Short Technical Reports

DNAOPT: A Computer Program to Aid Optimization of DNA Gel Electrophoresis and SDS-PAGE

ABSTRACT

Several methods and computer programs have been developed for estimating the size of DNA fragments from gel electrophoresis. However, methods are lacking that may facilitate in optimization of gel conditions. In this article, a computer program called DNAOPT is described, which was developed to assist researchers in tuning the gel conditions of gel electrophoresis. The DNA OPT program fits the reciprocal of the migration distance vs. the size of the DNA fragments using the hyperbolic regression method and computes the hyperbolic parameters such as signal, flatness and capacity (optimization parameters). The program further manipulates these parameters obtained by running gel electrophoresis under various conditions (i) to determine the relationship between the gel conditions (temperature, buffer concentration, electric field strength, etc.) and optimization parameters; (ii) to demonstrate gel electrophoresis curves and optimization parameters graphically; and (iii) to represent the optimizing parameters at different gel conditions in tabular form. The above-mentioned program options aid the users in selecting optimum gel conditions by running the gel repeatedly under various conditions in which the agarose con centration, electric field strength, temperature, buffer concentration and so on are varied in a systematic way for each set of gel conditions. Similarly, this program can also be used to optimize gel conditions of sodium dodecyl sulfate polyacrylamide gel electrophoresis.

INTRODUCTION

Electrophoresis has become an important tool for studying macromolecules, DNA, RNA and proteins. It is most commonly used for determining the length of DNA fragments from their electrophoretic mobilities. DNA fragment size is estimated by comparing the migration distance of the unknown fragment with those of known size (standards). Several methods have been proposed for approximating the relationship between the migration distance and the DNA fragment size, such as the *(i)* semilogarithmic method; *(ii)* Southern method, based on reciprocal relationship (1,9); *(iii)* global method, based on reciprocal relationship (8); *(iv)* hyperbolic regression method (10,11); and *(v)* graphical method (4,6). Based on these approximations, many computer programs have been developed for computing fragment lengths from gel electrophoresis data (2-4).

One of the common problems faced by researchers who run gel electrophoresis is adjusting gel conditions so that *(i)* the large fragments can be measured; *(ii)* bands of fragments having marginal size difference can be distinguished; and *(iii)* usable dynamic ranges can be increased. Presently, available programs on gel electrophoresis are not appropriate to overcome the above problem. Thus, there is need to develop a program that may support the researchers in optimizing gel conditions.

In this report a computer program is described that facilitates the researchers in optimizing gel conditions by running the gel repeatedly under various conditions in a systematic way. The optimization of gel conditions is accomplished by using the hyperbolic regression, which calculates optimizing parameters, such as signal, capacity and other characteristics of the response curve. In addition, the program computes the relationship between gel conditions and optimizing parameters so that users may select gel conditions according to their requirements. The program also has options to graphically express electrophoresis curves on the screen or on an HP-GL-compatible plotter.

ALGORITHM

Computation of Optimizing Parameters

DNAOPT used the hyperbolic regression method for fitting the standard data, which is the efficient method for determining the relationship between the migration distance and the size of DNA fragments over the entire range

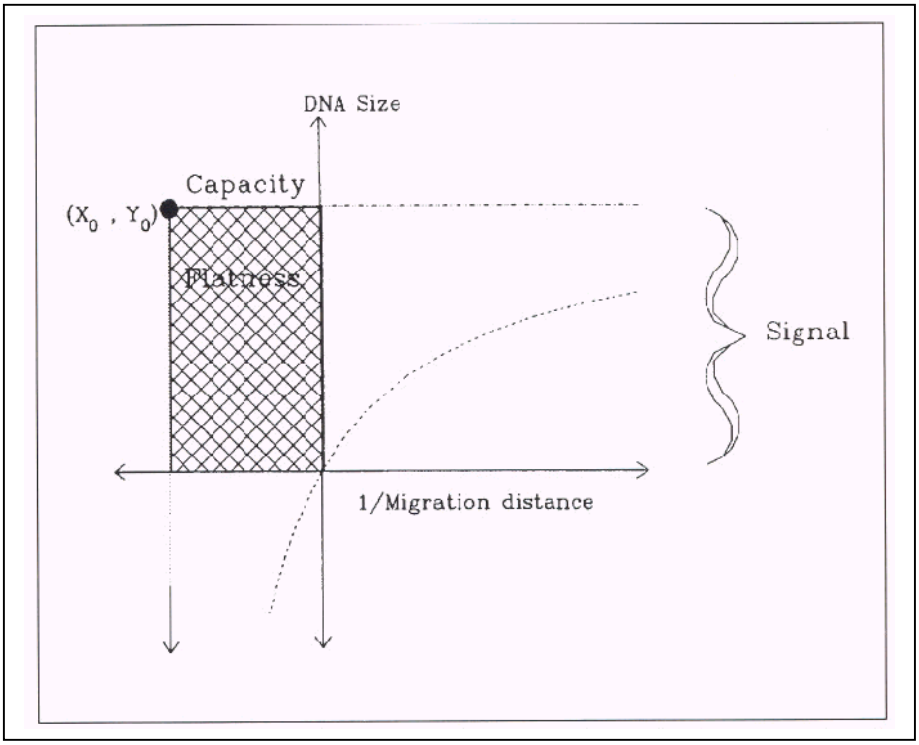

Figure 1. Graphical representation of rectangular hyperbola for DNA gel electrophoresis with new mathematical origin at *(Xo,Yo)*, where *Xo* and *Yo* represent capacity and signal, respectively. Flatness of the curve is shown by shadow area.

(10, 11). The program fits the size of the DNA fragments vs. the reciprocal of the migration distance using the hyperbolic regression method and derives a hyperbolic equation. The simplest hyperbola is described by the equation $XY = Co$, where Co is a constant, and it has asymptotes at X = 0 and Y = O. The equation (X - *Xo) (Y - Yo)* also represents a hyperbola that has asymptotes at *(Xo,Yo)*. The reciprocal of the migration distance vs. the size of the fragments of the gel electrophoresis hyperbolic curve is shown in Figure I, which illustrates the graphical significance of terms such as signal, capacity and flatness.

The upper dotted line in Figure I is defined as a signal which is the maximum size of a DNA fragment that can be measured for that set of conditions.

The capacity of the gel represents the shape of the response curve whether the curve becomes easily saturated or not. The flatness of the curve is the product of signal and capacity. In the general equation of rectangular hyperbola, the variables that represent the experimental data values are X (for reciprocal of migration distance) and *Y* (for fragment size),

$$(X - Xo)(Y - Yo) = Co \qquad \text{[Eq.l]}$$

The regression constants that define the curve are *Xo, Yo* and Co representing capacity, signal and flatness. The constants can be determined by fitting standard data to rectangular hyperbola using the least-squares curve-fitting method. The detailed derivation of the hyperbolic interpolation formula has been described previously (5,7,10,11).

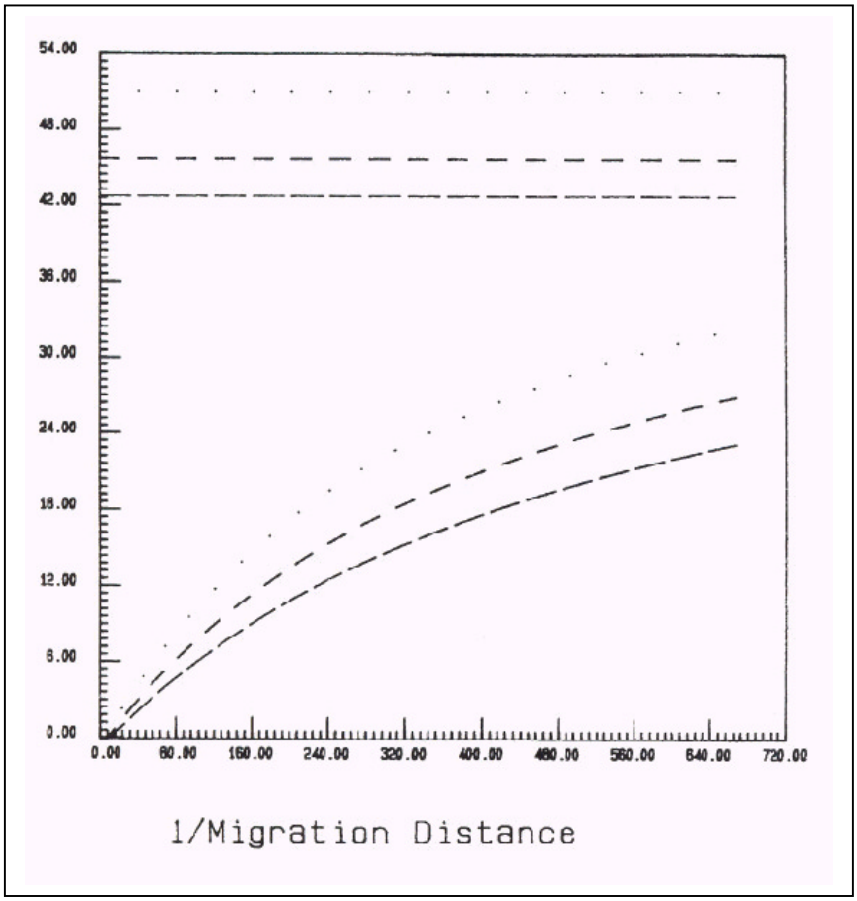

Figure 2. The optimization curves of Ie DNA digested with HindIII, electrophoresed in gels of different agarose concentrations. The line (- . '), (- -) and (- -) represent the standard curve and signal of Ie DNA electrophoresed on 0.8%, 1.0% and 1.2% agarose gels, respectively. The curves were fitted using the hyperbolic regression method and plotted on an HP-GL-compatible plotter using the DNAOPT program.

Relationship Between Optimizing Parameters and Gel Conditions

The program exploits the linear regression method for computing the relationship between optimizing parameters and gel conditions. It first calculates the optimizing parameters from equation 1 for various gel conditions; then it fits the optimizing parameters vs. the gel conditions data by following linear equation:

$$P = AO + AIX(G), \qquad \text{[Eq.2]}$$

where *P,* G, *Ao* and *Al* represent the optimizing parameter, gel condition, the constant and slope of the curve, respectively. The values of *Ao* and *Al* were calculated by fitting the data using the least-squares curve-fitting method from the computer program. Equation 2 represents the linear interpolation formula that determines the relationship between gel conditions and optimizing parameters.

COMPUTER PROGRAM

DNAOPT computer program was developed to run on an IBM@-compatible computer under the MS-DOS@ operating system. It is written in the C language and compiled using TURBO C. This is an entirely menu-driven program and provides the following major options: (i) Input/edit: The program has its own editor to enter/update the unknown/standard (DNA marker) data and can import data from the ASCII file. (ii) Optimization: DNAOPT computes the optimization parameters by fitting the standard data using the hyperbolic regression method and then compares these optimization parameters for different sets of gel conditions. The program has the ability to demonstrate the electrophoresis curve and to graphically optimize the parameters. These curves can be shown on the screen and can be plotted on the HPGL-compatible plotter (Figure 2). *(iii)* Fragment size estimation: This option allows the user to determine the size of DNA fragments of unknown samples by using the standard data file. The program fits the standard data (reciprocal of mobility vs. fragment size) by using hyperbolic curve fitting and calculates

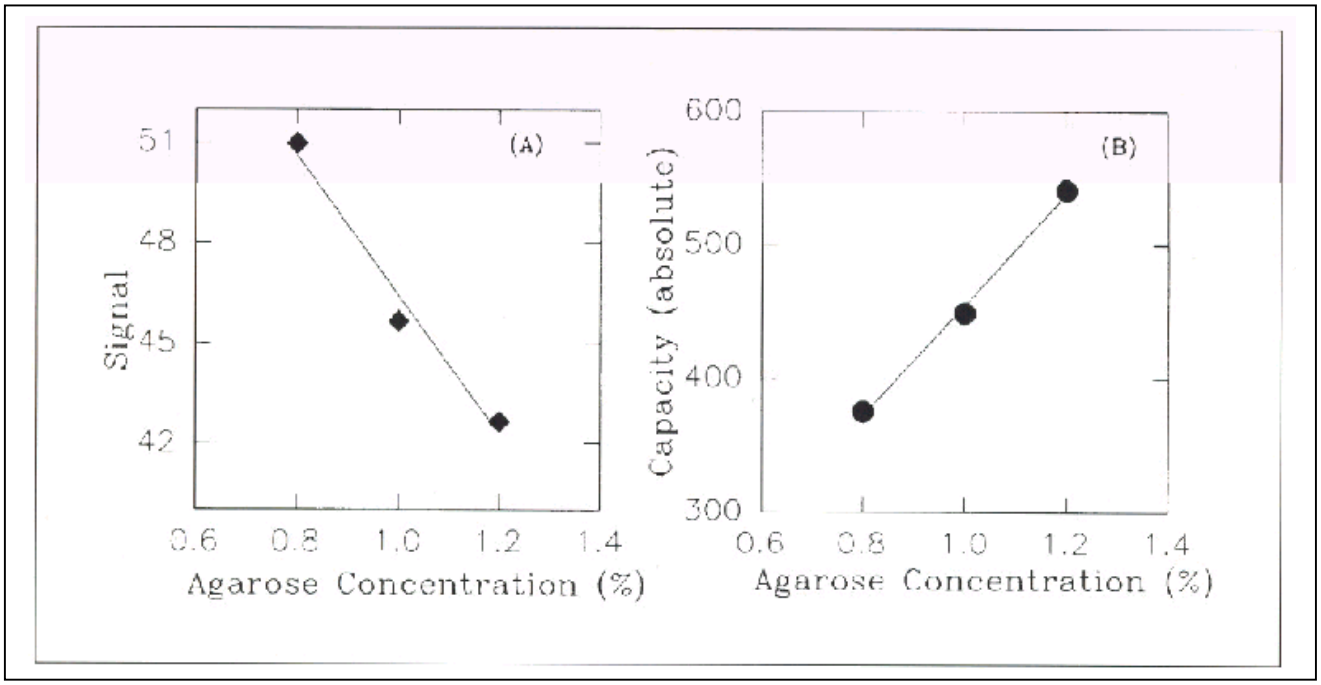

Figure 3. (A) Shows the percentage of agarose gel concentration vs. signal (maximum size of DNA fragments having sufficient resolution). (B) Shows the percentage of agarose gel concentration vs. capacity (absolute).

the DNA fragment size of unknown fragments. *(iv)* Relationship between optimizing parameters and gel conditions: The program fits the optimizing parameters vs. the gel conditions by using the linear regression method as described in the Algorithm section and computes the interpolation formula (equation 2)'0 This option assists the user in estimating the most appropriate gel conditions according to the user's requirements. *(v)* Print the data: This option allows the user to print the data file on the printer. The user should enter the old filename (previously saved in the standard data file) from the directory displayed on the screen when the program asks for the file name.

RESULTS

Lambda DNA was digested with *HindIII,* and the gel was run at a 0.8% agarose gel concentration at room temperature. In the second set of gel conditions, 'A DNA was digested with *HindIII,* and the gel was run at a 1.0% agarose concentration keeping the other conditions intact. Similarly, in the third set of gel conditions, the gel was run at a 1.2% agarose concentration keeping the other conditions intact. The fragment size was determined exactly by the GMAP program (6) from known sequences, and migration distances were measured as closely as possible with a ruler.

DNAOPT was used to analyze the electrophoresis data corresponding to each set of gel conditions. The electrophoresis curve and optimization parameters are shown in Figure 2, representing three sets of gel conditions, which are examples of the output of the DNAOPT program that is plotted on the HP-GL-compatible plotter. This figure demonstrates the impact of the agarose concentration on the electrophoresis curve. The program automatically computes the optimizing parameters, which were used for finding the relationship between the optimizing parameter and agarose concentration. Figure 3A shows the signal (maximum size of a DNA fragment that can be measured by a particular set of gel conditions at sufficient resolution) vs. the agarose gel concentration, which indicates that the signal is inversely proportional to the agarose gel concentration. The user may calculate the gel concentration required for measuring the large fragment of DNA using the interpolation formula (equation 2). Figure 3B shows the capacity vs. the agarose gel concentration, which demonstrates that the capacity (or slope) of the curve is proportional to the agarose gel concentration. The user can calculate the agarose concentration for obtaining the desired capacity using equation 2; for example, the user needs a high capacity to distinguish bands of the fragment having a marginal size difference. On the basis of equation 2, described in the Algorithm section, the user can adjust the agarose gel concentration for the required capacity. Similarly, the user can regulate other gel conditions depending on the requirements.

DISCUSSION

The main purpose for developing the DNAOPT program is to provide an

efficient, reliable and fast technique for optimizing gel conditions. The optimization of gel conditions is a prerequisite for utilizing the full efficiency and sensitivity of gel electrophoresis. To enhance the sensitivity and usable dynamic range of gel electrophoresis, it is mandatory to adjust the experimental conditions, namely, buffer concentration, electric field strength, temperature and agarose concentration. Hyperbolic regression is a mathematical technique that gives the researcher a quantitative analysis of gel electrophoresis data and could be utilized to fine-tune gel conditions.

To optimize gel conditions, the user should run the electrophoresis for different sets of conditions, and then use the program to analyze the data. The program provides the relationship between gel conditions and optimizing parameters that could be utilized to tune gel conditions according to the user's requirements. The program also computes the interpolation formula for gel conditions using equation 2, described in the Algorithm section, that allows the user to calculate gel conditions for using gel electrophoresis for a specific use. Thus, the user may tune the gel conditions using this program by running the gel repeatedly under various conditions in which the agarose concentration, electric field strength, temperature and buffer concentration are varied in a systematic way for each set of gel conditions. It has been shown earlier that the protein molecular weight standard data obtained from sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) at different experimental conditions of polyacrylamide gel, temperature, width of gel, etc., fit the hyperbolic curve very well so that this method can be applied to optimize SDS-PAGE (4,10,11).

The program is written in TURBO C and can be run on an IBM-compatible computer under MS-DOS Version 2.0 or higher and requires a CGA or EGA or VGA card. It runs without a math co-processor, and a hard copy of the graph can be obtained on an HP GL-compatible plotter or printer. A copy of the DOS version of DNAOPT program can be obtained from netserver at EMBL, from author (raghava @imtech.emet.in), or by downloading the program from BioTechNet@.

REFERENCES

I. Elder, J.K. and E.M. Southern. 1987. Computer-aided analysis of one dimensional restriction fragment gels, p. 165-172. *In* MJ. Bishop and CJ. Rawlings (Eds.), Nucleic Acid and Protein Sequence Analysis: A Practical Approach; IRL Press, Oxford.

2. Grabner, M. and R. Hofbauer. 1991. Acomputer program for molecular weight determination of DNA fragments (HOWBIG). Comput. Appl. Biosci. 7:317-319.

3. Lacroix, J. 1994. DNA fragment size determinator on agarose gel by using the application gel. Comput. Appl. Biosci. *10:185-187.*

4. Raghava, G.P.S. 1994. Improved estimation of DNA fragment length from gel electrophoresis. BioTechniques 17:100-104.

5. Raghava, G.P.S. and J.N. Agrewala. 1994. Method for determining the affinity of monoclonal antibody using non-competitive ELISA: A computer program. J. Immunoassay *15:115-128.*

6. Raghava, G.P.S., A.K. Joshi and J.N. Agrewilla. 1992. Calculation of antibody and antigen concentrations from ELISA data using a graphical method. J. Immunol. Methods 153:263-264.

7. Raghava, G.P.S. and G. Sahni. 1994. GMAP: A multipurpose computer program to aid synthetic gene design, cassette mutagenesis and the introduction of potential restriction sites into DNA sequences. BioTechniques 16: 11161123.

8. Schaffer, H.E. and R.R. Sederoff. 1981. Improved estimation of DNA fragment lengths from agarose gels. Anal. Biochem. *115:113*122.

9. Southern, E.M. 1979. Measurement of DNA length by gel electrophoresis. Anal. Biochem. *100:319-323.*

IO. Studnicka, G.M. 1987. Hyperbolic regression analysis for kinetics, electrophoresis, ELISA, RIA, Bradford, Lowry, and other applications. Comput. Appl. Biosci. 3:9-16.

II. Studnicka, G.M. 1991. ELISA assay optimization using hyperbolic regression. Comput. Appl. Biosci. 7:217-224.

This is a communication No. 024/94 from the Institute of Microbial Technology, Chandigarh, supported by grants from the Council of Scientific and Industrial Research and Department of Biotechnology, Government of India. The author is grate ful to Dr. *Nisha Dogra for her technical assistance. Address correspondence to G.p.s. Raghava, Institute of Microbial Technology, Sector 39A, Chandigarh 160 014, INDIA.*

Received 27 July 1994; accepted 20 August 1994.

G.P.S. Raghava
Institute of Microbial Technology
Chandigarh, India