Self-avoiding walk model for proteins: a real space renormalisation group treatment

1986 J. Phys. A: Math. Gen. 19 L357

(http://iopscience.iop.org/0305-4470/19/6/011)

J. Phys. A: Math. Gen. 19 (1986) L357-L364. Printed in Great Britain

LETTER TO THE EDITOR

Self-avoiding walk model for proteins: a real space renormalisation group treatment

A Christou and R B Stinchcombe

Department of Theoretical Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, UK

Received 9 January 1986

Abstract. Diffusion on the model for proteins introduced by Helman, Coniglio and Tsallis (HCT) is discussed using real space renormalisation group ideas. We perform cell-to-cell transformations in two dimensions for both the HCT model and the self-avoiding walk. The latter system is quasi-linear and so we expect trivial diffusive behaviour given by $d_w = 2d_f = \frac{8}{3}$. Diffusion on the HCT model however is complicated by the possibility of jumps across nearest-neighbour (hydrogen bond) 'bridges': our results are not consistent with the prediction that $d_w = 2$ for the HCT model; rather they support the assertion that d_w^{HCT} is close to but slightly greater than d_w^{SAW}.

In this letter we investigate diffusion on a model for proteins introduced by Helman *et al* (1984). We use a single cell real space renormalisation group method to determine the asymptotic time-dependent behaviour of a de Gennes (1976) myopic ant on the non-trivial HCT model.

Interest in the problem arises in the context of spin-lattice relaxation rates of the Fe^{3+} ions in low spin haemoproteins and ferredoxin. The dominating two-phonon process leads to a rate $1/T_1$ related to the temperature T by the following scaling form:

$$1/T_1 \propto T^{3+2d_s} f(T/\theta, d_s) \tag{1}$$

with θ the Debye temperature and d_s the spectral dimension (Alexander and Orbach 1982) for the proteins.

The experimental work of Stapleton *et al* (1980) and Allen *et al* (1982) finds the following low temperature behaviour for $1/T_1$:

$$1/T_1 \propto T^n. \tag{2}$$

This observation is consistent with (1) since in an appropriate low temperature regime the scaling function f takes a constant value and so the T dependence of $1/T_1$ is given simply by $1/T_1 \propto T^{3+2d_s}$.

These authors find $n \approx 6.3$ for haemoproteins and $n \approx 5.67$ for ferredoxin. In attempting to reconcile (2) with (1) Stapleton *et al* equated d_s in (1) with the fractal, or *Hausdorff*, dimension (Mandelbrot 1982) of the protein, d_f. This relates the radius of gyration of the protein $\langle R_f \rangle$ to the number of alpha-carbons by the power law

$$\langle R_f \rangle \sim N^{1/d_f}. \tag{3}$$

The numerical value for d_f for the proteins is accessible from the x-ray scattering data. From this, Stapleton *et al* obtained $d_f = \frac{4}{3}$ for ferredoxin and $d_f = \frac{5}{3}$ for the haemoproteins.

0305-4470/86/060357+08$02.50

These results are consistent with the relaxation rate data for n and also with the Flory theory result for the excluded volume problem (self-avoiding walk) $d_f = (d+2)/3$ for $d = 2, 3$ respectively.

Now, information on the dynamic excitations on general fractal spaces may be obtained by considering diffusion on the network. The diffusing particle is represented by a 'myopic ant'. Such a random walker successfully takes a step at every time interval, as opposed to a blind ant which, if it discovers its chosen direction to be blocked, will remain stationary until the next time interval when it tries again.

When considering diffusion on fractals it is important to distinguish between two limiting regimes. If the RMS displacement of the ant is $\langle R_w \rangle$ then in the regime $1 \ll \langle R_w \rangle \ll \langle R_f \rangle$ diffusion phenomena are self-similar and so are amenable to treatment using RSRG. In this regime $\langle R_w \rangle$ is given asymptotically by

$$\langle R_w \rangle \sim t^{1/d_w} \tag{4}$$

for fixed though large N, where t is the time or, equivalently, the number of steps of the ant. Expression (4) defines the *dimensionality of the walk* d_w. On fractal space d_w is not 2 in general but takes on some anomalous system-dependent value. In the regime where the number of steps t of the ant is large compared to the number of bonds of the system N, $\langle R_w \rangle$ will saturate to

$$\langle R_w \rangle \sim N^{1/d_f}. \tag{5}$$

These regimes are bridged by a crossover region in which $\langle R_w \rangle$ is given by

$$\langle R_w \rangle \sim N^{1/d_f} g(N^\theta / t) \tag{6}$$

where g is a universal scaling function and θ is the crossover exponent.

We confine our attention to the self-similar limit $1 \ll \langle R_w \rangle \ll \langle R_f \rangle$.

According to the Alexander-Orbach theory, knowledge of d_w yields the scaling behaviour of the low frequency density of states $\rho(\omega) \propto \omega^{d_s - 1}$ through the relation

$$d_s = 2d_f / d_w. \tag{7}$$

Consequently a simple SAW model for the proteins leads to a contradiction. A straightforward scaling argument (see footnote in Helman *et al* (1984)) shows that for the quasi-linear SAW $d_w = 2d_f$ and so from (7) one obtains $d_s = 1$ in all dimensions. This result is inconsistent with the assumption of Stapleton *et al* (1980), $d_s = d_f$. In order to overcome this difficulty Helman *et al* introduced conducting but massless nearest-neighbour crosslinks into the SAW model. They predicted that, provided the density of crosslinks or bridges was sufficiently high, the ant would 'see' the embedding Euclidean lattice and diffuse as if on Euclidean space. Their objective was therefore to create a protein model with SAW fractal dimension but with $d_w = 2$, thereby restoring consistency with the experimental results. The 'massless bridges' were to be interpreted as hydrogen bonds.

Subsequent numerical work on the HCT model in two dimensions (Yang *et al* 1985, Chowdhury and Chakrabarti 1985) however does not support the suggestions of Helman *et al.* Yang *et al* obtain $d_w = 2.6$ and indicate that they would expect $d_w = 2d_f = \frac{8}{3}$, as for the SAW, with more accurate calculations. Chowdhury and Chakrabarti however obtain an even higher result $d_w = 2.78$. Furthermore the latter authors argue that $d_w > \frac{8}{3}$ strictly, this being a direct consequence of the ant's steps over the bridges which, in terms of chemical distance travelled measured along the chain, constitute Lévy flights.

We attempt to clarify the situation using analytic renormalisation techniques. We consider the stochastic master equation governing the motion of the ant. Assigning a probability $P_n(t)$ that the ant is at site n of the HCT model at time t subject to $P_n(0) = \delta_{n,0}$ we obtain the time evolution from

$$\frac{dP_n(t)}{dt} = -VP_n(t) + \sum_{n+\delta} WP_m(t) \tag{8}$$

where W is the transition rate for the ant to step from a site onto a *particular* nearest neighbour. The parameter V is simply the step rate given by WZ_n with Z_n the local coordination number. The summation is over Z_n occupied nearest neighbours.

During the renormalisation process, the embedding square lattice is mapped onto one isomorphic to it together with an accompanied dilatation of space by a factor b and elimination of short-range SAW degrees of freedom. The motion of the ant on the renormalised system is governed by an equation analogous to (8) but with a renormalised hopping rate V'. The asymptotic, scale-invariant, phenomena are observed at a critical rate $0 < V = V' = V^* < \infty$. Clearly we expect $V^* = 1$ since our ant is myopic, not blind.

In deriving the explicit real space renormalisation group recursion relation for V, the step 'fugacity' (Stanley *et al* 1982) for the random walker, it becomes evident that one has to introduce a second fugacity K, associated with each step of the SAW. The parameter K, together with its RSRG transformation $K' = R_b(K)$, characterises the underlying fractal space. Similar approaches have been used in studying random walks on percolation clusters (Sahimi and Jerauld 1984), lattice animals (Family 1983) and Witten-Sander aggregates (Christou and Stinchcombe 1986).

Firstly we derive the transformation for K following de Queiroz and Chaves (1980) and Redner and Reynolds (1981). One generates the set of SAW starting in the lower left corner of a $b \times b$ cell and finishing on one of the b sites at the top of the cell. Each traversing SAW configuration of n steps is weighted by a factor K^n. A grand canonical partition function $Z^b_{\mathrm{SAW}}(K)$ is defined for the cell thus:

$$Z^b_{\mathrm{SAW}}(K) = \sum_n C_b(n) K^n \tag{9}$$

where $C_b(n)$ is the number of spanning SAW of n steps on a $b \times b$ cell. We set $R_b(K) = Z^b_{\mathrm{SAW}}(K)$. In general, the $b \times b$ cell is mapped onto a $b' \times b'$ cell via a 'cell-to-cell' transformation (Reynolds *et al* 1978), the renormalised partition function being given by

$$Z^{b'}_{\mathrm{SAW}}(K') = \sum_m C_{b'}(m)(K')^m. \tag{10}$$

In particular, for a cell-to-bond transformation, $Z^{b'=1}_{\mathrm{SAW}}(K') = K'$. If we stipulate that the partition function be conserved under renormalisation we have an implicit transformation $K \to K'$. This has a critical fixed point at $0 < K^* < \infty$ where one evaluates the eigenvalue of the linearised transformation, $\lambda_k = (\partial K'/\partial K)_{K^*}$, from which follows the fractal dimension of the SAW, d_f, in the usual way.

The procedure is analogous to that outlined above when one introduces random walks on the HCT models. Here, we define a new partition function $Z^b_{rw}(K, V)$ given for a $b \times b$ cell by

$$Z^b_{rw}(K, V) = \sum_{n,t} C_b(n, t) K^n V^t [(1/z_0)(1/z_1) \ldots (1/z_{t-1})] \tag{11}$$

with $C_b(n, t)$ being the number of random walks of t steps spanning each HCT model configuration of n steps and z_i being the coordination number of the site visited by the ant after i steps. The factor $[(1/z_0)(1/z_1)\ldots(1/z_{t-1})]$ arises from the fact that the weight associated with a particular ant step from a site a is V/z_a, where z_a is the local coordination number of a. Upon renormalisation the transformed partition function is

$$Z^{b'}_{rw}(K', V') = \sum_{m,s} C'_b(m, s)(K')^m(V')^s[(1/z_0)(1/z_1)\ldots(1/z_{s-1})]. \tag{12}$$

For $b' = 1$, we set $Z^{b'=1}_{RW}(K', V') = K'V'$.

Evidently, for a SAW without crosslinks, this 'kinetic' view (Nakanishi and Family 1984) of the random walk may be replaced by the more familiar 'static' interpretation $V/2 \rightarrow W$ as the local coordination number is constant.

A problem arises associated with the truncation of the infinite series (11) and (12). We note that at the critical fixed point (K^*, V^*) only random walks of length $t \leqslant \xi^{d_w}$ are important (see, e.g., Sahimi and Jerauld 1984), ξ being the end-to-end distance. However, given that all we know of d_w is

$$d_f \leqslant d_w \leqslant 2d_f \tag{13}$$

our choice of truncation point will necessarily be rather *ad hoc*. In our computations we first treat only walks satisfying $t \leqslant \xi^2$, this being the mid-point of the interval (13) since we know that $d_f = \frac{4}{3}$. We then repeat the process considering longer walks $t \leqslant \xi^{8/3}$. The results for d_w for the HCT model are affected significantly by the choice of truncation point. However, if we repeat the renormalisation for random walks on SAW (with no cross-linking allowed) we observe that the value d_w^{HCT}/d_w^{SAW} is much less affected by the truncation rule used.

We derive the recursion relations for the simplest of transformations $b = 2$, $b' = 1$ for the HCT model. Here, from de Queiroz and Chaves (1980),

$$K' = K^2 + 2K^3 + K^4 \tag{14}$$

and enumerating all random walks such that $t \leqslant \xi^{8/3}$,

$$\begin{aligned} K'V' = {} & K^2(\tfrac{1}{2}V^2 + \tfrac{1}{2}V^4 + \tfrac{1}{2}V^6) + K^3(\tfrac{1}{4}V^3 + \tfrac{5}{16}V^5 + \tfrac{21}{64}V^7) \\ & + K^3(\tfrac{1}{4}V^3 + \tfrac{5}{16}V^5 + \tfrac{21}{64}V^7) + K^4(\tfrac{1}{6}V^2 + \tfrac{7}{36}V^4 + \tfrac{43}{216}V^6). \end{aligned} \tag{15}$$

We define $\lambda_k = (\partial K'/\partial K)_{K^*}$ and $\lambda_v = (\partial V'/\partial V)_{V^*,K^*}$. One obtains

$$d_f = \log \lambda_k / \log b$$

and

$$d_w = \log \lambda_v / \log b. \tag{16}$$

Hence for the $b = 2$ cell-to-bond transformation (with $t \leqslant \xi^{8/3}$) for the HCT model one obtains $(K^*, V^*) = (0.466, 0.961)$ and $d_f = 1.398$ and $d_w = 2.100$. A $b = 2$ cell-to-bond transformation has also been very recently applied to the HCT model by Chowdhury (1985) who, though only considering walks up to $t = \xi^2$ and not properly allowing for the V/z_a factors required by the governing diffusion equation, obtained very similar results in this small cell calculation.

The corresponding recursion relations for random walks on the SAW (in the static interpretation) are

$$K' = K^2 + 2K^3 + K^4$$

$$K'W' = K^2(W^2 + 2W^4 + 4W^6) + K^3(W^3 + 3W^5 + 8W^7) + K^3(W^3 + 3W^5 + 8W^7) + K^4(W^4 + 4W^6). \tag{17}$$

This system of equations yields $(K^*, W^*) = (0.466, 0.577)$ and $d_f = 1.398$ and $d_w = 2.028$.

Similarly, for a cell-to-cell transformation $1 < b' < b$, one obtains for K, say, an implicit transformation (Reynolds *et al* 1978)

$$K'(b) = R_b[R_{b'}^{-1}(K'(b'))] \equiv R_{b/b'}(K'(b')) \tag{18}$$

given in terms of the cell-to-bond transformation function $K'(b) = R_b(K)$. The value of K at which $K'(b) = K'(b')$ or equivalently $R_b(K) = R_{b'}(K)$ corresponds to K^*. It then follows that the fractal dimension, d_f, is obtained from the relation

$$d_f = \frac{\log(\lambda_k(b)/\lambda_k(b'))}{\log(b/b')}. \tag{19}$$

Here $\lambda_k(b)$ is the K eigenvalue of the $b \times b$ cell-to-bond transformation. However, it is evaluated at the fixed point (K^*, V^*) of the cell-to-cell transformation. Similarly for V'

$$V'(b) = G_b(G_{b'}^{-1}(V'(b')) \equiv G_{b/b'}(V'(b')) \tag{20}$$

where G_b is the V cell-to-bond transformation function given by $V' = G_b(K; V) = Z^b_{\mathrm{SAW}}(K, V)/K'$, and the inverse is taken treating K only as a parameter.

From (20) it follows that

$$d_w = \frac{\log(\lambda_v(b)/\lambda_v(b'))}{\log(b/b')} \tag{21}$$

where the eigenvalues λ_v are again evaluated at the cell-to-cell fixed point.

A now familiar problem with small cell RSRG is associated with cell interfacing. This phenomenon leads to errors both in the fixed points and in the eigenvalues. It is reasonable to expect, however, that such interfacing problems will vanish as $b \to \infty$ for a cell-to-bond transformation and $b \to \infty$ together with $b/b' \to 1$ for a cell-to-cell transformation (Reynolds *et al* 1978, Redner and Reynolds 1981). In principle therefore, by evaluating that RSRG recursion equations for larger b and also for $b/b' \to 1$ one will obtain proved estimates for the exponents. Assuming that the resulting sequences of exponents are monotonic in b it would then be possible to obtain reliable extrapolations to $b \to \infty$ (Reynolds *et al* 1980).

We therefore proceed to evaluate the larger cell transformations and the results up to $b = 6$ are summarised in tables 1 and 2. The fixed point K^* and exponent d_f are given by Redner and Reynolds. For the HCT model one must generate all spanning SAW configurations and for each one enumerate exactly all random walks, allowing for steps over nearest-neighbour bridges, with $t \leq \xi^2$ and then repeat for $t \leq \xi^{8/3}$. For a $b = 4$ cell, for example, with $t \leq \xi^{8/3}$ one must enumerate random walks of up to $t = 73$ steps for each of the 649 spanning HCT SAW. The corresponding calculation for a $b = 5$ cell would involve counting and weighting the spanning random walks with

Table 1. Results of the cell-to-bond and cell-to-cell RSRG for random walks on the HCT model for proteins. The upper number in each pair of figures corresponds to having enumerated exactly all random walks on the HCT model of length t such that $t \leqslant \xi^2$; the lower ones are for $t \leqslant \xi^{8/3}$.

		b'		
b		1	2	3
2	V^*	1.128		
		0.961		
	d_w^{HCT}	1.815		
		2.100		
3	V^*	1.051	1.009	
		0.947	0.941	
	d_w^{HCT}	1.879	1.998	
		2.237	2.458	
4	V^*	1.090	1.082	1.135
		1.013	1.021	1.060
	d_w^{HCT}	1.927	2.042	2.128
		2.368	2.680	3.153

Table 2. Results of cell-to-bond and cell-to-cell RSRG for random walks on SAW. The upper number in each pair of figures corresponds to having enumerated exactly all random walks on the SAW with number of steps t such that $t \leqslant \xi^2$; the lower ones are for $t \leqslant \xi^{8/3}$.

		b'				
b		1	2	3	4	5
2	W^*	0.650				
		0.572				
	d_w^{SAW}	1.784				
		2.038				
3	W^*	0.581	0.538			
		0.531	0.512			
	d_w^{SAW}	1.748	1.692			
		2.117	2.136			
4	W^*	0.553	0.531	0.526		
		0.520	0.512	0.511		
	d_w^{SAW}	1.848	1.910	2.205		
		2.247	2.361	2.663		
5	W^*	0.539	0.526	0.523	0.521	
		0.514	0.510	0.509	0.508	
	d_w^{SAW}	1.916	2.022	2.277	2.379	
		2.357	2.521	2.813	3.001	
6	W^*	0.530	0.522	0.519	0.518	0.515
		0.510	0.508	0.507	0.506	0.505
	d_w^{SAW}	1.960	2.086	2.313	2.401	2.442
		2.438	2.625	2.901	3.077	3.169

number of steps t up to $t = 141$ for 36 368 spanning configurations. This required too much CPU time for our algorithm: a 180 min run on a VAX 11/780 was insufficient. However, knowing the transfer matrix elements $T_i(K)$ (Redner and Reynolds 1981), where T_i is the weight associated with SAW spanning a $b \times b$ cell from the origin to the ith ($i = 1, \ldots, b$) site on the uppermost layer of the cell, one is able to carry out the $b = 5$ and $b = 6$ transformations for SAW without crosslinks for which diffusion is confined to the quasi-linear chains, i.e. not dependent on SAW conformation. This serves as a good check for the reliability of the lower b exponents for the SAW and hence for those of the HCT model.

The results for d_w^{HCT} and d_w^{SAW} obtained from the small cells considered are not sufficiently smooth to allow reliable extrapolations. It is possible that these non-monotonicities are the effect of pathologies of the RSRG 'interfacing', problems which might not vanish in the limit $b \to \infty$. The probability that a spanning SAW terminates on a particular cell site is $1/b$ (in $d = 2$). Thus the 'corner rule' used would indicate that interfacing problems indeed worsen as b increases. However, if one notes that the corner rule may be interpreted (Redner and Reynolds) as representing the problem of a SAW in a wedge (apex angle $\pi/4$) with one end of the SAW fixed at the apex then we are left with an example of the problem of SAW in confined geometries. Scaling approaches (see de Gennes 1979) indicate that the confining geometry becomes irrelevant, as $b \to \infty$, to the critical behaviour. We therefore expect the accuracy of the exponents d_f and d_w to improve as larger cells are considered. This may be done using Monte Carlo renormalisation. As only relatively few spanning SAW configurations are sampled, longer random walks may then also be enumerated. We also expect the dependence of the exponent d_w on random walk length to vanish as t increases.

In table 3 we show $d_w^{HCT}(b/b')/d_w^{SAW}(b/b')$ for various cell-to-cell transformations. The predictions of Helman *et al* (1984) would require $d_w^{HCT}/d_w^{SAW} = 0.75$. Our results are not consistent with this prediction. They are, however, in qualitative agreement with the Monte Carlo work of Yang *et al* (1985) who obtained $d_w^{HCT}/d_w^{SAW} = 0.97$ and that of Chowdhury and Chakrabarti (1985) who obtained $d_w^{HCT}/d_w^{SAW} = 1.04$. Despite the observed non-monotonic behaviour of the exponents, however, our results would seem to favour the argument put forward by Chowdhury and Chakrabarti for $d_w^{HCT} > d_w^{SAW}$.

In conclusion therefore we have developed and implemented a two-parameter small cell RSRG for random walks on SAW, for which diffusion is trivial, and on the model

Table 3. The results for $d_w^{HCT}(b/b')/d_w^{SAW}(b/b')$ for various cell-to-cell transformations. Monte Carlo work of Chowdhury and Chakrabarti indicates $d_w^{HCT}/d_w^{SAW} \approx 1.04$. The upper number in each pair of figures corresponds to evaluating all random walks with $t \leq \xi^2$; the lower ones are for $t \leq \xi^{8/3}$.

	b'		
b	1	2	3
2	1.0173		
	1.0306		
3	1.0747	1.1806	
	1.0563	1.1509	
4	1.0428	1.0688	0.9654
	1.0538	1.1348	1.1840

of Helman *et al* for proteins. Our results are not consistent with the requirement that $d_w^{HCT}/d_w^{SAW} = 0.75$ as speculated by Helman *et al.* However they do seem broadly in agreement with the available numerical data which suggest that $d_w^{HCT}/d_w^{SAW} \approx 1$. On the other hand, they would appear to support the conclusions of Chowdhury and Chakrabarti that $d_w^{HCT} > d_w^{SAW}$. In order to establish whether the HCT models are in a distinct dynamic universality class to SAW one would have to introduce a new step fugacity X for jumps across the bridges and determine whether the variable is relevant in the RG sense. Strictly, one would expect d_w^{SAW} to be affected by the presence of crosslinks only if these links were distributed along the SAW in some hierarchical fashion. This is in fact what seems to happen as loops created by crosslinks have smaller loops within them and so on all length scales leading to the scale invariance that we have exploited in our renormalisation scheme above.

Finally, we point out that saturation effects for the non-scale-invariant limit where the number of random walk steps t is much greater than SAW length N are not observable within this RG framework since it is implicit that the SAW or HCT model remains invariant under space dilatations.

We thank Dr P M Duxbury for helpful discussions. The support of the SERC, in the form of a Research Studentship, is gratefully acknowledged by AC.

References

Alexander S and Orbach R 1982 *J. Phys. Lett.* **43** L625
Allen J P, Colvin J T, Stinson D G, Flynn C P and Stapleton H J 1982 *Biophys. J.* **38** 299
Chowdhury D 1985 *J. Phys. A: Math. Gen.* submitted
Chowdhury D and Chakrabarti B K 1985 *J. Phys. A: Math. Gen.* **18** L377
Christou A and Stinchcombe R B 1986 *J. Phys. A: Math. Gen.* to be published
de Queiroz S L A and Chaves C M 1980 *Z. Phys.* B **40** 99
Family F 1983 *J. Phys. A: Math. Gen.* **16** L97
de Gennes P G 1976 *La Recherche* 7 919
—— 1979 *Scaling Concepts in Polymer Physics* (Ithaca, NY: Cornell University Press)
Helman J S, Coniglio A and Tsallis C 1984 *Phys. Rev. Lett.* **53** 1195
Mandelbrot B B 1982 *The Fractal Geometry of Nature* (San Francisco: Freeman)
Nakanishi H and Family F 1984 *J. Phys. A: Math. Gen.* **17** 427
Redner S and Reynolds P J 1981 *J. Phys. A: Math. Gen.* **14** 2679
Reynolds P J, Stanley H E and Klein W 1978 *J. Phys. A: Math. Gen.* **11** L199
—— 1980 *Phys. Rev.* B **21** 1223
Sahimi M and Jerauld G R 1984 *J. Phys. A: Math. Gen.* **17** L165
Stanley H E, Reynolds P J, Redner S and Family F 1982 *Real-Space Renormalisation* ed T Burkhardt and J M J van Leeuwen (New York: Springer) p169
Stapleton H J, Allen J P, Flynn C P, Stinson D G and Kurtz S R 1980 *Phys. Rev. Lett.* **45** 1456
Yang Y S, Liu Y and Lam P M 1985 *Z. Phys.* B **59** 445