Volume 15 Number 23 1987 Nucleic Acids Research

The mechanism of cruciform formation in supercoiled DNA: initial opening of central basepairs in salt-dependent extrusion

Alastair I.H.Murchie and David M.J.Lilley*

Department of Biochemistry, The University, Dundee DD1 4HN, UK

Received October 28, 1987; Accepted November 11, 1987

ABSTRACT.

There are two alternative pathways by which inverted repeat sequences in supercoiled DNA molecules may extrude cruciform structures, called C-type and S-type. S-type cruciforms, which form the great majority, are characterised by absolute requirement for cations to promote extrusion, which then proceeds at higher temperatures and with lower activation parameters than for C-type cruciforms. The mechanism proposed for S-type extrusion involves an initial opening of basepairs limited to the centre of the inverted repeat, formation of intra-strand basepairing and a four-way junction, and finally branch migration to the fully extruded cruciform. The model predicts that central sequence changes will be more kinetically significant than those removed from the centre. We have studied the kinetics of cruciform extrusion by a series of inverted repeats related to that of pIRbke8 by either one or two mutations in the symmetric unit. We find that mutations in the central 8 to 10 nucleotides may profoundly affect extrusion rates - the fastest being 2000-fold faster than the slowest, whereas mutations further from the centre affect rates to a much smaller extent, typically up to ten-fold. These data support the proposed mechanism for extrusion *via* central opening.

INTRODUCTION.

DNA is a structurally polymorphic molecule [1-15], and the processes by which different conformers interconvert are imperfectly understood. We have made a particular study of the cruciform structure [4-6], the twin hairpin conformation adopted by sequences of two-fold symmetry [16,17], and the mechanistic pathways of the extrusion reaction by which they are formed [18-21].

In general, there are significant kinetic barriers to cruciform extrusion [19-26]. Cruciform structures may be classed in two groups, according to the kinetic character of their extrusion [19]. The great majority of sequences fall into the S-type class, requiring salt to be present for extrusion to proceed, and having moderate activation entropies and enthalpies. A much smaller group, the C-type cruciforms, extrude in the absence of added salt, and are extremely temperature-dependent with very large enthalpies and entropies of activation. The critical factor which controls the selection of extrusion mechanism is the nature of the DNA which forms the environment of the inverted repeat. Certain A+T rich sequences confer the C-type character on almost any inverted repeat placed nearby in *cis* [20], and in the absence of such sequences the kinetics default to S-type. It is self-evident that any process by which a cruciform is extruded must involve an initial unpairing

© IRL Press Limited, Oxford, England.

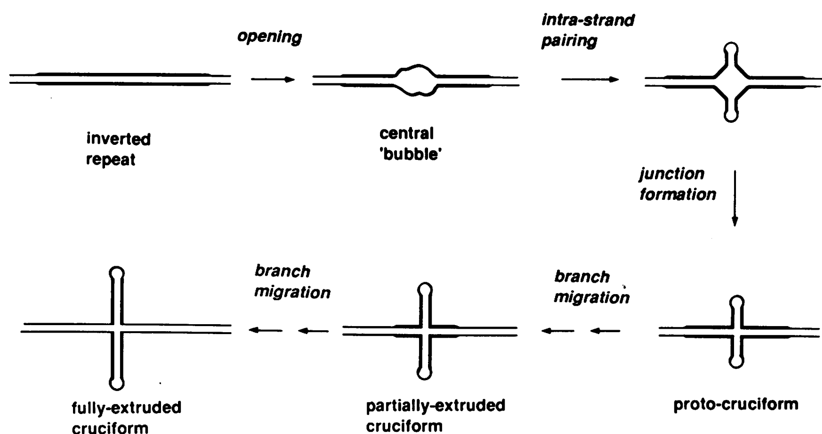

Figure 1. The proposed mechanism for S-type cruciform extrusion. The extent of the inverted repeat is indicated by the bold lines. The first stage of the process is the opening of basepairs around the sequence two-fold axis, followed by the formation of intra-strand basepairing to generate the incompletely extruded proto-cruciform structure. This may then branch migrate by a series of sequential transfers of basepairing to form the fully extruded cruciform structure.

of a region in the DNA. Analysis of the kinetic properties of representative C- and S-type cruciforms suggested that the size of this opening is the critical difference between the two molecular mechanisms of isomerisation - that of the C-type cruciforms being substantially larger, due to a cooperative destabilisation arising from the nature of the flanking DNA.

The proposed mechanism of the S-type extrusion process is [18,19,21] represented in Figure 1. Its fundamental features are as follows:

1. A relatively restricted initial opening of the helix in the centre of the inverted repeat. The opening event is probably further divisible into initiation and expansion events.
2. Formation of a structure in which intra-stand basepairs are formed. Again, this is likely to be divisible into a number of sub-steps required before the formation of the proto-cruciform is complete.
3. Sequential transfer of basepairs from the unperturbed helix into the growing cruciform stems by a branch migration process. Once the the entire inverted repeat is intra-strand paired, the extrusion will be complete.

Rates of branch migration have been measured, and are relatively fast [27]. Therefore steps 1 and 2 are most probably rate-limiting. The model therefore clearly predicts that the kinetic properties will be sensitive to the sequence of the centre of the inverted repeat, but relatively less sensitive to sequence changes made outside this region. For example, if we change an A.T basepair for G.C in the centre of a given S-type inverted repeat, this is likely to raise the free energy of the transition state, since this is the region which must be opened. However, if the same mutation is made

elsewhere in the inverted repeat, this basepair may remain fully paired in the transition state, and hence the kinetic consequences are likely to be less significant.

In this study we have tested this prediction. We have taken the bke inverted repeat of the S-type plasmid pIRbke8, the extrusion kinetics of which have been extensively studied in this laboratory [19], and constructed mutated sequences which differ from the parent bke sequence in one or two locations within the symmetrical unit. The extrusion kinetics of each plasmid have been studied, and in this way the kinetic properties have been related to the sequence of the inverted repeat in a systematic manner. We find that central mutations do indeed affect extrusion rates very much more than those located outside this region, and we regard this as further confirmation of the model originally proposed for the S-type process.

EXPERIMENTAL.

Construction of plasmids. All the plasmids employed in these studies were made in a manner analogous to the original construction of pIRbke8 [18]. In each case two 13 nucleotide oligonucleotides of the type

5'GATCCGGTACCAG3' and 5'AATTCTGGTACCG3'

were synthesised. The second was phosphorylated and an equimolar mixture of the two was hybridised and ligated using T4 DNA ligase. This produced unambiguous inverted repeat dimers, which were then purified using preparative polyacrylamide gel electrophoresis. After electroelution and purification on DEAE cellulose columns, the dimer fragments were phosphorylated, ligated to *Bam*HI cleaved pAT153 [28] which had been dephosphorylated using calf intestinal alkaline phosphatase, and transformed into *E.coli* HB101 [29] cells with selection for Ap^r. In general colonies could be screened by restriction enzyme cleavage of alkali mini-preparations of plasmid DNA [30]. The sequences in and around the inverted repeats of all plasmids were subsequently confirmed using chemical cleavage sequence analysis [31]. Enzymes used in the manipulation of DNA sequences were obtained from Amersham or Bethesda Research Labs.

Synthesis of oligonucleotides. All oligonucleotides employed in the construction of the plasmids for these studies were synthesised by phosphoramidite chemistry [32] implemented on an Applied Biosystems 381A DNA synthesiser. It was found unnecessary to purify these prior to the ligation step to make the inverted repeat dimers.

Kinetic methods. Cruciform-free plasmids were prepared as described before [19]. In brief, plasmid DNA prepared by lysozyme, EDTA, SDS lysis of *E.coli* cells was subjected to two rounds of CsCl-ethidium bromide isopycnic centrifugation. Supercoiled plasmid was recovered by side puncture and the ethidium bromide removed by multiple extraction with butan-1-ol at 0°C, followed by dialysis against 10 mM Tris.HCl (pH 7.5), 0.1 mM EDTA at 4°C. This DNA was used directly in kinetic experiments without further manipulation.

Cruciform extrusion experiments were performed using 29 µg/ml cruciform-free plasmid DNA in 10 mM Tris.HCl (pH 7.5), 0.1 mM EDTA with a given concentration of added NaCl. In the case

of experiments designed to measure optimal salt concentrations the plasmids were incubated with various NaCl concentrations for fixed times at 37°C before placing on ice for subsequent cruciform assays. For the determination of rate constants for the extrusion process, DNA was incubated in the above buffer with the indicated concentration of NaCl at 37°C unless otherwise indicated, and aliquots were removed to ice at various times. These were subsequently assayed for the extent of cruciform extrusion. In these experiments temperature was controlled to an accuracy of ±0.05 degrees using either Haake D8 or Grant LTD6 water baths.

The extent of cruciform extrusion was assessed using cleavage by either S1 nuclease, which cleaves the single-stranded loops or T7 endonuclease I [33], which cleaves the four-way junction. Following the cruciform-specific cleavage the DNA was purified by ethanol precipitation and digested to completion with *Hin*dIII, which cuts the molecules at a single site 350 bp from the inverted repeat. The DNA was then electrophoresed in 1% agarose gels in 10 mM Tris.borate (pH 8.3), 10 mM EDTA at room temperature for 15 h. Gels were stained in 1 µg/ml ethidium bromide and photographed under UV illumination using Kodak Tri-X Pan film.

Quantification of relative extent of cruciform extrusion was achieved by laser densitometry of gel photographic negatives using an LKB Ultroscan 2202 laser densitometer interfaced to a computer for curve fitting and integration of peak areas. Relative extent of extrusion was calculated as the intensity of the band due to site-specific cleavage at the cruciform, divided by the total intensity. Rate constants were calculated by linear regression analysis from the gradient of ln(1-extruded fraction) plotted against time. Errors are standard errors from the regression analysis.

Arrhenius activation energies were calculated from the gradients of plots of lnk against 1000/T, according to

$$k = Ae^{-Ea/RT}$$

where k is the rate constant, Ea the Arrhenius activation energy, R the gas constant and T the absolute temperature.

S1 nuclease and T7 endonuclease I digestions. S1 nuclease (Bethesda Research Laboratories) digestions were performed in 50 mM Na acetate (pH 4.6), 50 mM NaCl, 1mM $ZnCl_2$. Fast cruciforms were digested using 7 U S1 nuclease at 0°C for 45 m, while others were digested with 2 U at 15°C for 30 m. T7 endonuclease (a kind gift from P. Sadowski) digestions were performed in 50 mM Tris.HCl (pH 8.0), 10 mM $MgCl_2$, 1 mM dithiothreitol, 50 µg/ml BSA and 4 mM spermidine. The enzyme was diluted 1 in 10 as supplied. pIRbke16T was digested using 2 µl of diluted enzyme in a final 20 µl volume at 0°C for 120 m, while pIRbke8, pIRbke16C and pIRbke15C16C were digested using 1 µl diluted enzyme at 10°C for 60 m.

RESULTS.

Construction of a series of inverted repeats closely related to pIRbke8.

One of the best characterised inverted repeat sequences with regard to S-type cruciform extrusion is that of the plasmid pIRbke8 [19]. The plasmid contains the bke inverted repeat of total length 32

Figure 2. The sequences employed in these studies. At the top is drawn the parent bke sequence, shown in its cruciform conformation on the right. The sequences of all the mutant inverted repeats are shown below that of bke, with alterations from bke highlighted as open letters.

bp, the centre of which is an *Eco*RI site. This sequence forms a stable cruciform at native superhelix density, and exhibits a maximal rate of extrusion at 50 mM NaCl, with an Arrhenius activation energy of 42 kcal mol^{-1}. At 37°C in 50 mM NaCl extrusion proceeds with a halftime of 70 m.

In order to investigate the effects of sequence changes on the kinetics of cruciform extrusion we constructed the series of related inverted repeats indicated in Figure 2. These were made by cloning oligonucleotides of appropriate sequence into the *Bam*HI site of pAT153. All were subsequently confirmed by DNA sequencing. Each mutant sequence is related to the bke sequence by either one or two base changes in the symmetrical unit of the inverted repeat. They are named according to the nature of the changed nucleotide and its position measured from the 5' end of the inverted repeat - for example pIRbke14A has an adenine replacing the guanine at base 14.

Rates of cruciform formation by pIRbke8 and mutant sequences.

Cruciform-free DNA was prepared for each plasmid, by a standard procedure [19] involving positive supercoiling in the presence of ethidium bromide, followed by extraction of the intercalator into cold butan-1-ol. This DNA was exposed to a fixed temperature (37°C in most experiments) in the presence of 10 mM Tris.HCl pH 7.5 and sodium chloride between 35 and 75 mM. Samples were removed at various times, and immediately cooled on ice. The extent of cruciform formation was measured using S1 nuclease, which cleaves the single-stranded loop of the cruciform structure [34], followed by complete restriction cleavage at the unique *Hin*dIII site in each plasmid.

mutant	central sequence	S1 nuclease data k (error),	exp	T1/2 (m)	endonuclease I data k (error),	exp	T1/2 (m)	NaCl optimum (mM)
bke	GAATTC	-1.7 (.02)	-4	70	-1.6 (.18)	-4	72	50
16T	GATATC	-1.3 (.14)	-3	9	-0.8 (.02)	-3	13	50
14T16T	ATATATAT	v fast		<1				35
14A	AAAATTTT	-3.4 (.2)	-4	34				50
15C	GCATGC	-1.7 (.04)	-4	67				75
16C	GACGTC	v slow		>300	-1.7 (.18)	-5	692	50
15C16C	GCCGGC	v slow		>300	-7.0 (.6)	-6	1649	50
13G	GGAATTCC	-1.0 (.09)	-4	114				75
11T	GAATTC	-1.5 (.07)	-3	8				50
10G	GAATTC	-1.4 (.12)	-4	81				50
8A	GAATTC	-2.0 (.13)	-3	6				50
8T	GAATTC	-6.8 (.5)	-4	17				50
6T	GAATTC	-2.9 (.2)	-4	40				50

Table 1. Kinetic data for pIRbke8 and mutant sequences. Salt optima and rate constants measured for cruciform extrusion at 37°C. Rate constants are expressed as k (±standard error) x 10^{exp}. Half times have been calculated from $T_{1/2} = 0.69/k$. The rate of extrusion at 37°C for pIRbke14T16T was too fast to measure with accuracy using our procedures, although extrapolation of the Arrhenius plot (Figure 7) gives a rate constant at 37°C of -1.29 x 10^{-2} ($T_{1/2}$ = 0.9 m). pIRbke16C and pIRbke15C16C were poorly cleaved by S1 nuclease, and the resulting low intensities made quantification difficult. For this reason T7 endonuclease I was employed to measure these rate constants.

In initial experiments we determined the salt concentration for maximal cruciform extrusion rate for each sequence. These optima were taken as the centre of fairly broad peaks (a change in salt concentration of 10 mM typically changed the rate by less than 5%) [21], and are presented in Table 1. In general, central mutations which replaced AT basepairs by GC resulted in a requirement for higher salt concentrations. All subsequent kinetic experiments were performed at the salt optimum for the sequence being studied.

We performed time-courses of cruciform extrusion for each sequence at 37°C. Typical results of the experiments are shown in Figure 3, for the extrusion of pIRbke16T. The increase in intensity of the cruciform-specific band with time is clear. By densitometry of such gel negative photographs it was possible to quantify the relative extent of cruciform formation. Figure 4 shows some examples of plots of ln [unextruded DNA] as a function of time for the sequences. It is clear that the different sequences result in markedly different rates of cruciform extrusion. From the gradient of these data the rate constants have been determined, and these are collected together in Table 1. These are discussed below.

Rates of cruciform extrusion measured using a Holliday resolvase.

One concern to us in these experiments was the possibility of complications due to the use of S1 nuclease, for several reasons. First it has been suggested [23,35] that S1 nuclease may participate directly in the formation of cruciforms, due to its activity as a single-stranded DNA binding protein. While we have disproved this in thermodynamic terms [36] it was desirable to ensure that this was

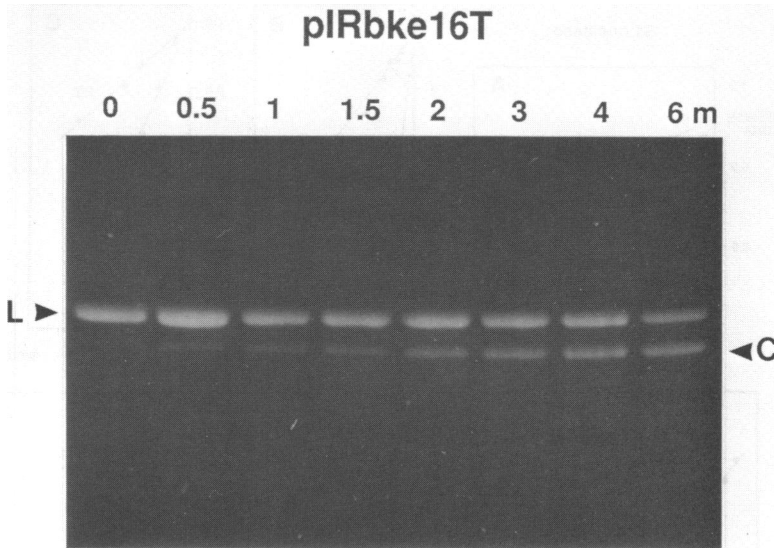

Figure 3. The extrusion of pIRbke16T as a function of time of incubation at 37°C - a representative example of the data obtained in these studies. Supercoiled cruciform-free pIRbke16T was incubated in 10 mM Tris.HCl (pH 7.5), 0.1 mM EDTA, 35 mM NaCl, and aliquots removed to ice at the times indicated. These were subsequently digested with S1 nuclease at 0°C, followed by complete cleavage with *Hin*dIII, and electrophoresed on a 1% agarose gel, the photograph of which is shown. Full length linear plasmid is indicated by the arrow on the left (L), while S1 nuclease cleavage of an extruded cruciform generates the shorter fragment which migrates as the band indicated by the arrow on the right (C). The tracks are labelled with the time (m) of incubation of the plasmid at 37°C. Laser densitometry of negative photographs permits quantification of the data.

not a factor in kinetic terms. Second, many of our mutant pIRbke8 sequences are altered in the sequences which form the loops of the extruded cruciforms, ie the region to be cleaved by S1 nuclease. In the case of the more G+C rich loops it becomes relatively difficult to cleave these.

For these reasons we employed the Holliday resolving enzyme T7 endonuclease I [33] in some of our kinetic experiments. The resolvases cleave at the four-way junctions of cruciforms [33,37-40], and thus their mode of interaction with the cruciform structure is totally different from that of the single-strand specific nucleases such as S1 nuclease. We performed time-courses of cruciform extrusion analogous to those above, except in the replacement of the S1 nuclease cleavage by digestion with endonuclease I, on pIRbke8 and three mutant sequences chosen to cover a large range of extrusion rates. These data are presented in Figure 5, and the rate constants measured are incorporated into Table 1.

A comparison of the data obtained using S1 nuclease and T7 endonuclease I shows them to be very similar, and those for the parent sequence bke, which extrudes on a timescale where it is easiest to

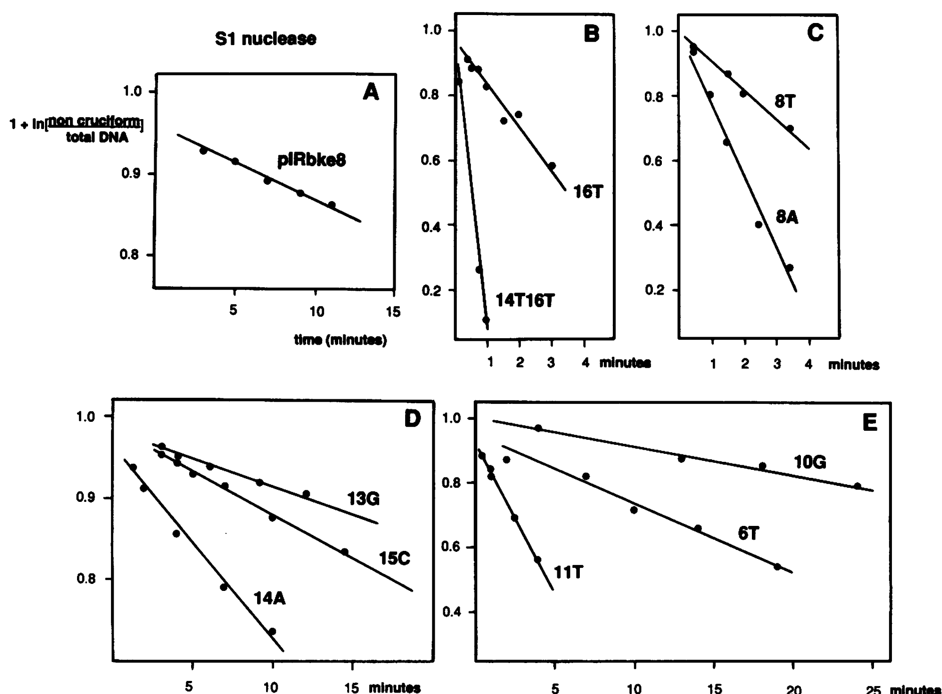

Figure 4. The kinetics of cruciform extrusion by pIRbke8 and mutant sequences. Plots of the disappearence of unextruded DNA with time for (A) pIRbke8, (B) pIRbke16T and pIRbke14T16T, (C) pIRbke8T and pIRbke8A, (D) pIRbke13G, pIRbke15C and pIRbke14A, and (E) pIRbke10G, pIRbke6T and pIRbke11T.

obtain good data, are almost identical. This agreement removes any anxiety about rate constants measured using S1 nuclease, and we may analyse the data of Table 1 with confidence.

Central sequence changes have the greatest consequences.

The rate constants measured for cruciform extrusion by the various sequences cover a wide range, with half times from less than a minute up to over 27 hours - a factor of almost 2000. With one exception, mutations of AT basepairs to GC result in slower rates of extrusion, and all mutations of GC to AT result in enhanced rates. These are represented graphically in Figure 6, and it is very clear from this that the most kinetically significant mutants are those in and around the centre of the inverted repeat. Thus central mutations may be at least ten times more effective than changes elsewhere in the inverted repeat. This is well in accord with predictions based on the proposed S-type mechanism.

Pairwise comparisons of the mutant sequences reveals the importance of base sequence - as opposed simply to base composition - in the rate of cruciform extrusion. If we compare the

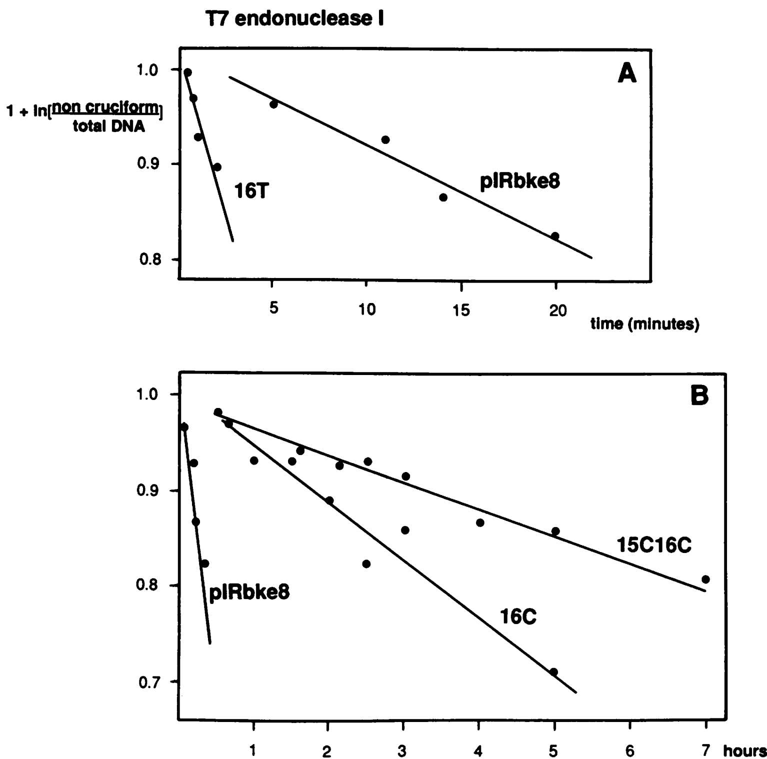

Figure 5. The kinetics of cruciform extrusion measured using the resolvase T7 endonuclease I. Plots of the disappearence of unextruded DNA with time for pIRbke8, pIRbke16T, pIRbke16C and pIRbke15C16C. Note that the data for pIRbke8 are plotted on both graphs, covering the timescales 25 m (A) and 7 h (B) respectively.

sequence 16T with the parent sequence pIRbke8,

bke	GGATCCGGTACCAGAATTCTGGTACCGGATCC	T1/2 70m
16T	GGATCCGGTACCAGATATCTGGTACCGGATCC	T1/2 9m

we see that simply reversing the order of the central dinucleotide, to generate the alternating ATAT sequence results in an 8-fold acceleration of rate. Even more dramatically, if we compare 14T16T with 14A,

14T16T	GGATCCGGTACCATATATATGGTACCGGATCC	T1/2 0.9m
14A	GGATCCGGTACCAAAATTTTGGTACCGGATCC	T1/2 34m

we find that there is a 40-fold difference in rate between these isomeric sequences, with the alternating AT sequence being the faster. Similar effects may be seen on replacement of AT basepairs by GC. Comparing 15C and 16C,

15C	GGATCCGGTACCAGCATGCTGGTACCGGATCC	T1/2 67m
16C	GGATCCGGTACCAGACGTCTGGTACCGGATCC	T1/2 692m

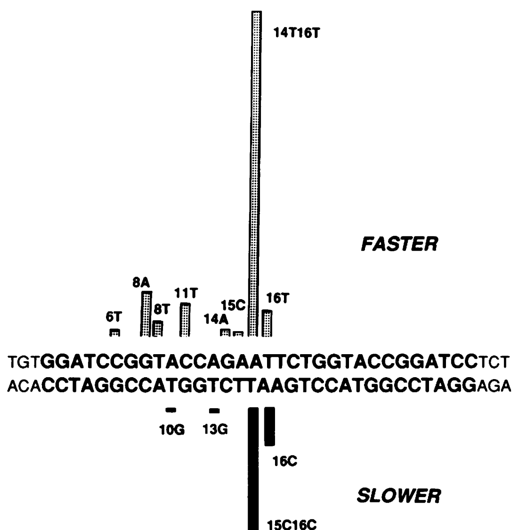

Figure 6. Schematic illustration of the variation in cruciform extrusion rates for the sequence mutants of pIRbke8. The sizes of the bars for each mutant are proportional to the ratio of the extrusion rate relative to the parent bke sequence. Bars above the sequence (speckled) represent rate enhancements relative to pIRbke8, while those below (black) represent slower relative rates.

we see that there is a 10-fold difference in rate between the two. Indeed, 15C is actually slightly faster than pIRbke8, possibly due to the alternation of purine and pyrimidine bases in the centre.

pIRbke14T16T extrusion at lower temperatures.

We chose to study the extrusion of pIRbke14T16T at lower temperatures for two reasons. First, this was the fastest cruciform we had made, and it was therefore of interest to study the kinetic

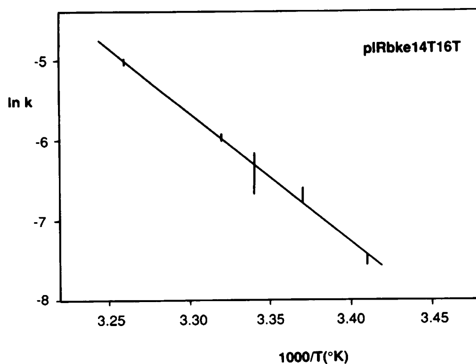

Figure 7. Arrhenius plot of the temperature dependence of cruciform extrusion by pIRbke14T16T. Plot of ln k (k is the rate constant) as a function of reciprocal temperature.

parameters in greater detail. Second, the extrusion rate at 37°C was really too fast to measure with accuracy by our techniques.

We performed time-courses of extrusion at temperatures between 15° and 33°C and calculated rate constants. This sequence exhibited rapid rates of extrusion at relatively low temperatures, and indeed we note that the rate for this plasmid at 15°C was actually faster than that of pIRbke8 at 37°C. For this reason the S1 nuclease and T7 endonuclease I incubations of the faster cruciforms are routinely performed on ice. The temperature dependence of the extrusion rate is presented as an Arrhenius plot in Figure 7, from which we have measured an Arrhenius activation energy (Ea) of

$$Ea = 32.8 \pm 0.2 \text{ kcal mol}^{-1}$$

which is about 10 kcal mol^{-1} less than pIRbke8 [19]. It appears that both enthalpy and entropy changes contribute to the observed rate enhancement for pIRbke14T16T.

Kinetic effects of non-central mutations.

While we have noted that mutations at non-central locations are less significant than those close to the sequence dyad, some have effects which are far from negligible. Thus, while 10G,

10G	GGATCCGGTGCCAGAATTCTGGCACCGGATCC	T1/2	80m

is only 1.1 times slower than pIRbke8, certain GC to AT changes may have much greater kinetic consequences. This may be seen clearly for 11T and 8A,

11T	GGATCCGGTATCAGAATTCTGATACCGGATCC	T1/2	7m
8A	GGATCCGATACCAGAATTCTGGTATCGGATCC	T1/2	6m

where 10-fold increases in rate were found. It may be seen that the effect of these mutations is to produce short alternating ATA or TAT runs, which could be responsible for the kinetic effects. This is supported by a comparison of 8A with 8T,

8A	GGATCCGATACCAGAATTCTGGTATCGGATCC	T1/2	6m
8T	GGATCCGTTACCAGAATTCTGGTAACGGATCC	T1/2	17m

where the alternation results in a three-fold acceleration in the rate of extrusion. We note that this ATA run is separated from the sequence dyad by six nucleotides. Mutation 6T,

6T	GGATCTGGTACCAGAATTCTGGTACCAGATCC	T1/2	40m

which produces no alternation, causes a relatively small rate enhancement of 1.8-fold relative to pIRbke8.

DISCUSSION.

Systematic study of the kinetics of extrusion by a salt-dependent cruciform as a function of sequence changes in the inverted repeat has revealed two main classes of mutation;

1. Mutations of the inverted repeat which affect the sequence of the central region may have an enormous influence on the rate of cruciform extrusion.
2. Mutations outside this central region have reduced effects, but may nevertheless be significant.

In general these results are very much in accord with predictions on the basis of the proposed

mechanism for the S-type extrusion process [19]. In this model the initial events are confined to a central region, and thus sequence changes within this region are expected to have the largest kinetic consequences. This is seen very clearly by the variation in rate constant for extrusion by a factor of 2000 for changes in the centre, compared with about ten outside this region. In addition, only central sequence changes bring about an alteration in the optimal salt concentration.

The importance of purine-pyrimidine alternation, particularly adenine-thymine alternation, is clearly revealed by these studies. Changes which lead to the formation of A-T alternation lead to considerable acceleration of cruciform extrusion, and when this is central the effect can be very large indeed. This must result from weaker stacking interactions in the alternating sequences, and we note that TpA doublets have the smallest dinucleotide stability constants [41]. These effects may be related to earlier demonstrations that $(AT)_n$ cruciforms appear to extrude cruciforms without a detectable kinetic barrier even at low temperatures [42,43], and we have speculated that these sequences may be special classes of C-type cruciforms [44], where the entire alternating segment forms a cooperatively melting unit which is subject to relatively easy denaturation. Thus the shorter $(AT)_n$ stretches may act as nuclei for the opening process to start extrusion by the S-type pathway. The ATA and TAT sequences outside the central region are interesting. These may increase the extrusion rate by an order of magnitude, despite being 6 to 8 nucleotides away from the centre of the inverted repeat, perhaps as a consequence of telestability effects [45]. Perhaps the alternating A-T sequences, which we know to be torsionally deformable [13], may be better able to accomodate an interface with the opened bubble, and the overall energy of the transition state may therefore be reduced. It may be significant that a common feature of the Pribnow boxes found -10 bp upstream of most prokaryotic promoters [46] is the sequence ATA or TAT. The initial opening events of cruciform extrusion may well turn out to have much in common with open complex formation by RNA polymerase in the initiation of transcription.

The kinetic data may be used to estimate the size of the initial opening in S-type cruciform extrusion. The large effects caused by mutations in the central hexanucleotide sequence, coupled with the small influence of the 13G mutation suggests that the initial process is the unpairing of approximately six basepairs. This opening must then presumably expand, normally a much easier process in thermodynamic terms [47], before it becomes possible to form the first intra-strand basepairs. However any estimate of initial bubble size can only be an approximation, as it is almost impossible to disentangle primary and nearest-neighbour effects, such as those of the remote ATA sequences. It is possible that the size of the first unit to be opened is not fixed, and depends on sequence, and thus eccentric ATA sequences may become incorporated into the primary bubble, for example.

We have demonstrated before that the efficiency of ions in promoting S-type extrusion is dependent both on the charge and the ionic radius, and we have interpreted the latter in terms of binding to a tertiary folded structure, ie a junction [21]. The optimum in extrusion rate obtained at 50 mM NaCl probably represents a balance between facilitating initial opening, which will be promoted by low

ionic strength, and stabilising the four-way junction. In the case of divalent ions the profile reaches a plateau at about 200 μM, reflecting the higher binding constants for these ions. By studying the extrusion at the optimal salt concentration we lower the free energy of the species already containing junctions and the kinetics become sensitive to the earlier opening events.

In conclusion, the results presented here are a further confirmation of the proposed mechanism for the extrusion of salt-dependent cruciforms. The analysis of sequence changes in the inverted repeat indicates that the initial opening events are restricted to the central region of the inverted repeat. This is consistent with the relatively small activation energy for this process, by comparison with the C-type mechanism for which we believe a much larger opening is involved. Thus cruciform extrusion may be employed as a probe of opening events of different amplitude and frequency in supercoiled DNA molecules, as a function of many parameters of which one of the most important is base sequence. The study of these processes continues to yield valuable insights into the dynamic nature of torsionally stressed DNA.

ACKNOWLEDGEMENTS

We thank Dr F Schaeffer for valuable discussion, Dr P Sadowski for a kind gift of T7 endonuclease I, and the MRC, the Royal Society and the Wellcome Trust for financial support.

*To whom correspondence should be addressed

REFERENCES

1. Dickerson, RE & Drew, HR (1981) *J. Molec. Biol.* **149**, 761-786
2 Wang, AH-J, Quigley, GJ, Kolpack, FJ, Crawford, JL, van Boom, JH, van der Marel, G & Rich, A (1979) *Nature* **282**, 680-686
3 Shakked, Z, Rabinovich, D, Kennard, O, Cruse, WBT, Salisbury, SA & Viswamitra, MA (1983) *J. Molec. Biol.* **166**, 183-201
4 Gellert, M, Mizuuchi, K, O'Dea, MH, Ohmori,H & Tomizawa,J (1979) *Cold Spring Harbor Symp. Quant, Biol.* **43**, 35-40
5 Lilley,DMJ (1980) *Proc.Natl.Acad.Sci. USA* **77**, 6468 - 6472
6 Panayotatos,N & Wells,RD (1981) *Nature* **289**, 466 - 470
7 Drew, HR & Travers, AA (1984) *Cell* **37**, 491-502
8 Cantor, CR & Efstratiadis, A (1984) *Nucleic Acids Res.* **12**, 8059-8072
9 Lyamichev, VI, Mirkin, SM & Frank-Kamenetskii, MD (1985) *J. Biomolec. Structure & Dynamics* **3**, 327-338
10 Pulleyblank, DE, Haniford, DB & Morgan, AR (1985) *Cell* **42**, 271-280
11 Arnott, S. & Selsing, E (1974) *J.Molec.Biol.* **88**, 509-521
12 Marini, JC, Levene, SD, Crothers, DM & Englund, PT (1982) *Proc. Natl. Acad. Sci. USA* **79**, 7664-7668
13 McClellan,JA, Palecek,E & Lilley, DMJ (1986) *Nucleic Acids Res.* **14**, 9291-9309
14 Suggs,JW & Wagner,RW (1986) *Nucleic Acids Res.* **14**, 3703-3716.
15 McClellan,JA & Lilley,DMJ (1987) *J.Molec.Biol.* **197**, 707-722
16 Platt,JR (1955) *Proc.Natl.Acad.Sci. USA* **41**, 181 - 183
17 Gierer,A (1966) *Nature* **212**, 1480-1481
18 Lilley,DMJ & Markham,AF (1983) *EMBO J.* **2**, 527 - 533
19 Lilley,DMJ (1985) *Nucleic Acids Res.* **13**, 1443 - 1465
20 Sullivan,KM & Lilley,DMJ (1986) *Cell* **47**, 817-827

21 Sullivan,KM & Lilley,DMJ (1987) *J.Molec.Biol* **193**, 397-404
22 Mizuuchi,K, Mizuuchi,M & Gellert,M (1982) *J.Molec.Biol.* **156**, 229 - 243
23 Gellert, M, O'Dea,MH & Mizuuchi, K (1983) *Proc. Natl. Acad Sci, USA* **80**, 5545 - 5549
24 Courey, AJ & Wang, JC (1983) *Cell* **33**, 817-829
25 Panyutin,I, Klishko,V & Lyamichev,V (1984) *Biomolec. Struc. & Dynamics* **1**, 1311 - 1324
26 Sinden,RR & Pettijohn,DE (1984) *J. Biol. Chem.* **259**, 6593 - 6600
27 Thompson,BT, Camien,MN & Warner,RC (1976) *Proc.Natl.Acad.Sci. USA* **73**, 2299-2303.
28 Twigg,AJ & Sheratt,D (1980) *Nature* **283**, 216-218
29 Boyer,HW & Roullard-Dussoix,D (1969) *J. Molec. Biol.* **41**, 459-472
30 Birnboim,HC & Doly,J (1979) *Nucleic Acids Res.* **7**, 1513-1523
31 Maxam, A & Gilbert, W (1980) *Methods Enzymol.* **65**, 499 - 560
32 Beaucage,SL & Caruthers,MH (1981) *Tetrahedron Lett.* **22**, 1859-1862
33 deMassey,B, Studier,FW, Dorgai,L, Appelbaum,F & Weisberg, RA (1984) *Cold Spring Harbor Symp. Quant. Biol.* **49**, 715-726
34 Lilley,DMJ (1981) *Nucleic Acids Res.* **9**, 1271-1289
35 Wang,JC, Peck,LJ & Becherer, K (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 85-91
36 Lilley,DMJ & Hallam,LR (1984) *J.Molec.Biol.* **180**, 179 - 200
37 Mizuuchi,K, Kemper,B, Hays,J & Weisberg,RA (1982) *Cell* **29**, 357 - 365
38 Lilley,DMJ, & Kemper,B (1984) *Cell* **36**, 413 - 422
39 West,SC & Korner,A (1985) *Proc. Natl.Acad.Sci. USA* **82**, 6445 - 6449
40 Symington,L & Kolodner,R (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7247 - 7251
41 Gotoh,O & Tagashira,Y (1981) *Biopolymers*, **20**, 1033-1042
42 Greaves,DR, Patient,RK & Lilley,DMJ (1985) *J. Molec. Biol.* **185**, 461-478.
43 Haniford,DB & Pulleyblank,DE (1985) *Nucleic Acids Res.* **13**, 4343-4363
44 Lilley,DMJ, Sullivan,KM, Murchie,AIH & Furlong,JC (1987) in *Unusual DNA Structures*, (Eds, Wells RD & Harvey, SC) Springer-Verlag, In the press
45 Burd,JF, Wartell,RM, Dodgson,JB & Wells,RD (1975) *J.Biol.Chem.* **250**, 5109-5113
46 Rosenberg,M & Court,D (1979) *Ann. Rev. Genetics*, **13**, 319-353
47 Gotoh, O (1983) *Adv. Biophys.* **16**, 1-52.