Biophysical Chemistry 13 (1981) 213–222
North-Holland Publishing Company

THE AVERAGE NUMBER OF DISTINCT SITES VISITED BY A ONE-DIMENSIONAL RANDOM WALKER AND ITS APPLICATION TO ISOTOPE EXCHANGE IN POLYPEPTIDES *

Yoshinori NAGAI
Department of Applied Physics, Waseda University, Tokyo 160, Japan

Hiroshi ASAI
Department of Physics, Waseda University, Tokyo 160, Japan

and

Takashi TSUCHIYA
Laboratory of Molecular Biology, National Institute of Arthritis, Metabolism and Digestive Diseases,
National Institutes of Health, Bethesda, Maryland 20205, U.S.A.

Received 19 May 1980
Revised manuscript received 10 July 1980

The average number of distinct sites visited by a random walker moving with arbitrary transition probability on a one-dimensional lattice is calculated. Asymptotic forms of this quantity for both asymmetric and symmetric random walks are determined, and an exact solution for the latter case is also given for any number of steps. The average number of sites visited is then analyzed for intermediate numbers of steps by introducing an exponent. This approach is applied to explain the results of isotope exchange experiments in polypeptides, and applications of asymmetric random walks to other biological problems are briefly discussed.

1. Introduction

We consider here the problem of the average number of distinct sites visited by a random walker on a one-dimensional lattice. As long as the transition probability of the walker is symmetric, i.e., $p = q$, where p is the probability of stepping to the right and $q = 1 - p$ is that to the left, this is not a new problem at all. Dvoretzky and Erdös [1] are generally recognized as the first workers to consider the average number of distinct sites S_n visited by a random walker and they obtained asymptotic forms of S_n in the limit of a large number of steps n, for some lattices with $d \geqslant 2$, where d represents the space dimension. They showed that $S_n \sim cn$ where c is a constant for $d \geqslant 3$ and $S_n \sim n/\ln n$ for $d = 2$. Vineyard [2] gave the asymptotic form of S_n for a one-dimensional lattice as $(8n/\pi)^{1/2}$, and gave values of c for several kinds of 3-d lattices. Montroll summarized all these results [3] and furthermore, Montroll and Weiss calculated [4] exact values of S_n for small numbers of steps n. All of these authors focused their attention on the symmetric random walk case.

Our main reason for reconsidering this rather old problem is the need to extend the above results to asymmetric cases ($p \neq q$) when we apply the concept of the numbers of distinct sites visited by the random walker to some biophysical problems. In applying various results (not only S_n) of random walk problems to biophysical systems such as linear chains of enzymes and biopolymers, it is necessary to consider asymmetric cases and not asymptotic forms because such biophysical systems are generally not as homogeneous as those treated in, e.g., solid state physics, and the number of steps cannot be *a priori* regarded as infinite.

---

* The authors would like to dedicate the present paper to Professor Nobuhiko Saitô on the occasion of his sixtieth birthday.

0 301–4622/81/0000–0000/S 2.50

A mathematical derivation of the quantity S_n for asymmetric random walks on a one-dimensional lattice is presented in this paper as a basis for applications of the results to biological systems. The application to linearly aggregated enzyme systems will be described elsewhere [5] and the problem of isotope exchange in polypeptides will also be considered precisely in a separate paper [6] in which the quantity S_n plays a crucial role. However, we report in the present paper elementary but essential treatment of the latter problem in order to show how useful the concept of the average number of sites visited by an asymmetric random walker is in certain biophysical systems.

In sect. 2 we consider the average number of distinct sites S_n (actually $L_n = S_n - 1$) and its asymptotic forms in the limit of large n for both asymmetric and symmetric cases. For the latter case we also derive an exact solution of L_n for any n. In sect. 3 we examine the behavior of L_n for intermediate values of n; this is of interest in biophysical applications. The singular nature of the symmetric random walk ($p = q$) is also analyzed. In sect. 4 we consider the problem of isotope exchanges in polypeptides in terms of the average number of distinct sites visited by an asymmetric random walker. We show theoretically the existence of the so-called frozen region in a polypeptide which was observed experimentally by Ikegami and Kono [7]. Discussion is given in sect. 5 where various other possibilities of application of the results of this paper to biological systems are also suggested.

2. The average number of sites visited

We restrict our discussion here to random walks on a one-dimensional lattice. The average number of distinct lattice points visited by a random walker after n steps is represented by S_n, and the probability that the lattice point l has been visited at least once in the first n steps is denoted by $F_n(l)$. $F_n(l)$ can be written in terms of $f_j(l)$, the probability that the random walker visits the site l for the first time on the jth step, or the first passage time probability, as

$$F_n(l) = \sum_{j=1}^{n} f_j(l) . \tag{1}$$

If we sum up this quantity over all the lattice points but the origin, we can obtain the quantity of interest S_n, namely,

$$S_n = \sum_{l\neq 0} F_n(l) . \tag{2}$$

Following Montroll [3,4] we define Δ_n as

$$\Delta_n \equiv S_n - S_{n-1}$$

$$= \sum_{l\neq 0} [F_n(l) - F_{n-1}(l)] = \sum_{l\neq 0} f_n(l) . \tag{3}$$

Then S_n is given as

$$S_n = 1 + \Delta_1 + \Delta_2 + \Delta_3 + \cdots + \Delta_n \tag{4}$$

with obvious

$$S_0 = 1 . \tag{5}$$

The generating function of Δ_n is defined as

$$\Delta(z) = \sum_{n=1}^{\infty} z^n \Delta_n , \tag{6}$$

which can be related to the generating function of the first passage time probability, i.e.,

$$F(z,l) \equiv \sum_{n=1}^{\infty} z^n f_n(l) , \tag{7}$$

as

$$\Delta(z) = \sum_{l\neq 0} \sum_{n=1} z^n f_n(l) = \sum_{l\neq 0} F(z,l) . \tag{8}$$

Since the generating function $F(z,l)$ for the 1-d case is already given by Feller [8] as

$$F(z,l) = \begin{cases} \left(\dfrac{1-(1-4pqz^2)^{1/2}}{2qz}\right)^{|l|} & \text{for } l > 0 \\ \left(\dfrac{1-(1-4pqz^2)^{1/2}}{2pz}\right)^{|l|} & \text{for } l < 0 \end{cases} , \tag{9}$$

we obtain the generating function $\Delta(z)$ as

$$\Delta(z) = \frac{(1-4pqz^2)^{1/2}}{1-z} - 1 . \tag{10}$$

Let us define a quantity L_n as

$$L_n = S_n - 1 \,. \tag{11}$$

Then its generating function $L(z)$ is

$$L(z) = \sum_{n=1}^{\infty} z^n L_n \,. \tag{12}$$

Taking into account (4), (6) and the range of summation, we have

$$L(z) = \sum_{i=1}^{\infty} z^i \Delta_i \sum_{n=i}^{\infty} z^{n-i} = \frac{\Delta(z)}{1-z} \,. \tag{13}$$

Expanding $L(z)$ in terms of z, we find the coefficient L_n of the term z^n as,

$$L_n = n - 2 \sum_{r=1}^{[\frac{n}{2}]} (n - 2r + 1) \frac{{}_{2r-1}C_r}{2r-1} (pq)^r$$

$$= n - \sum_{r=1}^{[\frac{n}{2}]} (n - 2r + 1) \frac{{}_{2r}C_r}{2r-1} (pq)^r \quad \text{for } n \geqslant 2, \tag{14}$$

where

$$[\tfrac{n}{2}] = \begin{cases} \dfrac{n}{2} & \text{for } n\text{: even} \\[2ex] \dfrac{n-1}{2} & \text{for } n\text{: odd} \end{cases} \tag{15}$$

and obviously $L_1 = 1$. Numerical values of (14) are plotted in fig. 1 and fig. 2 for various combinations of p and q.

In random walk problems one is usually interested

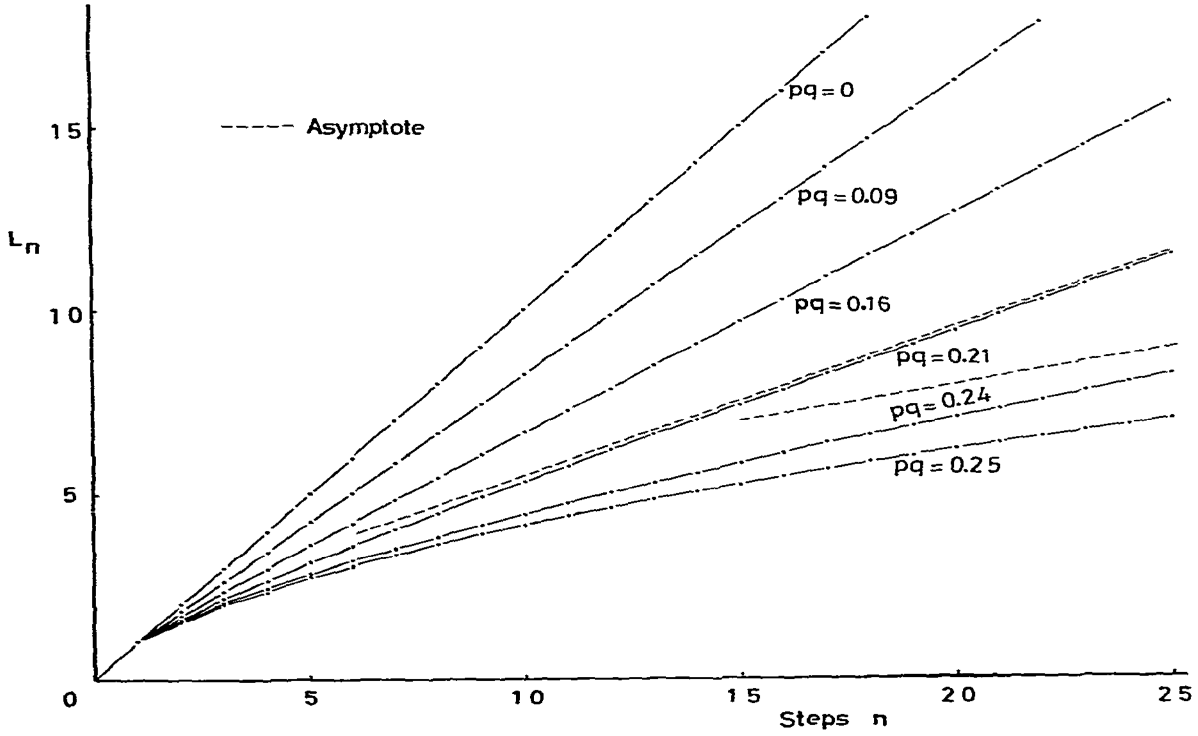

Fig. 1. Numerical values of L_n for various combinations of p and q calculated from (14). Broken lines are asymptotes (21).

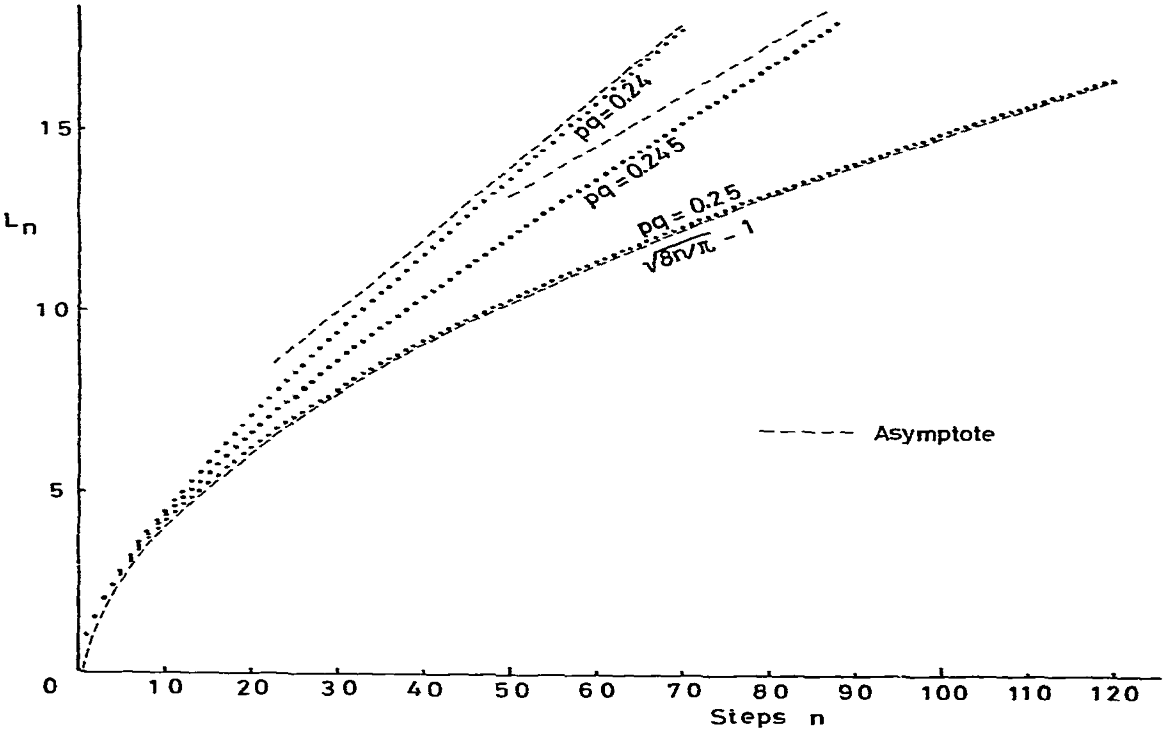

Fig. 2. Numerical values of L_n and asymptotes for nearly symmetric random walks.

in asymptotic forms of quantities of interest in the limit of large n. To derive the asymptotic form of L_n we must separate the symmetric case ($p = q$) from asymmetric ones ($p \neq q$).

Firstly we consider the asymmetric case ($p \neq q$). Using the power series expansions

$$(1-4pq)^{1/2} = 1 - \sum_{r=1}^{\infty} \frac{{}_{2r}C_r}{2r-1} (pq)^r \tag{16}$$

and

$$\frac{1}{(1-4pq)^{1/2}} = 1 + \sum_{r=1}^{\infty} {}_{2r}C_r (pq)^r , \tag{17}$$

the expression (14) becomes

$$L_n = (1-4pq)^{1/2} n + n \sum_{r=[\frac{n}{2}]+1}^{\infty} \frac{{}_{2r}C_r}{2r-1} (pq)^r + \frac{1}{(1-4pq)^{1/2}} - 1 - \sum_{r=[\frac{n}{2}]+1}^{\infty} {}_{2r}C_r (pq)^r . \tag{18}$$

Since $(p-q)^2 = 1 - 4pq$, (18) can be written in the form

$$L_n = |p-q| n + \frac{1}{|p-q|} - 1 + \delta_n , \tag{19}$$

where

$$\delta_n = \sum_{r=[\frac{n}{2}]+1}^{\infty} {}_{2r}C_r \left(\frac{n}{2r-1} - 1 \right) (pq)^r . \tag{20}$$

As shown in appendix A, δ_n defined above approaches

0_- for large n. Hence we have the asymptotic form of L_n for the asymmetric random walker ($p \neq q$), as

$$L_n \sim |p-q|\,n + \frac{1}{|p-q|} - 1 \tag{21}$$

which is obviously a linear function of n. The asymptotes (21) are shown in fig. 1 and fig. 2 for several combinations of p and q.

For the symmetric case ($p = q$), we can reproduce from (14) the well-known asymptotic form of L_n as follows. The expression (14) for this case becomes

$$L_n = n\left(1 - \sum_{r=1}^{[\frac{n}{2}]} \frac{{}_{2r}C_r}{2r-1}\left(\tfrac{1}{4}\right)^r\right) + \sum_{r=1}^{[\frac{n}{2}]} \frac{{}_{2r}C_r}{2^{2r}}\,. \tag{22}$$

Since

$$\frac{{}_{2r}C_r}{2^{2r}} = \frac{(2r-1)!!}{(2r)!!}\,, \tag{23}$$

eq. (22) can be rewritten as

$$L_n = n\left(\frac{1}{2} - \sum_{r=1}^{[\frac{n}{2}]-1} \frac{(2r-1)!!}{(2r+2)!!}\right) + \sum_{r=1}^{[\frac{n}{2}]} \frac{(2r-1)!!}{(2r)!!}\,. \tag{24}$$

The well-known formula [9]

$$\sum_{r=1}^{n} \frac{(2r-1)!!}{(2r+2)!!} = \frac{1}{2} - \frac{(2n+1)!!}{(2n+2)!!} \tag{25}$$

is used for the second term. In order to perform the summation of the third term we construct the following relation

$$\sum_{r=1}^{n} \frac{(2r-1)!!}{(2r)!!} = \frac{(2n+1)!!}{(2n)!!} - 1 \tag{26}$$

which can be easily proved by mathematical induction. Consequently we have

$$L_n = \frac{{}_{2[\frac{n}{2}]}C_{[\frac{n}{2}]}}{2^{2[\frac{n}{2}]}}\left(n + 2[\tfrac{n}{2}] + 1\right) - 1\,. \tag{27}$$

This is an exact expression of L_n for the symmetric random walk; such a relation has never previously been reported. When the number of steps n is very large, $[\frac{n}{2}]$ is regarded as $n/2$ and Stirling's formula can be used. Hence we obtain the asymptotic form of L_n as

$$L_n \sim \sqrt{8n/\pi} - 1\,, \tag{28}$$

which has already been reported by Vineyard [2], Montroll [3] and Montroll and Weiss [4]; this is also plotted in fig. 2.

3. The exponent for intermediate numbers of steps

For systems with high asymmetry, i.e., for which

$$x = pq \tag{29}$$

is small, L_n approaches the asymptote (21) quickly. For almost symmetric systems with $x \approx 1/4$, however, it takes a long time (large number of steps) to approach the asymptote. In applying the results of this paper to biological systems it is necessary to consider almost but not exactly symmetric random walks at intermediate numbers of steps, since it may not be possible or convenient in biological experiments to wait a sufficiently long time to permit the use of asymptotes for analysis of the results.

For this purpose, fig. 1 and fig. 2 are obviously inappropriate, since at small n the curves do not change appreciably for a slight variation of x where $x \approx 1/4$. We therefore introduce an exponent for intermediate steps as follows. We can express an arbitrary monotonous function $y(t)$ in terms of an approximate function of the form

$$y = \alpha t^{\lambda} + \beta \tag{30}$$

locally. If two points (y_1, t_1) and (y_2, t_2) are close enough, the exponent λ can be written as

$$\lambda = \frac{t_2 y_2' - t_1 y_1'}{y_2 - y_1}\,. \tag{31}$$

For $t_1 = t$ and $t_2 = t + \Delta t$ we have

$$\lambda = 1 + t\,\frac{y''}{y'} + O(\Delta t)\,. \tag{32}$$

We define the exponent at t, using (32) as

$$\lambda(t) \equiv \lim_{\Delta t \to 0} \lambda\,,$$

hence

$$\lambda(t) = 1 + t\,\frac{\mathrm{d}}{\mathrm{d}t} \ln y'\,. \tag{33}$$

When the independent variable is discrete, we define

$$y'_n = \frac{y_{n+1} - y_n}{2} + \frac{y_n - y_{n-1}}{2} \tag{34}$$

and then the exponent at the point n can be expressed as

$$\lambda(n) \equiv \frac{1}{2}\left(\frac{(n+1)y'_{n+1} - ny'_n}{y_{n+1} - y_n} + \frac{ny'_n - (n-1)y'_{n-1}}{y_n - y_{n-1}}\right) . \tag{35}$$

We can now use this definition of $\lambda(n)$ for the function L_n. From eq. (14) we have

$$\delta L_n \equiv L_n - L_{n-1} = 1 - \sum_{r=1}^{[\frac{n}{2}]} \frac{{}_{2r}C_r}{2r-1} (pq)^r , \tag{36}$$

then the exponent becomes

$$\lambda(n) = \frac{1}{4}\left(2 + \frac{(n+1)\delta L_{n+2} - n\delta L_n}{\delta L_{n+1}} + \frac{n\delta L_{n+1} - (n-1)\delta L_{n-1}}{\delta L_n}\right) . \tag{37}$$

Using (36) and (37), we can calculate $\lambda(n)$ for intermediate values of n. The results are shown in fig. 3; it is clear that the shape of the plot changes considerably for a small variation in x near $1/4$.

If we write eq. (37) as

$$\lambda(n) = 1 - \frac{n}{4} \frac{{}_{2[\frac{n}{2}]+2}C_{[\frac{n}{2}]+1}}{2[\frac{n}{2}]+1} \times (pq)^{[\frac{n}{2}]+1}\left(\frac{1}{\delta L_{n+1}} + \frac{1}{\delta L_n}\right) \quad \text{for } n\text{: odd} \tag{38}$$

and

$$\lambda(n) = 1 - \tfrac{1}{4}\, {}_{2[\frac{n}{2}]}C_{[\frac{n}{2}]} (pq)^{[\frac{n}{2}]} \times \left\{1 + \left(1 - \frac{1}{2[\frac{n}{2}]+1}\right)(4pq)\right\} \frac{1}{\delta L_n} \quad \text{for } n\text{: even} , \tag{39}$$

we can determine the value of $x = pq$ easily from (38) even when x is very close to $1/4$.

Since δL_n can be rewritten as

$$\delta L_n = (1 - 4pq)^{1/2} + \epsilon_n , \tag{40}$$

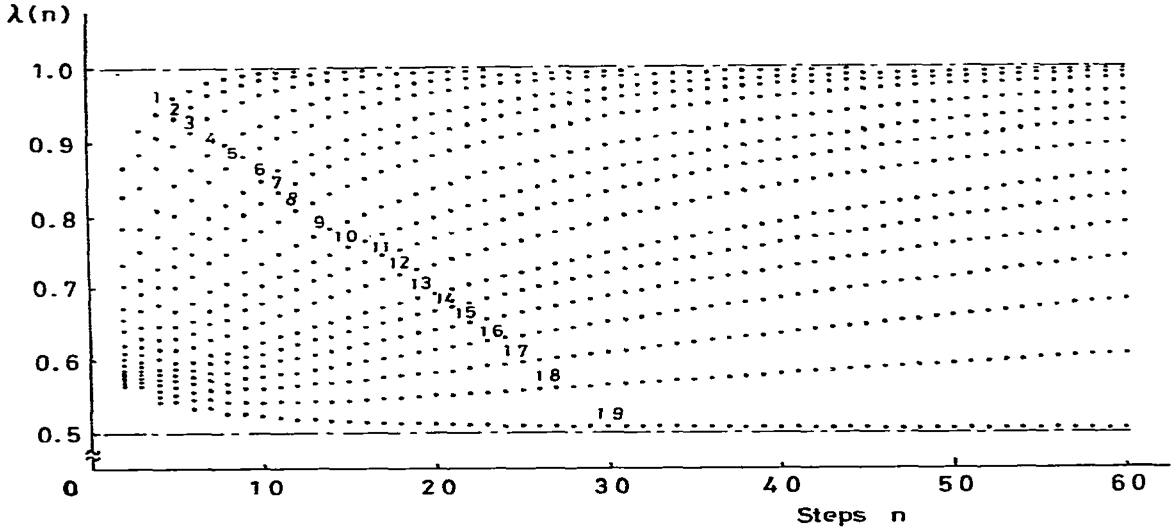

Fig. 3. Numerical values of the exponent $\lambda(n)$ calculated from eq. (37). The numbers from 1 to 19 denote the following values of pq; 1 : 0.14, 2 : 0.16, 3 : 0.18, 4 : 0.20, 5 : 0.21, 6 : 0.22, 7 : 0.225, 8 : 0.23, 9 : 0.235, 10 : 0.238, 11 : 0.24, 12 : 0.242, 13 : 0.244, 14 : 0.245, 15 : 0.246, 16 : 0.247, 17 : 0.248, 18 : 0.249, 19 : 0.25.

where

$$\epsilon_n = \sum_{r=[\frac{n}{2}]+1}^{\infty} \frac{{}_{2r}C_r}{2r-1}(pq)^r \tag{41}$$

and

$$0 \leqslant \epsilon_n < \frac{{}_{2[\frac{n}{2}]}C_{[\frac{n}{2}]}}{2^{2[\frac{n}{2}]}}(4pq)^{[\frac{n}{2}]+1} , \tag{42}$$

it is straightforward to show that

$$\lim_{n\to\infty} \lambda(n) = 1 \qquad \text{for } p \neq q \tag{43}$$

and

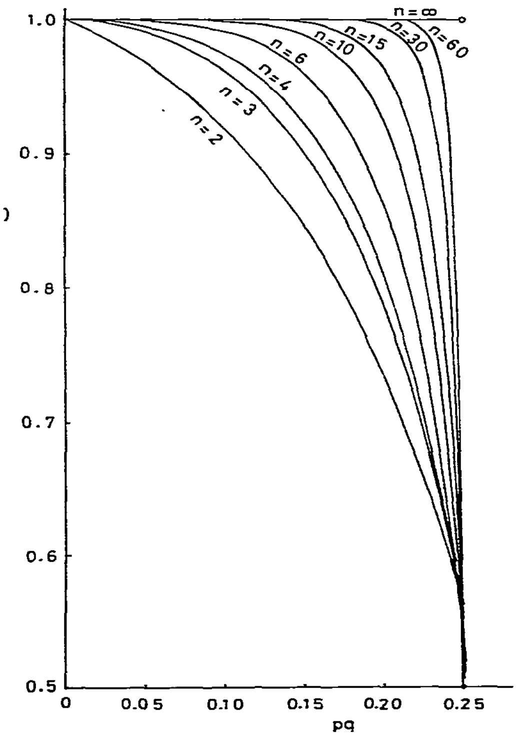

Fig. 4. Numerical values of the exponent $\lambda(n)$ as a function of pq calculated from eq. (37). The asymptotic results (43) and (44) are also depicted.

$$\lim_{n\to\infty} \lambda(n) = \tfrac{1}{2} \qquad \text{for } p = q . \tag{44}$$

These asymptotic results (43) and (44) which are depicted in fig. 4 indicate that our definition $\lambda(n)$, the exponent for intermediate steps, is reasonable.

4. Application to isotope exchanges in polypeptides

In this section we consider the problem of isotope exchanges in polypeptides using the concept of the average number of distinct sites S_n (or L_n) visited by an asymmetric random walker. We will give an example in which the term (19) other than the first term contribute significantly to the experimental results.

First, we outline the experimental procedure and results very briefly. All the hydrogen atoms (H) of NH groups of each monomer unit (amino acid residue) of a polypeptide in an aqueous solution are first replaced by isotopes, i.e., deuterium (D) or tritium (T). Only when a monomer is in the coil state can the isotope be exchanged with an H atom in the solution. The amount of isotope remaining in the polypeptide can be determined at intervals. If the pH of the solution is set at a value that corresponds to the helix-favored side of the helix–coil transition region, almost all of the monomer units will be in the helix state. Therefore, we would expect little isotope exchange to occur. The experimental results given by Ikegami and Kono for copolymer LGA III, for example, indicate that isotope exchange occurred in about 35% of the monomers, and about 65% of the monomers were in a rigid helix state, forming a so-called frozen region [7]. We now pay attention to the fact described in [7] that a polypeptide has two kinds of helix region, i.e., one is rigid (or frozen) and another is flexible, in a helix-favored solution. It is considered that a helical structure in a flexible region can easily change its state to coil state for a moment and return to helix state again, i.e., the state of a monomer unit in a flexible region fluctuates from time to time although the ensemble average of helix content is constant. We will explain these experimental results using a model in which the marginal point between helix and coil part moves like a weakly asymmetric random walker on a polypeptide chain.

We can apply the concept discussed in the previous sections of the present paper to this problem. First we assume that helix to coil transition can start only at both ends of a polypeptide chain, i.e., helix content is about 100%. We regard the marginal point between the helix part and coil part as a random walker with asymmetric transition probabilities, p in the direction of the helix region and q in that of the coil region. We furthermore assume that an isotope is released only when the monomer changes its state to the coil state for the first time. Hence the problem of determining the amount of released isotope is reduced to the problem of calculating the average number of distinct sites visited by a random walker; this was treated in sect. 2.

Let the quantity L_n defined in (11) be divided as follows

$$L_n = L_n^+ + L_n^- \tag{45}$$

where

$$L_n^+ = \sum_{l=1}^{\infty} F_n(l)\,, \tag{46}$$

$$L_n^- = \sum_{l=-1}^{-\infty} F_n(l)\,. \tag{47}$$

Corresponding generating functions $L^+(z)$ and $L^-(z)$ can easily be obtained in the same way as

$$L_n^\pm(z) = \frac{1}{2}\left(\frac{\Delta(z)}{1-z} \pm (p-q)\frac{z}{(1-z)^2}\right)\,, \tag{48}$$

where $\Delta(z)$ has been already derived in (10). $L_n^\pm$ can be easily obtained in the same way as

$$L_n^\pm = \tfrac{1}{2}L_n \pm \tfrac{1}{2}(p-q)n \qquad \text{for } n \geqslant 2$$

$$L_1^\pm = \tfrac{1}{2}[1 \pm (p-q)]\,. \tag{49}$$

For $p > q$

$$L_n^+ = |p-q|n + \frac{1}{2}\left(\frac{1}{|p-q|} - 1 + \delta_n\right) \tag{50}$$

and for $p < q$

$$L_n^+ = \frac{1}{2}\left(\frac{1}{|p-q|} - 1 + \delta_n\right) \tag{51}$$

where δ_n was defined in (20). The quantity (51) goes to a finite value for large n;

$$L_\infty^+ = \frac{1}{2}\left(\frac{1}{|p-q|} - 1\right)\,. \tag{52}$$

Table 1
Numerical values of constant frozen region (%) given by eq. (53). p is the transition probability of an asymmetric random walker and N is the degree of polymerization. A frozen region of 65% is obtained at a value of $p \approx 0.486$ for $N = 100$ and at a value of $p \approx 0.493$ for $N = 200$.

p	N	
	100	200
0.495	1.0%	50.5%
0.493	29.6%	64.8%
0.490	51.0%	75.5%
0.488	59.4%	79.7%
0.486	65.4%	82.7%

This result suggests that only a part of a polypeptide can undergo isotope exchange for H atoms in helix-favoured surroundings, when twice L_∞^+ is smaller than the degree of polymerization N. Thus, the existence of a constant frozen region can be accounted for by asymmetric random walk considerations. A constant frozen region (%) in a polypeptide is given by

$$\frac{N - 2L_\infty^+}{N} \times 100\,(\%)\,. \tag{53}$$

Numerical values of the frozen region obtained by (53) are tabulated in table 1 for two degrees of polymerization ($N = 100$ and 200).

5. Discussion

We have considered the average number of distinct sites visited by a random walker with arbitrary transition probability on a one-dimensional lattice. One of the main results is that the asymptote of the quantity is approached as a function of the number of steps n as long as the transition probability is asymmetric ($p \neq q$) without any other terms of lower powers of n. This emphasizes the singular character of the symmetric random walk which has the asymptotic form $n^{1/2}$. For asymmetric random walks with $pq \leqslant 0.21$ it is clear from fig. 1 that the quantity of interest behaves like the asymptote even for n of about 15. For random walks with $pq \sim 0.25$, however, many more steps are required before the quantity approaches the asymptote (see fig. 2). Furthermore, for a small number of steps n the quantity for the case $pq = 0.245$ behaves like the symmetric random walk. Hence it becomes of interest to investigate the case of nearly symmetric random walks with finite n, as was done in sect. 3 by introducing an exponent λ which is obviously 1/2 for the completely symmetric case, $pq = 0.25$, and 1 for the completely asymmetric case, $pq = 0$, i.e., a zipper-like case.

We have considered the problem of isotope exchange in polypeptides as an application of the concept of the average number of distinct sites visited by a random walker in sect. 4. It was possible to account for the existence of the so-called frozen region discovered in experiments on isotope exchange in polypeptides. In this derivation we incorporated some assumptions into the model such that the helix–coil transitions occur only from the ends of the polypeptide and that the marginal point between the helix and the coil parts which is regarded as a random walker can freely go through the ends of the polypeptide. Therefore, it might not be a precise description of the experimental situation of isotope exchange in polypeptides. We will present more accurate treatment of the problem free from these assumptions elsewhere [6]. All we wanted to emphasize in the derivation given in sect. 4 is that regarding the marginal point as an asymmetric random walker we can elucidate the existence of the frozen region which may not be fully explained yet. This simple but essential interpretation remains intact even in the more elaborate treatment of the problem.

Random walks with slightly asymmetric transition probability and with a finite number of steps n are applicable to many other situations, especially in biological systems, e.g., possible asymmetric movement of an enzymatically active subunit in an F-actin polymer at elevated temperatures [5], movement of the repressor on DNA, the cross-bridge movement of actin and myosin systems during muscular contraction, and the folding–unfolding of DNA.

Another possible extension of this approach is to incorporate the probability r of remaining at a site besides those of jumping to the right p and the left q. If this can be done successfully we can discuss further the problem of the span of a polymer chain which was considered by Rubin [10] and others [11], i.e., it will be possible to treat a stiff polymer chain.

Finally, the importance of the exact solution (27) should be noted. It holds for any number of steps n if the random walk is symmetric. This type of solution is rare in random walk theory, where only series expressions or asymptotic solutions are usually obtainable. We were able to obtain the exact solution (27) because of equation (26), which is a special case of the following expression,

$$\sum_{r=1}^{n} \frac{(2r+2m-1)!!}{(2r)!!} = \frac{1}{2m+1}\left(\frac{(2n+2m+1)!!}{(2n)!!} - (2m+1)!!\right), \quad m \geqslant 0 . \tag{54}$$

If we put $m = 0$, we recover equation (26). Equation (54) can be proved by the method employed by Hall and Knight [7]. We also present another formula

$$\sum_{r=m}^{n} \frac{(2r-2m+1)!!}{(2r+2)!!} = \frac{1}{2m-1}\left(\frac{1}{(2m)!!} - \frac{(2n-2m+3)!!}{(2n+2)!!}\right), \quad m \geqslant 1 , \tag{55}$$

which is a generalization of formula (25), as can be easily seen by putting $m = 1$.

Acknowledgements

The authors are indebted to Prof. Saitō for valuable discussions and encouragement. The authors would also like to express their gratitude to Dr. K. Moro and Dr. M. Fujiwara for their critical comments and discussions.

Appendix A

We show in this appendix that δ_n defined by (20) vanishes for large n, except $p = q$. The definition is

$$\delta_n \equiv \sum_{r=[\frac{n}{2}]+1}^{\infty} \left(\frac{n}{2r-1} - 1\right) \frac{{}_{2r}C_r}{2^{2r}} (4pq)^r . \tag{A.1}$$

Noting the relation

$$\frac{{}_{2r+2}C_{r+1}}{2^{2r+2}} = \left(1 - \frac{1}{2r+2}\right) \frac{{}_{2r}C_r}{2^{2r}} , \tag{A.2}$$

we can rewrite (A.1) as follows

$$\delta_n = \frac{{}_{2[\frac{n}{2}]+2}C_{[\frac{n}{2}]+1}}{2^{2[\frac{n}{2}]+2}}\,(4pq)^{[\frac{n}{2}]+1}\left\{\left(\frac{n}{2[\frac{n}{2}]+1}-1\right)\right.$$
$$+\sum_{i=1}^{\infty}\left(\frac{n}{2[\frac{n}{2}]+2i+1}-1\right)$$
$$\left.\times\prod_{l=1}^{i}\left(1-\frac{1}{2[\frac{n}{2}]+2l+2}\right)(4pq)^i\right\}$$

$$\equiv\frac{{}_{2[\frac{n}{2}]+2}C_{[\frac{n}{2}]+1}}{2^{2[\frac{n}{2}]+2}}\,(4pq)^{[\frac{n}{2}]+1}\gamma\,, \tag{A.3}$$

where the expression in braces is written as γ.

Since

$$\left(\frac{n}{2[\frac{n}{2}]+2i+1}-1\right)$$

is negative for any i and tends to -1 for large i, we have the following inequalities for the second term in the braces

$$\sum_{i=1}^{\infty}\left(\frac{n}{2[\frac{n}{2}]+2i+1}-1\right)\prod_{l=1}^{i}\left(1-\frac{1}{2[\frac{n}{2}]+2l+2}\right)(4pq)^i$$
$$<\left(\frac{n}{2[\frac{n}{2}]+3}-1\right)\sum_{i=1}^{\infty}\left(1-\frac{1}{2[\frac{n}{2}]+4}\right)^i(4pq)^i$$
$$=\left(\frac{n}{2[\frac{n}{2}]+3}-1\right)\frac{\left(1-\frac{1}{2[\frac{n}{2}]+4}\right)(4pq)}{1-\left(1-\frac{1}{2[\frac{n}{2}]+4}\right)(4pq)}\,, \tag{A.4}$$

and

$$\sum_{i=1}^{\infty}\left(\frac{n}{2[\frac{n}{2}]+2i+1}-1\right)\prod_{l=1}^{i}\left(1-\frac{1}{2[\frac{n}{2}]+2l+2}\right)(4pq)^i$$
$$>-\sum_{i=1}^{\infty}\prod_{l=1}^{i}\left(1-\frac{1}{2[\frac{n}{2}]+2l+2}\right)(4pq)^i$$
$$>-\sum_{i=1}^{\infty}(4pq)^i=-\frac{4pq}{1-4pq}\,. \tag{A.5}$$

From (A.4) and (A.5) it can be seen that for $p \neq q$ the quantity γ defined in (A.3) is negative definite and

$$\left(\frac{n}{2[\frac{n}{2}]+1}-1\right)+\left(\frac{n}{2[\frac{n}{2}]+3}-1\right)$$
$$\times\frac{\left(1-\frac{1}{2[\frac{n}{2}]+4}\right)(4pq)}{1-\left(1-\frac{1}{2[\frac{n}{2}]+4}\right)(4pq)}$$
$$>\gamma>\left(\frac{n}{2[\frac{n}{2}]+1}-1\right)-\frac{4pq}{1-4pq}\,. \tag{A.6}$$

For other factors of (A.3) it is easy to show that

$$\frac{{}_{2[\frac{n}{2}]+2}C_{[\frac{n}{2}]+1}}{2^{2[\frac{n}{2}]+2}}\sim\frac{1}{\{\pi([\frac{n}{2}]+1)\}^{1/2}} \tag{A.7}$$

for large n and

$$0\leqslant 4pq<1 \qquad \text{for } p\neq q\,. \tag{A.8}$$

Therefore we have

$$\lim_{n\to\infty}\delta_n=0_-\,. \tag{A.9}$$

References

[1] A. Dvoretzky and P. Erdös, Proc. 2nd Berkeley Symp. Math. Statist. and Prob. (1950) (Univ. of California Press, Berkeley, Calif., 1951) p. 353.
[2] G.H. Vineyard, J. Math. Phys. 4 (1963) 1191.
[3] E.W. Montroll, Proc. Symp. App. Math. Am. Math. Soc. 16 (1964) 193.
[4] E.W. Montroll and G.H. Weiss, J. Math. Phys. 6 (1965) 167.
[5] Y. Nagai, T. Tsuchiya and H. Asai, in preparation.
[6] Y. Nagai, H. Asai and T. Tsuchiya, to be published.
[7] A. Ikegami and N. Kono, J. Mol. Biol. 29 (1967) 251; M. Nakanishi, M. Tsuboi, A. Ikegami and M. Kanehisa, J. Mol. Biol. 64 (1972) 363.
[8] W. Feller, An introduction to probability theory and its applications, Vol. I (2nd ed., John Wiley & Sons, 1957) p. 318.
[9] H.S. Hall and S.R. Knight, Higher algebra (4th ed., Macmillan, 1906) pp. 333; Solutions of the examples in higher algebra (Macmillan, 1926) p. 161.
[10] R.J. Rubin, J. Chem. Phys. 56 (1972) 5747.
[11] K. Moro and M. Saitô, Polymer J. 8 (1976) 377.