Recognition of the DNA Minor Groove by Thiazotropsin Analogues


Solution-phase self-association characteristics and DNA molecular-recognition properties are reported for three close analogues of minor-groove-binding ligands from the thiazotropsin class of lexitropsin molecules; they incorporate isopropyl thiazole as a lipophilic building block. Thiazotropsin B (AcImPy™ThDp) shows similar self-assembly characteristics to thiazotropsin A (FoPy™ThDp), although it is engineered, by incorporation of imidazole in place of N-methyl pyrrole, to swap its DNA recognition target from 5'-ACTAGT-3' to 5'-ACGCGT-3'. Re-placement of the formamide head group in thiazotropsin A by nicotinamide in AIK-18/S1 results in a measureable difference in solution-phase self-assembly character and substantially enhanced DNA association characteristics. The structures and associated thermodynamic parameters of self-assembled ligand aggregates and their complexes with their respective DNA targets are considered in the context of cluster targeting of DNA by minor-groove complexes.

Introduction

The rational design of DNA-targeting drugs with novel biological activities and enhanced affinities for defined DNA sequence recognition requires knowledge of the structures of complex molecular assemblies. Thermodynamic signatures are also of significance together with information derived from allied biophysical techniques when considering designed molecular recognition processes. NMR spectroscopy in particular, when applied to the solution-phase study of ligand–DNA interactions, is an excellent source of structural information and many reports, including those of ground-breaking discoveries concerned with the mode of ligand binding, have shown the utility of this key technique for assessing the architecture of ligand–DNA complexes.

To date, little has been reported concerning the relationship between ligand self-assembly and its potential influence on the process of ligand–DNA binding. Examples in which DNA binding ligands aggregate include Hoechst 33258. Hydrophobic and electrostatic components of this structure combine to create strong association between ligand molecules. This leads to large aggregates that are readily detected by NMR spectroscopy and through isothermal titration calorimetry (ITC). DNA complex formation occurs readily for this ligand in a 1:1 ligand/DNA duplex binding mode, but the binding is influenced by ligand self-association. The occurrence of ligand aggregation has led to speculation about the energetic and structural consequences of ligand self-assembly in relation to DNA recognition; this is of particular relevance when considering 2:1 ligand/DNA duplex binding where ligand-on-ligand assembly is implicit in the structure of the ligand/DNA complex.

Our work in this field has focussed on the thiazotropsin class of DNA minor-groove-binding ligands (MGBs). These were developed within the framework of the so-called “Strathclyde strategy”. Similar to many classical examples of lexitropsin/DNA complexes, our studies show this class of ligand (which contains the isopropyl-thiazole building block for enhanced lipophilicity) to bind as antiparallel, face-to-face dimers in the DNA minor groove. These ligands and their later generation analogues aggregate in free aqueous solution. The extent of aggregation depends on ligand structure. Studies show that ligand assembly occurs in a similar antiparallel, face-to-face fashion to that found in the DNA minor-groove complexes formed by the ligands. Such a coincidence cannot be ignored. The extent and generality of this assembly process and its influence on DNA recognition should be understood in the context of drug-design strategies.

To address this, a comprehensive biophysical study was conducted by using three related molecules from the thiazotropsin class of ligands: thiazotropsin A, thiazotropsin B and AIK-
Results and Discussion

Isothermal titration calorimetry of free ligands

Heats of dilution derived from control titrations of thiazotropsins A and B into buffer alone were measured by using ITC. The results were compared with those obtained previously for AIK-18/51.[18] Raw data were processed by using the IC-ITC software package,[22] which generates values for binding constants and enthalpy changes for ligand dimerisation and step-wise self-aggregation models. The heat capacity changes were determined from the temperature dependence of the observed binding enthalpy measured at 25, 35, and 45 °C. By using the equilibrium constants at 35 °C as a reference together with the enthalpies and heat capacities of self-association, it was possible to determine values for \( \ln K_\text{ref} \) between 25 and 45 °C by using a van’t Hoff treatment according to Clarke and Glew[23,24] [Eq. (1)].

\[
\ln K_T = \ln K_{\text{ref}} + \left( \frac{\Delta H_{\text{ref}}}{R} \times \left[ \frac{1}{T_{\text{ref}}} - \frac{1}{T} \right] \right) + \left\{ \frac{\Delta C_P}{R} \times \left[ \frac{T}{T_{\text{ref}}} - 1 + \ln \left( \frac{T}{T_{\text{ref}}} \right) \right] \right\}
\]

(1)

Here \( K_T \) is the binding constant at temperature \( T \), \( K_{\text{ref}} \) is the binding constant at the reference temperature, \( \Delta H_{\text{ref}} \) is the enthalpy change at the reference temperature, \( R \) is the universal gas constant, \( T_{\text{ref}} \) is the reference temperature and \( \Delta C_P \) is the heat capacity change. The results are summarised and compared in Table 1. Heats of dilution for each ligand were determined in piperazine-N,N′-bis(2-ethane sulfonic acid) (PIPES) buffer at pH 6.8. In all cases, ITC dilution experiments showed an inconsistent heat of dilution, thus indicating ligand aggregation. The possible influence of ligand protonation/deprotonation on the enthalpies of dilution was investigated by conducting parallel experiments under different buffer conditions (PIPES vs. 2-(carbamoylmethylamino)ethansulfonic acid (ACES) at pH 6.8) as different buffer salts have different ionisation enthalpies. The \( pK_a \) of ACES is near to that of PIPES but ACES has a different enthalpy of (de)protonation. The results suggested that, within experimental error, the enthalpies of ligand self-association in PIPES and ACES buffers were equal; this is consistent with aggregation not being accompanied by (de)protonation (Table 2).

All dilution processes were found to be endothermic, with heats of dilution decreasing consistently as more ligand was injected into the sample cell. Representative ITC data (Figure 1) are shown for the dilution of 0.5 mM thiazotropsin A in PIPES buffer at pH 6.8. The variation of the dilution heat effects for thiazotropsin A and its analogues was consistent with aggregation. Single dilution experiments cannot distinguish between dimerisation only and extended step-wise self-association, but a measure of the temperature dependence of self-association has the potential to provide some distinction between the two.[7]

The results (Figure 2) revealed that a step-wise self-association model reproduces the observed data for thiazotropsins A

18/51. Their free solution behaviour and their respective DNA complexes have been studied to establish how structural and functional alterations to the ligand affect the formation and stability of their DNA complexes. DNase I footprinting,[13] ITC[20] complexes have been studied to establish how structural and energetic demands. Here comparisons are made between the NMR-derived, solution-phase, three-dimensional structures for these complexes. When combined with results from formamide (thiazotropsin A) to nicotinamide (AIK-18/51) maintaining 5′-ACTAGT-3′ recognition[18] but subtly changes the structural and energetic demands. Here comparisons are made between the NMR-derived, solution-phase, three-dimensional structures for these complexes. When combined with results from self-assembly studies of each ligand and comprehensive ITC measurements, the results begin to provide a context for designing mixed-ligand systems within the framework of DNA cluster site targeting.

Results and Discussion

Isothermal titration calorimetry of free ligands

Heats of dilution derived from control titrations of thiazotropsins A and B into buffer alone were measured by using ITC. The results were compared with those obtained previously for AIK-18/51.[18] Raw data were processed by using the IC-ITC software package,[22] which generates values for binding constants and enthalpy changes for ligand dimerisation and step-wise self-aggregation models. The heat capacity changes were determined from the temperature dependence of the observed binding enthalpy measured at 25, 35, and 45 °C. By using the equilibrium constants at 35 °C as a reference together with the enthalpies and heat capacities of self-association, it was possible to determine values for \( \ln K_\text{ref} \) between 25 and 45 °C by using a van’t Hoff treatment according to Clarke and Glew[23,24] [Eq. (1)].

\[
\ln K_T = \ln K_{\text{ref}} + \left( \frac{\Delta H_{\text{ref}}}{R} \times \left[ \frac{1}{T_{\text{ref}}} - \frac{1}{T} \right] \right) + \left\{ \frac{\Delta C_P}{R} \times \left[ \frac{T}{T_{\text{ref}}} - 1 + \ln \left( \frac{T}{T_{\text{ref}}} \right) \right] \right\}
\]

(1)

Here \( K_T \) is the binding constant at temperature \( T \), \( K_{\text{ref}} \) is the binding constant at the reference temperature, \( \Delta H_{\text{ref}} \) is the enthalpy change at the reference temperature, \( R \) is the universal gas constant, \( T_{\text{ref}} \) is the reference temperature and \( \Delta C_P \) is the heat capacity change. The results are summarised and compared in Table 1. Heats of dilution for each ligand were determined in piperazine-N,N′-bis(2-ethane sulfonic acid) (PIPES) buffer at pH 6.8. In all cases, ITC dilution experiments showed an inconsistent heat of dilution, thus indicating ligand aggregation. The possible influence of ligand protonation/deprotonation on the enthalpies of dilution was investigated by conducting parallel experiments under different buffer conditions (PIPES vs. 2-(carbamoylmethylamino)ethansulfonic acid (ACES) at pH 6.8) as different buffer salts have different ionisation enthalpies. The \( pK_a \) of ACES is near to that of PIPES but ACES has a different enthalpy of (de)protonation. The results suggested that, within experimental error, the enthalpies of ligand self-association in PIPES and ACES buffers were equal; this is consistent with aggregation not being accompanied by (de)protonation (Table 2).

All dilution processes were found to be endothermic, with heats of dilution decreasing consistently as more ligand was injected into the sample cell. Representative ITC data (Figure 1) are shown for the dilution of 0.5 mM thiazotropsin A in PIPES buffer at pH 6.8. The variation of the dilution heat effects for thiazotropsin A and its analogues was consistent with aggregation. Single dilution experiments cannot distinguish between dimerisation only and extended step-wise self-association, but a measure of the temperature dependence of self-association has the potential to provide some distinction between the two.[7]

The results (Figure 2) revealed that a step-wise self-association model reproduces the observed data for thiazotropsins A

18/51. Their free solution behaviour and their respective DNA complexes have been studied to establish how structural and functional alterations to the ligand affect the formation and stability of their DNA complexes. DNase I footprinting,[13] ITC[20] and NMR spectroscopy[14] studies previously showed thiazotropsins A and B to bind to the hexanucleotide 5′-ACTAGT-3′. Replacing one of the N-methylpyrrole groups of thiazotropsin A with N-methylimidazole, to form thiazotropsin B, changes the preferred binding sequence to 5′-(A/T)CGCCG(T/A)-3′.[21] The nitrogen atoms of the imidazole and thiazole rings dictate specificity towards G-C DNA base pairs. Alteration of the head group from formamide (thiazotropsin A) to nicotinamide (AIK-18/51) maintains 5′-ACTAGT-3′ recognition[18] but subtly changes the structural and energetic demands. Here comparisons are made between the NMR-derived, solution-phase, three-dimensional structures for these complexes. When combined with results from self-assembly studies of each ligand and comprehensive ITC measurements, the results begin to provide a context for designing mixed-ligand systems within the framework of DNA cluster site targeting.
and B much better than the dimerisation model. AIK-18/51 shows a different fit that potentially indicates a greater propensity for dimer formation over extended aggregation, an effect that might be attributed to the nicotinamide head group versus a formamide (thiazotropsin A) or acetamide head group (thiazotropsin B). Models of the aggregation state of AIK-18/51 predict a significant fraction of the ligand in a dimeric state in aqueous solution. [18] The substantially higher associa-

Table 1. ITC-derived thermodynamic data for the step-wise self-association and dimerisation models of thiazotropsin A (ThzA), thiazotropsin B (ThzB) and AIK-18/51.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ThzA</td>
<td>25</td>
<td>−13.4</td>
<td>−8.73</td>
<td>2.64</td>
<td>−637.2</td>
<td>−26.4</td>
<td>−22.14</td>
<td>1.32</td>
<td>−1076</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>−20.0</td>
<td>−15.91</td>
<td>0.79</td>
<td>−38.7</td>
<td>−35.05</td>
<td>0.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>−26.5</td>
<td>−22.94</td>
<td>0.28</td>
<td>−47.9</td>
<td>−44.73</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ThzB</td>
<td>25</td>
<td>−17.0</td>
<td>−12.25</td>
<td>3.04</td>
<td>−721</td>
<td>−35.3</td>
<td>−30.96</td>
<td>1.52</td>
<td>−1219</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>−23.8</td>
<td>−19.53</td>
<td>1.06</td>
<td>−50.2</td>
<td>−46.38</td>
<td>0.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>−31.4</td>
<td>−27.60</td>
<td>0.41</td>
<td>−59.9</td>
<td>−56.49</td>
<td>0.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIK-18/51</td>
<td>25</td>
<td>−9.60</td>
<td>−1.50</td>
<td>49.9</td>
<td>−328.1</td>
<td>−19.56</td>
<td>−8.56</td>
<td>24.9</td>
<td>−733.1</td>
</tr>
<tr>
<td></td>
<td>35</td>
<td>−10.80</td>
<td>−5.27</td>
<td>16.2</td>
<td>−21.90</td>
<td>−16.39</td>
<td>8.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>−16.17</td>
<td>−9.55</td>
<td>2.45</td>
<td>−33.87</td>
<td>−24.74</td>
<td>1.22</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[a] Data reported previously in ref. [18], but included here for comparison purposes.

Table 2. Association enthalpies for thiazotropsin A and thiazotropsin B compared with AIK-18/51 in PIPES and ACES buffer.

<table>
<thead>
<tr>
<th>Ligand</th>
<th>ΔH (PIPES) [kcal mol⁻¹]</th>
<th>ΔH (ACES) [kcal mol⁻¹]</th>
<th>(ΔHₚᵢᵖᵉˢ−ΔHₚᵢᵖᵉˢ) [kcal mol⁻¹]</th>
</tr>
</thead>
<tbody>
<tr>
<td>thiazotropsin A</td>
<td>−13.4</td>
<td>−15.2</td>
<td>1.8</td>
</tr>
<tr>
<td>thiazotropsin B</td>
<td>−17.0</td>
<td>−15.5</td>
<td>−1.5</td>
</tr>
<tr>
<td>AIK-18/51</td>
<td>−19.56</td>
<td>−21.7</td>
<td>2.14</td>
</tr>
<tr>
<td>buffer</td>
<td>−2.73</td>
<td>−7.50</td>
<td>4.77</td>
</tr>
</tbody>
</table>

[a] Enthalpy of (de)protonation of PIPES and ACES buffers. [25]

Figure 1. A) Representative example of the heats of dilution of 0.5 mM thiazotropsin A in PIPES buffer at 25 °C. B) Comparison between experimental (*) and calculated (△) heats of dilution using ITC-ITC for data fitting.

Figure 2. Van’t Hoff plots associated with ligand aggregation in buffered aqueous solution. A) Thiazotropsin A; B) Thiazotropsin B; C) AIK-18/51. Stepwise self-association (●) and dimerisation (▲) models are shown together with two different treatments of the experimental data points: stepwise self-association fit (●); dimerisation fit (▲).
tion constant \( K \) of AIK-18/51 compared with those of thiazotropsins A and B (Table 1) also supports this assertion. The negative enthalpy values observed for both stepwise aggregation and dimerisation models suggest that ligand aggregation is driven by a combination of van der Waals, dipole–dipole and hydrophobic forces between the rings of the ligand and/or hydrogen bonding between the amide linkages. Modelling studies support this notion for AIK-18/51 and show that stabilisation occurs through van der Waals and hydrophobic forces in concert with the creation of new vibrational degrees of freedom and electrostatic solvation. A large part of this stabilisation is hydrogen bonding between the amide linkages. Modelling studies support this notion for AIK-18/51 and show that stabilisation occurs through van der Waals and hydrophobic forces in concert with the creation of new vibrational degrees of freedom and electrostatic solvation. A large part of this stabilisation is hydrogen bonding between the amide linkages.

**Isothermal titration calorimetry of ligand–DNA complex formation**

ITC was used also to characterise the binding of thiazotropsin B and AIK-18/51 to a selection of DNA sequences under similar conditions and to allow thermodynamic parameters to be compared with those reported previously for the binding of thiazotropsin A to DNA. For thiazotropsin B, six dodecamer DNA duplexes were used that possessed sequence variations at the central six base pair binding site; the results are summarised in Table 3.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>( \Delta H ) ( \text{[kcal mol}^{-1})</th>
<th>( \Delta S ) ( \text{[kcal mol}^{-1}\text{K}^{-1})</th>
<th>( K ) ( \text{[M}^{-1})</th>
<th>( \Delta G ) ( \text{[kcal mol}^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCGACTAGTCGC</td>
<td>-3.555</td>
<td>3.963</td>
<td>2.34 ( \times ) 10^5</td>
<td>-7.318</td>
</tr>
<tr>
<td>GCTGCTAGGCGC</td>
<td>-3.797</td>
<td>3.397</td>
<td>1.92 ( \times ) 10^5</td>
<td>-7.194</td>
</tr>
<tr>
<td>GCGGCTAGGCGC</td>
<td>-3.868</td>
<td>3.665</td>
<td>3.28 ( \times ) 10^5</td>
<td>-7.533</td>
</tr>
<tr>
<td>GCGCTTAGGCGC</td>
<td>-3.350</td>
<td>4.648</td>
<td>7.24 ( \times ) 10^5</td>
<td>-7.998</td>
</tr>
<tr>
<td>GCGAGCGCTGCG</td>
<td>-7.420</td>
<td>1.755</td>
<td>5.30 ( \times ) 10^5</td>
<td>-9.175</td>
</tr>
<tr>
<td>GCTGCGGACGC</td>
<td>-6.683</td>
<td>2.068</td>
<td>2.60 ( \times ) 10^5</td>
<td>-8.751</td>
</tr>
</tbody>
</table>

[a] Number of ligand equivalents per DNA duplex is 2.0 in each case.

These data showed large reductions in the binding affinity of thiazotropsin B to the binding site in which the central base sequence was TA when compared with previous similar data for thiazotropsin A. Low binding affinities stem from a large reduction in the binding enthalpy caused by functional group changes introduced into the structure. Replacement of TA by GC at the midpoint of each DNA binding site led to a substantial increase in DNA binding affinity. The 2-amino hydrogen bond donor of guanine is capable of forming specific hydrogen bonds with both imidazole and thiazole nitrogens, the basis by which the change in recognition sequence is defined. Thus changing TA to GC (from 5'-ACTAGT-3' to 5'-ACGCGT-3' in one case and from 5'-TCTAGA-3' to 5'-TCCGCA-3' in the other) resulted in gains in binding enthalpy of 3.87 and 2.89 kcal mol\(^{-1}\) and losses of 2.208 and 1.329 kcal mol\(^{-1}\) K\(^{-1}\) of binding entropy, respectively. Both raw ITC data and the resulting enthalpograms (corrected for the effects of the heat of dilution) show how the heat response changes for the two types of sequence (Figure 3). Interestingly, the binding affinity of thiazotropsin B for 5'-ACGCGT-3' (Figure 3 E) was higher than its affinity for 5'-TCGCGA-3' (Figure 3 F). This might be related to an adjustment of van der Waals interactions between ligand and DNA to provide an optimised arrangement of hydrogen-bond interactions between ligand pairs and the DNA minor groove.

In contrast, under the same experimental conditions, the ITC profile of complex formation between AIK-18/51 and the dodecamer duplex d(GCGACTAGTCGC)\(_2\) was considerably different (Figure 4).

The enthalpogram showed two distinct binding modes, one being exothermic while the other was endothermic. The exothermic event is attributed to minor-groove binding, whereas the endothermic event in the context of drug–DNA interac-
tions arises from hydrophobic and/or electrostatic interactions, that is, the entropically favourable release of water and counter ions. In a related study, in which a similar two-phase binding event was observed but with less pronounced data overlap arising from the two processes, it was shown that the second non-specific event could be reduced by loading the solution with excess salt (to a concentration of 100 mM NaCl; Figure S1 in the Supporting Information). This altered the ITC profile to one showing only minor-groove binding of ligand to DNA. Being acutely affected by the concentration of Na$^{+}$ ions strongly suggests that the second nonspecific binding event is electrostatic in nature. One may speculate that, for AIK-18/51, the positively charged ligand competes with Na$^{+}$ counter ions in neutralising the negatively charged DNA backbone. In solution, positive ions are condensed around the polyanionic DNA backbone helix to reduce the net charge and help stabilise the duplex. Interaction between the positively charged molecule and DNA could have the effect of dismissing the cations around the DNA. The positively charged ligand could therefore provide competing backbone neutralisation. Such data are not seen for thiazotropsin A under similar experimental conditions; this suggests that the observed effect arises through the hydrophobic influence of the pyridine ring in the nicotinamide head group of AIK-18/51. The driving force for such nonspecific electrostatic binding might arise through the lipophilic head group’s being buried within the DNA structure rather than remaining solvent exposed in a polar environment. The nonspecific binding event occurs at a ligand/duplex stoichiometry close to 2:1; this contrasts with examples of ligands with alternative structures that bind at stoichiometries of 5:1 or greater. It suggests that the driving force for the secondary binding event in the case of AIK-18/51 is greater than for other related ligands in the thiazotropsin series.

Due to data overlap caused by the simultaneous occurrence of the two binding events (separated only at high salt concentrations), it was impractical to extract thermodynamic parameters for AIK-18/51.

**Comparison of structure and behaviour of free ligand assembly**

The differences in behaviour observed through ITC for thiazotropsin A and thiazotropsin B compared with AIK-18/51 were also reflected in NMR studies carried out on the ligands alone. Previously AIK-18/51 was shown to assemble as a head-to-tail, face-to-face dimer in aqueous solution. The ligand yielded NMR data consistent with behaviour commonly observed for biomacromolecules, namely large negative NOEs and broadened $^1$H NMR resonances. Examination of thiazotropsin A and thiazotropsin B by the same token showed insignificant differences in the observed data. Neither molecule gave negative NOEs. Additionally, $^1$H NMR resonance line widths were substantially narrower than those of AIK-18/51 (Figure S2). NMR-measured diffusion coefficients determined as a function of ligand concentration were also acquired and compared (Figure 5).

Consistent with the differences noted for the $^1$H NMR spectra, the diffusion data revealed AIK-18/51 to be substantially less mobile than either thiazotropsin A or B. Diffusion measured in this way shows a weighted average based on populations of differently sized aggregates in the fast-exchange limit on the NMR chemical-shift timescale. One explanation for the observations shown in Figure 5 is that the equilibrium between free ligand and assembled aggregates is more weighted towards assembled aggregates where AIK-18/51 is concerned compared with the other two ligands, which apparently associate less with one another. The data are consistent with tighter self-association between AIK-18/51 molecules than with the other two ligands, for which the self-association constants, $K_{ass}$, are more than an order of magnitude lower than that for AIK-18/51. In all cases, dilution gives rise to an increase in the self-
diffusion coefficient, as expected for a fast-exchange equilibri-
um tending towards monomer formation at the dilution limit
at the expense of aggregate formation. This result is important
for thiazotropsins A and B. Neither of these compounds gives
rise to observable intermolecular NOEs, but the effect of dilu-
tion on their self-diffusion coefficients lends weight to sugges-
tions that these compounds aggregate in aqueous solution, be
it to a lesser degree than AIK-18/51. This difference in self-asso-
ciation on binding to DNA cannot unfortunately be evaluated
numerically due to the issues noted earlier regarding a second-
dary DNA binding mode for AIK-18/51 binding to d(CGAC-
TAGTCGCG). Nevertheless, a qualitative comparison can be
drawn as detailed below.

Structure elucidation and analysis of ligand–DNA complexes

Solution-phase NMR analysis of thiazotropsin A with the DNA
duplex d(CGACGTAGTCGCG) previously led to a detailed under-
standing of the structural features that stabilise this ligand–
DNA complex through the DNA minor groove. Further struc-
tural and energetic investigations by NMR spectroscopy and
ITC for the same ligand with other DNA sequences revealed
additional key features that were responsible for stable com-
plex formation and factors (particularly steric) that readily de-
stabilise complex formation. In an extension to these studies,
we report here structures of the DNA complexes formed by
both thiazotropsin B and AIK-18/51, thus allowing comparisons
between all three complexes to be made.

A classical NMR approach to studying ligand–DNA com-
plexes in solution was adopted for both thiazotropsin B and
AIK-18/51. Ligands were titrated against their cognate DNA se-
quences, and the process was monitored by 1D 1H and
31P(1H) NMR spectroscopy. In all cases, slow ligand exchange
was noted on the NMR chemical shift timescale when free
DNA was in excess in solution. Data indicating the titration
end point for thiazotropsin B with d(CGACGTAGTCGCG) might be
compared with similar data for AIK-18/51 binding to d(CGAC-
TAGTCGCG) (Figure 6).

To complement the previously reported structure of the
ligand-free form of d(CGACGTAGTCGCG), NMR data were
used to determine the average solution structure of free
d(CGACGTAGTCGCG). Fingerprint regions of the 2D
[1H,1H] NOESY NMR data for both free d(CGACGTAGTCGCG) and
its complex with thiazotropsin B (Figures S3 and S4, respecti-
vely) along with that of the complex between d(CGACGTAGTCGCG)
and AIK-18/51 (Figure S5) were used as a basis for defining
DNA resonance assignments in free and ligand-bound forms
(Tables S1–S6). Changes in 1H NMR chemical shifts were typical
of such complexes (Figures S6 and S7).

The 31P(1H) NMR data recorded for the thiazotropsin B/
d(CGACGTAGTCGCG) complex were similar to those reported for
the thiazotropsin A/d(CGACGTAGTCGCG) complex (Table 4). In no-

Figure 6. Regions of 1D 1H NMR spectra acquired at 600 MHz monitoring complex formation between d(CGACGTAGTCGCG) and AIK-18/51 (A–E) and between
d(CGACGTAGTCGCG) and thiazotropsin B (F–J). Resonances for labile NH protons are shown to the left-hand side of each portion of the figure and are annotated
for fully formed complex (top) and free DNA (bottom). Resonances for aliphatic methyl groups in both DNA and ligand are shown to the right-hand side of
each portion of the figure. AIK-18/51 titration: A) Free DNA; B) 0.5:1 ligand/DNA; C) 1:1 ligand/DNA; D) 1.8:1 ligand/DNA; E) 2:1 ligand/DNA. Thiazotropsin B
titration: F) Free DNA; G) 0.5:1 ligand/DNA; H) 1:1 ligand/DNA; I) 1.5:1 ligand/DNA; J) 2:1 ligand/DNA.
Three dimensional structures of both ligand-free DNA and ligand/DNA complexes were determined for both ligands (Figure 8; thiazotropsin B; Figure 9, below, Aik-18/51). Structural details were evaluated in each case by using CURVES (Table 5). Statistics relating to the structure refinements are presented in tables S7 and S8. For the thiazotropsin B/DNA complex, the two ligands fitted snugly into the DNA minor groove, stacking face-to-face in a classical antiparallel orientation (Figure 8C). The hydrogen-bonding scheme for the complex was deduced indirectly on the basis of the calculated solution structure (Table 6).

The structure in the ligand free state (Figure 8D); this follows from the inherently wide minor groove (7.9 Å), which is sufficient for the binding of the ligand to the DNA duplex is clear from the increased dispersion of signals that occurs on complex formation.

<table>
<thead>
<tr>
<th>Table 4. Comparison of $^{31}$P chemical shifts for DNA complexes of thiazotropsin B and Aik-18/51 in the absence (free) and presence (bound) of two equivalents of ligand per duplex.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Base</strong></td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>C1</td>
</tr>
<tr>
<td>G2</td>
</tr>
<tr>
<td>A3</td>
</tr>
<tr>
<td>C4</td>
</tr>
<tr>
<td>G5</td>
</tr>
<tr>
<td>C6</td>
</tr>
<tr>
<td>G7</td>
</tr>
<tr>
<td>T8</td>
</tr>
<tr>
<td>C9</td>
</tr>
<tr>
<td>G10</td>
</tr>
</tbody>
</table>

[a] Labeled according to phosphates attached at position 5' with respect to the base. [b] $\Delta \delta^{31}$P = ($\delta^{31}$Pbound - $\delta^{31}$Pfree). [c] $\delta^{31}$P with ligand bound. [d] $\delta^{31}$P for native DNA.

Table 5. Parameters associated with the structure determination of complexes between thiazotropsin B and Aik-18/51 and their cognate DNA sequences.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>d(CGACGTGGTCG)2</th>
<th>d(CGAGCGGTCTG)2</th>
<th>Aik-18/51 with d(CGACGTGGTCG)2</th>
<th>Aik-18/51 with d(CGAGCGGTCTG)2</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of NOEs</td>
<td>207</td>
<td>740</td>
<td>391</td>
<td>757</td>
</tr>
<tr>
<td>minor-groove width (Å)</td>
<td>7.9</td>
<td>8.4</td>
<td>6.6</td>
<td>7.5</td>
</tr>
<tr>
<td>minor-groove depth (Å)</td>
<td>3.6</td>
<td>4.9</td>
<td>4.6</td>
<td>4.6</td>
</tr>
<tr>
<td>axis curvature (°)</td>
<td>17.8</td>
<td>17.8</td>
<td>18.1</td>
<td>24</td>
</tr>
<tr>
<td>base-pair tip angle (θ) (°)</td>
<td>-1.05</td>
<td>0.72</td>
<td>1.48</td>
<td>-0.33</td>
</tr>
<tr>
<td>base-pair x-displacement (dx) (Å)</td>
<td>1.78</td>
<td>-0.97</td>
<td>-1.57</td>
<td>-1.30</td>
</tr>
<tr>
<td>base-pair y-displacement (dy) (Å)</td>
<td>0.25</td>
<td>0.16</td>
<td>0.19</td>
<td>0.28</td>
</tr>
<tr>
<td>base-pair inclination (ϕ) (°)</td>
<td>12.9</td>
<td>5.1</td>
<td>5.3</td>
<td>8.8</td>
</tr>
<tr>
<td>base-pair buckle (κ) (°)</td>
<td>-6.17</td>
<td>-4.45</td>
<td>-16.1</td>
<td>-16.1</td>
</tr>
<tr>
<td>propeller twist (ω) (°)</td>
<td>3.98</td>
<td>-3.16</td>
<td>-2.30</td>
<td>-2.30</td>
</tr>
<tr>
<td>sugar pucker</td>
<td>C1′-exo or C2′-endo</td>
<td>C2′-endo or C3′-endo</td>
<td>C2′-endo or C3′-endo</td>
<td>C1′-exo</td>
</tr>
</tbody>
</table>
cient to accommodate the ligand without excessively disrupting the DNA structure. The global shape of the DNA structure in the complex agreed well with that of the free DNA, and average global base-pair axis parameters generally showed only minor deviations from the values calculated for canonical B-DNA.

The energy-minimised structure of AIK-18/51 in complex with d(CGACTAGTCG)₂ is shown in Figure 9. A head-to-tail arrangement of ligands in the minor groove was deduced through NOE contact between the Dp tail of one ligand and the pyridine head of the partnering ligand. The hydrogen-bonding scheme for the complex was deduced indirectly on the basis of the calculated solution structure (Table 6). The thiazole nitrogen in AIK-18/51 was implicated in the formation of hydrogen bonds with the exocyclic NH₂ of G⁷. A hydrogen bond was assigned between H33 of the Dp tail and A¹³N₃ of the second strand of DNA, which has not previously been reported for any of the complexes studied in our laboratory. Several NOEs were observed between the Dp tail of thiazotropsin B and A¹³H₂ in d(CGACGCGTCG)₂, thus indicating close contact between the two moieties. However, the distance between H27 in the Dp tail of thiazotropsin B and A¹³N₃ (> 2.5 Å) was not close enough for a hydrogen bond to be assigned. As with the thiazotropsin B-DNA complex, the labile (amide) protons of the ligand, which form hydrogen bonds with DNA bases, appeared to be inaccessible to solvent based on 2D [¹H,¹H] NOESY NMR data.

The global shape of the DNA structure upon ligand binding agreed well with that of the free DNA, but the overall axis curvature increased from 18.1° in the calculated ligand-free B-DNA structure to 24° in the calculated ligand-bound B-DNA structure. The central T·A base pairs showed significant inclination (22°) compared with the other base pairs, and average values of base-pair shear, stretch, stagger, buckle and opening were similar to those of canonical B-DNA. Some variations in the base-pair propeller twists were noted, especially at the central T·A base pairs, which showed significant propeller twist (30°) compared with the other base pairs.

Commentary on ligand self-association and DNA recognition

Ligand aggregation in the context of DNA recognition has become a focus of attention since evidence suggested that molecules self-assemble prior to DNA binding.⁻¹⁸ This phenomenon might be more widespread than has been previously reported.⁻¹⁹ Qualitative comparison of ITC and NMR data for the ligands studied here suggests underlying differences in their self-association properties. The additional aromaticity and hydrophobic potential of the nicotinamide head group could reduce ligand-on-ligand slippage more than in the other two ligand/DNA complexes described here; this is consistent with the measured ligand self-association data. It might therefore impose an additional structural restraint, which could be supplemented through the hydrogen-bond potential of the nitrogen in the pyridine ring. Incorporating aromatic functionality into such ligands is known to influence their biological effectiveness significantly.⁻¹⁹
Commentary on thermodynamic comparisons of the formation of ligand–DNA complexes

Titrations of ligands with their cognate DNA dodecamers are associated here in all cases with exothermic processes together with endothermic effects observed upon dilution; this is consistent with ligand aggregation prior to DNA binding. Favourable enthalpies of interaction agree with observations that exothermic processes occur for the vast majority of ligands binding to DNA at room temperature.[2] Unfavourable entropies indicate that association might induce conformational changes in the ligand or DNA, restrict the conformational freedom of the complex and/or result in insignificant desolvation of the binding site.[2−26] Large favourable ΔH and unfavourable or only slightly favourable ΔS suggest that water might assist complex formation. Entropic contributions to binding might not be evident if these are offset by greater opposing factors, such as the loss of conformational freedom. Minor-groove recognition by small molecules can be driven by enthalpy or entropy or both, and the thermodynamic signature of MGBs is highly dependent on ligands’ structures and the sequences of their binding sites.[17]

Analysis of the binding isotherms resulting from the titration of thiazotropsin A and its analogues reveals that when r ≤ 2, ΔHf remains constant; this agrees with a detailed comparative study on the thermodynamics of distamycin and netropsin binding to DNA,[31,32] and shows that binding free energy is dominated by a combination of noncovalent interactions and the hydrophobic transfer of ligand from its surrounding solution to the DNA minor groove. Enthalpy changes reflect the strength of the noncovalent interactions between molecules relative to those existing with the solvent.[32] Different molecular forces, such as hydrogen bonding, electrostatic and van der Waals interactions between the ligand and DNA or ligand and solvent, can contribute to the observed enthalpy. Ligand–DNA interactions, which have large contributions from the hydrophobic and electrostatic forces, are mainly driven by entropy due to the release of water and counter ions from the polyanion DNA duplex upon ligand binding.[26,33] Those interactions in ligand–DNA complexes driven by entropy are characterised by a negative or small positive enthalpy and a large positive entropy. If electrostatic interactions play a role in binding, their contribution is expected to counteract the observed enthalpy by increasing the entropy and reducing the enthalpy. In general, the systems studied in this work are characterised by large negative enthalpy changes with unfavourable entropies, thus suggesting that electrostatic effects are not a major driving force in the interaction. Electrostatic and van der Waals interactions with the solvent are inevitable, however, because the heat of dilution of the ligand into the buffer alone is subtracted from the heat of complex formation, thereby ensuring that the measured ΔH is due to ligand–DNA interactions only.

Overall, the distinct thermodynamic signature of ligand–DNA interactions allows the molecular forces responsible for binding to be differentiated. Establishing a link between the energetics of binding and structure is important when trying to understand biomolecular interactions and so improve the binding affinity. However, there are some drawbacks: improving binding enthalpy does not necessarily lead to a higher binding affinity because of enthalpy−entropy compensation. Enthalpy gains are countered by entropy losses, leading to no net increase in affinity. One major cause of this compensation mechanism is the nature of noncovalent interactions. For example, the enthalpic gain through hydrogen bond formation within a complex is often accompanied by entropic loss as these new bonds limit movement within the complex.

Perspective on complex formation in a mixed ligand framework

Ligand binding within the DNA minor groove results in allosteric effects that are suggested to cause modulation of DNA–protein recognition in biological systems.[34] The large effect caused by a relatively small molecule recognising short (6−8 bp) sequences of DNA has led to suggestions that, in order to improve the effectiveness of small-molecule interference with biological response processes, a synergistic approach could be used in which locally similar clusters of DNA structures or sequences could be targeted.[35−37] We considered whether mixtures of ligands could be used to target subtly different but related sequences at clusters of DNA recognition sites. Synergistic action between ligands might then exert wider, more efficient gene-silencing effects through allosteric activity. The results reported here provide the impetus for us to begin exploring mixed ligand/mixed DNA sequence systems to investigate how different ligands might chaperone one another towards DNA recognition through heteroassembly. Initial indications from a mixture of thiazotropsin A and thiazotropsin B targeted towards the DNA duplex d(CGACGCGTCG)₂ d(CGACGCGTACTAGTCG) suggest that ligand heteroassembly might assist in delivering ligands to clusters of DNA recognition sites. Self-assembled dimers of each ligand then appear to bind to their respective recognition sites. These intriguing preliminary insights suggest a role for designing complementarity between different ligands that could be exploited in targeting clusters of DNA recognition sites.

Table 6. Summary of proposed hydrogen-bonding scheme in the 2:1 DNA complexes formed with thiazotropsin B and AIK-18/51 based on labile-proton-exchange characteristics and solution-structure information.

<table>
<thead>
<tr>
<th>Ligand atom</th>
<th>DNA atom</th>
<th>Distance [Å]</th>
<th>Ligand atom</th>
<th>DNA atom</th>
<th>Distance [Å]</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH-2</td>
<td>G5 N3</td>
<td>1.9</td>
<td>NH-7</td>
<td>T5 O2</td>
<td>1.9</td>
</tr>
<tr>
<td>NH-8</td>
<td>C6 O2</td>
<td>2.1</td>
<td>NH-14</td>
<td>A6 O2</td>
<td>2.2</td>
</tr>
<tr>
<td>NH-14</td>
<td>G7 N3</td>
<td>2.4</td>
<td>NH-21</td>
<td>G7 N3</td>
<td>2.5</td>
</tr>
<tr>
<td>NH-22</td>
<td>T8 N3</td>
<td>2.3</td>
<td>NH-29</td>
<td>T8 O2</td>
<td>2.2</td>
</tr>
<tr>
<td>imidazole N</td>
<td>G5 H22</td>
<td>1.9</td>
<td>NH-33</td>
<td>A13 H2</td>
<td>1.9</td>
</tr>
<tr>
<td>thiazole N</td>
<td>G7 H22</td>
<td>2.1</td>
<td>thiazole N</td>
<td>G7 H22</td>
<td>1.9</td>
</tr>
</tbody>
</table>
Conclusion

The studies described here provide clear evidence that the lexitropsins investigated bind in a sequence-specific manner to the DNA minor groove in a 2:1 ratio through face-to-face binding in a head-to-tail fashion. Comparisons of ligand-free and ligand-bound forms of DNA solution structures show that overall shape is retained, but that ligand binding widens the minor groove and bends the DNA helix towards the major groove, making it narrower and deeper. Such alterations to DNA structure form the molecular basis of allosteric inhibition of protein–DNA interactions by small molecules.

Isothermal titration calorimetry studies reveal complete thermodynamic profiles for lexitropsin–DNA associations that include binding affinity ($K$), stoichiometry ($N$), enthalpy ($\Delta H$), entropy ($\Delta S$) and free energy of binding ($\Delta G$) for the interaction. For the investigated ligands, the lexitropsin–DNA interactions are predominantly enthalpically driven. The unfavourable entropies associated with these interactions are indicative of “induced-fit” binding and conformational changes in either of the reactants. Comparing the thermodynamic binding characteristics of closely related ligand structures to specific binding sites has helped to establish how modifications in the structure influence binding affinity. The results of these studies can be used to improve the binding selectivity and the physical properties of such ligands without sacrificing binding affinity. Lipophilic functional groups can be introduced at specific positions of the heterocyclic units to modify essential physicochemical properties without affecting binding capability, and in some cases, improve the binding affinity.

Ring slippage of the face-to-face minor-groove binders investigated here allows coverage of up to six DNA base pairs. The dynamics of this process can be reduced or arrested when ligand self-association is increased, as implied for the AIK-18/51-DNA complex, a factor that should be accounted for in ligand design. In terms of gene targeting, such small-molecular-weight ligands with enhanced lipophilicity could be used, for example, to disrupt transcription factor binding to response elements of a target gene composed of 6 bp sequences, such as hormone response elements of the androgen receptor (AR), glucocorticoid receptors (GRs), mineralocorticoid receptor (MR) and progesterone receptor (PR). The concept of targeting clusters of recognition sites to boost the therapeutic effectiveness of such ligands is now a focus of attention. Our initial biophysical studies exploring the recognition of side-by-side binding sites with mixtures of different ligands show promise, and the results from these findings will be reported in due course.

Accession codes

Coordinates have been deposited in the Protein Data Bank under accession numbers 2mnb (ligand-free d(CGACGGTCG)$_2$), 2nmd (thiazotropsin B·d(CGACGGTCG)$_2$ complex), 2mne (AIK-18/51·d(CGACGGTCG)$_2$ complex) and 2mnf (d(CGACGGTCG)$_2$). Chemical-shift assignments have been deposited in the BioMagResBank (BMRB) under accession numbers 19886 (ligand-free d(CGACGGTCG)$_2$), 19888 (thiazo-
tropsin B-d(CGACGCTGCG), complex), 19889 (AIK-18/51-d(CGACGCTGCG), complex) and 19890 (d(CGACGCTGCG)).

Experimental Section

NMR spectroscopy

Materials: The self-complementary ODNs d(CGACGCTGCG), and d(CGACTAGTCG), were supplied by Alpha DNA Ltd. (Montréal, Canada) as desalted, cartridge-purified, ethanol-precipitated, lyophilised powders, which were used without further purification. N-3-(Dimethylamino)propyl)-5-isopropyl-2-(1-methyl-4-(nicotinamido)-1H-pyrole-2-carboxamido)-1H-pyrole-2-carboxami
diothiazole-4-carboxamide (AIK-18/51) and 2-(4-(4-acetamido-1-methyl-1H-imidazole-2-carboxamido)-1-methyl-1H-pyrole-2-carboxami
dio-N-3-(dimethylamino)propyl)-5-isopropylthiazole-4-carboxami
die (thiazotropsin B) were prepared as the TFA salts, as described previously.21

General NMR sample preparation: For reference data sets, ODNs were typically dissolved (concentration = 2 mm) in phosphate buffer (1100 μL, pH 7.4, H2O/D2O 90:10). Half of this solution was admitted to a 5 mm NMR tube (Wilmad, USA, tube code: 528-PP-7) together with a quantity of stock 3-(trimethylsilyl)propionic-2,2,3,3-d4 acid sodium salt solution (TSP; 0.5 μL, 100 mm in D2O). The remaining solution was freeze-dried to remove the solvent and suspended in D2O phosphate buffer (550 μL) to examine the free ODN in D2O solution.

Following titration with the respective ligand (thiazotropsin B or AIK-18/51) of the ODNs dissolved in H2O/D2O (90:10) in phosphate buffer (pH 7.4), the solution was frozen, freeze-dried and then dissolved in 99% D2O to allow NMR examination of only the non-exchangeable protons of the complex.

Complex formation between thiazotropsin B and d(CGACGCGTGC): Thiazotropsin B (3.70 mg, 6.6 μmol) was dissolved in H2O/D2O (90:10; 100 μL containing phosphate buffer) to provide a stock solution of ligand (55.0 mm). d(CGACGCGTGC), was prepared as a 2 mm solution in H2O/D2O (90:10, 550 μL containing phosphate buffer). Thiazotropsin B solution (40 μL) was added to the solution of d(CGACGCGTGC), in aliquots (10 μL each) by using a Hamilton syringe. The resulting solution was gently mixed to disperse the precipitate that formed immediately upon contact between ligand and DNA solutions. 1H NMR spectroscopy was used to confirm the status of the sample at each stage of ligand addition. 2D ligand/DNA complex formation was typically achieved following addition of 40 μL stock ligand solution. The titration end point was apparent when complete disappearance of free DNA imino proton 1H NMR resonances was noted together with simultaneous replacement by a new set of imino proton 1H NMR signals. In this state, the sample was used in the accumulation of complete NMR data sets for the purposes of structure determination.

Complex formation between AIK-18/51 and d(CGACTAGTGC): AIK-18/51 (4.04 mg, 6.5 μmol) was dissolved in H2O/D2O (90:10, 100 μL containing phosphate buffer) to provide a stock solution of ligand (55.0 mm). d(CGACTAGTGC), was also prepared as an aqueous solution (2 mm in 550 μL H2O/D2O (90:10) phosphate buffer). The procedure described above for thiazotropsin B was then repeated.

NMR spectroscopy experiments: NMR data were acquired on a Bruker Avance-III 600 NMR spectrometer operating at 600.13 MHz for 1H resonance. A standard geometry triple-reso
nance probe head equipped for z-pulsed field gradients was used. Data were acquired in an identical manner for free ODNs (d(CGACGCTGCG), and d(CGACGCTGCG),) and for the complexes formed between thiazotropsin B and AIK-18/51. The probe temperature was 298 K in all instances. One-dimensional 1H NMR data were acquired by using either a presaturation (zpgr) pulse program or a double-pulsed-field-gradient-spin-echo (dpfgse)29 approach to eliminate the solvent resonance. At 600 MHz, data were typically acquired by using digital signal processing with 128 scans over a 1H frequency width equivalent to δ = 20.0275 ppm centred at 4.692 ppm into 32K data points (acquisition time: 1.09 s) using a 90° hard pulse (P1 = 10.16 μs) and a relaxation delay of 2.0 s. For the dpfgse routine, rectangular soft pulses (bandwidth = 125 Hz) were used for selective inversion at the solvent frequency together with sine-shaped gradient pulses (1 ms duration) in a ratio of 31:11.

Two-dimensional NMR data sets: 2D [1H,1H] DQF COSY NMR data (pulse program: cosydfphpr) were acquired by using solvent signal presaturation with 16 transients for each of 512 states t1 increments over a 1H frequency width of δ = 12.0166 ppm in both ω2 and ω1 into 4K complex data points (acquisition time 372 ms) with a recycle time of 2 s for a total data accumulation time of about 4.5 h.

2D [1H,1H] TOCSY NMR data were acquired by using the dip
si2egphp pulse program with 16 transients for each of 512 states–TPPI t1 increments over an ω3 frequency width of 20.04 ppm and an ω1 frequency width of δ = 12 ppm into 8K complex data points (acquisition time 341 ms) with a recycle time of 2 s and a pulsed spin-lock mixing time of 55 ms for a total data accumulation time of 4.5 h.

2D [1H,1H] NOESY NMR data used for assignment purposes were acquired by using the noesysgphp pulse program with 32 transients for each of 1024 states–TPPI t1 increments over an ω2 frequency width of δ = 20.04 ppm in both ω2 and ω1 into 8K complex data points (acquisition time 341 ms) with a recycle time of 2 s and a mixing time of 100 ms for a total accumulation time of about 22 h.

One-dimensional 31P{1H} NMR data were acquired at 161.977 MHz (9.4 T magnetic field) by using 64 transients over a frequency width of 810 Hz (ΔΩ = 5 ppm) into 512 data points (acquisition time 316 ms) with a recycle time of 0.5 s and centred close to the centre of the DNA phosphate resonance envelop. GARP composite pulse decoupling was used for 3H decoupling during the acquisition time only.

2D [1H,31P] correlations were acquired by HSQC. The evolution period 1/(4ΔΩ) was varied for different experiments to select for differently sized couplings. 31P NMR data were referenced indirectly according to Maurer and Kalbitzer.48

Natural-abundance 2D [1H,13C] HSQC NMR data were acquired at 14.1 T by using gradient coherence selection. Data were acquired with 128 transients over a 1H (ω2) frequency width of 7 kHz into 2K complex data points (acquisition time 146 ms) for each of 512 states–TPPI t1 increments over a 13C (ω3) frequency width of 16 kHz using a recycle time of 2 s between transients. All NMR data were processed on a Dell Precision 340 workstation running under Mi
crosoft Windows XP using Topspin (version 2.1, Bruker Biospin, Karlsruhe, Germany) with appropriate processing parameters and imported into SPARKY (version 3.114) for full data analysis.

NMR data assignment strategy: 2D [1H,1H] NOESY NMR data were assigned for both free DNA duplexes and the ligand-DNA complexes by using established assignment strategies49 for right-handed B-form DNA. 2D [1H,31P] DQFCOSY NMR data were used to
assign resonances for specific protons of the ligands in their complexes with the DNA duplexes. 31P resonances were assigned on the basis of observed correlations between 1H and H3', H4' and H5/H5' resonances. 2D (1H/H15C) HSQC NMR data were used to resolve assignment ambiguities and establish geminal proton resonance pairs.

Molecular modelling

Generation of starting models: Molecular models for thiazotropsin B and AIK-18/51 complexes with d(CGACCGCTGC)2 and d(CGACCGGCTGC)2, respectively, were generated in Sybyl 6.3 from standard B-form DNA; this is consistent with the NMR data (as indicated by the presence of imino proton 1H NMR resonances in the NMR data of free and bound DNA).

Generation of input files: NMR distance restraints were derived from 2D NOESY data. Sparky was used to generate an intensity file (INT.1) by using volumes of NOESY crosspeaks. This file was used with the PDB file of the starting model as input to Mardigrais to generate the NOE distance restraints file (dist) used as input for restraints applied during molecular-dynamics simulations.

The starting model PDB files were used to construct the necessary input files; parameter/topology (prmtop) and coordinate (inpcrd) files were required to perform minimisation and restrained molecular-dynamics simulations with Amber, the main program supplied with Amber version 10. Input files were created by using Leap, the basic preparation program for Amber simulations. Antechamber was used to generate prep (internal coordinate file) and frcmod (force field file) files for the ligands (thiazotropsin B and AIK-18/51) to be read by Leap using GAFF.

NMR structure refinement: To obtain accurate starting structures for the restrained molecular-dynamics calculations in explicit water, free decamer duplexes and the ligand-DNA complexes were first refined in vacuo by using Sander. Prior to the restraint MD production phase in vacuo, minimisation was carried out to relax the systems and relieve any unfavourable clashes between atoms. Here minimisation (imin = 1) was performed for 500 steps (maxcyc; 250 steps of steepest descent, SD; (ncyc = 250) followed by 250 steps of conjugate gradient, CG) by using a nonbonded cutoff of 12 Å (cut) and no periodic boundary (ntb= 0; the Sander input files can be found in the Supporting Information). During the molecular-dynamics production phase, the Langevin dynamics approach with a collision frequency of 1 ps−1 was used at a constant temperature (300 K) for 50 ps with a time step of 1 fs. The production phase was allowed to continue for a total time of 400 ps. During molecular-dynamics simulations, the distance restraints file derived from the NOE data was used to refine the structure. The final restart file was used to create a PDB file, which was used as a starting structure for carrying out the final NMR structure refinement in explicit solvent. The simulated systems were neutralised by the addition of 16 Na+ and 18 Na+ counterions for the complex and free DNA, respectively. Each system was placed in a periodic octahedral box solvated with TIP3P water with outer edges approximately 10 Å in any direction from the closest solute atom. Periodic boundary conditions with a 12 Å cutoff for nonbonded interactions were applied, with the particle mesh Ewald (PME) method applied to account for the long-range electrostatic interactions. Further details are supplied in the Supporting Information.

Structure analysis: The program CURVES (version 5.2) was used to analyse the resulting structures by applying the helicoidal parameters used to describe a nucleic acid duplex, as defined according to the EMBO workshop on DNA curvature and bending.

Isothermal titration calorimetry

Materials: The self-complementary oligodeoxy(x)ucleotides (ODNs) 5'-GGCGACTAGTCG-3', 5'-GGCTCTGAGGCG-3', 5'-GGCGCTCTAGGC-3', 5'-GGCGGCTAGGCC-3', 5'-GGCGCGCGTGC-3' and 5'-GGCTCCGGACGC-3' were purchased from MWG-Biotech AG (Ebersberg, Germany) as HPLC-purified, salt-free ODNs, custom synthesised on the 1 μmol scale. The MGBs were prepared as described previously. Millipore-filtered water was used in the preparation of all solutions. PIPES, ACES, ethylenediaminetetraacetic acid (EDTA), and NaCl (all from Sigma–Aldrich) were used to prepare buffers.

Sample preparation: DNA was prepared by annealing the ODN samples at 90 °C for 12 min, then gradually cooling them to room temperature. The ligands and the DNA dodecamers were dissolved in degassed buffer (0.01 M PIPES, 0.02 M NaCl, 0.001 M EDTA, pH 6.8). All solutions were degassed for 20 min before use in a desiccator or in a Microcal Thermovac sample degasser to decrease the noise and obtain a stable baseline. The concentrations of the DNA solutions were determined spectrophotometrically at 260 nm by using the OD values supplied by the manufacturer. For all DNA sequences, aliquots were taken and diluted to achieve the concentration required for the ITC experiments (15 μM). The ligand solution was prepared at 0.5 mM. For the ligand self-association study, the MGBs were dissolved in degassed PIPES (0.01 M PIPES, 0.02 M NaCl, 0.001 M EDTA) or ACES (0.01 M ACES, 0.02 M NaCl, 0.001 M EDTA) buffers, and the analyses were performed at ligand concentrations of 0.5 mM.

ITC experiments

Ligand–DNA titrations: ITC was performed at 25 °C by using a Microcal VP-ITC instrument (Microcal Inc., Northampton, USA). The control units were interfaced to PCs equipped with the Origin software package for data manipulation and instrument control. The DNA concentration in the sample cell was 15 μM, and the ligand concentration in the syringe was 0.5 mM. Mixing was carried out by stirring the sample cell at 329 rpm. A 280 μL rotating syringe with an impeller profiled needle was used to perform 25 repeat injections of the ligand (10 μL) with a 300 s delay between the first five injections, a 600 s delay between the subsequent 14 injections and a 300 s delay between the last six injections. To correct for the heat of dilution of the ligand, control experiments were performed at the same temperature under similar conditions with buffer only. The heats of ligand dilution were subtracted from the subsequent heat obtained for the titration of the DNA dodecamers with the ligand, thereby yielding the heat of binding for the ligand-DNA complexes.

Ligand dilution experiments: ITC dilution studies were performed at 25, 35 and 45 °C by using a Microcal VP-ITC instrument (Microcal Inc., Northampton, USA). All dilution experiments were set up so that ligand solutions (10 μL, 0.5 mM) were added to buffer in the sample cell every 300 s up to a total of 25 injections. Mixing was carried out by stirring the sample cell at 329 rpm. The binding constant, K, and the enthalpy, ΔH, of the ligand self-association could not be obtained from the Origin package coupled with the ITC instrument because it is not capable of treating data for self-assem-
bly directly. For this reason the software package IC-ITC was used for data analysis.

Data analysis: The heat change upon addition of ligand to the DNA solution is proportional to the power represented by the area under each peak, and is given by integration of this peak with respect to time; this gives the enthalpy change as a result of injection. After correction for the heat of dilution, the binding isotherms were fitted to models by nonlinear least-squares analysis (Origin 7.0, Microcal). The Origin algorithm allows fitting to a one or two independent binding sites model. The model for one independent set of binding sites works for any number of sites, N, if all sites have the same binding constant (K) and enthalpy change (ΔH). If a macromolecule has two separate sites with different K and ΔH values, then the two independent sets of binding sites model must be used. The model for one independent set of sites was applicable for all the data generated from the ITC titrations performed in this study.

In the ligand self-assembly study, the raw data were treated with Origin to generate both integrated heat effects per injection (dh) and molar heat effects per injection (ndh). These heat effects were used to generate a dh file (containing the dh and ndh data in two parallel columns) as a notepad file format in a Microsoft Windows environment. This file, together with the file containing information on the injected volumes used during the dilution experiment (volume file, vol), served as input for Preplictic (preparation program for IC ITC) to generate the necessary files for running IC-ITC to analyse the data.

Acknowledgements

We thank Dr. Niek Buurma (University of Cardiff, UK) for access to and advice in using the software application IC-ITC.

Keywords: DNA recognition · isothermal titration calorimetry · minor-groove binders · NMR spectroscopy · self-assembly · thiazotropsins


Received: April 29, 2014
Published online on: [EMB], 0000

These are not the final page numbers!
A snug fit: How does ligand–ligand affinity influence sequence-specific molecular recognition for the DNA minor groove? Energetic, structural and assembly characteristics for three related ligands and their cognate DNA recognition complexes were compared by NMR spectroscopy and ITC. Careful engineering of ligands can modify physicochemical properties such as aggregation characteristics without compromising DNA binding ability.