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).

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.