Computer assisted microdensitometric analysis of footprinting autoradiographic DATA

James C.Dabrowiak^{1*}, Andrew Skorobogaty¹, Ned Rich², Calvin P.H.Vary² and John N.Vournakis²

Departments of ¹Chemistry and ²Biology, Syracuse University, Syracuse, NY 13210, USA

Received 15 April 1985

ABSTRACT

A system comprised of a linear scanning microdensitometer interfaced to a personal computer has been developed to facilitate analysis of ligand-DNA footprinting autoradiograms.[†] The system, which can be used to record density and sequence information from autoradiographic films, enables the user to relate the area under an autoradiographic band to the concentration of radiolabeled molecules present in the electrophoresis gel. This report describes the computer program which performs the calculations and discusses the ability of the system to accurately determine oligonucleotide concentration, as a function of band separation, photographic response, and the computational algorithm used to calculate band areas.

INTRODUCTION

The identification of binding sites of equilibrium binding ligands bound to heterogeneous DNA's of defined sequence has been made possible by the development of footprinting methods (1-7). Footprinting analysis involves partial digestion of a uniquely radiolabeled DNA molecule in the presence and absence of a DNA binding ligand. If the ligand is able to inhibit DNA cleavage at the site of binding, the locations of ligand-DNA contact can be identified by analyzing the digest products using gel electrophoresis and autoradiography. In addition, footprinting analysis can, under the proper conditions, be used to evaluate ligand binding constants as a function of the sequence of DNA bases. This procedure entails a series of experiments involving digestion of DNA at various values of the ligand/DNA base-pair ratio, r_t . Thus, electrophoretic analysis of the digest products using high resolution polyacrylamide gel electrophoresis followed by autoradiography results in a two dimensional array of density data. The electrophoresing dimension corresponds to sequence while the second dimension, perpendicular to the direction of electrophoresis, shows the distribution of digest products as a function of

[†]The source program will be provided for a small fee to all interested parties.

Microdensitometer – Computer System

Figure 1. Block-diagram portraying the microdensitometer-computer system and the attendant software used in the analysis of footprinting autoradiograms.

r_L . With the proper experimental controls, densitometric measurements on the autoradiographic bands can be used to obtain the fraction of a particular site which is occupied by ligand, v_b (8). With knowledge of v_b and a quantity which is not directly measured in a footprinting experiment, namely the free ligand concentration, footprinting data can be used to construct a binding isotherm which allows evaluation of a ligand's binding constant (8,9).

Although linear scanning microdensitometry is often used to analyze footprinting autoradiographic data, few details concerning this technique have been published. In this report we describe a microdensitometer-computer system and supporting software capable of scanning and quantitatively analyzing footprinting autoradiograms. A representation of this system is shown in block-diagram form in figure 1. In addition to calibrating the system, we also examine its ability to determine concentrations of ^{32}P -labeled oligo-

nucleotides present in a polyacrylamide gel matrix from the corresponding autoradiogram. Aside from direct application to footprinting experiments, the system described is useful in a variety of biochemical applications requiring quantitative analysis of large amounts of autoradiographic data.

HARDWARE

The data acquisition system for the linear scanning densitometer (Jarrell Ash Model 23-100) was based around a Nippon Electric Corporation (NEC) Advanced Personal Computer. The display on the graphic screen could be transferred to a Hewlett-Packard plotter (model 7475A) or to an NEC dot-matrix printer (model PC-8023A-C) to provide a "hard copy" at the various stages of the analysis.

The normal signal-processing electronics of the densitometer were bypassed and a current to frequency converter (I-F) was connected directly to the output from the phototube. This produced a train of pulses with a frequency proportional to the intensity of the light striking the phototube. A 68705P3 microprocessor was employed to count the pulses produced by the I-F converter and send the values via an RS-232 serial interface link to the computer. The timing in the microprocessor was adjusted such that a value to the computer was sent every 280 milliseconds (approximately 3.5 values per second). Thus, each value represented the integral of the response from light striking the phototube in the previous 280 milliseconds. The high voltage for the phototube was supplied by the normal power supplies within the microdensitometer. This supply was set by the operator just prior to scanning by adjusting the sensitivity settings of the microdensitometer, and remained constant throughout the scanning period.

SOFTWARE

The software for the computer was written in Digital Research Personal Basic and stored on floppy-disks. The computer program was functionally divided into the following sections: data acquisition and smoothing, peak identification, baseline establishment, peak integration, and processed data output.

Data Acquisition

The data acquisition section of the program processed the values sent from the microprocessor, corrected them for the dark-current offset and converted them to absorbance units (from percentage transmission units). The signal output was a series of transmittance values recorded as a function of

position on the autoradiographic film.

The interactive data-acquisition routine involved, on the operator's part, single-key operations on the keyboard to instruct the computer when to take the dark-current reading. The sensitivity settings on the densitometer were then adjusted as the light beam traversed a segment of processed x-ray film having no autoradiographic images. The light beam was then blocked with an opaque object to obtain the dark current reading. All subsequent absorbance data were measured relative to this value.

The computer displayed the data on its graphics screen as it was acquired to allow the operator to monitor the scanning process. At the end of a scan, the computer stored the data in a disk file for future reference and entered the data analysis section of the program. This section retrieved the data from the disk file, and initially passed it through a nine point gaussian smoothing function to remove data acquisition related noise without measurably degrading the desired peak data.

Peak Identification

The peak identification section of the program allowed the operator to label the peaks in the scanned autoradiogram. If required, the computer could move the cursor indicating the peak location left or right up to ten increments to locate the maximum absorbance value, A_m , for a given peak. This allowed the operator to assign the position of the peak without having to locate its exact maximum. The automatic maximum search subroutine of the program could be disabled to allow the marking of shoulders and poorly resolved peaks. The operator assigned numbers to any or all of the peaks and the computer used these numbers as reference points for later calculations. The length of an oligonucleotide and its sequence was in turn determined through employment of DNA sequencing protocols on the radiolabeled fragment (10). The operator could mark all peaks in the display or, after a minimum of two peaks were marked, have the computer do an automatic peak identification. The automatic method relied on comparison of a scanned autoradiogram with an averaged set of previously scanned "library" autoradiograms of the same series of oligonucleotides using the same electrophoresing conditions. After the left-most and right-most peaks of the scanned autoradiogram were assigned by the operator and matched with the identical peaks in the library autoradiogram, the remaining peaks were assigned by the computer by comparison to the corresponding peaks in the library autoradiogram. The precise location of absorbance maxima were treated in the same way as operator-marked peaks with automatic position adjustment for maximum absorbance if that mode was enabled.

Once the computer assigned these peak positions, the operator could edit the peak assignments if so desired.

Baseline Establishment

The establishment of a baseline for a scan was mandatory for any peak area calculations. The assignment of a baseline for a particular scan involved, on the part of the operator, the choice as to which points of the spectrum would be labeled as "nodes" - or points for which it was reasonably certain lay on the "true" baseline. On the other hand, a single-key default instruction to the computer was occasionally used to signal it to assign the first and last point of the scan as nodes and to draw a vector between these two points -which then became the computed baseline for the scan.

Although only two nodes are necessary for the construction of a baseline, it was often desirable to use several nodes for baseline assignment. Ideally, the local minimum between each pair of peaks could be assigned as a node. For scans containing several ill-resolved peaks, however, the minima between these peaks could not sensibly be thought to lie on the true baseline.

For any scan, the first and last minima between bands were assigned as nodes by having the cursor on the screen moved to the required position and depressing a single-function key which instructed the computer to register that point as a node through which the baseline vector would pass. Additional nodes were assigned in a similar fashion. The criterion for these, however, required that the local minimum chosen as a node rested between two well resolved peaks separated by at least several peak-halfwidths. Again, in such a case, it was reasonably certain that this local minimum represented a true baseline point. If desired, the computer flattened the baseline by fitting the assigned baseline nodes to an equation of the form $a + bx + cx^2$. The coefficient "c" was generally close to zero.

The computer was then instructed to subtract the calculated baseline from the recorded scan and to plot the recalculated scan with the baseline flattened on the computer graphics screen complete with identification markers for the peaks and the assigned baseline nodes.

Peak Integration

Peak integration was achieved using the "trapezoidal rule" (11). The separations of each consecutive pair of peaks were calculated, and the average of these values was used as a typical peak width. Then, starting at the center position for each peak, the computer found the minimum absorbance value on each side of the center, up to a distance equal to the average peak width. This allowed for peaks up to twice as wide as the average width. The minimums

Figure 2. Three dimensional type hidden-line plot illustrating the juxtaposition of linear scans of adjacent lanes on a footprinting autoradiogram. The data was taken from a footprinting experiment involving actinomycin-D and a self-complementary dodecanucleotide duplex.

found were used as the start and end points of the integration. The calculated values were subsequently stored in a file and were also displayed on the screen.

Processed Data Output

The results of the peak analysis were re-presented on the graphics screen after each series of calculations to allow the operator to monitor the process. In addition, all the information displayed on the screen could be transferred to a dot matrix printer to obtain a hard-copy at various stages of the analysis, accompanied by a listing of the calculated peak areas if desired.

Alternatively, the original data could be directed to the HP-7475A plotter to obtain a higher quality printout. Furthermore, a provision existed enabling a pseudo three dimensional plot of a series of densitometrically obtained data to be juxtaposed, in the hidden-line format, on the same graph. Such plots displayed the variables of Absorbance, Sequence or Distance from the Origin on the Gel and drug-DNA base pair ratio, r_t or r_t^{-1} . A typical three dimensional plot is shown in figure 2.

SYSTEM CALIBRATION

Reproducibility

The bands appearing on footprinting autoradiograms were rectangular in shape and had a constant length (also referred to as the long dimension) of about 15 mm perpendicular to the direction of electrophoresis. The width of these bands (also referred to as the short dimension) parallel to the direction of electrophoresis, however, were found to be dependent on two factors. Firstly, the concentration of radiolabeled oligonucleotide within the polyacrylamide gel that gave rise to the photographic image and secondly, the distance the oligonucleotide migrated from the origin, or loading well, of the gel. In examining the reproducibility of the scanning process, a series of well isolated bands having absorbance maxima in the range $0.05 < A_m < 1.20$, were scanned by passing the densitometer beam repetitively (10 times) through the width of the band. The disparities between similar scans were small, with the largest standard deviation of approximately 0.002 being associated with the weakest bands.

Density Distribution Within a Given Band

The optical density distribution within a band was examined by linear scanning along both the long and short dimension of the band. Using slit dimensions of $2\mu \times 0.1$ mm and with the long dimension of the slit perpendicular to the direction of scanning, the absorbance versus distance curve for a scan through the center of a typical band along its long dimension is shown in figure 3. It was demonstrated that within the central portion of most bands, the absorbance maximum exhibited only a small variation, $\pm 1\%$, about a mean value. In some extreme cases, this variation was found to be as large as $\pm 3\%$. Upon scanning through the narrow dimension of a band at various points along its length, it was found that each scan was Gaussian in shape.* Since the absorbance maxima for all of the passes through the central portion of the band were constant, the optical density distribution along the band's long axis was probed by measuring the band widths at half height, $W_{1/2}$, for each of the various scans through the band's short dimension. As shown in figure 3, $W_{1/2}$ (in arbitrary units) was invariant through the central portion of the band indicating that the cross sectional areas produced by linear scanning in this region of the band are constant.

(* From computer fitting studies of the absorbance curves, it was found they were best described by a Gaussian rather than Lorentzian density distribution.)

Area Calculation as a Function of Peak Separation

In order to ascertain the precision with which the densitometer-computer

Figure 3. (a). The upper portion of the figure depicts a plot of the ratio of the peakwidth at half-height ($W_{1/2}$) to the amplitude (A_m) obtained by scanning through the narrow dimension of the band. These data are presented as a function of the distance along the long dimension of the band portrayed in part (b) of the figure.

system could determine the areas of autoradiographic bands, a study was undertaken in which the calculated peak areas for a pair of bands was correlated with the distance separating these bands on the film and also as a function of the relative intensity of one band with respect to the other. Four separate bands from different regions of an autoradiogram were cut out of a film as rectangular strips. The bands had absorbance maxima, A_m , and widths at half height, $W_{1/2}$, (in parentheses) of: band A, 0.365(0.078); band B, 0.223(0.068); band C, 0.191(0.084); and band D, 0.095(0.150).

Three pairs of bands, namely A-D, B-C, and B-D were chosen to study the integrity of the peak area calculations as a function of the band-to-band separation, $d(\text{mm})$. Two sections of autoradiographic film, each containing the appropriate band, were superimposed and subjected to densitometry. The overlaid film strips were repositioned after each scan to decrease the distance between the band centers. The "true" area under any band was that determined at a peak-to-peak separation of 15 mm. The ratio of the area of a band in the film overlay (a_x) to the "true" area of that band (a_0) as a function of the band separation for the pairs A-D, B-C and B-D is presented in figure 4.

Oligonucleotide Concentrations From Scanned Autoradiographic Data

The ability of the system to determine the relative concentrations of

Figure 4. A plot of the ratio of the calculated area of a band (a_x) to the "true" area (a_0) as a function of the peak-to-peak separation ($d(\text{mm})$) for the pairs of bands defined in the text.

- (a) Peak pair A-D. Amplitude ratio 3.8:1.0
- (b) Peak pair B-D. Amplitude ratio 2.3:1.0
- (c) Peak pair B-C. Amplitude ratio 1.2:1.0

oligonucleotides present in the gel was also examined. Various concentrations of different ^{32}P singly end-labeled oligomers (8 to 12 base pairs in length) were loaded into wells (10 mm wide) at the top of a 20%, 0.3 mm thick polyacrylamide gel (12) and electrophoresed. After electrophoresis, the gel was autoradiographed for periods of 2 and 10 hours using Kodak X-Omat AR x-ray film at -20°C without preflashing or the use of an intensifying screen. The bands were a uniform ~ 15 mm in length but were of varying width depending on the concentration of oligomer present in the gel, the exposure time and the distance the oligomer had migrated from the origin. These bands were scanned in the direction of electrophoresis to determine their areas and absorbance maxima, A_m . The values of the absorbance maxima for the bands on the 2 hr autoradiograph (lower curve) were $0.12 \leq A_m \leq 0.95$, while those for the 10 hr exposure (upper curve) fell in the range $0.11 \leq A_m \leq 1.58$. Following autoradiography, the bands were excised from the gel, counted using a scintillation counter, and plots of the scanned area versus number of counts for the two exposure times were constructed. As is shown in figure 5, a linear relationship exists between scanned area and counts in the absorbance range $0 < A_m < \sim 1.1$.

DISCUSSION

The computer-assisted microdensitometric system described herein has proven invaluable in the rapid analysis of footprinting autoradiograms.

Figure 5. Plots of the integrated peak areas determined densitometrically versus the counts/min evaluated by scintillation counting for ^{32}P labeled DNA restriction fragments. The nonlinear (plateau) region corresponds to $A_m > 1.0$ D unit.

The upper and lower curves represent data taken from film exposures of 10 and 2 hrs. respectively.

Densitometric scans of adjacent lanes in an autoradiogram of a footprinting electrophoretic gel can be recorded, compared and assessed for subtle changes in the enzymatic digestion pattern. Given the proper conditions, the system can be used to quantitatively ascertain the concentration of radiolabeled oligonucleotides present in the electrophoretic gel. Such information is necessary for the construction of classical binding isotherms from footprinting data (8).

When linear scanning as opposed to "whole-area" scanning (9) is employed to analyze autoradiographic data, it is critically important that the bands being compared have similar shapes and density distribution within the shape. If these criteria are met, scanning yields areas which are proportional to the concentration of molecules in the gel giving rise to the photographic image. Furthermore, if bands are rectangular in shape having one dimension constant, the case with DNA sequencing gels, the areas determined by linear scanning are essentially independent of the location of the scanning path of the densitometer beam through the variable dimension of the band. This condition obviates the need to precisely locate the scanning path of the densitometer beam through a band when areas (concentrations) between bands in different lanes on the gel are to be compared.

Although the simple trapezoidal algorithm is useful for calculating the areas of well-resolved peaks, problems arise when significant peak overlap occurs. It is evident from figure 4 that a sizeable overestimation of the area of a peak with a small area occurs as it nears a peak having a larger area. In the case of two peaks of similar magnitude, the respective peak areas are both overestimated as the band-to-band separation is diminished. The data in figure 4 can be used as correction curves to obtain more accurate estimates of peak areas as a function of their relative areas and their separation distance. Alternatively, the problem can be circumvented by deconvolving ill-resolved peaks by fitting Gaussian curves to the data. However, this alternative algorithm is impractical for footprinting titration experiments involving data from restriction fragments where 80 to 100 bases are represented on the autoradiogram.

ACKNOWLEDGMENTS

We are grateful to Drs. B. Ware and L. Plank for helpful discussions and to E. L. Fish, R. Reh fuss, M. J. Lane and L. Vriesinga for assistance and advice throughout the course of this work. The work described herein was supported by a grant from the National Institutes of Health, GM31895. One of us (J. C. D.) is grateful to the American Cancer Society for a Scholar in Cancer Research Award for 1985.

*Author to whom correspondence should be addressed.

REFERENCES

1. Lane, M. J., Dabrowiak, J. C., and Vournakis, J. N. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3260-2364.
2. Dabrowiak, J. C. (1983) *Life Sci.* **32**, 2915-2931, and references therein.
3. Fox, K. R., Waring, M. J. (1984) *Nucleic Acids Res.* **12**, 9271-9285.
4. Lane, M. J. Vournakis, J. N. and Dabrowiak, J. C. (1985) in *Molecular Basis of Cancer, Part B, Macromolecular Recognition, Chemotherapy and Immunology*, Ed. R. Rein, 145-154 Alan R. Liss, New York.
5. Van Dyke, M. W. and Dervan, P. B., (1984), *Nature (London)*, **225**, 1122-1127.
6. Vary, C. P. H., Lane, M. J., Dabrowiak, J. C. and Vournakis, J. N., (1984), in *Electrophoresis '84*, 312-314, Ed. Neu hof f, V., Verlag Chemie, Weinheim.
7. Ackers, G. K., Johnson, A. D. and Shea, M. A. (1982), *Proc. Natl. Acad. Sci. USA* **79**, 1129-1133.
8. Goodisman, J. and Dabrowiak, J. C. (1985) *J. Biomol. Str. and Dynamics*, **2** (1985) 967-980.
9. Borowitz, M., Senear, D. F., Shea, M. A. and Ackers, G. K. (1985) *Methods in Enzymology*, in press.
10. Maxam, A. and Gilbert, W. (1980) *Meth. Enzymology* **65**, 499-559.
11. Carnahan, B., Luther, H. A. and Wilkes, J. O. (1969) *Applied Numerical Methods*, John Wiley and Sons, New York.
12. Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) *Molecular Cloning- A Laboratory Manual*, Cold Spring Harbour Laboratory, p. 173.