



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

	Page
THESIS SUMMARY	i
STATEMENT	v
ACKNOWLEDGEMENTS	vi
ABBREVIATIONS	viii
GENERAL INTRODUCTION	1
 <u>SECTION I - THE MOLECULAR GENETICS OF DNA CODING</u> <u>FOR AVIAN KERATINS</u>	
<u>CHAPTER 1 - Introduction</u>	3
A. Introductory Comments	3
B. The Developing Keratinocyte	3
1. The developing embryonic feather	3
2. The role of the dermis in feather development	5
3. Other factors affecting embryonic feather development	6
a) Vitamin A	6
b) Hormones	6
c) Epidermal growth factor	7
4. Developing keratinocytes other than the chick feather	7
C. Gene Expression In The Developing Feather	8
1. The time course of keratin protein synthesis	8
2. The time course of mRNA synthesis	9
3. The nature of embryonic feather keratin proteins	9
4. The nature of embryonic feather keratin mRNA	12

5.	The nature of keratin-coding DNA	14
6.	Summary	14
D.	Other Multi-gene Systems	15
1.	Faithfully repeated genes	15
a)	Histone genes	15
b)	5S genes	16
2.	Heterogeneous multiple genes	18
a)	Silkmoth chorion proteins	18
b)	Immunoglobulin genes	19
3.	The molecular genetics of gene expression	21
a)	Prokaryote gene expression	21
b)	Eukaryote gene expression	22
E.	Aims of the Project	24
1.	General aims	24
2.	Specific aims of the project	25
<u>CHAPTER 2 - Materials and Methods</u>		27
A.	Materials	27
1.	Chemicals	27
a)	General chemicals	27
b)	Chemicals for electrophoresis	28
c)	Chemicals for <i>in vitro</i> DNA synthesis	28
2.	Proteins and enzymes	28
3.	Other materials	29
a)	Tissues	29
(i)	Embryonic feathers	29
(ii)	Adult erythrocyte nuclei	30
(iii)	Sperm	30
b)	Materials for enzyme preparations	30
(i)	Bacterial strains	30

(ii) Column materials	30
c) Other materials	30
B. Methods	31
1. General methods	31
a) Media	31
b) Buffers and the preparation of solutions	31
c) Determination of radioactivity	32
2. Preparation of purified keratin messenger RNA	32
3. Preparation of DNA	33
a) Chick erythrocyte DNA	33
b) Embryonic chick feather and chick sperm DNA	33
c) Plasmid DNA	34
4. Restriction endonuclease preparation and digestion conditions	35
a) Preparation of restriction endonucleases	35
b) Digestion conditions	38
c) Preparation of digested DNA for electro- phoresis	38
5. Gel electrophoresis	39
a) Agarose gel electrophoresis	39
b) Aqueous polyacrylamide gel electrophoresis	39
c) Formamide polyacrylamide gel electro- phoresis	40
6. Transfer of restricted and fractionated DNA to nitrocellulose	40
7. Preparation of labelled nucleotides	41
8. <i>In vitro</i> synthesis of labelled DNA	44
a) Oligo dT primed reverse transcription	44

b)	Random primed reverse transcription	44
c)	Nick translation of double-stranded DNA	45
9.	Filter annealing	45
a)	Pre-annealing	45
b)	Annealing	45
(i)	Single-stranded DNA	45
(ii)	Double-stranded DNA	46
c)	Washing and exposing filters	46
d)	The kinetics of annealing	46
10.	Preparation and tailing of double-stranded cDNA	47
a)	Synthesis of the first strand	47
b)	Synthesis of the second strand	47
c)	S ₁ cleavage of the dscDNA	48
d)	Size fractionation of the dscDNA	48
e)	Tailing the dscDNA	48
11.	Annealing tailed cDNA and vector DNA and transformation of <i>E. coli</i>	49
a)	Annealing to vector	49
b)	Procedure for transfection/transformation	49
12.	Detection and purification of recombinants	50
a)	Plasmid recombinants	50
b)	Phage λ recombinants	50
c)	Purification of recombinants	51
d)	Miniscreen detection of plasmid recombinants	51
13.	Electron microscopy	52
a)	Formation of heteroduplexes	52
b)	Visualization	52
14.	Containment Facilities	53

<u>CHAPTER 3 - Total 12S cDNA as a Probe For Complementary</u>	
<u>Sequences in Chick DNA</u>	54
A. Introduction	54
B. Results	57
1. No sequence repeat unit is apparent for the genes coding for 12S mRNA	57
2. Keratin gene expression does not involve DNA rearrangement	61
C. Discussion	62
<u>CHAPTER 4 - Pure 12S cDNA Species as Probes for</u>	
<u>Complementary Sequences in Chick DNA</u>	65
A. Introduction	65
1. Purification of individual mRNA species	65
2. The application of pure 12S mRNA species to the study of the molecular genetics of keratins	69
B. Results	70
1. Development of a rapid technique for the molecular cloning of embryonic feather 12S mRNA	70
a) Synthesis of the first strand	70
b) Synthesis of the second strand	70
c) Fractionation of the dscDNA	72
2. Molecular cloning of 12S dscDNA	73
a) Tailing, annealing and transforming the dscDNA	73
b) Detection of recombinants	74
c) 'Miniscreen' length determinations of recombinants	75

3.	Analysis of the 12S cDNA recombinants	76
a)	Resection with <i>Pst</i> I	76
b)	Heteroduplex analysis	77
c)	Restriction cleavage of the recombinant DNA	78
d)	Mapping restriction sites	79
4.	Pure 12S mRNA species as probes to restricted chick chromosomal DNA	81
a)	12S mRNA is comprised of two groups of species	81
(i)	Group I species	81
(ii)	Group II species	82
b)	The properties of Group I species	83
(i)	Many chick DNA sequences are homo- logous to Group I species	83
(ii)	The Group I species pK12 and pK23 have heterogeneous flanking and/or intervening sequences	85
C.	Discussion	87
1.	Isolation of pure 12S mRNA species	87
2.	12S mRNA species are heterogeneous in sequence	88
3.	Embryonic feather 12S mRNA is comprised of two groups of species	89
a)	Group I species	89
b)	Group II species	90
4.	Group I species represent a family of homo- logous sequences	92
5.	Sequences adjacent Group I species genes are heterogeneous	92

6. Are the Group I species linked?	93
7. Do the natural Group I genes contain intervening sequences?	95
8. Other comments	96
<u>CHAPTER 5 - General Discussion</u>	98
A. Current Status of the Process of Keratinization in Embryonic Feathers	98
1. Keratins: a multi-gene system	98
2. Evolution of keratin genes	100
3. Molecular genetics of keratin coding genes	101
B. Future Experiments of Clarification	102
C. Future Approaches	102
1. The number of species expressed in the embryonic feather	102
2. Sequence analysis of cDNA species	103
3. Isolation of chromosomal DNA segments carrying keratin genes	104
4. <i>In vitro</i> transcription analysis	104
<u>SECTION II - A RESTRICTION CLEAVAGE MAP OF COLIPHAGES</u> <u>186 AND P2.</u>	
<u>CHAPTER 6 - Introduction</u>	105
A. General Introduction	105
B. Restriction Endonucleases	105
C. Procedures for Mapping Restriction Endonuclease Sites on DNA	108
1. Double digestion and fragment redigestion	108
2. Physical characteristics of the DNA	108
3. Pulse-chase labelling of DNA	109

4.	End-labelling and partial digestion	109
5.	Differential protection of cleavage sites	110
6.	DNA transfer to nitrocellulose	110
D.	The Application of Restriction Mapping Techniques	
	To The Study of Viral Genomes	111
1.	DNA replication	112
2.	DNA transcription	112
3.	Localization of genes	114
4.	DNA sequencing	114
E.	The Coliphages 186 and P2	115
F.	The Aim of the Work Described In This Section	117
<u>CHAPTER 7 - Materials and Methods</u>		118
A.	Materials	118
1.	Bacterial strains	118
	a) <i>Escherichia coli</i> K12 strains	118
	b) Other strains	118
2.	Bacteriophage strains	118
3.	Chemicals	119
4.	Liquid media and buffers	120
5.	Solid media	121
B.	Methods	121
1.	Storage of bacteria and bacteriophage	121
2.	Growth of indicator bacteria	122
3.	Phage assays	122
4.	Preparation of phage stocks	123
	a) Heat induction	123
	b) Liquid infection	123
	c) Plate stocks	124
5.	Preparation of bacteriophage DNA	125

6. Restriction enzyme digestion, gel electrophoresis, fragment size determination and end-fragment determination	125
7. Isolation and redigestion of restriction fragments	126
8. Construction of a λ 186 phage	127
<u>CHAPTER 8 - Results</u>	128
A. Introduction	128
1. The use of two properties of the coliphages 186 and P2	128
a) Complementary single-stranded ends of 186 and P2 DNA	128
b) Viable hybrid 186-P2 phage	129
2. The use of double digests and redigests of fragments	129
3. The use of partial digests and annealing to a characterized fragment	130
B. Results	130
1. Mapping from end-renaturation and 186-P2 hybrid phage fragment comparisons	130
a) <i>Bam</i> HI	131
b) <i>Bgl</i> II	132
c) <i>Eco</i> RI	132
d) <i>Hind</i> III	133
e) <i>Pst</i> I	134
f) <i>Sal</i> I	134
g) <i>Xba</i> I	135
h) <i>Xho</i> I	136
2. Mapping from co-digestion and redigestion	136

a) <i>Bam</i> HI	136
b) <i>Pst</i> I	138
3. Mapping from partial digestion analysis	138
a) <i>Bam</i> HI	139
b) <i>Pst</i> I	141
4. An analysis of 186 insertion and deletion mutants	142
a) 186 <i>ins</i> 2 and 186 <i>ins</i> 3 each carry an IS3 element	142
(i) Restriction endonuclease evidence	142
(ii) DNA annealing evidence	143
b) Detection of a cryptic inserted sequence in a 186 deletion mutant	144
5. Molecular cloning of <i>Hind</i> III and <i>Eco</i> RI generated fragments of 186 DNA	148
<u>CHAPTER 9 - Discussion</u>	151
A. A Restriction Cleavage Map of 186, P2 and two 186 Insertion Mutants	151
B. Characterization of two 186 Deletion Mutants	152
C. Use of the Restriction Map in the Cloning and Analysis of 186 DNA	154
REFERENCES	156
APPENDIX - PUBLICATIONS	169

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*III, *Eco*RI, or *Hind*III, 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 *EcoRI*, *HindIII*, or *BglIII* 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 l2S mRNA sequences from which they were derived, annealed to the same set of between 20 and 30 *Bgl*III, *Hind*III 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 *Hind*III 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, *Hind*III, *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 *cI* gene and *int* gene respectively. 186*del*1 and 186*del*2 were shown to carry the same deletion affecting the *cI* gene, but 186*del*2 carried a cryptic insert in the repressor binding site (operator).

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.