8

Gel Electrophoresis and the Structure of RNA Molecules

DAVID M.J. LILLEY, ANAMITRA BHATTACHARYYA* AND SEAN McATEER

Department of Biochemistry, The University, Dundee, UK

RNA structure

RNA is central to molecular biology, as the 'working substance' of genetics. The base sequence of genes is transcribed into messenger RNA, which may be processed, spliced and even edited to various extents. The message is then translated into the synthesis of a protein, requiring the combined actions of many RNA and protein species. RNA also has a great many other roles in gene expression, exemplified by transfer RNA and ribosomal RNA. Naturally occurring RNA molecules are formally single-stranded, yet it has long been appreciated that their sequences contain the propensity for considerable secondary structure formation (see for example Dahlberg's review of rRNA; Dahlberg, 1989). In some respects the structures of large RNA molecules probably resemble globular proteins more closely than double-stranded DNA, with both secondary and tertiary interactions determining the overall shape of the molecule. This complex structure is undoubtedly extremely important in the function of RNA, as components of the translation apparatus and in the processing of transcripts. It is now appreciated that RNA is the main player in many of these processes, rather than playing a mere subsidiary role to proteins, as RNA can catalyse important chemical reactions. Of these, the best understood are the transesterification reactions of RNA splicing (Cech, 1987), but recently it has become apparent that the peptidyl transferase reaction of translation may also be RNA-catalysed (Noller, Hoffarth and Zimniak, 1992). In general it has been observed that catalytic activity of RNA does not survive treatments that cause a loss of structure (Cech and Bass, 1986; Uhlenbeck, 1987; Celander and Cech, 1991), and clearly the activity of ribozymes must be intimately connected with structure. RNA

Abbreviations: FRET, fluorescence resonance energy transfer; NOE, nuclear Overhauser effect; RRE, Rev response element; Tat, transactivator protein; TAR, transactivator response element.

---

* *Present address: Department of Microbiology and Immunology, UCSF, San Francisco, CA 94143 0414, USA.*

Biotechnology and Genetic Engineering Reviews — Vol. 10, December 1992
0264–8725/92/10/379–401 $20.00 + $0.00 © Intercept Ltd, P.O. Box 716, Andover, Hampshire SP10 1YG, UK

secondary structure has been shown to play an important role in ribosomal frameshifting (Brierley, Digard and Inglis, 1989).

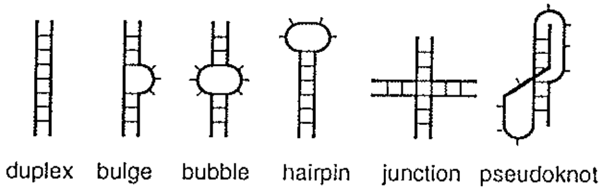


Figure 1. Schematic of some important elements in RNA structure. Most complex folded RNA structures may be subdivided into a number of elements of secondary and tertiary structures, including those shown. These are discussed in the text.

One way of studying the structure of complex folded RNA molecules is to seek smaller elements of the structure that may act as building blocks upon which the overall conformation is constructed (recently reviewed by Tinoco; Chastain and Tinoco, 1991; Tinoco, Puglisi and Wyatt, 1990). In fact, a number of such elements are readily identified in almost any sizeable RNA molecule, summarized in *Figure 1*. Most such RNA molecules contain extensive secondary structure, formed by long-range interactions based on Watson–Crick base pairs, and so the first element is simply the double-stranded duplex. This adopts an A-form helix (Dock-Bregeon *et al.*, 1988), in which the helical axis lies in a large central hole, creating a deep and narrow major groove. The backbone geometry is typically characterized by *C3′-endo* ribose pucker and *anti* glycosyl bonds. Stretches of duplex are frequently interrupted by mismatches, and these can take various kinds. Single-base mismatches comprise non-Watson–Crick pairings, such as an A–G pairing; in DNA such mismatches are relatively non-disruptive, causing minimal deviation of the helix axis (Bhattacharyya and Lilley, 1989b), and crystallographic (Kneale *et al.*, 1985; Brown *et al.*, 1986; Hunter *et al.*, 1986) and NMR (Gao and Patel, 1987, 1988) studies have indicated that such mispairings may be accommodated within the double helical geometry. There is no reason to suppose that RNA will be greatly different in this regard, and this seems to be confirmed by a crystal structure of an RNA duplex containing a number of consecutive U · G and U · C base mismatches (Holbrook *et al.*, 1991). Multiple mismatches are of two types, namely ‘bubbles’ and ‘bulges’. Bubbles (sometimes called internal loops) are a set of consecutive mismatched bases; once again studies in DNA suggest that these do not cause a large change in the trajectory of the helical axis (Bhattacharyya and Lilley, 1989a). Bulges are created where one strand contains one or more ‘extra’ bases, unopposed by bases (of any kind) on the other strand. In contrast to the other mismatched structures, bulges seem to generate a significant axial distortion in DNA, causing a kinking the magnitude of which depends on the number and kind of bases in the bulge. This is supported by electrophoretic studies (Bhattacharyya and Lilley, 1989; Hsieh and Griffith, 1989; Rice and Crothers, 1989), fluoresence resonance energy transfer (FRET) studies (R. Clegg, A.I.H. Murchie and D.M.J. Lilley, in preparation) and recent NMR studies (Aboul-ela *et al.*, 1992; Rosen, Shapiro and Patel, 1992). Sections of duplex are frequently formed by the formation of hairpin loops by sequences of inverted repeat symmetry; the number of bases within the unpaired loop can vary from four (or even less for special situations) to more than 10. The loops themselves are likely to be extensively structured, involving both stacking and hydrogen bonding interactions, as has been found in the stable tetraloop structure by NMR (Cheong, Varani and Tinoco, 1990; Tuerk *et al.*, 1988).

Additional long-range interactions are important in the determination of overall three-dimensional structure; once again these are based on hydrogen bonding and stacking interactions. A good example that illustrates both is the formation of pseudo-knots, first noted in plant viral RNA species (Pleij, Rietveld and Bosch, 1985); in these structures there are base pairing interactions between a single-stranded region and the loop of a hairpin loop (*see Figure 1*). Pseudo-knots are significant structural elements in a number of important RNA species, and are believed to be important in aspects of RNA splicing and translation. Base triple interactions are another important element in the mediation of long-range interactions, based in general on the additional hydrogen bonding capacity of purines in the Hoogsteen positions (the purine 6 and 7 positions). A number of examples of such triples are known in tRNA species, and they are undoubtedly important in the tertiary structures of other RNA molecules. The 2′-hydroxyl group of the ribose moiety may also be an important source of additional interactions; Pyle and Cech (1991) have shown that these groups provide a significant part of the binding energy for *Tetrahymena* ribozyme–substrate interaction.

Another important element of folded RNA structure is the helical junction, a location at which three or four helical segments interact at a point with or without spacing by unpaired bases. The four-way junction without additional bases would be the RNA equivalent of a cruciform, or Holliday junction, in DNA, and an example of a four-way RNA junction is found in the U1 snRNA (Branlant, Krol and Ebel, 1981). These structures have not been extensively studied in RNA, and our understanding of the equivalent DNA structures is very much better (discussed below, and reviewed in Duckett *et al.*, 1992). We would expect that RNA junctions, like their DNA counterparts, will maximize base stacking by undergoing pairwise coaxial helical stacking, as exemplified by tRNA, but other folding principles established for the DNA junction will not apply to RNA.

Another element of nucleic acid folding that should not be neglected is the important aspect of metal ion interactions. Nucleic acids are highly charged polyelectrolytes, and electrostatic repulsion is an important element of the folding process. Tertiary folding frequently brings about the close juxtaposition of mutually repulsive phosphate groups. Electrostatic repulsion is a relatively long-range interaction, and unless this is screened in some way by the intervention of cations, complete folding may be prevented. A very clear example of this is seen in the case of the four-way DNA junction, in which the folded form (the stacked X-structure generated by coaxial stacking of arms) is only observed in the presence of sufficient metal cations (Duckett *et al.*, 1988; Duckett, Murchie and Lilley, 1990). Similarly, cations appear to be required for the correct folding of tRNA (Holbrook *et al.*, 1977; Jack *et al.*, 1977; Quigley, Teeter and Rich, 1978) and the *Tetrahymena* ribozyme (Celander and Cech, 1991). Interactions with metal ions can take two forms; ions may be site-bound with slow exchange, or they may interact as a diffuse ion cloud (ion atmosphere) with fast exchange. Ions located in electronegative clefts generated by the folded structure of an RNA molecule may participate in chemical reactions, such as the lead-catalysed cleavage of tRNA (Brown, Dewan and Klug, 1985).

The solution of the complete structure of complex folded RNA molecules is far from easy, and it is likely to require the combined application of many different techniques and approaches. The molecules are usually too large to yield to the 'classical' approaches of crystallography and nuclear magnetic resonance (NMR), and an eclectic approach is necessary. Individual elements, including duplex, hairpins, bulges and pseudo-knots, are being successfully studied by conventional methods, especially NMR which has been made much more powerful for RNA with the introduction of heteronuclear methods (Nikonowicz and Pardi, 1992). But in general the information derived from NMR is inherently short-range, since much is derived from nuclear Overhauser effects (NOE) which can be seen up to distances of about 5 Å. Thus the extra conformational information must be 'filled in' from other methods, and one very important source of information is gel electrophoresis.

Gel electrophoresis of nucleic acids

Over the last decade gel electrophoresis has been extensively applied to the study of nucleic acid structure, especially for DNA. It provided the first indication of one of the most celebrated examples of local DNA structure, curved or bent DNA, and has continued to provide great insight into DNA structures of various kinds. The persistent problem with this approach is the lack of an agreed physical basis of the method, yet despite this drawback great progress has been made in the study of important structures such as curved DNA and the four-way junction. In the process our understanding of the technique itself has gradually improved, and it is now gradually becoming applied to the area of RNA structure.

It has been known for several years that DNA containing a sequence that generates a local curvature exhibits a pronounced retardation in electrophoretic mobility in polyacrylamide gels (Marini *et al.*, 1982; Diekmann and Wang, 1985; Hagerman, 1985; Koo, Wu and Crothers, 1986); the relative extent of retardation becomes greater as the gel concentration is increased. Furthermore, the fragments migrate most slowly when the sequence causing the curvature is centrally located (Wu and Crothers, 1984). Attempts to obtain a quantitative explanation of the electrophoretic anomaly have not been completely successful, and not lacking in controversy, but various theories can provide at least a qualitative agreement with experimental data (Lerman and Frisch, 1982; Lumpkin and Zimm, 1982; Levene and Zimm, 1989). Most are based upon the idea of the nucleic acid reptation (de Gennes, 1971), in which the DNA is considered to move through the gel under the influence of the electric field rather like a worm constrained to a tube created by the gel matrix. Using these assumptions, Lumpkin and Zimm (1982) derived a relationship between the rate of migration (μ) and the end-to-end distance of the molecule:

$$\mu = \frac{Q}{\zeta} \langle h_x^2/L^2 \rangle$$

where Q is the charge on the molecule, ζ is the frictional coefficient for translation along the tube, L is the contour length of the molecule and h_x is the component of the end-to-end vector **h** in the direction of the electric field. The brackets indicate an average over an ensemble of configurations. Thus the mobility is inversely proportional to contour length, which is the basis of the everyday use of gel electrophoresis to separate nucleic acids according to size. However it is the proportionality to end-to-end distance that explains the dependence on shape; curvature or kinking will clearly reduce this parameter. Calladine, Drew and co-workers (Calladine, Drew and McCall, 1988; Calladine *et al.*, 1991) have taken a different approach to explain the lowered mobility of curved DNA, in which the probability of the cylindrical envelope of a superhelix intersecting randomly located gel fibres is calculated; the greater the curvature the larger the cylindrical radius becomes, and therefore the greater is the probability of obstruction to forward motion. Levine and Zimm (1989) calculated the behaviour of curved DNA fragments under electrophoresis in polyacrylamide, using Monte Carlo methods to generate an ensemble of chain trajectories. They found that they had to include cross-interaction between the bendability of the DNA and the elastic properties of the gel matrix to obtain a good fit with experimental data. As longer DNA molecules are studied (where the contour length is very much longer than the persistence length) it seems probable that the electrophoretic behaviour takes on new characteristics, in which substantial reorganization of overall conformation takes place during the electrophoresis process (Deutsch, 1988).

Base bulges in DNA generate a kinking of the helical axis (Bhattacharyya and Lilley, 1989; Hsieh and Griffith, 1989; Rice and Crothers, 1989), that causes a very marked retardation in electrophoretic mobility in polyacrylamide (*Figure 2*). Like A-tract curved DNA, the magnitude of the retardation (normally expressed in terms of R_L, the apparent fragment size calibrated against 'normal' DNA markers divided by the true size from sequencing) is dependent on the position of the bulge, being greatest when the bulge is located at the centre of the fragment. In principle, such retardation might be due to a precise kinking of the helix axis, or to a point of flexibility. This is resolved by introducing two such bulges into a series of molecules of fixed length, and varying the spacing between the bulges. If the bulges generate a defined kinking, two such bulges should generate a molecule in which the outer arms have a precise angular relationship, and the dihedral angle between them should vary according to the spacing between the bulges and the helical periodicity of the spacer DNA. The end-to-end distance should thus vary in a sinusoidal way with the spacing, and this is just what is found experimentally (*Figure 3*). Lumpkin–Zimm (1982) theory can provide a reasonable agreement with the experimental data on the variation of gel mobility, although the amplitudes are not perfectly reproduced. Similar phasing experiments have been used to study A-tract curving (Drak and Crothers, 1991), protein binding (Zinkel and Crothers, 1991) and cisplatin adducts (Rice *et al.*, 1988). These results are strong evidence for the formation of axial kinks at base bulges, which is further supported by recent experiments in which end-to-end distances in a series of bulge-containing DNA molecules were studied by fluorescence resonance energy transfer (A.I.H. Murchie, C. Goehlke, R.M. Clegg and D.M.J. Lilley, unpublished data). The agreement is reassuring, and shows the power of gel ectrophoresis in the study of nucleic acids.

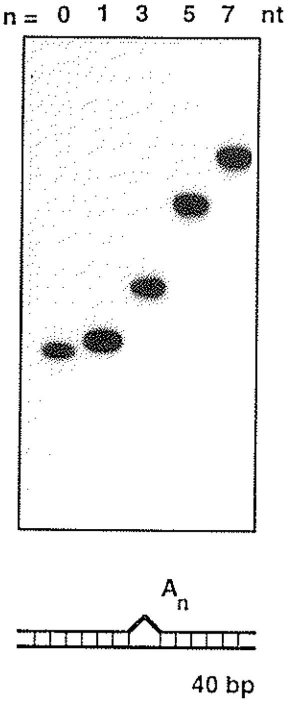


Figure 2. Retardation of electrophoretic mobility of double-stranded deoxyribooligonucleotides in polyacrylamide gels by bulged bases. [5′-^{32}P]-labelled 40 bp oligonucleotides containing central bulges of 0, 1, 3, 5 or 7 adenine bases were electrophoresed in 15% polyacrylamide. The autoradiograph is shown.

RNA duplexes, bulges and bubbles

By comparison with DNA, electrophoretic experiments on RNA structure are only just beginning. Perhaps the major reason for this lies in the greater difficulty in preparing monodisperse RNA samples of defined length and sequence. However, this can now be done relatively routinely. Chemical synthesis of RNA is beginning to become reasonably straightforward, but for electrophoretic experiments, typically requiring molecules of 40–100 bp in length, the usual method is to synthesize individual strands by *in vitro*

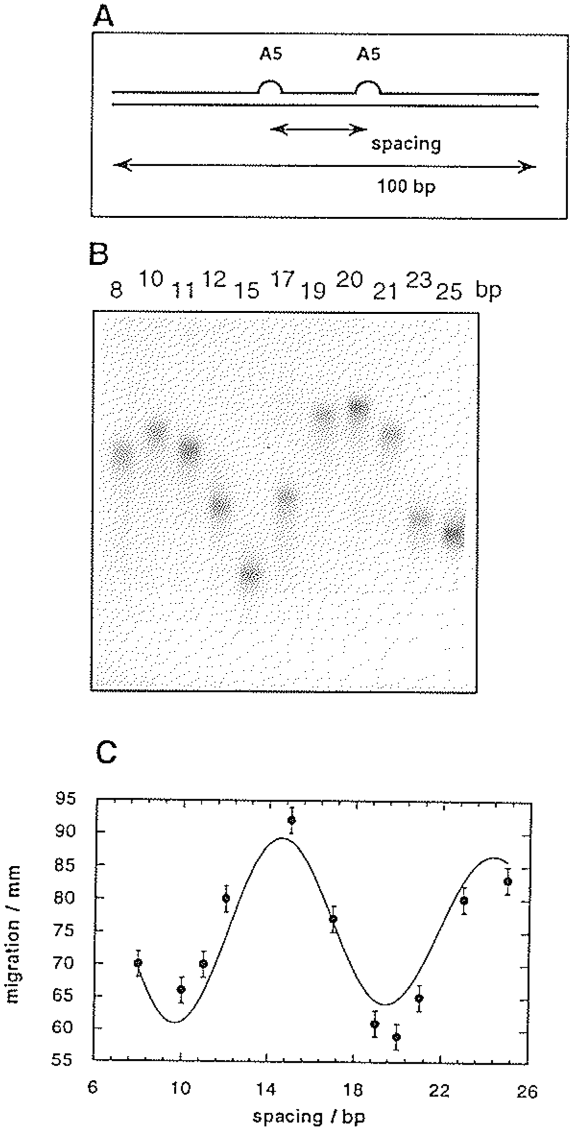


Figure 3. Variation in electrophoretic mobility with the spacing between two bulges in duplex DNA indicates axial kinking.

(A) A series of [5′-^{32}P]-labelled 100 bp duplexes containing two A_5 bulges were constructed, in which the spacing between the bulges was varied from 8 to 25 bp, keeping the bulges symmetrically disposed about the centre of the molecule. These were electrophoresed in polyacrylamide in 90 mM Tris–borate (pH 8·3), 1 mM $MgCl_2$, in 10% polyacrylamide at 15±0·1°C at 15 V cm^{-1}, followed by autoradiography. Note the clear sinusoidal modulation of electrophoretic mobility with bulge spacing.

(B) Plot showing the variation in electrophoretic mobility with the spacing between the A_5 bulges. The error bars on the data points represent the widths of the bands measured on the autoradiograph. The line was generated by non-linear regression to the model of Lumpkin and Zimm (1982).

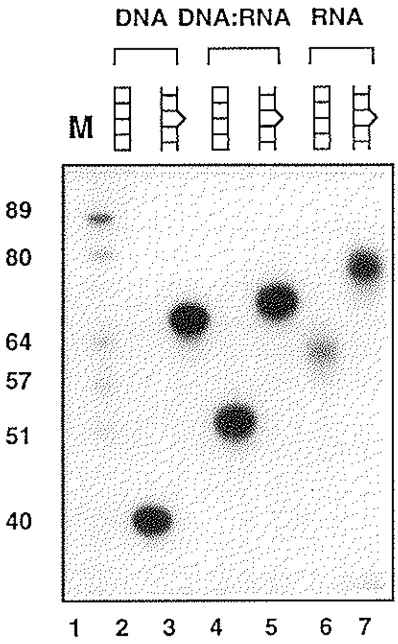


Figure 4. Gel electrophoretic mobility of double-stranded DNA, RNA–DNA and RNA, and the effect of bulge formation. Radioactively labelled 40 bp duplexes of ds DNA, hybrid RNA.DNA and ds RNA of identical sequence were electrophoresed in 15% polyacrylamide, alongside equivalent duplexes containing a bulge of five adenine bases. The hybrid molecule was constructed such that the bulge was located on the RNA strand. RNA species were synthesized by *in vitro* transcription of synthetic DNA templates by T7 RNA polymerase (Milligan *et al.*, 1987). Tracks: 1, DNA restriction fragments as size markers (pAT153 digested with *Hae*III)—the sizes are indicated on the left; 2, ds DNA duplex; 3, DNA duplex with bulge; 4, RNA–DNA hybrid duplex; 5, RNA–DNA duplex with bulge; 6, ds RNA duplex; 7, RNA duplex with bulge.

transcription from a synthetic DNA template using T7 RNA polymerase (Milligan *et al.*, 1987).

Perhaps the simplest RNA species to study are double-stranded duplexes of defined length and sequence. When we compared the mobility of 40 bp RNA and DNA duplexes of identical sequence (Bhattacharyya, Murchie and Lilley, 1990) we discovered that there was a substantial difference in mobility between the two polymers (*Figure 4*; compare tracks 2 and 6). In an earlier study of longer RNA species synthesized by Qβ replicase, Biebricher and co-workers had noted a slower electrophoretic migration compared to DNA (Biebricher, Diekmann and Luce, 1982). The origin of these differences must lie in differences of overall shape of dsDNA and RNA. This could affect the net charge on the molecules (due to differential counterion interactions) and also the frictional properties. Due to its A-form helix, RNA is significantly shorter and wider than DNA, and this could be reflected in its electrophoretic mobility. We were further surprised to discover that the mobility of the equivalent DNA–RNA hybrid was very much intermediate between those of pure DNA and RNA (*Figure 4*; track 4). It has always been thought that a

DNA–RNA hybrid adopts the A-form geometry of pure RNA, and the hybrid would therefore be expected to have very similar frictional properties to dsRNA. However, these gel electrophoresis experiments suggested that the hybrid may have a geometry resulting in a shape that is much more intermediate between the A- and B-structures. These observations have been supported by recent studies from Gast and Hagerman (1991), and by Crothers and co-workers (R.W. Roberts and D.M. Crothers, submitted for publication) who have shown that the mobility of a DNA–RNA hybrid may even reflect which strand is DNA, and which is RNA.

On introduction of base bulges into RNA duplexes we (Bhattacharyya, Murchie and Lilley, 1990), and Tang and Draper (1990), observed a marked decrease in electrophoretic mobility (*Figure 4*; compare tracks 6 and 7). the retardation of the RNA depended on the position of the bulge, being greatest when the bulge was located at the centre. The extent of the retardation also depended on the size and composition of the bulge; the retardation increased as the number of bulged bases was increased, and adenine bases were more effective than uracil bases (Bhattacharyya, Murchie and Lilley, 1990). This is clearly strongly suggestive of axial kinking, closely akin to the equivalent DNA experiments. This was established by phasing experiments (Bhattacharyya, Murchie and Lilley, 1990; Tang and Draper, 1990) in which the spacing was varied between two bulges (*Figure 5A*). As expected, a sinusoi-

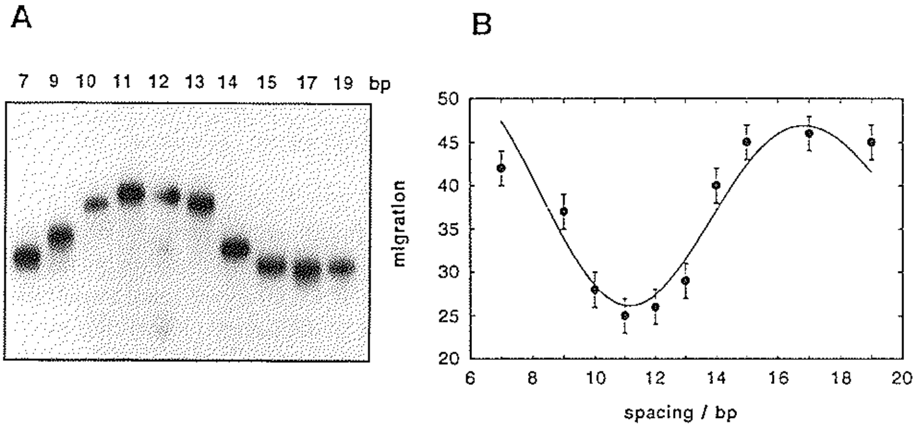
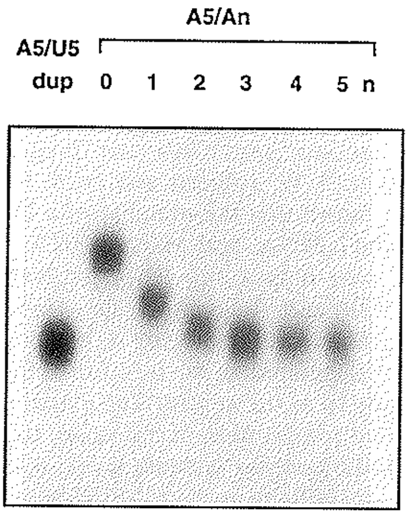


Figure 5. Variation in electrophoretic mobility as the spacing between two bulges in duplex RNA is varied; estimate of the helical repeat of dsRNA.

(A) A series of isomeric 60 bp RNA and hybrid RNA–DNA duplexes were constructed, containing two A_5 bulges on one strand, separated by 7–19 bp. These species were electrophoresed in 15% polyacrylamide. The modulation of mobility with spacing is clear.

(B)Plot showing the variation in electrophoretic mobility with the spacing between the A_5 bulges. The error bars on the data points represent the widths of the bands measured on the autoradiograph. The line was generated by non-linear regression to the model of Lumpkin and Zimm (1982), from which a helical periodicity of 11·3 bp per turn was calculated.



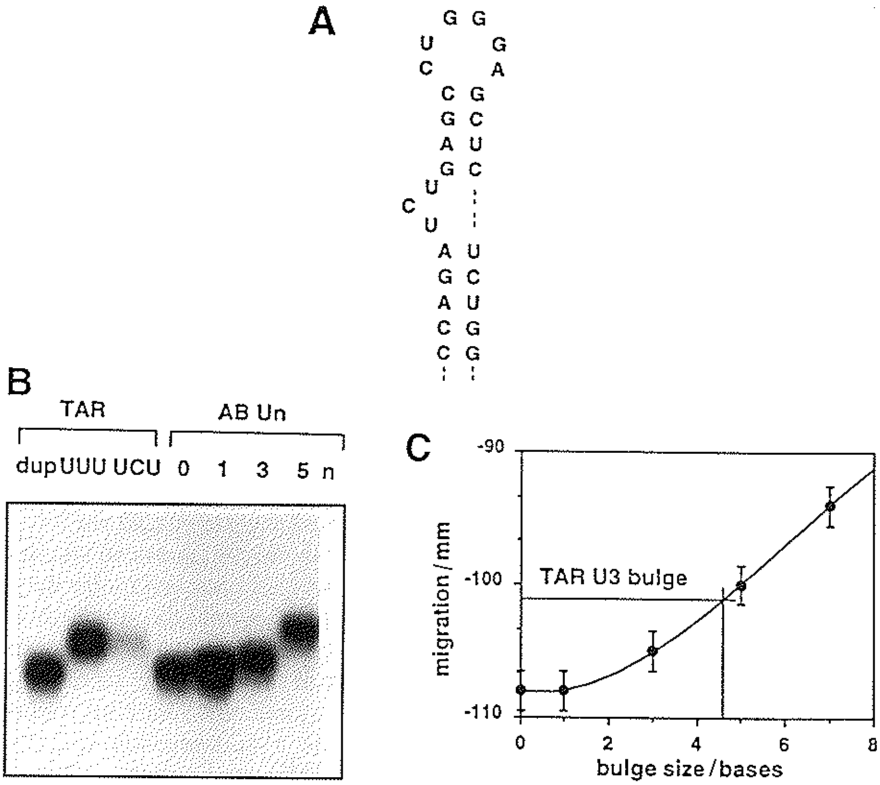
5' GGGCGUCCUGUGGAUCCAGGA_5UCGGAUCCUCUACGCCGCCC 3'
3' CCCGCAGGACACCUAGGUCCA_nAGCCUAGGAGAUGCGGCGGG 5'

Figure 6. Mismatched bases opposing a bulge reduce axial kinking. A series of double-stranded RNA species based on the AB sequence (Bhattacharyya and Lilley, 1989a) (shown below gel) containing an A_5 bulge were generated by *in vitro* transcription and hybridization. The sequence of the lower strand was varied in order to oppose the A_5 bulge with 0, 1, 2, 3, 4 or 5 adenine bases. These species were electrophoresed in 20% polyacrylamide in 90 mM Tris–borate, pH 8·3, 50 mM NaCl at 15±0·1°C, and the gel subjected to autoradiography. The left-most track contains a perfect duplex species in which the A_5 bulge is opposed by a U_5 sequence, thereby generating five consecutive A–U base pairs. The remaining tracks contain the A_5 bulged species, where the bulge is opposed by the indicated number of adenine bases. The effect of opposing the A_5 bulge is a marked increase in electrophoretic mobility, despite the increased size of the molecules.

dal modulation was found, indicative of defined axial kinking. From the least-squares fit of the whole data set to the Lumpkin–Zimm theory (*Figure 5B*), we obtained a helical periodicity of 11·3 bp per turn, consistent with an A-form helix. In a similar study Tang and Draper (1990) deduced a somewhat higher value of 11·8 bp per turn. We do not know the origin of the difference—this may reflect sequence dependence. Equivalent experiments on a DNA–RNA hybrid duplex indicated a lower periodicity, again consistent with an intermediate geometry for this duplex.

A related feature to the bulge that is found in many RNA species is the internal loop or bubble, formed by consecutive base mismatches; they may be symmetrical, or they may have more unpaired bases on one strand than the other. One such asymmetrical bubble has recently been demonstrated to undergo specific lead-induced cleavage (Pan and Uhlenbeck, 1992), and the Rev response element (RRE) that binds the HIV control protein Rev (regulator of virion expression) is a hairpin loop containing another example of an RNA bubble (Heaphy *et al.*, 1990). Our earlier studies indicated that symmetrical bubbles in DNA generated much less kinking of the helical axis (Bhattacharyya and Lilley, 1989), and we therefore examined the effect of opposed but non-complementary bases on the electrophoretic mobility of an RNA bulge-containing fragment (Riordan *et al.*, 1992). An A_5 bulge was progressively opposed by addition of adenine bases on the opposite strand, until a bubble of five consecutive A · A mispairs was generated, and the electrophoretic mobility in polyacrylamide was examined (*Figure 6*). The result of opposing the bulge was clear—despite the increase in size of the molecule, the mobility increased as the number of bases opposing the bulge was increased. Even addition of a single opposing base markedly increased the mobility, and two or more adenines conferred virtually the same mobility as the unbulged duplex. This indicates that the opposed bulge in RNA is considerably less effective in generating a local kinking of the helix axis.

Sequence specificity in DNA binding proteins is normally achieved principally by placing amino acid side-chains in the major groove, where they may donate and accept hydrogen bonds to and from the edges of the base pairs. However, in an A-form helix the axis is displaced several Ångstrom into the major groove, generating a very deep and narrow major groove in which the bases are much less accessible. This suggests that backbone geometry may be more important for RNA binding proteins, and of course the much greater structural repertoire of RNA molecules will generate many suitable candidates. What is sometimes called indirect readout is likely to be more important in RNA than for DNA. The well-defined axial kinking found in RNA bulges suggests that this might provide a recognizable target for selective protein binding (Bhattacharyya, Murchie and Lilley, 1990), and such bulges have been identified as protein binding sites in a number of cases, including the coat protein of bacteriophage R17 (Wu and Uhlenbeck, 1987). One example of special interest is the binding site for the transactivator protein (Tat) of HIV, the TAR element (transactivator response element) found attached to the 5′ end of HIV mRNA (Rosen, Sodroski and Haseltine, 1985; Feng and Holland, 1988; Jakobovits *et al.*, 1988). The sequence f the TAR element is shown in *Figure 7A*. It is a hairpin loop, the stem of which is disrupted by the presence of a number of bulges (Emmerman *et al.*, 1987; Muesing, Smith and Capon, 1987; Berkhout and Jeang, 1989). It has been found that the three-pyrimidine (either UUU or UCU) bulge closest to the loop is essential for both TAT binding and transactivation, and that the minimal protein binding site is the bulge and two basepairs on either side (Feng and Holland, 1988; Dingwall *et al.*, 1989, 1990; Roy *et al.*, 1990; Weeks *et al.*, 1990; Weeks and Crothers, 1991). We studied a duplex RNA species containing the TAR bulge and context (Riordan *et al.*, 1992), and showed that it migrated slowly in polyacrylamide consistent with an axial kink (*Figure 7B*). Indeed, it appeared as if this sequence was predisposed to being especially kinked, as a U_3 bulge in the TAR context migrated almost as slowly as a U_5 bulge in a different context. Kinking, as opposed to flexibility, was demonstrated by phasing the TAR sequence with an A_5 bulge (Riordan *et al.*, 1992), whereupon the expected modulation of electrophoretic mobility with the spacing between the bulges was observed (*Figure 8*). NMR studies of the



AB series

5' GGGCGUCCUGUGGAUCCAGG$^{U}_{n}$UCGGAUCCUCUACGCCGCCC 3'
3' CCCGCAGGACACCUAGGUCC..AGCCUAGGAGAUGCGGCGGG 5'

TAR sequences

5' GGGCGUCCUGUGGAUCCAGA^{UUU}GAGCAUCCUCUACGCCGCCC 3'
3' CCCGCAGGACACCUAGGUCU...CUCGUAGGAGAUGCGGCGGG 5'

Figure 7. The TAR sequence pyrimidine bulge causes retarded electrophoretic mobility in polyacrylamide.

(A) Central part of the base sequence of the HIV-1 TAR element drawn as a hairpin loop (Emmerman *et al.*, 1987; Muesing, Smith and Capon, 1987; Berkhout and Jeang, 1989). The full stem contains more than 20 bp, and additional single base bulges. The sequence of the three-pyrimidine bulge can be either UCU or UUU.

(B) Comparison of electrophoretic mobility of the TAR-bulge molecules and the AB series in 15% polyacrylamide, in 90 mM Tris–borate, pH 8·3, 50 mM NaCl with recirculation of buffer at 1 l h^{-1}. The gel was maintained at 15±0·1°C, and electrophoresed at 15 V cm^{-1}. The AB series comprised 40 bp duplexes interrupted by U_n bulges on the top strand, where n was 0, 1, 3, 5 or 7. We have previously characterized bending by these bulged molecules (Bhattacharyya and Lilley, 1989a). The TAR sequences were 40 bp duplexes of closely similar sequence, where the sequence flanking the bulge was that of the natural TAR RNA. The bulge sequences were either UUU or UCU. Tracks L to R: dup, TAR sequence without the three base bulge (i.e. perfect duplex species); UUU and UCU, TAR sequence with UUU or UCU bulges respectively; AB U_n, AB series of bulged duplexes containing the indicated number of bulged uracil bases.

(C) Plot of electrophoretic mobility against uridine bulge size for the AB series of molecules. The mobility of the U_3 bulge in the TAR sequence context is interpolated. Note the relatively slow migration of this species compared with an identical bulge in the AB sequence context. The error bars represent the widths of the bands measured on the autoradiograph.

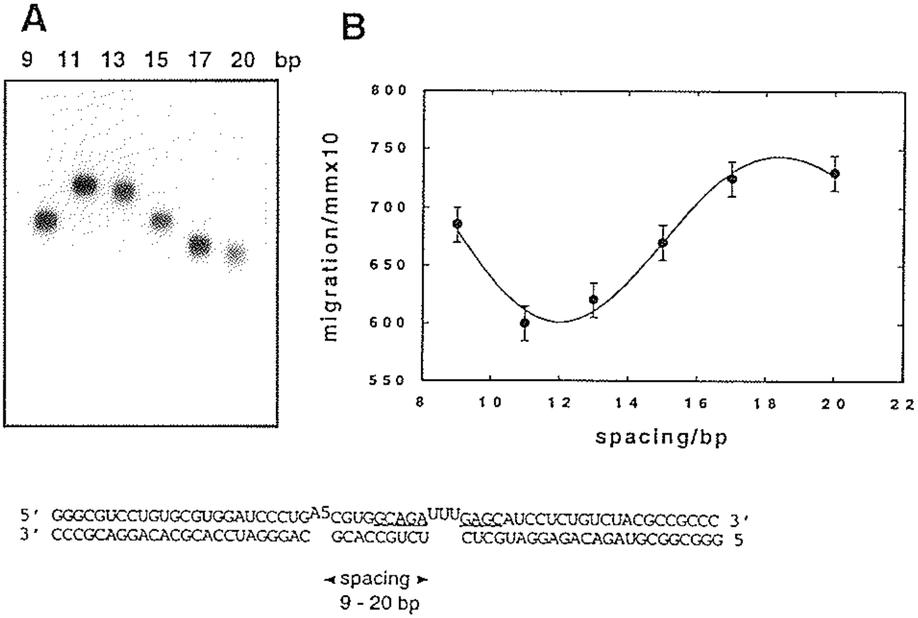


Figure 8. Phasing of base bulges shows that the TAR bulge generates a kinked helix axis. A series of 57 bp duplex species was generated in which the spacing between an A_5 bulge and the TAR U_3 bulge (underlined) was varied between 9 and 20 bp. The overall size of the molecule was kept constant through the series.

(A) Gel electrophoresis of the double-bulge series of RNA duplexes in 20% polyacrylamide in 90 mM Tris–borate, pH 8·3, 50 mM NaCl at 15±0·1°C. The spacing between bulges (bp) is indicated for each species above the track.

(B) Plot showing the variation in electrophoretic mobility with the spacing between the A_5 and TAR bulges in this series of isomeric molecules. The error bars on the data points represent the widths of the bands measured on the autoradiograph. The line was generated by non-linear regression to the model of Lumpkin and Zimm (1982). The data show clearly the expected sinusoidal variation between electrophoretic mobility and bulge spacing, confirming the presence of a defined kink at the TAR bulge.

TAR sequence by Puglisi *et al.* (1992) indicate that in the absence of a ligand the bulged pyrimidine bases are stacked into the helix, consistent with the formation of a pronounced kink. However, upon addition of arginine (a kind of ultrasimple model for the Tat protein) the bases flip out to be extrahelical, with the formation of a U · A · U base triple by one uridine. It would clearly be of interest to repeat the gel measurements in the presence of arginine.

Two groups have examined the effect of phased tracts of adenines on the electrophoretic mobility of duplex RNA (Gast and Hagerman, 1991; Wang *et al.*, 1991). In marked contrast to DNA, virtually normal mobility was observed, indicating that the A-tracts did not cause curvature of RNA. X-ray crystallographic studies of oligoadenine tracts in DNA suggested that they adopt a B′ conformation in which the A · T base pairs exhibit a large propeller twist with bifurcated hydrogen bonding, and a significant roll angle at the ends of the tract (Coll *et al.*, 1987; Nelson *et al.*, 1987). It seems probable that this conformation is not available to RNA, which is more straight as a consequence.

Nucleic acid junctions

Gel electrophoresis has been immensely valuable in the study of the conformation of DNA junctions. A number of years ago we showed that the introduction of a four-way junction into the centre of a DNA fragment reduced the electrophroetic mobility in polyacrylamide (Gough and Lilley, 1985), and that the relative retardation increased with the percentage of polyacrylamide. This was similar to the properties of kinetoplast DNA (Marini *et al.*, 1982), and was consistent with the presence of a pronounced bend or kink at the position of the junction. The electrophoretic mobility was found to be dependent on the concentration and type of cation present (Diekmann and Lilley, 1987), indicating a role for ion interactions in the structure. Cooper and Hagerman (1987) developed a technique based on the study of changes in electrophoretic mobility following extension of pairs of arms, and concluded that the symmetry of the junction was lower than tetrahedral.

We employed a closely related gel electrophoretic technique of comparing the six possible isomeric junctions with two long arms and two short arms generated by restriction cleavage. In the presence of metal cations (e.g. 100 μM Mg^{2+}), a pattern of migration was obtained in which the electrophoretic mobilities were grouped into three pairs of bands—the 2:2:2 pattern (*Figure 9*). This suggested that the junction adopts an X-shape in the presence of added cations, since there are two ways by which the long arms may be related to each other across an acute angle, two ways they may be related by an obtuse angle and lastly there are two possibilities that they may be co-linear. An X-shaped junction may be formed from a four-way junction by the pairwise coaxial stacking of helical arms, followed by a rotation in the manner of opening a pair of scissors. This generates an increase in base pair stacking interactions, while reducing steric and electrostatic interaction between the stacked pairs of arms. Upon alteration of the sequence at the point of strand exchange we observed that the electrophoretic pattern of our long-short arm junctions changed, with an exchange of mobilities for the fast and slow species (Duckett *et al.*, 1988). This is indicative that the stacking partners depend upon the local sequence at the junction, and indeed two isomers of the stacked X-structure are possible, differing in the choice of helical stacking partners.

Making the assumption that electrophoretic mobility is determined by end-to-end distance, then we assign the slower species to those in which the longer arms are related by the smallest angle of the X-structure. This means that the structure must be approximately antiparallel, i.e. the two continuous strands run in opposite directions. This is in contrast to the normal depiction of Holliday junctions, and to the model of Sigal and Alberts (1972). In the absence of added ions a quite different gel pattern was obtained (Duckett *et al.*, 1988), comprising four slow and two fast species (4:2 pattern); it is

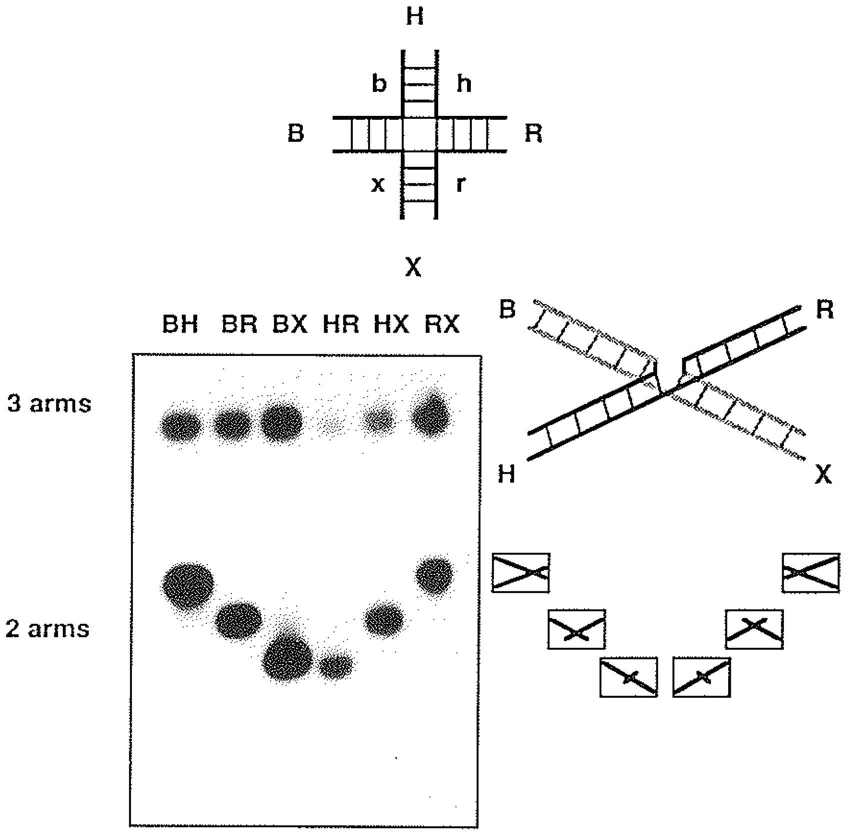


Figure 9. Gel electrophoretic analysis of a four-way junction. A junction was constructed from four 80 base oligonucleotides (strands b, h, r and x; Duckett *et al.*, 1988). Each arm (B, H, R or X) contained a unique site for a restriction enzyme, and could thus be selectively shortened. The gel electrophoretic mobility of the six double digests of the junction were compared in 5% polyacrylamide in 90 mM Tris–borate (pH 8·3) containing added cations. The six tracks contain the junctions after restriction digestion; the names refer to the uncleaved arms (e.g. BR is junction that has been cleaved in the H and X arms). The digests are incomplete, and the slower species in each track (3 arms) have suffered cleavage in one arm only. Those that have been cleaved in two arms (2 arms) exhibit three types of mobility, consistent with a two-fold symmetrical X-shape. Assuming that gel mobility is inversely proportional to the distance between the ends of the long arms (i.e. the smaller the angle between the long arms, the slower the species), this leads to the antiparallel X-shape as shown. In this junction there is coaxial stacking of the B and X arms, and a similar pairing of the H and X arms.

consistent with an extended structure in which the four arms are directed towards the corners of a square. The antiparallel, stacked X-structure has been confirmed in all essential details by other techniques, notably by comparison of fluorescence resonance energy transfer between the termini of the four arms (i.e. by comparison of the six end-to-end vector lengths; Murchie *et al.*, 1989; Clegg *et al.*, 1992). It is also consistent with results of hydroxyl radical probing (Churchill *et al.*, 1988), and transient electric birefringnce (Cooper and Hagerman, 1989).

Simple modelling suggests (Murchie *et al.*, 1989) that the most effective way to avoid stereochemical clash is by means of a right-handed, antiparallel structure, as observed experimentally, and this is confirmed by molecular mechanical energy minimization (von Kitzing, Lilley and Diekmann, 1990). In this conformation, there is a favourable alignment of strands and grooves that helps avoid steric clash between backbones, which is most effective if the small angle is 60°. Such experimental determinations of the angle in the four-way junction that have been made are all close to the theoretical expectation of 60°, and a similar packing has been observed in the crystal structure of a double-stranded oligonucleotide (Timsit *et al.*, 1989). The right-handedness of the X-structure was confirmed using FRET (Murchie *et al.*, 1989). The location of the continuous strands in the major grooves 3′ to the point of strand exchange is consistent with observed protection against cleavage by DNaseI (Lu *et al.*, 1989; Murchie *et al.*, 1990).

Three-way DNA junctions are formed from three strands of DNA, but their folding appears to be quite different from that of the four-way junction. Our long–short arm gel electrophoretic experiments (Duckett and Lilley, 1990) suggested that the angles between the three arms were more similar to each other than the inter-arm angles of the four-way junction. This indicated that the arms do not undergo the pairwise stacking of the four-way junction, and this was consistent with the chemical reactivity of thymine bases under all conditions. Model building indicates that a three-way junction in which two arms are coaxially stacked is not consistent with full base pairing.

RNA junctions

Gel electrophoretic experiments on RNA junctions are much more difficult to perform than those with DNA junctions. Apart from the general difficulties of preparing long RNA molecules for assembly into junctions, restriction enzymes cannot be used to shorten arms, and thus all species must be prepared individually. We have carried out this analysis for one three-way junction, in which individual arms were sequentially reduced from 40 bp to an 8 bp hairpin loop. The electrophoretic mobility of these three species in polyacrylamide was compared, whereupon it was observed that there was close to equal mobility whatever the ionic conditions (S. McAteer, A. Bhattacharya and D.M.J. Lilley, unpublished data). This suggested that the three-way junction is an extended, unstacked Y-structure similar to the three-way DNA junction.

The four-way junction in RNA is likely to be built on different sterochemical principles from its DNA equivalent. The structures of tRNA and pseudo-knots indicate that coaxial stacking is favourable in RNA where possible, but the overall folding is unlikely to be the right-handed, antiparallel structure found in DNA. An important aspect of the DNA stacked X-structure is the juxtaposition of strands and grooves (discussed above), but this will be quite different for an A-form helix. As a consequence, the RNA junction will probably fold according to a different principle. The RNA junction is harder to study than DNA by electrophoresis because a full long–short arm gel electrophoretic analysis requires the synthesis of many long RNA transcripts. However, we have performed a preliminary analysis on a junction of one sequence, and the results confirm that the folding appears to be quite different from that of the DNA junction. However, this requires more detailed study before a definite structure can be deduced.

Four-stranded structures

Considerable interest has been generated by the discovery of four-stranded DNA structures, based on a guanine tetrad, which can be parallel (Sen and Gilbert, 1988, 1990) or antiparallel (Sundquist and Klug, 1989; Williamson, Raghuraman and Cech, 1989). These conformations have been associated with immunoglobulin gene switching and telomere function respectively, and have been the subject of investigation by NMR (Jin *et al.*, 1990; Wang, Howard and Griffiths, 1991; Aboul-ela, Murchie and Lilley, 1993). Formation of these structures is strongly cation-dependent, and is associated with a marked alteration in gel electrophoretic mobility. Wherever oligoguanine stretches occur there seems to be a propensity to form the tetraplex structure, and this has been demonstrated for the retinoblastoma susceptibility (Murchie and Lilley, 1992) and the insulin genes (Hammond-Kosack and Docherty, 1992). This raises the question of whether these sequences might undergo tetraplex formation at the RNA level. Kim, Cheong and Moore (1991) used a gel electrophoretic retardation technique to demonstrate potassium-dependent tetraplex formation by an RNA molecule derived from the *Escherichia coli* 5S RNA, that contains a GGGG sequence. Tetraplex formation was found to be very slow, but the complexes once formed were very stable. Such interactions might form the basis of the dimerization of HIV RNA in the virion (Marquet *et al.*, 1991).

Comparison with other techniques

Gel electrophoresis has been immensely valuable in the study of the shapes of nucleic acids, especially DNA. It has generated significant insight into a number of interesting DNA structures, such as curved DNA (Diekmann and Wang, 1985; Hagerman, 1985; Koo, Wu and Crothers, 1986; Marini *et al.*, 1982; Wu and Crothers, 1984), bulges (Bhattacharyya and Lilley, 1989a; Hsieh and Griffith, 1989; Rice and Crothers, 1989; Wang and Griffith, 1991) and junctions (Cooper and Hagerman, 1987; Duckett *et al.*, 1988; Gough and Lilley, 1985). On a different size scale, agarose gel electrophoresis has been indispensable in the study of supercoiled DNA circles (Pulleyblank *et al.*, 1975; Wang, Peck and Becherer, 1983) and related topological problems such as knots and catenanes (Kanaar, van de Putte and Cozzarelli, 1988). Yet the greatest drawback to its use remains the lack of a good theoretical framework (*see* above). The most productive way to use gel electrophoresis in the study of nucleic acid structure is probably in combination with other techniques. This has certainly been true for DNA, and is likely to be even more so in the study of RNA structure. We have found that FRET has been extremely important in complementing other studies of DNA structure (Clegg *et al.*, 1992; Murchie *et al.*, 1989); it is hard to think of another technique that provides such long-range distance information in solution, and the methods for applying this to nucleic acids are improving all the time. We have made preliminary studies of RNA structure using FRET, and the number of studies will greatly increase in the immediate future. Hydrodynamic methods can be applied to RNA structure, such as light scattering, and Gast and Hagerman have recently applied transient electric birefringence to the study of RNA molecules (Gast and Hagerman, 1991). Chemical probing can provide an indication of the accessibility of bases within the secondary and tertiary structure of RNA, and a range of such probes are available (Ehresmann *et al.*, 1987). A different approach to the problem, which can generate important insight if used properly, is modelling. Clearly this can never prove a structure, but it can be useful in testing different stereochemistries, and can suggest new avenues for experimentation. Modelling RNA tertiary structure may be strengthened by the inclusion of phylogenetic data.

As a significant component of this armoury of techniques, gel electrophoresis has an important role to play in the analysis of the almost limitless number of interesting and important RNA structures in the cell.

Acknowledgements

We thank our colleagues for discussion and collaboration, especially Alastair Murchie, Derek Duckett, Fareed Aboul-ela, Fiona Riordan and Bob Clegg. We gratefully acknowledge provision of unpublished data by Professor D.M. Crothers and Dr J.D. Puglisi, and the financial support of the MRC, SERC and CRC.

References

ABOUL-ELA, F., MURCHIE, A.I.H., AND LILLEY, D.M.J. (1993). NMR study of parallel-stranded tetraplex formations by the hexadeoxynucleotide d (TG_4T). *Nature* **360**, 280–282.

ABOUL-ELA, F., MURCHIE, A.I.H., HOMANS, S.W. AND LILLEY, D.M.J. (1992). NMR study of a deoxyribonucleotide duplex containing a three base bulge. *J. Mol. Biol.* (in press).

BERKHOUT, B. AND JEANG, K.-T. (1989). *Trans*-activation of human immunodeficiency virus type 1 is sequence specific for both the single stranded bulge and loop of the *trans*-acting responsive hairpin: a quantitative analysis. *J. Virol.* **63**, 5501–5504.

BHATTACHARYYA, A. AND LILLEY, D.M.J. (1989a). The contrasting structures of mismatched DNA sequences containing looped-out bases (bulges) and multiple mismatches (bubbles). *Nucleic Acids Res.* **17**, 6821–6840.

BHATTACHARYYA, A. AND LILLEY, D.M.J. (1989b). Single base mismatches in DNA: long- and short-range structure probed by analysis of axis trajectory and local chemical reactivity. *J. Mol. Biol.* **209**, 583–597.

BHATTACHARYYA, A., MURCHIE, A.I.H. AND LILLEY, D.M.J. (1990). RNA bulges and the helical periodicity of double-stranded RNA. *Nature* **343**, 484–487.

BIEBRICHER, C.K., DIEKMANN, S. AND LUCE, R. (1982). Structural analysis of self-replicating RNA synthesised by Qβ replicase. *J. Mol. Biol.* **154**, 629–648.

BRANLANT, C., KROL, A. AND EBEL, J.-P. (1981). The conformation of chicken, rat and human U1A RNAs in solution. *Nucleic Acids Res.* **9**, 841–858.

BRIERLEY, I., DIGARD, P. AND INGLIS, S.C. (1989). Characterisation of an efficient coronovirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell* **57**, 537–547.

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

Brown, T., Hunter, W.N., Kneale, G. and Kennard, O. (1986). Molecular structure of the G.A. base pair in DNA and its implications for the mechanism of transversion mutations. *Proc. Natl. Acad. Sci. USA* **83**, 2402–2406.

Calladine, C.R., Drew, H.R. and McCall, M.J. (1988). The intrinsic curvature of DNA in solution. *J. Mol. Biol.* **201**, 127–137.

Calladine, C.R., Collis, C.M., Drew, H.R. and Mott, M.R. (1991). A study of electrophoretic mobility of DNA in agarose and polyacrylamide gels. *J. Mol. Biol.* **221**, 981–1005.

Cech, T.R. (1987). The chemistry of self-splicing RNA and RNA enzymes. *Science* **236**, 1532–1539.

Cech, T.R. and Bass, B.L. (1986). Biological catalysis by RNA. *Ann. Rev. Biochem.* **55**, 599–629.

Celander, D.W. and Cech, T.R. (1991). Visualizing the higher order folding of a catalytic RNA molecule. *Science* **251**, 401–407.

Chastain, M. and Tinoco, J.I. (1991). Structural elements in RNA. *Progr. Nucleic Acid Res. Mol. Biol.* **41**, 131–177.

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

Churchill, M.E., Tullius, T.D., Kallenbach, N.R. and Seeman, N.C. (1988). A Holliday recombination intermediate is twofold symmetric. *Proc. Natl. Acad. Sci. USA* **85**, 4653–4656.

Clegg, R.M., Murchie, A.I.H., Zechel, A., Carlberg, C., Diekmann, S. and Lilley, D.M.J. (1992). Fluorescence resonance energy transfer analysis of the structure of the four-way DNA junction. *Biochemistry* **31**, 4846–4856.

Coll, M., Frederick, C.A., Wang, A.H.-J. and Rich, A. (1987). A bifurcated hydrogen-bonded conformation in the d(A.T) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc. Natl. Acad. Sci. USA* **84**, 8385–8389.

Cooper, J.P. and Hagerman, P.J. (1987). Gel electrophoretic analysis of the geometry of a DNA four-way junction. *J. Mol. Biol.* **198**, 711–719.

Cooper, J.P. and Hagerman, P.J. (1989). Geometry of a branched DNA structure in solution. *Proc. Natl. Acad. Sci. USA* **86**, 7336–7340.

Dahlberg, A.E. (1989). The functional role of ribosomal RNA in protein synthesis. *Cell* **57**, 525–529.

de Gennes, P.G. (1971). Reptation of a polymer chain in the presence of fixed obstacles. *J. Chem. Phys.* **55**, 572–578.

Deutsch, J.M. (1988). Theoretical studies of DNA during gel electrophoresis. *Science* **240**, 922–924.

Diekmann, S. and Lilley, D.M.J. (1987). The anomalous gel migration of a stable cruciform: temperature and salt dependence, and some comparisons with curved DNA. *Nucleic Acids Res.* **14**, 5765–5774.

Diekmann, S. and Wang, J.C. (1985). On the sequence determinants and flexibility of the kinetoplast DNA fragment with abnormal gel electrophoretic mobilities. *J. Mol. Biol.* **186**, 1–11.

Dingwall, C., Ernberg, I., Gait, M.J., Green, S.M., Heaphy, S., Karn, J., Lowe, A.D., Singh, M., Skinner, M.A. and Valerio, R. (1989). Human immunodeficiency virus 1 tat protein binds trans-activation-responsive region (TAR) RNA *in vitro*. *Proc. Natl. Acad. Sci. USA* **86**, 6925–6929.

Dingwall, C., Ernberg, I., Gait, M.J., Green, S.M., Heapohy, S., Karn, J., Lowe, A.D., Singh, M. and Skinner, M.A. (1990). HIV-1 *tat* protein stimulates transcription by binding to a U-rich bulge in the stem of TAR RNA structure. *EMBO J.* **9**, 4145–4153.

DOCK-BREGEON, A.C., CHEVRIER, B., PODJARNY, A., MORAS, D., DEBEAR, J.S., GOUGH, G.R., GILHAM, P.T. AND JOHNSON, J.E. (1988). High resolution structure of the RNA duplex $[U(U\text{-}A)_6A]_2$. *Nature* **335**, 375–378.

DRAK, J. AND CROTHERS, D.M. (1991). Helical repeat and chirality effects on DNA gel electrophoretic mobility. *Proc. Natl. Acad. Sci. USA* **88**, 3074–3078.

DUCKETT, D.R. AND LILLEY, D.M.J. (1990). The three-way DNA junction is a Y-shaped molecule in which there is no helix-helix stacking. *EMBO J.* **9**, 1659–1664.

DUCKETT, D.R., MURCHIE, A.I.H. AND LILLEY, D.M.J. (1990). the role of metal ions in the conformation of the four-way junction. *EMBO J.* **9**, 583–590.

DUCKETT, D.R., MURCHIE, A.I.H., DIEKMANN, S., VON KITZING, E., KEMPER, B. AND LILLEY, D.M.J. (1988). The structure of the Holliday junction and its resolution. *Cell* **55**, 79–89.

DUCKETT, D.R., MURCHIE, A.I.H., BHATTACHARYYA, A., CLEGG, R.M., DIEKMANN, S., VON KITZING, E. AND LILLEY, D.M.J. (1992). The structure of DNA junctions, and their interactions with enzymes. *Eur. J. Biochem.* **207**, 285–295.

EHRESMANN, C., BAUDIN, F., MOUGEL, M., ROMBY, P., EBEL, J.-P. AND EHRESMANN, B. (1987). Probing the structure of RNAs in solution. *Nucleic Acids Res.* **15**, 9109–9128.

EMMERMAN, M., GUYADER, M., MONTAGNIER, L., BALTIMORE, D. AND MUESING, M.A. (1987). The specificity of the human immunodeficiency virus type 2 *trans*-activator is different from that of human immunodeficiency virus type 1. *EMBO J.* **6**, 3755–3760.

FENG, S. AND HOLLAND, E.C. (1988). HIV-1 *tat trans*-activation requires the loop sequence within TAR. *Nature* **334**, 165–168.

GAO, X. AND PATEL, D. (1987). NMR studies of A.C mismatches in DNA dodecanucleotides at acidic pH. Wobble A(*anti*).C(*anti*) pair formation. *J. Biol. Chem.* **262**, 16 973–16 984.

GAO, X. AND PATEL, D.J. (1988). G(*syn*).A(*anti*) mismatch formation in DNA dodecamers at acidic pH: pH-dependent conformational transition of G.A mispairs detected by proton NMR. *J. Am. Chem. Soc.* **110**, 5178–5182.

GAST, F.-U. AND HAGERMAN, P.J. (1991). Electrophoretic and hydrodynamic properties of duplex ribonucleic acid molecules transcribed *in vitro*: evidence that A-tracts do not generate curvature in RNA. *Biochemistry* **30**, 4268–4277.

GOUGH, G.W. AND LILLEY, D.M.J. (1985). DNA bending induced by cruciform formation. *Nature* **313**, 154–156.

HAGERMAN, P.J. (1985). Sequence dependence on the curvature of DNA: a test of the phasing hypothesis. *Biochemistry* **24**, 7033–7037.

HAMMOND-KOSACK, M.C.U. AND DOCHERTY, K. (1992). A consensus repeat sequence from the human insulin gene linked polymorphic region adopts multiple quadriplex DNA structures *in vitro*. *FEBS Lett.* **301**, 79–82.

HEAPHY, S., DINGWALL, C., ERNBERG, I., GAIT, M.J., GREEN, S.M., KARN, J., LOWE, A.D., SINGH, M. AND SKINNER, M.A. (1990). HIV-1 regulator of virion expression (Rev) protein binds to a stem-loop structure located within the Rev Response Element region. *Cell* **60**, 685–693.

HOLBROOK, S.R., CHEONG, C., TINOCO JR, I. AND KIM, S.-H. (1991). Crystal structure of an RNA double helix incorporating a track of non-Watson–Crick base pairs. *Nature* **353**, 579–581.

HOLBROOK, S.R., SUSSMAN, J.L., WADE WARRANT, R., CHURCH, G.M. AND KIM, S.-H. (1977). RNA–ligand interactions: (I) Magnesium binding sites in yeast tRNAPhe. *Nucleic Acids Res.* **8**, 2811–2820.

HSIEH, C.-H. AND 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.

HUNTER, W.N., BROWN, T., ANAND, N.N. AND KENNARD, O. (1986). Structure of an adenine-cytosine base pair in DNA and its implications for mismatch repair. *Nature* **320**, 552–555.

JACK, A., LADNER, J.E., RHODES, D., BROWN, R.S. AND KLUG, A. (1977). A crystallographic study of metal-binding to yeast phenyalanine transfer RNA. *J. Mol. Biol.* **111**, 315–328.

JAKOBOVITS, A., SMITH, D.H., JAKOBOVITS, E.B. AND CAPON, D.J. (1988). A discrete element 3′ of human immunodeficiency virus 1 (HIV-1) and HIV-2 mRNA initiation sites mediates transcriptional activation by an HIV *trans*-activator. *Mol. Cell. Biol.* **8**, 2555–2561.

JIN, R., BRESLAUER, K.J., JONES, R.A. AND GAFFNEY, B.L. (1990). Tetraplex formation of a guanine-containing nonameric DNA fragment. *Science* **250**, 543–546.

KANAAR, R., VAN DE PUTTE, P. AND COZZARELLI, N.R. (1988). Gin-mediated DNA inversion: product structure and the mechanism of strand exchange. *Proc. Natl. Acad. Sci. USA* **85**, 752–756.

KIM, J., CHEONG, C. AND MOORE, P.B. (1991). Tetramerization of an RNA oligonucleotide containing a GGGG sequence. *Nature* **351**, 331–332.

KNEALE, G., BROWN, T., KENNARD, O. AND RABINOVICH, D. (1985). G.T base pairs in a DNA helix: The crystal structure of d(G–G–G–G–T–C–C–C). *J. Mol. Biol.* **186**, 805–814.

KOO, H.-S., WU, H.-M. AND CROTHERS, D.M. (1986). DNA bending at adenine–thymine tracts. *Nature* **320**, 501–506.

LERMAN, L.S. AND FRISCH, H.L. (1982). Why does the electrophoretic mobility of DNA in gels vary with the length of the molecule. *Bipolymers* **21**, 995–997.

LEVENE, S.D. AND ZIMM, B.H. (1989). Understanding the anomalous electrophoresis of bent DNA molecules: a reptation model. *Science* **245**, 396–399.

LU, M., GUO, Q., SEEMAN, N.C. AND KALLENBACH, N.R. (1989). DNaseI cleavage of branched DNA molecules. *J. Biol. Chem.* **264**, 20 851–20 854.

LUMPKIN, O.J. AND ZIMM, B.H. (1982). Mobility of DNA in gel electrophoresis. *Biopolymers* **21**, 2315–2316.

MARINI, J.C., LEVENE, S.D., CROTHERS, D.M. AND ENGLUND, P.T. (1982). Bent helical structure in kinetoplast DNA. *Proc. Natl. Acad. Sci. USA* **79**, 7664–7668.

MARQUET, R., BAUDIN, F., GABUS, C., DARLIX, J.-L., MOUGEL, M., EHRESMANN, C. AND EHRESMANN, B. (1991). Dimerization of human immunodeficiency virus (type 1) RNA: stimulation by cations and possible mechanism. *Nucleic Acids Res.* **19**, 2349–2357.

MILLIGAN, J.F., GROEBE, D.R., WITHERALL, G.W. AND UHLENBECK, O.C. (1987). Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res.* **15**, 8783–8798.

MUESING, M.A., SMITH, D.H. AND CAPON, D.J. (1987). Regulation of mRNA accumulation by a human immunodeficiency virus *trans*-activator protein. *Cell* **48**, 691–701.

MURCHIE, A.I.H. AND LILLEY, D.M.J. (1992). Retinoblastoma susceptibility genes contain 5′ sequences with a high propensity to form guanine-tetrad structures. *Nucleic Acids Res.* **20**, 49–53.

MURCHIE, A.I.H., CLEGG, R.M., VON KITZING, E., DUCKETT, D.R., DIEKMANN, S. AND LILLEY, D.M.J. (1989). Fluorescence energy transfer shows that the four-way DNA junction is a right-handed cross of antiparallel molecules. *Nature* **341**, 763–766.

MURCHIE, A.I.H., CARTER, W.A., PORTUGAL, J. AND LILLEY, D.M.J. (1990). The tertiary structure of the four-way DNA junction affords protection against DNaseI cleavage. *Nucleic Acids Res.* **18**, 2599–2606.

NELSON, H.C.M., FINCH, J.T., LUISI, B.F. AND KLUG, A. (1987). The structure of an oligo(dA).oligo(dT) tract and its biological implications. *Nature* **330**, 221–226.

NIKONOWICZ, E.P. AND PARDI, A. (1992). Three dimensional heteronuclear NMR studies of RNA. *Nature* **355**, 184–186.

NOLLER, H.F., HOFFARTH, V. AND ZIMNIAK, L. (1992). Unusual resistance of peptidyl transferase to protein extraction procedures. *Science* **256**, 1416–1419.

PAN, T. AND UHLENBECK, O.C. (1992). A small metalloribozyme with a two-step mechanism. *Nature* **358**, 560–563.

PLEIJ, C.W.A., RIETVELD, K. AND BOSCH, L. (1985). A new principle of RNA folding based on pseudoknotting. *Nucleic Acids Res.* **13**, 1717–1731.

PUGLISI, J.D., TAN, R., CALNAN, B.J., FRANKEL, A.D. AND WILLIAMSON, J.R. (1992). Conformation of the TAR RNA-arginine complex by NMR spectroscopy. *Science* **257**, 76–80.

PULLEYBLANK, D.E., SHURE, M., TANG, D., VINOGRAD, J. AND VOSBERG, H.-P. (1975). Action of nicking-closing enzyme on supercoiled and nonsupercoiled closed circular DNA: Formation of a Boltzmann distribution of topological isomers. *Proc. Natl. Acad. Sci. USA* **72**, 4280–4284.

PYLE, M.A. AND CECH, T.R. (1991). Ribozyme recognition of RNA by tertiary interactions with specific 2′-OH groups. *Nature* **350**, 628–631.

QUIGLEY, G.J., TEETER, M.M. AND RICH, A. (1978). Structural analysis of spermine and magnesium ion binding to yeast phenylalanine transfer RNA. *Proc. Natl. Acad. Sci. USA* **75**, 64–68.

RICE, J.A. AND CROTHERS, D.M. (1989). DNA bending by the bulge defect. *Biochemistry* **28**, 4512–4516.

RICE, J.A., CROTHERS, D.E., PINTO, A.L. AND LIPPARD, S.J. (1988). The major adduct of the antitumor drug *cis*-diamminedichloroplatinum (II) with DNA bends the duplex by ~40° towards the major groove. *Proc. Natl. Acad. Sci. USA* **85**, 4158–4161.

RIORDAN, F.A., BHATTACHARYYA, A., MCATEER, S. AND 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.

ROSEN, C.A., SODROSKI, J.G. AND HASELTINE, W.A. (1985). The location of *cis*-acting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. *Cell* **41**, 813–823.

ROSEN, M.A., SHAPIRO, L. AND PATEL, D.J. (1992). Solution structure of a trinucleotide A-T-A-Bulge loop within a DNA duplex. *Biochemistry* **31**, 4015–4026.

ROY, S., PARKIN, N.T., ROSEN, C., ITOVICH, J. AND SONENBERG, N. (1990). Structural requirements for *trans* activation of human immunodeficiency virus type 1 long terminal repeat-directed gene expressin by Tat: importance of base pairing, loop sequence and bulges in the Tat-responsive sequence. *J. Virol.* **64**, 1402–1406.

SEN, D. AND GILBERT, W. (1988). Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* **334**, 364–366.

SEN, D. AND GILBERT, W. (1990). A sodium-potassium switch in the formation of four-stranded G4-DNA. *Nature* **344**, 410–414.

SIGAL, N. AND ALBERTS, B. (1972). Genetic recombination: the nature of crossed strand-exchange between two homologous DNA molecules. *J. Mol. Biol.* **71**, 789–793.

SUNDQUIST, W.I. AND KLUG, A. (1989). Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature* **342**, 825–829.

TANG, R.S. AND DRAPER, D.E. (1990). Bulge loops used to measure the helical twist of RNA in solution. *Biochemistry* **29**, 5232–5237.

TIMSIT, Y., WESTHOF, E., FUCHS, R.P.P. AND MORAS, D. (1989). Unusual helical packing in crystals of DNA bearing a mutation hot spot. *Nature* **341**, 459–462.

TINOCO JR, I., PUGLISI, J.D. AND WYATT, J.R. (1990). RNA folding. *Nucleic Acids Mol. Biol.* **4**, 205–226.

TUERK, C., GAUSS, P., THERMES, C., GROEBE, D.R., GUILD, N., STORMO, G., GAYLE, M., D'AUBERTON-CARAFA, Y., UHLENBECK, O.C., TINOCO JR, I. AND GOLD, L. (1988). CUUCGG hairpins: extraordinarily stable RNA secondary structure associated with various biochemical processes. *Proc. Natl. Acad. Sci. USA* **85**, 1364–1368.

UHLENBECK, U.C. (1987). A small catalytic oligoribonucleotide. *Nature* **328**, 596–600.

VON KITZING, E., LILLEY, D.M.J. AND DIEKMANN, S. (1990). The stereochemistry of a four-way DNA junction: a theoretical study. *Nucleic Acids Res.* **18**, 2671–2683.

WANG, J.C., PECK, L.J. AND BECHERER, K. (1983). DNA supercoiling and its effects on DNA structure and function. *Cold Spring Harbor Symp. Quant. Biol.* **47**, 85–91.

WANG, Y.-H. AND GRIFFITH, J. (1991). Effects of bulge composition and flanking sequence on the kinking of DNA by bulged bases. *Biochemistry* **30**, 1358–1363.

WANG, Y.-H., HOWARD, M.T. AND GRIFFITH, J.D. (1991). Phased adenine tracts in double-stranded RNA do not induce sequence-directed bending. *Biochemistry* **30**, 5443–5449.

WANG, Y.-H., DO LOS SANTOS, C., GAO, X., GREENE, K., LIVE, D. AND PATEL, D.J. (1991). Multinuclear nuclear magnetic resonance studies of Na cation-stabilised complex formed by d(G–G–T–T–T–C–G–G) in solution. *J. Mol. Biol.* **222**, 819–832.

WEEKS, K.M., AMPE, C., SCHULTZ, S.C., STEITZ, T.A. AND CROTHERS, D.M. (1990). Fragments of HIV-1 Tat protein specifically bind TAR RNA. *Science* **249**, 1281–1285.

WEEKS, K.M. AND CROTHERS, D.M. (1991). RNA recognition by Tat-derived peptides: interaction in the major groove? *Cell* **66**, 577–588.

WILLIAMSON, J.R., RAGHURAMAN, M.K. AND CECH, T.R. (1989). Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell* **59**, 871–880.

WU, H.-M. AND CROTHERS, D.M. (1984). The locus of sequence-directed and protein-induced DNA bending. *Nature* **308**, 509–513.

WU, H.-N. AND UHLENBECK, O.C. (1987). Role of a bulged A residue in a specific RNA protein interaction. *Biochemistry* **26**, 8221–8227.

ZINKEL, S.Z. AND CROTHERS, D.M. (1991). Catabolite activator protein-induced DNA bending in transcription initiation. *J. Mol. Biol.* **219**, 201–215.