Nucleic Acids Research, 1994, Vol. 22, No. 5 835–841

On the use of phasing experiments to measure helical repeat and bulge loop-associated twist in RNA

Roderick S.Tang+ and David E.Draper*
Department of Chemistry, Johns Hopkins University, Baltimore, MD 21218, USA

Received October 5, 1993; Revised and Accepted January 19, 1994

ABSTRACT

In a phasing experiment, two bends are introduced into a long duplex RNA or DNA and the number of base pairs between them varied. When electrophoresed in a gel, the set of molecules may show a periodic variation in mobility that contains information about the twist associated with the bends and the intervening helix. We show how a set of three phasing experiments can be used to extract this information, and apply it to an RNA helix bend at the bulge sequence A_2. The bulge introduces a negative (left-handed) twist of ~30°; at low temperatures, it is mostly confined to the 5′ side of the bulge. The apparent helical repeat of random sequence RNA measured in these experiments was 10.2 ± 0.1 base pairs, an unexpectedly low value. It is likely that moderate curvature of the RNA helix axis (30–40° over 80 bp) has affected the measurement.

INTRODUCTION

DNA bending has been a topic of considerable interest since the discoveries of 'A tract' sequences having intrinsic bend (1, 2) and the ability of many DNA binding proteins to induce bends (3). Similar issues arise when duplex RNA structure is considered: the helix axis may certainly be bent by some of the irregularities that are so frequently encountered (such as bulge loops or mismatches), and proteins may also distort the A form helix or alter the bending associated with an irregular feature. Indeed, protein binding sites in RNA are frequently centered on bulge and internal loops in helices; the HIV tat (4, 5) and rev proteins (6), R17 phage coat protein (7), and ribosomal protein S8 (8) are all recently studied examples. Bulge and internal loops in RNA may not only bend the helix axis, but may also alter the helical twist and introduce flexible joints. These alterations could facilitate tertiary interactions elsewhere in the RNA or allow particular conformational transitions. Though potentially important, these kinds of effects of bulge and internal loops on RNA helix structure and dynamics are relatively unexplored.

Gel-electrophoresis methods have been extensively applied in the study of DNA bends. Current theories predict that the bending of a stiff rod should reduce its electrophoretic mobility in gels (9), and the degree of bending has been experimentally correlated with the degree of retardation (10). In a 'phasing' experiment, two bends are placed in a helix and the number of base pairs between them varied. The sinusoidal variation of gel mobility with the number of base pairs reflects the relative orientation of the two bends (the *cis* configuration having a shorter end-to-end distance and slower gel mobility than the *trans*) and the helical repeat of the intervening base pairs. This kind of experiment has been used to detect the presence of a static or protein-induced bend and to measure the relative 'direction' of two bending sites (3).

Here we use an A_2 bulge sequence to show how information about RNA helix bend and twist may be extracted from a set of phasing experiments. Discussions of DNA bending frequently specify only two variables, the bend angle and bend direction, and implicitly assume that the bend-associated twist is not significantly different from that of the standard DNA helix. We find that this assumption is unwarranted for bulges in RNA, and show how phasing experiments can be used to measure the twist directly. In the course of this work we also made an attempt to measure the helical repeat of random sequence RNA, and obtained an answer that is almost certainly too small (10.2 ± 0.1 base pairs). We suggest that duplex RNAs may have significant curvature over the space of several helical repeats, and that this factor has biased the measurement.

MATERIALS AND METHODS

RNA synthesis

RNAs were prepared by transcribing derivatives of pT7/73-18 (BRL), a plasmid containing both T7 and T3 RNA polymerase promoters separated by multiple restriction sites. Synthetic DNA oligomers were cloned into pT7/T3-18 plasmids and the sequences confirmed by the dideoxy method (11). Plasmids were purified by a CsCl density gradient protocol which effectively removes contaminating RNA; the crude nucleic acid is dissolved in a concentrated solution of CsCl (6 gm CsCl plus 4.0 ml solution) and layered underneath a less dense solution (12 gm CsCl plus 12.0 ml buffer). During centrifugation the plasmid DNA rises to its equilibrium density, while the RNA remains behind in a pellet. Plasmids were cut with either *Pvu*II for transcription with T3 RNA polymerase or *Ssp*I for transcription

*To whom correspondence should be addressed

+Present address: Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309, USA

with T7 RNA polymerase (both polymerases purified in this laboratory). The linear plasmids were transcribed, and the full length RNA transcripts were isolated from an 8% non-denaturing polyacrylamide gel (12). Usually one of the transcripts was labeled by [α-^{32}P] ATP incorporation. The isolated RNA was purified by phenol extraction and ethanol precipitation.

Preparation of duplex RNAs

Duplex RNA was formed by mixing equimolar amounts of T7 and T3 transcripts in 10 μl of 0.2 M NaCl, 10 mM Tris pH 7.6 and 1 mM Na_2EDTA. The sample was then incubated for 15 minutes at 65°C and cooled to 37°C in 15 minutes. The RNAs transcribed from plasmids are longer than the duplex sequence, so the single stranded tails were trimmed by adding 20 units of T_1 RNase (Boehringer Mannheim) to the annealed RNA, and incubating at 37°C for 1 to 1.5 hours. Sequences of the duplex RNAs made by this transcription and nuclease digestion protocol are shown in Figure 1.

Preparation of duplex RNAs from PCR-generated DNA

To reduce the number of plasmid clones which must be isolated and sequenced for a phasing experiment, we devised a method to generate the needed DNA templates by the polymerase chain reaction (PCR). First, a plasmid was made that had the T7 RNA polymerase promoter sequence followed by the sequence for one of the arms, the bulge sequence, and a *SalI* restriction site; this DNA fragment could be purified in quantity from a *SalI*−*PvuII* digest of the plasmid. Next, a plasmid containing the other arm sequence, followed by the bulge and a fragment of pBR322 DNA, was amplified by PCR; by using various primers, the length of pBR322 sequence included was varied and terminated with an *XhoI* restriction site. This fragment was cleaved with *XhoI*, ligated to the *PvuII*−*SalI* fragment, and the ligation mixture re-cut with *XhoI* and *SalI*. The latter two enzymes share a four base overhang, so the result of this procedure is that only the two desired fragments become ligated. Upon transcription the desired length RNA is easily separated from the transcripts of any remaining *PvuII*−*SalI* DNA. A similar set of plasmid clones yielded the complementary strand lacking the A_2 sequences, and annealing of the two strands was carried out as described above. One of the two strands was transcribed with [α-^{32}P] ATP present. The annealed products have 7 and 9 base 3′ overhangs; digestion with T_1 RNase, which reduced the overhang lengths to 2 bases each, increased the mobilities of the RNA fragments slightly but did not alter the relative mobilities.

Electrophoresis and data analysis

The RNA samples were electrophoresed in 12% non-denaturing polyacrylamide gels which were 40 cm long and 0.4 mm thick. The buffer used in the electrophoresis was 75 mM Na_2HPO_4 and 5 mM Na_2EDTA adjusted to pH 7.0 with phosphoric acid; 2 mM $MgCl_2$ was also added in some cases. Gels were run at either room temperature or in a cold box, and were at 30° or 8°C, respectively, as measured by a thermometer attached to the glass plate. 150 V was used for 18 hours with circulation of the upper and lower buffer reservoirs to maintain constant pH. After autoradiography the distances of RNA migration were measured with a caliper. Control RNAs having the same number of base pairs were run alongside or in the same lane as the bulge-containing RNAs. The relative mobilities reported are the ratios of bulge-containing to control RNA migration distance.

For phasing experiments, the relative mobility data were plotted as a function of helix length, and fitted to a periodic function to determine the positions of the maxima and minima. All fits were made to averages of two independent experiments. An equation for the end-to-end distance of an RNA with two joints and three helix segments has been previously derived and used for this purpose (13). In some instances it has proved easier to use the equation

$$y = (mb + a)(1 - \cos(\phi)) + s \quad (1)$$

where b is the number of base pairs in the middle helix segment, the factor (mb + a) allows linear increase or decrease in the amplitude of the sine curve as the size of the RNA changes, and s offsets the curve on the y axis. The angle ϕ is a function of the helical repeat and bulge-associated twist, as discussed in Results. Least squares fitting of the curve to data sets was done by the Marquardt method (14).

RESULTS

Phasing observed with A_2 RNA bulges

Bulge loops reduce the gel-electrophoretic mobility of duplex RNAs (13, 15, 16). This observation alone does not necessarily mean that the RNA is bent, since a flexible joint may reduce the time averaged end-to-end distance of a duplex without making a static bend. A phasing experiment can help decide this question. When two bends are present in an RNA, the conformation of the helix axis will go from *nocis* to *trans* and back again as a full helical repeat (11−12 base pairs) of helix is added between the bends. There are two requirements of the bends for this behavior to be observed. First, they must have permanent bend; joints with hinge-like flexibility but no permanent bend will give a periodicity of ~6 base pairs instead of 12. (Hinge-like flexibility will not eliminate the periodicity, as long as the average conformation is bent.) Second, the bends must be torsionally stiff, otherwise the distinction between *cis* and *trans* will be lost.

In previous work we found that A_2, A_3, A_5, U_3, and U_5 bulges all gave approximately 11 base pair periodicity in phasing experiments (13). Similarly, Riordan *et al.* (16) found phasing of A_5 and UCU bulge sequences. All of these bulges must therefore induce some static bend in the helix. A U_2 bulge did not show any phasing, though there was a significant decrease in gel mobility for the RNAs containing two U_2 bulges (13); this is consistent with a flexible joint having negligible intrinsic bend. Although 3 and 5 base bulges show phasing, the curves are flattened and maxima and minima cannot be accurately determined; we have therefore chosen the A_2 bulge for an investigation of the phasing experiment.

A phasing experiment for RNAs with 21 to 38 base pairs between two A_2 bulges is shown in Figure 2; the RNA is diagrammed at the top of Figure 1. Similar data have been shown previously (13); in this experiment longer gels have been run in an attempt to obtain more accurate data. The diminishing amplitude of the gel retardation with increasing length is expected, since the lengths of the duplex 'arms' become a smaller proportion of the total RNA length. A phasing equation which allows for this attenuation has been fit to the curve (see Materials and Methods). While there is clearly a sinusoidal pattern to the graph, there are reproducible deviations from a smooth line which we attribute to sequence-dependent differences in the helical twist as the RNA is lengthened.

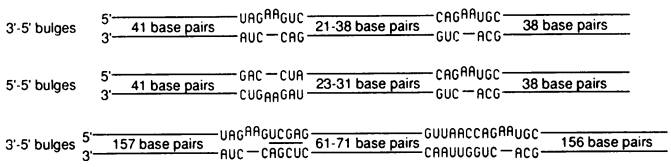

Figure 1. Sequences of bulge-containing RNA duplexes used for phasing experiments. Two duplexes have two A_2 bulges in the same strand (3′−5′ bulge) and differ only in length. The complete sequence of the shorter one is given in (13). The 71 base pair segment of the longer one is positions 445−375 of pBR322 plasmid DNA, and the underlined sequence indicates the position of the *SalI*/*XhoI* restriction sites used to construct the DNA template. The 5′−5′ bulge RNA has had one A_2 bulge and flanking base pairs switched to the opposite strand.

Table I. Phasing experiments with 5′−3′ A_2 bulges[a]

Conditions	R (bp)	ξ (deg)	Maximum (bp)	ξ (R = 11)
30°	11.4	−13	28.92	−46
30°, + Mg	11.1	−24	28.48	−32
8°	10.4	−81	28.34	−28
8° + Mg	11.9	+44	28.43	−30

[a]R (the helical repeat, expressed in base pairs per turn) and ξ (the bulge-associated twist, in degrees) are calculated from a least squares fit of equation (1) of ref. 13 to data as shown in Figure 2B, and are defined by equation (2) in Results. The fourth column is the number of base pairs inserted between the two bulges that gives the maximum relative gel mobility, as calculated from the fitted curve. The last column is a calculation of ξ made with R fixed at 11 base pairs. The reproducibility in the position of the maximum is ± 0.1 bp for these data sets. Four significant figures are shown for the maximum only for comparison with Tables II and III.

The set of RNAs with two A_2 bulges has been run under four solution conditions (30° or 8°C, ± 2 mM $MgCl_2$). For each experiment the position of the maximum gel mobility as fit from the phasing equation is listed in Table I, along with the helical repeat and bulge-associated twist. The latter two values come from the following consideration. The introduction of a bulge into a duplex RNA may alter the helix twist at that point, as well as cause bending. When two bulges are put into an RNA helix in a phasing experiment, this bulge-associated twist will add to the helical twist of the duplex between the bulges. Thus the torsion angle defined by the two arms of the three segment RNA is

$$\phi = \frac{2\pi b}{R} + \xi \qquad (2)$$

where b is the number of base pairs in the middle segment, R is the helical repeat of the duplex in base pairs per turn, and ξ is the twist associated with the two bulges. (The relation of ξ to bend geometry is discussed in the next section.) φ is π radians when the arms are in the *trans* conformation and the gel mobility is at a maximum. Changes in R alter the number of base pairs between the maxima, while ξ shifts the phasing curve along the x axis; both ξ and R could be determined from a phasing experiment of sufficient accuracy.

In practice, the uncertainties in ξ and R obtained from a single phasing experiment are quite large. R is determined by the number of base pairs separating two maxima or minima. Because the amplitude of the phasing curve diminishes rapidly, the position of the second maximum in our experiments is not accurately determined (see Figure 2B). We therefore do not find the apparent

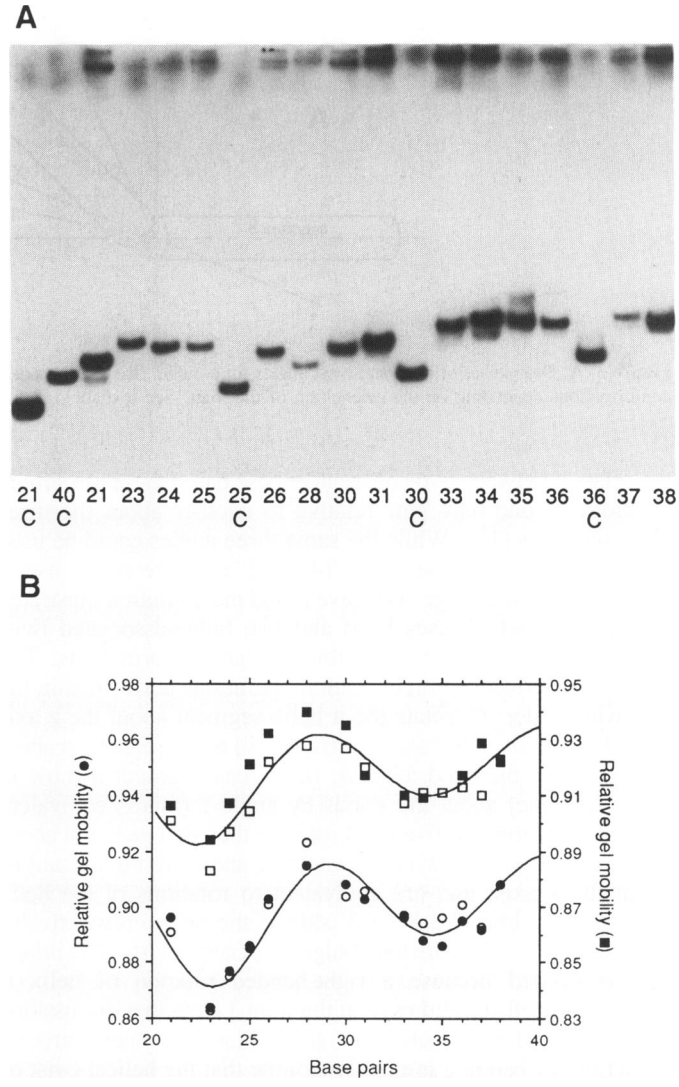

Figure 2. Phasing experiment with two A_2 bulges. **A**, gel run in sodium phosphate buffer at 30°C. The number of base pairs between the bulges is indicated; numbers followed by C are control RNAs lacking the two bulges. **B**, relative gel mobilities are plotted for gels run at 30° in sodium phosphate buffer (○,●) or 8° in the same buffer plus $MgCl_2$ (□,■). Open and closed symbols are the results of two independent experiments. Lines are least squares best fits of equation (1) of (13) to the averaged data; the best fit helical repeat and bulge-associated twist are listed in Table I.

1 bp/turn variation in the helical repeat with solution conditions to be significant, and below present data to show that it varies by less than 0.1 bp/turn over this range of conditions. ξ in turn depends on an accurate value for R. The last column of Table I therefore shows the bulge-associated twist (ξ) calculated if the helical repeat is fixed at 11.0 base pairs.

Description of bulge-associated twist

A bulge (or other helix defect) affects the three dimensional conformation of a helix and must be described by three parameters. Discussions of bending in DNA helix sequences usually use helical twist, tilt, and roll as the three angles defining

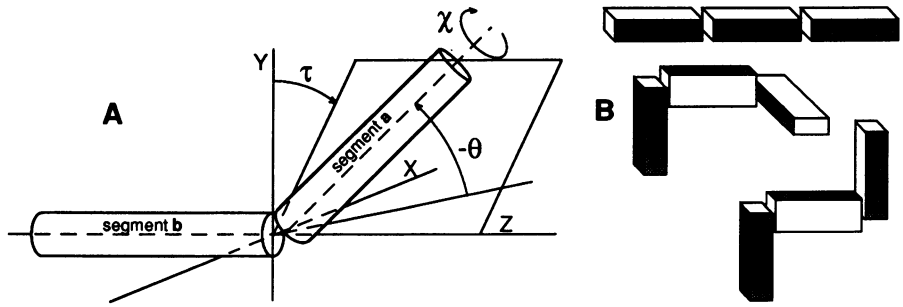

Figure 3. A, Definitions of bend and twist angles for a bulge. See text for a description of order of rotation operations. B, bend and twist operations produce different conformations depending on the orientation of the joint. See text for a discussion.

the relation between two consecutive base pairs; these are simply rotations of one base pair relative to another about the three Cartesian axes (17). While the same three angles could be used to describe a bulge, the terms 'tilt' and 'roll' are not as useful in the context of a bulge. We have found the definition illustrated in Figure 3, which uses bend and two bulge-associated twist angles, more useful for describing bulge conformations. The angles are defined as three rotation operations performed in the following order: (i) rotate the **a** helix segment about the z axis by χ (keeping the other segment fixed); (ii) rotate segment **a** about the x axis by the bend angle θ; (iii) rotate segment **a** (now in the y−z plane) about the z axis by angle τ (this is equivalent to keeping segment **a** fixed and rotating the other segment about its helical axis). The two twist angles χ and τ are both rotations about the z axis, and are equivalent to rotations of the helix segments on either the 5′ or 3′ side of the bulge, respectively. The sum $-(\chi + \tau)$ is the total bulge-associated twist. (The minus sign is needed because a right-handed rotation of helix **a** introduces a left-handed twist at the joint.) Note that discussions of DNA bending which refer to a minor or major groove 'direction' of bending implicitly assume that the helical twist of the DNA is not altered through the bend (e.g., ref. 18). This assumption is unwarranted for bulge-containing helices, for which substantial twisting is feasible.

When two bulges are present in the same strand of an RNA (as they are in the phasing experiment of Table I and Figure 2), the angle ξ that contributes to the torsion angle between the two end segments of helix is $-(\chi + \tau)$; the 5′ side of one bulge contributes $-\chi$ to the twist, and the 3′ side of the other bulge contributes $-\tau$. Thus the value of ξ calculated in Table I is the total bulge-associated twist.

To see how χ and τ contribute individually to the torsion angle ξ, the following exercise is helpful. Line up three blackboard erasers as shown at the top of Figure 3B (the top side is white, the bottom black). Starting with the right hand eraser, rotate it clockwise about its long axis by 90°, and then bend it upwards 90°. Repeat the operation with the middle eraser, to obtain the middle conformation shown. This corresponds to having two bends of 90° and $\chi = 90°$, $\tau = 0°$; ξ is $-(\chi + \tau) = -90°$. Now suppose the 5′−3′ orientation of the right joint is reversed. The rotation operations for this bend are also reversed in order: first bend the right eraser up by 90°, then rotate it 90° about the long axis of the middle eraser (bottom Figure 3B). ξ is now $-2\chi = -180°$. Thus the individual contributions of the two bulge-associated twist angles can be determined if the orientation of one bulge is reversed end for end.

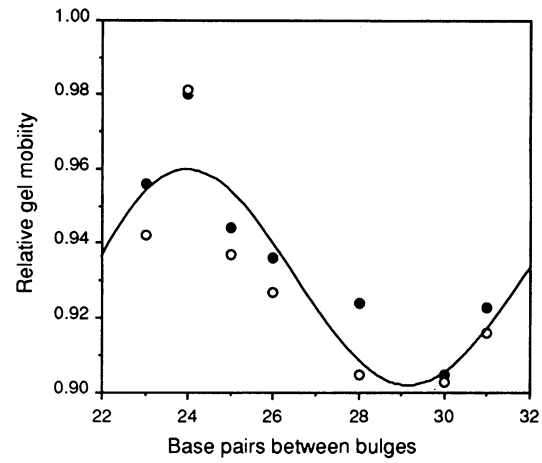

Figure 4. Phasing experiment with two A_2 bulges oriented 5′−5′. Data from duplicate runs at 30°C in sodium phosphate buffer have been fit as in Figure 2; parameters are listed in Table II.

Table II. Phasing experiment with 5′−5′ A_2 bulges[a]

conditions	minimum (bp)	$\xi-180°$	$-\chi$ (5′)	$-\tau$ (3′)
30°	29.28	−58	−29	−17
30°, + Mg	28.84	−44	−22	−10
8°	29.10	−52	−26	−2
8° + Mg	29.31	−59	−29	−1

[a]The gel mobility data were fit as described in Table I to data as shown in Figure 4; because the data span only 9 base pairs, the helical repeat was fixed at 11 base pairs. The second column is the number of base pairs inserted between the bulges that give a minimum relative gel mobility, as calculated from the fitted curve. 180° has been subtracted from the reported value of ξ to correct for the two bulges being in different strands; the value is directly comparable to ξ in Table I. χ and τ are twists (in degrees) associated with either the 5′ or 3′ side of the A_2 bulge and are calculated from ξ of Table I (with R = 11) and ξ of this Table, as described in the text. The standard error in determining the maximum is ±0.27 bp, and the overall uncertainty in the twist angles is ±6°.

A phasing experiment in which one of the two A_2 bulges has been placed on the opposite strand from the other is shown in Figure 4. The RNA is shown in Figure 1, and is named 5′−5′ since the 5′ sides of the bulges are facing each other. In this RNA $\xi + 180 = -2\chi$. (The 180° correction is necessary to account for the fact that the bulge is in the opposite strand, which, in the absence of bulge-associated twist, will change the *cis*

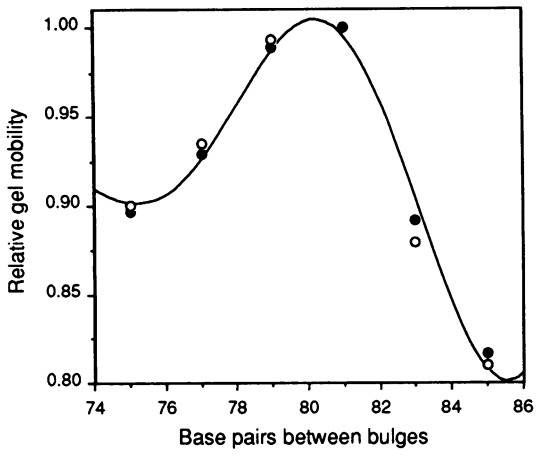

Figure 5. Phasing experiment for long RNAs. Gel mobilities are plotted for the long 3′−5′ bulge RNA shown in Figure 1; the mobilities have been normalized to the fastest running (81 bp) RNA. Data from duplicate gels run at 30°C in phosphate buffer are plotted, and fit to equation (1); the best fit maximum is listed in Table III.

conformation to *trans*.) A comparison of Figures 2B and 4 shows that the phasing has changed by roughly 180°. Calculated values for χ and τ are listed in Table II. From duplicate experiments, we estimate the error in determining these angles as about ± 6°. The variation of χ with conditions is therefore within experimental error, while the difference in τ between 30° and the two low temperature conditions is perhaps significant. While the differences between χ and τ are accurately determined by this experiment, the absolute values of the angles are subject to the assumption that the RNA helical repeat is 11 bp. Systematic error caused by uncertainty in the RNA helical repeat is considered in the Discussion.

Comparison of the two phasing experiments shown in Figures 2B and 4 reveals one curiosity. The amplitude of the mobility variation is about the same in each case, but the maximum relative mobility comes much closer to unity in Figure 4 than in Figure 2B. If the gel mobility were sensitive only to the effects of the bulged nucleotides on the helix axis, and not to the geometry of the bulged nucleotides themselves, the relative mobility should have reached the same maximum value in each case. There may well be some interaction of the bulged nucleotides with the gel matrix; indeed, others have observed different mobilities for left and right-handed supercoiled isomers of DNA (19).

Measurement of RNA helical repeat

The value of the helical repeat in phasing experiments is set by the distance between two minima or maxima in the phasing curve. As discussed above, a single phasing experiment has substantial error in measuring this distance, on the order of ±0.5 base pair. If a second phasing experiment were to find the *cis* or *trans* conformation for a long RNA with many helical repeats, then the same absolute error would be divided by the number of repeats, and a much more accurate value would be derived. The average helical repeat of DNA in solution turned out to be different than that determined by fiber diffraction (20, 21), and the same comparison for RNA would be of interest; the variation of the helical repeat with salt and temperature is also an important control for the above phasing experiments.

Table III. Calculation of the RNA helical repeat from phasing experiments[a]

conditions	maximum	difference	R (n = 5)	R (n = 4)
30°	79.64	50.72	10.14	12.68
30°, + Mg	79.51	51.03	10.21	12.76
8°	79.27	50.93	10.19	12.73
8° + Mg	79.31	51.06	10.21	12.77

[a]The number of base pairs between the two A_2 bulges giving the maximum gel mobility, as derived from fitting equation (1) in Materials and Methods to data of Figure 5, is shown in the second column. The number of base pairs giving a maximum gel mobility in the experiment of Table I has been subtracted (third column) and divided by either 5 or 4 to give the helical repeat values shown in the last two columns. The overall error in R is ±0.1 bp.

To carry out this experiment, we prepared longer RNA molecules that contain two A_2 bulges in the same sequence contexts as the RNAs discussed above, 75−85 base pairs separating the bulges, and correspondingly longer arms on either end (Figure 1). The DNA templates for transcribing the two RNA strands of each duplex were generated by ligation of PCR-generated DNA; the procedure is described in Materials and Methods. The results are graphed in Figure 5. A periodic variation in the gel mobility is observed, as expected. The fitted position of the maximum gel mobility under different conditions is listed in Table III, along with calculation of the helical repeat by comparison with the shorter molecules listed in Table I.

The estimated error of the helical repeat determination is ± 0.1 bp/turn, and the variation in R seen with temperature or added Mg^{2+} is less than this; thus we are justified in attributing any variations in ξ (measured in the phasing experiment of Table I) to changes in the bulge conformation. The calculated values are either 10.2 or 12.7, depending on whether four or five helical repeats are assumed to be present. 10.2 is closest to the range of values expected from x-ray diffraction experiments (see Discussion), but is unexpectedly small. There are two possible factors that may have affected our measurements. First, the sequence between the two bulges, which was picked randomly from pBR322, has a high G−C content (64%) and an excess of Y−R and R−Y dinucleotide sequences (59%) over R−R and Y−Y. We doubt this sequence bias can completely account for the difference from the expected value. Although crystal structures of DNA in the A form have shown that the twist between base pairs can vary widely, a large twist at one position is usually accompanied by a small twist nearby, so that the average twist values show little variation (see, for example, refs. 22−24).

A second factor to consider in this experiment is the possibility of static curvature of the RNA helix axis. The calculation assumes that the helix axis between the two bulges is straight; if it is not, then the minimum gel mobility will no longer occur at $\phi = 0°$. To see this, consider an RNA containing only bend at the two bulges and no bulge or helix-associated twist, i.e. a flat ribbon, and having a *cis* conformation. The molecule lies entirely in one plane. If the middle segment is now curved in a different plane (Figure 6A), rotation of one of the arms will decrease the end-to-end distance. α, the additional twist needed to attain the minimum gel mobility, will introduce a corresponding error in the determination of number of base pairs required for the *cis* conformation ($\phi = 0$). Calculations of α for different degrees of curvature and orientations of the plane of the curvature with respect to the end segments are shown in Figure 6B. The maximum error in finding the number of base pairs giving the

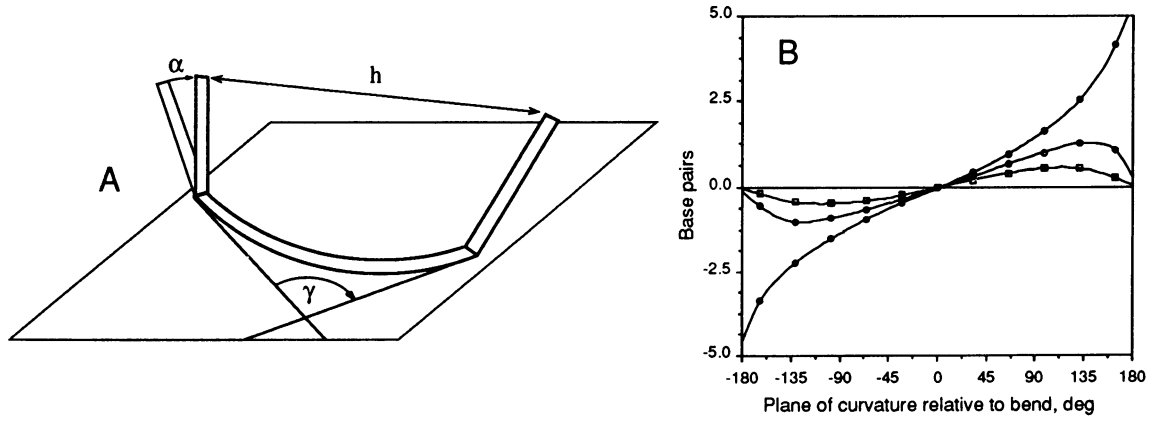

Figure 6. Effect of RNA curvature on the phasing experiment. **A**, effect of curving the helix between two bulges on the minimum end-to-end distance **h**. See text for discussion. γ is the angle of curvature defined by the two ends of the helix segment, and a is the difference between the position of the *cis* conformation and the minimum **h**. **B**, errors in the measurement of twist in a phasing experiment done with curved RNA. The minimum end-to-end distance was found by a simple program that computed the trajectory for RNAs with bulge-associated bend angles θ = 0.8 radians, various angles of the plane of curvature with respect to the plane in which the bend was made at one end (abcissa), and degrees of curvature (γ) 0.2 (□), 0.5 (●), and 0.8 (○) radians. (The program is similar to that described in ref. 17, only using the rotations as defined in Figure 3.) The ordinate is reported as the displacement of the gel mobility maximum from the expected position in base pairs, assuming a helical repeat of 11.

cis or *trans* conformation for an overall curvature of 30° is about ± 1 base pair. (The error varies with the magnitude of the bend angle.) When the curvature is at about 45° (0.8 radians) and oriented so that the RNA is a 'W' shape in the *cis* conformation, the end-to-end distance of the *trans* conformation is actually slightly smaller that that of the *cis* RNA, and the error is half a turn of helix. It is therefore quite feasible that fairly modest amounts of bend (30−40°) could lead to errors of several tenths base pairs in our determination of the helical repeat.

DISCUSSION

Structure of A_2 bulges

A complete description of the bend induced by a bulge has three components: the bend angle, the overall twist angle, and the distribution of the twist between the 5′ and 3′ sides of the bulge. These experiments define the latter value to about ±6°; the twist is about 12° more negative at the 5′ side at 30 °C, and 25° more negative at low temperatures. Mg^{2+}, which might be expected to stabilize specific conformations of a bulge by its interactions with phosphates, does not affect either the magnitude of the helix twist or its distribution between the 5′ and 3′ sides of the bulge.

The calculated magnitude of the bulge-associated twist depends on the assumed value for the RNA helical repeat, and consequently has a larger associated error. The calculations in Table I assume 11 base pairs per turn; a larger value of 11.4 changes the total bulge-associated twist from −30° to ~0°, while 10.6 bp/turn would make the twist more negative, −60°. As discussed below, it is unlikely that the RNA helical twist is outside this range, and an A_2 bulge therefore reduces the normal ~33° twist between base pairs substantially. The experiment does not distinguish between −30° and +330° twists, but only the former value is physically reasonable, since the extra backbone of two nucleotides does not allow more than ~90° positive twist without disrupting base pairs. In model building exercises, we have found that maintaining a continuous stack of bases on the strand with the bulge gives a positive twist, as the backbone tends to continue a right-handed twist through the bulged bases. Only if the bulged bases are pulled out of the helix and into one of the grooves do we start to observe negative twists. This sort of model for single-base bulges, with the base hydrogen bonding in the major groove, has recently been suggested for two different RNAs (25, 26). A single base bulge is accommodated in the groove with little perturbation of the normal helical twist between the flanking base pairs, but a two base bulge requires some disruption of the helix. Despite the qualitative nature of our conclusion concerning the bulge-associated twist, the observation that it is more negative than the normal ~33° twist of a helix is unexpected and should be an impetus to further model building and physical studies.

These electrophoresis experiments do not measure the bend angle (θ) directly, and we can only set rough limits on this parameter. Simple stacking of a single base bulge into an RNA helix cannot bend the axis by more than ~12°, and this is barely detectable as a gel retardation (13). A bend of 90° or more will probably behave peculiarly in a phasing experiment, since the *cis* conformation is a 'U' in which the largest dimension is no longer the end-to-end distance; this may be the situation with 3 and 5 base bulges (13). The A_2 bulge must be between these two extremes.

To date few NMR studies of bulges larger than one nucleotide have appeared, for either DNA or RNA. dA_2 and dA_3 bulges in DNA are stacked with each other and completely disrupt stacking between the flanking base pairs (27); in a second study of a dA_3 bulge the flanking pairs themselves were destabilized (28). The structure of a d(ATA) bulge was solved in detail by NMR, and causes a large 50−60° bend in the DNA, together with an 84° twist (29). This is a positive twist continuing the right handed sense of the backbone, opposite to what we deduce for an A_2 bulge. The only NMR study of a larger RNA bulge looked at the sequence UCU, in which the pyrimidines are also stacked with each other and inserted into the helix (30). The twist was not determined but is presumably also positive. The negative twist of the A_2 bulge suggests that it has a qualitatively different conformation than these other RNA and DNA bulge sequences.

Helical repeat of RNA

Although the helical repeat of RNA is commonly assumed to be 11 base pairs per turn, experiments in the literature give a range of possible values. In fiber diffraction experiments on random sequence double stranded RNA, it was difficult to distinguish 10 and 11 fold helical repeats (31, 32), and the value of 11 is based largely upon diffraction of homopolymers (33). The x-ray crystal structures of $tRNA^{phe}$ and $tRNA^{asp}$ provide the best structures of short RNA helices. The average twist angles calculated for the amino acid and anticodon stems of $tRNA^{phe}$, which are the least distorted by the tRNA tertiary structure, depend on the crystal form and refinement procedure. In one analysis the helical repeats are ~10.7 and 11.1 for the two stems in the orthorhombic form, respectively (34), while a refinement of the monoclinic form gave values of either 11.8 and 11.6, or 11.1 and 10.9, depending on the least squares method used (35). The same stems in $tRNA^{ala}$ have helical repeats between 10.5 and 11.0 (36). These values are clustered about the expected 11.0 bp/turn, but do not help in determining a more accurate repeat for random sequence RNA in solution. The one crystal structure of a Watson–Crick paired RNA helix, which is entirely A–U pairs, has a helical repeat near 11 (37). There have been a number of crystal structures of A form DNA solved which also give helical repeats on average near 11 (22, 23, 24).

Part of the motivation for carrying out the phasing experiment with longer RNAs was to determine a better value for the average RNA helical repeat in solution. The calculated value of 10.2 ±0.1 bp/turn is significantly lower than we expected from the studies cited above. (The alternative of 12.7 bp/turn is much too large, and even less likely to be correct). In considering ways in which the experiment could be biased, it seems to us that the most likely explanation is curvature of the RNA helix axis. As shown in Results, moderate amounts of curvature (30–40°) could have caused deviations in the measured repeat from the true value by several tenths of a base pair. There is little information which might decide whether random RNA sequences are commonly curved. The alternating sequence $U(UA)_6A$, which has been crystallized as a duplex, has distinct kinks at two positions separated by 5 bp and an overall bend of 22° (37). Whether this would be a static bend in solution is not known, but it does illustrate the ability of RNA to bend significantly within a short sequence. The sequence CGC is repeated three times with 11–12 bp periodicity in the longer 3′–5′ bulge RNA, and other dinucleotide sequences appear with approximately 11 bp periodicity, so there is the opportunity for small sequence dependent bends to add up to significant curvature in our experiments. If curvature is common, plots of duplex RNA electrophoretic mobilities or rotational diffusion times might be expected to show anomalous points when several sets of related sequences are compared. An insufficient number of RNAs have been carefully examined by these methods to draw any conclusions (38), but a larger collection of duplex RNAs has been prepared and is under investigation in this laboratory.

ACKNOWLEDGEMENTS

We thank L.Laing and Y.-X.Wang for preparation of T7 RNA polymerase, and P.Kebbekus for T3 RNA polymerase and synthesis of DNA oligomers. This work was supported by NIH grant GM37005.

REFERENCES

1. Marini, J.C., Levene, S.D., Crothers, D.M. and Englund, P.T. (1982) *Proc. Natl. Acad. Sci. USA*, 79, 7664–7668.
2. Hagerman, P.J. (1984) *Proc. Natl. Acad. Sci. USA*, 81, 4632–4636.
3. Zinkel, S.S. and Crothers, D.M. (1987) *Nature*, 328, 178–181.
4. Weeks, K.M. and Crothers, D.M. (1991) *Cell*, 66, 577–588.
5. Roy, S., Delling, U., Chen, C.-H., Rosen, C.A. and Sonnenberg, N. (1990) *Genes & Develop.*, 4, 1365–1373.
6. Bartel, D.P., Zapp, M.L., Green, M.R. and Szostak, J.W. (1991) *Cell*, 67, 529–536.
7. Wu, H.-N. and Uhlenbeck, O.C. (1987) *Biochemistry*, 26, 8221–8227.
8. Mougel, M., Ehresmann, B. and Ehresmann, C. (1986) *Biochemistry*, 25, 2756–2765.
9. Levene, S.D. and Zimm, B.H. (1989) *Science*, 245, 396–399.
10. Thompson, J.F. and Landy, A. (1988) , *Nucleic Acids Res.*, 16, 9687–9705.
11. Sanger, S., Nickelen, S. and Coulsen, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, 74, 5463–5467.
12. Draper, D.E., White, S.A. and Kean, J.M. (1988) *Methods Enzymol.*, 164, 221–237.
13. Tang, R.S. and Draper, D.E. (1990) *Biochemistry*, 29, 5232–5237.
14. Bevington, P.R. (1969) Data reduction and error analysis for the physical sciences. McGraw-Hill, New York.
15. Bhattacharyya, A., Murchie, A.I.H. and Lilley, D.M.J. (1990) *Nature*, 343, 484–487.
16. Riordan, F.A., Bhattacharyya, A., McAteer, S. and Lilley, D.M.J. (1992) *J. Mol. Biol.*, 226, 305–310.
17. Levene, S.D. and Crothers, D.M. (1983) *J. Biomol. Struc. Dynam.*, 1, 429–435.
18. Kerppola, T.K. and Curran, T. (1991) *Science*, 254, 1210–1214.
19. Drak, J. and Crothers, D.M. (1991) *Proc. Natl. Acad. Sci. USA*, 88, 3074–3078.
20. Wang, J.C. (1979) *Proc. Natl. Acad. Sci. USA*, 76, 200–202.
21. Rhodes, D. and Klug, A. (1980) *Nature*, 286, 573–578.
22. Bingman, C.A., Zon, G. and Sundaralingam, M. (1992) *J. Mol. Biol.*, 227, 738–756.
23. Eisenstein, M., Frolow, F., Shakked, Z. and Rabinovich, D. (1990) *Nucleic Acids Res.*, 18, 3185–.
24. Verdaguere, N., Aymamí, J., Fernández-Forner, D., Fita, I., Coll, M., Huynh-Dinh, T., Igolen, J. and Subirana, J.A. (1991) *J. Mol. Biol.*, 221, 623–635.
25. Yarus, M., Illangesekare, M. and Christian, E. (1991) *J. Mol. Biol.*, 222, 995–1012.
26. Wimberly, B., Varani, G. and Tinoco, I., Jr (1993) *Biochemistry*, 32, 1078–1087.
27. Rosen, M.A., Shapiro, L. and Patel, D.J. (1992) *Biochemistry*, 31, 4004–4014.
28. Aboul-ela, F., Murchie, A.I.H., Homans, S.W. and Lilley, D.M.J. (1993) *J. Mol. Biol.*, 229, 173–188.
29. Rosen, M.A., Shapiro, L. and Patel, D.J. (1992) *Biochemistry*, 31, 4015–4026.
30. Puglisi, J.D., Tan, R., Calnan, B.J., Frankel, A.D. and Williamson, J.R. (1992) *Science*, 257, 76–80.
31. Arnott, S., Wilkins, M.H.F., Fuller, W. and Langridge, R. (1967) *J. Mol. Biol.*, 27, 525–533.
32. Bloomfield, V.A., Crothers, D.M. and Tinoco, I., Jr. (1974) Physical Chemistry of Nucleic Acids. Harper & Row, New York.
33. Arnott, S., Hukins, D.W.L., Dover, S.D., Fuller, W. and Hodgson, A.R. (1973) *J. Mol. Biol.*, 81, 107–122.
34. Holbrook, S.R., Sussman, J.L., Warrant, R.W. and Kim, S.-H. (1978) *J. Mol. Biol.*, 123, 631–660.
35. Westhof, E. and Sundaralingam, M. (1986) *Biochemistry*, 25, 4868–4878.
36. Westhof, E., Dumas, P. and Moras, D. (1985) *J. Mol. Biol.*, 184, 119.
37. Dock-Bregeon, A.C., Chevrier, B., Podjarny, A., Johnson, J., de Bear, J.S., Gough, G.R., Gilham, P.T. and Moras, D. (1989) *J. Mol. Biol.*, 209, 459–474.
38. Gast, F.-U. and Hagerman, P.J. (1991) *Biochemistry*, 30, 4268–4276.