Resolution of DNA molecules greater than 5 megabases by contour-clamped homogeneous electric fields

Douglas Vollrath and Ronald W. Davis

Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305, USA

Received June 8, 1987; Revised and Accepted August 31, 1987

ABSTRACT

Excellent resolution of chromosomal DNA molecules from *Saccharomyces cerevisiae*, *Candida albicans* and *Schizosaccharomyces pombe* has been obtained using alternating contour-clamped homogeneous electric field (CHEF) gel electrophoresis. The largest of these molecules is greater than 5 Mb in size and is resolved after 130 hours in a 0.6% agarose gel at a field strength of 1.3 V/cm and a switching interval of 1 hour. Separation of concatamers of phage lambda DNA reveals four regions of resolution in alternating CHEF gel electrophoresis. There are two regions of good resolution in which mobility approximates a linear function of molecular weight. These are separated by a region of lower resolution and bounded at high molecular weights by a region of little or no resolution. The four regions are of practical and possibly theoretical importance.

INTRODUCTION

Alternating or pulsed field gel electrophoresis has been successfully applied to resolve whole chromosomal DNA molecules up to 2 mega-basepairs (Mb) in size from numerous single-celled eukaryotes [1-7]. Although current theory does not predict an upper limit to the size of molecules which can be separated by this technique [8], in practice, chromosomal DNA larger than 2 Mb has not been well resolved [1-4]. Extension of the technique to molecules of greater than 2 Mb would significantly expand the number of organisms to which it could be applied.

This paper addresses the problem of resolving molecules larger than 2 Mb. Two types of yeast have been used as sources of these large molecules: *Candida albicans*, a pathogenic yeast which contains a range of molecules starting at less than 2 Mb and extending to larger sizes [7]; and *Schizosaccharomyces pombe*, a fission yeast which has only three chromosomes, each of which should be greater than 2 Mb based on genetic mapping and genome size estimates [9,10].

These large DNA molecules have been resolved by employing a particular configuration of the alternating field technique known as contour-clamped homogeneous electric field (CHEF) gel electrophoresis [11]. It is shown here that a combination of low field strengths, long switching intervals and low agarose gel

concentrations can result in excellent resolution of molecules greater than 5 Mb in size.

MATERIALS AND METHODS

Preparation of Chromosomal DNA

Saccharomyces cerevisiae DNA samples were prepared as described [11]. *S. pombe* samples were prepared by growing strain 975 (h⁺) to stationary phase in YPD media (2% glucose, 2% bactopectone, 1% yeast extract). The cells were washed twice with 50 mM sodium EDTA (pH 8.0) and resuspended in 0.5 ml LET buffer (0.5 M sodium EDTA, pH 8.0, 10 mM Tris, pH 7.5) [1] to which 50-100 µl of 100 mg/ml Novozyme (Novo Laboratories) was added. The cells were warmed to 37°C and added to 0.835 ml of 1% low melting point agarose (FMC) in 125 mM sodium EDTA (pH 8.0) which had been cooled to 37°C. The mixture was allowed to solidify on ice. The agarose plug was then immersed in 1.4 ml LET buffer plus 1% beta-mercaptoethanol and incubated at 37°C overnight. The LET buffer was replaced with 1.4 ml of NDS buffer (0.5 M sodium EDTA, pH 8.0, 10 mM tris, pH 7.5, 1% sarkosyl) [1] containing 2 mg/ml proteinase K and the sample was incubated overnight at 50°C. The sample was then washed multiple times at 50°C with 50 mM sodium EDTA (pH 8.0) before storing at 4°C.

C. albicans (ATCC strain 14053) was grown to saturation in 5 ml of YPD media. The cells were washed twice with 50 mM sodium EDTA (pH 8.0), resuspended in 1 ml of 50 mM sodium EDTA plus 1% beta-mercaptoethanol and incubated at 30°C for 45 minutes. The cells were then washed twice with 1 M sorbitol and resuspended in 1 ml of 1 M sorbitol, 50 mM sodium phosphate (pH 7.4) and 50 mM sodium EDTA (pH 8.0). A 20 µl aliquot of 20 mg/ml Zymolyase 5000 (Kirin Breweries) was added and the cells were incubated overnight at 37°C. The resulting spheroplasts were carefully washed twice with 1 M sorbitol, 250 mM sodium EDTA (pH 8.0), and resuspended in 1.8 ml of this buffer. The cell suspension was warmed to 37°C and 2.5 ml of 1% low melting point agarose in 125 mM sodium EDTA was added. The mixture was allowed to solidify on ice, immersed in 4.3 ml of NDS buffer containing 2 mg/ml proteinase K, and incubated overnight at 50°C. The sample was then washed multiple times at 50°C with 50 mM sodium EDTA (pH 8.0). All three types of yeast chromosomal DNA preparations were stable for over 12 months when stored in 50 mM sodium EDTA at 4°C.

Human genomic DNA was prepared from peripheral blood leukocytes. Two volumes of fresh, EDTA-treated peripheral blood was gently overlaid on one volume of Ficoll-Paque (Pharmacia) and centrifuged at 900 x g for 12 minutes at room temperature. The "buffy coat," containing mononuclear leukocytes, was drawn off

with a pasteur pipet. The cells were washed twice with ice cold RSB (10 mM NaCl, 10 mM Tris, pH 7.5, 25 mM sodium EDTA, pH 8.0) [12], first with two volumes and then with one volume of buffer. The cells were counted in a hemocytometer and brought to a concentration of $5-7 \times 10^7$ cells/ml with RSB. The cell suspension was warmed to 37°C and immediately mixed with an equal volume of 1% low melting point agarose (also at 37°C) in RSB with 60 mg/ml proteinase K. The mixture was transferred into a teflon trough (14.5 cm x 0.6 cm x 0.5 cm) and allowed to solidify on ice. The agarose plug was then placed in a tube and completely immersed in an equal volume of 1% sodium dodecyl sulfate (SDS) in RSB and incubated at 50°C for 18-24 hours. The sample was washed extensively with 50 mM sodium EDTA (pH 8.0) at 50°C until no SDS precipitate was visible upon cooling to 40°C.

Digestion of Human Chromosomal DNA

A slice, with dimensions about equal to those of a gel well and containing about 5 µg of DNA, was cut from the agarose plug. The slice was placed in a 1.5 ml polypropylene tube with 1 ml of the appropriate restriction enzyme buffer and incubated at 37°C for 10-15 minutes to equilibrate the gel slice with the buffer. The buffer was then replaced. A total of three such incubations was performed, after which the slice was removed and blotted with a tissue. The slice was then placed in a 0.5 ml polypropylene tube containing 50 µl of restriction enzyme buffer. A two fold excess of restriction enzyme was added and the sample was incubated at the temperature recommended by the supplier (New England Biolabs) for 3 hours. This was repeated for a total incubation of 6 hours with a four-fold excess of enzyme. The reaction was terminated with 0.3 ml of 50 mM sodium EDTA. The sample could be stored for several days at 40°C with no deleterious effects.

Lambda Ladders

A total of 10 µg of bacteriophage lambda DNA (CI857S7) was added to 0.250 ml of 2% polyethylene glycol (M_r 8000), 2 mM ATP and 2 mM dithiothreitol. Three Weiss units of T4 DNA ligase (New England Biolabs) were added along with an equal volume of 1% low melting point agarose in 20 mM $MgCl_2$, 100 mM Tris (pH 7.5) which had been cooled to 37°C. The mixture was allowed to solidify on ice. The resulting plug was immersed in an equal volume of ligase buffer (1% PEG 8000, 1 mM ATP, 1mM DTT, 10mM $MgCl_2$, 50 mM Tris, pH 7.5) and incubated at room temperature for 18 to 24 hours. The sample was stored in 50 mM sodium EDTA (pH 8.0) at 40°C and remained stable for at least 3 months.

Gel Electrophoresis

CHEF gel electrophoresis was performed using a hexagonal array of electrodes as described [11], except that, for some gels, vertical rather than horizontal electrodes were employed. For all gels, the buffer was cooled to 90°C during electrophoresis and

the reorientation angle was 120°. Samples were either cut into blocks the size of a gel well and inserted, or they were melted by heating to 65-70°C for 2-5 minutes and loaded as a molten mixture using a 100 µl glass micropipet. In either case, the wells were then sealed with 1% low melting point agarose. Melting the samples can, at times, be more convenient. However, a larger proportion of DNA greater than 1 Mb in size seems to enter the gel when the samples are not melted.

Transfer and Hybridization

After staining with ethidium bromide (0.5 µg/ml), the gel was irradiated for 1 minute with 254 nm ultraviolet light (UV Products) to nick the DNA. The gel was then placed in 250 ml of denaturing solution (0.5 M NaOH, 1.5 M NaCl) for 30 minutes at room temperature with gentle shaking followed by a 30 minute incubation in 250 ml of neutralizing solution (0.5 M Tris, pH 7.5, 1.5 M NaCl) also at room temperature with shaking. The DNA was then transferred to a Genatran 45 filter (Plasco) using a reservoir of 500 ml of 10X SSPE [13] for 15-20 hours. The filter was baked for 1-2 hours at 80°C under vacuum and then prehybridized for 6 hours at 42°C in 15 ml of 3X SSPE, 50% formamide, 100 mg/ml salmon sperm DNA, 0.1% bovine serum albumin, 0.1% Ficoll (MW 400k), 0.1% polyvinyl pyrrolidone (MW 300k) and 1% SDS. Hybridization was carried out in 15 ml of the same solution with 5% dextran sulfate added for 24-36 hours at 42°C using 1-2 X 10⁶ cpm/ml of random hexamer primed probe [14]. The filter was washed once in 500 ml of 50% formamide, 3X SSPE, 0.2% SDS at 42°C for 15 minutes and twice in 500 ml of 0.5X SSPE, 0.1% SDS at 60°C for 30 minutes each.

RESULTS

Resolution Characteristics of CHEF

To achieve separation of molecules greater than 2 Mb in size, an understanding of the resolution characteristics of the CHEF system is essential. Excellent resolution of *S. cerevisiae* chromosomal DNA can be obtained using CHEF (Fig. 1A). However, these molecules are not uniformly distributed throughout the range of resolution and their sizes are not known with precision. A more accurate picture of the resolution characteristics can be obtained by electrophoresis of concatamers of bacteriophage lambda. Figure 1B shows the result obtained when a "ladder" of lambda concatamers is separated under conditions suitable for resolving most of the *S. cerevisiae* molecules. Four regions of separation are apparent. From bottom to top, they are region 1, an area of good resolution in which mobility appears to be a linear function of molecular weight; region 2, a small area of reduced resolution in the middle of the gel; region 3, an area of maximal resolution in which mobility is again a linear function of molecular weight; and region 4, an area at the top of the gel of little or no resolution.

Figure 1

Four regions of resolution in alternating CHEF gel electrophoresis.

(A) Chromosomal DNA from *S. cerevisiae* strain YNN295 was electrophoresed for 24 hours at 160 volts using an 80 second switching interval. The distance between driving electrodes was 26.5 cm. (B) and (C) Concatamers of phage lambda DNA molecules were subjected to two separate electrophoresis runs in 1% agarose gels using a field strength of 210 volts and a switching interval of either 60 or 80 seconds. The distance between driving electrodes was 30 cm. The 60 second gel is shown in (B) and arrows to the right of the gel indicate the approximate extent of the four regions described in the text. Data from both runs were used to plot the mobility versus molecular size curves shown in (C). Mobility was calculated relative to the lambda monomer (48.5 kb) to facilitate comparison between the two gels. Arrows to the right of the plot indicate the approximate extent of the four regions for the 80 second curve.

Figure 1C shows the effect of increasing the switching interval from 60 seconds (used in Fig. 1B) to 80 seconds on this resolution profile. Increasing the switching interval displaces region 3 to higher molecular weights. However, the slope of the curve and the extent of the region remain largely constant. This relationship holds

Figure 2

Four regions of resolution with different electrophoresis conditions and a variety of samples.

(A) *S. cerevisiae* and lambda samples were electrophoresed for 53.5 hours in 1% agarose at 120 volts using a switching interval of 4 minutes and an electrode spacing of 30 cm. Band 11, containing chromosomes VII and XV, is about 1.2 Mb in size and is shown as a point of reference [19]. (B) Human genomic DNA restricted with the infrequently cutting enzymes BssH II, Mlu I, and Not I was electrophoresed for 22 hours in 1% agarose at 160 volts using a switching interval of 60 seconds and an electrode spacing of 26.5 cm. Arrows to the right of both gels indicate the approximate extent of the four regions described in the text.

even when region 3 is shifted to higher molecular weights (Fig. 2A). Molecules differing by a single lambda monomer (48.5 kb) can be resolved in the 1.2-1.5 Mb range. In contrast, decreased resolution in region 1 is observed under these conditions. Thus, by properly positioning region 3, resolution of multi-megabase molecules which differ in length by only a few percent should be possible using CHEF.

Another noteworthy characteristic of region 3 is that DNA bands within it are

Figure 3

Effect of field strength and agarose gel concentration on the resolution characteristics of CHEF.

(A) and (B) *S. cerevisiae* and *C. albicans* samples were electrophoresed at 200 volts in a 0.7% agarose gel for 24 hours using a switching interval of (A) 90 seconds and (B) 120 seconds. (C) and (D) *S. cerevisiae* and *C. albicans* samples were electrophoresed at 100 volts in a 0.7% agarose gel for 44 hours using a switching interval of (C) 3.5 minutes and (D) 5 minutes. An *S. cerevisiae* rDNA [20] probe hybridizes to the lowest mobility *C. albicans* band in (D). (E) and (F) *S. cerevisiae* and *C. albicans* samples were electrophoresed at 200 volts using an 80 second switching interval and (E) a 1% agarose gel for 28 hours or (F) a 0.6% agarose gel for 24 hours. In all panels, the distance between driving electrodes was 26.5 cm.

more diffuse and thus appear fainter relative to bands in regions 1 and 2. This effect is well illustrated in genomic digests of human DNA. Figure 2B shows that all four regions are visible in such digests. Region 2 appears as an increase in staining intensity near the middle of the gel. Region 3 appears as a more faintly staining region between region 4 and region 2. These differences in intensity are not likely to be due to discontinuities in the distribution of fragment sizes in the samples since their presence is independent of both the restriction enzyme used and the switching interval.

Resolution of Multi-megabase Yeast Chromosomal DNA

The previous section establishes that CHEF gel electrophoresis displays a region of high resolution which, as with other alternating field techniques, can be shifted to higher molecular weights by increasing the switching interval. This knowledge was applied to the resolution of yeast chromosomal DNA of greater than 2 Mb in size.

Figure 4

Resolution of S. pombe chromosomal DNA molecules.

(A) *S. pombe* and *S. cerevisiae* samples were electrophoresed in a 0.6% agarose gel for 120 hours using a switching interval of 30 minutes and a field strength of 50 volts. (B) *S. pombe*, *C. albicans* and *S. cerevisiae* samples were electrophoresed in a 0.6% agarose gel for 130 hours using a switching interval of 60 minutes and a field strength of 40 volts. The distance between driving electrodes in both cases was 30 cm. Bands were identified by Southern blot using specific sequence probes for *S. pombe* chromosomes I (*mei2.8*) and II (*mei3.2*) [21,22]. *S. cerevisiae* rDNA [20] cross-hybridizes to *S. pombe* rDNA and is thus a probe for both *S. pombe* chromosome III and *S. cerevisiae* chromosome XII. CEN4 [23] was used to identify *S. cerevisiae* chromosome IV.

C. albicans is thought to possess larger chromosomal DNA than *S. cerevisiae* [7]. Figure 3A shows a CHEF gel run under conditions which optimize the resolution of the largest *S. cerevisiae* chromosomes and the smallest chromosomes of *C. albicans*. The intense ethidium bromide staining in the upper portion of the *C. albicans* lanes indicates that these large molecules are in region 4 of the gel. When the switching interval is increased from 80 seconds to 120 seconds to shift region 3 to a higher molecular weight, the larger chromosomal bands begin to smear in both samples (Fig. 3B).

The smearing effect observed for the largest DNA molecules could be a result of

requiring them to move under too strong an electric field. This hypothesis can be tested by resolving these same molecules at a lower voltage. However, in CHEF, lowering the voltage displaces region 3 to a lower molecular weight range. So, in order to resolve the same sized molecules at a lower voltage, the switching interval must be increased. In fact, a given range of molecular weights can be resolved using virtually any combination of field strength and switching interval as long as the product of these two values remains approximately constant (data not shown).

In Figure 3C, the field strength has been decreased to 100 volts and the switching interval increased to 3.5 minutes. The smearing effect is largely eliminated. When the switching interval is increased further to 5 minutes and the voltage held constant, good resolution of the larger *C. albicans* chromosomal DNA is obtained (Fig. 3D). Thus, large molecules which smear at one field strength are better behaved when the field strength is decreased. With this lesson in mind, resolution of the *S. pombe* chromosomal DNA was attempted.

The field strength was further decreased to 50 volts and the switching interval increased by a factor of six to 30 minutes. The agarose concentration was also decreased from 0.7% to 0.6%. Lowering the agarose concentration causes region 3 to shift to a higher molecular weight when the voltage and switching interval are held constant (Fig. 3E and F). However, the chromosomal bands tend to be more diffuse at lower agarose concentrations, limiting the usefulness of very dilute gels.

Figure 4A shows the results of this initial attempt to separate the *S. pombe* chromosomes. Good resolution of *S. cerevisiae* chromosomes IV and XII, which differ in size by about 1 Mb, is obtained. These chromosomes are also well resolved from the bulk of DNA in the *S. pombe* sample. However, the sample itself is still in region 4 of the gel.

To better resolve the *S. pombe* chromosomes, the voltage was decreased to 40 volts and the switching interval doubled to one hour. All other parameters were held constant. Figure 4B shows that good resolution is now obtained. Three DNA bands are evident. Probes specific for each chromosome identify these bands as the three individual chromosomal DNA molecules of *S. pombe* (data not shown). These assignments are consistent with those of other workers [15].

DISCUSSION

The resolution profile of CHEF gel electrophoresis displays four distinct regions: two regions of linear resolution differing in slope, each bounded above by a region of decreased resolution. This biphasic resolution profile is suggestive of the superposition of two different resolving mechanisms, one acting primarily in region 1 and another predominating in region 3. The mechanism acting in region 3 most likely results from the exquisite dependence on molecular weight of the time required

for a molecule to reorient when the direction of the electric field is changed, as proposed by Schwartz and Cantor [1]. The sensitivity of the position of region 3 to the switching interval supports this notion. Molecules in this region probably spend a large amount of time reorienting and only a small amount traveling through the gel.

In region 1, by contrast, the molecules probably reorient rapidly relative to the switching interval, so little resolution is obtained on the basis of reorientation time. Instead, alternating the fields may prevent the molecules from assuming a conformation in which mobility is independent of molecular weight, as when large DNA is subjected to conventional agarose gel electrophoresis [16]. This would allow sieving through the gel matrix to occur. In this model, the time required for a molecule to assume the non-sieving conformation would have to be long relative to the switching interval. The fact that the slope of region 1 approaches zero (i.e. no resolution) as the switching interval increases (see, for example, Fig. 2A) supports this idea.

The specific range of molecular weights present in region 3 of any given gel is a complex function of field strength, switching interval, agarose concentration and temperature. The magnitude of the force applied to the molecule by the electric field, the duration of that force, the resistance to reorientation caused by the agarose matrix and the thermal energy of the DNA all contribute to the reorientation time. In addition, there is an upper limit to the strength of the electric field that can be used to resolve a given large molecule when the agarose matrix and temperature are held constant. Above this limit, the molecules are not well behaved. For example, at field strengths normally used to resolve molecules less than 2 Mb, the largest *S. pombe* molecules do not even enter the gel. Presumably, these large molecules become trapped when required to move too rapidly through the gel matrix.

The low field strengths necessary to prevent trapping of multi-megabase DNA molecules require electrophoresis times of several days to a week. Thus, extension of CHEF gel electrophoresis to even larger molecules is limited by the extremely long run times involved. Perhaps this problem could be circumvented by the development of new gel matrices which reduce trapping. Alternatively, condensation of the molecules in a way that preserves the molecular weight dependence of reorientation time may prove fruitful.

By recognizing the importance of field strength in CHEF gel electrophoresis, it has been possible to resolve the three chromosomal DNA molecules of *S. pombe* which constitute the organism's 15 Mb haploid genome [10]. This has also recently been achieved using the pulsed field gradient system [15]. The ratio of the sizes of the two largest *S. pombe* molecules (chromosomes II and I) is estimated to be 1:1.4 from the genetic map and measurements of fluorescently stained condensed chromosomes [9,17]. Chromosome III, which contains over 1 Mb of the rDNA

cluster, is not adequately assessed by either of the above techniques because the rDNA exhibits unique condensation and recombination characteristics compared to the rest of the genome [17,18]. An independent estimate of the size of chromosome III is required.

Given the resolution profile in CHEF gel electrophoresis, extrapolations of molecular weight beyond the range of size markers must be done with caution. It is apparent, however, that chromosome III has a lower mobility in CHEF (Fig. 4B) than *S. cerevisiae* chromosome XII which is about 2.5 Mb in size (G. Carle and M. Olson, personal communication). Assuming chromosome III is about 3 Mb in size, chromosomes II and I should be 5 and 7 Mb respectively. If these estimates are reasonably accurate, then resolution of the chromosomes of many other organisms should be possible using this technique.

ACKNOWLEDGEMENTS

This work was supported by the Medical Scientist Training Program at Stanford (D.V.), Public Health Service grants GM21891 and AG02908, and the MacArthur Foundation (R.W.D.). We thank Adam Kuspa and Gilbert Chu for reading the manuscript, Gilbert Chu and Marjorie Thomas for helpful discussions and David Beach for providing the *S. pombe* DNA probes.

REFERENCES

1. Schwartz, D.C. and Cantor, C.R. (1984) Cell 37, 67-75.
2. Carle, G.F. and Olson, M.V. (1985) Proc. Natl. Acad. Sci. 82, 3756-3760.
3. Van der Ploeg, L.H., Schwartz, D.C., Cantor, C.R. and Borst, P. (1984) Cell 37, 77-84.
4. Van der Ploeg, L.H., Smits, M., Ponnudurai, T., Vermeulen, A., Meuwissen, J.H., and Langley, G. (1985) Science 229, 658-661.
5. Garvey, E.P. and Santi, D.V. (1986) Science 233, 535-540.
6. Bernards, A., Kooter, J.M., Michels, P.A., Moberts, R.M. and Borst, P. (1986) Gene 42, 313-322.
7. Snell, R.G. and Wilkins, R.J. (1986) Nucl. Acids Res. 14, 4401-4406.
8. Lumpkin, O.J., DeJardin, P. and Zimm, B.H. (1985) Biopolymers 24, 1573.
9. Gyax, A. and Thuriaux, P. (1984) Current Genetics 8, 85-92.
10. Egel, R., Kohli, J., Thuriaux, P. and Wolf, K. (1980) Ann. Rev. Genetics 14, 77-108.
11. Chu, G., Vollrath, D. and Davis, R.W. (1986) Science 234, 1582-1585.
12. Hieter, P.A., Hollis, G.F., Korsmeyer, S.J., Waldman, T.A., and Leder, P. (1981) Nature 294, 536-540.
13. Davis, R.W., Botstein, D. and Roth, J.R. (1980) Advanced Bacterial Genetics, pp. 174-176. Cold Spring Harbor Press.
14. Feinberg, A. and Vogelstein, B. (1983) Anal. Biochem. 132, 6-13.
15. Smith, C.L., Matsumoto, T., Niwa, O., Klco, S., Fan, J.-B., Yanagida, M. and Cantor, C.R. (1987) Nucl. Acids Res. 15, 4481-4489.

16. McDonell, M.W., Simon, M.N. and Studier, F.W. (1977) *J. Mol. Biol.* 110:119-146.
17. Umesono, K., Hiraoka, Y., Toda, T. and Yanagida, M. (1983) *Current Genetics* 7, 123-128.
18. Toda, T., Nakaseko, Y., Niwa, O. and Yanagida, M. (1984) *Current Genetics* 8, 93-97.
19. Mortimer, R.K. and Schild, D. (1985) *Micro. Rev.* 49, 181-212.
20. Philippsen, P., Thomas, M., Kramer, R.A. and Davis, R.W. (1978) *J. Mol. Biol.* 123, 387-404.
21. Beach, D.H., Rogers, L. and Gould, J. (1985) *Current Genetics* 10, 297-311.
22. McLeod, M., Stein, M., and Beach, D.H. (1987) *EMBO J.* 6, 729-736.
23. Hieter, P., Mann, C., Snyder, M. and Davis, R.W. (1985) *Cell* 40, 381-392.