ELSEVIER

Journal of Chromatography A, 806 (1998) 123–139

JOURNAL OF CHROMATOGRAPHY A

Kinetic properties of DNA migration under clamped homogeneous electric field conditions
DNA size, migration velocities and reorientation time determined in a single clamped homogeneous electric field run

L. Lopez-Canovas[a], L. Galan[b], E. Orozco[c,d], A.M. Riveron[a,d,*]

[a]*Department of Molecular Biology, Neurosciences Branch, National Center for Scientific Research (CNIC), P.O Box 6990, Havana, Cuba*

[b]*Department of Neurophysics, Neurosciences Branch, National Center for Scientific Research, P.O. Box 6990, Havana, Cuba*

[c]*Department of Experimental Pathology, CINVESTAV-IPN, A.P. 14-740, Mexico 07000 DF, Mexico*

[d]*CICATA-IPN, Molecular Biomedicine Program, Legaria 694, Col Irrigacion, Mexico D.F. 11500, Mexico*

Abstract

By analyzing DNA migrations per pulse at different pulse times we determined the reorientation times and migration velocities, during and after DNA reorientation of molecules separated under different clamped homogeneous electric field (CHEF) electrophoretic conditions. We obtained functions relating these parameters and DNA migration per pulse with running variables and DNA size. A statistical procedure that validates the predictions of these functions was also proposed using identical data for function fitness and validation. A method is proposed to estimate DNA size, reorientation time and migration velocities using the distance migrated by a DNA molecule in a single CHEF experiment. © 1998 Elsevier Science B.V.

Keywords: Clamped homogeneous electric field; DNA reorientation time; Migration velocities; Validation; Kinetic studies; DNA

1. Introduction

DNA molecules are resolved by pulsed field gel electrophoresis (PFGE) in patterns that depend on the system, sample, buffer, gel concentration, electric field, temperature, pulse time and running time [1,2]. Direct observation of DNA by fluorescence microscopy revealed different motion mechanisms when it moves in steady fields and when it reorientates itself in alternating fields [3,4]. This confirms that under any PFGE conditions, the motion mechanism and the velocity of a DNA molecule depend on the elapsed pulse time (t_p) and the molecule reorientation time (t_r) [1,4–6]. Mobility curves have been developed to mathematically describe DNA motion in PFGE as a function of pulse time [2,6]. However, if the elapsed portion of t_p surpasses t_r, DNA migrates as it does in steady field and, if t_p is shorter than t_r the molecule remains reorienting itself. Whereas, when the elapsed t_p is equal to t_r, the migration mechanism changes, and at this point the function describing DNA

*Corresponding author. Corresponding address: Department of Molecular Biology, Neurosciences Branch, National Center for Scientific Research (CNIC), P.O Box 6990, Havana, Cuba.

0021-9673/98/$19.00
PII S0021-9673(97)01122-9

mobility becomes discontinuous. Thus, two functions of pulse duration might be used for a better description of the motion of a DNA molecule during a single pulse, they may be separated at the point where the mobility curve is discontinuous. Functions based on these assumptions have been proposed [5–8] using as parameters the velocities during and after the molecule reorientation and its reorientation time. These parameters are affected by the running conditions and DNA size [8,9].

Considering the above mentioned postulates and functions, here, we obtained new mathematical expressions to accurately describe the migration of DNA molecules under distinct clamped homogeneous electric field (CHEF) experimental conditions. The analysis of DNA migration per pulse obtained at different pulse times and under identical electrophoretic conditions allows the determination of the two migration velocities and the reorientation time of the molecule [5,7]. In this paper, we obtained migration data for DNA molecules (230–1532 kb) separated under different CHEF and miniCHEF running conditions. Recently, we have found that under similar electrophoretic conditions CHEF chambers of different sizes provide equal migration per pulse of DNA molecules [10]. We studied here if (a) the two migration velocities could be described by a modified Lumpkin and Zimm's equation [11], (b) the DNA reorientation time could be described by its contour length and its reorientation velocity [9], and (c) two general linear functions that have as parameters the two migrations velocities and the reorientation time could describe migration per pulse of DNA molecules at any pulse time and electrophoretic conditions. If our functions describe well DNA motion they also would predict well both velocities, reorientation time and migration per pulse. Thus, we carried out the validation of these predictions by developing a statistical procedure that uses the complete set of experimental points for functions fitness and cross validation. The procedure is based on the deletion of each experimental point during the fitness and on its posterior prediction by the function. We also developed a procedure that estimates DNA size, reorientation time and migration velocities using the distance migrated by a DNA molecule in a single CHEF electrophoresis.

2. Theory, assumptions and statistical procedures

2.1. Identification of the variables

Yeast chromosome (YC_1, YC_2,..,YC_n) of contour lengths (L_1, L_2,..,L_n) were separated under different combinations of electric fields (E) and temperature (T). The selected combinations of E and T are shown in Table 2. We used the variable i, which identifies each row_i, to index E, T, η, YC and L as E_i, T_i, η_i, YC_i and L_i. The pulse times (t_p) and the numbers of pulses (N_p), applied to the CHEF experiments (a total of m runs) performed at a given E_i and T_i, were additionally indexed by the integer variable k; which ranges between 1 and m. Thus, for a molecule of contour length L_i, electrophoresed at the electric field E_i, temperature T_i, pulse times t_{pik} for the numbers of pulses N_{pik}, we will have the arrays of migrated distances D_{ik} and the migrations per pulses d_{ik}. The migration velocities during and after DNA reorientation and its reorientation time (obtained at a given E_i and T_i) will be named v_{ri}, v_{mi} and t_{ri}, respectively. The t_p will be also indexed by i because each molecule is resolved in a particular pulse time interval that depends on the used CHEF conditions.

2.2. Model for the migration per pulse of a DNA molecule as two functions of pulse time

Under constant electrophoretic conditions, the migrations per pulse of a DNA molecule [12] (called here d_r during the reorientation and d_m after the reorientation) separated in a selected pulse time interval are described by two linear functions of the pulse time. This description uses as parameters the velocities during and after the molecule reorientation and its reorientation time [5–7,10]. We have for d_r and d_m:

$$d_r = \nu_r t_p \text{ when } t_p <= t_r \quad (1)$$

$$d_m = \nu_m (t_p - t_r) \text{ when } t_p > t_r \quad (2)$$

where the migration per pulse is

$$d = D/N_p \quad (3)$$

2.3. Estimation of DNA reorientation time and migration velocities using migrations per pulse obtained at different pulse times

The parameters v_r, v_m and t_r of a DNA molecule separated under given experimental conditions can be determined by analyzing the relationship between DNA migration per pulse and pulse time [5,7,10]. For a DNA molecule separated by CHEF at different pulse times (a total of m runs) and at a given combination of electric field (E_i) and temperature (T_i) (Table 2, rows), we obtained the arrays of migrations per pulse (d_{ik}) and pulse times (t_{pik}). To determine v_{ri}, v_{mi} and t_{ri}, the arrays (d_{ik}, t_{pik}) were split at a particular t_{pik}, (see below) and fitted to two distinct linear functions of pulse duration (see Figs. 1 and 2,4 to note the two slopes in the data of each molecule). During reorientation we fitted:

$$d_{ik} = (a_0)_i + (a_1)_i t_{pik}$$

for a given i and k ranging from 1 to lim (4)

and after reorientation:

$$d_{ik} = (b_0)_i + (b_1)_i t_{pik}$$

for a given i and k ranging from (lim + 1) to m (5)

Here, a_0 and a_1, and b_0 and b_1 are the regression coefficients of Eqs. (4) and (5), respectively. The parameters v_{ri}, t_{ri} and v_{mi} were taken by analogy between Eqs. (1) and (4) and Eqs. (2) and (5), respectively. Thus, $v_{ri} = (a_1)_i$, $v_{mi} = (b_1)_i$ and $t_{ri} = (-b_0)_i/(b_1)_i$ [5,7]. The experimental pulse time that splits the arrays d_{ik}, t_{pik} to fit both linear equations is called $t_{p(i,lim)}$ and is automatically selected by the computer program as the one that separates Eqs. (4) and (5), rendering a minimum in the sum of their residual errors [7]. Linear fittings were done obtaining the M-estimates of the regression coefficients [13] (called here ϕ_0 the intercept and ϕ_1 the slope) as described in the accompanying paper [10]. This is a robust regression method that avoids the deleterious influence of data outliers in the estimates. For simplicity, in this section we rewrite the array (d_{ik}, t_{pik}) without the index i as (d_k, t_{pk}). M-estimates are the arguments that minimized the summation of the absolute deviations $|E_k|$, then ϕ_0 and ϕ_1 are:

$$(\phi_0, \phi_1) = \arg\min \sum |E_k|;$$

where $E_k = d_k - (\phi_0 + \phi_1 t_{pk})$ and $k = 1,..,n$

The number of ordinate pairs to fit the regression equation is n. The coefficient variances $S^2\phi_0$ and $S^2\phi_1$ were also calculated considering

$$S^2\phi_j = t_{pjj}^{-1}/\{4[f(0)]^2\} \text{ for } j = 0, 1$$

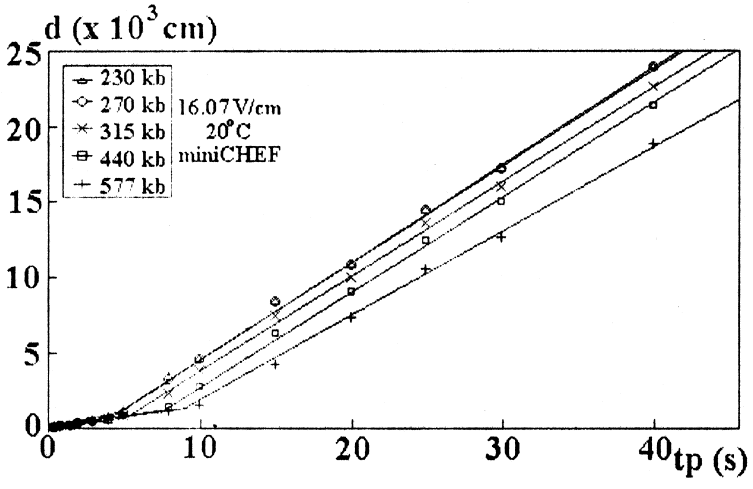

Fig. 1. Experimental migrations per pulses (d_{ik}) of distinct sized DNA molecules plotted as two linear functions of pulse time (t_{pik}).

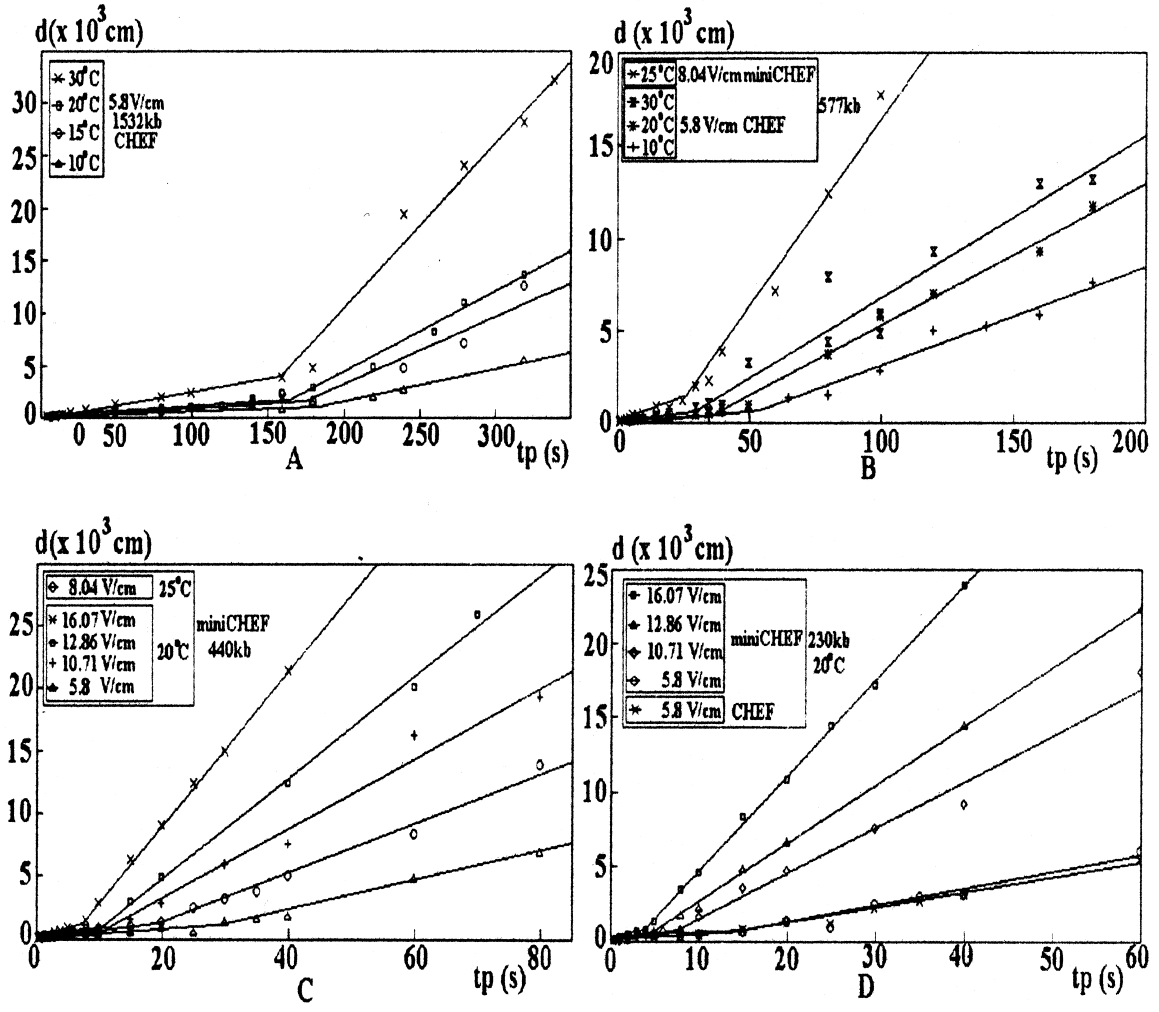

Fig. 2. Experimental migrations per pulses (d_{ik}) against pulse times (t_{pik}) of DNA molecules separated under different combinations of temperatures and electric fields. Molecules were electrophoresed (A and B) at constant electric field and different temperatures, and (C and D) at constant temperature and different electric fields. The graphics include data collected in CHEF and miniCHEF chambers.

where t^{-1}_{pjj} are the diagonal elements of the inverse of the design matrix; and f(0) depends on the error distribution function [10]. The calculations were done by programming iterative algorithms which used least square estimates as initial values [13].

Using the above described procedure we determined the v_{ri}, v_{mi} and t_{ri} for *Saccharomyces cerevisiae* (196-2, α, his, a kind gift of M. Luzzati, France) chromosomes separated under different combinations of electric field and temperature (Table 2) and we studied the dependence of v_r, v_m and t_r on running variables and DNA sizes.

2.4. Assumptions to obtain the expressions for DNA migration velocities and reorientation time

Based on the reptation theory, Lumpkin and Zimm [11] developed an equation describing DNA mobility in gel electrophoresis. According to the theory a tube is the trajectory that the leading end of the chain has chosen in the agarose gel and that the rest of the chain must follow. The model of tube-like-motion has been considered inadequate for describing gel electrophoresis of DNA [14,15]. This consideration was supported by the results obtained by direct

observation of DNA movement in the microscopy [3,16] and by studying DNA orientation using linear dichroism (LD) spectroscopy [14]. These studies revealed that the motion of DNA molecules in the gel displays an oscillation cycle of conformational changes characterized by an extension and a contraction phase. Such conformational changes, as the U conformation and the formation of kinks, are not considered in the original reptation model [14]. Recently, Akerman [17] analyzed again the available data from LD spectroscopy and microscopy studies, and considering the cyclic nature of DNA migration, he suggested that the fundamental reptation equation for DNA electrophoretic mobility also applies in the presence of strong fluctuations in the tube length [17]. It is known that the variables describing the velocities of different DNA molecules in PFGE are: the electric field (E), the particle net charge (Q), the DNA contour length ($L=0.34$ nm$\times$bp; where bp is the number of base pairs), and the buffer viscosity (η) that depends on the experimental temperature. The Lumpkin and Zimm equation gives a qualitatively correct description of the dependence of mobility on field strength, gel concentration and DNA length [14]; thus, we used it for describing the macroscopic velocities (v_r and v_m) of DNA molecules in CHEF electrophoresis. We fitted for v (v_r and v_m) a modified equation for the steady state velocity of DNA during conventional agarose gel electrophoresis [11]

$$v = \beta Q E^x / f \tag{6}$$

In Eq. (6), β is a proportionality constant and f is the friction coefficient. It was taken here as $f=8\pi\eta L^y$, where L^y resumes the size information of the DNA molecule. According to previous reports [6,8,18], the exponents x and y should be greater than one. We expected that E and L would have different exponents x and y in v_r and v_m expressions. Then, for v_r the exponents will be called x_r and y_r and for v_m x_m and y_m. In PFGE, DNA molecules migrate in zigzag, and along the diagonal the velocity is reduced after the reorientation by a factor of $\cos(\varphi/2)$, and during reorientation by a factor of $\cos^3(\varphi/2)$, where φ is the angle between the field force lines [8]. As β includes these factors we will use β_r and β_m for v_r and v_m expressions, respectively.

DNA molecules of distinct sizes separated at different electric fields and temperatures spend distinct times to align with the new field direction (t_r) [1–8]. The effect of running variables on t_r was proposed to be contained in the v_r expression [9]. Thus, if during t_r the molecule of contour length L passed from an agarose path to another in the new field direction at given v_r

$$t_r = \gamma L^s / v_r \tag{7}$$

where γ is a proportionality constant, and L^s indicates the apparent length of DNA molecules during the reorientation. We expect $s\neq 1$, as reported previously for the dependence of t_r on DNA size [4,6].

2.5. Processing data to fit the equations describing v_m, v_r and t_r

To fit the equations proposed here we used the reported sequenced sizes (kb) of yeast chromosomes [19]. The initial estimates for the exponents x_r, y_r, x_m and y_m were obtained by using a logarithmic transformation of Eq. (6) and fitting the following multivariate linear functions

$$\log v_m = c_0 + c_1 \log E + c_2 \log L \tag{8}$$

$$\log v_r = c_0' + c_1' \log E + c_2' \log L \tag{9}$$

We did not use log Q because it contains size information and gave correlation among the independent variables. To fit Eqs. (8) and (9) we used the arrays of the experimentally determined v_{mi} and v_{ri} as the dependent variables, and Q_i, L_i, E_i and η_i as the independent ones. The slopes (c_j and c_j') of Eqs. (8) and (9) were our initial exponent estimates ($x_m=c_1$; $y_m=c_2-1$ and $x_r=c_1'$; $y_r=c_2'-1$). We subtracted 1 from c_2 and c_2' after fitting Eqs. (8) and (9) to compensate for the presence of L in the variable Q (Eq. (6)). Finally, we fitted the experimental parameters v_{ri} and v_{mi} as linear functions of the variables h_{ri} and h_{mi}, where $h_{ri}=Q_i E_i^{xr}/[8\pi\eta_i L_i^{yr}]$, and $h_{mi}=Q_i E_i^{xm}/[8\pi\eta_i L_i^{ym}]$, respectively. To obtain the final values of x_m, y_m, x_r and y_r in Eq. (6) we iterated around these initial exponent estimates. The iterations ended when v_m and v_r expressions gave a minimum in the residual error.

The function for describing tr was obtained by

fitting the logarithms of t_{ri} as linearly depending on the logarithms of (L_i^s/v_{ri}). According to previous reports [4,6] we started the iterations with the initial value of $s=1.17$. The iterations around the exponent s ended when the expression gave a minimum in the residual error.

2.6. Unit system

To fit equations all physical magnitudes were expressed in the cgs system of units. That is, E in statvolt/cm; L in cm; Q in statcoulomb; η in cP, v_m and v_r in cm/s, t_p and t_r in s.

2.7. Statistical cross validation of functions predictions

If an experimental datum is used to fit a regression function, it can not be used to validate the prediction of the function for its value. To do it, the datum has to be excluded from the original fit. Considering this, we fitted the studied function without a single experimental point each time. Further, we predicted the deleted point and its confidence interval, and finally, we determined if the deleted experimental point is inside this interval. If we have a prediction efficiency greater than 90% we accept that the function describes well the relationship among the variables. That is, we have an array of n experimental pairs of points $[r_1,z_1]$, $[r_2,z_2]$,... $[r_n,z_n]$ where j indexes the array and varies from 1 to n, and the pair r,z represents two any related variables. We need to fit the function $z=c_0+c_1r$ and validate its predictions for all r values. If from the original array we delete a pair $[r_i,z_i]$, where i is comprised between 1 and n, and use the remaining $n-1$ data to fit the equation $z_{(-i)}=c_{0(-i)}+c_{1(-i)}r_{(-i)}$, we can replace the deleted r_i in this equation and obtain the corresponding prediction $z_{(-i)}=c_{0(-i)}+c_{1(-i)}r_i$. If the 0.975 percentile confidence interval of $z_{(-i)}$ is also predicted, we can determine if the corresponding experimental z_i falls into $z_{(-i)}\pm$CL (CL: confidence limits), and know the concordance between the experimental and predicted values. This procedure was repeated with all points i between 1 and n. The index $-i$ in the coefficients and predictions means that they were obtained without the point of the ordinal i in the array $[r_j,z_j]$.

3. Experimental

3.1. Samples

Agarose plugs containing immobilized *S. cerevisiae* DNA were prepared as described [20,21]. The plugs had the same cell concentration ($6.25\cdot10^9$ cells/ml) and the samples were prepared using the 4 *M* urea non-enzymatic procedure that renders deproteinized DNA [21]. The plugs were cast in special molds that gave samples of $0.25\times0.4\times0.1$ cm (length×width×thickness) for miniCHEF and $0.85\times0.85\times0.2$ cm for CHEF.

3.2. CHEF experiments

CHEF and miniCHEF chambers provide equal migrations per pulse of each DNA molecule at the same driving forces and pulse times, but the miniCHEF resolves the molecules in a shorter time [10]. Thus, yeast chromosomes were separated in CHEF (28.4 cm of electrode separation) [22,23] and in our miniCHEF (11.2 cm of electrode separation) laboratory-made chambers [20]. The pulse time and the number of pulses were set and checked by an automatic controlled switching unit [24]. All gels (length×width×thickness: $3.9\times3.9\times0.5$ cm for miniCHEF and $10.4\times10.4\times1$ cm for CHEF) were cast at 1.5% agarose (Lachema, Czech Republic) concentration in 0.5×TBE electrophoretic buffer (1×TBE: 89 m*M* Tris, 89 m*M* boric acid, 2 m*M* EDTA, pH 8.3). Before casting the gels, the volume of the buffer plus the agarose was measured, the mix was heated and the evaporated water was replaced. The gel was formed at room temperature. Buffer height on the gels during the electrophoresis was 0.9 cm in the CHEF and 0.5 cm in the miniCHEF. Each experimental set of electrophoretic runnings was done under identical conditions but at different pulse times. We performed four sets of CHEF runnings at 10, 15, 20 and 30°C and 5.8 V/cm. Meanwhile, in miniCHEF, we varied the electric field (5.8, 10.71, 12.86 and 16.07 V/cm) and the running temperature was maintained at 20°C. Another set of experiments was done in miniCHEF at 8.04 V/cm and 25°C (the pulse time, electric field and temperature are shown in Figs. 1 and 2,4 Table 2). The electric field was estimated by dividing the voltage read from the

power supply by the distance between the electrodes. The initial buffer ionic strength was checked by measuring the conductivity at 30°C, giving an average value of 959 μS. The running temperature, buffer conductivity and pH were maintained constant using a peristaltic pump to recirculate the buffer (7 l for CHEF and 1.5 l for miniCHEF) through a heat exchanger that was connected to a regulated temperature bath (Beckman Instruments, CA, USA). The variation of the temperature during each experiment was less than ±0.2°C. Buffer viscosity (η) was approached using published data of water viscosity (in cP) in the temperature (T) range of 5–45°C ($n=9$) [25]. We fitted viscosity data to the polynomy:

$$\eta_{H_2O} = g_0 + g_1 T + g_2 T^2 + g_3 T^3 + g_4 T^4$$

and we obtained the regression coefficients (g_j) and variances ($S^2 g_j$) that are shown in Table 1. Each η was calculated by replacing the corresponding experimental temperature in °C in the above expression. The gels were stained in ethidium bromide solution (0.5 μg/ml) during 1 h, photographed, blotted and hybridized [26] with [α^{32}P] dATP radiolabeled ura3 and trp1 gene probes (markers for chromosomes V and IV of *S. cerevisiae*, respectively). The chromosomes I, VI, III, IX, V and IV of yeast were identified in the photographs by comparing their migrations with those of λ-mers and with the positions of chromosomes V and IV in the autoradiographies.

3.3. Calculations of DNA migrations per pulse

Migrated distances (D_{ik}) were measured directly from the photography negatives of the CHEF gels with the aid of a digitizer tablet (Apple II, Graphic Tablet). Each measurement was carried out at least three times with a mean error of 0.001 cm. They were corrected by the photography magnification factor, and were divided by the corresponding number of pulses (N_{pik}) to obtain the migrations per pulse (d_{ik}). The measurements were done exclusively on gels where the bands were unequivocally identified. Overloaded gels or those having smearing tracks were discarded.

Table 1
Coefficients and variances of the polynomial function relating water viscosity and the experimental temperature

Subindex (j) of coefficients g_j	Coefficients (g_j)	Variances ($S^2 g_j$)
0	$1.7844 \cdot 10^{-2}$	$2.122 \cdot 10^{-7}$
1	$-5.9388 \cdot 10^{-4}$	$3.134 \cdot 10^{-11}$
2	$1.3494 \cdot 10^{-5}$	$2.180 \cdot 10^{-13}$
3	$-1.9278 \cdot 10^{-7}$	$2.303 \cdot 10^{-16}$
4	$1.2455 \cdot 10 - 9$	$2.817 \cdot 10^{-20}$

3.4. Software and the regression method to fit equations

The straight lines were fitted using the robust regression method described in the accompanying article [10]. Laboratory-written and public domain software were used to process the data. laboratory-written programs were written in Turbo-Pascal, version 7.0, running on MS-DOS operating system for an AcerMate 466d IBM compatible microcomputer.

4. Results and discussion

4.1. Experimental v_r, v_m and t_r of S. cerevisiae chromosomes under different electric fields and temperatures

We previously demonstrated that the migrations per pulse of a molecule electrophoresed at similar conditions in CHEF and miniCHEF are equal independently of the chamber dimension [10]. This permitted one to fit our equations on velocities, migration per pulse and reorientation time using the migration data obtained in CHEF experiments performed in chambers of different sizes.

The array of migrations per pulse of a DNA molecule (d_{ik}) separated under given experimental conditions and different pulse times, was fitted to two different linear functions of pulse duration (t_{pik}) as described [10] (see, Section 2). The experimental migration data and the straight lines for $d_{ik}=(a_0)_i+(a_1)_i t_{pik}$, for $t_{pik} \leq t_{p(i,\lim)}$, and $d_{ik}=(b_0)_i+(b_1)_i t_{pik}$, for $t_{pik} > t_{p(i,\lim)}$ (Eqs. (4) and (5), respectively) are shown in Figs. 1 and 2. In all trials $(a_0)_i$ was zero, whereas $(b_0)_i$ and the slopes $(a_1)_i$ and $(b_1)_i$ differed

from zero to a significant level of 0.05 [the values of $(a_0)_i$, $(a_1)_i$, $(b_0)_i$ and $(b_1)_i$ can be calculated using the data presented in Table 2]. For any molecule, and independently of the experimental conditions, we always obtained that the slope after the reorientation, $(b_1)_i$, was greater than the slope during the reorientation, $(a_1)_i$, (Figs. 1 and 2), evidencing that during reorientation DNA migrated slower. This slower movement appreciated at the macroscopic level may be explained based on the microscopic behavior of DNA molecules. Gurrieri et al. [27] observed the movement of stained DNA [27]. Using a 120° reorientation angle, they found that DNA molecules usually reverse their path after the application of a new pulse; and what was the tail of the molecule during the previous pulse, becomes the head and turns to the direction of the new field [27].

According to the relationships among $(a_1)_i$, $(b_0)_i$ and $(b_1)_i$ with v_{ri}, t_{ri}, v_{mi} [5,7], we determined the velocities and reorientation times for *S. cerevisiae* chromosomes separated under the CHEF conditions showed in Figs. 1 and 2 Table 2. At any selected combination of electric field and temperature, the v_r and v_m of the chromosome sized molecules slightly

Table 2
Reorientation times (t_{ri}) and migration velocities (v_{ri} and v_{mi}) of different sized DNA molecules determined from experiments in CHEF[a] and miniCHEF under distinct running conditions

Row (i)	Size (kb)	E (V/cm)	T (°C)	v_{mi} ($\cdot 10^5$ cm/s)	v_{ri} ($\cdot 10^5$ cm/s)	t_{ri} (s)
1	230	5.80	20.0	11.99	4.09	9.81
2	230	10.71	20.0	31.06	9.37	4.93
3	230	12.86	20.0	39.47	10.70	3.38
4	230	16.07	20.0	62.97	19.28	2.68
5	270	5.80	20.0	12.18	4.09	10.57
6	270	10.71	20.0	32.23	7.38	5.76
7	270	12.86	20.0	39.81	10.70	3.69
8	270	16.07	20.0	63.35	16.63	2.84
9	315	5.80	20.0	11.54	4.35	14.35
10	315	16.07	20.0	61.94	16.63	3.80
11	440	5.80	20.0	11.56	4.27	20.05
12	440	10.71	20.0	26.81	7.38	7.77
13	440	12.86	20.0	38.17	9.73	7.29
14	440	16.07	20.0	62.81	17.25	5.54
15	577	5.80	20.0	10.32	3.32	30.17
16	577	10.71	20.0	27.21	6.79	15.61
17	577	12.86	20.0	42.12	8.74	13.92
18	230	8.04	25.0	17.31	7.55	4.92
19	270	8.04	25.0	17.85	6.73	6.01
20	315	8.04	25.0	17.74	6.44	8.24
21	440	8.04	25.0	18.61	5.65	12.84
22	577	8.04	25.0	17.86	5.65	18.16
23	577	5.80[a]	10.0	5.30	1.57	38.81
24	577	5.80[a]	15.0	7.30	1.41	32.90
25	577	5.80[a]	30.0	9.76	3.15	24.74
26	1532	5.80[a]	10.0	3.48	0.67	160.27
27	1532	5.80[a]	15.0	5.56	1.99	150.61
28	1532	5.80[a]	20.0	7.75	1.06	142.46
29	1532	5.80[a]	30.0	15.66	2.50	134.77
30	230	5.80[a]	20.0	10.62	4.24	8.29
31	270	5.80[a]	20.0	10.90	3.61	10.59
32	315	5.80[a]	20.0	10.83	3.61	14.63
33	440	5.80[a]	20.0	10.56	4.09	18.82
34	577	5.80[a]	20.0	10.31	3.74	28.16

Examples of the original migration data are shown in Figs. 1 and 2,4.

depended on the DNA size, and sometimes DNA size and velocities seemed to be unrelated (Table 2 Fig. 1), although v_r had a higher dependence on DNA size. Increments of the temperature (Fig. 2A Fig. 2B), and electric field (Fig. 2C Fig. 2D), or of both simultaneously (Fig. 2B Fig. 2C) augmented v_r and v_m (Table 2). Variations of the electric field from 8.04 to 10.71 V/cm increased the v_r and v_m of the molecules more than increasing the temperature from 20 to 25°C (Table 2).

In the plots d_{ik} against t_{pik} (Figs. 1 and 2), t_r corresponds to the t_p value where the straight line, fitted after the molecule reorientation, would cross the t_p axis. For yeast chromosomes, I, VI, III, IX, V and IV, the crossing points were at different abscissa values (Figs. 1 and 2), and depended on the DNA size (Fig. 1) and on the applied CHEF running conditions (Fig. 2A Fig. 2B Fig. 2C Fig. 2D). As it was previously established [1,2,4,5,8,9,16,17,28], t_r was larger for DNA molecules of larger sizes, whereas for a given DNA size, it decreases at higher electric fields, or higher temperatures, or when both running variables are increased (Table 2). We agree with previous reports [4,16,17,28–30] that the velocities (v_r and v_m here) sharply depend on the buffer viscosity and the electric field and slightly from DNA size; whereas t_r is related to all these variables [1,2,4,8,9,16,17,28].

4.2. Quantitative relations between v_m, v_r and t_r and DNA size and running variables

The migration velocity during a single oscillating cycle has been reported as independent of DNA size [17,28]. However, these two variables are unrelated up to a possible logarithmic transformation of the data [28]. The results of our experiments partially agree with these reports. The migration velocities slightly depended on DNA size (Table 2); and to fit v_r and v_m equations (Eq. (6)) it was impossible to exclude the factor L^y, even when the variable Q was removed. Initial estimates of x_m, y_m, x_r and y_r for v_m and v_r (Eq. (6)) were obtained from the multivariate regression analyses of the logarithmic transformed variables (Eqs. (8) and (9)). After fitting, the coefficients c_j and c'_j significantly differed from zero (significant level of 0.05). They and the variances S^2c_j are shown in Table 3. Thus, the initial estimates

Table 3
Variances (S^2c_j and $S^2c'_j$) and multiple regression coefficients (c_j and c'_j) for Eqs. (8) and (9)

Subindex (j) of the coefficients	Coefficients (c_j)	Variances (S^2c_j)
Eq. (8): $\log v_m = c_0 + c_1 \log E + c_2 \log L$		
0	−3.38	$2.82\cdot10^{-1}$
1	1.74	$1.35\cdot10^{-2}$
2	−0.24	$5.90\cdot10^{-3}$
Eq. (9): $\log v_r = c'_0 + c'_1 \log E + c'_2 \log L$		
0	−7.13	$4.19\cdot10^{-1}$
1	1.43	$1.99\cdot10^{-2}$
2	−0.57	$8.79\cdot10^{-3}$

of the exponents of Eq. (6) for v_r and v_m were $x_r = 1.43$, $y_r = -0.57 - 1.0$; $x_m = 1.74$, $y_m = -0.235 - 1.0$. Using these exponents we fitted v_{ri} and v_{mi} as linear functions of the new independent variables h_{ri} and h_{mi}, where $h_{ri} = [Q_i E_i^{1.43}/[8\pi\eta_i L_i^{1.57}]$, and $h_{mi} = Q_i E_i^{1.74}/[8\pi\eta_i L_i^{1.24}]$. After the iterations around these initial exponent estimates, we obtained for v_m a straight line (Fig. 3A) with a non-significant intercept of $-6.65\cdot10^{-6}$ (variance = $3.45\cdot10^{-11}$), a significant slope $\beta_m = 0.665$ ($S^2\beta_m = 1.74\cdot10^{-4}$) and exponents $x_m = 1.76$ and $y_m = 1.08$. A statistical difference between β_m and cos (60°) was obtained to a significant level of 0.05. The v_r presented a non-significant intercept of $-2.52\cdot10^{-6}$ (variance = $8.98\cdot10^{-12}$), a significant slope $\beta_r = 0.0207$ ($S^2\beta_r = 5.58\cdot10^{-7}$), and exponents $x_r = 1.45$ and $y_r = 1.35$ (Fig. 3B). β_r also differed from $\cos^3$ (60°). The final expressions of v_m and v_r were:

$$\nu_m = 0.665\left[QE^{1.76}/(8\pi\eta L^{1.08})\right] \quad (10)$$

$$\nu_r = 0.0207\left[QE^{1.45}/(8\pi\eta L^{1.35})\right] \quad (11)$$

As we expected, we obtained for v_m and v_r (Eqs. (10) and (11)) different exponents x_m and x_r and y_m and y_r. Microscopy studies of the behaviour of DNA migration demonstrated that during the reorientation kinks may appear in the molecules [3,27]. They, and the current leading DNA segment, are simultaneously pulled by the field [28], and E could act as driving and retarding forces. These kinks and the reversion of the movement of DNA molecules during the reorientation [3,27] should be macroscopically ap-

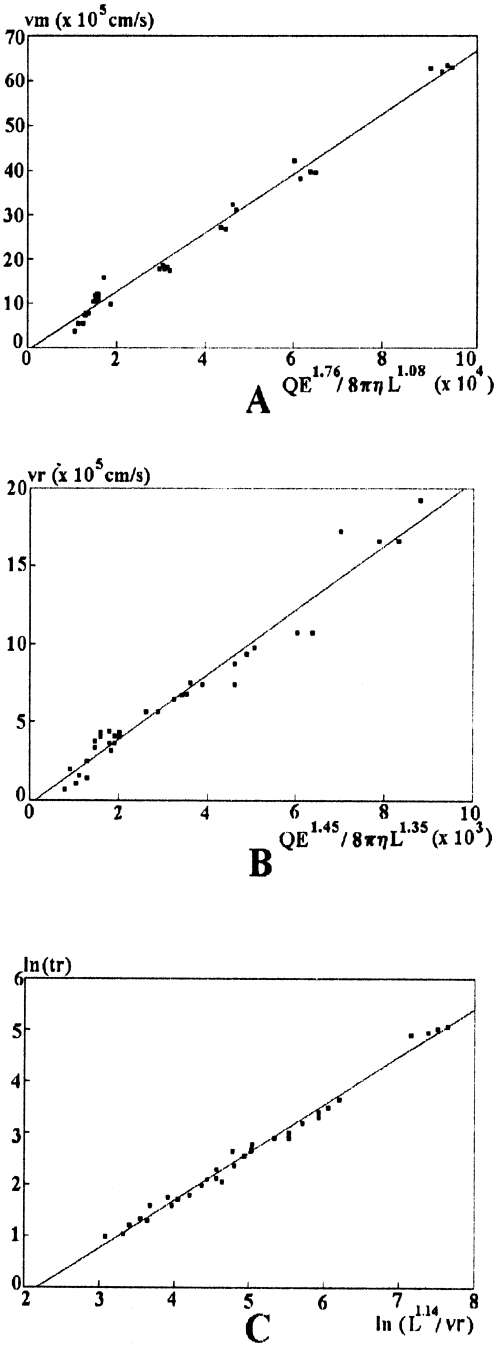

Fig. 3. Linear fittings of the parameters v_m (A), v_r (B) and the logarithm of t_r (C) as functions of the transformations of running variables and DNA size. The transformations are shown in the horizontal axis.

preciated as different average apparent lengths of DNA molecules during the reorientation and after the reorientation.

The differences between β_m and cos (60°) and β_r and $\cos^3$ (60°) could be due to the fact that the net charge of the DNA molecules was calculated as the formal charge (2bp1.0e^-, where e^- is the electron charge, and is $4.806 \cdot 10^{-10}$; bp: base pairs) and that the effect of the ions of the buffer on the net charge [31,32] might be included in β_r and β_m as an average scaling factor. β_m and β_r should be also affected by another scaling factor related to the gel concentration (C_g). We estimated its value using previously published data [16] relating DNA velocity and gel concentration (in %), obtaining that the scaling factor is roughly $1/0.2C_g$. Correcting β_m and β_r for this factor and for cos (60°) and $\cos^3$ (60°), the electrophoretic DNA charge approached 0.4 and 0.05, respectively, which are similar to those reported for free and gel electrophoresis [31,32]. The results achieved with these punctual estimations suggest that the electrophoretic DNA charge may be estimated from kinetic migration data by analyzing the β_m and β_r obtained from the appropriate CHEF experiments. The use of 1.0e^- to obtain Eqs. (10) and (11) simplifies these estimations.

Using the arrays t_{ri}, L_i, v_{ri} the parameter t_r was described by a function of DNA contour length and v_r (Fig. 3C). We obtained $\gamma = 0.134$, $s = 1.14$, and the ratio $(L^{1.14}/v_r)^{0.926}$. Thus,

$$t_r = 0.134\left(L^{1.14}/v_r\right)^{0.926} \quad (12)$$

According to LD spectroscopy and microscopy results, the relation $<L>/<v>$ was previously found useful to describe $<T>$, or the time needed to reach the undershoot in LD spectroscopy, considering that the averages $<L>$ and $<T>$ varied proportionally during a DNA oscillating cycle [17]. However, our macroscopic description (Eq. (12)) slightly differed from this proposition that contains L/v [9,17]. In agreement with our results, Oana et al. [16] found that the characteristic time (τ) of the DNA oscillating cycle, determined by microscopy studies, does not fit well to L/v. Thus, in Section 4.3 we will analyze if Eq. (12) describes well the relationship between the reorientation times obtained in CHEF, the running variables and the DNA sizes.

The parameter t_r may correspond to the time required to accomplish a complete oscillating cycle. The parameters t_r, τ and $<T>$ were similarly related to the electric field; τ and $<T>$ depend on $E^{-1.1}$ and $E^{-1.2\pm0.1}$, respectively [16,17]; whereas t_r depended on $E^{-1.34}$. However, Carlsson and Jonsson [33] recently found that in 1% agarose gels, the time needed to reach the overshoot in LD studies (t_p) was proportional to $E^{-1.26}$, whereas the time needed to reach the undershoot (t_u) was proportional to $E^{-1.21}$. They are interpreted as the time required by DNA to complete the extension phase, or the oscillating cycle, respectively. Thus, we can not rule out that t_r corresponds to the time required to accomplish the extension phase. Up to now, our calculation procedure for t_r, v_r and v_m does not detect the fine changes reported in microscopy and LD studies. Therefore, the observed similarities are not sufficient information to determine a reliable correspondence among these parameters.

We should add that in macroscopic studies of DNA migration, the reorientation time was reported to depend proportionally on $kb^{1.17}$ [4,6]. Here, operating with the exponents of L and Q in Eq. (12) as $(y_r+s-1)0.93$, we calculated that t_r depends on $L^{1.38}$, which has a similar dependence. Our results are also in good agreement with other earlier reports where $t_r \propto E^{-1.4}$, and η^1. [34,4]. The v_m equation contains the electric field raised to 1.76, which is close to the exponent reported for E in conventional electrophoresis [18]. If we take v_m as DNA velocity in the steady state, it is easy to see that v_r and t_r can be expressed as functions of v_m, as it has been proposed [8]. However, our data also show that v_r is not a simple linear reduction of v_m.

4.3. Validation of functions predictions

We validated the predictions of the equations that describe v_m, v_r and t_r (Eqs. (10)–(12), respectively) with data obtained from experimental CHEF runs. Eq. (10) $(-i)$, Eq. (11) $(-i)$ and Eq. (12) $(-i)$ were fitted as described in Section 2. The deleted variables $E_{(-i)}$, $\eta_{(-i)}$, $L_{(-i)}$ of the row$_i$ (Table 2) were replaced in the equations to obtain the confidence intervals and the respective predictions, $v_{m(-i)}$, $v_{r(-i)}$ and $t_{r(-i)}$ (Table 4). Most of the experimental parameters are inside the corresponding confidence interval of the predictions (Tables 2 and 4). Only the experimental parameters, v_{m29}, t_{r29} (for chromosome IV separated at 5.8 V/cm, 30°C), v_{r14} (for chromosome IX separated at 16.05 V/cm, 20°C), t_{r12} (for chromosome IX separated at 10.71 V/cm, 20°C), v_{r6} (for chromosome VI separated at 10.71 V/cm, 20°C) and v_{r3} (for chromosome I separated at 16.07 V/cm, 20°C) were out of their corresponding confidence intervals of the predictions (Tables 2 and 4). These results presented a general prediction efficiency (or agreement between experimental and predicted values) of 94%. For v_m it was 97%, for v_r 91%, and for t_r 94%; and only a single row (the molecule of 1532 kb separated at 5.8V/cm and 30°C) had two outliers representing the 3% of all studied molecules and experimental conditions (Table 4). This demonstrated that Eqs. (10)–(12) (Fig. 3A Fig. 3B Fig. 3C) predict tolerably well our experimental v_{mi}, v_{ri} and t_{ri}.

Data outliers are commonly found in experimentation. Despite the fact that we used a robust regression method to avoid their deleterious influence on t_r, v_r and v_m estimates, we are aware that our procedure or functions could fail in particular situations. Thus, we used Eq. (12) to predict DNA reorientation times obtained by other authors, for different experimental conditions and using different estimation procedures [16,35,36]. One of these procedures renders a parameter named T^* that approaches the molecule reorientation time. Eq. (12) predicted for conditions of 4 V/cm and 14°C, a t_r of 2.03 s for λDNA, of 11.4 s for a 170 kb molecule, of 75.8 s for a 670 kb molecule and of 152.2 s for a 1110 kb molecule. Under the above conditions and using 1% of agarose concentration, Hutson et al. [35] obtained T^* values of 1.9, 13, 72 and 150 s for each one of the mentioned molecules, respectively. The other procedure gives the parameter τ, which is determined from data obtained observing the movement of stained DNA in the microscope [16]. For phage T4 DNA, electrophoresed at 8.0 V/cm, 20°C in 1.5% of agarose gel, Eq. (12) predicted $t_r=3.7$ s. In the experiment τ was 3.4 s [16]. Thus, Eq. (12) also predicted the reorientation times calculated by other authors using other procedures and other experimental conditions.

Two general approaches have been widely used for the validation of the function predictions: (a) the

Table 4
Predictions for $v_{m(-i)}$, $v_{r(-i)}$ and $t_{r(-i)}$ of Eq. (10) $(-i)$, Eq. (11) $(-i)$ and Eq. (12) $(-i)$, respectively, each fitted without the data of row$_i$ of Table 1

Predicted row $(-i)$	$v_{m(-i)}\pm$CI ($\cdot 10^5$ cm/s)	$v_{r(-i)}\pm$CI ($\cdot 10^5$ cm/s)	$t_{r(-i)}$ (s)	Confidence limits	
				$t_{r(-i)}-$CI	$t_{r(-i)}+$CI
1	10.63±4.30	4.16±2.08	9.11	7.12	11.66
2	31.18±4.34	10.13±2.08	3.95	3.11	5.02
3	43.28±4.26	13.44±1.94·	3.11	2.42	4.01
4	63.73±4.70	17.73±2.20	2.29	1.78	2.94
5	10.50±4.28	3.93±2.08	11.42	8.93	14.60
6	30.75±4.27	9.62±1.95·	4.96	3.88	6.33
7	42.65±4.34	12.61±2.03	3.93	3.05	5.05
8	62.59±4.66	17.30±2.24	2.91	2.25	3.75
9	10.36±4.31	3.73±2.06	14.08	10.99	18.04
10	62.00±4.66	16.16±2.20	3.57	2.78	4.59
11	10.09±4.29	3.33±2.04	22.43	17.56	28.66
12	29.65±4.27	8.06±2.07	9.88·	7.84	12.46
13	41.00±4.33	10.53±2.08	7.69	6.00	9.85
14	59.76±4.45	14.00±1.77·	5.70	4.44	7.31
15	9.86±4.35	3.02±2.08	32.60	25.43	41.79
16	29.00±4.32	7.32±2.07	14.22	11.12	18.17
17	39.81±4.25	9.57±2.07	11.07	8.77	13.98
18	21.17±4.16	7.50±2.07	5.30	4.13	6.80
19	20.91±4.24	7.09±2.07	6.61	5.17	8.46
20	20.65±4.25	6.72±2.07	8.14	6.35	10.44
21	20.11±4.33	5.98±2.07	12.92	10.09	16.55
22	19.68±4.32	5.44±2.07	18.80	14.68	24.08
23	7.53±4.35	2.30±2.09	41.73	32.50	53.58
24	8.66±4.37	2.63±2.06	36.77	28.71	47.08
25	12.35±4.30	3.77±2.08	26.36	20.57	33.77
26	6.95±4.25	1.63±2.09	159.74	122.21	208.80
27	7.99±4.33	1.89±2.10	138.79	106.60	180.69
28	9.10±4.37	2.13±2.08	122.16	94.43	158.05
29	11.51±3.95·	2.69±2.10	97.07·	77.04	122.32
30	10.61±4.36	4.16±2.08	9.16	7.17	11.71
31	10.48±4.35	3.92±2.08	11.42	8.93	14.60
32	10.35±4.35	3.72±2.09	14.07	10.99	18.02
33	10.08±4.35	3.33±2.05	22.49	17.69	28.58
34	9.86±4.35	3.03±2.06	32.71	25.63	41.74

Symbol · tags statistical differences between experimental and predicted values in the cross-validation. CI is the 95% confidence interval.

experimental data are split in two sets, and one set is used for function fitness, and the other to compare the experimental values with the function predictions. This approach usually deteriorates the fitness of the functions, requests more data, and consumes additional time and resources. (b) The parameters and migrations are determined by a different experimental procedure, for example, using LD spectroscopy [6,17,33] or direct observation by microscopy of the movement of stained DNA in conditions similar to those applied in PFGE [3,16,27,35,37]. However, the migrations of stained and unstained DNA differ, and identical experimental conditions have to be set to carry out accurate comparisons. The validation procedure that we proposed here is simple, save time, reagents and efforts and does not deteriorate function fitness.

4.4. Migrations per pulse of all molecules at different running conditions as two functions of t_p, v_r, v_m and t_r

We separated the arrays of migration data d_{ik} in two sets. Any d_{ik} obtained at t_{pik} less than t_{ri} belongs

to a single set (that contains d_{rik}), and any d_{ik} obtained at t_{pik} greater than t_{ri} belongs to another single set (that contains d_{mik}). Further, using the set d_{rik} we fitted:

$$d_{rik} = \nu_{ri} t_{pik} \quad \text{for } t_{pik} \leq t_{ri} \tag{13A}$$

and using the other set we fitted

$$d_{mik} = \nu_{mi}(t_{pik} - t_{ri}) \quad \text{for } t_{pik} > t_{ri} \tag{13B}$$

where $\nu_{ri} = (\beta r Q_i E_i^{1.45}/[8\pi\eta_i L_i^{1.35}]$; $t_{ri} = \gamma(L_i^{1.14}/\nu_{ri})^{0.926}$ and $\nu_{mi} = \beta_m Q_i E_i^{1.76}/[8\pi\eta_i L_i^{1.08}]$.

Using the robust regression method described in the accompanying paper [10] we fitted Eq. (13A), taking $\nu_{ri} t_{pik}$ as the independent variable. For 194 pairs of points and to a significant level of 0.05 we obtained the intercept equal to zero ($p_0 = 2.88 \cdot 10^{-5}$, $S^2 p_0 = 1.23 \cdot 10^{-7}$) and the slope statistically equal to one ($p_1 = 1.02$, $S^2 p_1 = 4.54 \cdot 10^{-4}$). Eq. (13B) was obtained with 296 pairs of points by fitting d_{mik} as a robust linear function of $\nu_{mi}(t_{pik} - t_{ri})$. To a significant level of 0.05 the intercept was equal to zero ($p_0' = 1.36 \cdot 10^{-4}$, $S^2 p_0' = 3.81 \cdot 10^{-7}$) and the slope was statistically equal to one ($p_1' = 0.99$, $S^2 p_1' = 4.13 \cdot 10^{-5}$). We also validated if Eqs. (13A) and (13B) predicted the migrations per pulse of the molecules. When we deleted a row$_i$ of Table 2 we also eliminated from the data the arrays d_{ik}, t_{pik} of the particular molecule separated under the conditions shown in the row$_i$. Thus, replacing the $t_{r(-i)}$, $\nu_{r(-i)}$ and $\nu_{m(-i)}$ estimates (Table 4) in Eqs. (13A) and (13B), we predicted the array of migrations per pulse $d_{(-ik)}$ that each yeast chromosome should have at the assayed pulse times $t_{p(-ik)}$ and the conditions given in the row$_i$ of Table 2. The d_{ik}, obtained in CHEF or miniCHEF experiments at different t_{pik} were always similar to the corresponding predictions $d_{(-ik)}$ of Eq. (13A) $(-i)$ and Eq. (13B) $(-i)$ (Fig. 4), demonstrating that they predicted well the migrations per pulses of yeast chromosomes electrophoresed in CHEF under our experimental conditions.

The slopes p_1 and p_1' of Eqs. (13A) and (13B) were statistically equal (significant level of 0.05). Therefore, the migrations per pulse obtained during and after DNA reorientation converged in a single straight line. We obtained this single linear function taking into consideration the discontinuity at $t_p = t_r$. To model it we used a function $\Gamma(t_p - t_r)$ that takes values of 0 or 1 according to:

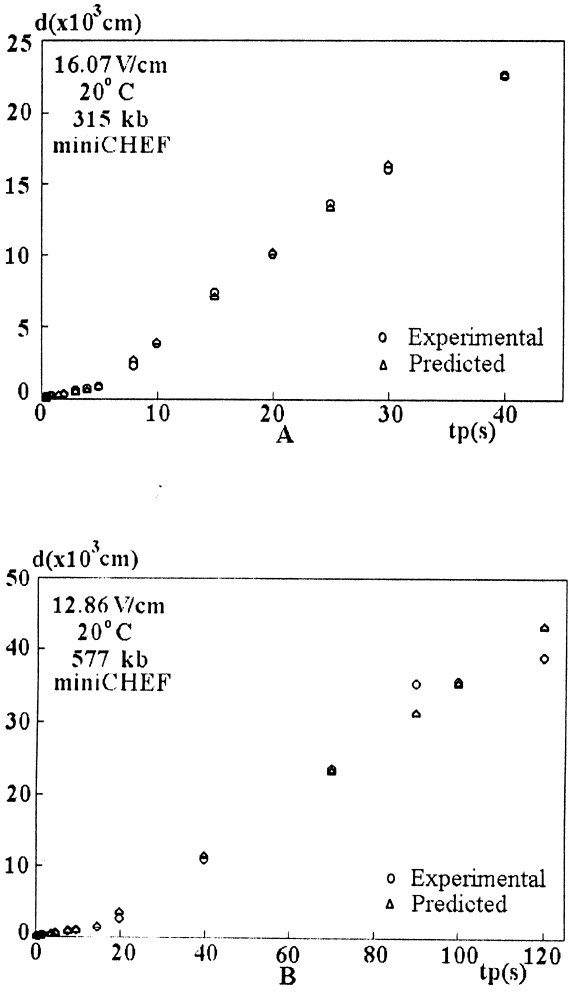

Fig. 4. Contrast among the experimental (d_{ik}) (○) and the predicted ($d_{(-ik)}$) (△) migrations per pulses of the 315 (A) and 577 (B) kb chromosomes at all assayed pulse times (t_{pik}) and the experimental conditions shown in the plots. Predictions were done with Eq. (13A) $(-i)$ and Eq. (13B) $(-i)$, using $\nu_{m(-i)}$, $\nu_{r(-i)}$ and $t_{r(-i)}$ estimates shown in Table 4.

$$\Gamma(t_p - t_r) = 1 \quad \text{if } (t_p - t_r) \leq 0$$

$$\Gamma(t_p - t_r) = 0 \quad \text{if } (t_p - t_r) > 0$$

Then, using $\Gamma(t_p - t_r)$, we rewrite together Eqs. (13A) and (13B) and describe the migration per pulse by the general function

$$d_{ik} = \nu_{ri} t_{pik} \Gamma(t_{pik} - t_{ri}) + \nu_{mi}(t_{pik} - t_{ri})[1 - \Gamma(t_{pik} - t_{ri})] \tag{14}$$

To obtain Eq. (14) we fitted $d_{ik}=p_0''+p_1''w_{ik}$; where w_{ik}, is the right member of Eq. (14) and v_{mi}, v_{ri}, t_{ri} are as for Eqs. (13A) and (13B). Thus $w_{ik} = \{\beta_r Q_i E_i^{1.45} / [8\pi\eta_i L_i^{1.35}]\}\{t_{pik}\Gamma(t_{pik} - t_{ri})\} + \{\beta_m Q_i E_i^{1.76}/[8\pi\eta_i L_i^{1.08}]\}\{(t_{pik}-t_{ri})[1-\Gamma(t_{pik}-t_{ri})]\}$. As can be observed in Figs. 1 and 2,5, the original migration data were scattered, whereas if we used the variable w_{ik} the data converged into a line. This general straight line (Fig. 5, continuous line) had non-significant intercept $p_0''=3.92\cdot10^{-5}$, $S^2p_0''=1.58\cdot10^{-8}$, and a significant slope $p_1''=0.998$ and $S^2p_1''=2.79\cdot10^{-6}$. Thus, we obtained a function that describes the migration per pulses that DNA molecules should have at different electric fields and temperatures in CHEF and miniCHEF electrophoresis. It was proposed that the dynamic of large DNA segments during electrophoresis can be described by a set of simple mechanical equations [28]. Our results support this proposition.

4.5. Calculation of DNA size, t_r, v_r and v_m using a single migrated distance in CHEF electrophoresis

We also analyzed if, by solving Eq. (14), we could estimate DNA size (and L), migration velocities and the reorientation time of a molecule from the value of its migrated distance (D) in a single CHEF run under known experimental conditions. The analytical solution of Eq. (14) can be obtained, but we solved it using a simple numerical method, that is based on increasing or decreasing an initial estimate of the solution until the sign of the function changes. Repeating these steps the solution is calculated with the precision that the analytical method would provide. Thus, giving to our computer program the experimental conditions (E, η, t_p), the experimental value of d (D/N_p), and an initial small size value (1 kb), Eq. (14) is evaluated by replacing these values in it. Then, the theoretic migration per pulse (d_t) for the initial size, is obtained and compared with the experimental one (d). While $d_t \neq d$, the size value is increased (or decreased) by Δkb (initial Δkb = 1), and a new d_t is recalculated and compared with d. The loop is finished when $(d-d_t)\approx0$ and the final size value (kb_u) is obtained. By replacing kb_u in Eqs. (10)–(12), we obtained the corresponding v_r, v_m and t_r (Table 5). We called these new estimations t_{ru}, v_{ru} and v_{mu}, because they were calculated from a single CHEF run.

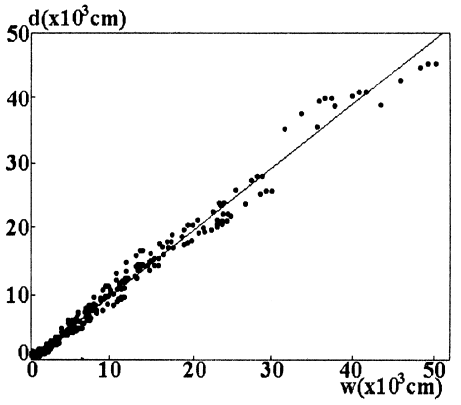

Fig. 5. Migrations per pulse of all studied DNA molecules (see examples in Figs. 1 and 2,4) ($n=490$) plotted together against the new independent variable w_{ik} (see Eq. (14)). Continuous line shows the best fit of Eq. (14) as a single straight line.

To validate if the method proposed in the above paragraph gives suitable estimates for kb_u, t_{ru}, v_{ru} and v_{mu} we compared the kb_u with the sizes of yeast chromosomes determined by sequencing [19], and compared the t_{ru}, v_{mu} and v_{ru} with the t_r, v_m and v_r obtained in separated CHEF experiments using the procedure described in Section 2. We did not use these later estimates for the fitness of Eqs. (1)–(12), (13A), (13B), (14).

The distances migrated by chromosome III of *S. cerevisiae* in miniCHEF electrophoresis at 10, 15, 20 and 30 s of pulse times, at 20°C and 10.71 V/cm were used to determine in each experiment the corresponding kb_u, t_{ru}, v_{ru} and v_{mu} (Table 5). The sizes (kb_u) estimated at these pulse times were close, giving a mean size of 344 kb, a mean error of 5.5 kb and a small relative standard deviation (R.S.D.) of 1.6% (Table 5). Similar results were obtained using migrations per pulse of this chromosome separated in miniCHEF at a higher electric field. The kb_u estimates were more disperse, but the R.S.D. [100(mean standard error/mean)] was less than 10%. In addition, the estimated mean sizes of chromosome III, either the one calculated from data obtained at 10.71 V/cm or the one at 12.86 V/cm did not significantly differ. However, the mean of kb_u obtained at 10.71 V/cm statistically differed from the sequenced size

Table 5
Estimations of DNA size (k_{bu}), t_{ru}, v_{ru} and v_{mu} solving Eq. (14) for the migrated distance of a molecule in a single CHEF run at $t_p > t_r$

t_p (s)	Yeast chromosome III					Yeast chromosome III					Yeast chromosome V				
	d ($\cdot 10^3$)	$E = 10.71$ V/cm				d ($\cdot 10^3$)	$E = 12.86$ V/cm				d ($\cdot 10^3$)	$E = 16.07$ V/cm			
		Estimations					Estimations					Estimations			
		kb_u	t_{ru}	$v_{ru} \cdot 10^5$	$v_{mu} \cdot 10^5$		kb_u	t_{ru}	$v_{ru} \cdot 10^5$	$v_{mu} \cdot 10^5$		kb_u	t_{ru}	$v_{ru} \cdot 10^5$	$v_{mu} \cdot 10^5$
10	0.985	336	6.76	8.85	30.26	1.161	420	7.19	10.67	41.02	1.592	529	7.33	13.60	59.60
15	2.528	333	6.68	8.88	30.28	4.024	340	5.37	11.49	41.72	4.182	561	7.95	13.32	59.32
20	3.856	353	7.24	8.70	30.14	5.676	381	6.29	11.04	41.34	7.328	547	7.68	13.44	59.44
25	–	–	–	–	–	–	–	–	–	–	10.568	527	7.30	13.61	59.62
30	6.861	354	7.26	8.69	30.13	–	–	–	–	–	12.665	592	8.57	13.07	59.07
Exp.Pr	–	–	6.25	6.53	28.89	–	–	5.65	10.71	39.56	–	–	8.40	14.56	63.37
S^2Pr	–	–	0.20	1.20	2.59	–	–	0.40	1.68	9.23	–	–	1.07	0.87	47.55
kb_u	–	344*	–	–	–	–	380	–	–	–	–	551	–	–	–
S kb_u	–	5.5	–	–	–	–	23	–	–	–	–	12	–	–	–
R.S.D. (%)		1.6					6.1					2.2			
Seq. size	–	315	–	–	–	–	315	–	–	–	–	577	–	–	

Comparison of the mean of k_{bu} with the sequenced size. Comparison of t_{ru}, v_{ru} and v_{mu} with the experimental t_r, v_r and v_m.
The units of d, t_r, v_r and v_m are as in Table 1. S^2Pr: Variances of parameters: $S^2 v_r$ and $S^2 v_m$ are multiplied by 10^{10}. Exp. Pr: t_r, v_r and v_m, determined as described in Section 2. (k_{bu}): Mean of the estimated sizes. Seq. size: Sequenced size. The asterisk indicates significant differences to a significant level of 0.05. R.S.D.: relative standard deviations of the estimated mean sizes.

of this chromosome. The other estimation did not significantly differ from the sequenced size (significant level 0.05). We also obtained size estimates kb_u for chromosome V separated in miniCHEF electrophoresis at 16.07 V/cm that are comparable to the real sequenced size (Table 5). The mean size did not significantly differ from the real size (Table 5). We concluded that the procedure did not render the exact sequenced size, as it happens when determining sizes by comigration with λ-mers; however, it gives estimates resembling the real sizes, with relatively small mean errors and the R.S.D.s were always less than 10%, indicating that the procedure provides size estimates with low variability.

We further compared the t_{ru}, v_{ru}, v_{mu} with the t_r, v_r and v_m of these molecules, determined as described in Section 2. In all tests, the calculated t_{ru}, v_{ru} and v_{mu} of chromosomes III and V did not significantly differ (significant level of 0.05) from the corresponding experimental t_r, v_r and v_m (Table 5). Thus, using the above described procedure, a single CHEF run is informative for approaching DNA size, t_r, v_r and v_m. This procedure, or another based on this principle, could give valuable information of migration velocities and reorientation times of the molecules to design CHEF runs.

4.6. Limitations and advantages of our approach

The functions presented here provided expressions for the velocities during a single pulse that were obtained considering how DNA spends the pulse time and how the driving and the frictional forces act on it. They have the limitation that the coefficients were determined only for 0.5×TBE and 1.5% agarose concentration. We do not know how they will behave at different conditions or for different size intervals. However, Eq. (12), describing t_r, predicted well the parameter τ for T4 DNA at 8 V/cm, 20°C and 1.5% agarose gel; and for various molecules the values of T^* obtained in 1% agarose gel, 0.5×TBE at 4 V/cm and 14°C, even for λ-DNA whose size is very far from the sizes of our analyzed molecules (230–1532 kb). As our functions were also able to reproduce DNA migration in CHEF, they can be used to develop efficient software to simulate DNA migration in the CHEF system. They should be also valuable to select a priori the experimental conditions required to separate a desired particular set of molecules.

We expect that changes in the ionic strength of the buffer and in the agarose concentration will modify the coefficients of our functions. Our functions will not reproduce migrations in equipments generating non homogeneous electric fields along the gel, up to a precise determination of the reorientation angle gradients. If these determinations are done, this kinetic approach, which is independent of the molecular model to explain the movement, may help to describe the migration of DNA in any PFGE system.

The method proposed to validate the predictions of the functions is general. Thus, it could be used in migrations studies carried out using other PFGE system. The procedure developed to approach DNA size, t_r, v_m and v_r using a migrated distance of a DNA molecule in a single CHEF run is also simple and permits the determinations of these parameters in a short time saving reagent and efforts. Finally, in the accompanying article [10] we proposed that migration data, collected in CHEF and miniCHEF can be used together to study DNA migration. The functions presented here corroborated this. Thus, by combining both, the data already collected in large chambers and new data obtained by fast separations in minichambers, we should be able to complete the description of DNA migration in CHEF in a very short time.

Acknowledgements

We thank Dr. Sergio Gurrieri and Dr. Marzella Baez-Camargo for kindly helping us to get the papers requested by the referees. E.O. is an International Fellow of the Howard Hughes Institute (USA).

References

[1] D.C. Schwartz, C. Cantor, Cell 37 (1984) 67.
[2] B. Birren and E. Lai, Pulsed Fields Gel Electrophoresis: A Practical Guide, Academic Press, New York, 1993, pp. 107, 111, 129, 131, 135.
[3] S. Gurrieri, E. Rizzarelli, D. Beach, C. Bustamante, Biochemistry 29 (1990) 3396.
[4] C. Bustamante, S. Gurrieri, S.B. Smith, Tibtech 11 (1993) 23.

[5] A.M. Riveron, D. Higginson, L. Lopez-Canovas, G. Perez, H.M. Garcia, M. Reyes, R. Manresa, Studia Biophys. 133 (1989) 73.
[6] G. Chu, Proc. Natl. Acad. Sci. USA 88 (1991) 11071.
[7] A.M. Riveron, L. Lopez-Canovas, J.A. Herrera, L. Ruiz, D. Higginson, M.D. Noa, F. Valdes, Arch. Med. Res. 25 (1994) 193.
[8] M. Burmeister and L. Ulanovsky (Editors), Methods in Molecular Biology, Vol. 12: Pulsed-Field Gel Electrophoresis. Protocols, Methods and Theories, Part IV: Theories and Observations, Humana Press, Totowa, NJ, 1992, pp. 347–467.
[9] O.J. Lumpkin, P. Dejardin, B.H. Zimm, Biopolymers 24 (1985) 1573.
[10] L. Lopez-Canovas, R. Biscay, M.D. Noa, E. Orozco and A.M. Riveron, J. Chromatogr. A (1998) in press.
[11] O.J. Lumpkin, B.H. Zimm, Biopolymers 21 (1982) 2315.
[12] E.M. Southern, R. Anand, W.R.A. Brown, D.S. Fletcher, Nucl. Acids Res. 15 (1987) 5925.
[13] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, Numerical Recipes in PASCAL, Press Syndicate of the University of Cambridge, Cambridge, 1989, pp. 548–598.
[14] B. Norden, Ch. Elvingson, M. Jonsson, B. Akerman, Quart. Rev. Biophys. 24 (1991) 103.
[15] B.H. Zimm, S.D. Levene, Quart. Rev. Biophys. 25 (1992) 171.
[16] H. Oana, Y. Masubuchi, M. Matsumoto, M. Doi, Y. Matsuzawa, K. Yoshikawa, Macromolecules 27 (1994) 6061.
[17] B. Akerman, Electrophoresis 17 (1996) 1027.
[18] C.R. Calladine, C.M. Collis, H.R. Drew, M.R. Mott, J. Mol. Biol. 221 (1991) 981.
[19] A. Goffeau, B.G. Barrell, H. Bussey, R.W. Davis, B. Dujon, H. Feldmann, F. Galibert, J.D. Hoheisel, C. Jacq, M. Johnston, E.J. Louis, H.W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, S.G. Oliver, Science 274 (1996) 546.
[20] A.M. Riveron, L. Lopez-Canovas, J.A. Herrera, M. Baez-Camargo, D. Higginson, E. Orozco, Anal. Lett. 28 (1995) 1973.
[21] L. Lopez-Canovas, D. Higginson, E. Orozco, A.M. Riveron, Anal. Lett. 29 (1996) 2079.
[22] G. Chu, D. Vollrath, R.W. Davis, Science 234 (1986) 1582.
[23] M. Reyes, A.M. Riveron, R. Martinez, G. Perez, Rev. CNIC. Ciencias Biologicas 21 (1990) 40.
[24] A.M. Riverón, J.A. Herrera, L. Ruiz-Esquivel, M. Báez-Camargo, L. López-Cánovas, M.D. Noa, E. Orozco, Anal. Lett. 28 (1995) 845.
[25] R.C. Weast (editor), Handbook of Chemistry and Physics, CRC Press, 57th ed., 1976–1977, CRC Press, Cleveland, OH, p. F-49.
[26] J. Sambrook, E.F. Fritsch and T. Maniatis, Molecular Cloning–A Laboratory Manual, Cold Spring Harbor Laboratory Press, New York, 2nd ed., 1989, Book 2, pp. 9.38, 9.52, 10.6.
[27] S. Gurrieri, S.B. Smith, K.S. Wells, I.D. Johnson, C. Bustamante, Nucl. Acids Res. 24 (1996) 4759.
[28] N. Lee, S. Obukhov, M. Rubinstein, Electrophoresis 17 (1996) 1011.
[29] H. Hervet, C.P. Bean, Biopolymers 26 (1987) 727.
[30] C.R. Cantor, C.L. Smith, M.K. Mathew, Ann. Rev. Biophys. Chem. 17 (1988) 287.
[31] J.A. Schellman, D. Stigter, Biopolymers 16 (1977) 1415.
[32] D. Stigter, Biopolymers 31 (1991) 169.
[33] Ch. Carlsson, M. Jonsson, Macromolecules 29 (1996) 7802.
[34] K. Gunderson, G. Chu, Mol. Cell. Biol. 11 (1991) 3348.
[35] M.S. Hutson, G. Holzwarth, T. Duke, J.L. Viovy, Biopolymers 35 (1995) 297.
[36] R.W. Whitcomb, G. Holzwarth, Nucl. Acids Res. 18 (1991) 6331.
[37] S.B. Smith, C. Heller, C. Bustamante, Biochemistry 30 (1991) 5264.