SUPPLEMENTARY INFORMATION

1. The duplex state of the ligand-free 5'-CGAATTCGTCTCCGAATTCG-3' oligonucleotide is not significantly populated at temperatures below the major transition.

It follows for Fig. 2 (main text), that the observed melting process can be considered as a reversible monomolecular two-state transition. This fact together with the shape of the pre-transitional DSC, UV and CD signals suggests that at lower temperatures hairpin form is the sole DNA form in solution. The existence of the hairpin as the only DNA form in the solution is in accordance with the ITC DNA dilution experiments (performed at temperatures lower than the main transition) showing a complete absence of heat effects (Fig. 1-SI) characteristic for dissociation events like duplex–to-single strands transition. Since species (hairpins, duplexes…) having similar thermodynamic properties at low temperatures as the underlying hairpin may not be visible by calorimetric and optical melting experiments, the final confirmation that higher oligomeric DNA forms (duplex, triplex…) are not significantly populated at the conditions applied for the binding studies should come from other methods. In our case this confirmation comes from PAGE (Fig. 4d-main text) since characteristic band(s) of the ligand-free model oligonucleotide show the same mobility as the self-complementary duplex formed from two 5'-CGGAATTCCG-3' strands.

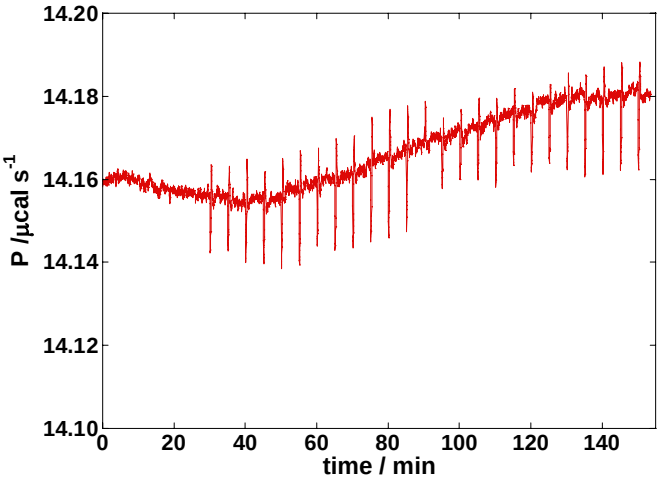

Figure 1-SI. Typical ITC dilution experiment with the model oligonucleotide 5'-CGAATTCGTCTCCGAATTCG-3'. Raw ITC signal accompanying the titration of the model DNA solution ($c_{syringe}$ = 55 µM, T = 41 °C) into the titration cell containing only the corresponding buffer solution. The heat effects do not change (decrease) with increasing concentration and are practically negligible. It should be mentioned that this is true also for some DNA sample preparations for which we observed a gradual peak in DSC thermograms before the main transition. This additional peak may be indicative for existence of another hairpin conformation observed by PAGE. Investigation of this phenomenon, however, remains our future perspective.

2. Thermodynamic analysis resulting from global χ^2 fitting of the model function (eq. 1 in the main text based on the mechanism presented there in Fig. 1) to the sets of ITC curves measured at various temperatures.

The standard Gibbs free energy of the j-th association step $\Delta G^{o}_{j,T_o} = -RT_o \ln K_j$ and the corresponding $\Delta H^{o}_{j,T_o}$ at T_o = 25 °C and $\Delta C^{o}_{P,j}$ (assumed to be temperature independent) defines $\Delta G^{o}_{j,T}$ (K_j) and $\Delta H^{o}_{j,T}$ at any T by the Gibbs-Helmholtz relation (integrated form)

$$\Delta G_{j,T}{}^{\circ} = T\left\{\Delta G_{j,T_o}{}^{\circ}/T_o + \Delta H_{j,T_o}{}^{\circ}\left[1/T - 1/T_o\right] + \Delta C^{o}_{P,j}\left[1 - T_o/T - \ln(T/T_o)\right]\right\} \qquad \text{(1-SI)}$$

and Kirchhoff's law (integrated form)

$$\Delta H^{o}_{j,T} = \Delta H^{o}_{j,T_o} + \Delta C^{o}_{P,j}(T - T_o). \qquad \text{(2-SI)}$$

Since total concentrations of ligand and DNA species are known at any ligand/DNA molar ratio, r, the temperature dependence of K_j enables the determination of equilibrium fractions of all species in the solution at any r and T and consequently their derivatives that define the model function (eq. 1 in the main text). Note that $\Delta H^{o}_{i,T}$ (eq. 1 in the main text) is not equal to $\Delta H^{o}_{j,T}$ (enthalpy change accompanying j-th step in the mechanism, as presented in Table 1-SI) but is the standard enthalpy of formation of the complex i = HL, D_2L, D_2L_2 from H and L. $\Delta H^{o}_{i,T}$ can be expressed at any T by the corresponding enthalpies $\Delta H^{o}_{j,T}$ (extrapolated to that T using Kirchhoff's law; eq. 2-SI) through the thermodynamic cycle presented in Fig. 1 of the main text.

The thermodynamic parameters ($\Delta G^{o}_{j,T_o} = -RT_o \ln K_j$, $\Delta H^{o}_{j,T_o}$, $T_o \Delta S^{o}_{j,T_o} = \Delta H^{o}_{j,T_o} - \Delta G^{o}_{j,T_o}$, $\Delta C^{o}_{P,j}$ at T_o = 25 °C) presented in Fig. 5 of the main text were obtained as average values from model analysis of three independent sets of ITC experiments (one set is presented in Fig. 4a of the main text). The corresponding errors (Fig. 5-main text) are therefore different to those obtained by the model analysis of each individual ITC data set. The best fit adjustable thermodynamic parameters obtained for each step of the proposed mechanism by using the iterative non-linear Levenberg –Marquardt regression procedure are presented in Table 1–SI.

Table 1-SI. Best fit thermodynamic parameters at 25 °C obtained from global fitting of the model function (eq. 1 in the main text) to the ITC data presented there in Fig. 4a [a]

process	parameter / unit	value (±s.d.)[b]
H + L ↔ HL	K_1 / M^{-1}	1.20 (±0.04)·10^8
	$\Delta H_1°$ / kcal mol^{-1}	-10.26 (±0.05)
	$\Delta C_{P,1}°$ / kcal mol^{-1} K^{-1}	-0.30 (±0.01)
2HL ↔ D_2L_2	K_2 / M^{-1}	1.10 (±0.10)·10^4
	$\Delta H_2°$ / kcal mol^{-1}	-7.77 (±0.27)
	$\Delta C_{P,2}°$ / kcal mol^{-1} K^{-1}	-0.08 (±0.02)
D_2L + L ↔ D_2L_2	K_3 / M^{-1}	3.48 (±0.28)·10^6
	$\Delta H_3°$ / kcal mol^{-1}	-18.68 (±0.27)
	$\Delta C_{P,3}°$ / kcal mol^{-1} K^{-1}	-0.49 (±0.01)
HL + L ↔ HL_2[c]	K_4 / M^{-1}	1.23 (±0.04)·10^4
	$\Delta H_4°$ / kcal mol^{-1}	-5.74 (±0.21)

[a] K_j, $\Delta H_j°$ and $\Delta C^{o}_{P,j}$ represent the binding constants, standard enthalpies and standard heat capacities of a given process at 25 °C that corresponds to a single step in the drug-DNA association mechanism presented in the main text.

[b] Standard deviations were obtained as square roots of diagonal elements of variance-covariance matrix.

[c] This binding step, that represents low-affinity non-specific binding, is not presented in the mechanism in the main text (Fig. 1). It was added to the model in order to achieve better agreement with the ITC curves at molar ratios $r > 1$. The values of thermodynamic parameters for the other binding events are not significantly influenced by inclusion of K_4 and $\Delta H_4°$ ($\Delta C_{P,4}°$ was set to zero), which can be seen also from the corresponding correlation matrix (Table 2-SI). Therefore, the fraction of HL (Fig. 4c in the main text) is corrected for the small contribution of HL_2 species and actually represents the sum of the HL and HL_2 fractions.

Qualitative information about the correlation between the adjustable parameters can be obtained through the variance-covariance matrix, $\boldsymbol{C}$ (consisting of elements, c_{kl}), which is defined as the inverse of the curvature matrix, $\boldsymbol{\beta}$, consisting of elements β_{kl}, where β_{kl} is given by the simplified relation [1,2]:

$$\beta_{kl} = \frac{\partial^2 \chi^2}{\partial B_k \partial B_l} \cong \sum_m \frac{1}{(\Delta(\Delta H_{T,m}))^2} \left(\frac{\partial \Delta H_T^{mod}}{\partial B_k} \frac{\partial \Delta H_T^{mod}}{\partial B_l} \right)_m , \qquad (3\text{-SI})$$

in which B_k and B_l are the adjustable parameters (K_j, $\Delta H_j°$, $\Delta C_{P,j}°$), ΔH_T^{mod} represents the model enthalpy and $\Delta(\Delta H_{T,m})$ the estimated experimental error of the m-th measured point. The

derivatives in eq. 3-SI are calculated numerically. Information about the correlation between the parameters B_k and B_l is contained in the correlation matrix ***G,*** the elements of which are defined as:

$$g_{kl} = \frac{c_{kl}}{\sqrt{c_{kk} c_{ll}}} , \qquad \text{where } -1 \le g_{kl} \le 1 \qquad \text{(4-SI)}$$

If the observed correlation $|g_{kl}|$ is very close to unity there is a high probability that the model would describe experimental data equally well by a different sets of parameters B_k and B_l, which means that their physical meaning may be lost.

In our case there are two pairs of adjustable parameters that are highly correlated K_2-K_3 and ΔH_2-ΔH_3 (Table 2-SI, marked red). We checked physical meaning of highly correlated and all other parameters by performing three sets (temperature dependence) of independent ITC experiments by using different DNA stocks from two different suppliers of DNA oligomers. Fitting of the model function (eq. 1-main text) to each ITC data-set resulted in significantly different values of the highly correlated parameters. Therefore, the obtained standard deviation from the average parameter values (different ITC data-sets) is for highly correlated parameters much higher (Fig. 5-main text) than the s.d. obtained as square roots of the diagonal elements of variance-covariance matrix resulting from fitting the model to the single data set (Table 1-SI). Thus, we consider the obtained thermodynamic parameters to be reliable within the presented (higher) error margins (Fig. 5-main text).

Table 2-SI. Correlation matrix corresponding to the model analysis of ITC data from Fig. 4a of the main text; values of adjustable parameters are given in Table 1-SI.

	K_1	$\Delta H_1°$	$\Delta C_{P,1}°$	K_2	$\Delta H_2°$	$\Delta C_{P,2}°$	K_3	$\Delta H_3°$	$\Delta C_{P,3}°$	K_4	$\Delta H_4°$
K_1	1.00										
$\Delta H_1°$	-0.37	1.00									
$\Delta C_{P,1}°$	0.30	-0.06	1.00								
K_2	0.13	0.57	-0.04	1.00							
$\Delta H_2°$	0.68	-0.17	0.44	0.49	1.00						
$\Delta C_{P,2}°$	-0.31	0.41	-0.79	0.17	-0.41	1.00					
K_3	-0.00	0.61	-0.05	0.98	0.41	0.15	1.00				
$\Delta H_3°$	0.54	0.20	0.34	0.76	0.90	-0.23	0.69	1.00			
$\Delta C_{P,3}°$	-0.10	0.54	-0.41	0.32	-0.19	0.78	0.27	0.10	1.00		
K_4	-0.46	0.23	-0.16	0.02	-0.28	0.20	0.05	-0.18	0.15	1.00	
$\Delta H_4°$	-0.01	-0.39	0.23	-0.48	-0.05	-0.31	-0.48	-0.24	-0.31	0.32	1.00

3. Structure based thermodynamic calculations.

Numerous studies on transfer of model compounds from a pure organic phase into water and on protein folding have shown that, for the observed transitions, both $\Delta C_{P,trans}°$ and $\Delta H_{trans}°$ can be parameterized in terms of the corresponding changes in solvent accessible non-polar (ΔA_N) and polar (ΔA_P) surface areas of solute molecules[3-13]. In this work we used such parameterization to correlate the heat capacity change ($\Delta C_P°$) and the enthalpy change ($\Delta H°$) with the ΔA_N and ΔA_P values that accompany ligand-DNA association. Solvent accessible surface areas for unbound ligand (netropsin), ligand-free DNA oligomer and the corresponding ligand-DNA complex were computed from the netropsin-DNA complex crystal structure (Fig. 3a in the main text; NDB entry = BD0078)[14] by the method introduced by Tsodikov et al.[15]. Exposed surfaces of carbon, carbon-bound hydrogen and phosphorus atoms were computed to give non-polar accessible surface areas, A_N, while those of all other atoms were classified as polar and are summarized in A_P. In these calculations we used a probe radius of 1.4 Å and the program default sets of atomic radii. The change in surface area on binding, ΔA, is the difference between the surface area of the complex and the summed surface areas for the ligand-free DNA oligomer and the free (unbound) ligand. In these calculations, the atomic coordinates of free ligand and ligand-free DNA were extracted from the structure of corresponding ligand-DNA complex by deleting the coordinates of either the duplex or the bound ligand. The estimated ΔA_N = -633 (± 40) Å^2 and ΔA_P = -356 (± 20) Å^2 are calculated as averages over different sets of atomic (van der Waals) radii available in the program.

By modelling the ligand-DNA binding as a rigid-body association of the ligand and DNA molecule in aqueous solution one can express $\Delta C_P°$ and $\Delta H°$ as the sum of non-polar (subscript N) and polar (subscript P) contributions[16-18]

$$\Delta C_P° = \Delta C_{P,N}° + \Delta C_{P,P}° = a\Delta A_N + b\Delta A_P \qquad (5\text{-SI})$$

and

$$\Delta H° = \Delta H°(T_H) + \Delta C_P°(T - T_H) = \Delta H_N° + \Delta H_P° = [c+a(T - T_H)]\Delta A_N + [d+b(T - T_H)]\Delta A_P, \quad (6\text{-SI})$$

Parameters a = 1.9 (±0.1) J mol^{-1} K^{-1} Å^{-2}, b = -1.1 (±0.1) J mol^{-1} K^{-1} Å^{-2}, c = -21.5 (±10) J mol^{-1} Å^{-2} and d = 205 (±20) J mol^{-1} Å^{-2} are obtained from solubility of the model cyclic dipeptides in water, while T_H = 60 °C is the selected reference temperature at which $\Delta H°(T_H)$ is parameterized as $\Delta H°(T_H) = c\Delta A_N + d\Delta A_P$.[18]

According to the model of ligand-DNA association described, the corresponding entropy of binding, $\Delta S°$, can be expressed as[10-13,18,19]

$$\Delta S° = \Delta S_{solv}° + \Delta S_{r+t}° + \Delta S_{pe}°, \qquad (7\text{-SI})$$

where $\Delta S_{solv}°$ is the entropy change resulting from the changes in solvation of drug and DNA molecules on their transfer from the unbound state to the ligand-DNA complex. It has been shown that the entropy change accompanying the processes involving transfer of non-polar groups (surface area) from solid, liquid and gas phases, as well as from protein interior to water, converges at temperature $T_S \approx 112$ °C to zero[17,20]. Accordingly, $\Delta S_{solv}°$ may be estimated as $\Delta S_{solv}° = \Delta C_P°\ln(T/T_S)$. $\Delta S_{r+t}°$ is the unfavourable entropy contribution resulting from losses in

translational and rotational degrees of freedom that accompany the formation of ligand-DNA complexes. There is a considerable debate in the literature over the acceptable approximation of the $\Delta S_{r+t}°$ term[7,10,11,21-30]. In our calculation we use the $\Delta S_{r+t}°$ value of -210 (±40) J/molK which, according to the theoretical consideration of bimolecular complex formation and from empirical consideration of specific cases that appear to represent the rigid-body association, seems to be the most appropriate one[7,21,25]. $\Delta S_{pe}°$ is the polyelectrolyte contribution derived from electrostatic polyelectrolyte theory as $\Delta S_{pe}° = -Z\Psi R\ln[Na^+]$.[31,32] In this equation $[Na^+]$ is the total concentration of Na^+ ions present in the solution ($[Na^+]$ = 0.23 M). Z is the charge on the drug molecule and Ψ is the proportion of monovalent counterions associated with each DNA phosphate group (Ψ = 0.88 for B-DNA). In effect, the product $Z\Psi$ is equivalent to the number of counterions released on binding of a ligand with charge Z.

The structure based energetics of drug-DNA rigid-body binding ($\Delta C_P°$, $\Delta H°$, $\Delta S°$, $\Delta G° = \Delta H° - T\Delta S°$) is now completely defined and can be compared to the corresponding experimental thermodynamic observations (Fig. 5-main text).

4. References-SI

1. Hallen, D. Data treatment: Considerations when applying binding reaction data to a model. *Pure & Appl. Chem.*, **65**, 1527-1532 (1993).
2. Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. *Numerical recipes,* Cambridge University Press, Oxford (1992).
3. Nichols, N., Sköld, R., Spink, C. & Wadsö, I. Additivity relations for the heat capacities of non-electrolytes in aqueous solution. *J. Chem. Thermodyn.,* **8**, 1081-1093 (1976)
4. Reynolds, J. A., Gilbert, D. B. & Tanford, C. Empirical correlation between hydrophobic free energy and aqueous cavity surface area. *Proc. Natl. Acad. Sci. USA,* **71**, 2925-2927 (1974).
5. Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H. & Zehfus, M. H. Hydrophobicity of amino acid residues in globular proteins. *Science,* **229**, 834-838 (1985).
6. Makhatadze, G. I. & Privalov, P. L. Energetics of protein structure. *Adv. Protein Chem.,* **47**, 307-425 (1995).
7. Spolar, R. S. & Record, M. T., Jr. Coupling of local folding to site-specific binding of proteins to DNA. *Science,* **263**, 777-784 (1994).
8. Myers, J. K, Pace, C. N. & Scholtz, J. M. Denaturant m values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding. *Protein Sci.,* **4**, 2138-2148 (1995).
9. Xie, D. & Freire, E. Molecular basis of cooperativity in protein folding. V. Thermodynamic and structural conditions for the stabilization of compact denatured states. *Proteins,* **19**, 291-301 (1994).
10. Baker, B. M. & Murphy, K. P. Dissecting the energetics of a protein-protein interaction: The binding of ovomucoid third domain to elastase. *J. Mol. Biol.,* **268**, 557-569 (1997).
11. Murphy, K. P., Xie, D., Thompson, K. S., Amzel, L. M. & Freire, E. Entropy in biological binding processes: Estimation of translational entropy loss. *Proteins,* **18**, 63-67 (1994).
12. Lee, K. H., Xie, D., Freire, E. & Amzel, L. M. Estimation of changes in side chain configurational entropy in binding and folding: General methods and application to helix formation. *Proteins,* **20**, 68-84 (1994).
13. Freire, E. Thermodynamics of protein folding and molecular recognition. *Pure & Appl. Chem.,* **69**, 2253-2261 (1997).
14. Van Hecke, K., Nam, P. C., Nguyen, M. T. & Van Meervelt, L. Netropsin interactions in the minor groove of d(GGCCAATTGG) studied by a combination of resolution enhancement and ab initio calculations. *Febs J.*, **272**, 3531 – 3541 (2005).
15. Tsodikov, O. V., Record, M. T., Jr. & Sergeev, Y. V. Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature. *J. Comput. Chem.,* **23**, 600-609 (2002).
16. Habermann, S. M. & Murphy, K. P. Energetics of hydrogen bonding in proteins: A model compound study. *Protein Sci.,* **5**, 1229-1239 (1996).
17. Murphy, K. P. & Freire, E. Thermodynamics of structural stability and cooperative folding behaviour in proteins. *Adv. Protein Chem.,* **43**, 313-361 (1992).

18. Robertson, A. D. & Murphy, K. P. Protein structure and the energetics of protein stability. *Chem. Rev.,* **97**, 1251-1268 (1997).
19. Spolar, R. S., Ha, J.-H. & Record, M. T., Jr. Hydrophobic effect in protein folding and other noncovalent processes involving proteins. *Proc. Natl. Acad Sci. USA,* **86**, 8382-8385 (1989).
20. Baldwin, R. L. Temperature dependence of the hydrophobic interaction in protein folding. *Proc. Natl. Acad. Sci. USA,* **83**, 8069-8072 (1986).
21. Chaires, J. B. Energetics of drug-DNA interactions. *Biopolymers,* **44**, 201-215 (1997).
22. Ren, J., Jenkins, T. C. & Chaires, J. B. Energetics of DNA intercalation reactions. *Biochemistry,* **39**, 8439-8447 (2000).
23. Page, M. I. & Jencks, W. P. Entropic contributions to rate acceleration in enzymic and intramolecular reactions and chelate effect. *Proc. Natl. Acad. Sci. USA,* **68**, 1678-1683 (1971).
24. Page, M. I. Entropy, binding-energy, and enzymatic catalysis. *Angew. Chem. Int. Edit.,* **16**, 449-459 (1977).
25. Finkelstein, A. V. & Janin, J. The price of lost freedom - entropy of bimolecular complex-formation. *Protein Eng.,* **3**, 1-3 (1989).
26. Janin, J. For Guldberg and Waage, with love and cratic entropy. *Proteins,* **24**, i-ii (1996).
27. Holtzer, A. The cratic correction and related fallacies. *Biopolymers,* **35**, 595-602 (1995).
28. Horton, N. & Lewis, M. Calculation of the free-energy of association for protein complexes. *Protein Sci.,* **1**, 169-181 (1992).
29. Brady, G. P. & Sharp, K. A. Energetics of cyclic dipeptide crystal packing and solvation. *Biophys. J.,* **72**, 913-927 (1997).
30. Tamura, A. & Privalov, P. L. The entropy cost of protein association. *J. Mol. Biol.,* **273**, 1048-1060 (1997).
31. Manning, G. S. Molecular theory of polyelectrolyte solutions with applications to electrostatic properties of polynucleotides. *Quart. Rev. Biophys.,* **11**, 179-246 (1978).
32. Record, M. T., Jr., Anderson, C. F. & Lohman, T. M. Thermodynamic analysis of ion effects on binding and conformational equilibria of proteins and nucleic acids - roles of ion association or release, screening, and ion effects on water activity. *Quart. Rev. Biophys.,* **11**, 103-178 (1978).