RNA bulges as architectural and recognition motifs

Thomas Hermann and Dinshaw J Patel

RNA bulges constitute versatile structural motifs in the assembly of RNA architectures. Three-dimensional structures of RNA molecules and their complexes reveal the role of bulges in RNA architectures and illustrate the molecular mechanisms by which they confer intramolecular interactions and intermolecular recognition.

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

E-mail: thermann@sbnmr1.ski.mskcc.org
pateld@mskcc.org

Structure 2000, **8**:R47–R54

0969-2126/00/$ – see front matter

Introduction

RNA molecules can form intricate three-dimensional architectures comparable to the complexity of protein folds (for reviews see [1,2]). The prevailing building blocks of large RNA molecules are double helices of base-complementary strands. Because of the limited number of packing arrangements, helices alone cannot account for the intriguing structural diversity observed in RNA folds. It is the combination of helical stems interspersed with non-helical motifs that confers structural diversity to the assembly of RNA duplexes (for reviews see [2–4]). Among these motifs, bulges stand out for their universal distribution in all types of structured functional RNAs. Bulges are unpaired stretches of nucleotides located within one strand of a nucleic acid duplex which is formed by hydrogen-bonded bases including canonical Watson–Crick and noncanonical base pairs. Bulge sizes can vary from a single unpaired residue up to several nucleotides that form frequently flexible extrusions from pseudo-continuous double helices. The strand of continuously base-paired residues, opposite to the bulged strand, facilitates the coaxial stacking of the flanking stems. RNA helices containing Watson–Crick base pairs adopt the A-conformation, with a deep but narrow major groove and a shallow minor groove. The discriminatory edges of Watson–Crick base pairs are buried within the inaccessible deep groove [5]. Bulges create unique recognition sites in RNA three-dimensional architectures both directly, by acting as molecular handles within otherwise uniform helical regions, and indirectly, by distorting the RNA backbone and allowing access to base pairs in a widened deep groove. Here, we provide an overview of the structural diversity of bulges in RNA architectures and their roles in intermolecular recognition.

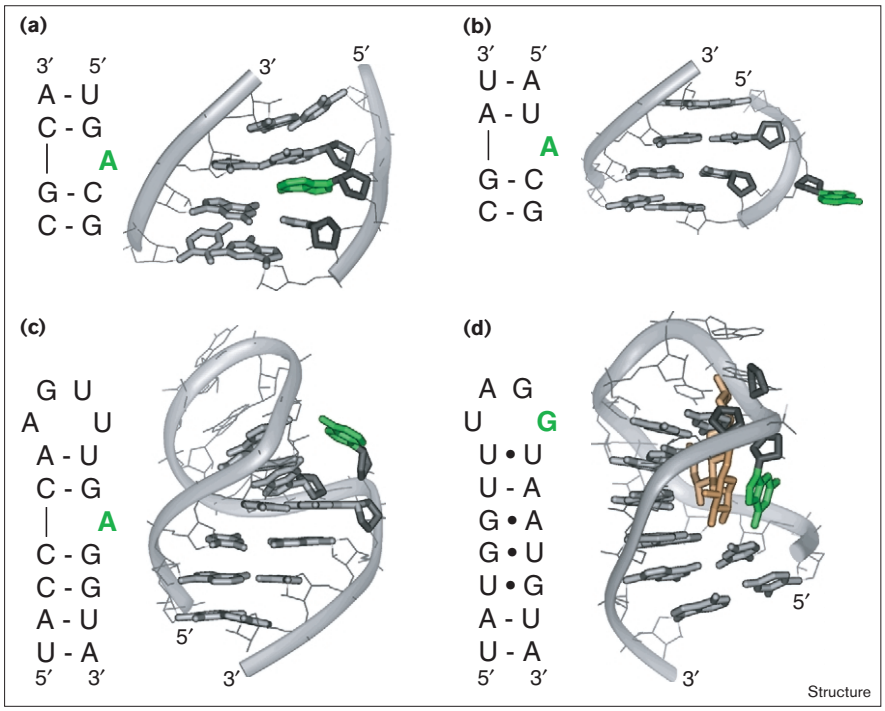
Conformations of bulged residues

The conformations of base-paired nucleotides are restricted by a limited number of possible hydrogen-bonding schemes and stacking interactions. In contrast, for bulged nucleotides a wide variety of arrangements are possible. Their conformations are governed by the competing interactions of both the unpaired residues and the surrounding base pairs. The preservation of continuous stacking between the segments flanking unpaired residues dominates the geometry of bulges. The stacking of bases contributes significantly to the stability of RNA folds as it minimizes the exposure of hydrophobic base surfaces to the polar solvent. Unpaired residues may participate in continuous stacking of the flanking regions, or they can be extruded from the duplex with the base pointing out into the solvent (Figure 1). Unpaired residues that are accommodated between base pairs favour kinking of the helix axis at the bulge site, as revealed, for example, by NMR studies on the RNA hairpins of a splicing regulatory element [6] and the binding motif of a bacteriophage coat protein [7] (Figure 1a). In contrast, when bulge nucleotides are looped out, the overall duplex geometries can remain close to the regular A-form, as observed in the crystal structure of a ten-base-pair duplex containing single bulged adenosines [8] (Figure 1b).

In addition to intercalative stacking, the folding back of unpaired residues onto adjacent hydrophobic helix surfaces, preferably in the shallow groove as seen in a spliced leader RNA hairpin [9] (Figure 1c), are occasionally observed as motifs of bulge stabilization. These non-intercalative packing interactions are labile and, therefore, dependent on subtle environmental changes. In three-dimensional solution structures of the complex formed between the bovine immunodeficieny virus (BIV) TAR (*trans*-activation response) RNA element and an arginine-rich Tat (*trans*-activating) peptide, a bulged uridine was found to be looped out on the shallow groove side of a duplex. Depending on the experimental conditions, the bulged nucleotide either pointed out into the solvent in a disordered manner [10] or packed against adjacent base pairs [11].

In crystals, packing effects can favour distinct bulge conformations. Nucleic acid duplexes prefer to arrange by end-to-end stacking in the crystal, giving rise to pseudo-continuous helices. As a result, bulges tend to be extruded from helices and stabilized by lateral intermolecular interactions. The bulged adenosine in a ten-base-pair RNA duplex (Figure 1b) participates in intermolecular crystal contacts, which fix the base in a completely looped-out geometry [8]. In the crystal structure of the matured

Figure 1



Bulges in RNA architectures. In the simplest case, RNA bulges consist of a single unpaired nucleotide within one strand of a base-paired region. **(a)** The bulged residue can be accommodated by stacking between the adjacent base pairs, inducing a considerable kink in the helix axis, as found in the stem of a 24-nucleotide hairpin loop which is the binding site of a bacteriophage coat protein [7]. **(b)** Alternatively, the unpaired nucleotide may be pushed out of the stem, giving rise to a quasi-continuous double helix, as observed in the crystal structure of an A-form duplex containing an adenosine bulge [8]. **(c)** Packing of the unpaired base into one of the helix grooves contributes to bulge stabilization, as seen in a spliced leader RNA hairpin containing a bulged adenosine that packs into the shallow groove of the stem [9]. **(d)** In some RNA complexes with peptides or small molecules, as in the tobramycin aptamer RNA complex shown here [21], bulged nucleotides participate as 'flap' residues closing over the ligand-binding site. Bulged residues are shown in green. In (d), tobramycin is shown in orange.

dimerization initiation site (DIS) of genomic human immunodeficiency virus 1 (HIV-1) RNA (see below) [12], bulged adenosines adopt two distinct conformations in the two different molecules of the unit cell: one folded back on the backbone of adjacent residues and one flipped-out. The looped-out adenosine forms intermolecular contacts with adjacent DIS RNA molecules in the crystal lattice. Conformational stabilization of single-stranded regions owing to packing effects is common in RNA crystal structures. An impressive example is provided by the lead-dependent ribozyme, which contains an internal loop rather than a bulge [13]. Intermolecular homopurine base-pair formation between three residues within each of the internal loops of two RNA molecules in the crystal gives rise to a parallel RNA helix, fixing the bases in a flipped-out conformation [14]. In solution, the bases of the internal loop are folded towards the inside of the molecule [15]. These examples suggest that some caution is required when comparing the conformations of RNA bulge residues in crystal structures with their counterparts studied in solution.

Just as crystal packing might impose preferred conformations on bulged nucleotides, tertiary contacts in RNA folds and intermolecular interactions in RNA complexes can have a similar effect. In the crystal structure of the P4–P6 domain of group I introns, an adenine-rich bulge docks into a duplex receptor motif which, together with Mg^{2+} ions, locks the bulge conformation in a narrow corkscrew turn [16,17] (see below). Bulged residues that participate in ligand-binding sites of RNA complexes frequently adopt geometries essential for molecular recognition and stabilization of the ligand. In the complex formed between bacteriophage MS2 coat protein and a bulge-containing RNA hairpin, an unpaired adenosine that is intercalated in the free RNA [7] anchors the nucleic acid on the coat protein by looping out and binding into a hydrophobic pocket on the protein surface [18,19]. This bulged adenine is one of three unpaired bases that mediate the specific recognition of RNAs by the MS2 coat protein.

Small molecules and peptides can be partially encapsulated by bulged bases that act as 'flaps' over the ligand-binding sites; this is especially common in RNA aptamers. Such flaps have been found in the solution structures of tobramycin [20,21] (Figure 1d) and neomycin [22] aptamer complexes in which a bulged cytosine or adenine, respectively, closes over the aminoglycoside-binding pocket in the deep groove of the RNA. Similarly, in two complexes of peptides derived from HIV-1 Rev [23] and human T-cell leukemia virus (HTLV-1) Rex [24] bound to RNA aptamers, the peptides are anchored by a guanine flap within the walls of the deep groove.

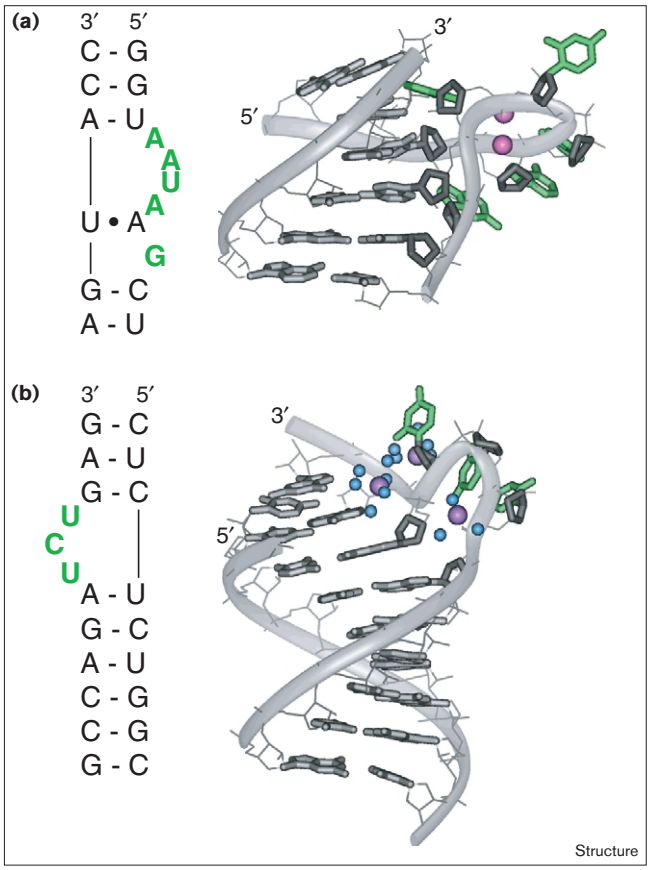
Bulge stabilization by metal ions

Distortions of the RNA backbone at bulge sites can force negatively charged phosphate groups into close proximity, thus creating preferred binding pockets for metal ions [25]. By binding to bulge regions, metal cations attenuate negative charge density and stabilize sharp

turns of the RNA backbone. The adenine-rich bulge, a structural hallmark motif of the P5 stem in the core of the P4–P6 domain of group I introns, adopts a narrow corkscrew turn in which the phosphate groups are turned inside the loop and coordinate two Mg^{2+} ions, while the bases are oriented towards the outside [16,17] (Figure 2a). This 'inside-out' geometry of the adenine-rich bulge allows for the formation of a network of hydrogen-bonding interactions involving the splayed-out bases, which anchor the P5 stem into the P4 subdomain. The prominent role of Mg^{2+} ions in the stabilization of this adenine-rich bulge has led to the suggestion that RNA folds around a 'metal ion core' [17], in analogy to the hydrophobic core of proteins.

The quasi-continuous arrangement of duplex RNA fragments in crystals strongly favours coaxial stacking of the stems that flank the bulge sites. In the crystal structure of an RNA duplex containing the HIV-1 TAR element, the three-nucleotide bulge, which is the recognition site of the viral regulatory Tat protein (reviewed in [26]), is looped out from a stem formed by linear stacking of the base-paired flanking regions [27]. The extruded bulge is stabilized by three Ca^{2+} ions which, through coordination to the phosphate groups, allow for a sharp turn in the RNA backbone (Figure 2b). The bases of the bulge are oriented away from the helix similar to those observed in the adenine-rich bulge of group I introns (see above). The RNA strand opposite to the bulge remains almost undistorted and adopts a conformation close to A-form. The crystallization conditions employed contain high $CaCl_2$ concentrations and Ca^{2+} ions have been found in the TAR bulge. It is also conceivable that Mg^{2+} ions could stabilize the bulge in a similar way, in line with observations from electric birefringence measurements on TAR in the presence of Mg^{2+} [28]. It remains to be established whether the metal-ion-stabilized conformation of the TAR crystal structure is biologically relevant with respect to Tat recognition. In solution, in the absence of high concentrations of divalent cations, the TAR bulge adopts a different conformation with the unpaired bases pointing inside the duplex and stacking on residues of the flanking stems which are kinked relative to each other [29]. Still another conformation of the bulge has been observed for TAR in complexes with either arginine amide or Tat peptides in solution [30,31]. The solution structure of the bound TAR contains elements observed in both the solution and crystal forms of free TAR. The U23 bulge base, which plays a pivotal role in the conformational transition induced by Tat binding, points inside the RNA duplex in both free and bound TAR in solution, positioning it for potential participation in U•A–U triple formation [30,32]. The overall shape of the bulge backbone in bound TAR is closer to that observed in the crystal structure with the characteristic extrusion of the bulge residues from the stem region [27].

Figure 2



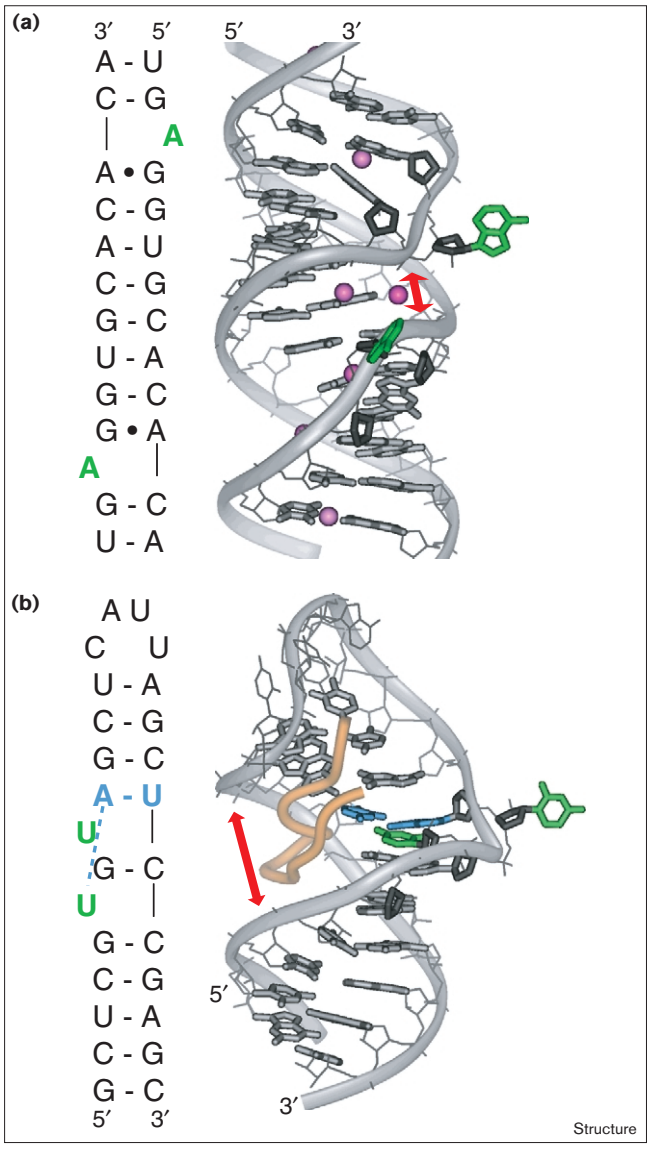
Coaxial stacking of the stems flanking medium-sized bulges can force strong distortions on the backbone causing negatively charged phosphate groups to move into close proximity. **(a)** Mg^{2+} ions (shown in pink) neutralize close backbone charges and stabilize the sharp corkscrew turn formed by the adenine-rich bulge in the P4–P6 domain of group I introns [16,17]. The adenine-rich bulge, extruding from the P5 stem, participates in tertiary interactions at the core of the P4–P6 domain, but, for clarity, it is shown here as an isolated motif. **(b)** Similarly, in the crystal structure of the HIV-1 TAR element, hydrated Ca^{2+} ions (purple) are found associated with the looped-out three-nucleotide bulge [27]. Note the undistorted conformation of the RNA strand opposite to the bulge in both structures. Bulged residues are shown in green and water molecules are depicted as blue spheres.

Bulge recognition by cationic ligands

Like metal cations, positively charged ligands have also been found to stabilize unfavourable electrostatic interactions resulting from the close proximity of phosphate groups in bulge regions. In the solution structures of complexes formed between HIV-1 Rev peptides and the Rev response element (RRE) RNA, a single-nucleotide bulge between two non-Watson–Crick base pairs forms backbone contacts with a positively charged arginine sidechain which attenuates the repulsion between two phosphate groups [33,34].

The electrostatic signature of RNA bulges can contribute to the specific recognition of RNA folds by cationic small

Figure 3



Bulged nucleotides can induce distortions in the conformation of the sugar–phosphate backbone leading to either widened or compressed grooves in RNA duplexes, both being hallmark sites for molecular recognition. **(a)** In the matured dimerization initiation site (DIS) of genomic HIV-1 RNA [12], the deep groove is compressed between two bulged adenosines (green) extruding from opposite strands of the helix. The close approach of the two strands is facilitated by Mg^{2+} ions (pink) which neutralize the negative charge on the sugar–phosphate backbone. **(b)** In the BIV TAR element [10,11], two consecutive bulged uracils, separated by a single canonical G–C base pair, organize the binding site of the arginine-rich Tat peptide (orange). One of the bulged uracil bases is oriented inwards, participating in a U•A–U triple (green•blue) that widens the deep groove of the RNA helix in order to allow the Tat peptide to enter. The other bulged uracil (green) is pushed out, enhancing backbone flexibility for the opening of the deep groove.

molecules through structural electrostatic complementarity between positively charged groups in the ligands and electronegative pockets in the RNA [35,36]. The HIV-1 regulatory RNA elements TAR and RRE, both of which contain bulges (see above) involved in recognition of the Tat and Rev regulatory proteins (for reviews see [37–39]), are specific targets for cationic aminoglycoside antibiotics (reviewed in [40]). The aminoglycosides inhibit the interaction of TAR [41] and RRE [42] with their cognate regulatory proteins. Three-dimensional models of aminoglycoside complexes of TAR [36] and RRE [43] suggest that, in both regulatory RNA elements, bulge nucleotides form contacts with the drugs. In addition to aminoglycosides, the binding of several other classes of cationic low molecular weight compounds to RRE [44] and TAR [45–48] have been shown to involve the recognition of bulge residues. These compounds include diphenylfuran derivatives [44,48], arginine-rich and lysine-rich peptoid/peptid oligomers [45], substituted bisbenzimidazoles [46], and quinoxalinediones [47]. The high specificity potential of drugs that recognize bulges in nucleic acid structures is attested by the enediyne antitumor drug neocarzinostatin, which selectively targets certain DNA bulges but not their RNA analogues [49,50].

Bulge-induced widening and compression of grooves

The dimensions of both the deep and the shallow groove in RNA duplexes can be altered by bulge-induced distortions of the nucleic acid backbone. Widening of the deep groove by non-Watson–Crick base pairs, triples and bulges, exposing the edges of base pairs, is frequently found at interaction sites where RNA domains dock into duplexes or ligands bind to RNA (for reviews see [2,51]). Structural motifs that compress the deep groove are less frequent in RNA architectures, probably because of the resulting close proximity of negatively charged phosphate groups across the groove. As in the case of sharp turns at bulges (see above), metal ions can stabilize such conformations with a compressed deep groove. An example is provided by the crystal structure of the matured DIS of genomic HIV-1 RNA [12], which contains two bulged adenine residues in opposite strands, separated by an eight-base-pair segment (Figure 3a). The extrusion of the unpaired adenines from the duplex along with the formation of adjacent non-Watson–Crick G•A pairs induces compression of the deep groove within the helix segment between the bulges. In contrast, the deep groove is widened in the upper and lower stem regions flanking the bulges. Mg^{2+} ions attenuate the electrostatic repulsion by binding to phosphate groups of the distorted duplex region. The backbone conformation between the two bulges creates a cation-binding site that has been termed the ‘magnesium clamp’ [12], because a Mg^{2+} ion is found bridging the two phosphate groups immediately 3′ of each bulge. The unique conformation of the RNA backbone of the DIS duplex (i.e., the bulged adenine residues and the widened deep groove flanking the bulges) provides potential recognition sites for viral proteins or tertiary contacts with the viral RNA itself [12].

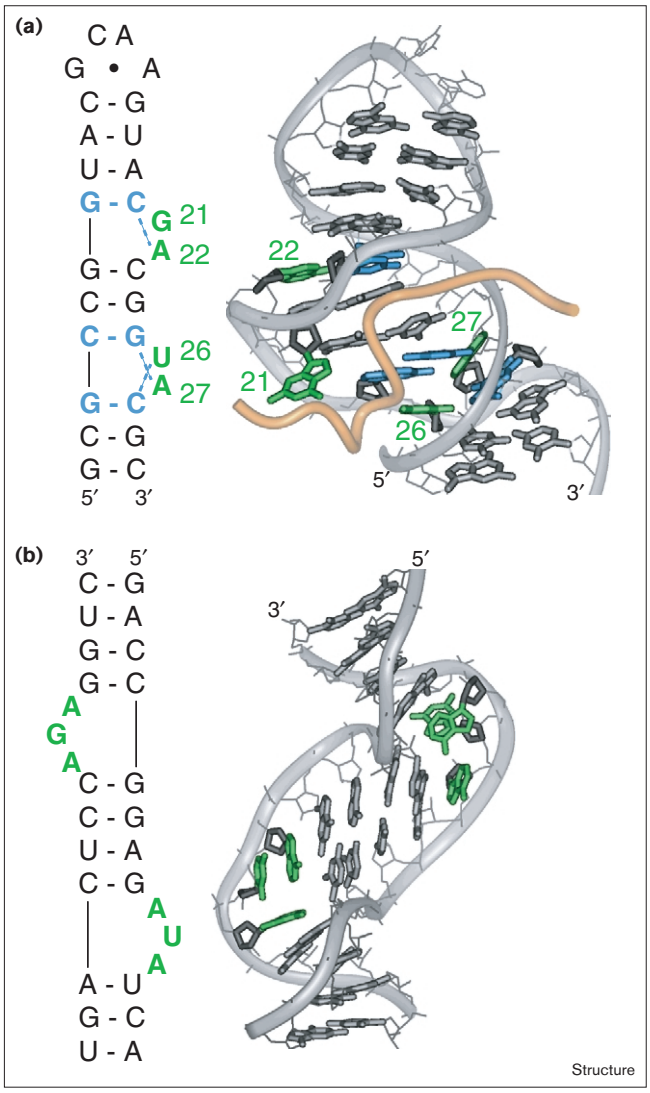
The role of bulges for widening RNA duplex grooves is twofold. Firstly, they provide flexible hinges in the RNA backbone as the unpaired residue is not restricted by base-pairing interactions. Secondly, they can act as 'wedges' pointing inside the deep groove and eventually interacting with base pairs. An additive interplay of these two mechanisms is seen in the three-dimensional structures of BIV TAR RNA bound to Tat peptides [10,11] (Figure 3b). In the TAR element, two bulged uracil residues, separated by a G–C pair, are involved in the organization of the peptide-binding site. One of the uracil bases is flipped out into the solvent and provides sufficient backbone flexibility to allow the participation of the second uracil in a U•A–U triple within the stem [11]. The Tat peptide binds as a β-hairpin fold to the deep groove of BIV TAR that is widened by the triple and bulge motifs. A similar scheme of base-triple formation by the docking of bulged residues into adjacent Watson–Crick pairs has been found in an aptamer RNA bound to an HTLV-1 Rex peptide [24] (Figure 4a). In this complex, three bulged nucleotides participate in base triples that widen the deep groove to accommodate the bound peptide, in addition the peptide is encapsulated by a bulged flap residue that is extruded from the duplex. The stacked pair of base triples at one junctional site, along with the combination of a base triple and an extruded flap residue at the second junction, induce a remarkable dislocation of the helix axes within duplex segments of the aptamer.

As in one of the bulges in BIV TAR RNA, in an RNA aptamer of a HIV-1 Rev peptide [23], a bulged uracil residue was seen to dock into an A–U pair giving rise to a U•A–U triple at the centre of the peptide-binding site in an expanded deep groove. In the complex between the same Rev peptide and its cognate RRE RNA, two looped-out single-base bulges provide flexibility for the widening of the deep groove peptide-binding site by interspersed homopurine non-Watson–Crick base pairs [33]. *In vitro* selection studies of the Rev–RRE interaction have demonstrated the non-conservation of the two bulged single bases in RRE [52,53] in line with their role as flexible hinges. Similarly, in an aptamer RNA complexed with flavin mononucleotide (FMN), the flexible looping out of a non-conserved base from an internal loop zippered up by non-Watson–Crick base pairs allows the formation of an adjacent base triple which functions as a stacking platform for the FMN substrate [54].

Interactions between bulged residues in complex recognition sites

In the previous sections, we have outlined the conformations and stabilization mechanisms of RNA bulges along with their contributions to tertiary interactions and recognition sites in RNA molecules. Here, two RNA architectures will be discussed for which intramolecular interactions between bulged residues have key roles in the formation of unique three-dimensional structures that form complex binding surfaces for proteins. In one of the structures, the bulge–helix–bulge RNA splice site of archaeal pre-tRNA splicing endonucleases [55] (Figure 4b), perpendicular

Figure 4



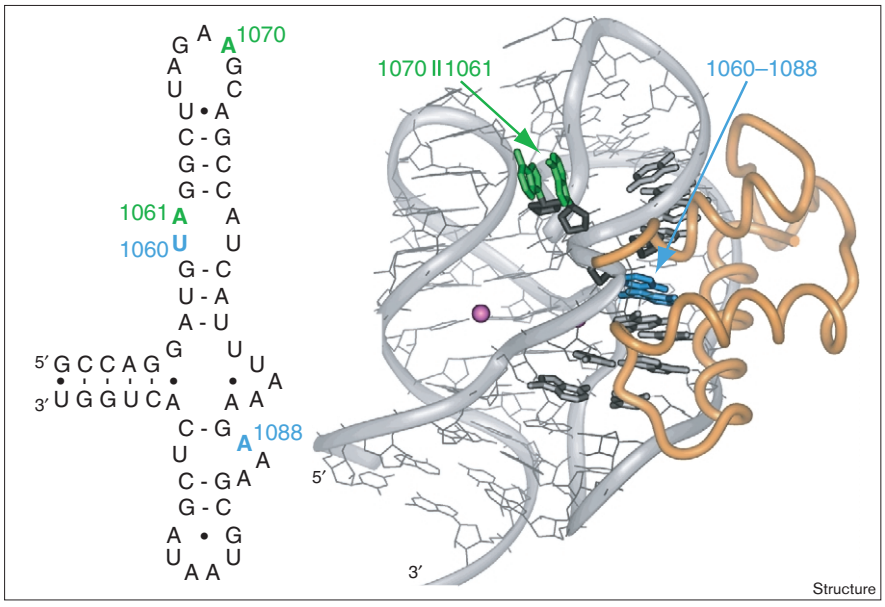
Bulges form constitutive elements of peptide- and protein-binding sites in RNA either indirectly, by shaping the architecture of the binding region, or directly, by providing the recognition motifs. **(a)** In a complex formed between an aptamer RNA and a HTLV-1 Rex peptide [24], three out of four bulged residues (A22, U26, A27) are oriented inside the duplex where they participate in base triples (green•blue) widening the deep groove to accommodate the bound peptide. The fourth bulged base (G21) is flipped-out to form a flap over the bound peptide. Note the dislocation of the helix axes of the three duplex segments joined by the base triples. **(b)** Extensive deformation of the RNA backbone by three-residue bulges is found in the bulge–helix–bulge RNA motif that constitutes the binding site of archaeal pre-tRNA splicing endonucleases [55]. The three unpaired nucleotides in each of the two bulges are stacked inside the double-strand zippered up to a motif in which a central helix is almost perpendicular to the flanking stems.

stacking interactions of adjacent bulge residues shape the protein-recognition site. The other RNA fold, the binding region of L11 protein in 23S rRNA [56,57] (Figure 5), is stabilized by a long-range stacking interaction and a long-range non-Watson–Crick base pair, each involving bulge residues.

The archaeal pre-tRNA intron–exon splice site contains a pseudo-twofold-symmetric bulge–helix–bulge motif that displays high conservation of secondary structure despite sequence variability [58–60]. Nuclear magnetic resonance (NMR) spectroscopy has revealed the three-dimensional structure of the bulge–helix–bulge motif [55], which serves as the recognition site for splicing endonucleases. The three unpaired bases in each bulge are oriented inside the RNA duplex where they stack partially on the adjacent helices (Figure 4a). Two consecutive bulge residues are stacked onto each other, almost perpendicular to the third unpaired base. In addition to these interactions, a network of hydrogen bonds involving the RNA backbone stabilizes both bulges [55]. The backbone conformation of the bulges leads to a Z-like shape of the three connected helix segments with abrupt changes in the helix direction at the bulged sites. In contrast to other bulges that introduce sharp turns in the RNA backbone, such as those observed in HIV-1 TAR and the adenine-rich bulge of group I introns, phosphate groups are not forced into close proximity in the bulge–helix–bulge motif. Hence, there is no requirement for neutralizing divalent metal cations in order to form the three-dimensional structure of the bulge–helix–bulge motif [55]. From the extension of the central helix by the two bulges, a flat recognition surface with a roughly twofold structural symmetry emerges on the shallow groove face that is recognized by the homodimeric splicing endonucleases [61]. It has been noted that the recognition surface of the bulge–helix–bulge motif is dominated by groups of the nucleic acid backbone, which might explain the sequence variations found in splice sites containing this motif [4].

The binding site of the L11 ribosomal protein within the 23S rRNA comprises approximately 58 nucleotides that fold into an intricate RNA architecture. The structure of this RNA in complex with the protein has been analyzed by X-ray crystallography [56,57] (Figure 5). The structural characteristics of the RNA fold and its recognition by the L11 protein, which binds to a shallow groove surface of the RNA, have been outlined in detail [56,57]. Here, we focus on two long-range interactions between sets of bulged nucleotides that organize the RNA fold by pairwise stacking and base pairing (Figure 5). The RNA duplex that constitutes the major binding surface for L11 consists of a number of Watson–Crick base pairs centred around a long-range *trans*-Hoogsteen A•U pair (A1088–U1060), which originates from two bulged nucleotides. This base pair has a key role in the specific binding of L11 as it participates in hydrogen bonding with two consecutive amino acids (Gly130 and Thr131) of the protein. A second crucial long-range interaction is formed by a bulged adenine (A1061) immediately following U1060 of the non-Watson–Crick A•U pair. A1061 loops out from a corkscrew-like turn and stacks on an adenosine (A1070) located in a different bulge within the RNA. This long-range stacking interaction, which has been termed a 'high-five' motif [57], is involved in tethering the loop on top of the L11-binding duplex inside the RNA fold. The two stacking adenines and the bases of the long-range

Figure 5



By stabilizing an adjacent interaction surface, bulges can participate in complex protein-binding sites. Two bulged adenine residues (A1070 || A1061; green), which are extruded from their flanking RNA regions, form a long-range stacking interaction (indicated by ||) in the crystal structure of the 58-nucleotide L11-binding region of 23S rRNA in complex with L11 protein (orange) [56,57]. The stacked bulged bases are part of an intricate network of tertiary interactions, including a long-range *trans*-Hoogsteen A•U pair (U1060–A1088; blue), which stabilize the L11-binding site. A bound magnesium ion is shown in pink.

trans-Hoogsteen A•U pair are highly conserved among eubacteria, archaea and eucarya [56], attesting their decisive role in stabilizing the three-dimensional architecture of the L11-binding domain in 23S rRNA.

Conclusions

The majority of RNA bulges known to date have been observed in the structural context of relatively small oligonucleotides. Nevertheless, from these studies, several distinct sets of preferred bulge conformations and stabilization mechanisms have emerged as candidates for general patterns of RNA three-dimensional architecture. The identification of bulge motifs as key structural elements in a wide range of RNAs underlines the importance and versatility of bulges in RNA architecture and molecular recognition.

The major challenges of research on three-dimensional structures of large RNAs and their protein complexes are currently being addressed at a rapid pace, and include recent progress towards the determination of higher resolution structures of the ribosome and its components. Electron-density maps for the 30S [62] and 50S [63] ribosomal subunits have reached 5 Å resolution and X-ray data for the complete ribosome is available to 7 Å resolution [64]. The refinement of these data to atomic resolution, anticipated in the near future, will reveal the precise three-dimensional structures of ribosomal RNAs and their interactions with proteins. Along with a plethora of novel elements of RNA structure, we can expect to see additional examples of RNA bulges as architectural and recognition motifs.

Acknowledgements

Funding was provided by NIH GM-54777 and NIH CA-49982.

References

1. Ferré-D'Amaré, A.R. & Doudna, J.A. (1999). RNA folds: insights from recent crystal structures. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 57-73.
2. Hermann, T. & Patel, D.J. (1999). Stitching together RNA tertiary architectures. *J. Mol. Biol.*, **294**, 829-849.
3. Batey, R.T., Rambo, R.P. & Doudna, J.A. (1999). Tertiary motifs in RNA structure and folding. *Angew. Chem. Int. Ed. Engl.* **38**, 2326-2343.
4. Moore, P.B. (1999). Structural motifs in RNA. *Annu. Rev. Biochem.* **67**, 287-300.
5. Weeks, K.M. & Crothers, D.M. (1993). Major groove accessibility of RNA. *Science* **261**, 1574-1577.
6. Varani, L., *et al.*, & Varani, G. (1999). Structure of tau exon 10 splicing regulatory element RNA and destabilization by mutations of frontotemporal dementia and Parkinsonism linked to chromosome 17. *Proc. Natl Acad. Sci. USA* **96**, 8229-8234.
7. Borer, P.N., *et al.*, & Pelczer, I. (1995). Proton NMR and structural features of a 24-nucleotide RNA hairpin. *Biochemistry* **34**, 6488-6503.
8. Portmann, S., Grimm, S., Workman, C., Usman, N. & Egli, M. (1996). Crystal structure of an A-form duplex with single-adenosine bulges and a conformational basis for site-specific RNA self-cleavage. *Chem. Biol.* **3**, 173-184.
9. Greenbaum, N.L., Radhakrishnan, I., Patel, D.J. & Hirsh, D. (1996). Solution structure of the donor site of a *trans*-splicing RNA. *Structure* **4**, 725-733.
10. 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.
11. Ye, X., Kumar, R.A. & Patel, D.J. (1995). Molecular recognition in the bovine immunodeficiency virus Tat peptide–TAR RNA complex. *Chem. Biol.* **2**, 827-840.
12. Ennifar, E., *et al.,* & Dumas, P. (1999). The crystal structure of the dimerization initiation site of genomic HIV-1 RNA reveals an extended duplex with two adenine-bulges. *Structure* **7**, 1439-1449.
13. Pan, T. & Uhlenbeck, O.C. (1992). A small metalloribozyme with a two-step mechanism. *Nature* **358**, 560-563.
14. Wedekind, J.E. & McKay, D.B. (1999). Crystal structure of a lead-dependent ribozyme revealing metal binding sites relevant to catalysis. *Nat. Struct. Biol.* **6**, 261-268.
15. Hoogstraten, C.G., Legault, P. & Pardi, A. (1998). NMR solution structure of the lead-dependent ribozyme: evidence for dynamics in RNA catalysis. *J. Mol. Biol.* **284**, 337-350.
16. Cate, J.H., *et al.*, & Doudna, J.A. (1996). Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science* **273**, 1678-1685.
17. Cate, J.H., Hanna, R.L. & Doudna, J.A. (1997). A magnesium ion core at the heart of a ribozyme domain. *Nat. Struct. Biol.* **4**, 553-558.
18. Vålegard, 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.
19. Convery, M.A., *et al.*, & Stockley, P.G. (1998). Crystal structure of an RNA aptamer–protein complex at 2.8 Å resolution. *Nat. Struct. Biol.* **5**, 133-139.
20. Jiang, L., Suri, A.K., Fiala, R. & Patel, D.J. (1997). Saccharide–RNA recognition in an aminoglycoside antibiotic–RNA aptamer complex. *Chem. Biol.* **4**, 35-50.
21. Jiang, L. & Patel, D.J. (1998). Solution structure of the tobramycin–RNA aptamer complex. *Nat. Struct. Biol.* **5**, 769-774.
22. Jiang, L., Majumdar, A., Hu, W., Jaishree, T.J., Xu, W. & Patel, D.J. (1999). Saccharide–RNA recognition in a complex formed between neomycin B and an RNA aptamer. *Structure* **7**, 817-827.
23. Ye, X., *et al.*, & Patel, D.J. (1999). RNA architecture dictates bound peptide conformations. *Chem. Biol.* **6**, 657-669.
24. Jiang, F., *et al.*, & Patel, D.J. (1999). Anchoring an extended HTLV-1 Rex peptide within an RNA major groove containing junctional base triples. *Structure* **7**, 1461-1472.
25. Hermann, T. & Westhof, E. (1998). Exploration of metal ion binding sites in RNA folds by Brownian-dynamics simulations. *Structure* **6**, 1303-1314.
26. Karn, J. (1999). Tackling Tat. *J. Mol. Biol.* **293**, 235-254.
27. Ippolito, J.A. & Steitz, T.A. (1998). A 1.3 Å resolution crystal structure of the HIV-1 trans-activation response region RNA stem reveals a metal ion-dependent bulge conformation. *Proc. Natl Acad. Sci. USA* **95**, 9819-9824.
28. Zacharias, M. & Hagerman, P.J. (1995). The bend in RNA created by the *trans*-activation response element bulge of human immunodeficiency virus is straightened by arginine and by Tat-derived peptide. *Proc. Natl Acad. Sci. USA* **92**, 6052-6056.
29. Aboul-ela, F., Karn, J. & Varani, G. (1996). Structure of HIV-1 TAR RNA in the absence of ligands reveals a novel conformation of the trinucleotide bulge. *Nucleic Acids Res.* **24**, 3974-3981.
30. 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.
31. Aboul-ela, F., Karn, J. & Varani, G. (1995). The structure of the human immunodeficiency virus type-1 TAR RNA reveals principles of RNA recognition by Tat protein. *J. Mol. Biol.* **253**, 313-332.
32. Tao, J., Chen, L. & Frankel, A.D. (1997). Dissection of the proposed base triple in human immunodeficieny virus TAR RNA indicates the importance of Hoogsten interaction. *Biochemistry* **36**, 3491-3495.
33. Battiste, J.L., *et al.*, & Williamson, J.R. (1996). α Helix–RNA major groove recognition in an HIV-1 Rev peptide–RRE RNA complex. *Science* **273**, 1547-1551.
34. Ye, X., Gorin, A., Ellington, A.D. & Patel, D.J. (1996). Deep penetration of an α-helix into a widened RNA major groove in the HIV-1 Rev peptide–RNA aptamer complex. *Nat. Struct. Biol.* **3**, 1026-1033.
35. Hermann, T. & Westhof, E. (1998). Aminoglycoside binding to the hammerhead ribozyme: a general model for the interaction of cationic antibiotics with RNA. *J. Mol. Biol.* **276**, 903-912.
36. Hermann, T. & Westhof, E. (1999). Docking of cationic antibiotics to negatively charged pockets in RNA folds. *J. Med. Chem.* **42**, 1250-1261.
37. Grate, D. & Wilson, C. (1997). Role REVersal: understanding how RRE RNA binds its peptide ligand. *Structure* **5**, 7-11.
38. Naryshkin, N.A., Gait, M.J. & Ivanovskaya, M.G. (1998). RNA recognition and regulation of HIV-1 gene expression by viral factor Tat. *Biochemistry (Moscow)* **63**, 489-503.

39. Patel, D.J. (1999). Adaptive recognition in RNA complexes with peptides and protein modules. *Curr. Opin. Struct. Biol.* **9**, 74-87.
40. Hermann, T. & Westhof, E. (1998). Saccharide–RNA recognition. *Biopolymers* **48**, 155-165.
41. Mei, H.-Y., *et al.*, & Czarnik, A.W. (1995). Inhibition of an HIV-1 Tat-derived peptide binding to TAR RNA by aminoglycoside antibiotics. *Bioorg. Med. Chem. Lett.* **5**, 2755-2760.
42. Zapp, M.L., Stern, S. & Green, M.R. (1993). Small molecules that selectively block RNA binding of HIV-1 Rev protein inhibit Rev function and viral production. *Cell* **74**, 969-978.
43. Leclerc, F. & Cedergren, R. (1998). Modeling RNA–ligand interactions: the Rev-binding element RNA–aminoglycoside complex. *J. Med. Chem.* **41**, 175-182.
44. Ratmeyer, L., *et al.*, & Wilson, W.D. (1996). Inhibition of HIV-1 Rev–RRE interaction by diphenylfuran derivatives. *Biochemistry* **35**, 13689-13696.
45. Hamy, F., *et al.*, & Klimkait, T. (1997). An inhibitor of the Tat/TAR RNA interaction that effectively suppresses HIV-1 replication. *Proc. Natl Acad. Sci. USA* **94**, 3548-3553.
46. Dassonneville, L., Hamy, F., Colson, P., Houssier, C. & Bailly, C. (1997). Binding of Hoechst 33258 to the TAR RNA of HIV-1. Recognition of a pyrimidine bulge-dependent structure. *Nucleic Acids Res.* **25**, 4487-4492.
47. Mei, H.-Y., *et al.,* & Czarnik, A.W. (1998). Inhibitors of protein–RNA complexation that target the RNA: specific recognition of human immunodeficieny virus type 1 TAR RNA by small organic molecules. *Biochemistry* **37**, 14204-14212.
48. Gelus, N., Bailly, C., Hamy, F., Klimkait, T., Wilson, W.D. & Boykin, D.W. (1999). Inhibition of HIV-1 Tat–TAR interaction by diphenylfuran derivatives: effects of the terminal basic side chains. *Bioorg. Med. Chem.* **7**, 1089-1096.
49. Kappen, L.S. & Goldberg, I.H. (1995). Bulge-specific cleavage in transactivation response region RNA and its DNA analogue by neocarzinostatin chromophore. *Biochemistry* **34**, 5997-6002.
50. Kappen, L.S. & Goldberg, I.H. (1997). Effect of ribonucleotide substitution on nucleic acid bulge recognition by neocarzinostatin. *Bioorg. Med. Chem.* **5**, 1221-1227.
51. Hermann, T. & Westhof, E. (1999). Non-Watson–Crick base pairs in RNA–protein recognition. *Chem. Biol.*, R335-R343.
52. 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.
53. Giver, L., Bartel, D., Zapp, M., Pawul, A, Green, M. & Ellington, A.D. (1993). Selective optimization of the Rev-binding element of HIV-1. *Nucleic Acids Res.* **21**, 5509-5516.
54. Fan, P., Suri, A.K., Fiala, R., Live, D. & Patel, D.J. (1996). Molecular recognition in the FMN–RNA aptamer complex. *J. Mol. Biol.* **58**, 480-500.
55. Diener, J.L. & Moore, P.B. (1998). Solution structure of a substrate for the archaeal pre-tRNA splicing endonucleases: the bulge-helix-bulge motif. *Mol. Cell.* **1**, 883-894.
56. Conn, G.L., Draper, D.E., Lattmann, E.E. & Gittis, A.G. (1999). Crystal structure of a conserved ribosomal protein–RNA complex. *Science* **284**, 1171-1174.
57. Wimberly, B.T., Guymon, R., McCutcheon, J.P., White, S.W. & Ramakrishnan, V. (1999). A detailed view of a ribosomal active site: the structure of the L11–RNA complex. *Cell* **97**, 491-502.
58. Thompson, L. & Daniels, C. (1990). Recognition of exon–intron boundaries by the *Halobacterium volcanii* tRNA intron endonuclease. *J. Biol. Chem.* **265**, 18104-18111.
59. Lykke-Andersen, J. & Garrett, R. (1994). Structural characteristics of the stable RNA introns of archaeal hyperthermophiles and their splicing junctions. *J. Mol. Biol.* **243**, 846-855.
60. Kleman-Leyer, K., Armbruster, D. & Daniels, C. (1997). Properties of the *H. volcanii* tRNA intron endonuclease reveal a relationship between the archeal and eukaryal tRNA intron processing systems. *Cell* **89**, 839-847.
61. Li., H., Trotta, C.R. & Abelson, J. (1998). Crystal structure and evolution of a transfer RNA splicing enzyme. *Science* **280**, 279-284.
62. Clemons, W.M., May, J.L.C., Wimberly, B.T., McCutcheon, J.P., Capel, M.S. & Ramakrishnan, V. (1999). Structure of a bacterial 30S ribosomal subunit at 5.5 Å resolution. *Nature* **400**, 833-840.
63. Ban, N., Nissen, P., Hansen, J., Capel, M., Moore, P.B. & Steitz, T.A. (1999). Placement of protein and RNA structures into a 5 Å-resolution map of the 50S ribosomal subunit. *Nature* **400**, 841-847.
64. Cate, J.H., Yusupov, M.M., Yusupova, G.Z., Earnest, T.N. & Noller, H.F. (1999). X-ray crystal structures of 70S ribosome functional complexes. *Science* **285**, 2095-2104.