AN ABSTRACT OF THE THESIS OF

Kelly L. Bennett for the degree of Doctor of Philosophy in Biochemistry and Biophysics presented on April 30, 1991.

Title: Sequence Conversion During Adenovirus DNA Replication.

Abstract approved: Redacted for Privacy
Dr. George D. Pearson

The work in this thesis has provided conclusive genetic evidence that "panhandle" intermediates form during adenovirus replication. Adenovirus chromosomes lacking 51 bp from their left-hand termini are infectious and capable of regenerating the missing origin sequence. Yet if an entire inverted terminal repeat is removed, the adenovirus chromosome is no longer viable. This first suggested, but did not prove, that "panhandles" formed during adenovirus replication. Homologous recombination or postreplicative overlap recombination could generate the same outcome. Analysis of the segregation of markers in the inverted repeats of adenovirus minichromosomes shows that homologous recombination does not mediate end repair. A special case was also found where postreplicative overlap recombination failed to transfer sequences between the inverted repeats, but similar molecules could exchange sequence information during "panhandle" formation. The exchange of information between inverted repeats is referred to as sequence conversion. A number of length and/or orientation constraints on sequence conversion during adenovirus DNA replication were

identified. A length- and orientation-dependent constraint was found for gap filling close to "panhandle" loops. Polymerization towards the loop could occur even when the gap was only 6 bp away. In contrast, polymerization away from the "panhandle" loop at a gap at 6 bp, did not take place. This steric constraint could reflect an asymmetry in the action of adenovirus DNA polymerase. A similar length and/or orientation dependent constraint was found for the removal of bulges (3 bp and 4 bp mismatches). Incision in the bulge of the 5' inverted repeat caused a block to the completion of sequence conversion at that site. When the bulge was in the 3' inverted repeat, a length requirement for successful removal was demonstrated. When 6 bp or 39 bp separated the bulge from the "panhandle" loop, removal of the bulge was not detected. When the distance was 79 bp, 184 bp, or 217 bp, bulges were successfully removed. The molecular basis for this obstruction remains to be determined. Moreover, incision in bulges located in the 3' inverted repeat triggers directional coconversion. Finally, small loops placed close to the site of polymerization did not cause the same length and orientation dependent constraints as did the "panhandle" loop.

Sequence Conversion During Adenovirus DNA Replication

by

Kelly L. Bennett

A THESIS

submitted to

Oregon State University

in partial fullfillment of
the requirements for
the degree of

Doctor of Philosophy

Completed April 30, 1991

Commencement June 1991

APPROVED:

Redacted for Privacy

Professor of Biochemistry and Biophysics in charge of major

Redacted for Privacy

Chairman of the Department of Biochemistry and Biophysics

Redacted for Privacy

Dean of the Graduate School

Date thesis is presented April 30, 1991

Typed by Kelly L. Bennett for Kelly L. Bennett

ACKNOWLEDGEMENTS

I would like to thank my colleagues for the advice, materials and friendship they provided me with over the past four years. This includes Dr. Kevin Ahern, Dr. Fu-Yun Xu, Kathy Preszler, Jasmine Ng, Margaret Ho, Mark Bonnlander, and a very special thanks to Dr. Chin-Hwa Hu for his completely altruistic approach in the laboratory.

I would like to acknowledge the advice, direction, and encouragement provide by my major professor, Dr. George Pearson. The support he provided for me is greatly appreciated. In addition, I would like to extend my gratitude to my committee members, Dr. Christopher Mathews, Dr. Pui Shing Ho, Dr. Dallice Mills, and Dr. Jane Aldrich, for their efforts. A special thanks is in order for Dr. Christopher Mathews for ensuring that I received a stipend, limiting the financial burden of graduate school.

Graduate school requires many sacrifices from the student's family. I would like to acknowledge and thank my husband, Joe Ewing, for the sacrifices he endured. The emotional and financial support my parents, Ben and Lorna Bennett, offered me has been greatly appreciated. Without their support the last four years would of been much more difficult. My son, Andrew Ewing, who was 22 months old at the outset of this adventure, has been a great source of delight and comfort when graduate school responsibilities were burdensome. I hope the sacrifices my family experienced will soon be replaced by the greater opportunities before us.

TABLE OF CONTENTS

	Page
CHAPTER I: BACKGROUND	1
Herpes simplex virus (HSV-1)	1
Adeno-associated virus (AAV)	4
Zygosaccharomyces rouxii plasmid pSR1	5
Adenovirus DNA replication	6
Adenovirus DNA replication proteins	9
The adenovirus origin of replication	11
Experimental strategy	12
Sequence conversion	17
Frameshift mutations (small insertions and deletions)	18
CHAPTER II: MATERIALS AND METHODS	23
Reagents and enzymes	23
Bacteria and plasmids	23
Growth and purification of adenovirus DNA	26
DNA transfection	28
DNA extraction from transfected cells	28
Southern blot analysis	29
Purification of the *cis* replication product from the input DNA	29
Restriction endonuclease cleavage of purified input and *cis* bands	30
Radioactive labeling of plasmid DNA by random priming	31
End-labeling of oligonucleotide probes	31

DNA hybridization 32

CHAPTER III: SEQUENCE CONVERSION OCCURS ON PANHANDLE INTERMEDIATES DURING ADENOVIRUS MINICHROMOSOME REPLICATION 34

INTRODUCTION 34

RESULTS 36

Marker segregation during adenovirus minichromosome replication 36

Marker segregation is not mediated by homologous recombination 40

"Panhandle" intermediates occur during adenovirus replication 41

DISCUSSION 45

Marker segregation is not mediated solely by postreplicative overlap recombination 46

Enzymes that are predicted to function during sequence conversion on "panhandle" intermediates 47

The limitations of sequence conversion on "panhandle" intermediates 47

The IR used as template by each mutant is different 48

The consequences of unsuccessful attempts of sequence conversion on "panhandle" intermediates 52

CHAPTER IV: SEQUENCE CONVERSION OCCURS DURING ADENOVIRUS POSTREPLICATIVE OVERLAP RECOMBINATION 71

INTRODUCTION 71

RESULTS 72

Intermolecular sequence conversion 73

Homologous recombination does not take place between the minichromosomes 74

Limitations to intermolecular sequence conversion between minichromosomes 75

A model for sequence conversion during postreplicative adenovirus overlap recombination 76

DISCUSSION 79

Homologous recombination is not detected between adenovirus minichromosomes in mammalian cells 80

Sequence conversion during postreplication overlap recombination 81

CHAPTER V: DIRECTIONAL COCONVERSION DURING ADENOVIRUS MINICHROMOSOME REPLICATION 90

INTRODUCTION 90

RESULTS 91

Unidirectional coconversion during adenovirus DNA replication 93

A model for unidirectional coconversion 94

DISCUSSION 95

The possible role of adenovirus replication proteins in preventing successful sequence conversion attempts 97

CHAPTER VI: SEQUENCE CONVERSION AT SITES CLOSE TO THE LOCATION OF ADENOVIRUS MINICHROMOSOME END-REPAIR 103

INTRODUCTION 103

RESULTS 104

Sequence conversion when the heterogeneity is only eight bp from the site of polymerization 105

Sequence conversion when the heterogeneity is 23 bp from the site of polymerization 110

Sequence conversion when two regions of heterogeneity are within 31 bp from the site of polymerization 113

Sequence conversion on adenovirus minichromosomes with three-bp deletions 114

DISCUSSION 114

Special properties of the tandem duplication 115

Further evidence that supports a steric *versus* a stability constraint on the duplex length for end-repair 116

Evolution of the ITR in view of the limitations of sequence conversion close to the `panhandle' loop 117

CHAPTER VII: CONCLUSIONS 126

A summary of the sequence conversion pathways 127

End-repair and sequence conversion are sterically constrained 130

Unexpected novel properties of adenovirus DNA replication enzymes 134

Adenovirus DNA polymerase and preterminal protein 134

Adenovirus DNA binding protein 135

All stable small bulges not located near the ends of small IR are always converted 135

Sequence conversion during adenovirus DNA replication functions to maintain the identity of the origin sequences and the length of the ITR 138

BIBLIOGRAPHY 145

LIST OF FIGURES

Figure		Page
I.1	The scheme for adenovirus DNA replication	20
I.2	Generation of a replication-proficient plasmid with two adenovirus termini from a plasmid bearing a single viral terminus	21
I.3	The assay and requirements for detection of a replication-proficient plasmid with two adenovirus termini from a plasmid bearing a single viral terminus	22
III.1	Structure of the pIR and pLIR series of plasmids	55
III.2	The location of the restriction sites and the products produced by restriction endonuclease cleavage	56
III.3	Endonuclease restriction analysis of sequence conversion in pIR3'S	57
III.4	Endonuclease restriction analysis of sequence conversion in pLIR3'S	58
III.5	Endonuclease restriction analysis of sequence conversion in pIR5'S	59
III.6	Endonuclease restriction analysis of sequence conversion in pLIR5'S	60
III.7	Endonuclease restriction analysis of transfected pIR3'S without helper DNA	61
III.8	Endonuclease restriction analysis of gel-purified pIR3'S dimer DNA	62
III.9	Diagram of adenovirus minichromosome sequence conversion mediated by homologous recombination	63
III.10	Mismatch repair (sequence conversion) of 3' mutants	64

III.11	Mismatch repair (sequence conversion) of 5' mutants	65
III.12	Summary of the IR used by each mutant	66
III.13	Analysis of the possible outcomes of sequence conversion during *trans* replication on pLIR5'S	67
III.14	Analysis of the possible outcomes of sequence conversion during *trans* replication on pIR5'S	68
III.15	Analysis of the possible outcomes of sequence conversion during *trans* replication on pLIR3'S	69
III.16	Analysis of the possible outcomes of sequence conversion during *trans* replication on pIR3'S	70
IV.1	Analysis of sequence conversion during postreplicative overlap recombination between two minichromosomes with an IR of 231 bp	85
IV.2	Control for homologous recombination	86
IV.3	Analysis of sequence conversion during postreplicative overlap recombination between two minichromosomes with an IR of 51 bp	87
IV.4	Analysis of the possible outcomes of sequence conversion during *trans* replication between pLIRWT and pLIRSD	88
IV.5	Analysis of the possible outcomes of sequence conversion during *trans* replication between pIRSD and pIRWT	89
V.1	Analysis of sequence conversion of the double mutants with the *Kpn*I and *Sph*I sites deleted	100
V.2	Analysis of coconversion on pIR5'K5'S	101
V.3	Directional coconversion away from panhandle loop	102
VI.1	Analysis of sequence conversion of the *Kpn*I mutants	120
VI.2	Analysis of sequence conversion of the *Xba*I mutants	121

VI.3	Analysis of sequence conversion of the double mutants with the *Kpn*I and *Xba*I sites altered	122
VI.4	Branch migration within the *Xba*I tandem duplication	123
VI.5	Analysis of sequence conversion on molecules with a three bp mutation	124
VI.6	The asymmetrical sequences at the *Bgl*I mutation site	125
VII.1	Summary of the orientation and/or length requirements for sequence conversion close to "panhandle" loops	140
VII.2	Mismatch repair toward panhandle loop	141
VII.3	Mismatch repair away from panhandle loop	142
VII.4	Adenovirus type 2 origin of DNA replication	143
VII.5	A mechanism for maintaining the length of the ITR in the absence of selection	144

SEQUENCE CONVERSION DURING ADENOVIRUS DNA REPLICATION

CHAPTER I: BACKGROUND

The goal of this work has been to gain insight into the mechanism that adenovirus employs to maintain the identity of its inverted terminal repeats (ITRs). All adenovirus genomes contain ITRs, which vary in length from 63 to 166 base pairs (bp) for different adenovirus serotypes (Stillman *et al.*, 1982a; Steenbergh *et al.*, 1977; Shinagawa *et al.*, 1983; for review see van Ormondt and Galibert, 1984). The ITR contains sequences for the origin of replication, and also provides a critical structural role for the replication of single-stranded replication intermediates. Maintaining the identity of the ITR is critical for the life cycle of adenovirus; deletion of one of the ITRs is lethal to the virus (Stow, 1982).

Other investigators studying a variety of organisms have also examined how the identity of inverted repeats (IRs) are maintained. A description of what is known about this process for herpes simplex virus, adeno-associated virus, and plasmid pSR1 of *Zygosaccharomyces rouxii* follows. The role of ITRs in adenovirus DNA replication is then discussed in detail.

Herpes simplex virus (HSV-1)

The herpes simplex virus (HSV-I) genome is a double-stranded DNA molecule of 96 x 10^6 Dal. The genome consists of two covalently linked components which are predominantly unique sequences L and S bracketed by large inverted repeats (Sheldrick and Berthelot, 1975). The repeated sequences which bracket the L component are designated **ab**

and the inversion **a'b'**, while the S inverted repeat sequences are referred to as **a'c'** and **ca** (Waldsworth, *et al.*, 1975). Sequence **a**, common to both L and S, is in the same orientation at each end of the genome, and the sequences that divide the L and S components are arranged in the order **b'a'a'c'**. Two regions within **a'c'** and **ca** have been differentiated (Roizman *et al.*, 1979; Knipe *et al.*, 1978; Roizman, 1979). The outermost regions, **a'c'1** and **c1a**, were shown to be obligatorily identical; insertions or substitutions of nucleotide sequences in this region of one sequence become reflected in the corresponding region of the other inverted sequence. This suggests that at least one of these regions acts as a template for the other through intramolecular interactions. Roizman *et al.* (1979) further argued in favor of a intramolecular repair model. First of all, the *ts*C75 mutant is believed to have an identical mutation in all four a sequences, an extremely rare event to have arisen independently. Moreover, the mutant was rescued by any one wild-type terminus, which further indicates that a copy mechanism exists which transposes any base substitution from one site to all other sites. The innermost regions, **c'2** and **c2**, can be, but are not required to be identical. The data also suggest that the main features of sequence identity for **ca** and **a'c'** apply for sequences **ab** and **b'a'** as well.

Further evidence that intramolecular repair maintains the identity of the inverted repeat comes from many observations used to develop a model for the replication of HSV-1. After infection, HSV-1 DNA is digested by a processive exonuclease to expose cohesive single-stranded DNA sequences. The genome then circularizes forming a

junction designated **bac**. The DNA is replicated by a rolling-circle mechanism which generates a concatemer that is cleaved only at the **bac** junction. Cleavage occurs at either the left or the right of sequence **a**; therefore, the excised molecules need to regenerate one of the ends which has lost sequence **a**.

Sequence **a** could be regenerated by using the internal inverted repeat as a template. A mechanism was proposed by Heumann (1976) for the replacement of terminal RNA primer sequences by DNA sequences, and Jacob *et al.* (1979) used this mechanism as a model for the repair of sequence **a**. A site-specific endonuclease could cleave the internal sequence **b'a'a'c'** creating a nick, followed by strand exchange between the inverted sequences, and finally copying sequence a using each strand of the internal inverted repeat as template. To regenerate the linear form, the **b'a'a'c'** junction would have to be nicked and resolved by branch migration. Depending on the location of the original cleavage in the concatemer and whether nicking occurred in either the repaired strand or the template strand, four isomers would arise from inversion of the L and S components. HSV-1 DNA in fact encompasses four equimolar populations of molecules differing solely in the orientations of L and S components relative to each other (Delius and Clements, 1976; Hayward *et al.*, 1975; Roziman *et al.*, 1974). These populations have been designated as P (prototype), IS (inversion of S component), IL (inversion of L component), and ISL (inversion of L and S component). Thus, if a P arrangement of HSV-1 DNA initiated the replication process and the cleavage of the concatemer occurred between the **a** and **c** sequences in the **bac** junction, cleavage and branch

migration of the repaired strand at the L-S junction would regenerate the P arrangement. However, cleavage and branch migration of the template strand and branch migration of its complement, would generate the IS arrangement. Similarly, cleavage of the concatemer between **b** and **a** sequences at the **bac** junction would ultimately generate the P or the IL arrangement (Jacob *et al.*, 1979).

In conclusion, the obligatory repair of the ends of the genome during HSV-1 replication provides a mechanism to invert the L and S components. It has been shown that sequence **a** is a site-specific inversion site (Mocarski *et al.*, 1980; Mocarski *et al.*, 1981; Smiley *et al.*, 1981). When sequence **a** is inserted into other regions of the genome, the new sequences now flanked by sequence **a** also invert. Furthermore, inversion appears to be caused by *trans*-acting factors present in infected, but not uninfected, cells (Mocarski and Roizman, 1982). This process ensures that the IRs remain identical.

Adeno-associated virus (AAV)

The genome of AAV is a linear, single-stranded DNA molecule 4.7 kilobases (kb) long with a 145-nucleotide inverted terminal repeat (Kozcot *et al.*, 1973; Berns and Kelly, 1974). The first 125 bases consist of a palindromic sequence (Lusby *et al.*, 1980), which is interrupted by two smaller palindromes. The terminal repeat is thought to serve as the primer for AAV DNA replication (Straus *et al.*, 1976; Hauswirth and Berns, 1977, 1979). Two possible hairpin structures can form to provide a 3'-OH primer essential for the initiation of DNA replication. One hairpin is folded on itself to obtain maximal base-pairing; each end forms a T-shaped structure, with the two T's separated by the

remaining single-stranded genome. Alternatively, the ends of the genome can interact to form a "panhandle-like" structure (Samulski *et al.*, 1983; Bohenzky and Berns 1989). Samulski *et al.* (1983) showed that viable wild-type virus could be recovered from ITR deletion mutants as long as some of the AAV terminal sequences were retained at both ends and that each mutant had at least one intact ITR. The recovery of wild-type virus occurred only by sequence conversion on the "panhandle-like" structure, where the ends of the genome interact, not the T-shaped structure. Bohenzky and Berns (1989) also created a chimeric viral chromosome by replacing an 11-bp symmetrical sequence at one end of the genome with a different 12-bp symmetrical sequence. All progeny molecules contained the wild-type sequence in both terminal repeats. Interactions between the ITRs clearly take place, but a mechanism involving homologous recombination cannot be ruled out in either of the experiments above.

***Zygosaccharomyces rouxii* plasmid pSR1**

Plasmid pSR1 is a double-stranded, circular DNA molecule consisting of 6,251 bp. It has a 959-bp inverted repeat dividing the plasmid molecule into two unique regions, one 2,654 bp long and the other 1,679 bp. Molecules exist in two isomeric forms, types A and B, arising by intramolecular recombination through the inverted repeats. The plasmid pSR1 encodes three large open reading frames designated P, R, and S. The R gene product is an essential *trans*-active factor for the intramolecular recombination (Araki *et al.*, 1985; Jearnpipatkul *et al.*, 1987). Matsuzaki *et al.* (1988) showed that any heterogeneity created at any region in the inverted repeats was restored to identity by

gene conversion. The group also demonstrated that a 58-bp sequence in the inverted repeat is essential for initiating intramolecular recombination, which is accompanied by high-frequency gene conversion.

Adenovirus DNA replication

The adenovirus genome is a linear, double-stranded DNA molecule of approximately 36 kb. The origin sequences controlling replication are embedded within an inverted terminal repeat (ITR) which ranges from 63 to 166 bp for different adenovirus serotypes (Stillman *et al.*, 1982a; Steenbergh *et al.*, 1977; Shinagawa *et al.*, 1983; for review see van Ormondt and Galibert, 1984). The ITR plays an important structural role, described in detail below, in adenovirus replication. A second striking structural feature of the virus chromosome is that a 55-kDal virus-encoded protein, called terminal protein (TP), is covalently linked to the 5' end of each DNA strand (Robinson *et al.*, 1973; Robinson, and Bellett, 1974). This discovery led to the suggestion that the protein was involved in priming DNA replication (Robinson and Bellett 1974; Rekosh *et al.*, 1977). The covalent link is formed by a phosphodiester bond between the 5'-deoxycytidine residue in the DNA and a serine residue in the protein (Rekosh *et al.,* 1977; Stillman, 1981; Desiderio and Kelly 1981; Challberg and Kelly, 1981; Smart and Stillman, 1982). Initiation of replication *in vitro* has been shown to require pTP (the 80-kDal precursor of terminal protein), Ad pol (adenovirus DNA polymerase), adenovirus TP-DNA complex as template, ATP, $MgCl_2$, and dCTP (Enomoto *et al.*, 1981; Lichy *et al.*, 1981; Pincus *et al.*, 1981; Stillman, 1981; Challberg *et al.*,

1982; Ikeda *et al.*, 1982; Tamanoi and Stillman, 1982; Lichy *et al.*, 1982). pTP is initially linked to replicated DNA, but pTP is processed by proteolysis to mature TP during packaging of the DNA into the virus particle (Challberg and Kelly, 1981).

Lechner and Kelly (1977) presented a model for adenovirus DNA replication (Fig I.1) based on an electron microscopic study of replicating viral chromosomes. Their experiments demonstrated that replication initiates equally from either, or both, molecular ends (Fig I.1b). Two types of replicating chromosomes were identified. Type I molecules consist of a double-stranded molecule with one or more single-stranded branches (Fig I.1b). Type II molecules are linear with a region of double-stranded DNA and a stretch of single-stranded DNA (Fig I.1g). Replication involves two phases. The first phase proceeds by a strand-displacement mechanism creating a double-stranded daughter molecule (Fig I.1c) and a single-stranded parental molecule (Fig I.1d). The second phase, called complementary strand synthesis, converts the single-stranded parental molecule into the other double-stranded daughter molecule (Fig I.1c). Daniell (1976), during a study of defective adenovirus genomes, proposed that a "panhandle" intermediate (Fig I.1e) formed by base-pairing between the ITRs could recreate a double-stranded origin to allow initiation of replication (Fig I.1f) on the displaced parental single strand. Furthermore, aberrant internal base-pairing provided an explanation for the structures of defective genomes. Lechner and Kelly embraced this idea, and included it in their model even though replicating "panhandle" structures had not been identified. Evidence supporting the existence

of "panhandles" has accumulated (Stow, 1982; Hay *et al.*, 1984; Wang and Pearson, 1985; Haj-Ahmad and Graham, 1986; Leegwater *et al.*, 1988; Lippe and Graham, 1989; Wang *et al.*, 1991), and will be discussed extensively in the following chapters.

More recent experiments (Ahern *et al.*, 1991) indicate a need to revise the Lechner and Kelly model for adenovirus DNA replication as shown in Fig. I.1. The displaced single strand (Fig. I.1d) can reenter replication by two pathways. One pathway involves "panhandles" (Fig. I.1e), and is called *cis* replication to emphasize the intra-strand base-pairing. The second pathway, termed *trans* replication (Ahern *et al.*, 1991), involves direct intermolecular hybridization between two displaced single strands to create a "parental-like" duplex.

The *trans* pathway was discovered by Ahern *et al.* (1991) when studying postreplicative overlap recombination in an adenovirus minichromosome system (Hay *et al.*, 1984; Wang and Pearson, 1991). Overlap recombination is a process by which overlapping terminal fragments recombine to generate a complete adenovirus genome (Chinnadurai *et al.*, 1979; Berkner and Sharp, 1983; Volkert and Young, 1983; McGrory *et al.*, 1988). Previously, it had been shown to occur through homologous recombination as a prereplicative process (Volkert and Young, 1983). More recently, it has been documented that overlap recombination can result from postreplicative events not involving homologous recombination (Ahern *et al.*, 1991). Postreplicative overlap recombination exhibits several salient features: (a) It is driven by adenovirus-directed strand-displacement replication. (b) Only one strand, the displaced strand, from each parental molecule,

participates in the formation of a heteroduplex intermediate (the pathway is therefore intrinsically nonreciprocal). (c) The heteroduplex arises from direct intermolecular hybridization of displaced strands, so single strands need not sequentially invade a duplex to form a recombinational intermediate (the process does not manifest polarity). (d) The intermediate structure is initially duplex only within the overlap. (e) Any heterologous sequences form single-stranded loops within the overlap or remain as unpaired 3' or 5' ends if located externally to the overlap. (f) Unpaired 3' ends, but not 5' ends, are excised exonucleolytically from the heteroduplex. (g) Large-scale internal nonhomologous regions (*e.g.*, large deletion or insertion loops) remain intact, but the removal of small bulges (3 or 4 bp) can occur, as documented in the following chapters. (h) Apparent crossovers at the outermost borders of the overlap are a consequence of repair DNA synthesis exclusively at those sites. (i) DNA repair regenerates adenovirus origins at both ends of the heteroduplex molecule. (j) Subsequent rounds of adenovirus-directed DNA replication segregate and amplify two recombinant molecules arising equally from each strand of the repaired heteroduplex.

Adenovirus DNA replication proteins

The development by Challberg and Kelly (1979a, 1979b) of a cell-free extract from adenovirus-infected cells has led to identification of all the factors required for adenovirus replication. Three viral-encoded proteins and three cellular proteins have been characterized. These proteins are adenovirus-encoded terminal protein precursor (pTP), single-stranded DNA-binding protein (DBP), and DNA polymerase (Ad

pol), and cellular proteins called nuclear factor I (NF I), nuclear factor II (NF II), and nuclear factor III (NF III). A discussion of the role of each of these proteins in adenovirus DNA replication follows.

As described above, pTP provides the primer for initiation of replication by forming a phosphodiester bond between the hydroxyl group of one of its serine residues and the phosphoryl group of dCMP, the first nucleotide in the new DNA chain (Challberg, *et al.*, 1980).

The discovery that DBP is required for replication came from the isolation of temperature-sensitive DNA mutants in the gene encoding DBP (Ensinger and Ginsberg, 1972; van der Vliet and Sussenbach, 1975; Ginsberg *et al.*, 1977; van Bergen and van der Vliet, 1983), which could be complemented by the addition of purified wild-type DBP (Kaplan *et al*, 1979; Friefeld *et al.*, 1983; Ostrove *et al.*, 1983; Prelich and Stillman 1986). *In vitro* studies have shown that DBP stimulates initiation by increasing pTP-dCMP formation three-fold (Nagata *et al.*, 1982), and is absolutely required for elongation (Field *et al.*, 1984; Prelich and Stillman, 1986). DBP is believed to stabilize the displaced strand by binding to it, and DBP has been shown to increase the processivity of the adenovirus DNA polymerase (Lindenbaum *et al.*, 1986).

The 140-kDal adenovirus DNA polymerase is required to form the pTP-dCMP initiation complex (Stillman *et al.*, 1982b; Friefeld *et al.*, 1983b; Ostrove *et al.*, 1983; van Bergen and van der Vliet, 1983). In the presence of DBP, Ad pol is a very processive polymerase, capable of synthesizing products ranging up to 40-kDal in size before dissociating from the template (Field *et al.*, 1984; Lindenbaum *et al.*, 1986). The

adenovirus DNA polymerase also has an intrinsic 3'-to-5' exonuclease activity (Field *et al* 1984).

Three cellular proteins are required for efficient adenovirus DNA replication. Nuclear factor I (NF I) is required to stimulate the formation of the pTP-dCMP complex (Nagata *et al.*, 1982). NF I binds to a specific sequence in the adenovirus origin. Nuclear factor III (NF III) is also a site-specific DNA-binding protein, and binds to a site adjacent to the NF I-binding site (Pruijn *et al.*, 1986; Rosenfeld *et al.*, 1987; O'Neill and Kelly, 1988). NF III stimulates the initiation of DNA replication, but whether NF III is absolutely required is as yet unresolved (Guggenheimer *et al.*, 1984; Pruijn *et al.*, 1986; Rosenfeld *et al.*, 1987; O'Neill *et al.*, 1988). The third cellular protein, nuclear factor II, has a type I topoisomerase activity and is required to complete elongation of full-length virus strands (Guggenheimer *et al.*, 1984a, b; Nagata *et al.*, 1983a).

The adenovirus origin of replication

Extensive deletions and substitutions within the inverted terminal repeats of adenovirus serotypes 2 and 5 have revealed three functionally distinct domains (reviewed by Challberg and Kelly, 1989; Stillman, 1989). The first 18 bp constitutes the minimum origin of replication. The 5'-terminal C residue is essential; however, the next 8 bp are required for length, not sequence, and are thought to provide a spacer function (van Bergen *et al.*, 1983; Graham *et al.*, 1989). Nucleotides 9 through 18 are absolutely required for initiation and are highly conserved among all serotypes (van Bergen *et al.*,1983; Tolun *et al.*, 1979; Stillman *et al.*, 1982a; Wang and Pearson, 1985). The second

and third domains adjacent to the minimal origin are not required for initiation, but increase the efficiency of initiation significantly. The second domain extends from nucleotide 19 through 39, and contains a highly conserved sequence between many, but not all, serotypes. The consensus sequence TGG(A/C)NNNNNGCCA within this domain is recognized and bound by nuclear factor I (Nagata *et al.*, 1983a; Rosenfeld and Kelly, 1986; Jones *et al.*, 1987). The third domain of the adenovirus origin extends from nucleotide 40 through 51, and also contains a consensus sequence AT(G/T)N(A/T)AAT recognized by nuclear factor III (Pruijn *et al.*, 1986; Rosenfeld *et al.*, 1987; O'Neill *et al.*, 1988). The adenovirus origin not only contains sequences essential for controlling replication, but it also plays an important structural role for the initiation of complementary strand synthesis. By separating the protein recognition sequences from the structural role of the origin, Wang *et al.* (1991) have demonstrated that the origin must consist of at least 31 bp of inverted sequences in order to recover genomes which were capable of participating in both phases of adenovirus replication.

Experimental strategy

Evidence for how adenovirus maintains the sequence identity of its ITRs first came from experiments performed by Stow (1982). He demonstrated that adenovirus chromosomes lacking 51 bp from their left-hand termini were infectious and capable of regenerating the missing origin sequences with high efficiency. If, however, an entire ITR was removed, the adenovirus chromosome was no longer viable. This end-repair could be mediated by three mechanisms: homologous

recombination, "panhandle" formation, or postreplicative overlap recombination. Experiments presented in this thesis, designed to study end-repair, were carried out in an adenovirus minichromosome system (Hay *et al.*, 1984; Wang and Pearson, 1985; Ahern *et al.*, 1991; Wang *et al.*, 1991). Plasmid molecules were constructed with a single adenovirus replication origin flanked by identical inverted repeats (IRs). Linearization of the plasmids exposes the origin, and the non-adenovirus repeated sequences are in an inverted orientation capable of substituting for the structural requirement of the adenovirus ITR during adenovirus replication (Fig. I.2a). Adenovirus minichromosomes can participate in both phases of adenovirus replication under appropriate conditions. Replication can initiate at the origin of the minichromosome (Fig. I.2b) and proceed by a strand displacement mechanism to create a double-stranded daughter molecule and a parental single strand (Fig. I.2c). Complementary strand synthesis on the displaced single strand can occur after forming a "panhandle" structure between the inverted sequences (Fig. I.2d), subsequent repair synthesis to regenerate a double-stranded origin (Fig. I.2e), and finally initiation of replication at the regenerated origin (Fig. I.2f) to produce the second double-stranded daughter molecule, now called the *cis* molecule (Fig. I.2h). Only the *cis* molecule is capable of amplification by further rounds of replication. This pathway exhibits several key features: (a) Only one strand, the displaced strand, from the input molecule participates in end-repair. (b) The intermediate, a panhandle structure, arises from direct intramolecular hybridization of complementary sequences (inverted

repeats) within the displaced strand. Since interacting sequences are initially single-stranded and hybridization does not involve the progressive invasion of a duplex structure, heterology surrounding the inverted repeats does not block the formation of panhandles. (c) Initially only the inverted repeats are double-stranded in the panhandle intermediate. (d) Unpaired 3' ends, but not 5' ends, are excised from the intermediate (Wang *et al.*, 1991). We have also observed the efficient removal of unpaired 3' ends from both strands of heteroduplex intermediates formed during adenovirus postreplicative overlap recombination (Ahern *et al.*, 1991). The mechanism for eliminating unpaired 3' tail is not yet known, but the potent 3'-to-5' exonuclease activity of the adenovirus DNA polymerase (Field *et al.*, 1984) may be involved. (e) The adenovirus origin is regenerated by DNA repair synthesis from the base-paired 3' end of the panhandle in the intermediate using 5' sequences as a template. Only panhandles 31 bp in size or longer can be repaired. The abrupt length dependence suggests a steric constraint. In fact, work presented in Chapter III supports this conclusion since the repair of otherwise isogeneic mismatches located within the duplex panhandle region is position- and orientation-dependent. Mismatches located at least 31 bp from the "panhandle" loop are not corrected by DNA repair synthesis proceeding away from the loop, although mismatches as close as 6 bp can be rectified by synthesis toward the loop. The molecular basis of the steric obstruction remains to be determined. (f) Subsequent rounds of adenovirus-specific DNA replication amplify the new, end-repair molecules.

Fig. I.3 summarizes the conditions required to recover minichromosomes capable of participating in both phases of adenovirus DNA replication. The minichromosome must have a structure like plasmid **a** (see also Fig. I.2) with an adenovirus origin of replication as well as identical sequences placed in a specific inverted orientation. Adenovirus replication proteins must also be supplied in *trans* by helper adenovirus DNA cotransfected with the plasmid DNA. When plasmid **a** is cotransfected with adenovirus helper DNA into human cells, incubated for 72 hours, and then the total DNA is extracted and analyzed by Southern blot hybridization with a plasmid probe, three adenovirus replication products can be identified: molecules which are the size of the input DNA, molecules which have undergone "panhandle" formation and recovered the second origin of replication (*cis*), and dimer molecules which form by end-to-end ligation during the transfection process.

Molecules which have undergone replication in mammalian cells can be identified by a method developed by Peden *et al.* (1980). The endonuclease restriction enzyme *Dpn*I cuts the site GATC only when it is methylated on both strands. Dam^+ *E. coli* hosts methylate GATC sites in plasmid DNA, but mammalian cells do not. Thus, plasmid DNA which has undergone replication in the human cells will be resistant to *Dpn*I treatment. As shown in the gel illustration in Fig. III.3, plasmid **a** transfected with helper is resistant to *Dpn*I, indicating that it has undergone replication in the mammalian cells. Only input size DNA is recovered when plasmid **a** is transfected without helper DNA. Moreover, it is sensitive to *Dpn*I treatment, indicating it was

unable to replicate in the mammalian cells. Plasmids **b** and **c** illustrate the other two requirements for an adenovirus origin of replication and inverted repeats. If a plasmid with a direct repeat (plasmid **b**), instead of an inverted repeat, is cotransfected with helper, *Dpn*I-resistant input and dimer-sized molecules are recovered, but no *cis* molecules are detected. If a plasmid lacking an adenovirus origin (plasmid **c**) is cotransfected with helper, only *Dpn*I-sensitive input-sized DNA is seen, demonstrating an inability to participate in adenovirus replication.

This assay provides a number of critical features essential for experiments in the thesis: (a) The ability to separate molecules which have only undergone strand displacement replication (input molecules), from those that have engaged complementary strand synthesis (*cis* molecules). (b) Isolation of the IR from sequences required for replication. (c) The plasmids have small genomes and are easy to manipulate by recombinant DNA techniques. (d) The IR is a polylinker which can be easily manipulated with restriction enzymes to introduce small regions of heterogeneity between the repeated sequences. (e) After transfection. the isolated DNA can be analyzed with endonuclease restriction enzymes to determine if sequence information has been exchanged between the IRs. In the following chapters, I will show that homologous recombination does not mediate the exchange of information between the IRs, but that "panhandles" and postreplicative overlap recombination both can. Since a special case was also found where postreplicative overlap recombination was unable to mediate the exchange of sequences but "panhandles" could,

these experiments provide direct and conclusive genetic evidence for "panhandle" intermediates during adenovirus DNA replication.

Sequence conversion

Evidence for gene conversion, a non-reciprocal exchange of genetic information, was first published by Lindegren (1953). Holliday (1962, 1964) and Whitehouse (1963) suggested that gene conversion occurs during recombination processes as a result of the formation of heteroduplex intermediates, which in turn are targets for subsequent excision and repair of mismatches. Initial studies of gene conversion events between allelic loci in fungi established a 50% rule; that is, approximately half of the conversion events were associated with reciprocal recombination of flanking markers. This rule has been shown more recently not to apply consistently. For example, intrachromosomal conversion between inverted duplicated HIS3 genes in yeast are not associated with reciprocal exchange (Klein, 1984). Furthermore, gene conversion events in higher eukaryotes are frequently not associated with reciprocal cross-overs. A thorough discussion of these and other examples is presented by Fink and Petes (1984). The term gene conversion has been adopted to describe a common outcome: a non-reciprocal transfer of DNA sequences, which possibly arises by more than one mechanism. The term "sequence conversion" has been limited here to describe the non-reciprocal transfer of sequences between the IRs of adenovirus minichromosome replication intermediates. An understanding of the mechanisms and enzymes involved in sequence conversion in the adenovirus minichromosome system will contribute to a general understanding of

gene conversion in eukaryotic cells. Our experiments predict the involvement of proteins encoded by both host and virus. Furthermore, unraveling the mechanism for sequence conversion on "panhandle" intermediates has revealed biologically significant consequences for adenovirus.

Frameshift mutations (small insertions and deletions)

All the sequence conversion events studied here involve three- and four-bp differences between the IRs of adenovirus minichromosomes. The addition and deletion of nucleotides account for the majority of observed mutations. The removal or addition of small numbers of nucleotides, known to occur spontaneously in many organisms, has been termed frameshift mutations (Streisinger *et al.*, 1966; Ames and Whitfield, 1966; Fitch, 1973; Roth, 1974; Stewart and Sherman, 1974; Farabaugh *et al.*, 1978; Pribnow *et al.*, 1981; Ripley and Shoemaker; 1983; Streisinger and Owen 1985). The primary DNA sequence has been shown to determine the frequency and specificity of frameshift errors; repeated sequences and quasipalindromic sequences are two well defined sources (Fitch, 1973; Stewart and Sherman, 1974; Farabaugh *et al.*, 1978; Pribnow *et al.*, 1981; Ripley and Shoemaker, 1983; De Boer and Ripley, 1984; Streisinger and Owen, 1985; Ripley *et al.*, 1986; Boer and Ripley, 1988). Yet other unexplainable complex frameshift mutations account for up to 8% of spontaneous frameshifts in bacteriophage T4 (Ripley *et al.*, 1986). *In vitro* experiments revealed that eukaryotic DNA polymerases α, β, and γ all generate frameshift mutations at an equal or lower frequency than base substitution errors (Kunkel, 1986). DNA polymerase β is the least

accurate enzyme with regard to frameshift errors, and DNA polymerase γ is the most accurate. The frequency of frameshifting by Ad pol has not yet been examined, but based on the results from other systems, it is highly likely that frameshifting occurs.

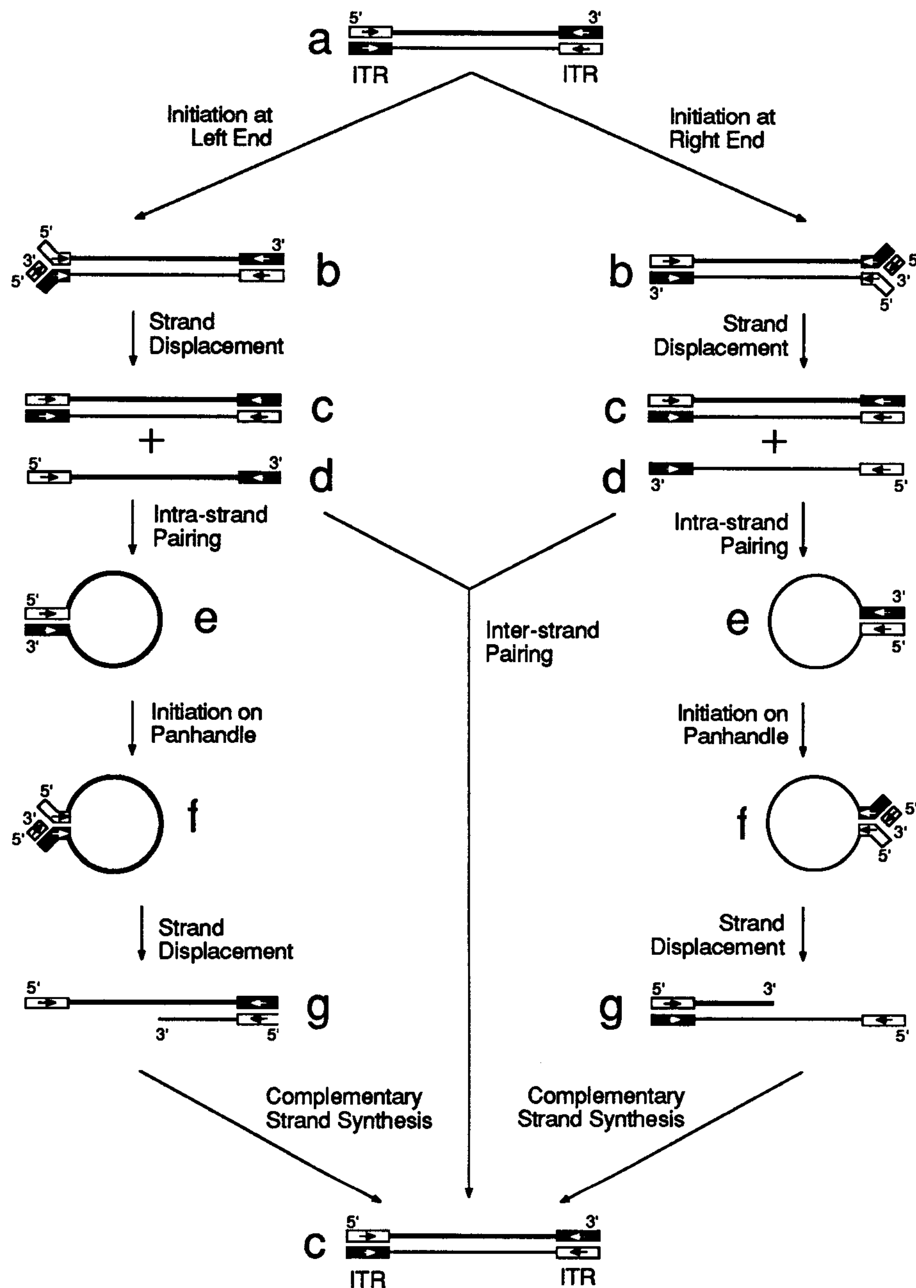


Figure I.1 The scheme for adenovirus DNA replication. Inverted terminal repeats (ITRs) are shown as open and closed boxes with black and white arrows to indicate orientation. Complementary DNA strands are depicted as thick and thin lines. The polarity of DNA strands are designated by 3' and 5'. The *cis* replication pathway involves intra-strand pairing. The *trans* pathway involves inter-strand pairing. (a) Parental double-stranded molecule. (b) Type I replication intermediate. (c) Daughter double-stranded molecule. (d) Displaced parental strand. (e) Panhandle intermediate. (f) Initiation on a panhandle intermediate. (g) Type II replication intermediate.

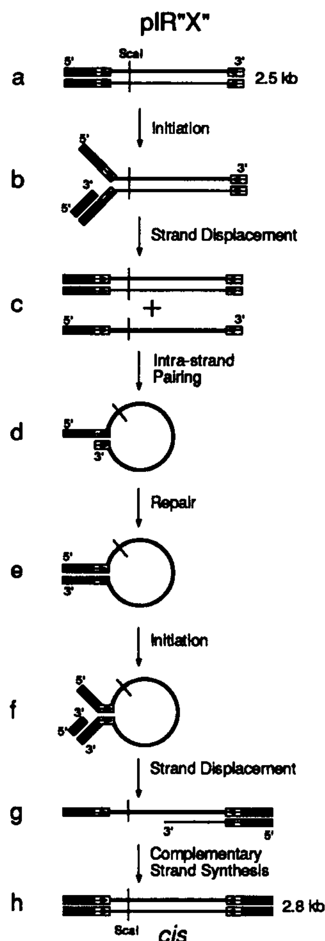


Figure I.2 Generation of a replication-proficient plasmid with two adenovirus termini from a plasmid bearing a single viral terminus. (a) *Eco*RI-linearized plasmid of 3.2-kb. The open boxes represent inverted repeat DNA and the arrows indicate the orientation. The thick lines correspond to adenovirus DNA and the thin line shows plasmid vector DNA. (b) Initiation of strand displacement replication at the single adenovirus terminus to form a type I replication intermediate. (c) Production of a daughter double-stranded molecule and a displaced parental strand at the conclusion of displacement replication. (d) Formation of a panhandle structure by base pairing between inverted repeat sequences on the displaced strand. (e) Extension of the 3' end of the panhandle molecule using adenovirus sequences as a template. (f) Initiation of complementary strand replication at the regenerated adenovirus origin on the panhandle molecule. (g) Complementary strand synthesis on a type II replication intermediate. (h) Production of a replication-proficient 3.6-kb daughter molecule with two adenovirus termini at the conclusion of complementary replication. *Eco*RI-linearized polylinker plasmids (2.5 kb) generate replication-proficient 2.8-kb daughter molecules by the same repair pathway.

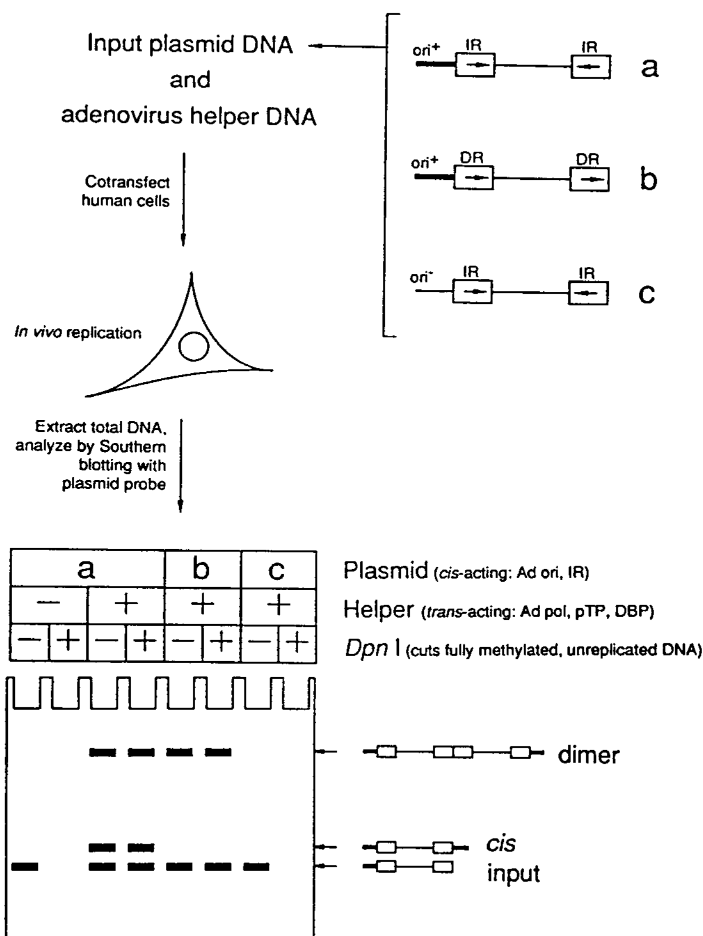


Figure I.3 The assay and requirements for detection of a replication-proficient plasmid with two adenovirus termini from a plasmid bearing a single viral terminus. The requirements include the presence of a single adenovirus origin, repeated sequences in an inverted orientation, and adenovirus helper DNA to provide viral replication proteins in *trans*. Plasmid **a** is cotransfected with helper DNA into human cells for 72 hours and the total extracted DNA is analyzed by Southern blotting with a plasmid probe. Three replication products are detected: input-sized molecules that can only participate in strand-displacement replication, *cis* molecules with a regenerated second adenovirus origin, and dimer molecules formed by end-to-end ligation during the transfection. Each band that has undergone replication in the human cells is resistant to *Dpn*I. When plasmid **a** is transfected without helper DNA, only input-sized, *Dpn*I-sensitive DNA is recovered, indicating that it did not replicate in human cells. When a plasmid with direct repeats (plasmid **b**) is contransfected into human cells, input and dimer, but not the *cis*, replication products are recovered. Plasmid **c** lacks an adenovirus origin and is unable to replicate in the human cells as seen by recovery of only *Dpn*I-sensitive input DNA.

CHAPTER II: MATERIALS AND METHODS

Reagents and enzymes

Restriction endonucleases, the Klenow fragment of *Escherichia coli* (*E. coli*) DNA polymerase I, bacteriophage T4 DNA ligase, calf intestinal alkaline phosphatase, ultrapure agarose, and phenol were purchased from Boehringer Mannheim Biochemicals, Bethesda Research Laboratories (BRL), or International Biochemicals, Inc. (IBI). Pronase was purchased from Calbiochem. Lysozyme and ribonuclease were from Sigma. All the enzymes were used as recommended by the manufacturers. [γ-^{32}P]dGTP (specific activity 3,000 Ci/mmol) and [α-^{32}P]ATP (specific activity 6,000 Ci/mmol) were purchased from New England Nuclear.

Bacteria and plasmids

Plasmids were maintained in *E. coli* JM107. These cells were made competent by a modified method of Hanahan (1983). JM107 colonies were grown on a minimal agar plate which included 5 mg/l thiamine. A single colony was used to inoculate 10 ml of YT broth (per liter: 5 g yeast extract, 10 g Casein enzyme hydrolysate, and 5 g NaCl) which was placed overnight in an incubator shaker at 37°C. The 10 ml culture was used to inoculate a liter of YT broth, and the bacteria were grown until the A_{560} was 0.6. The *E. coli* were chilled on ice for 10 min and then centrifuged at 3,500 x g for 5 min. The pellet was resuspended in 1/3 volume ice cold 0.1 M $CaCl_2$ by gently swirling and then held on ice for 4 hr. The bacteria were centrifuged as above and then resuspended in 1/25 the original volume into 0.1 M $CaCl_2$ containing 15% glycerol and held on ice overnight. The bacteria were

then pipetted into sterilized microfuge tubes (200 µl of cells) and frozen at -70°C until needed. Plasmid DNA was introduced into competent *E. coli* JM107 by the following method: 10 ng of supercoiled DNA or 0.1-0.6 µg of ligated DNA was added to 200 µl of competent JM107 cells and the mixture was incubated on ice for 1 hr. The cells were then heat shocked at 42°C for 90 sec, then transferred to 800 µl of SOC (2% Bacto ryptone, 0.5% Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) medium and shaken at 37°C for 1 hr, and finally 230 µl was spread onto a selective plate.

A series of plasmids was constructed which contained a single adenovirus origin flanked by identical sequences such that when the plasmids were linearized with *Eco*RI restriction endonuclease, the adenovirus origin was exposed and the identical sequences were in an inverted orientation (IR). Two sets of plasmids were made which consisted of 10 plasmids in each set. Each set of plasmids contained small sequence differences between the ITR, and the two sets differed by the length of the ITR. The pIR series had a 51-bp ITR which was composed of the polylinker from pUC19, while the pLIR series had a 232-bp ITR which also included the pUC19 polylinker plus juxtaposed sequences.

The construction of the pIR series started with the use of T4, a plasmid which was constructed by Kai Wang; the left 338 bp of the adenovirus type 2 genome was inserted into pUC19 and replaced the sequences between the *Eco*RI to *Ssp*I site. Three intermediate plasmids (T4-*Sph*I, T4-*Kpn*I, T4-*Xba*I) were made from the T4 plasmid, by deleting out the *Sph*I, *Kpn*I, restriction sites and filling in the *Xba*I

site. This was done by cutting T4 with the corresponding endonuclease restriction enzyme (*Sph*I, *Kpn*I, or *Xba*I) and then treating the linearized plasmid with Klenow, which will remove 3' overhangs and fill in 5' overhangs. Ten new test plasmids were constructed using the three intermediate plasmids plus T4. All four of the previously mentioned plasmids were cut with *Ssp*I and *Hin*dIII and the 392-bp (base pair) fragment was isolated by using low melt agar and phenol extraction. This fragment consists of the adenovirus origin plus the pUC19 polylinker. The same four intermediate plasmids were also cut with *Eco*RI and *Ssp*I, but this time the backbone (2112 bp) was isolated which contained only pUC19 sequences including the polylinker. Next the large and small fragments were put together such that the adenovirus sequences were flanked by the two polylinkers. When the new plasmids were linearized with *Eco*RI this exposed the adenovirus origin sequences, which were followed by the 5' polylinker, and the 3' polylinker was then at the opposite end of the plasmid and in an inverted orientation to the 5' polylinker. The plasmids were named according to the polylinker and the restriction site that was altered. For example, the molecule with the *Sph*I site deleted from the 5' polylinker was called pIR5'S, the plasmid with a perfect ITR was named pIRHin, and finally the plasmids with both restriction endonuclease sites removed were called pIRDS.

The construction of the pLIR series employed the above test plasmids plus plasmid pLIRST. The plasmids with the altered *Kpn*I and *Xba*I restriction endonuclease sites were made from cutting pLIRST with *Pst*I and purifying the large backbone (2,257 bp). This

was then ligated to the purified small fragment (418 bp) from the corresponding pIR plasmid also cut with *Pst*I. PLIRH was made in the same way by ligating the large backbone from pLIRST to the small fragment from pIRHin. *Sph*I long ITR plasmids were made by partial digestion of pLIRH with *Sph*I followed by treatment with Klenow and then religation to create the 3' and 5' mutants. The clone without any *Sph*I (pLIRDS) sites was constructed by cutting pLIRH to completion with *Sph*I and then treated as above.

A third set of plasmids with two altered restriction endonuclease sites in a single plasmid were constructed. These plasmids were all made from pLIR5'K. Partial and complete digestion of pLIR5'K were carried out using *Sph*I and *Xba*I followed by treatment with Klenow and then ligation to create the six possible combinations of nonidentical ITRs.

Growth and purification of adenovirus DNA

HeLa S_3 cells, used to propagate adenovirus, were grown as suspension cultures in ESM medium (GIBCO) containing 5% fetal bovine serum (FBS). When the HeLa S_3 cells reached a density of 5 x 10^5 cells\ml they were pelleted at room temperature at 2,400 x g in a Beckman J6 centrifuge. The cells were resuspended in 1/10 of the final infection volume and infected with adenovirus at 10 to 50 particles per cell. The infected cell suspension was incubated for 1 hr as a spinner culture to allow the virus to attach to the HeLa S_3 cells. Then the infected suspension culture was diluted to the final volume with the complete ESM medium (5% FBS) and incubated for 48 hr. To harvest the virus, the spinner cultures were first pelleted as described above,

then resuspended into 10 ml sterile PBSd (137 mM NaCl, 2.7 mM KCl, 1.5 mM KH_2PO_4, and 8 mM Na_2HPO_4, pH 8.0) per liter of infected cells. The cells were burst open by freezing the sample for 5 min in liquid nitrogen and then thawing at 37°C; this was repeated 3 times followed by sonication on ice for 3 min. The cells were spun down at 12,100 x g. The supernatant was saved and the cells resuspended in 5 ml PBSd per liter and sonicated and then pelleted once again. The supernatants were pooled and then extracted with 2/3 volume Freon (1,1,2-trichlorotrifluoroethane) by gently vortexing for 5 min at room temperature. The mixture was separated by centrifugation at 12,100 x g for 10 min at room temperature. The top layer from the Freon extraction was added to a CsCl gradient. The gradient was layered from bottom to top with 5 ml of 0.6 g/ml CsCl in 20% glycerol-PBSd, 15 ml 20% glycerol-PBSd. The crude cell lysate was added to the top of the gradient, and then mineral oil was added for balancing. The gradient was centrifuged for 90 min at 24,000 rpm at 5°C in a Beckman SW28 rotor and then the bottom band was collected and dialyzed against TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). The sample was transferred to a clean tube and treated with ribonuclease for 15 min at 37°C, then Pronase was added 1 mg/ml and 0.2% sodium dodecyl sulfate (SDS). This mixture was incubated overnight at 37°C. The DNA was then precipitated with 0.8M LiCl and one x volume isopropanol. The DNA was spun at 5,600 x g for 5 min for the previous step and for the steps to follow. The alcohol was decanted and the pellet resuspended in 400 µl TE buffer. The DNA was extracted twice with phenol/Tris and once with phenol/chloroform, then precipitated as

described above and dried. Finally, the DNA was resuspended in TE buffer and checked by gel electrophoresis.

DNA transfection

Linear plasmid DNA (5 µg) with a single adenovirus origin was transfected into the 293 cell line by a modified procedure of the DNA-calcium phosphate coprecipitation technique (Graham and van der Eb, 1973). Adenovirus DNA (5 µg) was cotransfected with the plasmid DNA to provide the adenovirus-encoded replication proteins to drive replication of the plasmid DNA by an adenovirus mechanism. The transfection was performed by gently mixing 125 µl of 0.5 mM $CaCl_2$ solution containing the plasmid and helper DNA with 125 µl of 2X HBS buffer; 1X HBS contains 280 mM NaCl, 50 mM Hepes, and 1.5 mM Na_2HPO_4, pH 7.05. After 20 min at room temperature, the mixture was added to the cells followed by incubation at 37°C in a CO_2 incubator for another 20 min. Then 5 ml of medium was added without disruption of the monolayer and incubated for 6 hr. As recommended by Frost and Williams (1978), the cells were subjected to a glycerol shock (15% glycerol in 1X HBS buffer for 2 min) to enhance the efficiency of DNA uptake. Cells were harvested after 66 hr at 37°C in the CO_2 incubator in order to isolate the total DNA.

DNA extraction from transfected cells

Three days after transfection, the total DNA was extracted from the 293 cells (Hirt, 1967; Wilkie, 1973). The monolayer of cells was scraped from the plate and then pelleted by spinning at 700 x g for 5 min. The cells were washed once with 1 ml of PBSd buffer and pelleted as above. The cells were then resuspended in 50 µl of Pronase (1

mg/ml), plus 0.4 ml of lysis solution (0.6% SDS, 10 mM EDTA, pH 8.0) and incubated at 37°C overnight. The lysate was then precipitated with 50 µl of 8 M LiCl and 1 ml of 95% ethanol and then dried. The recovered pellet was phenol extracted three times, twice with phenol/chloroform (1:1), and once with chloroform/isoamyl alcohol (24:1). The DNA was precipitated one more time as described above and then resuspended into 200 µl water.

Southern blot analysis

For detection of adenovirus replication products from extracted transfected cells Southern blot analysis using alkaline transfer was employed (Chomczyhske and Qasba, 1984; Reed and Mann, 1985). First the DNA was separated by 0.8% agarose gel electrophoresis in TBE buffer (89 mM Tris-boric, 89 mM Boric acid, 2 mM EDTA, pH 8.5) at 50 mA for 16 hr and then transferred from the gel onto a "Beta-Probe" nylon membrane (Bio Rad) in 0.4N NaOH overnight.

Purification of the *cis* replication product from the input DNA

Agarose gel electrophoresis (0.8%) was employed to separate the *cis* replication product from the input DNA. To detect the adenovirus replication products Southern blot analysis was required. Each purification gel was divided into two parts; one part was used to make an autoradiogram to serve as a template to locate the desired bands in the second part of the gel. A single lane was loaded with 16 µg of total DNA extracted from transfected 293 cells to serve as template while five lanes of the same DNA were run simultaneously and saved for the purification. Bands of agarose were cut out and minced with a razor blade then placed into a 5 ml Falcon tube and filled with ultrapure

distilled phenol. Each tube was vigorously vortexed for 10 min and then frozen at -20°C overnight. The next day the samples were centrifuged at 5,600 x g for 10 min, then vortexed a second time for 3 min, and frozen once again for 2 hr at -20°C. The samples were then extracted 3 times with phenol/Tris, and treated with phenol/chloroform once. Each sample was then brought up to a total volume of 2 ml with distilled water, 10 µg of tRNA was added, and then the sample was precipitated with 200 µl of 8 M LiCl and 2 ml isopropanol for 1 hr. The samples are spun for 10 min at 5,600 x g, then the alcohol supernatant was decanted, and the pellet was resuspended in 300 µl dH_2O. The sample was precipitated once again, then treated with a 70% ethanol wash, and finally dried.

Restriction endonuclease cleavage of purified input and *cis* bands

Each purified band was resuspended into 49 µl of distilled water, and 1 µl (1 ng/ml) of cleavage standard (see below) was added. The following was pipetted into a clean microfuge tube: 25 µl of the resuspended sample, 1 µl of 1 mM spermidine, 4 µl of Boehringer Mannheim recommended buffer, and 32-40 units of restriction endonuclease. The remaining 25 µl were saved and run beside the digested sample in the assay gel. The cleavage standard was made by cutting the plasmid T4 (made by Kai Wang) with *Dra*I. This creates a linear fragment of 1,739 bp with the polylinker 565 bp from the end. This cleavage standard was used for all *Sph*I, *Kpn*I and *Xba*I restriction analysis. The cleavage standard used for *Bgl*I restriction analysis was constructed by linearizing plasmid T4 with *Eco*RI. This

creates a 2,450-bp cleavage standard with the *Bgl*I site 1,424-bp from the end.

Radioactive labeling of plasmid DNA by random priming

The random priming method described by Feinberg and Vogelstein, (1983, 1984) was used to label CsCl-banded pUC19 with [α-^{32}P]dGTP. The reaction mixture was incubated overnight at room temperature and consisted of: 27 μl DNA (100 ng, heat denatured), 10 μl OLB reagent mix (250 mM Tris-HCl, pH 8.0; 250 mM $MgCl_2$, 0.1 mM each of dTTP, dATP, and dCTP, 1 M Hepes, pH 6.6, and hexadeoxyribonucleotides at 27 A_{260} units/ml), 6 μl bovine serum albumin (5 mg/ml), 5 μl [α-^{32}P]dGTP (50 μCi, 3,000 Ci/mmol), and 2 μl Klenow fragment of DNA pol I (2-5 units). After synthesis, unincorporated nucleotides were removed by a Sephadex G50 column as described by Maniatis (1982). The spins were carried out in an IEC clinical desk top centrifuge at room temperature at setting 4.

End-labeling of oligonucleotide probes

A 15-mer oligonucleotide identical to the left origin of adenovirus type 2 was synthesized by Dr. R.H. McParland in the Central Services Laboratory of the Center for Gene Research and Biotechnology at Oregon State University using an Applied Biosystems automated DNA synthesizer. The wildtype oligo (WT-oligo) has the sequence 3'-GATAAGATTATTATA-5'. The oligonucleotide was end-labeled with [γ-^{32}P]ATP (specific activity 6,000 Ci/mmol) by using T4 polynucleotide kinase as described by Maxam and Gilbert (1980). The oligomer molecules were then purified by chromatography through Nuclean-25 spin-columns (IBI).

DNA hybridization

Hybridization with the radiolabeled pUC19 probe was carried out as follows. The blotted membranes were rinsed with wash solution (2x SSC, 0.1% SDS, pH 7.0) until the pH was below 8.0; 1X SSC consists of 150 mM NaCl and 15 mM sodium citrate, pH 7.0. After neutralization, the filter was incubated in 50 ml of prehybridization solution (5X SSC, heat-denatured salmon sperm DNA at 0.25 mg/ml, 50 mM sodium phosphate pH 6.7, 50% formamide, 0.5% non-fat dry milk, and 1% SDS) for 4 hr at 41°C. Hybridization was carried out by incubating the prehybridized membrane with radiolabeled pUC19 (specific activity > 10^8 cpm/μg) overnight at 41°C in 12.5 ml of hybridization solution (5x SSC, heat-denatured salmon sperm DNA at 0.1 mg/ml, 20 mM sodium phosphate pH 6.7, 10% dextran sulfate, 50% formamide, and 1% SDS). After hybridization, the membrane was rinsed three times at room temperature with wash solution, rinsed for one hr at 65°C, then after a change of the wash solution, rinsed again for four hr at 65°C. Autoradiography was performed with X-ray film (Kodak X-OMAT XAR-5) for 8 to 120 hr using an intensifying screen at -20°C.

Hybridization with the end-labeled oligonucleotide probe was performed as modified from the method of Zeff and Geliebter (1987). The filter was neutralized as described above. It was then incubated in 50 ml of prehybridization solution (5x SSC, 20 mM sodium phosphate pH 6.7, 10X Denhardt's solution, 7% SDS, and heat-denatured salmon sperm DNA at 0.1 mg/ml) for 4 hr at 29°C (5°C to 7°C below the dissociation temperature). 1X Denhardt's solution contained Ficoll at 0.2 mg/ml, polyvinylpyrrolidone at 0.2 mg/ml, and bovine serum

albumin (Pentax, Fraction V) at 0.2 mg/ml. The dissociation temperature the oligonucleotide probe was determined by using the formula T_d(°C) = 2(A + T) + 4(G + C) where A, T, G, and C are the number of residues in the oligonucleotide (Suggs *et al.*, 1981). T_d for the WT-oligo was 36°C. The filter was then placed in 12.5 ml of the hybridization solution (5X SSC, 20 mM sodium phosphate pH 7.0, 10X Denhardt's solution, 7% SDS, 10% dextran sulfate, and heat-denatured salmon sperm DNA at 0.1 mg/ml), and incubated at the same temperature overnight. After hybridization the filter was rinsed three times with wash solution for 15 min at room temperature. If necessary, the membrane was washed again at a higher temperature for more stringent probe specificity.

CHAPTER III: SEQUENCE CONVERSION OCCURS ON PANHANDLE INTERMEDIATES DURING ADENOVIRUS MINICHROMOSOME REPLICATION

INTRODUCTION

The first indication that adenovirus contained an inverted terminal repeat came from electron microscopic studies of alkali denatured, formamide renatured genomes, where circular single-stranded molecules were seen (Garon *et al.*, 1972). In their initial report Garon *et al.* (1972) were unable to identify any "panhandle" projections from the circularized DNA molecules. Subsequently, Garon *et al.*, (1975), identified small projections when using a similar procedure with adenovirus type 18. It is believed that in the other serotypes tested that the "panhandle" projection is simply too short to be seen. The first suggestion for a role of the adenovirus ITR in replication came from Daniell (1976) when studying the structure of incomplete genomes. The proposal was made that displaced single-strands, could form a duplex between the ITR creating a "panhandle" structure for the initiation of replication. This mechanism could also explain the formation of some of the defective particles generated during adenovirus infection. The model was embraced and expanded upon by Lechner and Kelly (1977), who did an electron microscopic study of replicating adenovirus genomes even though they did not recover single-stranded circular molecules. Further suggestive evidence for "panhandle" formation during replication came from experiments conducted by Stow (1982). He demonstrated that adenovirus chromosomes lacking 51 bp from their left-hand termini were infectious and capable of regenerating the missing origin

sequence (end-repair). When an entire inverted terminal repeat (ITR) was removed the adenovirus chromosome was no longer viable. A number of papers using an adenovirus minichromosome system to examine origin sequence requirements also supports the hypothesis of "panhandle" formation during replication (Hay *et al.*, 1984; Wang and Pearson, 1985; Wang *et al.*, 1991). Characterization of adenovirus progeny from a mutant with a duplicated inverted terminal repeat also provided support for "panhandle" formation (Haj-Ahmad and Graham, 1986). None of the experiments described above provided definitive evidence for "panhandle" formation. Both homologous recombination or postreplicative recombination (Ahern *et al.*, 1991) could occur between the remaining base pairs of the ITR, and generate the same outcome. Lippe and Graham (1989) provided some evidence that homologous recombination did not mediate the recovery of a virus, with identical ITR, from a hybrid virus with an ITR that differed in the first 7 bp. Their analysis of plaques revealed pure populations of virus with identical ITRs, while the reciprocal recombinant was never found within the same plaque. One further piece of work suggests that replication of the displaced single-stranded molecules are initiated off duplexes formed by either inter- or intramolecular interactions between the ITR. *In vitro* experiments show that the same proteins are required for strand displacement replication as are for single-strand DNA synthesis (Leegwater *et al.*, 1988). The work presented in this chapter absolutely rules out the possibility that homologous recombination mediates the retrieval of lost bases within the ITR. Furthermore, molecules like the pIR series used in this chapter, when

assayed solely to detect intermolecular sequence conversion during postreplicative recombination, were shown to be unable to engage in that process (Chapter IV). In this chapter I show that the pIR series can undergo sequence conversion; therefore, when "panhandles" can form, they will form and are used for replication processes.

RESULTS

Marker segregation during adenovirus minichromosome replication

Failure to identify "panhandle" intermediates from replicating adenovirus chromosomes by electron microscopy has led researchers to develop other assays to substantiate the existence of such structures. Most of the work has involved using either the intact adenovirus genome or an adenovirus minichromosome system where the molecules lack some sequences from one of the ITRs. The limitation of these experiments has been the inability to rule out homologous recombination or postreplicative recombination (Ahern *et al.*, 1991) in their assays for measuring "panhandle" formation (Stow, 1982; Hay *et al.*, 1984; Wang and Pearson, 1985; Wang *et al.*, 1991; Haj-Ahmad and Graham, 1986; Lippe and Graham, 1989). To distinguish between these possible mechanisms I constructed two series of plasmids with markers in the IR that could be followed in an adenovirus minichromosome system (Hay *et al.*, 1984; Wang and Pearson, 1985).

The plasmids contain a single adenovirus origin which is flanked by identical sequences (Fig. III.1). When the plasmids were linearized with *Eco*RI the origin is exposed and the identical sequences were arranged as an adenovirus ITR, but with only a single origin of replication. The two series of plasmids differ in the length of the IR;

one set has a 51-bp IR and is designated the pIR series, while the pLIR set has a 229-bp IR. Each plasmid contains an easily manipulated polylinker as part of the IR. The four central nucleotides of the *Sph*I site have been removed from the 5' or the 3' polylinker, and were then named accordingly: pIR5'S has the *Sph*I site deleted from the 5' polylinker. Cotransfection of human cells with each plasmid and adenovirus helper DNA into human cells, extraction of the DNA, and finally Southern blot analysis revealed the presence of two bands. Band **b** is input-sized DNA which can only participate in strand displacement replication, and band **a** (*cis* DNA) consists of molecules that have end-repaired and now replicate like adenovirus chromosomes. All of the original molecules contained only one *Sph*I site, thus endonuclease restriction analysis of the replication products should reveal if information was exchanged between the IRs. Fig. III.2 illustrates the two basic types of mutants constructed; the 5' mutants with the restriction site deleted from the IR nearest the origin (A; marked by the X), or deleted from the 3' IR (B), and the products generated from restriction endonuclease digestion (Cut). Each molecule was labeled **a**, **b**, or **c** to indicate the position it will migarate to in the gel. After restriction endonuclease digestion of the *cis* products the 5' and the 3' mutants could potentially have molecules without the restriction site and will be retained at position **a** (marked with two Xs). After endonuclease restriction digestion position **b** will include the 5' mutant input; the restriction site is in the IR furthest from the origin, so only a few base pairs are removed. Also *cis* molecules for both the 5' and 3' mutants with only one restriction site

will migrate to position **b**. Position **c** is the location to which all *cis* product from both the 5' and 3' plasmids with two restriction sites will migrate. In addition, the input 3' mutants with the restriction site in the 5' ITR will migrate to position **c**. As can be seen, a number of sequence conversion events can be monitored from the restriction pattern. A *Sph*I-sensitive standard has been added to each sample to determine whether digestion goes to completion.

Comparison of the molecules treated with *Sph*I in lane 2 of Figs. III.3, III.4, III.5 and III.6 show that pLIR5'S generated minichromosomes without any *Sph*I sites (band **a**) while both 3' mutants and pIR5'S did not. Further analysis of the endonuclease restriction pattern for the 3' mutants is limited, since input (band **b**) molecules when cut migrate to position **c**, which covers up any minichromosomes that have obtained a second *Sph*I site (Fig. III.2). Furthermore, higher oligomers, which form by end-to-end ligation during the transfection, migrate to position **b,** further obscuring the results. The situation is slightly different for the 5' mutants; the input DNA was expected to appear to be resistant to *Sph*I since the site is only 41 bp from the 3' end of the molecule. Thus, the appearance of molecules which are doubly sensitive to *Sph*I cleavage (band **c**; Figs. III.5 and III.6; lane 2) appeared to indicate that the second *Sph*I site was restored. To acquire a clearer picture of the composition of each band **b** and **c**, bands **a** and **b** were gel purified. This was accomplished by electrophoresis of the transfection product; band **a** and band **b** were individually cut out of the gel, purified, then treated with the restriction enzyme, and finally run on the gel a second time (lanes 3, 5, 7, 9). As a

control for the appearance of artifactual bands, each gel was probed with two different probes: first , the plasmid backbone (pUC) of the minichromosome, and second an oligonucleotide (oligo) complementary to the very end of the adenovirus origin. Band **c** was not detected by the oligo probe since *Sph*I cleavage removed the adenovirus origin sequences from the molecule (lanes 5, 6, 9 and 10).

Analysis of pIR3'S revealed that some of the input-sized molecules (band **b**) had become resistant to *Sph*I, while none of the molecules from band **a** had lost or gained an *Sph*I site (Fig III.3, lanes 4 and 8). The input molecules were expected to migrate to position **c**, since the *Sph*I site is approximately 400 bp from the end. All the cleavage standard had cut to completion suggesting that the resistant band was correct. As a further control pIR3'S was transfected without helper DNA, after 72 hr the total DNA was purified and analyzed by Southern blot analysis (Fig. III.7). Only input-sized DNA (lane 1, band **b**) is recovered and is completely sensitive to sphI (lane 2, band **c**), and *Dpn*I (lane 3). *Dpn*I sensitivity indicates that the plasmid did not participate in replication while in the human cells (the *Dpn*I assay is discussed in the Introduction). One further experiment was done with pIR3'S; gel-purification and restriction analysis of the dimer DNA was carried out (Fig III.8). All of the purified dimer migrated to the expected lower position, indicating that both *Sph*I sites remained intact. None of the dimer DNA migrated to the location where the monomer input DNA resides. This indicates that only the symmetrical dimers were purified. During replication of the symmetrical dimers, displaced single-stranded foldback molecules are created, which

migrate as input-sized DNA (Hu, 1990). The purification of the dimer demonstrates that the dimer foldback input-sized molecules do not undergo sequence conversion and so do not contribute to the *Sph*I resistant input DNA.

Detailed analysis of pLIR3'S revealed a different set of outcomes. All the input molecules were sensitive to *Sph*I (Fig. III.4, lane 4, band **c**), and all the molecules purified from band **a** had two *Sph*I sites (lane 8, band **c**).

Plasmid pIR5'S provided still a different cleavage pattern; a mixture of sensitive and resistant molecules comprised the input band **b**, while molecules with either one or two *Sph*I sites resided in band **a** (Fig. III.5, lanes 3,4,7 and 8).

Finally, a fourth, and unique, composition of bands was uncovered by analysis of pLIR5'S minichromosomes (Fig. III.6). All the input-sized DNA (lanes 3 and 4, band **b**) was resistant to *Sph*I cleavage, while band **a** was composed of a population of three types of molecules; minichromosomes with no *Sph*I site, molecules with a single *Sph*I site, and also a small sample of molecules with two *Sph*I sites (lanes 5 and 6).

Marker segregation is not mediated by homologous recombination

The endonuclease restriction pattern of the mutants revealed lenght and orientation dependent constraints that are not consistent with a model of homologous recombination for the segregation of markers placed in the IR. Fig. III.9 diagrams the predicted outcomes if homologous recombination mediated the process. Summarized in Fig. III.9A are the predicted number of crossovers and conditions for

the segregation of markers of the 3' mutants, and for the 5' mutants in Fig. III.9B. The data for the 3' mutants predict a single cross versus a triple cross required to generate the observed outcomes , and the single cross is dependent on the length of homology of the flanking region. In contrast, the 5' mutants would require both single and triple crossovers, and the single crossover depends on length b and c but the triple crossover does not. Moreover, if a triple crossover can occur between 5' mutants, the identical cross (a, b, and d or a, c, and d) should be able to occur between the 3' mutants. Clearly homologous recombination is not the mechanism by which these molecules were formed.

"Panhandle" intermediates occur during adenovirus replication

The segregation of the markers in the IR can be accounted for by sequence conversion on "panhandle" intermediates formed during adenovirus replication. The possible scenarios for sequence conversion of the 3' mutants is shown in Fig. III.10. The IRs differ by 4 bp; the bulge represents the sequences for the *Sph*I restriction endonuclease site. The variable distance between the pLIR and pIR series of constructs is 184 bp and six bp, respectively. In the diagram, end-repair is depicted as the first event. An example will be shown for the 5' mutants where the origin must be added before sequence conversion can be completed. The mechanism, leading to sequence conversion, would require an endonuclease to cleave either within the bulge or within the strand opposite the bulge. Following the pathway where endonuclease cuts in the strand opposite the bulge, a gap would be created. The adenovirus polymerase is known to be able to add

deoxyribonucleotides to 3' ends of gapped DNA, (Lichy *et al.*, 1982; Longiaru *et al.*, 1979); thus, it could be involved from step c to d. And a ligase would be required to generate a continuous single stranded intermediate e. Only when X was 184 bp did we recover molecules that would be created by following this pathway, but not when X was six bp (compare Figs. III.3 and III.4, lane 8, band **c**). Previously, Wang et al., (1991) showed that end-repair on molecules with an IR of 31 bp or longer occurred. The current experiments indicate that gap filling requiring polymerization away from the "panhandle" loop is unable to take place when there is only a six bp duplex preceding the loop. Next, the 3' mutants are never resistant to *Sph*I (Fig. III.3 and III.4, lane 2, band **a**); the final molecule would have gone through an intermediate structure like j. This would require removal of the bulge, and when we examine the possible pathways we gain some insight into the ongoing process. Endonuclease action within the loop would generate both a 3'- and a 5'-overhang. To follow through path f to j only a 5'-exonuclease would be required plus a ligase, although a 3'-exonuclease may also function on the substrate, it is not required. After endonuclease action, replication could initiate at the origin (g) and displace the 5'-end of the strand (h') producing a "panhandle" intermediate now with only a 5'-overhang. A 5'-exonuclease would be required to remove the overhang to create intermediate i and then ligation would complete the repair process to generate intermediate j. While there is no evidence for either of the two 3' *Sph*I mutants participating in pathway f-j, there is evidence that sequence conversion attempts on pIR3'S (X=six bp) led to the creation of molecule n.

Molecule n is a molecule which has endswitched at the *SphI* site, is the size of the input DNA, and is no longer sensitive to *SphI* cleavage as is the original input DNA (Fig. III.3, lane 4, band **b**). Endswitched molecules are not recovered when X = 184 bp (Fig. 2, lane 4, band **b**); this indicates that molecule n does not result from strand displacement replication of intermediate h (h-m), but instead it appears to result from the lack of stability of the six-bp duplex created after endonuclease action in molecule f, thus opening up to molecule k. Since molecules with the bulge removed are not recovered on the 3' mutants, and endswitched molecule are recovered when X=six bp, this suggests that some aspect about structure h blocks completing the removal of bulges. Molecule k can be successfully turned into the double-stranded endswitched molecule by two possible mechanisms. One involves replication initiating at the origin and displacing the remaining 5'-end, this creates molecules m and h'. Alternatively, if a 3'-exonuclease acted on substrate k only molecule m would be created. Molecule m is then the substrate to generate the endswitched molecule n. The lack of stability of a six bp duplex also predicts that attempts at sequence conversion opposite the bulge, which creates intermediate c, will result in the production of input-sized single-stranded molecules q, and small double-stranded molecules q'. Further evidence for endswitched molecules when X=six bp arises when the 5' mutants are examined in detail.

The distance X is the same for the 5' mutants, and the possible pathways for sequence conversion are presented in Fig. III.11. The enzymes required to proceed from intermediate b to e include: an

endonuclease to cleave within the strand opposite the bulge, a polymerase to gap fill, and a ligase to repair the resulting nick. Both mutants with the long and short IR do generate the final molecule with two *Sph*I sites, indicating that pathway b to e could operate (Fig. III.5, lane 8; Fig. III.6 lane 6, band **c**). Analysis of the purified bands showed that input-sized DNA (Fig. III.5, lanes 3 and 4) from pIR5'S contained DNA that was both resistant and sensitive to *Sph*I, this is in contrast to pLIR5'S (Fig. III.6, lanes 3 and 4). The expectation had been that all the input-sized molecules would appear to be resistant to *Sph*I since the original plasmid had its *Sph*I site only 41 bp from the 3' end. The input molecules (band **b**) which migrate to position **c** after digestion are coming from molecules which have endswitched (c-n), as a result of the lack of stability of the remaining six bp duplex, following the creation of a gap. As stated above, this was also observed for pIR3'S. Only pLIR5'S and not pIR5'S produced minichromosomes without any *Sph*I sites (Figs. III.5 and III.6, lanes 1 and 2, band **a**). This observation helps to predict which pathway is followed by the mutants, pathway f-k, f-l, or f-i. Also, data from the 3' mutants suggest that it is unlikely that step f-k, or f-l is functioning. If loop removal could proceed through pathway f-k on the 5' mutants, then one would predict that loop removal on the 3' mutants would be successful, but its not (Fig. III.10, b-j). Also, there is no reason why step f-k should not function for pIR5'k. Pathway f-l requires strand displacement replication on intermediate l, and with the 3' mutants there is no evidence that such a mechanism could act on a similar substrate (Fig. III.10, h-m). Pathway g-i does predict that pIR5'S would be unable to

generate molecule j. This pathway requires that end-repair take place before sequence conversion can be completed. To proceed through this pathway, replication can initiate at the origin which will result in a displaced double-stranded molecule (h') and a panhandle intermediate h with a 3'-overhang. Removal of the 3'-overhang by an 3'-exonuclease (i) and then polymerization to the end of the molecule creates an intermediate without any *Sph*I sites (j). Intermediate h, is the structure which is predicted to prevent recovery of a resistant molecule to *Sph*I when X = six bp. As was observed with the 3' mutants gap filling did not occur when polymerization was required away from the "panhandle" loop at only six bp away. Previously Wang *et al.* (1991) demonstrated that end-repair required that the minichrosomes have a 31 bp IR. A very interesting result about pIR5'S is that it does confirm that polymerization away from the "panhandle" loop requires more than six bp, but polymerization towards the "panhandle" loop (c-e) with X=six bp does take place. This indicates that an asymmetry for polymerization towards and away from "panhandle" loops exists, which will be discussed in detail.

DISCUSSION

Markers in the IR of adenovirus minichromosomes were found to undergo a non-reciprocal exchange of sequence information, whereby sequences within one IR were replaced with a copy of the sequences from the other IR. This exchange of information did not occur as a result of homologous recombination (Fig. III.9). Later I show a special case was found where sequence conversion did not take place during intermolecular hybridization of two displaced strands, but

could when "panhandle" intermediates could mediate the sequence exchange. These experiments provide direct and conclusive genetic evidence for "panhandle' intermediates during adenovirus replication.

The remaining text will include a discussion of the following: (i) A special case of sequence conversion without a contribution of information exchange during postreplicative overlap recombination. (ii) Enzymes predicted to function during sequence conversion. (iii) Limitations of sequence conversion. (iv) the IR used as template is different for each mutant. (v) The consequences of unsuccessful sequence conversion attempts.

Marker segregation is not mediated solely by postreplicative overlap recombination.

Experiments, discussed in detail in Chapter IV, were carried out using the wild-type minichromosomes, with two *Sph*I sites, co-transfected with the double mutants (pIRSD and pLIRSD), which have no *Sph*I sites. Under these conditions only intermolecular hybridization of displaced strands will form heteroduplexes between the IRs. As described it Chapters I and IV, intermolecular hybridization of displaced strand occurs during adenovirus replication, and has been called the *trans* pathway. Ahern *et al.* (1991) discovered this process by studying postreplicative overlap recombination in an adenovirus minichromosome system. The interesting result was that the pLIR minichromosomes did undergo sequence conversion while the pIR molecules did not. Yet, as shown above, pIR5'S (Fig. III.5, lanes 7 and 8) produced progeny with two *Sph*I sites, demonstrating that this pIR mutant can engage in sequence conversion. These experiments indicate that even when sequence conversion between the

two displaced strands is not taking place, sequence conversion on "panhandle" intermediates can occur.

Enzymes that are predicted to function during sequence conversion on "panhandle" intermediates

The host-encoded proteins would include an endonuclease(s) which can cleave within a bulge and/or the strand opposite the bulge and a ligase. These are assumed to be host encoded since adenovirus is not known to encode either function. The adenovirus polymerase could provide the 3'-5' exonuclease (Field *et al.*, 1985) and all the polymerization functions which include gap filling and DNA synthesis during the removal of blocking structures (Fig. III.11, g-j). For the removal of the blocking structure, all the adenovirus replication proteins would be required. This includes two more adenovirus-encoded proteins, the adenovirus preterminal protein (pTP) and the DNA binding protein. In addition, the three host-encoded proteins, nuclear factors I, II, and III are required.

The limitations of sequence conversion on "panhandle" intermediates

We have shown that there are limitations to strand usage and template length requirements for sequence conversion during the process to restore the identity of the IRs. We have never observed the removal of small bulges (four bp) when positioned in the 5'-IR of the displaced strand (Fig. III.3 and Fig III.4, lane 8, no band **a**). But the evidence indicates that repair attempts are made that lead to the creation of a nonviable genome. It appears as though endonuclease incision does happen in the 5'-IR bulges. When the bulge is close to the "panhandle" loop endswitched double-stranded molecules are created when incision occurs in the bulge (Fig. III.3, lanes 3 and 4; Fig. III.10,

f-k-n). This process creates a molecule which is unable to end-repair, and so will be unable to become a fully replicating molecule. The molecules that do not endswitch appear to be unable to complete sequence conversion because a structure that blocks replication forms; cleavage within the loop generates a 3'- and a 5'-overhang. There is evidence that the 3'-overhang should be removable; an active 3'-exonuclease can work on similar substrates (Wang *et al.*, 1991), and adenovirus polymerase has a 3'-exonuclease associated with it (Field *et al.*, 1984). The 5'-overhang could be the source of the block, either a 5'-exonuclease is unable to remove the overhang or upon removal a non-ligatable structure is created. With regard to the requirements placed on template length, previously it has been shown that polymerization away from the "panhandle" loop requires at least 31 bp of duplex (Wang *et al.*, 1991). In addition, repair of heteroduplexes within the IR required more than six bp, most likely 31 bp also, when polymerization proceeds away from the loop. This was evident by the absence of resistant DNA (band **a**) to *Sph*I cleavage in Fig. III.5, lane 2, band **a**, and by the absence of molecules with two *Sph*I sites Fig. III.3, lane 8, band **c**. On the other hand, polymerization towards the "pan" to retrieve information within six bp of the large loop, on the other hand did take place (Fig. III.5, lanes 7 and 8, the doubly sensitive band **c**; Dig. III.11).

The IR used as template by each mutant is different

A chart is presented in Fig. III.12 to categorize which IR was used as a template and which mutants remained unconverted. Mutant pLIR5'S can use both IRs as template and some appear to remain

unconverted. Adenovirus minichromosome pIR5'S only uses the 3'-IR as template and some appear to remain unconverted. pLIR3'S only uses the 5'-IR as template, while pIR3'S is unable to successfully engage sequence conversion.

PLIR5'S is able to use both IRs as a template, and molecules with one *Sph*I site are detected (Fig. III.6, lane 6). The molecules with only one *Sph*I site did not necessarily arise from escaping sequence conversion. Repair during inter-strand hybridization can lead to the creation of molecules with one *Sph*I site (Chapter IV). Fig. III.13 lists the possible sequence conversion events that can arise during postreplicative overlap recombination. Four loci (heterologies) are created during inter-strand hybridization and are labeled a, b, c, and d. The plus sign (+) represents the *Sph*I restriction endonuclease site, the minus sign (-) indicates that the site has been deleted. During intermolecular pairing between two displaced strands from pLIR5'S, the *Sph*I sites will be positioned at loci b and c. Listed are all the possible outcomes when sequence conversion is initiated at all the possible loci. Incision at locus a, for example, will generate two molecule one with two *Sph*I sites (++) and one with only one *Sph*I site (+-). As can be seen from the possile outcomes of sequence conversion a number of possibilities exists that generate a molecule that appears unconverted. These include incision at a, b, c, d, a + b, and c + d. Thus, it is reasonable to think that all molecules are attacked, and none escape repair as was the case for pLIR3'S (Fig. III.12).

On pIR5'S sequence conversion was successful only when the 3'-IR is used as template (Fig. III.5, lane 8, band **c**). This is a result of an

asymmetrical steric obstruction to polymerization close to the loop. Polymerization away from the loop requires more than six bp, but synthesis toward the loop can proceed as close as six bp. The molecular basis of the steric obstruction remains to be determined. In addition, some pIR5'S mutants escape sequence conversion (Fig. III.5, lane 8, band **b**). At least some of the mutants that escape sequence conversion could arise from the four-bp bulge becoming part of the larger "panhandle" loop only six-bp away. The evidence for this possibility is provided by the endswitched molecules that are created from unstable gapped intermediates (Fig. III.5, lane 4, band **c**). The endswitched molecules indicate that the six-bp duplex might constantly be in flux with the hydrogen bonding constantly breaking and reforming. Above, the contribution of the sequence conversion during *trans* replication was discussed. It was reasoned that the molecules which survive with a single *Sph*I site could arise from sequence conversion attempts versus having escaped repair. Some sequence conversion attempts during inter-strand pairing between displaced pIR5'S strands are predicted to not take place (blocked). The loci that are predicted to not successfully sequence converted require polymerization away from the "panhandle" loop with only six bp at position X. Yet, as can be seen by the chart, a number of possible sequence conversion events could generate a molecule with a single *Sph*I site (a, b, c, d, a + b, a + d). Although it is possible that some of the pIR5'S molecules might escape sequence conversion, the possibility exists that *trans* replication could also contribute.

Transfected pLIR3'S only generates molecules with a second *SphI* site created by using the 5'-IR as template (Fig. III.4, lane 8, band **c**). Sequence conversion attempts which remove the *SphI* site from the 5' IR are not successful because an obstruction to completing the repair is formed. This is likely to be from the creation of a non-removable 5'-overhang, or the generation of a non-ligatable substrate (discussed above). As already mentioned above no molecules with one *SphI* site survive when pLIR3'S is transfected into human cells. Sequence conversion during inter-strand hybridization between the IR predicts that incision at loci a, b, c, d, a + b, and c + d can create a molecule with one *SphI* sites (Fig. III.15). Since none of these molecules are recovered, what could be going on? First, if it is assume that every site is attacked then only sequence conversion simultaneously at a + b and c + d could still contribute. Second, incision at a or d will place a 5'-overhang in the 5'-IR. We do not understand why this would prevent recovery of one of the possible viable minichromosomes, because replication initiating at the origin farthest away from the overhang should be able to remove the 5'-overhang. Since the molecule is not recovered, possibly this indicated that the presence of overhangs simply located in the 5'-IR blocks repair. With respect to the 5' mutants, overhangs are never generated in the 5'-IR, so even if every site is always attacked, there is still the possibility to create minichromosomes with one *SphI* site.

Mutant pIR3'S is unable to successfully participate in any sequence conversion attempts (Fig. III.3, lane 8, only band **b**). Sequence conversion attempts requiring polymerization away from the

loop also are inhibited on this mutant, providing further evidence for a steric obstruction. In addition, repair attempts using the 3'-IR as template are prevented. In this case, a non-removable 5'-overhang, or removal of the overhang creates a non-ligatable substrate (discussed in detail above). Also, this mutant with only six bp separating the bulge and the "panhandle" loop, is likely to escape repair by the bulge becoming part of the loop. Endswitched molecules were detected, suggesting a lack of stability once incision occurs within the site of heterogeneity (Fig. III.3, lane 4, band **c**). During intermolecular hybridization all repair attempts except those that create the molecule with one *Sph*I site are predicted not to function (Fig. III.16). As explained above for pLIR3'S, incision at locus a might also generate a block to successful sequence conversion. In contrast to pLIR3'S, single loci might be a substrate of attack for this mutant. This is possible since part of the time the bulge may become a part of the larger "panhandle" loop.

The consequences of unsuccessful attempts of sequence conversion on "panhandle" intermediates

The description above details sequence conversion attempts which lead to nonviable defective genomes. Isolation of viral particles of lower density than complete virions have been described for several human adenoviruses, including serotypes 2 (Maizel *et al.*, 1968; Rainbow and Mak, 1970; Rosenwirth *et al.*, 1974; Daniell, 1976), 12 (Mak, 1970; Rainbow and Mak, 1970), 3 (Prage *et al.*, 1972; Daniell, 1976), and 16 (Wadell *et al.*, 1973). The creation of incomplete Ad2 particles has been shown not to arise from degradation products (Rosenwirth *et al.*, 1974). Daniell (1976) carried out an analysis of the

defective genomes and found that a significant proportion of the DNA molecules contained extended regions of duplicated inverted sequences. From these data she developed a model for adenovirus replication which could explain the generation of the incomplete virions; this included the suggestion for "panhandle" formation of the displaced single-strands. A displaced single-strand which becomes nicked was proposed to be able to form a "panhandle" by the end base pairing with internal sequences which are complementary and in an inverted orientation. Polymerization out to the end would create an incomplete genome with a long IR. There was clearly a predominance of certain size classes of defective virions. Daniell's explanation for this included a constraint on packaging size, or by preferential sites for nicking and base pairing. While there are packaging constraints for the upper limits for genome length (reviewed by Grodzicker, 1980; Klessig, 1984; Jones and Shenk, 1978; Graham, 1984; Ghosh-Choudhury *et al.*, 1987), preferential sites for nicking are not necessary, though some may exist. Our experiments on "panhandle" formation have revealed a number of facts that can be applied to this question of class sizes. First, we know there is a very active 3'-exonuclease capable of removing over 4,000 bp of unpaired sequences. Second a 31-bp duplex, which can tolerate some heterogeneity, is needed between complementary inverted sequences for end-repair (Wang *et al.*, 1991). Thus, the presence of an imperfect internal inverted repeat (IR), of approximately 31 bp, and a 3'-exonuclease could remove any unpaired 3' sequences on nicked intermediates that form a "panhandle" between the internal IR. Then repair synthesis

extending the 3'-end could generate a size class of defective genome. The randomness of nicking is masked by the 3' exonuclease and a size requirement for "panhandle" formation. Moreover, a computer-directed search has revealed internal sequences capable of forming hairpin structures (Munroe 1983). Daniell argued that nonrandom nicking could generate the size classes, and one source of nonrandom nicking could come from sequence conversion attempts to restore the identity of the ITR. Nicks that occur in the 3' ITR within 31 bp of the panhandle create non-recoverable genomes, whereas nicks within six bp the duplex are unstable. Evidence was presented which showed that sequence conversion attempts within six bp of the "panhandle" loop created endswitched molecules (Fig. III3, lane 4, band **b**; Fig. III5, lane 4, band **c**). The presence of an internal IR, and a 3'-exonuclease could provide the template to generate a size class of defective genomes from the endswitched molecules. A second source could result from the contribution of the small defective genomes that are recovered. Four percent of the Ad2 defective particles isolated contain 5% or less of the total genome length. These small molecules could arise from sequence conversion attempts which produce 5'-overhangs (Fig. III.7). When the 5'-overhangs are in the 3' IR, evidence presented in Chapter V suggests that the 3' end of such molecules are removed by initiation of replication at the origin, which creates a small double stranded molecule. This small double-stranded molecule could be the source of the small defective genomes.

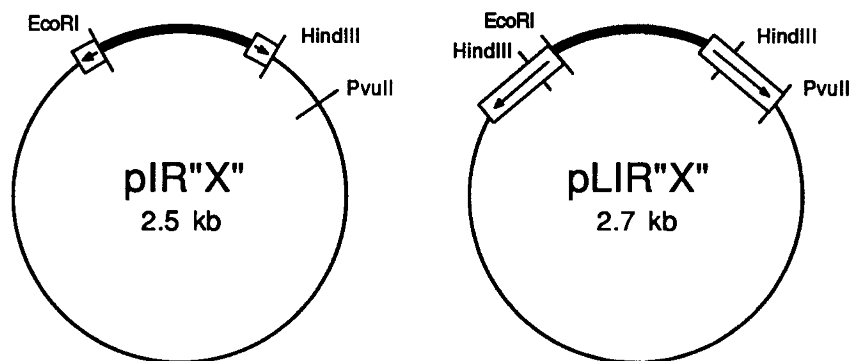


Figure III.1. Structure of the pIR and pLIR series of plasmids. The thick line represents the adenovirus origin sequences. The thin line is pUC19 backbone DNA. The boxes represent the identical sequences with the arrows designating the orientation. pIR"X" has 51 bp of identical sequence. pLIR"X" has the repeated sequence extended to the *Pvu*II site creating a 229 bp IR. Linearization with *Eco*RI exposes the origin sequence.

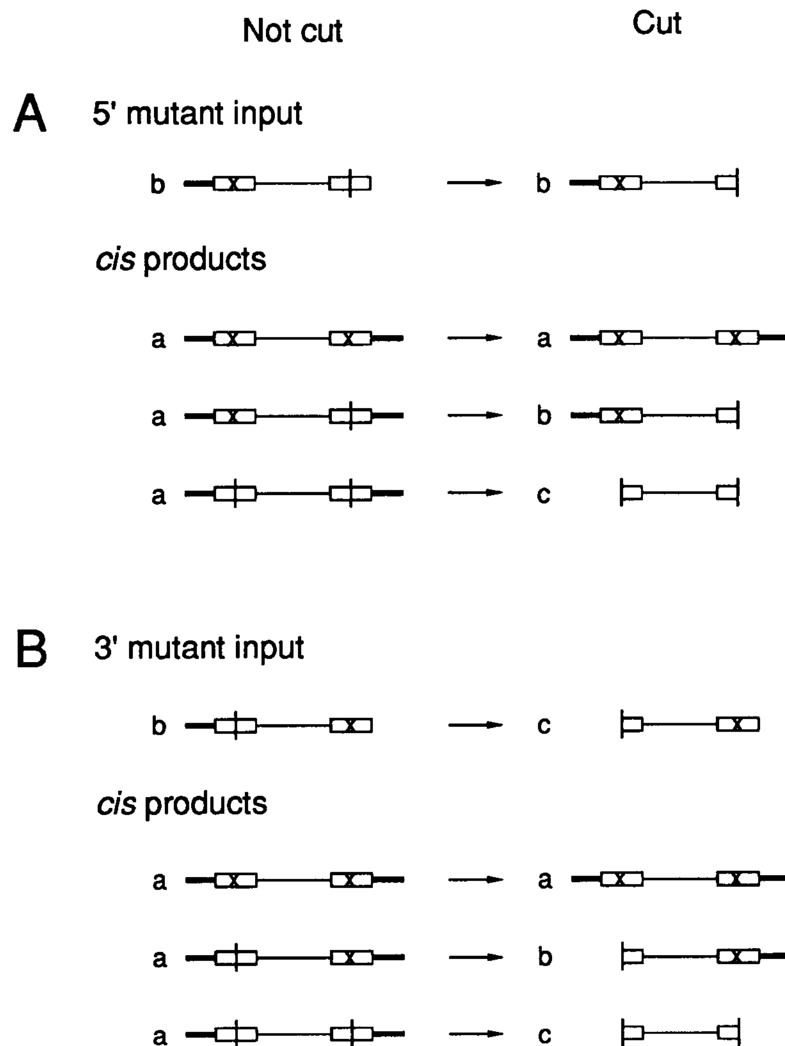


Figure III.2. The location of the restriction sites and the products produced by restriction endonuclease cleavage. Input-sized DNA can participate only in strand displacement replication, while *cis* molecules can engage both phases of adenovirus replication. (A) The uncut and cut replication products for the 5' mutant, with a, b, c labeling the position the band is found in the gel. Band **a** molecules are resistant to *Sph*I cleavage. Band **b** molecules consist of retained input-sized molecules, plus *cis* molecules with only one *Sph*I site. Molecules that will populate position **c** contain two *Sph*I sites. (B) The uncut and cut 3' mutants. Molecules that remain at position **a** will be resistant to *Sph*I cleavage. Band **b** will contain molecules with only one *Sph*I. Band **c** will have two sources; input molecules will migrate to position **c**, and *cis* molecules with two *Sph*I sites will occupy position **c**.

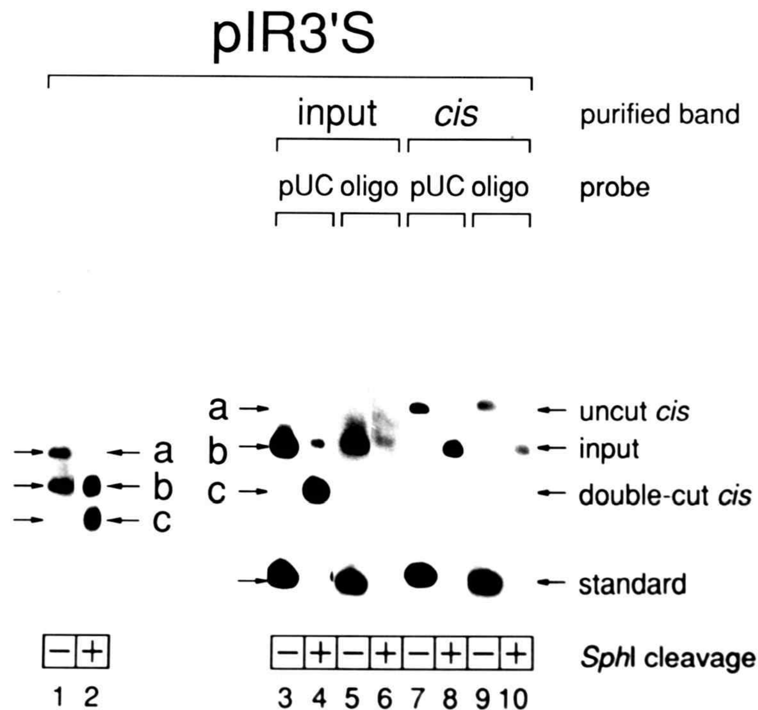


Figure III.3. Endonuclease restriction analysis of sequence conversion in pIR3'S. Sample in even numbered lanes were digested with *Sph*I prior to Southern blot hybridization. Unfractionated total DNA was isolated after transfection and hybridized with a plasmid probe in lanes 1 and 2. Input DNA was purified after transfection and analyzed using a plasmid probe (lanes 3 and 4) or an adenovirus origin-specific oligonucleotide probe (lanes 5 and 6). The *cis* replication product was similarly gel-purified and analyzed using a plasmid probe (lanes 7 and 8) or an adenovirus origin-specific oligonucleotide probe (lanes 9 and 10). a = uncut *cis* molecule, b = input, and c = doubly cut *cis* molecule. Standard = *Sph*I cleavage standard.

pLIR3'S

input cis purified band

pUC oligo pUC oligo probe

a b c

a b c

uncut *cis*

input

double-cut *cis*

standard

− + − + − + − + − + *Sph*I cleavage

1 2 3 4 5 6 7 8 9 10

Figure III.4. Endonuclease restriction analysis of sequence conversion in pLIR3'S. Sample in even numbered lanes were digested with *Sph*I prior to Southern blot hybridization. Unfractionated total DNA was isolated after transfection and hybridized with a plasmid probe in lanes 1 and 2. Input DNA was purified after transfection and analyzed using a plasmid probe (lanes 3 and 4) or an adenovirus origin-specific oligonucleotide probe (lanes 5 and 6). The *cis* replication product was similarly gel-purified and analyzed using a plasmid probe (lanes 7 and 8) or an adenovirus origin-specific oligonucleotide probe (lanes 9 and 10). a = uncut *cis* molecule, b = input, and c = doubly cut *cis* molecule. Standard = *Sph*I cleavage standard.

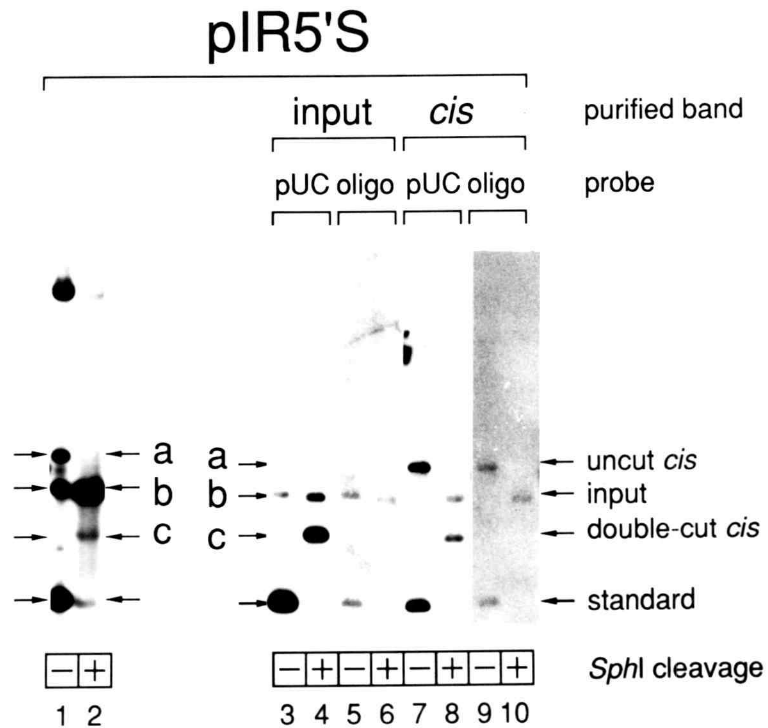


Figure III.5. Endonuclease restriction analysis of sequence conversion in pIR5'S. Sample in even numbered lanes were digested with *Sph*I prior to Southern blot hybridization. Unfractionated total DNA was isolated after transfection and hybridized with a plasmid probe in lanes 1 and 2. Input DNA was purified after transfection and analyzed using a plasmid probe (lanes 3 and 4) or an adenovirus origin-specific oligonucleotide probe (lanes 5 and 6). The *cis* replication product was similarly gel-purified and analyzed using a plasmid probe (lanes 7 and 8) or an adenovirus origin-specific oligonucleotide probe (lanes 9 and 10). a = uncut *cis* molecule, b = input, and c = doubly cut *cis* molecule. Standard = *Sph*I cleavage standard.

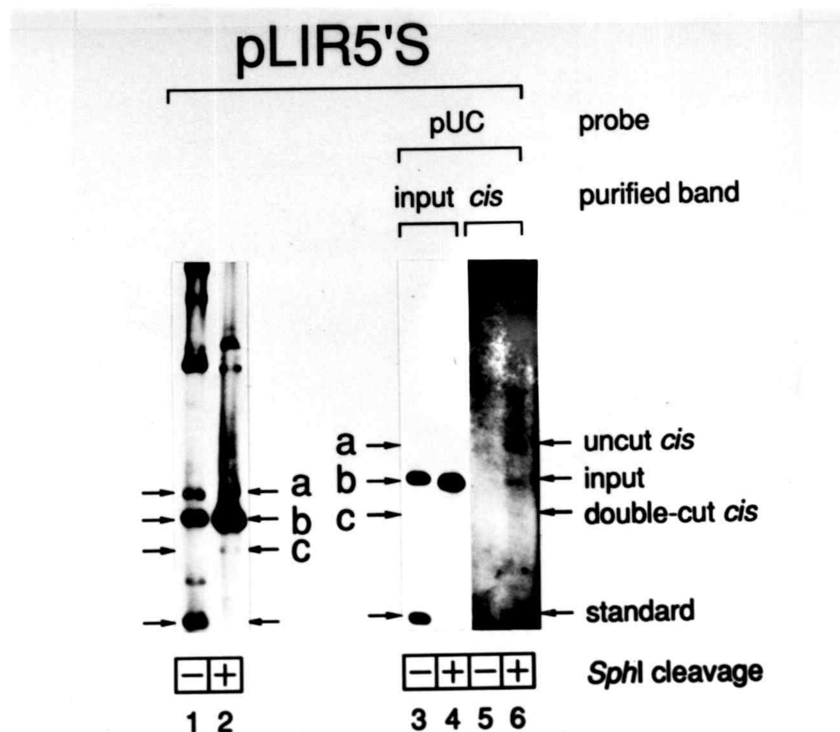


Figure III.6. Endonuclease restriction analysis of sequence conversion in pLIR5'S. Sample in even numbered lanes were digested with *Sph*I prior to Southern blot hybridization. Unfractionated total DNA was isolated after transfection and hybridized with a plasmid probe in lanes 1 and 2. Input DNA was purified after transfection and analyzed using a plasmid probe (lanes 3 and 4). The *cis* replication product was similarly gel-purified and analyzed using a plasmid probe (lanes 5 and 6). a = uncut *cis* molecule, b = input, and c = doubly cut *cis* molecule. Standard = *Sph*I cleavage standard.

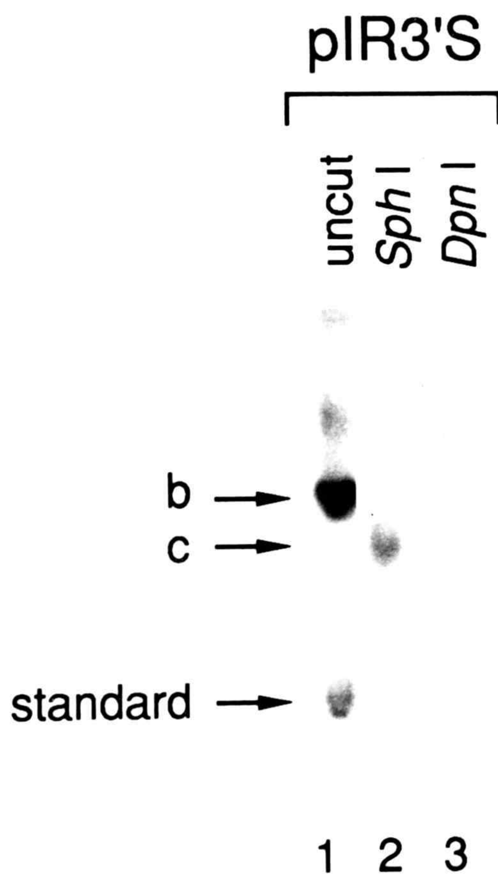


Figure III.7. Endonuclease restriction analysis of transfected pIR3'S without helper DNA. Sample in lanes 2 and 3 were digested with *Sph*I and *Dpn*I prior to Southern blot hybridization, respectively. Unfractionated total DNA was isolated after transfection and hybridized with a plasmid probe in lanes 1-3. b = input, c = singly cut input molecule. Standard = *Sph*I cleavage standard.

Figure III.8. Endonuclease restriction analysis of gel-purified pIR3'S dimer DNA. Sample in lane 2 was digested with *Sph*I prior to Southern blot hybridization. Gel-purified dimer DNA was hybridized with a plasmid probe in lanes 1 and 2. Standard = *Sph*I cleavage standard.

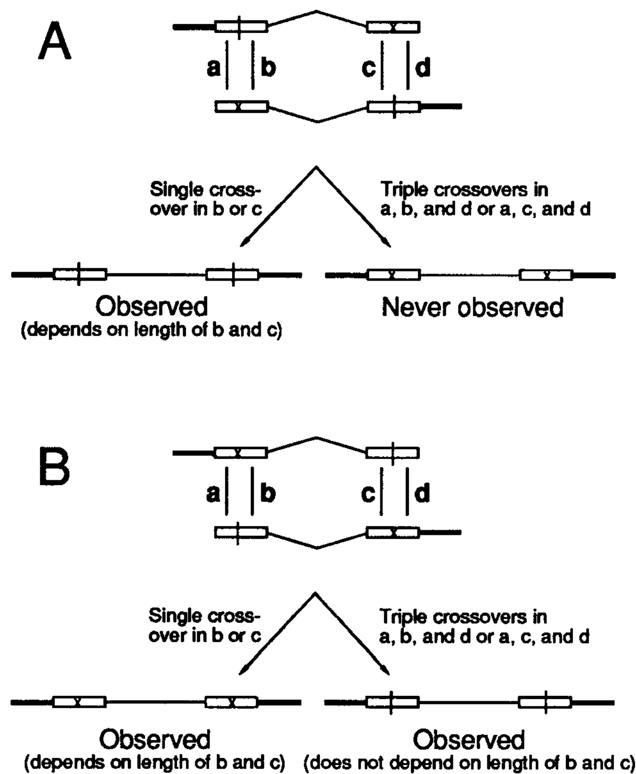


Figure III.9. Diagram of adenovirus minichromosome sequence conversion mediated by homologous recombination. The double-stranded molecules are indicated by single lines. The thick line is the adenovirus origin. The open boxes are inverted repeats and the small vertical lines indicate the presence of the *Sph*I restriction endonuclease cleavage site and an X denotes a mutated site. The thin line is plasmid DNA. Recombination can only occur through the inverted repeats. The crossover regions are identified as a, b, c, and d. (A) pIR3'S or pLIR3'S. (B) pIR5'S or pLIR5'S.

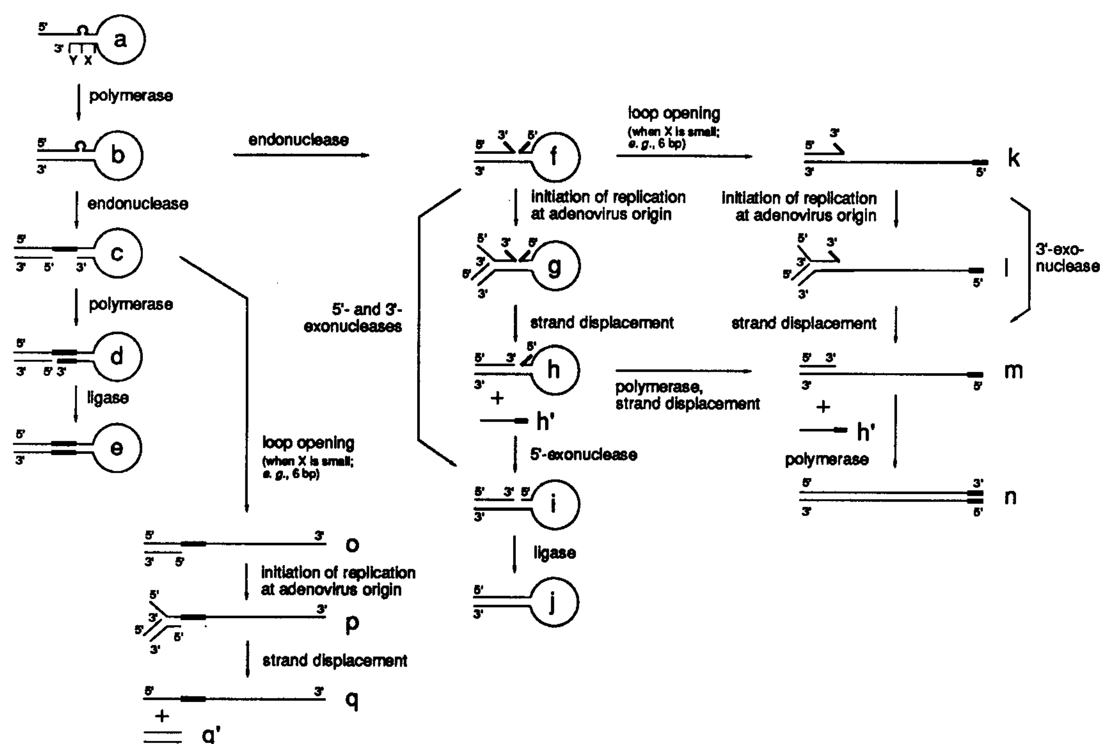


Figure III.10. Mismatch repair (sequence conversion) of 3' mutants. The single-stranded molecules are represented by single lines and the double-stranded molecules by two lines. The bulge or thick patch is the four central bp of the restriction endonuclease site. Y indicates the distance from the 3'-end of the displaced strand to the bulge. Y = 41 for both plasmids. X gives the distance from the bulge to the loop of the "panhandle". X = six bp for pIR3'S and X = 175 bp for pLIR3'S. (a) "Panhandle" of the displaced single-stranded molecule from a plasmid with a single viral origin. (b) Extension of a "panhandle" molecule using adenovirus sequences as a template. (c) Endonuclease action opposite the bulge creates a gap. (d) Polymerase gap fills. (e) Ligation generates a continuous single-stranded molecule. (f) Endonuclease action within the bugle creates a 5' and a 3' overhang. (g) Initiation of replication at the adenovirus origin. (h) "Panhandle" structure with the 3' overhang removed by displacement of single-strand h'. (i) 5'-exonuclease to remove 5'-overhang. (j) Ligation creates a continuous single-stranded molecule. (k) After endonuclease action pIR3'S opens up into a partial double-stranded molecule with a 3' overhang. (l) Initiation of replication at the origin. (m) Strand-displacement replication creates a partially double-stranded molecule without a 3' overhang. (n) Polymerization completes the creation of an endswitched double-stranded molecule. (o) Molecule c opens up after endonuclease action. (p) Initiation of replication at the adenovirus origin. (q) The displaced single-stranded molecule. (q') The small double stranded replication product.

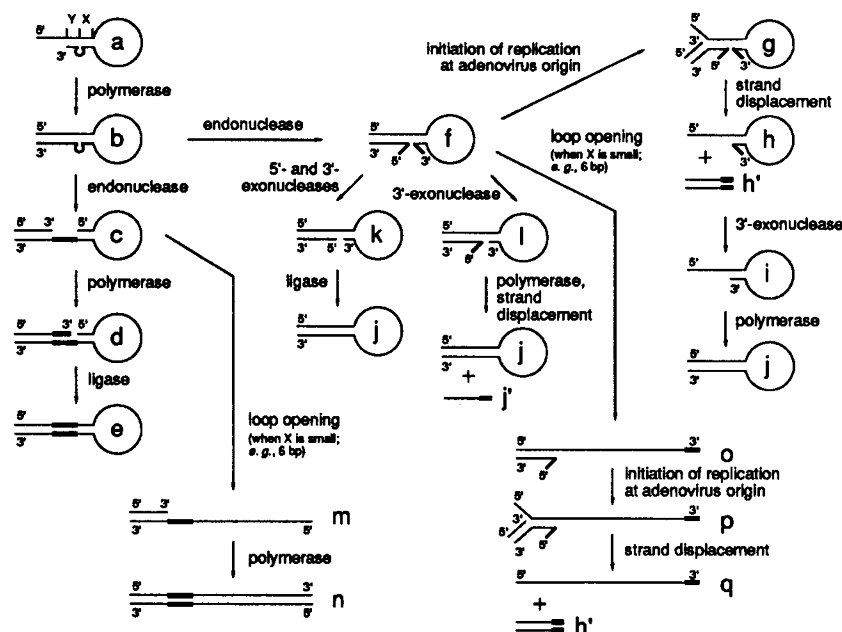
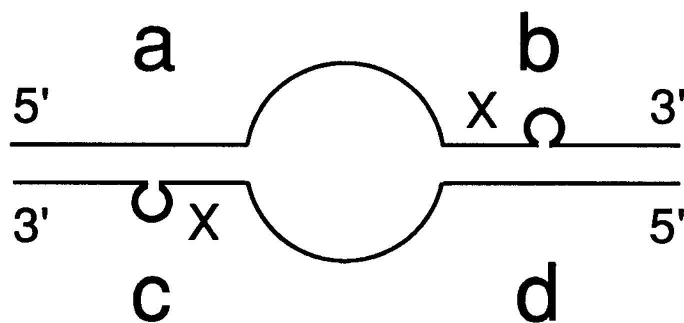


Figure III.11. Mismatch repair (sequence conversion) of 5' mutants. The single-stranded molecules are represented by single lines and the double-stranded molecules by two lines. The bulge or thick patch is the four central bp of the restriction endonuclease site. Y indicated the distance from the 3'-end of the displaced strand to the bulge. Y = 41 for both plasmids. X gives the distance from the bulge to the loop of the "panhandle". X = six bp for pIR3'S and X = 175 bp for pLIR3'S. (a) "Panhandle" of the displaced single-stranded molecule from a plasmid with a single viral origin. (b) Extension of a "panhandle" molecule using adenovirus sequences as a template. (c) Endonuclease action opposite the bulge creates a gap. (d) Polymerase gap fills. (e) Ligation generates a continuous single-stranded molecule. (f) Endonuclease action within the bulge creates a 5' and a 3' overhang. (g) Initiation of replication at the adenovirus origin. (h) "Panhandle" structure with the 5' overhand removed by displacement of double-strand molecule h'. (i) 3'-exonuclease to remove the 3'-overhang. (j) Ligation creates a continuous single-stranded molecule. (k) A 5'- and 3'-exonuclease creates a "panhandle" intermediate with a nick. (l) A 3'-exonuclease removes the 3'-overhang generating a "panhandle" intermediate that can be acted on by polymerase to produce a continuous single-stranded molecule j and displace single-strand j'. After endonuclease action pIR5'S opens up into a partial double-stranded molecule. (n) Polymerization creates a double-stranded endswitched molecule. (o) After endonuclease action pIR5'S opens up into a partial double-stranded molecule with a 5' overhang. (p) Initiation of replication at the adenovirus origin. (q) The displaced single-stranded molecule. (h') The small double stranded replication product.

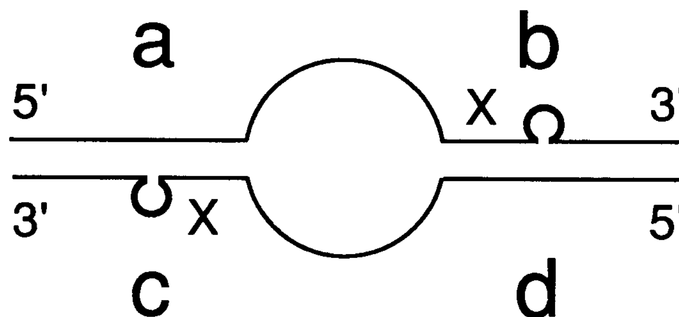
Plasmid	3' IR as template	5' IR as template	Unconverted *cis*
pLIR5'S	Yes	Yes	Yes
pIR5'S	Yes	No	Yes
pLIR3'S	No	Yes	No
pIR3'S	No	No	Yes

Figure III.12. Summary of the IR used by each mutant. Transfected pLIR5'S generates molecules which can use both IRs as template, plus molecules that appear to be unconverted. Transfected pIR5'S generates molecules that appear to be unconverted and molecules that use the 3'IR as template. Transfected pLIR3'S only produces molecules that have used the 5'IR as template. Transfected pIR3'S appears to be unable to engage sequence conversion.



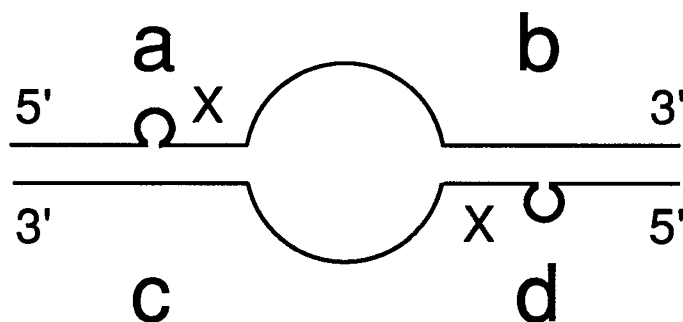
	Loci			
	a	b	c	d
pLIR5'S	-	+	+	-
<u>converted at:</u>				
a	+	+	+	-
b	-	-	+	-
c	-	+	-	-
d	-	+	+	+
a + b	+	-	+	-
a + d	+	+	+	+
b + c	-	+	-	+
c + d	-	+	-	+

Figure III.13. Analysis of the possible outcomes of sequence conversion during *trans* replication on pLIR5'S. Loci a, b, c, and d identifies the locations of the heterology. The plus sign (+) marks the presence of an *Sph*I site and the minus sign (-) the location of the deleted *Sph*I site. X = distance separating the bulge from the "panhandle" loop. The outcome of sequence conversion at each possible locus is listed.



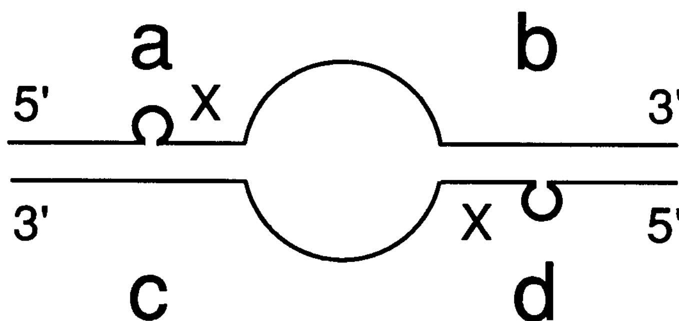
	Loci			
	a	b	c	d
pIR5'S	-	+	+	-
converted at:				
a	+	+	+	-
b	blocked		+	-
c	-	+	blocked	
d	-	+	+	+
a + b	+	-	+	-
a + d	+	+	+	+
b + c	blocked		blocked	
c + d	-	+	blocked	

Figure III.14. Analysis of the possible outcomes of sequence conversion during *trans* replication on pIR5'S. Loci a, b, c, and d identifies the locations of the heterology. The plus sign (+) marks the presence of an *SphI* site and the minus sign (-) the location of the deleted *Sph*I site. X = distance separating the bulge from the "panhandle" loop. The outcome of sequence conversion at each possible locus is listed.



	Loci			
	a	b	c	d
pLIR3'S	+	-	-	+
converted at:				
a	blocked		-	+
b	+	+	-	+
c	+	-	+	+
d	+	-	blocked	
a + b	blocked		-	+
a + d	blocked		blocked	
b + c	+	+	+	+
c + d	+	-	blocked	

Figure III.15. Analysis of the possible outcomes of sequence conversion during *trans* replication on pLIR3'S. Loci a, b, c, and d identifies the locations of the heterology. The plus sign (+) marks the presence of an *Sph*I site and the minus sign (-) the location of the deleted *Sph*I site. X = distance separating the bulge from the "panhandle" loop. The outcome of sequence conversion at each possible locus is listed.



	Loci			
	a	b	c	d
pIR3'S	+	-	-	+
<u>converted at:</u>				
a	blocked		-	+
b	blocked		-	+
c	+	-	blocked	
d	+	-	blocked	
a + b	blocked		-	+
a + d	blocked		blocked	
b + c	blocked		blocked	
c + d	blocked		blocked	

Figure III.16. Analysis of the possible outcomes of sequence conversion during *trans* replication on pIR3'S. Loci a, b, c, and d identifies the locations of the heterology. The plus sign (+) marks the presence of an *Sph*I site and the minus sign (-) the location of the deleted *Sph*I site. X = distance separating the bulge from the "panhandle" loop. The outcome of sequence conversion at each possible locus is listed.

CHAPTER IV: SEQUENCE CONVERSION OCCURS DURING ADENOVIRUS POSTREPLICATIVE OVERLAP RECOMBINATION

INTRODUCTION

In this Chapter, sequence conversion is shown to take place during adenovirus postreplicative overlap recombination. Overlap recombination is a processes by which overlapping terminal fragments recombine to generate a complete adenovirus genome (Chinnadurai *et al.*, 1979; Berkner and Sharp, 1983; Volkert and Young, 1983; McGrory *et al.*, 1988). Previously it had been shown to occur through homologous recombination as a prereplicative process (Volkert and Young, 1983). More recently it has been documented that overlap recombination can result from postreplication events, which do not involve homologous recombination (Ahern *et al.*, 1991). Two kinetic studies of replicating adenovirus genomes indicated that postreplicative overlap recombination occurs late during the infectious cycle (Bodnar and Pearson, 1980; D'Halluin and Milleville, 1984). Density shift experiments which labeled adenovirus DNA HH (heavy-heavy) with 5-bromodeoxyuridine (BRdU) and then shifted to medium lacking BRdU (light medium) showed that early (14 hr) during the infection that HH and HL genomes disappeared rapidly, then disappeared more slowly at 18 hr, and disappeared at the slowest rate at 22 hr in the replication cycle. A second observation was that at about 18 hr after infection accumulation of viral DNA began to deviate from logarithmic growth. The interpretation of the above results at the time was that the rate of initiation was changing during the course of the infection. However, the experiments are completely consistent with the process of

postreplicative overlap recombination, which involves hybridization of displaced single-stranded genomes. Early during the infection, when the concentration of displaced single-strands is low, intramolecular hybridization between the ITR ("panhandle" formation) would likely dominate in providing a duplex for initiation of complementary strand synthesis. Whereas, late during infection, as the concentration of single-stranded genomes increases, intermolecular hybridization (postreplicative overlap replication) between the genomes would be expected to increase. This process would have the effect of slowing the loss of HH and HL molecules, as well as cutting the production of daughter molecules, which would short-circuit logarithmic growth. This chapter demonstrates that the intermolecular sequence conversion detected is not mediated by homologous recombination, and provides further support for hybridization between displaced adenovirus strands.

RESULTS

It has been shown previously that adenovirus engages in postreplicative overlap recombination (Ahern et al. 1991) by using an adenovirus minichromosome system (Hay *et al.*, 1984; Wang and Pearson, 1985). We currently have used the same system to explore intermolecular sequence conversion of small (four bp) heteroduplexes between the IRs of adenovirus minichromosomes. pLIRWT and pLIRSD each has a single left adenovirus origin flanked by 229 bp of identical sequence which includes a polylinker (Fig. III.1). The only difference between the two plasmids is that pLIRSD has had the central four bp of the *Sph*I site delelted (SD) from both polylinkers.

When both plasmids are linearized with *Eco*RI the adenovirus origin is exposed. The duplicated flanking sequences are arranged as an ITR, but the replication origin sequences are missing from one end. Both linearized plasmids have been singly transfected with helper, total DNA purified, and hybridized with a plasmid probe during Southern blot analysis (Fig. VI.1, lanes 1-4). Prior to Southern blot analysis, a sample of the transfected products was digested with *Sph*I (lanes 2 and 4). Band **b** is the size of the input DNA, which can participate only in strand displacement replication. Band **a** is the cis replication product which has end-repaired, and can now participate in both phases of adenovirus replication. Band **c** is the product of cleavage when both *Sph*I sites are present. A cleavage standard was introduced before digestion with *Sph*I to monitor whether cutting proceeded to completion. As can be seen from the gel, when pLIRWT is cleaved with *Sph*I all bands migrate to position **c**, whereas when pLIRSD was treated with *Sph*I all replication products were resistant to cleavage. When the two plasmids are transfected together and digested with *Sph*I all three bands appear as expected (lanes 5 and 6).

Intermolecular sequence conversion

To determine whether sequence conversion took place between the two plasmids, bands **a** and **b** were gel-purified and separately cleaved with the endonuclease restriction enzyme *Sph*I (Fig. IV.1, lanes 7-10). Band **b** contains a mixture of molecules with *Sph*I sites (band **b**) and without (band **c**). The cis replication product (band **a**) contains molecules like pLIRSD, which are resistant to *Sph*I cleavage and remain at position **a**, as well as replication products like pLIRWT,

which are doubly sensitive to *Sph*I cleavage (band **c**), and finally, molecules which now contain one *Sph*I site (band **b**). This result shows that the two plasmids exchanged sequence information.

Homologous recombination does not take place between the minichromosomes

Two possible mechanisms could mediate the above result; homologous recombination, or adenovirus postreplicative overlap recombination. To distinguish the two possibilities, pLIRSD was cut with *Sst*I prior to transfection to remove the adenovirus origin of replication and then cotransfected with linearized pLIRWT (Fig. IV.2, lane 1). Minichromosome pLIRSD, under these conditions, was unable to participate in strand displacement replication and generate single strands, which is obligatory for postreplicative overlap recombination. Homologous recombination should be able to proceed with equal efficiency with the removal of the origin of replication from one of the plasmids. There is still an extensive region of homology between the two plasmids, and it has been shown in mammalian cells that homologous recombination between extrachromosomal molecules does not depend upon replication (Subramani and Berg, 1983; Ayares *et al.*, 1985). Fig. IV.2 displays the results of this experiment and demonstrates that homologous recombination does not take place between the two plasmids. Before treatment with *Sph*I three bands are seen; bands **a** and **b** are replication products from pLIRWT while band **c** is transfected pLIRSD (cut *Sst*I; lane 1). Lane 2 shows that all the bands were sensitive to *Sph*I, illustrating that homologous recombination did not take place when one plasmid was unable to undergo replication. pLIRWT still has two *Sph*I sites as when it was

transfected without pLIRSD (cut with *Sst*I; Fig IV.1, lane 1 and 2). In addition, pLIRSD did not end-repair to create a doubly resistant minichromosome. Lane 3, is a control lane, showing that no replication products are obtained when each plasmid is cut with *Sst*I and then transfected together with helper. The converse experiment was done with pLIRWT cut *Sst*I and pLIRSD cut *Eco*RI and the above result was verified (data not shown).

Limitations to intermolecular sequence conversion between minichromosomes

The next set of experiments was designed to determine whether adenovirus minichromosomes with a smaller IR, 51 bp rather than of 229 bp, could still engage in intermolecular sequence conversion. pIRWT and pIRSD each contains a single adenovirus origin which is flanked by 51 bp of identical sequence and contains the polylinker (Fig. III.1). When each is linearized, transfected, total DNA purified, and finally treated with *Sph*I, all replication products of minichromosome pIRWT migrate to position **c** while pIRSD molecules remain resistant (lanes 1-4). Gel-purification of bands **a** and **b**, and then digestion with *Sph*I shows that intermolecular sequence conversion did not take place between these two adenovirus minichromosomes. The purified input product (band **b**) when digested contains a mixture of band **b** and **c** as expected. The purified cis product (band **a**) when cut, only contains band **a** and band **c** but clearly no band **b**. So, the plasmid without the *Sph*I sites (pIRSD) and the plasmid with two *Sph*I sites (pIRWT) were unable to interact to create the hybrid molecule with only one *Sph*I site (band **b**). This result further confirms that homologous recombination is not the mechanism for intermolecular sequence conversion. There

is extensive homology through which homologous recombination can take place between the two plasmids. The only difference between the two sets of plasmids (pLIR series versus pIR series) is the missing sequences between the *Hin*dIII site and the *Pvu*II site to the left of the adenovirus origin sequences (Fig. III.1). The length of the IR has changed, while the major region of overlap has remained the same. This result indicates that either the length of the IR, or the location of the heterology relative to the end of the IR, or both, influence the potential for intermolecular sequence conversion during postreplicative overlap recombination.

A model for sequence conversion during postreplicative adenovirus overlap recombination

We have begun to delineate the steps in the mechanism involved in both intermolecular and intramolecular sequence conversion during strand hybridization of the IR through the course of adenovirus replication. Shown in chapter III is that there are definable limitations to strand usage, and template length requirements for sequence conversion, on substrates similar to the ones used in the above experiments. We have never observed the use of the 3'-IR as template when the extra sequences (four bp) are in the 5'-IR. We believe this to be a result of the formation of a structure that blocks replication; cleavage within the loop generates a 3'- and a 5'-overhang. Since there is evidence of an active 3'-exonuclease that can work on these substrates (Wang et al., 1991, Ahern *et al.*, 1991), and adenovirus polymerase has an active 3'-exonuclease associated with it (Field *et al.*, 1984), it is thought that the block results from the 5'-overhang, or removal of the 5'-overhang generates a non-ligatable structure (see

chapter III). With regard to the requirements placed on template length, it was shown that during "panhandle" formation end-repair requires at least 31 bp of duplex (Wang *et al.*, 1991). Also repair of heteroduplexes within the IR required more than six bp to polymerize away from the "panhandle" loop, but polymerization towards the loop to retrieve information within six bp of the large loop did occur (Chapter III). These observations can be extrapolated for building a model for intermolecular sequence conversion during adenovirus postreplicative overlap recombination.

When the two displaced parental strands of the adenovirus minichromosomes hybridize between their IRs, the rest of the molecules are unable to base pair, thus creating a large single-stranded looped region not unlike the "panhandle". Based on all the available data from past and present experiments, a model to explain these observations is put forth. During adenovirus replication of the cotransfected minichromosomes, two displaced strands can pair through their IR generating small regions of heteroduplex (Fig. IV.4). Loci a, b, c, and d are the sites of heterology between the two plasmids. Plasmid pLIRWT has two *Sph*I sites and is identified by two plus signs (+ +). Minichromosome pLIRSD has had both *Sph*I sites deleted and is marked with two minus signs (- -). All outcomes are listed for conversion at all possible loci. Incision at bulge a produces a 5'-overhang in a 5'-IR, and as described above, this creates a block to the completion of repair. In this case incision at bulge a should only prevent recovery of one of the molecules (pLIRWT). Whereas, pLIRSD could still be recovered by initiating replication at its origin, this

displaces the 5'-overhang. Data presented in Chapter III showed that pLIR3'S did not successfully convert any molecule with incision in the 5'-IR bulge. The meaning of this result is unclear, but it is assumed that if it inhibits successful sequence conversion on pLIR3'S it also must inhibit repair at that locus between pLIRWT and pLIRSD. Also, based on the result from pLIR3'S we assume that all heterologies within a stable duplex are targets for sequence conversion. In other words, both sites are always hit, not just one. With all these constraints imposed this leaves incision at loci b + c as the probable pathway for producing the minichromosome with one *Sph*I site (Fig. IV.1, lane 10, band **b**).

The *trans* replication product was not detected when cotransfection between two minichromosomes with only a 51 bp IR was carried out (Fig. IV.3, lane 8, no band **b**). The outcomes for sequence conversion at all possible loci are listed in Fig. IV.5. One additional constraint is placed on these mutants, that being sequence conversion that involves polymerization away from the "panhandle" loop does not take place. Because the heterology is only six bp away from the "panhandle" loop it is likely that part of the time the bulge is part of the larger loop. In this way these mutants might be able to escape repair. Locus d is the only potential site left that will generate a hybrid molecule with one *Sph*I site. Based on this model there is some possibility to recover minichromosomes with one *Sph*I site when the IR is only 51 bp, but the molecule was not recovered. Notice that sequence conversion at the one possible combination of loci (b + c) for the pLIR molecules will create 2 molecules with one *Sph*I site, while the one

possible site (d) for sequence conversion between the pIR molecules will only generate one molecule with one *Sph*I site. It is predicted that optimum sequence conversion between the pIR series should only happen 50% as frequently as it does between pLIR minichromosomes. Moreover, since the six-bp duplex is not stable one could predict that sequence conversion will be even less frequent. However, successful conversion is seen at a site like b' when "panhandle" formation can mediate the process. An alternative explanation for not recovering the hybrid molecule during transfection with the pIR plasmids, is that postreplicative overlap recombination might require an IR of longer then 51 bp.

DISCUSSION

This study confirms that there is adenovirus replication driven displaced-strand hybridization which leads to overlap recombination (Ahern *et al.*, 1991). Further, it was shown that small (four bp) heteroduplex regions between the two hybridized strands can be rectified, while previously it was shown that large heteroduplexes were not removed (Ahren *et al.*, 1991). This is in contrast to what Ayares *et al.* (1987) found. When they introduced extrachromosomal substrates into mammalian cells they found that both small and large heteroduplex loops were removed, with the small loops (8-10 nt) used a significant amount of the time as a template and the large loops (248-283 nt) as the acceptor site. It is proposed from the results of the study, that large loops are not attacked by an endonuclease because they are coated with adenovirus DNA-binding protein (DBP), whereas a small tightly constrained loop might be unable to be bound by DBP. Shown

also is that small heteroduplexes which are located close to the internal junction of the IR are not substrates for intermolecular sequence conversion.

Summarized below is the evidence that homologous recombination is not taking place at a detectable level between adenovirus minichromosomes in our *in vivo* replication system. Possibly as a result of having a system which is void of homologous recombination we have been able to detected a new process, sequence conversion during adenovirus postreplicative overlap recombination, which will be described in detail below.

Homologous recombination is not detected between adenovirus minichromosomes in mammalian cells

Two key experiments were preformed which unequivocally rule out homologous recombination as the mechanism to generate a hybrid molecule from two highly homologous adenovirus minichromosomes. One of the experiments involved using one of the two minichromosomes with the origin of replication removed; this prevents that minichromosome from participating in postreplicative overlap recombination. There was still extensive homology between the two molecules, and extrachromosomal homologous recombination in mammalian cells does not require replication (Subramani and Berg, 1983; Ayares *et al.*, 1985). Transfection of the two molecules, purification, and analysis revealed that under these conditions the hybrid molecule was not formed (Fig. IV.2, lane 2). The second experiment used two plasmids with a smaller IR which did not generate a recombinant, though the backbones of the plasmids were identical to those in the experiment where the hybrid molecule was

obtained (Fig. III.1, Fig. IV.1, lane 10 , compared to Fig. VI.3, lane 8, band **b**). The pLIR plasmids had 2,657 bp of homology, while the pIR plasmids had 2,486 identical bp, yet no recombinant was generated. This result is not consistent with homologous recombination as the mechanism for the intermolecular events that were recorded.

Sequence conversion during postreplication overlap recombination.

A model for sequence conversion during adenovirus replication driven overlap recombination does provide an explanation for the results obtained. Hybridization of displaced strands provides the opportunity for information exchange between two parental strands. Cotransfection of two molecules which differed by only four bp in each IR demonstrated that sequence conversion can be mediated by postreplicative recombination. Sequence conversion during "panhandle" formation could not contribute, since the IR of each molecule was identical. Two different sets of molecules were tested; one set exchanged sequence information (Fig. IV.1, lane 10, band **b**), the second set did not (Fig. IV.3, lane 8, no band **b**). The location of the heterology might be the crucial difference, or the length of the IR. Sequence conversion occurred when the IR was 229 bp with the heterology 184 bp away from the internal junction of the IR. Information was not exchanged when the IR was 51 bp and the heterology was only six bp from the internal junction. Previously, have limitations on sequence conversion attempts were seen (Chapter III). Cleavage within all possible substrates, bulges and strands opposite the bulges, do not always result in a viable replicating minichromosome. There are at least two possible explanations for why sequence

conversion did not occur between the minichromosomes with an IR of 51 bp. One explanation is that sequence conversion at only one of the possible sites of incision will be successful (Fig. IV.5, d). The contribution of this pathway may be below our level of detection, though detection did occur when intramolecular hybridization ("panhandle" formation) was a contributor to the outcome (Chapter III). The second possibility is that adenovirus chromosomes requiresan IR longer than 51 bp to participate in strand hybridization through the ITR. Many *in vivo* and *in vitro* experiments have defined the adenovirus origin of replication to consist of approximately the first 50 bp of the ITR with the first 18 bp containing the minimal origin (for reviews see Challberg and Kelly, 1989; Stillman, 1989). In addition, *in vivo* work has shown that the ITR provides a structural role for second strand synthesis, beyond providing the sequences for replication protein recognition; the length requirement for this structural role was shown to be 31 bp (Wang *et al.*, 1991). The results with the pIR minichromosomes, with a 51-bp IR, potentially opens a window to the question of the biological significance for evolutionarily conserved longer ITRs which range from 63 to 166 bp depending upon the serotype (Steebergh *et al.*, 1977; Shinagawa and Padmanabhan, 1980; Stillman *et al.*, 1982a; Shinagawa *et al.*, 1983; Sussenbach, 1984; for review see van Ormondt and Galibert, 1984). Discriminate is possible between the two reasons why sequence conversion did not take place between molecules with only a 51 bp IR. A similar set of minichromosome with a 51 bp IR, which have had the *KPN*I site deleted, has been constructed. These plasmids now have 39 bp between the site of heteroduplex and the single stranded

sequences which divide the repeat. Incision at c, or d, could produce the hybrid molecule (Chapter VI). Detection should not be a problem with a second site that can contribute to sequence conversion. If the recombinant is still not detected we could rule out the possibility that the location of the heteroduplex is the limiting factor. This would then lend support to the idea that a longer ITR is required for intermolecular hybridization between the two ITR.

Mautner and Mackay (1984) have previously looked for evidence of gene conversion between overlapping end fragments that contained considerable heterogeneity. They used the right-hand fragment from Ad5*ts*2, which has a lesion within the hexon structural gene, and the left-hand fragment from Ad2WT. Neither fragment contains all the information necessary for expression of the viral DNA replication proteins; thus, postreplicative overlap recombination is unable to contribute in the production of viable virus. The homology between Ad5 and Ad2 within the hexon gene is as low as 47% in some regions (Kinloch *et al.*, 1984). The cross gave rise to a recombinant with a crossover located within a 21-nucleotide tract of homology, flanked by single base heterogeneity. Branch migration did not extend at all into regions of heteroduplex. Shown was that some attempts at sequence conversion between the IR of adenovirus minichromosome replication intermediates result in the creation of non-viable virus. It is conceivable that during homologous recombination, that there are mechanisms to prevent branch migration into heterogeneous regions, and certainly for adenovirus when the outcome of repair can be deadly the constraints could be tight. In fact, in *E. coli* it has been

demonstrated that while unfacilitated branch migration is a rapid spontaneous process that the RecA protein is required to proceed efficiently past short DNA mismatches, and pyrimidine dimers, and is needed to drive through deletions or insertion (Das Gupta and Radding, 1982; Livneh and Lehman, 1982; Bianchi and Radding, 1983).

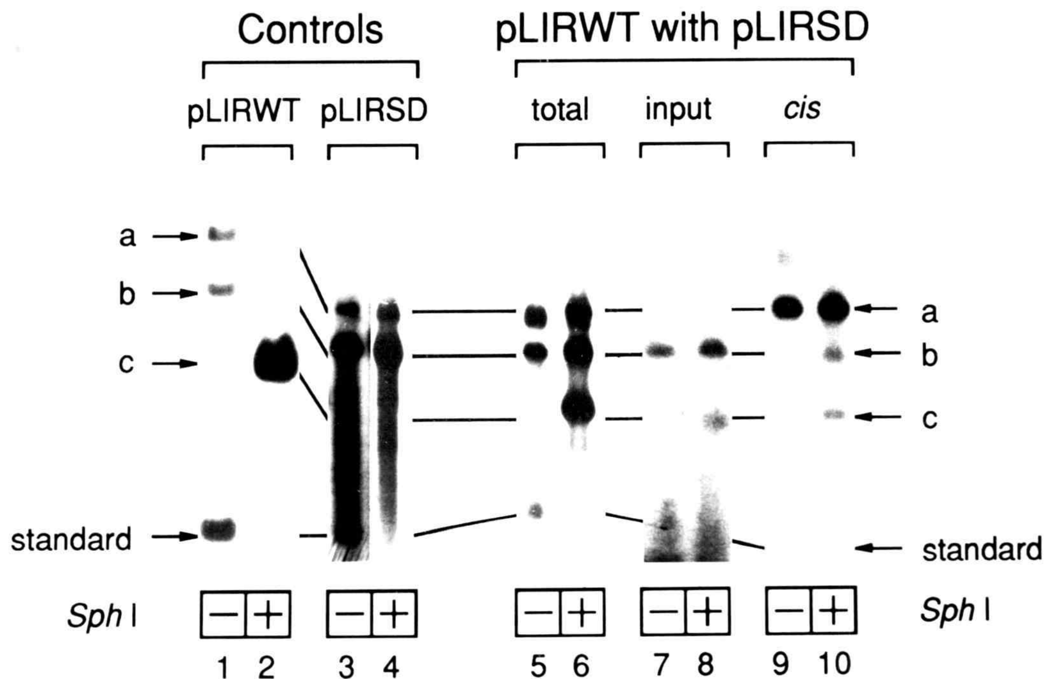


Figure IV.1. Analysis of sequence conversion during postreplicative overlap recombination between two minichromosomes with an IR of 231 bp. pLIRWT and pLIRSD were transfected alone (lanes 1-4), and together (lanes 5-10). Samples in even numbered lanes were digested with *Sph*I prior to Southern blot hybridization. Unfractionated total DNA was isolated after transfection and hybridized with a plasmid probe in lanes 1-6. Input DNA was purified after transfection and analyzed using a plasmid probe (lanes 7 and 8). The cis replication product was similarly gel-purified and analyzed using a plasmid probe (lanes 9 and 10). **a** = uncut cis molecule, **b** = input, and **c** = doubly cut cis molecule. Standard = *Sph*I sensitive cleavage standard.

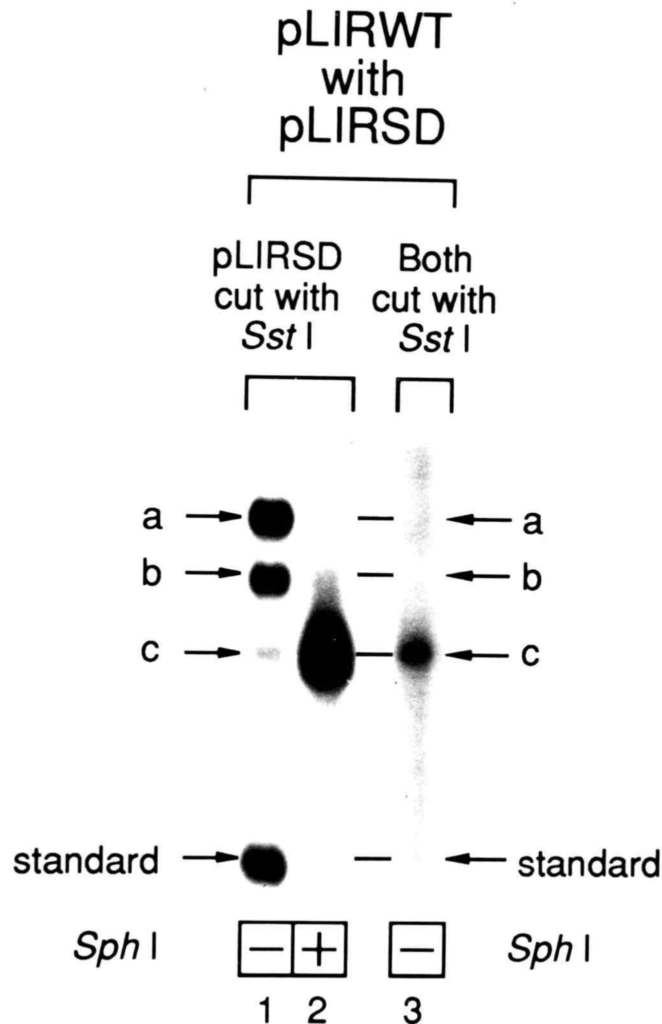


Figure IV.2. Control for homologous recombination. pLIRWT and pLIRSD were cotransfected. pLIRSD was cut with *Sst*I prior to transfection to remove the origin of replication (lanes 1 and 2). In lane 3 both pLIRSD and pLIRWT were cut with *Sst*I before transfection. Replication is obligatory for postreplicative overlap recombination, but not homologous recombination. The sample in lane 2 was digested with *Sph*I prior to Southern blot analysis. Unfractionated total DNA was isolated after transfection and hybridized with a plasmid probe in lanes 1-3. **a** = uncut cis molecule, **b** = input, and **c** = doubly cut cis molecule. Standard = *Sph*I sensitive cleavage standard.

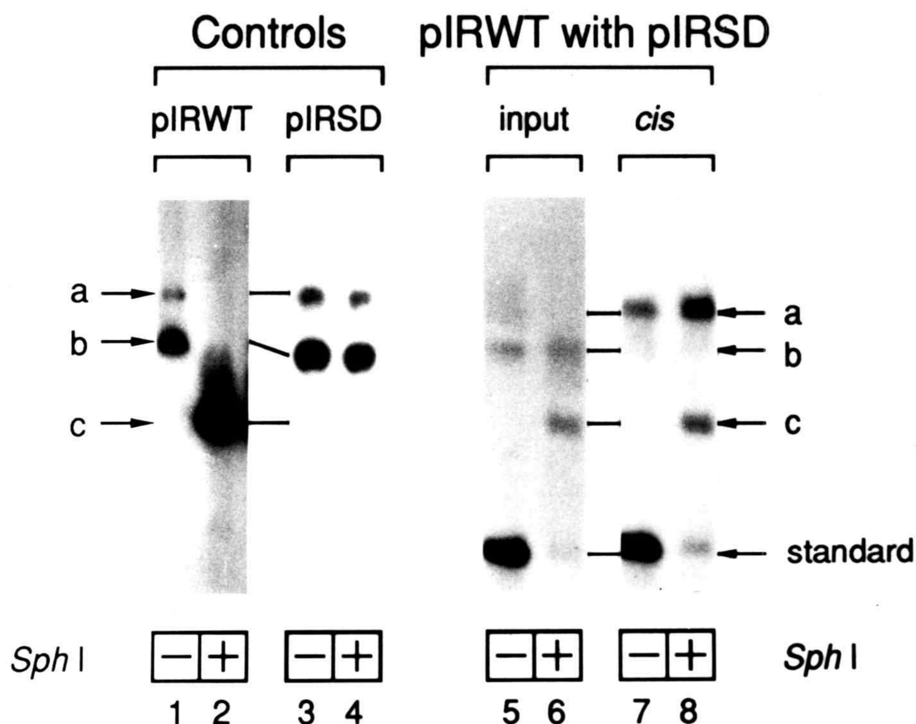
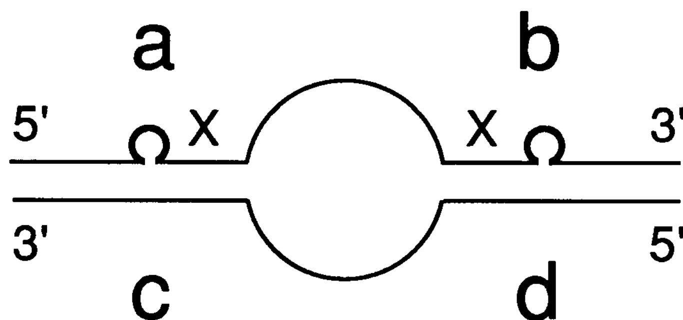
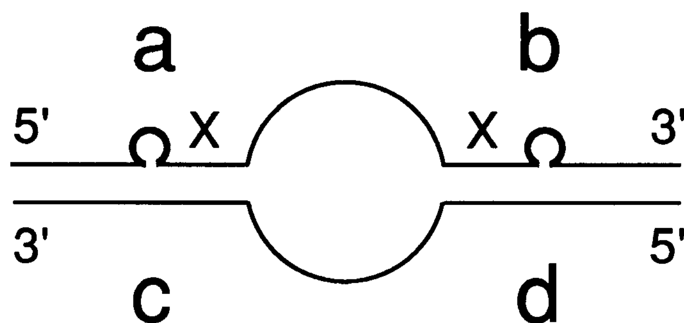


Figure IV.3. Analysis of sequence conversion during postreplicative overlap recombination between two minichromosomes with an IR of 51 bp. pIRWT and pIRSD were transfected alone (lanes 1-4), and together (lanes 5-8). Samples in even numbered lanes were digested with *Sph*I prior to Southern blot hybridization. Unfractionated total DNA was isolated after transfection and hybridized with a plasmid probe in lanes 1-4. Input DNA was purified after transfection and analyzed with a plasmid probe (lanes 5 and 6). The cis replication product was similarly gel-purified and analyzed with a plasmid probe (lanes 7 and 8). **a** = uncut cis molecule, **b** = input, and **c** = doubly cut cis molecule. Standard = *Sph*I sensitive cleavage standard.



	Loci			
	a	b	c	d
pLIRWT	+	+		
pLIRSD			-	-
<u>converted at:</u>				
a	blocked		-	-
b	+	-	-	-
c	+	+	+	-
d	+	+	-	+
a + b	blocked		-	-
a + d	blocked		-	+
b + c	+	-	+	-
c + d	+	+	+	+

Figure IV.4. Analysis of the possible outcomes of sequence conversion during *trans* replication between pLIRWT and pLIRSD. Loci a, b, c, and d identify the locations of the heterology. The plus sign (+) marks the presence of an *Sph*I site and the minus sign (-) the location of the deleted *Sph*I site. X = the distance between the bulge and the "panhandle" loop. The outcome of sequence conversion at each possible locus is listed.



	Loci			
	a	b	c	d
pIRWT	+	+		
pIRSD			-	-
converted at:				
a	blocked		-	-
b	blocked		-	-
c	+	+	blocked	
d	+	+	-	+
a + b	blocked		-	-
a + d	blocked		blocked	
b + c	blocked		blocked	
c + d	+	+	blocked	

Figure IV.5. Analysis of the possible outcomes of sequence conversion during *trans* replication between pIRSD and pIRWT. Loci a, b, c, and d identify the locations of the heterology. The plus sign (+) marks the presence of an *Sph*I site and the minus sign (-) the location of the deleted *Sph*I site. X = the distance between the bulge and the "panhandle" loop. The outcome of sequence conversion at each possible locus is listed.

CHAPTER V: DIRECTIONAL COCONVERSION DURING ADENOVIRUS MINICHROMOSOME REPLICATION

INTRODUCTION

In Chapters III and IV, sequence conversion between the IRs of adenovirus minichromosomes during complementary strand synthesis was documented. Unidirectional coconversion on replication intermediates is shown to occur in this Chapter. This verifies aspects of the model developed for the process of sequence conversion during adenovirus DNA replication. Evidence has been presented that sequence conversion attempts on "panhandle" intermediates which involved incision within small (four bp) bulges in the 5' IR created a structure that blocked repair. The removal of bulges has never been observed when positioned in the 5'-IR of the displaced strand (Fig. III.3 and Fig III.4, lane 8, no band **a**). However, the evidence indicates that repair attempts are made that lead to the creation of a nonviable genome. It appears as though endonuclease incision does happen in the 5'-IR bulges. When the bulge is close to the "panhandle" loop, endswitched double-stranded molecules are created when incision occurs in the bulge (Fig. III.3, lanes 3 and 4; Fig. III.10, f-k-n). This process creates a molecule which is unable to end-repair, and so will be unable to become a fully replicating molecule. The molecules that do not endswitch appear to be unable to complete sequence conversion because a structure that blocks repair forms. Cleavage within the loop generates a 3'- and a 5'-overhang. There is evidence that the 3'-overhang should be removable, because an active 3'-exonuclease can work on similar substrates (Wang *et al.*, 1991), and adenovirus

polymerase has a 3'-exonuclease associated with it (Field *et al.*, 1984). The 5'-overhang could be the source of the block; either a 5'-exonuclease is unable to remove the overhang or upon removal a non-ligatable structure is created. The relative orientation of the bulge to the ends of the molecules seemed to determine whether sequence conversion would be successful. Viable genomes were recovered when cleavage within a bulge in the 3'-IR was required. Shown in Chapter III are results that this process requires initiation of replication at the origin which would result in the generation of a small double-stranded molecule and a "panhandle" intermediate without the block (Fig. III.11, f-h). To test if the blocking structure was being removed I investigated the removal of two loops placed in the 3'-IR. The loop furthest away from the origin on a "panhandle" intermediate always should co-convert the other loop, but not *vice versa*. In this Chapter I show that this is the case, and discuss the possible role of the adenovirus replication proteins in creating the block.

RESULTS

Experiments were carried out to develop a better understanding of the mechanism for sequence conversion which takes place on "panhandle" intermediates and on hybridized displaced single strands during adenovirus replication. An adenovirus minichromosome system (Hay *et al.*, 1984; Wang and Pearson, 1985) was employed to study coconversion of deletions placed in a synthetic IR. Plasmids were constructed with a single left adenovirus origin of replication, flanked by 229 bp of identical sequence which includes a polylinker (Fig. III.1). When the plasmids are linearized with *Eco*RI the

adenovirus origin is exposed, and the duplicated flanking sequences are arranged as an ITR, but the origin sequences are missing from one end. The polylinker contains endonuclease restriction sites for *Sph*I and *Kpn*I. A set of molecules was constructed which had two restriction sites deleted. The *Kpn*I site was removed in every case from the 5' polylinker and the *Sph*I site was deleted from either the 3', 5' or both IRs, and named accordingly. For example, the plasmid with both restriction sites deleted from the 5'-IR was called pLIR5'K5'S. Each of these plasmids was cotransfected with adenovirus helper DNA. After 72 hours of incubation the DNA was extracted, a sample of each plasmid was digested with the appropriate restriction enzyme and analyzed by Southern blot hybridization. A cleavage standard was added to each lane before digestion, to ensure that digestion went to completion. The restriction pattern for the double mutants is presented in Fig V.I. The input DNA (band **b**) and the cis DNA (band **a**) are recovered (lanes 1, 4, 7). After endonuclease digestion of pLIR5'K5'S with *Kpn*I and *Sph*I, respectively, three bands are seen (lanes 2 and 3). Molecules which are resistant to both enzymes are retained at position **a**, minichromosomes with a single restriction site are at position **b** and molecules with two sites migrate to position **c**. Endonuclease digestion of pLIR5'K3'S reveals the same pattern for *Kpn*I digestion as above, bands **a**, **b** and **c** are generated (lane 5). *Sph*I digestion produces only molecules which are singly and doubly sensitive (lane 6). A control molecule was also tested; pLIR5'KDS has a single *Kpn*I site and no *Sph*I sites. It served as a control to establish that higher oligomers were not inadvertently migrating to position **c** when treated with *Sph*I.

As can be seen from Fig. V.1, lanes 7-9, this did not happen. From this analysis it cannot be determined if coconversion took place. Each of the plasmids produced the same restriction pattern as the respective corresponding plasmid with a single site deleted (Chapter III and Chapter VI).

Unidirectional coconversion during adenovirus DNA replication

To determine whether coconversion was taking place the following experiment was carried out. After transfection, pLIR5'K5'S was first treated with *Sph*I, and then bands **a** and **b** were gel-purified away from band **c** (compare Fig. V.2 lane 1 to Fig. V.1 lane 2), so that only molecules without *Sph*I sites were recovered. To determine if all of the purified molecules without *Sph*I sites also had lost the *Kpn*I site, the minichromosomes were then subjected to *Kpn*I cleavage (Fig. V.2, lane 2). It appears that all molecules which had undergone deletion of the *Sph*I site had also lost the *Kpn*I site; there are no molecules with two *Kpn*I sites (band **c**), and the intensity of band **b** remains the same before and after cutting, indicating that molecules from band **a** did not migrate to position **b**. When the reverse experiment is done, a different answer is obtained. It appears that coconversion did not take place, and that repair occurred independently at each site ("patchy"). In this experimetn, pLIR5'K5'S was first digested with *Kpn*I, then bands **a** and **b** were gel-purified from band **c**, followed by cutting with *Sph*I (Fig. V.2, lanes 3 and 4). Molecules which were resistant to *Kpn*I were sensitive to *Sph*I; minichromosomes without any *Kpn*I sites had two *Sph*I site (band **c**), and the molecules migrating to the position of band **b** demonstrates that a second population of molecules with one *SphI* sites

and no *Kpn*I sites had survived. In the first experiment, when coconversion took place the 5' polylinker was the template for both sites. In the second experiment only the population of molecules which had used the 5' polylinker as template to repair the *Kpn*I site (resistant molecules) were studied, and some molecules were observed with a single *Sph*I site (molecules dropping from band **a** to **b**), as well as other molecules had converted using the 3' polylinker as template (migrating from **a** to **c**). It appears these substrates can undergo both unidirectional coconversion and "patchy" repair.

A model for unidirectional coconversion

In Chapter III experiments involving sequence conversion on "panhandle" intermediates during adenovirus replication had indicated that endonuclease action within bulges in the 5'-IR created a block to the completion of repair. When the bulge was in the 3'-IR, recovery of a viable minichromosome was detected. As described in the introduction this can occur by removing the blocking structure by initiating replication at the origin (Fig. III.11, b-h). This model leads to the prediction that all conversion attempts to remove the *Sph*I site will also result in the coconversion of *Kpn*I, but not vice-versa. This is what the above experiment showed. The two contrasting events can be followed in Fig. V.3 (b-j versus b-j'). Cleavage within the *Kpn*I site (the bulge) would lead to intermediates f, followed by initiation of replication at the origin to generate intermediate h and h". The adenovirus polymerase could provide the next two functions to create intermediate j. This sequence conversion pathway will produce a molecule which is resistant to *Kpn*I but sensitive to *Sph*I. This is the result obtained in

Fig. V.2, lanes 3 and 4. In contrast endonuclease action within the *Sph*I site predicts the formation of the intermediates g' through j'. In this case every molecule which removes the *Sph*I site also removes the *Kpn*I site. The results presented in this paper confirm this prediction; all molecules that converted the *Sph*I site also converted the *Kpn*I site in the same direction, but the opposite did not take place. The same result is predicted for coconversion on hybridized displaced strands during postreplicative overlap recombination.

DISCUSSION

This Chapter demonstrates that unidirectional coconversion occurs on replication intermediates, which is consistent with the general model for sequence conversion. The prediction was made that if the loop was in the 3'-IR, then replication could initiate at the origin of replication and displace the blocking structure as a small double-stranded molecule (Fig. V.3). This also leaves a "panhandle" intermediate and/or intermolecular hybridized duplex through the IR that can now be repaired by the adenovirus polymerase using the 5'-IR as template to recover the 3' end sequences. This was testable, since any second mutations closer to the origin of replication should co-convert, with the 5'-IR serving as a template. Evidence was provided for this by showing that all molecules which have converted the *Sph*I site, when using the 5'-IR as template, also converted the *Kpn*I site by using the 5'-IR for the template (Fig. V.2, laness 1 and 2). The opposite, however, did not occur; when *Kpn*I converted using the 5' template not all *Sph*I sites co-converted (Fig. V.2, lanes 3 and 4). The directionality of coconversion supports predictions for our model of

sequence conversion on adenovirus replication intermediates. Furthermore, coconversion is triggered by the creation of a structure that blocks the repair processes at the initial site, thereby requiring the removal of the blocking structure, including all sequences between the origin of replication and the site of repair.

Studies of coconversion in mammalian cells have provided similar observations to those described in this report. Much of the work has been in analyzing sequence homogeneity in mammalian multigene families where gene conversion has been proposed to play an important role in their evolution (Edelman and Gally, 1970; Baltimore, 1981; Egel, 1981; Klein and Petes, 1981). For example, evidence for such a mechanism included the finding that the duplicated $^{A}\gamma$- and $^{G}\gamma$-fetal globin genes on the same chromosome were more alike than the allelic genes on homologous chromosomes (Slightom *et al.*, 1980). DNA sequence comparison of the human α-globin genes (Michelson and Orkin, 1983), and fetal globin genes (Stoeckert *et al.*, 1984) led the investigators to propose that sequence conversion events were "patchy", or constrained to the site of heterogeneity. There are two cases where sequence information prompted the investigator to propose that coconversion was likely, and that involved short tracts of 50 bp (Mellor *et al.*, 1983; Denaro *et al.*, 1984). This is in contrast to fungi, where coconversion has been demonstrated to occur over hundreds of nucleotides long (for review see Orr-Weaver and Szostak, 1985). Liskay and Stachelek (1986) analyzed intrachromosomal events between duplicated sequences. They found that coconversion could involve tracts as long as 358 bp under selection,

and at least 255 bp but not 450 bp when selection pressure was not on the second conversion site (silent site). One curious result was that there was an apparent directional constraint to coconversion of the silent site; coconversion 3' of the silent site was distance dependent, whereas coconversion 5' occurred only at a very low frequency, 3% versus 50% for a site 5' to the silent site at approximately the same distance away. This work, like ours, implies a constraint on the direction of coconversion. Furthermore, all the above work, in agreement with ours, indicates that conversion events can happen independently, involving only the mismatched sequences ("patchy"), but also can involve contiguous blocks of DNA leading to coconversion.

The possible role of adenovirus replication proteins in preventing successful sequence conversion attempts

The experiments described in this paper indicate that 5'-overhangs block sequence conversion attempts. This may be a result of the lack of 5' exonucleolytic action on the generated overhang, or creation of a non-ligatable structure. While there is no known 5'-exonuclease encoded for by adenovirus, the activity has been documented in mammalian cells (Ayares dt al., 1987; Guggenheimer et al., 1984a,b). Adenovius chromosomes which lack the terminal protein covalently attached to the 5' end are unable to support initiation of replication *in vitro* unless an host-encoded 5' exonuclease (factor pL) is present (Guggenheimer et al., 1984a,b). Factor pL activates adenovirus templates in vitro by degrading the 5' end of the DNA strand, which is normally the displaced strand, releasing 5' monodeoxynucleotides. If pL could act on the substrate with the 5'-overhang then an intermediate that is ligatable, or an intermediate

with extensive 5'-exonuclease action should be recoverable that is capable of replication. One possible explanation for why 5' ends are not excised is that 5'-exonucleases might be sequestered away from the replicating molecule. A second possibility is that the adenovirus DNA binding protein (DBP) might bind 5'-overhangs, thereby protecting 5'-overhangs from exonucleolytic activity, but unable to bind to 3'-overhangs. The loading of DBP onto single-stranded DNA might be unidirectional starting at the 5' end. The following suggest that this might be reasonable: (i) The 5' end is the initial end that is displaced during adenovirus strand displacement replication. (ii) Single-strand DNA binding by DBP is required for strand elongation, even for synthesis of the first 26 bases; thus, loading is normally unidirectional (Friefeld *et al.*, 1983; Prelich and Stillman 1986). (iii) DBP is an asymmetric molecule (Schechter *et al.*, 1980; van der Vliet *et al.*, 1978). (iv) binds to single stranded DNA cooperatively (van der Vleit *et al.*, 1978). (v) it has been shown DBP binds 3-11 bases per DBP molecule (van der Vliet *et al.*, 1978; Schechter *et al.*, 1980; van Amerongen *et al.*, 1987). Furthermore, single stranded DNA is asymmetric and DBP could potentially detect the polarity of the molecule. The fact that DBP binds to single-stranded DNA, double-stranded DNA, and RNA does not argue against unidirectional loading of DBP onto ssDNA. Binding of DBP to dsDNA and to RNA is quite different from that of binding ssDNA. DBP forms a stable complex only with the molecular ends of double-stranded DNA (Fowlkes *et al.*, 1979). When DBP is bound to RNA versus ssDNA, DBP sensitivity to trypsin is different, indicating that DBP binds RNA and ssDNA at different sites within the carboxyl

domain, or that the protein once bound is in a different conformation (Cleghon and Klessig, 1986). Therfore, DBP might load onto ssDNA in a unidirectional manner, protecting the 5'-overhang from exonuclease activity but not the 3'-overhang. As long as the 5'-overhang remains on the replication intermediate and without the terminal protein linked to that 5' end, strand displacement replication should not be capable of initiating at that site (Challberg and Kelly, 1979a).

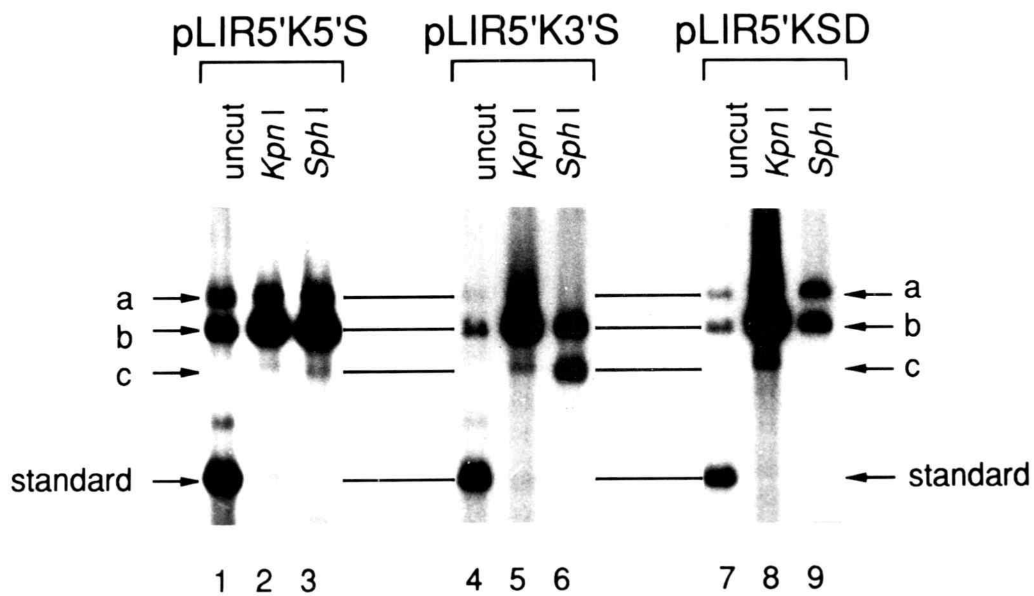


Figure V.1. Analysis of sequence conversion of the double mutants with the *Kpn*I and *Sph*I sites deleted. Each samples was singly digested with *Kpn*I and *Sph*I prior to Southern blot hybridization. Unfractionated total DNA was isolated after transection and hybridized with a plasmid probe; pLIR5'K5'S lanes 1-3, pLIR5'K3'S lanes 4-6, pLIR5'KSD lanes 7-9. **a** = uncut cis molecule, **b** = input, and **c** = doubly cut cis molecule. Standard = *Kpn*I and *Sph*I cleavage standard.

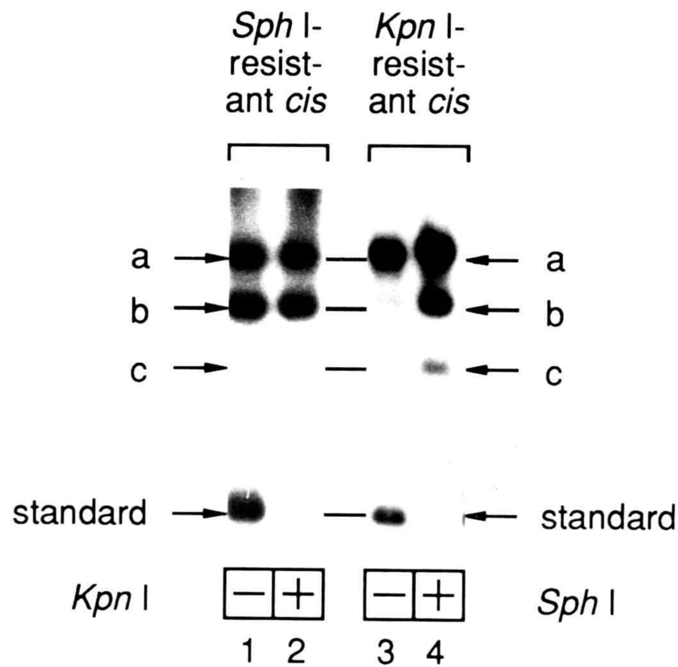


Figure V.2. Analysis of coconversion on pIR5'K5'S. Unfractionated total DNA was isolated after transection and hybridized with a plasmid probe. Endonuclease digestion with *Sph*I (lanes 1 and 2) and *Kpn*I (lanes 3 and 4) was carried out prior to gel-purification of band **a** and **b**. The gel-purified samples were then digested with the other enzyme; *Kpn*I (lanes 1 and 2), *Sph*I (lanes 3 and 4), preceeding Sothern blot hybridiztation. **a** = uncut cis molecule, **b** = input, and **c** = doubly cut cis molecule. Standard = *Kpn*I and *Sph*I cleavage standard.

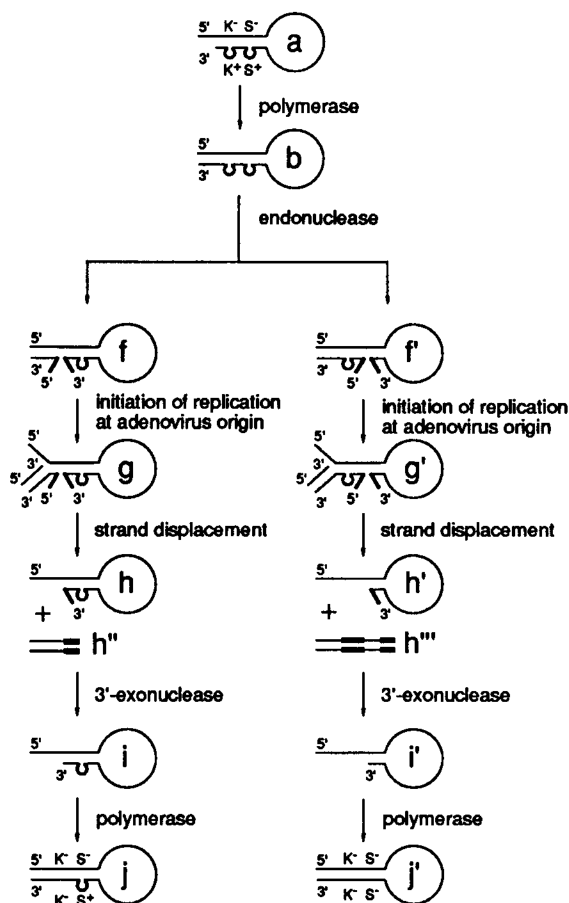


Figure V.3. Directional coconversion away from panhandle loop. The bulge or thick patch is the four central bp of the restriction endonuclease sites. K = *Kpn*I and S = *Sph*I. (a) Panhandle intermediate of transfected plasmid pLIR5'K5'S. (b) Polymerization to recover the 3'origin sequences. (f) Endonuclease cleavage within the *Kpn*I site. (g) Initiation of replication at the adenovirus origin. (h + h") Strand displacement replication removes 5'-overhang. (i) 3'-exonuclease cleavage removes 3'-overhang. (j) Polymerization creates an intermediate that underwent "patchy" repair. (f') Endonuclease cleavage in the *Sph*I site. (g') Initiation of replication at adenovirus origin. (h' + h"') Strand displacement replication removes both the *Kpn*I site and the 5'-overhang. (i') 3'-exonuclease removes 3'-overhang. (j') Polymerization creates an intermediate that coconverted both the *Sph*I site and the *Kpn*I site.

CHAPTER VI: SEQUENCE CONVERSION AT SITES CLOSE TO THE LOCATION OF ADENOVIRUS MINICHROMOSOME END-REPAIR

INTRODUCTION

In Chapter III gap filling within six bp of the "panhandle" loop could only be detected when polymerization was towards the loop but not away from the loop. Wang et al., (1991) have shown that end-repair on adenovirus minichromosomes required at least a 31 bp IR (Wang *et al.*, 1991). To gain a better understanding of this duplex requirement, three sets of mutants with small (four bp) insertion and/or deletions located within 31 bp from the site of end-repair were constructed. The basic question is whether a small bulge will prevent polymerization away from the loop, in the same way that the large "panhandle" loop does. It turns out that bulges close to the site of polymerization do not prevent replication away from that site. But a mutant was found that did not engage all sequence conversion possibilities when only 39 bp of duplex separated the bulge and the "panhandle" loop. The constraint placed on this mutant was both orientation and length dependent, much like the mutant with only six bp separating the two sites. One of the sets of mutants is unique in that the site of heterogeneity consisted of a four bp tandem duplication. The tandem duplication imparted some distinctive properties to the site of repair. For example, the tandem duplication appears to be able to branch migrate in order to facilitate the repair process.

Another set of mutants was constructed with a three bp mutation instead of four bp. This mutation behaved like all the other stable deletion mutants so far tested. Finally, a discussion follows of

how these results may provide insight into the evolution of the adenovirus ITR is presented.

RESULTS

These experiments were once again carried out in an adenovirus minichromosome replication system (Hay *et al.*, 1984; Wang and Pearson, 1985). The adenovirus minichromosomes were like those of the previous Chapters; they consisted of a single left adenovirus origin flanked by identical sequences which contained a polylinker (Fig III.1). When the plasmids are linearized with *EcoR*I the adenovirus origin is exposed, and the duplicated flanking sequences are arranged as an IR, but the replication origin sequences are missing from one end. The endonuclease restriction sites in the polylinker are used to create the heterogeneity between the IRs. The plasmids have been named to identify the restriction site *(Kpn*I=K, *Xba*I=X, *Bgl*I=B) and the polylinker (3' or 5') which had been cut and treated with Klenow; for example pLIR3'K, had the *Kpn*I site in the 3' polylinker acted on. Klenow treatment created three different types of mutants: the restriction endonuclease site for *Kpn*I has a four-bp deletion, *Bgl*I has a three-bp deletion, while the *Xba*I site has a four-bp insertion, which creates a tandem duplication. Each member of the pIR set of mutants has an IR of 51 bp, the pLIR series has an IR of 222 bp, and the *Bgl*I mutants have an IR of 144 bp. The first 51 bp of the IR for the *Bgl*I mutants (made by Jasmine Ng) are identical to the pIR, and pLIR series; the *Bgl*I mutants contain the pUC19 polylinker from *Eco*RI to *Hin*dIII. The next 70 bp are from *Nar*I to *Pvu*II in the *lac* Z gene of pUC19. The last 23 bp of the IRs consists of the sequence

between the *Bal*I and *Ava*I site from pBR322. This arrangement of sequences flanks the *Bgl*I site with 62 and 79 bp. One further set of molecules was constructed with a deletion made at the 5' *Kpn*I site, and insertions placed at the 3' or 5' *Xba*I endonuclease restriction site.

Sequence conversion when the heterogeneity is only eight bp from the site of polymerization

As pointed out in the introduction, it is of interest to know whetther a small bulge will induce a similar effect on polymerization as does the "panhandle" loop. A second question is whether or not stable sites of heterogeneity can escape repair. The *Sph*I mutants provided two examples where sequence conversion did not always take place, but the possibility for the site of heterogeneity to become a part of the large "panhandle" loop existed. All the *Kpn*I mutants had the site of heterogeneity eight bp from the 3'-OH used for end-repair. The input DNA (band **b**) and the cis DNA (band **a**) were gel purified away from each other before restriction endonuclease analysis. The results for two different size IRs are presented: the pIR series with an IR of 51 bp, and the pLIR set with a 229 bp IR, respectively. The restriction patterns for all the *Kpn*I mutants are displayed in Fig. VI.1. Band **b** from transfected plasmid pIR5'K appears to be resistant to *Kpn*I cleavage as expected, since the only *Kpn*I site is eight bp from the end of the molecule (lanes 1 and 2). This shows that the 39 bp duplex between the bulge and the "panhandle" loop is stable in contrast to pIR5'S (Fig. III.5, lane 4, band **c**). The purified band **a** consists of a population of minichromosomes with a single *Kpn*I site and with two *Kpn*I sites (lanes 3 and 4). Band **b** for plasmid pLIR5'K is also retained at position **b** after digestion, while band **a** contains a third constituent,

DNA which does not have any *Kpn*I sites (lanes 5-8). The restriction pattern for band **b** of the 3' mutants was the same; both were sensitive to *Kpn*I and migrate to position **c** as expected, since the *Kpn*I site is 346 bp form the end of the molecule (lanes 9, 10, 13, and 14). The restriction pattern for band **a** is different for the two 3' mutants; pIR3'K contains molecules with one *Kpn*I site and DNA with two *Kpn*I sites (lanes 11 and 12), while all the pLIR3'K minichromosomes are doubly sensitive to *Kpn*I (lanes 15 and 16).

To determine is the mutants can escape repair, the analysis of the restriction pattern requires consideration of contributions from both inter-molecular and intra-molecular sequence conversion. As described in Chapter III the transfected 5' molecules with a single restriction site do not necessarily indicate that they escaped repair. Minichromosomes which have undergone sequence conversion during postreplicative overlap recombination can generate molecules with one *Kpn*I site. Fig. III.13 lists the possible sequence conversion events that can arise between two displaced strands from pLIR5'S. This figure also applies for pLIR5'K. Two regions of heterology are created during inter-strand hybridization. The four possible sites of incision, within the bulge or in the strand opposite the bulge at each heterologous site, are labeled a, b, c, and d. The plus sign (+) now can represent the *Kpn*I restriction endonuclease site, while the minus sign (-) identified the site has been deleted. During intermolecular pairing between two displaced strands from pLIR5'K, the *Kpn*I sites will be positioned at loci **b** and **c**. Listed are all the possible outcomes when sequence conversion is initiated at all the possible loci. Incision at locus a, for

example will generate a molecule with two *Kpn*I sites (++) and a molecule with only one *Kpn*I site (+-). As can be seen from the figure a number of possibilities exist to generate a molecule that appears as if unconverted. These include incisions at a, b, c, d, a + b, and c + d. Since pLIR3'K does not generate any molecules with a single *Kpn*I site, this indicated that all sites of heterogeneity are converted and that both sites are always acted on. This leaves the possibility for sequence conversion initiating at a + b and a + d for pLIR5'K to generate a molecule with only one *Kpn*I site. Therefore, the regions of heterogeneity did not necessarily escape repair. For pIR5'K it is not as clear whether inter-strand sequence conversion can take place, and thereby contribute to the population of molecules with a single *Kpn*I site. In Chapter IV, when the heterogeneity was positioned close to the loop, inter-strand sequence conversion between small IRs did not take place. On pIR5'K the heterogeneity is 39 bp from the loop, so it is possible that inter-strand sequence conversion can take place, but this is not known. Therefore, it is not certain whether pIR5'K can escape repair, but as will be seen below, pIR3'K can and so it is likely this mutant also can.

A second important question is whether polymerization proceeds away from the small bulge without prior removal of the bulge. The indicator band for this is resistant DNA retained at position **a**. Diagram III.11 illustrates why the presence of band **a** diagnostic. In Chapters III and V we provided evidence that the creation of molecule j (Diagram III.11, band **a**) involved pathway f-i and not f-k or f-l. Thus, to produce molecule j, regeneration of the origin sequence must

happen before complete removal of the heterogeneity can take place. Adenovirus minichromosome pLIR5'K does generate band **a**, indicating that polymerization eight bp from a small bulge is not inhibited.

Plasmid pIR5'K does not produce band **a**. On both molecules the bulge is eight bp from the 3'-OH, but pLIR5'K has 217 bp of duplex separating the bulge from the "panhandle" loop, while pIR5'K has only 39 bp at that location (distance X). Previously Wang *et al.*, (1991) showed that only 31 bp of duplex is required for end-repair. Furthermore, removal of 3'-overhangs (14-1300 bp) on similar substrates has been demonstrated for molecules with only 37 bp of duplex (personal communication). In addition, a plasmid (pIRST39) with a 39-bp IR and a 12-bp overhang could efficiently replicate generating band **a**. pIR5'K is almost identical to this plasmid except that the last eight bp are complementary to the 5'-IR. The fact that pIRST39 works, but not pIR5'K, indicates that the last eight bp on pIR5'K forms a stable duplex, it is never removed by a 3'-exonuclease. What is blocking this pathway on pIR5'K? The steps required to generate a mutant without the restriction endonuclease site are diagramed in Fig. III.11 (a-i). The block must involve step b to g. We know that polymerization on substrate a to b can take place (evidence presented above), and that step h to i functions on pIRST39. We do not know what is blocking this pathway for pIR5'K, but it appears to be a function of the length X. First, when X is only 39 bp, and when y is eight bp, possibly the polymerase is stalled at the site of the 3'-OH before initiating end-repair, protecting the bulge from endonuclease action.

A second problem could also arise between step f to g, when X is 39 bp. The 3'-overhang could attract adenovirus polymerase, and when X is 39 bp, prevent completing the displacement of the 3'-end of the strand with the 5'-overhang (generation of h'). One further observation for pIR5'K is that the input DNA does not contain any endswitched molecules, as did pIR5'S (Compare Fig. III.5, lane 4, band **c**, to Fig IV.1, land 2, no band **c**). This demonstrates that a 39 bp duplex is fairly stable since some sequence conversion attempts are probably initiated but are unable to complete the repair (Fig. VI.1, no band **a**), yet the duplex did not unravel and generate the endswitched molecule. Finally both 5' mutants generated band **c** (Fig. III.1, lanes 4 and 8), which confirms that gap filling towards the "panhandle" loop takes place unhindered (Fig. III.9, a-e).

Plasmid pLIR3'K always repaired and in the same direction, as did pLIR3'S (Fig. III.4, lane 8, band **c**). Sequence conversion in the opposite direction is not detected on the 3'-mutants. This result is thoroughly discussed in Chapters III and V. Briefly, a block to completion of sequence conversion forms when incision within a bulge in the 5'-IR occurs (Fig. III.10). Since there is evidence of an active 3'-exonuclease that can work on these substrates (Wang *et al.,* 1991, Ahern *et al.,* 1991), and adenovirus polymerase has an active 3'-exonuclease associated with it (Field *et al.,* 1984), it is thought that the block results from the 5'-overhang or with removal of the 5' overhang a non-ligatable structure is generated. PIR3'K is different from both pIR3'S and pLIR3'K. Transfected pIR3'S did not generate any doubly sensitive molecules, yet pIR3'K did. The difference is that pIR3'S

requires polymerization away from the "panhandle" loop with only six bp separating the site of polymerization compared to 39 bp for pIR3'K. PIR3'K appears not always to undergo sequence conversion while pLIR3'K always does. The reason for this result is not clear. As described above ,pIR5'K might also be able to escape repair. Possibly, once again, this is a result of adenovirus polymerase stalled at the 3'-OH, eight bp away from the bulge, protecting the bulge from endonuclease attack.

Sequence conversion when the heterogeneity is 23 bp from the site of polymerization

The *Xba*I mutants have the site of heterogeneity 19 bp from the 3'-OH used for end-repair. Both series of plasmids (pIR and pLIR) were analyzed but bands **a** and **b** have not been gel purified away. The restriction pattern is basically the same for both sets of mutants. The 5' mutants consist of both resistant and doubly sensitive plasmids (bands **a** and **c** respectively) to *Xba*I cleavage (Fig. VI.2, lanes 1, 2 and 5, 6), while the 3' mutants appear not to have generated any molecules which are resistant (lanes 3, 4 and 7, 8). A close examination of the data reveal three unusual features. First, pIR5'X has undergone sequence conversion with gap filling away from the "panhandle" loop, even when there appears to be only 28 bp of duplex separating the bulge and the loop (Fig. VI.2, lane 2 band **a**). Second, both pIR3'X and pLIR3'X did not undergo sequence conversion when gap filling was required in the 5'-IR, which on all other molecules has always functioned even when the bulge is only six bp from the "panhandle" loop (Fig. VI.2 lanes 4 and 8, no band **a**). Third, repair that generates a 5'-overhang in the 5'-IR is successful, while in all other cases it has not

been (Chapters III and V). Upon careful examination of the sequences surrounding the altered *Xba*I site a number of special features can be identified (CCTCTAGCTAGAGT). CTAG are the four bp inserted at this site by Klenow treatment. The four-bp tandem duplication (CTAG) is flanked by a second duplication, a 2 bp duplication (CT and AG). This duplication can provide alternative base pairing at the edges of the heteroduplex, such that the first CT is not base paired, but the second CT is, thus creating two 2 bp mismatches separated by 2 bp. While we clearly have not determined what has occurred with these mutants, the base pairing at this site could be very dynamic. If this site is dynamic then explanations for the anomalies come forth. In fact, as will be discussed below, the strong possibility for branch migration through the whole region exists. Thus, when endonuclease cleavage opposite the bulge, at the farthest possible bp, occurs on plasmid pIR5'X, a 28 bp duplex with a 2 bp mismatch is created. Previously, Wang *et al.* (1991) showed that minichromosomes with a 31 bp IR could participate in end-repair but not with only a 28 bp IR. If the 31 bp duplex requirement is induced by a steric constraint on adenovirus polymerase, as is supported by pIR3'S and pIR5'S (six-bp duplex), as well as by pIR5'K (39-bp duplex), then it is possible that at 30 bp polymerization is favored over 3'-exonuclease activity. First of all, any molecules that remove the 2-bp overhang would no longer be able to end-repair. Second, if there is a steric constraint on polymerization away from the "panhandle" loop at 28 bp, there is also most likely a constraint on 3'-exonuclease activity. The crystal structure of the Klenow fragment revealed that the proposed active sites of

polymerization and of 3'-exonuclease activities are separated by 25 Å (approximately 7.4 bp of B DNA; Ollis *et al.*, 1985). The error correcting site extends in front of the polymerization site. Kinetic studies with T7 polymerase have provided a rational for this structural feature (Patel *et al.*, 1991; Wong *et al.*, 1991; Donlin *et al.*, 1991). The rate of transfer between the two active sites relative to the rate of polymerization on matched versus mismatch substrates determines which function will be active. On the 30-bp "panhandle", transfer to the exonuclease site may be inhibited, such that polymerization on the mismatched DNA becomes favored. Also, when endonuclease action happens at the same relative bp, but now in the 5'-IR, on pIR3'X and pLIR3'X a 2-bp 5'-overhang is created. I have already demonstrated that 5'-overhangs in the 5'-IR cause a block to sequence conversion (Fig. III.10, discussed in Chapters III and V). This explanation for the 3'-mutants is not completely satisfactory because incision at other positions will not create a 5'-overhang. However, this result is completely consistent with the fact that all orientation dependent blocks without a length constraint involve incision in the 5'-IR. The third unusual result is that sequence conversion attempts that involve creating 5'-overhangs in the 5'-IR on 5'-*Xba*I mutants succeed. There are differences between the 5'-overhangs created in the 5'-IR of the 3'- and 5'-*Xba*I mutants. Cleavage in one case is within the bulge (5'-mutant), and on the other substrate incision occurs in the strand opposite the bulge (3'-mutants). For the 5'-mutant a 5'-overhang that is complementary to the 3'-IR can be created, whereas, for the 3'-mutant endonuclease cleavage can create a two-bp mismatched 5'-

overhang. Consequently, the 5'-overhang in the 5'-*Xba*I mutants is special in that it consists of complementary base pairs to the 3'-IR. In conclusion, a dynamic site filled with alternative base pairing provides an explanation which is consistent with all the other results.

Sequence conversion when two regions of heterogeneity are within 31 bp from the site of polymerization

A set of molecules analyzed had the 5' *Kpn*I site deleted and the *Xba*I four-bp tandem duplication placed in either the 3', 5' or both IRs. The endonuclease restriction pattern has not changed as a result of the presence of two mutations within one molecule. The results are presented in Fig. VI.3. Uncut pLIR5'K5'X is in lane 1, cut with *Kpn*I lane 2, and cleavage with *Xba*I lane 3. PLIR5'K5'X when cut with each enzyme produces the identical patterns as do the single mutants (lane 1, 2, 3); each of the 5' mutants generate molecules which have both restriction sites (band **c**) and molecules with the restriction site deleted from both IRs (band **a**). The restriction pattern for pLIR5'K3'X matches that of the corresponding single mutant; *Kpn*I restriction pattern reveals molecules which are resistant to *Kpn*I and doubly sensitive, and the *Xba*I restriction pattern reveals that there are no molecules which lack both *Xba*I sites (no band **a**). Placing two bulges within 20 bp of the 3'-OH does not prevent end-repair. The *Kpn*I site is at eight bp from the end and the *Xba*I site is 19 or possibly 17 bp from the 3'-OH. As described above, the indicator band for end-repair without removing the bulge is the resistant band **a** for pLIR5'K. Therefore,*Kpn*I resistant **a** molecules created from pLIR5'K5'X and pLIR5'K3'X also confirm this result.

Sequence conversion on adenovirus minichromosomes with three-bp deletions

An attempt was made to gel-purify bands **a** and band **b** to analyze them separately for the three bp *Bgl*I mutants (Fig. VI.6). While the bands are not completely pure, these mutants behave exactly like the pLIR series with the *Sph*I and *Kpn*I sites altered. The *Sph*I and *Kpn*I mutants have an IR of 229 bp and the heterology is 184 bp and 217 bp from the "panhandle" loop, respectively. The *Bgl*I mutants have an IR of 144 bp and the heterology is 79 bp from the loop. The input DNA (band **b**) for pLIR5'B is resistant to *Bgl*I treatment as expected (lanes 1 and 2). The cis DNA (band **a**) consists of three constituents: DNA which is resistant to *Bgl*I digestion, molecules with one *Bgl*I site (band **b**), and plasmids with two *Bgl*I sites (band **c**; lanes 3 and 4). The transfected products from pLIR3'B are presented in lanes 5-6. Both input and cis DNA are completely sensitive to *Bgl*I (band **c**; lanes 6 and 8), and so it appears that three-bp bulges are handled the same way that four-bp mutations are. In addition, the *Bgl*I mutants are asymmetrical, such that a different sequence forms the bulge in the 3' and 5' mutants. This indicates that the sequence within the bulge does not influence the sequence conversion outcome.

DISCUSSION

The experiments above revealed a number of characteristics about the sequence conversion mechanism on adenovirus replication intermediates, which will be discussed below in detail. These include: (i) branch migration of tandem duplications, (ii) further evidence that supports a steric versus a stability constraint on the duplex length for

end-repair, and (iii) limitations to sequence conversion on replication intermediates might be reflected in the evolution of the ITR.

Special properties of the tandem duplication

Construction of the *Xba*I mutant resulted in the insertion of a tandem duplication. This tandem duplication provides the *Xba*I mutants with some special properties that result in the successful use of a pathway not available to the other mutants. The tandem duplication appears to be able to branch migrate (Fig. VI.4). Previous experiments had not detected the removal of extra sequences (three to four bp) when positioned in the 5'-IR. Incision in bulges in the 5'-IR creates a block to the completion of sequence conversion in Chapters III and V. The same result was also found here with the *Kpn*I and *Bgl*I mutants but not with the 5'-*Xba*I mutant (compare Fig. VI.1, lane 16, no band **a**, to Fig. VI.2 lanes 2 and 6, band **c**). The bulge in the 5'-*Xba*I mutants is special in that it consists of complementary base pairs to the 3'-IR. The complementarity of the base pairs provides the possibility for the bulge to branch migrate (Fig. VI.4A). When endonuclease cleavage occurs within the bulge as positioned in Fig VI.4B the 5'-overhang sequences are complementary to the sequences in the 3'-IR. The 5'-overhang can consequently base pair by simply displacing 3'-overhang sequences. As previoulsly mentioned, adenovirus polymerase has a 3'-exonuclease that can generate a ligatable substrate. In conclusion, branch migration provides an explanation for this otherwise anomalous result.

Further evidence supporting a steric *versus* a stability constraint on the duplex length for end-repair

Throughout this thesis I have discussed extensively a 31-bp duplex length requirement for the IR, in order for end-repair to occur. One possibility is that the probability of forming a duplex between the IR decreases as the length of the IR is shortened, in other words a stability constraint. A second possibility is that the enzymes responsible for end repair have a duplex requirement, and when not met cause a steric constraint. Even though we do not know why pIR5'K is unable to participate in the pathway outlined in Fig. III.11 b-g, the result is orientation-, and length-dependent. Molecules were recovered when pathway a-e was required; repair succeeds on the opposite IR. And when X = 217 bp but not when X = 39 bp pathway b-g was operative. Since end-repair can occur on molecules with an IR of 39 bp, as well as when it has a 12 bp 3'-overhang (personal communication), the low probability of duplex formation or the lack of stability of a 39-bp IR can not explain this result. In addition, the purified input band from pIR5'K does not contain any endswitched molecules, as did pIR5'S (Compare Fig. III.5, lane 4, band **c**, to Fig IV.1, land 2, no band **c**). This demonstrates that a 39-bp duplex is fairly stable since some sequence conversion attempts are probably initiated but are unable to complete the repair (Fig. VI.1, no band **a**), yet the duplex did not unravel and generate an endswitched molecule.

Plasmid pIR5'K and pIR3'K also exhibited a unique property which was also length dependent. We have shown that they are able to escape repair some of the time (Fig. VI.1, lanes 4 and 12, band **b**). This was not the case on the minichromosomes with a longer IR (Fig. VI.1,

lane 16, no band **b**). The distance that changed was the length between the heterology and the "panhandle" loop. Once again the cause of this limitation is uncertain, but it can not be a result of the reduced probability for "panhandle" formation, since "panhandle" formation is obligatory to obtain the result. These results do not absolutely rule out the possibility that the 31 bp length requirement is a function of the probability of forming a duplex, but they do show that a steric constraint is imposed on the various length requirements.

Evolution of the ITR in view of the limitations of sequence conversion close to the `panhandle' loop

To better understand the nature of the duplex length limitation for polymerization away from the "panhandle" loop, it was of interest to know whether a bulge would have a similar effect. Test plasmids (pLIRK series) were constructed with a bulge (four bp) only eight bp away from the site of end-repair. The analysis for this plasmid is the same as for pLIR5'S in Chapter III. Summarized in Fig. III.11 is how end-repair must happen before 5'-overhangs can be removed. The presence of the doubly resistant molecules (Fig. VI.1, lane 8, band **a**) from transfected pLIR5'K shows that the bulge was removed. This demonstrates that the bulge did not inhibit end-repair. The exact nature of the requirement for a 31 bp duplex for polymerization away from the "panhandle" is not yet understood, but it may have biological significance. The data from all the *Sph*I, *Kpn*I, and *Bgl*I mutants clearly indicate that small bulge that are clearly exposed are always substrates for sequence conversion. The exceptions have been when the bulges was close to the 3'-OH used for end-repair on molecules with small IRs (51 bp). One explanation for this involved the adenovirus

polymerase stalled at the bulge providing protection from endonuclease attack. Also, when the bulge was in an unstable region of duplex (pIR5'S and pIR3'S; Chapter III), such that it could transiently become part of the "panhandle" loop some minichromosomes escaped sequence conversion. While most adenoviruses have maintained the identity of their ITRs, there are a number of serotypes with some divergence at the internal junction of the ITR. First, there have been two different strains of avian adenovirus, Celo virus, with mismatches at the internal juction of the ITR. One contains two bp which are not homologous followed by the last eight bp of the ITR (Aleström *et al.,* 1982). The second strain contains the same two bp nonhomology, but is only followed by three more identical bp (Shinagawa et al., 1983). A third interesting divergence concerns a conserved hexamer (TGACGT) in all human adenovirus serotypes which is located at the internal end of the ITR. The function of the hexamer is unknown, but in the human adenovirus type 4 the left ITR hexamer has changed (TGATGT). Moreover, 4 of the last 10 bp, which includes the hexamer, are not homologous (Shinagawa *et al.,* 1983). The above examples could have arisen from insertions and/or deletions in the ITR close to the "panhandle" loop. In Chapter III, it was shown that four-bp bulges six-bp away from the "panhandle" loop could escape repair and were somewhat unstable. While we have also shown in this Chapter that bulges close to the ends of the "panhandle" could also escape repair, this only occurred on the minichromosome with an IR of 51 bp. All adenovirus ITRs range from 63 to 166 bp for different serotypes (Stillman *et al.*, 1982a; Steenbergh *et al.*, 1977; Shinagawa *et al.*, 1983;

for review see van Ormondt and Galibert, 1984). We have not tested repair of mutations placed close to the ends of the "panhandle" in this size range, but it is unlikely that they would be able to escape repair. When the IR was 229 bp all sites of heterology are substrates for sequence conversion. Moreover, base pair 9-18 is absolutely required for initiation and is highly conserved among all serotypes (van Bergen *et al.*, 1983; Tolun *et al.*, 1979; Stillman *et al.*, 1982; Wang and Pearson, 1985). While Stow (1983) has shown that adenovirus genomes that have one ITR with only 50 bp are able to recover the missing sequences, all known serotypes have longer ITRs. Possibly the selection pressure for ITRs longer than 50 bp is to be able to efficiently repair any mismatches close to the ends of the ITR.

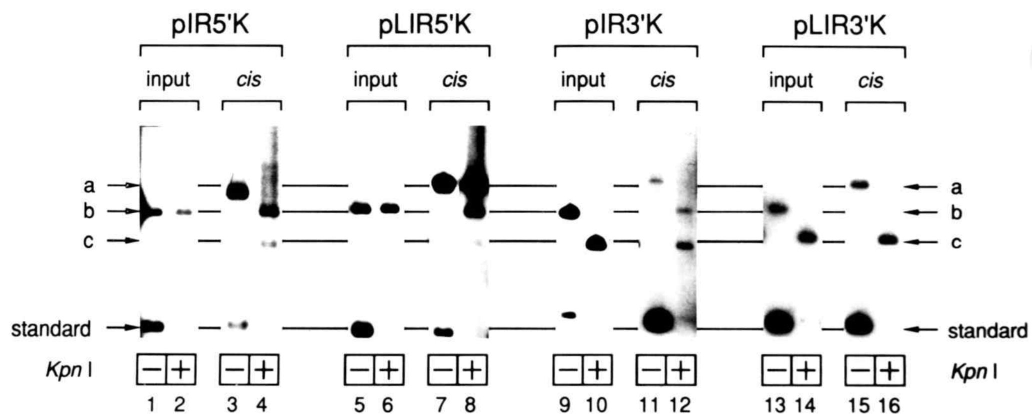


Figure VI.1. Analysis of sequence conversion of the *Kpn*I mutants. Samples in even numbered lanes were digested with *Kpn*I prior to Southern blot hybridization. Input and cis DNA was gel-purified after transfection and analyzed using a plasmid probe; pIR5'K lanes 1-4, pLIR5'K lanes 5-8, pIR3'K lane 9-12, and pLIR3'K lanes 13-16. **a** = uncut cis molecule, **b** = input, and **c** = doubly cut cis molecule. Standard = *Kpn*I sensitive cleavage standard.

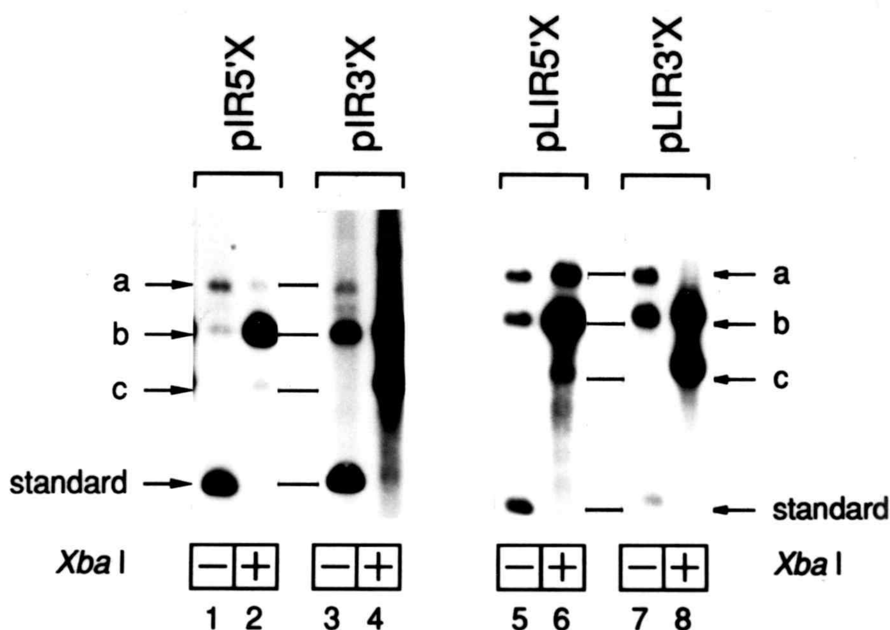


Figure VI.2. Analysis of sequence conversion of the *Xba*I mutants. Samples in even numbered lanes were digested with *Kpn*I prior to Southern blot hybridization. Unfractionated total DNA was isolated after transfection and hybridized with a plasmid probe; pIR5'X lanes 1-2, pIR3'X lanes 3-4, pLIR5'X lanes 5-6, pLIR3'X lanes 7-8. **a** = uncut cis molecule, **b** = input, and **c** = doubly cut cis molecule. Standard = *Xba*I sensitive cleavage standard.

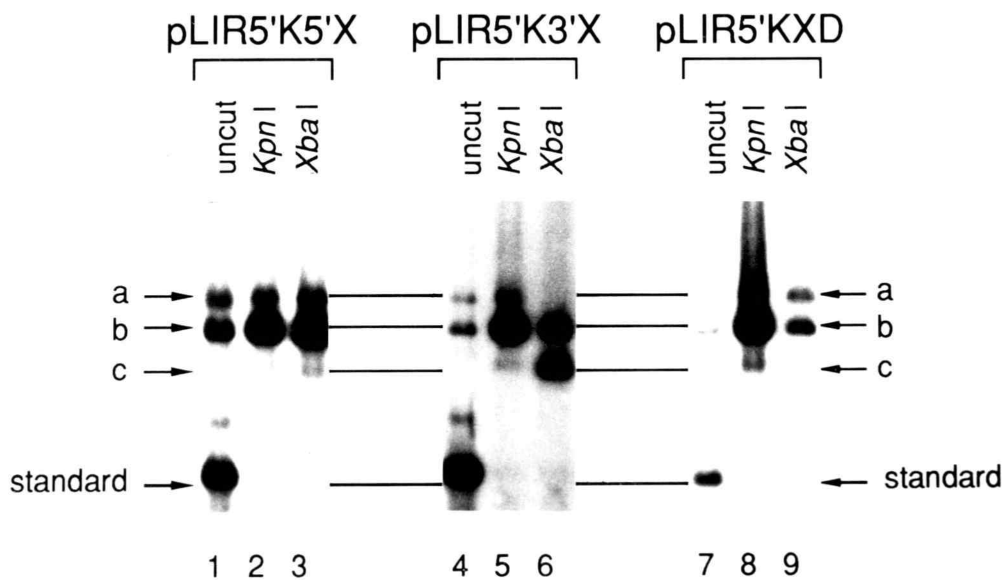


Figure VI.3. Analysis of sequence conversion of the double mutants with the *Kpn*I and *Xba*I sites altered. Each sample was singly digested with *Kpn*I or*Xba*I prior to Southern blot hybridization. Unfractionated total DNA was isolated after transfection and hybridized with a plasmid probe; pLIR5'K5'X lanes 1-3, pLIR5'K3'X lanes 4-6, pLIR5'KXD lanes 7-9. **a** = uncut cis molecule, **b** = input, and **c** = doubly cut cis molecule. Standard = *Kpn*I and *Xba*I cleavage standard.

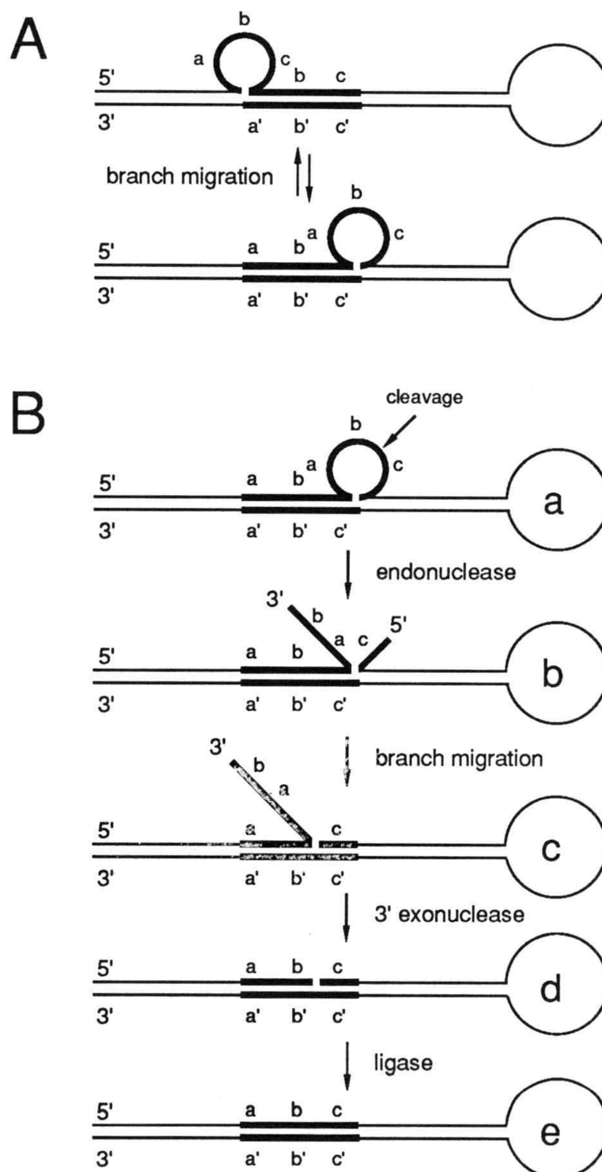


Figure VI.4. Branch migration within the *Xba*I tandem duplication. (A) Branch migration of the four bp duplicated sequences. (B) Endonuclease cleavage within the four bp bulge creates a 3'- and a 5'-overhang. Branch migration removes the 5'-overhang and 3'-exonuclease activity can remove the 3'-overhang. Ligation will restore the phosphodiester linkage.

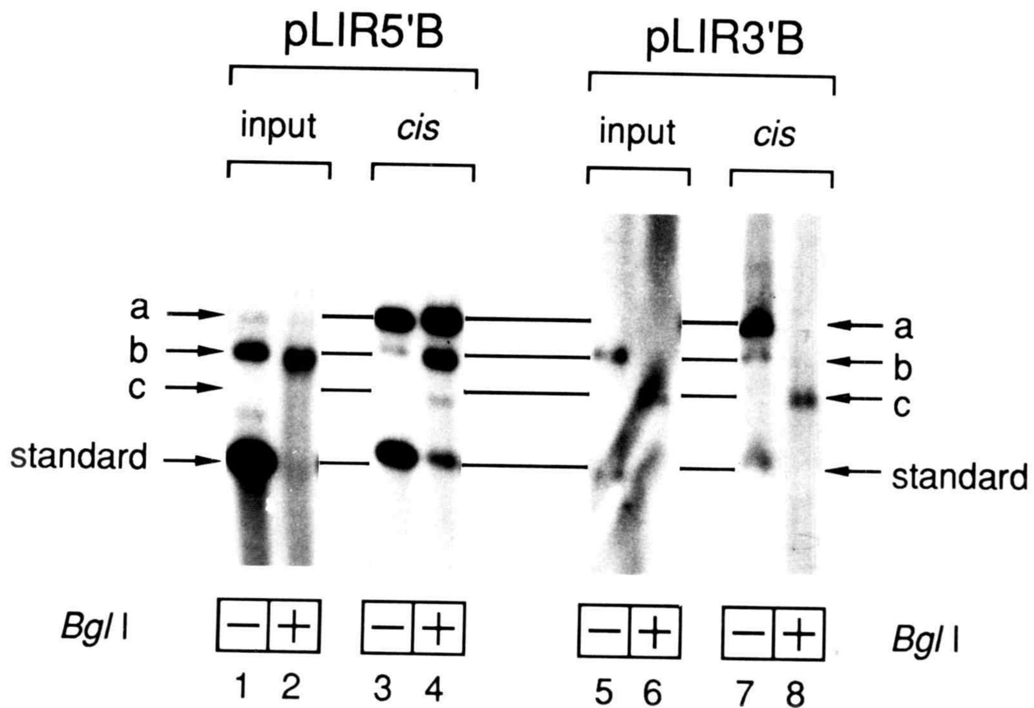


Figure VI.5. Analysis of sequence conversion on molecules with a three bp mutation. Samples in even numbered lanes were digested with *Bgl*I prior to Southern blot hybridization. Input and cis DNA was gel-purified after transfection and analyzed using a plasmid probe; pLIR5'B lanes 1-4, pLIR3'B lanes 5-8. **a** = uncut cis molecule, **b** = input, and **c** = doubly cut cis molecule. Standard = *Bgl*I sensitive cleavage standard.

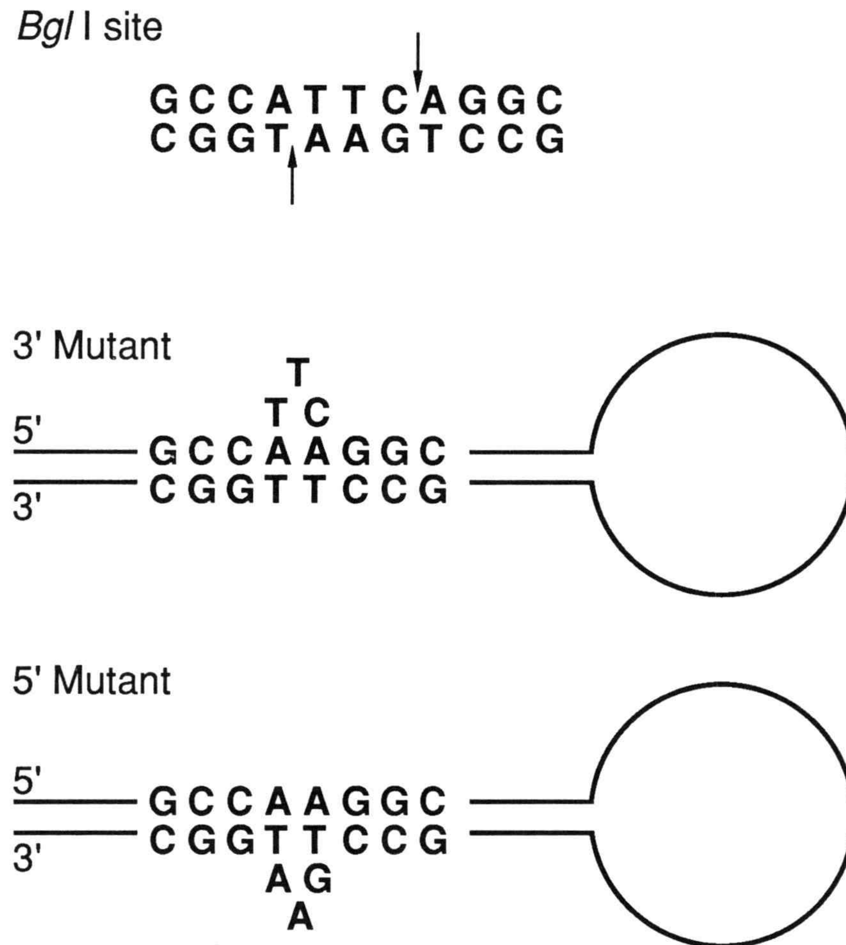


Figure VI.6. The asymmetrical sequences at the *Bgl*I mutation site. Top sequence is the full sequence of the *Bgl*I restriction endonuclease site. Arrows mark site of cleavage when digested with *Bgl*I. The 3'-mutant and 5'-mutant have different three-bp bulge sequences.

CHAPTER VII: CONCLUSIONS

The work in this thesis has provided conclusive genetic evidence that "panhandle" intermediates form during adenovirus replication. Stow (1982) demonstrated that adenovirus chromosomes lacking 51 bp from their left-hand termini were infectious and capable of regenerating the missing origin sequence. However, if an entire inverted terminal repeat (ITR) was removed, the adenovirus chromosome was no longer viable. This experiment provided the first suggestive evidence for "panhandle" formation during adenovirus replication, but it was not definitive. Homologous recombination or postreplicative overlap recombination (Ahern *et al.*, 1991) could generate the same outcome. Analysis of the segregation of markers in the IR show that homologous recombination does not mediate end-repair (Chapter III). A special case was found where postreplicative overlap recombination failed to transfer sequences between the IRs, but similar molecules could exchange sequence information during "panhandle" formation (Chapter IV). A number of constraints on sequence conversion during adenovirus DNA replication were identified. The length- and/or orientation-dependent limitations are summarized in Fig. VII.1. Replication intermediate **a** represents the orientation of the heterology (thick line bulge) in the IRs of the 5' mutants. Every 5' mutant tested was able to gap fill in the 5'-IR, even at only six bp from the "panhandle" loop (**a** to **b**). On the other hand, a length- and orientation-dependent constraint was found for the removal of bulges in the 3'-IR (**a** to **c**). When the distance X was six bp or 39 bp, bulges were not removed. When X was 79 bp, 184 bp, or 217 bp

they were successfully removed. The molecular basis for this obstruction remains to be determined. Intermediate **d** reflects the orientation of the bulge in the 3' mutants. Removal of the bulge in the 5'-IR was never seen (**d** to **e**). This constraint depended only on the orientation of the bulge and not on the length. In Chapter III, experiments showed that a block to repair at that site is created after incision of the bulge. Moreover, experiments in Chapter V show that incision of the bulge in intermediate **a** triggers directional coconversion. One final length- and orientation-dependent constraint was observed when gap filling was required in the 3' IR (**d** to **f**). When X was only six bp, the replication product requiring gap filling was not detected. Wang *et al.* (1991) have also shown that end-repair on adenovirus minichromosomes requires a 31 bp-IR. Since gap filling can occur when X is only six bp on intermediate **a**, but not on intermediate **d**, a steric constraint is imposed on repair close to the "panhandle" loop. Repair on intermediate **a** requires polymerization toward the loop, and synthesis away from the loop on intermediate **d**. It is very likely that this steric constraint reflects an asymmetry in the binding or activity of the adenovirus DNA polymerase. Below is a discussion of sequence conversion and the data that support the conclusions presented in Fig. VII.1.

A summary of the sequence conversion pathways

The proposed mechanisms for the exchange of sequences between the IRs during adenovirus replication are summarized in Figs. VII.2 and VII.3 . All 5'-mutants were able to gap-fill in the 5'-IR (Fig. VII.2). This was true even when the gap was only six bp from the

"panhandle" loop (Fig. III.5, lane 8, band **c**), and when only eight bp from the site of end-repair (Fig. VI.1, lanes 4 and 8, band **c**), as well as when the gap was approximately in the middle of the duplex (Fig. VI.5, lane 4, band **c**). These experiments show that there is no limitation to gap filling in the 5'-IR. The 3'-mutants, on the other hand, did not have bulges successfully removed from the 5'-IR (Fig. III.3, lane 8, no band **a**; Fig. III.4, lane 8, no band **a**; Fig. VI.1, lanes 12 and 16, no band **a**; Fig. VI.5 lane 8, no band **a**). Detectable endswitched molecules generated from the mutant with the bulge six bp from the loop indicate that endonuclease cleaved the bulge (Fig. III.10, f-k; Fig. III.3, lane 4, band **b**). When the bulge was within a stable part of the duplex, there was evidence that strand-displacement replication did not initiate at these nicks. An endswitched molecule would be created, but they were not found (Fig. III.4, lane 4, no band **b**; Fig. VI.I, lanes 10 and 14, no band **b**; Fig. IV.5, lane 6, no band **b**). The block to successful removal of the bulge in the 5'-IR is not known, but it most likely results from the creation of a 5'-overhang. The 5'-overhang may not be removed or a non-ligatable substrate may be created during removal. It is unlikely to be a consequence of the formation of a 3'-overhang. The adenovirus DNA polymerase is known to have a 3'-exonuclease activity (Field *et al.*, 1984), and we have previously shown that over 4000 bp of 3'-unpaired sequences are removed from structures similar to the ones used in these experiments (Wang *et al.*, 1991).

As mentioned above, gap filling in the 5'-IR can proceed unhindered. A length limitation, however, is imposed upon gap filling in the 3'-IR (Fig. VII.3). In the first case, gap filling requires

polymerization toward the "panhandle" loop; in the second case polymerization is away from the loop. A 3'-mutant with the gap at six bp did not produce sequence-converted minichromosomes (Fig III.3, lane 8, no band **c**). When the gap was at 39-bp, 79-bp, 184-bp, or 217-bp, sequence conversion occurred (Fig. VI.1, lane 12, band **c**; Fig. VI.5, lane 8, band **c**; Fig. III.4, lane 8, band **c**, Fig. VI.1, lane 16, band **c**, respectively). Wang *et al.* (1991) previously showed that end-repair happens when the IRs of adenovirus minichromosomes are 31-bp but not when the IR is 28-bp. End-repair also requires polymerization away from the "panhandle" loop. In summary, polymerization away form the "panhandle" loop is length-dependent. No length limitation was found when polymerization proceeded toward the "panhandle" loop.

The final pathway considered for sequence conversion is for the removal of bulges in the 3'-IR on the 5'-mutants (Fig. VII.2). This pathway is thought to proceed by endonuclease action within the bulge followed by adenovirus-specific replication initiating at the origin. As described above, when cleavage occurred in the bulge of the 3'-mutants, a block to sequence conversion was created. Moreover, we showed that strand-displacement replication cannot initiate at the site of the nick. This is evidence for the notion that replication must initiate at the origin to remove the block to produce the sequence converted molecule. A length limitation was also seen for proceeding through this pathway. When X was six bp or even 39 bp, this sequence conversion event was not detected (Fig. III.5, lane 8, no band **a**; Fig. VI.1, lane 4, no band **a**). When X was 62-bp, 184-bp, or 217-bp, this pathway was operative (Fig.

VI.5, lane 4, band **a**; Fig. III.6, lane 6, band **a**; Fig. VI.1, lane 8, band **a**). We do not know the exact cause for the block, but we do know that plasmid pIRST39, with a 39-bp IR and a 12-bp 3'-overhang, is able to end-repair. This plasmid is identical to pIR5'K except that the first eight bp of the 3'-overhang are complementary to the 5'-IR on plasmid pIR5'K. This suggests that if those eight bp were at all unstable and if pIR5'K could get to the stage of intermediate h, then the doubly resistant molecule should be produced. Thus, the block must take place between steps b to g. A thorough discussion of the possible causes of the block is presented in chapter VI and will not be repeated here.

End-repair and sequence conversion are sterically constrained

The first evidence for a constraint on end-repair was found when Wang *et al.* (1991) showed that a minichromosome with 28 bp IRs could not end-repair although one with 31 bp could. The nature of the constraint could not be determined from this experiment. The limitation on end-repair could be a result of a decreasing probability of forming a duplex between the IRs as the length of the IR is shortened; in other words, a stability constraint. A second possibility is that the enzymes responsible for end-repair require a minimum duplex length; for example, adenovirus DNA polymerase may not bind productively to panhandles shorter than 31 bp. Numerous experiments presented in this thesis have documented constraints on sequence conversion between the IRs. Since panhandles held together by as few as six bp can engage in sequence conversion (*e. g.*, pIR5'S), the instability of the panhandle duplex is absolutely excluded as the cause of the 31-bp length requirement. While the nature of some of the constraints are

still not understood, there is a common thread running through all of them. All constraints to sequence conversion are orientation- and/or length-dependent. Moreover, the orientation-dependent blocks always involve incision in the 5'-IR (Fig. VII.1, d-e). The length- and orientation-dependent blocks always involve incision in the 3'-IR (a-c, d-f). This orientation- and length-dependence is only consistent with steric constraints placed on using the "panhandle," not a stability constraint on forming the "panhandle". Below, I will discuss each of the experiments that demonstrate that minichromosomes were unable to complete a sequence conversion pathway successfully.

First the experiments that involved incision in the 3'-IR will be discussed (Fig. VII.1, a-c, d-f). These experiments strongly suggest a steric constraint on polymerization away from the "panhandle" loop, but not towards the loop. The two plasmids which demonstrate this asymmetrical constraint are pIR3'S and pIR5'S (Figs. III.3 and III.5). Sequence conversion attempts which involve polymerization away from the "panhandle" loop only six bp away do not take place on pIR3'S (Fig. III.3, lane 8, no band **c**). In contrast, polymerization toward the loop from six bp away does successfully proceed on pIR5'S (Fig. III.5, lane 8, band **c**). In addition, when polymerization away from the "panhandle" loop is required on pIR5'S, the sequence-converted molecule is not detected (Fig. III.5, lane 8, no band **a**). Sequence conversion on pIR5'K also revealed a constraint that depended on the distance between the bulge and the "panhandle" loop. Plasmid pIR5'K has a duplex of 39-bp (distance X) separating the site of repair from the "panhandle" loop. This is the longest length requirement observed.

Fig. VII.3 (b-j) diagrams the steps that can take place when X was 79 bp, 184 bp, or 217 bp, but not when X equaled 39 bp. We do not know the exact cause for this block, but we do know that plasmid pIRST39, that has a 39-bp IR and a 12-bp 3'-overhang, is able to end-repair. This plasmid is identical to pIR5'K except that the first eight bp of the 3'-overhang are complementary to the 5'-IR on plasmid pIR5'K. If those eight bp were at all unstable, then pIR5'K part of the time would look like intermediate h, which can end-repair. The data also demonstrate that polymerization without repair can take place on pIR5'K, as well as gap filling in the 5'-IR. This narrows the possible intermediates that could provide a block to sequence conversion to b-g. The possible nature of this block is discussed in Chapter VI, and will not be repeated here, except to point out again that successful repair appears to be dependent on the length of the duplex and the orientation of incision. One last length requirement was found. Escape from sequence conversion when the bulge was in a stable part of the duplex also depended on the distance between the bulge and the "panhandle" loop. When that length was 39 bp, sequence conversion did not happen, although it did when the distance was 79 bp, 184 bp, or 217 bp (Fig. VI.1, lane 12, band **b**; compared to Fig. VI.5, lane 8, no band **b**; Fig. III.4, lane 8, no band **b**; and Fig. IV.1, lane 16, no band **b**).

The constraints that depend only on the orientation, and not on the length of the IR, are not a result of a steric constraint imposed by the enzymes involved (V.II.1, d-e). Rather, they are a result of the asymmetry of the single-stranded DNA molecule combined with the asymmetrical direction of polymerization. Every 3'-mutant tested was

unable to complete sequence conversion attempts successfully that required incision in the 5'-IR (Fig. III.3, lane 8, no band **a**; Fig. III.4, lane 8, no band **a**; Fig. VI.1, lane 12 and 16, no band **a**; Fig. VI.2, lanes 4 and 8, no band **a**; Fig. VI.5, lane 8, no band **a**). Detectable endswitched molecules generated from the mutant with the bulge six bp from the loop indicates that endonucleolytic cleavage within the loop takes place (Fig. III.10, f-k; Fig. III.3, lane 4, band **b**). A 5'-overhang is clearly created in most cases (incision in a bulge), and for the 3'-*Xba*I mutants a 5'-overhang is probably created (Chapter III). It appears that either the 5'-overhang is not attacked by a 5'-exonuclease or a non-ligatable structure is formed after exonucleolytic removal of the overhang (Fig. VII.2, b-j). When the 5'-overhang is in the 5'-IR, sequence conversion is not completed, but it is in the 3'-IR (Chapter IV). Fig. VII.3 (b-j) shows the removal of a blocking structure in the 3'-IR by initiating strand-displacement replication at the adenovirus origin. We have shown that molecules with two bulges in the 3'-IR participate in asymmetrical coconversion. The bulge that is closest to the origin is always converted in the same direction as the second bulge when incision has taken place in the second bulge. But the reverse is not the case. An intermediate with a 5'-overhang in the 5'-IR cannot be rescued by replication initiating at the origin (Fig VII.3). These experiments strongly imply that the 5'-overhang blocks the completion of sequence conversion. The role of the adenovirus DNA replication enzymes in sequence conversion will be discussed below.

Unexpected novel properties of adenovirus DNA replication enzymes

Some sequence conversion attempts required strand-displacement replication to remove blocking structures. Those situations required all three adenovirus replication proteins. Some sequence conversion attempts may not occur as a result of an asymmetry in the adenovirus DNA polymerase (Ad pol). In addition, the presence of the adenovirus DNA binding protein (DBP) may have been responsible for the failure of some sequence conversion attempts. Finally, the absence of preterminal protein (pTP) may prevent the initiation of strand-displacement replication at sites of repair. The role of each of these enzymes in sequence conversion is now presented.

Adenovirus DNA polymerase and preterminal protein

Adenovirus DNA polymerase is thought to be involved in all steps of sequence conversion. It can provide the gap filling function, 3'-exonuclease activity, and DNA polymerization during strand-displacement replication. A number of experiments reviewed in this discussion strongly suggest that Ad pol is sterically constrained for DNA polymerization when close to large loops, but not small bulges.

The adenovirus DNA polymerase only initiates strand-displacement replication at the origin of replication. As described above, strand-displacement replication does not initiate at sites of repair that involve 5'-overhangs; an endswitched molecule would be created, but they are not found (Fig. III.4, lane 4, no band **b**; Fig. VI.I, lanes 10 and 14, no band **b**; Fig. IV.5, lane 6, no band **b**). Strand-displacement replication also does not occur at stable gaps; DNA polymerization is confined within the boundaries of the gap (Fig. III.6,

lane 4, no band **c**; Fig. VI.1, lanes 2 and 6, no band **c**; Fig. VI.5, lane 2, no band **b**). If strand-displacement replication could initiate at a gap, the 5' mutants would have generated input-sized, endswitched molecules sensitive to endonuclease restriction digestion (Fig. III.11, a-m). Origin sequences are treated differently than gaps or nicks during DNA polymerization, perhaps because only the full complement of replication proteins are assembled at the origin. The combined presence of these proteins may make the difference.

Adenovirus DNA binding protein

Endonucleases are unable to attack the "panhandle" loop. Ahern *et al.* (1991) previously showed that large single-stranded regions (greater than 500 nucleotides) are not removed during *trans* replication of adenovirus minichromosomes. It has been shown here that small bulges (three or four nucleotides) are cleaved during both *cis* and *trans* replication of adenovirus minichromosomes. One possibility is that DBP is able to bind to large loops but not small, tightly constrained ones, thereby protecting only sufficiently large loops. A thorough discussion of DBP binding to 5'-overhangs but not 3'-overhangs was also presented in Chapter V, and will not be repeated here. The binding of DBP to the 5'-overhang was proposed to provide protection from exonuclease attack.

All stable small bulges not located near the ends of small IR are always converted

As long as the heterogeneity is not protected, repair always happens. Two types of mutants were able to escape repair. The pIR *Sph*I mutants escaped repair by being close to the "panhandle" loop and transiently becoming part of the loop. The pIR *Kpn*I mutants

escaped repair by being close to the site of end-repair such that adenovirus DNA polymerase presumably protects the site. The pLIR 3' mutants, which never generate any *cis* molecules with a single endonuclease restriction site, demonstrate that all bulges are attacked (pLIR3'S, Fig. III.4, lane 8; pLIR3'K, Fig VI.1, lane 16; pLIR3'B, Fig. VI.6, lane 8, band **c**). The pLIR 5' mutants do produce *cis* molecules with a single restriction endonuclease site, but we assume that these arise during postreplicative overlap recombination by conversion at both heterologous sites (Chapter III, Figs. III.13-III.16). This analysis supports the view that all small unprotected sites of heterology are attacked.

It is likely that simple 3- and 4-bp insertion-deletion mutations, as well as tandem duplications, cause a range of perturbations in the helix. Until the individual bands have been purified for the *Xba*I mutants, we cannot be certain that every site on these mutants is always attacked, but the data thus far suggest that they are. Other repair systems do not act on every mutation. Moreover, different mutations are corrected with different frequencies. For example, different mismatches are corrected with varying frequencies by the methyl-directed mismatch repair system in *E. coli* (Lacks *et al.*, 1982; Claverys *et al.*, 1983; Lu *et al.*, 1984; Kramer *et al.*, 1984; Dohet *et al.*, 1985; Jones *et al.*, 1987; Su *et al.*, 1988; Dohet *et al.*, 1987; Dohet *et al.*, 1986; Fishel *et al.*, 1986). G-T mismatches and small insertion-deletion mutations (1-3 bp) are corrected the most efficiently (Learn and Grafstrom, 1989). To gain insight into why different mismatches are repaired with different frequencies and how the identification of the

mismatch contributes to this process, an enormous amount of work has gone into determining the structure and stability of the different mismatches. The G-T mismatch and small insertions and deletions are accommodated in the DNA duplex with striking differences. Nuclear magnetic resonance (NMR) studies (Patel *et al.,* 1982) and crystallographic studies (Kennard, 1985) have demonstrated that each base of a G-T mismatch stacks within the helix much as a conventional base-pair, but the two bases hydrogen bond as a "wobble" pair (Crick, 1966). Crystal structures of G-T pairs have also been shown to form a "wobble" pair when contained in DNA adopting the A (Brown *et al.,* 1985; Kneale *et al.,* 1985), B (Kennard, 1985), and Z (Ho *et al.,* 1985; Brown *et al.,* 1986) conformations. Furthermore, electrophoretic mobility studies have shown that the G-T mismatches do not alter the helix axis (Bhattacharyya and Lilley, 1989), whereas 1-4 bp deletions induce sharp bends or kinks (Hsieh and Griffith, 1989). Two-dimensional NMR studies have also suggested that the extra bases are stacked into the helix, creating a kink in the DNA (Hare *et al.,* 1986; Roy *et al.,* 1987). While the structures of the various mismatched bases have been correlated with repair efficiencies (Werntges *et al.,* 1986), the correlation does not extrapolate well for small deletions. So the question of why different mismatches are repaired with different efficiencies for even the very well defined system of mismatch repair in *E. coli* is still unanswered. In mammalian cells it is not known whether the repair pathway which corrects mismatched base pairs also corrects insertions and deletion. But it is known that in higher eukaryotes different mismatches are repaired with different

frequencies (Brown and Jiricny, 1988; Varlet *et al.,* 1990). Here we have shown that a variety of insertion and/or deletions are always repaired.

Sequence conversion during adenovirus DNA replication functions to maintain the identity of the origin sequences and the length of the ITR

The experiments presented in this thesis have shown that sequence conversion can operate to maintain the identity of the origin sequences for adenovirus replication. The origin of replication consists of several sequence elements with highly conserved positions. Some sequences in the elements are absolutely conserved. The adenovirus type 2 origin of DNA replication is displayed in Fig. VII.4. The first 50 bp of the ITR consists of the origin sequences for DNA replication. The remaining sequences of the ITR are of unknown function. Orientation- and length-dependent constraints on sequence conversion close to the "panhandle" loop have been defined. These constraints may play a role in minimizing the creation of subgenomic defective molecules, arising from aberrant panhandle intermediates, during adenovirus DNA replication. This suggests that the non-origin region of the ITR acts as a spacer to move the origin sequences away from the "panhandle" loop. Mutations within the origin sequences can still be completely repaired. A comparison of the ITR sequences between adenovirus serotypes reveals small clusters of GGGNGGAG tracts (G boxes) in the putative spacer region (Stillman *et al.*, 1982a). These G boxes are more closely spaced next to the "panhandle" loop; the farther away from the loop the greater the distance between the G boxes. Evidence has been presented in this thesis that demonstrates that removal of sequences close to the "panhandle" loop is inhibited. On the other hand, insertion of

sequences close to the "panhandle" loop can take place. The repeated G boxes closely clustered next to the "panhandle" loop provide a mechanism to maintain the length of the ITR by sequence conversion. Fig. VII.5 shows a single-stranded "panhandle" intermediate (a) with two repeated G boxes (**a1** and **a2**) separated by sequence **b**. These sequence elements are all within 39-bp of the "panhandle" loop. Slipped-strand mispairing during "panhandle" formation creates intermediate (b). The two bulges created by mispairing cannot be removed. Incision opposite the bulge closest to the "panhandle" loop creates a gap (c). Experiments presented in this thesis have shown that polymerization toward the "panhandle" loop can take place as close as 6 bp. In addition, it was shown that small bulges close to the site of repair do not inhibit polymerization (d). The sequence identity of the ITR can be restored after additional rounds of replication (e-g). Initiation of replication at the right origin instead of the left places the bulge in an orientation where gap filling can successfully take place (e and f). This process creates a third G box and a second sequence **b** (g). The result is a mechanism that maintains the length of the ITR in the absence of selection.

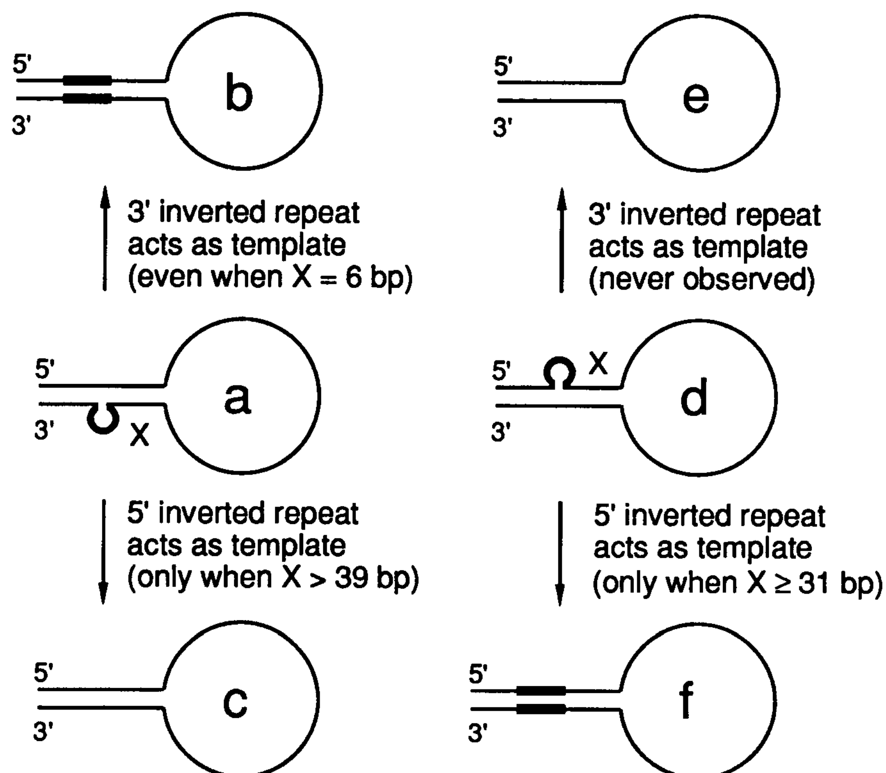


Figure VII.1. Summary of the orientation and/or length requirements for sequence conversion close to "panhandle" loops. Single-stranded adenovirus replication intermediates are shown as "panhandle" structures. The bulge (thick line) represents a restriction endonuclease site. X is the distance between the bulge and the "panhandle" loop. (a) Orientation of the bulge on 5' mutants. (b) Polymerization toward the "panhandle" loop occurs even when X = 6 bp. (c) Removal of the bulge from the 3'-IR requires more than 39 bp. (d) Orientation of the bulge on 3' mutants. (e) The removal of bulges in the 5'-IR was never observed. (f) Polymerization away from the loop requires 31 bp.

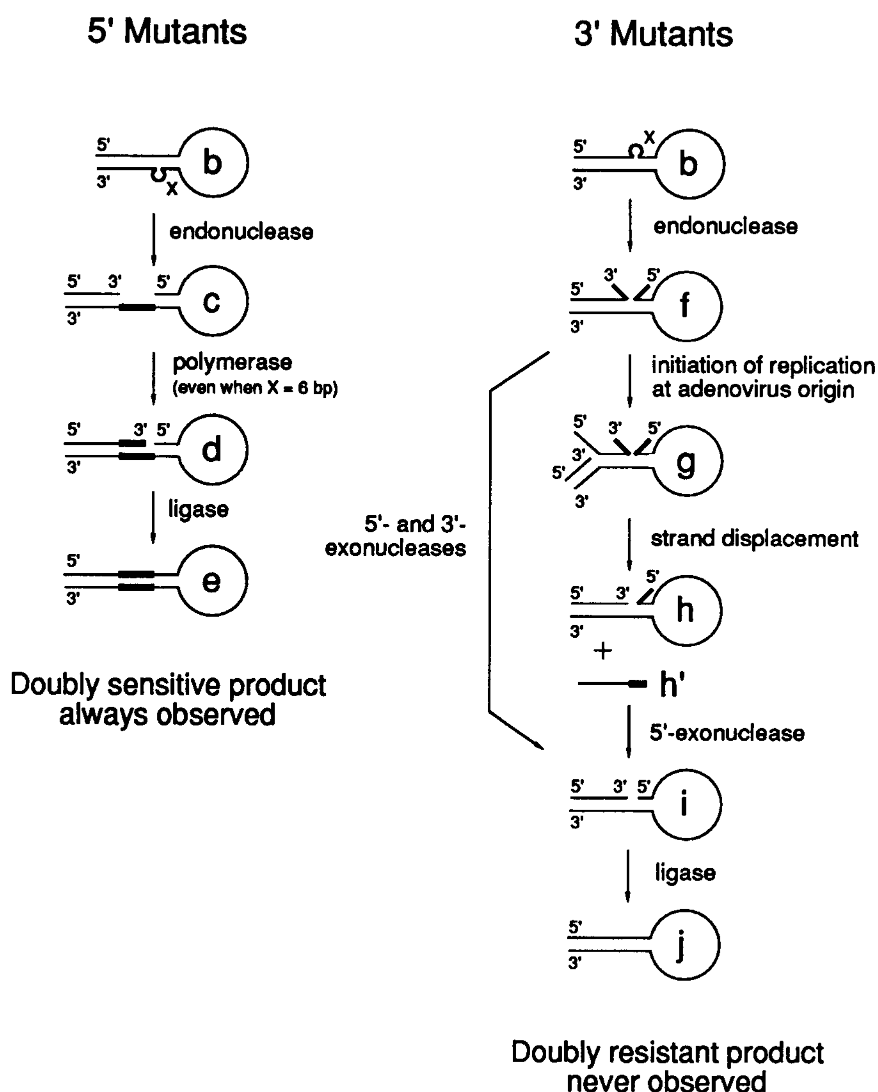


Figure VII.2. Mismatch repair toward panhandle loop. Single-stranded adenovirus replication intermediates are shown as a "panhandle" structures. The bulge (thick line) represents a restriction endonuclease site. X is the distance between the bulge and the "panhandle" loop. (b) Orientation of the bulge on 5' mutants. (c) Endonuclease creates a gap in the 5'-IR. (d) Gap filling even when X = 6 bp. (e) Ligation seals the nick. (b) Orientation of the bulge on 3' mutants. (f) Endonuclease cleavage within the bulge creates 3'- and 5'-overhangs. (g) Initiation of DNA replication at the adenovirus origin. (h + h') Strand-displacement replication removes the 3'-overhang. (i) 5'-exonuclease removes the 5'-overhang. (j) Ligation seals the nick.

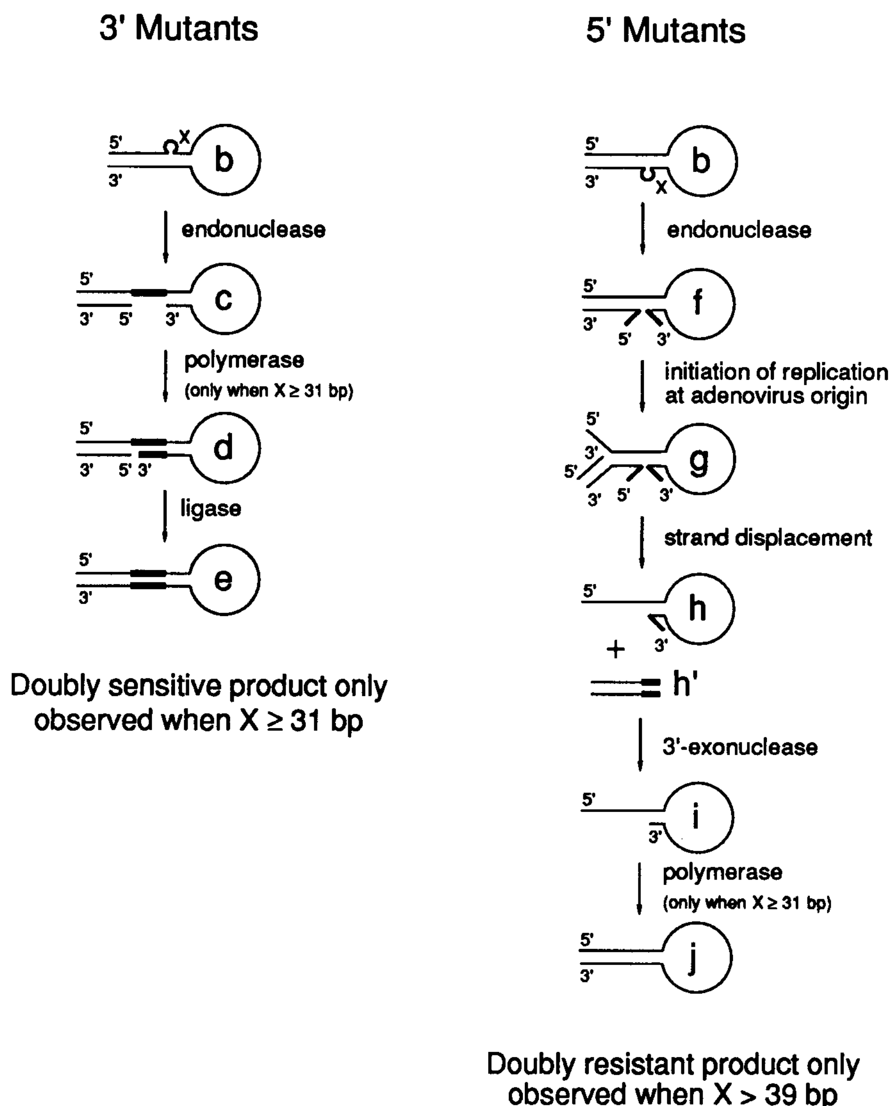


Figure VII.3. Mismatch repair away from panhandle loop. Single-stranded adenovirus replication intermediates are shown as a "panhandle" structures. The bulge (thick line) represents a restriction endonuclease site. X is the distance between the bulge and the "panhandle" loop. (b) Orientation of the bulge on 3' mutants. (c) Endonuclease cleavage creates a gap in the 3'-IR. (d) Polymerization only when X = 31 bp. (e) Ligation seals the nick. (b) Orientation of the bulge on 5' mutants. (f) Endonuclease cleavage within the bulge creates 3'- and 5'-overhangs. (g) Initiation of replication at the adenovirus origin. (h + h') Strand-displacement replication removes the 5'-overhang. (i) 3'-exonuclease removes the 3'-overhang. (j) Ligation seals the nick.

Adenovirus Type 2 Origin of DNA Replication

Adenovirus Inverted Terminal Repeat

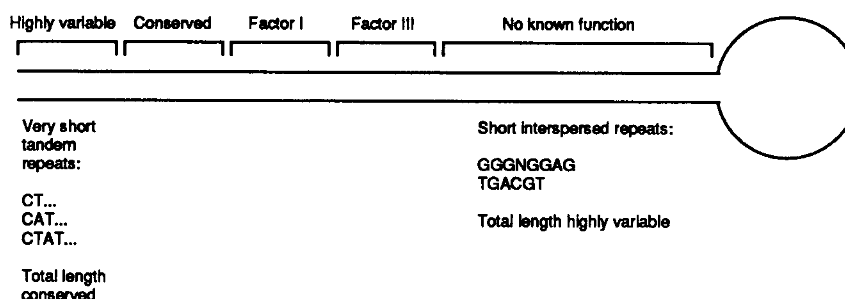


Figure VII.4. Adenovirus type 2 origin of DNA replication. The sequences for the origin of DNA replication of type 2 adenovirus resides in the first 50 bp of an approximately 100-bp ITR. Sequences 9-18 are absolutely conserved among all adenovirus serotypes. The location of these sequences is also conserved. Two cellular factors bind to the origin: nuclear factors I and III. The sequences that make up the contact points with the factors are conserved, but other sequences within this region diverge. The locations of the cellular binding sites are also conserved. The non-origin sequences in the ITR have no known function. Short, interspersed sequences are repeated in this region.

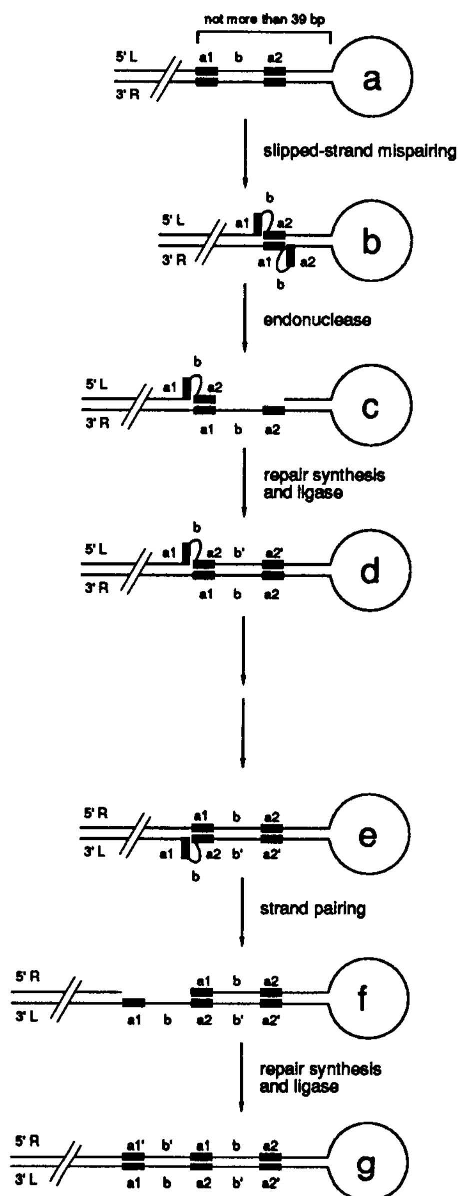


Figure VII.5. A mechanism for maintaining the length of the ITR in the absence of selection. (a) "Panhandle" intermediate initiated from the left (5' L) origin with two repeated sequences **a1** and **a2** located within 39 bp of the "panhandle" loop. (b) Slipped-strand mispairing during "panhandle" formation creates two bulges. (c) Incision in the top strand creates a repairable substrate. (d) Gap filling and ligation can take place even as close as 6 bp from the loop. (e-f) Subsequent rounds of replication with initiation at the right (5' R) origin orients the bulge such that incision in the top strand will lead to successful sequence conversion. (g) Gap filling generates a genome with a longer ITR due to the presence of a new **a**-like repeat and a new **b**-like sequence.

BIBLIOGRAPHY

Abastado, J.-P., Cami, B., Dinh, T. H., Igolen, J., and Kourilsky, P. (1984). Processing of complex heteroduplexes in *Escherichia coli* and Cos-1 monkey cells. *Proc. Natl. Acad. Sci. USA* **81**, 5792-5796.

Ahern, K. G., Wang, K., Xu, R.-Y., Mathews, C. Z., and Pearson, G. D. (1991). Strands hybridize in postreplicative adenovirus overlap recombination. *Proc. Natl. Acad. Sci. USA* **88**, 105-109.

Aleström, P., Stenlund, A., Peng, L., and Pettersson, U. (1982). A common sequence in the inverted terminal repetitions of human and avian adenoviruses. *Gene* **18**, 193-197.

Ames, B. N., and Whitfield, H. J., Jr. (1966). Frameshift mutagenesis in *Salmonella*. *Cold Spring Harbor Symp. Quant. Biol.* **31**, 221-225.

Araki, H., Jearnpipatkul, A., Tatsumi, H., Sakurai, T., Ushio, K., Muta, T., and Oshima, Y. (1985). Molecular and functional organization of yeast plasmid pSR1. *J. Mol. Biol.* **182**, 191-203.

Ayares, D., Ganea, D., Chekuri, L., Campbell, C. R., and Kucherlapati, R. (1987). Repair of single-stranded DNA nicks, gaps, and loops in mammalian cells. *Mol. Cell. Biol.* **7**, 1656-1662.

Ayares, D., Spencer, J., Schwartz, F., Morse, B., and Kucherlapati, R. (1985). Homologous recombination between autonomously replicating plasmids in mammalian cells. *Genetics* **111**, 376-388.

Baltimore, D. (1981). Gene conversion: some implications for immunoglobulin genes. *Cell* **24**, 592-594.

Berkner, K. L., and Sharp, P. A. (1983). Generation of adenovirus by transfection of plasmids. *Nucleic Acids Res.* **11**, 6003-6020.

Berns, K. I., and Kelly, T. J., Jr. (1974). Visualization of the inverted terminal repetition in adeno-associated virus DNA. *J. Mol. Biol.* **82**, 267-271.

Bhattacharyya, A., and Lilley, D. M. J. (1989). Single base mismatches in DNA: Long- and short-range structure probed by analysis of axis trajectory and local chemical reactivity. *J. Mol. Biol.* **209**, 583-597.

Bianchi, M. E., and Radding, C. M. (1983). Insertions, deletions and mismatches in heteroduplex DNA made by RecA protein. *Cell* **35**, 511-520.

Bodnar, J. W., and Pearson, G. D. (1980). Kinetics of adenovirus DNA

replication. II. Initiation of adenovirus DNA replication. *Virology* **105**, 357-370.

Bohenzky, R. A., and Berns, K. I. (1989). Interactions between the termini of adeno-associated virus DNA. *J. Mol. Biol.* **206**, 91-100.

Brown, T. C., and Jiricny, J. (1987). A specific mismatch repair event protects mammalian cells from loss of 5-methylcytosine. *Cell* **50**, 945-950.

Brown, T. C., and Jiricny, J. (1988). Different base/base mispairs are corrected with different efficiencies and specificities in monkey kidney cells. *Cell* **54**, 705-711.

Brown, T., Kennard, O., Kneale, G., and Rabinovich, D. (1985). High-resolution structure of a DNA helix containing mismatched base pairs. *Nature* **315**, 604-606.

Brown, T. Kneale, G., Hunter, W. M., and Kennard, O. (1986). Structural characterization of the bromouracilguanine base pair mismatch in a Z-DNA fragment. *Nucl. Acids Res.* **14**, 1801-1809.

Challberg, M. D., Desiderio, S. V., and Kelly, T. J., Jr. (1980). Adenovirus DNA replication *in vitro*: Characterization of a protein covalently linked to nascent DNA strands. *Proc. Natl. Acad. Sci. USA* **77**, 5105-5109.

Challberg, M. D., and Kelly, T. J., Jr. (1979a). Adenovirus DNA replication *in vitro*: origin and direction of daughter strand synthesis. *J. Mol. Biol.* **135**, 999-1012.

Challberg, M. D., and Kelly, T. J., Jr. (1979b). Adenovirus DNA replication *in vitro*. *Proc. Natl. Acad. Sci. USA* **76**, 655-659.

Challberg, M. D., and Kelly, T. J., Jr. (1981). Processing of the adenovirus terminal protein. *J. Virol.* **38**, 272-277.

Challberg, M. D., and Kelly, T. J., Jr. (1989). Animal virus DNA replication. *Annu. Rev. Biochem.* **58**, 671-717.

Challberg, M. D., Ostrove, J. M., and Kelly, T. J., Jr. (1982). Initiation of adenovirus DNA replication: detection of covalent complexes between nucleotide and the 80 Kd terminal protein. *J. Virol.* **41**, 265-270.

Chinnadurai, G., Chinnadurai, S., and Brusca, J. (1979). Physical mapping of a large-plaque mutation of adenovirus Type 2. *J. Virol.* **32**, 623-628.

Chomczyhske, P., and Qasba, P. K. (1984). Alkaline transfer of DNA to plastic membrane. *Biochem. Biophys. Acta* **782**, 67-75.

Claverys, J.-P., Mejian, V., Gasc, A.-M., and Sicard, A. M. (1983). Mismatch repair in *Streptococcus pneumoniae*: Relationship between base mismatches and transformation efficiencies. *Proc. Natl. Acad. Sci. USA* **80**, 5956-5960.

Cleghon, V. G., and Klessig, D. F. (1986). Association of the adenovirus DNA-binding protein with RNA both *in vitro* and *in vivo*. *Proc. Natl. Acad. Sci. USA* **83**, 8947-8951.

Crick, F. S. C. (1966). Codon-anticodon pairing: The wobble hypothesis. *J. Mol. Biol.* **19**, 548-555.

Daniell, E. (1976). Genome structure of incomplete particles of adenovirus. *J. Virol.* **19**, 685-708.

Das Gupta, C., and Radding, C. M. (1982). Polar branch migration promoted by recA protein: Effect of mismatched base pairs. *Proc. Natl. Acad. Sci. USA* **79**, 762-766.

de Boer, J. G., and Ripley, L. S. (1984). DNA secondary structures generate frameshift mutations. *Proc. Natl. Acad. Sci. USA* **81**, 5528-5531.

de Boer, J. G., and Ripley, L. S. (1988). An *in vitro* assay for frameshift mutations: Hot spots for deletions of 1 bp by Klenow-fragment polymerase share a consensus DNA sequence. *Genetics* **118**, 181-191.

Delius, H., and Clements, J. B. (1976). A partial denaturation map of herpes simplex virus tryp 1 DNA: evidence for inversions of the unique DNA regions. *J. Gen. Virol.* **2**, 357-364.

Denaro, M., Hammerling, U., Rask, L., and Peterson, P. A. (1984). The $E^{b}{}_{\beta}$ gene may have acted as the donor gene in a gene conversion-like event generating the $A^{bm12}{}_{\beta}$ mutant. *EMBO J.* **3**, 2029-2032.

Desiderio, S. V., and Kelly, T. J., Jr. (1981). Structure of the linkage between adenovirus DNA and the 55,000 molecular weight terminal protein. *J. Mol. Biol.* **145**, 319-337.

D'Halluin, J. C., and Milleville, M. (1984). Initiation rate of adenovirus DNA synthesis in infected cell. *Biochim. Biophys. Acta.* **782**, 67-75.

Dohet, C., Dzidie, S., Wagner, R., and Radman, M. (1987). Large non-homology in heteroduplex DNA is processed differently than single base pair mismatches. *Mol. Gen. Genet.* **206**, 181-184.

Dohet, C., Wagner, R., and Radman, M. (1985). Repair of defined single base pair mismatches in Escherichia coli. *Proc. Natl. Acad. Sci. USA* **82**, 503-505.

Dohet, C., Wagner, R., and Radman, M. (1986). Methyl-directed repair of frameshift mutations in heteroduplex DNA. *Proc. Natl. Acad. Sci. USA* **83**, 3395-3397.

Dong-Soo, I., and Muzyczka, N. (1990). The AAV origin binding protein Rep68 is an ATP-dependent site-specific endonuclease with DNA helicase activity. *Cell* **61**, 447-457.

Donlin, M. J., Patel, S. S., and Johnson, K. A. (1991). Kinetic partitioning between the exonuclease and polymerase sites in DNA error correction. *Biochemistry* **30**, 538-546.

Edelman, G. M., and Gally, J. A. (1970). Neuroscience : Second Study Program (Rockefeller Univ. Press, New York), 926-972.

Egel, R. (1981). Intergenic conversion and reiterated genes. *Nature* **290**, 191-192.

Ellens, D. J., Sissenbach, J. S., and Jansz, H. S. (1974). Studies on the mechanism of replication of adenovirus DNA III: electron microscopy of replicating DNA. *Virology* **61**, 427-442.

Enomoto, T., Lichy, J. H., Ikeda, J. E., and Hurwitz, J. (1981). Adenovirus DNA replication *in vitro*: Purification of the terminal protein in a functional form. *Proc. Natl. Acad. Sci. USA* **78**, 6779-6783.

Ensinger, M. G., and Ginsberg, J. S. (1972). Selection and preliminary characterization of temperature sensitive mutants of type 5 adenovirus. *J. Virol.* **10**, 328-339.

Farabaugh, P. J., Schmeissner, U., Hofer, M., and Miller, J. H. (1978). Genetic studies of the *lac* repressor. VII On the molecular nature of spontaneous hot spots in the *lac*I gene of *Escherichia coli*. *J. Mol. Biol.* **126**, 847-863.

Feinberg, A. P., and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochem.* **132**, 6-13.

Feinberg, A. P., and Vogelstein, B. (1984). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. (Addendum to Feinberg and Vogelstein, 1983). *Analytical Biochem.* **137**, 266-267.

Field, J., Gronostajski, R. M., and Hurwitz, J. (1984). Properties of the adenovirus DNA polymerase. *J. Biol. Chem.* **259**, 9487-9495.

Fink, G. R., and Petes, T. D. (1984). Gene conversion in the absence of reciprocal recombination. *Nature* **310**, 728-729.

Fishel, R. A., Siegel, E. C., and Kolodner, R. (1986). Gene conversion in *Escherichia coli*: resolution of heteroallelic nucleotides by co-repair. *J. Mol. Biol.* **188**, 147-157.

Fitch, W. M. (1973). Aspects of Molecular Evolution. *Annu. Rev. Genet.* **7**, 343-380.

Folger, K. R., Thomas, K., and Capecchi, M. R. (1985). Efficient correction of mismatched bases in plasmid heteroduplexes injected into cultured mammalian cell nuclei. *Mol. Cell. Biol.* **5**, 70-74.

Fowlkes, K. M., Lord, S. T., Linne, T., Pettersson, U., and Philipson, L. (1979). Interaction between the adenovirus DNA-binding protein and double-stranded DNA. *J. Mol. Biol.* **132**, 163-180.

Friefeld, B. R., Krevolin, M. D., and Horwitz, M. S. (1983a). Effects of the adenovirus H5*ts*125 and H5*ts*107 DNA binding proteins on DNA replication *in vitro*. *Virology* **124**, 380-389.

Friefeld, B. R., Lichy, J. H., Hurwitz, J., and Horwitz, M. S. (1983b). Evidence for an altered adenovirus DNA polymerase in cells infected with the mutant H5*ts*149. *Proc. Natl. Acad. Sci. USA* **80**, 1589-1593.

Frost, D., and Williams, J. (1978). Mapping temperature-sensitive and host range mutations of adenovirus type 5 by marker rescue. *Virology* **91**, 39-50.

Garon, C. F., Berry, K. W., and Rose, J. A. (1972). Arrangement of sequences in the inverted terminal repetition of adenovirus 18 DNA. *Proc. Nat. Acad. Sci. USA* **72**, 3039-3043.

Ghosh-Choudhury, G., Haj-Ahmad, Y., and Graham, F. L. (1987). Protein IX, a minor component of the human adenovirus capsid, is essential for the packaging of full length genomes. *EMBO J.* **6**, 1733-1739.

Ginsberg, J. S., Lundhoml, V., and Linne, T. (1977). Adenovirus DNA-binding protein in cells infected with wild type 5 adenovirus and two DNA-minus, temperature sensitive mutants, H5*ts*125 and H5*ts*149. *J. Virol.* **23**, 142-151.

Glazer, P. M., Sarkar, S. N., Chisholm, G. E., and Summers, W. C. (1987). DNA mismatch repair detected in human cell extracts. *Mol. Cell. Biol.* **7**, 218-224.

Graham, F. L. (1984). Covalently closed circles of human adenovirus DNA are infectious. *EMBO J.* **3**, 2917-2922.

Graham, F. S., Rudy, J., and Binkley, P. (1989). Infectious circular DNA of human adenovirus type 5: regeneration of viral DNA termini from molecules lacking terminal sequences. *EMBO J.* **8**, 2077-2085.

Graham, F. L., and van der Eb, A. J. (1973). A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52**,456-467.

Grodzicker, T. (1980) In Tooze, J. (ed.), The molecular biology of tumor viruses. Part 2. DNA Tumor Viruses. Cold Spring Harbor Laboratory Press, New York, 577-614.

Guggenheimer, R. A., Nagata, K., Kenny, M., and Hurwitz, J. (1984a). Protein-primed replication of plasmids containing the terminus of the adenovirus genome. II. Purification and characterization of a host protein required for the replication of DNA templates devoid of the terminal protein. *J. Biol. Chem.* **259**, 7815-7825.

Guggenheimer, R. A., Nagata, K., Lindenbaum, J., and Hurwitz, J. (1984b). Protein-primed replication of plasmids containing the terminus of the adenovirus genome. I. Characterization of an *in vitro* DNA replication system dependent on adenoviral DNA sequences. *J. Biol. Chem.* **259**, 7807-7814.

Haj-Ahamd, Y., and Graham, F. L. (1986). Characterization of an adenovirus type 5 mutant carrying embedded inverted terminal repeats. *Virology* **153**, 22-34.

Hanahan, D. (1983). Studies on transformation of Escherichia coli with plasmids. *J. Mol. Biol.* **166**, 557-580.

Hare, D., Shapiro, L., and Patel, D. J. (1986). Extrahelical Adenosine stacks into right-handed DNA: Solution conformation of the d(C-G-C-A-G-A-G-C-T-C-G-C-G) duplex deduced from distance geometry analysis of nuclear overhauser effect spectra. *Biochemistry* **25**, 7456-7464.

Hare, J. T., and Taylor, J. H. (1985). One role for DNA methylation in vertebrate cells is strand discrimination in mismatch repair. *Proc. Natl. Acad. Sci.* USA **82**, 7350-7354.

Hauswirth, W. W., and Berns, K. I. (1977). Origin and termination of adeno-associated virus DNA replication. *Virology*, **78**, 488-499.

Hauswirth, W. W., and Berns, K. I. (1979). Adeno-associated virus DNA replication: Nonunit-length molecules. *Virology*, **93**, 57-68.

Hay, R. T., Stow, N. D., and McDougall, I. M. (1984). Replication of adenovirus mini-chromosomes. *J. Mol. Biol.* **175**, 493-510.

Hayward, G. S., Jacob, R. J., Wadsworth, S. C., and Roizman, B. (1975). Anatomy of herpes simplex virus DNA: Evidence for four populations of molecules that differ in the relative orientations of their long and short components. *Proc. Natl. Acad. Sci.* **72**, USA 4243-4247.

Heumann, J. M. (1976). A model for replication of the end of linear chromosomes. *Nucl. Acids Res.* **3**, 3167-3171.

Hirt, B. (1967). Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* **26**, 365-395.

Ho, P. S., Frederick, C. A., Quigley, G. J., van der Marel, G. A., van Boom, J. H., Wang, A. H.-J., and Rich, A. (1985). G-T wobble base-pairing in Z-DNA at 1.0 Å atomic resolution: the crystal structure of d(CGCGTG). *EMBO J.* **4**, 3617-3623.

Holliday, R. (1962). Mutation and replication in *Ustilago maydis*. *Genet. Res.* **3**, 472-486.

Holliday, R. (1964). A mechanism for gene conversion in fungi. *Genet. Res.* **5**, 282-304.

Hsieh, C.-H., and Griffith, J. D. (1989). Deletions of bases in one strand of duplex DNA, in contrast to single-base mismatches, produce highly kinked molecules: Possible relevance to the folding of single-stranded nucleic acids. *Proc. Natl. Acad. Sci. USA* **86**, 4833-4837.

Hu, C.-H. (1990). Structural requirements for adenovirus DNA replication. Ph.D Thesis, Oregon State University, p. 26.

Ikeda, J.-E., Enomoto, T., and Jurwitz, J. (1982). Adenoviral protein-primed DNA replication *in vitro*. Proc. Natl. Acad, Sci. U.S.A. **79**, 2442-2446.

Jacob, R. J., Morse, L. S., and Roizman, B. (1979). Anatomy of herpes simplex virus DNA. XII. Accumulation of head to tail concatemers in nuclei of infected cells and their role in the generation of the four isomeric arrangements of viral DNA. *J. Virol.* **29**, 448-457.

Jearnpipatkul, A., Araki, H., and Oshima, Y. (1987). Factors encoded by and affecting the holding stability of yeast plasmid pSR1. *Mol. Gen. Genet.* **206**, 88-94.

Jones, K. A., Kadonaga, J. T., Rosenfeld, P. J., Kelly, T. J., and Tjian, R. (1987). A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. *Cell* **48**, 79-89.

Jones, N., and Shenk, T. (1978). Isolation of deletion and substitution mutants of adenovirus type 5. *Cell* **13**, 181-188.

Jones, M., Wagner, R., and Radman, M. (1987). Repair of a mismatch is influenced by the base composition of the surrounding nucleotide sequence. *Genetics* **115**, 605-610.

Kaplan, L. M., Ariga, H., Hurwitz, J., and Horwitz, M. S. (1979). Complementation of the temperature sensitive defect in H5*ts*125 adenovirus DNA replication *in vitro*. *Proc. Natl. Acad. Sci. USA* **76**, 5534-5538.

Kennard, O. (1985). Structural studies of DNA fragments: The G-T wobble base pair in A, B and Z DNA; The G-A base pair in B-DNA. *J. Biomol. Struct. Dyn.* **3**, 205-226.

Kinloch, R., Mackay, N., and Mautner, V. (1984). Adenovirus hexon: Sequence comparison of subgroup C serotypes 2 and 5. *J. Biol. Chem.* **259**, 6231-6436.

Klein, J. S. (1984). Lack of association between intrachromosomal gene conversion and reciprocal exchange. *Nature* **310**, 748-749.

Klein, H. L., and Petes, T. D. (1981). Intrachromosomal gene conversion in yeast. *Nature* **289**, 144-148.

Klessig, D. F. (1984). In Ginsberg, H.S. (ed.), The Adenoviruses. Plenum Press, New York, 399-449.

Kneale, G., Brown, T., Kennard, O., and Rabinovich, D. (1985). G-T Base-pairs in a DNA Helix: The crystal structure of d(G-G-G-G-T-C-C-C). *J. Mol. Biol.* **186**, 805-814.

Knipe, D. M., Ruyechan, W. T., Roizman, B., and Halliburton, I. W. (1978). Molecular genetics of herpes simplex virus: Demonstration of regions of obligatory and nonobligatory identity within diploid regions of the genome by sequence replacement and insertion. *Proc. Natl. Acad. Sci. USA* **75**, 3896-3900.

Knipe, D. M., Tuyechan, W. T., Honess, R. W., and Roizman, B. (1979). Molecular genetics of herpes simplex virus: The terminal a sequences of the L and S components are obligatorily identical and constitute a part of a structural gene mapping predominantly in the S component. *Proc. Natl. Acad. Sci. USA* **76**, 4534-4538.

Kozcot, F. J., Carter, B. J., Garon, C. F., and Rose, J. A. (1973). Self-complementarity of terminal sequences within plus or minus strands of adenovirus-associated virus DNA. *Proc. Natl. Acad. Sci. USA* **70**, 215-219.

Kramer, B., Kramer, W., and Fritz, H. J. (1984). Different base/base mismatches are corrected with different efficiencies by the methyl-directed DNA mismatch-repair system of *E. coli*. *Cell* **38**, 879-887.

Kunkel, T. A. (1986). Frameshift mutagenesis by eucaryotic DNA polymerases *in vitro*. *J. Biol. Chem.* **261**, 13581-13587.

Lacks, S. A., Dunn, J. J., and Greenberg, B. (1982). Identification of base mismatches recognized by the heteroduplex DNA-repair system in *Streptococcus pneumoniae*. *Cell* **31**, 327-336.

Learn, B. A., and Grafstrom, R. H. (1989). Methyl-directed repair of frameshift heteroduplexes in cell extracts from *Escherichia coli*. *J. Bact.* **171**, 6473-6481.

Lechner, R. L., and Kelly, T. J., Jr. (1977). The structure of replicating adenovirus 2 DNA molecules. *Cell* **12**, 1007-1020.

Leegwater, P. A. J., Rombouts, R. F. A., and van der Vliet. P. C. (1988). Adenovirus DNA replication *in vitro*: duplication of single-stranded DNA containing a panhandle structure. *Biochimica et Biophysica Acta*, **951**, 403-410.

Lichy, J. H., Field, J., Horwitz, M. S., and Hurwitz, J. (1982). Separation of the adenovirus terminal protein precursor from its associated DNA polymerase: role of both proteins in the initiation of adenovirus DNA replication. *Proc. Natl. Acad. Sci. USA* **79**, 5225-5229.

Lichy, J. H., Horwitz, M. S., and Hurwitz, J. (1981). Formation of a covalent complex between the 80,000 dalton adenovirus terminal protein and 5'-dCMP *in vitro*. *Proc. Natl. Acad. Sci. USA* **78**, 2678-2682.

Lindegren, C. C. (1953). Gene conversion in *Saccharomyces*. *J. Cell. Comp. Physiol.* **45**, 75-107.

Lindenbaum, J. O., Field, J., and Hurwitz, J. (1986). The adenovirus DNA binding protein and adenovirus DNA polymerase interact to catalyze elongation of primed DNA templates. *J. Biol. Chem.* **261**, 10218-10227.

Lippé, R., and Graham, F. L. (1989). Adenoviruses with nonidentical terminal sequences are viable. *J. Virol.* **63**, 5133-5141.

Liskay, R. M., and Stachelek, J. L. (1986). Information transfer between duplicated chromosomal sequences in mammalian cells involves contiguous regions of DNA. *Proc. Natl. Acad. Sci. USA* **83**, 1802-1806.

Livneh, Z., and Lehman, I. R. (1982). Recombinational bypass of pyrimidine dimers promoted by the recA protein of *Escherichia coli. Proc. Natl. Acad. Sci. USA* **79**, 3171-3175.

Lohman, T. M. (1986). Kinetics of protein-nucleic acid interactions: use of salt effects to probe mechanisms of interaction. *CRC Crit. Rev. Biochem.* **19**, 191-245.

Lu, A.-L., Clark, S., and Modrich, P. (1984). Methyl-directed repair of DNA base-pair mismatches *in vitro. Proc. Natl. Acad. Sci. USA* **80**, 4639-4643.

Lusby, E., Fife, K. H., and Berns, K. I. (1980). Nucleotide sequence of the inverted terminal repetition in adeno-associated virus DNA. *J. Virol.* **34**, 402-409.

Maizel, J. V., White, D. O., and Scharff, M. D. (1968). The polypeptides of adenovirus 2: soluble proteins, cores top components, and the structure of the virion. *Virology* **36**, 126-136.

Mak, S. (1970). Defective virions in human adenovirus type 12. *J. Virol.* **7**, 426-433.

Maniatis, T., Fritsch, D., and Sambrook, J. (1982). "Molecular cloning. A laboratory manual. " Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Matsuzaki, H., Araki, H., and Oshima, Y. (1988). Gene conversion associated with site-specific recombination in yeast plasmid pSR1. *Mol. Cell. Biol.* **8**, 955-962.

Mautner, V., and Mackay, N. (1984). Recombination in adenovirus: Analysis of crossover sites in intertypic overlap recombinants. *Virology* **139**, 43-52.

Maxam, A. M., and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.* **65**, 499-560.

Mazur, S. J., and Record, M. T., Jr. (1989). Association kinetics of site-specific protein-DNA interactions: Roles of nonspecific DNA sites and of the molecular location of the specific site. *Biopolymers* **28**, 929-953.

McGrory, W. J., Bautista, D. S., and Graham, F. L. (1988). A simple technique for the rescue of early region I mutations into infectious human adenovirus Type 5. *Virology* **163**, 614-617.

Mellor, A. L., Weiss, E. H., Ramachandran, K., and Flavell, R. A. (1983). A potential donor gene for the *bm1* gene conversion event in the C57BS mouse. *Nature* **306**, 792-795.

Michelson, A. M., and Orkin, S. H. (1983). Boundaries of gene conversion within the duplicated human α-globin genes. *J. Biol. Chem.* **258**, 15245-15254.

Miller, L. K., Cooke, B. E., and Fried M. (1976). Fate of mismatched base-pair regions in polyoma heteroduplex DNA during infection of mouse cells. *Proc. Natl. Acad. Sci. USA* **73**, 3073-3077.

Mocarski, E. S., Post L. T., and Roizman, B. (1980). Molecular engineering of the herpes simplex virus genome: insertion of a second L-S junction into the genome causes additional genome inversions. *Cell* **24**, 555-565.

Mocarski, E. S., Post, L. E., and Roizman, B. (1981). Site-specific inversion sequence of the herpes simplex virus genome: domain and structural features. *Proc. Natl. Acad. Sci. USA* **78**, 7047-7051.

Mocarski, E. S., and Roizman, B. (1982). Herpesvirus-dependent amplification and inversion of cell-associated viral thymidine kinase gene flanked by viral a sequences and linked to an origin of viral DNA replication. *Proc. Natl. Acad. Sci. USA* **79**, 5626-5630.

Müller, B., Jones, C., and West, S. C. (1990). T7 endonuclease I resolves Holliday junctions formed *in vitro* by RecA protein. *Nucl. Acids Res.* **18**, 5633-5636.

Munroe, S. H. (1983). A large inverted repeat sequence overlaps two acceptor splice sites in adenovirus. *Nucl. Acids Res.* **11**, 8891-8900.

Nagata, K., Guggenheimer, R. A., Enomoto T., Lichy, J. H., and Hurwitz J. (1982). Adenovirus DNA replication *in vitro*: Identification of a host factor that stimulates synthesis of the pre-terminal protein-dCMP complex. *Proc. Natl. Acad. Sci. USA* **79**, 6438-6442.

Nagata, K., Guggenheimer, R. A., and Hurwitz, J. (1983a). Specific binding of a cellular DNA replication protein to the origin of replication of adenovirus DNA. Proc. Natl. Acad. Sci. USA **80**, 6177-6181.

Nagata, K., Gugenheimer, R. A., and Hurwitz, J. (1983b). Adenovirus DNA replication *in vitro*: synthesis of full-length DNA with purified proteins. *Proc. Natl. Acad. Sci. USA* **80**, 4266-4270.

Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G., and Steitz, T. A. (1985). Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* **313**, 762-766.

O'Neill, E. A., Fletcher, C., Burrow, C. R., Heintz, N., Roeder, R. G., and Kelly, T. J. (1988). Transcriptional factor OTF-1 is functionally identical to the DNA replication factor NFIII. *Science* **241**, 1210-1213.

O'Neill, E. A., and Kelly, T. J. (1988). Purification and characterization of nuclear factor III (origin recognition protein C), a sequence-specific DNA binding protein required for efficient initiation of adenovirus DNA replication. *J. Biol. Chem.* **263**, 931-937.

Orr-Weaver, T. L., and Szostak, J. W. (1985). Fungal recombination. *Microbiol. Rev.* **49**, 22-58.

Ostrove, J. M., Rosenfeld, P., Williams, J., and Kelly, T. J., Jr. (1983). *In vitro* complementation as an assay for purification of adenovirus DNA replication proteins. *Proc. Natl. Acad. Sci. USA* **80**, 935-939.

Patel, D. J., Kozlowski, S. A., Marky, L. A., Rice, J. A., Broka, C., Dallas, J., Itakura, K., and Breslauer, K. (1982). Structure, dynamics, and energetics of deoxyguanosine-thymidine wobble base pair formation in the self-complementary d(CGTGAATTCGCG) duplex in solution. *Biochemistry* **21**, 437-444.

Patel, S. S., Wong, I., and Johnson, K. A. (1991). Pre-steady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant. *Biochemistry* **30**, 511-525.

Pearson, G. D. (1975). Intermediate in adenovirus type 2 replication. *J. Virol.* **16**, 17-26.

Pearson, G. D., and Hanawalt P. C. (1971). Isolation of DNA replication complexes from uninfected and adenovirus-infected HeLa cells. *J. Mol. Biol.* **62**, 65-80.

Peden, K. S. C., Pipas, J. M., Pearson-White, S., and Nathans, D. (1980). Isolation of mutants of an animal virus in bacteria. *Science* **209**, 1392-2986.

Pettersson, U. (1973). Some unusual properties of replicating adenovirus DNA. *J. Mol. Biol.* **81**, 521-527.

Pincus, S., Robertson, W., and Rekosh, D. M. K. (1981). Characterization of the effect of amphidicolin on adenovirus DNA replication: evidence in support of a protein primer model of initiation. *Nucl. Acids Res.* **9**, 4919-4938.

Prage, L. S., Hoglund, S., and Philipson, L. (1972). Structural proteins of adenoviruses. III. Charaterization of incomplete particles of adenovirus type 3. *Virology* **49**, 745-757.

Prelich, G., and Stillman, B. W. (1986). Functional characterization of thermolabile DNA-binding proteins that affect adenovirus DNA replication. *J. Virol.* **57**, 883-892.

Pribnow, D., Sigurdson, D. C., Gold, L., Singer, B. S., Napoli, C., Brosius, J., Dull, T. J., and Noller, F. H. F. (1981). rII cistrons of bacteriophage T4. DNA sequence around the intercistronic divide and positions of genetic landmarks. *J. Mol. Biol.* **149**, 337-376.

Pruijn, G. J. M., van Driel, S., and van der Vliet, P. C. (1986). Nuclear factor III, a novel sequence-specific DNA-binding protein from HeLa cells stimulating adenovirus DNA replication. *Nature* **322**, 656-659.

Ptashne, M. (1986). Gene regulation by proteins acting nearby and at a distance. *Nature* **322**, 697-701.

Radman, M., and Wagner, R. (1986). Mismatch repair in *Escherichia coli*. *Annu. Rev. Genet.* **20**, 523-538.

Rainbow, A., and Mak, S. (1970). Functional heterogeneity of virions in human adenovirus types 2 and 12. J. *Virology* **5**, 188-193.

Reed, K. G., and Mann, D. A. (1985). Rapid transfer of DNA from agarose gels to nylon membranes. *Nucleic Acids Res.* **13**, 7207-7221.

Rekosh, D.M.K., Russell, W. C., Bellett, A.J.D., and Robinson, A.J. (1977). Identification of a protein linked to the ends of adenovirus DNA. *Cell* **11**, 283-295.

Ripley, L. S., Clark, A., and de Boer, J. G. (1986). Spectrum of spontaneous frameshift mutations sequences of bacteriophage T4 rII gene frameshifts. *J. Mol. Biol.* **191**, 601-613.

Ripley, L. S., and Shoemaker, N. B. (1983). A major role for bacteriophage T4 DNA polymerase in frameshift mutagenesis. *Genetics* **103**, 353-366.

Robin, J., Bourgaux-Ramoisy, D., and Bourgaux, P. (1973). Single-stranded regions in replicating DNA of adenovirus type 2. *J.Gen. Virol.* **20**, 233-237.

Robinson, A. J., and Bellett, A. J. D. (1974). A circular DNA-protein complex from adenoviruses and its possible roll in DNA replication. *Cold Spring Harbor Symp. Quant. Biol.* **39**, 523-531.

Robinson, A. J., Younghusband, H. B., and Bellett, A. J. D. (1973). A circular DNA-protein complex from adenoviruses. *Virology* **56**, 54-69.

Roizman, B. (1979). The structure and isomerization of herpes simplex virus genomes. *Cell* **16**, 481-494.

Roizman, B., Hayward, G., Jacob, R., Wadsworth, S. W., and Honess, R. W. (1974). Human herpesvirus 1: a model for molecular organization of herpesvirus virions and their DNA. Proceedings of the XIth International Congress on Cancer, Excerpta Mecica, International Congress Series No. 350, Florence. (In) *Chemical and Viral Carcinogenesis* **2**, 188-198.

Roizman, B., Jacob, R. J., Knipe, D. M., Morse, L. S., and Rueychan, W. T. (1979). On the structure, functional equivalence, and replication of the four arrangements of herpes simplex virus DNA. *Cold Spring Harbor Symp. Quant. Biol.* **43**, 808-826.

Rosenfeld, P. J., and Kelly, T. J. (1986). Purification of nuclear factor 1 by DNA recognition site affinity chromatography. *J. Biol. Chem.* **261**, 1398-1408.

Rosenfeld, P.J., O'Neill, E. A., Wides, R. J., and Kelly T. J. (1987). Sequence-specific interactions between cellular DNA-binding proteins and the adenovirus origin of DNA replication. *Mol. Cell Biol.* **7**, 875-886.

Rosenwirth, B., Tjia, S., Westphal, M., and Doerfler, W. (1974).Incomplete particles of adenovirus. II. Kinetics of formation and polypeptide composition of adenovirus type 2. *Virology* **60**, 431-437.
Roth, J. R. (1974). Frameshift mutations. *Annu. Rev. Genet.* **8**, 319-346.

Roy, S., Sklenar, V., Appella, E., and Cohen, J. S. (1987). Conformational perturbation due to an extra adenosine in a self-complementary oligodeoxynucleotide duplex. *Biopolymers* **26**, 2041-2052.

Samulski, R. J., Srivastava, A., Berns, K. I., and Muzyizka N. (1983). Rescue of adeno-associated virus from recombinant plasmids: Gene correction within the terminal repeats of AAV. *Cell* **33**, 135-143.

Schechter, W. M., Davies, W., and Anderson, C. W. (1980). Adenovirus coded deoxyribonucleic acid binding protein. Isolation, physical properties, and effects of proteolytic digestion. *Biochemistry* **19**, 2802-2810.

Sheldrick, P., and Berthelot, N. (1975). Inverted repetitions in the chromosome of herpes simplex virus. *Cold Spring Harbor Symp. Quant. Biol.* **39**, 667-678.

Shinagawa, M., Ishiyama, T., Padmanabhan, R., K., Kamada, M., and Sato, G. (1983). Comparative sequence analysis or the inverted terminal repetition in the genomes of animal and avian adenoviruses. *Virology* **124**, 491-495.

Shinagawa, M., and Padmanabhan, R. (1980). Comparative sequence analysis of the inverted terminal repetitions from different adenoviruses. *Proc. Natl. Acad. Sci. USA* **77**, 3831-3835.

Slightom, J. L., Blechl, A. E., and Smithies, O. (1980). Human fetal $^{G}\gamma$- and $^{A}\gamma$-globin genes: complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. *Cell* **21**, 627-638.

Smart, J. E., and Stillman, B. W. (1982). Adenovirus terminal protein precursor: Partial amino acid sequence and site of covalent linkage to virus DNA. *J. Biol. Chem.* **257**, 13499-13560.

Smiley, J., Frong, B., and Leung, S. C. (1981). Construction of double jointed herpes simplex virus DNA molecule: inverted repeats are required for segment inversion and direct repeats promote deletion. *Virology* **111**: 345-362.

Steenbergh, P. H., Maat, J., van Ormondt, H., and Sussenbach, J. S. (1977). The nucleotide sequence at the termini of adenovirus type 5 DNA. *Nucl. Acids Res.* **4**, 4371-4389.

Stewart, J. W., and Sherman, F. (1974). Yeast frameshift mutations identified by sequence changes in iso-1-cytochrome c. Molecular and environmental aspects of mutagenesis (Prakash, L., Sherman, F., Miller M.W., Lawrence, C.W., Taber, H.W., eds) Charles C. Thomas, Springfield, IL, 102-127.

Stillman, B. W. (1981). Adenovirus DNA-replication *in vitro*: protein linked to the 5' end of nascent DNA strands. *J. Virol.* **37**, 139-147.

Stillman, B. W. (1989). Initiation of eukaryotic DNA replication *in vitro*. *Annu. Rev. Cell Biol.* **5**, 197-245.

Stillman, B. W., Tamanoi, F., and Mathews, M. B. (1982b). Purification of an adenovirus-coded DNA polymerase that is required for initiation of DNA replication. *Cell* **31**, 613-623.

Stillman, B. W., Topp, W. C., and Engler, J. A. (1982a). Conserved sequences at the origin of adenovirus DNA replication. *J. Virol.* **44**, 530-537.

Stoeckert, C.J., Jr., Collins, F. S., and Weissman, S. M. (1984). Human fetal globin DNA sequences suggest novel conversion event. *Nucl. Acids Res.* **12**, 4469-4479.

Stow, N. D. (1982). The infectivity of adenovirus genomes lacking DNA sequences from their left-hand termini. *Nucleic Acids Res.* **10**, 5105-5119.

Straus, S. E., Sebring, E. D., and Rose, J. A. (1976). Concatemers of alternating plus and minus strands are intermediates in adenovirus-associated virus DNA synthesis. *Proc. Natl. Acad. Sci. USA* **73**, 742-746.

Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E., and Inouye, M. (1966). Frameshift mutations and the genetic code. *Cold Spring Harbor Symp. Quant. Biol* **31**, 77-84.

Streisinger, G., and Owen, J. E. (1985). Mechanisms of spontaneous and induced frameshift mutation in bacteriophage T4. *Genetics* **109**, 633-659.

Su, S.-S., Lahue, R. S., Au, K. G., Modrich, P. (1988). Mispair specificity of methyl-directed DNA mismatch correction *in vitro*. *J. Biol. Chem.* **263**, 6829-6835.

Subramani, S., and Berg, P. (1983). Homologous and nonhomologous recombination in monkey cells. *Mol. Cell. Biol.* **3**, 1040-1052.

Suggs, S. V., Hirose, T., Miyake, T., Kawashima, E. H., Jonson M. J., Itakura, K., and Wallace, R. B. (1981). *ICN-UCLA symp. Mol. Cell. Biol.* **231**, 683-691.

Sussenbach, J. S. (1984). The structure of the genome. In "The Adenoviruses" (H.S. Ginsberg, Ed.). Plenum, New York, 271-308.

Sussenbach, J. S., Ellens, D. J., and Jansz, J. S. (1973). Studies on the mechanism of replication of adenovirus DNA II: the nature of single-stranded DNA in replicative intermediates. *J. Virol.* **23**, 1131-1138.

Sussenbach, J. S., van der Vliet, P. C., Ellens, D. J., and Jansz, H. S. (1972). Linear intermediates in the replication of adenovirus DNA. *Nature New Biol.* **239**, 47-49.

Tamanoi, F., and Stillman, B. W. (1982). Function of the adenovirus terminal protein in the initiation of DNA replication. *Proc. Natl. Acad. Sci. USA* **79**, 2221-2225.

Tolun, A., Aleström, P., Pettersson, U. (1979). Sequence of inverted terminal repetitions from different adenoviruses: demonstration of

conserved sequences and homology between SA7 termini and SV40 DNA. *Cell* **17**, 705-713.

van Amerongen, H., van Grondelle, R., and van der Vliet, P. C. (1987). Interaction between adenovirus DNA-binding protein and single-stranded polynucleotides studied by circular dichroism and ultraviolet absorption. *Biochemistry* **26**, 4646-4652.

van Bergen, B. G. M., van der Ley, P. A., van Priel, W., van Mansfield, A. O. M., and van der Vliet, P. C. (1983). Replication of origin containing adenovirus DNA fragments that do not carry the terminal protein. *Nucl. Acids Res.* **11**, 1975-1989.

van Bergen, B. G. M., and van der Vliet, P. C. (1983). Temperature sensitive initiation and elongation of adenovirus DNA replication *in vitro* with nuclear extracts from H5*ts*36, H5*ts*149, and H5*ts*125 infected HeLa cells. *J. Virol.* **42**, 642-648.

van der Eb, A. J. (1973). Intermediates in type 5 adenovirus DNA replication. *Virology* **51**, 11-23.

van der Vliet, P. C., Keegstra, S., and Jansz, J. S. (1978). Complex formation between the adenovirus type 5 DNA binding protein and single-stranded DNA. *Eur. J. Biochem.* **86**, 389-398.

van der Vliet, P. C., and Sussenbach, J. S. (1972). The mechanism of adenovirus DNA synthesis in isolated nuclei. *Eur. J. Biochem.* **30**, 584-592.

van der Vliet, P. C., and Sussenbach, J. S. (1975). An adenovirus gene function required for the initiation of viral DNA replication. *Virology* **67**, 415-427.

van Ormondt, H., and Galibert, F. (1984). Nucleotide sequences of adenovirus DNAs. *Curr. Topics Microb. Immunology* **110**, 73-142.

Varlet, I., Radman, M., and Brooks, P. (1990). DNA mismatch repair in *Xenopus* egg extracts: Repair efficiency and DNA repair synthesis for all single base-pair mismatches. *Proc. Natl. Acad. Sci. USA* **87**, 7883-7887.

Volkert, F. C., and Young, D. S. H. (1983). The genetic analysis of recombination using adenovirus overlapping terminal DNA fragments. *Virology* **125**, 175-193.

Wadell, G., Hammarskjold, M. L., and Varsanyi, T. (1973). Incomplete virus particles of adenovirus type 16. *J. Gen. Virol.* **20**, 287-302.

Wadsworth, S., Jacob, R. J., and Roizman, B. (1975). Anatomy of herpes virus DNA. II. Size, composition, and arrangement of inverted terminal repetitions. *J. Virol.* **15**, 1487-1497.

Wake, C. T., and Wilson, J. H. (1980). Defined oligomeric SV40 DNA: a sensitive probe of general recombination in somatic cells. *Cell* **21**, 141-148.

Wang, K., and Pearson, G. D. (1985). Adenovirus sequences required for replication *in vivo*. *Nucl. Acids Res.* **13**, 5173-5187.

Wang, K., Xu, F.-Y., Ahern, K. G., Pearson, G. D. (1991). Inverted repeats direct repair of adenovirus minichromosome ends. *Virology* **183**, in press.

Weingärtner, B., Winnacker, E.-L., Tolun, A., and Pettersson, U. (1976). Two complimentary strand-specific termination sites for adenovirus DNA replication. *Cell* **9**, 259-268.

Werntges, J. Gerhard, S., Riesner, D., and Fritz, H.-J. (1986). Mismatches in DNA double strands: thermodynamic parameters and their correlation to repair efficiencies. *Nucl. Acids Res.* **14**, 3773-3790.

Whitehouse, J. L. D. (1963). A theory of crossing-over by means of hybrid deoxyribonucleic acid. *Nature* **199**, 1034-1040.

Wiebauer, K., and Jiricny, J. (1989). *In vitro* correction of G-T mispairs to G-C pairs in nuclear extracts from human cells. *Nature* **339**, 234-236.

Wilkie, N. M. (1973). The synthesis and substructure of herpesvirus DNA: the distribution of alkali-labile single strand interruptions in HSV-1 DNA. *J. Gen. Virol.* **21**, 453-467.

Wong, I., Patel, S. S., and Johnson, K. A. (1991). An induced-fit kinetic mechanism for DNA replication fidelity: Direct measurement by single-turnover kinetics. *Biochemistry* **30**, 526-537.

Zeff, R. A., and Geliebter, J. (1987). Oligonucleotide probes for genomic DNA blots. *Focus* **9**, 1-2.