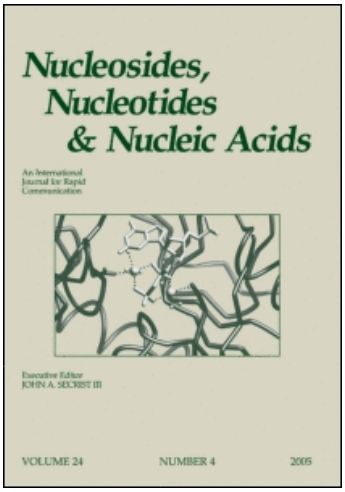

Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

Free Energy Calculations Predict Sequence Specificity in DNA-Drug Complexes

T. Härd[a]; L. Nilsson[a]
[a] Department of Medical Biophysics, Karolinska Institutet, Stockholm, Sweden

To cite this Article Härd, T. and Nilsson, L.(1992) 'Free Energy Calculations Predict Sequence Specificity in DNA-Drug Complexes', Nucleosides, Nucleotides and Nucleic Acids, 11: 1, 167 — 173
To link to this Article: DOI: 10.1080/07328319208021159
URL: http://dx.doi.org/10.1080/07328319208021159

NUCLEOSIDES & NUCLEOTIDES, 11(1), 167-173 (1992)

FREE ENERGY CALCULATIONS PREDICT SEQUENCE SPECIFICITY IN DNA-DRUG COMPLEXES

T. Härd and **L. Nilsson**, Department of Medical Biophysics, Karolinska Institutet, S-10401 Stockholm, Sweden.

Abstract. Molecular dynamics simulations of DNA-netropsin complexes in water were performed using the thermodynamic cycle-perturbation method to calculate the free energy difference between complexes with an adenine-containing binding site and corresponding complexes where adenines are replaced by 2,6-diaminopurines (dap). The calculations predict a free energy difference of 3.7±0.9 kcal/mol (at 300K) in favour of netropsin binding to an $(AATT)_2$ DNA sequence compared to a $(dapdapTT)_2$ sequence.

Introduction

The possibility to theoretically predict the effects of small structural changes on sequence-specific DNA-ligand equilibria would be very valuable in understanding the mechanisms of gene regulation or for designing sequence specific DNA-binding drugs. Recently, free energy calculations based on the thermodynamic cycle-perturbation approach and molecular dynamics (MD) simulations have been successfully employed to theoretically 'reproduce' experimentally observed differences in binding constants in various host-guest systems including enyme-substrate complexes[1], suggesting that similar calculations might be applied also to DNA complexes. The DNA-netropsin complex provides an excellent model system to test this possibility for two reasons: it is very well characterized with regard to structure[2] and thermodynamics[3] and there are large (free energy) differences between complexes involving different DNA sequences.

We have performed calculations of the free energy difference between two DNA-netropsin complexes (with netropsin bound to an AATT site) and the corresponding complexes with adenine replaced by 2,6-diaminopurine (dap). This particular modification was chosen because it represents a (relatively) small change in an AT base pair, i.e. the free energy change can be calculated without an exhaustive computational effort. However, the mutation might have an effect on netropsin binding that is similar to a complete AT to GC transformation, because a 2-amino group on a purine will protrude out into the netropsin binding site in the DNA minor groove forming a hydrogen bond with the pyrimidine carbonyl, as in a GC base pair (Fig. 1). Our calculations predict a

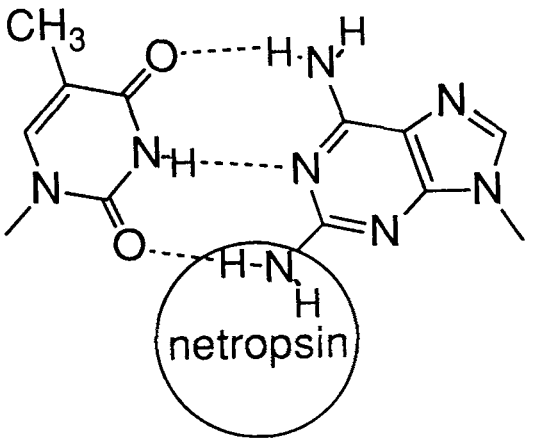

FIG. 1. Structure of the 2,6-diaminopurine-thymine base pair showing how the 'mutation' of an additional aminogroup at the adenine 6-position in an AT base pair might interfere with netropsin binding in the minor groove.

free energy difference of $\Delta\Delta G=3.7\pm0.9$ kcal/mole (at 300K) in favor of netropsin binding to an $(AATT)_2$ compared to a $(dapdapTT)_2$ DNA sequence, which is in agreement with experimentally observed free energy differences between closely related DNA-netropsin complexes[3].

Methods

The calculations are based on a thermodynamic cycle, where ΔG_1 and ΔG_2 denote (Gibbs) free energy changes upon binding of a ligand to two macromolecules (D and D'), in this case two double stranded DNAs with similar structures. The desired quantity $\Delta\Delta G=\Delta G_1-\Delta G_2$, which is directly related to the ratio of the equilibrium constants for the two complexes, is calculated as $\Delta\Delta G=\Delta G_4-\Delta G_3$, where ΔG_3 and ΔG_4 are free energy changes for the (non-physical) transformation of D into D' in the absence and presence of bound ligand, respectively. The free energy change when going from an initial to a mutated state can be calculated using statistical mechanical perturbation theory[4]:

$$\Delta G = \sum \Delta G(\lambda+\Delta\lambda) = \sum -RT \ln \langle \exp (U(\lambda) - U(\lambda+\Delta\lambda)) / RT \rangle_\lambda \qquad (1)$$

with $U(\lambda)=(1-\lambda)U_{init.}+\lambda U_{mut.}$, where λ is the coupling parameter ($0\leq\lambda\leq1$), $U_{init.}$ and $U_{mut.}$ are potential energies for inital and mutated states, respectively, and the brackets (<>) denote an ensemble average for conformations generated for the system at an intermediate state λ. A large number of conformations (MD trajectories) are generated at several λ values. Incremental free energy changes for intervals $\lambda+\Delta\lambda$ are then calculated using Eq. (1), choosing each $\Delta\lambda$ so that the full interval $0\leq\lambda\leq1$ is covered, and added to yield the total free energy change. In the present calculations the full potential function was chosen to be a linear function of λ, but other alternatives are also possible[4-6].

The free energy difference between netropsin-d$(AATT)_2$ and netropsin-(dapdapTT)$_2$ (dap=2,6-diaminopurine) complexes were calculated using two systems of different size (hexamer and octamer DNA) to check the consistency of the results. Starting structures for the simulated DNA-netropsin complexes in water (d$(GAATTC)_2$ and d$(CGAATTCG)_2$ with netropsin bound to the central AATT sequence) as well as reference DNA structures without bound netropsin were generated from published crystal structures[2,7] (Brookhaven Protein Data Bank). Terminal GC-base pairs were removed from the crystal structures to obtain hexamer and octamer DNA duplexes and polar hydrogens were added using equilibrium geometric parameters. The octamer-netropsin complex was centered in a 19Å-radius sphere of TIP3P waters[8]. 10 water molecules located within 3-5Å from phosphate oxygens were selected at random and replaced with sodium ions to yield a 75% neutralization of the total DNA charge. A "mean field stochastic deformable boundary"[9] was generated to maintain water density and to mimic the effect of solvent outside the boundary. The system (containing about 2600 atoms including 729 water molecules) was partitioned in a "reaction" region ($r\leq17$Å) for "full" molecular dynamics and a "buffer" region ($17\text{Å}<r\leq19\text{Å}$) for Langevin dynamics (300K heat bath and a Langevin friction coefficient of 50 ps^{-1}). The other three systems (octamer DNA without netropsin and hexamer DNA with and without bound netropsin) were set up using the same protocol. The hexamer DNA systems in 15Å water spheres (the outer 2Å being a Langevin region) each contained about 1200 atoms.

MD simulations and free energy calculations were performed using the CHARMM program[10]. The free energy simulation routine of the CHARMM program will be described in detail by Fleischman, Tidor, Brooks & Karplus (to be published).The CHARMM force field[11] was modified to reproduce correct base pairing energies and Watson-Crick geometries in vacuum without using an explicit hydrogen bonding potential. These modifications, which involved scaling of partial atomic charges and a (small) modification of nitrogen van der Waals parameters, will be described elsewhere. For netropsin, all partial atomic charges and geometric parameters involving the pyrrole

rings were those of Caldwell and Kollman[12] whereas other bonding and dihedral parameters (involving peptide bonds, and guanidinium and propylamidinium groups) were those of the currently supported CHARMM force field. All MD simulations were carried out using a (distance independent) dielectric constant of ε=1, and with non-bonded interaction potentials shifted to yield zero interaction energy and force at 10.5Å. The integration time step was 0.001 ps using the SHAKE algorithm[13] to constrain bonds to hydrogen atoms.

The simulations were carried out at 4 (hexamer) or 5 (octamer) intermediate λ values, where λ=0 represented the (initial) AATT sequence and λ=1 represented the (mutated) dapdapTT sequence. Thus, the mutation involved the (simultaneous) addition of four amino-groups, all in the DNA minor groove. Each trajectory involved an initial energy minimization preceding a 15 ps heating and equilibration period and "production" runs of 10 ps (hexamer) or 20 ps (octamer), which were used to calculate incremental ΔG values according to Eq. (1). Calculated (average) interaction energies did not change during the free energy production runs indicating that the systems had been well equilibrated.

It should be noted that the free energy change calculated here does not correspond to either Gibbs or Helmohotz free energies in a strict sense, because the pressure within the deformable boundary is expected to increase somewhat when the mutation involves an addition of atoms, and because the boundary potential does not have an "infinitely" steep cut-off. However, the difference between $\Delta\Delta$G and $\Delta\Delta$A is in this case expected to be negligible compared to the accuracy of the calculations.

Results and Discussion

The results obtained for the hexamer system are shown in Fig. 2. The free energy decreases (Fig. 2A) monotonically with λ, but the decrease is larger in the netropsin complex than in the reference DNA, indicating that the d(CAATTG)$_2$-netrospin complex has a lower free energy than the d(CdapdapTTG)$_2$-netropsin complex. The free energy difference (Fig. 2B) is negative for all values of λ, and levels off at $\Delta\Delta$G=-4.0±1.5 kcal/mol at λ=1. Calculations involving different number of configurations indicate that Eq. (1) converges (the error becomes smaller) as the number of generated conformations increases, as expected. Calculations for the larger octamer system yield results similar to the hexamer system ($\Delta\Delta$G=-3.4±0.9 kcal/mol), indicating that the observed free energy changes represent physically significant differences and that the results are independent of the size of the simulated systems. It has been shown that the netropsin binding affinity is determined by "local" interactions within the binding site[3]. It may therefore be

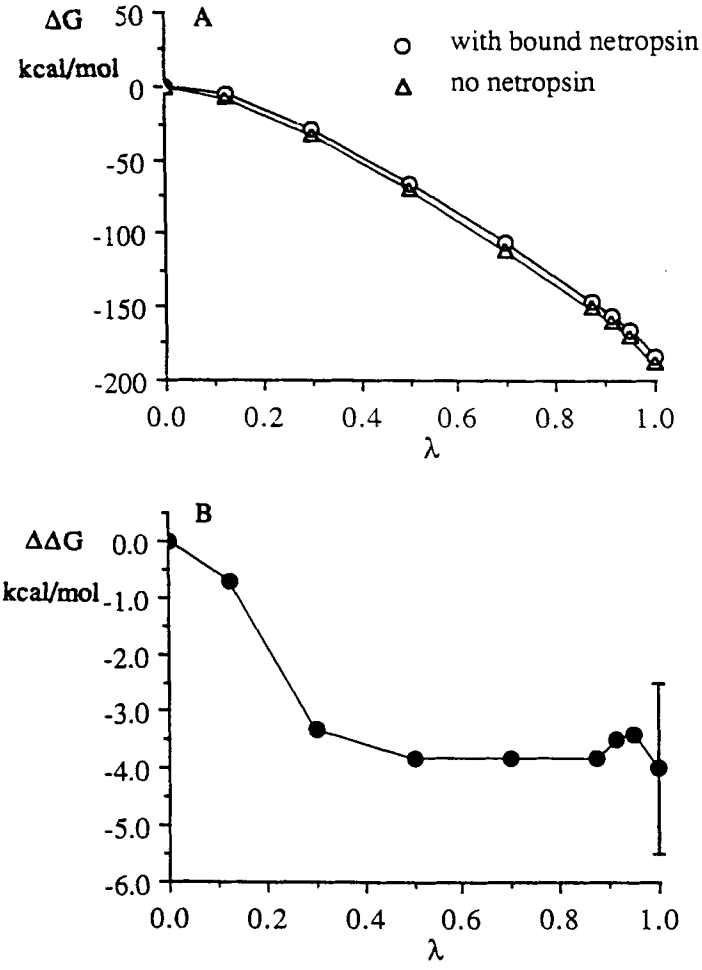

FIG. 2. (A) Free energy change (ΔG) and (B) free energy change difference (ΔΔG) as a function of the coupling parameter (λ) for the modification of $(5'GAATTC3')_2$ to $(5'GdapdapTTC3')_2$ in the absence and presence of bound netropsin. A negative ΔΔG at λ=1 corresponds to a lower free energy (higher binding constant) for the $(5'GAATTC3')_2$-netropsin complex. The error bar at λ=1 denotes the estimated standard deviation in ΔΔG calculated as described by Fleischman and Brooks[6]. The results shown in this figure are based on the 10^4 conformations generated during the time interval 15-25 ps of MD trajectories obtained for four λ values: 0.125, 0.5, 0.875 and 0.95, at a Langevin heat bath temperature of 300K. Incremental ΔG values in λ+Δλ intervals were calculated according to Eq. (1) using each trajectory to calculate two increments[6].

assumed that both calculations yield an estimate of the $\Delta\Delta G$ for changing the binding site from AATT to dapdapTT, and that the two results can be averaged ($\Delta\Delta G$=-3.7±0.9 kcal/mol).

The results compare well to experimentally observed energy differences for netropsin binding to related systems[3]. For instance, the measured $\Delta\Delta G$ for netropsin binding to poly(dAdT) versus poly(dGdC) is -5.6 kcal/mol, and $\Delta\Delta G$ for binding to a DNA oligomer with a sequence very similar to those studied here versus binding to poly(dGdC) is -4.4 kcal/mol, indicating that the AT to dapT mutation has an effect on netropsin binding that is similar to an AT to GC mutation. It may be argued that the present result cannot be compared to binding to AT versus GC sequences beacause the major groove of the dapT base pair differs from that of the GC base pair. However, studies of netropsin binding to poly(dAdT) and poly(dIdC) show that the structure of the DNA major groove is of less importance for the netropsin binding affinity[3]. (The major groove of an alternating diaminopurine-thymine polymer would be structurally similar to that of poly(dIdC)). It may also be argued that the present complexes, in which netropsin is bound to a pur-pur-pyr-pyr sequence cannot be compared to binding to alternating purine-pyrimidine sequences. However, it has also been shown that binding to the binding to the $d(CGCAATTGCG)_2$ DNA oligomer is thermodynamically similar to binding to poly(dAdT)[3]. Considering these arguments we suggest that the simulations reproduce the molecular interactions that determine netropsin discrimination between AT- and GC-rich DNA sequences, although further studies (including complete AT to GC mutations as well as calculations of contributions to ΔG due to hydrogen bonding, van der Waals interactions etc.) are required to elucidate the details of these interactions.

In summary, the present studies represent the first free energy calculations involving the sequence specificity of a DNA binding drug. The results indicate that the calculations reproduce a physically significant difference for netropsin binding to two DNA sequences, suggesting that calculations of this kind might be a valuable theoretical tool in studies of sequence specific DNA-ligand interactions.

Acknowledgements.

This research was supported by the Swedish Natural Sciences Research Council, the M. Bergvall Foundation and the Lisa and Johan Grönberg Foundation. We acknowledge the Department of Molecular Genetics at Karolinska Institutet and Dr. Ulf Skoglund for use of the Convex C210 computer and Professor Rudolf Rigler for valuable discussions.

REFERENCES

(1) McCammon, J. A. *Science* **1987**, *238*, 486-490. Lybrand, T. P.; McCammon, J. A.; Wipff, G. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 833-836. Wong, C. F.; McCammon, J. A. *J. Am. Chem. Soc.* **1986**, *108*, 3830. Bash, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A. *Science* **1987**, *235*, 574-576.

(2) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *J. Mol. Biol.* **1985**, *183*, 553-563.

(3) Marky, L. A.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 4359-4363.

(4) Mezei, M.; Beveridge, D. L. in *Computer Simulations of Chemical and Biomolecular Systems* (ed. by Beveridge, D. L. & Jorgensen, W. L.) *Annals of the New York Academy of Sciences 482*: New York 1986; pp 1-23.

(5) van Gunsteren, W. F. *Protein Engineering* **1988**, *2*, 5-13. Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184-189.

(6) Fleischman, S. H.; Brooks, C. L. *J. Chem. Phys.* **1987**, *87*, 3029-3034.

(7) Dickerson, R. E.; Drew, H. R. *J. Mol. Biol.* **1981**, *149*, 761-786.

(8) Jorgensen, W. L.; Chandrasekar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **79**, 926-935 (1983).

(9) Brooks, C. L.; Karplus, M. *J. Mol. Biol.* **1989**, *208*, 159-181.

(10) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. **1983**, *J. Comput. Chem. 4*, 187-217.

(11) Nilsson, L.; Karplus, M. *J. Comp. Chem. 1986*, *7*, 591-616.

(12) Caldwell, J.; Kollman, P. *Biopolymers* **1986**, *25*, 249-266.

(13) van Gunsteren, W. F.; Karplus, M. *Macromolecules* **1982**, *15*, 1528-1544.