MOLECULAR GENETICS OF DNA CODING FOR AVIAN FEATHER KERATINS AND FOR COLIPHAGES 186 AND P2.

A thesis submitted for the Degree
of Doctor of Philosophy in the
University of Adelaide.

by

ROBERT BRYCE SAINT, B.Sc.(Hons.)
Department of Biochemistry

April 1979.
Awarded November 1979

To Susie Kaye,

and her equally beautiful temperament.

CONTENTS

THESIS SUMMARY.

Restriction enzyme, molecular cloning and DNA annealing techniques have been used to study mRNA and DNA coding for the embryonic feather keratins of the chicken and the DNA genomes of coliphages 186 and P2. The coliphage DNAs were used to develop the techniques for application to the keratin system which awaited the availability of appropriate bio-hazard containment facilities before being undertaken. The following results were obtained.

1. Restriction endonuclease cleavage of chick DNA with *Bam*HI, *Bgl*II, *Eco*RI, or *Hin*dIII, fractionation on agarose gels, immobilization on nitrocellulose filters and annealing to DNA complementary to purified 12S mRNA isolated from the developing embryonic feather and coding for embryonic feather keratins, yielded a complex pattern of major and minor bands. These patterns consisted of 4 - 6 major bands and many minor bands. No simple repeat length could be deduced from these patterns, suggesting that keratin-coding DNA is heterogeneous in coding sequences, non-coding sequences or both.

2. Keratin gene expression was shown to be independent of DNA rearrangement, as the complex pattern of restriction fragments was identical in DNA isolated from germ-line tissue (sperm) the differentiated feather tissue and somatic tissue not synthesizing keratins (erythrocytes). Keratin gene expression must therefore involve the activation of pre-existing control regions in the DNA.

3. The purified 12S mRNA coding for feather keratin was transcribed into double-stranded DNA and individual species isolated by molecular cloning in *E. coli*. Sequence variation between species was confirmed by restriction enzyme analysis.

4. Preliminary analysis of the cloned species revealed the existence of two distinct groups of species comprising 12S mRNA: Group I (the more abundant group) and Group II (the less abundant). The fact that filter-bound DNA of individual Group I species bound more 12S cDNA than equal amounts of Group II species DNA and that pure Group I species and total 12S mRNA sequences (coding for keratins in cell-free translation systems) annealed to exactly the same complex set of *Eco*RI, *Hin*dIII, or *Bgl*II restricted chick DNA fragments, compels the conclusion that Group I species represent true keratin coding sequences.

Group II species annealed to restricted chick DNA fragments which were totally different to those annealing to either Group I species or total 12S mRNA sequences. Different Group II species appeared to anneal to certain common fragments, suggesting that this less abundant group was comprised of a family of sequence related species and were not simply contaminating mRNA species coding for 'housekeeping' functions. Their exact nature is at present, however, uncertain.

5. Group I species, the presumptive keratin-coding species, are members of a family of homologous species present in the chick genome. This is demonstrated by the fact that the two Group I species which have been examined so far,

shown to be non-identical by restriction analysis, and total 12S mRNA sequences from which they were derived, annealed to the same set of between 20 and 30 *Bgl*II, *Hin*dIII or *Eco*RI restricted chick DNA fragments under annealing and washing conditions of low stringency, (high salt). Under stringent (low salt) washing conditions, however, all except between 1 and 3 of the duplexes formed by these fragments and the Group I species were differentially lost from the filter, indicating that the majority of duplexes were mis-matched and therefore that these multiple copies were homologous and not identical. In addition the two non-identical Group I species annealed to *Eco*RI generated chick DNA fragments of different sizes under the stringent (low salt) washing conditions, demonstrating that differences must exist in the sequence of adjacent non-coding and/or intervening sequences (should they exist) for these two species.

6. Although the two Group I species discussed above annealed to different *Eco*RI generated chick DNA fragments under the stringent (low salt) washing conditions, they both annealed under these conditions to a *Hin*dIII generated chick DNA fragment of size 3.0 kb. Assuming that this is a single fragment and not two fragments co-electrophoresing by chance, sequences identical to or with very close homology to both of these species lie on the same fragment and are therefore linked in the genome. The exact nature of this linkage and of the extent of gene clustering, should it exist, was not determined.

7. Restriction cleavage maps of coliphages 186 and P2 were determined for the enzymes *Bam*HI, *Bgl*II, *Eco*RI, *Hin*dIII, *Pst*I, *Sal*I, *Xba*I, and *Xho*I. These maps were used to analyse four insertion or deletion mutants affecting the major control region of 186. 186*ins*2 and 186*ins*3 were shown to be insertions of an IS3 element in the *c*I gene and *int* gene respectively. 186*del*1 and 186*del*2 were shown to carry the same deletion affecting the *c*I gene, but 186*del*2 carried a cryptic insert in the repressor binding site (operator).

STATEMENT

This thesis contains no material which has been accepted for any award of any other degree or diploma in any university. To the best of my knowledge and belief, this thesis contains no material previously published or written by another person except where due reference is made in the text.

ROBERT SAINT.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the following people for their help during the course of my work. I thank each of them for their valuable assistance.

Professor W.H. Elliot: for permission to work in the Department of Biochemistry, University of Adelaide.

Professor G.E. Rogers: for the opportunity to work on the molecular biology of keratinocytes and for his enthusiasm, support and encouragement throughout the course of this work.

Dr. J.B. Egan: for the opportunity to work on the coliphages 186 and P2 and for the many stimulating discussions of the work described in Section II. I would also like to thank the members of his laboratory, in particular *Stephanie Hocking* and *Jean Finnegan* for their help.

Trevor Lockett, Dr. Peter Gibbs, Barry Powell, Phil Morris, Morry Frenkel and *Steve Wilton* for their many helpful discussions and for the enthusiastic atmosphere in which the research was carried out.

Professor K. Murray and members of his department, in particular *Dr. Keith Peden*, for their advice and their generosity in allowing me to use the facilities in the Department of Molecular Biology,

University of Edinburgh.

Dr. J Adams and members of his laboratory for their assistance and their generosity in allowing me to use the facilities in the Walter and Eliza Hall Institute for Medical Research.

Linda Mercer for her help with the electron microscopy described in Chapter 4.

Jeanette Brooker for the typing of this thesis.

Finally I would like in particular to express my thanks to *Lesley Crocker* for excellent technical assistance throughout this work and for the production of photographic plates and diagrams.

ABBREVIATIONS

A_{260}	-	absorbance of light of wavelength 260 nm (1 cm path length).
bp	-	nucleotide base pairs
BSA	-	bovine serum albumin
cDNA	-	single-stranded DNA complementary to RNA
cfu	-	colony-forming units
c.p.m.	-	counts per minute
dATP	-	adenine deoxyribose-5'-triphosphate
dCTP	-	cytosine deoxyribose-5'-triphosphate
dGTP	-	guanine deoxyribose-5'-triphosphate
DNA	-	deoxyribonucleic acid
DNase	-	deoxyribonuclease
dscDNA	-	double-stranded DNA complementary to RNA
DTT	-	dithiothreitol
dTTP	-	thymine deoxyribose-5'-triphosphate
EDTA	-	ethylenediaminetetracetic acid
kb	-	kilobase pairs = 1,000 base pairs
mRNA	-	messenger RNA
pfu	-	plaque-forming units
RNA	-	ribonucleic acid
RNase	-	ribonuclease
SDS	-	sodium dodecyl sulphate
TCA	-	trichloroacetic acid
Tris	-	tris(hydroxymethyl)aminomethane
UV	-	ultraviolet

GENERAL INTRODUCTION

GENERAL INTRODUCTION.

This general introduction is included to explain the structure of this thesis, which resulted directly from the circumstances surrounding recombinant DNA research and which therefore prevailed during the course of this project. The project incorporated studies on two completely different systems of gene expression, namely chick feather keratins and coliphages 186 and P2. The original project aims did not include the coliphage work and the keratin aspect maintained priority throughout. However considerable effort became channelled during the course of this work to the two coliphage genomes. This was a direct result of the absence of the necessary containment facilities in the Adelaide department to undertake recombinant DNA work on nucleic acids isolated from the chicken. Being denied the direct approach, Dr. J.B. Egan kindly agreed to the use of the coliphage DNAs to develop the necessary expertise.

In 1976 I was fortunate to have the opportunity to work in Professor K. Murray's laboratory at the University of Edinburgh and to use that laboratory's expertise and containment facilities for the molecular cloning of keratin genes. Soon after my arrival, however, a decision was made by the Adelaide department to postpone any recombinant DNA work pending the establishment of a properly constituted Adelaide University body to oversee such research. Again being denied the direct approach, the work on coliphages 186 and P2 was extended and studies of

keratin coded DNA were restricted to those using the total 12S cDNA (a mixture of species).

In August of last year the Walter and Eliza Hall Institute for Medical Research generously made available their C3 containment facility for the molecular cloning of keratin mRNA sequences. The results described in this thesis were obtained using the material I was able to prepare in Melbourne, as the containment facility nearing completion in the Adelaide department was incomplete at the time of writing this thesis.

As described, the initial adoption of the coliphage systems resulted directly from being unable to use the keratin system to develop the necessary techniques. This thesis is therefore divided into two sections; Section I describing studies of avian keratin genes and Section II describing studies of coliphages 186 and P2. The two sections are connected by the mutual theme of the study of gene organization and by common technical approaches.

SECTION I

THE MOLECULAR GENETICS OF DNA CODING FOR AVIAN KERATINS

CHAPTER 1

INTRODUCTION

CHAPTER 1.

INTRODUCTION.

A. Introductory Comments.

Keratins are a group of fibrous proteins found in many epithelial cells of higher organisms. Keratin-containing tissues are characteristically durable, pliable and insoluble, and are likely to have played a major role in the colonization of land by vertebrates. The general aim of the work described in this thesis is to understand the nature of the development and differentiation of avian keratinocytes, which synthesize keratins, and to elucidate the molecular mechanisms involved in keratin gene expression and its control. The embryonic feather has proved to be the most experimentally amenable tissue for the study of keratin gene expression and was the tissue used for the studies described in this thesis.

This introduction will describe current knowledge of the processes involved in feather development and keratinisation at the levels of both morphological development and of keratin gene expression and its control. Systems other than keratin that involve the expression of multiple sets of related genes will also be discussed. Finally a concise general consideration of the current knowledge of the nature of gene control will be made.

B. The Developing Keratinocyte.

1. The developing embryonic feather:

The events involved in the morphogenesis of the

chick embryonic feather have been studied extensively, and reviews of this work have been published by Romanoff (1960), Lillie (1965) and Voitkevich (1966). A more recent study using electron microscopy is described by Matulionis (1970).

After five days of incubation of fertilized chick eggs, the skin of the chick embryo consists of a layer of epithelial cells, overlain by peridermal cells and underlain by the mesoderm. Shortly afterwards the mesoderm condenses into regions of high cell density, and the epithelial cells begin to proliferate, so that by eight days the feather germs are visible as epidermal thickenings. By ten days the feather germs consist of cylindrical lumps of epidermal cells with mesodermal cores. Over the next two to three days the feather grows by cell division. Several cell types, all derived from the epidermal cells then become evident. The core of the feather is occupied by a few pulp cells of mesodermal origin and by two capillary vessels. Growth after 13 days is attributed to cell growth and movement, rather than further cell division so that cell division has largely ceased by day 13. Kemp *et al.*, (1974a) showed that keratin biogenesis proceeds after the final cell division, and keratinisation is eventually completed by about 18 days of development, three days before hatching. At this stage the cells are filled with the fibrous proteins and are dead. A more extensive review of keratin biogenesis is given below. Thus development and keratinisation of the embryonic feather is an example of a terminally differentiating tissue.

2. The role of the dermis in feather development:

Tissue grafting experiments involving the removal of the epidermis and dermis from various regions of chick embryos and recombination and maintenance *in vitro* have revealed that the epidermis from a given region is directed to develop structures normally associated with the region from which the dermis was derived (Sengel, 1957, 1971; Rawles, 1963, 1965; Wessels, 1962, 1965). Furthermore, the keratins produced in these recombinants are identical to those found in the same epidermal structures where no recombination had been performed (Dhouailly *et al.*, 1978). Thus the recombination of presumptive scale dermis from the leg of the embryo with presumptive feather epidermis from the back of the embryo causes the epidermis to form scales and to express the set of scale (and not feather) keratins normally found in scale tissues. The reverse recombination of presumptive feather dermis with presumptive scale epidermis leads to the formation of feathers and expression of the feather set of genes. These effects exerted by the dermis, only occur at specific stages of development (Rawles 1963, 1965) which are earlier than the onset of keratinisation in these tissues (Matulionis, 1970; Kemp *et al.*, 1974a; Beckingham Smith, 1973a).

In similar experiments, recombinants between duck dermis and feather epidermis (Dhouailly, 1967) and mouse foot-pad dermis and feather epidermis showed that the dermis can direct the morphology of the epidermis even in heterospecific recombinants. The keratins of the foot-pad structures in the latter example were shown to be those of

the scale type.

It is clear from these results that information of some description is transmitted either by cell-cell contact, extracellular substances influencing the environment of the epidermal cells, or a combination of both and that this information dictates the rate of mitosis of epidermal cells, thereby controlling epidermal morphology (Wessels, 1962; Dodson, 1963) as well as dictating the set of keratin genes expressed in the developing tissue (Dhouailly *et al.*, 1978).

3. Other factors affecting embryonic feather development:

a) Vitamin A:

It has been demonstrated that if immature (Fell and Mellanby, 1953) or highly differentiated (Fell, 1957) embryonic chick skin is cultured in the presence of excess Vitamin A, keratinisation is inhibited and the cells undergo a mucous metaplasia. The synthesis of keratin proteins is completely repressed under these conditions (Beckingham Smith, 1973b). For the older tissues at least, the effects are reversible on the removal of Vitamin A (Fitton-Jackson and Fell, 1963).

b) Hormones:

Several hormones have been implicated in the development of avian keratinocytes. The role of pituitary and thyroid glands has been discussed at some length by Voitkevich (1966), but few studies of these phenomena have been made at the molecular level. *In vivo* (Bartels, 1943) and *in vitro* experiments (Wessells, 1961; Kitano and

Kuroda, 1967) have indicated that thyroxine accelerates epidermal keratinisation. Hydrocortisone also hastens keratin synthesis in skin cultures (Fell, 1962; Sugimoto and Endo, 1969) and causes feather germs to abort.

c) Epidermal growth factor:

This polypeptide growth factor was first isolated from the submaxillary gland of the male mouse by Cohen (1962) and produced precocious epidermal development when administered to *in vivo* or *in vitro* organ culture (Cohen, 1965). The growth factor stimulated RNA and protein synthesis (Hoober and Cohen, 1967) and increased the proportion of ribosomes associated with polysomes (Cohen and Stastny, 1968). The material isolated from mouse is effective against skin tissue from a variety of animal sources, including the chicken, but an *in vivo* requirement for this factor is yet to be demonstrated.

It should be emphasized that there appears to be no simple mechanism for the initiation of feather development and of keratin synthesis, especially in view of the fact that the various factors described above may be affecting mitosis in general or specific gene expression or both. Clearly a molecular approach to the development of the feather and of the gene expression involved is required for a proper understanding of the role of these various factors.

4. Developing keratinocytes other than the chick feather:

Rheinwald and Green (1975a, 1975b) described serial cultivation of two classes of mammalian keratinocytes,

mouse keratinocytes of teratomal origin and human epidermal keratinocytes. Both of these types required support from 373 cells or other fibroblasts to initiate colony formation from single cells. Colonies thus formed, differentiated into structures resembling stratified squamous epithelia and contained epidermal keratin polypeptides amounting to 30 - 40% of total cell protein (Sun and Green, 1978). Cell division in growing colonies appeared to be confined to the basal layer (Rheinwald and Green, 1975b), as in skin, with the more superficial layers consisting of cells undergoing terminal differentiation. It is of interest that keratinocytes at all stages of differentiation, either in epithelium or in culture and including the basal cells, were stained by antiserum to a group of keratins purified from human stratum corneum (Sun and Green, 1978). This suggests that differentiation in this mammalian case is quite different to that of the feather described by Kemp *et al.*, (1974a), since the small multiplying cells of the epidermal colonies described by Sun and Green (1978) contained keratin proteins thus involving keratin gene expression before the final cell division and terminal differentiation. No molecular studies of the genes expressed in this system has yet been reported.

C. Gene Expression In The Developing Feather.

1. The time course of keratin protein synthesis:

Keratin synthesis begins at about the twelfth day of embryonic development, as judged by the appearance of keratin fibrils in the cells (Matulionis, 1970) and of keratin proteins detected by polyacrylamide gel electro-

phoresis (Kemp *et al.*, 1974a). Major feather keratin protein bands defined by a pH 9.5, polyacrylamide gel system were only present in trace amounts in 11- and 12-day feather extracts but rapidly increased in quantity after day 12 and had become the most abundant protein species in the feathers by day 15. Kemp *et al.*, (1974a) used labelling studies and gel electrophoresis of proteins to show that the keratin proteins were synthesized co-ordinately. As mentioned above, keratinisation is completed by about the eighteenth day of embryonic development.

2. The time course of mRNA synthesis:

Partington *et al.*, (1973) and Kemp *et al.*, (1974b) described the isolation of a purified fraction of mRNA coding for keratin proteins in a cell free system. The properties of this mRNA fraction are described below. Powell *et al.*, (1976) used labelled DNA complementary to this purified mRNA fraction to detect the presence of homologous mRNA in feather tissues of varying ages. They demonstrated that the 12S mRNA species are synthesized initially from about day 13, increasing to maximal levels at day 14 and 15 and decreasing after that. Thus keratin mRNA synthesis occurred at the same time as the synthesis of keratin proteins, demonstrating that the massive onset of synthesis of keratin proteins after day 12 is brought about by an increased rate of synthesis of keratin mRNA and not by differential translation of a pre-formed mRNA pool.

3. The nature of embryonic feather keratin proteins:

The keratin proteins synthesized during

the development of the embryonic feather finish as fibrous, insoluble, intracellular proteins which have a high cystine content and are aggregated into fibrillar masses, the structure of which has been extensively studied (see Fraser *et al.*, 1972 for a review). The fibrils are stabilized by inter-molecular disulphide bonds (Goddard and Michaelis, 1934) so that the biochemical study of the proteins requires disulphide bond breakage by reducing agents and stabilization by alkylation, e.g., carboxymethylation (Harrap and Woods, 1964a; Kemp and Rogers, 1972).

S-carboxymethylated chick feather keratins appear to be homogeneous in molecular weight, the estimates varying between 10,500 and 11,500 daltons (Harrap and Woods, 1964b; Walker and Rogers, 1976a), while reduced, non-alkylated keratin had a molecular weight of 10,000 daltons (Woodin, 1954). Despite this molecular weight homogeneity, the proteins appear to have different charge characteristics, as determined by ion exchange chromatography and electrophoresis (Harrap and Woods, 1964a; Kemp and Rogers, 1972; Walker and Rogers, 1976a). Walker and Rogers (1976a) found an estimated twenty-two protein chains in the embryonic feather. Sequence analysis of the proteins has suggested that each chain is the product of a separate gene, but that all the proteins are closely related in primary structure (Walker and Rogers, 1976b).

Similar observations have been made for the keratins expressed in other chick tissues, e.g., adult feather (Woods, 1971) and scales (Walker and Bridgen, 1976). However although the keratins of these tissues are related in

sequence (Walker and Bridgen, 1976), each tissue appears to involve the expression of a distinct set of keratin species (Harrap and Woods, 1964a; Kemp and Rogers, 1972; Walker and Bridgen, 1976).

The primary structure of the keratins suggests the basis for the heterogeneity in keratin proteins since analysis of chains from adult feather calamus (stalk) of the emu (O'Donnell, 1973b) and gull (O'Donnell and Inglis, 1974), from chick embryonic feather (Walker and Rogers, 1976b) and from chick scale (Walker and Bridgen, 1976) revealed the presence of major regions of sequence homology between the proteins from different tissues. This suggests that avian keratins evolved from a common ancestral protein by gene duplication and divergence leading to the observed group of sequence-related proteins differing only in a few amino-acid sites.

In addition to the group of keratins synthesized in the developing feather, Walker and Rogers (1976a) described a second group of proteins that were synthesized co-ordinately with, but were different in composition from the keratins. This group termed 'Fast' proteins because of their behaviour on pH 2.7 polyacrylamide gel electrophoresis, are slightly smaller than the keratins, having a molecular weight of 8,400, and are virtually devoid of cysteine but relatively rich in histidine, phenylalanine and tryosine. This group, although clearly distinguishable from keratins, may form a separate group of internally homologous species, since Walker and Rogers (1976b) showed that this fraction consisted of at least three species.

4. The nature of embryonic feather keratin mRNA:

The purified fraction of mRNA coding for keratin proteins in both the rabbit reticulocyte and wheat embryo cell-free systems (Partington *et al.*, 1973; Kemp *et al.*, 1974b and 1974c) and isolated from the polysomes of 14-day embryonic feathers maximally synthesizing this mRNA fraction (Powell *et al.*, 1976) sediments at about 12S in sucrose gradients and electrophoreses with a size of 760 nucleotides on acrylamide gels in 98% formamide (Kemp *et al.*, 1974b). From the wheat embryo cell-free translation studies and the observation that purified keratin mRNA electrophoresed as a single band on formamide gel electrophoresis (Kemp *et al.*, 1974b), 95% of the mRNA was judged to code for keratins.

This mRNA was found to have a poly(A) tract at the 3' end with an average length of 60 nucleotides (Morris and Rogers, 1979). The mRNA has a 'cap' structure at the 5' end consisting of 7-methylguanosine at the 5' terminus coupled by a triphosphate linkage to the penultimate residue (Morris and Rogers, 1979).

The hybridization kinetics of keratin mRNA with cDNA, prepared using the RNA dependent DNA polymerase of avian myeloblastosis virus, indicated that keratin mRNA is a heterogeneous mixture of species (Kemp, 1975). From this analysis it was estimated that there were 25 to 35 keratin mRNA species, an estimate consistent with the number of protein chains in the tissue (Walker and Rogers, 1976a). Lockett, Kemp and Rogers (manuscript in preparation) showed that each mRNA species appears to contain a sequence

of about 150 nucleotides adjacent to the 3' poly(A) tract which is unique, i.e., represented only once in the chick genome. A reiterated sequence, which is either short and faithfully conserved or longer and mis-matching was shown to be covalently attached to the 3' unique sequence and located further towards the 5' end of the mRNA. Since only 300 of the 760 bases of this mRNA can be attributed to the coding sequence, Lockett, Kemp and Rogers (in preparation) have suggested that the unique 3' end regions correspond to the untranslated region of the mRNA species and the repeated portions correspond to the coding regions, which must be homologous to some extent, as judged by protein sequence homologies.

It should be noted that no evidence has yet been obtained for the presence or absence of intervening sequences in the natural keratin-coding gene, so it is not yet known whether the 12S mRNA is processed from a precursor by splicing of RNA sequences as described recently for many eukaryotic mRNA species (Adenovirus: Berget *et al.*, 1977; Chow *et al.*, 1977; β-globin: Jeffreys and Flavell, 1977; Tilghman *et al.*, 1978a and 1978b; chick ovalbumin: Mandel *et al.*, 1978; Dugaiczyk *et al.*, 1978, Roop *et al.*, 1978).

Possible roles for intervening sequences or the RNA they code for in the control of gene expression have been proposed (Mandel *et al.*, 1978). No evidence has yet been obtained, however, suggesting that intervening sequences are involved in such processes, nor that they are present to satisfy requirements for mRNA processing or transport etc. In the absence of any such evidence, the suggestion

that they may have evolved to provide a means for rapid gene evolution by changes in mRNA processing (Gilbert, 1978) is an attractive one.

5. The nature of keratin-coding DNA:

The only three features of keratin-coding DNA thus far published are firstly the reassociation kinetic analysis suggesting that between 100 and 240 genes in the chick genome annealed to DNA complementary to the 12S mRNA fraction (Kemp, 1975). It should be noted that this estimate of gene number may include the keratins of other tissues e.g., scale, adult feather etc., as the 12S cDNA may be able to cross-anneal to these related sequences. Secondly, keratin-coding DNA has been shown not to be amplified in the embryonic feathers of thirteen day old chick embryos actively synthesizing the 12S mRNA (Gibbs, 1978). Thirdly, keratin-coding DNA is associated with long stretches of G-C rich DNA in the chick genome as detected by DNA density in isopycnic caesium chloride gradient centrifugation (Lockett, Kemp and Rogers, manuscript in preparation). All of these studies were made using DNA complementary to the purified 12S mRNA as a probe for complementary sequences.

As mentioned above, no evidence has been obtained for the presence or absence of intervening sequences in the natural keratin genes.

6. Summary:

The key features of the expression of keratin genes in the developing embryonic feather are therefore that this tissue is an example of a terminally differentiating

tissue. Cell division ceases at about day 12 or 13 of embryonic development and is followed by the co-ordinate expression of a set of keratin genes coding for a heterogeneous mRNA population and for a set of sequence homologous keratin proteins. Complexity of this set of mRNA and proteins has been observed and it has been suggested that as many as 20 or 30 different species are synthesized during keratinisation of the feather (Kemp, 1975; Walker and Rogers, 1976a).

D. Other Multi-gene Systems.

1. Faithfully repeated genes:

a) Histone genes:

The best studied example of faithfully repeated genes coding for a set of proteins is the set of histone genes expressed in the sea urchin. The five different histones, H1, H2a, H2b, H3 and H4 are clustered into a unit that is tandemly repeated in the genome. Restriction analysis of total sea urchin DNA (*S. purpuratus*, *L. pictus* and *Ps. miliaris*) revealed the existence of a 6 kb repeating unit containing sequences coding for all five histones (Kedes *et al.*, 1975; Weinberg *et al.*, 1975; Schaffner *et al.*, 1976). Gross *et al.*, (1976a) were able to isolate the individual mRNA species and determine the arrangement of these genes using restriction mapping (Schaffner *et al.*, 1976). The repeat unit was cloned into a λ vector and thermal denaturation used to examine A + T rich spacer sequences (Portmann *et al.*, 1976). Using restriction enzymes, controlled λ exonuclease digestion of cloned DNA and hybridization of purified histone mRNAs,

Gross *et al.*, (1976b) demonstrated that the polarity of the histone gene cluster in *Ps. miliaris* is 5'$H_1 \rightarrow H_2a \rightarrow H_3 \rightarrow H_2b \rightarrow H_4$3' with all mRNA species being synthesized from the one strand. Similar restriction (Cohen *et al.*, 1976) and electron microscopic (Wu *et al.*, 1976) studies on *S. purpuratus* have demonstrated that both the polarity of the histone repeat and the A + T richness of inter-gene regions are conserved between these two species. Sequencing studies by Schaffner *et al.*, (1978) have indicated that histones do not seem to be derived from longer precursor proteins, nor is there any evidence for intervening sequences within the coding regions. The A + T rich spacer sequences are made up of relatively simple nucleotide arrangements, but are not repetitious and apparently do not code for additional large proteins.

The histone genes of *Drosophila melanogaster* are repeated 100 times in the genome in two types of repeat unit of lengths 4.8 kb and 5.0 kb (Karp and Hogness, 1976; Lifton *et al.*, 1977). Both repeat units code for all histones, but the direction of transcription is not the same for all five species. H_3, H_2a and H_1 genes are transcribed from one strand while H_4 and H_2b are transcribed from the other (Lifton *et al.*, 1977). Thus co-ordinate expression requires at least two independent initiation events.

b) 5S genes:

DNA coding for the 5S RNA (5S DNA) can be isolated by repeated density gradient centrifugation (Brown *et al.*, 1971). The G + C rich 5S gene was associated

with an A + T rich spacer to form repeating units of 0.7 kb in length (Brown *et al.*, 1971). Restriction analysis of this DNA showed the presence of heterogeneity in the length of the A + T rich spacer sequences which differed by 15 base pair quanta. The A + T rich spacer was shown by sequence analysis to be comprised of tandem repeats of a 15 base pair sequence (Carroll and Brown, 1976a; Brownlee *et al.*, 1974). Variations in the repeat length are thus due to variations in the number of these subrepeats. Carroll and Brown (1976b) demonstrated by molecular cloning that adjacent 5S repeat units could be different in length, and sequence analysis revealed the presence of a 'pseudo-gene' structure (Jacq *et al.*, 1977) that is a perfect copy of 101 residues of the structural gene. Its sequence, however, is such as to be non-functional and perhaps an evolutionary relic of an earlier gene duplication event.

The fact that certain repeat units of these genes are not identical suggests that sudden correction mechanisms for the maintenance of tandemly linked genes (e.g., the 'master-slave' mechanism of Callan and Lloyd, 1960) as well as the expansion-contraction mechanism (Brown *et al.*, 1972) that require tandem homogeneity cannot explain these observations. Similarly, the hypothesis that the evolution of these tandem genes occurred by gene duplication and subsequent divergence is not supported, since variation should not be limited to the number of the 15 base pair subrepeats. (Carroll and Brown, 1976b). The regularity of the observed heterogeneity, however, suggests either that some correction mechanism exists and that the subrepeats

must be involved in this mechanism, or that some unknown selection pressure has prevented modifications other than tandem duplications of the 15 base pair subrepeats. A mechanism of unequal crossing-over between repeat units, similar to that proposed by Smith (1973) which leads to elimination or fixation of variants while permitting heterogeneity is compatible with the known information on 5S gene structure.

2. Heterogeneous multiple genes:

a) Silkmoth chorion proteins:

The silkmoth *Antheraea polyphemus* synthesizes at least 4 molecular weight classes of chorion proteins which are produced by the follicular cells surrounding the oocyte (Regier *et al.*, 1978a and 1978b). The most abundant classes are the A and B classes. Two dimensional gel electrophoresis of the proteins isolated from individual animals gave estimates of at least 18 different A species and 25 different B species (Kafatos *et al.*, 1977). Regier *et al.*, (1978a, 1978b) described the sequence of portions of several of these proteins. All A species showed extensive sequence similarities. The B species examined were distinct from the A species but homologous to each other.

Messenger RNA isolated from the silkmoth chorion and coding for the chorion proteins has been cloned into plasmid PML-21 (Sim Gek Kee *et al.*, 1978) and at least 20 clones corresponding to distinct species have been isolated. Sequence analysis identified these clones as coding for certain chorion proteins and confirmed the sequence homologies

described above. Sequence analysis has also revealed segments of B sequences that are homologous with certain A sequences, even though A and B species are distinct enough to be classified into two different families.

Kafatos *et al.*, (1978) have also used filter annealing procedures to examine the relatedness of different cDNA clones. They established conditions at which each sequence is annealed specifically to perfectly complementary sequences and not to homologous but non-identical sequences.

The nature of silkmoth chorion gene arrangement is yet to be determined and is awaited with great interest.

The properties of the silkmoth chorion system described above are remarkably similar to the keratin system described in this thesis. Chorion synthesis involves the expression of sets of homologous mRNA species and proteins, as does keratin synthesis. Thus a comparison of the two systems as they are elucidated may well be useful in understanding the biological processes involved in the expression of such multiple genes.

b) Immunoglobulin genes:

Immunoglobulin light and heavy chain molecules consist of two regions, a constant (C) region and a variable (V) region that is involved in antibody recognition. The occurrence of linked unique and repetitive sequences in keratin mRNA (Lockett, Kemp and Rogers, manuscript in preparation) was reminiscent of the immunoglobulin system, but recent results describing the gene arrangement of immunoglobulins suggest strongly that this system is unique.

The first set of these results were the nucleic acid hybridization studies of Tonegawa *et al.*, (1976) that indicated the presence of more germ-line sequences coding for the V region than for the C region for a given type of immunoglobulin chain. Restriction enzyme and molecular cloning showed the V and C regions of Kappa and Lambda chains to be separated in DNA from germ-line tissue, but to be close together in DNA from tissue actively synthesizing the antibodies (Hozumi and Tonegawa, 1976; Brack and Tonegawa, 1977; Rabbitts and Forster, 1978), although the final gene does include two intervening sequences within the coding region of the Kappa gene (Tonegawa *et al.*, 1978; Brack and Tonegawa, 1977). Thus DNA rearrangement during development appears to be necessary to form the gene to be expressed in the differentiated tissue.

Seidman *et al.*, (1978a) have recently generated variable region probes against two different Kappa light chain subgroups and with them identified two non-overlapping sets of *Eco*RI generated fragments of mouse DNA. Each set consisted of five to ten *Eco*RI-generated fragments and each fragment contained elements of variable region gene sequences. Since different sets of *Eco*RI-generated fragments carrying V gene elements were detected by probes to different Kappa light chain subgroups, and since 25 to 30 different subgroups have already been identified from amino acid sequence studies, there are probably a minimum of 125 to 150 distinct V genes in the mouse genome. Seidman *et al.*, (1978a) also demonstrated that extensive homology occurred between different V gene regions extending

beyond the structural gene, with sequence variations being mainly limited to those regions coding for the antigen binding site. Such homology was not observed in regions adjacent to two mouse βglobins (Tiemeier *et al.*, 1978). Seidman *et al.*, (1978b) suggest that this homology creates a large target for intragenic recombination. Thus the large number of V genes coupled with the possibility that further heterogeneity is generated during the DNA recombination event may explain the enormous diversity of V region antibody sequences.

Information on the arrangement of various V and C genes in relation to each other both in embryonic and mature tissues is at present very limited. For example it will be of interest to determine if different V genes are linked in an ordered cluster.

3. The molecular genetics of gene expression:

a) Prokaryote gene expression:

An extensive understanding of the molecular nature of gene expression is confined to a few prokaryote operons, such as the *lac* operon and *trp* operon of *E. coli* and the transcriptons of coliphage λ.

In general, the initiation of transcription occurs by specific binding of an RNA polymerase molecule to a recognition sequence (promoter) adjacent to the gene. Control is mediated by the interaction of various effector molecules (e.g., repressor proteins) with sequences (operators) that lead to interference with the RNA polymerase molecule, preventing its progression and thus inhibiting synthesis, or enhancing its affinity of binding

to the promoter and thus stimulating synthesis (Jacob and Monod, 1961).

New mechanisms of gene control in bacteria continue to be discovered. For example Bertrand *et al.*,(1975) have reported the presence of an attenuator region for the premature termination of RNA synthesis, located between the site of initiation and the structural genes of the *trp* operon. Such a mechanism on top of the normal initiation interactions described above allows control of expression at a very fine level.

Such studies of prokaryote gene expression provide models for use in the interpretation of the molecular processes involved in the expression of genes in eukaryotes.

b) Eukaryote gene expression:

As transcription of eukaryote genes must be initiated at some point, it is likely that transcription control regions such as those described above are present adjacent to the coding regions of eukaryotic genes. There is as yet, however, no clearly defined example of such a control region. In any case a more relevent question in terms of an organism capable of enormous morphological diversity is whether such control region models can explain the temporal development of different tissues and the genes expressed during this development. Britten and Davidson (1969) proposed such a model for gene control that is an extension of the bacterial operon form of control. The model suggests that all genes in the set required to be co-ordinately expressed have a common control element adjacent to these genes. Thus

each set of co-ordinately expressed genes would need to have a different control sequence. Expression of a gene on more than one occasion would therefore require more than one control element adjacent to it, each corresponding to the set of genes expressed at the different stages. No direct evidence has been found for this model, although certain fungal genes exhibit the characteristics required for it.

For example the protease of *Neurospora crassa* is synthesized during three metabolic conditions, *viz* limitation of sulphur, nitrogen or carbon. A regulatory mutant lacking a positive signal required for the synthesis of a family of related enzymes for sulphur metabolism could not synthesize the protease in response to a limitation for sulphur and yet the same mutant is capable of producing the enzyme when it is limited for either nitrogen or carbon (Hanson and Marzluf. 1975). A second regulatory mutant defective in the control of nitrogen metabolism fails to synthesize the protease only when it is limited for nitrogen. It should be noted that although several biochemical tests were performed to ensure that the protease synthesized in all cases was the same, a molecular analysis of the gene is required for definitive evidence that this effect is not due to multiple copies of the protease gene. In the absence of this, however, these results do suggest a series of control elements adjacent to the structural gene each one of which is also adjacent to the other genes expressed during the particular metabolic state.

This model should be directly testable once the regions adjacent to co-ordinately expressed genes are isolated.

Sequence analysis of these regions should expose common sequences. The avian keratin system is ideal for testing this model in this way, since it involves the co-ordinate expression of different sets of keratins in different tissues.

In summary, then, it appears that although the control of expression in certain bacterial operons is extremely well understood, similar eukaryotic processes are not. In particular the molecular processes that result in morphological differentiation and differential gene expression in the various tissues are very poorly understood.

E. Aims Of The Project.

1. General aims:

The broad aims of the project are to elucidate the processes involved in the formation and keratinisation of epithelial tissues. The two processes acting as markers for these processes are firstly the development and differentiation of the epidermis into the feather morphology and secondly the synthesis of the keratin proteins during the terminal differentiation of the tissue.

The study of the molecular biology of the developing feather is thus centred around two questions. Firstly, what factors direct the morphological development of the epidermis? Secondly, what are the molecular events and interactions required for the co-ordinate expression of the keratin proteins? A third question of biological interest is, what are the sequence relationships between different keratin species and what do they infer about

the evolution of this set of genes?

Since molecular approaches to the first of these questions is beyond the available technology, and as the second and third questions are becoming increasingly amenable to experimental approaches, particularly with the advent of recombinant DNA techniques enabling the purification and preparation of individual genes, it is these questions that are currently under examination.

2. Specific aims of the project:

The project described in this thesis was aimed at an examination of the nature of the molecular genetics of embryonic feather keratins with respect to gene arrangement, sequence relatedness and gene complexity. Three approaches were to be used for this study. Each of these used restriction DNA technology, that is, the capacity to cleave DNA sequences at specific sites, and the reader is referred to the introduction of Section II for a discussion of the principles and uses of this technology. The first of these was to use 12S cDNA derived from the purified 12S mRNA fraction coding for keratins as a probe to detect complementary sequences (i.e., the genes) in restricted and fractionated DNA. The second was to isolate individual keratin coding mRNA species from the 12S mRNA mixture by molecular cloning for use as probes as described for total 12S cDNA. The third was to isolate chromosomal DNA fragments coding for keratins and thus to examine directly the sequences expressed in the chick feather and sequences associated with them. Regrettably the problems encountered with

requirements for containment facilities have prevented an attempt at the last of these.

CHAPTER 2

MATERIALS AND METHODS

CHAPTER 2.

MATERIALS AND METHODS.

A. Materials.

1. Chemicals:

All chemicals used were of analytical reagent grade, or of the highest available purity. Chemicals listed below were obtained from the source indicated.

a) General chemicals:

Ammonium persulphate enzyme grade: Schwarz-Mann, Orangeburg, New York.

Coalfish DNA: Sigma Chemical Co., St. Louis, Missouri.

CsCl optical grade: Harshaw Chemical Co., Cleveland, Ohio.

Diethylpyrocarbonate: Sigma

Dimethylsulphoxide, acetonitrile and trichloroacetonitrile: purified as described by Symons (1974).

Ficoll 400: Pharmacia.

[^{32}P]orthophosphoric acid, high specific activity in dilute HCl: Australian Atomic Energy Commission, Lucas Heights, New South Wales.

Phenol: B.D.H. Chemicals Ltd., Poole, Dorsett, England, redistilled under N_2 and reduced pressure, stored at -15°C under N_2 prior to use.

Polyethyleneglycol: Union Carbide Corp.

Polyvinylpyrrolidone: May and Baker, West Footscray, Victoria.

POPOP (1,4-bis-(2,5-phenyloxazolyl)-benzene): Sigma

PPO (2,5-diphenyloxazole): Sigma

Sodium Dodecyl Sulphate (S.D.S.): Sigma

Streptomycin sulphate: Sigma

Sucrose-ultrapure: Schwarz-Mann, Orangeburg, New York.

Phage λ Charon 3A DNA to which oligo dA had been added to the 3' ends was the gift of Dr. D.J. Kemp. Chicken erythroblast ribosomal RNA was the gift of P. Krieg. pBR322 DNA carrying chick β-globin cDNA was the gift of R. Richards.

b) Chemicals for electrophoresis:

Acrylamide: Merck, twice recrystallized from $CHCl_3$

Agarose type I, low EEO: Sigma

Ethidium bromide: Sigma

Formamide: B.D.H., deionized as described by Pinder *et al.*, (1974).

N,N'-methylenebisacrylamide: B.D.H., recrystallized from $CHCl_3$

N,N,N',N'-tetramethylethylenediamine (TEMED): Eastman Organic Chemicals, Rochester, New York.

c) Chemicals for *in vitro* DNA synthesis:

Actinomycin D: a gift from Merck, Sharp and Dohme, Rahway, New Jersey.

dATP, dCTP, dGTP, dTTP: Sigma

dithiothreitol: Sigma

2-mercaptoethanol: Sigma

Oligo dT_{10}, free acid: P.L. Biochemicals Inc., Milwaukee, Wisconsin.

2. Proteins and Enzymes:

Proteins and enzymes used for this study were obtained from the source indicated below.

Bovine Serum Albumin, fraction V: Sigma

E. coli DNA polymerase I: Boehringer, Mannheim, Germany

E. coli Deoxyribonuclease I: Sigma

Myokinase: Sigma

Pancreatic ribonuclease type III: Sigma

Proteinase K: E. Merck, Darmstadt, Germany

Pyruvate kinase: Sigma

Restriction endonuclease *Pst*I: New England Biolabs, Beverley, Massachusetts.

Adult chicken erythrocytes were the gift of A. Keeves. Calf thymus terminal deoxynucleotidyl transferase was the gift of Dr. D.J. Kemp. RNA dependent DNA polymerase (reverse transcriptase) was the gift of Dr. J.W. Beard and the N.I.H. Cancer Program. The single-strand specific S_1 nuclease of *Aspergillus oryzae* was purified to the end of step 4 as described by Vogt (1973).

E. coli nucleotide kinase was the gift of Dr. R.H. Symons.

3. Other materials:

a) Tissues:

(i) Embryonic feathers:

Fertilized eggs of White Leghorn fowls (*Gallus domesticus*), strain Para 3, were obtained from the Parafield Poultry Research Station, Parafield, South Australia. The eggs were stored at 10°C for no more than seven days, and incubated at 37°C, 54% humidity in a forced draught incubator. 14 day old embryos were removed from the eggs, and washed with Hanks balanced salt solution (see Paul, 1970). They were then immersed in the same solution and the feathers removed by plucking with jewellers'

forceps. The feathers were then washed in Hanks solution by repeated low speed centrifugation.

(ii) Adult erythrocyte nuclei:

Blood was taken from adult chickens by heart puncture and the cells centrifuged for 5 min at 14,000 r.p.m. The plasma layer was removed and the erythrocytes resuspended in 0.15 M NaCl, 0.5 M KCl and 2 mM $MgCl_2$. The cells were washed twice by centrifugation and resuspension in this buffer and finally resuspended in 5 volumes of 2 mM $MgCl_2$ and incubated at 4°C for 10 min. The nuclei were centrifuged and resuspended in 2 mM $MgCl_2$ until the globin was largely removed.

(iii) Sperm:

Chicken semen provided by the Parafield Poultry Research Station was washed several times by centrifugation and resuspension in SSC.

b) Materials for enzyme preparations:

(i) Bacterial strains:

Bacillus amyloliquefaciens strain H, *Bacillus globigii* and *Escherichia coli* strain RY13 were all obtained from Professor K. Murray. *Haemophilus influenzae* exo$^-$ was obtained from Dr. H.O. Smith and *Streptomyces albus* strain G was obtained from Dr. R. Appels.

(ii) Column materials:

Sephadex G-50 was purchased from Pharmacia. Phosphocellulose P-11 and DEAE cellulose DE52 were purchased from Whatman.

c) Other materials:

Nitrocellulose was purchased from Sartorius.

Polyethyleneimine (PEI) thin layer, plastic backed plates were Macherey-Nagel polygram CEL300 PEI/UV_{254}.
The bacterial strain ED8654:$r_k^- m_k^+$ *supE supF trp R*, used for transformation and transfection was obtained from Professor K. Murray.

B. Methods.

1. General methods:

a) Media:

Media used for this study are described in Section II, Chapter 7 with the exception of modified M-9 medium which was made up as follows: 6 g of NaH_2PO_4, 3 g of KH_2PO_4, 0.5 g of NaCl, 1 g of NH_4Cl and 5 g of Casamino acids were dissolved in 1 litre of H_20, autoclaved and cooled. This medium was supplemented with 0.1 ml of 1 M $CaCl_2$, 1 ml of 1 M $MgSO_4$ 5 mls of 40% glucose, 1 ml of 0.2% vitamin B_1, 2 mls of 4% L-leucine and 2 mls of 4% L-proline.

b) Buffers and the preparation of solutions:

All solutions were prepared in glass-distilled water, and treated with diethylpyrocarbonate, followed by autoclaving, to eliminate any contamination with nuclease. Glassware was rendered nuclease-free by either autoclaving, incubation at 110° overnight, washing with 1 M KOH followed by rinsing with sterile glass-distilled water, or by a combination of these procedures. Spatulas, etc., were washed with alkali and rinsed in sterile water, as described. Pipettes and micropipettes were washed in glass-distilled water containing diethylpyrocarbonate and dried for 16 hours at 110°.

The buffers commonly used in this study were:

SSCE: 0.15 M NaCl, 0.015 M Na_3 citrate, 1 mM EDTA

TE: 10 mM Tris-HCl pH 7.5 (unless otherwise indicated), 1 mM EDTA.

c) Determination of radioactivity:

The level of radioactivity of ^{32}P-containing materials was generally determined by measuring the Cerenkov emissions using the ^{3}H settings of a Packard tri-carb scintillation counter. The level of radioactivity of ^{3}H-containing materials was determined in one of three ways. Total counts were determined by spotting the material onto a GF-A filter, drying the filter, immersing the filter in liquid scintillation fluid (3.5 g PPO, 0.35 g POPOP per litre of toluene) or in the case of volumes up to 10 µl, by dissolving in 2 ml of "toluene-triton" scintillation fluid (5 volumes of toluene scintillation fluid plus 3 volumes of triton X-100). The level of radioactivity incorporated into TCA insoluble material was determined by adding the sample to 5 ml of ice-cold 10% TCA containing 20 µg of carrier yeast tRNA, vortex mixing, filtering through GF-A filters, washing sequentially with 1 N HCl and ethanol, drying under an infra-red lamp and immersing the filter into toluene scintillation fluid.

2. Preparation of purified keratin messenger RNA:

12S mRNA coding for keratin proteins was isolated from the feathers of 14 day old chick embryos by the dissociation of polysomes with EDTA and repeated sucrose gradient centrifugation finally in the presence of SDS, essentially as described by Kemp *et al.*, (1974b).

3. Preparation of DNA:

a) Chick erythrocyte DNA:

DNA was isolated from chick erythrocyte nuclei obtained from a single chicken, essentially as described by Gross-Bellard *et al.*, (1973) as follows. Erythrocyte nuclei, prepared as described, were resuspended in 2 volumes of 0.15 M NaCl, 0.5 M KCl and 2 mM $MgCl_2$ and added dropwise to 100 ml of 1% SDS to lyse the nuclei. 25 ml of 4 M NaCl was added to the lysed nuclei and stirred gently until a homogeneous solution resulted. This solution was extracted twice with an equal volume of a 24:1 mixture of chloroform and n-octanol and dialyzed extensively against 10 mM Tris-HCl pH 8.0, 10 mM NaCl. Heat treated pancreatic RNase (80°C for 20 min to destroy any DNase contamination) was added to a final concentration of 50 µg/ml and incubated for 2 hr at 37°C. The solution was then brought to 10 mM in EDTA, 0.5% in SDS and 50 µg/ml in proteinase K and digested at 37°C overnight. The solution was extracted three times with an equal volume of water saturated phenol and dialyzed extensively against TE. The DNA was assayed for purity by measuring A_{260}/A_{280} and A_{260}/A_{230} which were always 1.8 - 2.0 and 2.0 - 2.3 respectively (results not shown).

b) Embryonic chick feather and chick sperm DNA:

Embryonic feathers prepared from 13 day old embryos as described and chick sperm from adult roosters prepared as described were suspended in 10 volumes of 10 mM Tris-HCl pH 8.0, 10 mM NaCl, 10 mM EDTA, 0.5% SDS and 100 µg/ml proteinase K and incubated at 37°C for 6 hours.

The digest was extracted 3 times with an equal volume of water-saturated phenol and dialyzed extensively against 0.1 x SSC. Pancreatic RNase (heat treated as described) was added to a final concentration of 100 µg/ml and incubated. at 37°C for 4 hours. The DNA was extracted several times with an equal volume of water saturated phenol and dialyzed extensively against TE.

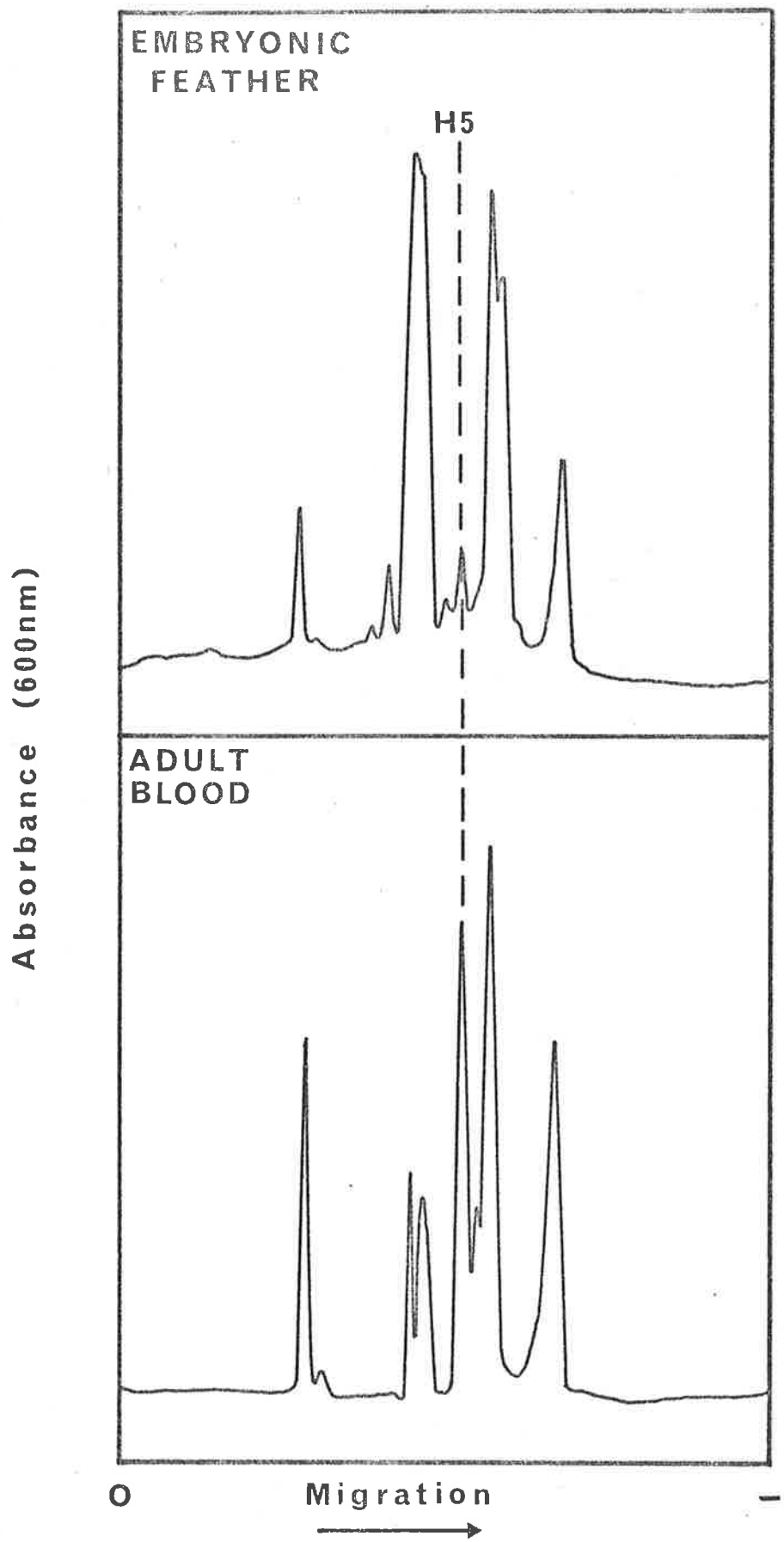
To test whether DNA isolated from 13 day feathers was contaminated with DNA from erythrocytes in the feather vein, the acid-soluble histones of the feather tissue were isolated from nuclei as described by Appels and Wells (1972) and electrophoresed on 15% polyacrylamide pH 2.7 gels (Panyim and Chalkley, 1969). Fig.2.1 shows that the erythrocyte specific histone H5 was present only as a minor contaminant of the histone preparation. Quantitation of the H5 band indicated that 15% of the isolated nuclei were erythocyte contaminants, although this is likely to be an overestimate, as this method of nuclei isolation favours the retention of the softer erythrocyte tissue and the loss of the tougher feather tissue. At least 85% of DNA from 13 day embryonic feathers was therefore derived from feather tissue, again confirming that had there been any gene rearrangement, it would have been readily detected by this method.

c) Plasmid DNA:

E. coli carrying plasmid pBR322 or any cloned derivative of it was grown at 37°C with aeration in 800 ml of modified M-9 medium to an A_{600} of 0.8. Chloramphenicol was added to a final concentration of

FIGURE 2.1. Analysis of Embryonic Feather Histones.

Nuclei and acid-soluble proteins were isolated from 13-day embryonic feathers and electrophoresed in 15% polyacrylamide at pH 2.7. The gels were stained with amido black and scanned in a Gilford spectrophotometer at 600 nm. Marker histones isolated from adult chick erythrocytes were the gift of A. Keeves and were electrophoresed under the same conditions. The position at which the erythrocyte specific histone H5 electrophoresed is marked by the dotted line.


EMBRYONIC
FEATHER
H5
ADULT
BLOOD
Absorbance (600nm)
O
Migration
–

50 µg/ml and the culture was left incubating with aeration at 37°C overnight. The cells were chloroform treated to kill the bacteria, centrifuged at 10,000 r.p.m. for 10 min, and washed in 40 ml of TE. The cells were centrifuged and resuspended in 12 mls of 25% sucrose in 0.05 M Tris-HCl pH 8.0 at 4°C 4 ml of fresh 10 mg/ml lysozyme was added and the mixture swirled gently for 5 min. 4 ml of 0.5 M EDTA pH 9.0 was then added and the mixture swirled gently for a further 5 min. 20 ml of a solution of 0.1% triton X-100, 0.0625 M EDTA and 0.05 M Tris-HCl pH 8.0 was then added and the mixture swirled gently for 10 min before being centrifuged at 40,000 r.p.m. for 30 min in a Beckman Ti50 head and L5-50 centrifuge. DNA in the supernatant was precipitated at 4°C overnight by the addition of solid NaCl to 0.5 M and solid polyethyleneglycol to 10% (w/v). The DNA was pelleted by centrifugation at 12,000 g for 10 min and the pellet resuspended in TE. One gm of solid caesium chloride and 0.1 ml of a 10 mg/ml solution of ethidium bromide were added per ml of TE and the solution centrifuged at 40,000 r.p.m. for 3 days. The lower (plasmid) band was removed by side puncture with a syringe, extracted with iso-amyl alcohol until the aqueous phase was free of ethidium bromide and finally dialyzed extensively against TE.

All DNA solutions were stored in TE at 4°C.

4. Restriction endonuclease preparation and digestion conditions:

a) Preparation of restriction endonucleases:

*Bam*HI was prepared by the method of Wilson

and Young (1975). *Bgl*II was prepared as follows: *Bacillus globigii* was grown in Luria broth (See Chapter 7) supplemented with 0.1% (w/v) glucose to late log phase and harvested by centrifugation. All subsequent steps were carried out at 4°C. 20 Grams of cells were resuspended in 100 ml of 50 mM Tris-HCl pH 7.4, 10 mM 2-mercaptoethanol and 10 mM $MgCl_2$ and passaged through a French Pressure Cell at 15,000 p.s.i. Cell debris was removed by centrifugation at 38,000 g for 30 min. Nucleic acids were precipitated by the addition of 1 ml of fresh 10% (w/v) streptomycin sulphate per 1500 A_{260} units and centrifugation at 12,000 g for 10 min. The supernatant was decanted and made 50% (w/v) with solid ammonium sulphate, stirred for 30 min and centrifuged at 12,000 g for 10 min. The supernatant was decanted and made 70% (w/v) with solid ammonium sulphate, stirred for 30 min and centrifuged at 12,000 g for 10 min. The pellet was resuspended in 20 ml of 20 mM $NaPO_4$ pH 7.4, 7 mM 2-mercaptoethanol and 2 mM $MgCl_2$ (buffer A), dialyzed extensively against this buffer and applied to a pre-washed phosphocellulose column (Whatman P-11) equilibrated with buffer A. The enzyme was eluted with a one litre 0 to 1 M NaCl gradient, active fractions pooled and dialyzed against buffer A in 50% glycerol. The enzyme isolated by this method is free of detectable levels of other double-strand nuclease activity. DNA cleaved with *Bgl*II prepared by this method has also been used effectively for ligation (Finnegan and Egan, 1979).

*Eco*RI was prepared essentially as described by Yoshimori (1971).

*Hin*dIII was prepared essentially as described by De Fillipes (1974) for *Hpa*II.

*Sal*I was prepared as follows: *Streptomyces albus* strain G was grown to late log phase as described for *Bacillus globigii* and harvested by centrifugation for 10 min at 12,000 g. All subsequent steps were carried out at 4°C. Cells (54 g wet weight) were resuspended in one volume of 50 mM Tris-HCl pH 7.4, 14 mM 2-mercaptoethanol and 10 mM $MgCl_2$ and passaged through a French Pressure Cell at 15,000 p.s.i. Cell debris was removed by centrifugation at 35,000 r.p.m. in a Beckman Ti50 head and L5-50 centrifuge. 1 ml of freshly prepared 10% (w/v) streptomycin sulphate was added to the supernatant per 1500 A_{260} units, stirred for 30 min and centrifuged for 10 min at 10,000 r.p.m. in a Beckman J21-B preparative ultracentrifuge. The supernatant was decanted, made 50% (w/v) with respect to ammonium sulphate by the addition of solid ammonium sulphate, stirred for 30 min and centrifuged for 10 min at 12,000 g.

The pellet was resuspended in 20 ml of 10 mM KPi pH 7.6 and dialyzed extensively against this buffer. This enzyme fraction was loaded onto a 3 x 6 cm DEAE-cellulose column flowing at 5 ml/hr and eluted with a 100 ml 0 - 0.8 M KCl gradient. The 3 ml fractions were assayed for cleavage of phage λ DNA. Active fractions were pooled and dialyzed against 50% glycerol, 0.2 M NaCl, 20 mM Tris-HCl pH 7.6. This preparation yielded over 20,000 units of enzyme activity, where 1 unit is defined as that amount required to digest 1 µg of λ DNA to completion in 1 hr at 37°C. The enzyme was also free of detectable levels of

other double-stranded DNA nucleases, although the preparation was never assayed for single-stranded nuclease activity.

b) Digestion conditions:

Conditions of restriction endonuclease cleavage of DNA were as follows: *Bam*HI and *Bgl*II digestions were carried out in 10 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, and 10 mM 2-mercaptoethanol at 30°C. *Hin*dIII digestions were as described for *Bam*HI, except that 50 mM NaCl was included and the digestions incubated at 37°C. *Eco*RI and *Sal*I digestions were as described for *Bam*HI, except that 100 mM NaCl was included and the digestions incubated at 37°C. Phage λ DNA was included in the total chick DNA digests to ensure that digestion had gone to completion. Fig.2.2 Tracks A and B show the results of including 0.2 µg of λ DNA in a 30 µg chick digest, electrophoresing and transferring it to nitrocellulose and annealing it to ^{32}P labelled λ DNA. Digestion was complete, as judged by the absence of partial digests. These λ fragments were used for molecular weight estimations of restricted chick DNA fragments annealing to the various probes.

c) Preparation of digested DNA for electrophoresis:

30 µg aliquots of chicken DNA restricted for agarose gel electrophoresis and transfer to nitrocellulose were extracted with an equal volume of a 1:1 mixture of phenol and chloroform, brought to 0.1 M with respect to NaCl and precipitated by the addition of three volumes of ethanol and chilling in dry ice-isobutanol for 10 min. The DNA was pelleted in an MSE bench centrifuge for 10 min, the supernatant drawn off and 50 µl of 0.1% SDS in TE added

to resuspend the DNA. In some cases it was necessary to incubate the DNA at 65°C for 5 - 10 min to achieve resuspension. This incubation, or a 5 - 10 min incubation in a vacuum dessicator was used to remove any remaining ethanol. 5 µl of 50% glycerol containing 0.025% bromophenol blue were added and the DNA loaded onto the gel.

5. Gel Electrophoresis:

a) Agarose gel electrophoresis:

Electrophoresis of restricted chicken DNA for transfer to nitrocellulose was carried out on vertical 20 x 30 x 1 cm slab 1% agarose gels. Aliquots of 30 µg of restricted DNA were loaded onto 1 cm x 1 cm slots of the vertical gel. Electrophoresis buffer consisted of 40 mM Tris-acetate pH 8.2, 20 mM sodium acetate and 1 mM EDTA and electrophoresis was carried out at 180 V for 16 hr. Ethidium bromide (0.4 µg/ml) was included in the gel and anode electrophoresis tank.

Electrophoresis of restricted DNA for analytical purposes was carried out on vertical 20 x 30 x 0.3 cm slab 1.0% agarose gels under the same conditions as described above, except that ethidium bromide was not included in the gel. Staining was achieved by immersing the gel in a 0.4 µg/ml solution of ethidium bromide for 1/2 - 1 hr and destaining in H_2O for 1/2 - 1 hr.

All ethidium bromide stained gels were photographed under short wave ultraviolet light on Kodak 2475 recording film using a red filter.

b) Aqueous polyacrylamide gel electrophoresis:

Electrophoresis of double-stranded DNA

species less than 1 kb in length was carried out on vertical 20 x 20 x 0.2 cm or 16 x 16 x 0.2 cm slab polyacrylamide gels of varying percentages exactly as described by Maniatis *et al.*, (1975b) using a 29.1 ratio of acrylamide to N,N',-methylenebisacrylamide and 0.09 M Tris-borate pH 8.3, 2.5 mM EDTA as the buffer system (Peacock and Dingman, 1968).

c) Formamide polyacrylamide gel electrophoresis:

Electrophoresis of denatured or single-stranded DNA was carried out on a vertical 20 x 20 x 0.2 cm slab apparatus containing 5% polyacrylamide in 98% formamide as described by Maniatis *et al.*, (1975b). 100 mg of ammonium persulphate, 170 mg of anhydrous Na_2HPO_4 and 45 mg of dihydrous NaH_2PO_4 were dissolved in 1 ml of H_2O. This solution was mixed with 74 ml of deionized formamide containing 3.19 g acrylamide and 0.56 g bis-acrylamide. 150 µl TEMED was added, mixed and the gel poured. The reservoir buffer consisted of 20 mM sodium phosphate pH 7.5 and was circulated during electrophoresis. Electrophoresis was carried out at 200 V. Samples to be electrophoresed in 98% formamide were ethanol precipitated as described above, resuspended in 98% formamide and incubated in a boiling water bath for 2 min before loading.

6. Transfer of restricted and fractionated DNA to nitrocellulose:

Restricted chick DNA fractionated on slab 1% agarose gels was transferred to nitrocellulose filter paper by a modification of the procedure described by Southern (1975). The entire gel was soaked in 0.5 M NaOH,

0.8 M NaCl for 1 hr to denature the DNA and then soaked in 0.5 M Tris-HCl pH 7.2, 1.6 M NaCl for at least 1 hr to neutralize the gel. The gel was then placed onto wet 3 MM paper which acted as a wick by dipping into 10 x SSC. A piece of pre-wet nitrocellulose filter paper was placed on the gel, followed by several layers of dry 3 MM filter paper and a wad of paper towels to blot the DNA onto the nitrocellulose filter. The transfer was allowed to proceed overnight. The filter was then rinsed briefly in 2 xSSC, air dried at 50°C for 30 min and baked in vacuo at 80°C for at least 90 min. Fig.2.2 shows the results of transferring λ DNA included in the 30 µg *Hin*dIII and *Eco*RI digests as described. The filters were annealed to ^{32}P-labelled λ DNA and autoradiographed. Fig.2.2 Tracks A and B show that there was no preferential loss of any size class of DNA during the procedure.

7. Preparation of labelled nucleotides:

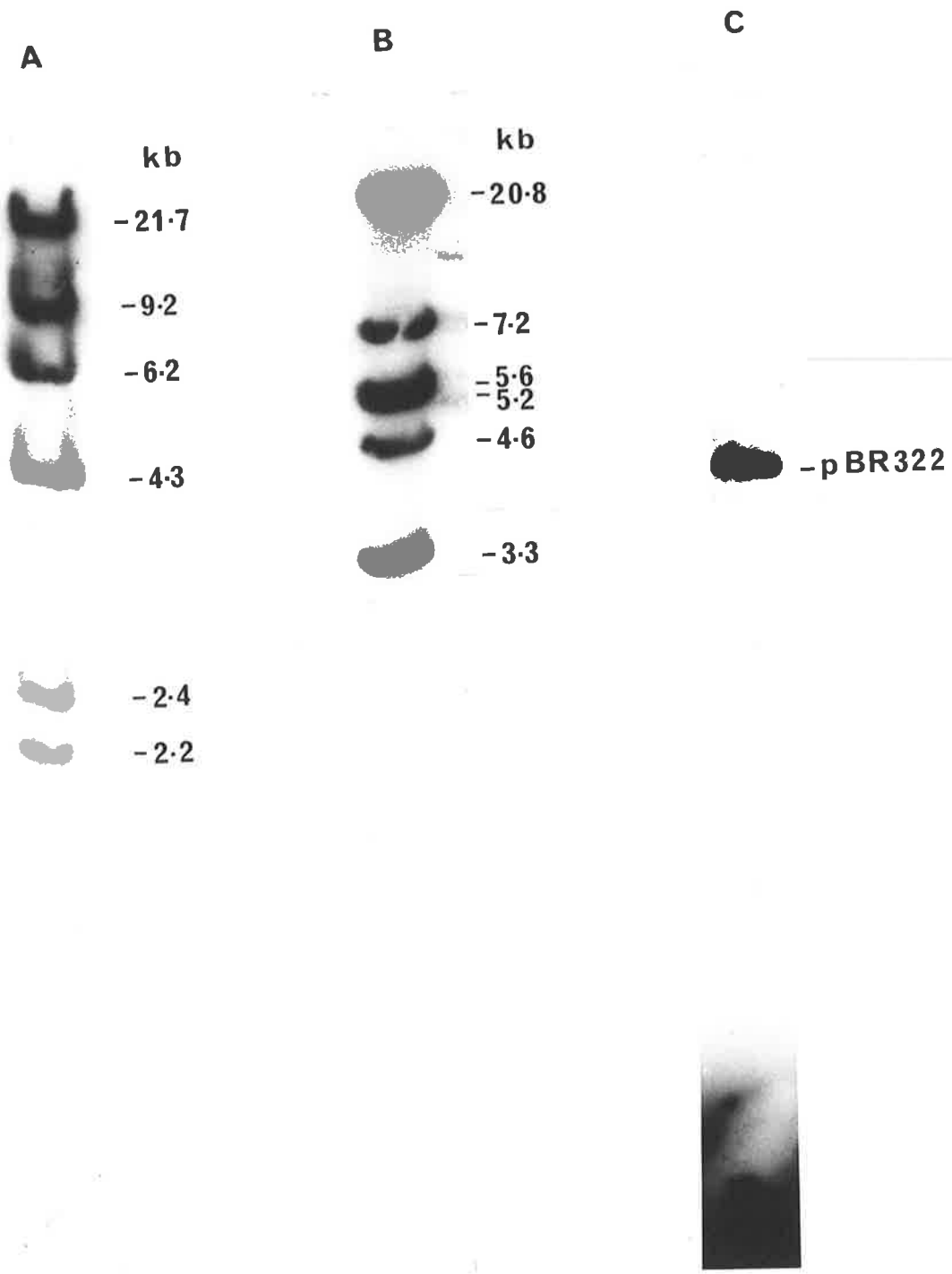
^{32}P-labelled dATP (initial specific activity 50 Ci/mmole) used in reverse transcription reactions was prepared as described by Symons (1977) and was the gift of Dr. R.H. Symons.

High specific activity nucleotides used in the nick-translation labelling of double-stranded DNA were prepared by a modification of the procedure described by Symons (1974) using the nucleosides deoxyadenosine, deoxycytidine and thymidine as a mixture. Deoxyguanosine was not included in the preparation because it is converted to the tri-phosphate form much more efficiently than the other nucleosides (R.H. Symons, personal communication).

FIGURE 2.2. Digestion of λ DNA and Nick-translation of pBR322 DNA.

0.1 μg of λ*c*I857 DNA was included in 30 μg of chick DNA digested with *Hin*dIII (Track A) or *Eco*RI (Track B), electrophoresed in 1% agarose, transferred to nitrocellulose, annealed to ^{32}P-labelled λ DNA, washed, dried and autoradiographed. Tracks A and B are the resulting autoradiogram.

Track C shows pBR322 DNA, labelled with ^{32}P using the 'nick translation' method, electrophoresed on a separate 1% agarose gel. The resulting product runs at the position of full length circular molecules.


A
kb
-21·7
-9·2
-6·2
-4·3
-2·4
-2·2
B
kb
-20·8
-7·2
-5·6
-5·2
-4·6
-3·3
C
-pBR322

4 mg each of deoxyadenosine, deoxycytidine and thymidine were weighed into a 5 ml round bottomed flask and dissolved in 15 mCi of carrier free ^{32}P and sufficient cold H_3PO_4 to give a final specific activity of 200 Ci per mmole. 2 - 3 drops of redistilled triethylamine were added to neutralize the solution and the neck of the flask was rinsed down with 0.5 ml of ethanol. The flask contents were taken to dryness on a rotary evaporator and further dried by the evaporation of two 0.5 ml lots of acetonitrile.

200 µl of dimethyl sulphoxide, 1 µl of triethylamine and 1 µl of trichloroacetonitrile were then added and the stoppered flask was well shaken to dissolve the contents. After incubation for 30 min at 37°C, a few drops of ethanol were added to the pale yellow reaction mixture and all volatile material removed on the rotary evaporator (the dimethyl sulphoxide remained). While the rotary evaporation was proceeding, the kinase reaction mix was prepared. 0.6 ml of a solution of 20 mM KCl, 10 mM magnesium acetate, 50 mM Tris-HCl pH 8.0 was supplemented with 25 µl of 0.1 M potassium phosphoenolpyruvate, 30 µl of 0.1 M ATP, 1 µl of 2-mercaptoethanol, 25 µl of *E. coli* kinase, 2 µl of pyruvate kinase (10 mg/ml) and 1 µl of myokinase (5 mg/ml). This mixture was then added rapidly to the reaction flask which was incubated at 37°C. After 15 - 20 min, a small sample was removed, diluted with water and analysed by thin-layer chromatography on a 1.5 x 5 cm strip of PEI-cellulose using 0.5 M NH_4HCO_3 as solvent. Markers of dAMP and dATP were dried onto the origin prior to application of the sample. After running for about 5 min, the thin-

layer was dried, the markers located under ultra-violet light and the regions corresponding to dATP, dADP and dAMP (plus the region in front of dAMP so as to include Pi) were cut out and counted by Cerenkov radiation.

Provided the kinase reaction was proceeding satisfactorily as indicated by the conversion of dNMP to dNTP, the reaction mixture was diluted with 4 ml of water (usually after a total incubation time of 30 - 40 min) and run onto a 0.85 x 3.0 cm column of DEAE-cellulose (HCO_3^-), prepared by successive elution with several column volumes of 0.1 M NaOH, water, 0.5 M NH_4HCO_3 and 0.07 M NH_4HCO_3. The column was washed with 20 ml 0.07 M NH_4HCO_3 to elute ^{32}Pi, [^{32}P]dNMP and α-[^{32}P]dNDP and then the α-[^{32}P]dNTP was eluted with 10 ml 0.25 M NH_4HCO_3 into a 50 ml Quickfit flask with a medium length neck. After the addition of 1.5 ml triethylamine and 3 ml ethanol, the flask contents were taken to dryness on the rotary evaporator. Two further 3 ml lots of ethanol were added and the flask contents taken to dryness. The final product was dissolved in 1.5 mls of 0.1 mM EDTA.

All evaporations were carried out using a rotary evaporator connected to a water pump and with a water bath at about 40° for heating.

For radiation protection, reaction flasks in Perspex beakers were handled with lead-impregnated gloves and all work was carried out behind Perspex screens and a lead impregnated apron.

8. *In vitro* synthesis of labelled DNA:

a) Oligo dT primed reverse transcription:

Oligo dT primed reverse transcription of mRNA was carried out in 25 µl reaction mixes containing up to 2.5 µg mRNA, 0.66 mM each of dCTP, dGTP, and dTTP, 0.06 M ^{32}P-labelled dATP (10 - 50 Ci/mmole), 8 mM DTT, 10 mM $MgCl_2$, 50 mM Tris-HCl pH 8.3, 10 µg/ml oligo dT and 100 µg/ml actinomycin D. 1 µl of reverse transcriptase was added and the solution incubated at 42°C for 15 min. The RNA template was removed by alkaline hydrolysis with 0.3 M NaOH for at least one hour at 37°C. The solution was neutralized by the addition of HCl to 0.3 M and Tris-HCl pH 7.5 to 0.1 M. 20 µg of sonicated, denatured coalfish DNA was added as carrier and the mix extracted with an equal volume of a 1:1 mixture of phenol and chloroform and finally with an equal volume of chloroform. The aqueous phase was removed and loaded onto a 0.6 x 30 cm Sephadex G-50 column and eluted with TE buffer to remove unincorporated nucleotides.

b) Random primed reverse transcription:

Priming of 18S and 28S ribosomal RNA lacking a 3' terminal poly A tract was achieved by the random hybridization of oligo nucleotides of salmon sperm DNA, prepared as described by Taylor *et al.*, (1976).

Conditions for the synthesis of this cDNA were as described for the oligo dT primed reaction, except that oligo dT was replaced by a final concentration of 2 mg/ml of oligonucleotide and the synthesis proceeded for 60 min at 37°C. The cDNA synthesized was isolated as described for oligo dT primed synthesis.

c) Nick translation of double-stranded DNA:

Labelling double-stranded DNA using *E. coli* DNA polymerase I (Rigby *et al.*, 1977) was carried out essentially as described by Maniatis *et al.*, (1975a). The 50 µl incubation mix contained 50 mM Tris-HCl pH 7.8, 5 mM $MgCl_2$, 10 mM 2-mercaptoethanol, 50 µg/ml of bovine serum albumin, 5 µM of each of ^{32}P-dATP, ^{32}P-dCTP, ^{32}P-dTTP (100 - 200 Ci/mmole each) and unlabelled dGTP. The DNA was nicked by the addition of 20 pg of *E. coli* DNase I and incubation at room temperature for 1 min before the reaction was started by the addition of 2 units of *E. coli* polymerase I. The solution was incubated at 15°C for 90 min. Plasmid pBR322 DNA labelled in this way was full length on a 1% agarose gel (Fig.2.2 Track C) and labelled to a specific activity of 1 - 5 x 10^7 c.p.m. per µg.

9. Filter annealing:

a) Pre-annealing:

Nitrocellulose filter papers to be incubated with labelled single-stranded DNA were pre-incubated in 2 x SSCE containing 0.2% each of bovine serum albumin, ficoll and polyvinylpyrrolidone (modified from Denhardt, 1966) overnight at 65°C.

b) Annealing:

The pre-annealed filters were blotted between 3 MM paper and dipped through the labelled probe in the minimum volume of 2 x SSCE, 0.5% SDS and annealed at 65°C for 16 - 24 hours.

(i) Single-stranded DNA:

When annealing cDNA, 0.5 x 10^6 c.p.m.

of probe per ml of solution was used to anneal to the restricted DNA and 100 µg per ml of unlabelled sonicated and denatured coalfish DNA was used as competitor for non-specific binding to the nitrocellulose.

(ii) Double-stranded DNA:

When annealing nick-translated, double-stranded DNA, 5×10^6 c.p.m. of probe per ml of solution was used to anneal to the restricted DNA and 10 µg per ml each of *E. coli* and pBR322 DNA were included as competitor. The probe in TE was denatured by heating in a boiling water bath for 5 min before being chilled on ice, made up to 2 x SSCE, 0.5% SDS and used to wet the filters as described.

c) Washing and exposing filters:

Filters were washed in several changes of 1.0 x SSCE, 0.5% SDS at 65°C (unless otherwise stated) over a 16 - 24 hr period. The washed filters were blotted with 3 MM paper, air dried and exposed to Fuji Rx medical X-ray film at -80°C using Ilford fast-tungstate intensification screens.

d) The kinetics of annealing:

The kinetics of annealing under these conditions was complex. Although theoretically in cDNA excess, the amount annealed could be increased by increasing the concentration of counts in the annealing solution (Fig.2.3(a)). The amount annealed could also be increased by increasing the amount of DNA bound to the filter (Fig. 2.3(b)). Although the number of counts used for the experiment shown in Fig.2.3 is below that used for other filter

FIGURE 2.3. Kinetics of Filter Annealings.

Varying concentrations of ^{3}H-labelled 12S cDNA were annealed to varying amounts of total chick DNA immobilized to nitrocellulose filter discs of 1 cm diameter. Annealing was carried out in 2 x SSCE at 65°C for the length of time indicated on the horizontal axis before being removed, washed for 24 hr in 2 x SSCE and the radioactivity determined as described in Chapter 2.

a) Effect of varying the amount of chick DNA bound to the filter:

□ 30 µg

■ 20 µg

● 10 µg

△ No DNA

Concentration of probe was 120 c.p.m./µl; 4 pg/µl. Specific activity was 3×10^{7} c.p.m./µg.

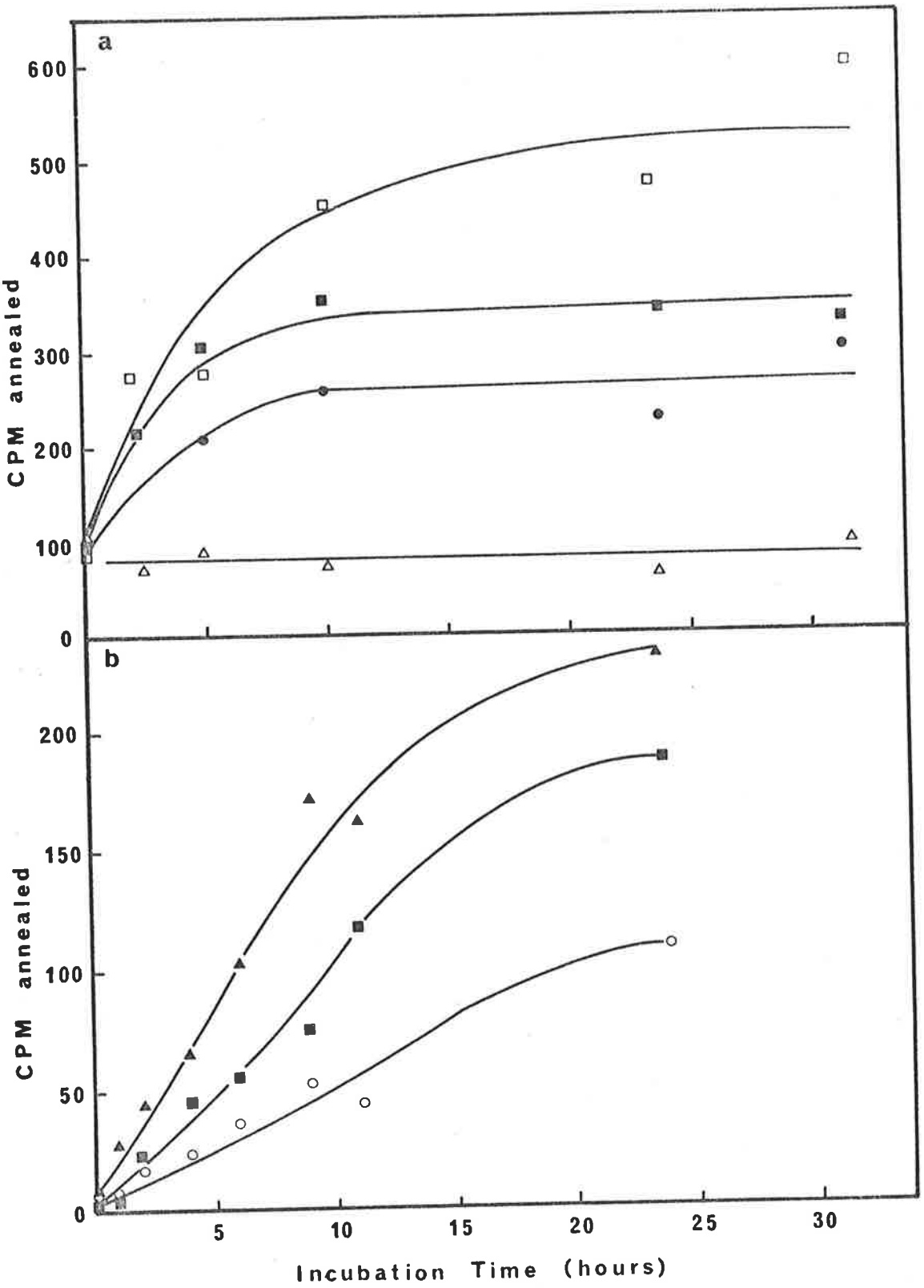
b) Effect of varying the concentration of probe binding to 20 µgm of immobilized chick DNA:

▲ 120 c.p.m./µl

■ 60 c.p.m./µl

○ 30 c.p.m./µl

The 12S cDNA was the same specific activity in both cases.


a
b
CPM annealed
CPM annealed
600
500
400
300
200
100
0
200
150
100
50
0
5
10
15
20
25
30
Incubation Time (hours)

annealings, the specific activity of the probe was less such that the cDNA concentration was of the same order.

Although difficult to explain, it should be noted that a similar observation on the effect of probe concentration when theoretically in excess was made by Grunstein and Hogness (1975) in filter hybridizations.

10. Preparation and tailing of double-stranded cDNA:

Details of the development of this rapid technique for synthesizing double-stranded cDNA are given in the text of Chapter 4. The final procedure used for the preparation of double-stranded cDNA for molecular cloning is described below and is a modification of the procedure described by Rougeon and Mach (1976).

a) Synthesis of the first strand:

Synthesis of the first strand on the 12S mRNA template was carried out as described in 2.11(b), except that actinomycin D was excluded from the mixture, the concentration of Tris-HCl pH 8.3 was 10 mM the concentration of the labelled nucleotide was increased to 0.1 mM and the specific activity of this nucleotide decreased to 1 - 2 Ci per mmole to ensure maximal synthesis of full length cDNA molecules containing a low proportion of labelled nucleotides that would have introduced single-stranded breaks upon decay.

b) Synthesis of the second strand:

Following synthesis of the first strand the reaction mixture was heated in a boiling water bath for 2 min, chilled on ice and a further 0.25 mM (final concentration) of nucleoside triphosphates and 0.001 M of DTT were

added. Reverse transcriptase (0.5 µl per 10 µl of reaction mix) was added and the solution incubated at 37°C for 5 hr. Fresh reverse transcriptase was added after 2.5 hr of incubation.

c) S_1 cleavage of the dscDNA:

The reaction mix was diluted ten fold into S_1 buffer (0.03 M NaOAc pH 4.6, 0.3M NaCl, 4.5 mM $ZnCl_2$) containing 0.5 units per µl and incubated at 37°C for 30 mins.

d) Size fractionation of the dscDNA:

The S_1 nuclease treated dscDNA was extracted with an equal volume of a 1:1 mixture of phenol and chloroform, chromatographed on a Sephadex G-50 column to remove unincorporated nucleotides, concentrated by rotary evaporation, loaded directly onto a 5 - 20% (w/v) sucrose gradient in TE buffer plus 0.1 M NaCl and centrifuged at 38,000 r.p.m. for 16 hr. Fractions were collected (see Chapter 4) pooled, precipitated by the addition of 3 volumes of ethanol, pelleted by centrifugation at 30,000 g for 30 min and resuspended in TE buffer.

e) Tailing the dscDNA:

Homopolymeric nucleotide tails were added to the dscDNA using calf thymus terminal deoxynucleotidyl transferase. 500 pmoles of ^{3}H-dCTP was dried down and resuspended in a solution of 0.14 M NaCacodylate, 0.03 M Tris (pH 7.6), 0.2 mM DTT, 4 mM $MgCl_2$ and 0.07 pmoles of dscDNA. 0.5 µl of the terminal deoxynucleotidyl transferase was added and the reaction at 4°C followed by the conversion of the ^{3}H-dCTP to a trichloroacetic acid insoluble form.

When an average of 10 - 20 nucleotides per end had been added, the reaction was stopped by the addition of an equal volume of 20 mM EDTA.

11. Annealing tailed cDNA and vector DNA and transformation of *E. coli*:

a) Annealing to vector:

0.04 pmoles of pBR322 DNA, cleaved with *Pst*I and tailed with deoxyguanosine as described above, or 0.02 pmoles of λ Charon 3A DNA tailed with deoxyadenosine at the RI cleavage site, were annealed to an equimolar amount of tailed cDNA in 0.2 M NaCl, 0.01 M Tris-HCl pH 8.2 by heating for 10 min at 65°C, incubating for 1 hr at 45°C and finally allowing the solution to cool slowly to 4°C at which it was stored.

b) Procedure for transfection/transformation:

E. coli strain ED8654:$r_k^- m_k^+$ *supE supF trpR* was grown overnight at 37°C in Luria broth and then diluted 1/50 into fresh Luria broth and grown at 37°C to an A_{600} of 0.6 - 0.65. The cells were chilled on ice for 30 min, pelleted by centrifugation and resuspended in 1/2 volume of ice-cold 0.1 M $MgCl_2$. The cells were pelleted immediately and resuspended in 1/20 volume of ice-cold 0.1 M $CaCl_2$. The cells were kept on ice for at least one hour. 0.2 ml of these competent cells was added to 0.1 ml of the DNA in 0.1 M Tris-HCl pH 7.1 and stirred occassionally at 4°C for 30 min. The cells were heat shocked at 42°C for 2 min, kept on ice for a further 30 min, and warmed slowly to room temperature.

L-broth (0.5 ml) was added to the pBR322 transformed

cells and incubated at room temperature for 30 min. The transformed or transfected cells were mixed with 3 ml of 0.7% T or L-agar (supplemented with 0.01 M $MgSO_4$ when plating λ Charon 3A recombinants) and plated on 1.5% T or L-agar plates which were incubated overnight at 37°C.

12. Detection and purification of recombinants:

a) Plasmid recombinants:

Detection of pBR322 plasmids carrying DNA complementary to the 12S mRNA of embryonic feathers was achieved by a modification of the procedure of Grunstein and Hogness (1975). Colonies from a transformation were transferred by toothpick to a master plate and to a sheet of nitrocellulose boiled in three changes of distilled H_2O, autoclaved for 10 min and lain onto an L-agar plate. The colonies were grown overnight on the nitrocellulose at 37°C. The colonies were lysed by transferring the nitrocellulose sequentially onto 3 MM paper saturated with 0.5 M NaOH for 7 min, 1 M Tris-HCl pH 7.4 for 2 min, 1 M Tris-HCl pH 7.4 for 2 min and 1.5 M NaCl, 0.5 M Tris-HCl pH 7.4 for 4 min. The lysed colonies were then sucked hard onto the nitrocellulose, washed with 95% ethanol and baked at 80°C in vacuo for 2 hr. Annealing conditions are described above.

b) Phage λ recombinants:

Detection of λ Charon 3A phage carrying DNA complementary to the 12S mRNA of embryonic feathers was achieved by the procedure described by Benton and Davis (1976). Untreated nitrocellulose filters were blotted directly onto phage plaques and dipped sequentially

through 1.5 M NaCl, 0.5 M NaOH for 30 seconds and 1.5 M NaCl, 0.5 M Tris-HCl pH 7.4 for 30 seconds before being dried between 3 MM paper and baked in vacuo at 80°C for 2 hr. Annealing conditions were as described above.

c) Purification of recombinants:

Phage λ recombinants were not purified or grown during the course of this work for reasons given in Chapter 4.

pBR322 plasmid recombinants were streaked out for single colonies and re-assayed by the procedure described above before being preparatively grown.

d) Miniscreen detection of plasmid recombinants:

Isolation of plasmids from small cultures for size estimations was carried out as follows. 1.5 ml cultures of the putative recombinant were grown overnight in L broth supplemented with 12.5 µg/ml tetracycline. The cells were pelleted by centrifugation for 2 min in an Eppendorf centrifuge, washed in TE pH 9.0, pelleted again and resuspended in 150 µl of 15% sucrose, 50 mM Tris-HCl pH 9.0 and 50 mM EDTA. 50 µl of a freshly prepared 4 mg/ml solution of lysozyme was added and the solution incubated for 15 min at room temperature and a further 30 min at 4°C. 150 µl of ice-cold H_2O was added, mixed and left on ice for 5 min. The solution was then transferred to 70°C for 15 min and the cell debris pelleted by centrifugation in an Eppendorf centrifuge for 15 min. The smeary pellet was removed using a hooked glass rod and the DNA precipitated by the addition of one-tenth volume of 3M

NaAcetate and 3 volumes of ethanol, chilled in dry ice-acetone and centrifuged for 5 min in an Eppendorf centrifuge. The DNA was resuspended in H_2O, ethanol precipitated again and resuspended in 100 µl of TE. A 50 µl aliquot of this was used in a 200 µl digestion mix using 1 unit of *EcoRI* and digesting at 30°C for 3 hr. The DNA was extracted with an equal volume of phenol, ethanol precipitated as described, resuspended in 20 µl of TE and electrophoresed on 1% agarose as described.

13. Electron microscopy:

a) Formation of heteroduplexes:

0.15 µg each of pBR322 and a cloned derivative were linearized by digestion with *SalI*, extracted with an equal volume of phenol and precipitated with ethanol as described. The DNA was taken up in 100 µl of 0.02 M EDTA and 0.1 N NaOH and left at room temperature for 10 min. The mix was neutralized by the addition of 10 µl of 2 M Tris-HCl pH 7.0 and formamide added to a final concentration of 50%. The DNA was allowed to anneal at room temperature for 1 hr and was stored for no more than 2 days at 4°C.

b) Visualization:

DNA in 25 µl of 65% formamide, 0.1 M Tris-HCl pH 7.5, 0.01 M EDTA and 0.1 µg/ml cytochrome C was run down a glass slide onto a hypophase of 20% formamide, 0.1 M Tris-HCl pH 7.5, 0.01 M EDTA and allowed to spread for 2 - 4 min. DNA was picked up onto a parlodion coated grid by touching the grid onto the water surface. Filmed grids had been prepared by placing 1 drop of 3.5% parlodion in iso-amyl acetate on a water surface, allowing the iso-

amyl acetate to vaporize and bringing the grids up from underneath.

The grids carrying the DNA were rinsed in 90% ethanol for 30 seconds and petroleum spirit for 30 seconds, allowed to dry and rotary shadowed with Platinum-Palladium. DNA was visualized in a Siemens Elmiskop 102 at 60 kV using 10,000 fold magnification.

14. Containment facilites:

All work involving recombinant DNA described in this thesis was carried out under C3, EK1 containment conditions for work involving viable organisms, and CO containment conditions for work not involving viable organisms, as defined and approved by the Australian Academy of Science Committee On Recombinant DNA and by the University Council of the University of Adelaide.

CHAPTER 3

TOTAL 12S cDNA AS A PROBE FOR COMPLEMENTARY SEQUENCES IN CHICK DNA

CHAPTER 3.

TOTAL 12S cDNA AS A PROBE FOR COMPLEMENTARY SEQUENCES IN CHICK DNA.

A. Introduction.

As described in Chapter 1 the embryonic feather keratin system is a multiple gene system. Protein sequence data (Walker and Rogers, 1976b) and DNA annealing kinetics (Kemp, 1975) suggested that keratins are a family of homologous, non-identical sequences that have probably evolved by gene duplication and sequence divergence. Kemp *et al.*, (1974a) showed that the keratin proteins are synthesized co-ordinately and Powell *et al.*, (1976) demonstrated that the onset of keratinization is limited by the availability of mRNA and therefore that control of gene expression is at the level of transcription of the DNA into mRNA. It was therefore of interest to study the molecular genetics of keratins to determine the nature of the arrangement of the genes coding for keratins. The absence of any observed polymorphism in the proteins of the embryonic feather as detected by gel electrophoresis (G.E. Rogers, personal communication) ruled out the analysis of these genes by genetic crosses. The approach therefore had to be at the molecular level.

The advent of a technique for the detection of sequences in restricted and fractionated DNA (Southern, 1975) made it possible to investigate the molecular genetics of keratin genes by using labelled DNA transcribed from and therefore complementary to the 12S mRNA coding for embryonic feather

keratins (12S cDNA). In particular this technique enabled the investigation of two aspects of the molecular genetics of keratins.

The first of these concerns the organization of the multiple genes coding for keratins. It is a common property of faithfully repeated multi-gene systems to be grouped into repeating units of identical or nearly identical sequence (see Chapter 1). For example the histone genes of the sea urchin and of *Drosophila* are organized into such repeat units, each unit carrying one copy of the gene for each of the five histones. The arrangement of the genes within the repeat unit is conserved, as is the sequence of the DNA between genes. Co-ordinately expressed non-identical genes, such as the globins, however, are not organized in such a rigid manner. Though the β gene and the α are at times co-ordinately expressed, the sequences adjacent to the gene appear not to be homologous. The α and β-globins are on separate chromosomes, so no requirement for linkage is necessary for co-ordinate expression. It was therefore of great interest to determine whether keratin genes are arranged in a sequence repeat unit, and if so, to determine the arrangement of genes within that repeat unit.

The second aspect concerns the stability or otherwise of the arrangement of DNA coding for keratins during development of the embryo and the developing feather.

Speculation about a role for genome rearrangement in the differentiation of somatic tissues of higher organisms was stimulated during the course of this work by the

discovery of prokaryotic insertion sequences capable of transposition within and between DNA molecules *in vivo* (Bukhari *et al.*, 1977). These elements are particularly intriguing as they transpose independently of the normal recombination pathways and may contain sequences for the termination or initiation of transcription of adjacent genetic material. Although no direct evidence exists for the presence of insertion type sequences in higher organisms, the genetic data of McClintock on the 'dissociator' element in maize is very suggestive of such behaviour (McClintock, 1956). Genetic evidence exists also for a similar element in *Drosophila melanogaster* (Greene, 1973).

The most spectacular results that promoted genome rearrangement as a potential means of differentiation and gene expression came from the immunoglobulin system. As described in Chapter 1 the constant and variable regions of immunoglobulin Lambda chain mRNA are on restriction fragments of different sizes in DNA from embryonic mouse tissue, whereas in DNA isolated from a myeloma actively synthesizing the Lambda chain mRNA, those regions are closely linked on a single restriction fragment.

Perhaps the best evidence against a role for genome rearrangement in differentiation is that provided by Gurdon, who demonstrated that nuclei from differentiated tissues of the frog *Xenopus laevis* can be transplanted to enucleated oocytes from which a mature adult can arise (Gurdon, 1964). While this functional assay clearly shows that the nuclei of these tissues have lost no information required for the growth and development of the organism, it does not strictly rule out the possibility

of DNA rearrangement, providing the process is reversible. It was of interest, then, to determine whether the genes coding for keratins also underwent rearrangement at some stage of development, particularly in view of the complexity of the keratin system and the similarity of immunoglobulin and keratin mRNA structure in that 12S mRNA contains covalently linked unique and repeated sequences (Lockett, Kemp and Rogers, in preparation and see Chapter 1).

These aspects were investigated using restriction endonuclease cleavage of genomic chicken DNA, fractionation on 1% agarose and immobilization onto nitrocellulose filter paper as described in Chapter 2. Sequences complementary to the 12S mRNA were detected by annealing 12S cDNA as described in Chapter 2.

B. Results.

1. No sequence repeat unit is apparent for the genes coding for 12S mRNA:

When 12S cDNA was annealed to restricted and fractionated genomic chicken DNA as described in Chapter 2, complex patterns of bands resulted on autoradiography. Fig.3.1 shows the patterns obtained when chick DNA was digested with *Bam*HI, *Bgl*II, *Eco*RI and *Hin*dIII. Each pattern consisted of several intense major bands and many other less intense bands.

This complex pattern is clearly different to the simple patterns obtained with faithfully repeated gene families. For example, *Eco*RI digested chick DNA fragments annealed to labelled DNA complementary to erythroblast 18S and 28S ribosomal RNA yielded the two bands seen in Fig.3.2.

FIGURE 3.1. The Patterns of Restricted and Fractionated Chick DNA Fragments annealing to 12S cDNA.

Track A: 30 µg of chick DNA was digested to completion with *Eco*RI, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light.

Tracks B to E: 30 µg of chick DNA was digested to completion with B: *Bam*HI

C: *Eco*RI

D: *Bgl*II

E: *Hin*dIII

electrophoresed in 1% agarose denatured *in situ*, transferred and immobilized onto nitrocellulose filter paper, annealed to ^{32}P-labelled 12S cDNA in 2 x SSCE at 65°C, washed in 1 x SSCE at 65°C, dried and autoradiographed.

Molecular weight estimates were obtained by transferring λ-*Eco*RI and λ-*Hin*dIII markers in adjacent Tracks (Fig.2.2).

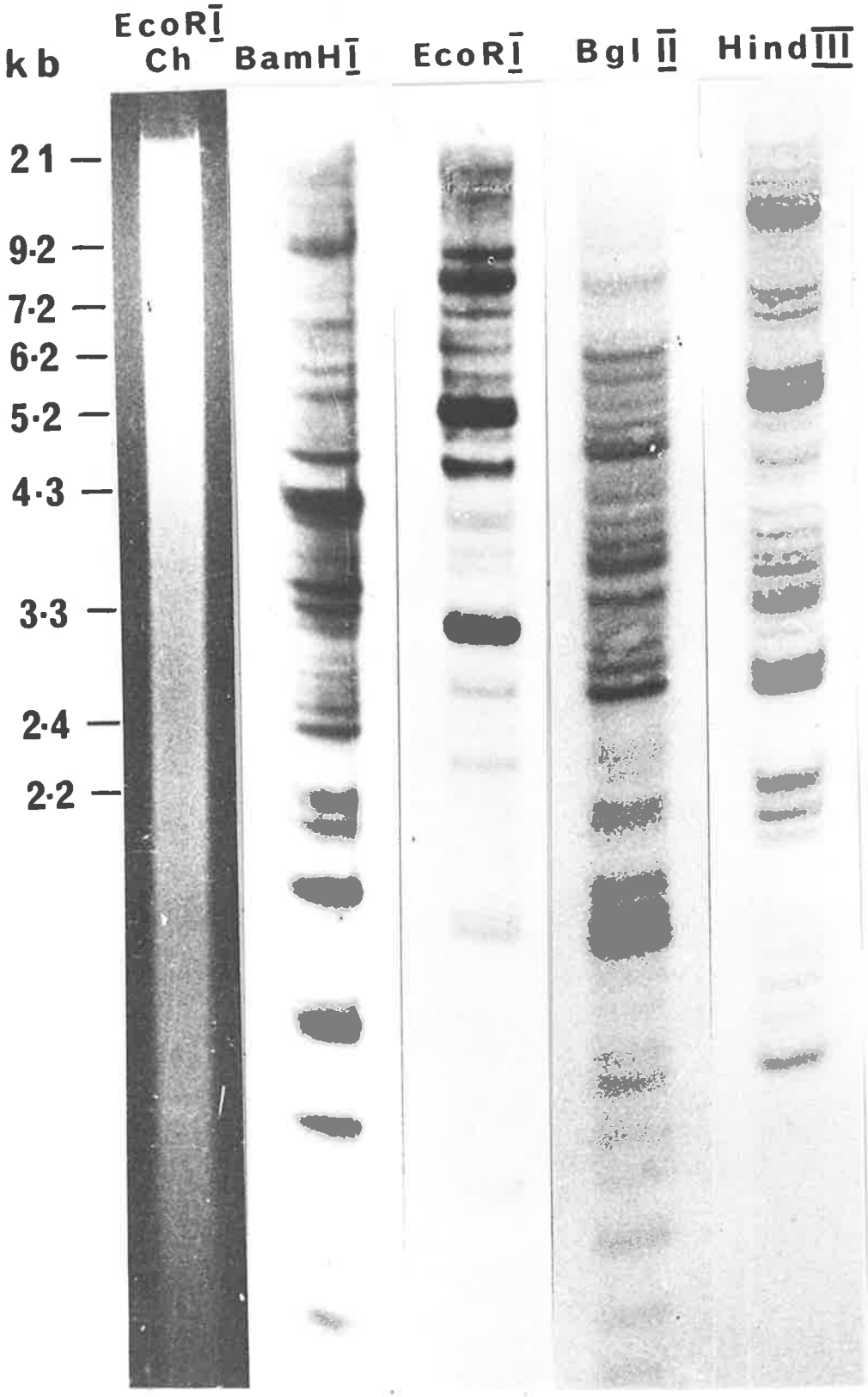

kb
EcoRI
Ch
BamHI
EcoRI
Bgl II
HindIII
21
9·2
7·2
6·2
5·2
4·3
3·3
2·4
2·2

FIGURE 3.2. The Patterns of Restricted and Fractionated Chick DNA Fragments Annealing to 12S cDNA or 18S and 28S Ribosomal cDNA.

Chick DNA was restricted with either *Eco*RI or *Bam*HI as indicated, fractionated in 1% agarose, denatured *in situ* transferred and immobilized onto nitrocellulose filter paper, annealed to ^{32}P-labelled 12S cDNA or a mixture of 18S and 28S ribosomal cDNA as indicated, washed, dried and autoradiographed. The range of fragment sizes fractionated by this procedure is approximately from 0.8 kb to 20 kb (see Fig.3.1).

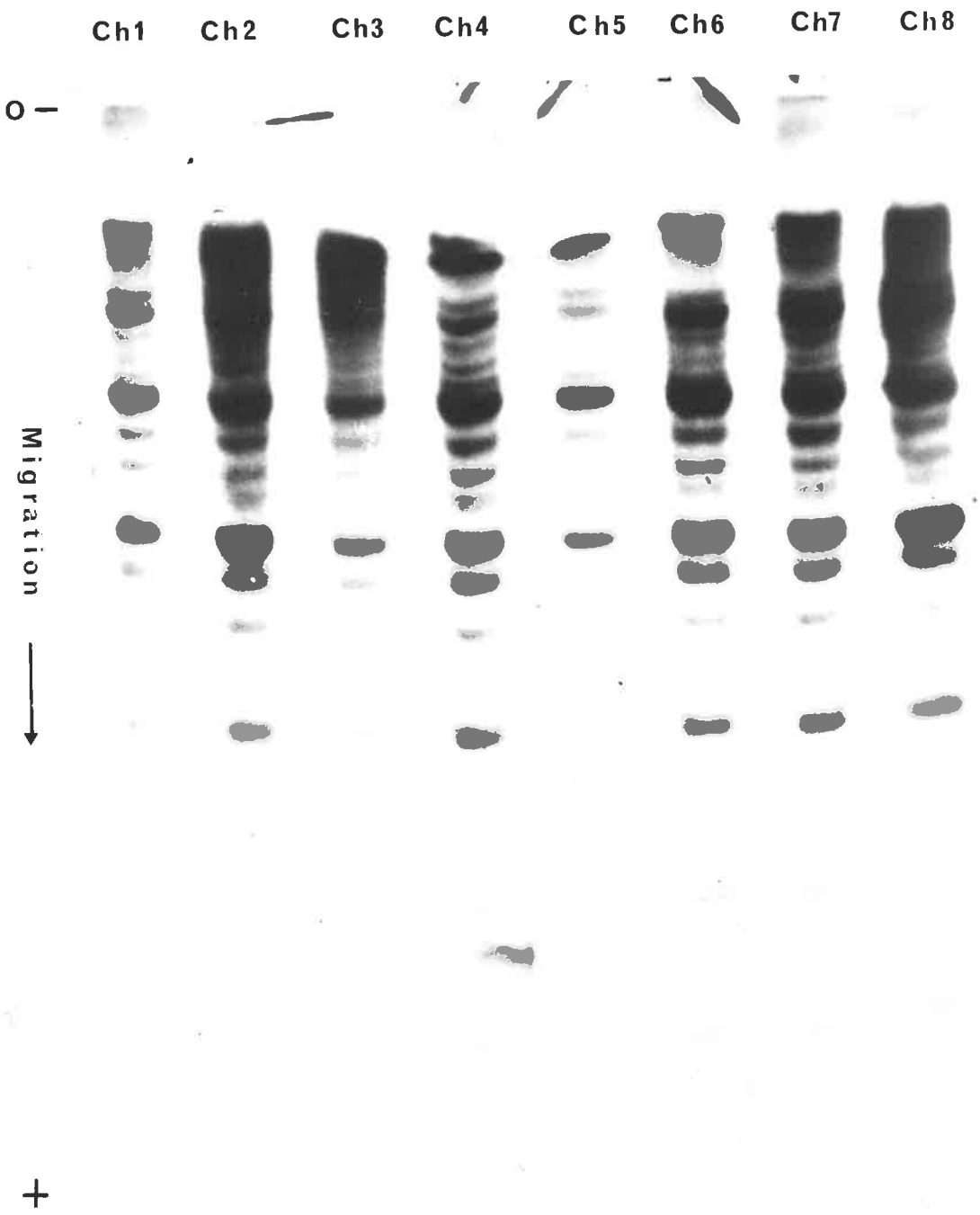
O –
Eco RI kcDNA
Eco RI Rib cDNA
BamHI kcDNA
BamHI Rib. cDNA
Migration
+

The observed complexity could have been due to one of four reasons. The simplest of these is that the mRNA coding for keratins was contaminated with non-keratin coding species. Results obtained with pure 12S mRNA species show that although there was a level of contamination, these contaminants did not contribute to the observed patterns (see the following Chapter). A second possibility is that chicken populations are genetically polymorphic at their keratin loci, i.e., that there are several alleles at a given locus that vary in the arrangement of restriction cleavage sites. This, however, was not a contributing factor to the observed complexity, since the DNA used for this study was always isolated from an individual, so that at most two alleles could have been present. In addition several preparations of DNA from different individuals, including those of different strains, were analysed and shown to be identical with respect to the fragments generated by *Eco*RI cleavage (Fig.3.3).

A third possibility is that the complexity of the pattern resulted from cross-annealing to sequence-related keratins not expressed in the embryonic feather, e.g., those expressed in the adult feather, beak, scale or claw. Walker and Bridgen (1976) showed by protein sequencing that there was considerable homology between certain peptides of scale keratins and embryonic feather keratins. Kemp (1975) showed that 12S cDNA annealed to chicken DNA in two phases. He concluded that the first duplexes formed were representative of repeated sequences in the cDNA. These sequences were mismatched and melted with a

FIGURE 3.3. Polymorphism Analysis of DNA Annealing to 12S cDNA.

Chick DNA isolated from the erythrocytes of several different individuals of different chick strains were restricted with *Eco*RI, fractionated in 1% agarose, denatured *in situ*, transferred and immobilized onto nitrocellulose filter paper, annealed to ^{32}P-labelled 12S cDNA, washed, dried and autoradiographed. The Track denoted 'Ch5' is less exposed because a smaller amount of restricted DNA was immobilized onto the filter, presumably because the ethanol precipitated DNA was not completely redissolved before electrophoresis. It should also be noted that marked differences were observed in the patterns of fragments annealing to 18S and 28S ribosomal cDNA for these different individuals (results not shown) so that the lack of polymorphism observed for these individuals with respect to the 12S cDNA sequences was not due simply to an overall lack of variation in genomes of domestic chickens.


Ch1
Ch2
Ch3
Ch4
Ch5
Ch6
Ch7
Ch8
O –
Migration
+

Tm of 73°C compared with a Tm of 93°C for the duplexes formed at high Cot values, i.e., for perfect duplexes. The mis-matched duplexes could represent cross-annealing between different embryonic keratin species as seen for 12S mRNA - 12S cDNA hybrids, but as well, it could represent cross-annealing of the 12S cDNA to related sequences coding for keratins not expressed in the feather, for example, beak, scale and claw keratins. It was possible to discriminate between these possibilities on the basis that the mis-matched duplexes described by Kemp (1975) should preferentially melt from the filter at elevated temperatures. If these mis-matched duplexes are the result of cross-annealing of embryonic feather sequences, then melting should not change the pattern of fragments, since a proportion of the duplexes in all the bands will be perfectly annealed. If, however, a portion of the mis-matched duplexes result from duplexes formed between the embryonic feather cDNA and keratin genes of the beak, scale, claw and adult feather, and these keratin genes are on different restriction fragments to those of the embryonic feather, then certain bands should preferentially disappear on incubation at temperatures above the Tm of these mis-matched duplexes, *viz.* 73°C.

Figure 3.4 shows a series of annealed transfers which were post-melted in SSC at temperatures of 65°C at 90°C at 5°C intervals for 30 min per filter before going through their normal washing procedure (see Chapter 2). It can be seen that no bands were selectively lost, even at 90°C. This result was supported by a similar experiment described in Chapter 4 using decreased salt concentrations rather

FIGURE 3.4. Post-melts of Chick DNA-12S cDNA Duplexes.

*Eco*RI digested chick DNA was electrophoresed in 1% agarose, denatured *in situ*, transferred and immobilized onto nitrocellulose filter paper and annealed to ^{32}P-labelled 12S cDNA in 2 x SSCE at 65°C. Each filter was post-washed in 1 x SSCE for 30 min at the temperature indicated above the Track, before being returned to 65°C, washed overnight, dried and autoradiographed. Fragment sizes fractionated by this procedure are as indicated in Fig.3.1.

O —
Eco RI
65°
70°
75°
80°
85°
90°
Migration
+

than increased temperature to increase the stringency of the washing conditions.

There are three possible explanations of this result. The first is that the mis-matched duplexes observed by Kemp (1975) were due to the annealing of 12S cDNA species to homologous, but non-identical chick sequences coding for other species expressed in the embryonic feather but not to those expressed in the beak, scale or claw. The second is that the genes coding for beak, scale or claw keratins reside on the same restriction fragments as the embryonic feather keratin genes and the third is that the beak, scale etc. proteins are expressed at a low level in the embryonic feather.

Whatever the interpretation, this result shows that cross-annealing, as described above, does not contribute to the observed complexity and that the fragments annealing to the 12S cDNA are transcribed to give the 12S mRNA during embryonic feather development.

Having thus discounted contamination, genetic polymorphism and cross-annealing to sequences other than those expressed in the embryonic feather, the complexity of the restricted chick DNA fragments annealing to 12S cDNA must be the result of sequence heterogeneity in the DNA coding for keratins. Studies of keratin protein sequences (Walker and Rogers, 1976b) and of the melting behaviour of 12S cDNA hybridized to 12S mRNA (Kemp, 1975), demonstrated that sequence heterogeneity exists both between species of keratin proteins and between the 12S mRNA species that code for them. The many restriction generated fragments

annealing to 12S cDNA further imply that sequences adjacent to the coding region (either flanking or intervening sequences) must vary for different keratin genes. The major conclusion from these results is therefore that the genes coding for embryonic feather keratins cannot be organised into discrete sequence repeat units.

2. Keratin gene expression does not involve DNA rearrangement:

The complexity of the keratin-coding DNA patterns, although a hindrance in the elucidation of keratin gene arrangement, provided an excellent system to assay for a keratin gene rearrangement which may occur during development of the organism and of the feather tissue itself. If the DNA coding for keratins is involved in any addition, deletion, translocation or transposition there is very little possibility of the same pattern re-emerging by chance.

DNA was isolated from chicken sperm, 13 day embryonic feather and adult blood tissue as described in Chapter 2. Restriction fragments were generated by either *Eco*RI, *Bgl*II or *Bam*HI and the pattern of DNA fragments annealing to 12S cDNA determined as described in Chapter 2. Figures 3.5, 3.6 and 3.7 show these patterns for *Eco*RI, *Bgl*II and *Bam*HI respectively. In no case were new fragments generated nor existing fragments lost during development, as detected by digestion with these three endonucleases. In addition 1:1 mixtures of feather DNA and sperm or blood DNA were digested with *Eco*RI, *Bgl*II and *Bam*HI, showing that the patterns of each type of DNA were truly identical, as no

FIGURE 3.5. Developmental analysis of *Eco*RI-digested Chick DNA Fragments Annealing to 12S cDNA.

Chick DNA isolated from different tissues was digested to completion with *Eco*RI, electrophoresed in 1% agarose, denatured *in situ*, transferred and immobilized onto nitrocellulose filter paper, annealed to ^{32}P-labelled 12S cDNA in 2 xSSCE at 65°C, washed in 2 xSSCE at 65°C, dried and autoradiographed. Tracks are as follows:

Sp: DNA isolated from chick sperm

F: DNA isolated from feathers of a 13-day old embryo.

Bl: DNA isolated from adult blood.

F/Sp: An equal mixture of feather and sperm DNA

F/Bl: An equal mixture of feather and blood DNA.

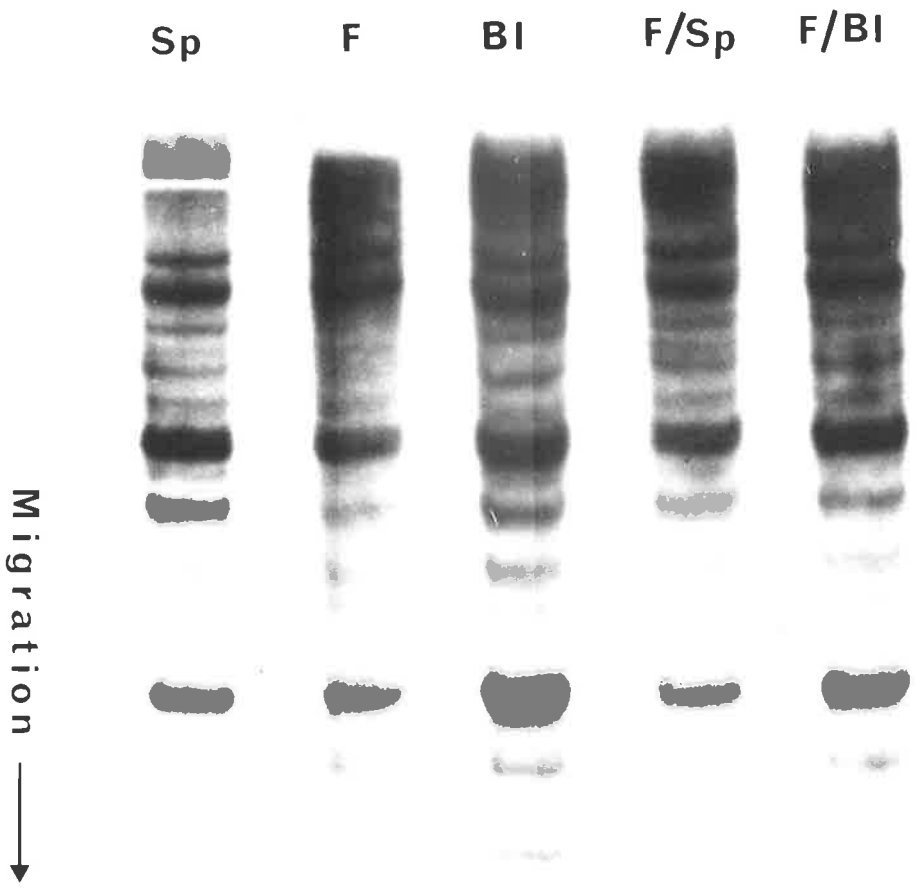

Sp
F
Bl
F/Sp
F/Bl
Migration

FIGURE 3.6. Developmental Analysis of *Bgl*II-digested Chick DNA Fragments Annealing to 12S cDNA.

Chick DNA isolated from different tissues was digested to completion with *Bgl*II, fractionated in 1% agarose, immobilized onto nitrocellulose filter paper, annealed to ^{32}P-labelled 12S cDNA, washed, dried and autoradiographed. Details of this figure are as described for Fig.3.5.

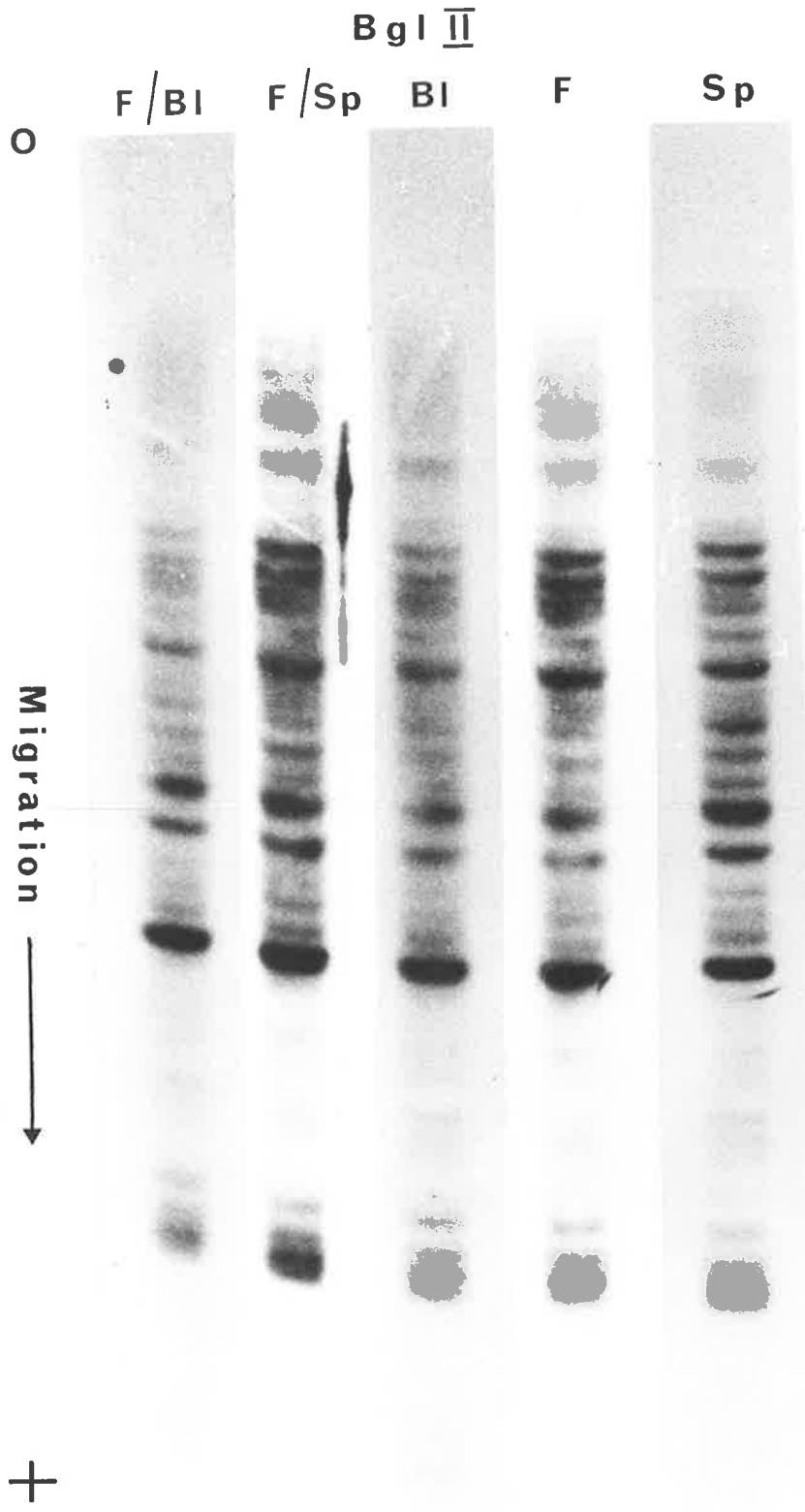
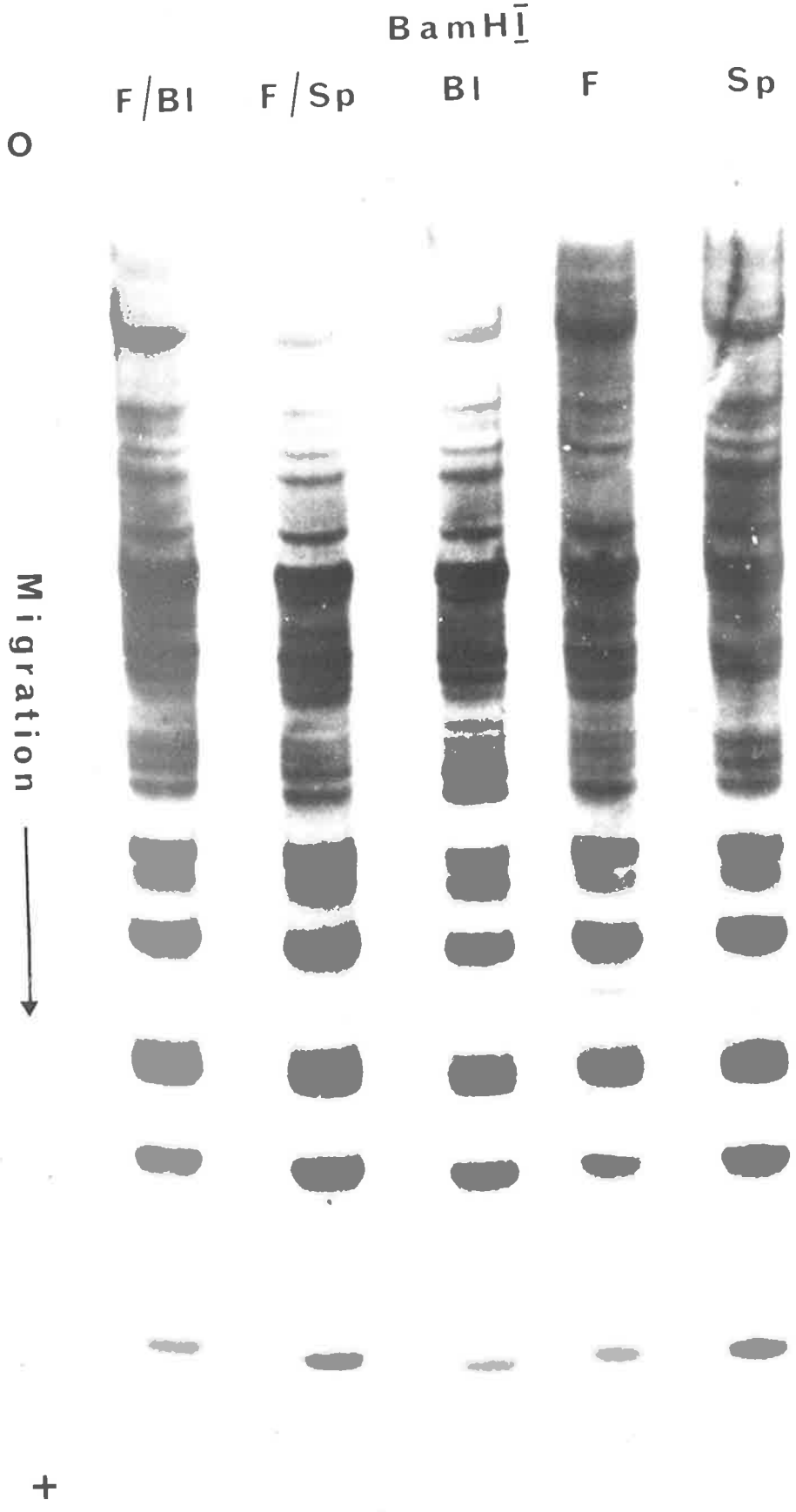

Bgl II
F/Bl
F/Sp
Bl
F
Sp
O
Migration
+

FIGURE 3.7. Developmental Analysis of *Bam*HI-digested Chick DNA Fragments Annealing to 12S cDNA.

Chick DNA isolated from different tissues was digested to completion with *Bam*HI, fractionated in 1% agarose, immobilized onto nitrocellulose filter paper, annealed to ^{32}P-labelled 12S cDNA, washed, dried and autoradiographed. Details of this figure are as described for Fig.3.5.


BamHI
F/Bl
F/Sp
Bl
F
Sp
O
Migration
+

new fragments were apparent in any of these tracks (Figures 3.5, 3.6 and 3.7).

The DNA isolated from 13 day embryonic feathers was true feather DNA and not contaminated to any significant extent with DNA from embryonic blood in the feather vein, as indicated by the relative absence of the erythrocyte specific histone H5 in nuclei of the feather tissue (see Chapter 2 and Fig.2.1). This tissue has largely ceased cell division by day 13 (Watterson, 1942; Matulionis, 1970) so that any rearrangement that occurred before or during the final cell division would have been obvious in the experiments described above. It is concluded, therefore, that keratin gene expression is independent of genome rearrangement.

C. Discussion.

This Chapter described the use of cDNA to total 12S mRNA coding for keratins as a probe to investigate the molecular genetics of chicken DNA coding for keratins. Two major conclusions resulted from this work.

The first of these is that DNA coding for keratins is heterogeneous in the sequence of both the coding and non-coding regions. The kinetics of 12S mRNA-cDNA hybridizations described by Kemp (1975) show that the coding region of DNA coding for keratins (i.e., the 12S mRNA sequences) are heterogeneous. However the large number of restricted chick DNA fragments annealing to 12S cDNA cannot be explained simply by coding sequence variability alone. For example even if the postulated 25 species (Kemp, 1975) each 700 bases

in length were entirely variable, then a restriction endonuclease with a six nucleotide specificity will cleave 4 - 5 times on a random basis (since we are assuming a total of 18 kb of heterogeneous coding region and since a 6 base specific restriction enzyme cleaves every 4^6 or approximately 4000 nucleotides). However as there appear to be of the order of 20 to 30 fragments involved, each annealing to the 12S cDNA, it must be concluded that either an organizational repeat unit, if it exists, is highly polymorphic in length, or heterogeneous in sequence. Either of these conclusions infers that there is considerable sequence heterogeneity both within the coding region and in adjacent sequences (either flanking or intervening).

Although the simplest interpretation of the data is that there exists no simple sequence repeat unit associated with DNA coding for keratins, this does not imply that keratin genes are not clustered, nor that there is no regular pattern of gene organization. In fact, it is precisely because of this sequence heterogeneity, that it is impossible to draw any conclusions about linkage relationships between the keratin genes.

The second conclusion from this study is that genome rearrangement is not involved in the expression of keratin genes. Expression must therefore involve the activation of pre-existing control regions rather than the insertion of efficient promoters adjacent to the gene. It also suggests that the DNA rearrangement observed for immunoglobulins must have evolved specifically as a means of generating the remarkable variability required by the immune system.

Given these two conclusions it becomes obvious that any major advance in our understanding of the molecular genetics of the keratin genes awaits the availability of isolated segments of chicken chromosomal DNA carrying the natural genes.

CHAPTER 4

PURE 12S cDNA SPECIES AS PROBES FOR COMPLEMENTARY SEQUENCES IN CHICK DNA

CHAPTER 4.

PURE 12S cDNA SPECIES AS PROBES FOR COMPLEMENTARY SEQUENCES IN CHICK DNA.

A. Introduction.

1. Purification of individual mRNA species:

Chapter 3 described the use of DNA complementary to total 12S mRNA coding for the family of embryonic feather keratins as a probe to detect DNA coding for keratins in the chicken genome. This work demonstrated that no apparent sequence repeat unit was associated with DNA coding for keratins, since 12S cDNA annealed to many fragments of different sizes. This complexity severely limited the amount of information that could be obtained using total 12S mRNA sequences as a probe. It was therefore necessary to isolate individual species from the mixed population of 12S mRNA and use them to investigate the properties of 12S mRNA species and of the genomic DNA coding for them.

Isolation and preparation of pure mRNA species from a mixture of species can be achieved by the use of the recently developed recombinant DNA techniques.

All approaches to the molecular cloning of mRNA species have used a three step procedure that starts with the mRNA and produces double-stranded DNA molecules whose sequence derives directly from the mRNA (dscDNA). Each dscDNA molecule is then joined to a vector DNA molecule that has the capacity to replicate autonomously in the cell, e.g., plasmid or bacteriophage DNA. These molecules are then

used to transform or transfect competent *E. coli* cells which are plated out for colonies or plaques. Since each colony or plaque derives from an initial single cell and conditions are used such that each cell carries only one recombinant molecule, each colony or plaque carries copies of the one recombinant molecule and thus a single cDNA species.

The first step in the preparation of the dscDNA involves the synthesis of cDNA on the mRNA template using Avian Myeloblastosis Virus RNA-dependent DNA polymerase (AMV reverse transcriptase). The second step involves extending the cDNA back along the molecule to form the dscDNA, using either *E. coli* DNA polymerase I (Maniatis *et al.*, 1976) or AMV reverse transcriptase which can also act as a DNA-dependent DNA polymerase (Rougeon and Mach, 1976). Both enzymes were investigated for use in the molecular cloning of 12S mRNA sequences. The two strands of this dscDNA molecule are covalently linked at one end, as the cDNA folds back to form its own primer. The final step, therefore, involves the cleavage of the end-loop, as well as of any other single-stranded portion of the molecule, with the single-strand specific nuclease S_1. The resulting double-stranded molecule has free 3'OH and 5'P residues at both ends.

In vitro joining of the dscDNA to a vector DNA molecule (phage or plasmid) has been achieved in two ways. The first was by the addition of homopolymeric 'tails' to the 3'OH residue of the DNA using terminal deoxynucleotidyl transferase (Jackson *et al.*, 1972; Lobban and Kaiser, 1973). When complementary base-pairing nucleotides are added to

donor and vector DNA respectively, the 'tails' anneal to form a circular molecule that will transform or transfect competent *E. coli* cells (see Fig.4.1). Rougeon *et al.*, (1975), Maniatis *et al.*, (1976) and Rabbitts (1976) used this joining procedure to clone rabbit β-globin sequences.

An alternative approach is to use the blunt end ligation activity of the enzyme T_4 polynucleotide ligase to join, to the DNA molecule, short polynucleotide segments ('linkers') carrying the recognition sequence of a restriction endonuclease (Fig.4.2). Ligation can then be effected using T_4 ligase to join the complementary ends generated by restriction cleavage of both vector and linkered donor DNA (Ullrich *et al.*, 1977).

It was decided to use the 'tailing' approach to clone the 12S mRNA species, principally because the 'linker' procedure involves a restriction cleavage of the dscDNA. As it has been estimated that there are as many as 20 - 30 species of mRNA in the 12S mRNA (Kemp, 1975), it is very likely that a given endonuclease will cleave some of these species. Such species would be rendered shorter by the cleavage and possibly disregarded in favour of longer molecules. This problem can be avoided by the use of linkers encoding the *Eco*RI cleavage site and by pre-methylating the DNA with *Eco*RI methylase to protect against internal cleavage of the donor DNA (Maniatis *et al.*, 1978). However, as *Eco*RI methylase was not available at the time and terminal deoxynucleotidyl transferase was, it was decided to use the 'tailing' procedure.

Having decided on the approach to the synthesis of the

FIGURE 4.1. Procedure for Cloning mRNA Sequences Using The 'Tailing' Procedure.

Diagrammatic representation of the steps involved in cloning the 12S mRNA sequences using the 'tailing' procedure described in Chapter 4.

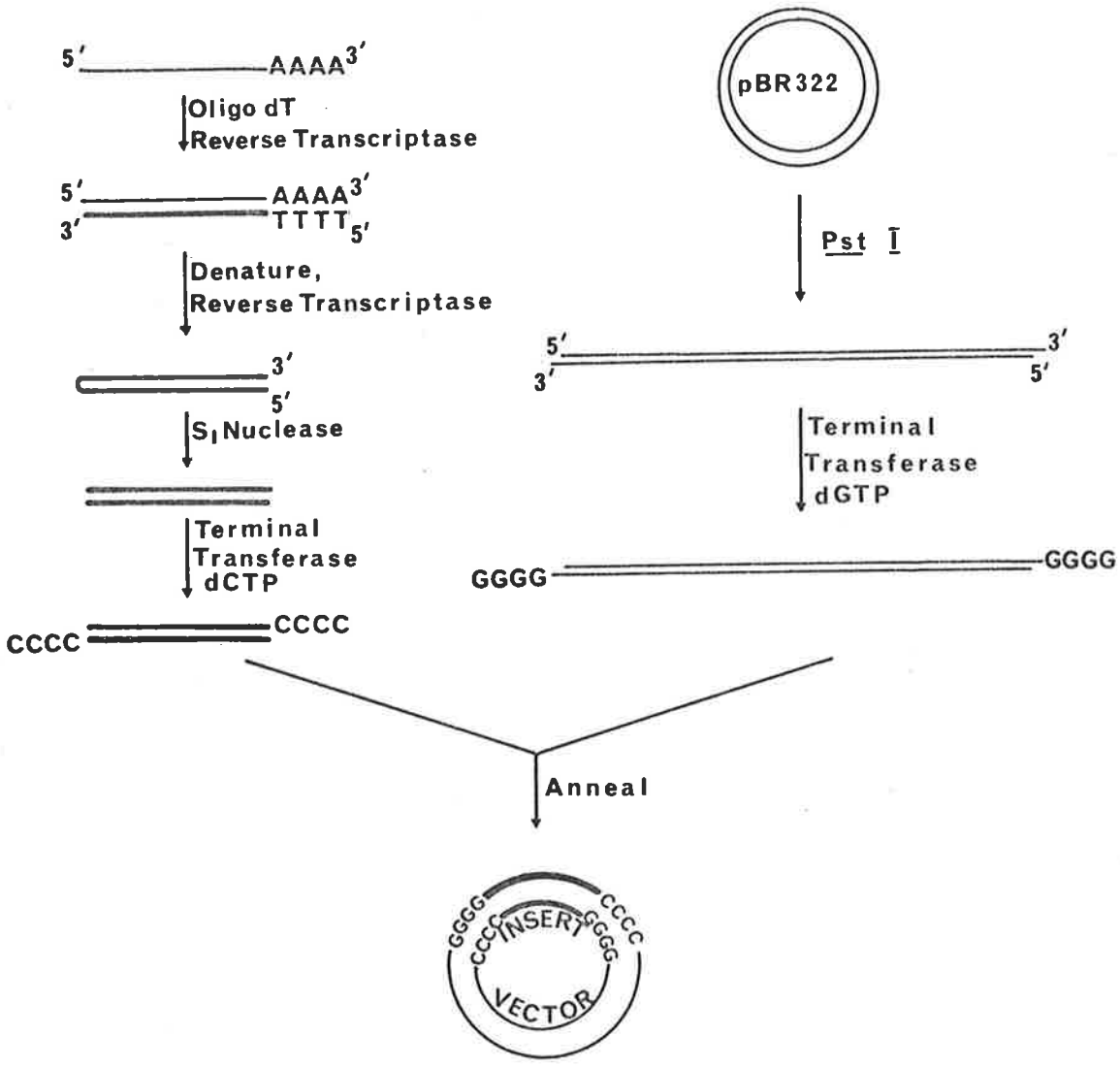
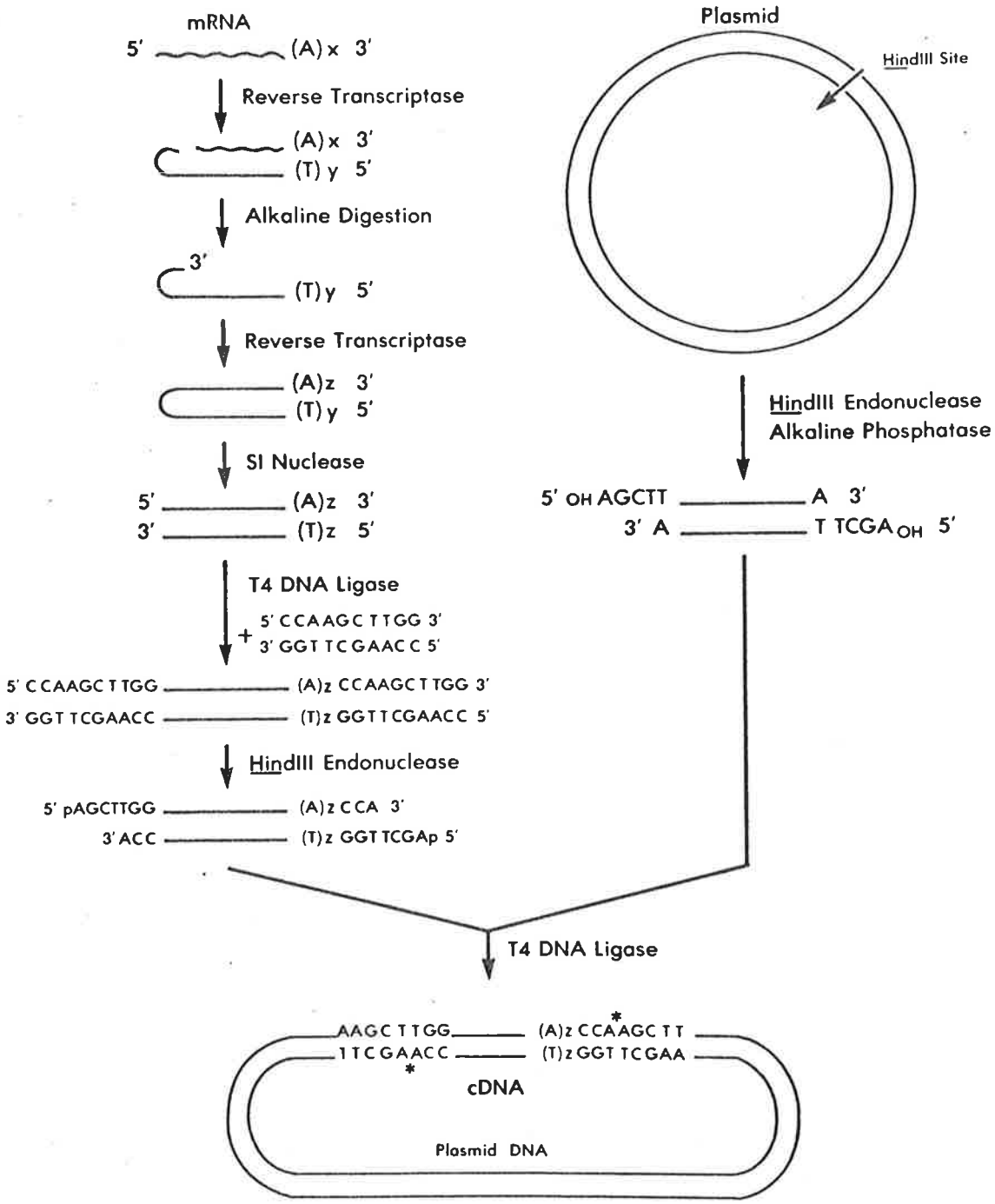

5′ AAAA 3′
Oligo dT
Reverse Transcriptase
5′ AAAA 3′
3′ TTTT 5′
Denature,
Reverse Transcriptase
3′
5′
S1 Nuclease
Terminal
Transferase
dCTP
CCCC
CCCC
pBR322
Pst I
5′
3′
3′
5′
Terminal
Transferase
dGTP
GGGG
GGGG
Anneal
GGGG
CCCC
INSERT
CCCC
GGGG
VECTOR

FIGURE 4.2. Procedure for Cloning mRNA Sequences Using The 'Linker' Procedure.

Diagrammatic representation of the steps involved in cloning mRNA species using the 'linker' procedure described in Chapter 4.

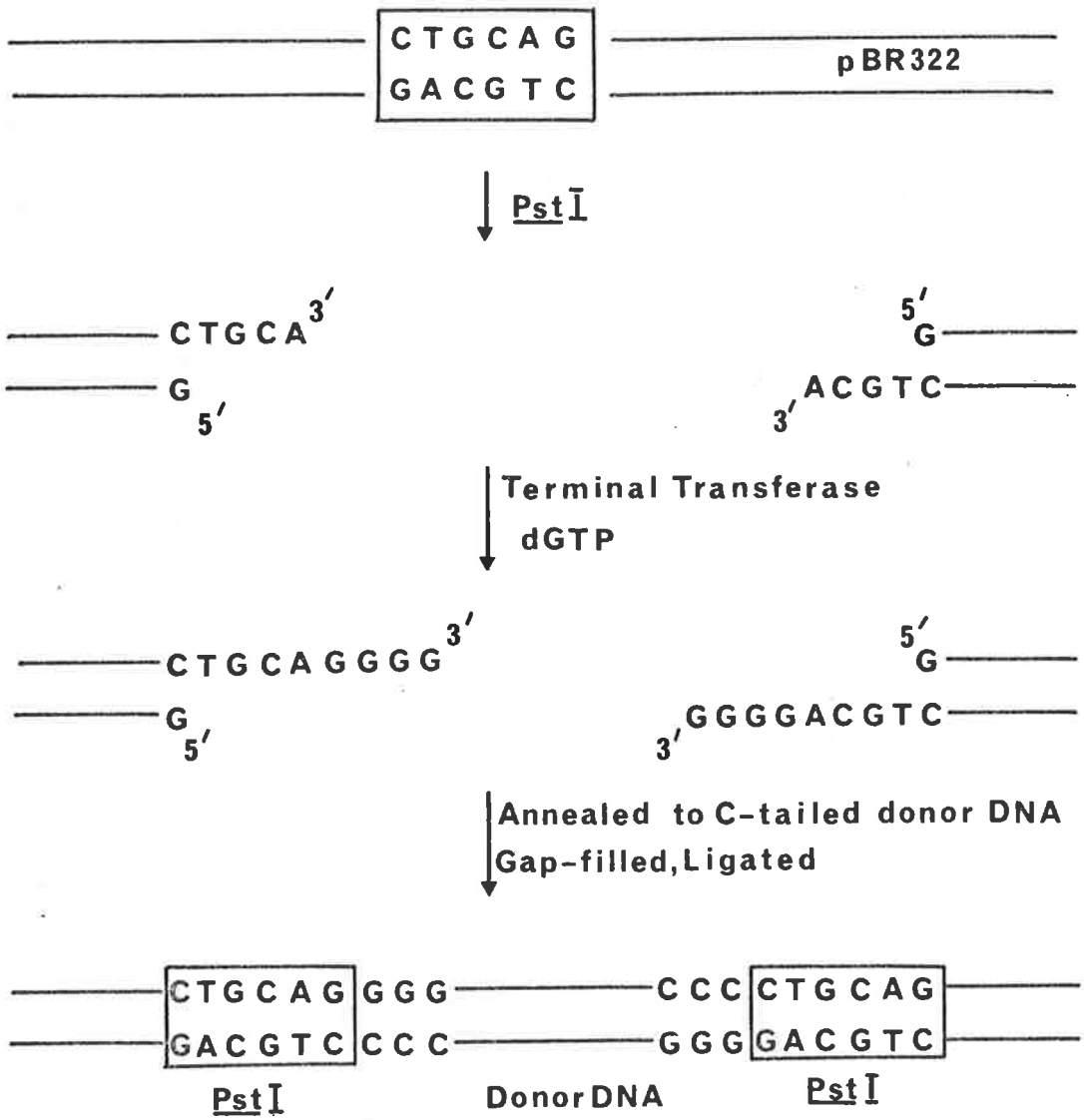

mRNA
5' (A)x 3'
Reverse Transcriptase
(A)x 3'
(T)y 5'
Alkaline Digestion
3'
(T)y 5'
Reverse Transcriptase
(A)z 3'
(T)y 5'
SI Nuclease
5' (A)z 3'
3' (T)z 5'
T4 DNA Ligase
+ 5' CCAAGCTTGG 3'
3' GGTTCGAACC 5'
5' CCAAGCTTGG (A)z CCAAGCTTGG 3'
3' GGTTCGAACC (T)z GGTTCGAACC 5'
HindIII Endonuclease
5' pAGCTTGG (A)z CCA 3'
3' ACC (T)z GGTTCGAp 5'
Plasmid
HindIII Site
HindIII Endonuclease
Alkaline Phosphatase
5' OH AGCTT A 3'
3' A T TCGA OH 5'
T4 DNA Ligase
AAGCTTGG (A)z CCAAGCTT
TTCGAACC (T)z GGTTCGAA
cDNA
Plasmid DNA

dscDNA and the method of joining to the vector, it only remained to decide which vector or vectors to use. As discussed in Chapter 1, one of the most interesting features of the 12S cDNA species coding for embryonic feather keratins is their ability to form mis-matched hybrids with 12S mRNA or chick DNA in solution hybridization (Kemp, 1975). To investigate this property thoroughly, it will be necessary to carry out cross-annealing studies between different species. Such studies would be greatly facilitated by the availability of 12S cDNA species cloned into vectors that are unrelated in sequence. For this reason the plasmid vector pBR322 (Bolivar *et al.*, 1977) and the phage vector λ Charon 3A (Blattner *et al.*, 1977) were used for the cloning. Plasmid pBR322 was cleaved at the single *Pst*I site and joined via G-C tails to the cDNA, while λ Charon 3A, with its cohesive termini ligated, was cleaved at the unique *Eco*RI site and joined via A-T tails to the cDNA. It should be noted that tailing the *Pst*I site of pBR322 with deoxyguanosine and annealing to DNA tailed with deoxycytidine regenerates the *Pst*I site, as shown in Fig.4.3. The inserts cloned into λ Charon 3A by A-T tails can also be resected by a method described by Goff and Berg (1978). This method uses the fact that denatured strands of the recombinant molecules will snap back to form a base-paired region between the A and T residues on either side of the inserted DNA. This leaves the λ arms exposed to cleavage with *E. coli* exonuclease VII. The insert is protected from cleavage by the A-T duplex at its ends.

Detection of recombinant molecules has been greatly

FIGURE 4.3. *Pst*I Cleavage and G Tailing Regenerates the *Pst*I Sites.

Diagrammatic representation of the regeneration of *Pst*I sites in *Pst*I cleaved pBR322 DNA by G tailing, annealing to a C tailed donor DNA molecule and transforming into a host bacterium.


CTGCAG
GACGTC
pBR322
PstI
CTGCA 3′
G 5′
5′ G
3′ ACGTC
Terminal Transferase
dGTP
CTGCAGGGG 3′
G 5′
5′ G
3′ GGGGACGTC
Annealed to C-tailed donor DNA
Gap-filled, Ligated
CTGCAG GGG CCC CTGCAG
GACGTC CCC GGG GACGTC
PstI
Donor DNA
PstI

facilitated by the development of two filter hybridization techniques. Plasmid recombinants are detected by growing colonies directly onto nitrocellulose filters lain onto nutrient plates, lysing the cells *in situ* and binding the released denatured DNA to the filter (Grunstein and Hogness, 1975). Detection of phage recombinants is achieved by laying nitrocellulose filter onto plaques grown in a bacterial lawn and denaturing and baking the DNA onto the filter (Benton and Davis, 1976). Recombinant colonies or plaques are then detected by probing the filters with radioactive sequences complementary to the required recombinant, in this case 12S cDNA.

2. The application of pure 12S mRNA species to the study of the molecular genetics of keratins:

The availability of pure species derived from 12S mRNA enables several facets of the molecular genetics of keratins to be examined. Perhaps the greatest difficulty in the initial interpretation of the results described in Chapter 3 is the fear that the 12S mRNA used was contaminated with non-keratin mRNA species. Molecular cloning ensures that the sequence isolated cannot be contaminated by other such species. Pure 12S mRNA species can also be used to test the properties of the total 12S mRNA or 12S cDNA described by both Kemp (1975) and in Chapter 3 of this thesis. In particular, the postulate that 12S mRNA is comprised of a family of homologous species can be tested once pure species have been isolated. The number of different species comprising 12S mRNA can also be determined with much greater accuracy.

This Chapter describes the molecular cloning of the embryonic feather mRNA sequences and the preliminary analysis of these sequences using restriction cleavage and filter annealing techniques.

B. Results.

1. Development of a rapid technique for the molecular cloning of embryonic feather 12S mRNA:

a) Synthesis of the first strand:

Synthesis of cDNA from the 12S mRNA template was carried out as described in Chapter 2 using AMV reverse transcriptase. For preparative synthesis, 10 µg of mRNA was used as template, yielding approximately 2 µg of cDNA. The cDNA was heterogeneous in length with a portion of the cDNA being full length copy (Fig.4.5).

b) Synthesis of the second strand:

Both *E. coli* DNA polymerase I and AMV reverse transcriptase were used to extend the cDNA into double-stranded form. In both cases the synthesis was carried out in the same tube as the first strand, following heating in a boiling water bath for 2 min to separate the cDNA from the mRNA template. The conditions used for the DNA polymerase I extension were those described by Kemp (personal communication) as follows: 20 µl of the cDNA mix described in Chapter 2 was diluted with 20 µl of a solution of 0.33 M Hepes pH 6.9, 0.66 M KCl and 300 mM dATP, dCTP, dGTP and dTTP. 2 µl of *E. coli* DNA polymerase I were added and the solution incubated at 15°C for 3 hr. The dscDNA thus formed was diluted into 200 µl of the S_1 buffer described in Chapter 2 and cleaved with S_1 nuclease

as described. Fig.4.4 shows the time course of incorporation of ^{32}P-dATP into the second strand, the first strand being labelled with ^{3}H-dCTP. It shows that nucleotides were being incorporated into the second strand, but that in addition, nucleotides were being lost from the first strand. This presumably resulted from the nuclease activities associated with the DNA polymerase I.

Fig.4.5 shows a denaturing 98% formamide gel of samples taken at various stages of this procedure. It shows that the molecular weight of the cDNA was increased after synthesis of the second strand and that S_1 cleavage did indeed reduce the molecular weight of the dscDNA to less than that of the cDNA. The size of the post-S_1 treated dscDNA was, however, relatively short, suggesting again that nuclease activity in the DNA polymerase I had shortened the dscDNA. A better approach would certainly have been to use the nuclease deficient 'Klenow' fragment of DNA polymerase I (Jacobsen *et al.*, 1974).

The alternative method was to use AMV reverse transcriptase to synthesize the second strand. For ease of manipulation it was decided to attempt a single incubation approach using this enzyme. The procedure used is detailed in Chapter 2. This method proved to be more efficient in the synthesis of the second strand of the dscDNA.

To test this method, 1 µg of 12S mRNA was used as template for the cDNA synthesis as described in Chapter 2. An aliquot was removed and assayed for cDNA synthesis, revealing that 18% copy had occurred, i.e., 0.18 µg of cDNA had been synthesized, as determined by the number of ^{32}P-nucleotides of known specific activity incorporated.

FIGURE 4.4. Synthesis of dscDNA with DNA Polymerase I.

DNA polymerase I was used to copy ^{3}H-labelled cDNA into dscDNA as described in Chapter 4. ^{32}P-labelled dATP was included to follow incorporation of nucleotides into the second strand. The amount of ^{3}H-labelled first strand was also followed. Radioactivity was determined by TCA precipitation as described in Chapter 2.

- O ^{32}P-nucleotides incorporated into an acid insoluble form.
- ■ ^{3}H-nucleotides remaining in an acid insoluble form.

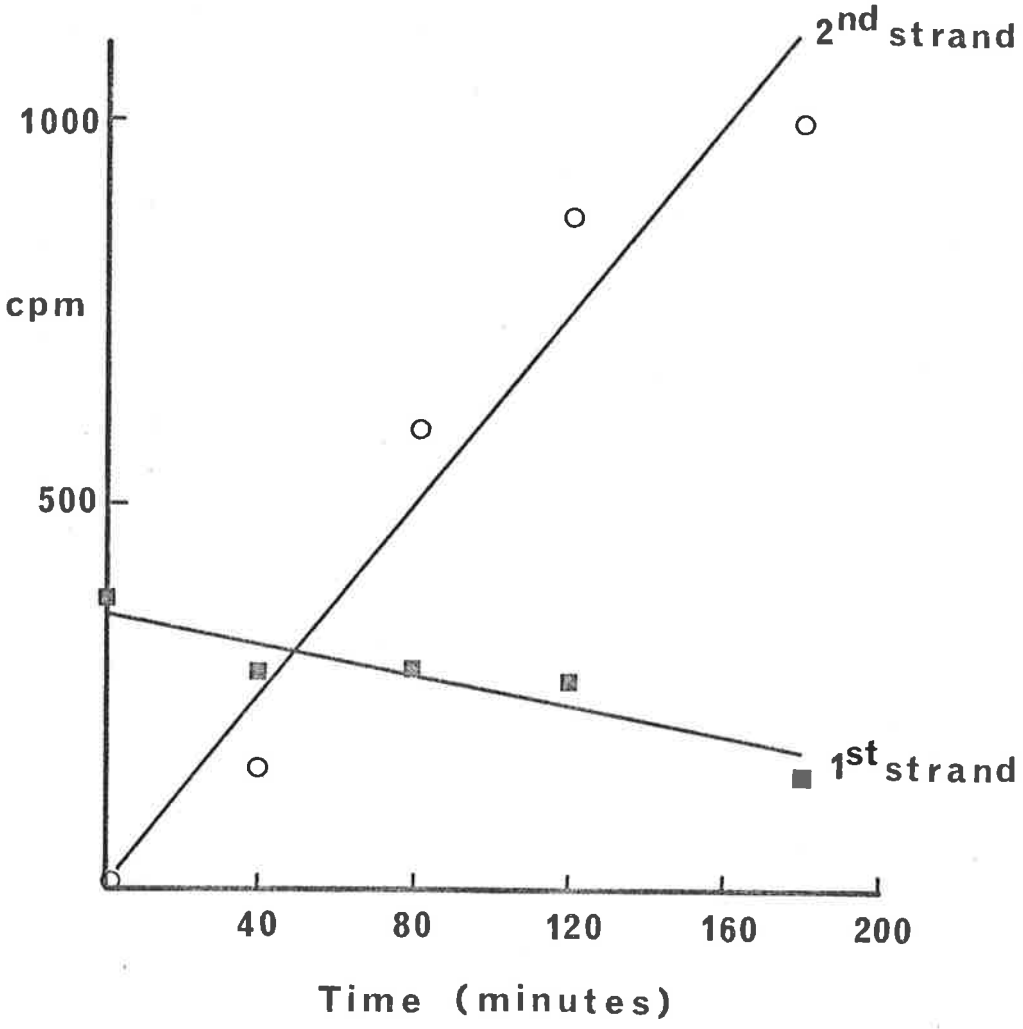

2nd strand
1000
cpm
500
1st strand
40
80
120
160
200
Time (minutes)

FIGURE 4.5. Formamide Gel Analysis of dscDNA Synthesized Using DNA Polymerase I.

Samples taken at various stages during the course of the dscDNA preparation that used *E. coli* DNA polymerase I to synthesize the second strand (see text) were ethanol precipitated, resuspended in formamide, heated in a boiling water bath, chilled, electrophoresed in 98% formamide, 5% polyacrylamide and autoradiographed. Tracks are as follows:

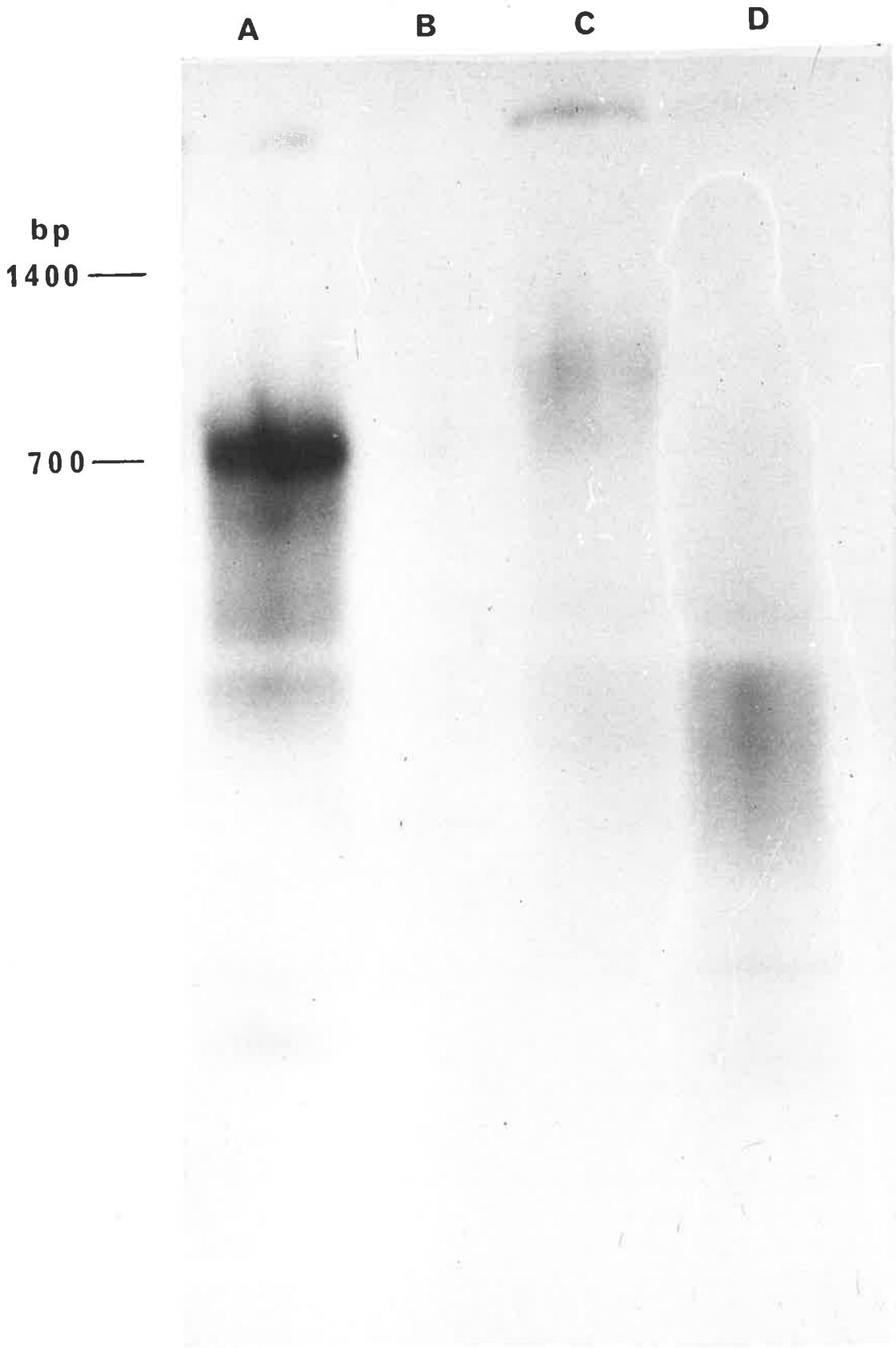
A: 12S cDNA synthesized using AMV reverse transcriptase

B: 12S cDNA digested with nuclease S_1

C: dscDNA synthesized using *E. coli* DNA polymerase I

D: dscDNA digested with nuclease S_1

The sizes of full length cDNA (700 bp) and full length dscDNA (1400 bp) are indicated.


A
B
C
D
bp
1400
700

An aliquot was also digested to determine the level of S_1 resistance in cDNA. 20% of the 0.18 µg of cDNA was resistant. Fig.4.6 shows, however that this resistant material must have been very small in size. The remainder of the cDNA was heated in a boiling water bath for 2 min to denature any cDNA-mRNA hybrids before more AMV reverse transcriptase was added for synthesis of the complementary second strand as described in Chapter 2. The dscDNA was then digested with S_1 nuclease either directly or after heating in a boiling water bath for 2 min. It was possible that during the second strand synthesis, residual mRNA could have been copied or could have reformed S_1 resistant hybrids. Boiling the dscDNA before the S_1 treatment would denature such hybrids, while the true dscDNA would snap back upon cooling because of the covalent attachment of strands at one end. The resistance of the dscDNA to S_1 nuclease with and without boiling was 50% and 55% respectively. Thus mRNA-cDNA hybrids were not contributing markedly to the S_1 resistance. Thus a final amount of 0.18 µg of S_1 nuclease-cleaved double-stranded DNA was synthesized from 1 µg of mRNA. Fig.4.6 shows that the cDNA was of increased molecular weight after second strand synthesis, as determined by formamide gel electrophoresis, and that S_1 cleavage produced a smear of dscDNA from very small species up to full length.

c) Fractionation of the dscDNA:

The dscDNA was fractionated on 5 - 20% sucrose gradients as described in Chapter 2. Fractions were pooled to give three molecular weight classes of dscDNA. Fig.4.7

FIGURE 4.6. Formamide Gel Analysis of dscDNA Synthesized Using AMV Reverse Transcriptase.

Samples taken at various stages during the course of the dscDNA preparation using AMV reverse transcriptase to synthesize the second strand (see text) were ethanol precipitated, resuspended in formamide, heated in a boiling water bath, chilled, electrophoresed in 98% formamide, 5% polyacrylamide and autoradiographed. Tracks are as follows:

A: 12S cDNA synthesized using AMV reverse transcriptase

B: 12S cDNA digested with nuclease S_1

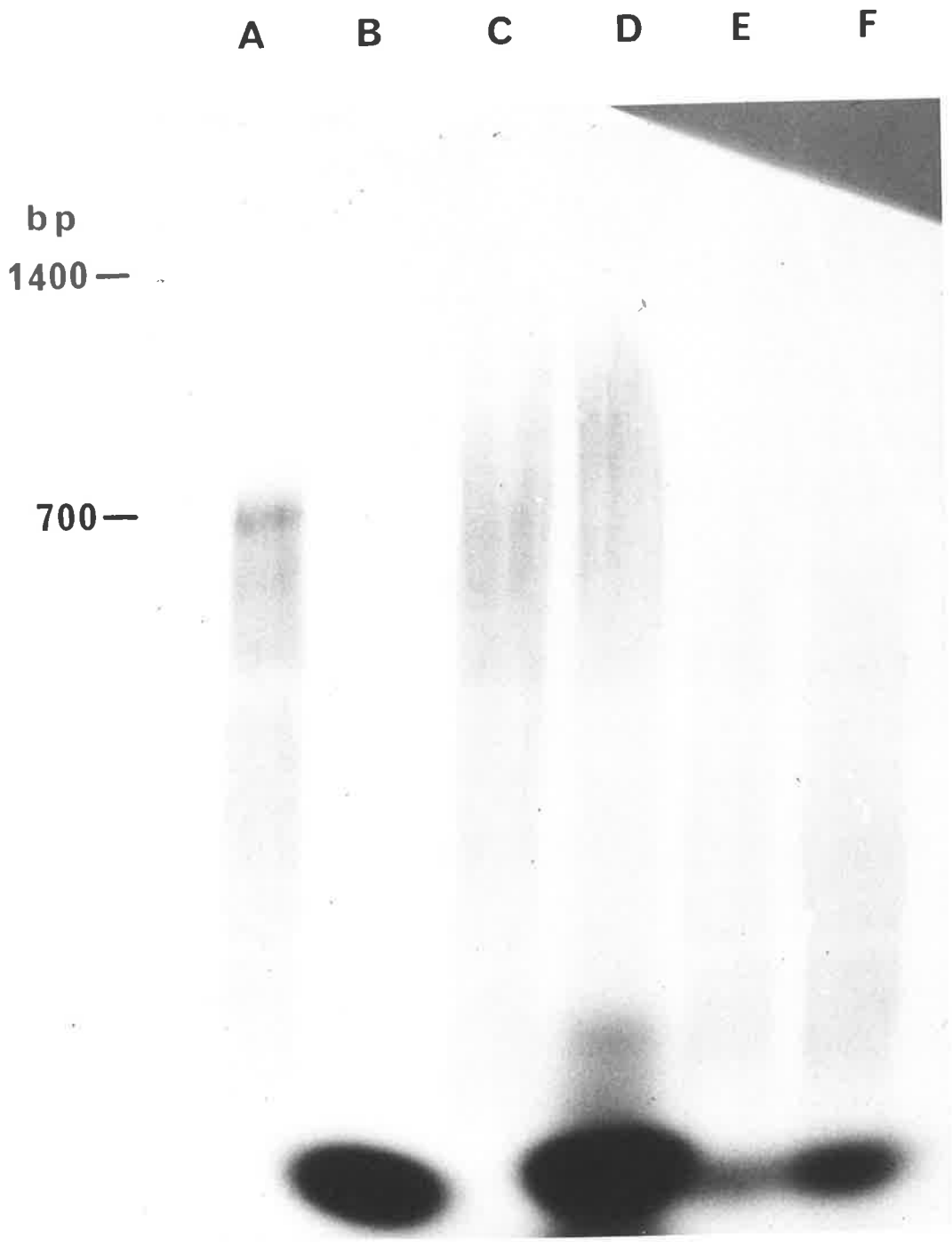
C: dscDNA synthesized using AMV reverse transcriptase in a 4 hr incubation for synthesis of the second strand

D: As for C, except that incubation was continued for an extra 2 hr.

E: dscDNA shown in Track C digested with S_1 nuclease

F: dscDNA shown in Track D digested with S_1 nuclease.

The sizes of full length cDNA (700 bp) and full length dscDNA (1400 bp) are indicated.


A
B
C
D
E
F
bp
1400
700

shows that sucrose gradient fractionation was effective in fractionating dscDNA as determined by electrophoresis of fractionated material in aqueous 3.5% polyacrylamide gels. It should be noted, however that even the high molecular weight fraction consisted of a range of different sizes. Fig.4.8 shows the sucrose gradient profile of the final preparative dscDNA preparation. The highest molecular weight fraction was pooled, ethanol precipitated as described, washed with cold 75% ethanol and resuspended in TE buffer. This technique, detailed in Chapter 2, is rapid since all incubations can be carried out in a single day. Fractionation of the dscDNA can then be carried out overnight on a sucrose gradient. An alternative method of fractionating the dscDNA is to chromatograph it on a large pore agarose column (e.g., Biogel A150M) (J. Adams, personal communication). Ethanol precipitation, tailing of the cDNA and annealing to the vector can be carried out on the following day, so that transformed colonies are ready for analysis three days after commencing the synthesis. The final yield of this high molecular weight fraction of 12S dscDNA was approximately 400 ng from 10 µg of initial 12S mRNA.

2. Molecular cloning of 12S dscDNA:

a) Tailing, annealing and transforming the dscDNA:

Tailing of the dscDNA was carried out as described in Chapter 2 using terminal deoxynucleotidyl transferase and dCTP or dTTP as substrate. $CoCl_2$ was used in the solution to obviate the requirement for free 3'OH

FIGURE 4.7. Gel Analysis of Fractionated 12S dscDNA.

Total 12S dscDNA prepared by the AMV reverse transcriptase method (see text) was sedimented on a 5 - 20% sucrose gradient at 170,000 g for 16 hours. Three broad cuts were made in the gradient, the material ethanol precipitated, electrophoresed on an aqueous 3.5% polyacrylamide gel and autoradiographed. Track A represents the lowest molecular weight material, Track B the material of intermediate molecular weight and Track C the highest molecular weight material. Although fractionation was clearly achieved, the size ranges of fragments in adjacent fractions overlapped.

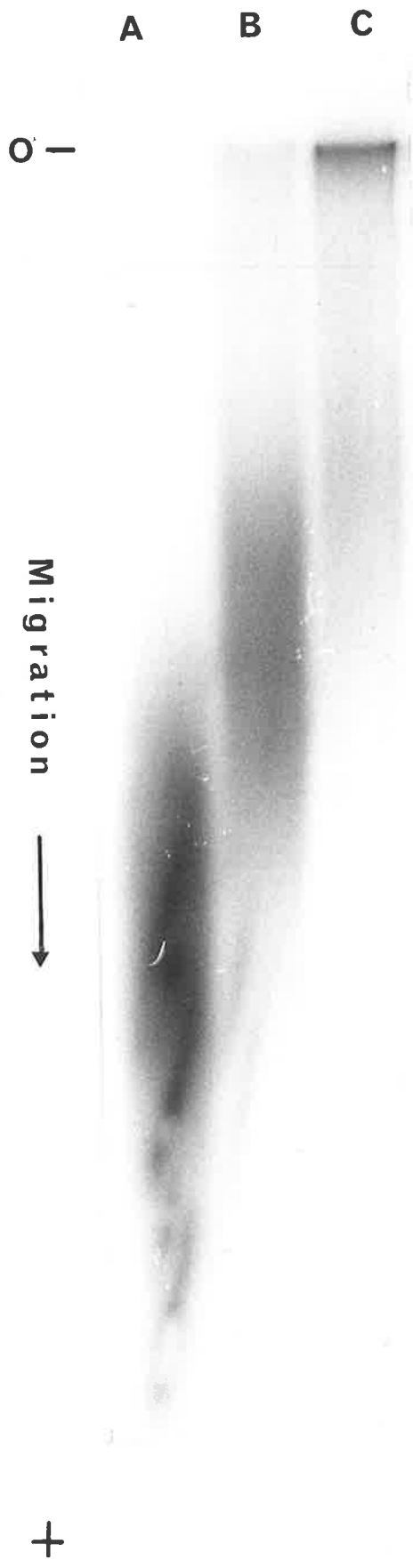
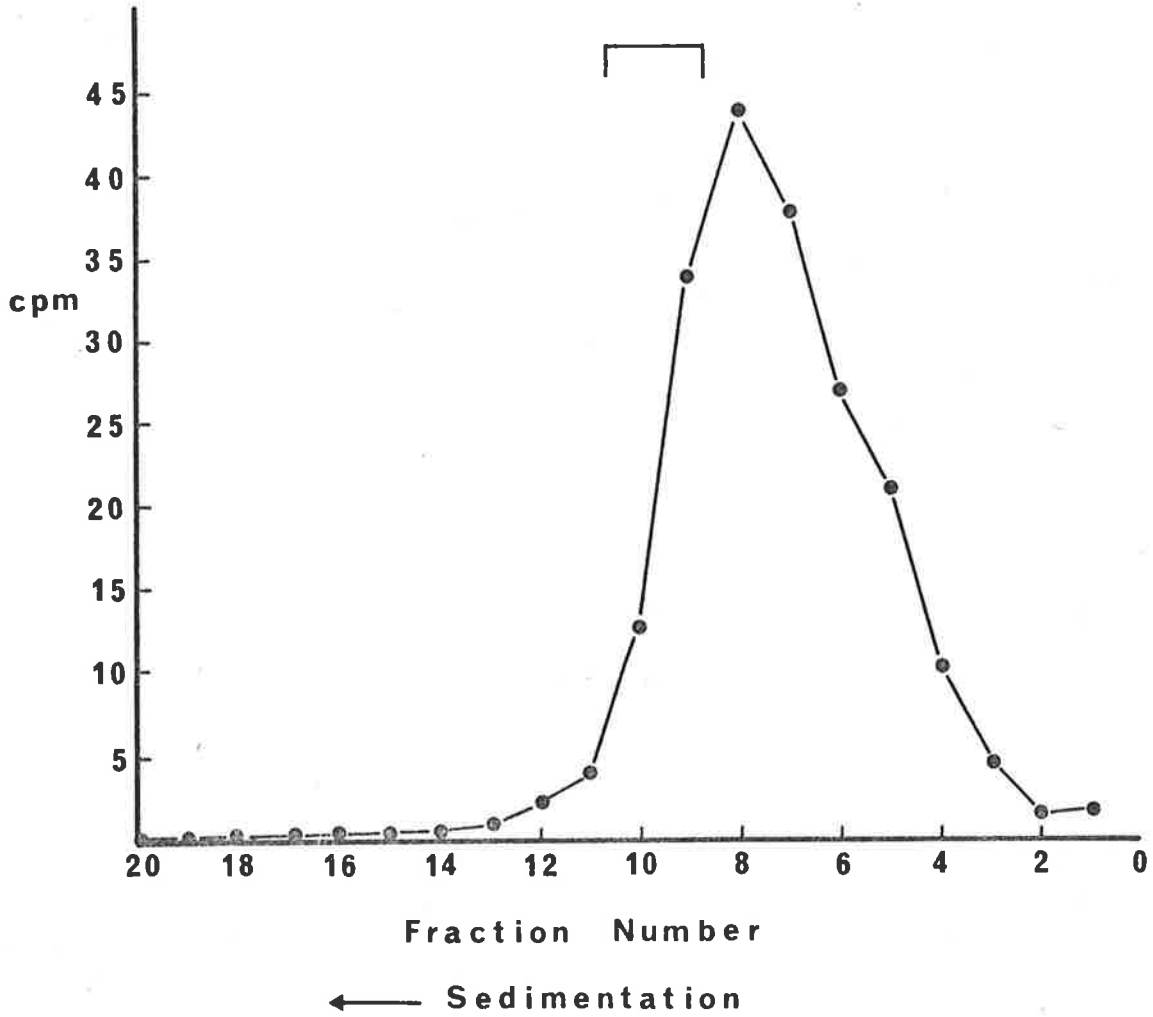

A
B
C
O
Migration
+

FIGURE 4.8. Fractionation of the Preparative 12S dscDNA Preparation.

The large scale 12S dscDNA preparation (see text) was sedimented on a 5 - 20% sucrose gradient at 170,000 g for 16 hrs. The gradients were fractionated and the 0.5 ml fractions containing the ^{32}P-labelled dscDNA were counted using Cerenkov radiation. The highest molecular weight material was pooled as indicated, ethanol precipitated and used for the subsequent cloning.


45
40
35
cpm
30
25
20
15
10
5
20
18
16
14
12
10
8
6
4
2
0
Fraction Number
Sedimentation

residues (Roychoudhury *et al.*, 1976). Figure 4.9 shows a typical time course for the addition of nucleotides to the DNA. When an average of 10 - 20 nucleotides had been added to each end of the molecules, the reaction was terminated by the addition of excess EDTA as described in Chapter 2.

The dscDNA tailed with dC or dT was annealed to pBR322 DNA tailed with dG or λ Charon 3A DNA tailed with dA respectively and used to transform and transfect *E. coli* as described in Chapter 2. The transformation efficiency of pBR322 was 2×10^7 colonies per μg, while *Pst*I digested and tailed pBR322 DNA transformed with an efficiency of 10^3 colonies per μg. When 4 ng of pBR322 DNA annealed to an equimolar amount of tailed dscDNA was used to transform *E. coli*, 200 colonies resulted, i.e., 50 fold over background and equivalent to 5×10^4 transformants per μg of pBR322 DNA annealed to the cDNA.

λ Charon 3A DNA transfected with an efficiency of 2×10^5 *pfu* per μg, while 10 ng of cleaved and tailed λ Charon 3A DNA gave no transfectants. Annealed λ Charon 3A DNA and 12S dscDNA yielded approximately 10^3 recombinants per μg of λ Charon 3A DNA.

b) Detection of recombinants:

Recombinant clones were detected using the procedures described by Grunstein and Hogness (1975) for pBR322 and Benton and Davis (1976) for λ Charon 3A as described in Chapter 2, using 12S cDNA as the probe. The results of these screenings are shown in Fig.4.10 along with a duplicate set of plaques annealed to 18S and 28S

FIGURE 4.9. Kinetics of dC Tailing of 12S dscDNA.

The dscDNA was incubated with terminal deoxynucleotidyl transferase in the presence of ^{3}H-dCTP as described in Chapter 2. 1 µl samples were taken at the times indicated and the amount of acid insoluble radioactivity determined. The dotted line indicates the number of counts incorporated equivalent to the addition of 15 dC residues on average per 3'OH end of the molecules.

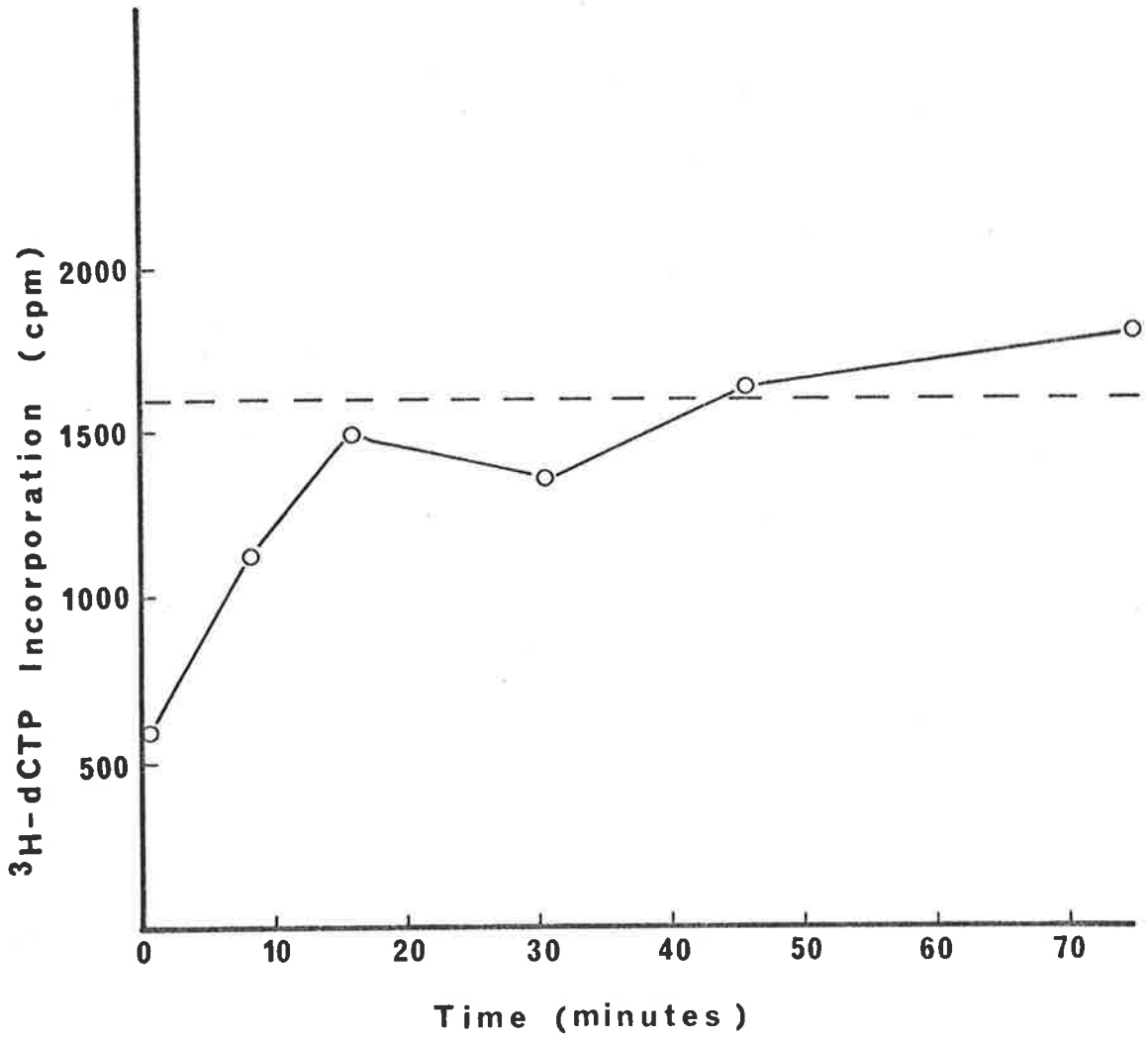

^{3}H-dCTP Incorporation (cpm)
2000
1500
1000
500
0
10
20
30
40
50
60
70
Time (minutes)

FIGURE 4.10. Detection of Plasmid and Phage Recombinants.

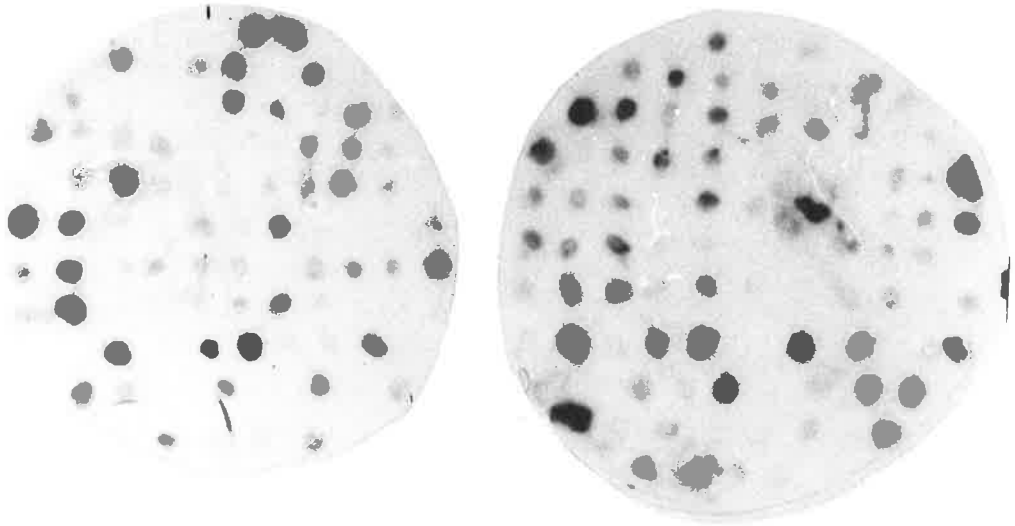
Colonies of recombinant plasmid transformed cells and plaques of recombinant phage transfected cells were denatured *in situ*, immobilized onto nitrocellulose filter paper, annealed to ^{32}P-labelled 12S cDNA or 18S and 28S ribosomal cDNA, washed, dried and autoradiographed.

a) Two sets of filters carrying immobilized colonies, annealed to ^{32}P-labelled 12S cDNA.

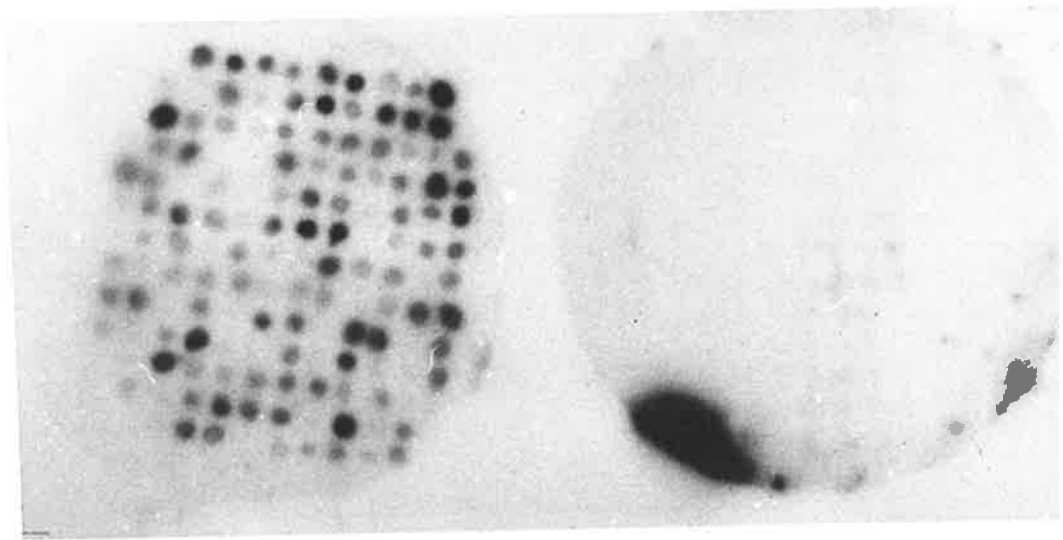
b) Replica filters of immobilized plaques annealed to ^{32}P-labelled 12S cDNA (left plate) or 18S and 28S ribosomal cDNA (right plate).

All recombinants were tested for the presence of ribosomal sequences, but none were detected.

a



b



ribosomal cDNA. All recombinants were tested for ribosomal sequences, but none were detected. The feature that was of most immediate interest was the variability in intensity of the different recombinants. As these could have represented different 12S cDNA species of differing abundance, it was decided to select colonies giving a range of intensities rather than concentrate on the most abundant species. The significance of this selection is discussed at the end of this Chapter. It should also be noted that for the reason described at the end of this Chapter, no further use could be made of the sequences cloned into λ Charon 3A, although use will be made of these in the future.

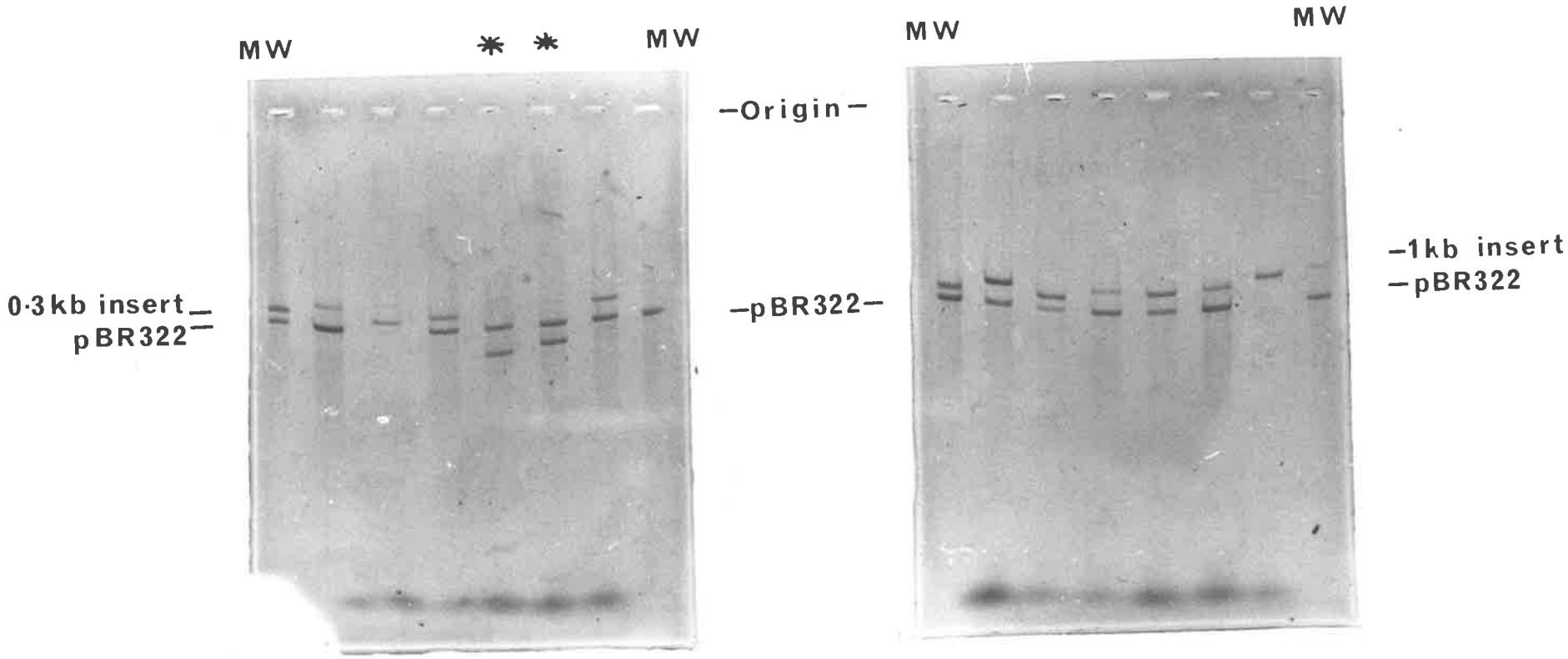
c) 'Miniscreen' length determinations of recombinants:

In order to gain an approximate estimate of the size of the inserted segment, small amounts of the recombinant plasmids were prepared as described in Chapter 2 and electrophoresed on 1% agarose gels. Figure 4.11 shows the results of two such 'miniscreen' preparations of recombinant plasmids linearized with *Eco*RI. Linearized pBR322 was included in each track as a marker, so that the presence of a second band of decreased mobility (increased size) indicated the presence of a recombinant plasmid. Two exceptions were the two recombinants that contained an *Eco*RI site within the insert and were therefore cleaved into two fragments. The presence of *Eco*RI sites within some of the inserted segments was the first indication of sequence heterogeneity between cloned species. The molecular weight markers used in the two side tracks were linearized pBR322

FIGURE 4.11. Miniscreen Analysis of Recombinant Plasmids.

Miniscreen amounts of plasmids annealing to 12S cDNA were prepared, by lysozyme and heat shock treatment, linearized with *Eco*RI, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Linearized parental plasmid pBR322 was included in all except the second track from the right. Thus recombinant plasmids are detected by the presence of a band of lower mobility (greater size) electrophoresing above the pBR322 DNA, except in the tracks marked with a star, denoting the presence of recombinant plasmids which have an *Eco*RI site within the inserted DNA. This results in the generation of two fragments electrophoresing with greater mobility than the pBR322 marker.

Linearized plasmids containing characterized inserts of 300 and 1000 base pairs respectively (N. Gough, personal communication) were electrophoresed in the end tracks, denoted by the letters MW.


MW
* *
MW
MW
MW
–Origin–
0·3kb insert
pBR322
–pBR322–
–1kb insert
–pBR322

plasmids carrying inserted immunoglobulin dscDNA sequences of approximately 300 base pairs and 1000 base pairs respectively (N. Gough, personal communication). The 12S cDNA inserts appeared to range in size from approximately 500 base pairs to approximately 800 base pairs. Several of these clones were chosen for study, purified and prepared as described in Chapter 2.

3. Analysis of the 12S cDNA recombinants:

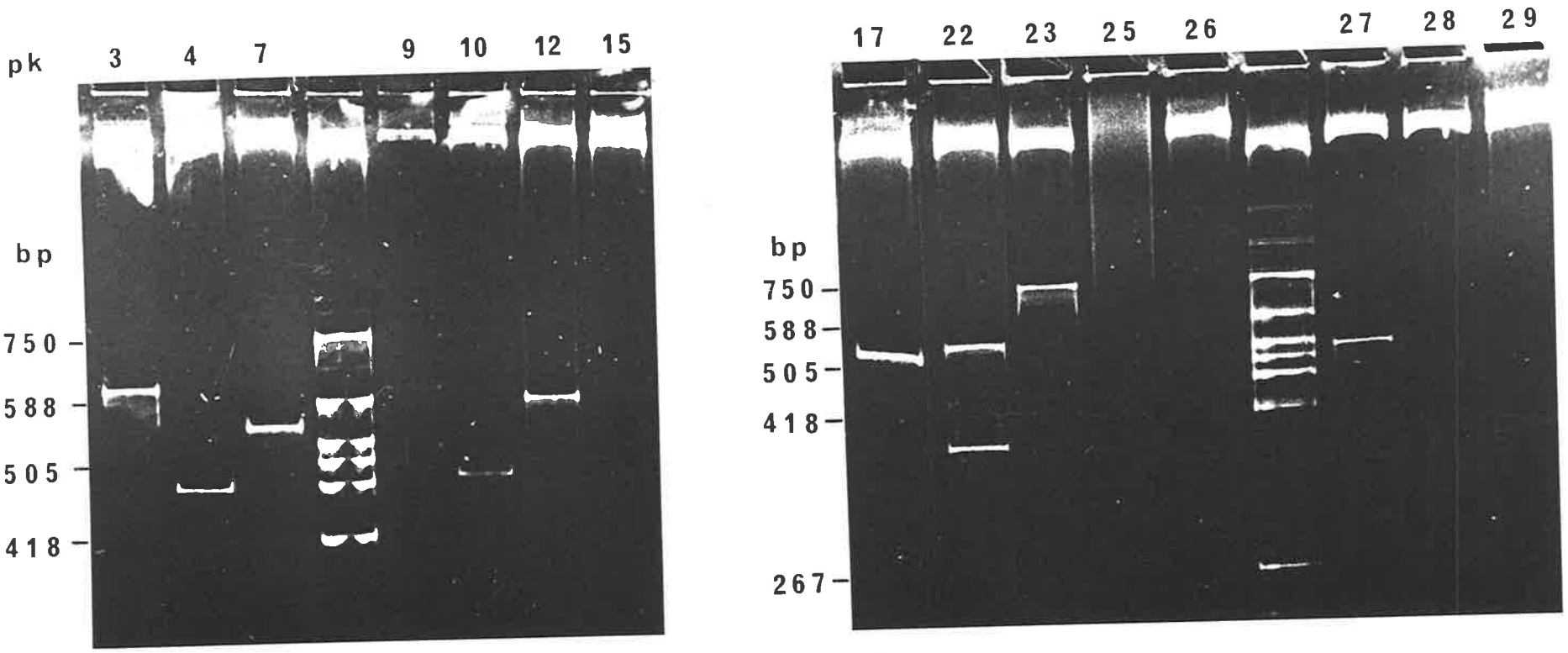
Fifteen recombinant plasmids which gave varying intensities in response to the colony screening described above were prepared as described in Chapter 2. These recombinants were analysed using restriction endonuclease digestion.

a) Resection with *Pst*I:

Each recombinant plasmid was digested with *Pst*I to determine if the *Pst*I sites had been regenerated as described above. Figure 4.12 shows the results of electrophoresing these digests on aqueous 6% polyacrylamide gels. Twelve of the fifteen recombinants could be resected with *Pst*I showing that in 80% of cases both of the *Pst*I sites had been regenerated. Figure 4.12 also shows that some of the recombinants contained *Pst*I recognition sequences in the inserted DNA, as shown by the presence of more than one band of low molecular weight, again suggesting heterogeneity of sequence between cloned species. The lengths of the inserts ranged from 400 base pairs to 935 base pairs with an average of approximately 600 nucleotides, (85% full length). *Hae*III-digested pBR322 DNA was used as the molecular weight markers in these gels (Sutcliffe, 1978).

FIGURE 4.12. Resection of Inserted DNA with *Pst*I.

DNA of the 15 recombinants prepared in bulk were each cleaved with *Pst*I, electrophoresed in aqueous 6% polyacrylamide, stained with ethidium bromide and photographed under UV light. *Hae*III cleaved pBR322 fragments were used as molecular weight markers. 80% of the recombinants could be resected with *Pst*I compared with 40% in the only reported study to date (Villa-Komaroff *et al.*, 1978).


pk
3
4
7
9
10
12
15
bp
750
588
505
418
17
22
23
25
26
27
28
29
bp
750
588
505
418
267

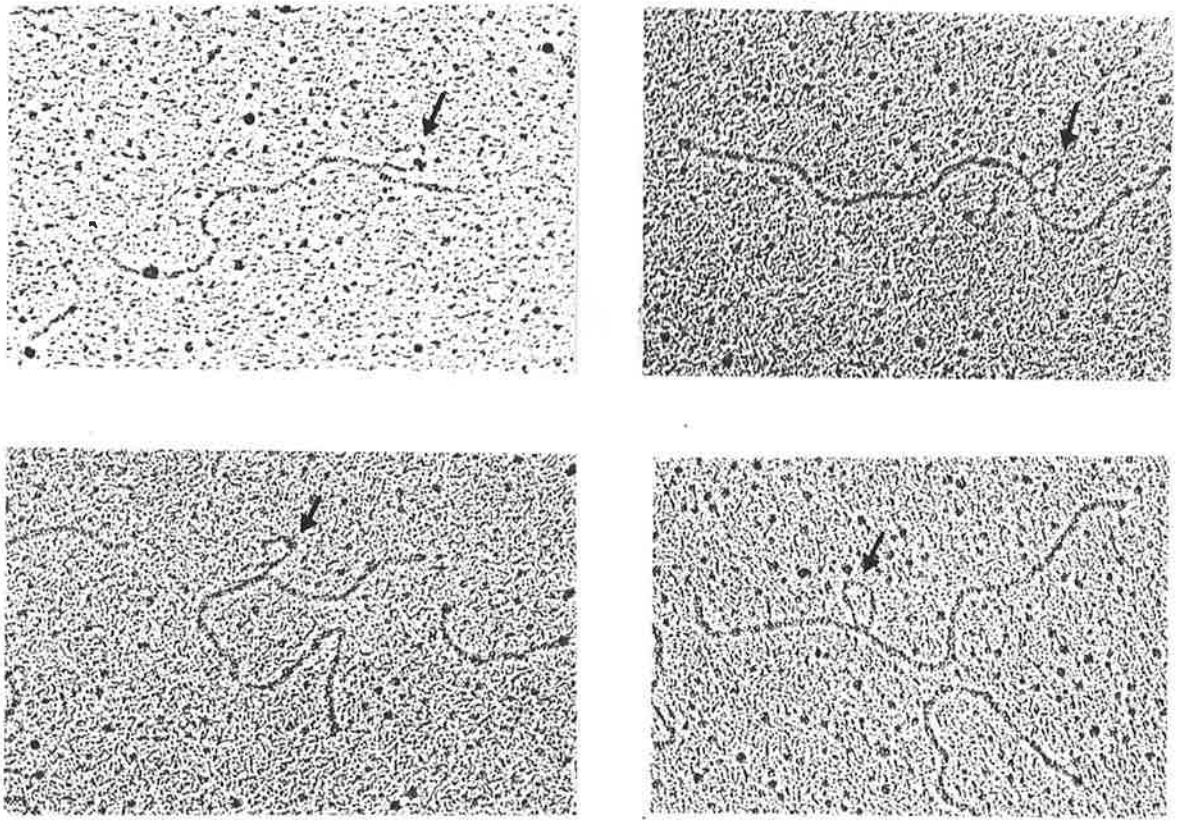
b) Heteroduplex analysis:

The size of 12S mRNA has been estimated to be approximately 760 nucleotides (Kemp *et al.*, 1974b) containing an average poly A length of 60 nucleotides (Morris and Rogers, 1979). Assuming that the average length of each set of G - C base pairs joining the insert to the vector is 20, resected inserts should be of the order of 740 - 800 base pairs long. It was therefore surprising that one recombinant plasmid, pK22, yielded, on *Pst*I cleavage two fragments of size 560 and 410 base pairs respectively (Fig.4.12). As it was possible that this DNA was a mixed population of recombinant plasmids, heteroduplexes were formed, as described in Chapter 2, between pK22 DNA and the parental pBR322 DNA so that an independent estimate of insert length could be made from visualization of the insert loops in the electron microscope.

Figure 4.13 shows several examples of the loop-outs observed with these heteroduplexes. Measurements on 51 such heteroduplexes, using single-stranded ϕX174 as the internal size standard, gave an estimate of the mean length of the insert as 1,008 base pairs with a standard deviation of 75 base pairs. Thus the pK22 DNA preparation must have consisted of a homogeneous population of recombinant plasmids carrying an insert of length 960 - 1000 base pairs and containing a *Pst*I recognition site within the insert. A close examination of the heteroduplexes shown in Fig.4.13 suggested a possible reason for the large size of this insert. Several of these heteroduplexes appeared to have a 'stalk' of duplex DNA at the base of the loop. In

FIGURE 4.13. pK22 Heteroduplex Analysis.

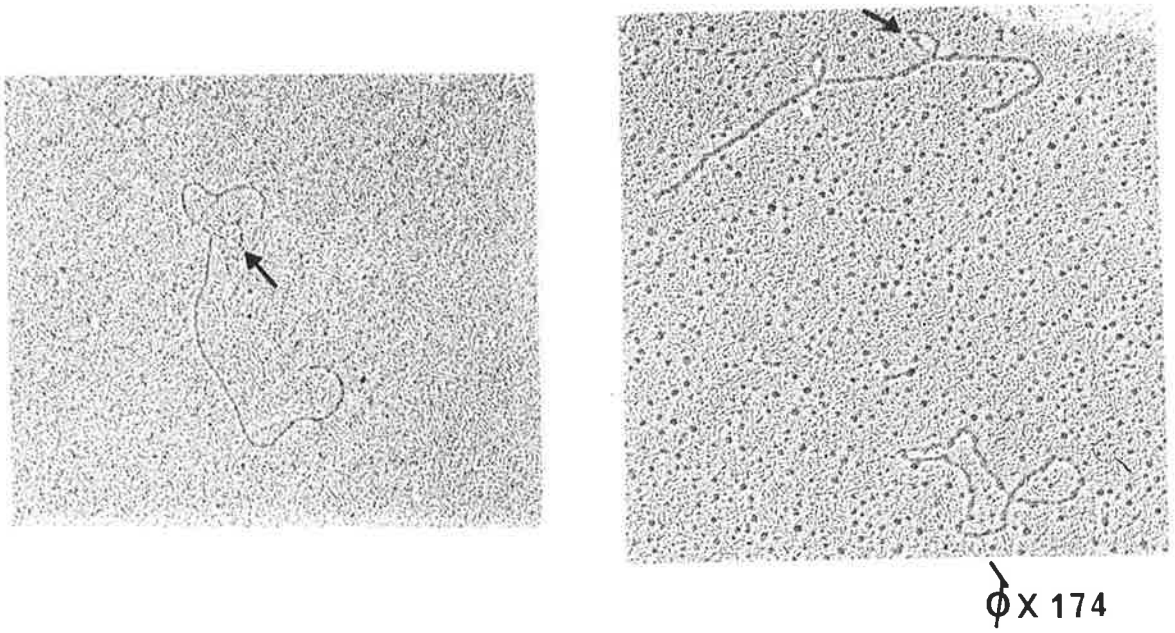
Four characteristic heteroduplexes formed between recombinant pK22 and the parental pBR322 molecule and visualized as described in Chapter 2 are shown. Magnification is 50,000 fold.

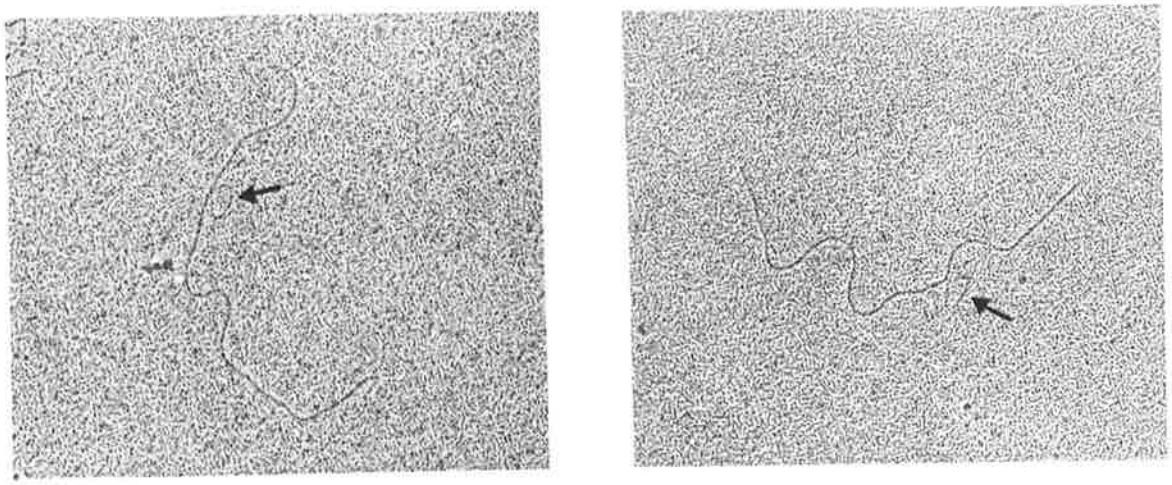


0 0·2 0·4 0·6 0·8 1·0 μ

FIGURE 4.14. pK10 Heteroduplex Analysis.

Four characteristic heteroduplexes formed between recombinant pK10 and the parental pBR322 molecule and visualized as described in Chapter 2 are shown. Magnification is 50,000 fold.


ϕX 174



0 0·2 0·4 0·6 0·8 1·0

contrast, heteroduplexes with pK10 had no such 'stalk' (Fig.4.14). Since the G and C stretches that join the insert to the vector are located at the base of the loop and can base-pair to form a duplex, it appears that pK22 had an unusually long stretch of these residues, which would account for the unusually large size of this particular recombinant. Labelling the 3'OH residues of the *Pst*I resected inserts by end filling with DNA polymerase I, depurination of the DNA with diphenylamine and gel electrophoresis would allow determination of the length of the G - C stretches and thus test this possibility.

c) Restriction cleavage of the recombinant DNA:

The enzymes *Eco*RI, *Hin*dIII, *Bam*HI, *Sal*I and *Bgl*II were all used to test for the presence of cleavage sites within the inserted DNA. *Eco*RI, *Hin*dIII *Bam*HI and *Sal*I all have a single cleavage site within the vector pBR322 DNA, so that the presence of a cleavage site within the insert will result in the DNA being cleaved into two fragments. No cleavage site within the insert results in linearization of the plasmid. Figures 4.15 and 4.16 show the digests with these four enzymes electrophoresed on 1% agarose gels. In each case at least one and at most four of the 15 recombinants were cleaved with each restriction enzyme.

Plasmid pBR322 does not have a *Bgl*II cleavage site, so that cleavage of a site within the inserted DNA results in linearization of the supercoiled and covalently closed circular forms of the plasmid. Fig.4.17(a) shows that

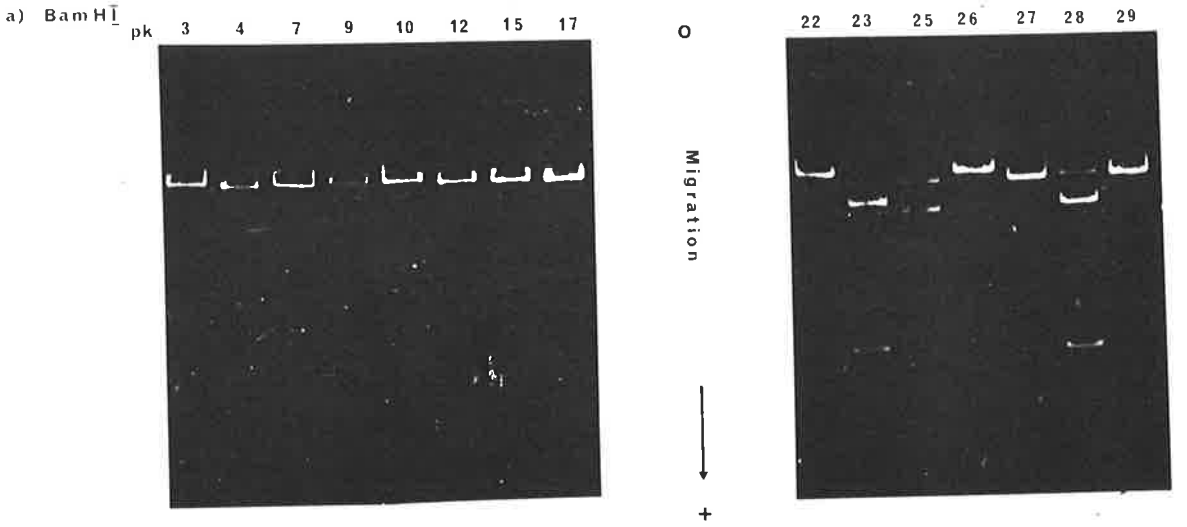
FIGURE 4.15. *Bam*HI and *Eco*RI Cleavage of Recombinant Plasmids.

Each recombinant plasmid was cleaved with

a) *Bam*HI

b) *Eco*RI,

electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Plasmid pBR322 contains one site for each of these endonucleases, so that presence of a site in the inserted DNA segment will result in the generation of two fragments. Thus *Bam*HI has cleavage sites within the inserts of pK23, pK25 and pK28 and *Eco*RI has cleavage sites within the inserts of pK10, pK15, pK22, and pK25.


a) BamHI
pk 3 4 7 9 10 12 15 17
22 23 25 26 27 28 29
O
Migration
+

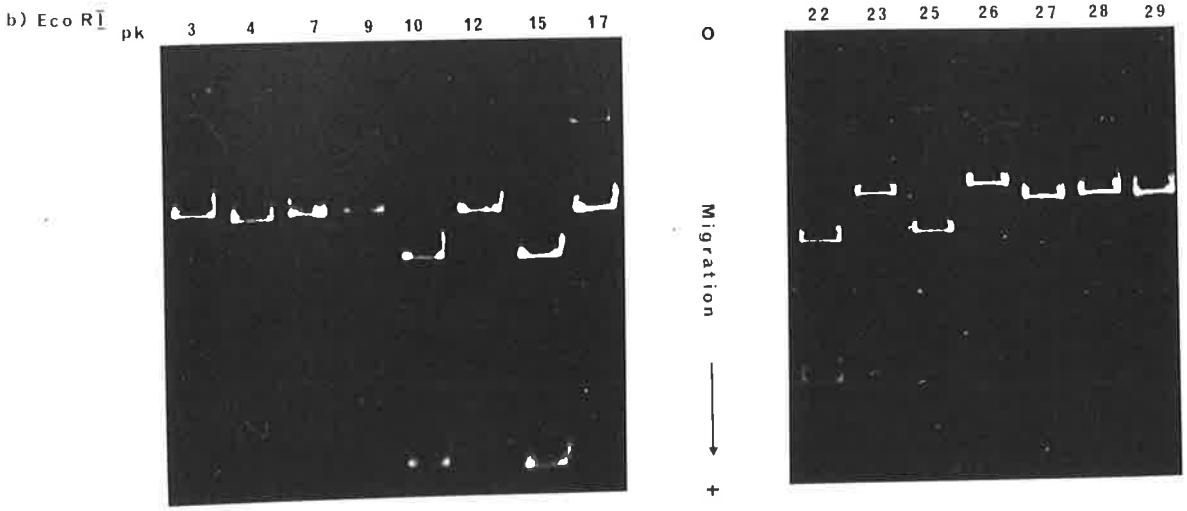

b) EcoRI
pk 3 4 7 9 10 12 15 17
22 23 25 26 27 28 29
O
Migration
+

FIGURE 4.16. *Hin*dIII and *Sal*I Cleavage of Recombinant Plasmids.

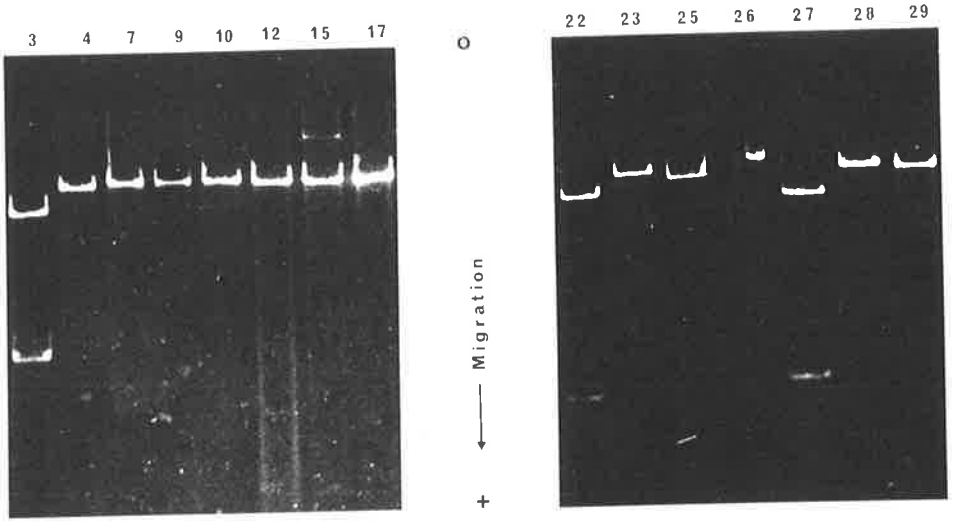
Each recombinant plasmid was cleaved with

a) *Hin*dIII

b) *Sal*I,

electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Plasmid pBR322 contains one site for each of these endonucleases, so that presence of a site in the inserted DNA segment will result in the generation of two fragments. Thus *Hin*dIII has cleavage sites within the inserts of pK3, pK22 and pK27 and *Sal*I has cleavage sites within the insert of pK26.

a) Hind III



b) Sal I

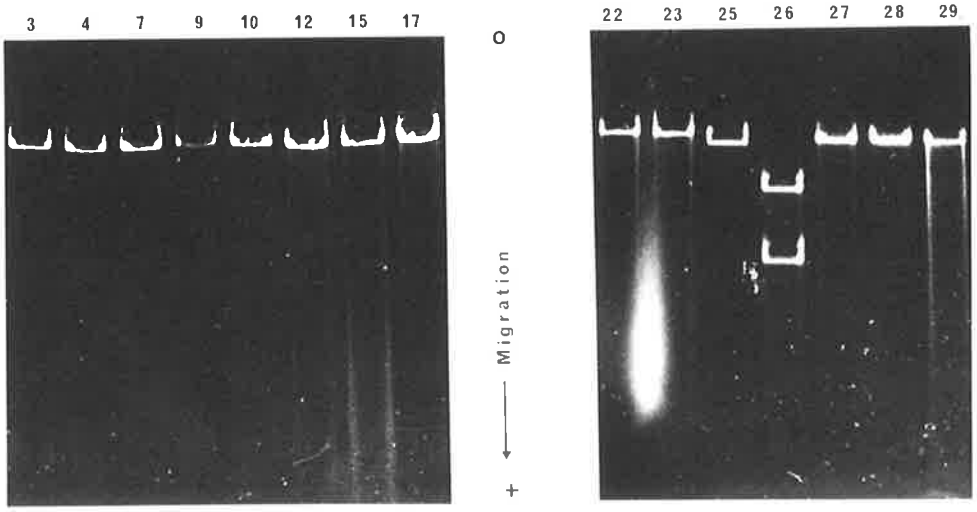
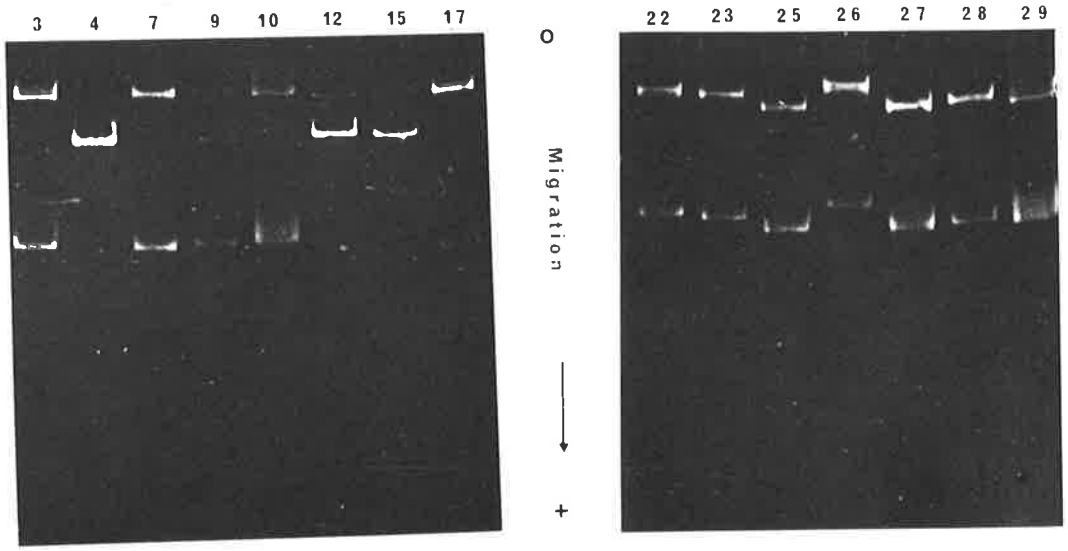


FIGURE 4.17. *Bgl*II Cleavage of Recombinant Plasmids:

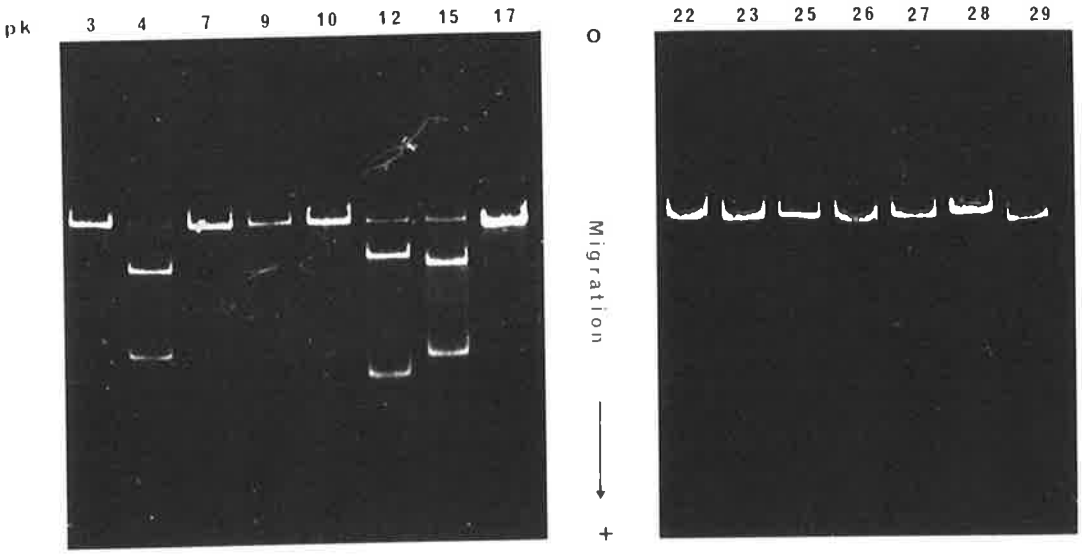
a) Each recombinant plasmid was cleaved with *Bgl*II, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. As pBR322 has no *Bgl*II cleavage site, presence of a site within the inserted DNA results in linearization of the superhelical and covalently closed circular forms of the plasmid. The absence of these forms for pK4, pK12, and pK15 suggested that they contained a *Bgl*II site within the inserted DNA.

b) Each recombinant plasmid was cleaved with *Bgl*II and *Sal*I except pK26 which was cleaved with *Bgl*II and *Bam*HI. In each case the second enzyme cleaved the plasmid once only to linearize it, so that presence of a *Bgl*II site within the inserted DNA resulted in the appearance of two fragments. Recombinants pK4, pK12 and pK15 were therefore confirmed to have a *Bgl*II recognition site within the inserted DNA.

a)


pk 3 4 7 9 10 12 15 17
O
Migration
+
22 23 25 26 27 28 29

b)


pk 3 4 7 9 10 12 15 17
O
Migration
+
22 23 25 26 27 28 29

this occurred with 3 of the recombinant plasmids, pK4, pK12 and pK15, since the two bands representing supercoiled DNA and covalently closed circular DNA were replaced by a single linear species. To test this interpretation all plasmids were digested with *Bgl*II and a second enzyme known to cleave at a unique site in the pBR322 part of the DNA and not the insert. Thus *Sal*I was used as the second enzyme in all cases except pK26 which has a *Sal*I site in the inserted DNA. *Bam*HI, having no site in the insert of pK26 was used as the second enzyme for this recombinant. In those cases where a *Bgl*II cleavage site exists within the insert the plasmid will be cleaved into two fragments, whereas the absence of a *Bgl*II cleavage site within the insert would result in linearization of the DNA by the second enzyme. Figure 4.17(b) shows the results of electrophoresing these double digests on 1% agarose, confirming that pK4, pK12 and pK15 contain a *Bgl*II cleavage site within the inserted DNA.

A summary of the cleavage patterns of these recombinants is given in Table 4.1.

d) Mapping restriction sites:

Double digestion of the recombinant DNAs with *Pst*I to resect the insert and a second enzyme that cleaves within the insert allowed the determination of the distance of that cleavage site from the *Pst*I ends. It should be stressed that orientation of the inserted DNA is unknown and that this 'mapping' does not presume orientation of inserts.

Figure 4.18 shows these digests electrophoresed on

TABLE 4.1. Analysis of 12S dscDNA Recombinants.

Summary of several features of the recombinants discussed in Chapter 4. These are the presence of regenerated *Pst*I sites at the two insert-pBR322 junctions, and the size of the inserts, both determined by aqueous acrylamide gel electrophoresis (see Fig. 4.12) as well as presence within the inserted DNA of cleavage sites for the restriction endonucleases indicated.

X indicates presence of the relevent cleavage site.

TABLE 4.1.

Recombinant (pK) Number	*Pst*I Sites Regenerated	Size (base pairs)	Cleavage sites present in inserted DNA.					
			*Bam*HI	*Bgl*II	*Eco*RI	*Hin*dIII	*Pst*I	*Sal*I
3	X	670				X		
4	X	500		X				
7	X	550						
9	X	610						
10	X	620			X			
12	X	590		X				
15	X	620		X	X		X	
17	X	530						
22	X	970			X	X	X	
23	X	740	X					
25		-	X		X			
26		-						X
27	X	560				X		
28		-	X					
29	X	400						

an aqueous 8% polyacrylamide gel. Only one site was present in each insert, so that the resulting two fragments add up to the original size of the insert. The only exception to this was pK10, which was partially digested by *Eco*RI to give fragments of size 330 and 160 base pairs as well as the original 490 base pair insert. Recombinants pK26 and pK28 were also double digested as described, since regeneration of one of the *Pst*I sites would lead to the appearance of a small fragment. In neither case, however was such a fragment present, indicating that neither *Pst*I site had been regenerated in these cases. The relative positions of the restriction sites are drawn diagrammatically in Fig.4.20. The recombinant pK22 has restriction sites of *Pst*I, *Eco*RI and *Hin*dIII within the insert, so it was more difficult to map. Figure 4.19 shows the results of digesting pK22 with *Pst*I and with combinations of *Pst*I, *Eco*RI and *Hin*dIII. Track B shows the complete digest of pK22 with *Pst*I electrophoresed on the aqueous 8% polyacrylamide gels resulting in two insert bands of 560 and 410 base pairs. The very large fragment electrophoresing at the top of the gel is the remainder of the recombinant DNA, that portion representing the pBR322 vector. Tracks C and G show *Pst*I-*Hin*dIII and *Pst*I-*Eco*RI double digests respectively. *Hin*dIII cleaved within the 560 base pair fragment to generate fragments of size 390 and 170 base pairs, while *Eco*RI cleaved within the 410 base pair fragment to generate fragments of size 340 and 70 base pairs. There are four possible orientations of these four fragments that would result in the distance between the *Hin*dIII and *Eco*RI sites

FIGURE 4.18. Double Digests of Recombinant Plasmids.

Recombinant plasmids were digested, electrophoresed on an aqueous 8% polyacrylamide gel, stained with ethidium bromide and photographed under UV light. Tracks are as follows:

A. *Hae*III digest of pBR322 DNA

B. *Pst*I/*Hin*dIII double digest of pK3 DNA

C. *Pst*I/*Bgl*II double digest of pK4 DNA

D. *Pst*I/*Eco*RI double digest of pK10 DNA

E. *Pst*I/*Bgl*II double digest of pK12 DNA

F. *Pst*I/*Bam*HI double digest of pK23 DNA

G. *Pst*I/*Sal*I double digest of pK26 DNA

H. *Pst*I/*Hin*dIII double digest of pK27 DNA

I. *Pst*I/*Bam*HI double digest of pK28 DNA

J. *Hae*III digest of pBR322 DNA.

Fragment sizes are expressed in base pairs.

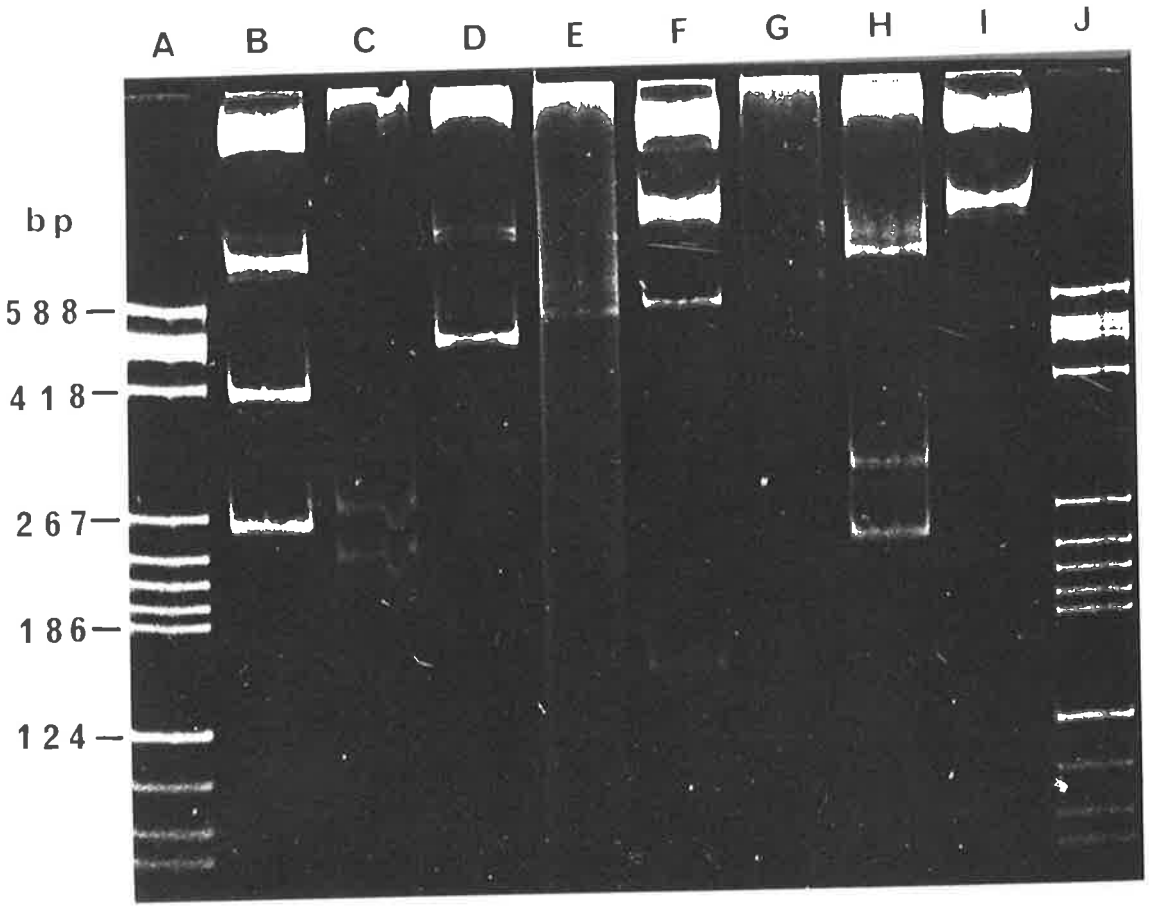

A
B
C
D
E
F
G
H
I
J
bp
588
418
267
186
124

FIGURE 4.19. Restriction Mapping of the pK22 Insert.

Restriction digests of pK22 were electrophoresed on an aqueous 8% polyacrylamide gel, stained with ethidium bromide and photographed under UV light. Tracks are as follows:

A. *Hae*III digest of pBR322

B. *Eco*RI/*Hin*dIII double digest of pK22

C. *Pst*I/*Eco*RI double digest of pK22

D. *Hin*dIII digest of pK22 redigested partially with *Pst*I (0.25 µl *Pst*I for 4 hr).

E. *Hin*dIII digest of pK22 redigested partially with *Pst*I (0.5 µl *Pst*I for 4 hr).

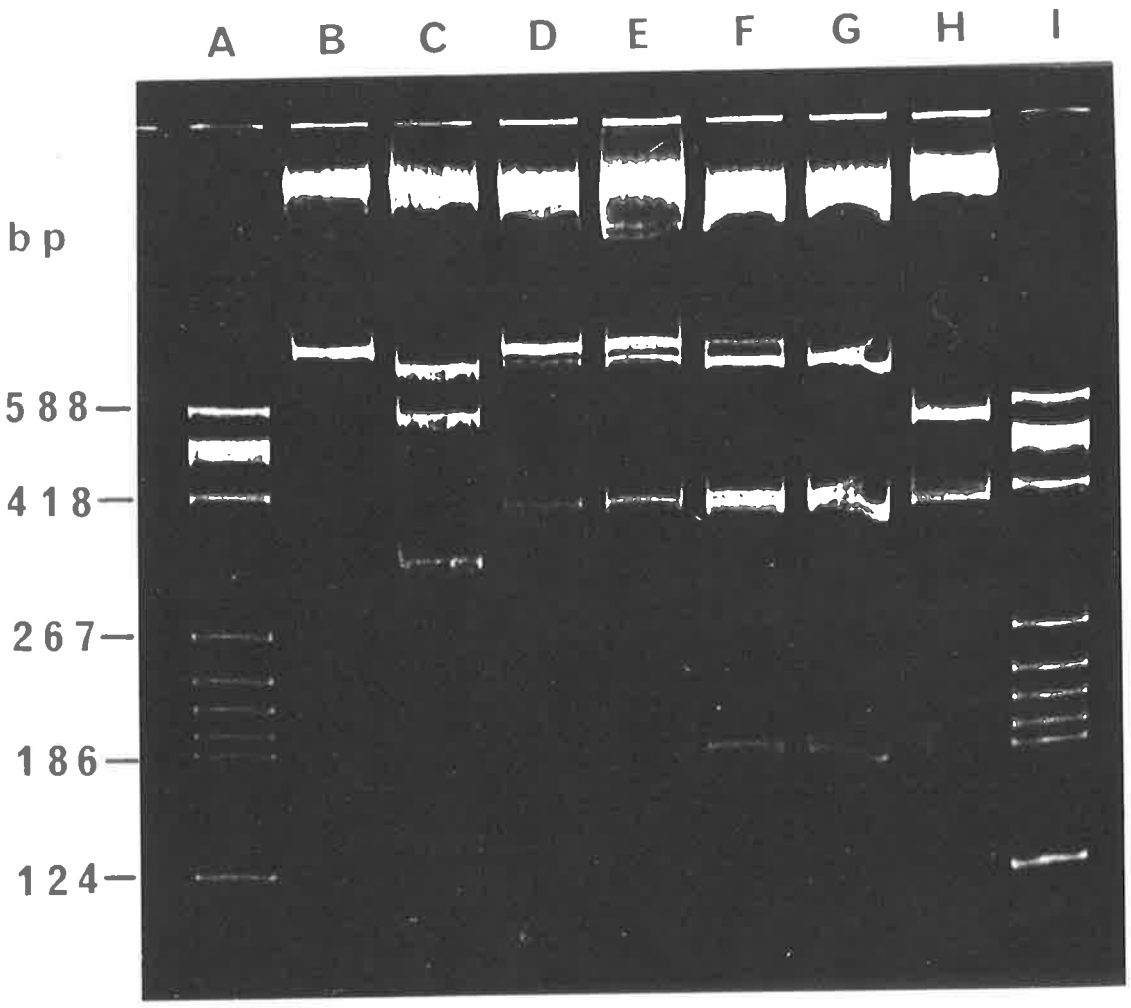
F. *Hin*dIII digest of pK22 redigested partially with *Pst*I (1.0 µl *Pst*I for 4 hr).

G. *Hin*dIII digest of pK22 redigested to completion with *Pst*I. (2.0 µl *Pst*I for 4 hr).

H. *Pst*I digest of pK22

I. *Hae*III digest of pBR322 DNA.

Fragment sizes are expressed in base pairs.


A
B
C
D
E
F
G
H
I
bp
588
418
267
186
124

being 730, 510, 460 or 240 base pairs respectively. Track H shows that the *Eco*RI-*Hin*dIII double digest generated a fragment of 730 base pairs and definitely no fragments of 510 base pairs or lower. Thus the 390 and 340 base pair fragments must lie between the *Hin*dIII and *Eco*RI sites. This final interpretation is presented in Fig.4.20.

4. Pure 12S mRNA species as probes to restricted chick chromosomal DNA:

a) 12S mRNA is comprised of two groups of species:

Several of the cloned 12S cDNA species were used to detect complementary sequences in restricted and fractionated chick DNA by the method of Southern (1975) as described in Chapter 2.

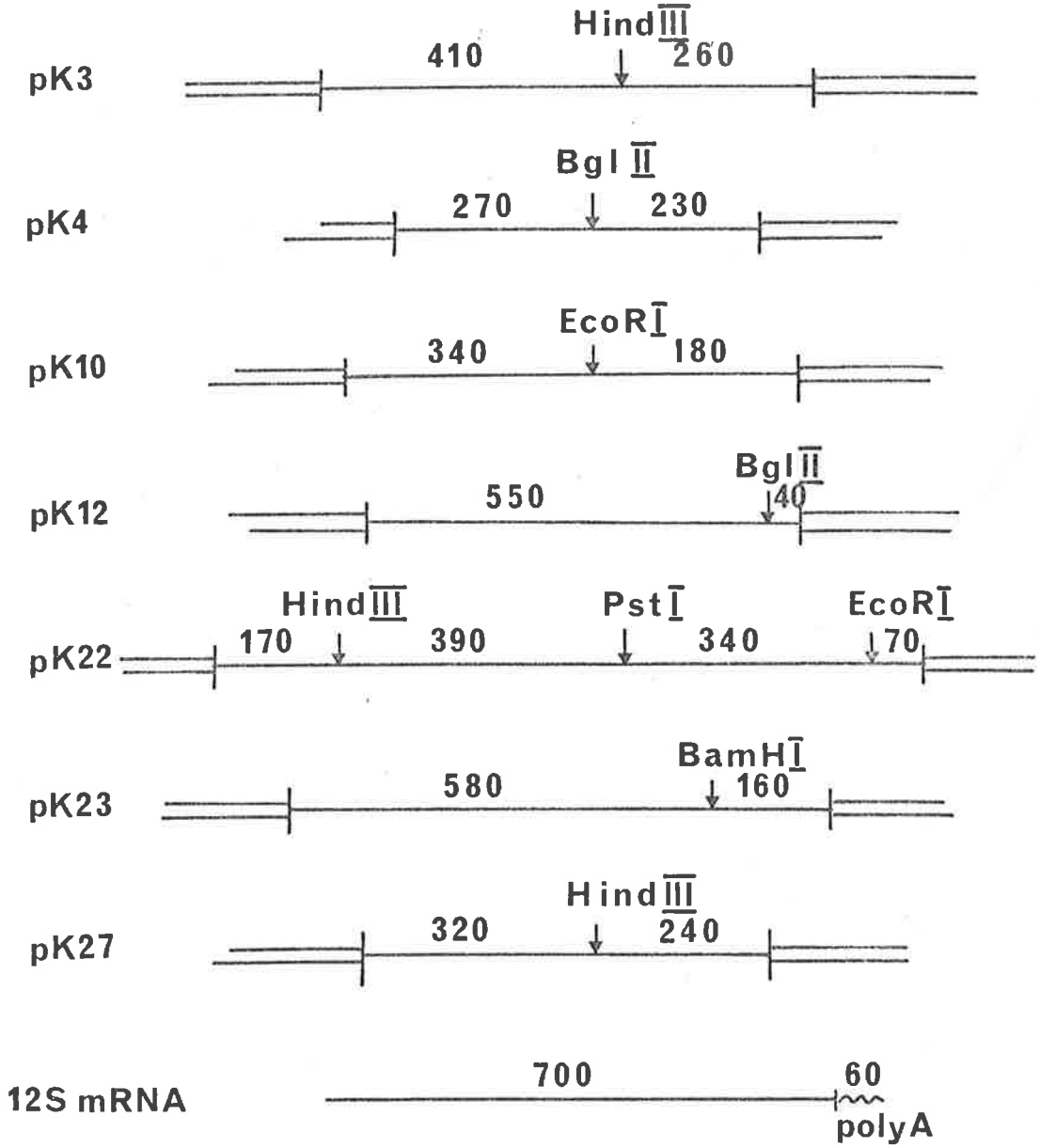
The results of annealing ^{32}P-labelled DNA of the recombinant plasmids pK3, pK10, pK12, pK15, pK22 and pK23 to *Hin*dIII digested and fractionated chick DNA are shown in Fig.4.21, together with the pattern of fragments annealing to total 12S cDNA as described in Chapter 3. The results of this experiment were unexpected but of great interest, as the recombinants used here as probes could be divided into two groups on the basis of their annealing patterns.

(i) Group I species:

The group represented by species pK12 and pK23 in Fig.4.21 exhibited the remarkable property that either of these pure species annealed to exactly the same fragments as the other as well as to exactly the same fragments as those annealing to total 12S cDNA. This is an extremely important finding, since the fact that a

FIGURE 4.20. Restriction Maps of Recombinant Plasmids.

Restriction cleavage sites within the inserts of several recombinant plasmids were mapped with respect to the terminal *Pst*I sites as described in Fig.4.18 and Fig.4.19. These are shown diagrammatically opposite. The thin line represents the inserted DNA while the double lines represent parental pBR322 DNA. It should be noted that orientation of the insert with respect to the pBR322 DNA is not known, nor is the orientation of the sequences of the inserted DNA known with respect to the mRNA sequence from which it was derived. Fragment sizes are expressed in base pairs.


pK3
Hind III
410
260
pK4
Bgl II
270
230
pK10
EcoR I
340
180
pK12
Bgl II
550
40
pK22
Hind III
Pst I
EcoR I
170
390
340
70
pK23
BamH I
580
160
pK27
Hind III
320
240
12S mRNA
700
60
polyA

single species isolated from 12S cDNA annealed to restricted chick DNA to produce the same complex pattern of fragments as that produced by total 12S cDNA under these annealing and washing conditions (2 x SSCE, 0.5% SDS 65°C) showed that these species were members of a family of homologous sequences in the chick genome. In addition the fact that the total 12S cDNA pattern of annealing was essentially identical to the pattern of annealing of each of these species suggested that the Group I sequences were more abundant than the Group II sequences. This was confirmed by annealing total 12S cDNA to equal amounts of each of the recombinant DNAs immobilized by spotting onto nitrocellulose. Figure 4.22 shows that under the high salt washing conditions (2 x SSCE, 0.5% SDS) and the low salt washing conditions (to be described below) that favour only the retention of perfect hybrids, the known Group I species pK12 and pK23 annealed considerably more 12S cDNA than did the known Group II species.

Group I thus is the more abundant of the two groups and, as is argued in the discussion at the end of this Chapter, must be sequences that code for embryonic feather keratins.

(ii) Group II species:

This second group of species is represented by pK3, pK10, pK15 and pK22. Figure 4.21 shows that these recombinants (with the exception of pK10 and pK22 which appear to be independent clones of the same mRNA species) annealed to *Hin*dIII restricted chick DNA to yield an entirely different pattern of fragments to that of

FIGURE 4.21. Detection of *Hin*dIII Digested Chick DNA Homologous To Recombinant Sequences.

Chick DNA was digested to completion with *Hin*dIII, electrophoresed in 1% agarose, denatured *in situ*, transferred and immobilized onto nitrocellulose, annealed to a pure 12S mRNA species cloned into pBR322 in 2 x SSCE at 65°C, washed in 2 x SSCE at 65°C, dried and autoradiographed. Tracks of restricted and fractionated chick DNA were annealed to ^{32}P-labelled pK3, pK10, pK12, pK15, pK22 and pK23 DNA as well as to total 12S cDNA as indicated. Fragment sizes are expressed in kb.

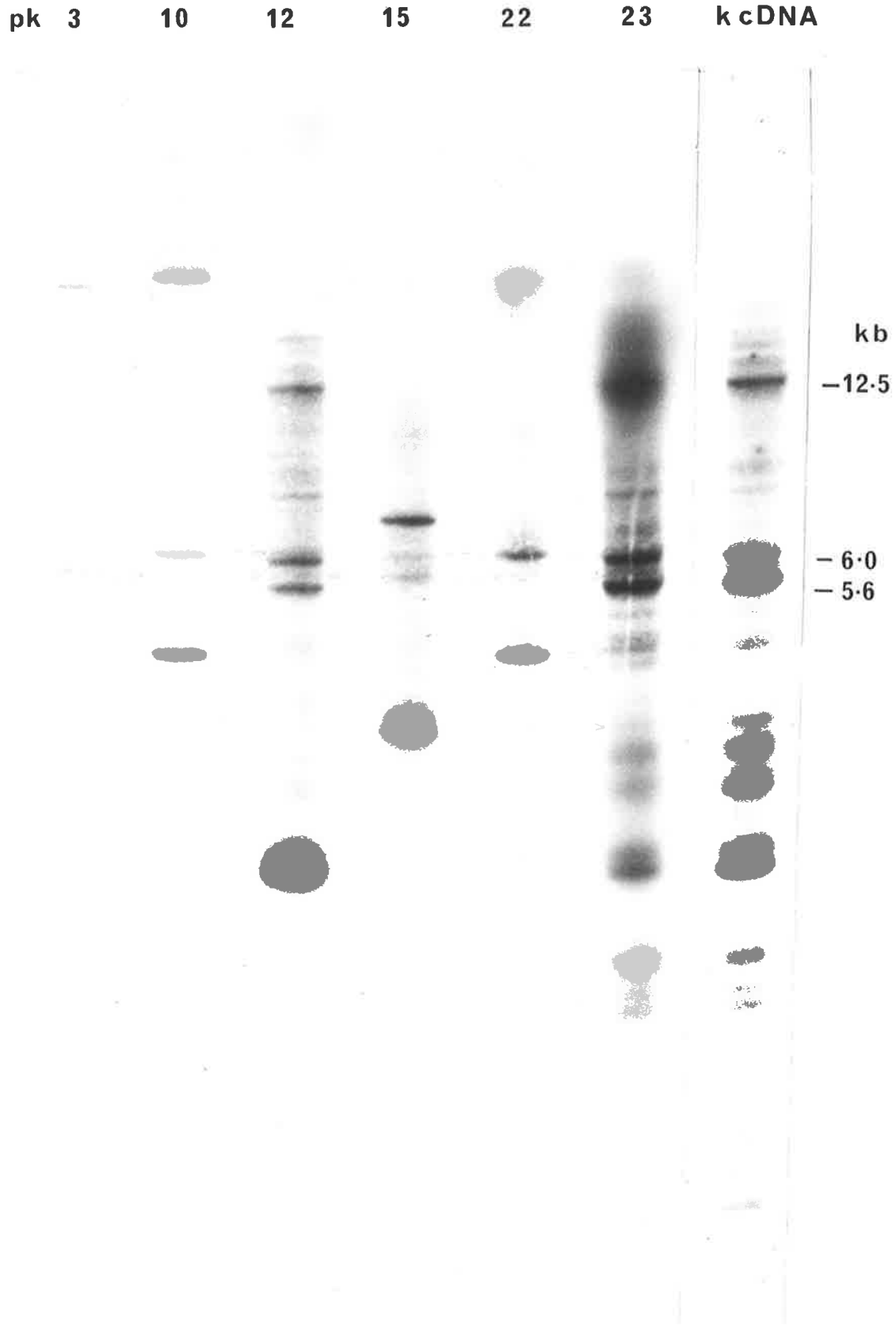

pk
3
10
12
15
22
23
k cDNA
kb
–12·5
– 6·0
– 5·6

FIGURE 4.22. Dot Analysis of Recombinant Plasmids.

10 ng of each recombinant was dotted and immobilized onto nitrocellulose filter paper and annealed to ^{32}P-labelled 12S cDNA as described in Chapter 4. Duplicate filters were then washed either in conditions of low stringency (2 x SSCE, 65°C) or of high stringency (0.05 x SSCE, 65°C) overnight before being dried and autoradiographed.

a) Shows a grid of the names of the recombinants spotted onto the nitrocellulose filter paper

b) Annealed filter washed in 2 x SSCE, 65°C overnight

c) Annealed filter washed in 0.05 x SSCE, 65°C overnight.

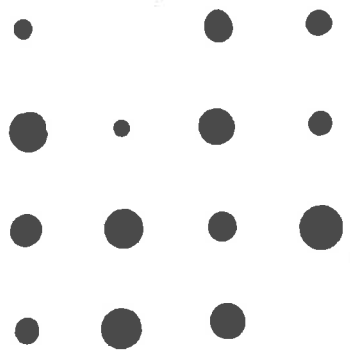
○ denotes a Group I species

□ denotes a Group II species

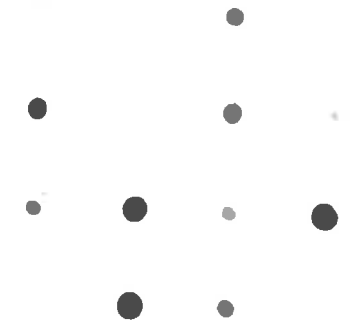
a)

3	7	9	10
12	15	17	19
22	23	25	26
27	28	29	pBR

b)



c)



either Group I species and to that of total 12S cDNA. The two fragments of approximate size 5.6 kb and 6.0 kb appear to be common to the Group I and Group II patterns but this is fortuitous, as both of the Group II fragments are marginally but definitely larger in size (i.e., have a detectably slower mobility) than the Group I counterparts. Each of the Group II species generates a different pattern from the other (again pK10 and pK22 being the exception), but certain minor bands can be seen to be in common. These common fragments suggest that this second less abundant group also is comprised of a family of homologous species.

It should be stressed that these conclusions are the result of preliminary studies of the recombinants and require much closer study to examine the nature and extent of this apparent homology. Speculation is made about the nature of the Group II sequences in the discussion at the end of this Chapter, but the examination of this group was carried no further than described above. The Group I species pK12 and pK23 were selected for further examination principally because they represented the most abundant group of species in 12S mRNA and because of the number of homologous sequences in the chicken genome.

b) The properties of Group I species:

(i) Many chick DNA sequences are homologous to Group I species:

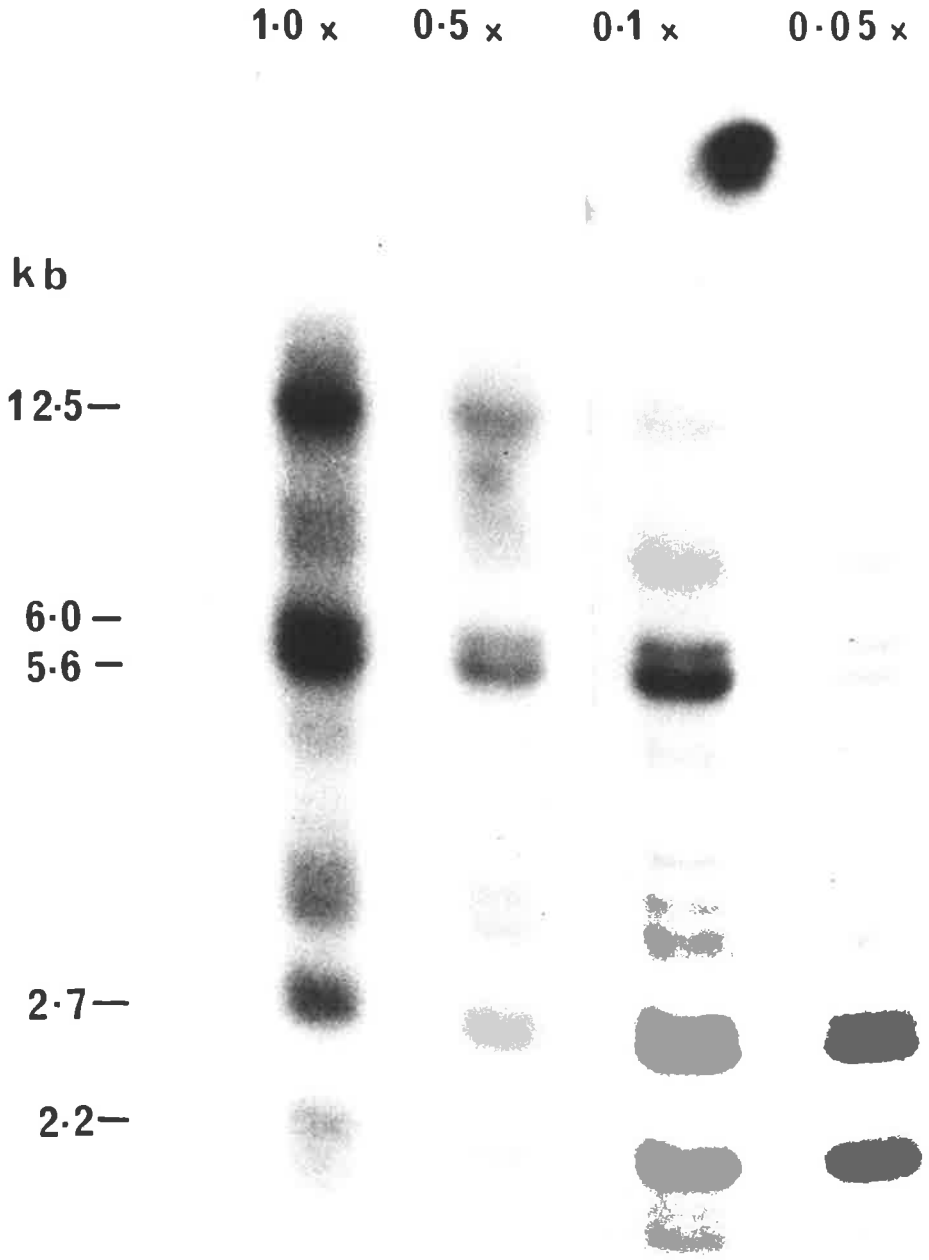
As described above, the pure Group I species, pK12 and pK23 each annealed to many different sized chick DNA restriction generated fragments. It was of interest to know whether these homologous sequences were due to the presence of many identical copies of the same

gene or to the presence of many related but non-identical sequences in the chick genome. To determine this, a series of *Hin*dIII restricted chick DNA filters were annealed to pK23 DNA and post-washed in 1 x SSCE, 0.5 x SSCE, 0.1 x SSCE and 0.05 x SSCE, i.e., in increasingly stringent washing conditions, since decreased salt concentrations decreases the stability of DNA duplexes. The rationale of this experiment is that mis-matched duplexes, being less stable than perfectly matched counterparts, will preferentially melt and thus be washed from the filter. Flavell *et al.*, (1978) used such an approach to distinguish perfect β-globin duplexes from the β - δ globin duplexes when using a cloned β-globin probe. They found that duplexes between β and δ globin sequences, the proteins of which differ in only 10 out of 146 amino acids (Dayhoff, 1972) melted preferentially on passing from 0.1 x SSC to 0.03 x SSC in the final 30 min wash. The approach in the experiments described here was slightly different in that the different salt washes were carried out at 65°C for 18 hours to ensure that any melted sequence was washed efficiently from the filter.

Figure 4.23 shows the results of this post-washing experiment. It can be clearly seen that the duplexes formed with the *Hin*dIII-chick fragments of size 2.7 kb and 2.2 kb preferentially remained while the other duplexes were preferentially lost from the filter. This result was supported by comparing the patterns of annealing of pK12 to *Bgl*II restricted DNA. Fig.4.24 shows that the duplexes formed with the *Bgl*II-chick fragments of size 2.6 kb and 1.6 kb are retained in the low salt wash, while the other

FIGURE 4.23. Post Melts of Chick DNA-pK23 DNA Duplexes.

Chick DNA was digested to completion with *Hin*dIII, electrophoresed in 1% agarose, denatured *in situ*, transferred and immobilized onto nitrocellulose filter paper and annealed to ^{32}P-labelled pK23 DNA in 2 x SSCE at 65°C. The filters were then washed for 16 hours in 1.0 x, 0.5 x, 0.1 x and 0.05 x SSCE at 65°C as indicated before being autoradiographed. The 0.05 x SSCE and 0.1 x SSCE tracks were exposed for approximately twice as long as the 1.0 x and 0.5 x SSCE washes, as even perfect duplexes are lost in the low-salt conditions, at a greater rate than in the high-salt conditions as shown by ^{32}P-labelled β-globin dscDNA-chick DNA duplexes examined under the same conditions (results not shown).


1·0 ×
0·5 ×
0·1 ×
0·05 ×
kb
12·5
6·0
5·6
2·7
2·2

duplexes are preferentially lost.

It is concluded, therefore, that many of the sequences annealing to pK12 and pK23 are homologous to them but have diverged in sequence to be non-identical. Thus pK12 and pK23 are confirmed to be members of a family of homologous sequences.

(ii) The Group I species pK12 and pK23 have heterogeneous flanking and/or intervening sequences:

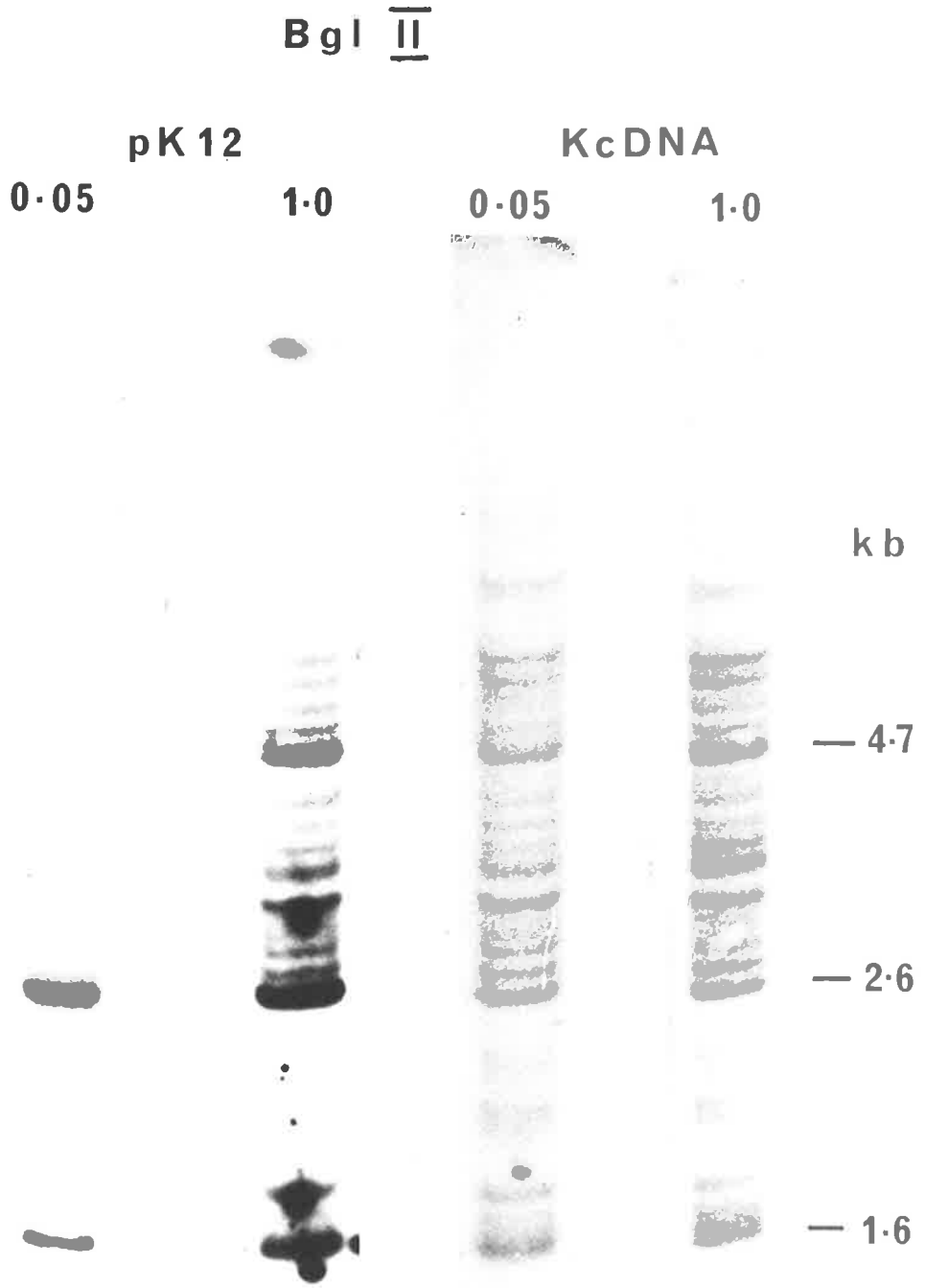
Having gained a certain amount of information about Group I species in general, the properties of the two species so far constituting the group were investigated.

As shown in Table 4.1, pK12 and pK23 have inserts of length 590 and 740 nucleotides respectively. Of the restriction endonucleases used to examine the inserts only *Bgl*II cleaved within the insert of pK12, while only *Bam*HI cleaved within the insert of pK23, suggesting that pK12 and pK23 were different species. This conclusion was confirmed by results obtained by C.P. Morris (personal communication) who used the restriction endonuclease *Hae*III to cleave pK12 and pK23. He demonstrated that pK12, carrying the shorter insert, had three cleavage sites within the insert, while the longer pK23 had only two. In addition the internal *Hae*III fragment of the pK23 insert had a different mobility to the internal fragments of the pK12 insert.

Having demonstrated that these species were different, it was of interest to determine the size of restricted chick DNA fragments that coded for each of these species.

FIGURE 4.24. High and Low Stringency Washing of *Bgl*II Digested Chick DNA-pK12 DNA Duplexes.

*Bgl*II-digested and fractionated chick DNA was immobilized onto nitrocellulose filter paper and annealed to ^{32}P-labelled pK12 DNA or 12S cDNA as indicated and as described in Fig.4.23. These filters were post-washed in 1 x SSCE or 0.05 x SSCE for 16 hr at 65°C as indicated before being autoradiographed.


Bgl II
pK12
KcDNA
0·05
1·0
0·05
1·0
kb
— 4·7
— 2·6
— 1·6

As described earlier, in the high salt washing conditions (2 x SSCE, 0.5% SDS) both of these recombinants annealed to the same restricted chick DNA fragments (Fig.4.21). The low-salt wash described above (0.05 x SSCE, 65°C overnight, Fig.4.23), however, favoured the retention of perfect duplexes and the loss of mis-matched duplexes. Thus a comparison of fragments remaining after this low salt wash was used to indicate which fragments carried perfect copies of either pK12 or pK23.

Figure 4.25 shows the results of washing *Eco*RI-digested chick DNA annealed to 12S cDNA, pK12 or pK23 in either 1.0 x SSCE or 0.05 x SSCE. Whereas the 12S cDNA pattern was relatively unchanged going from the 1 x SSCE to 0.05 x SSCE, the pK12 and pK23 patterns change considerably. The low salt wash preferentially removes pK12 duplexes with the 5.2 kb fragment, but pK23 duplexes with this 5.2 kb fragment are preferentially retained while all other duplexes are lost.

The Group I species, pK12 and pK23 therefore lie on *Eco*RI-generated restriction fragments of different sizes. This confirms the postulate made in Chapter 3 that a single gene coding for a keratin protein is not part of a sequence repeat unit, i.e., the flanking sequences (and intervening sequences, should they exist) are different for the different genes.

The above features applied also when pK12 and pK23 were used as probes to *Hin*dIII digested DNA. Figure 4.23 showed the results of washing pK23 duplexes formed with *Hin*dIII in the high or low salt conditions. Figure 4.26 shows 0.05

FIGURE 4.25. Low and High Stringency Washes of Duplexes Formed Between *Eco*RI-Digested Chick DNA and pK12 or pK23 DNA.

*Eco*RI-digested and fractionated chick DNA was immobilized onto nitrocellulose filter paper and annealed to ^{32}P-labelled pK12 DNA, pK23 DNA or 12S cDNA as indicated and as described in Fig.4.23. These filters were post-washed in conditions of low stringency (1 x SSCE, 65°C overnight) or high stringency (0.05 x SSCE, 65°C overnight) as indicated before being auto-radiographed.

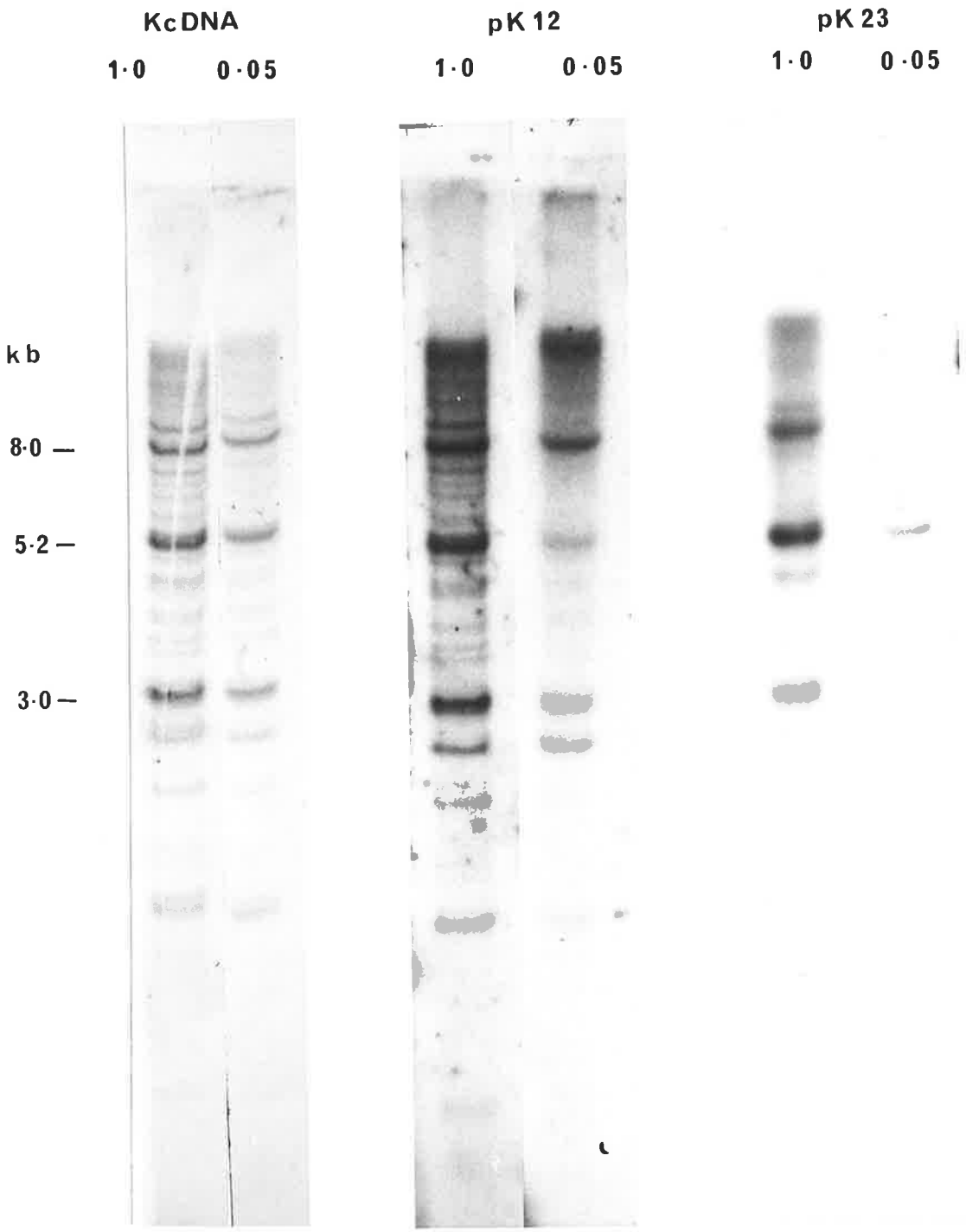
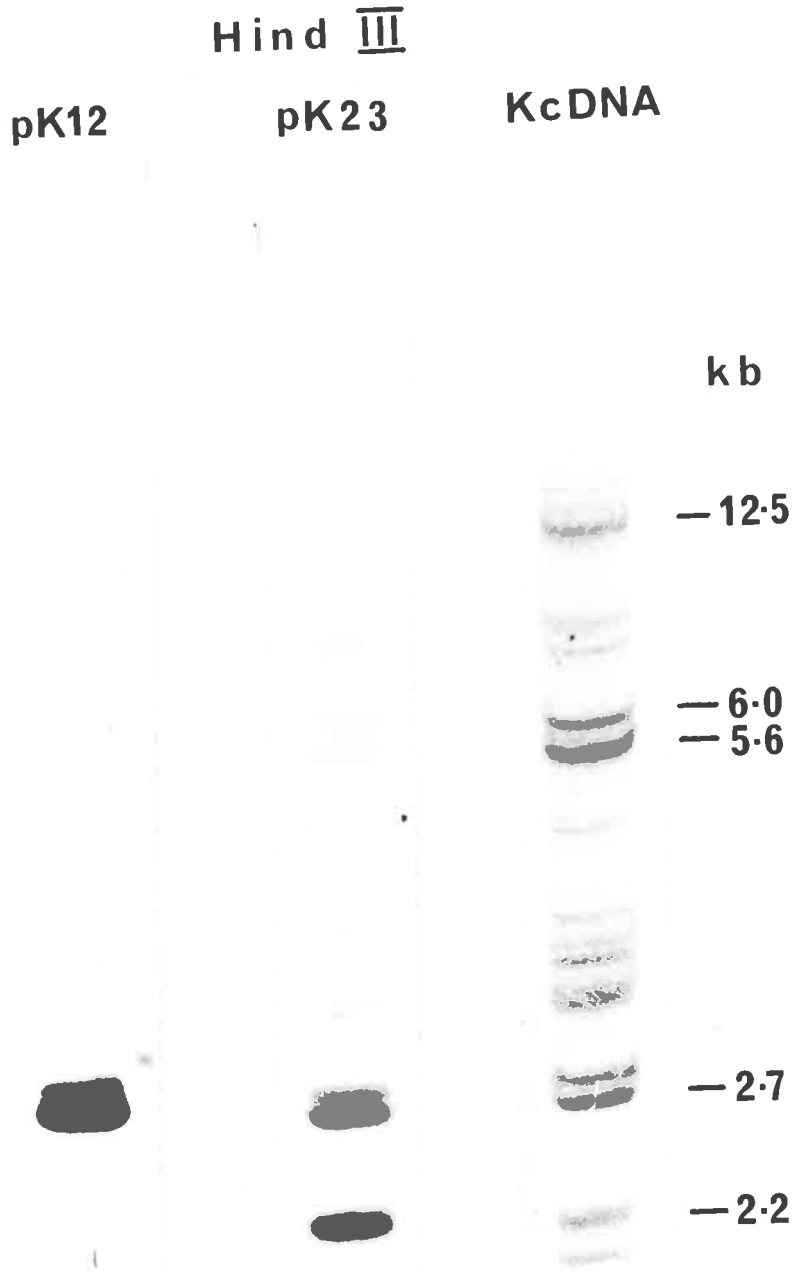

KcDNA
1·0
0·05
pK 12
1·0
0·05
pK 23
1·0
0·05
kb
8·0
5·2
3·0

FIGURE 4.26. High Stringency Washes of Duplexes Formed Between *Hin*dIII-digested Chick DNA and pK12 or pK23 DNA.

*Hin*dIII-digested and fractionated chick DNA was immobilized onto nitrocellulose filter paper and annealed to ^{32}P-labelled pK12, pK23 or 12S cDNA as indicated and as described in Fig.4.23. These filters were washed in conditions of high stringency (0.05 x SSCE, 65°C overnight) before being auto-radiographed.


Hind III
pK12
pK23
KcDNA
kb
12·5
6·0
5·6
2·7
2·2

x SSCE washes from that experiment, together with the 0.05 x SSCE wash of pK12 and total 12S cDNA duplexes formed with *Hin*dIII digested chick DNA. The 2.0 x SSCE washes of each of these probes is shown in Fig.4.21 and are essentially identical. Fig.4.26 shows again that preferential loss occurs to a much lesser extent for total 12S cDNA, while duplexes formed between pK12 and the 2.7 kb and 2.2 kb *Hin*dIII-chick fragments and between pK23 and the 2.7 kb *Hin*dIII-chick fragment are relatively stable. The significance of the observed annealing under these stringent conditions of both pK12 and pK23 to the 2.7 kb *Hin*dIII-chick fragment is discussed below.

C. Discussion.

1. Isolation of pure 12S mRNA species:

This Chapter describes the preparation of double-stranded DNA complementary to the 12S mRNA isolated from the embryonic feather and coding for the family of embryonic feather keratins. The dscDNA was cloned into the *Pst*I site of pBR322 using the complementary homopolymeric tails synthesized by terminal deoxynucleotidyl transferase to join the donor and vector DNA, followed by transformation of *E. coli* to generate individual clones carrying a single 12S mRNA species. The cloned DNA was tested for its ability to anneal to 12S cDNA using the Grunstein and Hogness (1975) technique for immobilizing the DNA of a colony to nitro-cellulose filter paper. Small amounts of recombinant plasmids were prepared by a rapid 'mini-screen' procedure and electrophoresed in 1% agarose to obtain approximate size estimates of the inserted segments.

Several recombinant plasmids were prepared and used to investigate two facets of the molecular genetics of DNA coding for keratins. The first of these concerned the proposed sequence heterogeneity of the multiple keratin proteins and the family of 12S mRNA species coding for them (Walker and Rogers, 1976b; Kemp, 1975). Restriction enzyme cleavage was used to assay the cloned sequences for sequence heterogeneity. The second facet concerned the nature of the chick DNA coding for these species. This was investigated using the cloned cDNA as a probe for complementary sequences in restricted, fractionated and immobilized chick DNA using the method described by Southern (1975).

2. 12S mRNA species are heterogeneous in sequence:

The dscDNA inserted into each of the recombinant clones was assayed for the presence of recognition sites for the restriction enzymes *Bam*HI, *Bgl*II, *Eco*RI, *Hin*dIII, *Pst*I and *Sal*I. Table 4.1 summarizes the results obtained with these enzymes. This study showed quite clearly that sequence heterogeneity exists between different 12S mRNA species, since the cloned species derived from them varied to such an extent in their susceptibility to restriction cleavage.

Interpretation of these results must, however, be subject to several qualifications. The first of these is that there may be further non-identical species in 12S mRNA that are not represented in this sample chosen for study. The second is that not all inserted DNA segments represent full length copies of the mRNA. Two identical species could thus be interpreted as being different,

simply because the inserted DNA was derived from two different regions of the same mRNA species and consequently carried different restriction cleavage sites. The third qualification is that any interpretation must take into account the results discussed below, showing that there are two distinct groups of sequences within 12S mRNA. A thorough investigation of sequence heterogeneity will therefore require a knowledge of the biological significance of each group, the categorization of cloned species into either group and estimates of the number of species within each group.

Given these limitations, however, the results clearly support the proposal by Kemp (1975) that sequences coding for keratins are heterogeneous in sequence. Of the 15 cloned species examined in this study, at least 9 of them appear to be different species on the basis of their susceptibility to restriction endonuclease cleavage.

3. Embryonic feather 12S mRNA is comprised of two groups of species:

The use of cloned species derived from 12S mRNA as probes to detect complementary sequences in restricted and fractionated chick DNA revealed the presence of two groups of sequences, termed Group I and Group II.

a) Group I species:

In the absence of sequence data or direct *in vitro* translation evidence, four properties of the Group I species and of the 12S mRNA from which they were derived, compel the conclusion that they represent true keratin-coding sequences. These properties are

that a single Group I species annealed to exactly the same restricted chick DNA fragments as did total 12S mRNA; that the majority of the duplexes formed were mis-matched and not perfect, as they were unstable in low-salt washing conditions; that the Group I species were the more abundant group of species in 12S mRNA as shown by filter annealing; and that 12S mRNA translated to yield 95% keratin proteins in a wheat-embryo cell free system (Kemp *et al* 1974b).

If the conclusion argued above is correct, keratin-coding sequences are confirmed to be a family of homologous sequences in the chick genome.

Further discussion of Group I species is given below.

b) Group II species:

Group II species are the less abundant group in 12S mRNA, since they anneal less cDNA when bound to a filter, and do not contribute to the patterns of 12S cDNA annealed to restricted chick DNA.

Group II species are also of interest, since they fit the description of a group of homologous species, as suggested by their ability to form mis-matched duplexes with certain restricted chick DNA fragments. The extent of homology and number of homologous sequences in the chick genome appear to be considerably lower than the Group I species. Nevertheless, its magnitude is sufficient to make it unlikely that these sequences correspond to contaminating mRNA species coding for the 'housekeeping' functions of the cell. It is therefore of great interest that Walker and Rogers (1976a) were able to subdivide the

proteins of the mature down feather into two distinct classes. The predominant class were the keratins, characterized by a high cystine content and constituting approximately 90% of the protein content. The other class, termed 'Fp', constituted the remaining 10% of the protein and were distinguishable from the keratins by a lack of cystine, a high proportion of tyrosine and histidine, and a lower molecular weight (8,400 compared to 10,500 - 11,500 for the keratins). This class of proteins migrated much faster than keratins on pH 2.7 acrylamide gels, hence the term 'fast protein' or 'Fp'. The 'Fp' class were shown to be heterogeneous, being comprised of at least three protein species, so that they also may represent a family of homologous proteins. Powell and Rogers (in preparation) have recently shown that 'Fp' mRNA is synthesized co-ordinately with keratin mRNA and is not fractionated from keratin mRNA by the mRNA isolation procedure used in this study. Although translation of the 12S mRNA in a wheat embryo cell-free system yielded an estimated 95% of products as keratin proteins (Kemp *et al.*, 1974c; B.C. Powell, personal communication) it is quite feasible that mRNA coding for keratin species translates more efficiently than mRNA coding for the Fp species so that the proportion of Fp mRNA could be considerably higher than the 5% predicted from the translation data. It is quite possible then, that the Group II species code for the Fp group of proteins. The possibility remains, of course, that they constitute a separate class of keratin-coding mRNA species.

4. Group I species represent a family of homologous sequences:

Recombinants pK12 and pK23, the two Group I species so far examined in detail, annealed to the same set of between 20 and 30 *Eco*RI, *Hin*dIII or *Bgl*II restricted chick chromosomal DNA fragments when the annealing and washing conditions were of sufficiently high salt to favour the formation of mis-matched duplexes (2 x SSCE, 65°C annealing; 1 x SSCE, 65°C washing). This result demonstrates that these Group I species concluded as being true keratin coding sequences are members of a family of homologous sequences in chick chromosomal DNA. Many of these sequences were shown to be homologous and not isologous (i.e., perfect copies) by washing the annealed filters in low-salt conditions that favoured the retention of stable perfect duplexes and the loss of less-stable mis-matched duplexes (0.05 x SSCE, 65°C, overnight).

These results support the proposal by Kemp (1975) that keratin-coding species are a family of sequence-related (homologous) species.

5. Sequences adjacent Group I species genes are heterogeneous:

As described above, low-salt washing conditions were used to melt preferentially the mis-matched duplexes from the filters. Duplexes which are preferentially retained must be highly stable and are likely to consist of, or at least include, perfect duplexes between the cDNA and the gene coding for the equivalent mRNA sequence. It should be noted that neither the pK12 nor the pK23 inserts

contain an *Eco*RI recognition site. The observation, then, that under stringent washing conditions, pK12 annealed to *Eco*RI restricted chick DNA fragments which were of different sizes to that of pK23 demonstrates clearly that sequences in equivalent regions adjacent to the sequences for the mature mRNA (flanking and/or intervening sequences, should they exist) must be different.

These results strongly support the proposal made in Chapter 3 that the genes coding for keratins are not organized into discrete sequence repeating units. Again it should be stressed that such results have no relevence to possible linkage relationships between the genes.

6. Are the Group I species linked?

As described in Chapter 3, the very complexity of the patterns of restricted chick DNA fragments annealing to the 12S cDNA or as described in this Chapter annealing to the Group I species pK12 and pK23, made it impossible to interpret the data in terms of gene arrangement. With the availability of the isolated Group I species it was possible, as described, to melt preferentially mismatched sequences and thus determine the sizes of restriction fragments that actually code for the particular species used as probe. It is possible, of course, that very closely related but non-identical sequences also remain under these conditions, although Flavell *et al.*, (1978) used similar procedures to differ between the very closely related β and δ globins. This possibility does not, however, interfere with the conclusion made in the preceding section nor with the following argument.

As described above, under the low-salt conditions, pK12 and pK23 annealed to completely different fragments of *Eco*RI digested chick DNA. Figure 4.26 shows, however, that both these sequences annealed to the *Hin*dIII-chick fragment of size 2.7 kb. Plasmid pK12 annealed to this fragment only, while pK23 annealed to both this and a 2.2 kb fragment. If the assumption is made that the 2.7 kb fragment annealing to pK12 and pK23 is the same fragment and not two different fragments which by chance are of the same size, then clearly the pK12 and pK23 related sequences lie on the same fragment and are therefore linked. It should be stressed that no data was obtained that supports the assumption that the 2.7 kb fragment represents a single fragment, so that this conclusion must remain tentative until analysed further by restriction analysis or molecular cloning of the fragments involved. Considering, however, that the chance of independent *Hin*dIII fragments being of identical sizes is very small, it is likely that these sequences are linked. The specific arrangement of these or any other of the Group I species on the DNA is, however, beyond the interpretation of the data presented here.

The only further conclusion that could be made, if the proposed linkage is correct, is that although regions of pK12 and pK23 related sequences are on the same 2.7 kb *Hin*dIII fragment, an *Eco*RI site must lie between them, since they are each associated with different *Eco*RI generated chick DNA fragments.

7. Do the natural Group I genes contain intervening sequences?

This topical question is impossible to answer from the data presented in this thesis. Evidence is certainly suggestive of intervening sequences, since under the stringent low-salt annealing conditions, pK12 which has no *Eco*RI site within the insert anneals to three different *Eco*RI generated chick DNA fragments, and pK23 which has no *Hin*dIII site anneals to two different *Hin*dIII generated chick DNA fragments. Interpreting these data in terms of intervening sequences, however, makes two unjustified assumptions, which fortunately are not necessary for the earlier interpretations. The first of these is that the conditions described favour only perfect duplexes. As mentioned earlier although perfect duplexes must be among the species favoured, there is no evidence that they are the only types of duplexes that remain. The second assumption is that the gene coding for either of these species is unique. If the gene was repeated, intervening sequences would be unnecessary to explain the data described above.

Thus although the data can be interpreted to imply the presence of intervening sequences in Group I genes, there is no evidence to support the two assumptions, described above, and necessary for this interpretation. Resolution of this question clearly awaits a study of the isolated natural genes themselves.

8. Other comments:

A feature of these results that merits discussion is the apparent enrichment for minor mRNA species during the preparation of the cloned cDNAs used in this study. Specifically, it is strange that the Group I species. described above, being apparently the most abundant species, are under-represented in the final clones selected for study, since 3 of the 5 clones examined in detail were from the less abundant Group II species. There are two possible reasons for this. The first is that the preparation of cloned cDNA used in this study enriched the minor fraction. It has been reported (Rougeon and Mach, 1975) that AMV reverse transcriptase synthesises the second strand of different mRNA species with different efficiencies (for example globin mRNA as against immunoglobulin mRNA). The second possibility brings this discussion to one of the major factors that shaped the course of this work. This is the fact that the preparation of the recombinant DNA was carried out in the containment facility of a host institution for subsequent characterization in the Adelaide laboratory. As the Adelaide department lacks the appropriate containment facilites for the growth of these organisms, the DNA brought back at the end of the trip was the only material available for subsequent use. This led directly to the introduction of a bias in the selection of recombinants. It was decided at the time to select a series of colonies with widely varying responses to the Grunstein and Hogness (1975) selection procedure, i.e., those which annealed different amounts of 12S cDNA. This decision was made on the basis that the different intensities could result

from differences in abundance and/or differences in extent of cross-annealing between keratin species. This may well occur to some extent, but it appears that the colonies selected for their low response were in fact Group II species since the cross-annealing of Group I species to many homologous chick DNA sequences described above suggests that every cloned Group I species may well give a dark response to the Grunstein and Hogness (1975) screening described.

It would be advantageous to use Group I or Group II sequences to screen cDNA cloned into a second unrelated vector to detect other homologous species. For this reason the dscDNA was also cloned into λ Charon 3A as described. Regrettably there was insufficient time to prepare these clones for use. The necessary conditions for such an analysis, including annealing procedures that discriminate between mis-matched and perfect duplexes, have been established, so that as soon as containment facilities are available it will be possible to screen for homologous but non-identical species and thus extend the catalogue of species which is at present so limited.

CHAPTER 5

DISCUSSION

CHAPTER 5.

GENERAL DISCUSSION.

This Chapter consists of a discussion of the results presented in Section I and is divided into three parts. The first of these considers the results in relation to previous studies of the keratin system; the second considers experiments of clarification that arise directly from the incompleted studies; and the third considers major approaches to the elucidation of the questions posed in Chapter 1 of this thesis, that arise from the results presented in this section.

A. Current Status of the Process of Keratinization in Embryonic Feathers.

1. Keratins: a multi-gene system:

The work described in this thesis strongly supports the proposal that the expression of keratin genes in the developing feather involves the co-ordinated expression of a multiple set of genes coding for feather keratins. The analysis of purified species of embryonic feather 12S mRNA coding for keratins and their use as probes to detect complementary sequences in chicken chromosomal DNA cleaved with restriction enzymes showed the existence of a group of species that annealed to many homologous sequences in the chick genome (Group I). Two different species have so far been placed into this class and a search for further species will give better estimates of the number of species of this type expressed in the

feather tissue. In addition, a second group of less abundant species, were shown to be present in 12S mRNA. Three distinct species have so far been placed into this class, but again a thorough search may reveal more species.

The significance of the results described in this thesis to the protein and mRNA studies described in Chapter 1 (in particular those described by Kemp, 1975 and Walker and Rogers, 1976a and 1976b) depends very much on the nature of these two groups of species. As argued in Chapter 4, the Group I species are most likely to be true keratin coding sequences, while the Group II species may be a further set of keratin coding sequences, or may code for the chemically different group of 'Fast' proteins or may simply be contaminating mRNA species coding for the 'housekeeping' functions of the cell (although the last of these is made very unlikely by cross-annealing properties suggesting this group is a family of homologous sequences).

Given this uncertainty, it is clear that the proposals of Kemp, (1975), and Walker and Rogers (1976b), that keratin mRNA and keratin proteins are families of homologous sequences are supported by the evidence presented in this thesis. The estimates of the number of protein and mRNA species as well as the number of genes themselves cannot be tested with the results obtained to this stage. The means of obtaining such estimates are now available and require the systematic categorization of recombinant plasmids carrying sequences derived from 12S mRNA as to which group they belong and as to whether they represent novel species or duplicate clones of species already categorized. The annealing procedures, established during the course of this

work, that can differentiate between an isologous (perfectly complementary) and homologous (similar but non-identical) species will be invaluable in extending this catalogue. Restriction enzyme cleavage and ultimately DNA sequencing will have to be used for a thorough analysis of these species.

One question that arises from the results described in this thesis is what effect the presence of the two classes of species within 12S mRNA has on the interpretation of the hybridization data of Kemp (1975). Rather than describe specific interpretations of his data, it is sufficient to say that they must be seen as a low resolution means of studying the sequences in 12S mRNA while the approach described in this thesis, made possible by the advent of recombinant DNA, permits very high resolution of the nature of keratin-coding sequences themselves. Thus the general conclusions made by Kemp (1975) are supported by this work, but it is clear that qualifications must be made as the high resolution studies yield information.

2. Evolution of keratin genes:

Interpretation of the results described in this section is hampered by a lack of knowledge of the nature of the Group I and Group II species. Group I species which are concluded to be keratin coding sequences exhibit the classical characteristics of a set of genes that have evolved by gene duplication and divergence. They are a family consisting of many gene loci that are related but not identical in sequence. No evidence was obtained for polymorphism of sequences within a given locus, so that

evolution of the many sequences and thus of the heterogeneity of keratin proteins presumably arose exclusively by gene duplication and divergence.

Group II species are markedly different and are thus likely to have evolved independently. Sequence analysis of these species will be invaluable in elucidating the nature of the Group II species.

3. Molecular genetics of keratin coding genes:

Assuming the logical conclusion that Group I species represent mRNA species coding for keratin proteins, the studies described in this thesis imply several features of chick DNA coding for keratins.

Firstly, several loci in the chick genome code for the homologous keratin species. Secondly, polymorphisms involving alleles of these loci were undetected in the studies described in this thesis. Thirdly, the sequence in the genome adjacent to those coding for the mature mRNA species differ for different species. That is, individual keratin genes expressed in the developing embryonic feather are not organized into discrete sequence repeating units.

Fourthly, although as yet tentative, evidence presented in this thesis suggests that at least two keratin genes are adjacent in the genome, since they were shown to reside on the same *Hin*dIII generated chick DNA fragment.

Finally, the particular arrangement of the keratin genes in the chick genome does not alter during the development of the organism, since the keratin genes of sperm embryonic feather and adult blood DNA were restricted to yield exactly the same pattern of fragments.

B. Future Experiments of Clarification.

The results and discussion presented in this section point to two immediate lines of experimentation that would clarify the situation dramatically. By far the most important of these is determining the nature of the Group I and Group II sequences. Thus it is essential as the next step in the study of this system, that Group I and Group II species be characterized either by DNA sequencing and a comparison of the predicted translation product with the known properties of proteins expressed in the developing feather, or by the use of hybridization to purify complementary mRNA species that could be identified by *in vitro* translation.

A second set of experiments of clarification would involve the analysis of pK12 and pK23 (the two Group I species) as probes in the low-salt washing conditions to restricted chick DNA. Should a chick DNA fragment generated by an enzyme other than *Hin*dIII form stable duplexes with pK12 and pK23 under the stringent washing conditions, the linkage proposal will be strongly supported.

C. Future Approaches.

The work reported here also leads to many longer term approaches to the elucidation of the molecular biology of keratinization.

1. The number of species expressed in the embryonic feather:

The first of these has been discussed earlier and involves a thorough characterization and categorization of the species expressed in the embryonic feather. The

approach to this will involve a three-step process. The first of these uses a characterized species as probe and a comparison of high-salt and low-salt annealing procedures to detect homologous but non-identical sequences in a random population of cloned species. The second will involve growth of the supposed novel species and characterization by length estimates and restriction enzyme cleavage analysis to confirm sequence differences. The final step may involve DNA sequencing for confirmation (if necessary) that the species is indeed novel. This process will give estimates of both the number of species expressed and of the relative proportions of these species present in the 12S mRNA, information that will be necessary for the investigation of gene expression in the developing feather.

2. Sequence analysis of cDNA species:

Apart from sequence identification described in section B above, DNA sequencing of cloned cDNA species will permit an examination of the evolutionary relationships between keratin proteins by comparison of the predicted translation products. This approach will be particularly valuable in view of the fact that direct amino acid sequence analysis is difficult to achieve because of the inherent problems in the purification of single proteins from the complex mixture of protein chains.

A comparison of sequences may also shed some light on sequences involved in the regulation of translation of the keratin mRNAs since sequences conserved in different species presumably represent biologically important regions.

It should also be noted that efforts are currently

being made to isolate scale and adult feather mRNA coding for keratins to extend the comparison, at all levels, to keratin sequences expressed in different tissues.

3. Isolation of chromosomal DNA segments carrying keratin genes:

In order to examine closely the nature and arrangement of the genes coding for embryonic feather keratins, it will be necessary to use recombinant DNA techniques to isolate chromosomal fragments that encode the keratin genes. The procedures for achieving this are available and will be attempted as soon as the appropriate containment facilities are available.

4. *In vitro* transcription analysis:

Ultimately a system will have to be established to study the interaction of potential effector molecules with the keratin genes and to assay for their effect on transcription. The biological characterization described above will provide much background information necessary for the interpretation of results gained with any such system. In addition a comparison of sequences in the chromosomal DNA adjacent to several of the co-ordinately expressed genes may point to regulatory regions that may be in common to these genes. Development of an *in vitro* transcription system using isolated chromosomal DNA segments or chromatin reconstituted from them is very likely to be the next quantal step required for the elucidation of the nature of keratin gene expression in the developing embryonic chick feather.

SECTION II

A RESTRICTION CLEAVAGE MAP OF COLIPHAGES 186 AND P2

CHAPTER 6

INTRODUCTION

CHAPTER 6

INTRODUCTION.

A. General Introduction.

The aim of the research project undertaken for this thesis was to elucidate the arrangement of the keratin coding genes in the chicken genome. For reasons given in the general introduction, however, considerable effort was devoted to establishing a restriction endonuclease cleavage map of the coliphages 186 and P2. This work could be justified solely as a basis for developing techniques of restriction endonuclease analysis of DNA that were to be later applied to keratin coding DNA. The coliphages 186 and P2 themselves are of such great interest that they further justify the work described below.

This introduction will consider the occurrence and nature of restriction endonucleases and describe various procedures for mapping the sites of cleavage of these enzymes on particular DNA species. The value of such maps to the study of viral DNA and viral functions will be discussed. Finally the nature of the coliphages 186 and P2 will be discussed in terms of the contribution to be made to our understanding of molecular biological processes by the study of these phages.

B. Restriction endonucleases.

The study of the restriction-modification properties of certain bacteria began with the observation that the host range of a given phage preparation depended on the

bacterial strain in which the phage had last propagated (Luria and Human, 1952, Bertani and Weigle, 1953). Arber and Dussoix (1962) formulated the hypothesis that this property was the result of two distinct enzymatic processes of the bacterial cell. One process acted to restrict the propagation of incoming phage while the other acted to modify the phage to be insensitive to the restriction. They demonstrated that the processes of restriction and modification acted at the level of the phage DNA, since when modified DNA replicated in a non-modifying host, the modification was maintained only in those progeny particles possessing one or both parental strands of DNA. They also demonstrated that unmodified phage DNA was rapidly degraded to nucleotides by a restricting host (Dussoix and Arber, 1962). Restriction, therefore, appeared to be a means of defense by the bacterial cell against foreign genetic material.

Meselson and Yuan (1968) used an *E. coli* K12 extract and an isolated enzyme fraction to demonstrate the *in vitro* digestion of unmodified DNA and protection of modified DNA. The digestion was shown to be by the introduction of double stranded breaks into the DNA. Linn and Arber (1968) demonstrated the same process with extracts from *E. coli* B.

The major advance in the use of restriction enzymes in the analysis of DNA came with the recognition of two distinct classes of restriction endonucleases, termed class I and class II. The class I endonucleases are represented by the *E. coli* K and B systems described above. These enzymes characteristically require S-adenosylmethionine,

ATP and Mg^{++} as cofactors and although they recognise specific sites on the DNA, the cleavage occurs at random sites (Murray *et al.*, 1973). Class II enzymes, however, require in general only Mg^{++} and cleave DNA at discrete, sequence specific sites. Smith and Wilcox (1970) and Kelly and Smith (1970) described the first of these, isolated from *Haemophilus influenzae* strain d and showed it to cleave specifically at the sequence $\begin{matrix} 5' \text{ GTPy}\downarrow\text{PuAC } 3' \\ 3' \text{ CAPu}\uparrow\text{PyTG } 5' \end{matrix}$. Since 1970 over eighty such site specific DNA endonucleases have been discovered (see Roberts, 1976 for a comprehensive review). These enzymes cleave DNA at specific nucleotide sequences of two, three, four, five or six nucleotides, depending on the particular enzyme. While the general term 'restriction' endonuclease is used for all these enzymes, it should be noted that although there is some evidence for a role in restriction-modification systems (Bron *et al.*, 1975 and Yoshimori, 1971) in most cases there is no justification for using the term 'restriction'. For example it is possible that such enzymes play a role in site specific recombination, as suggested by the results of Chang and Cohen (1977).

Thus restriction endonucleases are capable of generating fragments of DNA of discrete length and defined sequence. Mapping of the sites of cleavage determines the position occupied by each particular fragment in the original DNA molecule, enabling the information generated by the use of the individual fragments to be interpreted in terms of the linear genome.

C. Procedures for Mapping Restriction Endonuclease Sites on DNA.

1. Double digestion and fragment redigestion:

The first application of the cleavage of DNA by restriction endonucleases was by Nathans and his colleagues culminating in a restriction cleavage map of SV40 DNA with endonucleases isolated from *Haemophilus influenzae, H. parainfluenzae* and *E. coli* RY13 (Danna *et al.*, 1973). The order of fragments was determined by the redigestion of isolated partially digested fragments with the same enzyme and the identification by gel electrophoresis of the constituent fragments of that partial. The map was complete when overlapping partials accounted for all the fragments present in a limit digest. A second method involved digestion of the DNA with two enzymes, one cleaving at known positions and the other at unknown positions. Isolated fragments generated by one enzyme are then re-digested with the second enzyme. When the results of the double-digests are compared with those of the redigests, the unknown sites are mapped relative to the known sites.

2. Physical characteristics of the DNA:

A second approach to mapping restriction sites on the DNA is to exploit characteristics of the DNA itself. This is most applicable to viral genomes where various deletion and insertion mutants can be isolated and mapped by the visualization of DNA heteroduplexes by electron microscopy. The effect of various mutants on restriction fragment size is used to determine the position of fragments

in the DNA. For example, Allet and Bukhari (1975) used this approach to map restriction sites on λ DNA. Another feature that can be exploited is the occurrence of complementary single stranded ends that cohere under the appropriate conditions (Wu and Taylor, 1971). End fragments are identified by comparing digests of end-denatured DNA and end-annealed DNA. Alternative methods for the identification of end fragments is to end label the DNA using polynucleotide kinase and γ^{32}P-ATP prior to digestion or to digest the ends of the DNA prior to restriction with λ exonuclease or exonuclease III. In the first method, end fragments are those radioactively labelled after digestion (Carter *et al.*, 1975) and in the second method, end fragments are those with altered mobility on electrophoresis (Steinhart *et al.*, 1975).

3. Pulse-chase labelling of DNA:

Fragments isolated from double stranded DNA can be used as a source of primer to synthesize a neighbouring fragment using single-stranded DNA as the template (Seeburg and Schaller, 1975; Jeppesen *et al.*, 1976). When the neighbouring fragment of each primer is determined in this way, the restriction map is complete. This method is particularly suited to the mapping of single-stranded DNA bacteriophage for which a double-stranded replicative form DNA is available.

4. End-labelling and partial digestion:

Smith and Birnstiel (1976) have described a generally applicable method for mapping DNA restriction

sites. The DNA or isolated fragment of interest is end-labelled with polynucleotide kinase and γ^{32}P-ATP and digested with an enzyme that cleaves once. The resulting two fragments, each labelled at one end only, are isolated by gel electrophoresis. Partial digestion of either of these fragments with an uncharacterized set of cleavage sites results in a set of labelled partial fragments, each having the labelled end in common. The sequential size increments of the labelled set of partials defines the order of the fragments from the labelled end. since each partial consists of the end fragment plus one or more adjacent fragments.

5. Differential protection of cleavage sites:

Recently it has been shown that the action of endonucleases on certain sites can be differentially blocked by specific concentrations of DNA binding compounds such as distamycin A, actinomycin D and netropsin (Nosikov *et al.*, 1976; Nosikov and Sain, 1977). Such compounds are capable, therefore, of generating a partially digested fragment at the expense of its constituent fragments, due to the blockage of cleavage of a particular site. Neighbouring fragments are identified by the loss of certain fragments and the appearance of a fragment whose size is the sum of sizes of those fragments lost.

6. DNA transfer to nitrocellulose:

C.A. Hutchison III (unpublished results) has described a technique for mapping restriction sites that uses the transfer of DNA to nitrocellulose filter paper

by the method of Southern (1975). Unlabelled DNA fragments are electrophoretically separated on a slab gel and transferred to nitrocellulose. A radioactively labelled second digest is transferred at right angles to the first under conditions that allow hybridization of the labelled fragments to unlabelled ones bound to the nitrocellulose, but that do not allow the labelled fragments to bind to the nitrocellulose. Cross-hybridization between any one band of one dimension with any band of the other dimension will be detected by autoradiography. Thus for example the two different dimensions could be limit digests with characterized and uncharacterized enzymes respectively. Cross-hybridization localizes the unknown fragments to the position known for the other fragment. A somewhat less direct technique was developed independently during the course of this work and is presented in Chapter 8.

D. The Application of Restriction Mapping Techniques To The Study of Viral Genomes.

The most immediate advance facilitated by the discovery of type II restriction endonucleases occurred in the study of viral genomes. This was due to the fact that the genomes of many viruses (both eukaryotic and prokaryotic) are small enough to be cleaved into few fragments that can be identified and characterized by gel electrophoresis. For example, while the *E. coli* genome is cleaved by the enzyme *Eco*RI into hundreds of segments that cannot be individually separated by gel electrophoresis, the *Eco*RI generated fragments of the coliphage λ can each be separated and

identified. Since this section deals with the mapping of restriction sites on two coliphage genomes, and since the relevence of this mapping work lies in the use to which it is put, this introduction will consider general applications of restriction maps to the study of viral genomes.

1. DNA replication:

As described earlier, the great advantage of a restriction map is that it defines subsections of the genome that can be isolated and examined. This permits the localization of newly synthesized DNA, as defined by the uptake of radioactive thymidine into DNA regions. Presence or absence of labelled DNA in the various restriction fragments can then be used to follow the pattern of DNA synthesis. For example, following the elucidation of the first restriction map of SV40 (Danna *et al.*, 1973), Danna and Nathans (1972) used pulse labelling of SV40 infected monkey cells and an analysis of the distribution of label within the *Hin*dII fragments of the SV40 genome to show that SV40 replication begins at a specific site, proceeds bidirectionally and terminates about halfway around the circular molecule from the initiation point. Since then, of course, there has been an explosion of information on DNA replication derived from restriction mapping analysis.

2. DNA transcription:

The study of DNA transcription has benefited for the same reasons as those described above for DNA replication. The availability of defined sections of a DNA molecule

permits the use of RNA-DNA hybridization to detect RNA sequences transcribed from any restriction fragment. Khoury *et al.*, (1973) were the first to use this approach to map the sites of SV40 transcription during infection. Since then many such studies have been undertaken. These have been assisted by two major advances in restriction enzyme technology. Firstly, Southern (1975) described a technique for immobilizing DNA fragments separated by gel electrophoresis onto nitrocellulose filter paper and Alwyne *et al.*, (1977) described an analogous technique for covalently attaching RNA sequences to activated cellulose paper. The well characterized filter hybridization procedure (Gillespie and Spiegelman, 1965; Denhardt, 1966) can then be used to probe the immobilized DNA restriction fragments or RNA for complementary sequences to those of any particular RNA species or DNA restriction fragments respectively.

Restriction fragments can also be used to isolate RNA sequences derived from the fragment. This is achieved by specific hybridization to the DNA restriction fragment immobilized to nitrocellulose. (Prives *et al.*, 1974; Lewis *et al.*, 1976) or covalently bound to activated cellulose (Noyes and Stark, 1975). Translation of the RNA in a cell-free protein synthesizing system defines the products coded for by the DNA segment. Analysis of the products of particular restriction fragments can also be achieved using *in vitro* expression of the DNA in a linked transcription-translation system, as shown for coliphage f_1 replicative form DNA (Model *et al.*, 1975).

3. Localization of genes:

A very powerful method for the localization of genetically defined complementation units on viral DNA is by marker rescue of particular mutant loci (Edgell *et al.*, 1972; Seeburg and Schaller, 1975). The principal of this technique is that if a defective virus (i.e. one carrying a mutant in a particular gene) is accompanied in the cell by a segment of DNA carrying the wild type sequence for that gene, recombination (or mis-match repair in the case of single-stranded phage) permits the viral DNA to 'rescue' the wild type sequence and thus become fully functional. Of course if the accompanying DNA segment does not contain the sequence equivalent to that mutated in the defective virus, it will not be able to 'rescue' a functional gene and will therefore remain defective. By the use of a range of mapped restriction fragments, the various genes defined by these mutations can be located on a physical map of the viral genome.

4. DNA sequencing:

Determination of the nucleotide sequence of DNA molecules has been greatly facilitated by restriction endonuclease cleavage site mapping for two major reasons. Firstly such mapping allows the isolation and characterization of pure segments of the DNA and secondly, because cleavage is sequence specific, the isolated segments are of defined sequence and have defined end points. Both the enzymatic synthesis method of DNA sequencing (Sanger *et al.*, 1977) and the chemical degradation method (Maxam and Gilbert, 1976) rely on the availability in the DNA of defined restriction

sites to provide fragments of precise length and sequence. These techniques have been dramatically successful in the sequence analysis of viral genomes, culminating in the publication of the entire sequence of the ϕX174 genome (Sanger *et al.*, 1977) and the SV40 genome (Reddy *et al.*, 1978; Fiers *et al.*, 1978).

E. The Coliphages 186 and P2.

Bacteriophages 186 and P2 are temperate phages that infect *E. coli*. They were isolated by Jacob and Wollman (1956) and Bertani (1957) respectively and were described by Jacob and Wollman (1956) as 'non-inducible' phages as opposed to the 'inducible' phages characterized by bacteriophage λ. The 'inducible' phages (including λ) were a group with morphological and sequence homologies that were all induced from the lysogenic state by ultra-violet irradiation. Placing 186 and P2 in a different group to the 'λ' group was supported by the differences in morphology between 186/P2 and λ (Baldwin *et al.*, 1966; Yamagishi *et al.*, 1965; Wang and Schwartz, 1967) and the fact that the cohesive single-stranded ends of λ differ from those of 186 and P2 (Mandel and Berg, 1968). The cohesive ends of 186 and P2, however, differ in only 2 out of the 19 bases (Padmanabhan and Wu, 1972; Murray and Murray, 1973). The similarity between 186 and P2 was further emphasized by considerable DNA homology (Skalka and Hanson, 1972; Younghusband and Inman, 1974), the isolation of viable hybrid phage containing part of the P2 genome joined to part of the 186 genome (Bradley *et al.*, 1975) and the similarity of their genetic maps (Lindahl, 1969a; Lindahl,

1971; Sunshine *et al.*, 1971; Hocking, 1977) as summarized in Figure 6.1.

However with all these similarities, it is the differences between 186 and P2 that make these phages so interesting. As mentioned above, both these phage are temperate, i.e. they can exist in a lysogenic or dormant state where the phage genome is integrated with the *E. coli* chromosome. Although categorized by Jacob and Wollman (1956) as non-inducible, Woods and Egan (1974) demonstrated that 186 was, in fact, inducible. The terminology 'inducible' and 'non-inducible' for these two groups was therefore inaccurate, although the groups appear to represent two independent evolutionary groups of phage. Lysogenic 186 was shown to be induced by ultraviolet irradiation, mitomycin C and nalidixic acid (Woods and Egan, 1974). Lysogenic P2, however, was non-inducible. Ultraviolet irradiation did not inactivate the repressor (G. Bertani, 1968) and even with a heat-inactivated temperature sensitive repressor, P2 DNA could not excise from the *E. coli* genome (L.E. Bertani, 1968). Baldwin *et al.*, (1966) isolated a temperature sensitive repressor mutant of 186 and showed that the course of heat induction of the 186 lysogen was similar to that of a λ lysogen. Therefore 186 and P2 represent two very closely related phage that differ markedly in the nature of their respective lysogenic states.

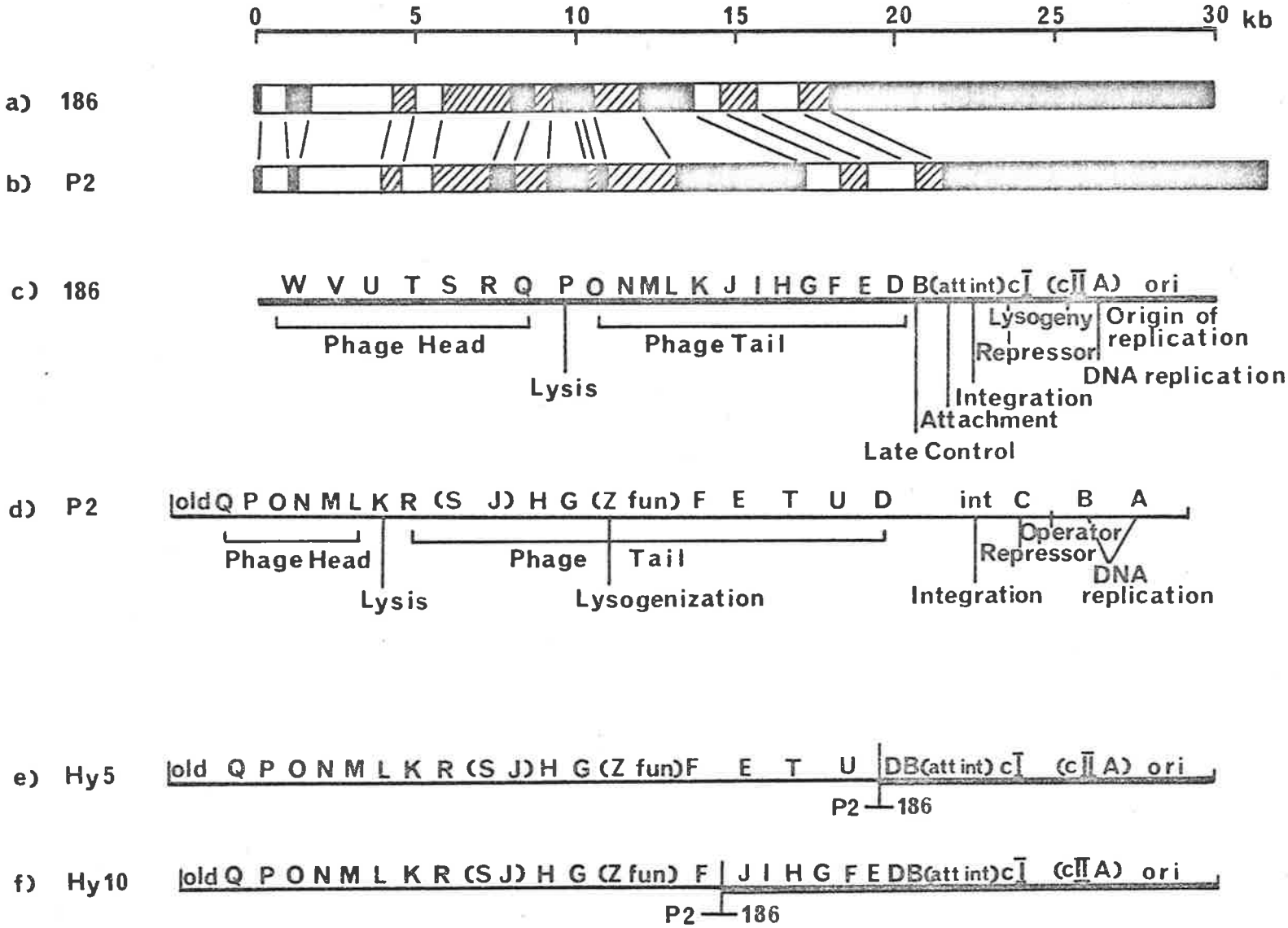
There are two major areas of molecular biology, that benefit from a study of 186 and P2. Firstly, knowledge of the processes involved in the regulation of expression of the phage genes during lytic infection will enrich our

FIGURE 6.1. Homology and Genetic Maps of 186, P2, and Hy10.

a) and b): Base sequence homology maps for 186 and P2 respectively, derived from electron microscopic heteroduplex data. The open areas represent regions of high homology, the striped areas are regions of partial homology and the filled areas are those regions which have very low homology (taken from Younghusband and Inman, 1974).

c) and d): Genetic maps for 186 and P2 respectively (Taken from Hocking, 1977 and Lindahl, 1971).

e) and f): The regions of 186 and P2 contributing to the hybrid phage Hy5 and Hy10 respectively (taken from Bradley, 1972 and Bradley *et al.*, 1975.).


0
5
10
15
20
25
30 kb
a) 186
b) P2
c) 186
W V U T S R Q P O N M L K J I H G F E D B(att int)cI (cII A) ori
Phage Head
Phage Tail
Lysis
Late Control
Attachment
Integration
Repressor
Lysogeny
Origin of replication
DNA replication
d) P2
oldQ P O N M L K R (S J) H G (Z fun) F E T U D int C B A
Phage Head
Phage Tail
Lysis
Lysogenization
Integration
Repressor
Operator
DNA replication
e) Hy5
old Q P O N M L K R (S J) H G (Z fun) F E T U DB(att int) cI (cII A) ori
P2
186
f) Hy10
old Q P O N M L K R (S J) H G (Z fun) F J I H G F E DB(att int) cI (cII A) ori
P2
186

understanding of the mechanisms used by living organisms to control their activity. A detailed understanding of the molecular processes involved are at present limited to a few bacteriophage (for example λ; Ptashne *et al.*, 1976) and bacterial operons (for example the *lac* operon, Bourgeois and Pfahl, 1976; and the *trp* operon, Brown *et al.*, 1978). Secondly, phages 186 and P2, because of the differences in inducibility, are an ideal system to study the relationship between a virus and its host, both during lytic infection, during lysogenic integration into the host chromosome and during induction of the prophage. For example the concept of the prophage state and the techniques developed for the study of temperate phages have been invaluable in the field of tumour virology and 'pilot' studies on the technically more amenable temperate phages should be profitable to studies on tumour viruses. A comparison of the processes of integration and excision (or lack of it), of host response to the presence of the prophage and of repressor inactivation and activation of the dormant phage will be facilitated by a restriction endonuclease cleavage map of the 186 and P2 genomes.

F. The Aim of the Work Described in this Section.

For the reasons given above, the aim of this study was to establish a restriction endonuclease map of the coliphages 186 and P2 in order to facilitate the elucidation of the molecular processes of DNA replication, DNA transcription and general phage expression during both the lytic cycle and the lysogenic state of 186 and P2.

CHAPTER 7

MATERIALS AND METHODS

CHAPTER 7.

MATERIALS AND METHODS.

A. Materials.

1. Bacterial strains:

The bacterial strains used in this study are described as follows:

a) *Escherichia coli* K12 strains:

W3350:F$^-$ *galK galT* (Campbell, 1965)

C600:F$^-$ *thr leu thi lacY tonA supE* (Appleyard, 1954)

ED8654:$r_k^- m_k^+$ *trpBE∇ met trpR recA*$_{56}$ *supE* (Murray *et al.*, 1977)

b) Other strains:

Bacillus amyloliquefaciens H, *Bacillus globigii* and *Streptomyces albus* G were provided by K. Murray. *Escherichia coli* RY13 was provided by J. May. *Haemophilus influenzae* Rdcom^{10} was provided by H.O. Smith.

2. Bacteriophage strains:

The bacteriophage strains used in this study are described as follows:

186*cIts*: A heat inducible mutant of 186, previously called 186p (Baldwin *et al.*, 1966)

P2*vir*1: A clear plaque mutant which is sensitive to immunity repression but which does not repress itself (Bertani, 1959)

Hy5 and Hy10: P2-186 hybrid phage isolated as described by Bradley *et al.*, (1975) and

Bradley (1972) respectively.

186*ins*2: A *c*I$^-$ mutant carrying an insertion of length 1.4 kb located 22.1 kb from the left end (Younghusband *et al.*, 1975)

186*ins*3: An *int*$^-$ mutant of 186 carrying an insertion of length 1.3 kb located 21.1 kb from the left end (Younghusband *et al.*, 1975)

186*del*1 and *del*2: Virulent phage with deletions at the operator site and *c*I gene respectively (Dharmarajah, 1975). Both phage derived from a temperature resistant mutant of 186*c*I*ts*.

λ*c*I857: A heat inducible mutant of phage λ (Sussman and Jacob, 1962)

λ728: A λ insertion vector constructed and described by Murray *et al.*, (1975)

λNN*c*I857*r*32: A derivative of λ*c*I857 carrying a copy of the IS2 insertion element (Hirsch *et al.*, 1972; Fiandt *et al.*, 1972)

λp*lac*5*c*I857S7MS505: A derivative of λ*c*I857 carrying a copy of the IS3 insertion element (Malamy *et al.*, 1977)

Strains of *E. coli* K12 C600 lysogenic for 186*c*I*ts*, λ*c*I857 and Hy5 were used in the preparation of these bacteriophage.

3. Chemicals:

All chemicals used in this study were of analytical

grade. CsCl, agarose, ^{32}P-dATP, nitrocellulose filter paper, restriction endonucleases *Bam*HI, *Bgl*II *Eco*RI, *Hin*dIII, *Pst*I and *Sal*I were all obtained as described in Section I. *Xho*I and *Xba*I were purchased from New England Biolabs Inc., Ma. U.S.A. T4 polynucleotide ligase was purchased from Miles Laboratories Inc., Elkhart, Ind., U.S.A. P.E.G.6000 was purchased from the Union Carbide Corporation.

Bacto-tryptone, Bacto-agar and yeast extract were obtained from Difco Laboratories, Detroit, U.S.A.

High titre stocks of bacteriophage were obtained from the following people: Hyl0: S.M. Hocking; 186*ins*2, 186*ins*3, 186*del*1, 186*del*2: R.M. O'Connor and J.B. Egan, λNN*c*I857*r*32, λp*lac*5*c*I857S7MS505: W. Szybalski. λ728 DNA was obtained from E.J. Finnegan.

Tris buffers were prepared from Trizma base (Sigma Chemical Co.) and the pH was adjusted with HCl unless otherwise indicated.

Scintillation fluid was as described in Section I.

4. Liquid media and buffers:

All media and solutions of chemicals were prepared in glass-distilled water and were sterilized by autoclaving for 25 min at a temperature of 120°C and a pressure of 15 lbs/inch^2. The compositions of the various media were as follows:

T broth - 2% Bacto-Tryptone, 0.5% NaCl.

L broth - 1% Bacto-Tryptone, 0.5% Yeast Extract, 1% NaCl.

LG broth - L broth supplemented with 0.1% glucose.

(The glucose was autoclaved separately as a 20% solution and added to sterile L broth).

LGC broth - L broth supplemented with 0.1% glucose and 0.0024 M $CaCl_2$. (Both the solutions of glucose (20%) and $CaCl_2$ (0.4 M) were autoclaved separately and added to sterile L broth).

TM - 0.01 M Tris pH 7.1, 0.01 M $MgSO_4$

TE - 0.01 M Tris pH 7.5, 0.001 M EDTA.

5. Solid media:

T soft agar - 1% Bacto-tryptone, 0.5% NaCl, 0.7% Bacto-agar.

Nutrient stabs - 2.5% Nutrient broth, 0.3% Yeast extract, 0.7% Bacto-agar

Plates - plates were prepared from 30 ml of the relevant mixture, dried (with lids on) at 37°C overnight and stored at 4°C until required.

T plates - 1% Bacto-tryptone, 0.5% NaCl, 1.2% Bacto-agar

LGC plates - LGC broth with 1.5% Bacto-agar.

B. Methods.

1. Storage of bacteria and bacteriophage:

Long term storage of bacteria was either in freeze-dried ampoules kept at 4°C or in 40% glycerol kept both in frozen form at -80°C and in liquid form at -15°C. Short term storage of bacteria was in nutrient stabs kept at 4°C. Bacteriophage stocks were stored over chloroform at 4°C.

2. Growth of indicator bacteria:

An overnight bacterial culture was prepared by inoculating broth (usually T broth) with bacteria from a stab and incubating with aeration in a New Brunswick gyratory water bath shaker (usually at 37°C) for approximately 16 hr. This overnight culture was used for inoculating further overnight cultures for approximately 2 weeks before being discarded. The next overnight culture was then prepared from the stab. Indicator bacteria were grown from fresh overnight cultures by diluting 1 in 30 into fresh broth (usually T broth) and incubating with aeration (usually at 37°C) in a New Brunswick gyratory water bath shaker to an A_{600} of approximately 1, measured on a Gilson flow-through spectrophotometer. This A_{600}, in a T broth culture, is equivalent to approximately 4×10^8 *cfu*/ml. The indicator bacteria were chilled and kept on ice until required.

3. Phage assays:

Phage were diluted in TM buffer, 0.1 ml samples were added to 0.2 ml samples of log phase indicator bacteria (4×10^8 *cfu*/ml) and incubated for 20 min at 37°C (for phage adsorption). 3 ml of T soft agar (held at 47°C) was added to each, mixed and then poured over an agar plate. Once the soft agar had solidified the plates were inverted and incubated at 37°C or 30°C overnight. Plaques were scored next day. For assays of 186 phage and λ phage, indicator bacteria were grown in T broth and T plates were used. For assays of phage λ the T soft agar was supplemented with 0.01 M $MgSO_4$. For assays of P2 phage and P2-186 hybrid

phage, indicator bacteria were grown in LGC broth and LGC plates were used.

4. Preparation of phage stocks:

a) Heat induction:

Phage stocks of λ*c*I857, 186*c*I*ts* and Hy5 were all prepared by heat induction of temperature sensitive lysogens. This method gave phage titres in the order of 10^{10} *pfu*/ml.

Overnight T broth cultures of the lysogens were diluted 1 in 30 into fresh LG broth and grown with aeration at 30°C to an A_{600} = 0.8. The cultures were transferred to 47°C for 10 min and then incubated at 37°C until lysis was complete, or for 4 hr if lysis did not occur. Bacterial debris was removed by centrifugation (10 min, 7800 g, 4°C) and the supernatants stored over chloroform at 4°C.

b) Liquid infection:

The stock of P2*vir*1 phage available for the preparation of phage stocks had been prepared from the non-modifying host *E. coli* C. It was anticipated that the P2 *vir*1 DNA used for these studies may have been required for transfection into a restricting host, for example an *E. coli* K12 strain. For this reason the P2*vir*1 phage were passaged through *E. coli* K12 before the phage stock was prepared by liquid infection as described below.

The unmodified P2*vir*1 were plated onto an *E. coli* K12 C600 lawn on LGC plates and were found to plate with an efficiency of 10^{-2} compared to the non-restricting *E. coli* C strain. A single plaque was selected and placed into 15 ml of LGC medium. An overnight T broth culture of

E. coli C600 was diluted 1 in 10 into LGC broth and grown to an A_{600} of 1.0. 0.2 ml of this culture was then added to the 15 ml of LGC containing a single plaque. The culture was incubated at 37°C until lysis occurred as indicated by a sharp drop in the A_{600}. 0.3 ml of 0.1 M EDTA were added to this culture immediately upon lysis. This small stock was assayed and contained approximately 10^{10} *pfu*/ml.

20 ml of a fresh overnight of *E. coli* C600 grown in T broth was added to 500 ml of LGC. The culture was grown to an A_{600} = 0.8 (approximately 3 x 10^8 bacteria/ml). P2*vir*1 from the small stock described above were added to the culture at a multiplicity of infection of 0.1. The culture was incubated until lysis occurred about 180 min later (as indicated by sharp drop in the A_{600}). 30 ml of 0.1 M EDTA were added to the 500 ml culture and the bacterial debris removed by centrifugation as described above. It should be noted that the pH of the LGC was 7.2. A pH significantly higher than this was found to inhibit infection, presumably due to the precipitation of Ca^{++}.

c) Plate stocks:

A single plaque of the desired phage was taken from a plate with a pasteur pipette, added to 1 ml of T broth and left for 30 min at room temperature to allow the phage to leach out of the agar. 0.1 ml of this solution was plated with 0.2 ml of a log phase culture (3 x 10^8 *cfu*/ml) of the appropriate host bacteria and incubated at 30°C or 37°C for 8 hr. The plate was chilled at 4°C for 30 min and then flooded with 3 ml of

ice-cold TM and left at 4°C overnight. The phage solution was removed with a pasteur pipette and stored over chloroform at 4°C.

5. Preparation of bacteriophage DNA:

Preparation of DNA from all phage stocks described above was carried out as follows:

Phage stocks prepared as described above were precipitated by the addition of polyethylene glycol (PEG6000) to 10% (w/v) and NaCl to 0.3 M. Phage were chilled at 4°C overnight, pelleted by centrifugation (10 min, 8000 g, 4°C) and resuspended in TM buffer. Solid CsCl was added to a final density of 1.45 gm/ml and centrifuged in a Ti50 rotor and Spinco Model L ultracentrifuge at 32,000 r.p.m., 4°C overnight.

The bacteriophage band was removed by side puncture and dialyzed against TM to remove the CsCl. The phage were extracted several times with water saturated phenol until no material was visible at the interface and then dialyzed extensively against TE to remove any remaining phenol. The DNA was stored at 4°C in TE.

6. Restriction enzyme digestion, gel electrophoresis, fragment size determination and end-fragment determination:

The preparation of restriction endonucleases, conditions for cleavage of the DNA, gel electrophoresis in 1% agarose, transfer to nitrocellulose and filter annealing with a preparation of DNA labelled *in vitro* are all described in Section I of this thesis, with the exception of *Xba*I and *Xho*I digestions, which were

carried out at 37°C in 10 mM Tris pH 7.5, 10 mM $MgCl_2$, 10 mM 2-mercaptoethanol and 150 mM NaCl.

The sizes of 186 and P2 restriction fragments were determined by co-electrophoresis of λ *Eco*RI and λ *Hin*dIII digests in adjacent tracks to those shown in the diagrams and by assuming the 186 and P2 genome size estimates of 30 and 32 kb respectively (Younghusband *et al.*, 1975). Fragments of size 0.5 kb or less will not be detected using the gel system described above so that two restriction sites less than 0.5 kb (1.5%) apart may be interpreted as a single site by the above mapping techniques.

In order to determine terminal fragments, the 20 µl digestion mix was incubated at 80°C for five minutes, rapidly chilled and kept on ice until loading. End denaturation and reannealing were never complete, but relative changes in intensity for the fragments involved readily identified them. Objective evidence from densitometer tracings of photographs using a Joyce Loebl double beam recording densitometer was used to confirm the assignments (not presented).

7. Isolation and redigestion of restriction fragments:

Digested DNA fragments to be redigested with a second restriction endonuclease were electrophoresed on a 20 cm x 20 cm x 0.3 cm slab gel apparatus using 1% agarose (Sea Plaque, Marine Colloids Inc.). Bands were visualized by briefly staining with ethidium bromide and were cut out with a sterile blade. The agarose was melted by heating to 70°C for 5 min and cooled to 37°C with the agarose remaining fluid. Digestion buffer and the second restriction

enzyme were added and the mixture incubated at 37°C until digestion was complete as judged by gel electrophoresis. The sample was then applied to a gel, allowed to set, and electrophoresed as described above.

8. Construction of a λpl86 phage:

The methods described in this section are based on those described by Murray *et al.*, (1977).

A mixture of λ728 DNA and 186*cIts* DNA was digested to completion with *Hin*dIII as described in Section I. A 50 µl ligation incubation was prepared containing 0.7 µg each of the *Hin*dIII digested λ728 and 186*cIts* DNA, 0.066 M Tris (pH 7.5), 0.01 M $MgCl_2$, 0.01 M DTT, 1 mM EDTA, 0.1 mM ATP and 0.5 units of T4 DNA ligase (one unit converts 1 nmole ^{32}P-phosphomonoesters into a phosphatase resistant form in 20 min at 37°C). The ligation mixture was incubated for 6 hr at 10°C and stored at 4°C for 24 hr before use. The ligated DNA was transfected into an *E. coli* C600(186*cIts*) host as described in Section I. The ligated DNA gave a 20 fold higher yield over the background of restricted, non-ligated DNA. Plaques to be studied (as described in Chapter 9) were taken off with a pasteur pipette into 1 ml of TM over chloroform and purified by plating as described in this Chapter. When all plaques on a plate exhibited the correct plaque morphology (i.e. clear plaque) a single plaque was picked and phage stocks were prepared by the plate stock method. DNA was prepared as described.

CHAPTER 8

RESULTS

CHAPTER 8.

RESULTS.

A. Introduction.

The DNA of coliphages 186 and P2 can be isolated as single full length molecules of size 30 kb and 32 kb respectively (Younghusband *et al.*, 1975). This chapter describes the mapping of sites of cleavage of the restriction enzymes *Bam*HI, *Bgl*II, *Eco*RI, *Hin*dIII, *Pst*I, *Sal*I, *Xba*I and *Xho*I on the DNA of these coliphages. Also described is the analysis of certain insertion and deletion mutants of 186 and the molecular cloning of a 186-*Hin*dIII fragment into an unrelated λ vector.

Three different approaches to the mapping of the sites of cleavage of the enzymes described above were used in this study. Each involved the analysis of restriction endonuclease generated fragments by electrophoresis in 1% agarose as described in Chapter 7. The rationale of each of these approaches is as follows:

1. The use of two properties of the coliphages 186 and P2:

   a) Complementary single-stranded ends of 186 and P2 DNA:

Linear DNA isolated from both 186 and P2 have short complementary single-stranded ends that can cohere to produce a circular molecule (Murray and Murray, 1973). It is possible, therefore, to identify the end fragments by comparing digests from DNA in the circular and linear forms. The end fragments will be present in a

digest of linear DNA, but will cohere in the circular form to produce a single fragment whose size is the sum of the end fragment sizes.

b) Viable hybrid 186-P2 phage:

The phages 186 and P2 are sufficiently related to form viable hybrid phage consisting of a section of the P2 genome combined with a section of the 186 genome (Bradley *et al.*, 1975). Two of these hybrids were used in the present study: Hy5 consisting of the left 70% (22.4 kb) of the P2 genome and the right 35% (10.5 kb) of the 186 genome (Younghusband *et al.*, 1975) and Hy10, consisting approximately of the left 50% of P2 (16 kb) and the right 50% (15 kb) of the 186 genome (Hocking, 1977). By comparing the digests of 186 and P2 DNA with those of the hybrids it was possible to determine which of the fragments derive from the left and which from the right of the two phages, since all the Hy5 or Hy10 DNA fragments (excluding the fragment with the P2-186 join) were present in either 186 or P2 digests. Those fragments common to 186 and Hy5 (Hy10) must be located in the right 35% (50%) of 186 while those fragments common to P2 and Hy5 (Hy10) phage must be located in the left 70% (50%) of P2. A possible dilemma might have arisen if the P2-186 join co-electrophoresed by chance with one of the P2 or 186 fragments, but fortunately this did not occur in the work reported here.

2. The use of double digests and redigests of fragments:

186 and P2 DNA were digested with two restriction enzymes or isolated restriction fragments were digested

with a second enzyme in order to locate certain restriction endonuclease cleavage sites.

3. The use of partial digests and annealing to a characterized fragment:

A mapping method was developed that involved annealing a ^{32}P-labelled fragment to a partial digest of DNA. The rationale for this method is that the size of any incompletely digested fragment (a 'partial') must be the sum of the sizes of the constituent fragments identifiable by gel electrophoresis of a complete digest. If one fragment can be identified as included in the partial, then the size of the partial must be the sum of the sizes of this fragment and of one or more adjacent fragments. Unfortunately a partial digest itself is usually far too complex to allow such interpretation, so gel electrophoresis, the Southern transfer technique (Southern, 1975) and annealing with a pure ^{32}P-labelled fragment was used to identify the set of partials containing that particular fragment. The neighbouring fragments to that used as probe can then be identified by correlating the size increments of the set of partials with the increments permissable from the fragment sizes of a complete digest.

B. Results.

1. Mapping from end-renaturation and 186-P2 hybrid phage fragment comparisons:

Figures 8.1 to 8.8 show the pattern of fragments obtained when the DNA of phages 186, P2 and their hybrids Hy5 and Hy10 were digested with *Bam*HI, *Bgl*II, *Eco*RI, *Hin*dIII, *Pst*I, *Sal*I, *Xba*I, and *Xho*I respectively. These figures also

include samples that were end-denatured as described in Chapter 7 thereby permitting the determination of the terminal fragments. These data yielded a great deal of information about the sites of cleavage on 186 and P2. This information was sufficient to completely map both P2 and 186 with respect to *Bgl*II, *Eco*RI, *Hin*dIII, *Xba*I and *Xho*I, and P2 alone with respect to *Pst*I and *Sal*I. It also mapped some of the sites of cleavage with the other enzymes. This information is detailed below and the results summarized in Fig.8.15(a) for 186 and Fig.8.16(a) for P2.

a) *Bam*HI:

186 DNA is cleaved seven times by *Bam*HI to give fragments of size 11.2, 6.5, 3.7, 3.0, 2.2, 1.3, 1.2 and 0.9 kb (Fig.8.1). The 1.2 kb and the 3.0 kb fragments reannealed to form a 4.2 kb fragment and were therefore the end fragments. The 1.2 kb fragment was present in Hy5 DNA, so must have been present in the right 35% of 186 and was therefore the right hand end fragment. The 3.0 kb piece was therefore the left hand end fragment. The 11.2 kb fragment present in the Hy10 digest was the 186 11.2 kb fragment (since the fragment containing the P2-186 join was clearly the 5.7 kb fragment) and as it was the only 186 fragment present besides the terminal 1.2 kb fragment, it must be adjacent to this fragment. No other sites of cleavage could be mapped by this procedure. However, it did establish sites of cleavage 10%, 59% and 96% from the left hand end of the 186 genome.

P2 DNA was cleaved three times by *Bam*HI to yield

FIGURE 8.1. *Bam*HI Digestion of 186 and P2 DNA.

186, Hy5, Hy10 and P2 DNAs were digested to completion with *Bam*HI, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks are as follows:

A: 186 DNA, end-denatured as described in Chapter 7.

A': 186 DNA

B: Hy5 DNA, end-denatured

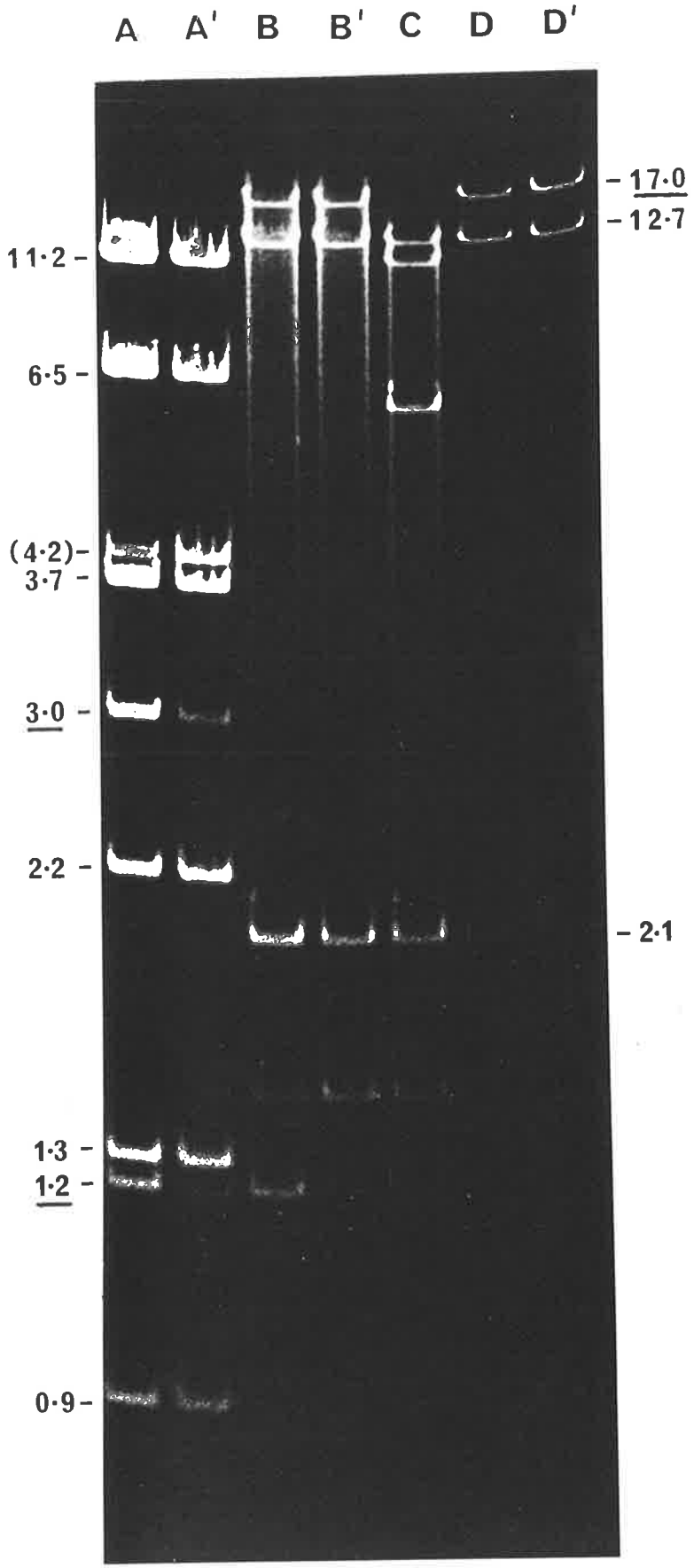
B': Hy5 DNA

C: Hy10 DNA

D: P2 DNA, end-denatured

D': P2 DNA

In Figures 8.1 to 8.8 the sizes of fragments expressed in kb were estimated from *Eco*RI-λ DNA and *Hin*dIII-λ DNA fragments co-electrophoresed in the adjacent tracks. In these figures, sizes of the end fragments of 186 and P2 DNA are underlined and the fragment size of the 'reannealed' end fragments is in brackets.


A
A′
B
B′
C
D
D′
17·0
12·7
11·2
6·5
(4·2)
3·7
3·0
2·2
2·1
1·3
1·2
0·9

fragments of 17.0, 12.7, 2.1 and 0.2 kb (Fig.8.1). The 0.2 kb fragment was inferred from the fact that in Hy5 the 1.3 kb right hand end fragment derived from 186 reannealed to give a 1.5 kb fragment. A 0.2 kb fragment must therefore have been the other end fragment, i.e., the left hand end fragment of P2. The 17.0 kb fragment was the right hand P2 end fragment, since both the 12.7 and 2.1 kb fragment were present in Hy5 and Hy10 and were therefore in the left half of P2 DNA. There was insufficient information to order the 12.7 and 2.1 kb fragments by this method, but it does map *Bam*HI sites 0.5% and 47% from the left hand end of P2 DNA.

b) *Bgl*II:

186 DNA was cleaved once with *Bgl*II to produce two fragments of size 5.7 and 24.3 kb (Fig.8.2). The 5.7 kb fragment was present in Hy5 and therefore must have been the right hand fragment. The site of *Bgl*II cleavage of 186 was therefore 81% from the left hand end.

P2 DNA was cleaved twice by *Bgl*II to yield fragments of 16.8, 7.8, and 7.4 kb (Fig.8.2). The end fragments were the 7.4 and 16.8 kb fragments. The 16.8 kb fragment was present in Hy5 and was therefore the left hand end fragment, leaving the 7.8 kb fragment as the middle one and the 7.4 as the right hand fragment. The sites of *Bgl*II cleavage of P2 DNA were therefore 53% and 77% from the left hand end.

c) *Eco*RI:

186 DNA was cleaved three times by *Eco*RI to yield four fragments of length 23.6, 3.3, 2.4 and 0.7 kb

FIGURE 8.2. *Bgl*II Digestion of 186 and P2 DNA.

186, Hy5 and P2 DNAs were digested to completion with *Bgl*II, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks were as follows:

A: 186 DNA end-denatured as described in Chapter 7

A': 186 DNA

B: Hy5 DNA end-denatured

B': Hy5 DNA

C: P2 DNA end-denatured

C': P2 DNA

In Figures 8.1 to 8.8 the sizes of fragments expressed in kb were estimated from *Eco*RI-λ DNA and *Hin*dIII-λ DNA fragments co-electrophoresed in the adjacent tracks. In these figures, sizes of the end fragments of 186 and P2 DNA are underlined and the fragment size of the 'reannealed' end fragments is in brackets.

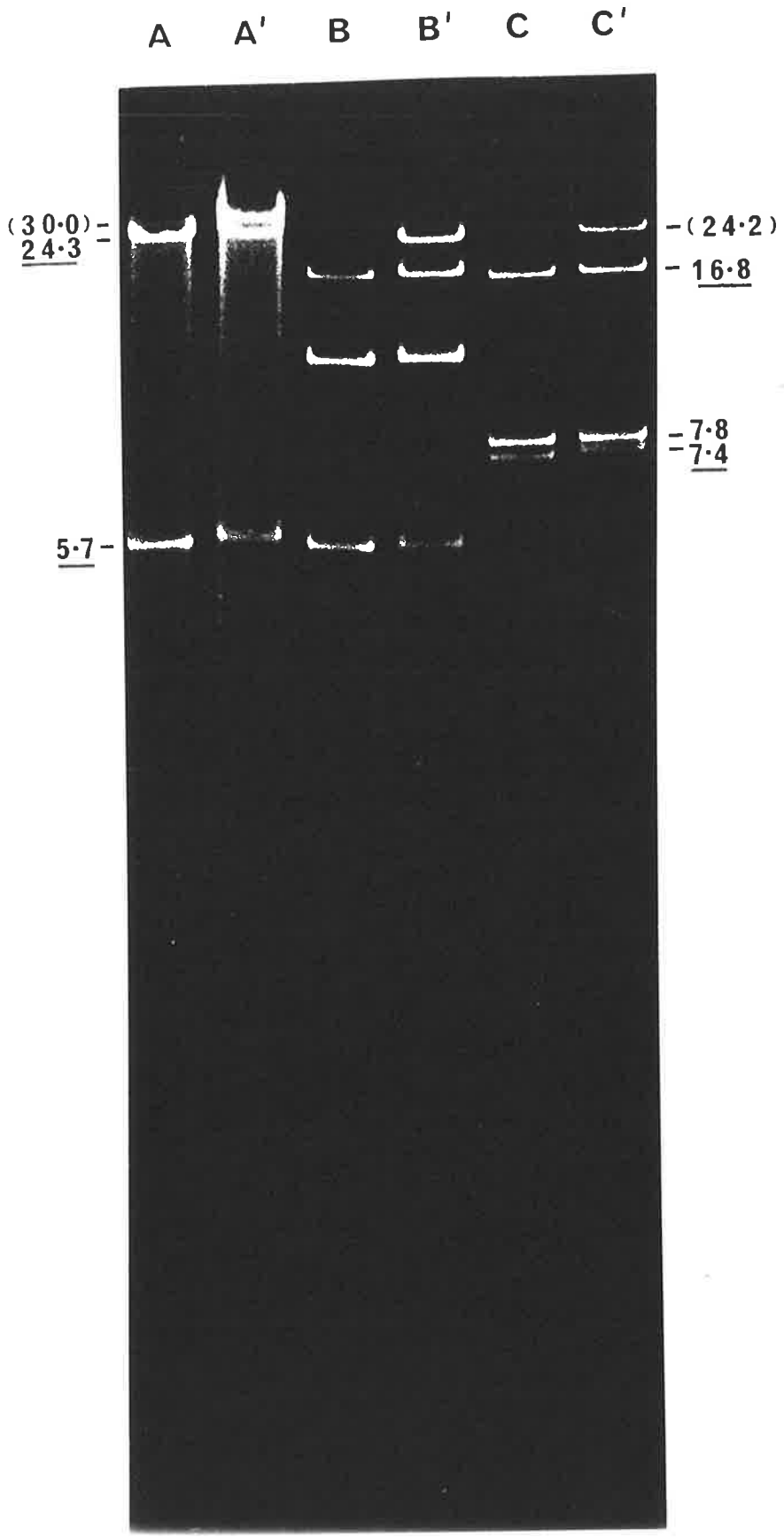
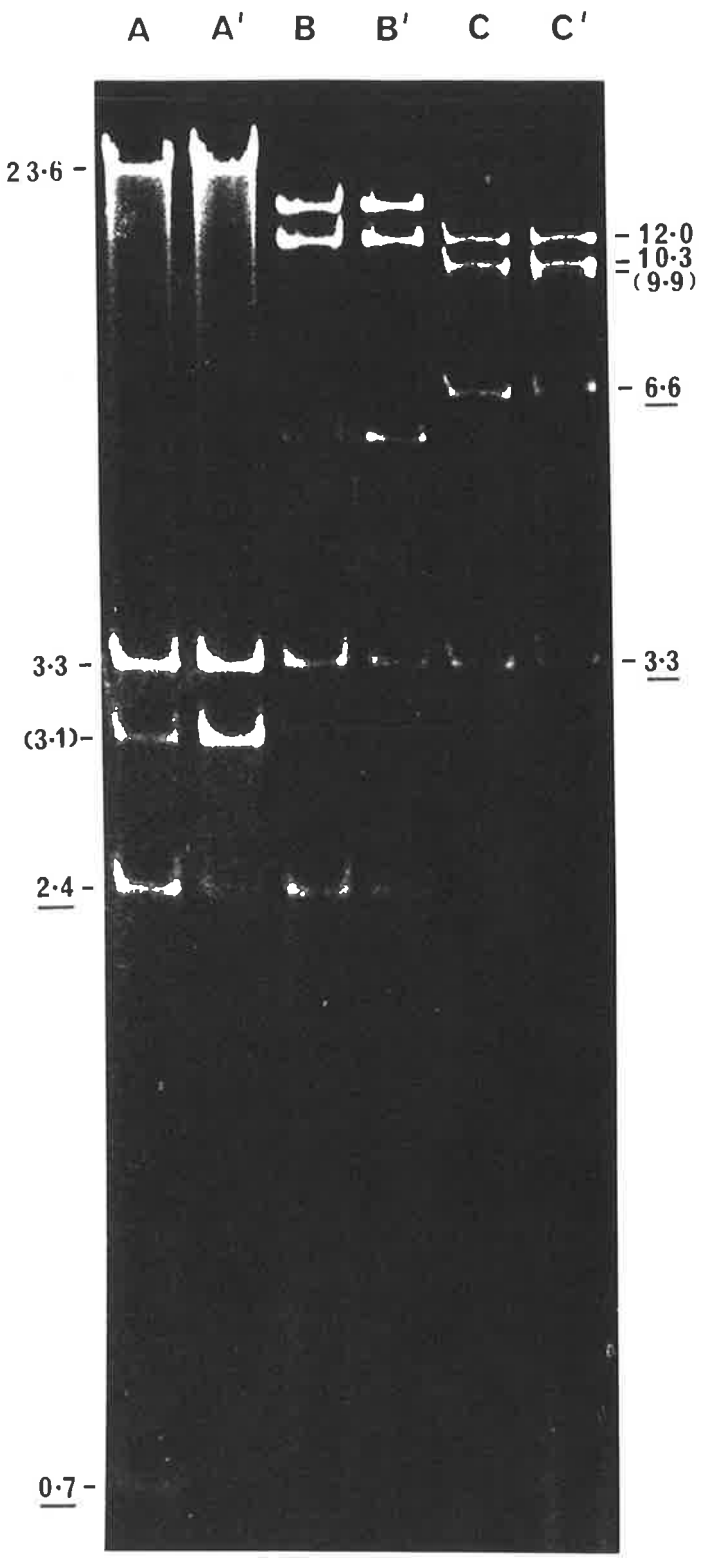

A
A'
B
B'
C
C'
(30·0)
24·3
(24·2)
16·8
7·8
7·4
5·7

FIGURE 8.3. *Eco*RI Digestion of 186 and P2 DNA.

186, Hy5 and P2 DNAs were digested to completion with *Eco*RI, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks etc., are as described for Figure 8.2.


A
A′
B
B′
C
C′
23·6
12·0
10·3
(9·9)
6·6
3·3
3·3
(3·1)
2·4
0·7

(Fig.8.3). The end fragments were the 2.4 and 0.7 kb fragments of which the 2.4 kb was present in Hy5 and was therefore the right hand end piece. The fragments of length 23.6 and 3.3 were the two internal fragments. If the 3.3 kb fragment was adjacent to the 2.4 kb end fragment, it would be present in Hy5 which contains the right 35% (≡ 10.5 kb) of 186. As it was not, then it must have been adjacent to the 0.7 kb fragment. The fragment order must therefore have been 0.7, 3.3, 23.6, 2.4 and the sites of *Eco*RI cleavage of 186 DNA were 2%, 13% and 92% from the left hand end.

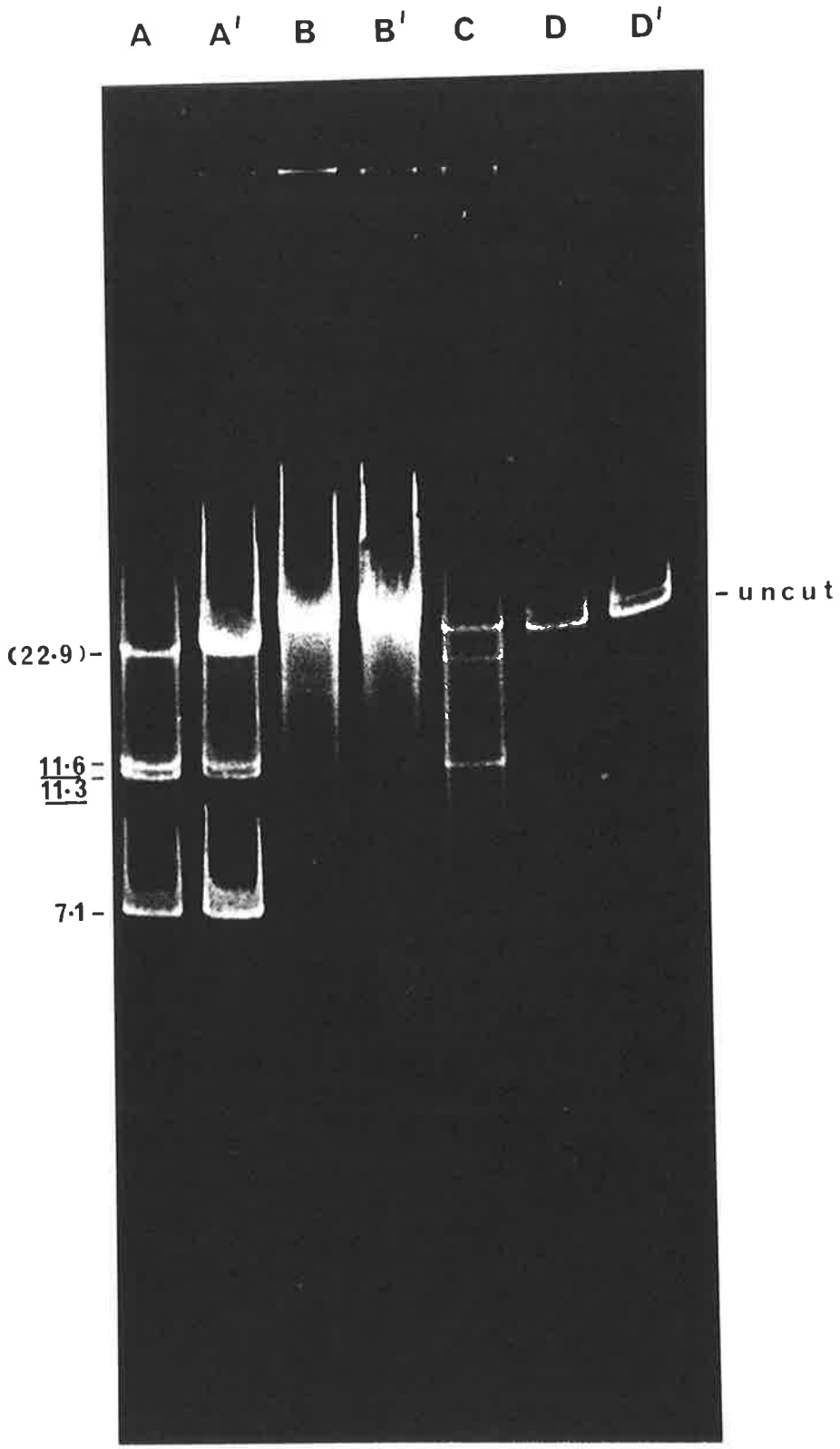
P2 is also cleaved three times by *Eco*RI to produce fragments of size 12.0, 10.3, 6.6 and 3.3, of which the 6.6 and 3.3 kb fragments were the ends (Fig.8.3). The presence of the 3.3 kb fragment in Hy5 showed that it was the left hand end fragment and the 6.6 kb fragment was therefore the right hand end fragment. The 12.0 kb fragment was also present in Hy5 showing that it was also in the left half of P2. The order of fragments was therefore 3.3, 12.0, 10.3, 6.6 and the sites of *Eco*RI cleavage of P2 10%, 48% and 79% from the left hand end, in agreement with those reported by Chattoraj *et al.*, (1977).

d) *Hin*dIII:

186 DNA cleaved with *Hin*dIII yielded fragments of size 11.3 kb, 11.6 kb and 7.1 kb (Fig.8.4). The 11.6 and 11.3 kb fragments reannealed to give a 22.9 kb fragment and the 11.6 kb fragment was on the right, as it was present in Hy10. The 11.3 kb fragment was therefore the left hand end fragment, leaving the 7.1 kb fragment as the central

FIGURE 8.4. *Hin*dIII Digestion of 186 and P2 DNA.

186, Hy5, Hy10 and P2 DNAs were digested to completion with *Hin*dIII, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks etc., are as described for Fig.8.1.


A
A′
B
B′
C
D
D′
– uncut
(22·9) –
11·6 =
11·3 =
7·1 –

one. The order of fragments was thus 11.3, 7.1, 11.6 and *Hin*dIII cleaved 186 DNA 38% and 61% from the left hand end.

P2 DNA was not cleaved by *Hin*dIII (Fig.8.4).

e) *Pst*I:

186 DNA was cleaved 22 times by *Pst*I, of which only 5 sites lay in the right 35%, i.e., were present in Hy5 (Fig.8.5). The fragments were 3.3, 2.3, 2.1, 1.9 and 1.0 kb of which the 1.9 kb fragment was the end fragment (deduced from the Hy5 pattern). As the prime objective of this mapping was to facilitate studies of the control region of 186 it was decided to map only those sites in the major control region, i.e., the region of 186 present in Hy5. However, the site of cleavage 1.9 kb from the right hand end and the 0.2 kb left hand end fragment inferred from the 2.1 kb annealed fragment, were the only two sites of cleavage mapped by these data.

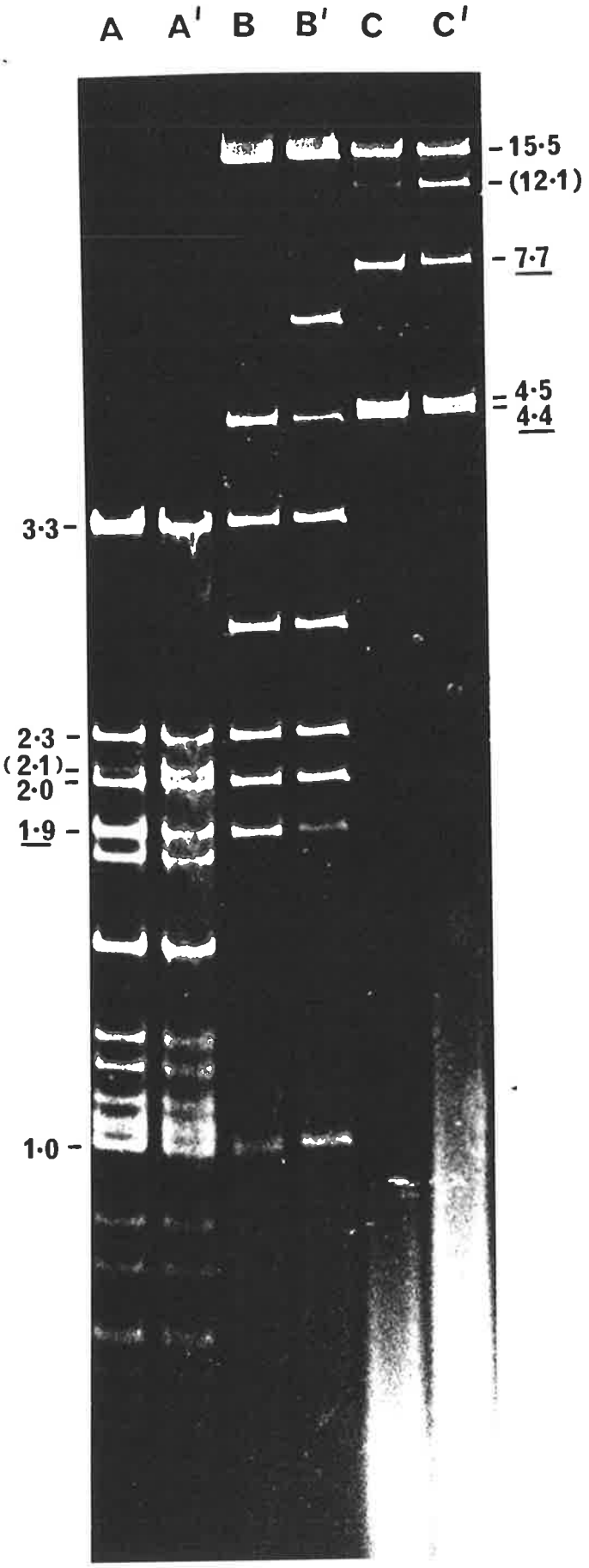
P2 DNA was cleaved three times by *Pst*I to yield fragments of 15.5, 7.7, 4.5 and 4.4 kb of which the 7.7 and 4.4 kb fragments were at the ends (Fig.8.5). The 4.4 and 15.5 kb fragments were present in Hy5 showing that the 4.4 kb fragment was the left hand end and the 15.5 kb fragment was adjacent to it. This left the 4.5 kb fragment as the right of the two internal fragments. The sites of *Pst*I cleavage of P2 DNA were 14%, 62% and 76% from the left hand end.

f) *Sal*I:

186 DNA was cleaved nine times by *Sal*I to yield fragments of 7.5, 6.3, 3.5, 3.5, 2.6, 2.2, 1.9, 1.0,

FIGURE 8.5. *Pst*I Digestion of 186 and P2 DNA.

186, Hy5 and P2 DNAs were digested to completion with *Pst*I, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks etc., are as described for Figure 8.2.


A
A′
B
B′
C
C′
15·5
(12·1)
7·7
4·5
4·4
3·3
2·3
(2·1)
2·0
1·9
1·0

0.8 and 0.7 kb (Fig.8.6(a)). The 1.0 and 7.5 kb fragments reannealed to give an 8.5 kb fragment. The 7.5 kb fragment was present in Hy5 and therefore was the right hand end fragment, the 1.0 kb fragment being at the left end. As no other fragments of 186 were present in Hy5, the fragment adjacent to the 7.5 kb end fragment must have been larger than 3 kb, as 10.5 kb of the right end of 186 is represented in the Hy5. Thus only the 6.3 kb fragment and either of the 3.5 kb fragments could have been adjacent. Fig.8.6(b) shows that one of the 3.5 kb fragments, but not the 6.3 kb fragment was present in Hy10 DNA which contains about 15 kb of the right end of 186. Therefore the fragment adjacent to the 7.5 kb end fragment must have been 3.5 kb long. Three of the nine *Sal*I cleavage sites on 186 DNA can thus be localized at 3%, 63% and 75% from the left hand end. The mapping of *Sal*I sites on 186 DNA was not carried any further.

P2 DNA was cleaved twice with *Sal*I to yield fragments of size 19.4, 10.8 and 1.8 kb of which the 1.8 and 19.4 kb fragments reannealed to give a 21 kb fragment and were thus at the ends (Fig.8.6(a)). The 1.8 kb fragment was present in Hy5 and was therefore the left hand end piece leaving the 19.4 kb fragment at the right end and the 10.8 kb fragment in the middle. The presence of the 10.8 kb fragment in the digest of Hy5 DNA further confirmed this order. The *Sal*I cleavage sites on P2 DNA were therefore 6% and 42% from the left hand end.

g) *Xba*I:

186 DNA was uncut by *Xba*I (Fig.8.7).

P2 DNA was cleaved once only to yield fragments of

FIGURE 8.6. *Sal*I Digestion of 186 and P2 DNA.

186, Hy5, Hy10 and P2 DNAs were digested to completion with *Sal*I, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light.
Tracks are as follows:

a) A: 186 DNA end-denatured as described in Chapter 7

A': 186 DNA

B: Hy5 DNA end-denatured

B': Hy5 DNA

C: P2 DNA end-denatured

C': P2 DNA

b) D: 186 DNA

E: Hy5 DNA

F: Hy10 DNA

Other details are as described in Fig.8.1.

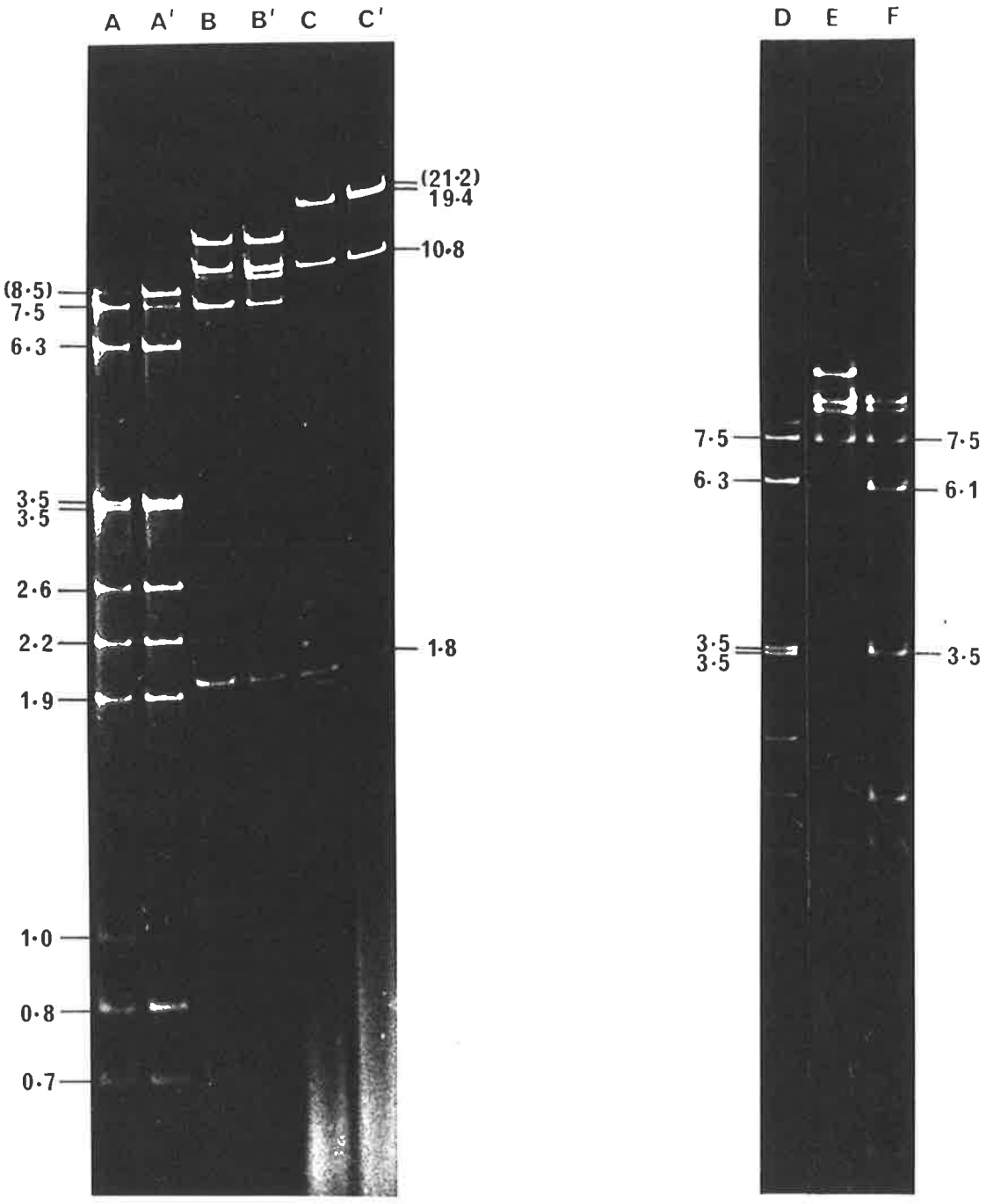
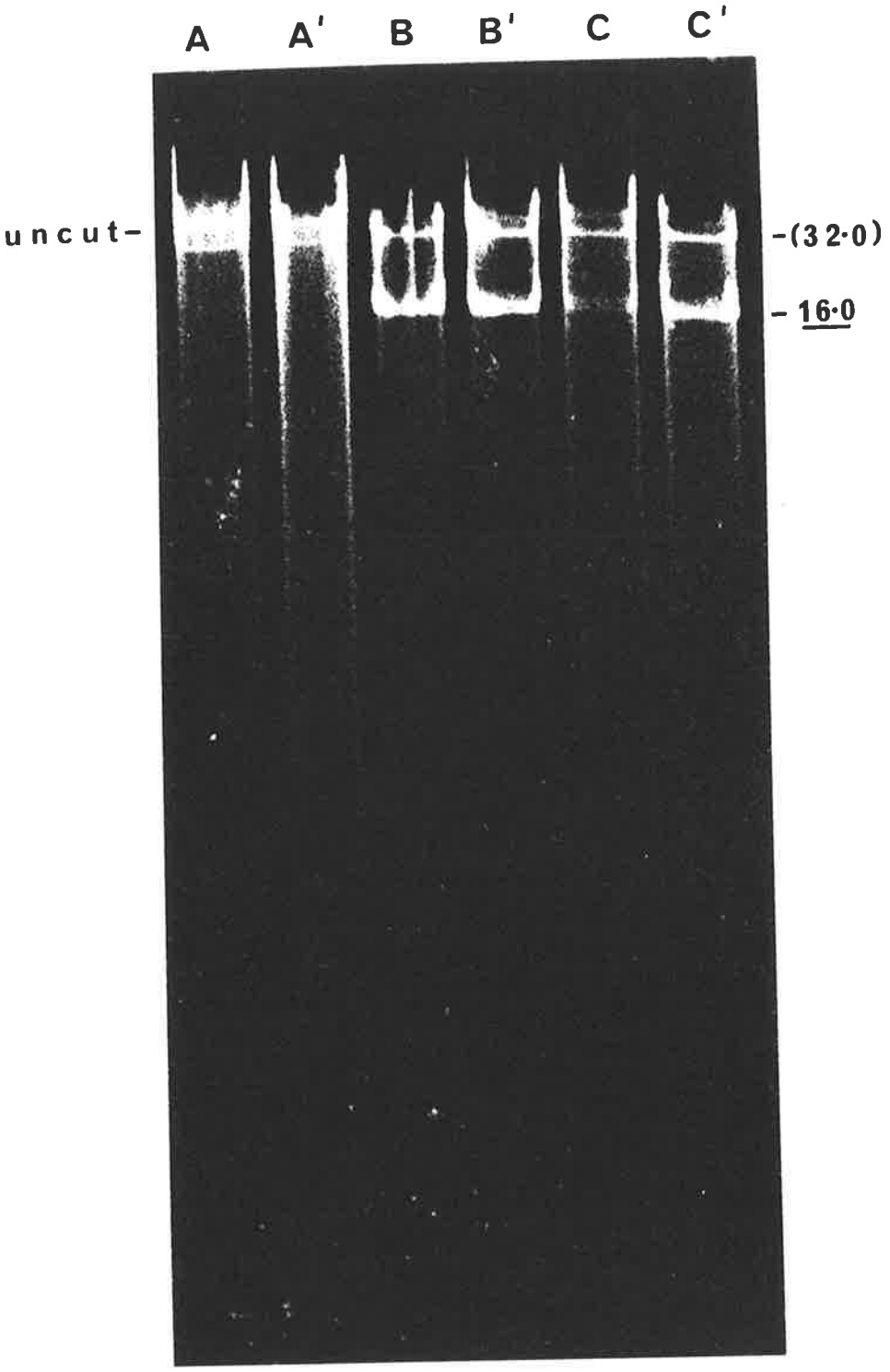

A
A'
B
B'
C
C'
(21·2)
19·4
10·8
(8·5)
7·5
6·3
3·5
3·5
2·6
2·2
1·9
1·8
1·0
0·8
0·7
D
E
F
7·5
6·3
7·5
6·1
3·5
3·5
3·5

FIGURE 8.7. *Xba*I Digestion of 186 and P2 DNA.

186, Hy5 and P2 DNAs were digested to completion with *Xba*I, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks etc., are as described for Fig.8.2.


A
A'
B
B'
C
C'
uncut
(32·0)
16·0

equal size (Fig.8.7). The site of *Xba*I cleavage of P2 DNA was therefore 50% from the left hand end.

h) *Xho*I:

186 DNA was cleaved once by *Xho*I to yield fragments of 9.5 and 20.5 kb (Fig.8.8). As the 9.5 kb fragment was present in Hy5, it must have been at the right hand end of 186. The site of cleavage of *Xho*I on 186 DNA was therefore 68% from the left hand end.

P2 DNA is not cleaved by *Xho*I (Fig.8.8).

2. Mapping from codigestion and redigestion:

The incompleted maps, *Bam*HI and *Pst*I - 186 and *Bam*HI-P2 were further mapped by examining DNA digested with two enzymes or an isolated fragment digested with a second enzyme. This method allowed the completion of the mapping of *Bam*HI-P2 and the *Pst*I-right 35% of 186 and extended the map of *Bam*HI-186 as described below and summarized in Fig.8.15(b) and Fig.8.16(b).

a) *Bam*HI:

186 DNA cleaved with *Bam*HI and *Hin*dIII showed that the *Bam*HI fragments of size 11.2 and 6.5 kb were both cleaved near an end to give slightly reduced size (Fig.8.9, Track B). From the mapping above we would expect the *Hin*dIII to cut the *Bam*HI 11.2 kb piece and generate a 10.4 kb fragment. This implies that the left *Hin*dIII cut was within the 6.5 kb fragment and very close to one end, as the cleaved fragment was 6.3 kb long. When the isolated *Hin*dIII 7.1 kb fragment was redigested with *Bam*HI, two fragments were detectable of size 6.3 and 0.8 (the latter of which was too faint to be visible in Fig.8.9, Track C, but was

FIGURE 8.8. *Xho*I Digestion of 186 and P2 DNA.

186, Hy5 and P2 DNAs were digested to completion with *Xho*I, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks etc., are as described for Fig.8.2.

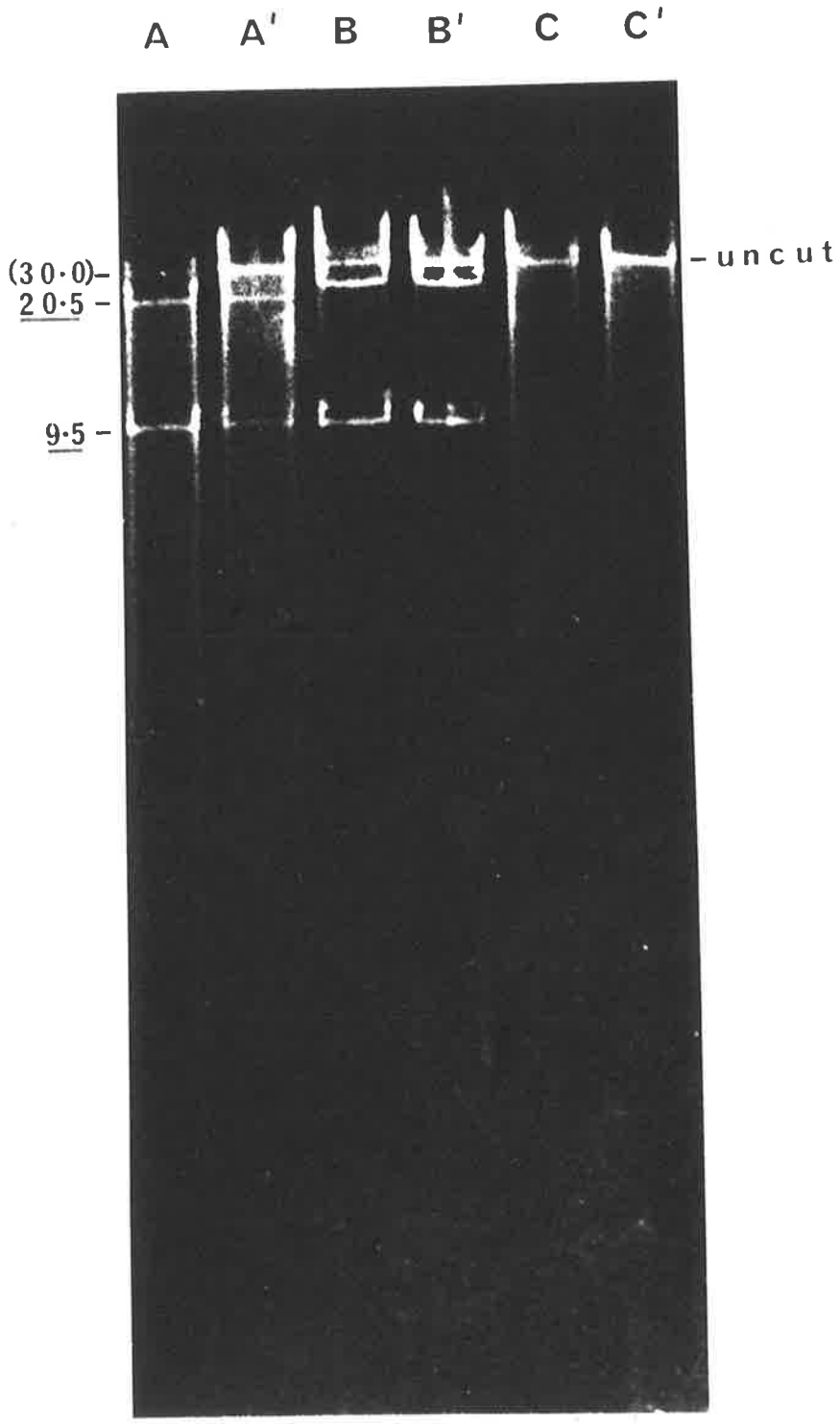

A
A′
B
B′
C
C′
(30·0)–
20·5 –
9·5 –
– uncut

FIGURE 8.9. *Bam*HI Double Digests of 186 DNA.

186 DNA was digested to completion with two enzymes or isolated fragments were redigested with a second enzyme, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks are as follows:

A: *Bam*HI digest

B: *Bam*HI/*Hin*dIII double digest

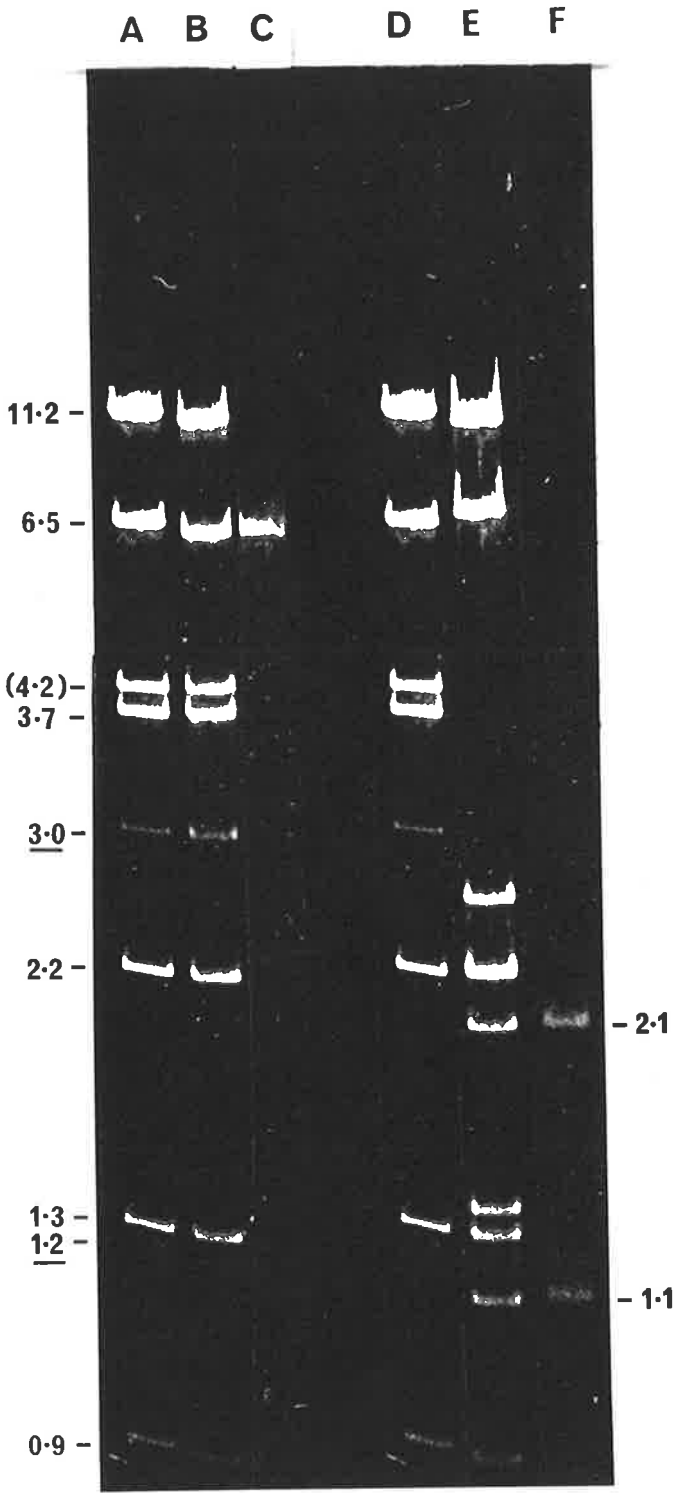
C: *Hin*dIII-186 7.1 kb fragment redigested with *Bam*HI

D: *Bam*HI digest

E: *Bam*HI/*Eco*RI double digest

F: *Eco*RI-186 3.3 kb fragment redigested with *Bam*HI.

Fragment sizes are expressed in kb.


A B C D E F
11·2 –
6·5 –
(4·2)–
3·7 –
3·0 –
2·2 –
– 2·1
1·3 –
1·2 –
– 1·1
0·9 –

apparent on the original). The 6.5 kb *Bam*HI fragment must therefore have been adjacent to the 11.2 kb fragment.

When 186 was digested by *Bam*HI and *Eco*RI, three *Bam*HI fragments were cleaved (Fig.8.9, Track E). These were the 11.2 kb fragment, the 3.0 kb end fragment and the 3.7 kb fragment. Again from the data presented earlier we could predict the cleavage of the 11.2 kb fragment into 10.0 kb and 1.2 kb fragments and the 3.0 kb fragment into 2.3 kb and 0.7 kb fragments. This leaves the 3.7 kb fragment which must have been cut at the *Eco*RI site 4.0 kb in from the left hand end. The isolated 3.3 kb *Eco*RI fragment was cleaved only once by *Bam*HI to yield fragments of apparent size 2.1 and 1.1 kb (Fig. 8.9, Track F). The *Bam*HI 3.7 kb fragment must therefore have been adjacent to the 3.0 kb end fragment, as there were no *Bam*HI sites between the *Bam*HI cut 3.0 kb and the *Eco*RI cut 4.0 kb from the left hand end (which is within the 3.7 kb *Bam*HI fragment).

P2 DNA cleaved with both *Bam*HI and *Eco*RI yielded fragments of sizes 11.7, 10.3, 6.6, 2.1 and 1.0 kb (Fig. 8.10). This implied that the 2.1 kb fragment and not the 12.7 kb fragment is adjacent to the left hand 0.2 kb fragment, since if it was, it should have been cleaved into 3.1 and 9.6 kb fragments and not 1.0 and 11.7 kb fragments as observed. This was confirmed when the *Eco*RI 12.0 kb fragment was redigested with *Bam*HI to give an 11.7 kb and no 2.1 kb fragment, while the *Eco*RI 3.3 kb fragment gave rise to 2.1 and 1.0 kb fragments (the 0.2 kb end fragment is not detected on these gels). The final

FIGURE 8.10. *Bam*HI Double Digests of P2 DNA.

P2 DNA was digested to completion with two enzymes or isolated fragments were redigested with a second enzyme, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks are as follows:

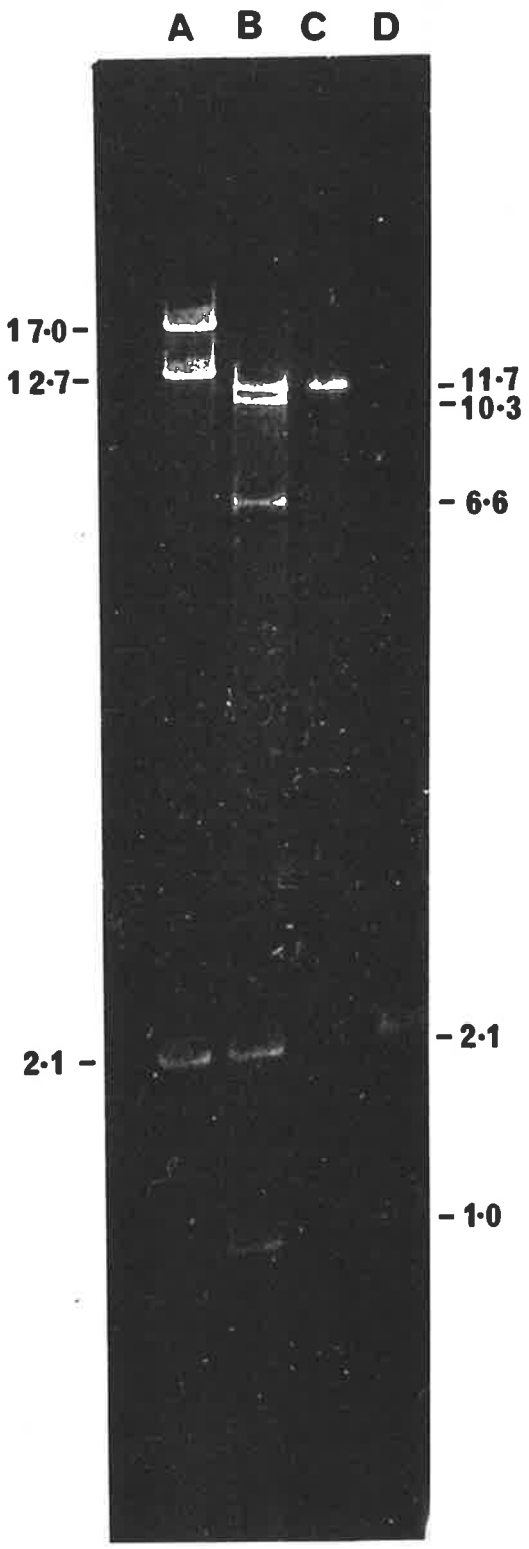
A: *Bam*HI digest

B: *Bam*HI/*Eco*RI double digest

C: *Eco*RI-P2 12.0 kb fragment redigested with *Bam*HI

D: *Eco*RI-P2 3.3 kb fragment redigested with *Bam*HI

Fragment sizes are expressed in kb.


A
B
C
D
17·0
12·7
11·7
10·3
6·6
2·1
2·1
1·0

order of *Bam*HI-P2 fragments is 0.2, 2.1, 12.7 and 17.0 with sites of cleavage 0.5%, 7% and 47% from the left hand end.

b) *Pst*I:

The presence of *Eco*RI and *Bgl*II sites 2.4 kb and 5.7 kb respectively from the right hand end were used to order the five *Pst*I-186 fragments of interest in the right 35% of 186 (i.e., those present in Hy5). Double digestion of Hy5 with *Pst*I and *Eco*RI affected the 2.1 kb fragment, whereas double digestion with *Pst*I and *Bgl*II affected only the 2.3 kb fragment (Fig.811). Thus the *Eco*RI site was within the 2.1 kb fragment and the *Bgl*II site was within the 2.3 kb fragment. From end fragment determination, we already know that the 1.9 kb fragment was the right hand end fragment. As the *Eco*RI site, only 0.5 kb further along, was within the 2.1 kb fragment, this fragment must have been adjacent to the end fragment as all other fragments are too large to fit in between. The *Bgl*II site that cleaved the 2.3 kb fragment into 1.6 and 0.7 kb fragments was a further 2.3 kb along, which necessitated the placement of the 1.0 kb fragment between the 2.1 kb and 2.3 kb fragments (the 3.3 kb alternative is too large). This conclusion was confirmed by partial digestion analysis as described below. The order of the fragments from the 65% mark of 186 must therefore have been 3.3, 2.3, 1.0, 2.1, 1.9 and the sites of cleavage 65%, 76%, 84%, 87% and 94% from the left hand end.

3. Mapping from partial digestion analysis:

Partial digestion analysis allowed the completion of the *Bam*HI map of coliphage 186 and confirmed the order of

FIGURE 8.11. *Pst*I Double Digests of 186 DNA.

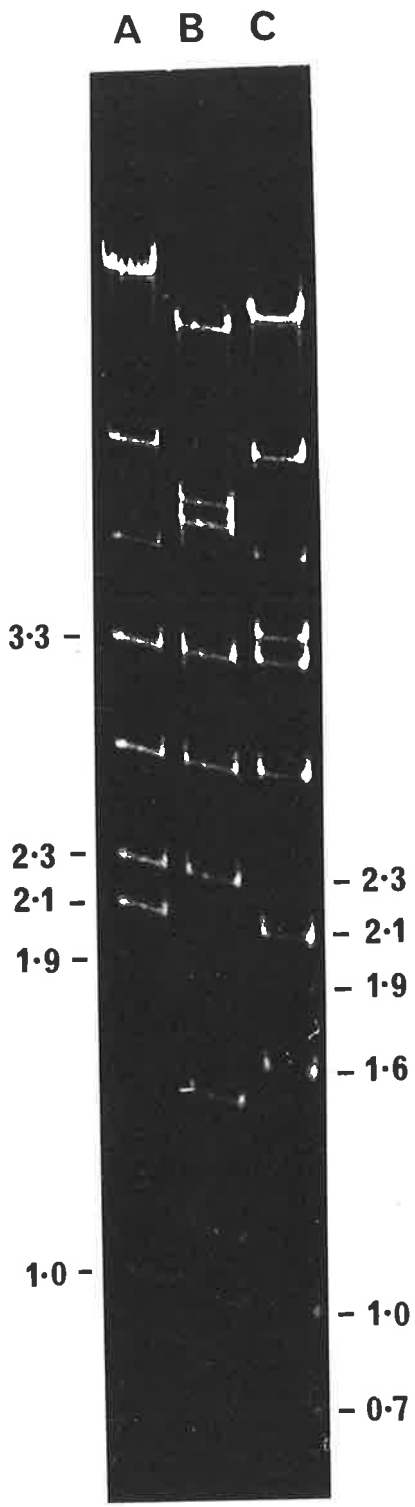
186 DNA was digested to completion with *Pst*I and either *Eco*RI or *Bgl*II, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks are as follows:

A: *Pst*I digest

B: *Pst*I/*Eco*RI double digest

C: *Pst*I/*Bgl*II double digest

Fragment sizes are expressed in kb.


A
B
C
3·3 –
2·3 –
2·1 –
1·9 –
1·0 –
– 2·3
– 2·1
– 1·9
– 1·6
– 1·0
– 0·7

*Pst*I sites in the right 35% of 186 DNA.

a) *Bam*HI:

Fig.8.15(b) shows the *Bam*HI sites on coliphage 186 DNA mapped by other methods. These approaches could not order the 0.9, 1.3 and 2.2 kb fragments, so the method described above was used to complete the mapping. Fig.8.12, Tracks A to E show a series of *Bam*HI digests of 186 DNA incubated for increasing lengths of time. Track A was obviously an incomplete digest while Track E was apparently complete digestion. The sizes of the complete digestion products are shown on the left and were determined by the co-electrophoresis of *Eco*RI and *Hin*dIII digests of λDNA as markers. The 1.2 kb and 3.0 kb fragments are the end fragments which can reanneal to form the 4.2 kb fragment, but this property is not relevant to this work. When the partials (Fig.8.12, Tracks A to E) were denatured *in situ*, neutralized, transferred to nitrocellulose filter paper by the method of Southern (1975) and annealed to a radioactively-labelled pure preparation of the *Bam*HI - 186 fragment of size 3.7 kb, the resulting autoradiography revealed the pattern of partials that included this fragment (Fig.8.12, Tracks A' to E'). Track E' showed the purity of the 3.7 kb fragment used as probe, since in this complete digest only the 3.7 kb fragment annealed to it. The other tracks revealed that the size of the two smallest partial fragments which included the 3.7 kb fragment were 4.6 kb and 6.7 kb. From the previous mapping (see Fig.8.15(b)) it is known that the smallest partial involving the 3.7 kb fragment and a fragment to the left (the 3.0 kb end fragment) would be of

FIGURE 8.12. Partial *Bam*HI-186 Digests Annealed to the *Bam*HI 3.7 kb Fragment.

A series of partial *Bam*HI digests of 186 DNA were electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks A to E are the fractionated digests of 0.5 µg of 186 DNA with sufficient *Bam*HI to give complete digestion after 45 min, incubated for :

A: 10 min

B: 15 min

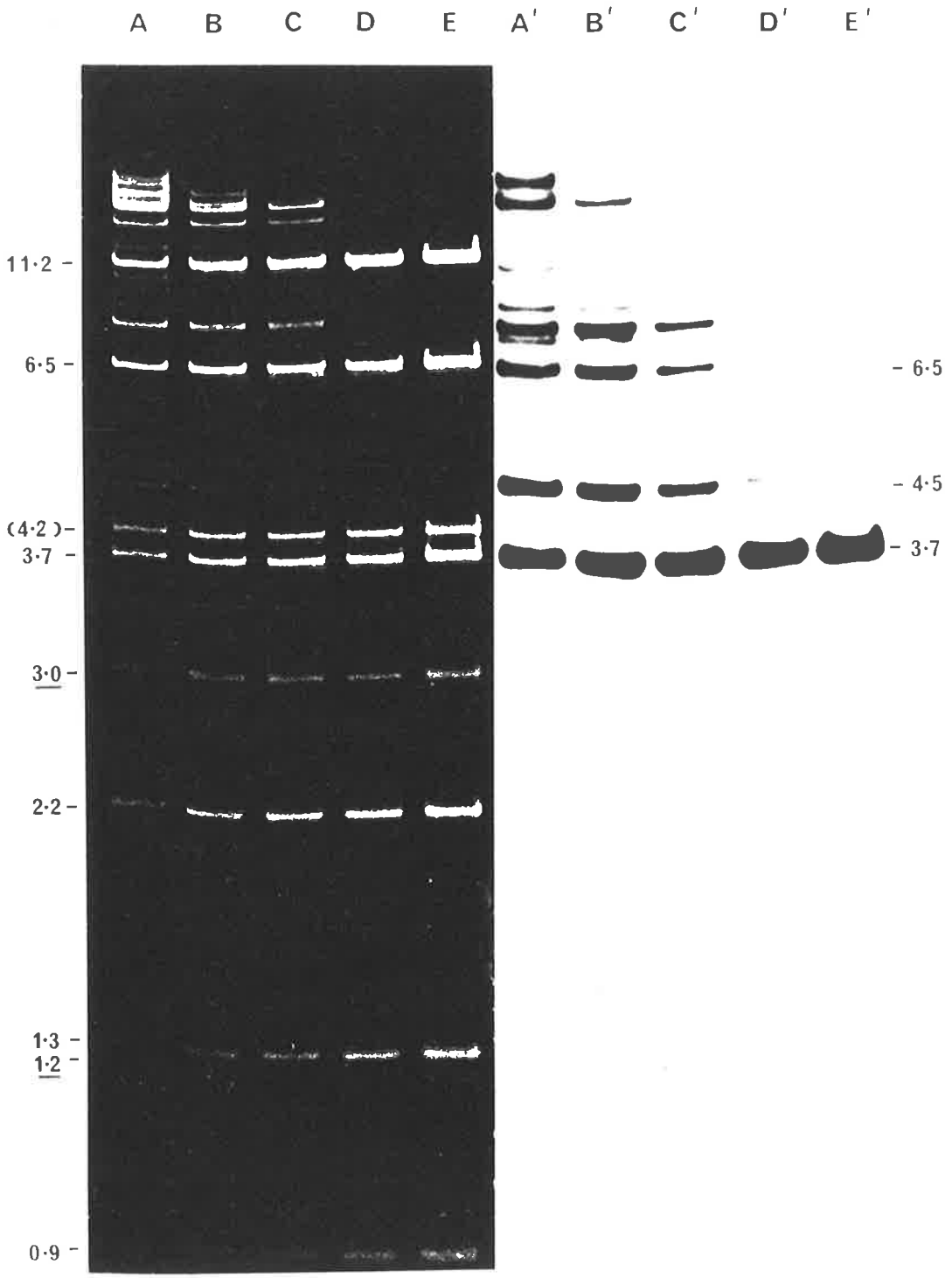
C: 20 min

D: 30 min

E: 60 min.

The DNA in the gel was denatured by alkali treatment, neutralized, transferred and immobilized to nitrocellulose filter paper, annealed to ^{32}P-labelled pure *Bam*HI-186 3.7 kb fragment, washed, dried and autoradiographed to produce the patterns shown in Tracks A' to E'. Fragment sizes are expressed in kb.

Pure *Bam*HI-186 3.7 kb fragment was obtained by cloning the fragment into the plasmid vector pBR322 (Finnegan and Egan, 1979) and was the gift of E.J. Finnegan.


A
B
C
D
E
A′
B′
C′
D′
E′
11·2
6·5
(4·2)
3·7
3·0
2·2
1·3
1·2
0·9
6·5
4·5
3·7

size 6.7 kb. Therefore the 4.6 kb partial must have been the sum of the 3.7 kb fragment and the fragment adjacent to it and to the right of it, i.e., the 0.9 kb fragment was the adjacent fragment. The next partial involving the 3.7 kb, the 0.9 kb and the third fragment to the right (i.e., the 1.3 kb or 2.2 kb fragment) would be either 5.9 kb or 6.8 kb long, depending on the order of these two fragments.

The absence of a 5.9 kb partial suggested that the 1.3 kb fragment could not have been adjacent, while the presence of a 6.7 kb fragment suggested that the 2.2 kb fragment was next to the 0.9 kb fragment. Unfortunately, this size was approximately the size of the partial that should have been generated by the 3.7 kb fragment and the 3.0 kb end fragment. Therefore, to confirm that the order of the remaining fragments from left to right was 0.9 - 2.2 - 1.3, a similar set of partials was hybridized to the labelled 6.5 kb fragment (Fig.8.13) which we know to be to the right of the unmapped fragments (Fig.8.15(b)). As the 11.2 kb fragment was to the right of the 6.5 kb fragment, partials involving these two fragments were far too large to interfere with the partials generated with the fragments to the left. Fig.8.13, Tracks A, B and C show increasing times of digestion of 186 DNA with *Bam*HI, and Tracks A', B' and C' show these annealed to the ^{32}P-labelled 6.5 kb fragment. Again the purity of the probe was shown in Track C' with only the 6.5 kb fragment of the complete digest annealing to the probe. Also Track A' of Fig.8.13 showed that, as expected, the set of partials annealing

FIGURE 8.13. Partial *Bam*HI-186 Digests Annealed to the *Bam*HI-186 6.5 kb Fragment.

A series of partial *Bam*HI digests of 186 DNA were treated as described in Fig.8.12, except that the fractionated and immobilized fragments were annealed to pure ^{32}P-labelled *Bam*HI-186 6.5 kb fragment, washed, dried and autoradiographed.

Tracks A, B and C show the fractionated digests as described in Fig.8.12 digested for:

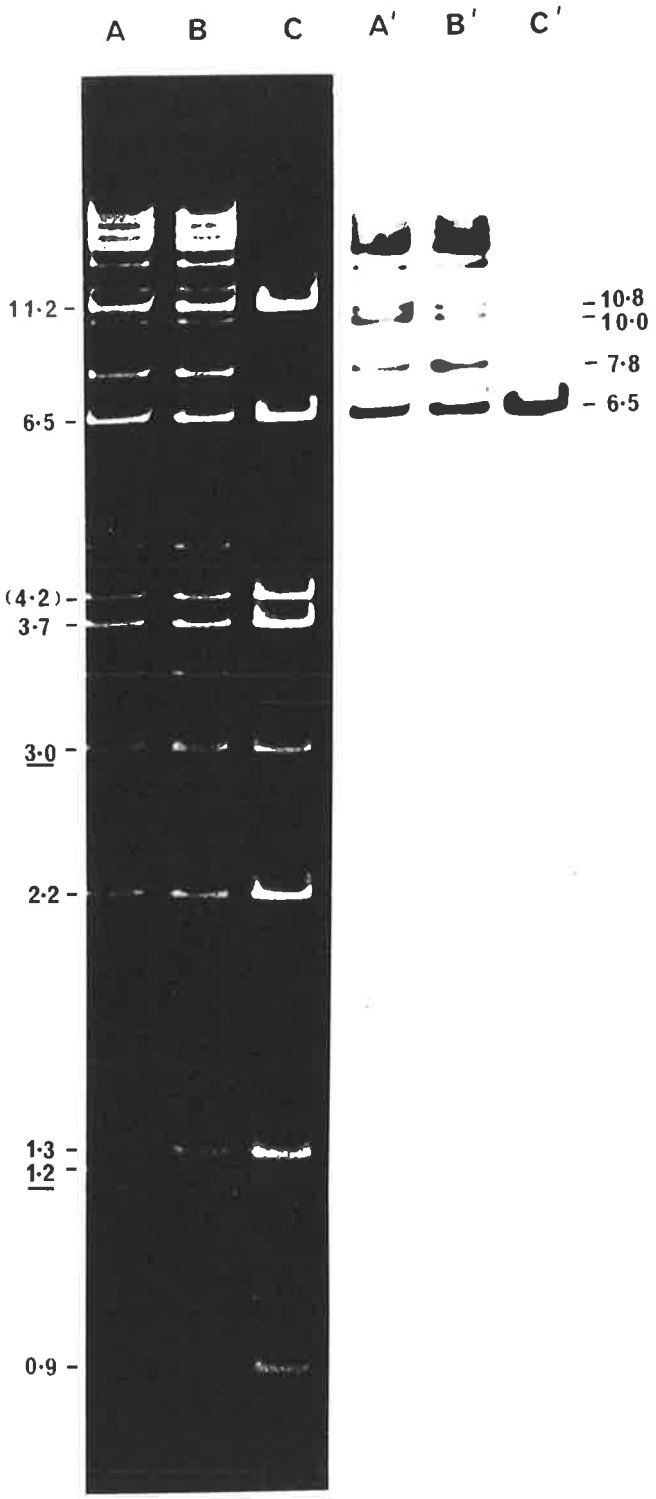
A: 10 min

B: 15 min

C: 60 min

Tracks A', B' and C' are the resulting autoradiogram. Fragment sizes are expressed in kb.

Pure *Bam*HI-186 6.5 kb fragment was obtained by cloning the fragment into the plasmid vector pBR322 (Finnegan and Egan, 1979) and was the gift of E.J. Finnegan.


A
B
C
A′
B′
C′
11·2
6·5
(4·2)
3·7
3·0
2·2
1·3
1·2
0·9
10·8
10·0
7·8
6·5

to the 6.5 kb fragment was different to the set annealing to the 3.7 kb fragment seen in Fig.8.12, Track A'. The sizes of the partials annealing to the 6.5 kb fragment were 7.8 kb, 10.0 kb and 10.8 kb confirming the above order of fragments. Thus the 1.3 kb fragment was adjacent to the 6.5 kb fragment, the 2.2 kb fragment was next and the 0.9 kb the third fragment from the 6.5 kb fragment. The final order of *Bam*HI fragments from the left of 186 was therefore 3.0, 3.7, 0.9, 2.2, 1.3, 6.5, 11.2, 1.2 and the sites of cleavage were at 10%, 22%, 25%, 33%, 37%, 59%, and 96%. This result is summarized in Fig.8.15(c).

b) *Pst*I:

The above method was used to confirm the order of some *Pst*I fragments of 186 indirectly inferred by double digestion (see above). Fig.8.16(b) shows the map of *Pst*I sites on the right 35% of coliphage 186 DNA as determined by the other methods. The isolated 1.0 kb fragment was used as probe to detect the set of partials containing that fragment. *Pst*I - 186 partial digests are shown in Fig.8.14, Tracks A, B and C and the autoradiograms of these annealed to the ^{32}P-labelled 1.0 kb piece are shown in Tracks, A', B' and C'. The size of the two smallest partials were 3.0 kb and 3.2 kb, suggesting that the adjacent fragments were of approximate size 2.0 kb and 2.2 kb. This agreed well with our prediction that the 1.0 kb fragment lay between the 2.1 kb and the 2.3 kb fragments.

The results of the mapping described above are summarized in Fig.8.15(c) for 186 and Fig.8.16(b) for P2.

FIGURE 8.14. Partial *Pst*I-186 Digests Annealed to the *Pst*I-186 1.0 kb Fragment.

A series of partial *Pst*I digests of 186 DNA were electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks A and B are partial digests while Track C is a complete digest.

The DNA in the gel was denatured *in situ* transferred and immobilized onto nitrocellulose filter paper, annealed to ^{32}P-labelled pure *Pst*I-186 1.0 kb fragment, washed, dried and autoradiographed to produce the patterns shown in Tracks A', B', and C'. Fragment sizes are expressed in kb.

Pure *Pst*I-186 1.0 kb fragment was obtained by cloning the fragment into the plasmid vector pBR322 (Finnegan and Egan, 1979) and was the gift of E.J. Finnegan.

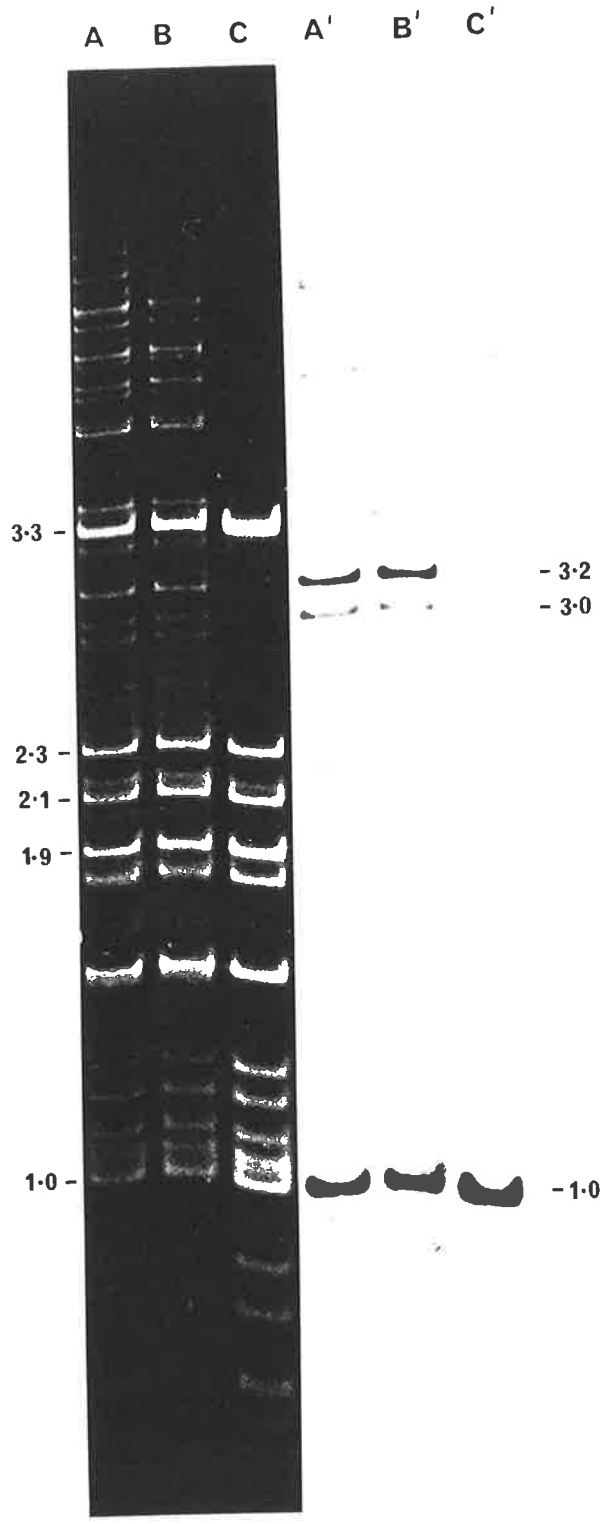

A
B
C
A'
B'
C'
3·3
2·3
2·1
1·9
1·0
3·2
3·0
1·0

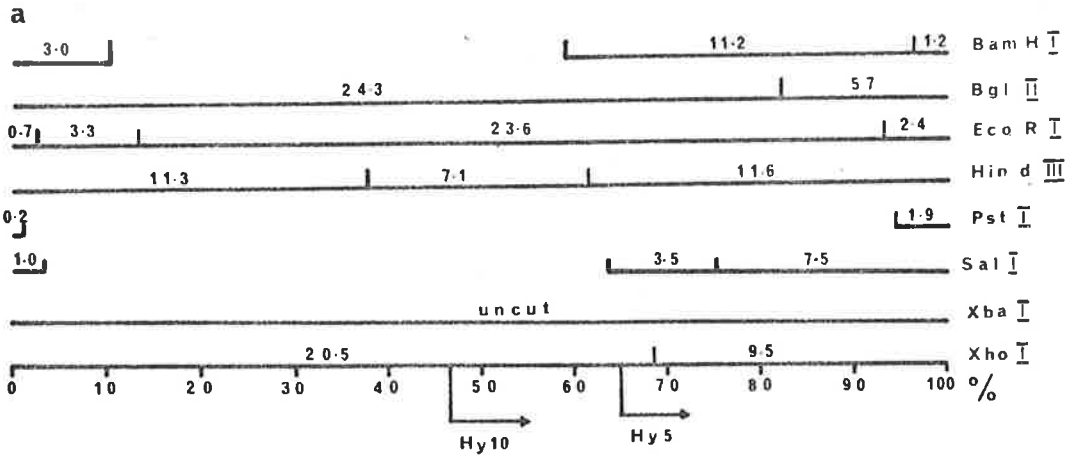
FIGURE 8.15. Restriction Endonuclease Cleavage Sites of 186 DNA.

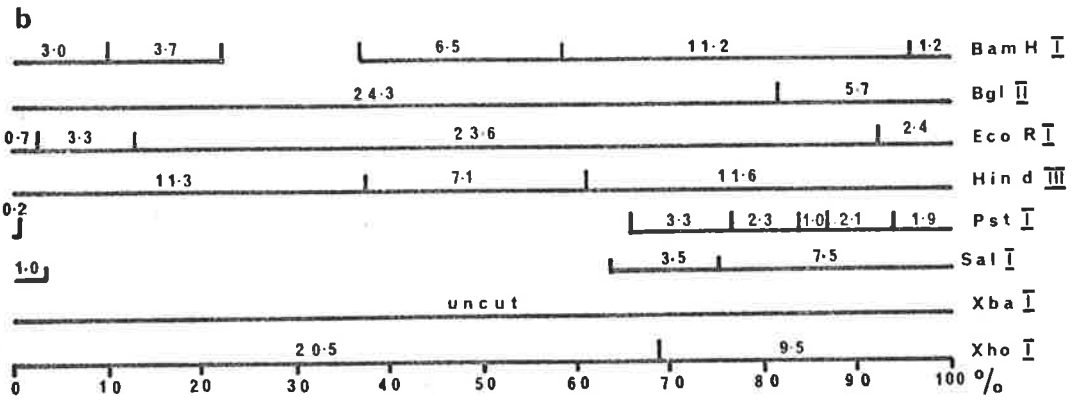
a) Sites determined by end-denaturation and by comparison with the hybrid phage.

b) Sites determined in a) and by double-digestion or redigestion of isolated fragments, and

c) The mapping determined in a) and b) and by partial digest analysis.

Fragment sizes in all cases are expressed in kb.


a
3·0
11·2
1·2
Bam H I
24·3
5·7
Bgl II
0·7
3·3
23·6
2·4
Eco R I
11·3
7·1
11·6
Hin d III
0·2
1·9
Pst I
1·0
3·5
7·5
Sal I
uncut
Xba I
20·5
9·5
Xho I
0
10
20
30
40
50
60
70
80
90
100
%
Hy10
Hy5


b
3·0
3·7
6·5
11·2
1·2
Bam H I
24·3
5·7
Bgl II
0·7
3·3
23·6
2·4
Eco R I
11·3
7·1
11·6
Hin d III
0·2
3·3
2·3
1·0
2·1
1·9
Pst I
1·0
3·5
7·5
Sal I
uncut
Xba I
20·5
9·5
Xho I
0
10
20
30
40
50
60
70
80
90
100
%

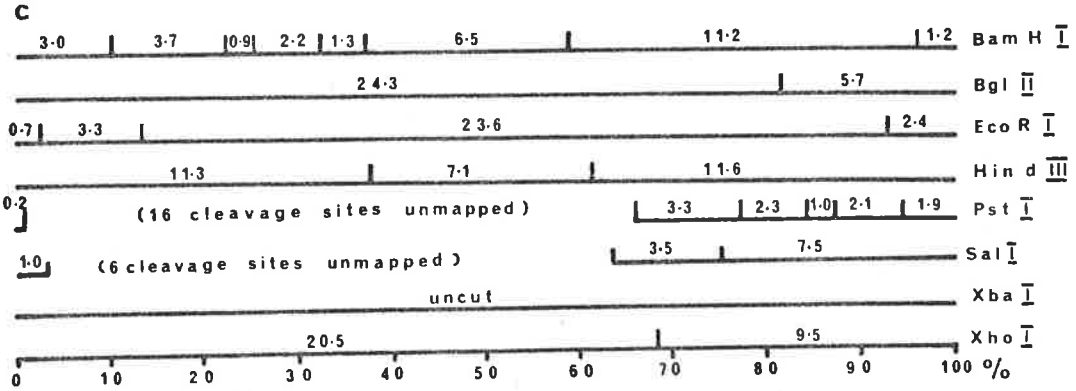

c
3·0
3·7
0·9
2·2
1·3
6·5
11·2
1·2
Bam H I
24·3
5·7
Bgl II
0·7
3·3
23·6
2·4
Eco R I
11·3
7·1
11·6
Hin d III
0·2
(16 cleavage sites unmapped)
3·3
2·3
1·0
2·1
1·9
Pst I
1·0
(6 cleavage sites unmapped)
3·5
7·5
Sal I
uncut
Xba I
20·5
9·5
Xho I
0
10
20
30
40
50
60
70
80
90
100
%

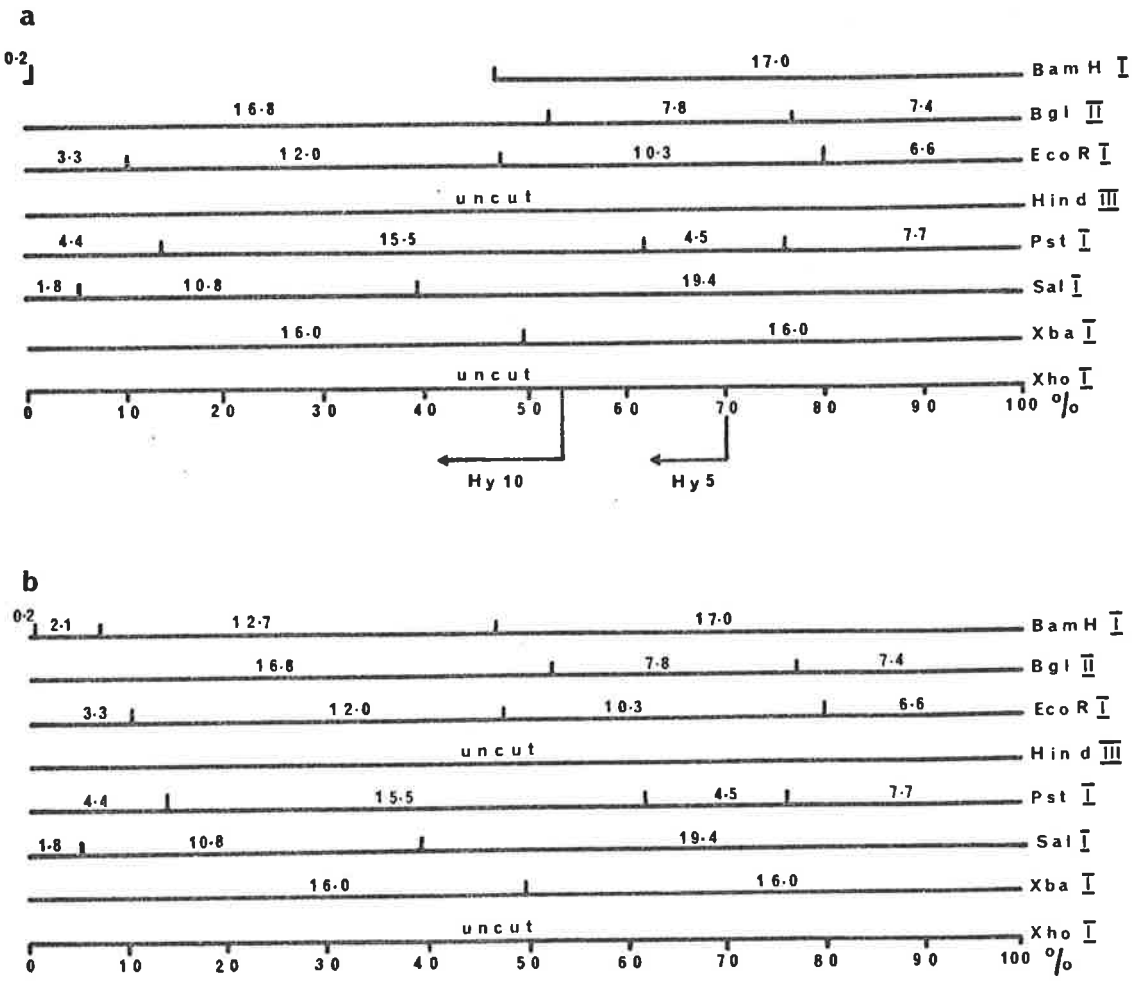
FIGURE 8.16. Restriction Endonuclease Cleavage Sites of P2 DNA.

a) Sites determined by end-denaturation and by comparison with the hybrid phage.

b) The completed map using double and redigestion of fragments.

Fragment sizes in both cases are expressed in kb.

P2 cleavage sites



4. An analysis of 186 insertion and deletion mutants:

Several insertion and deletion mutants of phage 186 have been isolated that affect the major control sites around the *c*I gene, approximately 22 kb from the left hand end (Younghusband *et al.*, 1975). Of these, the four mutants 186*ins*2, 186*ins*3, 186*del*1 and 186*del*2 have been characterized genetically and by heteroduplex analysis in the electron microscope, and were therefore chosen for further analysis by restriction cleavage.

a) 186*ins*2 and 186*ins*3 each carry an IS3 element:

The insertion mutant 186*ins*2 has been shown to be a $c\mathrm{I}^-$ mutant with a single insertion of 1.2 kb located 22.1 kb from the left end, while 186*ins*3 is an int^- mutant of 186 with an insertion of 1.2 kb located 21.1 kb from the left end (Younghusband *et al.*, 1975). The insertions were of unknown origin but in view of the recent discovery of 'insertion sequences' (IS elements) that can insert into and delete from DNA at remarkably high frequencies *in vivo* (see Bukhari *et al.*, 1977 for a general treatise) it was of interest to determine if either of the 186 inserts was an IS element.

(i) Restriction endonuclease evidence:

The restriction map of 186 DNA described above was used to analyse the inserted DNA sequences of 186*ins*2 and 186*ins*3, with respect to IS elements 1 to 5 whose susceptibility to restriction endonuclease cleavage has been determined for certain enzymes (Szybalski, 1977). Table 8.1 summarizes this information for the enzymes *Bam*HI, *Bgl*II, *Eco*RI, and *Hin*dIII. Fig.8.17

TABLE 8.1

SOME PROPERTIES OF INSERTION SEQUENCES 1 TO 5.

The lengths of Insertion Sequences 1 to 5 (IS1 to IS5) and the number of cleavage sites within these sequences for the restriction endonucleases indicated (nd = not determined). This table is taken from data presented by Szybalski (1977).

	IS1	IS2	IS3	IS4	IS5
Length (kb)	∿0.8	∿1.3	∿1.2	∿1.4	∿1.4
*Bam*HI	0	0	0	nd	nd
*Bgl*II	0	nd	nd	nd	nd
*Eco*RI	0	0	0	nd	≥1
*Hin*dIII	0	1	2	nd	nd

FIGURE 8.17. *Bam*HI, *Bgl*II and *Eco*RI Digests of 186*ins*2 and 186*ins*3 DNA.

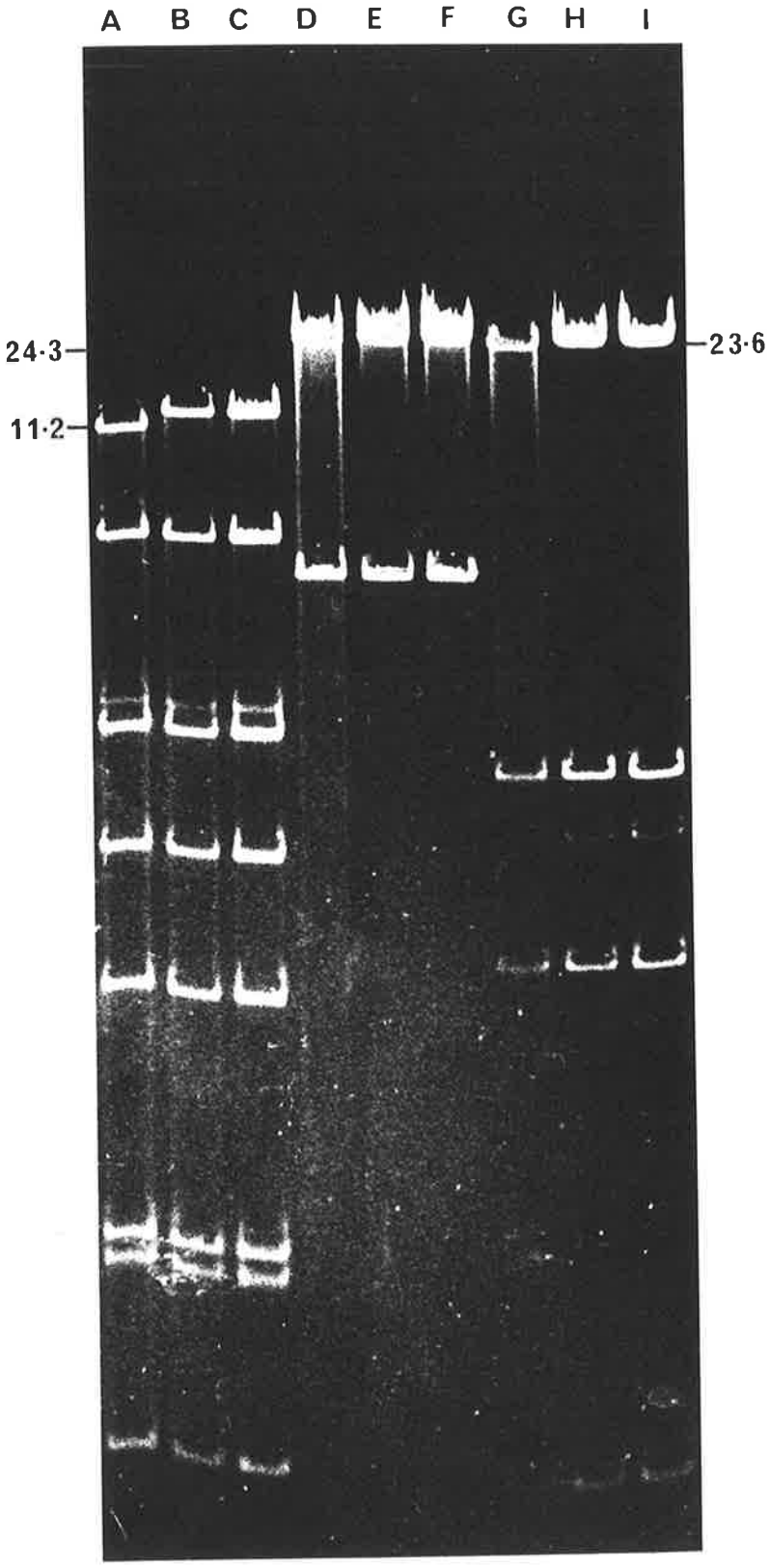
186, 186*ins*2 and 186*ins*3 DNAs were digested to completion with *Bam*HI, *Bgl*II and *Eco*RI, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks were as follows:

A, B and C: *Bam*HI digests of 186, 186*ins*2 and 186*ins*3 DNAs respectively

D, E and F: *Bgl*II digests of 186, 186*ins*2 and 186*ins*3 DNAs respectively

G, H and I: *Eco*RI digests of 186, 186*ins*2 and 186*ins*3 DNAs respectively.

The 186 fragments altered by the mutations are indicated by their size in kb.


A
B
C
D
E
F
G
H
I
24·3
11·2
23·6

shows the results of digesting 186, 186*ins*2 and 186*ins*3 DNA with three of these four enzymes. It can be seen that the number of fragments generated by *Bam*HI, *Bgl*II, and *Eco*RI did not change, but in each case the fragment carrying the inserted DNA had decreased mobility, i.e., greater length. Thus none of these three restriction endonucleases have cleavage sites within the inserted sequence. Digestion of 186*ins*2 and 186*ins*3 with *Hin*dIII, however, did generate new fragments. Fig.8.18 shows that the 11.6 kb right hand end fragment of 186 was cleaved into three fragments of size 8.0 kb, 4.1 kb and 0.9 kb for 186*ins*2 and 8.9 kb, 3.2 kb and 0.9 kb for 186*ins*3. Thus there were 2 *Hin*dIII sites within both these inserted segments of DNA and judging from the identical sizes of the fragment generated from within the inserts (the 0.9 kb fragments), these two inserts may have been identical sequences inserted at positions that appear to be 0.9 kb apart in the 186 genome (see Fig.8.19).

(ii) DNA annealing evidence:

From Table 8.1, therefore, the presence of two *Hin*dIII sites but no sites for *Bam*HI or *Eco*RI and the lengths estimated from heteroduplex analysis of 1200 nucleotides (Younghusband *et al.*, 1975) are all consistent with these inserts being IS3 elements. To test this, λNN*c*I857*r*32 and λp*lac*5*c*I857S7MS505 DNA carrying IS2 and IS3 elements respectively (Hirsch *et al.*, 1972; Fiandt *et al.*, 1972; Malamy *et al.*, 1972) were prepared, labelled by nick translation *in vitro* as described in Chapter 2 and annealed to filters to which either 10 μg of 186, 186*ins*2 or 186*ins*3 DNA had been bound.

FIGURE 8.18. *Hin*dIII Digests of 186*ins*2 and 186*ins*3.

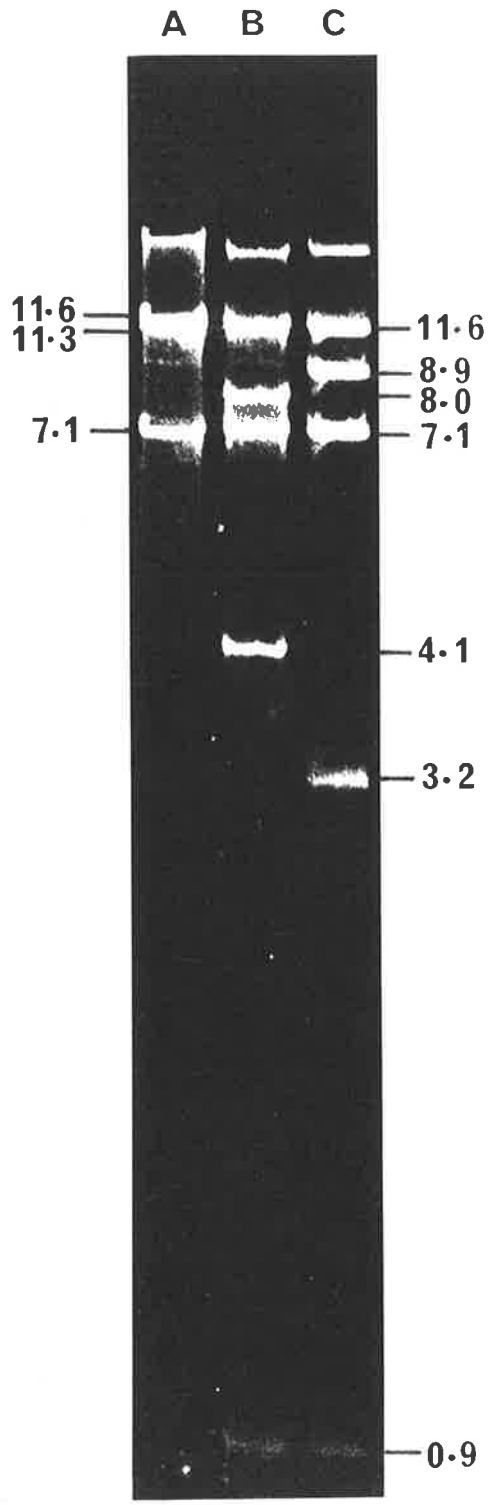
186, 186*ins*2 and 186*ins*3 DNAs were digested to completion with *Hin*dIII, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks are as follows:

A: 186 DNA

B: 186*ins*2 DNA

C: 186*ins*3 DNA

Fragment sizes are expressed in kb.


A
B
C
11·6
11·3
7·1
11·6
8·9
8·0
7·1
4·1
3·2
0·9

After hybridizing and washing as described in Section I, the filters were dried and counted using a toluene based scintillation fluid. The results are shown in Table 8.2 and clearly show sequence homology between λ DNA carrying IS3 (but not IS2) to 186*ins*2 and 186*ins*3, but not to 186 alone. It is therefore concluded that the inserted sequences in 186*ins*2 and 186*ins*3 are due to insertion of an IS3 element into the *c*I and *int* regions of the 186 genome respectively. A diagrammatic representation of these conclusions is shown in Fig.8.19.

b) Detection of a cryptic inserted sequence in a 186 deletion mutant:

The deletion mutant 186*del*1 is a *c*I$^-$ mutant with a net deletion of 1.8 kb while 186*del*2 is an operator defective (*vir*) mutant with a net deletion of 1.5 kb (Dharmarajah, 1975, R. O'Connor, unpublished observations). 186*del*2 was of particular interest, as the 5% deletion removed the operator or 186 that is under the control of the *c*I gene product (the repressor). That a 5% deletion could affect the operator (the repressor binding site) but not the promoter (the RNA polymerase binding site) or the essential gene A, prompted Dharmarajah (1975) to propose that the 186 promoter lay between the operator and structural gene as opposed to the normal situation where the operator overlaps and/or lies between the promoter and structural genes such that repressor bound to the operator presents a physical barrier to the progress of the RNA polymerase molecule along the DNA. Such an unorthodox arrangement has also been proposed for P2 on the same type of evidence

TABLE 8.2

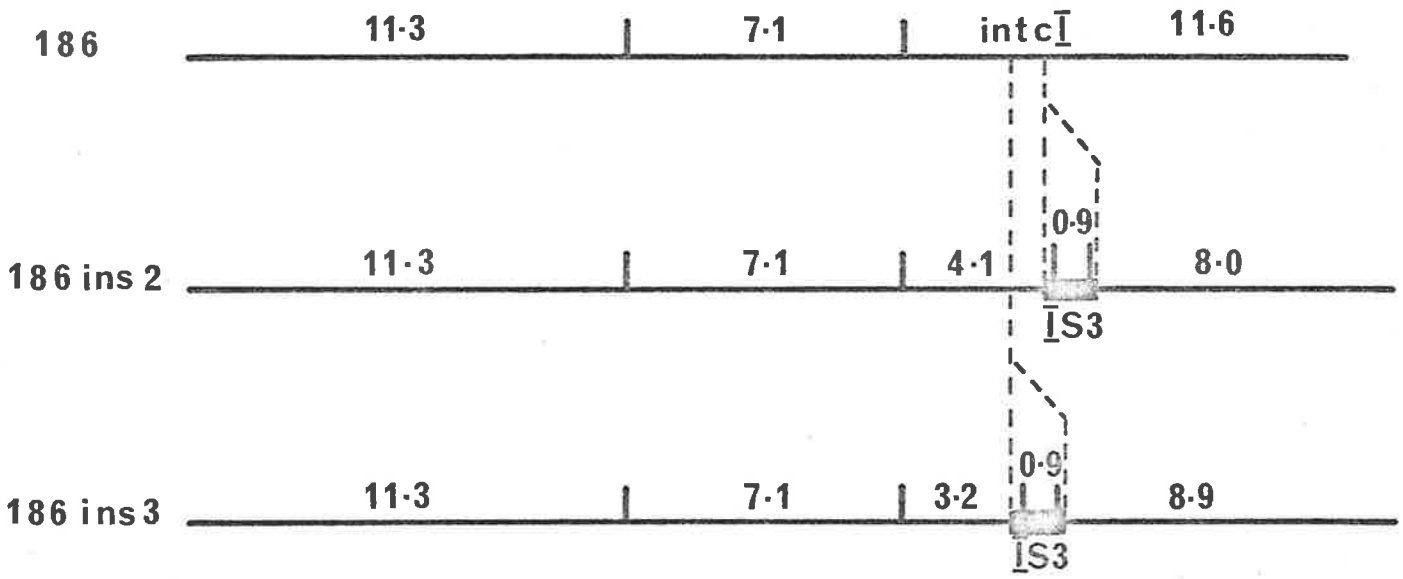
ANALYSIS OF 186 INSERTION MUTANTS.

^{3}H-labelled λ DNA carrying either an IS2 or an IS3 sequence annealed to a blank nitrocellulose filter disc, to 186 DNA bound to a filter disc and to 186*ins*2 and 186*ins*3 DNA bound to filter discs as indicated. The procedure used is described in Chapter 7. Duplicate samples were counted for 10 min each and the results averaged and expressed in c.p.m.

	IS2 ^{3}H-labelled λNN*c*I857*r*32	IS3 ^{3}H-labelled λp*lac*5*c*I857S7MS505
No DNA	33	41
186 DNA	44	38
186*ins*2 DNA	50	125
186*ins*3 DNA	44	174

FIGURE 8.19. Structures of the 186*ins*2 and 186*ins*3 Genomes.

The structures of 186*ins*2 and 186*ins*3 genomes, interpreted from the results described in Chapter 8 are diagrammatically represented by the horizontal lines. Vertical lines indicate sites of *Hin*dIII cleavage and the inserted segments are indicated by thick lines in the linear genomes. *Hin*dIII fragment sizes are expressed in kb.


186
11·3
7·1
intcI
11·6
0·9
186 ins 2
11·3
7·1
4·1
8·0
IS3
0·9
186 ins 3
11·3
7·1
3·2
8·9
IS3

(Bertani and Bertani, 1974). It was of interest, therefore, to map these deletions by restriction analysis and thus the regulatory sites they affect.

The repressor gene (*c*I) known to be affected by the 186*del*1 deletion had previously been localized to a region approximately 22 kb from the left end of 186 (Younghusband *et al.*, 1975). Two restriction sites of interest mapped close to this region, a *Sal*I site 22.5 kb from the left end and a *Pst*I site 22.7 kb from the left end of the genome.

Fig.8.20 shows the results of digesting 186, 186*del*1 and 186*del*2 with *Sal*I and *Pst*I. Tracks A, B and C show that the *Sal*I site was deleted in 186*del*1 and 186*del*2 DNA, as the 7.5 kb and 3.5 kb fragments of 186 (Track A) were replaced by a single fragment of 9.5 kb and 9.9 kb respectively (Tracks B and C). This suggested net deletions of 1.5 kb and 1.1 kb for 186*del*1 and 186*del*2 respectively (compared with the estimates of 1.8 kb and 1.5 kb from length measurements by electron microscopy mentioned earlier.) Both of these deletions, therefore, remove the *Sal*I site located 22.5 kb from the left end of 186 DNA.

Digestion of the deletion mutants with *Pst*I, however, gave a more complex result. Fig.8.20 Track D shows the pattern of *Pst*I generated 186 fragments. The adjacent Track E shows the *Pst*I digest of 186*del*1 DNA which behaved exactly as expected for a simple deletion that did not remove any *Pst*I cleavage sites. The same number of fragments was present in both 186 and 186*del*1, but the 3.3 kb fragment of 186 had been reduced to 1.8 kb, confirming

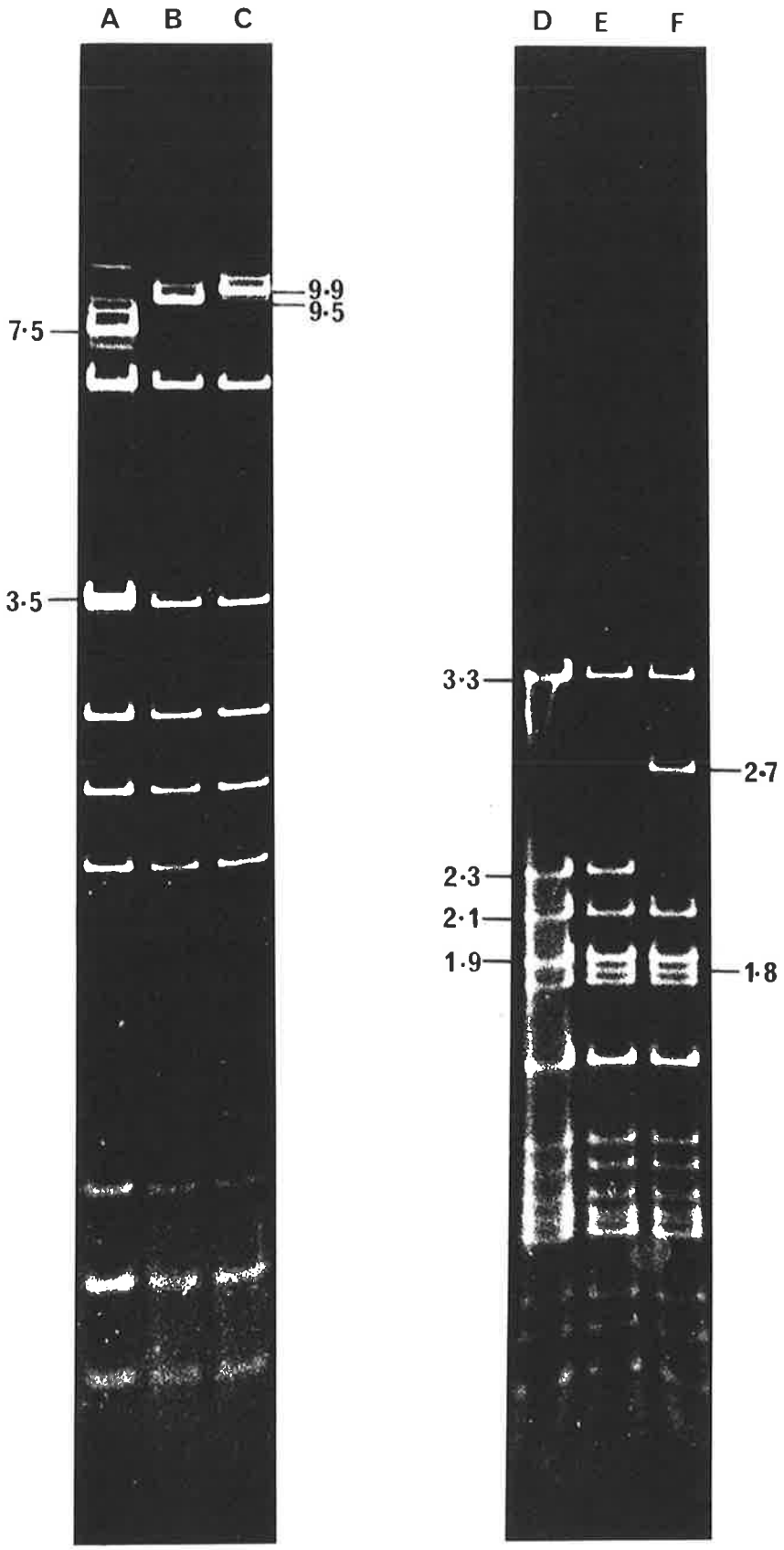
FIGURE 8.20. *Sal*I and *Pst*I Digests of 186*del*1 and 186*del*2.

186, 186*del*1 and 186*del*2 DNAs were digested to completion with *Sal*I or *Pst*I, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks are as follows:

A, B and C: *Sal*I digests of 186, 186*del*1 and 186*del*2 DNA respectively

D, E and F: *Pst*I digests of 186, 186*del*1 and 186*del*2 DNA respectively.

The fragments altered by the mutations are indicated by their size in kb.


A
B
C
9·9
9·5
7·5
3·5
D
E
F
3·3
2·7
2·3
2·1
1·9
1·8

the presence of a 1.5 kb deletion extending from 21.1 kb to 22.6 kb from the left end of 186, i.e., with a right end point between the *Sal*I site at 22.5 and the *Pst*I site at 22.7 kb from the left end.

Digestion of 186*del*2 with *Pst*I (Track F) showed that 186*del*2 DNA had more than a simple deletion. Again the same number of fragments was present as in 186, but in this case two fragments, the 3.3 kb fragment and the adjacent 2.3 kb fragment had been affected, resulting in two new fragments, one co-electrophoresing with the affected 1.8 kb band of 186*del*1 and a second novel band of size 2.7 kb. Presumably the *Pst*I site between the *Pst*I 3.3 kb and 2.3 kb fragments of 186 had not been removed, but both the 3.3 kb and the 2.3 kb fragments had been the subject of a DNA deletion or insertion.

There are two possible explanations for these results. The first is that 186*del*2 carried two deletions, one resulting in the reduction of the 3.3 kb fragment to 2.7 kb and the other resulting in the reduction of the 2.3 kb fragment to 1.8 kb. The 1.8 kb fragment could then have been the same size as that of 186*del*1 only by chance, as the 186*del*1 1.8 kb fragment has been shown to be derived from the 3.3 kb fragment. The more attractive explanation is that 186*del*2 carried the same deletion of the 3.3 kb fragment as did 186*del*1, resulting in the 1.8 kb fragments having identical mobilities, but that in addition there was a small segment of DNA inserted into the adjacent 2.3 kb fragment to increase its size to 2.7 kb. This would explain the same reduction in the 3.3 kb fragment to a 1.8 kb fragment as in 186*del*1 and in addition predicts that the 2.7 kb fragment derived from

the 2.3 kb fragment by the addition of a 0.4 kb DNA insert.

To resolve these possibilities, double digestion of 186*del*1 and 186*del*2 with *Pst*I and either *Xho*I or *Bgl*II were carried out. Fig.8.22 shows that the single *Xho*I cleavage site mapped 20.5 kb from the left end of 186 DNA, placing it within the 3.3 kb *Pst*I fragment, but to the left of the 186*del*1 insert beginning at 21.1 kb and extending rightward. It also shows that the single *Bgl*II site mapped within the adjacent 2.3 kb fragment.

The two alternative explanations described above could be tested by determining which of the new 1.8 kb and 2.7 kb fragments of 186*del*2 were cut by *Xho*I and *Bgl*II. The first explanation predicts that the 1.8 kb and 2.7 kb fragments derive from the 2.3 kb and 3.3 kb fragments respectively and therefore that *Bgl*II would cleave within the 1.8 kb fragment and *Xho*I within the 2.7 kb fragment. The second explanation predicts the opposite, i.e., that the 1.8 kb and 2.7 kb fragments derive from the 3.3 kb and 2.3 kb fragments respectively and therefore that *Xho*I would cleave within the 1.8 kb fragment and *Bgl*II within the 2.7 kb fragment.

Fig.8.21 shows clearly that the second explanation is correct. *Xho*I cleavage of both 186*del*1 and 186*del*2 DNA (Tracks B and C) occurred within the 1.8 kb fragment while *Bgl*II cleavage (Tracks E and F) occurred within the unaltered 2.3 kb fragment of 186*del*1 and within the 2.7 kb fragment of 186*del*2. Thus the 2.7 kb fragment derived from the 2.3 kb fragment by the insertion of 0.4 kb of DNA.

It is concluded, therefore that 186*del*2 contains the

FIGURE 8.21. Double Digests of 186*del*1 and 186*del*2.

186*del*1 and 186*del*2 DNAs were digested to completion with *Pst*I and either *Xho*I or *Bgl*II, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks are as follows:

A: *Pst*I digest of 186*del*1 DNA

B: *Pst*I/*Xho*I double digest of 186*del*1 DNA

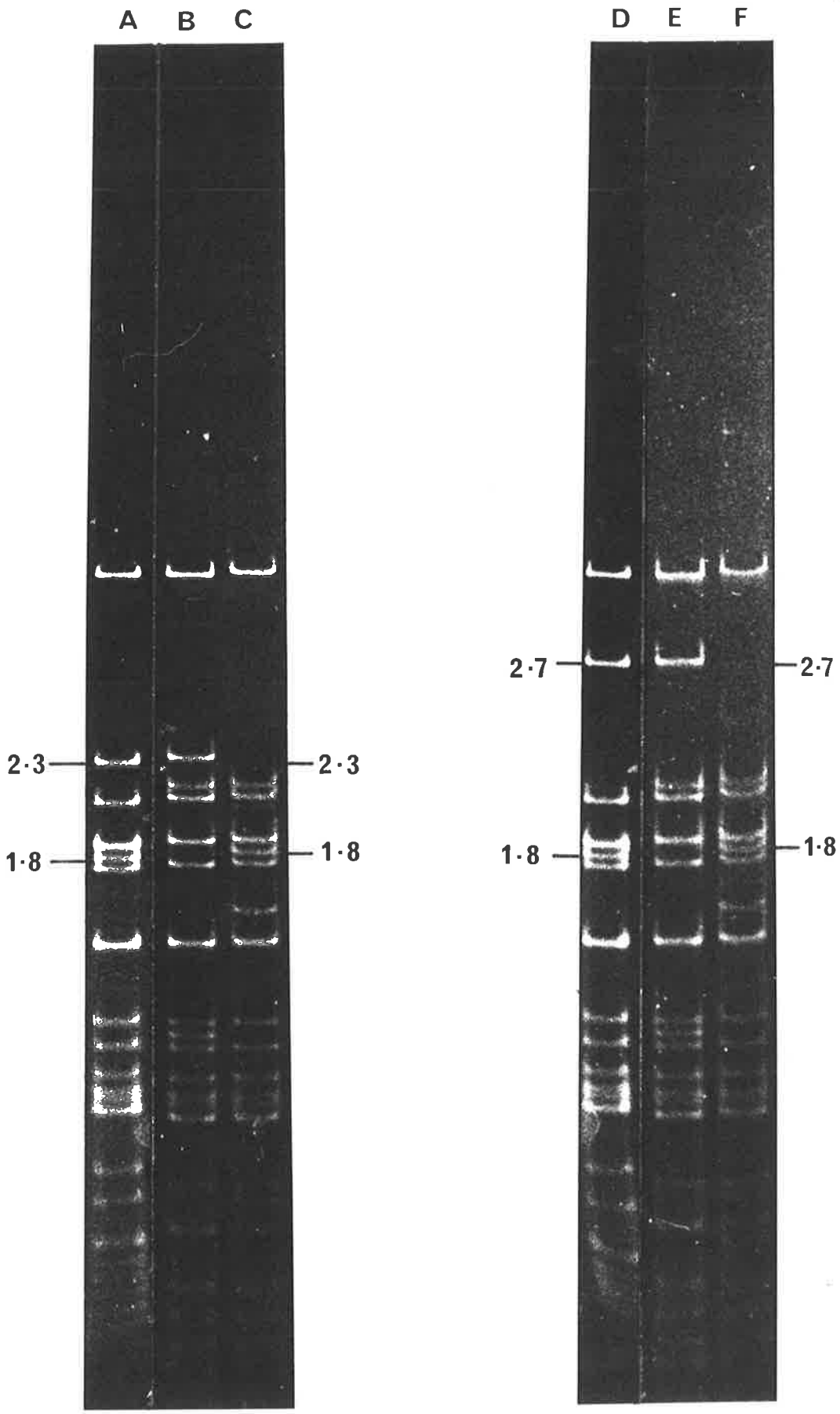
C: *Pst*I/*Bgl*II double digest of 186*del*1 DNA

D: *Pst*I digest of 186*del*2 DNA

E: *Pst*I/*Xho*I double digest of 186*del*2 DNA

F: *Pst*I/*Bgl*II double digest of 186*del*2 DNA.

The relevant fragments are indicated by their size in kb.


A
B
C
D
E
F
2·3
1·8
2·3
1·8
2·7
1·8
2·7
1·8

same deletion as 186*del*1, i.e., of 1.5 kb within the *Pst*I 3.3 kb fragment, and that in addition it carries a previously undetected 0.4 kb long DNA sequence inserted into the *Pst*I 2.3 kb fragment (summarized in Fig.8.22). The significance of this result is discussed in Chapter 9.

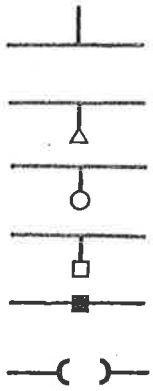
5. Molecular cloning of *Hin*dIII and *Eco*RI generated fragments of 186 DNA:

One of the major uses of restriction maps of DNA is that particular fragments can be cloned into unrelated vectors for independent analysis and large scale preparation. This approach is discussed at some length in Section I with reference to the study of feather keratins. Its usefulness is of no less value to the study of 186 and P2. For this reason the *Eco*RI and *Hin*dIII fragments of 186 were cloned into unrelated λ vectors. This work was done in co-operation with S.M. Hocking, who cloned the 186 *Eco*RI fragments into λ641 (Murray *et al.*, 1975). I shall describe the cloning of a 186 *Hin*dIII fragment into λ728 (Murray *et al.*, 1975).

Digestion of 186 DNA with *Hin*dIII generates four types of fragments, the 11.6 and 11.3 kb end pieces, the 7.1 kb middle piece and the 22.9 kb annealed end pieces. Both the 11.6 and the 11.3 kb fragments have at one end the 19 base single-stranded 186 end which cannot anneal and ligate with *Hin*dIII sites. The 7.1 kb and 22.9 kb fragments, however have the four base single-stranded *Hin*dIII generated ends, so can be joined by T_4 polynucleotide ligase to the *Hin*dIII generated ends of the λ vector. However because of the restrictions on the size of the λ molecule packaged into the phage head, a λ728 molecule

FIGURE 8.22. Structures of 186*del*1 and 186*del*2 Genomes.

Diagrammatic representation of the structures of the 186*del*1 and 186*del*2 genomes interpreted from the results presented in Chapter 8. The horizontal line represents the linear 186 genome of length 30 kb. Fragment sizes are expressed in kb. The relevant symbols are as follows:



site of *Sal*I cleavage

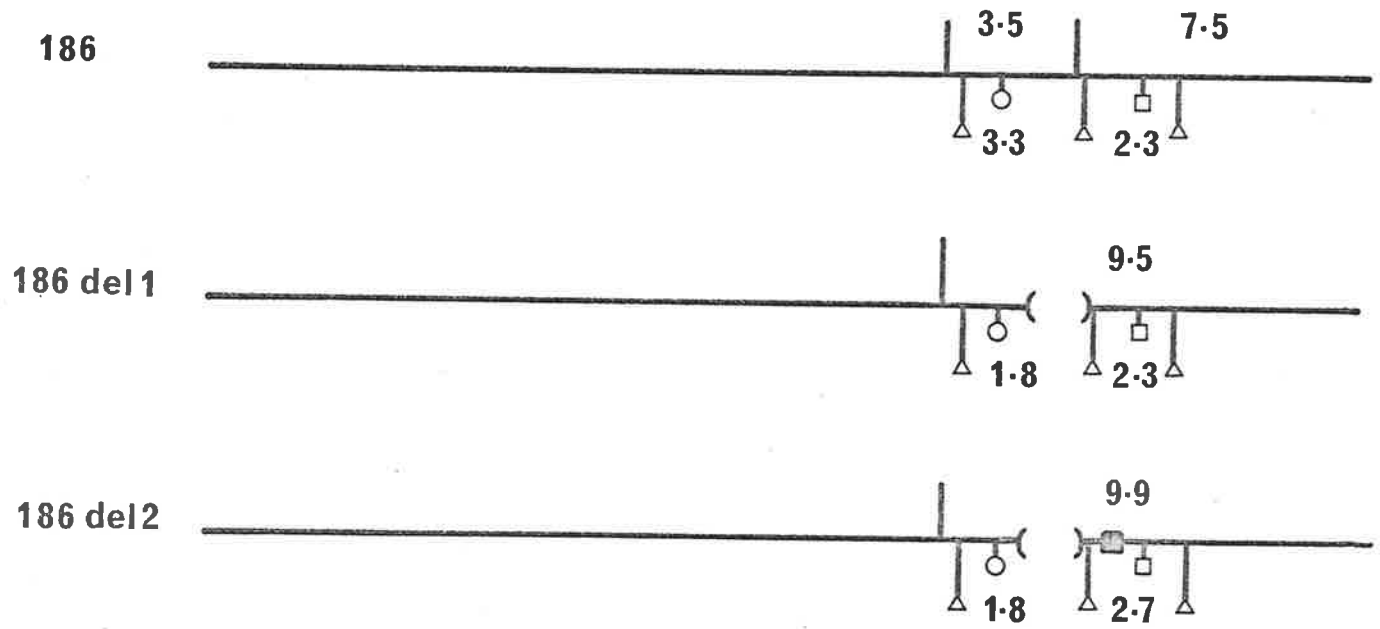
site of *Pst*I cleavage

site of *Xho*I cleavage

site of *Bgl*II cleavage

inserted segment

deleted segment


186
3·5
7·5
3·3
2·3
186 del 1
9·5
1·8
2·3
186 del 2
9·9
1·8
2·7

recombined with the 22.9 kb fragment would be inviable. Every viable phage, therefore, should either be a λ728 parental phage regenerated or a λ728 phage into which the 7.1 kb fragment has been inserted at the single *Hin*dIII site. The recombinants can be readily differentiated from parentals, as the *Hin*dIII site within λ728 lies within the *c*I gene. Incorporation of a fragment thus renders the repressor non-functional resulting in clear plaque morphology as opposed to the turbid plaque morphology of the parental phage.

186 DNA and λ728 DNA were digested with *Hin*dIII, ligated and transfected as described in Chapter 7. A single resultant clear plaque was selected from the transfectants, purified and grown as described in Chapter 7. To determine whether the recombinants carried the desired phage, DNA was prepared from this recombinant and from two recombinants isolated by Hocking (1977) thought to carry *Eco*RI generated 186 fragments. The DNA of these three putative recombinants and from λ728 or λ641 vector DNA and 186 DNA were digested with appropriate enzymes and electrophoresed as described above. Fig.8.23 shows the results of this analysis. Track A shows the two fragments of λ728 DNA generated by *Hin*dIII. (The largest fragment consists of the two fragments annealed via the cohesive single-stranded ends.). Track B shows that the putative recombinant containing the 7.1 kb fragment of 186 did have an inserted fragment co-electrophoresing with the *Hin*dIII 7.1 kb fragment of 186 in the adjacent Track C. This recombinant was named λp186*shnIII*1-2.

FIGURE 8.23. Restriction Analysis of λp186 Phage.

Vector λ phage and putative recombinant phage carrying *Hin*dIII or *Eco*RI generated 186 fragments were digested to completion with *Hin*dIII or *Eco*RI, electrophoresed in 1% agarose, stained with ethidium bromide and photographed under UV light. Tracks are as follows:

A: *Hin*dIII digest of λ728 DNA

B: *Hin*dIII digest of λ728 recombinant carrying the *Hin*dIII-generated 7.1 kb fragment of 186

C: *Hin*dIII digest of 186 DNA

D: *Eco*RI digest of 186 DNA

E: *Eco*RI digest of λ641 recombinant carrying the *Eco*RI-generated 3.1 kb fragment of 186

F: *Eco*RI digest of λ641 DNA

G: *Eco*RI digest of λ641 recombinant carrying the *Eco*RI-generated 3.3 kb fragment of 186

H: *Eco*RI digest of 186 DNA

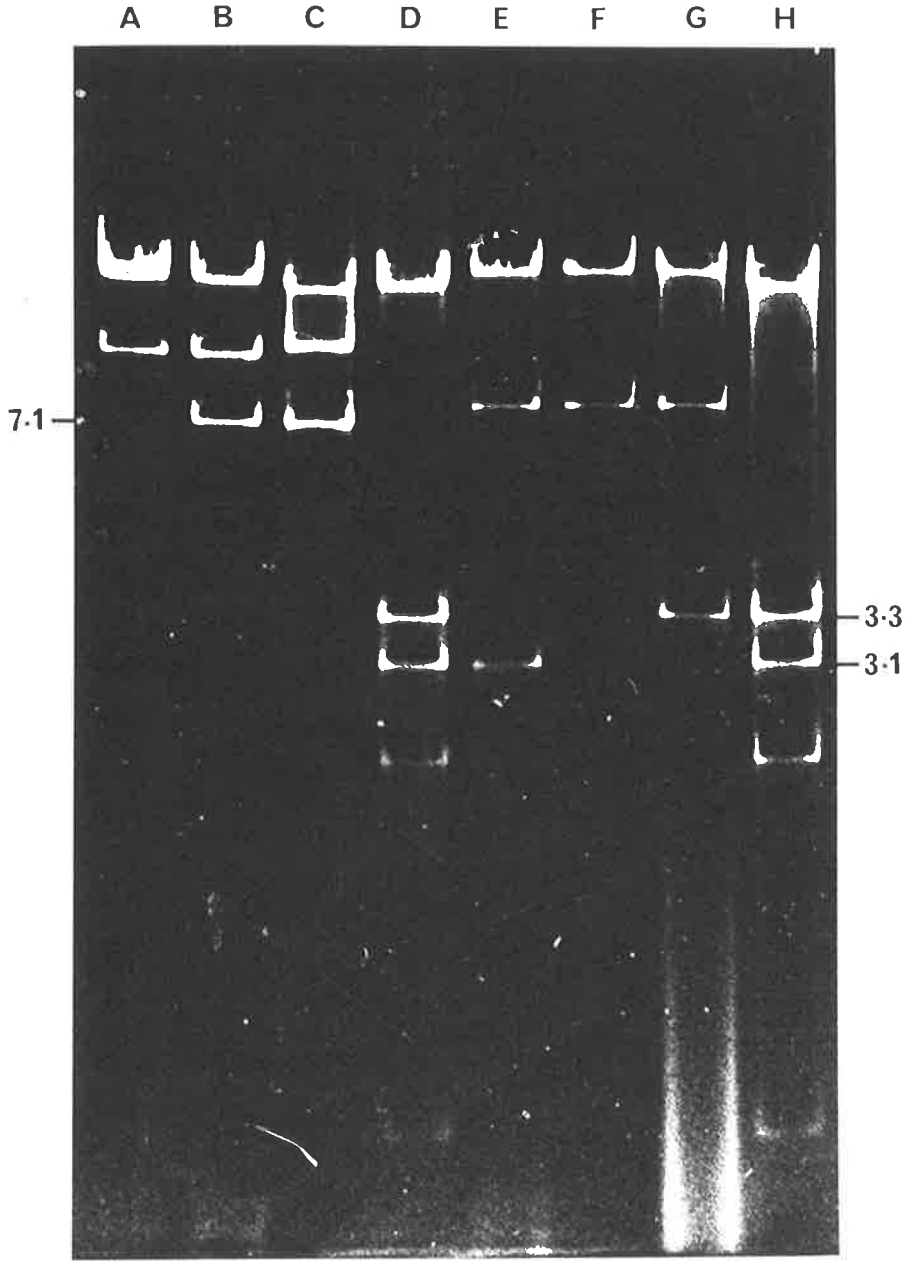

A
B
C
D
E
F
G
H
7·1
3·3
3·1

Fig.8.23 Track F shows the two fragments generated by *Eco*RI cleavage of the vector λ641. Tracks E and G show the two putative clones of *Eco*RI generated 186 fragments. Both recombinants carried a single inserted fragment co-electrophoresing with the *Eco*RI 3.1 kb annealed end-fragments (Track E) and the *Eco*RI 3.3 kb fragment (Track G) in the adjacent Tracks D and H of *Eco*RI digested 186 DNA.

These two recombinant phages were termed λp186*srI*3-1 and λp186*srI*1-2 respectively.

Since then further extensive cloning and analysis of restriction fragments has been carried out (E.J. Finnegan and J.B. Egan, 1979) and will be discussed in the following Chapter.

CHAPTER 9

DISCUSSION

CHAPTER 9.

DISCUSSION.

A. A Restriction Cleavage Map of 186, P2 and two 186 Insertion Mutants.

Section II of this thesis has described the establishment and some applications of a restriction endonuclease cleavage map of coliphages 186 and P2. A diagrammatic summary of these mapping data is given in Fig.8.15(c) and Fig.8.16(b) for 186 and P2 respectively. It can be clearly seen that use of the enzymes *Bam*HI, *Bgl*II, *Eco*RI, *Hin*dIII, *Pst*I, *Sal*I, *Xba*I and *Xho*I either individually or in combination can result in the generation of an extensive range of fragments of defined sequence and location from the 186 and P2 genomes respectively. Of particular interest to the study of 186 is the major control region known to be located in the right 35% of 186 (i.e., that portion represented in Hy5). Studies of this region will benefit greatly from this knowledge of sites of restriction cleavage, since 10 cleavage sites have been mapped within this 10.5 kb region. In addition, two insertion mutants of 186, *viz.* 186*ins*2 and 186*ins*3 have been shown by restriction cleavage analysis and DNA annealing to characterized probes to be carrying one copy each of the insertion sequence element IS3. This element is inserted 22.1 kb and 21.1 kb respectively from the left end of 186*ins*2 and 186*ins*3 and contains two *Hin*dIII restriction sites. This allows further subdivison of the control regions of 186, so that the entire 10.5 kb region can be dissected into specific

segments of 1.6 kb or less. This will facilitate the types of approaches to viral functions described in Chapter 7 and discussed briefly for 186 below.

B. Characterization of two 186 deletion mutants.

The analysis of 186*del*1 and 186*del*2 with the restriction endonucleases *Sal*I and *Pst*I revealed a great deal of information about these particular mutants. *Sal*I and *Pst*I digestion of 186*del*1 suggested that a 1.5 kb segment of the 186 genome had been deleted, as shown by the reduction in size of the affected *Pst*I fragment. In addition this deletion could be fairly precisely mapped, as it removed the *Sal*I site, 22.5 kb from the left end, but did not affect the *Pst*I site 22.7 kb from the left end. The deletion, therefore, extended from approximately the 21.1 kb mark to the 22.6 kb mark (given the limits of the restriction fragment size estimates.).

186*del*2 was not, however, a simple deletion. Digestion with *Pst*I showed that two of the 186 fragments (the 3.3 kb and 2.3 kb fragments) were affected and that two new fragments of size 2.7 kb and 1.8 kb were generated. Digestion of these fragments with enzymes cleaving either the 3.3 kb fragment or the 2.3 kb fragment showed that the 1.8 kb fragment derived from the 3.3 kb fragment (as in 186*del*1) and the 2.7 kb fragment derived from the 2.3 kb fragment. These results strongly suggest that 186*del*2 was deleted from the 21.1 kb mark to the 22.6 kb mark, as was 186*del*1, but in addition that a 0.4 kb segment was inserted somewhere to the right within the *Pst*I 2.3 kb fragment. This explanation is particularly attractive,

since 186*del*1 and 186*del*2 were isolated from the same culture (Dharmarajah, 1975). It is easy to envisage a two-step process of mutation, the first deleting 1.5 kb of DNA to produce 186*del*1 and the second inserting 0.4 kb DNA into this deleted DNA after at least one cycle of replication to produce 186*del*2. An alternative explanation can be formulated that proposes an extended deletion removing the *Pst*I site and a larger segment of DNA inserted adjacent to it which by chance contains a *Pst*I site, but this must be considered highly improbable, particularly as the 1.8 kb fragments derived from the deleted *Pst*I 3.3 kb fragments of both 186*del*1 and 186*del*2 electrophoresed with identical mobilities.

As described in Chapter 8, the presence of a 5% deletion, in 186*del*2, removing the repressor binding site (the operator) but not the RNA polymerase binding site (the promoter) prompted Dharmarajah (1975) to propose that the operator lay to the left of the promoter in a sequence operator-promoter-structural genes, with transcription proceeding rightward. Bertani and Bertani (1975) also proposed such an arrangement for the 'early' operon of P2.

The results presented in Chapter 8 make this proposal unnecessary for 186, as the deletion in 186*del*2, being the same as that in 186*del*1 affected only the *c*I gene and not the operator. Presumably the 0.4 kb addition to the right of this deletion in 186*del*2 specifically disrupts the operator. Such a mutation is independent of the location of the operator in the operon and hence does not

argue against the usual arrangement of promoter-operater-structural genes (with transcription proceeding rightward), where repressor binding to the operater presents a physical block to progress of the RNA polymerase molecule during transcription.

These results cannot lend support to the interpretation of the structure of the P2 'early' genes operon by Bertani and Bertani (1974) described above. Their interpretation relies on the sensitivity of heteroduplex analysis and electron microscopic visualization to determine the nature of the deletions and additions they used in the study. A more satisfactory approach would be the very direct one of using repressor and RNA polymerase binding to the DNA coupled with sequence analysis of the region to determine the sites of interaction between these proteins and the DNA at the molecular level.

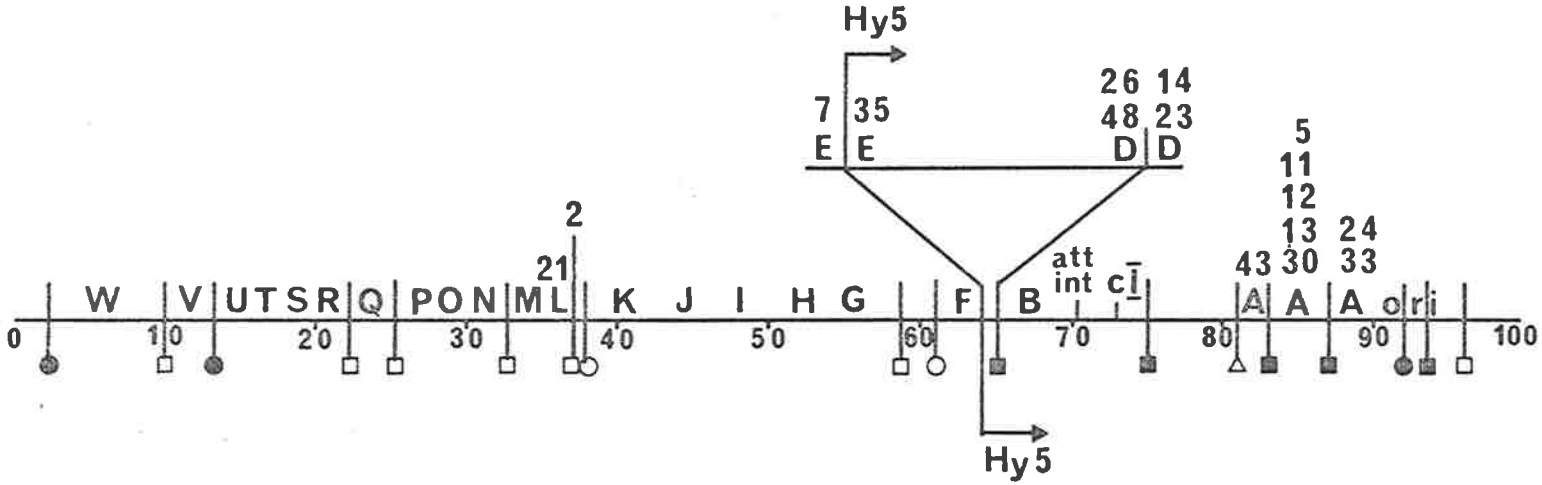
C. Use of the Restriction Map in the Cloning and Analysis of 186 DNA.

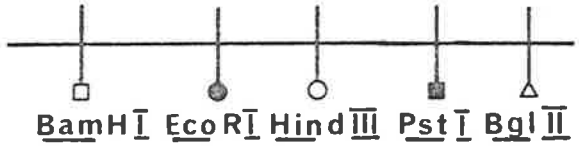
In vitro recombinant DNA technology, described in detail in Section I allows the molecular cloning of restriction fragments into unrelated vectors and the preparation of large amounts of the fragments free of any other fragments of the original DNA. A knowledge of the location and length of such fragments is provided by restriction site mapping and can be used in the physical localization of genes, RNA transcripts, replicated DNA etc., as described in Chapter 6.

Chapter 8 described the cloning and identification of the *Hin*dIII 7.1 kb fragment of 186 in the vector λ728.

FIGURE 9.1. Physical Map of the Genes of 186.

The results of physically mapping the 186 genes onto the linear 186 genome using molecular cloning and gene-rescue experiments (Finnegan and Egan, 1979) based on the restriction endonuclease cleavage mapping presented in this thesis. (Reproduced from Finnegan and Egan, 1979).





Cleavage sites

Two other recombinants prepared by Hocking (1977) were characterized and identified as the λ641 vector containing the *Eco*RI generated 186 3.1 kb and 3.3 kb fragments respectively. E.J. Finnegan and J.B. Egan (Manuscript in preparation) have cloned a series of *Bam*HI, *Bgl*II, and *Pst*I generated fragments into the plasmid vector pBR322. These fragments are now being used to analyse RNA transcription from the 186 genome (E.J. Finnegan, personal communication). The most recent example of the use of these fragments in the study of 186 is the physical mapping carried out by Finnegan and Egan (1979) of the genes of 186 (summarized in Fig.6.1) to the cloned restriction fragments described above. A summary of these results is shown in Fig.9.1 merely as an example of the major advances that have been made possible by the availability of a restriction site map and that will increase greatly our understanding of the molecular biology of these coliphages.

REFERENCES

REFERENCES

Allet, B. and Bukhari, A.I. (1975) *J. Mol. Biol.* 92, 529-540.

Alwine, J.C., Kemp, D.J. and Stark, G.R. (1977) *Proc. Nat. Acad. Sci., U.S.A.* 74, 5350-5354.

Appels, R. and Wells, J.R.E. (1972) *J. Mol. Biol.* 70, 425-434.

Appleyard, R.K. (1954) *Genetics* 39, 440-452.

Arber, W. and Dussoix, D. (1962) *J. Mol. Biol.* 5, 18-36.

Baldwin, R.L., Barrand, P., Fritsch, A., Goldthwait, D.A. and Jacob, F. (1966) *J. Mol. Biol.* 17, 343-357.

Bartels, E.A. (1943) *Arch. f. Entwicklmech.* 142, 763-816.

Beckingham Smith, K. (1973a) *Dev. Biol.* 30, 249-262.

Beckingham Smith, K. (1973b) *Dev. Biol.* 30, 263-278.

Benton, W.D. and Davis, R.W. (1976) *Science* 196, 180-182.

Berget, S.M., Moore, C. and Sharp, P.A. (1977) *Proc. Nat. Acad. Sci., U.S.A.* 74, 3171-3175.

Bertani, G. (1968) in *'Molecular Genetics'* (H. Wittman and N. Schuster, Eds.) Springer-Verlag, Berlin, pp. 180-186.

Bertani, G. and Weigle J.J. (1953) *J. Bacteriol.* 65, 113-121.

Bertani, L.E. (1968) *Virol.* 36, 87-103.

Bertrand, R., Korn, L., Lee, F., Platt, T., Squires, C.L., Squires, C. and Yanofsky, C. (1975) *Science* 189, 22-26.

Blattner, F.R., Williams, W.G., Blechl, A.E., Denniston-Thompson, K., Faber, H.E., Furlong, L., Grunwald, D.J., Kieper, D.O., Moore, D.D., Schumm, J.W., Sheldon, E.L. and Smithies, O. (1977) *Science* 196, 161-169.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heymeker, H.L. and Boyer, H.W. (1977) *Gene* 2, 95-113.

Bourgeois, S. and Pfahl, M. (1976) *Advan. Protein Chem.* 30, 1-99.

Brack, C. and Tonegawa, S. (1977) *Proc. Nat. Acad. Sci., U.S.A.* 74, 5652-5656.

Bradley, C. (1972) B.Sc.(Hons.) Thesis, University of Adelaide.

Bradley, C., Ong, P.L. and Egan, J.B. (1975) *Mol. Gen. Genet.* 140, 123-135.

Britten, R.J. and Davidson, E.H. (1969) *Science* 165, 349-357.

Bron, S., Murray, K. and Trautner, T. (1975) *Mol. Gen. Genet.* 143, 13-23.

Brown, D.D., Wensink, P.C. and Jordon, E. (1971) *Proc. Nat. Acad. Sci., U.S.A.* 68, 3175-3179.

Brown, D.D., Wensink, P.C. and Jordon, E. (1972) *J. Mol. Biol.* 63, 57-73.

Brown, K.D., Bennett, G.N., Lee, F., Schweingruber, M.E. and Yanofsky, C. (1978) *J. Mol. Biol.* 121, 153-177.

Brownlee, G.G., Cartwright, E.M. and Brown D.D. (1974) *J. Mol. Biol.* 89, 703-718.

Bukhari, A.I., Shapiro, J.A. and Adhya, S.L. (Eds) (1977) *'DNA Insertion Elements, Plasmids and Episomes'* Cold Spring Harbor Laboratory.

Callan, H.G. and Lloyd, L. (1960) *Phil. Trans. Roy. Soc. B.* 243, 135-219.

Campbell, A. (1965) *Virol.* 27, 329-339.

Carroll, D. and Brown, D.D. (1976a) *Cell.* 7, 467-475.

Carroll, D. and Brown, D.D. (1976b) *Cell.* 7, 477-486.

Carter, B.J., Khoury, G. and Denhardt, D.T. (1975) *J. Virol.* 16, 559.

Chang, S. and Cohen, S.N. (1977) *Proc. Nat. Acad. Sci., U.S.A.* 74, 4811-4815.

Chow, L.T., Gelinas, R.E., Broker, T.R. and Roberts, R.J. (1977) *Cell* 12, 1-8.

Cohen, S. (1962) *J. Biol. Chem.* 237, 1555-1562.

Cohen, S. (1965) *Dev. Biol.* 12, 394-407.

Cohen, S and Stastny, M. (1968) *Biochim. Biophys. Acta.* 166, 427-437.

Cohn, R.H., Lowry, J.C. and Kedes L.H. (1976) *Cell*, 9, 147-161.

Danna, K.J. and Nathans, D. (1972) *Proc. Nat. Acad. Sci., U.S.A.* 69, 3097-3100.

Danna, K.J., Sack Jr., G.H. and Nathans, D. (1973) *J. Mol. Biol.* 78, 363-376.

Dayhoff, M.O. (1972) *'Atlas of Protein Sequence and Structure'* (Washington D.C.: National Biomedical Research Foundation).

De Fillipes, F.M. (1974) *Biochem. Biophys. Res. Comm.* 58, 586-596.

Denhardt, D.T. (1966) *Biochem Biophys. Res. Comm.* 23, 641-646.

Dharmarajah, V.K. (1975) B.Sc.(Hons.) Thesis, University of Adelaide.

Dhouailly, D. (1967) *J. Embryol. Exptl. Morphol.* 18, 389-400.

Dhouailly, D., Rogers, G.E. and Sengel, P. (1978) *Dev. Biol.* 65, 58-68.

Dodson, J.W. (1963) *Exptl. Cell. Res.* 31, 233-235.

Dugaiczyk, A., Woo, S.L.C., Lai, E.C., Mace Jr., M.L., McReynolds, L. and O'Malley, B.W. (1978) *Nature*, 274, 328-333.

Dussoix, D. and Arber, W. (1962) *J. Mol. Biol.* 5, 37-49.

Edgell, M.H., Hutchison III, C.A. and Sclair, M. (1972) *J. Virol.* 9, 574-582.

Fell, H.B. (1957) *Proc. Roy. Soc. (Biol.)* 146, 242-256.

Fell, H.B. (1962) *J. Embryol. Exptl. Morphol.* 10, 389-409.

Fell, H.B. and Mellanby, E. (1953) *J. Physiol. (London)* 119, 470-488.

Fiandt, M., Szybalski, W. and Malamy, M.H. (1972) *Mol. Gen. Genet.* 119, 223-231.

Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., Van de Voorde, A., Van Heuverswyn, H., Van Herreweghe, J., Volckaert, G. and Ysebaert, M. (1978) *Nature*, 273, 113-120.

Finnegan, E.J. and Egan, J.B. (1979) *Mol. Gen. Genet.* in press.

Fitton-Jackson, F.L. and Fell, H.B. (1963) *Dev. Biol.* 7, 394-419.

Flavell, R.A., Kooter, J.M., De Boer, E., Little, P.F.R. and Williamson, R. (1978) *Cell* 15, 25-41.

Fraser, R.D.B., MacRae, T.P. and Rogers, G.E. (1972) *'Keratins, Their Composition, Structure and Biosynthesis.'*

Gibbs, P.E.M. (1977) Ph.D. Thesis, University of Adelaide.

Gilbert, W. (1978) *Nature*, 271, 501.

Gillespie, D. and Spiegelman, S. (1965) *J. Mol. Biol.* 12, 829-842.

Goddard, D.R. and Michaelis, L. (1934) *J. Biol. Chem.* 106, 605-614.

Goff, S.P. and Berg, P. (1978) *Proc. Nat. Acad. Sci., U.S.A.* 75, 1763-1767.

Greene, M.M. (1973) *Genetics* 73, (Suppl.), 187-194.

Gross, K., Probst, E., Schaffner, W. and Birnstiel, M.L. (1976a) *Cell* 8, 455-469.

Gross, K., Schaffner, W., Telford, J. and Birnstiel, M.L. (1976b) *Cell* 8, 479-484.

Gross-Bellard, M., Oudet, P. and Chambon, P. (1973) *Eur. J. Biochem.* 36, 32-38.

Grunstein, M.L. and Hogness, D.S. (1975) *Proc. Nat. Acad. Sci. U.S.A.* 72, 3961-3965.

Gurdon, J.B. (1964) *Adv. Morph.* 4, 1-43.

Hansen, M.A. and Marzluf, G.A. (1975) *Proc. Nat. Acad. Sci., U.S.A.* 72, 1240-1244.

Harrap, B.S. and Woods, E.F. (1964a) *Biochem. J.* 92, 8-18.

Harrap, B.S. and Woods, E.F. (1964b) *Biochem, J.* 92, 19-26.

Hirsch, H., Starlinger, P. and Brachet, (1972) *Mol. Gen. Genet.* 119, 191-206.

Hocking, S.M. (1977) Ph.D. Thesis, University of Adelaide.

Hoober, J.K. and Cohen, S. (1967) *Biochim. Biophys. Acta.* 138, 347-356.

Hozumi, N. and Tonegawa, S. (1976) *Proc. Nat. Acad. Sci., U.S.A.* 73, 3628-3632.

Jackson, D.A., Symons, R.H. and Berg. P. (1972) *Proc. Nat. Acad. Sci., U.S.A.* 69, 2904-2909.

Jacob, F. and Monod, J. (1961) *J. Mol. Biol.* 3, 318-356.

Jacob, F. and Wollman, E.L. (1956) *Ann. Inst. Pasteur* 91, 486-510.

Jacobsen, H., Klenow, H. and Overgaard-Hansen, K. (1974) *Eur. J. Biochem.* 45, 623-627.

Jacq, C., Miller, J.R. and Brownlee, G.G. (1977) *Cell* 12, 109-120.

Jeffreys, A.J. and Flavell, R.A. (1977) *Cell* 12, 1097-1108.

Jeppesen, P.G.N., Sanders, L. Slocombe, P.M. (1976) *Nucleic Acids Res.* 3, 1323-1339.

Kafatos, F.C., Maniatis, T., Efstratiadis, A., Sim Gek Kee, Regier, J.R. and Nadel, M. (1977) in *'Organization and Expression of the Eukaryotic Genome'* (Eds. E.M. Bradbury and K. Javaherian, Academic Press Inc., London) pp. 393-420.

Kafatos, F.C., Efstratiadis, A., Goldsmith, M.R., Jones, C.W., Maniatis, T., Regier, J.C., Rodakis, G., Rosenthal, N., Sim Gek Kee, Thireos, G. and Villa-Komaroff, L. (1978) In *'Differentiation and Development'*, Miami Winter Symposia Vol. 15 (F. Ahmed, T.R. Russell, J. Schultz and R. Werner, Eds., Academic Press, New York) pp. 299-315.

Karp, R. and Hogness, D.S. (1976) *Fed. Proc.* 35, 1623.

Kedes, L.H., Cohn. R., Lowry, J., Chang, A.C.Y. and Cohen, S. N. (1975) *Cell* 6, 359-370.

Kelly Jr., T.J. and Smith, H. . (1970) *J. Mol. Biol.* 51, 393-409.

Kemp, D.J. (1975) *Nature*, 254, 573-577.

Kemp, D.J. and Rogers, G.E. (1972) *Biochem.* 11, 969-975.

Kemp, D.J., Dyer, P.Y. and Rogers, G.E. (1974a) *J. Cell. Biol.* 64, 1114-1131.

Kemp, D.J., Partington, G.A. and Rogers, G.E. (1974b) *Biochem. Biophys. Res. Comm.* 60, 1006-1014.

Kemp, D.J., Schwinghamer, M.W. and Rogers, G.E. (1974c) *Mol. Biol. Reports*, 1, 441-446.

Khoury, G., Martin, M.A., Lee, T.N.H., Danna, K.J. and Nathans, D. (1973) *J. Mol. Biol.* 78, 377-389.

Kitano, Y. and Kuroda, Y. (1967) *Exptl. Cell. Res.* 48, 350-360.

Lewis, J.B., Atkins, J.F., Baum, P.R., Solem, R. and Gesteland, R.F. (1976) *Cell* 7, 141-151.

Lifton, R.P., Goldberg, M.L., Karp, R.W. and Hogness, D.S. (1977) *Cold Spring Harbor Symp. Quant. Biol.* 42, 1047-1052.

Lillie, F.R. (1965) *'Lillie's Development of the Chick'* Hamilton, H.L. Ed., Holt, Rinehart and Winston, New York) pp. 546-606.

Lindahl, G. (1969) *Virol.* 39, 839-860.

Lindahl, G. (1971) *Virol.* 46, 620-633.

Linn, S. and Arber, W. (1968) *Proc. Nat. Acad. Sci., U.S.A.* 59, 1300-1306.

Lobban, P.E. and Kaiser, A.D. (1973) *J. Mol. Biol.* 78, 453-471.

Luria, S.E. and Human, M.L. (1952) *J. Bact.* 64, 557-569.

McClintock, B. (1956) *Cold Spring Harbor Symp. Quant. Biol.* 21, 197-216.

Malamy, M.H., Fiandt, M. and Szybalski (1972) *Mol. Gen. Genet.* 119, 207-222.

Mandel, J.L., Breathnach, R., Gerlinger, P., Le Meur, M., Gannon, F. and Chambon, P. (1978) *Cell*, 14, 641-653.

Mandel, M. and Berg, A. (1968) *Proc. Nat. Acad. Sci., U.S.A.* 60, 265-268.

Maniatis, T. Jeffrey, A. and Kleid, D.G. (1975a) *Proc. Nat. Acad. Sci., U.S.A.* 72, 1184-1188.

Maniatis, T., Jeffrey, A. and Van de Sande, H. (1975b) *Biochemistry* 14, 3787-3794.

Maniatis, T., Hardison, R.C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G.K. and Efstratiadis, A. (1978) *Cell*, 15, 687-701.

Matulionis, D.H. (1970) *Z. Anat. Entwickl. Gesch.* 132, 107-157.

Maxam, A.M. and Gilbert, W. (1977) *Proc. Nat. Acad. Sci., U.S.A.* 74, 560-564.

Meselson, M. and Yuan, R. (1968) *Nature*, 217, 1110-1114.

Model, P., Horiuchi, K., McGill, C. and Zinder, N.D. (1975) *Nature* 253, 132-134.

Morris, C.P. and Rogers, G.E. (1979) *Mol. Biol. Reports*, In press.

Murray, N.E., Batten, P.L. and Murray, K. (1973) *J. Mol. Biol.* 81, 395-407.

Murray, N.E., Brammar, W.J. and Murray K. (1977) *Mol. Gen. Genet.* 150, 53-61.

Murray, K and Murray, N.E. (1973) *Nature New Biol.* 243, 134-139.

Nosikov, V.V., Braga, E.A., Karlishev, A.V., Zhuza, A.L. and Polyanovsky, O.L (1976) *Nucleic Acids Res.* 3, 2293-2301.

Nosikov, V.V. and Sain, B. (1977) *Nucleic Acids Res.* 4, 2263-2273.

Noyes, B.E. and Stark, G.R. (1975) *Cell* 5, 301-310.

O'Donnell, I.J. (1973) *Aust. J. Biol. Sci.* 26, 415-437.

O'Donnell, I.J. and Inglis, A.S. (1974) *Aust. J. Biol. Sci.* 27, 369-382.

Padmanabhan, R. and Wu, R. (1972) *J. Mol. Biol.* 65, 447-467.

Panyim, S. and Chalkley, R. (1969) *Arch. Biochem. Biophys.* 130, 337-346.

Partington, G.A., Kemp, D.J. and Rogers, G.E. (1973) *Nature New Biol.* 246, 33-36.

Paul, J. (1970) *'Cell and Tissue Culture'* 4th Edition (E. and S. Livingstone, Eds., Edinburgh and London) p. 91.

Peacock, A.C. and Dingman, C.W. (1968) *Biochem.* 7, 668-674.

Pinder, J.C., Staynov, D.Z. and Gratzer, W.B. (1974) *Biochem,* 13, 5373-5378.

Portmann, R., Schaffner, W. and Birnstiel, M. (1976) *Nature* 264, 31-34.

Powell, B.C., Kemp, D.J., Partington, G.A., Gibbs, P.E.M., and Rogers G.E. (1976) *Biochem. Biophys. Res. Comm.* 68, 1263-1271.

Prives, C.L., Aviv, H., Paterson, B.M., Roberts, B.E., Rozenblatt, S., Revel, M. and Winocour, E. (1974) *Proc. Nat. Acad. Sci., U.S.A.* 71, 302-306.

Ptashne, M., Backman, K., Humayun, M.Z., Jeffrey, A., Maurer, R., Meyer, B., and Sauer, R.T. (1976) *Science* 194, 156-161.

Rabbitts, T.H. and Forster, A. (1978) *Cell* 13, 319-327.

Rawles, M.E. (1963) *J. Embryol. Exptl. Morphol.* 11, 765-789.

Rawles, M.E. (1965) in *'Biology of Skin and Hair Growth'* (A.G. Lyne and B.F. Short, Eds., Angus and Robertson, Sydney). pp.105-128.

Reddy, V.B., Thimmappaya, B., Dhar, R., Subramanian, K.N., Zain, B.S., Pan, J., Ghosh, P.K., Celma, M.L., and Weissman, S.M. (1978) *Science* 200, 494-502.

Regier, J.C., Kafatos, F.C., Kramer, K.J., Heinrikson, R.L. and Kein, P.S. (1978a) *J. Biol. Chem.* 253, 1305-1314.

Regier, J.C., Kafatos, F.C., Goodfliesh, R. and Hood, L. (1978b) *Proc. Nat. Acad. Sci., U.S.A.* 75, 390-394.

Rheinwald, J.G. and Green, H. (1975a) *Cell* 6, 317-330.

Rheinwald, J.G. and Green, H. (1975b) *Cell* 6, 331-344.

Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P. (1977) *J. Mol. Biol.* 113, 237-251.

Roberts, R.J. (1976) *C.R.C. Critical Rev. in Biochem.* 4, 123-164.

Romanoff, A.L. (1960) *'The Avian Embryo'* MacMillan, New York.

Roop, D.R., Nordstrom, J.L., Tsai, S.Y., Tsai, M. and O'Malley, B.W. (1978) *Cell* 15, 671-685.

Rougeon, F. and Mach, B. (1976) *Proc. Nat. Acad. Sci., U.S.A.* 73, 3418-3422.

Roychoudhury, R., Jay, E. and Wu, R. (1976) *Nucleic Acids Res.* 3, 863-877.

Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, J.C., Hutchison III, C.A., Slocombe, P.M. and Smith, M. (1977) *Nature* 265, 687-695.

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

Schaffner, W., Gross, K., Telford, J. and Birnstiel, M.L. (1976) *Cell* 8, 471-478.

Schaffner, W., Kunz, G., Daetwyler, H., Telford, J., Smith, H.O. and Birnstiel, M.L. (1978) *Cell* 14, 655-671.

Seeburg, P.H. and Schaller, H. (1975) *J. Mol. Biol.* 92, 261-277.

Seidman, J.G., Leder, A., Edgell, M.H., Polsky, F.I., Tilghman, S., Tiemeier, D.C. and Leder, P. (1978a) *Proc. Nat. Acad. Sci., U.S.A.* 75, 3881-3885.

Seidman, J.G., Leder, A., Nau, M., Norman, B. and Leder, P. (1978b) *Science* 202, 11-17.

Sengel, P. (1957) *Experientia* 13, 177-182.

Sengel, P. (1971) *Advances in Morphogenesis* 9, 181-230.

Sim, G.K., Efstratiadis, A., Jones, C.W., Kafatos, F.C., Koehler, M., Kronenberg, H.M., Maniatis, T., Regier, J.C., Roberts, B.F. and Rosenthal, H. (1977) *Cold Spring Harbor Symp. Quant. Biol.* 42, 933-945.

Skalka, A. and Hanson, P. (1972) *J. Virol.* 9, 583-593.

Smith, G. (1973) *Cold Spring Harbor Symp. Quant. Biol.* 38, 507-513.

Smith, H.O. and Birnstiel, M.L. (1976) *Nucleic Acids Res.* 3, 2387-2398.

Smith, H.O. and Wilcox, K.W. (1970) *J. Mol. Biol.* 51, 379-391.

Southern, E.M. (1975) *J. Mol. Biol.* 98, 503-517.

Steinhart, W.L., Grafstrom, R.H. and Hill, C.W. (1975) *Biochem. Biophys. Res. Comm.* 67, 556-561.

Sugimoto, M. and Endo. H. (1969) *Nature* 222, 1270-1272.

Sun, T. and Green, H. (1978) *Cell* 14, 469-476.

Sunshine, M.G., Thorn, M., Gibbs, W., Calendar, R. and Kelly, B. (1971) *Virol.* 46, 691-702.

Sussman, R. and Jacob, F. (1962) *Compt. Rend. Acad. Sci. (Paris)* 254, 1517-1519.

Symons, R.H. (1974) in *Methods in Enzymology* (Grossman, L. and Moldare, K., Eds., Academic Press, New York), 29, 102-115.

Symons, R.H. (1977) *Nucleic Acids Res.* 4, 4347-4355.

Szybalski, W. (1977) in *'DNA Insertion Elements, Plasmids and Episomes'* (A.I. Bukhari, J.A. Shapiro and S.L.

Adhya) pp. 583-590.

Taylor, J.M., Illmensee, R. and Summers, J. (1976) *Biochim. Biophys. Acta.* 442, 324-330.

Tiemeier, D.C., Tilghman, S.M., Polsky, F.I., Seidman, J.G., Leder, A., Edgell, M.H. and Leder, P. (1978) *Cell* 14, 237-245.

Tilghman, S.M., Tiemeier, D.C., Seidman, J.G., Peterlin, B.M., Sullivan, M., Maizel, J.V. and Leder, P. (1978a) *Proc. Nat. Acad. Sci., U.S.A.* 75, 725-729.

Tilghman, S.M., Curtis, P.J., Tiemeier, D.C., Leder, P. and Weissman, C. (1978b) *Proc. Nat. Acad. Sci., U.S.A.* 75, 1309-1313.

Tonegawa, S., Hozumi, N., Matthyssens, G. and Schuller, R. (1976) *Cold Spring Harbor Symp. Quant. Biol.* 41, 877-889.

Tonegawa, S., Maxam, A.M., Tizard, R., Bernard, O. and Gilbert, W. (1978) *Proc. Nat. Acad. Sci., U.S.A.* 75, 1485-1489.

Villa-Komaroff, L., Efstratiadis, S., Broome, S., Lomedico, P. Tizard, R., Naber, S.P., Chick, W.L. and Gilbert, W. (1978) *Proc. Nat. Acad. Sci., U.S.A.* 75, 3727-3731.

Vogt, V.M. (1973) *Eur. J. Biochem.* 33, 192-200.

Voitkevich, A.A. (1966) *'The Feathers and Plumage of Birds'* (Sidgwick and Jackson, London).

Walker, I.D. and Bridgen, J. (1976) *Eur. J. Biochem.* 67, 283-293.

Walker, I.D. and Rogers, G.E. (1976a) *Eur. J. Biochem.* 69, 329-339.

Walker, I.D. and Rogers, G.E. (1976b) *Eur. J. Biochem.* 69, 341-350.

Wang, J.C. and Schwartz, H. (1967) *Biopolymers* 5, 953-966.

Watterson, R.L. (1942) *Physiol. Zool.* 15, 234-259.

Weinberg, E.S., Overton, G.C., Shutt, R.H. and Reeder, R.H. (1975) *Proc. Nat. Acad. Sci., U.S.A.* 72, 4815-4819.

Wessels, N.K. (1961) *Exptl. Cell Res.* 24, 131-142.

Wessels, N.K. (1962) *Dev. Biol.* 4, 87-107.

Wessels, N.K. (1965) *Dev. Biol.* 12, 131-153.

Wilson, C.A. and Young, F.E. (1975) *J. Mol. Biol.* 97, 123-125.

Woodin, A.M. (1954) *Biochem. J.* 57, 99-109.

Woods, E.F. (1971) *Comp. Biochem. Physiol.* 39A, 325-331.

Woods, W.H. and Egan, J.B. (1974) *J. Virol.* 14, 1349-1356.

Wu, M., Holmes, D.S., Davidson, N., Cohn, R.H. and Kedes, L.H. (1976) *Cell* 9, 163-169.

Wu, R. and Taylor, E. (1971) *J. Mol. Biol.* 57, 491-511.

Yamagashi, H., Nakamura, K. and Ozeki, H. (1965) *Biochem. Biophys. Res. Comm.* 20, 727-732.

Yoshimori, R.N. (1971) Ph.D. Thesis, University of California, San Francisco.

Younghusband, H.B., Egan, J.B. and Inman, R.B. (1975) *Mol. Gen. Genet.* 140, 101-110.

Younghusband, H.B. and Inman, R.B. (1974) *Virol.* 62, 530-538.

APPENDIX - PUBLICATIONS

APPENDIX - PUBLICATIONS

1. Papers Published

A Restriction Cleavage Map of Coliphages 186 and P2 (with J.B. Egan) (1979) *Mol. Gen. Genet.* 171, 79-89.

A Method Which Facilitates the Ordering of DNA Restriction Fragments (with J.B. Egan) (1979) *Mol. Gen. Genet.* 171, 103-106.

2. Papers Presented At Conferences

Restriction Endonuclease Cleavage Map of Coliphages P2 and 186 (with J.B. Egan) (1976) *Proc. Aust. Biochem. Soc.* 9, 47.

The Stability of Feather Keratin Gene Organisation During Development. (1978) Aust. Biochem. Soc. Specialist Conference on *'The Cell Nucleus and Cancer'*.

Molecular Cloning of Avian Keratin Sequences (with G.E. Rogers) (1979). *Proc. Aust. Biochem. Soc.* 12, 76.