Journal of Mathematical Chemistry Vol. 34, Nos. 3–4, November 2003 (© 2003)

On certain organic compounds with one mono-substitution and at least three di-substitution homogeneous derivatives

Valentin Vankov Iliev

Section of Algebra, Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria

E-mail: viliev@math.bas.bg, viliev@aubg.bg

Received 24 February 2003

In this paper parent substances with molecules which can be divided into a skeleton and six univalent substituents, and that have the properties mentioned in the title, are considered. Two instances are the molecules of benzene and cyclopropane. The Lunn–Senior's groups of substitution isomerism of these compounds are described and upper bounds of the numbers of their di-substitution and tri-substitution homogeneous derivatives are found. Lists of the possible simple substitution reactions among di-substitution homogeneous derivatives, on one hand, and di-substitution heterogeneous, and tri-substitution homogeneous derivatives, on the other, are given. These substitution reactions allow for some derivatives to be identified with their structural formulae.

KEY WORDS: substitution isomers, Lunn–Senior's permutation group of substitutional isomerism, simple substitution reactions, di-substituted and tri-substituted derivatives

AMS subject classification: 05E25, 20B35, 92E10

1. Introduction

A starting point of Lunn–Senior's theory of assigning a permutation group of symmetry of degree d to a given molecular structure divided into skeleton and d univalent substituents is the following old observation: the number of its substitution isomers does not depend on the nature of the ligants but only on the numbers λ_k of members of their different types x_k, $k = 1, 2, \ldots,$ and on the skeleton. The only natural restriction is that if the skeleton contains a univalent atom (or radical), then no univalent substituent is to be identical with this atom (radical). As far as the order of ligants is irrelevant, we obtain a *partition* $(\lambda_1, \lambda_2, \ldots, \lambda_d)$ of the number d, that is, $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_d \geqslant 0$, and $\lambda_1 + \lambda_2 + \cdots + \lambda_d = d$. Plainly, the monomial

$$x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_d^{\lambda_d}$$

is an exotic representation of substituents' *empirical formula* of the molecular structure under question. If Θ is the empirical formula of the skeleton, then

$$\Theta x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_d^{\lambda_d}$$

is the empirical formula of the molecule. The additional information that makes difference between its empirical and structural formulae consists of a set of lists A_k, $k = 1, 2, \ldots, d$, each one enumerating the unsatisfied valencies of the skeleton occupied by the identical ligants of type x_k. If a numeration $1, 2, \ldots, d$ of the unsatisfied valencies is fixed once and for all, then A_k are simply pairwise disjoint subsets of the integer-valued interval $[1, d]$, such that $[1, d] = \bigcup_k A_k$. Thus, the mathematical model of a *structural formula* of the substituents of a molecular structure with empirical formula $x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_d^{\lambda_d}$, is a *tabloid* $A = (A_1, A_2, \ldots, A_d)$ *with* d *nodes of shape* $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_d)$:

$$
A = \begin{array}{lllll}
a_{1,1}, & a_{1,2}, & \ldots & \ldots & a_{1,\lambda_1} \\
a_{2,1}, & a_{2,2}, & \ldots & a_{2,\lambda_2} & \\
\vdots & & & & \\
a_{t,1}, & a_{t,2}, & \ldots & a_{t,\lambda_t} &
\end{array}
\qquad
\begin{array}{l}
\text{the component } A_1 \\
\text{the component } A_2 \\
\vdots \\
\text{the component } A_t
\end{array}
$$

$$\downarrow \quad \varphi$$

$$
\lambda = \begin{array}{lllll}
\times & \times & \ldots & \ldots & \times \\
\times & \times & \ldots & \times & \\
\vdots & & & & \\
\times & \ldots & \ldots & \times &
\end{array}
\qquad
\begin{array}{l}
\lambda_1 \text{ nodes} \\
\lambda_2 \text{ nodes} \\
\vdots \\
\lambda_t \text{ nodes}
\end{array}
$$

Here $\varphi : T_d \to P_d$ is the natural projection of the set T_d of all tabloids with d nodes onto the set P_d of all partitions of d, that maps the tabloid A onto its *shape* λ: $\lambda_1 = |A_1|$, $\lambda_2 = |A_2|, \ldots, \lambda_d = |A_d|$.

The structural formula of a molecule encodes its "connexity data", and does not reflect in full so called "space configuration", because the latter is a special representation of the former. "Connexity" is a relation of order independent of considerations of space. The "structural" relations treated by chemists are relations of just this sort, and it is unfortunate that the word structure as used by engineers, etc., should carry with it geometrical connotations which are too special for chemistry" [1, p. 1030].

The inverse image $T_\lambda = \varphi^{-1}(\lambda)$ consists of all structural formulae of the substituents with empirical formula $x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_d^{\lambda_d}$. The fibers T_λ, $\lambda \in P_d$, of the map φ are the stages where the drama of isomerism is performed.

In [1], Lunn and Senior build in the phenomenon of isomerism of a certain type in the above mathematical model by means of action of a symmetry group G, consisting of permutations of the d unsatisfied valencies of the skeleton, and such that any isomer of the given empirical formula $\Theta x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_d^{\lambda_d}$ is represented by a G-orbit in T_λ. The group G acts on the set T_d of structural formulae by the rule

$$\sigma(A_1, A_2, \ldots, A_d) = \big(\sigma(A_1), \sigma(A_2), \ldots, \sigma(A_d)\big),$$

and produces the spaces $T_{\lambda;G}$ of G-orbits of the structural formulae from T_λ. The number $n_{\lambda;G}$ of these G-orbits is therefore an upper bound for the number $N_{\lambda;\Theta}$ of experimentally known derivatives with composition $\Theta x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_d^{\lambda_d}$:

$$N_{\lambda;\Theta} \leqslant n_{\lambda;G}$$

for any partition $\lambda \in P_d$. In the cases of mono-substituted derivatives ($\lambda = (d-1, 1)$), di-substituted homogeneous derivatives ($\lambda = (d-2, 2)$), and di-substituted heterogeneous derivatives ($\lambda = (d-2, 1^2)$), the experimenters, sometimes, are certain that the corresponding numbers $N_{\lambda;\Theta}$ attain their maximum values $n_{\lambda;G}$. In other words, all possible λ-*derivatives* are prepared. In the ideal (but unattainable) situation $N_{\lambda;\Theta} = n_{\lambda;G}$ for all partitions $\lambda \in P_d$, and these equalities define the symmetry group G up to so-called combinatorial equivalence (see [1, section 4; 2, section 26; 3, theorem 5.2.5]).

The *simple substitution reactions*

$$x_1^{\mu_1} \cdots x_i^{\mu_i} \cdots x_j^{\mu_j} \cdots \quad \longrightarrow \quad x_1^{\lambda_1} \cdots x_i^{\lambda_i} \cdots x_j^{\lambda_j} \cdots,$$

where $\lambda, \mu \in P_d$, and $\mu_1 = \lambda_1, \ldots, \mu_i = \lambda_i + 1, \ldots, \mu_j = \lambda_j - 1, \ldots, \mu_d = \lambda_d$, that is, the replacement of a ligant of type x_i by a ligant of type x_j, $j < i$, are encoded in the mathematical model via two partial orderings: on the level of empirical formulae we write $\lambda < \mu$, and on the level of the structural picture

$$B = (B_1, B_2, \ldots, B_d) \quad \longrightarrow \quad A = (A_1, A_2, \ldots, A_d),$$

$A, B \in T_d$, $\lambda = \varphi(A)$, $\mu = \varphi(B)$, of the above simple substitution reaction, where A is obtained from B by moving an element $s \in B_i$ in the set B_j, we write $A < B$. More generally, we write $\lambda < \mu$ if λ can be got from μ by a finite number of the above simple substitutions (this is the well-known *dominance order* of partitions, see [4, section 6.1]), and we write $A < B$ if A can be obtained from B via a finite sequence of the above simple movements of elements (see [3, section 3.2]. The latter ordering can be pulled down on the orbit-space $T_{d;G} = G \setminus T_d$: $a < b$ if there are $A \in a$, $B \in b$ with $A < B$ (see [3, section 4.1]). If $a < b$, $a, b \in T_{d;G}$, the product which corresponds to a can, in principle, be synthesized from the product which corresponds to b via a finite sequence of simple substitution reactions. Thus, the partially ordered set $T_{d;G}$ portrays the possible *genetic relations* among the derivatives of the molecule under consideration (see [3]).

In this paper we consider parent substances with molecules that can be divided into a skeleton and six univalent substituents, and have the properties mentioned in the title. Two instances are the molecules of benzene C_6H_6 and cyclopropane C_3H_6, which have one mono-substitution derivative, and three and four di-substitution homogeneous derivatives, respectively.

The paper is stratified as follows. In section 2, theorem 2.1 describes the Lunn–Senior's group G of substitution isomerism of our compounds and corollaries 2.7 and 2.8 give upper bounds of the numbers of their di-substitution, and tri-substitution homogeneous derivatives. In sections 3–5, we list the possible simple substitution reactions among di-substitution homogeneous derivatives, on one hand, and di-substitution heterogeneous, and tri-substitution homogeneous derivatives, on the other. These substitution reactions allow us to identify some derivatives with their structural formulae.

2. The Lunn–Senior's group of substitution isomerism

The theorem below gives a characterization of the Lunn–Senior's groups of substitution isomerism of the compounds from the title.

Theorem 2.1. If an organic compound consists of a skeleton with six univalent substituents and has one mono-substitution and at least three di-substitution homogeneous derivatives, then its Lunn–Senior's group of substitution isomerism is conjugated in S_6 either to the dihedral group

$$\big\langle(123)(456),\,(14)(26)(35),\,(14)(25)(36)\big\rangle$$

of order 12, or to the cyclic group

$$\big\langle(123456)\big\rangle$$

of order 6, or to the dihedral group

$$\big\langle(123)(456),\,(14)(26)(35)\big\rangle$$

of order 6.

Proof. Since there exists only one mono-substitution derivative, we have $n_{(5,1);G}=1$, so the Lunn–Senior's group $G\leqslant S_6$ of substitution isomerism is transitive (see [5, section 3.1.1]). The existence of at least three di-substitution derivatives means that

$$n_{(4,2);G}\geqslant 3. \tag{2.2}$$

Since the partition $(4,2)$ dominates the partition $(4,1^2)$ with respect to the dominance order, [3, corollary 5.3.2] implies

$$n_{(4,2);G}\leqslant n_{(4,1^2);G}. \tag{2.3}$$

In particular, $n_{(4,1^2);G}\geqslant 3$. Therefore [6, equation 6.1.1] and [6, equation 6.1.2] yield $g_{(4,2);G}=g_{(4,1^2);G}=g_{(3,2,1);G}=g_{(2,1^4);G}=g_{(3,1^3);G}=0$. Then the linear system [5, formula 3.2.1] becomes

$$g_{(6);G}+g_{(3^2);G}+g_{(2^3);G}+g_{(2^2,1^2);G}-\big(|G|-1\big)=0, \tag{2.4a}$$
$$2g_{(3^2);G}+4g_{(2^2,1^2);G}-\big(|G|n_{(3^2);G}-20\big)=0, \tag{2.4b}$$
$$3g_{(2^3);G}+3g_{(2^2,1^2);G}-\big(|G|n_{(4,2);G}-15\big)=0, \tag{2.4c}$$
$$6g_{(2^3);G}+6g_{(2^2,1^2);G}-\big(|G|n_{(2^3);G}-90\big)=0, \tag{2.4d}$$
$$2g_{(2^2,1^2);G}-\big(|G|n_{(5,1);G}-6\big)=0, \tag{2.4e}$$
$$2g_{(2^2,1^2);G}-\big(|G|n_{(4,1^2);G}-30\big)=0, \tag{2.4f}$$
$$4g_{(2^2,1^2);G}-\big(|G|n_{(3,2,1);G}-60\big)=0, \tag{2.4g}$$
$$4g_{(2^2,1^2);G}-\big(|G|n_{(2^2,1^2);G}-180\big)=0, \tag{2.4h}$$
$$\big(|G|n_{(3,1^3);G}-120\big)=0, \tag{2.4i}$$

$$\big(|G|n_{(2,1^4);G} - 360\big) = 0, \tag{2.4j}$$
$$\big(|G|n_{(1^6);G} - 720\big) = 0. \tag{2.4k}$$

Since $n_{(5,1);G} = 1$, equalities (2.4e) and (2.4f) yield $|G| \geqslant 6$, and

$$|G|(n_{(4,1^2);G} - 1) = 24. \tag{2.5}$$

Then the inequality $n_{(4,1^2);G} \geqslant 3$ implies $|G| \leqslant 12$. On the other hand, equalities (2.4c) and (2.4d) imply

$$|G|\big(n_{(2^3);G} - 2n_{(4,2);G}\big) = 60.$$

Thus, $|G|$ is a common divisor of 24 and 60, so we obtain two possibilities for the order $|G|$ of the group G: $|G| = 12$ or $|G| = 6$.

If $|G| = 12$, then from (2.5) we get $n_{(4,1^2);G} = 3$, and the inequalities (2.2) and (2.3) yield $n_{(4,2);G} = 3$. Now, equalities (2.4e) and (2.4c) imply $g_{(2^2,1^2);G} = 3$ and $g_{(2^3);G} = 4$. Hence equality (2.4a) yields $g_{(6);G} + g_{(3^2);G} = 4$. The equality $g_{(3^2);G} = 0$ is impossible since for every cycle $\sigma \in G$ of length 6 its square σ^2 has cyclic type (3^2). Therefore $g_{(6);G} = g_{(3^2);G} = 2$. Let σ be a cycle of length 6. After eventual conjugation, we can suppose that $\sigma^2 = (123)(456) \in G$. Now, consider the cyclic group $K = \langle\sigma\rangle$ of order 6 and its cyclic subgroup $H = \langle(123)(456)\rangle$ that contains the two elements of G of cyclic type (3^2). If ι is one of the the elements of G of cyclic type (2^3), then

$$\iota H \iota^{-1} = H, \tag{2.6}$$

so $L = H\langle\iota\rangle$ is a subgroup of G of order 6. Now, we choose $\iota \notin K$ (since $g_{(2^3);G} = 4$ there are three elements of cyclic type (2^3) outside K). If we suppose that L is cyclic, then we would have $L = K$ (the two elements of order 6 in G are in K), and in particular, $\iota \in K$: a contradiction. Hence L is isomorphic to the dihedral group of order 6. Further, the equality (2.6) and the considerations in [6, section 7.1] yield that we can set $\iota = (14)(26)(35)$, so $L = \langle(123)(456), (14)(26)(35)\rangle$. Now, in accord to [6, section 7.3.1], we get that the group G is conjugated to the group $\langle(123)(456), (14)(26)(35), (14)(25)(36)\rangle$.

If $|G| = 6$, then $n_{(4,1^2);G} = 5$ and equality (2.4e) implies $g_{(2^2,1^2);G} = 0$. Then equality (2.4a) becomes $g_{(6);G} + g_{(3^2);G} + g_{(2^3);G} = 5$.

If G is the cyclic group of order 6, then it is generated, up to conjugation, by the cycle (123456), and $g_{(6);G} = 2$, $g_{(3^2);G} = 2$, and $g_{(2^3);G} = 1$. Now, equality (2.4c) yields $n_{(4,2);G} = 3$.

If G is the dihedral group of order 6, then $g_{(3^2);G} = 2$, $g_{(2^3);G} = 3$, and in accordance to the equality (2.4c), we obtain $n_{(4,2);G} = 4$. Now, we apply [6, theorem 5.1.1]. □

Theorem 2.1 implies immediately the following two corollaries which yield the numbers of derivatives of the molecules under consideration.

Corollary 2.7. If an organic compound consists of a skeleton with six univalent substituents and has one mono-substitution and at least three di-substitution homogeneous derivatives, and if its Lunn–Senior's group of substitution isomerism has order 12, then this compound has exactly three di-substitution homogeneous derivatives, at most three di-substitution heterogeneous derivatives, and at most three tri-substitution homogeneous derivatives.

Corollary 2.8. If an organic compound consists of a skeleton with six univalent substituents and has one mono-substitution and at least three di-substitution homogeneous derivatives, and if its Lunn–Senior's group G of substitution isomerism has order 6, then this compound has exactly three di-substitution homogeneous derivatives, at most five di-substitution heterogeneous derivatives, and at most four tri-substitution homogeneous derivatives in case G is cyclic, or has three or four di-substitution homogeneous derivatives, at most five di-substitution heterogeneous derivatives, and at most four tri-substitution homogeneous derivatives in case G is dihedral.

3. Genetic relations: the group G has order 12

Here we consider the possible genetic relations among the derivatives of our molecule structure in the case when its Lunn–Senior's group G of substitution isomerism has order 12. An example is the benzen molecule C_6H_6 (see [1, section 6; 3, section 6.3; 7]). In accord to [3, section 6.3] and theorem 3.1, we may suppose $G = \langle(123456), (13)(46)\rangle$ and then we obtain $T_{(4,2);G} = \{a_{(4,2)}, b_{(4,2)}, c_{(4,2)}\}$, where:

- $a_{(4,2)}$ is the G-orbit
$$\big\{\big(\{1,2,4,5\},\{3,6\}\big),\ \big(\{2,3,5,6\},\{1,4\}\big),\ \big(\{1,3,4,6\},\{2,5\}\big)\big\}$$
of the tabloid $A^{(4,2)} = (\{1,2,4,5\},\{3,6\})$;
- $b_{(4,2)}$ is the G-orbit
$$\big\{\big(\{1,2,3,4\},\{5,6\}\big),\ \big(\{2,3,4,5\},\{1,6\}\big),\ \big(\{3,4,5,6\},\{1,2\}\big),\ \big(\{1,4,5,6\},\{2,3\}\big),\ \big(\{1,2,5,6\},\{3,4\}\big),\ \big(\{1,2,3,6\},\{4,5\}\big)\big\}$$
of the tabloid $B^{(4,2)} = (\{1,2,3,4\},\ \{5,6\})$;
- $c_{(4,2)}$ is the G-orbit
$$\big\{\big(\{1,2,4,6\},\{3,5\}\big),\ \big(\{1,2,3,5\},\{4,6\}\big),\ \big(\{2,3,4,6\},\{1,5\}\big),\ \big(\{1,3,4,5\},\{2,6\}\big),\ \big(\{2,4,5,6\},\{1,3\}\big),\ \big(\{1,3,5,6\},\{2,4\}\big)\big\}$$
of the tabloid $C^{(4,2)} = (\{1,2,4,6\},\{3,5\})$.

Further, we get $T_{(3^2);G} = \{a_{(3^2)}, b_{(3^2)}, c_{(3^2)}\}$, where:

- $a_{(3^2)}$ is the G-orbit
$$\big\{\big(\{1,2,4\},\{3,5,6\}\big),\ \big(\{2,3,5\},\{1,4,6\}\big),\ \big(\{3,4,6\},\{1,2,5\}\big),\ \big(\{1,4,5\},\{2,3,6\}\big),\ \big(\{2,5,6\},\{1,3,4\}\big),\ \big(\{1,3,6\},\{2,4,5\}\big),$$

$$\big(\{2,3,6\},\{1,4,5\}\big),\ \big(\{1,2,5\},\{3,4,6\}\big),\ \big(\{1,4,6\},\{2,3,5\}\big),$$
$$\big(\{3,5,6\},\{1,2,4\}\big),\ \big(\{2,4,5\},\{1,3,6\}\big),\ \big(\{1,3,4\},\{2,5,6\}\big)\big\}$$

of the tabloid $A^{(3^2)} = (\{1,2,4\},\{3,5,6\})$;

- $b_{(3^2)}$ is the G-orbit

$$\big\{\big(\{1,2,3\},\{4,5,6\}\big),\ \big(\{2,3,4\},\{1,5,6\}\big),\ \big(\{3,4,5\},\{1,2,6\}\big),$$
$$\big(\{4,5,6\},\{1,2,3\}\big),\ \big(\{1,5,6\},\{2,3,4\}\big),\ \big(\{1,2,6\},\{3,4,5\}\big)\big\}$$

of the tabloid $B^{(3^2)} = (\{1,2,3\},\{4,5,6\})$;

- $c_{(3^2)}$ is the G-orbit

$$\big\{\big(\{1,3,5\},\{2,4,6\}\big),\big(\{2,4,6\},\{1,3,5\}\big)\big\}$$

of the tabloid $C^{(3^2)} = (\{1,3,5\},\{2,4,6\})$.

Moreover, we obtain $T_{(4,1^2);G} = \{a_{(4,1^2)}, b_{(4,1^2)}, c_{(4,1^2)}\}$, where:

- $a_{(4,1^2)}$ is the G-orbit

$$\big\{\big(\{1,2,4,5\},\{3\},\{6\}\big),\ \big(\{2,3,5,6\},\{4\},\{1\}\big),\ \big(\{1,3,4,6\},\{5\},\{2\}\big),$$
$$\big(\{1,2,4,5\},\{6\},\{3\}\big),\ \big(\{2,3,5,6\},\{1\},\{4\}\big),\ \big(\{1,3,4,6\},\{2\},\{5\}\big)\big\}$$

of the tabloid $A^{(4,1^2)} = (\{1,2,4,5\},\{3\},\{6\})$;

- $b_{(4,1^2)}$ is the G-orbit

$$\big\{\big(\{1,2,3,4\},\{5\},\{6\}\big),\ \big(\{2,3,4,5\},\{6\},\{1\}\big),\ \big(\{3,4,5,6\},\{1\},\{2\}\big),$$
$$\big(\{1,4,5,6\},\{2\},\{3\}\big),\ \big(\{1,2,5,6\},\{3\},\{4\}\big),\ \big(\{1,2,3,6\},\{4\},\{5\}\big),$$
$$\big(\{1,2,3,4\},\{6\},\{5\}\big),\ \big(\{2,3,4,5\},\{1\},\{6\}\big),\ \big(\{3,4,5,6\},\{2\},\{1\}\big),$$
$$\big(\{1,4,5,6\},\{3\},\{2\}\big),\ \big(\{1,2,5,6\},\{4\},\{3\}\big),\ \big(\{1,2,3,6\},\{5\},\{4\}\big)\big\}$$

of the tabloid $B^{(4,1^2)} = (\{1,2,3,4\},\{5\},\{6\})$;

- $c_{(4,1^2)}$ is the G-orbit

$$\big\{\big(\{1,2,4,6\},\{3\},\{5\}\big),\ \big(\{1,2,3,5\},\{4\},\{6\}\big),\ \big(\{2,3,4,6\},\{5\},\{1\}\big),$$
$$\big(\{1,3,4,5\},\{6\},\{2\}\big),\ \big(\{2,4,5,6\},\{3\},\{1\}\big),\ \big(\{1,3,5,6\},\{4\},\{2\}\big),$$
$$\big(\{1,2,4,6\},\{5\},\{3\}\big),\ \big(\{1,2,3,5\},\{6\},\{4\}\big),\ \big(\{2,3,4,6\},\{1\},\{5\}\big),$$
$$\big(\{1,3,4,5\},\{2\},\{6\}\big),\ \big(\{2,4,5,6\},\{1\},\{3\}\big),\ \big(\{1,3,5,6\},\{2\},\{4\}\big)\big\}$$

of the tabloid $C^{(4,1^2)} = (\{1,2,4,6\},\{3\},\{5\})$.

Since

$$A^{(3^2)} < A^{(4,2)},\qquad A^{(3^2)} < B^{(4,2)},\qquad A^{(3^2)} < C^{(4,2)},\qquad B^{(3^2)} < B^{(4,2)},$$
$$B^{(3^2)} < (123456)C^{(4,2)},\qquad C^{(3^2)} < (123456)C^{(4,2)},$$

and since

$$A^{(4,1^2)} < A^{(4,2)}, \qquad B^{(4,1^2)} < B^{(4,2)}, \qquad C^{(4,1^2)} < C^{(4,2)},$$

we have the following inequalities

$$a_{(3^2)} < a_{(4,2)}, \qquad a_{(3^2)} < b_{(4,2)}, \qquad a_{(3^2)} < c_{(4,2)}, \qquad b_{(3^2)} < b_{(4,2)},$$
$$b_{(3^2)} < c_{(4,2)}, \qquad c_{(3^2)} < c_{(4,2)},$$

and

$$a_{(4,1^2)} < a_{(4,2)}, \qquad a_{(4,1^2)} < b_{(4,2)}, \qquad a_{(4,1^2)} < c_{(4,2)}.$$

The diagrams below represent "Körner like" relations between the homogeneous di- and tri-substitution products of our molecule structure, which can be used for complete identification of these six derivatives:

$$\begin{array}{c} a_{(4,2)} \\ \downarrow \\ a_{(3^2)} \end{array} \qquad \begin{array}{ccc} & b_{(4,2)} & \\ \swarrow & \downarrow & \\ a_{(3^2)} & b_{(3^2)} & \end{array} \qquad \begin{array}{ccc} & c_{(4,2)} & \\ \swarrow & \downarrow & \searrow \\ a_{(3^2)} & b_{(3^2)} & c_{(3^2)} \end{array} .$$

The diagrams

$$\begin{array}{c} a_{(4,2)} \\ \downarrow \\ a_{(4,1^2)} \end{array} \qquad \begin{array}{c} b_{(4,2)} \\ \downarrow \\ b_{(4,1^2)} \end{array} \qquad \begin{array}{c} c_{(4,2)} \\ \downarrow \\ c_{(4,1^2)} \end{array}$$

show that, as a consequence, the heterogeneous di-substitution derivatives can also be identified completely.

Here the arrow $a \to b$ means that $a > b$ and the product that corresponds to b can be obtained from the product that corresponds to a via a simple substitution reaction.

4. Genetic relations: the group G has order 6 and is cyclic

In this section we describe the genetic relations of the molecule structure under question when its Lunn–Senior's group G of substitution isomerism is cyclic of order 6. In accord with theorem 2.1, we can suppose $G = \langle(123456)\rangle$. Then $T_{(4,2);G} = \{a_{(4,2)}, b_{(4,2)}, c_{(4,2)}\}$, where:

- $a_{(4,2)}$ is the G-orbit
$$\big\{\big(\{1,2,4,5\},\{3,6\}\big),\ \big(\{2,3,5,6\},\{1,4\}\big),\ \big(\{1,3,4,6\},\{2,5\}\big)\big\},$$
of the tabloid $A^{(4,2)} = (\{1,2,4,5\},\{3,6\})$;
- $b_{(4,2)}$ is the G-orbit
$$\big\{\big(\{1,2,3,4\},\{5,6\}\big),\ \big(\{2,3,4,5\},\{1,6\}\big),\ \big(\{3,4,5,6\},\{1,2\}\big),$$
$$\big(\{1,4,5,6\},\{2,3\}\big),\ \big(\{1,2,5,6\},\{3,4\}\big),\ \big(\{1,2,3,6\},\{4,5\}\big)\big\}$$
of the tabloid $B^{(4,2)} = (\{1,2,3,4\},\{5,6\})$;

- $c_{(4,2)}$ is the G-orbit

$$\big\{(\{1,2,4,6\},\{3,5\}),\ (\{1,2,3,5\},\{4,6\}),\ (\{2,3,4,6\},\{1,5\}),\ (\{1,3,4,5\},\{2,6\}),\ (\{2,4,5,6\},\{1,3\}),\ (\{1,3,5,6\},\{2,4\})\big\},$$

  of the tabloid $C^{(4,2)} = (\{1,2,4,6\},\{3,5\})$.

We have $T_{(3^2);G} = \{a_{(3^2)}, b_{(3^2)}, c_{(3^2)}, d_{(3^2)}\}$, where:

- $a_{(3^2)}$ is the G-orbit

$$\big\{(\{1,2,4\},\{3,5,6\}),\ (\{2,3,5\},\{1,4,6\}),\ (\{3,4,6\},\{1,2,5\}),\ (\{1,4,5\},\{2,3,6\}),\ (\{2,5,6\},\{1,3,4\}),\ (\{1,3,6\},\{2,4,5\})\big\},$$

  of the tabloid $A^{(3^2)} = (\{1,2,4\},\{3,5,6\})$;
- $b_{(3^2)}$ is the G-orbit

$$\big\{(\{1,2,5\},\{3,4,6\}),\ (\{2,3,6\},\{1,4,5\}),\ (\{1,3,4\},\{2,5,6\}),\ (\{2,4,5\},\{1,3,6\}),\ (\{3,5,6\},\{1,2,4\}),\ (\{1,4,6\},\{2,3,5\})\big\},$$

  of the tabloid $B^{(3^2)} = (\{1,2,5\},\{3,4,6\})$;
- $c_{(3^2)}$ is the G-orbit

$$\big\{(\{1,2,3\},\{4,5,6\}),\ (\{2,3,4\},\{1,5,6\}),\ (\{3,4,5\},\{1,2,6\}),\ (\{4,5,6\},\{1,2,3\}),\ (\{1,5,6\},\{2,3,4\}),\ (\{1,2,6\},\{3,4,5\})\big\}$$

  of the tabloid $C^{(3^2)} = (\{1,2,3\},\{4,5,6\})$;
- $d_{(3^2)}$ is the G-orbit

$$\big\{(\{1,3,5\},\{2,4,6\}),\ (\{2,4,6\},\{1,3,5\})\big\}$$

  of the tabloid $D^{(3^2)} = (\{1,3,5\},\{2,4,6\})$.

We also obtain $T_{(4,1^2);G} = \{a_{(4,1^2)}, b_{(4,1^2)}, c_{(4,1^2)}, d_{(4,1^2)}, e_{(4,1^2)}\}$, where:

- $a_{(4,1^2)}$ is the G-orbit

$$\big\{(\{1,2,4,5\},\{3\},\{6\}),\ (\{2,3,5,6\},\{4\},\{1\}),\ (\{1,3,4,6\},\{5\},\{2\}),\ (\{1,2,4,5\},\{6\},\{3\}),\ (\{2,3,5,6\},\{1\},\{4\}),\ (\{1,3,4,6\},\{2\},\{5\})\big\}$$

  of the tabloid $A^{(4,1^2)} = (\{1,2,4,5\},\{3\},\{6\})$;
- $b_{(4,1^2)}$ is the G-orbit

$$\big\{(\{1,2,3,4\},\{5\},\{6\}),\ (\{2,3,4,5\},\{6\},\{1\}),\ (\{3,4,5,6\},\{1\},\{2\}),\ (\{1,4,5,6\},\{2\},\{3\}),\ (\{1,2,5,6\},\{3\},\{4\}),\ (\{1,2,3,6\},\{4\},\{5\})\big\}$$

  of the tabloid $B^{(4,1^2)} = (\{1,2,3,4\},\{5\},\{6\})$;

- $c_{(4,1^2)}$ is the G-orbit
$$\big\{\big(\{1,2,3,4\},\{6\},\{5\}\big),\ \big(\{2,3,4,5\},\{1\},\{6\}\big),\ \big(\{3,4,5,6\},\{2\},\{1\}\big),\\ \big(\{1,4,5,6\},\{3\},\{2\}\big),\ \big(\{1,2,5,6\},\{4\},\{3\}\big),\ \big(\{1,2,3,6\},\{5\},\{4\}\big)\big\}$$
of the tabloid $C^{(4,1^2)} = (\{1,2,3,4\},\{6\},\{5\})$;
- $d_{(4,1^2)}$ is the G-orbit
$$\big\{\big(\{1,2,4,6\},\{3\},\{5\}\big),\ \big(\{1,2,3,5\},\{4\},\{6\}\big),\ \big(\{2,3,4,6\},\{5\},\{1\}\big),\\ \big(\{1,3,4,5\},\{6\},\{2\}\big),\ \big(\{2,4,5,6\},\{1\},\{3\}\big),\ \big(\{1,3,5,6\},\{2\},\{4\}\big)\big\}$$
of the tabloid $D^{(4,1^2)} = (\{1,2,4,6\},\{3\},\{5\})$;
- $e_{(4,1^2)}$ is the G-orbit
$$\big\{\big(\{1,2,4,6\},\{5\},\{3\}\big),\ \big(\{1,2,3,5\},\{6\},\{4\}\big),\ \big(\{2,3,4,6\},\{1\},\{5\}\big),\\ \big(\{1,3,4,5\},\{2\},\{6\}\big),\ \big(\{2,4,5,6\},\{3\},\{1\}\big),\ \big(\{1,3,5,6\},\{4\},\{2\}\big)\big\}$$
of the tabloid $E^{(4,1^2)} = (\{1,2,4,6\},\{5\},\{3\})$.

We have
$$A^{(3^2)} < A^{(4,2)},\quad A^{(3^2)} < B^{(4,2)},\quad A^{(3^2)} < C^{(4,2)},\quad B^{(3^2)} < A^{(4,2)},\\ B^{(3^2)} < (153)(264)B^{(4,2)},\quad B^{(3^2)} < (123456)C^{(4,2)},\quad C^{(3^2)} < B^{(4,2)},\\ C^{(3^2)} < (123456)C^{(4,2)},\quad D^{(3^2)} < (123456)C^{(4,2)},$$
and
$$A^{(4,1^2)} < A^{(4,2)},\quad B^{(4,1^2)} < B^{(4,2)},\quad C^{(4,1^2)} < B^{(4,2)},\quad D^{(4,1^2)} < C^{(4,2)},\\ E^{(4,1^2)} < C^{(4,2)},$$
so
$$a_{(3^2)} < a_{(4,2)},\quad a_{(3^2)} < b_{(4,2)},\quad a_{(3^2)} < c_{(4,2)}, \tag{4.1}$$
$$b_{(3^2)} < a_{(4,2)},\quad b_{(3^2)} < b_{(4,2)},\quad b_{(3^2)} < c_{(4,2)}, \tag{4.2}$$
$$c_{(3^2)} < b_{(4,2)},\quad c_{(3^2)} < c_{(4,2)}, \tag{4.3}$$
$$d_{(3^2)} < c_{(4,2)}, \tag{4.4}$$
and
$$a_{(4,1^2)} < a_{(4,2)},\quad b_{(4,1^2)} < b_{(4,2)}, \tag{4.5}$$
$$c_{(4,1^2)} < b_{(4,2)},\quad d_{(4,1^2)} < c_{(4,2)},\quad e_{(4,1^2)} < c_{(4,2)}. \tag{4.6}$$

The inequalities (4.1)–(4.4) indicate the existence of the corresponding (simple) substitution reactions among the $(4,2)$- and the (3^2)-derivatives, and these substitution reactions can be used for complete identification of all $(4,2)$-derivatives. Indeed, two, three, and four (3^2)-products can be synthesized from the $(4,2)$-derivatives which correspond to $a_{(4,2)}$, $b_{(4,2)}$, and $c_{(4,2)}$, respectively.

The following sets of structural formulae of (3^2)-derivatives can be distinguished:

$$\{a_{(3^2)}, b_{(3^2)}\}, \qquad \{c_{(3^2)}\}, \qquad \{d_{(3^2)}\}.$$

Indeed, the products that correspond to the elements of these sets can be synthesized from three, two, and one $(4,2)$-derivatives, respectively.

The inequalities (4.5), (4.6) indicate the existence of the corresponding (simple) substitution reactions among $(4,2)$- and $(4,1^2)$-derivatives, and by means of these substitution reactions we can identify the following sets of $(4,1^2)$-derivatives:

$$\{a_{(4,1^2)}\}, \qquad \{b_{(4,1^2)}, c_{(4,1^2)}\}, \qquad \{d_{(4,1^2)}, e_{(4,1^2)}\}.$$

Indeed, the product that corresponds to $a_{(4,1^2)}$ can be synthesized only from the identifiable $a_{(4,2)}$, the products that correspond to $b_{(4,2)}$ and $c_{(4,1^2)}$ can be synthesized only from the identifiable $b_{(4,2)}$, and the products that correspond to $d_{(4,2)}$ and $e_{(4,1^2)}$ can be synthesized only from the identifiable $c_{(4,2)}$.

5. Genetic relations: the group G has order 6 and is dihedral

In this section we describe the genetic relations of the molecule structure under question when its Lunn–Senior's group G of substitution isomerism has order 6, and is dihedral. An instance is the molecule of cyclopropane C_3H_6 (see [6]). In accord with theorem 2.1, we can suppose $G = \langle(123)(456), (14)(26)(35)\rangle$. Then $T_{(4,2);G} = \{a_{(4,2)}, b_{(4,2)}, c_{(4,2)}, d_{(4,2)}, \}$, where:

- $a_{(4,2)}$ is the G-orbit

$$\big\{\big(\{1,2,3,4\},\{5,6\}\big),\ \big(\{1,2,3,5\},\{4,6\}\big),\ \big(\{1,2,3,6\},\{4,5\}\big),\ \big(\{2,4,5,6\},\{1,3\}\big),\ \big(\{3,4,5,6\},\{1,2\}\big),\ \big(\{1,4,5,6\},\{2,3\}\big)\big\}$$

  of the tabloid $A^{(4,2)} = (\{1,2,3,4\},\{5,6\})$;
- $b_{(4,2)}$ is the G-orbit

$$\big\{\big(\{1,2,4,5\},\{3,6\}\big),\ \big(\{2,3,5,6\},\{1,4\}\big),\ \big(\{1,3,4,6\},\{2,5\}\big)\big\}$$

  of the tabloid $B^{(4,2)} = (\{1,2,4,5\},\{3,6\})$;
- $c_{(4,2)}$ is the G-orbit

$$\big\{\big(\{1,2,4,6\},\{3,5\}\big),\ \big(\{2,3,4,5\},\{1,6\}\big),\ \big(\{1,3,5,6\},\{2,4\}\big)\big\}$$

  of the tabloid $C^{(4,2)} = (\{1,2,4,6\},\{3,5\})$;
- $d_{(4,2)}$ is the G-orbit

$$\big\{\big(\{1,2,5,6\},\{3,4\}\big),\ \big(\{2,3,4,6\},\{1,5\}\big),\ \big(\{1,3,4,5\},\{2,6\}\big)\big\}$$

  of the tabloid $D^{(4,2)} = (\{1,2,5,6\},\{3,4\})$.

We have $T_{(3^2);G} = \{a_{(3^2)}, b_{(3^2)}, c_{(3^2)}, d_{(3^2)}\}$, where:

- $a_{(3^2)}$ is the G-orbit

$$\big\{(\{1,2,4\},\{3,5,6\}),\ (\{2,3,5\},\{1,4,6\}),\ (\{1,3,6\},\{2,4,5\}),\\ (\{2,4,5\},\{1,3,6\}),\ (\{3,5,6\},\{1,2,4\}),\ (\{1,4,6\},\{2,3,5\})\big\},$$

  of the tabloid $A^{(3^2)} = (\{1,2,4\},\{3,5,6\})$;
- $b_{(3^2)}$ is the G-orbit

$$\big\{(\{1,2,5\},\{3,4,6\}),\ (\{2,3,6\},\{1,4,5\}),\ (\{1,3,4\},\{2,5,6\}),\\ (\{1,4,5\},\{2,3,6\}),\ (\{2,5,6\},\{1,3,4\}),\ (\{3,4,6\},\{1,2,5\})\big\}$$

  of the tabloid $B^{(3^2)} = (\{1,2,5\},\{3,4,6\})$;
- $c_{(3^2)}$ is the G-orbit

$$\big\{(\{1,2,6\},\{3,4,5\}),\ (\{2,3,4\},\{1,5,6\}),\ (\{1,3,5\},\{2,4,6\}),\\ (\{3,4,5\},\{1,2,6\}),\ (\{1,5,6\},\{2,3,4\}),\ (\{2,4,6\},\{1,3,5\})\big\}$$

  of the tabloid $C^{(3^2)} = (\{1,2,6\},\{3,4,5\})$;
- $d_{(3^2)}$ is the G-orbit

$$\big\{(\{1,2,3\},\{4,5,6\}),\ (\{4,5,6\},\{1,2,3\})\big\}$$

  of the tabloid $D^{(3^2)} = (\{1,2,3\},\{4,5,6\})$.

Moreover, we obtain $T_{(4,1^2);G} = \{a_{(4,1^2)}, b_{(4,1^2)}, c_{(4,1^2)}, d_{(4,1^2)}, e_{(4,1^2)}\}$, where:

- $a_{(4,1^2)}$ is the G-orbit

$$\big\{(\{1,2,3,4\},\{5\},\{6\}),\ (\{1,2,3,5\},\{6\},\{4\}),\ (\{1,2,3,6\},\{4\},\{5\}),\\ (\{2,4,5,6\},\{1\},\{3\}),\ (\{3,4,5,6\},\{2\},\{1\}),\ (\{1,4,5,6\},\{3\},\{2\})\big\}$$

  of the tabloid $A^{(4,1^2)} = (\{1,2,3,4\},\{5\},\{6\})$;
- $b_{(4,1^2)}$ is the G-orbit

$$\big\{(\{1,2,3,4\},\{6\},\{5\}),\ (\{1,2,3,5\},\{4\},\{6\}),\ (\{1,2,3,6\},\{5\},\{4\}),\\ (\{2,4,5,6\},\{3\},\{1\}),\ (\{3,4,5,6\},\{1\},\{2\}),\ (\{1,4,5,6\},\{2\},\{3\})\big\}$$

  of the tabloid $B^{(4,1^2)} = (\{1,2,3,4\},\{6\},\{5\})$;
- $c_{(4,1^2)}$ is the G-orbit

$$\big\{(\{1,2,4,5\},\{3\},\{6\}),\ (\{2,3,5,6\},\{1\},\{4\}),\ (\{1,3,4,6\},\{2\},\{5\}),\\ (\{1,2,4,5\},\{6\},\{3\}),\ (\{2,3,5,6\},\{4\},\{1\}),\ (\{1,3,4,6\},\{5\},\{2\})\big\}$$

  of the tabloid $C^{(4,1^2)} = (\{1,2,4,5\},\{3\},\{6\})$;

- $d_{(4,1^2)}$ is the G-orbit

$$\big\{\big(\{1,2,4,6\},\{3\},\{5\}\big),\ \big(\{2,3,4,5\},\{1\},\{6\}\big),\ \big(\{1,3,5,6\},\{2\},\{4\}\big),\\ \big(\{2,3,4,5\},\{6\},\{1\}\big),\ \big(\{1,3,5,6\},\{4\},\{2\}\big),\ \big(\{1,2,4,6\},\{5\},\{3\}\big)\big\}$$

  of the tabloid $D^{(4,1^2)} = (\{1,2,4,6\},\{3\},\{5\})$;

- $e_{(4,1^2)}$ is the G-orbit

$$\big\{\big(\{1,2,5,6\},\{3\},\{4\}\big),\ \big(\{2,3,4,6\},\{1\},\{5\}\big),\ \big(\{1,3,4,5\},\{2\},\{6\}\big),\\ \big(\{1,3,4,5\},\{6\},\{2\}\big),\ \big(\{1,2,5,6\},\{4\},\{3\}\big),\ \big(\{2,3,4,6\},\{5\},\{1\}\big)\big\}$$

  of the tabloid $E^{(4,1^2)} = (\{1,2,5,6\},\{3\},\{4\})$.

This yields the inequalities

$$A^{(3^2)} < A^{(4,2)},\quad A^{(3^2)} < B^{(4,2)},\quad A^{(3^2)} < C^{(4,2)},\quad B^{(3^2)} < (123)(456)A^{(4,2)},\\ B^{(3^2)} < B^{(4,2)},\quad B^{(3^2)} < D^{(4,2)},\quad C^{(3^2)} < (132)(465)A^{(4,2)},\quad C^{(3^2)} < C^{(4,2)},\\ C^{(3^2)} < D^{(4,2)},\quad D^{(3^2)} < A^{(4,2)},$$

and

$$A^{(4,1^2)} < A^{(4,2)},\quad B^{(4,1^2)} < A^{(4,2)},\quad C^{(4,1^2)} < B^{(4,2)},\quad D^{(4,1^2)} < C^{(4,2)},\\ E^{(4,1^2)} < D^{(4,2)},$$

so

$$a_{(3^2)} < a_{(4,2)},\quad a_{(3^2)} < b_{(4,2)},\quad a_{(3^2)} < c_{(4,2)}, \tag{5.1}$$
$$b_{(3^2)} < a_{(4,2)},\quad b_{(3^2)} < b_{(4,2)},\quad b_{(3^2)} < d_{(4,2)}, \tag{5.2}$$
$$c_{(3^2)} < a_{(4,2)},\quad c_{(3^2)} < c_{(4,2)},\quad c_{(3^2)} < d_{(4,2)}, \tag{5.3}$$
$$d_{(3^2)} < a_{(4,2)}, \tag{5.4}$$

and

$$a_{(4,1^2)} < a_{(4,2)},\quad b_{(4,1^2)} < a_{(4,2)}, \tag{5.5}$$
$$c_{(4,1^2)} < b_{(4,2)},\quad d_{(4,1^2)} < c_{(4,2)},\quad e_{(4,1^2)} < d_{(4,2)}. \tag{5.6}$$

The inequalities (5.1)–(5.4) indicate the existence of the corresponding (simple) substitution reactions among the $(4,2)$- and the (3^2)-derivatives, and the inequalities (5.5), (5.6) indicate the existence of the corresponding (simple) substitution reactions among the $(4,2)$- and the $(4,1^2)$-derivatives.

These substitution reactions can be used for distinguishing the products that correspond to different sets from the following sets of structural formulae of $(4,2)$-derivatives:

$$\{a_{(4,2)}\},\qquad \{b_{(4,2)}, c_{(4,2)}, d_{(4,2)}\},$$

and from the following sets of structural formulae of (3^2)-derivatives:

$$\{a_{(3^2)}, b_{(3^2)}, c_{(3^2)}\},\qquad \{d_{(3^2)}\}.$$

Indeed, it is enough to note that from the product which corresponds to $a_{(4,2)}$ can be synthesized four (3^2)-derivatives and from the products that correspond to the elements of the set $\{b_{(4,2)}, c_{(4,2)}, d_{(4,2)}\}$, can be synthesized two (3^2)-derivatives. The products that correspond to the sets $\{a_{(3^2)}, b_{(3^2)}, c_{(3^2)}\}$, and $\{d_{(3^2)}\}$ can be synthesized from two and one $(4, 2)$-derivatives, respectively.

Using the above substitution reactions, we also can identify the products corresponding to the following sets of structural formulae of $(4, 1^2)$-derivatives:

$$\{a_{(4,1^2)}, b_{(4,1^2)}\}, \qquad \{c_{(4,1^2)}, d_{(4,1^2)}, e_{(4,1^2)}\}.$$

This is because both product that correspond to $a_{(4,1^2)}$ and $b_{(4,1^2)}$ can be synthesized from the identifiable $a_{(4,2)}$, and the products which correspond to $c_{(4,1^2)}$, $d_{(4,1^2)}$, and $e_{(4,1^2)}$ can be obtained from the products that correspond to $b_{(4,2)}$, $c_{(4,2)}$, and $d_{(4,2)}$.

Acknowledgement

This work is partially supported by Grant MM-1106/2001 of the Bulgarian Foundation of Scientific Research.

References

[1] A.C. Lunn and J.K. Senior, J. Phys. Chem. 33 (1929) 1027.
[2] G. Pólya, Acta Math. 68 (1937) 145. English translation in: G. Pólya and R.C. Read, *Combinatorial Enumeration of Groups, Graphs and Chemical Compounds* (Springer, New York, 1987) pp. 1–95.
[3] V.V. Iliev, Commun. Math. Comput. Chem. MATCH 40 (1999) 153.
[4] A. Kerber, *Applied Finite Group Actions* (Springer, Berlin, 1999).
[5] V.V. Iliev, Commun. Math. Comput. Chem. MATCH 43 (2001) 67.
[6] V.V. Iliev, Commun. Math. Comput. Chem. MATCH 43 (2001) 79.
[7] W. Hässelbarth, J. Comput. Chem. 8 (1987) 700.