Biophysical Chemistry 4 (1976) 367–381

ASSOCIATION KINETICS WITH COUPLED DIFFUSIONAL FLOWS. SPECIAL APPLICATION TO THE LAC REPRESSOR–OPERATOR SYSTEM

Otto G. BERG and Clas BLOMBERG
Department of Theoretical Physics, Royal Institute of Technology, S-100 44 Stockholm 70, Sweden

Received 16 January 1976
Revised manuscript received 15 April 1976

The time development of the association of the lac repressor to the operator is considered in a model where the repressor is allowed to bind unspecifically to DNA and move along the DNA chain in a one-dimensional diffusion. The coupling to the three-dimensional diffusion outside the chain is introduced by letting the repressor associate and dissociate from the chain until it is finally bound to the operator. All distance correlations along the chain are included. The mean time of association is calculated and through a comparison with experimental data the molecular parameters are determined. The one-dimensional diffusion constant is found to be of the order of 10^{-9} cm^2 s^{-1}. The model is sufficiently general to be applicable to other similar systems.

1. Introduction

The suggestion that a diffusional search in a space can be considerably speeded up by confining the search to a subspace of lower dimensionality has recently attracted much attention [1,2]. Biological examples would be a substrate finding a membrane-bound enzyme through a two-dimensional diffusion along the membrane surface, or a regulatory protein finding its specific site on DNA through a one-dimensional diffusion along the DNA chain. In this paper we will consider the latter example, and apply it specifically to the much studied lac repressor–operator system. The experiments by Riggs and co-workers [3] have shown that the lac repressor finds its specific site, the operator, on DNA with a rate that is considerably higher than what would be expected from a conventional diffusion controlled process. Richter and Eigen [4] have analyzed two possible mechanisms for this enhancement; electrostatic attraction between the repressor and the operator, and the possibility that the repressor could be bound (for some time at least) unspecifically to DNA and then move along DNA. They showed that the electrostatic interaction cannot by itself explain the high association rate but that this is done by including the one-dimensional diffusion. In principle this is achieved by increasing the range of the operator from its actual length to an effective range which is the average distance a repressor molecule can move along DNA before dissociating. The idea being that a repressor coming within the effective range of the operator will eventually be captured. These calculations account well for the high association rate, but have not attempted to describe the time development of the association. In particular two effects were not regarded, namely the timelag while the repressor moves along the effective range of the operator to the specific site, and the fact that DNA, at distances from the operator larger than the effective range, will serve as a trap for the repressor and thereby reducing the association rate.

In the present work we shall take these effects into account by considering the whole length of the DNA chain and the correlated one- and three-dimensional diffusion of the repressor molecule along and outside this chain. In addition to giving the complete time development of the association process, this model also correlates the association rate for the repressor–operator complex with the equilibrium constant for repressor binding to non-operator DNA and is thereby able to provide definite values to molecular parameters like the one-dimensional diffusion constant.

This way of facilitating a difficult association by a restricted diffusion may be widely used in biochemical processes in which cases our developed model can be applied, possibly with some modification. The possibility for large molecules to find the right positions to fit together in the best way is normally small unless mechanisms of the proposed kind are used. Among such cases can be mentioned: other regulatory proteins with a general affinity for DNA as well as for their specific DNA sites, the association of tRNA at the ribosome–mRNA complex and the binding of enzymes to certain places of large molecule complexes.

In this way, we shall begin by discussing some aspects of diffusion controlled processes in section 2 in order to be able to formulate the general problem in an appropriate way in section 3. There we also show how to accomplish a general solution for the time development of the association. In section 4 we derive and discuss the important time quantities and the equilibrium constant for the unspecific binding. Then, a comparison with the experimental data in vitro is made in section 5 and the main parameters estimated. Finally, the important features are summed up and discussed in section 6.

2. Diffusion controlled processes in general; the "closed cell" approach

Consider a reaction vessel containing particles and sinks, i.e., the places where the particles can be bound or absorbed. Let the initial distribution of particles and sinks be homogeneous with n_0 sinks and c_0 particles per cm^3. Now, the main question is to describe the time development of the absorption process. Von Smoluchowski [5] solved this problem long ago for spherically shaped particles and sinks with the aid of a stationary solution to the diffusion equation. If there is initially an equal number of particles and sinks, $n_0 = c_0$, the probability $g(t)$ that a sink is not filled before time t is

$$g_1(t) = 1/(n_0 k_a t + 1) . \tag{2.1}$$

Here k_a, the association rate, is given by

$$k_a = 4\pi D b , \tag{2.2}$$

where D equals the sum of the diffusion constants and b, the reaction radius, equals the sum of the radii of the sinks and particles. A sink can absorb only one particle and every encounter between an empty sink and a particle is assumed to lead to absorption.

If there are N particles per sink, i.e., $c_0 = N n_0$ with $N > 1$, eq. (2.1) is changed into

$$g_N(t) = \{(1-1/N) \exp[-(N-1) n_0 k_a t]\}/\{1-(1/N) \exp[-(N-1) n_0 k_a t]\} . \tag{2.3}$$

$N < 1$ can be considered similarly.

As the probability for absorption between times t and $t + dt$ is $-g'(t)\,dt$, the mean time τ for absorption is given by

$$\tau = -\int_0^\infty t g'(t)\,dt = \int_0^\infty g(t)\,dt . \tag{2.4}$$

From eq. (2.1) this yields an infinite mean time and in this case a better measure is the half time $t_{1/2}$ defined such that $g_1(t_{1/2}) = 1/2$:

$$t_{1/2} = 1/n_0 k_a . \tag{2.5}$$

From eq. (2.3) the mean time is

$$\tau_N = -\frac{1}{n_0 k_a} \ln\left(1 - \frac{1}{N}\right). \tag{2.6}$$

To treat the absorption onto an elongated sink like a DNA chain, Richter and Eigen [4] extended

von Smoluchowski's scheme to sinks of spheroidal symmetry. This is hardly a feasible method for our purpose, as we want to consider the diffusion absorption onto the elongated sink coupled with a diffusion along it. To this end cylindrical coordinates are better adapted. As there is no stationary solution to the diffusion equation in cylindrical coordinates of the type used for spherical or spheroidal symmetry, the von Smoluchowski method is not directly applicable and we have to use a slightly different approach. In order to see the connections we first consider in some detail the spherically symmetrical problem.

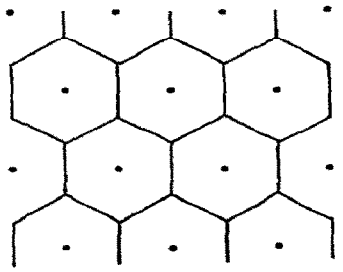

Fig. 1. The regular spacing of the sinks (indicated by dots) and the construction of closed cells around them, as it would look in two dimensions and hexagonal symmetry.

Assume that the sinks are distributed regularly such that the total volume can be filled by equal imaginary cells constructed around each sink (fig. 1). Every cell then has the volume n_0^{-1} cm^3. If the number of particles equals the number of sinks, and the initial distribution is homogeneous, there will be, on the average, one particle in each cell. Assume further that the absorption process in each cell is independent of the others. Then, a particle crossing one of the cell walls enters the neighbour cell at a point that is equivalent to the one it left in the first cell. Thus, the particle could be considered as simply being reflected at the boundary and staying in one cell, and it suffices to consider the process in only one cell subject to the boundary condition that the flux of particles across the cell walls equals zero. Also because of the initial homogeneity the average number of particles going in one direction across the walls would be equal to the average number going in the other, and the fluxes would cancel. For simplicity, the cell can be taken to be spheres of radius R such that $4\pi R^3/3 = n_0^{-1}$. This is not crucial since it is the size of the volume and not its shape that is of primary importance. Let $c(r, t)$ be the concentration of particles at the distance r from the center of the cell at time t. Then we have to solve the spherically symmetric diffusion equation

$$\frac{1}{D}\frac{\partial c}{\partial t} - \frac{\partial^2 c}{\partial r^2} - \frac{2}{r}\frac{\partial c}{\partial r} = 0 \tag{2.7a}$$

in the region $b < r < R$. At $r = b$ the particle is in contact with the sink, and assuming that such a contact leads to immediate absorption we have the traditional boundary conditions for an ideal sink

$$c(r = b, t) = 0 . \tag{2.7b}$$

The condition of zero particle flux across the cell walls is equivalent to the reflecting boundary condition

$$\left.\frac{\partial c}{\partial r}\right|_{r=R} = 0 . \tag{2.7c}$$

The initial condition is homogeneous

$$c(r, t = 0) = n_0 = 3/4\pi R^3 . \tag{2.7d}$$

The solution to the eqs. (2.7) is found in standard textbooks [6]:

$$c(r,t) = \sum_{n=1}^{\infty} A_n \frac{\sin[\lambda_n(r-b)]}{r} \exp(-\lambda_n^2 Dt) ,$$

which for $b \ll R$ is totally dominated by the lowest λ-value

$$\lambda_1 \approx \sqrt{3b/R^3} . \tag{2.8}$$

This gives the probability that a sink is still unfilled at time t

$$g_1(t) = \exp(-n_0 k_a t) , \tag{2.9}$$

with k_a as in eq. (2.2). The result is in agreement with eq. (2.1) only for times shorter than $1/n_0 k_a$. The discrepancy for longer times is due to the assumption that the cells are independent; when some sinks are filled and others are not, no such symmetry between the cells exists. It is important to note, however, that (2.9) is exactly what one gets with the von Smoluchowski scheme for the case where the sink can absorb an indefinite number of particles, i.e., the sinks do not fill up. [$g(t)$ then being the probability that a particle is not absorbed before the time t.] In this case there would be total symmetry between the cells at all times and the independence of the cells, i.e., the requirement of zero particle flux between them is a necessary consequence of this symmetry and the symmetry of the diffusion equation and its initial condition. The only questionable assumption in this case is the regular spacing of the sinks; seeing that this leads to exactly the same results as the traditional approach we conclude that it is satisfactory. This is of course what one would expect intuitively as it, in a sense, represents the average behaviour even for a random distribution.

If there are N particles per sink this "closed-cell" approach gives

$$g_N(t) = [g_1(t)]^N = \exp(-Nn_0 k_a t) , \tag{2.10}$$

which for large N is in good agreement with eq. (2.3) also asymptotically. The mean time is

$$\tau_N = \frac{1}{n_0 k_a} \frac{1}{N} , \qquad N \geq 1 , \tag{2.11}$$

in good agreement with (2.6) also for relatively small N. Especially we note that for $N=1$ there is a complete agreement between the mean time (2.11) and the half time (2.5); a similarity that we will invoke further on.

Summing up, the "closed cell" approach gives a good behaviour for times shorter than the mean time but it gives a good asymptotic behaviour only when there are several particles per sink in the system, or when the sinks are such that they do not fill up. This latter case, which describes a precipitation process or a quenching of fluorescent molecules, would in some cases be better covered by this approach than by the traditional one. We are therefore going to use this description as a starting point for the main problem.

3. The model

Consider an initially homogeneous distribution of particles (repressors) and sinks (operators). Every sink is imbedded in the middle of a long cylindrical chain (a DNA molecule), of length $2L$ and radius b.

A particle moves with a diffusion constant D in the medium and every time it encounters a chain it will be absorbed onto it, and move along the chain with a one-dimensional diffusion constant $\mathcal{D}_1$. With the probability $\lambda\,dt$ it will leave the chain during the short time interval dt, unless it has encountered the sink in which case it would be permanently absorbed. What is the time development of the absorption?

The diffusional flow of particles onto the chain, along, off and again onto the chain etc., ..., and finally into the sink is a complicated coupled three- and one-dimensional diffusion. There is, however, a simple way of separating the movements. Let

$u(z,t)\,dz =$ the probability that the particle is in the interval $(z, z+dz)$ on the chain at the time t,

$G(z,t)\,dz\,dt =$ the probability that a particle for the first time arrives at the chain in the interval $(z, z+dz)$ in the time interval $(t, t+dt)$,

$F(z,z',t-t')\,dz\,dt =$ the probability that a particle leaving the chain from z' at time t' again returns to the chain in the interval $(z, z+dz)$ and in the time interval $(t, t+dt)$.

For symmetry reasons we will only consider one half, $0 \leqslant z \leqslant L$, of the chain. The diffusion equation along the chain becomes

$$\frac{\partial u}{\partial t} = D_1 \frac{\partial^2 u}{\partial z^2} - \lambda u + \lambda \int_0^L dz' \int_0^t dt' F(z,z',t-t') u(z',t') + G(z,t) . \tag{3.1}$$

The first term on the right hand side is the change in the distribution due to the diffusional displacements along the chain. The second term comes from the probability that the particle dissociates from the chain. All the correlations are in the third term which describes the return to the chain of a particle that left it at some earlier time; the integrations make sure that all such returning particles are counted. The last term is the arrival of an uncorrelated particle, i.e., one that has not been associated with the chain previously.

The boundary condition at $z=0$ for an ideal sink

$$u(0,t) = 0 . \tag{3.2}$$

At the ends of the chain the choice of boundary condition is not crucial and we can use the simplest, reflecting boundaries

$$\left.\frac{\partial u}{\partial z}\right|_{z=\pm L} = 0 . \tag{3.3}$$

If there is initially no particle on the chain, we have the initial condition.

$$u(z,0) = 0 . \tag{3.4}$$

The quantity of greatest interest, the flux of particles into the sink

$$\Phi(t) = D_1 \left.\frac{\partial u}{\partial z}\right|_{z=0} \tag{3.5}$$

is in principle determined by the eqs. (3.1)–(3.5) once the functions F and G are known.

So far the formulation of the problem is quite general and independent of how the DNA chains are organized in space. The functions F and G can be calculated by considering the diffusion of particles outside the chains only, but to do this we need some specific assumptions about the chains. First we invoke the "closed cell" by assuming the chains to be straight, parallel and regularly spaced. Construct a cylinder around each chain, with the chain along its symmetry axis, such that the whole volume is filled by these equal imaginary cylinders. Their cross sections would be regular polygons (cf. fig. 1 for one possibility), but for simplicity we assume them to be circles of radius R. If there are n_0 chains per cm^3, every cylinder would have the length $2L$ (equal to the chain length), the radius R and the volume

$$2\pi L R^2 = n_0^{-1} . \tag{3.6}$$

The assumption about regular spacing is not crucial; as discussed in section 2 this is, in a sense, the average behaviour. More important is the assumption that the chains are straight; this can be justified by the fact that a particle close to the chain will experience it as approximately straight, and those straying far away will lose their correlations to the chain and then the assumption should not be significant.

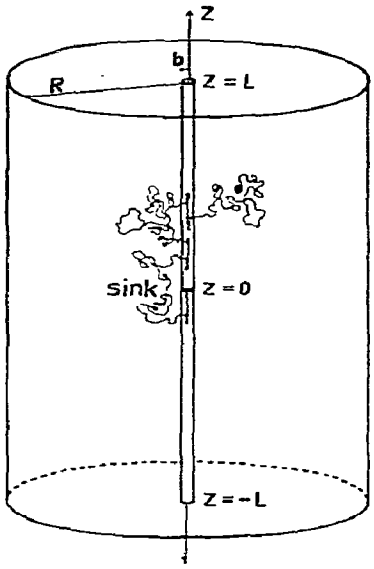

Fig. 2. The whole length, $2L$, of the chain (radius b) placed along the symmetry axis of a circular cylinder (radius R). The sink (operator site) is supposed to be in the middle on the chain. A possible trajectory for the particle before the final absorption is indicated by the irregular line.

As discussed in section 2, the symmetry between the cylinders, the "closed cells", makes it possible to consider the diffusion in one cylinder only. The functions F and G can now be calculated as follows:

Introduce cylindrical coordinates r and z with the z axis along the chain as before (fig. 2). For symmetry reasons the angular coordinate will not enter. Assuming that a particle which dissociates from the chain at $z=z'$ and time $t=0$ is "lifted" out the distance $r=a$ from the chain axis to start anew. When and where will it return to the chain?

The solution is given by the diffusion equation for the concentration $c(r, z, t)$ of particles

$$\frac{\partial^2 c}{\partial r^2}+\frac{1}{r}\frac{\partial c}{\partial r}+\frac{\partial^2 c}{\partial z^2}-\frac{1}{D}\frac{\partial c}{\partial t}=0\,, \tag{3.7}$$

with the boundary conditions

$$\left.\frac{\partial c}{\partial z}\right|_{z=L}=0\,, \tag{3.8}$$

$$\left.\frac{\partial c}{\partial z}\right|_{z=0}=0\,, \tag{3.9}$$

$$\left.kc\right|_{r=b}-\left.\frac{\partial c}{\partial r}\right|_{r=b}=0\,, \tag{3.10}$$

$$\left.\frac{\partial c}{\partial r}\right|_{r=R}=0\,, \tag{3.11}$$

and the initial condition

$$c(r, z, t=0)=\frac{1}{2\pi a}\delta(r-a)\,\delta(z-z')\,. \tag{3.12}$$

The symmetry between the cylinders, "closed cells", is expressed by the reflecting boundaries (3.8) and (3.11). They are consistent with the boundary condition (3.3) on the chain. The boundary condition (3.9) is equivalent to demanding that the total solution is symmetric with respect to $z=0$. This is necessary since we are considering only one half of the chain, cf. eq. (3.1); a particle entering the negative z-domain will be counted as having a positive z-coordinate, i.e., in principle reflected at $z=0$. In (3.10) a parameter k is introduced which will be of the nature of a reaction rate [7]. Actually, $2\pi Dbk$ would be the association rate onto the chains for a homogeneous distribution of particles, i.e., when the diffusion is not the time limiting factor, cf. eq. (3.13). A large value for k corresponds to the diffusion-limited reaction and gives essentially the boundary condition for an ideal sink, $c(r=b)=0$.

The introduction of the dissociation distance a may look somewhat artificial. In principle we could let $a \to b$ and the boundary condition (3.10) will ensure that a particle escaping from the chain is not immediately reassociated. In a real situation, however, the boundary conditions are provided by the detailed features, largely unknown, of the potential around the chain to which (3.10) can only be an approximation. We shall keep $a > b$ since this is more general and could give a better description of the dissociation. The results will not be very sensitive to the actual choice of the parameter a.

The particle flux, $F(z,z',t)$ back to the chain which enters eq. (3.1) is now given by

$$F(z,z',t) = 2\pi bD \left.\frac{\partial c}{\partial r}\right|_{r=b} = 2\pi bDk\, c(r=b)\,, \tag{3.13}$$

and can be calculated after solving eqs. (3.7)–(3.12).

The particle flux, $G(z,t)$, onto the chain for a particle that has not previously been associated with the chain can be calculated from the same equations. Only, the initial condition (3.12) should be replaced by

$$c(r,z,t=0) = \frac{1}{\pi(R^2-b^2)L}\,, \tag{3.14}$$

if we assume initially one particle homogeneously distributed in the cylinder. As the solution with (3.12) is a fundamental solution it is simpler to integrate

$$G(z,t) = \frac{1}{\pi(R^2-b^2)L} \int_0^L \mathrm{d}z' \int_b^R \mathrm{d}a\, 2\pi a\, F(z,z',t)\,, \tag{3.15}$$

where the parameter a has been used as a variable distance. Eq. (3.15) can be interpreted thus: with the homogeneous initial condition (3.14) the particle will with the probability $2\pi a\,\mathrm{d}a/\pi(R^2-b^2)$ start in the distance interval $(a,a+\mathrm{d}a)$ from the chain; with the probability $\mathrm{d}z'/L$ it will start in the interval $(z',z'+\mathrm{d}z')$. Integrating the fundamental solution over all z' and a gives the solution G.

Now, the mathematical formulation of the problem is complete. Due to the separation of the three-dimensional and on-dimensional diffusion we can first solve the eqs. (3.7)–(3.13) and (3.15) for F and G. They are in principle standard solutions which are known [6]. Entering them in eq. (3.1) gives an integro-differential equation which is also amenable to solution. To achieve a solution in closed form we have to use the Laplace transform. The Laplace transform $\tilde{f}(s)$ of a function $f(t)$ is defined as

$$\tilde{f}(s) = \int_0^\infty \mathrm{e}^{-st} f(t)\,\mathrm{d}t\,. \tag{3.16}$$

All calculations are referred to appendix 1 where it is shown that the Laplace transform of $\Phi(t)$, the flux of particles into the sink, is

$$\tilde{\Phi}(s) = \frac{L\,\tilde{G}(s)}{1+2\{\lambda[1-\tilde{\varphi}(s)]+s\}\sum_{n=1}^{\infty}\{D_1(n\pi/L)^2+s+\lambda[1-\tilde{\varphi}(s+Dn^2\pi^2/L^2)]\}^{-1}}\,, \tag{3.17}$$

where

$$\tilde{\varphi}(s) = k\,\frac{I_1(qR)\,K_0(qa) + K_1(qR)\,I_0(qa)}{I_1(qR)[k\,K_0(qb) + q\,K_1(qb)] + K_1(qR)[k\,I_0(qb) - q\,I_1(qb)]}\,, \tag{3.18}$$

$$\tilde{G}(s) = \frac{2b}{L(R^2-b^2)}\,\frac{k}{q}\,\frac{I_1(qR)\,K_1(qb) - K_1(qR)\,I_1(qb)}{I_1(qR)[k\,K_0(qb) + q\,K_1(qb)] + K_1(qR)[k\,I_0(qb) - q\,I_1(qb)]}\,, \tag{3.19}$$

$$q \equiv \sqrt{s/D}\,. \tag{3.20}$$

I_n and K_n are the nth order modified Bessel functions of the first and second kind respectively.

The complete time development of the association process is now determined by the Laplace inversion of (3.17).

Some interesting particular cases can be considered. For $\lambda = 0$, i.e., a particle cannot dissociate from the chain once it is caught, one gets simply

$$\tilde{\Phi}(s) = L\,\tilde{G}(s)\,\tanh(L^2 s/D_1)^{1/2}/(L^2 s/D_1)^{1/2}\,. \tag{3.21}$$

In summary the results above have been derived under the assumptions:

(1) A closed cylindrical cell around each chain, giving the restrictions discussed in section 2.

(2) Straight chains. As most of the particles dissociating from the chains will be reassociated fast, after small displacements only, they will experience the chains as approximately straight. Those straying too far will lose their correlations to the chains and this assumption should not be crucial.

(3) A particle dissociating from the chain is "lifted" out a distance $a-b$ from the chain where it starts again, thereby $a-b$ is introduced as a parameter of the order of one molecular diameter; $a \approx 2b$. We can let $a \to b$ and have the association governed by the traditional boundary condition (3.10) or let $k \to \infty$ and have the association entirely diffusion limited.

(4) Only one particle per sink; the general case of N ($N \geq 1$) particles per sink is, however, also covered by the results through

$$\Phi_N(t) = N\Phi(t)\left(1 - \int_0^t \Phi(t)\,dt\right)^{N-1}\,.$$

4. The mean times of association and the rate constants for the unspecific binding

Some important results can be derived directly without the explicit inversion of the Laplace transforms. The following relations hold

$$\tilde{\Phi}(0) = \lim_{s\to 0} \int_0^\infty e^{-st}\,\Phi(t)\,dt = \int_0^\infty \Phi(t)\,dt = 1\,, \tag{4.1}$$

$$\tilde{\Phi}'(0) = \lim_{s\to 0} \frac{d\tilde{\Phi}(s)}{ds} = \lim_{s\to 0} \frac{d}{ds} \int_0^\infty e^{-st}\,\Phi(t)\,dt = -\int_0^\infty t\,\Phi(t)\,dt = -\tau\,, \tag{4.2}$$

where τ is the mean time of association. One easily verifies that (4.1) holds for the expression (3.17) which implies that the particle is absorbed in the sink with the probability 1, i.e., with certainty. More interesting is the calculation of the mean time τ. Taking the derivative of (3.17) with respect to s and then letting $s \to 0$ gives

$$\tau = \tau_1 + N\tau_2 + N\tau_3 \approx \frac{1}{2}\frac{R^2}{D}\left(\ln\frac{R}{b} + \frac{1}{bk}\right) + \tfrac{1}{2}N\frac{R^2}{D}\left(\ln\frac{a}{b} + \frac{1}{bk}\right) + N\frac{1}{\lambda}\,, \tag{4.3}$$

where the quantity N is

$$N = \sum_{n=1}^{\infty} \frac{2\lambda}{D_1(n\pi/L)^2 + \lambda[1-\tilde{\varphi}(Dn^2\pi^2/L^2)]} . \tag{4.4}$$

The terms in (4.3) can be interpreted as follows:

$$\tau_1 = -L\tilde{G}'(0) = \frac{1}{2}\frac{R^2}{D}\left\{\frac{1}{1-(b/R)^2}\ln\frac{R}{b} + \left[1-\left(\frac{b}{R}\right)^2\right]\frac{1}{bk}\right\} - \frac{1}{8}\frac{R^2}{D}\left[3-\left(\frac{b}{R}\right)^2\right] \approx \frac{1}{2}\frac{R^2}{D}\left(\ln\frac{R}{b} + \frac{1}{bk}\right) \tag{4.5}$$

is the mean time for the first arrival at the chain.

$$\tau_2 = -\tilde{\varphi}'(0) = \frac{1}{2}\frac{R^2}{D}\left\{\ln\frac{a}{b} + \left[1-\left(\frac{b}{R}\right)^2\right]\frac{1}{bk}\right\} - \frac{1}{4D}(a^2-b^2) \approx \frac{1}{2}\frac{R^2}{D}\left(\ln\frac{a}{b} + \frac{1}{bk}\right) \tag{4.6}$$

is the mean time a particle stays outside the chain before being reabsorbed. Then $N\tau_2$ is the total mean time a particle spends off the chains and we can interpret N as the mean number of times the particle dissociates from the chain before the final absorption in the sink. This interpretation is also born out in the third term

$$\tau_3 = 1/\lambda , \tag{4.7}$$

which, obviously, is the mean time a particle spends on the chain each time, as λ is the molecular dissociation rate. $N\tau_3$ is then the total mean time the particle spends on the chain before the final absorption. With the aid of (3.6) τ can be rewritten as

$$\tau = \frac{1}{n_0 4\pi DL}\left[\ln\frac{R}{b} + \frac{1}{bk} + N\left(\ln\frac{a}{b} + \frac{1}{bk}\right) + Nn_0\frac{4\pi DL}{\lambda}\right]. \tag{4.8}$$

Comparing this to eqs. (2.5) and (2.11) we can identify the association rate k_a

$$k_a = 4\pi DL\left(\ln\frac{R}{b} + N\ln\frac{a}{b} + \frac{N+1}{bk} + Nn_0\frac{4\pi DL}{\lambda}\right)^{-1} . \tag{4.9}$$

This expression then is the extension of the von Smoluchowski result (2.2).

The association rate can be expressed in terms of the equilibrium constant, K_c, for the dissociation of particles from the chain. It is defined from the law of mass action and can be calculated by considering the particles and the chains as before but now without the specific sinks.

$$K_c \equiv \frac{[\text{conc. of free particles}]\,[\text{conc. of chains}]}{[\text{conc. of particles bound to the chains}]} .$$

In the "closed cell" one finds an expression for K_c which is identical to

$$K_c = \frac{\tau_2}{\tau_3}n_0 \approx \frac{1}{2}\frac{R^2}{D}\left(\ln\frac{a}{b} + \frac{1}{bk}\right)\lambda n_0 . \tag{4.10}$$

τ_2, eq. (4.6), the mean time the particle spends off the chain is evidently proportional to the concentration of free particles. Similarly, $\tau_3 = 1/\lambda$ is proportional to the concentration of bound particles. The relation (4.10) is reasonable and constitutes a proof that the mean times are correctly calculated. For the calculation of an equilibrium constant the assumption of a regular distribution of straight chains is totally insignificant as all distributions are homogeneous in equilibrium. Thus, the expressions for K_c and τ_2 (τ_3 also, of course) are correct irrespective of this assumption which instead enters the result through the factor N, eq. (4.4), the total number of visits to the chain.

We can also identify the rate of the unspecific association to the chains

$$k_{ass} = \frac{1}{n_0 \tau_1} = \frac{4\pi LD}{\ln(R/b)} \,. \tag{4.11}$$

Consequently, the unspecific dissociation rate is [13]

$$k_{diss} = K_c k_{ass} = \frac{\tau_2}{\tau_1 \tau_3} = \frac{\lambda[\ln(a/b) + 1/bk]}{\ln(R/b)} \,. \tag{4.12}$$

5. Comparison with experiments for the lac repressor-operator system

There is considerable experimental evidence to support the coupled diffusion model in this system. We have calculated the mean time of association, eq. (4.3),

$$\tau = \frac{1}{2}\frac{R^2}{D}\left(\ln\frac{R}{b} + \frac{1}{bk}\right) + \tfrac{1}{2}N\frac{R^2}{D}\left(\ln\frac{a}{b} + \frac{1}{bk}\right) + N\frac{1}{\lambda} \,.$$

In experiments at low concentrations [1], the first term contributes about 2% of the total mean time (under the assumption $\ln(R/b) \gg 1/bk$, i.e., a diffusion limited reaction), and we shall neglect it henceforth. This is equivalent to the assumption that the unspecific binding to the chain is equilibrated fast, which has been used successfully before [8]. Hence

$$\tau = \tfrac{1}{2}N\frac{R^2}{D}\left(\ln\frac{a}{b} + \frac{1}{bk}\right) + N\frac{1}{\lambda} = \frac{N[\ln(a/b) + 1/bk](1+n_0/K_c)}{4\pi DLn_0} \,. \tag{5.1}$$

The last equality is due to eq. (4.10). We see that in form this expression is identical to that given by Lin and Riggs [8] as the influence of the unspecific binding enters through the factor $(1+n_0/K_c)$, or $(1+D_t/K_{RD})$ in their notation. The difference being that we can interpret the association rate k_a in the molecular parameters. From eq. (4.9) one gets

$$k_a = \frac{4\pi DL}{N[\ln(a/b) + 1/bk](1+n_0/K_c)} \quad \text{cm}^3\,\text{s}^{-1} \,. \tag{5.2}$$

Under quite general conditions (see appendix 2) the sum in eq. (4.4) for N can be approximated as

$$N \approx L\left(\frac{\lambda}{D_1}\,\frac{\ln(2L/\pi b) - \gamma + 1/bk}{\ln(a/b) + 1/bk}\right)^{1/2} \,, \tag{5.3}$$

where $\gamma = 0.577...$ is Euler's constant. Hence

$$k_a = \frac{4\pi D}{1+n_0/K_c}\left(\frac{D_1/\lambda}{[\ln(a/b) + 1/bk]\,[\ln(2L/\pi b) - \gamma + 1/bk]}\right)^{1/2} \text{cm}^3\,\text{s}^{-1} \,. \tag{5.4}$$

The equilibrium constant K_c has been measured in the units μg/cm^3. As DNA weighs approximately 3×10^{-8} μg per cm of its length, we can write from (4.10) (DNA length = $2L$)

$$K_c = \frac{\lambda[\ln(a/b) + 1/bk]}{2\pi D} \times 3 \times 10^{-8} \ \mu\text{g/cm}^3 \,. \tag{5.5}$$

The findings [3,9] under standard experimental conditions (ionic strength $I = 0.05$ M) are

$$k_a = 7 \times 10^9 \ \text{M}^{-1}\,\text{s}^{-1} = 1.2 \times 10^{-11} \ \text{cm}^3\,\text{s}^{-1} \,, \tag{5.6}$$

$$K_c = 1.4 \ \mu\text{g/cm}^3 \,. \tag{5.7}$$

With the regular diffusion constant $D = 5 \times 10^{-7}$ cm^2 s^{-1}, the reaction radius $b = 6 \times 10^{-7}$ cm, the DNA length

$2L = 1.6 \times 10^{-3}$ cm, the DNA concentration $n_0 \ll K_c$, and the assumption $1/bk \ll \ln(L/b)$ (i.e., a diffusion limited reaction) we get from eqs. (5.4)–(5.7):

$$\lambda\left(\ln\frac{a}{b}+\frac{1}{bk}\right)=150\ \mathrm{s}^{-1}\ , \tag{5.8}$$

$$D_1 = 3 \times 10^{-9}\ \mathrm{cm}^2\,\mathrm{s}^{-1}\ , \tag{5.9}$$

$$\left(\frac{D_1/\lambda}{\ln(a/b)+1/bk}\right)^{1/2} = 4.5 \times 10^{-6}\ \mathrm{cm}\ . \tag{5.10}$$

$(D_1/\lambda)^{1/2}$ is a measure of the effective range of the operator, i.e., how far the repressor will diffuse along DNA before dissociating. The expression (5.10) gives the effective enhancement in the association rate from this increased range, i.e., a tenfold increase over the actual length of the operator.

Even stronger support for the picture with coupled diffusion comes from the dependence on the ionic strength. Lin and Riggs [9] have found experimentally

$$\ln K_c \approx \text{const.} + 21.5\, I^{1/2}, \tag{5.11}$$

where I is the ionic strength. Assuming that in eq. (5.5) only $\lambda[\ln(a/b)+1/bk]$ depends on the ionic strength, we have to conclude that

$$\lambda[\ln(a/b)+1/bk] \approx \text{const.} \times \exp(21.5\, I^{1/2})\ ,$$

and in view of (6.8)

$$\lambda[\ln(a/b)+1/bk] \approx 1.2 \exp(21.5\, I^{1/2})\ . \tag{5.12}$$

Then the dependence of the association rate, (5.4), on the ionic strength will be

$$\ln k_a \approx \text{const.} - \tfrac{1}{2}\ln\{\lambda[\ln(a/b)+1/bk]\} \approx \text{const.} - 11\, I^{1/2}\ . \tag{5.13}$$

(We have again neglected n_0/K_c, as $n_0 \ll K_c$ in all the experiments reported here.) This result is in excellent agreement with the experimental finding [3]. The slight curvature of the experimental curve could be explained by the decrease in the reaction radius b (in the factor $[\ln(2L/\pi b)-\gamma+1/bk]^{-1/2}$) with increasing ionic strength through the shielding of the electrostatic interaction as discussed by Richter and Eigen [4].

This dependence on the ionic strength of the association rate is the strongest evidence that the one-dimensional diffusion along the DNA chain is indeed part of the association mechanism. This was the main argument of Richter and Eigen [4] also. Our calculations have shown quantitatively how this dependence arizes from the dependence of the unspecific binding K_c on the ionic strength through the parameter $\lambda[\ln(a/b)+1/bk]$. We have also calculated the one-dimensional diffusion constant that enters this mechanism.

6. Discussion

The association kinetics with coupled three- and one-dimensional diffusion has been described from a fairly simple and quite general approach. The solution is achieved after a separation of the diffusional flows, each acting as a source and sink for the other. The precise coupling between the flows is determined by the boundary condition between them. The parameters describing this not exactly known boundary condition enter the result mainly through the combination $\lambda[\ln(a/b)+1/bk]$ which is connected to the rate of dissociation for the unspecific binding to the chains, k_{diss}, eq. (4.12). Its value is determined by the experimental data to be $k_{\mathrm{diss}} \sim 20\ \mathrm{s}^{-1}$ for the ionic strength $I = 0.05$ M. This is the "total" dissociation rate which enters the chemical rate equations and the law of mass action; it is not equal to the "molecular" dissociation rate, λ, which describes the dissociation of the molecular complex and can be followed by a more or less immediate reassociation.

The most important feature of the model is that it incorporates all correlations along the chain; a particle that has been unspecifically bound will always remember from where it comes after a dissociation. In this manner the entire length of the chain is important and not just the immediate neighbourhood of the operator. These correlations enter the association rate through the factor $[\ln(2L/\pi b) - \gamma + 1/bk]^{1/2}$ (~ 2.4 for the data used). It is to be expected that a random configuration of the chains will lead to a smaller influence of the distance correlations along the chain, which in turn would lead to a slightly lower estimate of D_1 than the one presented above. For the low concentrations of the in vitro experiments the difference from a straight chain should give a negligible influence on the result, but for the relatively tight packing of DNA in a cell the situation would be different.

The trick of considering each DNA chain as enclosed in separate cylindrical cells (which of course is not the case for the in vitro experiments) has the limitations, not very serious, discussed in section 2. In addition to these it seems important that the length and concentrations of chains should be such that they will not lead to a highly deformed cylinder, i.e., one whose length is much greater or much smaller than its diameter.

On the basis on this model it is concluded that the diffusion along DNA is indeed part of the association mechanism in the lac repressor–operator system. The strongest evidence for this conclusion is the quantitative dependence of the association rate on the ionic strength which can be derived from the ionic-strength dependence of the unspecific binding. Moreover, the one-dimensional diffusion constant entering this mechanism is found to be $D_1 \approx 3 \times 10^{-9}$ cm^2 s^{-1}.

In conclusion we would also like to comment on the situation in vivo where the picture seems less clear. It appears that the specific binding of repressor to operator relative to the unspecific binding to DNA is too weak to explain the regulation of the lac operon [9]. Also, the measurements by Jobe et al. [10] on different mutant operators have shown an unexpected curved dependence of the induction ratio on the in vitro repressor–operator dissociation constants. These discrepancies can be explained by the presence of a second binding site [11]. In a subsequent work [12] we will show this quantitatively and also consider in detail the relevant laws of mass action and show that they will differ for the in vivo and in vitro situations. This difference originates from the fact that in vivo there is a collection of closed systems (cells), each with its own operator and a specified number of repressors and no exchange between them.

Acknowledgement

We wish to thank Drs. M. Ehrenberg and R. Rigler for exciting our interest in this problem and for many stimulating discussions. This work has been supported by the Swedish Natural Science Research Council which we gratefully acknowledge.

Appendix 1

The solution of eqs. (3.7)–(3.12) is achieved through the standard separation of variables

$$c(r, z, t) = v(z, t)\, w(r, t)\,. \tag{A.1}$$

The z-dependent part is found [6] to be

$$v(z, t) = \frac{1}{L} + \frac{2}{L}\sum_{n=1}^{\infty} \cos(n\pi z'/L)\cos(n\pi z/L)\exp\left[-Dt(n\pi/L)^2\right]\,. \tag{A.2}$$

The r-dependent part, which is most easily handled in its Laplace-transformed version, is for $r < a$ [6]

$$\tilde{w}(r,s) = \frac{1}{2\pi D} \frac{I_1(qR)\,K_0(qa) + K_1(qR)\,I_0(qa)}{I_1(qR)[k\,K_0(qb) + q\,K_1(qb)] + K_1(qR)[k\,I_0(qb) - q\,I_1(qb)]}$$

$$\times \{[k\,K_0(qb) + q\,K_1(qb)]\,I_0(qr) - [k\,I_0(qb) - q\,I_1(qb)]\,K_0(qr)\}\,, \tag{A.3}$$

where $q \equiv \sqrt{s/D}$.

I_n and K_n are the nth order modified Bessel functions of the first and second kind respectively.

From (A.3) we can calculate the z-independent flux, $\varphi(t)$, of particles onto the chain. Its Laplace transform is

$$\tilde{\varphi}(s) = 2\pi b D \left.\frac{\partial \tilde{w}}{\partial r}\right|_{r=b} = k\,\frac{I_1(qR)\,K_0(qa) + K_1(qR)\,I_0(qa)}{I_1(qR)[k\,K_0(qb) + q\,K_1(qb)] + K_1(qR)[k\,I_0(qb) - q\,I_1(qb)]}\,.$$

The z-dependent flux in eq. (3.13) is equal to

$$F(z, z', t) = \varphi(t)\,v(z, t)\,.$$

Hence, the Laplace transform of F is

$$\tilde{F}(z, z', s) = \frac{1}{L}\tilde{\varphi}(s) + \frac{2}{L}\sum_{n=1}^{\infty} \tilde{\varphi}(s + D(n\pi/L)^2)\cos(n\pi z'/L)\cos(n\pi z/L)\,. \tag{A.4}$$

The Laplace transform $\tilde{G}(z, s)$ of the flux $G(z, t)$ can be calculated through the integration of $\tilde{F}(z, z', s)$ as indicated in eq. (3.15). The result which is independent of z is

$$\tilde{G}(s) = \frac{2b}{L(R^2 - b^2)}\,\frac{k}{q}\,\frac{I_1(qR)\,K_1(qb) - K_1(qR)\,I_1(qb)}{I_1(qR)[k\,K_0(qb) + q\,K_1(qb)] + K_1(qR)[k\,I_0(qb) - q\,I_1(qb)]}\,. \tag{A.5}$$

Now we know all the necessary functions entering eq. (3.1). As this equation contains a convolution integral it is also most easily handled in its Laplace–transformed version which takes care of the time integral:

$$D_1 \frac{\partial^2 \tilde{u}}{\partial z^2} - s\tilde{u} - \lambda\tilde{u} + \lambda \int_0^L \mathrm{d}z'\,\tilde{F}(z, z', s)\,\tilde{u}(z', s) + \tilde{G}(s) = 0\,, \tag{A.6}$$

with the boundary conditions

$$\tilde{u}(0, s) = 0\,, \qquad \left.\frac{\partial \tilde{u}}{\partial z}\right|_{z=L} = 0\,. \tag{A.7}$$

The flux of particles into the sink

$$\tilde{\Phi}(s) = D_1 \left.\frac{\partial \tilde{u}}{\partial z}\right|_{z=0} \tag{A.8}$$

can be calculated after a tedious but straightforward solution of (A.6) and (A.7).

An ansatz satisfying the boundary conditions (A.7) is

$$\tilde{u}(z, s) = \sum_{n=0}^{\infty} \tilde{f}_n(s)\sin[(2n+1)\pi z/2L]\,. \tag{A.9}$$

Entering this and the expression (A.4) into the eq. (A.6) gives

$$-\sum_{n=0}^{\infty}\left(D_1\frac{\pi^2}{4L^2}(2n+1)^2 + s + \lambda\right)\tilde{f}_n(s)\sin\frac{(2n+1)\pi z}{2L} + \lambda\sum_{m=0}^{\infty}\frac{2L}{\pi}\tilde{g}_m(s)\cos\frac{m\pi z}{L}\sum_{n=0}^{\infty}\tilde{f}_n(s)\frac{2n+1}{(2n+1)^2 - 4m^2} + \tilde{G}(s) = 0\,, \tag{A.10}$$

where we have defined

$$\tilde{g}_m(s) = \frac{1}{L}\tilde{\varphi}(s)\,, \qquad m=0\,,$$
$$= \frac{2}{L}\tilde{\varphi}\left(s + D\frac{m^2\pi^2}{L^2}\right), \qquad m>0\,, \tag{A.11}$$

in the expression (A.4).

Multiplying (A.6) by $\cos(m'\pi z/L)$ and integrating from $z=0$ to $z=L$, (Partial integration over $\partial^2\tilde{u}/\partial z^2$ before entering the expressions (A.9) and (A.4) and then completing the integration), gives a relation containing the difficult sum $\Sigma_n (2n+1)\tilde{f}_n(s)/[(2n+1)^2-4m^2]$. Similarly, multiplying (A.10) by $\sin[(2n'+1)\pi z/2L]$ and integrating from $z=0$ to $z=L$ gives another relation containing this sum (m' and n' are arbitrary positive integers or zero). Out of these two relations the sum above can be excluded and after rearranging one gets

$$\tilde{f}_{n'}(s) = \frac{4}{\pi}\frac{\tilde{G}(s)}{(2n'+1)^2 D_1\pi^2/4L^2+s+\lambda}\frac{1}{2n'+1} + \lambda\sum_{m=0}^{\infty}\frac{4}{\pi}\tilde{g}_m(s)\frac{(\pi D_1/2L)\Sigma_{n=0}^{\infty}(2n+1)\tilde{f}_n(s) - L\tilde{G}(s)\delta_{m0}}{\frac{1}{2}L\lambda\,\tilde{g}_m(s)(1+\delta_{m0}) - s - \lambda - D_1(m\pi/L)^2}$$
$$\times\frac{2n'+1}{(2n'+1)^2-4m^2}\frac{1}{(2n'+1)^2 D_1\pi^2/4L^2+s+\lambda}\,. \tag{A.12}$$

Multiplying (A.12) by $2n'+1$ and summing n' from zero to infinity we can solve for $\Sigma_{n=0}^{\infty}(2n+1)\tilde{f}_n(s)$. As

$$\tilde{\Phi}(s) = D_1\left.\frac{\partial\tilde{u}}{\partial z}\right|_{z=0} = D_1\frac{\pi}{2L}\sum_{n=0}^{\infty}(2n+1)\tilde{f}_n(s)$$

we finally arrive at the expression (3.17).

Extensive use has been made of some summation formulas which are difficult to find:

$$\sum_{k=0}^{\infty}\frac{1}{(2k+1)^2+x^2} = \frac{\pi}{4}\frac{\tanh(\pi x/2)}{x}\,,$$

$$\sum_{k=0}^{\infty}\frac{(2k+1)^2}{(2k+1)^2-4m^2}\frac{1}{(2k+1)^2+x^2} = \frac{\pi}{4}\frac{x\tanh(\pi x/2)}{x^2+4m^2}\,,$$

$$\sum_{k=1}^{\infty}\frac{1}{k^2+x^2} = \frac{\pi}{2}\frac{\coth(\pi x)}{x} - \frac{1}{2x^2}\,.$$

Appendix 2

$$N = \sum_{n=1}^{\infty}\frac{2\lambda}{D_1(n\pi/L)^2 + \lambda[1-\tilde{\varphi}(Dn^2\pi^2/L^2)]}\,,$$

where according to (3.18)

$$\tilde{\varphi}(Dn^2\pi^2/L^2) =$$
$$= \frac{I_1(n\pi R/L)K_0(n\pi a/L) + K_1(n\pi R/L)I_0(n\pi a/L)}{I_1(n\pi R/L)[K_0(n\pi b/L) + (n\pi/kL)K_1(n\pi b/L)] + K_1(n\pi R/L)[I_0(n\pi b/L) - (n\pi/kL)I_1(n\pi b/L)]}\,.$$

For $n \lesssim L/a \gg 1$, $b \leqslant a$ and $R \approx L$ we have $K_1(n\pi R/L) \approx 0$, $I_0(n\pi a/L) \approx I_0(n\pi b/L) \approx 1$, $I_1(n\pi R/L) \gg 1$ and $K_0(n\pi a/L) \gtrsim 0.03$.

Then

$$1 - \tilde{\varphi}(Dn^2\pi^2/L^2) \approx 1 - \frac{K_0(n\pi a/L)}{K_0(n\pi b/L) + (n\pi/kL)\, K_1(n\pi b/L)}$$

$$\approx 1 - \frac{-\ln(\frac{1}{2}n\pi a/L) - \gamma + O[(n\pi a/L)^2]}{-\ln(\frac{1}{2}n\pi b/L) - \gamma + 1/bk + O[(n\pi b/L)^2]} = \frac{\ln(a/b) + 1/bk}{\ln(2L/\pi b) - \gamma - \ln(n) + 1/bk},$$

where $\gamma = 0.577...$ is Euler's constant.

As long as the sum above is dominated by the terms for which $n \ll L/b$ the approximation (5.3) holds.

References

[1] G. Adam and M. Delbrück, in: Structural chemistry and molecular biology, eds. A. Rich and N. Davidson (Freeman, San Francisco, 1968) p. 198.
[2] M. Eigen, in: Quantum statistical mechanics in the natural sciences, eds. B. Kursunoglu, S.L. Mintz and S.M. Widmayer (Plenum Press, New York, 1974) p. 37.
[3] A.D. Riggs, S. Bourgeois and M. Cohn, J. Mol. Biol. 53 (1970) 401.
[4] P.H. Richter and M. Eigen, Biophys. Chem. 2 (1974) 255.
[5] M. von Smoluchowski, Z. Phys. Chem. 92 (1917) 129.
[6] H.S. Carslaw and J.C. Jaeger, Conduction of heat in solids (Oxford University Press, London, 1947).
[7] F.C. Collins and G.E. Kimball, J. Colloid. Sci. 4 (1949) 425.
[8] S. Lin and A.D. Riggs, J. Mol. Biol. 72 (1972) 671.
[9] S. Lin and A.D. Riggs, Cell 4 (1975) 107.
[10] A. Jobe, J.R. Sadler and S. Bourgeois, J. Mol. Biol. 85 (1974) 231.
[11] W.S. Reznikoff, R.B. Winter and C.K. Hurley, Proc. Nat. Acad. Sci. USA 71 (1974) 2314.
[12] O.G. Berg and C. Blomberg, J. Mol. Biol., submitted.
[13] Eqs. (4.11) and (4.12) will take quite general forms (independent of the specific assumptions of the "closed cell" approach) if the factor $\ln(R/b)$ is replaced by $\frac{1}{2}\ln(V/v)$; where $V = 2\pi LR^2 = n_0^{-1}$ is the free volume for each chain and $v = 2\pi Lb^2$ is the volume actually occupied by each chain.