub>, and p = R-CH(CH<sub>3</sub>)CH<sub>2</sub>CH<sub>3</sub>, and when the X is replaced by a solid circle, the planted 3D-tree (**1**) is converted into a planted promolecule (**2**) with a root (●). The proligands A and B are achiral, while the proligand p is chiral. Note that the S-configuration is tentatively assigned by considering a vacant bond (incident to the X originally) as the lowest priority (p > B > A > ●).

Each of the proligands can be further regarded as a planted promolecule, if a root is attached to its central vertex (as a new principal vertex). For example, the proligand p (i.e., a 2-butyl ligand) can be regarded as a planted promolecule, if a root is attached to the p through the vacant bond and the three achiral substituents (H, A = CH<sub>3</sub>, and B = CH<sub>2</sub>CH<sub>3</sub>) are replaced by the corresponding achiral proligands. The resulting planted promolecule is represented by ●—CHAB, which has the same skeleton as the starting planted promolecule (**2**). The proligand B (i.e., an ethyl ligand) can be regarded as a planted promolecule, if a root is attached to the B through the vacant bond and the three achiral substituents (two H’s and A = CH<sub>3</sub>) are replaced as above. The resulting planted promolecule is represented by ●—CH<sub>2</sub>A, which has the same skeleton as **2**.

This procedure is repeated to reach terminal vertices of the original monosubstituted alkane (or the planted 3D-tree). Thereby, the original monosubstituted alkane (or the planted 3D-tree) is recognized to have a nested structure of such planted promolecules.

### 2.2 Sphericities of Cycles

It is to be noted that each planted promolecule contained in a nested fashion is a 3D-object (not a graph) which has a common skeleton belonging to C<sub>3v</sub>. Hence, such planted promolecules can be enumerated by means of Fujita’s proligand method [12–14].

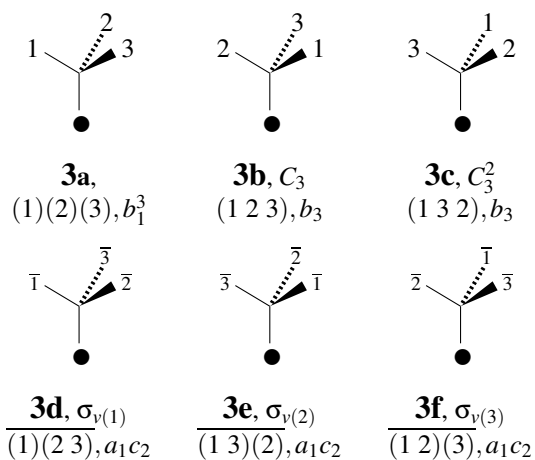


Figure 2. Symmetry operations for a C<sub>3v</sub>-skeleton (**3**). The identity operation (*I*) converts **3** into itself, where the resulting skeleton is denoted as **3a**.

For this purpose, we examine the three positions of the  $C_{3v}$ -skeleton under all operations of  $C_{3v}$ . In terms of the USCI approach [16], the three positions other than the root construct an orbit governed by a right coset representation (RCR) represented as  $(C_s \setminus)C_{3v}$ , whose concrete form is shown as a product of cycles as follows:

$$\begin{aligned} (C_s \setminus)C_{3v} &= \\ &= \frac{\{(1)(2)(3), (1\ 2\ 3), (1\ 3\ 2), \\ &\quad \overline{(1)(2\ 3)}, \overline{(1\ 3)(2)}, \overline{(1\ 2)(3)}\}}{(1)} \end{aligned} \quad (1)$$

The numbering used in the RCR  $(C_s \setminus)C_{3v}$  corresponds to the numbering shown in the skeleton (3), where the product of cycles corresponding to each improper rotation ( $\sigma_{v(1)}$ ,  $\sigma_{v(2)}$ , or  $\sigma_{v(3)}$ ) is attached by an overbar, which shows the inverse of chirality.

The sphericity of each cycle is determined according to its shape, so that a sphericity index (SI) is assigned to the cycle. The chirality fittingness of each orbit is determined by the corresponding SI.

1. A cycle belonging to an improper rotation is determined to be *homospheric*, if the length  $d$  is odd; and a sphericity index  $a_d$  is assigned to it. The chirality fittingness of a homospheric cycle permits the transitivity of achiral (pro)ligands only.
2. If the length  $d$  is even, on the other hand, the cycle in an improper rotation is determined to be *enantiospheric*; and a sphericity index  $c_d$  is assigned to it. The chirality fittingness of an enantiospheric cycle permits the transitivity of achiral (pro)ligands or the transitivity of pairs of enantiomeric ligands (i.e.,  $d/2$  of chiral proligands of the same kind and  $d/2$  of their enantiomeric proligands).
3. Any cycle belonging to a proper rotation is determined to be *hemispheric* and a sphericity index  $b_d$  is assigned to it, where the subscript  $d$  is the length of the cycle. The chirality fittingness of a hemispheric cycle permits the transitivity of achiral and chiral (pro)ligands.

Thereby, each symmetry operation acting on the skeleton is characterized by a product of such sphericity indices, as collected along with the corresponding product of cycles in Figure 2.

To count achiral planted promolecules and enantiomeric pairs of chiral planted promolecules, we apply Theorem 2 of [14] to the enumeration of planted promolecules without any modification. Hence, the cycle index with chirality fittingness (CI-CF) for this case is calculated by using the product of SIs collected in Figure 2:

$$\text{CI-CF}(C_{3v}; a_d, c_d, b_d) = \frac{1}{6}(b_1^3 + 2b_3 + 3a_1c_2). \quad (2)$$

To count achiral and chiral planted promolecules, Theorem 3 of [14] for the enumeration of ligands under the action of the maximum chiral subgroup is applied to this case so as to derive the following CI-CF:

$$\text{CI-CF}(C_3; b_d) = \frac{1}{3}(b_1^3 + 2b_3), \quad (3)$$

To count achiral planted promolecules only, the first proposition of Theorem 4 for the enumeration of achiral ligands [14] is applied to this case so as to derive the following CI-CF<sub>A</sub>:

$$\begin{aligned} \text{CI-CF}_A(C_{3v}; a_d, c_d, b_d) &= \\ &= 2\text{CI-CF}(C_{3v}; a_d, c_d, b_d) - \text{CI-CF}(C_3; b_d) \\ &= a_1c_2, \end{aligned} \quad (4)$$

To count chiral planted promolecules only, the second proposition of Theorem 4 for the enumeration of chiral ligands [14] is applied to obtain the following CI-CF<sub>C</sub>:

$$\begin{aligned} \text{CI-CF}_C(C_{3v}; a_d, c_d, b_d) &= \\ &= \text{CI-CF}(C_3; b_d) - \text{CI-CF}(C_{3v}; a_d, c_d, b_d) \\ &= \frac{1}{6}(b_1^3 + 2b_3 - 3a_1c_2), \end{aligned} \quad (5)$$

## 2.3 Recursive Procedures

By using eq. 3, let us consider the enumeration of achiral and chiral planted promolecules, where two enantiomers of each pair are counted separately. For this purpose, the SI appearing in the right-hand side of eq. 3 is replaced by the following ligand inventory:

$$b(x^d) = b_d = \sum_{k=0}^{\infty} \beta_k x^{dk}, \quad (6)$$

where achiral and chiral proligands for substitution are itemized by their carbon contents  $k$  ( $k = 1$  to  $\infty$ ) and we put  $\beta_0 = 1$ . If the ligand inventory is introduced into eq. 3, the resulting equation is a generating function for counting achiral and chiral planted promolecules. The generating function can be regarded as a new ligand inventory for the hemispheric cycle. This means that this process is recursive, so that we can obtain the following functional equation:

$$b(x) = 1 + \frac{x}{3}(b(x)^3 + 2b(x^3)), \quad (7)$$

where the principal node of the parent planted promolecule is taken into consideration by multiplying by  $x$  and the initial (trivial) planted promolecule is considered by adding 1. This type of equation was first noted by Pólya [4, 5], although the sphericity concept was not taken into consideration.

According to the chirality fittingness of an enantiospheric cycle, diploids (ordered pairs) of achiral

(pro)ligands or enantiomeric ligands satisfy the following ligand inventory:

$$c(x^d) = c_d = \sum_{k=0}^{\infty} \gamma_k x^{dk} \quad (8)$$

where achiral and chiral proligands for substitution are itemized by their carbon contents  $k$  ( $k = 1$  to  $\infty$ ) and we put  $\gamma_0 = 1$ . The concept of *diploid* is here used without definition, which will be reported elsewhere [22]. Because the diploids for this case are characterized by  $C_3$ , we modify eq. 7 into the following functional equation:

$$c(x^2) = 1 + \frac{x^2}{3} (c(x^2)^3 + 2c(x^6)), \quad (9)$$

where the principal nodes of the parent planted promolecules contained in a diploid are characterized by multiplying by  $x^2$  and the initial (trivial) planted promolecule is taken into consideration by adding 1.

By using eq. 4, we are able to enumerate achiral planted promolecules only. Thus, the SI appearing in the right-hand side of eq. 4 is replaced by the following ligand inventory:

$$a(x^d) = a_d = \sum_{k=0}^{\infty} \alpha_k x^{dk} \quad (10)$$

where achiral and chiral proligands for substitution are itemized by their carbon contents  $k$  ( $k = 1$  to  $\infty$ ) and we put  $\alpha_0 = 1$ . If the ligand inventories (eq. 10) along with eq. 6 are introduced into eq. 4, the resulting equation is a generating function for counting achiral planted promolecules only. Because the generating function can be regarded as a new ligand inventory for the homospheric cycle, this process is recognized to be recursive, so that we obtain the following functional equation:

$$a(x) = 1 + xa(x)c(x^2), \quad (11)$$

where the principal node of the parent planted promolecule is taken into consideration by multiplying by  $x$  and the initial (trivial) planted promolecule is considered by adding 1. Although this type of equation was first noted by Robinson *et al.* [17], their treatment did not take the sphericity concept into consideration.

Let  $B(x)$  be a generating function for counting achiral and chiral planted promolecules, where each pair of enantiomers is counted just once. Let  $C(x)$  be a generating function for counting chiral planted promolecules only, where each pair of enantiomers is counted just once. They are represented by the following equations:

$$B(x) = \sum_{k=0}^{\infty} B_k x^k \quad (12)$$

$$C(x) = \sum_{k=0}^{\infty} C_k x^k, \quad (13)$$

where we put  $B_0 = 1$  and  $C_0 = 0$ . Because we can place  $B_k = \alpha_k + C_k$  and  $\beta_k = \alpha_k + 2C_k$ , eq. 7 (for  $b(x)$ ) and eq. 11 (for  $a(x)$ ) give the following equations:

$$B(x) = \frac{1}{2} (a(x) + b(x)) \quad (14)$$

$$= 1 + \frac{x}{6} (b(x)^3 + 2b(x^3) + 3a_1c_2) \quad (15)$$

$$C(x) = \frac{1}{2} (b(x) - a(x)) \quad (16)$$

$$= \frac{x}{6} (b(x)^3 + 2b(x^3) - 3a_1c_2). \quad (17)$$

It should be noted that  $a(x)$  (eq. 10),  $c(x^2)$  (eq. 9), and  $b(x)$  (eq. 7) have recursive nature, while  $B(x)$  (eq. 14 or eq. 15) and  $C(x)$  (eq. 16 or eq. 17) have no recursive nature.

## 2.4 Preliminary Calculations for Implementation

To collect data for implementing the process of recursive calculation, suppose that the coefficients of  $x^k$  ( $k \leq 5$ ) have been already calculated to give intermediate generating functions as follows:

$$a(x) = 1 + x + x^2 + 2x^3 + 3x^4 + 5x^5 + \alpha_6 x^6 + \alpha_7 x^7 + \alpha_8 x^8 \quad (18)$$

$$c(x^2) = 1 + x^2 + x^4 + 2x^6 + 5x^8 + 11x^{10} + \gamma_6 x^{12} + \gamma_7 x^{14} + \gamma_8 x^{16} \quad (19)$$

$$b(x) = 1 + x + x^2 + 2x^3 + 5x^4 + 11x^5 + \beta_6 x^6 + \beta_7 x^7 + \beta_8 x^8 \quad (20)$$

where we add the coefficients of the terms  $x^6$ ,  $x^7$ , and  $x^8$  for calculation and omit the terms for  $k \geq 9$  tentatively. The intermediate generating functions (eqs. 18–20) are introduced into the respective right-hand sides of eqs. 11, 9, and 7. The resulting equations are expanded to give the following equations:

$$a(x) = 1 + x + x^2 + 2x^3 + 3x^4 + 5x^5 + 8x^6 + (6 + \alpha_6)x^7 + (9 + \alpha_7)x^8 + \dots \quad (21)$$

$$c(x^2) = 1 + x^2 + x^4 + 2x^6 + 5x^8 + 11x^{10} + 28x^{12} + (46 + \gamma_6)x^{14} + (69 + 2\gamma_6 + \gamma_7)x^{16} + \dots \quad (22)$$

$$b(x) = 1 + x + x^2 + 2x^3 + 5x^4 + 11x^5 + 28x^6 + (46 + \beta_6)x^7 + (69 + 2\beta_6 + \beta_7)x^8 + \dots, \quad (23)$$

which are compared with eqs. 18–20. Thereby, the coefficients of the terms  $x^6$  (for  $a(x)$  and  $b(x)$ ) or  $x^{12}$  (for  $c(x)$ ) give the following relationships:

$$\alpha_6 = 8 \quad (24)$$

$$\gamma_6 = 28 \quad (25)$$

$$\beta_6 = 28, \quad (26)$$

which show that they are determined as distinct numbers, if the generating functions up to  $x^5$  (for  $a(x)$  and  $b(x)$ ) or to  $x^{10}$  (for  $c(x)$ ) are given.

**Rule 1:** In general, if such generating functions up to  $x^n$  (for  $a(x)$  and  $b(x)$ ) or to  $x^{2n}$  (for  $c(x)$ ) are given, the coefficients ( $\alpha_{n+1}$ ,  $\gamma_{n+1}$ , and  $\beta_{n+1}$ ) are recursively generated as distinct numbers.

Comparison of the coefficients of the terms  $x^7$  (for  $a(x)$  and  $b(x)$ ) or  $x^{14}$  (for  $c(x)$ ) gives the following relationships:

$$\alpha_7 = 6 + \alpha_6 = 14 \quad (27)$$

$$\gamma_7 = 46 + \gamma_6 = 74 \quad (28)$$

$$\beta_7 = 46 + \beta_6 = 74. \quad (29)$$

These equations (eqs. 27–29) mean that the coefficients ( $\alpha_7$ ,  $\gamma_7$ , and  $\beta_7$ ) can be calculated by using the coefficients ( $\alpha_6$ ,  $\gamma_6$ , and  $\beta_6$ ), which have been already obtained by eqs. 24–26.

**Rule 2:** In general, if such generating functions up to  $x^n$  (for  $a(x)$  and  $b(x)$ ) or to  $x^{2n}$  (for  $c(x)$ ) are given, the coefficients ( $\alpha_{n+2}$ ,  $\gamma_{n+2}$ , and  $\beta_{n+2}$ ) are recursively calculated by using the generated coefficients ( $\alpha_{n+1}$ ,  $\gamma_{n+1}$ , and  $\beta_{n+1}$ ).

Comparison of the coefficients of the terms  $x^8$  (for  $a(x)$  and  $b(x)$ ) or  $x^8$  (for  $c(x)$ ) gives the following relationships:

$$\alpha_8 = 9 + \alpha_7 = 23 \quad (30)$$

$$\gamma_8 = 69 + 2\gamma_6 + \gamma_7 = 199 \quad (31)$$

$$\beta_8 = 69 + 2\beta_6 + \beta_7 = 199. \quad (32)$$

These equations (eqs. 30–32) mean that the coefficients ( $\alpha_8$ ,  $\gamma_8$ , and  $\beta_8$ ) can be calculated by using the coefficients ( $\gamma_6$  and  $\beta_6$ ) and the coefficients ( $\alpha_7$ ,  $\gamma_7$ , and  $\beta_7$ ), which have been already obtained by eqs. 25 and 26 as well as by eqs. 27–29.

**Rule 3:** In general, if such generating functions up to  $x^n$  (for  $a(x)$  and  $b(x)$ ) or to  $x^{2n}$  (for  $c(x)$ ) are given, the coefficients ( $\alpha_{n+3}$ ,  $\gamma_{n+3}$ , and  $\beta_{n+3}$ ) are recursively calculated by using the generated coefficients ( $\gamma_{n+1}$  and  $\beta_{n+1}$ ) and the calculated coefficients ( $\alpha_{n+2}$ ,  $\gamma_{n+2}$ , and  $\beta_{n+2}$ ).

## 2.5 Implementation and Results

In the preceding subsection, we have obtained three rules, each of which is capable of working as a basis of recursive calculation.

First, we implement Rule 1 as a Maple program for calculating the numbers of planted 3D-trees (monosubstituted alkanes) which are itemized with respect to carbon content (1 to 100). The upper limit of the carbon content (`ccntt`) is tentatively fixed to be 100, but it is freely selected.

Maple program for Rule 1, "EnumAlky11-100RR.mpl":

```
ax := 1 + x*a1*c2;
cx := 1+ (1/3)*x^2*c2^3 + (2/3)*x^2*c6;
bx := 1 + (1/3)*x*b1^3 + (2/3)*x*b3;

"Initial Values";
a1 := 1;
c2 := 1;
b1 := 1;
b3 := 1;
c6 := 1;

for ccntt from 1 to 100 by 1 do
ccntt:
Cbx:= coeff(bx,x^ccntt):
Cax:= coeff(ax,x^ccntt):
Ccx:= coeff(cx,x^(ccntt*2)):
a1 := a1 + Cax*x^ccntt:
c2 := c2 + Ccx*x^(ccntt*2):
b1 := b1 + Cbx*x^ccntt:
b3 := b3 +Cbx*x^(ccntt*3):
c6 := c6 + Ccx*x^(ccntt*6):
end do:

BBx := (1/2)*(a1 + b1):
CCx := (1/2)*(b1 - a1):

for ccntt from 1 to 100 by 1 do
printf("%d & %d & %d & %d \\\n",
ccntt,
coeff(a1,x^ccntt),
coeff(CCx,x^ccntt),
coeff(BBx,x^ccntt));
end do;
```

This code is stored in a file named "EnumAlky11-100RR.mpl" tentatively. In this code, the abbreviated symbols for functional equations are used as follows: `a1` for  $a(x)$ , `c2` for  $c(x^2)$ , `c6` for  $c(x^6)$ , `b1` for  $b(x)$ , and `b3` for  $b(x^3)$ . Such symbols as `a1` and `c2` are used to remember that such functions as  $a(x)$  and  $c(x^2)$  correspond to the SIs  $a_1$ ,  $c_2$  and so on. The first paragraph of this code declares three functional equations (eqs. 11, 9, and 7). In the 2nd paragraph, the initial values  $a(x) = 1$ ,  $c(x^2) = 1$ , and  $b(x) = 1$  are set by encoding `a1 := 1; c2 := 1; ... c6 := 1;`, which means that we put  $\alpha_0 = 1$ ,  $\gamma_0 = 1$ , and  $\beta_0 = 1$  for the initial (trivial) planted promolecule. The 3rd paragraph of the code shows a `do` loop, in which the next coefficients are calculated by using a Maple command `coeff` and added to the end of respective functional equations so as to generate intermediate generating functions. After escaping from the `do` loop, the 4th paragraph declares the calculation of  $B(x)$  (eq. 14) and  $C(x)$  (eq. 16). The 5th paragraph of the code (the final `do` loop) shows the print-out of the calculation results.

We execute the code by inputting the following command on the Maple inputting window:

```
read "EnumAlky11-100RR.mpl";
```

Table 1. Numbers of Monosubstituted Alkanes as Stereoisomers<sup>a</sup>

$k$	$\alpha_k$ (Achiral)	$C_k$ (Chiral)	$B_k$ (Achiral + Chiral)
1	1	0	1
2	1	0	1
3	2	0	2
4	3	1	4
5	5	3	8
6	8	10	18
7	14	30	44
8	23	88	111
9	41	255	296
10	69	742	811
11	122	2157	2279
12	208	6312	6520
13	370	18563	18933
14	636	54932	55568
15	1134	163479	164613
16	1963	489264	491227
17	3505	1471692	1475197
18	6099	4447896	4453995
19	10908	13500689	13511597
20	19059	41140608	41159667
21	34129	125818217	125852346
22	59836	386050543	386110379
23	107256	1188093392	1188200648
24	188576	3666547089	3666735665
25	338322	11344058829	11344397151
26	596252	35180323336	35180919588
27	1070534	109339097119	109340167653
28	1890548	340508394528	340510285076
29	3396570	1062419370490	1062422767060
30	6008908	3320666310903	3320672319811
31	10801816	10395996250010	10396007051826
32	19139155	32596713516873	32596732656028
33	34422537	102354659356690	102354693779227
34	61074583	321832884381903	321832945456486
35	109894294	1013230630964436	1013230740858730
36	195217253	3193818534035050	3193818729252303
37	351404205	10078773177202180	10078773528606385
38	624913284	31840095204449021	31840095829362305
39	1125291874	100689937510048385	100689938635340259
40	2003090071	318728146981844679	318728148984934750
41	3608175239	1009849492883144730	1009849496491319969
42	6428430129	3202399570304430261	3202399576732860390
43	11582995444	10163870352469178059	10163870364052173503
44	20653101216	32284255140318002679	32284255160971103895
45	37223637886	102625414714074250195	102625414751297888081
46	66420162952	32646531262854125327	326465312694961416234
47	119740546576	1039258352830707100287	1039258352950447646863
48	213802390264	3310561173602401804486	3310561173816204194750
49	385525375648	10552597199971538944446	10552597200357064320094
50	688796847976	33657783567745352117595	33657783568434148965571
51	1242290116020	107416128908320489280104	107416128909562779396124
52	2220789746752	343005708844148436621963	343005708846369226368715
53	4006120350070	1095901875016767892361183	1095901875020774012711253
54	7165331870036	3503253701348594788382175	3503253701355760120252211
55	12928003493212	11204506183709361691881774	11204506183722289695374986
56	23134244857388	3585313432749875133777601	35853134327521885582634989
57	41746843378398	114780284499516103424710066	114780284499557850268088464
58	74738504448344	367625900577928930637205709	36762590057800369141654053
59	134890383077280	1177976917802372651276354835	1177976917802507541659432115
60	241593618271624	3776181956727075855700156899	3776181956727317449318428523
61	43690806635266	12110118510764409721547578931	12110118510764845821354214197
62	781379784134280	38852360869430891549389142546	38852360869431672929173276826
63	1410660331905744	124696551878602256306336431059	124696551878603666666336803
64	2528481353408387	400363258120183679194038099086	40036325812018620767538407473
65	4565372792411543	1285913332058160943022006622567	1285913332058165508394799034110
66	818586868237027	4131636451627223217905517665230	413163645162723140377420002257
67	14782032579934260	13279472925491136452203415411282	13279472925491151234235995345542
68	26513372758159909	42695640515935232895574700033255	42695640515935259408947458193164
69	47883315932669529	137317535156294699435253007511481	137317535156294747318568940181010
70	85911079008919616	44177814668085655381196140461152	441778146680856641292275149380768
71	155172762932069894	142172102187291518643461224418285	1421721021872915341607375056488179
72	278488395251454441	4576703637505875523107359575732692	4576703637505875801595754827187133
73	503057870689711695	14737232019966626470926756674550461	14737232019966626973984627364262156
74	903086480091036027	4746794100812136027677418552412327	47467941008121361179863898643448354
75	1631482280676136454	152933618235177166865457366439774884	152933618235177168496939647115911338
76	2929592464502366968	492856526906819180501066288772624509	492856526906819183430658753274991477
77	529297972572170684	1588729810112513938828615985394313591	158872981011251394358159577966484275
78	9506763767988688492	5122579180406227728593146855624013995	5122579180406227738099910623612702487
79	17177649111373079244	16520897953013100680478350495809265213	1652089795301310069765599607182344457
80	30860176979309781255	53294486738307619893870635366392366797	53294486738307619924730812345702148052
81	55765496914378281449	171961814511179425854645791758391916817	17196181451117942591041128867270198266
82	10020686580368330935	554983895358241328803983143348747181065	5549838953582413289041900808929115512000
83	181091912344282520796	179153371698949081248363857574020950270	1791533716989490812664730448808449203498
84	325479416128007866823	5784469517534031671596027732477042276730	5784469517534031671921507148605050143553
85	588244550847580438141	18680728838224735253871870332518735488272	18680728838224735254460114883366315926413
86	105747629038222641392	60341164324086142450338425947342014158465	6034116432408614245139590223724236799857
87	191133289853080790442	194949030688973491805812076982184909280356	194949030688973491807723409880719990070798
88	3436634060302329898084	629961494518672050471790513680381903001241	629961494518672050475227147740684232899325
89	6211963598216168485372	2036059814107841938917943348390976037599701	2036059814107841938924155311988292206085073
90	11171367303208512523636	6581862425540511747351642675384991026248219	6581862425540511747362814042688199538771855
91	20194381412648318673600	21280753827037894634102485750210174037092208	2128075382703789463412268013162282355765808
92	36323270806296768489016	68818216387594067927142141836032850951612302	68818216387594067927178465106839147720101318
93	65665395161399373233104	22258522679763459441793142423987786427153211	22258522679763459441858807819149185800386315
94	118131205680564887503088	720052435262314352511798946720114468071962449	72005243526231435251191707925795032959465337
95	213571105479833409623678	2329727991559244716140309972474067355488212719	2329727991559244716140523543579547188897836597
96	384274151006716654639812	75390752851895918718762265049972022532851893	7539075285189591871876606924650728739187491705
97	69477503076787167713632	24400640566533283392478279276091469455998528189	24400640566533283392478974051122237330166241821
98	1250289517934799230260092	7898656403619995818013298227853048427084400480	78986564036199958180134232568052969626314660573
99	226067318078343257290736	255724695264031766423233519459641127061308852807	255724695264031766423235780132821910493881643542
100	4068826452827659622280676	828053158706032477813824555141644137978761399059	828053158706032477813828623968096965638383679735

<sup>a</sup>In each value of  $C_k$  (Chiral) or  $B_k$  (Chiral + Achiral), each enantiomeric pair is counted just once for chiral monosubstituted alkanes.

Thereby, we obtain the coefficients  $\alpha_k$  (eq. 10) for  $a(x)$  (eq. 11),  $B_k$  (eq. 12) for  $B(x)$  (eq. 14), and  $C_k$  (eq. 13) for  $C(x)$  (eq. 16), as collected in Table 1.

Rule 2 is implemented as a Maple program for calculating the numbers of planted 3D-trees (monosubstituted alkanes), which are itemized with respect to carbon content (1 to 100).

Maple program for Rule 2, "Enum2by1-100RR.mpl":

```
ax := 1 + x*a1*c2;
cx := 1 + (1/3)*x^2*c2^3 + (2/3)*x^2*c6;
bx := 1 + (1/3)*x*b1^3 + (2/3)*x*b3;

a1 := 1 + x + x^2;
c2 := 1 + x^2 + x^4;
b1 := 1 + x + x^2;
b3 := 1 + x^3 + x^6;
c6 := 1 + x^6 + x^12;

for ccntt from 3 to 100 by 2 do
ccntt:
ccntt1:=ccntt+1:
Cbx:= coeff(bx,x^ccntt):
Cax:= coeff(ax,x^ccntt):
Ccx:= coeff(cx,x^(ccntt*2)):
Cbxn:= Cbx + coeff(bx,x^ccntt1):
Caxn:= Cax + coeff(ax,x^ccntt1):
Ccxn:= Ccx + coeff(cx,x^(ccntt1*2)):
a1 := a1 + Cax*x^ccntt + Caxn*x^ccntt1:
c2 := c2 + Ccx*x^(ccntt*2) + Ccxn*x^(ccntt1*2):
b1 := b1 + Cbx*x^ccntt + Cbxn*x^ccntt1:
b3 := b3 + Cbx*x^(ccntt*3) + Cbxn*x^(ccntt1*3):
c6 := c6 + Ccx*x^(ccntt*6) + Ccxn*x^(ccntt1*6):
end do;

BBx := (1/2)*(a1 + b1):
CCx := (1/2)*(b1 - a1):

for ccntt from 1 to 100 by 1 do
printf("%d & %d & %d & %d \\\n",
ccntt,
coeff(a1,x^ccntt),
coeff(CCx,x^ccntt),
coeff(BBx,x^ccntt));
end do;
```

Thereby, we alternatively obtain the coefficients  $\alpha_k$  (eq. 10) for  $a(x)$  (eq. 11),  $B_k$  (eq. 12) for  $B(x)$  (eq. 14), and  $C_k$  (eq. 13) for  $C(x)$  (eq. 16), which turn out to be identical to those collected in Table 1. It should be noted that a pair of enantiomers is counted just once for the values of  $B_k$  and  $C_k$ .

Rule 3 is further implemented as a Maple program for calculating the numbers of planted 3D-trees (monosubstituted alkanes), which are itemized with respect to carbon content (1 to 100).

Maple program for Rule 3, "Enum3by1-100RR.mpl":

```
ax := 1 + x*a1*c2;
cx := 1 + (1/3)*x^2*c2^3 + (2/3)*x^2*c6;
bx := 1 + (1/3)*x*b1^3 + (2/3)*x*b3;

a1 := 1 + x + x^2 + 2*x^3;
c2 := 1 + x^2 + x^4 + 2*x^6;
b1 := 1 + x + x^2 + 2*x^3;
```

```
b3 := 1 + x^3 + x^6 + 2*x^9;
c6 := 1 + x^6 + x^12 + 2*x^18;

for ccntt from 4 to 100 by 3 do
ccntt:
ccntt1:=ccntt+1:
ccntt2:=ccntt+2:

Cbx:= coeff(bx,x^ccntt):
Cax:= coeff(ax,x^ccntt):
Ccx:= coeff(cx,x^(ccntt*2)):

Cbxn:= Cbx + coeff(bx,x^ccntt1):
Caxn:= Cax + coeff(ax,x^ccntt1):
Ccxn:= Ccx + coeff(cx,x^(ccntt1*2)):

Cbxnn:= 2*Cbx + Cbxn + coeff(bx,x^ccntt2):
Caxnn:= Caxn + coeff(ax,x^ccntt2):
Ccxnn:= 2*Ccx + Ccxn + coeff(cx,x^(ccntt2*2)):

a1 := a1 + Cax*x^ccntt + Caxn*x^ccntt1
+ Caxnn*x^ccntt2:
c2 := c2 + Ccx*x^(ccntt*2) + Ccxn*x^(ccntt1*2)
+ Ccxnn*x^(ccntt2*2):
b1 := b1 + Cbx*x^ccntt + Cbxn*x^ccntt1
+ Cbxnn*x^ccntt2:
b3 := b3 + Cbx*x^(ccntt*3) + Cbxn*x^(ccntt1*3)
+ Cbxnn*x^(ccntt2*3):
c6 := c6 + Ccx*x^(ccntt*6) + Ccxn*x^(ccntt1*6)
+ Ccxnn*x^(ccntt2*6):
end do;

BBx := (1/2)*(a1 + b1):
CCx := (1/2)*(b1 - a1):

for ccntt from 1 to 100 by 1 do
printf("%d & %d & %d & %d \\\n",
ccntt,
coeff(a1,x^ccntt),
coeff(CCx,x^ccntt),
coeff(BBx,x^ccntt));
end do;
```

We execute this code to obtain the coefficients  $\alpha_k$  (eq. 10) for  $a(x)$  (eq. 11),  $B_k$  (eq. 12) for  $B(x)$  (eq. 14), and  $C_k$  (eq. 13) for  $C(x)$  (eq. 16), which are identical with those collected in Table 1.

### 3 Enumeration of Monosubstituted Alkanes as Graphs

It is worthwhile to mention the difference between the present work and previous results based on Pólya's theorem. For this purpose, we implement the Pólya's treatment [4, 5](Section 40) by using the Maple programming language on the same line as the present treatment based on Fujita's proligand method.

#### 3.1 Implementation of Pólya's Treatment

Pólya's treatment has been implicitly based on the condition that substituents for a given skeleton are limited to atoms (or graphs), because it uses the following functional equation:

Table 2. Numbers of Monosubstituted Alkanes as Graphs

$k$	$\beta_k$	$R_k$
1	1	1
2	1	1
3	2	2
4	5	4
5	11	8
6	28	17
7	74	39
8	199	89
9	551	211
10	1553	507
11	4436	1238
12	12832	3057
13	37496	7639
14	110500	19241
15	328092	48865
16	980491	124906
17	2946889	321198
18	8901891	830219
19	27012286	2156010
20	82300275	5622109
21	251670563	14715813
22	772160922	38649152
23	2376294040	101821927
24	7333282754	269010485
25	22688455980	712566567
26	70361242924	1891993344
27	218679264772	5034704828
28	681018679604	13425117806
29	2124842137550	35866550869
30	6641338630714	95991365288
31	20792003301836	257332864506
32	65193446172901	690928354105
33	204709353135917	1857821351559
34	643665829838389	5002305607153
35	2026461371823166	13486440075669
36	6387637263287353	36404382430278
37	20157546705808565	98380779170283
38	63680191033811326	266158552000477
39	201379876145388644	720807976831447
40	637456295966779429	1954002050661819
41	2019698989374464699	5301950692017063
42	6404799147037290651	14398991611139217
43	20327740716521351562	39137768751465752
44	64568510301289106574	106465954658531465
45	205250829465372138276	289841389106439413
46	652930625323502669516	789642117549095761
47	2078516705781154747150	215281494597165556
48	6621122347418605999236	5873225808361331954
49	21105194400328603264540	16033495247557039074
50	67315567136179501083166	4379755494193757760
51	214832257817883268676228	119710153806425838814
52	686011417690517662990678	327387061440387339008
53	2191803750037541905072436	895843085951178143691
54	7006507402704354908634386	2452637333887058055384
55	22409012367431651387256760	6718266278929859733622
56	71706268655020636920412590	18411771398697833823000
57	229560568999073953692798530	50482497809955340308577
58	735251801155932599778859762	138479602828722190112039
59	2355953835604880192935786950	380035071896650874765695
60	7552363913454393305018585422	1043393111652308735060544
61	24220237021529255542901793128	2865843809192062819497928
62	77704721738862564478562419372	7874653977974760120343666
63	249393103757205923273004767862	21646080934050220557892722
64	800726516240369886869415206559	59523793603181888124081873
65	2571826664116326451416805656677	163742112144282534714835522
66	8263272903254454621679717667487	450592339345898249869218319
67	26558945850982287686439410756824	1240384226853570945094452814
68	85391281031870492304522158226419	341564928884560981955252130
69	274635070312589446753821947692491	940872333538775252898214212
70	883556293361713196673471289841920	25925341104562842786074383001
71	2843442043745830528041987180906464	71457982691643838449175520936
72	9153407275011751324703114402919825	1970179583330028225614336794
73	29474464039933253444911384038812617	543358066498116791491048793218
74	94935882016242721456641317195860681	1498953849616752546689802091439
75	305867236470354335362397013555686222	4136270918924306555319895187994
76	985713053813638363931725042047615986	11416819269749659903244393933095
77	3177459620225027881870211763360797866	31520543436975671843005866727635
78	10245158360812455466693057479236716482	87046595485364266327991021865784
79	33041795906026201378134350102991609670	240445481264024116481889057815169
80	106588973476615239818601447712094514849	664332507296471677736372953010226
81	343923629022358851765057080431162115083	1835928910350418500278990976469861
82	1109967790716482657708173152277862693065	5074873286445647518042353897671634
83	358306743397898162514836963824701526200	14031081066302277328140290178287812
84	11568939035068063343517534881082092420283	38801752647739183194830502890941088
85	37361457676449470508331985215885051414685	107325668225681414089071298904099717
86	120682328648172284901734328185066250958322	296924311570321654884671456351866990
87	389898061377946983613535486862904899351154	8216289590391066552460647711971525658
88	1259922989037344100947017661421066135900566	2274005628877041705221344665002373781
89	407211962821568387784209860378368243684774	6294934468640373568081967026378999337
90	13163724851081023494714456718073190565020074	1742901725231375322020339177696372341
91	42561507654075789268225165881832996392858016	48265274930187086026934877354369433592
92	137636432775188135854320606942871998671713620	133682671606330111681925831160058336852
93	445170445359526918883651950243136972227539526	37033281904758223646063855048514321953
94	1440104870524628705023716024645909501031427986	1026087505221182547255970631028299069618
95	4659455983118489432280833516053614544386049116	2843479637566109306175945756955049227258
96	15078150570379183743752829575150450761720343598	7881118057903303046879598021785329267924
97	48801281133066566784957253327213706786164770010	21847212898278527375129113460976162456395
98	157973128072399916360267214846588004453399061052	60572193491813719514851889145800693998935
99	511449390528063532846469299592463037555190496350	167964775240641674296692581041350933570844
100	1656106317412064955627653179109741103617145078794	465832107463665143875912088515974576774802



$$r(x) = 1 + \frac{x}{6} (r(x)^3 + 2r(x^3) + 3r(x)r(x^2)) \quad (33)$$

in place of eq. 15. Obviously, eq. 33 is a special case of eq. 15, where we put  $r(x) = B(x)$  in the left-hand side and  $r(x) = b(x) = a(x) = c(x)$  in the right-hand side. As a result, eq. 33 turns out to have recursive nature. The functional equation eq. 33 was first noted by Pólya [4] by using the symmetric group of degree 3 ( $S^{[3]}$ ), where an equation equivalent to eq. 7 was also derived by using the alternating group of degree 3 ( $A^{[3]}$ ).

We implement the functional equation (eq. 33) as a Maple program for calculating the numbers of planted trees (monosubstituted alkanes as graphs), which are itemized with respect to carbon content (1 to 100).

Maple program for counting planted trees, "EnumPTr1-100RR.mpl":

```
bx := 1 + (1/3)*x*b1^3 + (2/3)*x*b3;
rx := 1 + (1/6)*x*r1^3 + (1/3)*x*r3
+ (1/2)*x*r1*r2;

"Initial Values";
b1 := 1;
b3 := 1;
r1 := 1;
r2 := 1;
r3 := 1;

for cntt from 1 to 100 by 1 do
  cntt:
  Cbx:= coeff(bx,x^cntt):
  Crx:= coeff(rx,x^cntt):
  b1 := b1 + Cbx*x^cntt:
  b3 := b3 + Cbx*x^(cntt*3):
  r1 := r1 + Crx*x^cntt:
  r2 := r2 + Crx*x^(cntt*2):
  r3 := r3 + Crx*x^(cntt*3):
end do:
for cntt from 1 to 100 by 1 do
  printf("%d & %d & %d \\\ \n",
  cntt, coeff(b1,x^cntt), coeff(r1,x^cntt));
end do;
```

We execute this code to obtain the coefficients  $\beta_k$  for  $b(x)$  (eq. 7) and  $R_k$  for  $r(x)$  (eq. 33), as collected in Table 2. It should be noted that the two enantiomers of each pair are counted separately in the evaluation of  $\beta_k$ , while structural isomers (constitutional isomers) are counted as graphs in the evaluation of  $R_k$ . The results of  $R_k$  up to 20 are identical to those of Henze and Blair [3], the results from carbon content 21 to 30 are identical to those of Perry [27], and the results from carbon up to 50 are identical to those collected in [28](pages 153–154 of Vol. II).

## 4 Discussions

### 4.1 Stereoisomers vs. Graphs

Let us exemplify the difference between the enumeration results based on Fujita's proligand method (Table 1) and those based on Pólya's theorem (Table 2).

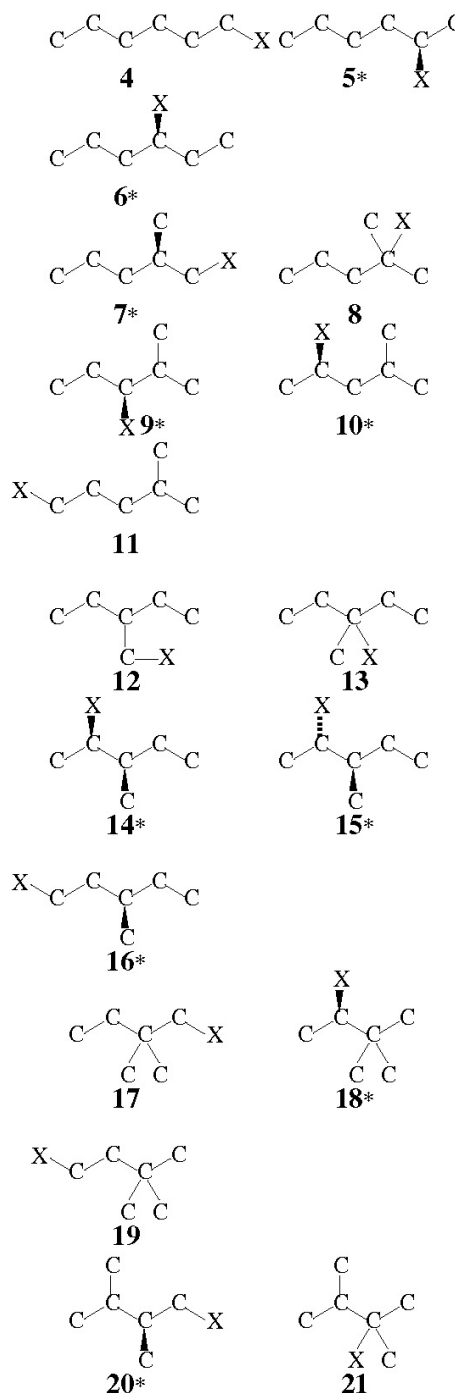


Figure 3. Planted 3D-trees or monosubstituted alkanes of carbon content 6. A planted 3D-tree with an asterisk is chiral, where the chirality is shown by a wedged bond and/or a hashed dash bond.

The  $k = 6$  row of Table 1 indicates the presence of 8 achiral monosubstituted alkanes and of 10 enantiomeric pairs of chiral monosubstituted alkanes. The structural formulas of monosubstituted alkanes of carbon content 6 are shown in Figure 3, where either one selected from each pair of enantiomers is depicted as a representative, which is designated by an asterisk. In agreement with the calculation result shown in Table 1, there exist 8 achiral monosubstituted alkanes (**4**, **8**, **11**, **12**, **13**, **17**, **19**, and **21**) and 10 enantiomeric pairs of chiral monosubstituted alkanes (**5**, **6**, **7**, **9**, **10**, **14**, **15**, **16**, **18**, and **20**).

On the other hand, Table 2 ( $R_6 = 17$ ) indicates the presence of 17 monosubstituted alkanes as graphs (constitutional isomers). The difference between the value and the value of Table 1 ( $B_6 = 18$ ) comes from the presence of diastereoisomers, **14** and **15**, which are equalized in the enumeration based on Pólya's theorem so as to be counted once as one graph. Chemically speaking, **14** and **15** are so congruent as to be regarded as one constitutional (structural) isomer. The value  $\beta_6 = 28$  in the  $k = 6$  row of Table 2 can be calculated from the data collected in the  $k = 6$  row of Table 1, i.e.,  $8 + 10 \times 2 = 28$ , which are the number of stereoisomers, where enantiomers are counted separately.

The  $k = 7$  row of Table 1 indicates the presence of 14 achiral monosubstituted alkanes and of 30 enantiomeric pairs of chiral monosubstituted alkanes. On the other hand, Table 2 indicates the presence of 39 monosubstituted alkanes as graphs. Manual enumeration shows that there are 14 achiral stereoisomers, 20 pairs of enantiomers with one asymmetric carbon, and 5 quartets of diastereoisomers with two asymmetric carbons. Thereby, we obtain the relationships  $14 + 20 \times 2 + 5 \times 4 = 74$  ( $= \beta_7$ ) and  $14 + 20 + 5 = 39$  ( $= R_7$ ). Note that  $\beta_7$  can be alternatively calculated by the data of Table 1, i.e.,  $14 + 30 \times 2 = 74$  ( $= \beta_7$ ).

The evaluation of  $R_7$  from  $B_7$  and  $C_7$  would require the evaluation of the number of diastereomeric quartets. In general, the evaluation of  $R_k$  from  $B_k$  and  $C_k$  requires the evaluation of the number of diastereomeric quartets and higher stereoisomeric sets. The latter evaluation is open to further investigation.

## 4.2 Comments on Earlier Works

Robinson *et al.* [17] reported the enumeration of monosubstituted alkanes by modifying Pólya's cycle indices (CIs) as follows:

$$B(x) = 1 + \frac{x}{6} (r(x)^3 + 2r(x^3) + 3p(x)r(x^2)), \quad (34)$$

where the  $r(x)$  in eq. 33 is replaced by  $p(x)$  for achiral ligands. However, the transitivity between a chiral ligand and its enantiomeric ligand is specified by the term  $r(x^2)$ , which is the same as the term for two chiral ligands of

the same chirality. By comparison between eq. 15 and eq. 34, the latter presumes  $p(x) = a(x)$  and  $r(x) = b(x) = c(x)$  in the present context, where the transitivity for  $p/\bar{p}$  (or  $\bar{p}/\bar{p}$ ) is mixed up with the transitivity for  $p/p$  (or  $\bar{p}/\bar{p}$ ) if  $p$  and  $\bar{p}$  represent a pair of enantiomeric proligands. In other words, their treatment also lacks the concept of sphericities.

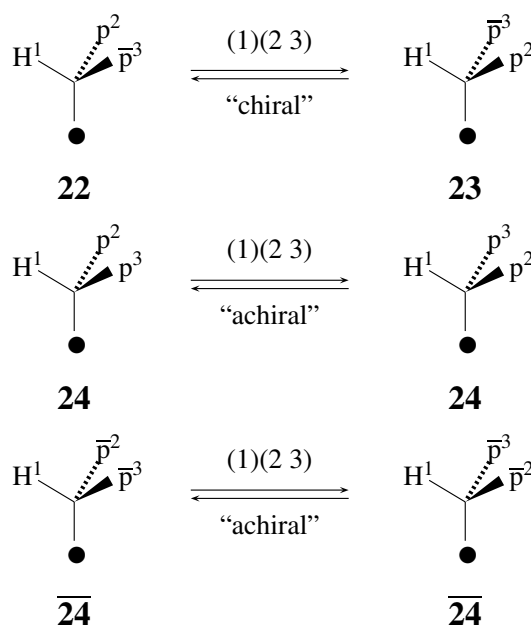


Figure 4. Erroneous interpretation of transitivity which comes from the permutation (1)(2 3) assigned erroneously to  $\sigma_{v(1)}$ .

An example of erroneous interpretation of transitivity by using eq. 34 is illustrated in Figure 4, where the term  $p(x)r(x^2)$  in eq. 34 corresponds to the action of the permutation (1)(2 3). When we put  $p = R\text{-CH}(\text{CH}_3)\text{CH}_2\text{CH}_3$  and  $\bar{p} = S\text{-CH}(\text{CH}_3)\text{CH}_2\text{CH}_3$ , the formula **22** represents pseudoasymmetry for carbon content 9. The action of the permutation (1)(2 3) converts **22** into its diastereomer **23**. This means that **22** and **23** are equalized so as to be regarded as one isomer, which is counted just once.

On the other hand, **24** is converted into the same **24** by the action of the permutation (1)(2 3). This means that **24** is counted just once. In a similar way, its enantiomer **24** is converted into the same **24** by the action of the permutation (1)(2 3). Thereby, **24** is counted just once, independent of **24**.

Consequently, in the enumeration based on eq. 34, individual isomers to be counted are **22** (or **23**), **24**, and **24**, which contribute by three to the coefficient of the term  $x^9$ . Obviously, the pseudoasymmetry between **22** and **23** is not treated properly.

Moreover, Robinson *et al.* [17] used the following functional equation:

$$p(x) = 1 + xp(x)r(x^2) \quad (35)$$

to enumerate achiral planted 3D-trees. Although this equation is akin to eq. 11, its connotation is entirely different from that of eq. 11. The misunderstanding connotation of eq. 35 stems from an unconscious assumption that the permutation (1)(2 3) can be assigned to the reflection operation  $\sigma_{v(1)}$ . The term  $r(x^2)$  in eq. 35 means that **24** is “achiral” and independently  $\overline{24}$  is also “achiral” because the permutation (1)(2 3) is erroneously assigned to the reflection operation  $\sigma_{v(1)}$ , as found in Figure 4. Note that an object superposable to its mirror image is achiral and that the permutation (1)(2 3) ( $\sim \sigma_{v(1)}$  erroneously) converts **24** (or  $\overline{24}$ ) into itself, which is inevitably but erroneously regarded as a mirror image (Figure 4). Hence, in the enumeration based on eq. 35, individual isomers to be counted as “achiral” are **24** and  $\overline{24}$ , which contribute by two to the coefficient of the term  $x^9$ .

By the same reason, we are forced to say that **22** and **23** are “chiral” so as to be “enantiomeric” to each other under the action of the permutation (1)(2 3) ( $\sim \sigma_{v(1)}$  erroneously). This means that the set of **22** and **23** is regarded as one isomer and counted just once as an “enantiomeric” pair of “chiral” stereoisomers. Note that the term *pseudoasymmetry* has been coined to explain this type of phenomenon (**22** vs. **23**). It can be said that chemical combinatorics based on eq. 35 using the criteria of Figure 4 still remains before the coinage of the term *pseudoasymmetry*.

In consequence, so long as we rely on eqs. 34 and 35, “achiral” and “chiral” are reversely recognized in cases of pseudoasymmetry, as shown in Figure 4.

In contrast to eqs. 34 and 35, the present enumeration based on Fujita’s proligand method uses eqs. 15 and 11. Because the permutation (1)(2 3) is assigned to the reflection operation  $\sigma_{v(1)}$ , the case of pseudoasymmetry described above is interpreted in agreement with geometry and point-group theory, as shown in Figure 5.

The action of the permutation (1)(2 3) converts **22** into itself as shown in Figure 5, because the one-cycle (1) is homospheric and the two-cycle (2 3) is enantiospheric. This means that **22** is achiral so as to be regarded as one stereoisomer, which is counted just once. On the same line, **23** is converted into itself by the action of the permutation (1)(2 3), so that it is achiral so as to be regarded as one stereoisomer, which is counted just once.

On the other hand, **24** is converted into its enantiomer  $\overline{24}$  by the action of the permutation (1)(2 3). This means that **24** and  $\overline{24}$  are regarded as one pair of enantiomers, which is counted just once.

Consequently, in the enumeration based on eq. 15, individual isomers to be counted are **22**, **23**, **24** (as one enantiomeric pair with  $\overline{24}$ ), which contribute by three to

the coefficient of the term  $x_9$ . Thus the pseudoasymmetry between **22** and **23** is treated properly.

In Fujita’s proligand method, the term  $c(x^2)$  in eq. 11 corresponds to the enantiosphericity, so that **22** is achiral and independently **23** is also achiral, because the permutation (1)(2 3) is correctly assigned to the reflection operation  $\sigma_{v(1)}$ , as found in Figure 5. Hence, in the enumeration based on eq. 11, individual isomers to be counted as being achiral are **22** and **23**, which contribute by two to the coefficient of the term  $x^9$ .

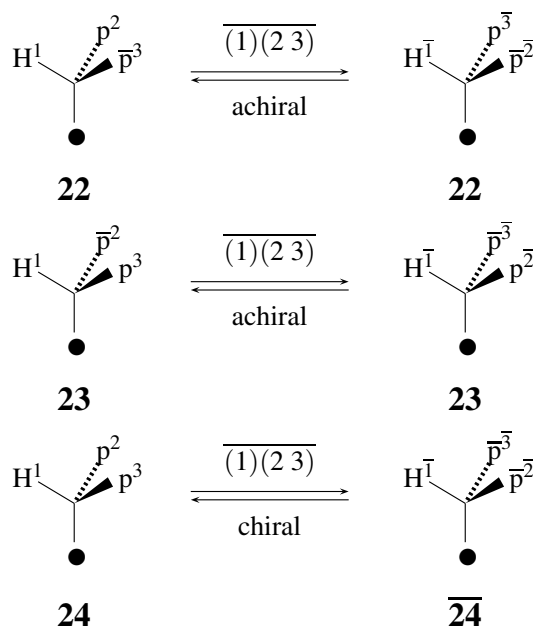


Figure 5. Correct interpretation of transitivity which comes from the permutation (1)(2 3) assigned properly to  $\sigma_{v(1)}$ .

In the enumeration based on eq. 17, a set of **24** and  $\overline{24}$  is counted just once as one pair of enantiomers, which contribute by one to the coefficient of the term  $x^9$ .

The enumeration based on the modified Pólya’s CIs (eq. 34 etc.) fortunately gives the same isomer numbers as obtained by means of Fujita’s CI-CFs (eq. 15 etc.). Enumeration of the two types provides the apparently same contribution (e.g., three to the coefficient of the term  $x^9$ ) because such a relationship between Figure 4 and Figure 5 holds true for any case of pseudoasymmetry. It should be noted, however, that the content of the former enumeration (cf. Figure 4) is quite different from the content of the latter enumeration (cf. Figure 5). Thus, the former enumeration erroneously shows the presence of one “chiral” isomer **22** (or **23**) and two “achiral” isomers **24** and  $\overline{24}$ , which contribute by three to the coefficient of the term  $x^9$  (Figure 4). The latter enumeration correctly shows the presence of two achiral isomers **22** and **23** and one chiral isomer **24** (as one enantiomeric pair with  $\overline{24}$ ),

which contribute by three to the coefficient of the term  $x^9$  (Figure 5).

The discussions described in this section have clearly shown that Fujita's proligand method [12–14] is concerned with the enumeration of stereoisomers (3D-objects), while Pólya's theorem [4, 5] is concerned with the enumeration of constitutional isomers (graphs). The difference between the two methodologies stems from the presence/absence of the sphericity concept so that Fujita's proligand method involves Pólya's theorem as a special case without sphericity. Moreover, it has been shown that the modification of Pólya's theorem by Robinson *et al.* [17] has failed in the proper treatment of pseudoasymmetric cases because of the lack of the sphericity concept.

### 4.3 Comments on a Stereochemical Convention

It is worthwhile to mention a stereochemical convention in comparison with the methodology adopted in the present approach. The IUPAC nomenclature (1996) [29] defines terms on stereoisomerism as follows:

1. (Def. 1) The term *stereoisomerism* is defined as "Isomerism due to differences in the spatial arrangement of atoms without any differences in connectivity or bond multiplicity between the isomers" [29](page 221). The related term *stereoisomers* is defined as "Isomers that possess identical constitution, but which differ in the arrangement of their atoms in space. See enantiomer, diastereoisomer, cis-trans isomers".
2. (Def. 2) The term *enantiomerism* is defined as "The isomerism of enantiomers" [29](page 2207). The related term *enantiomer* is defined as "One of a pair of molecular entities which are mirror images of each other and non-superposable. See also "enantiomorphism" [29](page 2207).

Because the term "stereoisomeric" (Def. 1) is concerned with a relationship between two compounds (as 3D-objects), the relevant term "stereoisomer" is incapable of specifying a compound (a 3D-object) without another 3D-object stereoisomeric to it. For example, **22** is not a "stereoisomer" by itself; **22** can be called a "stereoisomer" just when **23** is taken into consideration. Thereby, we are able to say that the compounds **22** and **23** are "stereoisomers". In a similar way, **24** cannot be a "stereoisomer" nor an "enantiomer" by itself; the **24** can be called a "stereoisomer" or an "enantiomer" just when **22** (or **23**) and/or **24** is taken into consideration. In the literal sense of the word "stereoisomer" defined in Def. 1, therefore, the expression "the number of stereoisomers" is permitted if a reference compound (e.g., **22**) is given or if a common skeleton for generating stereoisomers is

specified. When there exist two or more skeletons to be considered, therefore, the consideration of "the number of stereoisomers" is unsuitable for a systematic approach so as to require complicated treatments of various cases. In particular, the present enumeration problem for counting monosubstituted alkanes must consider an infinite number of skeletons which grow in their varieties in accord with the increase of carbon contents from zero to infinite. Then, what is aimed at by the expression "Numbers of Monosubstituted Alkanes as Stereoisomers"?

Obviously, Def. 1 contains two viewpoints, i.e., one viewpoint to recognize *differences* which are referred to by the expression "the differences in the spatial arrangement of atoms" and the other viewpoint to recognize *equivalences* which are referred to by the expression "without any differences in connectivity or bond multiplicity". However, a stereochemical convention tends to put emphasis on the viewpoint of "differences". On the same line, Def. 2 is mostly used in a context that enantiomers are non-superposable (i.e., different) instead of referring to their mirror-image equivalence. So long as we obey the tendency of the stereochemical convention, it is difficult to count "different" objects, as found easily. In contrast, the present approach puts emphasis on equivalence classes (e.g., sets of compounds equivalent under Def. 1 or Def. 2) so that such equivalence classes can be counted by employing group theory.

From a viewpoint of the group theory which the present approach employs, the emphasis of "differences" in the stereochemical convention implies that the alternating group of degree 3 ( $\mathcal{A}^{[3]}$ ) is used to recognize stereochemical "equivalence". Under the action of  $\mathcal{A}^{[3]}$ , two enantiomers of each pair are different from each other, so that they are counted separately, as found in the  $\beta_k$ -column of Table 2. Although the values collected in the  $\beta_k$ -column of Table 2 are adopted as "the numbers of stereoisomers" in terms of the stereochemical convention or as "the numbers of steric trees" in terms of the combinatorial enumeration by Pólya [4] and by Robinson *et al.* [17], we should mention several comments:

1. The use of  $\mathcal{A}^{[3]}$  is incapable of recognizing the chiral or achiral nature of each counted isomer. For example, the achirality of **22** (or **23**) and the chirality of **24** are not recognized properly under the action of  $\mathcal{A}^{[3]}$ . The recognition of achirality/chirality necessitates subsequent manual examination by stereochemists.
2. The use of  $\mathcal{A}^{[3]}$  is incapable of recognizing enantiomeric relationships. For example, the enantiomeric relationship between **24** and **24** is not recognized properly under the action of  $\mathcal{A}^{[3]}$ . The recognition of enantiomeric relationships again necessitates subsequent manual examination by stereochemists.

3. What is the subsequent manual examination described in the above-mentioned comments? In the stereochemical convention, the manual examination for recognizing achirality (e.g., **22** or **23**), chirality (e.g., **24** or  $\overline{\mathbf{24}}$ ), and enantiomeric relationship (e.g., between **24** and  $\overline{\mathbf{24}}$ ) is to consider the mirror image of each 3D-object. However, concrete processes of considering mirror images have not been fully formulated from the viewpoint of group theory.
4. Even if we adopt the symmetric group of degree 3 ( $S^{[3]}$ ) which is a supergroup of  $\mathcal{A}^{[3]}$ , the recognition of achirality/chirality and enantiomeric relationships has proved futile, as discussed in Section 3. The enumeration of this type has given the numbers of graphs, as shown in the  $R_k$ -column of Table 2.
5. Even by the attempt in which the symmetric group of degree 3  $S^{[3]}$  was used after modification [17], insufficient interpretations on achirality/chirality and on enantiomeric relationships took place, as discussed in Subsection 4.2.

These comments reveal why the stereochemical convention has adopted the values  $\beta_k$  (Table 2) as the numbers of stereoisomers. That is to say, the stereochemical convention lacks a mathematical formulation for recognizing achirality/chirality and enantiomeric relationships so that the convention is forced to rely on the alternating group of degree 3 ( $\mathcal{A}^{[3]}$ ).

In contrast, the present approach puts emphasis on “equivalences” or more precisely on “equivalence classes”, which are frequently called *orbits* in a mathematical context. The crux of the present approach is to use RCRs  $(C_s \setminus)C_{3v}$  and  $(C_1 \setminus)C_3$  (cf. Figure 2) in place of  $S^{[3]}$  and  $\mathcal{A}^{[3]}$ . As shown in Figure 5, for example, the use of  $\overline{(1)(2\ 3)}$  assigned to  $\sigma_{v(1)}$  is capable of treating the achirality of **22** (or **23**), which is superposable on its mirror image so as to be self-enantiomeric. The use of  $\overline{(1)(2\ 3)}$  is also capable of treating the enantiomeric relationship between **24** and  $\overline{\mathbf{24}}$ , which is characterized as being in a mirror-image relationship. Thereby, the manual examinations described above are formulated by the RCR  $(C_s \setminus)C_{3v}$ .

In terms of the present methodology, stereoisomer enumeration is concerned with the numbers of equivalence classes under the action of the RCR  $(C_s \setminus)C_{3v}$ . Note that the achiral compound **22** (or **23**) constructs an equivalence class under the action of  $(C_s \setminus)C_{3v}$ , where the achiral compound is counted just once. On the same line, the pair of **24** and  $\overline{\mathbf{24}}$  constructs an equivalence class under the action of  $(C_s \setminus)C_{3v}$ , where the pair is counted just once. It follows that each value in the  $C_k$ -column and the  $B_k$ -column of Table 1 represents the number in

which each achiral compound is counted just once and each enantiomeric pair is counted just once.

Consequently, the expression “Numbers of Monosubstituted Alkanes as Stereoisomers” (i.e., the title of this article) is permissible if the phrase “as Stereoisomers” is used as a simplified expression of the phrase “under the criterion of stereoisomerism” (cf. Def. 1) or more precisely “under the criterion of enantiomerism” (cf. Def. 2). The phrase “under the criterion of enantiomerism” means the action of the RCR of a point group, e.g.,  $(C_s \setminus)C_{3v}$ . This viewpoint means that enumeration concerned with stereoisomers should be done on the basis of equivalence classes provided “under the criterion of stereoisomerism” (cf. Def. 1) or more precisely “under the criterion of enantiomerism” (cf. Def. 2). More mathematically speaking, the enumeration concerned with stereoisomers should be done on the basis of equivalence classes provided by point groups through such RCRs as  $(C_s \setminus)C_{3v}$ .

When two enantiomers of each pair are counted separately according to the stereochemical convention, the value  $2C_k$  is used for counting chiral compounds; and the value  $\alpha_k + C_k (= \beta_k)$  is used for counting achiral and chiral compounds. These values should be adopted under the implicit restrictions described above for the stereochemical convention.

## 5 Conclusion

Planted three-dimensional (3D) trees or equivalently monosubstituted alkanes as stereoisomers are regarded as planted promolecules, which are enumerated by Fujita’s proligand method [12–14]. Cycle indices with chirality fittingness (CI-CFs), which are composed of three kinds of sphericity indices (SIs), i.e.,  $a_d$  for homospheric cycles,  $c_d$  for enantiospheric cycles, and  $b_d$  for hemispheric cycles, are used to obtain generating functions for calculating planted 3D-trees. Functional equations  $a(x)$ ,  $c(x^2)$ , and  $b(x)$  for recursive calculations are derived from the CI-CFs and programmed by means of the Maple programming language. Thereby, recursive calculations of the numbers of planted 3D-trees (or equivalently those of monosubstituted alkanes as stereoisomers) are executed and collected up to 100 carbon content in a tabular form. The results are compared with the enumeration of planted trees (as graphs).

We gratefully acknowledge the financial support given to our recent project by the Japan Society for the Promotion of Science: Grant-in-Aid for Scientific Research B (No. 18300033, 2006).

## References

- [1] A. Cayley, *Philos. Mag.*, **47**, 444-446 (1874).
- [2] A. Cayley, *Rep. Brit. Assoc. Advance. Sci.*, **45**, 257-305 (1875).
- [3] H. R. Henze and C. M. Blair, *J. Am. Chem. Soc.*, **53**, 3042 (1931).
- [4] G. Pólya, *Acta Math.*, **68**, 145-254 (1937).
- [5] G. Pólya and R. C. Read, *Combinatorial Enumeration of Groups, Graphs, and Chemical Compounds*, Springer-Verlag, New York (1987).
- [6] H. Hosoya, *Kagaku no Ryoiki*, **26**, 21-33 (1972).
- [7] D. H. Rouvray, *Chem. Soc. Rev.*, **3**, 355-372 (1974).
- [8] O. E. Polansky, *MATCH Commun. Math. Comput. Chem.*, **1**, 11-31 (1975).
- [9] *Chemical Applications of Graph Theory*, A. T. Balaban, ed., Academic Press, London (1976).
- [10] G. Pólya, R. E. Tarjan, and D. R. Woods, *Notes on Introductory Combinatorics*, Birkhäuser, Boston (1983).
- [11] K. Balasubramanian, *Chem. Rev.*, **85**, 599-618 (1985).
- [12] S. Fujita, *Theor. Chem. Acc.*, **113**, 73-79 (2005).
- [13] S. Fujita, *Theor. Chem. Acc.*, **113**, 80-86 (2005).
- [14] S. Fujita, *Theor. Chem. Acc.*, **115**, 37-53 (2006).
- [15] S. Fujita, *J. Am. Chem. Soc.*, **112**, 3390-3397 (1990).
- [16] S. Fujita, *Symmetry and Combinatorial Enumeration in Chemistry*, Springer-Verlag, Berlin-Heidelberg (1991).
- [17] R. W. Robinson, F. Harary, and A. T. Balaban, *Tetrahedron*, **32**, 355-361 (1976).
- [18] J. H. van't Hoff, *Archives Néerlandaises des Sciences exactes et naturelles*, **9**, 445-454 (1874).
- [19] J. H. van't Hoff, *La Chimie Dans L'Espace*, P. M. Bazendijk, Rotterdam (1875).
- [20] E. Fisher, *Ber. Dtsch. chem. Ges.*, **24**, 1836 (1891).
- [21] E. Fisher, *Ber. Dtsch. chem. Ges.*, **24**, 2683 (1891).
- [22] S. Fujita, *Theor. Chem. Acc.* (2006), in press.
- [23] S. Fujita, *Theor. Chem. Acta*, **91**, 315-332 (1995).
- [24] S. Fujita, *Theor. Chem. Acc.*, **99**, 224-230 (1998).
- [25] S. Fujita, *Bull. Chem. Soc. Jpn.*, **71**, 1587-1596 (1998).
- [26] M. B. Monagan, K. O. Geddes, H. M. Heal, G. Labahn, S. M. Vorkoetter, J. McCarron, and D. DeMarco, *Maple 9. Introductory Programming Guide*, Maplesoft, Waterloo (2006).
- [27] D. Perry, *J. Am. Chem. Soc.*, **54**, 2918-2920 (1932).
- [28] N. Trinajstić, *Chemical Graph Theory, Vol. I and II*, CRC, Boca Raton (1983).
- [29] IUPAC Recommendations 1996. Basic Terminology of Stereochemistry, *Pure Appl. Chem.*, **68**, 2193-2222 (1996).