

COMPUTER SYNTHESIS OF BENZENE DERIVATIVES FROM CYCLIC AND ACYCLIC PRECURSORS

A.V. Efimov and E.V. Babaev*

Chemical Dept., Moscow State University, Moscow 119899, Russia

Abstract

General methodology for computer assisted synthesis of benzene derivatives from another benzenes (by polar substitution reactions) or from acyclic precursors (by polar cyclocondensations) is briefly discussed. The approach is based on the further development of the consonant and dissonant relationship and its application to the cyclic benzene structures. New dissonant index is proposed for computer assisted synthesis planning. The program AROMAT, capable to predict reasonable polar types of cyclic and acyclic precursors of given benzene, is briefly discussed.

Although benzene derivatives play major part in organic chemistry, computer-assisted synthesis of benzenes is usually limited by substitution reactions. However, great number of reactions, where the benzene nucleus is constructed by polar reactions of ring formation from acyclic precursors (e.g., cyclocondensations) or via ring transformation of other cyclic precursors (e.g., recyclizations of heterocycles), was out of the scope of modern computer synthesis.

The goal of this communication is to describe general methodology for computer synthesis of benzene derivatives independently of *cyclic or acyclic* structure of precursors for the given substituted benzene. We shall strictly limit ourselves only by the class of polar reactions (e.g., electro- and/or nucleophilic addition and/or elimination processes). We classify the polar nature of cycles, chains or small reagent particles by the modified definition of **consonant** and **dissonant** types (first proposed by D.A.Evans in 1971^{1,2}).

DEFINITIONS

As the first approximation we consider two sets of substituents D and A:

D — **Donors and ortho-/para-orientants** with exact number of $(8N-1)$ valence electrons, like methyl, amino, hydroxyl, and halogen groups,

A — **Acceptors and meta-/ipso-orientants** with exact number of $(8N+1)$ valence electrons, like hydrogen, carboxyl, nitro, cyano, sulfo, — MgX (as in Grignard reagents), NN^+ (as in diazonium salts).

Nucleophiles and electrophiles are considered to be either **consonant** (with $8N$ electrons, i.e. of the type D^- or A^+) or **dissonant** (D^+ with $8N-2$ or $8N+6$ electrons or A^- with $8N+2$ electrons).

Substituted benzenes are classified according to the arrangement of D- and A-groups around the ring (note that H is the A group). It is postulated, that the presence of an even-membered carbon chain between two A groups (as far as between two D groups) or presence of an odd-membered chain between any A and D groups leads to some sort of conflict, called **dissonance**, and the resulting benzene is called dissonant. Another bifunctional relationship — odd between DD (or AA) or even — between A and D groups causes no conflict (consonant fragment). Here our definition slightly differs from the early terms (normal and umpolung reactivity³, consonant and dissonant^{1,2,4,5}, conjoint and disjoint arrangement⁶) since we consider the conflicts including heteroatoms and (formally) electropositive hydrogen atom. This is more closely related to first definition of "homogenous and heterogeneous" molecules proposed by A.Lapworth⁷.

Qualitatively, the absence of conflict reflects some sort of stabilization, usual conjugation (or cross-conjugation) and hence, consonant orientation of two or more groups in respect to external electro- or nucleophilic attack. On the contrary, appearance of a conflict means the uncertainty of such orientation as well as some destabilization (cf. refs. 1-3). It is noteworthy that every consonant molecule, being mentally dissociated into ions, can be presented as the set of *only octet ions with charge alternation*; this is impossible for dissonant ones.

The number of conflicts K is calculated by use of computer program for any benzene with groups of the above types A and D. It is easy to prove that there are only 4 possibilities for K -value, namely 0, 5, 8, and 9 (if only one shortest route between two groups is taken into account). Meanwhile, these conflicts have different origin (note that H is considered as the A group!), and may also be written as the vector (ZZ, ZH, HH), where:

ZZ - the number of conflicts between "heavy" (non-hydrogen) atoms Z (Z is either A or D),

ZH - conflicts between Z atoms and hydrogens,

HH - conflicts between hydrogen atoms.

Evidently, $ZZ + ZH + HH = K$. Therefore, any substituted benzene is suited in the one of 4 planes or triangles in the space (ZZ, ZH, HH), as it is shown at the Scheme 1. (Zero is degenerated plane for $K=0$.)

The above definition permits one to change usual structural formula of a benzene derivative to its simple *polar image*, reflected by the vector (ZZ, ZH, HH). By this manner, any predecessor with benzene ring of the target structure can be analogously described by its own vector. By the other hand, it is not a problem to calculate such vector to any reasonable open chain, that may be the precursor of a given benzene target.

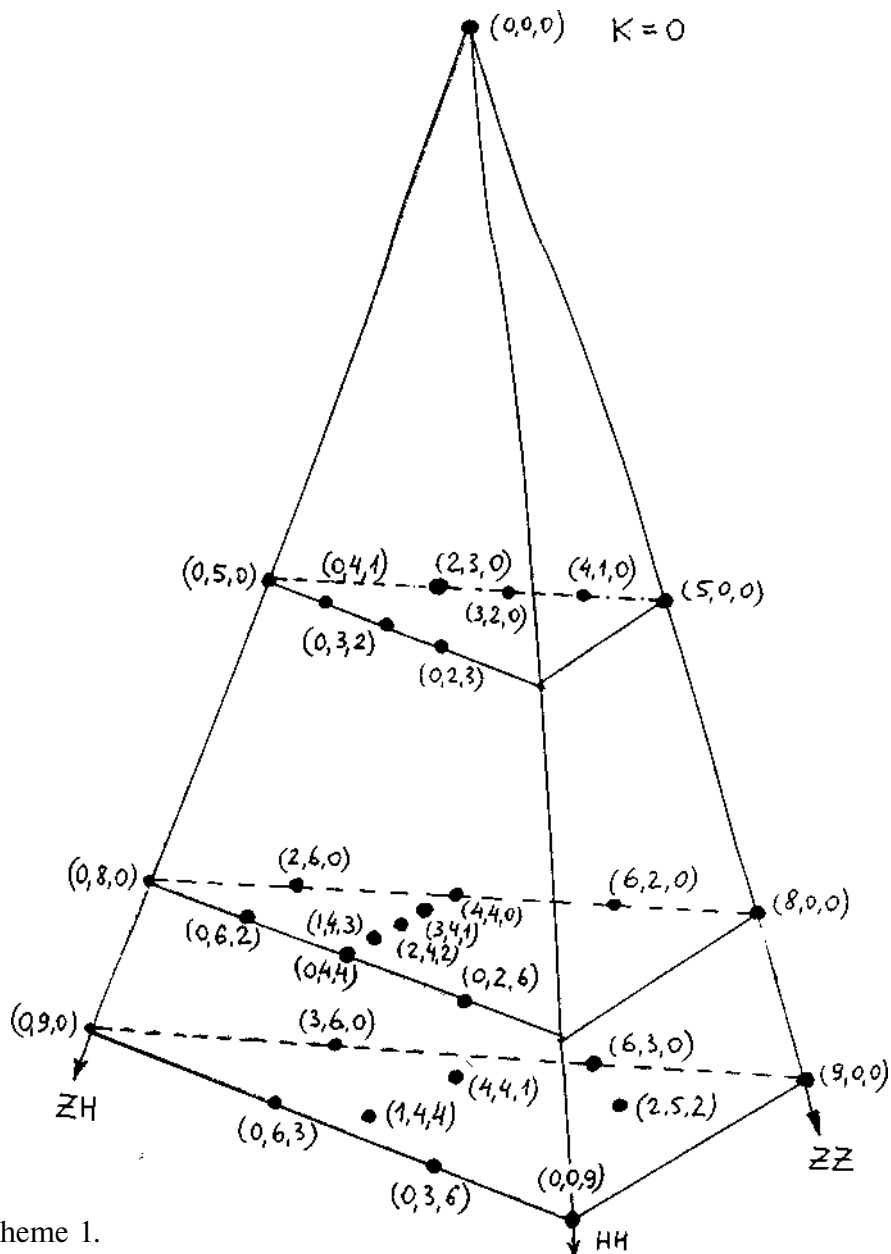
Our main idea is that the dissonant index K and the vector (ZZ, ZH, HH) are new interesting and prospectful structural indices that describe behavior of polar systems in polar reactions, and may be useful for computer-assisted synthesis planning.

Our observations for reactions of substituted benzenes with consonant electrophiles (A^+) or consonant nucleophiles (D^-) are: 1) initial K-value tends to conserve, 2) the ZZ part of vector tends to conserve, 3) if two isomeric products are possible, they should have the same vectors (ZZ, ZH, HH), 4) if orientation effects of two substituents D and A are in conflict, the product with minimal ZZ part is related to the effect of D group (i.e. D is stronger than A). These four rules permit us to develop different *consonant strategies* of benzene synthesis in almost complete agreement with chemical "common sense". Examples are given at the Fig.1. By this strategy every predicted precursor should have the same arrangement of polar groups as the target structure. Concrete examples of groups interchange may be taken from every textbook. Further development of this idea includes differentiation of the nature of conflicts: we can attribute some "weight" W_{ij} to every conflict between atoms i and j. The total dissonant index will be:

$$K = ZZ * W_{ZZ} + ZH * W_{zh} + HH * W_{hh}$$

and the methodology will include analogous minimization.

The consonant strategy can be applied not only for the cases of interchange of consonant groups (A^+ by A^+ or D^- by D^-), but also for the processes of substitution of one dissonant group by another (i.e. D^+ by D^+ or A^- by A^-). This opens possibility to include reactions of



Scheme 1.

Arrangement of vectors of substituted benzenes in the space (ZZ, ZH, HH). Every point belongs to the plane $ZZ+ZH+HH=K$.

electrophilic ipso-substitution or use of dissonant leaving groups (like H^- in the S_NH reactions or N_2 in diazonium salts) at the usual consonant strategy.

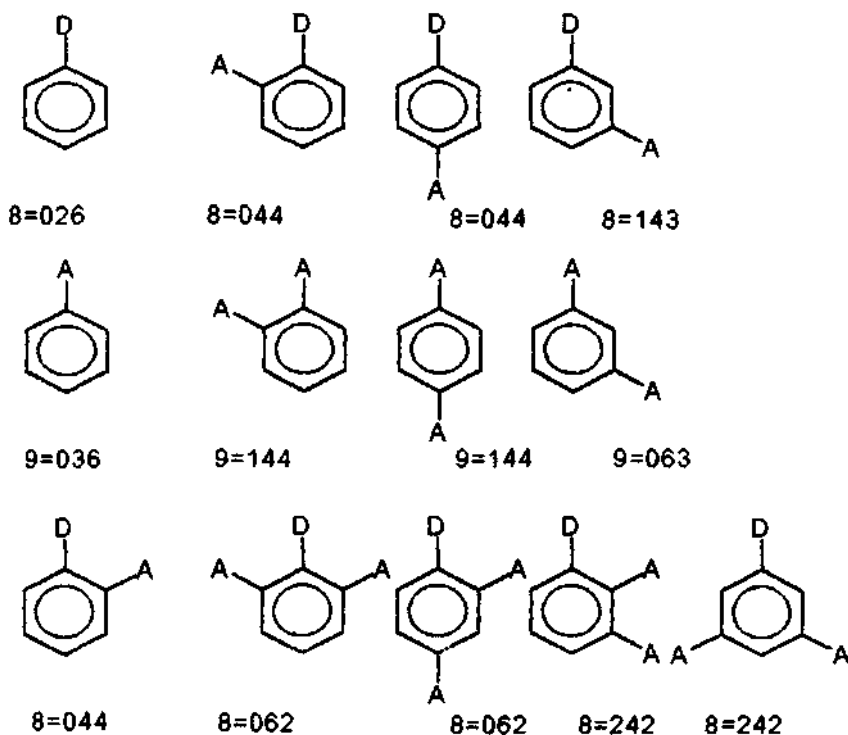


Fig.1. Calculating of K value and the sum ZZ+HZ+HH for isomers in the electrophilic substitution reactions

The *dissonant strategy* means changing of the initial polar nature of the group via reaction or retro-reaction, e.g. A^+ to D^+ , D^+ to A^+ , A^- to D^- , and D^- to A^- . In this case the following "trick" is proposed. Any substituent in the ring

is claimed to be of the opposite nature (thus, amino group is considered as pseudo-acceptor and nitro-group as pseudo-donor). Then the usual consonant strategy is applied to this "illogical" structure. As the result, the application of above 4 rules again gives reasonable predictions.

The last step is the search of reasonable acyclic precursors of the given benzene target. This type of reaction (closely related to heterocyclizations) plays most important role in the synthesis of completely consonant benzene structures ($K=0$), where the usual synthetic methods include many steps. For synthesis planning of such structures two different methodologies are applied:

1) *the consonant strategy*: in this case two different Kekule structures are considered for the target. Every single bond is broken, and its ends are changed to the atoms X (leaving group of the D-type) and H so that initial K-value (or at least the ZZ part) should be equal to those of the target structure. The same operation is applied to cleavage of every double bond (two X-atoms and two H-atoms appear). Resulting chain can be further broken to give multicomponent ensemble of precursors.

2) *the dissonant strategy*: in this case only one step of bond cleavage should be in conflict with initial alternation, while the other dissections should conserve the initial polarity in the ring. Our experience in application of the discussed program (preliminary called AROMAT) to design of consonant natural compounds - orsine and orsellinic acid will be published elsewhere.

REFERENCES

1. D.A.Evans, UCLA Physical Organic Chemistry Seminar, 6.5.1971
2. D.A.Evans, G.C.Andrews, Acc.Chem.Res., N7, 147 (1974) .
3. D.Seebach, Angew.Chem. (Int.Ed.Engl.), N18, 239 (1979).
4. F.Serratos, Organic Chemistry in Action. The Design of Organic Synthesis (Elsevier, 1990).
5. RDoenges, B.-T.Groebel, H.Nickclsen, J.Sander, J. Chem. Inf. Comput. Sci., Vol.25, 425 (1985).
6. T.-L.Ho, Polarity Control for Organic Synthesis (Chichester, 1991)
7. A.Lapworth, Mem.Manchester Lit. Phil. Soc., iii, 3 (1920).