

# New Approaches to Degenerate Interconversions of Organic Structures: 1. Levels of Structural Degeneracy and Computer-aided Search for Degenerate Rearrangements\*

Serge S. Tratch,<sup>a,\*\*</sup> Marina S. Molchanova,<sup>b</sup> and Nikolai S. Zefirov<sup>a</sup>

<sup>a</sup>Department of Chemistry, Moscow State University, Leninskie Gory, Moscow 119992, Russia

<sup>b</sup>Zelinskii Institute of Organic Chemistry, Russian Academy of Sciences, Leninskii prospekt 47, Moscow 119991, Russia

RECEIVED OCTOBER 4, 2005; REVISED NOVEMBER 8, 2005; ACCEPTED JANUARY 26, 2006

*Keywords*  
degenerate interconversions  
graph-based classification  
reaction design problems  
ARGENT–1 software  
new "Cope-like" rearrangements

The general methodology of systematic search for degenerate interconversions must be based on a certain hierarchy of degeneracy. The levels and sublevels of the proposed hierarchy are represented by topological and structural types and subtypes of degeneracy, which correspond to specific equations and to graphs identifying them. Graph operations that enable an efficient search for arbitrary types of organic reactions are formulated, and problems specific to generation of degenerate transformations are considered. In conclusion, results of computer-aided search (using the ARGENT–1 software) for new heteroderivatives and structural analogs of the Cope reaction are discussed. Many of the rearrangements thus found (especially [3,3]-shifts in cations and anions, which have not been studied to date) can be subjects of experimental investigations.

## INTRODUCTION

Degenerate transformations (DTs) are reversible interconversions in which chemical and/or spatial structures of the initial reagents (educts) and reaction products cannot be distinguished from each other. Degenerate transformations of individual species (neutral molecules, cations, anions, or radicals) with breaking/formation of chemical bonds are referred to as degenerate rearrangements. In the discussion of such rearrangements, the spatial structure of the rearranging particle is usually disregarded. From the standpoint of their mechanism, DTs can represent elementary stages of reactions, concerted

one-stage processes, or final results of more complicated, two- or multistage interconversions.<sup>1</sup> Examples are such well-known DTs as the [1,2]-shift in the 2-norbornyl cation (Figure 1a), which is one of the stages of the Wagner-Meerwein rearrangement; the Cope rearrangement in hexadiene-1,5 (Figure 1b), which is a concerted six-centered process; and prototropic tautomerizations of carboxylic acids (Figure 1c), for which a dissociative mechanism with intermediate formation of carboxylate anions is usually postulated. The degenerate (at R' = R'') isomerization of  $\alpha$ -pyrones<sup>2</sup> (Figure 1d) can serve as an example of a multistage reaction whose primary products (ketenes)

\* The paper represents the "chemist-oriented" part of the investigation reported at MATH/CHEM/COMP 2005 (Dubrovnik, June 20–25, 2005). The main ideas of the second, more mathematically involved part of our report are briefly outlined in the last section of the paper.

\*\* Author to whom correspondence should be addressed. (E-mail: tratch@org.chem.msu.su)

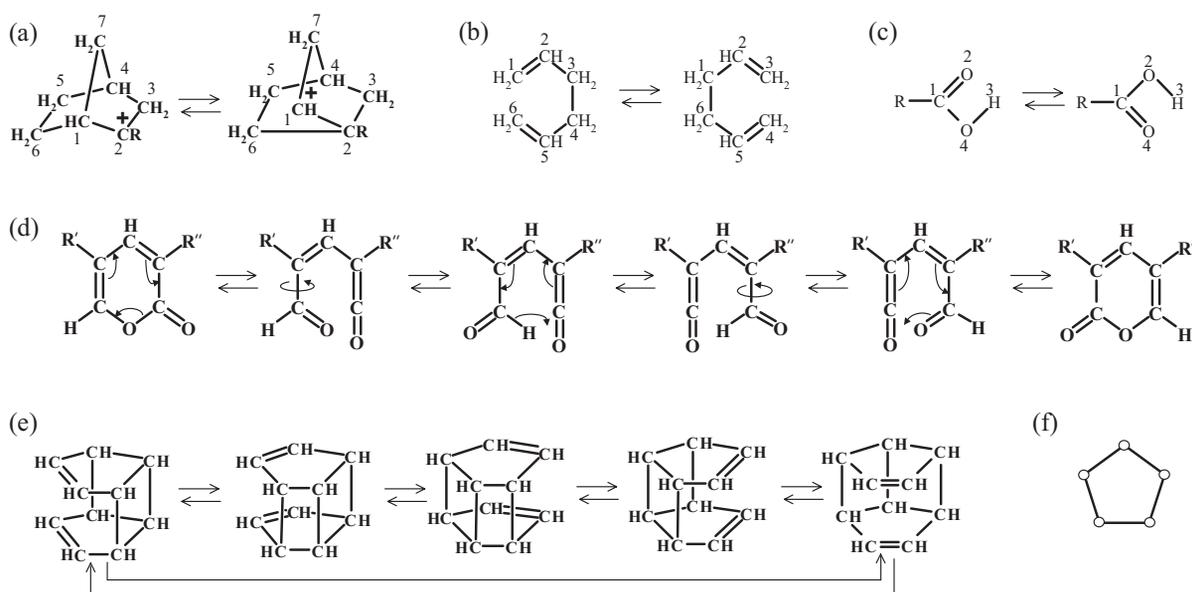


Figure 1. Interconversions of identical structures in an (a) elementary act of a skeletal rearrangement, (b) concerted process, (c) two-stage dissociation–recombination reaction, and (d) series of constitutional isomerization and stereoisomerization processes. The (e) multiply degenerate rearrangement of hypostrophene and its (f) reaction graph.

are formed and then disappear as a result of a nondegenerate [6]-electrocyclic process.

Methods used for the experimental study of DTs make it possible either to distinguish educts from reaction products (owing to the presence of isotopic labels or specially selected substituents) or to recognize degeneracy according to a decrease in the number of signals in NMR spectra. For example, the process in Figure 1a is nondegenerate at  $R \neq H$  and degenerate at  $R = H$ ; in the latter case, degeneracy must be manifested by merging of signals from protons of the  $C_1H$  and  $C_2H$  groups (and also from protons of the  $C_3H$  and  $C_7H$  groups), which become equivalent in the case of a DT. Note that in the case of a multiply or highly degenerate rearrangement (a reversible process involving more than two identical interconvertible particles), even all atoms or groups of atoms can become equivalent. As an example, Figure 1e depicts five structures of hypostrophene,<sup>3</sup> which consecutively pass into each other due to the Cope reaction. As easily seen, the multiply degenerate process in this case leads to the equivalence of all ten CH groups, a fact that enabled the authors of Ref. 3 to recognize degeneracy while studying the PMR spectra of dideuterio-substituted hypostrophene.

Although the phenomenon of degeneracy has been receiving significant attention of chemists over recent decades, the main theoretical results were obtained just for multiply degenerate rearrangements, such as the one in Figure 1e. More specifically, the principal objects of studies in mathematical chemistry were reaction graphs<sup>4</sup> of such rearrangements, *i.e.*, graphs whose vertices correspond to the rearranging particles while edges correspond to transitions between them. One of the possible

examples is the graph in Figure 1f, whose five vertices correspond to interconverted hypostrophene structures and five edges correspond to rearrangement processes in Figure 1e.

There are presently many published papers discussing the methods of construction, the metric properties, and the automorphism groups of reaction graphs that correspond to multiply degenerate rearrangements and especially to stereoisomerization processes (in the latter processes, the chemical constitution remains unchanged and degeneracy is associated with molecular motions that convert the stereoisomers in question into themselves). Since analysis of these results lies beyond the scope of this work, here we provide only some references to papers on the best studied reaction graphs,<sup>5–10</sup> as well as to more special studies on the general properties of these graphs.<sup>11–14</sup>

Among the works dealing with DTs, irrespective of the properties of their reaction graphs, we can distinguish numerous studies of permutational isomerization processes; see several fundamental papers on this topic.<sup>15–20</sup> In these (not necessarily degenerate) processes, the conversion of the initial structural or spatial isomer into the final one is associated with a certain permutation, which determines the renumbering of atoms. For example, let us consider the permutation  $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 1 & 6 & 5 & 4 \end{pmatrix}$ , which interconverts atoms 1 and 3, as well as atoms 4 and 6, but leaves atoms 2 and 5 as they are. When applied to the Cope rearrangement (Figure 1b), this permutation converts the "initial" double bonds 1–2 and 5–6, as well as single bonds 2–3, 3–4, and 4–5, into "final" bonds of the same

multiplicity (3–2, 5–4 and 2–1, 1–6, 6–5, respectively). This implies that the permutation in question really transforms the initial structure of hexadiene-1,5 in Figure 1b into an identical final structure. It is significant that many results obtained by theoretical analysis of "permutamer" interconversions (the term is taken from Ref. 17) turned out to be extremely helpful for dynamic stereochemistry<sup>19,20</sup> but have rarely been applied to degenerate transformations of constitutional isomers. In this connection, mention should be made of the use of permutation theory for solving several systematization problems. Classification of degenerate processes according to their chirality or achirality<sup>21</sup> and self-inverseness<sup>22</sup> or non-self-inverseness, as well as classification considering both factors,<sup>23</sup> have been illustrated in literature almost exclusively by examples of stereoisomer interconversions.

On the other hand, the permutation group theory is in principle an unsuitable tool for solving problems that must explicitly consider the chemical constitution of rearranging particles. Hence, it is no wonder that many degeneracy problems have received almost no attention until now. Among them, we can mention recognition of specific features responsible for degenerate interconversions of molecular structures, rigorous classification of structural changes during these interconversions, and systematic design of organic structures for which one can expect degenerate processes in general and rearrangements in particular.<sup>24</sup> The solution to all these problems necessarily requires the use of graph theory, because graphs are the very objects that can be used for unambiguously describing both the chemical constitution of the species in question and its change in any constitutional isomerization process.

#### SYSTEMATIC SEARCH FOR DEGENERATE INTERCONVERSIONS: OVERVIEW AND PROSPECTS

Despite being highly attractive, the problem of systematic search for new types of DTs has received only minor attention in literature. We know only three publications<sup>25–27</sup> addressing it, and in two of them DTs based on the Cope rearrangement are considered.<sup>28–30</sup> In the first of the cited papers,<sup>25</sup> A. T. Balaban used the previously constructed<sup>31</sup> complete lists of cubic multigraphs with 2 to 10 vertices and 1 to 4 loops. Substitution of four looped vertices (in connected as well as disconnected graphs) by the hexadiene-1,5 fragment enabled him to list all possible structural isomers of  $C_6H_6$ ,  $C_8H_8$ ,  $C_{10}H_{10}$ , and  $C_{12}H_{12}$  that can, in principle, undergo a degenerate Cope rearrangement. As an illustration, Figures 2a,b depict two DTs, namely the best investigated bullvalene rearrangement and the unusual isomerization of *cis*-1,2-dihydronaphthalene<sup>32</sup> predicted by Balaban.<sup>25</sup> Yet another example of multiple degeneracy based on the [3,3]-sigmatropic shift was discussed above, see Figure 1e.

In the second of the aforementioned papers,<sup>26</sup> its authors made an attempt to find such new modes of the Cope reaction that would lead to "averaging" of all carbon atoms in the rearranging particle. For this purpose, they developed a heuristic criterion, which enabled them to reveal (among valence isomers of annulenes  $C_nH_n$ ,  $n \leq 18$ , and polycarbons  $C_n$ ,  $n = 6, 8$ , obtained by various techniques) quite a number of structures for which multiply degenerate [3,3]-shifts are possible from the formal standpoint. Unfortunately, it turned out that none of such structures, except for those mentioned earlier in Ref. 25, can undergo the concerted process of the [3,3]-sigmatropic shift because of their extremely unfavorable rigid geometries.

In our opinion, systematic search for DTs should by no means be reduced to designing new concerted processes belonging to a single type selected in advance. Neither should it be aimed at finding only and specifically multiply degenerate rearrangements. On the contrary, new degenerate sigmatropic reactions different from the [3,3]-shift can undoubtedly be of interest (*e.g.*, [1,3]-, [2,2]-, or [1,5]-sigmatropic rearrangements), just like other degenerate pericyclic reactions (*e.g.*, intramolecular substitution or group transfer processes) and new types of degenerate interconversions of ions and radicals. In addition, one should bear in mind that a chemist can be interested not only in concerted DTs but also in multistage reactions, such as the rearrangement of  $\alpha$ -pyrones in Figure 1e. In the latter case, there is no need to make

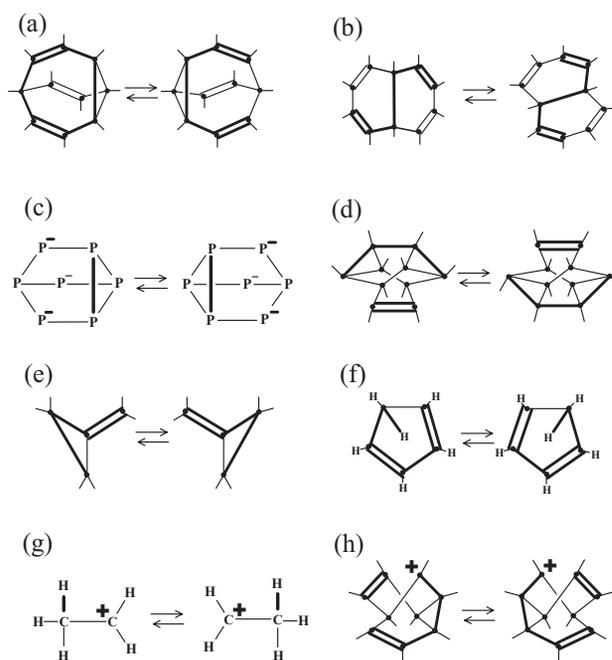


Figure 2. Examples of multiple degeneracy based on the (a,b) Cope rearrangement; (c,d) oxidation–reduction and cyclodismutation reactions; (e,f) [1,3]- and [1,5]-sigmatropic shifts; and (g,h) rearrangement processes in organic cations. In all cases, the broken bonds and bonds with changed multiplicity are marked by thick lines.

special allowances for steric factors, which often prevent the formation of the transition state required for a concerted process.

In Figures 2c–h, several well-known DTs that are essentially distinct from the Cope rearrangement are shown. The first two examples represent intramolecular oxidation–reduction processes in the heptaphosphide trianion<sup>33</sup> (Figure 2c; note the resemblance to the rearrangement of bullvalene) and [2+(1,2,1)]-substitution in the pentacyclic molecule of snoutene<sup>34</sup> (Figure 2d; in the used notation,<sup>35</sup> each number corresponds to a fragment that behaves as a single unbreakable unit during the reaction). The next two reactions can be regarded as [1,3]- and [1,5]-sigmatropic shifts (Figures 2e,f, respectively) in unsubstituted methylenecyclopropane<sup>36</sup> and cyclopentadiene.<sup>37</sup> The first of these rearrangements is symmetry-forbidden and probably proceeds via a biradical intermediate, whereas the second one, a symmetry-allowed reaction, is also typical of numerous monosubstituted derivatives of cyclopentadiene. Finally, the last two examples in Figure 2 illustrate interconversions of organic cations; the degenerate [1,2]-shift in the ethyl cation (Figure 2g) was the first DT whose reaction graph was investigated by Balaban.<sup>5</sup> For the rearrangement of the barbaralyl cation,<sup>38</sup> three types of DTs are usually postulated, and one of them is shown in Figure 2h. Note that the last two examples may be regarded as degenerate [1+(1,1)]- and [2+(3,2+)]-substitution processes, and both have been analyzed<sup>39,40</sup> from the standpoint of the permutation group theory.

The above data convincingly show that the Cope rearrangement is far from being the only type of DTs to which multiply degenerate transformations can correspond.<sup>41</sup> Nevertheless, even for the specific case of the [3,3]-sigmatropic shift, not all possibilities have been exhausted: the search in Refs. 25, 26 was limited to rearrangements of annulenes and polycarbons only. Below, considering just the same (*i.e.*, Cope) reaction, we present examples of both investigated and still unknown DTs illustrating the possible ways of designing derivatives and analogs of a given degenerate interconversion.

The first two examples in Figure 3 represent hetero-derivative [3,3]-sigmatropic shifts: degenerate interconversions of allyl carboxylates in Figure 3a have actually been mentioned in literature (see, *e.g.*, Ref. 43), whereas similar rearrangements of vinyloxymethyl substituted carbonyl compounds (see Figure 3b) are probably still unknown.<sup>44</sup> Note that, by analogy with heteroderivatives, one can also consider substituted derivatives of DTs with one or more hydrogen atoms replaced by suitable univalent substituents; this trivial procedure of degeneracy design is beyond the scope of this paper.

Another pair of examples demonstrates new modes of degeneracy obtained by addition of bonds that are not associated with substituents but remain completely unaffected by the DT. To proceed from the equation of the

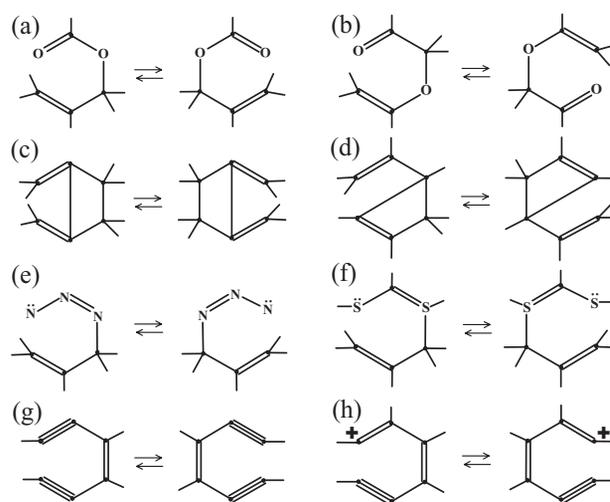


Figure 3. (a,c,e,g) Experimentally studied and (b,d,f,h) hypothetical [3,3]-sigmatropic shifts; all derivatives and analogs of the Cope rearrangement are represented in a uniform manner. Examples illustrate the introduction of (a,b) heteroatoms, (c,d) bonds unaffected by the reaction; (e,f) atoms changing their valence; and (g,h) bonds that change their multiplicity by two units.

"original" Cope reaction (Figure 1b) to equations expressing the investigated<sup>45</sup> rearrangement of 1,2-dimethylenecyclobutane (Figure 3c) and the still unknown rearrangement of 3-vinylcyclobutene (Figure 3d), one must add a single unchanged bond in both cases. Note that the DTs in Ref. 25 were generated just by "expansion" of the initial and rearranged hexadiene-1,5 structures; all bonds that do not change their multiplicity (and the corresponding atoms, here referred to as structural centers) are also present in the above examples of Figures 2a–h.

Even more interesting possibilities are revealed upon analysis of the structural analogs of the Cope reaction, *i.e.*, processes that also represent [3,3]-shifts but with somewhat different bond redistributions. One of the modifications is associated with the participation of a fragment containing two atoms which can change their valence by two units each. From the formal standpoint, the azide group represents just such a fragment, and the degenerate rearrangement of allyl azide (Figure 3e) is well-known.<sup>46</sup> (It is interesting that the authors of Ref. 46 believed that the first stage of the reaction is an attack of the "nitrene" nitrogen atom on a double bond; in a more recent paper,<sup>47</sup> this DT was already classified as [3,3]-sigmatropic.) Another unusual rearrangement with two valence-changing atoms can be represented by a [3,3 $\alpha\gamma$ ]-sigmatropic shift in *S*-allyl substituted sulfonium ylides (Figure 3f); the Greek letters denote the positions of sulfur atoms in the fragment. This "Cope-like" reaction was mentioned in our paper<sup>48</sup> as early as in 1980, but no attempts to implement it have been made to this day, probably because the synthesis of the initial ylide is too difficult.

Finally, the last two examples express DTs where, in contrast to the ordinary Cope rearrangement, either all (in Figure 3g) or some of the bonds (in Figure 3h) change their multiplicity by two units. The degenerate rearrangement of enediynes (Figure 3g) is well-known,<sup>49</sup> and its application (Bergman cyclization) is often used in contemporary synthetic chemistry. In contrast, the process in Figure 3h appears to be more unusual; this example shows that [3,3]-shifts, in principle, can proceed not only in uncharged systems but also in cationic, anionic, or radicalic ones.

Although this has not been specially mentioned, the Cope rearrangement was considered above as the one determining the type of the DT, different from types corresponding to other degenerate sigmatropic shifts, substitution processes, *etc.* However, the examples in Figures 3a–h indicate that classification of degeneracy is not so simple: rigorous characterization of DTs should take into account heteroatoms or substituents, bonds unaffected by the DT or changing their multiplicity by more than unity, and atoms changing their valence or carrying a charge. Hence, to describe and systematize degenerate transformations, one should elaborate a hierarchy whose levels would reflect certain information both about the structural changes involved (bond redistributions, transfers of charge or unpaired electron symbols) and on the specific structural features of rearranging particles (unaffected bonds, rings, the nature of atoms involved). Further, each level can be put in correspondence with a graph that adequately represents all information on the aforementioned structural factors. Finally, to generate such graphs efficiently, one needs to substantiate formal models, elaborate combinatorial algorithms, and write computer programs; these programs would be capable of a systematic search for the types of degenerate transformations that are actually interesting to a chemist. The main notions needed for clear formulation of all the above problems are briefly considered in the next two sections.

#### CLASSIFICATION OF DEGENERATE TRANSFORMATIONS ON THE BASIS OF THE GENERAL FORMAL-LOGICAL APPROACH

For arbitrary chemical processes (not necessarily degenerate), the problems related to their classification and design have already been solved. Our Formal-Logical Approach to organic reactions (see Refs. 50–53 and references therein) makes it possible to analyze all bond redistributions from the standpoint of graph theory and also to classify reactions according to their structural characteristics. Moreover, we have created a highly efficient computer program aimed at systematic search for new types of organic reactions. The main specific features of this program (named ARGENT-1) and the underlying mathematical models are considered in our papers.<sup>54,55</sup> It is

important that the working version of this software also enables the user to solve quite a number of problems related to degeneracy. Some results obtained by means of the ARGENT-1 program are briefly discussed below.

Let us initially illustrate the use of the main notions of the Formal-Logical Approach on a simple example: rearrangement of methylenecyclobutane and its deuterated derivatives.<sup>56</sup> This rearrangement, shown in Figure 4a, is degenerate only at  $R' = R'' = H$  or  $R' = R'' = D$ . Evidently, Figure 4a represents the complete chemical equation. All atoms in this equation can be partitioned into (a) reaction centers, *i.e.*, atoms with one or several bonds changing their multiplicity; (b) structural centers, *i.e.*, atoms with only unchanged bonds, of which at least two participate in the formation of chemically significant rings; and (c) atoms constituting (typically acyclic) substituents.<sup>57</sup> Further, removal of substituents only or of substituents together with structural centers leads to skeletal (*e.g.*, in Figure 4b) or reaction (*e.g.*, in Figure 4c) equations, respectively. In the obtained equations, the symbols of all chemical elements are still explicitly presented. If these symbols are substituted by some abstract symbols, then the reaction, skeletal, and chemical equations are transformed into the equations in Figures 4d–f, *i.e.*, into symbolic, structural equations, and equations without any special name. In the examples under consideration, substituents are denoted by squares (in Figure 4d); structural centers by small circles (in Figures 4d,e); and reaction centers by heavy dots (in Figures 4d,e,f). Note that other designations are used in the Formal-Logical Approach for reaction centers with unequal initial and final

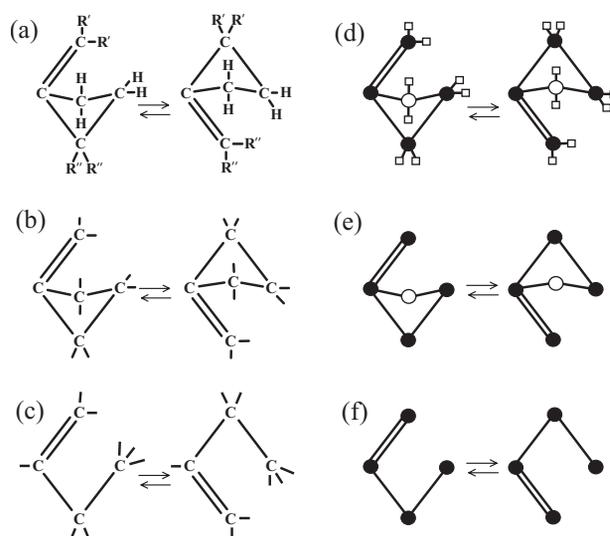


Figure 4. Main objects of the Formal-Logical Approach: (a) chemical, (b) skeletal, (c) reaction equations and (d–f) equations where the element symbols are replaced by abstract symbols corresponding to reaction centers, structural centers, and also atoms constituting the substituents. Symbolic equation (f) consists only of reaction centers, and structural equation (e) consists of both reaction and structural centers.

valences; the designation rules are discussed in detail in other publications.<sup>50,52</sup>

The equations considered here play a significant role in the multilevel hierarchical system<sup>53</sup> that was suggested for classification of organic reactions. The first three levels and the corresponding sublevels of this system (categories/subcategories, classes/subclasses, and interconversions types/subtypes of reaction systems) are of no special value for the systematization of DTs. (Maybe a single important remark is that, among 12 subcategories corresponding to reactions in charged and radicalic systems, only three can express degenerate reactions, namely, transformations of monocationic, monoanionic, or monoradicalic systems.) In contrast, the last two levels, *i.e.*, topological type/subtype and reaction type/subtype, are directly determined by the equations in Figure 4 (symbolic/structural and skeletal/chemical equations, respectively).

Since classification of structural changes in degenerate interconversions has not been specially discussed earlier,<sup>58</sup> it seems natural to apply the main ideas of the general classification to the particular case of DTs. To make the classification scheme as simple as possible, we consider only the three essential degeneracy levels that are associated with the types of chemical bonds actually taken into account; within individual levels, the degenerate processes are then partitioned into subtypes according to the nature of the atoms involved.

It is worthwhile to associate the upper (least detailed) level of the proposed system, the topological type, with the symbolic equation of a degenerate process. Accordingly, it is reasonable to characterize the corresponding degeneracy subtypes by reaction equations. One can easily see that the notions of topological type and subtype are formulated only on the basis of structural changes (see Figures 4f,c) but make no allowance either for cycles or for substituents. Further, at the intermediate level, the notions of the structural type and subtype of degeneracy can be associated with structural and skeletal equations, respectively (see Figures 4e,b). In this case, evidently, the cyclic structure of the rearranging particle(s) and even of the transition state (see Ref. 53) is taken into account.

At the lower, most detailed level of classification (determined by equations like those in Figures 4a,d), the actual processes with all substituents are considered. This means that chemical equations completely characterize degenerate transformations, whereas equations similar to that in Figure 4d describe some "pre-reactions," where all abstract symbols can be substituted by symbols of any chemical elements of suitable valences. Note that in the general Formal–Logical Approach, substituents are never explicitly taken into account, and therefore neither complete equations (such as those in Figures 4a,d) nor the corresponding graphs (see Figures 5a,d,g) are ever constructed during the search for new reactions. On the

contrary, in the case of DTs, these equations are not only classification attributes: as shown below, in some instances, exhaustive generation of multiply degenerate rearrangements requires all substituents to be explicitly taken into account.

## GRAPH-THEORETICAL FORMULATION OF DEGENERACY DESIGN PROBLEMS

Now let us show how equations similar to those of Figures 4a–f can be used in the search for new types of reactions in general and of degenerate transformations in particular. For this purpose, it is convenient to substitute each of the aforementioned kinds of equations (chemical, skeletal, reaction, the one corresponding to the "pre-reaction," structural, and symbolic) by a labeled graph. Here, these graphs are denoted by  $G_{\text{CEQ}}$ ,  $G_{\text{SKE}}$ ,  $G_{\text{REQ}}$ ,  $G_{\text{XEQ}}$ ,  $G_{\text{STR}}$ , and  $G_{\text{SEQ}}$ , respectively (see Figures 5a–f), and each of them provides the same information as the corresponding equation (*cf.* Figures 4a–f). Vertex labels in these graphs coincide with those in the equations, and edge labels are represented by double designations  $a/b$ , where  $a$  and  $b$  denote the multiplicities of the same edge in the left-hand and right-hand parts of the equation. In Figures 5a–f, only the labels 0/1, 1/2, 2/1, 1/0 are explicitly shown; the remaining labels (1/1), which correspond to unchanged single bonds, are omitted for simplicity. Note that since actual directions of all reactions are not taken into account in the Formal–Logical Approach,<sup>52</sup> graphs that are interconverted by replacement of all labels  $a/b$  by labels  $b/a$  (*i.e.*,

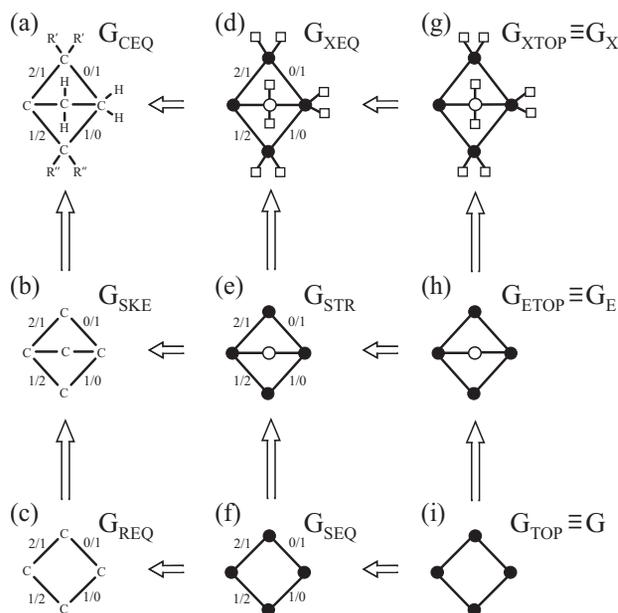


Figure 5. Graphs identifying the (a) chemical, (b) skeletal, (c) reaction, (d) corresponding to the "pre-reaction," (e) structural, and (f) symbolic equations, as well as (g–i) topology graphs from which they can be obtained. Short arrows signify operations of assigning edge and vertex labels; long arrows signify operations of graph expansion and attachment of acyclic substituents.

0/1 by 1/0, 1/0 by 0/1, etc.) are also regarded as indistinguishable.

Comparing Figures 5a–c with Figures 5d–f, one can see that to generate graphs  $G_{\text{CEQ}}$ ,  $G_{\text{SKE}}$ , or  $G_{\text{REQ}}$  from a given graph  $G_{\text{XEQ}}$ ,  $G_{\text{STR}}$ , or  $G_{\text{SEQ}}$ , respectively, one should just analyze all possible ways of substituting abstract symbols at the graph vertices by symbols of selected chemical elements. During this operation, denoted by three leftmost short arrows in Figure 5, one should naturally make allowance both for valences of actual elements and for symmetry considerations. One can similarly construct graphs  $G_{\text{XEQ}}$ ,  $G_{\text{STR}}$ , or  $G_{\text{SEQ}}$  from a given topology graph  $G_{\text{XTOP}}$ ,  $G_{\text{ETOP}}$ , or  $G_{\text{TOP}}$ , whose edges contain no information on the multiplicities of the corresponding bonds. The total number of possible edge labels needed to generate graphs  $G_{\text{XEQ}}$  and  $G_{\text{STR}}$  is equal to 15 (0/1, 1/2, 2/3, 0/2, 1/3, 0/3, 3/0, 3/1, 2/0, 3/2, 2/1, 1/0, 1/1, 2/2, and 3/3); however, only the first 12 labels are allowed for graphs  $G_{\text{SEQ}}$ . Transitions from graphs  $G_{\text{TOP}}$ ,  $G_{\text{ETOP}}$ , and  $G_{\text{XTOP}}$  (see Figures 5g–i) to graphs in Figures 5d–f are denoted by other three short arrows; note that vertices corresponding to reaction centers, structural centers, and substituents need not be actually distinguished in topology graphs.

The use of graph-theoretical considerations makes it possible to define three more operations that are important in the search for new nondegenerate and degenerate transformations. One of these operations, denoted by the lower three long arrows in Figure 5, represents the "expansion" of graph  $G_{\text{REQ}}$ ,  $G_{\text{SEQ}}$ , or  $G_{\text{TOP}}$  by addition of unaffected bonds and possibly new graph vertices, *i.e.*, structural centers. The second operation (see the upper three long arrows in Figure 5) is more of theoretical interest, because variation of substituents is in no case the key problem in any reaction generation task. The last operation, which is not shown in Figure 5, arises due to the fact that topology graphs ( $G_{\text{TOP}}$ ,  $G_{\text{ETOP}}$ , and  $G_{\text{XTOP}}$ ) for processes involving charged or radicalic species must also be generated from somewhat simpler graphs. More specifically, two vertices in the topology graph (according to the principles of the Formal–Logical Approach) can carry a sign of charge or an unpaired electron;<sup>52</sup> hence, one needs a procedure that can distribute two sign labels between vertices of a given "unsigned" graph ( $G$ ,  $G_{\text{E}}$ , or  $G_{\text{X}}$ , respectively). Note that, in the just considered rearrangement of methylenecyclobutane, charged and radicalic centers are absent, and therefore  $G \equiv G_{\text{TOP}}$ ,  $G_{\text{E}} \equiv G_{\text{ETOP}}$ , and  $G_{\text{X}} \equiv G_{\text{XTOP}}$ .

Let us illustrate the specific features of the processes involving charged species by degenerate [1,2]- (see Figure 6a) and [1,3]-shifts (Figure 6b) in the homotetrahydryl cation and by two nonequivalent (degenerate in Figure 6c and nondegenerate in Figure 6d) [1,2]-shifts in the homoprismyl cation. These transformations have been widely discussed in literature,<sup>4,6,7,11,13,14,39</sup> mainly in connection with their reaction graphs; similar known rearrangements

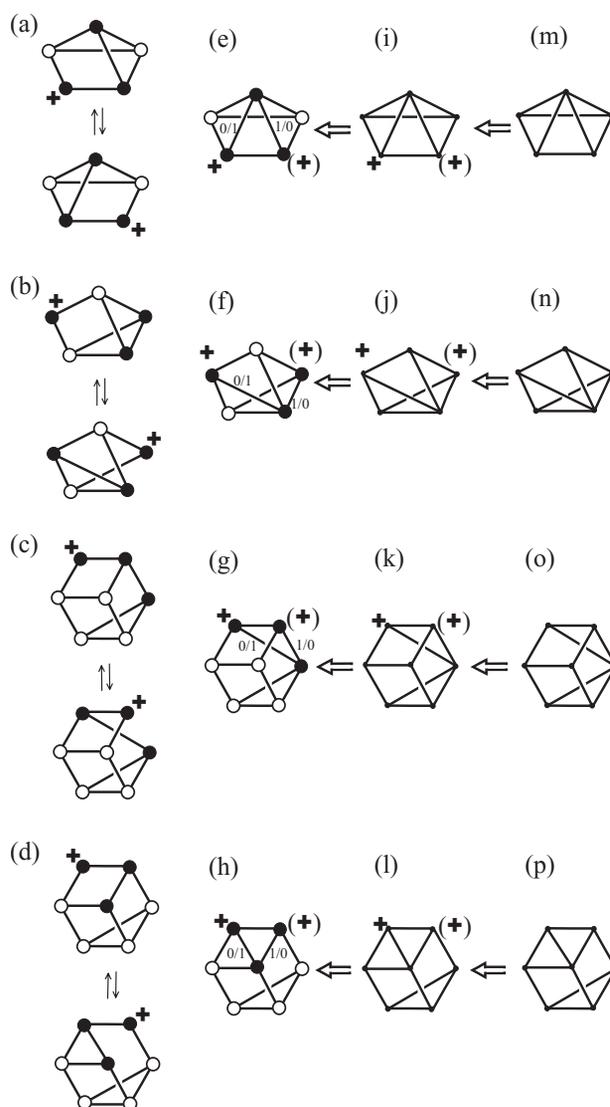


Figure 6. Structural equations for the processes of (a) [1,2]- and (b) [1,3]-shifts in the homotetrahydryl cation; (c) degenerate and (d) nondegenerate [1,2]-shifts in the homoprismyl cation; and also the corresponding graphs (e–h)  $G_{\text{STR}}$ , (i–l)  $G_{\text{ETOP}}$ , and (m–p)  $G_{\text{E}}$ . Note that graphs (m) and (n) are isomorphic and graphs (o) and (p) are not isomorphic.

of other carbonium ions have been extensively reviewed as well.<sup>59</sup> Structural equations in Figures 6a–d completely characterize the processes in question, because both reaction and structural centers are represented by CH groups only. Evidently, the examples in Figures 6a–d belong to one topological type (since they all represent the same three-centered electrophilic substitution and hence correspond to the same symbolic equation) but to different structural types. Note that the  $[i,j]$ -notation frequently used by chemists makes it possible to distinguish between [1,2]- and [1,3]-shifts in the homotetrahydryl cation (see Figures 6a,b), but in contrast to structural equations, it does not reveal the difference between two possible modes of the [1,2]-shift (Figure 6c,d) in the homoprismyl cation.

Graphs  $G_{STR}$  that identify the structural equations in Figures 6a–d are depicted in Figures 6e–h. A specific feature of these graphs is the presence of two sign labels "+" and "(+)," with the first one denoting the location of the sign in the upper part of the equation and the second one denoting its location in the lower part. The same labels are present in the "signed" topology graphs  $G_{ETOP}$  in Figures 6i–l, from which graphs  $G_{STR}$  can be generated, but absent in the "unsigned" graphs  $G_E$  in Figures 6m–p. It is easily seen that interconversion of sign labels, such as "+" and "(+)," in any graph  $G_{ETOP}$  leads to equations (corresponding to graphs  $G_{STR}$  and  $G_{SKE}$ ) whose parts are interchanged.

Now let us summarize the main specific features of operations that either successively produce graphs corresponding to the same hierarchy level (*e.g.*,  $G \Rightarrow G_{TOP} \Rightarrow G_{SEQ} \Rightarrow G_{REQ}$ ; see Figure 5) or enable passing to a lower-level, more detailed level (*e.g.*, from  $G_{TOP}$  to  $G_{ETOP}$  and then to  $G_{XTOP}$ ). Evidently, there are five such operations:

(1) *Assignment of sign labels to two vertices of graph  $G$ ,  $G_E$ , or  $G_X$ .* In the case of search for DTs, there are two possibilities. The first one is to skip this operation; in this case, the signed topology graph is identical to the unsigned one. The second possibility is to assign labels "+" and "(+)," or "–" and "(–)," or "•" and "(•)" to two graph vertices in all possible ways; in this case, the equations produced from any signed graph represent cationic, anionic, or radicalic processes, respectively.

(2) *Assignment of labels to all edges of graph  $G_{TOP}$ ,  $G_{ETOP}$ , or  $G_{XTOP}$ .* Labels corresponding to changes in bond multiplicities (0/1, 1/0, *etc.*) are allowed only for edges that join two reaction centers, whereas labels 1/1, 2/2, and 3/3 (associated with unaffected bonds) can be also used for other edges of the graph. In the case of graphs  $G_{TOP}$ , there are no unaffected bonds, and the corresponding labels 1/1, 2/2, and 3/3 are therefore absent.

(3) *Assignment of labels to all vertices of graph  $G_{SEQ}$ ,  $G_{STR}$ , or  $G_{XSEQ}$ .* If DTs are to be constructed, the graph whose vertices must be assigned atom symbols should itself correspond to a degenerate transformation. An important corollary: as a result of this operation, just as in the case of operation (2), degeneracy can disappear but it can never appear.

(4) *Expansion of graph  $G$ ,  $G_{TOP}$ ,  $G_{SEQ}$ , or  $G_{REQ}$  by adding unchanged bonds (and structural centers).* In contrast to operations (2) and (3), this operation can produce graphs corresponding to DTs even if, in principle, no DTs can be generated for the initial graphs; an example is discussed below.

(5) *Addition of substituents unaffected by the reaction.* This operation is not presently used in designing new reaction types. The only promising field of its use seems to be associated with the search for such multiply degenerate interconversions where just the presence of identical

substituents is the factor responsible for the appearance of multiple degeneracy.

Our discussion of degenerate interconversions would be incomplete without analyzing certain features that are specific to each hierarchy level. At the upper level, during successive generation of graphs  $G \Rightarrow G_{TOP} \Rightarrow G_{SEQ} \Rightarrow G_{REQ}$  (corresponding to equations in which all bonds change their multiplicity), an interesting problem arises for graphs  $G_{SEQ}$ . Let us explain its essence by the example of the symbolic equation in Figure 7a, which involves two centers X changing their valences by 2 units and one center X' changing its valence by 4 units. These centers are referred to as specific and dispecific, respectively<sup>50,52</sup> (such centers and their notations have not been considered above). At first sight, the transformation in Figure 7a cannot be regarded as degenerate, because, for example, the isolated center is differently (*i.e.*, as X' and X) represented in the left-hand and right-hand parts of the equation. However, when we pass to the reaction equation in Figure 7b (corresponding to the unrealistic interaction of a carbonyl compound with disulfurmonoxide and elementary sulfur), this distinction disappears. The fact that the hypothetical process in Figure 7b is undoubtedly degenerate shows that designations of reaction centers in symbolic equations (•, X, X', *etc.*) should not be taken into account in degeneracy checking, or otherwise the list of generated DTs will be incomplete. Actually, any such designation is determined only by labels of edges that are incident to the reaction center in a graph  $G_{SEQ}$ .

At the intermediate level of hierarchy, consideration of cycles in structural and skeletal equations can also lead to unexpected results. Addition of unchanged bonds (and possibly structural centers as well) to the symbolic or reaction equation representing a nondegenerate process can sometimes lead to the structural or skeletal equation corresponding to a degenerate process. For example, the undoubtedly nondegenerate addition–elimination reaction in Figure 7c is transformed into a degenerate rearrangement (see Figure 7d) after the addition of one unchanged bond.<sup>60</sup> A possible example is the hypothetical isomerization of cyclopropenylcarbene in Figure 7e, which could become real if cycloaddition of carbenes to a double bond were reversible. This specific feature is of value for practical implementation as well: in the systematic search for DTs, one should perform expansion of graphs  $G_{SEQ}$  and  $G_{REQ}$  corresponding not only to degenerate but also to nondegenerate processes. In addition, all equations constructed from expanded graphs  $G_{ETOP}$  and  $G_E$  must be checked for degeneracy even if no DTs can, in principle, correspond to initial graphs  $G_{TOP}$  or  $G$ .

A specific feature of the lower hierarchy level is that explicit consideration of substituents may be required for generation of all theoretically possible multiply degenerate interconversions of a given type. It is easily seen (in contrast, *e.g.*, to the [3,3]-shift in bullvalene in Figure 2a)

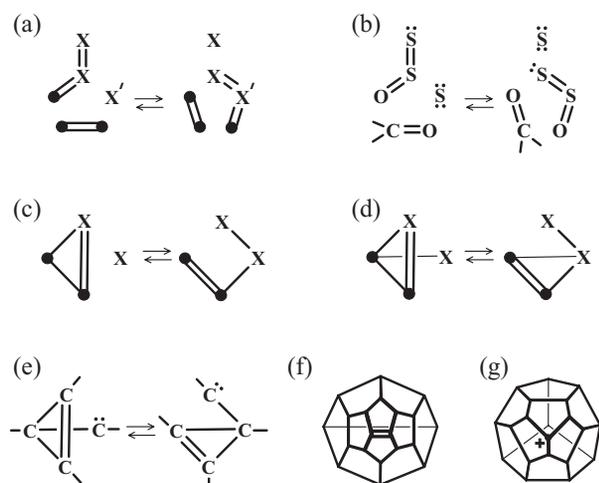


Figure 7. Examples illustrating the (a,b) possible inconsistency between the types of reaction centers in a degenerate rearrangement; (c–e) appearance of degeneracy in structural and skeletal equations obtained from a nondegenerate symbolic equation; and (f,g) necessary participation of substituents (all hydrogen atoms) in hypothetical multiply degenerate [1,3]-shifts in dodecahedrene and [1,2]-shifts in the dodecahedryl cation.

that the multiplicity of degeneracy for the [1,2]-shift in the ethyl cation in Figure 2g actually depends on the nature of substituents. For example, only two interconverting structures are involved in the degenerate [1,2]-sigmatropic shift of protonated tetramethylethylene  $\text{Me}_2\text{CH}-\text{C}^+\text{Me}_2$ . In our opinion, the most interesting "substituent-dependent" multiply degenerate transformations are those of dodecahedrene (see Figure 7f) and dodecahedryl cation (Figure 7g); apparently, these processes have not been observed yet.<sup>61</sup> The upper bounds for the numbers of vertices in reaction graphs of these rearrangements are  $30 \cdot 18!/2 \approx 9.6 \cdot 10^{16}$  and  $20 \cdot 19!/3 \approx 8.1 \cdot 10^{17}$ , respectively. Note that, in both cases, no extremely high degeneracy can be observed in the corresponding skeletal, structural, reaction, and symbolic equations. For example, each elementary stage of the hypothetical rearrangement in Figure 7g is described by the symbolic equation of an "ordinarily" degenerate  $[1^+(1,1)]$ -substitution process.

### COMPUTER-AIDED DEGENERACY DESIGN EXEMPLIFIED BY SEARCH FOR NEW [3,3]-SIGMATROPIC REARRANGEMENTS

The above methodology has been implemented in a computer program only for the least detailed level of the degeneracy hierarchy. Specifically, the just mentioned ARGENT-1 program enables its users to construct graphs  $G_{\text{TOP}} \Rightarrow G_{\text{SEQ}}$  or  $G_{\text{REQ}}$  and the corresponding (symbolic and reaction) equations starting from any graph in the chain  $G \Rightarrow G_{\text{TOP}} \Rightarrow G_{\text{SEQ}} \Rightarrow G_{\text{REQ}}$ . During generation, one can apply various constraints associated with the choice of labels, their allowed numbers, presence or absence of preselected fragments, required symmetry, *etc.*, and also

degeneracy. The use of the ARGENT-1 software in systematic searching for new types of nondegenerate as well as degenerate reactions was exemplified in Ref. 62 and also in our new paper being currently prepared for publication.<sup>63</sup> Here, we demonstrate the potential of the program for the solution of problems associated with the Cope rearrangement, which is, in fact, the best known degenerate transformation.

First, we discuss heteroderivatives of the hexadiene-1,5 rearrangement in Figure 1b or, in other words, graphs  $G_{\text{REQ}}$  generated starting from a given  $G_{\text{SEQ}}$ . If the number of heteroatoms, which are represented by nitrogen and oxygen atoms, is limited to four, the ARGENT-1 program produces 18 reaction equations, shown in Figures 8a–r. Naturally, this list includes equations that represent the aforementioned oxaderivatives of the Cope reaction<sup>43,44</sup> (see Figures 8d,c) and also the unusual rearrangement of diacyl peroxides<sup>64</sup> in Figure 8o. Among 9 possible azaderivatives of the Cope rearrangement (see Figures 8a,b, i–n,r), the first two seem to be actually implemented; degenerate [3,3]-shifts in tetraphenyl-substituted 2-azahexadiene<sup>65</sup> and 2,4-diazahexadiene<sup>66</sup> may serve as typical examples. Thus, in contrast to numerous nondegenerate [3,3]-sigmatropic shifts (*e.g.*, see reviews<sup>67,68</sup>), only a few degenerate analogs of the Cope reaction have apparently been investigated, and examples in Figures 8a–d,o are probably among the best known ones.

Of more interest are the modes of the hetero-Cope rearrangement that have not yet been experimentally implemented but seem quite likely. They constitute almost one half of all results in Figure 8. Above all, we should mention the theoretically possible [3,3]-shifts in allyl derivatives of compounds belonging to well-known classes: allyl nitrite (see Figure 8g), allylamidines (Figure 8j), and allyltriazenes (Figure 8m). A highly promising group of "candidates" for the search for new degenerate rearrangements consists of azomethines, which can be obtained from amines (or, with more difficulty, from amides) and carbonyl compounds. The possible examples are degenerate transformations of Schiff bases in Figures 8a,e (corresponding to butene-3-ylamines<sup>65</sup> and acetoxymethylamines, respectively) and rearrangements of condensation products synthesized from *O,N*-diacylhydroxylamines or *N,N'*-diacylhydrazines (see Figures 8p,q,r). Surely, one can also devise more specific synthetic routes to other rearranging structures of Figure 8. For example, 1,3,4,6-tetraazahexadienes-1,5 in Figure 8r can presumably be obtained from imidoyl chlorides and *N,N'*-disubstituted hydrazines or by coupling of appropriate amidine species.

The second example of the use of the ARGENT-1 software refers to a more complicated task, which (in contrast to the generation problem for heteroderivatives) cannot be manually solved. This problem involves the search for such structural analogs of the Cope rearrangement that contain atoms with a changed valence (see

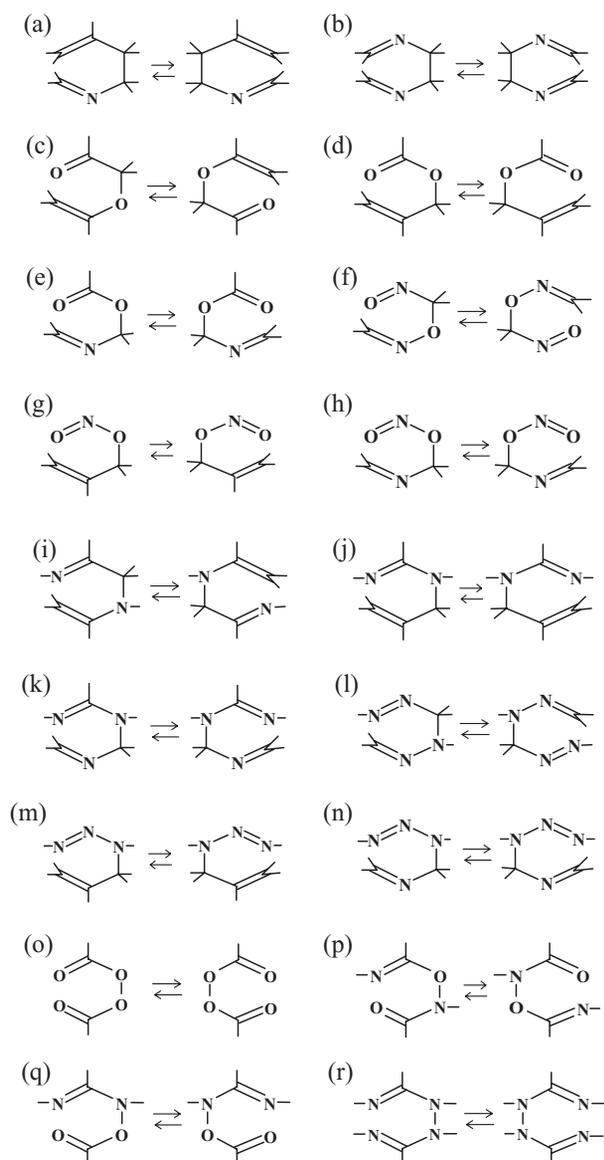


Figure 8. Heteroderivatives of the Cope rearrangement: reaction equations characterizing degenerate [3,3]-shifts with 1 to 4 nitrogen and oxygen atoms.

Figures 3e,f); bonds with multiplicity changes exceeding unity (Figures 3g,h); or signs of charge or an unpaired electron (Figure 3h). Within the framework of the above methodology, the solution to this problem is reduced to generation of graphs  $G_{SEQ}$  starting from the unsigned graph  $G_{TOP}$  and nine possible signed graphs  $G_{TOP}$  containing label pairs "+" and "(-)," "-" and "(-)," or "•" and "(•)". Evidently, the construction of graphs  $G_{TOP}$  is reduced to assignments of sign labels to two vertices of the preselected unsigned graph  $G$ , which is a six-membered cycle. To reduce the number of produced results, we restrict generation to symbolic equations corresponding to sigmatropic shifts in uncharged (74) and cationic (156) species; the remaining equations of anionic and radicalic processes can easily be obtained from cationic ones by substitution of "+" by the "-" or "•" symbol.

In addition, to reduce the number of unfeasible results to a minimum, some constraints are used: all bonds may change their multiplicities only by 1 or 2 units; all centers may have valence numbers only from 1 to 4; and all valence changes may assume only special values.<sup>69</sup> Making allowance for these constraints, we obtained 32 symbolic equations, among which 10 equations (see Figures 9a–j) representing [3,3]-shifts were manually selected. Note that the left-hand and right-hand parts of equations in Figure 9 (and also in Figure 10) are in no case associated with the possible geometries of the rearranging species. That is why angular forms of the "propargyl-like" and similar groups (just like in Figure 3g,h) are used only to conserve uniformity in representation of various [3,3]-shifts.

Among symbolic equations in Figure 9, only the first three represent rearrangements of uncharged species. The processes corresponding to the equations in Figures 9a,c have already been discussed (see Figures 3e,g). Among DTs that have not yet been implemented, we should mention the possible [3,3]-sigmatropic shifts in sulfonium, phosphonium, and similar ylides (*cf.* reaction equation in Figure 3f), as well as unusual rearrangements of azodicarbonitrile  $NC=N=CN$  and propargylenecyanamide  $HC\equiv C-CH=N-C\equiv N$ , which correspond to the equation in Figure 9c. Another topological type of degeneracy is determined by the symbolic equation in Figure 9b; the corresponding topological subtypes express, for example, the hypothetical transformations of propargyllallene  $HC\equiv C-CH_2-CH=C=CH_2$  and cyanoxyisocyanate  $N\equiv C-O-N=C=O$ .

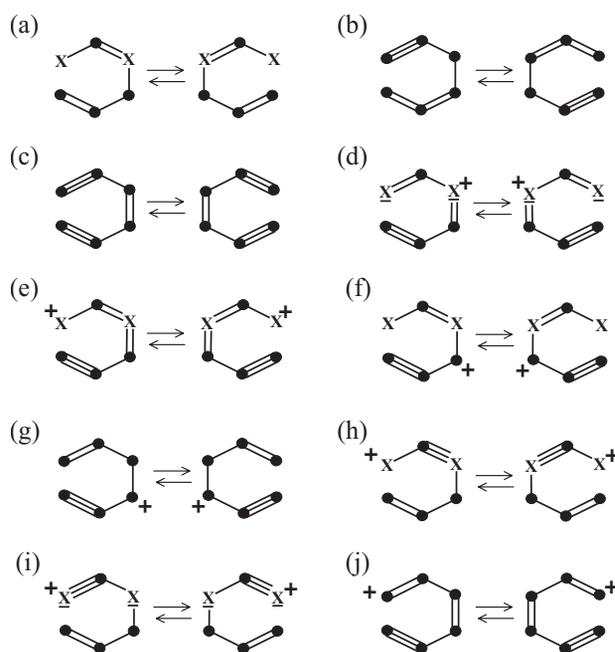


Figure 9. Structural analogs of the Cope rearrangement: symbolic equations characterizing degenerate [3,3]-sigmatropic rearrangements of (a–c) uncharged and (d–j) positively charged species.

Turning to [3,3]-shifts in charged systems, we must first note that acid- and base-catalyzed Cope reactions are well-known; the corresponding processes are often referred to as charge-induced or charge-accelerated ones, see reviews.<sup>68,70</sup> However, for the great majority of such transformations, formal analysis shows that the rearrangement step in the charged particle is either a substitution process or a nondegenerate conversion. As a result, the following problem arises: which types of positively charged (negatively charged or even radicalic) structures can, in principle, undergo "actually" degenerate [3,3]-shifts? Symbolic equations of Figures 9d–j represent the generalized solution to this problem; conversion of these equations (corresponding to topological types of degeneracy) into reaction equations was performed by means of the ARGENT-1 program, and it revealed many new topological subtypes of degeneracy.

Only some of the numerous results thus obtained are considered here together with probable synthetic routes to the rearranging species. However, it should be noted that the expected reliability of the predictions is not too high in this case, because cationic (anionic or radicalic) particles can undergo various competing reactions and the choice of the actual direction must be strongly dependent both on the nature of substituents and on the reaction conditions. That is why the proposed synthetic methods are represented in Figures 10a–j in a detailed form, *i.e.*, with all substituents and functional groups explicitly shown; for brevity, the deprotonating reagents and groups removed as anions are denoted by B<sup>-</sup> and Z, respectively. Another important feature of the rearrangements in charged (and radicalic) systems is the fact that they are often described by several resonance structures, and different structures always correspond to nonidentical symbolic equations.

More specifically, consideration of resonance makes it possible to partition seven degenerate [3,3]-shifts into three groups (see Figures 9d,e,g; 9h,i; and 9f,j), as demonstrated below, first for cationic rearrangements:

Figure 10a. – Interaction of allyl ketones (R = Alk) with monosubstituted acetylenes and subsequent transformation of the OH group into removable group Z (*e.g.*, into the tosyloxy group under the action of TsCl) provides a highly attractive route to a cation that can undergo the [3,3]-shift shown in Figure 9g. Other "Cope-like" bond redistributions are impossible in this case, because an alternative degenerate transformation of the mesomeric hexatrien-1,2,5-yl cation must be considered as an intramolecular [1<sup>+</sup>+(3,1)]-substitution but not as a [3,3]-sigmatropic shift.

Figure 10b. – A somewhat similar approach can result in the bis-aza analog of the cation in Figure 10a; it involves the interaction of *N,N'*-disubstituted amidines with propargyl aldehyde and subsequent transformation of the OH group into group Z. Note that the structure of the resultant cation (at R' = R'') must not depend on the

actual amidine nitrogen atom that initially attacks the carbonyl group. It can be also seen that each of the three resonance structures of the cation gives rise to its own corresponding symbolic equation (see Figures 9g,d,e).

Figure 10c. – Nitrogen-containing cations associated with rearrangements in Figures 9h,i can presumably be obtained by alkylation of substituted allylcarbodiimides or allylcyanamides. Naturally, to ensure degeneracy, both alkyl groups must be identical (R' = R'').

Figure 10d. – From the formal standpoint, the aforementioned cation in Figure 3h is the product of "anti-Markovnikoff" protonation of hexene-3-diene-1,5. Another possible route to this cation is the solvolytic removal of an appropriate (*e.g.*, tosyloxy: Z = TsO) group. Since the positive charge is localized in this case, the degenerate [3,3]-shift can be associated only with the symbolic equation in Figure 9j.

Figure 10e. – In the last example, the positive charge in the cation must be highly delocalized. This cation can be probably obtained by addition of ethynylcarbene to a substituted thioacetal and subsequent removal of an appropriate anion (*e.g.*, at Z = Cl). In this case, the [3,3]-sigmatropic shifts are theoretically possible for two out of five resonance structures; the corresponding symbolic equations can be found in Figures 9f,j.

The symbolic equations produced by the ARGENT-1 program can be also applied to the search for anionic analogs of the Cope rearrangement. In this case, naturally, one should replace the "+" sign by "-" in all equations of Figures 9d–j. Here are some proposed synthetic methods relating to anionic [3,3]-shifts; minimum comments are supplied:

Figure 10f. – Deprotonation (by means of base B<sup>-</sup>) of the imino group in propargyleneamidines, which presumably can be obtained as products of alkylethynylketone–amidine condensation, R ≠ H.

Figure 10g. – Interaction of mercaptals with substituted ethynylcarbene (or with propargyl chloride) and subsequent deprotonation of the reaction product. Just as in the previous case, it is assumed that R ≠ H.

Figure 10h. – Deprotonation of allylamidines (with all substituents R' = R'' and the allyl at nitrogen atoms) can, from the formal standpoint, be succeeded by a degenerate [3,3]-shift in the resultant anion. However, this anion is expected to be extremely unstable, because none of its resonance structures are energetically favorable.

Figure 10i. – In contrast, deprotonation of allylcyanamide must result in an essentially stable anion, which, in principle, can undergo the "Cope-like" rearrangement. Since the initial cyanamide seems to be quite available, this model is the most attractive one for experimental investigations.

Figure 10j. – The last synthetic possibility considered here includes addition of ethynylcarbene to substituted

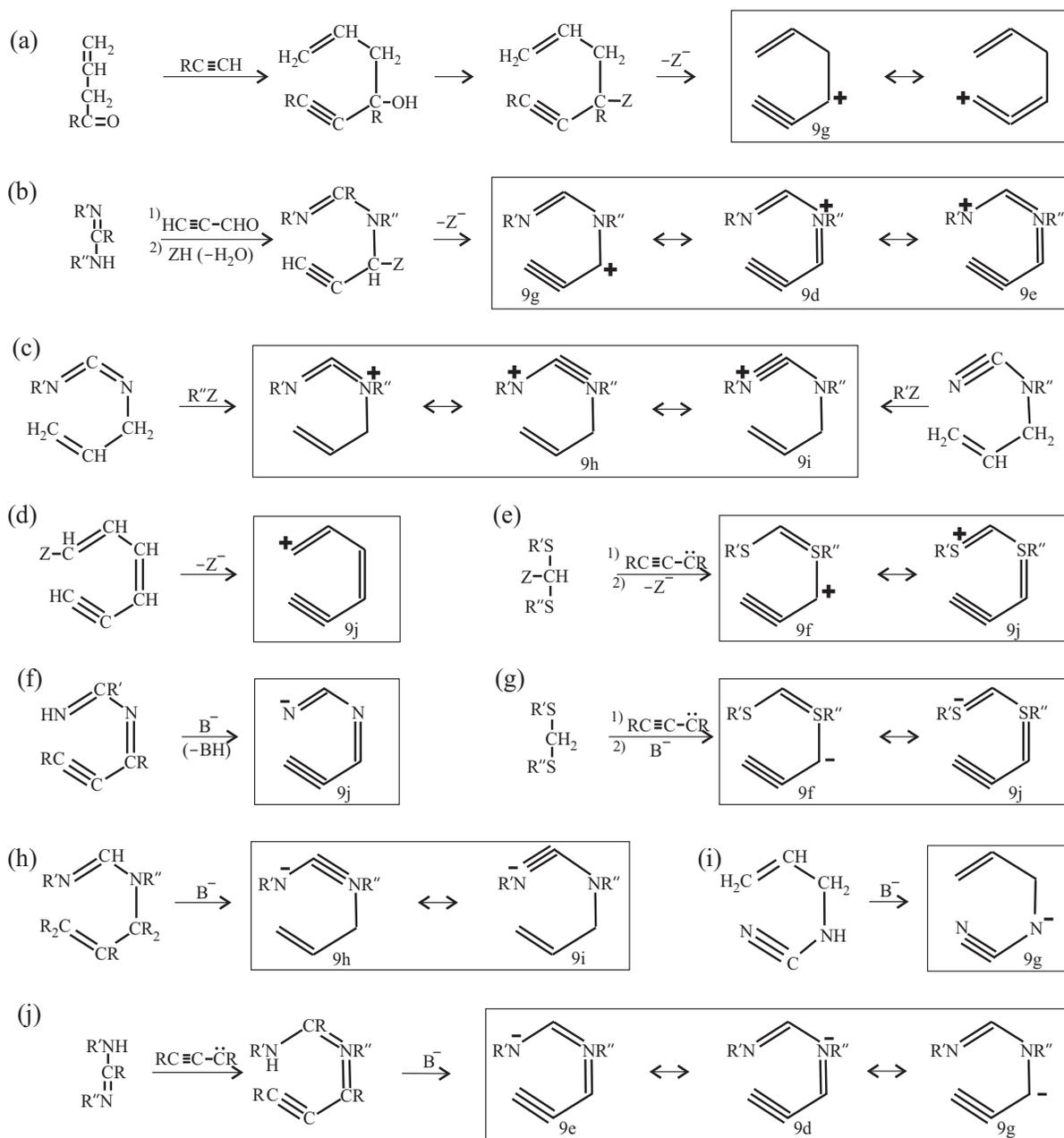


Figure 10. Possible synthetic approaches to (a–e) cationoid and (f–j) anionoid species that, in principle, can undergo degenerate [3,3]-shifts. References to the symbolic equations of Figure 9 are given near the corresponding resonance structures of the cations or anions.

amidines ( $\text{R}' = \text{R}''$ ). Deprotonation of the hypothetical reaction product must result in a sufficiently stable (due to the high degree of charge delocalization) anion, whose structure allows as many as three different degenerate [3,3]-shifts, namely, the anionic analogs of those represented in Figures 9d,e,g.

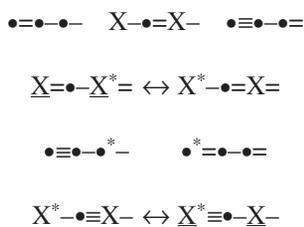
## CONCLUDING REMARKS

The systematic generation of symbolic and reaction equations just considered is not the only possible application of our ARGENT-1 software to theoretical investigation of degeneracy. There are several other related problems

that can be successfully solved using the computer-generated lists of hypothetical DTs. One of such problems involves analysis of "degeneracy synthons," *i.e.*, multi-atomic fragments that are converted into themselves by DTs of a given type.

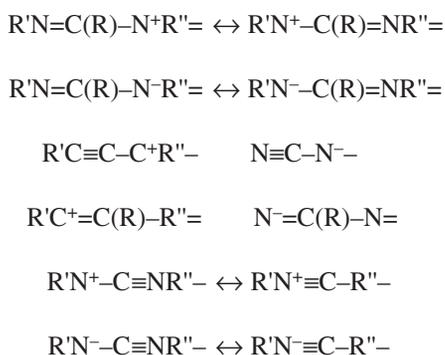
As a simple example, here we discuss the set of three-centered synthons that were extracted from the symbolic equations in Figures 9a–j. From the standpoint of the Formal-Logical Approach, the considered synthons represent a special case of reaction fragments<sup>51</sup> (note that equations involving such fragments characterize the so-called interconversion types of reaction systems, which represent an important level in the general classification of organic

transformations<sup>53</sup>). More specifically, inversion fragments of the third order ( $I_3$ ) are three-centered groups with the "external" bond (of any multiplicity) migrating from one end of this group to the other. Considering the equations in Figures 9a–j, one can easily extract the fragments that are converted into themselves as a result of the [3,3]-shift:<sup>71</sup>



Please note that two out of seven fragments are represented by resonance hybrids; in the last four fragments, the asterisk is a generalized symbol of charge or unpaired electron.

Evidently, one can easily obtain representations of each  $I_3$  fragment if the designations of centers are substituted by symbols of chemical elements of suitable valences. Possible examples of representations for three "unsigned" fragments are the allyl group  $CR'_2=C(R)-CR''_2-$ , the carbalkoxy group  $O=C(R)-O-$ , the azide group  $:N=N=N-$ , the thiolsulfuranyl group  $R'S-C(R)=SR''-$ , the propargylene group  $CR'\equiv C-R''=$ , and the cyanoimine group  $N\equiv C-N=$ . All these groups have been mentioned in the above examples. For the remaining four fragments, below we supply one example of a cation and one example of an anion capable of converting into themselves as a result of a [3,3]-shift:



Note that representations of "degenerate" fragments can be successfully applied not only to [3,3]-shifts but also to the search for other types of DTs, such as [1,3]- or [3,5]-sigmatropic rearrangements, processes of substitution, metathesis, etc.

In conclusion, we should outline yet another problem, which is associated both with fundamental characteristics of degeneracy and with the graphs considered above. In fact, for all the studied degenerate transformations (and for almost all processes mentioned in this paper), each

of the graphs  $G_{SEQ}$ ,  $G_{REQ}$ ,  $G_{STR}$ , and  $G_{SKE}$  can be characterized by two automorphism groups, which are necessarily nonidentical. In all such cases, the permutations that ensure degeneracy (above, we have provided an example for the Cope rearrangement) belong only to the larger of the groups and always represent symmetries of graphs  $G_{TOP}$  or  $G_{ETOP}$ . We refer to such DTs as regular. However, our investigations<sup>72</sup> proved also the existence of semiregular and irregular DTs, for which the permutations in question do not belong to the group of graphs  $G_{SEQ}$ ,  $G_{REQ}$ ,  $G_{STR}$ , or  $G_{SKE}$ , although they can still be symmetries of the topology graph. Examples of degenerate rearrangements pertaining to these essentially new classes are the chemically unfeasible processes represented by the symbolic, reaction, structural, and skeletal equations in Figures 7a,b,d,e. Formalization of the notion of regularity in terms of the permutation group theory, investigation of the relationship between the properties of regularity and self-inverseness, and also new results generated by the ARGENT-1 program will be described in the next paper of this series. In that paper, we are also planning to discuss several semiregularly and irregularly degenerate multistage transformations that seem to be promising targets for experimental implementation.

Readers interested in supplementary information on the current state of the ARGENT-1 project or willing to install our freeware program on their computer are welcome to contact the authors of this paper.

## REFERENCES

1. In certain cases, an alternative to degenerate rearrangement can consist of the existence of stable intermediate species with nonclassical bonds (such as the bridged cation in the case of Figure 1a, the symmetrical intermediate in the case of Figure 1b, or the structure with equalized CO and OH bonds in the case of Figure 1c). In fact, the choice between these two possibilities is not always trivial and may require additional experiments.
2. W. H. Pirkle, H. Seto, and W. V. Turner, *J. Am. Chem. Soc.* **92** (1970) 6984–6985.
3. J. S. McKennis, L. Brener, J. S. Ward, and R. Pettit, *J. Am. Chem. Soc.* **93** (1971) 4957–4958.
4. A. T. Balaban, *Reaction Graphs*, in: D. Bonchev and O. Mekenyan (Eds.), *Graph Theoretical Approaches to Chemical Reactivity*, Kluwer Acad. Publ., Dordrecht, 1994, pp. 137–180.
5. A. T. Balaban, D. Fărcașiu, and R. Bănică, *Rev. Roum. Chim.* **11** (1966) 1205–1227.
6. A. T. Balaban, *Rev. Roum. Chim.* **22** (1977) 243–255.
7. M. Randić, *Int. J. Quant. Chem., Quant. Chem. Symp.* **14** (1980) 557–577.
8. M. Randić, D. O. Oakland, and D. J. Klein, *J. Comput. Chem.* **7** (1986) 35–54.
9. B. M. Gimarc and A. R. Brant, *J. Chem. Inf. Comput. Sci.* **34** (1994) 1167–1173.
10. T. P. Živković, *Croat. Chem. Acta* **69** (1996) 215–222.

11. G. A. Jones and E. K. Lloyd, *Automorphism Groups of Some Chemical Graphs*, in: R. B. King (Ed.), *Chemical Applications of Topology and Graph Theory*, Stud. Phys. Theor. Chem., Vol. 28, Elsevier, Amsterdam, 1983, pp. 252–267.
12. M. H. Klin, S. S. Tratch, and N. S. Zefirov, *J. Math. Chem.* **7** (1991) 135–151.
13. M. H. Klin and N. S. Zefirov, *MATCH Commun. Math. Comput. Chem.* **26** (1991) 171–190.
14. E. K. Lloyd and G. A. Jones, *Acta Appl. Math.* **52** (1998) 121–147.
15. W. G. Klemperer, *J. Chem. Phys.* **56** (1972) 5478–5489.
16. W. Hässelbarth and E. Ruch, *Theor. Chim. Acta* **29** (1973) 259–267.
17. D. J. Klein and A. H. Cowley, *J. Am. Chem. Soc.* **97** (1975) 1633–1640.
18. J. G. Nourse, *J. Am. Chem. Soc.* **99** (1977) 2063–2069.
19. J. Brocas, M. Gielen, and R. Willem, *The Permutational Approach to Dynamic Stereochemistry*, McGraw-Hill, New York, 1983, 713 pp.
20. I. Ugi, J. Dugundji, R. Kopp, and D. Marquarding, *Perspectives in Theoretical Stereochemistry*, Lecture Notes in Chemistry, Vol. 36, Springer-Verlag, Berlin, 1984, 247 pp.
21. W. G. Klemperer, *J. Am. Chem. Soc.* **94** (1972) 8360–8371.
22. J. G. Nourse, *J. Am. Chem. Soc.* **102** (1980) 4883–4889.
23. J. Brocas and R. Willem, *J. Am. Chem. Soc.* **105** (1983) 2217–2220.
24. Just as in the case of nondegenerate reactions, many degenerate processes invented by a chemist or designed by a computer cannot be experimentally accomplished. Estimation of the feasibility of new DTs may require special investigations such as calculation of thermodynamic and kinetic parameters, modeling of the stereochemistry for transition states, consideration of the possible mechanisms for multistage processes, *etc.*
25. A. T. Balaban, *Rev. Roum. Chim.* **17** (1972) 883–896.
26. D. E. Lushnikov, E. V. Gordeeva, and N. S. Zefirov, *Tetrahedron Comput. Methodol.* **3B** (1990) 417–427.
27. E. V. Babaev and N. S. Zefirov, *J. Math. Chem.* **11** (1992) 65–78.
28. In both cases, research was stimulated by the unique rearrangement of bullvalene (Figure 2a), predicted by Doering and Roth in 1963<sup>29</sup> and experimentally implemented by Schröder in 1964.<sup>30</sup> In this rearrangement, just like in the process shown in Figure 1e, multiple repetition of the [3,3]-sigmatropic shift (the Cope reaction) leads to "averaging" of all carbons and all hydrogens in the molecule. However, the number of structures rearranging into each other is much greater in the case of bullvalene (1,209,600) than in the case of hypostrophene (five; see the graph in Figure 1f).
29. W. v. E. Doering and W. R. Roth, *Tetrahedron* **19** (1963) 715–737.
30. G. Schröder, *Chem. Ber.* **97** (1964) 3140–3149.
31. A. T. Balaban, *Rev. Roum. Chim.* **15** (1970) 463–485.
32. L. A. Paquette, *J. Am. Chem. Soc.* **93** (1971) 7110–7112.
33. M. Baudler, H. Ternberger, W. Faber, and J. Hahn, *Z. Naturforsch.* **34b** (1979) 1690–1697.
34. L. A. Paquette and J. C. Stowell, *J. Am. Chem. Soc.* **93** (1971) 2459–2463.
35. N. S. Zefirov and S. S. Tratch, *Zh. Org. Khim.* **12** (1976) 7–18 [in Russian].
36. J. P. Chesick, *J. Am. Chem. Soc.* **85** (1963) 2720–2723.
37. V. A. Mironov, E. V. Sobolev, and A. N. Elizarova, *Tetrahedron* **19** (1963) 1939–1958.
38. J. C. Barborak, J. Daub, D. M. Follweiler, and P. v. R. Schleyer, *J. Am. Chem. Soc.* **91** (1969) 7760–7761.
39. T. D. Bouman, C. D. Duncan, and C. Trindle, *Int. J. Quant. Chem.* **11** (1977) 399–413.
40. T. D. Bouman and C. Trindle, *Theor. Chim. Acta* **37** (1975) 217–231.
41. The numbers of vertices in reaction graphs for multiply degenerate DTs in Figures 2a–h are 1,209,600; 10; 1680; 6; 3; 3600; 20; and 6, respectively. In cases where DTs of two or several types proceed simultaneously, the number of interconverted structures can sharply increase. For example, for the barbaralyl cation in Figure 2h and the norbornyl cation in Figure 1a, the numbers of vertices in their reaction graphs are equal to  $2 \cdot 9! = 725,760$  and  $2 \cdot 7! \cdot 11! \approx 4 \cdot 10^{11}$ , respectively.<sup>40,42</sup>
42. C. K. Johnson and C. J. Collins, *J. Am. Chem. Soc.* **96** (1974) 2514–2523.
43. E. S. Lewis, J. T. Hill, and E. R. Newman, *J. Am. Chem. Soc.* **90** (1968) 662–668.
44. A thorough search unexpectedly showed that, in contrast to simple "1,4-dioxa-Cope" reactions, at least two such [3,3]-shifts have been observed in complex polycyclic systems. Rearrangement of the  $\alpha$ -tocopherol spirodimer (see M. Chauhan, F. M. Dean, K. Hindley, and M. Robinson, *Chem. Commun.* (1971) 1141–1143) is an impressive example; this multistep process is acid-catalyzed and probably proceeds via the intermediate phenoxylium ion.
45. J. J. Gajewski and C. N. Shih, *J. Am. Chem. Soc.* **94** (1972) 1675–1682.
46. A. Gagneux, S. Winstein, and W. G. Young, *J. Am. Chem. Soc.* **82** (1960) 5956–5957.
47. R. Curci, V. Lucchini, G. Modena, P. J. Kocienski, and G. Ciabattini, *J. Org. Chem.* **38** (1973) 3149–3153.
48. N. S. Zefirov and S. S. Tratch, *Chem. Scr.* **15** (1980) 4–12.
49. R. R. Jones and R. G. Bergman, *J. Am. Chem. Soc.* **94** (1972) 660–661.
50. S. S. Tratch and N. S. Zefirov, *Zh. Org. Khim.* **18** (1982) 1561–1583 [in Russian].
51. N. S. Zefirov and S. S. Tratch, *Zh. Org. Khim.* **20** (1984) 1121–1142 [in Russian].
52. S. S. Tratch and N. S. Zefirov, *J. Chem. Inf. Comput. Sci.* **38** (1998) 331–348.
53. S. S. Tratch and N. S. Zefirov, *J. Chem. Inf. Comput. Sci.* **38** (1998) 349–366.
54. N. S. Zefirov, S. S. Tratch, and M. S. Molchanova, *MATCH Commun. Math. Comput. Chem.* **46** (2002) 253–273.
55. S. S. Tratch, M. S. Molchanova, and N. S. Zefirov, *MATCH Commun. Math. Comput. Chem.* **46** (2002) 275–301.
56. W. v. E. Doering and J. C. Gilbert, *Tetrahedron Suppl.* (1966) 397–414.
57. In the general case, classification of atoms into reaction centers, structural centers, and substituents depends on the choice of resonance structures for reagents and products; the distinction of structural centers and substituents can also be nontrivial (see ref. 53 for details). Here is a simple example: carbon atoms of the phenyl group are not structural centers in most cases; the entire group  $C_6H_5$  is typically regarded as a univalent substituent of a special kind.

58. As far as we are aware, no attempts to elaborate a hierarchical classification of degenerate bond redistributions have been made. The very notion of type has been systematically applied to DTs in only one paper: in Ref. 13, all DTs characterized by the same reaction equation were attributed to the same type. It is important that designations of reactions often used by organic chemists ( $[i,j]$ -sigmatropic shift,  $[i,j]$ -migration in cations, *etc.*) should not serve as the basis for classification either: these designations are not universal, can be just as well applied to nondegenerate transformations, and sometimes fail to distinguish between completely dissimilar processes (see examples of [3,3]-shifts in Figures 3a–h and examples of [1,2]- and [1,3]-shifts in carbonium ions in Figures 6a–d).
59. R. E. Leone and P. v. R. Schleyer, *Angew. Chem., Int. Ed. Engl.* **9** (1970) 860–890.
60. The resulting transformation evidently represents a certain structural type of degeneracy, which is unusual in the sense that no topological degeneracy type corresponds to it. (Otherwise, the process characterized by the symbolic equation in Figure 7c would also be degenerate.) Thus, in the general case, the interrelationship between topological types/subtypes and structural types/subtypes of degeneracy is not as trivial as it may seem at first sight.
61. Similar degenerate [1,2]-shifts in the dodecahedryl radical and anion seem to be unknown as well. The possible stabilization of anionic analog of Figure 7g owing to the "rolling charge" effect was considered as far back as 40 years ago; see H. P. Schultz, *J. Org. Chem.* **30** (1965) 1361–1364.
62. M. S. Molchanova, S. S. Tratch, and N. S. Zefirov, *J. Phys. Org. Chem.* **16** (2003) 463–474.
63. M. S. Molchanova, S. S. Tratch, and N. S. Zefirov, unpublished paper.
64. M. J. Goldstein and H. A. Judson, *J. Am. Chem. Soc.* **92** (1970) 4119–4120.
65. D. Brombach and F. Vögtle, *Synthesis* (1977) 800–802.
66. F. Vögtle and E. Goldschmitt, *Angew. Chem., Int. Ed. Engl.* **13** (1974) 480–481.
67. T. L. Gilchrist and R. C. Storr, *Organic Reactions and Orbital Symmetry*, Cambridge Univ. Press, Cambridge, 1971, 271 pp.
68. S. Blechert, *Synthesis* (1989) 71–82.
69. For unsigned centers, the valence (or reaction number, in the more strict terminology)<sup>48,50</sup> is allowed to change here only by two units (see centers X in Figures 9a,f); the sum of valence changes over all atoms cannot exceed 4. In principle, charged centers can change their valence by one (such centers are denoted by • and •<sup>+</sup> if the charge corresponds to the lower valence of the center, or by  $\underline{X}$  and  $\underline{X}^+$  if the charge corresponds to its higher valence), by 3 (denoted by X and X<sup>+</sup> if the charge corresponds to the lower valence, or by  $\underline{Y}$  and  $\underline{Y}^+$  if the charge corresponds to the higher valence), or by 5 units. The program makes it possible to distinguish underlined (referred to as pseudospecific) centers from non-underlined ones. In the example under consideration, equations with centers whose valence is changed by 5 units and equations with centers  $\underline{Y}$  are not generated.
70. R. P. Lutz, *Chem. Rev.* **84** (1984) 205–247.
71. Strictly speaking, degeneracy in [3,3]-shifts is also possible if two different fragments are converted into each other. Among the symbolic equations in Figure 9, only the equation in Figure 9b contains such fragments. Possible examples are the propargyl and allenyl ( $R'C\equiv C-CR''_2-$  and  $-R'C=C=R''_2$ ), cyanamide and carbodiimide ( $N\equiv C-NR-$  and  $-N=C=NR$ ), or cyanate and isocyanate ( $N\equiv C-O-$  and  $-N=C=O$ ) groups.
72. S. S. Tratch, M. S. Molchanova, and N. S. Zefirov, in: *Third All-Russian Conference on Molecular Modeling*, Moscow, 2003, p. 123 [in Russian].

## SAŽETAK

### Novi pristup degeneriranoj interkonverziji organskih struktura: 1. Stupnjevi strukturne degeneracije i računalna potraga za degeneriranim interkonverzijama

Serge S. Tratch, Marina S. Molchanova i Nikolai S. Zefirov

Opća metodologija sustavne potrage za degeneriranim interkonverzijama mora se temeljiti na hijerarhiji degeneracija. Stupnjevi i podstupnjevi predložene hijerarhije predstavljeni su topološkim i strukturnim tipovima i podtipovima degeneracije, koji odgovaraju određenim jednadžbama i pridruženim grafovima. Formulirane su grafičke operacije koje omogućavaju učinkovitu potragu za proizvoljnim tipom organskih reakcija i razmatrani su problemi specifični za provedbu degeneriranih transformacija. Raspravlja se o rezultatima potrage pomoću računala (koristeći softver ARGENT-1) za heteroderivate i strukturne analoge Copeove reakcije. Mnoge tako otkrivene pregradnje (posebno [3,3]-pomaci u kationima i anionima, koji dosada nisu razmatrani) mogle bi se eksperimentalno proučavati.