# A statistical theory of directional isomerism in polymer chains and its application to polyvinylidene fluoride

**R. E. Cais and N. J. A. Sloane**
*Bell Laboratories, Murray Hill, New Jersey 07974, USA*
(*Received* 22 *March* 1982; *revised* 6 *May* 1982)

A formal statistical analysis is presented of structural isomerism in a polymer chain having head–head, head–tail and tail–tail orientations of a directional monomer unit (regioisomerism). We derive general relations between the probabilities of regiosequences with up to seven elements (heptads), and treat explicitly Bernoullian and first-order Markov models for chain growth. We then illustrate our theory by analysing 188 MHz $^{19}F$ n.m.r. spectra of two samples of polyvinylidene fluoride ($PVF_2$), Kynar 961 and Kureha KF-1100, and show that these polymers have regiosequence distributions which are described by first-order Markov statistics and not Bernoullian statistics. The implication is that the structural disorder in $PVF_2$ cannot be characterized accurately by a single parameter such as the 'percent of head–head, tail–tail defect content'. We find that the defect content in Kynar 961 is 5.0% and in Kureha KF-1100 3.7%, and that the appropriate reactivity ratio pairs for the first-order Markov model are $r_0 = 0.003$, $r_1 = 18$ and $r_0 = 0.003$, $r_1 = 24$, respectively.

**Keywords** Head–head; head–tail; tail–tail; isomerism; sequence statistics; polyvinylidene fluoride; Kynar 961; Kureha KF-1100; $^{19}F$ n.m.r.; reactivity ratios

## INTRODUCTION

Since the advent of high resolution nuclear magnetic resonance (n.m.r.) spectroscopy it has been possible to examine the microstructure of polymers in great detail. For example it is now a routine matter to determine the stereosequence distribution of homopolymers and the monomer sequence distribution of copolymers by this technique. Given the sequence distribution one can then formulate a statistical description of the chain growth process (e.g. Bernoullian, Markovian, etc.), and obtain a quantitative understanding of the influence of variables such as temperature and monomer feed ratio on polymer microstructure[1].

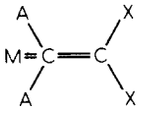
The statistical description of stereoconfigurational sequences (tacticity) in vinyl homopolymers has been formulated in two classic papers, by Coleman and Fox (1963)[2] and Frisch, Mallows and Bovey (1966)[3]. Likewise the theory of monomer sequence distributions in copolymers has been treated exhaustively by Price (1962)[4], Ito and Yamashita (1965)[5], and Pyun (1970)[6].

In contrast, scant theoretical attention has been paid to another form of structural isomerism in polymers, namely regioisomerism, i.e. directional isomerism arising from head–tail, head–head, and tail–tail additions of an asymmetrical monomer unit. Most polymers have a pure head–tail regioregular structure. However polymers from certain fluoroethylenes and dienes are notable exceptions, and a proper understanding of regiosequence statistics is required for their structural analysis[7].

Although a statistical description of regiosomerism can be viewed as a special case of monomer sequence statistics for binary copolymers, there are subtleties which warrant the formal treatment given here. These aspects have often been overlooked in polymers examined previously owing to their small fraction (typically less than 0.1) of inverted monomer units, which makes certain regioirregular sequences very difficult to observe. We illustrate our treatment by a quantitative analysis of $^{19}F$ n.m.r. spectra of polyvinylidene fluoride ($PVF_2$), and show that the regiosequence distribution follows first-order Markov statistics, and not Bernoullian statistics as recently reported[8].

## DEFINITIONS AND STRUCTURAL NOTATION

We restrict our attention to infinite linear chains constructed from a vinylidene monomer M, which is directional by virtue of an asymmetrical arrangement of substituents about the double bond, as illustrated below.




We do not consider stereoconfigurational isomerism, chain end effects, and branching, so the only structural defects involve regioirregular placements. According to conventional notation one end of M can be designated the 'head' (e.g. $=CX_2$), the other end the 'tail' (e.g. $=CA_2$). Thus a sequence of M units may have 'head–head' ($-CX_2-CX_2-$), head–tail' ($-CX_2CA_2-$), or 'tail–tail' ($-CA_2-CA_2-$) junctions[7].

By analogy with the established nomenclature for steroisomerism, we use the designations *isoregic*, *syndioregic*, and *aregic* for sequences in which the directional sense of successive monomer units is the same, alternating, and random, respectively. For convenience we represent the 'head' of M by 1 and the 'tail' by 0.

0032-3861/83/020179-09$03.00

*Table 1* List of all possible *n*-ads for $n \leq 7$. The complement of $s_n$, obtained by interchanging 0's and 1's, is denoted by $\bar{s}_n$

| Length *n* | Regiosequence $s_n$ | Regiosequence $\bar{s}_n$ |
|---|---|---|
| 1 | 0 | 1 |
| 2 | 00 | 11 |
| | 01 | 10 |
| 3 | 001 | 110 |
| | 100 | 011 |
| | 101 | 010 |
| 4 | 0010 | 1101 |
| | 0100 | 1011 |
| | 0011 | 1100 |
| | 1010 | 0101 |
| | 1001 | 0110 |
| 5 | 01001 | 10110 |
| | 01010 | 10101 |
| | 10010 | 01101 |
| | 11010 | 00101 |
| | 10011 | 01100 |
| | 01011 | 10100 |
| | 11001 | 00110 |
| 6 | 010010 | 101101 |
| | 010100 | 101011 |
| | 110100 | 001011 |
| | 100110 | 011001 |
| | 010110 | 101001 |
| | 110010 | 001101 |
| | 010011 | 101100 |
| | 010101 | 101010 |
| | 100101 | 011010 |
| | 110101 | 001010 |
| | 110011 | 001100 |
| 7 | 0100110 | 1011001 |
| | 0010101 | 1101010 |
| | 0100101 | 1011010 |
| | 0010110 | 1101001 |
| | 0110010 | 1001101 |
| | 1010010 | 0101101 |
| | 1010100 | 0101011 |
| | 0110100 | 1001011 |
| | 0110011 | 1001100 |
| | 1010011 | 0101100 |
| | 1010101 | 0101010 |
| | 0110101 | 1001010 |
| | 1100110 | 0011001 |
| | 1010110 | 0101001 |
| | 1100101 | 0011010 |

Consequently an *isoregic* ('head–tail') polymer has the regular binary sequence ....010101010101...., a *syndioregic* ('head–head, tail–tail') polymer has the sequence ....011001100110...., and an *aregic* polymer has an irregular sequence of 01 and 10 pairs[9].

## SEQUENCE OCCURRENCE PROBABILITIES AND WINDOW PROBABILITIES

We assume that the infinite sequences of 0's and 1's described in the previous section are generated by some random process, and will investigate the statistical properties of finite segments of these sequences. Since the infinite sequences are formed from 01's and 10's, not every finite segment can occur (e.g. 000 is forbidden).

Let $S_n$ denote the set of possible segments of $n$ 0's and 1's, or $n$-ads, that can occur. A typical $n$-ad will be denoted by $s_n$. The members of $S_1, \ldots, S_7$ are given in *Table 1*. It is simple to test if a particular $n$-ad $x_1x_2 \ldots x_n$ ($x_i = 0$ or 1) is in $S_n$. If it is then either all the pairs $x_1x_2, x_3x_4, x_5x_6, \ldots$ belong to the set $\{01,10\}$, or else all the pairs $x_2x_3, x_4x_5, x_6x_7, \ldots$ belong to $\{01,10\}$. There are two ways of dividing the $n$-ad into pairs, i.e. punctuating it, and at least one of them must produce 01's and 10's if the $n$-ad is in $S_n$. Most $n$-ads can be punctuated in only one way, i.e. they are unambiguous. There are only two ambiguous members of each $S_n$, namely 01010.... and 10101....

It is straightforward to see that the number of distinct $n$-ads in $S_n$ is:

$$N(S_n) = 2^{k+2} - 2 \text{ if } n = 2k+1$$

or

$$N(S_n) = 3.2^k - 2 \text{ if } n = 2k .$$

These numbers are given in *Table 2* for $n \leq 10$.

We shall associate three different probabilities with each $n$-ad $s_n = x_1x_2 \ldots x_n$, all $x_i = 0$ or 1. These are (i) the probability of occurrence of a punctuated $n$-ad $x_1x_2, x_3x_4, \ldots$, which will be denoted by $\pi(s_n)$; (ii) the probability of observing an unpunctuated $n$-ad through a window of length $n$ placed randomly on the infinite sequence, assuming that the direction of the sequence is known, which will be denoted by $p(s_n)$; and (iii) the probability $P_{obs}(s_n)$ of observing an unpunctuated $n$-ad through a randomly placed window when the direction of the sequence is unknown. These quantities will be defined in more detail in what follows.

Our basic assumption is that the sequences satisfy the following stationarity condition. If $X = \ldots, X_{-1}X_0, X_1X_2, X_3X_4, \ldots$ is one of our (punctuated) random sequences, then the probability that a particular $n$-ad $s_n = x_1x_2, x_3x_4, \ldots, x_{n-1}x_n$ ($n$ even) occurs in X beginning in position $t$ is independent of $t$. In other words we assume that the probability:

$$\text{Prob}\{X_t = x_1, X_{t+1} = x_2, \ldots, X_{t+n-1} = x_n\}$$

is a function

$$\pi(x_1x_2, x_3x_4, \ldots, x_{n-1}x_n) = \pi(s_n)$$

which is independent of $t$. Thus $\pi(s_n)$ is the unconditional probability of occurrence of the punctuated $n$-ad $s_n$. There is then a natural way to define $\pi(s_n)$ for any punctuated $n$-ad. For example, $\pi(1,10) = \pi(01,10)$, $\pi(1,10,1) = \pi(01,10,10)$, and so on.

If we examine a segment of length $n$ from a sequence X, through a window of length $n$ placed randomly on the sequence, the $n$ 0's and 1's seen through the window are equally likely to be punctuated like **, **, ⋯ or like *, **, *⋯. We therefore define the window probability $p(s_n)$ by:

$$p(x_1x_2 \ldots x_n) = \frac{1}{2}\pi(x_1x_2, x_3x_4, \ldots) + \frac{1}{2}\pi(x_1, x_2x_3, x_4 \ldots). \quad (1)$$

*Table 2* The number $N(s_n)$ of possible *n*-ads

| *n* | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $N(s_n)$ | 1 | 2 | 4 | 6 | 10 | 14 | 22 | 30 | 46 | 62 | 94 |

The probabilities $p(s_n)$ must satisfy various conditions, among which are:

$$0 \leqslant p(s_n) \leqslant 1, \tag{2}$$

$$\sum_{s_n \in S_n} p(s_n) = 1 \tag{3}$$

and

$$p(s_n) = p(0s_n) + p(1s_n) = p(s_n 0) + p(s_n 1). \tag{4}$$

Without making any further assumptions we can now deduce the following properties of the window probabilities of monads ($n=1$), dyads ($n=2$), triads ($n=3$), ... etc.

*Monads.* Since X is composed of 01's and 10's we have:

$$p(0) = p(1) = \frac{1}{2}. \tag{5}$$

*Dyads.* From (4) and (5) it follows that:

$$\begin{aligned} p(0) &= p(00) + p(10) = p(00) + p(01) \\ &= p(1) = p(01) + p(11) = p(10) + p(11) \end{aligned}$$

so that:

$$p(00) = p(11) \tag{6}$$

and:

$$p(01) = p(10) = \frac{1}{2} - p(11). \tag{7}$$

*Triads.* Similarly from (4), (6) and (7) we obtain:

$$p(001) = p(100) = p(011) = p(110) = p(11), \tag{8}$$

$$p(010) = p(101) = \frac{1}{2} - 2p(11). \tag{9}$$

*Tetrads.* Equations (4), (8) and (9) imply that:

$$p(0010) = p(1011), \tag{10}$$

$$p(0100) = p(1101), \tag{11}$$

$$\begin{aligned} p(0110) &= p(1001) = p(0010) + p(0011) \\ &= p(0100) + p(1100) = p(11), \end{aligned} \tag{12}$$

$$\begin{aligned} p(0010) + p(1010) &= p(0100) + p(0101) \\ &= \frac{1}{2} - 2p(11). \end{aligned} \tag{13}$$

*Pentads.* We state only the most important identities. From (4), (10), (11) and (12) we find:

$$p(00101) = p(01011) = p(10010) = p(10110) = p(0010), \tag{14}$$

$$p(01001) = p(01101) = p(10100) = p(11010) = p(0100) \tag{15}$$

$$p(00110) = p(10011) = p(0011), \tag{16}$$

$$p(01100) = p(11001) = p(1100), \tag{17}$$

$$\begin{aligned} p(00110) + p(10110) &= p(01100) + p(01101) \\ &= p(01001) + p(11001) \\ &= p(10010) + p(10011) \\ &= p(0110), \end{aligned} \tag{18}$$

$$\begin{aligned} p(01010) + p(01011) &= p(10101) + p(00101) \\ &= p(0101), \end{aligned} \tag{19}$$

$$\begin{aligned} p(01010) + p(11010) &= p(10101) + p(10100) \\ &= p(1010). \end{aligned} \tag{20}$$

Similar, but increasingly more complicated identities hold for hexads, heptads, ... etc.

The following conclusions may be drawn from these identities. We denote the complement of $s_n$ by $\bar{s}_n$ (obtained by interchanging 0's and 1's), and the transpose by $s_n^*$ (obtained by reversing the order of the symbols)[8]. The complement and transpose operations commute, so there is no ambiguity in writing $\bar{s}_n^*$. Firstly, since (4) and (5) are symmetric in 0 and 1, any identity satisfied by $p(s_n)$ remains true if all sequences involved are complemented, e.g. (10) is transformed into (11). Secondly, we see that the 'principle of sequence reversibility':

$$p(s_n) = p(s_n^*) \tag{21}$$

holds for $n \leqslant 3$, but need not hold for $n \geqslant 4$, since in general $p(0010) \neq p(0100)$. Thirdly:

$$p(s_n) = p(\bar{s}_n^*) \tag{22}$$

holds for $n \leqslant 4$, and for all pentads except that $p(01010)$ need not equal $p(10101)$. For $n \geqslant 6$ (22) need not hold (the failure of (19) and (20) in general is mentioned implicitly by Frisch *et al.*[3]). Conditions which ensure that (21) and (22) hold for all $n$ will be given in Section V.

## SEQUENCE OBSERVATIONAL PROBABILITIES

Since n.m.r. and other spectrometric techniques cannot determine the direction of a sequence, we define the unconditional probability of observing a sequence $s_n$ by:

$$P_{obs}(s_n) = p(s_n) + p(s_n^*), \quad \text{if } s_n \neq s_n^*, \tag{23}$$

$$= p(s_n), \quad \text{if } s_n = s_n^*. \tag{24}$$

Our goal is to use n.m.r. spectroscopy to estimate certain of the $P_{obs}(s_n)$, and from this to deduce information about the laws governing the formation of the polymer.

Since n.m.r. spectral measurements depend on the signal from a given element which is influenced by a certain number ($k$ say) of adjacent elements on each side, we can only measure $P_{obs}(s_n)$ when $n = 2k+1$ is odd. When n.m.r. signals can be obtained from both 0's and 1's (as in the analysis of $PVF_2$ by $^{13}C$ n.m.r.[10,11]), estimates of $P_{obs}(s_n)$ can be obtained for all $s_n$ of a given odd length (which is limited by the resolution of the probe). On the other hand, if only 1's produce a signal (as in the present

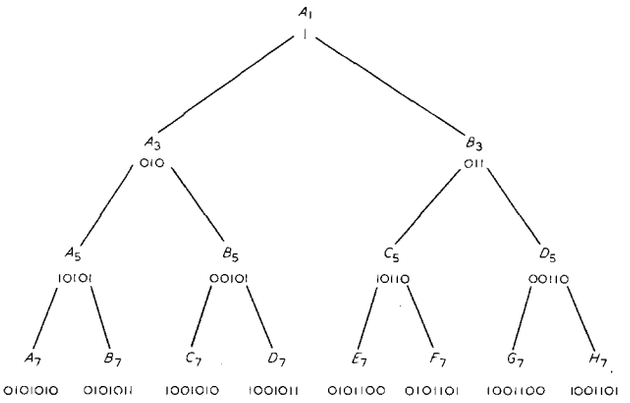



*Figure 1* Tree structure relating all observationally distinct 1-centred regiosequences from monad to heptads

analysis of $PVF_2$ by $^{19}F$ n.m.r.), estimates are obtained for precisely half of the $s_n$, namely those in which the middle element is a 1. For the complementary sequences $\bar{s}_n$, $P_{obs}(\bar{s}_n)$ cannot be estimated. However, if equation (22) holds, no information has been lost.

We now express the observational probabilities in terms of the window probabilities for all the observationally distinct 1-centred sequences up to heptads. Every observationally distinct 1-centred sequence of length $2k+1$ is contained in precisely two such sequences of length $2k+3$, and so the total number of such sequences of length $n=2k+1$ is $2^k$. These sequences for $k \leqslant 3$ are shown in the tree given in *Figure 1*.

*Monads* ($k=0$). In this trivial case there is only one n.m.r. signal $A_1$ from the monad 1. No information is obtained, since:

$$P_{obs}(1)=p(1)=\frac{1}{2}. \tag{25}$$

*Triads* ($k=1$). At the next higher level of resolution there are two n.m.r. signals (see *Figure 1*) corresponding to the events:

$$A_3=010 \ ,$$
$$B_3=011 \quad \text{or } 110 \ .$$

Then, from (8) and (9),

$$P_{obs}(A_3)=p(010)=\frac{1}{2}-2p(11), \tag{26}$$

$$P_{obs}(B_3)=p(011)+p(110)=2p(11). \tag{27}$$

*Pentads* ($k=2$). If all pentads are resolved, there are four n.m.r. signals (*Figure 1*). Then, from (14) and (15),

$$P_{obs}(A_5)=p(10101) \ , \tag{28}$$

$$P_{obs}(B_5)=p(00101)+p(10100)$$
$$=p(0010)+p(0100) \ , \tag{29}$$

$$P_{obs}(C_5)=p(10110)+p(01101)$$
$$=P_{obs}(B_5) \ , \tag{30}$$

$$P_{obs}(D_5)=p(00110)+p(01100) \ , \tag{31}$$

and, from (13), (18) and (19),

$$P_{obs}(A_5)+P_{obs}(B_5)=P_{obs}(A_3)=\frac{1}{2}-2p(11) \ , \tag{32}$$

$$P_{obs}(C_5)+P_{obs}(D_5)=P_{obs}(B_3)=2p(11) \ . \tag{33}$$

(32) and (33) also follow from the tree structure in *Figure 1*.

*Heptads* ($k=3$). There are eight observationally distinct heptads (see *Figure 1*), and we have:

$$P_{obs}(A_7)=p(0101010) \ ,$$
$$P_{obs}(B_7)=p(0101011)+p(1101010) \ ,$$
$$P_{obs}(C_7)=p(1001010)+p(0101001) \ ,$$
$$P_{obs}(D_7)=p(1001011)+p(1101001) \ ,$$
$$P_{obs}(E_7)=p(0011010)+p(0101100) \ ,$$
$$P_{obs}(F_7)=p(0101101)+p(1011010) \ ,$$
$$=2p(0101101) \ ,$$
$$P_{obs}(G_7)=p(1001100)+p(0011001) \ ,$$
$$=2p(1001100) \ ,$$
$$P_{obs}(H_7)=p(1001101)+p(1011001) \ . \tag{34}$$

The following identities can be easily proven:

$$P_{obs}(B_7)=P_{obs}(C_7) \ ,$$
and
$$P_{obs}(E_7)=P_{obs}(H_7) \ . \tag{35}$$

In any case we have:

$$P_{obs}(A_7)+P_{obs}(B_7)=P_{obs}(A_5) \ , \tag{36}$$

$$P_{obs}(C_7)+P_{obs}(D_7)=P_{obs}(B_5) \ , \tag{37}$$

$$P_{obs}(E_7)+P_{obs}(F_7)=P_{obs}(C_5) \ , \tag{38}$$

$$P_{obs}(G_7)+P_{obs}(H_7)=P_{obs}(D_5) \ . \tag{39}$$

For each value of $k$ the values of $P_{obs}(s_{2k+1})$ sum to $p(1)=\frac{1}{2}$, as required.

## SEQUENCE FORMATIONAL PROBABILITIES WITH A FIRST-ORDER MARKOV MODEL

The most widely applicable model for a binary copolymerization of monomer $A$ with monomer $B$ involves a first-order Markov process, which is characterized by four conditional probabilities $\pi(A|A)$, $\pi(A|B)$, $\pi(B|A)$ and $\pi(B|B)$, where $\pi(i|j)$ is the probability that the monomer $i$ is followed by the monomer $j$ (note that this is the reverse of the usual mathematical notation for conditional probabilities[12]).

In the present instance we assume that regiosequences are generated by a 'copolymerization' of the two monomers '01' and '10'. Thus there are four ways in which the chain can grow[7]:

| | probability | rate constant |
|---|---|---|
| $-01+01\rightarrow -01$ | $\pi(01\|01)$ | $k_{11}$ |
| $-01+10\rightarrow -10$ | $\pi(01\|10)$ | $k_{10}$ |
| $-10+01\rightarrow -01$ | $\pi(10\|01)$ | $k_{01}$ |
| $-10+10\rightarrow -10$ | $\pi(10\|10)$ | $k_{00}$ |

For simplicity, we set:

$$\pi(01|10)=x, \quad \pi(10|01)=y \ ,$$

so that $\pi(01|01)=1-x$ and $\pi(10|10)-1-y$. Our stationarity assumption implies that $x$ and $y$ satisfy:

$$\pi(01,)=\pi(01,)(1-x)+\pi(10,)y \ ,$$
$$\pi(10,)=\pi(01,)x+\pi(10,)(1-y) \ ,$$

and, since $\pi(01,)+\pi(10,)=1$, we find that:

$$\pi(01,)=\frac{y}{x+y} \ , \quad \pi(10,)=\frac{x}{x+y} \ . \tag{40}$$

The conditional probabilities, rate constants and concentrations of reacting species (enclosed within square brackets) are related by[4]:

$$x=\frac{k_{10}[-01][10]}{k_{11}[-01][01]+k_{10}[-01][10]} \ ,$$

$$y=\frac{k_{01}[-10][01]}{k_{01}[-10][01]+k_{00}[-10][10]} \ .$$

In terms of the reactivity ratios[13]:

$$r_0=\frac{k_{00}}{k_{01}} \ , \quad r_1=\frac{k_{11}}{k_{10}} \ ,$$

and using $[01]=[10]$, these formulas reduce to:

$$x=\frac{1}{1+r_1} \text{ and } y=\frac{1}{1+r_0} \ . \tag{41}$$

With this model it is straightforward to prove that the 'principle of sequence reversibility' (equation 21) holds for all $n$ if and only if $x=y$. This condition is analogous to the steady-state assumption for binary copolymerization[13], which states that the rates of interconversion between the two growing chain terminii (i.e. $-01$ and $-10$) are equal. We also note that equation (22) is always true for a first-order Markov model, as well as the simpler zeroth-order Markov or Bernoulli model, for which $x+y=1$.

The sequence observational probabilities may now be expressed in terms of $x$ and $y$, beginning with the heptads. Since the resulting expressions are symmetric functions of $x$ and $y$, it is convenient to set:

$$\alpha=x+y \ ,$$
$$\beta=xy \ .$$

Then $x,y$ can be determined from $\alpha,\beta$ by:

$$x=\frac{1}{2}(\alpha\pm\sqrt{\alpha^2-4\beta}) \ , \tag{42}$$

$$y=\frac{1}{2}(\alpha\mp\sqrt{\alpha^2-4\beta}) \ . \tag{43}$$

We find that:

$$\begin{aligned}\mathrm{P}_{\mathrm{obs}}(A_7)&=p(0101010)\\&=\frac{1}{2}\pi(01,01,01,01)+\frac{1}{2}\pi(10,10,10,10)\\&=\frac{1}{2}\frac{y}{x+y}(1-x)^3+\frac{1}{2}\frac{x}{x+y}(1-y)^3\\&=\frac{1}{2}\left(1+3\beta-\alpha\beta+\frac{2\beta^2}{\alpha}-\frac{6\beta}{\alpha}\right) \ ,\end{aligned} \tag{44}$$

and similarly:

$$\mathrm{P}_{\mathrm{obs}}(B_7)=\mathrm{P}_{\mathrm{obs}}(C_7)=\frac{\beta}{\alpha}\left(1-\alpha-\beta+\frac{\alpha^2}{2}\right) \ , \tag{45}$$

$$\mathrm{P}_{\mathrm{obs}}(D_7)=\frac{\beta}{2}\left(1-\alpha+\frac{2\beta}{\alpha}\right) \ , \tag{46}$$

$$\mathrm{P}_{\mathrm{obs}}(E_7)=\mathrm{P}_{\mathrm{obs}}(H_7)=\frac{\beta}{2}\left(1-\frac{2\beta}{\alpha}\right) \ , \tag{47}$$

$$\mathrm{P}_{\mathrm{obs}}(F_7)=\frac{\beta}{\alpha}(1-\alpha+\beta) \ , \tag{48}$$

$$\mathrm{P}_{\mathrm{obs}}(G_7)=\frac{\beta^2}{\alpha} \ . \tag{49}$$

The expressions for pentads and triads are then obtained from (36)–(39) and (32), (33):

$$\mathrm{P}_{\mathrm{obs}}(A_5)=\frac{1}{2}\left(1+\beta-\frac{4\beta}{\alpha}\right) \ , \tag{50}$$

$$\mathrm{P}_{\mathrm{obs}}(B_5)=\mathrm{P}_{\mathrm{obs}}(C_5)=\beta\left(\frac{1}{\alpha}-\frac{1}{2}\right) \ , \tag{51}$$

$$\mathrm{P}_{\mathrm{obs}}(D_5)=\frac{\beta}{2} \ , \tag{52}$$

$$\mathrm{P}_{\mathrm{obs}}(A_3)=\frac{1}{2}-\frac{\beta}{\alpha} \ , \tag{53}$$

$$\mathrm{P}_{\mathrm{obs}}(B_3)=\frac{\beta}{\alpha} \ . \tag{54}$$

From (26) we note:

$$p(11)=\frac{1}{2}\frac{\beta}{\alpha} \ . \tag{55}$$

The pentad relations (50)–(52) have been derived by Wilson using a different notation[14].

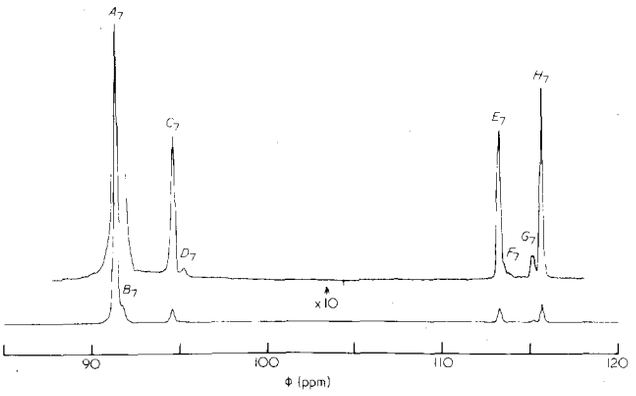



*Figure 2* 188.22 MHz $^{19}F$ n.m.r. spectrum of polyvinylidene fluoride (Kynar 961) at 50°C in dimethylformamide-$d_7$. Other experimental conditions and identification of the peaks are given in the text

For the Bernoullian model the conditional probabilities $\pi(i|j)$ are independent of $i$, so that $x+y=1$ and $\alpha=1$, $\beta=x(1-x)$. This model has been applied to heptad data obtained by Ferguson and Brame from the $^{19}F$ n.m.r. spectrum of $PVF_2$,[8] but as we show in the next section a first-order Markov model is more appropriate. In any case resolution of the two triad peaks $A_3$ and $B_3$ is sufficient to solve for the Bernoulli model, since estimates for $\beta$ are then directly obtained (from (53), (54)).

For the first-order Markov model resolution of triads gives only the ratio $\beta/\alpha$, and so pentads must be resolved in order to estimate $x$ and $y$, and hence $r_0$ and $r_1$. If heptads are resolved, then one can test for any deviation from the first-order model, and if necessary fit a second-order model (which requires the specification of four reactivity ratios[4]).

It should be pointed out that there is always an ambiguity in our equations, which only give information about the symmetric functions $\alpha$ and $\beta$, from which $x$ and $y$ must be derived according to (42) and (43) with an ambiguity of sign. This ambiguity exists because our observations do not discern the direction of chain growth. Once this direction is specified, then only one solution for $x$ and $y$ is admissable.

## APPLICATION TO POLYVINYLIDENE FLUORIDE

We illustrate the foregoing theory by an analysis of $^{19}F$ n.m.r. spectra of $PVF_2$. Two representative samples were chosen: Kynar 961 manufactured by the Pennwalt Corporation, and Kureha KF-1100 manufactured by Kureha Chemical Industry Company. They were observed as 8% solutions by weight in dimethylformamide-$d_7$ or dioxane-$d_8$ on a Varian XL-200 spectrometer operating at 188.22 MHz for $^{19}F$.

Initially several $PVF_2$ samples (both commercial and from our laboratory) were examined in the solvents dimethylformamide, dioxane, acetophenone and hexamethylphosphoramide at various temperatures from 17° to 160°C) to determine suitable conditions for the experiment. Some $PVF_2$ samples, notably those with low molecular weight, gave anomalous peaks (see later), and therefore were unsuitable for this work. Both Kynar 961 and Kureha KF-1100 are satisfactory in this regard. Best resolution of regiosequence heptads was achieved with the solvent dimethylformamide, but the relative spacing of these heptads was temperature dependent. Low temperature particularly favoured the separation of peaks $A_7$ and $B_7$, but as we show later, a reliable estimate of $P_{obs}(A_7)$ and $P_{obs}(B_7)$ could not be obtained using the solvent dimethylformamide. Dioxane was the preferred solvent in this regard. For these preliminary trials protons were decoupled from fluorine nuclei for optimum resolution to aid in peak identification.

The final experimental conditions for acquiring $^{19}F$ n.m.r. spectra were carefully chosen to ensure quantitative measurements. An interval of 15 s (well in excess of $5T_1$) separated 90° (9.0 $\mu$s) pulses to allow all $^{19}F$ magnetization to return to equilibrium, thereby avoiding potential distortion of relative peak intensities from differential spin–lattice relaxation times ($T_1$s), which range from 0.39 to 0.46 s at 21.1 kG[10]. Likewise the problem of differential nuclear Overhauser enhancement factors did not arise because protons were not irradiated at 200 MHz as required during decoupling from the observed fluorine nucleus. The spectral sweep width was 12 kHz (1.33 s acquisition time) with 32 K data points, and 1000–3000 transients were accumulated to ensure an adequate signal-to-noise ratio for even the weakest peaks ($D_7$,$F_7$). The detection limit is 1 part in 10000. A floating point Fourier transform of the time domain data was performed, with a minimal digital line broadening of 0.05 Hz.

Peak areas, proportional to $P_{obs}(s_7)$, were determined by expanding individual resonances on a scale of 24 Hz/cm and tracing the peaks, which were then cut out and weighed. Four spectra were recorded independently with different batches of Kynar 961 under identical conditions, and were exactly superimposable. Repeat weighings of a given peak were reproducible to ±0.01%. The only significant source of error lay in the cutting-out procedure owing to a slight degree of arbitrariness in separating overlapping peaks (notably $E_7$ from $F_7$, and $G_7$ from $H_7$).

The entire $^{19}F$ n.m.r. spectrum of $PVF_2$ dissolved in dimethylformamide is shown in *Figure 2*, with detailed expansions of the $A_3$ and $B_3$ regions in *Figures 3* and *4*, respectively. The heptads $A_7$-$H_7$ are indicated on these spectra. The assignments follow those given by Ferguson and Brame[8], and have been the subject of recent

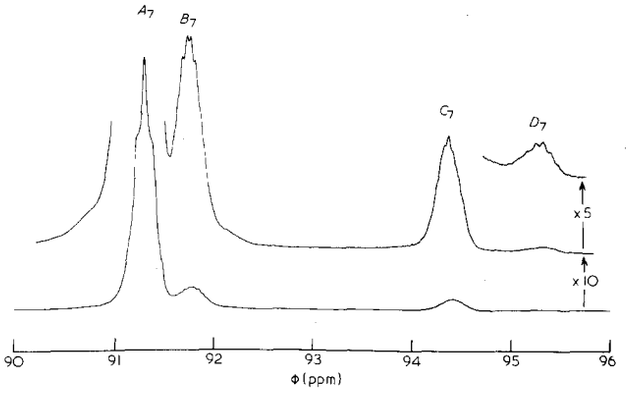



*Figure 3* Detailed expansions of $^{19}F$ resonances derived from the $A_3$ sequence (010) in Kynar 961 observed at 18°C. Other conditions are the same as for *Figure 2*. Note that at this lower temperature the peak $B_7$ is slightly better separated from $A_7$ than in *Figure 2*. Peak areas were measured at the following vertical amplifications: $A_7$, X1; $B_7$, X10; $C_7$, X10; and $D_7$, X50

theoretical calculations of Tonelli *et al.*[15] In some $PVF_2$ spectra obtained in dimethylformamide we have detected anomalous $^{19}F$ resonances at 95.32, 97.47, 98.88, 107.17, 109.92 and 114.15 ppm. Certain of these resonances can be seen in the spectra published by Ferguson and Brame[8], and Liepins *et al.*[16] We have determined that several of these peaks relax more slowly than the main peak $A_7$, and therefore are probably associated with mobile end groups and oligomeric residues. In one report these peaks, as well as multiplicity from long-range homonuclear spin coupling and probable spinning side-bands, have been incorrectly attributed to fine structure from sequences longer than heptads (i.e. four or more monomer units)[11]. Care must be taken to differentiate the anomalous peak at 114.15 ppm (marked X in *Figure 4*) from $F_7$, as these peaks have similar intensities.

Chemical shifts for the eight individual heptads $A_7$-$H_7$ are given in *Table 3*, with estimates for the corresponding values of $P_{obs}$ ($s_7$) for both polymers as measured in dimethylformamide. An internal check on these results is given by equation (35), which states that $P_{obs}(B_7) = P_{obs}(C_7)$ and $P_{obs}(E_7) = P_{obs}(H_7)$ no matter what statistical law governs the chain growth process. Similarly, according to equations (30), (37) and (38), $P_{obs}(C_7) + P_{obs}(D_7) = P_{obs}(E_7) + P_{obs}(F_7)$. The last two identities are satisfied by the present results to within the experimental error, but we find unexpectedly that $P_{obs}(B_7)$ is nearly 50% larger than $P_{obs}(C_7)$.

The exaggerated value of $P_{obs}(B_7)$ is an artifact caused by the solvent dimethylformamide. *Figure 5* shows a detailed expansion of the $A_3$ region observed at 90°C in dioxane-$d_8$. Under these different conditions the heptads shift relative to each other and have chemical shift values as follows: $A_7$, 89.88; $B_7$, 89.05; $C_7$, 94.62; $D_7$, 93.91; $E_7$, 112.24; $F_7$ and $G_7$, not resolved; and $H_7$, 114.48 ppm. Two new peaks appear on the high-field shoulder of $A_7$ at 90.63 and 91.23 ppm (*Figure 5*), with a combined probability equal to $P_{obs}(B_7) - P_{obs}(C_7)$ as measured for solutions in dimethylformamide (*Table 3*). Clearly these new peaks overlap with $B_7$ when the solvent is dimethylformamide. Their probability is too high to be associated with end groups (both Kynar 961 and Kureha KF-1100 have molecular weights in excess of 60 000), but too low to be nonad fine structure (i.e. $B_9 = 101010100$). However they must be derived from the $A_7$ sequence so we have included their area in the correct estimate for $P_{obs}(A_7)$.

*Tables 4* and *5* give the correct measured values for all $P_{obs}(s_7)$ for Kynar 961 and Kureha KF-1100, respectively. The values for $P_{obs}(A_7)$ and $P_{obs}(B_7)$ were measured from spectra obtained in dioxane solution, and the remaining probabilities were measured from spectra obtained in dimethylformamide solution, where the heptads $F_7$ and

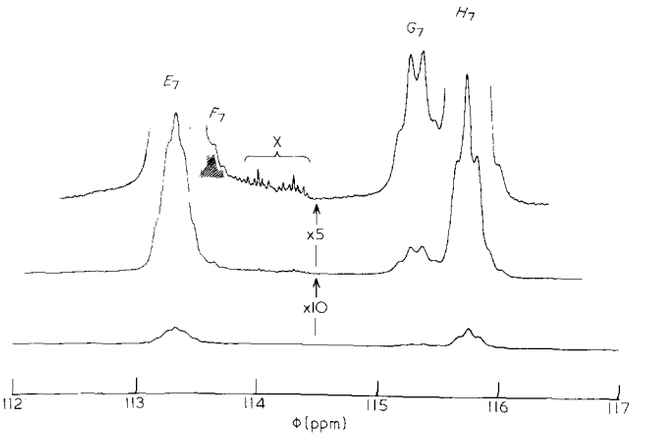



*Figure 4* Detailed expansions of $^{19}F$ resonances derived from the $B_3$ sequence (011) in Kynar 961 observed at 18°C. Other conditions are the same as for *Figure 3*. X = probable end-group resonance (note well-resolved spin-coupling multiplicity on this peak). Peak areas were measured at the following vertical amplifications: $E_7$, X10; $F_7$, X50; $G_7$, X50; and $H_7$, X10

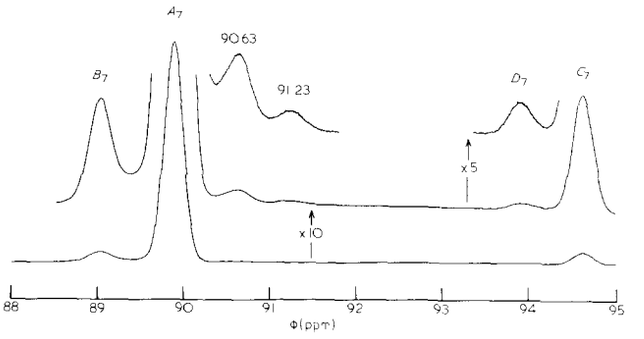



*Figure 5* Detailed expansions of $^{19}F$ resonances derived from the $A_3$ sequence (010) in Kynar 961 observed at 90°C in dioxane-$d_8$. Other conditions are the same as for *Figure 3*. The two peaks at 90.63 and 91.23 ppm seen here are hidden under $B_7$ in *Figure 3*, where they exaggerate the measured value of $P_{obs}$ ($B_7$)

*Table 3* Chemical shifts for the heptads observed in the $^{19}F$ n.m.r. spectra of $PVF_2$ in dimethylformamide-$d_7$ and the corresponding $P_{obs}$ values. The values of $P_{obs}$ ($A_7$) and $P_{obs}$ ($B_7$) are underestimated and overestimated, respectively (see text)

| Heptad | Sequence* | Φ (ppm)† ±0.05 | $P_{obs}$‡ | |
|---|---|---|---|---|
| | | | Kynar 961 | Kureha KF-1100 |
| $A_7$ | 0101010 | 91.31 | 0.393 ± 0.004 | 0.421 ± 0.004 |
| $B_7$ | 0101011 | 91.79 | 0.031 ± 0.001 | 0.022 ± 0.001 |
| $C_7$ | 1001010 | 94.43 | 0.023 ± 0.001 | 0.018 ± 0.001 |
| $D_7$ | 1001011 | 95.37 | 0.00130 ± 0.00005 | 0.00080 ± 0.00005 |
| $E_7$ | 0101100 | 113.33 | 0.024 ± 0.001 | 0.018 ± 0.001 |
| $F_7$ | 0101101 | 113.62 | 0.0004 ± 0.00008 | 0.0001 ± 0.00005 |
| $G_7$ | 1001100 | 115.34 | 0.0040 ± 0.0005 | 0.0020 ± 0.0005 |
| $H_7$ | 1001101 | 115.76 | 0.023 ± 0.001 | 0.018 ± 0.001 |

* 0 = $CH_2$, 1 = $CF_2$

† At 18°C in dimethylformamide-$d_7$ with internal hexafluorobenzene reference assigned 163.00 ppm

‡ Mean values from four independent trials. The errors arise solely from the imprecision of cutting out peaks. Note that $P_{obs}$ values are normalized to ½ as required by (25)

*Table 4* Comparison of measured $P_{obs}(s_n)$ values for heptads and pentads with those calculated for a first-order Markov model with $\alpha = 1.049$ and $\beta = 0.0522$ for Kynar 961

| Measured $P_{obs}(s_n)$ | | Calculated $P_{obs}(s_n)$ | |
|---|---|---|---|
| $n = 7$ | $n = 5$ | $n = 7$ | $n = 5$ |
| $A_7 = 0.400$<br>$B_7 = 0.024$ | $A_5 = 0.424$ | $A_7 = 0.404$<br>$B_7 = 0.022$ | $A_5 = 0.427$ |
| $C_7 = 0.023$<br>$D_7 = 0.0013$ | $B_5 = 0.024$ | $C_7 = 0.022$<br>$D_7 = 0.0013$ | $B_5 = 0.023$ |
| $E_7 = 0.024$<br>$F_7 = 0.0004$ | $C_5 = 0.024$ | $E_7 = 0.024$<br>$F_7 = 0.0002$ | $C_5 = 0.024$ |
| $G_7 = 0.0040$<br>$H_7 = 0.023$ | $D_5 = 0.027$ | $G_7 = 0.0026$<br>$H_7 = 0.024$ | $D_5 = 0.026$ |

$G_7$ are better resolved. These values are used to derive the statistical law governing the formation of $PVF_2$.

The simplest statistical law for chain growth is zeroth-order Markov or Bernoulli, for which $\alpha$ is unity. If this law is applicable we find from (46), (48) and (49) that:

$$P_{obs}(D_7) = P_{obs}(F_7) = P_{obs}(G_7) = \beta^2 \ ,$$

and, from (45) and (47),

$$P_{obs}(B_7) = P_{obs}(E_7) = \beta\left(\frac{1}{2} - \beta\right) \ .$$

Similarly at the pentad level we find:

$$P_{obs}(B_5) = P_{obs}(C_5) = P_{obs}(D_5) = \frac{\beta}{2} \ .$$

These identities are not satisfied by either our observations (*Tables 4* and *5*) or those reported by Ferguson and Brame[8]. Clearly the regiosequence distribution in $PVF_2$ cannot be described accurately by Bernoullian statistics.

*Tables 4* and *5* show that the $P_{obs}(s_5)$ values obtained according to (36)–(39) are described by a first-order Markov model with $\alpha = 1.049$ and $\beta = 0.0522$ for Kynar 961, and $\alpha = 1.037$ and $\beta = 0.0398$ for Kureha KF-1100. As previously noted, pentad data can always be described by a first-order Markov model, while the appropriateness of this model can only be tested by heptad or longer sequence probabilities. Likewise the heptad data could be modelled by a second-order Markov model or any statistical model with four or more adjustable parameters, but it is unnecessary to pursue these higher-order models since the present data show that the first-order Markov model is appropriate.

*Tables 4* and *5* show that the $P_{obs}$ values for two heptads, namely $F_7$, and $G_7$, deviate slightly from a first-order Markov model within the precision of our measurements. The apparent deviation of $F_7$ and $G_7$ may not be significant, because the absolute error in their probabilities is certainly much larger than for other heptads, since these two weak peaks are partially overlapped by more intense neighbours.

Given that a first-order Markov model is physically reasonable, we can solve for $x$ and $y$, and hence $r_0$ and $r_1$ from (42), (43) and (41). The solutions are either:

$$x = 0.997, \quad y = 0.052 \ , \tag{56}$$

or

$$x = 0.052, \quad y = 0.997 \ , \tag{57}$$

for Kynar 961, and

$$x = 0.997, \quad y = 0.040 \ , \tag{58}$$

or

$$x = 0.040, \quad y = 0.997 \ , \tag{59}$$

for Kureha KF-1100.

It is generally accepted for $VF_2$ polymerization that the propagating radical structure is more likely $-CH_2\cdot$ than $-CH_2{}^{\cdot}$,[14] so the appropriate solutions are (57) and (59). Thus:

$$r_0 = 0.003, \quad r_1 = 18 \ , \tag{60}$$

for Kynar 961, and

$$r_0 = 0.003, \quad r_1 = 24 \ , \tag{61}$$

for Kureha KF-1100.

Throughout the literature on $PVF_2$ the polymer structure has been characterized by the 'per cent head–head, tail–tail' addition. This is simply 100 $(p(11) + p(00))$, which is equal to 100 $P_{obs}(B_3)$ according to (27) and (6). Thus the per cent 'head–head', tail–tail' additions for Kynar 961 and Kureha KF-1100 are 5.0 and 3.7, respectively. This is the same as the 'per cent inverted monomer units', or 100 $\pi(10)$, as given by equation (40). This number cannot characterize the sequence microstructure of $PVF_2$ however except in the case of 'single parameter' Bernoullian statistics, which we have shown to be inapplicable in the present instance. The microstructure of $PVF_2$ is described instead by a 'two-parameter' first-order Markov model. Both Kynar 961 and Kureha KF-1100 are products of a free-radical addition polymerization, but, because of its higher defect content, the Kynar 961 must have been prepared at significantly higher temperature than that employed for the Kureha KF-1100.

$PVF_2$ provides the best example for illustrating regioirregular sequences in homopolymers, since it does not involve the stereoconfigurational irregularity that complicates the analysis of other fluoropolymers such as polyvinyl fluoride and polytrifluoroethylene. However, since $PVF_2$ has few defects, its n.m.r. analysis is not a routine experiment and may be likened to the task of determining reactivity ratios for a copolymer having only 5% of one type of monomer unit. This is partially offset by

*Table 5* Comparison of measured $P_{obs}(s_n)$ values for heptads and pentads with those calculated for a first-order Markov model with $\alpha = 1.037$ and $\beta = 0.0398$ for Kureha KF-1100

| Measured $P_{obs}(s_n)$ | | Calculated $P_{obs}(s_n)$ | |
|---|---|---|---|
| $n = 7$ | $n = 5$ | $n = 7$ | $n = 5$ |
| $A_7 = 0.424$<br>$B_7 = 0.019$ | $A_5 = 0.443$ | $A_7 = 0.426$<br>$B_7 = 0.018$ | $A_5 = 0.443$ |
| $C_7 = 0.018$<br>$D_7 = 0.0008$ | $B_5 = 0.018$ | $C_7 = 0.018$<br>$D_7 = 0.0008$ | $B_5 = 0.018$ |
| $E_7 = 0.018$<br>$F_7 = 0.0001$ | $C_5 = 0.018$ | $E_7 = 0.018$<br>$F_7 = 0.0001$ | $C_5 = 0.018$ |
| $G_7 = 0.0020$<br>$H_7 = 0.018$ | $D_5 = 0.020$ | $G_7 = 0.0015$<br>$H_7 = 0.018$ | $D_5 = 0.020$ |

the sensitivity of the $^{19}F$ nucleus to its structural environment, so that regiosequence heptads can be discerned, which allows one- and two-parameter statistical models to be tested.

## ACKNOWLEDGEMENT

The authors are grateful to Dr. A. J. Lovinger for donation of $PVF_2$ samples and Mrs. J. M. Kometani for assistance with the n.m.r. analysis.

## REFERENCES

1 Randall, J. C. 'Polymer Sequence Determination. Carbon-13 NMR Method', Academic Press, New York, 1977
2 Coleman, B. D. and Fox, T. G. *J. Polym. Sci. A* 1963, **1**, 3183
3 Frisch, H. L., Mallows, C. L. and Bovey, F. A. *J. Chem. Phys.* 1966, **45**, 1565
4 Price, F. P. *J. Chem. Phys.* 1962, **36**, 209
5 Ito, K. and Yamashita, Y. *J. Polym. Sci. A* 1965, **3**, 2165
6 Pyun, C. W. *J. Polym. Sci. A-2* 1970, **8**, 1111
7 Koenig, J. L. 'Chemical Microstructure of Polymer Chains', Wiley–Interscience, New York, 1980, Ch. 9
8 Ferguson, R. C. and Brame Jr., E. G. *J. Phys. Chem.* 1979, **83**, 1397
9 The terms 'antorientic', heteroorientic' and 'postorientic' have also been proposed for 'head–head', 'head–tail' and 'tail–tail' structures—see for example ref 7
10 Bovey, F. A., Schilling, F. C., Kwei, T. K. and Frisch, H. L. *Macromolecules* 1977, **10**, 559
11 Lin, F-T. *PhD Thesis* Univ. of Akron, 1979. Also see Lin, F-T. and Wilson, C. W. *Bull. Am. Phys. Soc.* 1981, **26**, 806, abstract CE5
12 Feller, W. 'An Introduction to Probability Theory and Its Applications', John Wiley, New York, 3rd Edn., 1968, Ch 5
13 Goldfinger, G. and Kane, T. *J. Polym. Sci.* 1948, **3**, 462
14 Wilson, C. W. *J. Polym. Sci. A* 1963, **1**, 1305
15 Tonelli, A. E., Schilling, F. C. and Cais, R. E. *Macromolecules* 1982, **15**, 849
16 Liepins, R., Surles, J. R., Morosoff, N., Stannett, V. T., Timmons, M. L. and Wortman, J. J. *J. Polym. Sci. Polym. Chem. Edn.* 1978, **16**, 3039