Theor Chem Acc (2007) 117:353–370
DOI 10.1007/s00214-006-0162-9

REGULAR ARTICLE

Graphs to chemical structures. 4: Combinatorial enumeration of planted three-dimensional trees as stereochemical models of monosubstituted alkanes

Shinsaku Fujita

Received: 19 April 2006 / Accepted: 18 July 2006 / Published online: 7 September 2006

Abstract Planted three-dimensional (3D) trees, which are defined as a 3D version of planted trees, are enumerated by means of Fujita's proligand method formulated in Parts 1–3 of this series [Fujita in Theor Chem Acc 113:73–79, 80–86, 2005; Fujita in Theor Chem Acc 115:37–53, 2006]. By starting from the concepts of proligand and promolecule introduced previously [Fujita in Tetrahedron 47:31–46, 1991], a *planted promolecule* is defined as a 3D object in which the substitution positions of a given 3D skeleton are occupied by a root and proligands. Then, such planted promolecules are introduced as models of planted 3D-trees. Because each of the proligands in a given planted promolecule is regarded as another intermediate planted promolecule in a nested fashion, the given planted promolecule is recursively constructed by a set of such intermediates planted promolecules. The recursive nature of such intermediate planted promolecules is used to derive generating functions for enumerating planted promolecules or planted 3D-trees. The generating functions are based on 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. For the purpose of evaluating c_d recursively, the concept of *diploid* is proposed, where the nested nature of c_d is demonstrated clearly. The SIs are applied to derive functional equations for recursive calculations, i.e., $a(x)$, $c(x^2)$, and $b(x)$. Thereby, planted 3D-trees or equivalently monosubstituted alkanes as stereoisomers are enumerated recursively by counting planted promolecules. The resulting values are collected up to 20 carbon content in a tabular form. Now, the enumeration problem initiated by mathematician Cayley [Philos Mag 47(4):444–446, 1874] has been solved in such a systematic and integrated manner as satisfying both mathematical and chemical requirements.

Keywords Planted 3D-tree · Monosubstituted alkane · Planted promolecule · Cycle index · Sphericity index · Stereoisomer enumeration

1 Introduction

Although Pólya's theorem [1,2] has been widely applied to chemistry [3–8], its scope and limitations in chemical applications have not been mentioned until recently. This is because conventional ways for enumerating isomers by Pólya's theorem have implicitly adopted a condition that substituents for a given skeleton are limited to atoms (or graphs) even whether the skeleton is regarded as being achiral or chiral.

A typical example for implying the scope and limitations of Pólya's theorem [1,2] is an enumeration of aliphatic alcohols (monosubstituted alkanes more generally or planted trees mathematically), which has been a classical problem initiated by a mathematician Cayley [9,10]. To do this task, Pólya has used the symmetric group of degree 3 ($\mathcal{S}^{[3]}$) for topological isomers (i.e., isomers as graphs) and the alternating group of degree 3 ($\mathcal{A}^{[3]}$) for spatial isomers (i.e., isomers as 3D structures). The Pólya's treatment has been implicitly based on the condition that substituents for a given skeleton are limited to atoms (or graphs), because the

---

S. Fujita (✉)
Department of Chemistry and Materials Technology,
Kyoto Institute of Technology,
Matsugasaki, Sakyoku,
Kyoto 606-8585, Japan
e-mail: fajitas@chem.kit.ac.jp

terms $2\mathcal{S}^{[3]} - \mathcal{A}^{[3]}$ and $\mathcal{S}^{[3]} - \mathcal{A}^{[3]}$ were used to count topologically different isomers having asymmetric carbon centers (cf. Sect. 42 of [1,2]).

So long as the substituents have been considered to be atoms, graphs, or at most achiral ligands, it can be said that Pólya's theorem and its chemical applications have disregarded the inner structure of the substituents. Under this condition, fundamental problems on stereochemistry such as *meso*-compounds and pseudoasymmetry had not emerged as obvious difficulties. Even when Pólya's theorem took account of asymmetric carbon centers (cf. Sect. 42 of [1,2]), it overlooked the problems on *meso*-compounds and pseudoasymmetry. It should be noted that the stereochemical problems have once been major difficulties at the beginning of stereochemistry [11,12] and solved in the last quarter of the nineteenth century in a descriptive or non-mathematical fashion [11–14]. Although Robinson et al. [15] has accomplished the enumeration of monosubstituted alkanes by modifying Pólya's cycle indices (CIs), it is still desirable to develop a more systematic method for comprehending the problems on *meso*-compounds and pseudoasymmetry.

To characterize stereoisomers as three-dimensional (3D) objects having inner structure, we have developed the USCI (unit-subduced-cycle-index) approach originally by means of algebraic derivation [16–19] and then we have alternatively formulated the USCI approach in a diagrammatical fashion [20–22]. The Fujita's USCI approach has emphasized the concept of *sphericities* of orbits governed by coset representations. In addition, the concepts of *proligands* and *promolecules* have been proposed to characterize the chirality/achirality of ligands as substituents in a more abstract fashion [23–25].

From a viewpoint brought about by Fujita's USCI approach, it has been found that Pólya's theorem is deficient in the concepts described above. This finding has indicated that an alternative method compatible with Fujita's USCI approach should be formulated to characterize and enumerate stereoisomers as three-dimensional objects having inner structure. For this purpose, we have proposed *the proligand method* as reported recently in Parts 1–3 of this series [26–28], where the original sphericity concept (i.e., sphericity of orbits) has been transformed into a new concept *sphericities of cycles* [26–28] through an intermediate concept "sphericities of orbits for cyclic subgroups" [29–32].

In the preset paper, Fujita's proligand method will be applied to the enumeration of monosubstituted alkanes as 3D structures, where planted 3D-trees will be formulated by using the concepts of proligands and promolecules.

2 Planted three-dimensional trees

Cayley introduced the concept of a tree [10] and enumerated all rooted trees in connection with isomers of monosubstituted alkanes [9,10], as reviewed in a book [33]. Pólya developed a more systematic approach to the enumeration of trees and rooted trees, after he established the main theorem ("Hauptsatz") bearing his name (cf. Sect. 16 of [1,2]).

In the present paper, we use the terms *graph*, *tree*, *root*, *rooted tree*, *vertex*, *edge*, and *degree* in accord with the common usage adopted in textbooks [34,35]. The term *planted tree* is used in the meaning of Pólya (cf. Sect. 31 of [1,2]) to designate a kind of rooted tree in which the root is a vertex which is an endpoint playing a special role. The vertices of a planted tree which are different from the root are called *nodes*. Among such nodes, the node directly connected to the root through a single edge (called *a stem*) is referred to as *a principal node*. Chemically speaking, the term *stem* corresponds to the valence bond of an alkyl ligand and the term *root* corresponds to the terminal of the valence bond. Because the relationship between the mathematical term "planted trees" and the chemical term "monosubstituted alkanes" (or "aliphatic alcohols" or "alkyl ligands") can be clearly specified in this manner, we will put a special focus on the enumeration of planted trees.

Because planted trees have been defined as a kind of graphs [1,2], they are insufficient to treat monosubstituted alkanes (aliphatic alcohols, alkyl ligands, or alkyl radicals equivalently) as stereoisomers. Obviously, the concept of chirality/achirality cannot be introduced into the concept of graphs so long as we remain the conventional methodology. It follows that we shall introduce the concepts of *rooted and planted three-dimensional* (*3D*) *trees* as advanced concepts for treating stereoisomers properly.

1. A *three-dimensional* (*3D*) *tree* is defined as a 3D object which is composed of N_v vertices and N_e edges, where each edge is bounded by two vertices; the relationship $N_\mathrm{v} = N_\mathrm{e} + 1$ is satisfied; and each vertex and its incident edges construct a 3D arrangement exhibiting chirality or achirality. The number of incident edges on each vertex is referred to as *degree*.
2. A *rooted three-dimensional (3D) tree* is defined as a 3D-tree in which one vertex is discriminated as a *root* from the other vertices.
3. A *planted three-dimensional* (*3D*) *tree* is defined as a kind of rooted 3D-tree in which there emerge only one special vertex named *a root* which is an endpoint, only one special edge named *a stem* which is bound

by the root, and only one special vertex named *a principal node* which is linked by the stem.

Sometimes, 3D-trees are referred to as *free* 3D-trees, when they are discriminated from rooted or planted 3D-trees.

As found easily, the definitions described above come from those employed for trees, rooted trees, and planted trees as graphs. However, the main objective is to take account of the 3D structure exhibiting chirality or achirality. When such 3D structures are considered, the free rotation around each edge causes multiple conformational changes. As a result, it is not an easy task to test the congruence between two 3D trees, two rooted 3D-trees, or two planted 3D-trees.

To formulate the criteria for judging such congruence, let us consider achiral stereoisomers of 4,6-dimethylnonan-5-ol having a molecular formula $C_{11}H_{23}OH$, which are designated by the 3D structural formulas (**1** and **2**) shown in Fig. 1. Their 3D properties are expressed by wedges and hashed bonds attached to the carbon chain. Note that each carbon atom is designated by a joint of two or more bonds, each hydrogen atom is designated by an open circle, and a hydroxyl group is designated by a solid circle. Strictly speaking, the hydroxy group (OH) contains two atoms, but it is regarded as one vertex in order to formulate planted 3D-trees.

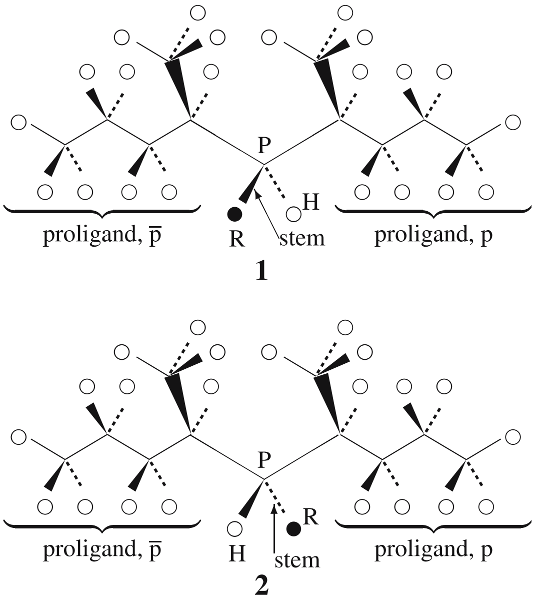


Fig. 1 Planted three-dimensional (3D) trees (**1** and **2**). Each joint of two or more edges represents a node. Each *open circle* represents an end vertex. The vertex P is referred to as a principal node. The vertex R designated by a *solid circle* is called *root*, which is an exceptional vertex and usually regarded as being not contained in the rooted 3D tree. The edge between the principal node and the root is called *stem*

The two structural formulas (**1** and **2**) can be regarded as planted 3D-trees, when the open circles and the joints are regarded as nodes as well as the solid circle is regarded as a root. According to the definition of planted trees, the vertex P is referred to as a principal node in **1** or **2** and the vertex R designated by a solid circle is the root of **1** or **2** as a planted 3D-tree. In contrast to the corresponding graphs (i.e., planted trees without 3D structural information), the planted 3D-trees (**1** and **2**) maintain the 3D properties expressed by wedges and hashed edges.

3 Planted promolecules

3.1 Planted promolecules as models of planted 3D-trees

Because the free rotation of each bond is permitted in agreement with chemical compounds, there emerge multiple conformational changes for each compound. On the same line, the corresponding planted 3D-tree contains one or more conformers, which exhibit different spatial arrangements due to rotations around edges (bonds). To enumerate planted 3D-trees, such conformers should be recognized to be congruent with each other so that they should be regarded as one "stereoisomer".

To assure this type of congruence, the concepts of proligands and promolecules [16,23] are applied to this case. A *proligand* is defined as a hypothetical ligand with chirality or achirality. Any ligand can be regarded as a chiral proligand or an achiral one in an abstract fashion without considering full information on spatial arrangements (i.e., conformation as stereochemical properties). A *skeleton* of $\mathbf{G}$-symmetry is defined as a 3D object having one ore more substitution positions. A *promolecule* is defined as a 3D object which is generated by placing a set of proligands on the positions of the skeleton. Among the substitution positions, a set of equivalent positions are numbered sequentially to give an orbit governed by a right or left coset representation (RCR or LCR), i.e., $(\mathbf{H}\backslash)\mathbf{G}$ or $\mathbf{G}(/\mathbf{H})$. Note that either right or left coset representation may be used and gives equivalent results.

To apply these concepts to a planted 3D-tree, we should define a *planted promolecule* by a "planting" procedure as follows:

1. We first select a skeleton and substitution positions. Let P be the principal node of a given planted 3D-tree. Suppose that each node linked to P is regarded as a substitution position. Then such substitution

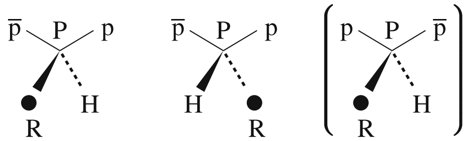
positions, the root (R), and the principal node (P) construct a *skeleton* of **G**-symmetry, in which the relevant edges (and the stem) are involved.

2. Next, we derive proligands by starting from the planted 3D-tree. When the P, the R, and the stem are deleted from the planted 3D-tree, the remaining part is partitioned into several segments, each of which is linked to the node chosen as each of the substitution positions in the selection of the skeleton. These segments can be regarded as proligands.
3. After the specification of the skeleton and the proligands, they are considered to construct a kind of promolecule having a root. To put emphasis on this kind of promolecule, a *planted promolecule* is defined as a promolecule in which one position is substituted by a special proligand called a *root*.

By following the planting procedure, for example, the planted 3D-tree (**1**) gives a planted promolecule (**3**) shown in Fig. 2, where a skeleton of $\mathbf{C}_{3v}$-symmetry is selected by the first step of the above procedure. The segments, i.e., an open circle (H), an *R*-CH(CH_3)($CH_2CH_2CH_3$) segment (p), and an *S*-CH(CH_3)($CH_2CH_2CH_3$) segment ($\overline{\text{p}}$), are regarded as proligands, where H is an achiral proligand, while p and $\overline{\text{p}}$ are chiral proligands of an enantiomeric pair. Similarly, the planted 3D-tree (**2**) gives a planted promolecule (**4** or equivalently **4**′) shown in Fig. 2.

Each segment recognized as a proligand (e.g., p and $\overline{\text{p}}$) can be further treated in terms of a planted 3D-tree, which is nested in the parent planted 3D-tree described above. This recognition of recursive nature is tentatively called "grafting" in order to be differentiated from the term "pruning". Thus, the grafting procedure is summarized for the sake of further discussions:

1. Suppose that the segment is modified as follows: the node corresponding to the principal node P of the parent planted 3D-tree is regarded as a new root, an edge incident to P is considered to be a new stem, and the alternative terminal node of the stem is considered as a new principal node. Thereby, the modified segment is regarded as a child planted 3D-tree.
2. The child planted 3D-tree is, then, converted into a planted promolecule by the planting procedure described above.



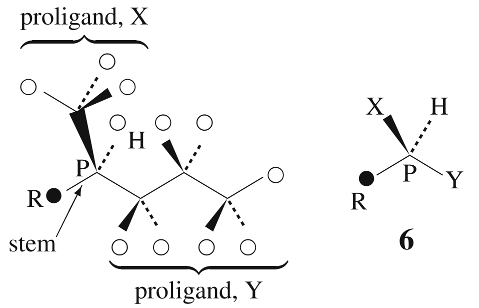
3 **4** (**4**′)

Fig. 2 Planted promolecule as models of the planted three-dimensional trees (**3** and **4**). The symbols p and $\overline{\text{p}}$ represent a pair of enantiomeric proligands. The symbol H represents an achiral proligand. Note that such a hydrogen atom (H) is also regarded as a proligand

Following the grafting procedure, for example, the right-hand segment denoted by p, i.e., *R*-CH(CH_3)($CH_2CH_2CH_3$), in the parent planted 3D-tree (**1**) is converted into a child planted 3D-tree (**5**), as shown in Fig. 3. The resulting planted 3D-tree (**5**) is regarded as a planted promolecule (**6**), where achiral proligands designated by H, X, and Y correspond to a hydrogen atom, a methyl ligand, and an *n*-propyl ligand, respectively.

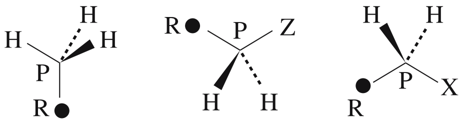
By applying the grafting procedure further, the proligand X (i.e., the methyl ligand) in **6** is regarded as a grandchild planted 3D-tree, which is ascribed to a subsidiary planted promolecule (**7**). On the other hand, the proligand Y (i.e., the *n*-propyl ligand) in **6** is regarded as another grandchild planted 3D-tree, which is ascribed to a planted promolecule (**8**). The symbol Z represents an achiral proligand corresponding to an ethyl ligand, which is further regarded as a planted 3D-tree.

The recursive application of the grafting procedure allows us to recognize that the proligand Z (i.e., the ethyl ligand) in **8** can be regarded as a further subsidiary planted 3D-tree. This is ascribed to a planted promolecule (**9**), which is of the same type as **8**, where the proligand Z is replaced by X corresponding to a methyl ligand. As a result, we are able to apply the grafting procedure recursively to a given planted 3D-tree until



5

Fig. 3 Planted 3D-tree derived from the proligand (p) and the corresponding planted promolecule



7 **8** **9**

Fig. 4 Planted promolecules corresponding to planted 3D-trees we reach its endpoints through intermediate planted 3D-trees. It should be emphasized that each intermediate planted 3D-tree can be regarded as a planted promolecule.

3.2 Symmetries of planted promolecules

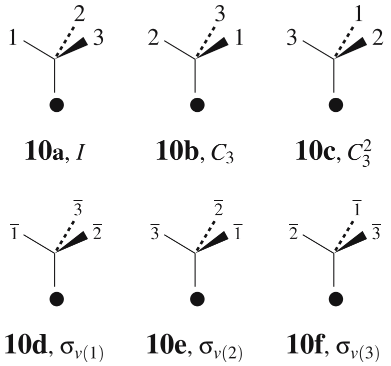
A given planted 3D-tree and any intermediate planted 3D-trees generated recursively can be ascribed to planted promolecules of the same skeleton, if all of the nodes except endpoints have the same degree (valency). For example, the planted promolecules relevant to the planted 3D-tree (**1** or **2**), i.e., **3**, **4**, **6**, **7**, and **8**, are based on the same skeleton of $\mathbf{C}_{3v}$-symmetry, although they are depicted in various ways.

Suppose that an orbit (equivalence class) of substitution positions in a skeleton of **G**-symmetry is governed by the RCR $(\mathbf{H}\backslash)\mathbf{G}$. A planted promolecule generated by placing a set of proligands on the substitution positions belongs to a subgroup **K** of **G**. The operations of **G** transform the planted **K**-promolecule into permuted ones which are congruent with the original planted **K**-promolecule. Under usual conditions, the skeleton corresponding to a planted 3D-tree belongs to $\mathbf{C}_{kv}$-point group, where the line from the root R to the principal node P through the stem is a k-fold axis of rotation. Moreover, the set of k positions of the skeleton is governed by the RCR $(\mathbf{C}_s\backslash)\mathbf{C}_{kv}$.

For example, the skeleton (**10** = **10a**) corresponding to the planted promolecules (**3**, **4**, **6**, **7**, and **8**) belongs to $\mathbf{C}_{3v}$, as shown in Fig. 5. The three positions other than the root construct an orbit governed by the RCR $(\mathbf{C}_s\backslash)\mathbf{C}_{3v}$, whose concrete form is shown as a product of cycles in Table 1. The numbering used in the the RCR $(\mathbf{C}_s\backslash)\mathbf{C}_{3v}$ corresponds to the numbering shown in the skeleton (**10**). It should be noted that 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.

Table 1 Right coset representation (RCR), cycle indices with chirality fittingness (CI-CF) and CI for a planted $\mathbf{C}_{3v}$-promolecule

Symmetry operation	RCR $(\mathbf{C}_s\backslash)\mathbf{C}_{3v}$	Product of SI	Product of variables
I	(1)(2)(3)	b_1^3	s_1^3
C_3	(1 2 3)	b_3	s_3
C_3^2	(1 3 2)	b_3	s_3
$\sigma_{v(1)}$	$\overline{(1)(2\ 3)}$	a_1c_2	s_1s_2
$\sigma_{v(2)}$	$\overline{(1\ 3)(2)}$	a_1c_2	s_1s_2
$\sigma_{v(3)}$	$\overline{(1\ 2)(3)}$	a_1c_2	s_1s_2


10a, I **10b**, C_3 **10c**, C_3^2

10d, $\sigma_{v(1)}$ **10e**, $\sigma_{v(2)}$ **10f**, $\sigma_{v(3)}$

Fig. 5 Symmetry operations for a $\mathbf{C}_{3v}$-skeleton (**10**). The identity operation (I) converts **10** into itself, where the resulting skeleton is denoted as **10a**

It is an easy task to obtain the product of cycles diagrammatically by applying the symmetry operations of $\mathbf{C}_{3v}$ to the reference skeleton (**10a**). Thereby, the three positions of the skeleton (**10a**) are permuted into congruent skeletons (**10a**–**10f**) in accord with the RCR $(\mathbf{C}_s\backslash)\mathbf{C}_{3v}$ (Fig. 5).

According to the concept of *sphericities of cycles* [26–28], d-cycles are classified into three types:

1. a homospheric cycle (a_d) if the corresponding symmetry operation is improper and d is odd,
2. an enantiospheric cycle (c_d) if the corresponding symmetry operation is improper and d is even, and
3. a hemispheric cycle (b_d) if the corresponding symmetry operation is proper and even whether d is odd or even.

Each of them is characterized by a sphericity index (SI), i.e., a_d, c_d, or b_d. Then, each symmetry operation is characterized by a product of sphericity indices, as shown in Table 1.

The planted promolecule (**3**) is considered to be generated from the skeleton (**10**) by placing an achiral proligand H on position 1, a chiral proligand p on position 2, and a chiral proligand $\overline{\mathrm{p}}$ on position 3. By the action of each improper operation (reflection), the achiral proligand H is converted into itself (i.e., $\overline{\mathrm{H}} = \mathrm{H}$), while the chiral proligand p (or $\overline{\mathrm{p}}$) is converted into its enantiomeric proligand $\overline{\mathrm{p}}$ (or $\mathrm{p} = \overline{\overline{\mathrm{p}}}$). The action of every symmetry operation of $\mathbf{C}_{3v}$ generates equivalent planted promolecules (**3a**–**3f**), as shown in Fig. 6.

Because **3a** and **3d** are identical except their numbering, the promolecule **3** is fixed by I and $\sigma_{v(1)}$. This means that **3** belongs to $\mathbf{C}_s$-symmetry. The $\mathbf{C}_s$-symmetry of **3** is interpreted by the concept of *chirality fittingness*, which has been originally formulated to characterize the accommodation capability of an orbit before and after subduction (i.e., desymmetrization) [16,36]. Thus, the orbit governed by the RCR $(\mathbf{C}_s\backslash)\mathbf{C}_{3v}$ is a three-membered homospheric orbit, which is capable of accommodating three achiral ligands (or proligands) of the same kind. On the other hand, the $(\mathbf{C}_s\backslash)\mathbf{C}_{3v}$-orbit is subduced into the subgroup $\mathbf{C}_s$ according to the following equation:

$$(\mathbf{C}_s\backslash)\mathbf{C}_{3v} \downarrow \mathbf{C}_s = (\mathbf{C}_s\backslash)\mathbf{C}_s + (\mathbf{C}_1\backslash)\mathbf{C}_s. \tag{1}$$

The resulting $(\mathbf{C}_s\backslash)\mathbf{C}_s$-orbit is a one-membered homospheric orbit so as to accommodate an achiral ligand or proligand. The other $(\mathbf{C}_1\backslash)\mathbf{C}_s$-orbit is a two-membered enantiospheric orbit so as to accommodate two achiral ligands (or proligand) or a pair of enantiomeric ligands (or proligands). The achiral proligand H of **3a** is in agreement with a_1 and the pair of chiral proligands p and $\overline{\text{p}}$ is in agreement with c_2 in the original meaning described in [16,36].

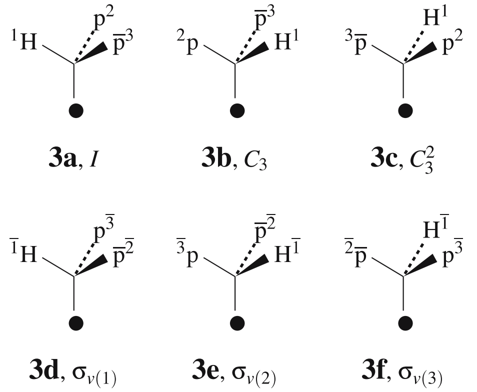
3.3 Enumeration of planted promolecules

Recently, the concept of *chirality fittingness* has been modified to characterize the transitivity of a cycle [26–28]:

1. A homospheric d-cycle (characterized by the SI a_d) interchanges d achiral proligands in a transitive fashion.
2. An enantiospheric d-cycle (characterized by the SI c_d) interchanges $d/2$ of chiral proligands of the same kind and $d/2$ of the corresponding enantiomeric chiral proligands of the same kind in a transitive fashion. Such an enantiospheric d-cycle is also capable of exchanging d achiral proligands of the same kind, where the half and the other half are exchanged by the action of the corresponding improper rotation.
3. A hemispheric d-cycle (characterized by the SI b_d) interchanges d proligands of the same kind in a transitive fashion, whether the proligands are achiral or chiral.

In terms of the chirality fittingness of a cycle, we have developed *the proligand method*, as reported in the previous papers of this series [26–28]. Thus, Theorem 2 of Part 3 of this series [28] has been introduced to accomplish the enumeration of ligands under the action of a point group. This theorem can be applied to the enumeration of planted promolecules without any modification. Hence, by using the product of SIs collected in Table 1, the cycle index with chirality fittingness (CI-CF) for this case is calculated as follows:

$$\text{CI-CF}(\mathbf{C}_{3v}, \$_d) = \frac{1}{6}(b_1^3 + 2b_3 + 3a_1c_2), \tag{2}$$



3a, I **3b**, C_3 **3c**, C_3^2

3d, $\sigma_{v(1)}$ **3e**, $\sigma_{v(2)}$ **3f**, $\sigma_{v(3)}$

Fig. 6 Planted $\mathbf{C}_s$-promolecules equivalent under the operations of $\mathbf{C}_{3v}$

which counts achiral planted promolecules and enantiomeric pairs of chiral planted promolecules. Theorem 3 of [28] for the enumeration of ligands under the action of the maximum chiral subgroup can be applied to this case so as to derive the following CI-CF:

$$\text{CI-CF}(\mathbf{C}_3, b_d) = \frac{1}{3}(b_1^3 + 2b_3), \tag{3}$$

which counts achiral planted promolecules and chiral planted promolecules, where two enantiomers of each pair are counted separately. The first proposition of Theorem 4 for the enumeration of achiral ligands [28] can be applied to this case so as to derive the following CI-CF$_A$:

$$\begin{aligned}\text{CI-CF}_A(\mathbf{C}_{3v}, \$_d) &= 2\text{CI-CF}(\mathbf{C}_{3v}, \$_d) - \text{CI-CF}(\mathbf{C}_3, b_d)\\ &= \frac{1}{3} \times (3a_1c_2) = a_1c_2,\end{aligned} \tag{4}$$

which counts achiral planted promolecules only. The second proposition of Theorem 4 for the enumeration of chiral ligands [28] can be applied to obtain the following CI-CF$_C$:

$$\begin{aligned}\text{CI-CF}_C(\mathbf{C}_{3v}, \$_d) &= \text{CI-CF}(\mathbf{C}_3, b_d) - \text{CI-CF}(\mathbf{C}_{3v}, \$_d)\\ &= \frac{1}{6}(b_1^3 + 2b_3 - 3a_1c_2),\end{aligned} \tag{5}$$

which counts chiral planted promolecules only.

Suppose that a set of proligands selected from the following proligand warehouse:

$$\mathbf{X} = \{\text{H}, \text{X}, \text{Y}; \text{p}, \overline{\text{p}}; \text{q}, \overline{\text{q}}; \text{r}, \overline{\text{r}}\}, \tag{6}$$

where the symbols H, X, and Y represent achiral proligands, while the pairs, p/$\overline{\text{p}}$, q/$\overline{\text{q}}$, and r/$\overline{\text{r}}$, denote enantiomeric pairs of chiral proligands. Then, the following ligand inventories are calculated:

$$a_d = \mathrm{H}^d + \mathrm{X}^d + \mathrm{Y}^d \tag{7}$$

$$c_d = \mathrm{H}^d + \mathrm{X}^d + \mathrm{Y}^d + 2\mathrm{p}^{d/2}\overline{\mathrm{p}}^{d/2} + 2\mathrm{q}^{d/2}\overline{\mathrm{q}}^{d/2} + 2\mathrm{r}^{d/2}\overline{\mathrm{r}}^{d/2} \tag{8}$$

$$b_d = \mathrm{H}^d + \mathrm{X}^d + \mathrm{Y}^d + \mathrm{p}^d + \overline{\mathrm{p}}^d + \mathrm{q}^d + \overline{\mathrm{q}}^d + \mathrm{r}^d + \overline{\mathrm{r}}^d. \tag{9}$$

These ligand inventories are introduced to Eqs. 2–5 to give generating functions for counting the respective numbers of objects. The generating functions derived from Eqs. 2–5 are essentially equivalent to those described in Part 3 of this series [28], i.e., Eq. 40 of Example 2, Eq. 45 of Example 3, Eq. 52 of Example 4, and Eq. 53 of Example 4.

For further discussions, Fig. 7 collects planted promolecules of various types. Note that an appropriate representative is depicted for every type of planted promolecules. For example, **11** (H^3) is a representative of achiral promolecules having H^3, X^3, Y^3, and so on, even when the warehouse (Eq. 6) is expanded to have other achiral proligands.

It should be noted that a chiral planted promolecule is characterized by a combined term to represent an enantiomeric pair. For example, the enantiomeric pair for **17** is characterized by the term, $\frac{1}{2}(\mathrm{Hp}^2 + \mathrm{H}\overline{\mathrm{p}}^2)$. On the other hand, the achiral planted promolecule **18** (or **19**) is represented by the term $\mathrm{Hp}\overline{\mathrm{p}}$. If we place p = XYZ for a *R*-proligand and $\overline{\mathrm{p}}$= $\overline{\mathrm{XYZ}}$ for an *S*-proligand, the two terms become equal, i.e., $\frac{1}{2}(\mathrm{H(XYZ)}^2 + \mathrm{H}\overline{\mathrm{(XYZ)}}^2) = \mathrm{H(XYZ)}^2$ and $\mathrm{H(XYZ)}\overline{\mathrm{(XYZ)}} = \mathrm{H(XYZ)}^2$, because the achiralities of X, Y, and Z mean that $\overline{\mathrm{X}} = \mathrm{X}$, $\overline{\mathrm{Y}} = \mathrm{Y}$, and $\overline{\mathrm{Z}} = \mathrm{Z}$. The discrimination of these two cases is a key to comprehend the pseudoasymmetry concerning **18** (or **19**).

4 Recursive enumeration of planted 3D-trees

4.1 Principle of determining chirality/achirality

It is worthwhile here to mention how multiple conformational changes are treated in the present formulation of planted 3D-trees as planted promolecules.

Suppose that each of the two proligands (X) in the planted promolecule (**12**) is replaced by an achiral ligand $\mathrm{CH_2Y}$ (e.g., Y = chlorine and H = hydrogen). Then, the resulting planted promolecule has at least two extreme modes of conformation (**34** and **35**), as shown in Fig. 8. The conformer (**34**) as a fixed 3D object belongs to $\mathbf{C}_1$-point group, while the conformer (**35**) as a fixed 3D object belongs to $\mathbf{C}_s$-point group. The former is chiral and the latter is achiral, although they are interchangeable by considering free rotations around an edge (bond).

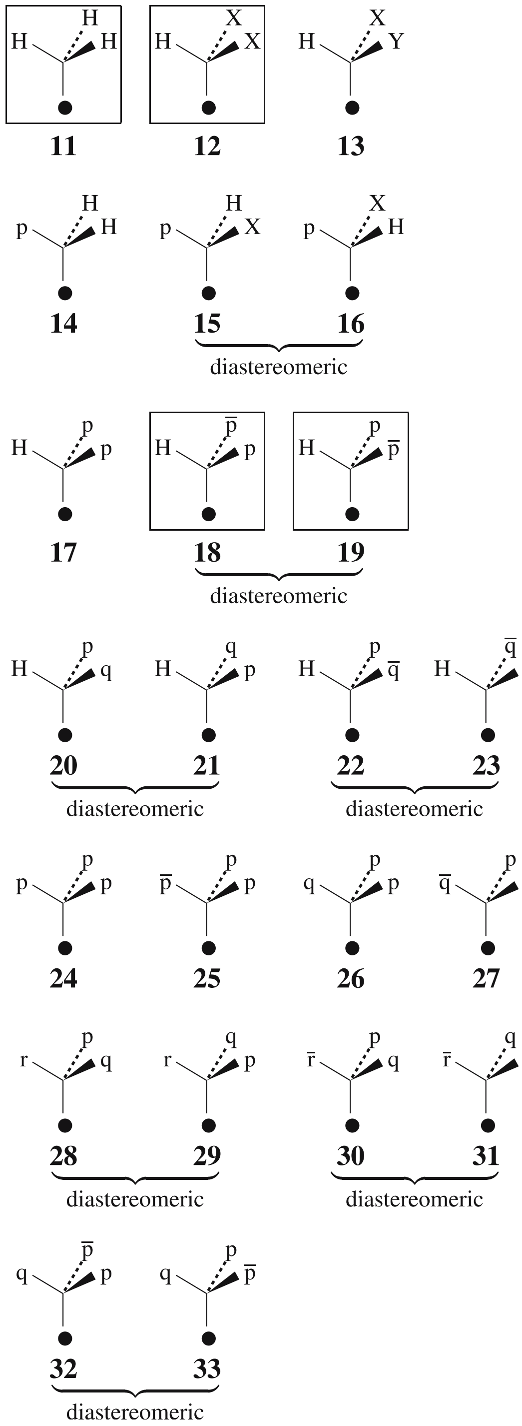
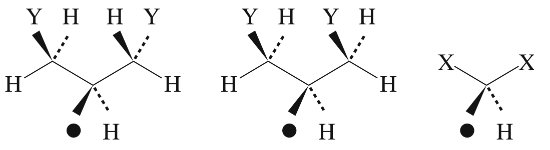


Fig. 7 Various types of planted promolecules based on the $\mathbf{C}_{3v}$-skeleton of a methyl type. The symbols H, X, and Y represent achiral proligands, while the pairs, p / $\overline{\mathrm{p}}$, q / $\overline{\mathrm{q}}$, and r / $\overline{\mathrm{r}}$, denote enantiomeric pairs of chiral proligands. The planted promolecules surrounded by a *box* are achiral. The other planted promolecules are chiral so that an appropriate enantiomer is depicted as a representative for each enantiomeric pair

So long as we take account of such free rotations, the chirality/achirality of every conformer (e.g., **34**, **35**, and others) should be considered. This treatment, however,



34 ($\mathbf{C}_1$) **35** ($\mathbf{C}_s$) **36** (= **12**)

Fig. 8 Fixed conformations (**34** and **35**) and the corresponding planted promolecule (**36**)

would give no fruitful results in combinatorial enumeration, because there are an infinite number of intermediate conformers.

In the present approach, on the other hand, the conformers (**34** and **35**) are replaced by a single planted promolecule **36** (= **12**). As a result, the infinite nature of conformational changes is embedded into proligands of finite nature, where the information on chirality/achirality is preserved. Although this feature is obvious in terms of the definition of a planted promolecule, it is summarized as a theorem:

Theorem 1 (Principle of determining chirality/achirality) *The chirality/achirality of a planted promolecule in isolation is determined if the chirality/achirality of its skeleton and that of its proligands are determined.*

This principle allows us to identify a planted 3D-tree as a single 3D object without explicit consideration of infinite conformational changes. The crux of this principle is that the non-rigidity of a planted 3D-tree is replaced by the rigidity of the corresponding planted promolecule without losing generality so that the chirality/achirality of the planted 3D-tree is determined by examining the planted promolecule.

It should be noted that the planted promolecule described in this principle is regarded as being in isolation. Without being in isolation, achirality might be restricted to exhibit chirality. For example, the two achiral proligands X (i.e CH_2Y ligand) exhibit chirality (i.e., local chirality) when restricted to the promolecule (**36**). When each of the proligands X is regarded as a planted promolecule, nevertheless, the planted promolecule (i.e., $\bullet$–CH_2Y) should be determined to be achiral in isolation.

Chemically speaking, one of the most important conclusions of this principle is that the identity of a compound having several extreme conformers (in fact, infinite intermediate conformers) should be confirmed in the level of promolecules after an appropriate skeleton is selected. Note that the selection of such a skeleton depends upon the purpose of our discussions (e.g., on atropisomers without free rotations).

4.2 Diploids for characterizing enantiosphericity

To develop a general method of enumeration, we shall examine the chirality fittingness of an enantiospheric cycle characterized by the SI (c_d) in more detail. The enantiospheric occupation mode of such an enantiospheric cycle has been obtained as generating functions in Lemma 1 of [28], where the aim of the lemma has been limited within the treatment of the relationship between proligands as substituents in a promolecule. Hence, the modes should be formulated so as to be capable of treating the relationship between promolecules, in particular, the relationship between planted promolecules as models of planted 3D-trees.

In order to treat the relationship between proligands as substituents in a promolecule, the enantiospheric occupation modes of an enantiospheric *orbit* referred to as “compensated chiral packing” have been introduced [16,36]. They can be modified to explain the action of an enantiospheric *cycle* as depicted schematically:

$$\overbrace{\begin{array}{|c|c|c|c|}\hline \mathrm{p} & \mathrm{p} & \cdots & \mathrm{p} \\ \hline \overline{\mathrm{p}} & \overline{\mathrm{p}} & \cdots & \overline{\mathrm{p}} \\ \hline \end{array}}^{d/2} \quad \text{and} \quad \overbrace{\begin{array}{|c|c|c|c|}\hline \overline{\mathrm{p}} & \overline{\mathrm{p}} & \cdots & \overline{\mathrm{p}} \\ \hline \mathrm{p} & \mathrm{p} & \cdots & \mathrm{p} \\ \hline \end{array}}^{d/2}$$

where every proper rotation acts distinctly on the upper row and on the the lower row of each scheme, while every improper rotation is concerned with the exchange between the upper row and the lower row as well as the chirality alternation of each proligand. Each mode of packing is fixed by the action of the enantiospheric d-cycle, because the proligand p (or $\overline{\mathrm{p}}$) is transformed into $\overline{\mathrm{p}}$ (or p) and at the same time the upper row containing $d/2$ of p (or $\overline{\mathrm{p}}$) and the lower row containing $d/2$ of $\overline{\mathrm{p}}$ (or p) are exchanged. These two schemes represent the occupation modes in diastereomeric promolecules. On the other hand, the enantiospheric occupation mode for an achiral proligand X is depicted as follows:

$$\overbrace{\begin{array}{|c|c|c|c|}\hline \mathrm{X} & \mathrm{X} & \cdots & \mathrm{X} \\ \hline \mathrm{X} & \mathrm{X} & \cdots & \mathrm{X} \\ \hline \end{array}}^{d/2}$$

To integrate the modes for chiral proligands and the mode of an achiral proligand, we shall define *a diploid* as a two-membered ordered set containing the enantiomeric objects (enantiomeric proligands in this step of discussion) of a given pair, i.e., $\{\mathrm{p}, \overline{\mathrm{p}}\}$ and $\{\overline{\mathrm{p}}, \mathrm{p}\}$, where every proper rotation transforms each of the diploids into itself, while every improper rotation causes the exchange of the two components of each diploid as well as the change of the chiralities of the components. The term $(\mathrm{p}\overline{\mathrm{p}})^{1/2} (= \sqrt{\mathrm{p}\overline{\mathrm{p}}})$ is assigned to the two diploids.

These diploids are presumed to be different from each other. On the other hand, an achiral proligand (X) generates one diploid, i.e. $\{\mathrm{X},\mathrm{X}\}$, because of $\overline{\mathrm{X}}=\mathrm{X}$.

Because each proper rotation transforms p (or $\overline{\mathrm{p}}$) into itself, each diploid $\{\mathrm{p},\overline{\mathrm{p}}\}$ (or $\{\overline{\mathrm{p}},\mathrm{p}\}$) is transformed into itself by a proper rotation. At the same time, the diploid $\{\mathrm{p},\overline{\mathrm{p}}\}$ (or $\{\overline{\mathrm{p}},\mathrm{p}\}$) is converted into itself even by an improper rotation, because we can presume $\{\overline{\overline{\mathrm{p}},\mathrm{p}}\} = \{\mathrm{p},\overline{\mathrm{p}}\}$ (or $\{\overline{\mathrm{p},\overline{\mathrm{p}}}\} = \{\overline{\mathrm{p}},\mathrm{p}\}$) in terms of chirality alternation after exchange. As a result, the schemes depicted above for p and $\overline{\mathrm{p}}$ are transformed into the following schemes:

$$\overbrace{\begin{array}{|c|c|c|c|}\hline \{\mathrm{p},\overline{\mathrm{p}}\} & \{\mathrm{p},\overline{\mathrm{p}}\} & \cdots & \{\mathrm{p},\overline{\mathrm{p}}\} \\ \hline \{\mathrm{p},\overline{\mathrm{p}}\} & \{\mathrm{p},\overline{\mathrm{p}}\} & \cdots & \{\mathrm{p},\overline{\mathrm{p}}\} \\ \hline \end{array}}^{d/2} \quad \text{and}$$

$$\overbrace{\begin{array}{|c|c|c|c|}\hline \{\overline{\mathrm{p}},\mathrm{p}\} & \{\overline{\mathrm{p}},\mathrm{p}\} & \cdots & \{\overline{\mathrm{p}},\mathrm{p}\} \\ \hline \{\overline{\mathrm{p}},\mathrm{p}\} & \{\overline{\mathrm{p}},\mathrm{p}\} & \cdots & \{\overline{\mathrm{p}},\mathrm{p}\} \\ \hline \end{array}}^{d/2}$$

It should be noted that these schemes for the diploids have the same appearance as the scheme for the diploid of the achiral proligand (X):

$$\overbrace{\begin{array}{|c|c|c|c|}\hline \{\mathrm{X},\mathrm{X}\} & \{\mathrm{X},\mathrm{X}\} & \cdots & \{\mathrm{X},\mathrm{X}\} \\ \hline \{\mathrm{X},\mathrm{X}\} & \{\mathrm{X},\mathrm{X}\} & \cdots & \{\mathrm{X},\mathrm{X}\} \\ \hline \end{array}}^{d/2}$$

After the formulation of diploids (i.e., $\{\mathrm{p},\overline{\mathrm{p}}\}$, $\{\overline{\mathrm{p}},\mathrm{p}\}$, etc.), the evaluation of such a ligand inventory as c_d (Eq. 8) is accomplished by using rather artificial terms such as $(\mathrm{p}\overline{\mathrm{p}})^{1/2}$, $(\mathrm{q}\overline{\mathrm{q}})^{1/2}$, and $(\mathrm{r}\overline{\mathrm{r}})^{1/2}$ for chiral proligands as well as by employing usual terms such as H, X, and Y for achiral proligands. Note that the subscript d for c_d represents an even number so as to give $d/2$ as a positive integer. Because the equivalence between p and $\overline{\mathrm{p}}$ by the action of improper rotations is concealed by the formulation of diploids, we can use the maximum chiral point group (e.g., $\mathbf{C}_3$ in the case of the $\mathbf{C}_{3v}$-point group at issue) to test the congruence of diploids.

This formulation allows us to reread Lemma 1 of [28] as aiming at the enumeration of such diploids, because the original derivation of the lemma [28] turns out to be essentially the same thing as the derivation described in the preceding paragraphs. It follows that the following CI-CF$_D$ is obtained:

$$\text{CI-CF}_D(\mathbf{C}_3, c_d) = \frac{1}{3}(c_2^3 + 2c_6). \tag{10}$$

The concept of diploids is capable of describing a pair of enantiomers for planted promolecules. Let Q be a chiral planted promolecule and $\overline{\mathrm{Q}}$ its enantiomeric planted promolecule. Then, there emerge two diploids, i.e., $\{\mathrm{Q},\overline{\mathrm{Q}}\}$ and $\{\overline{\mathrm{Q}},\mathrm{Q}\}$, which are presumed to be different from each other.

$$\overbrace{\begin{array}{|c|c|c|c|}\hline \{\mathrm{Q},\overline{\mathrm{Q}}\} & \{\mathrm{Q},\overline{\mathrm{Q}}\} & \cdots & \{\mathrm{Q},\overline{\mathrm{Q}}\} \\ \hline \{\mathrm{Q},\overline{\mathrm{Q}}\} & \{\mathrm{Q},\overline{\mathrm{Q}}\} & \cdots & \{\mathrm{Q},\overline{\mathrm{Q}}\} \\ \hline \end{array}}^{d/2} \quad \text{and}$$

$$\overbrace{\begin{array}{|c|c|c|c|}\hline \{\overline{\mathrm{Q}},\mathrm{Q}\} & \{\overline{\mathrm{Q}},\mathrm{Q}\} & \cdots & \{\overline{\mathrm{Q}},\mathrm{Q}\} \\ \hline \{\overline{\mathrm{Q}},\mathrm{Q}\} & \{\overline{\mathrm{Q}},\mathrm{Q}\} & \cdots & \{\overline{\mathrm{Q}},\mathrm{Q}\} \\ \hline \end{array}}^{d/2}$$

Suppose that each proper rotation transforms Q (or $\overline{\mathrm{Q}}$) into itself. Then each diploid $\{\mathrm{Q},\overline{\mathrm{Q}}\}$ (or $\{\overline{\mathrm{Q}},\mathrm{Q}\}$) is transformed into itself by a proper rotation. On the other hand, suppose that each improper rotation causes the exchange of the two components of a diploid as well as the change of the chiralities of the components. Thereby, the diploid $\{\mathrm{Q},\overline{\mathrm{Q}}\}$ is converted into itself even by an improper rotation, because we can presume $\{\overline{\overline{\mathrm{Q}},\mathrm{Q}}\} = \{\mathrm{Q},\overline{\mathrm{Q}}\}$.

On the other hand, an achiral planted promolecule A generates one diploid, i.e. $\{\mathrm{A},\mathrm{A}\}$, because of $\overline{\mathrm{A}}=\mathrm{A}$. The resulting diploid exhibits the following mode of packing:

$$\overbrace{\begin{array}{|c|c|c|c|}\hline \{\mathrm{A},\mathrm{A}\} & \{\mathrm{A},\mathrm{A}\} & \cdots & \{\mathrm{A},\mathrm{A}\} \\ \hline \{\mathrm{A},\mathrm{A}\} & \{\mathrm{A},\mathrm{A}\} & \cdots & \{\mathrm{A},\mathrm{A}\} \\ \hline \end{array}}^{d/2}$$

It follows that Lemma 1 of [28] is applicable to the enumeration of diploids extended to characterize planted promolecules. This means that Eq. 10 is also effective to the planted promolecules.

4.3 Generating functions for enumerating planted 3D-trees

Any planted 3D-tree (or any planted promolecule) can be characterized by the number n of non-terminal nodes so as to be represented by the term x^n. For example, we assign x^{11} to **1** (or **3**) and **2** (or **4**), x^5 to **5** (or **6**), x to **7**, x^3 to **8** and x^2 to **9**.

According to this type of expressions, let $a(x)$ be a generating function for enumerating achiral planted promolecules (cf. Eq. 4); $c(x)$ a generating function for enumerating diploids (ordered enantiomeric pairs, cf. Eq. 10); and $b(x)$ a generating function for enumerating achiral planted promolecules and chiral planted promolecules, where two enantiomers of each pair counted separately (cf. 3). Thus, we place the following generating functions:

$$a(x) = \sum_{n=0}^{\infty} \alpha_n x^n \tag{11}$$

$$c(x) = \sum_{n=0}^{\infty} \gamma_n x^n \tag{12}$$

$$b(x) = \sum_{n=0}^{\infty} \beta_n x^n, \tag{13}$$

where we put $\alpha_0 = 1, \gamma_0 = 1, \beta_0 = 1$ for the trivial cases; and α_n, γ_n, and β_n represent the numbers of planted promolecules at issue, which have the molecular formula x^n. Note that these equations can be evaluated by Eqs. 4, 10, and 3, respectively.

Let $B(x)$ be a generating function for enumerating achiral plus chiral planted promolecules (cf. Eq. 2), where each enantiomeric pair of chiral planted promolecules is counted just once. Let $C(x)$ be a generating function for enumerating chiral planted promolecules (cf. Eq. 5), where each enantiomeric pair of chiral planted promolecules is counted just once. Then, we find the following generating functions:

$$B(x) = \sum_{n=0}^{\infty} B_n x^n \tag{14}$$

$$C(x) = \sum_{n=0}^{\infty} C_n x^n \tag{15}$$

where we put $B_0 = 1$ and $C_0 = 1$ for the trivial cases and B_n and C_n represent the numbers of planted promolecules at issue, which have the molecular formula x^n.

4.4 Ligand inventories as generating functions

By using the grafting procedure, a planted 3D-tree having n non-terminal nodes (chemically carbons) as a parent is divided into three planted 3D-trees as children, which have totally $n-1$ nodes. Every child planted 3D-trees are represented by the term x^n recursively and can be regarded as proligands (or planted promolecules recursively), which are involved in the ligand inventories shown in Eqs. 7–9.

Let us first consider Eq. 9, where the warehouse (Eq. 6) grows in agreement with the grafting procedure. This equation is recognized to be a generating function of counting achiral proligands plus individual chiral proligands (not enantiomeric pairs), each of which is now further regarded as a child planted promolecule in the grafting procedure. Note that such an achiral proligand as a child planted promolecule takes either one of the types collected in Fig. 7. Moreover, the parent planted promolecule takes also either one of the types collected in Fig. 7. This situation can be treated in terms of Theorem 3 of [28] for the enumeration of ligands under the maximum chiral subgroup ($\mathbf{C}_3$) so that the CI-CF shown in Eq. 3 and the generating functions shown in Eq. 13 can be used recursively. Because three children to which the generating function (Eq. 13) is assigned are permuted in accord with the CI-CF (Eq. 3), Theorem 3 of of [28] shows that the series for enumerating achiral proligands plus individual chiral proligands (not enantiomeric pairs) is given by the following expression:

$$\frac{1}{3}\left(b(x)^3 + 2b(x^3)\right), \tag{16}$$

where the subscript d in Eq. 9 is treated as the power of x, i.e., x^d. Because Eq. 16 ignores the principal node of the parent planted promolecule tentatively, the following functional relation is obtained by multiplying by x and adding 1 for the initial (trivial) planted promolecule:

$$b(x) = 1 + \frac{x}{3}\left(b(x)^3 + 2b(x^3)\right). \tag{17}$$

This type of equation was first noted by Pólya [1,2], although the sphericity concept was not taken into consideration.

Next, let us consider Eq. 8, where the warehouse (Eq. 6) again grows in agreement with the grafting procedure. This equation is recognized to be a generating function of counting diploids (achiral proligands plus ordered enantiomeric pairs), each of which is now further regarded as a child planted promolecule. This situation can be treated in terms of Lemma 1 of [28] for counting enantiomeric occupation modes so that the CI-CF$_D$ shown in Eq. 10 and the generating functions shown in Eqs. 12 can be used recursively. Because three children to which the generating function (Eq. 12) is assigned are permuted according to the CI-CF$_D$ shown in Eq. 10, Lemma 1 of of [28] shows that the series for enumerating diploids is given by the following expression:

$$\frac{1}{3}\left(c(x^2)^3 + 2c(x^6)\right), \tag{18}$$

which is derived from Eq. 8 by the subscript d is treated as the power of x, i.e., x^d.

Because Eq. 18 ignores the principal nodes of the parent planted promolecules contained in a diploid, the following functional relation is obtained by multiplying by x^2 and adding 1 for the initial (trivial) planted promolecule:

$$c(x^2) = 1 + \frac{x^2}{3}\left(c(x^2)^3 + 2c(x^6)\right). \tag{19}$$

Finally, let us examine Eq. 7, where the warehouse (Eq. 6) again grows in agreement with the grafting procedure. By an intimate examination, Eq. 7 is recognized to be a generating function of achiral proligands, each of which is now further regarded as a child planted promolecule. This generating function exhibits the same property as the one evaluated by Eq. 4 which aims at counting parent planted promolecules. Thus, there emerges a recursive process.

Note that such an achiral parent planted promolecule takes either type of **11**, **12**, **18**, or **19** and that such a child planted promolecule takes either type selected from Fig. 7 according to chirality fittingness.

This recursive situation can be treated in terms of the first proposition of Theorem 4 of [28] for the enumeration of achiral ligands so that the CI-CF$_A$ shown in Eq. 4 and the generating functions shown in Eqs. 11 and 12 can be used recursively. Because three children to which the generating function (Eq. 11) is assigned are permuted in accord with the CI-CF$_A$ (Eq. 4), Theorem 4 of of [28] shows that the series for enumerating is expressed as follows:

$$a(x)c(x^2). \tag{20}$$

where the subscript d in Eq. 7 is treated as the power of x, i.e., x^d. Because Eq. 20 ignores the principal node of the parent planted promolecule tentatively, the following functional relation is obtained by multiplying by x and adding 1 for the initial (trivial) planted promolecule:

$$a(x) = 1 + xa(x)c(x^2). \tag{21}$$

Although this type of equation was first noted by Robinson et al. [15], their treatment did not take the sphericity concept. As a result, the functions $b(x)$ and $c(x)$ were not differentiated so that both of the functions, i.e., $b(x)$ in Eq. 17 and $c(x)$ in Eq. 21, were replaced by a single function $s(x)$. In contrast, the present formulation indicates that the function $c(x)$ shown in Eq. 21 should correspond to the function $c(x)$ shown in Eq. 19 and not to the function $b(x)$ shown in Eq. 17. This point will be discussed later in detail.

4.5 Planted 3D-trees and chiral planted 3D-trees

Because the generating function $B(x)$ (Eq. 14) is evaluated by using Eq. 2, Theorem 2 of of [28] shows that the series for enumerating is expressed as follows:

$$\frac{1}{6}\left(b(x)^3 + 2b(x^3) + 3a(x)c(x^2)\right) \tag{22}$$

where the subscript d in Eq. 7 is treated as the power of x, i.e., x^d. Because Eq. 22 ignores the principal node of the parent planted promolecule tentatively, the following functional relation is obtained by multiplying by x and adding 1 for the initial (trivial) planted promolecule:

$$B(x) = 1 + \frac{x}{6}\left(b(x)^3 + 2b(x^3) + 3a(x)c(x^2)\right) \tag{23}$$

On the same line, the generating function $C(x)$ (Eq. 15) is expressed as follows:

$$C(x) = 1 + \frac{x}{6}\left(b(x)^3 + 2b(x^3) - 3a(x)c(x^2)\right) \tag{24}$$

by starting from Eq. 5.

By introducing Eq. 17 (for $b(x)$) and Eq. 21 (for $a(x)$) into Eqs. 23 and 24, we obtain more convenient equations as follows:

$$B(x) = \frac{1}{2}\left(a(x) + b(x)\right) \tag{25}$$

$$C(x) = \frac{1}{2}\left(b(x) - a(x)\right), \tag{26}$$

which are represented by linear combinations of $a(x)$ and $b(x)$. It follows that each coefficient in the generating functions for $B(x)$ and $C(x)$ satisfies either of the following relationships.

$$B_n = \frac{1}{2}(\alpha_n + \beta_n) \tag{27}$$

$$C_n = \frac{1}{2}(\beta_n - \alpha_n) \tag{28}$$

These relationships can be obtained intuitively because we can put $B_n = \alpha_n + C_n$ and $\beta_n = \alpha_n + 2C_n$.

4.6 Recursive calculation

By using Eq. 21 for $a(x)$, 17 for $b(x)$, and 19 for $c(x)$, we can evaluate the coefficients α_n, β_n, and γ_n recursively. The results for $n \leq 20$ are summarized in Table 2.

To illustrate the process of recursive calculation, suppose that the coefficients of x^n ($n \leq 8$) have been already calculated. Note that the coefficients of x^{2n} for $c(x)$ are taken into consideration. Then, we presume the following generating functions:

$$\begin{aligned} a(x) &= 1 + x + x^2 + 2x^3 + 3x^4 + 5x^5 + 8x^6 + 14x^7 \\ &\quad + 23x^8 + \alpha_9 x^9 + \alpha_{10}x^{10} + \alpha_{11}x^{11} \end{aligned} \tag{29}$$

$$\begin{aligned} b(x) &= 1 + x + x^2 + 2x^3 + 5x^4 + 11x^5 + 28x^6 + 74x^7 \\ &\quad + 199x^8 + \beta_9 x^9 + \beta_{10}x^{10} + \beta_{11}x^{11} \end{aligned} \tag{30}$$

$$\begin{aligned} c(x^2) &= 1 + x^2 + x^4 + 2x^6 + 5x^8 + 11x^{10} + 28x^{12} + 74x^{14} \\ &\quad + 199x^{16} + \gamma_{18}x^{18} + \gamma_{20}x^{20} + \gamma_{22}x^{22} \end{aligned} \tag{31}$$

where the terms for $n \geq 12$ are omitted tentatively. These intermediate generating functions are introduced into the respective right-hand sides of Eqs. 21, 17, and 19. The resulting equations are expanded to give the following generating functions recursively:

Table 2 The numbers of planted 3D-trees or monosubstituted alkanes

n	α_n	β_n	γ_n	B_n	C_n	R_n
0	1	1	1	1	0	1
1	1	1	–	1	0	1
2	1	1	1	1	0	1
3	2	2	–	2	0	2
4	3	5	1	4	1	4
5	5	11	–	8	3	8
6	8	28	2	18	10	17
7	14	74	–	44	30	39
8	23	199	5	111	88	89
9	41	551	–	296	255	211
10	69	1,553	11	811	742	507
11	122	4,436	–	2,279	2,157	1,238
12	208	12,832	28	6,520	6,312	3,057
13	370	37,496	–	18,933	18,563	7,639
14	636	110,500	74	55,568	54,932	19,241
15	1,134	328,092	–	164,613	163,479	48,865
16	1,963	980,491	199	491,227	489,264	124,906
17	3,505	2,946,889	–	1,475,197	1,471,692	321,198
18	6,099	8,901,891	551	4,453,995	4,447,896	830,219
19	10,908	27,012,286	–	13,511,597	13,500,689	2,156,010
20	19,059	82,300,275	1,553	41,159,667	41,140,608	5,622,109

The numbers of planted 3D-trees are obtained under several conditions, i.e., α_n: achiral planted 3D-trees; β_n: achiral and chiral planted 3D-trees, where two enantiomers of a pair are counted separately; γ_n: diploids; B_n: achiral and chiral planted 3D-trees, where a pair of enantiomers is counted just once; C_n: chiral planted 3D-trees, where a pair of enantiomers is counted just once; and R_n: planted trees as graphs

$$\begin{aligned} a(x) = {} & 1 + x + x^2 + 2x^3 + 3x^4 + 5x^5 + 8x^6 + 14x^7 \\ & +23x^8 + 41x^9 \\ & +(28 + \alpha_9)x^{10} + (53 + \alpha_{10}x^{11}) + \cdots \end{aligned} \quad (32)$$

$$\begin{aligned} b(x) = {} & 1 + x + x^2 + 2x^3 + 5x^4 + 11x^5 + 28x^6 + 74x^7 \\ & +199x^8 + 551x^9 + (1,002 + \beta_9)x^{10} \\ & +(1,781 + 2\beta_9 + \beta_{10})x^{11} + \cdots \end{aligned} \quad (33)$$

$$\begin{aligned} c(x^2) = {} & 1 + x^2 + x^4 + 2x^6 + 5x^8 + 11x^{10} + 28x^{12} + 74x^{14} \\ & + 199x^{16} + 551x^{18} + (1,002 + \gamma_{18})x^{20} \\ & + (1,781 + 2\gamma_{18} + \gamma_{20})x^{22} + \cdots \end{aligned} \quad (34)$$

By comparing the original generating functions (Eqs. 29–31) with the expanded ones (Eqs. 32–34), we find that the coefficients of the terms x^9, x^{10}, and x^{11} satisfy the following relations:

$$\begin{aligned} & \alpha_9 = 41, \alpha_{10} = 28 + \alpha_9, \alpha_{11} = 53 + \alpha_{10} \\ & \beta_9 = 551, \beta_{10} = 1,002 + \beta_9, \beta_{11} = 1,781 + 2\beta_9 + \beta_{10} \\ & \gamma_{18} = 551, \gamma_{20} = 1,002 + \gamma_{18}, \gamma_{22} = 1,781 + 2\beta_{18} + \gamma_{20} \end{aligned} \quad (35)$$

Thereby, the coefficients α_9–α_{11}, β_9–β_{11}, and γ_{18}–γ_{22} are calculated as found in Table 2.

The resulting values are used to derive further intermediate generating functions $a(x)$, $b(x)$, and $c(x)$, which contain the calculated coefficients up to the term x^{11} and unknown coefficients for terms up to x^{14}. Then, the calculation process described above is repeated recursively. The values obtained up to 20 carbon content are listed in Table 2.

After the calculation of α_n and β_n, they are introduced into Eqs. 27 and 28 to give B_n and C_n. These values are also listed in Table 2.

4.7 Diagrammatical tracing

The recursive process of calculation can be traced diagrammatically as shown in Figs. 9, 10, 11 and 12, which depict 41 achiral planted 3D-trees (stereoisomers of mono-substituted alkanes) corresponding to the term x^9 appearing in $a(x)$ (Eq. 32). Although the configuration of each carbon node is not depicted except two pseudoasymmetric cases, these planted 3D-trees shown in Figs. 9, 10, 11 and 12 have 3D structures.

First, Fig. 9 shows 23 planted 3D-trees having CH_2 as a principal node and Z as a root (or primary monosubstituted alkanes in chemical meaning). The occurrence of the 23 stereoisomers is explained by the factorization $x^8 \cdot x^0x^0 \times x$ for the term x^9. The term x^8 of the factorization $x^8 \cdot x^0x^0 \times x$ represents an achiral proligand having 8 non-terminal nodes (carbons) for $a(x)$ in $a(x)c(x^2)$ (or a_1 in a_1c_2). The term x^0x^0 of the factorization $x^8 \cdot x^0x^0 \times x$ represents two terminal nodes (hydrogens) for $c(x^2)$ in $a(x)c(x^2)$ (or c_2 in a_1c_2). The last term x corresponds to an additional node as a new principal node. It follows that the coefficient 23 of the term x^8 in Eq. 32 and the coefficient 1 of the term x^0x^0 (= 1) in Eq. 34 are multiplied to give 23. Thus, there emerge 23 primary planted 3D-trees or primary monosubstituted alkanes (**37**–**59**), as depicted in Fig. 9. As found easily, the 23 planted 3D-trees correspond to the achiral planted promolecule of the type **12** (Fig. 7), in which H is replaced by an achiral R (alkyl) and each X is replaced by H (hydrogen), i.e., H_2RC– in chemical meaning. Chemically speaking, the primary monosubstituted alkanes **37**–**59** are represented by the formula R–CH_2–Z, where R contains eight carbons.

The primary planted 3D-trees or primary monosubstituted alkanes (**37**–**59**) collected in Fig. 9 are divided into subcategories by considering further factorization of the term x^8 appearing in the factorized term $x^8 \cdot x^0x^0 \times x$. It should be note that the itemization with respect to term x^8 corresponds to the itemization of 23 achiral planted 3D-trees or monosubstituted alkanes with x^8 (cf. the x^8-row ($n = 8$) of Table 2).

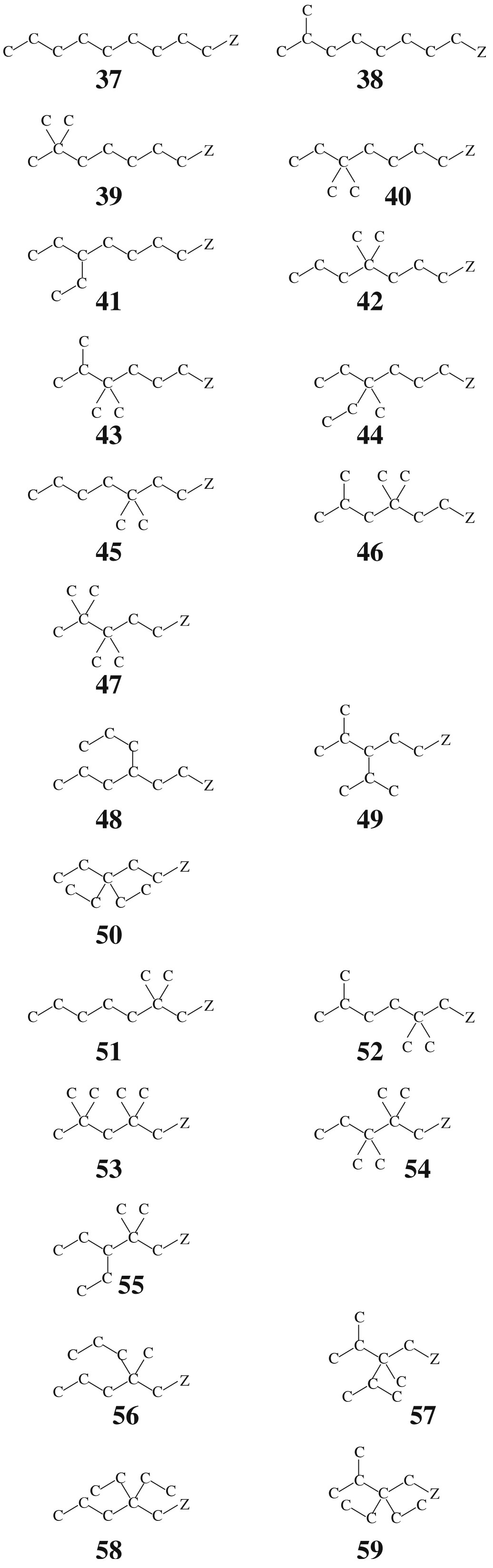


Fig. 9 Achiral planted 3D-trees or monosubstituted alkanes, which are assigned to the term x^9 (23 primary monosubstituted primary alkanes among 41 stereoisomers)

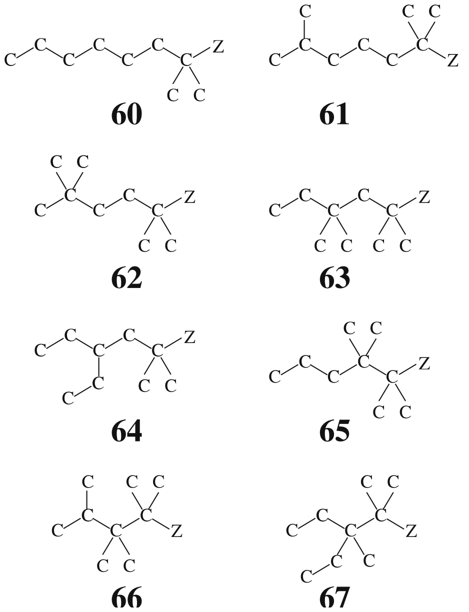


Fig. 10 Achiral planted 3D-trees or monosubstituted alkanes, which are assigned to the term x^9 (8 tertiary monosubstituted alkanes among 41 stereoisomers)

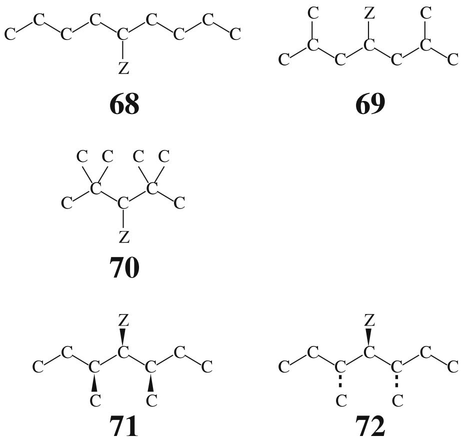


Fig. 11 Achiral planted 3D-trees or monosubstituted alkanes, which are assigned to the term x^9 (5 secondary monosubstituted alkanes among 41 stereoisomers)

1. According to the further factorization $x^7 \cdot x^0x^0 \times x$ of the term x^8, the coefficient 14 of the term x^7 in Eq. 32 and the coefficient 1 of the term x^0x^0 (=1) in Eq. 34 are multiplied to give 14. Thus, there emerge 14 primary planted 3D-trees or primary monosubstituted alkanes of this type, i.e., **37–50**. The factorization $x^7 \cdot x^0x^0 \times x$ again corresponds to the achiral planted promolecule of the type **12** (Fig. 7), in which H is replaced by an achiral R (alkyl) and each X is replaced by H (hydrogen), i.e., H_2RC– in chemical meaning. It follows that the primary monosubstitut-

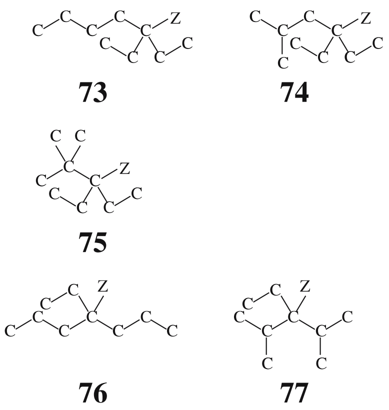


Fig. 12 Achiral planted 3D-trees or monosubstituted alkanes, which are assigned to the term x^9 (5 tertiary monosubstituted alkanes among 41 stereoisomers)

ed alkanes **37–50** are represented by the formula R–CH_2CH_2–Z, where R contains seven carbons.

2. According to another factorization $x^5 \cdot x^1x^1 \times x$ of the term x^8, the coefficient 5 of the term x^5 in Eq. 32 and the coefficient 1 of the term x^1x^1 $(= x^2)$ in Eq. 34 are multiplied to give 5. Thus, there emerge five primary planted 3D-trees or primary monosubstituted alkanes of this type, i.e., **51–55**. The factorization $x^5 \cdot x^1x^1 \times x$ corresponds to the achiral planted promolecule of the type **12** (Fig. 7), in which H is replaced by an achiral R (alkyl) and each X is replaced by Me (methyl), i.e., RMe_2C– in chemical meaning. It follows that the primary monosubstituted alkanes **51–55** are represented by the formula RMe_2C–CH_2–Z, where R contains five carbons.
3. According to further factorization $x \cdot x^3x^3 \times x$ of the term x^8, the coefficient 1 of the term x in Eq. 32 and the coefficient 2 of the term x^3x^3 $(= x^6)$ in Eq. 34 are multiplied to give 2. Thus, there emerge two primary planted 3D-trees or primary monosubstituted alkanes of this type, i.e., **56** and **57**. The factorization $x \cdot x^3x^3 \times x$ corresponds to the achiral planted promolecule of the type **12** (Fig. 7), in which H is replaced by Me (methyl) and each X is replaced by R (propyl), i.e., MeR_2C– in chemical meaning. It follows that the primary monosubstituted alkanes **56** and **57** are represented by the formula MeR_2C–CH_2–Z, where R contains three carbons.
4. According to the factorization $x^3 \cdot x^2x^2 \times x$ of the term x^8, the coefficient 2 of the term x^3 in Eq. 32 and the coefficient 1 of the term x^2x^2 $(= x^4)$ in Eq. 34 are multiplied to give 2. Thus, there emerge two primary planted 3D-trees or primary monosubstituted alkanes of this type, i.e., **58** and **59**. The factorization $x^3 \cdot x^2x^2 \times x$ corresponds to the achiral planted promolecule of the type **12** (Fig. 7), in which H is replaced by R (propyl) and each X is replaced by Et (ethyl), i.e., REt_2C– in chemical meaning. It follows that the primary monosubstituted alkanes **58** and **59** are represented by the formula REt_2C–CH_2–Z, where R contains three carbons.

On the other hand, Fig. 10 can be explained as follows. Thus, the term x^9 coming from the factorization $x^6 \cdot x^1x^1 \times x$ corresponds to eight planted 3D-trees having a tertiary principal node and Z as a root (i.e., tertiary monosubstituted alkanes in chemical meaning). The term x^6 of the factorization $x^6 \cdot x^1x^1 \times x$ represents an achiral proligand with six non-terminal nodes (carbons), which meets $a(x)$ in $a(x)c(x^2)$ (or a_1 in a_1c_2). The term x^1x^1 of the factorization $x^6 \cdot x^1x^1 \times x$ represents two proligands each having one node, which meet $c(x^2)$ in $a(x)c(x^2)$ (or c_2 in a_1c_2). The last term x corresponds to an additional node as a new principal node. It follows that the coefficient 8 of the term x^6 in Eq. 32 and the coefficient 1 of the term x^1x^1 $(= x^2)$ in Eq. 34 are multiplied to give 8. Thus, there emerge eight tertiary planted 3D-trees or tertiary monosubstituted alkanes (**60–67**), as depicted in Fig. 10.

The factorization $x^6 \cdot x^1x^1 \times x$ corresponds to the achiral planted promolecule of the type **12** (Fig. 7), in which H is replaced by an achiral R (alkyl) and each X is replaced by Me (methyl), i.e., RMe_2C– in chemical meaning. It follows that the tertiary monosubstituted alkanes **60–67** shown in Fig. 10 are represented by the formula RMe_2C–Z, where R contains six carbons.

There are five planted 3D-trees having a *secondary* principal node and Z as a root (i.e., secondary monosubstituted alkanes in chemical meaning), which are shown in Fig. 11, where the term *secondary* is used chemically. They can be explained by the factorized term $x^0 \cdot x^4x^4 \times x$ for the term x^9. The term x^0 of the factorization $x^0 \cdot x^4x^4 \times x$ represents no carbon node (but a terminal hydrogen node) for $a(x)$ in $a(x)c(x^2)$ (or a_1 in a_1c_2). The term x^4x^4 of the factorization represents two proligands having four carbon nodes, which meet $c(x^2)$ in $a(x)c(x^2)$ (or c_2 in a_1c_2). The last term x corresponds to an additional node as a new principal node. It follows that the coefficient 1 of the term $x^0 (= 1)$ in Eq. 32 and the coefficient 5 of the term x^4x^4 $(=x^8)$ in Eq. 34 are multiplied to give 5. Thus, there emerge five secondary planted 3D-trees or secondary monosubstituted alkanes (**68–72**), as depicted in Fig. 11.

The factorization $x^0 \cdot x^4x^4 \times x$ corresponds to the achiral planted promolecule of the type **12**, **18**, or **19** (Fig. 7). For the correspondence to **12**, H is replaced by a hydrogen atom and each X is replaced by R (butyl), i.e., R_2HC– in chemical meaning. It follows that the secondary monosubstituted alkanes **68–70** shown in Fig. 11

are represented by the formula R_2CH–Z, where each R contains four carbons.

For the correspondence to **18** (or **19**), H is replaced by a hydrogen atom, p (or $\overline{\mathrm{p}}$) is replaced by $R^{[R]}$ (chiral *sec*-butyl of *R*-configuration), and $\overline{\mathrm{p}}$ (or p) is replaced by $R^{[S]}$ (chiral *sec*-butyl of *S*-configuration), i.e., $R^{[R]}R^{[S]}HC$– in chemical meaning. It follows that the secondary monosubstituted alkanes **71** and **72** shown in Fig. 11 are represented by the formula $R^{[R]}R^{[S]}CH$–Z, which indicates the first appearance of pseudoasymmetry in the series of planted 3D-trees or monosubstituted alkanes.

The tertiary planted 3D-trees (**73**–**75**) depicted in Fig. 12 are characterized by the factorized term $x^4 \cdot x^2x^2 \times x$. The term x^4 of the factorization $x^4 \cdot x^2x^2 \times x$ represents 4 non-terminal nodes (carbons) for $a(x)$ in $a(x)c(x^2)$ (or a_1 in a_1c_2). The term x^2x^2 represents two plus two nodes for $c(x^2)$ in $a(x)c(x^2)$ (or c_2 in a_1c_2). The last term x corresponds to an additional node as a new principal node. It follows that the coefficient 3 of the term x^4 in Eq. 32 and the coefficient 1 of the term x^2x^2 $(= x^4)$ in Eq. 34 are multiplied to give 3. Thus, there emerge three tertiary planted 3D-trees or tertiary monosubstituted alkanes (**73**–**75**).

The factorization $x^4 \cdot x^2x^2 \times x$ of the term x^9 corresponds to the achiral planted promolecule of the type **12** (Fig. 7), in which H is replaced by an achiral R (butyl) and each X is replaced by Et (ethyl), i.e., REt_2C– in chemical meaning. It follows that the tertiary monosubstituted alkanes **73**–**75** shown in Fig. 12 are represented by the formula REt_2C–Z, where R contains four carbons.

The set of tertiary planted 3D-trees (**76** and **77**) depicted in Fig. 12 is characterized by the factorized term $x^2 \cdot x^3x^3 \times x$. The term x^2 of the factorization $x^2 \cdot x^3x^3 \times x$ represents two non-terminal nodes (carbons) for $a(x)$ in $a(x)c(x^2)$ (or a_1 in a_1c_2). The term x^3x^3 represents three plus three nodes for $c(x^2)$ in $a(x)c(x^2)$ (or c_2 in a_1c_2). The last term x corresponds to an additional node as a new principal node. It follows that the coefficient 1 of the term x^2 in Eq. 32 and the coefficient 2 of the term x^3x^3 $(= x^6)$ in Eq. 34 are multiplied to give 2. Thus, there emerge two tertiary planted 3D-trees or tertiary monosubstituted alkanes (**76** and **77**).

The factorization $x^2 \cdot x^3x^3 \times x$ of the term x^9 corresponds to the achiral planted promolecule of the type **12** (Fig. 7), in which H is replaced by an achiral R (ethyl) and each X is replaced by R (propyl), i.e., EtR_2C– in chemical meaning. It follows that the tertiary monosubstituted alkanes **76** and **77** shown in Fig. 12 are represented by the formula EtR_2C–Z, where R contains three carbons.

In summary, the diagrammatical tracing reveals the recursive nature of calculation. Thus, the number 41 is obtained by summing up the numbers of the planted 3D-trees shown in Figs. 9, 10, 11 and 12 as follows:

$$23 + 8 + 5 + (3 + 2) = 41, \tag{36}$$

where the respective numbers in the left-hand side appear as the coefficients of the recursive calculations (Eqs. 32 and 34), i.e., $23x^8$ in $a(x)$ (Eq. 32), $8x^6$ in $a(x)$ (Eq. 32), $5x^8$ in $c(x)$ (Eq. 34), $3x^4$ in $a(x)$ (Eq. 32), and $2x^6$ in $c(x)$ (Eq. 34). The total number 41 appears recursively as the coefficient of the term x^9 in $a(x)$ (Eq. 32).

Moreover, the number 23 of the planted 3D-trees shown in Fig. 9 is obtained as follows:

$$14 + 5 + 2 + 2 = 23, \tag{37}$$

where the respective numbers in the left-hand side appear as the coefficients of the recursive calculations (Eqs. 32 and 34), i.e., $14x^7$ in $a(x)$ (Eq. 32), $5x^5$ in $a(x)$ (Eq. 32), $2x^6$ in $c(x)$ (Eq. 34), an $2x^3$ in $a(x)$ (Eq. 32). The total number 23 appears recursively as the coefficient of the term x^8 in $a(x)$ (Eq. 32).

5 Graphs versus chemical 3D-structures

5.1 Comparison with Pólya's Theorem

Before we start our discussions on graphs versus chemical 3D-structures, we shall refer to several pioneering accomplishments on the enumeration of planted trees or their equivalents. Henze and Blair [37] obtained the number of alkyl radicals (or aliphatic alcohols) of a given carbon content by an elaborate method other than Polya's method, where the alkyl radicals were regarded as graphs, not as 3D-objects. Pólya [1,2] applied his main theorem (Hauptsatz) to the evaluation of the number of planted trees, which were mathematical counterparts of alkyl radicals. Pólya's results [1,2] were also limited within graphs and were not concerned with 3D-objects from the present viewpoint.

According to Pólya (Eq. 2.13 of Sect. 42 [1,2]), the number (R_n) of planted trees or alkyl radicals as graphs is represented by the following generating function:

$$s(x) = \sum_{n=0}^{\infty} R_n x^n, \tag{38}$$

where we put $R_0 = 1$ and obtain the following functional equation:

$$s(x) = 1 + \frac{x}{6}\left(s(x)^3 + 2s(x^3) + 3s(x)s(x^2)\right). \tag{39}$$

Obviously, Eq. 39 is a special case of Eq. 23 in which we place $s(x) = a(x) = b(x) = c(x)$ in the right-hand side and $s(x) = B(x)$ in the left-hand side. The results based

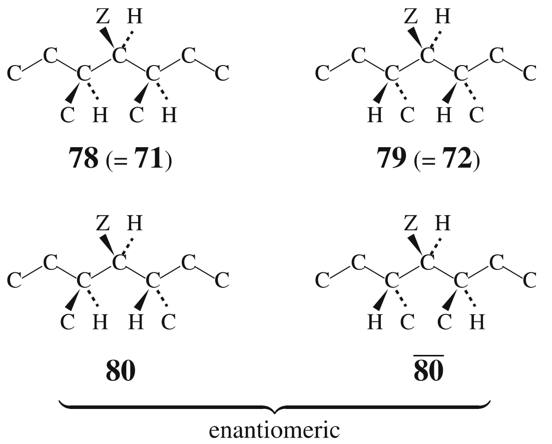


Fig. 13 Pseudoasymmetry versus enantiomeric relationship in planted 3D-trees or monosubstituted alkanes, which are assigned to the term x^9

on Eq. 39 are shown in the last column of Table 2. They are identical with those of Henze and Blair [37].

To discuss pseudoasymmetric cases appearing in the case of $n = 9$, a manual enumeration is done by considering itemization with respect to the number of stereogenic configurations:

Planted 3D-trees with x^9	Number
Achiral (except pseudoasymmetric cases)	39 (as graphs)
Chiral (with one stereogenic configuration)	102 (as graphs)
Chiral (with two stereogenic configurations)	62 (as graphs)
Chiral (with three stereogenic configurations)	7 (as graphs)
pseudoasymmetric case	1 (as graphs)
Total number	211 (as graphs)

By starting from these values, the row of $n = 9$ in Table 2 is confirmed as follows:

$$\alpha_9 = 39 + 2 = 41 \tag{40}$$

$$\beta_9 = (39 + 2) + (102 \times 2 + 62 \times 2^2 + 7 \times 2^3 + 2) = 551 \tag{41}$$

$$B_9 = (39 + 2) + (102 + 62 \times 2 + 7 \times 2^2 + 1) = 296 \tag{42}$$

$$C_9 = 102 + 62 \times 2 + 7 \times 2^2 + 1 = 255 \tag{43}$$

$$R_9 = 39 + 102 + 62 + 7 + 1 = 211, \tag{44}$$

where the first and the second pair of parentheses in the calculation of β_9 or B_9 are concerned with achiral and chiral stereoisomers, respectively.

It is informative to examine Fig. 13, which depicts the pseudoasymmetric cases (**78** and **79**) and a relevant enantiomeric pair of chiral stereoisomers (**80** and $\overline{\textbf{80}}$). They contribute variously to the recursive calculations shown in Table 2, where the effect of pseudoasymmetry can be clearly demonstrated as follows.

All of the stereoisomers (**78**, **79**, **80**, $\overline{\textbf{80}}$) are equalized in the recursive enumeration based on Eq. 39 so as to be counted as one graph (cf. R_9). They are categorized into an achiral stereoisomer **78**, an achiral stereoisomer **79**, and an enantiomeric pair of chiral stereoisomers (**80** and $\overline{\textbf{80}}$). Hence, the contribution of these stereoisomers to each coefficient is summarized as follows:

α_9: 2 (achiral)	**78**, **79**
β_4: 2 (achiral) + 2 (chiral)	**78**, **79** + **80**, $\overline{\textbf{80}}$
B_9: 2 (achiral) + 1 (enantiomeric pair)	**78**, **79** + **80**/$\overline{\textbf{80}}$
C_9: 1 (enantiomeric pair)	**80**/$\overline{\textbf{80}}$
R_9: 1 (graph)	**78**/**79**/**80**/$\overline{\textbf{80}}$

where each symbol with slashes (**80**/$\overline{\textbf{80}}$ or **78**/**79**/**80**/$\overline{\textbf{80}}$) indicates that the stereoisomers at issue are regarded as a single isomer so as to be counted just once under a criterion for calculating each coefficient. In place of this manual itemization, a more general combinatorial itemization is desirable but open to further investigation.

5.2 Merits of Fujita's proligand method

To show the merits of Fujita's proligand method, it is worthwhile to mention the previous results of Robinson et al. [15], who investigated the enumeration of planted trees or monosubstituted alkanes after modifying the terms in Pólya's CIs. They gave the number of planted trees having up to 14 carbon content, which are identical with our results for α_n and β_n ($n \leq 14$) collected in Table 2. Their modification of Pólya's CIs, however, did not contain the concepts of sphericity, proligand, promolecule, which are keys of the present approach based on Fujita's proligand method. This point is essential and shall be discussed in detail.

The concept of sphericity [16] indicates that an enantiospheric cycle (c_2) causes a transitive permutation between p and $\overline{\mathrm{p}}$, where the action of c_2 brings about the exchange between p and $\overline{\mathrm{p}}$ with chirality alternation. The transitive permutation corresponds to the property of an enantiospheric orbit called the compensated chiral packing of a chiral ligand p and its enantiomer $\overline{\mathrm{p}}$.

Although Robinson et al. [15] seemed to take such compensated chiral packing into implicit consideration, they considered the action of b_2 in place of the action of c_2. Thus, in place of Eq. 23 of the present approach, they derived their Eq. 13 [15], which can be written as

$$B(x) = 1 + \frac{x}{6}\Big(b(x)^3 + 2b(x^3) + 3a(x)b(x^2)\Big), \tag{45}$$

where their notation is changed to meet the present one. By combining this with their counterpart of Eq. 25, they derived their counterpart of the following equation:

$$a(x) = 1 + xa(x)b(x^2). \tag{46}$$

They applied their counterpart of Eq. 46 to recursive calculations in place of Eq. 21 of the present approach.

Note that the action of b_2 (i.e., $b(x^2)$) causes the exchange between p and p (or between $\overline{\mathrm{p}}$ and $\overline{\mathrm{p}}$) without chirality alternation and that b_2 does not exchange p and $\overline{\mathrm{p}}$ because of transitivity. If the action of b_2 is considered to exchange p and $\overline{\mathrm{p}}$, two modes of packing (i.e., p/$\overline{\mathrm{p}}$ and $\overline{\mathrm{p}}$/p) cannot be differentiated from each other so as to become degenerate as a single mode of packing. It follows that their treatment using Eq. 46 unconsciously overlooked such subtle but essential situations as concerned with b_2 and c_2, so that the mixing-up of b_2 with c_2 took place, contrary to their own intention. In fact, the two modes of packing (i.e., p/$\overline{\mathrm{p}}$ and $\overline{\mathrm{p}}$/p) according to the compensated chiral packing were replaced by other two modes of packing, i.e., p/p and $\overline{\mathrm{p}}$/$\overline{\mathrm{p}}$, in their treatment. Because the molecular formulas of p and $\overline{\mathrm{p}}$ are equal by considering their carbon contents only, the usage of Eq. 46 fortunately resulted in the same results as the present ones using Eq. 21.

From the viewpoint which is brought about by the sphericity concept, we can say that Robinson et al. [15] regarded the ligand inventory for c_d (e.g., Eq. 8) as being equal to the one for b_d (e.g., Eq. 9). In other words, their modified method is effective only on condition that the ligand inventory c_d can be equalized to b_d. This condition is satisfied in the recursive enumeration described in Sect. 4. However, their modified method is incapable of solving enumeration of promolecules with considering p and $\overline{\mathrm{p}}$, because the action of c_d is inevitable to be considered, as found easily by the discussions described in Sect. 3.3. It should be emphasized that Fujita's proligand method gives a common mathematical framework to solve both kinds of enumeration problems.

To evaluate c_d ($c(x)$) as a generating function, we have introduced the concept of *diploid* in Sect. 4.2. The concept provides us with a deep insight into the relationship between the SI c_d for enantiospheric cycles and the SI b_d for hemispheric cycles.

A ligand inventory for an SI c_d (e.g., Eq. 8) can be regarded as a generating function for counting such diploids of promolecules (or proligands) separately (e.g., $\{\mathrm{A}, \mathrm{A}\}$, $\{\mathrm{Q}, \overline{\mathrm{Q}}\}$, $\{\overline{\mathrm{Q}}, \mathrm{Q}\}$, and others). On the other hand, a ligand inventory for an SI b_d (e.g., Eq. 9) is regarded as a generating function for counting achiral, chiral promolecules (or proligands) and their enantiomers separately (e.g., an achiral promolecule A, a chiral promolecule Q, its enantiomer $\overline{\mathrm{Q}}$, and others). As found easily by the definition, there appears one-to-one correspondence for each of the following pairs: $\mathrm{A} \leftrightarrow \{\mathrm{A}, \mathrm{A}\}$, $\mathrm{Q} \leftrightarrow \{\mathrm{Q}, \overline{\mathrm{Q}}\}$, $\overline{\mathrm{Q}} \leftrightarrow \{\overline{\mathrm{Q}}, \mathrm{Q}\}$ under the maximum chiral subgroup (e.g., the subgroup $\mathbf{C}_3$ of $\mathbf{C}_{3v}$). This correspondence allows us to evaluate $c(x^2)$ (due to c_2) by using Eq. 19.

6 Conclusions

Planted three-dimensional (3D) trees, which are defined as a 3D extension of planted trees, are enumerated by Fujita's proligand method formulated in Part 1–3 of this series [26–28]. A planted 3D-tree is regarded as a planted promolecule, which is defined as a 3D object, where the substitution positions of a given 3D skeleton are occupied by a root and proligands. Each of the proligands is further regarded as another planted promolecule in a nested fashion. Thus, the nested character of intermediate planted promolecules is used to derive generating functions for enumerate planted 3D-trees. The generating functions are based on 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. For the purpose of evaluating c_d, the concept of *diploid* is proposed, where the nested nature of c_d is demonstrated clearly. The SIs are applied to derive functional equations for recursive calculations, i.e., $a(x)$, $c(x^2)$, and $b(x)$. Thereby, recursive calculations of planted 3D-trees or equivalently those of monosubstituted alkanes as stereoisomers are conducted and collected up to 20 carbon content in a tabular form.

References

1. Pólya G (1937) Acta Math. 68:145–254
2. Pólya G, Read RC (1987) Combinatorial enumeration of groups, graphs, and chemical compounds. Springer, Berlin Heidelberg New York
3. Hosoya H (1972) Kagaku no Ryoiki 26:989–1001
4. Rouvray DH (1974) Chem Soc Rev 3:355–372
5. Polansky OE (1975) MATCH Commun Math Comput Chem 1:11–31
6. Balaban AT (ed) (1976) Chemical applications of graph theory. Academic, London
7. Pólya G, Tarjan RE, Woods DR (1983) Notes on introductory combinatorics. Birkhäuser, Boston
8. Balasubramanian K (1985) Chem Rev 85:599–618
9. Cayley A (1874) Philos Mag 47(4):444–446
10. Cayley A (1875) Rep Brit Assoc Adv Sci 45:257–305
11. van't Hoff JH (1874) Archives Néerlandaises des Sciences exactes et naturelles 9:445–454
12. van't Hoff JH (1875) La Chimie Dans L'Espace. P. M. Bazendijk, Rotterdam
13. Fisher E (1891) Ber Dtsch Chem Ges 24:1836
14. Fisher E (1891) Ber Dtsch Chem Ges 24:2683
15. Robinson RW, Harary F, Balaban AT (1976) Tetrahedron 32:355–361
16. Fujita S (1991) Symmetry and combinatorial enumeration in chemistry. Springer, Berlin Heidelberg New York
17. Fujita S (1989) Theor Chim Acta 76:247–268
18. Fujita S (1990) J Math Chem 5:121–156
19. Fujita S (1990) Bull Chem Soc Jpn 63:203–215
20. Fujita S (2005) MATCH Commun Math Comput Chem 54:251–300

21. Fujita S (2006) MATCH Commun Math Comput Chem 55:5–38
22. Fujita S (2006) MATCH Commun Math Comput Chem 55:237–270
23. Fujita S (1991) Tetrahedron 47:31–46
24. Fujita S (1992) J Chem Inf Comput Sci. 32:354–363
25. Fujita S (1993) Polyhedron 12:95–110
26. Fujita S (2005) Theor Chem Acc 113:73–79
27. Fujita S (2005) Theor Chem Acc 113:80–86
28. Fujita S (2006) Theor Chem Acc 115:37–53
29. Fujita S (1995) Theor Chim Acta 91:291–314
30. Fujita S (1995) Theor Chim Acta 91:315–332
31. Fujita S (1998) Bull Chem Soc Jpn 71:1587–1596
32. Fujita S (1999) Bull Chem Soc Jpn 72:2409–2416
33. Biggs NL, Lloyd EK, Wilson RJ (1976) Graph theory 1736–1936. Oxford University Press, Oxford
34. Harary F (1969) Graph theory. Addison-Wesley, Reading
35. Foulds LR (1992) Graph theory applications. Springer, Berlin Heidelberg New York
36. Fujita S (1990) J Am Chem Soc 112:3390–3397
37. Henze HR, Blair CM (1931) J Am Chem Soc 53:3042–3046