Carbohydrate Polymers 3 (1983) 299–312

Water-soluble (1→3), (1→4)-β-D-Glucans from Barley (*Hordeum vulgare*) Endosperm. III. Distribution of Cellotriosyl and Cellotetraosyl Residues

R. G. Staudte*, J. R. Woodward†, G. B. Fincher† and B. A. Stone†

Departments of Mathematical Statistics* and Biochemistry†, La Trobe University, Bundoora, Victoria 3083, Australia

(Received: 25 November 1982)

ABSTRACT

Cellotriosyl and cellotetraosyl residues, linked by single (1→3)-β-linkages, account for more than 90% of the 40°C water-soluble (1→3),(1→4)-β-D-glucan from barley flour. We have analysed their sequence dependence by treating the polymer as a two-state Markov chain with stationary distribution. Quantitation of the penultimate oligosaccharides released during hydrolysis of the (1→3),(1→4)-β-D-glucan with (1→3),(1→4)-β-D-glucan 4-glucanohydrolase (EC 3.2.1.73) by analytical gel filtration chromatography enabled the relative abundance of two adjacent cellotriosyl, two adjacent cellotetraosyl and adjacent cellotetraosyl/cellotriosyl residues to be estimated and the sequence dependence to be evaluated.

Within the theoretical and practical constraints of the method it is concluded that the cellotriosyl and cellotetraosyl residues are arranged in an essentially independent (random) fashion. Thus, any mechanism proposed for the biosynthesis of the molecule should explain this apparently random distribution of cellotriosyl and cellotetraosyl residues as well as the presence, in relatively low frequency, of blocks of up to 10 or more adjacent (1→4)-linkages.

1. INTRODUCTION

The (1→3),(1→4)-β-D-glucan (hereafter referred to as β-glucan) extracted from barley flour at 40°C with water is an important constituent

Carbohydrate Polymers 0144-8617/83/$03.00

of endosperm cell walls. The polysaccharide is composed predominantly of cellotriosyl and cellotetraosyl residues, which account for more than 90% (mol/mol) of the molecule, separated by single (1→3)-β-linkages (Woodward *et al.*, 1983*b*). Thus the (1→3)- and (1→4)-linkages are distributed neither in a statistically random arrangement nor in a strictly repeating sequence. This type of arrangement is also observed for oat β-glucan and lichenin (Clarke & Stone, 1963). Longer blocks of up to 10 contiguous (1→4)-linkages in small but significant proportions are also present but contiguous (1→3)-linkages were not detected in the sample examined here (Woodward *et al.*, 1983*b*).

In the present work the distribution of cellotriosyl and cellotetraosyl residues in the 40°C water-soluble barley β-glucan is investigated. For this purpose the polysaccharide is assumed to form a two-state Markov chain in which the two states are cellotriosyl and cellotetraosyl units. The Markov property dictates that given any sequence of units (states), the probability that the next unit is a cellotriosyl or a cellotetraosyl residue depends only on the type of the previous, and hence adjacent, unit in the sequence. (For an introduction to Markov chains see Hoel *et al.*, 1972.) Furthermore, it is assumed that the Markov chain has a stationary distribution; that is, the occurrence of any given sequence is independent of its position in the long polysaccharide chain. Under these assumptions, a reasonable measure of the dependence in the sequence is the autocorrelation of the Markov chain, which may be estimated from the relative abundance of specific oligosaccharides released during enzymic hydrolysis of the β-glucan.

2. MATERIALS AND METHODS

2.1 Materials

The 40°C water-soluble β-glucan was prepared from barley (*Hordeum vulgare* cv. Clipper) flour as described by Woodward *et al.* (1983*a*). The (1→3),(1→4)-β-D-glucan 4-glucanohydrolases (EC 3.2.1.73) were enzymes I and II purified by Woodward & Fincher (1982*a*).

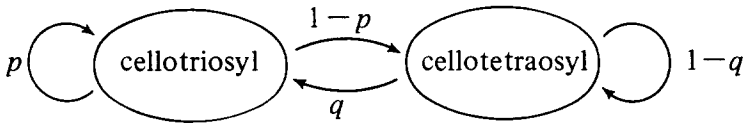
2.2 Mathematical methods

2.2.1 Theory

We have assumed that the cellotriosyl and cellotetraosyl residues, which account for 65% (mol/mol) and 27% (mol/mol) of the β-glucan respec-

tively and which are linked by single (1→3)-β-linkages (Woodward *et al.*, 1983*b*), form a two-state Markov chain with stationary distribution in the linear β-glucan molecule. The longer blocks of adjacent (1→4)-linkages, comprising approx. 8% (mol/mol) of the molecule, are considered only in relation to error analysis.

The dependence of the sequence in the Markov chain can be calculated from the transition probabilities p and q, where p is the probability that a cellotriosyl unit is followed by another cellotriosyl unit and q is the probability that a cellotetraosyl unit is followed by a cellotriosyl unit.



If the transition probabilities are as shown above, with $0 < p < 1$ and $0 < q < 1$, then the stationary distribution $[\pi, 1-\pi]$ is given by

$$\pi = \frac{q}{(1-p+q)} \tag{1}$$

where the probability π can be interpreted as the relative abundance of cellotriosyl units (approx. 0·71) and $1-\pi$ the relative abundance of cellotetraosyl units (approx. 0·29) in a long section of the molecule.

A measure of dependence is given by the correlation between adjacent cellodextrin units, the *autocorrelation* ρ. It can be shown that

$$\rho = p - q \tag{2}$$

If ρ is near 1, the sequence will contain large groups of adjacent cellotriosyl units and large groups of adjacent cellotetraosyl units; a value of ρ near -1 indicates that the cellotriosyl and cellotetraosyl units are arranged in an alternating sequence and a value of $\rho = 0$ is obtained if, and only if, the cellotriosyl and cellotetraosyl units are arranged independently, or randomly.

We want to estimate ρ, given a section of the sequence. If the section is cut into individual units it is possible to measure π and $1-\pi$; that is the frequency of cellotriosyl and cellotetraosyl units respectively in the polysaccharide. This has been achieved by complete enzymic hydrolysis with (1→3),(1→4)-β-D-glucan endohydrolase followed by quantitation of oligomeric products using Bio-Gel P-2 gel filtration chromatography (Woodward *et al.*, 1983*b*). However, the measurement of π and $1-\pi$

does not allow the estimation of p and q, and hence ρ, since the information on the frequency of adjacent units is lost on complete enzymic hydrolysis.

However, p, q and ρ can be estimated if the polysaccharide is cut into segments each containing two units, where a unit is a cellotriosyl or a cellotetraosyl residue. In this situation we have four possible two-unit segments forming a four-state Markov chain. This chain has a stationary distribution $[\pi_0, \pi_1, \pi_2, \pi_3]$ where π_0 is the relative abundance of two adjacent cellotriosyl residues in a long section of the molecule, π_1 is the relative abundance of a cellotriosyl residue followed by a cellotetraosyl residue, π_2 is the relative abundance of a cellotetraosyl residue followed by a cellotriosyl residue, and π_3 is the relative abundance of two adjacent cellotetraosyl residues. It can be shown (Appendix I) that

$$\pi_0 = \frac{pq}{(1-p+q)} \tag{3}$$

$$\pi_1 = \frac{(1-p)\,q}{(1-p+q)} = \pi_2 \tag{4}$$

Note that $\pi_0 + \pi_1 + \pi_2 + \pi_3 = 1$. Solving for p and q in eqns (3) and (4) gives

$$p = \frac{\pi_0}{(\pi_0 + \pi_1)} \tag{5}$$

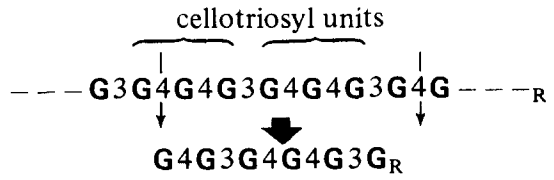
$$q = \frac{\pi_1}{(1-\pi_0-\pi_1)} \tag{6}$$

and the autocorrelation ρ can thus be estimated.

2.2.2 Practice

The values for π_0, π_1, π_2 and π_3 can be estimated from the 'penultimate' oligosaccharides released from the barley β-glucan during its depolymerisation with two (1→3),(1→4)-β-D-glucan endohydrolases. Both enzymes hydrolyse (1→4)-β-linkages in (1→3),(1→4)-β-D-glucans where the glucosyl residue is substituted at C(O)3 (Woodward & Fincher, 1982*b*). The frequency of two adjacent cellotriosyl residues (π_0) can be obtained from the relative abundance of the oligosaccharide **G4G3G4G4G3G**$_R$, where **G** denotes a glucosyl residue and subscript

R the reducing end. This oligosaccharide has a degree of polymerisation (*DP*) 6 and is released from the polysaccharide as shown.



It is referred to as a penultimate oligosaccharide because it is finally hydrolysed to form two molecules of 3-O-β-D-cellobiosyl-D-glucose (**G4G3G**$_R$), which are not degraded further.

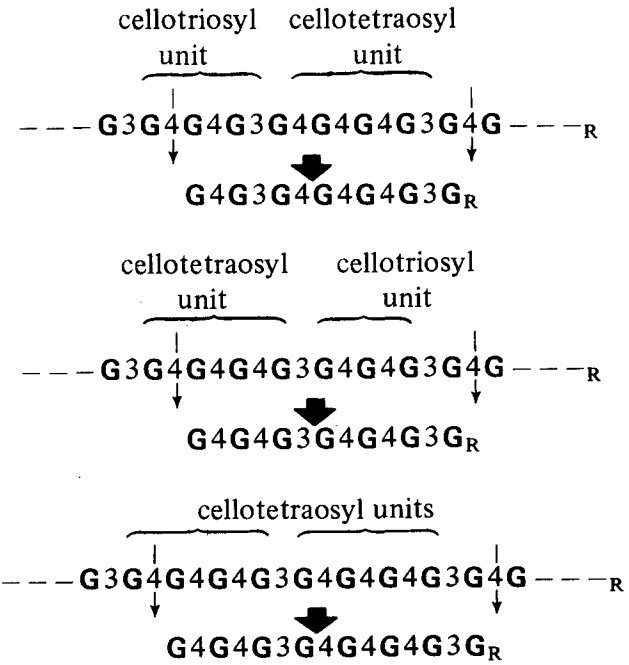
The relative abundance of oligosaccharides of

DP 7 (**G4G3G4G4G4G3G**$_R$ and **G4G4G3G4G4G3G**$_R$)

and

DP 8 (**G4G4G3G4G4G4G3G**$_R$),

released from the polysaccharide as shown below, provides values of π_1, π_2 and π_3 respectively.



Thus quantitation of oligosaccharides of *DP* 6, 7 and 8 in the enzymic hydrolysate of the barley β-glucan enables the estimation of π_0, $\pi_1 + \pi_2$, and π_3 respectively. The probabilities π_1 and π_2 are assigned as one-half the proportion of oligosaccharides of *DP* 7 in the hydrolysate (eqn. 4).

2.3 Enzymic hydrolysis

Barley β-glucan (0·2% w/v) in 2·5 mM sodium acetate–acetic acid buffer, pH 4·8 (containing 400 μg/ml bovine serum albumin and 5 mM sodium azide) was incubated at 38°C with (1→3),(1→4)-β-D-glucan endohydrolases at a final concentration of 0·8 μg/ml (0·22 nkat enzyme I, 0·27 nkat enzyme II) (Woodward & Fincher, 1982*a*). Aliquots (1 ml) were removed hourly, heated at 100°C for 10 min and 80 μl, containing approx. 160 μg carbohydrate, was fractionated on Bio-Gel P-2 as described by Woodward & Fincher (1982*b*). Oligosaccharides in the eluate were detected quantitatively with an Optilab Multiref 902 differential refractometer (Optilab, Vallingby, Sweden).

For error analysis, samples were applied to the gel filtration column in triplicate and peak area estimates were repeated three times for each profile. Median values were used in the estimation of peak areas. The hydrolysates of both enzyme I and enzyme II were examined.

3. RESULTS

The time course of the release of oligosaccharides during enzymic hydrolysis of barley β-glucan is shown in Fig. 1. Values for the abundance of penultimate oligosaccharides (*DP* 6, 7 and 8) are those obtained at 3 h (enzyme I) and 4 h (enzyme II) when the oligosaccharides account for 31% and 34% respectively of the total polysaccharide, and are expressed below on a molar basis

	DP 6 (π_0)	*DP* 7 ($\pi_1 + \pi_2$)	*DP* 8 (π_3)
Enzyme I	0·512	0·408	0·080
Enzyme II	0·558	0·365	0·077

The autocorrelation ρ calculated from these values using eqns (5), (6) and (2) is −0·003 (standard error 0·016) for the hydrolysate of

enzyme I and 0·050 (standard error 0·046) for the enzyme II hydrolysate. (For detailed error analysis see Appendix II.) The differences observed between enzymes I and II are attributed to subtle differences in the kinetics of hydrolysis (Woodward & Fincher, 1982*b*), which are presumably related to binding site specificities.

4. DISCUSSION

The autocorrelation values (ρ) of −0·003 and 0·050 suggest that little or no autocorrelation in the sequence of cellotriosyl and cellotetraosyl residues exists in the barley β-glucan. Thus the residues are arranged in an apparently independent (random) fashion. The presence of oligosaccharides of *DP* 7 confirms that the β-glucan preparation does not consist of two discrete populations of molecules, one containing cellotriosyl units exclusively and the other cellotetraosyl residues, since in this situation only oligosaccharides of *DP* 6 and *DP* 8 would be released and the values of ρ would approach unity.

The conclusion that the cellotriosyl and cellotetraosyl units are randomly arranged is based on four assumptions. Firstly, it is assumed that their arrangement in penultimate oligosaccharide hydrolysis products, which account for approx. one-third of the polysaccharide at the time of analysis, is representative of their arrangement in the intact polysaccharide.

Secondly, it is necessary to assume that the rates of enzymic hydrolysis of the three penultimate oligosaccharides are approximately equal if a true indication of their relative abundance in the hydrolysate is to be obtained. Thirdly, the rates of production of the oligosaccharides from the polysaccharide are assumed to be approximately the same. Although precise kinetic data on these rates are not available, the oligosaccharides reach maximum levels at the same time and disappear from the hydrolysate at approximatey equal rates (Fig. 1). Furthermore, all values for the autocorrelation calculated from the relative abundance of oligosaccharides at hourly intervals between 2 h and 6 h hydrolysis are close to zero (data not shown). Differences in relative abundance of the oligosaccharides in hydrolysates of enzymes I and II have been noted (see Results section) but do not affect the conclusion, since for both enzymes the values for the autocorrelation are near zero.

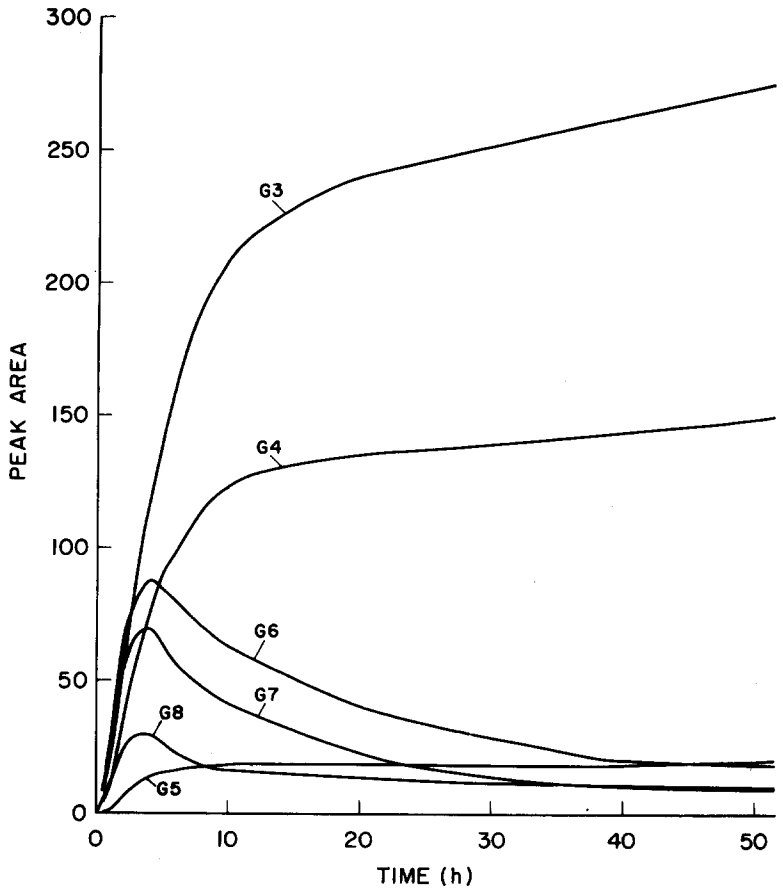


Fig. 1. Time course of release of oligosaccharides from barley β-glucan by barley (1→3),(1→4)-β-D-glucan 4-glucanohydrolase II. **G3**–**G8** represent oligo-β-glucosides of *DP* 3–8 respectively.

The fourth assumption relates to the enzymic release of oligosaccharides containing long blocks of contiguous (1→4)-β-glucosyl residues with single (1→3)-linked residues at their reducing termini (Woodward *et al.*, 1983*b*). It is assumed that these products, which include oligosaccharides of *DP* 6–8, do not significantly affect the relative abundance of the penultimate oligosaccharides. We have re-calculated values of ρ of −0·011 and 0·028 for enzymes I and II respectively after subtracting the maximum possible proportions of oligosaccharides of *DP* 6–8

attributable to the long blocks of contiguous (1→4)-linkages (values taken from Woodward & Fincher, 1982*b*). The re-calculated values of ρ remain close to zero and indicate that this assumption does not invalidate the conclusion that cellotriosyl and cellotetraosyl residues are arranged independently.

It should be emphasised that the mathematical analysis used in the present study assumes a two-state Markov chain and only defines the dependence in the sequence of adjacent cellotriosyl and cellotetraosyl residues. The method would not detect repeating sequences of higher *DP*, for example of the type envisaged by Pattee (1961).

The molecular consequence of a random arrangement of cellotriosyl and cellotetraosyl residues in the barley β-glucan sample is likely to be an irregular conformation, with little tendency for extensive or close alignment of the molecules, and this is indeed consistent with its solubility in water and its physicochemical properties (Woodward *et al.*, 1983*a*,*b*). However, relatively long blocks of adjacent (1→4)-linkages are present in significant proportions and may allow the formation of intermolecular junction zones (Woodward *et al.*, 1983*b*).

The fine structural detail of the β-glucan raises questions regarding the mechanism for its biosynthesis. It is now apparent that proposed mechanisms should account for a random distribution of the cellotriosyl and cellotetraosyl units, and the presence of blocks of up to 10 or more adjacent (1→4)-linkages. If the molecule is assembled by polymerising, in (1→3)-β-linkage, cellodextrins of variable length, then their abundance in the polysaccharide could be controlled by cellodextrin or cellodextrin-intermediate concentrations in a precursor pool, or by the relative affinities of the synthase for precursor molecules. Brett & Northcote (1975) and Hopp *et al.* (1978) suggest that lipid- or protein-linked intermediates might be involved in the elongation of β-glucan chains.

Henry & Stone (1982*a*) could find no evidence for the involvement of lipid intermediates in β-glucan synthesis, which they suggested occurs by direct addition of single glucosyl residues from uridine diphosphate glucose to the elongating polysaccharide chain. The mechanisms for insertion of (1→3)- or (1→4)-linkages and control of their distribution are not known, although the elaboration of cellodextrin blocks on the enzyme itself and their subsequent addition, *en bloc*, to the chain through (1→3)-linkages is not ruled out. An important observation in their work was that the relative abundance of cellotriosyl and cello-

tetraosyl residues in the β-glucan synthesised *in vitro* is significantly affected by the concentration of uridine diphosphate glucose in the assay mixture (Henry & Stone, 1982*b*). Thus the availability of substrate may influence the composition of β-glucans synthesised *in vivo*; whether substrate concentration would affect the sequence dependence of cellotriosyl and cellotetraosyl units is not known.

ACKNOWLEDGEMENTS

This work was supported by grants from the Australian Research Grants Scheme. We thank Dr A Bacic, Dr D. R. Phillips, Dr G. M. Polya and Mr S. Sheather for critically reviewing the manuscript.

REFERENCES

Brett, C. T. & Northcote, D. H. (1975). *Biochem. J.* **148**, 107.

Clarke, A. E. & Stone, B. A. (1963). *Rev. Pure Appl. Chem.* p. 134.

Henry, R. J. & Stone, B. A. (1982*a*). *Biochem. J.* **203**, 629.

Henry, R. J. & Stone, B. A. (1982*b*). *Plant Physiol.* **69**, 632.

Hoel, P. G., Port, S. C. & Stone, C. J. (1972). *Introduction to stochastic processes*, chapter 1, Houghton Mifflin Co., Boston, USA.

Hopp, H. E., Romero, P. A., Daleo, G. R. & Pont Lezica, R. (1978). *Eur. J. Biochem.* **84**, 561.

Pattee, H. H. (1961). *Biophys. J.* **1**, 683.

Woodward, J. R. & Fincher, G. B. (1982*a*). *Eur. J. Biochem.* **121**, 663.

Woodward, J. R. & Fincher, G. B. (1982*b*). *Carbohydr. Res.* **106**, 111.

Woodward, J. R., Phillips, D. R. & Fincher, G. B. (1983*a*). *Carbohydrate Polymers* **3**, 143.

Woodward, J. R., Fincher, G. B. & Stone, B. A. (1983*b*). *Carbohydrate Polymers* **3**, 207.

APPENDIX I

In what follows we represent the sequence of cellotriosyl and cellotetraosyl units by a Markov chain of 0's and 1's respectively. Let X_n denote the nth state (unit) of the chain, and let $\underline{\pi}_n = [\pi_n, 1-\pi_n]$, where $\pi_n = P\{X_n = 0\}$, be its distribution vector. When $\underline{\pi}_n$ is free of n, we say the chain has a stationary distribution, and call it $\underline{\pi} = [\pi, 1-\pi]$. Let the transition matrix of the chain be denoted $P = [P_{ij}]$, where

$P_{ij} = P\{X_{n+1} = j \mid X_n = i\}$, $i = 0, 1$; $j = 0, 1$. In the notation section of 2.2.1,

$$P = \begin{bmatrix} p & 1-p \\ q & 1-q \end{bmatrix}$$

When $0 < p < 1$ and $0 < q < 1$ it can be shown (see Hoel *et al.*, 1972) that there exists a unique stationary distribution which is given by the solution of the matrix equation $\underline{\pi} = \underline{\pi}P$; the solution is $\pi = q/(1-p+q)$. It may also be shown that the relative abundance of 0's in a sequence of the chain approaches π as the sequence grows without bound.

The autocorrelation of a Markov chain with stationary distribution is defined to be the correlation between X_n and X_{n+1}:

$$\rho = \frac{E[X_n X_{n+1}] - E(X_n)\,E(X_{n+1})}{\sqrt{\text{Var } X_n}\sqrt{\text{Var } X_{n+1}}}$$

In our case, $EX_n = 1-\pi$, $\text{Var } X_n = \pi(1-\pi)$, and $E[X_n X_{n+1}] = (1-\pi)(1-q)$, so $\rho = p - q$. It is easy to see that X_n, X_{n+1} are independent if and only if $p = q$ (i.e. $\rho = 0$).

Now if we are given a long sequence of the chain which has been broken down into individual 0's and 1's, we may estimate π by the proportion of 0's in the sample, but this does not help us estimate p, q, and ρ. For example, the following three chains all have the same stationary distribution (with $\pi = 0{\cdot}7$), but ρ is different in the three cases:

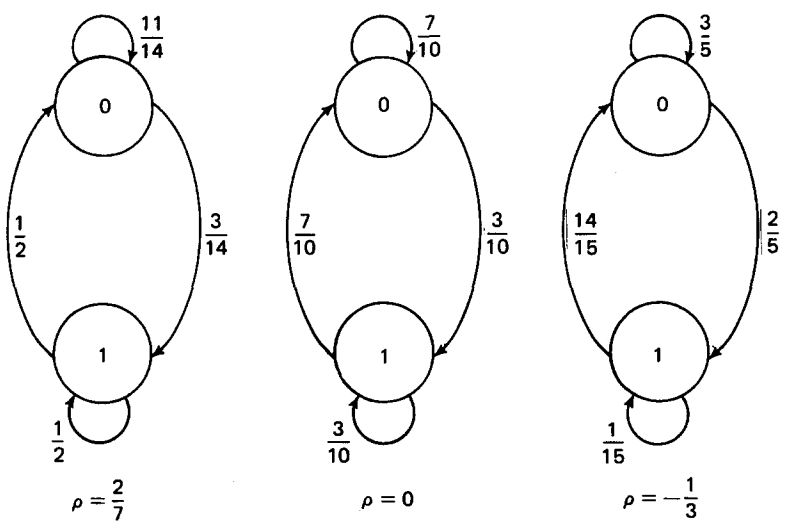


Fig. A.1.

Suppose now that the long sequence were broken down into segments of length two units. Even if we cannot distinguish the '01' pieces from the '10' pieces, we have enough information to estimate p, q and ρ, as shown below.

Formally, we have a new four-state Markov chain given by $Y_n = (X_{2n-1}, X_{2n})$, $n = 1, 2, \ldots$, where the possible states are the pairs (0, 0), (0, 1), (1, 0) and (1, 1). Let P_Y and $\underline{\pi}_Y$ denote the transition matrix and stationary distribution for the chain $\{Y_n\}$. Then P_Y may be derived from P, and $\underline{\pi}_Y$ from $\underline{\pi}_Y = \underline{\pi}_Y P_Y$. This is rather messy to solve for $\underline{\pi}_Y$. However, let $Z_n = (X_{n-1}, X_n)$, $n = 2, 3, \ldots$; then $\underline{\pi}_Z = \underline{\pi}_Z P_Z$ is easy to solve for $\underline{\pi}_Z$ and moreover $\underline{\pi}_Z = \underline{\pi}_Y$. To see this, note that $Z_{2n} = Y_n$, so $P_Z^2 = P_Y$. It then follows from $\underline{\pi}_Z = \underline{\pi}_Z P_Z^2 = \underline{\pi}_Z P_Y$ and the uniqueness of the stationary distribution that $\underline{\pi}_Z = \underline{\pi}_Y$. The solution of $\underline{\pi}_Y = \underline{\pi}_Y P_Y$ is

$$\underline{\pi}_Y = [\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11}]$$

where $\pi_{00} = (pq)/(1-p+q)$, $\pi_{01} = (1-p)q/(1-p+q) = \pi_{10}$, and π_{11} is determined by $\pi_{00} + \pi_{01} + \pi_{10} + \pi_{11} = 1$.

Example

If $\{X_n\}$ is given by the third chain of Fig. A.1, the corresponding four-state chains $\{Y_n\}$ and $\{Z_n\}$ are shown below in Fig. A.2. Note that only four of the 16 transition probabilities are shown for $\{Y_n\}$.

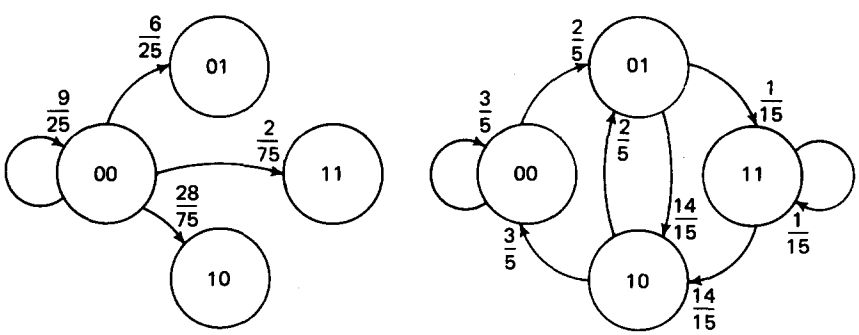


Fig. A.2. $\underline{\pi}_Y = [0{\cdot}42, 0{\cdot}28, 0{\cdot}28, 0{\cdot}02] = \underline{\pi}_Z$.

To simplify the notation we hereafter and in the text proper refer to the four states (0, 0), (0, 1), (1, 0), and (1, 1) by 0, 1, 2 and 3, and replace $[\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11}]$ by $[\pi_0, \pi_1, \pi_2, \pi_3]$. In this notation

$$p = \frac{\pi_0}{\pi_0 + \pi_1}$$

and

$$q = \frac{\pi_1}{1 - \pi_0 - \pi_1}$$

APPENDIX II

Error analysis

In Appendix I it was shown that

$$\rho = p - q = \frac{\pi_0}{\pi_0 + \pi_1} - \frac{\pi_1}{1 - \pi_0 - \pi_1}$$

so that ρ could be estimated by

$$\hat{\rho} = \frac{\hat{\pi}_0}{\hat{\pi}_0 + \hat{\pi}_1} - \frac{\hat{\pi}_1}{1 - \hat{\pi}_0 - \hat{\pi}_1}$$

where $\hat{\pi}_0$ is the estimate of π_0 based on the relative proportion of *DP* 6 in the sample, and $\hat{\pi}_1$ is $\frac{1}{2}$ the relative proportion of *DP* 7 in the sample.

We assume that $\hat{\pi}_0 = \pi_0 + \epsilon_0$, $\hat{\pi}_1 = \pi_1 + \epsilon_1$, where ϵ_0, ϵ_1 are random errors with means $E(\epsilon_0) = 0 = E(\epsilon_1)$, variances $\sigma_0^2 = \mathrm{Var}(\epsilon_0)$, $\sigma_1^2 = \mathrm{Var}(\epsilon_1)$, and covariance $\sigma_{01} = \mathrm{Cov}(\epsilon_0, \epsilon_1)$. Then $\hat{\rho}$ may be written

$$\hat{\rho} = \rho + \frac{\pi_1\epsilon_0 - \pi_0\epsilon_1}{(\pi_0 + \pi_1)^2} - \frac{(\pi_0 - 1)\,\epsilon_1 - \pi_1\epsilon_0}{(1 - \pi_0 - \pi_1)^2} + \Delta$$

where Δ is a sum of higher order terms in ϵ_0, ϵ_1 which may be neglected. It follows that

$$E\hat{\rho} \approx \rho$$

and

$$\mathrm{Var}\,\hat{\rho} \approx \frac{\begin{array}{r}[1 - 2(\pi_0 + \pi_1)]^2\sigma_0^2 + [\pi_0 - \pi_0^2 + \pi_1^2]^2\sigma_1^2 \\ -2\sigma_{01}[1 - 2(\pi_0 + \pi_1)]\,[\pi_0 - \pi_0^2 - \pi_1^2]\end{array}}{(\pi_0 + \pi_1)^4(1 - \pi_0 - \pi_1)^4}$$

The standard error of $\hat{\rho}$ is by definition $\sqrt{\text{Var}\,\hat{\rho}}$, and this can be estimated from the data.

For example, there were three independent pairs of estimates for enzyme I, namely

π_0	0·510	0·510	0·517
π_1	0·2055	0·2045	0·2015

These three pairs may be used to estimate σ_0^2, σ_1^2, σ_{01}, and $\hat{\pi}_0$, $\hat{\pi}_1$ are taken as the respective averages of the pairs. The values are then substituted into the expression for the standard error. A similar analysis was used on enzyme II.