Structural basis of DNA folding and recognition in an AMP–DNA aptamer complex: distinct architectures but common recognition motifs for DNA and RNA aptamers complexed to AMP

Chin H Lin and Dinshaw J Patel

Background: Structural studies by nuclear magnetic resonance (NMR) of RNA and DNA aptamer complexes identified through *in vitro* selection and amplification have provided a wealth of information on RNA and DNA tertiary structure and molecular recognition in solution. The RNA and DNA aptamers that target ATP (and AMP) with micromolar affinity exhibit distinct binding site sequences and secondary structures. We report below on the tertiary structure of the AMP–DNA aptamer complex in solution and compare it with the previously reported tertiary structure of the AMP–RNA aptamer complex in solution.

Results: The solution structure of the AMP–DNA aptamer complex shows, surprisingly, that two AMP molecules are intercalated at adjacent sites within a rectangular widened minor groove. Complex formation involves adaptive binding where the asymmetric internal bubble of the free DNA aptamer zippers up through formation of a continuous six-base mismatch segment which includes a pair of adjacent three-base platforms. The AMP molecules pair through their Watson–Crick edges with the minor groove edges of guanine residues. These recognition G·A mismatches are flanked by sheared G·A and reversed Hoogsteen G·G mismatch pairs.

Conclusions: The AMP–DNA aptamer and AMP–RNA aptamer complexes have distinct tertiary structures and binding stoichiometries. Nevertheless, both complexes have similar structural features and recognition alignments in their binding pockets. Specifically, AMP targets both DNA and RNA aptamers by intercalating between purine bases and through identical G·A mismatch formation. The recognition G·A mismatch stacks with a reversed Hoogsteen G·G mismatch in one direction and with an adenine base in the other direction in both complexes. It is striking that DNA and RNA aptamers selected independently from libraries of 10^{14} molecules in each case utilize identical mismatch alignments for molecular recognition with micromolar affinity within binding-site pockets containing common structural elements.

Address: Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10021, USA.

Correspondence: Dinshaw J Patel
E-mail: pateld@mskcc.org

Key words: adaptive binding, AMP-G recognition alignment, mismatch stabilized duplex, rectangular-shaped minor groove, two bound AMPs

Received: **7 August 1997**
Revisions requested: **9 September 1997**
Revisions received: **18 September 1997**
Accepted: **22 September 1997**

Chemistry & Biology November 1997, 4:817–832
http://biomednet.com/elecref/1074552100400817

Introduction

The methodology of *in vitro* selection and amplification of nucleic acids (also called SELEX) [1–3] has made it possible to select from random single-stranded RNA and DNA libraries unique nucleic acid folds that can target ligands of interest. (These ligands range from small cofactors to proteins and are bound with affinities in the micromolar to nanomolar range (reviewed in [4–7]).) These RNA and DNA aptamer complexes offer a unique opportunity of gaining new insights into nucleic acid architecture and recognition (reviewed in [8]) and the structural elements that contribute to catalytic activity of nucleic acids [6,7]. RNA aptamer methodology has led to the selection of novel ribozymes exhibiting kinase [9], ligase [10], amide bond cleavage [11] and amino-acid transfer [12] activities, and RNA-cleaving [13] and DNA-cleaving [14] divalent cation-dependent DNA enzymes have been identified on the basis of DNA aptamer methodology. Here we address the similarities and/or differences in the architectures, binding pockets and recognition elements of RNA [15] and DNA [16] aptamers with different binding-site sequences that target the same ligand (ATP in our case). Such an analysis is now possible after our determination of the solution structure of the AMP–DNA aptamer complex and comparison with the solution structure [17] and dynamics [18] of the AMP–RNA aptamer complex reported previously from our laboratory (see also [19]). The fundamental insights into molecular recognition that can be gained should have an impact on processes ranging from the design of ribozymes and DNAzymes to their potential therapeutic applications [20,21].

Huizenga and Szostak [16] have applied *in vitro* selection techniques to identify a DNA aptamer that binds ATP

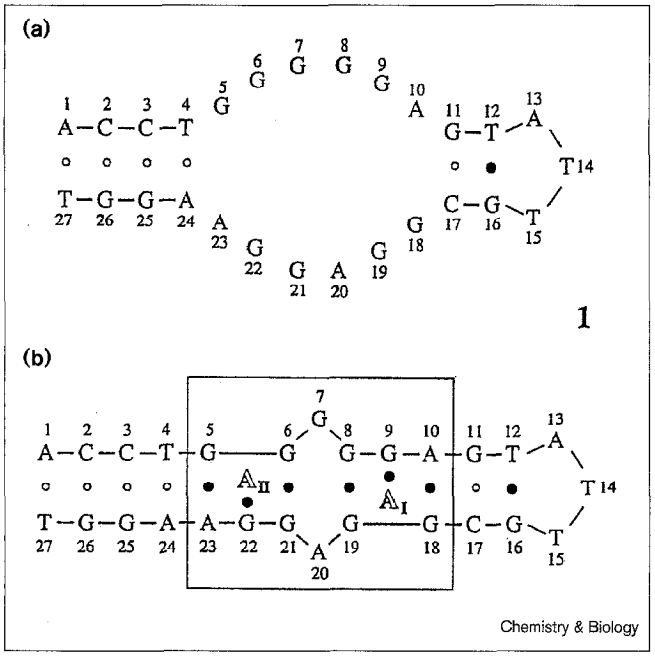
(attached through its C^8 position to an agarose column) with 6 ± 3 μM affinity. This DNA aptamer, which also binds AMP and adenosine with similar affinity, contains two highly conserved guanine-rich regions, two invariant adenine residues and two regions of predominant Watson–Crick covariation (Figure 1a). Our nuclear magnetic resonance (NMR)–molecular dynamics structure determination of the AMP–DNA aptamer complex reported below identifies the formation of a tertiary fold of the DNA aptamer containing pairs of related G·G and G·A mismatches and G·AMP recognition sites, shown schematically in Figure 1b. This architecture is different from the Huizenga–Szostak model [16] of the complex based on site-directed mutagenesis and base analog substitution experiments, which proposed that the ATP-binding pocket was centered about two bulges containing adenine residues, which flank a G-tetrad and helical stems that emanate from this G-quadruplex platform (see Figure S1 in the Supplementary material provided with the internet version of this paper).

Results

Stoichiometry of the complex

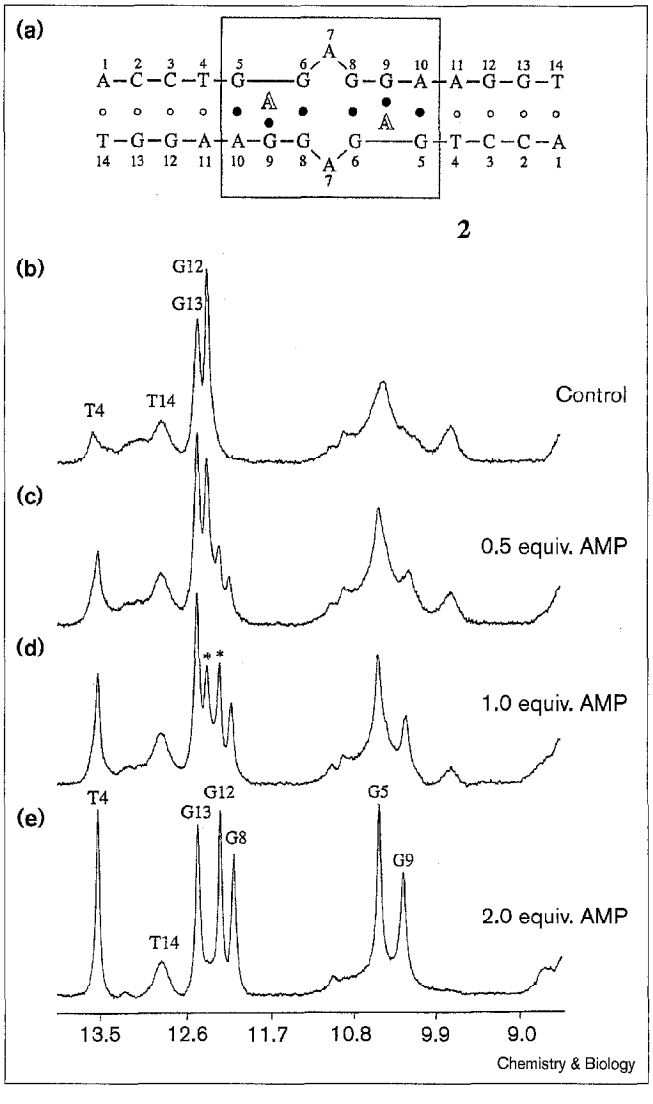
The fold of the AMP-binding 27-mer DNA aptamer **1** complex shown schematically in Figure 1b contains two nonequivalent ligand-binding sites. A symmetric version of this secondary fold can be generated through pairing of self-complementary 14-mers as shown for sequence **2** in Figure 2a. The imino proton spectra (8.5 to 14.5 ppm) of the self-complementary 14-mer AMP-binding DNA aptamer **2** in the free state and bound to two equivalents of AMP in H_2O buffer are plotted in Figure 2b and e, respectively. The imino proton spectra sharpen dramatically on complex formation with the twofold symmetry of the free duplex retained in the complex (Figure 2e). Separate resonances are observed for the free 14-mer DNA aptamer **2** and the complex containing two bound AMP molecules at substoichiometric additions of AMP, with no

Figure 1



(a) The sequence and Watson–Crick pairing alignments for the 27-mer ATP-binding DNA aptamer **1** identified through *in vitro* selection [16]. **(b)** The secondary fold determined in the present study for the ATP-binding 27-mer DNA aptamer **1** complex. There are two nonequivalent AMP-binding sites designated I and II in this complex.

Figure 2



(a) The secondary fold determined in this study for the AMP-binding self-complementary 14-mer DNA aptamer **2** in its complex with two equivalents of bound AMP. The exchangeable proton NMR spectra (8.5–14.5 ppm) of **(b)** the free self-complementary 14-mer DNA aptamer **2**, and **(c–e)** on gradual addition of 0.5, 1.0 and 2.0 equivalents of AMP to generate the AMP–14-mer DNA aptamer **2** complex in H_2O buffer, pH 6.3 at 4°C. The imino proton assignments in the spectra of the complexes are listed over the resolved resonances.

evidence for intermediate states containing a single bound ligand (Figure 2c,d). The NMR titration spectra described above (Figure 2) were recorded at a 2 mM (in strands) DNA concentration and a similar stoichiometry for complex formation was observed for titration experiments at a 0.4 mM (in strands) DNA concentration.

AMP–27-mer DNA aptamer 1 complex and single-site substitution analogs

The corresponding imino proton spectrum of the 27-mer DNA aptamer **1** complex (containing excess AMP corresponding to the addition of three equivalents of ligand) in H_2O buffer is plotted in Figure 3a. The imino proton resonances remain narrow and well dispersed in this complex, which contains two nonequivalent AMP-binding sites henceforth designated I and II. The NMR spectra of this 27-mer DNA aptamer **1** complex are of a quality necessary for structural characterization. The corresponding imino proton spectrum of the 27-mer DNA aptamer **1** analog containing a G8 for I8 substitution in H_2O buffer is plotted in Figure 3b. The spectral patterns are consistent with complex formation for this analog with the imino proton of G8 at 12.65 ppm (Figure 3a) shifting downfield to 14.39 on inosine substitution (Figure 3b). The data analysis leading to proton assignments and resolution of overlapping cross peaks was greatly aided by collecting NMR data on AMP complexes of DNA aptamer **1** analogs containing nine inosine, two adenine and one O^6-methyl guanine for guanine substitutions, three N^6-methyl adenine for adenine substitutions and one deletion in the sequence as shown schematically in Figure 3c. Single base substitutions that resulted in complex formation are represented by [++], those that resulted in weak binding are represented by [+/-] and those that resulted in no binding by [-] in the schematic in Figure 3c.

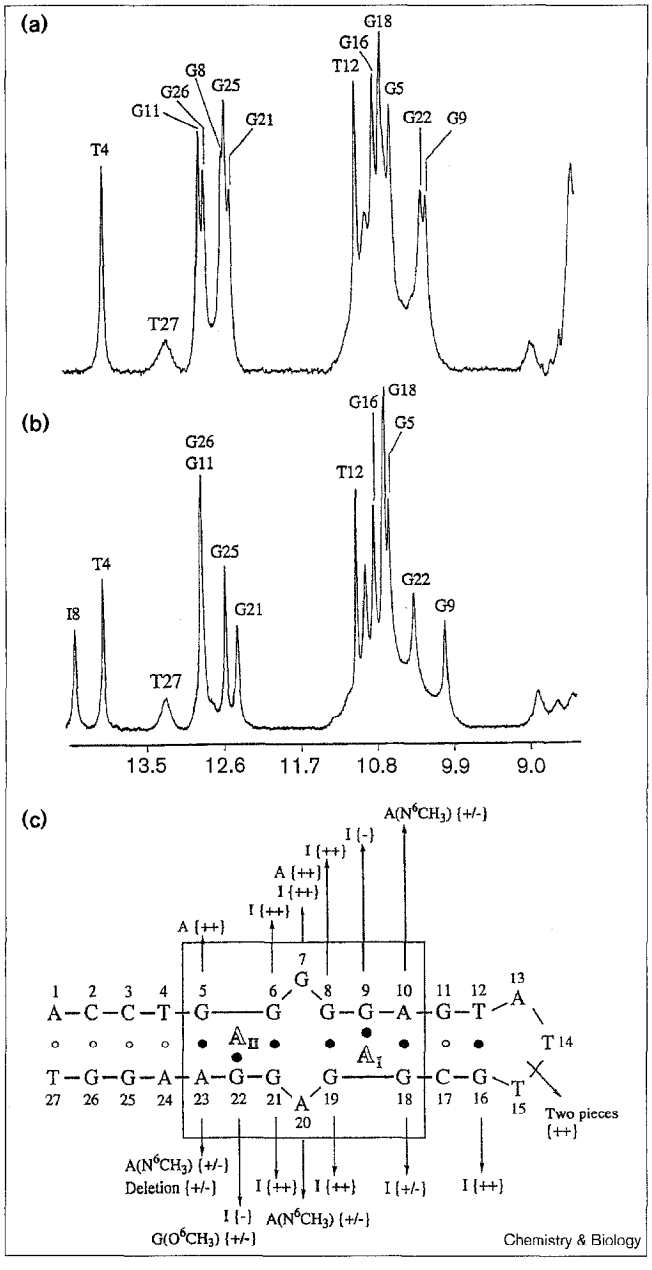
Exchangeable proton spectra and restraints

The exchangeable and nonexchangeable protons in the 27-mer DNA aptamer **1** complex have been assigned from analysis of two-dimensional NMR data sets on complexes containing both unlabeled and uniformly labeled $^{13}C,^{15}N$-AMP. The excellent quality of the two-dimensional NMR data sets can be observed in expanded NOESY (140 ms mixing time) contour plots of the corresponding unlabeled AMP–27-mer DNA aptamer **1** complex in H_2O buffer (Figure 4). These include cross-strand nuclear Overhauser effects (NOEs) between the imino protons (10.0 to 14.2 ppm) and base and amino protons (5.0 to 9.0 ppm) across Watson–Crick pairs (Figure 4a) and between hydrogen-bonded guanine amino protons (8.2 to 9.2 ppm) and nonexchangeable base protons (6.0 to 8.4 ppm) of nearby purine residues across mismatch pairs (Figure 4b) within the DNA aptamer in the complex.

Nonexchangeable proton spectra and restraints

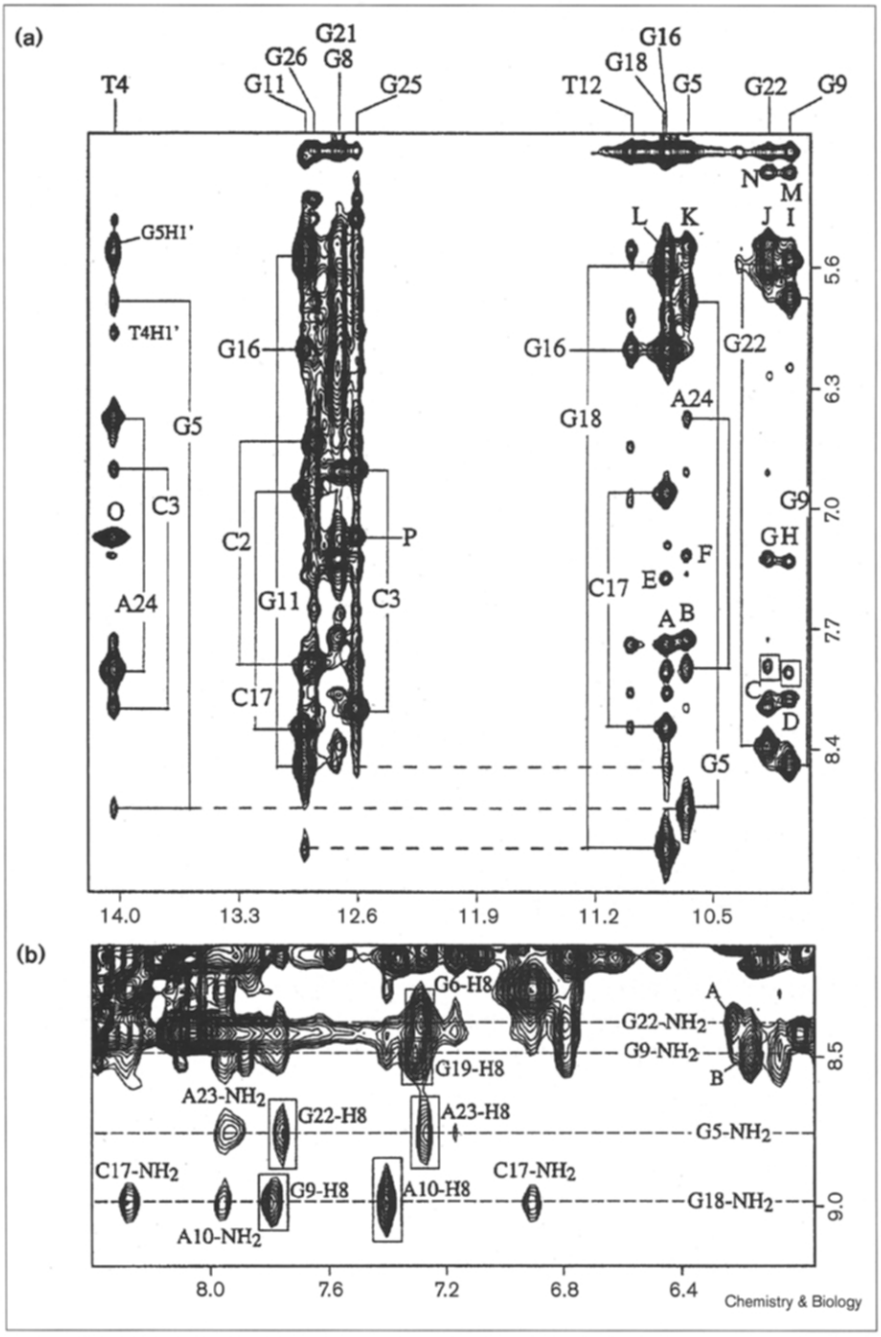
The NOESY (200 ms mixing time) contour plots of the unlabeled AMP–27-mer DNA aptamer **1** complex in D_2O

Figure 3



The exchangeable proton NMR spectra (8.5–14.5 ppm) of **(a)** the AMP–27-mer DNA aptamer **1** complex (3.0 equivalents AMP per DNA) and **(b)** the AMP–27-mer DNA aptamer **1** analog (G8 to I8 substitution) complex (3.0 equivalents AMP per DNA) in H_2O buffer, pH 6.3 at 4°C. The imino proton assignments in the spectra of the complexes are listed over the resolved resonances. Note the downfield shift in the imino proton of G8 (12.65 ppm) (a) on I8 (14.39 ppm) (b) substitution. **(c)** A schematic summary of single-base substitution analogs of the AMP–27-mer DNA aptamer **1** complex that were investigated in the present study to confirm proton assignments and resolve potential ambiguities in the assignments. These include inosine for guanine substitutions at G6, G7, G8, G9, G16, G18, G19, G21 and G22, adenine for guanine substitutions at G5 and G7, O^6-methyl guanine for guanine substitution at G22, N^6-methyl adenine for adenine substitutions at A10, A20 and A23 and deletion of A23 in the sequence. Substitutions that resulted in complex formation are represented by {++}, those that resulted in weak binding are represented by {+/–} and those that resulted in no binding by {–}.

Figure 4



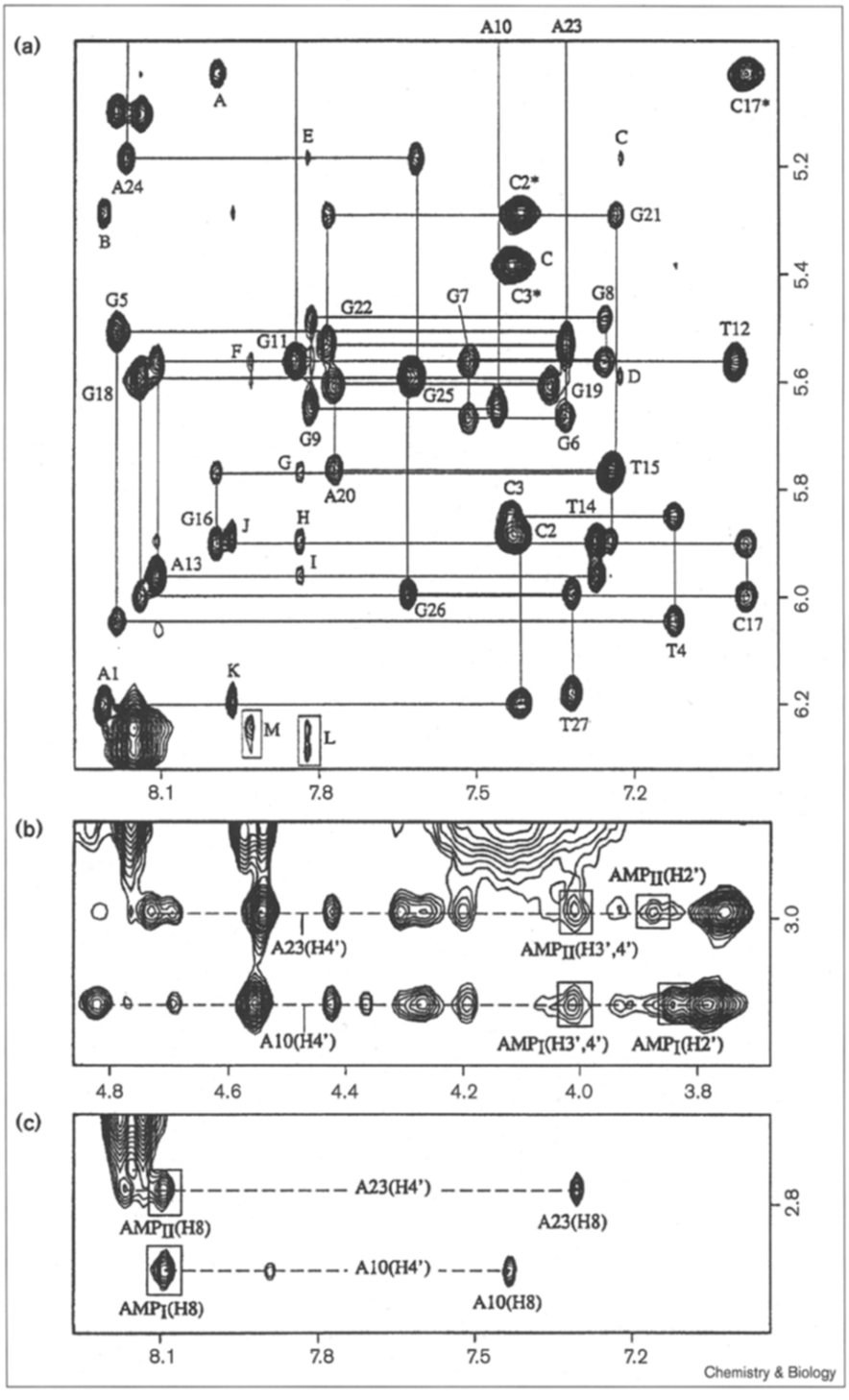
Expanded NOESY contour plots (140 ms mixing time) of the AMP–27-mer DNA aptamer **1** complex (3.0 equivalents AMP per DNA) in H_2O buffer, pH 6.3 at 4°C.
(a) Expanded region correlating NOEs between the imino protons (9.9–14.2 ppm) and the amino, base and sugar H1′ protons (4.8 to 9.2 ppm). The labeled cross peaks correspond to NOEs between the imino protons and the resolved hydrogen-bonded and exposed amino protons of guanine and cytidine residues. The labeled peaks A to P are assigned as follows:
A, G18(NH1)–G9(H8);
B, G5(NH1)–G22(H8);
C, G22(NH1)–G5(H8);
D, G9(NH1)–G18(H8);
E, G18(NH1)–A10(H8);
F, G5(NH1)–A23(H8);
G, G22(NH1)–G6(H8);
H, G9(NH1)–G19(H8);
I, G9(NH1)–G18(H1′);
J, G22(NH1)–G5(H1′);
K, G5(NH1)–G22(H1′);
L, G18(NH1)–G9(H1′);
M, G9(NH1)–G18(H3′);
N, G22(NH1)–G5(H3′);
O, T4(NH3)–A24(H2);
P, G25(NH1)–A24(H2).
(b) Expanded region correlating intermolecular NOEs between the base and sugar H1′ protons (6.0–8.4 ppm) and the hydrogen-bonded guanine amino protons (8.2– 9.2 ppm). The assignments of the NOEs between the hydrogen-bonded guanine amino protons (assigned chemical shifts indicated by dashed horizontal lines) and the base protons (boxed peaks with assignments) are labeled along with intermolecular NOE cross peaks A and B whose assignments are as follows.
A: G22(NH_2)–AMP_{II}(H2) and
B: G9(NH_2)–AMP_{I}(H2).

buffer solution at 10°C also exhibit extremely well resolved cross peaks as shown for the expanded region correlating NOEs between the base protons (6.9 to 8.3 ppm) and the sugar H1′ protons (4.9 to 6.3 ppm) within the DNA aptamer in the complex (Figure 5a). We can trace the connectivities between base protons and their own and 5′-flanking sugar H1′ protons from residue A1 to T27 except at G5–G6 and G18–G19 steps in the sequence of the DNA aptamer in the complex (Figure 5a). Exchangeable and nonexchangeable proton chemical shifts for the AMP–27-mer DNA aptamer **1** complex are given in Table S1 in the Supplementary material available with the internet version of this paper.

Intermolecular restraints in the complex

The two nonequivalent AMP molecules bound to the 27-mer DNA aptamer **1** exhibit distinct chemical shifts, permitting intermolecular NOEs to be identified within each ligand-binding site in the complex. Some examples of intermolecular NOEs are outlined in Figure 5b and c

Figure 5



An expanded NOESY contour plot (200 ms mixing time) of the AMP–27-mer DNA aptamer **1** complex (3.0 equivalents AMP per DNA) in D_2O buffer, pH 6.3 at 10°C. **(a)** Expanded region correlating NOEs between the base protons (6.9–8.3 ppm) and the sugar H1′ and cytidine H5 protons (5.0–6.3 ppm). The lines trace the NOE connectivities between the base protons and their own and 5′-flanking sugar H1′ protons along the length of the DNA aptamer sequence. The cytidine H6–H5 cross peaks are designated by asterisks. Intermolecular NOEs are shown as boxed cross peaks. The labeled peaks A to M are assigned as follows:
A, G16(H8)–C17(H5);
B, A1(H8)–C2(H5);
C, A24(H2)–A24(H1′);
D, A24(H2)–G25(H1′);
E, A23(H2)–A24(H1′);
F, A10(H2)–G11(H1′);
G, A13(H2)–T15(H1′);
H, A13(H2)–T14(H1′);
I, A13(H2)–A13(H1′);
J, A1(H2)–C2(H1′);
K, A1(H2)–A1(H1′);
L, A23(H2)–AMP_{II}(H2);
M, A10(H2)–AMP_{I}(H2).
(b) Expanded region correlating intermolecular NOEs (boxed cross peaks) between the A10 and A23 sugar H4′ protons and the bound AMP^{I} and AMP^{II} sugar H2′, H3′ and H4′ protons (3.7–4.9 ppm), respectively. **(c)** Expanded region correlating intermolecular NOEs (boxed cross peaks) between the A10 and A23 sugar H4′ protons and the bound AMP_{I} and AMP_{II} base H8 protons (3.7–4.9 ppm), respectively.

with additional examples available in the Supplementary material. The assigned intermolecular NOEs between the bound AMP_I and AMP_{II} ligands and the DNA aptamer protons in the AMP–27-mer DNA aptamer **1** complex are listed in Table 1 with their magnitudes defined as s (strong), m (medium), w (weak) and vw (very weak).

We detect exchange of the bound AMP molecules between binding sites and with free AMP as monitored by the presence of exchange cross peaks in NOESY and ROESY (rotating frame NOESY) spectra of the 27-mer DNA aptamer **1** complex (containing excess AMP corresponding to the addition of three equivalents of ligand) in D_2O solution. The bound AMP molecules exhibit NOEs

to protons associated with their individual binding sites at short mixing times and, in addition, weaker NOEs to protons associated with the other binding site at longer mixing times as a result of chemical exchange.

Formation of mismatched pairs on complex formation

The analysis of intramolecular and intermolecular NOE patterns identified a range of mismatch alignments on AMP–27-mer DNA aptamer **1** complex formation. The alignment of the sheared G18·A10 (also G5·A23) mismatch pair is defined by strong NOEs between the amino protons of G18 and the H8 proton of A10 and G9, strong NOEs between the amino protons of A10 and the sugar H1′ proton of G18 and an upfield chemical shift of ≈ 10.5 ppm for the exposed imino proton of G18 in the complex.

The alignment of the reversed Hoogsteen G8·G19 (also G21·G6) mismatch pair is defined by a strong NOE between the imino proton of G8 and the H8 proton of G19, a downfield chemical shift of ≈ 12.3 ppm for the hydrogen-bonded imino proton of G8 and a broadened-out resonance for the exposed imino proton of G19 in the complex.

The AMP_I·G9 (also AMP_{II}·G22) recognition alignment is defined by a strong NOE between the hydrogen-bonded amino proton of G9 and the H2 proton of AMP_I, a strong NOE between the hydrogen-bonded amino proton of AMP_I and the H1′ proton of G9 and an upfield chemical shift of ≈ 10.3 ppm for the exposed imino proton of G9 in the complex.

Structure calculations

The solution structure of the complex was solved starting from distance geometry followed by distance-restrained molecular dynamics calculations guided by a restraints list which included 44 intermolecular NOEs (Table 1), with the computational protocols outlined in the Materials and methods section. We incorporated 15 experimentally identified hydrogen-bonding distance restraints for the stem Watson–Crick and wobble T12·G16 pairs (Figure 1a) in the complex during the distance geometry phase of the computations. An additional 12 experimentally identified hydrogen-bonding distance restraints for the six mismatches in the central core of the complex (Figure 1b) were incorporated at the start of the distance-restrained molecular dynamics computations in one set of calculations (protocol 1) but excluded in a second set of calculations (protocol 2). It should be noted that similar refined structures of the complex resulted from using protocols 1 and 2, implying that the distribution of intramolecular and intermolecular NOEs in the core of the complex defines the mismatch pairing alignments.

Structure analysis

Seven superpositioned distance-refined structures of the central segment of the 27-mer DNA aptamer **1** complex (excluding poorly defined phosphates of bound AMPs) based on computational protocol 1 are plotted in stereo in Figure 6a and exhibit root-mean-squared deviation (r.m.s.d.) values of 1.40 ± 0.27 for the entire complex and 1.17 ± 0.18 for the core (G5–G10, G18–G23 and bound AMPs) of the complex (Table 2a). Individual residues in the Watson–Crick and mismatch-paired stem segment are well defined amongst the refined structures of the complex (shown schematically in Table 2b). A representative distance-refined structure of the central segment of the 27-mer DNA aptamer **1** complex containing two bound AMP molecules is shown in stereo in Figure 6b. The refined structures of the complex are quite similar using computational protocol 2 (see Figure S2 in the Supplementary material) with corresponding r.m.s.d. values in the range 1.48 ± 0.33 for the entire complex and 1.21 ± 0.18 for the core of the complex (Table 2a).

The two AMP-binding sites, which are located in the minor groove of the DNA helix, span a continuous six-base mismatch segment with all mismatches containing guanine and adenine residues adopting *anti* glycosidic torsion angle alignments in the complex (Figure 6b). The

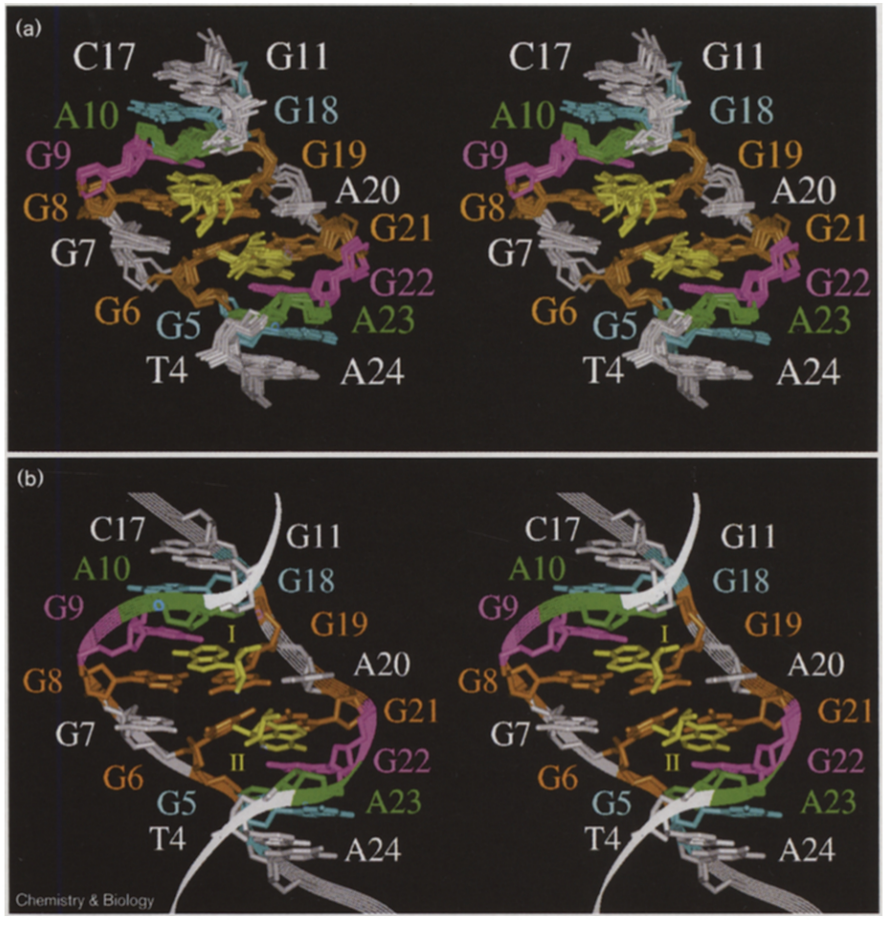
Table 1

Intermolecular NOEs between the bound AMP_I and AMP_{II} ligands and the DNA aptamer protons associated with their binding sites in the AMP–27-mer DNA aptamer 1 complex.

AMP_I protons	Intermolecular NOEs in binding pocket I
H8 (8.07 ppm)	A10-H4′ (m), A10-H5′,H5″ (m)
H2 (6.25 ppm)	G8-NH1 (m), G9-NH1 (w), G9-NH_2 (s), A10-H2 (m), G19-H8 (w), G19-H1′ (vw), G19-H4′ (w), G19-H5′,H5″ (w)
H1′ (5.45 ppm)	A10-H1′ (w), A10-H4′ (w), I19-H2 (m)*
H2′ (3.85 ppm)	A10-H4′ (m)
H3′ (4.03 ppm)	A10-H4′ (w)
H4′ (4.00 ppm)	A10-H4′ (w)
NH_2 (6.78, 7.46 ppm)	G8-NH1 (m), G8-NH_2 (m), G9-H1′ (s), A10-H8 (m)
AMP_{II} protons	Intermolecular NOEs in binding pocket II
H8 (8.06 ppm)	A23-H4′ (m), A23-H5′,H5″ (m)
H2 (6.20 ppm)	G21-NH1 (m), G22-NH1 (w), G22-NH2 (s), A23-H2 (m), G6-H8 (w), G6-H1′ (vw), G6-H4′ (w),G6-H5′,H5″ (w)
H1′ (5.56 ppm)	I6-H2 (m)*, A23-H1′ (w), A23-H4′ (w)
H2′ (3.81 ppm)	A23-H4′ (m)
H3′ (4.03 ppm)	A23-H4′ (w)
H4′ (4.00 ppm)	A23-H4′ (w)
NH_2 (6.79, 7.76 ppm)	G21-NH1 (m), G21-NH_2(m), G22-H1′ (s), A23-H8 (m)

*NOEs observed for the I6 and I19 substituted AMP–DNA aptamer 1 complex.

Figure 6



(a) Superposed stereo view of seven distance-refined structures (using protocol 1) of the central segment (T4–G11, C17–A24 and bound AMPs) of the AMP–27-mer DNA aptamer **1** complex. The bases involved in the G5·A23 and G18·A10 mismatches are colored blue-green, the G22·AMP_{II} and G9·AMP_{I} recognition sites are colored magenta-yellow and the G6·G21 and G19·G8 mismatches are colored orange-orange. The G7 and A20 bases which participate in base platforms at G6-G7 and G19-A20 steps are colored white. The phosphate groups of the two bound AMPs, which are poorly defined amongst the refined structures, have been deleted for clarity. **(b)** The corresponding stereo view of one representative distance-refined structure (using protocol 1) of the central segment (T4–G11, C17–A24 and bound AMPs) of the AMP–27-mer DNA aptamer **1** complex. A ribbon is drawn through the backbone. The color code is the same as in (a).

two nonequivalent AMP-binding sites in the 27-mer DNA aptamer **1** complex (Figure 1b) exhibit features in common with the AMP_{I}·G9 recognition mismatch flanked by sheared A10·G18 and reversed G8·G19 mismatches at site I (pairing alignments shown in Figure 7b) and the AMP_{II}·G22 recognition mismatch flanked by sheared A23·G5 and reversed G21·G6 mismatches at site II (Figure 6b).

The helix is unwound within individual (G5–G6)·(G21–G22–A23) and (G18–G19)·(G8–G9–A10) AMP-binding sites and is also bent slightly towards the major groove centered about the adjacently bound AMP molecules in the complex. The sugar H1′ protons on the two bound AMP molecules are directed towards and approach within 3.6 Å of each other, as do the sugar H4′ protons (within 3.0 Å of each other) among the refined structures of the complex.

There is extensive cross-strand stacking between bases on adjacent mismatch pairs as observed between G9 and G18 and between AMP_{I} and A10 for binding site I (Figure 7a) and as observed for the corresponding cross-strand stacking between residues associated with the binding site II counterpart in the complex.

In addition, adjacent G19 and A20 purine residues are approximately coplanar, as are adjacent G6 and G7 residues (Figure 7c), resulting in the formation of adjacent mutually stacked (G8·G19)·A20 and (G21·G6)·G7 three-base platforms (Figure 8a). There is both cross-strand (between G6 and G19) and same-strand stacking between these platforms (Figure 8b), with a significant displacement in the helix axis associated with the two halves of the complex at this central step (Figure 6).

The generation of this unique adjacent three-base platform architecture is associated with an expanded rectangular minor groove binding pocket (in which segments of the phosphodiester backbone are aligned normal to the helix axis) capable of accommodating the two adjacently bound AMP molecules, as shown in space-filling views of the complex in the right panels of Figure 9a and b.

Discussion

Huizenga and Szostak model for the AMP–DNA aptamer complex

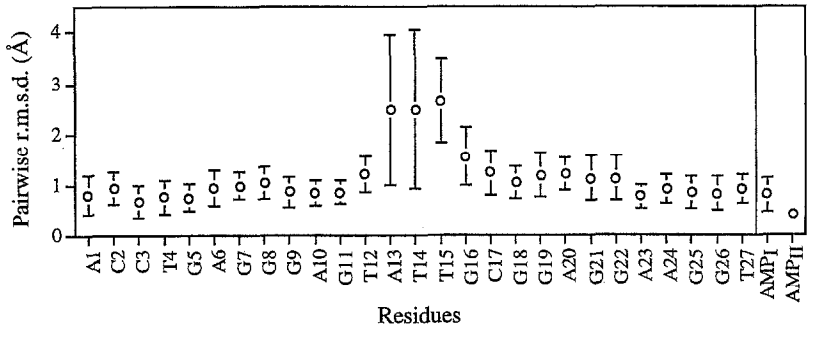
Huizenga and Szostak [16] have proposed a model for the nucleic acid fold and the ligand-binding site in the complex of the ATP-binding DNA aptamer. On the basis

Table 2

(a) NMR refinement statistics for the AMP–27-mer DNA aptamer 1 complex.

NMR distance restraints for the entire complex		
DNA aptamer distance restraints	716	
Torsion restraints (5 each for 2 AMP residues in complex)	10*	
Hydrogen bond restraints	15† + (12)‡	
AMP distance restraints	16	
Intermolecular distance restraints	44	
Structural statistics for the entire complex	Protocol 1‡	Protocol 2‡
NOE violations		
Number > 0.2 Å	18.5 ± 3.5§	15.4 ± 1.2#
Maximum violations (Å)	0.45 ± 0.22	0.42 ± 0.22
r.m.s.d. of violations	0.073 ± 0.006	0.072 ± 0.002
Deviations from the ideal covalent geometry		
Bond length (Å)	0.012 ± 0.003	0.012 ± 0.003
Bond angle (°)	3.62 ± 0.02	3.68 ± 0.02
Impropers (°)	0.37 ± 0.03	0.36 ± 0.02
Pairwise r.m.s.d. values (Å) among distance-refined structures		
Entire AMP–DNA aptamer complex		
Entire complex	1.40 ± 0.27	1.48 ± 0.33
Entire complex less phosphates of bound AMP	1.34 ± 0.27	1.37 ± 0.32
Core complex (G5–A10 and G18–A23 plus AMP)		
Core complex	1.17 ± 0.18	1.21 ± 0.18
Core complex less phosphates of bound AMP	0.99 ± 0.17	1.05 ± 0.17

(b) Pairwise individual r.m.s.d.s in the AMP–27-mer DNA aptamer 1 complex based on protocol 1.

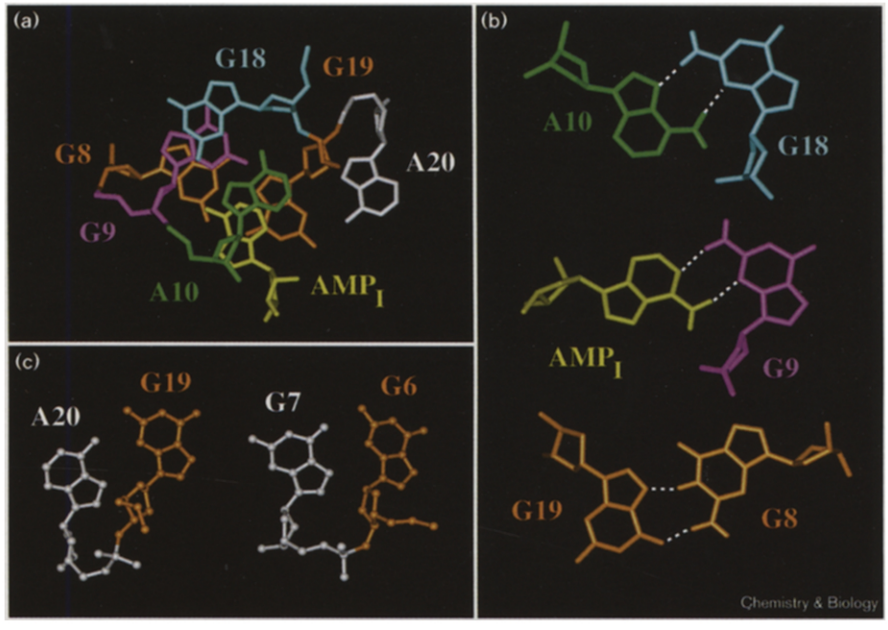


*The bound AMPs adopt a C3′-endo conformation in the complex based on strong NOEs between the H8 and their own H3′ protons. Torsion angle restraints were incorporated to retain this sugar pucker of the bound AMP molecules in the complex during the refinement. †A total of 15 experimentally identified hydrogen-bonding restraints for the stem Watson–Crick and wobble T12·G16 pairs in the complex were used during the distance geometry phase of the computations. ‡An additional 12 experimentally identified hydrogen-bonding distance restraints (± 0.15 Å) for the six mismatches within the central core of the complex (G5–A10 and G18–A23 plus AMP) have been incorporated at the beginning and retained during the distance-restrained molecular dynamics phase of the computations in protocol 1. In contrast, no non-canonical pairing restraints were used during the distance-restrained molecular dynamics computations in protocol 2. §NOE violations > 0.2 Å in the core of the complex involving protocol 1: 11.1 ± 1.4. #NOE violations > 0.2 Å in the core of the complex involving protocol 2: 11.2 ± 1.5.

of site-directed mutagenesis and base analog substitution experiments [16], they proposed that the ATP-binding pocket was centered about two adenine bulges that flank a G-tetrad and helical stems that emanate from this G-quadruplex platform.

We have observed only *anti* glycosidic torsion angles for the guanine and adenine residues in the core (G5–G10 and G18–A23) of the AMP-27-mer DNA aptamer **1** complex as monitored by the magnitude of the NOE between the base and its own sugar H1′ protons at low mixing times [22]. The

Figure 7



(a) Overlap geometry for the (G8–G9–A10)·(G18-G19-A20) segment plus bound AMP_I in a representative distance-refined structure of the AMP–27-mer DNA aptamer **1** complex. Note the cross-strand stacking between G9 and G18 and between the purine ring of AMP_I and the base and sugar rings of A10. **(b)** Pairing alignments for the sheared G18·A10 mismatch pair, the G9·AMP_I mismatch pair involved in ligand recognition and the reversed Hoogsteen G8·G19 mismatch pair. **(c)** Views of the G19–A20 and G6–G7 base platforms looking down the helix axis.

absence of *syn* glycosidic torsion angles rules out formation of a G-quadruplex with antiparallel alignments of adjacent strands as proposed in the Huizenga–Szostak model [16] for the solution structure of the ATP-binding DNA aptamer complex, because this would require alternating G(*syn*)-G(*anti*) alignments along individual strands and alternating G(*syn*)·G(*anti*)·G(*syn*)·G(*anti*) alignments around individual G-tetrads [23–26]. Further, A10 and A23 are not bulged residues as proposed in the Huizenga–Szostak model [16], but rather participate in sheared G·A mismatch formation as established in the present study of the AMP–27-mer DNA aptamer **1** complex.

The DNA aptamer complex contains two adjacently bound AMP molecules

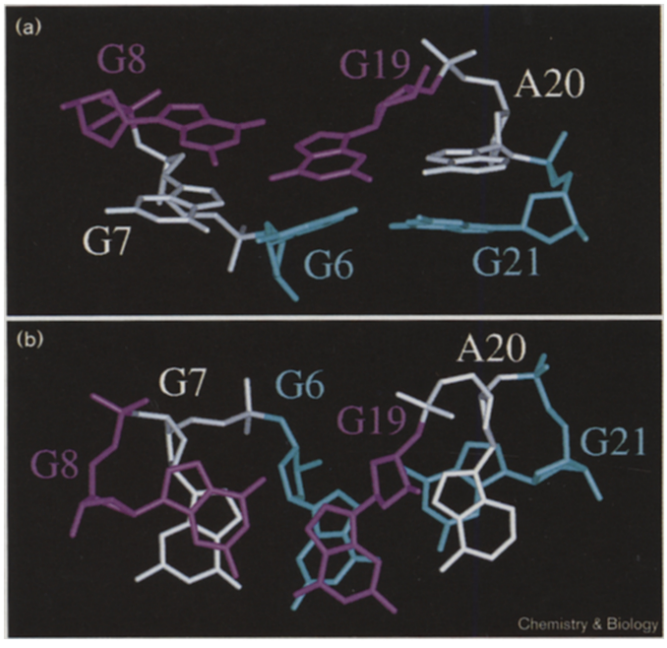
All studies of ligand–RNA aptamer and ligand–DNA aptamer complexes reported previously have either assumed or demonstrated one ligand bound per nucleic acid aptamer (reviewed in [5,6]). This stoichiometry has been confirmed from available structural studies on ligand–RNA aptamer and ligand–DNA aptamer complexes in solution (reviewed in [8]). Our identification of two AMP molecules bound to the DNA aptamer represents an unexpected and striking result that opens up new avenues to molecular recognition by nucleic acid binding sites.

The pair of AMP ligands must cooperatively target their binding sites on the DNA aptamer as our NMR experimental data found no evidence for complexes containing one bound AMP per DNA aptamer during the gradual addition of the cofactor to the nucleic acid (Figure 2).

The two bound AMP molecules are positioned in close spatial proximity to each other within a rectangular binding pocket in our NMR-based solution structure of the AMP–DNA aptamer **1** complex (Figures 6,9,10). We have been unable to directly define potential contacts between the bound AMPs in the complex, however. This is because the individual base and sugar protons of bound AMP_I and AMP_{II} have only slightly different chemical shifts precluding the unambiguous identification of potential NOEs between the same proton type for the AMP molecules bound at sites I and II in the complex. Further, the phosphates of the bound AMPs are poorly defined among the refined structures of the complex, precluding the identification of potential sugar–phosphate interactions between bound AMP molecules in the complex.

We address here the issue of the unexpected stoichiometry of two AMPs bound per DNA aptamer at the millimolar DNA concentrations used in our NMR studies. The selection experiments of Huizenga and Szostak [16] were undertaken on 1–3 mM ATP covalently attached through its C^8 position using a nine-atom linker to an agarose column. It is conceivable, given the long length of the linker, that adenine rings of adjacent covalently bound ATP molecules were brought into proximity on the agarose column through stacking interactions of their purine rings. Such alignments could conceivably facilitate the selection of DNA aptamer sequences that target a pair of immobilized ATP molecules on the agarose column. The bound DNA aptamers were subsequently eluted from the agarose column with ATP-containing buffer [16].

Figure 8



(a) A view normal to the helix axis and looking into the minor groove binding site defining the stacking between adjacent (G8·G19)·A20 and (G21·G6)·G7 base triads containing G19–A20 and G6–G7 base platforms in a representative distance-refined structure of the AMP–27-mer DNA aptamer **1** complex. The G8·G19 mismatch is shown in magenta, the G6·G21 mismatch is shown in cyan and the G7 and A20 bases involved in platform formation are shown in white. **(b)** Base overlaps between adjacent (G8·G19)·A20 and (G21·G6)·G7 base triads. Note the cross-strand stacking between the purine rings of G6 and G19 on partner strands.

Quality of the refined structures

We determined the same family of solution structures of the AMP–DNA aptamer **1** complex (as defined by r.m.s.d. values for both the entire and core complexes, Table 2), independently of whether experimentally identified non-canonical base pairs were included (protocol 1) or excluded (protocol 2) during the distance-restrained molecular dynamics calculations. The number of > 0.2 Å violations were 11.1 ± 1.4 (protocol 1) and 12.6 ± 2.2 (protocol 2) for the central core (G5–A10 and G18—A23 plus two bound AMPs) of the complex (Table 2). The majority of these violations are centered about the G7 and A20 residues, which participate in base platform formation (Figure 8) and correspond to the least well-defined segment of the complex on the basis of the available distribution of restraints to guide the computations.

Adaptive binding

We observe broad imino proton NMR spectra for residues in the asymmetrical internal bubble (Figure 1a) of the free 27-mer DNA aptamer **1** and the free self-complementary 14-mer DNA aptamer **2** duplex (Figure 2b). In contrast, narrow imino proton NMR spectra are observed for residues in the same segment of the AMP–27-mer DNA aptamer **1** complex (Figure 3a) and the AMP-self-complementary 14-mer DNA aptamer **2** complex (Figure 2b). Our results establish that the unstructured asymmetric internal bubble of the free AMP-binding DNA aptamer zippers up through base mismatch and three-base platform alignment on complex formation. Such adaptive binding is a recurring feature in previous reports of other RNA and DNA aptamer complexes (reviewed in [8]).

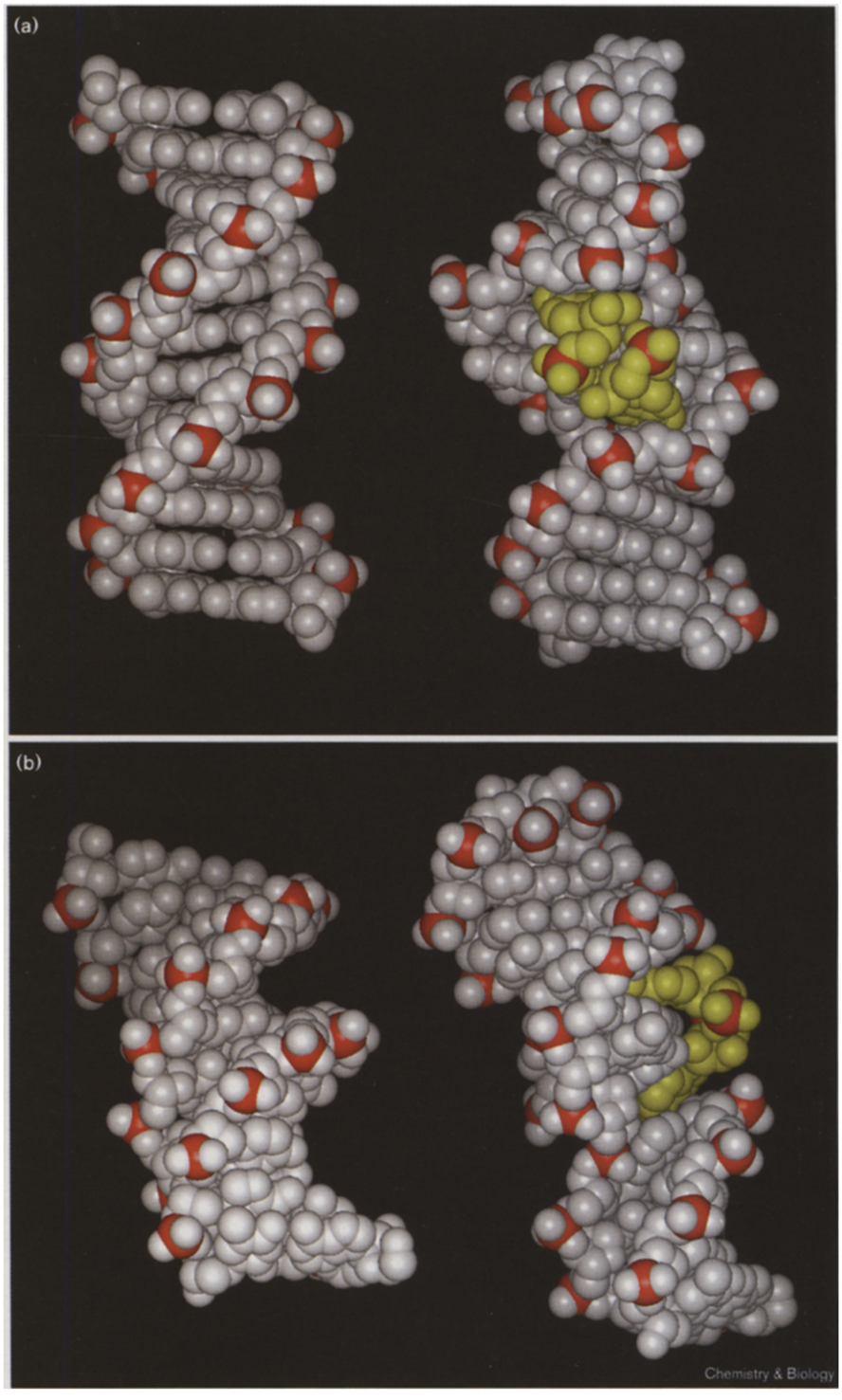
The present study of the binding of a pair of AMPs to an asymmetric internal bubble-containing DNA aptamer through mismatch and three-base platform zippering up of the helix has striking parallels with our previous demonstration of the binding of a single molecule of the cofactor flavin mononucleotide (FMN) to an asymmetric internal bubble-containing RNA aptamer through mismatch and base-triple zippering up of the helix [27]. The AMP molecules pair with guanines and are intercalated into the zippered-up mismatch-containing helix of the DNA aptamer complex reported here, while the isoalloxazine ring of the bound FMN pairs with an adenine and is intercalated into the zippered-up mismatch-containing helix of the RNA aptamer complex described in [27].

Mismatch-lined binding pocket

Six consecutive base mismatches, two of which also participate in three-base platforms, constitute the binding pocket of the AMP–27-mer DNA aptamer **1** complex (Figure 6b). The outermost mismatches (G5·A23 and G18·A10) are of the sheared G·A type, which involve pairing of the minor groove of the guanine and the major groove of the adenine (Figure 7b, top). Such sheared G·A base pairs were initially postulated from computational approaches [28] and subsequently identified experimentally in DNA [29] and RNA [30]. Such sheared G·A mismatches tend to be involved in extensive cross-strand stacking with flanking pairs and hence it is not surprising that extensive cross-strand stacking of G9 and G18 and of A10 and AMP_I has been observed between the adjacent sheared G18·A10 and recognition G9·AMP_I mismatches for binding site I in the AMP–27-mer DNA aptamer **1** complex (Figure 7a). Indeed, the stacking is manifested in the large upfield shifts observed for the base and sugar protons of A10 (H8: 7.42 ppm, H1′: 4.85 ppm, H2′: 1.06 ppm, H4′: 3.16 ppm) and A23 (H8: 7.29 ppm, H1′: 4.78 ppm, H2′: 1.08 ppm, H4′: 3.02 ppm) on complex formation, because of their position directly over the purine rings of the bound AMPs (Figure 7a).

The innermost mismatches (G6·G21 and G19·G8) are of the reversed Hoogsteen G·G type involving pairing of the Watson–Crick edge of one guanine (G8 and G21) and the Hoogsteen edge of its partner guanine (G19 and G6) (Figure 7b, bottom). Such reversed Hoogsteen G·G mismatches have been previously postulated [31] and identified [32] within the Rev binding site on Rev response

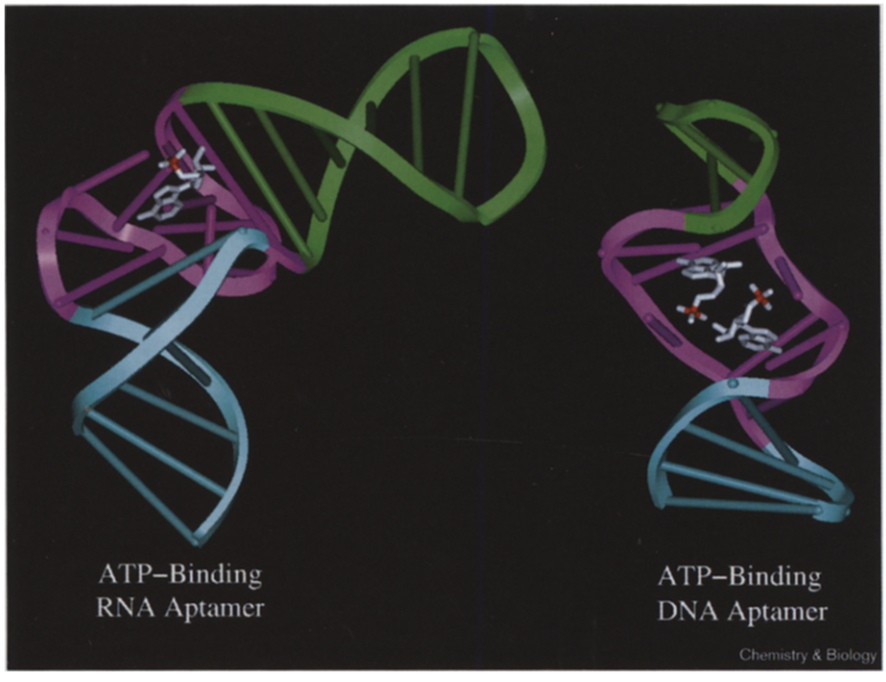
Figure 9



(a) Space-filling views comparing B-form DNA (left) and one representative distance-refined structure of the AMP–27-mer DNA aptamer **1** complex (right). The DNA is in white except for the backbone phosphorus atoms, which are in red. The two bound AMP molecules are in yellow except for their phosphorus atoms, which are in red. Note the expanded and rectangular-shaped minor groove with a segment of the backbone aligned normal to the helix axis in the complex (right). **(b)** Views rotated by approximately 90° along the helical axis relative to those shown in (a). Note the increased separation between the walls of the minor groove in the complex (right).

element (RRE) RNA [32] and in AMP–RNA aptamer [17,19] and arginine/citrulline–RNA aptamer [33] complexes. An important consideration relates to the incorporation of reversed mismatches whose glycosidic bonds are oriented in opposite directions (associated with parallel-stranded duplexes) adjacent to regular mismatches whose glycosidic bonds are oriented in the same direction (associated with antiparallel-stranded duplexes) in a DNA helix. This can be achieved by the looping out of intervening bases, as has been reported in the Rev peptide–RRE RNA complex [32,34] or through looping out of adjacent bases (G7 and A20) to generate three-base

Figure 10



Ribbon representations of the solution structures of the ATP-binding RNA aptamer complex (left; one equivalent of bound AMP) [17] and of the ATP-binding 27-mer DNA aptamer **1** complex (right; two equivalents of bound AMP). The backbones are shown as ribbons and the bases and base pairs are shown as cylindrical segments. The AMP-binding pockets on the nucleic acids are shown in pink with flanking stem segments shown in green and blue. The bound AMP molecules in a stick representation are shown in white with their phosphorus atoms in red.

platforms as observed here for the AMP–27-mer DNA aptamer **1** complex.

The recognition mismatches (G9·A_I and G22·A_{II}) involve a novel alignment that pairs the minor groove of guanine with the Watson–Crick edge of adenine (Figure 7b, center). This G·A mismatch alignment differs from the sheared G·A mismatch pair [29,30] in that the minor groove of the guanine pairs with the Watson–Crick edge of the adenine in the former alignment (Figure 7b, center) whereas it pairs with the Hoogsteen edge of the adenine in the latter alignment (Figure 7b, top). The same recognition G·AMP alignment has been observed in the corresponding AMP–RNA aptamer complex [17,19].

Binding-pocket architecture

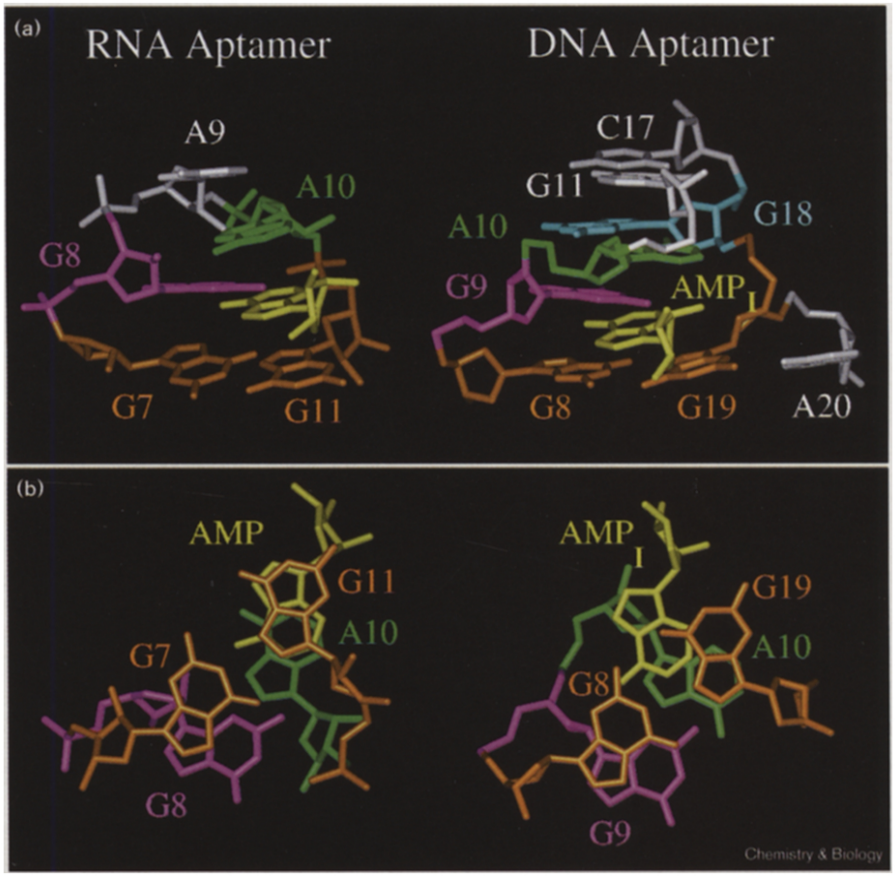
The concept of dimeric ligands targeting the minor groove of DNA has been previously reported for complexes of distamycin dimers bound to dA·dT-rich sites [35] and divalent cation-coordinated chromomycin dimers bound to dG·dC-rich sites [36] on duplex DNA. The widening of the minor groove in these complexes is achieved without distorting the sugar–phosphate backbone at the binding site, which remains inclined by ≈ 45° relative to the helix axis. By contrast, segments of the sugar–phosphate backbone are normal to the helix axis in the AMP-27-mer DNA aptamer **1** complex, generating an unprecedented rectangular widened minor groove (Figure 9a, right), in stark contrast to its counterpart in B-form DNA (Figure 9a, left). The separation between the walls of the minor groove at the binding site for bound AMPs essentially doubles in the AMP-27-mer DNA aptamer **1** complex (Figure 9b, right) relative to the separation in B-form DNA (Figure 9b, left).

The two intercalated AMP ligands are separated by a pair of stacked three-base platforms in our solution structure of the AMP–27-mer DNA aptamer **1** complex (Figure 6b). A somewhat parallel result has been reported for the binding of two actinomycin D molecules at adjacent overlapping G–C binding sites in the d(G–C–G–C)·d(G–C–G–C) sequence context where the intercalated phenoxazone chromophores are separated by two base pairs [37,38]. For both complexes, there is unwinding of the helix associated with complex formation and a bending of the helix towards the major groove, resulting in an opening of the minor groove to relieve potential steric clashes and accommodate the adjacently bound ligand molecules.

The concept of base platforms, where adjacent bases are positioned in the same plane rather than stacked on top of each other, emerged initially from computational models of triad DNA [39]. Such models proposed that triplet repeat disease sequences could adopt triad DNA structures where antiparallel duplexes were constructed of base triads rather than base pairs. The X-ray structure of the P4–P6 domain of the *Tetrahymena* group I ribozyme provided the first experimental demonstration of base platforms at three A–A steps [40] and this was followed by NMR identification of a triad containing an A–A platform

Figure 11

A comparison of the AMP-binding site in the AMP–40-mer RNA aptamer complex [17] with the AMP_I binding site in the AMP–27-mer DNA aptamer **1** complex. **(a)** Views normal to the helix axis are shown for the G7–G8–A9–A10–G11 segment plus AMP of the RNA aptamer complex [17] (left) and for the (G8–G9–A10–G11)·(C17–G18–G19–A20) segment plus AMP_I of the DNA aptamer complex (right). **(b)** Views down the helix axis are shown for the G7–G8 and A10–G11 segments plus AMP of the RNA aptamer complex [17] (left) and for the G8–G9–A10 and G19 segments plus AMP_I of the DNA aptamer complex (right).



in the RNA-binding site for the S8 protein [41] and a triad containing a T–A platform in the DNA quadruplex formed by the single-repeat *Bombyx mori* telomeric sequence [42]. Our study presents additional examples of base platforms at G6–G7 and G19–A20 steps (Figure 7c) in the AMP–27-mer DNA aptamer **1** complex. It should be noted that there is an important distinction between base triads, where the three bases are positioned in a triangular alignment [39,41,42], and the three-base (G8·G19)·A20 and (G21·G6)·G7 platforms in the AMP–27-mer DNA aptamer **1** complex, where the three bases are positioned in a linear alignment (Figure 8a,b).

The adjacent three-base (G8·G19)·A20 and (G21·G6)·G7 platforms contribute two critical structural components to the architecture of the binding pocket. First, the linear array of the three bases in the individual platforms results in a dramatic widening of the minor groove and, together with the sheared G·A mismatch alignments, contributes to the rectangular shape of the expanded minor groove binding pocket (Figure 9a, right panel). Second, they contribute to the ≈ 4 Å displacement in the helix axis, which is centered between the adjacently stacked (G8·G19)·A20 and (G21·G6)·G7 platforms in the AMP–27-mer DNA aptamer **1** complex (Figure 6).

Uniqueness of the AMP–DNA aptamer fold

We have reported here the structure of a new secondary and tertiary fold for the AMP–DNA aptamer complex containing two cofactor molecules bound per DNA aptamer **1**. On the basis of the published selection strategy [16], it is conceivable that AMP could also target other DNA folds with comparable micromolar affinity. Huizenga and Szostak [16] identified the ATP-binding DNA aptamer **1** investigated in this paper by using a two-step selection protocol. They identified 17 clones after eight rounds of selection, all of which had different sequences and appeared to be of independent origin. One of these clones was chosen for a second round of selection, from which emerged DNA aptamer **1**. It is conceivable that the choice of other clones for second-round selection could have resulted in the identification of AMP-binding DNA aptamers with different stoichiometries and/or tertiary folds.

Distinct tertiary structure architectures for the AMP–RNA and AMP–DNA aptamer complexes

The ATP-binding RNA [15] and DNA [16] aptamers have very different binding-site sequences and secondary structure folds. In addition, we have established here that the ATP-binding RNA and DNA aptamers also exhibit

different ligand-binding stoichiometries. It would therefore be reasonable to anticipate that the ATP-binding RNA and DNA aptamer complexes will exhibit distinct tertiary structure architectures; this has now been verified through the determination of the structures of the complexes in solution (Figure 10).

The ATP-binding RNA aptamer binds one equivalent of AMP within an S-shaped pocket formed by its asymmetric internal loop on complex formation (Figure 10, left) [17,19]. The AMP is integrated through intercalation in an extremely stable GNRA-like hairpin fold (where N is any nucleotide and R is a purine), with the ligand aligned through AMP·G mismatch formation. The flanking helical stems are aligned at an angle of $\approx 110°$, resulting in an L-shaped structure for the AMP–RNA aptamer complex.

By contrast, the ATP-binding DNA aptamer binds two equivalents of AMP within a continuous six-mismatch binding pocket adopted by its asymmetric internal bubble on complex formation (Figure 10, right). The AMPs intercalate into a rectangular minor groove binding pocket adopted by the zippered-up mismatch-containing helix, with the ligands aligned through AMP·G mismatch formation. The flanking helical stems are slightly bent in the AMP–DNA aptamer complex.

Similarities in the recognition alignments and ligand binding pockets of AMP–RNA and AMP–DNA aptamer complexes

Despite the distinct architectures of the solution structures of the AMP–RNA aptamer [17,19] and AMP–DNA aptamer complexes (Figure 10) and their distinct stoichiometries, there are striking common features associated with the three-dimensional topology of the binding pocket and the recognition alignment. Thus, for both AMP–RNA aptamer [17,19] and AMP–DNA aptamer complexes, the recognition alignment involves AMP·G mismatch formation through a pair of hydrogen bonds between the Watson–Crick edge of AMP and the minor groove edge of guanine (Figure 11a). Further, the AMP·G mismatch alignment stacks over a reversed Hoogsteen G·G mismatch in one direction in the structures of the AMP–RNA aptamer (Figure 11a, left) and AMP–DNA aptamer (Figure 11a, right) complexes. The AMP·G alignment stacks with an adenine in the opposite direction in both complexes, though the alignment of this adenine is distinct in AMP–RNA (Figure 11b, left) and AMP-DNA aptamer (Figure 11b, right) complexes.

The stacking patterns down the helix axis centered about the recognition AMP·G mismatch site demonstrate that the purine base of the intercalated AMP exhibits a greater overlap with flanking purines in the AMP–RNA aptamer complex (Figure 11b, left) than in the AMP–DNA aptamer complex (Figure 11b, right).

Significance

Structural analyses of nucleic acid aptamer complexes have already provided a wealth of information on nucleic acid architecture and molecular recognition. Such complexes generally involve nucleic acids up to 40-mers in length making them ideal candidates for structural characterization in solution by nuclear magnetic resonance (NMR). The major focus to date has been on RNA aptamer complexes, with only two examples of the structures of DNA aptamer complexes published to date. One of these is an X-ray structure of the α-thrombin–DNA aptamer complex [43] and the other is an NMR structure of the L-argininamide–DNA aptamer complex which involves adaptive binding [44]. The current structure of the AMP–DNA aptamer complex adds to this limited data set and provides the first opportunity of comparing the solution structures of complexes of RNA and DNA aptamers targeted to the same ligand. The solution structure of the AMP–DNA aptamer complex unexpectedly contains two equivalents of bound AMP positioned in an unprecedented expanded and rectangular minor groove binding site composed solely of base mismatches and three-base platforms with extensive cross-strand stacking. This complex is distinctly different from the structures of antitumor drug dimers bound to widened DNA minor grooves composed of Watson–Crick base pairs reported previously [35–38] and represents a new paradigm for molecular recognition involving minor groove targeting of DNA. A comparison with the previously solved solution structure of the AMP–RNA aptamer complex [17,19] establishes that the AMP–DNA and AMP–RNA aptamer complexes, with their different binding site sequences, secondary structure folds and ligand stoichiometries, adopt distinct architectures, but, surprisingly, have identical recognition pairing alignments and common structural elements that define their binding pockets.

Materials and methods

Sample preparation

Unlabeled AMP (Sigma) was used without further purification. Uniformly $^{13}C,^{15}N$-labeled AMP (Cambridge Isotopes) was used without further purification. The 27-mer DNA aptamer 1 (Figure 1) and 14-mer DNA aptamer 2 (Figure 2a) sequences were synthesized on a 10 μM scale on an Applied Biosystems 392 DNA synthesizer using solid phase β-cyanophosphoramidite chemistry and purified by reversed-phase HPLC. The AMP was added gradually to the DNA aptamers and complex formation monitored by recording imino proton spectra at 4°C; separate resonances were observed for the free and bound DNA aptamers.

NMR data collection and processing

NMR spectra of the AMP–DNA aptamer complexes were collected in aqueous buffer (50 mM KCl, 10 mM phosphate pH 6.3) at 4°C (exchangeable protons) and 10°C (nonexchangeable protons) on Varian 600 MHz Unity INOVA NMR spectrometers. Two-dimensional data sets included NOESY, COSY, TOCSY and ROESY homonuclear experiments on the unlabeled AMP-containing complexes and $^{1}H,^{13}C$-filtered and $^{1}H,^{15}N$-filtered NOESY heteronuclear experiments on uniformly

^{13}C,^{15}N-labeled AMP-containing complexes. Data sets were processed using Varian VNMR version 5.3 software and analyzed using the FELIX program (Molecular Simulations Inc.).

Nonexchangeable interproton distance restraints were obtained from the buildup of NOE cross-peak volumes in NOESY data sets (50, 90, 150 and 200 ms mixing times) on the AMP–27-mer DNA aptamer 1 complex in D_2O buffer at 10°C and bounds were set between ± 10% and ± 20% of the calculated distances using the fixed cytidine H5–H6 reference distance of 2.45 Å. Non-stereospecific assignments were treated with r^{-6} averaging. Interproton distance restraints involving exchangeable protons of the same complex were obtained from NOESY spectra (90 and 120 ms mixing times) in H_2O buffer at 4°C with bounds set between ± 15% and ± 20% of the calculated distance using the thymine imino to adenine H2 reference distance of 2.91 Å across an A·T base pair.

Distance geometry and molecular dynamics calculations

A set of 300 initial structures of the AMP-27-mer DNA aptamer 1 complex were generated by the X-PLOR [45] based metric matrix distance geometry (DG) protocol guided by the available distance restraints. Hydrogen-bonding distance restraints were imposed to align the experimentally identified stem Watson–Crick pairs (± 0.10 Å) and stem wobble T12·G16 pair (± 0.15 Å) during the DG calculations. These structures were quantitatively scored to select 21 structures with the least NOE violations, acceptable covalent geometry, and favorable van der Waals energy.

The X-PLOR-based [45] restrained molecular dynamics (MD) calculations were carried out in two cycles using the simulated annealing protocol and the CHARMM force field with reduced phosphate charges. Additional hydrogen-bonding distance restraints (± 0.15 Å) were imposed at this stage in protocol 1 to align the experimentally identified sheared G5·A23 and G18·A10 mismatch pairs, G9·AMP, and G22·AMP recognition pairs and reversed Hoogsteen G8·G19 and G6·G21 mismatch pairs within the core binding region of the complex. Each cycle of restrained MD simulations were initially carried out at 300K with a force constant of 1 kcal $mol^{-1}Å^{-2}$ on all experimentally obtained distance restraints. The structure was subjected to 500 cycles of energy minimization and was slowly heated to 1000K in 5 ps (0.5 ps per 50K increase). The force constants of the experimentally obtained distance restraints were slowly scaled up to 32 (nonexchangeable protons) and 16 (exchangeable protons) kcal $mol^{-1}Å^{-2}$, over a period of 10 ps. The system was allowed to evolve for another 10 ps at 1000K and next cooled gradually to 300K over 5 ps (0.5 fs time step) with retention of the full scale of distance restraints and subsequently equilibrated for 12 ps at 300K. The coordinates were averaged over the last 5 ps and the resulting coordinates subjected to 20 cycles (100 steps each cycle) of conjugate gradient energy minimization. Seven final distance-refined structures were selected on the basis of the criterion of low restraints violation and low total energy.

A second set of computations were undertaken that avoided the potential biasing of the resulting refinements through introduction of restraints involving experimentally defined mismatch pairs at the beginning of the distance-restrained MD calculations on the complex. These computations (designated protocol 2) did not introduce hydrogen-bonding distance restraints involving the mismatch pairs during any stage of the distance-restrained MD calculations.

Distance-restrained MD protocols 1 and 2 independently yielded the same family of refined structures exhibiting similar r.m.s.d. values (Table 2) for the AMP–27-mer DNA aptamer 1 complex.

Coordinates deposition

The coordinates of the AMP–27-mer DNA aptamer 1 complex have been checked for correct chirality and deposited (Accession number: 1aw4) in the Protein Data Bank.

Supplementary material

Supplementary material available with the on-line version of this paper includes one table (Table S1) of proton chemical shifts of the AMP–27-mer DNA aptamer complex, one figure (Figure S1) showing the proposed Huizenga–Szostak secondary fold of the ATP-binding DNA aptamer [16], one stereo figure (Figure S2) of seven superpositioned distance-refined structures of the core of this complex using protocol 2 and three figures (Figures S3–S6) outlining intermolecular NOEs in the AMP–27-mer DNA aptamer complex in H_2O and D_2O solution.

Acknowledgements

This research was supported by start up funds to D.J.P. We thank John Hubbard for help with the preparation of the color graphics.

References

1. Ellington, A.D. & Szostak, J.W. (1990). *In vitro* selection of RNA molecules that bind specific ligands. *Nature* **346**, 818-822.
2. Tuerk, C. & Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505-510.
3. Robertson, D.L. & Joyce, G.F. (1990). Selection *in vitro* of an RNA enzyme that specifically cleaves single-stranded DNA. *Nature* **344**, 467-468.
4. Joyce, G.F. (1994). *In vitro* evolution of nucleic acids. *Curr. Opin. Struct. Biol.* **4**, 331-336.
5. Gold, L., Polisky, B., Uhlenbeck, O.C. & Yarus, M. (1995). Diversity of oligonucleotide functions. *Annu. Rev. Biochem.* **64**, 763-797.
6. Lorsch, J.R. & Szostak, J.W. (1996). Chance and necessity in the selection of nucleic acid catalysts. *Accounts Chem. Res.* **29**, 103-110.
7. Breaker, R.R. (1997). DNA aptamers and DNA enzymes. *Curr. Opin. Chem. Biol.* **1**, 26-31.
8. Patel, D.J. (1997). Structural analysis of nucleic acid aptamers. *Curr. Opin. Chem. Biol.* **1**, 32-46.
9. Lorsch, J.R. & Szostak, J.W. (1996). *In vitro* evolution of new ribozymes with polynucleotide kinase activity. *Nature* **371**, 31-36.
10. Cuenoud, B. & Szostak, J.W. (1995). A DNA metalloenzyme with DNA ligase activity. *Nature* **375**, 611-614.
11. Dai, X.C., de Mesmaeker, A. & Joyce, G.F. (1995). Cleavage of an amide bond by a ribozyme. *Science* **267**, 237-240.
12. Lohse, P.A. & Szostak, J.W. (1996). Ribozyme catalyzed amino acid transfer reactions. *Nature* **381**, 442-444.
13. Santoro, S.W. & Joyce, G.F. (1997). A general purpose RNA-cleaving DNA enzyme. *Proc. Natl Acad. Sci USA* **94**, 4262-4266.
14. Carmi, N., Shultz, L.A. & Breaker, R.R. (1996). *In vitro* selection of self-cleaving DNAs. *Chem. Biol.* **3**, 1039-1046.
15. Sassanfar, M. & Szostak, J.W. (1993). An RNA motif that binds AMP. *Nature* **382**, 183-186.
16. Huizenga, D.E. & Szostak, J.W. (1995). A DNA aptamer that binds adenosine and ATP. *Biochemistry* **34**, 656-665.
17. Jiang, F., Kumar, R.A., Jones, R.A. & Patel, D.J. (1996). Structural basis of RNA folding and recognition in an AMP-RNA aptamer complex. *Nature* **382**, 183-186.
18. Nonin, S., Jiang, F. & Patel, D.J. (1997). Imino proton exchange and base pair kinetics in the AMP–RNA aptamer complex. *J. Mol. Biol.* **268**, 359-374.
19. Dieckmann, T., Suzuki, E., Nakamura, G.K. & Feigon, J. (1996). Solution structure of an ATP-binding RNA aptamer reveals a novel fold. *RNA* **2**, 628-640.
20. Osborne, S.E., Matsumura, I. & Ellington, A.D. (1997). Aptamers as therapeutic and diagnostic reagents: problems and prospects. *Curr. Opin. Chem. Biol.* **1**, 5-9.
21. Osborne, S.E. & Ellington, A.D. (1997). Nucleic acid selection and the challenge of combinatorial chemistry. *Chem. Rev.* **97**, 349-370.
22. Patel, D.J., Kozlowski, S.A., Nordheim, A. & Rich, A. (1982). Right-handed and left-handed DNA: studies of B-DNA and Z-DNA by using proton nuclear Overhauser effect and phosphorus NMR. *Proc. Natl Acad. Sci. USA* **79**, 1413-1417.
23. Kang, Z., Zhang, X., Ratliff, R., Moyzis, R. & Rich, A. (1992). Crystal structure of four-stranded *Oxytricha* telomeric DNA. *Nature* **356**, 126-131.
24. Macaya, R.F., Schultze, P., Smith, F.W., Roe, J.A. & Feigon, J. (1993). Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc. Natl Acad. Sci. USA* **90**, 3745-3749.

25. Wang, K.Y., Krawczyk, S.H., Bischofberger, N., Swaminathan, S. & Bolton, P.H. (1993). The tertiary structure of a DNA aptamer which binds to and inhibits thrombin determines activity. *Biochemistry* **32**, 11285-11292.
26. Kettani, A., Kumar, R.A. & Patel, D.J. (1995). Solution structure of a DNA quadruplex containing the fragile X syndrome triplet repeat. *J. Mol. Biol.* **254**, 638-656.
27. Fan, P., Suri, A.K., Fiala, R., Live, D. & Patel, D.J. (1990). Molecular recognition in the FMN–RNA aptamer complex. *J. Mol. Biol.* **258**, 480-500.
28. Westhof, E., Romby, P., Romaniuk, P.J., Ebel, J.P., Ehresmann, C. & Ehresmann, B. (1989). Computer modeling from solution data of spinach chloroplast and of *Xenopus laevis* somatic and oocyte 5S RNAs. *J. Mol. Biol.* **207**, 417-431.
29. Li, Y., Zon, G. & Wilson, W.D. (1991). NMR and molecular evidence for a G•A mismatch base pair in a purine-rich DNA duplex. *Proc. Natl Acad. Sci. USA* **88**, 26-30.
30. Heus, H.A. & Pardi, A. (1991). Structural features that give rise to unusual stability of RNA hairpins containing GNRA loops. *Science* **253**, 191-194.
31. Bartel, D.P., Zapp, M.L., Green, M.R. & Szostak, J.W. (1991). HIV-1 REV regulation involves recognition of non-Watson–Crick base pairs in viral RNA. *Cell* **67**, 529-536.
32. Battiste, J.L., Tan, R., Frankel, A.D. & Williamson, J.R. (1994). Binding of an HIV Rev peptide to Rev response element RNA induces formation of purine–purine base pairs. *Biochemistry* **33**, 2741-2747.
33. Yang, Y., Kochoyan, M., Burgstaller, P., Westhof, E. & Famulok, M. (1996). Structural basis of ligand discrimination by two related RNA aptamers resolved by NMR spectroscopy. *Science* **272**, 1343-1347.
34. Ye, X., Gorin, A., Ellington, A.D. & Patel, D.J. (1996). Deep penetration of an α-helix into the widened RNA major groove in the HIV-1 Rev peptide–RNA aptamer complex. *Nat. Struct. Biol.* **3**, 1026-1033.
35. Pelton, J.G. & Wemmer, D.E. (1989). Structural characterization of a 2:1 distamycin A•d(CGCAAATTGGC) complex by two dimensional NMR. *Proc. Natl Acad. Sci. USA* **86**, 5723-5727.
36. Gao, X. & Patel, D.J. (1989). Solution structure of the chromomycin–DNA complex. *Biochemistry* **28**, 751-762.
37. Scott, E.V., Zon, G., Marzilli, L.G. & Wilson, W.D. (1988). 2D NMR investigation of the binding of the anticancer drug actinomycin D to duplexed dATGCGCAT: conformational features of the unique 2:1 adduct. *Biochemistry* **27**, 7940-7951.
38. Chen, H., Liu, X. & Patel, D.J. (1996). DNA bending and unwinding associated with actinomycin D antibiotics bound to partially overlapping sites on DNA. *J. Mol. Biol.* **258**, 457-459.
39. Kuryavyi, V.V. & Jovin, T.M. (1995). Triad DNA: a model for trinucleotide repeats. *Nat. Genet.* **9**, 339-341.
40. Cate, J.H., *et al.* & Doudna, J.A. (1996). RNA tertiary structure mediation by RNA platforms. *Science* **273**, 1696-1699.
41. Kalurachchi, K., Uma, K., Zimmermann, R.A. & Nikonowicz, E.P. (1997). Structural features of the binding site for ribosomal protein S8 in *Escherichia coli* 16S rRNA defined using NMR spectroscopy. *Proc. Natl Acad. Sci. USA* **94**, 2139-2144.
42. Kettani, A., Bouaziz, S., Wang, W., Jones, R.A. & Patel, D.J. (1997). *Bombyx mori* single repeat telomeric DNA sequence forms a G-quadruplex capped by base triads. *Nat. Struct. Biol.* **4**, 382-389.
43. Padmanabhan, K., Padmanabhan, K.P., Ferrara, J.D., Sadler, J.E. & Tulinsky, A. (1993). The structure of α-thrombin inhibited by a 15-mer single-stranded DNA aptamer. *J. Biol. Chem.* **268**, 17651-17654.
44. Lin, C.H. & Patel, D.J. (1996). Encapsulating an amino acid in a DNA fold. *Nat. Struct. Biol.* **3**, 1046-1050.
45. Brunger, A.T. (1992). *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR.* Yale University Press, New Haven, CT, USA.