Vol. 34 (2003) *ACTA PHYSICA POLONICA B* No 7

GEL ELECTROPHORESIS AT HIGH FIELDS*

P. Paściak[a], M.J. Krawczyk[a], M. Kopeć[a], J. Dulak[b]
and K. Kułakowski[a]

[a]Faculty of Physics and Nuclear Techniques, University of Mining and Metallurgy
Mickiewicza 30, 30-059 Kraków, Poland
[b]Faculty of Biotechnology, Jagellonian University
Gronostajowa 7, 30-387 Kraków, Poland

(Received November 6, 2002)

We investigate the band velocity v and the diffusion coefficient D of DNA in gel for the geometration mechanism. Here we treat the geometration as a one-step process. The velocity dispersion shows a maximum for some small number of hookings. This allows to expect that the diffusion constant behaves in a similar way. On the other hand, the average number of hookings should increase with the molecule length. Summarizing, the diffusion constant should decrease for very long molecules. However, we do not find this effect in experimental data. This contradiction can be resolved by a conclusion that multiple hookings do not occur.

PACS numbers: 05.60.Cd, 87.15.Vv

1. Introduction

Gel electrophoresis is an experimental technique commonly used to separate DNA molecules by length. The idea is to measure the distance passed by molecules in gel under action of an applied electric field. The friction force depends on the molecule length in such a way that shorter molecules pass a greater distance than the longer ones. After some time the molecules form separated bands in gel, and each band contains DNA of different length. The technique attracts much attention and effort of theoreticians and computational physicists, who have found there several interesting problems to solve. For reviews of these problems and efforts see Refs. [1,2].

Here we concentrate on the geometration effect. Geometration is a specific kind of motion, which occurs for long molecules in gel under action of large electric field [3]. During this motion, molecules are successively

---

* Presented at the XV Marian Smoluchowski Symposium on Statistical Physics, Zakopane, Poland, September 7–12, 2002.

stretched and compressed because of hookings at the gel fibres. Theory of this effect was given by Popelka *et al.* [4], and generalized in Refs. [5,6]. The goal of this paper is an extrapolation of the existing models to the case of multiple hookings of a molecule at gel fibres. The results are compared to our experimental data.

2. Analytical extrapolation

Our aim is to investigate the case of a few of hookings: two or more. We develop a system of coupled fundamental equations for the probability $p(n,t)$ that a molecule is hooked n times. The equations are

$$\begin{aligned}\frac{dp(n,t)}{dt} &= -p(n,t)[w(n,n-1)+w(n,n+1)] \\ &\quad +p(n-1,t)w(n-1,n)+p(n+1,t)w(n+1,n)\,, \qquad (2.1)\end{aligned}$$

where $k,n=1,\ldots,N$, $w(n,k)$ is the probability of state change from n to k. As $w(n,k)$ vanishes for $n-k\neq\pm1$, the problem belongs to the family of one-step processes [7]. The probability of a hooking within the time period $(t,t+dt)$ is $v(n)dt/\lambda(n)$, where $v(n)$ is the velocity of molecules in an n-th state, and $\lambda(n)$ is the mean free path of molecules. Our postulate is to write $v(n)/\lambda(n)=\gamma u/(\lambda(n+1))$, where u,λ are constants: u is the velocity without hooking and γ is a phenomenological parameter. The justification of this postulate is that certainly $v(n)$ decreases with n, but not the mean free path.

The problem remaining is to propose $w(n+1,n)$. In other words, we have to evaluate how quickly the number of hookings decreases for various n. Two arguments should be brought about here. First is that for high values of n, the average length of the part of molecule which is to be unhooked is short. Second is that for the same reasons, the velocity of unhooking is small, because the driving arm is short. These two agents cancel at least partially. Then we are going to consider two model propositions:

I. $w(n+1,n)=1/\tau=\mathrm{const.}(n)$, where τ is the mean time of unhooking;

II. $w(n+1,n)=n/\tau$, where τ is the mean time of unhooking from the state $n=1$.

With these two possibilities we are going to find the stationary distribution $p(n,\infty)$, denoted for the sake of brevity $p(n)$ from now on. The upper limit of the summation is bounded by the number N of reptons, *i.e.* the molecule length. For the model I we get

$$p(n) = A(x,N)\frac{x^n}{n!}\,, \tag{2.2}$$

where $x = \gamma u\tau/\lambda$, and $A(x,N)$ is the normalization constant. In the limit of infinite N we obtain the Poisson distribution, and x is just the average number of hookings. If x is not too large, the distribution of n can be approximated by this limit form, because the probabilities of higher n are small. In this limit, the mean values of the velocity and its second moment can be expressed for both choices by means of some special functions. However, for our purposes it is much simpler to rely on numerical calculations. The exception is that for the model I we get a close formula for the mean velocity

$$\frac{v}{u} = \frac{1-\exp(-x)}{x} \approx 1 - \frac{x}{2} + \frac{x^2}{6} - \frac{x^3}{24} + \ldots\,. \tag{2.3}$$

For both models, the interpretation of the quantity x should be consistent with our previous analytical results for one or two hookings (n=1,2) [5, 6]

$$\begin{aligned} \frac{v}{u} &= \frac{1}{1+\frac{3s}{4\lambda}+\frac{s^2}{3\lambda^2}}\,, && (2.4)\\ \tau(1) &= \frac{5s}{4u}\,, && (2.5) \end{aligned}$$

where s is the molecule length and $\tau(1)$ is the time of sliding from a single point of hooking. This consistency is equivalent to a condition that the slope of the curve v/u *versus* s/λ at $x = 0$ should be $-3/4$ in both models, I and II. The point is that at $x = 0$, the slope is determined by the time τ, which is the same in both models. For the model I we get

$$\frac{x}{2} \equiv \frac{\gamma u\tau}{2\lambda} = \frac{3s}{4\lambda} \equiv \frac{3u\tau(1)}{5\lambda}\,. \tag{2.6}$$

The condition $\tau(1) = \tau$ gives the coefficient $\gamma = 6/5$. For simplicity, we keep the velocity $u = 1$ from now on.

In Figs. 1, 2 we show the plots of the mean velocity $v = \langle 1/n\rangle$ and its second cumulant $\sigma_v = \langle 1/n^2 - v^2\rangle^{1/2}$, calculated in models I and II as dependent on the dimensionless ratio x. The dispersion σ_v should give an information on the diffusion constant D. Note that the latter is connected with the bandwidth w at a given time t of measurement, by the relation $w^2 = w_0^2 + 2Dt$. Although there is no direct relation between σ_v and D, we believe that D must decrease when the velocity distribution gets narrow.

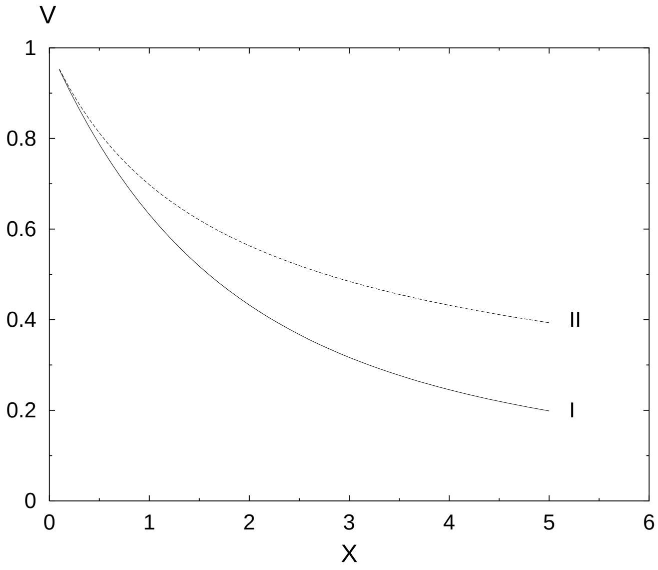

Fig. 1. The velocity v/u as dependent on $x = u\tau/\lambda$ for the models I (full line) and II (dotted line). The parameter x is expected to increase monotonically with the molecule length. In the model I, x is just the average number of hookings.

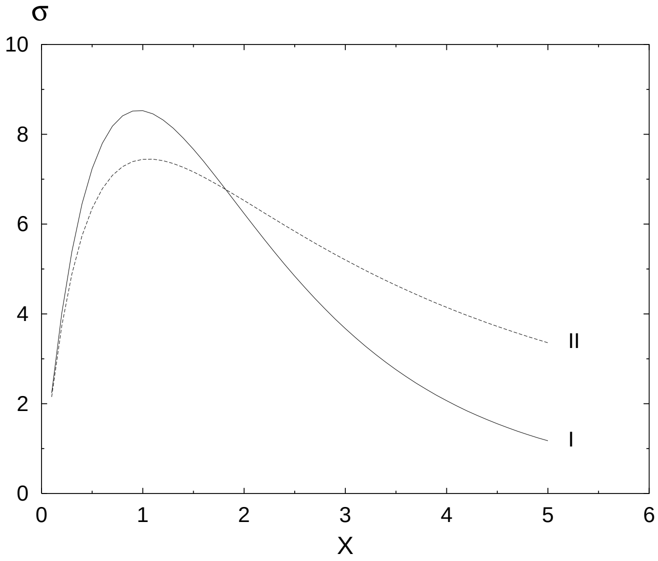

Fig. 2. The velocity dispersion σ_v as dependent on $x = u\tau/\lambda$ for the models I (full line) and II (dotted line).

The conclusion of this chapter is that for a wide set of decreasing functions $\tau(n)$ we can expect a reduction of the bandwidth, when the average number of hookings is above one. The origin of the maximum is that for large number n of hookings, the velocity $v(n) \propto 1/n$ changes within a small period near zero. The above conclusion becomes a question, when we pass to next chapters. Can the maximum be seen in the experiment? The answer depends on the ratio s/λ. However, we do not know a priori how large the mean free path λ is. Then we do not know if we are able to observe multiple hookings in the experimental data.

3. Experimental data

Only preliminary results can be presented here, and more data will be published elsewhere. The method of measurement is the same as in Ref. [5]. The electric field is 7.5 V/cm, and the agarose gel concentration is 1 percent. The bacteriophage lambda 48.5 kbp is cut by the restriction enzyme Hind III and the longest three fragments obtained in this way are of 23.13, 9.42 and 6.56 kbp. The velocity and diffusion constant dependences on the molecule length are shown in Figs. 3 and 4, respectively. The diffusion coefficient D is obtained with an assumption, that the observed band is a Gaussian dependence $\rho(x)$ of the density of DNA on the coordinate x. We note that the geometration does not occur for the shortest length 6.56 kbp, or at least the effect is not typical there [9]. For long molecules, the diffusion coefficient increases with the molecule length. We do not observe any manifestation of a maximum of D with s.

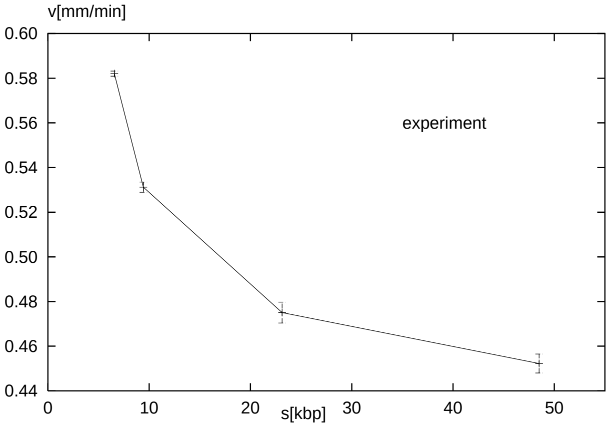

Fig. 3. Experimental data on the velocity v of DNA against the molecule length s.

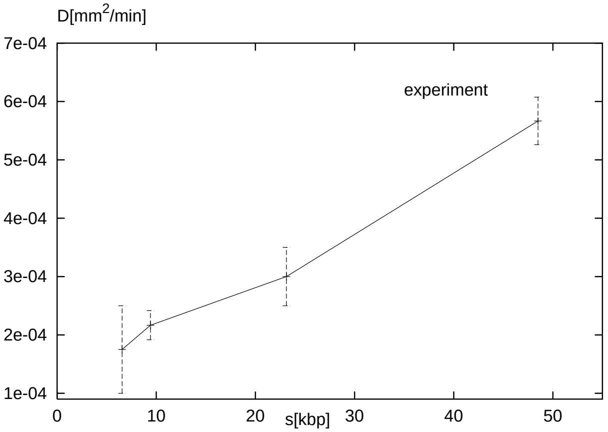

Fig. 4. Experimental data on the diffusion coefficient D of DNA against the molecule length s.

4. Discussion

The results described in Sections 2 and 3 can be summarized as follows. Analytical considerations based on the fundamental equations suggest, that the diffusion coefficient shows a maximum with the average number n of hookings. The result does not depend much on a particular form of the transition probabilities $w(n,n')$, and therefore it seems to be generic. It seems also natural to expect that the average number of hookings is proportional to the molecule length. Then, both the diffusion coefficient and the bandwidth should display a maximum with the molecule length. However, the experiment does not confirm this maximum. A possible solution of this contradiction is that the actual number of hookings is not greater than one. Then, what we see from experiment (Fig. 4) is just the left, ascending part of the curves obtained from the fundamental equations (Fig. 2). In other words, the mean free path λ is longer than the molecule length s.

This result can be understood in frames of the geometration picture. Once hooked, the molecule is stretched along the field and its arms are unlikely to hook again. On the contrary, the conformation of unhooked molecules is far from a straight line because of entropy forces, and in this state a hooking becomes probable.

We note that the accordance between the results of the two methods is only qualitative. In particular, the experimental variation of the velocity with the molecule length (Fig. 3) is smaller than this one calculated (Fig. 1). The list of possible origins of this discrepancy includes thermal fluctuations, which are absent in the analytical calculation.

REFERENCES

[1] G.W. Slater, C. Desruisseaux, S.J. Hubert, J-F. Mercier, J. Labrie, J. Boileau, F. Tessier, M.P. Pepin, *Electrophoresis* **21**, 3873 (2000).

[2] J.-L. Viovy, *Rev. Mod. Phys.* **72**, 813 (2000).

[3] J.M. Deutsch, *Science* **240**, 22 (1988).

[4] S. Popelka, Z. Kabatek, J.-L. Viovy, B. Gas, *J. Chromatography A* **838**, 45 (1999).

[5] M.J. Krawczyk, J. Dulak, K. Kułakowski, *Electrophoresis* **23**, 182 (2002).

[6] M.J. Krawczyk, K. Kułakowski, *Theory in Biosciences* **121**, 231 (2002).

[7] N.G.van Kampen, *Stochastic Processes in Physics and Chemistry*, North Holland, Amsterdam 1987.

[8] M.E.J. Newman, G.T. Barkema, *Monte Carlo Methods in Statistical Physics*, Clarendon Press, Oxford 1999.

[9] S. Magnusdottir, B. Akerman, M.Jonsson, *J. Phys. Chem.* **98**, 2624 (1994).