ENTROPIC BINDING MODE PREFERENCE IN COOPERATIVE HOMO-DIMERIC DRUG-DNA RECOGNITION

Rubinson M.A.[a], Parkinson J.A.[b], Evstigneev M.P. [a,c*]

[a] Department of Physics, Sevastopol State University, Universitetskaya str.33, Sevastopol 299053, Russia

[b] WestCHEM Department of Pure and Applied Chemistry, University of Strathclyde, 295 Cathedral Street, Glasgow G1 1XL, United Kingdom

[c] Department of Biology and Chemistry, Belgorod State University, Belgorod 308015, Russia

*Corresponding author:

Prof. M. P. Evstigneev, Department of Biology and Chemistry,
Belgorod State University
308015 Belgorod, Russia
Tel. +79787881869; email: max_evstigneev@mail.ru

Abstract

The present work reveals the entropic preference in the two-step binding process of small molecule DNA minor groove binders (MGBs), involving the formation of a dimer in free solution followed by the binding of that dimer to DNA, which contrasts with the sequential or simultaneous cooperative binding of two unbound MGB molecules with DNA that is implicitly assumed in the majority of reported DNA-MGB binding studies. The results are generalized for the *n*-mer binding case and the methodological outcomes are discussed.

Key words: entropy of binding, dimerization, homo-dimeric binding, DNA minor-groove binders

Introduction

Small molecule recognition of nucleic acids has been the subject of extensive research since the elucidation of the DNA double-helix structure in 1953 by Watson and Crick. It is currently recognized as a key mechanism responsible for medico-biological properties of some drugs, in particular, those exerting antitumor properties [1,2]. One of the main difficulties in drug-based chemotherapy is the toxicity associated with nucleic acid targeted drug intervention. These toxic side effects associated with DNA binding drugs, at least in part, are directly linked to their relatively weak DNA sequence targeting specificity. The essential strategy in DNA-targeted drug design has been to find or create drugs with high specificity and cooperativity of binding to particular DNA sequences [2,3]. One of the rare examples when this strategy has been known to succeed was in the discovery of the lexitropsin class of DNA minor groove binders (to be referred to here as MGB ligands). These exert a characteristic homo-dimeric type of complexation by which two drug molecules simultaneously occupy a particular binding site [4,5]. The discovery of this DNA recognition mode opened up an important page in the history of DNA-targeted drug design and to date remains an active field of research (e.g. see [6] for review).

So far the strategy and the methods used for investigating homo-dimeric binding have been completely focused on structural and/or thermodynamic investigations of 2:1 drug-DNA complexes in the crystalline or solution phases whilst ignoring the possible route by which homo-dimer formation occurs. However, a recent question was raised concerning whether the formation of drug homo-dimer occurs at the heart of the DNA binding site through sequential or simultaneous binding of ligands (i.e. *via* a DNA-dependent process), or whether it occurs in free solution due to drug dimerization as a prelude to DNA complexation during a pre-assembly stage (i.e. *via* a DNA-independent process) (**Fig. 1**) [7]. If the DNA-independent process is operating, the method of investigation concerning homo-dimeric binding should be complemented by investigation of ligand self-assembly in the absence of DNA as a pre-requisite to further ligand-DNA binding studies. This is an important methodological consequence not so far considered in this field of research. In the present work we provide the statistical-thermodynamic rationale for addressing this question.

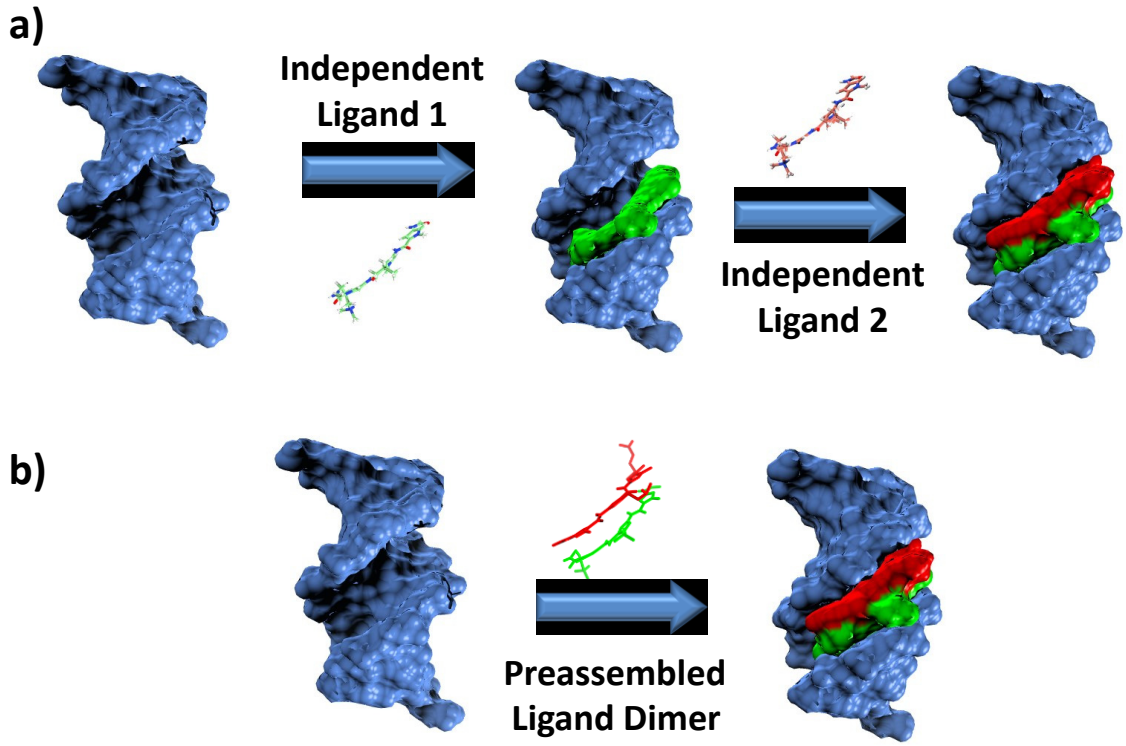

Figure 1. Schematic representation of DNA-dependent (a) and DNA-independent (b) routes of the MGB ligands binding with DNA. For (a) Ligand 1 and Ligand 2 may also bind at the same time but this does not involve association between the ligands prior to DNA binding.

Results and Discussion

Evidence supportive of binding mode preference for homo-dimeric binding with DNA.

Data available in the literature on the dimerization of MGB ligands are quite scarce and limited to the self-association of some typical DNA binders (e.g. [7-10]). Nevertheless some preliminary conclusions may be drawn, which indirectly support some preference for a DNA-independent mechanism of homo-dimeric binding to DNA:

(i) MGB ligands that exert homo-dimeric binding typically also exert quite strong self-association in solution in the absence of DNA (e.g. [7-10]);

(ii) the structure of known MGB dimers in free solution appears to be very similar to that found in ligand/DNA complexes for 2:1 binding. In particular, the anti-parallel orientation and maintenance of stacking between aromatic moieties of the molecules

comprising MGB dimers and the dimers bound to DNA seem to be preserved (e.g. [7,10]);

(iii) the homo-dimeric 2:1 binding of MGB ligands with DNA typically features high cooperativity with very minor contributions from the intermediate 1:1 stage [11,12], suggesting that the binding process is not sequential.

Under the assumption that the conclusions listed above are valid for the homo-dimeric binding discussed in this paper, it follows that homo-dimeric binding occurs through either of two mechanisms. This can be on DNA through simultaneous binding of two MGB ligands with DNA (i.e. a DNA-dependent process in which DNA acts as the former or template onto which the ligands assemble), or it can go *via* a two-step process whereby dimerization in free solution is then followed by the preformed dimer binding with the DNA (i.e. a DNA-independent process). To the best of our knowledge no experimental or theoretical evidence has been put forward to date which enables these two binding routes to be distinguished from one another. Valuable information on binding route can be gained from kinetic studies, *e.g.* those pioneered by D.Crothers with respect to homo-dimeric binding of distamycin passing through sequential (DNA-dependent) 1:1 to 2:1 route [13,14]. However, in the case of the minor contribution from 1:1 binding discussed in the present work, it would be very difficult, if not impossible, to distinguish in kinetic terms dimer formation in free solution followed by binding with DNA from dimer formation at the DNA followed by dimer adaptation to DNA.

In summary, no apparent structural or energetic preference for one or other route can be postulated *a priori*. In the context of this work it is assumed that these two routes are enthalpically indistinguishable. If so, it is reasonable to further explore the entropic change associated with the DNA-dependent and DNA-independent homo-dimeric binding processes.

Entropic cost of selecting a particular route of binding.

Let us generalize the task and consider the case of an aggregate binding with DNA and containing n molecules (the n-mer binding model), for which the homo-dimeric binding (n=2) constitutes a partial case. The ligand self-association required to form n-mers is commonly described within the framework of an isodesmic aggregation model, which assumes that the self-association occurs by sequential addition of monomers and the corresponding equilibrium self-association constant, K_X, is identical for each binding step [15,16].

For the DNA-independent mechanism the two binding steps, namely self-assembly into the *n*-mer state followed by binding of the *n*-mer to DNA, are statistically independent events and can be described as

$$\left(C^1_{x_1}\cdot\ldots\right)\cdot K_N K_X^{n-1} N_1 \qquad (1)$$

where K_N is the equilibrium binding constant of the *n*-mer with DNA; x_1 and N_1 are the numbers (or concentrations) of free (monomeric) ligand molecules and DNA binding sites, respectively and $C^1_{x_1}=\frac{x_1!}{1!(x_1-1)!}$ is the number of ways to select one molecule out of x_1 monomers. Taking into account that $x_1 >> n$, the product of the $C^1_{x_1}$ factors in Eq. 1 can be further simplified as $C^1_{x_1}\cdot\ldots \approx x_1^n$, the common law of mass action typical of the independent binding events is yielded thus:

$$K_N K_X^{n-1} x_1^n N_1 \qquad (2)$$

For the DNA-dependent mechanism, no distinct binding steps exist, because the binding with DNA and the formation of the *n*-mer occur simultaneously. Hence, the self-association and the DNA binding are statistically dependent events and the use of the law of mass action in the standard form of Eqs.1 and 2 is not valid. Recalling that the DNA-dependent and the DNA-independent processes are enthalpically equivalent (i.e. the magnitude of the product $K_X \cdot K_N$ is similar for these two routes) the following binding term may be written

$$C^n_{x_1} K_N K_X^{n-1} N_1 \qquad (3)$$

where $C^n_{x_1}=\frac{x_1!}{n!(x_1-n)!}$ describes the overall number of ways to make up an *n*-mer from x_1 monomers.

Further simplification of Eq. 3 employing the condition $x_1 \times n$ results in

$$\frac{1}{n!}\cdot K_N K_X^{n-1} x_1^n N_1 \qquad (4)$$

The principal difference between Eq. 4 (DNA-dependent mechanism) and Eq. 2 (DNA-independent mechanism) is the quantity $1/n!$ originating from different probabilities of

independent formation of the *n*-mer in free solution as compared with the formation of the *n*-mer on the DNA template. This $1/n!$ quantity lowers the magnitude of the DNA binding constant and gives unfavorable entropic contribution to the net Gibbs free energy change in the DNA-dependent route of binding, equivalent to

$$\Delta S_n = -R \ln n! \tag{5}$$

As a consequence, it results in overall thermodynamic preference of the two-step binding process involving the *n*-mer pre-assembly stage. Note that this result is general. It does not introduce any assumptions regarding the nature of the ligands, their number in the binding domain and the type of the binding site. It is also only limited by the three empirical conditions (i)-(iii) listed above, which are typically valid for homo-dimeric binding of MGB ligands. If the same set of conditions also operates for other types of ligands assembling on the DNA template, so the conclusions drawn regarding the preferred route of binding may also be transferred to that type of ligand as well.

For the DNA-dependent homo-dimeric binding (n=2) of MGB ligands, the entropic contribution in Eq. 5 is the smallest: lowering of the K_N value occurs just by the factor 2 as compared with the DNA-independent route. Hence, the probabilistic nature of the preference for a DNA-independent mechanism of homo-dimeric binding does not completely rule out the contribution from the DNA-dependent mechanism, i.e. in reality the mixture of both routes may operate on MGB binding with DNA. Nevertheless, the results of the present work for the first time reveal an entropic preference for the two-step process of DNA binding involving pre-assembly (dimerization) of MGB ligands in free solution.

One may wonder on whether the factor of 2 ($RTln2 \approx 1.7$ kJ/mol) may be important in terms of the net Gibbs free energy change on DNA binding, ΔG_{bind}? The value of ΔG_{bind} for monomeric DNA binding averaged over the set of typical MGB ligands is relatively high and equals to -(9.1±1.4) kJ/mol (calculated from the data given in [17]). However, the resultant standard deviation, 1.4 kJ/mol, is relatively small and commensurate with the 1.7 kJ/mol equivalent of the factor of 2 in binding affinity. Hence, relatively small differences in binding affinities may be more important in terms of comparative analysis of MGB ligand binding with DNA than the absolute value of ΔG_{bind}.

Methodological consequences of the entropic preference of a DNA-independent binding mechanism.

Two principal methodological outcomes could be formulated regarding the revealed preference for a two-step homo-dimeric binding process.

Firstly, as already stated in the introductory section, the approach for investigating homo-dimeric binding should not be limited to structural and/or thermodynamic exploration of the 2:1 binding process alone. The dimerization stage must be explored in detail in the absence of DNA in terms of structure and thermodynamics, which may throw light on the specificity of the formation of the given 2:1 complex in comparative analysis with other MGB ligands.

Secondly, the binding model should contain in explicit form the dimerization of ligand with the self-association constant K_X. If this is not introduced (as is the case in the majority of reported studies of homo-dimeric ligand binding to DNA), the magnitude of the experimentally measured apparent DNA binding constant, K'_N ($K'_N = K_N \cdot K_X$), will be strongly overestimated by the value of K_X (which may take $(1..10)\cdot 10^3$ M^{-1} [7,9]), which may lead to erroneous conclusions when comparing binding affinities of various MGB ligands to DNA.

In conclusion it may be stated that the route towards cooperative homo- or *n*-mer drug binding with DNA appears to be an important issue which must be addressed prior to any quantitative thermodynamic analysis of such processes. As a consequence of this work, the self-association properties of MGB ligands must be investigated in detail as a means of further deepening understanding of the highly cooperative homo-dimeric binding of clinically-important drugs with DNA. It also creates a challenge in searching for ways to experimentally distinguish between different routes of MGB binding with DNA.

Acknowledgements

This work was supported by Russian Fund for Basic Researches (project no.15-04-03119)

References

1. Ihmels, H.; Otto, D. Top Curr Chem 2005, 258, 161-204.
2. Wilson, W. D.; Tanious, F. A.; Mathis, A.; Tevis, D.; Hall, J. E.; Boykin, D. W. Biochimie 2008, 90, 999-1014.
3. Nelson, S. M.; Ferguson, L. R.; Denny, W. A. Mutat Res 2007, 623, 24-40.
4. Dervan, P. B. Bioorg Med Chem 2001, 9, 2215-2235.
5. Wemmer, D. E. Biopolymers 2001, 52(4), 197-211.
6. Suckling, C. J. J Phys Org Chem 2008, 21(7-8), 575-583.
7. Salvia M.-V.; Addison, F.; Alniss, H. Y.; Buurma, N. J.; Khalaf, A. I.; Mackay, S. P.; Anthony, N. G.; Suckling, C. J.; Evstigneev, M. P.; Hernandez-Santiago, A. A.; Waigh, R. D.; Parkinson, J. A. Biophys Chem 2013, 179, 1-11.
8. Armitage, B. A. Top Curr Chem 2005, 253, 55-76.
9. Buurma, N. J.; Haq, I. J. J Mol Biol 2008, 381, 607-621.
10. Parkinson, J. A.; Scott, F. J.; Suckling, C. J.; Wilson, G. Med Chem Comm 2013, 4, 1105-1108.
11. Anthony, N. G.; Johnston, B. F.; Khalaf, A. I.; Mackay, S. P.; Parkinson, J. A.; Suckling, C. J.; Waigh, R. D. J Am Chem Soc 2004, 126, 11338-11349.
12. Alniss, H. Y.; Anthony, N. G.; Khalaf, A. I.; Mackay, S. P.; Suckling, C. J.; Waigh, R. D.; Wheate, N. J.; Parkinson, J. A. Chem Sci 2012, 3, 711-722.
13. Baliga, R.; Crothers, D. M. Proc Natl Acad Sci USA 2000, 97, 7814-7818.
14. Baliga, R.; Baird, E. E.; Herman, D. M.; Melander, C.; Dervan, P. B.; Crothers, D. M. Biochemistry 2001, 40, 3-8.
15. Martin, R. B. Chem Rev 1996, 96, 3043-3064.
16. Evstigneev, M. P.; Buchelnikov, A. S.; Evstigneev, V. P. Phys Rev E 2012, 85, 061405.
17. Kostjukov, V. V.; Hernandez Santiago, A. A.; Rojas Rodriguez, F.; Rosas Castilla, S.; Parkinson, J. A.; Evstigneev, M. P. Phys Chem Chem Phys 2012, 14, 5588-5600.