COMPUTER DESIGN
OF HETEROCYCLIC RING-OPENING REACTIONS.

D.A. **Maiboroda, V.V. Scherbukhin, E.V. Babaev***

Chemical Dept., Moscow State University, 119899 Russia

Abstract

Novel approach to the problem of computer design of heterocychc ring-opening reactions is discussed. Few empirical rules are proposed to estimate the most probable bond to be cleaved in azole structures. This rules implemented as the algorithm into computer program FROG permit one to predict reasonable non-equivalent products of heterocyclic ring opening.

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. Although many reviews and books are devoted to this topic, we failed to find simple, easily formalized rules, that may permit one to enumerate (or even estimate) the direction of bond cleavage for given heterocyclic structure.

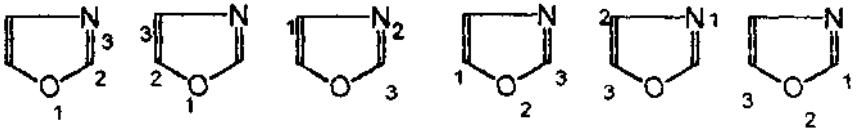
The goal of this communication is to propose simple rules for different mechanisms of ring opening of heteroaromatic structures. These rules permit us to order the bonds in given azole structure according to the highest probability of their cleavage. We shall consider the case of azoles ring opening, since they are better represented in literature

Let us roughly classify the mechanisms of heterocyclic ring opening into three broad categories: polar (thermal processes under the action of polar species), Red-Ox (reductive and oxidative conditions) and photochemical reactions Here we consider polar ring opening processes. Polar reactions can be divided into two groups: *endo- and exo- ring opening.* Since the cleavage of ring bond requires appearance of new bond either between the (endo) atoms of initial cycle, or between initial cycle and external (exo) group, the first case is treated as endo-opening, and the another as exo-opening:

Endo-Ring Opening Exo-Ring Opening

General Methodology and Algorithms

1. As the first approximation we consider that the bond to be broken should be single bond. We consider 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 2 double and 3 single bonds) there are either 3 or 6 (depend on symmetry) possibilities to find such 1-2-3 triads.



2. Every atom of the triad 1-2-3 is labelled as donor (D) or acceptor (A) center. As the first approximation we consider any heteroatom (N,S,O - either of pyrrole or pyridine type) to be donor center, and any carbon -- to be acceptor center.

3. The possibility of ring opening according to exo- and endo-type depends on the ***polar nature of all atoms in triad 1-2-3.*** Experimental data show clear consequence of priority of bonds cleavage depending on the D or A nature of every atom in triad. We avoid here the discussion, that lead to results discussed below. Complete review and theoretical analysis will be published elsewhere, and here we present only the main qualitative result.

RULE FOR ENDO-TYPE RJNG OPENING. Empirical data show, that the most probable consequence of atoms for endo type of opening is DDD (i.e. neigborhoud of three heteroatoms), while the less probable is the combination AAA, (i.e. fragment with three carbons).

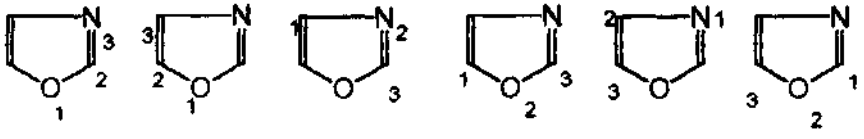
For simplicity, let us match every combination by binary code, where the categories of binary number are the positions 1,2,3. The values 0 or 1 are attributed to A and D respectively. Thus, the most probable combination DDD will get the code 111 (or decimal 7), while the less probable consequence AAA will be written as 000 (or decimal 0). Another combinations of polar centers will have the following empirical order of priority

Atoms 1,2,3	DDD	DDA	DAD	DAA	ADD	ADA	AAD	AAA
Binary code	111	110	101	100	011	010	001	000
Decimal code	7	6	5	4	3	2	1	0

RULE FOR EXO-TYPE RING OPENING. Empirical data show, that the most probable consequence of atoms for exo type of opening is another consequence, namely D-A-D, while the less probable is the combination A-D-A. In this case the other combinations are ordered analogously by binary code, however the matching should follow another consequence:

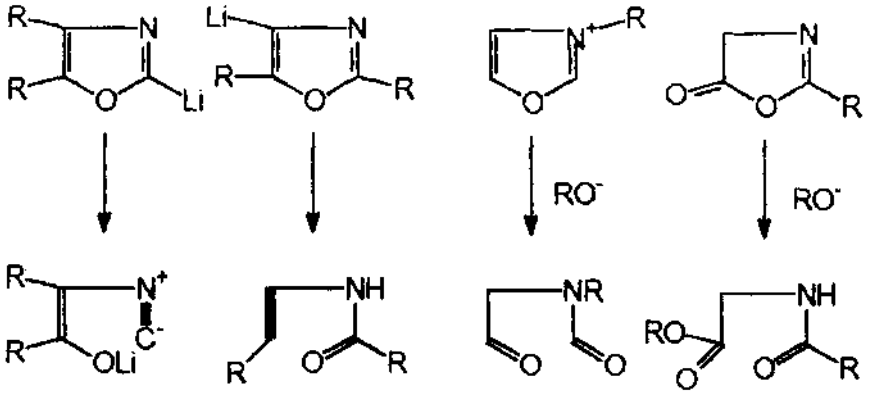
Atoms 1,2,3	DAD	DAA	DDD	DDA	AAD	AAA	ADD	ADA
Binary code	111	110	101	100	011	010	001	000
Decimal code	7	6	5	4	3	2	1	0

4. Now it is possible to attribute the code for every triad of any azole structure. This algorithm is realized in the computer program FROG. After input of azole structure all the triads are labelled, and every triad is characterized by two decimal values (first - raiting for endo-type opening, second - for exo-type opening). For example



	DAD	DAA	ADA	ADA	DAA	ADA
Endo	101	100	010	010	100	010
	5	4	2	2	4	2
Exo	111	110	000	000	110	000
	7	6	0	0	6	0

In all cases the higher raiting corresponds to the most probable direction of bond cleavage. Examples of structures and their transformation for oxazoles, that can be found elsewhere, correspond to the highest raiting:



5. At the last step (which is now in progress) the choice of diagram should give to user possibility to obtain the resulting structure of open chain product (or products, if the cycle dissociate to more than one component). Even for the case of such "simple" structures as azoles, the problem of generating the final structure seems to be complicated. The most serious question is to take into account leaving groups and external nucleo- and electrophilic species, as well as to predict quantitatively the role of polar substituents in the ring.

Nevertheless, even at the first level discussed above, the program really select (by raiting criteria) just those ring openings that are found experimentally. This gives the hope that such algorithm may also predict previously unknown directions of ring cleavage.