Physical Chemistry of Biological Morphogenesis

By L. G. Harrison
DEPARTMENT OF CHEMISTRY, UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER, B.C., CANADA V6T 1Y6

1 Introduction

Morphogenesis is the creation of a complicated shape out of a simpler one by chemical processes in living organisms. To the physical scientist, the essence of it is symmetry-breaking. Yet the precept that 'asymmetry begets asymmetry' is not necessarily violated. Natural disturbances, which are continually present everywhere, contain adequate asymmetry to serve as antecedent for any shape, however complex. But how do living organisms go about making a rather precise selection of what parts of the available asymmetry to amplify? We know that the specification for the development of an organism is written in the genetic code. For some thirty years, the most rapidly advancing divisions of biological science have been those dealing with DNA, RNA, and proteins. The powerful theories here have been those of molecular geometry, statically perceived in terms of building block or jigsaw puzzle fitting together of parts.

Genetic information does not, however, specify how much of a protein is to be produced at any time, nor where it is to go. No-one today believes that the nucleus contains a reduced spatial map from which the organism is built, as a building is constructed from an architect's plans, or as a human sperm was once believed from vague microscopical resemblance to contain a 'homunculus'. There is a wide gap in understanding between macromolecular geometry and the large-scale shape of the whole organism, and this gap is not being closed very quickly. This review is intended, therefore, to show physical chemists a field in which experiment and theory are rather far apart, and in which they might be able to make some contribution towards bringing them closer together.

All physical chemistry is divided into three parts: structure, equilibrium, and kinetics. In approaching a set of phenomena for which there are no definitely proved or generally accepted explanations, the physical chemist must ask himself in which of these fields the answers are likely to lie. Two accounts with almost identical titles have no overlap whatever in outlook. These are a paper written in 1952 by Turing,[1] entitled 'The Chemical Basis of Morphogenesis', and a chapter of Lehninger's 'Biochemistry',[2] entitled 'The Molecular Basis of Morphogenesis'. The former expounds a model, in terms of rate equations for reaction and diffusion, showing how spatially inhomogeneous arrangements of material might be established and maintained kinetically in a system in which the equilibrium state is a homogeneous uniform distribution. This type of theory has later been

[1] A. M. Turing, *Philos. Trans. R. Soc. London, Ser. B*, 1952, **237**, 37.
[2] A. L. Lehninger, 'Biochemistry', (2nd Edn.), Worth Publishers, 1975, chap. 36.

extensively elaborated in the theory of 'dissipative structures' of Prigogine and his collaborators,[3–5] and more specifically tied to biological examples in the work of Gierer and Meinhardt.[6–8]

The structural approach[2] is based on precise evidence for the geometry of assemblies of quite large numbers of protein molecules in viruses,[2,9] microtubules,[10] and microfilaments.[11] Its successes stop short of the cellular level. Studies of microstructure with the aid of optical and electron microscopes are, however, providing increasing evidence that microtubules and microfilaments are often present at places and times of morphogenetic significance, especially at the onset of cell division in both plants ('pre-prophase band' of microtubules[12]) and animals (contractile ring of microfilaments[13]). Many biologists expect that the 'cytoskeleton' of these structural proteins will prove to be the geometrical link between the molecular and the macroscopic scale.

Living organisms are never at equilibrium, and are probably farthest away from it when undergoing morphogenesis. Nevertheless, there may be cases in which the appearance of disparate parts is to be accounted for principally in terms of a drive towards minimum free energy, rather than in terms of geometrical fitting, or imbalances in rates of reaction and transport processes. There is then some correspondence between the process concerned and phase transitions. In respect of some particular kinds of interaction, the 'equilibrium' state is heterogeneous. Classical theories of phase nucleation, critical supersaturation, critical micelle concentration, and spinodal decomposition have been used in discussion of pattern formation.[14,15] The behaviour of mixtures of cells sometimes mimics molecular mixtures in ways that suggest a concept of cell-as-molecule. Such mixtures frequently 'sort out' into two separate aggregates, each of a single type of cell. Steinberg[16] showed that this phenomenon could be explained in some detail by differential adhesions between cells. This leads to an analogue of surface tension, measurable by the sessile drop method,[17] for an assembly of

[3] I. Prigogine, in 'Fast Reactions and Primary Processes in Chemical Kinetics', 5th Nobel Symposium, ed. S. Claesson, Interscience, New York, 1967, p. 371.
[4] P. Glansdorff and I. Prigogine, 'Thermodynamics of Structure, Stability, and Fluctuations', Wiley-Interscience, New York, 1971.
[5] G. Nicolis and I. Prigogine, 'Self-Organization in Non-equilibrium Systems', Wiley, New York, 1977.
[6] A. Gierer and H. Meinhardt, *Kybernetik*, 1972, **12**, 30.
[7] H. Meinhardt, *J. Cell Sci.*, 1977, **23**, 117.
[8] A. Gierer, *Prog. Biophys. Mol. Biol.*, 1981, **37**, 1.
[9] A. Klug, *Fed. Proc., Fed. Am. Soc. Exp. Biol.*, 1972, **31**, 40.
[10] 'Microtubules', ed. K. Roberts and J. S. Hyams, Academic Press, 1979.
[11] 'Cell Motility: Molecules and Organization', ed. S. Hatano, H. Ishikawa, and H. Sato, University Park Press, Baltimore, 1979.
[12] J. D. Pickett-Heaps and D. H. Northcote, *J. Cell Sci.*, 1966, **1**, 109.
[13] R. Rappaport, *Int. Rev. Cytol.*, 1971, **31**, 169; T. Schroeder, *Z. Zellforsch.*, 1970, **109**, 431, *J. Cell Biol.*, 1972, **53**, 419, *Proc. Natl. Acad. Sci. USA*, 1973, **70**, 1688; R. E. Kane, ref. 11, p. 639.
[14] T. C. Lacalli and L. G. Harrison, *J. Theor. Biol.*, 1978, **74**, 109.
[15] J. W. Cahn, *Acta Metallurg.*, 1961, **9**, 795; *J. Chem. Phys.*, 1965, **42**, 93.
[16] M. S. Steinberg, *J. Exp. Zool.*, 1970, **173**, 395.
[17] H. M. Phillips and M. S. Steinberg, *Proc. Natl. Acad. Sci. USA*, 1969, **64**, 121.

cells, and is in general closely analogous to classical concepts of solution thermodynamics and immiscibility.

The term 'self-assembly' has commonly specified the structural approach, but is nowadays being extended to include sorting-out by differential adhesion. The dangerously similar term 'self-organization' is used by Prigogine[5] for the kinetic approach. It is perhaps least confusing to avoid these terms, unless one uses them to cover the whole topic, regardless of mechanism.

In the 1930's, a term 'morphogenetic substance' was current, and it later became clear[18] that what was then envisaged is now known as messenger RNA, and that it does not solve the problems of morphogenesis. From the 1950's onwards, the term 'morphogen' has been used, sometimes rather vaguely. Its most specific meaning is to specify either of two substances, an 'activator' and an 'inhibitor', commonly designated X and Y, which appear in Turing's model[1] and most later versions of reaction-diffusion. Candidates for the style and title of morphogen range from ammonia *via* cyclic adenosine monophosphate (cAMP) to some rather large proteins and glycoproteins. In non-living chemical systems, cerium-catalysed oxidation of malonate by bromate (Belousov–Zhabotinski reaction) shows time and space periodicities, for which the analogues of three morphogens 'X, Y, and Z' are probably BrO_2^-, Ce^{4+}, and Br^-.[5,19–21] No substance is yet definitely established as a morphogen in a living system.

2 The Kinetic Approach: Reaction-diffusion

A. Symmetry-breaking: Optical Resolution as an Example.—Reaction diffusion mechanisms for morphogenesis, when expressed as rate equations, commonly contain simple autocatalytic terms [equation (1)] where X represents a displacement from equilibrium and may therefore assume both positive and negative values. Hypothetical chemical mechanisms, such as Prigogine's 'Brusselator'[3–5] commonly show a bimolecular autocatalytic step [equation (2)], assumed to have orders corresponding to molecularities. The essence of this concept of autocatalysis was first mentioned, and is most simply illustrated, in regard to a different but closely related problem, the origin of optical activity in nature, *e.g.* why living material contains, in general, only L-amino-acids and D-sugars.

$$\partial X/\partial t = kX \qquad (1)$$

$$2X + Y = 3X \qquad (2)$$

Mills,[22] in a presidential address on inorganic stereochemistry to a regional meeting of the British Association in 1932, pointed out that, if a reaction in which a prochiral reactant A yields a chiral product D or L is autocatalysed bimolecularly and stereospecifically [equation (3)], then the racemic state is unstable.

[18] J. Brachet, preface to S. Puiseux-Dao, 'Acetabularia and Cell Biology', transl. P. Malpoix-Higgins, Springer-Verlag, New York Inc., 1970.
[19] A. M. Zhabotinski, *Biofizika*, 1964, **9**, 306.
[20] A. T. Winfree, *Science*, 1972, **175**, 634; *Sci. Am.*, 1974, **230**, 82.
[21] R. J. Field and R. M. Noyes, *Faraday Symp. Chem. Soc.*, 1974, **9**, 21; J. D. Murray, *J. Theor. Biol.*, 1976, **56**, 329.
[22] W. H. Mills, *Chem. Ind.* (*London*), 1932, 750.

$$A \xrightarrow{2D} D \text{ and } A \xrightarrow{2L} L \ , \qquad (3)$$

If, for example, the D:L ratio is at any time 2:1, the rates of formation of D and L (with order as expected from molecularity) are in the ratio 4:1 and the system is moving closer to resolution as D. Mills showed that fluctuations could be expected to be present in adequate amount to start a system on the road to resolution. Mills' concept has been rediscovered or elaborated several times.[23–27] At least two accounts[23,26] indicated that an intermediate stage should be an assemblage of separate territories, each occupied by pure D or pure L. This is a rudimentary morphology, but an unstable one, in that large areas will ultimately surround and destroy smaller ones.

If we now use X to represent the measure of asymmetry [equation (4)], and not

$$X = D - L \qquad (4)$$

the concentration of a single substance [as one might do in relation to equation (2)], the rate of growth of asymmetry is:

$$\partial X/\partial t = k_f AD^2 - k_f AL^2 = k_f A(D - L)(D + L) = k_f A(D + L)X \qquad (5)$$

There are many circumstances, especially when X is small, in which the total product (D + L) is likely to be varying much more slowly than the asymmetry X. Equation (5) is then, approximately, equivalent to equation (1); bimolecular autocatalysis gives first-order exponential growth of asymmetry. The effective rate parameter is, however, a pseudo rate-constant, containing both the total reactant concentration A and the total product concentration (D + L), both assumed to be held roughly constant by external supply and removal.

Figure 1 shows a more complete model of this kind. Catalysis takes place by adsorption of D and L on to sites on a catalytic surface. Arrangement of D and L is random, site activation is bimolecular and stereospecific, there are sufficient concentrations of D and L in solution to maintain adsorption saturation, and adsorption–desorption is rapid. Autocatalysis of D production then depends on D and L through the ratio D/(D + L). $\mathscr{D}$ is diffusivity along a co-ordinate s.

$$\partial D/\partial t = k_f AD^2/(D + L)^2 - k_r D^3/(D + L)^2 - k_{ext} D + \mathscr{D}\partial^2 D/\partial s^2 \qquad (6)$$

The equation is similar for $\partial L/\partial t$, with D and L interchanged throughout. This model introduces the reverse reaction on the catalyst, an important feature neglected in Mills' original suggestion. Since this is a cubic term, it is capable of preventing the asymmetrising effect of the squared term for the forward reaction. The symmetry-breaking is, indeed, made possible only because the non-stereospecific removal, giving the linear term k_{ext} D, opposes the effect of the cubic term. I showed[27] that, in a well-stirred system with the diffusion term omitted from equation (6), symmetry-breaking will occur from the racemic steady state only if:

[23] F. C. Frank, *Biochim. Biophys. Acta*, 1953, **11**, 459.
[24] F. F. Seelig, *J. Theor. Biol.*, 1971, **31**, 355; 1971, **32**, 93.
[25] P. Decker, *Nature (London), New Biol.*, 1973, **241**, 72.
[26] L. G. Harrison, *J. Theor. Biol.*, 1973, **39**, 333.
[27] L. G. Harrison, *J. Mol. Evol.*, 1974, **4**, 99.

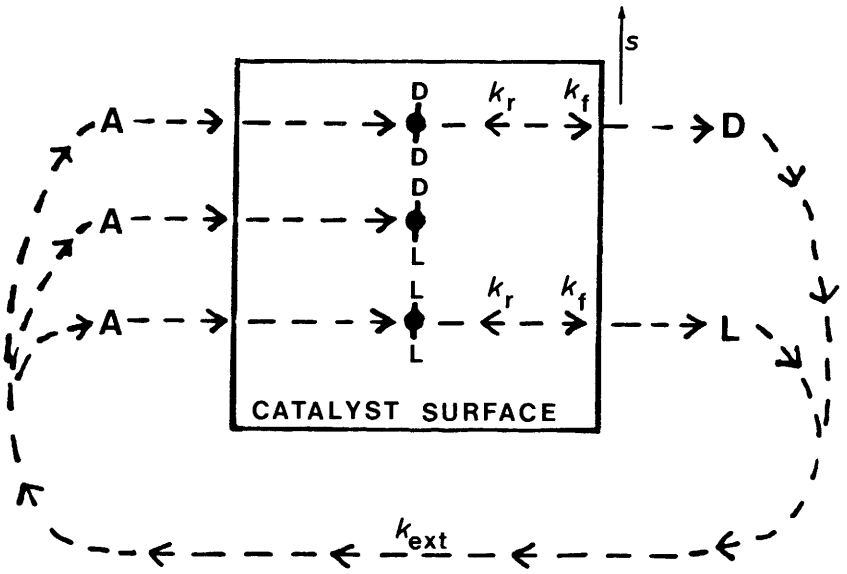


Figure 1 *The simplest irreversible (dissipative) cycle for symmetry breaking by bimolecular autocatalysis. When diffusion is considered, the system is usually envisaged as essentially one-dimensional, elongated along the direction s*

$$k_{ext} > k_r/4 \tag{7}$$

If we put in the diffusion term, and examine what happens if a simple sinusoidal pattern of asymmetry is superimposed on the racemic steady state, then we find

$$X = a(t)\sin(2\pi s/\lambda) \tag{8}$$

that for long wavelengths the amplitude a increases with time, while for short wavelengths it decreases. The threshold wavelength for growth is:[28]

$$\lambda_0 = 2\pi\mathscr{D}^{\frac{1}{2}}/(k_{ext} - k_r/4)^{\frac{1}{2}} \tag{9}$$

This simple example illustrates many of the important general features of kinetic mechanisms for the generation of spatial pattern:

(*a*) A threshold rate of interference from the rest of the universe must be exceeded for symmetry-breaking to occur. The term k_{ext} is a measure of the rate of entropy increase, or dissipation of energy, in the surroundings, necessary to create and maintain a low-entropy ordered state in the system. Hence the Prigogine school use the term 'dissipative structure' for what I prefer to call 'kinetically maintained structure'.

(*b*) Reaction-diffusion mechanisms can set up quantitative scales of distance. This is seen in the dimensionality of $(\mathscr{D}/k)^{\frac{1}{2}}$, where k is a first-order

[28] L. G. Harrison, in 'Origins of Optical Activity in Nature', ed. D. C. Walker, Elsevier, 1979, chap. 10.

constant. This form is exemplified, for a threshold spacing, in equation (9). More complex expressions for 'chemical wavelengths', *e.g.* the Turing wavelength given in equation (18) below, are variations on the same theme. This quantitative aspect is a particular strength of the reaction-diffusion mechanism that is not easily matched in structural or equilibrium models.

(*c*) Reaction-diffusion tends to amplify long-range order. For example, the simplest possible autocatalysis-diffusion rate equation is:

$$\partial X/\partial t = kX + \mathscr{D}\partial^2 X/ds^2 \tag{10}$$

If we put into this the sinusoidal disturbance (8), we find that X grows according to a simple exponential growth law:

$$X = X_0 \exp(k_g t) \tag{11}$$

where

$$k_g = k - \mathscr{D}(4\pi^2/\lambda^2) \tag{12}$$

Figure 2(a) shows how k_g increases with wavelength. If we use the more complicated model of equation (6), start from a small disturbance to the spatially uniform racemic steady state, and confine attention to the early stages of growth or decay of the disturbance, this is again exponential, and gives equation (13),

$$k_g = (k_{ext} - k_r/4) - \mathscr{D}(4\pi^2/\lambda^2) \tag{13}$$

which corresponds to Figure 2(a) with a different value for the threshold. This monotonic increase explains the eventual dominance of one product over the other, as in spontaneous optical resolution. Formation of a stable pattern requires a different dependence, *e.g.* as in Figure 2(b), in which some finite-wavelength disturbance grows fastest and hence is eventually selected by the system out of whatever nature happens to provide at all wavelengths. Turing's model, described in Section 2B, is designed to produce this dependence of k_g on wavelength.

Spontaneous resolution has been discussed here as an analogy, which provides the quickest route to understanding some basic features of the kinetic approach to all kinds of symmetry breaking. Two questions more directly related to optical asymmetry immediately arise, however. First, since the kinetic discussion shows the possibility, in a non-living chemical system, of spontaneous resolution in converting a prochiral substrate into a chiral product, without any preliminary doping of the catalyst with optically asymmetric material, has this ever been observed? In 1958, there was a claim[29] that such spontaneous resolution had been found in the reduction of α-ketoglutaric acid and its oxime to α-hydroxyglutaric acid and glutamic acid by hydrogenation on Raney nickel [equation (14)].

[29] T. Isoda, A. Ichikawa, and T. Shimamoto, *Rikagaku Kenkyusho Hokoku* (*J. Inst. Phys. Chem. Res., Tokyo*), 1958, **34**, 134 (In Japanese; translation by Y. Koga available from L. G. Harrison).

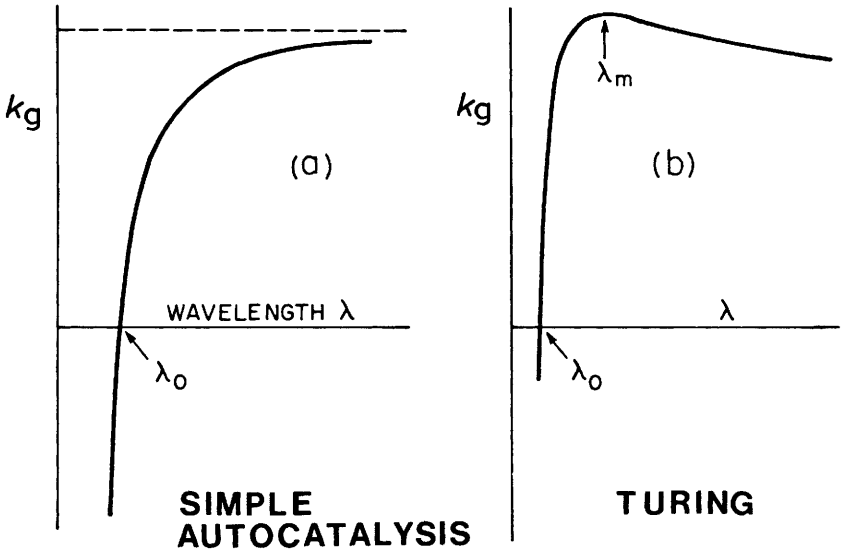
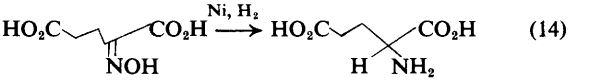


Figure 2 *Exponential growth rate constants for pattern* (i.e. *non-uniform concentration distribution*) *sinusoidal in position* s: (*a*) *for simple autocatalysis and diffusion of one substance, equation* (10)*; more rapid amplification of longer wavelengths;* (*b*) *for Turing's two-morphogen model, showing most rapid amplification at* λ_m. *For* (*a*), $\lambda_0 = 2\pi(/\mathscr{D}k)^{\frac{1}{2}}$
(Reproduced by permission from ref. 28)



(14)

I have tried to repeat this (unpublished; McGinnis, thesis[30]) and obtained equivocal data on resolution but definite indications of autocatalysis. This topic deserves further study on modern catalysts more sensitive to asymmetrising influences.

Second, the question arises of just how close any symmetry-breaking in living systems may be to this model. It is well known that small chiral molecules are usually found in only one enantiomeric form in living organisms, and that this asymmetry carries through to their assembly into structures as large as proteins, including the assembly of several protein subunits into a polymer. I have pointed out, however, that on the large scale both enantiomers are present in structures as obvious as our right and left hands, and I have made a 'wild surmise' that the smallest spatial scale at which right- and left-handed structures appear might be an important scale to consider, and might be not far above the scale of organisation of a few protein subunits into a polymer.[28,31]

B. Turing's Equations.—Turing's proposal[1] was principally in the form of linear

[30] M. J. McGinnis, M.Sc. thesis, University of British Columbia, 1977.
[31] L. G. Harrison and T. C. Lacalli, *Proc. R. Soc. London, Ser. B* 1978, **202**, 361.

differential rate equations involving two measures of displacement from equilibrium, X and Y. These are, for diffusion in one spatial dimension s:

$$\partial X/\partial t = k_1X + k_2Y + \mathscr{D}_X\partial^2X/\partial s^2 \quad (15a)$$

$$\partial Y/\partial t = k_3X + k_4Y + \mathscr{D}_Y\partial^2Y/\partial s^2 \quad (15b)$$

These equations, being linear, can be solved exactly in analytical form. The solutions are summations of real or complex exponentials in the time and distance variables. For them to represent morphogenetic behaviour, there are various restrictions on the values of the four rate constants and two diffusivities. The essence of the mode of operation of the Turing model has been described qualitatively by Maynard-Smith.[32] We suppose k_1 and k_3 to be positive, k_2 to be negative, and k_4 to be zero; also $\mathscr{D}_Y > \mathscr{D}_X$; *i.e.* X catalyses its own growth and that of Y, Y inhibits X, and the inhibitor diffuses faster than the activator (X). This implies that, from a localised centre of activity, Y will, at least in early stages of development, spread out further than X. The useful slogan 'short-range activation; long-range inhibition' is therefore often used to characterise mechanisms of the Turing type. But, in the full operation of a Turing model to give a precisely ordered pattern, both X and Y are periodically distributed over the whole region concerned.

In Maynard-Smith's illustration (Figure 3), we suppose that the system, an elongated one along one spatial dimension s, is initially at equilibrium through-

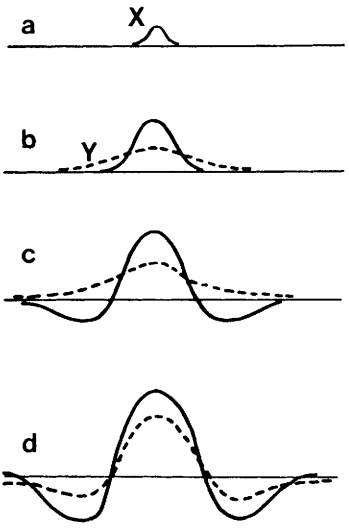


Figure 3 *Qualitative picture of how the Turing model leads to* X *and* Y *waves in phase, redrawn following Maynard-Smith's illustration*[32] *(schematic only; not computed). See text for explanation*

[32] J. Maynard-Smith, 'Mathematical Ideas in Biology', Cambridge University Press, 1968.

out with respect to both X and Y. Somewhere, a small and fairly localised positive X-disturbance is introduced [Figure 3(a)]. It grows (k_1 catalysis) and spreads ($\mathscr{D}_X$ diffusion), but a positive Y displacement grows within the X displacement and spreads faster (k_3 and $\mathscr{D}_Y$ terms) [Figure 3(b)]. Where Y has spread beyond the effective boundaries of the X peak, its inhibitory effect (k_2 term) produces X troughs [Figure 3(c)]. In these regions of negative X, the k_3 term becomes an inhibition of Y, *i.e.* where X has been 'pushed down' it can 'pull Y down after it' [Figure 3(d)]. The implication is that X and Y are on the way to settling down into wavelike patterns in phase with each other. In this model, we do not find an activator at one end of the system and an inhibitor at the other; X and Y have their maxima and minima together.

A convenient abbreviation of the mathematical treatment of Turing's equations[33] may be obtained by going directly to the apparent final state in Maynard–Smith's illustration, and taking initial X and Y disturbances to be sinusoidal in s, and exactly in phase or 180° out of phase, so that $\theta = Y/X$ is everywhere the same and depends only on time, $\theta = \theta(t)$. Insertion of such disturbances into equations (15) confirms the possibility of this type of solution by yielding an ordinary differential equation for the evolution of θ with time. Eventually, the X and Y waves will settle down to a constant ratio of amplitudes, and will grow or decay together in simple exponential fashion with a rate constant k_g. How k_g varies with wavelength depends on the values of the kinetic and diffusion parameters. A typical plot in the region of morphogenetic significance is shown in Figure 2(b). A maximum for k_g occurs at a finite wavelength; given sufficient time, and sufficient concentration range available for exponential growth of both X and Y on both sides of equilibrium, a pattern of this wavelength will finally dominate the system.

The growth rate constant (doubled, for convenience) is given by:

$$2k_g = k_1 + k_4 - (2\pi/\lambda)^2(\mathscr{D}_X + \mathscr{D}_Y) + (b^2 + 4k_2k_3)^{\frac{1}{2}} \qquad (16)$$

where

$$b = k_4 - k_1 - (2\pi/\lambda)^2(\mathscr{D}_Y - \mathscr{D}_X) \qquad (17)$$

The maximum of k_g is at:

$$\lambda_m = 2\pi(\mathscr{D}_Y - \mathscr{D}_X)^{\frac{1}{2}}/[k_4 - k_1 + (\mathscr{D}_Y + \mathscr{D}_X)(-k_2k_3/\mathscr{D}_X\mathscr{D}_Y)^{\frac{1}{2}}]^{\frac{1}{2}} \qquad (18)$$

which is dimensionally of the form $(\mathscr{D}/k)^{\frac{1}{2}}$ and in fact reduces for reasonable values of the parameters to $2\pi[\mathscr{D}_X\mathscr{D}_Y/(-k_2k_3)]^{\frac{1}{4}}$. Since k_2k_3 is negative in the region of interest, k_g may be complex, indicating time-oscillatory behaviour. When k_g is real and positive, spatial pattern can be established without any time-oscillatory behaviour. For fairly wide ranges of the kinetic constants and diffusivities, k_g varies with λ as shown in Figure 2(b); k_g is real at all finite positive wavelengths, and passes through a maximum. For the maximum to be sharp, with k_g going negative at long wavelengths, k_1 and k_4 must be chosen within a rather restricted range, and the latter must be negative (Y self-inhibitory).

[33] T. C. Lacalli and L. G. Harrison, *J. Theor. Biol.*, 1978, **70**, 273.

All the required conditions were given by Turing,[1] very briefly and without indication of derivation. Lacalli and I[33,34] gave simple derivations of these, and showed how the conditions could be plotted on a diagram of k_1 and k_4.

Turing's model has the obvious defect of unrestricted exponential growth, which must always become unrealistic after some finite time, because concentrations grow to exceed any possible value or, in the other direction, become negative. More realistic models give non-linear differential equations, often corresponding to practical limitations on concentrations, such as saturation effects. The strength of the Turing model is that, for the initial development of pattern out of uniformity, the mathematical process of linearisation casts most of these non-linear models back into the Turing form.[34] The Turing model therefore shows how pattern should start to grow out of uniformity for most reaction-diffusion mechanisms; but it does not give a full account of how the pattern may change as it develops, because of non-linearities.

Mathematical analysis of reaction-diffusion has so far been carried out mainly for fixed boundary conditions. Biological development is concerned largely with moving boundaries, because organisms grow. Experiments in this field often also involve changes in boundaries, *e.g.* when one chops an organism into pieces and watches what happens next. The term 'regulation' is used to describe the ability of an organism to maintain or re-establish pattern in the face of such natural or artificial changes in boundaries. Turing's model has been criticised[35,36] for lack of regulatory capacity. A major objection has been that the Turing model requires a precise fit between some multiple or half-multiple of the 'chemical wavelength' corresponding to maximum k_8 and the size of the system. The former is fixed by the dynamic constants of the mechanism; the latter is continually changing. This type of objection is invalid, for two reasons. First, even if the fit of wavelength to system size is inexact, a pattern can develop, and one particular pattern may arise for quite a wide range of system size. Figure 4 shows a computation of the development of a half-wave of X and Y along an elongated system. The calculations were carried out for possible relevance to the morphogenesis of slime moulds (Plate 4). Computation of the effect of cutting the specimen into two unequal pieces showed re-establishment of the pattern in both pieces. Amplitudes are, however, shown normalised in the diagrams. In this, as in all calculations on the linear equations, there is a great increase in amplitude as time goes by, and it is a valid objection that regulation usually takes enough time that the system must have moved into a region of concentrations in which one should be using more realistic non-linear models.

Second, as mentioned above in relation to equation (5), autocatalytic rate parameters may often turn out to be pseudo rate-constants, containing concealed concentrations that happen to be, in practice, constant, if one looks beyond the rate equations to the possible chemical mechanisms giving rise to them. In the

[34] T. C. Lacalli and L. G. Harrison, *J. Theor. Biol.*, 1979, **76**, 419.
[35] J. Bard and I. Lauder, *J. Theor. Biol.*, 1974, **45**, 501.
[36] C. H. Waddington, 'Principles of Embryology', George Allen and Unwin, London, 1956, pp. 422—423.

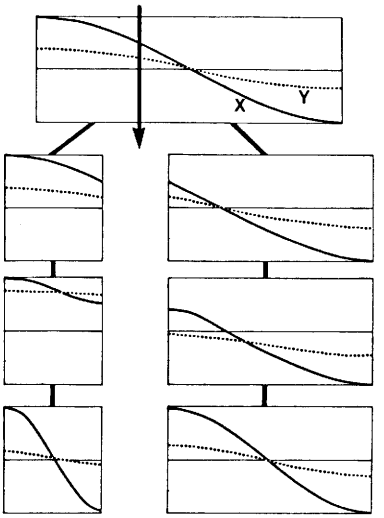


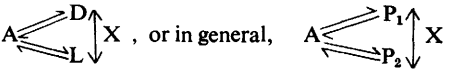
Figure 4 *Computation of restoration of pattern in both parts of an elongated system cut into two unequal pieces, on the basis of the Turing model, for comparison with what happens in slime moulds (Plate 4). Amplitudes are shown normalized; there is a very large increase as time goes on*

case of equation (5), if the total reaction product is P = D + L, the autocatalytic rate constant is k_fAP; for the adsorption saturation model on a heterogeneous catalyst [equation (6)], the corresponding expression is k_fA/P. If this type of expression is used for the constant k in equations (10) and (12), the threshold wavelength for growth of a disturbance is:

$$\lambda_0 = 2\pi(\mathscr{D}\mathrm{P}/k_f\mathrm{A})^{\frac{1}{2}} \qquad (19)$$

and is thus controlled by the reactant/product ratio A/P, which depends upon supply and removal and can certainly change as a system grows. It is, for example, quite easy to envisage circumstances[37] in which reactant is supplied to the bulk of a system across its boundaries and used up in such a way that its steady-state concentration falls as the system grows. Equation (19) then shows λ_0 increasing as the system grows. I have shown[31] that the same kind of argument can be made for the 'chemical wavelength' λ_m, and that this can in fact grow in proportion to the size of the system, so that a pattern can be stable for indefinite increase in size. The important point here seems to be that the reaction scheme should have two sequences in parallel starting from the same reactant A, with the 'morphogen variable' X or Y being a measure of imbalance between the concentrations produced along the two branches, *e.g.* equation (20).

[37] L. G. Harrison, in 'Developmental Order: Its Origin and Regulation', 40th Annual Symposium Society Developmental Biology, ed. S. Subtelny, Alan R. Liss Inc., 1981 (in press).



(20)

where the products P_1 and P_2 and the parameters of rate processes leading to them need not be symmetrically related as closely as enantiomers.

The model as so far described shows how non-uniform concentration distributions in space could arise for two hypothetical substances X and Y, or for some greater number of substances P_1, P_2, *etc.*, with the symbols X and Y serving for quantitative measures of imbalance. The question remains of how X and Y are supposed to influence morphogenesis. In single-celled plants, which have rigid cell walls and are capable of converting differential reaction rates into a complicated shape (Plates 2 and 3), it is reasonable to envisage X or Y as a catalyst for growth processes at the cell surface. Embryology of multicellular animals is discussed very differently, in terms of a variety of levels of gene expression achieved by something analogous to the throwing of switches to turn on or off the activity of particular genes in certain individual cells or groups of cells. Kauffman[38–40] has suggested, first, that the notation 1 or 0 could be used to indicate on or off for each switch, leading to a 'binary epigenetic code' in which, for example, the notation 0100 on a region of an embryo would mean that of four genes A, B, C, D, only B is switched on in that region. Second, he suggested that the function of a chemical morphogen is to throw a switch, *e.g.* positive X turns A on while negative X turns A off. Third, he proposed that when a morphogen system goes through a succession of patterns as a system grows, each pattern can serve to switch on or off a new gene in some definite sequence, and that each switching is permanent. If we consider, for instance, two-dimensional Turing patterns on a growing ellipse, with the rate parameters truly constant so that increasingly complex patterns are favoured as the ellipse becomes larger, the expected succession of patterns has nodal lines and positive and negative regions as shown in Figure 5. Kauffman envisaged each pattern as switching one gene in

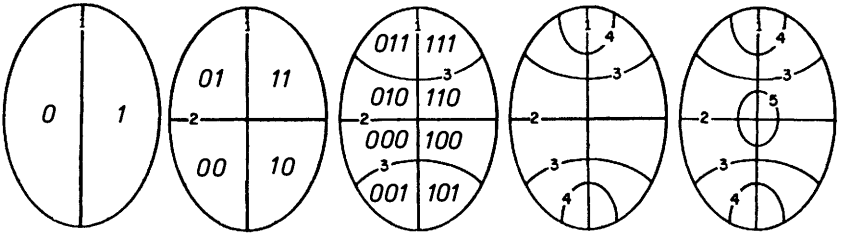


Figure 5 *Nodal lines of successive Turing waves on a growing ellipse (size shown normalized). At each stage, positive* X *switches a particular gene on* (1) *and negative* X *switches it off* (0). *The resulting binary epigenetic code is shown for the first three stages. Nodes* 1 *and* 2 *may appear to have been established in the wrong order; for a full account, see Kauffman,*[39,40] *This model is hypothetical*

[38] S. A. Kauffman, *Science*, 1973, **181**, 310; 'Cell Patterning', CIBA Foundation Symposia, new series, no. 29, Elsvier, Amsterdam, 1975, p. 201.
[39] S. A. Kauffman, *Am. Zool.*, 1977, **17**, 631.
[40] S. A. Kauffman, R. Shymko, and K. Trabert, *Science*, 1978, **199**, 259.

Plates 1—10 *Some examples of morphogenesis*

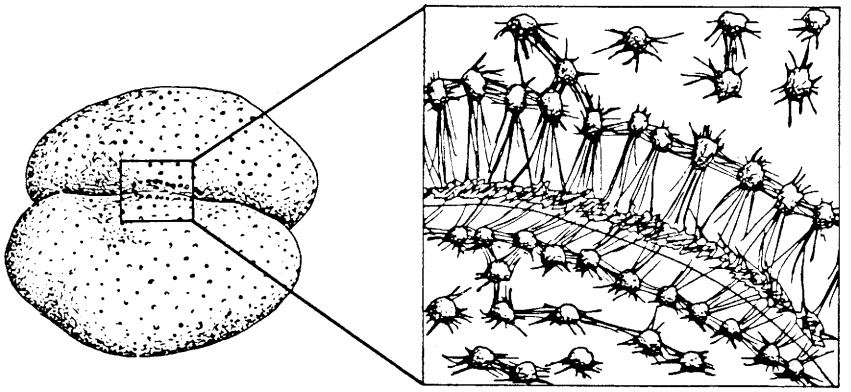
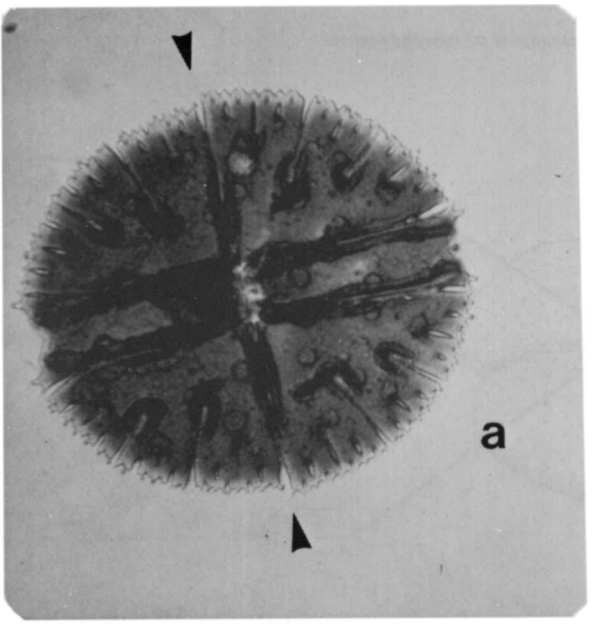


Plate 1 *Pores in the cellulosic cell wall of a single-celled alga,* Cosmarium botrytis. *The array of pores is not quite regular, but far from random. Nucleation-depletion model, section* 3A *and Figures* 8 *to* 10
(Reproduced by permission from ref. 14)



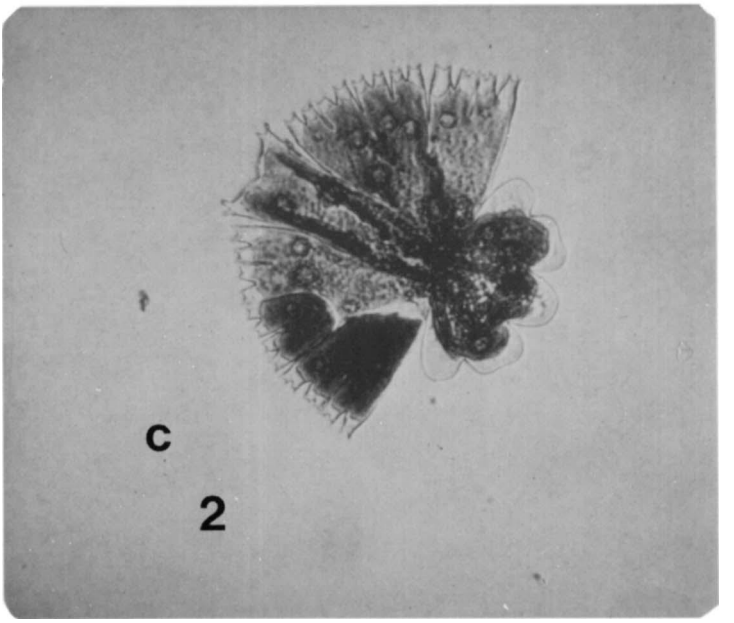
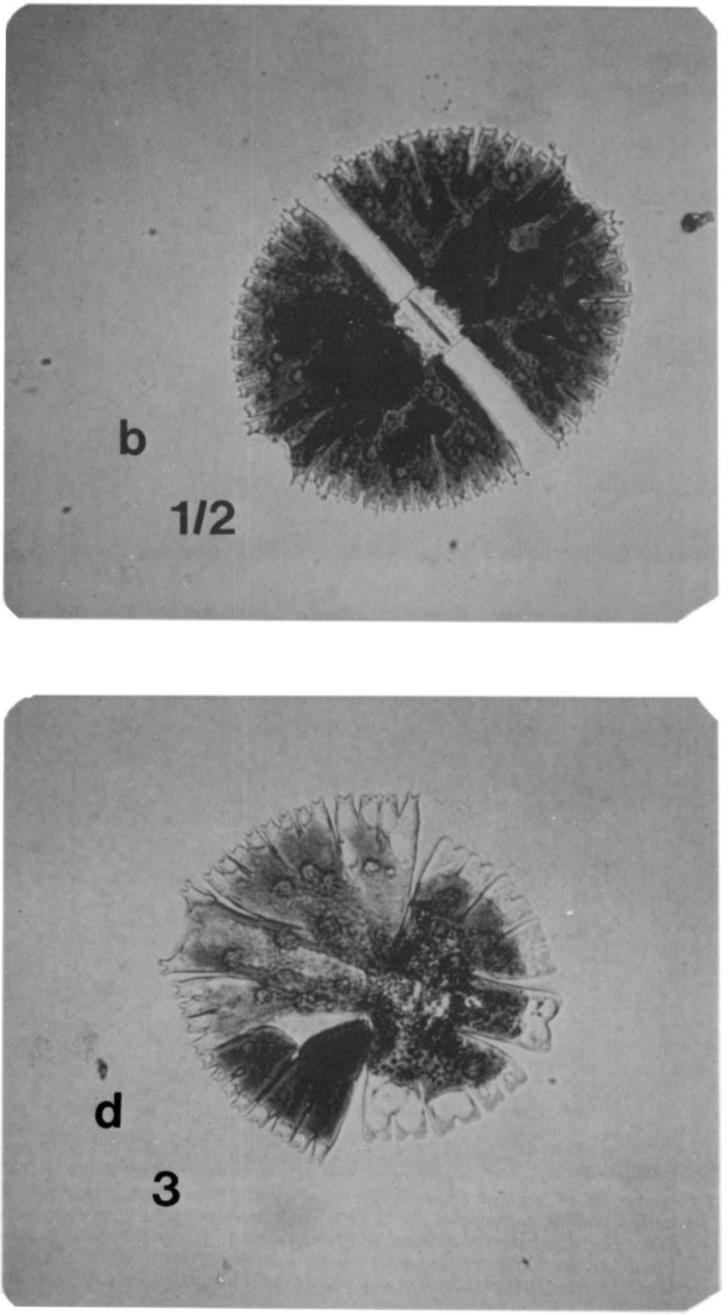


Plate 2 *Semicell morphogenesis in a single-celled alga,* Micrasterias rotata. *The cell has the shape of a biconvex lens, almost split in half (arrows) and with the semicircular margin of each half deeply indented. In vegetative reproduction, the cell splits at the isthmus between the two halves. A 'bubble' of cell wall grows out from each half, and develops a*



wavy outline, the lobes of which repeatedly square off and bifurcate in a manner suggesting the presence of non-linear concentration waves, sections 2B *and* C. *Fully-grown, diameter* $\sim 200\mu$m. *Numbers are hours after cell division*

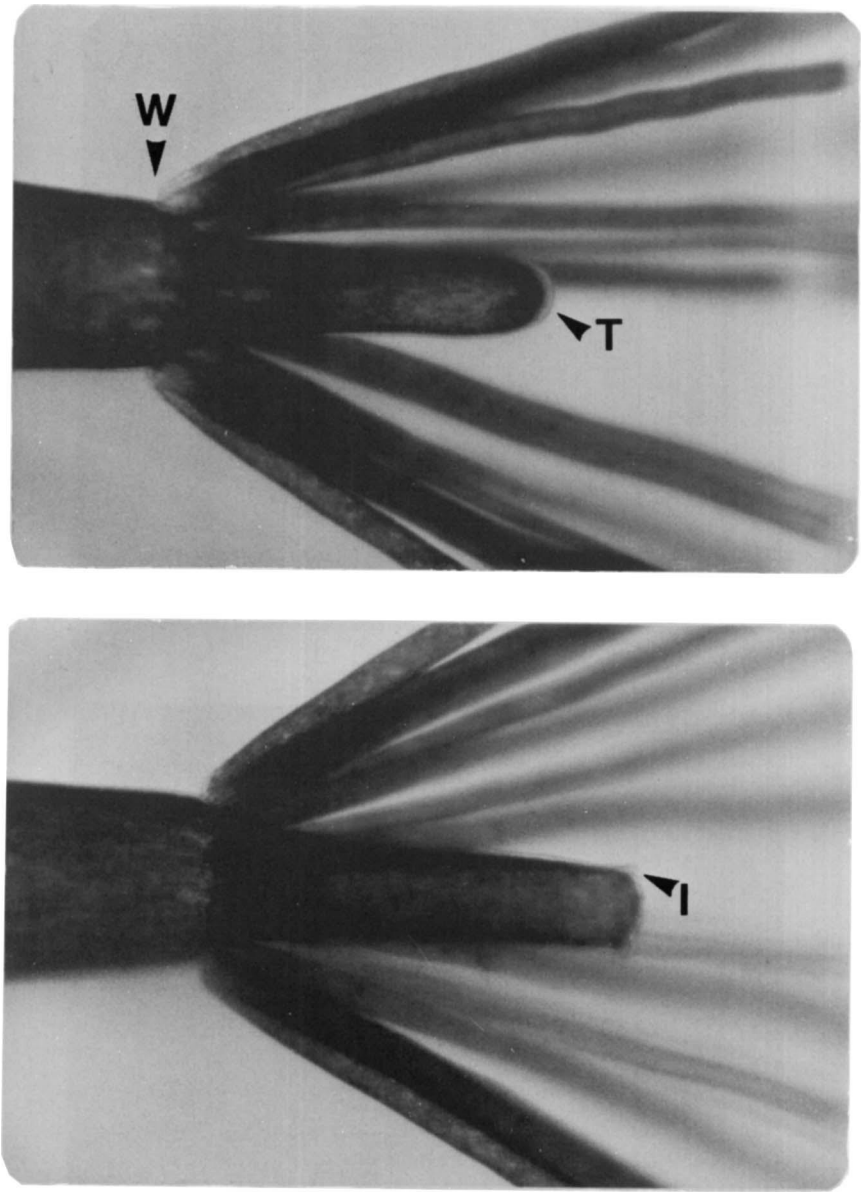


Plate 3 *Growing tip of a giant single-celled alga,* Acetabularia mediterranea. *The cell grows, by action mainly at rounded tip* T, *into a cylinder perhaps* 4cm *long* 400μm *in diameter. Every few days, the tip flattens and forms a ring of hair initials* I *from which a whorl* W *of hairs grows* (W *marks previous whorl; another is just forming at* I). *Reaction diffusion and Arrhenius-type temperature dependence of hair spacing at first appearance of initials, section* 2D

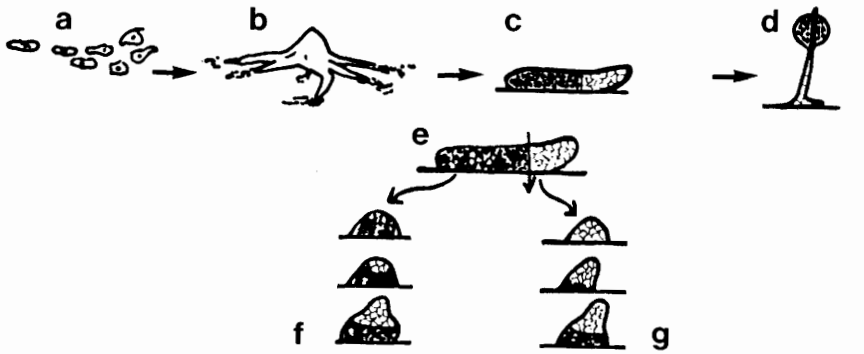


Plate 4 *Multicellular morphogenesis without cell division or feeding: the cellular slime mould* Dictyostelium discoideum. *Thousands of independent amoebae, dividing and feeding on decaying vegetation, aggregate into an elongated* (~ 1 mm) *assembly* (a *to* b *to* c) *and differentiate into 'pre-spore' and 'pre-stalk' cells* (c *or* e), *which rearrange into the stalk and mass of spores of the fruiting stage* d. *Treatment of the differentiation by reaction-diffusion, section* 2B. *Restoration of pattern after cutting* (e *to* f *and* g), *computation shown in Figure* 4

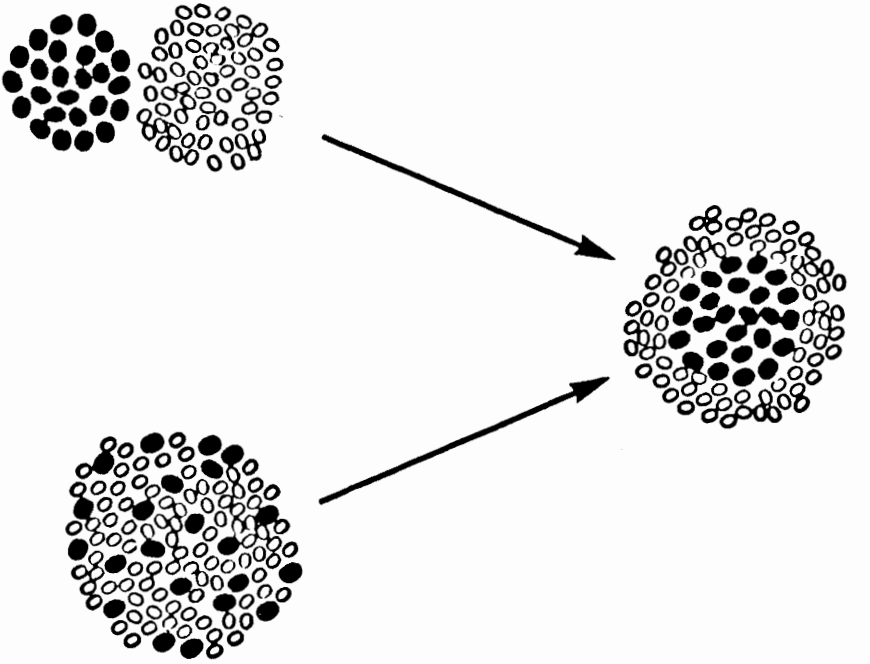


Plate 5 *Schematic representation of 'engulfment' and 'sorting out' experiments on two types of living cells from an animal embryo. Cells as analogues of immiscible molecules and differential adhesion, sections* 3B *and* 4A

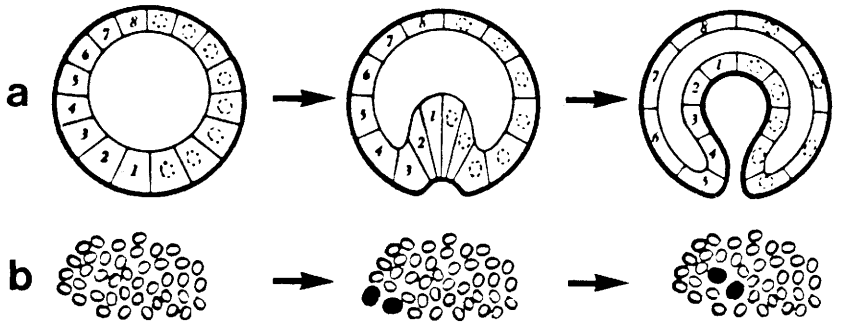


Plate 6 *Gastrulation: (a) Schematic cross-section through a hollow sphere of one layer of cells, showing shape changes and cell movements by which the gut starts to form. Given in many textbooks as typical of processes in the vertebrates, actually seen in a form close to the idealised diagram only in* Amphioxus *and echinoderms.*

(Reproduced by permission of McGraw-Hill Book Co. from Gilchrist, A Survey of Embryology)

(b) In nematode worms, two cells differentiate and move inward in a manner like the sorting out in Plate 5. From these cells, the gut starts to form. Turing regarded (a) as symmetry breaking from a sphere to be explained by reaction-diffusion

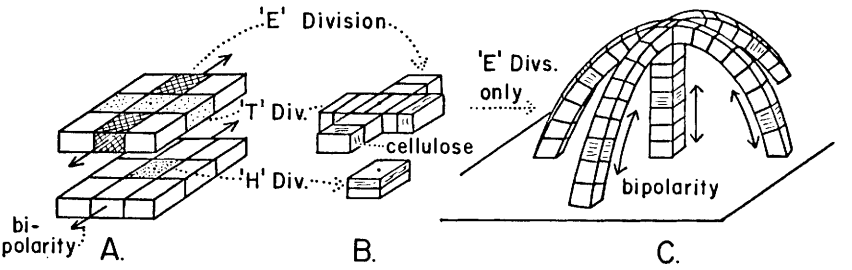


Plate 7 *Initiation of branches or leaves from a plant stem depends on changes in direction of cell division.* A: *two layers of cells lying along the stem.* B, C: *Growth of a branch upwards needs all three possible directions of cell division, here called* E, H, *and* T *because line of division completes the shape of one of those letters on top or front of diagram of cell.*

(Reproduced by permission from ref. 68)

Reaction-diffusion control may lie in tubulin precursors, section 4C

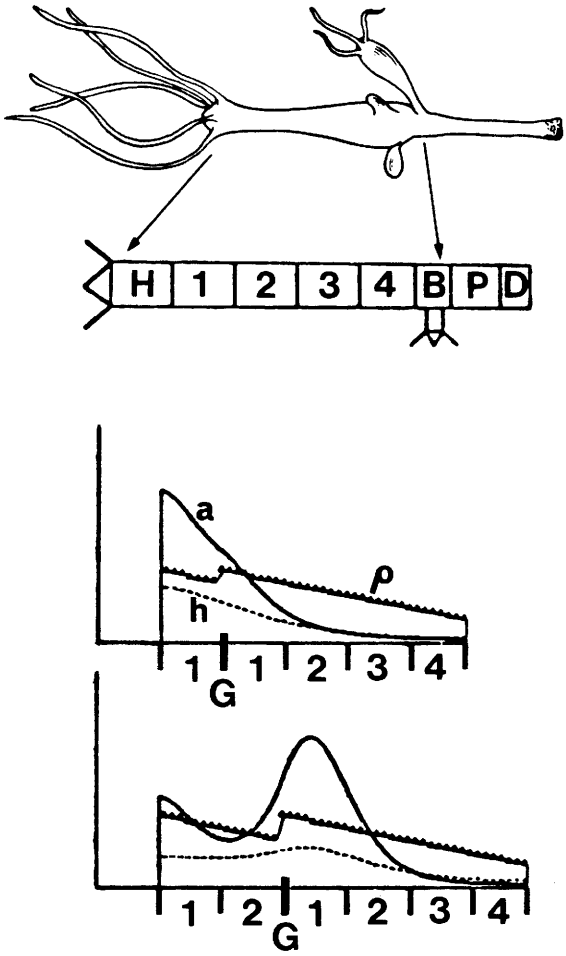


Plate 8 *Hydra is about* 1 mm *long and has about* 10^5 *cells of about* 15 *different types.* H: *hypostome;* 1, 2, 3, 4: *gastric segments;* B: *budding area;* P: *peduncle;* D: *basal disc. Grafting experiments using pieces from two specimens show that a graft of* 1 *to* 1234 *does not grow an extra ring of tentacles, but a graft of* 12 *to* 1234 *does form tentacles at the graft* G. *Gierer–Meinhardt reaction-diffusion theory, section* 2C, *accounts for this by producing a large activator peak out of the discontinuity in source gradient* ρ. a: *activator;* h: *inhibitor*

(Parts of diagram reproduced by permission from refs. 47 and 6)

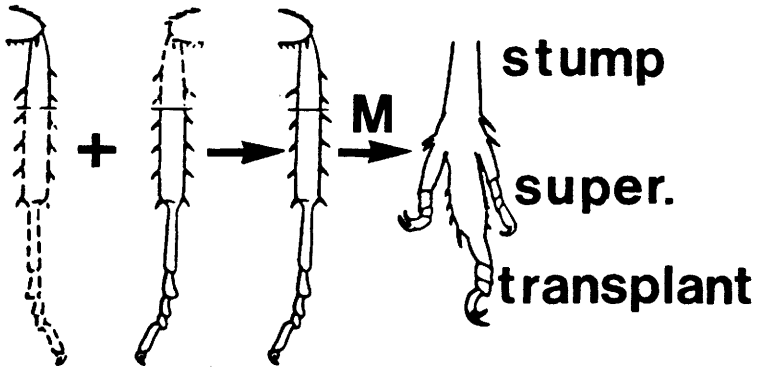


Plate 9 *The stump of an amputated left leg of a cockroach has grafted on to it an amputated right leg, rotated through* 180°. *The parts fuse, and at the next moult, two supernumerary limbs, both left, grow from the graft. Clockface gradient which may need either reaction diffusion or differential adhesion of both to explain it, Figure* 6
(Reproduced by permission from ref. 57)

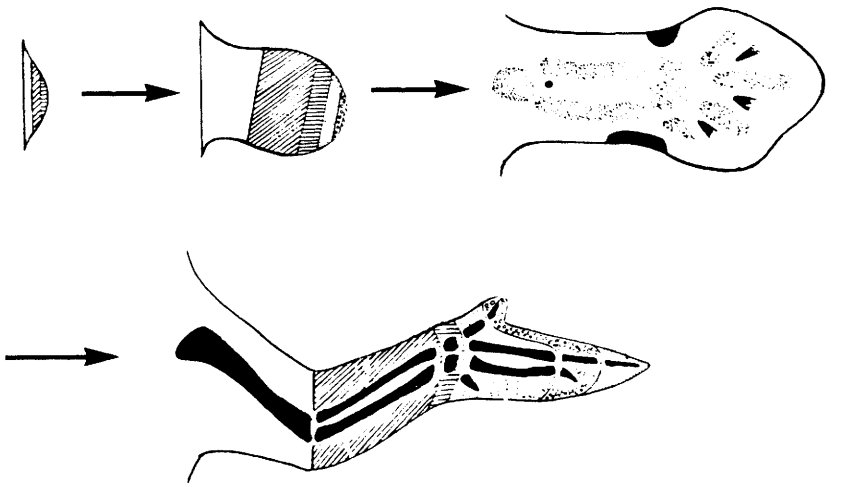


Plate 10 *Four stages in growth of vertebrate limb (chick wing) from a limb bud, showing early origin of several regions within which parts of skeletal structure form by differentiation of cartilage cells (black in last stage shown). In some amphibia (newts), limb graft behaviour closely parallels that in insects, Plate* 9. *Positive feedback in cartilage formation and glycosaminoglycans, section* 4B
(Reproduced in modified form by permission from ref. 57)

a manner which is not cancelled by the next pattern, leading to an increasingly complicated compartmentalisation of the region. He related this to experimental observations on insect development, first to the formation of some boundaries known as 'clone restriction lines' in the development of an insect wing, and second to the origin of segmentation at an early stage of embryonic development.

These examples serve to indicate that, if the hypothetical morphogens exist at all, they are likely to be very diverse both in respect of chemical nature and in respect of the location of their morphogenetic activity. For the single-celled plants, this activity must be near to the cell surface and concerned with either addition of lipids to the cell membrane or addition of cellulose to the wall. For multicellular animals, they may be concerned with nuclear DNA and inhibitors of gene expression attached to it at specific points.

C. Hypothetical Reaction Mechanisms and Non-linear Models.—Great diversity is possible in chemical reaction mechanisms that yield systems of non-linear rate equations equivalent, upon linearisation around equilibrium, to the Turing equations. Only a few schemes have been studied extensively by mathematical analysis and computation.[5,6,31,40,41] From these, however, it is already evident that different models have quite different characteristics in the regions of non-linear behaviour.

The Prigogine 'Brusselator' model is the reaction:

$$A + B \rightleftharpoons D + E \tag{21}$$

proceeding with the aid of morphogen intermediates X and Y according to:

$$A \rightleftharpoons X \tag{22a}$$

$$B + X \rightleftharpoons Y + D \tag{22b}$$

$$2X + Y \rightleftharpoons 3X \tag{22c}$$

$$X \rightleftharpoons E \tag{22d}$$

Mathematical analysis has been mainly for time-independent A, B, D, and E, as fixed by supply and removal, and diffusible X and Y with, as in Turing's model, $\mathscr{D}_Y > \mathscr{D}_X$.

The presence of two reactants, one of which converts into X while the other removes X, is an important feature of this model. As usual, there are threshold conditions for the development of kinetically maintained structure: A must exceed a threshold value, but B must be *less than* a threshold value. This makes it rather easy to devise models for control of A and B that would allow interesting things to happen over only part of a system, as quite commonly happens in morphogenesis. For instance, if B diffuses into a system from its boundaries, and is used up in an overall first-order decay [apart from its destruction in reaction (22b)], a steady state can arise in which B is higher at the boundary than at the

[11] J. D. Murray, 'Lectures on Nonlinear-differential-equation Models in Biology', Clarendon Press, Oxford, 1977.

centre of the system. Morphogenetic action may then be 'switched on' only in some central region. This could be relevant, for example, to tip growth of some algae, fungal hyphae, and root hairs of higher plants, in which the growth action is concentrated in a roughly hemispherical tip of a cell, leaving behind it a cylinder which grows not at all or very slowly. Sometimes, more complicated patterns can form at the tip (Plate 3).

Variants of the Brusselator scheme have been devised by Tyson and co-workers,[42,43] for example:

$$\text{Reactants} \rightleftharpoons Y \tag{23a}$$

$$Y \rightleftharpoons X \tag{23b}$$

$$2X + Y \rightleftharpoons 3X \tag{23c}$$

$$X \rightleftharpoons \text{Products} \tag{23d}$$

For this class of models, computations show another important feature that is not shared by some other types of non-linear model (*e.g.*, the model of Gierer and Meinhardt discussed below). Morphogen peaks can rather easily arise, decay, and move around as pattern develops, in a manner somewhat reminiscent of Ostwald ripening of a precipitate, but of course with better spatial control and with the spatial heterogeneity sustained only by kinetic effects. Lacalli[44] has shown by computations using the rate equations equivalent to equations (23), for a two-dimensional region with random input and no-flux boundaries, that a very irregular pattern of X peaks at first arises and gradually changes into a perfect hexagonal array. Not all models have the adaptability to find the kinetic route to this structure from an irregular one.

Reaction schemes containing three intermediates X, Y, and Z are capable of generating both spatial pattern and time-oscillatory behaviour without the need for bimolecular autocatalysis in a single step. In some instances, the distinction is trivial, depending simply on the degree of approximation made in writing down the kinetic scheme. Consider, for example, the sequence:

$$C + X \rightleftharpoons Y \tag{24a}$$

$$X + Y \rightleftharpoons Z \tag{24b}$$

$$A + Z \rightleftharpoons X + Z \tag{24c}$$

$$X \rightleftharpoons B \tag{24d}$$

If the first two steps are fast, so that equilibrium is maintained in them at all times, they are merely a long-winded way of writing the addition of 2X to a catalytic site C to give an activated site CX_2 which is Z. The first three steps are then equivalent to:

$$A + 2X = 3X \tag{25}$$

A more complicated example of a three-intermediate system is the scheme of

[42] J. J. Tyson and J. C. Light, *J. Chem. Phys.*, 1973, **59**, 4164.
[43] J. J. Tyson and S. A. Kauffman, *J. Math. Biol.*, 1975, **1**, 280.
[44] T. C. Lacalli, *Philos. Trans. R. Soc. London, Ser. B*, 1981, **294**, 547.

Field and Noyes[21] for the Belousov–Zhabotinski reaction. This is the cerium-catalysed oxidation of malonate by bromate, overall:

$$3CH_2(CO_2H)_2 + 4BrO_3^- \rightarrow 9CO_2 + 4Br^- + 6H_2O \quad (26)$$

Spatial pattern and temporal oscillations are made visible in this reaction by adding an Fe^{2+}/Fe^{3+} couple with an indicator such as ferroin, giving orange and blue colours. For a stirred solution, one sees oscillations, in which the colours alternate in time. For an unstirred system in a tall cylinder, stripes of alternating colour are seen moving vertically. In a Petri dish, concentric rings of alternating colour expand from several centres, and interact where they meet in a manner not characteristic of interference of waves.[20] It remains unclear whether this reaction is closely analogous to important phenomena in biological systems, or whether it is a chemical curiosity. *Stationary* spatial order is not a common feature of the reaction, except in flow systems.[45] Some doubt has also been cast on the existence of diffusive coupling in the system. In a tall cylinder, the coloured stripes continue to move uninterrupted when the cylinder of solution is cut by a horizontal Plexiglas plate.[46]

The Field and Noyes scheme involves bromous acid, bromide, and cerium(IV) as the three intermediates. Out of an overall mechanism with many steps, the significant ones, with their X, Y, Z designations as used in the 'Oregonator' scheme, are probably:

$$\begin{array}{rl} BrO_3^- + Br^- + 2H^+ \rightarrow & HBrO_2 + HOBr \\ A \quad + \quad Y \quad\quad \rightarrow & X \end{array} \quad (27a)$$

$$\begin{array}{rl} HBrO_2 + Br^- + H^+ \rightarrow & 2HOBr \\ X \quad + \quad Y \quad\quad \rightarrow & P \end{array} \quad (27b)$$

$$\begin{array}{rl} BrO_3^- + HBrO_2 + 2Ce^{3+} + 3H^+ \rightarrow & 2HBrO_2 + 2Ce^{4+} + H_2O \\ X \quad + \quad B \quad\quad\quad\quad \rightarrow & 2X \quad\quad + \quad\quad Z \end{array} \quad (27c)$$

$$\begin{array}{rl} 2HBrO_2 \rightarrow & BrO_3^- + HOBr + H^+ \\ 2X \rightarrow & \text{products Q} \end{array} \quad (27d)$$

$$\begin{array}{l} 4Ce^{4+} + BrCH(CO_2H)_2 + H_2O + HOBr \rightarrow 2Br^- + 4Ce^{3+} + 3CO_2 + 6H^+ \\ 2Z \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \rightarrow 2Y \end{array} \quad (27e)$$

In contrast to the rather changeable behaviour of this reaction, many biological systems show evidence for the existence of remarkably persistent gradients. Much of this evidence arises from cutting organisms and grafting so as to juxtapose pieces which nature would not have thought of putting together. (Plates 8, 9, Figure 6). The model of Gierer and Meinhardt[6–8] gives a good account of some of these situations, and has probably attracted more attention among experimental biologists than any other model of the reaction-diffusion type (this whole field of modelling being still regarded with much scepticism among biologists). The model envisages a gradient, fixed in time or changing only very slowly, of sources for the morphogens in an activator–inhibitor scheme. The latter is such that it tends to amplify small differences between regions, *e.g.* to

[45] M. Marek and E. Svobodová, *Biophys. Chem.*, 1975, **3**, 263.
[46] N. Kopell and L. N. Howard, *Science*, 1973, **180**, 1171.

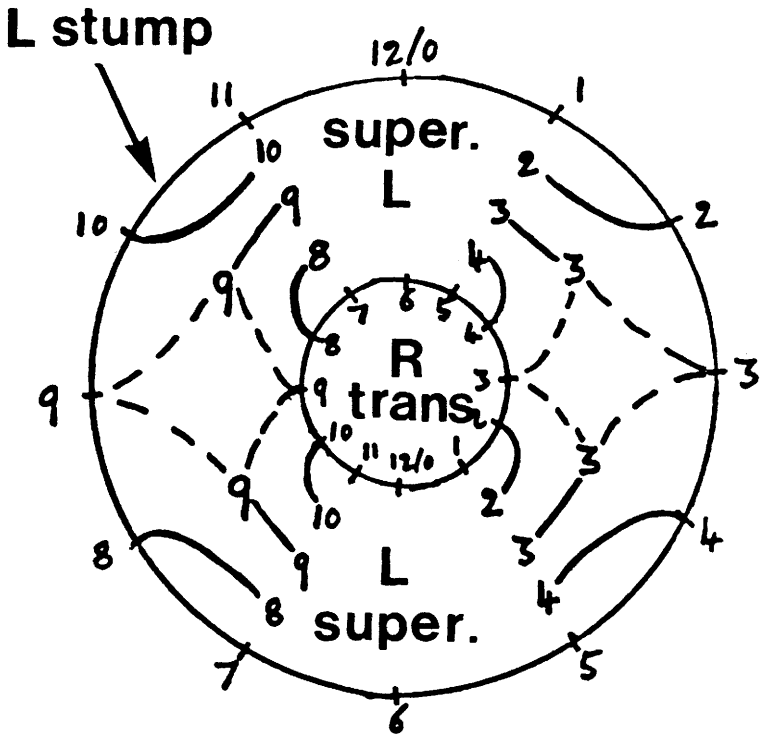


Figure 6 *The clockface angular co-ordinate model of French, Bryant, and Bryant,*[57] *which generalises results of insect leg transplant experiments (Plate 9). Both reaction-diffusion and differential adhesion are being used in attempts to account for this behaviour. Here, supernumerary limbs grow wherever there is a complete circle of values, 1 to 12, and clockwise specifies left leg. Solid and broken lines are contours of positional information*

change a linear gradient into a curved one with a large peak at the high end, and to stabilise the result of this amplification fairly strongly against disturbing influences. Thus it contrasts markedly with the adaptability above-mentioned for the Brusselator, and is not a likely model for growing a two-dimensional hexagonal array out of chaotic input. The source gradient might be something on the multicellular scale of organisation, such as variation in the fractions of two types of cell in a tissue. *Hydra* has a greater proportion of nerve cells at the 'head' end, where the tentacles form, than at the base.[47]

Gierer and Meinhardt actually proposed a number of related models, but the one they used most extensively uses two morphogens a and h (activator and inhibitor), corresponding roughly to Turing's X and Y. The symbols a and h in the following equations will, however, signify complete concentrations, not

[47] P. Grant, 'Biology of Developing Systems', Holt, Rinehart, and Winston, 1978, p. 457; L. Wolpert, J. Hicklin, and A. Hornbruch, *Symp. Soc. Exp. Biol.*, 1971, XXV, 391.

deviations from equilibrium. As usual, $\mathscr{D}_h > \mathscr{D}_a$. The parameters ρ and ρ', both functions of distance s, represent the source gradients for a and h; ρ_0, c, c', μ and ν are rate constants.

$$\partial a/\partial t = \rho_0\rho + c\rho a^2/h - \mu a + \mathscr{D}/\partial^2 a/\partial s^2 \quad (28a)$$

$$\partial h/\partial t = c'\rho' a^2 - \nu h + \mathscr{D}_h \partial^2 h/\partial s^2 \quad (28b)$$

This model was first applied, with notable success, to the results of grafting experiments in *Hydra* (Plate 8).[6,47] More recently, it has been applied to insect morphogenesis.[7]

D. Temperature Sensitivity of Spacing in a Pattern.—If the spacing between repeated parts in a pattern is indeed to be regarded as a chemical wavelength of the form $2\pi(\mathscr{D}/k)^{\frac{1}{2}}$, it follows that spacing is a combination of chemical rate parameters and should itself show the attributes of a rate parameter, among them its well known temperature dependence in the Arrhenius form. An activation energy of the order 10 kJ mol^{-1} is likely for $\mathscr{D}$; that of k may be much more variable, but for processes occupying a time scale of a few hours something like 50 kJ mol^{-1} is the most likely value. The apparent activation energy of the spacing, from a plot of ln λ *versus* $1/T$, should then be

$$E_\lambda = (1/2)(E_D - E_k) \sim (1/2)(10 - 50) = -20 \text{ kJ mol}^{-1} \quad (29)$$

In studies of the single-celled alga *Acetabularia* in my laboratory,[48] spacings have been measured, as a function of temperature, between hairs in the whorls that are from time to time formed through a large part of the growth of the organism (one whorl every few days for some months; Plate 3). Figure 7 shows plots of this temperature dependence, which is just as expected from the above very generalised argument. This good agreement is probably fortuitous. The simplest expression for a properly controlled spacing is $2\pi[\mathscr{D}_X\mathscr{D}_Y/(-k_2k_3)]^{\frac{1}{2}}$ in which the diffusivity and rate constant have been replaced by geometric means of two such quantities, and uncertainties in the estimate of what E_λ should be increase with all such increases in complexity of the theoretical expression. Nevertheless, I hope that this example may serve to encourage further studies along these lines, which have the potential to give fairly clear indications of whether or not a reaction-diffusion mechanism is operating.

3 The Equilibrium Approach: Phase Transitions and their Cell-as-molecule Analogues

A. Inhibitory Fields.—Various scattered structures, both on the surface of a single cell (Plate 1) and in lines or sheets of cells, form arrays that are neither random nor fully ordered. Statistical analysis of the distribution of such structures usually shows that they pack as if each structure were much larger than it appears to be, so that around each visible structure there is a region (a circle, in

[48] L. G. Harrison, J. Snell, R. Verdi, D. E. Vogt, G. D. Zeiss, and B. R. Green, *Protoplasma*, 1981, **106**, 211.

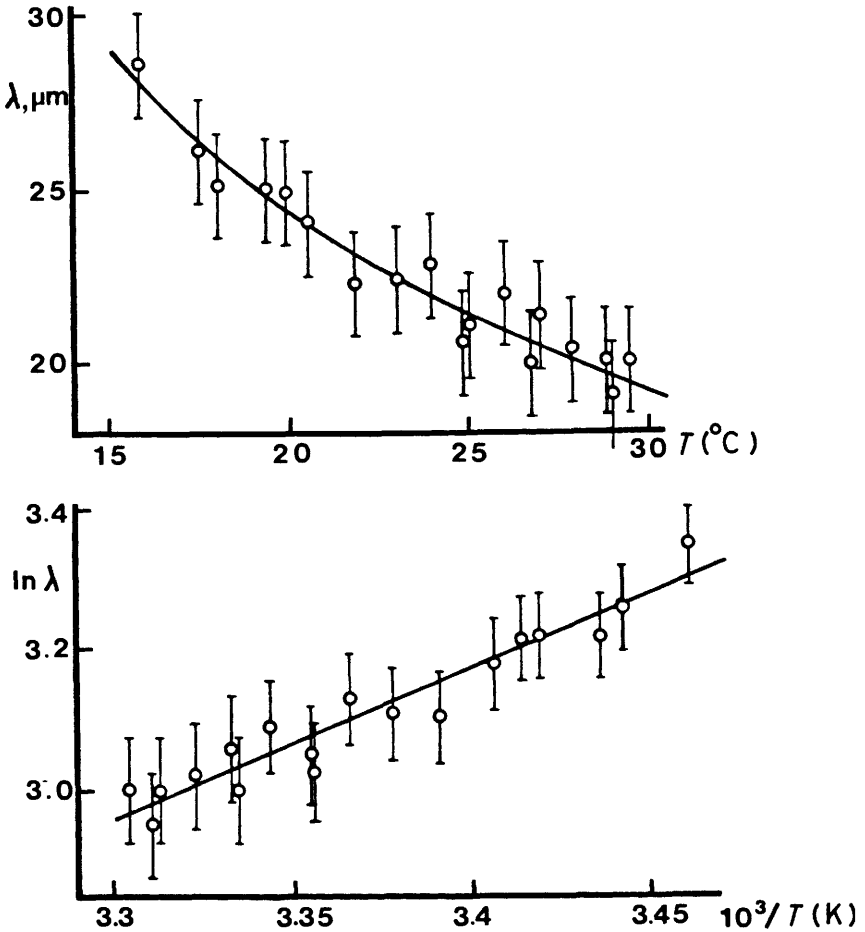


Figure 7 *Spacing λ between parts of a pattern* (Acetabularia *hairs, Plate* 3) *has an Arrhenius-type temperature dependence with a slope corresponding to an apparent activation energy of* -20 kJ mol^{-1}
(Reproduced by permission from ref. 48)

the two-dimensional cases) within which the formation of a similar structure is forbidden. This region is referred to as an 'inhibitory field'. Attempts to account for the inhibitory field have usually invoked a diffusible substance, either an inhibitor moving out from the structure or, more simply, a substance needed for the structure moving towards it and consequently being depleted in the surrounding region below some threshold for formation of similar structures. In such a 'depletion model', the possible relevance of classical ideas of phase transitions is obvious, since we are concerned with nucleation of a structure above some

threshold concentration, which might be analogous to a critical supersaturation.

Among the methods of statistical analysis used for partly ordered patterns, one of the simplest and most popular is the Clark and Evans R parameter.[49] For each pattern point in a two-dimensional array, the distance to its nearest neighbour is found. The average value of this distance is divided by $\rho^{-\frac{1}{2}}/2$, where ρ is the area density of points. The result, R, is unity for a random array, 2 for a perfect square array, and 2.1491 for a perfect hexagonal array. It is striking that R values between 1.62 and 1.70 are found for structures as diverse as: hair follicles on Australian sheep,[50] cone cells in a monkey's retina,[51] and pores in the cell wall of a single-celled alga.[14] (See Figure 8 and Plate 1). Computation[50] for an inhibitory field of fixed radius around each pattern point yields $R = 1.757$.

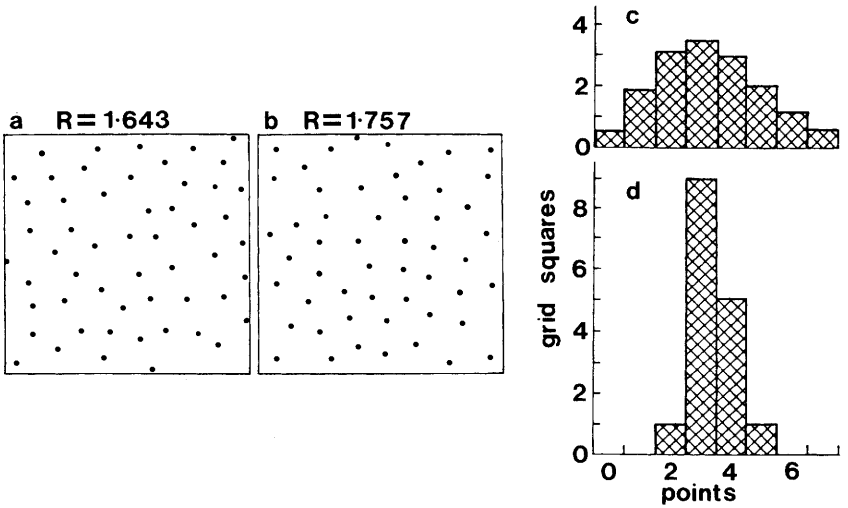


Figure 8 *Partly-ordered patterns: (a) Pores in the cell wall of a desmid* (cf. *Plate* 1, *for a different species but almost identical pattern); (b) computed pattern for an invisible 'inhibitory field' of fixed size around each pattern point.* R *is Clark and Evans' measure of order (see text); (c) for division of pattern into* 16 *smaller squares* (4 × 4 *grid*), *Poisson distribution of number of points in each square,* i.e. *random distribution; (d) actual distribution for pattern (a), showing that it is far from random*
(Reproduced by permission from ref. 14)

There is a serious physiochemical problem in reconciling the concept of a fixed-radius field with the suggestions that the field is generated by inward or outward diffusion. One might expect such diffusion to be time dependent, with any particular concentration contour probably expanding in radius with $t^{\frac{1}{2}}$. Lacalli and I[14] computed patterns on this basis, and obtained very low order: $R = 1.352$.

[49] P. J. Clark and F. C. Evans, *Ecology*, 1954, **54**, 445.
[50] J. H. Claxton, *J. Theor. Biol.*, 1964, **7**, 302.
[51] H. Wässle and H. J. Riemann, *Proc. R. Soc. London, Ser. B*, 1978, **200**, 441.

A second difficulty concerns the nature of the critical concentration condition for nucleation of new pattern points. The computations just mentioned involve a constant rate of nucleation in all parts of the system not covered by inhibitory fields. In classical nucleation theory, the nucleation rate varies extremely rapidly with concentration above the critical supersaturation. We concluded that this model of the critical concentration would not work.

Another type of critical concentration commonly found in solutions is the critical micelle concentration.[52] Detergent solutions commonly contain monomeric species A up to a concentration C_c at which a polymerisation of some fairly definite number m of species A to form micelles M occurs:

$$mA = M \tag{30}$$

The ideal chemical equilibrium equation for this process shows that, if m is a

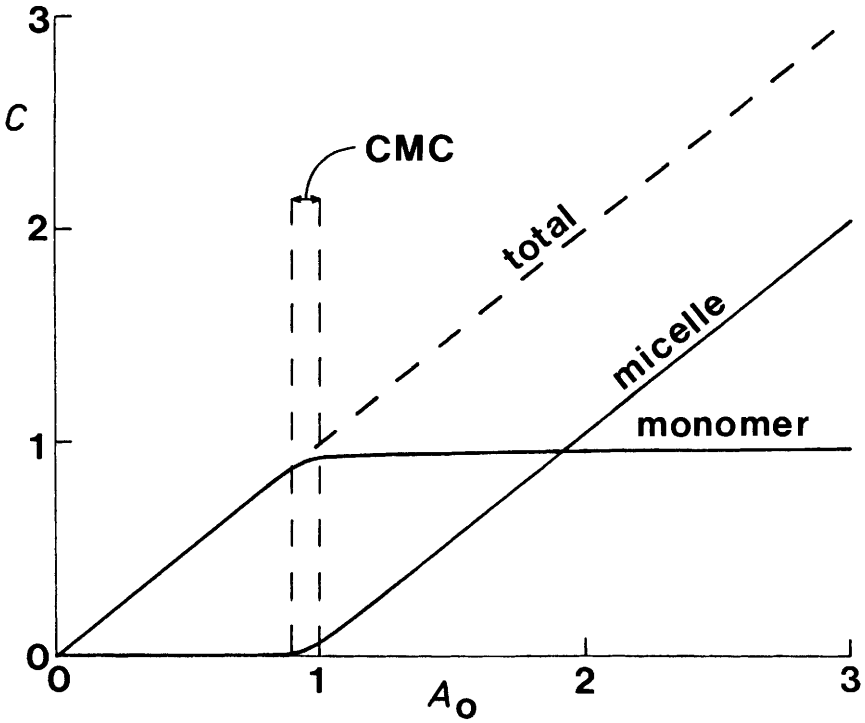


Figure 9 *Ideal micelle equilibrium for* 100 *monomers per micelle. This is a type of 'critical concentration' effect in which the system is well controlled above the critical concentration, in contrast to critical supersaturation effects.* A_0, *total concentration (as monomers);* CMC, *critical micelle concentration*
(Reproduced by permission from ref. 14)

[52] L. R. Fisher and D. G. Oakenfull, *Chem. Soc. Rev.*, 1977, **6**, 25.

large and constant number, if A is added to solution until C_c is reached, any further addition of A leads to micelle production while the monomer concentration stays at C_c (Figure 9, for $m = 100$). This constancy of A concentration beyond the critical value at which micelles start to form is exactly what is needed to give uniform behaviour in the whole system outside the inhibitory fields. Our[14] model for cell wall pores is shown in Figure 10. As the cell grows to its full size,

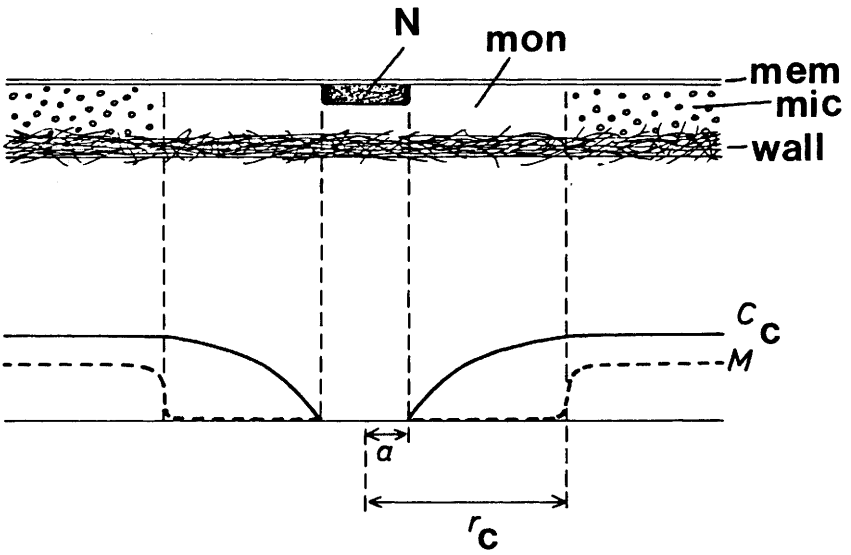


Figure 10 *The 'micelle' model for the boundary of an inhibitory field at a plant cell surface.* N, *nucleus of pore initial;* mem, *cell membrane (interior of cell is above);* wall, *primary cellulosic cell wall;* mon, *solution of monomers;* mic, *solution of micelles*; C_3, *critical micelle concentration; M, monomer concentration; solid line is micelle concentration* (Reproduced by permission from ref. 14)

the cytoplasm is bounded, as usual, by a lipid bilayer membrane. Outside (below, in the diagram) there is a thin layer of solution bounded by the primary cellulosic cell wall. As the cell reaches full size, there is a period of some minutes before the thicker secondary cell wall forms between membrane and primary wall. In this short period, plugs of an unknown material (N) are laid down to form the 'pore initials', which later disappear to leave the pores in their place. We envisage the unknown material as being present in the solution layer as monomers A and micelles, or polymers, M, with fraction f of all the material (calculated as monomers) in the polymerised state. From time to time, a micelle may attach to the membrane and change into a state in which attachment of additional material takes place at an equilibrium concentration $C_1 \ll C_c$. N has a definite radius a (~ 50 nm) and additional material goes to thicken N downwards without increasing a. This diffusion of material towards N depletes the surrounding region

below C_c, so that there are no micelles present up to some radius r_c, of the order of 10^3nm but changing with time. By considering diffusion of both monomers and micelles, we were able to show that, if the micelle diffusivity is $\mathscr{D}_M$ and that of monomers is roughly equal:

$$r_c = a(3.02\mathscr{D}_M t/a^2)^{(1-f)/2} \tag{31}$$

[From equations (23) and (24) of our paper[14] with the approximations indicated below those equations.] This model allows for r_c to vary with time very much more slowly than $t^{\frac{1}{2}}$, and hence is capable of accounting for the observed degree of order in the pore pattern.

B. Sorting Out: Cells as Molecules.—In the embryonic development of animals, cells differentiate into a number of different types, and each type ends up occupying a different spatial region to constitute the various tissues. (This description is oversimplified; a tissue may contain more than one cell type.) A major question, still incompletely resolved, is whether some chemical influence, *e.g.* a reaction-diffusion pattern of concentration, directs the cells in a particular region to differentiate in a particular way, or whether individual cells can differentiate anywhere, followed by migration of like cells to the same region. It has long been known that, if two types of embryonic tissue are taken apart into separate cells, and the cells are intimately mixed, they tend to 'sort out' into two aggregates, each of one cell type. Steinberg[16] studied this phenomenon in such a way as to place it on a basis comparable to the sorting out of two types of molecule in a pair of immiscible liquids.

Steinberg took six tissues from chick embryos (A, B, C, D, E, F, respectively: back epidermis, pigmented epithelium of the eye, heart ventricle, liver, cores of limb cartilage, and neural tube) at early enough stages to ensure that each contained only one cell type. He showed that these tissues, taken in pairs, tended to arrange themselves as a sphere of one tissue entirely enclosing a sphere of the other, in a reproducible order (*e.g.* A always goes inside B). Essentially the same final configuration is reached in two types of experiment: 'engulfment', in which two pieces of tissue are skewered in adjacent positions on one skewer, and 'sorting-out' as described in the preceding paragraph (Plate 5). This suggests that the final configuration is an equilibrium one.

He showed that if two pair combinations yield A > B and B > C (where > means 'goes inside'), the transitive property is present, *i.e.* the A/C experiment yields A > C. From such study of all 15 possible pair combinations, it was possible to prove the existence of a hierarchy A > B > C > D > E > F. The chance of this occurring 'by accident' for a list of n items is $n!/2^{n(n-1)/2} = 6!/2^{15} = 0.022$ for $n = 6$. Hence the existence of the hierarchy is established to about 98 % confidence. Such a series suggests the existence of a quantifiable property measuring position in the series. (As, for example, one might find a replacement series for metals in solution, and assert the existence of a quantifiable property, which is of course oxidation potential.) For the corresponding phenomenon in a series of immiscible liquids, this property would be surface tension. (A question

of relative volumes comes in here. If the two assemblies of cells have the same volume, the contact surface between A and B has the same area whether A goes inside B or the reverse. The interfacial tension γ_{AB} then makes the same contribution to the free energy of A-inside-B and B-inside-A, and cancels out of the free energy difference between those configurations. One need then think only of the relative surface tensions γ_A and γ_B between each aggregate and the medium in which they are suspended.) By placing a piece of each tissue on a flat plate, Steinberg[17] essentially carried out a 'sessile drop' surface tension determination He did not quantify the result, but showed simply that the final shape was flatter as the tissue lay lower in the hierarchy.

This led to the concept of cohesive forces between adjacent cells, or 'differential adhesion' between cells of different types. Many kinds of model have been proposed for this. In the simplest (homophilic), a molecule attached to the exterior of one cell surface adheres to a like molecule on the surface of another cell. The same adhesive molecule might be present throughout the hierarchy, with the cell types differing only in the fraction of the surface covered by these molecules. Suppose that these fractions are a and b for cell types A and B. If the molecules are in fixed positions, the adhesive fraction of any contact area between two cells will be a^2, b^2, and ab for the contacts AA, BB, and AB. Thus, if the cells are all geometrically similar so that contact areas are the same for all pair combinations, the binding energy for AB is the geometric mean of those for AA and BB. This, curiously enough, is the same rule suggested by Berthelot in 1898 for van der Waals forces in mixtures of fluids, and later justified in the theory of London dispersion forces, approximately.[53]

Quantitative aspects of adhesion between cells remain rather obscure. If cells, suspended in an aqueous medium, had bare lipid membranes, they would have *actual* van der Waals attractive forces at the contact of adjacent cells. The resulting binding energy has been estimated at 330 $kT\mu m^{-2}$ of contact area. For contact area of only $10\mu m^2$, this gives a binding energy of $3300kT$. Colloidal particles are considered to form stable aggregates if binding forces exceed about $10kT$ between particles. Thus cells with bare lipid surfaces should be bound together into very rigid masses; neither sorting out nor the suspension of cells in media such as the blood should be possible.[54] Evidently the glycoprotein cell coat of animal cells must, among other functions, cancel out the membrane-to-membrane adhesions, leaving it possible for specific molecules to re-establish much smaller adhesions. To avoid the same problem of excessive adhesion, these specific interactions must be quite weak or quite sparsely distributed. This may be illustrated by considering the quantities involved in the immiscibility criterion for two types of cell. Two pieces of tissue, each about 1 mm^3, would each contain of the order of $N = 10^6$ cells. If each cell is regarded as a 'molecule', the unmixing of

[53] J. H. Hildebrand and R. L. Scott, 'Regular Solutions', Prentice-Hall, 1962.

[54] D. E. Brooks, personal communication, using data from V. A. Parsegian and D. Gingell, *J. Adhesion*, 1972, **4**, 283; 'Recent Advances in Adhesion', ed. L.-H. Lee, Gordon and Breach, New York, 1972; D. Gingell and S. Vince, 'Adhesion and Motility of Cells', ed A. S. G. Curtis, Cambridge University Press, 1979.

an intimate mixture of these cells into two separate aggregates involves an unfavourable entropy change:

$$\Delta S = -2kN\ln 2 = -1.4 \times 10^6 k \tag{32}$$

Immiscibility therefore requires a favourable energy change for unmixing of the order of a million times kT for the whole assembly, or about kT per cell. More precisely, if a cell–cell contact leads to binding energy:

$$w_{AA} = ckTa^2;\ w_{BB} = ckTb^2;\ w_{AB} = ckTab \tag{33}$$

then for $10\mu m^2$ contact, one adhesive molecule per nm^2 at $a = 1$ (*i.e.* 10^7 as maximum number of molecule-to-molecule adhesions per contact), and adhesive energy of order kT (about 2.5 $kJ\,mol^{-1}$) between molecules, $c = 10^7$. For close packing of molecules (12 contacts) in both mixed and unmixed assemblies, the energy of unmixing would be:

$$\Delta E = 6Nw_{AA} + 6Nw_{BB} - 3N(w_{AA} + w_{BB} + 2w_{AB}) = -3cNkT(a - b)^2 \tag{34}$$

This, together with equation (31) indicates that separation is favoured for:

$$|a - b| > (2\ln 2/3c)^{\frac{1}{2}} = 2 \times 10^{-4} \tag{35}$$

i.e. the discussion concerns parts in ten thousand of the surface covered with adhesive molecules.

The above considerations are on the basis of a static picture of the cell as a rigid object. This is clearly incorrect, because Brownian motion of rigid objects of size of the order of $10\mu m$ would be much too slow to account for sorting out on a time scale of hours. The cell surface must be seen as continually changing in shape as a result of (i) molecular collisions, in the manner of Brownian motion, producing temporary small-scale deformations of the cell surface, and (ii) similar deformations being produced from inside the cell by, for example, the action of contractile microfilaments, which might still be random in relation to any directional effect on motion of the cells.[55] Such motions would of course tend to diminish the effective contact area between cells at any instant, so that much greater surface coverages of adhesive molecules would be needed to produce time-average interaction energies as discussed above. One way to express this kind of effect might be to use rigid-cell geometrical pictures, but introduce a fictitious temperature, much higher than the real temperature, to represent the enhanced random motion. Computer modelling of sorting out is an active field.[56]

Some remarkable phenomena occur in insect and amphibian (newt) limb regeneration, especially in grafting experiments, indicating the existence of persistent gradients which can be represented by an angular co-ordinate around the limb and a linear co-ordinate along it (Plate 9, Figure 6).[57] These are un-

[55] M. S. Steinberg and L. L. Wiseman, *J. Cell Biol.*, 1972, **55**, 606.

[56] R. Gordon, N. S. Goel, M. S. Steinberg, and L. L. Wiseman, 'Mathematical Models for Cell Rearrangement', ed. G. D. Mostow, Yale Univ. Press, 1975, p. 196; N. S. Goel and G. Rogers, *J. Theor. Biol.*, 1978, **71**, 103 and 141; R. J. Matela and R. J. Fletterick, *J. Theor. Biol.*, 1980, **84**, 673.

[57] V. French, P. J. Bryant, and S. V. Bryant, *Science*, 1976, **193**, 969.

explained, but there is definite evidence for gradients of adhesiveness of cells in animal embryos.[58]

4 The Molecular Basis of Morphogenesis

This review does not include a thorough presentation of the structural approach for two reasons. First, the review is intended to indicate the areas of specific physicochemical interest, and much structural information is of more purely biological or organochemical interest. Second, the balance of space allotted to different topics in this review reflects my bias on the extent to which the various types of concept are likely to turn out to be important in the crucial control steps in morphogenesis: kinetic > equilibrium > structural. The chemical reader should recognise that this is probably precisely the opposite order to that which the majority of experimental biologists would currently give. I have selected a few problems for brief discussion below, because I believe that they give a perspective on how structural aspects may fit in with the other matters discussed above.

A. Heterophilic or Lock-and-key Adhesion.—Adhesion between two cells might involve bonding between two different molecules. This is often referred to as the lock-and-key model, probably with a geometry in mind similar to that of a substrate fitting into a suitably shaped infold of an enzyme; except that here the substrate will not be a free-moving small molecule, but probably held as terminal to some larger structure as the key is held in the hand. A cell with 100 % locks would bond best to a cell with 100 % keys. This model has been used particularly in relation to some features of the assembly of the vertebrate nervous system, especially the joining of the optic nerve to the brain. In the lower vertebrates, the junction is made at the optic tectum, which can be thought of very roughly as a square of a million cells, 1000 × 1000. Each retinal axon seems to find precisely the correct cell to connect to, so that there is a specificity of about one part in 1000 in each of two rectangular co-ordinate directions. Rival theories for how this happens are numerous and diverse,[59,60] and include an activation–inhibition model (Willshaw and von der Malsburg) with interesting correspondences to the Turing model. Roth[60] has proposed a lock-and-key model in which there is a gradient from 100 % locks to 100 % keys across each direction on the tectum, and matching gradients across the array of axons advancing to meet the tectum. Accounts of this model have generally failed to point out, however, that, if the locks and keys are in fixed positions on the cell surface and cannot adjust to find each other, a cell with 50 % of each will join up equally well with an axon regardless of the fractions of locks and keys on the latter. The model is not fully specified without a precise account of how far the locks and keys are able to migrate to find their partners, and the model doesn't work without some limited mobility.

[58] M. S. Steinberg, in ref. 37.

[59] R. W. Sperry, 'Organogenesis', ed. R. L. DeHaan and H. Ursprung, Holt, Rinehart and Winston, 1965, p. 161; R. A. Hope, B. J. Hammond, and R. M. Gaze, *Proc. R. Soc. London, Ser. B*, 1976, **194**, 447; D. J. Willshaw and C. von der Malsburg, *Proc. R. Soc. London, Ser. B*, 1976, **194**, 431; *Philos. Trans. R. Soc. London, Ser. B*, 1979, **287**, 203.

[60] R. B. Marchase, A. J. Barbera, and S. Roth, in 'Cell Patterning', (CIBA Foundation Symposia, new series, no. 29), Elsevier, Amsterdam, 1975, p. 315.

Roth, in addition to proposing the model and finding some experimental evidence for an adhesive gradient (using tecta and retinal cells *in vitro*), has devised plausible biochemical models for the nature of the locks and keys. The animal cell coat, outside the lipid membrane, consists of glycoproteins, which are long polypeptide chains carrying a number of oligosaccharide side-chains, each having perhaps ten monosaccharide molecules in it. The side chains are built up with the aid of glycosyl transferase enzymes, which can attach to the end of the growing chain, wait for the next required sugar to come along, in the form of a sugar nucleotide, and attach it to the chain, at which stage the enzyme is released from attachment to the chain. Roth's model is that the glycoprotein is attached to one cell, the enzyme to the other, and when the two have joined it happens that the next sugar substrate is absent and the enzyme is never released. This gives the lock-and-key junction. Roth devised various forms of this model, and among them pointed out that, if there were three types of side chain, each of which could exist in ten different stages of partial completion, 10^3 chemically different arrangements could arise from just a few monosaccharides, quite enough to account for a specificity of one part in 1000. To my mind, this sort of model is very useful but perhaps merely pushes the crucial question back one stage. The production of the required gradients of incomplete side chains would require some gradients of sugar substrates, and how do those gradients arise? I am led back towards reaction-diffusion.

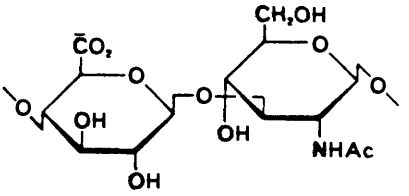
B. Positive Feedback on the Multicellular Scale.—The model of Gierer and Meinhardt described in Section 3C above used autocatalysis (positive feedback) on the molecular scale, with a fixed source gradient which might involve relative numbers of two types of cells. If any substance is found which is produced by a particular type of cell and which tends to make other cells differentiate into that same type, then there is a positive feedback direct into the source gradient. Such a substance is not a morphogen in the strict Turing sense. It might, however, give the differentiation process some partial mathematical analogy to reaction-diffusion theory, depending upon how closely the subsequent movements of differentiated cells resemble diffusion.

Schaller[61] (in the same institute as Gierer and Meinhardt) found an oligopeptide hormone, produced by nerve cells of *Hydra*, which when added to *Hydra* tissue gave it an increased tendency to form 'heads', *i.e.* rings of tentacles. The 'head' contains a higher fraction of nerve cells than the rest of the animal. There is certainly a positive feedback loop here, but this substance is not regarded (Meinhardt, personal communication) as the morphogen a in equations (28).

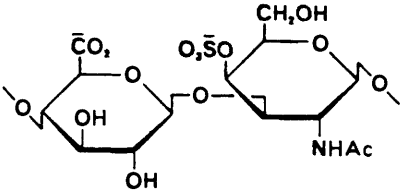
Glycosaminoglycans (GAGs) are polysaccharides composed of alternating units of a sugar acid (*e.g.* glucuronic acid) and an amino-substituted sugar. They are widespread in animals, but are found particularly as constituents of the extracellular matrix of connective tissue.[62] Hyaluronate (1) and the chondroitin

[61] H. C. Schaller, *J. Embryol. Exp. Morph.*, 1973, **29**, 27 and 39.
[62] J. E. Scott, *Chem. Br.*, 1979, **15**, 13.

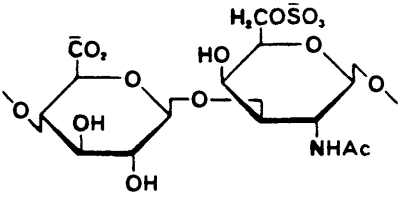
sulphates (2) and (3) are especially prevalent in cartilage. In the embryogenesis of vertebrates, differentiation into cartilage precursors is the first stage in formation of the skeleton, determining its complicated geometry. The structure of a limb, for example, is controlled mainly by activity in the tip region of a limb bud, in which



(1)



(2)



(3)

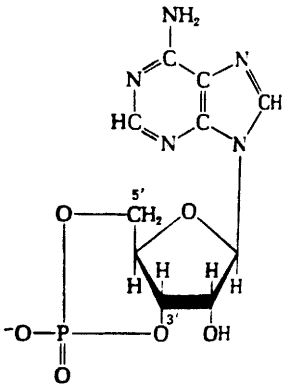
a mass of primary mesenchyme cells is source material for differentiation towards cartilage and other tissues, and is controlled by adjacent structures such as the apical ectodermal ridge and zone of polarising activity (Plate 10).[63]

Urist *et al.*[64] showed that muscle tissue from newborn rats could be made to differentiate into cartilage by the influence of bone matrix gelatin, and that this change was enhanced by chondroitin sulphate but depressed by hyaluronate. Both substances are produced by the differentiated cells, so that there is some evidence here for both self-activation and self-inhibition. These authors, however, mention Turing's theory and believe that a morphogen is present but that it is not to be identified with either of these polysaccharides. They write: 'Morphogens are

[63] L. Wolpert, J. Lewis, and D. Summerbell, in ref. 60, p. 95.
[64] M. S. Urist, Y. Terashima, M. Nakagawa, and C. Stamos, *In Vitro*, 1978, **14**, 697.

short-lived, low-molecular mass, rapidly diffusable hydrophobic proteins that have not yet been isolated, purified, and identified but might be found in and on differentiating cell surfaces.'

Cyclic adenosine monophosphate (cAMP) (4) is a small molecule best known because its production just inside the surface of liver cells is the immediate



(4)

response to an extracellular signal from the hormone epinephrine (adrenaline) and the start of a chain of reactions inside the cell, which ends with greatly enhanced release of glucose to the blood.

CyclicAMPalsoappears,however,inthemorphogenesisofcellularslimemoulds (Plate 4). At the end of the stage of independent amoeboid single cells, some of these cells send out pulses of cAMP which act as signals to all the others to gather into aggregates.[65] The aggregates then differentiate into two types of cells, which rearrange into a stalk and a mass of spores. It is believed that, during the final rearrangement (culmination stage) cAMP may be produced by stalk cells and act as a chemotactic signal for the movements of these cells leading to stalk formation.[66] This is a kind of positive feedback; but on another level, the relation of cAMP to certain enzymes (adenylate cyclase, which produces cAMP, and phosphodiesterase, which destroys it), there is a possibility of finding a morphogen pair X and Y, consisting of a small molecule and a large one, quite likely to show the required imbalance of diffusivity.

C. Growth of Cell Surfaces.—Plant cell morphogenesis (Plates 2 and 3) involves expansion in area of both the lipid membrane and the cellulosic cell wall outside it. The lipids are neither formed at the surface nor added to it molecule by

[65] G. Gerisch, D. Hülser, D. Malchow, and U. Wick, *Philos. Trans. R. Soc. London, Ser. B*, 1975, **272**, 181. (A paper in a discussion on the physics and chemistry of biological recognition, which is all relevant to the present topic.) P. C. Newell, *Endeavour* (new series), 1977, **1**, 63.

[66] M. Sussman and R. Brackenbury, *Annu. Rev. Plant Physiol.*, 1976, **27**, 229.

molecule. They are conveyed from the interior of the cell in the form of vesicles up to 100nm in diameter.[67] These vesicles carry, as the electron microscope indicates, rosette-shaped structures from each of which a cellulose microfibril of a definite length grows after incorporation of the vesicle into the cell membrane. Thus the amount of cellulose produced goes in lock-step with the amount of material added to the membrane. The key to morphogenetic control is therefore, to my mind, the control rate of addition of vesicles. This still leaves a very wide field, chemically. Movement of vesicles through the interior of the cell is likely to be related to cytoplasmic streaming and could be controlled by the cytoskeleton of protein microfilaments. On the other hand, it may be that vesicles are always present at the surface in excess, and that most of them bounce off the membrane but some stick and fuse to it because of chemical differences in the surface of the membrane and the vesicles in different regions. One must then look for morphogens among the substances which act to set up gradients of these surface modifiers. The latter could be something as simple as Ca^{2+}, which is known to have a strong influence on electrical potentials at membranes and to affect, for example, the fusion of the sperm to the egg, and the fusion of neurotransmitter-filled vesicles to synaptic membranes in nerve cells. But the surface modifiers could also be something as complicated as a glycoprotein.

In multicellular plants, because of the rigidity of their cell walls and inability of the cells to sort out, control of direction of cell division is the primary morphogenetic control in development of branches, leaves, *etc.* (Plate 7).[68] The plane of division of a cell is first marked, before anything significant has happened to the geometry of the nucleus, by the appearance of a 'pre-prophase band' of microtubules, girdling the equator which is to be the plane of division.[12] This suggests that morphogenetic control mechanisms may reside at the cell surface rather than the nucleus, and leads me again to look a stage or more back from the microtubules and to enquire what substances and reactions determine the positions of nucleation sites for tubulin polymerisation. Chemicals that control plant growth were described in an earlier review in this series.[69] They include the cytokinins, which are purine and adenine derivatives, and which promote cell division and differentiation to the extent that an entire tobacco plant can be grown from a pith segment with the aid of 6-benzyloxypurine. Such compounds are promising candidates for the name of morphogen.

In general, structural studies by electron microscopy and other techniques such as immunofluorescence are providing extensive and spectacular evidence for the occurrence of microtubules and microfilaments in important places during development. I am inclined to believe, however, that these structures are manifestations of morphogenetic control, rather than the controllers themselves. But they are important as pointers towards the control mechanisms, which will probably ultimately be found in the biosynthesis of nucleation sites for these macromolecules. Either or both of nucleation-depletion theory and reaction-

[67] L. A. Staehelin, in ref. 37.
[68] P. B. Green, *Annu. Rev. Plant Physiol.*, 1980, **31**, 51.
[69] R. L. Wain, *Chem. Soc. Rev.*, 1977, **6**, 261.

diffusion theory of the Turing type could turn out to be the proper explanation of control.

Note added in proof: A recent discussion of the Royal Society on 'Theories of biological pattern formation' contains the most recent information on several of the topics in this review.[70]

[70] *Philos. Trans. R. Soc. London, Ser. B*, 1981, **295**, 425—617