Combinatorial Models and Polarity Control Rules in Heterocyclic Design

Eugene V. Babaev

Chemistry Department, Moscow State University, Moscow, 119899, Russia

Abstract. *The review is devoted to applications of combinatorial models for classification and design of heterocyclizations, recyclizations, and ring opening reactions of heterocycles. These models are based on the combination of concepts of reaction graphs and various polarity control rules. The existing models of reaction graphs and polar classifications of reagents are critically overviewed. The model of polar cyclization graphs is suggested for design of novel routes to the consonant heterocycles. The models of local and global polarity control in the heterocyclic ring opening reactions and their applications are discussed. The model of the ring bonds redistribution graphs for the recyclization reactions and its applications to the computer-assisted search of novel ring transformations are reviewed. Elaborated computer programs and examples of experimental confirmation of some predictions are discussed.*

1. Introduction

The ring with heteroatoms is the central subject of heterocyclic chemistry. Since the cycle is the fundamental topological concept, the processes of cycle formation and destruction are among the central topics in the chemistry of heterocycles. The reactions with such essential topological changes are the well-known processes of heterocyclizations (*appearance* of a cycle), ring opening reactions (*disappearance* of a cycle), and recyclizations (*transformation* of one cycle to another). Many generations of chemists have contributed to understanding peculiarities of these three fundamental types of reactions, but even nowadays the discovery of new examples of such reactions is occasional. Due to the lack of exact mathematical models the chemistry of heterocycles is rather an art than a strict science. Of course, there are many attempts to rationalize the logic of heterocyclic chemistry by applying various theoretical models and implement them as computer programs (see e.g., the comprehensive review[1] on this topic). In spite of these efforts, the spacious factual material on heterocyclic rings (trans)formation and cleavage still requires better classification.

In our early works we suggested few combinatorial models helpful for classification and design of reactions in heterocyclic chemistry. These models, based on the combination of two important concepts of the *reaction graphs* and of the *polarity control rules,* have been applied to heterocyclizations, recyclizations and ring opening reactions of heterocycles. The goal of this paper is to review our methodology and to prove its usefulness in solving some selected problems of heterocyclic chemistry. The plan of the review is the following. In Section 2 we shall briefly review the concept of reaction graphs. Then in Section 3 we shall

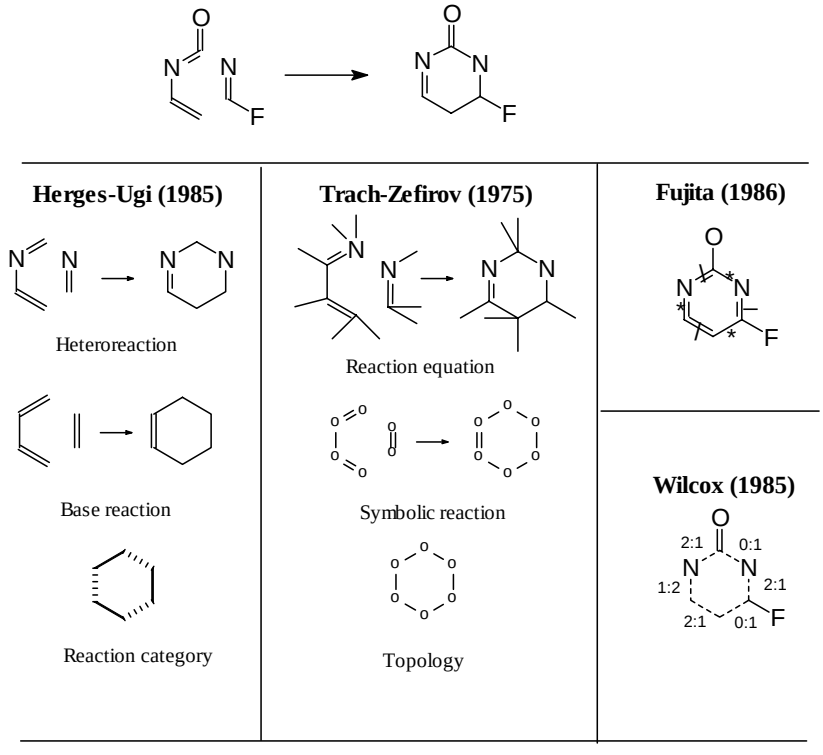
critically review the concept of polarity control in the heterocyclization reactions and suggest the model of polar cyclization graphs as a useful tool in classification and design of novel cyclizations. In Section 4 the problem of local and global polarity control in the ring opening reactions will be discussed. In the last Section 5 the model of the ring bonds redistribution graphs for the recyclization reactions will be reviewed, and our attempts of the computer-assisted search of novel ring transformations will be discussed.

2. Reaction Graphs as the Diagrams of Bonds Redistribution

2.1. What is the reaction graph?

During the last two decades methods of graph theory became useful instruments for the analysis of various problems in theoretical and experimental chemistry[2,3]. A traditional application of graphs in chemistry is the description of molecular structures. Indeed, a graph, as a set of vertices and edges, corresponds to the classical notion of a structural formula as a sequence of atoms and bonds. An unusual important application of graph theory is the use of a graph as a representation of a chemical reaction, namely, for the description of bond redistribution in the course of an organic reaction. The first mathematical and graph-theoretical approaches to the description of structural changes during a chemical reaction were developed in the 70's by Balaban[4], Ugi and Dugundji[5,6], Hendrickson[7], and Trach and Zefirov[8,9]. Later publications in this direction are represented by the research of Fujita[10,11] and others, see reviews[1,12,13].

How can the graph be used as an image of chemical reaction? The central idea of virtually all the approaches (see examples in Scheme 1) consists in the establishment of a correlation between the atoms of the



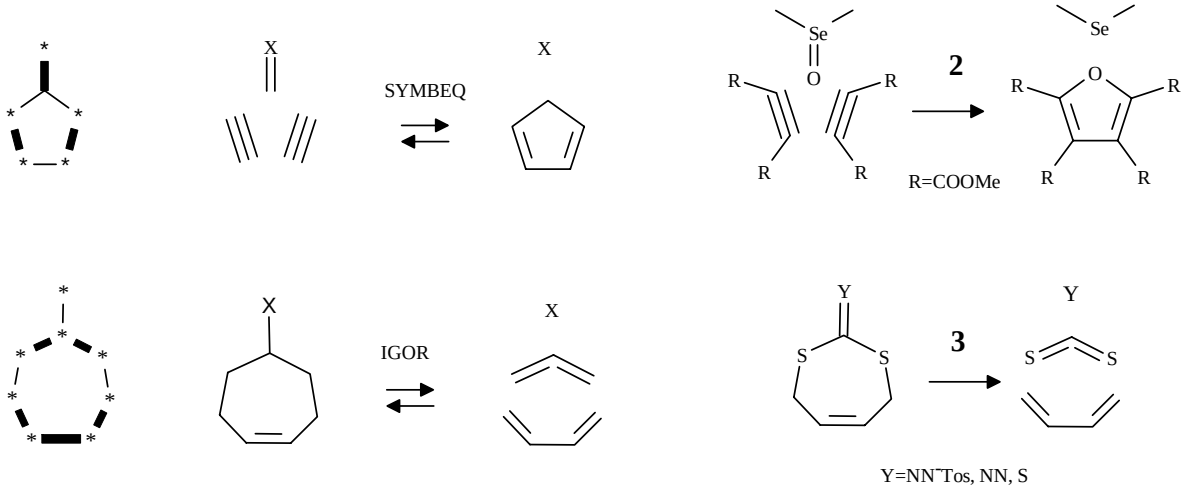
Scheme 1. Examples for description of the same reaction by various reaction graphs[1].

reactants and the products by means of their "superimposition" upon one another. After this confluence (or superposition), one can unambiguously describe the *changes in the bonds* during a chemical reaction, expressing this redistribution of bonds by a graphical diagram. Such diagrams have different names (for example, symbolic equations, reaction categories, and even virtual transition states), and we shall use below the term *reaction graphs*. As an example of how the diagrams of the same reaction look in the languages of the various approaches, one can consider a conceivable heterocyclization 1 and the corresponding reaction graphs in Scheme 1. (For various aspects of representing the reactions see reviews[12,14].)

2.2. The use of reaction graphs for classification and design of new reactions

The reaction graphs (like those reported in Scheme 1) substantially simplify the usual chemical equation which is changed to the graphical aspect ("skeleton") of the reaction. Such skeleton, as a rule, contains atoms that change the oxidation state (or any environment) and bonds that change their multiplicity during the reaction. Thus chemical reactions can be reduced to a finite number of bond redistribution diagrams. A comparison of such diagrams for various reactions makes it possible, first and foremost, to reveal the similarity or the degree of novelty of the reactions themselves[5,8].

Another way to use bond redistribution diagrams is the prediction of new reactions. For this purpose one is required to build on real substituents, chains, rings, etc. to these diagrams. In a number of cases this sort of predictions was heuristic. The indicated principle was the cornerstone of the computer programs SYMBEQ[15,16] and IGOR[17] capable to generate the reaction graphs. The first example of computer-assisted discovery of new reaction was the novel synthesis of the furan ring (equation 2 in Scheme 2) predicted by the SYMBEQ program[15]. Another example of new reaction, that has been predicted by means of the IGOR program and discovered experimentally, is the novel diene synthesis as the result of unknown heterocyclic ring opening reaction (equation 3 in Scheme 2)[18].



Scheme 2. Examples of prediction of novel reactions by use of reaction graphs methodology. Left: the reaction graphs; middle: “skeletal” equations restored from reacttion graph; right: actually found examples of new reactions[15,17].

2.3. Advantages and disadvantages of the reaction graph models

The concepts of reaction graphs have introduced into the chemistry a novel idea of a *function* (or operator) that is acting on the set of molecular structures. Evidently, this function can be equally applied to either starting ensemble of reactants (predicting the products) or to the final set of products (predicting the reactants). Hence, having been first claimed, the concept of reaction graphs looked as a new promising tool for resolving key problems of organic synthesis planning. This forecast, however, turned out to be too optimistic, and the reaction graphs still have not become the natural language of chemistry.

Although the idea of reaction graphs was developed to rigorous mathematical level[15], it is of rather limited use in the chemistry of heterocycles. Thus, in many important reactions, e.g., usual cyclocondensations and recyclizations, it is either difficult to built simple reaction graphs or to apply them fruitfully for classification and design of novel reactions. The first problem is that the multicomponent reactions (that are widespread in heterocyclic synthesis) involve many leaving groups, and considering of all such groups may dramatically complicate the structure of the resulting bond redistribution diagram. The second serious problem lies in the tautomerism of heterocycles (and some their acyclic precursors), which makes difficult to assign certain reaction graphs for such processes. Indeed, taking into account various tautomeric pairs results in many “waste” reaction graphs, while the important information about the bonds from the heterocyclic rings is often lost.

It is therefore not surprising that the concept of reaction graphs may find (and has really found) most useful application in the specific area of classification and design of the concerted processes. In such reactions the bonds are redistributed along a cyclic framework (actual or imaginary), and the location of all formed/broken bonds along the monocyclic reaction graph is the definite procedure. For more complex reactions the (potentially fruitful) idea of reaction graphs requires some natural modifications.

3. THE GRAPHS OF HETEROCYCLIZATION REACTIONS

3.1. Classical disconnection schemes of heterocycles as the specific type of reaction graphs

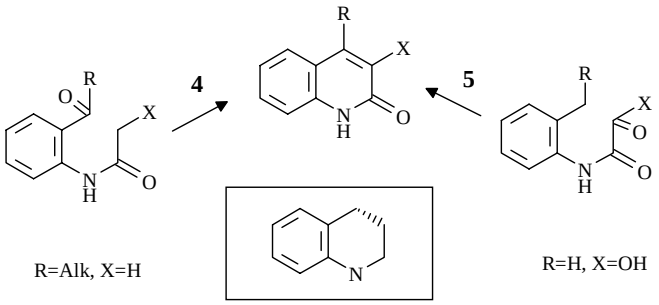
Although heterocyclic chemists have no habit to use the reaction graphs for classification and design of heterocyclization reactions, they do intuitively use quite similar principle of *mapping* the atoms of a heterocyclic product into the atoms of starting reactants. Indeed, the standard graphic method used to demonstrate the diversity of possible cyclizations leading to the given heterocycle is the disconnection of the skeleton of the heterocycle into fragments. On paper, this is done by simply removing (or marking by a dashed line) the skeletal bonds that arise during the course of the reaction. By now it is difficult to establish the authorship of this construction, which is intuitively obvious: as far back as the 1950s, the use of such "diagrams" or "schemes of synthesis" was extremely popular[19]. Today it is simply impossible to review the methods of synthesis of a specific heterocycle without such disconnection schemes (see for example[20]); and

in the classic handbook[21] these diagrams are methodically enumerated and used for a comprehensive classification of the synthesis of diverse heteroaromatic systems.

Drawing of a disconnection scheme means direct and reverse mapping of matching atoms of the acyclic reactants and cyclic product. Therefore these schemes resemble the reaction graphs, although they are in some sense "bad reaction graphs." Indeed, such diagrams bear incomplete information on the total redistribution of bonds and changes in atoms' environment and display only those cyclic bonds that are formed. (Thus, the leaving groups of reactants eliminated in the course of reaction usually are not presented in the disconnection schemes.) Nevertheless, let us call them the *cyclization graphs*.

All of the advantages of describing heterocyclic syntheses by such graphs are obvious at first glance. If there are no permutations of atoms via cyclization it is easy to establish the structural kinship between different syntheses of a given heterocycle, or even the syntheses of different heterocycles. Hence, analogously to the reaction graphs, the oversimplified cyclization graphs are helpful in classification of reactions. The combinatorial character of such diagrams is also obvious, and it is easy to enumerate either manually (or by a programmable microcalculator) all the conceivable schemes for the synthesis of simple heterocycles. Theoretically deduced cyclization graphs can be also compared with real syntheses, and for the still unknown schemes an attempt can be made to devise likely reagents that might be suitable for use in a new type of synthesis. A discussion of such attempts (particularly for synthesis of thiazoles and indazoles) can be found in the reviews[1,12].

Meanwhile, an obvious disadvantage of the cyclization graphs is that they do not offer any information whatever on the chemistry involved in the formation of the skeletal bond. In fact, a single diagram may describe processes that are basically different. For example, in the synthesis of quinolines[22] exactly the same skeletal bond C3-C4 is formed in the Camps reaction 4 and in the Madelung reaction 5 (Scheme 3).



Scheme 3. Synthesis of quinolines by the Camps reaction 4 and the Madelung reaction 5. Both reactions are described by the same disconnection scheme (cyclization graph) shown in rectangle

It would appear that the synthetic schemes should be expressed by a common diagram. However, it is obvious to any organic chemist that in these intramolecular condensations the carbonyl and methylene components (i.e., the electrophilic and nucleophilic centers) have simply changed places relative to the heteroatom (or the benzene ring). As a result, the heuristic and classifying function of cyclization graphs is

seriously limited, and it is advisable to modify them by supplementing with information on the polar nature of the reaction centers.

3.2. How to account for the nature of reaction centers in the reagents?

3.2.1. Classifications of reagents according to the polarity of terminal atoms. In order to avoid paradoxical identification of different reactions (as in the example above) by the same cyclization graph it is necessary to distinguish the electro- and nucleophilic nature of the reaction centers in reactants. This problem has attracted special attention in the reviews on heterocyclic synthesis. Thus, Potts' classification[23] of reagents involved in heterocyclizations included the concept of bielectrophiles (in particular, 1,1-, 1,2-, and all the way up to a 1,5-bielectrophile) and also the concept of binucieophiles (in particular. 1,2-, 1,3-, and l,4-binucleophiles). In addition, Potts mentioned in passing the importance of a third category of reagents that are simultaneously electro- and nucleophiles, but focused primary attention on the design of bielectrophiles and binucleophiles by varying the degree of unsaturation (hybridization) of the reaction center. This type of classification in regard to the importance of the electronic nature of the reagents' reaction centers in describing heterocyclization was developed further, for example, in the handbook of Katritzky[24] and paper of Jorgensen[25].

In order to provide more heuristic utilization of this concept in the design of heterocyclic structures or reactions, let us generalize the Potts' classification in two directions. First, let us add to the existing classification of 1,*k*-bielectrophiles and 1,*k*-binucleophiles the no less rigorous concept of a 1,*k*-"ambiphile," that is a *k*-atomic chain with the opposite polarities of terminal atoms. (Natural particular case should be a 1,1-ambiphile with simultaneous combination of electro- and nucleophilicity in one and the same atom, like in carbenes, or halogen derivatives of CH-acids.) Second, we should consider the case concept of a 1,1-binucleophile, which for unknown reason was missing in early classification tables [23-25]. Typical 1,1-binucleophiles (like water, hydrogen sulfide, amines, and aliphatic nitro compounds) are often used in heterocyclizations with 1,4- or 1,5-bielectrophiles, forming five- or six-membered hetarenes or a benzene ring.

As a result of this slight modification[26], the original Potts' classification of reagents for heterocyclic synthesis acquires the previously lacking harmony and completeness. For simplicity, we will name a reagent with *k* atoms and a given *E* or *N* nature of the terminal reaction centers as a reagent of a particular *EN* type (*E*-electrophilic, *N*-nucleophilic). Therefore, any reagent can be assigned to either 1,*k*-*EE*, or 1,*k*-*NN*, or 1,*k*-*EN* type. We should emphasize the main feature of reagent classification on the basis of *EN* type: this classification reflects the behavior of the particular reagent *in a given reaction only*. For example, hydroxylamine is of a l,2-*NN*-type in syntheses of isoxazoles, but it will naturally be of a 1,1-*NN* type in syntheses of N-oxides of pyridines. Acetone, depending on the reaction, may act as reagent of a 1,1-*EE*, l,2-*EN*, or 1,3-*NN* type. The flexibility of this sort of classification consists mainly of the following: the potential

ambiguity (multiplicity) of reactivity modes of a given reagent is compensated by the absolutism of definition of its *EN* type if the reaction mechanism is known.

3.2.2. Classifications of reagents according to the polarity distribution along the chain.

Additionally to the *EN*-classification (based on the polarity of the terminal atoms in reactants) there is another general classification of reagents based on the arrangement of all polar centers along the reagent's chain. Such arrangement may be either consonant (with alternating sequence of donor and acceptor centers along the chain) or dissonant (with violation of such an alternation). The initial terminology first suggested by Evans in 1971 was further used in the works related to the computer-assisted organic synthesis[27-33]. Initial idea of such a classification was claimed in 1920 by Lapworth in his principle of alternating polarities induced by a heteroatom along the chain[34]. (Lapworth classified reagents in similar sense using the terms "homogeneous" and "heterogeneous".) An analogous classification of reagents was discussed by Seebach in his reactivity umpolung concept[35]. He used the terms of "normal" structures (i.e., with an alternating sequence of donor-acceptor centers in the chain) and structures, containing an "umpolung" (inversion of polarity of some donor or acceptor carbon center in relation to the heteroatom at the end of the chain). The same classification principle (in terms of "conjoint and disjoint" structures) has been also discussed by Ho[36,37]. More references can be found in the author's review[38].

3.3. Statistical analysis of polar types of reagents used in pyridine ring synthesis

The existence of rational polar classifications of open chain reagents for heterocyclic synthesis leads to an intriguing problem. Suppose, we have a certain heterocycle and select a certain disconnection scheme of its synthesis. The question arises: can we arbitrarily choose the polarity of the open chain reagents (say, varying their *EN*-type or the consonant and dissonant properties)? Theoretically, there are no restrictions for polar structure of reagents, but in practice there are.

After examining the rich and extremely diverse material on the syntheses of six-membered heterocycles, McKillop and Boulton[39] concluded that in this field of synthesis the concept of *E*- and *N*- centers (readily identifiable in the original reagents) plays an important role: in many cases, electrophilic and nucleophilic centers *alternate around the ring*. Unfortunately, these authors never formulated clearly the exact content of the "alternation rule." Instead, without any further reference to the alternation phenomena, the cited review covered many examples of disconnection schemes with the participation of an electrophilic heteroatom, making unclear the initial thesis.

It is necessary to clarify this statement. Consider example of possible ring synthesis of the pyridine, the simplest representable member of the heteroaromatic azines. Let us analyze the following problem: how are reagents distributed by their *EN*-type in the ever reported syntheses of the pyridine ring? Are there any *polarity control* rules for the pyridine ring synthesis that can be definitively verified by careful statistical analysis of empirical facts?

Resolving of such a problem has required extremely precise criteria for the selection of the literature sources that are used, the choice of actual structures that are classed as "pyridines," and the methods of pyridine ring construction. Details and rigid restrictions we finally used are published[26,40]. As the result we have set up a unique computer database on those syntheses that are one- and two-component heterolytic cyclizations of maximally unsaturated pyridines, including those containing annelated alicycles[40]. Following the selection principles we classed as heterolytic cyclizations only those examples of cycloaddition and electrocyclization in which the influence of the substituents clearly indicated the *EN* type of the reagent. (The recyclizations, frequently used in design of pyridines, have been ignored.)

Every reagent used for the synthesis was classified according to its *EN*-type in given reaction. For the total 638 reactions the distribution of reagents to polar types was the following[26]:

Polar type of reagent(s)	Number of examples
(1,1-*NN*)+(1,5-*EE*)	130
(1,3-*NN*)+(1,3-*EE*)	289
(1,4-*EN*)+(1,2-*EN*)	81
(1,5-*NN*)+(1,1-*EE*)	8
(1,6-*EN*)	130

As it is evident, there are no examples of reagents that serve as 1,2- or 1,4-bielectro- or binucleophiles, and as 1,1- or 1,3-ambiphiles. Therefore, there is a real polarity control rule that connects oddness of the reactants chain and the polarities of terminal atoms. It is easy to understand that the nature of the terminal reaction centers of the reactants that are used for the synthesis of pyridines *can be established by truncating the hypothetical chain with alternating charges*. For the best of our knowledge there are no exceptions from *this* alternation rule.

The next generalization arises from the analysis of permissible position of the nitrogen heteroatom in the skeleton of "allowed" reagents. From total 638 reagents with nitrogen atom (as the precursor of the pyridine ring heteroatom) in 634 cases the N atom was either nucleophilic center or was located in consonant position to the polar (donor or acceptor) terminal atoms of reagent. In only 4 reported cases this "heteroalternation rule" (but not the initial alternation rule!) was violated.

An attempt to generalize the rule to the multicomponent syntheses of pyridines and more hydrogenated structures has resulted in the similar trend[26]. In some multicomponent synthesis of saturated structures the "forbidden" dissonant open chains (like alpha-haloketones of 1,2-*EE*-type or alpha-aminocarbonyl compounds of ambiphilic 1,3-*EN* type) were used. However, just for these cases the pyridine skeleton appeared as a by-product of the competitive formation of odd membered rings. Moreover, the

heteroalternation rule was strictly valid for one- and two-component synthesis of the quinoline[26], see Scheme 4. Perhaps the only exception is the above mentioned reaction 5 (Scheme 3), and it is remarkable that the main direction of this "forbidden" reaction is the formation of an odd-membered ring. (This is actually the Madelung synthesis of indoles, not of quinolines!) The rule is also applicable (although with more exceptions) for synthesis of the isoquinoline, and the exceptions are explainable[26]. Qualitative analysis of the polarity of reagents traditionally used in the synthesis of other azines (and their benzoderivatives) gives rise to very simple conclusion: the heteroalternation rule is most strictly valid for those six-membered rings where the heteroatoms are located in the alternating sequence around the ring.

The provided analysis defines more accurately the content of the alternation rule. However, it can be further generalized. Following Evans and others[27,30,31], let us expand the consonant or dissonant classification from open chains to the cyclic structures. The consonant/dissonant property of the cycle depends on the existence of odd and even chains between the heteroatoms and oddness of the ring. Odd-membered cycles (e.g. azoles) are always dissonant. Even-membered cycles (e.g. azines) may be either consonant or dissonant, depending on mutual arrangement of heteroatoms around the ring as well as polar exo-substituents. Thus, pyridine, pyrimidine, pyridone-2, and resorcinol are consonant, while pyrazine, pyridazine, 3-hydroxypyridine, and pyrocatechole are dissonant.) From this viewpoint, the observed literature trends for heterocyclizations clearly confirm the following thesis[41]: the predominant precursors of the consonant heterocycles are the acyclic consonant structures. This rule was confirmed even for the synthesis of carbocyclic benzene ring[42] by polar cyclocondensations. In turn, by combination of only consonant chains it is impossible to obtain the dissonant heterocycle, say 3-hydroxypyridine or pyrrole, and the dissonant precursors are required.

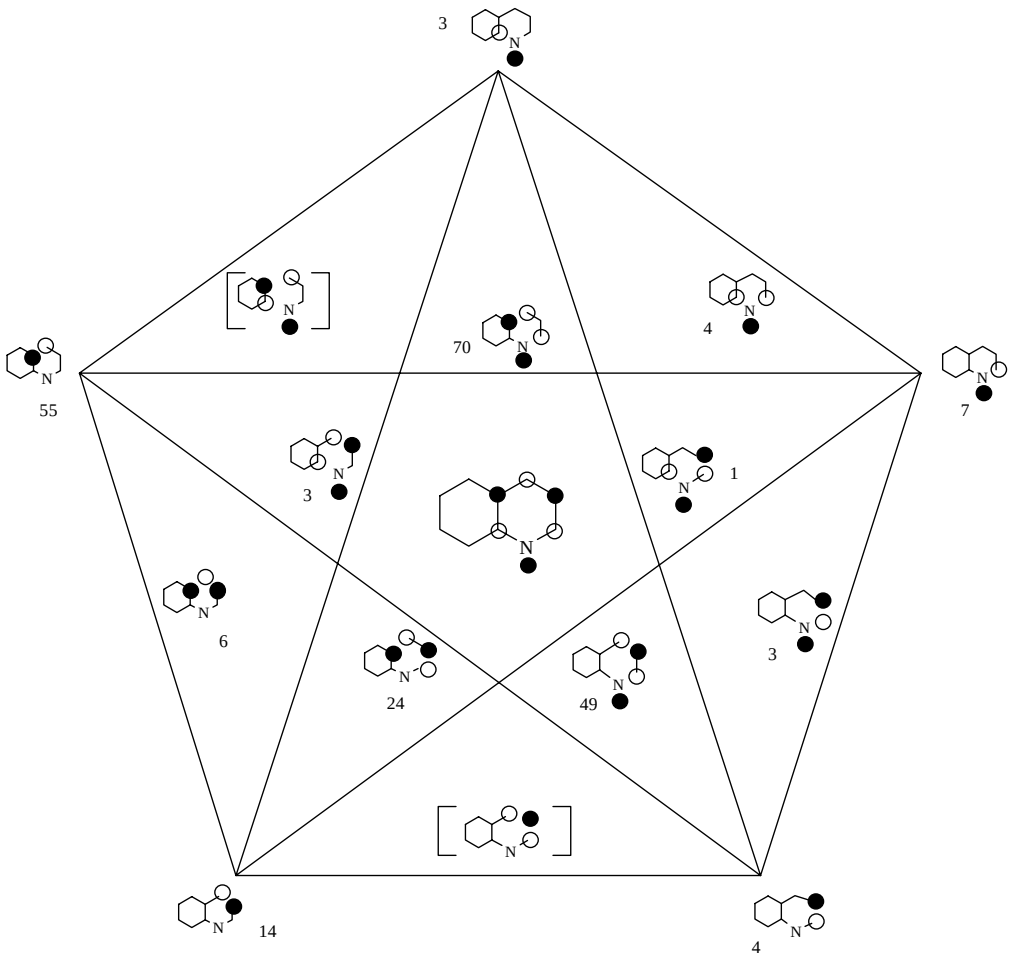
3.4. Polar cyclization graphs

Let us emphasize that the discussed rules accentuate combinatorial properties of the initial and final *structures*. Then the question is how to introduce the concept of electro- and nucleophilic centers into classical disconnection schemes, that, as we mentioned, are simple reaction graphs (cyclization graphs). We do not have the right to place E- and N-labels on the *mapping* until we have at our disposal counter-information on the nature of the reaction centers in the final product of cyclization. For consonant cyclic structures the arrangement of polarities around the ring is well known: it is simply alternation of electro- and nucleophilic centers (cf. known reactivity of positions 2, 3, and 4 of the pyridine nucleus). Hence, in the course of heterocyclizations from consonant chains to consonant cycles the initial "philicity" of centers (electro- or nucleophilicity) is conserved. Therefore, it is possible to define (only for the consonant heterocycles) the *polar cyclization graphs* as the new sort of reaction graphs by placing *E* and *N* labels in the corresponding cyclization graphs. Examples of such diagrams are presented in Scheme 4 for the quinoline ring, where the *EN* type of reaction centers is reflected by hollow and black dots.

Of course, the black and white (or *N* and *E*) labels can be arbitrarily located on any cyclization graph of any (consonant or dissonant) heterocycle, thus helping to enumerate possible combinations of reactants' polarities. Such diagrams, however, have nothing common to the idea of reaction graphs, and they are not heuristic in estimating the optimal polarity of the precursors. Only in the case of consonant heterocycles the polar cyclization graphs are the desired tool to predict optimal polarities of precursors for the unknown synthesis schemes: the required polarity is manifested by the diagram itself.

3.5. Heuristic utilization of the rules in novel quinoline synthesis

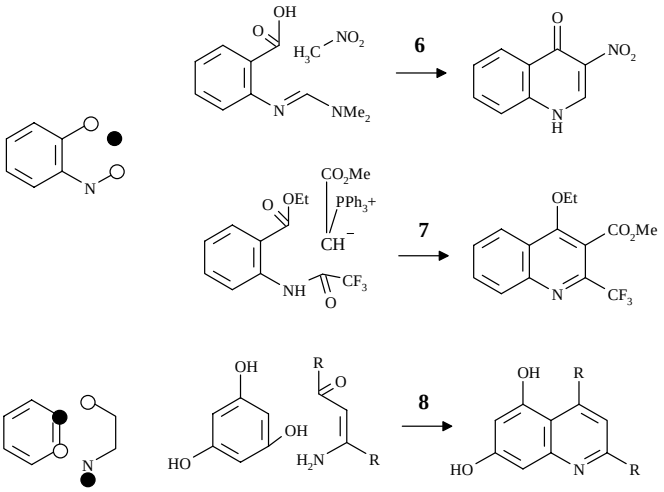
Early we elaborated the computer program "Heterocycland"[1a,b] capable for generating and ranking polar cyclization graphs for the given consonant target. Results of generation of such diagrams for one- and two-component synthesis of the quinoline ring are presented in Scheme 4. Comparison of theoretically possible and experimentally discovered synthetic pathways allows one to conclude, which reaction types are poor or better investigated. The numbers of known reactions near every diagram in Scheme 4 are taken from reviews[21,22,26].



Scheme 4. Actual number of examples and polarity of terminal atoms in reagents (*N* - black and *E* - white dots) used for one- and two-component cyclizations leading to the quinolines illustrated by the polar cyclization graphs. Unknown schemes are in square brackets.

It is rather elusive to consider that all possible combinations of consonant precursors have been exhausted in the synthesis of consonant quinolines. Thus, there are still no examples for two cyclization graphs, namely for the CCCNC + C combination of reagents (such cyclization graph is also still unknown for pyridine) and for the combination CCCN + benzene fragment.

The concept of polar cyclization graphs immediately suggests that for the first case the mono-carbon fragment should be of 1,1-*NN* type, and the other fragment must be of 1,5-*EE* type[43]. One possible combination of reagents was suggested to be the 1,1-binucleophilic nitromethane and 1,5-bielectrophilic amidine of anthranilic acid, see reaction 6 in Scheme 5. This predicted reaction -- the novel synthesis of the quinoline nucleus -- was indeed confirmed experimentally in 1993 by the author[43]. In 1994 another research group independently rediscovered the same cyclization graph with the same (1,1-*NN* + 1,5-*EE*) arrangement of polarities of reactants[44] as is shown by reaction 7 in Scheme 5.



Scheme 5. The uknown polar cyclization graphs for the quinoline synthesis (left), the recently discovered novel strategies of qunoline ring synthesis (reactions 6,7), and the predicted combination of consonant reactants for the still unknown reaction 8.

The "last" still unknown two-component cyclocondensation leading to the quinoline ring (more precisely, to the pyridine fragment of the nucleus) is the combination "CCCN fragment + benzene fragment". The heteroalternation rule predicts the 1,4-ambiphilic type for the CCCN fragment. (Such reactivity type is well known and is typical, say for enamine of 1,3-dicarbonyl compounds and enamides.) The benzene fragment, obviously, should serve as 1,2-ambiphile. Although it is somewhat difficult to imagine a benzene with opposite reactivity of atoms in ortho-position, such reactivity type is realized in the chemistry of phloroglucinole. (The last compound readily reacts with electrophiles and is also able to undergo substitution of a hydroxyl group with amines.) The final hypothetical reaction 8 is presented in Scheme 5, and the author offers the confirmation of this prediction to any interested heterocyclic chemist.

4. COMBINATORIAL MODELS AND POLARITY CONTROL FOR RING OPENING REACTIONS

One actual branch of organic synthesis is the use of heterocyclic structures as masked precursors of open chain structures with necessary arrangement of polar groups along the chain. The topic is well reviewed (see, e.g.[45-48]), and there are two still unanswered general questions associated with the ring opening reactions. The first one (related to synthesis planning) is: which heterocycles may serve as precursors of the given open chain fragment. The second question (related to the reaction planning) is: how to predict the direction of bonds cleavage for a given heterocycle with few heteroatoms and/or fused rings. An ideal solution of both problems would be the idea of reaction graphs with mapping of the starting cycle to the final open chain(s). However, except the prediction of the ring opening reaction 3 (see Section 2), there are no such approaches in the literature.

It is rather easy to suggest how the reaction graphs for ring opening reactions may look like. One can simply label the bond(s) to be broken in the initial cycle, say, by dashed line(s), and this will be just the oversimplified example of mapping of the atoms in the initial cycle into the atoms of final open chain. Defined in such a manner the ring opening graphs are simply reversal of the idea of the cyclization graphs discussed in the previous chapter. Such graphs can result useful in two respects: first, to enumerate all possible ring opening reactions (say, for a complicated heterocycles with many heteroatoms) and second, to compare similarity of such reactions for different heterocycles. However, a care should be taken in attempts to displace polarity labels on such reaction graphs.

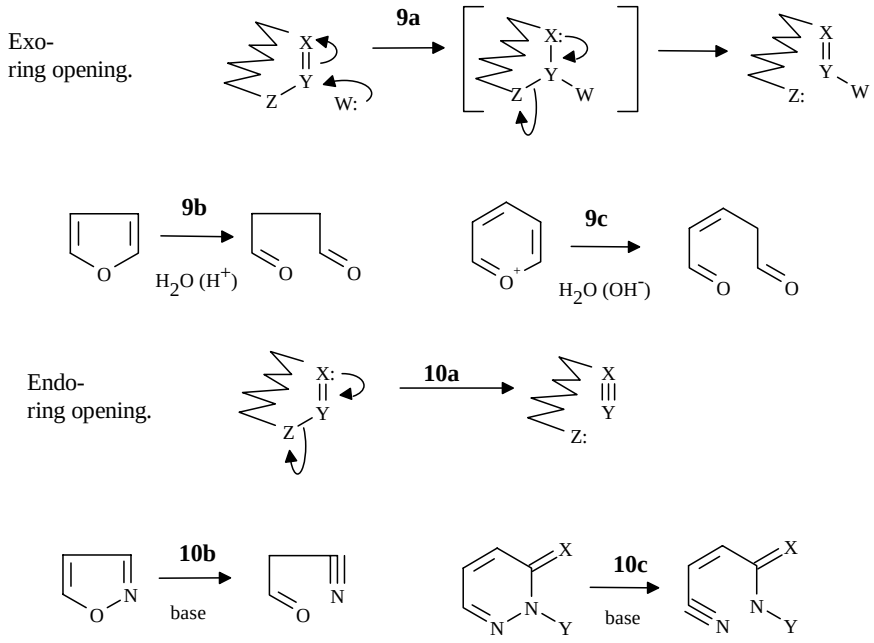
4.1. Correspondence of polarities in heterocyclic precursor and open chain structures

As we proved above in the Section 3, the global arrangement of polarities in the acyclic precursors and final cycle appeared important for better understanding of heterocyclizations. Hence, we may put reversed question: what is the correspondence in polarities of atoms in the initial cycle and the resulting open chain in the ring opening reactions? Let us avoid discussion of photochemical, oxidative and reductive processes of ring cleavage and consider only thermal polar reactions (usually under action of polar species) of heteroaromatic (may be fused) azoles and azines.

Developing the early idea of Stirling[47b], let us subdivide the polar ring opening processes into two groups[49]. Actually, the cleavage of a ring bond is accompanied by the formation of a new bond. This can be either a new bond between initial cycle and external (exo) group or a new bond between (endo) atoms of initial cycle. The first group will be called *exo-ring opening* (equation 9a) and the second group -- *endo-ring opening* reactions (equation 10a in Scheme 6). The exo-processes resemble usual nucleophilic addition/substitution reactions, while the endo-processes resemble the elimination reactions.

Polarity preservation in the exo-ring opening reactions. As a rule, the exo-ring opening reactions (cleavage of cyclic Y-Z) bond occur as the result of nucleophilic attack on the carbon atom (Y is carbon and

Z is heteroatom on Scheme 6). Of course, the carbon atom is the electropositive center of the polar bond C-Z. If the external nucleophilic center W is more electronegative than the carbon atom, then the polar nature of the cyclic atoms C and Z should conserve in the formed open chain. Hence, in exo-ring opening reactions the appropriate *dissonant or consonant type of an open chain should be the same as was in the initial cycle* [41]. Thus, hydrolysis of consonant pyrillium salt results in the consonant glutaconic dialdehyde, while dissonant furane under hydrolysis result in the dissonant open chain, see reactions 9b, 9c in Scheme 6.



Scheme 6. Examples of exo- and endo-ring opening and polarity changes via reaction with corresponding changes or conservation of the initial consonant and dissonant nature of a cycle.

Polarity changes in the endo-ring opening reactions. The endo-ring opening reactions often require strong bases and frequently lead to the cleavage of bonds between heteroatoms. Consider azine and azole with Y-Z bond (Y, Z - electronegative heteroatoms like N and O), for instance pyridazine and isoxazole. Both initial heterocycles are evidently dissonant. It is easy to imagine a ring opening with the cleavage of the cyclic bond Y-Z, as in examples 10b, 10c in Scheme 6, so that the atoms Y and Z of the open chain are the nucleophilic centers. Ring opening of the *dissonant* oxazole results in the *consonant* alpha-ketonitrile, while the *dissonant* pyridazine is converted into the *dissonant* gamma-aminonitrile.

In the two above examples the initial polar nature of the cycle is either conserved or changed *via* ring opening reaction. The only way to elaborate rational combinatorial model for this phenomenon is to consider the dissonance to be not only qualitative but quantitative property. Let us call a "polar conflict" an even chain between a pair of heteroatoms. The neighborhood of heteroatoms (even-membered “zero size chain”) is the particular case of the polar conflict. There is one polar conflict in isoxazole (indeed, only the bond NO is dissonant, while the NCCCO fragment is consonant), and there are two such conflicts (NN and NCCCCN even fragments) in pyridazine. Hence in the case of oxazole the polar conflict disappeared via ring opening giving rise to the consonant product. However, for pyridazine the ring opening causes disappearance of only

one conflict from two, and the final product is still dissonant. This "arithmetic" of conflicts and their changes is under our extensive current investigation (see[41,42]), and the just discussed examples illustrate particular importance of this idea for heterocyclic chemistry.

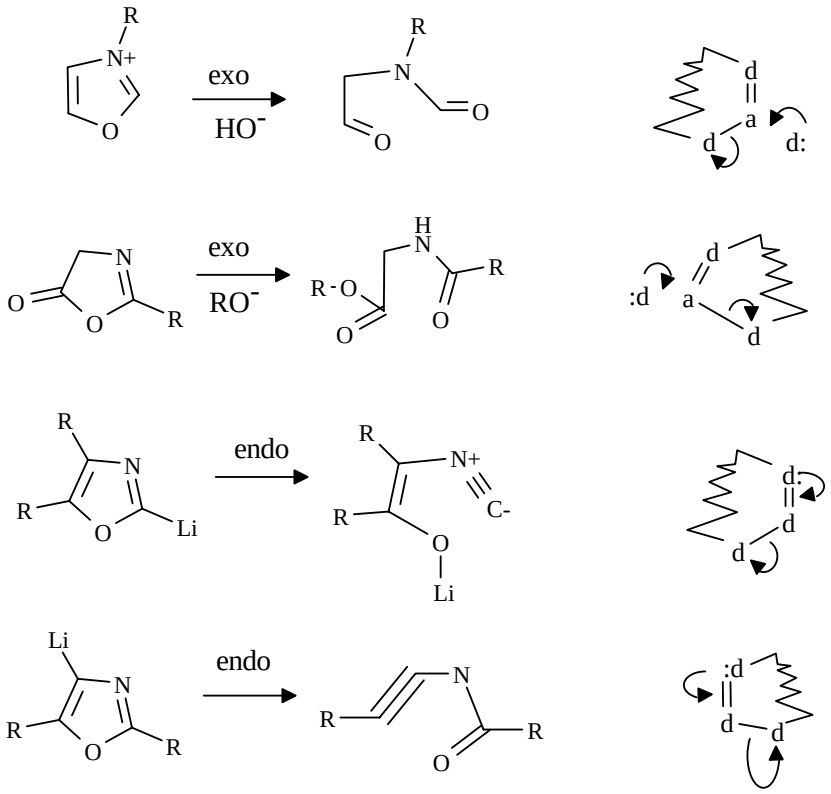
Difference in polarity changes via exo- and endo-ring opening reactions is essential. It is possible to put polarity labels on the centers of a ring opening graph (as the mapping) only for the case of exo-ring opening reaction.

4.2. Selectivity of azole ring opening reactions: the "polar triads" rule

A heteroaromatic structure (monocyclic or fused) with few heteroatoms may undergo cleavage of different bonds, and the type of cleavage may be either the endo- or exo-ring opening. Thus, two C-O bonds in oxazoles may undergo cleavage according to exo- or endo-type, and all four possibilities have been found, see Scheme 7. Although the topic is well reviewed for separate heterocycles, we failed to find simple, easily formalized rules, that may permit one to predict (or even estimate) the direction of bond cleavage for given heterocyclic structure. Early we have collected and analyzed available empirical data on the selectivity of ring opening for azoles (for azines the data are less representable). As we proved early, the fruitful approach to explain (and predict) selectivity of ring opening in azoles is to consider *local* environment of an atom in the cycle, i.e. to analyze the *polarity in triads* of cyclic atoms. (Although the exo-cyclic groups are ignored, one can reasonably estimate their influence on the adjacent atom of a cyclic triad.) Ranking the triads permits us to order the bonds in given azole structure according to the highest probability of their cleavage. The following algorithm can be suggested[49]:

As the first approximation, the bond to be broken was considered as the single bond. We also considered that any ring opening needs at least three ring atoms (numbered 1,2, and 3) responsible for this reaction. The single bond to be broken is the 1-2 bond (bond 2-3 may be either single or double). For the case of azoles (with two double and three single bonds) there are either 3 or 6 (depend on symmetry) possibilities to find such 1-2-3 triads. Every atom of such a triad is labeled as a donor (d) or an acceptor (a) center. Thus, any heteroatom (N, S, O - either of pyrrole or pyridine type) is (d) center, while carbon atom is usually (a) center. Of course, an exo-substituent may emphasize either (a) or (d) nature of the adjacent endocyclic carbon atom.

The possibility of ring opening according to exo- and endo- type depends on the *polar nature of all atoms in triad 1-2-3*, namely on the arrangement of (d) and (a) labels along a vector triad. Empirical data confirm the following qualitative rules[49]. Most probable consequence of atoms for endo-opening to be (d)(d)(d) (i.e. neighborhood of three heteroatoms), while the less probable to be the combination (a)(a)(a), (i.e. fragment with three carbons). Most probable consequence of atoms for exo-type of opening is another one, namely



Scheme 7. Different selectivity for endo- and exo-ring opening reactions observed for the oxazole ring qualitatively treated by different arrangements of polarities in triads (d - donor, a - acceptor center).

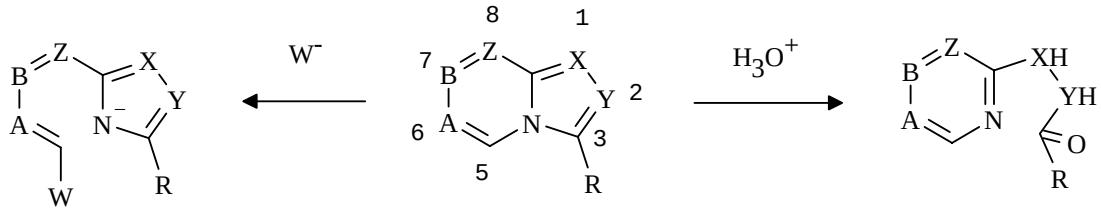
(d)(a)(d), while the less probable is the combination (a)(d)(a). Other possible combinations have intermediate priority, and for every such combination the numerical value of probability (one set of numbers for endo- and another for exo-case) has been assigned.

The discussed principles illustrate an example of the *local polarity control rule*. The suggested algorithm permits one (by selecting all triads and comparing their rating) to provide immediate conclusion on the most probable selectivity of ring opening. These simple rules have been implemented into the computer program FROG[49]. Input of a heterocyclic structure results in two rating lists of most probable bond cleavage: for exo- and for endo-ring opening. As an example, the highest rating suggested by the program for the selectivity of the oxazole ring opening coincides with the experimentally observed directions of exo- and endo-reactions (see Scheme 7). Hence, the program really selects (by rating criteria) just those ring openings that are found experimentally and may be also useful to predict previously unknown directions of ring cleavage.

4.3. Selectivity of ring opening reactions of bridgehead azolo-azines: the alternation effect in the non-alternant system

Due to our long year experience in the area of ring opening and transformations of bridgehead azolopyridines (particularly, indolizines)[50,51] we have been interested in the theoretical analysis of selectivity of their ring opening reactions. Cleavage of both 5- and 6-membered rings is possible for this class. The general trends of the exo- and endo-ring opening reactions of bridgehead azoloazines and the influence of

substituents on the selectivity of these reactions have been recently reviewed[52]. As it was proved, the exo-type of ring opening is the most representative for azaindolizines, and especially with the cleavage of a CN bond adjacent to bridgehead nitrogen atom, i.e. either C_5-N bond of the azine fragment or C_3-N bond of the azole ring. (The numbering we use corresponds to the conventional numbering of atoms in the prototype indolizine ring.) These two directions of rings cleavage are presented in general form in Scheme 8.



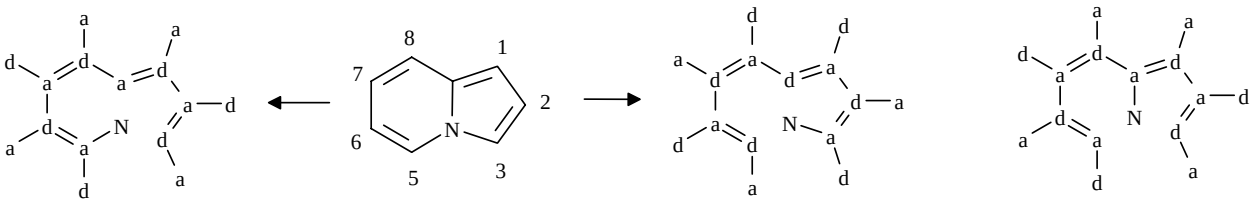
Scheme 8. Schematic representation of possibility of C_5-N (left) and or C_3-N (right) bonds cleavage in the series of bridgehead azoloazines. The ring opening products may be either isolated structures or intermediates in further rearrangements. (Numbering of atoms as in the parent indolizine nucleus).

Although both endo- and exo-types of ring cleavage are observed for this family, let us concentrate over the exo-processes. The observed qualitative trends of substituent influence on the selectivity of exo-ring opening are the following[52]:

(i) The C_5-N bond cleavage is promoted by the donating nature (azasubstitution, the presence of an external acceptor) of positions 6, 8, 1, and 3 in the bicycle and by the accepting nature (the presence of external donating and oxo substituents at positions 5, 7, and 2). The chain atoms closest to the opening assembly (primarily 5 and 6 and then 7 and 8 and so forth) have the greatest effect. Quaternisation at positions 1 and 3 also promotes this bond cleavage.

(ii) Cleavage of the C_3-N bond is observed more rarely. It occurs in the structures with donating groups at positions 2, 5, and 7 and the accepting nature of atoms 3, 1, 8, and 6. As in the previous case the polarity of the centers closest to the opening assembly (i.e. of atoms in the five-membered ring) is most significant.

Some observed trends can be easily explained with the help of polar triads rule, i.e. in terms of more or less preferable (d)(a)(d) fragments in the bicycle. However, it was somewhat difficult to treat the influence of polar substituents located far along from the bonds to be broken. A simple approach helpful to resolve this problem [52,53] is based on the idea of alternating effect of substituent along the parent indolizine framework. One can separate in the parent indolizine bicycle three types of chains A, B, and C (as shown on the Scheme 9). These chains are obtained by formal removal of two (from three) CN-bonds. One may assume that the nitrogen atom adjacent to the carbon tetraene fragment induces an alternation of the donor (d) and acceptor (a) centers along the chain. The arrangement of (d) and (a) centers in the chains A and C match each other, and this clockwise alternation will be called the direct alternation chain. The arrangement of polarities in the chain B is opposite to the cases A and C, and this counter-clockwise alternation will be called the reversed alternation chain.



Scheme 9. Qualitative model of substituent influence along the parent indolizine framework. The direct and reversed alternation chains help to understand the influence of donor and acceptor groups on the polarity of atoms C_5 and C_3, responsible for the ring cleavage in the azaindolizines.

One may assume that any polar substituent attached to the indolizine bicycle (heteroatom or exo-group) according to its own polarity should intensify one or another alternation chain. Thus, the direct alternation chain is intensified by heteroatoms and exo-acceptor groups in positions 6, 8, 1, 3 and exo-donor groups in positions 5, 7, 2. In contrast, the reverse alternation chain is picked out by heteroatoms and exo-acceptor groups in positions 5, 7, 2 and exo-donor groups in positions 6, 8, 1, 3. In turn, intensifying of one or another chain means increase of the positive charge on either C_5 or C_3 atom and, therefore, assist either C_5-N or C_3-N bond cleavage.

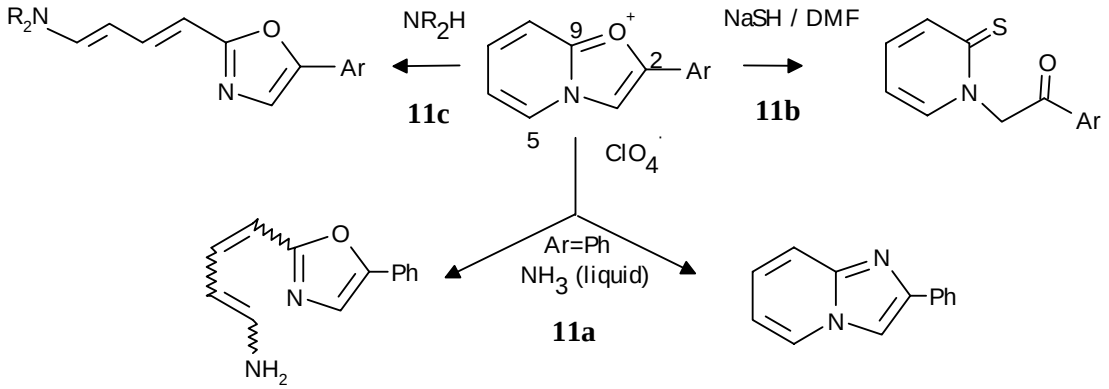
One may conclude that experimentally observed trends (i) and (ii) of substituent influence on the ability of C_5-N and C_3-N bonds cleavage can be quite clearly explained in the terms of intensifying of either direct or reversed alternation chains. Namely, the substituents that emphasize the direct alternation chain promote C_5-N bond cleavage, whereas the substituents that increase the reverse alternation chain assist C_3-N bond cleavage.

The semiempirical SINDO1 calculations[53] for the families of isomeric aza-, nitro-, and methoxy-substituted indolizines and azaindolizinium cations support this simple model. The changes of charges of skeletal atoms caused by the insertion of substituent displayed pronounced alternation in accordance with the simple model of strengthening of either direct or reverse alternation chain. The final confirmation of this simple rule follows from the analysis of the energies for pairs of 3- and 5-OH anionic sigma-complexes calculated by SINDO1 for isomeric azaindolizines. The data of energies for isomers display the same principle: the azaindolizines with a nitrogen atom at positions 6, 8, 3, 1 (increase of direct alternation chain) prefer to form the C_5-adducts, whereas those aza-substituted at positions 2, 5, 7 (increase of reverse alternation chain) favor to form the C_3-adducts.

4.4. Experimental study of unknown ring cleavage reactions

The models suggested in Sections 4.2 and 4.3 may serve as examples of *local* and *global polarity control rules* useful for explaining and predicting the selectivity of ring opening reactions. It was attractive to use these rules in the experimental search of unusual selectivity of ring opening reactions for a certain

heterocycle. Our choice was the bridgehead aromatic system of oxazolo[3,2-a]pyridinium cation (Scheme 10). Both fused rings have known capacity to undergo exo-ring opening, however the only reported reactions for this bicycle were the exo-cleavage of C_9-O bond for cation and of C_2-O bond for mesoionic 2-oxoderivatives[52]. The local triad rule confirmed the observed direction of cleavage and suggested possibility of exo-cleavage for any of three C-N bonds, although with lower rating. The common sense excluded the possibility of C_9-N cleavage (due to the loss of aromaticity of both fused rings). The global alternation rule (strengthening of the direct alternation chain in bicycle by oxygen heteroatom) favored to the C_5-N cleavage and excludes the possibility of C_3-N cleavage.



Scheme 10. Ambident reactivity of the oxazolo[3,2-a]pyridinium salts in the ring opening reactions.

The efforts, therefore, were concentrated over the discovery of the exo-opening of fused pyridine ring at the C_5-N bond. The quantum chemical calculations have predicted[54] that for O- and S-nucleophiles the energy of adducts favored exclusively the C_9-O ring opening, whereas the C_5-N cleavage could be awaited for nitrogen nucleophiles. Indeed, the use of liquid ammonia as reactant resulted in the mixture of products of opening of both 5- and 6-membered rings (reaction 11a in Scheme 10). On the other hand, the action of hydrosulfide anion caused cleavage of C_9-O bond (reaction 11b) Finally, the use of secondary amines gave high yields of aminobutadienes as the result of the desired selective C_5-N bond cleavage (reaction 11c in Scheme 10).

5. Recyclization Graphs - New Language for Description and Classification of Heterocyclic Ring Transformations

5.1. Early attempts to classify recyclizations

An important class of organic reactions that play major part in the entire heterocyclic chemistry are the ring transformations (or recyclizations) of heterocycles. These elegant reactions (frequently discovered by a lucky chance) often lead to heterocycles with unusual disposition of heteroatoms and substituents, or products not available by any other synthetic methods. Well-known examples of recyclizations are the "named" reactions[55] of Yur'ev, Zinke-Koenig, Hafner type, or the Dimroth, Cornforth, Boulton-Katritzky, Kost-Sagitullin rearrangements. Nowadays the ability to undergo ring transformations is found for the

overwhelming majority of known heterocycles, particularly for heteroaromatic cycles. The mechanisms of transformations of this sort are diverse, and it is often difficult to propose the structural scheme of bonds redistribution, by considering only the initial and final heterorings. Despite the enormous amount of factual material in this area and the profusion of reviews on this topic (see e.g.,[55]), one must conclude that there is no general rational classification of heterocyclic ring transformations in accordance with a structural principle. As a result, it is often difficult, for example, to establish the real degree of novelty of a recyclization or rearrangement declared as being "novel". Evidently, it is necessary first to define somehow the degree (or measure) of structural similarity of the existed reactions.

In most of reviews devoted to recyclizations the general question on the structural classification of such reactions has not arisen; only unique papers touched upon this problem. For instance, in the classical monograph[55a] of Van der Plas the heterocyclic recyclizations were not classified, but conveniently ordered, according to the size of initial and final cycles, and the nature and number of heteroatoms. Meanwhile, in other reviews the idea of structural classification has appeared in different forms. Schwaika suggested[56] the most important structural feature for comparison of recyclizations to be the size of fragment of initial ring incorporated into the final heterocycle, and similar principle was applied to classification of various pyrillium ring transformations[57]. In the review of L'abbé[58] convenient classification of azoles rearrangements was proposed, taking into account another criterion - the size of the side chain of initial heterocycle, that is incorporated into the final cycle. (Following this idea, the size of side chain for the Dimroth, Cornforth, and Boulton-Katritzky rearrangements should be considered as 1, 2, and 3, respectively.) In the other reviews only some general reaction schemes of ring transformations for particular heterocycles (or some families of analogous heterocycles differed by heteroatoms or annelated rings) were discussed.

We would like to mention, that even in the most developed of the approaches cited above chemists utilized the principle of structural similarity of the molecules (or their different substructures), but not the similarity of reactions. As we mentioned in the Section 2, the reaction graphs are just the appropriate objects for comparison the similarity of reactions. In our early series of papers[59-62] we suggested novel type of reaction graphs for description of heterocyclic rearrangements and ring transformations. These graphs -- the recyclization graphs (or graphs of cyclic bonds redistribution) -- have been utilized to classify recyclizations in the hierarchic system, to compare structural similarity of recyclizations, and to predict interesting examples of recyclization of previously unknown structural types.

5.2. Definition of recyclization graphs

Let us briefly recall our definition of recyclization graphs. We call recyclization (or ring transformation) any transformation of a heterocycle, that include the steps of ring opening and ring closure in any sequence. We shall consider mainly monocyclic systems, and among condensed heterocycles - only those containing one joint bond between the annelated rings (as it is in indole or acridine). Let us call ring transformation to be *simple ring transformation* (or SRT) if: (a) there is no formation of any transient cycles via recyclization,

except the final one; (b) there is no permutation of atoms or groups via ring transformation; (c) the initial cycle is transformed into no more than one cycle of the final structure. Most of known recyclizations are SRTs, particularly, the large family of ANRORC-reactions; boundaries of the term are discussed early[59,60]. Below we shall be limited only by SRTs.

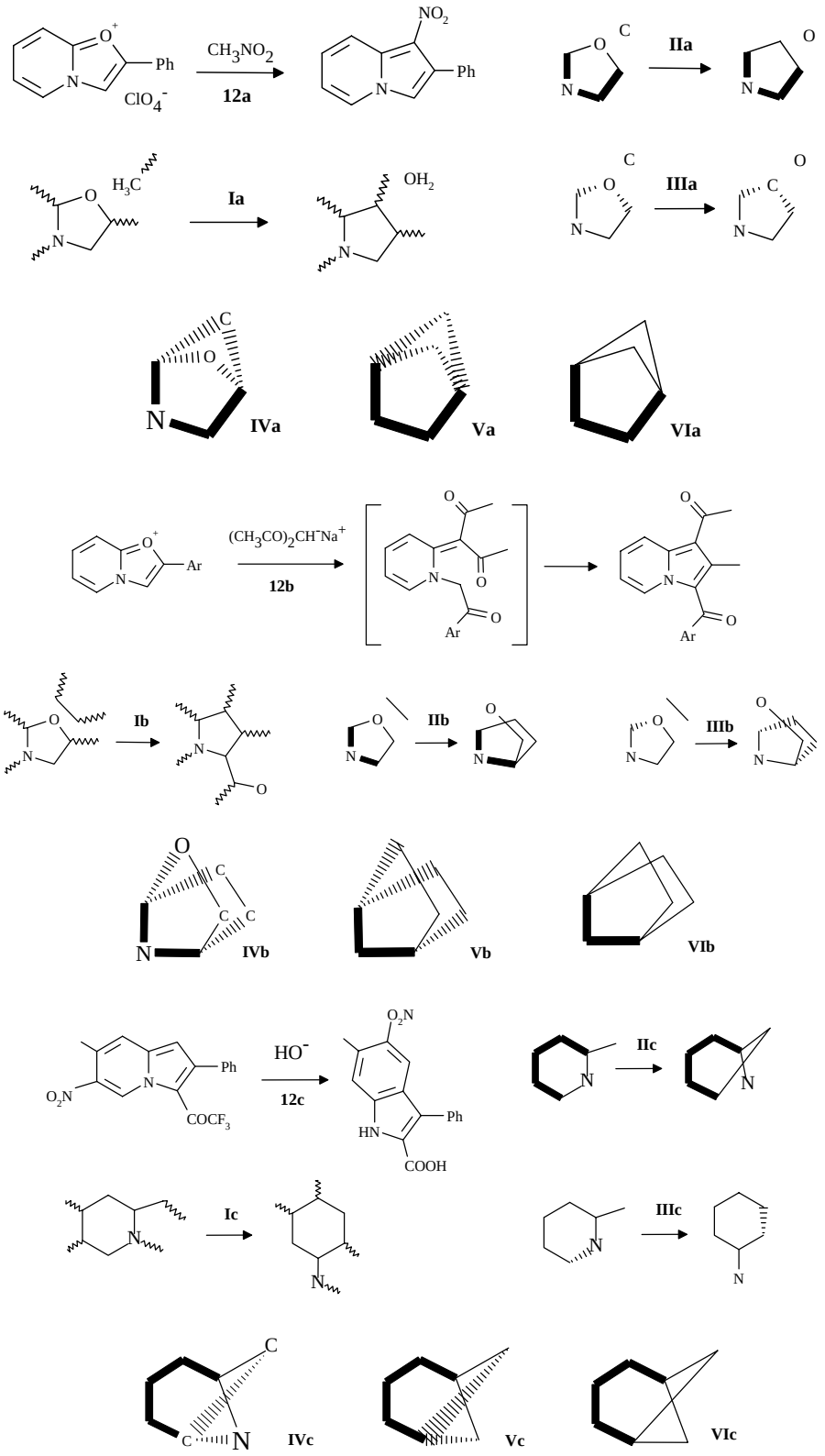
The main idea of our approach is rather simple. Since the mechanism of a recyclization is established, the skeleton of the initial cycle can be easily found among the atoms and bonds of final products and vice versa, the skeleton of the final cycle can be found among the atoms and bonds of the starting reagents. Paying all the attention only to these cyclic skeletal substructures (and ignoring all other details of molecular structure), we can get significant simplification of the chemical equation. The reaction mechanism establishes the correspondence between matching atoms and bonds of the initial and final cycles. Therefore, just these cyclic (initial and final) substructures can be chosen for superposition to define the graph of ring transformation reaction.

Recyclization graph as superposition of molecular graphs. Let us give more strict definitions following Ref. 60. Let us define for any SRT (with the mechanism known beforehand) two types of graphs - the molecular graphs (that are not coincide with initial and final structures and are determined only by the given type of transformation) and the recyclization graph (for any reaction that is SRT). Let us define the molecular graph M_S of starting reagents as the set of vertices and edges, that correspond only to those V atoms and R_S skeletal bonds that either exist in the initial cycle or are incorporated in the final cycle. By analogous manner, the molecular graph of final products M_f (with V vertices and R_f edges) contains the same V atoms and only those skeletal bonds, that either exist in the final cycle or were presented in the initial cycle. Let us keep the symbols of heteroatoms as the labels of vertices in the graphs M_S and M_f and omit all other symbols (like hydrogen atoms, multiple bonds, and any substituents including annelated rings). Although the same V vertices are chosen to construct both graphs M_S and M_f , the number and/or distribution of edges in these graphs is obviously different.

Let us enumerate the vertices of M_S -graph; then the mechanism of SRT immediately permits us to enumerate by corresponding numbers the matching vertices of the M_f -graph. Let us make mental superposition of the structures of graphs M_S and M_f according to the matching vertices (and edges) with identical numbers. The pairs of vertices with the same number should be identified into one new vertex, as well as the pairs of edges - into one new edge. Let us define such superposition as the new recyclization graph with V vertices. Let us designate the edges of recyclization graph (presented also in the graphs M_S and/or in M_f) by the following manner:

(a) by the *dashed* line (edge) if the edge is presented only in one (but not another) M-graphs;
(b) by the *bold* line if the edge belongs to both (initial and final) cycles;
(c) by the *usual (solid)* line if the edge belongs to only one cycle of M-graphs.

Examples. In order to illustrate the idea of recyclization graphs in pictorial form let us use examples of SRTs discovered by the author's research group[63-65] and presented in Scheme 11.



Scheme 11. Assignment of recyclization graphs to reactions 12a-12c. Ia-c: Choice of skeletal atoms that appeared in both (initial and final) cycles. IIa-c: M-graphs with bold fragment common to both cycles. IIIa-c: The cyclization and ring opening graphs. IVa-c: G_2-graph as the result of superposition of M-graphs. Va-c: G_1-graphs. VIa-c: G_0-graphs. -graphs.

For every reaction 12a-12c the pairs of M-graphs II and the resulting recyclization graph IV are shown. In the first reaction[63] the 1,1-binucleophilic carbon atom substitutes oxygen atom in the oxazole ring, and both M_S and M_f graphs are disconnected. (The result is the appearance of four dashed edges in the recyclization

graph IVa.) For the second reaction[64] 12b formally the same conversion of one heterocycle to another occurs. However in this case the incorporated 1,2-ambiphilic fragment consists of two atoms, and initially disconnected M_S graph is converged to the connected M_f graph with quite different recyclization graph. Of course, in both reactions the fused pyridine fragment is considered as an "inessential substituent" in the 5-membered ring and, therefore, is not shown in the recyclization graph. In the third example[65] of "haloformic recyclization" 12c the transformation of the indolizine ring to indole (actually, of the pyridine fragment to benzene) is followed by the haloformic reaction. However, the resulting recyclization graph IVc does not reflect any changes in the "inessential" substituent, since only atoms of initial and final six-membered rings are used to construct the diagram.

From the formal viewpoint, the recyclization graph is nothing but a labeled graph. The labels are on the vertices (symbols of the heteroatoms) and on the edges (three above mentioned types or "colors" of edges - dashed, bold, and solid). Including of so many labels into the same recyclization graph permits one to follow the changes in any cyclic skeletal bond of heterocycle via SRT at the same -- united -- diagram. As it follows from the above definition, the structure of recyclization graph does not depend on the order of superposition of M-graphs; hence, the recyclization graphs of direct and reversed reactions coincide. With evidence, skeletal pairs of M-graphs can be immediately restored from a recyclization graph (by drawing two diagrams and reversed procedure of restoring/removal of corresponding cyclic bonds from dashed lines in each of them).

Recyclization graph as superposition of the ring opening graph and cyclzation graph. Let us compare recyclization graph with the discussed above ring formation graphs (disconnection scheme) and ring opening graphs. The disconnection scheme (the superposition of cyclic and acyclic molecular graphs) is evidently a labeled monocycle with two sorts of labeled edges (formed and retained via cyclization). The ring opening graph is the exact reversal of the ring formation graph, and is again the labeled monocycle. In contrast, the recyclization graph is the superposition of a cycle to cycle, so it is by definition a bicycle (since the fragment of one cycle is identified with fragment of another cycle).

Actually, the mechanism of recyclization is the combination of ring opening and ring closure (generally, in any sequense). Each step, as we discussed above, can be described by a graph of reaction, namely by ring opening graph and by cyclization graph, where only the cyclic bonds (formed or broken) are presented. Therefore, the total redistribution of cyclic bonds -- recyclization graph -- is the *superposition of these two reaction graphs*. This is evident from Scheme 11, where the pair of reacton graphs IIIa-c is shown for every reaction 12a-c. Such an alternative definition of recyclization graphs is extremely important. First, we may "extract" not only skeletal M-graphs, but also elementary steps from the final superposition diagram. Second, the polarity rules, separately defined for the cyclization graphs and ring opening graphs, can be jointly utilized for understanding the polarity control via recyclizations.

The idea of superposition of two reaction graphs into a new one allows us to simplify the procedure of obtaining the recyclization graph. Let us draw the skeleton of the initial cycle only with that side chain (or a

disconnected external fragment) that will participate in the final heteroring. Let us draw the ring opening graph, indicating by dashed line(s) those bond(s) that are broken in the initial cycle. Let us draw at the same diagram the cyclization graph by simply adding new bonds (also labeling them by dashed lines) that are the novel bonds of the forming cycle. To emphasize the fragment common to both cycles let us also label it, say by bold line. The resulting diagram is the recyclization graph.

5.3. Recyclization graphs and the structural similarity of reactions

The bicyclic feature of any recyclization graph is essential. It is not a problem to compare labeled monocycles (for heterocyclizations or ring opening processes), thus establishing kinship or difference of the corresponding reactions. It is, however, somewhat difficult to compare labeled bicycles in order to establish classification of recyclizations.

Recyclization kinds. Analogously to cycizations, the recyclizations can be multicomponent, and this is the first criteria for classification. M-Graphs are connected graphs only in the case of rearrangements. In all other cases (i.e. for reactions where an external reagent is included into the final cycle and/or a fragment is eliminated from the initial cycle) at least one of M-graphs should be disconnected and consists of more than one component. The number of components in M-graphs is the first natural criteria to classify SRTs into kinds. Designating the initial and final heterocycles by symbols A and B, and the external reagents (or eliminated fragments) -- by symbols X,Y..., one can immediately find three main kinds of recyclizations, presented by the following schemes:

A => B (i.e. rearrangements);

A + X => B or A => B + X (reactions with elimination *or* incorporation of ring fragments);

A + X => B + Y (reactions with elimination *and* incorporation of ring fragments).

Recyclizations of the same kind evidently have the same number of dashed edges in the structure of their recyclization graphs. Examples 12a-12c are of different kinds, and the last reaction has the same kind as any rearrangement.

Recyclization families, classes and types.

Comparison of various recyclization graphs for transformations of azoles[59,60] and azines[61,62] led to conclusion, that recyclization graphs of SRTs differ from one another either by the topological structure of graphs, or by the number and/or distribution of labeled vertices and dashed edges. These characteristics have been used as the most important ones to establish the general structural similarity of SRTs. The following three different types of recyclization graphs have been introduced[60]:

G_2-graph is the initially defined recyclization graph with labels on vertices and edges.

G_1-graph is G_2-graph where the symbols of heteroatoms are omitted.

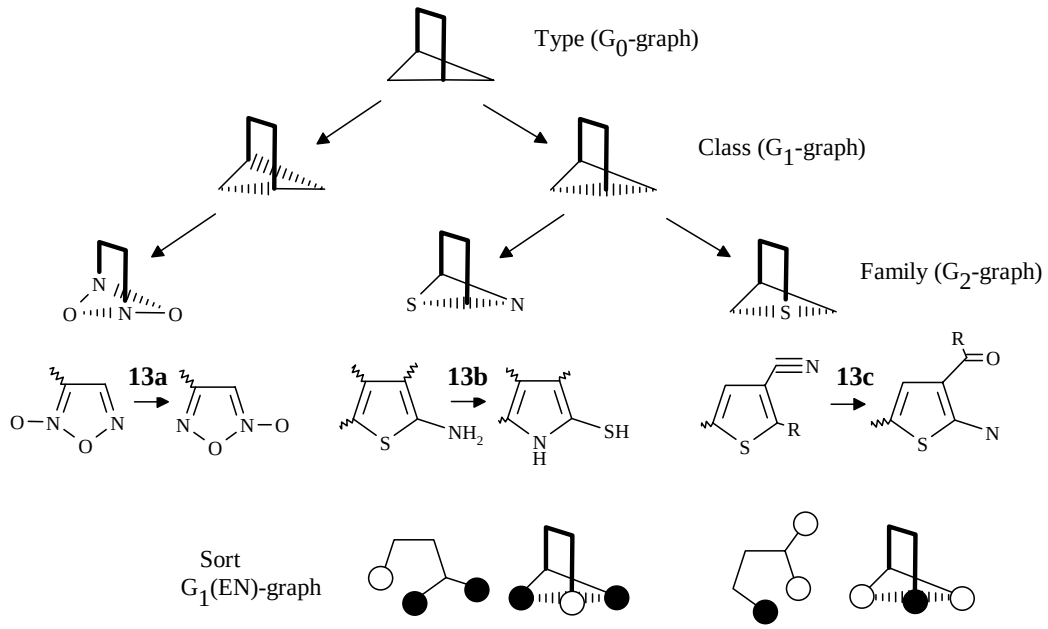
G_0-graph is G_1-graph where the dashed labels of edges are omitted.

Examples of G_0-, G_1- and G_2-graphs (diagrams IV, V, and VI) are shown for recyclizations on Scheme 11.

The identity of G_2 -graphs corresponds to the closest structural similarity of SRTs; let us say in this case that the SRTs belong to the same *family*. The family is determined uniquely by the skeletons of initial and final heterocycles and the position of heteroatoms (i.e. by the pair of M-graphs). Theefore, any variation in the degree of unsaturation, tautomerism, presence of any substituents (including condensed rings) should keep the reaction at the same family of SRTs. Another type of structural similarity of SRTs is the identity (more strictly, isomorphism) of their G_1-graphs; let us say SRTs with the same G_1-graphs belong to the same *class*. In classes the heterocycles are different, but the structural diagram of their recyclizations (without heteroatoms) is the same. At least, let us say reactions belong to the same *type* if their G_0-graphs coincide. Recyclization type is the crudest similarity level. Evidently, reactions of the same type have the same bold-labeled fragment in the structure of their G_0-graphs, i.e. the fragment presented both in the initial and final cyclic structures. The idea of classification into types is important if one compares SRTs of different kinds, say inter- and intramolecular recyclizations. Many illustrations of similarity and dissimilarity of famous recyclizations are collected in[59-61]. Particularly, the Dimroth-, Cornforth-, and Boulton-Katritzky-type rearrangements of azoles (see early classification[58]) belong to different types, while the furoxane rearrangement differs from the Dimroth-type transformation in the class, not in the type (see Schemes 12, 13).

The hierarchical classification of recyclization graphs. The graphs G_0, G_1, and G_2 are combinatorial objects: all they are bicyclic graphs, that are differently labeled. The manner of labeling can be treated in terms of a hierarchy[60] (see Scheme 12). Indeed, at the top there are G_0-graphs. Next level is formed by G_1-graphs, that are labeled G_0-graphs, and the labels are dashed edges that can be differently placed on the solid edges of G_0-graphs. On the lowest level there are G_2-graphs, that are labeled G_1-graphs with the set of heteroatoms as the labels. From the last level the corresponding pair of M-graphs can be immediately restored, and familiar equation of recyclization can be obtained (of course, only with skeletal atoms of cycles and heteroatom labels). The symmetry of a graph on higher level rigidly determines the possibility of non-equivalent labeling at the lower level.

The suggested idea of the hierarchical classification seems very fruitful: all theoretically possible skeletal equations for SRTs can be once and forever enumerated and classified by arranging their recyclization graphs into structural types, kinds, classes, and families. The classification, in turn, can be used as an organizational principle to create database on known recyclizations, to establish degree of structural similarity between SRTs and real novelty of discovered examples, and to assist search of really new recyclizations in respect to the chosen hierarchical level. This ideal project has indeed been realized in practice[1,60] with the help of our computer program GREH (Graphs of REcyclizations of Heterocycles). The program is easily available from WWW (see Ref. 1b).



Scheme 12. Illustration of the hierarchic classification of recyclizations (see text). For reactions 13b and 13c (labels: the polarity of centers in open chains and the corresponding G_1(EN)-graphs are shown black for *N*-centers, white for *E*-centers).

The enumeration and codes for recyclization graphs. The concept of recyclization graphs was further developed in few directions. Particularly, two combinatorial problems has been resolved, namely:

(i) How to assign a specific code to every recyclization graph, so that the chemist can immediately call any recyclization without referencing to the diagram.

(ii) How much ring transformations are possible for every branch of the classification tree

Solution of the first problem lied in the assignment of a unique code for every type and class. The unique code for G_0-graph is the vector XYZ (X and Y are the size of initial and final heterocycles and Z - number of vertices in the bold-labeled bridge of G_0-graph). The codes for recyclization types of reactions 12a-12c are 554, 553, and 665 respectively. The codes for classes or G_1-graphs require precise definition of arrangement of the dashed edges in the ancestoring graph G_0. For rearrangements there are only two dashed edges. The position of dashed edges in the smaller cycle of the G_1-graph is indicated in the following manner: the edge closest to the bold-labeled bridge should be designated by letter a, the next one - by letter b etc. The solid edges in the greater cycle should be numbered analogously by the same symbols a,b,c..., starting from the edge adjacent to the a-edge of the smaller cycle. The resulted code for rearrangements looks like XYZ-$i_X j_Y$, (i and j - relative position of dashed labels in cycles X and Y) and the above example of recyclization 12c has the code 665-aa. For recyclization graphs of various kinds the notation XYZ-($i_X j_X$)($i_Y j_Y$) was suggested[62]. The letters in the brackets (defined in the manner similar to rearrangements) reflect the relative disposition of broken/formed bonds in the initial and final cycle. For SRTs 12a, 12b the codes of classes are 554-(ab)(ab) and 553-(a)(ac).

The problem of enumeration of recyclizations was resolved in different ways. Combinatorial formulas for enumeration of recyclization classes, and families have been suggested; the formulas were obtained by

application of the Cauchy-Frobenius-Burnside lemma to this specific sort of graphs. One particular result of such enumeration is that for combinations of three heteroatoms (C,N,S) the number of theoretically possible Dimroth rearrangements for azoles is 486. In addition to the enumeration formulas, the full list of all possible codes for recyclizations of various kinds for 5-, 6-, and 7-membered rings was generated with the help of computer program for those SRTs with no more than 4 bonds changed via reaction. Therefore, this was the desired full structural classification of any (known and still unknown) SRTs. Meanwhile, the most important result was the program GREH, that permits chemist to visualize the branches of the classification tree.

How the program GREH works. The program GREH was designed[60] in order to generate, visualize, and investigate the SRT hierarchic tree. Each new branch of the tree is generated when the user attempts to display it. All the branches are stored in separate files on the hard disk for the subsequent viewing and search. For the user's convenience, recyclization graphs are shown only at the top level (G_0-graphs); all the deeper levels present information as differently labeled pairs of molecular graphs. Thus, starting from any chosen G_0 graph the user can go down the levels with the unique possibility to scan all pairs of molecular graphs on the given level and to search the necessary (hetero)cycle and its side chain for reactants and products. At the bottom level the degree of saturation is varied, and it is possible to edit the structures and add references, i.e. to create the database entries. After such an input of a recyclization the information is saved in such a manner that on every higher level it is clear how much data exists in the lower level(s). This methodology permits one to immediately conclude on the real degree of novelty of any SRT in respect to its type, family, and class.

Degenerated rearrangements. The model of recyclization graphs has appeared to be extremely useful for the description and enumeration of the specific type of SRTs -- the degenerated and quasi-degenerated heterocyclic rearrangements[62]. One example of such reaction is the well-known (quasi)degenerated Dimroth-type rearrangement reported for many heterocycles, particularly for 1-alkyl-2-iminopyridine derivatives, imidazo[1,2-a]pyridines etc.[19,20,52,55]. Specific of recyclization graphs G_1 and G_2 for this case of SRTs is that the initial skeletal structures (M-graphs) are isomorphic. Therefore, superposition of two isomorphic M-graphs results in the diagram that with necessity should have the certain elements of symmetry. As we proved early, *only three types of symmetry groups* (more precisely, the automorphism groups of a graph) are permitted for these recyclization graphs. These symmetry rules have been used to enumerate all theoretically possible examples of degenerated rearrangements for 6- and 5-membered rings with one heteroatom[62]. Analogous atlas for the still unknown degenerated rearrangements of quinoline was discussed recently[66].

Recyclization sorts. In many cases the SRTs are polar on nature. The opening of the initial ring occurs as the heterolytic exo-process of bond cleavage (see Section 3), and the formation of novel ring is the usual heterocyclization between appropriate *E*- and *N*-reaction centers. One can consider the arrangement of polarity labels in the initial cycle, open chain intermediate (real or hypothetical), and final cycle. If in all cases the polaritiy is the matching property, then the appropriate labels may be included into the structure of recyclization graphs.

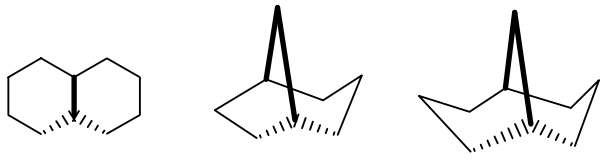
Therefore, it appeared reasonable to introduce one more type of reaction graphs -- G_1(*EN*)-graphs[59-61], that are the G_1-graphs with *E* and *N* labels on the vertices. Examples of G_1(*EN*)-graphs are presented in Scheme 12. These G_1(*EN*)-graphs can play the same role as the above discussed polar cyclization diagrams. Of course, these graphs are combinatorial objects, and they can be arranged as an independent level of the hierarchical tree.

5.4. The model and reality

Actual arrangement of SRTs into types and classes. In order to understand actual distribution of SRTs into types, classes, and families, and with the goal to precise the criteria of "novelty" of recyclizations we performed analysis of the literature data. We have limited ourselves by the simplest cases of SRTs - the heterocyclic rearrangements of azoles to azoles, azines to azines, and interconversion reactions between azoles and azines, according to the strict limitation "one cyclic bond is broken, one cyclic bond is formed". Following the logic of the hierarchic classification it was extremely interesting to understand what types and classes of SRTs are still unknown.

For rearrangements of azoles the comprehensive analysis was published[59], and it resulted in conclusion that all possible types (G_0-graphs) have already been discovered. For interconversions of azines and azoles there are few recyclization types that have never been realized[61]. These still unknown types have the codes 563, 663, and 662 (see definition above), and their diagrams are presented in Scheme 12.

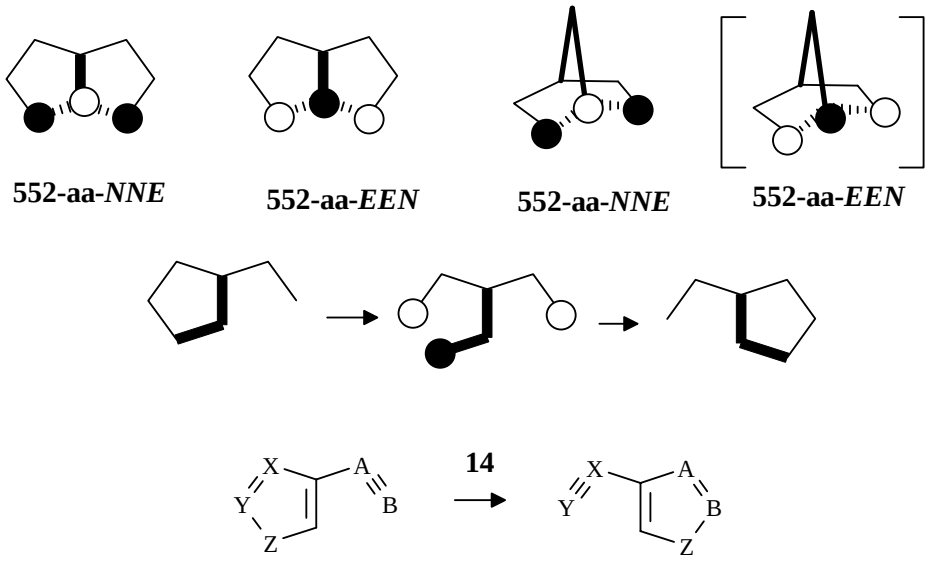
Analysis of rearrangements' distribution into classes (G_1-graphs) resulted in rather surprising picture. For overwhelming majority of all rearrangements *the unique predominating class has the code* XYZ-aa. By other words, in the pattern of bonds redistribution for rearrangements the formed and broken bonds prefer to be adjacent[59-61]. It is noteworthy, that the letter code aa for G_1-graphs corresponds to some unique feature of recyclization graphs: only for this code the total number of reaction centers (between which the skeletal ring bonds are redistributed) is equal to three. (It is clear, that for any other permitted combination of letters in the code the total number of the centers should increase up to four.) Hence, the high frequency of aa-classes appearance among the heteroaromatic rearrangements seems to follow from peculiar magic rule of "minimal number of reaction centers". The violation of this rule may cause the concurrent processes and formation of by-products, for example due to intermolecular condensations, oligomerizations etc. Although some classes with the code other than XYZ-aa are known, they appeared rather rarely[59]. To the best of our knowledge, among the classes with the aa-code there are three still unknown classes presented in Scheme 13.



Scheme 13. Unknown types of SRTs with aa-code for rearrangements of 5- and 6-membered rings.

Prediction of unknown sorts of rearrangements. As we mentioned above, it is difficult to suggest a new rearrangement for azoles at the top of hierarchical tree: all the types and all aa-classes are already discovered. Then the question is: how the azole rearrangements are distributed in sorts?

As it follows from the combinatorial considerations, any G_1-graph with the code XYZ-aa can give rise only to the pair of G_1(*EN*) graphs: the possible structures should contain either two *N* and one *E* label or vice versa. (Indeed, either *E*-center is rearranged between two *N*-centers, or vice versa, the *N*-center is redistributed between two *E*-centers.) Let us add these indices (either *NNE* or *EEN*) to the codes of G_1-graphs in order to establish the codes of recyclization sorts. The rearrangements of the first *NNE*-sort turn out to be more widespread in azoles' chemistry, and they are also better studied. In particular, the well-known examples of the Dimroth (see Scheme 12), Cornforth, and Boulton-Katritzky rearrangements have the codes 552-aa-*NNE*, 553-aa-*NNE*, and 554-aa-*NNE*, see diagrams in Scheme 14.



Scheme 14. Examples of some known and unknown recyclization sorts. For unknown sort (shown in square brackets) the skeletal structures and arrangement of polar centers in the initial and final cycles (and intermediate open chain) can be restored from the recyclization graph, and the concrete reaction 14 of unknown sort can be predicted[60].

Two from these three named reactions have “opposite” sort, and examples of 552-aa-*EEN* and 554-aa-*EEN* rearrangements can be found in literature. However, the sort, opposite to the Cornforth rearrangement, with the code 553-aa-*EEN* is still unknown. It is easy to extract from the code the total structural equation and the appropriate polarities required for the design of recyclization (see Scheme 14) and hence, to predict the novel elegant rearrangement sort illustrated by reaction 14. The choice of various concrete initial heterocycles able to undergo the rearrangement 14 was discussed early[60].

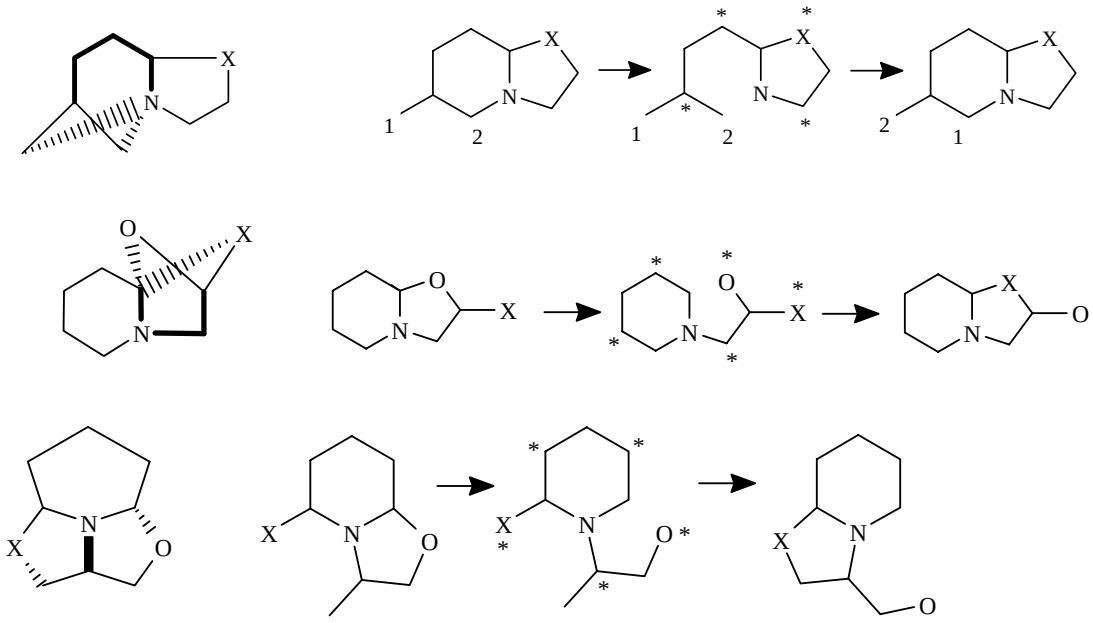
The sorts of still unknown rearrangements “azole-azine”, “azine-azole”, and “azine-azine” (for which we failed to find at least one reference in the literature), are the following [61]: 563-aa-*NNE*, 563-aa-*EEN*, 562-aa-

EEN, 653-aa-*NNE*, 663-aa-*NNE*, 663-aa-*EEN*, 662-aa-*EEN*, 662-aa-*NNE*. These codes are just the predictions of novel recyclization sorts recommended for experimental studies.

5.5. Computer-assisted predictions and experimental discovery of unknown recyclization families

Our interest in the experimental chemistry of the bridgehead azoloazines stimulated our efforts in predicting novel examples of recyclizations in this class with the help of program GREH and to verify the predictions experimentally. The bridgehead azoloazines (particularly the heteroaromatic cations) seemed to be attractive models for study new recyclizations, due to their ability to open either 5- or 6-membered ring (see ref. (52) and Section 4.3) and variety of modes for ring closure.

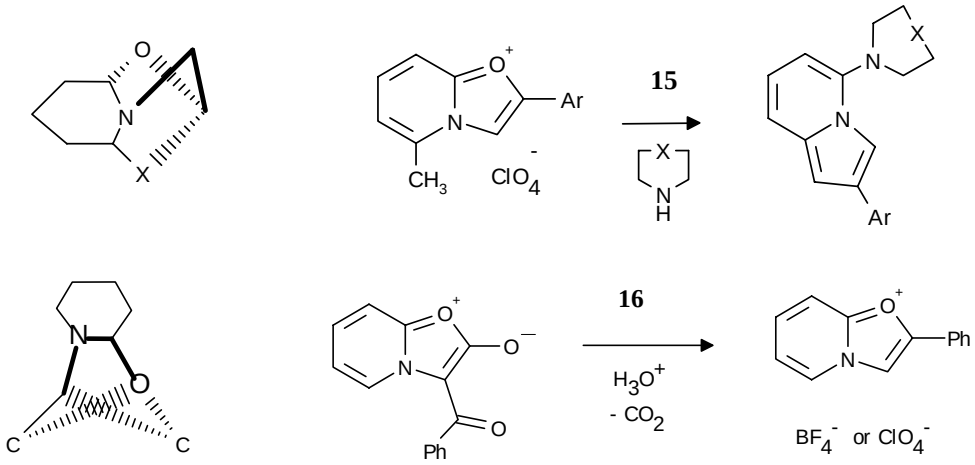
The methodology we used has been recently reviewed[67]. After careful analysis of the literature data the structures selected to study recyclizations were the derivatives of 1-heteroindolizines (neutral, mesoionic, and cationic). The GREH program has suggested for every cycle of the bicyclic 1-heteroindolizines the appropriate location of certain functions around the ring that are necessary for closure of new ring.



Scheme 15. Selected predictions of novel recyclization families for 1-heteroindolizines by the GREH program. From the recyclization graph (left, shown with annelated ring) the skeletal equation is restored. The design of polarity for exo- and endo-cyclic atoms achieved by application of the principle of the direct alternation chain. Labels in the open chain structures indicate the nucleophilic centers.

In order to avoid “combinatorial explosion”, the strict limitations were used, particularly, invariance of cycle size, minimal number of broken/formed bonds, and the aa-classes of recyclization graphs. The last requirement was the cleavage of only the most probable bonds in 1-X-indolizines (X - heteroatom), i.e. of C_9-X, C_2-X, and C_5-N bonds (see Section 4.3). After applying of such restrictions the predicted number of recyclizations decreased from few hundreds to few tens. Selected sixteen examples were published as the forecast[67], and few of them are presented in Scheme 15.

Design of substituent polarity for the concrete structures, capable to undergo the predicted transformations, was achieved by applying the principle of the *direct* polarity chain to 1-heteroindolizines (see Section 4.3) as it is illustrated in Scheme 15 this Synthesis of some structures with the desired polarities of substituents around the ring was achieved, particularly of 5-methyloxazolo[3,2-a]pyridinium salts[68]. The obtained structure undergoes elegant recyclization to 5-aminoindolizines (reaction 15 in Scheme 16) with complete agreement of the initial forecast[69,70]. It should be mentioned that neither for monocyclic nor for fused oxazoles such structural class has never been observed. We also failed to find any related analog with exactly the same sort. Another example of confirmation the forecast[67] was the recyclization of the mesoionic structure to the cationic oxazolopyridine[71], reaction 16 in Scheme 16. The feature of this reaction is that the final cationic structure, in turn, is able to undergo further recyclizations (see e.g. Scheme 11), and the principle of such "recyclization tandems" is rather rare case of heterocyclic chemistry.



Scheme 16. Examples of novel recyclization families for fused oxazoles predicted early and discovered recently. On the left: initial recyclization graph used for verification.

6. Conclusion

As we tried to prove, many qualitative observations in heterocyclic chemistry can be more strictly reformulated with the help of the reaction graphs. These graphs are not something external, that comes from outside the topic of the chemistry of heterocycles: these models are inside the discipline. The well reviewed idea of polarities arrangement (along the cycles and chains) should be supplemented by simple considerations on the conservation and change of polarities via heterolytic reactions. Adding combinatorics to this idea results in helpful tool for education and rational classifications of reactions, particularly of heterocyclic formation, cleavage and transformation reactions. As we proved, the discussed models, implemented as the computer programs, may be really useful in rational *reactions planning* -- search and discovery of unknown reactions.

Acknowledgements

Research leading to the results discussed in this review was generously funded by International Science Foundation (in 1994-95), Volkswagen - Stiftung (in 1996-97), GRACENAS foundation (St. Petersburg, Grant Grant No.95-0-9.4-222), and Russian Foundation of Basic Research (Grant No.96-03-32953). The author wish to thank his co-authors for contribution.

References

1. (a) Lushnikov, D. E.; Babaev, E. V. *Chem. Heteroc. Compounds (Engl. Transl.)* **1993**, *10*, 1111. Original: *Khim. Geterotsikl. Soedin.* (in Russian), 1299 (*Chem. Abstr.* **120**: 216194k). (b) Computer programs mentioned in the review (programs GREH and Heterocycland for IBM PC; DOS or MS Windows 3.x) are now free available from WWW (http://org.chem.msu.su/~babaev).
2. Trinajctic, N. *Chemical Graph Theory.* CRC Press: Boca Raton, FL, 1991; second edition.
3. *Chemical Graph Theory: Introduction and Fundamentals;* Bonchev, D.; Rouvray D., Eds.; Abacus Press/Gordon and Beach: New York, 1991.
4. Balaban, A.T. *Rev. Chim.* **1967**, *12*, 875.
5. Dugundji, J.; Ugi I. *Topics Curr. Chem.* **1973**, *39*, 19.
6. Ugi, I.; Gillespie P. *Angew. Chem. (Int. Ed. Engl.)* **1971**, *10*, 914.
7. Hendrickson, J. B. *Angew. Chem. (Int. Ed. Engl.)* **1974**, *13*, 47.
8. Zefirov, N. S.; Trach S. S. *Zh.Org.Khim.* **1975**, *11*, 1785 (in Russian).
9. Zefirov, N. S. *Acc. Chem. Res.* **1987**, *20*, 237.
10. Fujita, S. *J. Chem. Inf. Comput. Sci.* **1986**, *26*, 205.
11. Fujita, S. *J. Chem. Soc. Perkin Trans. II.* **1988**, 597.
12. Barone, R.; Chanon, M. In *Computer Aids to Chemistry.* Vernin, G.; Chanon M., Eds.; Chichester: E. Horwood, 1986.
13. Wilcox, C. S.; Levinson, R. A. In *Artificial Intelligence Applications in Chemistry;* Pierce, T.H.; Hohne, B.A., Eds.; Washington: Amer. Chem. Soc. 1986, p.209.
14. Hendrickson, J. B. *Rec. Trav. Chim.* **1992**, *111*, 323.
15. Zefirov, N. S.; Tratch, S. S. *Anal. Chim. Acta.* **1990**, *235*, 115.
16. Zefirov, N. S.; Baskin, I. I.; Palyulin, V. A. *J. Chem. Inf. Comput. Sci.* **1994**, *34*, 994.
17. Bauer, J. *Tetrahedron Comput. Methodol.* **1989**, *2*, 269.
18. Herges, R.; Hoock C. *Science.* **1992**, *255*, 711.
19. *Heterocyclic Compounds;* Elderfield, R. C., Ed.; Wiley: NY, 1950; Vols. 1-7.
20. *Comprehensive Heterocyclic Chemistry;* Katritzky, A. R.; Rees, C. W., Eds.; Pergamon Press: Oxford, 1984; Vols. 1-8.
21. *Houben-Weil Methoden der organischen Chemie. Bd. E 7a (Hetarene II);* Kreher, R. P., Ed.; Stuttgart, N.Y.: Thieme Verl., 1991; Teil 1.

22. Jones, G. In *Quinolines;* Jones, G. Ed.; L.:Wiley, 1977; Part 1, p. 93
23. Potts, K. In *Comprehensive Heterocyclic Chemistry;* Katritzky, A. R.; Rees, C. W., Eds.; Pergamon Press: Oxford, 1984; Vol. 5, p. 111.
24. Katritzky, A. R. *Handbook of Heterocyclic Chemistry;* Oxford: Pergamon, 1985; p. 382.
25. Bures, M.G.; Jorgensen, W.L. *J. Org. Chem.* **1988**, *53*, 2504.
26. Babaev, E.V. *Chem. Heteroc. Compounds (Engl. Transl.)* **1993**, *7*, 796. Original: *Khim. Geterotsikl. Soedin.* (in Russian), 937. (*Chem. Abstr.* **120**: 77095p)
27. Evans, D. A. *UCLA Physical Organic Chemistry Seminar*, May 6, 1971. (Manuscript 242A: Evans, D. A. An Organizational Format for the Classification of Functional Groups).
28. Evans, D. A.; Andrews G. C. *Acc. Chem. Res.* **1974**, *7*, 147.
29. Ugi I., Gillespie, P. *Angew. Chem.* **1971**, *23*, 985.
30. Doenges, R.; Groebel, B.-T.; Nickelsen, H; Sander J. *J. Chem. Inf. Comput. Sci.* **1985**, *4*, 425.
31. Serratosa, F. *Buttleti Societat Catalana Cienc. Fis., Quim., Mat.* **1977**, *1(2)*, 75 (In Catalan). (*Chem. Abstr.* **89** 196341s.)
32. Wagener, M.; Gasteiger, J. In *Software Developments in Chemistry*; Gasteiger, J., Ed.; Springer-Verlag: Berlin, 1990; Vol. 4, p. 265.
33. Serratosa, F.; Xicart, J. *Organic Chemistry in Action. The Design of Organic Synthesis*. 2nd Edition, Elsevier: Amsterdam, 1996.
34. Lapworth, A. *Manchester Memoirs Lit. Phil. Soc.* **1920**, *64(3)*, 1.
35. Seebach, D. *Angew. Chem. (Int. Ed. Engl.)* **1979**, *4*, 239.
36. Ho, T.-L. *Reviews of Chem. Intermediates.* **1988**, *9*, 117.
37. Ho, T.-L. *Polarity Control for Synthesis*. Wiley: Chichester, 1991.
38. Babaev, E. V. *Mendeleev Chemistry Journal (Engl.Tansl.)*, **1994**, *38(6)*, 78. Original: *Ross. Khim. Zh. [Zhurnal Ross. Khim. Ob-va im. Mendeleeva]* (in Russian), 54 (*Chem. Abstr.* **122**: 186630w).
39. McKillop, A.; Boulton, A. J. In *Comprehensive Heterocyclic Chemistry;* Katritzky, A. R.; Rees, C. W., Eds.; Pergamon Press: Oxford, 1984; Vol. 2, p. 67.
40. Babaev, E. V.; Cytovskij, S. V. *Moscow University Chem. Bull. (Engl. Transl.)* **1994**, *49(1)*, 82. Original: *Vestnik Mosk. Univ. Ser.2: Khim.* (in Russian), *35(1)*, 85 (*Chem. Abstr.* **120**: 2298437d).
41. (a) Babaev, E. V. *AIP Conference Proceedings. (ECCC 1. Computational Chemistry. Nancy, France May 1994.)*. AIP Press: Woodbury, NY; **1995**, *330*, 746. (b) Babaev, E. V., *7-th International Conference on Mathematical Chemistry, Girona (Spain), May 26-31, 1997* (Plenary lecture); Proceedings **1997**, 17.
42. Efimov, A. V.; Babaev, E. V. *ibid.* 751.
43. (a) Babaev, E. V. *Chem. Heteroc. Compounds (Engl. Transl.)* **1993**, *7*, 818. Original: *Khim. Geterotsikl. Soedin.* (in Russian), 962 (*Chem. Abstr.* **120**: 106734d).

44. Latham, E. J.; Murphy, S. M.; Stanforth, S. F. *Tetrahedron Lett.* **1994**, *35*, 3395.
45. Meyers, A. I. *Heterocycles in Organic Synthesis*; N.Y.: Wiley Interscience, 1974.
46. Kozikowski, A. P. In *Comprehensive Heterocyclic Chemistry;* Katritzky, A. R.; Rees, C. W., Eds.; Pergamon Press: Oxford, 1984; Vol. 1, p. 413.
47. (a) Lipshutz, B. H. *Chem. Rev.* **1986**, *86*, 795. (b) Stirling, Ch. J. M. *ibid.* **1978**, *78*, 517.
48. Gilchrist, T. L. *Adv. Heterocycl. Chem.* **1987**, *41*, 41.
49. (a) Maiboroda, D. A.; Scherbukhin, V. V.; Babaev, E. V. *AIP Conference Proceedings. (ECCC 1. Computational Chemistry. Nancy, France May 1994.).* AIP Press: Woodbury, NY; **1995**, *330*, 757. (b) Simple computer program FROG (for DOS, IBM PC) described in this paper is available upon request from the author.
50. Babaev, E. V.; Bobrovskii S. I.; Bundel Yu. G. *Chem. Heteroc. Compounds (Engl. Transl.)* **1988**, *11*, 1307. Original: *Khim. Geterotsikl. Soedin.* (in Russian), 1570 (*Chem. Abstr.* **111**: 97031g).
51. (a) Babaev, E. V. *Bull. Soc. Chim. Belg.* **1992**, *9*, 823. (b) Babaev, E. V.; Torocheshnikov V. N.; Bobrovskii S. I. *Chem. Heteroc. Compounds (Engl. Transl.)* **1995**, *9*, 1079. Original: *Khim. Geterotsikl. Soedin.* (in Russian), 1235 (*Chem. Abstr.* **124**: 342267g).
52. Maiboroda, D. A., Babaev E. V. *Chem. Heteroc. Compounds (Engl. Transl.)* **1995**, *11*, 1251. Original: *Khim.Geterotsikl.Soedin.* (in Russian), 1445 (*Chem. Abstr.* **125**: 33500r).
53. (a) Maiboroda, D. A.; Babaev, E. V.; Jug K. *J. Org. Chem.* **1997** (in press). (b) For SINDO method see Nanda, D. N.; Jug, K. *Theor. Chim. Acta* **1980**, *57*, 95.
54. Babaev, E.V.; Pasichnichenko K. A.; Maiboroda, D. A. *Khim. Geterotsikl. Soedin.* (in Russian), **1997**, *3*, 397.
55. (a) van der Plas, H. C. *Ring Transformation of Heterocycles;* Academic Press: London, N.Y., 1973; Vols. 1, 2. (b) Other comprehensive reviews on the “named” reactions are Refs. 17, 24, 58 - 60.
56. Schvaika, O. P.; Artyomov, V. N. Usp. Khimii (in Russian) **1972**, *41*, 1788.
57. *Pyrillium Salts. Syntheses, Reactions and Physical properties*. Balaban, A. T., Ed.; Academic Press: N.Y., 1982.
58. L’abbé, G. *J. Heterocycl. Chem.* **1984**, *21*, 627.
59. Babaev, E. V.; Zefirov, N. S. *Bull. Soc. Chim. Belg.* **1992**, *101*, 67.
60. Babaev, E. V.; Lushnikov D. E.; Zefirov N. S. *J. Amer. Chem. Soc.* **1993**, *115(6)*, 2416.
61. Babaev, E. V.; Zefirov, N. S. *Chem. Heteroc. Compounds (Engl. Transl.)* **1992**, *6*, 671. Original: *Khim. Geterotsikl. Soedin.* (in Russian), 808 (*Chem. Abstr.* **118**: 101252u).
62. Babaev, E. V.; Zefirov, N. S. *J. Math. Chem.* **1992**, *11*, 65.
63. Babaev, E. V.; Bozhenko, S. V.; Maiboroda, D. A. *Russ. Chem. Bull. (Engl.Transl.)*, **1995**, *11*, 2203. Original: *Izv. Akad. Nauk, Ser. Khim.* (in Russian), 2298. (*Chem. Abstr.* **124**: 316926b).
64. Babaev, E. V.; Bozhenko, S. V. *Khim. Geterotsikl. Soedin.* (in Russian), **1997**, *1*, 141.

65. Terenin, V. I.; Babaev, E. V.; Yurovskaya, M. A.; Bundel, Yu. G. *Chem. Heteroc. Compounds (Engl. Transl.),* **1992**, *6*, 658. Original: *Khim. Geterotsikl. Soedin.* (in Russian), 792 (*Chem. Abstr.* **118**: 168899d).
66. Babaev, E. V.; Jernigan, K. In *4-th Internatonal Conference on Chemical Structures. June 2-6, 1996*; Noordwijkerhout. Netherlands, 1996;. Abstr. of papers, 31.
67. Babaev, E. V.; Zefirov, N. S. *Khim. Geterotsikl. Soedin.* (in Russian), **1996**, *11-12*, 1564.
68. Babaev, E. V.; Efimov, A. V.; Maiboroda, D. A. *Chem. Heteroc. Compounds (Engl. Transl.)*, **1995**, *8*, 962. Original: *Khim. Geterotsikl. Soedin.* (in Russian), 1104 (*Chem. Abstr.* **124**: 232301f).
69. Babaev, E. V.; Bozhenko, S. V.; Efimov, A. V. In *12-th Symposium on Chemistry of Heterocyclic Compounds and 6-th Blue Danube Symposium on Heterocyclic Chemistry*. Brno, Czech Republik, Sept. 1-4, 1996; Conference Proceedings, P21.
70. Babaev, E. V.; Efimov, A. V.; Maiboroda, D. A.; Jug, K. *Liebigs Ann.* **1997** (in press).
71. Babaev, E. V.; Orlova, I. A. *Khim. Geterotsikl. Soedin.* (in Russian), **1997**, *4* (in press).