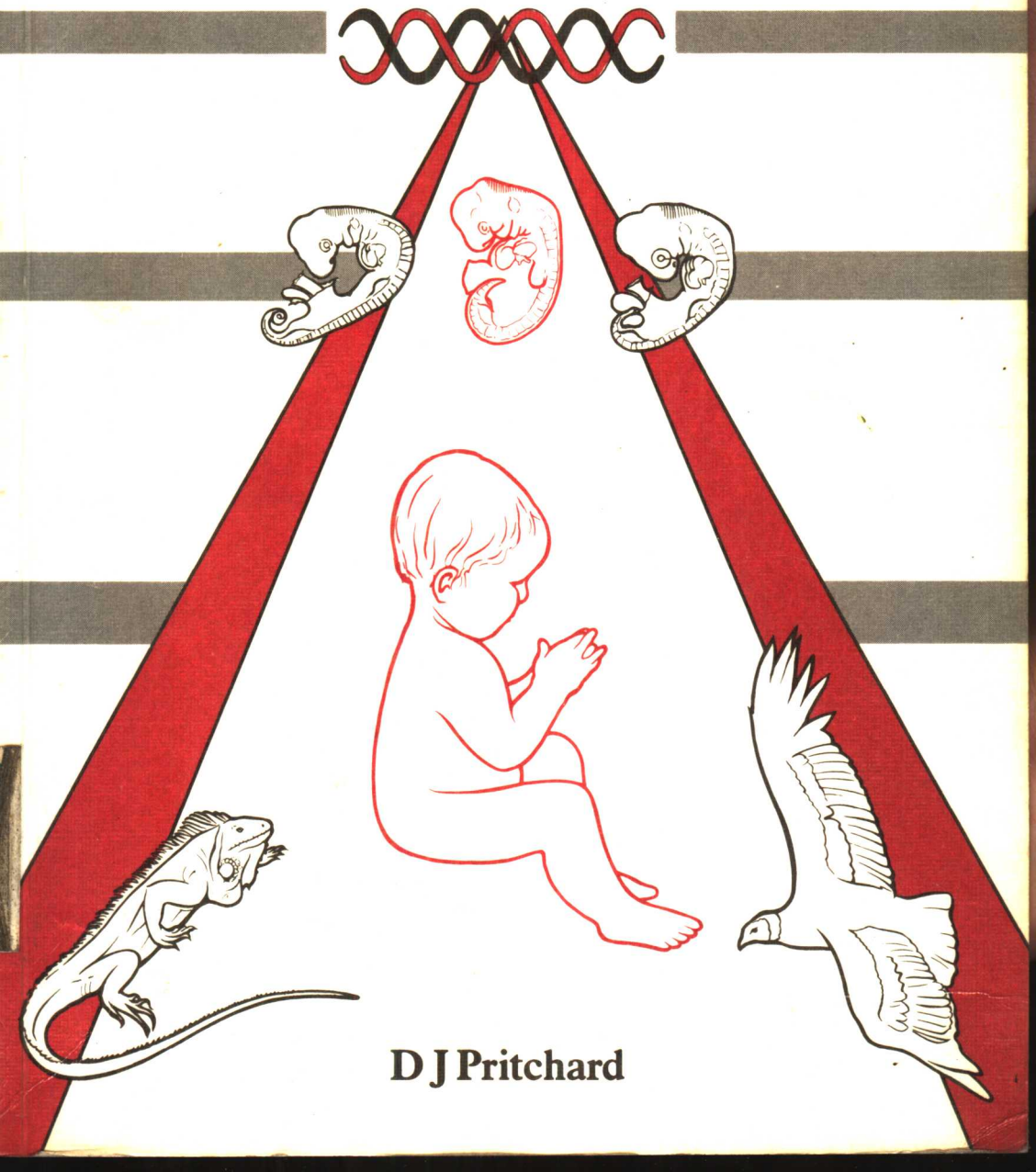
Foundations of developmental genetics



D J Pritchard

Foundations of Developmental Genetics

D. J. Pritchard

Department of Human Genetics
University of Newcastle upon Tyne, England

Taylor & Francis
London and Philadelphia
1986

UK Taylor & Francis Ltd, 4 John St., London WC1N 2ET

USA Taylor & Francis Inc., 242 Cherry St., Philadelphia, PA 19106–1906

British Library Cataloguing in Publication Data

Pritchard, D. J.
Foundations of developmental genetics.
1. Genetics
I. Title
575.1 QH430

ISBN 0-85066-356-3
ISBN 0-85066-287-7 Pbk

Library of Congress Cataloging in Publication Data

Pritchard, D. J. (Dorian J.)
Foundations of developmental genetics.

Bibliography: p.
Includes index.
1. Developmental genetics. I. Title.
QH453.P75 1986 591.3 86-5884
ISBN 0-85066-356-3
ISBN 0-85066-287-7 (pbk.)

Illustrated by D. J. Pritchard
Typeset by Mathematical Composition Setters Ltd, Salisbury
Printed by Taylor & Francis (Printers) Ltd, Rankine Road, Basingstoke, Hants.

Foundations of Developmental Genetics

Preface

In the 'Aphorismen' of Georg C. Lichtenberg is the following quotation: "The thing that astonished him was that cats should have two holes cut in their coat exactly at the place where their eyes are." The individual to whom this refers might well also have wondered why, when he looked in the mirror, the right side of his face was so very much like the left, but very different from everyone else's; or when he found a centipede or millipede in the garden, whether he might one day find an 'infinipede', with an unlimited number of segments and legs. It is for persons with just this kind of ingenious curiosity about biology, albeit tempered by the maturity of wider experience, that this book is written. It has grown from a course in developmental genetics that I teach to second-year science students at the University of Newcastle upon Tyne, and while its direction has derived very largely from the questions research workers are currently asking about how animals develop, the emphasis placed on the various topics has been influenced to a considerable extent by the interests and background of my students. For example, I have assumed all readers to have a basic knowledge of transmission genetics, biochemistry and biology, and I have felt it necessary to emphasize at the outset the enormous gulf between eukaryotes and prokaryotes. I have dealt briefly with invertebrate embryogenesis, but have paid a lot of attention to the details of molecular biology. At all times I have tried to relate the happenings among molecules to those that proceed on grosser levels. My main aim, however, has been to strengthen and establish bridges between embryology and molecular biology, to facilitate mutual understanding and traffic between the two fields. Scattered throughout the text are pointers to and suggestions for research projects which should reinforce these links.

Although primarily intended for undergraduates, I have tried to develop the subject matter sufficiently to interest postgraduate students and research workers in other subjects, and where they have presented themselves I have deliberately chosen intellectually attractive examples, that I believe inspire new interest in both young heads and old. Somewhat to my surprise my analysis has led to what amounts to a new theory relating development, adaptation and evolution, although unbeknown to me evolution theorists have been following a convergent course for some considerable time.

The genotypes and phenotypes of animals in general are already quite well

described elsewhere, but between the two lies another realm, the 'epigenetic space', into which numerous exploratory sorties have been made, but which still remains a largely uncharted region. This book is intended as a guide to this area, to the epigenetic phenomena by which during development genotype becomes converted into phenotype. In this area we must tackle the literature of many disciplines: anatomy, biochemistry, cell biology, embryology, evolution theory, genetics, molecular biology, physiology and zoology. But the developmental biologist is somewhat like the architect who respects and incorporates the skills of many other trades yet adds something else of his own. Epigenetics thus provides a very special kind of intellectual challenge.

Of all scientific disciplines this must be the most difficult in which to make generalizations. One of the tenets of the scientific approach is that if we dig deep enough we will eventually discover common elements in various systems, which when identified and described will allow laws to be written and predictions made. However, in developmental biology we are bedevilled by the problems of the levels of organization at which we should expect to find similarities and where we can safely draw boundaries around groups of phenomena with common elements. A powerful camp currently supports the idea that body patterning in vertebrates and insects is probably defined in a similar fashion. To my mind this is a highly dubious notion. At the molecular level there are still hopes that punctuation of the genetic message will prove to be carried out in similar fashions in bacteria and higher organisms, but this cherished belief must probably soon be abandoned. Where experimental work is concerned it is very difficult to be sure which of the results one obtains in modified situations can be extrapolated to intact organisms. We still do not know what kinds of experiments can validly be carried out with cells taken out of the body and cultured in man-made vessels.

While writing this book I have at times felt the need to adopt a determined stance in order to correct what seems to me a misunderstanding, or bias of emphasis in recent literature. For example, one currently popular idea is that a species' DNA encodes *all* the instructions necessary to bring about development of its members. This is in fact untrue. Not only are many other components of the fertilized egg essential precursors of elements of body cells, but also, in order to form a normal viable animal, the information encoded in its DNA must be interpreted in the correct molecular environmental context. This is perhaps best illustrated by an analogy. For example, imagine that the orders issued to a soldier on manoeuvres include the instruction: "Using locally available materials build a shelter to protect yourself from the prevailing elements". Depending on circumstances this could be interpreted as a direction to construct a windbreak of snowblocks, a canopy of dried leaves, or a platform of mangrove roots. An alternative instruction might say "Using the kit provided assemble the building shown on the enclosed plan", in which case only one outcome would be correct. Interpretation by a developing embryo of the message contained in its DNA falls somewhere between these situations. Individuals of similar genotypes will utilize essentially the same raw materials, but may incorporate them in rather different

proportions, or, taking cues from their surroundings, they may over- or under-emphasize particular bodily features. The resultant derived phenotypes may therefore differ considerably even though encoded by similar genotypes. The plumage of flamingos is typically a brilliant pink, but they develop this colour only if their food contains a high concentration of carotenoids.

Conversely, similar phenotypes can be derived from radically different genotypes since the genes do not contain a complete programme for development in the way that concept is generally understood. Development occurs in a predictable fashion largely because the particular set of conditions present at the outset naturally gives rise to a specific consequence. An analogy for this principle would be the natural colonization of a newly formed island. The different plant and animal species become established in a sequence defined by natural laws, but without external direction. Similarly, a fertilized egg begins its ontogenetic journey along certain developmental pathways because there is only a limited number of open alternatives. It proceeds towards a defined phenotypic outcome by a series of cause-and-effect events, only some of which are actually dictated by the genome. In elucidating this progression genetic mutants are of particular value since they allow us to distinguish the genetically coded aspects from those that are merely the automatic consequences of preceeding situations.

A more fundamental problem faced by the developmental biologist relates to the very basis of what is widely accepted as 'the scientific approach'. Many investigators operate on the principle that analysis of natural phenomena must necessarily involve subdividing the object of investigation into ever smaller elements and describing those elements in the minutest detail, on the assumption that if this kind of analysis is taken far enough, all there is to know will eventually be revealed. This is equivalent to pulling a bicycle pump to pieces and examining the composition and structure of its parts in order to find out how it performs its major function of pumping air. Clearly all the answers are not to be found in that way, but the exercise of examining the relationships between the diverse parts of living organisms and the impact of external non-biological forces upon them is something many of us are not trained to do.

Throughout most of this book I also have followed the reductionist approach, first examining aspects of gross phenotype, then moving down through cell physiology and the properties of proteins, through translation and RNA, to the DNA and transcription. However, I have then tried to pull together some of the threads revealed in previous chapters, in order to elucidate the principles by which animal development is achieved, and to look at whole organisms in a developmental and evolutionary context. It is at this stage more than any other that the text departs from the less imaginative style of reading matter often recommended to undergraduates. It is hoped that the latter chapters particularly may contribute something to the more theoretical approach to living systems which has traditionally played such an important part in the development of our subject.

One topic I have not covered, but which promises to contribute valuable ideas for the future, is the study of viruses. As yet it is difficult to discern to what extent

viruses are prokaryote infiltrators, fragments of eukaryotes that have broken loose, something in between, or something different altogether. It is therefore at present impossible to know whether the rules that govern expression of viral genes integrated into eukaryote DNA should also apply to the rest of the genome, or vice versa. Inter-species transfer of genetic material as viral particles has apparently occurred on many occasions and when we know the full extent of this exchange we may well need to revise some of our best-loved theories.

Each of the disciplines I have entered has its own standards of excellence and its own style of communication, but I have tried where possible to unify the different approaches. In the interest of encouraging the flow of positive thinking so essential in undergraduate teaching I have avoided the critical appraisal of all observations that would be essential in a more advanced text. For the same reason, references and the names of individual workers are rarely included in the text, although all relevant sources, plus fuel for arguments to counter those I advance, should be accessible by an intelligent perusal of the reading lists.

In attempting to cover such a wide range of subject matter errors and misunderstandings are bound to arise, and I am very grateful to colleagues more knowledgeable than myself who have taken the trouble to read and criticize sections of the manuscript. In this respect I owe particular thanks to Professors Ken Burton and Stuart Glover and to Doctors Donald Ede, Tim Horder, Monica Hughes, Alec Panchen, Surinder Papiha, Tony Samson, Clarke Slater and Robert Whittle. However, I must admit I have not always taken their advice and must accept full responsibility for those errors that remain. I wish to express my grateful thanks to Mrs Valerie Webb and Mr Ian Munro for the diagram annotation and to our secretarial staff, especially Mrs Tessa Havelock and Miss Debbie Hayles. My family, Penny, Ceri and Hamish earn my sincere gratitude for their interest and support and for foregoing trips and holidays they might otherwise have had. I gratefully acknowledge permission granted by the McGraw-Hill Book Company Ltd to reproduce Figures 2.11 and 4.7, and the Company of Biologists Ltd to reproduce a diagram from one of my own papers as Figure 7.9.

This book could not have come into existence without the additional stimulus afforded by the thought-provoking words of many other authors, in particular those of my late Professor, C. H. Waddington, and by discussions with many people, including numerous students, my colleagues in Newcastle, and my former teachers and colleagues at the Department of Genetics in Edinburgh. I owe a special debt of gratitude to Professor Derek Roberts for his constant guidance and enthusiastic support and for initiating the idea for this work.

Dorian Pritchard
Newcastle upon Tyne
January 1986

Dedication

To my parents, who gave me not only life itself, but also a love of it in its many aspects and who, despite extreme provocation, never discouraged me from asking questions. Also to Penny, for providing some of the most important answers.

Contents

Chapter 1 Prokaryotes and the origins of eukaryotes

1.1. Genotype and phenotype

Most of the readers of this book will already be aware of Darwin's theory of evolution through natural selection. From the point of view of the evolution of genetic material, an important consideration is that although mutational forces introduce variation into the genes, the forces of selection do not act upon the genes themselves, but rather upon their products: they act upon the living organism, the 'survival machine', that carries those genes. The genetic information which is inherited by an organism from its parents becomes translated into physiological, anatomical and behavioural properties as it develops, and it is these properties which may ensure survival of the organism to the point where it can reproduce and pass its genes on to the next generation.

The identifiable or measurable properties of an organism are what we call its phenotype. Development involves the translation of the genetic blueprint, or genotype, into phenotype, which can be considered to be the product of interaction between genes and environment:

$$\text{Phenotype} = \text{Genotype} \times \text{Environment}$$

This book is about the nature of that interaction, the way in which an organism is produced from its genes.

Classical genetic theory deals with many aspects of the structure of the genetic material, its replication, its mutation and its distribution between progeny, or between members of a population. But what classical genetics has largely avoided is the way in which genes are selectively expressed at different times in development, or in different parts of the body. For example, if a *Drosphila* fruit fly carries a gene for scarlet eye colour, why is this gene not expressed in all the cells of the body? Why is a fly with a scarlet gene not scarlet all over? Why is the larva not abnormally pigmented also? All over the world many laboratories are tackling just this type of problem, using a wide variety of living materials and some of the most advanced molecular genetic techniques. Many of the observations of the early embryologists, which at the time seemed inexplicable, are now falling into place in the minds of present-day developmental biologists. It therefore seems an opportune time to draw together some of the early observations, and to try to see

them in the context of the new knowledge acquired from the very high-powered techniques which have arisen in the field of molecular biology, founded upon the simpler systems of bacteria.

As a starting point to our investigation, we will consider a major type of genetic control which operates in bacteria.

1.2. Prokaryote control systems

The living world is divided into two major groups: the eukaryotes, those with a 'true nucleus', which includes all the plants, animals, fungi and protozoa, and the prokaryotes (literally 'before a nucleus') — the blue-green algae (more correctly called Cyanobacteria), and the true bacteria. (The viruses do not easily fall into either major category and their origin is uncertain.)

In the prokaryotes there are various systems for the control of gene expression. We are not concerned here with the intricacies of prokaryote genetics and for our purpose it will be sufficient to consider a simplified treatment of the negative and positive control of transcription. In both cases there are induction or inducible systems, and repression, or repressible systems. Inducible systems characteristically cope with the situation where the organism needs to respond appropriately to the occasional presence of a particular biochemical in the medium, such as a food substance like lactose. In contrast, repressible systems come into operation to cope with scarcity of an essential substance, such as tryptophan, which is normally present in adequate concentration in the environment.

Figure 1.1 illustrates the two types of negative control. In both there is a linked array of structural genes of related function. The term structural gene, or cistron, describes a gene which codes for an enzyme or structural protein utilized by an organism in carrying out its normal life processes. In each system three cistrons are illustrated, labelled A, B and C, and D, E and F. Close to the 'upstream' or 3′ end of the A gene is what is called an operator site (O) and upstream from this a promoter site (P). The promoter site is the part of the DNA to which the RNA polymerase (the enzyme which carries out transcription) first becomes attached. The cistrons in each group have related functions and this group of linked genes, with associated operator and promoter sites, is termed an operon. In both the inducible and repressible systems there is also a regulator gene (R) which may or may not be structurally linked to the operon. The regulator gene codes for a repressor protein, that controls the transcription of the structural genes.

In a negative inducible system, the repressor protein, in its native state, actively binds to the operator site. If RNA polymerase becomes attached to the promoter site, its passage along the DNA is then blocked by the repressor protein in its path. However, the repressor protein has the capacity to change its shape and its binding properties should it become associated with a specific regulatory metabolite. This property of a protein which allows it to change shape is known as allostery, and the protein is said to be an allosteric protein.

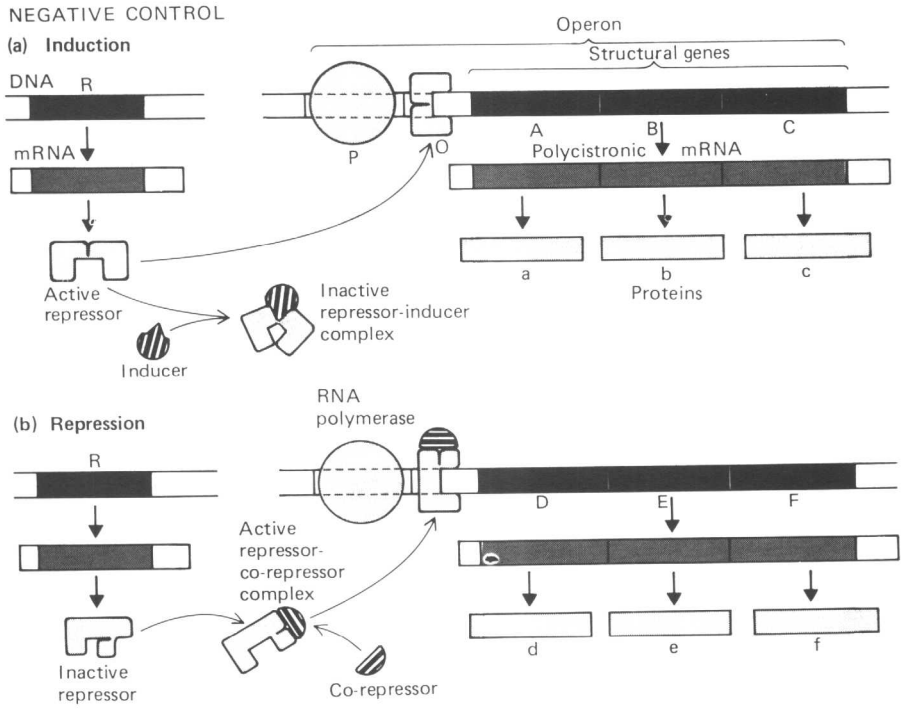


Figure 1.1. Negative control of transcription in prokaryotes.
(a) Induction. A regulator gene, R, codes for a repressor protein, that binds directly on to the operator site, O, upstream of the three structural genes, A, B and C. RNA polymerase attaches to the promotor site, P, but its transcription of the structural genes is blocked by bound repressor. If inducer molecules are present in the cell, they bind to the repressor and reduce its binding affinities for O, so that A, B and C can be transcribed. One polycistronic messenger RNA molecule is produced and translated into three enzymes with related functions. P, O, A, B and C constitute an operon. (b) Repression. In this case the repressor protein is inactive in its native state, so that the structural genes D, E and F are transcribed constitutively. The repressor becomes active in the presence of co-repressor. It then binds to O, blocks the passage of RNA polymerase from P and so prevents the expression of D, E and F.

The operon model was proposed by Jacob and Monod as an explanation of some features of what has become considered to be the classic inducible system, the lac operon. This is concerned with the uptake and metabolism of the sugar lactose. In this system allolactose is the regulatory metabolite, which is formed in the cell when lactose is present. The lac repressor protein responds to allolactose, changing its shape so that it can no longer bind to the lac operator site (Figure 1.1(a)). RNA polymerase which binds to the promoter, P, now passes down the DNA and transcribes A, B and C as a single long molecule of RNA—a polycistronic message

— which is translated into three enzymically active proteins in the ribosomes. When the cytoplasmic concentration of lactose diminishes, due to the action of the enzymes which are produced by the operon, insufficient allolactose is available to inactivate the repressor protein. The repressor therefore binds back on to the operator site and blocks further transcription, pending a later increase in lactose.

In negatively repressible systems (Figure 1.1(b)) the regulator gene codes for a repressor protein which in its native state is *inactive*. RNA polymerase that binds to the promoter site can thus transcribe D, E and F without inhibition. The repressor protein becomes activated by combining with a specific co-repressor, which in some cases is the end-product of the synthetic pathway. The repressor/co-repressor complex binds on to the operator site and blocks further transcription. When cytoplasmic levels of the end-product (i.e., co-repressor) diminish, the complex dissociates, the repressor becomes inactive and transciption of D, E and F proceeds once more.

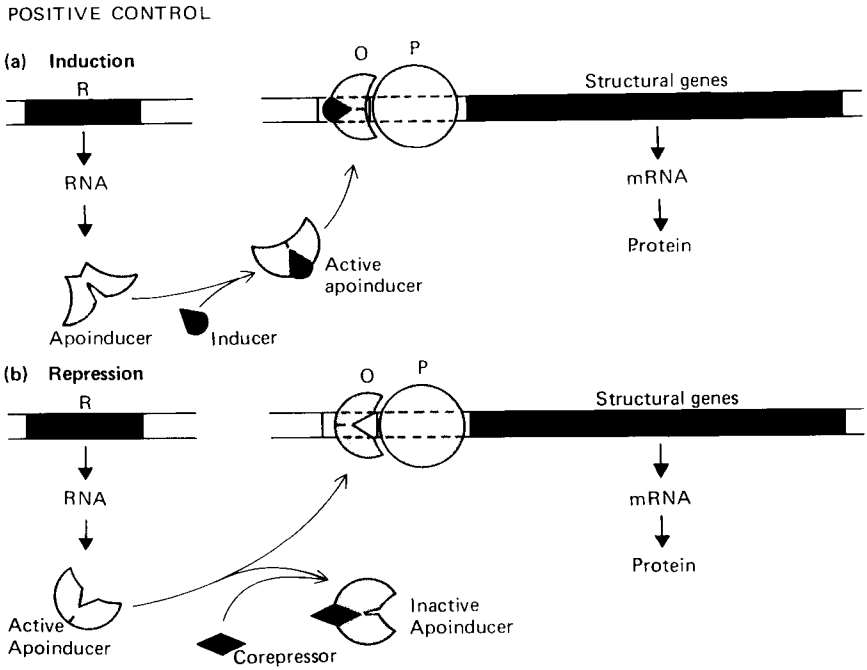


Figure 1.2. Positive control of transcription in prokaryotes.
(a) Induction. An inactive apoinducer protein is produced by expression of a regulatory gene sequence. This becomes active on combination with an inducer molecule and the complex binds to an operator site beside the structural gene(s). This facilitates binding of RNA polymerase to the promoter, allowing transcription to take place. (b) Repression. In this case the apoinducer is active in its native state, but will not bind to the operator site if it becomes complexed with a co-repressor molecule.