RING-TRANSFORMATION GRAPHS IN HETEROCYCLIC CHEMISTRY

Eugeny V. Babaev* and Nikolai S. Zefirov

Department of Chemistry, Moscow State University, Moscow 119899, USSR

A novel graph-theoretical and topological approach is suggested for the description and classification of a great number of *simple* ring-transformation reactions of heterocycles. The suggested *ring-bond-redistribution graphs* (which can be easily obtained from the known mechanisms of rearrangements or recyclizations) visualize the degree of similarity of different ring transformations. A simple type of structure of these labelled graphs and the clear chemical sense of the used labels permit us to classify a lot of known *simple* heterocyclic ring transformations in an hierarchic system, which can be used to predict unknown *types*, *classes* and *sorts* of these reactions.

INTRODUCTION

The methods of graph theory during the last ten years have become a useful instrument for analyzing various problems in theoretical and experimental chemistry (1-9), especially in such fields as isomer enumeration (2,4), structure and reactivity of conjugated π-systems (1-5), "magic" electron-counting rules in cluster chemistry (1,6,7) and QSAR-researches (8,9). Traditionally applications of graphs in chemistry are based on the description of molecular structure by the graph as a set of points and lines in simple accord with the intuitive chemical image of a molecule as consequence of atoms and bonds.

One of the important applications of graph theory, suggested more than 15 years ago (10), is the use of a graph as an image of a chemical reaction. In a series of papers, Zefirov and Tratch (11-15) (and later Fujita (16)) used graphs for the description of bond redistributions in the course of an organic reaction. If the initial and final molecules in a reaction are described as labelled multigraphs, then any reaction is characterized by an edge redistribution in the initial multigraph. The redistribution is represented by a symbolic equation in which only the bonds changing their order are presented. The relative disposition of these bonds is uniquely characterized by the topology identifying graph; the vertices and edges of this graph also correspond to reaction centres and to bonds changing their order in the course of the reaction, respectively. For example, one can transform the chemical equation of the Diels-Alder reaction (1) into the symbolic equation (2) and then to the corresponding topology identifier (Chart 1). This approach (the so-called "formal-logical approach" (10,13,14) or FLA) has been developed up to a rigorous mathematical level (15). The computer programs based on FLA have been elaborated (15,17); they permit to generate all possible symbolic equations (corresponding to a given topology identifier) and, hence, they are the powerful tools of search for new types of organic reactions (12).
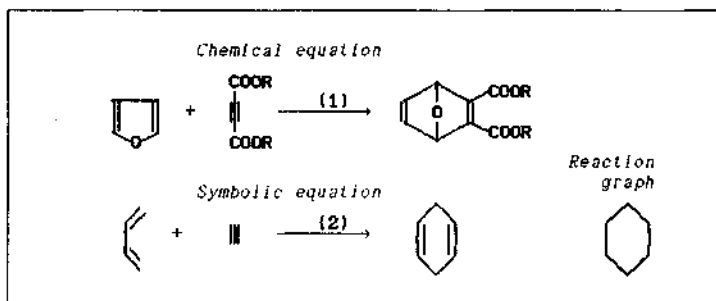

CHART 1.

Due to our interest both in graph theory (18) and in the chemistry of heterocycles (19,20) we present in this paper our graph-theory approach to one particular branch of organic reactions - ring transformations and/or rearrangements of heterocycles.

Since the classical work of Dimroth (21), this type of reactions received great attention from organic chemists and this area has been extensively studied and reviewed (see, for example, (22-28)). These reactions frequently lead to unexpected products, they often involve unusual mechanisms and sometimes they prove to be the only routes to the target heterocyclic structures. In recent times the examples of such rearrangements, including ring transformations and/or recyclizations, have been found for almost every heterocyclic system of any ring size and any type, number or distribution of heteroatoms (29). Indeed, the reactions presented on Charts 2-6 can be taken as good examples of such rearrangements. Many of them have special names and one can easily recall the well-known examples of the Dimroth rearrangement of azoles (21-23) (eq. (8)), azines (22,24) (eq. (29)) and fused azoloazines (29,24,20), Cornforth (22,23) (eq. (6)), Alberth (23,25) or Boulton-Katritzky (28,29) (eq. (3), (4)) rearrangements of five-membered heterocycles and the Kost-Sagitullin transformations of a pyridine nucleus to benzene (20,24) (eq. (15)). Other examples of "named" ring transformations are the well-known Hafner (24,26) and Zinke-Konig (24) reactions, as well as the processes of pyrrole, furan and thiophen interconversions (eq. (20)), generally accepted in the Russian literature as Yur'ev reactions (19).

However, in spite of the great experimental development of this area of heterocyclic chemistry, ring transformation reactions are still poorly classified. Because of the great variety of recyclization types, it is sometimes difficult to justify the real novelty of a claimed "new ring transformation" published in the literature. There are only a few works concerned with the general classification of ring transformations.

In his first classical review on ring transformations Van der Plas (22) classified the reactions by the size of starting and final heterocyclic rings. L'abbé (23) suggested a classification of monocyclic rearrangements for 5-membered heterocycles taking into account the number of side-chain atoms in the different recyclization processes. For instance, in the case of the Dimroth reaction (8), the size of the side chain corresponding to an amino group, is equal to one; for the Cornforth rearrangement (6), the side chain is a carbonyl group with size 2 and for the Boulton-Katritzky rearrangement (3) the size corresponding to an oxime group is equal to 3. Schwaika (25) examined a large number of azole ANRORC-reactions with hydrazine, and classified the resulting ring transformations by the size of the unchanged ring fragments in the starting heterocycle. The classification of pyrillium transformations (26) has been based on the distribution of the fragments of the starting pyrillium ring between the final cyclic structure and its side chains. As it was shown by Balaban in the only announced approach (26c), further rationalization of this classification could be performed on the basis of simple graph-theoretical considerations; the cited work remains the only application of the graph theory in the field of heterocyclic ring transformations.

Thus, there is no universal approach capable of generalizing the recyclization processes into a hierarchical system and, what is more important, to predict new unknown examples of ring transformations on the basis of such a classification. The goal of this paper is to suggest an approach for the solution of these problems. The main idea of our approach consists of constructing a special graph of ring-bonds redistribution for any ring transformation process according to the following analysis of the structure of these graphs.

Basic notions of the ring-bond-redistribution graph approach

Each reaction is characterized by the corresponding redistribution of bonds. Let us take, for instance, a rearrangement (3) of 1,2,4-oxadiazole oxime 1 to acylaminofurazane 2 (Chart 2; this is an example of the well-known Boulton-Katritzky rearrangement (23,28). First, we can describe the bond redistribution for reaction (3) by means of the FLA-formalism (Chart 2). The interconversion of the labelled multigraphs relating to the starting and final molecules (eq. (3a) on Chart 2) as well as the symbolic equation (3b) permits us to classify reaction (3) as an $((1,1) + (3,1))$-cyclodismutation (12,13). Being correct in principle, this simple classification, however, brings no information about the starting and final heterocyclic rings, because the corresponding chains between the atoms 1-4, 1-8 and 4-8, which determine the cyclic structure (see equation (3c)), are omitted in the FLA-formalism. A more detailed treatment needs the additional notion of the structural
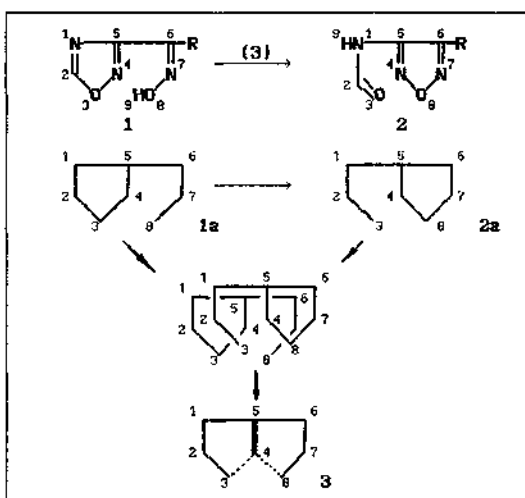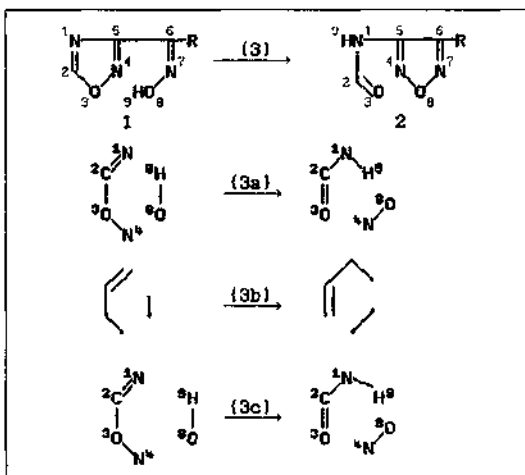
CHART 2.

equation (see eq. **3d**) (15), which reflects the cyclic structures of reactants and products (the additional structural centres are denoted by circles in eq. **3c**). The computer procedures for the constructive generation of structural equations have not been elaborated to date, but a rigourous mathematical model of their generation has been established (15).

The other specific feature leads to more serious difficulties in the applications of FLA to the problem considered. It is clear that the hydrogen atom (or another easily migrating function) is not really fixed at a single atom (e.g. at N^1 in structure **2** of Chart 2), but can migrate (to O^3 and N^4, see Chart 2) due to tautomeric interconversions. For this reason, the equations similar to (**3b**) or (**3d**) contain non-necessary information about the location of some atoms. Thus, there is a need for some alternative objects of a graph-theoretic nature applicable to the specific field of heterocyclic chemistry.

We shall use the term *heterocyclic ring transformation* (HRT) for any reaction containing the steps of ring opening and ring closure in any sequence. We call the heterocyclic ring transformation to be *simple* (SHRT)

1) if only one ring of the starting heterocycle is transformed into only one ring of the final molecule, 2) if there is no transient formation of cycles except the formation of the final cycle and 3) if there is no transient permutation of the atoms belonging either to the starting or the final heterocyclic rings. For instance, the rearrangements of Charts 2,3,5 are *simple*, however the HRT (12), (13) of Chart 4 are not because the above conditions 1), 2) or 3) are violated.

Now we can determine a new type of graphs responsible for HRT-reactions and concentrate all our attention exclusively on the ring atoms and bonds belonging either to the starting or the final cycle. Let us use the following algorithm to construct this new type of graph of *ring bonds redistribution* (RBR-graph) :

1. The molecular structures are replaced by unmarked monocyclic molecular graphs. The vertices and the edges of the graphs are only the atoms and the bonds, entering either in the final or in the starting ring. (If condensed rings are present in the starting system, then only the ring with minimal size is considered for transformation.)

2. The symbols of heteroatoms, as well as hydrogen atoms, multiple bonds, substituents (including condensed rings) must be omitted in the graph structure.

3. The atoms (and corresponding graph vertices) of the final structure are numbered in accordance with the numbers of the starting structure and the mechanism of HRT.

4. The graphs of the final and starting structures are "superposed" according to the matching vertices with identical numbers. The resulting ring transformation graph contains edges of different types, designated by solid lines (if the edge is present in both molecular graphs), dashed lines (if the edge is present only in one molecular graph) and bold lines (if the edge belongs to both cycles).

The corresponding molecular graphs 1a, 2a, their superposition and the resulting RBR-graph 3 are shown in Chart 2. We have to emphasize, that the resulting graphs 1a, 2a differ in principle from the graphs used in FLA (cf. eqs 3a-d) both in the vertices and the edges. Due to the algorithm requirements, graphs 1a and 2a contain vertices corresponding to both heterocyclic ring systems as well as to the oxime or acylamino groups, present as fragments in the final or starting rings. There are no vertices corresponding to hydrogen atoms (as well as atoms of substituents R in 1 and 2) since they do not enter into the ring structure. We also ignore the multiplicity of the bonds; for example, bonds 1-2 and 2-3, which change their orders. The resulting graph 3 consists only of edges, responsible for the ring bonds, and they are marked either as solid lines (corresponding to the unchanged bonds 3-2-1-5-6-7-8 of final and starting heterocyclic rings), or as dashed lines (corresponding to the broken and formed bonds 3-4 and 4-8) or as bold line (corresponding to the bond 4-5 mutual to both cycles). Obviously, the RBR-graphs of the direct and reverse reaction are identical. This type of graphs we call a G_1-graph.

Heterocyclic rearrangements. Following the above mentioned algorithm, one can draw the G_1-graphs for the transformations (4-9) of five-membered heterocycles (31a-d), Chart 3 (RBR-graphs 3,4a-9a of Charts 2,3). Comparing the structures of G_1-graph, it is easy to establish the similarity of different ring transformation reactions : the HRT's belong to the same class if their G_1-graphs are identical. As one can see from Chart 3, a very similar reactions (7) and (8) corresponding to thiadiazole-triazole rearrangement (31a,d) are nevertheless different and belong to separate classes, due to the differences in their G_1-graphs 7a, 8a. Reaction (7) is of the same class as the Cornforth rearrangement (31d) (6) due to the identity of the graphs 6a, 7a. In reaction (4) the starting molecule contains an isoxazole condensed to benzene ring. It is obvious that the benzene nucleus can be considered as a substituent and only the isoxazole ring (but not the 9-membered cycle of the bicyclic structure) undergoes transformation. The resulting G_1-graph 4a is identical with graph 3 of the Boulton-Katritzky rearrangement (28,31d).

A more detailed comparison of the G_1-graphs of Chart 3 gives another way to classify the HRT. If the dashed lines are substituted by solid ones in the G_1-graphs 4a-9a, one obtains the bicyclic graphs 4b-9b which are called as G_0-graphs (Chart 3). The structure of any G_0-graph contains a pair of annelated 5-membered cycles (responsible for the size of the starting and final cycles), and the graphs 4b-9b differ from each other only by the size of a bridge between the annelated cycles (the edges of this type are shown as bold lines in G_0-graphs as well as in G_1-graphs, see Chart 3). Simple presentations using G_0-graphs can now be used to demonstrate the global similarity of different heterocyclic rearrangements : the reactions are of the same type if their G_0-graphs are identical. For example, the transformations (3)-(9) are divided into three types ((3-5), (6-7), (8-9)), due to the difference in bridge size of the G_0-graphs. Inside the reaction type one can sub-classify
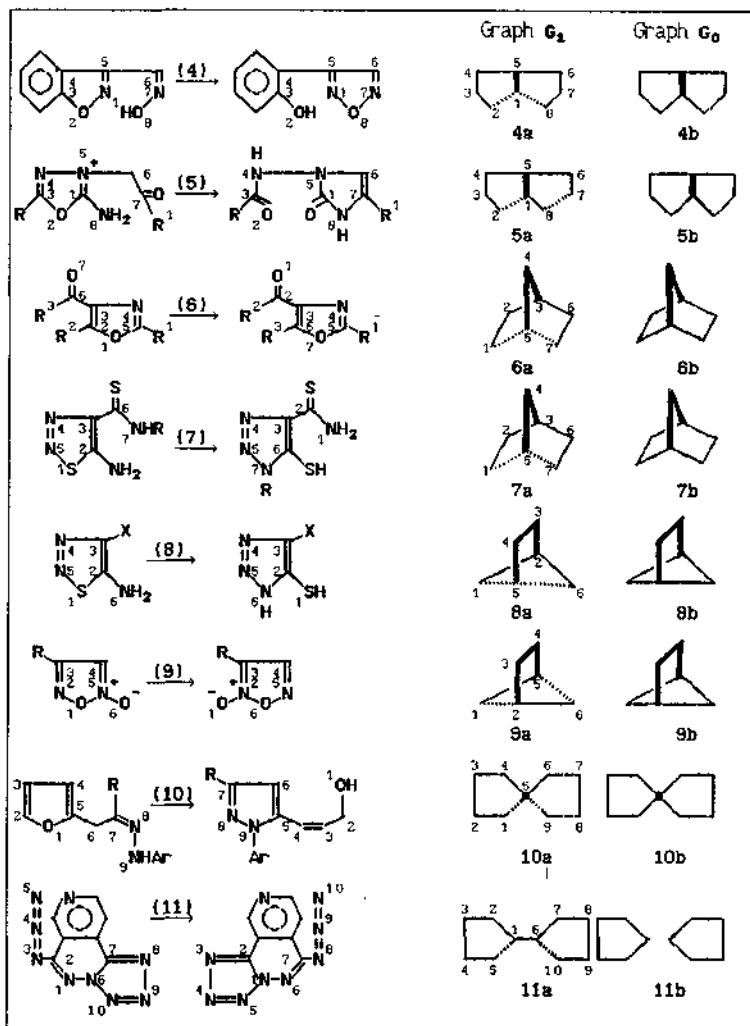
CHART 3.

certain reaction classes, responsible for the same G_1-graph. For example reactions (8), (9) are of the same type but belong to different classes.

In G_0-graphs 4b-9b the size of the bridge, common to both cycles, varies from two to four vertices. One can draw another kind of G_0-graph (see the graphs 10b,11b on Chart 3) as a pair of 5-membered rings, containing one or no common vertices. However, the chemical prototypes responsible for the G_0-graphs 10, 11 - comparatively rare reactions (10) and (11) - have also been found in the literature (31b,c).

Since the given examples of RBR-graphs for the reactions (3)-(11) prove to be of the bicyclic structure, it becomes important to question whether any ring transformation reaction can be represented exclusively by bicyclic G_0- and G_1-graphs? It is easy to prove that the bicyclic character of the graphs will take place only if the HRT is an SHRT (32). Fortunately, most of the known ring transformations of heterocycles prove to be SHRT (see the above-mentioned examples), and, hence, the bicyclic structure of G_0-graphs should be regarded empirically as a general case in heterocyclic chemistry. Some rare counter-examples of non-*simple*

reactions are shown on Chart 4. These HRT's are concerned with photochemical rearrangements occuring *via* intermediate cycle formation (equation (12)) or permutation of ring atoms as in reaction (13) ("no-side-chain" rearrangements (23)).
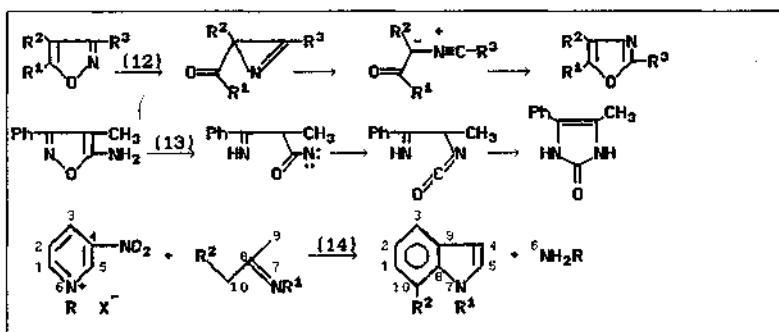

CHART 4.

Ring contraction/extention examples. One can use ring transformation graphs G_1 and G_0 to classify not only the above mentioned azoles transformations (Charts 2-3) but the rearrangements of heterocyclic systems of any ring size. Examples (15)-(19) shown of Chart 5 include reactions of extention, contraction and retention of ring size from 3- to 7-membered heterocycles (34). All the graphs of Chart 5, as in the previous examples, consist of a pair of annelated cycles and a bridge, whose size is, by definition, determined by the fragment, common to both heterocyclic rings. A new peculiarity is shown up in the G_0-graphs 18 and 19 where the bridges coincide with the smaller cycle of the annelated rings.
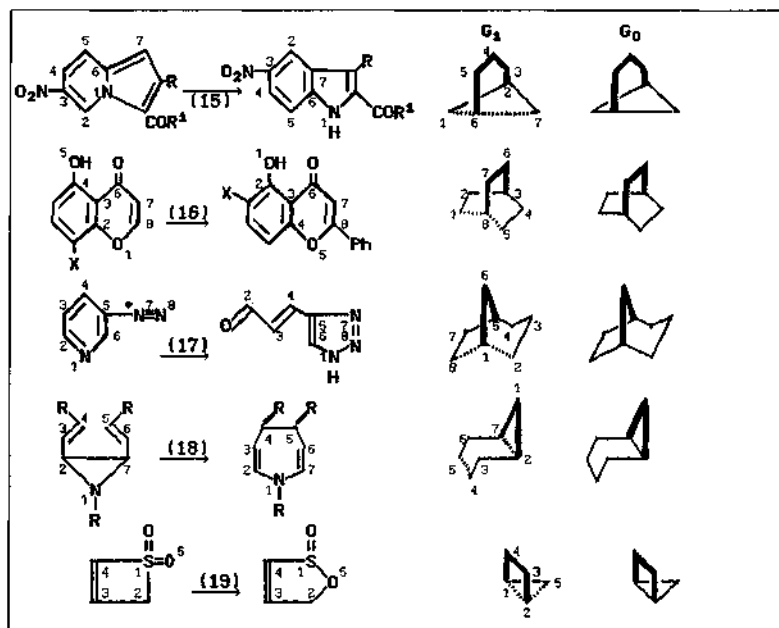

CHART 5.

Fragment including/eliminating reactions. The developed approach is wide enough to include not only rearrangements but also recyclizations in which reagents are included in the final ring and/or in which some fragments are eliminated from starting ring. There are many well-known examples of this kind; for instance, the reactions of heteroatom exchange, especially in pyrillium and furan chemistry (19,26,29) see, for example eqns (20), (21) of Chart 6. We have to mention that the terms HRT and SHRT are also applicable to this kind of reactions (reactions (20), (21) illustrate the SHRT, but reaction (14) (35) in Chart 4 is not the simple one). This new *kind* of HRT should be distinguished from the previously examined rearrangements. If A,B ... are the labels of heterocyclic ring systems and X_1, X_2 ... are the labels for any linear external fragments incorporated into the structure of a final or starting ring, we can roughly classify the *kind* of HRT by the number of independent fragments (A,B,X). For example all the rearrangements (3)-(11), (15)-(19) are of the A → B - kind. However, the reaction (24) (25) can be represented by the equation A + X→B and the reactions (20), (21) and (25) - by equation $A + X_1 \rightarrow B + X_2$. We also mention that equation A→B + X describes the reaction (22) (31d), which represents a ring transformation with elimination of a starting ring fragment. Obviously, the number of symbols A,B,X in the classification equation is not the reaction order but only the number of fragments related to the starting or final ring.

The algorithm for G_1-graph construction can be easily adapted to reactions of every kind (see the corresponding RBR-graphs in Chart 6). However, in cases of SHRT which are not rearrangements the resulting molecular graphs prove not to be connected graphs, because one must include in molecular graphs
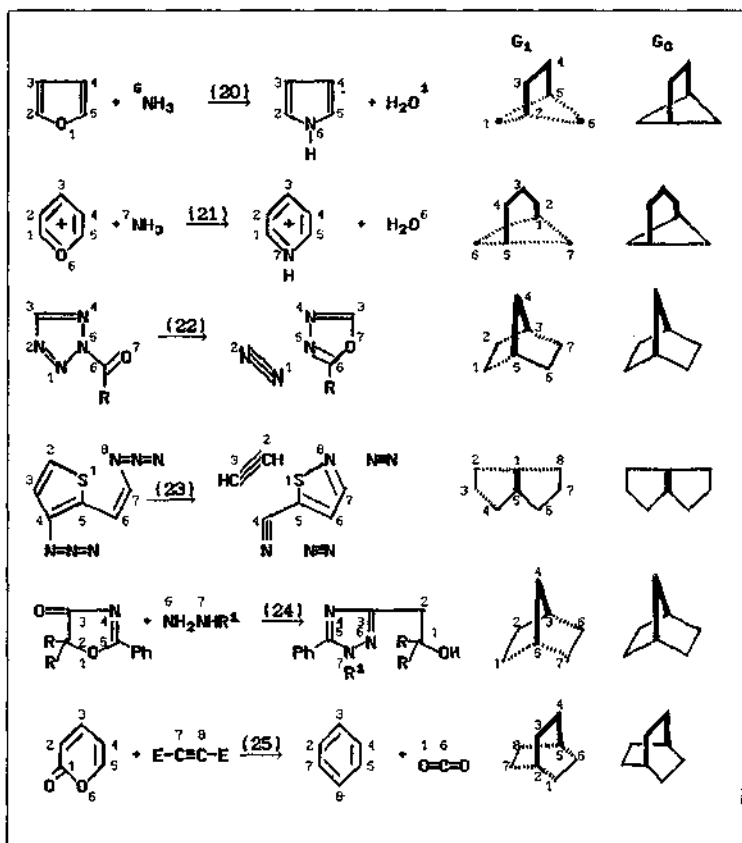

CHART 6.

the linear fragments of reagents (or eliminated groups). In fact, the only difference between the G_1-graphs of HRT in Chart 6 and the graphs of rearrangements in Charts 2,3 and 5 is the number of dashed lines corresponding to the formed or broken bonds in heterocyclic rings. All G_0-graphs obtained from G_1-graphs possess the usual bicyclic structure in accordance with the size of the starting and final rings (see Chart 6). It is of interest that the unusual reaction (23) (36) should be classified as $A \Rightarrow B + X$ and its G_1-graph is also bicyclic. We also have to mention that reactions like (25), usually regarded (29) as examples of cycloaddition - elimination processes, may now be considered as simple ring transformations corresponding to G_1-graph.

Evidently, all the examples on Chart 6 are SHRT. An example of a *non-simple* HRT of the kind $A + X_1 \rightarrow B + X_2$ is a rather complex reaction (14), see Chart 4, where the starting pyridine ring appears in both the cycles of the final indole nucleus (35).

Direction of bond heterolysis. Following the algorithm of RBR-graphs construction, it is easy to determine the degree of similarity of any SHRT. Nevertheless, there is still another *chemical* reason for a further classification of similar SHRTs. Let us compare the reactions (26) and (27) (37a,b), Chart 7. Both reactions
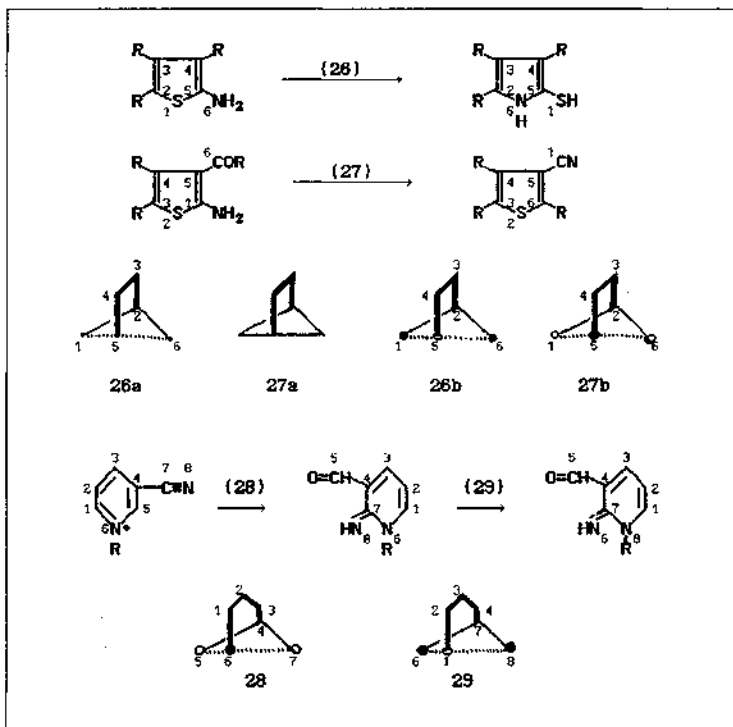

CHART 7.

are represented by the identical G_0-graph 26a or G_1-graph 27a. The real difference between these SHRTs, however, consists in the number of electrophilic and nucleophilic centres, participating in the reaction mechanism. Indeed, for the Dimroth rearrangement (26), electrophilic carbon C-2 leaves the nucleophilic sulphur atom and forms a bond with the nucleophilic nitrogen of the exo-amino group (37a). On the contrary, in reaction (27) the nucleophilic sulphur atom goes away from the electrophilic C-1 atom to the electrophilic C-6 atom of the carbonyl group (37b). Thus, we have to take into consideration the number and the distribution of potentially nucleophilic or electrophilic centres both in the ring and in the side chain. Let us designate the nucleophilic ("donor") atoms by a black heavy dots and the electrophilic ("acceptor") atoms

by a hollow circles at the end of the corresponding vertices in G_1-graph 26a. The resulting new graphs, 26b and 27b contain useful additional information on the direction of bond heterolysis in reactions (26) and (27). We call this kind of graph a G_2-graphs and determine one more step in our classification : the SHRTs are of the same *sort* if their G_2-graphs coincide. For instance, the reactions (26) and (27) are of the same *type*, *kind* and *class*, but of different *sorts*. In fact, the rearrangement (26) is of the same sort as the Dimroth rearrangement (8). Another suitable illustration of the application of G_2-graphs is the interesting consequence of SHRT (28) and (29) (38), see Chart 7. From the structure of the corresponding G_2-graphs 28a and 29a, one can more easily grasp the resulting mechanism of this double HRT, as well as the difference in direction of bond heterolysis in both pyridinium rings.

It is obvious that the construction of G_2-graphs (which are, in fact, labelled graphs) is applicable only to the reactions with heterolytic mechanism of bond redistribution. Indeed, the whole family of the ANRORC-reactions can be treated in the framework of our approach using G_2-graphs. However, if HRT describes the process with a homolytic or synchronous mechanism, only G_0- or G_1-graphs can be used. It should be also mentioned that any RBR-graph of a direct reaction is identical to the corresponding RBR-graph of the reverse reaction because the algorithm for RBR-graph construction includes both the starting and final molecules.

As we have discussed, one can construct G_0-, G_1- and G_2-graphs for any HRT, including not only the *simple* reactions. In the cases of reactions which are *not* SHRT type, it can lead to a more complex RBR-graph type. For instance, the corresponding graphs of the reactions in Chart 4 would have more than two cycles in their G_0-graphs and will evoke more difficulties in classification. In this paper we shall limit our analysis exclusively to SHRTs, because they are, in fact, the most widely spread HRT.

The suggested notion of RBR-graphs G_0, G_1 and G_2 is novel and constitute a simple image of HRT. Moreover, it can be developed into an exact mathematical instrument in the field of heterocyclic chemistry, which, at present is a descriptive rather than an exact science.

SYMMETRY PROPERTIES OF RBR-GRAPHS AND THE HIERARCHIC CLASSIFICATION OF SHRT

Now it is necessary to examine more carefully the structure of the suggested graphs G_0, G_1 and G_2 and relationships between them.

G_0-**graphs**. As has been pointed out, G_0-graphs are bicyclic graphs (32) with a labelled chain; for the purpose of classification one can omit this label to get parent unlabelled graphs (abstract graphs). It should be mentioned that the resulting bicyclic abstract graphs are simply the graphs with cyclomatic number 2 (33) and there are different possibilities for the topology of cycle annelation, symbolically illustrated by pictures 30 (θ-topology), 31 (spiro-topology) and 32 (Chart 8).
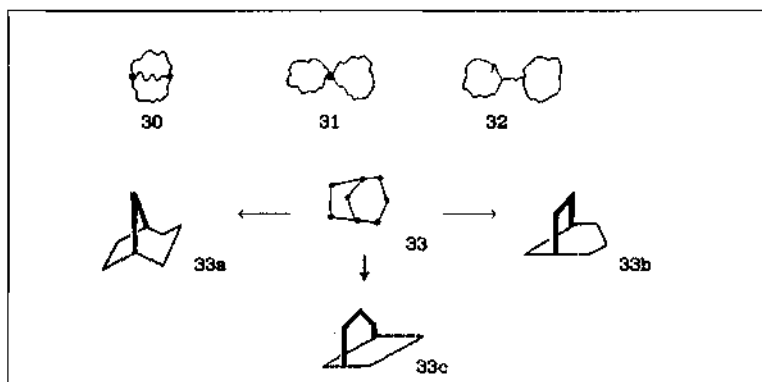

CHART 8.

The example of the correspondence between G_0-graph **33a** and its parent abstract graph **33** is given in Chart 8. The reversible procedure of labelling a chain in **33** gives three different G_0-graphs **33a-c**. These RBR-graphs, from the chemical point of view are correspond to quite different types of ring transformations. This is obvious from a comparison of the sizes of annelated rings in **33a-c**, *i.e.* the sizes of the starting and final heterocyclic systems.

The last consideration gives another reason for classifying the structure of G_0-graphs by the size of the pair of cycles (starting and final heterocycles) which could be annelated only by a certain and finite numbers of combinations; in this case the only variable value will be the size of the labelled bridge, common to both cycles. Chart 9 illustrates all the nonequivalent possibilities for the annelation of cycles with a sizes from 5 to 7; in each row the only variable is the size of the mutual bridge.
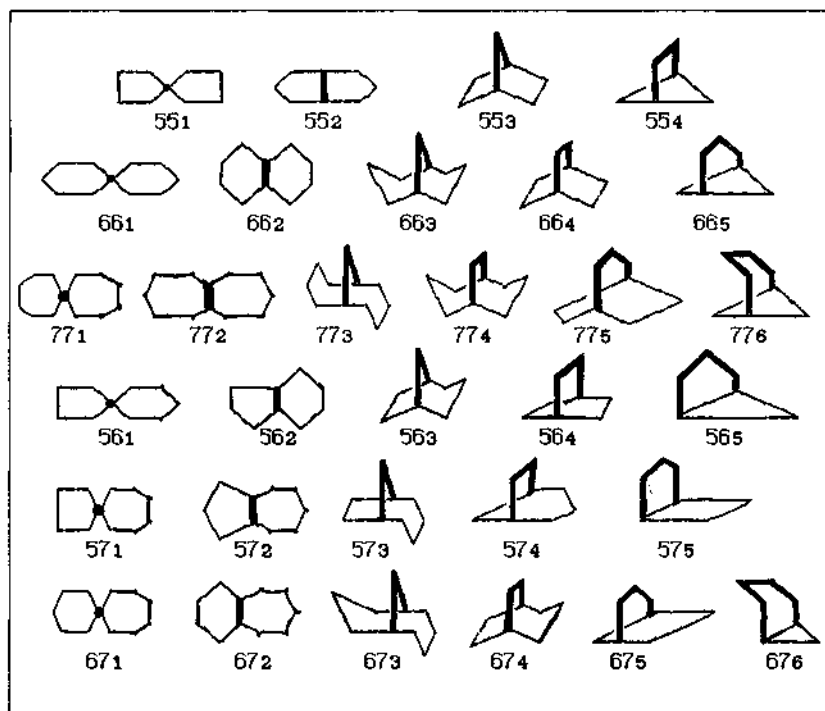

CHART 9.

In general, we can codify any G_0-graph by the size of the annelated cycles and the size of the bridge as follows : if M and N are the sizes of the smallest and the largest cycles, and K the number of vertices in the labelled bridge, then the vector (M, N, K) codifies the only one G_0-graph. In the general case $O < K < N$; the corresponding examples for K = O (unconnected cycles in **11b**, FIG. 3), K = 1 (spiro-cycles in **10b**, Chart 3) or K = N (G_0-graphs **18** and **19**, Chart 5) have been shown above. Now one can see the chemical difference between G_0-graphs **33a-c** on Chart 8, which are correspond to the abstract graph **33** : they differ in their codes (563), (574) and (675), i.e. they are responsible for HRT with different sizes of starting and final rings, namely 5-6, 5-7 and 6-7.

The suggested code (rather analogous to IUPAC-rules for bicycloalkanes) permits us easily to tabulate all the possible combinations of ring transformation of M-membered heterocycles to N-membered ones. All the types of SHRT with their codes from 5- to 7-membered rings are shown in Chart 9; particularly, this is the complete classification of SHRT-types for usual heteroaromatics.

G₁-graphs. As we defined above, G_0-graphs were constructed from G_1-graphs by changing the dashed edges into solid ones (see Chart 3); in this way SHRT's have been classified into types. If we try to reverse the procedure, i.e. to sub-classify or pick out the different classes in the given type of SHRT, we have to label some edges in a given G_0-graph *vice versa* to dashed ones. As one can see, the number of dashed labels is determined by the reaction kind, namely by the number of components A, B, X in the SHRT (see Chart 6). Let us limit the SHRT by the simplest (A→B) kind, i.e. for rearrangements. In this case the only two dashed labels (responsible for the broken or formed bonds) are needed. These labels can be distributed between the annelated cycles of G_0-graphs in different ways, due to the structure and symmetry of the given G_0-graph. Some examples are given in Chart 10. (It is obvious, that the bold-marked edges of the bridge in G_0-graph
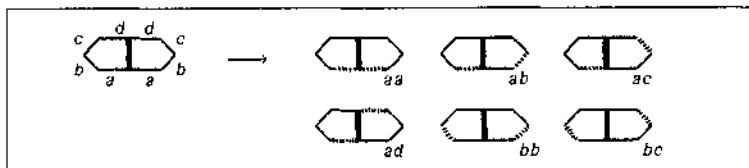

CHART 10.

should not be labelled, since they are responsible for the unchanged bonds in both heterocyclic rings). Simple combinatorial considerations permit to enumerate the number of possible non-equivalent G_1-graphs, corresponding to the graph G_0, *i.e.* all theoretically possible classes of rearrangements in given reaction type. For instance, all the classes of the type (552) are presented in Chart 10. Let m and n are the number of solid edges in each cycle of G_0-graph; for G_0-graph with the code (MNK) $m = M-K+1$ and $n = N-K+1$. The combinatorial equations for the full number of non-equivalent classes of rearrangements (G_1-graphs), responsible for the given type (G_0-graph) are shown in Table 1. As one can see from the table, only the

TABLE 1.

G_0-graph topology	θ-Topology (as in 30)					Spiro-topology (as in 31)				
Equality of annelated cycles size	Equal		Non-equal			Equal		Non-equal		
Odness*) of the cycles	Even Even	Odd Odd	Even Even	Even Odd	Odd Odd	Even Even	Odd Odd	Even Even	Even Odd	Odd Odd
Number of G_1 graphs	$\dfrac{m^2+2m}{4}$	$\dfrac{(m+1)^2}{4}$	$\dfrac{mn}{2}$	$\dfrac{mn}{2}$	$\dfrac{mn+n}{2}$	$\dfrac{m(m+2)}{8}$	$\dfrac{(m+1)(m+3)}{8}$	$\dfrac{mn}{4}$	$\dfrac{mn+n}{4}$	$\dfrac{(m+1)(n+1)}{4}$
					$(m<n)$				n-Even m-Odd	

*) The oddness of the number of solid (not bold) lines in every annelated cycle of G_0-graph.

topology and evenness or oddness of m and n values in the G_0-graph are necessary to enumerate the corresponding G_1-graphs. There is no serious problem to extend the combinatorial examination from rearrangements to SHRT of any kind; the only difference will lie in the number of dashed labels, responsible for the formed or broken bonds.

A simple convention can be used to codify any class of rearrangements on the basis of the distribution of dashed labels in the bicyclic structure. Since G_0-graph contains two annelated cycles, let us denote the unlabelled edges in the lesser cycle by the letters a, b, c..., beginning from the edge which is the nearest to the bridge (see Chart 10); we also denote by the same letters the edges of the annelated cycle, beginning from the edge which is adjacent to the a-edge of the lesser cycle. As a consequence, the expression (MNK-ij), as for instance (552-ab), will codify a certain G_1-graph or a certain class of SHRT (the first letter i is responsible for the lesser cycle; for equal ring size the letters i, j should be lexically regulated). For instance, we can construct and codify certain G_1-graphs from each G_0-graph, which contains annelated 5-membered cycles (i.e. of (55K)-type); in Chart 11 the correspondence between the resulting G_1-graphs and their codes is shown. As one can see, the given graphs on Chart 11 exhaust all possible classes of rearrangements of 5-membered heterocyclic rings. Moreover, in Table 2 the codes of all theoretically possible classes of reciprocal rearrangements between heterocycles with ring sizes from 5 to 7 are tabulated.
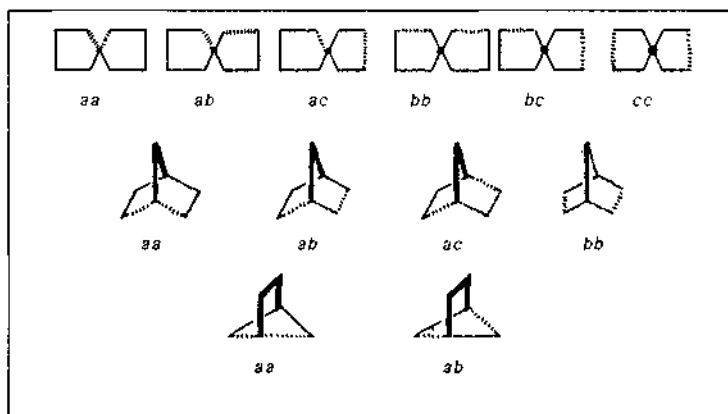

CHART 11.

G_2**-graphs.** As we have shown above, the SHRTs of a given class (reactions with identical G_1-graphs) can differ from each other by the sort of ring transformation due to the distribution of electrophilic and nucleophilic centres in the starting molecule, i.e. by their G_2-graph structure. Now the problem of enumeration and construction of all possible G_2-graphs (reaction sorts) for given G_1-graph (reaction class) arises. This problem is equivalent to the enumeration of labelled G_1-graphs of a given structure, where the label is (see Chart 7) a hollow circle or a heavy point for the vertices adjacent to the dashed labels. (Since the colours show the direction of bond heterolysis, the only requirement is that the ends of any dashed edge should have different colours). For the examples examined in Chart 7, there are only two variants of G_2-graphs responsible for G_1-graphs both of the codes (554-aa) and (665-aa). It is obvious that in any case a G_1-graph of (MNK-aa)-structure should correspond only to a pair of G_2-graphs. In the general case, the number of G_2-graphs depends on the symmetry of the initial G_1-graph. As one can see from Chart 12, the possible number of G_2-graphs for heterocyclic rearrangements would vary from one to four. Simple group-theoretical considerations (see Appendix and Chart 13) can be used to predict the number of possible G_2-graphs corresponding to the G_1-graph of a given symmetry, Chart 13. The suggested rule permits us to enumerate the number of the sorts of heterocyclic rearrangements just by a first glance on the reaction class. It is useful for any SHRT with any structure and symmetry of a G_1-graph with the topology of θ-type (one can for instance examine the groups of the graphs on Chart 12); the only exceptions are the above mentioned G_1-graphs of the aa-classes.

From the chemical point of view, an important problem is how to get a common chemical equation from the structures of RBR-graphs. As we have mentioned above, any RBR-graphs of the direct reaction is

TABLE 2.

| Possible codes for G_0-graphs	Possible codes for the corresponding G_1-graphs													
561	*aa*	*ab*	*ac*	*bb*	*bc*	*cc*								
562	*aa*	*ab*	*ac*	*ad*	*ae*	*bb*	*bc*	*bd*	*be*					
563	*aa*	*ab*	*ac*	*ad*	*ba*	*bb*								
564	*aa*	*ab*	*ac*											
565	*aa*													
571	*aa*	*ab*	*ac*	*ad*	*bb*	*bc*	*bd*	*cc*	*cd*					
572	*aa*	*ab*	*ac*	*ad*	*ae*	*af*	*ba*	*bb*	*bc*	*bd*	*be*	*bf*		
573	*aa*	*ab*	*ac*	*ad*	*ae*	*ba*	*bb*	*bc*	*bd*	*be*				
574	*aa*	*ab*	*ac*	*ad*										
575	*aa*	*ab*												
661	*aa*	*ab*	*ac*	*bb*	*bc*	*cc*								
662	*aa*	*ab*	*ac*	*ad*	*ae*	*bb*	*bc*	*bd*	*cc*					
663	*aa*	*ab*	*ac*	*ad*	*bb*	*bc*								
664	*aa*	*ab*	*ac*	*bb*										
665	*aa*	*ab*												
671	*aa*	*ab*	*ac*	*ad*	*ba*	*bc*	*bd*	*cc*	*cd*					
672	*aa*	*ab*	*ac*	*ad*	*ae*	*af*	*ba*	*bb*	*bc*	*bd*	*be*	*bf*	*ca*	*cb* *cc* *cd*
673	*aa*	*ab*	*ac*	*ad*	*ae*	*ba*	*bb*	*bc*	*bd*	*be*				
674	*aa*	*ab*	*ac*	*ad*	*ba*	*bb*								
675	*aa*	*ab*	*ac*											
676	*aa*													
771	*aa*	*ab*	*ac*	*ad*	*bb*	*bc*	*bd*	*cc*	*cd*	*dd*				
772	*aa*	*ab*	*ac*	*ad*	*ae*	*af*	*bb*	*bc*	*bd*	*be*	*cc*	*cd*		
773	*aa*	*ab*	*ac*	*ad*	*ae*	*bb*	*bc*	*bd*	*cc*					
774	*aa*	*ab*	*ac*	*ad*	*bb*	*bc*								
775	*aa*	*ab*	*ac*	*bb*										
776	*aa*	*ab*												
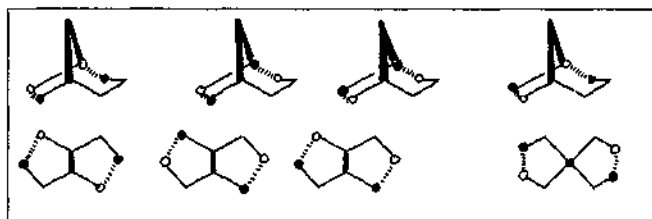
CHART 12.

identical with the corresponding graph of the reverse reaction; that's why the given reaction graph (G_1 or G_2) will determine only the pair of molecular graphs (corresponding to the structure of both the starting and the final molecules), but not the direction of the reaction. It is easy to reconstruct the corresponding pair of molecular graphs (see for instance 1a, 2a in Chart 2) starting from the G_1-graph (e.g. 3) : one must remove one or other dashed edge in the G_1-graph. In the general case, when the G_1-graph contains more than two dashed edges, the first molecular graph should be constructed by removal of the dashed edges from the first cycle of the G_1-graph and the second molecular graph - correspondingly by the same procedure in the second annelated cycle.
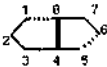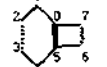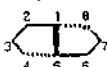
Example of G_1-graph	Elements of permutation group	Class of the permutation (a, b, c)	Number of G_2-graphs
	(1)(2)(3)(4)(5)(6)(7)(8) (or E)	a	4
	E, (17)(26)(35)(4)(8)	a c	3
	E, (14)(23)(58)(67)	a b	2
	E, (15)(26)(37)(48)	a c	3
	E, (19)(28)(37)(46)(5)(10), (16)(27)(38)(49)(5 10), (14)(23)5 10)(69)(78)	a c c b	2

CHART 13.

Hierarchic classification. Any rearrangement (as well as SHRT) with given mechanism belongs to a certain type, class and sort of ring transformation reaction, and is described by a certain G_0-, G_1- and G_2-graphs. The resulting classification proves to be the hierarchic one : the types, the classes and the sorts should represent different levels of the classification tree; and besides, each lower level determines the structure of a higher level. Really, on the first level, only the ring size M and N of the starting and final heterocycles is defined; every pair (M, N) determines at the second level the inequivalent (MNK)-vectors which can be represented by the possible ways of annelation of M- and N-membered cycles or labelled G_0-graphs. Every reaction type can be divided into different reaction kinds and classes, which form the third level of the classification tree. We want to mention, that the kind of a reaction is defined only by the number of dashed edges in the G_0-graph, while the class - by the dashed edges distribution; the possible number of kinds is determined by the full number of unlabelled edges in the starting G_0-graph, the number of classes (or G_1-graphs) - both by the reaction kind and by the structure and symmetry of the G_0-graph. At last, the symmetry and the distribution of the dashed edges in every G_1-graph of the third level determines the possible structures of G_2-graphs (or reaction sorts) on the last (fourth) level. From a given G_2-graph one can get the corresponding pair of molecular graphs which would contain the labels responsible for potentially electrophilic and nucleophilic centres. The resulting molecular graphs are the images of concrete starting and final heterocycles.

In principle, one can continue the classification to deeper levels by including as new parameters the number and the distribution of heteroatoms, multiple bonds and condensed rings. It is obvious, however, that this kind of problem would be examined better by the use of computers. Some simple combinatorial rules for these new levels as well as the computer program, including different levels of classification, are now in progress.

The suggested classification permits us to define the structural similarity of different ring transformations in accordance with their disposition in the classification tree. In this way one can satisfactorily classify the already existing SHRT. On the other hand, the hierarchic structure of the classification and the usefulness of combinatorial methods turn us to the reversible problem of prediction of new probable sorts, classes and types of SHRT. Although in general this problem seems to be rather complex, it is possible to give several remarks on the existing rearrangements of 5-membered heteroaromatics with the purpose of predicting new ring transformations. In the present paper we have to be limited only by the qualitative examination of the problem and the details will be given in further communications.

We were limited only by the SHRT-rearrangements of 5-membered aromatic heterocycles (including the condensed aromatic rings) only of the $A \Rightarrow B$ - kind which have been mentioned in the reviews [22, 23, 25, 27, 28]. We have examined 170 examples of different rearrangements (consider rearrangements to be identical when the difference was only in the nature of substituents). It has been found that these rearrangements are distributed among the types, classes and sorts of SHRT in a rather irregular way, see Chart 14. Between the five possible reaction types (see G_0-graphs on Charts 3,9) the most important (97% of all rearrangements) are G_0-graphs with the codes (552), (553) and (554) i.e. of θ-topology. Most interesting is the distribution into classes of every type : the aa-class turns to be the most widely spread class (more than 90% in every type) among the azoles rearrangements. In particular, the well known examples of Dimroth, Cornforth and Boulton-Katritzky rearrangements belong to the aa-classes.

As we have already mentioned (see Chart 7) there are only two possible sorts of G_2-graphs corresponding to G_1-graphs of aa-classes : the possible structures should contain either two nucleophilic (N) and one electrophilic (E) centre (NNE-sort) or *vice versa* (EEN-sort). The rearrangements of the first NNE-sort turn out to be more wide-spread in azole chemistry and they were also better examined. (The named rearrangements of Dimroth (554-aa-NNE), Cornforth (553-aa-NNE) and Boulton-Katritzky (552-aa-NNE) are just of this sort). The contrary EEN-sort, nevertheless, has also been found and its part varies from 20% in the (554-aa)-class to 10% in the (552-aa)-class. Due to the opposition in the distribution of donor/acceptor centres this reaction sorts could be conventionally named as "Anti"-Dimroth sort and "Anti"-Boulton-Katritzky sort (see Chart 14) respectively. To our knowledge, the corresponding example of the rather interesting "Anti"-Cornforth rearrangement of the sort (553-aa-EEN) is unknown yet (see G_2-graph 34 on Chart 14). Probably, this sort of SHRT is waiting to be really the new one in SHRT.

Name of HRT	$N^{*)}$	G_2-graph	$N^{*)}$	G_1-graph	Code of G_1-graph
Dimroth reaction	44		59		554-aa
Opposite sort	15				
Cornforth reaction	32		32		553-aa
Opposite sort	-				
Alberth-Boulton-Katritzky reaction	45		50		552-aa
Opposite sort	5				
Holla, Ambekar reaction	5		5		551-aa
Opposite sort	-				
Others	24	-	24	-	Others

A new sort of HRT ("Anti-Cornforth" reaction)

G_2-graph: 34 Equation:
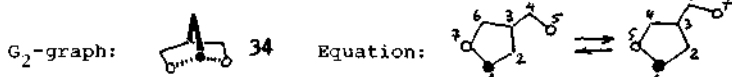
$^{*)}$ N - the number of known unequivalent HRTs

CHART 14.

APPENDIX

Let us limit ourselves by the graphs with θ-topology (like 30 on FIG. 8), which is more widespread in SHRT than the other topologies like 31 or 32. The symmetry of a G_1-graph (as of anyother graph), can be characterized by the group of automorphisms of the graph, which is isomorphic to a certain permutation group, determined, for instance, on the graph vertices [33]. Examples of ordinary permutation groups of the vertices of some G_1-graphs and the possible number of corresponding G_2-graphs are shown in Chart 13. As one can see, the permutations can be divided into three classes : a) the identical one, b) those which permute the vertices only in the same annelated

cycle in the G_1-graph and c) those which permute the vertices of different annelated cycles in the G_1-graph. A simple rule can be used to correlate the number of G_2-graphs (L) with the presence of certain permutations in the permutation group of a given G_1-graph :

1. If all the permutations are of a)-class, then $L = 4$.
2. If there exists in the group any permutation of c)-class, then $L = 2$.
3. In all other cases, $L = 3$.

ACKNOWLEDGEMENT

REFERENCES AND NOTES

1. Chemical Applications of Topology and Graph Theory, R.B. King, ed, Elsevier, Amsterdam, 1983.
2. Chemical Applications of Graph Theory, A.T. Balaban, ed, Academic Press, London, 1976.
3. R.E. Merrifield, H.E. Simmons, Topological Methods in Chemistry, Wiley Interscience, New York, 1989.
4. Advances in the Theory of Benzenoid Hydrocarbons, M.J.S. Dewar e.a. eds, *Top. Curr. Chem.*, Vol. 153, Springer-Verlag, Berlin, Heidelberg, 1990.
5. A. Graovac, I. Gutman, N. Trinajstic, in "Lecture Notes in Chemistry", Vol. 4, Springer-Verlag, Berlin, Heidelberg, 1977.
6. D.M.P. Mingos, R.L. Johnston, in "Structure and Bonding", Vol. 68, Springer-Verlag, Berlin, Heidelberg, 1987, p. 29.
7. R.B. King, D.H. Rouvray, *J. Amer. Chem. Soc.* 99, 7834 (1977); R.B. King, *Inorg. Chim. Acta*, 116, 99, 126 (1986); B.K. Teo, *Inorg. Chem.*, 23, 1251 (1984).
8. M.I. Stankevitch, I.V. Stankevitch, N.S. Zefirov, *Uspekhi Khimii (Russ.)* 57, No. 3, 337 (1977).
9. Drug Design : Fact or Fantasy, G. Jolles, K.R.H. Wooldridge, eds, Academic Press, London, 1984.
10. N.S. Zefirov, S.S. Tratch, *Zh. Org. Khim.*, 11, 225, 1785 (1975).
11. N.S. Zefirov, S.S. Tratch, *Chem. Scripta*, 15, 4 (1980).
12. N.S. Zefirov, *Acc. Chem. Res.*, 20, 237 (1987).
13. N.S. Zefirov, S.S. Tratch, *Zh. Org. Khim.*, 20, 1121 (1984).
14. N.S. Zefirov, S.S. Tratch, G.A. Gamziani, *Zh. Org. Khim.*, 22, 1341 (1986).
15. N.S. Zefirov, S.S. Tratch, *Analytica Chimica Acta*, 235, 115 (1990).
16. S. Fujita, *J. Chem. Inf. Comput. Sci.*, 26, 205 (1986); *ibid.* 27, 111 (1987); *ibid.* 28, 128 (1988).
17. E.V. Gordeeva, Candidate Diss., Moscow University, 1986; I.I. Baskin, Candidate Diss., Moscow University, 1989.
18. a) E.V. Babaev in "Principles of Symmetry and Systemology in Chemistry", N.F. Stepanov, ed, Moscow University, 1987, p. 30 (Ch. Abstr. 109, No. 27902b); b) S.S. Tratch, N.S. Zefirov, *ibid.*, p. 54.
19. Yu.K. Yur'ev, N.S. Zefirov in "Development of Organic Chemistry in USSR", Moscow, Nauka, 1967.
20. a) E.V. Babaev, S.I. Bobrovskii, Yu.G. Bundel, *Khim. Geterotsikl. Soedin.*, 11, 1570 (1988); VI Intern. IUPAC Congress on Org. Synth., Moscow, 1987, Thesa A-341.
21. O. Dimroth, *Ann. Chem.*, 364, 183 (1909).
22. H.C. Van der Plas, Ring Transformation of Heterocycles, Academic Press, London, New York, 1973, Vol. 1, 2.
23. G. L'abbé, *J. Heteroc. Chem.* 21, 627 (1984).
24. A.N. Kost, S.P. Gromov, R.S. Sagitullin, *Tetrahedron*, 37, 3423 (1981).
25. O.P. Shwaika, V.N. Artyomov, *Uspekhi Khimii (Russ.)* 41, No. 19, 1788 (1972).
26. a) Pyrillium Salts. Syntheses, Reactions and Physical properties, A.T. Balaban e.a., eds, Academic Press, N.Y., 1982; b) A.T. Balaban in "New Trends in Heterocyclic Chemistry", R.B. Mitra e.a., ed, Elsevier, Amsterdam, 1979, p. 79; c) A.T. Balaban, Reports of VI Intern. IUPAC Congress on Org. Synth., Moscow, 1986, Plenary Lecture.
27. A. Padwa in "Rearrangements in Ground and Exited States", P. de Mayo, ed, Academic Press, New York, Vol. 3, 1980, p. 501.
28. A.J. Boulton, "Lectures in Heterocyclic Chemistry", Vol. 2, R.N. Castle, L.B. Townsend, eds, Heterocorporation, Orem, Utah, 1974, p. S-45; M. Ruccia, N. Vivona, D. Spinelli, *Adv. Heteroc. Chem.*, 29, 141 (1981).
29. Comprehensive heterocyclic chemistry, A.R. Katritzky, C.W. Rees, eds, Pergamon Press, Oxford, 1984, vol. 1-8.
30. Ref. 202-204 in Heterocyclic compounds, K. Shofield, ed, vol. 4, 1973.

31. a) Reaction (7) V.A. Bakulev, E.F. Dankova, V.S. Mokrushin, E.O. Sidorov, A.T. Lebedev, *Khim. Geterotsikl. Soedin.*, **6**, 845 (1987); b) reaction (10) B.S. Holla, S.I. Ambekar, *J. Chem. Soc. Chem. Comm.*, 221 (1979); c) reaction (11) B. Stanovnik, M. Tisler, B. Stepanov, *J. Org. Chem.*, **36**, 3812 (1971); d) reactions of general type (4, 6, 8, 9, 22) from reviews [22,23,28]; reaction (5) ref. 765 in [22]; reaction (8) T. Kindt-Larsen, C. Pedersen, *Acta Chem. Scand.* **16**, 1800 (1962).

32. Really, for simple ring transformations the G_0-graph is constructed from the molecular graph of the starting system (which contains only one cycle) by addition of only one new edge into the final cycle. Independently of the place of the new edge arrangement, the resulting G_0-graph has to contain only two cycles [33].

33. F. Harary, Graph Theory, Addison-Wesley, Reading, MA, 1969.

34. Reaction (15) ref. [20]; reaction (17) ref. [24]; reaction (16) P.E. McCuscer, E.M. Philbin *J. Chem. Soc.*, **4**, 2374 (1963); reaction (18) A. Lattes in "Lectures in Heteroc. Chem.", M. Tisler, ed, Ljubljana, Vol. 3, 1975, p. 93; reaction (19) R.M. Dorson, P.D. Hammen, R.A. Davis, *J. Org. Chem.* **36**, 2693 (1971).

35. M.A. Yurovskaja, A.Z. Afanas'ev, V. Chertkov, E.M. Gizatullina, Yu. G. Bundel, *Khim. Geterotsikl. Soedin.*, **12**, 1625 (1987); M.A. Yurovskaja, Doctor diss., Moscow University, 1990.

36. C.J. Moody, C.W. Rees, S.C. Tsoi, *J. Chem. Soc. Chem. Comm.*, **11**, 550 (1981).

37. a) W.G. Middleton, V.A. Enhelhardt, B.S. Fischer, *J. Amer. Chem. Soc.*, **80**, 2822 (1958); b) O. Meth-Cohn, B. Narine, *J. Chem. Res.* (S), 294 (1977).

38. S.L. Johnson, K.A. Rumon, *Tetrahedron Letters*, 1721 (1966).