DNA Secondary Structure

in Vivo

ANGUS DAVISON

Thesis presented for the Degree of Doctor of Philosophy

University of Edinburgh

1994

Declaration

I hereby declare that this thesis was composed by myself, and that the work described is my own, unless otherwise stated.

/

ANGUS DAVISON
March 1994

Abstract

Long DNA palindromes cannot be propagated in wild-type *Escherichia coli*. Either a replicon with a long DNA palindrome is so poorly replicated that it is inviable, or the palindrome itself is unstable so that it undergoes partial or complete deletion. The deleterious effect of long DNA palindromes is alleviated in strains mutant for *sbcC* and *D*: each mutation by itself is necessary and sufficient to allow the plating of a λ phage containing a long palindrome. In this thesis, the mechanism by which long DNA palindromes cause inviability is studied, and it is discussed whether inviability is a consequence of an unusual DNA structure.

An analysis of the effects of palindromic central asymmetry on the propagation of a λ phage across a broad range of *E. coli* K-12 host strains is described. Palindromes carrying an 8bp asymmetry confer a less severe phenotype than do perfect palindromes, arguing that a centre-dependent pathway for palindrome-mediated inviability exists that is independent of the host strain. For further study, a set of long DNA palindromes with paired changes in the central sequence was constructed in bacteriophage λ. Identical palindrome centres were previously used by others to test the S-type model for cruciform extrusion *in vitro*. The plaque areas produced by the palindrome-containing λ phage were compared on an *E. coli sbcC* lawn. Central sequence changes had a greater effect upon the plaque area than peripheral changes, implying that the residual palindrome-mediated inviability in *E. coli sbcC* is centre-dependent and could be due to the formation of a cruciform structure. The results argue strongly that intrastrand pairing within palindromes is critical in determining their effects *in vivo*.

The effect on viability of the presumed DNA loop structure *in vivo* is also studied. A correlation was identified between sequences that may have two base hairpin loops *in vitro*, and the plaque size of a λ counterpart with the same central palindromic sequence. A tentative consensus for the central sequence that may form a two base loop *in vivo* is identified as Py_ _Pu (Py=C,T; Pu=A,G), where the full sequence that may be required is PuPy_ _PuPy.

Finally, a metastable intermediate hypothesis to account for the mechanism of palindrome-mediated inviability is proposed.

Acknowledgements

Abbreviations

AMPS	ammonium persulphate
ATP	adenosine-5'-triphosphate
att	λ attachment site
BBL	Baltimore Biological Laboratories
BIME	bacterial interspersed mosaic element
bp	base pair(s)
BSA	bovine serum albumin
CA	casitone-agar
CAS	casitone-agar, variable salt
cos	cohesive λ termini
χ	*chi* site
(d)dATP	2'(3'-di)-deoxyadenosine-5'-triphosphate
(d)dCTP	2'(3'-di)-deoxycytidine-5'-triphosphate
(d)dGTP	2'(3'-di)-deoxyguanosine-5'-triphosphate
(d)dTTP	2'(3'-di)-deoxythymidine-5'-triphosphate
(d)dNTP	2'(3'-di)-deoxynucleoside-5'-triphosphate
DNase	deoxyribonuclease
DTT	dithiothreitol
eop	efficiency of plating
exoIII	exonuclease III
ΔG	Gibbs free energy change
EDTA	diaminoethanetetra-acetic acid
ERIC	enterobacterial repetitive intergenic consensus
γ	*gam* protein of λ
I	inosine
IHF	integration host factor
IRU	intergenic repeat unit
IS	insertion sequence
k	reaction constant

kb	kilobase pairs
krpm	kilorevolutions per minute
l	litre(s)
λ	bacteriophage lambda
MOI	multiplicity of infection
NMR	nuclear magnetic resonance
OD	optical density
ori	origin of replication
pal ( )	DNA palindrome (length, including any asymmetry)
PEG	polyethylene glycol
PFU	plaque forming units
Pu	purine (G, A)
Py	pyrimidine (C, T)
PU	palindromic unit
REP	repetitive extragenic palindrome
RIB	reiterative IHF BIME
RIP	repetitive IHF-binding palindrome
RNase	ribonuclease
$T_{\frac{1}{2}}$	reaction half time
T4 endoVII	T4 endonuclease VII
TAE	tris/acetate EDTA
TBE	tris/borate EDTA
TEMED	N-N-N'-N'-tetra-methyl-1,2-diamino-ethane
TM	tris-magnesium buffer
T_{MN}	dinucleotide stability constant
Tn	transposon
Tris	2-amino-2-(hydroxymethyl)-1,3-propandiol
Triton X-100	Octylphenoxypolyethoxyethanol
ts	temperature sensitive
WT	wild-type
Xgal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

Table of Contents

CHAPTER 1:

Introduction

Figure 1.1 bke palindrome in linear and cruciform conformations (Murchie and Lilley, 1987).

```
TGTGGATCCGGTACCAGAA:TTCTGGTACCGGATCCTCT
ACACCTAGGCCATGGTCTT:AAGACCATGGCCTAGGAGA
```

```
              ↓ ↑

              AT
              A  T
              GC
              AT
              CG
              CG
              AT
              TA
              GC
              GC
              CG
              CG
              TA
              AT
              GC
              GC
          TGT   TCT
          ACA   AGA
              CG
              CG
              TA
              AT
              GC
              GC
              CG
              CG
              AT
              TA
              GC
              GC
              TA
              CG
              T  A
              TA
```

2

1.1 Introduction

A DNA palindrome is a sequence in which complementary strands read the same in the 5'-3' direction. A cruciform structure may be formed by the denaturation of a DNA palindrome and subsequent intrastrand base pairing in a process known as extrusion (Figure 1.1). The intersection between interstrand and intrastrand pairing is the four-way junction, the intrastrand paired region is the stem, and the single-strand region at the end is the loop. The equivalent structure formed by a single strand DNA palindrome is a hairpin. Throughout this thesis, "palindrome length" refers to the whole sequence; the palindrome in Figure 1.1 is 32 base pairs (bp). If there is a non-palindromic spacer in the centre of a palindrome, the whole sequence is an inverted repeat (or imperfect palindrome) and the spacer an asymmetry.

The existence of palindrome and cruciform DNA has been considered for many years, almost since the discovery of the structure of DNA itself (Watson and Crick, 1953; Platt, 1955). Despite this awareness, DNA has too often been considered a "rigid-rod" type structure. However, within the last 10-15 years there has been an explosion in knowledge of unusual (non-B-) DNA structures, including cruciforms, H-DNA, Z-DNA, triple, and tetrastrand helixes (reviewed by Wells, 1988). To complement this, unusual DNA structures have been implicated in various genetic diseases including fragile X syndrome (Kremer *et al.*, 1991) and Kennedy's disease (La Spada *et al.* 1991).

Many of the sequences responsible for these structures have proven difficult to clone and study in *Escherichia coli* (Collins and Hohn, 1978), and may also cause lesser problems in eukaryotes (Henderson and Petes, 1993; Ruskin and Fink, 1993; Gordenin *et al.*, 1993). Long DNA palindromes introduced into *E. coli* cause two related phenomena known as palindrome-mediated instability and inviability (Courey and Wang, 1983; Leach and Stahl, 1983). Chalker *et al.* (1988) discovered that mutations in *sbcC* are sufficient to alleviate the inviability associated with long DNA palindromes, and this has enabled their study in *E. coli* (Chalker,

3

1990; Allers, 1993). This thesis is concerned with elucidating the mechanism for inviability in *E. coli*, including:

1. The means by which long DNA palindromes cause inviability, and whether it is a consequence of the formation of an unusual DNA secondary structure.
2. The extent to which the inviability caused by long DNA palindromes may be modelled through cruciform extrusion *in vitro*.
3. The secondary structure of DNA loops *in vivo* and their effect upon inviability.

These objectives are reiterated in more detail later in this introduction, in addition to a short description of the experimental design. The remainder of this chapter is an introduction to DNA cruciforms, palindrome-mediated inviability, and other topics related to this thesis.

1.2 Cruciform Extrusion *in Vitro*

The most unusual aspect of cruciform extrusion *in vitro* is possibly the slow reaction kinetics. DNA sequence, ionic conditions, and temperature all exert an effect, but the reaction half-time ($T_{1/2}$) may still be in the minutes-hours range even under optimal conditions (Gellert *et al.*, 1979; Mizuuchi *et al.*, 1982; Courey and Wang, 1983; Gellert *et al.*, 1983). Two principal pathways for *in vitro* extrusion have been formulated, the first described is S-type.

S-type Extrusion

S-type cruciform extrusion (Lilley and Markham, 1983; Lilley, 1985; Murchie and Lilley, 1987; Sullivan and Lilley, 1987; Murchie *et al.*, 1992) is characterized by a requirement for cations (Salt dependent), relatively moderate activation entropies and enthalpies, and a profound effect of central sequence changes on the reaction kinetics. In S-type extrusion (Figure 1.2a), it is postulated that an initial melting limited to the central base pairs is followed by protocruciform formation and intrastrand base pairing, and then branch migration to the fully extruded cruciform.

4

The transition state for the reaction is probably the central melted bubble or protocruciform (Murchie and Lilley, 1987). Central base opening is believed to involve 6-10bp and the whole process is facilitated by negative supercoiling.

Cations are required for specific ion binding in the transition state, specifically to shield phosphate-phosphate repulsions in the anionic cavity of the protocruciform. Sullivan and Lilley (1987) studied the efficacy of a variety of metal ions and related species to promote extrusion. Group IA cations and tetramethylammonium are optimal in the 50mM to 60mM range, whereas Group IIA cations (and selected transition metal ions, notably manganese (II)) are effective over a wide range, down to 200μM. Hexamminecobalt (III) and the polyamines promote extrusion at concentrations as low as 15-40μM. All the other ions tested are ineffective. A positive correlation was found between the rate of extrusion and the ionic radius. The moderate enthalpies and entropies of activation for S-type extrusion indicate that the reaction involves the breaking and reforming of a relatively small number of base pairs during the transition step (Courey and Wang, 1983; Gellert *et al.*, 1983; Sinden and Pettijohn, 1984).

Sequence changes have been informative in determining the pathway of extrusion. Murchie and Lilley (1987) made a systematic study on the effect of central sequence changes to the extrusion of a short palindrome, bke. The difference between the sequences that have the greatest and least effect on the extrusion kinetics was 2000-fold, whereas mutations outside the central 8-10 nucleotides typically affected the rate up to 10-fold. This confirms the central base opening hypothesis. Courey and Wang (1988) and Zheng and Sinden (1988) also established that palindromes with a greater central A/T content are kinetically more active. The degree of methylation of a sequence is also relevant. Murchie and Lilley (1989) reported that methylation at the N^6-adenine of the sequence GGATCC increases the rate of extrusion, whereas C-5-cytosine methylation decreases the rate. This is in agreement with the known effects of adenine and cytosine methylation on DNA stability (Collins and Myers, 1987).

Figure 1.2a. S-type pathway for cruciform extrusion *in vitro* (from Murchie *et al.*, 1992). There is an initial melting of central base pairs (step 1), followed by intrastrand pairing and protocruciform formation (step 2), then finally branch migration to the fully extruded cruciform (step 3). The transition state is probably the central melted bubble or the protocruciform. Central base opening is believed to involve about 8-10bp.

b. C-type pathway for cruciform extrusion *in vitro* (from Murchie *et al.*, 1992). There is a large scale opening of the A/T-rich DNA (step 1), and the unwinding is propagated through the palindrome. Intrastrand pairing of the entire palindrome to the fully extruded cruciform occurs in a single step (step 2).

a.

S-type Extrusion

Step 1

Step 2

Step 3

b.

C-type Extrusion

Step 1

Step 2
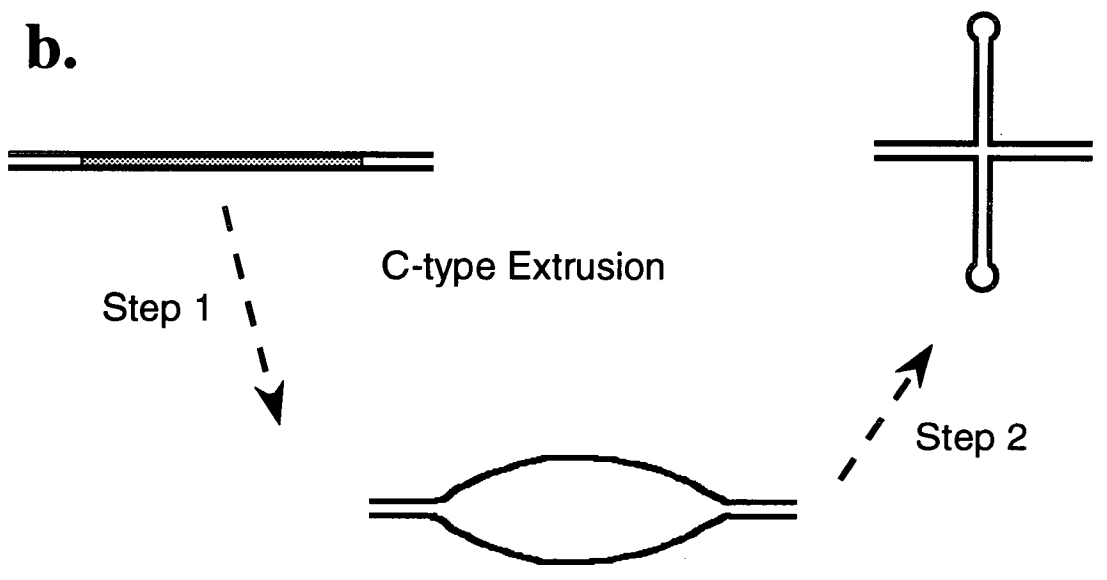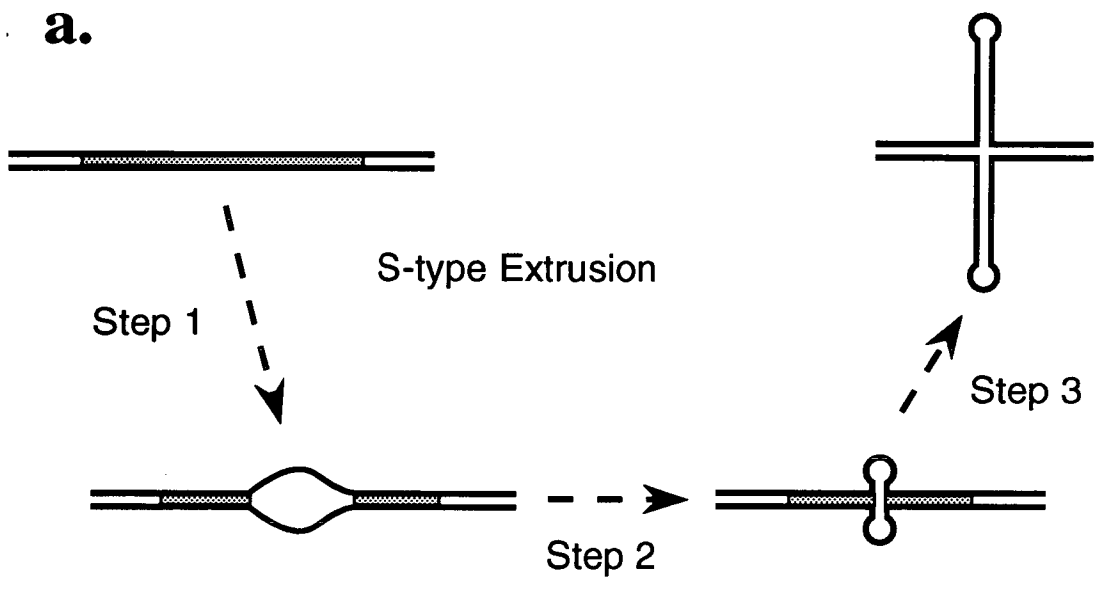
C-type Extrusion

C-type cruciform extrusion is exhibited by a small group of DNA palindromes that extrude in low salt. (Lilley, 1985; Sullivan and Lilley, 1986; Sullivan *et al.*, 1988; Bowater *et al.*, 1991; Murchie *et al.*, 1992). The extrusion process has very large enthalpies and entropies of activation (typically 180kcal mol^{-1}), and the palindromes must be flanked by very A/T-rich DNA (typified by the inverted repeat of ColE1, after which they are named; Oka *et al.*, 1979). In C-type extrusion, the A/T tract is believed to initiate the base opening; the absence of salt helps destabilise the helix (Figure 1.2b). The melted region is propagated through the palindromic DNA and the intrastrand base pairing of each cruciform arm occurs in one step. The large scale opening is responsible for the high enthalpies and entropies of the reaction. A/T-rich sequences from many sources confer C-type kinetics on any palindrome in *cis* (Sullivan and Lilley, 1986; Sullivan *et al.*, 1988). Cruciform extrusion of alternating (AT)$_n$ sequences (that are themselves palindromic) may occur via an intermediate C- and S-type mechanism (Greaves *et al.*, 1985; Haniford and Pulleyblank, 1985; Murchie *et al.*, 1992). The difference is that (AT)$_n$ sequences melt cooperatively at low temperatures without any detectable kinetic barrier.

1.3 Cruciform DNA Structure

In the following section, the individual components of DNA cruciforms are reviewed in detail.

Four-way Junctions

Cruciform junctions are structurally analogous to the Holliday junction (Holliday, 1964) by virtue of the common feature of a four-way junction. Furthermore, it is the formation of a protocruciform junction that may be implicated in the transition state for the cruciform extrusion reaction (Murchie and Lilley, 1987).

Four-way junctions have a compact, highly ordered structure (Sigal and Alberts, 1972), although only recently has the molecular geometry been analysed in detail. They migrate aberrantly on polyacrylamide gels because their bent DNA

Figure 1.3 (reproduced with permission from Allers, 1993). Schematic representation of the stacked X-structure of the four-way junction (Murchie *et al.*, 1989). The DNA strands are a right handed cross of antiparallel molecules that minimise the unfavourable steric interactions. Note that two of the strands are continuous.
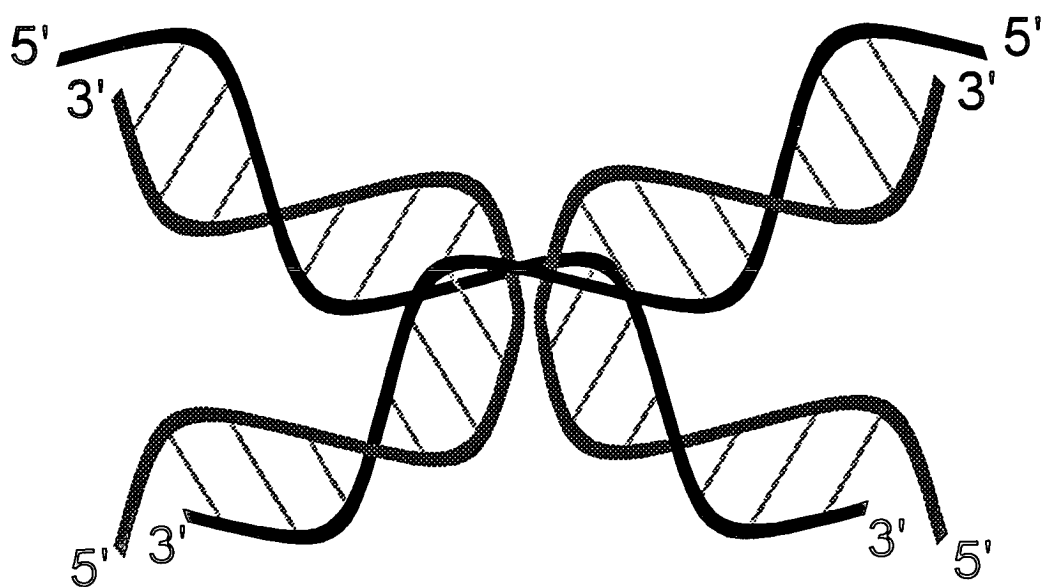
structure reduces the mobility in relation to the angle of the bend (Gough and Lilley, 1985; Cooper and Hagerman, 1987; Duckett et al., 1988). The structure has also been studied by fluorescence energy transfer (Murchie et al., 1989; Clegg et al., 1992), various enzymatic methods (Murchie et al., 1990, 1991), chemical probing (Gough et al., 1986; Churchill et al., 1988), and NMR analysis (Wemmer et al., 1985; Chen et al., 1991).

The optimal conformation is a stacked X-structure (Cooper and Hagerman, 1987; Duckett et al., 1988; Clegg et al., 1992), that minimises the unfavourable electrostatic interactions by allowing helix-helix stacking and an increase in the base pair stacking (Cooper and Hagerman, 1989; Murchie et al., 1989; see Figure 1.3). Within a four-way junction two of the strands are continuous and full base pairing is observed. Two isomers of each junction are possible, the dominant one being determined by the stability of the stacking interactions, although the precise structure is also strongly influenced by the nature and concentration of cations (Diekmann and Lilley, 1987; Cooper and Hagerman, 1989; Duckett et al., 1990; Duckett et al., 1992). Junction folding is driven by the unfavourable entropy of exposing the bases to the solvent, which brings the charged phosphate groups closer together. Cations are therefore required to screen the phosphate-phosphate repulsions. Group IIA metal ions (Mg^{2+}, Ca^{2+}) enable folding at concentrations as low as 80µM. Hexamminecobalt (III) and spermine are as effective at 2µM and 25µM, respectively. However, the monovalent cations of group IA (Na^+) are only partially effective in folding the junctions and much higher concentrations are necessary. Even then, a less symmetrical conformation is adopted and the thymidine bases at the junction remain reactive to OsO_4 (Duckett et al., 1990).

Hairpin Loops

DNA loops are conventionally represented as four unpaired bases. While Haasnoot et al. (1985, 1986, 1987) predicted an optimum loop length of four bases, the exact details of hairpin folding have been found to depend upon the base sequence of the loop and stem. It has also become clear that a single loop may adopt several distinct conformations (Erie et al., 1992). Unfortunately, all published

experiments were performed under differing conditions so that direct comparisons between them are difficult.

Senior *et al.* (1988) and Paner *et al.* (1990) both studied the effect of base composition on the thermal stability of hairpins. It was found that a hairpin with four thymines in the loop is the most stable and four adenines least stable. The extra stability of the thymine loop may be due to stacking between unpaired T residues of the loop (also between T residues and the stem; Senior *et al.*, 1988; Paner *et al.*, 1990) or a wobble T•T pair forming between the two terminal thymines of the loop (Haasnoot *et al.*, 1986; Blommers *et al.*, 1991). Orbons *et al.* (1986, 1987), reporting on "extrastable" loops, determined that the octanucleotide C*GC*GTGC*G (C* are C-5-methylated cytosines) forms a minihairpin in which the central -GT- part is a loop of two nucleotides only.

More work has been published on RNA hairpins and it is well established that a large proportion of rRNA tetra-loop sequences are of the UNCG or GNPuA family (N= A, C, G, U; Pu= A, G) and that they form unusually stable loops. In the case of the former sequence this is due to the closure of the loop by a wobble U•G base pair (Cheong *et al.*, 1990; Antao *et al.*, 1991; Hirao *et al.*, 1992; James and Tinoco, 1993). Therefore, attempts have been made to formulate an equivalent DNA consensus for unusually stable hairpins.

Antao *et al.* (1991) reported that DNA hairpins with the sequences C(GNPuA)G or G(CTTG)C at the centre are unusually stable, but that C(TTCG)G has no specific internucleotide loop interactions. Blommers *et al.* (1989, 1991) conducted a more detailed study on DNA hairpins and concluded that PyTTPu loop sequences (Py= C, T) enhance hairpin stability whereas PuTTPy sequences do not, agreeing in part with Antao *et al.* (1991). It was found that if the sequence is TTTA then a Hoogsteen T•A base pair is formed, stacked on the 3' end of the B-helical stem, and the second base of the loop is folded into the minor groove. If the sequence is CTTG then a normal Watson-Crick C•G base pair is formed (Blommers *et al.*, 1989, 1991). Additionally, it is not important whether the four membered sequence is closed by a GC or AT base pair. However, replacing the central TT by AA limits the hairpin to a four-membered loop (Blommers *et al.*, 1989, 1991). These results

corroborated in part by Pieters *et al.* (1989, 1990), Rinkel and Tinoco (1991), and Kallick and Wemmer (1991).

Pieters *et al.* (1989, 1990) determined that the sequences CG(CTAG)CG and CG(C^{a}TAG)CG (C^a arabinose) both have two residue hairpin loops, differing from Blommers *et al.* (1989) in that the central base sequences are -TA- rather than -TT-. Rinkel and Tinoco (1991) found that a DNA dumbbell with a loop sequence of CGTG also has a two residue loop (ie PyGTPu). Kallick and Wemmer (1991) identified the loop sequence ATAT as four-membered (ie PuTAPy). Howard *et al.* (1991) and Raghunathan *et al.* (1991) studied five DNA sequences with putative loops of the sequence: CGCG, CGAG, CIAG, AGCG, AGCI. The first three form two-residue loops, confirming the Py_ _Pu consensus identified by Blommers *et al.* (1989), except that CGAG and CIAG will have two membered *purine* loops. Xodo *et al.* (1988, 1991) found that hairpins with the sequence (CG)$_3$AT(CG)$_3$ or (CG)$_3$TA(CG)$_3$ have melting profiles that are consistent with two residue loops, also contrary to the prediction of Blommers *et al.* (1989).

1.4 DNA Supercoiling

Negatively supercoiled DNA is strained so that the B-helix is destabilized (Majumdar and Thakur, 1985; Benham, 1992; Sen, 1992). DNA opening at the weakest site is accompanied by a decrease in the superhelical stress of the whole molecule (Frank-Kamenetskii, 1990). Therefore, in a negatively supercoiled molecule the unfavourable free energy of cruciform extrusion (ΔG_C) may be offset by the partial global relaxation of the whole molecule. ΔG_C is unfavourable because it includes the cost of any unpaired bases in the loop and four-way junction; it has been estimated to be 15-20 kcal mol^{-1} (Courey and Wang, 1983; Gellert *et al.*, 1983; Lilley and Hallam, 1984). The requirement for supercoiling in cruciform extrusion was first demonstrated *in vitro* (Lilley, 1980; Panayotatos and Wells, 1981). Many experiments *in vitro* have confirmed that a threshold negative superhelical density is required for cruciform extrusion to occur by demonstrating a large shift in the rate of extrusion between two superhelical density values (Singleton and Wells, 1982;

Courey and Wang, 1983; Gellert *et al.*, 1983; Lilley and Hallam, 1984; Naylor *et al.*, 1986; Courey and Wang, 1988).

1.5 Branch Migration

Branch migration is essentially a random walk process since each step on its own is energetically neutral in linear DNA due to two base pairs merely changing branches (Lee *et al.*, 1970; Thompson *et al.*, 1976). However, for cruciform extrusion in negatively supercoiled DNA there will be a thermodynamic gain for each step due to the partial relaxation of the whole molecule. In most cases, linear DNA has been used as a substrate for experiment, so the relevance to cruciform extrusion may not be direct.

Rates of branch migration in linear DNA are relatively fast: in the order of several thousand base pairs per second. Base mismatches disrupt the helical stacked structure of junctions, and consequently impede branch migration. In the presence of Mg^{2+} ions, branch migration through single base pair heterologies is almost completely blocked, although this is not the case with Na^+ ions (Panyutin and Hsieh, 1993). This supports the notion that the DNA melting at the junction during branch migration involves a few base pairs only (Robinson and Seeman, 1987). *In vivo*, a relevant enzyme could accelerate the effective rate of branch migration by preventing movement in one direction. In this way the junction would be constrained at a heterology until it passes through (Robinson and Seeman, 1987; Panyutin and Hsieh, 1993).

1.6 DNA Melting

DNA melting is a complex and poorly understood process. A brief review of the literature is given below as an introduction to the subject because melting is relevant to cruciform extrusion.

Many studies have attempted to determine the nearest neighbour stacking interactions that are necessary to melting, primarily Gotoh and Tagashira (1981), but

more recently Wartell and Benight (1985), Breslauer *et al.* (1986) and Doktycz *et al.* (1992). The possible opening pathways have been analysed by Ramstein and Lavery (1988, 1990), Briki *et al.* (1991a and b), and Giradet and Ramstein (1992). Base pair opening lifetimes have been measured to be in the region of 10ms at room temperature (Guéron *et al.*, 1987).

(AT)$_n$ DNA is unusual in that it undergoes cruciform extrusion (and therefore melting) without any detectable kinetic barrier. Also, Kowalski *et al.* (1988) found that stable (rather than transient) DNA unwinding accounts for single strand nuclease hypersensitivity in supercoiled A/T-rich regions at low salt concentrations.

1.7 Cruciform DNA in *E. coli*

Whether DNA cruciforms exist *in vivo* has been in dispute until recently. This section is devoted to: (a) the conditions that may determine whether cruciforms exist *in vivo* (b) some examples where cruciforms have been detected.

Conditions for Cruciform Formation *in Vivo*

A threshold level of negative supercoiling is necessary to make extrusion energetically favourable, as already mentioned. Circular plasmid DNA extracted from a cell has a superhelical density similar to this threshold level, but does not reflect the unconstrained levels in the cell since *in vivo* half of the supercoiling is believed to be constrained by interaction with nucleoproteins such as HU (Broyles and Pettijohn, 1986). *In vivo* assays have confirmed that the unconstrained superhelical density is about a half that of naked supercoiled DNA, representing about one quarter of the required energy due to the quadratic relationship between superhelical density and supercoiling (Greaves *et al.*, 1985; Lilley, 1986; Bliska and Cozzarelli, 1987). Also, early studies predicted that the kinetics of cruciform extrusion may be too slow compared to the chromosome replication rate for an appreciable proportion of palindromes to be extruded (Courey and Wang, 1983; Gellert *et al.*, 1983).

The first efforts to detect cruciforms *in vivo* possibly failed because insensitive and intrusive assays were used and the palindromes were insufficiently long. Typical

experiments have used palindromes of less than 100bp, coupled with low temperature extraction (Courey and Wang, 1983; Lyamichev *et al.*, 1983) or psoralen crosslinking (Sinden *et al.*, 1983). Perhaps not surprisingly, evidence for the formation of cruciform structures *in vivo* has been sparse for many years.

Increasingly, it has been realised that the level of unconstrained supercoiling in the cell may be dynamic, and it is known that in *E. coli* there are 40-50 independent domains of supercoiling (Broyles and Pettijohn, 1986; Lilley, 1986). Also, the superhelical density of cellular DNA responds directly to environmental factors such as osmotic shock (McClellan *et al.*, 1990; Dayn *et al.*, 1991). The histone-like protein H1 (H-NS) is implicated in having a regulatory role on the level of chromosomal supercoiling that underlies the osmotic dependent regulation of gene expression (Hulton *et al.*, 1990). Transcription also effects DNA supercoiling. Liu and Wang (1987) proposed the twin domain model of transcription induced supercoiling where DNA upstream of the RNA polymerase accumulates negative supercoils and DNA downstream positive supercoils. Negative supercoiling would be highest proximal to the polymerase. If there are multiple and opposing transcripts causing local variations in supercoiling, and the chromosome is divided into domains, then the global level of supercoiling in the cell is irrelevant with respect to the formation of cruciforms.

In addition, there is a small single stranded bubble in the DNA template during transcription which could perhaps overcome the kinetic barrier to cruciform formation. Processes such as DNA replication or RecBCD-mediated unwinding also generate single stranded DNA in the replication fork or the unwound bubble.

In summary, given that extrusion may be catalysed *in vivo*, and that the level of supercoiling is not constant, it is possible that many palindromes exist as cruciforms *in vivo*, but that they do so only transiently because they are relatively small. Long DNA palindromes have not been investigated so thoroughly because of their associated inviability problems.

Evidence for Cruciforms *in Vivo*

Panayotatos and Fontaine (1987) mapped the cleavage site *in vivo* of T7 endonuclease I to the ColE1 palindrome. Since the ColE1 palindrome is flanked by highly A/T-rich DNA, helix destabilisation could occur easily *in vivo*. McClellan *et al.* (1990) and Dayn *et al.* (1991) assayed the appearance of OsO_4/chloroacetaldehyde sensitivity under different stresses (eg. osmotic shock) which increase the cellular negative supercoiling levels. Chemical sensitivity was greatest at the centre of the $(AT)_{34}$ palindrome studied, consistent with cruciform formation *in vivo*. Dayn *et al.* (1992) were able to observe cruciforms in relation to transcription induction. Sinden *et al.* (1991) and Zheng *et al.* (1991) designed "torsionally-tuned" palindromes and probed for cruciforms *in vivo* with a psoralen cross-linking assay. For their relatively A/T-rich centred palindromes, they concluded that cruciform formation is dependent upon DNA supercoiling *in vivo*, and the location within a plasmid with respect to transcription. Allers (1993) has recently demonstrated that long DNA palindromes in λ phage adopt a methylation resistant structure that is consistent with the presence of an unusual DNA secondary structure *in vivo*.

1.8 Palindromes in Prokaryotes and Eukaryotes

There is a distinct contrast between the nature and distribution of DNA palindromes in prokaryotes and eukaryotes. In prokaryotes, naturally occurring long DNA palindromes are rare and tend to be imperfect. In comparison, DNA palindromes in eukaryotes are relatively common and frequently perfect, comprising up to several percent of the genome (Wilson and Thomas, 1974).

Transcription

The inverted repeat sequence of DNA palindromes means that they are attractive sites for specific protein interactions since cooperative binding of two protein monomers may increase the specificity of reactions (Park *et al.*, 1993). Consequently, short DNA palindromes are common in regulatory regions (Müller and Fitch, 1982), and implicated in the control of gene expression. Dimri and Das (1990)

identified a 14bp inverted repeat in the DNA gyrase promoter of *Klebsiella pneumoniae* between the consensus -10 and -35 sequence. A cruciform could interfere with promoter recognition where the stimulus for extrusion could be provided by DNA gyrase introducing negative supercoils, and making the reaction favourable. The result would be the homeostatic control of gene expression.

DNA palindromes are also implicated in transcription termination, but indirectly after transcription of the sequence to RNA. The hairpin structure causes RNA polymerase to pause and terminate transcription (for a review see Platt, 1986).

Replication

DNA palindromes at origins of replication are implicated in the initiation of DNA synthesis, and probably serve as binding sites for the replication complex. *E. coli* has inverted and direct repeats at the origin (Meijer *et al.*, 1979; Hirota *et al.*, 1979) as does λ, with a 28bp near perfect inverted repeat to the right of the λ origin (Hobom *et al.*, 1979). The *E. coli* inverted repeats have central asymmetries, perhaps suggesting that they are active as protein binding sites rather than cruciform structures because a large asymmetry may make extrusion kinetically forbidden under physiological conditions. Prokaryotic plasmids also have palindromic sequences at their replication origins (Lin and Meyer, 1987; Wang *et al.*, 1993).

Mobile Genetic Elements

Bacterial insertion sequences (IS elements) are flanked by almost perfect inverted repeats, each repeat of which is up to 40bp long (eg. IS*1*, ~30bp inverted repeats; Ohtsubo and Ohtsubo, 1978). However, they flank large non-palindromic sequences. More relevantly, composite (class I) transposons consist of two copies of an IS element in inverted orientations, generally separated by a gene coding for drug resistance (see Kleckner, 1989, for a review of Tn*10*). If the resistance gene is removed *in vitro*, a long, almost perfect palindrome is generated. The instability of this palindrome, and the inviability that it confers on its host plasmid have been studied *in vivo* (Collins, 1981; Collins *et al.*, 1982).

Repetitive Palindromic DNA

Prokaryotic repetitive palindromic DNA has been well characterized by DNA sequence analyses in *E. coli*, *Salmonella typhimurium*, and other members of the *Enterobacteriaceae*. Repetitive Extragenic Palindromes (REP sequences; Higgins *et al.*, 198 ; also known as PU sequences (Gilson *et al.*, 198)) are ~40bp, highly conserved palindromes where a G/C-rich inverted repeat is generally separated by a more variable A/T-rich spacer. They exhibit species specificity and occur either singly or in clusters of up to six elements. The clusters themselves are conserved, and have been termed BIMEs (Bacterial Interspersed Mosaic Elements; Gilson *et al.*, 1991). REPs tend to be found in extragenic sequences or the 3' termini of genes, and although they are not transcription terminators, they may have a role in stabilising mRNA from degradation (Stern *et al.*, 1988). However, it is more likely that the primary role of REPs is in chromosomal structure and organisation. REPs bind DNA gyrase, and the histone-like protein, HU, enhances this activity. REPs may also be the site of gyrase action on the chromosome (Yang and Ames, 1990). Additionally, DNA polymerase I forms a stable complex with a REP sequence (Gilson *et al.* 1990), implying that the REP could be the preferred entry site for DNA polymerase I.

Recently, RIP or RIB (repetitive IHF-binding palindrome, Oppenheim *et al.*, 1993; reiterative IHF BIME, Boccard and Prentki, 1993) elements have been described. RIP/RIB elements consist of two inverted REPs separated by a single IHF (Integration Host Factor, a histone related protein) recognition site. It is possible that their function is to insulate domains from changes in supercoiling.

Intergenic Repeat Units (IRUs) or Enterobacterial Repetitive Intergenic Consensus sequences (ERICS) have also been identified in *E. coli* (Sharples and Lloyd, 1990; Hulton *et al.*, 1991). They are 124-127bp imperfect inverted repeats that are less frequent than REPs. No function has yet been assigned to them, although there is speculation that they are a form of "selfish" DNA.

Human Genetic Disease

Increasingly, unusual DNA structures are believed to have a role in human genetic disease. A molecular analysis of the products of illegitimate recombination linked them to palindromes, A/T-rich regions, alternating purine/pyrimidine sequences

and Alu family repeats (Stary and Sarasin, 1992). Fragile X-syndrome (Kremer *et al.*, 1991) and Kennedy's disease (LaSpada *et al.*, 1991) are associated with a variation in length of DNA triplet repeats, many of which have the potential to adopt non-B-DNA type structures. The triplet expansion of the repeats is believed to be caused by a DNA replication block inducing replication slippage. DNA structure and its relation to disease is reviewed by Sinden and Wells (1992).

1.9 Long DNA Palindromes in *E. coli*

Naturally occuring palindromes longer than ~40bp are rare in *E. coli*, and perfect DNA palindromes longer than 100-150bp cannot generally be propagated in wild-type (WT) *E. coli*. Either a replicon containing a long DNA palindrome is so poorly replicated that it is inviable, or the palindrome itself is unstable so that it undergoes partial or complete deletion (inviability and instability; Collins, 1981; Leach and Stahl, 1983; Leach and Lindsey, 1986; Shurvinton *et al.*, 1987; Lindsey and Leach, 1989). The inviability and instability is relevant because it has caused continuing problems for the cloning of eukaryotic DNA in *E. coli*. Collins and Hohn (1978) first noted the absence of palindrome sequences in a cosmid library, and subsequent work by Nader *et al.* (1985), and Wyman *et al.* (1985) has confirmed this. Similar problems using *Streptococcus sp.* indicate that the problem could be widespread amongst prokaryotes (Behnke *et al.*, 1979). Long DNA palindromes clearly exert a strong biological effect in *E. coli*, but there is still little direct evidence that this is due to the formation of cruciform structures *in vivo*.

Over the past ten years, many of the cloning problems caused by palindromes have been overcome in hosts mutant for *recB*, *recC*, *sbcB*, and particularly *sbcC* (Collins *et al.*, 1982; Leach and Stahl, 1983, Chalker *et al.*, 1988). The *sbcC* mutation by itself is sufficient to allow the plating of λ phage with long palindromes; *sbcC* is now known to be one partner in a two gene system with *sbcD* (Gibson *et al.*, 1992). Together they have been implicated in palindrome-mediated inviability, in addition to their role in the cosuppression of the recombination deficiency of *recBC* mutants (Lloyd and Buckman, 1985; Gibson *et al.*, 1992).

Chalker *et al.* (1988) reported that λ *red gam pal* phage plate on *E. coli sbcC* efficiently, and that further mutations in *recB/recC/recD* increase the λ burst size. However, the increase is not specific to palindrome-containing λ because it results from permitting rolling circle replication. The mutation *sbcC* also permits as good a recovery of supercoiled λ DNA molecules as from a *recBC sbcBC* host. Very long palindromes (>700bp) remain completely unstable and delete rapidly to more stable, shorter palindromes in *E. coli sbcC* (Leach and Stahl, 1983).

1.10 Palindrome-Mediated Instability and Inviability

It is likely that inviability and instability (see 1.9) are linked by a common cause: unusual DNA secondary structure formation *in vivo*. The following section is a review of the factors that govern instability and inviability.

Palindrome Length

Generally, longer palindromes have greater levels of associated inviability and instability, although the exact levels of each are highly context dependent. Collins (1981) and Mizuuchi *et al.* (1982) showed that a 976bp Tn5 derivative palindrome is unstable and a head to head palindromic plasmid dimer (>few kb) causes inviability. Sinden *et al.* (1983) and Gellert *et al.* (1983) did not observe any deletions with palindromes of 76 and 114bp. However, Lilley (1981) identified a 260bp palindrome that was unstable and caused inviability. Commonly, inviability tends to occur above a threshold length of approximately 100-200bp, and although smaller palindromes tend to be viable, they may be unstable. Palindromes may be both stable and viable, but still have other effects. Warren and Green (1985) identified two 146/147bp palindromes that are stable, but have reduced copy number due to plasmid multimerisation. Single strand vectors tend to be less tolerant of palindromes and the threshold is reached at shorter sequence lengths (Müller and Turnage, 1986).

Asymmetry at the Palindrome Centre

Large asymmetries alleviate inviability and instability. If the 2750bp central region of Tn5 is removed (creating a palindrome of inverted IS elements), then there is a five-fold increase in *recA*-independent tandem IS element excision (Collins *et al.*, 1982). However, the asymmetry need not be so large to alleviate instability or inviability. Peeter *et al.* (1988) found that in *Bacillus subtilis* a 173bp asymmetry had an equal effect on deletion frequency compared with a 1.4kb asymmetry. Warren and Green (1985) identified 57bp as the shortest sequence necessary to alleviate inviability in an *E. coli* plasmid, although lengths greater than 150bp were required before instability was completely overcome. Zheng *et al.* (1991) and Sinden *et al.* (1991) studied palindromes with asymmetries of between 6bp and 14bp. A correlation was identified between the stability of base pairing in the hairpin stem and the deletion frequency in wild type cells.

Direct Repeats

Direct repeats are deletion endpoints which can influence the degree of instability of a palindrome. However, a palindrome is not a prerequisite for deletion between direct repeats. Egner and Berg (1981) reported that nine base-pair direct repeats were used as the deletion endpoint for the *recA*-independent Tn5 excision between terminal inverted repeats. Usually, the frequency of deletion between direct repeats (no palindrome present) is related to the repeat length (Peeters *et al.*, 1988; Pierce *et al.*, 1991). However, palindrome-mediated stimulation may promote the preferential use of shorter (4-5bp) sequences as deletion end points. The position of the repeat is relevant because the deletion often utilises one direct repeat in the potential hairpin arm (Pierce *et al.*, 1991; Weston-Hafer and Berg, 1991a). Most models for deletion postulate that the direct repeats are brought closer by the formation of a hairpin structure and the palindrome is deleted during replication (Ripley, 1982; Ripley and Glickman, 1983; Glickman and Ripley, 1984; Trinh and Sinden, 1993). There is some dispute as to whether a lagging strand bias to deletion between direct repeats exists that could arise from a deletion during replication (Trinh and Sinden, 1991; Weston-Hafer and Berg, 1991b).

Sequence Context

The sequence context may have a variable effect upon instability. DasGupta *et al.* (1987) reported that a single base position shift increases the instability of a 90bp palindrome 3000-fold. Kazic and Berg (1990) found large variations in instability for a palindrome inserted in either a mutant *bla* gene at λatt, the (in *E.coli*) *E.coli* chromosomal *lac* or F'*lac*. Some of the differences in deletion may be due to variations in local levels of negative supercoiling induced by transcription (Liu and Wang, 1987). Zheng *et al.* (1991) observed higher levels of cruciforms *in vivo* between divergent promoters than downstream of genes. The influence of sequence context on viability is less well documented. Hagan and Warren (1983) determined that a palindrome did not alter its lethality to the vector in various carrier plasmids. In contrast, λ vectors tolerate long DNA palindromes (in an *E. coli sbcC* host) whereas plasmids do not (Chalker, 1990).

DNA Replication

DNA replication is considered necessary to palindrome-mediated inviability and instability. If λ with an 8.4kb palindrome infects an *E. coli rec⁺ dnaB^{ts}* at the restrictive temperature (no DNA replication), there is no loss of the palindrome sequence or carrier replicon relative to a palindrome free λ (Shurvinton *et al.*, 1987). Likewise, if an *E. coli rec⁺* λ lysogen host is superinfected with a palindrome-containing λ, there is no λ DNA replication and no concomitant reduction in the levels of intracellular supercoiled λ DNA (again, relative to a palindrome free λ; Leach and Lindsey, 1986). Lindsey and Leach (1989) used demethylation to study the fate of λ DNA in an *E. coli rec⁺ dam* host and found that a decrease in the yield of λ *pal* with two newly synthesised (unmethylated) strands may occur without any loss of input (hemimethylated) strands. This result implies that inviability may be a consequence of slow replication.

In vitro evidence that palindromes halt DNA synthesis also exists (LaDuca *et al.*, 1983; Weaver and DePamphilis, 1984), and there has been some speculation as to whether the effect of palindromes is lagging strand biased: either due to the formation of hairpins between Okazaki fragments or the inability of the polymerase to reinitiate synthesis from within a lagging strand hairpin. Weston-Hafer and Berg

(1991b) did not observe a difference in the deletion frequency depending upon the direction of DNA replication, whereas Trinh and Sinden (1991) reported that lagging strand deletion was 10-20-fold higher.

Host Genotype

As mentioned in 1.9, *sbcC* was isolated (with *sbcB*) as a cosuppressor of defects in the cell associated with the *recBC* genotype. Mutation of *sbcC* on its own has no reported effects other than the alleviation of the instability and inviability effect of palindromic sequences; otherwise the cell is effectively wild-type. *sbcC* is now known to be one partner in a two gene system with *sbcD* (Gibson *et al.*, 1992). *sbcC* and *sbcD* are (individually) necessary and sufficient to alleviate palindrome-mediated inviability, in addition to their role in the cosuppression of the recombination deficiency of *recBC* mutants (Lloyd and Buckman, 1985; Gibson *et al.*, 1992). Clues to the function of *sbcC* and *sbcD* are covered below.

Evidence that SbcCD may be a nuclease was reported by Leach *et al.* (1992) who observed a resemblance of *sbcC* with gene *46* of T4 and *D13* of T5 and a similar resemblance of *sbcD* with genes *47* and *D12* from T4 and T5, respectively. Both the T4 and T5 genes have a similar genomic organisation to *sbcCD* (Gram and Rüger, 1985; Kaliman *et al.*, 1988) and all these proteins are members of the UvrA-related superfamily of nucleotide-binding proteins (Gorbalenya and Koonin, 1990). The T4 genes *46* and *47* are responsible for bacterial chromosome degradation after infection (Kütter and Wiberg, 1968), and T4 recombination is deficient in 46^- 47^- ^{T4} infections (Cunningham and Berger, 1977). Speculation that SbcCD is a nuclease is not inconsistent with the slow replication of palindrome DNA as a cause of replicon inviability, since slow replication could be because SbcCD acts on one arm only of a replication fork (Lindsey and Leach, 1989). Further confirmation that SbcCD is an exonuclease of double strand DNA *in vitro* has recently been reported (J. Connelly, D. Leach, personal communication).

The *sbcC* gene and λ *gam* are believed to interact (Kulkarni and Stahl, 1989). γ protein inhibits RecBCD enzyme to permit rolling circle λ replication in WT *E. coli* (Unger and Clark, 1972). Kulkarni and Stahl (1989) reported that whereas λ *gam pal* requires an *sbcC* mutation to plate, λ *gam*[+] *pal* plate on all the hosts tested, including

WT. The implication is that γ may inhibit SbcC in a similar manner to its inhibition of RecBCD. This observation is strengthened given that the *E. coli* genome may have arisen from two duplications so that homologous genes should be a quarter or half the way around the chromosome from one another (Riley and Anilionis, 1978). *sbcC* is located 180° from the *recBCD* genes on the genetic map (Lloyd and Buckman, 1985). This observation may be no more than a coincidence, but the hypothesis is further corroborated by analysis of another gene, *sbcD*, which lies downstream of *sbcC* and is a cosuppressor of *recBC*.

Palindrome-mediated inviability is also alleviated in *E. coli sbcD* (Naom *et al.*, 1989; Gibson *et al.*, 1992). *sbcD* has no independent promoter and overlapping stop/start codons with *sbcC*, implying that it is probably cotranscribed with *sbcC*. Sequence analysis has revealed that *sbcC* and *sbcD* may make up a multi-subunit enzyme akin to RecBCD since *sbcC* and *sbcD* have limited (but interrupted) homology to *recB* and *recC*, respectively. However, the homology is very limited. Furthermore, a putative *Bacillus subtilis* homologue of *sbcD* has been discovered downstream of *addAB*, the *B. subtilis* analogue of *recBC* (Sharples and Lloyd, 1993).

Instability is *recA*-independent and so considered to be an illegitimate recombination event. Other genes that have an effect are primarily those involved in recombination pathways such as RecF (Ishiura *et al.*, 1989), but the effects are pleiotropic and difficult to study in isolation. DNA repair genes (Lundblad and Kleckner, 1984; Schaaper, 1989) and DNA gyrase (Saing *et al.*, 1988) also influence the maintenance of palindromes, but are similarly pleiotropic in their effects.

1.11 Hypotheses to Account for Palindrome-Mediated Instability

There are two principal models:

Replication Slippage

Many schemes of replication slippage have been proposed to account for instability (Streisinger *et al.*, 1966; Ripley, 1982; Ripley and Glickman, 1983).

Glickman and Ripley (1984) and DasGupta *et al.* (1987) proposed that hairpin structures bring pairs of direct repeats closer together during DNA replication. This model is illustrated in Figure 1.4i (after DasGupta *et al.*, 1987). It is proposed that there is a transient dissociation of the nascent strand during DNA replication, followed by reannealing at the second copy of the direct repeat, and resumption of DNA synthesis. The remaining hairpin would be lost during post replicational repair or during subsequent rounds of replication. Replication slippage is analogous to copy choice recombination (Lederberg, 1955) and is believed to be also responsible for frameshift mutations (Ripley, 1982), gene amplification in eukaryotes (Hyrien *et al.*, 1988) and involved in specific human diseases (Sinden and Wells, 1992).

Secondary Structure Cleavage

The other model for instability proposes that the hairpin/cruciform structure is a substrate for a cleavage at the base by a conformation specific nuclease (DasGupta *et al.*, 1987; see Figure 1.4ii). Four-way junctions are structurally equivalent to Holliday junctions (Holliday, 1964; Sigal and Alberts, 1972; Duckett *et al.*, 1988) and are potential sites for recombination. After cleavage it is proposed that at the double-strand break produced there is limited 3'-5' exonucleolytic digestion, enabling the complementary strands of the direct repeats to anneal, ligate and be repaired. Many putative enzymes have been identified for the cleavage reaction. The action of RecBCD is proposed to be limited to four-way junctions in whose formation it participates (Taylor and Smith, 1990). RuvC has a cleavage specificity for homologous DNA which may limit its activity to partly extruded cruciforms (Dunderdale *et al.*, 1991). Many other four-way junction specific enzymes have been discovered in prokaryotes and eukaryotes: T4 endonuclease VII (Kemper *et al.*, 1981), T7 endonuclease I (deMassy *et al.*, 1985) and activities in *S. cerevisae* (West and Körner, 1985), HeLa cell extracts (Waldman and Liskay, 1988), and calf thymus (Elborough and West, 1990).

25

Figure 1.4. Hypotheses to account for palindrome-mediated instability. Reproduced with permission from Allers (1993).

i. Replication slippage a. Palindromic DNA flanked by direct repeats. b. Hairpin formation during replication. c. DNA synthesis impeded at hairpin. d. Transient dissociation of the nascent strand and reannealing at the second copy of the direct repeat. e. Resumption of DNA synthesis. f. Hairpin lost during post-replicational repair or subsequent rounds of replication.

ii. Secondary structure cleavage a. Palindromic DNA flanked by direct repeats. b. Cruciform (or hairpin) formation. c. Double-strand break introduced at junction. d. Limited 3'-5' exonucleolytic digestion, exposing direct repeats. e. Annealing of complementary sequences from the direct repeats. f. Ligation and repair synthesis.

i.

ii.

a.

b.

c.

d.

e.

f.

b.

c.

d.

e.

27
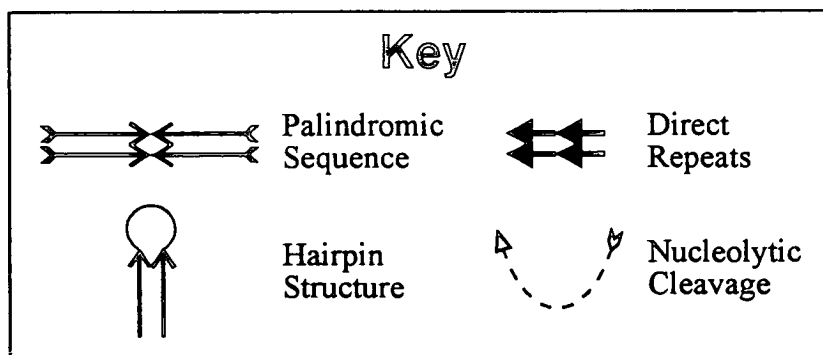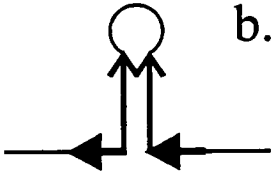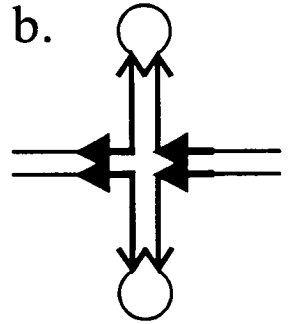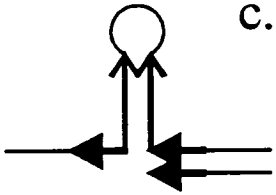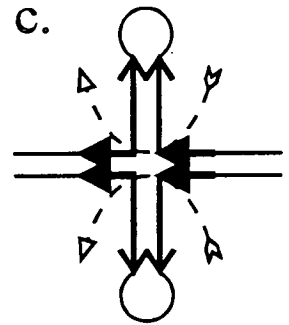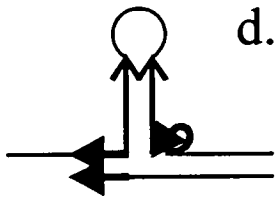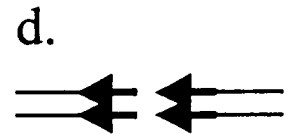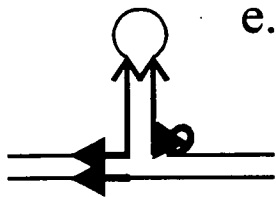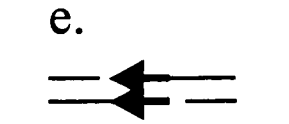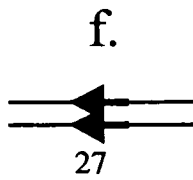
1.12 Hypotheses to Account for Palindrome-Mediated Inviability

The two inviabilty models are analogous to the respective instability models.

Secondary Structure Cleavage

As with instability, one inviability model also suggests that a conformation specific cleavage at the four-way junction is responsible for the loss of the replicon, except the religation step does not occur. Again, no nuclease that fits all of the conditions has yet been discovered. The most likely enzyme (RuvC) requires homologous DNA sequences at the junction point, that will be present during the (for cleavage) ×
extrusion of a cruciform only (Dunderdale *et al.*, 1991). Additionally, cruciform cleavage is inconsistent with the results of Lindsey and Leach (1989) which show that one of the parent strands of a λ *pal* is not degraded.

Replication Dependent Cleavage

D. Leach (personal communication) has proposed a replication dependent model that is illustrated in Figure 1.5. The leading strand of DNA replication encounters one of the stem loop structures as a continuation of the double helix and replication proceeds unhindered. However, on the lagging strand, replication cannot prime each new Okazaki fragment because of the persistence of the DNA hairpin. Repeated failure of initiation makes the hairpin a site for nuclease action, a double strand break is introduced, and the DNA degraded by SbcCD.

This model is consistent with the results of Lindsey and Leach (1989), in that one of the parental strands remains intact following replication. It is also compatible with a speculative role of SbcCD (D. Leach, personal communication), to prevent error prone DNA synthesis by removing the products of replication slippage before they cause permanent mutation. It is proposed that small, imperfect palindromes only rarely persist sufficiently long to cause a problem to replication. Long perfect palindromes always cause a replication block. In either case SbcCD removes the products of blocked or slipped replication.

28

Figure 1.5 Replication dependent cleavage hypothesis to account for palindrome-mediated inviability (D. Leach, personal communication). Reproduced with permission from Allers (1993).

a. Palindrome DNA in replicon. b. Cruciform extrusion. c. DNA synthesis at origin initiates. d. Synthesis towards cruciform. e. and f. Leading strand replication encounters hairpin as a continuation of double helix, replication proceeds without error. Lagging strand replication cannot prime each new Okazaki fragment because of the persistence of the DNA hairpin. g. Repeated failure of initiation on lagging strand (or persistence of hairpin) make it a site for SbcCD (nuclease?) action. h. Double strand break introduced. i. Degradation of one of the products of replication. j. No net gain in DNA synthesis.

Key

Palindromic Sequence		Origin of Replication
Cruciform Structure		Leading Strand Synthesis
Hairpin Structure		Lagging Strand Synthesis
SbcCD Nucleolytic Cleavage		Replication Fork
Exonucleolytic Degradation		

a.

b.

c.

d.

e.

f.

g.

h.

SbcCD

i.

j.
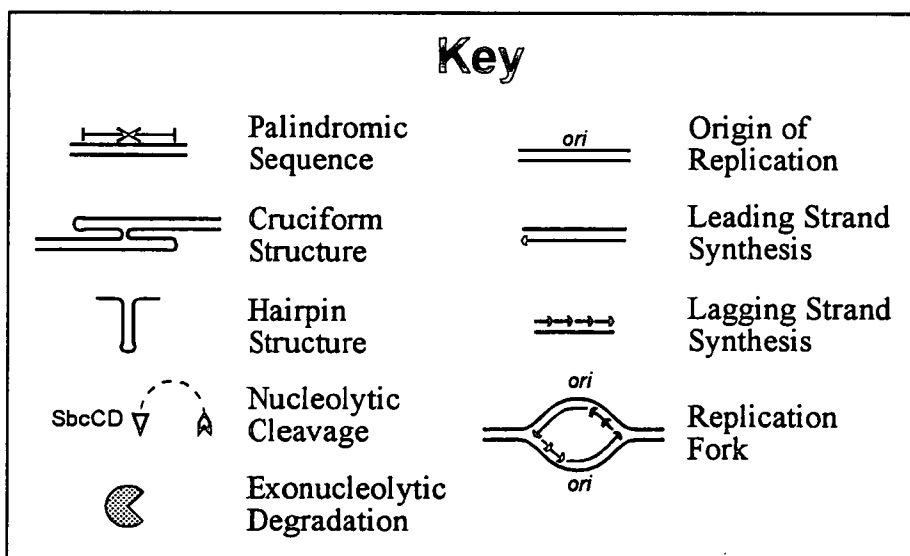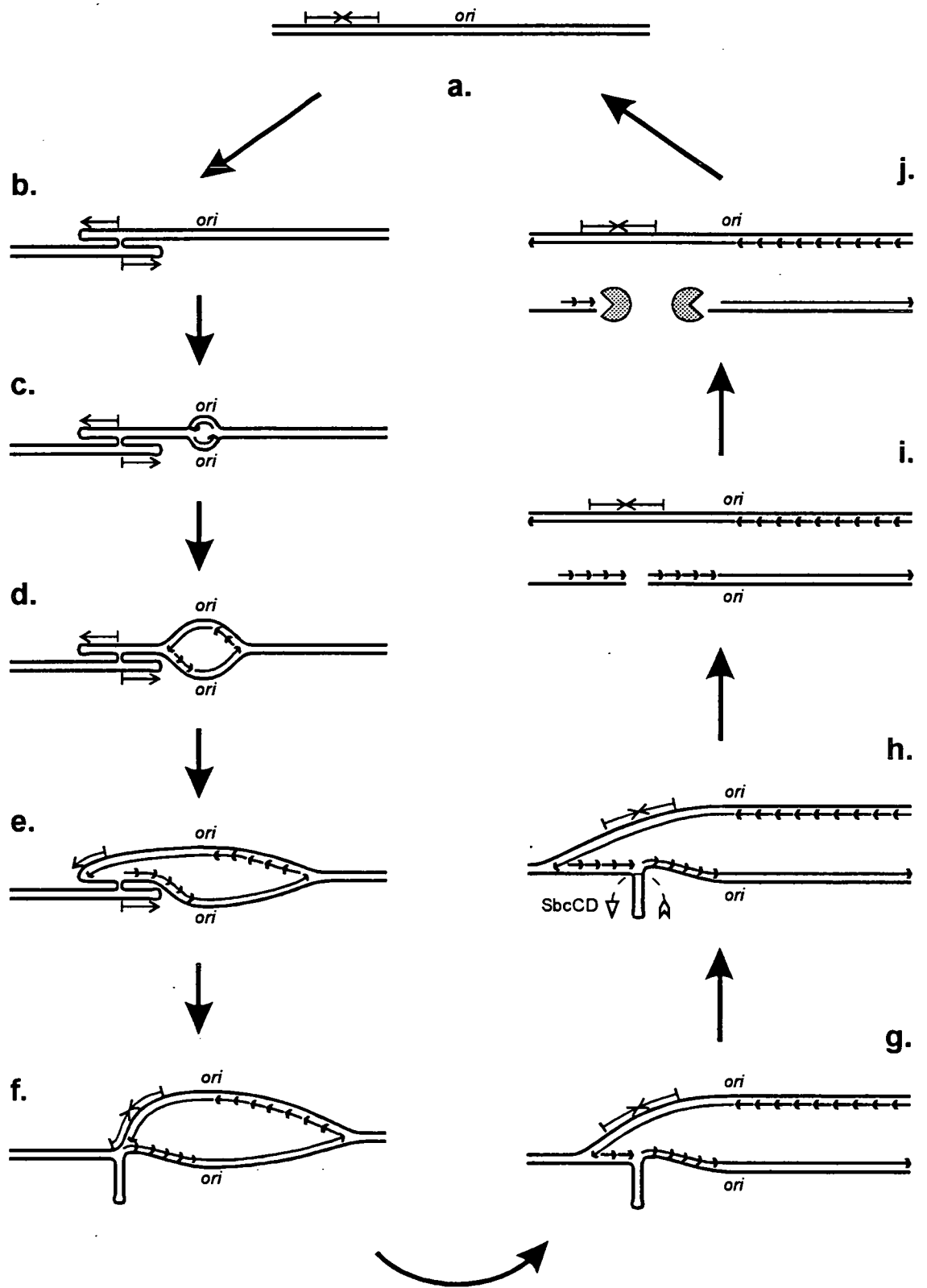
1.13 Aims of this Thesis

The intention of this thesis is to increase the understanding of unusual DNA secondary structures *in vivo*, specifically cruciforms, and the pathway by which they cause inviability in *E. coli*. The topics that will be addressed are:

1. The means by which long DNA palindromes cause inviability, and whether it is a consequence of the formation of an unusual DNA secondary structure.
2. The extent to which the inviability caused by long DNA palindromes may be modelled through cruciform extrusion *in vitro*.
3. The secondary structure of DNA loops *in vivo* and their effect upon inviability.

These aims are achieved by:

1. The development of a plaque size assay to measure the viability of palindrome-containing bacteriophage λ on an *E. coli sbcC* host. Under the standardised conditions used, plaque size is assumed to be an indirect product of the λ burst size, and thus a measure of its viability. Using this assay, the effects of nucleotide sequence changes on DNA secondary structure formation in *E. coli sbcC* are measured. Attempts to assay the effect of osmotic stress and orientation of insertion in λ are also described.
2. Analysis of the effects of central asymmetry on the propagation of palindromic DNA in bacteriophage λ across a broad range of *E. coli* hosts (in conjunction with A. F. Chalker).
3. Identification of *in vitro* extrusion conditions that partly model the plating of palindrome-containing λ phage on *E. coli sbcC*.

The results described in Chapter 3 have been published (Chalker, A. F., E. A. Okely, A. Davison, and D. R. F. Leach, 1993 *Genetics* **133**: 143-148; Appendix 2) and the results in Chapter 4 accepted for publication (Davison, A. and D. R. F. Leach, 1994 *Genetics* **137**: 1-8).

CHAPTER 2:

Materials & Methods

Materials

2.1 Strains

All *E. coli* strains, bacteriophage λ strains and plasmid vectors used in this work are described in Tables 2.1, 2.2, and 2.3.

2.2 Media

Bacteriological Media

The following quantities are for 1l final volumes made up in distilled water and sterilised by autoclaving for 20 minutes at 15 lb in^{-2}.

BBL Agar: 10g trypticase (Baltimore Biological Laboratories), 5g NaCl, 10g Bacto-agar (Difco). BBL Top Agar: As BBL agar, but containing only 6.5g Bacto-agar (Difco) per litre.

CA Agar (Plaque Area): 10g Bacto-casitone (Difco), 14.7g NaCl, 10g Bacto-agar (Difco). CA Top Agar (Plaque Area): As CA agar (plaque area), but containing only 5g Bacto-agar (Difco) per litre. CAS Agar (Plaque Area): As CA Agar (Plaque Area), except 0-20g NaCl. CAS Top Agar (Plaque Area): As CA Top Agar (Plaque Area), except 0-20g NaCl.

L Agar: 10g Bacto-tryptone (Difco), 5g yeast extract (Difco), 10g NaCl, 15g Bacto-agar (Difco). L Broth: 10g Bacto-tryptone (Difco), 5g yeast extract (Difco), 10g NaCl. L Broth p/c: L broth supplemented with 0.2% maltose and 5mM $MgSO_4$. L Broth (Liquid Lysis): 10g Bacto-tryptone (Difco), 5g yeast extract (Difco), 10g NaCl, 10mM $MgSO_4$. Seakem Agar: L broth containing 12g Seakem agarose (FMC Bioproducts) per litre, and supplemented with 0.3% glucose, 2mM $MgSO_4$, 0.8mM $CaCl_2$, 4μM $FeCl_3$, 10mM Tris/HCl (pH 7.5), 10mg vitamin B_1, after cooling to 46°C. Seakem agarose is a pure form of agarose containing no contaminating enzyme inhibitors. Seakem Top Agar: L broth containing 4g Seakem agarose (FMC

33

Table 2.1 *E. coli* strains used in this thesis.

Strain	Description	Source	Notes
BHB2688	*recA* (λ*imm*⁴³⁴ *cIts b2 red3 Eam4 Sam7*)/λ	N. Murray	1
BHB2690	*recA* (λ*imm*⁴³⁴ *cIts b2 red3 Dam15 Sam7*)/λ	N. Murray	1
DL494	*sbcC201*	D. Leach	2
JC7623	*recB21 recC22 sbcB12 sbcC201*	R. Lloyd	3
N2361	*rec⁺ sbcC⁺*	R. Lloyd	3
N2362	*recB21*	R. Lloyd	3
N2364	*sbcC201*	R. Lloyd	3
N2365	*recB21 sbcC201*	R. Lloyd	3
N2678	*recD1009*	R. Lloyd	3
N2680	*recD1009 sbcC201*	R. Lloyd	3
N2692	*recD1009 recA::Tn10*	R. Lloyd	3

1. Hohn (1979).

2. A derivative of JM83, F⁻ *ara* Δ(*lac-proAB*) *rpsL* φ80 *lacZ* ΔM15 (Yanisch-Perron *et al.*, 1985).

3. A derivative of AB1157, F⁻ Δ(*gpt-proA*)62 *argE3 his-4 leu-6 thr-1 ara-14 galK2 lacY1 xyl-5 thi-1 supE44 rpsL31 tsx33* (Howard-Flanders and Theriot, 1966).

Table 2.2 Bacteriophage λ strains used in this thesis. For full details of λ phage construction, see Methods (2.5) plus the results chapters, 3, 4, 6, and 7.

Phage	Description	Source	Notes
λAD1	ΔB *pal472(StuI) spi6 cI857*	This work	1,2,3
λAD2	ΔB *pal486(StuI) spi6 cI857*	This work	1,2,3
λDRL112	*spi6 cI857*	D. Leach	1
λDRL116	ΔB *pal571 spi6 cI857*	D. Leach	1,2,3
λDRL133	ΔB *pal462(SacI) spi6 cI857*	E. Okely	1,2,3
λDRL134	ΔB *pal476(XbaI) spi6 cI857*	E. Okely	1,2,3
λDRL137	ΔB *pal579 spi6 cI857*	E. Okely	1,2,3
λDRL148	ΔB *pal516(SalI) spi6 cI857*	E. Okely	1,2,3
λDRL152	*spi6 cI857 χC153*	D. Leach	1,4
λDRL167	ΔB *pal462(SacI) spi6 cI857 χC153*	This work	1,2,3,4
λADbke	ΔB *pal480(EcoRI) spi6 cI857 χC153*	This work	1,2,3,4
λAD16T	ΔB *pal480(EcoRI) spi6 cI857 χC153*	This work	1,2,3,4
λAD15C	ΔB *pal480(SphI) spi6 cI857 χC153*	This work	1,2,3,4
λAD16C	ΔB *pal480(AatII) spi6 cI857 χC153*	This work	1,2,3,4
λAD15C16C	ΔB *pal480(NaeI) spi6 cI857 χC153*	This work	1,2,3,4
λAD13G	ΔB *pal480(EcoRI) spi6 cI857 χC153*	This work	1,2,3,4
λAD11T	ΔB *pal480(EcoRI) spi6 cI857 χC153*	This work	1,2,3,4
λAD10G	ΔB *pal480(EcoRI) spi6 cI857 χC153*	This work	1,2,3,4

(cont'd over)

35

Table 2.2 (cont'd)

Phage	Description	Source	Notes
λADbkb	ΔB *pal*480(*Bam*HI) *spi*6 *cI*857 χC153	This work	1,2,3,4
λADbkc	ΔB *pal*480(*Nhe*I) *spi*6 *cI*857 χC153	This work	1,2,3,4
λADbkd	ΔB *pal*480(*Sno*I) *spi*6 *cI*857 χC153	This work	1,2,3,4
λAD-CGCG	ΔB *pal*480(*Xmn*I) *spi*6 *cI*857 χC153	This work	1,2,3,4
λAD-AGCG	ΔB *pal*480(*Xmn*I) *spi*6 *cI*857 χC153	This work	1,2,3,4
λAD-CGAG	ΔB *pal*480(*Xmn*I) *spi*6 *cI*857 χC153	This work	1,2,3,4
λAD-AGAG	ΔB *pal*480(*Xmn*I) *spi*6 *cI*857 χC153	This work	1,2,3,4
λAD-CTTG	ΔB *pal*480(*Xmn*I) *spi*6 *cI*857 χC153	This work	1,2,3,4
λAD-GTTC	ΔB *pal*480(*Xmn*I) *spi*6 *cI*857 χC153	This work	1,2,3,4
λAD-TTTA	ΔB *pal*480(*Xmn*I) *spi*6 *cI*857 χC153	This work	1,2,3,4
λAD-ATTT	ΔB *pal*480(*Xmn*I) *spi*6 *cI*857 χC153	This work	1,2,3,4
λAD-TTTT	ΔB *pal*480(*Xmn*I) *spi*6 *cI*857 χC153	This work	1,2,3,4

1. λ *spi*6 are *red gam*. *cI*857 codes for a temperature sensitive repressor.

2. ΔB is a deletion of the *Eco*RI B fragment of λ.

3. *pal*NNN ( ) indicates that the phage carries a long palindrome. NNN is the length, including any asymmetry. The brackets describe the restriction site at the centre of the palindrome (if any). The palindromic sequences of these phage are described in the Chapters 3, 4, 6, and 7.

4. λ χC153 have a chi site at bp 38481-38488.

Table 2.3 Plasmid vectors used in this thesis. Full details of plasmid constructions are included in the Methods (2.6) and Chapter 4.

Plasmid	Derivation	Source	Notes
pMS2B	pUC18	M. Shaw	1
pADbke	λADbke/pMS2B	This work	2
pAD16T	λAD16T/pMS2B	This work	2
pAD15C	λAD15C/pMS2B	This work	2
pAD16C	λAD16C/pMS2B	This work	2
pAD15C16C	λAD15C16C/pMS2B	This work	2
pAD13G	λAD13G/pMS2B	This work	2
pAD11T	λAD11T/pMS2B	This work	2
pAD10G	λAD10G/pMS2B	This work	2

1. Derivative of pUC18 which stabilises long DNA palindromes (M. Shaw, C. Blake, and D. R. F. Leach, unpublished observations).

2. Derivative of pMS2B containing the central 32bp *Taq*I fragment from the palindrome centre of the cited λ strain.

Bioproducts) per litre.

Phage Buffer: 3g KH_2PO_4, 7g Na_2HPO_4, 5g NaCl, 1mM $MgSO_4$, 1mM $CaCl_2$, 1ml gelatin (1% w/v). TM Buffer: 10mM Tris/HCl (pH 7.5), 10mM $MgSO_4$.

Media Additives

0.5M $CaCl_2$ and 1M $MgSO_4$ Stocks, made up in distilled water, autoclaved. 0.01M $FeCl_3$, 20% Glucose, and 20% Maltose Stocks, made up in distilled water, filter sterilised.

Ampicillin (100mg ml^{-1}; Beecham Pharmaceuticals) was stored at −20°C and used at 100µg ml^{-1}. Xgal (5-bromo-4-chloro-3-indolyl-β-D-galactoside; 25mg ml^{-1}; Northumbria Biologicals) was made up in dimethylformamide and stored at −20°C. It was used at 40µg ml^{-1}. Vitamin B_1 (5mg ml^{-1}; Sigma Chemical Company) was made up in distilled water, filter sterilised and stored at 4°C.

2.3 Materials for DNA Purification and Manipulation

General Solutions and Materials for DNA Purification

Unless otherwise stated, general laboratory chemicals were purchased from Sigma Chemical Company, Fisons or BDH.

1M Tris/HCl Stock (pH 6.5 to 8.0): 1M Tris base, pH adjusted with concentrated HCl, autoclaved. 0.5M EDTA Stock (pH 8.0): 0.5M EDTA disodium salt, adjusted to pH 8 with glacial acetic acid, autoclaved. TE Buffer Stock (10×): 100mM Tris, 10mM EDTA, adjusted to pH 7.5 with concentrated HCl, autoclaved.

Solutions for Phenol, Phenol-Chloroform and Chloroform Extraction

Phenol: Distilled, liquefied (88%) phenol (Rathburn Chemicals) was stored in 25ml aliquots at −20°C, and protected from light in foil-wrapped 25ml polypropylene tubes. After thawing, 0.1% (w/v) 8-hydroxyquinoline (Sigma Chemical Company) was added and the phenol equilibrated using TE buffer. Initially, this was performed

by emulsifying the phenol with 10× TE buffer (100mM Tris/HCl, pH 7.5, 10mM EDTA) for 5 minutes, followed by centrifugation at 4.5krpm in an MSE Centaur-2 bench centrifuge. The aqueous layer was removed, and equilibration repeated another two times using 1× TE buffer. The aqueous layer was removed and

replaced with 2ml 1× TE buffer, 0.2% (v/v) β-mercaptoethanol (Sigma Chemical Company). The phenol was protected from light, stored at 4°C in a 25ml polypropylene tube, and used within one month. Phenol-Chloroform: As above, except after final equilibration the phenolic phase was combined with an equal volume of 24:1 chloroform-isoamylalcohol, the aqueous layer removed and replaced with 2ml 1× TE buffer, 0.2% (v/v) β-mercaptoethanol. The phenol-chloroform was protected from light, stored at 4°C and used within one month. 24:1 Chloroform-isoamylalcohol: 24 volumes of chloroform were combined with 1 volume of isoamylalcohol and protected from light in a brown glass bottle.

Solutions for Ethanol and Isopropanol Precipitation

3M Sodium Acetate (pH 5.3) was prepared by adding 0.19 volumes of sterile 3M acetic acid to 0.81 volumes of sterile 3M sodium acetate, autoclaved. 70% Ethanol: Prepared by adding 0.3 volumes of sterile distilled water to 0.7 volumes of ethanol.

Preparation of Dialysis Tubing

Dialysis tubing was cut into lengths of 20cm and boiled for 10 minutes in 2l of 2% sodium bicarbonate and 1mM EDTA. It was then rinsed thoroughly in distilled water and boiled for 10 minutes in distilled water. After cooling, the tubing was stored at 4°C, submerged in 1mM EDTA 50% (v/v) ethanol. Before use, the tubing was rinsed inside and out with sterile distilled water.

Solutions for Purification of Bacteriophage λ Particles

DNase I (10mg ml[-1]; Sigma Chemical Company) was made up in sterile 50mM Tris/HCl (pH7.5), 10mM NaCl, 1mM DTT, 100μg ml[-1] BSA, 50% (w/v) glycerol and stored at −20°C. RNase A (10mg ml[-1]; Sigma Chemical Company) was made up in sterile 10mM Tris/HCl (pH7.5), 15mM NaCl, then placed in a boiling

water bath for 10 minutes and allowed to cool slowly to room temperature. This treatment denatures contaminating DNases but not RNase A. RNase A was stored at −20°C. Pronase (20mg ml^{-1}; serine protease from *Streptomyces griseus*; Sigma Chemical Company) was prepared in 10mM Tris/HCl (pH 7.5), 10mM NaCl and autodigested for 1 hour at 37°C. Stored at −20°C. Pronase Dialysis Buffer: 20mM Tris/HCl (pH 7.5), 100mM NaCl, 1mM EDTA, 0.002% Triton X-100.

30% PEG 6000, 3M NaCl: PEG 6000 (Fluka Chemie) was made up in sterile 3M NaCl and stored at 4°C. CsCl Solutions (Boehringer Mannheim) were made up in sterile phage buffer to concentrations of 31.2%, 45.4% and 56.2% (w/w).

Solutions for *in Vitro* Packaging of λ DNA

Buffer A: 20mM Tris/HCl (pH 7.5), 3mM MgCl$_2$, 0.05% (v/v) β-mercaptoethanol, 1mM EDTA, stored at −70°C. Buffer M1: 6mM Tris/HCl (pH 7.5), 30mM spermidine, 60mM putrescine, 18mM MgCl$_2$, 15mM ATP, 0.2% (v/v) β-mercaptoethanol, stored at −70°C.

Solutions for Transformation of *Escherichia coli*

0.1M CaCl$_2$ was made up in distilled water, autoclaved.

Solutions for Plasmid DNA Purification

Small Scale Method and Large Scale Method (for subsequent *in vitro* extrusion): Lysozyme: 10mg ml^{-1} in 50mM Tris/HCl (pH 8.0), freshly prepared. 10% Triton X-100 (in sterile distilled water). Large Scale Method only (*In Vitro* Extrusion): Butan-1-ol saturated with a equal volume of 1× TF. TF Buffer Stock (10×): 100mM Tris, 1mM EDTA, adjusted to pH 7.5 with concentrated HCl, autoclaved.

Large Scale Method (General): Plasmid DNA preparation was performed using a QIAGEN plasmid midi kit. The solutions used were those described by the manufacturer.

Solutions for *in Vitro* Extrusion

Standard extrusion buffer: 10mM Tris/HCl (pH 7.5), 0.1mM EDTA, NaCl (35mM to 75mM), and 29μg ml^{-1} plasmid DNA. "Physiological" extrusion buffer (buffer P): 10mM Tris/HCl (pH 7.5), 0.1mM EDTA, 150mM potassium glutamate, 4mM magnesium acetate, and 29μg ml^{-1} plasmid DNA. T4 endonuclease VII reaction buffer: 0.5M Tris/HCl (pH 8.0), 50mM NaCl, 10mM MgCl$_2$, 1mM dithiothreitol (DTT), 5mM spermidine, 250μg ml^{-1} bovine serum albumen (BSA), and T4 endoVII at a concentration of 1U μl^{-1} (Picksley *et al.*, 1990). T4 endoVII was a gift from B. Kemper.

Solutions for UV Melting Analysis

Buffer UV1: 0.5mM Tris/HCl (pH7.0), 0.1mM EDTA, 10mM NaCl. UV2: 2mM Tris/HCl (pH 7.5), 0.5mM NaClO$_4$, 0.1mM EDTA. UV3: As UV2 except with the addition of 20% formamide.

Enzymes and Buffers for DNA Manipulation

Restriction Endonucleases: All restriction endonucleases used in this work and their incubation buffers were purchased from Northumbria Biologicals, New England Biolabs, or Boehringer Mannheim, and used according to the manufacturers instructions.

Solutions for DNA Sequencing

With the exception of the sequencing primer, the solutions used were those provided by the manufacturer.

Other Enzymes

Incubation buffers were generally supplied by the manufacturer. Bacteriophage T4 DNA Ligase (New England Biolabs) was incubated in 50mM Tris/HCl (pH 7.5), 10mM MgCl$_2$, 10mM DTT, 1mM ATP, 25μg ml^{-1} BSA. Klenow Enzyme (Boehringer Mannheim) was incubated in restriction endonuclease buffer, as it was generally used subsequent to a restriction digest. Exonuclease III (Boehringer Mannheim) was used

41

in 50mM Tris/HCl (pH 8.0), 5mM MgCl$_2$, 1mM DTT. All solutions were stored at –20°C.

Other Solutions

dNTP Stocks (Sigma Chemical Company) were prepared in sterile distilled water at a concentration of 50mM and 2mM. Stored at –20°C. BSA (20mg ml^{-1}; Boehringer Mannheim) was stored at –20°C.

Solutions for Gel Electrophoresis

Agarose Gel Electrophoresis. TAE Gel Buffer Stock (20×): 0.8M Tris/acetate, 20mM EDTA (pH 8.0). TAE Gel-Loading Sample Buffer Stock (5×): 0.2M Tris/acetate, 0.25M EDTA (pH 8.0), 0.2% bromophenol blue, 15% Ficoll 400. Ethidium Bromide (10mg ml^{-1}): Prepared in 1× TE buffer, protected from light at 4°C. Used at 0.5μg ml^{-1}.

Polyacrylamide Gel Electrophoresis. TBE Gel Buffer Stock (10×): 0.89M Tris/borate, 20mM EDTA (pH 8.0). Formamide-EDTA Gel-Loading Sample Buffer: 98% formamide, 0.1% bromophenol blue, 0.1% xylene cyanol FF, 10mM EDTA (pH 8.0). 40% Acrylamide Stock: 38% (w/v) acrylamide, 2% (w/v) N-N-methylene bis-acrylamide (Northumbria Biologicals), protected from light at 4°C. 0.5× TBE, 6% Acrylamide Solution for DNA Sequencing Gels: 59.5 g urea, 7ml 10× TBE buffer, 21ml 40% acrylamide stock, distilled water to 140ml total. Stored at 4°C in foil-wrapped bottle. 5× TBE, 6% Acrylamide Solution for DNA Sequencing Gels: 12.75g urea, 15ml 10× TBE buffer, 4.5ml 40% acrylamide stock, 6g sucrose, and a few grains of bromophenol blue. Stored at 4°C in foil-wrapped bottle. TEMED (N-N-N'-N'-tetra-methyl-1,2-diamino-ethane; Sigma Chemical Company) was stored at 4°C and protected from light. 10% AMPS (ammonium persulphate) was freshly prepared in distilled water. Gel Fix: 10% (v/v) glacial acetic acid, 10% (v/v) methanol.

Methods

2.4 Bacterial Methods

Storage of Bacteria

For short term storage, bacteria were kept at 4°C on L agar plates sealed with parafilm. Permanent stocks were prepared by adding 5 drops of sterile 100% glycerol to 1ml of a stationary phase culture in an Eppendorf. This was sealed with parafilm and stored at −70°C.

Growth of Bacteria

Temporary stocks were generated by streaking permanent stocks to single colonies on L agar plates, which were incubated at 37°C overnight unless otherwise indicated. Overnight cultures were grown by inoculating a single colony into L broth and shaking at 37°C, unless otherwise indicated.

Preparation of Plating Cultures

Plating cultures of the appropriate bacterial strain were made by diluting an overnight culture 10-fold in L broth p/c. This was grown shaking at 37°C for 2-2½ hours, depending upon the viability of the strain, and diluted with an equal volume of TM buffer. Plating cultures were stored at 4°C for up to 3 days.

2.5 Bacteriophage λ Methods

Titration of Bacteriophage λ Stocks

Accurate Titration: The phage stock was diluted serially in L broth, and 100µl of each appropriate dilution was added to 250µl of the plating culture and left to adsorb for 20 minutes at room temperature. 2.0ml of molten BBL top agar at 46°C were then added and the mixture was poured immediately onto the surface of a fresh,

dried BBL plate. The plaques were counted after incubating the plates overnight at 37°C.

Spot Tests: If a rough estimate of phage titre was required, 250μl of the plating culture was briefly mixed with 2.0ml of molten BBL top agar (at 46°C) and poured onto the surface of a fresh, dried BBL plate. After this had set, 2μl or 10μl aliquots of the phage dilutions were spotted onto the bacterial lawn and allowed to dry before incubating the plates overnight at 37°C. This technique is particularly useful for comparing the plating efficiencies of different phage in the cloning procedure (see below), as up to 16 (10μl) and 52 (2μl) spots may be compared on a single plate.

λ Cloning Procedure

Annealed palindromic oligonucleotides were inserted into the unique, central *Sac*I site of a 462bp palindrome in λ (either DRL133, Chapter 3 or DRL167, Chapters 4, 6, and 7). To aid the identification of new clones, the inserts were designed so that the cloning procedure disrupted the *Sac*I site, and for positive identification another restriction site was introduced into the centre. The oligonucleotides were not phosphorylated so as to avoid insertion of multimers. Potential new λ phage were identified by their ability to plate on a lawn of JC7623 *recBC sbcBC*, but not on N2361 *rec*⁺. The palindrome sequence was confirmed by comparing its length and restriction site at the centre against the parent λ on a polyacrylamide gel. The palindrome centres of λAD1, λAD2, λADbke, λAD16T, λAD15C, λAD16C, λAD15C16C, λAD13G, λAD11T, and λAD10G were sequenced after cloning into pMS2B.

Plating Behaviour Assay

Aliquots of 250μl of plating cells were incubated for 15 minutes at room temperature with an appropriate dilution of phage. 2ml of BBL top agar was added and the mixture poured onto BBL plates. The relative efficiency of plating (eop) on a given strain, x, was calculated (from the number of plaques on different hosts) as:

$$\frac{(pal \text{ phage on strain x})/(pal \text{ phage on JC7623 } recBC \ sbcBC)}{(\text{DRL112 on strain x})/(\text{DRL112 on JC7623 } recBC \ sbcBC)}$$

44

DRL112 is an isogenic palindrome free phage used to normalize the plating efficiencies on various hosts for minor differences not attributable to the palindromes.

Plaque Area Assay

Plaque area (under the standardised conditions used here) is assumed to be an indirect product of the λ burst size, and thus a measure of the viability of the palindromic phage.

Plate Pouring: CA agar (Plaque Area) and CAS agar (Plaque Area) were freshly prepared using the exact amounts set out in the Materials section. Plates were poured at 46°C using exactly 50ml of agar and allowed to set in large stacks (~20 plates). The plates were used after three days, although the terminal two plates of each stack were discarded. These precautions were necessary because small variations between plates, which are primarily due to drying conditions, can markedly affect the size of plaques on a lawn.

Phage Plating: A single plaque of the appropriate phage was picked using a sterile Pasteur pipette into 1ml L broth. The mixture was vortexed shortly with 10μl chloroform (to lyse remaining host cells) and left for one hour at 4°C to allow diffusion of phage from the agar plug. It was then diluted serially and titrated accurately on N2364 *sbcC* plating cells. The volume of the phage suspension corresponding to 200-300 PFU (plaque forming units) was calculated from the titre, added to 250μl of a fresh N2364 *sbcC* plating culture and left to adsorb for 20 minutes at 37°C. These conditions were chosen to maximise preadsorption of the phage. Exactly 2.0ml of molten CA or CAS top agar at 46°C were then added and mixed. This was poured immediately onto the surface of a CA or CAS plate, and incubated at 37°C overnight.

Plaque Area Quantification: Each plate was checked visually, then the area of the plaques on the plate was measured using a Quantimet 970 digital image analyser (Cambridge Instruments). The images of the plaques were presented to the image analyser via a Chalnicon video camera. Optimum illumination was provided by placing a lightbox (20cm × 20cm) at 45cm beneath the plate. The intensity of the illumination was automatically adjusted by the image analyser such that the peak brightness of the plaque images was equivalent to a video signal of 1V. Field width

was set at approximately 20mm, and the image was digitised into 896 by 704 pixels with 256 grey levels per pixel. The analysis of the image was performed by a Quips routine "PLAQUE" written by Dr C. E. Jeffree (Science Faculty EM Facility, University of Edinburgh). The plaques were first detected by setting appropriate grey level threshold values. Following detection, single-pixel noise was removed by means of an erode procedure followed by a dilate procedure to restore surviving objects to their original size. An editing loop was provided to enable deletion of spurious objects, or the manual separation of overlapping plaques. Measurements of the plaques were obtained based on the detected region.

Plaque Analysis: A graph of cumulative frequency plotted against plaque area was plotted to analyse the results. The median was taken as an indicator of the viability of the λ. The median was the most reproducible parameter since it minimises the effects of outliers. Outliers were relatively common in this analysis for two reasons: λ phage that are initially unadsorbed produce pinprick plaques on the cell lawn, while revertants (phage that have lost the palindrome) produce very large plaques.

Bacteriophage λ Cross

A fresh overnight culture of JC7623 *recBC sbcBC* was diluted 100-fold in L broth p/c and grown to OD_{650} = 0.45 (~2 × 10⁸ cells ml⁻¹). The culture was centrifuged at 4.5krpm for 5 minutes in an MSE Centaur-2 bench centrifuge, resuspended in a equal volume TM buffer and starved shaking at 37°C for 30 minutes. 250µl of the starved cell culture was mixed with 250µl of a phage suspension containing DRL133 *spi6 cI857 pal462* and DRL152 *spi6 cI857 χC153* in a 2:1 ratio and incubated at 37°C for 30 minutes. The phage were added at a combined multiplicity of infection (MOI) of 10. The cell culture with the adsorbed phage was then diluted 100-fold with pre-warmed L broth and grown shaking at 37°C for 1½ hours. Chloroform was added to a final concentration of 0.2% and shaking was continued for a further 5 minutes to complete lysis. Cell debris was pelleted by centrifugation at 4.5krpm for 5 minutes (MSE Centaur-2 bench centrifuge). The supernatant was stored at 4°C and titrated on permissive and restrictive indicator strains. The recombinant phage, DRL167 *spi6 cI857 χC153 pal462* was identified by

its comparatively large plaque size on N2364 *sbcC*, and inability to plate on N2361 *rec⁺*. It was further checked by restriction analysis.

Preparation of Bacteriophage λ Stocks by Plate Lysates

The required phage were titrated and single plaques were picked using a sterile Pasteur pipette into 1ml phage buffer. The mixture was vortexed shortly with 10μl chloroform and left for one hour at 4°C to allow diffusion of phage from the agar plug. 0.1ml, and 0.5ml volumes of the phage suspension were added to 250μl of a fresh plating culture and left to adsorb for 20 minutes at room temperature. 2.0ml of molten Seakem top agar at 46°C was added. This was poured immediately onto the surface of a very fresh (< 1 hour old) Seakem agar plate and incubated at 37°C for 6-8 hours until the lawn reached confluent lysis. 4ml of TM buffer was poured onto the lawn and the plate agitated gently overnight to allow diffusion of the phage from the agar. The TM buffer was removed and vortexed with 50μl chloroform, then centrifuged at 4.5krpm for 10 minutes in an MSE Centaur-2 bench centrifuge to sediment agar and cell debris. The supernatant was titrated, and stored at 4°C.

Preparation of Bacteriophage λ Stocks by Liquid Lysates

This method was used when large quantities of phage were required, such as in the preparation of high quality λ DNA for subcloning and sequence analysis. A fresh overnight culture was diluted 50-fold in pre-warmed L broth (liquid lysis), and grown shaking at 37°C. The flask was no more than one-eighth full, as efficient λ lysis is dependent upon good aeration. When the culture had reached $OD_{650} = 0.5$ ($\sim2.5 \times 10^8$ cells ml^{-1}) after $\sim1\frac{1}{4}$ hours, phage were added at a multiplicity of infection (MOI) of 0.1. The flasks were returned to 37°C with vigorous shaking and the OD_{650} was monitored periodically. After 2-4 hours of growth, cell lysis occurs and the OD_{650} begins to fall (from ~2.5). However, it can take another 2-4 hours of growth to reach the minimum OD_{650} ($0.4 \sim 0.7$). When two successive OD_{650} readings were similar, indicating that the minimum had been reached, 1ml of chloroform per 500ml of lysate was added and shaking was continued for a further 10 minutes to complete lysis. The lysate was then centrifuged at 10krpm for 15 minutes at 4°C

(Sorvall centrifuge, GSA rotor) to remove cell debris. The supernatant was stored at 4°C.

Purification of Bacteriophage λ Particles from Plate Lysates

This method was used to purify λ particles from plate lysates in order to extract λ DNA by the small scale method. 1ml of a plate lysate was transferred to an Eppendorf. 5μl DNase and 5μl RNase A (both 10mg ml^{-1}) were added and incubated at 37°C for 30 minutes to digest bacterial DNA and RNA. The phage suspension was centrifuged for 1 minute at 15krpm (Sorvall Microspin 24 centrifuge) and the supernatant added to 0.5ml ice-cold 30% PEG 6000, 3M NaCl in a pre-cooled Eppendorf. The Eppendorf was inverted five times to mix the suspensions and incubated on ice for 1 hour. It was then centrifuged at 15krpm for 15 minutes at 4°C (Sorvall Microspin 24 centrifuge) and the supernatant was removed carefully. The Eppendorf was centrifuged again for 30 seconds and any remaining supernatant removed. The pellet was gently resuspended in 200μl TM buffer, 200μl of chloroform was added and the Eppendorf vortexed very briefly three times. The chloroform causes PEG 6000 to precipitate, leaving the λ particles in aqueous suspension. The phases were separated by centrifugation at 15krpm for two minutes (Sorvall Microspin 24 centrifuge) and the aqueous (upper) layer was transferred to a fresh Eppendorf and stored at 4°C.

Purification of Bacteriophage λ Particles from Liquid Lysates

Precipitation of λ Particles with PEG 6000: DNase and RNase were added to the liquid lysate to a final concentration of 1μg ml^{-1} and incubated at room temperature for 30 minutes, swirling occasionally. NaCl was then added to the liquid lysate to a final concentration of 1M, and PEG 6000 to a final concentration of 10% (w/v). Both were dissolved by stirring gently. The suspension was incubated on ice for four hours or overnight and then centrifuged at 10krpm for 15 minutes at 4°C (Sorvall centrifuge, GSA rotor). The supernatant was discarded and the centrifuge bottle allowed to drain for 10 minutes. The pellet was resuspended at room temperature in 5ml phage buffer by gentle shaking for one hour. The suspension was then transferred to a glass bottle and a further 2ml phage buffer used to rinse the

inside of the centrifuge bottle. An equal volume of chloroform was added and the suspension vortexed gently for 10 seconds. The phases were separated and the PEG 6000 pelleted by centrifugation at 4.5krpm for 10 minutes (MSE Centaur-2 bench centrifuge). The aqueous layer was transferred to a fresh bottle and stored at 4°C.

Purification of λ Particles by CsCl Step Gradient Centrifugation: A CsCl step gradient was made by successively underlayering 2.5ml of 31.2%, 2.0ml of 45.4% and 1.5ml of 56.2% CsCl solutions in a 13ml Beckman Ultra-Clear™ tube using a Pasteur pipette. The phage suspension was then carefully loaded on top of the CsCl solution layers to within 2mm of the top of the tube. The step gradients were balanced to within ±10mg and centrifuged at 35krpm for 35 minutes at 20°C (Sorvall ultracentrifuge, TH641 titanium swing-out rotor). Two bands could be seen in visible light: the lower was phage particles and the upper was debris. The phage band (~1ml) was collected through the side of the tube using a hypodermic needle (19G1½ 1.1 × 40mm) and syringe. The band from the step gradient was dialysed against 2l of phage buffer for 2 hours at 4°C. The dialysed suspension was then loaded onto a second CsCl step gradient, which had been prepared as before. In order to minimise phage loss, the inside of the dialysis tubing was rinsed with 1ml phage buffer; this was also loaded onto the CsCl step gradient. The tubes were balanced and centrifuged as before. Normally, only one band was visible after the second round of centrifugation. This was collected as before using a hypodermic needle and syringe, and stored at 4°C.

In Vitro Packaging of Bacteriophage λ DNA

Sonicated Extract: Lysogenic *E. coli* BHB2690 was grown shaking in 500ml L broth at 30°C to OD_{650}= 0.3. Lysogens were induced by transferring the culture to a 43°C shaking waterbath for 15 minutes, followed by vigorous shaking at 37°C for 1 hour. After cooling on ice/water for 10 minutes, the cell suspension was centrifuged at 6krpm for 6 minutes at 4°C (Sorvall centrifuge, GSA rotor). Cell pellets were resuspended in 1/500 volume buffer A, transferred to a single Nalgene polypropylene centrifuge tube and diluted with 2.6ml buffer A. The suspension was sonicated in ~3 second bursts (to prevent foaming) until no longer viscous and centrifuged at 6krpm for 6 minutes at 4°C (Sorvall centrifuge, SS34 rotor). 50µl of supernatant was

aliquoted to pre-cooled Eppendorf tubes, frozen in a dry ice/methanol bath, and stored at −70°C.

Freeze-Thaw Lysate: Lysogenic *E. coli* BHB2688 was grown shaking in 500ml L broth at 30°C to $OD_{650} = 0.3$. Lysogens were induced by transferring the culture to a 43°C shaking waterbath for 15 minutes, followed by vigorous shaking at 37°C for 1 hour. After cooling on ice/water for 10 minutes, the cell suspension was centrifuged at 6krpm for 6 minutes at 4°C (Sorvall centrifuge, GSA rotor). Cell pellets were resuspended in 1/100 volume cold 10% sucrose, 50mM Tris/HCl (pH 7.5), pooled and dispensed into a Nalgene ultracentrifuge tube. 150μl freshly prepared lysozyme (10mg ml^{-1} in 0.25M Tris/HCl, pH 7.5) was added and gently mixed. The suspension was frozen in dry ice/methanol, thawed first at room temperature and then more gradually at 4°C. The suspension was kept on ice, 150μl of buffer M1 was added and gently mixed, and then centrifuged at 40krpm for 1 hour at 4°C (Sorvall ultracentrifuge, Ti50 rotor). 55μl of supernatant was aliquoted to pre-cooled Eppendorf tubes, frozen in dry ice/methanol, and stored at −70°C.

Packaging: Buffer A and buffer M1 were thawed at room temperature, while sonicated extract and freeze-thaw lysate were thawed on ice. The components were mixed in this order: Buffer A (7μl), λ DNA (5μl), Buffer M1 (2μl), Sonicated Extract (10μl), Freeze-Thaw Lysate (10μl). The reaction was incubated at 25°C for 1 hour, then diluted with 500μl of phage buffer. 100μl, and 10μl aliquots of this phage suspension were titrated.

2.6 Plasmid Methods

Maintenance of Plasmids

Plasmids were maintained by supplementing the media with ampicillin.

Transformation

This method is derived from Mandel & Higa (1970).

Preparation of Competent Cells: A fresh overnight culture was diluted 100-fold in L broth and grown shaking at 37°C to an $OD_{650} = 0.4$. The culture was

chilled on ice for 15 minutes, and 20ml were centrifuged at 4.5krpm for 5 minutes at 4°C (MSE Centaur-2 bench centrifuge). The cell pellet was resuspended in 2ml ice-cold 0.1M CaCl$_2$ and incubated on ice for 20 minutes. The suspension was again centrifuged at 4.5krpm for 5 minutes at 4°C (MSE Centaur-2 bench centrifuge), and resuspended gently in 400µl of ice-cold 0.1M CaCl$_2$. The competent cells were used immediately.

Transformation: 50-100ng plasmid DNA (suspended in a minimum volume of 1× TE buffer or distilled water) was added to a 200µl aliquot of competent cells and incubated on ice for 30 minutes. The cells were heat-shocked for 2 minutes in a 42°C waterbath and rapidly cooled on ice. 10µl and 100µl aliquots of the transformed cells were spread on L agar plates containing ampicillin and Xgal and incubated overnight at 37°C.

Cloning of Palindromic Centres from λ into pMS2B

The centre of the palindrome in λAD1, λAD2, λADbke, λAD16T, λAD15C, λAD16C, λAD15C16C, λAD13G, λAD11T, and λAD10G was subcloned into pMS2B and sequenced. To do this, the palindrome was gel purified after restriction with *Eco*RI (or *Ava*I and *Pvu*I for the palindromes with an *Eco*RI centre). A 32bp *Taq*I fragment containing the palindrome centre was ligated into the *Acc*I site of pMS2B (a derivative of pUC18), thereby destroying it. The plasmid DNA was transformed into DL494 (*E. coli sbcC*) and the cells plated on L agar supplemented with Xgal and ampicillin. Plasmids with inserts at the *Acc*I site gave rise to white colonies while those without gave rise to blue colonies. The plasmids with inserts corresponding to the centre of the palindrome were further identified using DNA minipreps and restriction analysis.

In Vitro Cruciform Extrusion

Kinetic studies were performed according to Murchie and Lilley (1987), with some variations. The major difference was that T4 endoVII (a gift from B. Kemper) was used in place of SI nuclease as the cruciform probe. T4 endoVII cleaved at the four-way junction of the DNA cruciform under the conditions used.

Plasmid DNA was isolated by the Large Scale (*In Vitro* Extrusion) protocol, and thawed slowly at 6°C prior to the experiment. Cruciform extrusion experiments were initially carried out in 10mM Tris-HCl (pH7.5), 0.1mM EDTA buffer containing a known concentration of NaCl. To determine the extrusion salt optimum the plasmid DNA was incubated in the buffer for a given time period with NaCl added to a concentration of between 35mM and 75mM. Subsequently, *in vitro* cruciform extrusion experiments were performed in a more physiological buffer (buffer P). After incubation, the sample was rapidly transferred to ice to terminate the reaction. The reaction constant for each extrusion was ascertained by incubating the plasmid DNA in the appropriate buffer, and then withdrawing aliquots onto ice over a given time period. All extrusions took place at 37°C (±0.01) in a Grant LTD6 waterbath. The samples were then diluted into T4 endoVII reaction buffer and incubated with 40U T4 endoVII at 16°C for 40 minutes. No cruciform extrusion was detectable in this buffer at 16°C. After this, the DNA samples were ethanol precipitated, resuspended in a suitable restriction buffer, and digested to completion with *Sca*I.

The DNA samples were loaded onto a 1.5% agarose gel vertical electrophoresis unit and run for ~2 hours at 40V. After staining the gel in ethidium bromide, the proportion of fragments arising from the T4 endoVII specific cleavage was determined by assaying the relative intensity of bands on a gel negative with a Shimadzu CS-930 dual- wavelength Chromato Scanner (densitometer) and DR-2 data recorder. The wavelength used was 550nm. Where necessary, several exposures of the film were taken to ensure that the response was approximately linear. A graph of 1+ ln(unextruded fraction) against time was plotted using linear regression analysis. The reaction constant (k) was calculated from the gradient of the graph, and the reaction half time as ln2/k.

2.7 UV Melting Analysis

Oligonucleotides were synthesised and purified following standard methods by Oswel DNA service (T. Brown, Edinburgh). Absorbance versus temperature melting curves were measured at 260nm on a Perkin-Elmer Lambda 15 ultraviolet

spectrophotometer equipped with a Peltier block and controlled by an IBM PS2 microcomputer. A heating rate of $0.9°C$ min^{-1} was used. Data were collected and processed using the PECSS2 software package (Perkin-Elmer Ltd.) Absorbance versus temperature data were converted to the first derivative, and the melting temperature was defined as the temperature at which the curve attained its maximum point. UV melting took place in buffer UV1, UV2, or UV3. This work was assisted by T. Brown (Edinburgh; Ebel *et al.*, 1992).

2.8 General DNA Purification and Manipulation Methods

Phenol, Phenol-Chloroform and Chloroform Extraction

Successive phenol-chloroform and chloroform extractions were generally sufficient to purify DNA from contaminating protein. An initial phenol extraction was used where the purity of DNA was paramount, or where the purpose of the phenol was to disrupt λ phage heads (see large scale method for purification of λ DNA).

Phenol Extraction was carried out in Eppendorf tubes. An equal volume of equilibrated phenol was added to the DNA solution and the two were homogenised by vortexing briefly. The phases were separated by centrifugation at 15krpm for 5 minutes (Sorvall Microspin 24 centrifuge). The aqueous (upper) layer was transferred to a fresh Eppendorf. Phenol-Chloroform Extraction was carried out identically to phenol extraction, except that equilibrated 25:24:1 phenol-chloroform-isoamylalcohol was used in place of phenol. Chloroform Extraction was carried out identically to phenol extraction, except that 24:1 chloroform-isoamylalcohol was used in place of phenol, and the phases were separated by centrifugation at 15krpm for only one minute.

Ethanol and Isopropanol Precipitation

Ethanol precipitation was normally used to concentrate and purify DNA. Isopropanol precipitation was used where a large volume was impractical, or where it was desirable to avoid the coprecipitation of RNA.

Ethanol Precipitation was carried out in 15ml Corex tubes or Eppendorf tubes. 0.1 volumes of 3M sodium acetate pH 5.3 were added to 0.9 volumes of DNA solution, followed by 2 volumes freezer-cold (−20°C) ethanol. The tube contents were mixed by inversion and incubated at on ice for at least 20 minutes to allow precipitation. The DNA was then pelleted by centrifugation at 15krpm for 20 minutes at 4°C (Sorvall centrifuge, SS34 rotor, or Sorvall Microspin 24 centrifuge). The ethanol was discarded, the tube filled with freezer-cold 70% ethanol, and recentrifuged at 15krpm for 10 minutes at 4°C. The supernatant was discarded, the pellet was vacuum dried for 5 minutes and redissolved in 1× TE buffer or distilled water.

Isopropanol Precipitation was carried out identically to ethanol precipitation, except that 0.7 volumes of isopropanol (at room temperature) were used in place of 2 volumes freezer-cold ethanol, and centrifugation was carried out at room temperature. The pellet was rinsed with freezer-cold 70% ethanol and recentrifuged as described above.

Purification of DNA from Agarose Gels and Solutions

Purification of DNA from agarose gels and solutions was performed using GENECLEAN® and MERmaid™ kits (BIO 101) according to the manufacturer's instructions. The MERmaid™ procedure was used to purify small DNA fragments from aqueous solutions. Both kits are based on the method of Vogelstein & Gillespie (1979).

2.9 Bacteriophage λ DNA Purification

Small Scale Method

This method was used to purify λ DNA from plate lysates. The λ particles were purified from plate lysate phage stocks as described earlier (see: Purification of Bacteriophage λ Particles from Plate Lysates). The λ suspension was extracted with 200μl of phenol to disrupt the phage heads, and then extracted with 200μl of phenol-chloroform and chloroform to purify the λ DNA from denatured capsid

proteins. The DNA was precipitated with ethanol, rinsed with 70% ethanol and resuspended in 10-40µl sterile distilled water.

Large Scale Method

This method was used to purify λ DNA from liquid lysates. The λ particles were purified from liquid lysate phage stocks as described earlier (see: Purification of Bacteriophage λ Particles from Liquid Lysates). The phage band from the second round of CsCl step gradient centrifugation was dialysed overnight against $2l$ of $1\times$ TE buffer at 4°C. Pronase was added to the phage band (in the dialysis tubing) to a final concentration of 1mg ml^{-1}. This was then dialysed for 2 hours against $1l$ of pronase dialysis buffer at 37°C. The phage DNA solution was transferred to Eppendorf tubes and the inside of the dialysis tubing was rinsed with 500µl $1\times$ TE buffer to minimise loss of phage DNA. The pooled DNA solutions were extracted with an equal volume of phenol to disrupt remaining phage heads, and then extracted with equal volumes of phenol-chloroform and chloroform to purify the λ DNA from denatured protein. The DNA was precipitated with ethanol, rinsed with 70% ethanol and resuspended in 100µl sterile distilled water and stored at −20°C. The DNA concentration was estimated from the absorbance at 260 nm as determined by a Perkin-Elmer Lambda 15 spectrophotometer.

2.10 Plasmid DNA Purification

Small Scale Method

1.5ml of a fresh overnight culture were transferred to an Eppendorf tube and centrifuged for 30 seconds at 15krpm (Sorvall Microspin 24 centrifuge). The supernatant was discarded and the tube centrifuged again for 5 seconds at 15krpm. The remaining liquid was removed and the pellet was resuspended in 70µl ice-cold 50mM Tris/HCl (pH 8.0). The cell suspension was incubated on ice for 5 minutes after addition of 6µl lysozyme solution. 1µl RNase solution and 5µl EDTA (pH 8.0) were added. After 5 further minutes incubation on ice, 1.5µl 10% Triton was mixed gently into the suspension, and the cell debris was removed by centrifugation at

15krpm for 5 minutes at 4°C (Sorvall Microspin 24 centrifuge). Approximately 70μl of supernatant containing the plasmid DNA were transferred to a fresh Eppendorf tube and extracted with equal volumes of phenol-chloroform and chloroform. The sample was ethanol precipitated, washed with 70% ethanol and resuspended in distilled water, and stored at -20°C.

Large Scale Method

The large scale purification of plasmid DNA was performed using a QIAGEN plasmid midi kit, as described by the manufacturer. For pMS2B and its derivatives, 40ml fresh overnight culture was generally sufficient.

Large Scale Method (*In Vitro* Extrusion)

Supercoiled, cruciform free plasmid DNA was isolated from an overnight half litre culture, following a scaled up procedure of the Small Scale Method up to the triton lysis step. It was further purified through two rounds of CsCl-ethidium bromide isopycnic centrifugation. The ethidium bromide was removed with seven or more rounds of butan-1-ol extraction on ice, followed by extensive dialysis in 1× TF at 6°C. The plasmid DNA was stored at −70°C.

2.11 General DNA Manipulations

Annealing of Oligonucleotides

As the majority of oligonucleotides used in this work were palindromic, special precautions were taken to promote the formation of double-stranded DNA over self-annealed hairpins. Oligonucleotides (Oswel DNA Service) were annealed in 1× TE buffer that had been made 10mM NaCl. This was prepared by adding 40μl of an aqueous solution of the oligonucleotide to 8μl of 10× TE buffer and 32μl of 25mM NaCl in an Eppendorf tube. The mixture was heated to 100°C for 1 minute in a waterbath containing a large volume (> 500ml) of water and the oligonucleotide solution allowed to cool to room temperature gradually. It was then diluted 100-fold in 1× TE buffer, 10mM NaCl and stored at −20°C.

Restriction Digests

Restriction digests were carried out in the appropriate incubation buffer, and used according to the manufacturer's instructions, except that a five-fold excess of restriction endonuclease (5U μg^{-1} DNA) was generally added. Incubations were for two hours at 37°C. Where necessary, the reaction was stopped by phenol-chloroform and chloroform extractions, or heating to 70°C for 10 minutes. The latter method was used in the case of λ DNA restriction digests to denature the cohesive (cos) ends. However, this method was not used where the palindrome had been excised by an EcoRI digest, as the heating may promote hairpin formation and the loss of the palindromic DNA fragment.

DNA Sequencing

Standard double strand DNA sequencing (Sanger *et al.*, 1977) was performed with a Sequenase® v2.0 DNA sequencing kit (United States Biochemical) following a method derived from Manfioletti and Schneider (1988). 5-10μg λ DNA or 2-5μg plasmid DNA were purified from aqueous solution using the GENECLEAN® procedure, and resuspended in 9μl distilled water. 1μl (~0.5μg) sequencing primer (Oswel DNA Service) was added to the DNA in an Eppendorf tube. The DNA was then denatured by incubation in a waterbath at 100°C for 4 minutes, and rapid cooling on dry ice/ethanol. The labelling reaction was prepared in a Eppendorf tube and included: DTT (2μl), [α^{35}S]dATP (Amersham; 0.5μl, 5μCi), diluted labelling mix (1× in distilled water; 0.7μl), Sequenase® buffer (5×; 2μl), diluted Sequenase® enzyme (1/8 in ice-cold enzyme dilution buffer; 2.5μl). The labelling reaction was kept on ice while the template-primer mix was thawed and immediately centrifuged at 15krpm for 10 seconds to remove the DNA solution from the walls of the tube. The template-primer mix was then added to the labelling reaction and incubated at room temperature for 5 minutes. 4μl aliquots of the labelling reaction were then transferred to four Eppendorf tubes, each containing 2μl of either ddG, ddA, ddT or ddC termination mix; the tubes had been pre-warmed at 37°C for 1 minute before the addition of 4μl labelling reaction. The termination reactions were then incubated at 37°C for 3 minutes. Where the template was capable of forming secondary structures (and thereby eliciting band compression) the termination reactions were performed

at 42-45°C. The reactions were halted by adding 4μl stop solution to each tube and placing the sample on ice. The sequencing reactions were stored at −20°C until gel electrophoresis.

The sequencing of long DNA palindromes was a more arduous task (see Chapter 8). Various methods were attempted, including the use of: a TAQUENCE™ DNA sequencing kit (United State Biochemicals; dideoxy DNA sequencing using Taq DNA polymerase); a double strand DNA cycle sequencing system (GIBCO BRL); exonuclease III to digest double strand DNA from the 3' end and prepare single strand DNA template. dITP and 7-deaza-dGTP sequencing were also attempted.

DNA Ligation

DNA ligation was carried out in a total volume of 10-20μl bacteriophage T4 DNA ligase buffer. This was supplied as a 10× stock of which 0.1 volumes were added to the DNA solution. 40 units of bacteriophage T4 DNA ligase (New England Biolabs) were then added and the reaction was incubated for 4 hours or overnight at 15°C. The reaction was terminated by heating to 70°C and the solution was allowed to cool to room temperature slowly to promote the reassociation of double-stranded DNA. DNA fragments were generally ligated with the vector molecule at a 3-5-fold molar ratio of cohesive ends.

Radiolabelling of DNA

Radiolabelling of DNA was generally carried out directly after an *Eco*RI restriction digest. As Klenow enzyme functions adequately in virtually all restriction endonuclease buffers, the incubation buffer was not changed. 0.5μl (5μCi) of [α^{32}P]dATP or [α^{35}S]dATP (Amersham) was added to the restriction digest. This was supplemented by 1μl of 2mM stock solutions of appropriate non-radioactive dNTPs, as determined by the composition of the 5' overhang produced by the restriction endonuclease. 1μl Klenow enzyme (1U) was added and the reaction was incubated at room temperature for 15 minutes. A further 1μl Klenow enzyme was added and the incubation continued for a further 10 minutes at room temperature. 1μl of 50mM stock solutions of dGTP, dATP, dCTP and dTTP were added as a chase (to ensure flush termini were produced) and the incubation continued for a further 5 minutes at

room temperature. The DNA was purified by phenol-chloroform and chloroform extractions and ethanol precipitation, and resuspended in 1× TE buffer.

2.12 Gel Electrophoresis

Agarose Gel Electrophoresis

Agarose gels were made and run in 1× TAE gel buffer. The concentration of agarose was varied according to the size of the relevant DNA fragment. 0.7-0.8% was generally used, and 1.5% for isolation of small (~450bp) fragments. DNA samples were prepared by the addition of 0.2 volumes of TAE gel-loading sample buffer, and loaded into the gel slots. Gels were generally run at ~5 V cm^{-1} (~80 V) and examined by 304nm UV light on a C-62 Blak-Ray transilluminator (Ultraviolet Products Incorporated) after exposure to 0.5μg ml^{-1} ethidium bromide and destaining. Photographs were taken on Polaroid 667 film using a Wratten 25 red filter with an exposure of 1 second and an aperture of f/5.6½. Negatives were produced in a similar manner using Ilford HP5 film. DNA fragments for purification were excised using a sterile scalpel blade.

Polyacrylamide Gel Electrophoresis

Polyacrylamide gels (40cm × 21cm × 0.4 mm) were made and run in TBE gel buffer using a Sequi-Gen® Nucleic Acid Sequencing Cell (BioRad). The apparatus was assembled according to the manufacturer's instructions.

DNA sequencing gels consisted of 0.5-5× TBE buffer gradient 6% polyacrylamide denaturing gels. The glass plates were assembled using 0.4mm spacers and clamps provided by the manufacturer (BioRad). 50μl TEMED and 125μl 10% AMPS were added to 10ml 0.5× TBE 6% acrylamide solution, and the mixture was promptly used to impregnate a 25cm × 5cm strip of blotting paper in a casting tray. The bottom edge of the glass plate sandwich was pushed firmly against the blotting paper and the catalysed acrylamide was allowed to enter the mould by capillary action. After this had set (~2 minutes), thereby sealing the bottom edge, the glass plate sandwich was laid at an angle of ~20° to the horizontal. 25μl TEMED and

70μl 10% AMPS were added to 25ml 0.5× TBE 6% acrylamide for the "top" mix. 6μl TEMED and 16μl 10% AMPS were added to 5ml 5× TBE 6% acrylamide for the "bottom" mix. Polyacrylamide was poured into the gel mould using a 25ml glass pipette. A 16-tooth comb or 24-tooth "sharks-tooth" comb was inserted. The top of the mould was covered in Saran wrap and left overnight at 4°C. The gel mould was then assembled in the apparatus according to the manufacturer's instructions. 350ml 1× TBE buffer were poured into the cathodic buffer chamber, the gel mould was clamped into place, and the anodic buffer chamber was filled with 1× TBE buffer. The comb was removed and the gel was run at a constant power of 40W until a temperature of 50-55°C had been attained. Samples were denatured by boiling in a waterbath for 3 minutes and cooling rapidly on ice. 5μl of the sample were then loaded onto the gel using a micropipetter, and the gel was run at a constant power of 38W. After electrophoresis, the gel mould was dismantled and the glass plate with the polyacrylamide gel placed in 2½l gel fix for 15 minutes. The gel was transferred to wet blotting paper (Ford Goldmedal), covered in Saran wrap, and dried in a BioRad Model 583 Gel Drier for 1 hour at 80°C.

2.13 Autoradiography

^{32}P Isotopes

Autoradiography was carried out at −70°C in Cronex (DuPont) cassettes containing "xtra-life" intensifying screens, using Cronex 4 (DuPont) X-ray film (30cm × 40cm) preflashed to an OD_{540} of 0.1. The exposure time was varied from 4-16 hours, and films were developed in an X-OGRAPH Compact X2 automatic film processor.

^{35}S Isotopes

Autoradiography was carried out at room temperature in Cronex (DuPont) cassettes, using Cronex 4 (DuPont) X-ray film (30cm × 40cm). The exposure time was varied from 1-7 days, and films were developed in an X-OGRAPH Compact X2 automatic film processor.

CHAPTER 3:

The Effect of Central Asymmetry on the Propagation of Long DNA Palindromes

3.1 Introduction

DNA palindromes longer than 100 to 150bp cannot be propagated in wild type *E. coli* and replicons carrying such sequences suffer two fates. They are very poorly replicated leading to the phenotype described as inviability; or the palindromes they contain are partially or completely deleted, a phenotype denoted instability (see introduction to this thesis; Collins, 1981; Lilley, 1981; Leach and Stahl, 1983; Leach and Lindsey, 1986). It is not yet known whether the mechanism of inviability is related to the mechanism of instability. However, both inviability and instability may be due to the formation of an unusual DNA secondary structure *in vivo*, a cruciform. This chapter is concerned with the means by which central palindrome asymmetry affects inviability and the implications for cruciform formation *in vivo*.

Palindrome asymmetry is relevant because it may reveal details of the mechanism for inviability in *E. coli*. The introduction to this thesis reports that an asymmetry to the centre of a palindrome may alleviate the effects of palindrome-mediated inviability and instability. It is possible that large asymmetries could prevent cruciform extrusion and, hence, increase viability. Warren and Green (1985) determined that 57bp is the shortest sequence necessary to alleviate inviability, although lengths greater than 150bp were required before instability was completely overcome. Peeters *et al.* (1988) confirmed that in *Bacillus subtilis* sequences longer than ~150bp do not necessarily have any greater effect.

Despite the above mentioned evidence, the question of whether palindromic DNA sequences can adopt cruciform structures *in vivo* has been difficult to answer conclusively. Recently, sensitive assays and specifically designed A/T-rich palindromes, coupled to the use of osmotic conditions expected to favour extrusion, have permitted the detection of significant amounts of cruciform structures *in vivo* (Panayotatos and Fontaine, 1987; McClellan *et al.*, 1990; Dayn *et al.*, 1991). Also, Zheng *et al.* (1991) and Sinden *et al.* (1991) have shown that low levels of cruciform structures may be detected *in vivo*, where a cruciform with 6bp asymmetry is present at a higher level than a cruciform with 14bp asymmetry. The results described in this chapter support the idea that cruciform structures can form *in vivo* and extend the

evidence to normal conditions and to palindromes of average base composition. This is achieved by comparing the propagation of λ phages carrying long perfect palindromes with that of phages carrying palindromes with small regions of asymmetry on a range of *E. coli* host strains.

3.2 Results

A Set of Palindromes with Different Central Sequences

λ phages carrying palindromes that differ only in their central sequence were constructed. These phages were derived from λDRL116, that has previously been used to study the behaviour of palindrome bearing phage (Chalker *et al.*, 1988). All of the phages carry a similar inverted repeat but have different central sequences that vary in their degree of symmetry. Four of the phages have perfect centres and three have imperfections at or near the centre, 8, 15, and 27bp in length. These palindromes are depicted in Figure 3.1. The construction of λDRL133, λDRL134, λDRL137, and λDRL148 is described in detail by Chalker (1990). λAD1 and λAD2 were constructed as described in the Materials and Methods.

Plating Behaviour of Palindrome Phage

A subset of the above phages (λDRL116, λDRL133, λDRL134, λDRL137 and λDRL148) was initially assessed for plating on strains of *E. coli* known to affect the propagation of DNA palindromes (Chalker, 1990). The plating procedure and relative efficiency of plating (eop) calculation are described in the Materials and Methods. The results of Chalker (1990) are reproduced in Table 3.1. It can be seen that the eop of the different phages varied considerably and that the phages with perfect palindromes showed a lower eop than those with imperfect centres. This poor plating was particularly marked for phages λDRL133 and λDRL134. λDRL148, which also has a perfect centre, gave an eop intermediate between this pair and the phages with imperfect centres. Of the imperfect phages, λDRL137, with a 27bp (off-set) asymmetry, was just able to plate on N2362 *recB* and N2692 *recD recA* (the least permissive hosts that enabled plating), whereas λDRL116, with a 15bp asymmetry

did not. Chalker (1990) defined a hierarchy of plating that was repeated and confirmed for this chapter following the same plating procedure. It is as follows (number of base pairs of asymmetry, if any, is bracketed):

$$\lambda DRL137(27) > \lambda DRL116(15) > \lambda DRL148 > \lambda DRL133, \lambda DRL134.$$

This hierarchy applied across the set of host strains tested which argues that it is independent of host genotype.

Recovery of Supercoiled DNA

To determine whether the plating effect was reflected by an effect at the DNA level, experiments were performed to determine the recovery of supercoiled DNA (Chalker *et al.*, 1988; Chalker, 1990). After a period of 15 minutes postinjection, recovery of $\lambda DRL116$ DNA was close to 100% from an *sbcC* mutant and 30% from a wild-type host, but reduced to 30% and 2%, respectively, for $\lambda DRL134$. The recovery of $\lambda DRL148$ DNA was intermediate between the asymmetric λ phages and $\lambda DRL133/\lambda DRL134$, comparable with its plating behaviour.

Plating Behaviour of a Phage with an 8bp Asymmetry

To extend the observations of Chalker (1990) to a palindrome with a shorter asymmetry, $\lambda AD2$, a phage with an 8bp central insertion, was constructed and its behaviour was compared with that of an isogenic phage with a perfect palindrome ($\lambda AD1$). The control phage with the perfect palindrome behaved similarly to phages $\lambda DRL133$ and $\lambda DRL134$, but $\lambda AD2$ showed a phenotype intermediate between that conferred by the perfect palindromes and the palindrome with a 15bp asymmetry. These results are incorporated into Table 3.1.

64

Figure 3.1

The structure and derivation of the palindromes studied. The restriction sites involved in the construction of the palindromes are indicated as follows: E (*Eco*RI), S (*Sac*I), Sa (*Sal*I), St (*Stu*I), X (*Xba*I). λDRL133, λDRL134, λAD1, and λDRL148 are perfect palindromes, whereas λDRL116, λAD2, and λDRL137 are imperfect. λAD2 has an asymmetric centre of 8bp, λDRL116 has an asymmetric centre of 15bp and λDRL137 has an asymmetry of 27bp which is located 15bp from the centre of the palindrome. The left hand side of the figure depicts the central 70bp of each of the palindromes.

TTATTTACTATCACAAGATTAGTAGATGTCGTTGGCGAAAAGTCATCTACTAATCTTGTGATAGTAAATA

DRL116

  E   S S   E

15

SstI

TATGACAATATCCAGAAATATCTGCTTCGAATGAGCTCATTCGAAGCAGATATTTCTGGATATTGTCATA

DRL133

  E   S   E

XbaI

ATATCCAGAAATATCTGCTTCGAATGAGCTGATCTAGATCAGCTCATTCGAAGCAGATATTTCTGGATAT

DRL134

  E   X   E

StuI

CAATATCCAGAAATATCTGCTTCGAATGAGCTAGGCCTAGCTCATTCGAAGCAGATATTTCTGGATATTG

AD1

  E   St   E

StuI   StuI

CAGAAATATCTGCTTCGAATGAGCTAGGCCTCTAGACGAAGGCCTAGCTCATTCGAAGCAGATATTTCTG

AD2

  E   St St   E

8

SalI   SalI

AGTCGACGTCACGCGTCCATGGAGATCTCGAGGCCTGCAGGCATGCAAGCTTGCATGCCTGCAGGTCGAC

DRL137

  E   SaSa   E

27
15

pMTL 24
polylinker

SalI

CGAATGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACTCTAGAGGATCCCCGGGTACCGAGCTCATTCG

DRL148

  E   Sa   E

Table 3.1

Plating behaviour of palindrome-containing phages on mutant *E. coli* hosts.

E. coli strain and genotype		λ Phage				DRL133 DRL134
		DRL137	DRL116	AD2	DRL148	AD1
N2361	*rec⁺ sbcC⁺*	-	-	-	-	-
N2362	*recB*	±^{a}	-	-	-	-
N2692	*recD recA*	+^{b}	-	-	-	-
N2678	*recD*	+	+	±^{c}	±^{c}	-
N2364	*sbcC*	+	+	+	+	±^{d}
N2365	*recB sbcC*	+	+	+	+	+
N2680	*recD sbcC*	+	+	+	+	+

+ indicates an eop of $>10^{-1}$, - indicates an eop of $<10^{-2}$. Certain combinations of phage and bacteria are on the borderline between plating and not plating and the precise eop in these situations is very sensitive to minor variations in environmental conditions, such as the thickness of the plates, the age of the plates, the batch of the agar, and the temperature of the incubation. Platings were therefore done under standard conditions but in some cases it has still been necessary to report the observed range of eop that was observed in repeated experiments.

a. eop of 10^{-1}-10^{-2} (small plaques).

b. eop of $>10^{-1}$ (small plaques).

c. eop of 10^{-1}-10^{-2} (very small plaques).

d. eop of 1-10^{-2} (very small plaques).

3.3 Discussion

A Centre-Dependent Pathway May Exist *in Vivo*

In vitro, it has been shown that the central sequence dictates the cruciform extrusion rate for palindromes of average base composition (Murchie and Lilley, 1987; Courey and Wang, 1988; Zheng and Sinden, 1988). However, *in vivo*, the effects of central sequence have been less well documented. Warren and Green (1985) argued that the structure responsible for inviability was generated by a mechanism involving interaction of the arms of the palindrome and was independent of the centre, because relatively large asymmetries of >50bp were required to alleviate inviability. In contrast, Chalker (1990) and this chapter together demonstrate that when more sensitive assays are used, effects of much smaller changes in the central DNA sequence can be detected. As expected, and in agreement with Warren and Green (1985), all seven phages described above are unable to plate on wild type hosts. However, a marked effect of central asymmetry on the propagation of λ *pal* across a range of mutant *E. coli* hosts was observed. A similar concordance was noticed even in wild type hosts when the recovery of supercoiled DNA was analysed (Chalker, 1990). This suggests that a centre-dependent pathway for the phenotypic effect of palindromes exists in both wild type and mutant *E. coli.*

Palindromes carrying small central asymmetries confer a less severe phenotype than do perfect palindromes and the severity of phenotype correlates inversely with the length of asymmetry. Thus, a central asymmetry of only 8bp (λAD2) clearly reduces the severe phenotype caused by the perfect palindromes in λAD1, λDRL133, and λAD134. *In vitro* work has indicated that cruciform extrusion of perfect palindromes requires denaturation of the central 8-10bp (Murchie and Lilley, 1987; Courey and Wang, 1988). That a difference can be detected *in vivo* between perfect palindromes and a palindrome with 8 bases of asymmetry is consistent with a centre-dependent pathway of extrusion.

Catalysed Cruciform Extrusion *in Vivo*?

Cruciform extrusion *in vitro* occurs so slowly that, in normal palindromic DNA sequences under physiological conditions, it was predicted that cruciform extrusion *in vivo* may not occur (Courey and Wang, 1983, 1988; Gellert *et al.*, 1983). This kinetic argument against cruciform extrusion *in vivo* has been supported by several studies in which cruciform structures have failed to be detected *in vivo* (Courey and Wang, 1983; Sinden *et al.*, 1983). Only recent work has revealed that in certain instances cruciforms may form (McClellan *et al.*, 1990; Zheng *et al.*, 1991; Sinden *et al.*, 1991).

The results described above are most easily explained if cruciform extrusion does occur *in vivo* by perfect palindromes of normal composition, and even in palindromes with asymmetric centres of 8, 15, and 27bp. To reconcile the disparity with the S-type cruciform extrusion data of Murchie and Lilley (1987), it is tempting to suggest that cruciform formation is catalysed *in vivo*. The palindromes employed here are longer than 400bp, whereas other workers have tended to use palindromes less than ~150bp to avoid the propagation problems of longer sequences in plasmid hosts. It is, however, this phenotype of inviability that is likely to correlate with the accumulation of cruciform structures *in vivo*. It is predicted that (as for palindromes *in vitro*) the centre affects the kinetics of extrusion in a catalysed reaction, but the length of the palindrome is critical in determining the lifetime of the cruciform structure. Additionally, if the pathway to cruciform extrusion *in vivo* is related to S-type extrusion *in vitro* (Murchie and Lilley, 1987), then specific base pair changes will have predictable effects on the phenotype of palindrome containing phage *in vivo*.

CHAPTER 4:

Central Sequence of a Long Palindrome Affects DNA Secondary Structure Formation *in Vivo*

4.1 Introduction

For DNA palindromes of average base composition under moderate salt concentrations, the S-type pathway for cruciform extrusion *in vitro* has been proposed. It is illustrated in the introduction to this thesis. Changes to the central sequence of a palindrome affect the *in vitro* extrusion kinetics to a greater degree than more peripheral changes (Murchie and Lilley, 1987; Courey and Wang, 1988; Zheng and Sinden, 1988). A plausible explanation is that a central AT to GC change raises the free energy of formation of the transition state by discouraging DNA melting, whereas outside the centre a similar change has little effect upon transition state formation; the altered DNA usually remains unmelted during that step (Murchie and Lilley, 1987). It follows that if palindrome-mediated inviability is a result of a pathway involving cruciform extrusion, then specific base pair changes may have a predictable affect on the phenotype of palindrome-containing λ phage (λ *pal*).

Whether the inviability is a result of a cruciform structure is not yet known, and whether cruciforms do exist *in vivo* has been in dispute until relatively recently (Panayotatos and Fontaine, 1987; McClellan *et al.*, 1990; Zheng *et al.*, 1991; Sinden *et al.*, 1991). Chapter 3 describes the centre-dependent plating of palindromic λ strains on a range of *E. coli* hosts, providing evidence for the mechanism of inviability. In this chapter, the centre-dependence of palindrome-mediated inviability is further tested in an *sbcC* host by measuring the effect of central nucleotide sequence changes in a palindrome on the relative plaque size of the carrier λ, and on the presumed secondary structure formation *in vivo*. The *sbcC* mutation by itself is sufficient to allow plating (Chalker *et al.*, 1988); *sbcC* is now known to be one partner in a two gene system with *sbcD* (Gibson *et al.*, 1992).

The results argue that intra-strand pairing between the arms of a palindrome is critical in determining its effect on inviability, and are consistent with cruciform extrusion *in vivo*. The same data suggest that the structure of a DNA loop *in vivo* is sequence dependent, and in some cases the loop may consist of two residues only.

71

4.2 Results

Bacteriophage λ Cross

λDRL133 *spi6 cI857 pal462* was crossed with λDRL152 *spi6 cI857 χC153* by incubation of both phage in a culture of JC7623 *recBC sbcBC*. The recombinant phage, DRL167 *spi6 cI857 pal462 χC153*, was used for the cloning and plaque size analysis because it has comparatively large plaques on an *E. coli sbcC* lawn (due to the χ site).

Construction of a Set of Palindromes with Paired Changes in the Central Sequence

A series of long perfect palindromes, differing only in their central sequence, was constructed in λ phage to test the prediction that palindrome-mediated inviability is centre-dependent. The palindrome centres of the λ phage were related to the bke series of palindromes in pAT153 that had previously been used to test the S-type model for cruciform extrusion *in vitro* (Murchie and Lilley, 1987). The phage were named according to the corresponding bke pAT153 palindromes for ease of comparison between them (λADbke, λAD16T, etc.). As shown in Figure 4.1a, the central 12bp of the palindromes constructed in λDRL167 were identical to the equivalent bke centres in pAT153, but the adjoining palindromic arms of the λ constructs were different and over ten times longer.

Plating Behaviour of Palindrome Phage

The phenotype conferred upon the λ by the palindrome was assessed by plating on a lawn of N2364 *sbcC*, where differences in plaque size between the various λ phage were visible. These palindromic phage were inviable on wild-type hosts. Plaque area (under the standardised conditions used here) is assumed to be an indirect product of the λ burst size, and thus a measure of the viability of the palindromic phage. The plaque area was initially assessed on CA plates with a standard NaCl concentration (5g/l). A higher salt concentration (14.7g/l) was subsequently used for all the plaque assays reported here because it was found to

amplify the differences between λ *pal* phage. The areas of individual plaques were assessed accurately for each λ strain using an image analyser. Approximately 60-100 plaques per plate were analysed and twelve plates for each bacteriophage strain. Graphs of cumulative frequency plotted against plaque area for the eight λ phage are presented in Figures 4.2, 4.3, and 4.4. The data for these graphs are reproduced in Appendix 1. The median was the most reproducible parameter since it minimises the effects of outliers. Outliers were relatively common in this analysis for two reasons: λ phage that are initially unadsorbed produce pinprick plaques on the cell lawn, while revertants (phage that have lost the palindrome) produce very large plaques. The plaque assay median determined for each strain is presented in Table 4.1 and schematically illustrated in Figure 4.5a.

Central base-pair changes had the greatest effect on the plaque size. Figures 4.2, 4.3, 4.4, and particularly 4.5a clearly demonstrate this centre-dependence: λAD13G had smaller plaques than λAD10G, and both had smaller plaques than λADbke. Significantly, this pattern also included λAD15C and λAD15C16C. However, the precise determinants of the degree of inviability of a palindrome-containing λ phage are different from those that govern the rate of cruciform extrusion *in vitro* and seem inconsistent with a simple S-type mechanism. The central sequence of the palindrome had the predicted effect on plaque size only for changes to the very central two base pairs. Those λ phage with a 5'-AT-3' or a 5'-TA-3' at the very centre of the palindrome (λADbke and λAD16T, plaque sizes 0.36 and 0.33mm^2) had smaller plaques than the two phage with 5'-CG-3' at the centre (λAD16C, plaque size 0.70mm^2; λAD15C16C, plaque size 0.46mm^2), consistent with greater levels of the structure causing inviability for the former λ phage. For changes outside the central two base pairs the direction of change was the reverse of that predicted by the S-type model, implying that the intrastrand pairing (rather than DNA melting) between the arms of a palindrome may be critical in determining the level of inviability. Changes that *slowed in vitro* extrusion *decreased* the plaque size, and vice versa. For instance, λAD13G (central sequence GGAATTCC) had smaller plaques than λADbke (central sequence AGAATTCT).

73

Figure 4.1

Bacteriophage λ and plasmid constructs. **a.** Top: Partial restriction map of λDRL167 (not to scale). The full length of the palindrome *Eco*RI restriction fragment is 454bp. Bottom: The structure and derivation of the palindromes studied, showing part of the potential cruciform structure on the right. The top sequence λADbke is considered the parent and differences from it are highlighted as open letters. To ease comparisons, each λ was named similarly to the equivalent Murchie and Lilley (1987) construct in pAT153. At the bottom is the entire parent palindrome sequence that Murchie and Lilley (1987) employed and it is presented for comparison.

b. A simplified map of pMS2B, a derivative of pUC18. *Taq*I restriction fragments containing the centre of the palindrome were subcloned into the *Acc*I site of pMS2B. The resultant plasmids were named similarly to their λ parents (pADbke, pAD16T, *etc*).

Figure 4.1a

DRL167   EcoRI   SacI   EcoRI   (spi6)   cI857

TaqI   TaqI

Lambda AD Strain	5' ... 3'	Restriction Site
bke	GAATGAGCTTCCAGAA:TTCTGGAAGCTCATTC	EcoRI
16T	GAATGAGCTTCCAGAT:ATCTGGAAGCTCATTC	EcoRV
15C	GAATGAGCTTCCAGCA:TGCTGGAAGCTCATTC	SphI
16C	GAATGAGCTTCCAGAC:GTCTGGAAGCTCATTC	AatII
15C16C	GAATGAGCTTCCAGCC:GGCTGGAAGCTCATTC	NaeI
13G	GAATGAGCTTCCGGAA:TTCCGGAAGCTCATTC	EcoRI
11T	GAATGAGCTATCAGAA:TTCTGATAGCTCATTC	EcoRI
10G	GAATGAGCTGCCAGAA:TTCTGGCAGCTCATTC	EcoRI

```
          C,
           T
          C  AT
           A   T
              GC
          G  AT
              CG
          T  CG
         G,A TA
              TA
              CG
              GC
           : :
```

GGATCCGGTACCAGAA:TTCTGGTACCGGATCC

b.

TaqI   TaqI

ScaI   AccI
2177   417
pMS2B
2.7kb
0

Table 4.1

Kinetics of Cruciform Extrusion and Plaque Assay Data for the Palindrome Plasmids and Phage.

	Cruciform Extrusion: bke series in pAT153, (NaCl buffer) from Murchie and Lilley (1987)					Cruciform Extrusion: pAD series, (NaCl buffer)				Cruciform Extrusion: pAD series, (buffer P)			Median Plaque Area
Mutant	k(error)	exp	$T_{1/2}$ (min)	NaCl (mM) Optimum		k(err)	exp	$T_{1/2}$ (min)	NaCl(mM) Optimum	k(err)	exp	$T_{1/2}$ (min)	Area (/mm²)
bke	-1.7(.02)	-4	70	50		-1.9(.34)	-3	6.1	45	-1.3(.17)	-3	8.9	0.36
16T	-1.3(.14)	-3	9	50		-4.0(.41)	-3	2.9	45	-2.1(.48)	-3	5.5	0.33
15C	-1.7(.04)	-4	67	75		-2.0(.26)	-3	5.8	45	-2.1(.29)	-3	5.5	0.17
16C	-1.7(.18)	-5	692	50		-1.9(.20)	-5	608	45	-5.2(.78)	-6	2222	0.70
15C16C	-7.0(.60)	-6	1649	50		-1.2(.14)	-5	963	45	-6.9(.69)	-6	1674	0.46
13G	-1.0(.09)	-4	114	75		-8.9(.15)	-4	13.0	60	N/D	-	-	0.24
11T	-1.5(.07)	-3	8	50		-4.8(.53)	-3	2.4	45	-1.2(.12)	-3	9.6	0.41
10G	-1.4(.12)	-4	81	50		-2.4(.26)	-4	48.1	45	N/D	-	-	0.33

Rate constants are expressed as k(± standard error) × 10^{exp}. Half times were calculated from ln2/k. N/D: k not determined.
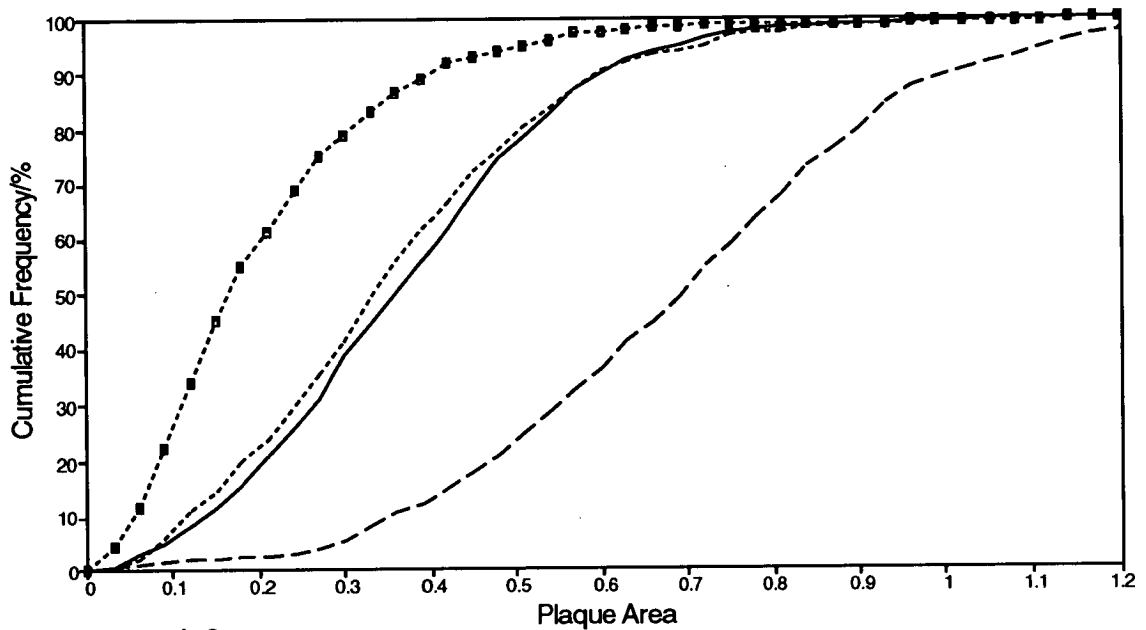
Figure 4.2

Cumulative frequency distribution of plaque sizes for four of the λ phages (plaque area in mm^2). The differences between λAD15C, λADbke, and λAD16C were clearly visible to the eye, but the difference between λADbke and λAD16T was only apparent after quantification by image analysis. The median was taken as the point at which the 50% cumulative frequency intersected the line on the x-axis.
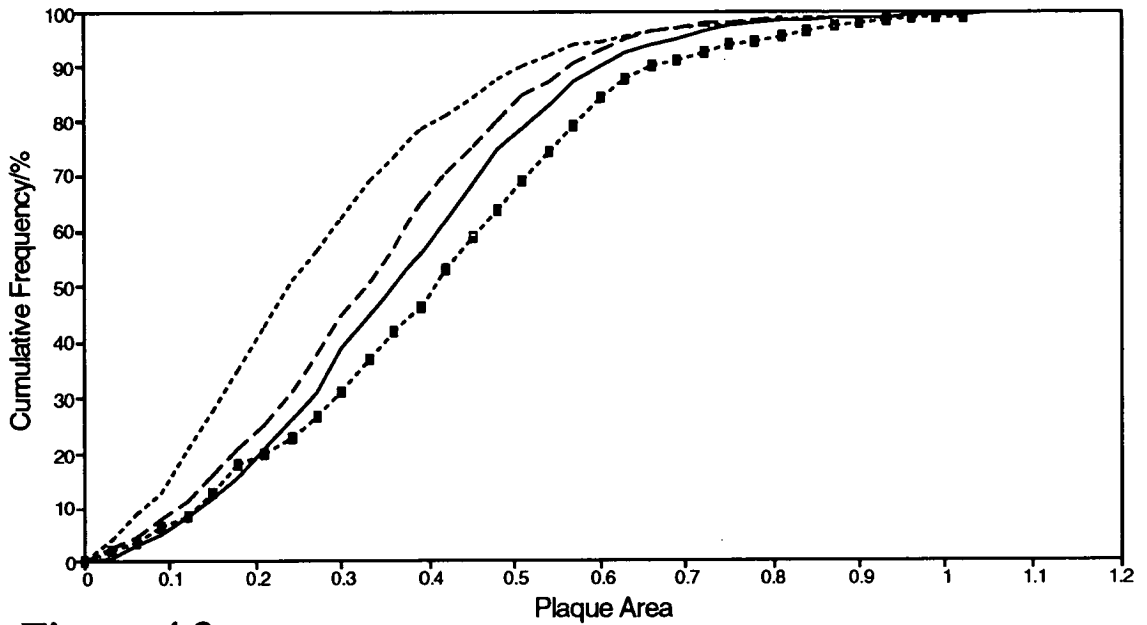
Figure 4.3

Legend: ADbke ···· AD13G ·✦· AD11T − − AD10G

Cumulative frequency distribution of plaque sizes for four of the λ phages (plaque area in mm^2). The median was taken as the point at which the 50% cumulative frequency intersected the line on the x-axis. λAD13G and λAD10G both have extra CG pairs outside the central six bases. λAD11T has extra AT pairs outside the centre compared with λADbke.

78
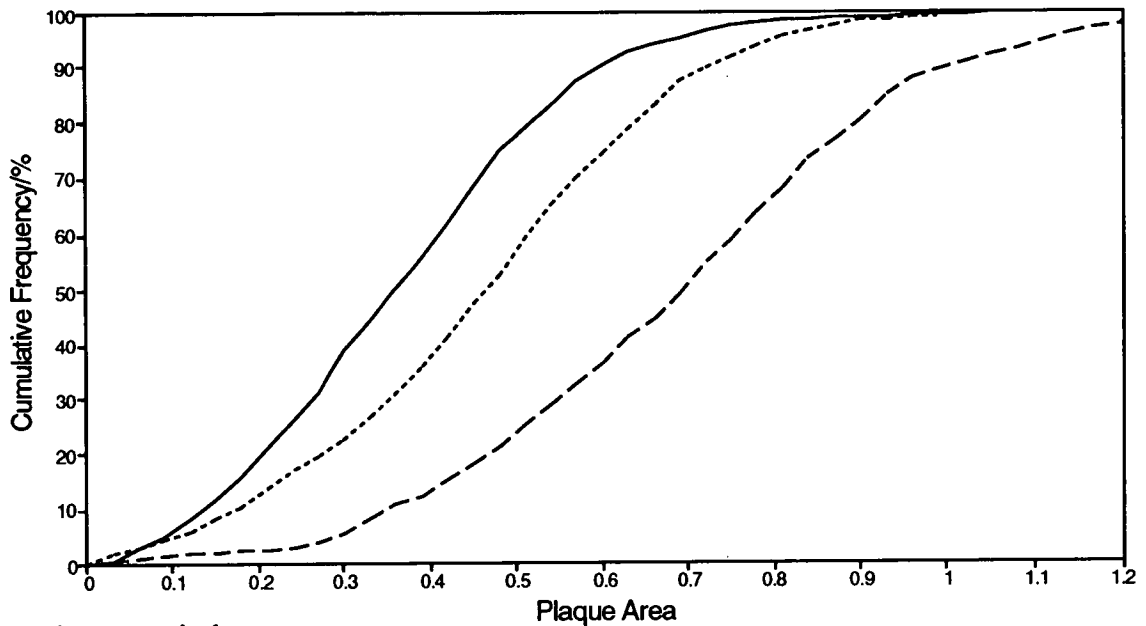
Figure 4.4

Legend: ADbke ···· AD15C16C — — AD16C

Cumulative frequency distribution of plaque sizes for three of the λ phages (plaque area in mm^2). The median was taken as the point at which the 50% cumulative frequency intersected the line on the x-axis. λAD16C and, λAD15C16C have one and two more GC-pairs, respectively, in the centre than λADbke, yet λAD15C16C plaques were smaller than those of λAD16C.

79

Figure 4.5

a. Schematic illustration of the variation in plaque sizes for the λAD mutants. The size of the bar is proportional to the difference in plaque area compared with λADbke. λAD15C16C is dotted to make it easily visible. Mutations that decreased plaque size are shown above the line and mutations that increased it, below. Only mutations to the very central two base pairs had an effect *in vivo* that was predicted by the *in vitro* results under optimal salt conditions (see 4.5b and c), assuming that increased cruciform extrusion corresponds to decreased plaque size and vice versa. All other mutations altered the plaque size in a reverse manner compared with the *in vitro* data, the effect lessening further out from the centre.

b., c., and d. Schematic illustration of the variation in cruciform extrusion rates *in vitro*. The size of the bar is proportional to the change in rate for that sequence compared with the parent bke/pADbke (breaks indicate that insufficient space was available). Mutations that enhanced the *in vitro* extrusion rate are shown above the line and mutations that depressed it are shown below the line. The results for 16C are overlaid and dotted so that they are easily visible. **b.** and **c.** Data from Murchie and Lilley (1987) and the pAD mutants, respectively, in the optimal salt (NaCl) buffer. Mutations to A/T base pairs tended to enhance extrusion and mutations to C/G had the reverse effect. Changes to the centre had the greatest effect, although changes in the peripheral base sequence were not always insignificant (see 11T in 4.5b and 10G in c). The same changes had similar effects in the two plasmid systems, although mutations to C/G pairs had a much greater relative effect in pMS2B. In both instances the 15C mutation caused extrusion to be unexpectedly fast.

d. Extrusion in Buffer P (150mM potassium glutamate/ 4mM magnesium acetate) for the pAD mutants. This buffer was selected to mimic physiological conditions. For six out of the eight palindromes, cruciform extrusion *in vitro* was found to follow the same relative order compared with the plaque area data (evident from the similar shapes of 4.5a and d).

Figure 4.5

a.

Smaller

Larger

b.

Faster

Slower

c.

Faster

Slower

d.

Faster

Slower
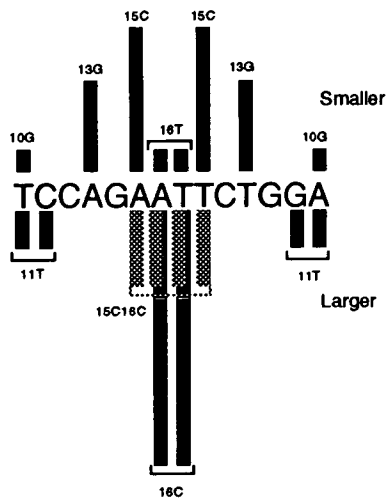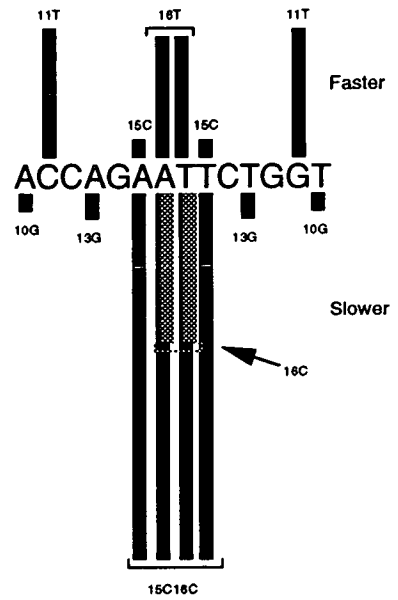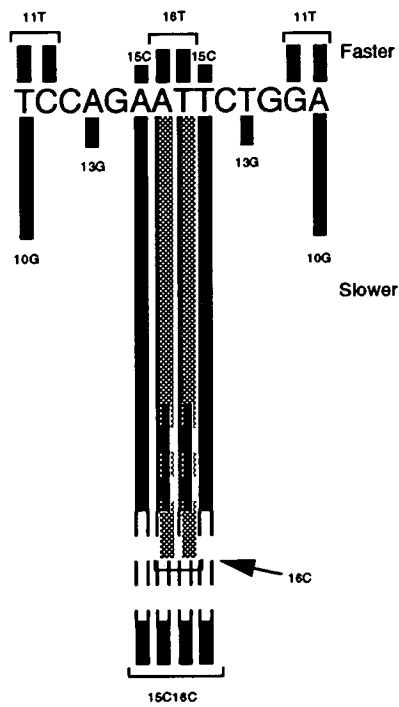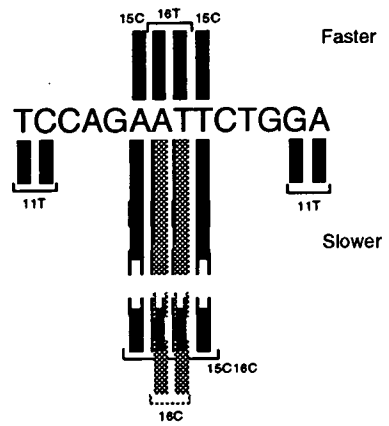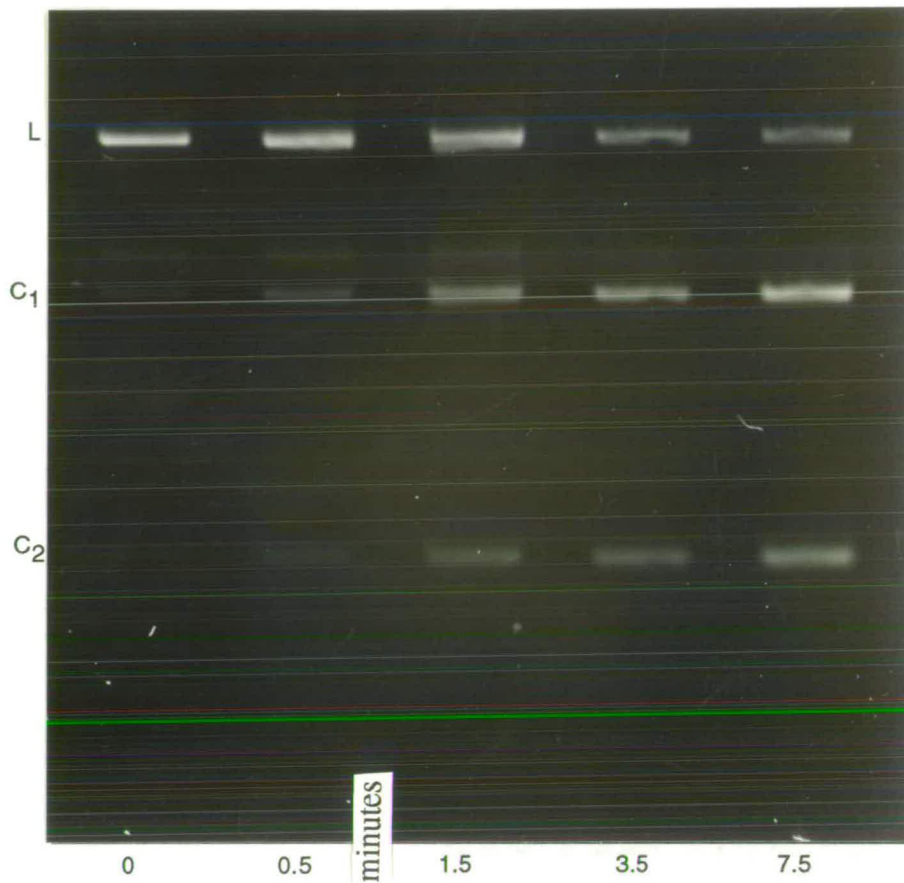
Figure 4.6

An example of cruciform extrusion at 37°C. pAD11T was incubated in 45mM NaCl, 10mM Tris-HCl (pH7.5), 0.1mM EDTA, and sampled between 0 and 7.5 minutes as shown. Full length linear plasmid DNA is indicated by the letter L. The fragments resulting from T4 endoVII cleavage at the cruciform are labelled C_1 and C_2.

Cruciform Extrusion *in Vitro* of Two Related Series of Palindromes

It was necessary to confirm that the central changes in λ had the same effect on *in vitro* extrusion as the bke centres in pAT153. Since long DNA palindromes are unstable in plasmid vectors (even with the *sbcC* mutation), it was not possible to use the full length palindromes. Instead a 32bp *Taq*I fragment containing the λ palindrome centre was subcloned into the *Acc*I site of pMS2B (forming the pAD series of plasmids), then sequenced to check that the constructs were as predicted. A partial map for these constructs is illustrated in Figure 4.1b. The extrusion kinetics of the palindrome centres in the pAD plasmids were compared with the similar centres of Murchie and Lilley (1987) in pAT153.

The cruciform extrusion salt (NaCl) optima that were determined for the pAD series are indicated alongside the equivalent bke series salt optima in Table 4.1. For most of the reactions the salt optimum was found to be 45mM, compared with 50mM in the experiments of Murchie and Lilley (1987). For the two bke sequences that had a raised salt optimum of 75mM (15C and 13G), only one of the pAD equivalents had a similarly raised optimum (pAD13G, 60mM). To determine the reaction constant a time course was performed at the optimum salt concentration. A sample extrusion time course experiment is illustrated in Figure 4.6. The extrusion data are detailed in Table 4.1 and further compared with those of Murchie and Lilley (1987) in Figure 4.5b and Figure 4.5c. In general, the changes that were made to the palindromes had comparable effects in both vectors, although the absolute values varied considerably.

Cruciform Extrusion *in Vitro* in a "Physiological" Buffer

Cruciform extrusion of the pAD series of palindromes was carried out in a more physiological buffer (buffer P; containing potassium glutamate and magnesium acetate, suggested by G. Smith). The extrusion half times that were determined are detailed in Table 4.1 and compared with the plaque assay results in Figure 4.5a and 4.5d. It was found that buffer P had a significant effect upon the relative extrusion kinetics of the palindromes, and that the reaction half times for six out of the eight plasmids could be compared to the equivalent order of plaque areas. The reaction constants for the two other plasmids remained relatively unchanged in this buffer.

Various other buffers were also used, but none were found to have as significant an effect as buffer P.

UV Melting Analysis of Short Oligonucleotides

Ultraviolet melting experiments were carried out (with the assistance of T. Brown, Edinburgh) on short oligonucleotides of the palindrome centres that could form hairpin stem-loop structures. UV melting transitions of DNA hairpins potentially give information on the thermal stability of the loop and stem, and are thus an indicator of its structure. Since the DNA sequences involved are all perfect palindromes it was found that double strand DNA was always favoured over a hairpin loop structure, and a hairpin melting transition was not be observed. This was the case even at low oligonucleotide concentrations, and in very low salt buffers.

4.3 Discussion

Palindrome-Mediated Inviability in an *E. coli sbcC* Host is Centre-Dependent

Evidence for the existence of DNA cruciforms *in vivo* has been comparatively lacking until recently (Panayotatos and Fontaine, 1987; McClellan *et al.* 1990; Dayn *et al.* 1991; Zheng *et al.* 1991; Sinden *et al.* 1991) and the pathway for extrusion is only beginning to be investigated. Warren and Green (1985) demonstrated that large insertions of at least 50bp in the centre of a palindrome were required before inviability was overcome. This led them to argue that the structure responsible for inviability was dependent upon interactions between the palindrome arms and did not involve the centre. However, Chalker (1990) and Chapter 3 of this thesis discuss the detection of the increased viability produced by much smaller asymmetries in long DNA palindromes. The conclusion was that a centre-dependent pathway for inviability operates and moreover, that it is independent of *E.coli* genotype for the strains studied. Zheng *et al.* (1991) came to a similar conclusion regarding centre-dependence using a physical assay for cruciforms, and Sinden *et al.* (1991) reported that the stability of base pairing in the hairpin stem and ease of cruciform formation

affect the deletion frequency. Of direct relevance to the plating assay described here, Allers (1993) has recently demonstrated that the central sequence of a set of long DNA palindromes in λ phage adopt a methylation resistant structure.

If a centre-dependent pathway for the formation of a cruciform structure exists in *E.coli* then this predicts that palindrome-mediated inviability will be particularly sensitive to the base sequence at the centre of the palindrome, and progressively less sensitive to more peripheral sequence changes. The results reported here confirm that a centre-dependence exists in *E. coli sbcC*; the relevant data are presented in Table 4.1 and Figure 4.5.

Intrastrand Pairing is Critical to Viability in *sbcC* Hosts

The plaque size of the λ phage was progressively less affected by changes further outside the centre of the palindrome, implying that palindrome-mediated inviability is centre-dependent (evident from the triangular shape of Figure 4.5a). Such effects are consistent with a cruciform causing inviability *in vivo*. However, the effects of base sequence changes on viability do not correlate with S-type extrusion kinetics at the optimal salt concentration *in vitro*. In fact, it was found that changing GC to AT base pairs (outside the central 2bp) increased the plaque size, and vice versa. These results argue that intrastrand pairing between bases in the palindrome centre, as opposed to central melting, is the major determinant of viability. The effects of sequence changes on viability can only be interpreted in terms of DNA melting for the central two base pairs. Thus, the λ phage with a central 5'-AT-3' or 5'-TA-3' sequence (λADbke, λAD16T) both have considerably smaller plaques than the comparable phages with a 5'-CG-3' sequence at the centre (λAD16C, λ15C16C). Alternatively, the differences between these four phages could be explained in terms of altered loop structure and stability, rather than DNA melting. The fact that sequence changes up to 7bp away from the centre have an effect upon the plaque size (compare λAD10G with λADbke), suggests that the structure causing inviability may not be fully stabilised unless the stem is at least 5-6bp. Whether or not cruciform structures are responsible for the observed effects on viability, the results argue strongly for a role of intrastrand base pairing at the palindrome centre.

85

Cruciform Extrusion *in Vitro* in a "Physiological" Buffer Models *in Vivo* Behaviour More Closely

Cruciform extrusion *in vitro* in a "physiological" buffer models the effects of long palindromes on λ viability more closely than S-type extrusion under optimal salt (NaCl) conditions. The extrusion reactions in buffer P suggest that, under these conditions, intrastrand base pairing (as opposed to DNA melting) becomes relatively more important in determining the rate of *in vitro* extrusion. That it is possible to partly model the viability by cruciform extrusion *in vitro* provides further evidence that a cruciform structure may be responsible for the effects measured *in vivo*. It is likely that the added magnesium is responsible for much of the altered extrusion kinetics observed since magnesium is known to specifically effect DNA supercoiling, Holliday junction structure, and branch migration. However, if palindrome-mediated inviability is due to the presence of cruciform structures *in vivo*, then it will be determined by either the process of getting there (cruciform extrusion) or staying there (cruciform stability). Clearly, the cruciform extrusion reactions model only part of this. Additionally, it is not possible to rule out the effect of protein-DNA interactions that might occur as a consequence of processes such as transcription, replication, or homologous recombination.

DNA Loop Structures *in Vivo* May Contain Only Two Residues

Most information on DNA loops has come from studies of DNA hairpins *in vitro*. Early reports suggested that a loop of four residues was normal (Haasnoot *et al.*, 1983; Hilbers *et al.*, 1985), but recently several studies have concluded that hairpin loops of only two residues may exist with specific base sequences (Xodo *et al.*, 1988; Blommers *et al.*, 1991; Kallick and Wemmer, 1991; Raghunathan *et al.*, 1991). Conclusions on loop size for the cruciforms of the λAD phage can be reached by comparing the plaque areas of λADbke (0.36mm^2), λAD15C (0.17mm^2), λAD16C (0.70mm^2), and λAD15C16C (0.46mm^2). λAD15C had the smallest plaques among the phages constructed, yet its extra CG base pairs were in positions that would be unpaired in a four residue cruciform loop, and should therefore not have affected intrastrand pairing to a great extent. Similarly, λAD15C16C had CG base pairs at these same positions in place of the AT base pairs present in λAD16C. It is noted

that the equivalent plasmid palindrome pAD15C (and the corresponding Murchie and Lilley (1987) bke palindrome) also extruded unexpectedly rapidly *in vitro*. Furthermore, in buffer P, both pAD15C and pAD15C16C extruded particularly rapidly, relative to pADbke and pAD16C, respectively. Murchie and Lilley (1987) suggested that the small discrepancy in their data was due to the alternating purine/pyrimidine structure (present in pAD15C) reducing the DNA stability at the centre of the inverted repeat. Whilst this is a possibility *in vitro*, a more likely explanation of the severe effect *in vivo*, (and in buffer P *in vitro*) is that the DNA loops of the cruciform structures formed in λAD15C and λAD15C16C contain two residues only (or perhaps were four base loops with some two base character or altered loop stacking). The data suggest that the extra CG pairs in λAD15C and λAD15C16C stabilise the loop structure *in vivo*, and therefore contribute to palindrome-mediated inviability.

This explanation is supported by the work of Blommers *et al.* (1989) who implied that the sequence 5'-CATG-3' could theoretically form a two base loop in a DNA hairpin *in vitro*. Lowered loop stability could also explain why λAD16T (5'-ATAT-3') had only slightly smaller plaques than λADbke (5'-AATT-3') despite the pAD16T palindrome extruding more rapidly *in vitro*. The loop of λADbke could be more stable *in vivo* because the central sequence was 5'-AT-3' with the purine in the 5' position rather than the reverse (Blommers *et al.*, 1989). A possibility that cannot be discounted is that the *in vivo* melting stabilities of dinucleotide base pairs are not identical to the values that have been determined *in vitro*. The observed differences between the palindrome-mediated inviability and cruciform extrusion *in vitro* may not be surprising in view of the highly complex environment of the cell.

Implications for Cruciform Extrusion of Naturally Occurring Palindromes in Wild-Type Hosts

The results presented here demonstrate the existence of a centre-dependent pathway of palindrome-mediated inviability in *sbcC* mutants, and therefore suggest that a pathway for cruciform formation exists *in vivo*. The results of Chalker (1990) and Chapter 3 together imply that cruciform extrusion of long palindromes is not dependent on the *sbcC* genotype and therefore that SbcC may act in the processing

of cruciform structures rather than in their formation. This is supported by two other observations both arguing indirectly that *sbcC* may encode a nuclease. The first is that the *gam* gene of bacteriophage λ (known to encode a nuclease inhibitor) permits the propagation of long palindromes via an interaction with SbcC (Kulkarni and Stahl, 1989). The second is that the SbcC polypeptide is distantly related to genes responsible for the major exonuclease activity of bacteriophages T4 and T5 on host chromosomal DNA, T4 *gp46* and T5 *gpD13* (Leach *et al.*, 1992). Furthermore, it has been shown that *sbcC* mutants do not have an altered level of intracellular DNA supercoiling which could facilitate the extrusion reaction (J. Lindsey and D. Leach, unpublished). Recently, evidence that SbcC and D constitute an exonuclease has been obtained (J. Connelly, D. Leach, personal communication).

The implication of the results presented here is that a centre-dependent pathway for palindrome-mediated inviability, consistent with cruciform extrusion, also exists for wild-type hosts. At a given threshold length or stability of intrastrand base pairing the structure may become a substrate for the SbcC and D proteins. The short range of the base sequence changes studied here (maximum seven bases from the centre) argues that similar transient reactions are likely to occur in short, naturally occuring DNA palindromes. However, in these cases the equilibrium will normally be significantly in favour of the unextruded conformation for thermodynamic reasons, so that no inviability results.

CHAPTER 5:

The Effect of Osmotic Stress on λ and λ *pal* Viability *in Vivo*

5.1 Introduction

DNA supercoiling is implicated in the osmotic regulation of gene expression (Eshoo, 1988; Higgins *et al.*, 1988; Ni Bhriain *et al.*, 1989; Hsieh *et al.*, 1991). More specifically, the linking number of the intracellular DNA is related to the degree of osmolarity, although the relationship is poorly understood. Also, the response to osmotic stress throughout the chromosome may not be equal, although recent evidence has indicated that the independently supercoiled *E. coli* chromosomal domains do not necessarily differ in functional superhelical density (Miller and Simons, 1993).

In this chapter, the effect of osmotic stress (NaCl) on the plaque size of a set of palindrome-containing and palindrome-free λ phage is measured. Since negative supercoiling is necessary to cruciform extrusion, osmotic stress may have an indirect effect on λ *pal* viability.

5.2 Results

Construction of a Set of Palindromes with Paired Changes in the Central Sequence

A series of long perfect palindromes was constructed in λ phage. The same λ *pal* phage were used as those described in Chapter 4 and Figure 4.1. Two other phage lacking a palindrome were also used so that palindrome specific effects of osmotic stress could be isolated from any non-specific effects. λDRL152 and λDRL112 are otherwise isogenic to the λ *pal* phage, except that the latter also lacks a χ site.

Plating Behaviour of Palindrome Phage

The phenotype conferred upon the λ by the palindrome was assessed by plating on a lawn of N2364 *sbcC*, where differences in plaque size between the various λ phage were visible. This has been described previously in Chapter 4. The

plaque area was assessed accurately on CAS plates with added NaCl concentrations between 85mM (5g per litre; standard for λ plating) and 300mM using an image analyser. Approximately four plates for each λ phage were measured at a given NaCl concentration. Graphs of plaque area against NaCl concentration are presented in Figures 5.1, 5.2, 5.3, 5.4, and the median plaque areas detailed in Table 5.1.

The effect of NaCl was the same for all of the λ and λ *pal* phages examined. Similar differences between the plating of the various λ *pal* phage at different salt concentrations were observed as those described in Chapter 4 with 250mM added salt.

Table 5.1

Median plaque areas of λAD *pal* phage ($/mm^2$) on CAS agar with specified quantities of added NaCl. N/D: not determined.

λ Phage	{NaCl}/mM			
	85	140	250	300
bke	0.65	1.24	0.36	0.16
16T	0.56	1.26	0.33	0.16
15C	0.53	1.10	0.17	0.02
16C	0.75	1.69	0.70	0.70
15C16C	0.71	1.46	0.46	0.34
13G	0.60	1.09	0.24	0.08
11T	0.77	1.31	0.41	0.22
10G	N/D	1.12	0.33	0.14
DRL112	0.38	0.79	N/D	0.08
DRL152	3.4	4.0	3.1	2.9
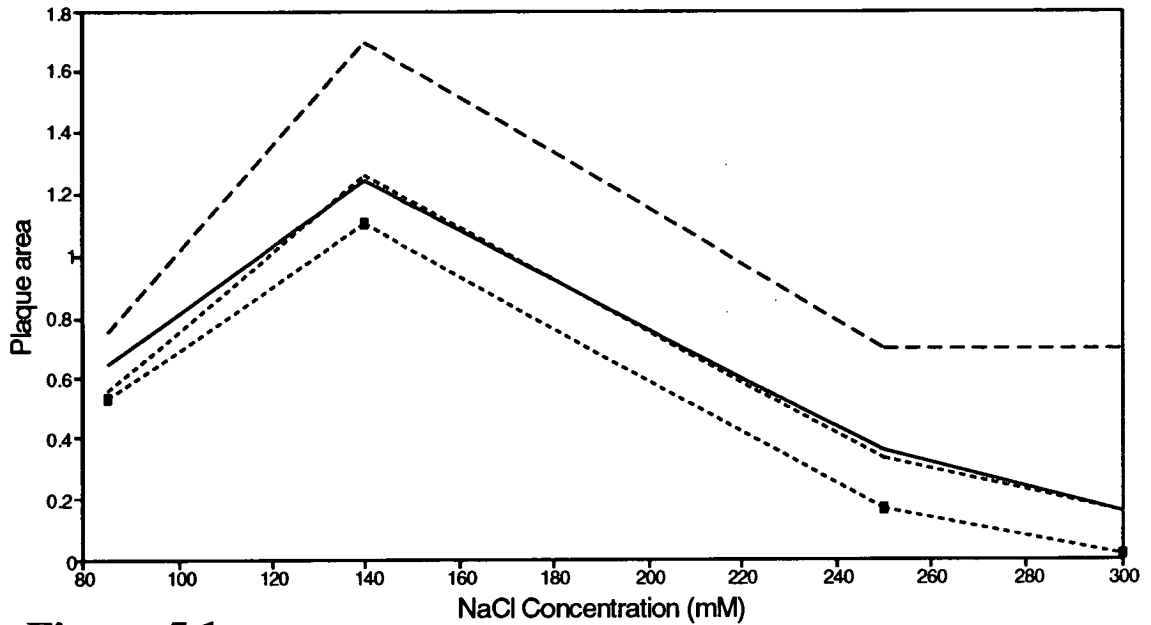
Figure 5.1

Median plaque area of λAD *pal* phages with varying quantities of added NaCl compared against λADbke (plaque area in mm^2). For all of the NaCl concentrations used λAD16C plaques were larger than λADbke, and λAD16T plated similarly to λADbke. λAD15C plaques were barely visible with 300mM added salt.

Figure 5.2

Legend: ADbke — · · · · AD15C16C — ✖ · DRL112 — – – AD16C
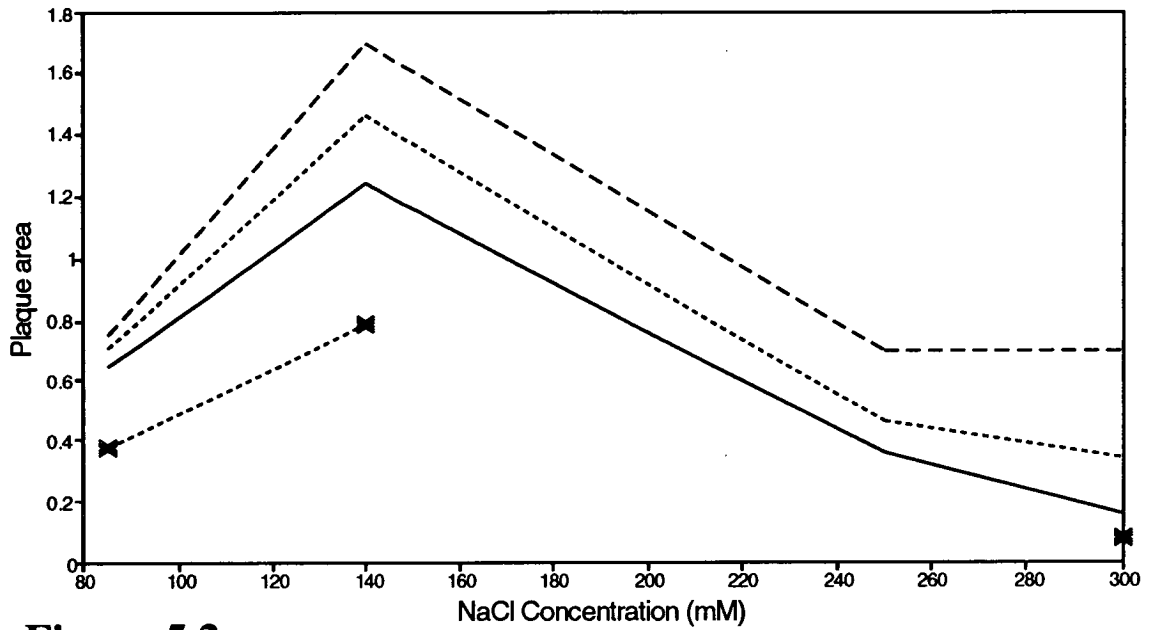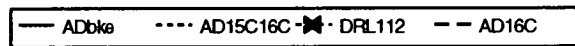
Median plaque area of λAD *pal* phages and λDRL112 with varying quantities of added NaCl compared against λADbke (plaque area in mm^2; value not determined for DRL112, 250mM). For all of the salt concentrations used λAD15C16C had plaques intermediate in size between λADbke and λAD16C.
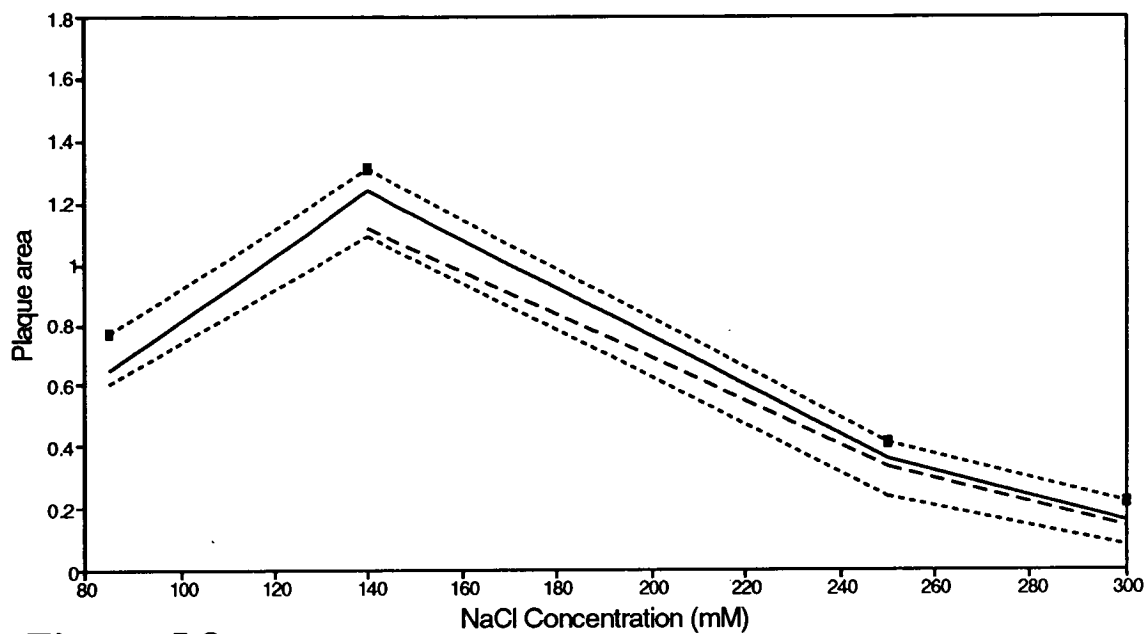
Figure 5.3

Median plaque area of λAD *pal* phages with varying quantities of added NaCl compared against λADbke (plaque area in mm²). The value for λAD10G with 85mM added salt was not determined. For the salt concentrations used, replacing an AT pair with a GC pair in the palindrome arms decreased the plaque size (λAD10G and λAD13G) and *vice versa* (λAD11T).

Figure 5.4
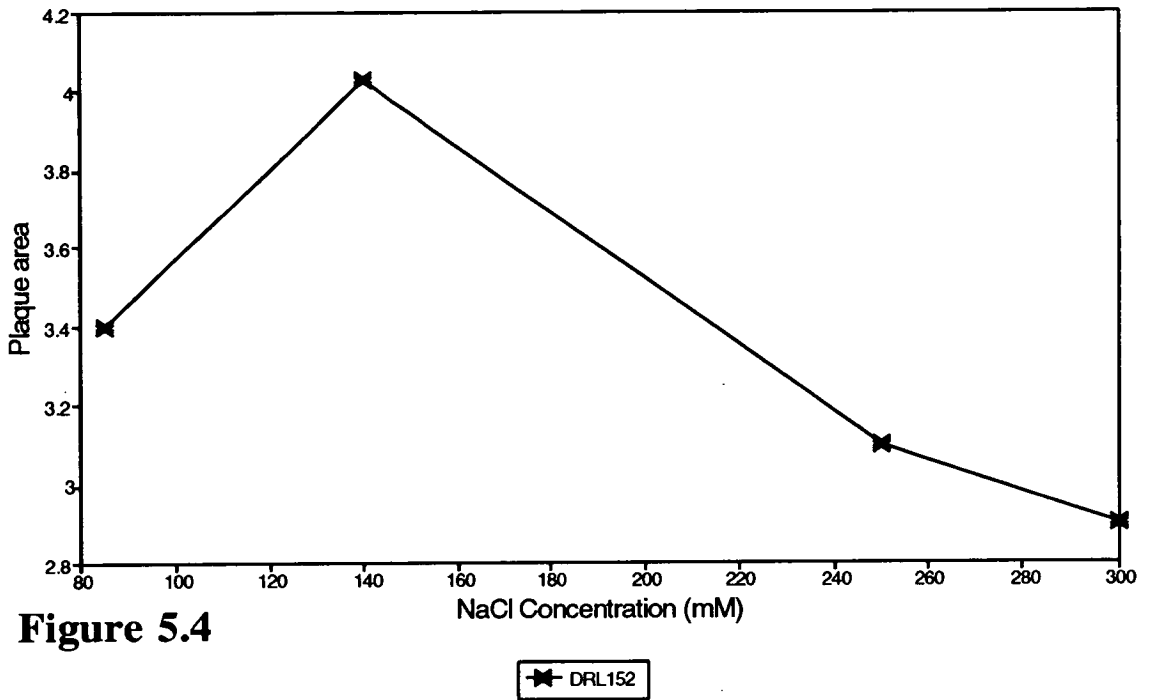
Median plaque area of λDRL152 with varying quantities of added NaCl (plaque area in mm^2). The effect of osmotic stress on λDRL152 plating was similar to the effect on the λ *pal* phage.

5.3 Discussion

λ Phage Plating is Affected by Osmotic Stress

High NaCl concentrations have a quantifiable effect on the cellular DNA supercoiling levels via the osmotic stress that is exerted (Higgins *et al.*, 1988), although the effect on λ supercoiling is not known. In this chapter, the effect of osmotic stress on viability was assayed with a series of palindrome-containing and palindrome-free λ phage in *E. coli sbcC*.

Increased salt concentration initially increased the viability (as measured by plaque size), but above ~140mM NaCl the viability decreased. At all of the salt concentrations assayed each λ phage maintained a similar relative viability, but λ phage lacking a palindrome were equally affected. It is likely that high salt concentrations affected λ plaque size due to a general effect on viability.

A final observation is that depending upon the salt concentration used, the plaque area of an isogenic λ (λDRL152) was between ~1.4 and 145 times larger on N2364 *sbcC* than for some of the λAD *pal*. Clearly, although *sbcC* fully alleviates inviability in terms of plating efficiency, there remains a variable degree of inviability that depends on the extremeness of the palindrome phenotype.

CHAPTER 6:

The Effect of Further Nucleotide Sequence Changes on DNA Secondary Structure Formation *in Vivo*

6.1 Introduction

Chapter 4 reports that central nucleotide sequence changes to a palindrome affect the relative plaque size of the carrier λ in a centre-dependent manner. The results argue that intrastrand pairing between the arms of a palindrome is critical in determining the effect on viability, and may be consistent with cruciform extrusion *in vivo*. It is also suggested that the structure of a DNA loop *in vivo* may be sequence dependent, and in some cases, the loop may consist of two residues only. The two sequences with suspected two residue loops described in that chapter are λAD15C (CATG central four bases) and λAD15C16C (CCGG). Both have the sequence C_ _G in common. In this chapter, the DNA sequence required for the formation of two residue loops *in vivo* is explored further.

Additionally in Chapter 4, it was found that the effects of sequence changes on viability may only be interpreted in terms of DNA melting for the central two base pairs, a conclusion arising from a comparison between λ *pal* phage with AT, TA, or CG sequences in the central two positions. This chapter describes the construction of a λ *pal* phage with a GC centre to test whether this hypothesis applies to all the possible palindrome centres.

6.2 Results

Construction of Palindromes with Paired Changes in the Central Sequence

Three long perfect palindromes, differing only in their central sequence, were constructed in λ phage (λADbkb, λADbkc, and λADbkd). The palindrome centres of the λ phage were related to the bke series of palindromes in pAT153 (Murchie and Lilley, 1987, 1989), and the central sequences of the λ phage are described in Figure 6.1.

The central sequence of λADbkb (GGATCC) is identical to a sequence used by Murchie and Lilley (1989) to test whether disrupting the alternating

purine/pyrimidine sequence present at the centre of the bke15C palindrome (GCATGC ie PuPyPuPyPuPy) decreases the rate of cruciform extrusion *in vitro*. The sequences of λADbkc and λADbkd do not have exact equivalents in the sequences investigated by Murchie and Lilley (1987, 1989).

The palindrome of λADbkc is identical to λAD15C except the central 5'–AT–3' is replaced by 5'–TA–3'. This construct tests whether switching of the position of the bulky purine residue (A) and disrupting the alternating purine/pyrimidine sequence has an effect on loop structure and stability (as judged by altered viability). λADbkd has a palindrome sequence with a central GC to complete the set of λ phage with AT, TA, CG, or GC at the centre, and further tests whether inviability is interpretable in terms of DNA melting at this position. Unfortu--nately, it was also necessary to alter the next pair from the centre (compared with λADbke) because otherwise the restriction site, *Sac*I, (necessary for the cloning procedure) would have been formed.

Plating Behaviour of Palindrome Phage

The phenotype conferred upon the λ by the palindrome was assessed by plating on a lawn of N2364 *sbcC* on CA plates. This has already been described in detail in Chapter 4 and Materials and Methods. Graphs of cumulative frequency plotted against plaque area for the three new λ phage and three controls from Chapter 4 are presented in Figures 6.2, 6.3, and 6.4 and the medians are detailed in Table 6.1. The data for these graphs are reproduced in Appendix 1.

Table 6.1

Plaque assay data for λADbkb, λADbkc, and λADbkd. The other λ are controls.

Phage:	ADbke GAATTC	AD15C GCATGC	AD16C GACGTC	ADbkb GGATCC	ADbkc GCTAGC	ADbkd GTGCAC
Median Plaque Area (/mm²)	0.61	0.20	1.15	0.63	0.20	0.58

DRL167   EcoRI   SacI   EcoRI   cI857

        Taq l   Taq l

Lambda AD Strain	5' ... 3'	Restriction Site
bke	GAATGAGCTTCCAGAA:TTCTGGAAGCTCATTC	EcoRI
15C	GAATGAGCTTCCAGCA:TGCTGGAAGCTCATTC	SphI
16C	GAATGAGCTTCCAGAC:GTCTGGAAGCTCATTC	AatII
bkb	GAATGAGCTTCCAGGA:TCCTGGAAGCTCATTC	BamHI
bkc	GAATGAGCTTCCAGCT:AGCTGGAAGCTCATTC	NheI
bkd	GAATGAGCTTCCAGTG:CACTGGAAGCTCATTC	SnoI

Figure 6.1

Bacteriophage λ constructs. Top: Partial restriction map of λDRL167 (not to scale). Bottom: The structure and derivation of the λ *pal* used in this chapter. The top sequence λADbke is considered the parent and differences from it are highlighted as open letters. λADbke, λAD15C, and λAD16C are the controls for this chapter. λADbkb, λADbkc, and λADbkd are new constructs.
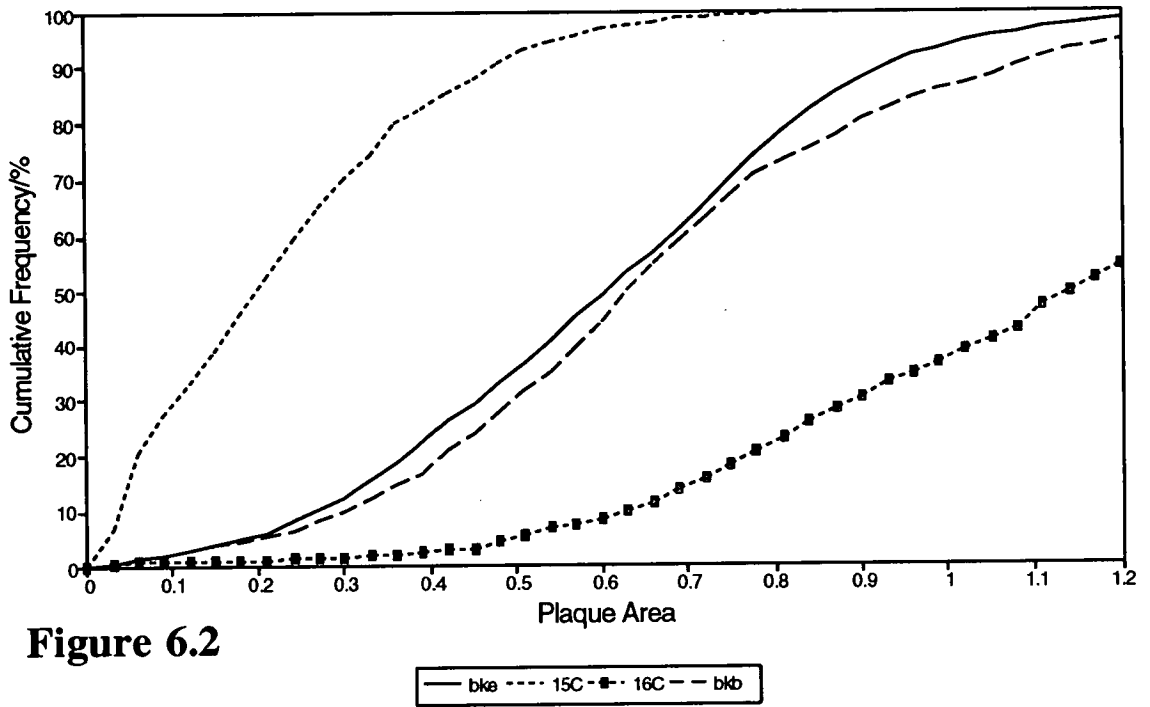
Figure 6.2

Cumulative frequency distribution of plaque size for λADbkb compared against λADbke, λAD15C, and λAD16C (plaque area in mm^2).

Figure 6.3

Cumulative frequency distribution of plaque size for λADbkc compared against λADbke, and λAD15C (plaque area in mm^2).
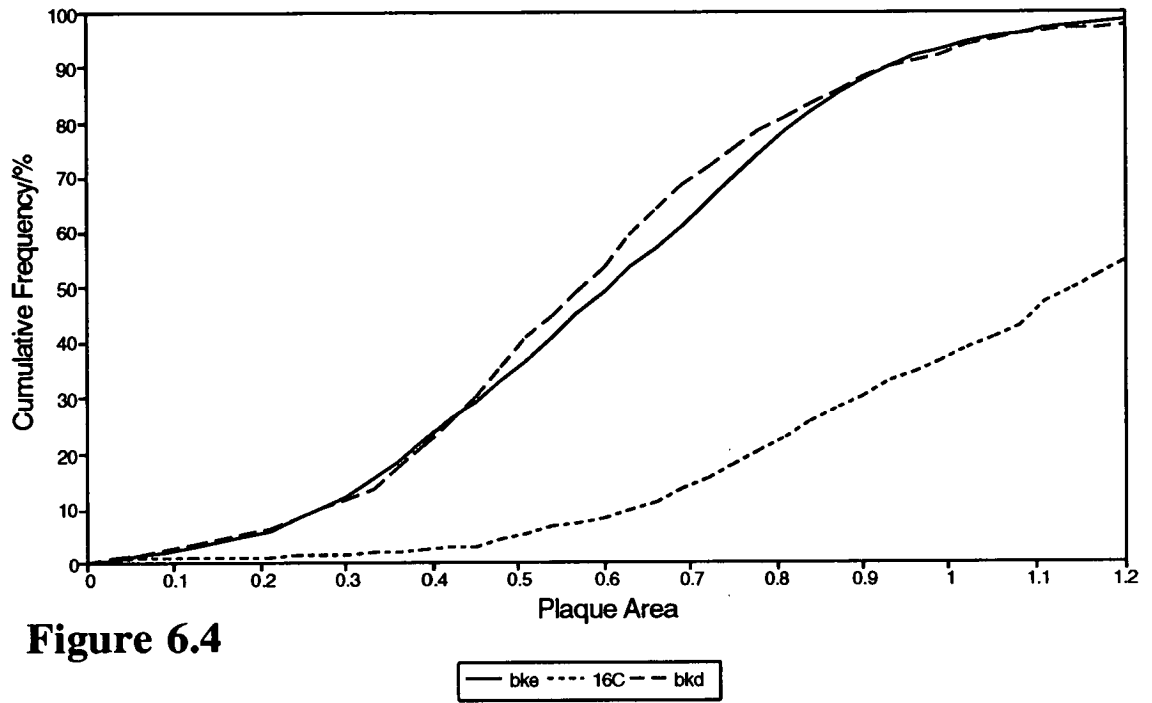
Figure 6.4

Cumulative frequency distribution of plaque size for λADbkd compared against λADbke and λAD16C (plaque area in mm^2).

The overall plaque sizes of all of the phages were somewhat larger than that detailed in Chapter 4. Although the exact reason for this variation is unknown, it should be unimportant since the plating behaviour was suitably controlled by the inclusion of λADbke, λAD15C, λAD16C. The base pair changes of the control λ phage had the same relative effect as those reported in Chapter 4 (compare Figure 4.2 with Figure 6.2).

Of the new λ phages, λADbkb and λADbkd had plaques of a similar size to λADbke. Yet the Murchie and Lilley (1989) equivalent of the palindrome in λADbkb (pIRbkb) extruded ~2.4× more slowly *in vitro* than pIRbke, and λADbkd has an extra GC sequence in the centre compared with λADbke. λADbkc had plaques of an identical size to λAD15C; switching of the central AT to TA had no effect upon the plaque size.

6.3 Discussion

The Effect of Disrupting the Alternating Purine/Pyrimidine Sequence at the Centre is Complex

Chapter 4 confirmed that the comparable λ *pal* phage to pIRbke15C, λAD15C has small plaques *in vivo*, corresponding to increased formation of an unusual DNA secondary structure *in vivo*. Murchie and Lilley (1989) proposed that the palindrome of pIRbke15C extrudes rapidly due to the alternating purine/pyrimidine sequence reducing DNA stability at the centre of the palindrome. The results of the plaque assay for λADbkb may be consistent with the hypothesis of Murchie and Lilley (1989), since λADbkb has larger plaques than λAD15C (see Figure 6.2). However, Blommers *et al.* (1989, 1991) reported that PyTTPu sequences at the centre of a palindrome may form two residue loops *in vitro*, and that the first T residue *may* be replaced by a purine without disrupting the extrastable loop. Unfortu-nately, since this partial consensus sequence is disrupted in λADbkb (compared with λAD15C), it is possible that λAD15C has small plaques because the DNA loop of a secondary structure *in vivo* is of two residues only.

However, another factor to consider is that a GATC site may be potentially methylated *in vivo*. Murchie and Lilley (1989) reported that the palindrome of pIRbkb (GGATCC centre) extrudes ~2.4x more slowly than pIRbke (GAATTC), but with N^6-methylation of the adenine in the sequence GATC, extrusion is 1.6x more rapid than pIRbke. λ replicates rapidly so that the GATC dam methylase sites maintain a partly methylated state (Allers, 1993). For the λADbkb construct *in vivo*, there will presumably be an equilibrium between the extrusion of N^6-methylated and non-methylated palindromes. Furthermore, the central sequence of DNA palindromes are undermethylated compared with normal λ DNA (Allers, 1993). The plaque area observed for λADbkb may reflect a balance between cruciform extrusion from unmethylated and methylated DNA. This possibly explains the reason that λADbkb plaques are a similar size to those of λADbke.

The Formation of Two-Residue Loops *in Vivo* May be More Resistant to Base Sequence Changes than *in Vitro*

λAD15C (CATG) has the PyATPu consensus and the other λ phage with a putative two residue loop (λAD15C16C) has the Py_ _Pu consensus sequence of Blommers *et al.* (1989, 1991). This chapter reports that λADbkc (GCTAGC ie _PyTAPu_) has plaques that are as small as those of λAD15C. Presuming that the plaque size differences are due to differences in the formation of a DNA secondary structure, then the conclusion is that λADbkc also has altered loop structure and stability *in vivo*, and possibly two residue loops. λADbkc differs from the consensus sequence of Blommers *et al.* (1989, 1991) in that it has a central TA, not AT. Taken with λAD15C16C (GCCGCC, _PyCGPu_) that also has relatively small plaques (compared with λAD16C), these results imply that the formation of two residue loops *in vivo* may be more resistant to base sequence changes than *in vitro*. This presumes that the effect observed on viability is due to altered loop structure. The alternating purine/pyrimidine structure of λAD15C is disrupted in λADbkc, implying that this does not contribute to the extreme phenotype *in vivo*.

The exact determinants for loop structure *in vivo* are explored in the following chapter.

The Influence of Melting in the Centre is Elusive

λADbkd (GTGCAC) was constructed to test the hypothesis that the effect on λ viability may be interpretable in terms of DNA melting for the central two base pairs. The analysis of λ *pal* phage with centres AT, TA, and CG is described in Chapter 4, where it emerged that CG-centred palindromes have larger plaques than those with AT or TA centres. The explanation may be that the thermal stability at the centre of the palindrome is raised for CG centres so that palindrome-mediated inviability is lessened. Unexpectedly, λADbkd has slightly smaller plaques than λADbke. Therefore, it is possible that the observed differences between AT, TA, CG, or GC centres may not be due to altered thermal stability, but could better be explained in terms of altered loop structure and stability.

However, it was necessary to make two paired changes (compared with λADbke) to construct λADbkd to avoid reforming a *Sac*I site. The central sequence of λADbkd is therefore GTGCAC ie _PyGCPu_. As will be described in Chapter 7, if the central sequence is Py_ _Pu in a palindrome (and as has already been demonstrated *in vitro*), it could critically determine the structure of the loop. It is possible that, whilst the GC at the centre of λADbkd may raise the thermal stability (compared with AT or TA), the enhanced viability may be offset by the lowered viability of a λ phage with a more stable or two residue cruciform loop.

CHAPTER 7:

Sequence Determinants of DNA Loop

Structure *in Vivo*

7.1 Introduction

The influence of base sequence on DNA hairpin loop structure and thermal stability *in vitro* is reviewed in Chapter 1. In Chapters 4 and 6, loop structure is discussed in relation to its effect on λ *pal* viability. Four DNA palindromes are tentatively mentioned as having two residue loops. The central four bases of these palindromes are: CATG, CCGG (both Chapter 4), CTAG, TGCA (both Chapter 6). A consensus common to these sequences is Py_ _Pu, agreeing in part with the PyTTPu sequence for two residue loops that Blommers *et al.* (1989, 1991) determined *in vitro*. Possibly, bulky purine residues do not disrupt the loop structure as greatly *in vivo* as was predicted *in vitro* so that the central two bases may be varied more.

In this chapter, the hypothesis that λ *pal* viability may be loop structure dependent is tested and the sequence determinants for two residue loops are identified. Thus, further evidence is accumulated for the existence of DNA secondary structure *in vivo* causing λ *pal* inviability. The plaque size assay is modified so that in effect, it is a probe for DNA loop structure *in vivo*. Following a similar strategy to Chapter 4, the central sequences used are identical to two sets that have previously been used to investigate DNA hairpin loop structure *in vitro* (Blommers *et al.*, 1989, 1991, see preceding paragraph; Howard *et al.*, 1991; Raghunathan *et al.*, 1991). Howard *et al.* (1991) and Raghunathan *et al.* (1991) determined that C_ _G loop sequences may have an extra base pair, in agreement with the Py_ _Pu consensus of Blommers *et al.* (1991), but that two purines in the centre may well be conformationally feasible. Blommers *et al.* (1989, 1991) suggested that if the central sequence is CAAG then a two residue loop may not be formed. This chapter asks whether the same sequences may form two residue loops *in vivo*.

In addition, the effect of orientation in λ of a palindrome with an asymmetric insert is analysed. Some controversy exists as to whether palindrome-mediated inviability and instability is lagging strand biased in DNA replication (see Chapter 1; Trinh and Sinden, 1991; Weston-Hafer and Berg, 1991b). This aspect is

investigated first since it is a prerequisite to the loop structure studies which involve palindromes with asymmetries.

7.2 Results

Construction of a Long Palindrome Inserted into λ in Two Different Orientations

A long palindrome was constructed in a series of independent λ clones, and the resultant phages named λAD-CTTG-1 through to λAD-CTTG-5. The central sequence of the λ palindrome (described in Figure 7.1) is a perfect palindrome except for the central two bases. Note that in Figure 7.1, a single strand only is shown. Blommers *et al.* (1989, 1991) reported that the sequence CTTG at the centre of a palindrome may form a two residue loop *in vitro*, but that CAAG may not (although no direct evidence was offered for the latter). Both of these sequences are present at the palindrome centre in λAD-CTTG. The strategy was to isolate λ phage with the full length palindrome in opposing orientations, such that if a strandedness to λ *pal* inviability exists, it would be evident as a difference in plaque size between λ with the palindrome in different orientations. This experiment is dependent upon the sequence CTTG possessing a two residue loop and CAAG a four residue loop. Five independent λ phage were isolated; the likelihood that at least one of the palindromes is in a different orientation is >96.5%.

Plating Behaviour of λAD-CTTG Series
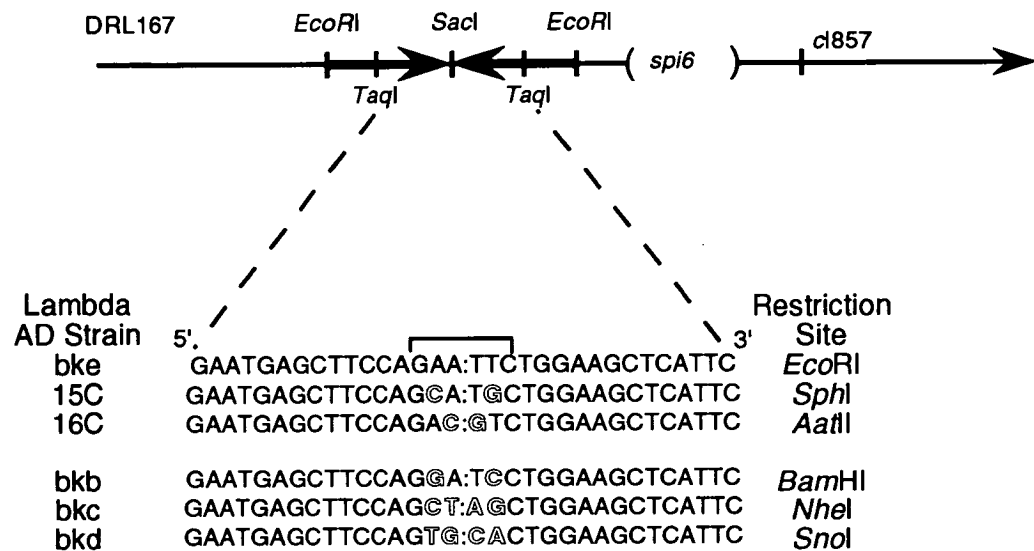
The phenotype conferred upon the λ by the palindrome was assessed by plating on a lawn of N2364 *sbcC* with CA plates. This has already been described in detail in Chapter 4 and the Materials and Methods. The analysis was varied in that eight plates were used for each independent λAD-CTTG phage. A graph of cumulative frequency plotted against plaque area is plotted for the λAD-CTTG phage in Figure 7.2 and the medians are detailed in Table 7.1.

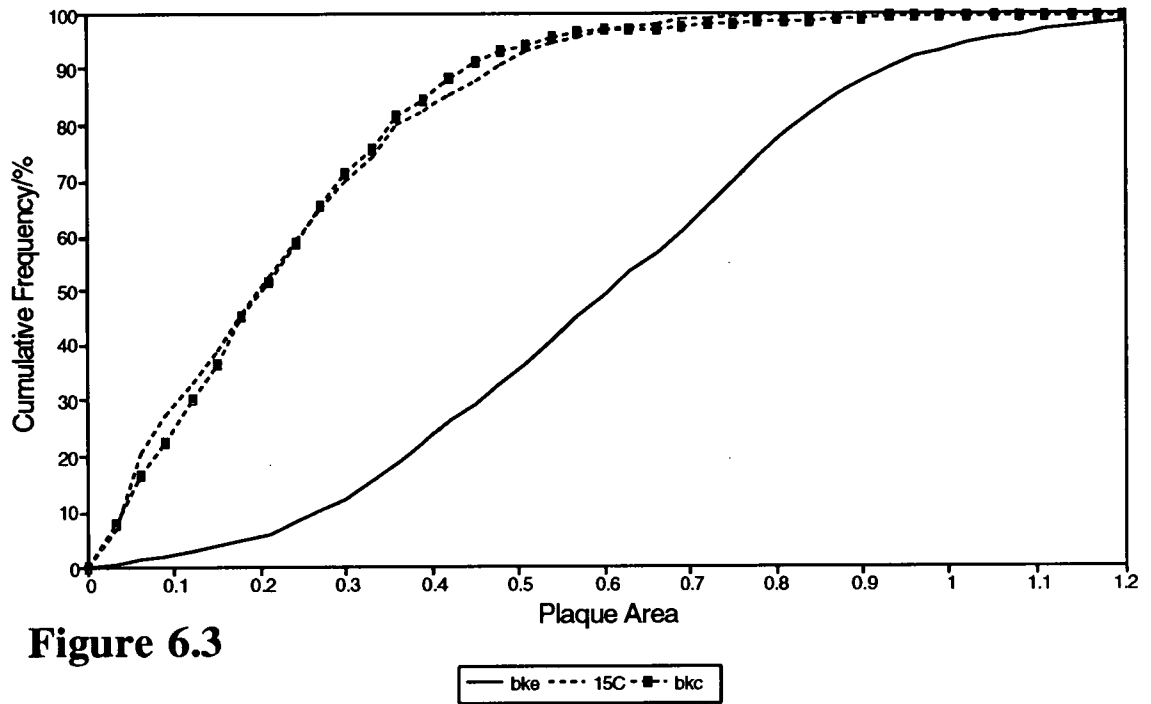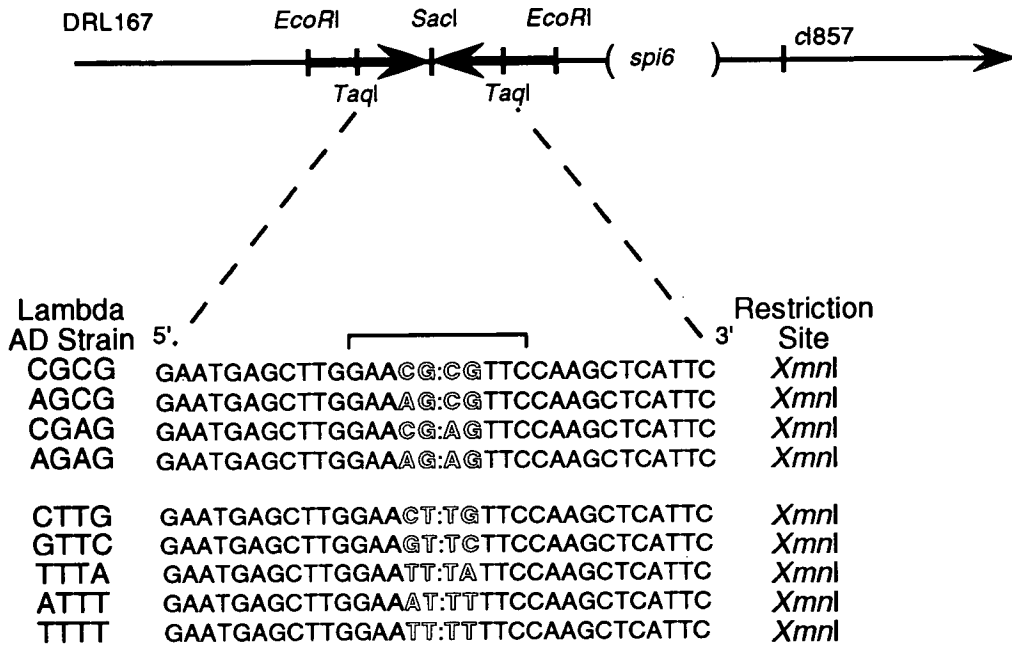**Figure 7.1**

Bacteriophage λ constructs. Top: Partial restriction map of λDRL167 (not to scale). Bottom: The structure and derivation of the palindromes used in this chapter. Differences between the phages are highlighted as open letters. All but one of the phage have a small asymmetric centre. λAD-CTTG was used to test the effect of palindrome orientation on the λ plaque size. The upper three sequences are derivatives of sequences used by Howard *et al.* (1991), and the fourth sequence is closely related. The lower five sequences are derivatives of sequences used by Blommers *et al.* (1987, 1989). All of the sequences have *Xmn*I restriction identification sites, GAANNNNTTC (N= any base).

110
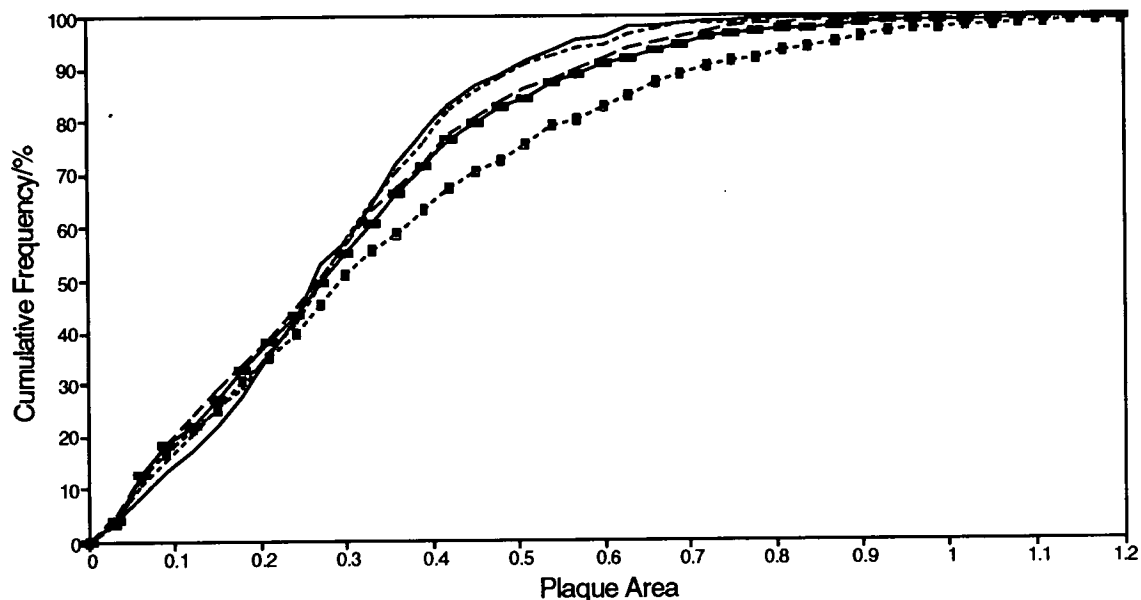
Figure 7.2
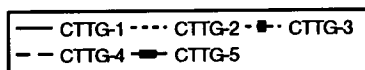
Cumulative frequency distribution of plaque sizes for five independently isolated λAD-CTTG constructs (plaque area in mm^2). The median plaque area of each λ is similar, although λAD-CTTG-3 has more plaques greater than ~0.5mm^2.

Table 7.1

Plaque assay data for λAD-CTTG phage.

Phage:	CTTG-1	CTTG-2	CTTG-3	CTTG-4	CTTG-5
Median Plaque Area (/mm^2)	0.26	0.27	0.30	0.27	0.27

111

Figure 7.2 and Table 7.1 show that the λ phage analysed have almost identical plaque size medians, but that λAD-CTTG-3 has a greater proportion of plaques larger than ~0.5mm². It is unlikely that this difference is significant, since the plaque size differences between the λ phage cover a relatively small 0.04mm². This is in contrast to the 0.53mm² range of plaque size differences that is reported in the second part of this results section. Possible explanations are addressed in the discussion. With regard to the following results section it is probably sufficient that the orientation of insertion of an asymmetric palindrome (at least for the sequence examined) does not have a large effect upon the plaque size.

Construction of Two Sets of Long DNA Palindromes with Small Asymmetries in the Central Sequence

Nine long palindromes, differing only in their central sequence were constructed in λ phage (including λAD-CTTG). One strand of the central six bases for each palindrome was identical to a member of two separate series of DNA sequences used to study the structure of DNA hairpin loops *in vitro* (Blommers *et al.*, 1989; Howard *et al.*, 1991). The bases further out were different due to constraints imposed by the cloning. The central sequences of the λ phage are described in Figure 7.1.

Sequences with the same central six bases to λAD-CGCG (the only perfect palindrome), λAD-AGCG, and λAD-CGAG were studied by Howard *et al.* (1991). Thermal transition (UV melting) and molecular modelling studies were interpreted to indicate that the hairpins with CGCG and CGAG centres have two base loops, and the AGCG centre a four base loop. The extra sequence λAD-AGAG was also included in this study.

The central palindromic sequences of λAD-CTTG, λAD-GTTC, λAD-TTTA, λAD-ATTT, and λAD-TTTT were studied by Blommers *et al.* (1989). In this case, NMR and UV melting experiments were interpreted so that CTTG and TTTA centres (ie Py_ _Pu) may have a two residue loop, but that GTTC and ATTT centres (ie Pu_ _Py) have four base loops. Earlier studies showed that the centre TTTT is extrastable due to stacking between the unpaired T residues (also between the T residues and the stem; Senior *et al.*, 1988; Paner *et al.*, 1988) or due to the formation of a wobble

T•T pair between the terminal thymidines of the loop (Haasnoot *et al.*, 1986; Blommers *et al.*, 1991).

Plating Behaviour of Two Sets of Long DNA Palindromes in λ with Small Asymmetries

The phenotype conferred upon the λ by the palindrome was assessed by plating on a lawn of N2364 *sbcC* with CA plates. This has already been described in detail in Chapter 4 and the Materials and Methods. Eight plates were analysed for each λ phage. Graphs of cumulative frequency plotted against plaque area are plotted for the two sets of λ phage in Figure 7.3 and 7.4 and the medians are detailed in Table 7.2 and 7.3. The data are reproduced in Appendix 1 alongside further control λ from Chapter 4 that are not described here. The difference in plaque area between λAD-CTTG in Table 7.1 and Table 7.3 may simply be explained due to differences arising from plating on different days, and consequent critical differences in the plates (eg moisture content, exact temperature of incubation, etc.). Each experiment was internally controlled so this factor should be irrelevant. Additionally, the platings were made alongside better characterised λ phage (ie those used in Chapter 4; data not presented here, but reproduced in Appendix 1)

Figure 7.3 and Table 7.2 together show that λAD-CGCG and λAD-CGAG have much smaller plaques than λAD-AGCG and λAD-AGAG. Figure 7.4 and Table 7.3 show again that the plaque data may be divided into two groups of large and small plaque-forming phage. λAD-CTTG and λAD-TTTA have much smaller plaques than λAD-GTTC and λAD-ATTT, respectively. Whilst uncertainties about the orientation of insertion of the palindromes into λ and the effect of the other strand exist, it is probably significant that the four λ phage with palindromes predicted to have two base loops *in vivo* both plated with relatively small plaques. The relatively large plaque area of λAD-TTTT is perhaps the only result that is contrary. The significance of both of these sets of results is addressed in the following discussion.
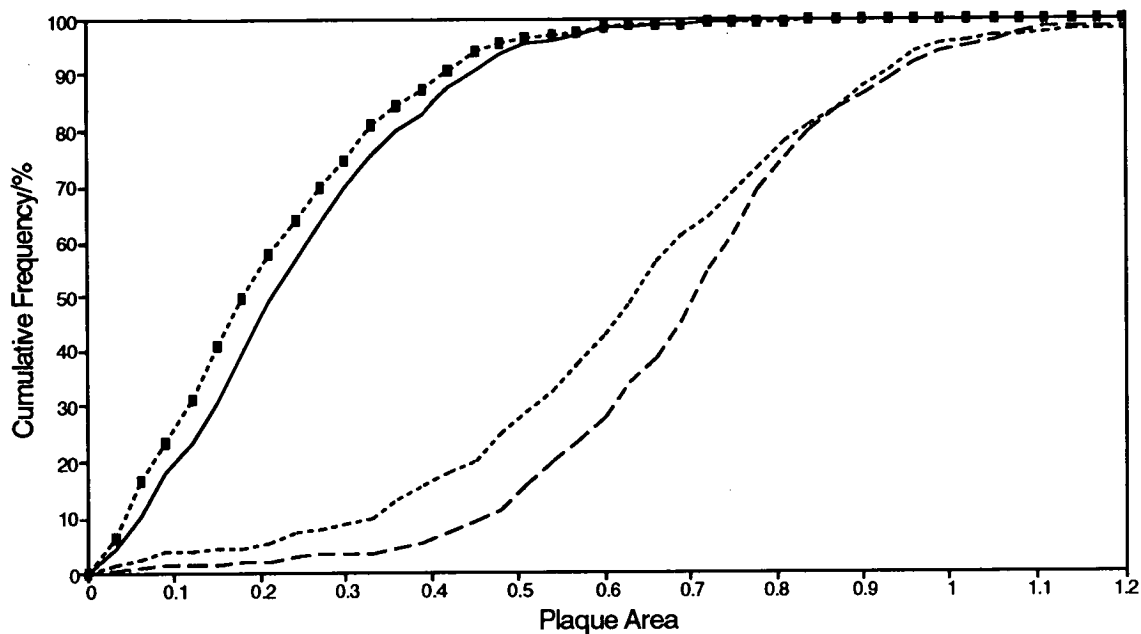
113

Figure 7.3

Cumulative frequency distribution of plaque sizes for four λAD- constructs (sequences based on Howard *et al.*, 1991; plaque area in mm^2). Two distinct separations in plaque size are visible between λAD-CGCG, λAD-CGAG (small plaques) and λAD-AGCG, λAD-AGAG (large plaques).

Table 7.2

Plaque assay data for λAD derivatives of sequences studied by Howard *et al.* (1991).

Phage:	CGCG	AGCG	CGAG	AGAG
Median Plaque Area (/mm²)	0.22	0.63	0.18	0.71
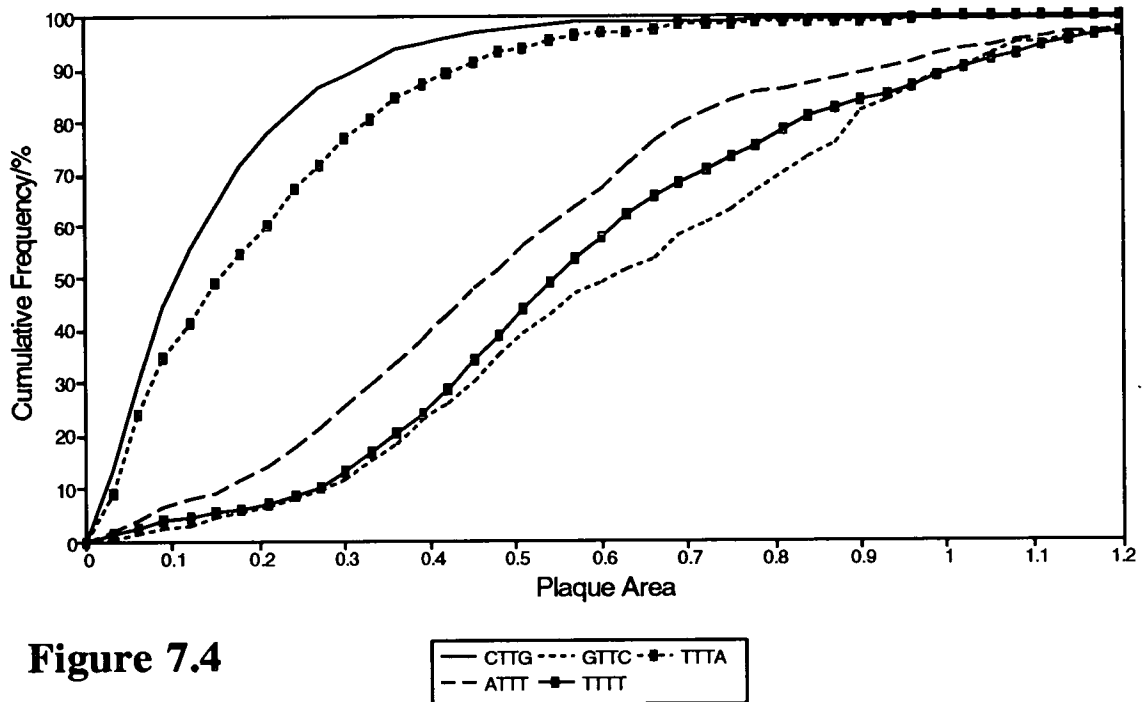
Figure 7.4

Legend: CTTG ···· GTTC ·■· TTTA ── ATTT ─■─ TTTT

Cumulative frequency distribution of plaque sizes for five λAD- constructs (sequences based on Blommers *et al.*, 1987, 1989; plaque area is in mm^2). Two distinct separations in plaque size are visible between λAD-CTTG, λAD-TTTA, (small plaques) and λAD-GTTC, λAD-ATTT, λAD-TTTT (large plaques).

Table 7.3

Plaque assay data for λAD derivatives of sequences studied by Blommers *et al.* (1987, 1989).

Phage:	CTTG	GTTC	TTTA	ATTT	TTTT
Median Plaque Area (/mm^2)	0.10	0.61	0.16	0.47	0.55

115

7.1 Discussion

The Orientation of Insertion of an Asmmetric Palindrome in λ has No Discernable Effect

A long DNA palindrome with a small central asymmetry was cloned in λ in two different orientations. The central sequence of the palindrome was based on a similar sequence used by Blommers *et al.* (1989) to study hairpin loop structure. It was designed so that one strand of the palindrome may exist as a two base loop when extruded as a cruciform, and the other a four base loop. With this construct, it was predicted that if a strandedness to palindrome-mediated inviability exists, then the preferential catalysis of extrusion from one strand may affect viability to varying extents depending upon whether the loop on that strand (and thus ease of formation of the transition state) has two or four bases. Possible reasons for a strandedness and preferential catalysis will not be speculated on here.

There may be a number of reasons as to why no strandedness was observed in this experiment and the results would only have been unambiguous if a strandedness was observed. The most obvious reason is that no strandedness to inviability exists in λ, but also in the light of other experiments described, it is possible that both of the strands (ie CTTG and CAAG) form two residue loops *in vivo* or at least have some two residue character. In Chapters 4 and 6, reasonable evidence is presented that the similar sequences CATG and CTAG form two base loops, and Raghunathan *et al.* (1991) have indicated that loops with two purines are conformationally feasible. A further possibility is the small chance (<3.5%) that all of the λ phage isolated had palindromes in the same orientation.

A Correlation Exists Between Sequences that have Two Base Hairpin Loops *in Vitro* and the Plaque Size of the λ *pal* Counterpart

The structure of certain DNA hairpin loops has been reasonably well studied *in vitro*, and the subject is reviewed in the introduction to this thesis. Almost nothing is known about the same DNA loop structures *in vivo*, and how they may differ. Previous chapters have described an assay for λ viability, and its use to measure the

effect of paired nucleotide sequence changes on DNA secondary structure formation *in vivo*. In this chapter the same assay is used as an indirect measure for the effect of small nucleotide changes at the centre of a palindrome on DNA loop structure *in vivo*.

Nine palindromic sequences were constructed in λ with identical centres to two sets that had previously been used to investigate DNA hairpin loop structure *in vivo* (Blommers *et al.*, 1989; Howard *et al.*, 1991). The major difference is that the palindromes used here are much longer and, more importantly, double stranded, so that the combined effect on viability of both strands was measured. Since all but one of the palindromes used is asymmetric, the effect of orientation of the λAD-CTTG palindrome in λ was measured and interpreted as probably not significant (see preceding discussion).

The results show that the four λ *pal* equivalents that are proposed to have two base loops *in vitro*, all have small plaques on an *E. coli sbcC* lawn. This correlates with increased DNA secondary structure formation *in vivo* and thus greater levels of inviability. The decrease in viability is presumably because the transition state bubble or protocruciform is more easily reached. Therefore, λAD-CGCG has smaller plaques than λAD-AGCG, despite a theoretically lower melting temperature for the palindrome centre for the latter. Since the effect of the additional A•T pair in λAD-AGCG (compared with λAD-CGCG) did not decrease the plaque size, the greater influence is presumably upon altered loop structure and stability. However, it is possible that a base mispair may form for the sequence in λAD-AGCG, but that it is less stable than the extra base pair in λAD-CGCG. Two possible reasons have been proposed for the extra stability of the central sequence TTTT *in vitro*. λAD-TTTT was the only phage with unexpectedly large plaques.

The probability that the data sets of Blommers *et al.* (1989), Howard *et al.* (1991), and the plaque assay results would correlate without an underlying reason is slight. It is likely that the correlation is evidence that sequences that may have two base loops *in vitro* also have two base loops *in vivo*. However, the relative contribution of the opposing palindrome strand on inviability (not present in the DNA hairpin loop studies) is undetermined in this assay, and this is discussed in the following part. In conclusion, this chapter presents evidence (although indirect) for

117

the formation of two residue DNA loops *in vivo*, and is further evidence that an unusual DNA secondary structure causes palindrome-mediated inviability in λ.

Sequence Determinants of Loop Structure *in Vivo*

It is likely that three palindromic sequences already described in this thesis may have two base loops *in vivo*. The central six bases of these perfect palindromes are: GCATGC, GCTAGC, and GCCGGC. Additionally, a two base loop may explain the unexpectedly small plaques of a λ phage with the sequence GTGCAC at the palindrome centre. In this chapter, four further sequences are described as possibly forming two base loops *in vivo*: ACGCGT, ACGAGT, ACTTGT, and ATTTAT. Common to each one of these sequences is the Py_ _Pu motif. It has already been established *in vitro* that a Py_ _Pu palindrome sequence is likely to form a two residue loop (Blommers *et al.*, 1989, 1991; Howard *et al.*, 1991; Raghunathan *et al.*, 1991). The exact requirements for the outer bases were not fully tested by Blommers *et al.* (1989, 1991), nor in this chapter. It is possible that the continuation of stacking in the stem is critical to loop structure, and that a more complete consensus will be PuPy_ _PuPy. The possibility remains that the alternating purine/pyrimidine structure is responsible for the decreased viability of some of these phage via a decreased thermal stability in the centre of the palindrome. However, the experiments of Chapters 4 and 6 have indicated that λ inviability may be primarily determined by the strength of intrastrand base pairing, not thermostability, and also that disrupting the alternating purine/pyrimidine structure present in λAD15C (GCATGC) to that of λADbkc (GCTAGC) had no effect upon viability.

The central two bases of these putative two base loop sequences are: AT, TA, CG, GC, GA (or TC on the other strand), and TT (or AA). That such a wide group of sequences did not effect the apparent loop size *in vivo*, as judged by the viability assay, is evidence that the central two bases may be varied more than studies *in vitro* have previously indicated without inhibiting two base loop formation. Thus, if the central two bases of a palindrome are relatively unimportant to loop structure then the greater effect of varying them may well be on altered thermal stability, as discussed in Chapter 4.

118

The flexibility of the central two bases in terms of loop structure also provides a possible explanation for the relative failure of the assay to detect differences in viability induced by the orientation of an asymmetric palindrome. It is proposed that the sequence CGAG at a palindrome centre has a two residue loop *in vitro* despite two purines in the central positions (Howard *et al.*, 1991; Raghunathan *et al.*, 1991). Therefore, it is possible that the palindrome of λAD-CTTG, used for the orientation assay, has two base loops on both strands (CTTG and CAAG centres) *in vivo*.

In conclusion, this chapter describes evidence for the effect of altered loop structure *in vivo* on λ *pal* viability, and is thus further evidence for the mechanism by which long DNA palindromes cause inviability.

CHAPTER 8:

Sequencing DNA Palindromes

8.1 Introduction

Sequencing DNA palindromes is problematic due to the tendency of the single stranded DNA to form a hairpin, thus preventing DNA polymerase entry, or removing the DNA polymerase from the extension reaction. The accuracy of the results in this thesis depend upon confirming the sequence at the centre of the palindromes. This chapter describes how this was achieved.

8.2 Results

Palindrome Length and Centre Restriction Site

Samples of each λ *pal* DNA were restricted with *Eco*RI, Klenow labelled with ^{35}S, and half of the sample retained. The remaining half of the sample was then restricted with the relevant enzyme for the centre of the palindrome. All samples were then run on a polyacrylamide gel alongside suitable controls to confirm the size of the palindromes and the restriction site at the centre. The restriction map of all of the λ phage was as predicted.

Subcloning and Sequencing of DNA Palindromes

The centres of λAD1, λAD2, λADbke, λAD16T, λAD15C, λAD16C, λAD15C16C, λAD13G, λAD11T, and λAD10G were subcloned and sequenced to confirm that they were as predicted. The palindrome was gel purified after excision with *Eco*RI (or *Ava*I and *Pvu*I for the palindromes with an *Eco*RI centre). A 32bp *Taq*I fragment containing the palindrome centre was then ligated into the *Acc*I site of pMS2B (a derivative of pUC18; D. Leach, M. Shaw, C. Blake). It is illustrated in Figure 4.1b. The sequences of all of the plasmids with inserts were as predicted, with no multiple inserts being detected.

Direct Sequencing Methods

In an attempt to bypass the above method for sequencing the palindrome centres, several more direct methods were attempted. Direct, double strand λ sequencing with Sequenase® v2.0 gave readable λ sequence but terminated after reading 50-100bp into the DNA palindrome. The TAQUENCE® double strand sequencing kit and a GIBCO BRL double strand cycle sequencing kit (extension reactions at >70°C) also failed to sequence through the palindrome. Evidently, even at 70°C the palindrome remains as a DNA hairpin. Exonuclease III was used to attempt to remove one arm of the palindrome in a controlled manner, preventing hairpin formation, but without success.

8.3 Discussion

Sequencing the centre of the DNA palindromes was necessary to confirm the accuracy of the results described in this thesis. A method was devised whereby it was possible to subclone and sequence the centre of the palindrome. More direct methods of sequencing the entire palindrome failed. 8 out of the 22 palindromes constructed in this work were sequenced. The remainder were thoroughly checked by restriction analysis. Since the primary chance of error is in the initial synthesis of the oligonucleotide, a method is currently being devised to check the DNA sequence of oligonucleotides via mass spectrometry (T. Brown, Edinburgh).

CHAPTER 9:

Discussion

9.1 Summary

The intention of this thesis is to investigate the formation of unusual DNA secondary structures *in vivo*, and the mechanism by which they cause inviability. The results described argue that the existence of an unusual DNA secondary structure in *E. coli* causes palindrome-mediated inviability, and suggest that the structure may be DNA cruciform. However, if cruciform structures are responsible for inviability, the stability of intrastrand base pairing (as opposed to DNA melting) must be the critical determinant of their formation or stability. The results also suggest that DNA loop structure *in vivo* has an important effect on inviability.

In **Chapter 3**, an analysis of the effects of central asymmetry on the propagation of palindromic DNA across a broad range of *E. coli* K-12 host strains is described. Palindromes carrying small asymmetries confer a less severe phenotype than do perfect palindromes, and the severity of phenotype correlates inversely with the length of asymmetry. Furthermore, the effect is visible with an 8bp asymmetry, approximately equal to the size of the denatured bubble in S-type extrusion *in vitro*. Together, this evidence argues that a centre-dependent pathway for palindrome-mediated inviability exists that is independent of *E. coli* host strain and is consistent with cruciform extrusion *in vivo* for palindromes of average base composition (even with asymmetries).

The major part of this thesis is concerned with the development of a plaque size assay to measure the viability of palindrome-containing bacteriophage λ on an *E. coli sbcC* host. Under the standardised conditions used, plaque size is assumed to be an indirect product of the λ burst size and a measure of its viability. Using this assay the effects of nucleotide sequence changes on the formation of an unusual DNA secondary structure in *E. coli sbcC* are measured (**Chapters 4, 6, and 7**). Attempts to assay the effect of osmotic stress (**Chapter 5**) and orientation of insertion in λ (**Chapter 7**) are also described.

In **Chapter 4**, it is confirmed that palindrome-mediated inviability is centre-dependent in an *sbcC* host by measuring the relative plaque size of the carrier λ in relation to the effects of central nucleotide sequence changes in a palindrome. The

results argue that intrastrand pairing between the arms of a palindrome is critical in determining its effect on inviability, and are consistent with cruciform extrusion *in vivo*. The implication for cruciform extrusion in other *E. coli* hosts is discussed. The same data suggest that the structure of a DNA loop *in vivo* is sequence dependent, and in some cases the loop may consist of two residues only. Simultaneously, *in vitro* cruciform extrusion conditions are identified that partly model the plating of palindrome-containing λ phage on an *E. coli sbcC* lawn, providing further evidence that cruciform structures may be responsible for the measured *in vivo* effects.

Chapter 6 describes the investigation of the effect of further sequence changes on λ *pal* viability (plaque size). The influence of melting at the palindrome centre was difficult to isolate. However, it is reported that the ability of a sequence to form two residue DNA loops *in vivo* may be more resistant to base changes than *in vitro*. Therefore, the sequence requirements for two residue DNA loop structure *in vivo* were determined and are reported in **Chapter 7**.

The plaque size assay was modified so that it was a probe for DNA loop structure *in vivo* (**Chapter 7**). The central sequences used are identical to two sets that have previously been used to investigate DNA hairpin loop structure *in vitro*. It is reported that a correlation exists between sequences that may have putative two base hairpin loops *in vitro*, and the plaque size of the λ counterpart. In addition, a tentative consensus for the central sequence that may form two base loops *in vivo* is identified as Py_ _Pu, where the full sequence that may be required is PuPy_ _PuPy. These results provide further evidence for the existence of an unusual DNA secondary structure *in vivo*.

Various other results are reported. In **Chapter 5**, the effect of osmotic stress on the plaque size of palindrome-containing λ phage is described. However, λ phage lacking a palindrome were equally affected by osmotic stress so no firm conclusions are possible without further investigation. The orientation of insertion of an asymmetric palindrome in λ had no discernable effect (**Chapter 7**). In **Chapter 8**, a method for the sequencing of long DNA palindromes is described.

9.2 Discussion

The topics that are summarised above have already been addressed in detail in the discussion of each results chapter. The intention of this discussion is to formulate a hypothesis for the mechanism of palindrome-mediated inviability. The exact structure causing inviability is considered in detail first of all.

DNA Cruciforms or Hairpins?

Together, the results described in this thesis are relatively consistent, and invoke a hypothesis where it is an unusual DNA secondary structure that causes inviability. Throughout the text, the structure responsible is generally referred to as a DNA cruciform. However, it is possible that the structure responsible for inviability is the single strand equivalent, a DNA hairpin. The significant effect of intrastrand pairing is arguably strong evidence for a DNA hairpin causing inviability, and the loop structure studies may be interpreted either due to cruciforms or hairpins. Only the effect of the central two base pairs may be interpreted in terms of DNA melting having a significant effect on inviability. As will be discussed later, it is feasible that both DNA hairpins and cruciforms coexist in the cell. Interconverting pathways between cruciforms and hairpins *in vivo* may blur the effect on inviability so that the causative structure may not be isolated. Possibly, both cause inviability to varying degrees depending upon the levels of each present.

If cruciform extrusion *per se* is responsible for the reported inviability effects then some attempt must be made to explain the difference between the results that Murchie and Lilley (1987) obtained and the demonstrated importance of intrastrand pairing *in vivo* (Chapter 4). S-type cruciform extrusion *in vitro* is dependent upon the A/T (or G/C) content of the palindrome centre, and it is probable that the transition state for S-type cruciform extrusion is either the central base opening or the protocruciform structure (Murchie and Lilley, 1987). Murchie and Lilley (1989) empirically established that for the palindromes studied, a natural log relationship exists between predicted melting temperature of the central 10bp and extrusion half-time. However, notable exceptions were excluded from the calculation (pbke15C,

pbke11T), indicating that the temperature dependence was not interpreted as the sole influence. Therefore, the apparent discrepancy between the *in vitro* and *in vivo* data sets may be because the transition state for each reaction is subtly different.

The *in vitro* reactions were performed under salt conditions (~50mM NaCl), designed to represent a balance between facilitating initial opening (promoted by low ionic strength) and stabilising the four-way junction (Murchie and Lilley, 1987). Also, the ability of varied salt concentrations to alter *in vitro* extrusion when modelling the λ *pal* plating was confirmed in Chapter 4. Cruciform extrusion in *E. coli* does not necessarily occur under "optimal" salt conditions; the salt concentration is higher and more complex. Therefore, if the transition state is shifted towards the protocruciform *in vivo* then inviability may be determined by a contribution from intrastrand pairing and melting. Additionally cruciform extrusion *in vivo* may be catalysed, for instance by the waves of negative supercoiling induced by transcription, affecting the transition state for the reaction (Liu and Wang, 1987).

Implications for the Mechanism of Inviability Arising from a Methylation Protection Assay *in Vivo*

Allers (1993) describes the development of a non-intrusive assay for non-B-DNA structures *in vivo*. Dam methylase was used to probe for the formation of a methylation-resistant structure at the centre of long DNA palindrome in bacteriophage λ. Like the studies described in this thesis, the palindrome centres were varied and the effect determined. Four perfect palindromes plus a palindrome with a ten base central asymmetry were constructed in λ. The use of dam methylase constrained all of the palindromes used to a central GATC sequence.

The central sequence of these palindromes is illustrated in Table 9.1 alongside the median plaque area of the palindrome-containing λ phage and the percentage methylation at the palindrome centre. Methylation was measured against a series of internal control GATC sites and the plating procedure was as described in this thesis. Differences between the sequences (compared with the first) are highlighted in bold.

127

Table 9.1 Methylation at palindrome centre and plaque size of λ *pal* (data reproduced from Allers, 1993)

Phage/ Centre	Central Palindromic Sequence	Predicted[a] Thermal Stability (°C)	% Methylation[b] N2364 *sbcC*	Plaque Area (mm^2)
DRL176/Perfect	GTGGATCCAC	78.6	51.3	0.59
DRL177/Perfect	GTTGATCAAC	71.6	36.2	0.61
DRL178/Perfect	GTAGATCTAC	68.5	26.1	0.81
DRL179/Perfect	GTCGATCGAC	82.6	23.3	0.28
DRL180/Asymm.	GTGGATCTCA	-	40.8	1.85

a. Calculated from the mean of the dinucleotide melting stabilities of the central ten base pairs (Gotoh and Tagashira, 1981). Value for DRL180 not determined.

b. Mean methylation at a series of control sites was 69.1%.

Analysing the plaque size results first, the data are almost fully consistent with the results of this thesis, despite the complicating factor that all the centres are methylated to varying degrees *in vivo*. The λ *pal* with an asymmetry has by far the largest plaques. Also, DRL177 and DRL178 (greater A/T content in palindrome arms) have larger plaques than DRL176 (although only slightly larger for the former), confirming that intrastrand pairing in the palindrome determines the mechanism of inviability. Allers (1993) proposed that DRL179 has the smallest plaques because the sequence at the centre may form a two residue loop *in vivo*. Although, the central sequence differs from the Py_ _Pu consensus described in Chapter 7 (G_ _C ie Pu_ _Py), a combination of methylation and the particular stability of stacking in the stem may overcome the normal two base loop sequence requirement since the central six bases are CGATCG. Also, a similar central sequence studied by Xodo *et al.* (1988,

1991) shows a UV melting transition consistent with the formation of a two base hairpin loop.

The principal conclusion of the methylation data is that all of the palindromes examined are significantly undermethylated *in vivo* compared with the control sites, implying that an unusual DNA secondary structure may be responsible. However, even a cursory analysis reveals that the palindrome centres that cause the least viability problems do not have sequences that are least undermethylated *in vivo*, as might have been expected. Also, the methylation of the asymmetric centre was unexpectedly low, and intermediate between two λ phage with perfect palindromes. Three of the perfect palindromes show an undermethylation that may be inversely related to the melting temperature (DRL176, DRL177, DRL178), a relationship that may also include DRL179 if it is assumed that the structure causing undermethylation for this construct has a two base loop *in vivo*. Clearly, the differences between the undermethylation and plaque assay results require an explanation.

Both the plaque size assay and the undermethylation assay appear to be indicators for the degree of non-B-DNA secondary structure formation *in vivo*. Therefore, a central question is what determines the structure(s) that causes inviability? Since it is probable that the differences are a result of both assays reacting to a different structure, a comparison of the two is potentially more informative than each in isolation. Additionally, it may reveal intimate details of the mechanism for palindrome-mediated inviability. Hypotheses that may account for the discrepancy are covered in the following section.

The Structure Causing Inviability May Arise via a Series of "Metastable" Intermediates

A hypothesis for palindrome-mediated inviability is suggested where cruciform extrusion *in vivo* may involve two steps (or more), that occur via a series of metastable intermediates (Figure 9.3). The principal metastable intermediate may be either a melted bubble at the centre of the palindrome, or more likely a protocruciform. Since its formation is the initial step of cruciform extrusion, the limiting step will be DNA melting. Rapid and frequent DNA melting/protocruciform formation may occur, where the reaction is fully reversible and the protocruciform

129

is normally reabsorbed. Therefore, the melted bubble/protocruciform at the palindrome centre will be reflected in the high degree of undermethylation. Only when the protocruciform is relatively stable/persistent (as a result of intrastrand pairing) will full cruciform extrusion occur, causing inviability. A central prediction of this hypothesis is that short palindromes (less than ~20bp) will exhibit a relatively high degree of undermethylation that is DNA melting dependent. A further prediction is that the absolute level of extruded cruciforms in the cell will be relatively low. A number of lines of evidence may corroborate this hypothesis.

Two sources have recently postulated that either HU protein or IHF may bind to an intermediate in cruciform formation, preventing extrusion (Pontiggia *et al.*, 1993; Oppenheim *et al.*, 1993). HU protein binds with high affinity to extruded cruciforms and when present in sufficient concentration, interferes with their formation *in vitro* (Pontiggia *et al.*, 1993). HU also forms nucleosome-like structures with DNA which reduce the free energy available for extrusion, but this does not necessarily exclude the possibility that it binds to a metastable intermediate *in vivo*, preventing extrusion. Intracellular levels of HU are relatively low, and insufficient for all the DNA to be in nucleosome-like structures (Pettijohn and Hodges-Garcia, 1990), and local supercoiling levels may temporarily exceed normal constrained levels under the effect of transcription (Liu and Wang, 1987). Additionally, HU protein recognises a non-sequence specific element of a cruciform (the four-way junction?) so may mistake long DNA palindromes for REP elements *in vivo* (presuming that REP elements exist as cruciforms). REPs are probably involved in chromosomal structure and organisation, and may be sites for DNA gyrase and DNA polymerase I entry. It is tempting to speculate that if DNA strand breaks are introduced into long DNA palindromes for gyrase/pol I entry, then it is possible that the extreme persistence of the non-B-DNA structure (compared with REPs) causes their inviability through a nuclease action at the strand break. SbcCD may fulfil this function in wild-type *E. coli*. In *E. coli sbcC* the remaining inviability may be a result of slow replication caused by the cruciform, the action of SbcD (and residual SbcC activity?) or the reaction of other cruciform specific enzymes,.

The metastable intermediate hypothesis may even be extended to explain the relatively high degree of undermethylation of the λ phage with an asymmetric
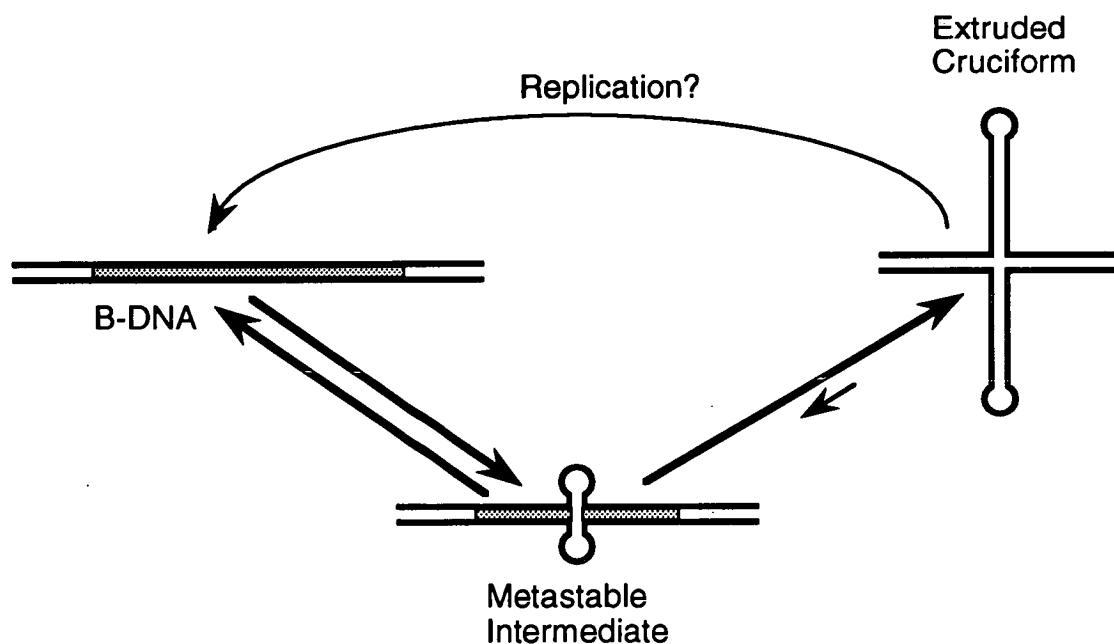
Figure 9.3

Proposed metastable intermediate hypothesis for cruciform formation *in vivo*. A single intermediate is pictured; other metastable intermediates such as the melted bubble may exist. It is proposed that once a full length cruciform is formed from a long DNA palindrome, the reaction is essentially irreversible. Cruciform structures may be removed by DNA replication, and in some cases the structure may be persistent enough to slow replication or cause inviability. The combined level of metastable intermediate will be relatively high. Full length cruciforms will be present at relatively low levels and palindromic B-DNA will make up the remainder.

palindrome. Despite the asymmetry, it is possible that the central ten bases, when melted, may form some intrastrand base pairs that are sufficient to allow the formation of a metastable intermediate. Therefore, formation of a protocruciform structure may be relatively rapid and frequent because the predicted melting temperature may be lower (the central ten bases of DRL180 are more A/T-rich than DRL176), but the hairpin is the least stable (or requires a longer stem before it is stabilised) so the associated inviability is partly alleviated.

Finally, a metastable intermediate hypothesis may help to explain the small discrepancies between the *in vitro* extrusion data of Murchie and Lilley (1987) and this thesis (Chapter 4). Murchie and Lilley (1987) primarily used SI nuclease for the extrusion assay, whereas for this thesis T4 endoVII was the enzyme of choice. Therefore, both experiments potentially measured different species. It is probable that SI nuclease could cleave at the DNA melted bubble, and loops of both a protocruciform and cruciform, whereas T4 endoVII cleaves the four-way junction present in the protocruciform/cruciform.

Several other hypotheses may explain palindrome-mediated viability. They are covered briefly in the following parts.

Hairpins and Cruciforms May Coexist *in Vivo*

The distinction between DNA cruciforms and hairpins *in vivo* will lie in the different pathways for their formation. Perhaps inviability results from hairpin formation in the replication fork only, where the products of intrastrand pairing are removed relatively rapidly. Palindromes with a large asymmetry form hairpins with a lower frequency, and so have lower levels of associated inviability. In this scheme undermethylation may be a result of cruciform extrusion (from B-DNA), where the cruciform is relatively persistent in the cell (compared with the hairpin). In this way, the effect on undermethylation will be DNA melting dependent. It is possible that if the asymmetry is sufficiently large, cruciform extrusion does not occur. In this hypothesis, it is difficult to conceive the reason that cruciforms do not also cause inviability (with the effect evident as a melting dependence), and also the reason that the asymmetric palindrome showed a relatively high level of undermethylation (Allers, 1993).

Nonetheless, a dual pathway hypothesis (where both pathways lead to inviability) may explain the influence that mutations in genes for homologous recombination have on inviability (see Chapter 3). If hairpin formation *in vivo* occurs via RecBCD-mediated unwinding, then in *E. coli rec*[+] all phage with long DNA palindromes are inviable as a result of cruciform extrusion and hairpin formation. As described in Chapter 3, DRL137, a phage with a 37bp asymmetry is just able to plate on *E. coli recB*. In this host, RecBCD-mediated unwinding is presumably prevented, and since the asymmetry is large, cruciform extrusion from B-DNA may barely occur. In an *E. coli recD* host, RecBCD-mediated unwinding may still occur and λ phage with asymmetric palindromes are able to plate. Possibly, this is because the primary pathway for inviability for these λ *pal* is via hairpin formation, and that homologous recombination is stimulated in *recD* hosts. Homologous recombination may be responsible for "rescuing" a proportion of the structure that would otherwise cause inviability. Alternatively, unwinding may be closely followed by rewinding in a *recD* host, tending to prevent hairpin formation. Therefore, on an *E. coli recA recD* host (where homologous recombination is prevented), most λ *pal* are unable to plate.

Perhaps homologous recombination that removes cruciform/hairpin structures may also have an influence on the degree of methylation *in vivo*? RecA may mistake single strand loops for the invading strand of normal recombination. This may be reflected as a balance between ease of access to the loop (large asymmetries, greatest access), and the ease of cruciform/hairpin formation. Alternatively, long palindrome sequences may have a "self-self" recognition as quadruplex DNA, influencing their methylation, but not their associated inviability. Many other intriguing hypotheses for the mechanism of inviability and undermethylation of palindrome centres are possible, all of which await investigation.

BIBLIOGRAPHY

Allers, T. (1993) *PhD Thesis* (University of Edinburgh).

Antao, V. P., S. Y. Lai, and I. Tinoco Jr (1991) *Nucleic Acids Res.* **19**: 5901-5905.

Behnke, D., H. Malke, M. Hartmann, and F. Walter (1979) *Plasmid* **2**: 605-616.

Benham, C. J. (1992) *J. Mol. Biol.* **225**: 835-847.

Bliska, J. B., and N. R. Cozzarelli (1987) *J. Mol. Biol.* **194**: 205-218.

Blommers, M. J. J., C. A. G. Haasnoot, C. W. Hilbers, J. H. van Boom, and G. A. van der Marel (1987) *NATO Asis Ser. E. Appl. Sci.* **133**: 78-91.

Blommers, M. J. J., F. J. M. van de Ven, G. A. van der Marel, J. H. van Boom, and C. W. Hilbers (1991) *Eur. J. Biochem.* **201**: 33-51.

Blommers, M. J. J., J. A. L. I. Walters, C. A. G. Haasnoot, J. M. A. Aelen, G. A. van der Marel, J. H. van Boom, and C. W. Hilbers (1989) *Biochemistry* **28**: 7491-7498.

Boccard, F., and P. Prentki (1993) *EMBO J.* **12**: 5019-5027.

Bowater, R., F. Aboul-ela, and D. M. J. Lilley (1991) *Biochemistry* **30**: 11495-11506.

Breslauer, K. J., R. Frank, H. Blöcker, and L. A. Marky (1986) *Proc. Natl. Acad. Sci. USA* **83**: 3746-3750.

Briki, F., R. Lavery, D. Genest, and J. Ramstein (1991a) *J. Chim. Phys.* **88**: 2567-2572.

Briki, F., J. Ramstein, R. Lavery, and D. Genest (1991b) *J. Am. Chem. Soc.* **113**: 2490-2493.

Broyles, S. S., and D. E. Pettijohn (1986) *J. Mol. Biol.* **187**: 47-60.

Chalker, A. F. (1990) *PhD Thesis* (University of Edinburgh).

Chalker, A. F., D. R. F. Leach, and R. G. Lloyd (1988) *Gene* **71**: 201-205.

Chen, S., F. Heffron, W. Leupin, and W. J. Chazin (1991) *Biochemistry* **30**: 766-771.

Cheong, C., G. Varani, and I. Tinoco Jr (1990) *Nature* **346**: 680-682.

Churchill, M. E. A., T. D. Tullius, N. R. Kallenbach, and N. C. Seemann (1988) *Proc. Natl. Acad. Sci. USA* **85**: 4653-4656.

Clegg, R. M., A. I. H. Murchie, A. Zechel, C. Carlberg, S. Diekmann, and D. M. J. Lilley (1992) *Biochemistry* **31**: 4846-4856.

Collins, J. (1981) *Cold Spring Harbor Symp. Quant. Biol.* **45**: 409-416.

Collins, J., and B. Hohn (1978) *Proc. Natl. Acad. Sci. USA* **75**: 4242-4246.

Collins, M., and R. R. Myers (1987) *J. Mol. Biol.* **198**: 737-744.

Collins, J., G. Volckaert, and P. Nevers (1982) *Gene* **19**: 139-146.

Cooper, J. P., and P. J. Hagerman (1987) *J. Mol. Biol.* **198**: 711-719.

Cooper, J. P., and P. J. Hagerman (1989) *Proc. Natl. Acad. Sci.* **86**: 7336-7340.

Courey, A. J., and J. C. Wang (1983) *Cell* **33**: 817-829.

Courey, A. J., and J. C. Wang (1988) *J. Mol. Biol.* **202**: 35-43.

Cunningham, R. P., and H. Berger (1977) *Virology* **80**: 67-82.

DasGupta, U., K. Weston-Hafer, and D. E. Berg (1987) *Genetics* **115**: 41-49.

Dayn, A., S. Malkhosyan, D. Duzhy, V. Lyamichev, Y. Panchenko, and S. Mirkin (1991) *J. Bacteriol.* **173**: 2658-2664.

Diekmann, S., and D. M. J. Lilley (1987) *Nucleic Acids Res.* **15**: 5765-5774.

Dimri, G. P., and H. K. Das (1990) *Nucleic Acids Res.* **18**: 151-156.

Doktycz, M. J., R. F. Goldstein, T. M. Paner, F. J. Gallo, and A. S. Benight (1992) *Biopolymers* **32**: 849-864.

Duckett, D. R., A. I. H. Murchie, A. Bhattacharyya, R. M. Clegg, S. Diekmann, E. von Kitzing, and D. M. J. Lilley (1992) *Eur. J. Biochem.* **207**: 285-295.

Duckett, D. R., A. I. H. Murchie, S. Diekmann, E. von Kitzing, B. Kemper, and D. M. J. Lilley (1988) *Cell* **55**: 79-89.

Duckett, D. R., A. I. H. Murchie, and D. M. J. Lilley (1990) *EMBO J.* **9**: 583-590.

Dunderdale, H. J., F. E. Benson, C. A. Parsons, G. J. Sharples, R. G. Lloyd, and S. C. West (1991) *Nature* **354**: 506-510.

Ebel, S., A. N. Lane, and T. Brown (1992) *Biochemistry* **31**: 12083-12086.

Egner, C., and D. E. Berg (1981) *Proc. Natl. Acad. Sci. USA* **78**: 459-463.

Elborough, K. M., and S. C. West (1990) *EMBO J.* **9**: 2931-2936.

Erie, D. A., A. K. Suri, K. J. Breslauer, R. A. Jones, and W. K. Olson (1992) *Biochemistry* **32**: 436-454.

Eshoo, M. W. (1988) *J. Bact.* **170**: 5208-5215.

Frank-Kamenetskii, M. D. (1990). *Cold Spring Harbor Monograph* **20**: 185-215. DNA Topology and Its Biological Effects, Eds, Cozzarelli, N. R. and J. C. Wang.

Gellert, M., K. Mizuuchi, M. H. O'Dea, H. Ohmori, and J. Tomizawa (1979) *Cold Spring Harbor Symp. Quant. Biol.* **43**: 35-40.

Gellert, M., M. H. O'Dea, and K. Mizuuchi (1983) *Proc. Natl. Acad. Sci. USA* **80**: 5545-5549.

Gibson, F. P., D. R. F. Leach, and R. G. Lloyd (1992) *J. Bacteriol.* **174**: 1222-1228.

Gilson, E., J. M. Clément, D. Perrin, and M. Hofnung (1987) *Trends Genet.* **3**: 226-230.

Gilson, E., W. Saurin, D. Perrin, S. Bachellier, and M. Hofnung (1991) *Nucleic Acids Res.* **19**: 1375-1383.

Gilson, E., D. Perrin, and M. Hofnung (1990) *Nucleic Acids Res.* **18**: 3941-3952.

Giradet, J. L., and J. Ramstein (1992) *Biochim. et Biophys. Acta* **1130**: 127-132.

Glickman, B. W., and L. S. Ripley (1984) *Proc. Natl. Acad. Sci. USA* **81**: 512-516.

Gorbalenya, A. E., and E. U. Koonin (1990) *J. Mol. Biol.* **213**: 513-591.

Gordenin, D. A., K. S. Lobachev, N. P. Degtyareva, A. L. Malkova, E. Perkins, and M. A. Resnick (1993) *Mol. Cell. Biol.* **13**: 5315-5322.

Gotoh, O., and Y. Tagashira (1981) *Biopolymers* **20**: 1033-1042.

Gough, G. W., and D. M. J. Lilley (1985) *Nature* **313**: 154-156.

Gough, G. W., K. M. Sullivan, and D. M. J. Lilley (1986) *EMBO J.* **5**: 191-196.

Gram, H., and W. Rüger (1985) *EMBO J.* **4**: 257-264.

Greaves, D. R., R. K. Patient, and D. M. J. Lilley (1985) *J. Mol. Biol.* **185**: 461-478.

*Haasnoot, C. A. G., M. J. J. Blommers, and C. W. Hilbers (1987) *Structure, Dynamics and Functions of Biomolecules, Springer Ser. Biophys.* **1**: 212-216.

Haasnoot, C. A. G., S. H. deBruin, C. W. Hilbers, G. A. van der Marel, and J. H. van Boom (1985) *Proc. Int. Symp. Biomol. Struct. Interact., Suppl. J. Biosci.* **8**: 767-780.

Haasnoot, C. A. G., S. H. deBruin, R. G. Berendsen, H. G. J. M. Janssen, T. J. J. Binnendijk, C. W. Hilbers, and G. A. van der Marel (1983) *J. Biomol. Struct. Dyn.* **1**: 115-129.

Haasnoot, C. A. G., C. W. Hilbers, G. A. van der Marel, J. H. van Boom, V. C. Singh, N. Pattabiraman, and P. A. Kollman (1986) *J. Biomol. Struct. Dyn.* **3**: 843-857.

Hagan, C. E., and G. J. Warren (1982) *Gene* **19**: 147-151.

Haniford, D. B., and D. E. Pulleybank (1985) *Nucleic Acids Res.* **13**: 4343-4363.

Henderson, S. T., and T. D. Petes (1993) *Genetics* **133**: 57-62.

Higgins, C. F., G. F. Ames, W. M. Barnes, J. M. Clement, and M. Hofnung (1982) *Nature* **298**: 760-762.

Higgins, C. F., C. J. Dorman, D. A. Stirling, L. Waddell, I. R. Booth, G. May, and E. Bremer (1988) *Cell* **52**: 569-584.

Hilbers, C. W., C. A. G. Haasnoot, S. H. deBruin, J. J. M. Joordens, G. A. van der Marel, and J. H. van Boom (1985) *Biochemie* **67**: 685-695.

Hirao, I., Y. Nishimura, Y. Tagawa, K. Watanabe, and K. Miura (1992) *Nucleic Acids Res.* **20**: 3891-3896.

Hirota, Y., S. Yasuda, M. Yamada, A. Nishimura, K. Sugimoto, H. Sugisaki, A. Oka, and M. Takanami (1979) *Cold Spring Harbor Symp. Quant. Biol.* **43**: 129-138.

Hobom, G., R. Grosschedl, M. Lusky, G. Scherer, E. Schwarz, and H. Kössel (1979) *Cold Spring Harbor Symp. Quant. Biol.* **43**: 165-178.

Hohn, B. (1979) *Methods Enzymol.* **68**: 299-309.

Holliday, R. (1964) *Genet. Res.* **5**: 282-304.

Howard, F. B., C. Chen, P. D. Ross, and H. Todd Miles (1991) *Biochemistry* **30**: 779-782.

Howard-Flanders, P., and L. Theriot (1966) *Genetics* **53**: 1137-1150.

Hsieh, L., J. Rouviere-Yaniv, and K. Drlica (1991) *J. Bacteriol.* **173**: 3914-3917.

Hulton, C. S. J., C. F. Higgins, and P. M. Sharp (1991) *Molecular Micro.* **5**: 825-834.

Hulton, C. S. J., A. Seirafi, J. C. D. Hinton, J. M. Sidebotham, L. Waddell, G. D. Pavitt, T. Owen-Hughes, A. Spassky, H. Buc, and C. F. Higgins (1990) *Cell* **63**: 631-642.

Hyrien, O., M. Debatisse, G. Buttin, and B. R. de Saint Vincent (1988) *EMBO J.* **7**: 407-417.

Ishiura, M., N. Hazumi, T. Koide, T. Uchida, and Y. Okada (1989) *J. Bacteriol.* **171**: 1068-1074.

James, J. K., and I. Tinoco Jr (1993) *Nucleic Acids Res.* **21**: 3287-3293.

Kaliman, A. V., V. M. Kryukov, and A. A. Bayev (1988) *Nucleic Acids Res.* **16**: 10353-10354.

Kallick, D. A., and D. E. Wemmer (1991) *Nucleic Acids Res.* **19**: 6041-6046.

Kazic, T., and D. E. Berg (1990) *Genetics* **126**: 17-24.

Kemper, B., M. Garabett, and U. Courage (1981) *Eur. J. Biochem.* **115**: 133-141.

Kleckner, N. (1989). *Mobile DNA*: 227-268. Eds, Berg, D. E. and M. M. Howe, (American Society
✳   for Microbiology, Washington, D. C.).

Kremer, E. J., M. Pritchard, M. Lynch, S. Yu, K. Holman, E. Baker, S. T. Warren, D. Schlessinger,
        G. R. Sutherland, and R. I. Richards (1991) *Science* **252**: 1711-1714.

Kulkarni, S. K., and F. W. Stahl (1989) *Genetics* **123**: 249-253.

Kutter, E. M., and J. S. Wiberg (1968) *J. Mol. Biol.* **38**: 395-411.

LaDuca, R. J., P. J. Fay, C. Chuang, C. S. McHenry, and R. A. Bambara (1983) *Biochemistry* **22**:
        5177-5188.

LaSpada, A. R., E. M. Wilson, D. B. Lubahn, A. E. Harding, and K. H. Fischbeck (1991) *Nature* **352**:
        77-79.

Leach, D. R. F., R. G. Lloyd, and A. F. W. Coulson (1992) *Genetica* **87**: 95-100.

Leach, D., and J. Lindsey (1986) *Mol. Gen. Genet.* **204**: 322-327.

Leach, D. R. F., and F. W. Stahl (1983) *Nature* **305**: 448-451.

Lederberg, J. (1955) *J. Cell. Comp. Physiol.* **45**: 75-107.

Lee, C. S., R. W. Davis, and N. Davidson (1970) *J. Mol. Biol* **48**: 1-22.

Lilley, D. M. J. (1980) *Proc. Natl. Acad. Sci. USA* **77**: 6468-6472.

Lilley, D. M. J. (1981) *Nature* **292**: 380-382.

Lilley, D. M. J. (1985) *Nucleic Acids Res.* **13**: 1443-1465.

Lilley, D. M. J. (1986) *Nature* **320**: 14-15.

Lilley, D. M. J., and L. R. Hallam (1984) *J. Mol. Biol.* **180**: 179-200.

Lilley, D. M. J., and A. F. Markham (1983) *EMBO J.* **2**: 527-533.

Lin, L., and R. J. Meyer (1987) *Nucleic Acids Res.* **15**: 8319-8331.

Lindsey, J. C., and D. R. F. Leach (1989) *J. Mol. Biol.* **206**: 779-782.

Liu, L. F., and J. C. Wang (1987) *Proc. Natl. Acad. Sci. USA* **84**: 7024-7027.

Lloyd, R. G., and C. Buckman (1985) *J. Bacteriol.* **164**: 836-844.

Lundblad, V., and N. Kleckner (1984) *Genetics* **109**: 3-19.

Lyamichev, V. I., I. G. Panyutin, and M. D. Frank-Kamenetskii (1983) *FEBS Lett.* **153**: 298-302.

Majumdar, R., and A. R. Thakur (1985) *Nucleic Acids Res.* **13**: 5883-5893.

Mandel, M., and A. Higa (1970) *J. Mol. Biol* **53**: 159-162.

Manfioletti, G., and C. Schneider (1988) *Nucleic Acids Res.* **16**: 2873-2884.

Massay, B. d., F. W. Studier, L. Dorgai, E. Appelbaum, and R. A. Weisberg (1985) *Cold Spring
        Harbor Symp. Quant. Biol.* **49**: 715-726.

McClellan, J. A., P. Boublíkova, E. Palecek, and D. M. J. Lilley (1990) *Proc. Natl. Acad. Sci. USA*
        87: 8373-8377.

Meijer, M., E. Beck, F. C. Hansen, H. E. N. Bergmans, W. Messer, K. von Meyenburg, and H.
        Schaller (1979) *Proc. Natl. Acad. Sci. USA* **76**: 580-584.

Miller, W. G., and R. W. Simons (1993) *Mol. Microbiol.* **10**: 675-684.

Mizuuchi, K., M. Mizuuchi, and M. Gellert (1982) *J. Mol. Biol* **156**: 229-243.

Müller, U. R., and W. M. Fitch (1982) *Nature* **298**: 582-585.

Müller, U. R., and M. A. Turnage (1986) *J. Mol. Biol.* **189**: 285-292.

Murchie, A. I. H., R. Bowater, F. Aboul-ela, and D. M. J. Lilley (1992) *Biochimica et Biophysica
✳   Acta* **1131**: 1-15.

Murchie, A. I. H., R. M. Clegg, E. von Kitzing, D. R. Duckett, S. Diekmann, and D. M. J. Lilley
✳   (1989) *Nature* **341**: 763-766.

Murchie, A. I. H., and D. M. J. Lilley (1989) *J. Mol. Biol.* **205**: 593-602.

Murchie, A. I. H., J. Portugal, and D. M. J. Lilley (1991) *EMBO J.* **10**: 713-718.

Nader, W. F., T. D. Edlind, A. Heuttermann, and H. W. Sauer (1985) *Proc. Natl. Acad. Sci. USA* **82**:
        2698-2702.

Naom, I. S., S. J. Morton, D. R. F. Leach, and R. G. Lloyd (1989) *Nucleic Acids Res.* **17**.

Naylor, L. H., D. M. J. Lilley, and J. H. van der Sande (1986) *EMBO J.* **5**: 2407-2413.

Ni Bhriain, N., C. J. Dorman, and C. F. Higgins (1989) *Mol. Microbiol.* **3**: 933-942.

Ohtsubo, H., and E. Ohtsubo (1978) *Proc. Natl. Acad. Sci. USA* **75**: 615-619.

Oka, A., N. Nomura, M. Morita, H. Sugisaki, K. Sugimoto, and M. Takanami (1979) *Molec. Gen. Genet.* **172**: 151-159.

Oppenheim, A. B., K. E. Rudd, I. Mendelson, and D. Teff (1993) *Mol. Microbiol.* **10**: 113-122.

Orbons, L. P. M., A. A. van Beuzekom, and C. Altona (1987) *J. Biomol. Struct. Dyn.* **4**: 965-987.

Orbons, L. P. M., G. A. van der Marel, J. H. van Boom, and C. Altona (1986) *Nucleic Acids Res.* **14**: 4187-4196.

Panayotatos, N., and A. Fontaine (1987) *J. Biol. Chem.* **262**: 11364-11368.

Panayotatos, N., and R. D. Wells (1981) *Nature* **289**: 466-470.

Paner, T. M., M. Amaratunga, M. J. Doktycz, and A. S. Benight (1990) *Biopolymers* **29**: 1715-1734.

Panyutin, I. G., and P. Hsieh (1993) *J. Mol. Biol.* **230**: 413-424.

* Peeter, B. P. H., J. H. de Boer, S. Bron, and G. Verona (1988) *Mol. Gen. Genet.* **212**: 450-458.

Pettijohn, D. E., and Y. Hodges-Garcia (1990) *The Bacterial Chromosome*: 241-245, Eds, Drlica, K. and M. Riley (American Society for Microbiology, Washington, D. C.).

Picksley, S. M., C. A. Parsons, B. Kemper, and S. C. West (1990) *J. Mol. Biol.* **212**: 723-735.

Pierce, J. C., D. Kong, and W. Masker (1991) *Nucleic Acids Res.* **19**: 3901-3905.

Pieters, J. M. L., E. deVroom, G. A. van der Marel, J. H. van Boom, and C. Altona (1989) *Eur. J. Biochem.* **184**: 415-425.

Pieters, J. M. L., E. deVroom, G. A. van der Marel, J. H. van Boom, T. M. G. Koning, R. Kaptein, and C. Altona (1990) *Biochemistry* **29**: 788-799.

Platt, J. R. (1955) *Proc. Natl. Acad. Sci. USA* **41**: 181-183.

Platt, T. (1986) *Ann. Rev. Biochem.* **55**: 339-372.

Pontiggia, A., A. Negri, M. Beltrame, and M. E. Bianchi (1993) *Mol. Microbiol.* **7**: 343-350.

Raghunathan, G., R. L. Jernigan, H. Todd Miles, and V. Sasisekharan (1991) *Biochemistry* **30**: 782-788.

Ramstein, J., and R. Lavery (1988) *Proc. Natl. Acad. Sci. USA* **85**: 7231-7235.

Ramstein, J., and R. Lavery (1990) *J. Biomol. Struct. Dynam.* **7**: 915-933.

Riley, M., and A. Anilionis (1978) *Ann. Rev. Microbiol.* **32**: 519-560.

Rinkel, L. J., and I. Tinoco Jr (1991) *Nucleic Acids Res.* **19**: 3695-3700.

Ripley, L. S. (1982) *Proc. Natl. Acad. Sci. USA* **79**: 4128-4132.

Ripley, L. S., and B. W. Glickman (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47**: 851-861.

Robinson, B. H., and N. C. Seeman (1987) *Biophys. J.* **51**: 611-626.

Ruskin, B., and G. R. Fink (1993) *Genetics* **133**: 43-56.

Saing, K. M., H. Orii, Y. Tanaka, K. Yanagisawa, A. Miura, and H. Ikeda (1988) *Mol. Gen. Genet.* **214**: 1-5.

Sanger, F., S. Nicklen, and A. R. Coulson (1977) *Proc. Natl. Acad. Sci. USA* **74**: 5463-5467.

Schaaper, R. M. (1989) *Genetics* **121**: 205-212.

Sen, S., A. Lahiri, and R. Majumdar (1992) *Biophys. Chem.* **42**: 229-234.

Senior, M. M., R. A. Jones, and K. J. Breslauer (1988) *Proc. Natl. Acad. Sci. USA* **85**: 6242-6246.

Sharples, G. J., and R. G. Lloyd (1990) *Nucleic Acids Res.* **18**: 6503-6508.

* Sharples, G. J., and R. G. Lloyd (1993) *Nucleic Acids. Res.* **21**: 2010.

Sigal, N., and B. Alberts (1972) *J. Mol. Biol.* **71**: 789-793.

Sinden, R. R., and D. E. Pettijohn (1984) *J. Biol. Chem.* **259**: 6593-6600.

Sinden, R. R., and R. D. Wells (1992) *Current Opinion in Biotechnology* **3**: 612-622.

Sinden, R. R., S. S. Broyles, and D. E. Pettijohn (1983) *Proc. Natl. Acad. Sci. USA* **80**: 1797-1801.

Sinden, R. R., G. Zheng, R. G. Brankamp, and K. N. Allen (1991) *Genetics* **129**: 991-1005.

Singleton, C. K., and R. D. Wells (1982) *J. Biol. Chem.* **257**: 6292-6295.

Stary, A., and A. Sarasin (1992) *Nucleic Acids Res.* **20**: 4269-4274.

Stern, M. J., E. Prossnitz, and G. F. Ames (1988) *Molec. Mirobiol.* **2**: 141-152.

Streisinger, G., Y. Okada, and J. Emrich (1966) *Cold Spring Harbor Symp. Quant. Biol.* **31**: 77-84.

Sullivan, K. N., and D. M. J. Lilley (1986) *Cell* **47**: 817-827.

Sullivan, K. M., and D. M. J. Lilley (1987) *J. Mol. Biol.* **193**: 397-404.

Sullivan, K. M., A. I. H. Murchie, and D. M. J. Lilley (1988) *J. Biol. Chem.* **263**: 13074-13082.

Taylor, A. F., and G. R. Smith (1990) *J. Mol. Biol.* **211**: 117-134.

Thompson, B. J., M. N. Camien, and R. C. Warner (1976) *Proc. Natl. Acad. Sci. USA* **73**: 2299-2303.

Trinh, T. Q., and R. R. Sinden (1991) *Nature* **352**: 544-547.

APPENDIX 1

Table 1

Numerical cumulative frequency data from Chapter 4.
Lambda AD clones bke to 10G

Clone:	bke	16T	15C	16C	15C16C	13G	11T	10G
Plaque Area (/mm2) Interval	Cumulative Frequency							
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.03	0.80	0.85	4.64	0.80	1.95	3.95	1.84	2.67
0.06	3.18	2.18	11.91	1.01	3.19	8.66	3.39	4.66
0.09	4.80	6.19	22.21	1.63	4.33	12.79	6.24	7.94
0.12	8.42	11.03	34.19	1.94	5.93	21.11	8.19	11.03
0.15	11.76	14.78	45.28	2.34	8.46	27.74	12.62	16.40
0.18	15.55	19.73	54.67	2.54	10.56	35.06	17.88	20.84
0.21	21.13	23.80	61.36	2.66	13.78	42.73	20.09	25.27
0.24	26.30	30.43	68.70	3.18	16.96	50.83	22.95	30.88
0.27	31.12	35.68	75.18	4.05	19.37	56.48	26.99	37.69
0.30	38.73	41.38	78.99	5.59	22.49	62.31	31.29	44.61
0.33	44.88	49.36	83.56	8.15	26.79	68.68	36.99	50.54
0.36	50.13	55.78	86.97	10.93	30.65	73.27	41.87	56.67
0.39	55.86	61.52	89.47	12.32	35.89	78.49	46.20	65.32
0.42	61.57	66.64	92.05	15.20	41.45	81.18	53.01	70.41
0.45	68.52	72.35	93.27	17.87	47.61	84.65	58.76	75.09
0.48	74.47	76.27	93.84	20.86	52.70	88.03	63.64	79.80
0.51	78.66	80.59	94.95	24.79	58.58	90.26	69.14	85.05
0.54	83.16	83.89	95.92	28.86	64.89	92.37	74.15	87.39
0.57	87.45	87.47	97.43	32.66	69.90	94.29	78.93	90.98
0.60	90.08	90.91	97.75	36.67	74.10	94.75	84.58	93.22
0.63	92.71	92.48	97.89	41.22	78.71	95.57	87.62	95.17
0.66	94.25	93.51	98.37	44.56	82.78	96.46	89.99	96.48
0.69	95.04	94.17	98.50	48.78	87.56	97.20	91.22	97.06
0.72	96.58	95.01	98.76	54.76	89.80	97.37	92.55	97.90
0.75	97.30	96.90	98.76	58.60	91.76	97.56	94.27	98.15
0.78	97.86	97.41	98.87	63.87	93.75	98.22	94.83	98.43
0.81	98.29	97.65	98.87	68.01	95.59	98.72	95.65	98.86
0.84	98.58	98.26	98.87	73.48	96.62	98.80	96.67	99.28
0.87	98.80	98.72	98.95	76.76	97.69	98.86	97.30	99.37
0.90	98.80	98.92	99.20	80.25	98.22	99.06	97.94	99.48
0.93	99.11	98.92	99.20	85.12	98.44	99.25	98.49	99.67
0.96	99.28	99.13	99.32	87.60	98.80	99.42	98.93	99.76
0.99	99.42	99.36	99.46	89.18	98.89	99.59	98.93	99.76
1.02	99.42	99.46	99.46	90.74	99.40	99.59	99.00	99.76
1.05	99.64	99.68	99.46	92.02	99.58	99.85	99.24	99.85
1.08	99.71	99.68	99.46	93.04	99.58	100.00	99.47	99.92
1.11	99.83	99.68	99.59	94.77	99.68	100.00	99.69	99.92
1.14	99.90	99.68	99.66	95.98	99.68	100.00	99.76	99.92
1.17	99.90	99.86	99.66	96.97	99.68	100.00	99.76	99.92
1.20	99.90	99.86	99.81	97.72	99.91	100.00	99.91	99.92
1.23	100.00	100.00	99.88	98.44	100.00	100.00	100.00	99.92
1.26	100.00	100.00	100.00	98.73	100.00	100.00	100.00	99.92
1.29	100.00	100.00	100.00	99.52	100.00	100.00	100.00	99.92
1.32	100.00	100.00	100.00	99.65	100.00	100.00	100.00	99.92
1.35	100.00	100.00	100.00	99.65	100.00	100.00	100.00	99.92
1.38	100.00	100.00	100.00	99.88	100.00	100.00	100.00	99.92
1.41	100.00	100.00	100.00	99.88	100.00	100.00	100.00	99.92
1.44	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.92
1.47	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.92
1.50	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.92
1.53	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.92
1.56	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table 2

Numerical cumulative frequency data from Chapter 6.
Lambda AD bkb to 16C

Clone: Plaque Area (/mm2) Interval	bkb	bkc	bkd	bke	15C	16C
		Cumulative Frequency				
0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.03	0.76	7.83	1.15	0.64	6.87	0.65
0.06	1.42	16.55	1.74	1.44	20.41	0.97
0.09	2.14	22.55	2.61	2.34	27.46	1.14
0.12	3.17	30.16	3.43	2.93	33.05	1.21
0.15	4.02	36.37	4.33	3.81	38.90	1.21
0.18	4.44	45.27	5.14	4.93	45.76	1.21
0.21	5.46	51.44	6.37	5.97	52.63	1.21
0.24	6.25	58.91	8.37	8.09	59.01	1.44
0.27	8.26	65.40	10.09	10.13	64.89	1.78
0.30	9.75	71.20	11.59	12.39	69.87	1.78
0.33	12.32	75.73	13.79	15.71	74.45	1.87
0.36	14.73	81.22	17.34	18.53	79.78	2.14
0.39	16.78	84.41	21.38	22.10	82.66	2.58
0.42	20.65	88.50	25.73	26.26	85.26	2.76
0.45	23.58	91.36	30.05	29.13	87.95	3.11
0.48	27.61	93.14	35.31	33.03	90.58	4.22
0.51	31.60	94.28	40.82	36.46	93.14	5.57
0.54	34.99	95.56	44.58	40.95	94.45	6.62
0.57	39.50	96.38	49.02	44.92	95.70	7.15
0.60	44.33	96.88	53.59	49.03	96.92	8.28
0.63	50.01	96.94	59.14	53.20	97.74	9.91
0.66	54.67	97.17	64.09	56.91	98.12	11.35
0.69	58.80	97.61	68.36	60.85	98.94	13.69
0.72	63.00	98.01	72.02	65.61	99.23	15.75
0.75	66.81	98.08	75.23	69.92	99.28	18.25
0.78	70.68	98.22	78.52	74.18	99.47	20.61
0.81	73.50	98.65	80.86	78.77	99.66	23.07
0.84	75.25	98.65	83.74	82.07	99.66	25.58
0.87	77.62	98.99	85.81	85.38	99.66	28.30
0.90	80.44	98.99	88.26	87.77	99.81	29.95
0.93	82.42	99.27	90.11	90.11	99.81	32.98
0.96	84.26	99.27	91.30	92.46	99.86	34.48
0.99	85.71	99.27	92.52	93.43	99.86	36.57
1.02	86.69	99.47	93.95	94.67	99.86	38.89
1.05	88.15	99.61	94.97	95.75	99.86	40.59
1.08	90.32	99.61	95.91	95.98	99.86	42.73
1.11	91.82	99.61	96.44	97.10	99.86	47.06
1.14	93.06	99.61	96.73	97.30	99.86	49.58
1.17	93.50	99.61	97.23	98.00	99.86	51.73
1.20	94.47	99.61	97.59	98.50	99.86	54.32
1.23	95.53	99.61	98.05	98.77	99.86	58.13
1.26	96.33	99.61	98.45	99.08	99.86	60.96
1.29	97.50	99.61	98.54	99.15	99.86	63.38
1.32	97.96	99.61	98.84	99.42	99.86	67.29
1.35	98.29	99.84	99.00	99.59	99.86	68.86
1.38	98.46	99.84	99.29	99.59	99.86	70.70
1.41	98.62	99.84	99.29	99.59	99.86	72.66
1.44	98.96	99.84	99.39	99.69	99.86	74.29
1.47	99.12	99.94	99.49	99.69	99.86	76.99
1.50	99.12	99.94	99.49	99.69	99.86	78.88
1.53	99.29	99.94	99.49	99.69	99.86	81.04
1.56	99.29	100.00	99.59	99.69	99.86	82.07
1.59	99.35	100.00	99.64	99.69	99.86	84.19
1.62	99.51	100.00	99.74	99.69	99.86	84.67
1.65	99.51	100.00	99.74	99.76	99.86	85.74
1.68	99.51	100.00	99.84	99.76	99.86	87.09
1.71	99.51	100.00	99.84	99.76	99.86	87.70
1.74	99.51	100.00	99.84	99.76	99.86	88.87
1.77	99.59	100.00	99.90	99.76	99.86	89.69
1.80	99.59	100.00	99.90	99.76	99.86	90.07
1.83	99.76	100.00	99.90	99.76	100.00	91.17
1.86	99.76	100.00	99.90	99.76	100.00	91.40
1.89	99.76	100.00	99.90	99.76	100.00	91.63
1.92	99.92	100.00	100.00	99.76	100.00	92.09
1.95	99.92	100.00	100.00	99.86	100.00	93.33
1.98	99.92	100.00	100.00	99.86	100.00	93.96
2.01	99.92	100.00	100.00	99.86	100.00	94.27

Table 2 (cont'd)

2.04	99.92	100.00	100.00	99.86	100.00	95.20
2.07	99.92	100.00	100.00	99.86	100.00	95.60
2.10	99.92	100.00	100.00	99.86	100.00	96.30
2.13	99.92	100.00	100.00	99.86	100.00	96.30
2.16	99.92	100.00	100.00	99.86	100.00	96.76
2.19	99.92	100.00	100.00	99.86	100.00	96.76
2.22	99.92	100.00	100.00	99.86	100.00	97.45
2.25	99.92	100.00	100.00	99.86	100.00	97.69
2.28	99.92	100.00	100.00	99.86	100.00	98.38
2.31	100.00	100.00	100.00	99.93	100.00	98.38
2.34	100.00	100.00	100.00	99.93	100.00	98.38
2.37	100.00	100.00	100.00	99.93	100.00	98.61
2.40	100.00	100.00	100.00	99.93	100.00	98.61

Table 3

Numerical cumulative frequency data from Chapter 7.
Lambda AD CTTG-1 to 5

Clone:	CTTG-1	CTTG-2	CTTG-3	CTTG-4	CTTG-5
Plaque Area (/mm2) Interval:		Cumulative Frequency			
0.00	0.00	0.00	0.00	0.00	0.00
0.03	4.06	4.57	3.53	4.91	3.86
0.06	8.46	10.60	12.47	12.71	12.55
0.09	13.26	15.17	16.99	18.42	18.05
0.12	17.09	20.73	21.36	24.11	22.21
0.15	21.76	25.56	25.16	29.30	26.92
0.18	27.81	29.16	30.44	33.70	32.61
0.21	35.66	36.19	35.15	38.85	38.06
0.24	42.67	41.84	39.82	44.96	43.37
0.27	53.35	49.54	45.25	50.42	49.65
0.30	56.90	57.68	50.95	56.52	54.99
0.33	64.04	64.51	55.40	62.96	60.86
0.36	72.16	70.49	58.77	67.53	66.04
0.39	78.00	75.93	63.43	72.14	71.30
0.42	83.07	81.92	67.15	77.26	76.24
0.45	86.63	85.47	70.23	80.73	79.43
0.48	88.93	88.43	72.65	83.53	82.76
0.51	91.50	90.85	75.35	86.10	84.28
0.54	93.32	92.50	78.88	87.88	86.98
0.57	95.27	93.59	80.15	89.68	88.90
0.60	96.03	94.45	82.70	91.62	90.91
0.63	97.71	96.44	84.55	93.86	91.85
0.66	97.85	97.69	87.17	94.75	93.31
0.69	98.33	98.25	88.57	95.79	94.50
0.72	98.81	98.30	90.38	96.74	95.99
0.75	99.16	98.55	91.39	97.80	96.39
0.78	99.23	99.02	92.00	98.36	97.11
0.81	99.51	99.38	93.12	98.47	97.25
0.84	99.51	99.64	93.71	98.72	97.68
0.87	99.58	99.69	94.95	98.98	98.05
0.90	99.58	99.69	96.07	98.98	98.43
0.93	99.65	99.69	96.87	98.98	98.83
0.96	99.65	99.69	97.19	98.98	98.83
0.99	99.72	99.69	97.66	99.19	98.83
1.02	99.79	99.69	97.82	99.28	98.83
1.05	99.79	99.69	97.99	99.74	98.83
1.08	100.00	99.69	98.62	99.79	99.26
1.11	100.00	99.90	98.62	99.79	99.26
1.14	100.00	99.90	98.94	99.79	99.26
1.17	100.00	99.90	99.11	99.79	99.26
1.20	100.00	99.90	99.11	99.79	99.26
1.23	100.00	99.90	99.27	99.79	99.32
1.26	100.00	99.90	99.27	99.79	99.32
1.29	100.00	99.90	99.27	99.79	99.40
1.32	100.00	99.90	99.27	99.79	99.46
1.35	100.00	99.90	99.39	99.79	99.46
1.38	100.00	99.90	99.56	99.79	99.46
1.41	100.00	99.90	99.72	99.79	99.46
1.44	100.00	99.90	99.72	99.79	99.60
1.47	100.00	99.90	99.72	99.79	99.60
1.50	100.00	99.90	99.86	99.79	99.60
1.53	100.00	99.90	99.86	100.00	99.60
1.56	100.00	99.90	99.86	100.00	99.60
1.59	100.00	99.90	100.00	100.00	99.60
1.62	100.00	100.00	100.00	100.00	99.60
1.65	100.00	100.00	100.00	100.00	99.68
1.68	100.00	100.00	100.00	100.00	99.68
1.71	100.00	100.00	100.00	100.00	99.68
1.74	100.00	100.00	100.00	100.00	99.68
1.77	100.00	100.00	100.00	100.00	99.68
1.80	100.00	100.00	100.00	100.00	99.68
1.83	100.00	100.00	100.00	100.00	99.68
1.86	100.00	100.00	100.00	100.00	99.68
1.89	100.00	100.00	100.00	100.00	99.68
1.92	100.00	100.00	100.00	100.00	100.00

Table 4

Numerical cumulative frequency data from Chapter 7.
Lambda AD-CGCG to AGAG.

Clone:	CGCG	AGCG	CGAG	AGAG	bke	15C
Plaque Area (/mm2) Interval	Cumulative Frequency					
0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.03	4.22	1.40	6.57	0.56	1.73	15.65
0.06	10.19	2.56	16.50	1.12	5.41	32.17
0.09	17.93	3.73	23.50	1.40	9.52	44.35
0.12	23.37	3.73	31.09	1.68	12.55	55.36
0.15	30.76	4.66	40.88	1.68	16.67	65.51
0.18	39.37	4.66	49.78	1.96	21.43	74.20
0.21	49.03	5.59	57.66	2.24	26.19	80.29
0.24	56.94	7.23	64.09	2.80	31.82	84.93
0.27	63.62	7.93	69.78	3.36	35.71	88.70
0.30	69.95	9.09	74.60	3.36	41.34	91.30
0.33	75.75	9.79	81.02	3.64	46.75	93.33
0.36	79.79	13.05	84.53	4.20	53.46	95.07
0.39	82.95	15.38	87.45	5.32	61.26	96.52
0.42	87.87	17.72	90.95	7.56	67.97	97.10
0.45	91.04	19.81	93.87	9.24	73.59	97.68
0.48	93.50	24.71	· 95.47	11.48	79.65	98.55
0.51	95.78	28.67	96.64	15.69	85.50	99.71
0.54	96.13	32.63	97.08	20.17	88.31	99.71
0.57	97.01	37.30	97.81	23.25	90.26	99.71
0.60	98.24	42.66	98.54	27.73	92.21	99.71
0.63	98.42	48.72	98.83	33.89	93.94	99.71
0.66	98.95	56.18	98.98	38.66	95.67	99.71
0.69	99.12	61.31	99.27	45.10	96.32	99.71
0.72	99.30	64.34	99.56	54.62	97.40	99.71
0.75	99.47	69.00	99.56	61.34	97.84	99.71
0.78	99.65	73.43	99.56	69.19	97.84	99.71
0.81	99.82	78.09	99.56	75.35	98.70	99.71
0.84	99.82	81.12	99.85	80.11	98.92	100.00
0.87	99.82	83.92	99.85	84.03	99.13	100.00
0.90	99.82	88.11	99.85	86.55	99.13	100.00
0.93	99.82	90.91	99.85	89.08	99.13	100.00
0.96	99.82	93.94	99.85	92.16	99.13	100.00
0.99	100.00	95.34	99.85	94.12	99.13	100.00
1.02	100.00	96.27	99.85	95.24	99.13	100.00
1.05	100.00	96.74	99.85	96.08	99.13	100.00
1.08	100.00	97.20	99.85	97.48	99.13	100.00
1.11	100.00	97.67	99.85	98.32	99.13	100.00
1.14	100.00	98.14	99.85	98.60	99.35	100.00
1.17	100.00	98.14	99.85	98.60	99.35	100.00
1.20	100.00	98.14	99.85	98.60	99.57	100.00
1.23	100.00	98.83	100.00	98.60	99.57	100.00
1.26	100.00	98.83	100.00	99.16	99.57	100.00
1.29	100.00	99.07	100.00	99.44	99.57	100.00
1.32	100.00	99.30	100.00	99.44	99.57	100.00
1.35	100.00	99.53	100.00	99.72	99.57	100.00
1.38	100.00	99.53	100.00	99.72	99.57	100.00
1.41	100.00	99.53	100.00	99.72	99.57	100.00
1.44	100.00	99.53	100.00	100.00	99.57	100.00
1.47	100.00	99.77	100.00	100.00	99.57	100.00
1.50	100.00	100.00	100.00	100.00	99.57	100.00
1.53	100.00	100.00	100.00	100.00	99.57	100.00
1.56	100.00	100.00	100.00	100.00	99.57	100.00
1.59	100.00	100.00	100.00	100.00	99.57	100.00
1.62	100.00	100.00	100.00	100.00	99.57	100.00
1.65	100.00	100.00	100.00	100.00	99.57	100.00
1.68	100.00	100.00	100.00	100.00	99.57	100.00
1.71	100.00	100.00	100.00	100.00	99.57	100.00
1.74	100.00	100.00	100.00	100.00	99.57	100.00
1.77	100.00	100.00	100.00	100.00	99.57	100.00
1.80	100.00	100.00	100.00	100.00	99.57	100.00
1.83	100.00	100.00	100.00	100.00	99.57	100.00
1.86	100.00	100.00	100.00	100.00	99.57	100.00
1.89	100.00	100.00	100.00	100.00	99.57	100.00
1.92	100.00	100.00	100.00	100.00	99.57	100.00
1.95	100.00	100.00	100.00	100.00	99.57	100.00
1.98	100.00	100.00	100.00	100.00	99.57	100.00
2.01	100.00	100.00	100.00	100.00	99.57	100.00

Table 4 (cont'd)

2.04	100.00	100.00	100.00	100.00	99.78	100.00
2.07	100.00	100.00	100.00	100.00	99.78	100.00
2.10	100.00	100.00	100.00	100.00	99.78	100.00
2.13	100.00	100.00	100.00	100.00	99.78	100.00
2.16	100.00	100.00	100.00	100.00	99.78	100.00
2.19	100.00	100.00	100.00	100.00	99.78	100.00
2.22	100.00	100.00	100.00	100.00	99.78	100.00
2.25	100.00	100.00	100.00	100.00	99.78	100.00
2.28	100.00	100.00	100.00	100.00	99.78	100.00
2.31	100.00	100.00	100.00	100.00	99.78	100.00
2.34	100.00	100.00	100.00	100.00	99.78	100.00
2.37	100.00	100.00	100.00	100.00	100.00	100.00

Table 5

Numerical cumulative frequency data from Chapter 7.
Lambda AD-CTTG to TTTT.

Clone: Plaque Area (/mm2) Interval	CTTG	GTTC	TTTA	ATTT	TTTT	bkg	1SC
		Cumulative Frequency					
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.03	13.89	0.69	9.21	2.17	1.57	1.27	12.74
0.06	29.66	1.66	24.04	4.34	2.79	3.11	23.89
0.09	44.53	2.42	34.49	6.67	3.94	5.17	32.29
0.12	55.39	3.11	41.28	7.93	4.71	6.68	38.66
0.15	63.58	4.46	48.83	9.37	5.56	8.74	45.47
0.18	71.64	5.68	54.81	11.71	6.21	11.12	51.55
0.21	78.10	6.52	60.08	14.57	7.20	13.50	56.62
0.24	82.80	8.19	67.01	17.88	8.85	16.52	61.54
0.27	86.75	9.80	71.98	21.56	10.53	18.67	66.60
0.30	88.87	11.93	76.96	25.69	13.12	21.89	69.64
0.33	91.14	15.04	80.53	29.45	16.87	25.00	72.68
0.36	93.97	18.52	84.59	33.86	20.72	28.76	77.02
0.39	94.89	22.83	87.28	38.52	24.02	33.66	79.19
0.42	95.64	26.09	89.35	42.92	28.56	38.23	81.79
0.45	96.71	30.26	91.37	48.14	34.33	41.05	84.10
0.48	97.48	35.18	93.15	51.74	38.73	45.70	87.71
0.51	97.86	39.11	93.94	55.95	43.94	50.04	91.18
0.54	98.24	43.07	95.39	59.90	48.92	53.67	92.34
0.57	98.70	46.87	96.34	63.57	53.74	55.86	94.07
0.60	98.70	48.75	96.79	67.44	57.40	58.65	95.23
0.63	98.92	51.33	97.12	71.78	62.32	61.54	96.39
0.66	99.07	53.71	97.66	76.21	65.53	64.43	96.96
0.69	99.07	58.16	98.11	79.74	68.22	67.39	98.12
0.72	99.07	60.87	98.23	81.90	70.98	70.97	98.70
0.75	99.07	63.45	98.60	84.25	73.49	74.63	98.70
0.78	99.38	66.92	98.69	85.53	75.30	77.46	98.99
0.81	99.38	70.40	98.69	86.07	78.47	80.12	99.28
0.84	99.38	73.42	98.93	87.35	81.05	82.66	99.42
0.87	99.38	76.17	98.93	88.44	82.62	85.60	99.42
0.90	99.38	81.78	99.05	89.35	84.16	87.42	99.71
0.93	99.38	84.17	99.18	90.53	84.90	89.45	99.71
0.96	99.38	86.55	99.55	91.54	86.71	91.47	99.71
0.99	99.38	89.08	99.67	92.90	88.85	92.28	99.71
1.02	99.38	90.55	99.75	93.72	90.12	93.70	99.71
1.05	99.38	92.65	99.75	94.35	92.11	94.51	99.71
1.08	99.38	94.75	99.75	95.26	92.78	94.51	99.71
1.11	99.38	94.96	99.75	95.90	94.06	95.72	99.71
1.14	99.38	95.59	99.75	96.63	95.26	96.03	99.71
1.17	99.38	97.06	99.75	97.00	96.25	96.84	99.71
1.20	99.38	97.69	99.75	97.18	96.85	97.76	99.71
1.23	99.69	98.11	99.75	97.54	97.32	98.16	99.71
1.26	99.69	98.53	99.75	97.91	97.59	98.77	99.71
1.29	99.69	98.53	99.75	98.09	97.59	98.77	99.71
1.32	99.69	98.53	99.75	98.27	97.99	99.18	99.71
1.35	99.69	98.53	99.75	98.82	98.59	99.38	99.71
1.38	99.69	98.74	99.75	98.82	98.79	99.38	99.71
1.41	99.69	98.74	99.75	99.00	98.79	99.38	99.71
1.44	99.69	99.16	99.88	99.00	98.79	99.58	99.71
1.47	99.69	99.58	99.88	99.00	98.99	99.58	99.71
1.50	99.69	99.58	99.88	99.00	98.99	99.58	99.71
1.53	99.69	99.58	99.88	99.00	98.99	99.58	99.71
1.56	99.69	99.58	99.88	99.19	99.19	99.58	99.71
1.59	99.69	99.58	99.88	99.19	99.39	99.58	99.71
1.62	99.69	99.79	99.88	99.37	99.39	99.58	99.71
1.65	99.69	99.79	99.88	99.46	99.46	99.58	99.71
1.68	99.69	99.79	99.88	99.46	99.66	99.58	99.71
1.71	99.69	99.79	99.88	99.46	99.66	99.58	99.71
1.74	99.69	99.79	99.88	99.64	99.66	99.58	99.71
1.77	99.69	99.79	99.88	99.64	99.73	99.58	99.71
1.80	100.00	99.79	99.88	99.64	99.73	99.58	99.71
1.83	100.00	99.79	99.88	99.64	99.80	99.58	100.00
1.86	100.00	99.79	99.88	99.64	99.80	99.58	100.00
1.89	100.00	99.79	99.88	99.64	99.80	99.58	100.00
1.92	100.00	99.79	99.88	99.64	99.80	99.58	100.00
1.95	100.00	99.79	99.88	99.64	99.80	99.78	100.00
1.98	100.00	99.79	99.88	99.73	99.80	99.78	100.00
2.01	100.00	99.79	99.88	99.73	99.80	99.78	100.00

Table 5 (cont'd)

2.04	100.00	100.00	99.88	99.73	99.80	99.89	100.00
2.07	100.00	100.00	99.88	99.73	99.80	99.89	100.00
2.10	100.00	100.00	99.88	99.73	99.80	99.89	100.00
2.13	100.00	100.00	99.88	99.73	99.80	99.89	100.00
2.16	100.00	100.00	99.88	99.73	99.80	99.89	100.00
2.19	100.00	100.00	99.88	99.73	99.80	99.89	100.00
2.22	100.00	100.00	99.88	99.73	99.80	99.89	100.00
2.25	100.00	100.00	99.88	99.73	99.80	99.89	100.00
2.28	100.00	100.00	99.88	99.73	99.80	99.89	100.00
2.31	100.00	100.00	99.88	99.73	99.80	99.89	100.00
2.34	100.00	100.00	99.88	99.73	99.80	99.89	100.00
2.37	100.00	100.00	99.88	99.73	99.80	100.00	100.00
2.40	100.00	100.00	99.88	99.73	99.80	100.00	100.00

APPENDIX 2

Cruciform extrusion *in vitro* is a first order reaction where the rate is proportional to the concentration of unextruded species. The rate of the reaction,

$$U \longrightarrow C$$
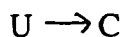
will be given by the first order rate equation:

$$\frac{-d[U]}{dt} = k[U]$$

where [U] is the molar concentration of U, $-d[U]/dt$ is the rate at which U disappears, and k is the rate constant. Integration results in the equation

$$\ln\left(\frac{[U_0]}{[U]}\right) = kt$$

where $[U_0]$ is the concentration of U at time $=0$, and [U] is the concentration at time t. Therefore, the half-time $(t_{1/2})$ of the reaction is

$$t_{1/2} = \frac{\ln 2}{k}$$

Relative extent of cruciform extrusion was calculated as the intensity of the band due to cleavage at the cruciform, divided by the total intensity. Rate constants were calculated using linear regression analysis of the gradient of $1+\ln(1\text{-extruded}$ fraction). A sample graph is illustrated overleaf. The kinetic data used for the calculations are also listed.

t = time (seconds) at which sample removed.

y = 1+ ln(1- proportion unextruded DNA).
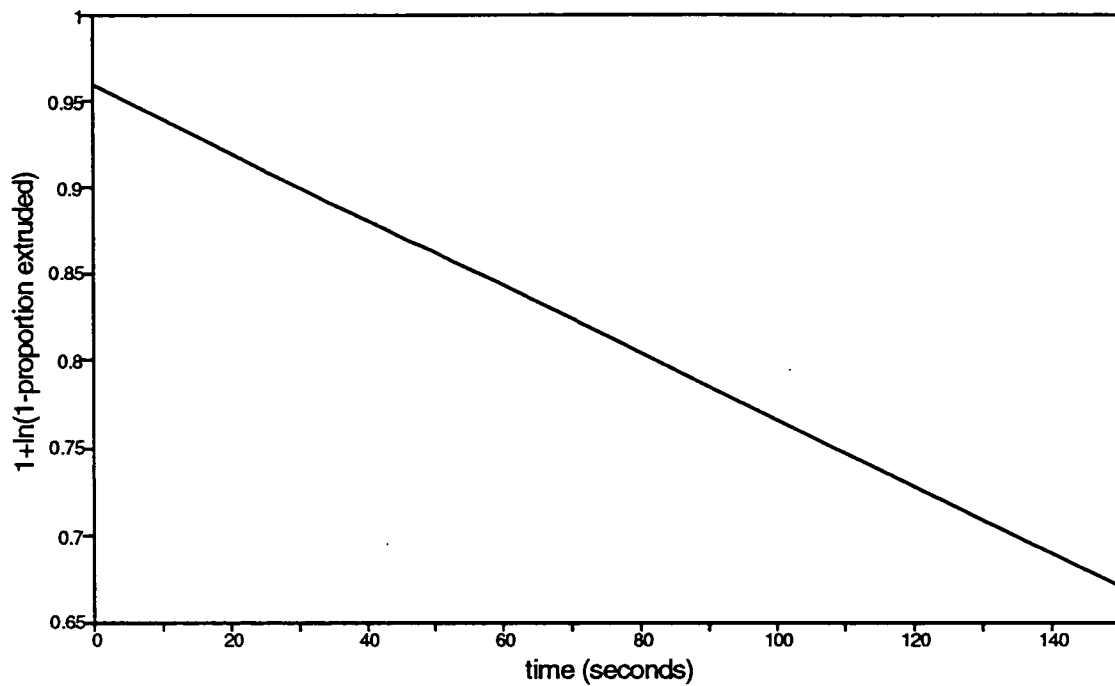
Figure A The kinetics of cruciform extrusion by pADbke, measured using T4 endoVII. Graph plotted using linear regression analysis to calculate the reaction constant.

Buffer: 50mMNaCl
Sequence: bke

t (sec)	y		Regression Output:	
30	0.91		Regression Output:	
60	0.83	Constant		0.9588
90	0.76	Std Err of Y Est		0.072565
120	0.77	R Squared		0.578007
150	0.73	No. of Observations		25
30	0.93	Degrees of Freedom		23
60	0.89			
90	0.81	X Coefficient(s)	-0.00192	
120	0.69	Std Err of Coef.	0.000342	
150	0.62	T1/2(min)	6.1	
30	0.94			
60	0.9			
90	0.84			
120	0.72			
150	0.74			
30	0.9			
60	0.93			
90	0.81			
120	0.78			
150	0.79			
30	0.83			
60	0.73			
90	0.63			
120	0.64			
150	0.53			

Buffer: 50mMNaCl
Sequence: 15C

t (sec)	y		Regression Output:	
30	0.78		Regression Output:	
60	0.73	Constant		0.781351
90	0.64	Std Err of Y Est		0.046024
120	0.56	R Squared		0.789266
150	0.49	No. of Observations		18
30	0.65	Degrees of Freedom		16
60	0.6			
90	0.54	X Coefficient(s)	-0.00201	
120	0.47	Std Err of Coef.	0.00026	
150	0.43	T1/2(min)	5.8	
30	0.72			
60	0.68			
90	0.62			
120	0.56			
30	0.7			
90	0.63			
120	0.57			
150	0.5			

Buffer: 50mMNaCl
Sequence: 16T

t (sec)	y		Regression Output:	
30	0.77		Regression Output:	
60	0.49	Constant		0.8076
90	0.4	Std Err of Y Est		0.086489
120	0.15	R Squared		0.811922
150	0.06	No. of Observations		24
30	0.62	Degrees of Freedom		22
60	0.53			
90	0.4	X Coefficient(s)	-0.00397	
120	0.24	Std Err of Coef.	0.000408	
150	0.18	T1/2(min)	2.9	
30	0.79			
60	0.56			
90	0.44			
120	0.41			
150	0.29			
30	0.75			
60	0.6			
90	0.47			
120	0.31			
150	0.35			
30	0.67			
60	0.57			
120	0.38			
150	0.37			

Buffer: 50mMNaCl
Sequence: 16C

t (sec)	y		Regression Output:	
600	0.96		Regression Output:	
3600	0.82	Constant		0.940004
7200	0.81	Std Err of Y Est		0.036503
10800	0.76	R Squared		0.873842
14400	0.73	No. of Observations		14
1800	0.95	Degrees of Freedom		12
3600	0.89			
7200	0.79	X Coefficient(s)	-1.9E-05	
10800	0.71	Std Err of Coef.	2.03E-06	
14400	0.63	T1/2(min)	608	
1800	0.92			
3600	0.82			
7200	0.8			
14400	0.69			

Buffer: 50mMNaCl
Sequence: 15C16C

t (sec)	y		Regression Output:	
600	0.98		Regression Output:	
3600	0.9	Constant		0.955122
7200	0.88	Std Err of Y Est		0.022692
10800	0.84	R Squared		0.875349
14400	0.8	No. of Observations		14
1800	0.95	Degrees of Freedom		12
3600	0.9			
7200	0.81	X Coefficient(s)	-1.2E-05	
10800	0.83	Std Err of Coef.	1.35E-06	
14400	0.77	T1/2((min)	963	
1800	0.93			
3600	0.9			
7200	0.86			
10800	0.81			

Buffer: 50mMNaCl
Sequence: 13G

t (sec)	y		Regression Output:	
		Constant		0.951139
30	0.97	Std Err of Y Est		0.087732
300	0.73	R Squared		0.724573
420	0.61	No. of Observations		15
30	0.95	Degrees of Freedom		13
180	0.87			
300	0.79	X Coefficient(s)	-0.00089	
420	0.73	Std Err of Coef.	0.000153	
30	0.93	T1/2(min)	13	
180	0.63			
300	0.6			
420	0.5			
30	0.93			
180	0.71			
300	0.67			
420	0.49			

Buffer: 50mMNaCl
Sequence: 11T

t (sec)	y	Regression Output:	
30	0.58		
60	0.37	Constant	0.766811
90	0.27	Std Err of Y Est	0.099489
120	0.01	R Squared	0.829171
150	0.01	No. of Observations	19
30	0.61	Degrees of Freedom	17
60	0.43		
90	0.21	X Coefficient(s)	-0.00483
120	0.11	Std Err of Coef.	0.000531
150	0.02	T1/2(min)	2.4
30	0.74		
60	0.46		
90	0.38		
120	0.24		
150	0.18		
30	0.79		
60	0.46		

Buffer: 50mMNaCl
Sequence: 10G

t (sec)	y	Regression Output:	
30	0.99	Constant	0.984165
540	0.89	Std Err of Y Est	0.063495
1020	0.82	R Squared	0.880941
1500	0.75	No. of Observations	14
30	1	Degrees of Freedom	12
540	0.76		
1020	0.65	X Coefficient(s)	-0.00024
1500	0.59	Std Err of Coef.	2.57E-05
1980	0.53	T1/2(min)	48.1
30	1		
540	0.83		
1020	0.7		
1500	0.6		
1980	0.46		

Buffer: buffer P
Sequence: blc

t (sec)	y	Regression Output:	
0	0.96		
120	0.82	Constant	0.942805
0	0.94	Std Err of Y Est	0.038398
30	0.91	R Squared	0.704024
60	0.85	No. of Observations	26
90	0.82	Degrees of Freedom	24
120	0.76		
0	0.96	X Coefficient(s)	-0.0013
30	0.95	Std Err of Coef.	0.000171
60	0.88	T1/2(min)	8.9
90	0.85		
135	0.77		
0	0.97		
30	0.98		
60	0.9		
90	0.86		
120	0.76		
0	0.91		
30	0.81		
60	0.8		
90	0.77		
0	0.91		
38	0.9		
60	0.84		
90	0.85		
120	0.81		

Buffer: buffer P
Sequence: 15C

t (sec)	y	Regression Output:	
0	0.69		
30	0.59	Constant	
60	0.54	Std Err of Y Est	
90	0.48	R Squared	
120	0.45	No. of Observations	
30	0.67	Degrees of Freedom	
60	0.52		
90	0.45	X Coefficient(s)	-0.00213
120	0.39	Std Err of Coef.	0.000288
0	0.59	T1/2(min)	5.5
30	0.53		
60	0.39		
120	0.36		
0	0.64		
30	0.62		
60	0.47		
90	0.46		
120	0.37		

Buffer: buffer P
Sequence: 16T

t (sec)	y	Regression Output:	
0	0.82		
30	0.79	Constant	0.782353
60	0.7	Std Err of Y Est	0.08536
90	0.65	R Squared	0.537079
120	0.64	No. of Observations	19
0	0.84	Degrees of Freedom	17
30	0.77		
60	0.73	X Coefficient(s)	-0.00211
90	0.69	Std Err of Coef.	0.000476
120	0.58	T1/2(min)	5.5
30	0.75		
60	0.63		
90	0.56		
120	0.47		
0	0.64		
30	0.63		
60	0.53		
90	0.53		
120	0.38		

Buffer: buffer P
Sequence: 16C

t (sec)	y	Regression Output:	
0	0.99		
3600	0.96	Constant	
7200	0.92	Std Err of Y Est	
0	1	R Squared	
3600	0.95	No. of Observations	
7200	0.94	Degrees of Freedom	
10800	0.91		
14400	0.9	X Coefficient(s)	-5.2E-06
0	1	Std Err of Coef.	7.8E-07
3600	0.96	T1/2(min)	2222
10800	0.93		
14400	0.89		
0	1		
7200	0.87		
12600	0.91		
22020	0.91		
28800	0.83		

Buffer: buffer P
Sequence: 15C16C

t (sec)	y	Regression Output:	
0	1		
3600	0.94	Constant	0.979584
7200	0.9	Std Err of Y Est	0.022508
0	1	R Squared	0.860194
3600	0.95	No. of Observations	18
7200	0.92	Degrees of Freedom	16
10800	0.92		
14400	0.88	X Coefficient(s)	-6.9E-06
0	1	Std Err of Coef.	6.91E-07
3600	0.95	T1/2(min)	1674
7200	0.95		
10800	0.88		
14400	0.84		
0	1		
7200	0.9		
12600	0.89		
22020	0.86		
28800	0.8		

Buffer: buffer P
Sequence: 11T

t (sec)	y	Regression Output:	
0	0.85		
30	0.83	Constant	
60	0.81	Std Err of Y Est	
120	0.73	R Squared	
0	0.88	No. of Observations	
30	0.84	Degrees of Freedom	
60	0.8		
120	0.71	X Coefficient(s)	-0.0012
0	0.92	Std Err of Coef.	0.000199
30	0.86	T1/2(min)	9.6
60	0.81		
90	0.75		
105	0.75		
0	0.85		
30	0.85		
60	0.81		
90	0.73		
120	0.72		
0	0.82		
30	0.79		
90	0.7		
120	0.67		
30	0.91		
60	0.87		
90	0.87		
120	0.8		

The Effects of Central Asymmetry on the Propagation of Palindromic DNA in Bacteriophage λ Are Consistent With Cruciform Extrusion in Vivo

Alison F. Chalker,[1] Ewa A. Okely, Angus Davison and David R. F. Leach

Institute of Cell and Molecular Biology, University of Edinburgh, Edinburgh EH9 3JR, Scotland

ABSTRACT

The propagation of λ phages carrying long perfect palindromes has been compared with that of phages carrying imperfect palindromes with small regions of central asymmetry. The perfect palindromes confer a more deleterious phenotype than those with central asymmetry and the severity of the phenotype declines with the length of asymmetry in the range from O to 27 base pairs. These results argue that a center-dependent reaction is involved in the phenotypic effects of palindromic DNA sequences, consistent with the idea that cruciform extrusion occurs *in vivo*.

DNA palindromes longer than 100–150 base pairs (bp) cannot be propagated in wild-type *Escherichia coli* and replicons carrying such sequences suffer two fates. They are very poorly replicated leading to the phenotype described as inviability or the palindromes they contain are partially or completely deleted, a phenotype denoted instability (COLLINS 1981; LILLEY 1981; COLLINS, VOLCKAERT and NEVERS 1982; HAGAN and WARREN 1982; MIZUUCHI, MIZUUCHI and GELLERT 1982; HAGAN and WARREN 1983; LEACH and STAHL 1983; LEACH and LINDSEY 1986; SHURVINTON, STAHL and STAHL 1987; LINDSEY and LEACH 1989). These two fates are interrelated since replicons that have deleted their palindromes will have a selective advantage over their parents. However, it is not yet known how, or indeed whether, the mechanism of inviability is related to the mechanism of instability.

The propagation of palindromic sequences is also affected by the host genotype. It was initially demonstrated (LEACH and STAHL 1983) that a λ phage carrying a long palindrome could plate on *recBC sbcB* hosts which were subsequently shown to harbor an additional mutation in the *sbcC* gene (LLOYD and BUCKMAN 1985). Further work demonstrated that it was the *sbcC* mutation that specifically relieved the inviability of this phage (CHALKER, LEACH and LLOYD 1988). We now know that *sbcC* is part of a two gene system (the second gene has been named *sbcD*) implicated in the inhibition of propagation of palindromic DNA and the cosuppression of recombination deficiency of *recBC* mutants (LLOYD and BUCKMAN 1985; GIBSON, LEACH and LLOYD 1992). In addition to these specific interactions, mutations that improve the replication or packaging of λ contribute indirectly to its

ability to propagate palindromes (CHALKER, LEACH and LLOYD 1988; CHALKER 1990). In the case of *red gam* phage, mutations in *recB* or *recC* perform this function by permitting rolling-circle replication (see ENQUIST and SKALKA 1975) and mutations in *recD* do so, both by permitting rolling circle replication and by stimulating recombination (CHAUDHURY and SMITH 1984; AMUNDSEN *et al.* 1986; BIEK and COHEN 1986). These indirect effects have been discussed previously (CHALKER, LEACH and LLOYD 1988). Finally, it has been shown that the *gam* gene of λ facilitates the propagation of palindromes (LEACH, LINDSEY and OKELY 1987; KULKARNI and STAHL 1989) and that this effect occurs via a novel mechanism in addition to the known inhibition of RecBCD by Gam since it can be observed in a *recBCD* deletion mutant (KULKARNI and STAHL 1989). The latter authors suggest that Gam may work by inactivating the SbcC protein.

The question of whether palindromic DNA sequences can adopt cruciform structures *in vivo* has been difficult to answer conclusively. Initially, work suggested that the palindromes tested could not be detected in a cruciform conformation (COUREY and WANG 1983; SINDEN, BROYLES and PETTIJOHN 1983). These studies however used palindromes which were short (68 and 66 bp) and did not cause a severe phenotype *in vivo*. It might therefore be predicted that they would not reside predominantly in a cruciform conformation. More recently, different assays and the use of specifically designed palindromes with AT rich centers, coupled to the use of physiological conditions expected to favor extrusion, have permitted the detection of cruciform structures inside cells (PANAYATATOS and FONTAINE 1987; McCLELLAN *et al.* 1990; DAYN *et al.* 1991; ZHENG *et al.* 1991; SINDEN *et al.* 1991). The work reported below adds support

---

[1] Present address: Smith Kline Beecham, Brockham Park, Betchworth, Surrey RH3 7AJ, England.

to the idea that cruciforms can form *in vivo* and extends the evidence to normal physiological conditions and to palindromes that do not contain particularly AT-rich centers. Since such palindromes have been shown to have high kinetic barriers to cruciform extrusion *in vitro* (COUREY and WANG 1983, 1988; GELLERT, O'DEA and MIZUUCHI 1983; MURCHIE and LILLEY 1987) it seems likely that their extrusion *in vivo* is catalyzed.

MATERIALS AND METHODS

Bacterial strains: The bacterial strains used are derivatives of AB1157 F⁻ DEL(*gpt-proA*)62 *argE3 his-4 leu-6 thr-1 ara-14 galK2 lacY1 xyl-5 thi-1 supE44 rpsL31 tsx33* and were obtained from R. G. LLOYD.

Bacteriophage strains: The palindrome containing phages are all derived from DRL116, λ *pal spi6 cI857*, as shown in Figure 1. Initially, DRL133 was constructed by cleavage of DRL116 with *Sst*I and religation to delete the central 109 bp. The central *Sst*I site of DRL133 was then used for the insertion of oligonucleotide sequences obtained from the Oswel DNA service (T. BROWN, Department of Chemistry) to construct DRL134, AD1 and AD2. To facilitate the isolation and identification of the desired clones, the oligonucleotides used introduced a new restriction target sequence and inactivated the *Sst*I site. The oligonucleotides were not phosphorylated. This should prevent the insertion of multimers and no multimers were detected in any of the constructions. DRL137 was constructed by insertion of the polylinker sequence of the plasmid pMTL24 (CHAMBERS *et al.* 1988) into the *Sst*I site of DRL133, and DRL148 was made by cleavage of DRL137 with *Sal*I and religation. All manipulations were by standard methods, as outlined in SAMBROOK, FRITSCH and MANIATIS (1989).

Bacteriophage plating: Overnight cultures of cells were diluted 1:10 into L broth with 5 mM MgSO₄ and 0.2% maltose and grown for 2 hr at 37°. An equal volume of 10 mM Tris-HCl (pH 8), 10 mM MgSO₄ was then added. Aliquots of 0.25 ml of these plating cells were then incubated for 15 min at room temperature with an appropriate dilution of phage and poured onto Baltimore Biological Laboratories (BBL) plates in 2 ml of BBL top agar. The relative efficiency of plating on different hosts was calculated as:

(*pal* phage on strain *x*)/
[*pal* phage on JC7623 (*recBC sbcBC*)]
(DRL112 on strain *x*)/
[DRL112 on JC7623 (*recBC sbcBC*)]

DRL 112, λ Δ*B spi6 cI857*, is an isogenic palindrome free phage used to normalize the plating efficiencies on various hosts for minor differences not attributable to the palindromes.

Recovery of supercoiled DNA and its analysis: The methods used were as described in LEACH and LINDSEY (1986) with the following modifications. Infection was allowed to proceed for 10 min before DNA extraction. Two successive CsCl gradients instead of one were employed to ensure minimal contamination of supercoils with linear and relaxed DNA. Densitometry of bands was performed on a Shimadzu CS-930 densitometer. The control, palindrome-free, phage used in these experiments was MMS659, λ Δ*b1453 cI857*, obtained from F. STAHL.

EXPERIMENTAL RESULTS

A set of palindromes with different central sequences: λ Phages carrying palindromes that differ only in their central sequence were constructed. These phages were derived from DRL116, that has previously been used to study the behavior of palindrome bearing phage (CHALKER, LEACH and LLOYD 1988). All of the phages carry the same 446 bp (2 × 223 bp) inverted repeat but have different central sequences that vary in their degree of symmetry. Four of the phages have perfect centers and three have imperfections at or near the center, 8, 15 and 27 bp in length. These palindromes are depicted in Figure 1.

Plating behavior of palindrome phage: A subset of the above phages (DRL116, DRL133, DRL134, DRL137 and DRL148) was initially assessed for plating on strains of *E. coli* known to affect the propagation of DNA palindromes. The results of this experiment are shown in Table 1. It can be seen that the efficiency of plating (eop) of the different phages varied considerably and that the phages with perfect palindromes showed a lower eop than those with imperfect centers. This poor plating was particularly marked for phages DRL133 and 134. DRL148, which also has a perfect center, gave an eop intermediate between this pair and the phages with imperfect centers. Of these DRL137, with a 27-bp (off-set) asymmetry, plated with more facility than DRL116, with a 15-bp center asymmetry, as shown using N2362 and N2692, the least permissive hosts. We were in fact able to define a hierarchy of plating which is as follows (the number of base pairs of asymmetry is shown in brackets):

DRL137 (27) > DRL116 (15) > DRL148 (0) >
DRL133 (0), DRL134 (0).

This is a hierarchy that applies across the set of host strains tested which argues that it is independent of the host genotype.

Recovery of supercoiled DNA: In order to determine whether the effect, observed at the level of plating, was reflected by an effect at the DNA level, experiments were performed to determine the recovery of supercoiled DNA. Previous work had shown that the recovery of supercoiled DNA of the phage with the 15-bp asymmetry at the center of the palindrome (DRL116) was affected by the *sbcC* genotype of the host (CHALKER, LEACH and LLOYD 1988). We now confirm that, after a period of 15 min postinjection, recovery of DRL116 DNA in supercoiled form is close to 100% from an *sbcC* mutant but reduced to 30% when purified from a wild-type host. However, when DRL134, one of the phage with a perfect palindrome conferring the severe phenotype was tested, recovery from an *sbcC* mutant was reduced to 30%

TTATTTACTATCACAAGATTAGTAGATGTCGTTGGCGAAAAGTCATCTACTAATCTTGTGATAGTAAATA

DRL116

E   S S   E

274 15   274

SstI

TATGACAATATCCAGAAATATCTGCTTCGAATGAGCTCATTCGAAGCAGATATTTCTGGATATTGTCATA

DRL133

E   S   E

227   227

XbaI

ATATCCAGAAATATCTGCTTCGAATGAGCTGATCTAGATCAGCTCATTCGAAGCAGATATTTCTGGATAT

DRL134

E   X   E

234   234

StuI

CAATATCCAGAAATATCTGCTTCGAATGAGCTAGGCCTAGCTCATTCGAAGCAGATATTTCTGGATATTG

AD1

E   St   E

232   232

StuI   **StuI**

CAGAAATATCTGCTTCGAATGAGCTAGGCCTCTAGACGAAGGCCTAGCTCATTCGAAGCAGATATTTCTG

AD2

E St St E

235 8 235

SalI   **SalI**

AGTCGACGTCACGCGTCCATGGAGATCTCGAGGCCTGCAGGCATGCAAGCTTGCATGCCTGCAGGTCGAC

DRL137

E   SaSa   E

257 27 272
15

SalI

CGAATGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACTCTAGAGGATCCCCGGGTACCGAGCTCATTCG

DRL148

E   Sa   E
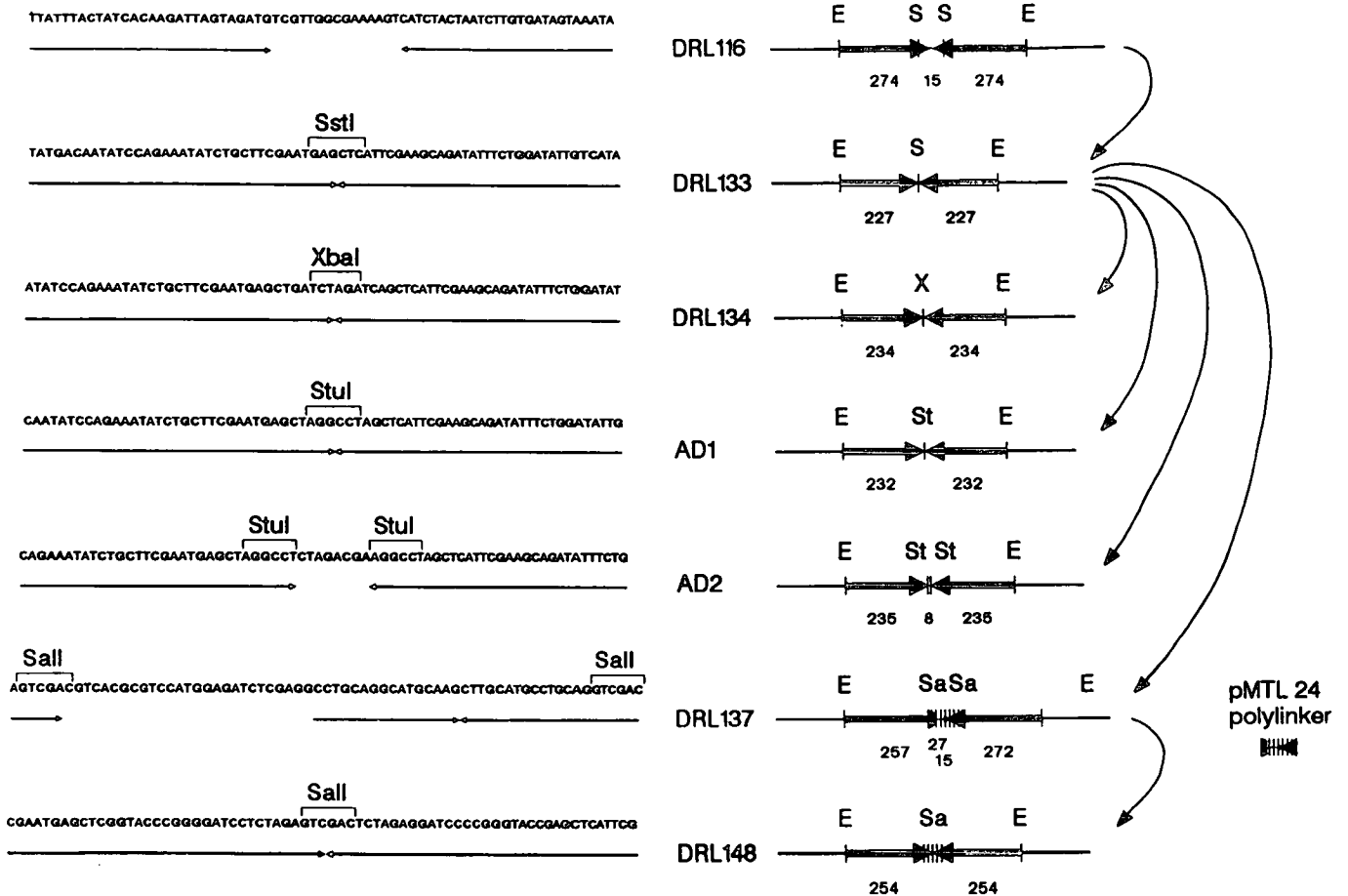
254   254

pMTL 24
polylinker

FIGURE 1.—The structure and derivation of the palindromes studied. The restriction sites involved in the construction of the palindromes are indicated as follows: E (*Eco*RI), S (*Sst*I), Sa (*Sal*I), St (*Stu*I), X (*Xba*I). The numbers of base pairs in the arms and (if asymmetric) the centers of the palindromes are listed under each representation. DRL133, DRL134, AD1 and DRL148 are perfect palindromes, whereas DRL116, AD2 and DRL137 are imperfect. AD2 has an asymmetric center of 8 bp, DRL116 has an asymmetric center of 15 bp and DRL137 has an asymmetry of 27 bp which is located 15 bp from the center of the palindrome. The left hand side of the figure depicts the central 70 bp of each of the palindromes.

and that from wild type to 2%. The behavior of DRL148, the phage with a perfect palindrome conferring an intermediate plating phenotype, was also intermediate at the DNA level. It showed a recovery of 90% from an *sbcC* mutant and 6% from wild type.

Plating behavior of a phage with an 8-bp asymmetry: To extend these observations to a palindrome with a shorter asymmetry, AD2, a phage with an 8-bp central insertion, was constructed and its behavior was compared with that of an isogenic phage with a perfect palindrome (AD1). The control phage with the perfect palindrome behaved similarly to phages DRL133 and DRL134, but AD2 showed a phenotype intermediate between that conferred by the perfect palindromes and the palindrome with a 15-bp asymmetry. These results have been incorporated into Table 1.

DISCUSSION

A center-dependent pathway exists *in vivo*: *In vitro*, it has been shown that the central sequence dictates

the cruciform extrusion rate for palindromes of average base composition (MURCHIE and LILLEY 1987; COUREY and WANG 1988; ZHENG and SINDEN 1988). However, *in vivo*, the effects of central sequence have been less well documented. The experiments of WARREN and GREEN (1985) demonstrated that inviability could not be overcome unless large insertions of 50 bp or more were made at the center of the palindrome. They argued, therefore, that the structure responsible for inviability was generated by a mechanism involving interaction of the arms of the palindrome and was independent of the center. By contrast, our results demonstrate that when more sensitive assays are used, effects of much smaller changes in the central DNA sequence can be detected. In agreement with WARREN and GREEN, all six phages studied here are unable to plate on wild-type hosts. We can, however, see a marked effect of central asymmetry even in wild-type hosts when we look at the recovery of supercoiled DNA. Effects are also

TABLE 1

Plating behavior of palindrome-containing phages on mutant *E. coli* hosts

E. coli strain and genotype	λ Strain				
	DRL137	DRL116	AD2	DRL148	DRL133 DRL134 AD1
N2361 rec⁺ sbcC⁺	−	−	−	−	−
N2362 recB21	±[a]	−	−	−	−
N2692 recD1009 recA269::Tn10	+[b]	−	−	−	−
N2678 recD1009	+	+	±[c]	±[c]	−
N2364 sbcC201	+	+	+	+	±[d]
N2365 recB21 sbcC201	+	+	+	+	+
N2680 recD1009 sbcC201	+	+	+	+	+

+ indicates an eop of >10⁻¹, − indicates an eop of <10⁻². Certain combinations of phage and bacteria are on the borderline between plating and not plating and the precise eop in these situations is very sensitive to minor variations in environmental conditions, such as the thickness of the plates, the age of the plates, the batch of agar, and the temperature of incubation. Platings were therefore done under standard conditions but in some cases it has still been necessary to report the observed range of eop that was observed in repeated experiments.

[a] eop of 10⁻¹–10⁻² (small plaques).
[b] eop of >10⁻¹ (small plaques).
[c] eop of 10⁻¹–10⁻² (very small plaques).
[d] eop of 1–10⁻² (very small plaques).

evident when we examine plating behavior on mutant hosts. This suggests that a center-dependent pathway for phenotypic effects of palindromes does exist in both wild-type and mutant *E. coli*.

We have shown that palindromes carrying small central asymmetries confer a less severe phenotype than do perfect palindromes and the severity of the phenotype correlates inversely with the length of asymmetry. Thus a central asymmetry of only 8 bp (AD2) clearly alleviates the severe phenotype caused by the perfect palindromes in AD1, DRL133 and DRL134. *In vitro* work has indicated that cruciform extrusion of perfect palindromes requires denaturation of the central 8–10 bp (MURCHIE and LILLEY 1987; COUREY and WANG 1988) and that most cruciform loops include 4 bases (GOUGH, SULLIVAN and LILLEY 1986). That a difference can be detected *in vivo* between perfect palindromes and a palindrome with 8 bases of asymmetry argues strongly for a center-dependent pathway of extrusion. It is interesting that one of the perfect palindromes (DRL148) confers an intermediate phenotype, similar to that of the palindrome with the 8 bp asymmetry and less severe than those of the other perfect palindromes. It is true that the center of DRL148 is slightly more G/C rich than those of the other perfect palindromes. However, there are other features of the DRL148 structure which may share responsibility for its more moderate phenotype. The central region carries several palin-

dromic recognition sequences for restriction enzymes which might abortively attempt to nucleate extrusion and thus inhibit extrusion from the central sequence; additionally, there is a sequence 5'-CCCGGGG-3', 13-bp on either side of the center, which may interfere with extrusion. Finally, perhaps the central sequence itself is inherently less prone to initiate extrusion. Further work is required to determine why DRL148 behaves differently from the other perfect palindromes.

If cruciform extrusion occurs *in vivo*, it must be catalyzed: It has previously been demonstrated that cruciform extrusion *in vitro* occurs so slowly, in normal DNA sequences under physiological conditions, that no *in vivo* effects would be expected (COUREY and WANG 1983, 1988; GELLERT, O'DEA and MIZUUCHI 1983; MURCHIE and LILLEY 1987). This kinetic argument against cruciform extrusion *in vivo* has been supported by several studies in which cruciform structures have not been detected *in vivo* (COUREY and WANG 1983; SINDEN, BROYLES and PETTIJOHN 1983). Recent work, however, has revealed that under conditions of elevated superhelical density, palindromes that have central sequences favoring cruciform extrusion *in vitro* can extrude such structures *in vivo* (MCCLELLAN *et al.* 1990; DAYN *et al.* 1991; ZHENG *et al.* 1991; SINDEN *et al.* 1991). In addition, the deletion frequencies of certain palindromes argue for a pathway involving a cruciform intermediate (SINDEN *et al.* 1991). Nevertheless, the sum of these observations has been taken to argue against significant extrusion of palindromes of *average* base composition under *normal* physiological conditions.

Our results are, however, most easily explained if cruciform extrusion does occur *in vivo* in perfect palindromes with centers of normal base composition and even in palindromes with asymmetric centers of 8, 15 and 27 bp. In order to reconcile the *in vitro* data with our results we must suggest that cruciform extrusion is catalyzed *in vivo*. If so, why have previous studies *in vivo* required unusual palindromes and/or unusual physiological conditions to detect extrusion? We suggest that the relevant difference between the palindromes studied here and those studied previously is length. The palindromes employed in the present work are between 446 and 571 bp in length, whereas those used by other workers had been kept shorter than 150 bp to avoid the propagation problems of the longer sequences. It is, however, this phenotype of inviability that is likely to correlate with the accumulation of cruciform structures *in vivo*. We propose therefore that (as for palindromes *in vitro*) the center affects the kinetics of cruciform extrusion in a catalyzed reaction, but the length of the palindrome is critical in determining the lifetime of the cruciform structure.

In vitro, cruciform extrusion occurs either via the C-type or the S-type pathway (LILLEY 1985). The C-type pathway involves a large denaturation bubble which is favored by highly A/T-rich sequences at low salt concentrations and is therefore unlikely to relate to extrusion *in vivo*. The S-type pathway proceeds via a small central denaturation bubble which folds into a protocruciform that in turn extrudes to the full cruciform by branch migration. The transition state, which is probably the protocruciform, is thought to be stabilized by salt (SULLIVAN and LILLEY 1987). If the pathway to cruciform extrusion *in vivo* is related to S-type extrusion *in vitro*, we predict that specific base-pair changes that alter the rate of *in vitro* extrusion will have predictable effects on the phenotype of palindrome containing phage *in vivo*. Experiments to test this prediction are under way.

LITERATURE CITED

AMUNDSEN, S. K., A. F. TAYLOR, A. M. CHAUDHURY and G. R. SMITH, 1986 *recD*: the gene for an essential third subunit of exonuclease V. Proc. Natl. Acad. Sci. USA **83**: 5558–5562.

BIEK, D. P., and S. N. COHEN, 1986 Identification and characterization of *recD*, a gene affecting plasmid maintenance and recombination in *Escherichia coli*. J. Bacteriol. **167**: 594–603.

CHALKER, A. F., 1990 *SbcC* and palindrome-mediated inviability in *Escherichia coli*. Ph.D. Thesis, University of Edinburgh.

CHALKER, A. F., D. R. F. LEACH and R. G. LLOYD, 1988 *Escherichia coli sbcC* mutants permit stable propagation of replicons containing a long DNA palindrome. Gene **71**: 201–205.

CHAMBERS, S. P., S. E. PRIOR, D. A. BARSTOW and N. P. MINTON, 1988 The pMTL *nic* cloning vectors. I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing. Gene **68**: 139–149.

CHAUDHURY, A. M., and G. R. SMITH, 1984 A new class of *Escherichia coli recBC* mutants: implications for the role of RecBC enzyme in homologous recombination. Proc. Natl. Acad. Sci. USA **81**: 7850–7854.

COLLINS, J., 1981 The instability of palindromic DNA. Cold Spring Harbor Symp. Quant. Biol. **45**: 409–416.

COLLINS, J., G. VOLCKAERT and P. NEVERS, 1982 Precise and nearly-precise excision of the symmetrical inverted repeats of Tn*5*: common features of *recA*-independent deletion events in *E. coli*. Gene **19**: 139–146.

COUREY, A. J., and J. C. WANG, 1983 Cruciform formation in a negatively supercoiled DNA may be kinetically forbidden under physiological conditions. Cell **33**: 817–829.

COUREY, A. J., and J. C. WANG, 1988 Influence of DNA sequence and supercoiling on the process of cruciform formation. J. Mol. Biol. **202**: 35–43.

DAYN, A., S. MALKHOSYAN, D. DUZHY, V. LYAMICHEV, Y. PANCHENKO and S. MIRKIN, 1991 Formation of (dA-dT)$_n$ cruciforms in *Escherichia coli* cells under different environmental conditions. J. Bacteriol. **173**: 2658–2664.

ENQUIST, L. W., and A. SKALKA, 1975 Replication of bacteriophage lambda DNA dependent on the function of host and viral genes. I. Interaction of *red gam* and *rec*. J. Mol. Biol. **75**: 185–212.

GELLERT, M., M. H. O'DEA and K. MIZUUCHI, 1983 Slow cruciform transitions in palindromic DNA. Proc. Natl. Acad. Sci. USA **80**: 5545–5549.

GIBSON, F. P., D. R. F. LEACH and R. G. LLOYD, 1992 Identification of *sbcD* mutations as cosuppressors of *recBC* that allow propagation of DNA palindromes in *Escherichia coli* K12. J. Bacteriol. **174**: 1222–1228.

GOUGH, W. G., K. M. SULLIVAN and D. M. J. LILLEY, 1986 The structure of cruciforms in supercoiled DNA: probing the single-stranded character of nucleotide bases with bisulphite. EMBO J. **5**: 191–196.

HAGAN, C. E., and G. J. WARREN, 1982 Lethality of palindromic DNA and its use in selection of recombinant plasmids. Gene **19**: 147–151.

HAGAN, C. E., and G. J. WARREN, 1983 Viability of palindromic DNA is restored by deletions occurring at low but variable frequency in plasmids of *Escherichia coli*. Gene **24**: 317–326.

KULKARNI, S. K., and F. W. STAHL, 1989 Interaction between the *sbcC* gene of *Escherichia coli* and the *gam* gene of phage lambda. Genetics **123**: 249–253.

LEACH, D. R. F., and J. C. LINDSEY, 1986 *In vivo* loss of supercoiled DNA carrying a palindromic sequence. Mol. Gen. Genet. **204**: 322–327.

LEACH, D. R. F., J. C. LINDSEY and E. A. OKELY, 1987 Genome interactions which influence DNA palindrome mediated instability and inviability in *Escherichia coli*. J. Cell Sci. Suppl. **7**: 33–40.

LEACH, D. R. F., and F. STAHL, 1983 Viability of lambda phages carrying a perfect palindrome in the absence of recombination nucleases. Nature **305**: 448–451.

LILLEY, D. M. J., 1981 *In vivo* consequences of plasmid topology. Nature **292**: 380–382.

LILLEY, D. M. J., 1985 The kinetic properties of cruciform extrusion are determined by DNA base-sequence. Nucleic Acids Res. **13**: 1443–1465.

LINDSEY, J. C., and D. R. F. LEACH, 1989 Slow replication of palindrome-containing DNA. J. Mol. Biol. **206**: 779–782.

LLOYD, R. G., and C. BUCKMAN, 1985 Identification and genetic analysis of *sbcC* mutations in commonly used *recBC sbcB* strains of *Escherichia coli* K-12. J. Bacteriol. **164**: 836–844.

MCCLELLAN, J. A., P. BOUBLIKOVA, E. PALECEK and D. LILLEY, 1990 Superhelical torsion in cellular DNA responds directly to environmental and genetic factors. Proc. Natl. Acad. Sci. USA **87**: 8373–8377.

MIZUUCHI, K., M. MIZUUCHI and M. GELLERT, 1982 Cruciform structures in palindromic DNA are favoured by DNA supercoiling. J. Mol. Biol. **156**: 229–244.

MURCHIE, A. I. H., and D. M. J. LILLEY, 1987 The mechanism of cruciform formation in supercoiled DNA: initial opening of central basepairs in salt dependent reaction. Nucleic Acids Res. **15**: 9641–9654.

PANAYATATOS N., and A. FONTAINE, 1987 A native cruciform DNA structure probed in bacteria by recombinant T7 endonuclease. J. Biol. Chem. **262**: 11364–11368.

SAMBROOK, J., E. F. FRITSCH and T. MANIATIS Editors, 1989 *Molecular Cloning: A Laboratory Manual*, Ed. 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

SHURVINTON, C. E., M. M. STAHL and F. W. STAHL, 1987 Large palindromes in the λ phage genome are preserved in a *rec*⁺ host by inhibiting DNA replication. Proc. Natl. Acad. Sci. USA **84**: 1624–1628.

SINDEN, R. R., S. BROYLES and D. PETTIJOHN, 1983 Perfect palindromic *lac* operator DNA sequence exists as a cruciform structure in supercoiled DNA *in vitro* but not *in vivo*. Proc. Natl. Acad. Sci. USA. **80**: 1797–1801.

SINDEN, R. R., G. ZHENG, R. G. BRANKAMP and K. N. ALLEN,

1991   On the deletion of inverted repeated DNA in *Escherichia coli*: effects of length, thermal stability and cruciform formation *in vivo*. Genetics **129**: 991–1005.

SULLIVAN, K. M., and D. M. J. LILLEY, 1987   Influence of cation size and charge on the extrusion of a salt-dependent cruciform. J. Mol. Biol. **193**: 397–404.

WARREN, G. J., and R. L. GREEN, 1985   Comparison of physical and genetic properties of palindromic DNA sequences. J. Bacteriol. **161**: 1103–1111.

ZHENG, G., and R. R. SINDEN, 1988   Effects of base composition at the center of inverted repeated DNA sequences on cruciform transitions in DNA. J. Biol. Chem. **263**: 5356–5361.

ZHENG, G., T. KOCHEL, R. W. HOEPFNER, S. E. TIMMONS and R. R. SINDEN, 1991   Torsionally tuned cruciform and Z-DNA probes for measuring unrestrained supercoiling at specific sites in DNA of living cells. J. Mol. Biol. **221**: 107–129.