Formation of gels and complexes by pairwise interacting polymers
P.G. Higgs, R.C. Ball

To cite this version:

HAL Id: jpa-00211143
https://hal.archives-ouvertes.fr/jpa-00211143
Submitted on 1 Jan 1989

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Formation of gels and complexes by pairwise interacting polymers

P. G. Higgs and R. C. Ball
Cavendish Laboratory, Madingley Road, Cambridge, CB3 0HE, G.B.

(Reçu le 28 mars 1989, accepté le 14 juin 1989)

Abstract. — Pairwise bonding between polymers in solution may lead to network formation — e.g., double helical junction zones in biopolymer gels, or association between polyacids and polybases. The equilibrium length distribution of bonded sequences to be expected in a branched network structure is obtained here by modification of helix-coil transition theory for linear chains. In the limit of infinite chains a « pairing transition » is observed at a finite temperature. At this point the mean length of the bonded sequences diverges and the system changes from a network to a solution of perfectly paired molecules. Subsequent aggregation of these pairs is then probable. Kinetic equations are developed to represent the helix formation reactions in gels and in single chains. These are compared to Monte Carlo simulations. In a network neighbouring bonded sequences will share only one of their two chains and may therefore not merge. Whilst the fraction of chain which is bonded rises smoothly to its equilibrium value, the number of bonded sequences per molecule (which is related to the network modulus) is found to pass through a maximum if the equilibrium bonded fraction is large. This suggests the possibility of high modulus gels as transient states during gelation.
1. Introduction. Pairwise interactions.

This paper addresses aggregation in polymer solutions in which pairs of molecules may form extended junction zones as illustrated in figure 1. Polyacids and polybases form complexes in aqueous solution which are stabilized by hydrogen bonds, e.g. poly-acrylic acid (PAA) and polyvinylpyrrollidone (PVP) or polyoxyethylene (PEO). Polysaccharides such as carrageenans and agarose form double helical structures, while gelatin gels contain triple helices of the protein collagen. These are again bound by hydrogen bonds. Similar structures occur between pairs of oppositely charged polyelectrolytes. The double helical structure of DNA is another familiar example of pairwise interacting polymers. The standard use of the terms « helix » and « coil » for the bound and unbound sections of a molecule will be followed, although not all of the paired structures mentioned above are actually helical in shape.

Many of these materials may exist either as gels or solutions depending on the concentration, temperature, length of the molecules and other factors. The strength of interaction between PAA and PEO is influenced by the pH of the solution. Addition of NaOH partially neutralizes the polyacid, shifting the dissociation equilibrium of the COOH groups. The larger the number of ionized COO⁻ groups the fewer hydrogen bonds may form.

Fig. 1. — Examples of pairwise interacting polymers.

Helical junction zones -

eg. Carrageenans
   Agarose
   Gelatin

Polyacid-Polybase complexes -

eg. PAA + PEO
    PAA + PVP

Polyelectrolytes -

eg. PMAA + PVBMA

...
and the weaker the interaction with the polybase. Iliopoulos and Audebert [1, 2] have identified three types of behaviour in the PAA-PEO system. At high degrees of neutralization the viscosity of the combined solution is what would be expected for two species of non-interacting polymers. At moderate degrees of neutralization the viscosity increases to several orders of magnitude above its non-interacting value — a gel is formed. At low degrees of neutralization the viscosity decreases to a fraction of the non-interacting value due to the formation of tightly bound compact clusters. Almost 100% of the length of the polymer is in the bound configuration in these clusters, whereas in the gels a much smaller proportion is bound.

Similar behaviour is observed in carrageenan solutions by Hermansson [3]. At temperatures greater than 50 °C the storage modulus $G'$ of the solution is negligible. In the range 45-50 °C a gel is formed and $G'$ passes through a sharp maximum. For $T < 45$ °C $G'$ is again much less than its maximum value. A qualitative explanation of this is illustrated in figure 2. At high temperatures (weak interactions) the solution contains mostly single molecules and a few loosely bound clusters. As $T$ is reduced the number of junction zones between molecules increases until the Flory-Stockmayer gelation point [4] is reached (mean functionality $> 2$). A highly branched gel is formed having a high modulus due to the large number of « elastically active strands ». Further decrease in $T$ causes the typical length of a helix to grow until it approaches the length of a molecule. The modulus and the mean functionality of a molecule again decrease. A series of electron micrographs in reference [3] clearly shows this process happening. At very low $T$ it would be expected that a solution of paired molecules should form, however in the carrageenan system it appears that the double helices themselves have a tendency to aggregate into « superstrands ». The superstrands form a low modulus gel with relatively few branch points.

Double helical pairs are likely to be much less flexible than single chains. Since the pairs have very little configurational entropy, if two pairs are placed side by side there is little configurational entropy to be lost. It therefore requires only weak attractive forces (such as Van der Waals forces) to cause the rod-like pairs to aggregate. However for the carrageenan it appears that this phase separation occurs only after the pairing reaction. In this paper only specific pairwise interactions will be considered, and no account will be taken of more general phase separation phenomena such as aggregation of rods into superstrands or the precipitation of polymers from solution.

A rod-like pair is expected to have a much larger hydrodynamic radius than a flexible random coil. A pair forming reaction might therefore lead to an increase in the viscosity of the solution. However Iliopoulos and Audebert [2] find a decrease in the viscosity under conditions where the binding is the strongest. This again suggests that pairs have aggregated into clusters. If aggregation of the single chains had occurred before the pair forming reaction there might be compact clusters formed, but these would not have the near 100% bonded fraction observed. Here again the phase separation appears to occur only after the pairing reaction, and it seems reasonable to consider the latter independently.

As a further experimental example, Tsuchida et al. [5] have studied the aggregation of the polyelectrolytes poly-methacrilic acid (PMAA) and poly-vinylbenzyltrimethylammonium chloride (PVBMA) which again leads to either gels or compact clusters. Light scattering experiments have been performed by ter Meer et al. [6] on gelatin at a concentration below that required for gelation. They conclude that almost linear aggregates are present. This is to be expected in the light of the above since if there were any appreciable branching the system would have gelled. A review of associating biopolymers has been given by Clark and Ross-Murphy [7], and many types of complex formation have been discussed by Tsuchida and Abe [8].
2. Factors limiting helix growth.

Information on the distribution of lengths of helix in a gel can be obtained by differential scanning calorimetry (DSC). The shorter the mean helix length the lower the mid-point transition temperature will be and the broader the melting curve. Djabourov et al. [9] have taken DSC measurements on gelatin gels which have been allowed to reach equilibrium at a range of different setting temperatures. Broader melting curves with lower mid point temperatures are observed at lower setting temperatures, indicating shorter helices exist at
lower temperatures. The same trends are found by Busnel et al. [10]. These gels typically take a very long time to reach equilibrium. Djabourov [9] has observed the increase in helix fraction during gelation by optical rotation measurements, and finds it is still increasing slowly after $10^3$ hours. Long time scale processes are also shown up by the appearance of double peaks in the DSC curves [10] which are strongly dependent on the temperature history of the sample.

Four factors which could govern the helix growth process are discussed below in the light of the above experiments.

(i) Chain Ends — Busnel et al. [10, 11] have done a simple computer simulation in which helices are initiated at random points on two molecules (or three for triple helices) and allowed to grow until the end of one of the chains is reached. Thus is similar to the « one sequence approximation » discussed in the books on helix coil transformations by Poland and Scheraga [12] and Poland [13]. The chance of finding a helical sequence of length $j$ linking two molecules of length $N$ is

$$p_j \propto (N - j + 1)^2 s^j$$  \hspace{1cm} (2.1)

where $s$ is the equilibrium constant for the bonding of one extra pair of monomer units at the end of a helix. If $s > 1$ the distribution is sharply peaked with the maximum very close to $j = N$. This may be relevant to the pair formation which we expect when the interactions are strong, but is not relevant to gels, for which there must be more than one helix per molecule. The one sequence model predicts the opposite trend for the melting curve to that observed, since at lower temperatures $s$ is larger and the mean $j$ moves closer to $N$.

(ii) Interaction between neighbouring helices — if several helices have been initiated on one molecule they will impede each other's growth. Two helices growing toward each other as in figure 3 may not merge since they share only one chain. In a single chain model (e.g. formation of $\alpha$ helix in a protein molecule) such merging would be allowed. The above experiments on gelatin suggest that helix initiation is favoured by lowering the temperature, and that a larger number of shorter helices is formed at lower $T$. It therefore appears that competition between neighbouring helices is an important feature.

The Zimm-Bragg model [14] assigns a statistical weight $\sigma s^j$ to a helix of length $j$ relative to a coil sequence of the same length. The unfavourable nature of the end sites of a helix is modelled by the initiation parameter $\sigma$ which is usually much less than 1. The mean helix length is a decreasing function of $\sigma$. However this model deals only with linear chains. It will be extended below to deal with branched networks.

(iii) Non-bonding regions on a chain — Molecules may contain specific monomers or sequences which are unable to bond. In some cases this may simply act to lower the average bonding energy per unit length. In other cases the non bonding points may act as blocks over which the paired sequence may not continue. Iliopoulos and Audebert [15] have produced a

![Fig. 3. — Growing helices in a network may not merge if they share only one molecule between them.](image)
semi-quantitative model to explain the effect of the number of ionized COO$^-$ groups on the aggregation between polycarboxylic acids and polybases (experiments discussed above). Bonded sequences may only exist between non-ionized groups. A minimum sequence length is required for stability, and the number of possible uninterrupted sequences decreases rapidly with the number of ionized groups. The helix fraction is therefore very sensitive to the proportion of ionized groups. A full explanation of this effect would need to account for the dissociation equilibrium of the acid groups. The proportion of ionized groups must be calculated self-consistently with the helix fraction. Since ionized groups only exist in non-bonded sequences they cannot be assumed to be randomly distributed along the whole chain.

Gelatin contains a relatively high proportion of proline residues. These may exist in either a low energy trans configuration or a high energy cis configuration. The cis proline introduces a "kink" into the molecule beyond which the helix may not continue until the proline happens to flip to the trans state. It has been proposed [10, 11] that this reaction plays a large part in controlling helix growth in gelatin. It is probably very important in the gelation kinetics, but the following rough argument suggests that this is not a dominant effect at equilibrium.

Suppose that possible blockage points exist on a chain which are in thermal equilibrium. The number in the higher energy non-bonding state is therefore of the form

$$N(T) = \frac{N_0}{1 + \exp(\Delta E/kT)}$$

where $\Delta E$ is the energy difference between the states. The mean spacing between blockage points is therefore

$$L(T) = L_0(1 + \exp(\Delta E/kT)).$$

If these points are the dominant factor limiting helix growth then the mean helix length would also be expected to be a decreasing function of $T$, which is contrary to observation. For the rest of the paper it will be assumed that the effect of single non-bonding monomers can be simply represented by reducing the mean bonding energy per unit length.

(iv) Reduction of coil entropy — In single chain helix-coil models the statistical weight of a coil sequence is usually set to 1 independent of length. In models for double stranded nucleic acid this becomes length dependent due to the entropy change when a loop is closed. (See discussion of "perfect matching" and "imperfect matching" models in [13, 14]. Loop entropy also has an important effect on the "two-chain coiled coils" models for tropomyosin pairs considered by Skolnick [16].) In a gel the weights for the coil sequences may again be length dependent even if these small closed loops are not considered. A coil between two helices may be "reeled in" in both directions as the helices grow, producing a reduction in coil entropy per unit length. This effect depends on the physical spacing of the junction zones and is therefore difficult to treat mathematically. It is likely that in most gels there is sufficient freedom for the molecules to rearrange themselves to prevent such "pulling tight" from occurring. This effect is probably unimportant at equilibrium but may be significant in the kinetics. The presence of such reel-like junction zones has been shown to have important effects on the stress-strain behaviour of the gel [17]. The unwinding of junction zones may provide a mechanism for gel failure.

Thus the most important feature considered above appears to be the interaction of the helices on any one molecule. A model will now be developed which illustrates these effects on the gelation kinetics and equilibrium network structure.
3. Equilibrium solution for infinite chains.

The equilibrium distribution for a single chain may be derived from a combinatorial representation of the partition function \([12, 13]\). Let there be \(m_j\) coil sequences of length \(j\) each of which contributes a factor \(u_j\) to the partition function, and \(n_j\) helical sequences each contributing a factor \(v_j\). In the Zimm-Bragg model these are \(u_j = 1\) and \(v_j = \sigma s^j\). The problem is equivalent to a 1 dimensional Ising model in a magnetic field. Let the length of the molecule be \(N\), and let the total number of sequences of each type be \(G\) (which has yet to be determined). The \(m_j\) and \(n_j\) must satisfy the constraints

\[
\begin{align*}
\sum_j m_j &= G \\
\sum_j n_j &= G \\
\sum_j j(m_j + n_j) &= N .
\end{align*}
\]

The partition function for a given distribution of sequence lengths is

\[
Z(G, m_j, n_j) = \left( G! \prod_j \frac{u_j^{m_j}}{m_j!} \right) \left( G! \prod_j \frac{v_j^{n_j}}{n_j!} \right).
\]

The total partition function is the sum over all possible distributions.

\[
Z_{\text{tot}} = \sum_G \sum_{\{m_j, n_j\}} Z(G, m_j, n_j) .
\]

In the limit \(N \to \infty\) and \(G \to \infty\) this sum is dominated by its maximum term. The solution for the infinite chain is found by maximizing (3.2) subject to constraints (3.1). This gives the exponential distribution of sequence lengths

\[
\begin{align*}
m_j &= G (\lambda - 1) \left( \frac{1}{\lambda} \right)^j \\
n_j &= G (\frac{\lambda - s}{s}) \left( \frac{s}{\lambda} \right)^j
\end{align*}
\]

where \(\lambda\) is the largest root of

\[
(\lambda - 1)(\lambda - s) = \sigma s .
\]

A branched network structure will now be treated in the same way. In the limit where all the chains in a network are long the ends of the chains become unimportant and it is sufficient to consider one long chain of length \(N \to \infty\) linked to itself many times. If there are \(G\) single bonding sequences then \(G/2\) double helix sequences will be formed. The number of distinct ways of pairing the sequences of length \(j\) is

\[
\text{number of pairings} = \frac{n_j!}{(n_j/2)! 2^{n_j/2}} .
\]

We assume that the sections to be paired are not close together along the chain and therefore neglect the entropy of closure of small loops. There is a factor \(a^3/V\) for every pair, where
$V$ is the system volume and $a$ is some molecular length scale, which is the entropy penalty for forcing one member of the pair to be localized close to the other. The network partition function is now

$$Z_\text{net} = \left( \frac{a^3}{V} \right)^{G/2}$$  \hspace{1cm} (3.6)

assuming again that the maximum term dominates the sum equivalent to equation (3.3). We will set $u_j$ to 1 again, but $v_j$ will be set to $(\sigma_0 s^j)^{1/2}$ so that the pair has weight $\sigma_0 s^j$ relative to its two constituent coils. The constant $\sigma_0$ includes the helix end effects as in the single chain model. Other factors which occur once per helix are incorporated into a definition of an effective $\sigma$ value below. In appendix A equation (3.6) is maximized with the same constraints (3.1), resulting in the distributions

$$m_j = G(\lambda - 1) \left( \frac{1}{\lambda} \right)^j$$

$$n_j = G \left( \frac{\lambda^2 - s}{s} \right)^j \frac{1}{\lambda}$$  \hspace{1cm} (3.7)

where $\lambda$ is the largest root of

$$\lambda (\lambda - 1)(2\lambda^2 - \lambda - s) = \sigma s.$$  \hspace{1cm} (3.8)

The factor $1/\lambda$ derives from the chemical potential for monomers, and therefore occurs once per monomer. The length distribution of paired helices varies as $(s/\lambda^2)^i$ instead of $(s/\lambda)^i$ as in the single chain case. The same feature occurs in some models representing double stranded nucleic acids [12] but these still represent one dimensional structures rather than the branched network above. In equation (3.8) $\sigma = \sigma_0 Na^2/V$ and is the concentration dependent effective $\sigma$ value which represents all factors which occur once per helix (see the derivations in appendices A and B). In addition to the end effects it includes the loss of translational entropy which occurs when two molecules are connected, and is proportional to the concentration in the present model. The $s$ factor remains independent of concentration. It represents the bonding energy and the configurational entropy change per monomer pair.

The helix fraction $\theta$ and the number helices per unit length $g = G/N$ may be determined from (3.7) and (3.8) as shown in appendix A. In figure 4 $\theta$ and $g$ are plotted as functions of $s$ for one constant value of $\sigma$. For the single chain model $\theta$ follows a « sigmoid » curve, tending to 1 as $s \to \infty$ and to 0 as $s \to 0$, whereas the network model shows a transition to a 100% helical state at a finite value of $s$. In both models $g(s)$ has a maximum at $s = 1$. The single chain model is symmetric in that $g(s) = g(1/s)$, whereas for the network model $g \to 0$ at the transition point.

The transition occurs when the largest root of (3.8) is at $\lambda = s^{1/2}$. From (3.7) it can be seen that the mean helix length then diverges, whilst the mean coil length remains finite. Hence $\theta \to 1$. Formation of a new helical sequence is intrinsically unfavourable due to the small $\sigma$ factor. However there is a certain amount of entropy to be got back due to forming the helix pairs. As $s$ increases the system wishes to form fewer sequences. The fewer sequences there are the fewer ways of forming the pairs, and the less entropy can be regained. Thus the longer the helices become the greater is the penalty for forming a new one. It therefore seems reasonable that at some finite $s$ the system becomes one single helix pair.

Substituting $\lambda = s^{1/2}$ into equation (3.8) implies that the pairing transition occurs at

$$s = (1 + \sigma^{1/2})^2.$$  \hspace{1cm} (3.9)
Fig. 4. — Variation of the helix fraction $\theta$ and the number of helical sequences per unit length $g$ as a function of $s$ at constant $\sigma = 2 \times 10^{-2}$. The network model shows a pairing transition beyond which $\theta = 1$. In both models $g$ is maximum at $s = 1$.

In the limit $\sigma \to 0$, $\theta(s)$ becomes a step function at $s = 1$. If $s$ is held constant at some value greater than $1$ then the paired state exists for $\sigma^{1/2} < s^{1/2} - 1$, whereas for the single chain $\theta \to 1$ only as $\sigma \to 0$. In one dimension phase transitions may only occur if the interactions are
of sufficiently long range, as demonstrated by Poland and Scheraga [18, 12]. For instance the perfect matching model shows a phase transition at a value of $s$ less than 1, below which $\theta = 0$. In this model long range interactions occur due to the entropy of loop closure. Skolnick has considered models in which each molecule in a pair may form helices independently, but in which there is an additional attractive reaction between helices on the two chains [16]. When loop closure entropy is accounted for then loops again become negligible if the interactions are too weak. The network model given here, however, should not be considered as a one dimensional model with long range interactions. The transition is between a branched structure and a truly one dimensional structure.

From figure 4 it can be seen that both $1 - \theta$ and $g$ tend linearly to zero at the transition point. If in the neighbourhood of the transition point $s = (1 + \sigma^{1/2})^2 - \Delta s$ then it is found from (3.8), (A3) and (A4) that

$$\theta \approx 1 - \frac{\Delta s}{\sigma^{1/2}}$$

$$g \approx \frac{\Delta s}{2}.$$  

For simplicity we have also assumed that $\sigma^{1/2} \ll 1$. The classification of this phase transition is unclear. If we identify $1 - \theta$ as an order parameter then the critical exponent $\beta = 1$. As a mean field value this is quite distinct from magnetic transitions ($\beta = 1/2$), but matches percolation on a Bethe lattice [19].

Thus for infinite chains the system forms a gel if $s < (1 + \sigma^{1/2})^2$ and forms pairs otherwise. As $s$ is reduced below 1 the helices become shorter and farther apart, however an infinite chain will always have an infinite number of helices. Finite chain length has an important effect for low $s$ where the gel is close to « melting ». The effect of finite chains is investigated in section 5, but first the gelation kinetics will be considered.


Skolnick [20] and Tanaka et al. [21] have considered the kinetics of helix coil transitions using the method of Glauber [22] for 1d Ising models. A spin $\mu_i$ is assigned to each site $i$ and differential equations may be written relating the average spin $\langle \mu_i \rangle$ and the correlations between nearby spins $\langle \mu_i \mu_{i+1} \rangle$ and $\langle \mu_{i-1} \mu_i \mu_{i+1} \rangle$ etc. as functions of time. An alternative method is to write down equations linking the rate of change of the number of helices $n_i$ of a given length $i$ to all the other $n_j$. This has been done by Chay and Stevens [23], who considered unzipping of polynucleotides. Both methods produce an infinite series of coupled equations, and it is necessary to make approximations to reduce this to a manageable set. In the spin models this involves neglecting higher order correlations. Craig and Crothers [24] have used a « matrix iteration » method in which the probability of occurrence of each possible sequence of length three is stored, and all lengths greater than three are treated only in a mean field way. This produces surprisingly good agreement with Monte Carlo simulations of the complete chain.

Most of the above papers have concentrated on the breakup of helices when the temperature is raised. Here the kinetics of helix formation and gelation is considered. The spin models are not readily generalizable to gel formation which is no longer one dimensional. The approach used here will be closer to the second method. The model below will be termed the First Order Initiation (FOI) model since the rate of helix initiation will be proportional to the number of initiation sites available. This model correctly describes the inability of helices in a gel to merge (Fig. 3) but remains a purely one dimensional model. The FOI model may
be easily simulated by Monte Carlo methods. The MC program will be used to check the validity of the approximations used in solving the differential equations, and these same approximations will then be used in a model for the branched network which it is not possible to simulate directly.

The rules for the MC simulation of the FOI model are summarized here.

\[
\begin{align*}
\text{CCC} & \Rightarrow \text{CHC} \\
& \quad \text{with probability } k_1 \\
\text{CCH} & \Rightarrow \text{CHH} \\
& \quad \text{with probability } k_2 \\
\text{HCH} & \Rightarrow \text{HHH}.
\end{align*}
\]

The rules have been chosen so as to produce the Zimm Bragg model for the single chain as the equilibrium solution. The constants \( k_1 \) and \( k_2 \) convert the equilibrium constants \( s \) and \( \sigma \) into reaction rates. The equilibrium solution does not depend on \( k_1 \) and \( k_2 \). In the simulations they have been set to 1 to reduce the number of parameters. A site is chosen at random and its state is changed with relative probabilities given by (4.1) which depend on the states of its nearest neighbours. Sites at the end of the chain are treated as though there were an extra site next to them which is permanently in the coil state.

It is assumed that at \( t = 0 \) the system is quenched from a very high temperature at which \( \theta = 0 \) to a temperature at which the reaction may proceed. In figure 5 the results of the simulation for chains of length 60 are shown. The helix fraction is plotted against \( \log_{10} \) time. The helix fraction increases smoothly to its maximum value, whilst the number of helices per unit length passes through a maximum for \( s > 1 \). This overshooting in \( g \) is easily understandable. Initially there is a large expanse of chain available and therefore many helices are initiated. These grow rapidly until they begin to interact with each other. If \( s \) and \( \sigma \) are such that the mean helix length at equilibrium is large then the only way a helix may grow is to wait until its neighbours happen to shrink to zero length by random fluctuations. The rate of increase of \( \theta \) then becomes very small, whilst \( g \) will gradually decrease.

Differential equations representing this same FOI model may be written down as follows. Let the coil fraction be \( f = 1 - \theta \). There are \( g \) helices and \( g \) coil sequences per unit length on average. The rate of initiation of helices in the FOI model is directly proportional to the number of available sites. For each helix there is one \( C \) site at either end on which initiation may not occur. However if two helices are separated by only one \( C \) site then this site will be counted twice. Therefore the fraction of available sites is \( (f - 2g + gq_1) \), where \( q_1 \) is the probability that a randomly chosen coil sequence is of length 1. A helix may always shrink, but it may only grow if the next coil sequence is of length 2 or more. Therefore the growth rate is proportional to \( 1 - q_1 \). If \( p_1 \) is the probability that a randomly chosen helical sequence is of length 1, then the following equations represent the FOI model of the infinite chain.

\[
\begin{align*}
\frac{df}{dt} &= -\sigma sk_1(f - 2g + gq_1) + g((1 - p_1)k_2 + p_1k_1) - gsk_2(1 - q_1) \\
\frac{dg}{dt} &= +\sigma sk_1(f - 2g + gq_1) - gp_1k_1.
\end{align*}
\]
Fig. 5. — Monte Carlo simulation of the FOI model with no helix merging. System is quenched to a constant temperature at time zero. Chain length \( N = 60, \sigma = 2 \times 10^{-2} \). The \( s \) values are labeled. A max in \( g(t) \) occurs for \( s \gg 1 \).

In appendix B it is demonstrated that the Zimm Bragg model is the equilibrium solution of (4.2). In addition there is an infinite series of equations of the form

\[
\frac{dp_j}{dt} = -p_j k_2 (1 + s (1 - q_1)) + p_{j-1} k_2 s (1 - q_1) + p_{j+1} k_2 .
\]  (4.3)
We wish to use simplest possible set of equations which will suffice. At equilibrium the mean helix length is \( \lambda / (\lambda - s) \) and the mean coil length is \( \lambda / (\lambda - 1) \).

Therefore

\[
f = \frac{g\lambda}{\lambda - 1}, \quad 1 - f = \frac{g\lambda}{\lambda - s},
\]

but at equilibrium \( p_1 = (\lambda - s) / \lambda \) and \( q_1 = (\lambda - 1) / \lambda \), so

\[
p_1 = \frac{g}{1 - f}, \quad q_1 = \frac{g}{f}.
\]

It will be assumed that the relationships (4.5) are approximately true at all times. Substitution of these into (4.2) gives a closed pair of coupled equations in \( f \) and \( g \). This substitution is exact at \( t = 0 \) and at \( t \to \infty \), so it is to be expected that it will not be too bad an approximation at intermediate times. The functions \( g(t) \) and \( \theta(t) \) can be found by standard numerical methods from equations (4.2) using the initial conditions \( f = 1 \) and \( g = 0 \). The result is compared to the M.C. model in figure 6. Equations (4.2) represent an infinite chain. An approximate set of equations for finite chains is discussed in section 5. These results are also shown in figure 6. There is a small difference in the equilibrium values of \( \theta \) and \( g \) for the different length chains. The equilibrium values for \( N = 60 \) lie within the uncertainty of the M.C. results, which are also for \( N = 60 \). All the differential equation solutions differ from the MC method slightly at intermediate times due to the approximation of equation (4.5). The difference is, however, small, and the curves all show the same basic shape. Differential equation methods will therefore be used in the rest of the paper since they require only a fraction of the time to compute.

The third equilibrium of (4.1) may be altered to permit merging and break-up of helices. This corresponds to a truly single chain model such as formation of \( \alpha \) helix on a protein molecule.

\[
\frac{k_3 s}{k_3 \sigma} \quad \text{HCH} \quad \Rightarrow \quad \text{HHH}.
\]

The corresponding differential equations for this FOI model with merging helices are

\[
\frac{df}{dt} = -\sigma sk_1 (f - 2g + gq_1) + g((1 - p_1)k_2 + p_1k_1) - gs((1 - q_1)k_2 + q_1k_3) + \sigma k_3(1 - f - 2g + gp_1)
\]

\[
\frac{dg}{dt} = \sigma sk_1 (f - 2g + gq_1) - gp_1k_1 - gsq_1k_3 + \sigma k_3(1 - f - 2g + gp_1)
\]

This again has the Zimm-Bragg model as its equilibrium solution (appendix B).

To represent the branched network properly it is necessary to use a Second Order Initiation (SOI) model. Merging may not occur in a network, therefore the equations are similar to (4.2). For a helix to be initiated an available site must be found on two separate chains (assumed independent), therefore the \((f - 2g + gq_1)\) factor is squared. The initiation reaction is second order, therefore the rate per site is proportional to the concentration \( C \). Growth of a helix may only occur if both the coil sequences leading from its end are longer than 1. The \((1 - q_1)\) factor is therefore squared. Hence

\[
\frac{df}{dt} = -\sigma_0 sk_0 C(f - 2g + gq_1)^2 + g((1 - p_1)k_2 + p_1k_1) - gs k_2(1 - q_1)^2
\]

\[
\frac{dg}{dt} = +\sigma_0 sk_0 C(f - 2g + gq_1)^2 - gp_1k_1.
\]
Fig. 6. — FOI model with no helix merging, \( s = 1.1, \sigma = 2 \times 10^{-2} \). Comparison of Monte Carlo simulation \((N = 60)\) with differential equation solutions for \(N = 60, 600, \infty\). The approximations used in the kinetic equations appear reasonably good for the whole time range.

It is shown in appendix B that the equilibrium solution of these equations is the sequence length distribution for the network derived in section 3, if we define a concentration dependent value of \(\sigma\) by

\[
\sigma = \sigma_0 k_0 C / k_1 .
\]  

(4.9)
This distribution may therefore be obtained both from a partition function and as the equilibrium solution of a set of kinetic equations. The approximation (4.5) is still appropriate for equations (4.7) and (4.8).

Figure 7 compares the three models for infinite chains. In the SOI model $k_1$ and $k_2$ have again been set to 1, and $k_0$ is included in the definition of $\sigma$. The FOI models with and

---

**Fig. 7.** — Comparison of the FOI model for single chains (with and without merging) to the SOI model for the network, $s = 1.15$, $\sigma = 2 \times 10^{-2}$. 
Fig. 8. — The SOI model for $\sigma = 10^{-3}$, $s$ values labelled. The behaviour of $\theta$ and $g$ is very sensitive to $s$ at smaller $\sigma$ values. The $s = 1.08$ curve has $\theta \to 1$, since this is beyond the pairing transition at this value of $\sigma$. The approach to equilibrium is in general slower, and the overshoot in $g$ is more pronounced for the SOI model than the FOI models.

without merging appear to differ very little. The curve with merging reaches equilibrium earlier than the one without, as would be expected. The overshoot in $g$ is less pronounced when merging is allowed. It would be wrong to conclude that these two models are always so similar. Adjusting the rate constants $k_1$ and $k_2$ could produce large differences between the
two at intermediate times without affecting the equilibrium values. In principle these constants could be found by fitting experimental measurements. The qualitative conclusion that inability to merge slows down the attainment of equilibrium remains valid.

In general in the SOI model equilibrium is reached more slowly than in the FOI models, and the overshoot in \( g \) is more pronounced. Figure 8 shows the SOI model for \( \sigma = 10^{-3} \). The curves are much more sensitive to \( s \) for smaller \( \sigma \) values. The curve for \( s = 1.08 \) has \( \theta \rightarrow 1 \) since this is above the pairing transition at this value of \( \sigma \).

In summary — in all the models \( g \) increases steadily towards its equilibrium value when the equilibrium helix fraction is small and overshoots when the equilibrium helix fraction is large. The crossover occurs at about \( s = 1 \) in each case. The helix fraction always increases smoothly towards its equilibrium value.

Experiments on carrageenans by Hermansson [3] have measured the storage modulus \( G' \) during gel formation. In this case the interaction strength is governed by the concentration of potassium ions in solution. In some cases \( G' \) passes through a maximum. The strong highly-branched gel is formed as a transient state under conditions where the weak gel (formed from superstrands of aggregated pairs) is the equilibrium state. In a Gaussian network the modulus is proportional to the number of elastically active strands and hence to the number of crosslinks. The elasticity of biopolymer gels probably has little in common with that of ideal Gaussian networks, however it is to be expected that a gel with a large number of effective crosslinks will have a larger modulus than one with a smaller number. It is therefore reasonable to draw an analogy between the behaviour of \( g \), the number of « crosslinks » in the network model above, and the behaviour of the experimental modulus. Not all biopolymer experiments find \( G \) passing through a maximum. It should be remembered that in the carrageenans superstrand formation is probably proceeding in some regions at the same time as the first helix formation reaction is occurring in others. Predicting the modulus of such a mixed system is not easy. We cannot necessarily expect a direct relation between the structure of a gel on the molecular scale (as revealed by electron micrographs, for example) and the mechanical properties. In a gel there are likely to be inhomogeneities in the structure on a much larger scale which may also significantly affect the modulus.

The kinetics of gelatin aggregation at low concentrations has been studied by light scattering [6] and optical rotation [11]. These experiments demonstrate that the reaction mechanism is strongly dependent on concentration, and suggest that helices are formed by a first order intramolecular folding process at low concentration and by a second order intermolecular process at higher concentrations. The increase in helix fraction in gelatin gels has been followed by Djabourov et al. [9] and Chatelier et al. [23] over a long time range. After an initial rapid increase, \( \theta \) continues to increase in proportion to \( \log t \). An equilibrium value of \( \theta \) is never reached within the experimental time scale. This very slow increase at long times cannot be explained by any of the models considered here since they all decay exponentially towards a constant value as \( t \rightarrow \infty \). The experiments suggest that topological effects are important — the achievement of a state with high helix fraction requires substantial rearrangement of the molecules in space. This rearrangement would take a long time if the molecules were already strongly bound in several places.

5. Finite chains.

Sites at the end of finite chains always have a lower chance of being bound than those in the centre, since unwinding from the ends is relatively easy. Therefore the helix fraction decreases as the length decreases. This effect is shown clearly by the polyacid-polybase
complexes studied by Osada [26]. The equilibrium solution for a single finite chain may be obtained exactly by the matrix method [12]. Equation (3.5) is the characteristic equation of a matrix with eigenvalues \( \lambda \). Both the larger and smaller eigenvalues become important for finite chains. Since no matrix formulation of the network model is available, an approximate treatment of finite length chains is used here which works for both single chain and network.

Rather than fixing the absolute length of all the chains to be \( N \), let there be an exponential distribution of chain lengths with mean length \( N \). Let the mean number of C sites be \( F = N f \) and the mean number of helical sequences per chain be \( G = N g \). It is always possible for a helix to come unwound from its ends therefore a typical finite chain will always be terminated by coil sequences. Therefore assume the mean number of coil sequences per molecule is \( G + 1 \).

The FOI model with no merger becomes

\[
\begin{align*}
\frac{dF}{dt} &= -\sigma sk_1(F - 2G + (G + 1)q_1) + G((1 - p_1)k_2 + p_1k_1) - Gsk_2(1 - q_1), \\
\frac{dG}{dt} &= +\sigma sk_1(F - 2G + (G + 1)q_1) - Gp_1k_1
\end{align*}
\]

and the approximation for \( p_1 \) and \( q_1 \) is

\[
p_1 = \frac{G}{N - F}, \quad q_1 = \frac{G + 1}{F}.
\]

The results from these equations have already been shown in figure 6 and are in reasonable agreement with the Monte Carlo method (see Sect. 4). It therefore seems reasonable to use the same approximation procedure for the network. The equilibrium solution of the equivalent finite chain SOI equations is found in appendix C. This result may be used to construct a phase diagram of pairwise interacting polymers. In appendix C \( G \) is calculated as a function of \( s \), \( \sigma \), and \( N \). The Flory-Stockmayer gelation criterion — \( G \geq 2 \) — is used to provide a simple estimate of the boundaries of the gel region. For sufficiently large \( \sigma \) and \( N \) there are two values of \( s \) at which \( G = 2 \). Between these values \( G \geq 2 \) and there is sufficient connectivity for a gel to exist. For smaller \( \sigma \) the gel region becomes narrower, until for \( \sigma \leq \sigma_{\text{min}} \) \( G < 2 \) for all \( s \). A numerical calculation of the boundary line for two values of \( N \) is shown in figure 9. There is no well defined phase transition between gel and paired structures for finite chains. However, \( G \) falls sharply close to the point where \( g \to 0 \) for the infinite chain. The position of the gel-pairs boundary is calculated approximately in appendix C. The leading term of this approximation is independent of \( N \), indicating that the upper boundary of the gel region is fairly insensitive to chain length. The lower boundary has the approximate equation

\[
s \approx \frac{1}{1 + \sigma N/2}
\]

providing \( \sigma \gg \sigma_{\text{min}} \). The point at which the gel melts is therefore strongly dependent on \( N \). Since \( \sigma \) is an increasing function of concentration the existence of a minimum \( \sigma \) implies a minimum concentration for gelation. The value of \( \sigma_{\text{min}} \) is proportional to \( 1/N^2 \) for large \( N \).

In this theory it has been assumed that the entropy loss when forming a new helix can simply be represented by one constant factor in the partition function proportional to \( C \), which is incorporated into the definition of \( \sigma \). In reality the entropy loss depends on the concentration in a more complicated way and also on the cluster size distribution. To calculate this properly would require an elaborate self-consistent argument. Converting from the
Fig. 9. — The behaviour of pairwise interacting polymers for different $s$ and $\sigma$. The symbols show the numerically calculated sol-gel boundary line for $N = 60$ and $N = 600$. The minimum $\sigma$ values are $8.6 \times 10^{-3}$ and $8 \times 10^{-5}$, respectively. The solid curves show the approximately calculated boundary line for $N = 600$ (Eqs. (C4) and (C6)).

Fig. 10. — Schematic phase diagram for changing interaction strength and concentration. $AF = $ free energy change on binding one monomer pair.
variables $\sigma$ and $s$ to the variables concentration and temperature or concentration and interaction strength is rather difficult to do beyond the approximations used here. Figure 10 summarizes the expected behaviour in terms of the free energy change $\Delta F_m$ per monomer and the concentration. The paired region, though labelled as a solution in figure 10, may form a gel or precipitate due to the aggregation of pairs.

It should also be remembered that formation of small loops by the reconnection of chains which are already connected has been neglected in this paper. If they were included we would still expect a pairing transition, but the pair would contain small unbonded loops within it (in the manner of two chain DNA models) and would therefore not have a 100% helix fraction. No attempt has been made to take account of the spacial arrangement of the chains in the gel. Chains in the centre of a cluster of connected molecules will have a reduced chance of reacting with new chains due to the screening of the outer molecules. These effects await a more detailed treatment.

**Conclusions.**

A model for the aggregation of pairwise interacting polymers has been considered. The fraction of bonded monomers is a decreasing function of temperature, whilst the number of bonded sequences per molecule passes through a maximum at intermediate temperature. For sufficiently long chains and sufficiently high concentrations the connectivity becomes large enough for gel formation at these intermediate temperatures.

The mean helical sequence length increases sharply at low temperatures. In the limit of infinite chains a pairing transition is observed at a finite temperature at which point the mean helix length diverges and the molecules become 100% bonded.

At temperatures above the pairing transition there is an exponential distribution of sequence lengths at equilibrium. This distribution has been obtained both from a combinatorial formulation of the partition functions, and as the equilibrium solution of a set of kinetic equations.

Several sets of kinetic equations have been presented which illustrate the difference between helix-coil transitions in a single chain and in a branched network. Attainment of the equilibrium state in a network is slowed due to the inability of neighbouring helical sequences on a molecule to merge, and because the helix initiation reaction is second order. Under some conditions the number of helices per molecule passes through a maximum before reaching equilibrium. This suggests the possibility of formation of high modulus gels as transient states.

**Appendix A.**

**Maximizing the network partition function.**

It is required to maximize equation (3.6) for $Z_{\text{net}}$ subject to constraints (3.1). Proceed by maximizing the function $F_{\text{net}}$.

$$F_{\text{net}} = \ln Z_{\text{net}} + \text{constraints}$$

$$= \sum_j \left\{ m_j \ln (u_j) - m_j \ln (m_j/e) + n_j \ln (v_j) - \frac{n_j}{2} \ln (n_j/2e) - \frac{n_j}{2} \ln 2 \right\}$$

$$+ 2G \ln (G/e) + \alpha \left( G - \sum m_j \right) + \beta \left( G - \sum n_j \right) + \gamma \left( N - \sum j(m_j + n_j) \right)$$

$$+ \frac{G}{2} \ln \left( a^3/\Delta \right)$$
where $\alpha, \beta, \gamma$ are Lagrange multipliers to be determined, and Stirling's approximation has
been used for the factorials.

\[
\frac{\partial F_{\text{net}}}{\partial m_j} = \ln u_j - \ln m_j - \alpha - \gamma j = 0
\]
\[
\frac{\partial F_{\text{net}}}{\partial n_j} = \ln v_j - \frac{1}{2} \ln \frac{n_j}{2} - \beta - \gamma j - \frac{1}{2} \ln 2 = 0
\]

\[
m_j = u_j e^{-\alpha - \gamma j}
\]
\[
n_j = v_j^2 e^{-2\beta - 2\gamma j}.
\]

Set $u_j = 1$, and $v_j^2 = \sigma_0 s^j$, and define $\lambda = e^\gamma$. From the first two constraints:
\[
e^{-\alpha} = G(\lambda - 1)
\]
\[
e^{-2\beta} = \frac{G(\lambda^2 - s)}{\sigma_0 s}
\]

which gives the distributions (3.7). From the third constraint:
\[
N = G \left( \frac{\lambda}{\lambda - 1} + \frac{\lambda^2}{\lambda^2 - s} \right). \tag{A1}
\]

So far $G$ has been assumed fixed. $Z_{\text{net}}$ must now be maximized with respect to $G$. This is done by maximizing in $Z_{\text{net}}$ w.r.t. $\lambda$ whilst maintaining relationship (A1).

\[
\frac{\partial \ln Z_{\text{net}}}{\partial \lambda} = \sum_j \frac{\partial m_j}{\partial \lambda} \left( \ln u_j - \ln m_j \right) + \sum_j \frac{\partial n_j}{\partial \lambda} \left( \ln v_j - \frac{1}{2} \ln n_j \right) +
\]

\[
+ 2 \frac{\partial G}{\partial \lambda} \ln G + \frac{1}{2} \ln \left( \frac{a^3}{V} \right) \frac{\partial G}{\partial \lambda} = 0
\]

\[
\frac{\partial G}{\partial \lambda} \left( \alpha + \beta + 2 \ln G + \frac{1}{2} \ln \left( \frac{a^3}{V} \right) \right) + \gamma \frac{\partial G}{\partial \lambda} \text{ « constant »} = 0
\]

\[
\alpha + \beta + 2 \ln G + \frac{1}{2} \ln \left( \frac{a^3}{V} \right) = \frac{1}{2} \ln \left( \frac{1}{G^2(\lambda - 1)^2 \frac{\sigma_0 s}{G(\lambda^2 - s) \frac{G^4 a^3}{V}}} \right) = 0
\]

\[
\lambda (\lambda - 1)(2\lambda^2 - \lambda - s) = \frac{N\sigma_0 a^3}{V} s = \sigma s. \tag{A2}
\]

This equation defines $\sigma$, which is concentration dependent and includes all factors in the partition function which occur once per helix. An alternative derivation of equation (A2) is given in appendix B.

From (A1) and (A2)
\[
g = \frac{G}{N} = \frac{(\lambda - 1)^2 (\lambda^2 - s)}{\sigma s} \tag{A3}
\]
\[
\theta = \frac{g \lambda^2}{\lambda^2 - s} = \frac{\lambda^2(\lambda - 1)^2}{\sigma s}. \tag{A4}
\]

The corresponding results for the single chain are
\[
N = G \left( \frac{\lambda}{\lambda - 1} + \frac{\lambda}{\lambda - s} \right)
\]
\[
g = \frac{\sigma s}{\lambda (2\lambda - s - 1)}, \quad \theta = \frac{\lambda - 1}{2\lambda - s - 1}.
\]
If $s > 1$ the largest root of (A2) at $\sigma = 0$ is the larger root of $2 \lambda^2 - \lambda - s = 0$. If $s < 1$ the largest root is $\lambda = 1$. In both cases the largest root increases with increasing $\sigma$, and all the other roots are always $\leq 1$. Therefore there is no ambiguity about which root to take.

Appendix B.

Equilibrium solutions of kinetic equations.

FOI MODEL (Eq. (4.2)). — At equilibrium detailed balance applies, therefore

$$n_j s(1 - q_1) = n_{j+1}.$$  

The ratio $n_j/n_{j+1}$ is independent of $j$ so there must be an exponential distribution of sequence lengths. Propose that

$$q_j = \frac{m_j}{G} = (\lambda - 1) \left( \frac{1}{\lambda} \right)^j$$

$$p_j = \frac{n_j}{G} = (\lambda' - 1) \left( \frac{1}{\lambda'} \right)^j$$

for some $\lambda$ and $\lambda'$ to be determined. At equilibrium $df/dt = dg/dt = 0$. Adding (4.2a) and (4.2b) gives

$$1 - p_1 - s(1 - q_1) = 0$$

from which $1/\lambda' = s/\lambda$ as required in (3.4). From (4.2b)

$$\sigma s g \left( \frac{\lambda}{\lambda - 1} - 2 + \frac{\lambda - 1}{\lambda} \right) - g \frac{\lambda - s}{\lambda}$$

which gives (3.5). An exactly similar procedure works with equations (4.7).

SOI MODEL (Eq. (4.8)). — Again propose (B1) because of detailed balance. Adding (4.8a) and (4.8b) implies $1/\lambda' = s/\lambda^2$. From (4.8b)

$$\sigma_0 s k_0 C g^2 \left( \frac{\lambda}{\lambda - 1} - 2 + \frac{\lambda - 1}{\lambda} \right)^2 - g k_1 \frac{\lambda^2 - s}{\lambda^2} = 0$$

$$\sigma_0 s k_0 C \frac{k_1}{k} = \frac{1}{g} (\lambda - 1)^2 (\lambda^2 - s)$$

but (mean helix length + mean coil length) = $1/g$ which is equivalent to (A1). Hence, defining $\sigma = \sigma_0 k_0 C/k_1$ we obtain equation (3.8) or (A2).

Appendix C.

Network of finite chains.

By analogy with (5.1) the SOI model for finite chains of mean length $N$ is

$$\frac{dF}{dt} = -\frac{\sigma_0 s k_0 C}{N} (F - 2G - (G + 1)q_1)^2 + G((1 - p_1)k_2 + p_1 k_1) - Gsk_2(1 - q_1)^2$$ (C1a)

$$\frac{dG}{dt} = +\frac{\sigma_0 s k_0 C}{N} (F - 2G - (G + 1)q_1)^2 - Gp_1 k_1$$ (C1b)
into which (5.2) may be substituted. In this case
\[ N = (G + 1) \frac{\lambda}{\lambda - 1} + G \frac{\lambda^2}{\lambda^2 - s} \]
\[ (N (\lambda - 1) - \lambda) (\lambda^2 - s) = G \lambda (2 \lambda^2 - \lambda - s). \] (C2)

Here exponential distributions of the form (3.7) have been assumed. This is valid if the molecules themselves have an exponential length distribution, but not if they are all strictly of length $N$. Defining $\sigma$ as in appendix B, we obtain from (C1b)
\[ \frac{\sigma s}{N} (G + \lambda^2 + (\lambda - 1)^2) = G (\lambda - 1)^2 (\lambda^2 - s). \] (C3)

To find the sol-gel boundary line pairs of values of $\sigma$ and $s$ need to be found which satisfy (C2) and (C3) with $G = 2$. This has been done numerically in figure 9. Approximate solutions can be found as below.

The boundary for $s < 1$ occurs when $\theta \to 0$ and $\lambda \to 1$. Let $\lambda = 1 + \varepsilon$, then from (C2)
\[ \varepsilon \approx \frac{3}{N} \quad \text{for} \quad N \gg 1, \]
and from (C3)
\[ \varepsilon^2 \approx \frac{9 \sigma}{2 N} \frac{s}{1 - s} \]
therefore
\[ \sigma \approx \frac{2}{N} \frac{1 - s}{s} \] (C4)

which is the same as (5.3). This is valid if $1 - s$ is not a small parameter. If $s = 1$, then (C3) gives
\[ \varepsilon^3 \approx \frac{9 \sigma}{4 N}, \quad \sigma_{\text{min}} \approx \left( \frac{16}{3 N} \right)^2. \] (C5)

The minimum $\sigma$ value does not correspond precisely to $s = 1$, since the maximum value of $G$ is not precisely at $s = 1$ for finite $N$. This is visible in figure 9 for $N = 60$. Therefore (C5) is only approximate.

On the boundary at $s > 1$, $\theta \to 1$ and $\lambda = s^{1/2} + \varepsilon$. (C2) implies $\varepsilon \approx s^{1/2}/N$, and combining this with (C3) gives
\[ \sigma \approx \frac{4 (s^{1/2} - 1)^2 \left( 1 + \frac{2 s^{1/2}}{N (s^{1/2} - 1)} \right)}{(3 - 2 s^{1/2} + 2 s)^2 \left( 1 + \frac{4 s^{1/2} (2 s^{1/2} - 1)}{N (3 - 2 s^{1/2} + 2 s)} \right)}. \] (C6)

This is not strongly dependent on $N$ when $N \gg 1$, unlike (C4). In figure 9 (C4) and (C6) are shown to provide good approximations when $N = 600$ over the range of $\sigma$ considered, and only fail for $\sigma$ close to $\sigma_{\text{min}}$.

Acknowledgments.

Paul Higgs wishes to thank Dr. Simon Ross-Murphy and Dr. Mike Cates for their suggestions, and is grateful for the support of the Agricultural and Food Research Council and Unilever Research.
References