Edited by H. MESSEL

THE BIOLOGICAL MANIPULATION OF LIFE

The Biological Manipulation of Life

edited by

H. Messel
C.B.E., B.A., B.Sc, Ph.D.
Professor of Physics and
Head of the School of Physics,
The University of Sydney, Australia

A course of lectures contributed
in the 21st International Science School
for High School Students,
sponsored by the Science Foundation for Physics
within The University of Sydney,
at The University of Sydney,
31 August - 11 September, 1981

PERGAMON PRESS

SYDNEY • OXFORD • NEW YORK • PARIS • TORONTO • FRANKFURT

Pergamon Press (Australia) Pty Ltd,
19a Boundary Street, Rushcutters Bay, N.S.W. 2011, Australia.
Pergamon Press Ltd,
Headington Hill Hall, Oxford OX3 0BW, England.
Pergamon Press Inc.,
Maxwell House, Fairview Park, Elmsford, N.Y. 10523, U.S.A.
Pergamon Press Canada Ltd,
Suite 104, 150 Consumers Road, Willowdale, Ontario M2J 1P9, Canada.
Pergamon Press GmbH,
6242 Kronberg-Taunus, Hammerweg 6, Postfach 1305,
Federal Republic of Germany.
Pergamon Press SARL,
24 rue des Ecoles, 75240 Paris, Cedex 05, France.

First published 1981

Cover design by Ingrid Padina
Typeset in Australia by Savage and Co. Pty Ltd, Brisbane
Printed in Australia by Macarthur Press Pty Ltd, Parramatta

National Library of Australia Cataloguing in Publication Data:

International Science School for High School Students
(21st: 1981: University of Sydney)
The biological manipulation of life.

ISBN 0 08 024825 X.
ISBN 0 08 024824 1.

1. Genetic engineering—Addresses, essays, lectures.
I. Messel, Harry, 1922- II. Title.

575.1

Preface

The Science Foundation for Physics within the University of Sydney is proud to present at its 21st International Science School for High School Students once again an interdisciplinary subject, namely genetic engineering.

We have chosen 'The Biological Manipulation of Life' as the general heading for the twenty lectures because they cover selected topics on DNA technology, the use of which opens a wide range of potential practical applications in industry, medicine and agriculture.

The lecturers are specialists in their fields and the material has been prepared, written and edited for fifth-year high school students. I feel, however, that the lectures will be of interest not only to the students but also to the widest section of the public which, world-wide, is becoming increasingly aware of this field of research on the frontiers of science. It is felt that the material presented will be appreciated by the increasingly science-conscious layman, and also, in areas other than their own, by specialised scientists.

The Foundation's 1981 International Science School, and, indeed, this book are intended to stimulate and develop scientific awareness in Australia and throughout the world.

On behalf of the Foundation I wish to take this opportunity of thanking the lecturers—Dr W. F. Bodmer, Dr K. D. Brown, Professor D. M. Danks, Sir Otto Frankel, Dr J. W. Goding, Professor C. B. Kerr, Dr J. Langridge, Dr W. J. Peacock, Professor A. J. Pittard, Dr W. R. Scowcroft, Dr J. Shine and Dr R. H. Symons—for having given so generously of their time and effort. Thanks are also due to Professor L. C. Birch, Challis Professor of Biology at this University, who helped so greatly in the planning of the lecture series.

The overseas students, as well as those selected from throughout Australia, deserve the applause of all, and the Science Foundation for Physics is happy to honour and reward the ability and diligence of these young people.

H. MESSEL

Sydney, August 1981

Contributors

DR W. F. BODMER, Director of Research, Imperial Cancer Research Laboratories, London, England.

DR K. D. BROWN, School of Biological Sciences, The University of Sydney.

PROFESSOR D. M. DANKS, Professor of Paediatrics, Department of Paediatrics, Royal Children's Hospital, Melbourne.

SIR OTTO FRANKEL, C.S.I.R.O. Division of Plant Industry, Black Mountain, Canberra.

DR J. W. GODING, Walter and Eliza Hall Institute of Medical Research, Royal Melbourne Hospital.

PROFESSOR C. B. KERR, Professor of Preventive and Social Medicine, Commonwealth Institute of Health, The University of Sydney.

DR J. LANGRIDGE, C.S.I.R.O. Division of Plant Industry, Black Mountain, Canberra.

DR W. J. PEACOCK, Chief of the C.S.I.R.O. Division of Plant Industry, Black Mountain, Canberra.

PROFESSOR A. J. PITTARD, Professor of Microbiology, Department of Microbiology, University of Melbourne.

DR W. R. SCOWCROFT, C.S.I.R.O. Division of Plant Industry, Black Mountain, Canberra.

DR J. SHINE, Research School of Biological Sciences, Australian National University, Canberra.

DR R. H. SYMONS, Department of Biochemistry, The University of Adelaide.

The Sponsors

The Science Foundation for Physics within the University of Sydney gratefully acknowledges the generous financial assistance given by the following group of sponsors, without whose help the 1981 International Science School for High School Students and the production of this book would not have been possible.

The Government of New South Wales

The Australian Government

The Nell and Hermon Slade Trust

The James N. Kirby Foundation

Alexander Boden, Esq.

DR W. J. PEACOCK

Chief of the C.S.I.R.O. Division of Plant Industry
Canberra

Contents

CHAPTER 1

The Rise of Molecular Biology and Gene Manipulation

W. J. Peacock

Every one of you will have heard the words 'genetic engineering'. You will have heard these words not only because of human insulin being produced in bacterial cells, or because of the even greater excitement in the press about interferon as a possible anticancer agent, but also because, especially a couple of years ago, it was supposed that man could now bring about genetic manipulations to create new organisms, possibly very dangerous ones.

The concern about genetic engineering was generated partly because we now have the ability to be specific about what we do with the genes of organisms and because we are now in a position to bring segments of DNA from one organism into another. We are in an age of genetic surgery rather than in an age of genetic social science. Also it was thought, wrongfully as it turned out, that the danger of genetic meddling by scientists was enhanced by the fact that they used, as one of their major tools, the bacterium *Escherichia coli* which lives in our intestines. It was argued that there was always a chance that any gene that was being multiplied in *E. coli* in the laboratory could ultimately escape and cause an epidemic in the human population.

Virtually all of this worry about *E. coli* was based on ignorance and mis-statement. A great deal of work has now established the safety of the manipulations used in recombinant DNA technology and genetic engineering. Rather than concern or horror, we should feel a sense of wonder and anticipation for the potential of this field of science.

Let me remind you of some of the basic facts and events which have surrounded this area so that you might be better equipped to judge what is being done and to form your own opinions.

The molecular biology revolution

At the beginning of this century we knew that the determinants of heredity, the genes, were located in chromosomes, and in the first two decades it was clearly established that genes were organised in a linear fashion along the chromosomes. It was not until almost the middle of the century that we knew that the gene was composed of deoxyribonucleic acid, DNA, and that a gene coded for a protein product which subsequently was responsible for a particular effect or phenotype in a cell or organism. Some proteins are structural proteins like collagen, so important in cartilage and bone formation, and myosin, one of the proteins important in muscle structure. Other proteins are enzymes which bring about various metabolic processes of the cell. All proteins in all organisms are determined by genes and there is a one-gene-one-protein relationship.

Not long after the half-way point in this century—in 1953 to be precise—a short paper was published in the journal *Nature*, which really was clearly the starting point of one of the major revolutions in the development of biological knowledge, the revolution of molecular biology. This paper by an American, James Watson, and an Englishman, Francis Crick, described the structure of DNA. This was a feat in itself, but what was even more important was that the particular structure, a double helix, had very important connotations biologically. The molecular architecture of DNA really suggested to us the means by which genetic information was coded and the means by which that coded information was passed from one cell generation to another.

The central dogma

Following the structural description of the molecule came a remarkable series of research discoveries. First of all it was shown how the parental double helix was duplicated so that it passed identical information to two daughter molecules. It was then shown how genetic words were spelt out in the structure of the molecule and how that genetic message was read by the cell and transformed into the sequence of amino acids making up a protein.

Watson and Crick showed that the DNA molecule consisted of two chains wound around each other to form a double helix. The backbones of the chains were made up of repeating units of sugar residues connected together by phosphate bridges. Attached to this backbone were nitrogenous bases of two types, purines and pyrimidines, which are in fact the elements that make up the genetic code and the genetic words. There are four different nitrogenous bases; adenine and guanine are the purines, and cytosine and thymine are the two pyrimidines. We usually write these A, G, C and T respectively. Watson and Crick realised that one critical point of the structure was that the bases associated in pairs to form the steps of the double helical stairway and that always an A was with a T and a G with a C. In fact, these were the only ways in which these four bases could be fitted together. The important point about this is that if we know the sequence of bases along one strand we automatically know the complementary sequence in the other helical strand. This feature is the essence of the replication mechanism which relies on these rules of base pairing. Each strand acts as a template upon which a new strand is assembled. This is called semi-conservative replication. Very elegant experiments were done in the late 1950s showing that the DNA molecules of bacteria, as well as the DNA molecules of higher organisms, either plants or animals, obey the same semi-conservative rule of replication. We now know that the detailed biochemical mechanisms differ slightly in higher organisms and in bacteria, but, nevertheless, the same basic rules for the DNA molecules apply.

Now, how do these two types of base pairing, or four, if we consider the left and right differentiation, provide enough infor-

mation to code for the many thousands of different proteins that occur in animals and plants? The secret is that three successive sets of base pairs are required to code for each amino acid of the protein molecule. The genetic code is a triplet code. This means that there are four times four times four, or sixty-four ways of arranging the bases, so there are sixty-four potentially different code letters. As a result of a truly remarkable series of experiments we know that sixty-one of the sixty-four code for amino acids. Since there are only twenty amino acids this means there is considerable redundancy in the code. Another way of saying this is that for some amino acids more than one triplet has the appropriate code for a particular amino acid. The remaining three triplets are signals which are used to indicate the end of a genetic word.

Another remarkable finding, and when one thinks about it, one that makes a lot of sense, is that the genetic code is universal. Triplets have the same meaning in bacteria as in broadbeans, as in man. Thus, the sequence of bases in three unit letters defines a genetic word which determines the sequence of amino acids in a protein.

The deciphering of the genetic code forms the first element of the central dogma. Elegant as the experiments were to demonstrate each letter of the genetic language, other experiments may have surpassed them in elegance. These were ones which considered the problem of how the code words as they appeared in the DNA structure were ultimately read and translated into the amino acid sequence of a protein. It was shown that this is achieved through nucleic acid molecules slightly different in structure to the DNA molecules we have discussed. These molecules, ribonucleic acid, RNA, have a different sugar in their backbone and for the most part they remain single stranded. They are read or transcribed from one of the strands of a DNA double helix, the sense strand. In higher organisms this transcription occurs within the nucleus and the resultant RNA molecule, the messenger RNA (mRNA), travels from the nucleus into the cytoplasm where it associates with ribosomes and directs the formation of a particular protein.

The molecular mechanisms involved in the translation of the nucleotide word in the mRNA molecule into the amino acid sequence of the end product protein are understood in considerable detail.

The information explosion from microorganisms

All of the essential elements of the central dogma of how genes reproduce and work were elucidated in bacteria and their viruses, the bacteriophages. I can remember how, in the early 1960s, the 'Science and the Citizen' pages of *Scientific American* were dominated by the findings in molecular biology of microorganisms. I used to rush for the latest copy to learn of the unravelling of the genetic code, the demonstration of mRNA, the mechanism of recombination of genes and so on. The years from 1953 to about 1970 belonged in genetic biology to research on *E. coli* and its lambdaphage. It was these organisms and others like them which gave us the key of simplification to our understanding of the complexities of the processes of life. The one millimetre of DNA in the *E. coli* chromosome is remarkably well-known to us. We know the sequences of many of its 3000 or so genes, and we know in some cases the way in which these genes are switched on and off.

A property of the microbial system which enabled so much progress to be made was that small segments of the chomosome could be manipulated individually and studied in detail. The much greater size of the genomes of eukaryotes was a factor which daunted many investigators and prevented incisive attacks on gene organisation and regulation in higher organisms. It was not possible to isolate a specific gene from the great mass of DNA.

Recombinant DNA technology

The development of recombinant DNA technology transformed the molecular biology of eukaryotes. There were two essential discoveries which made it possible to isolate particular DNA segments in a pure form and in large quantity. One was restriction enzymes (endonucleases) which cut DNA into discrete segments, dependent on the distribution of certain base sequences,

and the other was plasmids, small DNA molecules which can enter bacterial cells and multiply. The combination of these two discoveries enabled molecular biologists to cut chromosomal DNA into small pieces which could be inserted into plasmid DNA and multiplied into large quantity in a bacterial culture.

Restriction enzymes, isolated from bacteria, cut DNA at their recognition sequences which, because they are only four or six base pairs long, occur in most higher organism DNA with appreciable frequency and are dependent on the base composition of the DNA. Some of the enzymes cut the two strands of the helix straight across to form flush ends, while others make staggered breaks leaving ends with single stranded tails. These single stranded tails are identical regardless of whether the DNA source is bacterial, plant or animal. Complementary single stranded tails are formed by the restriction enzymes at the end of each cut segment. Any two DNA molecules cut with the same restriction enzyme will, when mixed together, be able to join by virtue of the complementarity of the cut ends—the cohesive ends.

When a lot of cut fragments are mixed together in conditions where some of the cut fragments are plasmid molecules, individual hybrid plasmid molecules can be isolated by cloning in a bacterial cell. The hybrid plasmid molecule or chimeric plasmid is one which now contains an inserted piece of foreign DNA. The cloning process in bacteria separates each chimeric plasmid from the general population and, by amplifying it, makes it possible to isolate and study in detail the inserted segment. Absolutely crucial to the cloning procedures are genetic characteristics which make it possible to select for a plasmid which has an insert at a particular place along its length. You will see later in the lecture series that this is generally achieved with the use of antibiotic resistance genes. Since *E. coli* has a doubling time of about twenty minutes, many millions of bacteria containing the chimeric plasmid can be grown in culture in a period of a day. The plasmid molecules act as vectors for introducing and multiplying DNA segments within the host bacterial cells. Viruses can be used as vectors in animal cells and there is hope that a vector system will soon be developed for plant cells.

Gene banks

Using recombinant DNA techniques it is now possible to clone in a population of bacteria, the entire DNA of any organism. The total DNA is cleaved into segments of a few thousand base pairs with a restriction enzyme, the segments joined individually to plasmid vectors and introduced into *E. coli*. Each bacterial colony contains one DNA segment from the donor organism. The population of colonies forms a library, or bank, of all the segments of the donor's genome. It is possible with a variety of techniques to identify particular genes contained in the bank. We can, for example, select from the gene banks of different animals the genes coding for the haemoglobin molecules and ask what have been the evolutionary pathways of the gene sequence and of the control regions.

Sometimes where a mRNA for a particular gene can be readily isolated, the cloned gene sequence is obtained by producing in a test tube a double stranded DNA copy of the gene message and inserting this into a plasmid for cloning. These copy-genes can then be used as a probe for selecting the natural gene sequences from a gene bank.

The new eukaryote biology

By enabling researchers to handle precise segments of DNA, recombinant DNA methodologies have transformed the molecular biology of higher organisms. It has been possible to examine particular genes in great molecular detail, to map their location in the chromosomes, to determine the sequence of regions important for their regulation, and to study their genetics by manipulating their structure in the test tube. Researchers have now begun to examine the differences in the suites of genes which are expressed in different tissues and to seek out the regulatory signals which are responsible for the ordered development of the complex organism.

Already there have been some surprising findings. For example, it was assumed that colinearity of DNA sequence, mRNA sequence, and protein amino acid sequence would be true just as it is in microorganisms. But this is not the case in the majority of genes that have been analysed in animals and

plants. The DNA coding sequences are interrupted by other DNA sequences called introns which are clipped out of the primary RNA transcript. The coding regions of the RNA are then joined together again to form the functional mRNA molecule! Split genes and RNA splicing reflect a level of gene regulation not present in microorganisms.

One of the major achievements of recombinant DNA techniques has been the unravelling of the mechanisms used by animals to generate the staggering array of antibodies produced by the immune system. It turns out that DNA segments are actually rearranged in the chromosome so that in different antibody-producing cells different gene sequences are generated. Because there are a number of alternative sequence forms for several regions of the overall antibody gene sequence, the number of rearrangements possible is astronomical, enough to generate the huge diversity needed for the immune protection response.

Other segments of DNA move about the genome–jumping genes have been identified in bacterial systems before and it now seems that they are also a feature of higher organisms. They have already been identified in *Drosophila*, yeast and maize and they may be in all higher organisms. The role of these mobile sequences is yet to be established, but clearly it will soon be worked out and it will bring us closer to an understanding of the molecular basis of heredity.

Genetic engineering

An understanding of the how, when and why of the workings of genes is not going to be the only result of recombinant DNA technology. It will also help us to understand many biological processes and already it is having impact on such diverse fields as plant physiology and diagnostic medicine. For example, we can identify particular alterations in gene sequences which lead to an inherited disease. The power of recombinant DNA detection systems is such that with only a small sample of DNA, say from some white blood cells, it is possible to diagnose such diseases far more readily than has been possible in the past. Many