AN ABSTRACT OF THE THESIS OF

Kevin Ahern for the degree of Doctor of Philosophy in

Biochemistry and Biophysics presented on May 1, 1986.

Title: ADENOVIRUS REPLICATION IN TRANS: A NEW REPLICATION PATHWAY

Redacted for privacy

Abstract approved: _____
Dr. George D. Pearson

Replication of plasmid molecules containing either a left-end or a right-end adenovirus origin was investigated. When the plasmids were linearized to expose their respective adenovirus termini, mixed, and transfected with helper adenovirus DNA into the human 293 cell line, a new molecule was detected. The new molecule was the size predicted for a recombinant between the two input plasmids. The recombinational event, however, was not due to simple homologous recombination and was totally dependent upon adenovirus replication. No molecules of the same size were detected in the absence of helper adenovirus DNA, indicating that detection of the new molecule was enabled by a function provided by the adenovirus DNA, most likely replication. The molecules appeared to be replicating by an adenovirus-driven mechanism as indicated by their covalent attachment to protein and their accumulation from 35-72 hours in a time course experiment. The available evidence indicates that the molecules were produced by replication followed by recombination rather than vice-versa.

A mechanism, called *trans* replication, is postulated whereby the displaced complementary strands produced during replication interact via base pairing to form a duplex. Kinetic data of adenovirus replication previously obtained by other workers is consistent with the process functioning in the normal adenovirus replication cycle. Such a process is a hitherto unknown means for adenovirus to finish complementary replication. As such, it represents the first clear violation of the Meselson-Stahl principle of semi-conservative replication that has been observed in any biological system.

Adenovirus Replication in Trans: A New Replication Pathway

by

Kevin Ahern

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Completed May 1, 1986

Commencement June 1986

APPROVED:

Redacted for privacy

Professor of Biochemistry and Biophysics in charge of major

Redacted for privacy

Chairman of Department of Biochemistry and Biophysics

Redacted for privacy

Dean of Graduate School

Date thesis is presented May 1, 1986

Typed by Barbara Hanson for Kevin Ahern

ACKNOWLEDGMENTS

I would like to acknowledge and thank my colleagues who contributed to this work. In my laboratory they are Kai Wang, Kate Mathews, and Fu Yun Xu who contributed plasmid constructs, buffers, and help whenever needed, Jerry Myers who was of great assistance with the electron microscopic studies, and Dr. Bob Enns who helped in orienting me to the electron microscope in the beginning.

Gael Kurath in Dr. Joane Leong's laboratory at Oregon State University contributed the nucleic acids for the R-loop mapping study of IHNV. David Irwin in Dr. Ross MacGillivray's laboratory at the University of British Columbia contributed the DNAs used in the prothrombin work. Dr. C.A. Ryan at Washington State University contributed the DNAs used in the Proteinase Inhibitor I work. Robert Durst co-wrote the ASAP computer program for DNA sequence searches and developed the extremely fast binary arithmetic that enabled its speed of operation. Thanks also to Greg Bertoni in Dr. Dallice Mills' laboratory at Oregon State University and Dr. Kathy Berkner at Zymogenetics for contributing plasmid constructs used in this work.

The support and efforts of my major professor, Dr. George Pearson, throughout this work were greatly appreciated. In addition, thanks are in order for the time and efforts of my other committee members, Dr. Ken van Holde, Dr. Chris Mathews, Dr. Dallice Mills, and Dr. Gene Craven. Thanks also to Barbara Hanson for expert typing of this dissertation. This work was supported by grants from the National Cancer Institute (CA 17699) and a Predoctoral Training Grant in Molecular and Cellular Biology by the National Institute of General

and Medical Sciences (5T 32 GM0774-05).

I also would like to acknowledge the moral support provided by all my family members, especially my wife Julie, whose adventures to Antarctica I could never imagine duplicating and my mother Shirley, whose endless source of energy is envied.

Finally, I dedicate this work to my grandfather Albert Ralph O'Harrow, of Fowler, Illinois, who died of cancer in November of 1985. Ab, as he was called, was the most fun person I have ever met. My memories of him are filled with wonderful vignettes of his life: Playing baseball and being mistaken for his brother by people wanting to hire a pitcher; watching Jackie Gleason and screaming "How sweet it is!"; his endless comic rendition of the song "Don't Let the Rains Come Down"; his love of the card game "Pitch" and his tendency to slam down the cards and exclaim "Good God Evelyn!" when his wife bid too low for his tastes; his ability to be able to laugh at almost anything and to infectioiusly spread that laughter to those around him; and finally the graceful manner in which he died. Up until the very last day he never lost his dignity nor his sense of humor. In memory of his life this work is dedicated.

TABLE OF CONTENTS

		Page
1.	INTRODUCTION	1
	Adenovirus Replication	2
	In vivo Studies	3
	Electron Microscopic Analyses	7
	In vitro Adenovirus Replication	11
	The Proteins in Adenovirus Replication	12
	Precursor Terminal Protein (pTP)	12
	Adenovirus DNA Polymerase	15
	Adenovirus DNA Binding Protein	16
	Nuclear Factor I	17
	Nuclear Factor II	19
	Sequence Dependence of Adenovirus Replication	19
	Other Sequences in the ITR and their Role in Replication	24
	Unexplained Aspects of Adenovirus Replication	25
	Complementary Strand Synthesis	27
2.	MATERIALS AND METHODS	31
	Reagents/Enzymes	31
	Media	32
	DNA Isolation	33
	Preparation of Blunt-Ended DNA	34
	Nick Translation of DNAs	34
	Spin-Column Purification of DNA	35
	Preparation of Competent *E. coli* Cells	35
	Preparation of DNA for Bacterial Transformation	36
	Transformation of Bacterial Cells	37
	Growth and Purification of Adenovirus DNA	37
	Transfection of 293 Cells	38
	Gel Preparation	39
	Southern Blots	39
	DNAs used in Construction of Clones	40
	Construction of Clones	41
3.	A NOVEL RECOMBINATIONAL ACTIVITY IN ADENOVIRUS INFECTED CELLS	51
	Strategy	53
	Detection of *Trans* Replication	57
	Repair of Origin Sequences	64
	Assay of Homologous Recombination by *Cis* Repair	64

4. CONCLUSIONS 71

Summary of Results 71
Possible Interpretations 72
Known Recombination Mechanisms of Adenovirus 73
Models for Recombination 75
Interpretation of Observed Phenomena 78
Postulated Mechanism of *Trans* Replication 81
A New Mode of Replication 82
Intracellular Complementary Strand Pairing 85
Relationship to Previous Studies 87
Kinetic Studies 88
Implications of *Intermolecular* Pairing 90
Previous Work 90
Homologous Recombination 91
Trans Replication and ITR Size 91
Trans Replication and ϕ29 92
Summary 93

5. BIBLIOGRAPHY 94

AI. R-LOOP MAPPING THE INFECTIOUS HEMATOPOIETIC NECROSIS VIRUS 113

Introduction 114
Materials and Methods 114
Results 116
Discussion 121
References 122

AII. HETERODUPLEX CHARACTERIZATION OF THE PROTEASE INHIBITOR I GENE OF POTATO AND TOMATO 124

Introduction 125
Materials and Methods 125
Results and Discussion 126
References 128

AIII. CHARACTERIZATION OF THE BOVINE PROTHROMBIN GENE BY ELECTRON MICROSCOPY 129

Introduction 130
Materials and Methods 131
Results 132
Discussion 138
References 141

AIV. A FAST MODULAR DNA SEARCH PROGRAM FOR ANALYSIS OF LONG DNA SEQUENCES ON APPLE II COMPUTERS 143

Abstract 144
Introduction 144

Hardware Requirements 145
Software Description 146
Program Operation 146
Conversion of GenBank Files 149
Discussion 151
References 151

AV. CONSTRUCTION OF CLONES 153

Introduction 154
References 187

LIST OF FIGURES

		Page
Chapter 1		
Figure 1	Summary of the scheme for adenovirus replication	8
Chapter 2		
Figure 2	Construction of p44x	42
Figure 3	Construction of pKT14	43
Figure 4	Construction of pKAT21	44
Figure 5	Construction of pD25	45
Figure 6	Construction of pRS17	47
Figure 7	Construction of pBP34	48
Figure 8	Construction of clone 342	49
Figure 9	Construction of pADIRK2	50
Chapter 3		
Figure 10	Summary of results with EcoRI-cut clone 7 and EcoRI-cut pXD6 in the presence of helper Ad2 DNA	54
Figure 11	Southern blot analysis for trans replication	55
Figure 12	Trans replication assayed from 10 to 72 hours as indicated at the top of the figure	58
Figure 13	Summary of the results with EcoRI-cut clone 7 (exposed origin) and PstI-cut pXD6 (embedded origin) in the presence of helper Ad2 DNA	60
Figure 14	Summary of results with BamHI-cut clone 7 (embedded origin) and EcoRI-cut pXD6 (exposed origin) in the presence of helper Ad2 DNA	61

Figure 15 Postulated homologous recombinational intermediates when EcoRI-cut clone 7 and EcoRI-cut pXD6 (A), BamHI-cut clone 7 and EcoRI-cut pXD6 (B), or EcoRI-cut clone 7 and PstI-cut pXD6 (C) were transfected with helper Ad2 DNA 63

Figure 16 Generation of a 6.7 kb molecule from an 8.0 bk molecule with two embedded origins 65

Figure 17 Postulated recombinational pathways for synthesis of a cis-repairable adenovirus origin when EcoRI-cut clone 7 and PvuII-cut pRS17 were transfected with helper Ad2 DNA 67

Figure 18 Postulated products of homologous recombination between EcoRI-cut clone 7 and EcoRI-cut clone 342 68

Figure 19 Results of trans replication with molecules marked with an insertion and a deletion 70

Chapter 4

Figure 20 The Meselson-Radding model for recombination 76

Figure 21 Double-strand-break repair model of recombination 77

Figure 22 Lin model for homologous recombination 79

Figure 23 Postulated mechanism of trans replication involving pairing of displaced strand, repair by DNA polymerase, and amplification by replication 83

Figure 24 Trans replication in the adenovirus replication cycle 84

Figure 25 Pairing of displaced adenovirus strands through complementary sequences 86

Appendix AI

Figure 26 Analysis by electron microscopy of R-loops formed between IHNV genomic RNA and plasmids containing cloned IHNV cDNA sequences 117

Figure 27 Analysis by electron microscopy of R-loops formed between IHNV genomic RNA and pairs of plasmids containing cloned IHNV cDNA sequences 118

Figure 28 Physical map of the IHNV genome 123

Appendix AIII

Figure 29 Analysis by electron microscopy of heteroduplexes formed between cloned bovine genomic DNA (λBII3) and cloned prothrombin cDNA (pBII111) 134

Appendix AV

Figure 30 Construction of clone 11 155

Figure 31 Construction of clone 11y 157

Figure 32 Construction of clone 17 159

Figure 33 Construction of clone 4 161

Figure 34 Construction of p7x 163

Figure 35 Construction of pIR12 165

Figure 36 Construction of clones pAT(7-67) 167

Figure 37 Construction of pRL102 169

Figure 38 Construction of pRLKAN 171

Figure 39 Construction of pRL12 173

Figure 40 Construction of pRBAM 175

Figure 41 Construction of clones pRL(x) 177

Figure 42 Construction of pBN7 179

Figure 43 Construction of pSP38 181

Figure 44 Construction of pADSMA14 183

Figure 45 Construction of pADSPH61 185

LIST OF TABLES

Page

Appendix AI

Table 1 Measurements of single R-loops and location of hybridizations on the viral genome 119

Table 2 Gap measurements for double R-loops 120

Appendix AII

Table 3 Comparison of gene maps of the tomato and potato proteinase inhibitor I genes generated by heteroduplex analysis and DNA sequencing 127

Appendix AIII

Table 4 Subclones of cloned phage DNA used in the heteroduplex analysis 133

Table 5 Lengths of exons and introns in the prothrombin gene 136

Table 6 Location of restriction sites within the prothrombin gene 137

Table 7 Lengths and locations of inverted repeat sequences within introns 139

ADENOVIRUS REPLICATION *IN TRANS*: A NEW REPLICATION PATHWAY

CHAPTER 1: INTRODUCTION

The adenovirus system is one of the most thoroughly studied eukaryotic viral replication systems. Adenovirus was first described by Rowe *et al.* (1) and later by Hilleman and Werner (2). It is a class of ubiquitous mammalian viruses that have been isolated from cattle, sheep, pigs, monkeys, tree shrews, mice, horses, dogs, and even birds (see 3 for reviews). At least 42 distinct serotypes have been identified among the human population and these have been placed into 6 subgroups using a variety of criteria including hemagglutination properties (4), GC content (5), DNA homology (6), restriction maps (7), and oncogenicity (8). The genome of the human serotype designated adenovirus 2 (Ad2) has been completely sequenced and it contains 35,937 base pairs. The adenovirus system has been widely used as a model for the study of oncogenic transformation, gene expression, and DNA replication.

Adenovirus 12 was the first human virus shown to be oncogenic when tumors developed following injection of the virus into newborn hamsters (9). Various human adenovirus serotypes appear to have differing potential for transformation in rodents, ranging from the non-oncogenic types 1,2,4,5,6,8,9,10,13,15,17,19,20 and 22-30 and weakly oncogenic types 3,7,11,14,16 and 21 to the highly oncogenic types 12, 18, and 31 (8,10,11). Although the reasons for the differing capacities for oncogenicity between the serotypes are not known, the region responsible for transformation has been localized,

initially by Graham et al. (12), to the E1 region of the virus. Analysis at the DNA sequence level is providing information about the multiple E1 proteins and their role in transformation (13, see 14 for reviews).

Research on adenoviruses spans the entire realm of eukaryotic gene expression. Electron microscopic examinations of heteroduplexes between adenovirus mRNAs and adenovirus DNA revealed the phenomenon of splicing for the first time (15-17). Splicing of adenovirus mRNAs is being investigated for insight into the process itself (18,19) and as a control mechanism (20). Adenovirus has also provided a wealth of information about the control of eukaryotic transcription. Expression of the adenovirus E1a region has trans-acting effects on the transcription of a wide variety of genes, including other adenovirus genes (21-25), SV40 genes (26), and cellular genes (27). For a review of adenovirus transcription see reference 28. Adenovirus also has effects on the cellular translation process. These are exerted as positive controls upon viral RNA sequences (29) and negative effects upon at least some cellular sequences (30,31). Two viral RNAs play a role in at least part of the process at the level of initiation of protein synthesis (32-35). Other aspects of adenovirus gene expression under study include polyadenylation (36-38), recombination (39-45), the role and mechanism of enhancers (46,47), and virus structure (48-50).

ADENOVIRUS REPLICATION

This study concerns the adenovirus replication process.

Adenovirus has an interesting molecular structure that makes it an attractive model system for replication processes. All adenoviruses examined to date have two distinctive features, an inverted terminal repeat (ITR) at each end of the linear virus DNA (51,52) and a protein designated 'Terminal Protein' (TP) covalently attached to the 5' end of each strand (53-55). ITR sizes range from 63 bases for the avian CELO virus (56) to 166 bases for tree shrew adenovirus (57). The ITR is thought to play at least two roles in replication. Sequences contained within the ITR are required for initiation of replication (58-61). In addition, the ITR is widely believed to play a role in the replication of the complementary strand, probably through formation of a panhandle-like structure (62). For other reviews of the process of adenovirus replication see Winnacker (63), Horwitz et al. (64), Kelly (65), Challberg and Kelly (66), Stillman (67), Fütterer and Winnacker (68), Tamanoi and Stillman (69), Sussenbach and van der Vliet (70), and Friefeld et al. (71).

IN VIVO STUDIES

Early work on adenovirus replication was done primarily in infected cells (in vivo), and it established many of the unique features of the viral replicative cycle. Adenovirus infection is initiated when virion fibers bind specific cellular receptors (72). After entering the cell, viral capsids are removed mainly in the cytoplasm (72) and virion cores enter the nucleus (73) where they are uncoated, resulting in essentially virion-free genomes (74). Replication is temporally regulated, first being detected at 6-8 hours

post-infection, peaking at about 18-20 hours (75).

Adenovirus transcription can be roughly divided into an early (0-8 hours) and a late phase (8-36 hours) (for a review see 28). Although the regulation mechanisms for the control of transcription are not totally established, it appears that at least the viral E1a and adenovirus DNA binding protein (AdDBP) genes play roles in the process (22-25,47,76, 77). The E1a transcription unit codes for at least three mRNAs whose expression can be detected within the first hour of infection (78). Viruses defective in the E1a genes fail to transcribe the E1b, E2a, E2b, E3, E4, and L1 regions of the genome (22,79). In an in vitro system AdDBP repressed transcription from the E4 transcriptional unit (80). Incubation of single stranded DNA in this system relieved the repression, probably through competition for AdDBP. Other experiments indicate that AdDBP may regulate late gene expression. Rice and Klessig (81) have proposed that AdDBP regulates its own transcription, but does not affect E1a, E1b, E3, or E4 mRNA levels in infected cells.

The three known adenovirus coded proteins required for replication are contained within the E2a and E2b transcription units (see below). The E2A promoter controlling transcription of the AdDBP also controls E2B transcripts (82) and may itself be regulated by E1a gene products in a sequence-independent manner (76,83,84). Differential expression of the genes is probably controlled at least partly at the level of splicing (82), a feature of other adenovirus transcriptional units (20,85). Although all of the ways by which adenovirus controls gene expression are not yet known, the end result clearly is effective. Up to 500,000 copies of Ad2 per cell are produced by late in infection

(86) although only about 20% are ultimately incorporated into viral particles (75). Meanwhile, during the adenovirus infection, cellular DNA replication is inhibited beginning at about 6-8 hours post-infection and inhibition is essentially complete by 12 hours (87, 88).

Initial characterizations of the replication process utilized pulse labelling techniques in adenovirus-infected cells (89,90) and nuclei from infected cells (91-93). These results suggested that adenovirus replication was semi-conservative in nature (89-91). Careful analysis of ^{3}H-labelled DNA from these cells revealed that replicative intermediates, which could be chased into mature DNA, had a much greater buoyant density than mature viral DNA (89,94-97). The density difference disappeared upon treatment with the single-strand specific nuclease S1, but not RNase, indicating the presence of single-stranded DNA in the replicative intermediates (95, 96). Sedimentation analysis in alkaline sucrose gradients revealed that the length of newly synthesized strands never exceeded the length of the linear virus DNA (90,95), suggesting that replication did not occur by a rolling-circle process. In addition, mapping with pulse-labelling techniques established that both initiation (98) and termination (98-100) sites existed at both ends of the linear DNA molecule.

The AdDBP gene was the first adenovirus replication gene to be identified. It was initially characterized in studies of H5ts125, a temperature-sensitive mutant of Ad5 that is defective for viral DNA replication (101, 102). The 72 kilodalton (kDal) AdDBP is first detected about 4 hours post-infection and is synthesized continuously into late infection (30,103,104). Newly synthesized AdDBP is found in

the nucleus of infected cells and is the predominant early viral protein, present at about $5x10^6$ molecules per cell (105). A variety of mutant forms of AdDBP are available, and they point to functions in both replication and regulation of gene expression (25,106). AdDBP is analogous to the SV40 T-antigen in this dual role (107). AdDBP binds the displaced single strand of DNA generated by adenovirus replication (108,109) and functions in elongation of DNA replication (106,110,111). AdDBP is analogous to the bacteriophage T4 gene 32 single-strand-binding protein in its role in replication and its ability to stimulate denaturation of double-stranded DNA to form single-stranded DNA coated with protein (112). It is not known, however, whether AdDBP is like gene 32 in stimulating the reverse reaction, causing pairing of single strands into a duplex form (112).

The entire AdDBP gene has been sequenced (113). The protein has been purified to homogeneity by DEAE-cellulose chromatography, selective precipitation and gel filtration (114,115). The molecular weight of the protein predicted by the coding sequences is 59 kDal compared to estimates of 72 kDal by polyacrylamide SDS gel electrophoresis, a discrepancy that may be due to the presence of a large number of proline residues located in the amino end of the protein (113). It is phosphorylated and has a high charge heterogeneity due to differential phosphorylation (116,117). The role of phosphorylation in the function of the protein is not apparent, however, as it seems to have little effect on the ability of AdDBP to bind DNA (117).

ELECTRON MICROSCOPIC ANALYSES

The earliest electron microscopic studies of adenovirus molecules revealed that denatured strands of adenovirus could form a panhandle structure that was sensitive to exonuclease III treatment (51,52), demonstrating for the first time the presence of the adenovirus ITR. The presence of a protein on the viral terminus was also first detected by electron microscopic analysis of virion DNA isolated by a process that did not involve protease treatment or phenol extraction (53,54). Adenovirus molecules isolated in this manner appeared circular but could be linearized by treatment with protease (53,54).

Electron microscopic studies played a major role in focusing the individual data from in vivo work into the now established mechanism of adenovirus replication. Replicating adenovirus 5 (Ad5) molecules appeared clearly in the electron microscope as Y-shaped molecules with two double-stranded arms and a single-stranded arm that appeared to have originated from one of the ends of the virus (94,118). The most influential electron microscopic study of replicating Ad2 molecules was made by Lechner and Kelly (62, see Figure 1). Careful analysis of replicating molecules revealed Y-shaped molecules (designated as Type I molecules), partly single-stranded/partly double-stranded molecules (called Type II molecules), and rare composite molecules with both characteristics (named Type I/Type II molecules). They presented a model for adenovirus replication which proposed that the Type I and Type II molecules represented intermediates in DNA synthesis (Figure 1).

They proposed that Type I molecules arose from initiation events

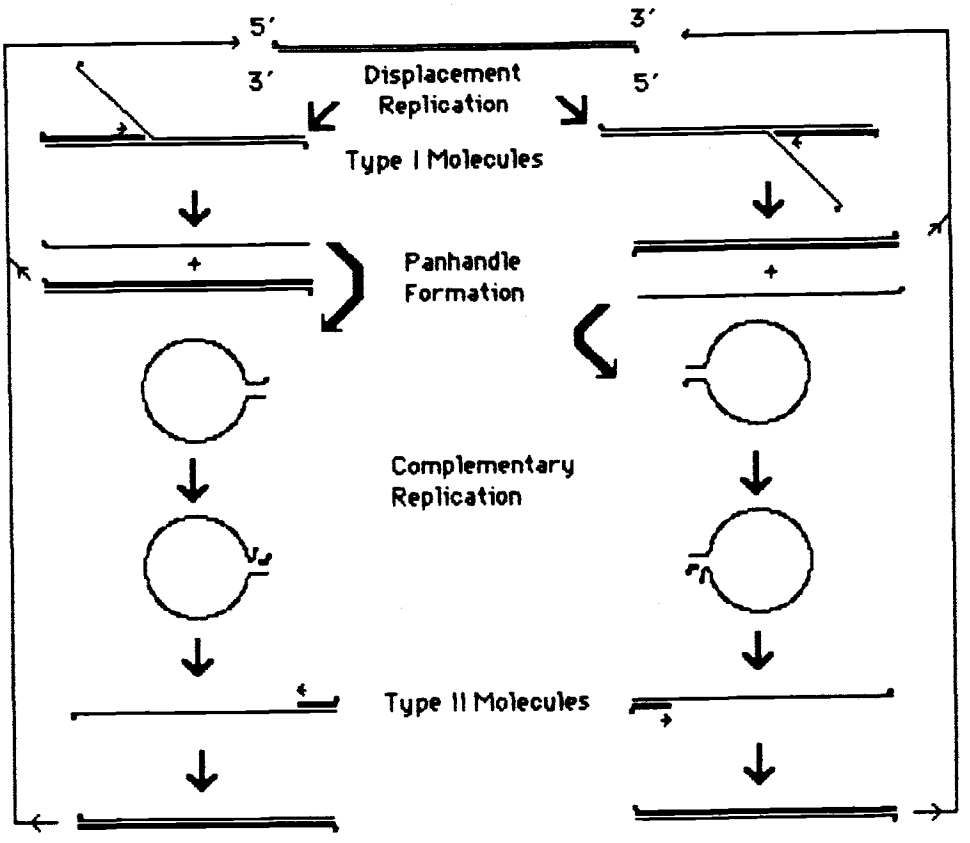


Figure 1: Summary of the scheme for adenovirus replication. See text for details. Adapted from Lechner and Kelly (1977).

near the 3' end of one strand with replication proceeding to displace the 5' end of the complementary strand. The predicted products of this reaction are one fully double-stranded molecule and one single-stranded molecule. Further, they proposed that the single-stranded molecule continued the replication process via a complementary pairing of the ITR sequences to form a panhandle-shaped molecule which served as substrate for initiation much like the linear duplex molecule. Elongation of the initiated panhandle intermediate was proposed as the mechanism of formation of Type II molecules. Elongation of Type II molecules would ultimately result in another double-stranded DNA molecule, completing the replication cycle. Combination Type I/Type II molecules were categorized as Type II molecules reinitiated for displacement strand synthesis before completion of complementary strand synthesis.

Although it is widely accepted today that adenovirus replication does utilize a panhandle intermediate in the replication cycle, there is no direct evidence for it. Electron microscopic analysis of replicating molecules that were photocrosslinked and then denatured revealed no panhandle-like structures (119). Indirect evidence in support of panhandle formation was obtained by Nigel Stow. He found that adenovirus molecules, forced to form a panhandle, apparently can do so (120). Stow made deletions to remove part or all of one molecular end of intact adenovirus DNA. Deletions removing as many as 51 base pairs at an end, when transfected into cells, resulted in production of infectious adenovirus. When examined more closely, these adenoviruses were found to contain both ends in a completely intact form, as if a complete end and a partial end had paired and the

missing part of the ITR was repaired by a DNA polymerase (120). Mutant adenoviruses that contained deletions of the entire ITR (which would be unable to form panhandles) produced no infectious virus when transfected into cells.

The panhandle intermediates proposed by Lechner and Kelly and suggested by Stow's work are what will be referred to as *cis* replication. *Cis* replication is defined here as intramolecular events, such as panhandle formation, which produce functional replicating molecules. By contrast, *trans* replication is defined as intermolecular events, such as recombination, which produce of functional replicating molecules. One example of *trans* replication has been shown by Chinnadurai *et al.* (121). They transfected overlapping terminal fragments of Ad2 and Ad5 into 293 cells (293 cells are the name of a human embryonic kidney cell line transformed with adenovirus DNA by Graham *et al.* (122), also see below). The recombinant viruses that they isolated contained the respective termini from each input fragment joined through what presumably was a recombinational event in the overlap. Similar results have been described by other laboratories (40,41) but *trans* replication has remained an understudied phenomenon.

Stow's circumstantial evidence for the panhandle intermediates in adenovirus replication was later supported by other *in vivo* work (123, 124) but the molecular mechanisms involved in the entire replication process were not elucidated until the development of an *in vitro* soluble nuclear extract system by Challberg and Kelly (125,126).

The Lechner and Kelly model for replication was an attractive interpretation of the structures observed in replicating adenovirus

molecules, but it did not suggest a solution to the adenovirus dilemma, how a linear virus initiates replication at a molecular end. All known DNA polymerases require a primer (127), yet all evidence suggested that adenovirus initiated replication at an end (90,95,98-100,128). It was not clear how or if such a structure was primed for replication. A resolution to the problem of priming adenovirus replication was originally proposed as an extraordinary property of the protein on the viral termini by Rekosh et al. (129). They hypothesized that terminal protein functions in initiation of replication, priming the synthesis of daughter strands (129). A great deal of evidence, both direct (130) and indirect (131-133), for the priming function of terminal protein has been presented in support of this novel hypothesis (also see below).

IN VITRO ADENOVIRUS REPLICATION

Direct tests of the idea of protein priming awaited suitable in vitro systems enabling the study of the initiation of replication. Previously, the only available systems permitted elongation of pre-initiated viral strands (91-93,134,135). The soluble nuclear system of Challberg and Kelly (125,126) was a landmark for the study of the biochemistry of adenovirus replication. It consisted of a nuclear extract made from adenovirus-infected HeLa cells that were treated with hydroxyurea 2 hours after infection. Extracts made at 20-22 hours post-infection contained the proteins required for replication, but were almost free of virus DNA. Maximum DNA synthesis required DNA-protein complex isolated from Ad5 virions, nuclear

extract from adenovirus-infected cells, ATP, Mg^{++}, and 4 dNTPs (125,126). Extracts made from uninfected cells did not work, nor did T7 or ϕx174 templates. Proteinase-treated adenovirus DNA was less than 5% as effective as adenovirus DNA with the 55 kDal terminal protein intact (125,126). Although modifications of the system allowed initiation of adenovirus origin sequences *in vitro* without a covalently attached terminal protein (136-138), the optimal reaction nevertheless did have such a substrate requirement. As a result, *in vitro* analysis using plasmids containing adenovirus sequences must be interpreted with caution (see below).

Electron microscopic characterization of the products synthesized *in vitro* revealed that up to 5% of the molecules were Type I, but none were unambiguously identified as Type II, suggesting that complementary synthesis did not occur readily in the system (125). Fractionation of the system ultimately enabled identification of two cellular factors and three adenovirus encoded factors now known to govern adenovirus replication (67,139). As a result, full-length adenovirus DNA can be synthesized using purified proteins only (140).

THE PROTEINS IN ADENOVIRUS REPLICATION

Precursor Terminal Protein (pTP)

At about the time of the development of Challberg and Kelly's *in vitro* system, evidence began accumulating that terminal protein was involved in adenovirus replication. Terminal protein was detected at the 5' termini of nascent replicating adenovirus DNA both *in vivo* (131,141-145) and *in vitro* (146,147). The reaction requires DNA

sequences at the adenovirus terminus (136,148,149) and is catalyzed by a complex between terminal protein and a 140 kDal protein now known to be an adenovirus encoded DNA polymerase (Adpol) (150, see below). The term 'terminal protein' is now used exclusively to designate the 55 kDal protein found on virion DNA while 'precursor terminal protein' (pTP) denotes the 80-87 kDal protein found on nascent adenovirus strands. The 55 kDal protein is derived from the 80-87 kDal protein by proteolytic processing (82, 151). Conversion of the 80-87 kDal form to the 55 kDal form is not required for replication and probably occurs just prior to packaging (152).

One of the standard assays for initiation of adenovirus replication *in vitro* measures the phosphate transfer from [α-^{32}P]-dCTP to pTP (148,149). The pTP-dCMP linkage is a phosphodiester bond between the β-hydroxyl of a serine in the pTP and the 5' hydroxyl of the terminal cytidine residue in the DNA (146,151). pTP covalently attaches to dCTP radioactively labelled at the α phosphate but not with α-labelled dGTP, dATP, or dTTP (136). Absolute requirements for formation of the pTP-dCMP complex (formed from joining pTP and dCTP), at least in crude extracts, include Mg^{++}, Adpol, ATP, and adenovirus DNA covalently bound to TP (132,136,139,148,149,153-158). The requirement for adenovirus DNA covalently bound to TP can be circumvented with terminal adenovirus sequences alone when different reaction conditions of extracts (136,137) are used.

A cellular factor designated Nuclear Factor I (NF I), not required for formation of pTP-dCMP but which greatly enhances its production, was isolated by Nagata *et al.* (153) and independently by Rawlins *et al.* (156, see below). Rijnders *et al.* (159) demonstrated specific

binding of the pTP-Adpol complex to the origin of adenovirus DNA replication and that a sequence of 14 nucleotides between positions 9-22 plays a crucial role in this binding. pTP alone appeared to be responsible for the binding (159).

After debate about its origin (160,161), pTP was shown to be encoded by adenovirus in three ways: mRNA for pTP was mapped on the adenovirus genome by electron microscopy (82), protein sequences from fragments derived from proteoltically-degraded terminal protein aligned with predicted adenovirus DNA sequences (162), and in vitro translation of early mRNA selected by hybridization to adenoviral DNa fragments spanning the region from 14.7 to 31.5 map units (163). pTP is part of the adenovirus transcription unit designated E2b which also contains Adpol (82,164). The majority of the coding sequence for pTP lies between coordinates 23.4 and 28.9 of the adenovirus genome (165,166).

It is not known how widespread the phenomenon of protein priming of DNA replication is. Other viruses, in addition to adenovirus, have protein(s) covalently attached to their genomes. They include parvovirus H1 (167), hepatitis B virus (168), plant viruses (169), the Bacillus subtilis phage ϕ29 (170), and several RNA viruses including poliovirus (171). The function of the covalently attached protein is not understood in any of these viral systems except ϕ29. ϕ29, like adenovirus, has a linear double stranded DNA molecule that is initiated for replication by protein priming (170). Only minor differences between replication of ϕ29 and adenovirus have been distinguished.

Adenovirus DNA Polymerase

It was initially unclear which DNA polymerase was involved in adenovirus DNA replication because inhibitors specific for cellular polymerase α and γ both inhibited replication (172-174). Only after it was shown that pTP co-purified with a 140 kDal protein did it begin to become apparent that adenovirus encoded its own polymerase (150,154). Orientation of pTP coding sequences in the adenovirus genome (175), demonstrated that the group N complementation mutants affecting DNA replication had mutations residing outside of the coding sequences of the terminal protein coding sequences (162,176). It is now apparent that N mutants define the adenovirus DNA polymerase. Purified fractions that contained the pTP-140 kDal complex complemented the replication defect in nuclear extracts made from cells infected with N group mutants (158,164,177). Finally, translation of coding sequences from the group N region resulted in a protein that had DNA polymerase activity. The protein was also immunoprecipitated with antibody made against the pTP-140 kDal DNA polymerase complex (178).

Adpol has an associated 3' to 5' exonuclease activity that can function in a proofreading capacity, like other viral and prokaryotic DNA polymerases (179). Adpol is a highly processive enzyme. DNA chains synthesized by Adpol can be up to 30,000 nucleotides in length under optimal conditions (179). By contrast, HeLa DNA polymerase α synthesized chains only 100 to 1000 nucleotides long. Adpol, unlike polymerase α, poorly elongates RNA-primed DNA. It was the only polymerase tested that was capable of labelling pTP with [α-^{32}P]-dCTP

(180). AdDBP stimulated DNA synthesis by Adpol using poly(dT) as the template and oligo(rA) or oligo(dA) as the primer, but did not stimulate DNA synthesis by HeLa DNA polymerase α on similar templates (179). E. coli single-stranded DNA binding protein cannot functionally replace AdDBP. Adpol is very tightly complexed with pTP, requiring glycerol gradients containing 1.7 M urea and 0.5 M NaCl for separation (154). In addition, Adpol is insoluble in the absence of terminal protein (178).

The region of the genome potentially encoding Adpol is designated E2b and has an open reading frame that would code for a protein of 120 kDal (164,175). Several messages for the E2b transcriptional unit have been mapped on the adenovirus genome (147). The only analysis of the E2b transcriptional unit has been at the electron microscopic level (82), so it is not yet clear where coding sequences for the 5' end of the pTP gene or the Adpol begin.

Adenovirus DNA Binding Protein

The abundance and availability of mutants enabled early in vivo characterizations of the functions of AdDBP in replication (see above). Nonetheless, the development of in vitro systems gave additional insight into the function of AdDBP in replication. AdDBP does not appear to be required for initiation of replication (106,149,181), and actually inhibits replication in the absences of the cellular NF I protein (153). Its primary role, as determined in in vitro systems, appears to be in elongation (58,140). This is probably mediated by its effect on Adpol (179). DNA synthesis by the

Adpol-pTP complex is stimulated up to 100-fold by the AdDBP (179), and AdDBP has an inhibitory effect on the 3'-5' exonuclease activity of the Adpol (179).

Nuclear Factor I

As the *in vitro* system was increasingly purified and fractionated, it was observed that nuclear extracts from uninfected cells appeared to stimulate replication (137,157,182). Ultimately the factor, called Nuclear Factor I (NF I) (153,156) present in uninfected cells was purified. NF I binds the sequence TGG($N_{6 \text{ or } 7}$)GCCAA (N refers to any nucleotide) found in the ITR of human adenovirus serotypes 2,3,5,7,12,31, and even the simian adenovirus SA7 (153,156). DNase I protection experiments indicate that NF I protects a 24-28 base sequence spanning the consensus binding sequence (155,156), but only the bases from 24-41 are required for binding of NF I (183).

NF I functions to enhance initiation of replication *in vitro*, but is not absolutely required for initiation (156). Indeed, some adenovirus serotypes do not contain a consensus NF I site within their ITRs. Ron Hay (184) compared rates of replication *in vivo* of plasmids carrying the terminal sequences of serotypes Ad2 and adenovirus 4 (Ad4) in either Ad2 or Ad4 infected cells. Ad4, like several other human serotypes, has no consensus NF I site within its ITR (185). Replication rates of Ad2 and Ad4 origins were equal in cells infected with Ad4, but the replication rate of the Ad4 origin was far below that of the Ad2 origin when incubated in Ad2 infected cells (184). It is not known how or why some serotypes have eliminated the requirement

for NF I.

Consensus NF I sites are located upstream of the promoter sequences of the human c-myc oncogene (186), within the immunoglobin μ gene (187), in the BK viral enhancer (188), flanking sequences of the chicken lysozyme gene (189,190), and in the MMTV LTR (188). It is not known, however, if NF I actually plays a role in the expression of these genes. NF I binding sites in the human c-myc and the chicken lysozyme gene correspond to DNase hypersensitive sites and may indicate a role for NF I in activation of chromatin (186,189). NF I-like proteins have now been identified in a wide variety of eukaryotic organisms, from chicken oviduct and mouse liver to humans, a range which also spans the evolutionary range of known adenoviruses (187-191). The location of the NF I site within the adenovirus ITR places it upstream of the E1a transcriptional unit, an arrangement that places consensus NF I sites upstream of at least two oncogenes, E1a and c-myc.

Finally, a sequence analysis of the adenovirus genome performed using BiT Software Specialties Apple Sequence Analysis Program revealed that the Ad2 genome contains three consensus NF I sites. Two are located within the terminal ITRs. The third is located at nucleotide 10,767, which places it directly between the template regions for two adenovirus viral associated RNAs (Ahern, K.G., unpublished observations). Perhaps more importantly, this sequence is immediately 5' to the pTP open reading frame, which is itself immediately upstream of the Adpol coding sequences. While it is possible that occurrence of a consensus NF I binding site upstream of

adenovirus sequences coding for replication proteins is by chance alone, its proximity to the replication genes and its location upstream of other transcriptional units (187-191) make it a sequence that should be examined more closely.

NUCLEAR FACTOR II

After the isolation of NF I from uninfected cells, efficient initiation of replication was made possible with purified proteins (153). DNA strands only 25% of the length of the genome were synthesized under these conditions (140,153). Purification of another activity, Nuclear Factor II (NF II), from uninfected cells enabled synthesis of full length adenovirus strands (139,140). Its involvement in adenovirus replication does not appear to be at the level of initiation of replication (139). NF II copurifies with a type I topoisomerase (topo I) activity, and NF II can be functionally replaced by purified topo I from HeLa cells or calf thymus but not by bacterial topoisomerase I (140). Aphidicolin treatment also blocks adenovirus DNA synthesis (192). The blocked intermediates produced by aphidicolin treatment have DNA synthesis halted in the same region of the viral DNA as when elongarion of replication occurs in the absence of NF II (192). The significance, if any, of these observations is unknown.

SEQUENCE DEPENDENCE OF ADENOVIRUS REPLICATION

For a review of adenovirus sequences see van Ormondt and Galibert

(185). The ITRs of the adenoviruses that have been sequenced have several features in common. First, all but the avian Chick Embryo Lethal Orphan (CELO) virus have a C at the 5' end of each strand covalently attached to pTP. CELO virus has a G (56). At positions 9-17 there occurs a sequence highly conserved among all adenoviruses examined. When only human sequences are compared, the homology extends to position 20 (193). These conserved sequences have been directly implicated in initiation of adenovirus replication by many investigators using a variety of systems. Analyses of sequence requirements for adenovirus replication have been done both in vitro and in vivo. Early studies were performed in vitro. Two basic strategies to detect initiation were used, the phosphate transfer reaction to pTP (148,149, see above) and assays which measured incorporation of labelled dNTPs into specific DNA templates (156).

Origin sequences required for replication can be mapped by mutations. Plasmid molecules carrying adenovirus origin sequences are more readily manipulated and mutated than viral molecules and have largely been used for analysis, both in vitro and in vivo. They carry the built-in disadvantage, however, that they do not contain the pTP required for optimal synthesis in vitro (125,126). While this has largely been ignored (see below), it is important to be aware of this when interpreting sequence requirements.

The in vitro system requires that plasmid sequences be linearized to expose an adenovirus terminus for efficient labelling of pTP with [α-^{32}P]-dCTP (136). Supercoiled plasmids containing the same adenovirus sequences or plasmids linearised with internal adenovirus sequences will not function in the assay (136). Results of in vitro

studies indicate that a conserved sequence between nucleotides 9 and 18 of the ITR is essential for replication. Deletions or mutations which affect any of the sequences between nucleotides 9 and 18 greatly inhibit labelling of pTP with [α-^{32}P]-dCTP (58-60,194) and production of displaced strands when viewed by electron microscopy (61,195). Sequences outside the strictly conserved region between nucleotides 9 and 18 also seem to be required for enhancement of replication *in vitro*. Rawlins *et al.* (156) showed that the first 18 bases of the viral genome were sufficient to support a limited degree of initiation, but that the presence of sequences between nucleotides 19 and 67 greatly improved its efficiency.

All of the work described above was performed on mutated plasmid sequences lacking pTP. Stillman *et al.* (193) assayed replication *in vitro* using as templates adenovirus virion DNAs from a variety of serotypes. The serotypes chosen had many regions of variation in the ITR sequences. The advantage of this approach is that pTP is not a limiting factor, since it was present on each of the virion DNAs. Their results indicate that when extracts were made from Ad2-infected cells, all of the serotype DNAs (2,4,7,9, and 31) analyzed supported initiation and elongation, but only poorly with serotypes 4, 9, and 31. Serotypes 4 and 9 diverge considerably from Ad2 after nucleotide 20, so it is not surprising that their replication efficiency was low. Serotype 31, on the other hand, is quite similar to Ad2 for the first 50 nucleotides, and it is interesting that its replication should be affected.

The sequence from nucleotides 25 to 45 contains the consensus NF I binding site found in serotypes 2, 3, 5, 7, 12, and 31 (185). When

this sequence is deleted, labelling of terminal adenovirus DNA fragments *in vitro* is greatly reduced (183). It has been proposed that the sequences from 9-18 and 19-67 constitute two domains necessary for efficient adenovirus replication (58,156,196). The first domain is necessary for binding of the pTP-Adpol complex to the end of the adenovirus DNA, and is absolutely required for initiatioin of replication. The second is the NF I binding site, and is not absolutely required for initiation *in vitro*, but instead serves to enhance initiation.

Recently, a new cellular protein designated Nuclear Factor III (NF III) has been shown to bind immediately adjacent (nucleotides 40-50) to the NF I site in the adenovirus ITR and to stimulate adenovirus replication moderately (P. van der Vliet, personal communication). Beyond these domains, no other sequence in the ITR has been determined to be essential for any aspect of adenovirus replication despite the fact that there are short regions of sequence conserved between serotypes. Some of these sequences are GC rich and have homology to sequences near the SV40 origin and to sequences in polyoma virus and BK virus. It is not known if they function similarly in adenovirus.

Analyses of sequence requirements for replication *in vivo* have been employed more recently (123,124,184,196). *In vivo* studies suffer from some of the same difficulties as *in vitro* studies. Plasmid molecules containing adenovirus origin sequences are typically used. The optimum template for replication *in vivo* is not yet known.

The approach taken to studying replication *in vivo* has been to transfect plasmid sequences containing putative origin

sequences into cells with helper virus DNA to provide replication proteins. The plasmids are then assayed for replication. Two methods have been used to identify replication of the plasmid sequences. The first detects the presence of a new molecular size produced by a hypothesized repair of a *cis* replication event (123,124). The other examines the production of *Dpn*I-resistant plasmid sequences as a measure of replication (154,184). This method is based on the presence in the plasmid of the *Dpn*I restriction endonuclease recognition sequence, $G^{m}ATC$ (197). Eukaryotic cells do not have a methylase for adenine and, thus, *Dpn*I-resistant plasmid DNA is produced as a consequence of replication. Results of these analyses containing varying lengths of the adenovirus ITR revealed that molecules containing only the terminal 45 base pairs of the Ad2 ITR replicate as efficiently as molecules with the entire ITR (196). Deleted molecules retaining 36 base pairs or less of the ITR failed to replicate detectably (196).

In contrast, Wang and Pearson (123) detected replication of plasmid sequences containing only 30 base pairs of the ITR. This discrepancy remains unexplained. One interpretation is that the plasmids used by Hay (196) were inefficient at complementary replication compared to the ones used by Wang and Pearson, perhaps due to differing lengths of inverted sequences capable of forming a panhandle (123). Nevertheless, there does seem to be a distinct difference between sequence requirements for replication *in vivo* and *in vitro*. This apparent discrepancy may be due to the fact that studies *in vitro* are only detecting a part of the replication cycle, displacement replication, while the situation in the infected cell is

more complicated. Intriguingly, only the adenovirus terminal 18 base pairs appear to be required for replication in Ad4-infected cells (184). It is not known at this time why this is so, but it has been speculated that since Ad4 does not have a consensus NF I site within its ITR, that it must have other mechanisms for replication which do not require NF I (184).

OTHER SEQUENCES IN THE ITR AND THEIR ROLE IN REPLICATION

The ITR sequences of all human serotypes have in common an AT-rich region of about 50 base pairs at the terminus and an internal GC-rich region (193). A role for the AT rich region for replication *in vivo* and *in vitro* is becoming apparent (see above), but a role for the GC-rich region is not. Chow and Pearson (198) described site-specific nicking in the GC-rich part of the Ad2 ITR, but its relationship, if any, to replication of adenovirus is unknown. At least one GC-rich sequence, GGGCGG, conserved among the adenovirus ITRs (185), has the consensus binding sequence of the SP1 transcription factor described by Tjian's laboratory (199,200). The ITRs of all the human adenoviruses that have been examined, with the exception of serotype 4, all terminate at or near the 3' end of the ITR with the sequence TGACG or one closely related to it (193,201). Ad4 has only one copy of the TGACG sequence. It is not known what role, if any, this sequence plays in the replication of adenovirus.

It is curious that the studies to date have only identified a requirement for about half of the ITR for adenovirus DNA replication. It is important to note, however, that all *in vitro* studies of

adenovirus DNA replication have focused on the roles of factors and sequences involved in displacement synthesis. Synthesis of the complementary strand may be much more complicated and have different sequence requirements than displacement strand synthesis. This may become more apparent as progress is made in understanding replication *in vivo*, or as *in vitro* systems that support complementary replication are developed.

UNEXPLAINED ASPECTS OF ADENOVIRUS REPLICATION

van der Vliet *et al.* (202) have obtained evidence that a heat-stable, ribonuclease-sensitive fraction from the cytosol of uninfected HeLa cells can stimulate initiation of adenovirus replication 3-fold. The fraction appears to contain at least one, and possibly two, RNA components which complement each other. The role of an RNA molecule in adenovirus replication is not apparent at this time, but small RNA molecules have been implicated in processes such as RNA splicing (203,204).

In contrast to the traditional view of monitoring adenovirus replication by measuring labelling of pTP by [α-^{32}P]-dCTP, Pearson *et al.* (195) and Enns *et al.* (61) demonstrated adenovirus dependent DNA replication of plasmids containing adenovirus origin sequences by electron microscopy using the *in vitro* system of Challberg and Kelly (125, 126). Replication was scored by counting plasmid molecules with a visible displayed single strand projecting from them. 5-10% of all plasmid molecules observed contained visible, displaced, single-stranded tails. One major and several minor origins were mapped (195). The

major origin was located at the boundary of the adenovirus-pBR322 junction. Replication proceeded in a direction consistent with the polarity of the adenovirus sequences. In apparent contrast to virtually every other study of adenovirus replication, strand displacement replication occurred even when the adenovirus sequences were not at the molecular terminus, but instead were bounded by flanking pBR322 plasmid sequences. No explanation has ever been given for this apparent discrepancy with other systems for measuring initiation of replication. It may be that the system of Pearson et al. is detecting an otherwise unobserved mechanism for adenovirus replication.

Results obtained by Frank Graham's laboratory (205,206) may be relevant to the observations of Pearson's laboratory. They detected what are apparently circular adenovirus molecules in a Ad5-infected BRK and HeLa cells. Restriction analysis that showed head-to-tail joints of a fraction of the viral DNA (205). Up to 15% of the DNA was detected in this form. It was detected as early as 3 hours and was still present as late as 120 hours post infection (205). In addition, Graham (206) transfected synthetic, circular Ad5 molecules into 293 cells and showed that they were just as infectious as transfected linear Ad5 DNA. In related experiments, Ahern et al. (Ahern, K.G. Wang, K. and Pearson, G.D. manuscript in preparation) and Hay (124) have observed perfect excision of adenovirus ITRs from flanking plasmid sequences replicated in vivo in the presence of adenovirus when the ITRs are in a separated, but inverted, orientation.

Replication of a circular adenovirus to generate linear adenovirus molecules and perfect excision of adenovirus termini from flanking

sequences means that a mechanism to liberate linear molecules must exist. The apparently high infectivity of circles (206) and the very efficient excision of adenovirus termini suggest that the processes are not rare events in the cell. If indeed the process responsible for these observations is an internal initiation, as has been proposed by Hay (124), then the data of Pearson's laboratory (61, 195) are provocative. It represents the first demonstration of adenovirus replication on a template that has been shown not to be labelled by pTP in vitro (136), an observation that invalidates the widespread assay of the labelling of pTP by α-^{32}P-dCTP as the first step in adenovirus replication.

COMPLEMENTARY STRAND SYNTHESIS

The adenovirus replication scheme is consistent with all other DNA replication systems in general. Initiation occurs at one or both ends of the virus in a sequence dependent manner (94,98,128,207). Replication proceeds by a strand displacement mechanism, ultimately liberating one of the two parental strands (109,118). Adenovirus replication is semi-conservative (134,135,208), but since only displacement synthesis has been rigorously studied, the semi-conservative nature of adenovirus replication is only partly understood. Unfortunately, the intense work that went into understanding initiation and elongation of adenovirus replication has not been followed by studies to determine the rest of the viral replicative cycle.

It is known that a single strand is displaced during replication

(62) but virtually nothing is known about how it is ultimately converted into double-stranded DNA. This phase of the replicative cycle will be referred to here as complementary strand synthesis. Complementary strand synthesis is postulated to proceed via an intramolecular pairing between the two complementary ITR sequences at the ends of the DNA molecule virus (62, see Figure 1). Stow's work suggesting that panhandles can form *in vivo* has been expanded by Wang and Pearson (123) and Hay *et al.* (124). They constructed molecules containing a single adenovirus ITR followed by a sequence repeated in an inverted orientation at the end of a plasmid molecule. When these molecules were linearized to expose the adenovirus origin and transfected into cells with helper adenovirus DNA, both the input plasmid and a slightly larger molecule were detected (123,124). The new molecule was the size predicted if strand displacement replication had occurred and the inverted sequences in the input molecule paired, followed by DNA polymerase filling in the ends to form a panhandle (123,124). Such a molecule would then have two origins, just like adenovirus and presumably could continue replication just like adenovirus. In the absence of helper adenovirus DNA, only the input plasmid was detected (124). Similarly, only the input plasmid was detected if the molecule was not linearized to expose the adenovirus origin or if a plasmid was used that did not have inverted sequences capable of forming a panhandle (123,124). The circumstantial evidence for panhandle formation provided by Stow (120), Wang and Pearson (123), and Hay (124) should be followed by investigations of alternative mechanisms for completing the adenovirus replication cycle to provide as comprehensive an analysis of adenovirus replication as

possible.

One obvious alternative to *cis* replication is *trans* replication, as detailed above. Although the mechanism(s) of *trans* replication is routinely characterized as recombination, there are other ways for joining isolated origins in *trans*, such as intermolecular pairing of displaced complementary strands. Surprisingly, this simple idea has not been tested in the adenovirus system. Determining the mechanism(s) of complementary strand replication is central to understanding how adenovirus replication works and its relationship to other DNA replication schemes. The data generated by a study of *trans* replication will likely also give information about recombination, and may tell how replication and recombination affect each other. A complete understanding of the mechanism(s) which adenovirus uses to complete its replication cycle is central to understanding adenovirus replication and its relationship to other DNA replication schemes.

The aim of the work described here is to study *trans* replication as an alternative mechanism to *cis* replication in completing the adenovirus replication cycle. Distinct adenovirus origin-containing plasmids from the left and the right ends of Ad2 will be separated in *trans* and replicated in the presence of a helper virus *in vivo*. Parameters for the ultimate joining of the separated origins will be examined. Recombination studies have clearly shown that such sequences can be put together in adenovirus infections (40,41,121). The purpose will be to attempt to elucidate the mechanism(s) by which these sequences are put together and understand how that impacts on the process of replication. By approaching replication in such a manner it may be possible to further define requirements for

recombination in adenovirus infected cells.

Since the sequences to be joined are on plasmids, it will be possible for the first time to study recombination of adenovirus termini in the absence of an adenovirus infection. Such a non-traditional approach to replication may ultimately give insight into *cis*-acting mechanisms in replication and provide information about the way in which replication occurs in *trans*.

CHAPTER 2: MATERIALS AND METHODS

REAGENTS/ENZYMES

Chemicals and supplies were purchased from the companies as listed below: β-mercaptoethanol (Sigma), sodium dodecyl sulphate {SDS} (BioRad), NaCl (Baker), diethyloxydiformate (Eastman), lysozyme (Sigma), [α-^{32}P]-dATP 3000 curies/millimole {ci/mmol} (New England Nuclear), dNTPs (Sigma), synthetic molecular linkers (BRL), ethidium bromide (Sigma), agarose (Sigma), EDTA (Sigma), $MgCl_2$ (Baker), KCl (Baker), potassium acetate (Baker), ultra-pure formamide (BRL), electron microscopic grids {200 mesh} (Pelco), Pt/Pl wire 80/20 (Pelco), chloroform (Baker), nuclease-free bovine serum albumin {BSA} (BRL), fetal calf serum (Hyclone), $CaCl_2$ (Baker), nick-translation kit (Cooper Biomedical), ammonium acetate, HEPES (Sigma), Sephadex G50 (Sigma), ampicillin (Sigma), tetracycline (Sigma), kanamycin sulfate (Sigma), chloramphenicol (Sigma), penicillin/streptomycin (Gibco), ammonium persulfate (BioRad), yeast extract (Difco), AG501-X8 ion exchange resin (BioRad), CsCl (BRL), and Gene Screen Plus (New England Nuclear).

All enzymes with the exception of RNase A (Schwarz/Mann), RNase T1 Sigma), lysozyme (Sigma), and Pronase (Calbiochem) were purchased from New England Biolabs, Bethesda Research Laboratories, Boehringer Mannheim Biochemicals, or Pharmacia and used according to the manufacturer's specifications. Typically 1-5 units of restriction enzyme were used per microgram of DNA to obtain complete cleavage. Reactions were for 1-2 hours at the temperature recommended by the

manufacturer and reaction mixtures were phenol-extracted to remove proteins after the cleavage.

Water-saturated phenol for phenol extractions was prepared by distilling crude phenol, and incubating the distillate with successive volumes of 1 M Tris (pH 8) until the pH of the aqueous layer was 7.5. The aqueous layer was removed and the phenol was either mixed with an equal volume of TE (10 mM Tris, 1 mM EDTA, pH 7.5) to make phenol saturated with water or the phenol was mixed with an equal volume of chloroform and then mixed with TE buffer (pH 7.5) to make $CHCl_3$/phenol (50/50) saturated with water.

MEDIA

Bacterial medium was modified YT broth which consisted of yeast extract (5 g/l), casein enzyme hydrolysate (10 g/l), and NaCl (5 g/l). Plates made with this medium also included 15 g/l of Bacto-Agar. Antibiotics were used at the following concentrations: ampicillin, 50 µg/ml; tetracycline, 15 µg/ml; kanamycin, 50 µg/ml; and chloramphenicol, 50 µg/ml. When plasmids were amplified, chloroamphenicol was added at a concentration of 1 ml of a 34 g/100 ml solution per liter of cells.

Tissue culture medium was Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% fetal calf serum, penicillin, streptomycin, and glucose (4.5 g/l). Supplementation of glucose to 4.5 g/l greatly augmented the ability of 293 cells to stick to the tissue culture plates during $CaPO_4$ transformations.

DNA ISOLATION

Plasmid DNAs used in these constructions were isolated from E. coli cells according to the procedures of Schlief and Wensink (209), or Maniatis et al. (210). DNA was either precipitated from crude preparations made from bacteria by using sodium trichloroacetic acid/95% ethanol (50/50) or directly suspended in cesium chloride and cesium banded as per Schlief and Wensink (209).

DNA was collected from the lower band in the cesium gradient by elution through the bottom of the centrifuge tube. Ethidium bromide was extracted from the cesium banded DNA by extracting at least eight times with an equal volume of butanol and diluting with two volumes of water. The DNA was precipitated by mixing with six times the original volume of 95% ethanol, freezing at -20°C overnight, and microfuging for 30 minutes. The pellet was drained, dried under vacuum and resuspended in TE buffer (pH 7.5).

DNA was isolated from transfected 293 cells (122) by the rapid nuclear isolation method (211). A 6-cm dish of cells 50-75% confluent was harvested by vigorously washing the dish with the suspension medium. These were transferred to a 15 ml Falcon tube and centrifuged at 1,000 revolutions per minute (rpm) at 4°C in an SS34 rotor. The supernatant was discarded. The pellet was resuspended in 500 µl of NP40 solution. (0.01M NaCl, 0.01M Tris pH 7.4, 0.0015M $MgCl_2$, 0.5% NP40), placed on ice for 10 minutes, then shaken vigorously for 10 seconds. This material was centrifuged at 8000 rpm in an SS34 rotor for 10 minutes at 4°C and the supernatant was saved as the cytoplasmic fraction. The pellet was resuspended in 200 µl of 0.6% SDS, 10 mM

EDTA overnight. The next morning the nuclear mixture was treated with 10 µl of RNase for 3 hours at 50°C, then with Pronase at 1 mg/ml for 3 hours at 37°C, then extracted with phenol, phenol/chloroform (1/1), and chloroform/isoamyl alcohol (24/1). Finally the DNA was precipitated by addition of 100 µl of 7.5 M ammonium acetate and 600 µl of 95% ethanol. After centrifugation, the pellet was resuspended in 200 µl of TE buffer (pH 7.5).

PREPARATION OF BLUNT ENDED DNA

For all of the descriptions below, blunting refers to the treatment of 1 µg of DNA in 50 µl of solution containing 20 µM dATP, 20 µM dTTP, 20 µM dCTP, 20 µM dGTP, 40 mM KPO_4, 6.6 mM $MgCl_2$, 1 mM β-mercaptoethanol, with 2 units of DNA Polymerase I (Klenow fragment) at 37°C for 15 minutes. Prior to ligation, blunted DNA was eluted through a spin-column prepared as described below.

NICK TRANSLATION OF DNAs

Labelled DNAs for probing Southern blots were made with the Cooper Biomedical nick translation kit according to the manufacturer's specifications. CsCl-banded DNA (1 µg) was incubated with 10 µl (200 µCi) of [α-^{32}P]-dATP (3000 Ci/mmol) in 50 µl of 20 µM dCTP, 20 µM dGTP, 20 µM dTTP, 1 µM unlabelled dATP, 50 mM Tris-HCl (pH 7.2), 10 mM $MgCl_2$, 0.1 mM dithiothreitol {DTT}, 50 µg/ml BSA, 5 units DNA polymerase I, 125 pg DNAse I. The reaction was at 16°C for 1 hour. The nick translated DNA was purified before use by eluting through a

spin-column (see below).

SPIN COLUMN PURIFICATION OF DNA

A one ml syringe was filled with Sephadex G50 suspended in TE buffer (pH 7.5), centrifuged at setting 4 on an IEC clinical desk top centrifuge at room temperature for 3 minutes, and washed 2 times at the same centrifuge setting with 200 µl of TE buffer (pH 7.5). Samples (100 to 200 µl) containing the DNA (prepared by diluting with TE buffer (pH 7.5)) was eluted through the column at setting 4 and collected for later use.

PREPARATION OF COMPETENT E. COLI CELLS

The E. coli strains used in these studies were C600, JM83, JC9387, or A716. JC9387 and A716 were generously provided by Dr. Franklin Stahl, University of Oregon. Preparation of competent cells for each strain was identical and is a modification of the method of Hanahan (213). Cells were maintained as single colonies on YT plates until used. An overnight incubation of one colony from a plate in 20 ml of YT broth was made at 37°C. The next morning, 10 ml of cells from this flask were inoculated into one liter of YT broth supplemented with 10 mM $MgCl_2$. Cells were allowed to grow at 37°C until reaching an optical density of 0.5 at 590 nm (about 2 hours).

At this time, the flask with the cells was sequentially placed on ice for 10 minutes, centrifuged at 4000 x g for 5 minutes at 4°C, resuspended in 500 ml of 50 mM $CaCl_2$, incubated on ice 20 minutes,

centrifuged at 4000 x g for 5 minutes, and resuspended in 20 ml of 50 mM $CaCl_2$/50% glycerol. The cell mixture was incubated on ice for 12 hours and 50 µl was pipetted into a sterile 0.5 ml microfuge tube and frozen at -80°C until used. Cells prepared in this way were stable for at least 18 months. To prepare the frozen cells for transformation, they were thawed on ice for 5 minutes, mixed with 150 µl of aqueous 50 mM $CaCl_2$, allowed to sit for 5 minutes on ice, then mixed with DNA.

PREPARATION OF DNA FOR BACTERIAL TRANSFORMATION

Ligated DNAs used to make the clones described below were prepared in the following manner, unless otherwise described. Deletion clones (made by removing a sequence by restriction cutting and blunting if necessary) utilized 100 ng of linearized DNA purified through a spin column which was ligated in a 20 µl reaction volume with T4 DNA ligase using the manufacturer's specifications at temperatures ranging from 9°C for 16 hours to 22° for 4 hours. No significant difference was observed in transformation efficiency between these two conditions.

Insertion reactions utilized 200 ng to 1 µg of linearized vehicle DNA (blunted if necessary) and a 5 to 1 molar ratio of DNA (also blunted if necessary) to be inserted. Occasionally DNA bands were purified from agarose gels in the following manner. The DNA band was cut from an agarose gel and the band was sliced into tiny slices. This slurry was mixed with an equal volume of phenol, vigorously shaken for 5 minutes at room temperature, frozen at -80°C for 15 minutes, then spun in a microfuge for 8 minutes. The top layer was

collected and subjected to multiple phenol extractions, followed by extraction with $CHCl_3$/Phenol (1/1), extraction with $CHCl_3$, and precipitation with one-half volume of ammonium acetate (7.5 M) and six volumes of ethanol (95%). All DNAs were spin-column purified before ligation.

TRANSFORMATION OF BACTERIAL CELLS

Transformation of competent E. coli cells described above involved mixing one tube of cells with DNA (1 ng of supercoiled DNA or 0.1 to 1 µg of ligated DNA) on ice for 30 minutes, heat shocking at 42°C for 2 minutes, adding 1 ml of YT broth, and incubating at 37°C for 1 hour. Aliquots (25 µl to 200 µl) were spread on YT plates with a selective drug. The plates with the cells were incubated at 37°C for 16 hours.

GROWTH AND PURIFICATION OF ADENOVIRUS DNA

Adenovirus DNA used in these studies was prepared as follows. One liter of actively growing HeLa cells ($4\text{-}5x10^5$ cells/ml) in DMEM with 5% fetal calf serum were suspended in 100 ml of DMEM medium in a small flask then mixed with virus at a multiplicity of infection (MOI) of 10-50 per cell. Incubation was for one hour at 37°C. Then 400 ml of original medium and 500 ml of new DMEM was added. The cells were incubated for 40-48 hours at 37°C then centrifuged at 5000 rpm for five minutes. All centrifugations except as noted utilized an HB4 rotor at 4°C. Twice, the pellet was resuspended in 10 ml of phosphate-buffered saline deficient buffer (PBSd) (160 g NaCl,

4 g KCl, 4 g KH_2PO_4, 23 g $NaHPO_4$ per liter solution), and centrifuged for five minutes at 5000 rpm. The cells were sonicated for three minutes and the mix was centrifuged at 10,000 rpm for 10 minutes. The supernatant was transferred to a new tube and the pellet was resuspended in 5 ml of PBSd. The pellet mix was resonicated, centrifuged, and the pellet and supernatant were separated as above. The supernatants were combined, extracted with 10 ml of freon, and centrifuged at 10,000 rpm for 10 minutes. The top layer was layered onto a step gradient that consisted from bottom to top of 5 ml of PBSd, 20% glycerol, 0.6 g/ml CsCl followed by 15 ml of PBSd, 20% glycerol. This was centrifuged in an SW28 rotor at 24,000 rpm for 90 minutes at 4°C. The bottom band was collected and dialyzed twice against TE buffer (pH 7.5), for one hour each. The dialyzed material was RNased for one hour at 37°C. Then SDS was added to 0.6% and Pronase was added to 1 mg/ml. This was incubated at 37°C for two hours then extracted once with phenol, once with chloroform/isoamyl alcohol (24/1), and once with ether. NaCl was added (0.2 M) and 2-2.5 volumes of 95% ethanol was added. The sample was frozen at -20°C for at least one hour and precipitated at 10,000 rpm for 30 minutes in an HB4 rotor at 4°C. Typically 250-300 µg or adenovirus DNA was obtained from one liter of cells.

TRANSFECTION OF 293 CELLS

The human embryonic kidney cell line, 293 cells, was transfected exactly according to the $CaPO_4$ precipitation technique as previously described (214). Transfections utilized 3 ug of each of the plasmid

species, 3 ug of Ad2 DNA, and sufficient salmon sperm DNA to make 15 µg per 6 cm Petri dish. DNAs were harvested at 57 hours except as noted.

GEL PREPARATION

Routine analysis of DNA fragments was by agarose gel electrophoresis with the vertical so-called "mini-gels" (Idea Scientific). Gels consisted of 1% agarose (W/V) in TBE buffer (10.8 g Tris Base, 5.5 g Boric Acid, 0.93 g EDTA per liter liquid). Agarose gels for Southern analysis utilized 200 ml of 0.8% agarose (W/V) in TBE buffer. This was poured into a slab gel apparatus (BRL) with dimensions 9 3/4 in. by 7 7/8 in. Samples were loaded in 0.5% Bromphenol Blue (Sigma) as a marker and electrophoresis was at 50-100 volts until the dye had traveled 90% of the length of the gel.

SOUTHERN BLOTS

All analyses of DNA sequences isolated from 293 cells was by Southern blotting (215). After electrophoresis (see above) 0.8% agarose gels were stained in ethidium bromide (1 µg/ml) for 20 minutes with gentle agitation then exposed to long wavelength UV light (Ultraviolet Products, Inc.) for 2 minutes. The gel was removed from the UV light, mixed with 400 ml of 0.4 M NaOH/0.6 M NaCl for 30 minutes, and then neutralized with 0.5 M Tris (pH 7.4)/0.6 M NaCl for 30 minutes. DNA was then eluted from the gel by capillary action for 16-24 hours onto GeneScreen Plus paper exactly as described by the

manufacturer. The GeneScreen Plus paper was denatured with 0.2 M NaOH for 30 seconds and neutralized with 0.2 M Tris-HCl (pH 7.4) for an additional 30 seconds and left to dry.

Prehybridization of the paper was for 16-72 hours at 65°C in 10-20 ml of a solution that consisted of 10% dextran sulphate (w/v), 1 M NaCl, 1% SDS (w/v), salmon sperm DNA (0.15 mg/ml), and heparin (1 mg/ml). Longer prehybridization times resulted in somewhat lower backgrounds. After the prehybridization reaction, the solution was replaced with fresh solution that was the same, except that it contained no heparin and 2-5x10^6 counts per minute (cpm) of nick-translated probe (pUC8 or pBR322). Hybridization was at 65°C for 12-24 hours. Then the paper was removed from the hybridization solution and washed exactly according to the specifications of the Genescreen Plus manufacturer. Autoradiography was performed with Kodak X-OMAT XAR-5 X-ray film for 24-120 hours utilizing an intensifying screen at -20°C.

DNAs USED IN CONSTRUCTION OF CLONES

Clones 7, 17, and dl18 have been described previously (61). pHSK is a plasmid carrying the kanamycin-resistance gene of Tn5 (212) between the HindIII and SalI sites of pBR322 constructed by Greg Bertoni in Dr. Dallice Mills' laboratory at Oregon State. pXD6 was constructed by Kai Wang in George Pearson's laboratory at Oregon State by treating pIB IV (provided by Kathy Berkner at Zymogenetics) with HindIII and religating. pXD6 contains of 1000 base pairs from the right end of Ad2 DNA inserted into the HindIII to EcoRI site of pBR322

in such a way that the ITR sequences terminate at the EcoRI site and the HindIII site at 34,937 on the Ad2 map is at the HindIII site of pBR. pADIR was also constructed by Kai Wang and it consists of 1364 and 1000 nucleotides from the left- and right-ends of Ad2 respectively, joined at the terminus of each adenovirus ITR sequence by an EcoRI linker. This fragment was cloned into the EcoRI-HindIII sites of pBR322 in the orientation that the HindIII site 1000 bases from the right terminal sequences is joined to the HindIII site of pBR322.

CONSTRUCTION OF CLONES

In the diagrams that follow, these abbreviations are used: tetracycline resistance (T^r); ampicillin resistance (A^r); kanamycin resistance (K^r). All blunting reactions utilized the Klenow fragment of DNA polymerase I and all four dNTPs as described above. p44x (Figure 2) is a derivative of pBR322 made by inserting an XhoI linker into the blunt-ended EcoRI site of pBR322. pKT14 (Figure 3) was made by cutting pHSK at the unique SmaI site, inserting an EcoRI linker, and cleaving with EcoRI to liberate a 1.3 kb fragment containing the kanamycin-resistance gene with sticky EcoRI ends. This was ligated into the EcoRI site of plasmid dL18. pKAT21 (Figure 4) is a derivative of pHSK made by inserting the small XhoI to SalI fragment containing the 5' end of the tetracycline-resistance gene of plasmid p44x into the unique SalI site of pHSK. pD25 (Figure 5) is a derivative of clone 17 made by cutting at the unique XbaI and HindIII sites in the molecule, blunting, ligating and transforming. pRS17

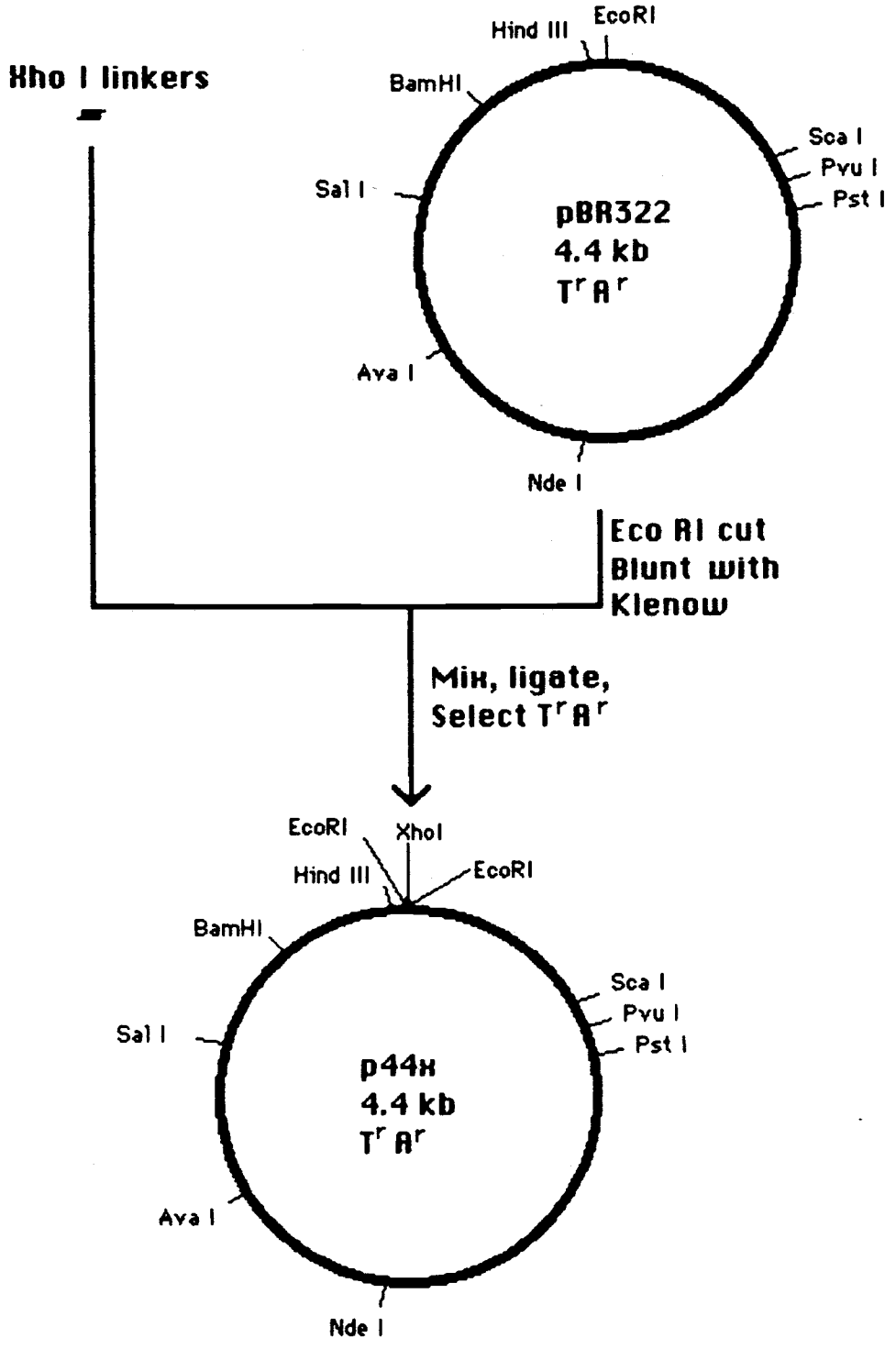


Figure 2: Construction of p44x

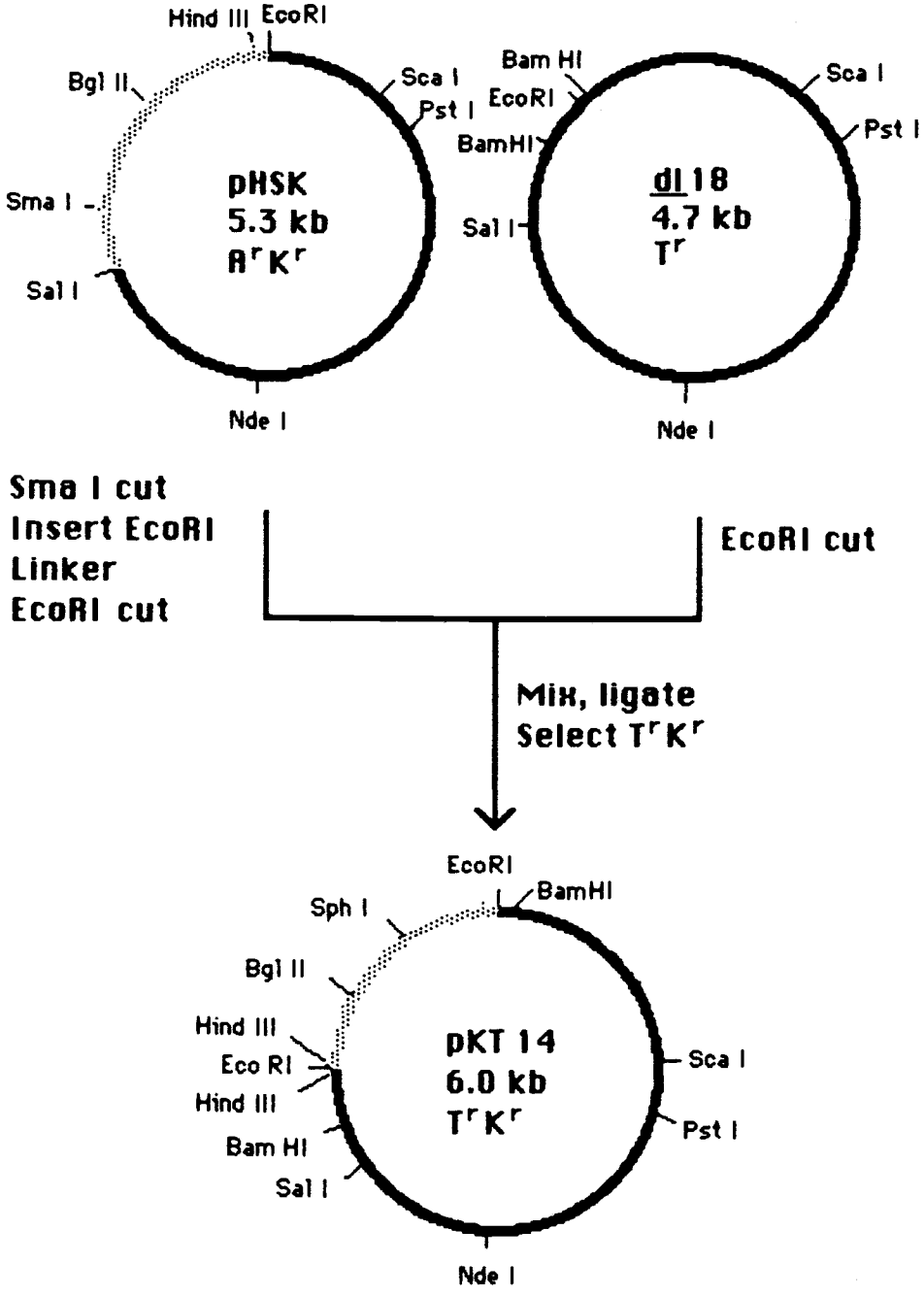


Figure 3: Construction of pKT14

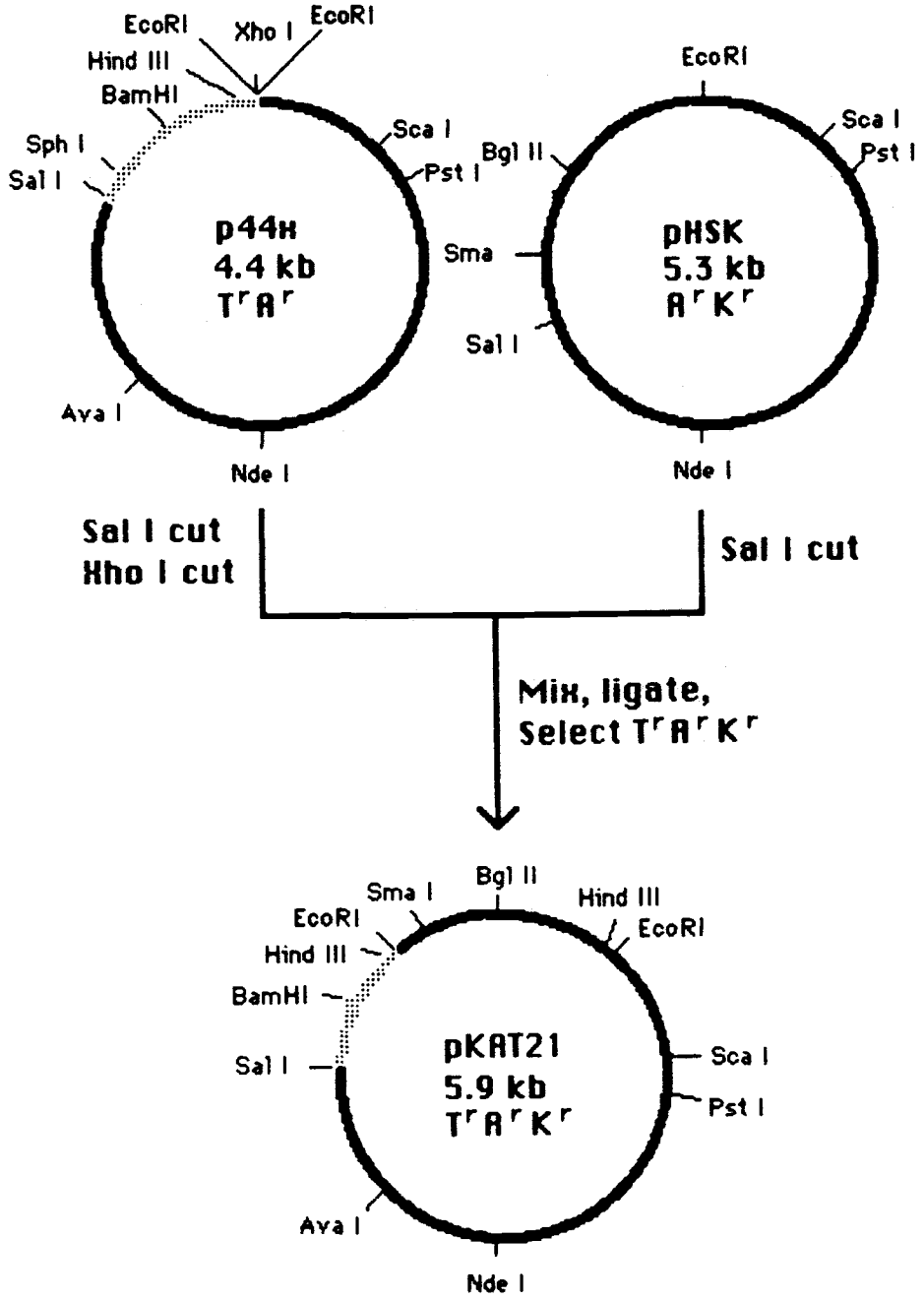


Figure 4: Construction of pKAT21

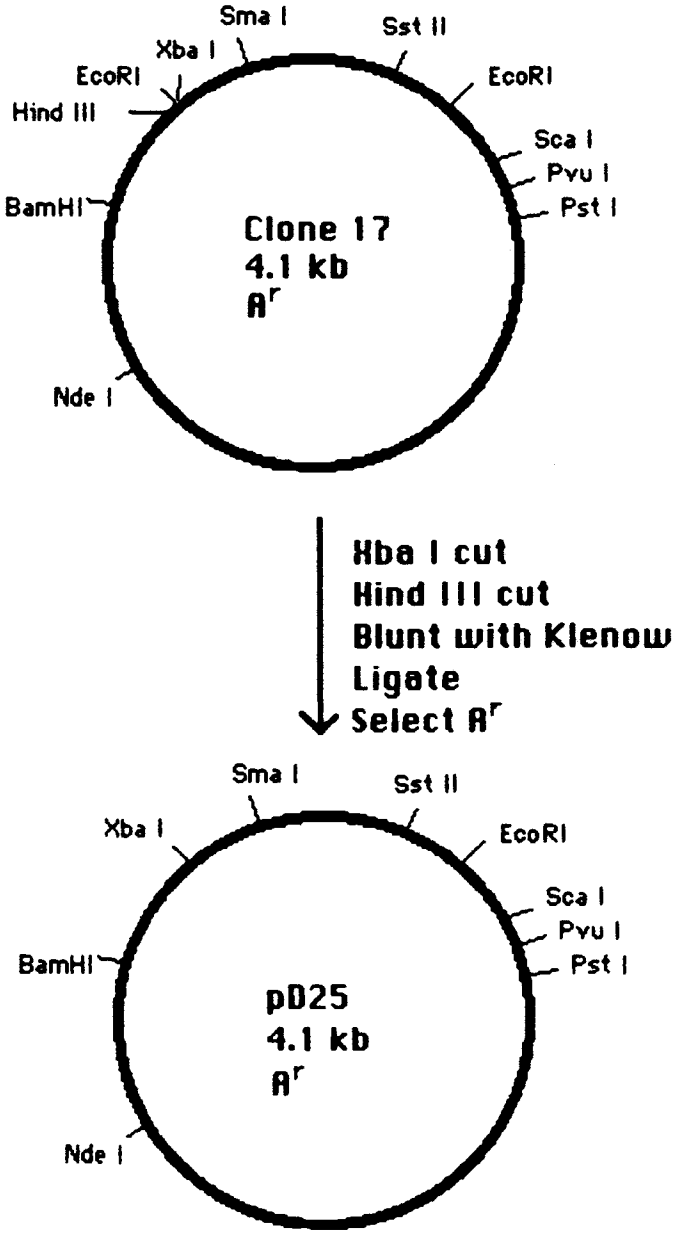


Figure 5: Construction of pD25

(Figure 6) is a derivative of pD25 made by cutting pD25 at the unique EcoRI and SstII sites in the molecule and blunting. pBP34 (Figure 7) is a derivative of pXD6 made by cutting at the unique BalI and PvuII sites in the molecule and ligating. clone 342 (Figure 8) is a derivative of pBP34 made by cutting at the unique ScaI site in the molecule and ligating in the presence of the blunted, gel purified, EcoRI fragment of pKAT21 that contains the kanamycin-resistance gene. pADIRK2 (Figure 9) is a derivative of pADIR made by ligating the EcoRI fragment of pKT14 containing the kanamycin-resistance gene into the EcoRI site of pADIR.

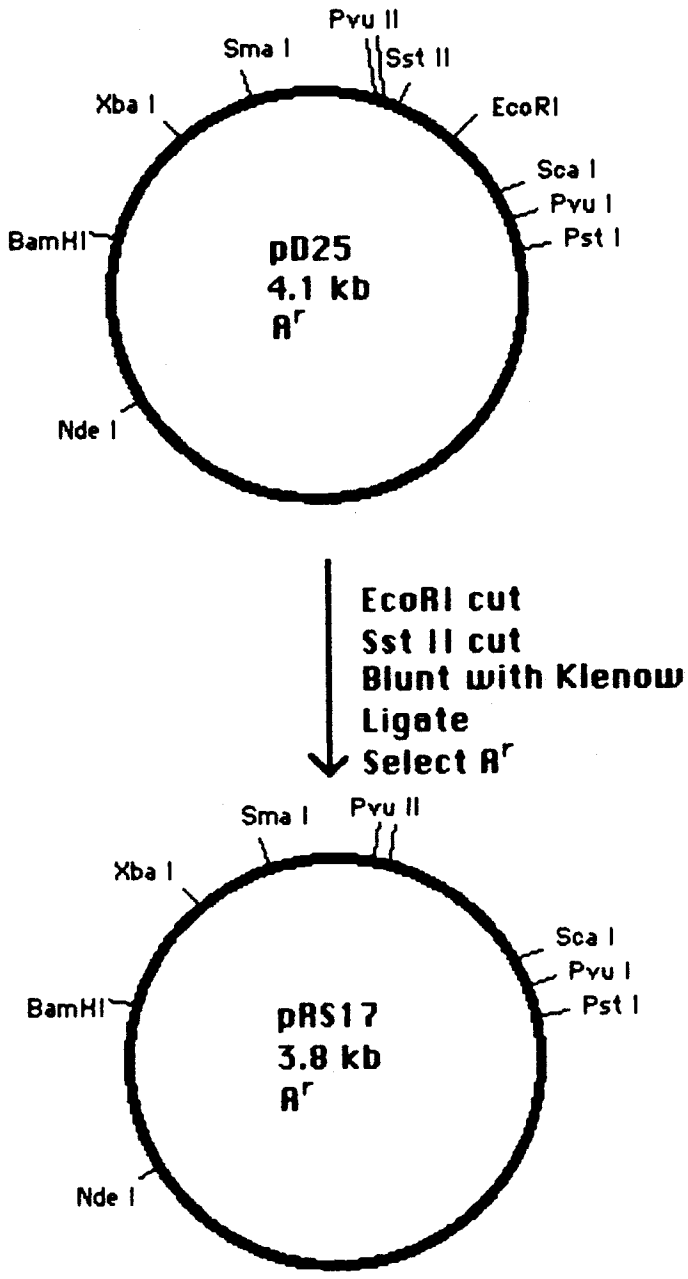


Figure 6: Construction of pRS17

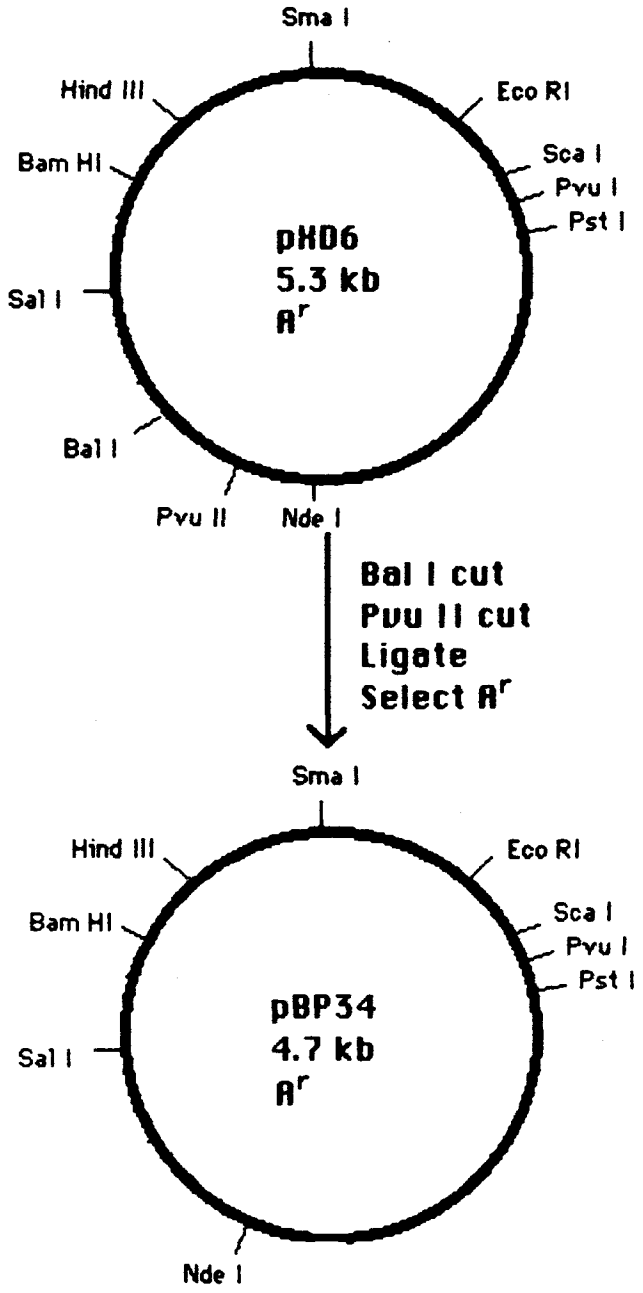


Figure 7: Construction of pBP34

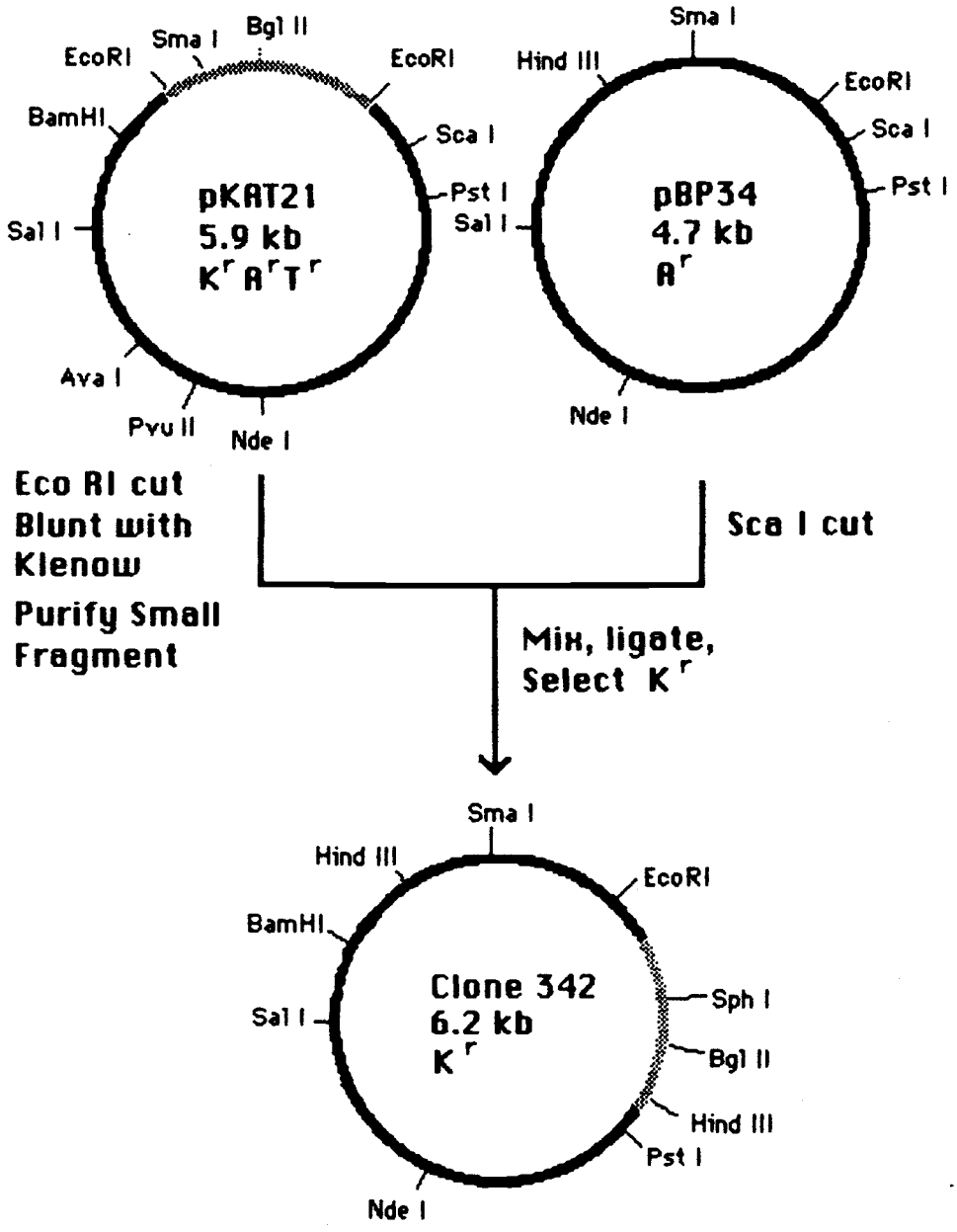


Figure 8: Construction of clone 342

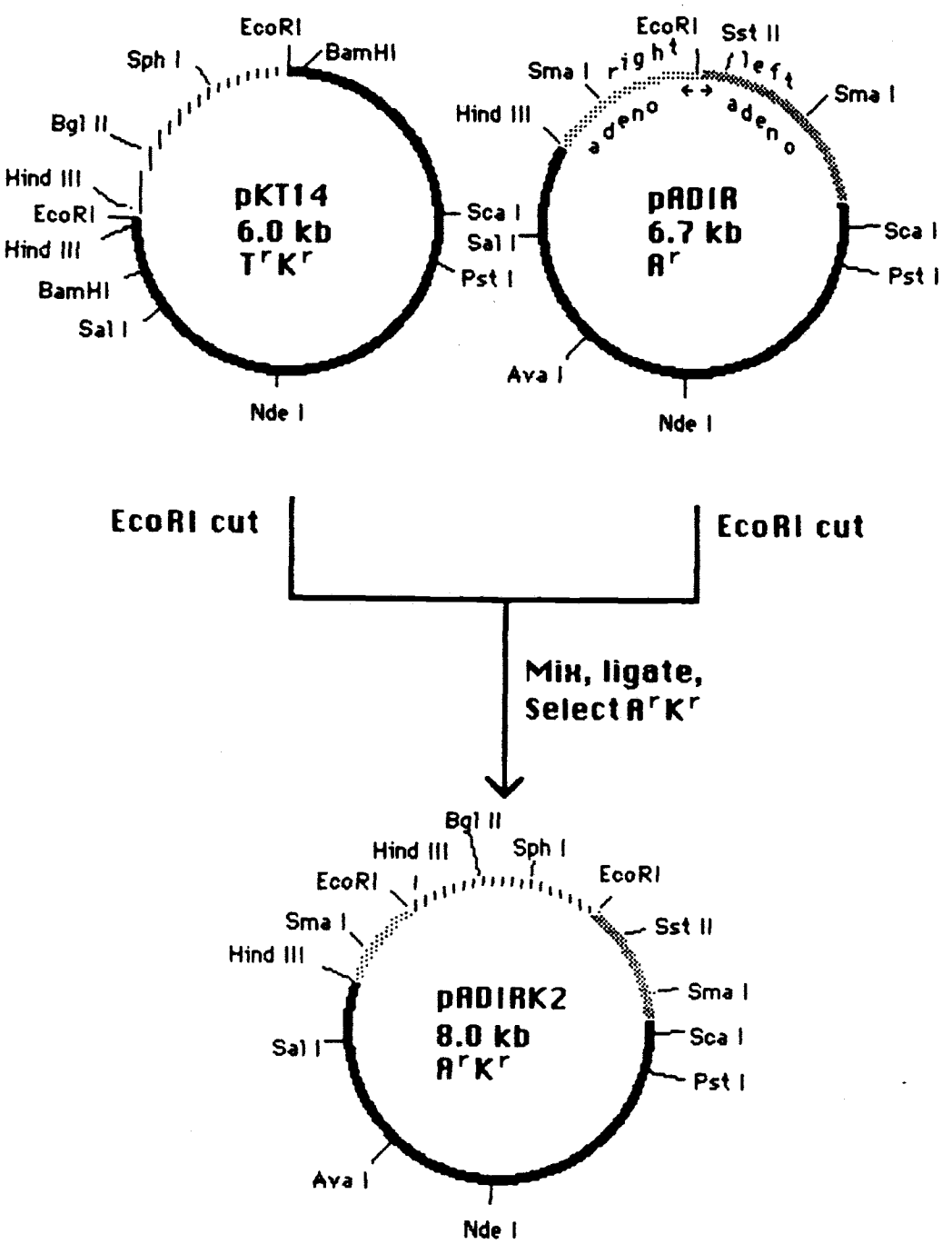


Figure 9: Construction of pADIRK2

determination is beyond the scope of this work.

CHAPTER 3: A NOVEL RECOMBINATIONAL ACTIVITY IN ADENOVIRUS INFECTED CELLS

A major obstacle to studying complementary strand replication is the difficulty in distinguishing complementary strand replication from displacement strand replication. The final product formed, double-stranded adenovirus DNA, depends upon both aspects of replication. To date no *in vitro* system has been described which gives clear insight into complementary strand replication. *Cis* replication is widely accepted as a mechanism in complementary replication, but there are at least two alternatives to *cis* replication that should be considered. One is that complementary replication is initiated on completely single-stranded displaced adenovirus strands in the absence of a panhandle. There are several observations in the literature that are relevant. A wide variety of adenovirus and non-adenovirus single-stranded DNA molecules support formation of pTP-dCMP complexes *in vitro*. These include denatured adenovirus DNA, denatured plasmids containing adenovirus terminal sequences, single-stranded ϕX174, M13, f1, or G4 DNAs, and even synthetic oligonucleotides containing the 15 terminal nucleotides at the 3' end of adenovirus DNA (136,157). One distinction about the utilization of single-stranded templates versus double-stranded templates by pTP is that pTP does not appear to require the presence of a specific nucleotide sequence for priming activity on single-stranded templates (59). As a result, it is not clear whether the phenomenon is relevant to adenovirus replication or not. Such a determination is beyond the scope of this work.

Another alternative to *cis* replication is a mechanism acting in *trans* to combine together single strands displaced during replication. Such a mechanism might be as simple as base pairing of complementary sequences or a more complicated scheme involving recombinational pathways. Previous work has demonstrated that recombination or a recombination-like process can link separated adenovirus origins into a single adenovirus molecule (40,121). Study of processes that appear to operate in *trans* may give insight to both replication and recombination during adenovirus infection. A unique strategy for studying these intracellular events is outlined below.

The complete cycle of adenovirus replication requires that origin sequences be at each end of the molecule. Molecules with only one origin at a terminus are capable of displacement replication, but in the absence of a mechanism to put an origin on the end of a molecule lacking one, the displaced strand cannot remain in the pool of replicating molecules and is lost (120). Several laboratories have employed a strategy *in vivo* whereby a missing origin is replaced by a postulated intramolecular pairing between self-complementary (inverted) sequences of the displaced strand followed by DNA repair to regenerate the missing origin (120,123,124). These repaired molecules are then fully functional for replication. When plasmid constructs were used, replacement of the missing origin was detected by the appearance of a molecule larger than the input plasmid (123,124).

A derivative of this idea is employed here. Left- and right-end adenovirus origins were linked to plasmid sequences. In contrast to the studies of *cis* replication, no self-complementary sequences existed within the plasmids. However, each left or right origin

sequence was in an inverted orientation relative to flanking pBR322 sequences in each molecule (see Figure 10). Events that join origin sequences will have to function in *trans* and can be detected by appearance of a new size of molecule (Figure 11). Requirements for replication in *trans* can be determined by altering the structures of the molecules used and assaying the effects on production of new molecules.

STRATEGY

Initial experiments were performed with plasmid molecules containing an adenovirus left-end origin (clone 7), an adenovirus right-end, origin (pXD6, figure 7), or both origins (pADIRK2, figure 9). Each adenovirus origin terminated at an *Eco*RI site so the adenovirus sequences necessary for replication could be exposed by cleavage with *Eco*RI. Adenovirus-directed strand displacement of *Eco*RI-cleaved clone 7 and pXD6 produce single stranded molecules terminating at the 5' end with exposed adenovirus origins and at the 3' end with complementary pBR322 sequences except for a 20-30 base stretch of non-complementary sequences at the 3' end of each. These molecules cannot proceed through the replication cycle in *cis* due to their inability to form a panhandle, and are dependent on *trans* acting events to bring the two ends together for completion of the replication cycle. *Eco*RI-cut pADIRK2 on the other hand is competent for both displacement replication and complementary replication due to the presence of 2 origins in inverted orientations.

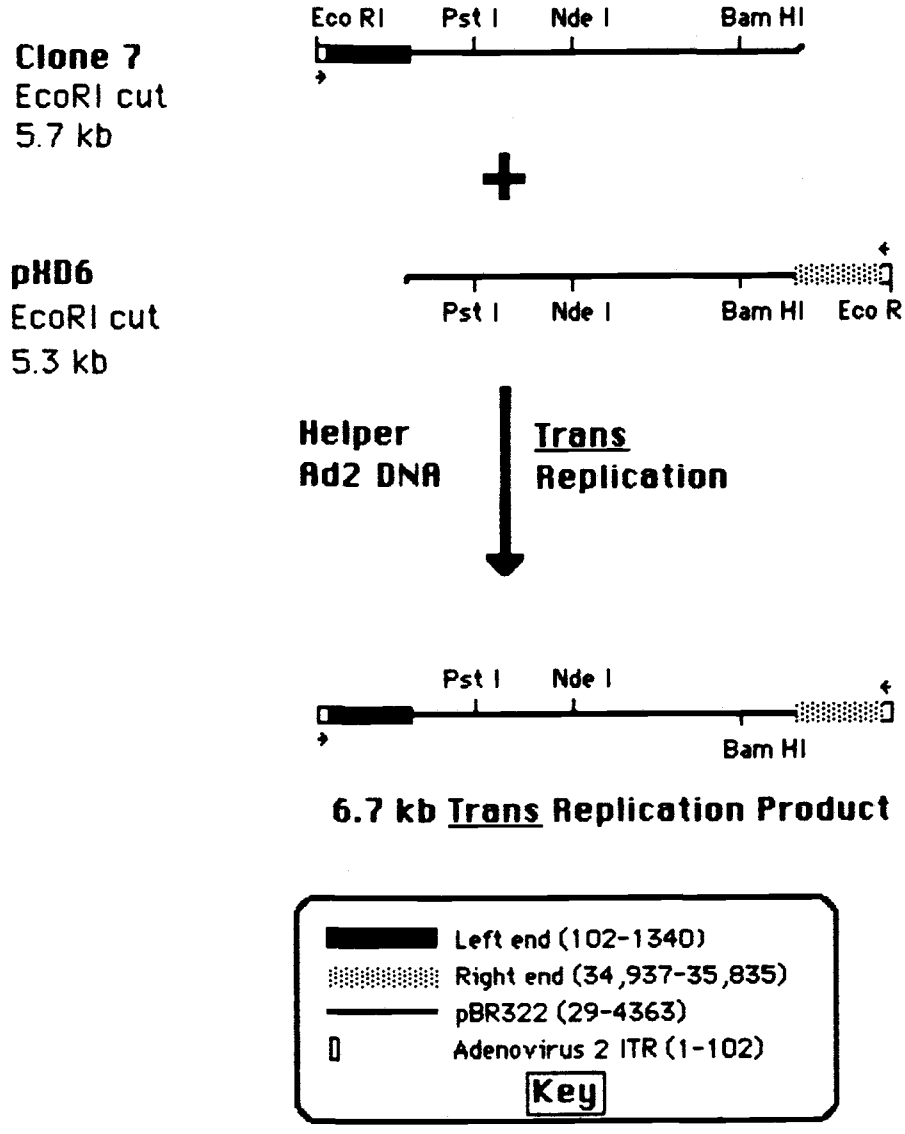


Figure 10: Summary of results with EcoRI-cut clone 7 and EcoRI-cut pXD6 in the presence of helper Ad2 DNA.

Figure 11: Southern Blot Analysis for Trans Replication. All lanes had EcoRI-cut pBR322 DNA. Presence or absence of helper Ad2 DNA is indicated at the bottom. DNAs used in each transfection as follows. Lane 1: EcoRI-cut clone 7, EcoRI-cut pXD6; Lane 2: EcoRI-cut clone 7, EcoRI-cut pXD6; Lane 3: BamHI-cut clone 7, EcoRI-cut pXD6; Lane 4: EcoRI-cut clone 7, PstI-cut pXD6; Lane 5: EcoRI-cut clone 7, PvuII-cut pRS17; Lane 6: Bgl11-cut pADIRK2. Molecular sizes in kilobase pairs (kb) are indicated between lanes 5 and 6.

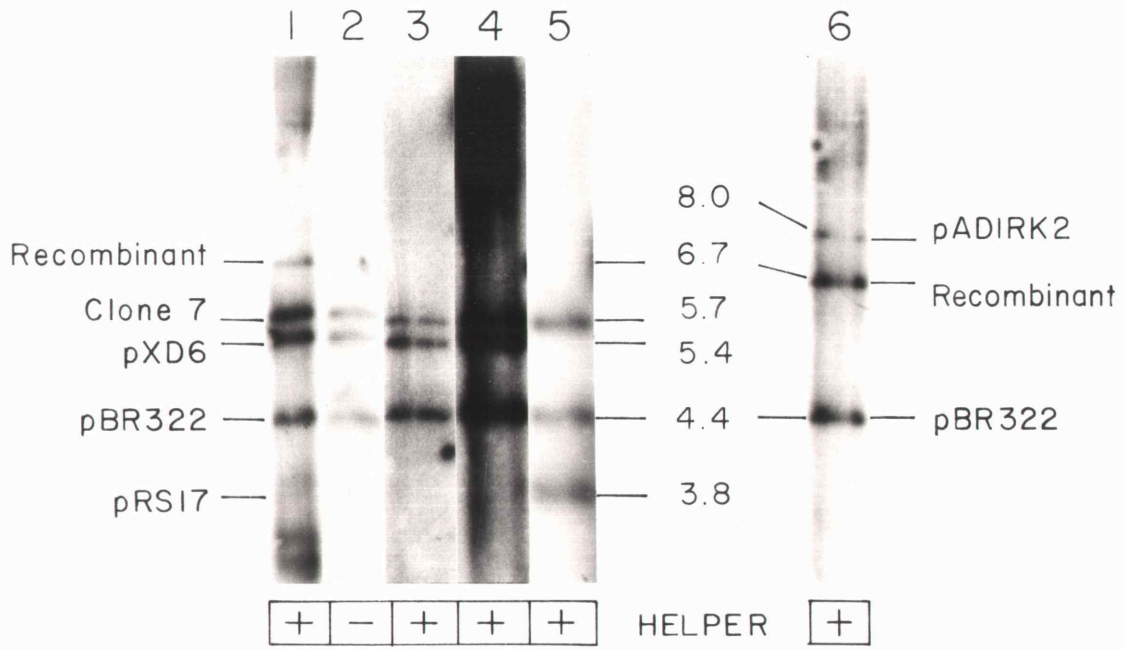

1 2 3 4 5
6
Recombinant
Clone 7
pXD6
pBR322
pRS17
8.0
6.7
5.7
5.4
4.4
3.8
pADIRK2
Recombinant
pBR322
+ − + + +
HELPER
+

Figure 11:

DETECTION OF TRANS REPLICATION

293 cells were co-transfected with Ad2 DNA, EcoRI-cleaved clone 7, and EcoRI-cleaved pXD6, then assayed. Replication in trans should generate a molecule 6.7 kb in size (Figure 10).

Extracts were analysed by Southern blot analysis (215) utilizing probes complementary to the pBR322 sequences in the plasmids. Results from this experiment revealed a prominent band at 6.7 kb (Figure 11, lane 1). When the same plasmid molecules were transfected without Ad2 DNA, no 6.7 kb molecule was detected (Figure 11, lane 2). These results indicated that appearance of the 6.7 kb molecule was dependent on functions provided by the helper Ad2 DNA. A likely function was expected to be replication, since replication could greatly increase copy number of the replication competent 6.7 kb molecules. An experiment was performed to determine whether the 6.7 kb molecules were replicating.

Molecules replicating by an adenovirus-driven mechanism should contain pTP covalently attached to their 5' ends like adenovirus DNA. On the other hand, molecules which recombine but do not replicate should not contain pTP. Since standard preparation of the extracts involves Pronase treatment that could destroy pTPs attached to replicating molecules, it was of interest to examine the electrophoretic mobility of extract molecules in the absence of Pronase treatment. A covalent pTP-DNA complex will not enter a native agarose gel. No band migrating at 6.7 kb was observed in the absence of pronase treatment, presumably due to the presence of pTP at each 5' end of each strand of the 6.7 kb DNA molecule (Figure 12, lane 8).

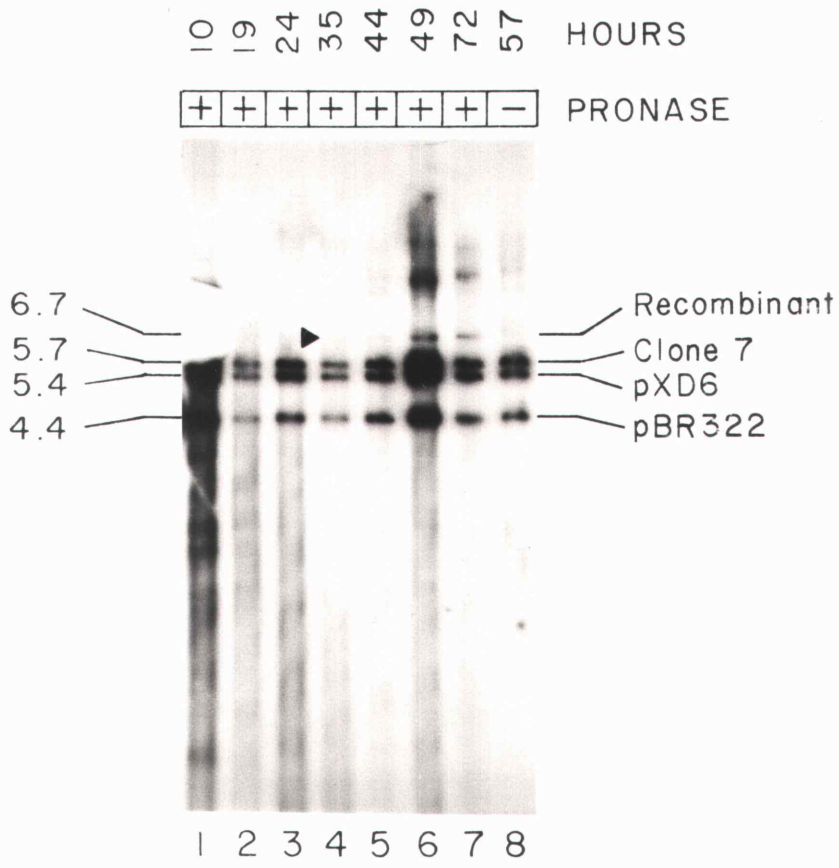


Figure 12: Trans replication assayed from 10 to 72 hours as indicated at the top of the Figure. Digestion with (+) or without (-) Pronase is also indicated. Each experiment utilized EcoRI-cut clone 7, EcoRI-cut pXD6, EcoRI-cut pBR322, and helper Ad2 DNA. The 6.7 kb trans recombinant is marked by a solid triangle. Sizes are given in kilobase pairs.

Thus, the 6.7 kb molecule appears to be replicating. A time course experiment detected the 6.7 kb band as early as 35 hours and its continued presence up to 72 hours post-transfection (Figure 12).

EcoRI-cut clone 7 and EcoRI-cut pXD6 are both competent for displacement replication and capable of homologous recombination. This meant that there were at least two possible mechanisms by which the 6.7 kb recombinant could have arisen. The first is that homologous recombination occurred as the first process followed by replication. The other is that displacement replication itself caused recombination to occur, followed by subsequent rounds of replication.

If displacement replication is required for *trans* replication, then the 6.7 kb *trans* recombinant should not be detected unless displacement replication-proficient molecules are used. On the other hand, if the *trans* replication process is the result of simple homologous recombination, then *trans* recombinants should be made regardless of whether displacement replication-proficient molecules are used. The only limitation on detection of *trans* recombinants would be if they replicate sufficiently. To attempt to distinguish between these two possibilities, experiments were performed in which one of the input plasmids was restricted so as to prevent displacement replication, yet allow homologous recombination to proceed to make replication competent molecules.

Plasmids with internally located adenovirus origins, such as BamHI-cut clone 7 or PstI-cut pXD6, undergo displacement replication very inefficiently at best in vitro (136). Experiments were performed with PstI-cut pXD6, EcoRI-cut clone 7, and Ad2 DNA (Figure 13) or BamHI-cut clone 7, EcoRI-cut pXD6, and Ad2 DNA (Figure 14). If

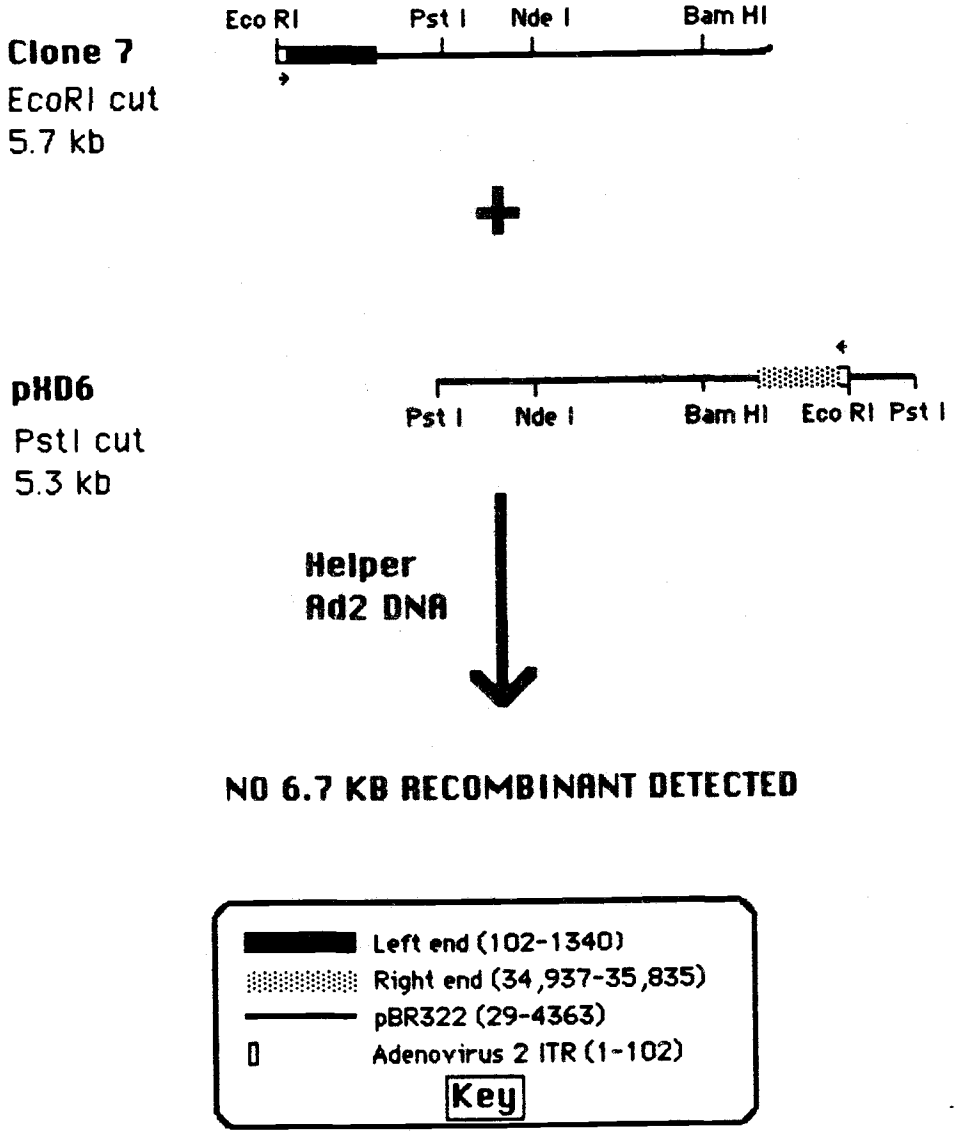
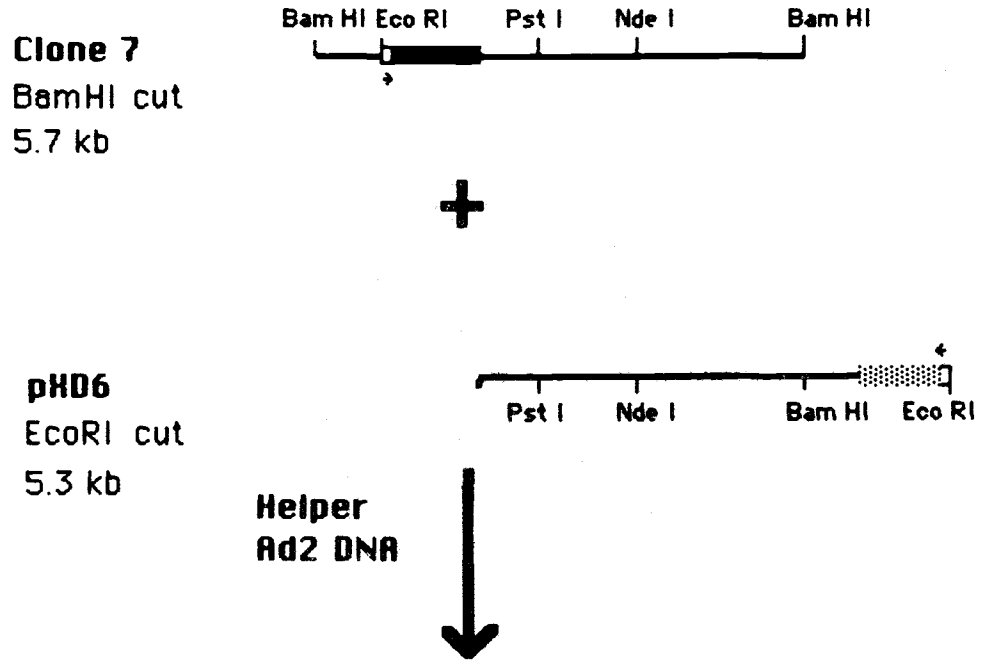


Figure 13: Summary of results with EcoRI-cut clone 7 (exposed origin) and PstI-cut pXD6 (embedded origin) in the presence of helper Ad2 DNA.



NO 6.7 KB RECOMBINANT DETECTED

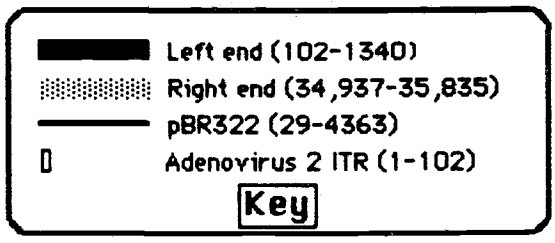


Figure 14: Summary of results with BamHI-cut clone 7 (embedded origin) and EcoRI-cut pXD6 (exposed origin) in the presence of helper Ad2 DNA.

homologous recombination were the first step in *trans* replication, primary recombinants 7.1 kb or 7.4 kb respectively would be formed (see Figure 15). Both of these primary recombinants have a dual origin structure that can be repaired in adenovirus-infected cells to make a secondary replicating recombinant of 6.7 kb with origins at each terminus (216). Therefore, if homologous recombination is the first step in the process of trans replication, a 6.7 kb recombinant should be made when either *Bam*HI-cut clone 7 or *Pst*I-cut pXD6 is transfected with a displacement replication partner and helper Ad2 DNA. A critical process relevant to this interpretation, however, is the efficiency with which the repair process occurs (see below).

Results of these experiments revealed that neither primary nor secondary recombinants were detected when either of the two origin-containing plasmids had a structure that prevented displacement replication *in vitro* (Figure 11 lanes 3 and 4). This suggests but does not prove that trans replication is a displacement replication-dependent process. An alternative interpretation is that homologous recombination between two displacement replication proficient molecules produces a molecule completely competent for continued replication (Figure 15A), whereas homologous recombination between two molecules, one proficient for displacement replication and one not (Figures 15B & 15C), produces a molecule which must be repaired before becoming fully replication-proficient. If the repair process were inefficient, then recombinants might not be detected for that reason alone. This alternative was not considered likely, in view of Nigel Stow's work (216) showing that adenovirus molecules with a structure like that of the homologous intermediates B&C in Figure 15

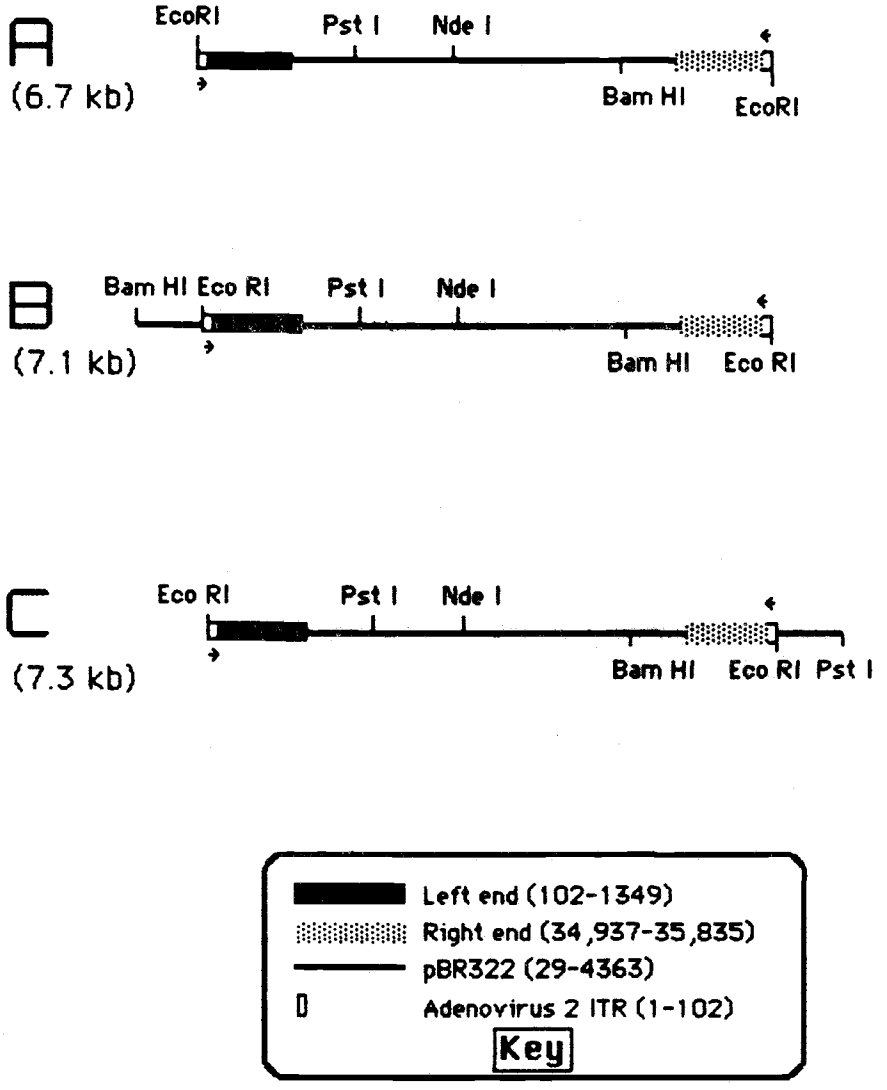


Figure 15: Postulated homologous recombinational intermediates when EcoRI-cut clone 7 and EcoRI-cut pXD6 (A), BamHI-cut clone 7 and EcoRI-cut pXD6 (B), or EcoRI-cut clone 7 and PstI-cut pXD6 (C) were transfected with helper Ad2 DNA.

are indeed readily repaired *in vivo*.

REPAIR OF ORIGIN SEQUENCES

In order to determine the efficiency of repair of disabled adenovirus origin sequences could be detected in this system, an experiment was performed which required repair of both origin sequences. *Bgl*II cut pADIRK2 was transfected into 293 cells with helper Ad2 DNA. pADIRK2 is an 8.0 kb plasmid that contains both the left and the right ends of adenovirus DNA in an inverted orientation separated by the 1.3-kb kanamycin-resistance gene. *Bgl*II cuts within the kanamycin gene (see Figure 16). Analysis of extracts from these cells revealed a 6.7 kb molecule, the size predicted if the adenovirus ITRs were released in linear form free of the 1.3 kb of sequence separating them (Figure 11). The intensity of the 6.7 kb molecule was greater than or equal to the 6.7 kb band produced by *trans* replication, demonstrating that removal of flanking sequences from origin sequences was not limiting in this system. In addition, these results confirm the observations of Hay *et al.* (124) who has described a similar process. A replication dependent mechanism to excise ITR sequences from molecules like pADIRK2 has been proposed (124).

ASSAY OF HOMOLOGOUS RECOMBINATION BY *CIS* REPAIR

Another approach to determine the role of homologous recombination in the trans replication process utilized *Eco*RI-cut clone 7 and *Pvu*II-cut pRS17 which were transfected into 293 cells in the presence

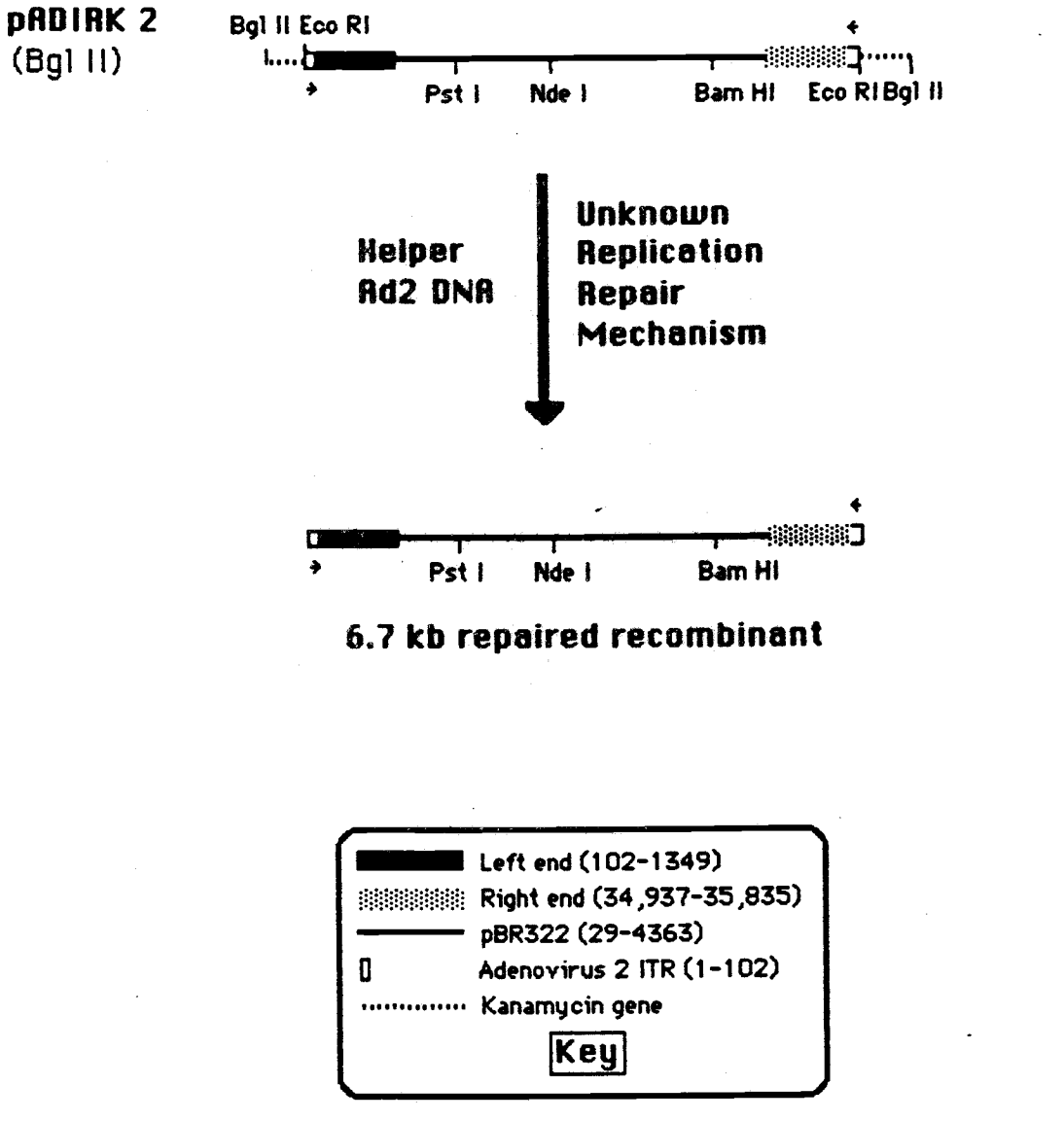


Figure 16: Generation of a 6.7 kb molecule from an 8.0 kb molecule with two embedded origins.

of helper Ad2 DNA. pRS17 contains adenovirus sequences complementary to clone 7 (but not the ITR) in an inverted orientation relative to the flanking pBR322 sequences (Figure 17). In addition, pRS17 is incapable of undergoing adenovirus-directed replication due to the lack of the adenovirus origin. Homologous recombination between EcoRI-cut clone 7 and PvuII-cut pRS17 should make a molecule, however, that is readily repairable by a *cis* replication mechanism (123,124). The primary recombination products between these molecules should be molecules 5.2 or 6.8 kb, depending on the location of the recombination. *Cis* repair of the primary recombinational products should make 5.4 and/or a 7.0 kb molecules depending upon the location of the initial recombinational event (see Figure 17). In these experiments, neither primary replication products (5.2 kb or 6.8 kb) nor secondary *cis* repaired products (5.4 kb or 7.0 kb) were detected (Figure 11 lane 5).

All of these data are consistent with the idea that simple homologous recombination is not the first process that occurs in production of the 6.7 kb recombinant. It is not conclusive, however, because replication-incompetent molecules do require an extra step to become replication-proficient, namely repair of the origins, that replication-competent molecules do not. To examine the role of homologous recombination in *trans* replication in a way that does not require repair of the origin sequence, a molecule marked by insertion and deletion of sequences was constructed from pXD6. Clone 342 lacks 600 bp between the unique PvuII and BalI sites of the pBR322 sequences and contains a 1.5-kb insertion of the kanamycin gene sequences in the Sca I site (see Figure 18).

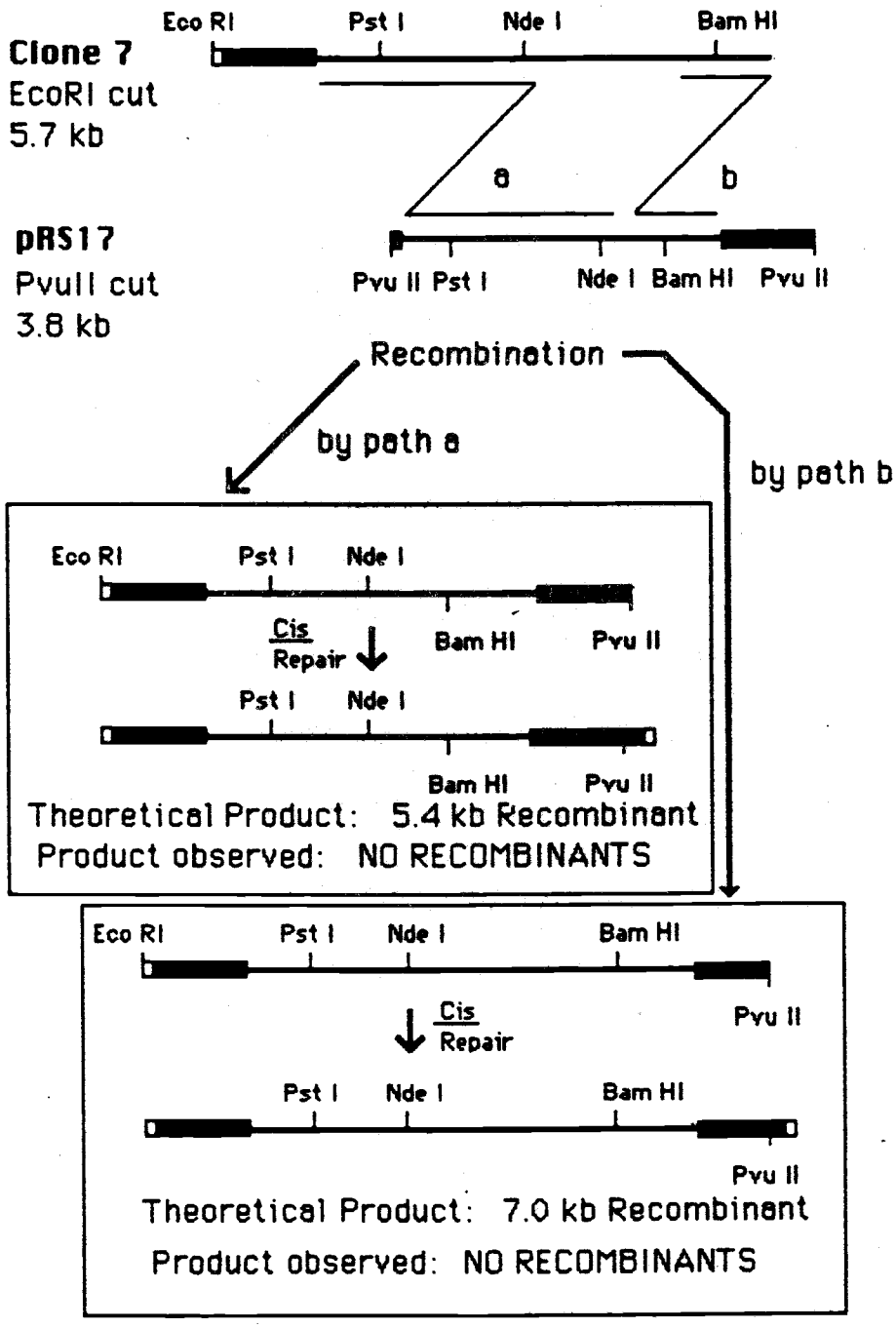


Figure 17: Postulated recombinational pathways for synthesis of a cis-repairable adenovirus origin when Eco-RI-cut clone 7 and PvuII-cut pRS17 were transfected with helper Ad2 DNA.

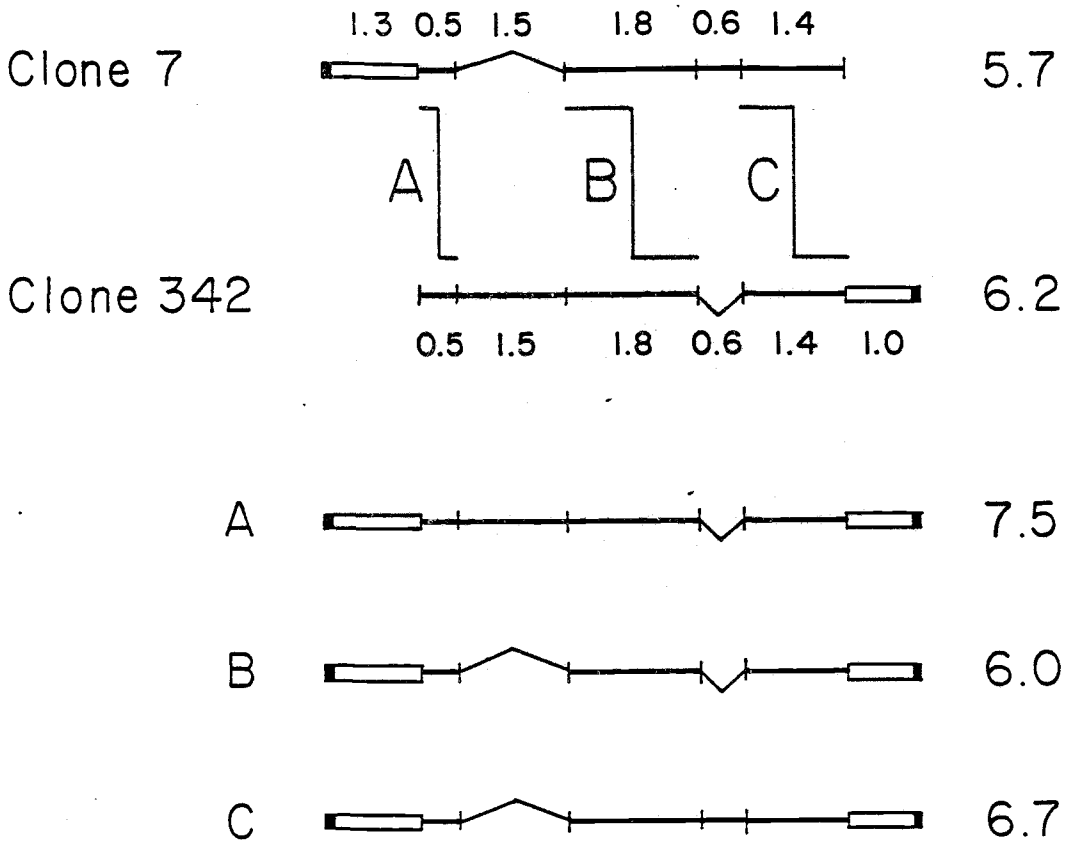


Figure 18: Postulated products of homologous recombination between EcoRI-cut clone 7 and EcoRI-cut clone 342.

Clone 342 and clone 7 were cleaved with EcoRI and transfected with Ad2 DNA into 293 cells. The strategy of this experiment is shown in Figure 18. If simple homologous recombination were the first step in the process of trans replication, then three recombinants of sizes 7.5 kb, 6.7 kb, and 6.0 kb should be produced due to crossovers occurring within the three regions of homology between clone 7 and clone 342. As shown in Figure 19, only the 7.5 kb and the 6.7 kb molecules were detected. Simple recombination involving crossover at homologous sequences is thus ruled out as the initial mechanism for trans replication.

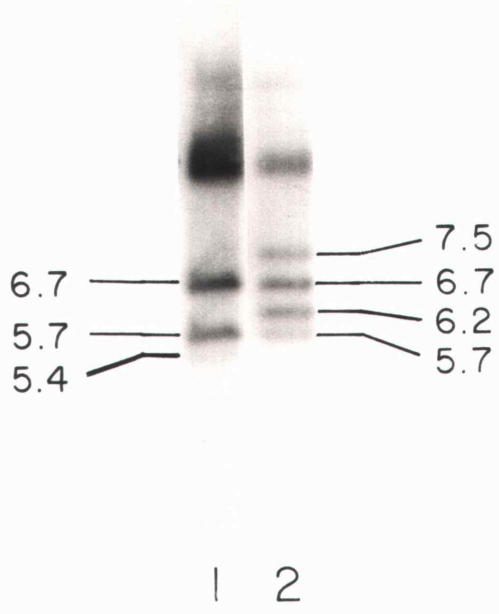


Figure 19: Results of *trans* replication with molecules marked with an insertion and a deletion. Helper Ad2 DNA was added to each transfection. Numbers indicate molecular sizes in kilobase pairs. Input DNAs as follows. Lane 1: *Eco*RI-cut clone 7, *Eco*RI-cut pXD6; Lane 2: *Eco*RI-cut clone 7, *Eco*RI-cut clone 342.

CHAPTER 4: CONCLUSIONS

The aim of this work was to elucidate mechanisms of adenovirus replication. Since adenovirus has a linear DNA molecule with origins at each end, removing the origin from one end of the molecule still enables the first phase of replication, displacement replication, but prevents the second phase, complementary replication, unless there is a means of regenerating the missing origin. Previous studies have suggested that *cis* replication is a legitimate means of regenerating a missing origin and that *cis* replication is likely to play a role in adenovirus replication (62,123,124). The results of this study showed that, in addition to *cis* replication, adenovirus origin-containing plasmids can utilize a process described as *trans* replication to complete the replication cycle.

SUMMARY OF RESULTS

Adenovirus left and right end origin sequences were localized separately in inverted orientations on respective pBR322 plasmids (clone 7 and pXD6). When both plasmids were cleaved with *Eco*RI to expose the origins and transfected into 293 cells with helper Ad2 DNA, a 6.7 kb recombinant predicted from a *trans* replication event was detected. Results from a gel retardation assay suggested that the 6.7 kb DNA was like replicating adenovirus DNA in that it appeared to be attached to a protein. A time course experiment detected the presence of the 6.7 kb recombinant as early as 35 hours and as late as 72 hours after transfection.

All of the data obtained suggested that displacement replication of both the input plasmids was required for production of the 6.7 kb recombinant. No 6.7 kb recombinant was detected when the same experiments were performed in the absence of helper Ad2 DNA or when one of the pairs of transfected plasmids (BamHI-cut clone 7 or PstI-cut pXD6) was present in a configuration that prevented displacement replication in vitro. Removal of flanking plasmid sequences from adenovirus origin sequences did not appear to be a limiting factor in the tests for recombination, as the 8.0 kb BglII-cut pADIRK2 plasmid with two embedded adenovirus origins efficiently produced a 6.7 kb recombinant when transfected with helper Ad2 DNA.

In addition, attempts to isolate a cis-repairable plasmid by recombination between a replication proficient molecule (EcoRI-cut clone 7) and an origin lacking plasmid (PvuII-cut pRS17) failed to produce any detectable recombinants. Finally origin-containing plasmids marked by deletion and insertion (clone 7 and clone 342) produced only two of the three recombinants expected by general recombination.

POSSIBLE INTERPRETATIONS

The dual ability of each of the initial plasmids (Eco-RI-cut clone 7 and pXD6) to replicate and recombine suggests two possible mechanism(s) for production of the 6.7 kb recombinant. One possibility is that simple homologous recombination between the plasmid sequences united the spatially separated adenovirus origins

into a single recombinant molecule. This molecule would be competent for both displacement and complementary replication. Amplification of the recombinant via the complete replication cycle could produce adequate amounts of the 6.7 kb recombinant to be detected by blotting, while in the absence of replication none would be detected, even though the same process would be occurring. This process will be referred to as the Recombination Amplification (RA) process.

An alternative explanation is that the actual process involved reversal of recombination and replication; that is, a recombination process was detected that was dependent upon replication of both of the input plasmids. In the absence of replication of both of the input plasmids, no recombinant was produced. This process will be referred to as the Replication Dependent Recombination (RDR) process.

KNOWN RECOMBINATION MECHANISMS OF ADENOVIRUS

Adenovirus has been shown to recombine in a variety of studies (40-44,121,217), but little has been done to characterize the mechanism(s) involved. The methods to assay recombination used in previous experiments differ from the methods employed in this work in that they are selective in nature, so it is not easy to correlate other conclusions with this work. Nevertheless, interpretation of the experimental evidence obtained to date does not rule out either the RA or the RDR process as a mechanism for recombination. Wolgemuth and Hsu (217) found structures in replicating Ad2 DNA molecules late in infection of HeLa cells that resembled the intermediates in recombination postulated by Meselson and Radding (218, see below).

Intriguingly, their data shows that the single-stranded branch in recombination is generated from one end of Ad2 during replication, elimating the need for an endonuclease nicking step proposed by Meselson and Radding. Whether such a means of recombination could function in replication in *trans* to put displaced single strands together has never been established.

Berkner and Sharp (40) cloned overlapping sections of the Ad5 genome into plasmid molecules where the respective termini were separated in *trans*. Transfection of these molecules into 293 cells resulted in production of full-length infectious virus DNA, albeit at an efficiency 100-fold lower than with intact adenoviral DNA. The only requirement for the production of infectious virus was that at least one of the ends had to be linearized to expose the adenovirus terminus. Of course, the fragments of adenovirus sequence in the respective plasmids had to overlap. Similarly, Chinnadurai et al. (121) transfected overlapping DNA restriction fragments of adenovirus serotypes 2 and 5 into 293 cells and recovered functional recombinant viruses joined at the junction of the overlap. They observed that the yield of recombinant virus obtained was directly related to the extent of overlap between sequences. Production of the recombinant virus was attributed to homologous recombination, but other possibilities, such as RDR were neither considered, nor ruled out. In order to consider fully the possible recombinational mechanism(s) involved in *trans* replication, it is instructive to examine the process of recombination in general.

MODELS FOR RECOMBINATION

Most models for homologous recombination predict a pairing of complementary sequences at some point in the process. The Meselson and Radding model (218, see Figure 20) postulates a nick in one strand followed by polymerization from the 3' end of the nick, generating a displaced single strand (Figure 20a). The displaced single strand is then postulated to invade a homologous sequence, forming a D-loop structure (Figure 20b) which is subsequently degraded, whereupon the invading strand is ligated into place (Figure 20c). The limited region of asymmetric heteroduplex DNA is expanded by DNA synthesis on the donor duplex and degradation on the recipient duplex (Figure 20c), ultimately forming a Holliday-like structure (219,220). Then resolution, after branch migration (Figure 20e), results in the crossover (Figure 20f) or noncrossover (Figure 20g) configuration.

The double-strand-break model for recombination (221) has been proposed to explain the stimulation of recombination by a double strand break in one DNA molecule relative to homologous sequences in another. In this model (Figure 21), both strands of one duplex are cut and an exonuclease makes a gap flanked by protruding 3' ends (Figure 21a). One 3' end invades a homologous duplex, forming a D-loop (Figure 21b). Polymerization from the 3' end of the invading strand enlarges the D-loop until the boundary of the other end of the linear DNA is reached (Figure 21c), where complementary sequences anneal. Repair synthesis from the 3' end of the newly annealed sequences results in a molecule with double Holliday junctions (Figure 21d). Subsequent resolution leads to non-crossover

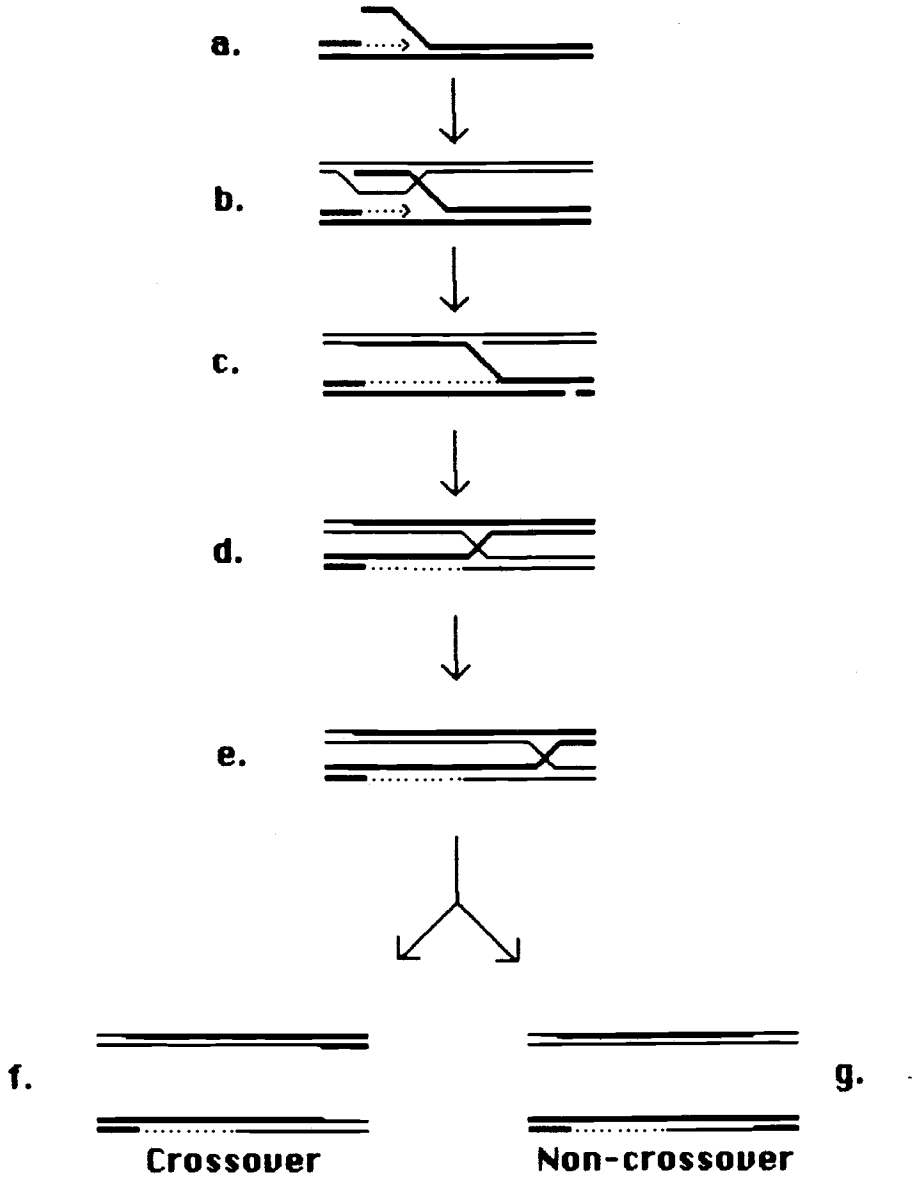


Figure 20: The Meselson-Radding model for recombination. See text for details. Adapted from Meselson and Radding (1975).

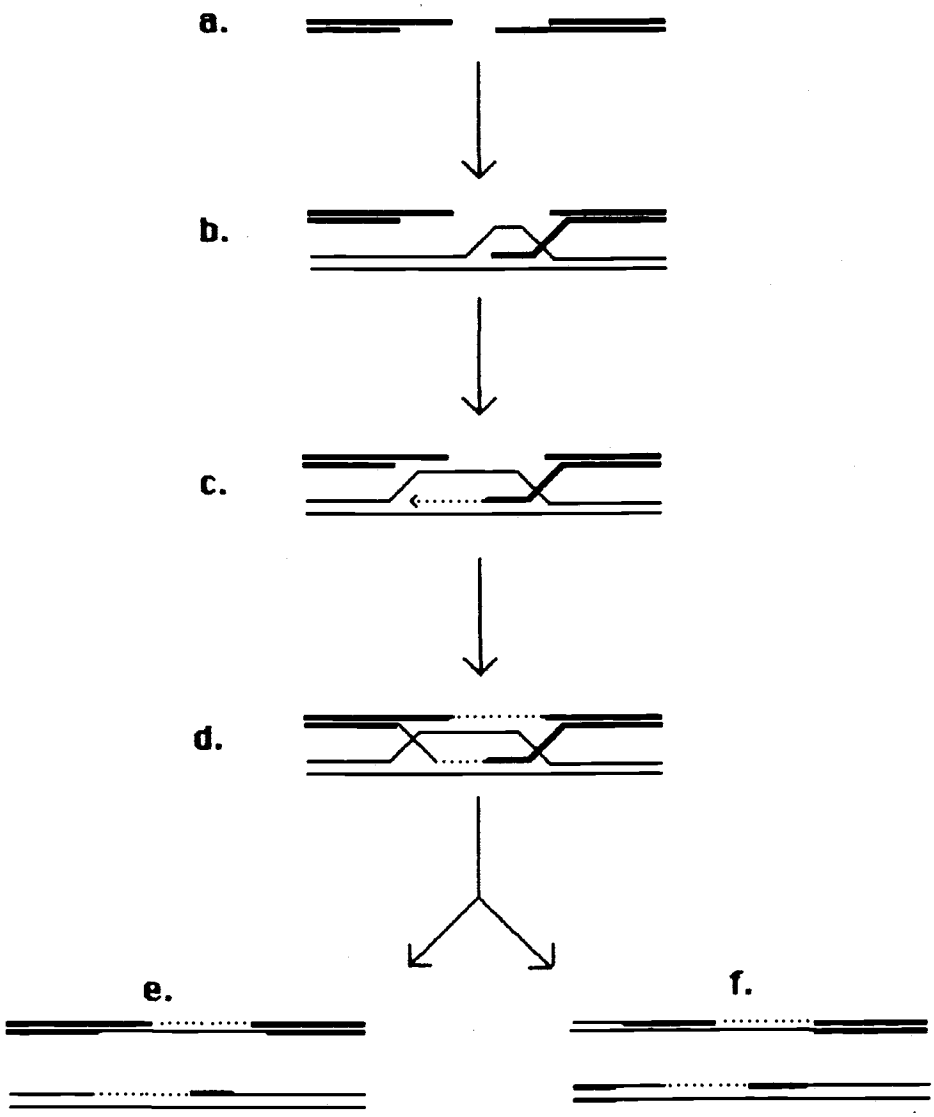


Figure 21: Double-strand-break repair model of recombination. See text for details. Adapted from Szostak et al. (1983).

(Figure 21e) or crossover (Figure 21f) configurations.

Both of these models for recombination involve strand invasion where a single strand invades a duplex region, creating a D-loop. Fwu-Lai Lin (222) has proposed an alternative model for homologous recombination in eukaryotic cells that relies strictly on pairing of complementary sequences during the recombination process, but does not require strand invasion. It represents an alternative mechanism for the double-strand-break model. His model is summarized in Figure 22. It postulates a double strand break at or near regions of homology (Figure 22a) followed by a 5' exonuclease to expose complementary sequences in each molecule (Figure 22b). These sequences pair through regions of complementarity (Figure 22c) and an endonuclease removes protruding 3' sequences (Figure 22d). Finally gap repair completes the process. Lin's model also differs from the others in that it does not involve a reciprocal exchange of sequences during the recombination (222).

INTERPRETATION OF OBSERVED PHENOMENA

It is not surprising that active recombination (trans replication) was detected in adenovirus-infected cells. The linear structure of adenovirus and its displacement mode of replication result in production of single-stranded molecules, a prerequisite intermediate in all of the models of recombination above. The recombination observed in trans replication, however, did not appear to be consistent with the recombination models described above. First, the recombination is likely not simply due to single

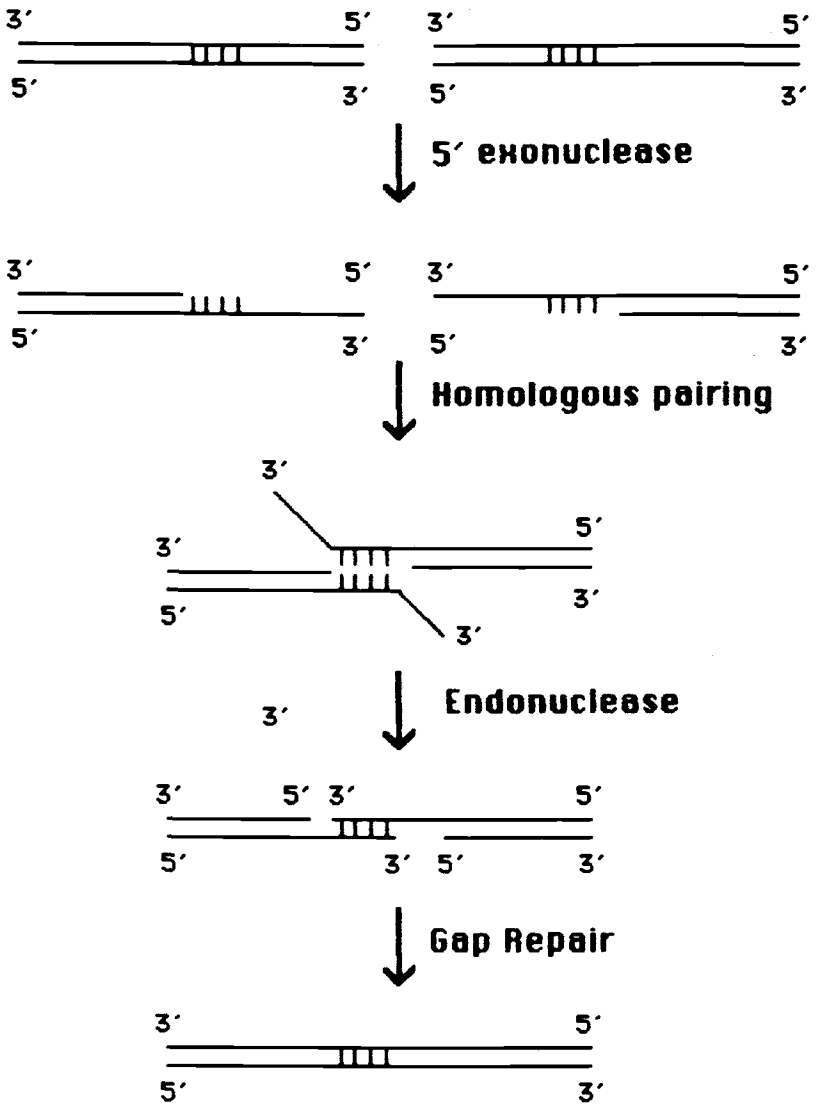


Figure 22: Lin model for homologous recombination. See text for details. Adapted from Lin et al. (1984).

strand invasion. Adenovirus displacement replication produces plentiful single-stranded molecules. If these were "hot" for recombination via strand invasion, then displacement replication should be required of one of the pair of input plasmids. The single strands produced by displacement replication should readily invade the double-stranded non-replicating plasmid. In fact, when either of the input plasmids was replication incompetent, *trans* replication was not detected.

Second, *trans* replication was not simply due to recombination stimulated by the presence of double-stranded breaks, since all of the molecules contained double-stranded breaks, yet *trans* replication was detected only when both molecules were capable of displacement replication. Finally, the first step in *trans* replication did not seem to be due to simple homologous recombination between related sequences. Homologous recombination between a replication-competent molecule (*Eco*RI-cut clone 7 or pXD6) and a molecule disabled for displacement replication (*Pst*I-cut pXD6 or *Bam*HI-cut clone 7) should make a 6.7 molecule that is able to replicate fully. No such molecule was detected in spite of the fact that excision of flanking plasmid sequences from adenovirus origin sequences appeared to be an efficient process in this system.

In addition, the *cis* repair mechanism of Hay *et al.* (124) and Wang and Pearson (123) is an efficient means of replacing a missing adenovirus origin from molecules with self-complementarity. In the presence of helper Ad2 DNA, recombinants between *Eco*RI-cut clone 7 and *Pvu*II-cut pRS17 should have a *cis*-repairable configuration. No trace was detected, however, of either recombinant molecules or resulting

cis repaired products.

Lastly, by using homologous molecules marked by insertion and deletion, only two of the three recombinants expected for recombination between homologous sequences were detected. Therefore, appearance of the 6.7 kb molecule does not appear to be due to simple homologous recombination followed by amplification by replication (an RA process). By contrast, an RDR process is consistent with all of the data. *Trans* replication simply requires replication of both input plasmids.

POSTULATED MECHANISM OF *TRANS* REPLICATION

One intriguing mechanism consistent with these data is that *trans* replication is mediated by base pairing between the complementary pBR322 sequences in the displaced single strands. Displacement strand replication of *Eco*RI-cut clone 7 and *Eco*RI-cut pXD6 produces single-stranded left and right adenovirus termini at the 5' end of each molecule and complementary pBR322 sequences at the 3' end, except for 20-30 bases of non-homology at the far 3' end. Complementary sequences could simply base pair to form an intermediate with double-stranded internal pBR322 sequences and flanking single-stranded adenovirus left- and right-end sequences. A paired molecule could continue replication by removal of the non-homologous sequences by DNA polymerase or an exonuclease followed by filling-in of ends. The process of filling-in of the ends has already been illustrated in the cis-repair process (123,124). By this scheme, *trans* replication would not be detected when a replication incompetent plasmid was used since

it could not produce single strands by displacement replication.

In addition, the experiment between EcoRI-cut clone 342 and EcoRI-cut clone 7 can be explained by pairing of the displaced strands to make the intermediate shown in Figure 23. Subsequent replication from each end would make two new molecules of 6.7 kb and 7.5 kb, the exact sizes predicted by the mechanism proposed for *trans* replication (Figure 23). Further, the absence of the 6.0 kb band predicted for homologous recombination is also completely consistent with the pairing hypothesis as the initial step in *trans* replication and inconsistent with homologous recombination as the initial step.

A NEW MODE OF REPLICATION

From the early work of Meselson and Stahl (223) until the present, DNA replication in every biologycal system examined has been described as semi-conservative. Each strand of the parental duplex molecule is used as a template to synthesize two daughter duplex molecules, each containing a parental strand and a newly synthesized strand. Parental strands may be modified to distinguish them from the newly replicated strands so that repair can be directed against errors in the newly replicated strand, increasing the fidelity of replication. This is a major advantage for semi-conservative replication. If adenovirus utilizes a *trans* pairing mechanism in its replication cycle, then a new mode of replication is realized (Figure 24). It will be referred to as semi-promiscuous replication (in contrast to the structly semi-conservative mode associated with *cis* replication). Semi-promiscuous replication is semi-conservative for displacement

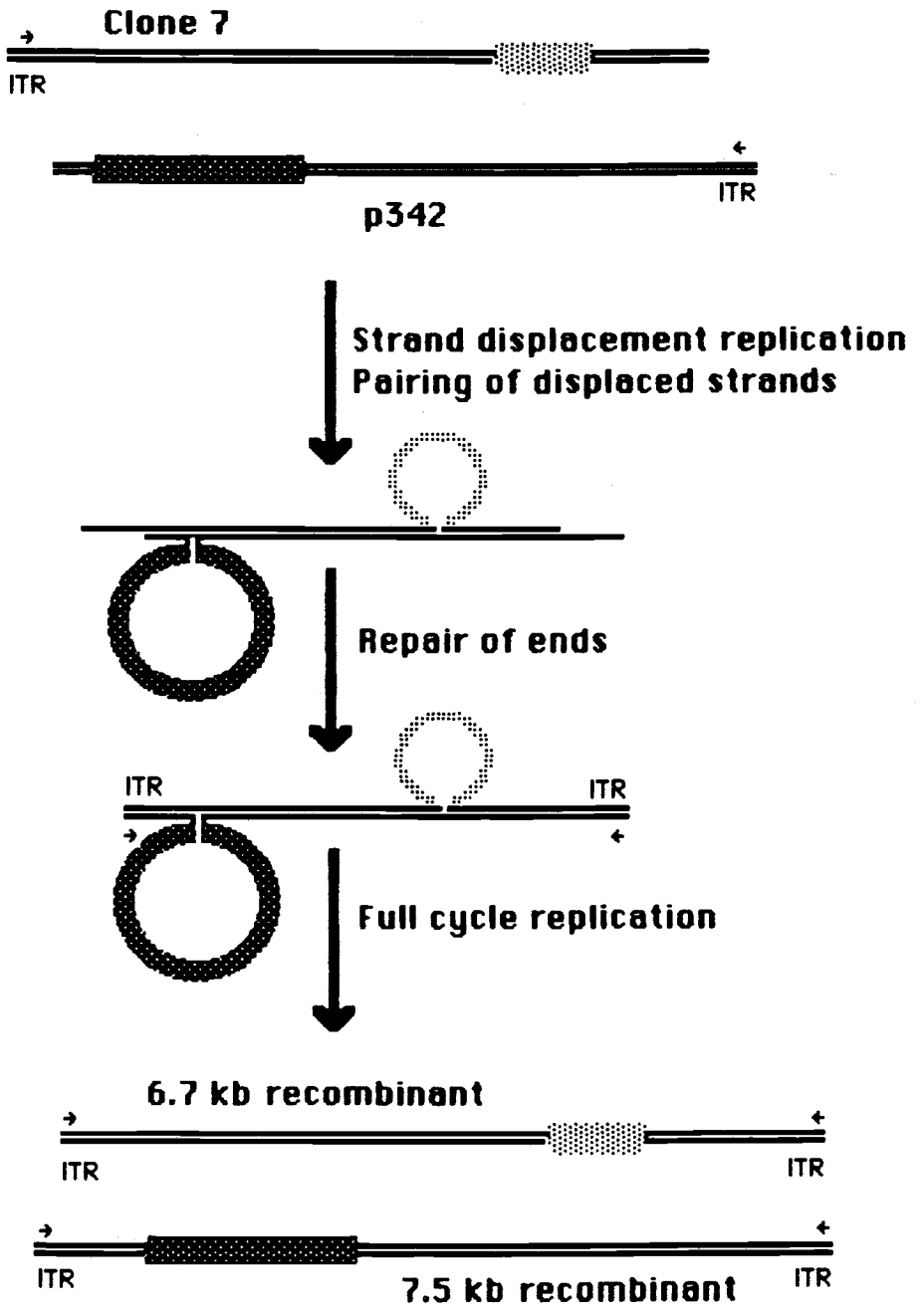


Figure 23: Postulated mechanism of trans replication involving pairing of siaplaced strands, repair by DNA polymerase, and amplification by replication.

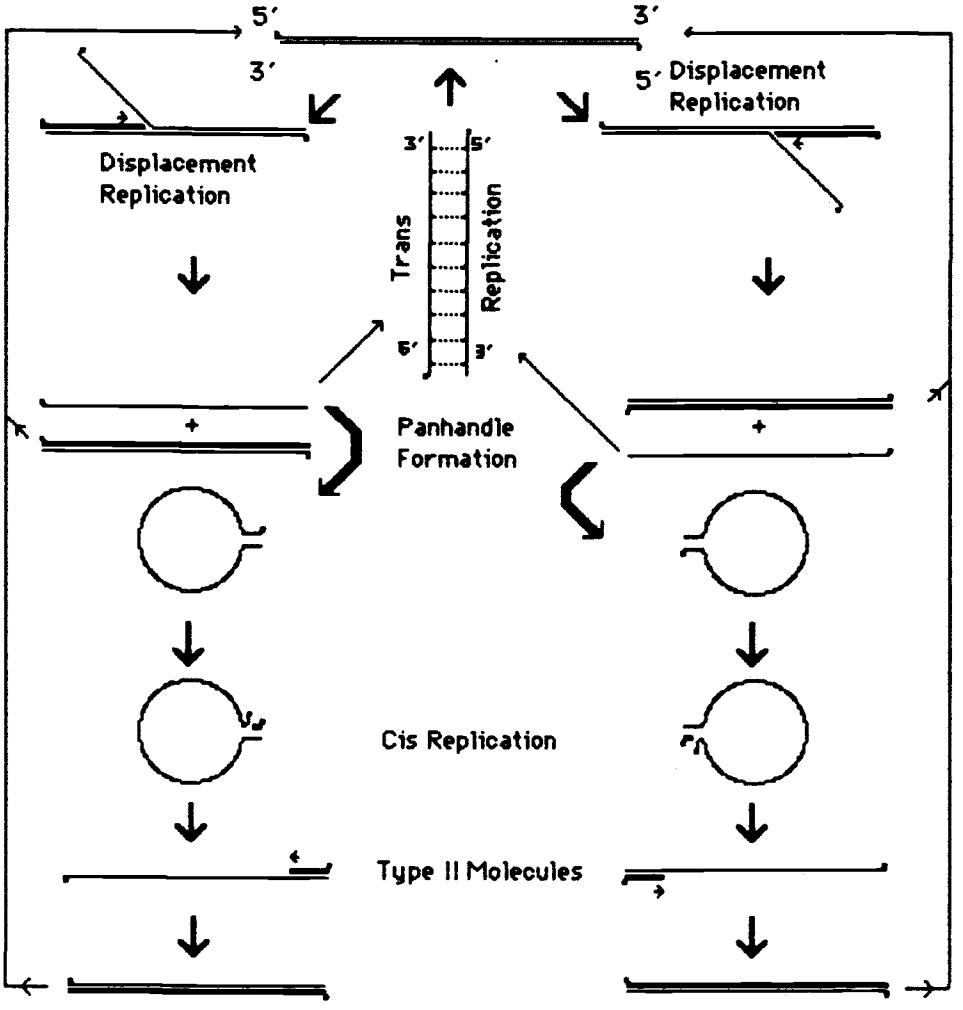


Figure 24: Trans replication in the adenovirus replication cycle.

synthesis (i.e. DNA polymerase uses any given strand as a template and to synthesize a new strand complementary to it), but promiscuous for complementary synthesis (i.e. any displaced single strand can pair with any other complementary displaced single strand). No evidence has been presented, but non-complementary strands could in theory also pair via the ITR sequences of each (see Figure 25). Such a base-paired molecule would be equivalent to a dual-handled panhandle and could contribute to the replicative cycle, probably at a low level. Molecules like these may not be an insignificant consideration, however, for such a molecule is an alternative to panhandle formation in repairing partially truncated ITR sequences in adenovirus (120) or plasmid panhandles (123,124). Because of this possibility, previous interpretations of data about *cis* replication should be reexamined.

INTRACELLULAR COMPLEMENTARY STRAND PAIRING

No direct evidence has been obtained yet in any system that proves that single-stranded nucleic acids actually pair with complementary sequences to form a duplex in the cell. Several laboratories, however, have obtained results consistent with the idea of pairing between complementary nucleic acid sequences. As explained above intramolecular pairing (*cis* pairing) is consistent with the observed repair of damaged adenovirus ITR sequences (120,123,124). Attempts to introduce nucleic acids that can pair in the cell such as "sense" and "anti-sense" mRNAs have in some cases resulted in the inhibition of the translation of selected mRNAs (224-228). Intramolecular pairing

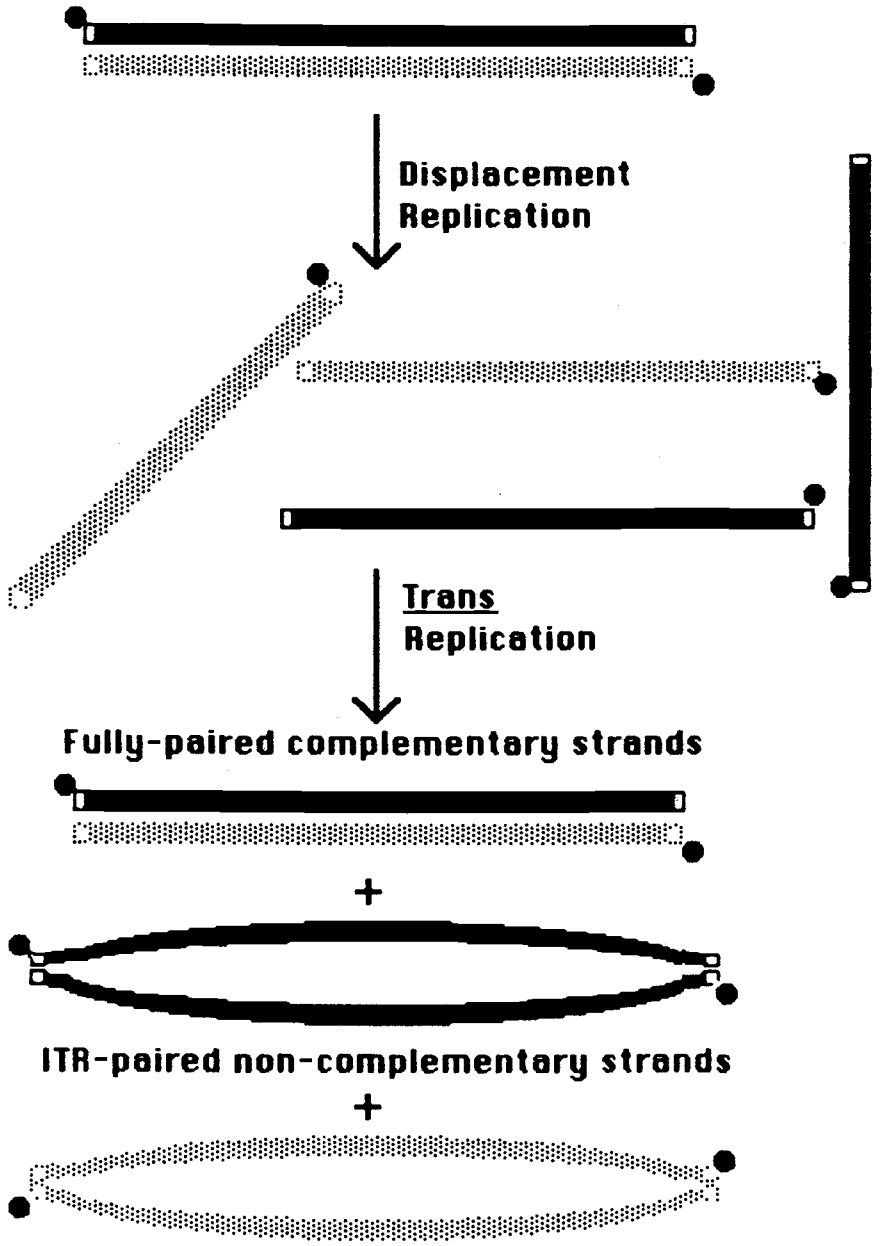


Figure 25: Pairing of displaced adenovirus strands through complementary sequences. See text for details.

(trans pairing) of "sense" and "anti-sense" mRNAs has been postulated as the mechanism responsible, presumably by inhibiting the translation process and/or the stability of mRNA molecules (224, 225). In bacterial systems, anti-sense/sense pairing of RNA molecules has been proposed as a mechanism controlling copy number and incompatibility of ColE1 type plasmids (229).

RELATIONSHIP TO PREVIOUS STUDIES

The data presented above strongly suggest that trans replication is a process distinct from homologous recombination and likely tied to replication. In fact, data consistent with homologous recombination was not obtained in any of the experiments. It would be premature, however, to suggest that homologous recombination is not occurring during an adenovirus infection (see below). Other investigators using other assays have obtained evidence of homologous recombination in an adenovirus system. Berkner and Sharp (40) showed that recombinant adenovirus molecules could be made by transfection of plasmids containing overlapping segments of adenovirus DNA if only one origin was exposed. Thus, recombination did not depend upon displacement replication of both input molecules. It should be noted, however, that the detection of homologous recombination by Berkner and Sharp and the lack of detection in the present work is easily reconciled by examining the different assays. First, their assay employed a selective procedure, namely plaque formation. Plaque production depends only on the amplification of a single recombination event. It is not clear how sensitive the blotting assay employed in this work

is, but if it is below the level of detection of rare recombinational events that are subsequently amplified by replication, then our results are totally consistent with those of Berkner and Sharp. By this interpretation, homologous recombination could indeed be occurring in the work described here, but far less frequently than trans replication and below the level of detection.

An alternative explanation is that since the plasmid DNAs described here are transfected with helper adenovirus DNA, the complete cycle of adenovirus replication (displacement and complementary synthesis) is actively occurring. In the system of Berkner and Sharp, only displacement replication can occur in the absence of a trans event to bring together the separated origins. It is possible that when complementary replication is occurring, as it is in the system described here, homologous recombination is inhibited while in its absence, it occurs at a normal level.

Kinetic Studies

Data from at least two kinetic studies of the replication of adenovirus DNA are consistent with the idea of trans replication. Bodnar and Pearson (230) and D'Halluin and Milleville (231) used density shift analysis of BUdR-marked DNA to measure initiation rates of adenovirus DNA replication. Bodnar and Pearson labeled adenovirus DNA HH (Heavy-Heavy) with BUdR and then shifted to light medium lacking BUdR. D-Halluin and Milleville started with LL (Light-Light) DNA and added BUdR, but their results were identical to those of Bodnar and Pearson which will be described here. The progress of DNA

from HH to HL (Heavy-Light) and LL was monitored by CsCl density gradient centrifugation. Measurements of HL DNA and LL DNA formed during the chase in light medium indicated that the rates of disappearance of HH and appearance of HL varied with time after infection. HH DNA disappeared at the fastest rate early (14 hr) in the replication cycle, slower midway (18 hr), and at the slowest rate late (22 hr) in the replication cycle. In contrast HL was produced at the fastest rate 14 hours after infection, slower at 18 hours and at the slowest rate 22 hours after infection.

Bodnar and Pearson interpreted this to mean that the rates of initiation were changing during the infectious cycle. These kinetics, however, are exactly what would be expected if displaced strands paired as a function of strand concentration. Early in infection, when the concentration of single strands from displacement synthesis is low, displaced strands would be more likely to undergo cis replication. This would result in production of two HL molecules from each HH molecule and would represent the fastest rate of HL synthesis. By contrast, late in the replication cycle, when displaced single H strands are at a high concentration, the pairing of complementary H strands will simply regenerate HH molecules. Thus, the rate of disappearance of HH strands late in infection will be slowed, due to their partial replenishment by the pairing of displaced strands.

In addition, Bodnar and Pearson observed that at about 18 hours after infection accumulation of viral DNA began to deviate significantly from logarithmic growth, and thereafter the kinetics of accumulation were roughly linear. This, too, is what might be predicted if pairing of displaced strands occurs as a function of

their concentration. At low concentrations of displaced strands, cis replication would be predicted to predominate, ultimately producing two double-stranded molecules from each parent. This doubling would occur (giving rise to logarithmic accumulation) until the concentration of displaced strands was at a high enough level such that pairing occurred. Pairing would have the effect of acting like a shunt, producing the equivalent of 1.5 double-stranded molecules from each parent, and preventing the full doubling reaction from occurring. Therefore, growth will deviate from logarithmic accumulation and become more linear.

IMPLICATIONS OF INTERMOLECULAR PAIRING

Previous Work

Previous investigations using density shift analysis with BUdR have indicated that adenovirus replication is semi-conservative (134, 135, 208). This is not surprising or inconsistent with the prediction of intermolecular pairing. The product of displacement replication is a double-stranded molecule that has been replicated semi-conservatively and a single-stranded molecule that can replicate semi-conservatively (cis replication) or promiscuously (trans replication). If single-stranded molecules replicated promiscuously, they would not be detected directly in a density shift analysis. This would be insignificant unless intermolecular pairing occurs as a significant part of the replicative cycle. If it does then it considerably alters the semi-conservative nature of the replication

and interpretations of data from earlier studies may have to be reexamined.

Homologous Recombination

The idea of *trans* replication occurring by a base pairing mechanism is not necessarily an alternative to homologous recombination. Although all the models described above require base pairing as an intermediate in the recombinational process, only the Lin model (222) suggests that single-stranded sequences might be recombinationally "hot". This is consistent with the data described here if homologous recombination proceeds by a process whereby the cellular recombinational machinery produces exposed single strands capable of pairing relatively infrequently (Figure 22). Once produced, however, homologous recombination could readily proceed. This might explain why *trans* replication appeared to require displacement replication. In its absence, the basal rate of production of exposed single strands by exonuclease activity and subsequent recombination was too low to be detected by Southern blotting whereas in the presence of displacement replication many highly recombinogenic, complementary, single strands were produced.

Trans Replication and ITR Size

The factors determining the choice between a *cis* or *trans* mode of replication are likely to interact dynamically. Obviously, pairing of single strands should be dependent on the concentration of strands.

The single strand concentration is probably determined by at least three factors, the efficiency of displacement strand synthesis, the length of ITR capable of *cis* pairing, and the efficiency of complementary strand synthesis on *cis*-paired (panhandle) molecules. Early in replication when the concentration of single strands is low, *cis* pairing would be favored over *trans* pairing due to the local concentration of ITR sequences. If *cis* pairing were inefficient due to a short ITR or if complementary synthesis on *cis*-paired molecules did not keep pace with displacement synthesis, then single strands would accumulate, increasingly favoring a *trans* mechanism later in the infection. Serotypes with a shorter ITR by this prediction might engage in *trans* replication more often than serotypes with a longer ITR.

One consequence of a high *trans* to *cis* replication mode is that the inverted terminal repeats would "see" each other less often (they do not "see" each other at all in the *trans* mode), and as a result diverge more rapidly. It is interesting, then, to note that the avian CELO virus, which is the adenovirus with the shortest known ITR, is also the only known adenovirus with a mismatch within the bounds of its ITR (56).

Trans Replication and ϕ29

An interesting virus to examine in light of *trans* replication is the *Bacillus subtilis* bacteriophage ϕ29. Like adenovirus, ϕ29 has a linear DNA molecule with a terminal protein covalently attached to the 5' end of each strand, and it replicates with a strand displacement

mechanism (170). Analysis of its molecular termini and the sequences surrounding them revealed that ϕ29 has an ITR of only 6 bp, much shorter than the adenovirus ITR (232,233). Wang et al. (K. Wang, personal communication) have shown that only 21-30 base pairs of inverted sequence are required for cis repair. If this is interpreted to indicate the sequence length required for panhandle formation in cis replication, then the 6 base pair ITR on ϕ29 DNA might need an additional pathway (such as trans replication) to complete the replication cycle. Obviously, other mechanisms might be used in ϕ29 replication. Nevertheless, if the arguments given above about the roles of cis and trans replication and their relationship to the size of the ITR are valid, then ϕ29 would be predicted as a likely system in which to find intermolecular pairing.

SUMMARY

In summary, the clear interpretation of these data is that adenovirus origin-containing molecules can use a trans mechanism, if necessary, to complete the replication cycle. The mechanism appears to be driven by displacement of both molecules, and has been postulated to be due to pairing between complementary sequences in the displaced strands. It is not clear, however, whether adenovirus actually utilizes either a cis or a trans mechanism during the viral replication cycle. The dynamics of cis versus trans replication ultimately may be determined using a competition assay currently under investigation.

CHAPTER 5: BIBLIOGRAPHY

1. Rowe, W.P., Huebner, A.J., Gilmore, L.K., Parrott, R.N., and Ward, T.G. (1953). Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. Proc. Soc. Exp. Biol. Med. 84, 570-573.

2. Hilleman, M.R., and Werner, J.R. (1954). Recovery of a new agent from patients with acute respiratory illness. Proc. Soc. Exp. Biol. Med. 85, 183-188.

3. Ishibashi, M. and Yasue, H. (1984). Adenoviruses of Animals. In The Adenoviruses, Harold S. Ginsberg, ed. (New York and London: Plenum Press), p. 497-562.

4. Rosen, L. (1960). A hemagglutination-inhibition technique for typing adenoviruses. Am. J. Hyg. 71, 121-128.

5. Piña, M., and Green, M. (1965). Biochemical studies on adenovirus multiplication. IX. Chemical and base composition analyses of 28 human adenoviruses. Proc. Natl. Acad. Sci. USA 54, 547-551.

6. Green, M., Mackey, J.K., Wold, W.S.M., and Rigden, P. (1979). Thirty-one human adenovirus serotypes (Ad1-Ad31) form five groups (A-E) based upon DNA genome homologies. Virology 93, 481-492.

7. Wadell, G., Hammarskjöld, M.-L., Winberg, G., Varsanyi, T.M., and Sundell, G. (1980). Genetic variability of adenoviruses. Ann. NY. Acad. Sci. 354, 16-31.

8. Green, M. (1970). Oncogenic viruses. Ann. Rev. Bioch. 39, 701-756.

9. Trentin, J.J., Yabe, Y., and Taylor, G. (1962). The quest for human cancer viruses. Science 137, 835-841.

10. Huebner, R.J. (1967). Adenovirus-directed tumor and T-antigens. In Perspectives in Virology, Vol. V, M. Pollard, ed. (New York: Academic Press), pp. 147-166.

11. Trentin, J.J., Van Hoosier, G.L., and Samper, L. (1968). The oncogenicity of human adenoviruses in hamsters. Proc. Soc. Exp. Biol. Med. 127, 683-689.

12. Graham, F.L., van der Eb., A.J., and Heijneker, H.L. (1974). Size and location of the transforming region in human adenovirus type 5 DNA. Nature 251, 687-691.

13. Montell, C., Courtois, G., Eng, C., and Berk, A. (1984). Complete transformation by adenovirus 2 requires both E1A proteins. Cell 36, 951-961.

14. Graham, F.L. (1984). Transformation by and oncogenicity of human adenoviruses. In The Adenoviruses, Harold S. Ginsberg, ed. (New York and London: Plenum Press), pp. 339-398.

15. Berget, S.M., Moore, C., and Sharp, P.A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. Proc. Natl. Acad. Sci. USA 74, 3171-3175.

16. Chow, L.T., Gelinas, R.E., Broker, T.R., and Roberts, R.J. (1977). An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. Cell 12, 1-8.

17. Kitchingman, G.R., Lai, S.-P., and Westphal, H. (1977). Loop structures in hybrids of early RNA and the separated strands of adenovirus DNA. Proc. Natl. Acad. Sci. USA 74, 4392-4395.

18. Padgett, R.A., Hardy, S.F., and Sharp, P.A. (1983). Splicing of adenovirus RNA in a cell-free transcription system. Proc. Natl. Acad. Sci. USA 80, 5230-5234.

19. Goldenberg, C.J., and Hauser, S.D. (1983). Accurate and efficient in vitro splicing of purified precursor RNAs specified by early region 2 of the adenovirus genome. Nucl. Acids Res. 11, 1337-1348.

20. Montell, C., Fisher, E.F., Caruthers, M.H., and Berk, A.J. (1984). Control of adenovirus E1B mRNA synthesis by a shift in the activities of RNA splice sites. Mol. Cell. Biol. 4, 966-972.

21. Gilardi, P. and Perricaudet, M. (1984). The E4 transcriptional unit of Ad2: Far upstream sequences are required for its transactivation by E1a. Nucl. Acids Res. 12, 7877-7889.

22. Nevins, J.R. (1981). Mechanism of activation of early viral transcription by the adenovirus E1A gene product. Cell 26, 213-220.

23. Hearing, P. and Shenk, T. (1985). Sequence-independent autoregulation of the adenovirus type 5 E1A transcription unit. Mol. Cell. Biol. 5, 3214-3221.

24. Jones, N. and Shenk, T. (1979). An adenovirus type 5 early gene function regulates expression of other early viral genes. Proc. Natl. Acad. Sci. USA 76, 3665-3669.

25. Rice, S.A. and Klessig, D.F. (1984). The function(s) provided by the adenovirus-specified, DNA-binding protein required for viral late gene expression is endependent of the role of the protein in viral DNA replication. J. Virol. 49, 35-49.

26. Velcich, A. and Ziff, E. (1985). Adenovirus E1A proteins repress transcription from the SV40 early promoter. Cell 40, 705-716.

27. Stein, R. and Ziff, E.B. (1984). HeLa cell β-Tubulin gene transcription is stimulated by adenovirus 5 in parallel with viral early genes by an E1a-dependent mechanism. Mol. Cell. Biol. 4, 2792-2801.

28. Sharp, P.A. (1984). Adenovirus Transcription. In The Adenoviruses, Harold S. Ginsberg, ed. (New York and London: Plenum Press), pp. 173-204.

29. Logan, J. and Shenk, T. (1984). Adenovirus tripartite leader sequence enhances translation of mRNAs late after infection. Proc. Natl. Acad. Sci. USA 81, 3655-3659.

30. Beltz, G.A. and Flint, S.J. (1979). Inhibition of HeLa cell protein synthesis during adenovirus infection: Restriction of cellular messenger RNA sequences to the nucleus. J. Mol. Biol. 131, 353-373.

31. Babich, A., Feldman, L.T., Nevins, J.R., Darnell, J.E. Jr., and Weinberger, C. (1983). Effect of adenovirus on metabolism of specific translational discrimination. Mol. Cell. Biol. 3, 1212-1221.

32. Schneider, R.J., Weinberger, C., and Shenk, T. (1984). Adenovirus VAI RNA facilitates the initiation of translation in virus-infected cells. Cell 37, 291-298.

33. Reichel, P.A., Merrick, W.C., Siekierka, J., and Mathews, M.B. (1985). Regulaton of a protein synthesis initiation factor by adenovirus virus-associated RNAI. Nature 313, 196-200.

34. Siekierka, J., Mariano, T.M., Reichel, T.A., and Mathews, M.B. (1985). Translational control by adenovirus: Lack of virus-associated RNA_I during adenovirus infection results in phosphorylation of initiation factor eIF-2 and inhibition of protein synthesis. Proc. Natl. Acad. Sci. USA 82, 1959-1963.

35. O'Malley, R.P., Mariano, T.M., Siekierka, J., and Mathews, M.B. (1986). A mechanism for the control of protein synthesis by adenovirus VA RNA I. Cell 44, 391-400.

36. Hart, R.P., McDevitt, M.A., Ali, H., and Nevins, J.R. (1985). Definition of essential sequences and functional equivalence of elements downstream of the adenovirus E2A and early simian virus 40 polyadenylation sites. Mol. Cell. Biol. 5, 2975-2983.

37. Bhat, B.M. and Wold, W.S.M. (1985). ATTAAA as well as downstream sequences are required for RNA 3' end formation in the E3 complex transcription unit of adenovirus. Mol. Cell. Biol. 5, 3183-3193.

38. Fraser, N.W., Baker, C.C., Moore, M.A., and Ziff, E.B. (1982). Poly(A) sites of adenovirus serotype 2 transcription units. J. Mol. Biol. 155, 207-233.

39. Doerfler, W., Gahlmann, R., Stabel, S., Deuring, R., Lichtenberg, U., Schulz, M., Eick, D., and Leisten, R. (1984). On the mechanism of recombination between adenovirus and cellular DNAs: The structure of junction sites. In Current Topics in Microbiology and Immunology: The Molecular Biology of Adenoviruses 1, Walter Doerfler, ed. (Berlin, Heidelberg, New York, Tokyo: Springer-Verlag), pp. 193-228.

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

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

42. Munz, P.L. and Young, C.S.H. (1984). Polarity in adenovirus recombination. Virology 135, 503-514.

43. Mautner, V. and Boursnell, M.E.G. (1983). Recombination in adenovirus: DNA sequence analysis of crossover sites in intertypic recombinants. Virology 131, 1-10.

44. Williams, J., Grodzicker, T., Sharp, P. and Sambrook, J. (1975). Adenovirus recombination: Physical mapping of crossover events. Cell 4, 113-119.

45. Boursnell, M.E.G., and Mautner, V. (1981). Recombination in adenovirus: Crossover sites in intertypic recombinants are located in regions of homology. Virology 112, 198-209.

46. Hen, R., Borrelli, E., Sassone-Corsi, P., and Chambon, P. (1983). An enhancer element is located 340 base pairs upstream from the adenovirus 2 E1A capsite. Nucl. Acids Res. 11, 8747-8760.

47. Imperiale, M.J. and Nevins, J.R. (1984). Adenovirus 5 E2 transcription unit: An E1A-inducible promoter with an essential element that functions independently of position or orientation. Mol. Cell. Biol. 4, 875-882.

48. Nermut, M.V. (1984). The Architecture of Adenoviruses. In The Adenoviruses, Harold S. Ginsberg, ed. (New York and London: Plenum Press), pp. 5-34.

49. Corden, J., Engelking, H.M., and Pearson, G.D. (1976). Chromatin-like organization of the adenovirus chromosome. Proc. Natl. Acad. Sci. USA 73, 401-404.

50. Brown, D.T., Westphal, M., Burlingham, B.T., Winterhoff, U., and Doerfler, W. (1975). Structure and composition of the adenovirus type 2 core. J. Virol. 16, 366.

51. Garon, C.F., Berry, K.W., and Rose, J.A. (1972). A unique form of terminal redundancy in adenovirus DNA molecules. Proc. Natl. Acad. Sci. USA 69, 2391-2395.

52. Wolfson, J. and Dressler, D. (1972). Adenovirus-2 DNA contains an inverted terminal repetition. Proc. Natl. Acad. Sci. USA 69, 3054-3057.

53. Robinson, A.J., and Bellett, J.D. (1975). A circular DNA-protein complex from adenoviruses and its possible role in DNA replication. Cold Spring Harbor Symp. Quant. Biol. 39, 523-531.

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

55. Carusi, E.A. (1977). Evidence for blocked 5' termini in human adenovirus DNA. Virology 76, 380-394.

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

57. Brinckmann, U., Darai, G., and Flügel, R.M. (1983). The nucleotide sequence of the inverted terminal repetition of the tree shrew adenovirus DNA. Gene 24, 131-135.

58. Guggenheimer, R.A., Stillman, B.W., Nagata, K., Tamanoi, F., and Hurwitz, J. (1984). DNA sequences required for the in vitro replication of adenovirus DNA. Proc. Natl. Acad. Sci. USA 81, 3069-3073.

59. Challberg, M.D. and Rawlins, D.R. (1984). Template requirements for the initiation of adenovirus DNA replication. Proc. Natl. Acad. Sci. USA 81, 100-104.

60. Tamanoi, F. and Stillman, B.W. (1983). Initiation of adenovirus DNA replication in vitro requires a specific DNA sequence. Proc. Natl. Acad. Sci. USA 80, 6446-6450.

61. Enns, R.E., Challberg, M.D., Ahern, K.G., Chow, K.-C., Mathews, C.Z., Astell, C.R., and Pearson, G.D. (1983). Mutational mapping of a cloned adenovirus origin. Gene 23, 307-313.

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

63. Winnacker, E.-L. (1978). Adenovirus DNA: Structure and function of a novel replicon. Cell 14, 761-773.

64. Horwitz, M.S., Kaplan, M., Abboud, J., Maritato, J., Chow, L.T., and Broker, T.R. (1978). Adenovirus DNA replication in soluble extracts of infected cell nuclei. Cold Spring Harbor Symp. Quant. Biol. 43, 769-780.

65. Kelly, T.J. Jr. (1984). Adenovirus DNA Replication. In The Adenoviruses, Harold S. Ginsberg, ed. (New York and London: Plenum Press), pp. 271-308.

66. Challberg, M.D. and Kelly, T.J. (1982). Eukaryotic DNA replication: Viral and plasmid model systems. Ann. Rev. Biochem. 51, 901-934.

67. Stillman, B.W. (1983). The replication of adenovirus DNA with purified proteins. Cell 35, 7-9.

68. Fütterer, J. and Winnacker, E.-L. (1984). Adenovirus DNA replication. In Current Topics in Microbiology and Immunology: The Molecular Biology of Adenoviruses 1, W. Doerfler, ed. (Berlin, Heidelberg, New York, Tokyo: Springer-Verlag), pp. 41-64.

69. Tamanoi, F. and Stillman, B.W. (1984). The origin of adenovirus replication. In Current Topics in Microbiology and Immunology: The Molecular Biology of Adenoviruses 1, W. Doerfler, ed. (Berlin, Heidelberg, New York, Tokyo: Springer-Verlag), pp. 75-88.

70. Sussenbach, J.S., and van der Vliet, P.C. (1984). The mechanism of adenovirus DNA replication and the characterization of replication proteins. In Current Topics in Microbiology and Immunology: The Molecular Biology of Adenoviruses 1, W. Doerfler, ed. (Berlin, Heidelberg, New York, Tokyo: Springer-Verlag), pp. 53-74.

71. Friefeld, B.R., Lichy, J.H., Field, J., Gronostajski, R.M., Krevolin, M.D., Nagata, K., Hurwitz, J., and Horwitz, M.S. (1984). The *in vitro* replication of adenovirus DNA. In Current Topics in Microbiology and Immunology: The Molecular Biology of Adenoviruses 1, W. Doerfler, ed. (Berlin, Heidelberg, New York, Tokyo: Springer-Verlag), pp. 221-256.

72. Philipson, L., Lonberg-Holm, K., and Pettersson, U. (1968). Virus-receptor interactions in an adenovirus system. J. Virol. 2, 1064-1072.

73. Morgan, G., Rosenkranz, H.S., and Medris, B. (1969). Structure and development of viruses as observed in the electron microscope. X. Entry and uncoating of adenovirus. J. Virol. 4, 777-788.

74. Lonberg-Holm, K., and Philipson, L. (1969). Early events of virus-cell interactions in an adenovirus system. J. Virol. 4, 323-335.

75. Green, M., Parsons, J.T., Piña, M., Fujinaga, K., Caffier, H., and Landgraf-Leurs, I. (1970). Transcription of adenovirus genes in productivity infected and in transformed cells. Cold Spring Harbor Symp. Quant. Biol. 35, 803-818.

76. Imperiale, M.J., Hart, R.P., and Nevins, J.R. (1985). An enhancer-like element in the adenovirus E2 promoter contains sequences essential for uninduced and E1A-induced transcription. Proc. Natl. Acad. Sci. USA 82, 381-385.

77. Babich, A., and Nevins, J.R. (1981). The stability of early adenovirus mRNA is controlled by the viral 72 kDal DNA-binding protein. Cell 26, 371-379.

78. Nevins, J.R., Ginsberg, H.S., Blanchard, J.M., Wilson, M.C., and Darnell, J.E. (1979). Regulation of the early adenovirus transcription units. J. Virol. 32, 727-733.

79. Berk, A.J., Lee, F., Harrison, T., Williams, J., and Sharp, P.A. (1979). A pre-early Ad5 gene product regulates synthesis of early viral mRNAs. Cell 17, 935-944.

80. Handa, A., Kingston, R.E., and Sharp, P.A. (1983). Inhibition of adenovirus early region IV transcription in vitro by a purified viral DNA binding protein. Nature 302, 545-547.

81. Rice, S.A. and Klessig, D.F. (1985). Isolation and analysis of adenovirus type 5 mutants containing deletions in the gene encoding the DNA-binding protein. J. Virol. 56, 767-778.

82. Stillman, B.W., Lewis, J.B., Chow, L.T., Mathews, M.B., and Smart, J.E. (1981). Identification of the gene and mRNA for the adenovirus terminal protein precursor. Cell 23, 497-508.

83. Kingston, R.E., Kaufman, R.J., and Sharp, P.A. (1984). Regulation of transcription of the adenovirus EII promoter by E1a gene products: Absence of sequence specificity. Mol. Cell. Biol. 4, 1970-1977.

84. Murthy, S.C.S., Bhat, G.P., and Thimmappaya, B. (1985). Adenovirus EIIA early promoter: Transcriptional control elements and induction by the viral pre-early E1A gene, which appears to be sequence independent. Proc. Natl. Acad. Sci. USA 82, 2230-2234.

85. Akusjärvi, G. and Persson, H. (1981). Controls of RNA splicing and termination in the major late as transcription unit. Nature 292, 420-426.

86. Fanning, E. and Doerfler, W. (1977). Intracellular forms of adenovirus DNA. Virology 81, 433-448.

87. Piña, M., and Green, M. (1969). Biochemical studies on adenovirus multiplication. XIV. Macromolecule and enzyme synthesis in cells replicating oncogenic and non-oncogenic human adenoviruses. Virology 38, 573-583.

88. Mantyjarvi, R., and Russell, W.C. (1969). Studies on DNA synthesis in cells infected with adenovirus. J. Gen. Virol. 5, 339-350.

89. 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.

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

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

92. Yamashita, T., Arens, M., and Green, M. (1977). Adenovirus deoxyribonucleic acid replication. J. Biol. Chem. 252, 7940-7946.

93. Kaplan, L.M., Kleinman, R.E., and Horwitz, M.S. (1977). Replication of adenovirus type 2 DNA in vitro. Proc. Natl. Acad. Sci. USA 74, 4425-4429.

94. Sussenbach, J.S., Ellens, D.J., and Jansz, H.S. (1973). Studies on the mechanism of replication of adenovirus DNA. I. The nature of single-stranded DNA in replicative intermediates. J. Virol. 12, 1131-1140.

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

96. Pettersson, U. (1973). Some unusual properties of replicating adenovirus type 2 DNA. J. Mol. Biol. 81, 521-527.

97. Bellett, A.J.D. and Younghusband, H.B. (1972). Replication of the DNA of chick Embryo Lethal Orphan Virus. J. Mol. Biol. 72, 691-709.

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

99. Schilling, R., Weingartner, B., and Winnacker, E.L. (1975). Adenovirus type 2 DNA replication: II. Termini of DNA replication. J. Virol. 16, 767-774.

100. Tolun, A. and Pettersson, U. (1975). Termination sites for adenovirus type 2 DNA replication. J. Virol. 16, 759-766.

101. van der Vliet, P.C., Levine, A.J., Ensinger, M., and Ginsberg, H.S. (1975). Thermo-labile DNA-binding proteins from cells infected with a temperature-sensitive mutant of adenovirus defective in viral synthesis. J. Virol. 15, 348-360.

102. van der Vliet, P.C. and Sussenbach, J.S. (1975). An adenovirus type 5 gene function for initiation of viral DNA replication. Virology 67, 415-426.

103. Gilead, Z., Sugawara, K., Shanmugam, G., and Green, M. (1976). Synthesis of the adenovirus-coded DNA binding protein in infected cells. J. Virol. 18, 454-460.

104. Chow, L.T., Lewis, J.B., and Broker, T.R. (1984). RNA transcription and splicing at early and intermediate times after Adenovirus-2 infection. Cold Spring Harbor Symp. Quant. Biol. 44, 401-414. (1980).

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

106. 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 H5ts36-, H5ts149-, and H5ts125-infected HeLa cells. J. Virol. 46, 642-648.

107. Alwine, J.C., Reed, S.I., and Stark, G.R. (1977). Characterization of the autoregulation of Simian Virus 40 gene A. J. Virol. 24, 22-27.

108. Jeng, Y.-H., Wold, W.S., Sugawara, K., Gilead, Z., and Green, M. (1977). Adenovirus type 2 coded single-stranded DNA binding protein: *in vivo* phosphorylation and modification. J. Virol. 22, 402-411.

109. Kedinger, C., Brison, O., Perrin, F., and Wilhelm, J. (1978). Structural analysis of replicative intermediates from adenovirus type 2 infected HeLa cell nuclei. J. Virol. 26, 364-383.

110. Levine, A.J., van der Vliet, P.C., Rosenwirth, B., Rabek, J., Frenkel, G., and Ensinger, M. (1975). Adenovirus-infected, cell-specific, DNA-binding Proteins. Cold Spring Harbor Symp. Quant. Biol. 39, 559-566.

111. Foster, D.A., Hantzopoulos, P., and Zubay, G. (1982). Resistance of adenovirus DNA replication to aphidicolin is dependent on the 72-kilodalton DNA-binding protein. J. Virol. 43, 679-686.

112. Alberts, B.M., and Frey, L. (1970). T4 bacteriophage gene 32: A structural protein in the replication and recombination of DNA. Nature 227, 1313-1318.

113. Kruijer, W., van Schaik, F.M.A., and Sussenbach, J.S. (1981). Structure and organization of the gene coding for the DNA binding protein of adenovirus type 5. Nucl. Acids Res. 9, 4439-4457.

114. Linné, T., Jörnvall, H., and Philipson, L. (1977). Purification and characterization of the phosphorylated DNA-binding protein from adenovirus-type 2 infected cells. Eur. J. Bioch. 76, 481-490.

115. Sugawara, K., Gilead, Z., and Green, M. (1977). Purification and molecular characterization of adenovirus type 2 DNA binding protein. J. Virol. 21, 338-346.

116. Axelrod, N. (1978). Phosphoproteins of adenovirus 2. Virology 87, 366-383.

117. Klein, H., Maltzman, W., and Levine, A.J. (1979). Structure function relationships of the adenovirus DNA-binding protein. J. Biol. Chem. 254, 11051-11060.

118. Ellens, D.J., Sussenbach, 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.

119. Revet, B. and Benichou, D. (1981). Electron microscopy of Ad5 replicating molecules after in vivo photocrosslinking with trioxsalen. Virology 114, 60-70.

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

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

122. Graham, F.L., Smiley, J., Russell, W.C., and Nairn, R. (1977). Characteristics of a human cell line transformed by DNA from human adenovirus 5. J. Gen. Virol. 36, 59-72.

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

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

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

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

127. Kornberg, A. (1980) In DNA Replication (San Francisco: W.H. Freeman and Company).

128. Horwitz, M.S. (1976). Bidirectional replication of adenovirus type 2 DNA. J. Virol. 18, 307-315.

129. 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.

130. Rijnders, W.M., van Bergen, B.G.M., van der Vliet, P.C., and Sussenbach, J.S. (1983). Immunological characterization of the role of adenovirus terminal protein in viral DNA replication. Virology 131, 287-295.

131. Robinson, A.J., Bodnar, J.W., Coombs, D.H., and Pearson, G.D. (1979). Replicating adenovirus 2 DNA molecules contain terminal protein. Virology 96, 143-158.

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

133. Stillman, B.W. and Bellett, A.J.D. (1979). An adenovirus protein associated with the ends of replicating DNA molecules. Virology 93, 69-79.

134. Yamashita, T., Arens, M., and Green, M. (1975). Adenovirus deoxyribonucleic acid replication II. Synthesis of viral deoxyribonucleic acid in vitro by a nuclear membrane fraction from infected KB cells. J. Biol. Chem. 250, 3273-3279.

135. Brison, O., Kedinger, C., and Wilhelm, J. (1977). Enzymatic properties of viral replication complexes isolated from adenovirus type 2 infected HeLa cell nuclei. J. Virol. 24, 423-435.

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

137. van Bergen, R.G.M., van der Ley, P.A., van Driel, W., van Mansfeld, A.D.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.

138. Guggenheimer, R.A., Nagata, K., Kenny, M., and Hurwitz, J. (1984). Protein-primed replication of plasmids containing the terminus of the adenovirus genome. J. Biol. Chem. 259, 7815-7825.

139. Lichy, J.H., Nagata, K., Friefeld, B.R., Enomoto, T., Field, J., Guggenheimer, R.A., Ikeda, J.-E., Horwitz, M.S., and Hurwitz, J. (1983). Isolation of proteins involved in the replication of Adenoviral DNA in vitro. Cold Spring Harbor Symp. Quant. Biol. 47, 731-740.

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

141. Cavalier-Smith, T. (1974). Palindromic base sequences and replication of eukaryote chromosome ends. Nature 250, 467-470.

142. Kelly, T.J. Jr. and Lechner, R.L. (1979). The structure of replicating adenovirus DNA molecules: Characterization of DNA-protein complexes from infected cells. Cold Spring Harbor Symp. Quant. Biol. 43, 721-728.

143. Coombs, D.H., Robinson, A.J., Bodnar, J.W., Jones, C.J., and Pearson, G.D. (1979). Detection of covalent DNA-protein complexes: The adenovirus DNA-terminal protein complex and HeLa DNA-protein complexes. Cold Spring Harbor Symp. Quant. Biol. 43, 741-752.

144. Stillman, B.W. and Bellett, A.J.D. (1979). An adenovirus protein associated with the ends of replicating DNA molecules. Virology 93, 69-79.

145. Van Wielink, P.S., Naaktgeboren, N., and Sussenbach, J.S. (1979). Presence of protein at the termini of intracellular adenovirus type 5 DNA. Bioch. Bioph. Acta 563, 89-99.

146. 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.

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

148. 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.

149. 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-kilodalton terminal protein. J. Virol. 41, 265-270.

150. 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.

151. 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, 349-337.

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

153. 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 preterminal protein-dCMP complex. Proc. Natl. Acad. Sci. USA 79, 6438-6442.

154. 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.

155. Nagata, K., Guggenheimer, R.A., and Hurwitz, J. (1983). Specific binding of a cellular DNA replication protein to the origin of replication of adenovirus DNA. Proc. Natl. Acad. Sci. USA 50, 6177-6181.

156. Rawlins, D.R., Rosenfeld, P.J., Wides, R.J., Challberg, M.D., and Kelly, T.J. Jr. (1984). Structure and function of the adenovirus origin of replication. Cell 37, 309-319.

157. Ikeda, J.-E., Enomoto, T., and Hurwitz, J. (1982). Adenoviral protein-primed initiation of DNA chains in vitro. Proc. Natl. Acad. Sci. USA 79, 2442-2446.

158. Stillman, B.W. and Tamanoi, F. (1982). Adenoviral DNA replication: DNA sequences and enzymes required for initiation in vitro. Cold Spring Harbor Symp. Quant. Biol. 47, 741-750.

159. Rijnders, A.W.M., van Bergen, B.G.M., van der Vliet, P.C., and Sussenbach, J.S. (1983). Specific binding of the adenovirus terminal protein precursor-DNA polymerase complex to the origin of DNA replication. Nucl. Acids Res. 11, 8777-8789.

160. Green, M., Brackmann, K., Wold, W.S.M., Cartas, M., Thornton, H., and Elder, J.H. (1979). Conserved primary sequences of the DNA terminal proteins of five different human adenovirus groups. Proc. Natl. Acad. Sci. USA 76, 4380-4384.

161. Rekosh, D. (1981). Analysis of the DNA-terminal protein from different serotypes of human adenovirus. J. Virol. 40, 329-333.

162. Smart, J.E. and Stillman, B.W. (1982). Adenovirus terminal protein precursor. J. Biol. Chem. 257, 13499-13506.

163. Binger, M.H., Flint, S.J., and Rekosh, D.M. (1982). Expression of the gene encoding the adenovirus DNA terminal protein precursor in productively infected and transformed cells. J. Virol. 42, 488-501.

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

165. Gingeras, T.R., Sciaky, D., Gelinas, R.E., Bing-Dong, J., Yen, C.E., Kelly, M.M., Bullock, P.A., Parsons, B.L., O'Neill, K.E., and Roberts, R.J. (1982). Nucleotide sequences from the ad-2 genome. J. Biol. Chem. 257, 13475-13491.

166. Aleström, P., Akusjärvi, G., Pettersson, M., and Pettersson, U. (1982). DNA sequence analysis of the region encoding the terminal protein and the hypothetical N-gene product of adenovirus type 2. J. Biol. Chem. 257, 13492-13498.

167. Revie, D., Tseng, B.Y., Grafstrom, R.H., and Goulian, M. (1979). Covalent association of protein with replicative form DNA of parvovirus H-1, Proc. Natl. Acad. Sci. USA 76, 5539-5534.

168. Gerlich, W.H. and Robinson, W.S. (1980). Hepatitis B virus contains protein attached to the 5' terminus of its complete DNA strand. Cell 21, 801-809.

169. Daubert, S.D., Bruening, G., and Najarian, R.C. (1978). Protein bound to the genome RNAs of cowpea mosaic virus. Eur. J. Bioch. 92, 45-51.

170. Salas, M. (1984). A new mechanism for the initiation of replication of ϕ29 and adenovirus DNA: Priming by the terminal protein. In Current Topics in Microbiology and Immunology: The Molecular Biology of Adenoviruses 1, W. Doerfler, ed. (Berlin, Heidelberg, New York, Tokyo: Springer-Verlag), pp. 89-106.

171. Golini, F., Nomoto, A., and Wimmer, E. (1978). The genome-linked protein of picornaviruses: IV. Difference in the VPg's of encephalomyocarditis virus and poliovirus as evidence that the genome-linked proteins are virus-coded. Virology 89, 112-118.

172. Longiaru, M., Ikeda, J.-E., Jarkovsky, Z., Horwitz, S.B., and Horwitz, M.S. (1979). The effect of aphidicolin on adenovirus DNA synthesis. Nucl. Acids Res. 6, 3369-3386.

173. van der Vliet, P.C. and Kwant, M.M. (1978). Role of DNA polymerase γ in adenovirus DNA replication. Nature 276, 532-534.

174. Krokan, H., Schaffer, P., and DePamphilis, M.L. (1979). Involvement of eucaryotic deoxyribonucleic acid polymerases α and γ in the replication of cellular and viral deoxyribonucleic acid. Biochemistry 18, 4431-4445.

175. Aleström, P., Akusjärvi, G., Pettersson, M. and Pettersson, U. (1982). DNA sequence analysis of the region encoding the terminal protein and the hypothetical N-gene product of adenovirus type 2. J. Biol. Chem. 257, 13492-13498.

176. Galos, R.S., Williams, J., Binger, M.-H., and Flint, S.J. (1979). Location of additional early gene sequences in the adenoviral chromosome. Cell 17, 945-956.

177. Friefeld, B.R., Lichy, J.H., Hurwitz, J., and Horwitz, M.S. (1983). Evidence for an altered adenovirus DNA polymerase in cells infected with the mutant H5ts149. Proc. Natl. Acad. Sci. USA 80, 1589-1593.

178. Rekosh, D., Lindenbaum, J., Brewster, J., Mertz, L.M., Hurwitz, J., and Prestine, L. (1985). Expression in E. coli of a fusion protein product containing a region of the adenovirus DNA polymerase. Proc. Natl. Acad. Sci. USA 82, 2354-2358.

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

180. Dekker, B.M.M. and van Ormondt, H. (1984). The nucleotide sequence of fragment HindIII-C of human adenovirus type 5 DNA Gene 27, 115-120.

181. Friefeld, B.R., Krevolin, M.D., and Horwitz, M.S. (1983). Effects of the adenovirus H5ts125 and H5ts107 DNA binding proteins on DNA replication in vitro. Virology 124, 380-389.

182. 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.

183. Leegwater, P.A.J., van Driel, W., and van der Vliet, P.C. (1985). Recognition site of nuclear factor I, a sequence-specific DNA binding protein from HeLa cells that stimulates adenovirus DNA replication. EMBO 4, 1515-1521.

184. Hay, R.T. (1985). Origin of adenovirus DNA replication: Role of the nuclear factor I binding site *in vitro*. J. Mol. Biol. 186, 129-136.

185. van Ormondt, H. and Galibert, F. (1984). Nucleotide sequences of adenovirus DNAs. In Current Topics in Microbiology and Immunology: The Molecular Biology of Adenoviruses 1, W. Doerfler, ed. (Berlin, Heidelberg, New York, Tokyo: Springer-Verlag), pp. 73-142.

186. Siebenlist, U., Hennighausen, L., Battey, J., and Leder, P. (1984). Chromatin structure and protein binding in the putative regulatory region of the c-myc gene in Burkitt lymphoma. Cell 37, 381-391.

187. Hennighausen, L., Siuebenlist, U., Danner, D., Leder, P., Rawlins, D., Rosenfeld, P., and Kelly, T. Jr. (1985). High-affinity binding site for a specific nuclear protein in the human IgM gene. Nature 314, 289-292.

188. Nowock, J., Borgmeyer, U., Püschel, A.W., Rupp, R.A.W., and Sippel, A.E. (1985). The TGGCA protein binds to the MMTV-LTR, the adenovirus origin of replication, and the BK virus enhancer. Nucl. Acids Res. 13, 2045-2061.

189. Nowock, J. and Sippel, A.E. (1982). Specific protein-DNA interaction at four sites flanking the chicken lysozyme gene. Cell 30, 607-615.

190. Borgmeyer, U., Nowock, J., and Sippel, A.E. (1984). The TGGCA-binding protein: A eukaryotic nuclear protein recognizing a symmetrical sequence on double-stranded linear DNA. Nucl. Acids Res. 12, 4295-4311.

191. Gronostajski, R.M., Adhya, S., Nagata, K., Guggenheimer, R.A., and Hurwitz, J. (1985). Site-specific DNA binding of nuclear factor I: Analyses of cellular binding sites. Mol. Cell. Biol. 5, 964-971.

192. Oguro, M., Yamashita, T., Ariga, H., and Nagano, H. (1984). Adenovirus DNA synthesized in the presence of aphidicolin. Nucl. Acids Res. 12, 1077-1086.

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

194. Lally, C., Dörper, T., Gröger, W., Antoine, G., and Winnacker, E.-L. (1984). A size analysis of the adenovirus replicon. EMBO 3, 333-337.

195. Pearson, G.D., Chow, K.-C., Enns, R.E., Ahern, K.G., Corden, J.L., and Harpst, J.A. (1983). *In vitro* replication directed by a cloned adenovirus origin. Gene 23, 293-305.

196. Hay, R.T. (1985). The origin of adenovirus DNA replication: Minimal DNA sequence requirement *in vivo*. EMBO 4, 421-426.

197. Geier, G. and Modrich, P. (1979). Recognition sequence of the *dam* methylase of *Escherichia coli* K12 and mode of cleavage of DpnI endonuclease. J. Biol. Chem. 254, 1408-1413.

198. Chow, K.-C. and Pearson, G.D. (1984). Site-specific nicking within the adenovirus inverted terminal repetition. Nucl. Acids Res. 12, 1489-1499.

199. Dynan, W.S. and Tjian, R. (1983). Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. Cell 32, 669-680.

200. Dynan, W.S. and Tjian, R. (1985). Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature 316, 774-778.

201. 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.

202. van der Vliet, P.C., van Dam, D., and Kwant, M.M. (1984). Adenovirus DNA replication *in vitro* is stimulated by RNA from uninfected HeLa cells. FEBS 171, 5-8.

203. Krainer, A.R. and Maniatis, T. (1985). Multiple factors including the small nuclear ribonucleoproteins U1 and U2 are necessary for Pre-mRNA splicing *in vitro*. Cell 42, 725-736.

204. Black, D.L., Chabot, B., and Steitz, J.A. (1985). U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. Cell 42, 737-750.

205. Ruben, M., Bacchetti, S., and Graham, F. (1983). Covalently closed circles of adenovirus 5 DNA. Nature 301, 172-174.

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

207. Flint, S.J., Berget, S.M., and Sharp, P.A. (1976). Characterization of single-stranded viral DNA sequences present during replication of adenovirus types 2 and 5. Cell 9, 559-571.

208. Frenkel, G.D. (1978). Adenovirus DNA synthesis *in vitro* in an isolated complex. J. Virol. 25, 459-463.

209. Schlief, R.F. and Wensink, P.C. (1981). In Practical Methods in Molecular Biology. (Berlin, Heidelberg, New York, and Tokyo: Springer-Verlag).

210. Maniatis, T., Fritsch, E., and Sambrook, J. (1982). In Molecular Cloning. (New York: Cold Spring Harbor Laboratory).

211. 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.

212. Beck, E., Ludwig, G., Auerswald, E.A., Reiss, B., and Schaller, H. (1982). Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn5. Gene 19, 327-336.

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

214. Wigler, M., Sweet, R., Simm, G.K., Wold, B., Pellicer, A., Lacy, E., Maniatis, T., Silverstein, S. and Axel, R. (1979). Transformation of mammalian cells with genes from procaryotes and eucaryotes. Cell 16, 777-785.

215. Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503-517.

216. Stow, N.D. (1981). Cloning of a DNA fragment from the left-hand terminus of the adenovirus type 2 genome and its use in site-directed mutagenesis. J. Virol. 37, 171-180.

217. Wolgemuth, D.J., and Hsu, M.-T. (1981). DNA replication-mediated recombination of molecules of adenovirus 2 DNA. Proc. Natl. Acad. Sci. USA 78, 5076-5080.

218. Meselson, M.S. and Radding, C.M. (1975). A general model for genetic recombination. Proc. Natl. Acad. Sci. USA 72, 358-361. Cell 33, 25-35.

219. Holliday, R. (1968). A mechanism for gene conversion in fungi. Genet. Res. Camb. 5, 282-304.

220. Holliday, R. (1968). Genetic recombination in fungi. In Replication and Recombination of Genetic Material, W.J. Peacock and R.D. Brook, eds. (Canberra: Australian Academy of Science), pp. 157-174.

221. Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J., and Stahl, F.W. (1983). The double-strand-break repair model for recombination. Cell 33, 25-35.

222. Lin, F.-L., Sperle, K., and Sternberg, N. (1984). Model for homologous recombination during transfer of DNA into mouse L cells: Role for DNA ends in the recombination process. Mol. Cell. Biol. 4, 1020-1034.

APPENDICES

223. Meselson, M. and Stahl, F.W. (1958). The replication of DNA in E. coli. Proc. Natl. Acad. Sci. USA 44, 671-682.

224. Izant, J.G. and Weintraub, H. (1984). Inhibition of thymidine kinase gene expression by anti-sense RNA: A molecular approach to genetic analysis. Cell 36, 1007-1015.

225. Izant, J.G. and Weintraub, H. (1985). Constitutive and conditional suppression of exogenous and endogenous genes by anti-sense RNA. Science 229, 346-352.

226. Crowley, T.E., Nellen, W., Gomer, R.H., and Firtel, R.A. (1985). Phenocopy of discoidin I-minus mutants by antisense transformation in Dictyostelium discoideum. Cell 43, 633-641.

227. Rosenberg, U.B., Schröder, C., Preiss, A., Kienlin, A., Côté, S., Riede, I., and Jäckle, H. (1986). Structural homology of the product of the Drosophila Krüppel gene with Xenopus transcription factor IIIA. Nature 319, 336-339.

228. Melton, D.A. (1985). Injected anti-sense RNAs specifically block messenger RNA translation in vivo. Proc. Natl. Acad. Sci. USA 82, 144-148.

229. Tomizawa, J-I. and Itoh, T. (1981). Plasmid ColE1 incompatibility determined by interaction of RNA I with primer transcript. Proc. Natl. Acad. Sci. USA 78, 6096-6100.

230. Rodnar, J.W. and Pearson, G.D. (1980). Kinetics of adenovirus replication. II. Initiation of adenovirus DNA replication. Virology 105, 357-370.

231. D'Halluin, J.C. and Milleville, M. (1984). Initiation rate of adenovirus DNA synthesis in infected cells. Bioch. Bioph. Acta 782, 67-75.

232. Yoshikawa, H., Friedmann, T., and Ito, J. (1981). Nucleotide sequences at the termini of ϕ29 DNA. Proc. Natl. Acad. Sci. USA 78, 1336-1340.

233. Escarmis, C. and Salas, M. (1981). Nucleotide sequence at the termini of the DNA of Bacillus subtilis phage ϕ29. Proc. Natl. Acad. Sci. USA 78, 1446-1450.

APPENDIX AI

R-LOOP MAPPING THE INFECTIOUS HEMATOPOIETIC NECROSIS VIRUS

INTRODUCTION

The Infectious Hematopoietic Necrosis Virus (IHNV) is a fish rhabdovirus coded for by a single stranded RNA genome of 10,900 bases (1). The virion consists of 5 proteins (2, 3): N, a nucleocapsid protein; M1, a matrix protein; M2, a matrix protein; G, a surface glycoprotein; and L, a polymerase. In this regard, IHNV appears to be related to the mammalian rhabdovirus prototypes, vesicular stomatitus virus (VSV) and rabies. A sixth non-virion protein designated NV has been described (1) which has not been detected in other rhabdoviruses.

In order to further characterize the relationship of IHNV to the prototypical rhabdoviruses, R-Loop mapping (4, 5) was used to establish the gene order and estimate gene size of the six IHNV genes on the IHNV genome. In this technique partial sequences of cDNAs made from mRNAs for the six proteins in various combinations were hybridized to the IHNV genome and examined by electron microscopy. By mapping each cDNA individually a relative position for each gene on the genome was determined. Analysis of pairs of cDNAs hybridized to the genome revealed the absolute position of each gene relative to the other IHNV genes in the genome.

MATERIALS AND METHODS

cDNAs were prepared as previously described (6). IHNV genomic RNA for R-loop mapping was prepared as described previously (1). Cloned

IHNV cDNA plasmids were prepared as described for large-scale isolation and linearized by cleavage with NdeI (Bethesda Research Laboratories). Formamide (Bethesda Research Laboratories) was deionized by treatment with an AG 501-X8 mixed bed ion-exchange resin (Bio-Rad Laboratories). R-loops were formed as previously described (4) with the following modifications. Linearized plasmid DNA (100 ng) in 18 µl of 78% formamide-5 mM EDTA was denatured at 80°C for 10 min. IHNV RNA (100 ng) was added to the denatured DNA, the solution was adjusted to 20 µl of 70% formamide-300 mM NaCl-4.5 mM EDTA, and R-loops were formed by incubation at 50°C for 12 to 16 hr.

Samples (4 µl) of the R-loop reaction mixture were spread by the urea-formamide method (6). The hyperphase (40 µl) consisted of 4 M urea, 80% formamide, 5 mM EDTA, and cytochrome C at a concentration of 40 µg/ml. The hypophase (20 ml) was 50% formamide. The DNA-protein film was adsorbed to a Parlodion-coated grid, stained with uranyl acetate (0.05M in 0.05M HCl), and rotary shadowed with platinum-palladium (80/20). Grids were examined with a Zeiss EM-10A electron microscope operating at 40 kV. Molecular lengths were measured by a calculator-driven digitizer on photographic prints enlarged to a final magnification of 136,000. DNA molecules of known sequence were used as length standards. Double-stranded DNA and RNA:DNA hybrid duplexes had identical contour lengths under the spreading conditions used in these experiments. IHNV RNA lengths were converted to nucleotides by using a unit length of 10,900 nucleotides and correcting for RNA in RNA:DNA hybrid duplexes.

RESULTS:

Association of each of the cDNA sequences with specific gene products has been demonstrated previously (1, 6). R-loop analysis of NdeI-cut plasmids pL262, pG480, pN144, pM1163, pM219 and pNV58 hybridized to the IHNV genome were used to locate coding sequences of the L, G, N, M1, M2, and NV genes respectively. Plasmid pN419 was also included since it failed to cross hybridize with other plasmids carrying cDNA from the N gene mRNA. Examples of the R-Loops obtained with the respective plasmids is shown in Figure 26. By hybridizing pairs of plasmid DNAs containing different cDNA coding sequences to genomic IHNV (Figure 27), it was possible to orient the respective cDNA sequences relative to each other in the N-ward to L-ward direction as N-M1-M2-G-NV-L.

The information in Table 1 was compiled from the relative positions of the individual cDNA sequences from each other and the absolute positions of each cDNA sequence from an end of the IHNV genome. This information gave the absolute position of each cDNA from the N and L ends of the IHNV genome. The relatively large uncertainty in the position of the genes in the middle of the genome, pG480 and pNV58, suggested that some of these sequences might overlap. This result was not consistent with previous tests for cross-hybridization between the cDNA sequences (1, 6). Measurements of the gap present between each of the doubly-paired cDNA sequences to the genome (Table 2) confirmed the measures in Table 1 and further established with a higher degree of accuracy that the coding sequences in pG480 and pNV58 do not overlap.

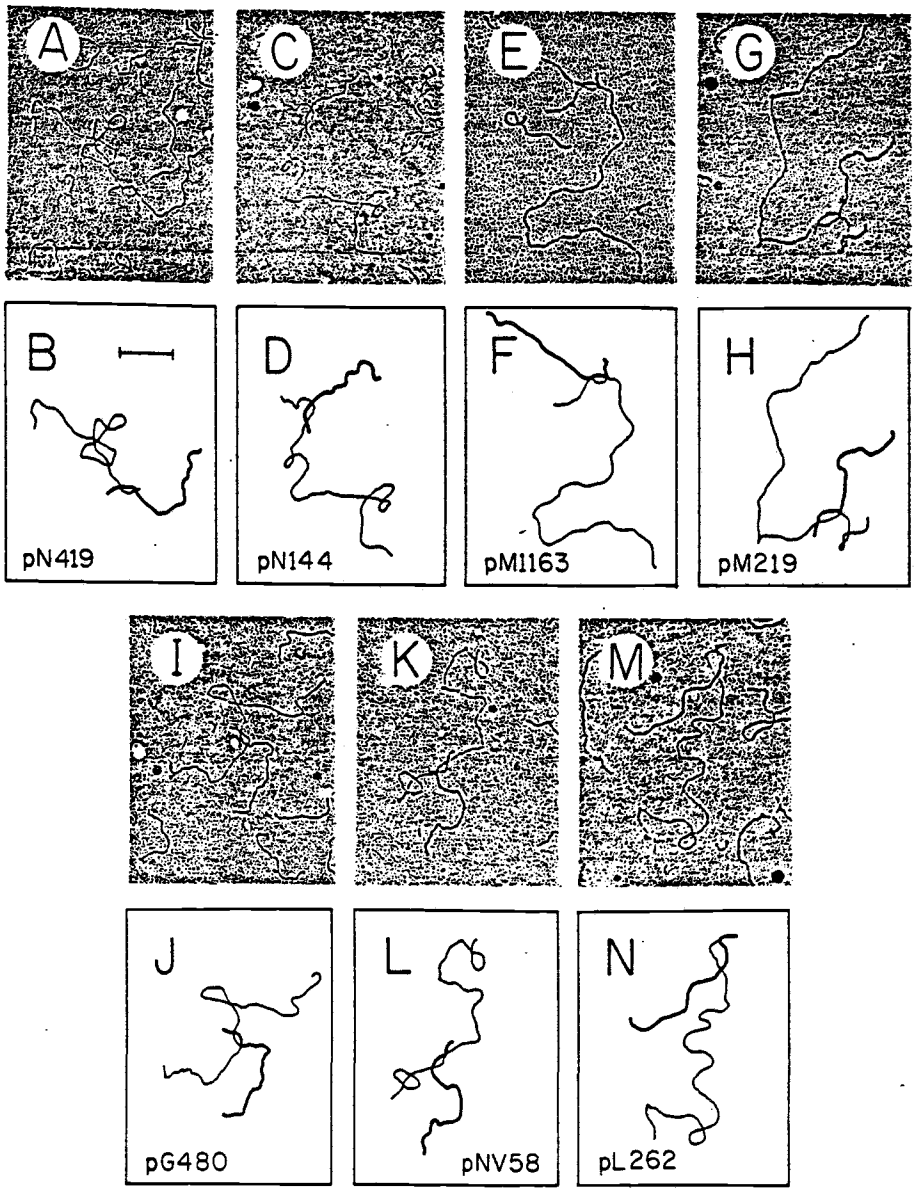


Figure 26: Analysis by electron microscopy of R-loops formed between IHNV genomic RNA and plasmids containing cloned IHNV cDNA sequences.

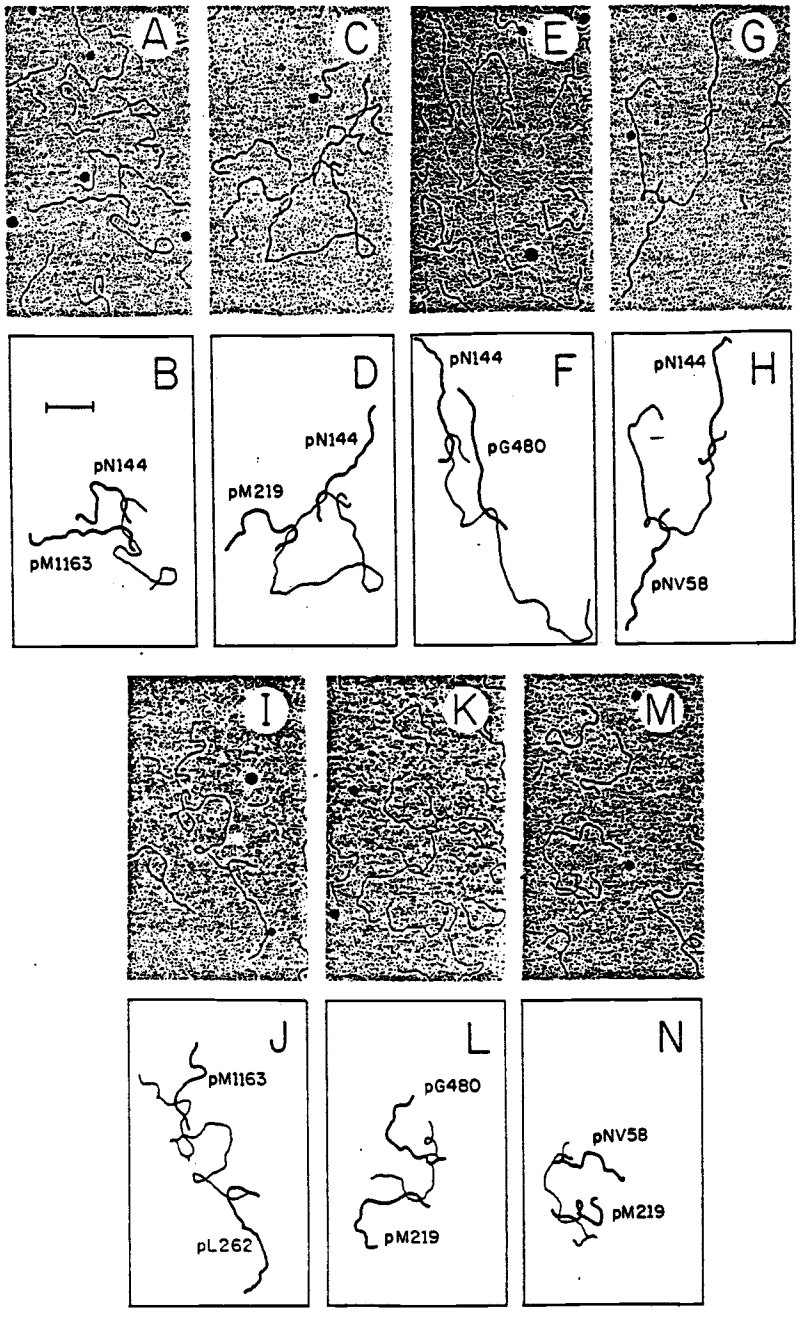


Figure 27: Analysis by electron microscopy of R-loops formed between IHNV genomic RNA and pairs of plasmids containing cloned IHNV cDNA sequences.

Table 1. Measurements of single R-loops and location of hybridizations on the viral genome.

	Measurements of *			Location of **	
PLASMID	SHORT END	LOOP	LONG END	N-WARD NUCLEOTIDE	L-WARD NUCLEOTIDE
pN419	--	490 ± 50	10,410 ± 500	--	490 ± 50
pN144	960 ± 110	440 ± 40	9500 ± 450	960 ± 110	1400 ± 150
pM1163	1520 ± 220	410 ± 60	8970 ± 390	1520 ± 220	1930 ± 285
pM219	2020 ± 250	530 ± 60	8350 ± 470	2020 ± 250	2550 ± 310
pG480	4180 ± 190	400 ± 40	6320 ± 240	4180 ± 190	4580 ± 230
pNV58	4430 ± 240	450 ± 40	6020 ± 270	4430 ± 240	4880 ± 280
pL262	--	710 ± 70	10,190 ± 500	10,190 ± 500	10,900 ± 570

* In nucleotides (± standard deviation) assuming 10,900 nucleotides as the complete length of the genome.

**Nucleotide position (± standard deviation) of R-loop boundaries from O(N-ward end) to 10,900 (L-ward end).

Table 2. Gap measurements for double R-loops

Plasmids	Gap*
pN144 and pM1163	440 ± 30
pN144 and pM219	1140 ± 110
pN144 and pG480	2710 ± 410
pN144 and pNV58	3260 ± 400
pM1163 and pL262	7630
pM219 and pG480	1470 ± 70
pM219 and pNV58	2160 ± 100

*In number of nucleotides (± standard deviation) between adjacent boundaries of the two R-loops.

DISCUSSION

The gene order of the IHNV genome was determined by single and double R-loop mapping to be N-M1-M2-G-NV-L. With the exception of the NV gene, this order is identical to that of the analogous genes of vesicular stomatitis (3') N-NS-M-G-L (5') (7, 8) and rabies (3') N-M1-M2-G-L (5') (9). The 3' to 5' polarity of the IHNV genes was not determined in this study, but is suggested if one assumes that the respective cDNA sequences (with the exception of the pN419 sequence) that were cloned for each gene ultimately came from the 5' end of each gene in the genome. This is a reasonable assumption if one assumes that the 5' end of each gene in the genome corresponds to the 3' end of the message made from it and that the method used for obtaining cDNAs in this experiment enriches in the 3' ends of the message. From this assumption, the 3' to 5' orientation of the IHNV genes is perfectly consistent with the N gene residing at the 3' end of the IHNV RNA and the L gene at the 5' end of the RNA.

Finally, if the another assumption is made that the 3' end of each gene in the genome begins immediately adjacent to the apparent 5' end of the next cDNA as mapped on the genome (a phenomenon established for the vesicular stomatitis virus genome) then the boundaries of the coding sequences for each gene are defined. The boundary of the N gene at the 3' terminus was assumed to be the end of the genome itself. With these assumptions, the estimated size of each of the mRNAs for the genes of IHNV was compared to the known sizes of the mRNAs made from each gene. The size estimated for each gene is well within the range expected for the known molecular weight of each gene

product (see Figure 28) and observed size of mRNA. These results provide additional indirect evidence that the 3' to 5' orientation suggested for the map of the genome is correct. Therefore at least in genomic organization IHNV is closely related to vesicular stomatitis virus and rabies. In addition IHNV has coding sequences for a gene product (NV) not observed in other rhabdoviruses.

REFERENCES

1. Kurath, G.K. and Leong, J.C. (1985). Characterization of infectious hematopoietic necrosis virus mRNA species reveals a nonvirion rhabdovirus protein. J. Virol. 53, 462-468.

2. Hill, B.J. (1975). Physico-chemical and serological characterization of five rhabdoviruses infecting fish. J. Gen. Virol. 27, 369-378.

3. McAllister, P.E., and Wagner, R.R. (1975). Structural proteins of two salmonid rhabdoviruses. J. Virol. 15, 733-738.

4. Chow, L.T. and Broker, T.R. (1981). Mapping RNA:DNA heteroduplexes by electron microscopy. In Electron microscopy in biology, Vol. 1, J.D. Griffith, ed. (New York: John Wiley & Sons, Inc.), pp. 139-188.

5. Moore, C.L. (1981). The electron microscopy of ribonucleic acids. In Electron microscopy in biology, vol. 1, J.D. Griffith ed. (New York: John Wiley & Sons, Inc.). pp. 67-88.

6. Kurath, G., Ahern, K.G., Pearson, G.D., and Leong, J.C. (1985). Molecular cloning of the six mRNA species of Infectious Hematopoietic Necrosis Virus, a fish Rhabdovirus, and Gene Order Determination by R-Loop Mapping. J. Virol. 53, 469-476.

7. Abraham, G., and Banerjee, A.K. (1976). Sequential transcription of the genes of vesicular stomatitis virus. Proc. Natl. Acad. Sci. USA 73, 1504-1508.

8. Ball, L.A., and White, C.N. (1976). Order of transcription of genes of vesicular stomatitis virus. Proc. Natl. Acad. Sci. USA 73, 442-446.

9. Flamand, A. and Delagneau, J.F. (1978). Transcriptional mapping of the rabies virus in vivo. J. Virol. 28, 518-523.

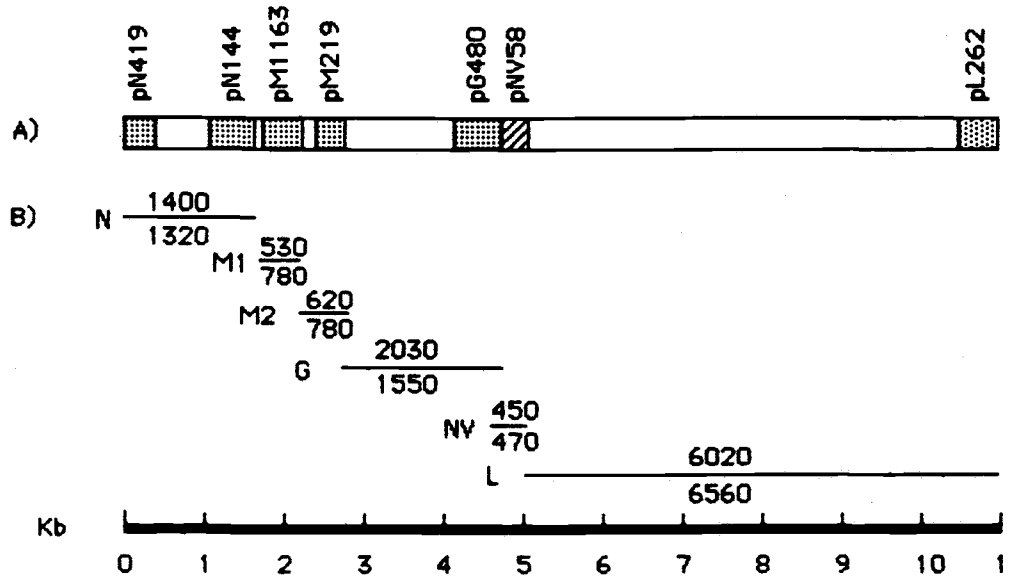


Figure 28: Physical map of the IHNV genome.

APPENDIX AII

HETERODUPLEX CHARACTERIZATION OF THE PROTEASE INHIBITOR I GENE OF POTATO AND TOMATO

INTRODUCTION

In plants, one mechanism for defense against disease, insect attacks, or environmental stress is the induction of protease inhibitors (1-4). The gene coding for proteinase Inhibitor I, whose expression is similarly induced in tomato leaves (Lycopersicum esculentum L. var. Bonny Best) was recently isolated and characterized by Lee et al. (5). In order to contribute to the complete characterization of the gene and provide an evolutionary comparison with a similar gene from potato, heteroduplex analysis was performed on the protease Inhibitor I gene from tomato and potato.

MATERIALS AND METHODS

Heteroduplexes were formed between NdeI-cut proteinase inhibitor cDNA clone pTI-24 and EcoRI-cut genomic subclone pTPI-1 (5) corresponding to the tomato inhibitor I gene and EcoRI-cut pPI-1 corresponding to the potato inhibitor I gene. DNA molecules (100 ng of each species) were denatured together in 10 µl of 80% formamide by heating at 70°C for 10 minutes. Hybridization occurred at room temperature for 1 hour in reaction mixtures (20 µl) containing 50% formamide and 100 mM NaCl. DNA spreading conditions were essentially as described (6). The entire heteroduplex mixture was spread as a hyperphase (40 µl) containing 50% formamide, 100 mM NaCl, 5 mM EDTA

and cytochrome C at a concentration of 40 µg/ml. The hypophase (20 ml) was 25% formamide.

The DNA-protein film was adsorbed to a Parlodion-coated grid, stained with uranyl acetate, (0.05M in 0.05M HCl) and rotary shadowed with platinum-palladium (80/20). Grids were examined with a Zeiss EM-10A electron microscope operating at 40 kV. Molecular lengths were measured on a video display using a Videoplan II image analysis system. pUC8 DNA was used as a length standard. Single-stranded DNA measurements were converted to double-stranded lengths using a factor of 1.16 to correct for compression during spreading.

RESULTS AND DISCUSSION

Heteroduplex analysis of the proteinase inhibitor gene revealed it to be a relatively simple gene, being composed of three exons and two short introns. The close evolutionary relationship between the proteinase inhibitor I genes of potato and tomato was demonstrated by the ability of the cDNA for the tomato to form heteroduplexes with both genomic DNAs. The exon sizes between the two species appeared to be virtually identical and the intron sizes were very close also (Table 3). Comparison of the DNA sequence of the gene and the heteroduplex measures described here provided an excellent check on the accuracy of analysis by heteroduplex mapping.

Table 3. Comparison of Gene Maps of the Tomato and Potato Proteinase Inhibitor I genes generated by Heteroduplex analysis and DNA sequencing.

Feature	Measurement by Heteroduplex Mapping (Potato)	Measurement by Heteroduplex Mapping (Tomato)	Measurement by DNA sequencing
5' Exon	73 ± 14 (52)	71 ± 13 (50)	Not determined
Intron I	466 ± 59 (53)	397 ± 37 (46)	454
Exon II	69 ± 14 (51)	74 ± 14 (50)	90
Intron II	372 ± 49 (86)	356 ± 50 (48)	404
Exon III	349 ± 25 (46)	351 ± 19 (46)	Not determined

*Measures are ± standard deviation with number of measurements in parentheses.

REFERENCES

1. Plunkett, G., Senear, D.F., Zuroske, G., and Ryan, C.A. (1982). Proteinase inhibitors I and II from leaves of wounded tomato plants: Purification and properties. Arch. Biochem. Biophys. 213, 463-472.

2. Melville, J.C. and Ryan, C.A. (1972). Chymotrypsin inhibitor I from potatoes. J. Biol. Chem. 247, 3445-3453.

3. Bryant, J., Green, T.R., Gurusaddaiah, T., and Ryan, C.A. (1976). Proteinase inhibitor II from potatoes: Isolation and characterization of its protomer components. Biochemistry 15, 3418-3424.

4. Brown, W.E. and Ryan, C.A. (1984). Isolation and characterization of a wound induced trypsin inhibitor from alfalfa leaves. Biochemistry 23, 3418-3422.

5. Lee, J.S., Brown, W.E., Pearce, G., Dreher, T.W., Ahern, K.G., Pearson, G.D., and Ryan, C.A. (1986). Complete sequence of a wound-inducible proteinase inhibitor I gene in tomato. J. Biol. Chem., submitted.

6. Chow, L.T. and Broker, T.R. (1981). Mapping RNA:DNA heteroduplexes by electron microscopy. In Electron microscopy in biology, Vol. 1, J.D. Griffith, ed. (New York: John Wiley & Sons, Inc.), pp. 139-188.

APPENDIX AIII

CHARACTERIZATION OF THE BOVINE PROTHROMBIN GENE BY ELECTRON MICROSCOPY

INTRODUCTION

Serine proteases are a family of structurally-related proteins involved in a variety of physiological processes including digestion, fertilization, the vertebrate immune response, and blood coagulation (for review see (1)). The clotting of blood involves at least seven serine proteases participating in two different pathways (2). Prothrombin is a plasma glycoprotein of M_r 65,000 (3) that has been detected in many species including lamprey (4). The complete amino acid sequences of bovine (3) and human (5, 6) prothrombin have been determined. The carboxy-terminal region of prothrombin contains the catalytic domain that shares amino acid sequence homology with trypsin (2).

Five phage were isolated from a genomic phage library that contain overlapping DNA spanning 42.4 kb of bovine genomic DNA. The complete coding sequences for the bovine prothrombin gene resides within this DNA segment (7). In order to further characterize the molecular structure of the bovine prothrombin gene and establish evolutionary relationships to other genes, an electron microscopic mapping of heteroduplexes formed between the phage genomic sequences and the cDNA sequences of the bovine prothrombin gene was performed.

MATERIALS AND METHODS

Construction of a genomic bovine phage library, isolation of the bovine prothrombin gene from the library, and characterization of the genomic clones has been described previously (7). The cloned bovine prothrombin cDNAs used in this study were constructed by MacGillivray and Davie (8). Heteroduplexes were formed between EcoRI-cut or PstI-cut prothrombin cDNA plasmids (pBII102 or pBII111) and DNA from either λ clones containing bovine genomic sequences (λBII1, λBII2, or λBII3) or from appropriately cleaved subclones of bovine genomic sequences (pHE2, pBGII3, or pGBII21; see Table 4).

DNA molecules (100 ng of each species) were denatured together in 10 µl of 80% formamide (Bethesda Research Laboratories) by heating at 70°C for 10 minutes. Hybridization occurred at 37°C for 1 hour in reaction mixtures (20 µl) containing 50% formamide and 200 mM NaCl. DNA spreading conditions were essentially as described (9). The entire heteroduplex mixture was spread as hyperphase (40 µl) containing 50% formamide/100 mM NaCl/5 mM EDTA, 100 ng of DNA added as a length standard, and cytochrome C at a concentration of 40 µg/ml. The hypophase (20 ml) was 20% formamide. The DNA-protein film was adsorbed to a Parlodion-coated grid, stained with uranyl acetate (0.05M in 0.05M HCl), and rotary-shadowed with platinum-palladium (80/20). Grids were examined with a Zeiss EM-10A electron microscope operating at 60 kV. Molecular lengths were measured on a video display using a Videoplan II image analysis system. Single-stranded DNA measurements were converted to double-stranded lengths using the factor 1.16 to correct for compression during spreading.

RESULTS

The positions of introns within the prothrombin gene were determined by heteroduplex analysis with the electron microscope. Heteroduplexes were formed between prothrombin cDNAs and either recombinant phage λ DNAs or sub-clones derived from the phage DNA (Table 4). Typical heteroduplexes are shown in Figure 29. Table 5 summarizes the salient features of the prothrombin gene as determined in the electron microscopic analysis. It is a complex gene, encoded by at least 14 exons extending over 15 kb of DNA. Intron sizes ranged from H and M that were less than 100 bp in length, but detectable, up to intron L which was almost 7000 bp. The exons varied in size from 28 to 317 bp. The total length of exons measured by electron microscopy was 1941 bp, accounting within experimental error for all of the cDNA sequence that has previously been determined (1998 bp).

The 5' to 3' orientation of the gene was determined from the observation of heteroduplexes between genomic prothrombin DNA and the cDNA pBII102 which lacks 5' cDNA sequences. These heteroduplexes did not contain intron loops A through D. Several restriction endonuclease cleavage sites within the gene were also mapped by electron microscopy (Table 6). This information allowed the alignment of the gene with the restriction map of the clone bovine genomic sequences. Based on these measurements, the 5' end of the gene was positioned at coordinate 14.3 and the 3' end at coordinate 29.2, a distance of 14.9 kb.

Regions of self-complementarity within the λ genomic sequences

Table 4. Subclones of cloned phage DNA used in the heteroduplex analysis.

Subclone	Flanking restriction sites		Location of cloned fragment[a]
pHE2	HindIII	EcoRI	11.7 to 16.6
pBGII21	BamHI	Sau3AI	15.4 to 20.4
pBGII3	BamHI	BamHI	15.4 to 21.9

[a]Location is given in kbp relative to the prothrombin map in reference 7.

Figure 29: Analysis by electron microscopy of heteroduplexes formed between cloned bovine genomic DNA (λB113) and cloned prothrombin cDNA (pBII111).

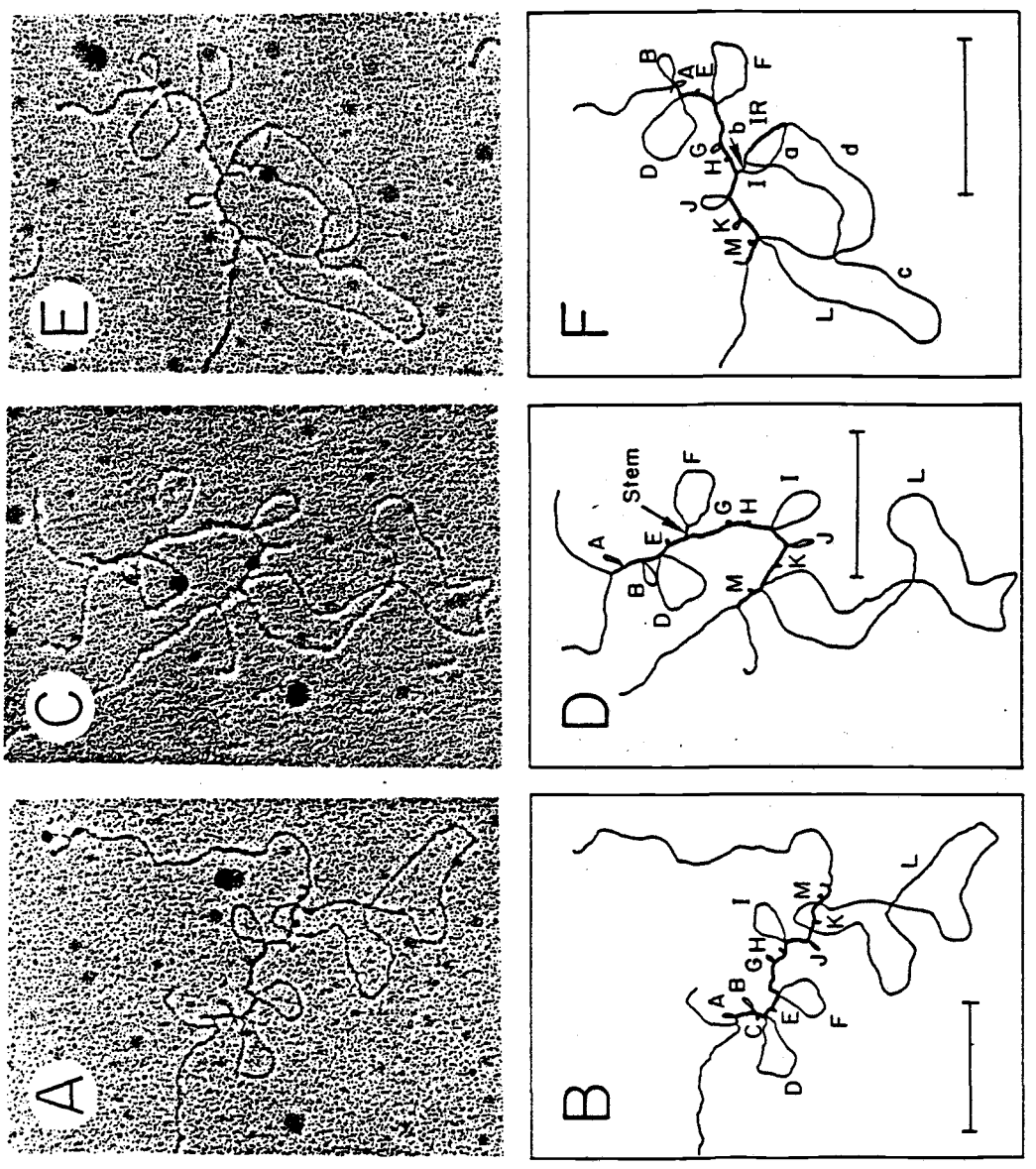


Figure 29:

Table 5. Lengths of exons and introns in the prothrombin gene.

Exon	Length[a]	Length[b]	Intron	Length[a]	Length[b]
1	98 ± 14 (43)	94	A	261 ± 46 (48)	342
2	168 ± 18 (46)	164	B	601 ± 62 (49)	---
3	28 ± 8 (52)	25	C	170 ± 39 (68)	227
4	53 ± 13 (68)	51	D	157 ± 73 (50)	---
5	103 ± 13 (56)	105	E	112 ± 19 (50)	98
6	139 ± 15 (54)	137	F	1381 ± 99 (54)	---
7	317 ± 26 (60)	315	G	235 ± 23 (58)	293
8	137 ± 15 (41)	135	H	<100	75
9	117 ± 16 (41)	127	I	1055 ± 94 (51)	---
10	170 ± 19 (56)	168	J	397 ± 46 (48)	---
11	159 ± 19 (56)	174	K	216 ± 29 (56)	242
12	160 ± 17 (62)	182	L	6940 ± 255 (41)	---
13	65 ± 10 (90)	71	M	<100	135
14	227 ± 17 (39)	266			

[a]Expressed as mean ± standard deviation in base pairs where the number of measurements are in parentheses.
[b]Length in base pairs as determined from DNA sequence analysis.

Table 6. Location of restriction sites within the prothrombin gene

Restriction Site	Coordinate[a]	Location in gene[b]
BamHI	15.4	<100 bp 5' from exon 3 in intron B
SstI	16.4[c]	673 ± 69 (38) bp 3' from exon 4 in intron D
Sau3AI	20.4	798 ± 69 (14) bp 3' from exon 8 in intron I
BamHI	21.9	124 ± 17 (48) bp 3' from intron K in exon 12

[a]See Table 4.

[b]Expressed as mean ± standard deviation where the number of measurements are in parentheses.

[c]SstI-cut pHE2.

were observed, an indication that inverted repeat sequences were present. A stem at the base of intron F is visible in Figure 25C. Figure 25E illustrates inverted repeat (IR) sequences shared between introns I and L. Table 7 summarizes the measurements of these features. Although the stem in intron F is roughly the size of an Alu-like element already described in the bovine genome (10), further characterization will require nucleotide sequence analysis. Occasionally, stem-like structures were observed at the bases of introns D and I, but were not measured.

From the data presented in Table 7, the position of the IR sequences within introns I and L was established. Direct measurement of the loop between IR sequences in "snapback" genomic DNA gave a separation of 5692 bp. This corresponded best to the arrangement 5'-b-10-J-11-K-12-c-3', a distance of 5684 bp. Therefore, the order in intron I is 5'-a-IR-b-3' and in intron L it is 5'-c-IR-d-3'. At least one other inverted repeat located between intron L and genomic sequences distal to the 3' end of the gene was visualized, but not measured.

DISCUSSION

Electron microscopic heteroduplex analysis between cloned genomic DNA and cloned cDNA demonstrated that the bovine prothrombin gene is approximately 15 kb in length and contains 14 exons and 13 introns. The cDNA clones used in this study contain only 5 bp of 5' untranslated sequence, so the presence of additional intron(s) in the 5' untranslated region of the gene cannot be excluded. It is also

Table 7. Lengths and locations of inverted repeat sequences within introns.

Feature[a]	Length[b]
Stem	119 ± 26 (9)
IR	387 ± 27 (45)
a	586 ± 57 (40)
b	129 ± 23 (45)
c	4456 ± 186 (46)
d	2117 ± 109 (30)
Loop[c]	5692 ± 234 (37)

[a]See Figure 25.

[b]Expressed as mean ± standard deviation in base pairs where the number of measurements are in parentheses.

[c]Separation between IR sequences in "snapback" DNA; see text.

possible that other small introns were not detected by the heteroduplex analysis because of the difficulty in visualizing introns smaller than 100 bp. Therefore the number of exons and introns are minimum estimates. Since there are few introns less than 80 bp in size (11), however, it is probable that there are no other introns in the prothrombin gene.

The prothrombin gene is 15 kb in size, yet its cytoplasmic mRNA measures only 2150±100 nucleotides in size, including a poly A tail. Thus, 86% of the prothrombin gene is comprised of introns. The numbers and sizes of the exons are similar between the bovine and human prothrombin genes within the regions of both that have been characterized (7). The two exons at the 3' end of the bovine gene are separated from the rest of the exons by a large intron that is 6.9 kb in size. The presence of a large intron in the corresponding position of the human gene may be why the 3' end of the human gene has not been characterized to date. There are no obvious changes apparent at the electron microscopic level in the sizes of any of the exons in the two genes.

The detection of repetitive elements in an inverted orientation within and between introns of the bovine prothrombin gene were an interesting by-product of this investigation. Repetitive sequences of the Alu I type are known to be located within the larger introns of the human gene (12). A repetitive element similar to the human Alu element has been isolated from the bovine genome (10) and is 120 nucleotides in length. This is much shorter than the 300 base pairs constituting the human Alu repeat (13) and may explain in part the difference in length of the introns between the two species. An

alternative explanation may be either the deletion of DNA from the bovine introns, or the insertion of DNA into the human introns.

REFERENCES

1. Neurath, H. (1984). Evolution of proteolytic enzymes. Science 224, 350-357.

2. Jackson, C.M. and Nemerson, Y. (1980). Blood coagulation. Ann. Rev. Bioch. 49, 765-811.

3. Magnussen, S., Peterson, T.E., Sottrup-Jenson, L. and Claeys, H. (1975). In Proteases and Biological Control, E. Reich, D.B. Rifkin, and E. Shaw, Eds. (Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratories) pp. 123-149.

4. Doolittle, R.F., Oncley, J.L. and Surgenor, D.M. (1962). Species differences in the interaction of thrombin and fibrinogen. J. Biol. Chem. 237, 3123-3127.

5. Butkowski, R.J., Ekion, J., Downing, M.R., and Mann, K.G. (1977). Primary structure of human prethrombin 2 and α thrombin. J. Biol. Chem. 252, 4942-4957.

6. Walz, D.A., Hewett-Emmett, D., and Seegers, W.H. (1977). Amino acid sequence of human prothrombin fragments 1 and 2. Proc. Natl. Acad. Sci. USA 74, 1969-1972.

7. Irwin, D.M., Ahern, K.G., Pearson, G.D., and MacGillivray, R.T.A. (1985). Characterization of the Bovine Prothrombin Gene. Biochemistry 24, 6854-6861.

8. MacGillivray, R.T.A. and Davie. E.W. (1984). Characterization of the bovine prothrombin mRNA and its translation product. Biochemistry 23, 1626-1634.

9. Chow, L.T. and Broker, T.R. (1981). Mapping RNA:DNA heteroduplexes by electron microscopy. In Electron microscopy in biology, Vol. 1, J.D. Griffith, ed. (New York: John Wiley & Sons, Inc.), pp. 139-188.

10. Watanabe, Y., Tsukada, T., Notake, M., Nakanishi, S., and Numa, S. (1982). Structural analysis of repetitive DNA sequences in the bovine corticotropin-β-lipoprotein precursor gene region. Nucl. Acids Res. 10, 1459-1469.

11. Weiringa, N., Hofer, E., and Weissmann, C. (1984). A minimal intron length but no specific internal sequence is required for splicing the large rabbit β-globin intron. Cell 37, 915-925.

12. Degen, S.J.F., MacGillivray, R.T.A., and Davie, E.W. (1983). Characterization of the complementary deoxyribonucleic acid and gene coding for human prothrombin. Biochemistry 22, 2087-2097.

13. Jelinek, W.R. and Schmid, C.W. (1982). Repetitive sequences in eukaryotic DNA and their expression. Ann. Rev. of Bioch. 51, 813-844.

APPENDIX AIV

A FAST MODULAR DNA SEARCH PROGRAM FOR ANALYSIS OF LONG SEQUENCES ON APPLE II COMPUTERS

ABSTRACT

An extremely fast DNA search program (ASAP) in modular form for analyzing up to 40,000 nucleotides is described. The program package, designed for the Apple II series of computers, requires at least one disk drive, a monitor, and 64 kilobytes of random access memory (k) of memory. It is written very compactly in BASIC and machine language so as to optimize memory usage. Nine subprograms are described. They enable input of sequences manually, proof-editing functions, printing, duplication of sequences, renaming of files, deletion of files, user-defined sequence searches, and pre-programmed restriction enzyme searches. A recent complete restriction analysis of the Ad2 genome (35,937 bp) for 47 enzymes required only 1 minute 54 seconds, less than 3 seconds per enzyme.

INTRODUCTION

The analysis of DNA sequences, a laborious task by hand, has been greatly facilitated by the development of software for microcomputers commonly found in researchers laboratories and homes. An increasing number of such programs for Apple and IBM computers has been described (1-11). The limiting amount of memory available on microcomputers is a major obstacle to DNA sequences analysis. A variety of approaches have been utilized in order to maximize the speed of operation,

including programming in compiled languages such as Pascal or FORTRAN. Since compiled programs require more memory than non-compiled, speed is often achieved at the expense of memory usage.

The programs described here are an attempt to maximize both memory usage and speed. They have been written in BASIC and 6502 assembler and utilize PRONTO-DOS (© Tom Weishaar) for fast disk access and memory rearrangement. With this format, 40,000 nucleotides can be analyzed at once on a 64k Apple II+, Apple IIe, or Apple IIc computer. Unlike other machine language programs that utilize vast amounts of memory, the BASIC programs described here utilize two machine language routines that together occupy less than 0.1k of memory. Nonetheless, searches are incredibly fast and flexible, requiring 3 seconds or less for files as large as 40,000 nucleotides.

HARDWARE REQUIREMENTS:

Hardware requirements for this system include an Apple II+, IIe, or IIc with at least 64k of random access memory (RAM), at least one disk drive, and a monitor. Two options helpful for operation include a second drive and a printer capable of printing 80 columns. No other hardware or software is required other than the one program disk. As well as is known, computer add-ons such as cards will not affect operation of the program, but non-standard devices for driving printers may affect output.

SOFTWARE DESCRIPTION

The program is driven by compact Applesoft BASIC routines that call 6502 assembler routines for searching or editing. The program operating system is PRONTO-DOS (© Tom Weishaar) and is included in the package. Disks used for storing sequences typed in by the user must be formatted in DOS 3.3. The program is completely menu driven and disk based. Starting up the program automatically loads the machine language routines so that they may be called upon immediately. Sequences are stored and loaded as binary files. In the scheme described here, A=0001, G=0010, C=0100, T=1000. See below for instructions on converting GenBank files to ASAP format. Each of the individual programs are loaded independently through the menu. Loading is rapid, requiring 2-5 seconds maximum. Graphic output is minimized in the interest of program compaction.

PROGRAM OPERATION

Control of the programs is through the main menu which begins operation after the system is booted from the program disk. Users operate the programs through the selection of options that are usually addressed with 1 or 2 key presses. There are two programs for searching sequence files. The first, SEARCH FOR A BASE SEQUENCE, allows the user to examine a previously stored sequence file. Operation is user-controlled. First the file is selected by the user and loaded by the program. A mini-menu then allows the options of searching for a sequence of the user's choice, determining G C content,

loading a different sequence, displaying a help screen, or returning to the main menu.

The search option in the mini-menu allows searches for sequences up to 125 nucleotides long and may consist of any combination of A, C, G, T, R (A or G), Y (T or C), S (G or C), W (A or T), M (A or C), K (G or T), V (A or C or G), H (A or C or T), D (A or G or T), B (C or G or T), or N (A or C or G or T). Thus variable sequences (as found in many consensus sequences) are allowed in the search.

The main advantage to the numerical strategy employed in encoding the nucleotides, however, is that all searching can be completed in one pass by the use of binary arithmetic. Searches for sequences that are variable in several positions are not perceptibly longer than searches for absolute sequences. This is a major difference between this software and other packages currently available that have a geometric progression in time with each variable position added to a search sequence. Program output includes sequence found, position of sequence found, and finally the number of such sequences found. Asymmetric sequences are analyzed automatically, first in the (+) strand and then in the (-) strand.

The other search program is SEARCH FOR RESTRICTION SITES. It contains all of the commercially available restriction enzymes available in the United States as of February 1, 1985. This constitutes 86 specificities. The enzymes are divided into three categories based on the number of bases that they recognize: fewer than 6, 6 or more, or asymmetric sequences of all lengths. Again, asymmetric sequences are automatically searched first in the (+) strand and then in the (-) strand. EcoRI (GAATTC) and XmnI

(GAANNNNTTC) are counted as recognizing six bases whereas HpaII (CCGG), NciI CC(C/G)GG, or AvaI (CYCGRG) are counted as recognizing 4, 4.5, or 5 bases respectively. Users have the option of searching for all the enzymes automatically, or being prompted for each one. Output can be directed to the printer or monitor.

The remaining 7 programs are for manipulation of the sequence files. INPUT A SEQUENCE enables users to type in their own sequences. Sequence files up to 40,000 nucleotides long are permitted. The program is equipped to beep if a letter other than A, C, G, or T is entered. An editing feature enables simple corrections from the point of the error onward. All sequences input are saved on diskettes initialized in DOS 3.3 in a name chosen by the user. Protection is built-in to prevent overwriting one sequence with another of the same name.

PROOF/EDIT provides a means for checking the accuracy of an existing sequence file by comparing input from a user typing in any part of the file with the stored file. When a discrepancy is found, it is sounded with a beep, freeing the typist from having to routinely view the screen. The PROOF program is merged with the EDIT program which enables editing as errors are found. The EDIT program can also be accessed independently of the PROOF program. EDIT is an elaborate editor for sequence files and enables substitutions, insertions, or deletions so long as the maximum length of 40,000 bases or the minimum length of 1 base is not exceeded. Upon completion of editing, the corrected file is stored in the place of the original file.

MERGE will sequentially link together multiple sequence files into one, so long as the maximum length of 40,000 nucleotides is not

exceeded. The merged file is then saved on disk according to a user-supplied name. This program is designed for typists who enter sequence files in segments to merge them together into one file later. PRINTFILE will send the complete output of a sequence file to a printer or monitor. The format is 80 columns wide for a printer or 40 columns wide for a monitor.

DUPLICATE A SEQUENCE FILE makes backup copies of sequence files onto disks formatted in DOS 3.3. DUPLICATE A SEQUENCE FILE has built-in safeguards to prevent overwriting files with the same name, or overfilling a disk. RENAME A SEQUENCE FILE allows users to change file names as long as the following rules for nomenclature are followed: No two files can have the same name; names must begin with a letter; names must be 25 characters or less in length. Any names which are entered that do not follow these rules are not accepted ;by the program. DELETE A SEQUENCE FILE drops the selected sequence file from the program and disk catalog, freeing up the space on the disk for other uses.

CONVERSION OF GENBANK FILES

The national DNA sequence bank, GenBank, operated by the National Institutes of Health has over 4,000,000 bases within its files. The following instructions are provided for the conversion of GenBank sequences into a format that ASAP can recognize so that any GenBank file can be analyzed by ASAP. The description that follows is for raw GenBank files and may or may not work on GenBank files translated for other programs, such as MicroGenie ®.

GenBank sequence files as downloaded by modem over the telephone lines are typically full of errors, due to extraneous noise. It is recommended that files be downloaded multiple times and compared before being translated here. GenBank sequences are transmitted with information in addition to the DNA sequences. This information should be stripped away (a word processor will suffice for this translation). The raw GenBank sequences are sent as the letters A, G, C, T, N, R, or Y. The bases A, G, C, T should be converted to the numbers 1, 2, 4, or 8 respectively. ASAP has no means of dealing with N, R, or Y and it is recommended that files be terminated at the point of occurrence of these letters. Only relatively few GenBank files have N, or R, or Y in them. If so desired, however, N, R, or Y can be encoded by a 0 in the file without catastrophic effects but the 0 will not be recognized as such. All files must terminate after the last base with the number 255. To have the file be recognized by the program, the file size and name must be entered in a file called S[E]EQFILES ([E] is a control-E). S[E]EQFILES contains all of the catalog information of the sequence disk as follows. First comes a number that tells how many sequence files are on the disk. This is followed by a return and a number that tells how many bases are in the first file. This is followed by a carriage return [CR] and the name of the first file. This is then repeated for each file. The overall scheme then is #Files[CR]#Bases [CR] File Name [CR] # Bases [CR] File Name [CR] etc. until the last name. The entered sequence file must fit into this scheme in the same manner. After this information is correctly stored in S[E]EQFILES, the ASAP program can act on the file.

DISCUSSION

The programs described here have been designed with compactness, speed of operation, and functionality in mind. The BASIC programs are modular and are each less than 5.3k long. The two machine language routines put together occupy less than 0.1k of total memory. Since the main menu controls program operation, users are not required to learn complicated commands or instructions. Individual program design is modular so that expansion by addition of other program modules will be relatively simple. By making the operation disk based, no additional memory will be required as more programs are added later. The ability to increase software capabilities without requiring additional memory is a major advance in the design of programs for microcomputers. The speed of operation, achieved through machine language routines for searches and PRONTO-DOS for file loading, is comparable with any softwarae currently available on the market. Lastly, since these programs are written in BASIC, which comes as standard equipment with Apple II computers, no additional software is required to run these programs.

REFERENCES

1. Gingeras, T.R., and Roberts, R.J. (1981). Computer assisted methods for nucleic acid sequencing. In Genetic Engineering, Vol. 3, J. Setlow and A. Hollaender, Eds. Plenum Press, New York. pp. 319-338.

2. Staden, R. (1980). A new computer method for the storage and manipulation of DNA gel reading data. Nucl. Acids Res. 8, 3673-3694.

3. Korn, L.J., Queen, C.L., and Wegman, M.N. (1977). Computer Analysis of nucleic acid regulatory sequences. Proc. Natl. Acad. Sci. USA 74, 4401-4405.

4. Larson, R. and Messing, J. (1982). Apple II software for M13 shot gun DNA sequencing. Nucl. Acids Res. 10, 39-49.

5. Fristensky, B., Lis, J. and Wu, R. (1982). Portable microcomputer software for nucleotide sequence analysis. Nucl. Acids Res. 10, 6451-6471.

6. Malthiery, B., Bellon, B., Giorgi, D. and Jacq, B. (1984). Apple II pascal program for molecular biologists. Nucl. Acids Res. 12, 569-579.

7. Queen, C.L., and Korn, L.J. (1984). A comprehensive sequence analysis program for the IBM personal computer. Nucl. Acids Res. 12, 581-599.

8. Schwindinger, W.F. and Warner, R. (1984). DNA sequence analysis on the IBM-PC. Nucl. Acids Res. 12, 601-604.

9. Lagrimini, L.M., Brentano, S.T. and Donelson, J. (1984). A DNA sequences analysis package for the IBM personal computer. Nucl. Acids Res. 12, 605-614.

10. Paolella, G. (1985). A fast DNA sequence handling program for Apple II computer in BASIC and 6502 assembler. CABIOS 1, 43-49.

11. Larson, R. and Messing, J. (1983). Apple II computer software for DNA and protein sequence data. DNA 2, 31-35.

APPENDIX AV

CONSTRUCTION OF CLONES

INTRODUCTION

The purpose of this section is to provide a comprehensive description of clones constructed during this work that were not used in the experiments described above. Clones dl17, dl12, dl18, dl20, dl21, dl30, dl31, dl36, dl67 (collectively referred to as pdl(7-67) here), 7, and 16z have been described previously (1). In the diagrams that follow, the following abbreviations are used: tetracycline resistance (T^r); ampicillin resistance (A^r); kanamycin resistance (K^r).

Figure 30: Construction of clone 11. Clone 11 was made by cutting XD7 with SalI, treating with BAL31 nuclease and selecting for loss of tetracycline resistance. Clone 11 was made in an attempt to delete a suspected background origin sequence suggested by electron microscopic studies of XD7 replication. It contains an asymmetrical 1.6 kb deletion spanning from about position 600 to 2200 on the pBR322 numbering map.

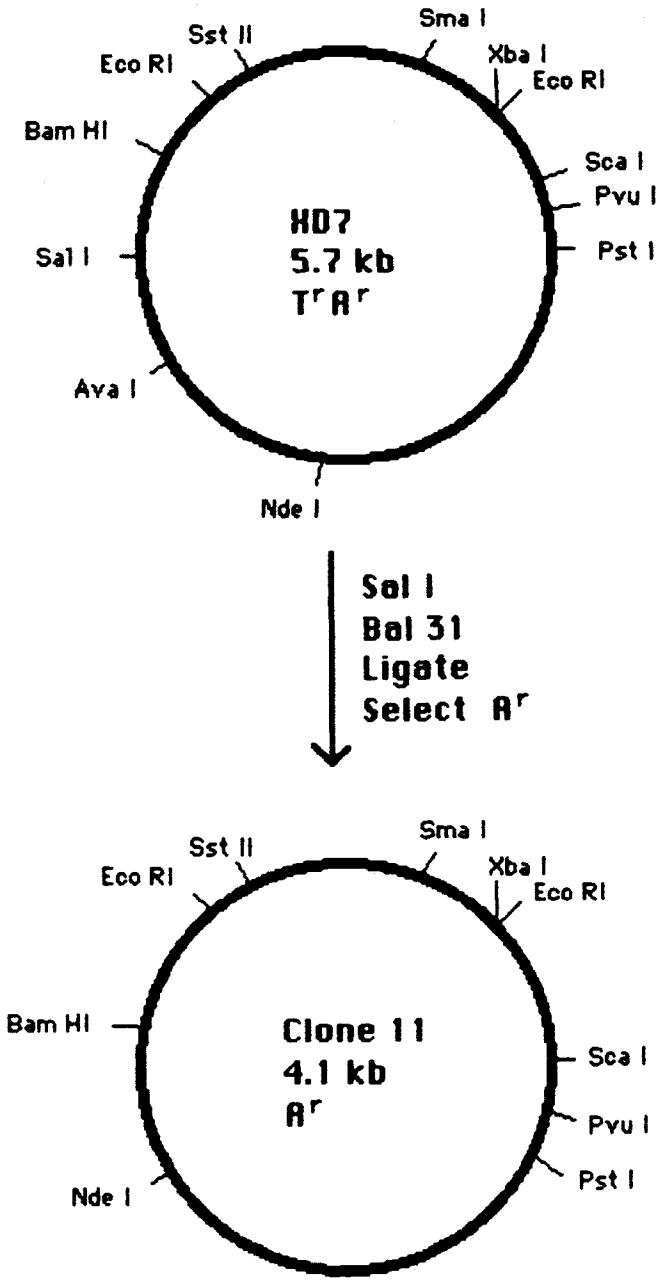

Sma I
Sst II
Xba I
Eco RI
Eco RI
Bam HI
Sca I
Pvu I
HD7
5.7 kb
TrAr
Sal I
Pst I
Ava I
Nde I
Sal I
Bal 31
Ligate
Select Ar
Sma I
Sst II
Xba I
Eco RI
Eco RI
Bam HI
Clone 11
4.1 kb
Ar
Sca I
Pvu I
Pst I
Nde I

Figure 30

Figure 31: Construction of clone 11y. Clone 11y is a derivative of clone 11 made by cutting clone 11 with EcoRI and religating. Clone 11y as a result contains no adenovirus DNA but has the same 1.6 kb deletion in the pBR322 sequences as clone 11.

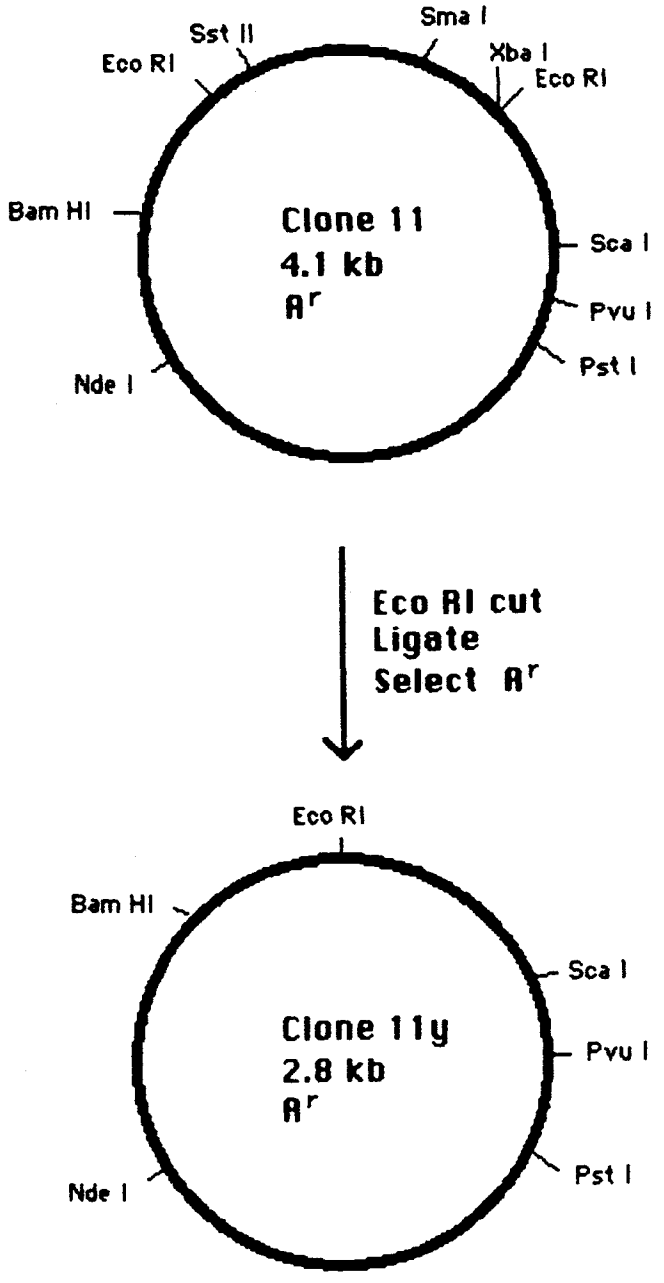

Sma I
Sst II
Xba I
Eco RI
Eco RI
Bam HI
Clone 11
4.1 kb
Ar
Sca I
Pvu I
Pst I
Nde I
Eco RI cut
Ligate
Select Ar
Eco RI
Bam HI
Sca I
Clone 11y
2.8 kb
Ar
Pvu I
Pst I
Nde I

Figure 31

Figure 32: Construction of clone 17. Clone 17 is a derivative of clone 11 made by cutting clone 11 with EcoRI, religating, and screening for inversion of the EcoRI fragment. Clone 17 thus has the 1.3 kb EcoRI adenovirus fragment in an inverted orientation relative to clone 11.

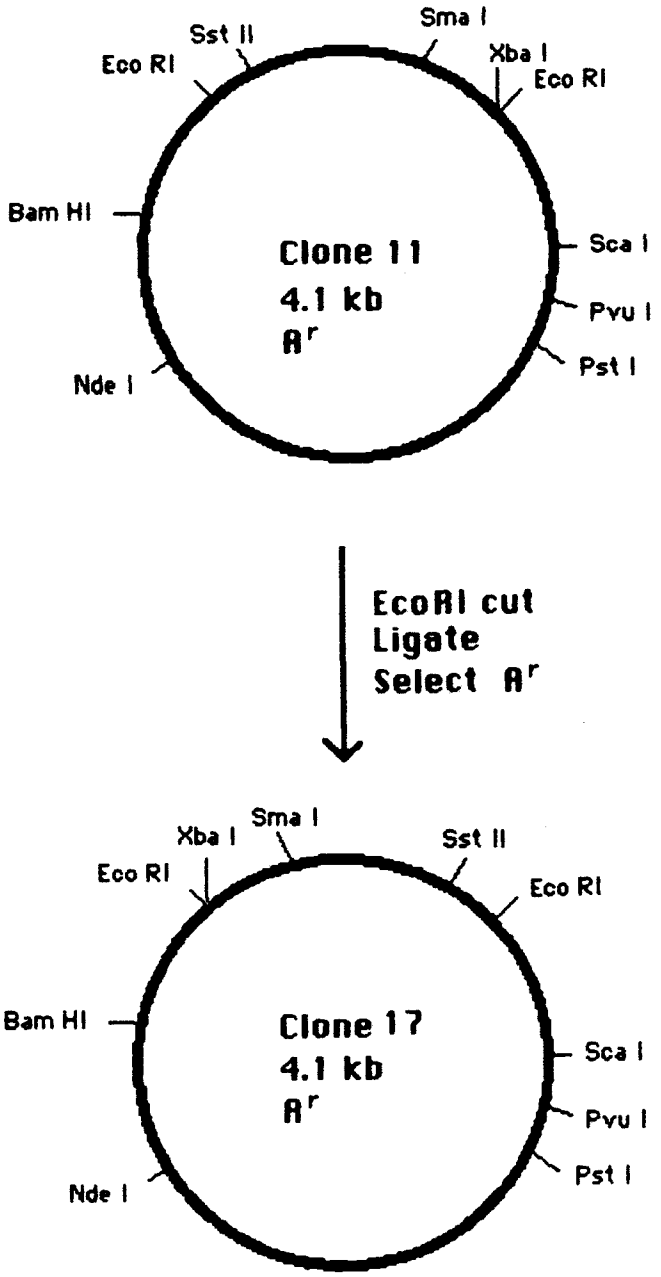

Sma I
Sst II
Xba I
Eco RI
Eco RI
Bam HI
Clone 11
4.1 kb
A^r
Sca I
Pvu I
Pst I
Nde I
EcoRI cut
Ligate
Select A^r
Sma I
Xba I
Sst II
Eco RI
Eco RI
Bam HI
Clone 17
4.1 kb
A^r
Sca I
Pvu I
Pst I
Nde I

Figure 32

Figure 33: Construction of clone 4. Clone 4 was made by an accidental nuclease activity at the BalI site of clone 17. It contains a 1000 base pair deletion spanning approximately from the BalI site to the pBR322 HindIII site.

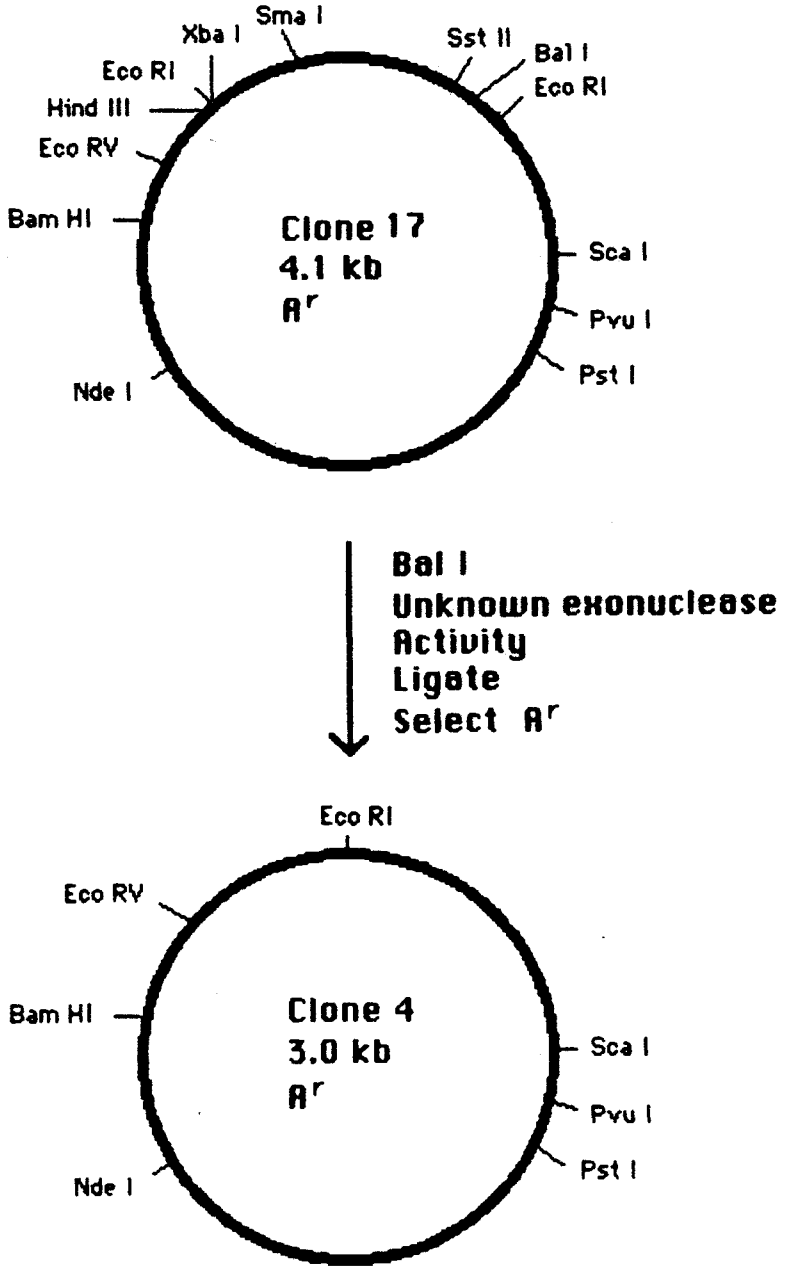

Sma I
Xba I
Sst II
Bal I
Eco RI
Eco RI
Hind III
Eco RV
Bam HI
Clone 17
4.1 kb
Ar
Sca I
Pvu I
Pst I
Nde I
Bal I
Unknown exonuclease
Activity
Ligate
Select Ar
Eco RI
Eco RV
Bam HI
Clone 4
3.0 kb
Ar
Sca I
Pvu I
Pst I
Nde I

Figure 33

Figure 34: Construction of p7x. p7x was constructed at the same time as p44x with the same intent. It was made by inserting an XhoI linker into the blunted EcoRI site of clone 7. DNA sequence analysis has not been performed, but the presence of an XhoI site is indicated by restriction mapping.

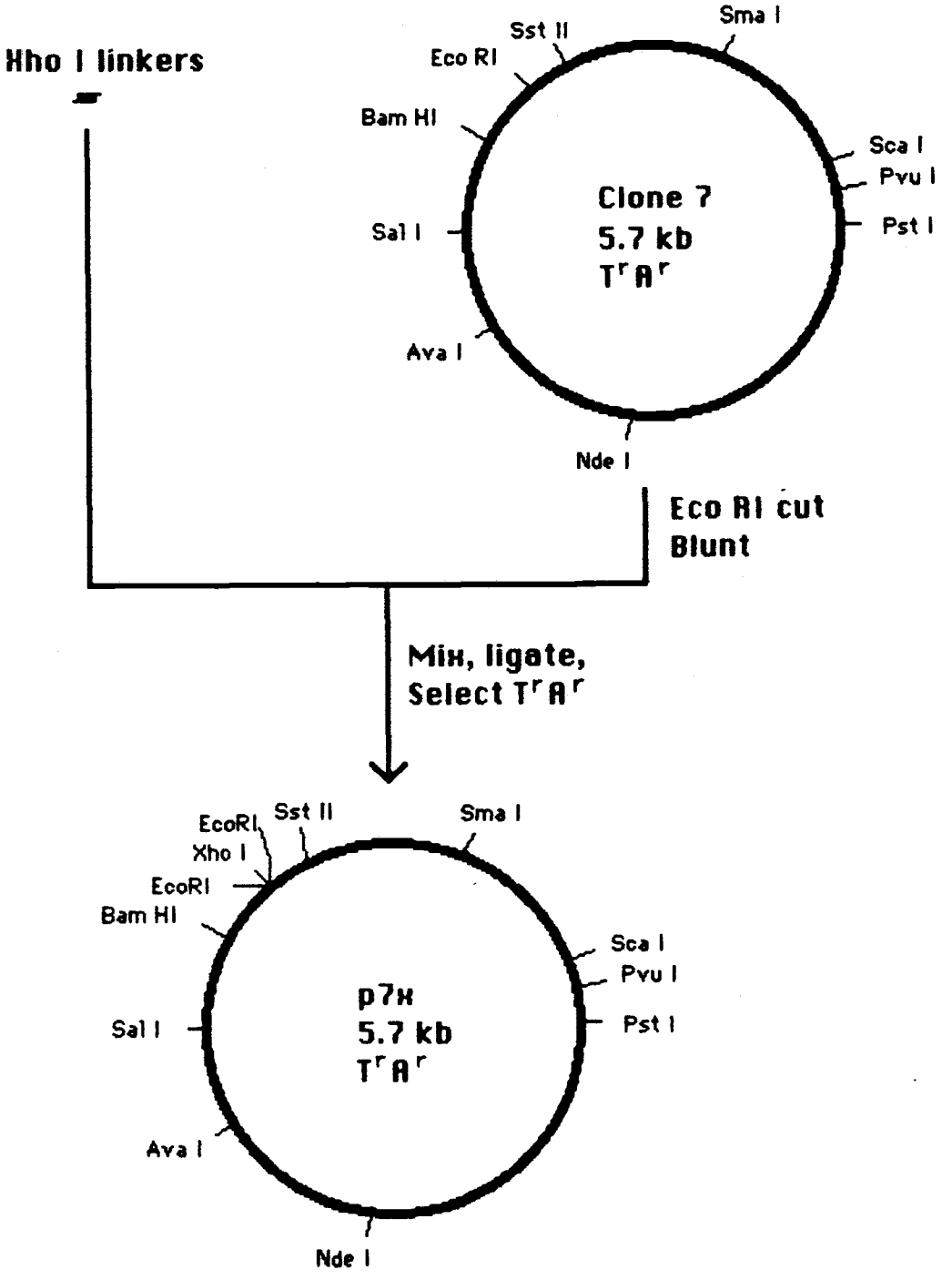

Xho I linkers
Sst II
Sma I
Eco RI
Bam HI
Sca I
Pvu I
Clone 7
5.7 kb
TrAr
Sal I
Pst I
Ava I
Nde I
Eco RI cut
Blunt
Mix, ligate,
Select TrAr
EcoRI
Sst II
Sma I
Xho I
EcoRI
Bam HI
Sca I
Pvu I
p7x
5.7 kb
TrAr
Sal I
Pst I
Ava I
Nde I

Figure 34

Figure 35: Construction of pIR12. pIR12 was made by inserting the small EcoRI to EcoRV fragment of clone 4 into the EcoRI to EcoRV sites of pDL12. This molecule at the EcoRI site has the first 12 base pairs of adenovirus 2 in an inverted orientation separated by an EcoRI site. The molecule has not been sequenced, but the sequence across the junction should be TTATTGATGATGAATTCATCATCAATAA, an inverted repeat of 28 base pairs. This molecule is somewhat unstable when grown in E. coli, forming many multimers and few monomers.

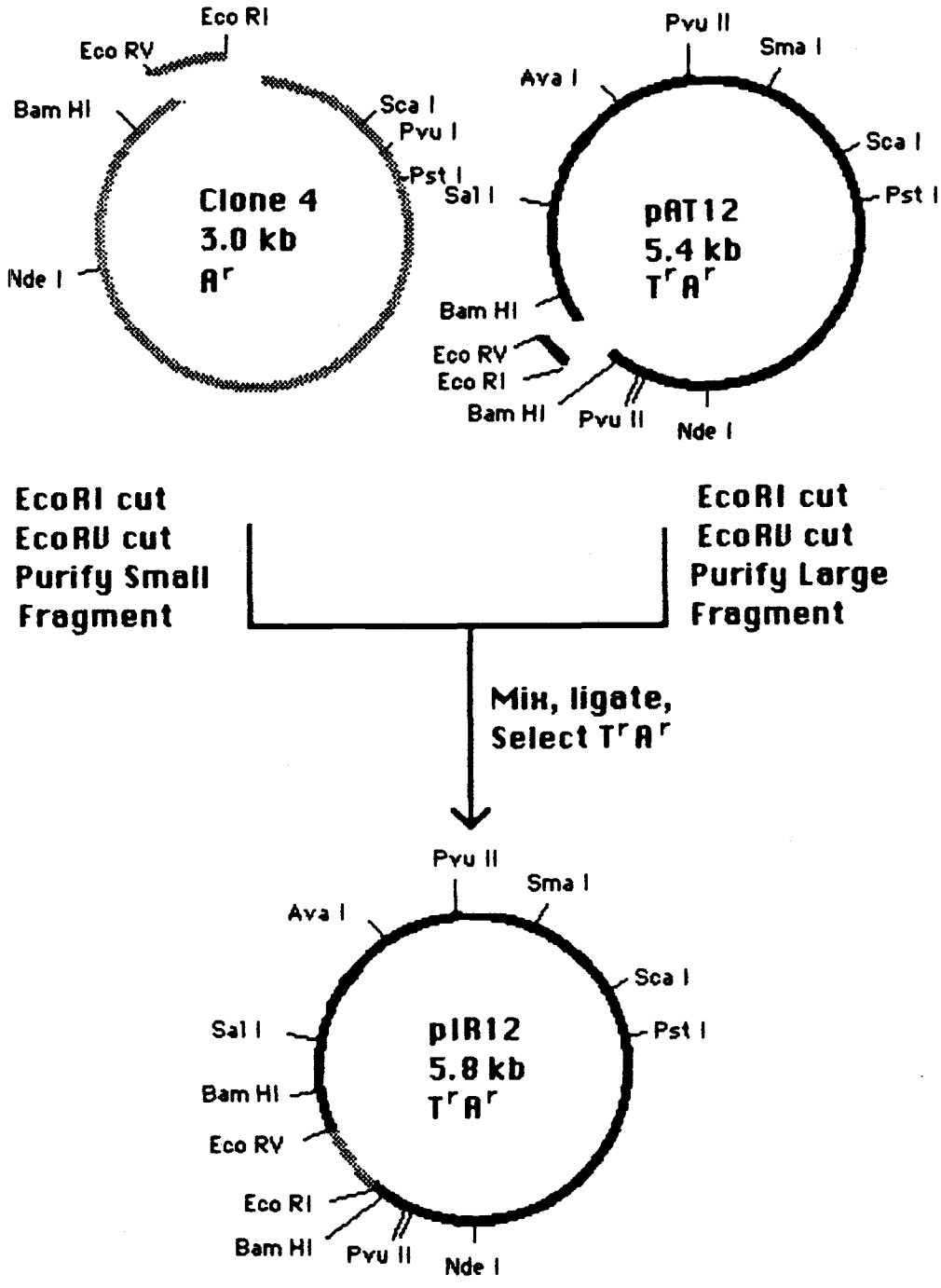

Eco RI
Eco RV
Bam HI
Sca I
Pvu I
Pst I
Sal I
Nde I
Clone 4
3.0 kb
Ar
Pvu II
Sma I
Ava I
Sca I
Pst I
pAT12
5.4 kb
TrAr
Bam HI
Eco RV
Eco RI
Bam HI
Pvu II
Nde I
EcoRI cut
EcoRV cut
Purify Small
Fragment
EcoRI cut
EcoRV cut
Purify Large
Fragment
Mix, ligate,
Select TrAr
Pvu II
Sma I
Ava I
Sca I
Sal I
pIR12
5.8 kb
TrAr
Pst I
Bam HI
Eco RV
Eco RI
Bam HI
Pvu II
Nde I

Figure 35

Figure 36: Construction of clones pAT(7-67). The series of clones designated pAT(7-67) were constructed by ligating individual PvuII-cut dL(7-67) clones with PvuII-cut p16z and selecting for resistance to tetracycline and ampicillin. The resulting molecules acquire the ampicillin resistance from p16z and the tetracycline resistance and adenovirus sequences from each respective dL clone. Although the orientations of the fragments when put together should be random, all clones examined had the orientation shown on the restriction map.

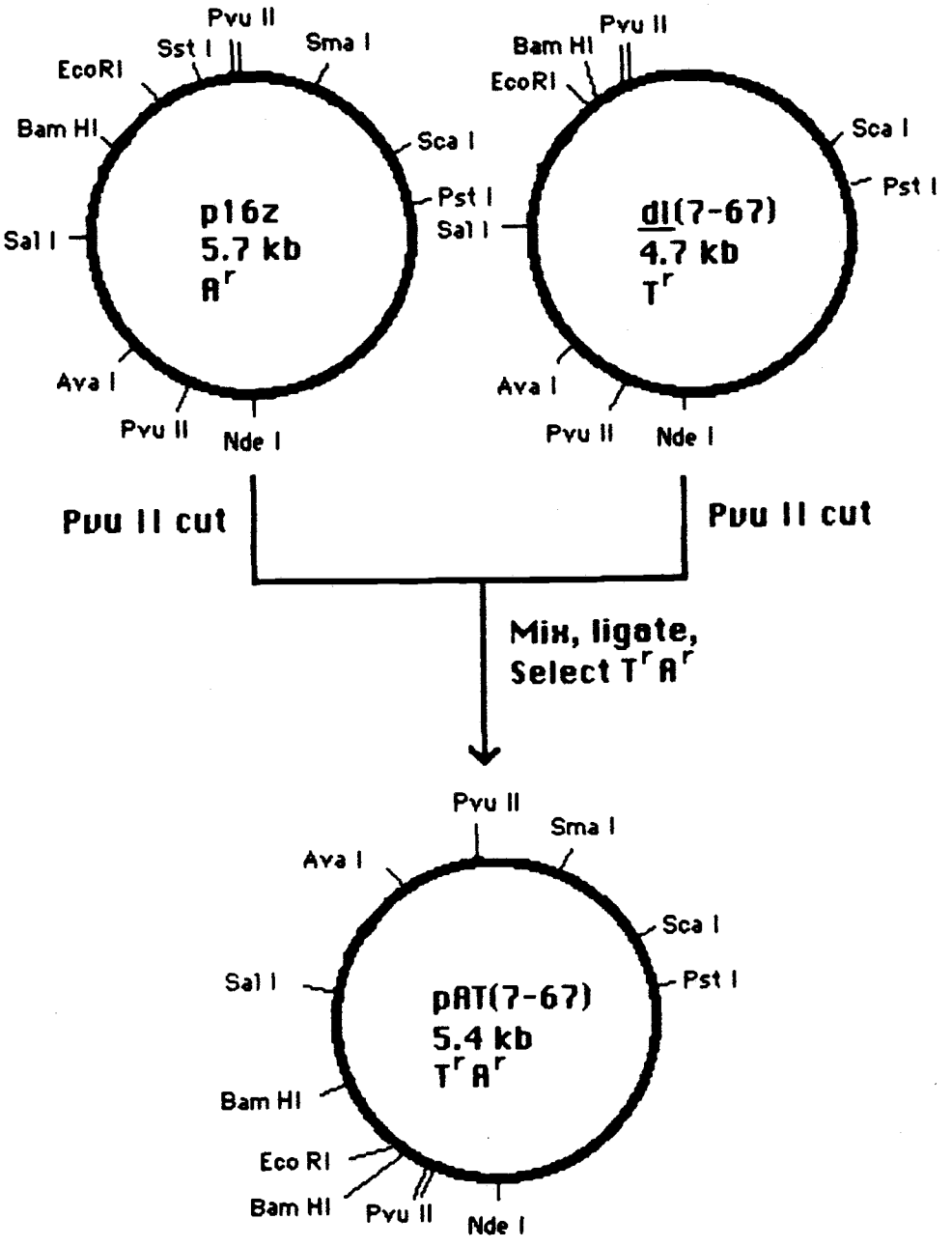

Pvu II
Sst I
Sma I
EcoRI
Bam HI
Sca I
Pst I
Sal I
p16z
5.7 kb
A^r
Ava I
Pvu II
Nde I
Bam HI
Pvu II
EcoRI
Sca I
Pst I
Sal I
dI(7-67)
4.7 kb
T^r
Ava I
Pvu II
Nde I
Pvu II cut
Pvu II cut
Mix, ligate,
Select T^r A^r
Pvu II
Sma I
Ava I
Sca I
Sal I
Pst I
pAT(7-67)
5.4 kb
T^r A^r
Bam HI
Eco RI
Bam HI
Pvu II
Nde I

Figure 36

Figure 37: Construction of pRL102. pRL102 is a derivative of pXD6 made by ligating the small SalI to SmaI fragment of clone 7 (containing the left adenovirus ITR) into the large SalI to SmaI fragment of pXD6. It contains the entire adenovirus ITR separated by ~900 bases of left end adenovirus DNA and ~500 bases of right-end adenovirus DNA. The boundary between left and right sequences is a SmaI site. The ITR sequences are exposed by cleavage with EcoRI.

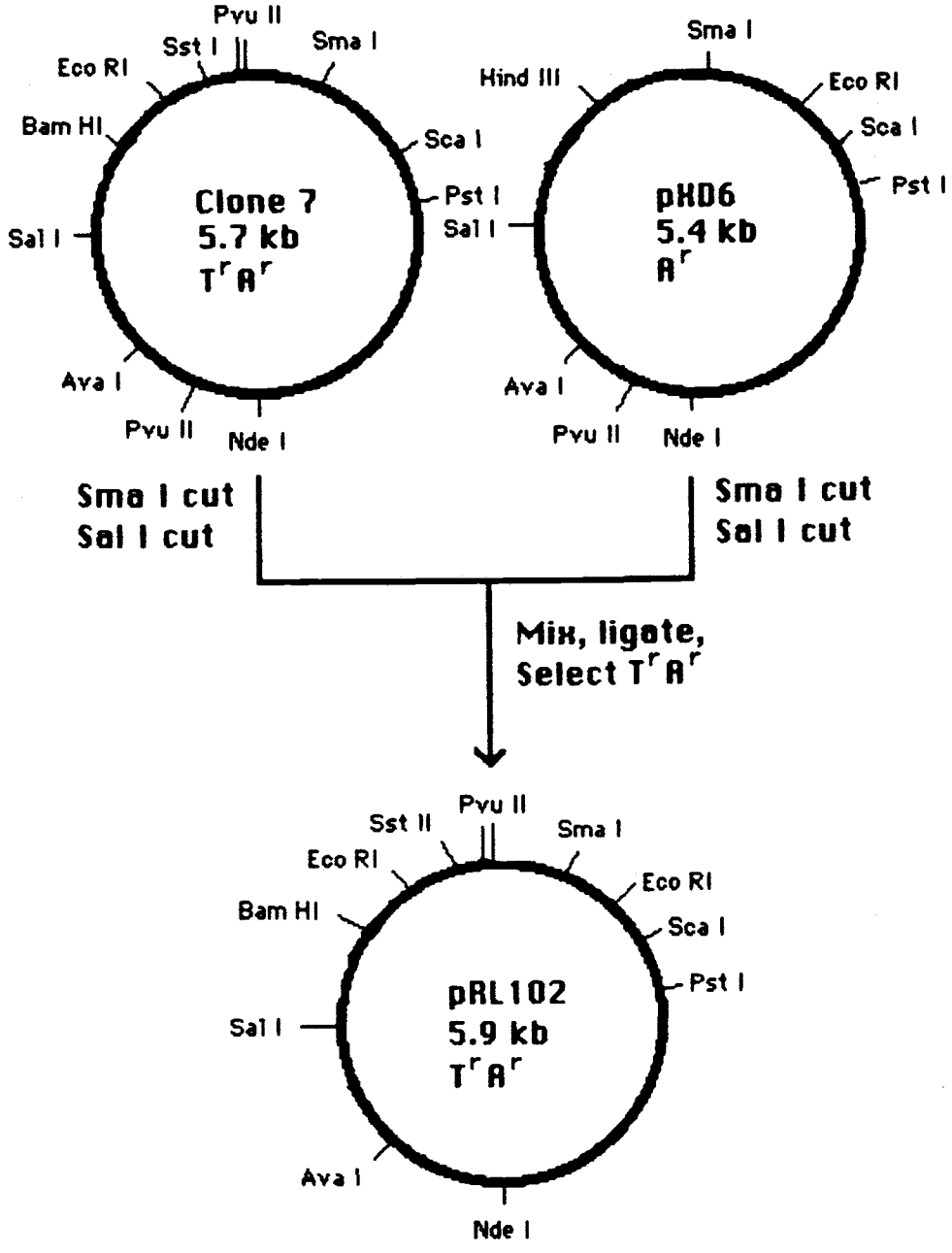

Pvu II
Sst I
Sma I
Eco RI
Bam HI
Sca I
Clone 7
5.7 kb
TrAr
Pst I
Sal I
Ava I
Pvu II
Nde I
Sma I
Hind III
Eco RI
Sca I
pHD6
5.4 kb
Ar
Pst I
Sal I
Ava I
Pvu II
Nde I
Sma I cut
Sal I cut
Sma I cut
Sal I cut
Mix, ligate,
Select TrAr
Pvu II
Sst II
Sma I
Eco RI
Eco RI
Bam HI
Sca I
pRL102
5.9 kb
TrAr
Pst I
Sal I
Ava I
Nde I

Figure 37

Figure 38: Construction of pRLKAN. pRLKAN is a derivative of pRL102 made by inserting a blunt-ended, gel-purified, EcoRI fragment of pKAT21 into the blunt-ended SstII site of pRL102. The resulting molecule is identical to pRL102 with an inserted kanamycin resistance gene at the SstII site and may be useful for study of integration of G418 resistance, since the kanamycin coding sequences are near the adenovirus E1a protomer.

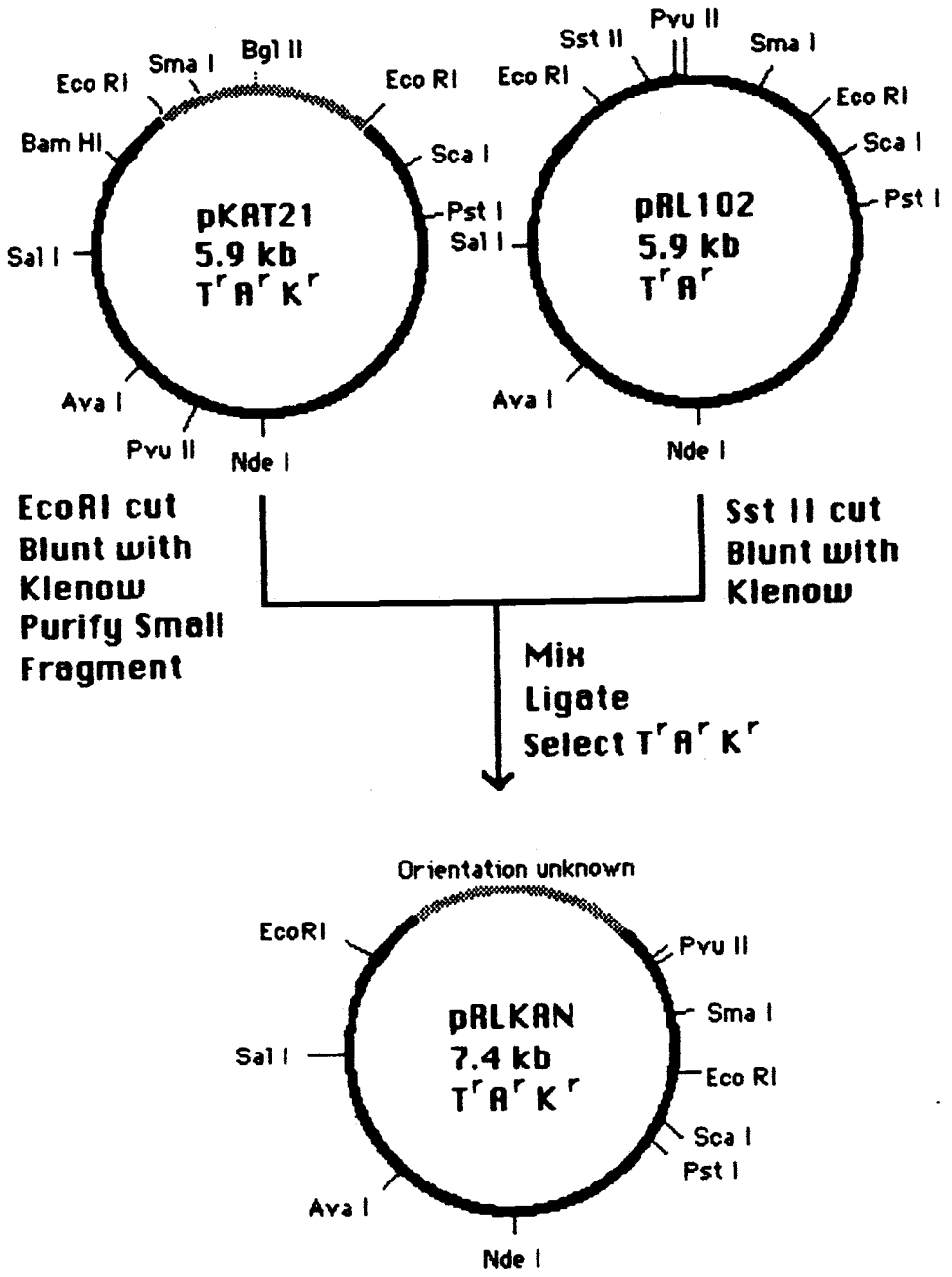

Bgl II
Sma I
Eco RI
Eco RI
Bam HI
Sca I
pKAT21
5.9 kb
T^r A^r K^r
Pst I
Sal I
Ava I
Pvu II
Nde I
Pvu II
Sst II
Sma I
Eco RI
Eco RI
Sca I
pRL102
5.9 kb
T^r A^r
Pst I
Sal I
Ava I
Nde I
EcoRI cut
Blunt with
Klenow
Purify Small
Fragment
Sst II cut
Blunt with
Klenow
Mix
Ligate
Select T^r A^r K^r
Orientation unknown
EcoRI
Pvu II
pRLKAN
7.4 kb
T^r A^r K^r
Sma I
Sal I
Eco RI
Sca I
Pst I
Ava I
Nde I

Figure 38

Figure 39: Construction of pRL12. pRL12 is a derivative of pXD6 made by ligating the small SalI to PvuII fragment of dL12 into the large SalI to SmaI fragment of pXD6. It contains the full right end of adenovirus at one EcoRI site and 12 base pairs to the left end of adenovirus at the other EcoRI site.

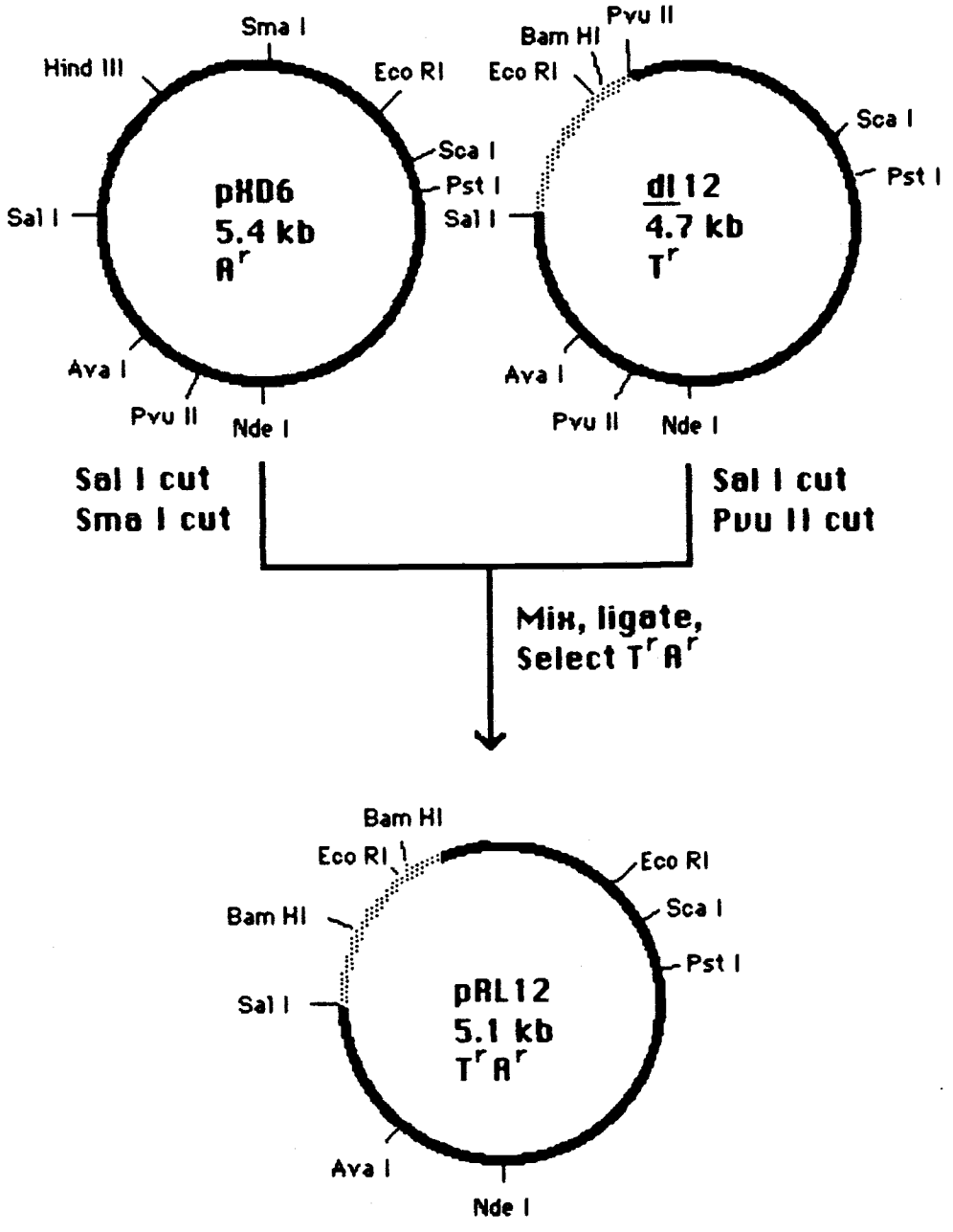

Sma I
Hind III
Eco RI
Sca I
Pst I
Sal I
pHD6
5.4 kb
A^r
Ava I
Pvu II
Nde I
Bam HI
Pvu II
Eco RI
Sca I
Pst I
Sal I
dl 12
4.7 kb
T^r
Ava I
Pvu II
Nde I
Sal I cut
Sma I cut
Sal I cut
Pvu II cut
Mix, ligate,
Select T^r A^r
Bam HI
Eco RI
Eco RI
Sca I
Bam HI
Pst I
Sal I
pRL12
5.1 kb
T^r A^r
Ava I
Nde I

Figure 39

Figure 40: Construction of pRBAM. pRBAM is a derivative of pRL12 made by cutting pRL12 with BamHI and ligating. It was a progenitor of plasmids pRL7, pRL18, pRL21, pRL31, pRL36, and pRL67, collectively referred to as pRL(x).

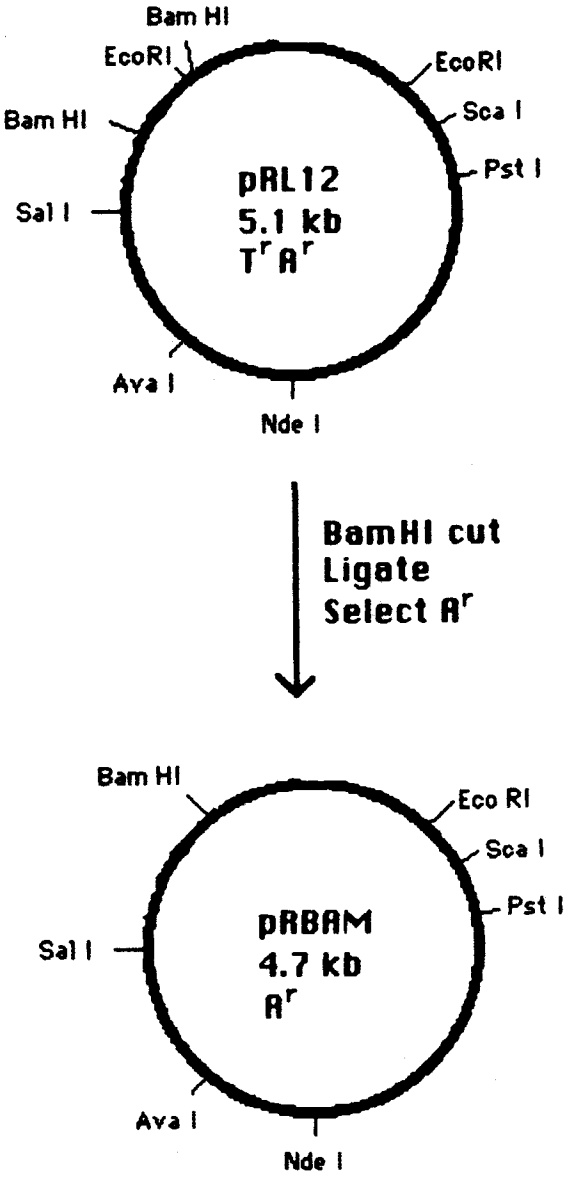
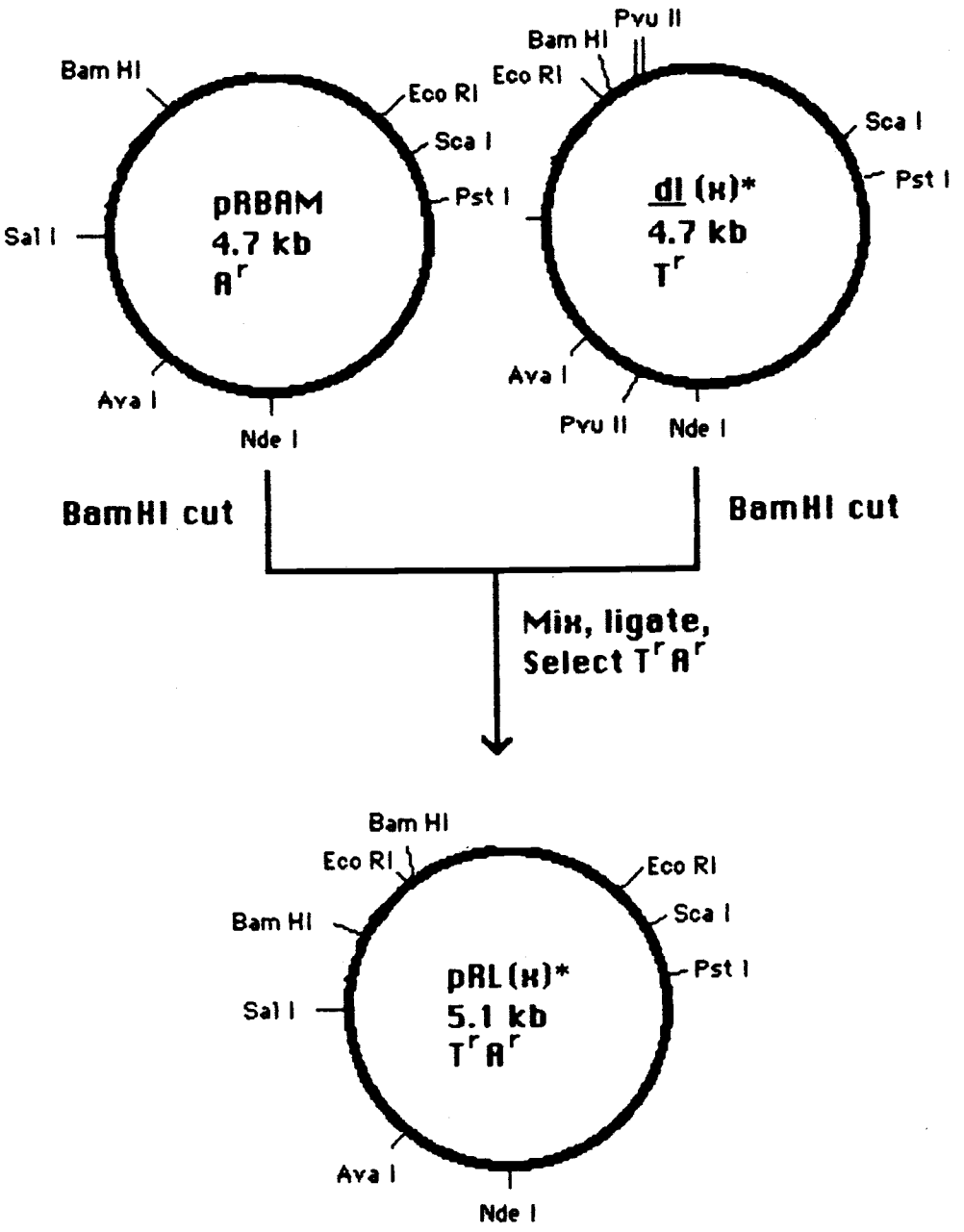

Bam HI
EcoRI
EcoRI
Bam HI
Sca I
Pst I
pRL12
5.1 kb
T^r A^r
Sal I
Ava I
Nde I
BamHI cut
Ligate
Select A^r
Bam HI
Eco RI
Sca I
Pst I
pRBAM
4.7 kb
A^r
Sal I
Ava I
Nde I

Figure 40

Figure 41: Construction of clones pRL(x). The clones designated pRL(x) constitute plasmids pRL7, pRL18, pRL21, pRL31, pRL36, and pRL67. In each case the number in the plasmid name refers to the number of bases of adenocirus left end DNA located at one EcoRI site in the plasmid. The left end sequences are inverted relative to an intact adenovirus right-end at the other EcoRI site. Each plasmid is identical to pRL12, except for the number of adenovirus left end bases contained. The plasmids were made by ligating the small BamHI fragment of the dL series of clones of the same numbers into the BamHI site of pRBAM and selecting for resistance to tetracycline and ampicillin.



* Note-κ=7, 18, 21, 31, 36, or 67

Figure 41

Figure 42: Construction of pBN7. pBN7 is a derivative of pADIRK2 made by cutting at the unique BsmI and NdeI sites within the molecule, blunting, and ligating. It was a progenitor of a series of clones designed to study the amount of sequence required for trans replication.

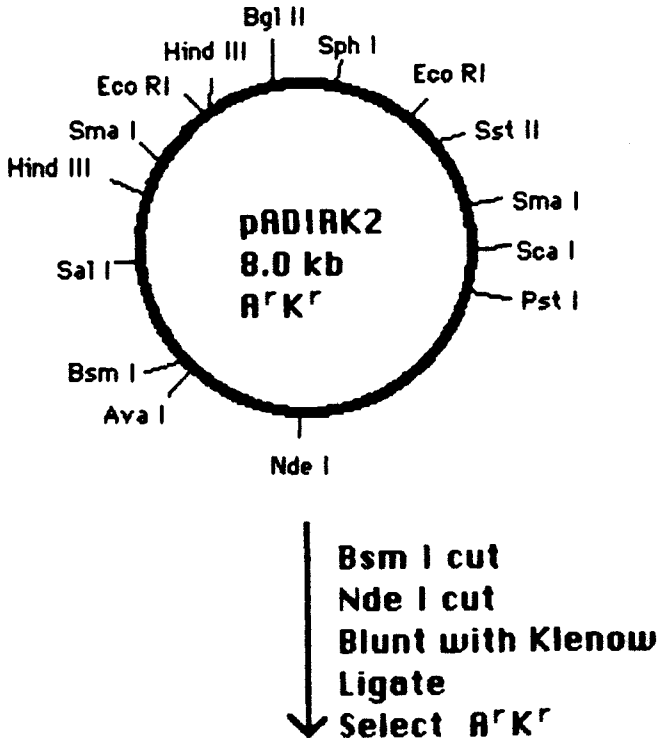

Bgl II
Hind III
Sph I
Eco RI
Eco RI
Sst II
Sma I
Hind III
Sma I
Sca I
Pst I
Sal I
Bsm I
Ava I
Nde I
pADIAK2
8.0 kb
ArKr
Bsm I cut
Nde I cut
Blunt with Klenow
Ligate
Select ArKr
Bgl II
Hind III
Sph I
Eco RI
Eco RI
Sst II
Sma I
Hind III
Sma I
Sca I
Pst I
Bam HI
Sal I
pBN7
7.1 kb
ArKr

Figure 42

Figure 43: Construction of pSP38. pSP38 is a derivative of pADIRK2 made by cutting at the unique ScaI and PvuI sites in the molecule, blunting, and ligating. It, too, was designed as a progenitor of a series of clones designed to study the amount of sequence required for trans replication.

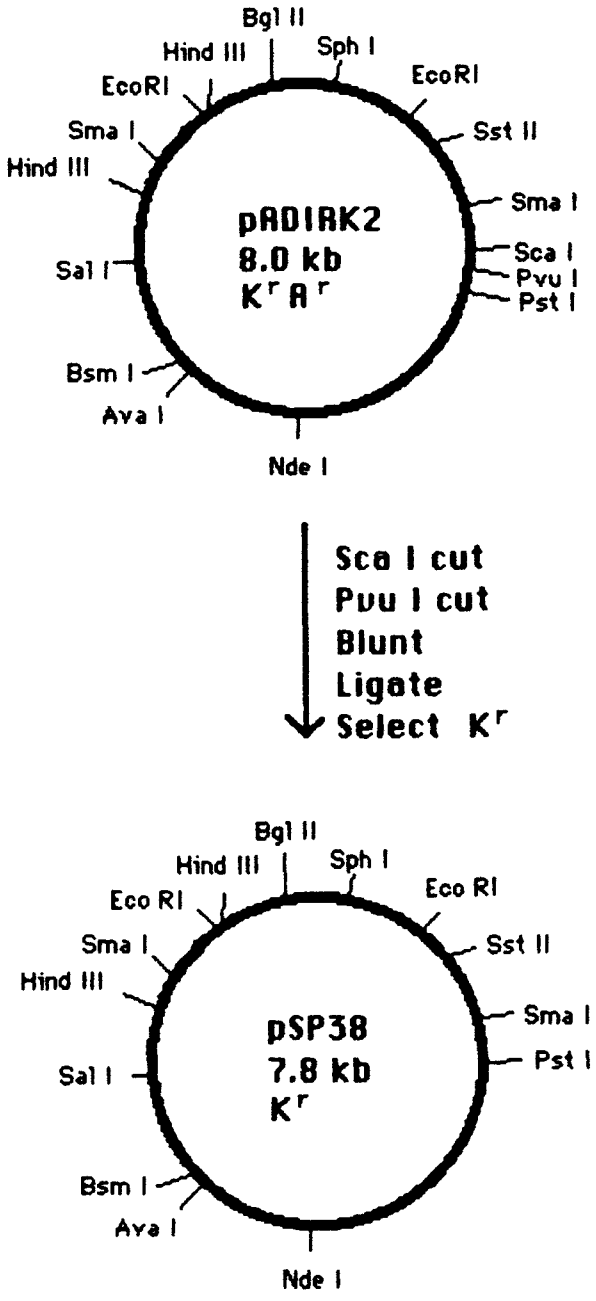

Bgl II
Hind III
Sph I
EcoRI
EcoRI
Sma I
Sst II
Hind III
Sma I
pADIRK2
8.0 kb
K^r A^r
Sca I
Pvu I
Sal I
Pst I
Bsm I
Ava I
Nde I
Sca I cut
Pvu I cut
Blunt
Ligate
Select K^r
Bgl II
Hind III
Sph I
Eco RI
Eco RI
Sma I
Sst II
Hind III
Sma I
pSP38
7.8 kb
K^r
Pst I
Sal I
Bsm I
Ava I
Nde I

Figure 43

Figure 44: Construction of pADSMA14. pADSMA14 was made by mixing adenovirus 5 DNA cut with SmaI with p44x DNA cut at the EcoRV site, ligating and selecting for ampicillin resistance and tetracycline sensitivity. It contains a 6480 base sequence from adenovirus 5 spanning nucleotides 6573 to 13053 on the adenovirus 2 numbering scheme. This region of adenovirus contains the entire pTP gene, the two VA RNAs, the entire gene designated 52/55, the i leader, the 3' end of the tripartite leader of the major late promoter, the 5' coding sequences of AdpoL, the 5' sequences of the virion IIIa gene, and the only consensus Nuclear Factor I binding site of adenovirus outside the ITR sequences.

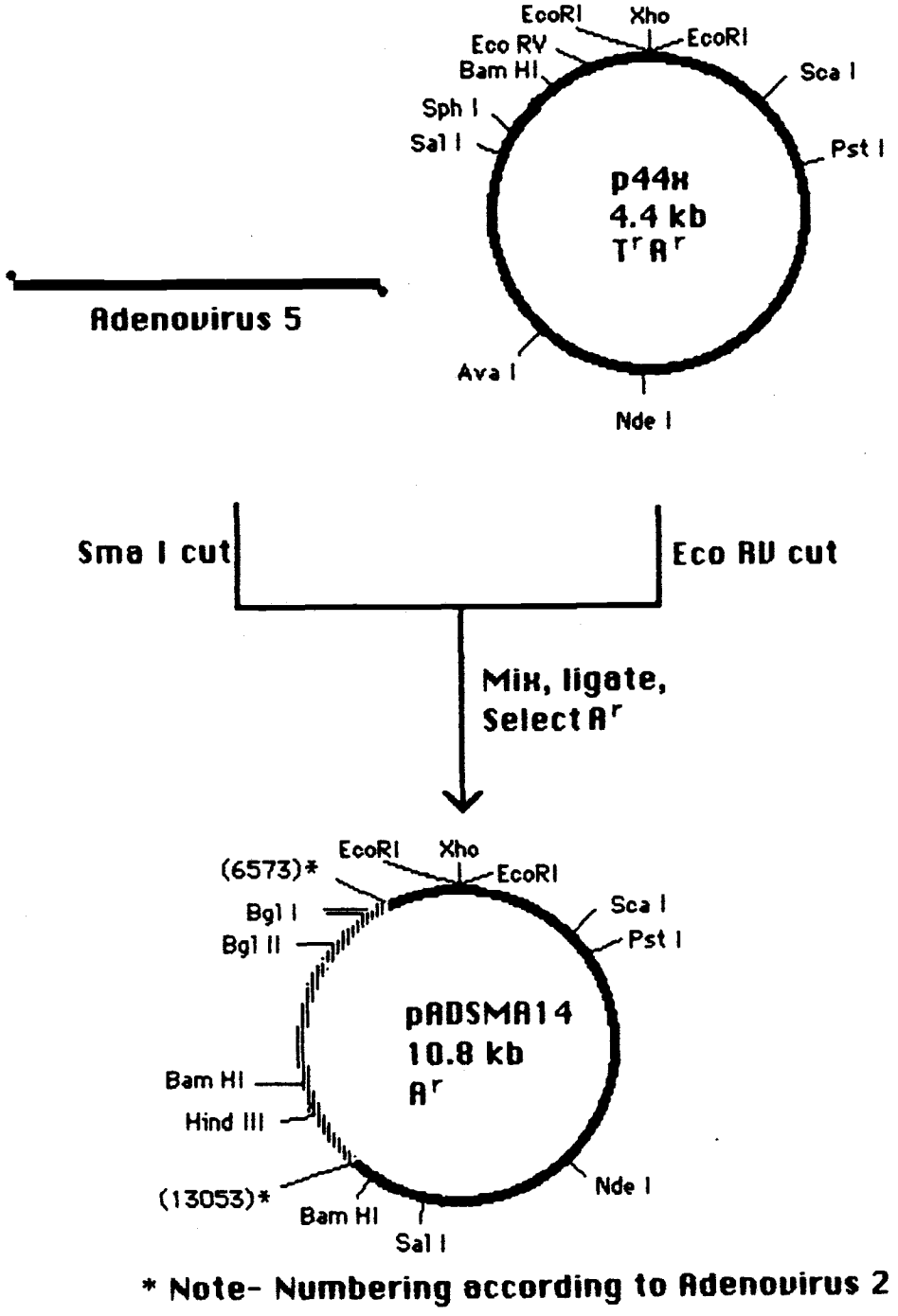

EcoRI
Xho
EcoRI
Eco RV
Bam HI
Sca I
Sph I
Sal I
Pst I
p44H
4.4 kb
TrAr
Adenovirus 5
Ava I
Nde I
Sma I cut
Eco RV cut
Mix, ligate,
Select Ar
EcoRI
Xho
(6573)*
EcoRI
Bgl I
Sca I
Bgl II
Pst I
pADSMA14
10.8 kb
Ar
Bam HI
Hind III
Nde I
(13053)*
Bam HI
Sal I
* Note- Numbering according to Adenovirus 2

Figure 44

Figure 45: Construction of pADSPH61. pADSPH61 was made by mixing adenovirus 5 DNA and p44x DNA each cut with SphI, ligating and selecting for ampicillin resistance and tetracycline sensitivity. It contains adenovirus DNa from positions 25140 to 31810 on the adenovirus 2 numbering scheme. It should be noted that the SphI site at position 31108 in adenovirus 2 is not present in adenovirus 5 and is the reason why the SphI fragment terminates at 31810.

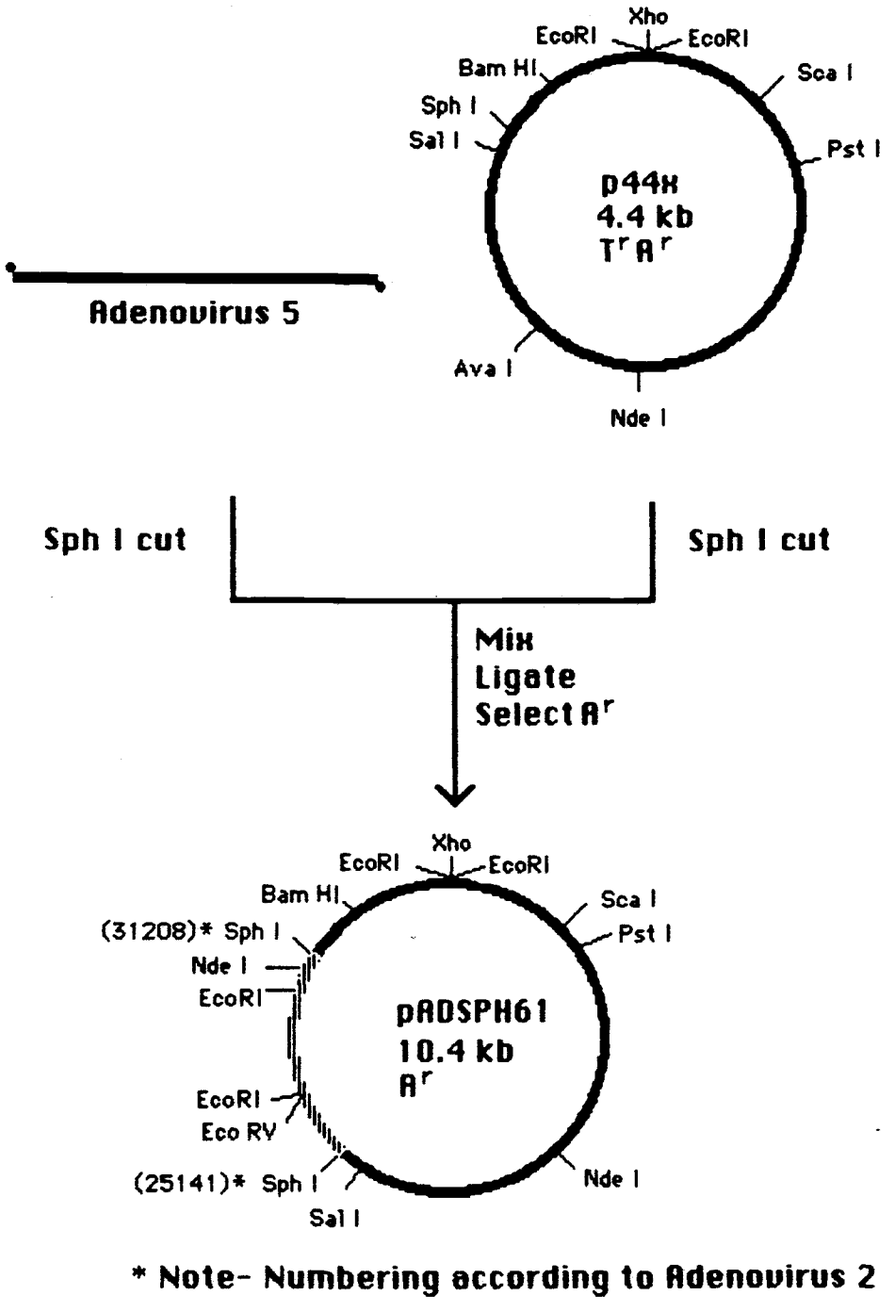

Xho
EcoRI
EcoRI
Bam HI
Sca I
Sph I
Sal I
Pst I
p44H
4.4 kb
TrAr
Adenovirus 5
Ava I
Nde I
Sph I cut
Sph I cut
Mix
Ligate
Select Ar
Xho
EcoRI
EcoRI
Bam HI
Sca I
(31208)* Sph I
Pst I
Nde I
EcoRI
pADSPH61
10.4 kb
Ar
EcoRI
Eco RV
(25141)* Sph I
Nde I
Sal I
* Note- Numbering according to Adenovirus 2

Figure 45