Birkbeck

UNIVERSITY OF LONDON

Birkbeck ePrints

Birkbeck ePrints: an open access repository of the research output of Birkbeck College

http://eprints.bbk.ac.uk

Evstigneev, Maxim P.; Davies, David B. and Veselkov, Alexei N. (2006). Stochastic models (cooperative and non-cooperative) for NMR analysis of the hetero-association of aromatic molecules in aqueous solution. *Chemical Physics* **321** (1-2) 25-33.

Citation for this version:
Evstigneev, Maxim P.; Davies, David B. and Veselkov, Alexei N. (2006). Stochastic models (cooperative and non-cooperative) for NMR analysis of the hetero-association of aromatic molecules in aqueous solution. *London: Birkbeck ePrints.* Available at: http://eprints.bbk.ac.uk/archive/00000380

Citation for the publisher's version:
Evstigneev, Maxim P.; Davies, David B. and Veselkov, Alexei N. (2006). Stochastic models (cooperative and non-cooperative) for NMR analysis of the hetero-association of aromatic molecules in aqueous solution. *Chemical Physics* **321** (1-2) 25-33.

STOCHASTIC MODELS (COOPERATIVE AND NON-COOPERATIVE) FOR NMR ANALYSIS OF THE HETERO-ASSOCIATION OF AROMATIC MOLECULES IN AQUEOUS SOLUTION

Alexei N. Veselkov[†], Maxim P. Evstigneev[‡] and David B. Davies[‡*]

[†] Department of Physics, Sevastopol National Technical University, Sevastopol 99053, Crimea, Ukraine.

[‡] School of Biological and Chemical Sciences, Birkbeck College, University of London, Malet Street, London WC1E 7HX, UK.

[*] Author to whom correspondence should be addressed.

Mailing address:

     Prof. D.B.Davies

     Department of Chemistry

     Birkbeck College,

     University of London,

     Malet street

     London WC1E 7HX, UK.

     Fax: +44(0)207-631-6246

     Tel.: +44(0)207-791-0791

     e-mail: davidbry@ndavies.co.uk

Summary

Stochastic cooperative (STOCH-C) and non-cooperative (STOCH-NC) models have been developed for NMR analysis of the hetero-association of aromatic compounds in solution, in order to take into account all physically-meaningful association reactions of molecules in which there are no limitations on the lengths of the aggregates and complexes. These algorithmical approaches are compared with previously published basic (BASE) and generalized (GEN) analytical statistical thermodynamical models of hetero-association of biologically-active aromatic molecules using the same sets of published NMR data measured under the same solution conditions (0.1M phosphate buffer, pD=7.1, T=298K). It is shown that, within experimental errors, the BASE analytical model may be used to describe molecular systems characterized by relatively small contributions of hetero-association reactions, whereas the GEN model may be applied to hetero-association reactions of any aromatic compound with different self-association properties. The STOCH-C computational algorithm enabled the effect on hetero-association of the interactions of molecules with different cooperativity parameters of self-association to be estimated for the first time and it is proposed that the algorithm for the stochastic models has great potential for detailed investigation and understanding of the interactions of aromatic molecules in solution.
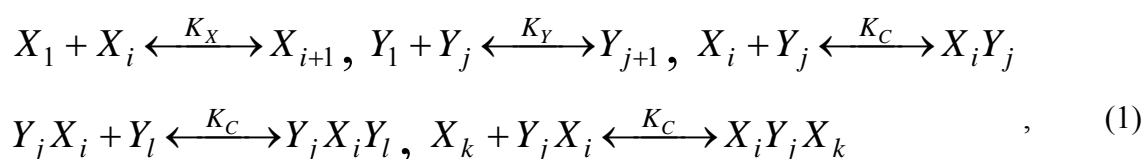
1. Introduction

Extensive use of aromatic compounds in clinical practice and in different biophysical studies at the molecular and cellular levels is due to their great biological and medical activity. Many biologically-active aromatic compounds act via their complexation with nuclear DNA [1]. In addition it has been shown that the biological activity of aromatic compounds may be substantially changed when combinations of such drugs are used. Thus it is known that caffeine alters the efficacy of a number of aromatic anticancer drugs, such as doxorubicin, novatrone, ellipticine and others [2-4] and affects the toxicity of a typical DNA intercalator, ethidium bromide [5]. For example, the use of novatrone in combination with other aromatic and non-aromatic antibiotics has been found to be a very effective therapy with different leukaemias [4]. The molecular mechanisms of such action include formation of hetero-complexes between the aromatic ligands and competition between the ligands for DNA binding sites [6-10]. It has been shown recently that the hetero-association of aromatic molecules may play a substantial role in modulating the biological activity of aromatic drugs under certain experimental conditions; for example, when the anticancer antibiotic daunomycin interacts with DNA in the presence of proflavine or ethidium bromide [10].

Hetero-association of aromatic molecules is also responsible for an increase in the solubility of different antibiotics and vitamins in the presence of hydrotropic agents, such as caffeine and nicotinamide [11,12]. In such cases the contribution of hetero-complexes to the total dynamic equilibrium may predominate at the physiological ratio of the concentrations of the interacting molecular components in solution [12]. Hence, it may be concluded that chemico-physical investigations of the hetero-association of aromatic molecules in solution are important for understanding the mechanisms of action of different combinations of biologically-active compounds in cellular systems.

Depending on the experimental method applied, different statistical-thermodynamical models have been used to analyze the hetero-association of aromatic compounds. Dimer models

are mostly used to interpret spectrophotometric data, *i.e.* they take into account the formation of self-aggregates and hetero-association complexes with no more than two molecules in the stack. The dimer model is valid only when the concentrations and the equilibrium association constants of the interacting molecules are relatively small [6-8,13]. When comparatively large concentrations of molecules are used in the experiment, such as for NMR analysis, more general models are considered as they take into account both the indefinite self- and hetero-association of aromatic molecules in solution [14,15]. The assumptions used in NMR modeling of the association of aromatic molecules in solution are well established and their scope and limitations discussed in a review on the comparison of indefinite self-association models [13].

Two generalized theoretical approaches have been used to analyze the hetero-association of aromatic molecules having different biological-medical and/or chemical- physical properties using NMR data: the basic [15,16] and the generalized [17] models. In these models the same reaction scheme has been used to describe the dynamic equilibrium of two aromatic components *X* and *Y* in solution:

$$X_1 + X_i \xleftrightarrow{K_X} X_{i+1} \, , \; Y_1 + Y_j \xleftrightarrow{K_Y} Y_{j+1} \, , \; X_i + Y_j \xleftrightarrow{K_C} X_i Y_j$$

$$Y_j X_i + Y_l \xleftrightarrow{K_C} Y_j X_i Y_l \, , \; X_k + Y_j X_i \xleftrightarrow{K_C} X_i Y_j X_k \qquad , \qquad (1)$$

where X_1, Y_1 are monomeric concentrations of *X* and *Y* molecular components in solution, respectively; X_i, X_k, Y_j, Y_l are aggregates, containing i, k monomers of *X* and j, l monomers of *Y*; K_X, K_Y, K_C are equilibrium constants of self- and hetero-association of the molecules; $i,j,k,l \in 1.\infty$. In the model in equation (1) the hetero-association complexes $X_i Y_j$ have one hetero-stack, whereas $X_i Y_j X_k$ *etc* have two hetero-stacks in the complex.

The difference between the two approaches is due to the contributions to observed chemical shifts of the "edge effects" of aromatic molecules in the formation of hetero-complexes; edge effects are taken into consideration in the generalized (GEN) model and make it more advantageous than the basic (BASE) model for investigations of the hetero-association of aromatic molecules with relatively high hetero-association constants [17]. A limitation of both models is that formation of

hetero-complexes with no more than two hetero-stacks are taken into account in the dynamic equilibrium in solution. However, it is found that the contributions to hetero-association reactions of more than two hetero-stacks to the total dynamic equilibrium may be essential for some combinations of biologically-active aromatic molecules [10,16] and, moreover, it may be predominant in the action of hydrotropic agents with aromatic molecules [12]. The appropriate analysis, which takes into account the probability of formation of hetero-complexes of aromatic molecules with more than two hetero-stacks, needs to be made.
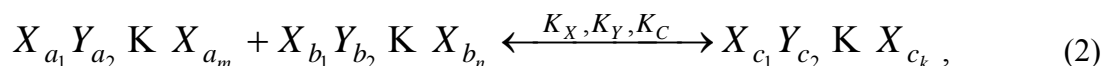
Models of the hetero-association of aromatic molecules developed previously [14-17] also do not take into account the possible cooperativity effects of the self-association of X and Y components in solution. Cooperativity effects in the self-association of aromatic molecules can be characterized by a cooperativity parameter σ and reactions may be cooperative ($\sigma<1$), non-cooperative ($\sigma\sim1$) and anti-cooperative ($\sigma>1$) [13,18]. Investigations have shown that the magnitude of the cooperativity parameter for self-association of aromatic molecules ranges from 0.4 (acridine dyes [18]) up to 1.9 (flavine-mononucleotide [12]), *i.e.* σ values for self-association may have some effect on the calculated hetero-association parameters. The importance of σ on the calculated values of the hetero-association parameters is not known and needs to be investigated.

In this work stochastic non-cooperative (STOCH-NC) and cooperative (STOCH-C) models have been developed for NMR analysis of the hetero-association of two aromatic compounds, taking into account all physically meaningful association reactions of molecules in solution. A comparison has been made of the scope and limitations of different models (BASE, GEN, STOCH) and their effect on the calculated parameters of hetero-association of biologically-active aromatic molecules. The influence of cooperativity effects in the self-association of the interacting molecules on the parameters of hetero-association has also been discussed.

2. Stochastic Model

2.1 *Non-cooperative association model.*

The general case of the association of two aromatic compounds, *X* and *Y*, in aqueous solution considers formation of complexes having any possible distribution of homo- and hetero-stacks according to a generalized hetero-association reaction:

$$X_{a_1} Y_{a_2} \text{ K } X_{a_m} + X_{b_1} Y_{b_2} \text{ K } X_{b_n} \xleftrightarrow{\quad K_X, K_Y, K_C \quad} X_{c_1} Y_{c_2} \text{ K } X_{c_k} \; , \qquad (2)$$

where a_i, b_i, c_i are the numbers of single type molecules (the length) for the *X* or *Y* aggregate; *m*, *n*, *k* are the number of aggregates within every complex; K_X, K_Y, K_C, are the equilibrium constants of self-association of *X* and *Y*, and their hetero-association, respectively. As only nearest neighbors are considered to affect the association of aromatic molecules [13-17], it is assumed that the magnitudes of the equilibrium constants K_X, K_Y, K_C are independent of the number of molecules in the aggregates and complexes. Depending on the length of the aggregates a_i, b_i, c_i, which in some cases may adopt zero values, reaction (2) summarizes formation of homo-stacks of components *X* and *Y* with association constants K_X or K_Y, and a hetero-stack between these molecules characterized by constant, K_C.

An *I* index, which can adopt two values *X* or *Y*: $I \in (X,Y)$, is introduced. All possible types of complexes in solution may be distributed in three groups: i) complexes of the *X…Y* type, flanked by an *X* molecule from one side, and *Y* molecule from the other side; ii) complexes of the *X…X* type, flanked by *X* molecules from both sides; and iii) complexes of the *Y…Y* type, flanked by *Y* molecules from both sides. Hence it is possible to introduce a *T* index, which designates the type of complex and adopts values $T \in (X…Y, X…X, Y…Y)$.

Consider an arbitrary complex, containing *L* molecules of *X* and *Y* type and the number of hetero-stacks between *X* and *Y* molecules in this complex is *h* ($h \leq L - 1$). Hence the lengths (*l*) of the aggregates of one type of molecule, which form the complex, are equal to l_i, where $i \in (1…h+1)$ is the number of an aggregate within the complex.

Taking into consideration the assumptions of the distinctive features of the association of aromatic molecules in solution [13], analytical expressions for the total concentration C_I and chemical shift δ_I of I-type molecules in the fast exchange condition of the NMR experiment can be written in terms of the I index in the following form:

$$\begin{cases} C_I = C_I^{self} + \sum_{h=1}^{\infty} \left[f(h)\left(C_I^{XKX} + C_I^{YKY}\right) + \left(1 - f(h)\right)C_I^{XKY} \right] \\ \delta_I = \delta_I^{self} + \sum_{h=1}^{\infty} \left[f(h)\left(\delta_I^{XKX} + \delta_I^{YKY}\right) + \left(1 - f(h)\right)\delta_I^{XKY} \right] \end{cases} \tag{3}$$

where

$$C_I^{self} = \frac{I_1}{\left(1 - K_I I_1\right)^2}, \qquad \delta_I^{self} = \frac{I_1}{I_0\left(1 - K_I I_1\right)}\left[2\delta_{mI} - 2\delta_{dI} + \frac{2\delta_{dI} - \delta_{mI}}{1 - K_I I_1} \right] \qquad \text{are} \qquad \text{the}$$

concentrations and chemical shifts due to self-association reactions [13,18]; I_0, I_1 are the initial and monomeric concentrations of I-type molecules in solution, respectively; δ_{mI}, δ_{dI} are the proton chemical shifts of an I-type molecule in the monomer and dimer form, respectively;

$f(h) = \frac{1}{2}\left[1 + (-1)^h\right]$ is a unity/zero function, separating even and odd values of h.

Using the mass conservation law and the additivity model for proton chemical shifts in the fast-exchange condition on the NMR timescale [13,18], the concentration C_I^T and the chemical shift δ_I^T of I-type molecules in T-type complexes, presented in eqns. (3), may be determined as:

$$\begin{cases} C_I^T = \sum_{l_1=1}^{\infty} K \sum_{l_{h+1}=1}^{\infty} N_I^T C_{l_1 K l_{h+1}}^T \\ \delta_I^T = \frac{1}{I_0} \sum_{l_1=1}^{\infty} K \sum_{l_{h+1}=1}^{\infty} D_I^T C_{l_1 K l_{h+1}}^T \end{cases}, \tag{4}$$

where N_I^T is the number of *I*-type molecules in a *T*-type complex with a distribution of the lengths

of the aggregates $l_1...l_{h+1}$; $C_{l_1 K l_{h+1}}^T$ is the concentration of the *T*-type complex with the same

distribution of lengths of the aggregates.

Taking into account the law of mass action for every *T*-type complex, the following equations may

be written:

$$
\begin{cases}
X K Y: \quad C_{l_1 K l_{h+1}}^{XKY} = K_C^h \prod_{i=1}^{\frac{h+1}{2}} x_{l_{2i-1}} y_{l_{2i}}; \quad N_X^{XKY} = \sum_{i=1}^{A_X^{XKY}} l_{2i-1}; \quad N_Y^{XKY} = \sum_{i=1}^{A_Y^{XKY}} l_{2i}; \\
\qquad A_X^{XKY} = \frac{h+1}{2}; \quad A_Y^{XKY} = \frac{h+1}{2} \\
X K X: \quad C_{l_1 K l_{h+1}}^{XKX} = \frac{1}{2} K_C^h x_{l_{h+1}} \prod_{i=1}^{h/2} x_{l_{2i-1}} y_{l_{2i}}; \quad N_X^{XKX} = \sum_{i=1}^{A_X^{XKX}} l_{2i-1}; \quad N_Y^{XKX} = \sum_{i=1}^{A_Y^{XKX}} l_{2i}; \\
\qquad A_X^{XKX} = \frac{h+2}{2}; \quad A_Y^{XKX} = \frac{h}{2} \qquad\qquad\qquad\qquad ,(5) \\
Y K Y: \quad C_{l_1 K l_{h+1}}^{YKY} = \frac{1}{2} K_C^h y_{l_{h+1}} \prod_{i=1}^{h/2} y_{l_{2i-1}} x_{l_{2i}}; \quad N_X^{YKY} = \sum_{i=1}^{A_X^{YKY}} l_{2i}; \quad N_Y^{YKY} = \sum_{i=1}^{A_Y^{YKY}} l_{2i-1}; \\
\qquad A_X^{YKY} = \frac{h}{2}; \quad A_Y^{YKY} = \frac{h+2}{2}
\end{cases}
$$

where $I_l = K_I^{l-1} I_1^l$ is the concentration of *I*-type aggregate with length *l*; A_I^T is the number of *I*-

type aggregates in the *T*-type complex.

The value of D_I^T in equations (4) represents the chemical shift of *I*-type molecules in a *T*-

type complex with a distribution of lengths of aggregates $l_1...l_{h+1}$ and is determined as the

difference between the total chemical shift of the *I*-type molecules in isolated aggregates and that

with extra shielding in hetero-stacks:

$$
D_I^T = \left[\delta_{mI} N_I^T - 2N_I^T(\delta_{mI} - \delta_{dI}) + 2A_I^T(\delta_{mI} - \delta_{dI})\right] - h(\delta_{mI} - \delta_{CI}), \qquad (6)
$$

where δ_{CI} is a chemical shift of *I*-type molecule in the hetero-stack.

Equation (6) can be reduced to the following expression:

$$D_I^T = \delta_{mI}\left(2A_I^T - N_I^T - h\right) + 2\delta_{dI}\left(N_I^T - A_I^T\right) + \delta_{CI}h \ . \tag{7}$$

The system of equations (3) is now completely determined. In the non-cooperative theoretical approach considered above there are no limitations on the lengths of aggregates and complexes and equations (3) take into account all physically possible formation reactions of hetero-complexes with all possible combination of aggregates of the interacting molecules in solution. Model (3) is based on the same physical assumptions which have already been used in previous models of molecular self-association [13,18] and hetero-association [14-17] and does not introduce any extra limitations to the dynamic equilibrium summarized in equation (2).

2.2 *Cooperative association model.*

Introduction of the self-association cooperativity parameter (σ) primarily results in changes to the expression for calculation of the concentration of aggregates with numbers of single-type molecules $l > 1$ [13,18]:

$$x_l = \sigma_X K_X^{l-1} x_1^l \ , \quad y_l = \sigma_Y K_Y^{l-1} y_1^l \ . \tag{8}$$

Utilization of eqn.(8) in eqns.(5) needs special examination of the concentrations $C_{l_1 K l_{h+1}}^T$ of all complexes containing at least one aggregate of length unity. In order to modify the analytical form for the cooperative model, the expression $I_l = K_I^{l-1} I_1^l$ in eqns.(5) may be replaced by the following

$$I_l = \sigma_I K_I^{l-1} I_1^l \ , \tag{9}$$

where $\sigma_I = \begin{cases} \sigma_I, & \text{if } l > 1 \\ 1, & \text{if } l = 1 \end{cases}$ is an alternating cooperativity parameter of the self-association of I-type molecules in solution.

In summary, the cooperative stochastic model of hetero-association of aromatic molecules is represented by eqns.(3)-(7) together with expression (9).

2.3 *Computational algorithm.*

Investigation of equations (3) indicates that derivation of analytical expressions suitable for computation of the association parameters is practically impossible due to the great complexity of the necessary mathematical manipulations. A thorough theoretical analysis of the hetero-association of two aromatic components with formation of no more than two hetero-stacks in the mixed complex was made previously for the GEN model [17]. Analysis of such a relatively simple case shows that rather complicated mathematical manipulations are needed to calculate the hetero-association parameters K_X, K_Y, K_C and δ_C from the concentration dependence of the proton chemical shifts. Hence, the theoretical approach summarized in equation (3), based on functional-analytical modeling of the experimentally-observed chemical shifts, requires the development of a special computational algorithm in order to apply this model to analyze the NMR data for the hetero-association of aromatic molecules.

It is convenient to present the algorithm of the stochastic model in the form of two subprograms: the first one calculates the concentrations of compound X or Y in solution (Fig.1), the second calculates the chemical shift of the corresponding proton of the X or Y compound (Fig.2). In order to embody the stochastic algorithm it is simply necessary to replace the references on the corresponding computational subprograms in the program code of the standard algorithm of data processing in the analytical models by the references to these procedures. Hence, the computational procedure in both the analytical and algorithmical approaches of the modeling of the hetero-association of aromatic molecules can be carried out using the same calculation strategy, described elsewhere [15-17].

Consider an arbitrary molecular complex containing L molecules (length of the complex) and any number of hetero-stacks ranging from 0 (a self-associate) up to the maximum possible value of L-1 (with alternation of X and Y molecules in the complex). Let a molecule of compound X correspond to unity and a Y molecule to zero. It follows that the given complex can be presented

in the form of a binary number *C* with *L* bits (a variable *Complex* in Figs.1,2). Testing every bit in *C* in a cycle for the state of 0 or 1 (cycle using *i* variable in Figs.1,2) enables the concentration of the complex *C* to be calculated by means of multiplication of the stack formation constant (K_X for homo-stack "11"; K_Y for homo-stack "00"; K_C for hetero-stack "01" or "10") on the monomeric concentration of the bit (x_1 for "1", y_1 for "0"; for example, a unity in the site "...01..." results in the coefficient $K_C \cdot x_1$) (see blocks 14-17 in Fig.1).

Calculation of the chemical shift depends on the neighboring bits, *i.e.* the contribution given by the two neighboring molecules: $\Delta\delta_C$ from a hetero-stack or $\Delta\delta_S$ from a homo-stack. Sequential summation of these contributions enables the chemical shift of the current complex and its concentration (see blocks 14-17 in Fig.2) to be determined. Cycling all numbers from 0 to 2^L-1 is equivalent to generating all possible complexes having length *L* (cycle using variable *Complex* in Figs.1,2). Finally, an outer cycle starting from the monomeric form 1 up to the initially given maximum length of the complex *N* fully completes the generating procedure (cycle using variable *l* in Figs.1,2).

Summation of the concentrations of the complexes inside the cycles (blocks 23,24 in Fig.1) results in the overall concentration being equivalent to the mass conservation law in the analytical models [15-17]. Sequential summation of the multiplications of the concentration of the generated complex *C* on its chemical shift (block 23 in Fig.2) enables the overall chemical shift to be obtained for either *X* (as in Fig.2) or *Y*. It follows that the physical meaning of the stochastic algorithm is quite straightforward: expressions (3) were obtained in an analytical form for the analytical models, whereas a similar procedure is provided algorithmically in the stochastic model, using a set of program cycles, describing the law of mass action, the mass conservation law and the additive model for chemical shift [13,18].

It should be noted that cycling over all possible complexes in the stochastic model results in a pair of equivalent complexes (*e.g.* $X_iY_jX_k \equiv X_kY_jX_i$) as in the analytical models [17]. Hence, if the current complex *C* and a reversed complex C^1 are consistent with $C \le C^1$, it means that the *C*

complex is generated for the first time and should be included in the calculations; otherwise the *C* complex is ignored. This condition is embodied in block 4 of the general algorithm (see Figs.1,2) and reflects the role of the coefficient 1/2 in the basic model [16] and the revised summation in the generalized model [17].

It is significant that the stochastic algorithm provides a programmed access to every generated complex, enabling any conditions to be applied to the method of calculation of concentrations and the chemical shift and hence to subsequent expansion of reaction schemes (1) or (2) without any limitations. An example of the advantage of the stochastic approach is the introduction of the cooperativity parameter (σ) into the computational scheme of the hetero-association analysis.

As a result of the programmed access to every generated complex it is possible to determine the number of aggregates of *X* or *Y* type and, consequently, to calculate the resultant cooperativity coefficient for the current complex (block 22 and variable *Koeff* in Figs.1,2). After that, the calculated coefficient is used as a multiplier for the concentrations (blocks 23,24 in Fig.1) and chemical shifts (block 23 in Fig.2) of the complex. It is evident that when $\sigma_X = \sigma_Y = 1$ the stochastic cooperative algorithm gives similar results to those for the non-cooperative model.

3. Discussion

3.1 *Analysis of the hetero-association parameters calculated using different non-cooperative models.*

The experimental concentration and temperature dependences of proton chemical shifts, obtained under similar solution conditions for all systems studied (*T*=298K, 0.1 mol/l Na-phosphate buffer, pH 7.1) [9,10,15-17,19,20], have been used to calculate the hetero-association parameters of aromatic molecules using the basic (BASE) [15,16], generalized (GEN) [17] and stochastic non-cooperative (STOCH-NC) and cooperative (STOCH-C) models in this work. A detailed description of the conditions of the NMR measurements and computational procedure for determining the

hetero-association parameters of aromatic molecules is given in ref. [15] and the results of the calculations of equilibrium hetero-association constants are summarized in Table 1 for different aromatic systems.

It is seen from Table 1 that the equilibrium constants (K_C) for hetero-association of aromatic ligands calculated using the basic model (BASE) are greater than K_C (GEN) values for all the molecular systems studied. This is obviously due to the different assumptions used in the two theoretical approaches; *i.e.* it was concluded previously [17] that the main reason for the lower value of K_C calculated using the GEN model compared with the BASE model results from inclusion of "edge effects" in the GEN model. Introduction of edge effects (*i.e.* the dependence of the chemical shift on the position of the molecule within an aggregate or hetero-complex) results in lowering the average contribution of hetero-association reactions to the overall dynamic equilibrium and so to a decrease in the K_C value [17]. Hence, the difference in the calculated parameters using the BASE and GEN models depends on the relative contribution of hetero-association reactions to the overall dynamic equilibrium in solution.

In order to estimate the contribution of hetero-association reactions to the overall dynamic equilibrium in solution, it is reasonable to introduce a numerical characteristic of the hetero-association factor, *i.e.* the relative weight (f_C) of the hetero-association parameter, K_C, as follows:

$$f_C = \frac{K_C}{K_X + K_Y + K_C}. \tag{10}$$

Values of f_C (%) are summarized in Table 2 using K_C values of the BASE model. Values of f_C vary from 1 to 78% indicating a range from a very small contribution of hetero-association to the equilibrium ($f_C \sim 1\%$) to a very large contribution ($f_C \sim 78\%$).

It is also possible by equation (11) to calculate the relative difference, ε, of the hetero-association constants between the basic and generalized (BASE/GEN), basic and non-cooperative stochastic (BASE/STOCH-NC), generalized and non-cooperative stochastic (GEN/STOCH-NC) models:

$$\varepsilon(Mod1 - Mod2) = \left| \frac{K_C(Mod1) - K_C(Mod2)}{K_C(Mod2)} \right| \cdot 100\% \qquad (11)$$

Using data in Table 1 the calculated values of ε as a function of f_C are summarized in Table 2 and those for the relative difference of the calculated parameters between the BASE and GEN models presented in Fig.3. It is shown in Fig.3 that the variation in ε depends substantially on the magnitude of the hetero-association constant and, on average, increases proportionally to K_C. The deviation of ε between the BASE and GEN models is not greater than ~30% for relatively small contributions of the hetero-association to the overall dynamic equilibrium ($f_C \leq 1/3$), e.g. for hetero-association of aromatic drugs with caffeine ("ligand+CAF"). On the other hand an increase of the hetero-association factor f_C results in much greater values of ε, being the most pronounced for hetero-complexes of aromatic drugs with daunomycin ("ligand+DAU") (see Table 2). Hence it may be concluded that utilization of the basic model of hetero-association is most likely to be correct for descriptions of the mixed solutions of aromatic molecules characterized by relatively small contributions of the hetero-association reactions ($f_C \leq 1/3$), when compared to those for relatively large contributions ($f_C > ca.0.4$).

It is worth noting that the difference between the results of calculations using the BASE and GEN models depends not only on the magnitudes of the equilibrium hetero-association constants but also on the magnitudes of the equilibrium constants of the self-association of the interacting molecules X and Y. In particular, the NOV+CAF and AO+CAF systems are characterized by relatively low values of the hetero-association factor $f_C < 6\%$ (Table 2), whereas the relative difference between the hetero-association constants, derived from the BASE and GEN models, reaches rather high values (~20%) for systems with a hetero-association factor $f_C \leq 1/3$. The latter may be due to the influence of edge effects in the aggregates of NOV and AO molecules forming hetero-complexes with CAF, because NOV and AO are characterized by the highest magnitudes of equilibrium self-association constants for all the molecular systems studied (see Table 1).

It is interesting to note that the average differences in the calculated parameter, $\varepsilon\%$, between the basic and stochastic models appear to be similar to those differences using the BASE and GEN models (Table 2). However, if we exclude molecular systems with high values of f_C (*i.e.* PF+DAU, EB+DAU and PI+DAU), which are probably stabilized by intermolecular H-bonds [10,16], then the value of ε (BASE/STOCH-NC) of 18.7% is approximately 1.3 times smaller than ε (BASE/GEN) of 24.7% for the hetero-complexes studied. Such an effect obviously results from the process of averaging in the STOCH-NC model, which uses cycling over all possible molecular associations in the mixed solution.

It is seen from Table 2 that, with respect to the STOCH-NC model, the GEN model gives more consistent results than the basic model. Comparison of the calculated values of the hetero-association constants using the GEN and STOCH-NC methods indicates that the magnitude of their relative difference, ε, has no systematic correlation with f_C values and is no greater than ~30% (Table 2), which approximately corresponds to the standard error of the determination of equilibrium association constants from NMR experiments (Table 1). These results indicate that there is only a relatively small contribution of hetero-complexes with more than two hetero-stacks in the dynamic equilibrium in solution for all the molecular systems studied. The difference between the hetero-association constants for "ligand-CAF" systems, characterized by low hetero-association factors ($f_C \leq 1/3$), does not exceed 20%, which confirms the assumptions made previously [9,15,19] that the effect of hetero-complexes of type *XYX* when $K_X >> K_Y$ and the effect of hetero-complexes with a number of hetero-stacks more than two are insignificant in mixed solutions of these aromatic molecules. However, a relatively high value of f_C in "dye-DAU" systems (Table 2) results in an increase in the difference between the GEN and STOCH-NC models up to ~30%.

3.2 *Effect of the self-association cooperativity parameter (σ) on the hetero-association parameters of aromatic molecules.*

Aromatic drugs studied in this work are characterized by different cooperativity parameters, σ [13,18], and may be classified as cooperative, non-cooperative and anti-cooperative as presented in Table 3. The cooperativity parameters of self-association of aromatic drugs are included in the calculations of K_C using the cooperative stochastic model (STOCH-C) and the results are summarized in Table 1. The relative differences of the hetero-association constants between non-cooperative and cooperative stochastic models, ε (STOCH-NC / STOCH-C), calculated from data on aromatic molecules in Tables 1 and 3 are presented in Table 2. Analysis of the results indicates that there is little correlation between the hetero-association factor f_C and the relative deviation of the hetero-association constants, ε, for the systems studied. The maximum value of ε is less than 30% and the mean deviation only *ca*. 7% giving support to the idea that the cooperativity factors of the self-association of aromatic molecules have little effect on the calculated hetero-association parameters. However, some conclusions may be drawn with respect to the observed correlations between the cooperativity parameters, σ, of the self-association of aromatic molecules and the deviations of the hetero-association parameters.

The molecular systems of hetero-association in Table 2 may be grouped using different combinations of cooperativity modes of the interacting molecules, according to the classification of aromatic drugs given in Table 3. Four different combinations of cooperativity modes may be considered for the different molecular systems investigated in mixed solutions: cooperative/non-cooperative; non-cooperative/anti-cooperative; cooperative/anti-cooperative; and non-cooperative/non-cooperative. The results have been arranged in Table 4 in terms of the decrease in the mean value of the relative difference of the hetero-association constants, ε (STOCH-NC / STOCH-C)%.

Introduction of a cooperativity parameter in the cooperative model of self-association results in multiplication of the concentration of the *i*-th aggregate by the cooperativity parameter [13,18] and so expressions for concentrations of hetero-complexes also contain components multiplied by the cooperativity parameters of the interacting molecules. Hence, one may expect the largest

differences in hetero-association parameters between non-cooperative and cooperative stochastic models for the molecular systems "*Non-coop./Anti-coop.*" and "*Coop./Non-coop.*". It is seen from Table 4 that the highest mean value of deviation is observed for "*Coop./Non-coop.*" hetero-association (15.8%) and the lowest one relates to "*Non-coop./Non-coop.*" hetero-association (1.3%) which is consistent with the assumption presented above. The molecular systems "*Non-coop./Anti-coop.*" and "*Coop./Anti-coop.*" have similar deviations within experimental error (Table 4), which indicate that the effect of "compensation" of the calculated hetero-association parameters will take place for "*Coop./Anti-coop.*" systems and the absence or practically very small deviation for non-cooperative hetero-association.

3.3 *Conclusions.*

1. Stochastic non-cooperative (STOCH-NC) and cooperative (STOCH-C) models have been developed for analysis of the hetero-association reactions of aromatic molecules using NMR data. The proposed approaches have no limitations on the types of associations of aromatic molecules in solution and may be considered as the most general models (within the limitations of the NMR experiment).

2. A comparative analysis of the indefinite non-cooperative models of hetero-association of aromatic molecules: basic (BASE) [15,16], generalized (GEN) [17] and stochastic (STOCH-NC) has shown:

   (i)   The BASE analytical model is mainly valid to describe molecular systems characterized by relatively low contributions of hetero-association reactions (hetero-association factor $f_C \leq 1/3$);

   (ii) The GEN analytical model gives results in agreement with the STOCH-NC algorithmic model within the error limits of $\leq 30\%$ for all the systems studied. It is concluded that GEN model of hetero-association of aromatic molecules may be applied to any aromatic compounds with different self-association properties.

3. The computational algorithm of the stochastic model (STOCH-C) enabled the hetero-association parameters to be calculated for the first time by taking into account the cooperativity factor (σ) of the self-association of aromatic molecules and to estimate its effect on the multi-component equilibrium in solution.

4. It has been found that differences between the calculated parameters using the STOCH-NC and STOCH-C models are not greater than ~30% for all the molecular systems investigated and the differences depend substantially on the type of cooperativity of the self-association of the interacting molecules.

The proposed algorithm of the stochastic model has a great potential for detailed investigations of the interactions of aromatic molecules in solution because it is not limited to reaction scheme (2).

 References

1. Gale E.F., Cundliffe E., Reynolds P.E., Richmond M.H., Waring M.J. The molecular basis of antibiotic action. 2nd Edn. L.: John Wiley, 1981. 500 p.

2. Iliakis G., Nusse M., Ganapathi R., Egner J., Yen A. Differential reduction by caffeine of adriamycin induced cell killing and cell cycle delays in Chinese hamster V79 cells. Int. J. Radiat. Oncol. Biol. Phys. 1986, **12**, 1987-1995.

3. Traganos F., Kapuscinski J., Darzynkiewicz Z. Caffeine modulates the effects of DNA-intercalating drugs in vitro: a flow cytometric and spectrophotometric analysis of caffeine interactions with novantrone, doxorubicin, ellipticine, and the doxorubicin analogue AD198. Cancer Res., 1991, **51**, 3682-3689.

4. Armitage J.O. The role of mitoxantrone in non-Hodgkin's lymphoma. Oncology, 2002, **16**, 490-560.

5. Kimura H., Aoyama T. Decrease in sensitivity to ethidium bromide by caffeine, dimethylsulfoxide or 3-aminobenzamide due to reduced permeability. J. Pharmacobiodyn., 1989, **12**, 589-595.

6. Larsen R.W., Jasuja R., Hetzler R.K., Muraoka P.T., Andrada V.G., Jameson D.M. Spectroscopic and molecular modelling studies of caffeine complexes with DNA intercalators. Biophys. J., 1996, **70**, 443-452.

7. Zdunek M., Piosik J., Kapuscinski J. Thermodynamical model of mixed aggregation of ligands with caffeine in aqueous solution. Part II. Biophys. Chem., 2000, **84**, 77-85.

8. Lyles M.B., Cameron I.L. Interactions of the DNA intercalator acridine orange, with itself, with caffeine, and with double stranded DNA. Biophys. Chem., 2002, **96**, 53-76.

9. Davies D.B., Veselkov D.A., Djimant L.N., Veselkov A.N. Hetero-association of caffeine and aromatic drugs and their competitive binding with a DNA oligomer. Eur. Biophys. J., 2001, **30**, 354-366.

10. Davies D.B., Veselkov D.A., Veselkov A.N. NMR determination of the hetero-association of phenanthridines with daunomycin and their competitive binding to a DNA oligomer. Eur. Biophys. J., 2002, **31**, 153-162.

11. Lim L.-Y., Go M.-L. Caffeine and nicotinamide enhances the aqueous solubility of the antimalarial agent halafantrine. Eur. J. Pharm. Sci., 2000, **10**, 17-28.

12. Veselkov A.N., Lantushenko A.O., Chubarov A.S., Veselkov D.A., Davies D.B. ^{1}H NMR analysis of the self-association of riboflavin-mononucleotide and its complexation with nicotinamide in an aqueous solution. Phys. Chem. (in Russian), 2002, **76**, 1350-1356.

13. Martin R.B. Comparisons of indefinite self-association models. Chem.Rev., 1996. **96**, 3043-3064.

14. Weller K., Shutz H., Petri I. Thermodynamical model of indefinite mixed association of two components and NMR data analysis for caffeine-AMP interaction. Biophys. Chem., 1984, **19**, 289-298.

15. Davies D.B., Veselkov D.A., Veselkov A.N. Structure and thermodynamics of the hetero-association of aromatic molecules in aqueous solution determined by NMR spectroscopy. Molecular Phys., 1999, **97**, 439-451.

16. Davies D.B., Veselkov D.A., Kodintsev V.V., Evstigneev M.P., Veselkov A.N. ^{1}H NMR investigation of the hetero-association of aromatic molecules in aqueous solution: factors involved in the stabilization of complexes of daunomycin and acridine drugs. Molecular Phys., 2000, **98**, 1961-1971.

17. Veselkov A.N., Evstigneev M.P., Veselkov D.A., Davies D.B. A generalized NMR-derived statistical-thermodynamical model of hetero-association of aromatic molecules in aqueous solution. J. Chem. Phys., 2001, **115**, 2252-2266.

18. Davies D.B., Djimant L.N., Veselkov A.N. ^{1}H NMR investigation of self-association of aromatic drug molecules in aqueous solution. Structural and thermodynamical analysis. J. Chem. Soc. Faraday Trans., 1996, **92**, 383-390.

19. Davies D.B., Veselkov D.A., Evstigneev M.P., Veselkov A.N. Self-association of the antitumour agent novatrone (mitoxantrone) and its hetero-association with caffeine. J. Chem. Soc. Perkin Trans. 2, 2001, 61-67.

20. Veselkov A.N., Lantushenko A.O., Evstigneev M.P., Veselkov D.A., Davies D.B. Comparative analysis of the hetero-association of azido-analogues of ethidium with antibiotic daunomycin in aqueous solution using ^{1}H NMR spectroscopy. Ukr. Chem. J., 2002, **68**, 78-83.

21. Veselkov A.N., Lantushenko A.O., Veselkov D.A., Davies D.B. Structural and thermodynamical analysis of the self-association of phenanthridine dyes in aqueous solution based on ^{1}H NMR spectroscopy data. J. Struct. Chem., 2002, **43**, 253-260.

22. Eaton R.J., Veselkov D.A., Baranovsky S.F., Osetrov S.G., Djimant L.N., Davies D.B., Veselkov A.N. ^{1}H-NMR studies of self-association of anthracycline antibiotic molecules in aqueous solution. Chem. Phys. Rep., 2000, **19**, 387-399.

Table 1. Parameters of hetero-association of aromatic molecules X and Y calculated using the basic (BASE), generalized (GEN), stochastic non-cooperative (STOCH-NC) and cooperative (STOCH-C) models from NMR spectroscopic data [a,b,c]

X+Y	K_X	K_Y	K_C			
			BASE [ref.]	*GEN* [ref.]	*STOCH/NC*	*STOCH/C*
PF+CAF	700±70	11.8±0.3	160±17 [9]	129±27	121±26	137±26
AO+CAF	4600±600		264±21 [9]	224±45	270±54	368±70
EB+CAF	305±14		62±4 [9]	57±11	56±11	57±11
PI+CAF	63±6		28±5 [9]	27±5	27±5	27±5
DAU+CAF	720±130		72±4 [9]	62±12	55±11	51±10
NOV+CAF	28000±8000		324±40 [19]	280±56	313±60	320±65
AMD+CAF	1440±160		246±48 [9]	226±45	245±50	260±50
PF+DAU	700±70	720±130	2080±150 [16]	1180±230	1000±200	1055±210
AO+DAU	4600±600		2910±520 [16]	1864±373	2024±400	2240±450
EB+DAU	305±14		3580±580 [10]	1740±348	1564±310	1480±300
PI+DAU	63±6		720±80 [10]	454±91	406±81	385±75
EMB+DAU	276±17		660±100 [20]	430±86	586±117	556±110
EDC+DAU	19±3		320±65 [20]	245±49	290±60	265±55
PF+EB	700±70	305±14	690±50 [17]	520±50 [17]	500±50	550±110
PI+EB	63±6	305±14	126±9 [15]	102±15	93±18	94±18

[a] 500 MHz NMR measurements made for solutions in 0.1M phosphate buffer, pD=7.1, T=298K [9,10,15-17,19,20].

[b] Abbreviations used: PF – proflavine, AO – acridine orange, EB – ethidium bromide, PI – propidium iodide, CAF – caffeine, DAU – daunomycin, NOV - novatrone, AMD – actinomycin D, EMB – ethidium monoazide, EDC – ethidium diazide.

[c] The experimental results are all taken from the literature where the differences between the experimental and predicted chemical shifts (cs) are calculated in terms of the discrepancy function (the sum of the square of the differences between calculated and measured cs in titration experiment comprising at least 15 dilution steps). The value of the discrepancy function for all systems studied and all models applied was not greater than *ca.* 10^{-5}. This corresponds to an average deviation between experimental and predicted cs per data point of *ca.* 0.0002ppm which is comparable to the experimental error in measurements of chemical shifts.

Table 2 Relative differences $\varepsilon\%$ of the hetero-association constants, calculated using different

models as a function of the hetero-association factor, f_C

System	f_C, %	ε, %			
		BASE-GEN	BASE-STOCH/NC	GEN-STOCH/NC	STOCH/NC-STOCH/C
PF+CAF	18.4	24.0	32.2	6.6	11.7
AO+CAF	5.4	17.9	2.2	17.0	26.6
EB+CAF	16.4	8.8	10.7	1.8	1.8
PI+CAF	27.2	3.7	3.7	0.0	0.0
DAU+CAF	9.0	16.1	30.9	12.7	7.8
NOV+CAF	1.1	15.7	3.5	10.5	2.2
AMD+CAF	14.7	8.8	0.4	7.8	5.8
PF+DAU	59.4	76.3	108.0	18.0	5.2
AO+DAU	35.4	56.1	43.8	7.9	9.6
EB+DAU	77.7	105.7	128.9	11.3	5.7
PI+DAU	47.9	58.6	77.3	11.8	5.5
EMB+DAU	39.9	53.5	12.6	26.6	5.4
EDC+DAU	30.2	30.6	10.3	15.5	9.4
PF+EB	40.7	32.7	38.0	4.0	9.1
EB+PI	25.5	23.5	35.5	9.7	1.1
Mean		35.5	35.9	10.8	7.1

Table 3 Cooperativity parameter (σ) for the self-association of different drugs in 0.1M phosphate

buffer, pD=7.1

Non-cooperative		Cooperative		Anti-cooperative	
Drug	$\sigma^{ref.}$	Drug	$\sigma^{ref.}$	Drug	$\sigma^{ref.}$
EB	0.89±0.06 [18]				
PI	0.98±0.05 [15]				
EMB	0.96±0.08 [21]	PF	0.42±0.06 [18]	DAU	1.34±0.06 [22]
EDC	0.97±0.04 [21]	AO	0.45±0.05 [18]	AMD	1.49±0.10 [18]
NOV	0.98±0.04 [19]				
CAF	1.08±0.02 [9]				

Table 4 Relative differences ε% of the hetero-association constants with respect to the mode of cooperativity of the interacting molecules

| Coop./Non-coop., ε %		Non-coop./Anti-coop. , ε %		Coop/Anti-coop. , ε %		Non-coop/Non-coop, ε %	
PF+CAF	11.7	EB+DAU	5.7	PF+DAU	5.2	EB+CAF	1.8
AO+CAF	26.6	PI+DAU	5.5	AO+DAU	9.6	PI+CAF	0.0
PF+EB	9.1	EMB+DAU	5.4			EB+PI	1.1
		EDC+DAU	9.4			NOV+CAF	2.2
		CAF+DAU	7.8				
		CAF+AMD	5.8				
Mean	~15.8		~6.6		~7.4		1.3

FIGURE LEGENDS

Fig.1 Algorithm for calculating the overall concentration of compounds X and Y

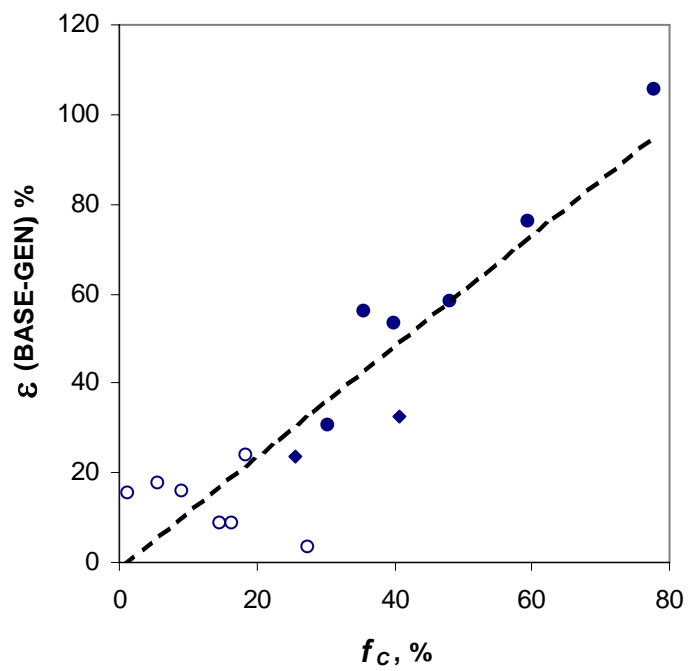Fig. 2. Algorithm for calculating the weighted average proton chemical shift of compound X and Y

Fig. 3. Dependence on the hetero-association factor f_C, % of the relative differences of the hetero-association constants, ε (BASE-GEN), %, calculated using the basic and generalized models: $\bullet$ – DAU+ligand; $\circ$ – CAF+ligand; $\blacklozenge$ – PF+EB, EB+PI.
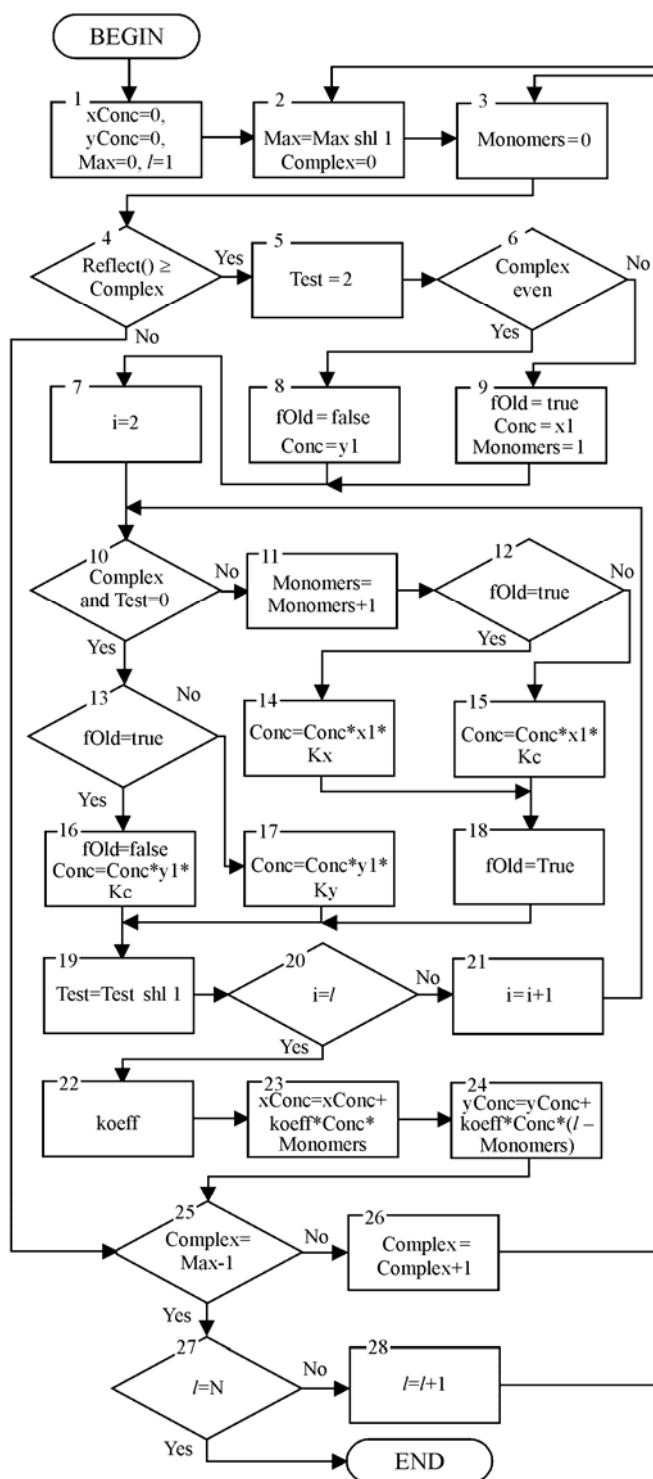
A.N. Veselkov et al., Fig.1
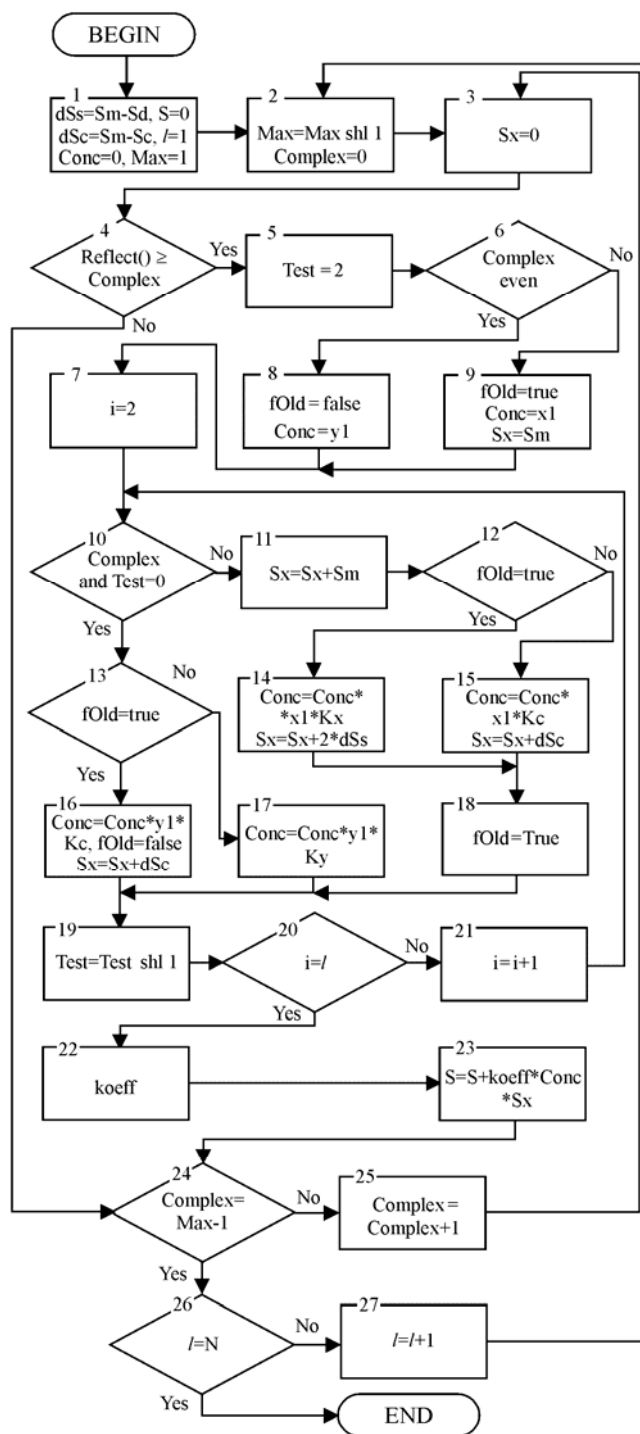
A.N. Veselkov et al., Fig.2

A.N. Veselkov et al., Fig.3