CrossMark

General statistical-thermodynamical treatment of one-dimensional multicomponent molecular hetero-assembly in solution

Anatoly S. Buchelnikov [a,b,*], Vladyslav P. Evstigneev [a], Maxim P. Evstigneev [a,b]

[a] Department of Physics, Sevastopol National Technical University, 33 Universitetskaya Str., Sevastopol 99053, Ukraine
[b] Department of Biology and Chemistry, Belgorod State University, 85 Pobedy Str., Belgorod 308015, Russia

A B S T R A C T

A general treatment has been developed for the multicomponent one-dimensional non-covalent molecular hetero-assembly in solution using transfer matrix and sequence generating function approaches. The main result is set of equations, which allows one to obtain any thermodynamical quantities of the multicomponent system and, in particular, experimental observable enabling one to get all equilibrium parameters of molecular interaction. The matrix form of presentation of the key equations allows their direct incorporation into matrix-oriented mathematical software, which leads to a two orders of magnitude increase in speed of calculation of as compared to other known approaches.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Non-covalent hetero-assembly of small molecules in solution is currently attracting great attention due to extensive use of this phenomenon in molecular electronics [1–3], laser physics [4,5], supramolecular chemistry [2,6–9], advanced materials [10–13], numerous biochemical applications [2,14,15] etc. The key feature of this process is the formation of nanoscale aggregates, which modify the macroscopic physical properties of solute resulting in outcomes of practical importance ranging from discovering of new perspective materials to treatment of human diseases [16].

The most well-known partial cases of such aggregations are one-dimensional, one-component self-assembly (see Refs. [9,17] for review) and two-component hetero-assembly (see Refs. [8,18] for review) leading to formation of aggregates, containing one type or two different types of molecules, respectively, ordered along a certain axis. Recently, a focus has been given to assemblies containing higher numbers of different molecules (multicomponent hetero-assembly) and direct applications of such approaches have been reported in different fields of chemical science. For instance, the hetero-assembly process can significantly improve the solubility of hydrophobic drugs [19,20]. It has also appeared to be useful for creation of various supramolecular structures [21,22] as well as

for regulation of medico-biological activity of drugs controlled by an investigator [23].

Although the set of physical laws and commonly introduced basic assumptions behind the description of multicomponent hetero-assembly makes such systems relatively clear in terms of analysis, there are two difficulties which need to be faced when dealing with molecular aggregation:

(i) setting up a link to the experimental observables, and
(ii) correct treatment of so-called 'reflected' complexes, i.e. physically identical non-symmetric complexes, resulting from infinite summation necessary to account for all possible types of complexes able to form in solution (for more discussion of the problem see Refs. [18,24]). An exception is in special cases of intermolecular assembly [9,22] or in account for cooperative effects [9].

Both difficulties have long been overcome for one-component self-assembly (reviewed in Ref. [17]) and have recently been overcome for two-component hetero-assembly [24]. The closest approaches to similar solutions for multicomponent systems have been suggested using the algorithmic approach [25], transfer matrix formalism [22] and sequence generating function formalism [26]. However, these solutions are all incomplete, viz. the algorithmic approach [25] is limited by the finite lengths of aggregates, the transfer matrix approach [22] does not treat 'reflected' complexes at all and therefore considers experimental observable incorrectly and the sequence generating function approach [26] fully accounts for the 'reflected' complexes but does not provide

a link to experimental observables. In the present work we provide a general treatment of the multicomponent one-dimensional molecular hetero-assembly in solution which fully accounts for the two aforementioned difficulties. The treatment uses various standard statistical-thermodynamical approaches which yield identical outcomes and which reduces to all major known partial cases previously described in literature.

2. Results and discussion

2.1. General approach to the treatment of the multicomponent hetero-assembly

Three basic statistical-thermodynamical approaches have so far been used in the literature for treating multicomponent systems and accounting for all possible types of molecular complexes in solution, viz. algebraic (or combinatorial), transfer matrix (TM) and sequence generating function (SGF) approaches. All are based on a similar set of physical assumptions [24]: the mass conservation law, the law of mass action and the independence of the equilibrium constant on the number of molecules in an aggregate (i.e. non-cooperativity). The algebraic approach [18], which is based on explicit treatment of all possible molecular complexes in solution, is inefficient for systems with more than two components and will not be used in the present work.

The basic idea of the TM and SGF approaches (recently reviewed in Ref. [27]) is a derivation of the partition function, Z_n, of linear aggregates (or linear lattice made of n cells) containing n molecules of N different types (to be further referred to as n-mer aggregate). The general form of Z_n can be given as [28]:

$$Z_n = \mathbf{u}\mathbf{W}^{n-2}\mathbf{v}^\mathrm{T}, \tag{1}$$

where the elements of $N \times N$ matrix, $\mathbf{W}$, and $1 \times N$ vectors, $\mathbf{u}$ and $\mathbf{v}$, depend on selection of a particular model [28]. In addition, the $\mathbf{u}$ and $\mathbf{v}$ vectors define the ends of growing aggregates.

The difference between TM and SGF approaches is basically the way in which the $\mathbf{W}$ matrix is filled, i.e. whether the cells of the matrix contain probabilities or statistical weights (TM, for instance, see Ref. [22]), or partition functions of oligomers made of single-type molecules (SGF, for instance, see Ref. [26]).

The specificity of the hetero-assembly of small molecules (unlike, for example, well-developed lattice models, reviewed in Ref. [27]) is the possibility of formally infinite aggregation resulting in a requirement to take into consideration the distribution of all possible types of molecular complexes ranging from unity (monomer) to infinity. In such a case the grand partition function, Ξ, of the N-component system taking into account all possible types of complexes can be obtained via summation of the partition functions (1) over all natural n:

$$\Xi = \sum_{n=1}^{\infty} Z_n. \tag{2}$$

The main goal of further application of Ξ to a particular system is the description of experimental concentration (or temperature) dependences of certain experimental observable, ξ (e.g. molecular absorption in spectrophotometry, heat effect in calorimetry, chemical shift in NMR), which enables one to obtain the equilibrium constants (or other thermodynamical parameters) of hetero-assembly, K_{ij}, between i- and j-type molecules, and other parameters, for instance, observable parameter, ξ_{ij}, of i-type molecule being in stack with j-type molecule. Once these quantities are known, the distribution of molecular aggregates in solution is considered to be fully determined. Such a link between equilibrium parameters and experimental observations has been used in the majority of models of molecular hetero-assembly, cited in the introductory section. In the present study we shall also follow this strategy.

2.2. Derivation of the grand partition function using the TM approach

Let us introduce the statistical weight, $K_{ij}C_j$, of the j-type molecule being in a hetero-stack with i-type molecule within an aggregate. These quantities arranged in the $N \times N$ matrix form give a transfer matrix, $\mathbf{M}$:

$$\mathbf{M} = \begin{pmatrix} K_{11}C_1 & K_{12}C_2 & \cdots & K_{1N}C_N \\ K_{21}C_1 & K_{22}C_2 & \cdots & K_{2N}C_N \\ \vdots & \vdots & \ddots & \vdots \\ K_{N1}C_1 & K_{N2}C_2 & \cdots & K_{NN}C_N \end{pmatrix}, \tag{3}$$

which bears the same meaning as $\mathbf{W}$ in Eq. (1).

Raising $\mathbf{M}$ to the power of $n-1$ enables one to enumerate all possible arrangements of n molecules of N different types within an n-mer aggregate. Hence, according to Eq. (1) the partition function of an n-mer aggregate takes the form:

$$Z_n = \mathbf{c}\mathbf{M}^{n-1}\mathbf{1}^\mathrm{T}, \tag{4}$$

where the vectors, $\mathbf{c} = (\begin{matrix} C_1 & C_2 & \ldots & C_N \end{matrix})$ and $\mathbf{1} = (\begin{matrix} 1 & 1 & \ldots & 1 \end{matrix})$, correspond to vectors $\mathbf{u}$ and $\mathbf{v}$ in Eq. (1), respectively.

Eq. (4) along with Eq. (2) can be further used in order to derive the grand partition function of the N-component hetero-assembly system:

$$\Xi = \sum_{n=1}^{\infty} Z_n = \sum_{n=1}^{\infty} \mathbf{c}\mathbf{M}^{n-1}\mathbf{1}^\mathrm{T} = \mathbf{c}(\mathbf{I} - \mathbf{M})^{-1}\mathbf{1}^\mathrm{T}, \tag{5}$$

where $\mathbf{I}$ is the $N \times N$ unity matrix. Infinite summation in Eq. (5) takes place only if the spectral radius of $\mathbf{M}$ by absolute value is less than unity. However, this condition is always valid if the monomer concentrations in $\mathbf{c}$ are determined from the mass conservation law, acting as a physical constraint for the infinite geometric series in Eq. (5).

The grand partition function (5) includes statistical weights of all possible complexes formed in solution, i.e. accounts for all possible arrangements of molecules of different types. The whole set of possible complexes also includes the pairs of mirror symmetrical complexes (for example, complexes of $C_jK_{ji}C_iK_{ii}C_i$ and $C_iK_{ii}C_iK_{ij}C_j$ type) which have been previously denoted as 'reflected' complexes [18,24,25]. They become physically indistinguishable under the condition $K_{ij} = K_{ji}$, which is the case most frequently met in one-dimensional hetero-assembly in solution (account of the distinguishability, $K_{ij} \neq K_{ji}$, was accomplished in Ref. [29] for two-component systems and is dependent on specific features of complexation of the i-th and j-th molecules). For that reason, we shall evaluate the grand partition function for the case of indistinguishable 'reflected' complexes.

Considering that the 'reflected' complexes are physically equivalent, the grand partition function (5) may be corrected by analogy to that previously reported for two-component systems [18,24]

$$\Xi_{\mathrm{corr}} = \frac{\Xi + \Xi_{\mathrm{symm}}}{2}, \tag{6}$$

where Ξ_{symm} is the partition function of symmetrical with respect to their geometrical centre complexes (for instance, complexes of $C_jK_{ji}C_iK_{ii}C_iK_{ij}C_j$ or $C_iK_{ii}C_jK_{ji}C_i$ type).

In order to derive Ξ_{symm} let us introduce supplementary matrices

$$\mathbf{M_S} = \begin{pmatrix} K_{11}C_1 & 0 & \cdots & 0 \\ 0 & K_{22}C_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & K_{NN}C_N \end{pmatrix}, \quad \mathbf{M_d} = \mathbf{M} \circ \mathbf{M} = \begin{pmatrix} K_{11}^2C_1^2 & K_{12}^2C_2^2 & \cdots & K_{1N}^2C_N^2 \\ K_{21}^2C_1^2 & K_{22}^2C_2^2 & \cdots & K_{2N}^2C_N^2 \\ \vdots & \vdots & \ddots & \vdots \\ K_{N1}^2C_1^2 & K_{N2}^2C_2^2 & \cdots & K_{NN}^2C_N^2 \end{pmatrix}, \tag{7}$$

where sign '∘' denotes Hadamard product yielding element-wise matrix multiplication [30]. Raising $\mathbf{M_d}$ to the power of $n-1$ enables

one to derive the partition functions of the two 'halves' of symmetrical complexes with total length $2(n-1)$. Taking into account that in the centre of symmetry may sit a monomer or a homo-dimer of molecules of certain type, the partition function of an n-mer, corrected for the 'reflected' complexes, takes the form:

$$Z_{n,\text{symm}} = \mathbf{c}\mathbf{M_d}^{n-1}\mathbf{1}^{\text{T}} + \mathbf{c}\mathbf{M_S}\mathbf{M_d}^{n-1}\mathbf{1}^{\text{T}} = \mathbf{c}(\mathbf{I} + \mathbf{M_S})\mathbf{M_d}^{n-1}\mathbf{1}^{\text{T}}. \quad (8)$$

In Eq. (8) vector $\mathbf{c}$ (or $\mathbf{c}\mathbf{M_S}$) acts as a vector of statistical weights of the central monomer (or dimer) of the symmetrical complex. Summation of Eq. (8) over n gives the partition function of the symmetrical complexes:

$$\Xi_{\text{symm}} = \sum_{n=1}^{\infty} Z_{n,\text{symm}} = \sum_{n=1}^{\infty} \mathbf{c}(\mathbf{I} + \mathbf{M_S})\mathbf{M_d}^{n-1}\mathbf{1}^{\text{T}}$$
$$= \mathbf{c}(\mathbf{I} + \mathbf{M_S})(\mathbf{I} - \mathbf{M_d})^{-1}\mathbf{1}^{\text{T}}. \quad (9)$$

Finally, substituting Eqs. (5) and (9) into Eq. (6) enables one to derive the fully corrected grand partition function of the N-component hetero-assembly system:

$$\Xi_{\text{corr}} = \frac{1}{2}\mathbf{c}((\mathbf{I} - \mathbf{M})^{-1} + (\mathbf{I} + \mathbf{M_S})(\mathbf{I} - \mathbf{M_d})^{-1})\mathbf{1}^{\text{T}}. \quad (10)$$

2.3. Derivation of the grand partition function using the SGF approach

The SGF approach, originally proposed by Lifson [31] and then fully reviewed for chain polymers in Ref. [32], is based on representation of an n-mer by sub-chains containing elements of a single type. The sub-chains are given by corresponding sequence-generating functions. In the context of the n-mer aggregates in the present work, the Lifson's sub-chains are homo-aggregates made of single-type molecules, and the Lifson sequence generating functions are the partition functions of these homo-aggregates. Recently, the full derivation of the grand partition function for the N-component system corrected for 'reflected' complexes has been performed [26]. We provide a brief re-derivation of the Ξ_{corr} quantity compliant with the set of notations used throughout the present work.

The statistical weight of a homo-aggregate, containing k molecules of type j, can be written as

$$\zeta_j(k) = C_j(K_{jj}C_j)^{k-1}. \quad (11)$$

Hence, the sequence generating function for homo-aggregates of any length from unity to infinity made of j-type molecules are given as

$$\gamma_j = \sum_{k=1}^{\infty} \zeta_j(k) = C_j \sum_{k=1}^{\infty} (K_{jj}C_j)^{k-1} = \frac{C_j}{1 - K_{jj}C_j}. \quad (12)$$

The quantities $K_{ij}\gamma_j$ arranged in a form of a $N \times N$ matrix give matrix, $\mathbf{G}$, which is analogous to matrix $\mathbf{W}$ in Eq. (1):

$$\mathbf{G} = \begin{pmatrix} 0 & K_{12}\gamma_2 & K_{13}\gamma_3 & \cdots & K_{1N}\gamma_N \\ K_{21}\gamma_1 & 0 & K_{23}\gamma_3 & \cdots & K_{2N}\gamma_N \\ K_{31}\gamma_1 & K_{32}\gamma_2 & 0 & \cdots & K_{3N}\gamma_N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ K_{N1}\gamma_1 & K_{N2}\gamma_2 & K_{N3}\gamma_3 & \cdots & 0 \end{pmatrix}. \quad (13)$$

Zeros on the principal diagonal appear as a result of merging two neighbouring homo-aggregates into a single aggregate, which means that expression of $K_{ij}\gamma_j$ type does not bear meaning within the context of the SGF method.

Let us introduce vector $\boldsymbol{\gamma}$, containing sequence generating functions of homo-aggregates, which may be located on the edge of an n-mer:

$$\boldsymbol{\gamma} = (\gamma_1 \quad \gamma_2 \quad \gamma_3 \quad \cdots \quad \gamma_N). \quad (14)$$

Hence, the expression $\boldsymbol{\gamma}\mathbf{G}^{p-1}\mathbf{1}^{\text{T}}$ represents the sum of partition functions of all possible p homo-aggregates arranged sequentially. It follows that the grand partition function taking into account all combinations of homo-aggregates within all possible complexes can be written as

$$\Xi = \sum_{p=1}^{\infty} \boldsymbol{\gamma}\mathbf{G}^{p-1}\mathbf{1}^{\text{T}} = \boldsymbol{\gamma}(\mathbf{I} - \mathbf{G})^{-1}\mathbf{1}^{\text{T}}. \quad (15)$$

Eq. (15) is similar to that derived in Ref. [21] (Eq. (13)) with respect to ligand binding to linear lattice (except that Eq. (13) contains an additional term 1 standing for the contribution of free from ligand binding sites on the lattice) which is a consequence of similarity in basic approach used to derive partition functions in these two works.

Eq. (15) contains the same weakness as discussed above with respect to Eq. (5), i.e. overestimation of the 'reflected' complexes. Correction of Eq. (15) can be accomplished following the same approach used above to correct Ξ in TM formalism and based on Eq. (6).

The statistical weight of symmetrical homo-aggregates located on both ends of an n-mer has the form

$$\zeta_j^{(2)}(k) = C_j^2 (K_{jj}^2 C_j^2)^{k-1}. \quad (16)$$

Hence, the sequence generating function of two 'halves' of the symmetrical complexes can be written as

$$\gamma_j^{(2)} = \sum_{k=1}^{\infty} \zeta_j^{(2)}(k) = C_j^2 \sum_{k=1}^{\infty} (K_{jj}^2 C_j^2)^{k-1} = \frac{C_j^2}{1 - K_{jj}^2 C_j^2}. \quad (17)$$

Any symmetrical complex may contain only odd number of homo-aggregates. It follows that one of them will always sit in the centre of an n-mer. Employing the vector $\boldsymbol{\gamma}$ in Eq. (14) in order to represent this central homo-aggregate, and introducing matrix $\mathbf{G_d}$ as

$$\mathbf{G_d} = \begin{pmatrix} 0 & K_{12}^2\gamma_2^{(2)} & K_{13}^2\gamma_3^{(2)} & \cdots & K_{1N}^2\gamma_N^{(2)} \\ K_{21}^2\gamma_1^{(2)} & 0 & K_{23}^2\gamma_3^{(2)} & \cdots & K_{2N}^2\gamma_N^{(2)} \\ K_{31}^2\gamma_1^{(2)} & K_{32}^2\gamma_2^{(2)} & 0 & \cdots & K_{3N}^2\gamma_N^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ K_{N1}^2\gamma_1^{(2)} & K_{N2}^2\gamma_2^{(2)} & K_{N3}^2\gamma_3^{(2)} & \cdots & 0 \end{pmatrix}, \quad (18)$$

it is possible to derive the corrected partition function for the symmetrical complexes:

$$\Xi_{\text{symm}} = \sum_{p=1}^{\infty} \boldsymbol{\gamma}\mathbf{G_d}^{p-1}\mathbf{1}^{\text{T}} = \boldsymbol{\gamma}(\mathbf{I} - \mathbf{G_d})^{-1}\mathbf{1}^{\text{T}}. \quad (19)$$

Substituting Eqs. (15) and (19) into Eq. (6) the fully corrected grand partition function for the N-component hetero-assembly system can be obtained as:

$$\Xi_{\text{corr}} = \frac{1}{2}\boldsymbol{\gamma}[(\mathbf{I} - \mathbf{G})^{-1} + (\mathbf{I} - \mathbf{G_d})^{-1}]\mathbf{1}^{\text{T}}. \quad (20)$$

2.4. Comparison of the grand partition functions derived using the TM and SGF approaches

Let us show that the grand partition functions, Ξ_{corr}, derived using the TM (Eq. (10)) and SGF (Eq. (20)) approaches are identical. In order for this to be accomplished it is necessary to introduce supplementary matrix objects, which will also be used later to derive the experimental observables:

$$\mathbf{K} = \begin{pmatrix} K_{11} & K_{12} & \cdots & K_{1N} \\ K_{21} & K_{22} & \cdots & K_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ K_{N1} & K_{N2} & \cdots & K_{NN} \end{pmatrix}, \quad \mathbf{\Delta_C} = \begin{pmatrix} C_1 & 0 & \cdots & 0 \\ 0 & C_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C_N \end{pmatrix}. \quad (21)$$

As shown in Appendix A, incorporation of these objects into Eqs. (10) and (20) results in two identical equations for Ξ_{corr} suggesting full equivalence of TM and SGF approaches in the description of molecular hetero-assembly in N-component systems.

Once the Ξ_{corr} is obtained, further derivation of the equilibrium parameters of molecular interaction in solution is based on application of partition function formalism is already not linked rather than the TM or SGF approaches. The key specificity in application of the partition function formalism to the N-component system with respect to that published by other authors is the need to use general differentiation of scalar function by matrix. We shall derive all necessary equations enabling one to set up a link between experimental observable and molecular equilibrium in solution with the aid of Ξ_{corr} obtained above using the TM approach (Eq. (10)) which is more convenient than the SGF approach for direct implementation in matrix-oriented mathematical software.

2.5. Derivation of an equation for the mass conservation law

The total concentrations of all types of molecules in solution, C_{0i}, are commonly known in experiments and therefore can be used as a constraint in obtaining equilibrium parameters of molecular complexes from experimental titration curves. The total concentrations, C_{0i}, can be found by differentiation of the grand partition function (10) by the natural logarithm of the corresponding monomer concentration, C_i [24,33]. In general matrix formulation this procedure can be written as differentiation of a scalar function by a matrix:

$$\mathbf{\Delta_T} = \frac{\partial \Xi_{corr}}{\partial \tilde{\mathbf{\Delta}}_C}, \quad (22)$$

where $\mathbf{\Delta_T}$ and $\tilde{\mathbf{\Delta}}_C$ are diagonal $N \times N$ matrices which contain C_{0i} and $\ln C_i$ elements on their principal diagonals, respectively.

Evaluation of Eq. (22) to closed analytical form is given in Appendix B; the end result is

$$\mathbf{\Delta_T} = \frac{1}{2}\left[(\mathbf{I}-\mathbf{M})^{-1}\mathbf{1}^T\mathbf{1}\Delta_C(\mathbf{I}-\mathbf{M})^{-1} + (\mathbf{I}-\mathbf{M_d})^{-1}\mathbf{1}^T\mathbf{1}\Delta_C(\mathbf{I}+2\mathbf{M_S}+\mathbf{M_d})(\mathbf{I}-\mathbf{M_d})^{-1}\right]. \quad (23)$$

2.6. Derivation of an equation for the experimental observables

In the majority of known models describing multicomponent complexations of small molecules, the experimental observable, ξ_{0i} (molecular absorption in spectrophotometry, heat effect in calorimetry, chemical shift in NMR etc.) of molecules of the ith type is represented as an additive quantity (for example, see Refs. [17–20,22–26]), which can be obtained by the averaging of particular observables, $\xi_i^{(s)}$, over all s types of complexes (including monomers) with respect to their amounts in solution, i.e. mole fractions, $f_i^{(s)}$:

$$\xi_{0i} = \sum_s \xi_i^{(s)} f_i^{(s)}, \quad (24)$$

Expansion of Eq. (24) yields

$$\xi_{0i} = \sum_s \xi_i^{(s)} f_i^{(s)} = \xi_i^{(m)} f_i^{(m)} + \sum_{s \neq m} \xi_i^{(s)} f_i^{(s)} = \xi_i^{(m)} - \sum_{s \neq m} \Delta\xi_i^{(s)} f_i^{(s)}, \quad (25)$$

where $\xi_i^{(m)}$ is the experimental observable of the i-type molecule in the monomer state; $\Delta\xi_i^{(s)} = \xi_i^{(m)} - \xi_i^{(s)}$.

Another assumption commonly used in analysis of a one-dimensional assembly is the approximation of 'nearest neighbour', which considers that only the nearest molecules (in direct contact with the i-type molecule in a complex) influence the ξ_{0i} of that molecule (for example, see Refs. [17,18,22,24]). Within the framework of a one-dimensional hetero-assembly two possibilities exist, viz. (i) only one of the nearest neighbours affects ξ_{0i} (used, for example, in Ref. [22] and usually termed 'head-to-tail' or 'head-to-head'-type aggregation), and (ii) both nearest neighbours affect ξ_{0i} (used, for example, in Refs. [17,18,24] and usually termed 'sandwich'-type aggregation). In order to distinguish these cases let us introduce the parameter $t \in 1, 2$ which is the number of neighbours affecting the experimental observable of given molecule. Let ξ_{ij} be the experimental observable of an i-type molecule in contact with a j-type molecule (following definitions formulated in Ref. [24] at $i = j$ it is a homo-stack and at $i \neq j$ it is a hetero-stack). Due to the 'nearest neighbour' approximation one represents $\xi_i^{(s)}$ in Eq. (25) as a particular combination of $\xi_i^{(m)}$ and ξ_{ij} [18]. Hence, the experimental observable in Eq. (25) can be written in a more definite form

$$\xi_{0i} = \xi_i^{(m)} - \frac{t}{2}\left(\sum_{j=1}^N \Delta\xi_{ij} f_{ij}^{(d)} + \sum_{j=1}^N \Delta\xi_{ij} f_{ji}^{(d)}\right), \quad (26)$$

where $\Delta\xi_{ij} = \xi_i^{(m)} - \xi_{ij}$; quantities $f_{ij}^{(d)}$ and $f_{ji}^{(d)}$ are the mole fractions of i-j and j-i stacks, respectively. They can be expressed via the total concentrations, S_{ij} and S_{ji}, of these stacks as $f_{ij}^{(d)} = \frac{S_{ij}}{C_{0i}}$ and $f_{ji}^{(d)} = \frac{S_{ji}}{C_{0i}}$, which, in their turn, can be directly obtained from the grand partition function [24,33]:

$$S_{ij} = \frac{\partial \Xi_{corr}}{\partial \ln K_{ij}},$$

$$S_{ji} = \frac{\partial \Xi_{corr}}{\partial \ln K_{ji}}. \quad (27)$$

Taking into account the symmetry of any stack with respect to interchange of i and j, i.e. $K_{ij} = K_{ji}$, Eq. (26) can be rewritten as

$$\xi_{0i} = \xi_i^{(m)} - \frac{t}{C_{0i}}\sum_{j=1}^N \Delta\xi_{ij} S_{ji}. \quad (28)$$

Further generalisation of Eq. (28) in matrix form requires introduction of the following matrix objects:

$$\xi_\mathbf{m} = \begin{pmatrix} \xi_{m1} & \cdots & \xi_{m1} \\ \vdots & & \vdots \\ \xi_{mN} & \cdots & \xi_{mN} \end{pmatrix}, \quad \xi = \begin{pmatrix} \xi_{11} & \cdots & \xi_{1N} \\ \vdots & \ddots & \vdots \\ \xi_{N1} & \cdots & \xi_{NN} \end{pmatrix}, \quad \mathbf{S} = \frac{\partial \Xi_{corr}}{\partial \tilde{\mathbf{K}}}, \quad (29)$$

where $\tilde{\mathbf{K}}$ is the $N \times N$ matrix containing natural logarithms of corresponding elements of matrix $\mathbf{K}$. Now Eq. (28) can be rewritten as

$$\xi_{0i} = \mathbf{1}_i\xi_\mathbf{m}\mathbf{1}_i^T - t\mathbf{1}_i\mathbf{\Delta_T}^{-1}(\xi_\mathbf{m} - \xi)\mathbf{S}\mathbf{1}_i^T, \quad (30)$$

where $\mathbf{1}_i$ is the $1 \times N$ vector which contains zeros excepting unity in ith position. As the left side of Eq. (30) is equal to the ith element on principal diagonal of an arbitrary $N \times N$ matrix, the full set of Eq. (30) can be readily generalised into the matrix form

$$\xi_0 = \xi_\mathbf{m} - t\mathbf{\Delta_T}^{-1}(\xi_\mathbf{m} - \xi)\mathbf{S}, \quad (31)$$

where ξ_0 is the $N \times N$ matrix, which contains the experimental observables of molecules of all types on its principal diagonal. The explicit form of matrix S is derived in Appendix C.

Eqs. (23) and (31) provide a complete solution of the problem of non-cooperative molecular hetero-assembly in N-component systems.

2.7. Utilisation of the derived equations

The matrix form of presentation of the key Eqs. (10), (23) and (31) of the N-component model of hetero-assembly developed in the present work allows their direct incorporation into matrix-oriented mathematical software, such as, for example, MATLAB. The use of such equations in the description of experimental data may follow the standard algorithm reviewed in Refs. [24,25] and currently being used by many authors in applications concerning molecular assembly. We shall briefly describe this algorithm without going into its details.

The input data are matrix ξ_0 of the experimental observables, matrix of total concentrations Δ_T, and the equilibrium parameters of self-assembly of each of N types of molecules commonly known from independent studies in one-component solution (i.e. the parameters K_{ii} and ξ_{ii} on the principal diagonal of matrices $\mathbf{K}$ and ξ, respectively). The unknown quantities are the equilibrium parameters of hetero-assembly (i.e. K_{ij} and ξ_{ij} in matrices $\mathbf{K}$ and ξ, respectively). Their search is accomplished by minimization of the discrepancy between experimental ξ_0 and theoretically calculated ξ_0 from Eq. (31). Inside the minimization cycle the solution of Eq. (23) is performed with respect to matrix Δ_C (or vector $\mathbf{c}$).

The listing of program code designed for use in MATLAB software along with brief explanation of how it works are given in Supplementary Material. The validation of Eqs. (23) and (31) and the program code have been performed using ^{1}H NMR data for three-component ($N = 3$ and $t = 2$) hetero-association of the antibiotics novatrone and daunomycin with caffeine published in Ref. [25]. The output results given in Supplementary Material were compared against the N-STOCH algorithmic model [25], which is the only approach available in the literature that provides a solution of the N-component hetero-assembly and allows comparison with the model developed in the present work. The calculated magnitudes of K_{ij} and ξ_{ij} became identical starting from the maximum lengths of complexes $L = 10$ (the key limitation of the N-STOCH approach [25]). Under such a limitation the overall time of computation using N-STOCH amounted to few hours whereas, with the model developed above, it has the order of several minutes. This result proves the apparent superiority of the N-component model developed in the present work.

It should be noted, however, that the present model is only applicable to cases of non-cooperative hetero-assembly. If cooperativity of one-dimensional hetero-assembly is involved, such models as the N-STOCH approach or the models from Refs. [8,9] are still the alternative. Nevertheless a potential possibility to introduce cooperativity of hetero-assembly into the equations of the model developed in the present work seems real and is a matter of further publication. For instance, a first step in such analysis could be the modification of the grand partition functions (10) and (20) with an aim to account for fundamental anticooperativity of one-dimensional assembly in a way recently accomplished with respect to self-association in Ref. [34].

The use of matrix equations, obtained in the present work, is not limited to the search for equilibrium parameters of complexation. Knowledge of the grand partition function (Eq. (10)) provides a link to any thermodynamical quantity of the N-component system.

2.8. Reduction of derived equations to partial cases known from the literature

The N-component model, developed in Ref. [22], does not take into account the 'reflected' complexes and therefore provides incorrect equations for the observable parameter. Excluding terms related to 'reflected' complexes in Eq. (23) and then taking trace of Δ_T one obtains $\mathbf{c}(\mathbf{I} - \mathbf{M})^{-2}\mathbf{1}^T$, which is identical to that derived in Ref. [22].

The two-component model developed in Ref. [24] can be obtained from Eqs. (23) and (31) by reduction of all matrices to 2×2 dimension and further element-wise writing down of the observable parameter at $t = 2$.

The N-component model developed in Ref. [26] provides full account of the 'reflected' complexes but does not report any link to observable parameter and does not provide a formula for the mass conservation law in close analytical form. As noted in subsection C the author of Ref. [26] used the SGF method to derive the grand partition function which leads to Ξ_{corr} identical to Eq. (20) when written in matrix form.

One-component self-association is a trivial case of molecular assembly. Eqs. (23) and (31) at $N = 1$ and $t = 2$ reduce to the well-known equations for the mass conservation law and observable parameters in Ref. [17].

A special case of molecular hetero-assembly is used when one of the components (for instance, i-type) is in such small concentrations that it is assumed to exist only as a monomer in solution as free molecules or in a hetero-complex with other j-type molecules ($i \neq j$) aggregating without any restrictions (see, for example, Refs. [35,36]). Partial derivative of Ξ_{corr} with respect to the monomer concentration C_i at $C_i = 0$ extracts from Ξ_{corr} all terms related to complexes containing i-type monomer, whereas the value of Ξ_{corr} at $C_i = 0$ gives the part of Ξ_{corr} not containing the i-type monomer. It follows that the grand partition function $\Xi_{corr}^{(1)}$ of the N-component system, in which the i-type molecules exist as monomer in either uncomplexed or complexed with other molecules state, takes the form

$$\Xi_{corr}^{(1)} = \Xi_{corr}|_{C_i=0} + C_i \frac{\partial \Xi_{corr}}{\partial C_i}\bigg|_{C_i=0}. \tag{32}$$

Substitution of Eq. (10) into Eq. (32) at $N = 2$ and $N = 3$ enables one to get $\Xi_{corr}^{(1)}$ identical to those derived in Refs. [35,36]. Detailed derivation of the grand partition function for above-mentioned case is outlined in Appendix D.

3. Conclusion

In the present work we provide a general treatment of the N-component one-dimensional molecular hetero-assembly in solution using various standard statistical-thermodynamical approaches with further reduction to all major partial cases described in literature. The main outcome is set of Eqs. (10), (23) and (31) which provide links to any thermodynamical quantities of the N-component system and to the experimental observable enabling the determination of all the equilibrium parameters of molecular complexation. The presentation of the key equations in matrix form allows their direct incorporation into matrix-oriented mathematical software, such as, for example, MATLAB, leading to increase both in speed and accuracy of calculations compared to solving a similar task by the algorithmic approach [25].

The matrix model along with the program code developed in the present work, provide a tool for quantitative investigation of a wide set of possible non-cooperative one-dimensional assemblies in N-component mixtures of interacting compounds. Several classes of practically-important interactions can be described within the framework of the matrix model, such as the aggregation of small aromatic molecules [17], multicomponent hydrogen-bonded linear supramolecular polymerizations [22], drug-hydrotrope interactions [20] etc. The next step in the development of the model would be the introduction of cooperativity of hetero-assembly, which is a matter of further publication.

Acknowledgements

Appendix A. Comparison of the grand partition functions derived using TM and SGF approaches

Matrix $\mathbf{K}$ in Eq. (21) can be decomposed as the sum of two matrices, $\mathbf{K} = \mathbf{K_h} + \mathbf{K_S}$, where

$$\mathbf{K_h} = \begin{pmatrix} 0 & K_{12} & \cdots & K_{1N} \\ K_{21} & 0 & \cdots & K_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ K_{N1} & K_{N2} & \cdots & 0 \end{pmatrix}, \quad \mathbf{K_S} = \begin{pmatrix} K_{11} & 0 & \cdots & 0 \\ 0 & K_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & K_{NN} \end{pmatrix}.$$

Introducing one more object,

$$\mathbf{K_d} = \mathbf{K}^{\mathrm{T}} \circ \mathbf{K} = \mathbf{K} \circ \mathbf{K}, \tag{A1}$$

it is now possible to rewrite the matrices used in derivation of Ξ_{corr} in the TM approach, viz.

$$\mathbf{c} = \mathbf{1}\Delta_{\mathbf{c}}, \quad \mathbf{M} = \mathbf{K}\Delta_{\mathbf{c}}, \mathbf{M_d} = (\mathbf{K}^{\mathrm{T}}\Delta_{\mathbf{c}}) \circ (\mathbf{K}\Delta_{\mathbf{c}}) = \mathbf{K_d}\Delta_{\mathbf{c}}^2,$$
$$\mathbf{M_S} = \mathbf{K_S}\Delta_{\mathbf{c}}. \tag{A2}$$

Substituting Eqs. (A2) into Eq. (10), the Ξ_{corr} in TM approach takes the form:

$$\Xi_{\mathrm{corr}} = \frac{1}{2}\mathbf{1}\left[(\Delta_{\mathbf{c}}^{-1} - \mathbf{K})^{-1} + (\Delta_{\mathbf{c}}^{-1} + \mathbf{K_S})(\Delta_{\mathbf{c}}^{-2} - \mathbf{K_d})^{-1}\right]\mathbf{1}^{\mathrm{T}}. \tag{A3}$$

In order to accomplish transformation of Ξ_{corr} in the SGF approach, it is necessary to introduce supplementary objects:

$$\Delta_{\gamma} = \begin{pmatrix} \gamma_1 & 0 & \cdots & 0 \\ 0 & \gamma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_N \end{pmatrix}, \quad \Delta_{\gamma}^{(2)} = \begin{pmatrix} \gamma_1^{(2)} & 0 & \cdots & 0 \\ 0 & \gamma_2^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_N^{(2)} \end{pmatrix},$$

which allows one to rewrite formally defined matrices as $\gamma = \mathbf{1}\Delta_{\gamma}$, $\mathbf{G} = \mathbf{K_h}\Delta_{\gamma}$, $\mathbf{G_d} = (\mathbf{K_d} - \mathbf{K_S}^2)\Delta_{\gamma}^{(2)}$.

Hence, Eq. (20) takes the form

$$\Xi_{\mathrm{corr}} = \frac{1}{2}\mathbf{1}\left[(\Delta_{\gamma}^{-1} - \mathbf{K_h})^{-1} + (\Delta_{\gamma}^{-1} - (\mathbf{K_d} - \mathbf{K_S}^2)\Delta_{\gamma}^{(2)}\Delta_{\gamma}^{-1})^{-1}\right]\mathbf{1}^{\mathrm{T}}. \tag{A4}$$

It is now easy to show that

$$\Delta_{\gamma} = \Delta_{\mathbf{c}}(\mathbf{I} - \mathbf{K_S}\Delta_{\mathbf{c}})^{-1} = (\Delta_{\mathbf{c}}^{-1} - \mathbf{K_S})^{-1} \tag{A5}$$

and

$$\Delta_{\gamma}^{(2)} = \Delta_{\mathbf{c}}(\mathbf{I} + \mathbf{K_S}\Delta_{\mathbf{c}})^{-1}\Delta_{\mathbf{c}}(\mathbf{I} - \mathbf{K_S}\Delta_{\mathbf{c}})^{-1}$$
$$= (\Delta_{\mathbf{c}}^{-1} + \mathbf{K_S})^{-1}(\Delta_{\mathbf{c}}^{-1} - \mathbf{K_S})^{-1}. \tag{A6}$$

Substituting Eqs. (A5) and (A6) into Eq. (A4), and making necessary transformations, one gets an equation identical to Eq. (A3):

$$\Xi_{\mathrm{corr}} = \frac{1}{2}\mathbf{1}\left[(\Delta_{\mathbf{c}}^{-1} - \mathbf{K_S} - \mathbf{K_h})^{-1} + \left(\Delta_{\mathbf{c}}^{-1} - \mathbf{K_S} - (\mathbf{K_d} - \mathbf{K_S}^2)(\Delta_{\mathbf{c}}^{-1} + \mathbf{K_S})^{-1}\right)^{-1}\right]\mathbf{1}^{\mathrm{T}}$$
$$= \frac{1}{2}\mathbf{1}\left[(\Delta_{\mathbf{c}}^{-1} - \mathbf{K})^{-1} + (\Delta_{\mathbf{c}}^{-1} + \mathbf{K_S})((\Delta_{\mathbf{c}}^{-1} - \mathbf{K_S})(\Delta_{\mathbf{c}}^{-1} + \mathbf{K_S}) - (\mathbf{K_d} - \mathbf{K_S}^2))^{-1}\right]\mathbf{1}^{\mathrm{T}}$$
$$= \frac{1}{2}\mathbf{1}\left[(\Delta_{\mathbf{c}}^{-1} - \mathbf{K})^{-1} + (\Delta_{\mathbf{c}}^{-1} + \mathbf{K_S})(\Delta_{\mathbf{c}}^{-2} - \mathbf{K_d})^{-1}\right]\mathbf{1}^{\mathrm{T}}. \tag{A7}$$

Appendix B. Derivation of matrix $\Delta_{\mathbf{T}}$

In general the grand partition function (10) can be represented as

$$\Xi_{\mathrm{corr}} = \frac{1}{2}\mathbf{1}\mathbf{F}\mathbf{1}^{\mathrm{T}}, \tag{B1}$$

where $\mathbf{F} = \mathbf{F}(\Delta_{\mathbf{c}}, \mathbf{K}, \mathbf{K_S}, \mathbf{K_d})$ is a matrix function of independent matrix arguments, $\Delta_{\mathbf{c}}, \mathbf{K}, \mathbf{K_S}, \mathbf{K_d}$. Within this work one has to find the closed form of partial derivatives of the grand partition function (B1) with respect to $\tilde{\Delta}_{\mathbf{c}}$ and $\tilde{\mathbf{K}}$, which are $N \times N$ matrices containing natural logarithms of corresponding elements of $\Delta_{\mathbf{c}}$ and $\mathbf{K}$, respectively. For instance, the partial derivative of Ξ_{corr} with respect to $\tilde{\Delta}_{\mathbf{c}}$ can be written as follows:

$$\frac{\partial \Xi_{\mathrm{corr}}}{\partial \tilde{\Delta}_{\mathbf{c}}} = \frac{1}{2}\left(\frac{\partial \mathbf{1}}{\partial \tilde{\Delta}_{\mathbf{c}}}\mathbf{F}\mathbf{1}^{\mathrm{T}} + \mathbf{1}\frac{\partial \mathbf{F}}{\partial \tilde{\Delta}_{\mathbf{c}}}\mathbf{1}^{\mathrm{T}} + \mathbf{1}\mathbf{F}\frac{\partial \mathbf{1}^{\mathrm{T}}}{\partial \tilde{\Delta}_{\mathbf{c}}}\right).$$

Derivatives $\frac{\partial \mathbf{1}}{\partial \tilde{\Delta}_{\mathbf{c}}}$ and $\frac{\partial \mathbf{1}^{\mathrm{T}}}{\partial \tilde{\Delta}_{\mathbf{c}}}$ when multiplied by vectors $\mathbf{F}\mathbf{1}^{\mathrm{T}}$ and $\mathbf{1}\mathbf{F}$, respectively, yield $N \times N$ zero matrices. Hence,

$$\frac{\partial \Xi_{\mathrm{corr}}}{\partial \tilde{\Delta}_{\mathbf{c}}} = \frac{1}{2}\mathbf{1}\frac{\partial \mathbf{F}}{\partial \tilde{\Delta}_{\mathbf{c}}}\mathbf{1}^{\mathrm{T}}. \tag{B2}$$

Let us expand Eq. (B2) according to Eqs. (22) and (A2):

$$\Delta_{\mathbf{T}} = \frac{1}{2}\mathbf{1}\begin{bmatrix} (\mathbf{I} - \Delta_{\mathbf{c}}\mathbf{K})^{-1}\frac{\partial \Delta_{\mathbf{c}}}{\partial \tilde{\Delta}_{\mathbf{c}}}(\mathbf{I} - \mathbf{K}\Delta_{\mathbf{c}})^{-1} - \Delta_{\mathbf{c}}^{-1}\frac{\partial \Delta_{\mathbf{c}}}{\partial \tilde{\Delta}_{\mathbf{c}}}\Delta_{\mathbf{c}}(\mathbf{I} - \mathbf{K_d}\Delta_{\mathbf{c}}^2)^{-1} \\ + (\mathbf{I} + \mathbf{K_S}\Delta_{\mathbf{c}})(\mathbf{I} - \mathbf{K_d}\Delta_{\mathbf{c}}^2)^{-1}\frac{\partial \Delta_{\mathbf{c}}}{\partial \tilde{\Delta}_{\mathbf{c}}}(\mathbf{I} - \mathbf{K_d}\Delta_{\mathbf{c}}^2)^{-1} \\ \Delta_{\mathbf{c}}^{-1}(\mathbf{I} + \mathbf{K_S}\Delta_{\mathbf{c}})(\mathbf{I} - \Delta_{\mathbf{c}}^2\mathbf{K_d})^{-1}\frac{\partial \Delta_{\mathbf{c}}}{\partial \tilde{\Delta}_{\mathbf{c}}}\Delta_{\mathbf{c}}(\mathbf{I} - \mathbf{K_d}\Delta_{\mathbf{c}}^2)^{-1} \end{bmatrix}\mathbf{1}^{\mathrm{T}}. \tag{B3}$$

Actually Eq. (B3) is a sum of matrix products $\mathbf{a}\frac{\partial \Delta_{\mathbf{c}}}{\partial \tilde{\Delta}_{\mathbf{c}}}\mathbf{b}^{\mathrm{T}}$, where $\mathbf{a}, \mathbf{b}$ are certain $1 \times N$ vectors. Consider this term in more detail:

$$\mathbf{a}\frac{\partial \Delta_{\mathbf{c}}}{\partial \tilde{\Delta}_{\mathbf{c}}}\mathbf{b}^{\mathrm{T}} = \sum_{i=1}^{N}\sum_{j=1}^{N}a_i\frac{\partial \|\Delta_{\mathbf{c}}\|_{ij}}{\partial \tilde{\Delta}_{\mathbf{c}}}b_j = \sum_{i=1}^{N}a_i\mathbf{1}_i^{\mathrm{T}}\mathbf{1}_i b_i\|\Delta_{\mathbf{c}}\|_{ii}$$
$$= (\mathbf{a}^{\mathrm{T}}\mathbf{b}) \circ \Delta_{\mathbf{c}}. \tag{B4}$$

Thus Eq. (B4) enables one to rewrite Eq. (B3) in the following form:

$$\Delta_{\mathbf{T}} = \frac{1}{2}\begin{bmatrix} (\mathbf{I} - \mathbf{K}\Delta_{\mathbf{c}})^{-1}\mathbf{1}^{\mathrm{T}}\mathbf{1}(\mathbf{I} - \Delta_{\mathbf{c}}\mathbf{K})^{-1} \\ + (\mathbf{I} - \mathbf{K_d}\Delta_{\mathbf{c}}^2)^{-1}\mathbf{1}^{\mathrm{T}}\mathbf{1}\Delta_{\mathbf{c}}(\mathbf{I} + \mathbf{K_S}\Delta_{\mathbf{c}})(\mathbf{I} - \mathbf{K_d}\Delta_{\mathbf{c}}^2)^{-1}\Delta_{\mathbf{c}}^{-1} \\ + \Delta_{\mathbf{c}}(\mathbf{I} - \mathbf{K_d}\Delta_{\mathbf{c}}^2)^{-1}\mathbf{1}^{\mathrm{T}}\mathbf{1}\Delta_{\mathbf{c}}(\mathbf{K_S}\Delta_{\mathbf{c}} + \mathbf{K_d}\Delta_{\mathbf{c}}^2)(\mathbf{I} - \mathbf{K_d}\Delta_{\mathbf{c}}^2)^{-1}\Delta_{\mathbf{c}}^{-2} \end{bmatrix} \circ \Delta_{\mathbf{c}}. \tag{B5}$$

Here one took into account the invariance of $\Delta_{\mathbf{T}}$ under transposition. As physically meaning elements of $\Delta_{\mathbf{T}}$ are only elements on the principal diagonal, Eq. (B5) takes the final form

$$\Delta_{\mathbf{T}} = \frac{1}{2}\left[(\mathbf{I} - \mathbf{M})^{-1}\mathbf{1}^{\mathrm{T}}\mathbf{1}\Delta_{\mathbf{c}}(\mathbf{I} - \mathbf{M})^{-1} + (\mathbf{I} - \mathbf{M_d})^{-1}\mathbf{1}^{\mathrm{T}}\mathbf{1}\Delta_{\mathbf{c}}(\mathbf{I} + 2\mathbf{M_S} + \mathbf{M_d})(\mathbf{I} - \mathbf{M_d})^{-1}\right]. \tag{B6}$$

Appendix C. Derivation of matrix S

Derivation of matrix $\mathbf{S}$ is similar to derivation of matrix $\Delta_{\mathbf{T}}$ described in Appendix B as it consists of partial derivatives

$$\frac{\partial \Xi_{\mathrm{corr}}}{\partial \tilde{\mathbf{K}}} = \frac{1}{2}\mathbf{1}\frac{\partial \mathbf{F}}{\partial \tilde{\mathbf{K}}}\mathbf{1}^{\mathrm{T}}$$

in terms of Appendix B. Expand Eq. (C1) accounting for Eq. (29) as it was accomplished above:

$$\mathbf{S} = \frac{1}{2}\mathbf{1}\begin{bmatrix} (\Delta_{\mathbf{c}}^{-1} - \mathbf{K})^{-1}\frac{\partial \mathbf{K}}{\partial \tilde{\mathbf{K}}}(\Delta_{\mathbf{c}}^{-1} - \mathbf{K})^{-1} + \frac{\partial \mathbf{K_S}}{\partial \tilde{\mathbf{K}}}(\Delta_{\mathbf{c}}^{-2} - \mathbf{K_d})^{-1} \\ + (\Delta_{\mathbf{c}}^{-1} + \mathbf{K_S})(\Delta_{\mathbf{c}}^{-2} - \mathbf{K_d})^{-1}\frac{\partial \mathbf{K_d}}{\partial \tilde{\mathbf{K}}}(\Delta_{\mathbf{c}}^{-2} - \mathbf{K_d})^{-1} \end{bmatrix}\mathbf{1}^{\mathrm{T}}. \tag{C1}$$

It is now necessary to define partial derivatives in Eq. (C1) taking Eq. (B4) as an example:

$$\begin{cases} \mathbf{a}\frac{\partial \mathbf{K}}{\partial \mathbf{K}}\mathbf{b}^{\mathrm{T}} = \sum_{i=1}^{N}\sum_{j=1}^{N}a_i\frac{\partial K_{ij}}{\partial \mathbf{K}}b_j = \sum_{i=1}^{N}\sum_{j=1}^{N}a_i\mathbf{1}_i^{\mathrm{T}}\mathbf{1}_j b_j K_{ij} = (\mathbf{a}^{\mathrm{T}}\mathbf{b})\circ\mathbf{K} \\[2mm] \mathbf{a}\frac{\partial \mathbf{K_S}}{\partial \mathbf{K}}\mathbf{b}^{\mathrm{T}} = \sum_{i=1}^{N}\sum_{j=1}^{N}a_i\frac{\partial K_{ij}}{\partial \mathbf{K}}b_j = \sum_{i=1}^{N}a_i\mathbf{1}_i^{\mathrm{T}}\mathbf{1}_i b_i K_{ii} = (\mathbf{a}^{\mathrm{T}}\mathbf{b})\circ\mathbf{K_S} \\[2mm] \mathbf{a}\frac{\partial \mathbf{K_d}}{\partial \mathbf{K}}\mathbf{b}^{\mathrm{T}} = \sum_{i=1}^{N}\sum_{j=1}^{N}a_i\frac{\partial (K_{ij}K_{ji})}{\partial \mathbf{K}}b_j = \sum_{i=1}^{N}\sum_{j=1}^{N}a_i(\mathbf{1}_i^{\mathrm{T}}\mathbf{1}_j + \mathbf{1}_j^{\mathrm{T}}\mathbf{1}_i)b_j K_{ij}K_{ji} = (\mathbf{a}^{\mathrm{T}}\mathbf{b} + \mathbf{b}^{\mathrm{T}}\mathbf{a})\circ\mathbf{K_d} \end{cases}$$
$$\text{(C2)}$$

It should be noted that all differentiations in Eq. (C2) were made without loss of generality, i.e. assuming all the matrices to be non-symmetrical. This takes place due to introduction of matrix $\mathbf{S}$ in Eq. (27) in such a way. Particularly, the differentiation of matrix $\mathbf{K_d}$ was made according to Eq. (A1). Application of Eq. (C2) to Eq. (C1) yields

$$\mathbf{S} = \frac{1}{2}\begin{bmatrix} \left((\Delta_{\mathbf{C}}^{-1} - \mathbf{K})^{-1}\mathbf{1}^{\mathrm{T}}\mathbf{1}(\Delta_{\mathbf{C}}^{-1} - \mathbf{K})^{-1}\right)\circ\mathbf{K} + \left(\mathbf{1}^{\mathrm{T}}\mathbf{1}(\Delta_{\mathbf{C}}^{-2} - \mathbf{K_d})^{-1}\right)\circ\mathbf{K_S} \\ + \left((\Delta_{\mathbf{C}}^{-2} - \mathbf{K_d})^{-1}(\Delta_{\mathbf{C}}^{-1} + \mathbf{K_S})\mathbf{1}^{\mathrm{T}}\mathbf{1}(\Delta_{\mathbf{C}}^{-2} - \mathbf{K_d})^{-1}\right)\circ\mathbf{K_d} \\ + \left((\Delta_{\mathbf{C}}^{-2} - \mathbf{K_d})^{-1}\mathbf{1}^{\mathrm{T}}\mathbf{1}(\Delta_{\mathbf{C}}^{-1} + \mathbf{K_S})(\Delta_{\mathbf{C}}^{-2} - \mathbf{K_d})^{-1}\right)\circ\mathbf{K_d} \end{bmatrix}.$$

After some simplifications one obtains the final analytical form of matrix $\mathbf{S}$:

$$\mathbf{S} = \frac{1}{2}\begin{bmatrix} \left((\mathbf{I} - \mathbf{M})^{-1}\mathbf{1}^{\mathrm{T}}\mathbf{1}\Delta_{\mathbf{C}}(\mathbf{I} - \mathbf{M})^{-1}\right)\circ\mathbf{M}^{\mathrm{T}} + \left(\Delta_{\mathbf{C}}(\mathbf{I} - \mathbf{M_d})^{-1}\mathbf{1}^{\mathrm{T}}\mathbf{1}\right)\circ\mathbf{M_S} \\ + \left((\mathbf{I} - \mathbf{M_d^T})^{-1}(\mathbf{I} + \mathbf{M_S})\Delta_{\mathbf{C}}\mathbf{1}^{\mathrm{T}}\mathbf{1}(\mathbf{I} - \mathbf{M_d^T})^{-1}\right)\circ\mathbf{M_d} \\ + \left((\mathbf{I} - \mathbf{M_d})^{-1}\mathbf{1}^{\mathrm{T}}\mathbf{1}\Delta_{\mathbf{C}}(\mathbf{I} + \mathbf{M_S})(\mathbf{I} - \mathbf{M_d})^{-1}\right)\circ\mathbf{M_d^T} \end{bmatrix}.$$
$$\text{(C3)}$$

Appendix D. Reduction of the grand partition function for a special case of molecular hetero-assembly

The grand partition function (10) is actually a sum of products, each of which contains certain combination of monomer concentrations, that are the markers of molecules, and equilibrium constants, that are the markers of stacks between molecules. Removing from the grand partition function (10) products which contain more than m factors of given type, i.e. monomer concentration of, say, i-type molecules, C_i, one obtains the partition function of the system in which i-type molecules are presented in complexes m times only. Let us rewrite grand partition function (10) in the form of a power series in variable C_i:

$$\Xi_{\mathrm{corr}} = a_0 + a_1 C_i + a_2 C_i^2 + a_3 C_i^3 + \cdots = \sum_{k=0}^{\infty}a_k C_i^k, \tag{D1}$$

where a_k is a coefficient, which contains the sum of products of all equilibrium constants and monomer concentrations excepting C_i. From the definition of a_k one has

$$a_0 = \Xi_{\mathrm{corr}}\big|_{C_i=0}, \quad a_1 = \frac{\partial \Xi_{\mathrm{corr}}}{\partial C_i}\bigg|_{C_i=0}, \quad a_2 = \frac{1}{2}\frac{\partial^2 \Xi_{\mathrm{corr}}}{\partial C_i^2}\bigg|_{C_i=0}, \text{etc}$$

or in the general form

$$a_k = \frac{1}{k!}\frac{\partial^k \Xi_{\mathrm{corr}}}{\partial C_i^k}\bigg|_{C_i=0}. \tag{D2}$$

Substituting Eq. (D2) into Eq. (D1) one obtains

$$\Xi_{\mathrm{corr}} = \sum_{k=0}^{\infty}\frac{C_i^k}{k!}\frac{\partial^k \Xi_{\mathrm{corr}}}{\partial C_i^k}\bigg|_{C_i=0},$$

which is an explicit expansion of Ξ_{corr} into a Maclaurin series.

The partition function of interest is

$$\Xi_{\mathrm{corr}}^{(m)} = a_0 + a_1 C_i + \cdots + a_m C_i^m = \sum_{k=0}^{m}\frac{C_i^k}{k!}\frac{\partial^k \Xi_{\mathrm{corr}}}{\partial C_i^k}\bigg|_{C_i=0}. \tag{D3}$$

Eq. (D3) at $m = 1$ readily yields Eq. (32).

Appendix E. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.chemphys.2013.06.006.

References

[1] F.J.M. Hoeben, P. Jonkheijm, E.W. Meijer, A.P.H.J. Schenning, Chem. Rev. 105 (2005) 1491.
[2] H. Sun, K. Ye, C. Wang, H. Qi, F. Li, Y. Wang, J. Phys. Chem. A 110 (2006) 10750.
[3] S.S. Babu, S. Prasanthkumar, A. Ajayaghosh, Angew. Chem. Int. Ed. 51 (2012) 1766.
[4] G.W.M. Heckl, in: H. Figger, D. Meschede, C. Zimmermann (Eds.), Laser Physics at the Limits, Springer, Berlin, 2002.
[5] I. Last, J. Jortner, Phys. Rev. A 73 (2006) 063201.
[6] T.F.A. de Greef, M.M.J. Smulders, M. Wolffs, A.P.H.J. Schenning, R.P. Sijbesma, E.W. Meijer, Chem. Rev. 109 (2009) 5687.
[7] C. Wang, Z. Wang, X. Zhang, Acc. Chem. Res. 45 (2012) 608.
[8] A.J. Markvoort, H.M.M. ten Eikelder, P.A.J. Hilbers, T.F.A. de Greef, E.W. Meijer, Nat. Commun. 2 (2011) 509.
[9] H.M.M. ten Eikelder, A.J. Markvoort, T.F.A. de Greef, P.A.J. Hilbers, J. Phys. Chem. B 116 (2012) 5291.
[10] A. Katsuhiko, P.H. Jonathan, V.L. Michael, V. Ajayan, C. Richard, A. Somobrata, Sci. Technol. Adv. Mater. 9 (2008) 014109.
[11] L. Mao, I. Shinsuke, J. Qingmin, A. Misaho, P.H. Jonathan, A. Katsuhiko, Sci. Technol. Adv. Mater. 13 (2012) 053001.
[12] M.P. Evstigneev, A.S. Buchelnikov, D.P. Voronin, Yu.V. Rubin, L.F. Belous, Yu.I. Prylutskyy, U. Ritter, Chem. Phys. Chem. 14 (2013) 568.
[13] Z. Liu, A.C. Fan, K. Rakhra, S. Sherlock, A. Goodwin, X. Chen, Q. Yang, D.W. Felsher, H. Dai, Angew. Chem. Int. Ed. 48 (2009) 7668.
[14] H.-J. Schneider, Angew. Chem. Int. Ed. 48 (2009) 3924.
[15] A. Lakshmanan, S. Zhang, C.A.E. Hauser, Trends Biotechnol. 30 (2012) 155.
[16] V.R. Horowitz, L.A. Janowitz, A.L. Modic, P.A. Heiney, P.J. Collings, Phys. Rev. E 72 (2005) 041710.
[17] R.B. Martin, Chem. Rev. 96 (1996) 3043.
[18] A.N. Veselkov, M.P. Evstigneev, D.A. Veselkov, D.B. Davies, J. Chem. Phys. 115 (2001) 2252.
[19] L. Ribeiro, R.A. Carvalho, D.C. Ferreira, F.J.B. Veiga, Eur. J. Pharmaceut. Sci. 24 (2005) 1.
[20] M.P. Evstigneev, V.P. Evstigneev, A.A. Hernandez Santiago, D.B. Davies, Eur. J. Pharmaceut. Sci. 28 (2006) 59.
[21] Y. Kong, J. Phys. Chem. B 105 (2001) 10111.
[22] F.G.J. Odille, S. Jónsson, S. Stjernqvist, T. Rydén, K. Wärnmark, Chem. Eur. J. 13 (2007) 9617.
[23] M.P. Evstigneev, V.P. Evstigneev, D.B. Davies, Chem. Phys. Lett. 432 (2006) 248.
[24] V.P. Evstigneev, A.A. Mosunov, A.S. Buchelnikov, A.A. Hernandez Santiago, M.P. Evstigneev, J. Chem. Phys. 134 (2011) 194902.
[25] M.P. Evstigneev, V.P. Evstigneev, D.B. Davies, J. Chem. Phys. 127 (2007) 154511.
[26] Y. Kong, J. Chem. Phys. 135 (2011) 156101.
[27] V.B. Teif, K. Rippe, Brief. Bioinform. 13 (2012) 187.
[28] D. Poland, J. Chem. Phys. 58 (1973) 3574.
[29] M.P. Evstigneev, D.A. Beshnova, A.O. Rozvadovskaya, Russ. J. Phys. Chem. B 26 (2007) 60.
[30] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge Press, New York, 1991.
[31] S. Lifson, J. Chem. Phys. 40 (1964) 3705.
[32] Y. Kong, J. Am. Stat. Assoc. 101 (2006) 1253.
[33] K. Weller, H. Shütz, I. Petri, Biophys. Chem. 19 (1984) 289.
[34] M.P. Evstigneev, A.S. Buchelnikov, V.P. Evstigneev, Phys. Rev. E 85 (2012) 061405.
[35] A.S. Buchelnikov, A.A. Hernandez Santiago, M. Gonzalez Flores, R. Vazquez Ramirez, D.B. Davies, M.P. Evstigneev, Eur. Biophys. J. 41 (2012) 273.
[36] A.S. Buchelnikov, A.F. Khrustalev, M.P. Evstigneev, Appl. Spectrosc. 67 (2013) 29.