



CHICK

EMBRYONIC FEATHER

GENES

A thesis submitted to the University of Adelaide, South Australia,
for the degree of Doctor of Philosophy.

by

CHARLES PHILLIP MORRIS B.Sc.(Hons.)

Adelaide Centre for Gene Technology,
Department of Biochemistry,
University of Adelaide,
South Australia.

October, 1984.

TABLE OF CONTENTS

	PAGE
THESIS SUMMARY	i
STATEMENT	v
ACKNOWLEDGEMENTS	vi

CHAPTER 1

INTRODUCTION.

1.1	PRELIMINARY COMMENTS.	1
1.2	CHICK EMBRYONIC FEATHER DEVELOPMENT.	2
1.2.1	FEATHER MORPHOGENESIS.	2
1.2.2	THE ROLE OF THE DERMIS.	4
1.2.3	OTHER FACTORS AFFECTING FEATHER DIFFERENTIATION.	6
1.3	THE NATURE OF EMBRYONIC FEATHER PROTEINS.	8
1.3.1	FEATHER SPECIFIC β -KERATINS.	8
1.3.2	FAST PROTEINS.	10
1.3.3	SCALE LIKE β -KERATINS.	12
1.3.4	OTHER PROTEINS FOUND IN EMBRYONIC FEATHERS.	13
1.3.5	SUMMARY.	14
1.4	GENE EXPRESSION IN THE EMBRYONIC FEATHER.	14
1.4.1	THE TIME COURSE OF PROTEIN SYNTHESIS.	14
1.4.2	THE TIME COURSE OF mRNA SYNTHESIS.	15
1.4.3	THE NATURE OF THE mRNAs EXPRESSED IN EMBRYONIC FEATHERS.	15
1.4.4	THE NATURE OF THE GENES EXPRESSED IN EMBRYONIC FEATHERS.	17

1.4.5	SUMMARY.	20
1.5	THE MAMMALIAN EPIDERMAL TISSUES.	20
1.6	AIMS OF THE PROJECT.	23
1.6.1	GENERAL AIMS.	23
1.6.2	SPECIFIC AIMS OF THIS THESIS.	24

CHAPTER 2

MATERIALS AND GENERAL METHODS.

2.1	MATERIALS.	25
2.1.1	FINE CHEMICALS FOR SPECIFIC PROCEDURES.	25
2.1.1a	ELECTROPHORESIS.	25
2.1.1b	CELL-FREE TRANSLATION.	25
2.1.1c	CELL-FREE TRANSCRIPTION.	25
2.1.1d	COLUMN CHROMATOGRAPHY.	26
2.1.1e	FLUOROGRAPHY.	26
2.1.1f	SCINTILLATION COUNTING.	26
2.1.2	ENZYMES.	27
2.1.3	RADIOCHEMICALS.	27
2.1.4	BUFFERS.	28
2.1.5	BACTERIAL MEDIA.	28
2.1.5a	LIQUID MEDIA.	28
2.1.5b	SOLID MEDIA.	29
2.1.6	ANTIBIOTICS.	29
2.1.7	BACTERIAL STRAINS.	29
2.1.8	BACTERIOPHAGE STRAINS.	30
2.1.9	TISSUE.	30
2.1.10	MISCELLANEOUS MATERIALS.	30
2.1.11	MISCELLANEOUS FINE CHEMICALS.	31

2.2	GENERAL METHODS.	31
2.2.1	PREPARATION OF GLASSWARE AND SOLUTIONS.	31
2.2.2	PREPARATION OF FEATHER LYSATE.	32
2.2.3	PREPARATION OF FEATHER RNA.	32
2.2.3a	PREPARATION OF FEATHER POLYSOMES.	32
2.2.3b	PHENOL EXTRACTION OF POLYSOMES.	33
2.2.3c	EDTA DISSOCIATION OF POLYSOMES.	33
2.2.3d	OLIGO(dT) CELLULOSE CHROMATOGRAPHY OF RNA.	33
2.2.4	PREPARATION OF PURIFIED KERATIN mRNA.	34
2.2.5	CELL-FREE TRANSLATION OF RNA.	34
2.2.6	S-CARBOXYMETHYLATION OF TRANSLATION PRODUCTS.	35
2.2.7	GEL ELECTROPHORESIS OF PROTEINS AT pH 2.7.	36
2.2.8	SDS GEL ELECTROPHORESIS OF PROTEIN.	36
2.2.9	ETHANOL PRECIPITATION.	36
2.2.10	PREPARATION OF PLASMID DNA.	37
2.2.11	RESTRICTION OF DNA.	38
2.2.12	AGAROSE GEL ELECTROPHORESIS OF DNA.	38
2.2.13	POLYACRYLAMIDE GEL ELECTROPHORESIS OF DNA.	39
2.2.14	OLIGO(dT) PRIMED REVERSE TRANSCRIPTION.	39
2.2.15	DETECTION OF RECOMBINANT PLASMIDS.	40
2.2.16	TRANSFER OF DNA TO NITROCELLULOSE.	40
2.2.17	HYBRIDIZATION OF RADIOACTIVE PROBES TO IMMOBILIZED DNA.	40
2.2.18	ISOLATION OF RESTRICTION FRAGMENTS.	41
2.2.18a	FROM POLYACRYLAMIDE GELS.	41
2.2.18b	FROM LOW MELTING POINT AGAROSE GELS.	42
2.2.19	RADIOLABELLING OF DNA.	42

2.2.19a	5'-TERMINAL RADIOLABELLING.	42
2.2.19b	3'-TERMINAL RADIOLABELLING.	42
2.2.19c	NICK TRANSLATION OF DOUBLE-STRANDED DNA.	43
2.2.19d	EXTENSION OF UNIVERSAL PRIMER ON M13 SUB-CLONES.	43
2.2.20	SEPARATION OF RADIOLABELLED TERMINI.	44
2.2.20a	STRAND SEPARATION.	44
2.2.20b	SECONDARY RESTRICTION.	45
2.2.21	CHEMICAL CLEAVAGE DNA SEQUENCING REACTIONS.	45
2.2.22	THE M13 CLONING SYSTEM.	47
2.2.22a	PREPARATION OF M13 REPLICATIVE FORM DNA.	48
2.2.22b	PREPARATION OF M13 VECTORS.	49
2.2.22c	LIGATION CONDITIONS.	49
2.2.22d	TRANSFORMATION OF <i>E. coli</i> , JM101.	50
2.2.22e	HARVESTING (+) STRAND OF RECOMBINANT BACTERIOPHAGE.	50
2.2.23	DIDEOXY SEQUENCING REACTIONS.	50
2.2.23a	HYBRIDIZATION.	50
2.2.23b	POLYMERIZATION.	50
2.2.23c	SEQUENCING GEL ELECTROPHORESIS.	51
2.2.24	CONTAINMENT FACILITIES.	51
2.2.25	COMPUTER ASSISTED DNA SEQUENCE ANALYSIS.	51

CHAPTER 3

PRELIMINARY CHARACTERIZATION OF PUTATIVE mRNAs IN EMBRYONIC FEATHER.

3.1	INTRODUCTION.	53
3.2	SPECIFIC METHODS.	56

3.2.1	RANDOM PRIMED REVERSE TRANSCRIPTION.	56
3.2.2	RNA HYBRIDIZATION REACTIONS.	56
3.2.3	POLYACRYLAMIDE GEL ELECTROPHORESIS OF RNA.	57
3.2.3a	98% FORMAMIDE GELS.	57
3.2.3b	8 M UREA GELS.	58
3.2.4	DETECTION OF RNA SAMPLES ON POLYACRYLAMIDE GELS.	58
3.2.5	³ H-BOROHYDRIDE LABELLING OF KERATIN mRNA.	59
3.2.6	SIZING OF POLY(A) TRACTS.	60
3.2.7	DETECTION OF POLY(A) USING POLY(U).	60
3.2.8	NUCLEOTIDE SEQUENCING OF RNA.	61
3.2.8a	PLUS-MINUS SEQUENCING.	61
3.2.8b	DIDEOXY SEQUENCING.	62
3.3	RESULTS.	62
3.3.1	ANALYSIS OF THE RNAs PRESENT IN 14-DAY FEATHERS.	62
3.3.1a	PURIFICATION OF 9, 12 AND 14S RNA FROM RIBONUCLEOPROTEIN PARTICLES.	62
3.3.1b	TRANSLATION OF RNA IN A CELL-FREE SYSTEM.	63
3.3.1c	ANALYSIS OF TRANSLATION PRODUCTS BY ELECTROPHORESIS AT pH 2.7.	64
3.3.1d	PURIFICATION OF RNA FROM SDS DISSOCIATED POLYSOMES.	65
3.3.1e	OLIGO(dT) CELLULOSE CHROMATOGRAPHY OF RNA.	66
3.3.1f	TRANSLATION OF RNA IN A CELL-FREE SYSTEM.	66
3.3.1g	ANALYSIS OF TRANSLATION PRODUCTS BY ELECTROPHORESIS AT pH 2.7.	66

3.3.1h	14S RNA cDNA SYNTHESIS AND HYBRIDIZATION KINETICS.	68
3.3.2	DETAILED STUDIES OF 12S KERATIN mRNA.	69
3.3.2a	12S mRNA POLY(A) LENGTH.	69
3.3.2b	ATTEMPTS TO OBTAIN STRUCTURAL INFORMATION ON KERATIN mRNA BY RESTRICTION OF KERATIN cDNA.	72
3.3.2c	ATTEMPTS TO ANALYSE KERATIN mRNA 3'-SEQUENCES.	75
3.4	DISCUSSION.	77
3.4.1	THE NATURE OF 14-DAY EMBRYONIC FEATHER mRNAs.	77
3.4.2	KERATIN mRNA FINE STRUCTURE.	79

CHAPTER 4

KERATIN mRNA SEQUENCES.

4.1	INTRODUCTION.	83
4.2	SPECIFIC METHODS.	86
4.2.1	MINISCREEN EXAMINATION OF PLASMID RECOMBINANTS.	86
4.3	RESULTS.	86
4.3.1	DETERMINATION OF RECOMBINANT PLASMID INSERT SIZE.	86
4.3.2	DIGESTION OF RECOMBINANTS WITH Hae III.	87
4.3.3	SELECTION OF A GROUP I cDNA CLONE FOR FURTHER STUDY.	87
4.3.4	RESTRICTION MAPPING OF pCFK23.	88
4.3.5	RADIOLABELLING RESTRICTION FRAGMENTS FOR DNA SEQUENCING.	89

4.3.6	SEPARATION OF ³² P-LABELLED TERMINI.	90
4.3.7	CHEMICAL CLEAVAGE DNA SEQUENCING REACTIONS.	91
4.3.8	M13 CLONING AND DIDEOXY SEQUENCING.	92
4.3.9	THE NUCLEOTIDE SEQUENCE OF pCFK23.	93
4.3.10	SECONDARY STRUCTURE IN pCFK23.	95
4.3.11	SELECTION OF A FULL LENGTH cDNA CLONE.	96
4.3.12	THE PARTIAL NUCLEOTIDE SEQUENCE OF pCFK26.	97
4.3.13	pCFK26 - A TERATOGENY.	97
4.4	DISCUSSION.	99
4.4.1	GROUP I SEQUENCES CODE FOR EMBRYONIC FEATHER β -KERATIN.	99
4.4.2	THE NEW MODEL FOR KERATIN mRNA STRUCTURE.	99
4.4.3	THE CODING SEQUENCE OF KERATIN mRNA.	100
4.4.4	THE 3'-UNTRANSLATED SEQUENCE OF KERATIN mRNA.	101
4.4.5	THE 5'-UNTRANSLATED REGION OF KERATIN mRNA.	102

CHAPTER 5

NON-KERATIN mRNA SEQUENCES.

5.1	INTRODUCTION.	104
5.2	SPECIFIC METHODS.	105
5.2.1	ELECTROPHORESIS OF DENATURED RNA.	105
5.2.2	TRANSFER OF GLYOXYLATED RNA TO NITROCELLULOSE.	105
5.2.3	HYBRID SELECTION.	106
5.2.3a	BINDING DNA TO NITROCELLULOSE.	106
5.2.3b	HYBRIDIZATION AND ELUTION OF RNA.	106
5.3	RESULTS.	107

5.3.1	SELECTION OF A GROUP II CLONE FOR SEQUENCING.	107
5.3.2	THE NUCLEOTIDE SEQUENCE OF pCFK22.	108
5.3.3	pCFK22 SEQUENCES IN THE cDNA LIBRARY.	109
5.3.4	THE NUCLEOTIDE SEQUENCE OF pCFK10.	110
5.3.5	THE LENGTH OF THE pCFK22 mRNA.	111
5.3.6	HYBRID RELEASE TRANSLATION.	112
5.3.7	NUCLEOTIDE SEQUENCE OF pCFK3.	113
5.3.8	THE NUCLEOTIDE SEQUENCE OF pCFK15.	115
5.3.9	FREQUENCY OF SCALE SEQUENCES IN THE FEATHER cDNA LIBRARY.	116
5.3.10	THE PRESENCE OF pCFK15 SEQUENCES IN FEATHER mRNA.	117
5.4	DISCUSSION.	117
5.4.1	THE PRESENCE OF SCALE KERATIN mRNA IN EMBRYONIC FEATHERS.	117
5.4.2	GROUP II SEQUENCES CODE FOR FAST PROTEIN.	119

CHAPTER 6

IDENTIFICATION OF FAST PROTEIN GENES.

6.1	INTRODUCTION.	121
6.2	SPECIFIC METHODS.	123
6.2.1	SCREENING THE GENOMIC LIBRARY.	123
6.2.2	PREPARATION OF RECOMBINANT PHAGE DNA.	124
6.2.3	EXTENSION OF DNA PRIMERS ON FEATHER mRNA.	124
6.2.4	PRIMER EXTENDED SEQUENCING.	125
6.2.5	FPLC ANION EXCHANGE CHROMATOGRAPHY.	126
6.2.6	FPLC CATION EXCHANGE CHROMATOGRAPHY.	126

6.2.7	HPLC REVERSE PHASE C18 CHROMATOGRAPHY.	126
6.2.8	HPLC REVERSE PHASE DIPHENYL CHROMATOGRAPHY.	127
6.2.9	AUTOMATED PROTEIN SEQUENCING.	127
6.3	RESULTS.	128
6.3.1	CHARACTERIZATION OF THE pCFK22 GENE.	128
6.3.1a	SELECTION OF RECOMBINANTS FROM THE GENOMIC LIBRARY.	128
6.3.1b	LOCALIZATION OF THE pCFK22 GENE IN λ CFP1.	129
6.3.1c	POSITIVE IDENTIFICATION OF THE pCFK22 GENE.	131
6.3.1d	RESTRICTION MAPPING OF λ CFP1.	132
6.3.1e	THE NUCLEOTIDE SEQUENCE OF THE pCFK22 GENE.	133
6.3.1f	COMPARISON OF THE pCFK22 GENE WITH THE KERATIN GENES.	137
6.3.2	PURIFICATION AND SEQUENCING OF FAST PROTEIN.	139
6.3.2a	SEPARATION OF FAST PROTEINS BY ION EXCHANGE FPLC.	139
6.3.2b	PURIFICATION OF FAST PROTEINS BY REVERSE PHASE HPLC.	143
6.3.2c	PARTIAL AMINO ACID SEQUENCING OF FAST PROTEIN COMPONENTS.	145
6.4.	DISCUSSION.	147
6.4.1	ISOLATION OF SEQUENCES CODING FOR FAST PROTEINS, A FAMILY OF SEQUENCE RELATED GENES.	147
6.4.2	THE NATURE OF FAST PROTEINS.	149
6.4.3	THE BIOLOGICAL ROLE OF FAST PROTEINS.	150
6.4.4	CONTROL OF FAST PROTEIN GENE EXPRESSION.	151

CHAPTER 7

CONCLUDING DISCUSSION.

7.1	FEATURES OF THE EMBRYONIC FEATHER SYSTEM.	153
7.1.1	PROTEINS EXPRESSED DURING DIFFERENTIATION.	153
7.1.2	CONTROL OF GENE EXPRESSION.	156
7.1.3	EVOLUTION OF FEATHER GENES.	158
7.2	THE EMBRYONIC FEATHER SYSTEM — FUTURE PROSPECTS.	159

BIBLIOGRAPHY

THESIS SUMMARY.

The chicken embryonic feather, an epidermal tissue, undergoes a process of terminal differentiation typified by the synthesis of large amounts of β -keratins and an associated set of proteins called fast proteins. The fast proteins are characterized by a high electrophoretic mobility and a high content of histidine. The research presented in this thesis was directed toward the molecular characterisation of the major genes expressed during the differentiation of embryonic chick feather tissue. In the absence of recombinant DNA research facilities, the early work was directed toward the direct physical characterisation of the RNAs expressed in feathers by using classical RNA separation and *in vitro* translation techniques. When recombinant DNA research became possible, a more precise description of the genes expressed in feathers was achieved by the analysis of feather cDNA clones and chicken genomic DNA clones. The following results were obtained.

1. Sucrose gradient fractionation of feather RNA revealed several size classes in the range 9 to 14S, all of which were shown to contain some β -keratin translational activity. A low level of fast protein translational activity was demonstrated for both 9 and 12S size classes. A consistently observed 14S RNA peak was shown to contain little translation activity and it was identified as being largely derived from ribosomal RNA sequences by R_0t analysis.
2. Detailed studies on the keratin coding 12S RNA species were undertaken to examine its fine structure. The size of the keratin mRNA poly(A) tract was estimated to be 65 nucleotides by direct detection of stained poly(A), 3'-terminal RNA labelling and hybridization of poly(A) sequences to ^3H -poly(U). Restriction enzyme digestion of single-stranded keratin cDNA revealed that the keratin mRNA consisted of a complex family of RNA species. However, detection of 3'-terminally derived cDNA restriction fragments by end-labelling, prior to restriction cleavage, revealed only a few

fragments, indicating some degree of RNA sequence conservation at the 3'-end.

Attempts to directly sequence the 3'-ends of some keratin mRNAs revealed that the majority of the RNA molecules contained 30 to 40 residues at their 3'-termini, which were conserved and which contained the typical AAUAAA polyadenylation sequence, while the atypical AUUAAA sequence was shown to be present at low frequency in 12S feather mRNA.

It was concluded that classical RNA separation techniques were unable to effectively resolve the RNA species of the embryonic feather and the available techniques for the direct analysis of RNA could not successfully distinguish components of the complex mixture of molecular species present in keratin mRNA.

3. The availability of a cDNA library made from embryonic feather mRNA allowed a more detailed molecular analysis of mRNAs expressed in the feather. Two classes of sequences were present in the library (termed Group I and Group II). DNA sequence analysis of several Group I clones revealed that this sequence class coded for embryonic feather β -keratins. The sequence data extended from the poly(A) sequence at the 3'-end to the last nucleotide at the 5'-end of the RNA and allowed the first elucidation of the major proportion of a feather keratin protein amino acid sequence as well as providing a very high resolution description of the structure of keratin mRNA molecules. A comparison of the coding and 3' and 5'-untranslated sequences between particular keratin mRNAs and between keratin gene sequences revealed a pattern of gene structure consistent among these sequences, as follows. Keratin genes were found to be a family of 8 to 40 genes, each gene containing coding regions for 97 amino acids which are strongly conserved. The long (430 to 440 bp) 3'-untranslated regions of these genes were found on average to be less conserved than the coding regions, although two

strongly conserved blocks of sequence were observed, one of which was found at the 3'-end of the mRNA. The 5'-untranslated region, which contained the single intron, revealed rigid homology between the genes only in the region 5' to the position of the intron. It is proposed that the rigidly conserved region of the 5'-untranslated region is the site of interaction of a tissue or stage specific effector molecule involved in the initiation of transcription of stage or tissue specific sets of genes in chick epidermal tissues.

4. The nucleotide sequences of four Group II cDNA clones are presented, one of these, pCFK15, was identified as coding for the typical β -keratin protein found in embryonic scale tissue. This was confirmed by Northern transfer analysis using pCFK15 as a probe. The nucleotide sequences of two of the remaining Group II cDNA clones pCFK10 and 22 revealed that they were derived from the same gene and the sequence data indicated that they coded for a protein with a similar amino acid composition to that of fast proteins. A positive identification was not possible since these clones were not full length copies of the mRNA and no protein sequencing data was available for fast proteins, however, hybrid release data indicated that these clones and another clone, pCFK3, selected mRNAs coding for a protein which co-migrated with fast protein on SDS gels. The fourth Group II cDNA clone, pCFK3, which was presumed to be derived from a fast protein mRNA was different in sequence from pCFK10 and 22 and contained only 3'-untranslated sequence.
5. The full protein coding potential (120 amino acids) of the gene from which pCFK10 and 22 were derived was revealed by the isolation and DNA sequencing of a portion of a λ -Charon 4A chick genomic clone. The gene was shown to contain an intron in the 5'-untranslated region at the same position as the keratin genes. Strong sequence homology between the keratin genes and this gene was also shown for the 5'-untranslated region, while the

remainder of the gene showed no such homology to the keratin genes.

6. The gene described above was positively identified as a fast protein gene when the sequence of the protein for which it coded was compared to the N-terminal twenty amino acids of a fast protein component which, for the first time, was isolated in a pure state from embryonic feather protein using FPLC and HPLC techniques and then sequenced using a gas-liquid solid phase protein sequenator.