STUDIES IN THE GENERAL THEORY OF DEVELOPMENT AND EVOLUTION.

by

Brian C. Goodwin

B.Sc. (Hons. Biol.) McGill

B.A. (Hons. Math.) Oxford

Submitted to the University of Edinburgh as a Thesis in fulfilment of the requirements for the degree of Doctor of Philosophy.

Institute of Animal Genetics

University of Edinburgh

May, 1961.

CONTENTS

	Page
INTRODUCTION..	1
SECTION I - The Epigenetic Variables	12
SECTION II - System and Environment	19
SECTION III - The Components	37
SECTION IV - The Differential Equations	40
SECTION V - The Statistical Mechanics	52
SECTION VI - The Canonical Ensemble	61
SECTION VII - The Epigenetic Temperature	67
SECTION VIII - The Epigenetic Functions	75
SECTION IX - An Epigenetic Equation of State	85
SECTION X - The Epigenetic Forces	90
SECTION XI - Relaxation Time and Parametric Change . . .	112
CONCLUSIONS	125
APPENDIX 1	136
APPENDIX 2	151
APPENDIX 3	152
APPENDIX 4	154
APPENDIX 5	158
ACKNOWLEDGMENTS	159
BIBLIOGRAPHY	160

INTRODUCTION

The possibility of comprehending a large class of biological processes within the conceptual framework of Darwinian theory is one which has occurred to many biologists since 1859. Roux was probably the first to introduce into embryology the ideas of competition and survival when he talked of a struggle between the different parts of a developing embryo. Physiological and developmental regulation have often been described (e.g. Waddington, 1940; Raven, 1938; Weiss, 1939) as the result of an underlying competitive interaction between organs and tissues for nutrients or, more generally, for 'physiological space' (i.e. the proper environment for survival and growth). More recently, these concepts have been applied to the analysis of the learning process (Pringle, 1951), and to problems of cultural development (Gerard, Kluckholn and Rapoport, 1956).

Underlying these extensions of Darwin's original theory there is the idea of a generalised system which, because of its fundamental properties, is subject to a particular type of uni-directional change. In different areas of enquiry, this change has been called evolution, development, learning, and the general theory seeks to describe these as particular instances of what has been called a Darwinian process (Russell, 1958).

The central phenomenon for which an explanation is sought

in all of these disciplines and which is the distinctive feature of Darwinian systems is the emergence of discrete states of 'organic stability' (Bateson, 1894; Waddington, 1957) which arise during their motion and are characterised by the property of being 'adapted' to the existing environmental conditions. Discontinuities occur between different parts of such systems as they expand into slightly differing environments, and these discontinuities lead to the appearance of discrete entities which are called species in evolution, cell types in embryology, behaviour patterns in learning, and cultures in anthropology. The problem is to account for the process which brings these entities into existence.

The expanding potential of Darwinian systems is due to a replicative mechanism whereby their components are multiplied, allowing for exponential increase under 'favourable' conditions (by which is meant the absence of any limiting terms in the equations of the system which could prevent unbounded increase). However, under real conditions, this replicative potential is not realised because the components are dependent for their replication upon a common and limited amount of 'biological space' (conditions allowing for replication) for which they are in competition. The interactions between the components will therefore be defined by the conditions in this space, which is regarded as the 'environment' of the biological system.

Darwinian analysis is concerned primarily with changes wherein the relative numbers of the different species of component are altered. Such changes of state are caused by disturbances which may be external to the system (changing its external parameters), or internal (changing the species of component present). If neither of these change, the

system will approach a steady state in which the relative numbers of the species of components remains constant. This steady state may involve unbounded motion, such as balanced exponential growth in a bacterial population in a constant environment and without mutation. Although such a population 'moves', the only changes which occur in it are quantitative ones, the relative values of the components remaining constant.

The disturbances acting on Darwinian systems are usually defined as random, but this description is very relative to the sort of knowledge we have of the process, and the type of observation which can be made on it. The system responds in some manner to the disturbances, which may be temporary or permanent. When the response involves a change in the composition of the system (changes in the relative numbers of different species of components), then it is said to change its state. The process whereby different stable states of the system arise following the application of a certain set of disturbances has been called selection. Darwin first used this word because he was led to some of his ideas through an analogy with domestic breeding procedure, where selection has a clear connotation. But he came to regret his use of the term natural selection, and indeed the phrase is confusing in its implications. Any system subjected to a sufficient disturbance will move to a new state, which is then 'selected' by the disturbance. In the present study, selection will be used in this sense, but applied to a Darwinian system. It will mean no more than that under a particular set of

disturbances, such a system will undergo a particular process of change leading to some new steady state.

Darwinian analysis is dominated by a certain point of view. Given a system as complex as a biological one, the appropriate question to ask is not What is the detailed motion of the parts?, but What stable states (if any) will survive the application of a given set of disturbances or stimuli? The analysis is a statistical one which seeks to describe the behaviour of the whole system without assuming more than is necessary about the interactions of the components (either because detailed knowledge is not available, or because only certain interactions are important in determining the statistical properties of the whole). This is exactly the position which one is in when attempting to describe the properties of a gas in terms of the motion and interactions of its molecules. This physical problem has been solved by the elegant mathematical apparatus of statistical mechanics. In the present study, an exactly analagous method will be used in an attempt to explain the fundamental behaviour of embryonic cells during development in terms of the interactions of their components.

The particular model appropriate to a given field of biological study will be determined by the type of observation which can be made, and the sort of behaviour investigated. The distinction between a system and its environment (i.e. which variables are to be treated a dependent and which as independent, respectively) is not always an obvious one in biological contexts, and must be made on the basis of the experiments which define the subject. Even in

the most familiar case of individual organisms and their external 'environment' there are interactions which make an analysis of the evolutionary process much more complicated than is uaually assumed to be the case (see, e.g. Waddington, 1957). The organism may change its environment, and this changed environment may then change the organism. If these changes are in any way correlated, there will occur a spiral-type of process whereby organism and environment are carried through a series of changes which must really be looked upon as unitary, the distinction between the two parts being artificial. The environmental parameters then become absorbed into the system as variables.

When one comes to the analysis of developmental processes, which will be our primary concern, this problem becomes acute. No distinction between system and environment within the developing embryo can be made which allows of a complete causal analysis of the form: The environmental parameters have certain values at a certain time, and the 'system' responds by moving to a new state. This new state induces certain changes in the parameters, which then remain constant at their new values until the system has responded by moving to a further state, etc. However, this is in fact the pattern of classical analysis, and it has been extensively used in experimental embryology. One speaks of inducers which act upon competent tissue, causing a particular response. This tissue in turn interacts with others to produce a new environment characterised by a further set of determining conditions. Needham's (1951) series of cones is a geometrical model of this resolution of development

into sequences of transient motion between points of 'stable equilibrium', the transients being 'caused' by the inductive stimuli of organisers.

This method of causal analysis is valid only if it can be shown that the changes of state in the system can occur much more rapidly than the changes which occur in its environment. This important question is discussed in section II, and again more fully in section XI.

Darwinian system are in general open: they exchange matter and energy with the environment. This is because the components of such systems are themselves open, requiring a flow of matter and energy for their own maintenance and replication. Equilibration in such systems occurs when there is a balance between the rate of production and the rate of destruction of each different species; this is the steady state. The changes which may occur in these systems are therefore not restricted to those allowed in closed systems by the second law of thermodynamics. The final state is not completely determined by the initial conditions in open systems (i.e. the initial values of certain variables are unimportant in determining their final values), and a decrease of entropy is no longer forbidden. At the same time, 'opening' the system seems to make inapplicable the powerful and elegant analytical tools which have been developed for the study of systems which conserve energy and obey the classical requirement of weak energy interaction between components. Certainly

biological systems are not conservative in the energy sense, and there is strong interaction between the components.

However, it will be found that under certain assumptions concerning the interactions of the components in embryonic cells, a set of differential equations may be derived for their synthesis and interactions which have the important property of possessing an integral. This integral is regarded as analagous to the Hamiltonian (total energy) of physical systems, and with it a new statistical mechanics may be constructed (the necessary conditions for this will be shown to be satisfied). For this purpose, much use will be made of the very important studies made by Volterra (1931) in the mathematical theory of the 'Struggle for Life' in ecological groups, and the recent extensions of this theory in two very fine papers by Kerner (1957, 1959) who used, as we do, Volterra's integral as the starting point of a statistical mechanics. Kerner's interest, like Volterra's, is demographical. Ours is embryological, so that the results of the theory will have quite different interpretations in the different contexts. However, it is the great power of a statistical mechanics that it gives qualitative predictions of great generality, showing the main features of a statistical ensemble by smoothing out the enormous variety of individual motions, so that systems with the same fundamental 'microscopic' structure will show similar 'macroscopic' properties. There is no reason why the same general qualitative features should not be observed in different biological contexts. Indeed, it is precisely the assumption of general Darwinian analysis that such similarities

are to be observed at different levels of organisation in biological systems. The two subjects which are concerned with relatively short time-changes (fractions of individual life-times), embryology and learning, have the advantage over evolution and anthropology (time-changes of the order of many individual life-times) in that the macroscopic properties of the systems are directly observable experimentally. In embryology these macroscopic phenomena are observed as the behaviour of developing cells, and an extensive language has been formulated to describe the causal relationships between the responses of these cells and the forces which operate upon them in embryonic tissues. It is the language of classical embryology: competence, determination, individuation, organiser, inducer, etc. The present study attempts to account for some of these phenomena as statistical properties arising from certain types of interactions assumed to occur between molecular components of embryonic cells. The cell thus becomes a systemic unit whose behaviour is described by variables of state analogous to the temperature, entropy, free energy, pressure, etc., of statistical thermodynamics. Analogues of all these variables will be found for the embryonic cell; and it will be shown how many of the concepts of classical embryology may be reformulated in terms of 'epigenetic temperature', 'epigenetic free energy', etc. The epigenetic temperature will be found to occupy a particularly important place in discussing systems of the type defined in Section IV (obeying 'Volterra dynamics', whose statistical properties were first studied by Kerner). The chief concerns of the present study will be:

(1) The demonstration that embryonic cells may obey Volterra dynamics at the molecular level,

(2) the study of the consequences of this for the statistical behaviour of the system, and

(3) the interpretation of this behaviour in the embryological context.

There is a final feature of general Darwinian theory which requires mention. The fundamental assumption underlying this way of analysing biological systems is the continuity of the living process within itself and with matter. The variables in Darwinian systems are always functional (they define an activity rather than a structure). Now the discreteness of the points of organic stability which arise in evolution and embryology are observed at the structural level, and so all attempts to explain their appearance must involve the reduction of structure to function. This resolution of substance into process is the program of contemporary science and the foundation of modern cosmology, (see e.g. Whitehead, 1929) so it is entirely consistent to assume the validity of this principle in biological theory. However, we are a very long way from achieving a sufficiently complete description of the forces operating in cells to account for their structural and functional features. There must be many spatial parameters and historical properties entering as initial conditions before morphogenesis can be even approximately accounted for analytically.

In this study, we shall assume that the perpetuation of

function during the replicative process implies a perpetuation of structure. The interdependence of structure and function in enzymes goes some way towards resolving the problems which underly the duality encountered in the description of biological systems. And the properties of macromolecules in general must be expected to account ultimately for much of the hetergeneity and complexity of the cell interior. However, our primary concern will be to consider the functional relations within a particular set of macromolecular constituents of the cell, and to show how they give rise to a process which may underly major developmental changes. Such a study can define some of the necessary conditions for morphogenesis, perhaps the main ones; but this is still far from defining the sufficient conditions.

The point of view taken in this work, then, is that development is to be understood as a particular type of Darwinian process. The components of the system will be defined in such a way that they satisfy the conditions necessary for participation in such a process (replication and competitive interaction), and then the exact type of interactions assumed to occur between them will be given. From these assumptions a set of differential equations will be derived expressing the 'motion' of the system. An integral will be derived for these equations, and with this integral a statistical mechanics of the embryonic cell will be constructed. The properties of the system so obtained will be studied with a view to explaining the characteristic behaviour of developing cells in terms of the new variables of state deriving

from the statistical mechanics: epigenetic temperature, entropy, free energy, etc. The epigenetic temperature will be shown to be a measure of the 'excitation' of the embryonic system (i.e., the size and the nature of the fluctuations occurring in the macromolecular synthetic activities of embryonic cells). The main result of the analysis will be that the intracellular aspects of development may be described as a process of slow 'cooling' in an initially excited system, which brandes into a number of parts whose states diverge progressively until they become stabilised at low epigenetic temperature in mutually exclusive epigenetic states. The gradual 'cooling' of the system is a result of the 'frictional' forces which necessarily accompany irreversible change in a complex dynamic system. Finally, the use of the Darwinian concept of adaptability in the embryological context is discussed, and a relationship between the epigenetic temperature and the adaptive properties of the theoretical epigenetic system is examined. The 'cooling' of the developmental system is then found to correspond to a gradual change in the responsiveness of embryonic cells, such that as the epigenetic temperature decreases towards zero, the epigenetic system loses its capacity for adaptive response to environmental stimuli and becomes stabilised in some differentiated condition which is determined by its developmental history.

SECTION I

THE EPIGENETIC VARIABLES

Classical embryology regards the cell as its unit of structure and function. The concepts of differentiation, individuation, induction, competence, are all defined in terms of the movements and responses of cells and cell masses. We have seen that the first application of Darwinian ideas to embryological phenomena was by Roux, and this view-point has been taken up more recently by others (e.g. Waddington, 1948) who speak of a struggle or competition between the various tendencies of cells in a piece of developing tissue. The idea of competitive interactions between cells underlies many theories about the forces which operate in differentiating tissues, such as those which use the concepts of gradient and field, and speak of physiological competition and dominance (Child, 1941; Weiss, 1939). No doubt a more complete statement of the process could be given in these terms, pursuing the evolutionary analogy and looking upon tissues, organs, and cell-types as families, genera and species. However, it is becoming increasingly clear that development must be understood in molecular terms if it is to satisfy the stringent requirement of continuity between the organic and the inorganic realms. Weiss (1950) has strongly emphasised this view, and has correctly insisted that cellular differentiation must be considered in terms of populations of

chemical substances and their 'molecular ecology' within cells.

The cell from this view-point becomes a system rather than an 'atomic unit', its properties being sought in the statistical behaviour of large numbers of small components. However, in studying the behaviour of large cell masses, a different order of statistics is appropriate, and the cell does become a unit, an atom, with its own pattern of interactions with adjacent cells. Thus, different levels of organisation give rise to different types of statistical behaviour, higher order phenomena not appearing significantly in lower level behaviour. For example, in development, certain processes of differentiation are not observed unless there is an aggregate of cells above a certain critical size (Grobstein and Zwilling, 1953). However, we shall be primarily concerned in the present study with the behaviour of single cells in the developmental process, and for this purpose we require a theory which begins at the molecular level.

Weiss and Kavanau (1957) have recently proposed a model which treats two levels simultaneously, cell and organism, in terms of molecular properties. Their primary interest is growth regulation in development, and their model puts physiological competition into molecular terms. Competition is for 'physiological space' (the proper conditions for development), defined in terms of the density of 'antitemplates', specific inhibitors of the templates which are responsible for growth and differentiation. Whenever 'space' becomes available by the concentration of

antitemplates in the environment dropping below a particular value, tissues will grow into it until, by their own synthetic activities, they fill up their environment with a critical density of antitemplates again. This model treats embryonic growth as a unitary process, whose only point of stability is the adult organism. The equations and curves describing such a process are already fairly complex ones with negative feedback characteristics, whose prototype is the familiar logistic equation. No attempt is made to provide an explanation for the differentiation process which occurs 'within' the general equation of growth which they derive. In terms of their model, this would have to be done by distinguishing different species of template, and studying the changing spatial distributions of these templates and their antitemplates in various parts of the developing embryo. Such a programme may be possible. But the complete solution of any set of equations seeking to describe the whole of embryonic development as a single transient would be unmanageably complicated. It is in the face of such difficulties that one is forced to divide the organism into organisational levels of cells and tissues, and also to break down the time process of development into a sequence of relatively short transients between quasi-equilibrium states. Classical embryology has itself followed this method of analysis.

An epigenetic theory in molecular terms which uses a thoroughly Darwinian argument is that of Spiegelman (1948). He suggested that the enzymes should be regarded as the primary

observables or variables of the differentiation process within cells, and that the subcellular replicative components are the 'plasmagenes', different species of which synthesise different enzymes. The plasmagenes compete for precursors for their own replication, and the developmental process is the 'selection' of different species of plasmagenes (hence synthesis of different enzymes) by different cytoplasmic environments in various parts of the embryo. There are many details of this process which Spiegelman does not go into, and current evidence indicates that some of his assumptions require modification. Waddington's (1948) analysis now seems to be a more accurate representation of the epigenetic process in these terms. It is not sufficient to take only the enzymes into account in defining the developmental state of a cell. Rather, this definition must be made in terms of the concentrations of all the different proteins present in a cell at a particular stage of development, so that epigenesis becomes 'the ontogeny of the proteins' (Shen, 1955). The term 'plasmagene' would then be replaced by a particle or system responsible for protein synthesis. The present hypothesis is that the protein synthetic activities are carried out by the microsomes or ribosomes. These may be looked upon as a type of gene-initiated plasmagene (Waddington, 1956). As to their replicative capacity, the evidence at present available indicates that this is not a property of the particles themselves, but of the coupled system of ribosomes and protein, which together show autosynthetic activity.

There are many different types of autosynthetic system (Hinshelwood, 1953), but the essentials of the process are given by the simplest case, where the equations are,

$$\frac{dP}{dt} = k_P R$$

$$\frac{dR}{dt} = k_R P \qquad \ldots\ldots\ldots (1)$$

The equations mean that some species of molecule R is required for the synthesis of another species P (R is a necessary but not a sufficient condition for synthesis of P, i.e. R is a catalyst), and P is required for the synthesis of R. The solutions are,

$$P = a_1 e^{\alpha t} + a_2 e^{-\alpha t}$$

$$R = b_1 e^{\alpha t} + b_2 e^{-\alpha t} \qquad \alpha = \sqrt{k_P k_R}$$

After a long time ($t \to \infty$), the negative exponential terms will be very small, and the equations will approach the conditions of exponential increase:

$$\left.\frac{dP}{dt}\right]_{t\to\infty} = \alpha P, \qquad \left.\frac{dR}{dt}\right]_{t\to\infty} = \alpha R \qquad \ldots\ldots\ldots (2)$$

If, now, P represents protein and R ribonucleic acid (RNA), the above solution should hold for these components. Such evidence as there is in single cell studies (Mitchison, 1957)

indicates that the solution (2) is never quite reached, or that the 'constant' α takes different values during the period of one cell division. It is, however, clear that the deviation from equations (2) is a periodic one since the value of each cellular component is doubled after each division interval. A more accurate form of these equations will be derived in Section IV.

A further important feature of a system which generally obeys equations (2) is that the ratio of P to R remains constant. That this is in fact roughly the case for RNA and protein in many different types of cells has frequently been observed. Brachet (1959) has accumulated evidence that an increase in protein synthesis is always accompanied by increased RNA synthesis in embryonic development; and we may cite Maaløe (1960) for a particular example of this in microorganisms. Maaløe's observations show that RNA and protein vary together in rough proportion to the growth rate under a variety of environmental conditions as we expect in an autosynthetic system. The close coupling between RNA - and protein-synthesis has been demonstrated by Sands and Roberts (1956) and by Pardee and Prestidge (1956) while Ycas (1959) has put forward a particular model to account for these and his own observations on the intimate connections between the two synthetic mechanisms. His model does not imply a strict stoichiometry between the quantities of RNA and protein synthesized, the coupling being via a common nucleic acid-amino acid precursor complex. But the autosynthetic process does not

imply stoichiometry at the primary synthetic level either, there being only an asymptotic correlation between the concentrations due to some cyclic feed-back system whereby the synthesis of each depends upon some activity of the other. The significant observations for our purpose are those of Brachet (1959) cited above, and of Dalq (1957), for example, on the correlation between RNA and protein gradients in mammalian cells during development. Although these are perhaps only suggestive experiments, it is difficult to avoid the conclusion that RNA and protein are mutually dependent for their synthesis and that they always show comparable concentrations in developing cells. These species of macromolecule have become the primary observables in biochemical embryology. They can be regarded as the biochemical indices of development, and in the subsequent study the epigenetic variables will be defined in these terms.

SECTION II

SYSTEM AND ENVIRONMENT

We must consider now the most useful way in which to draw a distinction between system and environment in a developmental theory conceived as a Darwinian process. Mathematically there is no criterion whereby environmental parameters and system variables can be distinguished. The mathematical description merely states certain functional relationships between a set of symbols, and we are at liberty to transform the functions so that any set of symbols are independent variables (parameters), while the others are dependent variables (system variables).

The individual cell is in some sense a single system, and its surroundings are the environment whose properties can be in some degree altered independently of the state of the cell. However, there are certain cellular constituents which fall most naturally into the category of external parameters, insofar as they conform to certain conditions of relative constancy and furthermore may in a particular class of experiments be manipulated independently. Among these are the genes, and the associated experimental studies concerned are those of developmental genetics, The genes are, essentially, a partial set of initial conditions for the developmental system, defining

a certain range of potential catalytic activity. Genetic methods can be used to alter these initial conditions resulting in altered patterns of development.

There are many other essential determinants of development which enter the system as initial conditions. The embryo starts with a very complex functional architecture built into the egg, a microscopic morphology whose basic organisational features are perpetuated in descendant cells (see Kacser, 1960). These initial conditions must be accommodated within any set of equations seeking to describe development, and we will consider how they may enter a theoretical scheme in later pages. The genes, however, occupy a somewhat different position as external parameters of the developmental system, and must be considered as important determinants throughout the epigenetic process. Exactly what their relationship is to the developmental system will be a primary concern of this section. Our argument will begin with a consideration of how the cell is organised metabolically and what the pattern of interactions is among its various constituents.

It has been argued that protein and ribonucleoprotein (RNP) are the essential variables of the embryological system, and that together they have autosynthetic properties. However, as well as the cycle of synthetic activities described roughly by equations (1), an embryonic cell may be synthesising protein which has no direct functional relation to the metabolic activities leading to growth. These proteins, with the associated species of RNP, constitute the 'differentiated mass' of a cell which Weiss (1957) has distinguished from the 'generative mass' or reproductive part.

We shall extend this distinction slightly using Swann's important generalisation, which follows from the dissociability of growth and division (Swann, 1957), that a cell may become differentiated for division or for some other particular function.

The 'generative mass' is that set of protein and RNA species which define a closed catalytic cycle of synthetic activities having the autosynthetic property. This system, which brings about its own replication, we call the autosynthetic system of the cell. All other specialised synthetic activities, including those directly involved in cell division, fall into the category of heterosynthetic processes, as we shall call them. The relationship between autosynthetic and heterosynthetic processes may be represented as follows (Figure 1).

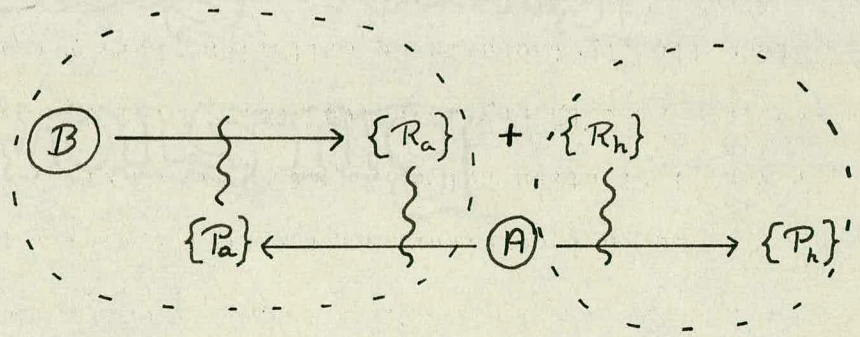


Figure 1.

{Ra} and {Pa} are the different species of RNA and protein which form the autosynthetic system. Their catalytic activities are shown as wavy lines, {Ra} catalysing the synthesis of {Pa} from precursors A, and {Pa} involved catalytically in the formation of {Ra} from precursors B. These substances B are also used in the synthesis of RNA species {Rh} which

produce 'heterosynthetic protein' $\{P_h\}$. Thus the autosynthetic system is that enclosed in the left circle, and the heterosynthetic system is that enclosed in the right. They both compete for precursors A and B, a point which will be discussed later. Together they form what we shall call the epigenetic system. The important distinction between the two is that there is no direct feed-back from $\{P_h\}$ onto $\{R_h\}$ in the heterosynthetic process. This does not imply that $\{R_h\}$ may not synthesise more of itself simultaneously with the synthesis of $\{P_h\}$, as has been suggested by Pardee (1956). It means only that $\{P_h\}$ is 'sterile' protein from the point of view of replication.

Heterosynthetic activities are dependent upon the autosynthetic system, but in the 'ideal' steady-state of growth (e.g. as (2)) all components increase at the same exponential rate. This follows from the argument given below. If R_h is any RNP-species belonging to the heterosynthetic system, its rate of synthesis is given by

$$\frac{dR_h}{dt} = k_1\{P_a\} = k_1\{P_a\}_0 e^{\alpha t}$$

$$\therefore \quad R_h]_{t\to\infty} = \frac{k_1 (R_h)_0 \{P_a\}_0}{\alpha} e^{\alpha t}$$

whence $$\frac{dR_h}{dt}\Big]_{t\to\infty} = k_1 (R_h)_0 \{P_a\}_0 e^{\alpha t} = \alpha R_h]_{t\to\infty}$$

The associated species of protein, P_h , has the rate equation

$$\frac{dP_h}{dt} = k_2 R_h = \frac{k_1 k_2 (R_h)_0 \{P_a\}_0 e^{\alpha t}}{\alpha}$$

$$\therefore \quad P_h]_{t\to\infty} = \frac{k_1 k_2 (P_h)_0 (R_h)_0 \{P_a\}_0 e^{\alpha t}}{\alpha^2}$$

whence $$\frac{dP_h}{dt}\Big]_{t\to\infty} = \frac{k_1 k_2 (P_h)_0 (R_h)_0 \{P_a\}_0 e^{\alpha t}}{\alpha} = \alpha P_h]_{t\to\infty}$$

The autosynthetic system can thus support exponential increase in both parts of the epigenetic system. This may be observed in tissue culture, in a culture of cells taken from the periosteal tissue of the chick, for example (Willmer, 1953). Even when these cells are growing and dividing rapidly, they lay down fibres between and within the cells in a manner characteristic of partially differentiated bone cells. The full exponential potential of the system can only be demonstrated in such tissues when optimal conditions are maintained in the culture medium, and it will be necessary to enquire into the factors which control the autosynthetic potential and lead to steady state conditions in embryonic cells.

It is the heterosynthetic system whose activity leads to the dramatic morphological changes which occur during differentiation. P_h may be myosin, collagen, haemoglobin, fibrin, lens protein, cholinesterase, etc. -protein characteristic of the differentiated state of a particular cell. They are involved in some specialised activity whose immediate result

is not an increase in the cell's 'survival potential' (i.e. this activity does not feed-back directly onto the basic replicative cycle of the autosynthetic system, which is the 'life' of a cell). In fact, these activities are often such as to lead ultimately to the death of the cell. Thus the contractability of an isolated muscle cell in no way extends its own life-time; the deposition of a collagenous matrix about a bone cell cannot serve its own immediate advantage; and the erythrocyte eventually loses its autosynthetic mechanism altogether, ensuring its early death.

The feed-back from heterosynthetic activities to the fundamental synthetic processes of cell survival and growth is slow and indirect compared with the immediate cycle of autosynthesis. The distinction between auto- and heterosynthetic systems is not an absolute one, but is relative to the time intervals for these feed-back cycles. Cell division may be looked upon as the specialised activity of a differentiated cell only if we agree to take time-periods of seconds or a few minutes, definitely less than the time between successive cell divisions, as significant intervals for observations on the epigenetic system. For if we extend the interval beyond the division time, then growth and division becomes a unitary process, the intimate coupling of the two activities preventing any reasonable separation between them. But if the processes which we decide to follow take place over a sufficiently small time interval, then our equations can describe the synthesis of spindle protein,

for example, as part of a heterosynthetic activity, with no significant feed-back in time intervals less than those between successive cleavages.

The specialised activities of differentiated cells tend to be mutually exclusive. Thus there are groups of heterosynthetic processes which lead to a particular differentiated state of a cell, and differentiation is the gradual emergence of one such group of processes, to the exclusion of others. Autosynthetic processes are also modified during differentiation, but the major discontinuities occur between the heterosynthetic activities of developing cells.

Consider now the intracellular localisation of the epigenetic system. This system must be shown to have some degree of autonomy in its replicative activities. There is some evidence that the cytoplasm of many cell types has an autonomous capacity for synthesis and growth in the absence of the nucleus. The most striking case of this is probably the giant alga Acetabularia whose enucleated cytoplasm showed active sunthesis of protein and RNA 14 days after nuclear removal (Brachet, et al., 1955). However, there are other cases where enucleation immediately results in a cessation of normal cytoplasmic activities, as in amoebae. And in all cases, the removal of the nucleus interferes with the synthetic processes of the cytoplasm sooner or later, so that there is certainly a question of indirect nuclear influence over cytoplasmic activities, if not immediate participation in them. It must therefore be concluded that the

cytoplasm is not entirely self-sufficient, although it may have a considerable synthetic autonomy in many cases. Its continued function always depends to some extent upon synthetic activities taking place within the nucleus.

There is the further fact that some heterosynthetic processes occur in the nucleus. For example, haemoglobin is synthesised in the nuclei of erythroblasts. Therefore, although the cytoplasm may be the primary site of autosynthetic activities, it certainly cannot be taken always to contain the whole of the epigenetic system.

We come now to the genes. There is a fundamental difference between the metabolic activities of the deoxyribonucleoprotein (DNP) of embryonic cells and those of the protein-RNP fraction. Whereas the latter species of macromolecule vary greatly in concentration between different types of cell in the developing organism, and within the same cell at different stages of development, the concentration of DNP per cell remains remarkably constant (Brachet, 1957). The metabolic stability of DNP as contrasted with the lability of RNP and protein has been further shown with the use of radio-isotopes (Brown and Roll, 1955). There is also a considerable amount of evidence for the view that the replicative mechanism for DNP, which brings about a doubling of the genetic material for each cell division, is independent of the 'epigenetic' replicative mechanism for RNP and protein (see e.g. Schaecter et al., 1958).

These biochemical facts show that the genetic

determinants, which are assumed to consist of DNP, do not form an integral part of the synthetic processes of the epigenetic system. However, it is known that the genes play an essential role in the control of developmental processes (Waddington, 1940), and so they must somehow influence the synthetic activities of the epigenetic system. The metabolic inertness of DNP relative to RNP and protein does not mean that different genetic determinents may not have differential 'activities' at different developmental stages within a cell. (The nature of this 'activity' and its changes will be discussed later). Some such hypothesis is a necessary feature of any theory of gene action during development. There is in fact direct evidence that some kind of chromosomal differentiation does occur in differentiating cells (Daly, Allfrey and Mirsky, 1952). Pavan (1954) has shown interesting polytene chromosomes differences at different stages of development; while the experiments of Briggs and King (1955) have also been cited in favour of some kind of nuclear differentiation during development. These observations are interpreted to mean changes in gene activities, which somehow cause changes in the concentrations of different species of RNP and protein.

The mechanism whereby changes in gene activity are transmitted to the epigenetic system remains thoroughly obscure. It is usually assumed that a genetic determinant, C_i, carries potential information which is expressed by the catalytic synthesis of an RNP template, R_i, whose base sequence is somehow determined

by that of C_i. R_i belongs to a class of ribosomal particles, all having the same base sequence, which is catalytically active in the synthesis of specific protein P_i.

We have seen that the epigenetic system, consisting of the classes R_i and P_i, is not entirely independent of nuclear components for its own maintenance. This observation is an important one. It shows that the replicative system of the cytoplasm is always subject to some decay and eventually dies when isolated from the nucleus. This means that the autosynthetic system, as we have defined it, is not in fact capable of indefinite replication in the manner described roughly by equations (2), even when all the normally required nutrients are made abundantly available (optimal conditions for growth).

One factor which might be expected to account for this is that there is always some destruction of the protein and RNP by inactivation and hydrolysis, so that these are removed from the autosynthetic system at finite rates: their instantaneous concentrations depend upon a balance between synthesis and degradation. However, it can be shown (Appendix 2) that the addition of destruction terms to the autosynthetic equations (2) is not a sufficient condition for the eventual decay of the system. Only if the decay rate is 'larger' in a particular sense than the autosynthetic coefficients will the system die out. The inequality which determines this decay of the autosynthetic system imposes a rather artificial restriction upon the coefficients, and it seems unlikely that this is the real

solution of the problem.

It is clear that equations (1) are only the roughest of approximations to the complete equations of the self-replicating cytoplasmic mechanism, expressing no more than its overall capacity for exponential increase under the proper conditions and over long time periods (greater than the division time). One of these conditions is now seen to be the presence of the nucleus. Therefore the smallest autonomous unit in the organism would seem really to be a single cell, and the proper representation of the self-perpetuating metabolic system is not Figure 1, but a double cyclic scheme such as the one which Waddington (1956) has considered in a discussion of nucleo-cytoplasmic interactions: one cycle is confined to the cytoplasm, and has direct feed-back on itself; the other cycle includes the genes in its feed-back loop. The 'system' thus becomes the whole cell, which certainly has a much greater degree of autonomy than the cytoplasm alone. However, there is no fully autonomous unit in the organism, and we face here the problem which was anticipated in the introduction: what is the biological system, and what is the environment? The finding that the autosynthetic system, as we have defined it, does not have complete autonomy, and that the degree of autonomy may vary greatly among different cell types, does not mean that we must redefine the epigenetic system so that it includes those factors on which its continued survival depends. In order to treat the genes as controlling parameters of the epigenetic system, as all the evidence

indicates we should, it is necessary to show that changes occurring in the epigenetic system are considerably more rapid than those occurring in gene activities. Thus the feed-back loop from the epigenetic system through the genes must involve a time lag many times greater than the 'relaxation time' of the epigenetic system. (The relaxation or characteristic time of the epigenetic system is defined in terms of the mean frequency of oscillations which the epigenetic system undergoes about its equilibrium state (steady state). It determines the observation time required in order to obtain good time averages for the epigenetic variables of state.) There are good reasons for believing that this condition may be satisfied in embryonic cells, as will be discussed in Section XI. It will then be possible to treat the two cycles of the cellular system independently, the first (not including the genes) defining the interactions within the epigenetic system, and the second (containing the genes) as 'information feed-back' from system to parameters. Such a treatment will require a very extensive modification of the original equations (1).

This procedure allows one to analyse in some detail the relations between the genes and the rest of the cell. There is compelling evidence, both biochemical and genetic, for the view that the genes are indeed distinct in some manner from other cellular constituents. They represent, essentially, a very stable source of specific 'information' (i.e. stable and specific catalytic potential) for the cell. The inability of an enucleated cell to maintain itself over prolonged time periods indicates

some intrinsic instability in the cytoplasmic replicative mechanism, such that there is an information 'leak' of some kind, making it dependent upon a flow of information from the nucleus. This information is usually assumed to be in the form of RNP-species, R_i, as stated above. It is presumably by controlling the flow of information to the cytoplasm that the genes can influence the composition and metabolic activity of a cell.

The cell membrane is a natural outer boundary of the molecular system which we are defining by protein and RNP concentrations. These macromolecular species include the catalytic elements which order the complex chemical substratum of the cell interior into characteristic metabolic patterns. Certain chemical substances will be entering a cell at particular rates, and others will be leaving and the cell membrane will determine what substances these may be. With their flow rates fixed, and the protein and RNA concentrations held constant (steady state of the epigenetic system), there will be established within the cell a particular steady state of metabolites. It is this steady state which is usually considered in studies of open metabolic systems, where the rate constants (enzyme activities) are regarded as the external parameters and the metabolic intermediates the variables. Such systems have been the object of much interesting study (e.g. Kacser, 1957) and they provide analogical explanations for many biological phenomena (overshoot, equifinality, variation in size with temperature,

etc.). However, this type of study gives no indication of how the rate constants themselves will vary, which must be one of the primary concerns of a developmental theory. The properties of the metabolic system are nevertheless of considerable importance in interpreting the consequences of changes in the epigenetic system, and how it interacts with its environment.

The embryonic cell has thus been divided into three major categories of constituents. One we have called the epigenetic system, consisting of different species of RNA and protein. This has been taken to determine the basic structural and functional properties of a cell at any moment of its development. It is dependent for its survival upon a complex substratum of chemical substances, an internal milieu, which is defined by general factors such as inorganic ions, sugar, amino acids, nucleic acids, and other intermediary metabolites. This substratum is metabolically ordered by the catalytic and structural properties of the epigenetic system into the 'metabolic system'. Some of the constituents of the metabolic system are continuous with the external milieu, outside the cell membrane, and thus become part of the environment of the epigenetic system, affecting its composition. Besides these general environmental factors, there are specific factors such as the matrix in which the cells grow, hormones, inductive substrates, organisers and embryonic inducers. The final set of environmental parameters, the genes, are the third major category of cellular constituents.

The epigenetic system, the metabolic system, and the genes do not by any means exhaust the enormous molecular spectrum of the cell. There are many important constituents such as fatty acids, sterols, lecithins, polysaccharides, etc., substances which are neither proteins nor nucleo-proteins nor metabolic intermediates of relatively small molecular weight. However, their concentrations are determined by the concentrations of the enzymes responsible for their synthesis and degradation, and by the concentrations of metabolites such as 'active acetate', ATP, glucose, amino acids, etc. They are therefore derivative constituents whose concentrations are in principle calculable from a complete knowledge of the boundary conditions of the cell (supply of nutrients), the state of the epigenetic system (which includes all the enzymes), and the initial conditions. We need not, then, include these biochemical species in a definition of the epigenetic state of the cell.

We turn now to the much discussed but little understood question of gene 'activities'. The metabolic system is a part of the environment of the genes, and it is normally assumed that gene activity depends largely on the metabolic composition of the cell. There are many possible mechanism, all consistent with orthodox kinetic principles, whereby the quantity of gene product can be varied continuously or discontinuously, in a manner dependent upon the state of the metabolic system, and hence, ultimately, upon the state of the epigenetic system and the boundary conditions at the cell surface (external milieu of the cell). As catalysts, the genes depend upon a supply of substrate for the expression of their

catalytic potential, and the metabolic system will determine the rate and the nature of this supply. Waddington (1957) has considered in detail how a competition between different genes for a common precursor can lead to a progressive divergence of their relative activities so that eventually the products of some genes will be produced at a much higher rate than others. This is a case of continuous and divergent changes in the activities of groups of genes, arising from differential rates of supply of different substrates from the cytoplasm.

A quite different model suggesting discontinuous gene activity has been proposed by Jacob, Schaeffer, and Wollman (1960). This model arises from a comparison of the properties of a class of genetic elements which have been called episomes (Ibid., 1960), and those of repressible and non-repressible metabolic mutants, both observed in microorganisms (E. coli, Salmonella). Without going into detail, the essentials of the hypothesis are that certain biochemical entities exist (episomes, repressors) which have a specific affinity for some part of the genome. These elements can exist in one of two states, either joined to a specific site on the genetic material, or free in the cytoplasm. In the first state they block the activity of certain genes or groups of genes, while in the second they do not interfere with gene activity. Gene control would in this case be not via substrate but via catalytic surface. However, the state of the episomes must be determined by the metabolic state of the cell, and so we are led once again

to the conclusion that gene activity is ultimately dependent upon the state of the epigenetic system and the boundary conditions of the cell.

We have now a complete cycle of interactions between the three major categories of cellular constituent, the epigenetic system determining the metabolic system, which determines the state of the genes, which affects the epigenetic system. There may also be feed-back cycles between any two of these, such as the direct (non-genetic) repression mechanisms that have been postulated for feed-back control of enzymes by the metabolites they produce. Our main concern will be with one part of this cycle: the intrinsic properties of the epigenetic system which account for its characteristic behaviour in response to stimuli from the genes and from the external milieu.

SECTION III

THE COMPONENTS

Within the epigenetic system there may be groups of protein species which form a functional metabolic unit, wherein the concentrations of the individual protein species are always adjusted to the same level. An example of such a 'functional unit' could be the enzymes of the Krebs cycle. These enzymes are involved in the generation of ATP from ADP and 'active acetate' (acetyl-coenzyme A). They form one of the primary energy generating systems of the cell. It is possible that there is some regulatory mechanism within this set of proteins which keeps their concentrations (more accurately, their enzyme activities) in rough proportion to one another throughout changes in the physiological state of the cell. In this case they are to be regarded collectively as defining a single component of the epigenetic system. In general, then, the proteins of the epigenetic system are to be partitioned into groups of species in such a way that the 'concentration' of each group can be varied independently of the others; while within a group the protein concentrations always increase or decrease together. These functional groups, consisting of one or more protein species (and the associated species of RNP), are the epigenetic components. The 'concentration' of a composite component will be an average over the concentrations of the component protein species,

whose concentrations are in any case assumed to be maintained at roughly similar values.

The Krebs cycle enzymes have never been shown to vary together in the manner suggested above. The fact that they are involved in one primary metabolic activity, the production of ATP, does not necessarily imply an internal regulatory mechanism which adjusts all enzyme activities to about the same level. However, the discovery of feed-back repression of enzyme activities by the terminal product of an enzyme sequence in micro-organisms shows that such metabolic control devices do exist in certain cells, and provide a very efficient method of adjusting the concentrations of groups of enzymes to their required levels (i.e. to the required level of enzyme product). By the strategic placing of repression points, the complex metabolic network of living cells can be partitioned into functional units concerned with particular metabolic activities which can be independently varied to produce different physiological states (see, e.g. Magasanik, 1959).

Whether or not such a mechanism is used in metazoon embryos for the control of enzyme activities remains to be demonstrated. It may be that each protein can be independently varied in concentration, there being no functional units within which all members have approximately the same activity. However, in order to accommodate the possible existence of composite functional units, which on circumstantial grounds seems quite likely, we shall define components in the general sense given above. They are groups of

protein species forming functional metabolic units within which the concentrations of the constituent proteins are not independently variable, but whose average concentrations may vary independently of those of other groups. Some of these may, of course, consist of a single protein species.

Out of these components the functional algebra of the cell is constructed; i.e. these are the units with which we may construct a 'block diagram' showing the flow pattern of matter and energy through the epigenetic system. There is a close formal similarity between this representation of cellular organisation and that which Rosen (1959) calls an M,R system. M is a metabolic unit in Rosen's formulation, corresponding to our more general functional unit of protein, P; while R is a 'repair unit', which we associate with a species of ribosome. However, Rosen's self-reproducing or self-repairing automaton is a static or algebraic representation of cellular organisation, and our primary concern is with the dynamics of its change. We must therefore consider now the differential equations which describe the synthesis and interactions of the epigenetic components.

SECTION IV

THE DIFFERENTIAL EQUATIONS

We donate by X_i the concentration of the ith. component of our epigenetic system. Under optimal conditions, with no limitations on precursors, all componenets will increase exponentially, and we can write

$$\frac{dX_i}{dt} = \alpha X_i \quad (i = 1, 2, \ldots n) \qquad \ldots\ldots\ldots\ldots\ldots (3)$$

α being the exponential coefficient.

However, there are many types of interactions in embryonic cells which prevent the full realisation of the autosynthetic potential so that α will vary in some way dependent upon the environment and the concentrations of the components.

$$\alpha = \alpha(X_1, X_2, \ldots X_n ; a_1, a_2, \ldots a_m)$$

the a_i's being the external parameters.

We are interested now in the possible steady state solutions which the equations (3) may have (defined by $\alpha(X_1, \ldots X_n ; a_1, \ldots a_m) = 0$) when we make different assumptions about the interactions occurring between the components, and about the manner in which the parameters a_i enter the function α. These steady states are the points of 'stable equilibrium' (Needham, 1950) which are assumed to occur in embryonic tissues between the transient motion caused by the

action of 'organisers', 'inducers', and other stimuli. They are therefore of primary importance, giving the epigenetic state of the tissue at such points. And the way the genes enter the function α as external parameters shows to what extent the occurrence of different steady states is under gene control.

For example, we could assume that the components compete for a limited supply of precursors for their own synthesis in such a way that an increase in any one reduces the synthetic activity of all, by reducing the available supply. Then α would take the form

$$\alpha = \alpha_i(a) - f(X;a) \qquad \begin{aligned} a &= (a_1, a_2, \cdots a_m) \\ X &= (X_1, X_2, \cdots X_n) \end{aligned}$$

where $\alpha_i(a)$ is a different function of the a_i's for each component, but $f(X;a)$ is the same function of X and a for all.

This is a sort of generalised logistic equation for many variables, its prototype being

$$\frac{dX}{dt} = X\ (\alpha - X) \qquad \text{where} \quad f(X) = X$$

However, it is shown readily that a system described by equations of the form

$$\frac{dX_i}{dt} = X_i \left[\alpha_i(a) - f(X;\ a) \right] \qquad (X_i < \infty)$$

is unstable in the sense that the components cannot coexist: one will always increase and eliminate the others. For, considering any two, we have,

$$\left(\frac{1}{Xi}\ \frac{dXi}{dt} - \frac{1}{X_K}\ \frac{dX_K}{dt}\right) = (\alpha_i - \alpha_K)$$

the function, f, being eliminated between them. Integrating, this gives,

$$\log \frac{Xi}{X_K} = (\alpha_i - \alpha_K)\, t + c, \quad \text{or} \quad \frac{Xi}{X_K} = A\, e^{(\alpha_i - \alpha_K)t}$$

If $(\alpha_i - \alpha_K) < 0$, then $\frac{Xi}{X_K} \to 0$ as $t \to \infty$, and so we see that $Xi \to 0$, ($X_K \not\to \infty$ because of limited precursors).

By considering all pairs of components, we find that the one with the largest α, say α_K, will eliminate all the others. Which one this may be is dependent upon the choice of $(a_1, a_2, \ldots a_m)$; but this sytem clearly does not describe the behaviour of an embryonic cell, wherein steady states occur with many components coexisting. Only if all the α's are equal will there be such steady states, given by solutions of the equation,

$$f(X_1, X_2, \ldots X_n; a_1, a_2, \ldots a_m) = \alpha \quad \ldots (4)$$

This equation (4) defines some n-dimensional surface with an infinite number of steady states, and no means of discriminating between them by different values of external parameters. Suppose, for example, that f is a linear function and that equation (4) takes the form,

$$a_1 X_1 + a_2 X_2 + \ldots + a_n X_n = \alpha$$

Then the steady states all lie in a hyperplane which is determined by the external parameters. But the particular point

on the hyperplane representing a steady state solution $(\bar{X}_1, \bar{X}_2, \cdots \bar{X}_n)$ will depend upon the initial conditions of the system, so that there is no possibility of the system travelling to some point which is independent of the starting point; there is no regulation. Such equations cannot, therefore, represent the motion of the epigenetic system which is known to be regulatory.

Williamson (1957) has discussed in the ecological context the general problem of obtaining points of stability for populations of many different species competing for limited food supply. Mathematically, it is necessary to find a set of n equations which will allow of discrete solutions for the Xi's. One way of doing this is to add a set of 'controlling factors' or 'preference factors' g_i to the equation (3) so that we get equations of the form,

$$\frac{dX_i}{dt} = X_i[\alpha_i(a) - g_i(X_i, a) f(X, a)] \quad \cdots \cdots \quad (5)$$

The functions $g_i(X_i, a)$ are monotonic increasing in Xi such that when Xi is small the bracketed expression is positive, while when Xi is large it is negative. This function therefore controls the sign of $\frac{dX_i}{dt}$ in a manner depending upon the size of Xi and upon the parameters a_i. In general there must be as many controlling factors as there are species (components) in the system. The steady states are now defined by the equations,

$$\alpha_i(a) - g_i(X_i, a) f(X, a) = 0 \qquad (i = 1, 2, \cdots n)$$

The solutions of these equations will be isolated if the functions gi and f are analytic, and the parameters ai will then determine what those solutions are. The genes would therefore, in such a system, define a discrete set of points (steady states) to which the system could move, each point being surrounded by a region under its domination ('confluent set' Ashby, 1957). Regulation now occurs within a confluent set so that when the state of the system falls within such a set, it will eventually reach the corresponding steady state. This is much more like the behaviour of a developing cell, the genes determining a certain number of possible steady states, the one to which the system moves being determined by its initial conditions. However, a further analysis of the system could not proceed without specific knowledge of the nature of the functions gi4f, and this we do not have at present. Thus, the representation of the epigenetic system in the above manner, with a whole set of control factors superimposed upon the equations expressing competition for the raw material of protein synthesis, does not give any obvious way into a general statistical treatment of the system's behaviour analogous to that of statistical mechanics in the physical sciences. There are many other possible representations of the manner in which components and external parameters enter differential equations of the general type (3) consistent with what we know of interactions and environmental influences in embryonic cells (which is very little). However, there is one general method of analysis which leads us directly

into a situation just like that in statistical mechanics, and this we will now investigate.

Consider the possibility of a 'binary analysis' of the interactions between epigenetic components. By this we mean taking any two components of the system, X_1 & X_2, say, and considering their effects to one another, ignoring all the other components for the moment. It is important to remember here how the components have been defined: they are functional units which are independently variable and in competition for precursors for their own replication. We now argue that under any given set of environmental conditions, there will be a dominance relation between any two components, such that one will be synthesised at the expense of the other when precursors becoming limiting for protein and RNA synthesis. The direction and the magnitude of this dominance (which may be zero) will be determined by the genes and by the other environmental factors. Supposing that X_1 dominates X_2 in this manner, the effect of X_1 on X_2 can be written, in a first order approximation, as

$$\frac{dX_2}{dt} = (\alpha'_{21} - \alpha_{12} X_1) X_2 \qquad (\alpha_{12} > 0)$$

where $\alpha_2(X_1, X_2, \ldots X_n; a_1, a_2, \ldots a_m)$ has been replaced by $(\alpha'_{21} - \alpha_{12} X_1)$, α'_{21} being still a function of the a_i's and of the components other than X_1, while α_{12} is now determined by the a_i's only (i.e. we have expanded α_2 to the first order in X_1). Since whatever reduction has occurred in

the rate of X_2 -synthesis is added to X_1 -synthesis, we have,

$$\frac{dX_1}{dt} = (\alpha'_{10} + \alpha_{12} X_2)X_1 ,$$

α_1 being increased by the factor $\alpha_{12} X_2$, where again α'_{10} is a function of the parameters a_i and components other than X_2, while α_{12} is the same term as appears in the expression $\frac{dX_2}{dt}$. These equations imply that if some X_2 is added to the system when precursors are in limited supply, then some of it will be converted to X_1; while if X_1 is added to the system, it will bring about a further conversion of X_2 into X_1: the system is 'weighted' in favour of X_1-synthesis. The quadratic terms occur because of the autosynthetic nature of the underlying synthetic activities.

If we extend the above arguments to all pairs of components, expanding the functions α_i to the first order in all the variables and assuming a dominance relation between each pair, we get equations of the form,

$$\frac{dX_i}{dt} = X_i\left[\alpha_{io} + \sum_j \alpha_{ij} X_j\right] \quad \text{- - - - - - - - -} \quad (6)$$

$$(i = 1, 2, \text{- - - } n)$$

$$\alpha_{io} = \alpha_{io}(a)$$

$$\alpha_{ij} = -\alpha_{ji}$$

This analysis selects a particular type of interaction between pairs of components as primary for the description of how the whole system behaves. It assumes that the interactions which occur between components can be reduced to effects of the type

described by equations (6). If no such dominance relation holds between a pair of components, then they are said to show no binary interaction, and the binary coefficient α_{rs} for this pair will be 0. In this case there is no conversion of component r into component s (or vice versa). This does not mean that these components can have no other effects within the system. The binary coefficients α_{ij} may themselves be functions of the components, as will be seen later. But the implication is that, whatever the influence which one component may have upon another, that influence can be represented as affecting binary interactions of the type described.

This argument is valid only if one protein or RNA species can be converted into another in embryonic cells. The dynamic state of these macromolecular species in living cells is a well known fact, a given concentration being maintained only by constant synthesis and degradation. The question now is whether or not the amino-acids and nucleic acids derived from the hydrolysis of protein and RNA are recycled through the pool of precursors for new syntheses, or whether they are discarded by the cell as waste products. In rapidly growing cells such as those in bacterial cultures there seems to be little or no conversion of one protein species into another, the synthesis being almost irreversible (e.g. Spiegelman and Halvorson, 1957). However, Dubnoff (1955) has produced evidence that there is a reversible cycle of synthesis, inactivation, and hydrolysis in some adaptive enzymes of E. coli (maltase, lactase, β-galactosidase

and others), the enzymes being stabilised by the presence of the appropriate substrate, with a resulting increase in its concentration. This is precisely the sort of labile behaviour which is here postulated for the proteins of the embryonic cell. These cells grow and divide slowly compared with bacteria, and it has been suggested (Paul, 1958), that an internal reversible turnover of these macromolecules is necessary for differentiation in embryonic tissues. Some very important observations bearing on this point were reported by Ebert (1950) and confirmed and extended by Clayton (1954), showing that certain proteins (e.g. myosin, lens protein) which can be detected immunologically in chick embryos appear in wide distribution in embryonic tissues at early stages of development, later disappearing from all tissues except those whose differentiated state is characterised by high concentrations of the particular proteins. Another example of this type of behaviour has been given by Mulnard (1955) for alkaline phosphatase. There is no direct evidence that the specific protein which disappears from tissues not ultimately developing into heart or eye or nervous system is actually hydrolysed and converted into other species of protein; it may disappear simply by dilution, and be ultimately discarded. However, it would be surprising if a system as complex and dynamic as an embryonic cell did not show 'overshoot' phenomena in protein and RNA synthesis as well as in other types of activity, (overshoot in the metabolic system is well known) with fluctuations occurring in the quantities of different species. Cycles of protein

synthesis and hydrolysis have been observed by Kavanau (1954) for the sea-urchin and there seems good reason to believe that there is a fairly free interconversion of different macromolecular species during embryonic development, as tissues respond to different stimuli.

The equations (6) imply strict equivalence between the 'values' of different components with respect to their interconvertibility. However, as we have defined them, the different components may vary greatly in the number of proteins which constitute them, and so we must introduce equivalent numbers β_i. These parameters are positive constants which are determined by the number and size (molecular weight) of the proteins in the ith component. If $\beta_i = 2$ while $\beta_j = 1$, then it takes 2 j-components to make one i-component. Taking $\alpha_{ij} > 0$, the incremental increase in X_i at the expense of X_j in time Δt is,

$$\frac{\alpha_{ij}}{\beta_i} X_i X_j \Delta t$$

while X_j decreases incrementally by an amount $\frac{\alpha_{ij}}{\beta_j} X_i X_j \Delta t$. The ratio of the quantity of X_j decreased to X_i increased is $\beta_i/\beta_j = 2$ in our example. The corrected equations now have the form,

$$\frac{dX_i}{dt} = X_i\left[\alpha_{io} + \frac{1}{\beta_i}\sum_j \alpha_{ij} X_j\right] \quad \cdots\cdots\cdots (7)$$
$$(i = 1, 2, \cdots n)$$

These equations are identical with those first studied by Volterra(193 for the general prey-predator relations in wild populations and

his important results will be the basis of our theory.

By assuming antisymmetry between the coefficients, $\alpha_{ij} = -\alpha_{ji}$ we are implying that $\alpha_{ii} = 0$. Thus the equations have no self-interaction terms, $-\alpha_{ii}X_i^2$, of Verhulst-Pearl type. Volterra showed that such terms give rise to a kind of frictional damping in the system bringing it ultimately through damped oscillations to a particular state. Nearly all equations of growth are based upon this type of 'logistic' behaviour for their stability, and it seems likely that such 'frictional' forces do occur in biological systems. However, we are interested in relatively short time changes in the system, from one disturbance to the next. Under such conditions, it would seem possible to ignore the slow damping forces, and to investigate the statistical behaviour of the undamped system.

It will be found that (7) cover a much larger range of possible interactions among components than is at first evident. This important extension of the Volterra equations, without altering the properties which gives them their importance for a statistical analysis, was obtained by Kerner (1959) in the second of his two papers. He showed that the coefficients α_{ij} could be practically any function of the variables X_i, subject only to a mild constraint which we will discuss later, providing only the antisymmetry properties are kept. With each choice of function, we get a different statistical distribution for the components, and this gives the basic equations a very great

flexibility in accommodating a wide range of possible observations. However, in the present study we shall restrict ourselves largely to the simplest case where the α_{ij}'s are functions of the external parameters only, while keeping in mind the greater power of the formulation.

A kinetic derivation of the differential equations (7) should be possible, showing what properties must be assumed and under what conditions they hold. However, the kinetics of macromolecule synthesis from a precursor pool are still unknown, and even with simple mass-action assumptions there are too many simultaneous equations to allow of any adequate kinetic treatment. In the face of these difficulties we are forced to use the more general analytical approach given above, and to justify our assumptions by the results they give. We shall now investigate the consequences of this model. (In the following four sections we will draw heavily upon the mathematical results of Volterra (1931, 1937), and Kerner (1957, 1959).

SECTION V

THE STATISTICAL MECHANICS

Of particular interest in the theory are the stationary states of the epigentic system, for which the derivatives $\frac{dXi}{dt}$ vanish. Calling the stationary values of the variables q_i, these quantities are defined by the equations,

$$(8) \quad \alpha_{i0}\beta_i + \sum_j \alpha_{ij} q_j = 0 \qquad (i = 1,2 \ldots n)$$

We shall assume that these equations have a unique solution with all q_i positive, possible only if the number n of different components is even, and if all the α_{i0} do not have the same sign. The first condition arises when we try to solve the matrix question,

$$(9) \quad Aq = p \qquad p = (-\alpha_{10}\beta_1, -\alpha_{20}\beta_2, \ldots -\alpha_{n0}\beta_n)'$$

$$A = (\alpha_{ij})$$

$$q = (q_1, q_2, \ldots q_n)'$$

A is an antisymmetric matrix which is nonsingular only if n is an even number. The second condition appears when we study the stability of the system in terms of the equations (7). Writing these in the form,

$$\beta_i \frac{\partial X_i}{\partial t} = \left[\alpha_{i0}\beta_i + \left(\sum_j \alpha_{ij} X_j\right)\right] X_i \ ,$$

if now we add over all i, the quadratic terms cancel in pairs, and we are left with,

$$\sum_i \beta_i \frac{dXi}{dt} = \sum_i \alpha_{io} \beta_i X_i$$

Suppose now $\alpha_{io} > 0$, all i ($\beta_i > 0$ also), and suppose α is a lower bound for the α_{io}'s. Then assuming that not all Xi are zero, we can write

$$\sum_i \beta_i \frac{dX_i}{dt} \geq \alpha \sum_i \beta_i X_i \text{ , or } \frac{\sum_i \beta_i dX_i}{\sum_i \beta_i X_i} \geq \alpha dt$$

Integrating, we have,

$$\sum_i \beta_i X_i \geq a e^{\alpha t} \qquad \text{..... (10)}$$

Therefore the system increases without limit, and there can be no stationary state. If all $\alpha_{io} < 0$, a similar argument shows that all components must eventually become zero. We must therefore assume that not all α_{io}'s have the same sign.

Let us now see if there is any functional basis for distinguishing between components with + and components with - coefficients. In order to do this, we consider the condition of the system when all components are present in very small concentration, so that the quadratic terms in the differential equations are much smaller than the linear ones. Referring to Figure 1, it seems possible that the autosynthetic components will have positive linear

coefficients, while the heterosynthetic components will have negative ones. The implications of this are that the heterosynthetic system will always lag slightly behind the autosynthetic in its response to changing environmental conditions. In view of the latter's dependence on the former, this seems highly probable. The magnitude of this effect can be seen if we write down the solutions of equations (7), using the formulation of infinitesimal operators for the differential equations (see Appendix 3). The solutions are, to first order in t,

$$Xi = (Xi)_o e^{\alpha_{io} t + \left[\frac{1}{\beta_i}\sum_j \alpha_{ij}(X_j)_o\right]t + O(t^2)}$$

where $(Xi)_o$ is the initial value of Xi, at t = 0. When the system is allowed to grow from its initial condition at t = 0, with all $(Xi)_o$'s small, the autosynthetic system is the first to increase and the heterosynthetic system will increase only when the autosynthetic components have increased to some extent (assuming that some of the heterosynthetic components can increase at the expense of some of the autosynthetic ones). It can be shown that, using the above distinction between +ve and -ve linear coefficients, any factor which tends to destroy the components of both systems indiscriminately will in general have the effect of shifting the steady-state in favour of the autosynthetic system, a property which is consistent with some observations on the general behaviour of cells (see Appendix 4). The autosynthetic system is the primary one in living cells, and this response to destructive stimuli may be the basis of a type of regulative behaviour which protects the fundamental autosynthetic cycle.

This assumption has the further implication that stability in the epigenetic system is dependent upon the coupled action of both the growing and the differentiating parts. The autosynthetic system alone would grow indefinitely, as shown by equation (10), whereas the heterosynthetic system alone would decay. Some uncoupling of the two systems may underly the occurrence of malignant growth in cells which have escaped the normal regulating mechanisms, and undergo an expansion process without any stable (stationary) points.

The distinction between the auto- and the heterosynthetic systems need not be as rigid as we have suggested. There may be autosynthetic components with negative linear coefficients, and heterosynthetic components with positive ones; we need assume only the existence of both positive and negative coefficients without specifying which components are which. In fact, the requirement for positive and negative terms may well be something of a mathematical artefact, a stability condition imposed by the equations (7) which in the biological system may be satisfied in another way. This may also be true of the parity condition, that there be only even numbers of components. This last requirement is clearly a very artificial one, arising from the idealised mathematical treatment.

Two factors, ignored by equations (7), may be expected to account for the difference between the real situation and the highly abstracted mathematical one which we are considering. The first of these is the occurrence of frictional damping, which makes the system

stable independently of the signs of the α_{ij}'s and removes the parity condition. This stability due to self-damping terms is most probably a long-term result, its effect on the short-term processes which we want to study being, effectively, the occurrence of positive and negative linear coefficients. Which of these coefficients are positive and which negative may then depend upon the particular point in the long-term process which is being considered.

The other feature of the real system which is not considered in equations (7) is the spatial element: the time-lags involved in the diffusion of metabolites and macromolecules, and the spatial separation of the different reaction sites. The Volterra equations apply only to a system whose components are uniformly distributed in space, and wherein the interactions are instantaneous. These conditions will certainly not be realised within a cell, and the approximation given by the assumptions of homogeneity will be progressively less valid as differentiation proceeds, and the morphology of a cell becomes more significant for its activity. We must therefore be prepared to modify the original equations in order to accommodate these spatial factors, and the manner in which this may be done will be discussed more fully in Section XI. Suffice it for the present to say that the mathematical requirements of the Volterra equations are not to be taken too literally as indicating significant features of the biological system, and their interpretation may be fairly loose.

Assuming that the conditions are satisfied for a unique solution of equations (8), it is possible to obtain an integral of

the Volterra equations which will form the basis of the statistical theory. Rewriting equations (7) in the form,

$$\frac{\beta_i}{X_i}\frac{dXi}{dt} = \alpha_{io}\beta_i + \sum_j \alpha_{ij} X_j \qquad \ldots\ldots\ldots (11)$$

and eliminating $\alpha_{io}\beta_i$ between these and equations (8), we get the equations,

$$\frac{\beta_i}{X_i}\frac{dXi}{dt} = \sum_j \alpha_{ij}(X_j - q_j)$$

We now introduce the new variables,

$$x_i = \log \frac{Xi}{qi}$$

as suggested by Kerner (1957) which puts the equations in the form,

$$\beta_i \frac{dx_i}{dt} = \sum_j \alpha_{ij} q_j (e^{x_j} - 1) \qquad \ldots\ldots\ldots (12)$$

Multiply both sides by $qi(e^{x_i}-1)$, and add over i ;

$$qi(e^{x_i}-1)\frac{dx_i}{dt} = \sum_i \sum_j \alpha_{ij} q_i q_j (e^{x_i}-1)(e^{x_j}-1)$$

The right-hand side vanishes, because of the antisymmetry of the coefficients α_{ij} , and so we can integrate the left-hand side to give us the expression,

$$qi(e^{x_i}-x_i) \equiv G = \text{constant} \qquad \ldots\ldots\ldots (13)$$

This integral was found by Volterra, and it is the only one which is apparent. Volterra showed that the Xi are variable between finite limits, and that they fluctuate continuously without damping.

Furthermore, the time average of these variables are the qi's, which are therefore independent of the initial conditions. This is seen by integrating equations (7) between 0 and t, giving,

$$\frac{\beta_i}{t} \log \frac{Xi(t)}{Xi(0)} = \alpha_{i0}\beta_i + \sum_j \alpha_{ij} \frac{1}{t} \int_0^t X_j(t)\,dt \quad \ldots\ldots(14)$$

As $t \to \infty$, remembering that Xi(t) remains bounded, we find that the time averages of Xi , namely

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t X_i(t)\,dt$$

satisfy the same equations (8) as qi.

In order to construct a statistical mechanics with the Volterra equations, the first requirement is to show that Liouville's theorem is satisfied. Using the usual hydrodynamical argument, we represent a point of an n- dimensional fluid by $(x_1, x_2, \ldots x_n)$, its density by $\rho(x_1, x_2, \ldots x_n)$, and its velocity by $\underline{v} = (\dot{x}_1, \ldots \dot{x}_n)$. The purpose of introducing the variables $x_i = \log \frac{Xi}{qi}$ is precisely to obtain a Liouville theorem for the system. The hydrodynamical equation of continuity is,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \underline{v}) = \frac{\partial \rho}{\partial t} + \sum_i \frac{\partial (\rho \dot{x}_i)}{\partial x_i} = 0$$

Expanding the sum, we have,

$$\frac{\partial \rho}{\partial t} + \sum_i \dot{x}_i \frac{\partial \rho}{\partial x_i} + \sum_i \rho \frac{\partial \dot{x}_i}{\partial x_i} = 0$$

The second sum vanishes since $\dot{x}_i$ is independant of x_i ,

and we are left with,

$$\frac{D\rho}{Dt} = \frac{\partial \rho}{\partial t} + \sum_i \frac{\partial \rho}{\partial x_i} \dot{x}_i = 0$$

which is the required theorem.

There are other variables besides x_i which provide a Liouville theorem for the Volterra system, and from them a class of 'canonical' transformations could be constructed. However, the x_i are convenient for the present purpose, varying as they do between $-\infty$ and $+\infty$ (unlike the Xi which are everywhere +ve), and giving an integral which is decomposable into components

$$G_i = \beta_i q_i (e^{x_i} - x_i)$$
$$G = \sum_i G_i$$

We can now see how general the Volterra equations can be if all that is required of them is a Liouville theorem. If we introduce a new parameter $\gamma_{ij} = \frac{\alpha_{ij}}{\beta_i \beta_j} = -\gamma_{ji}$, then the differential equations (12) may be written in the partial differential form,

$$\dot{x}_i = \sum_j \gamma_{ij} \frac{\partial G}{\partial x_j} = \sum_j \frac{\alpha_{ij}}{\beta_i} q_j (e^{x_j} - 1)$$

The continuity equation is, again,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \underline{v}) = \frac{\partial \rho}{\partial t} + \underline{v} \cdot \nabla \rho + \rho (\nabla \cdot \underline{v})$$
$$= \frac{D\rho}{Dt} + \rho (\nabla \cdot \underline{v})$$

Then,

$$\frac{D\rho}{Dt} = 0 \quad \text{only if} \quad \nabla \cdot \underline{v} = 0$$

$$\text{i.e.} \quad \sum_i \frac{\partial \dot{x}_i}{\partial x_i} = \sum_i \sum_j \left(\frac{\partial \gamma_{ij}}{\partial x_i} \frac{\partial G}{\partial x_j} + \gamma_{ij} \frac{\partial^2 G}{\partial x_i \partial x_j} \right)$$

$$= \sum_i \sum_j \frac{\partial \gamma_{ij}}{\partial x_i} \frac{\partial G}{\partial x_j} = 0$$

This is a constraint upon the γ_{ij}'s and upon G, and it is not very restrictive. With arbitrary G, we could choose γ_{ij} to be independent of Xi and Xj . Then the original equations would be defined by,

$$\dot{x}_i = \sum_j \gamma_{ij} \frac{\partial G}{\partial x_j}$$

which could be any one of a large range of possibilities, chosen either to fit the statistics (through G) or the kinetics (through γ_{ij}'s and G) of the experimental observations. We restrict ourselves in the following to the case of constant γ_{ij}'s, for in this form we can study most closely the steady-state solutions and how they vary with changes in the external parameters.

SECTION VI

THE CANONICAL ENSEMBLE

The integral G is made up of a sum of terms which we call Gi, giving a convenient separation of the system into components,

$$G \equiv \sum_i Gi = \sum_i \beta_i q_i (e^{\chi_i} - \chi_i)$$

Each Gi has a minimum value $\beta_i q_i$ occurring when $\chi_i = 0$ (Xi = qi, at the steady-state). The variables χ_i have a range of variation from $-\infty$ to $+\infty$, and the surface G encloses a simply connected region of phase-space which increases monotonically for both +ve and -ve increases in the χ_i. G therefore has all the properties required for constructing a statistical mechanics (see Khinchin, 1949).

Following Gibbs, we define an ensemble in statistical equilibrium as one for which $\frac{\partial \rho}{\partial t} = 0$. The density, ρ, is usually a function of G alone, and then the ensemble average of any phase function $f(\chi_1, \chi_2, \ldots \chi_n)$ is defined as

$$\bar{f} = \frac{\int \rho f \, d\chi}{\int \rho \, d\chi} \qquad d\chi = d\chi_1 d\chi_2 \cdots d\chi_n$$

the integrals being taken over phase space from $-\infty$ to $+\infty$. ρ is therefore of the nature of a probability density, and our assumption

concerning the ergodic nature of the Volterra system allows us to replace time averages by phase averages.

We consider now the analogue, in the Volterra statistical mechanics, of the conceptual model which Gibbs called the canonical ensemble (Gibbs, 1902). We assume that an epigenetic system within a cell does not have its G constant, but that there is allowed some 'G-exchange' between the cell and adjacent cells which are taken to have the same mean G. (Just what 'G-exchange' means for embryonic cells is something that will be discussed in detail below). The system is now only one component of a much larger ensemble. It is no longer constrained to move only on the surface G = constant, but can move freely in the phase space. The question now is how the point representing the system will be distributed in phase. According to a basic proposition in statistical mechanics, the distribution law is,

$$\rho = e^{\frac{\psi - G}{\theta}}$$

defining the Gibbs canonical ensemble. The probability distribution is normalised,

$$\int e^{\frac{\psi - G}{\theta}} \, dx = 1$$

whence we get the Gibbs phase integral,

$$Z \equiv e^{-\psi/\theta} = \int e^{-G/\theta} \, dx \quad (\psi \text{ being independent of } x)$$

The canonical ensemble is a very important concept in physics because it allows one to study systems which are not isolated but are

in thermal equilibrium with their surroundings. The many duplicates of the initial system, which surround it and with which it can exchange energy may be regarded as a 'heat bath' in which the system is immersed. θ is the thermodynamic temperature of the ensemble while ψ is the free energy of the system in thermodynamic equilibrium. This same construction can be used to study the 'thermodynamic' properties of the epigenetic system. In the following, all the thermodynamic functions and parameters will be prefixed by the term 'epigenetic' when we are discussing their analogues for the biological association. Thus the familiar meanings of words like temperature, free energy, work, will not be lost, but will be transposed to this new situation, where their interpretation will be different but many of their basic properties unaltered.

We have already noted that in the Volterra statistical mechanics the integral G is decomposable. Unlike the situation in physics, where decomposability necessitates the introduction of a small energy of interaction between components in order to provide for energy exchange between them, the Volterra components may interact even when they are rigorously separated mathematically, the interactions being inherent in the dynamics.

The canonical ensemble is the appropriate tool for the study of one part of the epigenetic system, consisting of say ν components out of the total of n in the cell. This part will not have its G fixed, but will exchange G with the rest of the cell. This means that the ν components under consideration interact freely

with the other components of the cell in the manner described by the equations, competing for precursors, being synthesized and destroyed, and generally taking part in the cellular processes of growth and differentiation. G-exchange can take place only between components which are interacting in this manner, and so it would seem in general to be confined to the components within a single cell. However, if under certain conditions there is cytoplasmic continuity between adjacent cells, allowing a free exchange of metabolites and macromolecules between them, then G-exchange could occur inter-cellularly, with the establishment of a mean G over two or more cell systems. It seems doubtful that this condition is ever an important one in embryonic systems, intercellular forces being generally of a somewhat different type from intracellular ones, and requiring therefore a different analytical treatment.

Returning to the subsystem of ν components within a single cell, the other $n-\nu$ components are now regarded as forming a 'heat-bath' in which the subsystem is immersed. The canonical ensemble shows that if ν is large, then the fluctuations of the subsystem about its mean G will be small; but when ν is small, then these fluctuations may be quite substantial. In any experimental work, the latter situation is certainly the more likely. It is not yet possible to observe a few protein species continuously in a single embryonic cell; indeed, there are very few techniques whose quantitative resolution reaches the single cell. However, it is not unreasonable to envisage a further refinement of technique

allowing for such close observation, and then the observations will most likely be made on a very few proteins. The fluctuations occurring in their mean G value will then be expected to be relatively large, reflecting the pattern of their individual fluctuations. The experimental side of the present theory will be taken up again in a later section.

The first phase average which we want to investigate is that of

$$Q_i \equiv \frac{\partial G}{\partial x_i} = \beta_i q_i \left(\frac{x_i}{q_i} - 1 \right) \qquad \ldots\ldots\ldots (15)$$

$$\bar{Q}_i = \int \frac{\partial G}{\partial x_i} e^{-G/\theta} \, dx \Big/ \int e^{-G/\theta} \, dx$$

$$= -\theta \int_{-\infty}^{\infty} \frac{\partial}{\partial x_i} \left(e^{-G_i/\theta} \right) dx_i \Big/ \int_{-\infty}^{\infty} e^{-G_i/\theta} \, dx_i$$

$$= -\theta \left[e^{-\frac{\beta_i q_i}{\theta} (e^{x_i} - x_i)} \right]_{-\infty}^{\infty} \Big/ \int_{-\infty}^{\infty} e^{-G_i/\theta} \, dx_i$$

$$= 0$$

This means that the phase average of Xi is qi , and this we have already shown to be the time average of the system, by equation (14). This suggests that the system is ergodic, with the phase trajectory of a representative phase point travelling over the surface $G(x_1, x_2, \cdots x_n)$ = constant, in such a way as to cover all of it but a set of measure zero, given sufficient time. Volterra showed that for small (linearised) fluctuations about the steady state, all components show undamped oscillations. We will adopt the position that this is also true for the non-linear system; i.e. that the system is ergodic with phase averages equal to time averages.

SECTION VII

THE EPIGENETIC TEMPERATURE

Consider now the phase average of Qi^2,

$$\overline{Qi^2} = (\beta_i q_i)^2 \overline{\left(\frac{X_i}{q_i} - 1\right)^2} = \int \left(\frac{\partial G}{\partial x_i}\right)^2 e^{-G/\theta} dx \Big/ \int e^{-G/\theta} dx$$

$$= \left\{ -\theta \left[\frac{\partial G}{\partial x_i} e^{-G_i/\theta} \right]_{-\infty}^{\infty} + \theta \int_{-\infty}^{\infty} e^{-G_i/\theta} \frac{\partial^2 G_i}{\partial x_i^2} dx_i \right\} \Big/ \int_{-\infty}^{\infty} e^{-G_i/\theta} dx_i$$

Now

$$\frac{\partial^2 G_i}{\partial x_i^2} = \beta_i q_i e^{x_i} = \frac{\partial G}{\partial x_i} + \beta_i q_i$$

and so the above expression becomes,

$$\overline{Qi^2} = \theta \left\{ \int_{-\infty}^{\infty} e^{-G_i/\theta} \frac{\partial G_i}{\partial x_i} dx_i \Big/ \int_{-\infty}^{\infty} e^{-G_i/\theta} dx_i \right\} + \theta \beta_i q_i$$

$$= \theta \beta_i q_i$$

Therefore, $$\theta = \beta_i q_i \overline{\left(\frac{X_i}{q_i} - 1\right)^2} \qquad \text{..... (16)}$$

$$= \beta_i q_i \frac{\overline{(X_i - q_i)^2}}{q_i^2} \quad \text{for all components i.}$$

This is a result of considerable interest in the theory, for it shows that the epigenetic temperature, θ , is a measure of the mean square deviations of the components from their stationary concentrations, qi . This general parameter is an indicator of the level of excitation of the system, being zero only when there are no fluctuations and all components stay at their stationary values qi. The equilibrium condition of the epigenetic system is now defined as that for which the expression (15) has the same value for all components, so that the quantity,

$$\sum_{i=1}^{n} \frac{Q_i^2}{\beta_i q_i}$$

is equally partitioned among all components on the average. There is another quantity which is equally partitioned at equilibrium. We write,

$$\begin{aligned} Ti &= x_i \frac{\partial G}{\partial x_i} \\ &= \beta_i \, qi \left(\frac{X_i}{q_i} - 1 \right) \log\left(\frac{Xi}{qi}\right) \end{aligned}$$

which is seen to be a loose analogue of kinetic energy. Taking its phase average, we have,

$$\begin{aligned} \overline{T_i} &= \int x_i \frac{\partial G}{\partial x_i} e^{-G/\theta} \, dx \Big/ \int e^{-G/\theta} \, dx \\ &= \left\{ -\theta \left[x_i e^{-G_i/\theta} \right]_{-\infty}^{\infty} + \theta \int_{-\infty}^{\infty} e^{-G_i/\theta} \, dx_i \right\} \Big/ \int_{-\infty}^{\infty} e^{-G_i/\theta} \, dx_i \\ &= \theta \qquad \text{..... (17)} \end{aligned}$$

Thus the quantity $\sum_i T_i$ is equipartitioned at θ per degree of freedom in the equilibrium situation. From the two expressions (15)

and (16), we get the result,

$$\frac{\theta}{\beta_i q_i} = \frac{\overline{(Xi - qi)^2}}{qi^2} = \frac{\overline{(Xi - qi)}}{qi} \log\left(\frac{Xi}{qi}\right) \quad \text{.....(18)}$$

With some refinement in experimental technique for determining protein concentrations, Xi, this relationship should be susceptible of verification. The mean quantities in (17) make $\frac{\theta}{\tau_i}$ ($\tau_i = \beta_i q_i$) directly available observationally through Xi and qi. In the theory of the canonical ensemble, θ is defined only to within a scale factor which is arbitrarily fixed by two reference points and the number of divisions assigned to the interval (degrees Fahrenheit, Centigrade, etc., in physical thermometry). In physical systems, the zero point of absolute temperature is a completely 'quiet' state with no random motion of particles: their energy states are fixed and can be determined with certainty. Similarly, in the present theory, the state where $\theta = 0$ is characterised by an absence of random motion, or fluctuation, each component having $G_i = \tau_i$, which we may call the zero point Gi or the intrinsic temperature. The expression $\frac{\theta}{\tau_i}$ is a pure number (see (18)), and thus τ_i ($= \beta_i q_i$) has the same units as θ. Since qi is a concentration, the parameter β_i will have units determined by those of θ .

In a statistical mechanics, the parameter θ is also of significance in indicating the direction of preferred flow of G from one system to another weakly coupled to it. The one with higher θ will tend to lose G to the one with lower θ , so that when the two systems have equilibrated they will both have the same $\bar{G}$, hence the same θ . This means that there will be an adjustment of concentrations and oscillations

among the different components of both systems until,

$$\beta_i q_i \frac{\overline{(X_i - q_i)^2}}{q_i^2} = \beta_K q_K \frac{(X_K - q_K)^2}{q_K^2} \quad \ldots\ldots\ldots (19)$$

for all pairs of subscripts i, K .

Consider now the probability distribution of a single component, i.e. the probability that it will have a value in the range x_i, $x_i + dx_i$. This is given by the Boltzmann distribution, which is

$$P_i\, dx_i = \frac{e^{-G_i/\theta}}{Z_i}\, dx_i = \frac{e^{-\alpha G_i}}{Z_i}\, dx_i \qquad (\alpha = \tfrac{1}{\theta})$$

(Z_i is the Gibbs phase integral for one component; see Appendix 5 for its derivation).

Rewriting this in terms of a new variable $\xi_i = \frac{X_i}{q_i}$, and using $\tau_i = \beta_i q_i$, this becomes,

$$P_i(\xi_i)\, d\xi_i = \frac{\xi_i^{\alpha\tau_i - 1}\, e^{-\alpha\tau_i \xi_i}}{(\alpha\tau_i)^{-\alpha\tau_i}\, \Gamma(\alpha\tau_i)}\, d\xi_i \quad \ldots\ldots (20)$$

The moments of order n of ξ_i are,

$$\overline{\xi_i^n} = \frac{\Gamma(\alpha\tau_i + n)}{\Gamma(\alpha\tau_i)} (\alpha\tau_i)^{-n},$$

and in particular $\overline{\xi_i} = 1$.

As $\theta \to 0$ ($\alpha \to \infty$), all the moments approach unity, as is required for the completely 'quiet' state with no fluctuations of

components.

When $\alpha\tau_i > 1$, the most probable ξ_i, denoted by $[\xi_i]$, is,

$$[\xi_i] = 1 - \frac{1}{\alpha\tau_i} = 1 - \frac{\theta}{\tau_i},$$

which is always less than $\bar{\xi}_i = 1$. When $\alpha\tau_i \leq 1$,

$$[\xi_i] = 0;$$

in fact, $P(\xi_i = 0) = \infty$ if $\alpha\tau_i < 1$.

This means that the most probable values of the variables x_i ($= \log \xi_i$), will be negative ones. The reason for this is that surfaces of constant G have larger lobes in the region $\{x_i < 0\}$ than in the region $\{x_i > 0\}$, and the ergodic hypothesis implies that the system will therefore spend more of its time in the first region than in the second.

If now the components are ordered in sequence such that,

$$\tau_1 \leq \tau_2 \leq \tau_3 \leq \cdots \leq \tau_n$$

then there will be two categories according as $\tau_i \leq \theta$ or $\tau_i > \theta$. In the first, the most probable observed value of the concentration will be a very small one; in the second, concentrations will be appreciably larger, and never very far from the stationary values. These two categories are of fundamental interest in this study, providing a theoretical basis for an understanding of the discontinuities which arise in embryonic development. A descriptive interpretation of the above result is that any system with competitive interactions defined in the manner we have given will show a division of its components

into major and minor groups, these being defined relative to the value of the parameter θ . The major group, with $\tau_i > \theta$, competes successfully for precursors and maintains itself near its steady state concentration, with relatively small oscillations. The other group is unsuccessful in its competition for precursors, and only occasionally manages to show bursts of synthetic activity which bring significant quantities of the component into the system. This latter group does not actually vanish from the system, but its members are unable to maintain themselves constantly at significant concentrations. We shall see later that they tend to be practically pushed out of existence by the dominant components when the system undergoes spontaneous irreversible change.

From equation (16), the relation $\alpha\tau_i > 1$ can be written,

$$\frac{\overline{(Xi - qi)^2}}{qi^2} < 1, \text{ and } \alpha\tau_\kappa \leq 1 \text{ becomes}$$

$$\frac{\overline{(X_\kappa - q_\kappa)^2}}{q_\kappa^2} \geq 1, \text{ for the system in equilibrium.}$$

We see therefore that the size of the oscillation in the two groups is different, the major components oscillating less than the minor ones.

It is the non-linearity of the initial differential equations which gives to the Volterra system this discontinuity in the behaviour of the components. The oscillations are non-linear ones, having the general characteristics shown in Figures 2 and 3, for major and minor components respectively.

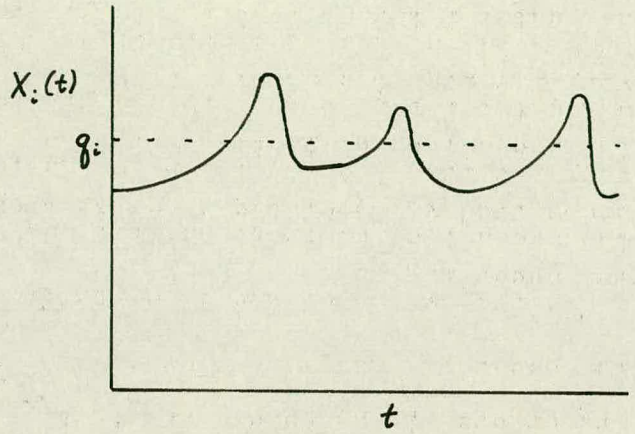


Figure 2.

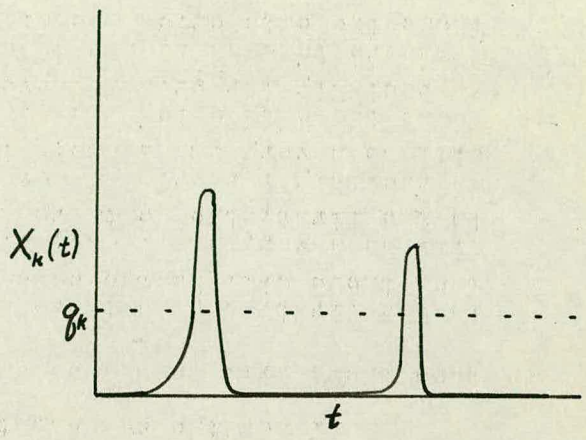


Figure 3.

The oscillations are not regular ones, but occur with irregular frequency and have the general form of a sort of overblown random noise. The major components oscillate in shallow troughs and cross their stationary values qi much more frequently than the minor components, which oscillate in sharp infrequent peaks. When θ is very small, nearly all the components will be in the major category ($\tau_i > \theta$), so that they will oscillate very nearly sinusoidally about their values qi. This was the case which Volterra considered in a linearised approximation to his equations. Under these conditions, with θ and the x_i's small, the Boltzmann distribution is itself normal:

$$P_i \, dx_i = \frac{e^{-\alpha\tau_i(e^{x_i} - x_i)}}{(\alpha\tau_i)^{-\alpha\tau_i}\Gamma(\alpha\tau_i)} \qquad (\alpha = \frac{1}{\theta})$$

$$\approx \frac{(\alpha\tau_i)^{\alpha\tau_i}}{\Gamma(\alpha\tau_i)} e^{-\alpha\tau_i} e^{-\frac{1}{2}\alpha\tau_i x_i^2}$$

$$\approx \left(\frac{\alpha\tau_i}{2\pi}\right)^{1/2} e^{-\frac{1}{2}\alpha\tau_i x_i^2}$$

Thus we see that it is only when θ has relatively large values that the peculiarities of the non-linear conditions emerge. Interpretating embryological phenomena in terms of the theory we may say that in early stages of development an embryo is in an excited condition (large θ) and its responses to environmental stimuli indicate non-linear behaviour (discontinuities); while as tissues become determined and differentiation proceeds, the system becomes less excited and much more stable over a large class of disturbances which would have caused 'switching' in earlier, more excited, states. It is thus possible to give interpretations of concepts such as competence, determination, and buffering in terms of the theory. But this must be done through a study of the analogues of the conventiona thermodynamic functions, to which we now turn.

SECTION VIII

THE EPIGENETIC FUNCTIONS

The epigenetic free energy, ψ , is defined by,

$$\psi = -\frac{\log Z}{\alpha} = \sum_{i=1}^{n}\left\{\tau_i \log \alpha\tau_i - \frac{\log \Gamma(\alpha\tau_i)}{\alpha}\right\} \quad \ldots\ldots (18)$$

$$\equiv \sum_{i=1}^{n} \psi_i \qquad \left(\alpha = \frac{1}{\theta}\right)$$

The epigenetic internal energy is,

$$\bar{G} = -\frac{\partial \log Z}{\partial \alpha} = \sum_{i=1}^{n} \tau_i\left[\log \alpha\tau_i + 1 - \varphi(\alpha\tau_i)\right] \ldots\ldots (19)$$

$$\equiv \sum_i G_i$$

where $\varphi(x) = \dfrac{\partial \log \Gamma(x)}{\partial x} = \dfrac{\Gamma'(x)}{\Gamma(x)}$, the digamma function.

The epigenetic entropy is,

$$\bar{S} \equiv -\overline{\log \rho} = \overline{\frac{G-\psi}{\theta}}$$

$$= \log Z - \alpha \frac{\partial \log Z}{\partial \alpha}$$

$$= \sum_{i=1}^{n}\left[\alpha\tau_i + \log \Gamma(\alpha\tau_i) - \alpha\tau_i \varphi(\alpha\tau_i)\right]$$

$$\equiv \sum_{i=1}^{n} \bar{S}_i \qquad \ldots\ldots\ldots (20)$$

The entropy has the important property that, for a system with given G, the function defined by $-\overline{\log \rho}$ takes a maximum for the canonical ensemble where $\rho = e^{\frac{\psi - G}{\theta}}$ than for other ensembles

(other distributions). The usual interpretation of this is that non-equilibrium states tend to decline into equilibrium ones of maximal entropy. This gives us a maximum principle for the epigenetic system, which is open to energy and matter but closed to G($\bar{G}$ being fixed, and hence θ). Accordingly, the epigenetic system with $\bar{G}$ held constant will move until $\bar{S}$ = maximum; and under spontaneous change when $\bar{G}$ and θ are allowed to change, it will move so that $\delta S > 0$. We shall study such irreversible processes in more detail later.

An epigenetic system with $\bar{G}$ held constant is one which is neither growing nor differentiating, but is simply maintaining itself. The purpose of using functions like S and ψ is to see what the 'spontaneous' motion of the system would be if it were allowed to change its concentrations of protein and RNA.

The functions ψ , $\bar{G}$, and $\bar{S}$ vary with θ and with the τ_i's (hence with qi, since β_i is a constant). We shall study first their variation with θ .

$\bar{G}$ is a monotonic increasing function of θ , as is immediately apparent from the following

$$\bar{G} = -\frac{\partial \log Z}{\partial \alpha}$$

$$\frac{\partial \bar{G}}{\partial \theta} = -\alpha^2 \frac{\partial \bar{G}}{\partial \alpha} = \alpha^2 \frac{\partial^2 \log Z}{\partial \alpha^2}$$

But

$$\frac{\partial^2 \log Z}{\partial \alpha^2} = \left(\frac{\partial \log Z}{\partial \alpha}\right)^2 - \frac{1}{Z}\frac{\partial^2 Z}{\partial \alpha^2}$$

$$= (\bar{G})^2 - \overline{(G^2)}$$

$$= \overline{(\bar{G}-G)^2} \geq 0$$

Therefore $$\frac{\partial \bar{G}}{\partial \theta} = \alpha^2 \overline{(\bar{G}-G)^2} \geq 0$$

From (19) we get expressions for $\bar{G}i$ in the limits of small and large θ. When θ is small (α large), we use the asymptotic expansion of $\varphi(x)$ (Copson, 1935):

$$\varphi(x+1) \sim \log x + \frac{1}{2x} + O\left(\frac{1}{x^2}\right)$$

Hence

$$\begin{aligned} \bar{G}i &= \tau_i \left[\log \alpha\tau_i + 1 - \varphi(\alpha\tau_i) \right. \\ &\sim \tau_i \left[\log \frac{\alpha\tau_i}{\alpha\tau_i - 1} + 1 - \frac{1}{2(\alpha\tau_i - 1)} \right. \\ &\sim \tau_i \left[\frac{1}{\alpha\tau_i} - \frac{1}{2\alpha\tau_i} + 1 \right] \\ &\sim \frac{1}{2}\theta + \tau_i \qquad \ldots\ldots\ldots (21) \end{aligned}$$

Therefore in this limit $\bar{G}i - \tau i$ approaches $\frac{1}{2}\theta$, and we see that,

$\sum_i [G_i - (G_i)_{min}] = \sum_i [G_i - \tau_i]$ is equipartitioned at $\frac{1}{2}\theta$ per degree of freedom in the limit of small θ. When θ is very large (α very small), we find for the limit of $\bar{G}i$,

$$\begin{aligned} \tau_i\left[\log \alpha\tau_i + 1 - \varphi(\alpha\tau_i)\right] &= \tau_i \left[\log\alpha\tau_i + 1 + \gamma + \frac{1}{\alpha\tau_i} - \sum_{n=1}^{\infty} \left(\frac{1}{n} - \frac{1}{n+\alpha\tau_i}\right)\right. \\ &= \tau_i \left[\log \frac{\tau_i}{\theta} + 1 + \gamma + \frac{\theta}{\tau_i} - \sum_{n=1}^{\infty} \left(\frac{1}{n} - \frac{1}{n+\alpha\tau_i}\right)\right. \\ &\sim \tau_i \left[\frac{\theta}{\tau_i}\left(1 - \frac{\tau_i}{\theta} \log \frac{\theta}{\tau_i}\right)\right] \\ &\sim \theta \qquad \ldots\ldots\ldots (22) \end{aligned}$$

In this limit G is equipartitional among all species at θ

per degree of freedom.

The curve for $\bar{G}i$ as a function of θ is given in Figure 4,

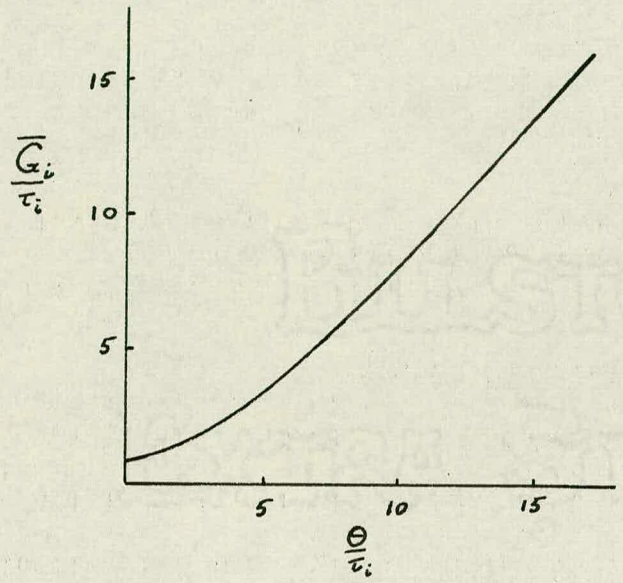


Figure 4.

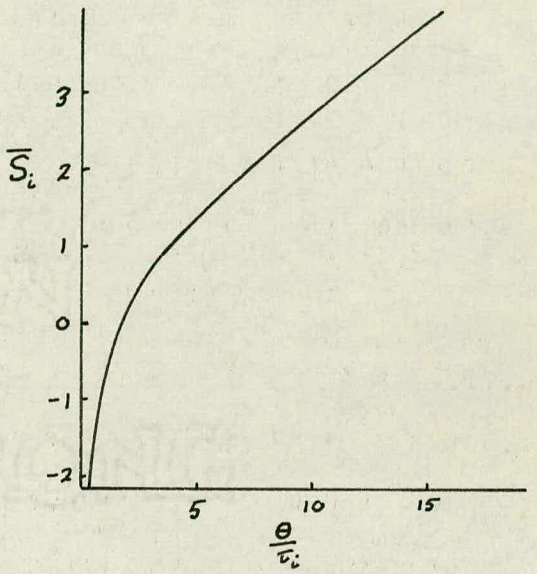


Figure 5.

Figure 5 shows how $\bar{S}_i$ varies with θ. The entropy increases without limit as $\theta \to \infty$ and decreases without limit as $\theta \to 0$. This behaviour is shown by the following limits.

$$\bar{S}_i = \alpha\tau_i + \log \Gamma(\alpha\tau_i) - \alpha\tau_i \varphi(\alpha\tau_i)$$

$$\alpha\tau_i \varphi(\alpha\tau_i) = \alpha\tau_i \left[\frac{1}{\alpha\tau_i} + \gamma - \sum_{n=1}^{\infty} \left(\frac{1}{n} - \frac{1}{n+\alpha\tau_i} \right) \right] \to 1$$

as $\alpha \to 0$

$$\Gamma(\alpha\tau_i) \to \infty \quad \text{as} \quad \alpha \to 0$$

hence $\bar{S}_i \to \infty$, as $\alpha \to 0$

When α is very large, we have the asymptotic expansions,

$$\log \Gamma(x) \sim \left(x - \tfrac{1}{2}\right) \log x - x + O\left(\tfrac{1}{x}\right)$$

$$x\varphi(x) \sim x \log x + \tfrac{1}{2} - O\left(\tfrac{1}{x}\right)$$

whence

$$\alpha\tau_i + \log \Gamma(\alpha\tau_i) - \alpha\tau_i \varphi(\alpha\tau_i) \sim -\frac{1}{2} \log \alpha\tau_i + 0\left(\frac{1}{\alpha\tau_i}\right)$$

$$\longrightarrow -\infty \text{ , as } \alpha \to \infty .$$

The free energy, ψ_i , does not vary monotonically with θ , but has the form shown in Figure 6,

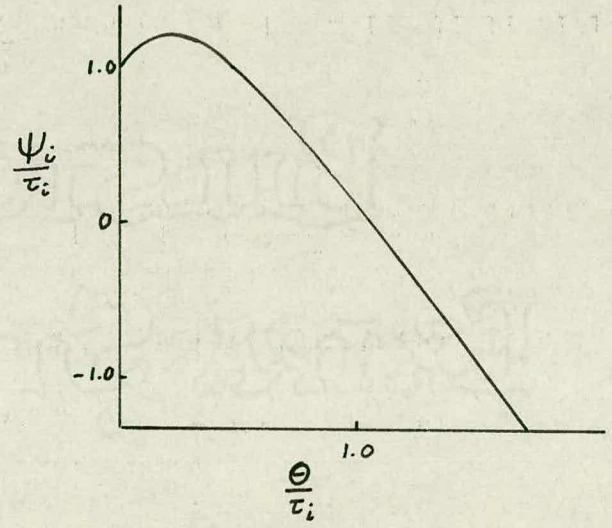


Figure 6.

$$\frac{\psi_i}{\tau_i} = \log \alpha\tau_i - \frac{\log \Gamma(\alpha\tau_i)}{\alpha}$$

For large α , this becomes, approximately,

$$\frac{\psi_i}{\tau_i} \sim \log \alpha\tau_i - \frac{1}{\alpha\tau_i}\left[\alpha\tau_i \log \alpha\tau_i - \frac{1}{2} \log \alpha\tau_i - \alpha\tau_i\right]$$

$$\sim 1$$

As $\alpha \to 0$, $\log \alpha\tau_i \to -\infty$

$$-\frac{\log \Gamma(\alpha\tau_i)}{\alpha} \longrightarrow -\infty$$

whence $\frac{\psi_i}{\tau_i} \longrightarrow -\infty$, as $\theta \to \infty$

The maximum occurs at the root of the expression,

$$\frac{\partial \psi_i}{\partial \theta} = -\alpha^2 \frac{\partial \psi_i}{\partial \alpha} = -\left[\alpha\tau_i + \Gamma(\alpha\tau_i) - \alpha\tau_i\,\varphi(\alpha\tau_i)\right] = 0$$

From (20), we see that this is simply,

$$\frac{\partial \psi_i}{\partial \theta} = -\bar{S}_i \quad , \text{ and } \bar{S}_i \text{ has a single zero, as}$$

shown in Figure 5.

The unbounded negative values of the epigenetic entropy, $\bar{S}_i$, as $\theta \to 0$ are quite unlike the lower limit which is assumed for the entropy in physical systems: $S \to 0$ as $T \to 0$ (Planck's formulation of the Nernst postulate). This limit depends upon the limits $C_p \to 0$, $C_v \to 0$ for the heat capacities as $T \to 0$. In the present theory, we have seen that

$$\bar{C}_i \equiv \frac{\partial \bar{a}_i}{\partial \theta} \to \frac{1}{2}, \quad \text{as} \quad \theta \to 0$$

Hence there is no analogue of the Nernst heat theorem, and the epigenetic entropy is unbounded in the lower limits of θ. For the upper limit of θ, $\frac{\partial \bar{a}_i}{\partial \theta} = 1$, which is the analogue of the law of Dulong and Petit for the limiting value of specific heats for high temperatures. However, the epigenetic heat capacity does not play the central role in the present theory that it does in the physical theory, because of the difficulty of controlling θ experimentally. This will be discussed more fully when we consider a possible equation of state for the epigenetic system.

We want now to see how the epigenetic entropy and free energy change when θ is kept constant, but the τ_i's are allowed to vary. From (20) it is clear that τ_i will affect $\bar{S}_i$ in exactly the same way as α and so we need only invert the graph in Figure 5 to get

the relationship between $\bar{S}_i$ and τ_i (Figure 7). $\bar{S}_i \to \infty$ when $\tau_i \to 0$, and $\bar{S}_i \to -\infty$ as $\tau_i \to \infty$.

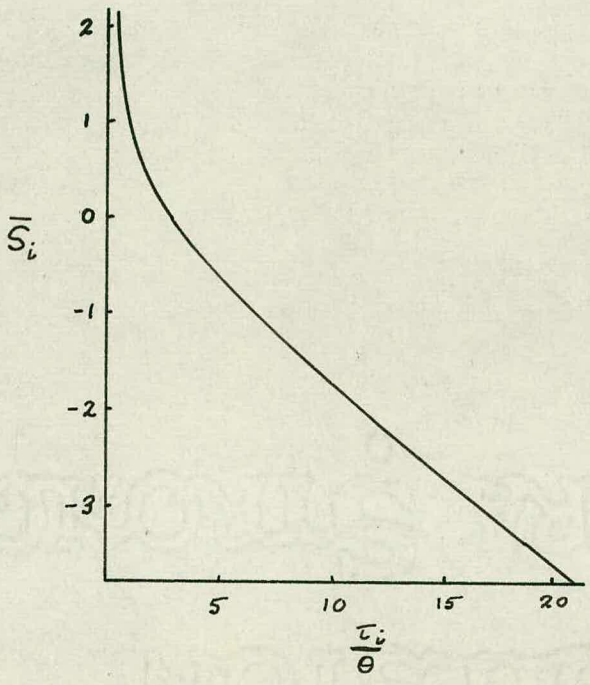


Figure 7.

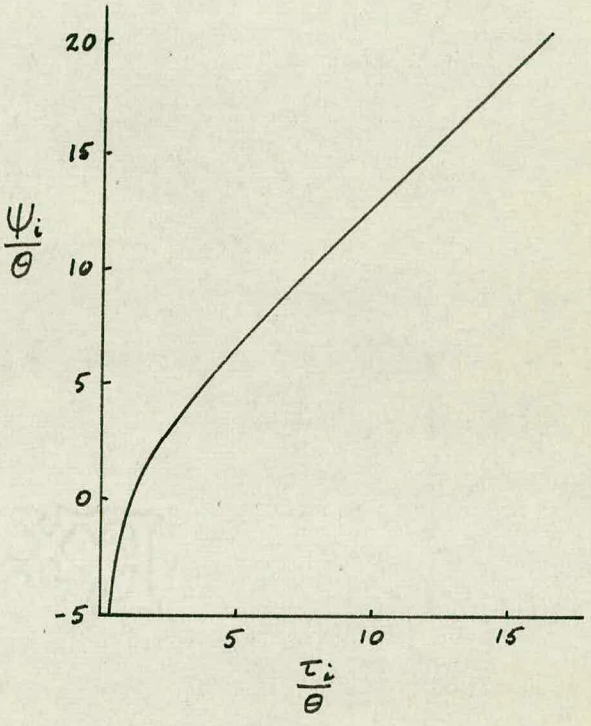


Figure 8.

Figure 8 shows how ψ_i varies from $-\infty$ to ∞ as τ_i goes from 0 to ∞.

Since $\tau_i \log \alpha\tau_i \to 0$

and $-\log \Gamma(\alpha\tau_i) \to -\infty$ as $\tau_i \to 0$,

therefore $\psi_i \to -\infty$ as $\tau_i \to 0$.

As $\tau_i \to \infty$, we have

$$\frac{-\log \Gamma(\alpha\tau_i)}{\alpha} \sim -\left[\tau_i \log \alpha\tau_i - \tau_i - \frac{1}{2\alpha} \log \alpha\tau_i\right]$$

and so

$$\psi_i = \tau_i \log \alpha\tau_i - \frac{\log \Gamma(\alpha\tau_i)}{\alpha}$$

$$\sim \tau_i \left(1 - \frac{1}{2\alpha\tau_i} \log \alpha\tau_i \right)$$

$$\sim \tau_i$$

These graphs show that an increase in τ_i causes a decrease in $\bar{S}_i$ and an increase in ψ_i. Now the conditions of spontaneous change in a 'thermodynamic' system are,

$$\delta\bar{S} > 0, \qquad \delta\psi < 0$$

Therefore an increase in τ_i can only occur spontaneously if the entropy decrease and free energy increase which this brings about in $\bar{S}_i$ and ψ_i respectively are more than compensated for by decrease in the τ_κ's of some other components. From Figures 7 and 8 we see that a small change $\Delta\tau_i$ in a component with small τ_i brings about much larger increments $\Delta\bar{S}_i$ and $\Delta\psi_i$ than a corresponding incremental change $\Delta\tau_\kappa$ for a large component (large τ_κ). This means that much more 'work' must be done in increasing the concentration of a small component by a given amount than in bringing about the same increase in the concentration of a large one (see Section X).

The concept of work enters a statistical mechanics in association with the external parameters. The τ_i are functions of these parameters, and it is through them that forces external to the system cause changes in $\bar{G}$, and hence changes in the state of the epigenetic system. In the embryological case these parameters are the genes, inducers, substrates, physical temperature, etc. The generalised force conjugate to the parameter a_r is defined as,

$$F_r = -\frac{\partial G}{\partial a_r} = -\sum_{i=1}^{n} \frac{\partial \tau_i}{\partial a_r} \left(e^{x_i} - x_i \right)$$

with the canonical mean given by,

$$-\overline{\frac{\partial G}{\partial a_n}} = -\int \frac{\partial G}{\partial a_n} e^{-\alpha G} dx \Big/ \int e^{-\alpha G} dx$$

$$= \frac{1}{\alpha} \int \frac{\partial (e^{-\alpha G})}{\partial a_n} dx \Big/ \int e^{-\alpha G} dx$$

$$= \frac{1}{\alpha Z} \frac{\partial Z}{\partial a_n}$$

$$= \frac{1}{\alpha} \frac{\partial \log Z}{\partial a_n}$$

$$= -\frac{\partial \Psi}{\partial a_n} = -\sum_{i=1}^{n} \frac{\partial \psi_i}{\partial a_n}$$

$$= -\sum_{i=1}^{n} \frac{\partial \tau_i}{\partial a_n} \left[\log \alpha \tau_i + 1 - \varphi(\alpha \tau_i) \right]$$

Hence

$$\overline{F}_n \equiv \sum_{i=1}^{n} \overline{F}_{in} \quad , \text{ where}$$

$$\overline{F}_{in} = -\frac{\partial \tau_i}{\partial a_n} \left[\log \alpha \tau_i + 1 - \varphi(\alpha \tau_i) \right]$$

$$= -\frac{\overline{G}_i}{\tau_i} \frac{\partial \tau_i}{\partial a_n} = -\overline{G}_i \frac{\partial \log \tau_i}{\partial a_n} \qquad \ldots\ldots\ldots(23)$$

The epigenetic work done by the system when the parameter a_n changes by an amount δa_n is defined as

$$\delta W \equiv \overline{F}_n \delta a_n = -\left(\sum_{i=1}^{n} \overline{G}_i \frac{\partial \log \tau_i}{\partial a_n} \right) \delta a_n$$

More generally, when several external parameters change incrementally, the work done is

$$\delta W = \sum_{n=1}^{m} \bar{F}_n \delta a_n = -\sum_{n=1}^{m} \left(\sum_{i=1}^{n} \bar{G}_i \frac{\partial \log \bar{\tau}_i}{\partial a_n} \right) \delta a_n$$

This quantity may be positive or negative, depending on whether epigenetic work is done by or on the system, respectively. We shall now see if these definitions can be made meaningful in an embryological context.

SECTION IX

AN EPIGENETIC EQUATION OF STATE

If there is any equation of state for the epigenetic system analogous to,

$$pV = nkT$$

for physical systems, then it will be found in terms of some general environmental parameter of the epigenetic system which plays a universal role in determining the state of the system. Such a parameter may be the physical temperature, T. A considerable amount of experimental study has gone into an investigation of the responses of single cells and embryos to changes in temperature. One general conclusion from such studies is that as the temperature is increased over a certain range, size is decreased. For example, in Chilomonas paramecium, Mucibabic (1956) has reported that from 5° to 26°C the size of cells decreases from giants which do not divide, to a minimum cell size which remains constant over the temperature range 26° to 30°C. Above 30°C size rapidly increases again. In poikilothermic embryos, the same general relationship holds: overall size of the embryo decreases with increasing temperature, over certain ranges, while the developmental rate

increases (Hartmann, 1918). It is not known however, whether their reduced size is due to smaller cells or to reduced numbers of cells. On the theoretical side, it has been shown (e.g. Kacser, 1957) that increased temperature reduces the steady state values of the reactants in open metabolic systems, at the same time increasing the rate at which they approach the steady state. This result depends upon the fact that diffusion constants have, in general, much lower temperature coefficients than reaction constants. A further assumption is that the concentrations of the catalysts remain constant. Now insofar as the steady state concentrations of the epigenetic components (proteins) are subject to these same theoretical considerations, the same conclusion holds: their concentrations will decrease with increased temperature. However, the epigenetic components include the catalysts of the metabolic system, and a decrease in these concentrations will decrease the reaction constants. This may reverse the temperature effect on the steady state of the metabolites; it will, in any event, work in the opposite direction, so that the overall metabolic response to temperature change may be quite small: as an open biochemical system, a cell tends to be buffered against temperature fluctuations.

Looking now at the epigenetic system from the point of view of the Volterra equations, we may consider what effect a temperature increase may have on the different parameters. The β_i 's will remain unaffected, but the α_{io}'s and the α_{ij} 's will change, being rate constants themselves, or functions of rate constants. However, it is not at all evident just how the kinetic rate constants enter these parameters, which are composite functions including both diffusion and reaction terms. The α_{io} 's will have a simpler form than the α_{ij} 's, the latter representing

the rate constants for a complex process of degradation, synthesis, and feed-back by diffusion, all compressed into a single expression; while the former are the autosynthetic coefficients, involving diffusion and reaction constants. Without further information about these coefficients, all that can be shown is that if the α_{ij}'s are increased in absolute value relative to the α_{io}'s as the result of a temperature increase, then the steady state concentrations of the epigenetic components will decrease (see below). In general terms, if the binary interactions in a Volterra system are increased in absolute value over the self-interaction terms, then the size of the system decreases (i.e. all steady state concentrations decrease).

Taking the most simplified case, this can be shown in the following manner. The steady state concentrations, qi, are determined by the solution of the matrix equation

$$Aq = p$$ (see page 52)

Suppose now all the coefficients α_{ij} are increased in absolute value relative to the α_{io}'s by an amount $r > 1$, so that the new equations take the form

$$r\sum_j \alpha_{ij} q'_j = p_i \qquad (i = 1, 2, \ldots n)$$

or, in the matrix form,

$$rAq' = p$$

The assumption that r is the same for all components eliminates any selective effect which a temperature change might have on the system.

The solution of the above equation is

$$q' = \frac{1}{r} A^{-1} p$$

whereas the solution of the original equations is

$$q = A^{-1} p$$

Since $r > 1$, it is clear that the steady state solutions q_i' are all smaller than the original ones, and by the same amount. This shows that any 'force' which increases the interaction coefficients α_{ij} relative to the linear coefficients α_{io} will cause a decrease in the steady state concentrations of the epigenetic system; and conversely, the values q_i increase when the α_{ij}'s are decreased relative to the α_{io}'s.

Suppose, then, that over some range of temperature

$$q_i(T) = c_i T^{-b_i}$$

From equation (23) we obtain the relationship

$$\bar{F}_{iT} = -\bar{G}_i \frac{\partial \log \beta_i q_i}{\partial T} = \bar{G}_i \frac{b_i}{T}$$

When θ is large, which is the condition under which an equation of state for the epigenetic system is to be sought, $\bar{G}_i$ is equipartitioned at θ per component, and the above relationship becomes

$$\bar{F}_{iT} = \theta \frac{b_i}{T}$$

Taking now the sum over all components, we have

$$\bar{F}_T = \sum_{i=1}^{n} \bar{F}_{iT} = \frac{\theta}{T} \sum_{i=1}^{n} b_i$$

or $T \bar{F}_T = n b \theta$, with $b = \frac{\sum b_i}{n}$

This could be regarded as an epigenetic analogue of the familiar relation (24), with T in place of p, and $\bar{F}_T$, the variable conjugate

to T, instead of V. As T increases, $\bar{F}_T$ decreases, when θ is held constant. Exactly what $\bar{F}_T$ is in the epigenetic system is not at all obvious although it is closely related to size. It is not a single datum of direct observation, as is V. Indeed, this derivation of a possible equation of state serves to emphasize what is already apparent from experimental biology: there are exceedingly few general quantitative relationships which have ever been shown to hold between macroscopic observables in biological systems. Even θ is not an obvious experimental observable, although it may become a useful parameter in embryological studies to describe the metabolic excitation of the system and the general character of the fluctuations of its components. However, this type of behaviour is not yet the concern of experimental embryology, and may require a further refinement of observational technique. Other relationships between general parameters may then begin to appear as well. But at the moment it does not seem fruitful to follow further the analytical pattern of classical thermodynamics, and pursue analogies such as the Carnot cycle in a hypothetical T, θ -plane. The difficulty in such an analysis is not the theory, but rather the experimental frame of reference in which to verify any derived relationships. θ is not under direct experimental control, as is T in the physical context; and $\bar{F}_T$ remains unidentified in biological terms, so either the relationship derived is meaningless, or its verification awaits a closer experimental analysis of the biological material.

SECTION X

THE EPIGENETIC FORCES

The epigenetic system of an embryonic cell is under the influence of forces arising from two quite different sources, one from the genes, and the other from the external milieu of the cell. The latter set of forces includes the organisers, inducers, and evocators of experimental embryology, to which we shall refer collectively as inductive stimuli. In this section we will consider the possibility that these two major categories of epigenetic forces, genes and inductive stimuli, may enter the Volterra equations in quite different ways, producing different effects upon the epigenetic state. It will be suggested that qualitative differences first arise in embryonic cells in response to quantitative differences in inductive stimuli because of the existence of threshold values in the amount of epigenetic work required to bring about certain changes of epigenetic state. Secondary changes are then assumed to occur in gene activities which carry the epigenetic system through another period of transient motion to a more fully differentiated condition. This argument leads to an interpretation of the concepts of competence and determination in terms of the present theory.

Due to the interactions between the variables of a system

obeying Volterra dynamics, any disturbance entering the system through a parameter will affect the instantaneous concentrations of all components, increasing some and decreasing others. We observe first that the steady state is stable towards small periodic fluctuations in the parameters α_{io}, α_{ij}, (the β_i's do not fluctuate), as was shown by Volterra (1931). Such parametric fluctuations are simply superimposed upon the intrinsic oscillations of the system. Moreover, a random disturbance entering the system through one or a few parameters spreads rapidly over the whole system, its effect on any single component being very small. Thus the system is 'buffered' against fluctuations in environmental factors.

The next point, also observed by Volterra (1937), is that the value of q_r for a particular component r is much more sensitive to changes in the α_{rj}'s than to changes in α_{ro}. A stimulus entering the system through α_{ro} tends to be distributed over all components in an unspecific manner, some mean concentrations increasing and some decreasing. But changes in the binary coefficients α_{rj} have a specific effect upon the relative mean concentrations of pairs of components r and j, and it will be suggested that the genes exercise a degree of selective control over the epigenetic state through these parameters.

We wish now to consider a possible parametric representation of the response of the epigenetic system to an inductive stimulus. In general, induction causes the preferential synthesis of a particular protein or set of proteins within a cell receiving the stimulus, with secondary changes occurring subsequently in other proteins.

Perhaps the clearest examples of this type of response are to be found in the induction of enzymes by their substrates. This phenomenon has been known for some time in microorganisms (e.g. Dubos, 1940), but only recently has it been shown to occur in embryonic cells (Stearns and Kostellow, 1958). The latter authors induced the synthesis of tryptophane peroxidase in dissociated cells of frog embryos by exposure to tryptophane. Enzyme activity in early gastrula cells rose from a low level to a higher stationary value in the course of twelve hours incubation with substrate.

A set of parametric changes in the Volterra equations producing an increase in the steady state value of a particular component, say r , while affecting the other components only slightly, is a decrease in the absolute values of the coefficients α_{rj} (r fixed, j = 1, 2.... n) (see Appendix 4). In effect, this decrease in the binary coefficients reduces the intensity of the interactions between the component r and the others, with the result that the steady state value, q_r , rises. The presence of substrate protects an enzyme from inactivation, at the same time encouraging its synthesis, and this may very plausibly be represented in the equations as a decrease in the absolute values of the interaction parameters α_{rj}, the effect of which is a release of component r from strong interaction with the others. Embryonic induction may involve quite different biochemical mechanisms from those of enzyme induction; but it seems not unreasonable to suppose that inductive stimuli in general act by releasing a particular component or set of components from a condition of repression resulting from the competitive interactions occurring in a cell, in consequence of

which the steady state concentrations of the released components increase. The actual shape of the curve of X_r during the induction depends very largely upon how quickly the parameters change in relation to the characteristic time for the system. What we can say, however, is that it will not have a simple shape, although long-term sampling, say every hour, as in the experiments of Stearns and Kostellow, will not reveal its complexity and the pattern of its fluctuations. The question of characteristic time and parametric changes is discussed in Section XI.

If we consider the decrease in the absolute values of the α_{rj}'s as a parametric change which is independent of gene action, brought about solely by the inducer, then the subsequent rise in q_r is a primary response of the epigenetic system, i.e. one which occurs before any change takes place in gene activities. If before the induction occurred, q_r was small, so that the inequality $\beta_r q_r < \theta$ was satisfied, then we know from the discussion in Section VII that the instantaneous value of the component, X_r, would be predominantly very low, near zero, and its oscillations would be highly irregular. Assuming that after the induction process q_r increases to a new value q_r' so that the inequality is reversed, $\beta_r q_r' > \theta$, then component r will show much greater stability in its oscillations, with X_r never far from the new mean value. Under the postulate that there is some kind of information feed-back to the genes, the continuing presence of a significant quantity of X_r in the system under the new conditions may bring about changes in gene activity which tend to stabilise the new high steady state value, q_r', so that even when

the inducing stimulus, s_r, is removed, the inequality is maintained, and component r does not fall back to its earlier low instantaneous concentrations. The component will then have become established as a major component of the epigenetic system: it has been 'selected' by the epigenetic forces. However, if there are no secondary changes in gene activity which stabilise the changed condition of component r, then a removal of the inducer, s_r, will reverse the earlier effect on the parameters α_{rj}, and the mean value will return to q_r. In this case, the system will have undergone a reversible change of state under a reversal of environmental conditions, a process which Weiss has called a 'modulation'.

The question now arises as to the nature of the postulated secondary gene changes in terms of the system parameters. If some of the α_{rj}'s were originally negative, say $\alpha_{rk} < 0$, then information feed-back from the induced high level of X_r might cause a change of sign in this parameter, so that, after a certain time, $\alpha_{rk} > 0$. This means that the dominance relation between components r and k is reversed, the result being that the synthesis of component r is favoured over that of some other component k. Such a reversal of sign in α_{rk} will lead to changes in the mean values of the other components, the result of the whole process being a fairly extensive change of epigenetic state. The details of these secondary changes in epigenetic state depend upon the interaction pattern of the whole system, and may be regarded as the theoretical counterpart of the complex and inter-related changes which have been observed to occur in genetic experiments on developing organisms, and go under the

name of gene interactions.

In the adjustment to the new parametric conditions, the steady states of certain components may be expected to decrease, and this may involve a reversal of the inequality relative to θ; i.e. certain components may move from the major to the minor category, so that their instantaneous concentrations change radically. For example, the reversal in the sign of α_{rk} could bring about a 'switch' in the mean concentrations of components r and k relative to θ so that r moves into the major category of components and k into the minor category. As the epigenetic state moves again under the influence of other stimuli which do not 'select' component k (i.e. cause it to increase sufficiently so that $\beta_k q_k > \theta$), then q_k will decrease more and more in the absence of significant feed-back from X_r, which continues to have predominantly very low values. (We saw in Section III that the components of the epigenetic system are subject to decay in the absence of some kind of information flow from the genes).

The irreversible loss of component k from the system may then be due to one or more of the following possibilities. The value of q_k becomes so small that the amount of epigenetic work required to bring the component into the major category, with $\beta_k q_k > \theta$, is so large (see graph of ψ_k for very small τ_k, Section VIII) that no normal embryonic stimulus will be sufficient to reverse the inequality, and the component is eventually lost from the system; the new steady state solution with the changed parameters gives a negative value to q_k, which is to be interpreted as a loss of the component from the system (see Volterra, 1931); or the genetic determinants which

might 'stabilise' component k at values $\beta_k q_k > \theta$ (by changing the signs of some α_{kj}'s) have been inactivated in some manner and so component k cannot remain permanently in the major category, and drops to very low values as soon as selective stimuli are removed, eventually disappearing.

The actual amount of work done upon the epigenetic system by an inducing stimulus s_r, acting over the time period t_0 to t_1, is given by

$$W = -\sum_i \int_{t_0}^{t_1} \bar{G}_i \frac{\partial \log \tau_i}{\partial s_r} \, \delta s_r \qquad s_r = s_r(t)$$

If during this time the steady state value of component i changes from q_{i0} to q_{i1} and θ is constant during the process, then this expression may be written as

$$W = -\sum_i \int_{q_{i0}}^{q_{i1}} \beta_i \left[\log \alpha \tau_i + 1 - \varphi(\alpha \tau_i) \right] dq_i$$

$$= -\sum_i \psi_i(q_{i1}) - \psi_i(q_{i0})$$

$$= \psi_0 - \psi_1$$

Thus we see that the epigenetic work done by s_r is equal to the increase in the epigenetic free energy, ψ, providing that this change takes place at constant θ. If we restrict our attention to q_r, and ignore the relatively small changes which occur in the other components (assuming that s_r acts upon the parameters α_{rj} in the manner described above), then we need calculate only the change in ψ_r to find the epigenetic work done by the inductive stimulus. The rest of the system is then regarded as an epigenetic 'heat bath'

('excitation bath') for component r, and we can assume θ = constant. But when the changes of state in the system are too large for this procedure to be valid, it is necessary to use a more elaborate mathematical technique for calculating the changes in the epigenetic functions during irreversible change.

From the graph of ψ_r on page 81 , it is clear that the amount of epigenetic work done by an inductive stimulus s_r which moves a component r from the minor to the major category is determined largely by the 'distance' of τ_r from θ: $(\theta - \tau_r)$ (considering only changes in component r, and with θ constant). The epigenetic free energy decreases very rapidly as τ_r gets small, and therefore when τ_r is much smaller than θ , the amount of epigenetic work required to produce the condition $\tau_r > \theta$ is very large. For $\tau_r > \theta$, the curve of ψ_r becomes nearly linear, so that the same amount of epigenetic work produces approximately the same change in τ_r .

Consider now the conditions which will determine whether or not a particular component, r, belonging initially to the minor category of components of the epigenetic system, may be moved permanently to the major category by the action of an inductive stimulus. A first point to notice is that the component must be present in finite quantity: $q_r > 0$. If $q_r = 0$, the amount of epigenetic work required to induce it is infinite, so that this component will be permanently absent from the system. This theoretical condition arises from the obligatory autosynthetic nature of the differential equations (7), and is therefore probably an exaggeration of the actual conditions of synthesis prevailing in cells. The epigenetic free energy of a

component, ψ_r, may not in fact plunge to $-\infty$ as $\tau_r \to 0$ because when $\tau_r = 0$, the information required for the synthesis of protein and RNA of type r is not necessarily lost from the epigenetic system; the genes may still be capable of supplying this information to the system, a possibility not recognised by the differential equations (7). However, the general shape of the curve of epigenetic free energy brings out an important feature of competitive and autosynthetic systems: the smaller the mean concentration of a component, the greater is the 'force' required to cause it to increase by any given amount. The theoretical system with which we are dealing adds the further requirement that a component must be present in the epigenetic system before it can be acted upon by an epigenetic force, a condition which may be interpreted as the necessary synthesis of any particular protein, hence the activity of the corresponding gene, even if very small, in order for the protein to be inducible in a cell. This condition is in fact believed by some to be satisfied by 'young' cells of both microorganisms (Dubnoff, 1955) and of embryos (Herrmann, 1959).

Secondly, the stimulus must be strong enough or last long enough to bring about a change of inequality from $\tau_r < \theta$ to $\tau_r > \theta$. Thus, for a given θ and a given τ_r, s_r must be sufficient in intensity and duration to cause the above reversal in the inequality. This process we call the induction of component r.

Thirdly, it must be possible for the postulated secondary action of the genes to occur before the inducing stimulus, s_r, is removed, so that the inequality $\tau_r > \theta$ is maintained even after the initial

parametric change is reversed (return of the α_{rj}'s to initial absolute values).

Assuming now that the inductive stimulus may be complex in the sense that it may act upon several components together, we can define the condition of competence in a cell relative to a particular stimulus and hence relative to the particular set of components affected, and under the postulates concerning induction and gene action introduced above. It is that state of the epigenetic system satisfying the first and the third conditions above: the components must be present in the cell, hence the corresponding genes must be active; and it must be possible for secondary changes to occur in gene activity following induction, so that the induced components remain in the major category and so exert a continuing influence upon the further development of the cell.

The cell may now be said to be 'determined' when the secondary parametric changes due to altered gene activities have occurred, following the induction, and the epigenetic system has responded to these changed environmental conditions by moving to a new state. The new conditions may have caused an extensive reorganisation of the epigenetic system, with a possible loss of some components (hence changed dimensions of phase space), and 'switching' between certain components of the major and minor categories. The only strictly irreversible change occurring (irreversible in the sense that no finite epigenetic force can cause a return to the original state, and not in the usual thermodynamic sense of the term) is the loss of a component from the system; but we shall see that continuously acting

'dissipative' forces in the system which tend to cause a reduction in θ bring about changes in the state of the epigenetic system which make it progressively more difficult to reverse the effects of 'induction' on 'competent' cells.

It is interesting to observe that there are four independently variable factors which are involved in deciding whether or not a set of components may be induced and determined in the manner described: the initial values of τ_i's and θ, the inducing stimulus, and the state of the genes. The above considerations make explicit the complex relations which must prevail between the conditions of the epigenetic system and its environment in order for induction and determination to take place. The inducing force must be commensurate with the epigenetic work required to move a particular set of components into the major category of the epigenetic system, and the genes must be capable of responding to the new epigenetic state in such a way that the components become stabilised as major components. If any of these factors fails to satisfy the necessary relationship with the others, the system will fail to respond in the normal manner.

The suggestions which have been made concerning the parametric changes which might correspond to the processes of induction and determination may well be wrong in detail. However, it is necessary that there be some specific alteration in the epigenetic state under the action of an inductive stimulus, and that this lead to changes of a more permanent nature in the values of the system parameters, which are not reversed upon removal of the stimulus. The existence of a statistical discontinuity in a system with strong competitive interaction between

components, such as we are considering, suggests that threshold values may play an important role in determining which proteins will be synthesised in increasing amounts by a differentiating cell, and which will be lost. Also, some kind of self-reinforcing mechanism must occur in conjunction with these induced changes of state, since, as Waddington (1948, 1954) has emphasised, the more a cell differentiates in a certain direction, the greater is its tendency to continue on the same path. In distinguishing between an initial, non-genic, response of the epigenetic system to an inductive stimulus, and a secondary parametric alteration due to feed-back from the new epigenetic state to the genes (the self-reinforcing process), we are continuing our analysis in the manner most consistent with the structure of the present theory, which is also, it is hoped, reasonably consistent with the actual function of the cell. The cytoplasm is the first part of the cell to receive an inductive stimulus, and so we have assumed that the primary response will occur in the epigenetic system. This response may be very specific, affecting primarily one or a few proteins initially, and not reflected immediately in other cellular activities: it is of the nature of an evocation. We have represented this as the release of certain components from strong competitive interaction with the others, the result being an increase in their steady state concentrations. In competent cells, the initial response leads to further changes of a more permanent kind, which may involve a considerable reorganisation of the cell's synthetic activities. These we have assumed to be gene-determined changes in the signs and magnitudes of the binary coefficients, α_{ij} , occurring as the result of feed-back

from the components which have been carried by the inductive stimulus through the statistical discontinuity and into the major category of epigenetic components, whose instantaneous concentrations are stable (i.e. small fluctuations about mean concentrations).

The initial response of the epigenetic system to the inducing stimulus, and its subsequent response to changes in gene activities are to some extent separable in time, the discontinuity between minor and major categories acting as a sort of boundary between these two processes. Only after this boundary has been crossed by induced components is feed-back from their stable instantaneous concentrations assumed to be significant. Then there may occur significant changes in gene activities, resulting in a switch in the signs and the magnitudes of the binary coefficients α_{ij}. The epigenetic system then responds to these gene-initiated changes. Certain of the major components may continue to increase in mean concentrations, and certain other components may be expected to decrease in steady state value, while those for which the new steady state solution gives negative values will be eliminated from the system. In order that a fairly close analysis of the responses of an embryonic cell to changes in its environment and in its gene activities be possible in terms of a theory such as we are considering, which describes the changes as sequences of transient motion between quasi-equilibrium states, some such time separation of the complex events taking place during induction and determination is necessary.

The experimental evidence bearing directly on this last point is very meagre, for the analysis which we are considering is a

rather fine one. However, from studies on induction and differentiation in the central nervous system of amphibians, Nieuwkoop (1955) has suggested a distinction between dependent and independent developmental processes, one following another in the same tissue during its differentiation. He has further distinguished between two types of dependent change, which he calls activation and transformation, the first having the character of an all-or-none response, and the second proceeding more gradually by a series of small irreversible steps. The time separation between the processes he makes quite explicit: "the processes of activation and transformation form successive steps in the induction process" (Nieuwkoop, 1955). The transforming influence appears to be in the nature of a field phenomenon, varying quantitatively along the cranio-caudal axis, and producing qualitatively different effects on presumptive neural cells. Between the reception of activating and transforming stimuli, embryonic tissue develops autonomously in a manner dependent upon the nature and the intensity of the stimuli. It is just such a step-wise description of developmental processes which can be most readily accommodated within the theoretical structure of the present statistical mechanics of cellular behaviour. The parametric representation of induction given above suggests that the primary response of the epigenetic system corresponds to a dependent phase of development, while the secondary changes due to information feed-back from epigenetic system to genes may be regarded as an independent developmental process.

The problem of how quantitative differences in an inducing stimulus can produce qualitatively distinct responses in cells has

been much discussed in embryology (e.g. Waddington and Yao, 1950). It has usually been assumed that a self-reinforcing process is operative in developing cells, whereby small initial differences in state will be progressively exaggerated as development proceeds. Such a hypothesis has been introduced in the present theory in ad hoc form as information feed-back from epigenetic system to genes, acting in such a way as to favour the increased synthesis of the major epigenetic components. However, it has often been suspected that some kind of threshold phenomena must occur in embryonic cells as well as a mechanism for the exaggeration of initial differences, in order to account for the 'switching' which occurs between metabolic states as the result of quantitative changes in an environmental parameter. Perhaps one of the clearest examples of this type of behaviour is to be found in the occurrence of different antigens on the surface of Paramecium under different environmental conditions (temperature, food; see Beale, 1954); and similar though less well-defined behaviour is shown by embryonic cells in their responses to quantitative differences in inductive stimuli. The statistical discontinuity of the present theory, arising from the strong (non-linear) competitive interaction between components, provides one theoretical demonstration of a very general type of threshold effect which could account for such switching phenomena. The intensity of an inductive stimulus can be defined in this theory in terms of the amount of epigenetic work which a stimulus can do on an epigenetic system, different amounts of 'work' being required to bring different components into the major category of the epigenetic system from the minor category. For each component

and for each set of initial conditions for the system, there is a particular 'threshold value' of the appropriate epigenetic force required to induce the component. When the component crosses the statistical discontinuity between minor and major categories, there is a switch in the condition of its instantaneous concentrations which will show up as a rather sudden appearance or disappearance of the component in question as an important constituent of the cell. Continuous information feed-back to the genes is assumed also to be switched on or off when the discontinuity in the epigenetic system is crossed by a component.

The question now arises whether it is possible to suggest an ageing process in terms of the present theory which could account for the gradual loss of competence in undetermined cells, at the same time suggesting how determined cells become progressively stabilised in one particular 'creode' (Waddington, 1957) as the result of their developmental histories. We shall show that slow changes in θ may reflect a gradually changing condition of protein synthesis and interaction in embryonic cells which could account for their 'canalisation' (Waddington, 1957) along distinct developmental pathways, making it increasingly difficult to 'switch' them from one creode to another.

Consider first the conditions in early embryonic cells. The submicroscopic morphology of such cells (e.g. early gastrula cells of amphibians) is much simpler than that of cells in later stages of development, after some differentiation has occurred (see e.g. Ranzi, 1958, Karasaki, 1959), and it has been suggested (Herrmann, 1959) that the RNA templates may not yet be incorporated into ribosomal

particles, but may be loosely distributed in the cytoplasm and at the cell surface. Herrmann (1959) speculates further that all possible proteins are being made in these early embryonic cells, although the predominant ones will probably be those concerned with growth and division. These two possibilities, the absence of a detailed submicroscopic morphology and the active synthesis of all possible proteins in young embryonic cells, suggest that the epigenetic temperature may be high in the early stages of development. This means that these cells are in an excited metabolic state, with high rates of turn-over and inter-conversion among the different species of protein and RNA, and relatively large fluctuations about the steady state values. There will probably also be strong competitive interaction between the components under these conditions, so that the interaction coefficients α_{ij} will be large. The major components of such cells may be expected to be the proteins involved in growth and division: the cells are differentiated for division, and those proteins involved in the division process constitute the heterosynthetic system; while the minor components will be all the other proteins which the cell can synthesise (defined by the genome). These latter components will perhaps be present in significant mean concentrations, with their τ's not much less than the value of θ for the cell, since we may assume that there has been as yet no selection 'against' them, and all the genes are assumed to be active. However, their instantaneous concentrations will all be predominantly very low, with occasional bursts of synthesis, as is characteristic of minor components.

With the τ's of the minor components at values not much

less than θ, the epigenetic system is in a reactive state, since relatively small stimuli will be sufficient to do the epigenetic work required for the induction of these components. Furthermore, in excited systems with large θ the τ's of the major components will probably not be very much greater than θ, so that these components may be fairly readily switched also by environmental influences. When such a system receives an inductive stimulus, it may be expected to undergo a fairly rapid change of state. Those components which are induced will become major components at large steady state values (because of the assumed large θ), and a high level of continuous information feed-back to the genes will be established. Secondary parametric changes due to changed gene activities will then result in further changes of epigenetic state which carry the cell along a particular developmental pathway. We must assume that the details of these secondary changes have been determined by the evolutionary history of the organism, so that they are integrated in such a manner as to produce developmental changes leading to a particular differentiated state of the cell (see e.g. Waddington, 1954). This problem is taken up again in the following section.

As development proceeds, several factors will cause the epigenetic temperature to decrease. One is the existence of forces causing slow 'frictional' damping in the system. This dissipative effect is produced by all factors which reduce the efficiency of the synthetic processes, such as the time lags involved in the diffusion and reaction of metabolites, and the decay of macromolecules. These damp the oscillations due to the inter-conversion of the different

species of protein and RNA, and hence reduce θ. Another factor is the increasing complexity of the submicroscopic morphology of differentiating cells, and the progressive organisation of the synthetic machinery into spatially separated units, the cell organelles (mitochondria, microsomes, membranes, vacuoles, etc.). This increasingly well-defined molecular ecology greatly reduces the interactions occurring between separated groups of macromolecular species, making their inter-conversions improbable, and reducing the absolute values of their interaction parameters. The cell tends thus to break down into subsystems, within which there may still be strong interactions, but between which the competitive forces are much reduced. The general effect of this increased morphology will be a reduction in θ and in the α_{ij}'s.

There may also be a change in the values of the β's, if larger functional groups of proteins are formed during the morphological changes. Such 'condensations' of components would reduce the dimensions of phase space, and somewhat reorganise the interaction forces. There would thus occur extensive changes in the original equations, and it seems very probable that at some point in development the binary interactions between independent components will cease to be important forces in determining the behaviour of the cell. When the molecular ecology of the cell is very well defined, then competitive interactions between the molecular 'niches' will be much reduced, and the intrinsic pattern of interaction determining the system's behaviour will involve other forces. In terms of equations (7), we may that as differentiation proceeds, the α_{ij}'s decrease and the β's may increase so that at

some point the binary terms may be ignored, while the self-interaction terms become increasingly important and can no longer be left out. There will then be a transition from a condition of strong binary interactions and weak damping, to weak interaction and significant damping, in which latter condition the equations take on a predominantly logistic form, with very weak coupling between components.

At this point θ will be very small, the statistical properties of the system will have become practically normal (see page 73), and the present statistical mechanics may not be applicable to the developmental process beyond this stage. We may say that the system at this point is fully determined, its developmental fate is fixed, and a statistical mechanics is no longer appropriate to a study of its motion.

Consider now how the stability of the epigenetic system towards general environmental stimuli changes as θ decreases during development. The 'distance' of a major component r, from the epigenetic temperature of the system, $(\tau_r - \theta)$, is a measure of its stability towards variations in environmental parameters. For if this distance is small, then a fluctuation in some parameter such as temperature, pH, glucose supply, etc., could cause τ_r to drop below θ temporarily, and this might set off secondary parametric changes through the genes which could lead to the loss of this component from the major category. Epigenetic states in which both major and minor categories have τ's near the epigenetic temperature are particularly sensitive to environmental changes and represent the 'epigenetic crises' of the present theory. In this condition, a temperature shock, for example,

could switch one or more components relative to θ, and alter the normal pathway of development. But as the distance of a major component from θ increases, the possibility of a temporary stimulus causing it to move to the minor category decreases. Those components of a cell which are selected by the epigenetic forces as members of the major category at some early point in development, when θ is quite large, will tend to become stabilised at correspondingly large steady state concentrations. As θ decreases during further development, their distances from the epigenetic temperature increase, and they become progressively more stable towards variations in environmental parameters. At the same time, the minor components tend to decay progressively in the absence of continuous information feed-back to the genes, and are gradually eliminated from the system. The two categories of epigenetic components tend to diverge from one another, making it increasingly difficult to cause any switching between major and minor components. The epigenetic system thus loses its reactivity to environmental stimuli during development, and becomes stabilised in some creode as the epigenetic temperature decreases, the creode being characterised by a particular set of components which have been selected by the epigenetic forces acting upon the system during its developmental history.

Within the theoretical confines of a statistical mechanics of intracellular processes, we cannot properly go further than this, into a discussion of the inter-cellular forces occurring during individuation and regionalisation (Waddington, 1956), which are, clearly, of very great importance in embryonic development. Indeed, we have already

gone outside the terms of our theory in introducing the ad hoc hypothesis of information feed-back from the epigenetic system to the genes. A proper analysis of this feed-back and its implications for the statistical mechanics is required for any further extension of the present developmental theory. This would carry us into irreversible processes proper, and would add much to a theoretical understanding of development. Without attempting to solve this problem, which can hardly be projected without a fuller experimental understanding of the nature of gene activities, we may nevertheless investigate the conditions which must be satisfied in embryonic cells in order that the whole theoretical approach developed in this study be valid for the analysis of embryological processes. This will form the content of the next section.

SECTION XI

RELAXATION TIME AND PARAMETRIC CHANGE.

One of the major difficulties encountered in the present theoretical treatment of intracellular embryonic events has been the distinction between dependent, coupled variables, and independent variables or parameters, whereby we distinguish a developmental system from its environment. The analysis of the properties of different cellular constituents given in Section II led to a definition of the epigenetic components, and in Section IV a set of differential equations was derived for the 'motion' of the epigenetic system in terms of a particular pattern of interactions between these components. The epigenetic system was thus given, in a certain measure, an independent existence, making it meaningful to speak of its state in terms of functions which describe the statistical condition of the system: G, ψ, θ, etc. These functions, the epigenetic variables of state, exist only for stationary or equilibrium states of the system, which are defined by particular sets of values of the β_i and g_i, and some mean value of G (which defines θ, and hence the mean size of the oscillations in the instantaneous concentrations of the epigenetic components). If some parameter changes to a new value, then a new stationary state is defined, and there will be a certain time lag before the statistical properties of the system settle down to new equilibrium values, thus giving new values to the epigenetic variables. If the parameters change continually, then it may happen that the epigenetic system cannot change its statistical properties rapidly enough for any equilibrium condition to be established in response to the parametric changes. In this case the epigenetic variables do not exist, and a statistical mechanics is of no use in describing the properties of the system. Only when the

parametric changes are such that they allow the system to exist in an equilibrium or quasi-equilibrium condition is it possible to analyse its behaviour in statistical mechanical terms. It is therefore of great interest for our present purposes to get some idea of the rate at which the epigenetic system can reach a condition of equilibrium after a small disturbance. This rate is determined by the relaxation time of the system, which quantity also allows us to estimate the period required for observation of the system in order to obtain a good time average for the epigenetic functions.

Clearly these are the fundamental time constants of the epigenetic system, which will give us the relations between microscopic and macroscopic events in embryonic cells, and the relative rates of different intracellular processes in development. Only the roughest of estimates can be made for these quantities, due to the difficulty of obtaining the relevant information from embryological material. However, we shall find that the relaxation time of the epigenetic system is certainly very much larger than that of physical systems, as might be expected in view of the complexity of the processes in the biological case, and the relative slowness of macromolecular syntheses as compared with the rapidity of particle interaction due to collisions in physical systems. There is a corresponding slowness of response in embryological systems, and a much longer observation period is required for the measurement of macroscopic quantities. These considerations will allow us to discuss the problem of irreversible change in the epigenetic system, the conditions which must be satisfied in order to treat the genes as controlling parameters of the system, and the tenability of the present theoretical treatment of developmental processes in the light of

available experimental evidence.

The relaxation time of a system is obtained from a study of the rate of change of the distribution function, ρ, after a small disturbance. Thus suppose that immediately after a disturbance the distribution function differs from the equilibrium function, ρ_o, by a small quantity, $\Delta\rho$:

$$\rho = \rho_o + \Delta\rho$$

The quantity we are interested in is $-\frac{d\Delta\rho}{dt}$, the rate at which the effect of the disturbance is annulled. Now for a small disturbance it is in general true that the rate of return toward equilibrium is proportional to the magnitude of the perturbation (see, e.g., ter Haar, 1954), and we can write

$$-\frac{d\Delta\rho}{dt} = k\Delta\rho$$

which gives us the relation

$$\Delta\rho = (\Delta\rho)_o e^{-kt}$$

The relaxation time is now defined as the time required for the disturbance to be reduced to $\frac{1}{e}$ of its original value. This quantity is, therefore, $t = \frac{1}{k}$, at which time $\Delta\rho = \frac{1}{e}(\Delta\rho)_o$. We are thus led to enquire into the nature of the rate constant, k, and the factors which determine its size.

The forces which cause ρ to return to an equilibrium value from non-equilibrium ones are just those forces which bring about an even distribution of G throughout all parts of the epigenetic system, producing the equilibrium relationships

$$\beta_i q_i \overline{\left(\frac{x_i}{q_i} - 1\right)^2} = \beta_j q_j \overline{\left(\frac{x_j}{q_j} - 1\right)^2} = \theta \qquad (\text{all } i,j)$$

This distribution of G is effected by the interactions of the components, whereby G is exchanged between different parts of the system. The rate at which G can be thus exchanged is therefore dependent upon the intensity of the binary interactions between the components. The 'motion' of the epigenetic system is defined in terms of the synthesis and degradation

of different protein and RNP species, and the competitive interactions between epigenetic components relate to protein and RNP metabolism. Local fluctuations in G are due to local variations in the concentrations of different protein species, and the rate at which these fluctuations die out will be determined by the rate of protein synthesis, the mean density or concentration of the proteins and ribonucleo proteins in the system, and the intensity of their interactions. The essential time factor in the 'motion' of the epigenetic system, the limiting factor which determines how rapidly all other epigenetic processes can take place, is the time required for the biosynthesis of protein from its precursors. This is a fundamental constant in any mechanistic study of cellular processes, as has been pointed out by Dalgliesh (1957), but it is also difficult to estimate. Several complications, discussed by Dalgliesh, arise in the interpretation of the results of tracer experiments which have been designed to give information about the rate of protein synthesis, and the following calculations for the relaxation time are subject to errors which may involve a factor of 10 or more.

McQuillen, Roberts, and Britten (1959), working with E. coli, have estimated that the time required for the synthesis of one molecule of protein from precursors is 5 seconds. Zalokar (1961) substantiates this estimate with work on Neurospora, wherein he finds a biosynthetic time of 'a few seconds or less' for protein. From the data of Halvorson and Cohen (1958) on yeast, it is possible to deduce that this biosynthetic time is less than 10 seconds. However

the detailed studies of Loftfield and Eigner (1958) on the synthesis of ferritin in rat liver indicate that this protein, which they isolated in crystalline form, requires about 6 minutes for complete synthesis from precursors; while the work of Dintzis, Borsook, and Vinograd (1958) on haemoglobin synthesis in rabbit reticulocytes gives a similar result. Zalokar (ibid.) suggested that the difference between these two groups of results may be due to a delay caused by some secondary assemblage mechanism in the formation of the complex proteins, ferritin and haemoglobin, and not to a difference in the time required for the template synthesis of the component polypeptides. If we accept this explanation for the discrepancy between these two sets of data, then we may take a value of 6 seconds as a reasonable approximation to the biosynthetic time for elementary protein.

We require next some estimate of the 'concentration' of RNP templates involved in the synthesis of any particular species of protein. Using again the data of McQuillen et al. (ibid.), and also a calculation reported by Guild (1956), it would seem that a reasonable estimate for the 'concentration' of active ribosomal templates producing a particular species of protein in a cell with high protein synthetic activity is of the order of $\frac{1}{100}$th. of the steady state concentration of the respective protein species. For any stationary state of the epigenetic system, then, we assume that the concentration of the ith. species of ribosomal template remains relatively constant at the value of $\frac{1}{100} q_i$, where q_i is the stationary value of the respective protein species, while the instantaneous concentration of this protein, X_i, oscillates. Any particular ribosomal system will then have the

capacity to synthesise $\frac{1}{100} q_i$ units of protein in 6 seconds, which is a rate of $\frac{1}{10} q_i$ units of protein per minute.

The frequency with which the instantaneous concentration of the various protein species can oscillate depends upon the mean amplitude of the oscillations. If it is assumed that this mean amplitude is in the region of $\frac{1}{100} q_i$, then during the rising part of an oscillation $\frac{1}{100} q_i$ units of protein must be synthesised. The ribosomal system has the capacity to synthesise this amount of protein in 6 seconds, and so we take this time as the duration of the rising section of a single oscillation. Assuming that this represents roughly one-half of the whole oscillation (although for non-linear oscillations, such as those shown in Figures 2 and 3, the rising section is greater than one-half of the period of a single cycle), we obtain the result that the period of a complete oscillation is about 12 seconds. This is a frequency of 5 oscillations per minute. The relaxation time can be taken to be roughly the same order of magnitude as the period of one oscillation, 12 seconds, while the time required to obtain a good time average is about one hundred times the relaxation time, which is 20 minutes. Finally, in order that the epigenetic functions may be used to describe the state of the system as it is propelled through a sequence of quasi-stationary states by continual parametric change, it is necessary that the parameters change many times more slowly than the observation time of 20 minutes; i.e., there must be a period of a few hours between significant parametric changes.

These values serve only to give a very rough idea of the relative time scales to be expected of cellular events on the basis of the above

estimates and assumptions. Compared with physical systems, where the relaxation time is of the order of 10^{-7} seconds (helium gas under normal conditions), 'epigenetic' processes are very slow, the ratio of the two being approximately 10^{8}. There are of course certain classes of biological phenomena which are basically physico-chemical in nature, not involving the synthesis of macromolecules and so having a much smaller relaxation time than that calculated above. Examples of these are conduction in nerve fibres, and the responses of the 'metabolic system' to environmental disturbances (see, e.g., Bradley and Calvin, 1956). These processes have short time constants, more comparable with physical ones. However, it is often felt by biologists that such phenomena are not the most characteristic features of cellular systems, and that underlying them there is another level of organisation which determines in a more fundamental manner the behaviour of developing cells. Our analysis of the macromolecular organisation of cells has led us to define this more complex organisational level in terms of the epigenetic system, and it is not surprising to find that it has a relaxation time of quite a different order from those of classical molecular systems. The fundamental process which intervenes between these levels is the biosynthesis of macromolecules.

We must consider now the experimental and theoretical implications of these observations. One set of experimental facts bearing on the question of the oscillatory properties of the epigenetic system and its relaxation time is the observed protein turnover rate of cells. Mandelstam (1960) has recently reviewed much of the important data on this subject, and concludes that turnover in most protein fractions is a characteristic of all cells, protozoan and metazoan, especially when they are in a 'resting' or non-growing state. The stationary state of the epigenetic system corresponds to the non-growing condition of cells, for which Mandelstam quotes values of 7 percent turnover per hour for microorganisms, and 1 percent per hour for mature mammalian cells. These represent average values over all cellular protein, some fractions of which may be quite inert, and some very active, The question now is what relation these experimental turnover rates bear to the theoretical cyclic oscillations of the epigenetic system.

The binary postulates imply that the amino acids of a cell are well 'stirred' among the proteins of the different epigenetic components, the dominance relations producing a constant conversion of one component into another, without cyclic reincorporation of the amino acids of a degraded molecule into itself again. Furthermore, the higher the epigenetic temperature, the more active is this stirring process, since the quantities of amino acids converted from one component into another are increased with increased amplitude of the oscillations. This theoretical endogenous cycling of amino acids within a cell could be observed by the usual methods employed for the study of turnover rates in proteins only if equilibration occurs between the

internal amino acid pool from which protein is synthesized, and the external amino acids which contain the added tracers. It is usually assumed that such equilibration does occur, and that the calculated turnover values measure the equilibration rate for proteins and amino acids. However, the observations of Cowie and McClure (1959) on metabolic pools in yeast cells indicate that this assumed equilibration between different amino acid pools may not occur. They have discovered two functionally distinct pools in Candida utilis, having quite different properties, only one of which equilibrates rapidly with exogenous amino acids. The other, 'internal', pool, which is directly connected with the processes of protein synthesis, seems to consist of amino acids in a different state from those in the expandable reservoir which equilibrates with the environment. If such pools occur generally, then observed turnover rates would represent only the net flow of amino acids through the internal pool and the cellular proteins, and would tell us very little about any endogenous interconversion of different protein species which may occur within the more or less isolated internal system. This is an important consideration, for the true turnover rates of an oscillatory system such as the one we are considering are very much higher than those actually observed, as we shall now show.

Considering the quantities used earlier in the calculation of the relaxation time, we have a protein species present in concentration q_i and undergoing five oscillatory cycles per minute, with mean amplitude $\frac{1}{100} q_i$. In one cycle, $\frac{1}{100} q_i$ is synthesized and destroyed, so that in one minute $\frac{1}{20} q_i$ will have turned over. This is a rate of 5 percent turnover per minute for this species! Even if we assume a

smaller mean amplitude of oscillation, say $\frac{1}{1000} q_i$, we still get a value of .5 percent turnover per minute, which is 30 percent per hour. It is therefore clear that either the observed turnover values are very low, due to some consideration such as lack of equilibration between amino acid pools; or that the frequencies and amplitudes of the epigenetic oscillations are much smaller than we have assumed. That no endogenous cycling at all occurs seems improbable, from the evidence (Mandelstam, 1960). It is quite possible that the epigenetic temperature of mature mammalian cells is indeed very low, the amplitudes of the oscillations being small ($\sim 10^{-4} q_i$) and giving turnover values in the region of 1 percent per hour, as observed. It might then be the case that turnover measurements on young embryonic cells will give much higher values, in accordance with the theory that θ is large in the early stages of development and drops to low values as cells reach their mature, differentiated condition. The mean frequencies of the oscillations might also be much lower than we have assumed, with perhaps only one or two cycles per minute instead of five. This would then give a greater value for the relaxation time. Unfortunately, the experimental evidence is not yet available to settle these points. It would be of considerable interest in this respect to be able to make direct and continuous observations on protein concentrations in single cells, a possibility which seems now to be within the range of technical achievement. This would circumvent any equilibration barriers which make the usual turnover calculations subject to error.

From an estimate of the relaxation time of the epigenetic system, it was concluded that parametric change must be very slow during the developmental process if the state of the epigenetic system is to be

described in terms of particular values of θ, G, F, etc. The rate of irreversible change must be such that the epigenetic system remains always very close to equilibrium so that, over the observation period of 20 minutes, the system may be said to be stationary, and the values of the above functions may be calculated. This means that any changes in gene activities which may be caused by a feed-back process from the epigenetic system to the genes must be correspondingly slow, requiring a few hours for significant changes to occur. We must consider now what experimental information there is on this aspect of the argument.

The first point to observe is that embryonic events are in general very slow compared with ordinary physical and chemical processes. (We confine our attention to post-gastrula and homologous developmental stages, when protein and RNP synthesis become significant in embryonic cells, and intracellular differentiation supersedes cell division in the developmental process.) Thus in Rana pipiens, it takes 8-10 hours for the neural plate to appear after the main topographical features of the primary germ layers have been established in the late gastrula. Cell division, a process representing roughly a doubling of all cellular constituents, requires about 20 hours in post-gastrula stages of this organism. The formation of a lens vesicle under the inductive influence of the optic cup requires from 24 to 30 hours. (These examples are from Rugh (1950), for Rana pipiens at 18°C.)

These are clearly very complex embryonic events, and do not give us direct information on the rates of more elementary processes. We come closer to the latter when we consider Holtfreter's (1938) experiments on the duration of various competences in isolated amphibian ectoderm. Depending upon the particular competence considered (brain,

ear-vesicle, mesenchyme, etc.) the competences last from 12 to 48 hours. This represents a slow rate of decay in the condition of whatever cellular components are responsible for the respective competence (assumed to be the genes), and is perfectly compatible with the rate we require for parametric changes in the epigenetic system.

A close study of the temporal relationships in another elementary embryonic process, that of induction in the central nervous system of amphibians, has been made by Johnen (1956). She found that a minimum contact time of four hours was required for competent ectoderm of the newt to be 'activated' (she uses Nieuwkoop's (1955) phrase) by inducing tissue (notochord). The activated tissue then develops autonomously for several hours. These results were confirmed by Toivonen (1958). Although it is not possible to know in detail what is happening within single cells in such experiments (the explanation of these results put forward by the above authors differ significantly), it is nevertheless clear that some process of activation in competent ectoderm cells has occurred only after an exposure time of about 4 hours to an inducing stimulus. Assuming that the genes are the essential factors involved in this activation process, we are led to the conclusion that a significant change in gene activities required a period of 4 hours.

A much clearer example of gene activation by cytoplasmic states is afforded by studies with the protozoan, Paramecium. This organism carries surface antigens whose nature is determined both by the state of the cytoplasm and by the potentialities of the genome. By transferring a genetic factor from an organism whose cytoplasmic state does not allow it to be expressed into an organism whose cytoplasmic states causes it to become active, it is possible to observe the time interval required

for the genetic factor to be activated and expressed (Beale, 1954). The minimum time required for this process is 5 fissions, which, under favourable environmental conditions, is about 25 hours (5 hours for one fission). Whether the feedback from cytoplasm to genes has occurred continuously or discontinuously, this time period would seem to be quite sufficient for a system with a relaxation time of about 12 seconds to remain in a near-equilibrium condition throughout the process (assuming that the earlier considerations apply to Paramecium). In this case we would certainly be justified in treating the genes as external parameters of the epigenetic system. Whether or not we can argue from the properties of Paramecium to those of embryonic cells is of course another question.

The applicability of a statistical mechanics to developmental processes and the treatment of the genes as an important set of controlling parameters of a quasi-stationary intracellular system rest upon the considerations of the present section. We have found some evidence in support of the requirement that changes in gene activities must be much slower than the calculated relaxation time of the epigenetic system, but the facts presently available do not allow us to reach any definite conclusions. However, if the present treatment is valid, then it should be possible to proceed further with a theoretical study of the changes of state which the epigenetic system may undergo, and how a particular sequence of parametric changes leads to the 'canalisation' of the epigenetic system in a particular creode (Waddington, 1957). This involves the development of a transformation theory for the irreversible motion of the epigenetic system, with which little progress has as yet been made (see, however, Appendix 1). More definite experimental justification

of the present approach to the analysis of intracellular events in development is required, furthermore, before a further theoretical extension of this method can be pursued with confidence. In particular, it would be of great interest to know if protein concentrations do undergo constant oscillations within single cells, and if the pattern of these oscillations for a particular protein species changes abruptly at some point as the steady state of the protein is increased from a low level in embryonic cells to a considerably higher one. These observations must await some further refinements in the technique of studying the composition of single cells, when it will become clearer whether or not the present statistical mechanics is applicable to the problem of development at the cellular level.

SECTION XII

CONCLUSIONS

The main problem which has been our concern in this study was to find some explanation for the discontinuities of the developmental process in terms of a theoretical model incorporating only the essential properties which are believed to underly the molecular organisation of embryonic cells. These discontinuities are observed in embryology as the emergence of discrete species of cell, arising from one single cell by a process involving the expansion of the initial system and the divergence of its derivative parts. The fundamental biochemical process underlying intracellular development was assumed to be the replication of macromolecules in a competitive environment. There are many possible mathematical representations of systems composed of replicating components which interact competitively. The particular equations derived in this study depend upon the assumption of a relatively strong type of self-accelerating replication, known as autosynthesis, and also strong competitive interaction between the replicating components. The equations thus obtained are the same as those derived by Volterra (1931) in a study of prey-predator relations in animal populations. We have seen that a statistical mechanical analysis of such systems is made possible by the existence of an integral of the differential equations. Since these equations are non-linear, it is of very considerable interest to study the macroscopic properties of the system, and to see if any discontinuous behaviour is to be expected. A discontinuity in such systems was in fact discovered by Kerner (1957), and the derivation of this result was given in section VII. Together with certain other properties and assumptions relating to parametric change in developmental systems, this property was used in section X

as the basis for the theoretical explanation of some fundamental embryonic processes, especially those relating to 'threshold' phenomena. The demonstration of threshold properties in the theoretical model studied in this work is important because it makes unnecessary the assumption of discontinuous change in system parameters in order to account for discontinuous behaviour in biological systems. Such an assumption is often made in theoretical discussions (e.g., Ashby, 1952); but we are then driven immediately to ask what mechanism is responsible for the switch in parameter-values, and the primary cause of the discontinuity remains undiscovered. However, in a Volterra system, statistical discontinuities (i.e. discontinuities in the Boltzmann distribution of individual components) occur as a consequence of the non-linear structure of the differential equations, and the phenomenon is thus traced ultimately to the mathematical representation of the two basic assumptions used in constructing the equations: replication of components, and competitive interaction between them. It has been felt for some time by many biologists that these two fundamental aspects of the microscopic structure of biological systems may be sufficient to account for many of their essential macroscopic properties, especially those relating to discontinuous behaviour (see, e.g., Waddington (1948) and Spiegelman (1948) for developments of this argument in embryology), but this has been without actual demonstration. The present study therefore represents a further elaboration of this view, and presents specific mathematical support for it.

It is quite possible that the particular mathematical model used in this work to represent the macromolecular organisation of embryonic cells may be more non-linear than the system really is. Both autosynthesis and binary interaction are 'strong' assumptions, and it may be found

necessary to replace them by somewhat weaker ones. There are many ways in which the concepts of replication and competitive interaction may be translated mathematically, and each will lead to somewhat different results and predictions, some of which may approximate more closely to the properties of embryonic cells than the model used. But on the basis of the results obtained for Volterra systems, it may be expected that statistical discontinuities will occur in the behaviour of systems whose microscopic organisation involves replicating and competitively interacting components, features which almost certainly require a non-linear representation. The more generalised systems corresponding to weaker representations of replication and competitive interaction will probably be more complex mathematically than Volterra systems; and the addition of time-lags and diffusion terms would certainly destroy the simplicity of the Volterra equations, at the same time increasing the representational accuracy of the model. However, so long as such systems are conservative in some sense, so that an integral may be found for their differential equations, and a statistical mechanical analysis is possible, then the macroscopic properties of the system may be studied without too serious mathematical difficulty. The requirements which must be satisfied in real systems in order that a statistical mechanics be applicable to the analysis of their dynamics was discussed at some length in section XI, and the mathematical advantages of this technique were also pointed out. Evidence was presented in support of the applicability of this method to embryonic cells, and particularly the possibility of treating the genes as controlling parameters of the developmental system. However, the final decision on this question must await a finer experimental

examination of embryonic cells, and a more detailed quantitative analysis of their dynamics.

We have spoken of the developmental process as one which shares certain basic features with other types of biological process, with the suggestion that the common factors involved in the 'microscopic' organisation of a certain class of biological systems justifies the use of the descriptive term 'Darwinian' in reference to them. Our chief concern in this study has been to demonstrate how it is that one particular phenomenon, characteristic of the behaviour of such systems, may be accounted for in terms of their microscopic properties. Particular assumptions were made in an attempt to fit the analysis to the embryological case, and the system derived from these assumptions was one which obeyed Volterra dynamics. The existence of a general macroscopic parameter, θ, relative to which discontinuities occur in the distributions of single components in such a system, was found to be very important in suggesting how certain embryological phenomena could be interpreted in terms of the theoretical model. In particular, the possibility of explaining the molecular processes underlying embryonic induction and the threshold effects observed therein was investigated in some detail in section X. It was found that θ may be used as a measure of the reactivity of the epigenetic system to environmental stimuli, while the distance of the major epigenetic components from θ determines their stability relative to the various epigenetic forces which act upon the system. As differentiation proceeds and the molecular ecology of the cell becomes better defined, it was found that θ may be expected to decrease, and the epigenetic system then becomes less reactive and increasingly stable towards environmental stimuli. The 'random'

oscillatory motion of the epigenetic system gradually decreases as the system becomes stabilised in a well-defined creode which leads to a particular differentiated state, and a statistical mechanics such as the one used in this study ceases to be useful in describing the further motion of the 'determined' system.

The question now arises whether or not macroscopic parameters like θ may exist for more general types of Darwinian system, in terms of which their behaviour at different stages in their evolution or development may be described. There is one observation which suggests that this may be the case, an observation which relates the parameter θ to the 'adaptive' properties of a Volterra system.

The concept of adaptibility is a fundamental one in any study of Darwinian systems, but we have not used it so far in the developmental theory which has been developed. The major difficulity involved in using this concept is to define with some accuracy the criterion whereby a biological system may be said to be adapted to its environment, in terms somewhat finer than mere survival as against death. In the evolutionary context, a degree of precision has been achieved by the use of the criterion of 'fitness' to measure the adaptation of an organism or a population to its environment. But in embryology, no corresponding criterion has been suggested, and indeed there may be some question as to the usefulness of this concept in explaining the responses of embryonic cells to environmental stimuli. However, adaptability is such a central phenomenon in the behaviour of living systems that it seems inevitable that this concept must enter into the description of the behaviour of embryonic cells, even if we do

not yet have an experimental measure of adaptation in this case (see Appendix 1 for one way in which this concept can be used to obtain a result of some interest for a developmental theory, without actually defining adaptation explicitly).

There is a distinction between the two concepts, adaptation and adaptability, and it is really the latter in which we are interested at the moment: the capacity of a system for adaptive response. Without answering in detail the question of what an adaptive response is, we can still discuss the properties which a system must have in order that it can change its state with considerable versatility and rapidity in response to a wide variety of environmental conditions. In somewhat different form, this problem is being studied by control engineers whose task is to construct fully automatic machines which operate at optimum values under a variety of changes in plant (machine) parameters and noise level (these constituting the 'environmental conditions' of such a system). These 'self-optimising systems' function under an adaptive criterion which is usually based upon economic considerations, such as least overall cost, maximum profit, etc. The point which is of great interest to us is that adaptation in such systems has been found to involve a continuous system oscillation, caused primarily by the constant movement of the hunting device which continuously seeks the optimum operating conditions (Tsien, 1954; Milsum, 1960). We may say that any variable components in an adaptive system must constantly undergo cyclic variations in order that all possible states of the system be tested for optimum performance in the prevailing environmental conditions. In engineering design it is necessary to construct the machine in such a manner that oscillations will occur. In a Volterra system, such oscillations occur as the result

of the microscopic structure of the system, and there is no need to add forcing functions to such a system in order to make its components oscillate: they do so spontaneously. We thus see that a Volterra system has one of the properties which is essential for adaptive response to environmental stimuli, and indeed it is just this property which underlies the statistical mechanics developed in earlier pages. That quantity which is conserved in a Volterra system is in large part the dynamic, oscillatory motion of the components, and Θ is, as we have seen, a measure of this oscillation as the mean square deviation of the components from their stationary concentrations.

Now the frequency and the amplitude of the oscillations of an adaptive system determine the rate at which it can respond to stimuli, and how close it will remain to the optimum state once it has responded. When frequency and amplitude are large, the system responds rapidly, but it does not remain very close to the optimum state because of the large oscillations. The system may then be said to sacrifice accuracy of adaptive response to rapidity. In the Volterra system, this situation is defined by large values of Θ, which condition we suggested might correspond to the metabolic states of embryonic cells in early developmental stages. This condition was described as a highly reactive one, the cell being capable of rapid response to a wide variety of stimuli; i.e. it has a wide range of competences. A slow decrease in Θ during the developmental process would then correspond to a decreasing rapidity of adaptive response to epigenetic stimuli, and a closer approach to the state of maximum adaptation. When Θ is near zero, the system's capacity for adaptive response is very much reduced, and it becomes stabilised in some creode; while at $\Theta = 0$, the adaptibility of the epigenetic system vanishes. The study of adaptive systems

with many degrees of freedom has only recently begun, and the relationships which appear to hold between the epigenetic temperature and the adaptive properties of a Volterra system require further investigation before they can be regarded as rigorously established. Other important macroscopic relationships may then begin to emerge which could, for example, lead to an important interpretation of the epigenetic equation of state derived in Section IX, and an identification of the variable F_T. However, the possibility of using the fundamental macroscopic parameter, θ, to describe the adaptive behaviour of a system obeying Volterra dynamics suggests that similar parameters may exist for more general Darwinian systems, and that such parameters may occupy an equally central position in any discussion of their adaptive behaviour. The microscopic structure of such systems, involving the basic 'forces' of replication and competitive interaction, may produce continuous system oscillations, as in the Volterra system, which oscillations then form the basis for their adaptive properties. Further theoretical work on systems with weaker interaction than that assumed for the Volterra case will be required before this suggestion can be properly explored. But if such a result is obtained, then a number of diverse aspects of Darwinian theory might be drawn together within a comprehensive structure, and the power of theoretical biology would be greatly extended.

In conclusion, the theoretical considerations of this study are advanced as a contribution towards the general analysis of biological phenomena, with particular reference to the organisation and behaviour of embryonic cells. Certain features only of biological systems were studied, with a view towards the explanation of some fundamental behavioural characteristics which distinguish them from those systems

which have been successfully studied from the point of view of physical analysis. These peculiarly 'biological' characteristics include the occurrence of threshold phenomena, with consequent discontinuous behaviour in the system, and the complex property of adaptability. Most of the study has been concerned with the first of these, with the elaboration of the mathematics necessary for its statistical analysis, and with the implications of this analysis in the study of embryological development. How adaptability arises in the theoretical system as a result of its microscopic structure has been only briefly considered.

There is no doubt that these two phenomena, discontinuous behaviour and adaptability, are of very considerable biological interest and importance. However, it must be admitted that in the embryological context, they represent only part of the problem. Furthermore, the analysis as developed is applicable only to intracellular events. There remains the whole field of inter-cellular processes, and the enormously complicated problem of the physico-chemical forces underlying morphological changes in developing cells and tissues. Besides this, we have the question of gene-control and the mechanism of information feed-back which was introduced into our arguments without any theoretical elaboration. These questions have not been treated here, and they seem to represent the major part of the embryological puzzle. However, the arguments developed in this work may contribute to the understanding of development by showing how a statistical mechanics may be applied to the study of cellular phenomena, and what macroscopic consequences arise from particular microscopic assumptions, together with their implications for further experimental study. The mathematical analysis of biological phenomena is bound to be difficult, because of the non-linearities which occur

in the equations describing their motion. At the same time, these non-linearities have very interesting consequences, and our understanding of the behaviour of all real systems, physical as well as biological, will be greatly extended by a concerted attack upon the 'non-linear barrier' (Kovach, 1960). The statistical methods used in this study represent one approach to this problem; and it is to be hoped that the extension of these methods to more general types of Darwinian system may lead to an understanding of the basic organisation and behaviour of a wide class of biological systems within the structure of a single comprehensive theory.

APPENDIX I

The statistical theory used in this work to study some macroscopic aspects of embryonic cell behaviour is a theory of stationary processes. A stationary state defined by the values $(q_1, \ldots q_n)$ of the epigenetic components was first assumed to exist, and then the behaviour of the system in the neighbourhood of this state was investigated. Such a study cannot give any information about changes in the steady state values, q_i. These are defined by the values of the system parameters, ϵ_i and α_{ij}, and only if we have some knowledge about how these change can we determine what changes are to be expected in the q_i's. Some simple assumptions were made in Section X about the changes which might occur in the parameters α_{ij} during the process of induction, and the effects of these changes on the q's. However, the question we wish to consider now is the possibility of discovering some general principles governing the sequence of changes in the stationary states of a cell during development. This would considerably extend the theoretical analysis of the developmental process, and would be an important complement to the statistical mechanics used here.

There are two ways of proceeding with this question. We can formulate the problem in terms of the mathematical structure of the statistical theory already used. In this case we must ask what transformations may occur in the matrix T $(T = (\gamma_{ij})$, where $\gamma_{ij} = \frac{\alpha_{ij}}{\beta_i/\beta_j})$, and in the vector σ ($\sigma =$ $(-\alpha_{10}, \cdots -\alpha_{n0})$), satisfying the relations $T\tau = \sigma$ $(\tau > 0)$ and any further constraints which are imposed by the developmental mechanism. The set of all these possible transformations gives us the set of all possible developmental sequences, $\tau = T^{-1}\sigma$. The differentiation of any particular cell type

should then be given by one of these sequences.

Alternatively, we can proceed with the problem independently of the statistical mechanics, using any information we have about the properties of developing cells to set up a model for the sequence of changes in the q's. We are then free to make any assumptions which seem reasonable regarding the dynamics of development, providing only we do not contradict the assumptions already made for the statistical mechanics.

The first method would seem to offer certain advantages, and be more consistent with our whole approach. Thus we could use the assumptions already introduced regarding irreversible change, and interpret them in terms of parametric changes in T and σ. These assumptions include information feed-back from epigenetic state to genes, the assumed effects of inductive stimuli, slow decreases in $|\alpha_{ij}|$ and possible increases in the β's, as well as a slow decrease in θ, and other constraints such as those imposed by the basic architecture of the egg (as discussed by Kacser, (1960)). However, these assumptions are not easily translated into mathematical conditions governing changes in T and σ, especially the important one of information feed-back. This last question requires a very considerable extension of the theory, and a much more detailed knowledge of the biological facts underlying the process. Thus having stated the problem in this form, we will immediately leave it, and consider the alternative approach.

In constructing theories about the organisational principles of biological systems, much attention has recently been paid to the analogy between organisms and communications systems. The application of the ideas and concepts developed in connection with information transfer to the organisation and activities of organisms has led to the possibility

of some kind of 'information-theoretic' approach to the analysis of biological process. It thus seems possible that development may be profitably considered from this point of view, with the epigenetic components playing some part in a communication system, their concentrations obeying some general distribution law which can be deduced from certain communication criteria applying to the system. We shall now follow this suggestion, using an analogy between the 'informational' structure of cells and that of language, with proteins as the analogues of words and epigenetic components as the analogues of sentences. A major difficulty arises in this study when we come to define the criterion whereby cells organise their information resources to produce a system maximally adapted to its current environment. This criterion is the principle of 'maximum adaptation' which has been mentioned in the conclusion of the thesis, a principle which seems to characterise the behaviour of living systems but has as yet been given no satisfactory definition from developmental processes. According to such a criterion the concentrations of the epigenetic components at any state of development will be determined by their 'adaptive value' relative to the prevailing environmental conditions; and the epigenetic system will move, that is to say the concentrations of the epigenetic components will change, in such a manner as to always maximise the adaptation of the system to its environment. This implies some principle of least action in biology, but in the absence of an explicit definition of the biological analogue of 'action', the principle has intuitive value only. There is a way round this difficulty, however, leading to a result of some interest. In the following we shall use an extension of a very interesting method used by Mandelbrot (1954) in a study

of the 'informational' structure of language.

The transfer of information from template to product during macromolecular synthesis is now recognised as a very fundamental type of communication process in biological systems, involving some kind of biological code. The elementary signals in this code, analogous to the letters of the alphabet in language, are believed to be the chemical units out of which the macromolecules are synthesised, nucleotides in the case of DNA and RNA, and amino acids in the case of proteins. Information could then be carried in the linear sequence of the units making up any particular macromolecule. It is this observations which has led to analogies between the structure and function of macromolecules in cells, and those of words in language. If the analogy is a good one, then certain properties of words and language should be found in macromolecular organisation, and attempts have been made to demonstrate this, especially for proteins. For instance, it is known that the sequences of letters in the words of any language have certain constraints, such that some sequences are much more likely than others. The amino acid sequences are now known for a number of proteins, and a statistical analysis of the pattern of these sequences might be expected to reveal constraints similar to those between the letters of words. This has so far given negative results, but the reason for this is simply that the number of proteins whose amino acid sequences are known is still far too small: the sample is not yet sufficiently large to give statistically significant results.

The very existence of a biological code in macromolecules remains to be demonstrated, and this is the concern of those engaged in the well-known 'coding problem' for nucleic acids and proteins. It is

here that the theorems of information theory are most obviously applicable, and may lead to important conclusions. However, the coding problem is not directly relevant to our present problem, which is to discover something about the macroscopic structure of cells. Assuming that the microscopic evidence, although largely circumstantial, is sufficient to justify the use of an analogy between words and proteins, it is necessary now to pursue the analogy between language and cellular organisation at a higher level than that of letters and words in order to find possible theorems relating to the distribution of the epigenetic components. Since proteins are regarded as the analogues of words, the epigenetic components must be regarded as the analogues of phrases or sentences, the next level of structure in language above words. Because we will treat the epigenetic components as the uncorrelated symbols of a developmental message, thus allowing the use of the Shannon information function of their probabilities, we regard them as being analogous to sentences, between which constraints are considerably less strong than those between phrases.

The argument, to be developed more precisely below, is the following. At any stage of development in a single cell, a message is being transmitted from the genes and the environment of the cell. The message is carried by a sequence of symbols, the epigenetic components, whose statistics (i.e., their steady state concentrations) are determined by the 'epigenetic environment' (gene activities and cellular environment) according to some principle of 'maximum adaptation'. The information in the message is, in effect, a set of instructions for the metabolic and other activities of the cell, resulting in a state of maximum adaptation for the given environmental conditions. We thus assume

that the functional state of a cell is determined primarily by the nature and the quantities of the epigenetic components, as was argued in Section II, and that these change with changes in the 'epigenetic environment' so as to maintain a condition of maximum adaptation in the cell. This argument leads to certain conclusions about the distribution of the epigenetic components, as will now be shown.

Denote by E_i the ith symbol, called a species of epigenetic component, with i varying from 1 to n, the potential number of different symbols (components). Let p_i be the probability of occurrence of E_i. The E_i's are made up of more elementary symbols, the proteins, P_s,s varying from 1 to S, say. (These symbols may correspond to polypeptide units smaller than functional protein, pre-proteins, which are the generalisation of Halvorson's hypothetical pre-enzymes (Halvorson, 1960), out of which specific proteins are made. However, we shall continue to refer to the P_s as proteins). With each P_s we associate a single additive 'cost function', $C_s(a)$, which defined the 'price' which a cell must pay in producing the protein P_s under the environmental conditions defined by a ($a=(a_1,a_2,\ldots a_m)$). This cost function is assumed to bear an inverse relationship to a value function, $V_s(a)$, which defines the 'value' of the particular protein P_s to the cell under the environmental conditions a; i.e.,

$$C_s = f\left(\frac{1}{V_s}\right)$$

, f being monotonic in $\frac{1}{V_s}$ and having an inverse, f^{-1}. This means that the greater the value of a protein, the less does it cost the cell to produce it, this cost being measured chiefly in terms of the behaviour which results from having P_s in the cell.

Thus if P_s produces an activity of high adaptive value to the cell, then C_s is low, and vice versa. The reason for using a cost function rather than a value function is that it gives a simpler form to the equations. The existence of a single cost for each protein means that all incompatibilities have been resolved in assigning prices to the symbols, P_s. Assuming that the function f has an inverse, this condition is equivalent to the existence of a well-defined value function which assigns to each protein a particular adaptive value under the given environmental conditions, according to some undefined criterion. We might have as the functional relation between C_s and V_s, for example,

$$C_s = \log \frac{1}{V_s} \quad ; \quad V_s = e^{-C_s} \qquad \begin{matrix} 0 \leqslant V_s \leqslant 1 \\ 0 \leqslant C_s \leqslant \infty \end{matrix}$$

Then additivity of the C_s's becomes multiplicativity of the V_s's:

$$C_1 + C_2 = \log \frac{1}{V_1} + \log \frac{1}{V_2} = \log \frac{1}{V_1 V_2} .$$

However, very little can as yet be said about the possible nature of such functions as V_s and C_s, and we shall only assume their existence. The problem now is to discover the p_i's, using certain principles of information transfer which are assumed to be satisfied by the communication system.

The first result which we shall establish is a relationship between the cost of an epigenetic component and its 'rank', by which is meant the position of the component in a linear sequence when all the components are ordered according to increasing cost. Thus the component with rank 1 has least cost, then component with rank 2 follows with a greater cost, etc. This ordering also establishes the rank of a component in terms of decreasing adaptive value: the component with rank 1 has greatest adaptive value, etc.

The number of components of cost C is given by the difference

relation

$$N(C) = N(C-C_1) + N(C-C_2) + \ldots\ldots\ldots\ldots + N(C-C_S).$$

This means simply that all components costing C can be decomposed into the sum of those components costing C whose last symbol is P_s, $s = 1,2, \ldots\ldots\ldots S$. The general solution of this difference equation is

$$N(C) = \sum_{i=1}^{S} d_i M_i^C \quad , \text{ the } d_i\text{'s being determined by the}$$

initial conditions. The number of components which cost C or less is then

$$\sum_{c=1}^{C} N(c) = \sum_{c=1}^{C} \sum_{i=1}^{S} d_i M_i^c = \sum_{i=1}^{S} d_i \left(\sum_{c=1}^{C} M_i^c \right)$$

$$= \sum_{i=1}^{S} d_i M_i \frac{1 - M_i^C}{1 - M_i} = \sum_{i=1}^{S} B_i (M_i^C - 1)$$

$$\text{where } B_i = \frac{d_i M_i}{M_i - 1}$$

Thus $$N(C) = \sum_{i=1}^{S} B_i M_i^C - B_0$$

$$\text{where } B_0 = \sum_{i=1}^{S} B_i$$

Now it is known that such equations have only one real, positive root, which we denote by M. Thus for C large, we have, approximately,

$$N(C) = B_1 M^C - B_0$$

But we can write N(C) = r, the rank of a component, since this rank is simply the order of the component in terms of increasing cost.

We therefore find

$$r = B_1 M^C - B_0$$

or, $$C_r = \log_M (r + w) + \log_M b \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1).$$

$$w = B_0$$

$$b = \frac{1}{B_1}$$

where C_r is the cost of the component of rank r.

We want next to find a relation between p_r, the probability of occurrence of component E_r, and its cost, C_r. We assume that p_r varies with $\frac{1}{r}$, so that as rank (or cost) increases, probabilities decrease. In order to find a relation between p_r and C_r, we shall introduce a communication criterion in the following form. The average cost per component is $\sum_{r=1}^{n} p_r C_r = C$, and the information carried by the message with probabilities p_r is $H = -\sum_{r=1}^{n} p_r \log p_r$. We assume that the best weighted set of symbols (epigenetic components) is that one which minimises the average cost per unit of information; i.e.,

$$\frac{C}{H} = \frac{\sum p_r C_r}{-\sum p_r \log p_r}$$ is a minimum. One further condition

condition is that $\sum_{r=1}^{n} p_r = 1$.

Following the usual procedure in determining p_r from these conditions, we find that

$$P_r = P' M^{-BC_r} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2).$$

B is a positive quantity, since p_r decreases with increasing C_r, and its actual value is determined by H/C ;

$$B = \frac{H}{C \log M}$$

$\frac{1}{B}$ could be called the informational temperature of the weighted components. P' is fixed by the condition $\sum_{r=1}^{n} p_r = 1$.

From equations (1) and (2) we can eliminate the unknown function C_r, with the result

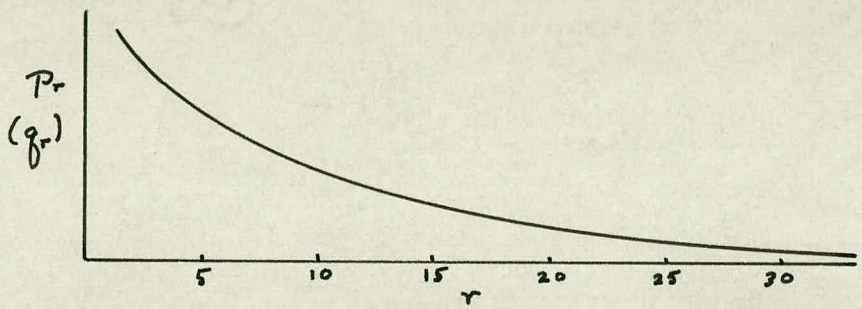
$$p_r = P' M^{-B} \left[\log_M (r + w) + \log_M b \right]$$

$$= P' (r + w)^{-B} b^{-B}$$

$$= P (r + w)^{-B} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3).$$

$(P = P' b^{-B}.)$

Thus we have a relation between the probability of occurrence of an epigenetic component and its rank in order of increasing cost. If one could identify all the different epigenetic components of a single cell and determine their concentrations, ordering the components according to decreasing concentrations (decreasing probabilities), then a graph of concentration against rank (position in the ordering) should give a curve of the form (3). The distributions (3) are all of skew type, the detailed shape of the curve being determined by the parameters B and w (about which more will be said below). The most significant feature of these curves is the rapidity with which p_r decreases over the first few values of r, and the very long 'tail' which occurs for large r, as is shown in the figure.



Since the ordinate could equally well be the steady state concentrations q_r, we see that the result predicted by our argument is that a few epigenetic components should be present in large and intermediate concentration, while the majority should occur in small concentration. Now the ordering of the q_r's in increasing concentration is just the opposite ordering from that given in Section VII, where the components were divided into two categories according to their relation to the epigenetic temperature, θ, except that the parameters β enter into the latter relations. Roughly, then, the distribution (3) implies that as the epigenetic system approaches the stationary state for a given set of

environmental conditions, the distribution of its components will be such that a few will be present in significant concentrations, but the majority will occur in small concentration. If for this steady state Θ has a significant value, so that $\Theta > \beta_r q_r$ for those components with small q_r, then the predicted distribution of the q_r's is such that the majority of the components will belong to the minor category. Those components which are selected are the ones with lowest 'cost', i.e., highest adaptive value, for the prevailing environmental conditions.

The derived distribution applies to the epigenetic components only after the 'developmental message' has been transmitted for some time, sufficient for the statistics to reach a stationary condition for the given environmental conditions. How long this may be depends upon the rate of transmission compared with the rate of change of the environmental parameters. This brings us once again to the question of the rate of macromolecular synthesis. This biosynthetic rate must be much more rapid than the rate of change of the epigenetic environment, for otherwise the statistics of the system will be constantly changing with changing $C_r(a)$, and the distribution (3) will not be reached until $(a_1, a_2, \ldots\ldots a_m)$ reaches a constant value; i.e., until development is complete. This question is essentially the same as that relating to the relaxation time of the epigenetic system and the rate of parametric change, discussed in Section XI. In view of the conclusions arrived at there, we shall assume that the transmission rate of the developmental message is sufficiently high that the distribution (3) applies to every stationary (or quasi-stationary) state of the epigenetic system, and not just to the final distribution of components in the fully differentiated cell.

The distribution (3) is a very efficient one in the sense that the cell channels its synthetic activities towards the production of a few components which are of greatest value in any particular environment. An example of this type of behaviour may be observed in microorganisms during adaptation to new energy sources, when rapid synthesis of a few key proteins occurs, the concentrations of these proteins soon reaching a level which accounts for a very considerable fraction of the total protein of the cell. In the embryological context, such distributions are suggested by the occurrence of large concentrations of a few species of antigen in developing heart and lens tissue (Ebert, 1950; Clayton, 1954), other species not being detectable; and on the theoretical side, it seems possible to regard the 'histogenetic key substances' (Waddington, 1956) around which the further development of particular tissues is mobilised as the components of low cost (high adaptive value) which are predicted to occur in high concentration. It is, furthermore, true that differentiated cells are characterised by high concentrations of certain proteins, while a great variety of proteins which are included in the genetic potential of these cells are undetectable. There may thus be an experimental basis for the distribution (3), although a complete justification for the use of such a distribution would require a much more complete knowledge of protein concentrations in single cells.

It is of interest to consider briefly some further properties of the family of distributions given by (3). The parameter w is a sort of 'degeneracy' parameter which relates to the state of lowest cost in the epigenetic system. If B > 1, so that the sum $\sum_1^n p_r = \sum_1^n p(r+w)^{-B}$ converges as $n \to \infty$, then w has an appreciable effect upon this sum, and hence affect $\frac{1}{p} = \sum_1^n (r+w)^{-B}$ as $n \to \infty$ ($\frac{1}{p}$ is a 'sum-over-states').

The convergence of this sum for $B > 1$ implies that the number of potential epigenetic components, n, is effectively unlimited. But when $B < 1$, we must take n finite in order to obtain a finite value for $\frac{1}{p}$. In this case w has much less effect in determining $\frac{1}{p}$ than does the finite value chosen for n. Since $B = \frac{H}{C \log M}$, we see that the above inequalities are

$$\frac{C}{H} < \frac{1}{\log M} \quad \text{(n in finite)} \quad \text{and} \quad \frac{C}{H} > \frac{1}{\log M} \quad \text{(n finite).}$$

The conditions of convergence are therefore relative to the average cost per unit of information carried by the epigenetic components, but it would be premature to attach embryological significance to these mathematical conditions, particularly in view of the difficulty of identifying C with any measurable quantity.

Distributions of the type (3) have been found to fit a surprisingly wide set of data accumulated in biological and sociological studies. Thus Willis (1922, 1940) found that the distribution

$$p_r = Pr^{-(\alpha + 1)} \qquad \text{.....(4)}$$

$$0 < \alpha < 1$$
$$1 \leqslant r < \infty$$

gave a good fit for the frequency of occurrence of genera having r species. (see also Yule, 1924). Zipf (1949) used such distributions to fit a very large variety of statistical data including word distributions in texts, income distributions, sizes of cities and towns, and a host of other examples. He suggested that these distributions are as central to the social and biological sciences as the Gaussian distribution is to random phenomena in the physical sciences. Mandelbrot (1956) has developed the theoretical basis for the use of the Willis distribution in taxonomic studies in a very original manner. He showed that distributions of type (4) belong to the category known as exceptional stable Cauchy-Levy distributions,

having as characteristic function the general form

$$\varphi(u) = \exp\left\{-a|u|^{\alpha}\left(1 + \frac{iu}{|u|}\tan\frac{\alpha\pi}{2}\right)\right\} \qquad 0 < \alpha < 1$$

The Willis distribution is obtained for $\alpha = \frac{1}{2}$. These distributions have the peculiarity that the expected or mean value of r is infinite. This is also true of the distributions (3) if $1 < \mathcal{B} < 2$. The tendency of such systems to a non-normal limit is characterised by the fact that each component is no longer negligible relative to the whole, a condition quite different from that in physical systems where the behaviour of a single component is of practically no significance for the statistical behaviour of the whole, and where random quantities all but disappear as the number of components gets very large. We may therefore expect that systems having exceptional distributions may show statistical behaviour quite different from that which we associate with physical systems. Of particular interest in the developmental context is the possibility that the distribution (3) may hold for the epigenetic components with $1 < B < 2$. In this case the expected rank of a component is infinite, hence $p_r = 0$ for the great majority of the epigenetic components, and thus only a few components will occur in significant concentrations. This observation is of considerable interest, for it shows that a system organised as a communication channel in the manner assumed above tends spontaneously towards a highly asymmetrical distribution of its components; i.e., it will always tend towards a differentiated condition. Such behaviour is again quite different from that obtained for physical systems, where the variables tend to be symmetrically distributed about some mean value. This result suggests that certain features of developmental systems may arise from fairly elementary properties of the macromolecular organisation of cells. However, these conclusions must be regarded as very tentative ones, subject to further

justification for the use of the functions $C_r(a)$, their identification with measurable biological quantities, and also some experimental test of the derived distribution law (3). It is nevertheless of some interest to see how the concepts of communication theory may be used to obtain results which are outside the scope of a statistical mechanics, but which are complementary to and an extension of the latter theory.

APPENDIX 2

To the autosynthetic equations

$$\left.\begin{aligned} \frac{dP}{dt} &= k_P R \\ \frac{dR}{dt} &= k_R P \end{aligned}\right\} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

are added terms representing the destruction and decay of protein and RNP,

$$\left.\begin{aligned} \frac{dP}{dt} &= k_p R - K_1 P \\ \frac{dP}{dt} &= k_R P - K_2 R \end{aligned}\right\} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

The solutions of these equations are given by the expressions

$$P = a_1 e^{\lambda_1 t} + a_2 e^{\lambda_2 t}$$

$$R = b_1 e^{\lambda_1 t} + b_2 e^{\lambda_2 t}$$

where λ_1, λ_2 are the roots of the characteristic equations

$$\lambda^2 + (K_1 + K_2)\lambda + (K_1 K_2 - k_p k_R) = 0$$

These roots are

$$\{\lambda_1, \lambda_2\} = \frac{-(K_1 + K_2) \pm \{(K_1 + K_2)^2 - 4(K_1 K_2 - k_p k_R)\}^{\frac{1}{2}}}{2}$$

The condition that the system eventually decay, i.e., that $P, R \rightarrow 0$ as $t \rightarrow \infty$, is that both the roots λ_1 and λ_2 be negative. This is the case if $(K_1 K_2 - k_p k_R) > 0$, or $K_1 K_2 > k_p k_R$. The destruction coefficients, K_1 and K_2, must therefore satisfy this inequality in order to ensure that the system decay.

APPENDIX 3.

Infinitesimal Operators.

For a general system of differential equations,

$$\left.\begin{aligned} \frac{dx_1}{dt} &= f_1(x_1, x_2, \ldots\ldots, x_n) \\ \frac{dx_2}{dt} &= f_2(x_1, x_2, \ldots\ldots, x_n) \\ &\\ \frac{dx_n}{dt} &= f_n(x_1, x_2, \ldots\ldots, x_n) \end{aligned}\right\} \quad \ldots\ldots\ldots (1)$$

the infinitesimal operator of the system is defined by

$$D = f_1 \frac{\partial}{\partial x_1} + f_2 \frac{\partial}{\partial x_2} + \ldots\ldots + f_n \frac{\partial}{\partial x_n} .$$

This operator represents the total time differentiator of the system, and the solution of the set of equations (1) can be written in the form

$$x_1 = e^{Dt} (x_1)_0$$

$$x_2 = e^{Dt} (x_2)_0$$

$$x_n = e^{Dt} (x_n)_0$$

where e^{Dt} is defined as the operator

$$1 + Dt + D^2 \frac{t^2}{2!} + \ldots\ldots + D^r \frac{t^r}{r!} + \ldots\ldots$$

and $(x_i)_0$ is the value of the function x_i at $t = 0$. The above solution, written in expanded form, is therefore

$$x_1 = (x_1)_0 + [Dx_1]_{(t = 0)} t + [D^2 x_1]_{(t = 0)} \frac{t^2}{2!} + \ldots\ldots$$

Now the Volterra equations are

$$\frac{1}{X_i}\frac{dX_i}{dt} = \alpha_{i0} + \frac{1}{\beta_i}\sum_{j=1}^{n}\alpha_{ij}X_j \qquad (i = 1, 2, \cdots n)$$

Writing $x_i = \log X_i$, this becomes

$$\frac{dx_i}{dt} = \alpha_{i0} + \frac{1}{\beta_i}\sum_{j=1}^{n}\alpha_{ij}e^{x_j} \equiv f_i(x_1, x_2, \cdots x_n)$$

The solution of these equations is therefore, according to the above formulation,

$$x_i(t) = e^{Dt}(x_i)_0$$

$$= (x_i)_0 + [Dx_i]_0 t + \cdots + [D^r x_i]_0 \frac{t^r}{r!} + \cdots$$

where

$$D = \sum_i f_i \frac{\partial}{\partial x_i}$$

The solution to the first order in t is simply

$$x_i(t) = (x_i)_0 + [Dx_i]_0 t + O(t^2)$$

$$= (x_i)_0 + [f_i]_0 t + O(t^2)$$

Returning to the variables X_i, this becomes

$$X_i(t) = (X_i)_0 e^{[f_i]_0 t + O(t^2)}$$

and

$$[f_i]_0 = \alpha_{i0} + \frac{1}{\beta_i}\sum_{j=1}^{n}\alpha_{ij}(X_j)_0$$

We therefore obtain the required result,

$$X_i(t) = (X_i)_0 e^{\alpha_{i0}t + \left[\frac{1}{\beta_i}\sum_{j=1}^{n}\alpha_{ij}(X_j)_0\right]t + O(t^2)}$$

APPENDIX 4

I. Effect of a non-specific destructive agent upon the epigenetic system.

In order to obtain a clear and relatively simple result, it is necessary to make some simplifying assumptions about the distinction between auto- and heterosynthetic components. Assume that the heterosynthetic components (n - r in number, say) dominate the autosynthetic components (r in number), and that within each group there are no binary interactions. Then the matrix of the system takes the form

$$\begin{bmatrix} 0 & \cdots & 0 & \alpha_{1,r+1} & \cdots & \alpha_{1,n} \\ 0 & \cdots & 0 & \alpha_{2,r+1} & \cdots & \alpha_{2,n} \\ \vdots & & & & & \\ 0 & \cdots & 0 & \alpha_{r,r+1} & \cdots & \alpha_{r,n} \\ \alpha_{r+1,1} & \cdots & \alpha_{r+1,r} & 0 & \cdots & 0 \\ \vdots & & & & & \\ \alpha_{n,1} & \cdots & \alpha_{n,r} & 0 & \cdots & 0 \end{bmatrix}$$

If components 1 to r are the autosynthetic components, while components r + 1 to n form the heterosynthetic group, then the terms in the upper right block of the matrix are negative, while those in the lower left block are positive, and all other terms are zero. Now it is readily seen that an antisymmetric matrix of this form will be singular unless $r = \frac{n}{2}$, and therefore we must assume further that there are equal numbers of auto- and heterosynthetic components.

The equations of the stationary state are now

$$\left.\begin{aligned} \alpha_{1,r+1} q_{r+1} + \alpha_{1,r+2} q_{r+2} + \cdots + \alpha_{1,n} q_n &= -\beta_1 \alpha_{10} \\ \vdots & \\ \alpha_{r,r+1} q_{r+1} + \alpha_{r,r+2} q_{r+2} + \cdots + \alpha_{r,n} q_n &= -\beta_r \alpha_{r0} \\ \alpha_{r+1,1} q_1 + \alpha_{r+1,2} q_2 + \cdots + \alpha_{r+1,r} q_r &= -\beta_{r+1} \alpha_{r+1,0} \\ \vdots & \\ \alpha_{n,1} q_1 + \alpha_{n,2} q_2 + \cdots + \alpha_{n,r} q_r &= -\beta_r \alpha_{r0} \end{aligned}\right\} \quad \cdots\cdots\cdots\cdots (1)$$

In these equations the terms $\alpha_{1o}, \cdots \alpha_{ro}$ are positive, being the linear coefficients of the autosynthetic components, while the terms $\alpha_{r+1,o} \cdots \alpha_{no}$ are negative. We interpret a non-specific destructive influence as a reduction of all these linear coefficients by an amount δ, say, which is small enough so that there is still a positive steady state solution of the disturbed system, the binary coefficients remaining unaltered. Since all the terms α_{ij} of the first r equations in (1) are negative, a reduction of the terms α_{io} (i=1,2 .....r) by a small amount δ must result in a decrease of at least one of the steady state values (q_{r+1}, ...q_n). On the other hand, because the terms α_{ij} of the last (n-r) equations in (1) are positive, while the α_{io}'s are negative, a decrease in the α_{io}'s (i=r+1, ...n) must result in an increase in at least one of the steady state values (q_1, q_2, ...,q_r). Under these conditions, then, a destructive influence will shift the steady state of the epigenetic system towards an increase in the stationary concentrations of the autosynthetic components, and a decrease in the stationary concentrations of the heterosynthetic components. This result was demonstrated by Volterra (1931).

It seems likely that a similar though weaker result might be obtained if the conditions assumed above for the rigid division of auto- and heterosynthetic systems are relaxed. However, such a demonstration is difficult and has not yet been obtained.

An example in cell behaviour which suggests this type of response to non-specific destructive stimuli is to be found in the effect of temperature shocks upon the division process in unicellular organisms and in cells in tissue culture (Lark and Maaløe, 1954; Chèvremont-Comhaire and Chèvremont, 1956). Sublethal shocks of heat or cold affect the division process in such cells much more than they do the processes of growth, somehow undoing

the preparations a cell may have made for division, but not, apparently, interfering with cellular activities not specifically related to division. In our terminology, temperature shocks affect heterosynthetic activities much more than they do autosynthetic activities, and bring about a shift in the physiological state of a cell away from heterosynthesis. It is usually assumed, in explanation of this phenomenon, that the division process involves a mechanism with a much higher temperature-sensitivity than the normal growth process (Swann, 1957). However, in view of our above result, it is possible to interpret this phenomenon as a characteristic response of the epigenetic system to a non-specific destructive influence, which tends to shift the steady state in the direction of decreased heterosynthetic activity and increased autosynthetic activity. It is in general true that cells change from more differentiated to less differentiated states when placed in unnatural environments as, for example, cells which are removed from an organism and grown in culture (Willmer, 1954). Such behaviour tends to protect the more fundamental activities of cells at the expense of more specialised ones; i.e., this behaviour protects the autosynthetic system at the expense of heterosynthetic processes, and may find an explanation in terms of the system response demonstrated above.

II. Effect of a specific stimulus on the epigenetic system, acting through the binary coefficients of a single component.

Suppose that the binary coefficients, α_{rj} (r fixed, j=1,2,.....n) are all reduced in absolute value by the same amount, their new values being given by $c\alpha_{rj}$ ($0 < c < 1$). We want to show that the result of this parametric change is an increase in q_r, the steady state concentration of

component r, while the other components are affected non-specifically and may increase or decrease in steady state concentrations by a small amount.

Starting with the system defined by the steady state equation

$$Aq = p,$$

we can write the individual steady state solutions in the form

$$q_i = \frac{1}{|A|}\sum_j A_{ij} p_j \quad ,$$ where $|A|$ is the determinant of the matrix A, and A_{ij} is the co-factor of α_{ji} in A. If now the coefficients α_{rj} are changed to $c\alpha_{rj}$ (r fixed, j=1,2,.....n), then the coefficients α_{jr} are also changed to $c\alpha_{jr}$, since $\alpha_{rj} = -\alpha_{jr}$. The determinant of the new matrix is now $c^2 A$, since a row and a column of the old matrix have been multiplied by c. Furthermore, the new co-factors A'_{ij} $(i,j, \neq r)$ are $c^2 A_{ij}$, while the new co-factors A'_{rj} are cA_{rj}. The new steady state values are then (for $i \neq r$)

$$q_i' = \frac{1}{c^2|A|}\sum_{j=1}^{n} A'_{ij} p_j = \frac{1}{|A|}\sum_{j\neq r} A_{ij} p_j + \frac{A_{ir} p_r}{c|A|}$$

and

$$q_r' = \frac{1}{c^2|A|}\sum_{j=1}^{n} A'_{rj} p_j = \frac{1}{c|A|}\sum_{j=1}^{n} A_{rj} p_j = \frac{1}{c} q_r$$

Thus $q_r/q'_r = c < 1$, while the new values q'_i are altered through the quantity $\frac{A_{ir}p_r}{c\,A}$, which is in general small and may be positive or negative, depending upon the signs of A_{ir} and p_r.

APPENDIX 5

The Gibbs phase integral is defined by

$$Z \equiv \int e^{-\alpha G} dx = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{-\alpha \sum_i \tau_i (e^{x_i} - x_i)} dx_1 \cdots dx_n$$

This integral breaks up into a product of integrals of the form

$$\int_{-\infty}^{\infty} e^{-\alpha \tau_i (e^{x_i} - x_i)} dx_i \qquad (i = 1,2, \ldots\ldots n.)$$

These integrals are readily evaluated. Writing first $\xi_i = e^{x_i}$, the integral takes the form

$$\int_0^{\infty} e^{-\alpha \tau_i (\xi_i - \log \xi_i)} \frac{d\xi_i}{\xi_i} = \int_0^{\infty} \xi_i^{\alpha \tau_i - 1} e^{-\alpha \tau_i \xi_i} d\xi_i$$

Now writing $t_i = \alpha \tau_i \xi_i$, we get

$$(\alpha \tau_i)^{-\alpha \tau_i} \int_0^{\infty} t_i^{\alpha \tau_i - 1} e^{-t_i} dt_i = (\alpha \tau_i)^{-\alpha \tau_i} \Gamma(\alpha \tau_i)$$

The phase integral is now the product of these expressions,

$$Z = \prod_{i=1}^{n} (\alpha \tau_i)^{-\alpha \tau_i} \Gamma(\alpha \tau_i)$$

ACKNOWLEDGMENTS.

It is with much gratitude that I thank Professor C. H. Waddington for providing the opportunity and the facilities for this study, and for his encouragement in the pursuit of a theoretical treatment of the developmental process.

I wish also to express my indebtedness to Dr. H. Kacser for discussions which clarified and added precision to the arguments developed in this work.

The author received financial assistance from the National Research Council of Canada during the course of these studies, and the generosity of this council is gratefully acknowledged.

BIBLIOGRAPHY

Ashby, W. R. (1952), Design for a Brain. Chapman and Hall Ltd., London.

" (1958), "The Mechanism of Habituation," paper delivered at Symposium, The Mechanisation of Thought Processes, National Physical Laboratory, Teddington, Middlesex.

Bateson, W. (1894), Materials for the Study of Variation. Macmillan, London.

Beale, G. H. (1954), The Genetics of Paramecium aurelia. Cambridge University Press, London and New York.

Bradley, D. F., and M. Calvin (1956), General Systems: Yearbook of the Society for General Systems Research. "Behaviour: imbalance in a network of chemical transformations," 1, 56.

Brachet, J. (1957), Biochemical Cytology. Academic Press, New York.

" (1960), The Biochemistry of Development. Pergamon Press.

Brachet, J., H. Chantrenne, and F. Vanderhaeghe (1955), "Recherches sur les intéractions biochimiques entre le noyau et le cytoplasme chez les organismes unicellulaires." Biochim. Biophys. Acta, 18, 544.

Briggs, R., and T. J. King (1955), in Biological Specificity and Growth, ed. E. Butler. Princeton University Press, p. 207.

Brown, G. B., and P. M. Roll (1955), in The Nucleic Acids, ed. E. Chargaff and J. N. Davidson. Academic Press, New York, 2, 341.

Chèvremont-Comhaire, S., and M. Chèvremont (1956), "Action de températures subnormales suivies de rechauffement sur l'activité mitotique en culture de tissus. Compt. rend. Soc. de biol. 150, 1046.

Child, C. M. (1941), Patterns and Problems of Development. Chicago University Press.

Clayton, R. M. (1954), "Localisation of embryonic antigens by antisera labelled with fluorescent dyes," Nature, 174, 1059.

Cowie, D. B., and F. T. McClure (1959), "Metabolic pools and the synthesis of macromolecules." Biochim. Biophys. Acta, 31, 236.

Dalgleish, C. E. (1957), "Time factors in protein biosynthesis." Science, 125, 271.

Daly, H. M., V. G. Allfrey, and A. E. Mirsky (1952), "Uptake of glycine-N^{15} by components of cell nuclei. J. Gen. Physiol. 36, 173.

Dalcq, A. M. (1957), Introduction to General Embryology. Oxford University Press.

Dintzis, H. M., H. Borsook, and T. Vinograd (1958), "The synthesis of haemoglobin by microsomes from rabbit reticulocytes," in Microsomal Particles and Protein Synthesis, ed. R. M. Roberts. Pergamon Press, New York.

Dubnoff, J. W. (1955), in Symposium on Amino Acid Metabolism, ed. W. D. McElroy and B. Glass. Johns Hopkins University Press.

Dubos, R. S. (1940), "The adaptive production of enzymes by bacteria." Bact. Rev. 4, 1.

Ebert, J. D. (1950),"An analysis of the effects of anti-organ sera on the development, in vitro, of the early chick blastoderm." J. exp. Zool. 115, 351.

Gerard, R. W., C. Kluckholn, and A. Rapoport (1956), "Biological and cultural evolution: some analogies and explanations." Behav.Sci., 1, 6.

Grobstein, C., and E. Zwilling (1953), "Modification of growth and differentiation of chorio-allantoic grafts of chick blastoderm pieces after cultivation at a glass-clot interface," J. exp. Zool. 122, 259.

Halvorson, H. O. (1960), "The induced synthesis of proteins," Adv. Enzymol. 22, 99.

Hartmann, M. (1918), "Uber den Einfluss der Temperatur auf Grosse und Beschaffenheit von Zelle und Kern im Zusammenhange mit der Beeinflussung von Funktion, Wachstum, und Differenzierung der Zellen und Organe," Arch. Entwick. Org. 44, 114.

Herrmann, H. (1959), "The embryonic cell as a protein forming system," in 4th Int. Cong. Biochem., 6, 171.

Hinshelwood, C. N. (1953), "Autosynthesis," J. Chem. Soc. pt.II, 1947.

Holtfreter, J. (1938), "Veränderungen der Reaktionsweise im alternden, isolierten Gastrulacktoderm," Arch. Entwick.Org. 138, 163.

Jacob, F., P. Schaeffer, and E. L. Wollman (1960), in 10th Symposium of Soc. for Gen. Microbiol., "Microbial Genetics," Cambridge University Press, p. 67.

Johnen, A. G. (1956), "Temporal relationships in the induction process," Proc. Acad. Sci. Amst., Ser. C, 59, 652.

Kacser, H. (1957), Appendix in C. H. Waddington's Strategy of the Genes. Allen and Unwin, London.

" (1960), "Kinetic models of development and heredity," Symp.Soc. exp. Biol. 14, 13.

Karasaki, S. (1959), "Electron microscope studies on the amphibian embryo," Embryologia, 4, 247.

Kavanau, J. L. (1954), "Amino acid metabolism in the early development of the sea urchin, Paracentrotus lividus," Exptl. Cell Res., 7, 530.

Kerner, E. H. (1957), "A statistical mechanics of interacting biological species," Bull. Math. Biophys. 19, 121.

" (1959), "Further considerations on the statistical mechanics of biological associations. Bull. Math. Biophys. 21, 217.

Khinchin, A. I. (1949), The Mathematical Foundations of Statistical Mechanics. Dover Publications, N. Y.

Kovach, L. D. (1960), "Life can be SO nonlinear," Aer. Sci. 48, 218.

Lark, K. G., and O. Maaløe (1954), "The induction of cellular and nuclear division in Salmonella typhimurium by means of temperature shifts," Biochim. Biophys. Acta, 15, 345.

Loftfield, R. B., and E. A. Eignar (1958), "The time required for the synthesis of a ferritin molecule in rat liver," J. biol. Chem. 231, 925.

Maaløe, O. (1960), "The nucleic acids and control of bacterial growth," in 10th Symp. of Soc. for Gen. Microbiol., Microbial Genetics, Cambridge University Press, p. 272.

Maganasik, B. (1958), "The metabolic regulation of purine interconversion and of histidine biosynthesis," in The Chemical Basis of Development, ed. W. D. McElroy and B. Glass, p. 485.

Mandelbrot, B. (1954), "Simple games of strategy occurring in communication through natural languages," Trans. Inst. Rad. Eng., P.G.I.T., 3, 124.

" (1956), "The language of taxonomy; systems with Willis (natural) structure" in Symposium on Information Theory. Butterworth, London, p. 135.

Maddelstam, J. (1960), "The intracellular turnover of protein and nucleic acids and its role in biochemical differentiation," Bact. Rev. 24, 289.

McQuillan, K., R. P. Roberts, and R. J. Britten (1959), "Synthesis of nascent protein by ribosomes in E. coli.," Proc. nat. Acad.Sci;Wash. 45, 1437.

Milsum, J. H. (1960), "Adaptive control of processes by an economic criterion," paper presented to 1st. Cong. of Int. Fed. of Automatic Control, Moscow.

Mitchison, J. M. (1957), "The growth of single cells," Exptl. Cell Res. 13, 244.

Mucibabic, S. (1956), "The effect of temperature on the growth of Chilomonas paramecium," J. exp. Biol. 33, 627.

Mulnard, J. (1955), "Les phosphomonoesterases acid et alcaline dans le développement du rat et de la souris," Arch. Biol. 66, 525.

Needham, J. (1951), Biochemistry and Morphogenesis. Cambridge University Press.

Nieuwkoop, P. D. (1955), "Independent and dependent development in the formation of the central nervous system in amphibians," Exp. Cell Res. Suppl., 3, 262.

Pardee, A. B., and L. S. Prestidge (1956), "The dependence of nucleic acid synthesis on the presence of amino acids in Escherichia coli," J. Bact. 71, 677.

Paul, J. (1958), Discussion in 4th Int. Cong. Biochem., 6, 14.

Pavan, C. (1954), "Salivary chromosome and differentiation," Proc. 9th Int. Cong. Genet., p.778.

Pringle, J. W. S. (1951), "On the parallel between learning and evolution," Behaviour, 3, 174.

Ranzi, S. (1958), "Electron microscopy of cells of the young amphibian embryo," Arch. Embryol. Morphol. exp. 2, 102.

Raven, C. P. (1938), "Der Potenzbegriff in der Entwicklungsmechanick," Acta Biother. 4, 51.

Rosen, R. (1959), "A relational theory of biological systems, " Bull. Math. Biophys. 21, 109.

Rugh, R. (1950), The Frog: Its Reproduction and Development. Blakiston, Philadelphia and Toronto.

Russell, W. M. S. (1958), "Evolutionary concepts in behavioural science," General Systems, 3, 18.

Sands, M. K. and R. B. Roberts (1952), "The effects of a tryptophane histidine deficiency in a mutant of E. coli," J. Bact. 63, 505.

Schaecter, M., O. Maaløe, and N. O. Kjeldgaard (1958), "Dependency on medium and temperature of cell size and chemical composition during balanced growth of Salmonella typhimurium, J. Gen. Microbiol. 19, 592.

Shen, S. C. (1955), in Biological Specificity and Growth, ed. E. G. Butler, Princeton University Press, p. 73.

Spiegelman, S. (1948), "Differentiation as the controlled production of unique enzymatic patterns," Symp. Soc. exp. Biol. 2, 286.

Spiegelman, S., and H. O. Halvorson (1953), in Adaptation in Microorganisms, ed. E. F. Gale and R. Davis. Cambridge University Press, p. 98.

Stearns, R. N., and A. B. Kostellow (1958), in The Chemical Basis of Development, ed. W. D. McElroy and B. Glass. Johns Hopkins University Press, p. 448.

Swann, M. M. (1957), "The control of cell division: A review." I. General Mechanisms," Cancer Res. 17, 727.

Toivonen, S. (1958), "The dependence of the transformation of competent ectoderm on temporal relationships in the induction process," J. Embryol. exp. Morphol. 6, 479.

Tsien, H. S. (1954), Engineering Cybernetics, McGraw Hill Book Co., N. Y.

Volterra, V. (1931), Leçons sur la Théorie Mathématique de la Lutte pour la Vie. Gautiers-Villars et Cie., Paris.

Waddington, C. H. (1940), Organisers and Genes, Cambridge University Press.

" (1948), "The genetic control of development," Symp. Soc. exp. Biol. 2, 145.

" (1954), "The integration of gene-controlled processes and its bearing on evolution," Proc. 9th, Cong. Gen. (Caryologia Suppl.), p. 232.

" (1956), Principles of Embryology, Allen and Unwin, London.

" (1957), The Strategy of the Genes, Allen and Unwin, London.

Waddington, C. H., and T. Yao (1950), "Studies on regional specificity in the amphibian organiser centre," J. exp. Biol. 27, 126.

Weiss, P. (1939), The Principles of Development, Holt, N. Y.

" (1950), "Perspectives in the field of morphogenesis," Quart. Rev. Biol. 25, 177.

Weiss, P. and J. L. Kavanau (1957), "A model of growth and growth control in mathematical terms," J. Gen. Physiol. 41, 1.

Whitehead, A. N. (1929), Process and Reality, Macmillan and Co., N. Y.

Willis, J. C. (1922), Age and Area, Cambridge University Press.

" (1940), The Course of Evolution, Cambridge University Press.

Willmer, E. N. (1953), Tissue Culture, Methuen and Co. Ltd., London.

Williamson, M. H. (1957), "An elementary theory of interspecific competition," Nature, 180, 422.

Ycas, M. (1959), in Symposium on Information Theory in Biology, ed. H. P. Yockey, A. C. Platzman, and H. Quastler. Pergamon Press, N. Y., p. 70.

Yule, G. U. (1924), "On a distribution of J. C. Willis for genus-species relations," Phil. Trans. B, 213, 21.

Zalokar, M. (1961), "Kinetics of amino acid uptake and protein synthesis in Neurospora," Biochim. Biophys. Acta. 46, 423.

Zipf, G. K. (1949), Human Behaviour. Addison Wesley, Cambridge (Mass.)