Biophysical Chemistry 15 (1982) 41–51
Elsevier Biomedical Press

ASSOCIATION KINETICS WITH COUPLED THREE- AND ONE-DIMENSIONAL DIFFUSION

CHAIN-LENGTH DEPENDENCE OF THE ASSOCIATION RATE TO SPECIFIC DNA SITES

Otto G. BERG [a] and Måns EHRENBERG [b]

[a] *Department of Theoretical Physics, Royal Institute of Technology, S-100 44 Stockholm, and* [b] *Institute of Molecular Biology, Uppsala University, S-751 24 Uppsala, Sweden*

Received 23rd June 1981
Revised manuscript received 1st October 1981

Key words: Diffusion; Association; Rate enhancement; DNA binding; Sliding mechanism

The effective target size of a specific binding site on a long polymeric chain can be considerably enhanced if the associating molecules slide along the chain in a one-dimensional diffusion. Such a mechanism has been proposed to explain the 'faster than diffusion controlled' association of the *lac* repressor protein with the operator site on a long DNA chain. An important experimental tool for investigating the properties of rate-enhancing mechanisms of this type is now becoming available. In these experiments the DNA chain length is varied and the association rate of the investigated protein is measured. Previous attempts to describe theoretically how such one-dimensional sliding processes behave under those conditions were either oversimplified or connected with fundamental errors. We present here a theory for the association rate of proteins with a specific site on DNA molecules with variable lengths. We use a steady-state approach as do Schranner and Richter (Biophys. Chem. 8 (1978) 47) but, in contrast, use boundary conditions which appear to be physically more reasonable.

1. Introduction

Many biological functions rely on the binding of a protein to a specific site on the chromosome. Some examples are repressor-operator binding which turns gene expression on and off, binding of RNA polymerase to the promoter site where transcription starts, and the action of restriction endonucleases which first bind to their recognition sequences before cleaving. These binding sites are defined by a specific DNA sequence. In order to speed up the search for such sites, which are present in the vast excess of structurally similar nonspecific sites on the chromosome, the proteins can be envisaged as sliding along the DNA in a one-dimensional diffusion. Such a process would be possible if the protein had some affinity for the nonspecific DNA, without being held too tightly so that thermal forces would drive a diffusive motion restricted to one dimension along the chain.

This sliding mechanism was first proposed to explain the anomalously fast *lac* repressor-operator association [1], which proceeds at a rate several orders of magnitude faster than that expected from an ordinary diffusion-controlled association between molecules and targets of the size of the operator. Several theoretical calculations have been presented to describe the effects of the sliding mechanism on the association rate constant under various conditions. Thus, Richter and Eigen [2] assumed that such sliding in effect would extend the operator target size to a length equal to the distance a repressor could slide along the chain without intermediate dissociation. The argument here is that a repressor landing within this distance from the operator site would also be able to slide into the site and be bound. Berg and Blomberg [3–5] extended the sliding model by explicity incorporating the coupling between the one-dimensional diffusion along the chain and the three-dimensional one in solution, thus being able to account both for the

0301-4622/82/0000–0000/$02.75

competitive effect of the nonspecific binding and the time delay due to the sliding.

The predictions of the Berg-Blomberg theory (extended and summarized in ref. 6) compare very well with experimental data for the *lac* repressor-operator association [7,8] where the ionic strength has been varied corresponding to a variation in the nonspecific binding strength over several orders of magnitude. The conclusion is that the *lac* repressor does appear to slide with a rate of about 10^{-9}–10^{-10} cm^2 s^{-1}.

The theoretical models discussed above were not intended to describe the association to shorter rod-like DNA fragments. For such a situation, Schranner and Richter [9] calculated a chain-length dependence of the association rate using a steady-state approach for a spheroidal geometry. However, Schranner and Richter used a coupling between the three-dimensional diffusion outside the DNA chain and the one-dimensional diffusion along the chain which seems unreasonable.

They neglected the contribution to the total flux onto the chain of protein molecules dissociated from the nonspecific sites on the DNA chain. This neglect of spatial correlations led them to predict a zero association rate constant to the specific site of an infinitely long DNA chain. On the surface their result is not surprising, since it is well known that an infinitely long absorbing cylinder together with a fixed concentration of diffusing particles at infinite distance leads to zero influx per unit length of the cylinder. In the present work, we show that a mechanical application of this result to the description of associating proteins to specific DNA sites leads to erroneous conclusions. We use a steady-state approach and we proceed by taking the appropriate spatial correlations into account when we formulate the boundary conditions which describe the coupling between the flux from the free state and diffusion flux along the DNA chain. The relevant chain-length dependence of the association rate to specific DNA sites is then derived. Next we show that there indeed exists a nonzero flux to specific sites embedded in infinitely long straight DNA chains and we present an expression in closed form for the association rate constant in this special case.

Finally, we discuss the formal equivalence between steady-state treatments, as in the present work, and approaches based on the calculation of mean association times as employed by Berg and co-workers [3–6].

2. The diffusion equations

To define the relevant molecular parameters, let us first consider an equilibrium situation with one DNA chain in a solution with a homogeneous concentration c_0 of protein. The density u of bound protein per unit length of the chain is assumed to be so small that all exclusion effects can be neglected. The association flux ϕ_{ass} of protein onto the chain is proportional to c_0:

$$\phi_{\mathrm{ass}} = kc_0 \tag{1}$$

This relationship defines the microscopic nonspecific association rate constant k. The corresponding dissociation flux ϕ_{diss} is proportional to binding density u:

$$\phi_{\mathrm{diss}} = \lambda u \tag{2}$$

This relationship defines the microscopic nonspecific dissociation rate constant λ. At equilibrium these two fluxes balance such that the nonspecific binding constant can be expressed as:

$$K_{\mathrm{RD}} = \frac{ulD_{\mathrm{T}}}{c_0 D_{\mathrm{T}}} = \frac{kl}{\lambda} \tag{3}$$

where l is the length of a nonspecific binding site * and D_{T} the total concentration of these sites, such that

---

* l is typically the length of a base-pair ($l \approx 3.4$ Å), since each base-pair is the potential start of a nonspecific binding site. However, the choice of l is not important here other than for the choice of units for the nonspecific binding constant which is normally given per mol base-pair. For simplicity of notation, we shall assume in the following that K_{RD} is given in units of cm^3 and all bimolecular association rate constants in cm^3 s^{-1}.

u/D_T is the total concentration of nonspecifically bound protein.

To derive a specific association rate constant, consider a DNA chain of length $2L$ situated along the z-axis, $-L \leqslant z \leqslant L$, with the specific site in the middle at $z=0$. The effective reaction radius of the chain is b. The concentration of free protein $c(r,z)$ outside the chain satisfies a steady-state diffusion equation which in cylindrical coordinates becomes:

$$\frac{\partial^2 c}{\partial r^2} + \frac{1}{r} \cdot \frac{\partial c}{\partial r} + \frac{\partial^2 c}{\partial z^2} = 0 \tag{4}$$

with the usual boundary condition at infinity

$$c(r,z) \to c_0 \text{ for } r \to \infty \text{ or } z \to \infty \tag{5}$$

When nonspecifically bound, a protein is assumed to be able to slide along the chain with a relative one-dimensional diffusion constant D_1, or dissociate with a rate constant λ. Thus, the density of nonspecifically bound protein $u(z)$ satisfies the diffusion equation:

$$D_1 \frac{d^2 u}{dz^2} - \lambda u + kc(b,z) = 0; \quad -L<z<L, z \neq 0 \tag{6}$$

where the source term kc is the association flux of free protein proportional to the concentration $c(b,z)$ outside the chain surface $r=b$. At the chain ends, we shall assume reflecting boudaries

$$\left.\frac{du}{dz}\right|_{z=\pm L} = 0 \tag{7}$$

whereas at the specific site complete absorption takes place

$$u(0,t) = 0 \tag{8}$$

In this way, the freely diffusing protein is used as a source term for the sliding (eq. 6). The coupling must also be reciprocal, such that the concentration $c(r,z)$ of free protein depends on the binding density u, and $c(r,z)$ from eq. 6 must give the net flux ϕ onto the chain:

$$\phi(z) \equiv 2\pi b D \left.\frac{\partial c}{\partial r}\right|_{r=b} = kc(b,z) - \lambda u(z); \quad -L<z<L \tag{9}$$

where D is the relative three-dimensional diffusion constant. Only the boundary conditions at the flat ends of the rod remain to be specified. To avoid some of the problems with a discontinuous change of boundary conditions, we shall assume that the chain extends beyond $|z|=L$ all the way to infinity, but that this part is totally inert for binding, i.e.

$$\phi(z) \equiv 2\pi b D \left.\frac{\partial c}{\partial r}\right|_{r=b} = 0; \quad |z|>L \tag{10}$$

We note that this inert extension effectively introduces an excluded volume effect, which for all practical purposes will be negligible. Furthermore, since we are unable in any case to specify what kind of association takes place over the flat ends of the rod, this assumption is as justified as any other.

Now, eqs. 4–10 completely determine the problem and the specific association rate constant, k_a, can be calculated from the steady-state flux into the sink at $z=0$

$$k_a c_0 \equiv 2 D_1 \left.\frac{du}{dz}\right|_{z=0} \tag{11}$$

or, equivalently, from the total net flux onto the chain

$$k_a c_0 = \int_{-L}^{L} dz\, 2\pi b D \left.\frac{\partial c}{\partial r}\right|_{r=b} \tag{12}$$

The main difference between our treatment and the formulation of Schranner and Richter [9] appears in eqs. 6 and 9. Our eq. 9 is not only the net source term in the one-dimensional diffusion equation (eq. 6) for the movement along the chain but also serves as the boundary condition at the chain surface for the three-dimensional diffusion equation (eq. 4). In contrast, Schranner and Richter [9] solve eq. 4 or an equivalent diffusion equation in spheroidal coordinates with an absorbing or partially absorbing boundary at the chain surface. By doing this they neglect the influence on the free particle distribution of the efflux from nonspecifically bound molecules on the chain. In effect, their association and dissociation fluxes are coupled only at infinite distance from the chain. This implies a neglect of the cross-fluxes between nonspecific sites where a protein dissociating from one such site can reassociate to one nearby without diffusion to infinity. This in turn leads to an underestimate of the concentration of molecules outside the chain and this underestimate becomes progressively more pronounced the longer the chain. In the limit of infinitely long chains, therefore, they inevitably end up with the result that the influx to the specific site is zero.

The essence of our treatment, where this artifact has been avoided by taking the efflux of particles from the chain into account when the inner boundary condition for the three-dimensional diffusion is formulated, may be illustrated in the following way. Far away from $z=0$ the influence of the absorbing barrier at the specific site is weak. This influence becomes progressively smaller the larger the z-coordinate. Therefore, the nonspecific DNA sites are, to a correspondingly better approximation, equilibrated with the protein molecules in the free state as the distance to the specific site becomes larger. In consequence, the net influx to the chain decreases with the distance to the sink and amounts to zero in the limit of infinite z. Accordingly, the influence of an absorbing barrier at the specific site on the distribution of particles in the free state is only local and does not extend to the whole DNA chain.

3. Solution

3.1. Integral equations

Since the problem is totally symmetric around $z=0$, we need only consider $z\geqslant 0$. The general solution to eq. 4 satisfying the boundary condition at infinite distance (eq. 5) is:

$$c(r,z)=c_0\left[1-\int_0^{\infty}\mathrm{d}\xi f(\xi)\cos(\xi z)K_0(\xi r)\right] \tag{13}$$

where K_0 is a modified Bessel function and $f(\xi)$ an unknown function to be determined from the boundary conditions at $r=b$ (eqs. 9 and 10). The general solution to the sliding equation (eq. 6) with boundary conditions (eqs. 7 and 8) can be expressed as:

$$u(z)=kD_1^{-1}\int_0^{L}\mathrm{d}z'c(b,z')g(z,z') \tag{14}$$

where $g(z,z')$ is the Green's function solution

$$g(z,z')=\begin{cases}\dfrac{\sinh(\mu z)\cosh(\mu L-\mu z')}{\mu\cosh(\mu L)}; & 0<z<z'<L\\[2ex] \dfrac{\sinh(\mu z')\cosh(\mu L-\mu z)}{\mu\cosh(\mu L)}; & 0<z'<z<L\end{cases} \tag{15}$$

and

$$\mu^2\equiv\lambda/D_1$$

Thus, inserting eqs. 13 and 15 into eq. 14 we can derive an expression for the binding density $u(z)$ and thus for the net flux ϕ of protein onto each unit length of chain

$$\begin{aligned}\phi(z)&\equiv kc(b,z)-\lambda u(z)\\ &=kc_0\frac{\cosh(\mu L-\mu z)}{\cosh(\mu L)}\left\{1-\int_0^{\infty}\mathrm{d}\xi f(\xi)K_0(\xi b)\left[1+\left(\frac{\xi}{\mu}\right)^2\right]^{-1}\right\}+kc_0\frac{\sinh(\mu z)}{\mu\cosh(\mu L)}\int_0^{\infty}\mathrm{d}\xi f(\xi)K_0(\xi b)\xi\sin(\xi L)\left[1+\left(\frac{\xi}{\mu}\right)^2\right]^{-1}\\ &\quad-kc_0\int_0^{\infty}\mathrm{d}\xi f(\xi)\cos(\xi z)K_0(\xi b)\left[1+\left(\frac{\mu}{\xi}\right)^2\right]^{-1};\quad 0<z<L\end{aligned} \tag{16}$$

The boundary conditions, eqs. 9 and 10, are transformed to

$$2\pi Dc_0\int_0^\infty \mathrm{d}\xi f(\xi)\, \xi bK_1(\xi b) \cos(\xi z)=\begin{cases}\phi(z): & 0<z<L\\ 0: & z>L\end{cases} \tag{17}$$

or, equivalently,

$$f(\xi)=(\pi^2Dc_0)^{-1}\{\xi bK_1(\xi b)\}^{-1}\int_0^L \mathrm{d}z\,\phi(z)\cos(\xi z) \tag{18}$$

Now, the problem is uniquely defined by the two coupled integral equations, eqs. 16 and 18. Although they can be combined to form a single integral equation in $f(\xi)$, it is for numerical reasons more convenient to keep them separate.

From eq. 12 the specific association rate constant is given by

$$k_a=2c_0^{-1}\int_0^L \mathrm{d}z\,\phi(z)=2\pi^2Df(0) \tag{19}$$

3.2. *Infinite chains*

Contrary to what has been claimed by Richter and Eigen [2] and by Schranner and Richter [9], there exists a steady-state solution for the absorption of proteins to a specific site on an infinitely long straight DNA chain. Furthermore, in this special case where $L\to\infty$, a closed expression for k_a can be found. One obtains

$$f^\infty(\xi)=\frac{\left[1/\left\{\xi^2K_0(\xi b)+(2\pi D/k)(\xi^2+\mu^2)\,\xi bK_1(\xi b)\right\}\right]}{\left[\pi/2\mu+\int_0^\infty \mathrm{d}\xi\,\mu^2/(\xi^2+\mu^2)\left\{\xi^2+(2\pi D/k)(\xi^2+\mu^2)\,\xi bK_1(\xi b)/K_0(\xi b)\right\}\right]} \tag{20}$$

Transforming to dimensionless entities

$$\xi b\equiv x,\quad 1/\mu b\equiv\gamma \tag{21}$$

$$k/2\pi D\mu^2b^2\equiv\alpha^2 \tag{22}$$

one has from eqs. 19 and 20 the association rate constant

$$k_a^\infty=\frac{2\pi bD}{\left[\gamma/2\alpha^2+\pi^{-1}\int_0^\infty \mathrm{d}xK_0(x)/xK_1(x)\left\{1+(\gamma x)^2\right\}\left\{1+(\gamma x)^2+\alpha^2xK_0(x)/K_1(x)\right\}\right]} \tag{23}$$

In fig. 1, $k_a^\infty/2\pi bD$ has been plotted as a function of α for small values of the parameter γ. The physical

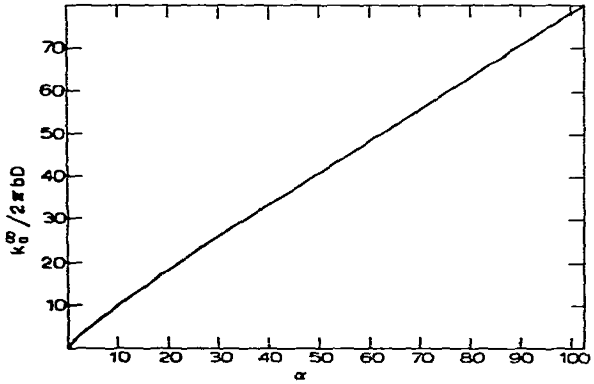

Fig. 1. The association rate constant for infinite chain length from eq. 23 as a function of the effective sliding distance α.

meaning of these parameters can be explained as follows:

$$\gamma b \equiv 1/\mu \equiv (D_1/\lambda)^{1/2} \tag{24}$$

is the microscopic sliding distance, i.e., the average distance the protein can slide without microscopic dissociation.

$$\alpha b \equiv (kD_1/\lambda 2\pi D)^{1/2} \equiv (K_{\mathrm{RD}} D_1/2\pi D l)^{1/2} \tag{25}$$

can be interpreted as a macroscopic (or effective) sliding distance, i.e., the distance the protein can slide along the chain before 'total' dissociation (i.e., microscopic dissociation and diffusion away from the chain). As discussed previously [5,6,10], in the diffusion-controlled limit, a total (or macroscopic) dissociation is preceded by a very large number of unsuccessful microscopic dissociations which leave the protein free in solution for a very short time before reassociation to the same or to a neighboring site on the chain. Then, $\gamma \ll \alpha$, or equivalently, $k \gg 2\pi D$, and in this limit one also finds numerically that the result of eq. 23 does not depend on γ. We can take this as a practical definition of diffusion control: choose λ and k sufficiently large (while $kl/\lambda = K_{\mathrm{RD}}$ is fixed) such that the result becomes independent of λ. All numerical results below are calculated in this limit.

The sliding distance α as defined by eq. 25 depends only on such physical characteristics as the diffusion constants and the nonspecific binding constant K_{RD}. As a consequence it is well suited as an experimentally measurable parameter of the theory. In contrast, a definition which relies on a macroscopic nonspecific dissociation rate constant fails, since such a rate constant cannot be uniquely defined in this system. In fact, for such a definition it must be specified exactly how far from the chain a protein has to reach before it is counted as dissociated [6].

4. Finite chains

4.1. Length dependence

To be more suitable for numerical integration, eqs. 16, 18 and 19 can be manipulated further to give (neglecting terms of the order $e^{-\beta/\gamma} = e^{-\mu L}$)

$$\Psi(z) = \frac{1}{\pi}\int_0^{\infty} \mathrm{d}x \cos(xz) \frac{1 + A\cos(\beta x) + U(x)\alpha^2 x K_0(x)/K_1(x)}{1 + (\gamma x)^2 + \alpha^2 x K_0(x)/K_1(x)}; \quad 0 < z < \infty \tag{26}$$

where

$$U(x) \equiv 2\int_{\beta}^{\infty} \Psi(z)\cos(xz)\,\mathrm{d}z \tag{27}$$

$$\beta \equiv L/b \tag{28}$$

and A is solved from

$$A = \frac{2}{\pi}\int_0^{\infty} \mathrm{d}x \frac{\sin(x\beta)}{x} \frac{\alpha^2 x K_0(x)}{K_1(x)} \frac{1 + A\cos(\beta x) - U(x)\{1 + (\gamma x)^2\}}{1 + (\gamma x)^2 + \alpha^2 x K_0(x)/K_1(x)} \tag{29}$$

Then, the specific association rate constant is given by

$$k_{\mathrm{a}} = \frac{2\pi b D}{\dfrac{\gamma}{2\alpha^2} + \dfrac{1}{\pi}\displaystyle\int_0^{\infty} \mathrm{d}x \frac{1}{1 + (\gamma x)^2} \frac{K_0(x)}{xK_1(x)} \frac{1 + A\cos(\beta x) - U(x)\{1 + (\gamma x)^2\}}{1 + (\gamma x)^2 + \alpha^2 x K_0(x)/K_1(x)}} \tag{30}$$

In the infinite chain limit, $L \to \infty$, $\beta \to \infty$, $A \to 0$, $U(x) \to 0$ and the result, eq. 23, is recovered. The equation system, eqs. 26–29, has been solved iteratively using fast Fourier-transform routines on a computer starting with the ansatz $A = U(x) = 0$ in eq. 26, i.e., starting with the infinite chain solution and then calculating new values for A and $U(x)$ from eqs. 27 and 29. The convergence of the procedure can be checked by an internal consistency requirement, $A \to U(0)$. It is also confirmed by the numerical result that

$$\frac{k_a}{2\pi bD} \cong \frac{2\beta}{\ln(2\beta)} \tag{31}$$

for all $\beta \gg 1$ when the sliding length is much greater than the chain length, i.e., when $\alpha \gg \beta$. This is the limit when the whole chain serves as an absorbing target. In this limit alone, our result agrees with that of Schranner and Richter. This is as expected, since in this case there can be no cross-fluxes like those that were neglected in their derivation. The approximate agreement in this limit also confirms that our different way of treating the chain ends is of no fundamental consequence; i.e., the inert chain extension is not significantly different from approximating the DNA cylinder by a spheroid.

In fig. 2 we have plotted k_a as a function of the sliding length α for various chain lengths β. In fig. 3 k_a is plotted as a function of β for various values of α. A direct comparison with the predicted length dependence from Schranner and Richter [9] is difficult, since their sliding distance is based on the existence of a uniquely defined dissociation rate constant. Although such a rate constant can be defined for the finite chains, it must be length dependent [6]. Consequently, their results [9] must carry an implicit length dependence not accounted for by their theory.

4.2. Direct association to a specific site without sliding

The present treatment allows us to derive in a very direct way the association rate constant to a specific DNA site in the case when the protein molecule is unable to diffuse along the DNA chain. It is somewhat surprising that this fundamental situation, i.e., when there is no rate-enhancing mechanism present, has been given such scarce attention in the literature. To our knowledge the only derivations published are those by Berg et al. [6] where the analysis was based on mean association time calculations. We are now in a position where we can offer a simpler solution to the problem. Since the inherent mathematical difficulties cannot be completely avoided we shall derive suitable approximations.

Consider a DNA chain which is much longer than the dimension of the specific site. To calculate the flux of particles onto the specific site we need the solution to the following diffusion problem. Along the

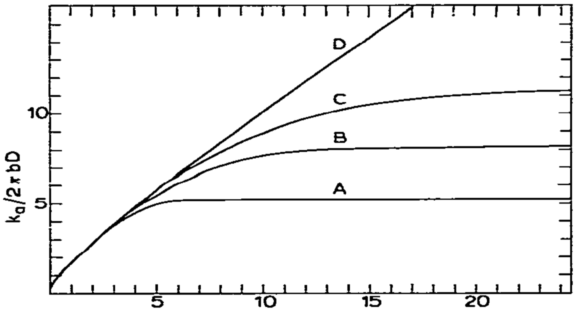

Fig. 2. The association rate constant as a function of the effective sliding distance α for various chain lengths. Curves: A, $\beta=6$; B, $\beta=12$; C, $\beta=20$; D, $\beta\to\infty$.

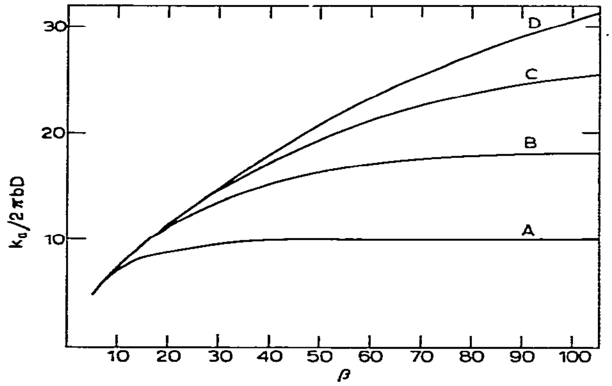

Fig. 3. The association rate constant as a function of chain length β for various values of the effective sliding distance. Curves: A, $\alpha=10$; B, $\alpha=20$; C, $\alpha=30$; D, $\alpha=40$.

z-axis there is a cylinder with radius b. It constitutes a reflecting barrier for free particles everywhere except in the interval $-L' \leqslant z \leqslant L'$ which defines the length of the specific site and where incoming particles are absorbed. At infinity the free protein concentration has the value c_0 and we look for the stationary flux over the absorbing barrier. It is clear that the solution to this problem follows as a special case of the present treatment if we take the limits $k \to \infty$ (diffusion control) and $\lambda \to 0$ (no return to the free state). The whole DNA chain of length $2L$ is now an absorber embedded in a reflecting cylindrical extension where no sliding takes place and we derive our result by making the identification $L = L'$ so that $2L$ now means the length of the specific site.

From eqs. 16, 18 and 19, one finds in the limit $k \to \infty$, $\lambda = 0$

$$\int_0^{\infty} \mathrm{d}\xi f(\xi) \cos(\xi z) K_0(\xi b) = 1; \quad 0 \leqslant z < L' \tag{32}$$

$$f(\xi) = \frac{1}{\pi^2 D c_0} \frac{1}{\xi b K_1(\xi b)} \int_0^{L'} \mathrm{d}z\, \phi(z) \cos(\xi z) \tag{33}$$

$$k_a = 2\pi^2 D f(0) \tag{34}$$

eq. 32 is simply the requirement that the surface concentration is zero for an ideal sink, and eq. 33 gives the net stationary flux zero onto the nonspecific part $z > L'$ of the chain.

One simple approximation can be derived by assuming the total net flux, $k_a c_0$, to be distributed homogeneously over the site such that $\phi(z) = k_a c_0 / 2L'$. Then, from eq. 33 one finds

$$f(\xi) = \frac{k_a}{2\pi^2 D L'} \frac{1}{\xi b K_1(\xi b)} \frac{\sin(\xi L')}{\xi} \tag{35}$$

Demanding that eq. 32 be valid at least at $z = 0$, i.e., zero surface concentration at least outside the middle of the site, one finds

$$k_a \cong \frac{4\pi D L'}{\dfrac{2}{\pi} \displaystyle\int_0^{\infty} \mathrm{d}\xi \frac{K_0(\xi)}{\xi b K_1(\xi b)} \frac{\sin(\xi L')}{\xi}} \tag{36}$$

For $L' \gg b$, this gives approximately (if $\xi b K_1(\xi b) \approx 1$ in the region where the integrand contributes most)

$$k_a \cong \frac{4\pi D L'}{\ln(2L'/b)} = 2\pi D b \frac{2\beta}{\ln(2\beta)}$$

which is the expected result for this limit as in eq. 31. However, the approximations employed would be even better for $L' < b$ and consequently eq. 36 will be a reasonable approximation also in this interesting limit.

For smaller L', the asymptotic parts of the Bessel functions contribute increasingly $K_0(x)/xK_1(x)_{x \to \infty} \approx 1/(x + \frac{1}{2})$.

A useful approximation here is $K_0(x)/xK_1(x) \approx K_0(x) + 1/(x + \frac{1}{2})$, giving from eq. 36

$$k_a \cong \pi D b \frac{\pi}{1 + \frac{1}{2} \ln(2b/L')} \tag{37}$$

for $L'/b \ll 1$. For the special case when the specific site is defined within one base-pair, $2L' = l$ and $L'/b \cong 0.1$, eq. 37 gives $k_a \cong 1.3\pi D b$ in good agreement with the result of Berg et al. [6].

The importance of employing appropriate boundary conditions, or equivalently of taking 'hopping' [6] into account, for the derivation of k_a is clearly seen from eq. 37. The fact that the absorbing specific site is embedded in a reflecting cylinder leads to the flux depending only very weakly on the length of the specific site for small values of L'. In fact, it is the cylinder radius b rather than the length L' which determines the cross-section for the reaction in this special case of $L'/b \ll 1$.

4.3. *A comparison between the mean-time approach and the steady-state treatement*

We can also compare the present steady-state result with the previous mean-time calculations [3–6]. By taking the appropriate infinite limit in ref. 3, exactly the same expression for k_a^∞ can be derived as that given in eq. 23, thus confirming the equivalence between the two approaches. The mean-time approach has the feature that it 'automatically' includes competition effects, i.e., the time delay due to nonspecific binding. In the steady-state picture above, even an infinite chain constitutes an infinitely dilute system and to include effects such as the nonspecific competition, one must make the proper identifications. In agreement with the Von Smoluchowski treatment [11], the bulk concentration c_0 of protein at infinity should be interpreted as the free concentration, i.e., that fraction of the total concentration which is not nonspecifically bound. If the concentration of chains in the experimental system is n_0, the average concentration of nonspecifically bound protein will be $n_0 \cdot 2\int_0^L u(z)\mathrm{d}z$ where u is the density of bound particles along the chain. Then, the total concentration of protein is given by $c_{\text{tot}} = c_0 + n_0 \cdot 2\int_0^L u(z)\mathrm{d}z$. With the relationships from section 3 above, this gives a competition factor

$$\frac{c_0}{c_{\text{tot}}} = \left\{1 + D_\text{T} K_\text{RD}\left[1 - \int_0^\infty d\xi f(\xi)\frac{\sin(\xi L)}{\xi L}\left(K_0(\xi b) + \frac{2\pi D}{k}\xi b K_1(\xi b)\right)\right]\right\}^{-1} \tag{38}$$

The competition factor is a multiplicative correction to the overall association rate constant which should be included in all results above. For weak nonspecific binding, $D_\text{T} K_\text{RD} \ll 1$, $c_0/c_{\text{tot}} \approx 1$ and it does not contribute. The part under the integral sign takes care of the reduced binding density along the chains within the sliding distance from the specific site. For short sliding lengths and long chains, $\alpha \ll \beta$, this reduction is small and the competition factor is the standard one: $c_0/c_{\text{tot}} \approx [1 + D_\text{T} K_\text{RD}]^{-1}$.

In the interesting limit of long chains and long sliding distances, ($\alpha > \beta \gg 1$) as well as strong nonspecific binding ($D_\text{T} K_\text{RD} \gg 1$), this competition factor gives an overall association rate which is determined by the sliding time along the whole chain, $k_a n_0 \propto D_1/L^2$. Thus, by taking the nonspecifically bound particles into account in this way, we arrive also in this extreme case at a result corresponding to what has been obtained from mean-time calculations [5,6].

5. Discussion

We have calculated the rate constant for association to specific DNA sites of protein molecules, which when nonspecifically bound, can diffuse along the DNA chain. In particular, we have derived how the length of the DNA chain surrounding the specific site influences the association rate. The derivations were based on the assumption of a steady state for the diffusional flux of particles onto the specific site. This method of deriving rate constants was first employed by Von Smoluchowski [11] and since then has had many applications. As far as the method is concerned, our approach is thus identical to that of Richter and Eigen [2] and of Schranner and Richter [9]. Yet, our results diverge from theirs and in some limits they even differ drastically. The source of this discrepancy can be traced to the coupling employed here between the diffusional flux from the free state onto the DNA chain and the sliding of the protein on the chain to its target at the specific site. Since this flux coupling is at the very heart of the present treatment and leads to questions of general physical interest, we shall comment further on this topic.

The DNA chain surrounding the specific site acts as a sink for the freely diffusing particles outside it. This led Richter and Eigen [2] to comment that: "It is well known that an infinitely long chain would be too powerful a sink as to allow for a nontrivial steady state solution". Subsequently, Schranner and Richter [9] derived in spheroidal coordinates the rate of association to a specific site on a long DNA chain and indeed in their calculation the association rate constant tended to zero as the length of the chain increased beyond limits. The reason why they obtained this result is that they used the whole stretch of the DNA chain as an absorber of particles from the free state without taking into account properly the reentry of molecules into the free state from the chain. In fact, the strength of the DNA molecule as a sink is determined by the net steady-state influx of particles onto the chain. This net influx depends strongly on the distance from the specific site and a very long DNA molecule will accordingly absorb particles from the free state preferentially around the sink at the specific site over a distance comparable to the macroscopic sliding distance α as defined above. Therefore, even if the chain length tends to infinity the effective sink for the particles in the free state will remain finite in length.

The boundary conditions employed in the present work automatically take these features into account and allow us to derive directly the proper rate constant for association to specific sites on DNA chains of finite and varying length (fig. 3). These boundary conditions also make the results independent of any assumed—and thereby, within the theory, undefined—macroscopic nonspecific dissociation rate constant. In addition, we have derived in a cylindrical coordinate system a close expression for the rate constant for association to a specific site embedded in a straight DNA chain of infinite length (eq. 23). In contrast to the result of Schranner and Richter [9], our calculations show that when the chain length increases, the association rate constant increases monotonically and reaches a plateau when the chain length becomes much larger than the macroscopic sliding distance α. Thus, the discrepancy between our results and those of Schranner and Richter will be small when the chain is much shorter than the sliding distance α (fig. 2). However, when the chain length increases, the importance of a correct assessment of the boundary conditions becomes more and more critical. Thus, at infinite chain length, we obtain the maximal rate constant k_a (figs. 2 and 3) while Schranner and Richter [9] obtain that for zero association.

From this it also follows that there is no fundamental reason to employ spheroidal coordinates for this diffusion problem, contrary to what was claimed by Richter and Eigen [2]. Indeed, it turns out that when correct boundary conditions are used the spheroidal coordinates increase the mathematical difficulties in relation to cylinder coordinates.

If we want to identify a sliding mechanism from experiments we must compare it with the limiting case of no sliding. The rate constant for association to a specific DNA site in a chain where no one-dimensional diffusion occurs, derived here as a side result, therefore constitutes an important reference point. One interesting feature of this direct association is that there is a very weak dependence on the length of the specific site. The rate constant is thus not proportional to the length itself but is related to the logarithm of the extension along the z-axis of the specific site. Instead, it is the radius of the DNA chain which enters linearly in the expression of the association rate constant. This is a purely geometrical effect. It follows directly from the boundary conditions employed here and this result of the present work is in accord with a derivation by Berg et al. [6] based on mean association times.

Acknowledgements

We thank Professor C.G. Kurland and Dr. K. Booth for their comments on the manuscript. This work was supported by the Swedish Natural Science Research Council.

References

1 A.D. Riggs, S. Bourgeois and M. Cohn, J. Mol. Biol. 53 (1970) 401.
2 P.H. Richter and M. Eigen, Biophys. Chem. 2 (1974) 255.
3 O.G. Berg and C. Blomberg, Biophys. Chem. 4 (1976) 367.
4 O.G. Berg and C. Blomberg, Biophys. Chem. 7 (1977) 33.
5 O.G. Berg and C. Blomberg, Biophys. Chem. 8 (1978) 271.
6 O.G. Berg, R.B. Winter and P.H. von Hippel, Biochemistry 20 (1981) 6929.

7 M.D. Barkley, Biochemistry 20 (1981) 3833.
8 R.B. Winter, O.G. Berg and P.H. von Hippel, Biochemistry 20 (1981) 6961.
9 R. Schranner and P.H. Richter, Biophys. Chem. 8 (1978) 135.
10 O.G. Berg, Chem. Phys. 31 (1978) 47.
11 M. von Smoluchowski, Z. Phys. Chem. 92 (1917) 129.