Crystal structures of an A-form duplex with single-adenosine bulges and a conformational basis for site–specific RNA self-cleavage

Stefan Portmann[1,2], Susan Grimm[3], Christopher Workman[3], Nassim Usman[3] and Martin Egli[1]*

Background: Bulged nucleotides are common secondary structural motifs in RNA molecules and are often involved in RNA–RNA and RNA–protein interactions. RNA is selectively cleaved at bulge sites (when compared to other sites within stems) in the presence of divalent metal cations. The effects of bulge nucleotides on duplex stability and topology have been extensively investigated, but no detailed X-ray structures of bulge-containing RNA fragments have been available.

Results: We have crystallized a self-complementary RNA–DNA chimeric 11-nucleotide sequence containing single-adenosine bulges under two different conditions, giving two distinct crystal forms. In both lattices the adenosines are looped out, leaving the stacking interactions in the duplex virtually unaffected. The bulges cause the duplex to kink in both cases. In one of the structures, the conformation of the bulged nucleotide places its modeled 2'-oxygen in line with the adjacent phosphate on the 3' side, where it is poised for nucleophilic attack.

Conclusions: Single adenosine bulges cause a marked opening of the normally narrow RNA major groove in both crystal structures, rendering the bases more accessible to interacting molecules compared with an intact stem. The geometries around the looped-out adenosines are different in the two crystal forms, indicating that bulges can confer considerable local plasticity on the usually rigid RNA double helix. The results provide a conformational basis for the preferential, metal-assisted self-cleavage of RNA at bulged sites.

Addresses: [1]Department of Molecular Pharmacology & Biological Chemistry, Northwestern University Medical School, 303 East Chicago Avenue, Chicago, IL 60611-3008, USA, [2]Laboratory of Organic Chemistry, ETH – Swiss Federal Institute of Technology, Universitätstrasse 16, CH-8092 Zürich, Switzerland and [3]Department of Chemistry and Biochemistry, Ribozyme Pharmaceuticals Inc., 2950 Wilderness Place, Boulder, CO 80301, USA.

*Corresponding author.

Introduction

RNA can form a number of diverse secondary structural elements, including double helices, hairpin loops, single- and multiple-base bulges, internal loops and pseudoknots. This structural flexibility provides the framework for the numerous biological functions of RNA molecules [1]. Bases that lack a pairing partner on the opposite strand, known as bulges, are an abundant feature of postulated and established secondary structures [2]. Bulge nucleotides have been shown to be important for protein binding [3–6], and they are involved in the tertiary folding of large RNAs [7]. A bulged adenosine in the translational operator of the bacteriophage R17 replicase gene is essential for specific binding to the R17 coat protein [8,9]. An unpaired adenosine in helix II of *Escherichia coli* 5S RNA, which also occurs in homologous RNAs from other prokaryotes, is believed to be important for the binding of the ribosomal protein L18 to the helix II–helix III region of the RNA (see [10] and papers cited therein).

Functionally important single-adenosine bulges are found at key sites in self-splicing group I, group II and nuclear pre-mRNA introns [11–18]. Interhelical and loop regions were observed to have preferential affinity for

Pb^{2+} binding, and it was suggested that flexible and dynamic portions of RNA molecules are highly sensitive to Pb^{2+} cleavage [19]. This is consistent with the locations of the two major and a few of the minor Pb^{2+} cleavage sites in the catalytic-core structure of group I intron RNA [20]. Efficient cleavage occurs adjacent to a bulge nucleotide of the substrate-binding site in the P7 stem, indicating that Pb^{2+} can cleave at locations where divalent metal ions involved in catalysis bind. Oligonucleotides that pair with an RNA single strand, but leave one or two bases unpaired, were observed to cause RNA self-cleavage 3' to the bulge residue in the presence of a range of divalent metal cations, providing an attractive extension of the antisense strategy (D. Hüsken and colleagues, unpublished data). This cleavage is mediated by the 2'-OH of the bulge nucleotide ribose, and produces a 2',3'-cyclic phosphate and a free 5'-hydroxyl group.

Bulges generally destabilize RNA duplexes, and the extent of destabilization is often dependent on the length of the bulge loop [21]. Bulges formed by A$_n$ or U$_n$ (where n indicates the length of the bulge loop) have similar stabilities in RNA duplexes. The actual loss of stability

caused by a bulge depends on the flanking sequence, but non-nearest-neighbor interactions also affect the free energy increments in RNA duplexes [21,22]. In DNA duplexes, bulges located on an all-purine adenosine strand and thus between two purines were observed to be more stable than those on an all-pyrimidine thymidine strand, between two pyrimidines [23]. It was suggested that this may be due to the smaller cost of energy for disrupting the pyrimidine–pyrimidine stacking interactions compared with those for disrupting the purine–purine stacking interactions. In RNA duplexes, however, no differential effect was observed between bulge residues in all-purine or all-pyrimidine strands [21]. Single-purine bulges seem to destabilize a duplex to a somewhat greater extent than single-pyrimidine bulges, both in DNA [23] and RNA [24]. This would imply that bulges affect RNA A-form and DNA B-form duplexes in a similar fashion. Similarly, both the DNA duplex [25,26] and the RNA duplex (e.g. [27]) can be bent as a consequence of bulges. In contrast, stacking in the DNA B-form duplex tends to occur between bases of the same strand, whereas interstrand stacking is much more important in the RNA A-form duplex, suggesting that the conformational behavior of bulges in RNA and DNA duplexes may be different.

Relatively little is known about the detailed effects of bulges on RNA stem conformation and about the influences of secondary structure on specific protein–RNA and RNA–RNA interactions. Gel retardation experiments have revealed a pronounced kinking at bulge sites [27–30], and chemical probing [31] and NMR experiments on RNA in solution ([32–34] and references therein) suggest that bulges probably cause a widening of the major groove. The actual conformations of unpaired bases — intrahelical or extrahelical — are affected by the nature of the base, the flanking sequence and interactions with proteins. NMR structural analyses in solution of DNA duplexes containing single-base bulges have demonstrated that an unpaired purine nucleotide is generally stacked into the helix [35–44], whereas an unpaired pyrimidine nucleotide is either stacked into [45] or looped out of the helix [46,47], with the two states occasionally in a temperature-dependent equilibrium [47,48]. An unpaired uridine within an eight-base-pair RNA duplex was found to be extrahelical in solution [49]. In the crystal structure of the bacteriophage MS2 coat protein–RNA operator complex, an unpaired adenosine is looped out and interacts with the protein. In the solution structure of the unliganded RNA operator, however, this nucleotide is intercalated between the neighboring base pairs ([9] and literature cited therein). The only X-ray crystal structure of a bulge-containing oligonucleotide that has been determined previously, that of the 13-base deoxyoligonucleotide d(CGCAGAATTCGCG) (bulged nucleotide underlined), revealed that the adenosine of one strand is looped out of the helix whereas that on the opposite strand is stacked

into the helix [50]. The looped-out base is intercalated between a cytosine and a guanine base that are ~6.8 Å apart due to the insertion of the extra A residue in one strand, thus straightening the B-form double helix. The amount of bending of the duplex was ~20°, comparable to that observed in the native Dickerson–Drew dodecamer d(CGCGAATTCGCG). In the duplex, the stacking interactions between base pairs adjacent to the bulge site were only marginally affected.

To elucidate the effects of an unpaired nucleotide on the topology of an RNA duplex, we crystallized a short, bulged-oligonucleotide fragment. As single unpaired adenosines occur very frequently in secondary structures of biologically important RNA molecules, an adenosine was selected as the bulge residue. We report here the X-ray crystal structures of an A-form double helix, with a single unpaired adenosine per strand, in two distinct crystal lattices. The double helix is formed by a self-complementary, 11-nucleotide RNA–DNA chimera, containing five central deoxyribonucleotides, including the bulged adenosine, flanked by three ribonucleotides on either side (Fig. 1).

A chimeric strand was chosen for two reasons. First, in crystal structures of chimeras the RNA portion was shown to conformationally dominate the DNA. This occurred whether the duplex was an all-RNA/all-DNA hybrid [51], or consisted of chimeric RNA–DNA strands [52–54]. Thus, the overall conformation of such fragments in the crystal is usually of the RNA A-form. Second, RNA fragments containing a bulge are relatively easily cleaved in vitro when compared to other stem structures (but not when compared to random coil) at the phosphate group 3' to the bulge nucleotide in the presence of Mg^{2+} or other divalent metal cations (D. Hüsken and colleagues, unpublished data). Exclusion of 2'-hydroxyl groups at and around the bulged site prevents cleavage of the oligonucleotide under crystallization conditions that include metal cations.

The unpaired adenosine could in principle be accommodated at two sites within each strand (Fig. 1), providing a larger variety of possible interactions between duplexes and thus an enhanced chance for formation of a good crystal lattice. In either conformation, strand cleavage is prevented by the presence of the flanking deoxyribonucleotides. Crystals grown in the absence of Mg^{2+}, but in the presence of relatively high concentrations of spermine, diffracted to 2.8 Å resolution, and crystals grown in the presence of moderate amounts of Mg^{2+} diffracted to 1.83 Å resolution (Table 1). The two crystal forms, termed spermine form and Mg form, feature different packing modes. Although the bulged adenosine is looped out in both of them, the backbone conformation around the unpaired residue differs fundamentally between the two structures. Here, we present the structures of the two

Figure 1
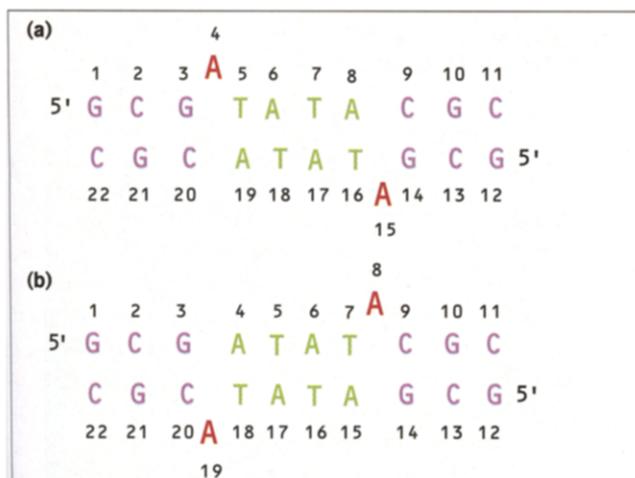
Two possible secondary structures of the bulge-containing (red) chimeric RNA (purple) – DNA (green) duplex and residue numbering. The bulged adenosine is wedged **(a)** between a purine and a pyrimidine, G3 and T5 (first strand), or **(b)** between two pyrimidines, T7 and C9 (first strand). In both crystal forms duplex (b) is observed.

bulge-containing duplex fragments, which reveal how the unpaired adenosines affect the topology of the A-form duplex. The structures also shed light on the conformational flexibility of bulge sites, and provide a structural basis for the observation that metal-assisted RNA strand cleavage occurs preferentially at bulge sites.

Results and discussion
The same loop site is seen in both crystal forms

We will use nucleotides 1 to 11 of one strand to describe the conformational features of the sugar-phosphate backbone and the bulge base. In both the spermine- and Mg-form crystals, the duplex strands are related by crystallographic two-fold symmetry. Of the two possible sites for the bulge base in the duplex, both crystal forms adopt the conformation with the bulge base between T7 and C9 (Fig. 1b). The preservation of the bulge site may simply reflect the higher thermodynamic stability of a bulge residing between the two pyrimidines T7 and C9 compared to that of a bulge between purine G3 and pyrimidine T5; the intrastrand stacking should be stronger in the latter case. Alternatively, adenosines looped out at positions 8 and 19 may result in interactions between duplexes that stabilize crystal lattices, shifting the equilibrium between the two possible conformations towards the one observed here.

Two arguments lead us to speculate that the observed bulged fragment is the thermodynamically more stable of the two. In RNA duplexes, single and multiple adenosine bulges flanked by two G residues were less stable than the corresponding bulges between two C residues [21]. Thus, the disruption or reduction of stacking interactions

between two purine bases presumably causes greater loss of stability than the changes in the interactions between the two pyrimidine bases. In terms of lattice interactions, the two conformations shown in Figure 1 may not be as different as they first appear. In three-dimensional space, the adenosine bulges would jut out on the opposite side of the duplex in Fig. 1a relative to the one in Fig. 1b. Thus, despite different relative orientations of bulges and grooves in the two cases, stacking interactions between terminal base pairs of adjacent duplexes in the lattice would be largely unaffected, supporting the notion that the bias in the sites of bulging is probably not due to mere lattice effects. The observed conformation also contains a purine–purine step (G3pA4), which provides stability through intrastrand stacking that is not present in the alternative conformation. The different sequences in the center of the two alternative structures (DNA portion) may account for further stability differences between them.

Alterations in duplex structure

The extrahelical orientations of the adenosine affect the duplex differently in the two crystal forms. The duplex portions around the bulge base are shown in Figure 2. In the spermine-form structure (Fig. 2a), the bulge causes a more severe distortion of the regular A-form sugar-phosphate backbone geometry (C3'-*endo* puckers) than in

Table 1

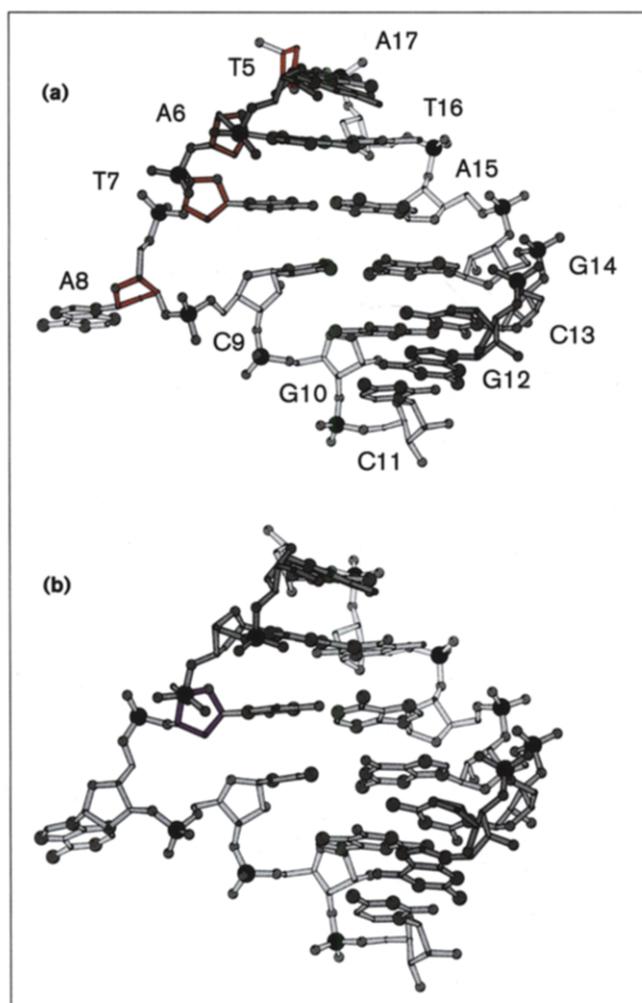**Crystal and refinement data for spermine- and Mg-form crystals.**

Parameter	Spermine form crystal	Mg-form crystal
Space group	Hexagonal P6$_5$22	Orthorhombic C222$_1$
Cell constants		
a (Å)	27.25	25.89
b (Å)	27.25	33.97
c (Å)	177.40	57.56
V/base pair (Å3)	1811	1205
Resolution (Å)	2.8	1.83
Reflections (F$_{obs}$ ≥ 2σ(F$_{obs}$)) (10 Å to full res.)	989	2170
Completeness (%)	82	94
R$_{merge}$ (%)	7.5	8.1
Asymmetric unit	Single-stranded 11-mer	Single-stranded 11-mer
DNA/RNA atoms	279	279
Water molecules	23	38
Spermine	–	2
Final R-factor (%)	18.7	15.9
R$_{free}$ (%)	–	23.6
RMS dev. bonds (Å)	0.015	0.013
RMS dev. angles (°)	1.95	3.28

the Mg-form structure. Beginning in the central portion of the spermine-form backbone, the sugar puckers are converted to C1'-*exo* (A6) and C2'-*endo* types (T7 and bulge A8). Conversely, the only sugar pucker in the Mg-form that deviates from the regular C3'-*endo* type observed in A-form duplexes is the one in residue T7 (C2'-*endo*), immediately preceding the bulge (Fig. 2b). This difference is also illustrated in Figure 3; the δ torsion angles in the spermine form deviate from the standard *sc*⁺ conformation for several nucleotides. The puckering modes observed here around the bulge probably represent common ways for an RNA backbone to react in a non-standard duplex situation, rather than a

Figure 2

The sugar-phosphate backbone geometries and the base stacking interactions are different in the spermine-form and Mg-form duplexes. Major groove views of the bottom half of the duplex, which is related to the rest via a crystallographic two-fold rotation axis between base pairs T5•A17 and A6•T16, are shown. **(a)** The spermine-form duplex, with sugars adopting C1'-*exo* (A6) or C2'-*endo* (T7 and A8) puckers, highlighted in red. **(b)** The Mg-form duplex, with the sugar of residue T7 adopting a C2'-*endo* conformation, highlighted in blue. Darker portions of the duplex are closer to the reader.

peculiar geometrical feature of the RNA–DNA chimeric duplex. C2'-*endo* type sugar puckers have also been observed in and around bulge and hairpin loops in other RNA molecules (e.g. [32,55]).

The backbone conformation of the other strand in the region opposite the bulge is also different in the two crystal forms. In the spermine form, residue G4 adopts an extended backbone conformation, with α and γ in the *ap* range, accompanied by subtle alterations in angles β and ε (Fig. 3) [56,57]. In contrast, the bulge adenosine does not affect the backbone conformation of the opposite strand in the Mg form. These differences are probably due to alternative adjustments to the base-stacking interactions at the bulge site in the two duplexes (see below). In both forms, the conformation of the backbone portion 3' to the bulge appears more regular (Fig. 3).

In addition to the distinct geometries of the backbones, several characteristic differences in the base-stacking interactions occur between the two crystal forms. To simplify, one might refer to the duplex regions around the bulge in the spermine- and Mg-form duplexes as the interstrand and intrastrand stacking models, respectively, as illustrated in Figure 2. In the spermine-form duplex, the slide (relative shift of base pairs along their long dimension) between base pairs T7•A15 and C9•G14, and the one between the latter base pair and G10•C13 are regular (Fig. 2a). This generates overlap between bases from opposite strands, a common feature in A-RNA. In the duplex region around the bulge in the spermine form, it appears that the bulge adenosine is pulling base pair T7•A15 out of the stack. By comparison, there is much more intrastrand stacking around the bulge site in the Mg-form duplex, where the two base pair steps preceding the adenosine bulge have considerably reduced sliding (by ~1 Å). This results in increased intrastrand stacking, which is visible in Figure 2b for the bases of purines A15 and G14. This increased intrastrand stacking may make an important contribution to the presumably higher stability of the A8/A19 bulge duplex relative to the alternative A4/A15 bulge (Fig. 1).

The different ways to maximize stacking between the base pairs flanking the bulge are also evident from the varying helical twists between base pairs T7•A15 and C9•G14 in the two duplexes. In the spermine form, the larger slide is compensated by a smaller twist (26°), whereas the more B-DNA-like slide value in the Mg-form is accompanied by a regular twist of 33°. In both duplexes, the greatest rise occurs at the A6pT7 step (3.3 Å), and the helical rise is generally increased in the central part of both duplexes. Thus, the two duplexes have the local conformational plasticity of an A-form duplex and represent two alternative ways to accommodate a bulged residue.

Figure 3

The backbone torsion angles in the spermine-form (red) and Mg-form (blue) duplexes demonstrate the differences in the sugar-phosphate backbone geometry. The bulge adenosine is located at position 8 in both cases.
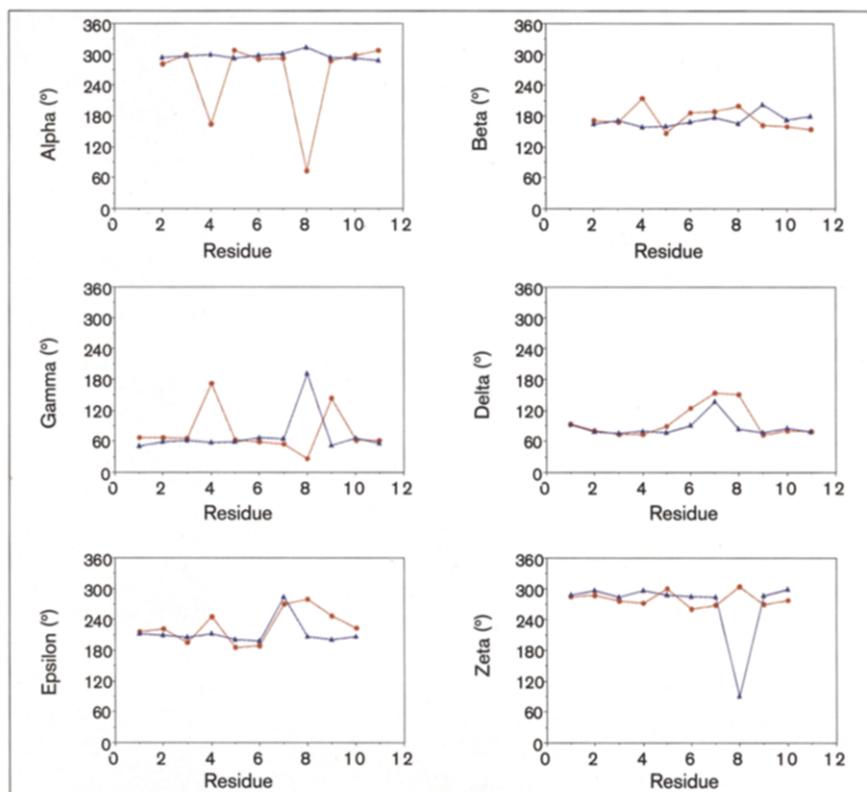
Two different bulge conformations

Except for torsion angle β, which lies in the *ap* range in both cases, the conformations of the unpaired adenosine are entirely different in the two crystal forms. As noted above, the bulge affects the geometry of the adjacent duplex region in the spermine form, but is incorporated into the Mg-form duplex with virtually no alteration in the neighboring backbone region (Fig. 3, e.g. angles γ and ϵ on the 3'-side). All other backbone torsion angles of the bulge adenosine adopt different conformations in the two structures. Not only are the sugar puckers of that nucleotide different, but its glycosidic torsion is in the *syn* range in the Mg form and is in the *anti* range in the spermine form. In the latter crystal form, the non-standard RNA puckering of the residues preceding the bulge results in an orientation of the bulge nucleotide deoxyribose that is roughly perpendicular to that in the Mg-form (Fig. 2). Thus, relative to the sugars of the nucleotides flanking the bulge adenosine, its own sugar is inserted into the oligonucleotide strand with more or less opposite polarity in the spermine form. By comparison, the sugar orientation of the adenosine in the Mg-form follows the orientation of the sugar moieties in the rest of the backbone in an almost seamless manner.

The conformation of the sugar-phosphate backbone at and around the bulge site in the spermine-form structure is remarkably similar to that observed for a uridine bulge (flanked by G residues) in an RNA duplex, based on a combination of model-building and NMR solution structures [49]. None of the torsion angles deviates by more than 25° between the two structures. In the NMR model the uridine is also looped out and its sugar and the one from the 5'-adjacent G have C2'-*endo* puckers. The structure trapped in the spermine-form lattice, with certain residues adopting B-DNA type sugar puckers, thus seems to be representative of the structures adopted by RNA duplexes containing an unpaired nucleotide, not a structure imposed upon the RNA by the central DNA portion.

Bulged duplexes are kinked with enlarged major grooves

In the crystallized fragments, the adenosine bulges are located on opposite strands, separated by four base pairs. This places them on the same side of the duplex, approximately above one another (Fig. 4). The base planes are protruding from the backbone nearly perpendicular to the axis of the helix, completely withdrawn from the grooves. The phosphate group of residue C9, adjacent to the adenosine bulge on the 3' side, is turned inward to the major groove in both crystal forms (Figs 2,4). Beginning at nucleotide A6 and including its phosphate group, the backbone to the 5'-side of the bulge in the spermine form is rolled into the minor groove,

resulting in a pronounced narrowing of that groove (Figs 4b, 5). This contraction of the minor groove between the two looped-out A nucleotides is accompanied by kinks of the helix between base pairs T7•A15 and C9•G14 (G3•C20 and A4•T18 at the opposite end) (Fig. 4a). The kink per adenosine bulge in both structures is about 10°; because the bulges are on the same side of the duplex, they bend the fragment along roughly the same trajectory and thus by a total of about 20°. This is evident from the almost straight appearance of the extended duplex in Figure 4b, and the prominent kink at the center between the bulges after rotating the duplex by 90° (Fig. 4a).

The amount of kinking per adenosine in these structures compares well to the findings from a number of other techniques. Using fluorescence resonance energy transfer (FRET), the kinking in DNA and RNA duplexes has been shown to be of similar magnitude and of the order of 90° for a bulge containing seven A nucleotides (A_7) [58]. Because the axial kinking has a roughly linear dependence on the bulge size, the estimated kink per unpaired nucleotide is ~15°. Similarly, the kinking per adenosine bulge of an RNA duplex, based on transient electric birefringence, was estimated to be ~16° [59], and the value of 85° observed for the kink in a 100-base-pair DNA fragment containing a central A_7 bulge [27] gives a kink of ~12° for a single base bulge, assuming that the relationship between the bulge length and the degree of kink is more or less linear.

An important consequence of the kink in the backbone adjacent to the looped-out bases is the marked opening of the major groove in that region (Figs 4,5). In the spermine-form duplex, the groove is widened to about 8 Å on average, or roughly twice its normal width, for almost a full turn between the bulges. Consistent with the more subtle conformational alterations caused by the bulge in the Mg-form duplex, the major groove is somewhat less enlarged (Fig. 5). Because the major groove is enlarged, the edges of bases lie completely exposed and are readily accessible to interacting molecules. Opening of the major groove is strongly asymmetric and occurs primarily on the 5'-side of the bulge. This is consistent with the observation that major groove accessibility extends further in the 5'-direction than in the 3'-direction in bulged RNA duplexes, based on purine N7 chemical modification studies [31]. The extent of major groove opening present in the spermine- and Mg-form duplexes, however, is greater than the relatively modest groove opening caused by a single nucleotide bulge as assessed by chemical alkylation experiments.

Lattice interactions of coaxially stacked RNA helices

The spermine-form and Mg-form crystal lattices are both characterized by extensive stacking interactions between adjacent duplexes. In the spermine form, duplexes are

Figure 4
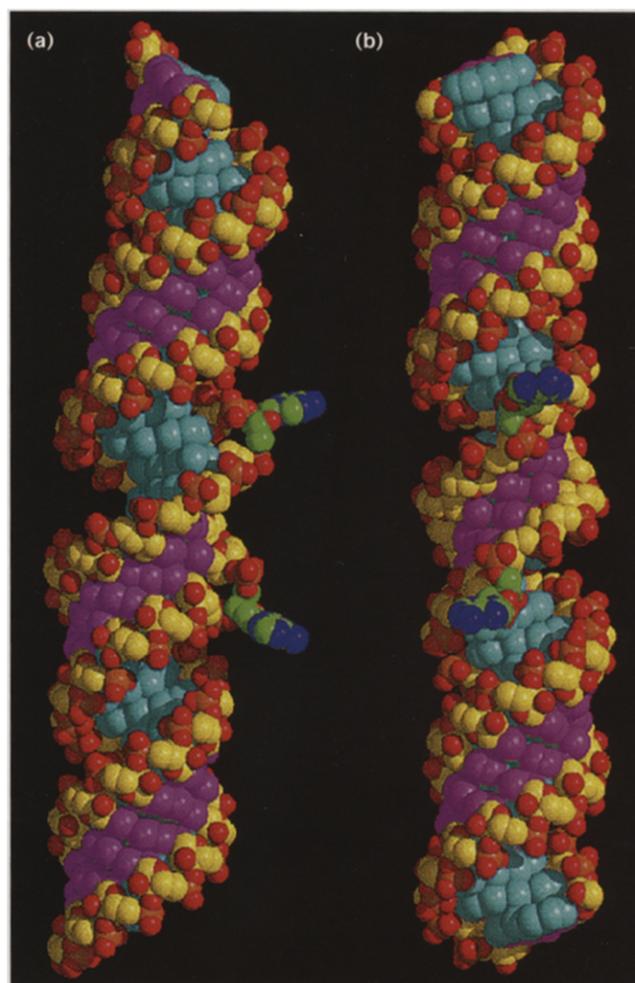
The bulged duplexes are kinked with enlarged major grooves and compressed minor grooves. (a) A view across the widened major groove (cyan) between the two bulges (green) opening to the left. Note the wider gap between the sugar phosphate backbones in the center compared with the standard topologies of the major groove above and below. To show the kink more clearly, we have generated an extended double helix by superpositioning three base pairs of standard A-form RNA duplexes on both termini of the bulged spermine-form fragment (the RNA portions of the Mg-form are similar). (b) The duplex rotated around the vertical by 90° and viewed into the compressed minor groove (purple) between the two bulges. The decamer duplex $(gcg/ATAT/cgc)_2$ can be thought of as a molecule containing two trimer duplexes and a central tetramer duplex, with the two bulge-induced kinks at the trimer–tetramer junctions.

stacked head to tail (Fig. 6), creating infinitely long columns along the crystallographic six-fold axis. In the Mg-form coaxial stacking of duplexes is also observed, but the helices are oriented head to head (Fig. 7). The stacking interactions between terminal base pairs in the spermine-form lattice resemble those occurring for a pyrimidine–purine step in a continuous duplex. Thus, although the twist angle at the junction between the two duplexes is considerably reduced in the spermine-form

Figure 5
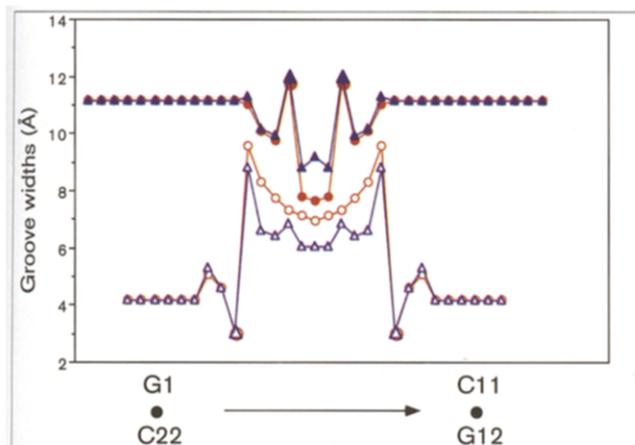
The duplexes have a widened major groove and a narrowed minor groove in the vicinity of the bulge. A plot of the major (O) and minor (Δ) groove widths in the extended RNA–DNA chimeric duplexes, based on the spermine-form (red) and the Mg-form (blue) crystal structures (see Fig. 2) is shown. The groove widths are defined as the shortest P···P distances minus 5.8 Å (the sum of the PO₂ van der Waals radii) and are plotted from one end of the duplex to the other. Bigger circles and triangles indicate P···P distances involving the phosphate of a bulge residue.

lattice (Fig. 6b), there is extensive interstrand stacking at this C(p)G step (phosphate group missing), common in RNA A-form duplexes. Coaxial stacking of helices within a single molecule makes an important contribution to the overall stability of large RNAs [60], and has been observed in several large RNA structures [51,61–63]. Coaxial stacking between helices also seems to be the common packing motif in most crystal structures of RNA oligonucleotide fragments [64]. In the crystal structure of an eight-nucleotide RNA, the interactions between the ends of stacked duplexes were found to mimic those in a regular GpC step [57]. The terminal base pairs featured an extensive intrastrand overlap, as expected for a purine–pyrimidine step. Coaxial stacking of helices in large RNA molecules and between duplexes in crystal lattices is thus governed by similar rules to those for base-pair stacking within the continuous duplex of an RNA stem.

In the Mg-form crystal, only the terminal guanine bases of duplexes stack on top of one another, with the cytosine portions involved in stacking interactions with the bulge adenosines from adjacent duplexes (Fig. 7b). In the spermine-form crystal, packing causes stacking between bulge adenosines from two adjacent duplexes (Fig. 6b). The N6/N7 edges of this A nucleotide face the minor groove formed at the interface of a third and a fourth duplex, which are stacked end to end so that the minor groove is nearly continuous (Fig. 6). This arrangement is also stabilized by a C–H···O hydrogen bond between C8 of the bulge adenine and the ribose 2'-hydroxyl group from

the terminal G residue (2.89 Å, Fig. 6b). A similar relative orientation between the bulge base and a terminal G residue from an adjacent duplex is present in the Mg-form lattice. The bulge base in this lattice sits on top of a terminal cytosine base, in contrast to its position in the spermine-form lattice, where it is stacked on a symmetry-related bulge adenosine (Fig. 7b). Compared with the spermine form, the bulge base in the Mg-form lattice is slightly rotated with respect to the terminal guanine and shifted away from its ribose, leading to a stabilization of the C•G–A base triple formed by a 2.97 Å hydrogen bond between N7(A) and N2(G) (Fig. 7b). The spermine-form crystals are much less tightly packed than the Mg-form crystals, allowing the more extensive conformational deviations from the standard geometry that were observed for the spermine form duplex. The well resolved electron density maps produced by the relatively high resolution of the Mg-form data revealed the presence of two spermine molecules in this lattice. The spermines are engaged in

Figure 6
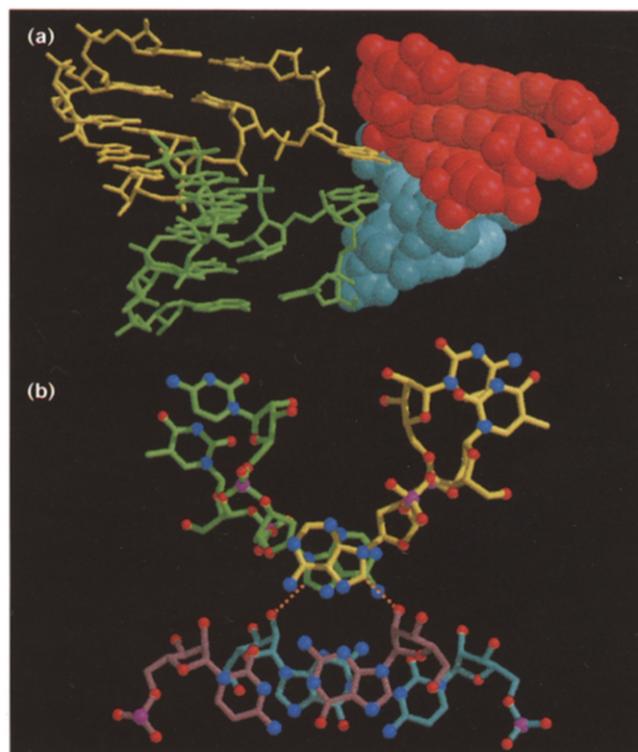
Molecular interactions in the spermine-form crystal lattice. **(a)** Stacked adenosine bulges from two adjacent duplexes (yellow and green; note the *anti* conformation of the bulge base) contact the minor groove formed by end-to-end stacked helices (having 5':3'/5':3' interactions), shown as red and cyan van der Waals models. **(b)** The interaction viewed roughly along the crystallographic six-fold rotation axis. Well stacked adenosine bulges form base triples with terminal G•C base pairs of duplex stacks. The terminal base pair of the red van der Waals duplex in (a) is drawn in pink, oxygen atoms are red, nitrogen atoms are blue and phosphorus atoms are magenta. The C–H···O hydrogen bond between C8 (A) and O2' (G) is shown as a dotted orange line.

Figure 7
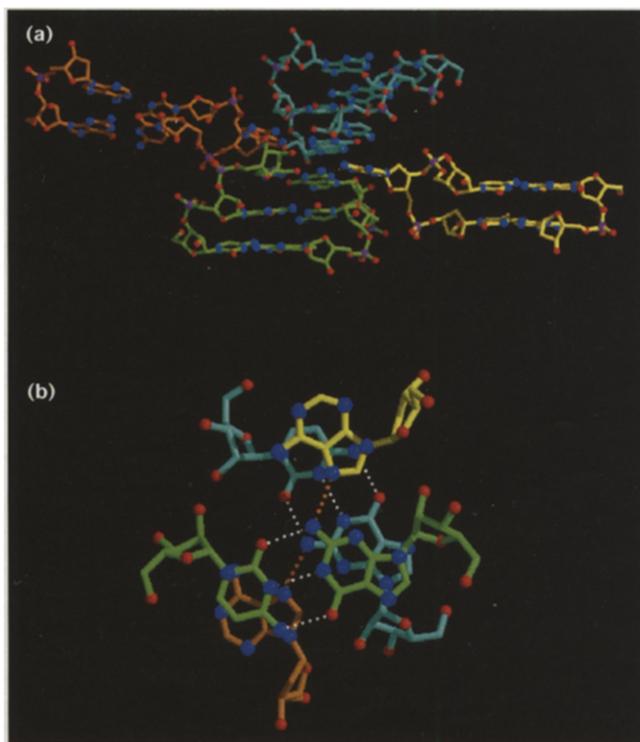
Molecular interactions in the Mg-form crystal lattice. **(a)** Looped-out adenosines from two symmetry-related duplexes (yellow and orange bonds; note the *syn* conformation of the bulge base) are shifted over the base planes of cytosines from terminal base pairs within end-to-end stacked duplexes (having 5':5'/3':3' interactions), shown in green and cyan. **(b)** The G•C–A base triples, which form the interface between stacked duplexes, viewed roughly along the normal to the bulged adenosine base plane. Color code as in Fig. 6; Watson–Crick hydrogen bonds are shown as dotted white lines, and the hydrogen bond between N7 of A and N2 of G is shown as a dotted orange line.

numerous hydrogen bonds to phosphate groups and base atoms as well as to water molecules (Table 2).

A conformational basis for RNA self-cleavage at bulge sites

RNA self-cleavage occurs at sites containing single- or double-base bulges or base mismatches in the presence of a variety of divalent metal cations, such as Mg^{2+} and Zn^{2+} (D. Hüsken and colleagues, unpublished data). With single-base bulges this cleavage proceeds via an attack of the 2'-hydroxyl group of the bulge nucleotide at the adjacent phosphate group on the 3' side, resulting in a 2',3'-cyclic phosphate at the bulge residue and a free 5'-hydroxyl terminus. The deoxynucleotide bulge in the structures determined here can be thought of as an inhibitor of this cleavage reaction. Modeling the bulge residue as a ribonucleotide by attaching an OH group at its 2'-position, however, can provide clues about the conformational requirements in and around the bulge for an in-line mechanism of attack of the 2'-O⁻ at the 3'-phosphate phosphorus atom. Phosphodiester transesterification can

proceed via either an in-line or adjacent mechanism; the latter requires a pseudorotation to position the attacking and leaving groups in axial or apical positions in the trigonal-bipyramidal transition state ([65,66] and references therein). A preorientation of the RNA backbone for an in-line displacement, such as that found for the formation of 2'–5' linkages, dramatically enhances the cleavage rate [67]. RNA self-cleavage is probably facilitated at specific sites by deviations from a standard double-helical conformation, such as at stem–hairpin loop junctions or within loops (e.g. [68,69])

Since acceleration of RNA cleavage is dependent on conformation, and RNA self-cleavage reactions occur at bulge sites, we examined the orientation of a 2'-oxygen modeled into the bulge, relative to the scissile P-O5' bond of the phosphate group 3' to the bulge adenosine. In the spermine-form duplex, the modeled bulge 2'-oxygen is roughly adjacent to the 3'-phosphate group (Fig. 8a). The O2'–P and P–O5' vectors form a right angle and the O2'···P distance is over 4 Å. In the Mg-form crystal lattice, however, the altered bulge conformation results in a more open O2'–P–O5' angle, with a slightly reduced O2'···P distance of 3.8 Å (Fig. 8b). This arrangement clearly represents the initial stage of an in-line replacement at the 3'-phosphate group; thus, a

Table 2

Spermine interactions in the Mg-form crystal.

Spermine atom	Nucleic acid atom/water	Distance (Å)
12 spm (N1)	C9 (O1P)	2.75
	G10 (O1P)	3.23
	G10ᵃ (O2P)	2.72
	23 (H₂O)	3.15
12 spm (N5)	G10 (O1P)	3.04
	C11ᵃ (O2P)	2.87
12 spm (N10)	G1ᵃ (N7)	2.92
	G1ᵃ (O4')	3.14
	C10ᵃ (O1P)	2.90
12 spm (N14)	C9ᵃ (N4)	3.23
	G10ᵃ (O6)	2.87
	17ᵃ (H₂O)	3.25
	51ᵃ (H₂O)	3.31
13 spm (N1)	C2ᵃ (O3')	3.08
	G3ᵃ (O2P)	2.83
	T5ᵃ (O1P)	2.75
	A8 (O1P)	2.63
13 spm (N5)	A6ᵃ (O2P)	2.78
	18 (H₂O)	3.08
13 spm (N10)	31ᵃ (H₂O)	2.66
	44ᵃ (H₂O)	3.38
13 spm (N14)	G3ᵃ (N7)	2.96
	28ᵃ (H₂O)	2.89

ᵃSymmetry-related atoms

Figure 8
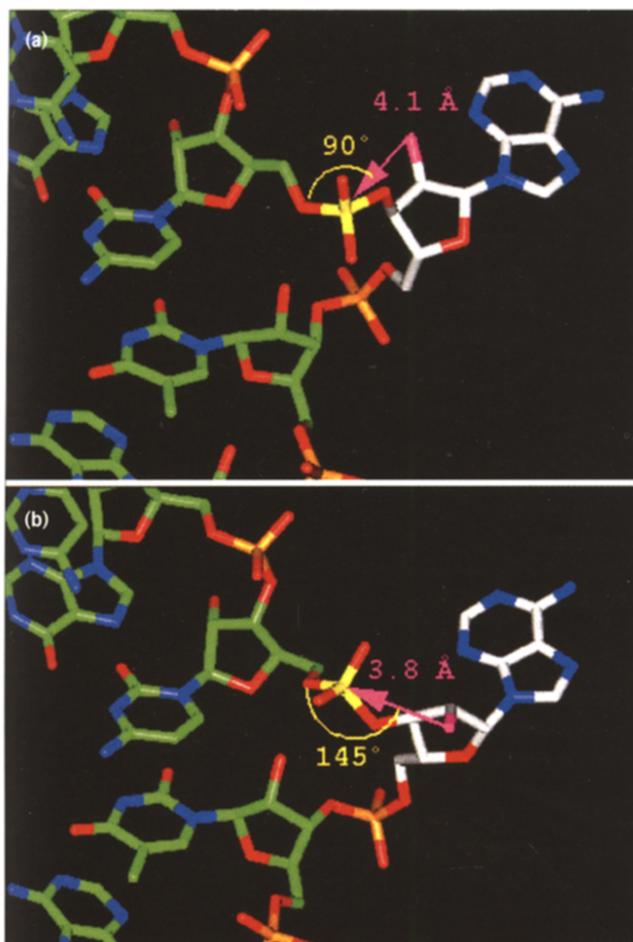
In the Mg-form duplex, the conformation of the bulge residue would preorient an RNA backbone for self-cleavage. **(a)** Adjacent arrangement of the modeled bulge 2'-oxygen (pink) and the scissile P–O5' bond in the spermine-form duplex. **(b)** Near in-line arrangement of the modeled 2'-oxygen and scissile bond in the Mg-form duplex. The carbon atoms of the bulged nucleotide are shown in white. The line of attack by the bulge 2'-oxygen (pink) at the adjacent phosphate group on the 3' side (yellow) is indicated with pink arrows and the hypothetical O2'–P–O5' angle is shown in yellow. Carbon atoms are green (except as noted above), oxygen atoms are red, nitrogen atoms are blue, phosphorus atoms are orange (except as noted above). The tight contacts between phosphate oxygens of residues C9 and G10 are visible in the upper portions of the two figures.

particular stable bulge conformation can preorient the RNA backbone for a self-cleavage reaction.

The accommodation of the bulge adenosine within the duplex causes an unusually close spacing between the phosphate oxygens from residues C9 and G10 on the 3' side of the bulge in both A-form duplexes (Figs 8,9). In the spermine form, the distance between O1P (C9) and O2P (G10) is 3.95 Å and thus well below the value commonly observed between phosphate oxygens for A-form duplexes. In the Mg-form duplex, the

corresponding distance is 4.14 Å. Consistent with the short distances between these phosphate oxygens, the P···P distances for the two residues are clearly below the average distance in both the spermine (5.55 Å, average 6.08 Å) and the Mg-form duplex (5.24 Å, average 5.98 Å). The negative potential built up by the two adjacent phosphates seems to be an attractive coordination site for cations. It is very plausible that the 3'-phosphate group is activated for attack by the (presumably deprotonated) 2'-OH of the bulge and that the 5'-leaving group is stabilized through ion coordination to that site.

Although the Mg-form crystals were grown from droplets that contained moderate amounts of magnesium chloride, we did not locate any Mg^{2+} ions in the electron density maps. One of the spermine molecules (residue 12 in Table 2) observed in the Mg-form crystal, however, coordinates to O1P phosphate oxygens of residues C9 and G10. It is possible that, at the relatively high concentrations of spermine which were used for growing these crystals, the polycation is effectively competing with Mg^{2+} for the coordination to the activation site. Nevertheless, the Mg-form crystal structure provides an excellent model for site-specific RNA self-cleavage at bulge sites. It also provides an explanation for the location of the cleavage to the 3'-side of the bulge, which results from the relative orientations of attacking group and scissile bond and from the activation and stabilization of phosphate and leaving groups, respectively, through cation coordination.

Figure 9
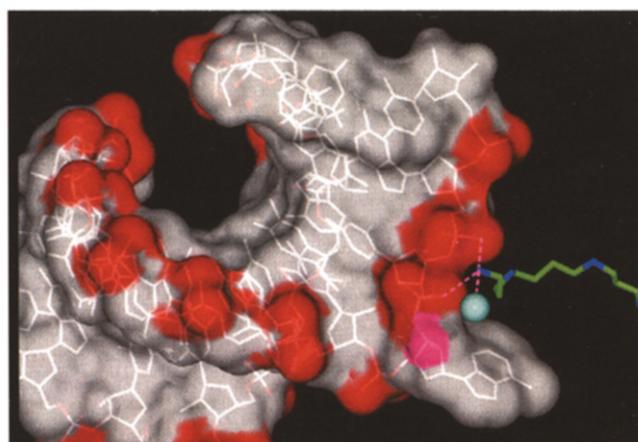
Surface drawing of the bulged Mg-form duplex with an underlying skeletal model viewed across the major and minor grooves. The bulge adenosine is located at the lower right. The red patches indicate the location of the bridging and non-bridging phosphate oxygens, but do not represent the electrostatic potential distribution in a strict sense. The drawing shows the two closely spaced, negatively charged phosphate groups adjacent to the bulge on the 3' side and the coordinating spermine molecule (residue 12, Table 2). The modeled bulge 2'-oxygen is highlighted in pink, the cyan sphere is a water molecule, positively charged spermine nitrogen atoms are blue and carbon atoms are green, and hydrogen bonds are dashed.

Significance

Our results reveal three possible roles for nucleotide bulges in the control of RNA topology and recognition. First, the bulge causes a kink in the stem region of an RNA and thus can influence long-range interactions in large RNAs. The kink of 10° per adenosine bulge observed in the A-type helices here is similar in magnitude to bulge-related kinking in RNAs as assessed by a variety of methods, both in solution and in the solid state. Second, the bulge widens the narrow major groove, predominantly on the 5'-side of the unpaired nucleotide, creating a conformation that allows RNA-binding molecules to probe the nucleotide sequence. The drastic expansion of the major groove between the bulged nucleotides in our crystal structures to almost twice its normal width is greater than we anticipated for a single-base bulge. This result is undoubtedly due in part to the location of the bulges on opposite strands, separated by roughly a half-turn, in our model duplex. Third, the bulged base is looped out of the helix, providing a structural basis for the involvement of these sites in interactions with proteins and in RNA folding and processing. Protein contacts can be mediated through both stacking interactions and hydrogen bonding to, for example, the adenine N1 and N6 atoms, which are prominently displayed above the duplex.

The Mg-form crystal structure provides a structural basis for the preferential ion-assisted self-cleavage of RNA in single-stranded regions, such as bulge loops. In addition, our crystallographic results are compatible with several bulge-related catalytic functions of RNA. The modeled 2'-OH of the bulge in the Mg-form structure is exposed on the surface of the duplex; the 2'-OH acts as the nucleophile in group II intron lariat formation and in the first step of the pre-mRNA splicing reaction, and a surface location would make sense for such functions. A looped-out conformation of the A263 bulge at the guanosine-binding site in the *Tetrahymena* group I intron, similar to the ones in the present crystal structures, would widen the major groove of the P7 helix. Although the A263 bulge is located adjacent to the helix terminus, where the groove is naturally more open, a further increased groove width would facilitate access for the substrate from the major groove side to base pair G264•C311 and nucleotide A265.

Materials and methods

Oligonucleotide synthesis and crystallizations
The chimeric oligonucleotide was synthesized by the solid-phase phosphoramidite method [70,71] and purified according to described methods [71]. Spermine-form crystals were grown from droplets containing 0.8 mM oligonucleotide, 15 mM sodium cacodylate buffer (pH 6.9) and 50 mM spermine tetrahydrochloride (drop size 20 μl), equilibrated against 25 ml of 40 % 2-methyl-2,4-pentanediol (MPD). A crystal of size 0.4 x 0.3 x 0.2 mm grown in 9 days was selected for data collection. Addition of small amounts of

magnesium chloride to the crystallization buffer resulted in immediate quantitative precipitation of the nucleic acid fragment. This indicates that the adenosines are also looped out in solution, and that clamping of stacked A nucleotides from two duplexes by the metal ions could generate insoluble aggregates of high molecular weight. Such simultaneous Mg^{2+} coordination to the N7 atoms from two stacked purine bases was observed previously in a high resolution RNA crystal structure [72]. Crystallization solutions supplemented with Mg^{2+} were therefore slightly heated prior to the crystallization setups to avoid strong precipitation, and the concentration of Mg^{2+} was kept lower than the one for spermine tetrahydrochloride. Mg-form crystals took several months to grow, and crystallization was generally preceded by precipitation. The typical crystallization conditions were 0.5 mM oligonucleotide, 10 mM sodium cacodylate buffer (pH 7.4), 33 mM spermine tetrahydrochloride, 2 mM magnesium chloride (drop size 30 μl), equilibrated against 25 ml 40 % MPD. The crystal with size 0.2 x 0.2 x 0.2 mm used for data collection was cut out of a cluster. Crystal data and solution parameters for both crystal forms are summarized in Table 1.

Crystal structure solution and refinement: spermine form
Data were collected at room temperature on a Rigaku R-axis II image plate system. A total of 18394 reflections were collected and merged to 1034 unique reflections with an overall R_{merge} of 7.5 %. The data is 82 % complete to 2.8 Å resolution (50 % completeness in the 3.0–2.8 Å shell). The structure was solved by the Molecular Replacement method, performing numerous rotation and translation searches around and along the four distinct two-fold rotation axes in the enantiomorphic space group pair $P6_122/P6_522$ with several bulged and unbulged A- and A'-form duplex models. Models with alternative sites for the A-bulge in the chosen sequence were also taken into account, for example 5'-gcgATATAcgc (A4 model) and 5'-gcgATATAcgc (A8 model) (Fig. 1, RNA lower case, bulge underlined). Model A4 could be refined to a crystallographic R-factor of 31.5 % (R_{free} = 47.4 %) [73,74], and refinement of model A8 resulted in an R-factor of 21.1 % (R_{free} = 27.8 %; neither model included solvent molecules). The correctness of the A8 model was supported by sum ($2F_{obs}-F_{calc}$) and difference ($F_{obs}-F_{calc}$) electron density in the vicinity of residues T7 and c9 (clearly indicating a looped-out nucleotide) in maps based on the A4 model (Fig. 1b). The structure was refined with X-plor [73] and the final R-factor, including 23 water molecules per single strand, is 18.7 % ($F_{obs} \geq 2\sigma(F_{obs})$ between 10 and 2.8 Å).

Crystal structure solution and refinement: Mg-form
A total of 9641 reflections were measured at room temperature on identical in-house equipment and merged to 2264 unique reflections with an overall R_{merge} of 8.1 %. The structure was solved with the Molecular Replacement program AMoRe [75,76], using the spermine form duplex as a model (resolution 30–3 Å, sphere radius 12 Å). Rigid-body fitting with a single strand lacking the bulge adenosine produced a correlation factor of 52.6 and an R-factor of 48 %. This solution was then subjected to refinement with X-plor and the resolution of the included data was gradually extended to 1.83 Å. At later refinement stages, the bulge nucleoside and the adjacent phosphate groups as well as 38 water molecules and two spermines were assigned to regions of superimposed sum and difference electron density. The final R-factor was 15.9 %. ($F_{obs} \geq 2\sigma(F_{obs})$ between 10 and 1.83Å)

The final coordinates for both structures have been deposited with the Nucleic Acid Database. The permanent NDB ID codes for the spermine- and Mg-form structures are UHK045 and UHK046, respectively.

Acknowledgements

References

1. Wyatt J.R. & Tinoco Jr, I. (1993). RNA structural elements and RNA function. In *The RNA World* (Gesteland, R.F. & Atkins, J.F., eds), pp. 465–496, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
2. Turner, D.H. (1992). Bulges in nucleic acids. *Curr. Opin. Struct. Biol.* **2**, 334–337.
3. Baudin, F. & Romaniuk, P.J. (1989). A difference in the importance of bulged nucleotides and their parent base pairs in the binding of transcription factor IIIA to *Xenopus* 5S RNA and 5S RNA genes. *Nucleic Acids Res.* **17**, 2043–2056.
4. Dingwall, C., *et al.*, & Skinner, M.A. (1990). HIV-1 Tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure. EMBO J. **9**, 4145–4153.
5. Roy, S., Delling, U., Chen, C.-H., Rosen, C. A. & Sonenberg, N. (1990). A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated *trans*-activation. *Genes Dev.* **4**, 1365–1373.
6. Harper, J.W. & Logsdon, N.J. (1991). Refolded HIV-1 Tat protein protects both bulge and loop nucleotides in TAR RNA from ribonucleolytic cleavage. *Biochemistry* **30**, 8060–8066.
7. Woese C.R. & Gutell, R.R. (1989). Evidence for several higher order structural elements in ribosomal RNA. *Proc. Natl. Acad. Sci. USA* **86**, 3119–3122.
8. Wu, H.-N. & Uhlenbeck, O.C. (1987). Role of a bulged A residue in a specific RNA–protein interaction. *Biochemistry* **26**, 8221–8227.
9. Valegård, K., Murray, J.B., Stockley, P.G, Stonehouse, N.J. & Liljas, L. (1994). Crystal structure of an RNA bacteriophage coat protein–operator complex. *Nature* **371**, 623–626.
10. Zhang, P., Popienick, P. & Moore, P.B. (1989). Physical studies of 5S RNA variants at position 66. *Nucleic Acids Res.* **17**, 8645–8656.
11. Michel, F., Hanna, M., Green, R., Bartel, D.P. & Szostak, J.W. (1989). The guanosine-binding site of the *Tetrahymena* ribozyme. *Nature* **342**, 391–395.
12. Cech, T.R. (1990). Self-splicing of group I introns. *Annu. Rev. Biochem.* **59**, 543–568.
13. Jacquier, A. (1990). Self-splicing group II and nuclear pre-mRNA introns: how similar are they? *Trends Biochem. Sci.* **15**, 351–354.
14. Schroeder, R., von Ahsen, U. & Belfort, M. (1991). Effects of mutations of the bulged nucleotide in the conserved P7 pairing element of the phage T4 td intron on ribozyme function. *Biochemistry* **30**, 3295–3303.
15. McPheeters, D.S. & Abelson, J. (1992). Mutational analysis of the yeast U2 snRNA suggests a structural similarity to the catalytic core of group I introns. *Cell* **71**, 819–831.
16. Moore, M.J., Query, C.C. & Sharp, P.A. (1993). Splicing of precursors to mRNA by the spliceosome. In *The RNA World* (Gesteland, R.F. & Atkins, J.F., eds), pp. 303–357, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
17. Wittop Koning T. H. & Schümperli, D. (1994). RNAs and ribonucleoproteins in recognition and catalysis. *Eur. J. Biochem.* **219**, 25–42.
18. Query, C.C., Moore, M.J. & Sharp, P.A. (1994). Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. *Genes Dev.* **8**, 587–597.
19. Gornicki, P., *et al.*, & Ehresmann, B. (1989). Use of lead(II) to probe the structure of large RNAs. Conformation of the 3' terminal domain of *E. coli* 16S rRNA and its involvement in building the tRNA binding sites. *J. Biomol. Struct. Dyn.* **6**, 971–984.
20. Streicher, B., von Ahsen, U. & Schroeder, R. (1993). Lead cleavage sites in the core structure of group I intron-RNA. *Nucleic Acids Res.* **21**, 311–317.
21. Longfellow, C.E., Kierzek, R. & Turner, D.H. (1990). Thermodynamic and spectroscopic study of bulge loops in oligoribonucleotides. *Biochemistry* **29**, 278–285.
22. White, S.A. & Draper, D.E. (1989). Effects of single-base bulges on intercalator binding to small RNA and DNA hairpins and a ribosomal RNA fragment. *Biochemistry* **28**, 1892–1897.
23. LeBlanc, D.A. & Morden, K.M. (1991). Thermodynamic characterization of deoxyribonucleotide duplexes containing bulges. *Biochemistry* **30**, 4042–4047.
24. Groebe, D.R. & Uhlenbeck, O.C. (1989). Thermal stability of RNA hairpins containing a four-membered loop and a bulge nucleotide. *Biochemistry* **28**, 742–747.
25. Hsieh, C.-H. & Griffith, J.D. (1989). Deletions of bases in one strand of

duplex DNA, in contrast to single-base mismatches, produce highly kinked molecules: possible relevance to the folding of single-stranded nucleic acids. *Proc. Natl. Acad. Sci. USA* **86**, 4833–4837.
26. Rice, J.A. & Crothers, D.E. (1989). DNA bending by the bulge defect. *Biochemistry* **28**, 4512–4516.
27. Tang, R.S. & Draper, D.E. (1990). Bulge loops used to measure the helical twist of RNA in solution. *Biochemistry* **29**, 5232–5237.
28. Bhattacharyya, A., Murchie, A.I.H. & Lilley, D.M.J. (1990). RNA bulges and the helical periodicity of double-stranded RNA. *Nature* **343**, 484–487.
29. Riordan, F.A., Bhattacharyya, A., McAteer, S. & Lilley, D.M.J. (1992). Kinking of RNA helices by bulged bases, and the structure of the human immunodeficiency virus transactivator response element. *J. Mol. Biol.* **226**, 305–310.
30. Lilley, D.M.J. (1995). Kinking of DNA and RNA by base bulges. *Proc. Natl. Acad. Sci. USA* **92**, 7140–7142.
31. Weeks, K.M. & Crothers, D.M. (1993). Major groove accessibility of RNA. *Science* **261**, 1574–1577.
32. Puglisi, J.D., Tan, R., Calnan, B.J., Frankel, A.D. & Williamson, J.R. (1992). Conformation of the TAR RNA–arginine complex by NMR spectroscopy. *Science* **257**, 76–80.
33. Puglisi, J.D., Chen, L., Blanchard, S. & Frankel, A.D. (1995). Solution structure of a bovine immunodeficiency virus Tat–TAR peptide RNA complex. *Science* **270**, 1200–1203.
34. Ye, X., Kumar, A. & Patel, D.J. (1995). Molecular recognition in the bovine immunodeficiency virus Tat peptide–TAR RNA complex. *Chemistry & Biology* **2**, 827–840.
35. Patel, D.J., *et al.*, & Breslauer, K.J. (1982). Extra adenosine stacks into the self-complementary d(CGCAGAATTCGCG) duplex in solution. *Biochemistry* **21**, 445–451.
36. Hare, D., Shapiro, L. & Patel, D. J. (1986). Extrahelical adenosine stacks into right-handed DNA: solution conformation of the d(C-G-C-A-G-A-G-C-T-C-G-C-G) duplex deduced from distance geometry analysis of nuclear Overhauser effect spectra. *Biochemistry* **25**, 7456–7464.
37. Gorenstein, D. G., *et al*, & Jones, C. R. (1987). [32]P NMR and two-dimensional NMR spectra of nucleic acids and 2D NOESY-constrained molecular mechanics calculations for structural solution of duplex oligonucleotides. *Bull. Magn. Reson.* **8**, 137–146.
38. Roy, S., Sklenar, V., Appella, E. & Cohen, J. S. (1987). Conformational perturbation due to an extra adenosine base in a self-complementary oligodeoxynucleotide duplex. *Biochemistry* **26**, 2041–2052.
39. Woodson, S. A. & Crothers, D. M. (1988). Preferential location of bulged-guanosine internal to a G-C tract by ^{1}H-NMR. *Biochemistry* **27**, 436–445.
40. Woodson, S. A. & Crothers, D. M. (1988). Binding of a 9-amino-acridine to bulged-base DNA oligomers from a frameshift hot spot. *Biochemistry* **27**, 8904–8914.
41. Kalnik, M. W., Norman, D. G., Swann, P. F. & Patel, D. J. (1989). Conformation of adenosine bulge-containing deoxytridecanucleotide duplexes in solution: extra adenosine stacks into duplex independent of flanking sequence and temperature. *J. Biol. Chem.* **264**, 3702–3712.
42. Woodson, S. A. & Crothers, D. M. (1989). Conformation of a bulge-containing oligomer from a hot spot sequence by NMR and energy minimization. *Biopolymers* **28**, 1149–1177.
43. Nikonowicz, E. P., Roongta, V., Jones, C. R. & Gorenstein, D. G. (1989). Two-dimensional ^{1}H and ^{31}P NMR spectra and restrained molecular dynamics structure of an extrahelical adenosine base tridecamer oligodeoxyribonucleotide duplex. *Biochemistry* **28**, 8714–8725.
44. Nikonowicz, E. P., Meadows, R. P. & Gorenstein, D. G. (1990). NMR structural refinement of an extrahelical adenosine tridecamer d(CGCAGAATTCGCG)$_2$ via a hybrid relaxation matrix procedure. *Biochemistry* **29**, 4193–4204.
45. van den Hoogen, Y. T., van Beuzekom, A. A., van den Elst, H., van der Marel, G. A., van Boom, J. H. & Altona, C. (1988). Extra thymidine stacks into the d(CTGGTGCGG)•d(CCGCCCAG) duplex. An NMR and model-building study. *Nucleic Acids Res.* **16**, 2971–2986.
46. Morden, K. M., Chu, Y. G., Martin, F. H. & Tinoco Jr., I. (1983). Unpaired cytosine in the deoxyoligonucleotide duplex dC(A)$_3$C(A)$_3$G-dC(T)$_6$G is outside of the helix. *Biochemistry* **22**, 5557–5563.
47. Kalnik, M. W., Norman, D. G., Li, B. F., Swann, P. F. & Patel, D. J. (1990). Conformational transitions in thymidine bulge-containing deoxytridecanucleotide duplexes. *J. Biol. Chem.* **265**, 636–647.
48. Kalnik, M. W., Norman, D. G., Zagorski, M. G., Swann, P. F. & Patel, D. J. (1989). Conformational transitions in cytidine bulge-containing

deoxytrideca-nucleotide duplexes: extra cytidine equilibrates between looped out (low temperature) and stacked (elevated temperature) conformations in solution. *Biochemistry* **28**, 294–303.

49.   van den Hoogen, Y. T., van Beuzekom, A. A., de Vroom, E., van der Marel, G. A., van Boom, J. H. & Altona, C. (1988). Bulge-out structures in the single-stranded trimer AUA and in the duplex (CUGGUGCGG)•(CCGCCCAG). A model-building and NMR study. *Nucleic Acids Res.* **16**, 5013–5030.

50.   Joshua-Tor, L., Frolow, F., Appella, E., Hope, H., Rabinovich, D. & Sussman, J. L. (1992). Three-dimensional structures of bulge-containing DNA fragments. *J. Mol. Biol.* **225**, 397–431.

51.   Pley, H. W., Flaherty, K. M. & McKay, D. B. (1994). Three-dimensional structure of a hammerhead ribozyme. *Nature* **372**, 68–74.

52.   Egli, M., Usman, N., Zhang, S. & Rich, A. (1992). Crystal structure of an Okazaki fragment at 2 Å resolution. *Proc. Natl. Acad. Sci. USA* **89**, 534–538.

53.   Egli, M., Usman, N. & Rich A. (1993). Conformational influence of the ribose 2'-hydroxyl group: crystal structures of DNA–RNA chimeric duplexes. *Biochemistry* **32**, 3221–3237.

54.   Ban, C., Ramakrishnan, B. & Sundaralingam, M. (1994). A single 2'-hydroxyl group converts B-DNA to A-DNA. Crystal structure of the DNA–RNA chimeric decamer duplex d(CCGGC)r(G)d(CCGG) with a novel intermolecular G•C base-paired quadruplet. *J. Mol. Biol.* **236**, 275–285.

55.   Cheong, C., Varani, G. & Tinoco Jr., I. (1990). Solution structure of an unusually stable RNA hairpin, 5'GGAC(UUCG)GUCC. *Nature* **346**, 680–682.

56.   Haran, T. E., Shakked, Z., Wang, A. H.-J. & Rich, A. (1987). The crystal structure of d(CCCCGGGG): a new A-form variant with an extended backbone conformation. *J. Biomol. Struct. Dyn.* **5**, 199–217.

57.   Portmann, S., Usman, N. & Egli, M. (1995). The crystal structure of r(CCCCGGGG) in two distinct lattices. *Biochemistry* **34**, 7569–7575.

58.   Gohlke, C., Murchie, A.I.H., Lilley, D.M.J. & Clegg, R.M. (1994). Kinking of DNA and RNA helices by bulged nucleotides observed by fluorescence energy transfer. *Proc. Natl. Acad. Sci. USA* **91**, 11660–11664.

59.   Zacharias, M. & Hagerman, P.J. (1995). Bulge-induced bends in RNA: quantification by transient electric birefringence. *J. Mol. Biol.* **247**, 486–500.

60.   Walter, A.E., *et al.*, & Zuker, M. (1994). Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding. *Proc. Natl. Acad. Sci. USA* **91**, 9218–9222.

61.   Kim, S.-H., *et al.*, & Rich, A. (1974). Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science* **185**, 435–440.

62.   Robertus, J.D., *et al.*, & Klug, A. (1974). Structure of yeast phenylalanine tRNA at 3 Å resolution. *Nature* **250**, 546–551.

63.   Scott, W.G., Finch, J.T. & Klug, A. (1995). The crystal structure of an all-RNA hammerhead ribozyme: a proposed mechanism for RNA cleavage. *Cell* **81**, 991–1002.

64.   Egli, M., Portmann, S., Tracz, D., Workman, C. & Usman, N. (1995). Crystallization and preliminary X-ray diffraction analysis of double-helical RNA octamers. *Acta Cryst. D* **51**, 1065–1070.

65.   Westheimer, F.H. (1968). Pseudo-rotation in the hydrolysis of phosphate esters. *Accounts Chem. Res.* **1**, 70–78.

66.   Brown, R.S., Dewan, J.C. & Klug, A. (1985). Crystallographic and biochemical investigation of the lead(II)-catalyzed hydrolysis of yeast phenylalanine tRNA. *Biochemistry* **24**, 4785–4801.

67.   Usher, D.A. & McHale, A.H. (1976). Hydrolytic stability of helical RNA: a selective advantage for the natural 3',5'-bond. *Proc. Natl. Acad. Sci. USA* **73**, 1149–1153.

68.   Dock-Bregeon, A.C. & Moras, D. (1987). Conformational changes and dynamics of tRNAs: evidence from hydrolysis patterns. *Cold Spring Harb. Symp. Quant. Biol.* **52**, 113–121.

69.   Hosaka, H., Sakabe, I., Sakamoto, K., Yokoyama, S. & Takaku, H. (1994). Sequence-specific cleavage of oligoribonucleotide capable of forming a stem and loop structure. *J. Biol. Chem.* **269**, 20090–20094.

70.   Usman, N., Egli, M. & Rich, A. (1992). Large scale chemical synthesis, purification and crystallization of RNA–DNA chimeras. *Nucleic Acids Res.* **20**, 6695–6699.

71.   Wincott, F., *et al.*, & Usman, N. (1995). Synthesis, deprotection, analysis and purification of RNA and ribozymes. *Nucleic Acids Res.* **23**, 2677–2684.

72.   Egli, M., *et al.*, & Frederick, C. A. (1990). Atomic-resolution structure of the cellulose synthase regulator cyclic diguanylic acid. *Proc. Natl.*

Acad. Sci. USA **87**, 3235–3239.

73.   Brünger, A.T. (1992). *X-PLOR: A System for X-ray Crystallography and NMR (Version 3.1),* Yale University, New Haven.

74.   Brünger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**, 472–475.

75.   Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* **50**, 760–763.

76.   Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Cryst. A* **50**, 157–163.