CHROMOSOMAL PROTEIN-DNA INTERACTIONS

Thesis by

Robert Francis Murphy

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1980

(Submitted July 17, 1978)

ACKNOWLEDGMENT

To Dr. James Bonner, for providing me with the means and opportunity to study whatever seemed interesting;

To Drs. Herschell Mitchell, Elias Lazarides, Jim Strauss, and Jack Richards, for agreeing to oversee this work;

To Dr. R. Bruce Wallace, for his invaluable help and advice;

To the members of the Bonner group, for their cooperation;

To my teammates on various and diverse athletic teams, for making life in this one-horse town a little less unbearable;

To Steve and Bonnie Petersen, for being my friends;

And to my parents and family, with all my love.

The preparation costs of this thesis were defrayed in part by the Weigle Fund.

ABSTRACT

The nature and function of some of the protein-DNA interactions in eukaryotic chromatin were investigated. The nucleosome structure of isolated template-active chromatin was determined. In vitro chemical acetylation of chromatin was shown to result in a structure similar to that of deproteinized DNA, which represents a shift towards the properties of isolated template active chromatin, and chromatin containing specific transcribed genes.

The process of chromatin replication was shown to include a shortening of the internucleosomal spacer, resulting in decreased nuclease sensitivity. Newly-replicated chromatin was separated from bulk chromatin in shallow metrizamide density gradients. Newly-synthesized histone and newly-acetylated protein were shown to be present predominantly in the unreplicated chromatin fraction.

The accuracy and reproducibility of non-linear least squares determinations of the thermal denaturation transitions of DNA and chromatin were determined using computer programs designed for ease of use and adaptability to mini-computer configurations. Direct fitting of melt data to a normalized error function gave results very similar to those obtained by fitting Gaussian curves to derivatized data. This approach avoids errors introduced by the derivatization method, and requires fewer data points.

CONTENTS

PART I

Chapter 1. Alkaline extraction of non-histone proteins from rat liver chromatin

Reprinted from

Biochimica et Biophysica Acta, 405 (1975) 62–66

BBA 37139

ALKALINE EXTRACTION OF NON-HISTONE PROTEINS FROM RAT LIVER CHROMATIN

ROBERT F. MURPHY and JAMES BONNER

Division of Biology, California Institute of Technology, Pasadena, Cal. 91125 (U.S.A.)

(Received March 18th, 1975)

SUMMARY

The dissociation of non-histone proteins from chromatin by alkaline pH was studied using sodium dodecyl sulfate-polyacrylamide gel electrophoresis. A significant degree of selectivity of dissociation was observed. This selectivity may provide a new method for fractionating non-histone proteins.

INTRODUCTION

Much work has been done on the selective dissociation of histones from nucleohistones at low pH [1–3]. Other investigations have reported on the dissociation of histones from DNA at high pH, particularly at pH 12 and above [4–7]. In this paper we describe experiments on the selective dissociation of chromosomal proteins by titration in the range pH 8–11.

MATERIALS AND METHODS

Preparation of chromatin. Chromatin was prepared from Sprague-Dawley rat livers (Pel-freeze), by the sucrose purification method of Marushige and Bonner [8]. After pelleting through a 1.7 M sucrose cusion, the chromatin was resuspended in four volumes of 10 mM sodium bisulfite, 10 mM Tris buffer, pH 8.0, and spun at 12 000 × *g* for 10 min. The chromatin pellet was resuspended at a concentration of 20 A_{260} nm units per ml and sheared for 90 s at 30 V in the Virtis homogenizer. The sheared chromatin was spun at 12 000 × *g* for 20 min to remove a small amount of material not solubilized by shearing.

Alkaline dissociation of chromatin. 10 ml of sheared chromatin was titrated to the desired pH (using a Radiometer 22 pH meter) by the dropwise addition of 1 M NaOH with stirring (pH values were checked by warming a small sample of the material to 25 °C and determining the pH again). The chromatin was then stirred for 30 min and centrifuged at 45 000 rev./min for 20 h in a Beckman Ti50 rotor (134 800 × *g*). The chemical composition of the material in the supernatants and pellets was determined, and the proteins of each sample were analyzed by polyacrylamide gel electrophoresis.

Chemical composition. RNA and DNA were separated by the Schmidt-

Tannhauser procedure. The amount of DNA was determined from the absorption at 260 nm and RNA was determined by the orcinol procedure [9]. Total protein was determined by the procedure of Lowry et al. [10] using bovine serum albumin as standard. Histones were extracted with 0.2 M H_2SO_4 and the insoluble non-histones removed by centrifugation at 12 000 × *g* for 10 min. Pellets from this extraction were resuspended in 1 M NaOH and the amount of non-histone proteins determined by the procedure of Lowry et al. [10], while the amount of histone protein was determined from the absorption at 230 nm using the relation 1 mg/ml = 4.15 $A_{230\ nm}$ units.

Polyacrylamide gel electrophoresis. Sodium dodecyl sulfate-10% acrylamide gels were run, stained and scanned as described previously, and the scans were resolved into a number of Gaussian curves such that the sum of those curves equaled the original data [11, 12].

RESULTS

Table I presents the results of protein and nucleic acid determination for the supernatants and pellets of chromatins titrated to pH 8–11 and pelleted as described above. As can be seen, more than 95% of the DNA is recovered in the pellets at all pH values. Increasing amounts of RNA, histone and non-histone are dissociated from chromatin as the pH is increased. The greatest increase in the release of all three components occurs between pH 10 and pH 11. This may be due to the titration of lysine residues (p*K* approx. 10.5). Indeed, in the titration of chromatin to pH 11, the buffering effect of the lysine residues is clear as evidenced by the need for a greater amount of NaOH to raise the pH from 10 to 11 than is needed to raise the pH from 8 to 9 or from 9 to 10. There is also a noticeable rise in viscosity of the preparation above pH 10 perhaps correlated with the dissociation of protein from DNA in some regions of the chromatin and concomitant unfolding of DNA.

To discover whether particular non-histone proteins are released as the pH of chromatin is increased, we have run sodium dodecyl sulfate gels on the various fractions. Fig. 1 shows the composite gel scans of the supernatants (dissociated material) at pH 8–11. Each Gaussian does not necessarily represent one peptide, but may represent one or more peptides of identical or similar molecular weight.

TABLE I

DISSOCIATION OF CHROMATIN CONSTITUENTS BY ALKALINE pH

Chromatin isolation, titration and centrifugation were as described in Materials and Methods. Amounts are expressed as percent of starting material of each component ± one standard deviation. Ratio given is the weight ratio of each component to DNA in whole chromatin. S, supernatant (dissociated); P, pellet (undissociated).

pH	DNA			RNA			Histone			Non-histone		
	S	P	±	S	P	±	S	P	±	S	P	±
8	1	99	0.5	12	88	5	3	97	1	5	95	3
9	1	99	0.7	26	74	9	3	97	1	8	92	3
10	2	98	0.8	36	64	10	7	93	1	12	88	4
11	2	98	0.8	55	45	10	18	82	3	22	78	6
Ratio:	1.0			0.03			1.1			0.8		

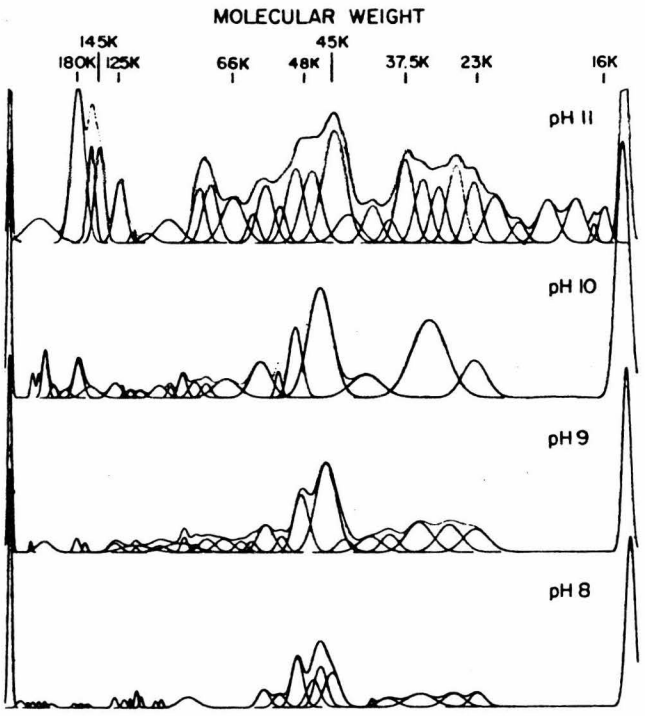


Fig. 1. Composite gel scans of protein dissociated at several alkaline pH values. Samples of 100 μg protein from each supernatant were separated on 10-cm sodium dodecyl sulfate gels, and then the absorbances normalized to represent equivalent amounts of starting chromatin. Protein to DNA ratios are as follows: pH 8, 0.07; pH 9, 0.09; pH 10, 0.17; pH 11, 0.35.

In the first place there is a small general increase in release of material from chromatin as the pH is increased from 8 to 11. In addition, however, there is a significant increase in the release of particular protein fractions over this pH range. This increase in dissociated material is particularly marked in the high (140 000–180 000) molecular weight and low (22 000–43 000) molecular weight regions of the gel.

Since mobility in these gels is proportional to log molecular weight, the molecular weight of each band can be determined using established standards [12]. The area under each peak is proportional to the mass of protein in that component. Therefore, the number of molecules in each band can be calculated. Table II presents the data in terms of molecules per genome ($3 \cdot 10^9$ base pairs) and in percent of total number of molecules per genome for each molecular weight class [12]. The fact that some classes of proteins are dissociated to a significantly higher degree or to a significantly lesser degree than the percent of total protein dissociated at that pH, is evidence of selectivity of dissociation.

Proteins which are dissociated at pH 8 (the pH at which the chromatin is prepared) may be released by the shearing process (no control for this is possible since unsheared chromatin is an aggregate of very high molecular weight). In addition, the proteins dissociated at pH 8 evidently have a very low affinity for chromatin and may indeed be bound to chromosomal proteins rather than to DNA itself.

TABLE II

DISSOCIATION OF NON-HISTONE PROTEINS BY ALKALINE pH

Results of calculations on data in Fig. 1. Total number of molecules column gives data from Garrard et al. [12] for number in each class per haploid genome ($1.8 \cdot 10^{12}$ mol. wt).

Molecular weight ($\times 10^3$)	Total number of molecules ($\times 10^3$)	Percent of total			
		pH 8	pH 9	pH 10	pH 11
>195	46	13	17	57	57
171–190	101	1	5	19	106
150–170	122	2	2	9	35
140–149	92	0	0	0	64
130–139	91	0	0	0	0
100–129	554	3	6	5	11
92–98	122	0	0	0	12
80–90	450	1	3	6	14
68–79	1460	2	4	4	7
65	690	0	5	3	18
56–63	1005	0	6	8	16
50–55	4790	1	2	3	5
48	660	5	5	5	13
45	1950	7	11	10	14
31–43	1620	23	34	41	70
23–30	2876	15	38	62	98
<22	6708	0	0	0	28
Total	22582	5	8	12	22

One of the principal proteins released at pH 8 is a major non-histone and one which appears to be identical with the muscle protein, actin. It is possible that F actin is attached or bound to chromatin at one point and protrudes from chromatin as an F actin fiber, and is thus easily released by hydrodynamic forces.

DISCUSSION

The significance of our findings is 2-fold. First, we have established that the non-histone proteins of rat liver chromatin are subject to selective dissociation at alkaline pH. Second, our results indicate that it should be possible to enrich for a rather specific subset of non-histone chromosomal proteins by first titrating chromatin to pH 10, removing the dissociated material and then titrating the chromatin again to pH 11 and removing the dissociated material. This procedure would provide still another method of fractionating and characterizing the non-histone chromosomal proteins.

ACKNOWLEDGMENTS

We wish to thank Dr Angeline Douvas, Mr William R. Pearson and Mr Joel Gottesfeld for advice and helpful counsel, as well as Ms Ryn Miake for technical assistance.

REFERENCES

1 Johns, E. W. and Butler, J. A. V. (1962) Biochem. J. 82, 15–18
2 Murray, K. (1966) J. Mol. Biol. 15, 409–419
3 Murray, K., Vidali, G. and Neelin, J. M. (1968) Biochem. J. 107, 207–215
4 Matsuyama, A., Tagashita, Y. and Nagata, C. (1971) Biochim. Biophys. Acta 240, 184–190
5 Ramm, E. I., Vorob'ev, V. I., Birshtein, T. M., Bolotina, I. A. and Volkenshtein, M. K. (1972) Eur. J. Biochem. 25, 245–253
6 Ohba, Y. and Hayashi, M. (1972) Eur. J. Biochem. 29, 461–468
7 Russev, G., Venkov, C. and Tsavev, R. (1974) Eur. J. Biochem. 43, 253–256
8 Marushige, K. and Bonner, J. (1966) J. Mol. Biol. 15, 160–174
9 Dische, Z. and Schwartz, K. (1937) Mikrochim. Acta 2, 13
10 Lowry, O. H., Rosebrough, N. J., Farr, A. L. and Randall, R. J. (1951) J. Biol. Chem. 193, 265–275
11 King, J. and Laemmli, V. K. (1971) J. Mol. Biol. 62, 465–477
12 Garrard, W. T., Pearson, W. R., Wake, S. K. and Bonner, J. (1974) Biochem. Biophys. Res. Commun. 58, 50–57

PART II

Chapter 2. Structure of transcriptionally active chromatin

Reprinted from
Proc. Nat. Acad. Sci. USA
Vol. 72, No. 11, pp. 4404–4408, November 1975
Biochemistry

Structure of transcriptionally active chromatin

(chromatin fractionation/deoxyribonuclease II/chromatin subunits/ν-bodies)

JOEL M. GOTTESFELD*, ROBERT F. MURPHY, AND JAMES BONNER

Division of Biology, California Institute of Technology, Pasadena, Calif. 91125

Contributed by James Bonner, September 10, 1975

ABSTRACT **Rat-liver chromatin has been fractionated into transcriptionally active and inactive regions [Gottesfeld *et al.* (1974) *Proc. Nat. Acad. Sci. USA* 71, 2193–2197] and the distribution of nuclease-resistant complexes in these fractions has been investigated. About half of the DNA of both fractions is resistant to attack by the endonuclease DNase II. The nuclease-resistant structures of inactive chromatin are DNA–histone complexes (ν-bodies) which sediment at 11–13 S. Template-active chromatin yields two peaks of nuclease-resistant nucleoprotein. These complexes sediment at 14 and 19 S, and contain DNA, RNA, histone, and nonhistone chromosomal proteins. Polyacrylamide gel electrophoresis reveals a complex pattern of chromatin proteins, suggesting that the complexes are heterogeneous in composition.**

A regular repeating unit in chromatin was first suggested from the x-ray diffraction studies of Pardon *et al.* (1): a series of reflections were observed in the x-ray patterns of native and reconstituted nucleohistones, but not in the x-ray diffraction patterns of DNA or histones by themselves. It was proposed that the chromatin fiber is organized into a regular supercoil of pitch 100–120 Å. This model, although widely accepted for some time, has now come under question. Olins and Olins (2) have observed regular spacings of chromatin particles (termed ν-bodies) in water-swollen nuclei centrifuged onto electron microscope grids. These particles are 60–80 Å in diameter and are joined by thin filaments 15 Å in diameter. These results have been confirmed and extended by other laboratories (3–5). Nuclease digestion studies also support a subunit or particulate structure for chromatin. Both endogenous (6–7) and exogenous (8–14) nucleases appear to recognize a repeating nucleoprotein unit along the chromatin fiber. Furthermore, chromatin particles isolated from nuclease-treated or sonicated chromatin resemble ν-bodies in the electron microscope (15–17). Thus many lines of evidence support the subunit or "beads-on-a-string" model of chromatin structure (2, 18).

Studies on chromatin structure have been generally carried out with whole chromatin, unfractionated with respect to transcriptional activity. Since only a minor portion of the DNA in any differentiated cell type is ever transcribed into RNA, the properties of unfractionated chromatin reflect primarily the structure of inactive regions. We are interested in whether template-active chromatin is organized as is the inactive region or whether it is in a different conformation. Previous work from this laboratory has shown that it is possible to separate chromatin into transcriptionally active and inactive fractions (19–21). In this communication we report that both fractions contain nuclease-resistant nucleoprotein complexes. However, the nuclease-resistant structures of inactive chromatin are DNA–histone complexes, while the nuclease-resistant structures of active chromatin are complexes of DNA, RNA, histone, and nonhistone chromosomal protein.

Abbreviation: DFP, diisopropylfluorophosphate.

* Present address: MRC Laboratory of Molecular Biology, Hills Road, Cambridge, England CB2 2QH.

MATERIALS AND METHODS

Chromatin Fractionation. Chromatin was prepared from rat liver by the method of Marushige and Bonner (22) and treated with diisopropylfluorophosphate (DFP) to inhibit endogenous protease activity (23). Fractionation was carried out as diagrammed in Fig. 1; details of this method have been published previously (21).

Preparation of Chromatin Subunits. Nuclease-resistant subunits of rat-liver chromatin were prepared as follows: DNase II was added to 10 units per A_{260} unit of chromatin in 25 mM sodium acetate (pH 6.6). Digestion was carried out at 24° and was terminated after 90 min by raising the pH to 7.5 with 0.1 M Tris·HCl (pH 11). Nuclease-resistant subunits from chromatin fraction P1 were prepared by homogenizing the pellet fraction in 25 mM sodium acetate (pH 6.6) and redigesting with DNase as described above for whole chromatin. Undigested chromatin (about 20% of the input DNA) was removed by centrifugation at 27,000 × *g* for 10–15 min. The supernatant was layered on isokinetic sucrose gradients in SW25.1 cellulose nitrate tubes. The gradients were formed according to Noll (24); the parameters were C_{TOP} = 15% (weight/volume), C_{RES} = 34.2% (weight/volume), and V_{MIX} = 31.4 ml. All solutions contained 10 mM Tris·HCl (pH 8). Centrifugation was at 25,000 rpm for 36–42 hr. Gradients were analyzed with an ISCO UV Analyzer and chart recorder. Fractions from these gradients were rerun on 5–24% isokinetic sucrose gradients. The parameters were C_{TOP} = 5.1% (weight/volume), C_{RES} = 31.4% (weight/volume) and V_{MIX} = 9.4 ml. Centrifugation was in the SW 41 rotor at 39,000 rpm at 4° for 16.5 hr.

Subunits were also prepared from chromatin devoid of histone I. Removal of this histone was accomplished by extraction of Virtis-sheared chromatin (45 V, 90 sec) with 0.5 M NaCl at 4°. The resultant nucleohistone was pelleted by centrifugation in the Ti 50 rotor at 50,000 rpm for 18 hr. The pellet was digested with nuclease as described above.

Redigestion of Chromatin Fraction S2. Chromatin of fraction S2 was redigested with nuclease in three different ways: the DNase II present in fraction S2 from the first nuclease treatment was reactivated by adding EDTA to 20 mM and lowering pH to 6.4 with dilute HCl. Alternatively, aliquots of chromatin fraction S2 were dialyzed against either 25 mM sodium acetate (pH 6.6) or 5 mM sodium phosphate (pH 6.7) containing 2.5×10^{-4} M $CaCl_2$ and 2.5×10^{-4} M $MgCl_2$. DNase II was added to the chromatin in sodium acetate buffer to 10 units/A_{260} unit of chromatin; staphylococcal nuclease was added to the chromatin in sodium phosphate buffer to 50 units/ml. Reactions were carried out at

24°. Aliquots were taken at various times to test for the production of trichloroacetic-acid-soluble material (measured by absorbance of the supernatant at 260 nm after centrifugation at 27,000 × *g* at 4° for 15 min).

DNA Size Estimation. Single-strand DNA lengths were estimated by velocity sedimentation in alkaline sucrose gradients. The parameters for the isokinetic gradients were C_{TOP} = 15.9% (weight/volume), C_{RES} = 38.9% (weight/volume), and V_{MIX} = 6.1 ml (24). All solutions contained 0.1 N NaOH. Chromatin samples were suspended in 0.1 N NaOH, 2% sodium dodecyl sulfate, 2 M urea, and 100–200 μl aliquots were layered on each gradient. Centrifugation was in the SW50.1 rotor at 48,000 rpm for 16 hr at 20°. DNA molecular weights were determined relative to a standard sized by electron microscopy (320 nucleotide-long, 5.4 S, calf thymus DNA; a gift of Ms. M. Chamberlin). Double-strand lengths were determined in the analytical ultracentrifuge (25).

Analysis of Chromatin Composition. Histone and nonhistone protein content was determined as described (26). Protein was analyzed by sodium dodecyl sulfate-disc gel electrophoresis (27) and by acid-urea gel electrophoresis (28). DNA and RNA were determined by the methods of Schmidt and Tannhauser (29).

Enzymes. DNase II (EC 3.1.4.6) and micrococcal nuclease (EC 3.1.4.7) from *Staphylococcus* were purchased from Worthington.

RESULTS

Chromatin Fractionation. DNase II preferentially attacks a select portion of chromatin DNA. The amount of DNA in this fraction varies depending upon the source of the chromatin, but corresponds quite closely to the measured template activity of the particular chromatin (20). The fractionation scheme used herein is diagrammed in Fig. 1. After 5 min exposure to DNase II, 15% of rat-liver chromatin DNA remains soluble after centrifugation (fraction S1). About 11% of the total DNA is Mg^{++}-soluble and is found in fraction S2. This DNA comprises a subset of whole genomal DNA sequences and is enriched 5-fold in transcriptionally active sequences (21). The DNA has a double-strand length of about 700 base pairs and a single-strand length of 200–600 nucleotides (range of observed values). About 1–3% of this DNA is acid-soluble. After 30 min exposure to DNase II, nearly 80% of the chromatin is found in fraction S1, and 20–24% is found in fraction S2. After prolonged nuclease digestion roughly half of the DNA of both fraction S1 and fraction S2 is acid-soluble. A more detailed description of the kinetics of DNase II action on chromatin is presented elsewhere (J. Gottesfeld, G. Bagi, B. Berg, and J. Bonner, manuscript submitted).

Table 1 lists some of the properties of the chromatin fractions: the composition of fraction P1 is similar to that of unfractionated rat-liver chromatin (21, 22). Fraction S2, however, is enriched in RNA and nonhistone protein and has a reduced content of histone protein. All the major histone species are present in fraction S2 isolated from DFP-treated chromatin. In a previous communication (21) we reported that fraction S2 lacked histone I; the absence of this histone was presumable due to its proteolytic degradation, since the chromatin was not treated with protease inhibitors.

Subunit Structure of Chromatin. Rat-liver chromatin and fraction P1 chromatin have been digested with DNase II for extended periods of time (90 min), and the resulting soluble chromatin has been analyzed by centrifugation in isokinetic sucrose gradients (Fig. 2). About 40% of the input DNA sediments extremely slowly; the bulk of this DNA is acid-soluble and hence has been reduced to oligonucleotides by the nuclease. Fig. 2 presents data for whole chromatin depleted of histone I (Fig. 2A) and for P1 chromatin (Fig. 2B). Most of the chromatin sediments at about 11–13 S, with some material sedimenting more rapidly. The properties of nuclease-resistant particles from unfractionated chromatin have been described in detail elsewhere (8–14, 16, 17). We find similar properties for the particles from P1 chromatin: they are composed of equal amounts of protein and DNA (by weight), and the protein complement is almost entirely histone (Table 1). Subunits from native chromatin sediment slightly more rapidly than subunits from histone-I-depleted chromatin (Fig. 2). The sedimentation coefficients (± SD) estimated from ten isokinetic 5–24% sucrose gradients were 12.6 ± 0.9 S for subunits of native chromatin and 11.4 ± 0.7 S for subunits of histone-I-depleted chromatin. The length of DNA contained in the subunits has been studied by many workers; values of 120 to 210 base pairs have been obtained

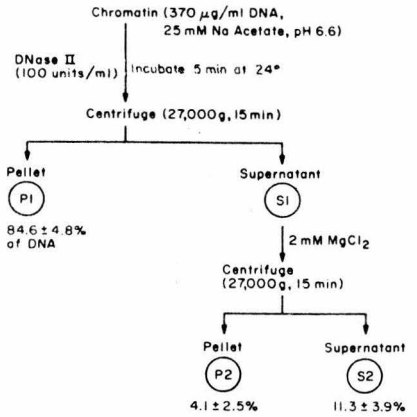


FIG. 1. Fractionation scheme. The yields of DNA in each fraction are the mean and standard deviation for 11 determinations.

Table 1. Chemical composition of rat-liver chromatin fractions

Sample	Composition relative to DNA (w/w)		
	Histone	Non-histone	RNA
Unfractionated chromatin	1.06	0.65	0.05
P1 chromatin*	1.15	0.58	0.05
11–13S subunits†	1.03	<0.05	—
S2 chromatin*	0.61	1.60	0.25
S2 subfractions			
3–5 S	0.24	0.60	—
14.0 ± 0.8S‡	0.72	1.35	0.3–0.4
18.7 ± 1.0S‡	0.54	3.2	0.3–0.7

* Fractionation carried out as in Fig. 1.
† Similar compositions were obtained for native and histone-I-depleted subunits.
‡ Sedimentation values ± SD determined from 24 gradients equivalent to those in Fig. 3.

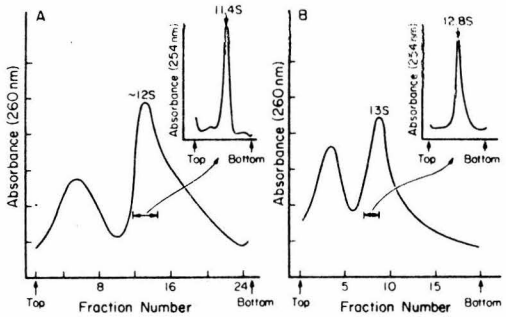


FIG. 2. Sucrose gradient sedimentation of DNase-II-treated chromatin samples. (A) Histone-I-depleted chromatin was digested and soluble chromatin was centrifuged for 42 hr as discussed. (B) Fraction P1 chromatin was digested and soluble chromatin was centrifuged for 36 hr. Fractions were pooled as indicated and rerun on 5–24% gradients (insets).

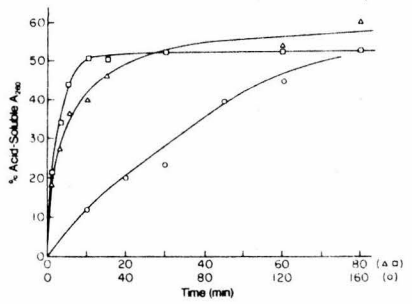


FIG. 4. Kinetics of digestion of chromatin fraction S2. Chromatin was fractionated as described (Fig. 1). S2 chromatin was incubated with nucleases in three ways: reactivation of DNase II (○); addition of fresh DNase II (△); addition of staphylococcal nuclease (□). Aliquots were taken at various times to test for the production of trichloroacetic-acid-soluble material.

(10, 13, 14, 16, 17). The 11–13S subunits resemble *ν*-bodies (2) in the electron microscope (17).

Subunit Structure of Active Chromatin. We now ask whether nuclease-resistant structures occur in transcriptionally active regions of chromatin. Chromatin from rat liver was treated with DNase II for 5 min, fractionated as before (Fig. 1), and S2 material was analyzed on isokinetic sucrose gradients (Fig. 3, curve A). About half of the UV-absorbing material applied to the gradient sediments at 3–5 S; greater than 90% of this material is acid-precipitable after the 5 min nuclease treatment. Two more rapidly sedimenting peaks are seen in the gradient of S2 chromatin, one at 14 S and another at 19–20 S. These gradients were calculated for particles of density 1.44 g/cm^3, and so the observed sedimentation coefficients could be in error if the particle densities are very different from 1.4 to 1.5 g/cm^3. About 6% of the input nucleic acid pelleted during the centrifugation.

To test whether the 14 and 19S nucleoprotein complexes might be multimers of the 11–13S subunit (i.e., dimers and trimers), chromatin fraction S2 was reincubated with DNase. Upon analysis in isokinetic sucrose gradients (Fig. 3, curve B) no significant changes were observed in the >10S region. Reincubation with DNase has been carried out in three ways (reactivation of DNase II, addition of fresh DNase II, addition of staphylococcal nuclease), and similar results were obtained with each of these methods. The kinetics of redigestion of S2 chromatin are presented in Fig. 4. With each method of redigestion, about 50–60% of the input A_{260} became acid-soluble. As the reaction approached completion the solutions first became turbid and eventually a precipitate developed.

S2 chromatin has been isolated after various times of nuclease treatment (30 sec to 30 min) and analyzed in sucrose

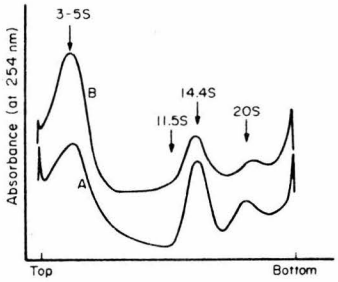


FIG. 3. Sucrose gradient sedimentation of template-active fraction S2 chromatin. Curve A: chromatin was fractionated (Fig. 1) and S2 material was centrifuged for 17.5 hr at 39,000 rpm in 5–24% sucrose gradient. Curve B: fraction S2 was isolated and DNase II reactivated by the addition of 20 mM EDTA (pH 6.4). Incubation was for 1 hr at 24°. The reaction was terminated by raising the pH to 8 with 0.1 M Tris·HCl (pH 11), and the sample was centrifuged as described above.

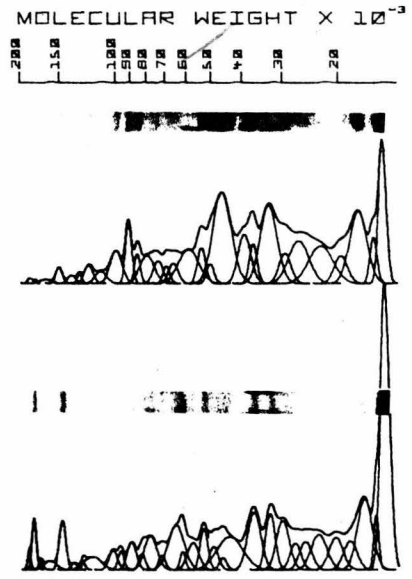


FIG. 5. Sodium dodecyl sulfate–polyacrylamide gels of S2 chromatin subfractions. Total protein of the 3–5S (upper gel) and 14S (lower gel) complexes was electrophoresed as described (27). The stained gels were scanned and the densitometer profiles were resolved into gaussian components by a least-squares computer analysis.

gradients. It was found that the 14 and 19S complexes appear in fraction S2 simultaneously. Thus there is no evidence for a precursor–product relationship between the 19 and 14S complexes. The nuclease-resistant S2 DNA (14 and 19S) was found to have a weight-average single-strand length of 170 nucleotides after the initial DNase treatment (5 min). After prolonged digestion (up to 120 min), the single-strand length was reduced to 120 nucleotides. Nearly all the DNA which sedimented at 3–5 S after the initial DNase digestion was rendered acid-soluble by redigestion.

We have investigated the chemical compositions of the subfractions of S2 chromatin (Table 1). The material at 14–19 S is enriched in both RNA and nonhistone chromosomal proteins. All the major histone species are present in the 14 and 19S complexes; however, these complexes exhibit reduced histone to DNA ratios relative to either unfractionated chromatin or 11–13S subunits (Table 1). The protein population of S2 chromatin has been investigated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (Fig. 5). Material from the 3–5S (upper gel) and from the 14S (lower gel) position in the sucrose gradient (Fig. 3) has been analyzed. Densitometer scans of the stained gels were resolved into gaussian components by a least squares computer analysis. Numerous quantitative and qualitative differences between the 3–5S and 14S proteins can be recognized in these gels. From the length of DNA contained in the 14S chromatin complex (170 nucleotides), the protein-to-DNA ratio (Table 1), and the complexity of the protein population (Fig. 5), we conclude that the 14S peak consists of a heterogeneous population of nucleoprotein species.

DISCUSSION

The template-active fraction of rat-liver chromatin is organized in a fashion similar to that of inactive chromatin; both fractions consist of regions of nuclease-sensitive and nuclease-resistant DNA. Nuclease-resistant segments in transcriptionally inactive chromatin are due to histone–DNA interactions, while the nuclease-resistant segments of active chromatin are due to DNA complexed with both histone and nonhistone proteins. Nuclease-resistant structures of inactive chromatin sediment at 11–13 S and resemble ν-bodies (2) in the electron microscope (17). The nuclease-resistant complexes of active chromatin sediment at 14–19 S and contain RNA as well as protein and DNA.

Our results shed new light on the findings of Felsenfeld's laboratory (8, 9, 31). These investigators have reported that portions of the globin gene are found in both nuclease-sensitive ("open") and nuclease-resistant ("closed") regions of reticulocyte chromatin. Our results suggest that nuclease sensitivity in a *limit digest* does not discriminate between active and inactive chromatin regions. Felsenfeld's data indicate that regions of the globin genes are always covered with protein (30) but make no distinction between histone and nonhistone protein. On the basis of our data we speculate that active genes are complexed with nonhistone as well as histone protein in the 14 and 19S structures.

Inactive genes are complexed mainly with histone and are in the ν-body structures. Although active, like inactive, chromatin contains nuclease-resistant and sensitive regions, there are other major differences. Thermal denaturation and circular dichroism studies (31–33) suggest that active chromatin is in a more extended, more DNA-like conformation than inactive chromatin. The electron microscope has revealed differences in the structure of transcriptionally active and inactive regions of chromatin. Ribosomal genes in the act of transcription are the length of their transcription product (pre-rRNA) (34). On the other hand, DNA complexed with histones in the ν-body configuration is one-seventh the length of the same deproteinized DNA (4). The basic fiber diameter of inactive chromatin is 100 Å. Active chromatin has a fiber diameter of about 30 Å (35, 36). Thus both physical chemical and electron microscopy studies suggest that DNA of active chromatin is more extended than is the DNA of inactive chromatin. It is probable that this is why active chromatin is more susceptible to nuclease attack. This differential sensitivity to nuclease forms the basis of our fractionation procedure (Fig. 1).

We wish to thank Dr. K. Peters and Mr. W. Wheatley for helpful discussions. This work was supported by the U.S. Public Health Service (GM 86 and GM 13762).

1. Pardon, J. F., Wilkins, M. H. F. & Richards, B. M. (1967) *Nature* **215**, 508–509.
2. Olins, A. L. & Olins, D. E. (1974) *Science* **183**, 330–332.
3. Woodcock, C. L. F. (1973) *J. Cell Biol.* **59**, 368a.
4. Griffith, J. D. (1975) *Science* **187**, 1202–1203.
5. Oudet, P., Gross-Bellard, M. & Chambon, P. (1975) *Cell* **4**, 281–300.
6. Hewish, D. R. & Burgoyne, L. A. (1973) *Biochem. Biophys. Res. Commun.* **52**, 504–510.
7. Burgoyne, L. A. Hewish, D. R. & Mobbs, J. (1974) *Biochem. J.* **143**, 67–72.
8. Clark, R. & Felsenfeld, G. (1971) *Nature* **229**, 101–106.
9. Clark, R. & Felsenfeld, G. (1974) *Biochemistry* **13**, 3622–3628.
10. Axel, R., Melchior, W., Jr., Sollner-Webb, B. & Felsenfeld, G. (1974) *Proc. Nat. Acad. Sci. USA* **71**, 4101–4105.
11. Rill, R. & van Holde, K. E. (1973) *J. Biol. Chem.* **248**, 1080–1083.
12. Sahasrabuddhe, C. G. & van Holde, K. E. (1974) *J. Biol. Chem.* **249**, 152–156.
13. Noll, M. (1974) *Nature* **251**, 249–252.
14. Oosterhof, D. K., Hozier, J. C. & Rill, R. L. (1975) *Proc. Nat. Acad. Sci. USA* **72**, 633–637.
15. van Holde, K. E., Sahasrabuddhe, C. G., Shaw, B., van Brugger, E. F. J. & Arnberg, A. C. (1974) *Biochem. Biophys. Res. Commun.* **60**, 1365–1370.
16. Senior, M. B., Olins, A. L. & Olins, D. E. (1975) *Science* **187**, 173–175.
17. Gottesfeld, J. M., Kent, D., Ross, M. & Bonner, J. (1975) in *Florida Colloquium on Molecular Biology*, eds. Stein, G. & Stein, J. (Academic Press, New York), in press.
18. Kornberg, R. (1974) *Science* **184**, 868–871.
19. Marushige, K. & Bonner, J. (1971) *Proc. Nat. Acad. Sci. USA* **68**, 2941–2944.
20. Billing, R. J. & Bonner, J. (1972) *Biochim. Biophys. Acta* **281**, 453–462.
21. Gottesfeld, J. M., Garrard, W. T., Bagi, G., Wilson, R. F. & Bonner, J. (1974) *Proc. Nat. Acad. Sci. USA* **71**, 2193–2197.
22. Marushige, K. & Bonner, J. (1966) *J. Mol. Biol.* **15**, 160–174.
23. Douvas, A., Harrington, C. & Bonner, J. (1975) *Proc. Nat. Acad. Sci. USA* **72**, in press.
24. Noll, H. (1967) *Nature* **215**, 360–363.
25. Studier, F. W. (1965) *J. Mol. Biol.* **11**, 373–390.
26. Bonner, J., Chalkley, R., Dahmus, M., Fambrough, D., Fujimura, F., Huang, R. C. C., Huberman, J., Jensen, R., Marushige, K., Olenbusch, H., Olivera, B. & Widholm, J. (1968) in *Methods in Enzymology*, eds. Grossman, L. & Moldave, K. (Academic Press, New York), Vol. 12, part B, pp. 3–65.
27. King, J. & Laemmli, U. K. (1971) *J. Mol. Biol.* **62**, 465–477.
28. Panyim, S. & Chalkley, R. (1969) *Arch. Biochem. Biophys.* **130**, 337–346.

29. Schmidt, G. & Tannhauser, S. J. (1945) *J. Biol. Chem.* **161,** 83–89.
30. Axel, R., Cedar, H. & Felsenfeld, G. (1973) *Cold Spring Harbor Symp. Quant. Biol.* **38,** 773–783.
31. McConaughy, B. L. & McCarthy, B. J. (1972) *Biochemistry* **11,** 998–1003.
32. Gottesfeld, J. M., Bonner, J., Radda, G. K. & Walker, I.O. (1974) *Biochemistry* **13,** 2937–2945.
33. Polacow, I. & Simpson, R. T. (1973) *Biochem. Biophys. Res. Commun.* **52,** 202–207.
34. Miller, O. L. & Hamkalo, B. A. (1972) in *Molecular Genetics and Developmental Biology*, ed. Sussman, M. (Prentice-Hall, Englewood Cliffs, N.J.), pp. 183–199.
35. Griffith, J. (1970) Ph.D. Dissertation, California Institute of Technology, Pasadena, Calif.
36. Miller, O. L. (1965) *Nat. Cancer Inst. Monogr.* **18,** 79–89.

PART III

Chapter 3. Physical properties of chemically acetylated rat liver chromatin

Reprinted from
Proc. Natl. Acad. Sci. USA
Vol. 74, No. 8, pp. 3244–3248, August 1977
Biochemistry

Physical properties of chemically acetylated rat liver chromatin

(acetic anhydride/nucleosomes/thermal denaturation/histones)

R. BRUCE WALLACE, THOMAS D. SARGENT, ROBERT F. MURPHY, AND JAMES BONNER

Division of Biology, California Institute of Technology, Pasadena, California 91125

Contributed by James Bonner, May 23, 1977

ABSTRACT **The physical properties of rat liver chromatin and nucleosomes acetylated with acetic anhydride were examined in order to clarify the mechanism by which chemical acetylation of histones increases template activity *in vitro* [Marushige, K. (1976) *Proc. Natl. Acad. Sci. USA* 73, 3937–3941]. Acetylation was found to have dramatic effects on the magnesium solubility, nuclease sensitivity, thermal denaturation, and sedimentation of chromatin and nucleosomes. The significance of the results to models of gene activation and chromatin replication is considered.**

The five principal species of histone are associated with DNA to form the eukaryotic chromosome. These basic proteins are involved in the repression of template activity of chromatin (1, 2). It has long been thought that a degree of transcriptional control might be achieved by the selective modification of certain basic residues of the histones (3, 4).

In the current model of chromatin structure, 200 base pairs of DNA are complexed with histones to form the nucleosomal subunit. Although the nucleosome is a general feature of all eukaryotic chromatin thus far examined, there appears to be heterogeneity within the nucleosomes of any single cell type (5). Some of this heterogeneity consists of different conformations of nucleosomes of active and inactive genes (6, 7), which may result from histone modification (4).

Marushige (8) has reported that chemical acetylation of calf thymus chromatin increases its template activity *in vitro* without resulting in significant removal of the histones from the DNA. These results suggest that acetylation of histones participates in activation of genes for transcription. In this report we describe the effect of chemical acetylation on the physical properties of rat liver chromatin and nucleosomes.

MATERIALS AND METHODS

Preparation of Nuclei. Rat liver nuclei were prepared from frozen rat livers (Pelfreeze). All manipulations were performed at 0° unless otherwise stated. Livers were thawed in 0.25 M sucrose/10 mM Tris·HCl, pH 7.4/6 mM KCl/5 mM $Mg(OAc)_2$/0.1 mM ethylene glycol-bis(β-aminoethyl ether)-*N*-*N'*-tetraacetic acid (EGTA). Livers were transferred to three volumes of the same buffer containing 1 mM phenylmethylsulfonyl fluoride (homogenization buffer) and homogenized with five to seven strokes in a motor-driven glass/Teflon homogenizer. The homogenate was diluted to 10 volumes with homogenization buffer, strained through cheesecloth, and centrifuged at 1500 rpm in an HG-4L rotor (Sorvall RC-3 centrifuge) for 10 min. The pellet was washed twice in 10 volumes of homogenization buffer containing 0.5% Nonidet-P40 (Particle Data Laboratories, Ltd), two to four times in homogenization buffer containing 0.1 mM phenylmethylsulfonyl fluoride and 0.5% Nonidet-P40, and finally once in homogenization buffer. The nuclear pellet was resuspended in homogenization buffer and an equal volume of 5 mM $Mg(OAc)_2$/98% glycerol was added. This suspension is stored at −20° for up to 1 month.

Preparation of Chromatin. Nuclei were recovered from the glycerol suspension by centrifugation at 3000 rpm in an HB-4 rotor (Sorvall RC-2 centrifuge) and washed once in 10 mM Tris·HCl, pH 7.4/6 mM KCl/5 mM $Mg(OAc)_2$ (TKM). The nuclei were lysed by vigorous homogenization in 2.5 mM Tris·HCl, pH 7.4/2.5 mM ethylenediaminetetraacetic acid (EDTA), followed by centrifugation at 10,000 rpm in the HB-4 rotor for 10 min. The chromatin was then washed once in the same buffer and twice in 5 mM sodium borate, pH 8.2/10 mM NaCl and finally was suspended in 5 mM sodium borate, pH 8.2/10 mM NaCl at an A_{260} of 10 (all chromatin absorbances are measured in 1 M NaOH).

Novikoff hepatoma cells [line NISI-67 adapted for growth in Swimms 210 medium (Gibco)] were cultured as described by Plagemann (9). Cells were labeled with [2-^{14}C]thymidine (56 mCi/mmol) at 0.05 μCi/ml for two cell generations. Preparation of nuclei will be described elsewhere (R. B. Wallace, S. K. Dube, and J. Bonner, unpublished data); they were mixed with 4-fold excess of rat liver nuclei. Chromatin was prepared as described above.

Preparation of Nucleosomes. Nuclei were recovered from the glycerol suspension as described above, washed once in TKM, and suspended in TKM at an A_{260} of 150–200. The suspension was warmed to 37°, brought to 0.25 mM $CaCl_2$, and digested for 5 min with 0.24 μg (7.5 units) of staphylococcal nuclease (P-L Biochemicals) per A_{260} unit. The nuclei were cooled to 0° and centrifuged at 3000 rpm in the HB-4 rotor for 5 min. The nuclear pellet was homogenized in 2.5 mM Tris·HCl, pH 7.4/2.5 mM EDTA and centrifuged at 10,000 rpm in the HB-4 rotor for 10 min. The supernatant was then passed over a Sephadex G-50 column equilibrated with 5 mM sodium borate, pH 8.2/0.1 mM EGTA; the excluded fraction, representing nucleosomes, was adjusted to an A_{260} of 10. Typical preparations were 5–15% acid soluble and 50–70% monomer nucleosomes.

Acetylation of Chromatin and Nucleosomes. Chromatin and nucleosomes were acetylated with acetic anhydride as described by Wong and Marushige (10). Chromatin solutions were acetylated with 0.14 mM, 0.7 mM, and 7 mM acetic anhydride at an A_{260} of 10. Acetylated chromatin samples were dialyzed overnight against 10 mM Tris·HCl, pH 7.4. The incorporation of acetate groups was monitored by acetylating chromatin with [^{3}H]acetic anhydride (Schwartz/Mann) diluted to 0.4 mCi/mmol. The amount of acetylation achieved with the three concentrations was essentially the same as that reported by Wong and Marushige.

The costs of publication of this article were defrayed in part by the payment of page charges from funds made available to support the research which is the subject of the article. This article must therefore by hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: EGTA, ethylene glycol-bis(β-aminoethyl ether)-*N*,*N'*-tetraacetic acid; TKM buffer, 10 mM Tris·HCl, pH 7.4/6 mM KCl/5 mM $Mg(OAc)_2$; EDTA, ethylenediaminetetraacetic acid.

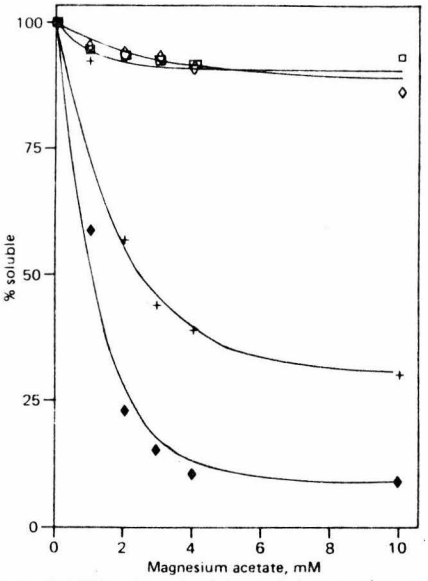


FIG. 1. Solubility of acetylated chromatin in magnesium acetate. Nucleosomes were acetylated with acetic anhydride at 0 (♦), 0.14 (+), 0.7 (◊), and 7 mM (□) (see *Materials and Methods*), dialyzed against 10 mM Tris·HCl, pH 7.4/0.1 mM EGTA, and diluted to an A_{260} of 1.0. Aliquots were adjusted to the desired magnesium acetate concentration and centrifuged at 0° for 10 min at 10,000 rpm in an SS-34 rotor (Sorvall RC-2B). The A_{260} of the supernatant was measured, and the values are expressed as percentage of total A_{260}. The lines represent nonlinear least-squares exponential fits to the data (13).

Thermal Denaturation. Acetylated and control chromatins were dialyzed exhaustively against 0.25 mM EDTA (pH 8) and adjusted to an A_{260} of approximately 1.5. Melting was performed with a Gilford 2400 spectrophotometer equipped with a thermal cuvette and digital absorbance meter under control of a Tektronix-31 programmable calculator. Absorbances were recorded every 0.4°, each point being the average of 10 readings. The data were corrected for the small hyperchromicity of the buffer and derivatized by using the least-squares method of Savitzky and Golay (11) (15-point cubic quartic polynomial first-derivative smooth). The derivative data were normalized to percentage hyperchromicity and resolved into gaussian components by using a nonlinear least-squares fitting program (12, 13).

Sucrose Density Gradient Centrifugation of Nucleosomes. Isokinetic sucrose gradients were prepared by the method of Noll (14), assuming a particle density of 1.44 g/cm^3 and using the following parameters: gradient volume, 11.6 ml; C_{top}, 5% (wt/wt) sucrose; C_{res}, 26.7% (wt/wt) sucrose; sample volume, 0.5 ml. The gradient buffer was 10 mM Tris·HCl, pH 8/0.1 mM EGTA. The gradients were run at 36,000 rpm in an SW41 rotor for 18 hr at 4°. The gradients were scanned by using an ISCO UA-5 ultraviolet monitor equipped with a digital absorbance meter, and data were collected by a Tektronix-31 programmable calculator. Fractions of the gradients containing nucleosomes acetylated with [^{3}H]acetic anhydride were collected and assayed for radioactivity directly in Aquasol 2 scintillation fluid (New England Nuclear).

Polyacrylamide Gel Electrophoresis. The products of DNase I digestion of control and acetylated chromatins were analyzed on 12% polyacrylamide/7 M urea gels as described by Maniatis *et al.* (15). Slab gels, 2 × 120 × 150 mm, were prepared and aged overnight. Samples were brought to 50% formamide, boiled for 3 min, and quenched on ice. Electrophoresis was at 200 V and continued until the bromophenol blue dye was 0.5 cm from the end of the gel. The gel was stained for 2 hr with ethidium bromide (1 μg/ml in 0.5 M ammonium acetate) and photographed under ultraviolet light with a yellow filter.

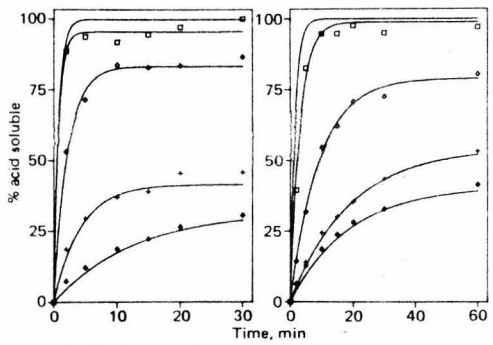


FIG. 2. Nuclease sensitivity of acetylated chromatin. Chromatin was prepared from [^{14}C]thymidine-labeled cells and acetylated as indicated in *Materials and Methods*. The samples were dialyzed against 10 mM Tris·HCl, pH 7.4, the A_{260} was adjusted to 0.4, and the sample was adjusted to 10 mM NaCl, 3 mM $Mg(OAc)_2$, and 1 mM $CaCl_2$. DNase I at 0.5 μg/ml (*Left*) or staphylococcal nuclease at 0.14 μg/ml (*Right*) was added, and the digestions were carried out at 24°. Aliquots (0.5 ml) were removed at intervals and precipitated with 1 ml of cold 0.3 M perchloric acid. Insoluble material was removed by centrifugation. The supernatant was neutralized with NaOH, adjusted to 1% sodium dodecyl sulfate, 50 mM Tris·HCl (pH 7.4), and assayed for radioactivity in Aquasol 2. After 1 hr of digestion, the remaining chromatin was adjusted to 1% sodium dodecyl sulfate, 50 mM Tris·HCl (pH 7.4), 0.2 M sodium perchlorate and assayed for total radioactivity. Acetic anhydride concentrations: 0 (♦), 0.14 mM (+), 0.7 mM (◊), and 7 mM (□). The lines represent nonlinear least-squares exponential fits to the data (13), with the uppermost line in each panel representing the digestion of deproteinized rat liver DNA (see Table 1).

RESULTS

One of the most obvious physical changes that accompanies chemical acetylation of rat liver chromatin is its solubilization. Under the conditions of the acetylation reaction (i.e., 0.15 M NaCl), chromatin is a condensed, insoluble precipitate. Chromatin is similarly condensed in the presence of 2 mM Mg^{2+}. Upon the addition of acetic anhydride at 0.7 or 7 mM, the chromatin precipitate becomes noticeably more soluble. We investigated this phenomenon by comparing the solubility of control and acetylated nucleosomes in various concentrations of magnesium (Fig. 1). The control nucleosomes were very insoluble, less than 10% remaining soluble in 10 mM magnesium. Acetylation increased the solubility of nucleosomes, the 0.7 mM and 7 mM acetylated material becoming almost completely soluble.

Template active regions of chromatin are more sensitive to DNase I (6, 7) than are inactive regions. If acetylation is responsible for gene activation *in vivo*, *in vitro* acetylated chromatin might be expected to show greater nuclease sensitivity than unacetylated chromatin. Fig. 2 *left* and Table 1 show the kinetics of digestion of acetylated ^{14}C-labeled chromatin with DNase I. Increased acetylation of chromatin dramatically al-

Table 1. Kinetic parameters from Fig. 2

	DNase I			Staphylococcal nuclease		
	k*	M†	R‡	k	M	R
Control	0.079	32.0	1.0	0.054	40.5	1.0
Acetylated chromatin:						
0.14 mM	0.237	41.6	3.9	0.054	53.9	1.3
0.7 mM	0.458	83.3	15.2	0.110	78.9	3.9
7 mM	1.300	95.6	49.4	0.322	98.8	14.4
DNA	1.290	100.0	51.3	0.653	100.0	29.6

* First-order rate constant (min^{-1}).
† Maximal percentage digested.
‡ Relative initial rate of reaction. The initial rate for control chromatin was 0.01 A_{260} unit/ml per min with DNase I and 0.0088 A_{260} unit/ml per min with staphylococcal nuclease.

tered the sensitivity of the DNA to the nuclease. Chromatin acetylated with 0.14 mM acetic anhydride was digested 4 times faster than control chromatin, and that acetylated at 0.7 mM was digested 15 times faster. Maximally acetylated chromatin (7 mM) was as sensitive as deproteinized DNA to DNase I. Fig. 2 *right* shows the results of a similar experiment with staphylococcal nuclease. It can be seen that 0.14 mM acetylated chromatin was slightly more sensitive than control chromatin to staphylococcal nuclease, and 0.7 mM acetylated chromatin was digested 4 times as fast as control chromatin; 7 mM acetylated chromatin was digested approximately half as fast as deproteinized DNA. Thus, chemically acetylated chromatin is moderately sensitive to staphylococcal nuclease but extremely sensitive to DNase I.

Weintraub and Groudine (6) reported that, during the digestion of active genes with DNase I, the digested DNA appears as multiples of 10 nucleotides. Fig. 3 shows that 0.14 mM and 0.7 mM acetylated chromatins were digested by DNase I to the same 10-nucleotide repeat pattern as control chromatin. This result suggests that the histone–DNA interaction that produces this periodic pattern is not disrupted by the acetylation. On the other hand, the 7 mM acetylated chromatin did not have this repeat pattern (although faint bands are apparent at 10-nucleotide intervals).

The DNA of the transcriptionally active regions of chromatin is believed to have a lower melting temperature than the DNA of the inactive regions. This belief is supported by the thermal fractionation experiments of McConaughy and McCarthy (16). In addition, DNase II-fractionated template-active chromatin melts at a lower temperature than does unfractionated chromatin (ref. 17; unpublished data).

FIG. 3. Polyacrylamide gel electrophoresis of DNase I-digested chromatin DNA. Control and acetylated chromatins were prepared as described in *Materials and Methods*. The samples were dialyzed against 10 mM Tris·HCl (pH 7.4) and adjusted to A_{260} of 10. Each was brought to 10 mM NaCl, 3 mM $Mg(OAc)_2$, and 1 mM $CaCl_2$, and DNase I, 10 μg/ml, was added. Each sample was digested for various times at 24° such that the percentage of acid-soluble A_{260} was 25, 30, 15, and 15 for control, 0.14 mM, 0.7 mM, and 7 mM acetic anhydride, respectively. Reactions were stopped by bringing each to 1% sodium dodecyl sulfate; then the samples were phenol/chloroform extracted as described by Marmur (29). DNA was precipitated with ethanol and dissolved in gel buffer (15); 25 μg of each was brought to 50% formamide, boiled for 3 min, and cooled on ice. Electrophoresis was performed as described in *Materials and Methods*. From left to right: DNA from unacetylated chromatin and 0.14 mM, 0.7 mM, and 7 mM acetylated chromatin.

Table 2. Summary of melting transitions from Fig. 4

	Tm1	%	Tm2	%	Tm3	%	Tm4	%
				Nucleosomes				
Control	64.5	34.4	76.3	60.6	81.1	5.0	—	—
Acetylated:								
0.14 mM	60.2	39.3	73.8	52.0	79.7	8.8	—	—
0.7 mM	48.3	48.0	65.7	39.0	75.7	13.0	—	—
7 mM	42.1	55.2	52.2	31.6	69.7	13.2	—	—
DNA	40.8	100.0	—	—	—	—	—	—
				Chromatin				
Control	62.0	29.1	71.1	18.7	77.1	23.4	83.4	28.8
Acetylated:								
0.14 mM	53.3	20.8	65.0	26.4	73.6	20.5	81.3	32.3
0.7 mM	47.6	44.9	63.3	28.7	—	—	79.8	26.4
7 mM	43.7	83.7	50.7	6.4	—	—	75.6	9.9
DNA	40.8	100.0	—	—	—	—	—	—

Temperatures are in degrees Celsius. Transitions are numbered in order of increasing temperature. Percentages refer to the fraction of the total hyperchromicity of each sample in a given transition.

The thermal denaturation of control and acetylated chromatins and nucleosomes, as well as deproteinized DNA, is shown in Fig. 4. The data are presented as first-derivative melting profiles, fitted to three or four gaussian components to quantitate the transitions observed. Table 2 summarizes the melting transitions and the contribution of each to the total hyperchromicity. The most obvious effect of chemical acetylation on thermal denaturation is the dramatic shift to lower melting temperatures of all the observed transitions. This is true for both chromatin and nucleosomes, although the melting profiles of control and acetylated chromatin were more complex.

Fig. 5 shows the effect of acetylation on the sedimentation behavior of nucleosomes. The sedimentation of acetylated

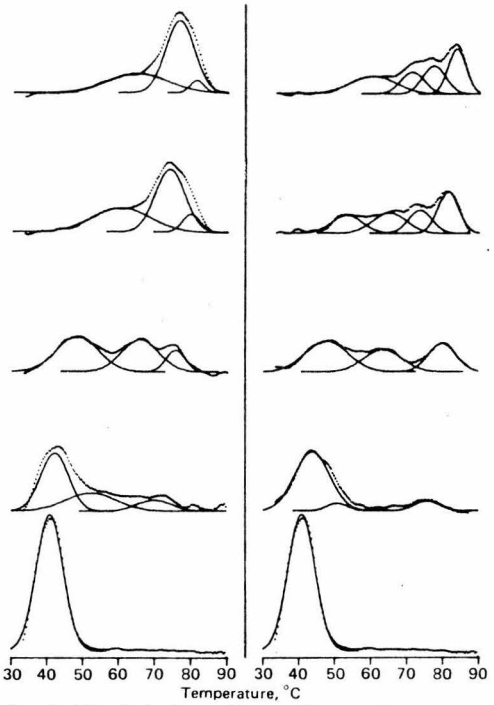


FIG. 4. First-derivative melting profiles (see *Materials and Methods*). Full scale for each profile is 3% change in hyperchromicity per degree. From top to bottom: unacetylated material, 0.14 mM, 0.7 mM, and 7 mM acetylated material, and deproteinized rat liver DNA. (*Left*) Nucleosomes. (*Right*) Chromatin. ···, dH/dT. —, Gaussian components.

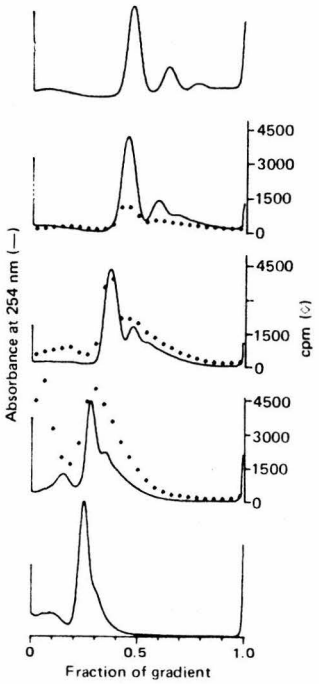


FIG. 5. Sucrose gradient sedimentation of nucleosomes. Sedimentation is from left to right. From top to bottom: nucleosomes, 0.14 mM acetylated nucleosomes, 0.7 mM acetylated nucleosomes, 7 mM acetylated nucleosomes, and sodium dodecyl sulfate-treated nucleosomal DNA (30). The sedimentation coefficients of the monomer peak are 11.3, 10.6, 8.6, 6.2, and 5.4 S, respectively. The specific activity of the labeled samples was 19.1, 67.4, and 119 acetates per 200 base pairs of DNA for 0.14 mM, 0.7 mM, and 7 mM acetylated nucleosomes, respectively. Under the conditions of acetylation, approximately half as much acetate was incorporated into nonhistones as into histones, and no detectable acetate was incorporated into DNA.

nucleosomes was dramatically retarded compared with that of control nucleosomes and approached that of sodium dodecyl sulfate-treated control nucleosomes. To determine whether all of these nucleosomes were acetylated equally and whether all of the acetate remained bound to the nucleohistone, the acetylation reaction was carried out with [^{3}H]acetic anhydride. Fig. 5 also shows the distribution of [^{3}H]acetate across the gradients. From 0.14 to 0.7 mM acetic anhydride there was a 4-fold increase in acetate incorporation per nucleosome whereas from 0.7 to 7 mM acetic anhydride there was only a 2-fold increase. This demonstrates the limit acetylation reported by Wong and Marushige (10). All nucleosomes, monomer through trimer, were acetylated. At 7 mM acetic anhydride, [^{3}H]acetate counts were found at the top of the gradient, perhaps representing dissociated histone H1.

DISCUSSION

We have demonstrated that chemical acetylation of nucleohistone *in vitro* results in dramatic changes in its physical properties. The results obtained suggest possible functions for acetylation *in vivo*.

The chemical acetylation of nucleosomes was found to increase their solubility in the presence of Mg^{2+}. Previous studies from this laboratory have shown that DNase II digestion followed by precipitation with 0.15 M NaCl (18) or 2 mM $MgCl_2$ (19) yields a fraction of rat liver chromatin enriched in sequences transcribed *in vivo* (19).

The sensitivity of chromatin to DNase I and, to a lesser extent, to staphylococcal nuclease is significantly increased after treatment with acetic anhydride. Similar nuclease sensitivity properties have been reported for transcriptionally active sequences (6, 7, 20), although the biochemical basis for this phenomenon has not been determined. Marushige (8) reported that the derepression of template activity by chemical acetylation with 0.7 mM acetic anhydride was a result of the acetylation of histones H2A and H2B, with H3 and H4 not contributing significantly. H3 and H4 have been implicated in the protection of multimers of 10 nucleotides of DNA from digestion by DNase I (21). Fig. 3 demonstrates that the 10-nucleotide repeat pattern is observed in 0.7 mM acetylated chromatin even though the rate of digestion by DNase I is dramatically accelerated (Fig. 2 *left*). This suggests that the interaction of H3 and H4 with the DNA has not been greatly perturbed, whereas the template activity measurements by Marushige suggest that the interaction of H2A and H2B with DNA has been functionally altered by acetylation. Higher levels of acetylation result in a disruption of the 10-nucleotide repeat pattern (Fig. 3).

Thermal denaturation of nucleohistone demonstrates the

dramatic stability of the histone–DNA complex. In order to make DNA sequences available for transcription (or replication), it might be necessary to decrease the stability of this complex (22). Although no direct correlation of nucleohistone thermal denaturation behavior with functional state has been demonstrated, isolated template-active chromatin has a lower melting temperature than does the bulk of chromatin (17), and active sequences of chromatin elute from hydroxylapatite at lower temperatures (16). The results presented in Fig. 4 and Table 2 demonstrate that chemical acetylation lowers the melting temperature of nucleohistone, as might be expected if acetylation were correlated with transcriptional activation. The transitions observed in the first-derivative melting profiles are similar to those reported by others (23–25). We have resolved the transitions into gaussian components in order to quantitate the contribution of each transition to the total hyperchromicity. From previous work (2, 23), as well as unpublished observations in our laboratory, tentative assignments may be made to the transitions of control nucleosomes. The two transitions at the highest temperatures are due to the denaturation of the histone octamer–DNA complex. The lowest transition is due to a more heterogeneous set of interactions as reflected by the breadth of this transition. One of these contributing interactions is probably H1–DNA complexes (26). Acetylation of nucleosomes with 0.14 mM acetic anhydride had little effect on the profile. The relative contribution of each transition remained unchanged, and the transitions were shifted to lower temperatures, T1 being shifted by the largest amount. This is consistent with the role of H1 in the T1 transition, in view of the extensive acetylation of this histone by acetic anhydride (10). The acetylation of nucleosomes with 0.7 mM acetic anhydride resulted in dramatic changes in the melting profile: all three transitions were shifted to lower temperatures and the relative contributions of T1 and T2 were increased at the expense of T3. This may reflect the differential effect of acetylation of H2A/H2B and H3/H4 as discussed above. Acetylation with 7 mM acetic anhydride virtually abolished the stabilization of DNA melting by histones, even though histones remain associated with the DNA (refs. 8 and 10; Fig. 5). The thermal denaturation profiles of chromatin show the same general features as those of nucleosomes.

At the lowest level of acetylation used (0.14 mM acetic anhydride), there were dramatic changes in magnesium solubility (Fig. 1) and DNase I sensitivity (Fig. 2 *left*). However, staphylococcal nuclease sensitivity (Fig. 2 *right*), DNase I digestion pattern (Fig. 3), and thermal denaturation (Fig. 4 and Table 2) were only slightly perturbed by the acetylation. Although the sites of chemical acetylation have not been determined, it is possible that this low level of acetylation destabilizes interactions that are qualitatively different from those affected by higher levels.

The demonstration of a direct role of acetylation of histones in gene activation, as suggested by Allfrey *et al.* (3), has been elusive. Jackson *et al.* (27) argued that histone acetylation is too extensive to be a specific mechanism of gene regulation. However, histone modification may be important in DNA replication and histone deposition, as well as in gene activation (4, 28). The many functions of acetylation might be accomplished through different classes of sites for acetylation on the histones. For gene activation, it may be sufficient to alter histone–histone or histone–DNA interactions slightly as in the 0.14 mM acetylation. This could result in the conversion of nucleosomes to a conformation more conducive to transcription of the associated DNA. On the other hand, DNA replication or histone deposition might require a more extensive destabilization of these histone–DNA interactions, as in the higher levels of acetylation reported here.

We thank Dr. S. K. Dube for many stimulating discussions and critical evaluation of this work and Ms. M. J. Johnson for valuable technical assistance. This work was supported by U.S. Public Health Service Grant GM-13762. R.B.W. is the holder of a Medical Research Council of Canada fellowship. T.D.S. is supported by a predoctoral fellowship from the National Science Foundation.

1. Huang, R. C. C., Bonner, J. & Murray, K. (1964) *J. Mol. Biol.* **8**, 54–64.
2. Shih, T. Y. & Bonner, J. (1970) *J. Mol. Biol.* **48**, 469–487.
3. Allfrey, V. G., Faulkner, R. & Mirsky, A. E. (1964) *Proc. Natl. Acad. Sci. USA* **51**, 786–794.
4. Ruiz-Carrillo, A., Wangh, L. J. & Allfrey, V. G. (1975) *Science* **190**, 117–128.
5. Weintraub, H. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1212–1216.
6. Weintraub, H. & Groudine, M. (1976) *Science* **193**, 848–856.
7. Garel, A. & Axel, R. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3966–3970.
8. Marushige, K. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 3937–3941.
9. Plagemann, P. G. W. (1969) *Biochim. Biophys. Acta* **182**, 46–56.
10. Wong, T. K. & Marushige, K. (1976) *Biochemistry* **15**, 2041–2046.
11. Savitzky, A. & Golay, M. J. E. (1964) *Anal. Chem.* **36**, 1627–1639.
12. Marquardt, D. W. (1963) *J. Soc. Ind. Appl. Math.* **11**, 431–441.
13. Pearson, W., Davidson, E. H. & Britten, R. J. (1977) *Nucleic Acids Res.*, in press.
14. Noll, H. (1967) *Nature* **215**, 360–363.
15. Maniatis, T., Jeffrey, A. & van deSande, H. (1975) *Biochemistry* **14**, 3787–3794.
16. McConaughy, B. L. & McCarthy, B. J. (1972) *Biochemistry* **11**, 998–1003.
17. Duerksen, J. D. & McCarthy, B. J. (1971) *Biochemistry* **10**, 1471–1478.
18. Marushige, K. & Bonner, J. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 2941–2944.
19. Gottesfeld, J. M., Garrard, W. T., Bagi, G., Wilson, R. F. & Bonner, J. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 2193–2197.
20. Reeves, R. (1976) *Science* **194**, 529–532.
21. Sollner-Webb, B., Camerini-Otero, R. F. & Felsenfeld, G. (1976) *Cell* **9**, 179–193.
22. Alberts, B. & Frey, L. (1970) *Nature* **227**, 1313–1318.
23. Li, H-J. & Bonner, J. (1971) *Biochemistry* **10**, 1461–1470.
24. Woodcock, C. L. F. & Frado, L-L. Y. (1975) *Biochem. Biophys. Res. Commun.* **66**, 403–410.
25. Li, H-J., Change, C. & Weiskopf, M. (1973) *Biochemistry* **12**, 1763–1771.
26. Olins, D. E. (1969) *J. Mol. Biol.* **43**, 439–460.
27. Jackson, V., Shires, A., Chalkley, R. & Granner, D. K. (1975) *J. Biol. Chem.* **250**, 4856–4863.
28. Louie, A. J., Candido, E. P. M. & Dixon, G. H. (1973) *Cold Spring Harbor Symp. Quant. Biol.* **38**, 803–819.
29. Marmur, J. (1961) *J. Mol. Biol.* **3**, 208–218.
30. Smith, C. A., Jordan, J. M. & Vinograd, J. (1971) *J. Mol. Biol.* **59**, 255–272.

PART IV

Chapter 4. Altered nucleosome spacing in newly replicated chromatin from Friend leukemia cells

Chapter 5. Properties of newly replicated chromatin: Analysis using shallow metrizamide gradients

Reprinted from
Proc. Natl. Acad. Sci. USA
Vol. 75, No. 12, pp. 5903–5907, December 1978
Biochemistry

Altered nucleosome spacing in newly replicated chromatin from Friend leukemia cells

(nucleosomes/chromatin replication/staphylococcal nuclease)

ROBERT F. MURPHY, R. BRUCE WALLACE*, AND JAMES BONNER

Division of Biology, California Institute of Technology, Pasadena, California 91125

Contributed by James Bonner, September 18, 1978

ABSTRACT **Chromatin from Friend leukemia cells labeled with [^{14}C]thymidine for 24 hr followed by [^{3}H]thymidine for 10 min is converted into nucleosomes by staphylococcal nuclease at only half the rate that total chromatin is converted. Polyacrylamide gel electrophoresis of nucleosomal DNA from cells labeled for 24 hr with [^{14}C]thymidine followed by 10 min with [^{3}H]thymidine demonstrates that the internucleosomal spacer of newly replicated chromatin is approximately 20 base pairs shorter than that of total chromatin. The implications of this difference for models of chromatin structure are discussed.**

Considerable evidence from a number of laboratories indicates that the bulk of eukaryotic DNA exists in a repeated, globular form [reviewed by Kornberg (1)]. The repeat unit (the nucleosome), contains 140 base pairs of DNA associated with an octamer of the "inner" histones (H2A, H2B, H3, and H4), and between 40 and 80 base pairs in an internucleosomal spacer. Recent experiments suggest that 30–50 base pairs of this spacer are complexed with histone H1 (2–4). Variations in the spacer length have been observed between species and cell types (5–11), and the breadth of the bands observed in DNA gels has led to the postulation of heterogeneity within a single cell type (12). Investigations have so far revealed no differences in nucleosome spacing between transcribed and untranscribed chromatin (13–18). In this report we present evidence of a difference in repeat length between newly replicated and total chromatin from Friend leukemia cells.

MATERIALS AND METHODS

Cell Culture and Labeling. Friend leukemia cells, clone FSD-3, were grown in suspension culture as described previously (19). Cells were labeled with [^{14}C]thymidine (61 Ci/mol, Moravek Biochemicals) at 0.1 μCi/ml for 24 hr (1.5 generations) followed by [^{3}H]thymidine (21 Ci/mmol, Amersham) at 10 μCi/ml for 10 min or at 4 μCi/ml for 30 min (1 Ci = 3.7×10^{10} becquerels).

Preparation of Nuclei and Nucleosomes. Cells were harvested at the end of the ^{3}H labeling period, washed once in phosphate-buffered saline, and suspended in 10 mM Tris·HCl, pH 8.0/10 mM NaCl/5 mM $Mg(OAc)_2$/0.5% Nonidet-P40 (Shell Chemical). After 10 min, a nuclear pellet was formed by centrifugation at $2000 \times g$ for 2 min. The nuclei were washed once in Tris/NaCl/Mg/Nonidet-P-40 and once in Tris/NaCl/Mg and were resuspended in Tris/NaCl/Mg/0.25 mM $CaCl_2$. The suspension was adjusted to approximately 50 μg of DNA per ml and brought to 37°C. Staphylococcal nuclease (P-L Biochemicals) was added to aliquots of the nuclear suspension at concentrations of 0.01–65 units/ml (1 unit produces 15 A_{260} units of acid-soluble material from DNA in 30 min at pH 8.8, 37°C), and digestion was stopped after 5 min by addition of 4 vol of ice-cold Tris/NaCl/Mg. The reaction mix was centrifuged at $2000 \times g$ for 2 min, and the soluble material was removed (this fraction was found to be equivalent to the fraction soluble in 5% perchloric acid). The pellet was suspended in 2.5 mM Tris·HCl, pH 8.0/2.5 mM ethylene glycol bis(β-aminoethyl ether)-*N,N,N',N'*-tetraacetic acid (EGTA). The Tris/EGTA-insoluble (undigested) fraction was removed by centrifugation at $2000 \times g$ for 3 min. The amount of DNA in the various fractions was determined by bringing aliquots to 1% sodium dodecyl sulfate and scintillation counting in Aquasol-2 (New England Nuclear). A ^{14}C to ^{3}H spilldown correction of 17% was performed (the labeling conditions produced roughly equal amounts of ^{3}H and ^{14}C cpm). Under these counting conditions, the sum of the cpm in the three fractions for each isotope was constant from low to high enzyme concentration, and equal to the total input cpm. Thus the differences observed were not due to preferential quenching of large [^{3}H]DNA. In addition, no differences between 24-hr ^{14}C-labeled and 24-hr ^{3}H-labeled DNA were seen (data not shown).

Preparation of DNA and Gel Electrophoresis. DNA was extracted from digested nuclei essentially by the procedure of Marmur (20) and dissolved in 89 mM Tris/89 mM boric acid/2.5 mM EDTA containing 5% (vol/vol) glycerol and 0.01% bromophenol blue. Approximately 10 μg of the DNA samples was applied to 15-cm 4% polyacrylamide gels [in Tris/borate/EDTA, (21)] and electrophoresed at 150 V for 2 hr. Gels were fractionated into 2-mm slices (Aliquogel fractionator, Gilson) and scintillation counted in Aquasol-2. Data were corrected for spilldown of ^{14}C and converted to mobility relative to the dye. The replicative form 2 DNA of phage ϕX174 was cleaved with restriction endonuclease *Hae* III and labeled by a slight modification of the procedure of Berkner and Folk (22).

RESULTS

We have examined the digestion by staphylococcal nuclease of DNA in isolated nuclei from cells that had been labeled for 24 hr with [^{14}C]thymidine followed by 10 or 30 min with [^{3}H]thymidine. The amount of radioactivity in acid-soluble, Tris/EGTA-soluble and Tris/EGTA-insoluble form was determined at various ratios of enzyme to substrate. [From analysis of sucrose gradient profiles, the Tris/EGTA-soluble fraction represents nucleosome multimers from 1 to about 20 (data not shown).] Fig. 1 and Table 1 show that 30-min-labeled chromatin is digested into nucleosomes at about $\frac{2}{3}$ the rate, and 10-min-labeled chromatin at less than $\frac{1}{2}$ the rate that total chromatin is digested. However, the fact that there is no difference in the production of acid-soluble material suggests a difference in the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviation: EGTA, ethylene glycol bis(β-aminoethyl ether)-*N,N,N',N'*-tetraacetic acid.

* Present address: Department of Biology, City of Hope National Medical Center, Duarte, CA 91010.

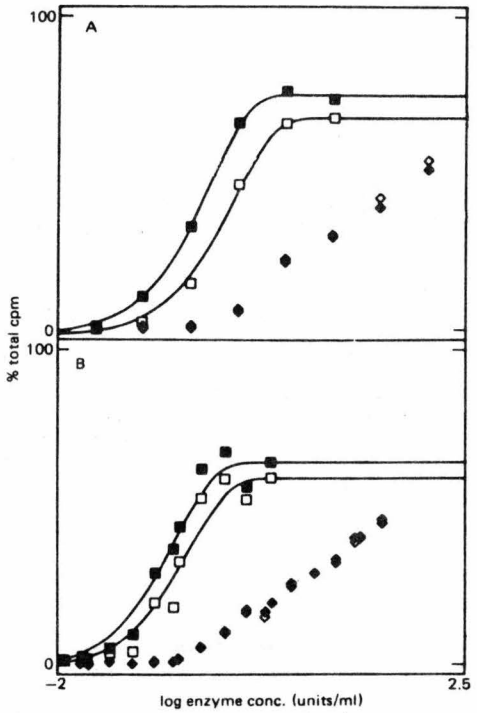


FIG. 1. Digestion of newly labeled and total chromatin by staphylococcal nuclease. The percent of total cpm in each fraction (P) is plotted versus the log of enzyme concentration (E, units/ml). (This semilogarithmic plot expands the low enzyme region.) (A) Ten-minute labeling, ^3H total cpm = 35,421 (±6.3%), ^{14}C total cpm = 13,710 (±8.3%) (uncertainties are % SD). (B) Thirty-minute labeling, data from three experiments. ^3H total cpm = 164,524, 110,084, 167,690; ^{14}C total cpm = 46,223, 100,724, 100,750 (all ±3–4%). ♦, ^{14}C (total) Tris/NaCl/Mg-soluble (acid-soluble); ◊, ^3H (new) Tris/NaCl/Mg-soluble (acid-soluble); ■, ^{14}C (total) Tris/EGTA-soluble (nucleosomes); □, ^3H (new) Tris/EGTA-soluble (nucleosomes). The lines represent fits to the equation $P = Me^{kE}$ drawn by using a nonlinear least-squares fitting program (R. F. Murphy, W. R. Pearson, and J. Bonner, unpublished and ref. 23) (see Table 1).

initial availability of internucleosomal DNA. The accuracy of these estimates depends in large part on the adherence of nuclease digestion to pseudo-first-order kinetics under our conditions. This problem may be avoided by plotting the same data as $^{14}C/^3H$ ratio (Fig. 2). It can be seen that the difference is

Table 1. Kinetic parameters from Fig. 1

Chromatin	k*	M†	R‡
10-min new	0.224	70.2	15.7
10-min total	0.416	76.9	32.0
Ratio new/total	0.539	0.912	0.491
30-min new	0.665	61.4	40.9
30-min total	0.956	66.8	63.9
Ratio new/total	0.696	0.920	0.640

* First-order rate constant (ml unit^{-1} min^{-1}).
† Maximum percent digested.
‡ Initial rate of reaction ($k \times M$).

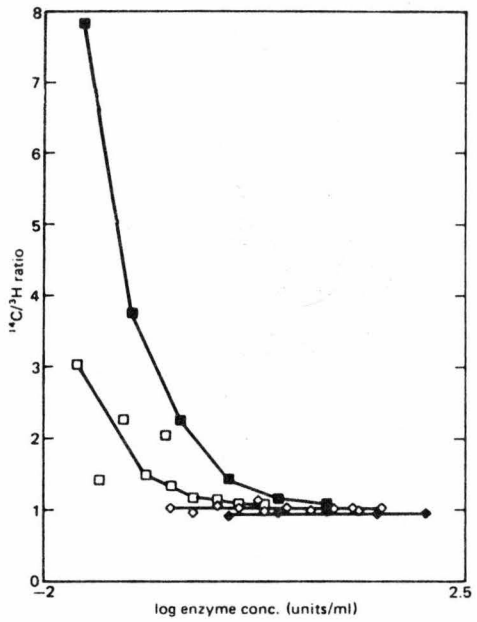


FIG. 2. Comparison of digestion of total and newly labeled chromatin. Data from Fig. 1 replotted as $^{14}C/^3H$ ratio for each sample. ♦, 10-min Tris/NaCl/Mg-soluble; ◊, 30-min Tris/NaCl/Mg-soluble; ■, 10-min Tris/EGTA-soluble; □, 30-min Tris/EGTA-soluble.

greatest at low enzyme-to-chromatin ratios, approaching 8-fold for a 10-min labeling.

The difference in digestion rates observed suggests a difference in the structure of newly replicated chromatin. In order to examine that possibility, total DNA from various digests of nuclei labeled for 24 hr with [^{14}C]thymidine followed by 10 min with [^3H]thymidine was prepared and electrophoresed in 4% polyacrylamide gels (Fig. 3). The monomer peaks in each gel are coincident, but a difference in the higher multimers is evident. The ^3H peaks appear broader, so that for dimers and trimers the ^{14}C peak is almost included in the ^3H peak. The panels on the right show the data plotted as $^3H/^{14}C$ ratio. Peaks are seen to the right of the position of the ^{14}C-labeled multimers (marked by the arrows), demonstrating a difference in the average size of the two types of multimer DNA.

To quantitate this difference, Gaussian curves were fit to the data from each gel for each isotope by using a nonlinear least-squares fitting program (R. F. Murphy, W. R. Pearson, and J. Bonner, unpublished; ref. 24). ^{32}P-Labeled fragments of ϕX174 replicative form 2 DNA cleaved with *Hae* III were either run on parallel gels or mixed with the ^3H and ^{14}C samples. These gels were also fractionated and their radioactivities were determined. The mobilities of these fragments were 1–2% higher at low DNA concentrations (approximately 100 ng/gel), compared to the same fragments mixed with the Friend cell DNA samples (approximately 10 μg/gel; data not shown). For the standardization, values from the latter gels were used. Fig. 4 shows the nucleotide length of each fragment [determined from the nucleotide sequence (25)] plotted semilogarithmically versus its mobility. This yields a straight line, from which the sizes of the ^3H and ^{14}C nucleosomes were calculated.

Fig. 5 shows the variation in nucleosome size with multimer number, from which the nucleosome repeat length can be de-

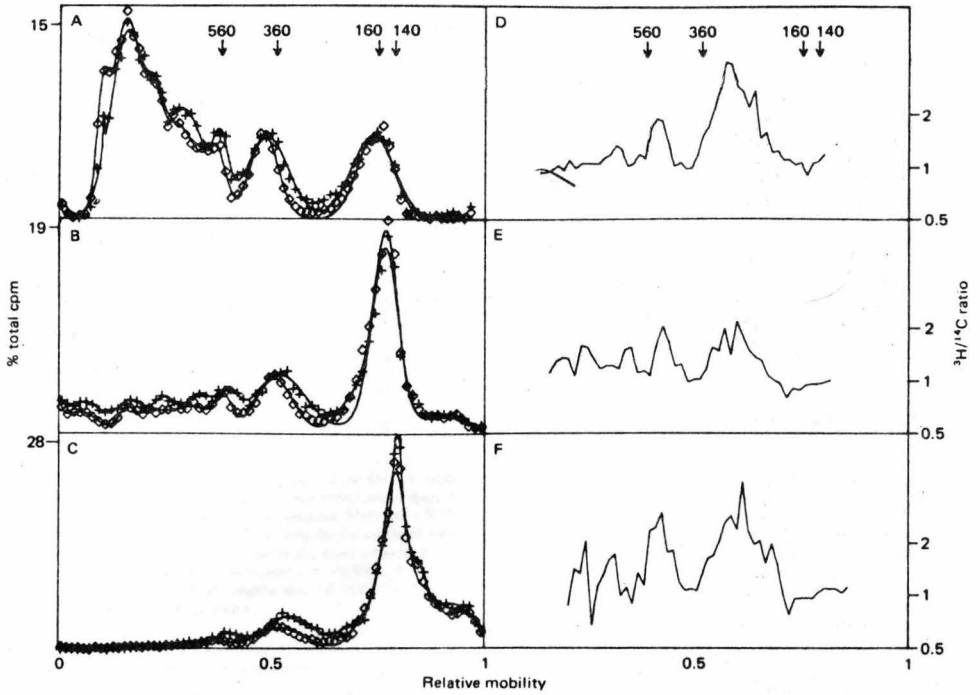


FIG. 3. Polyacrylamide (4%) gel electrophoresis of DNA from newly labeled and total nucleosomes. (*A–C*) Percent of total nuclear cpm for ^3H (+, new) and ^{14}C (◇, total) nucleosomal DNA is plotted versus mobility relative to bromophenol blue. ^3H total cpm: *A*, 4019; *B*, 4571; *C*, 3605. ^{14}C total cpm: *A*, 5772; *B*, 6259; *C*, 4873. The lines represent the sum of Gaussian curves fit to the data by using a nonlinear least-squares fitting program (R. F. Murphy, W. R. Pearson, and J. Bonner, unpublished). The arrows mark the approximate position of the ^{14}C core and multimers, and the associated numbers of base pairs were calculated from the standards as described in the text. The enzyme concentration increases 3.16-fold from *A* to *B* and from *B* to *C*. (*D–F*) Ratio of $^3H/^{14}C$ cpm from *A–C*.

termined independently of the extent of digestion. Table 2 summarizes this data. The difference in size is clearly due to a difference in spacer length, which varies from 10 base pairs in the least digested sample to 29 in the most. Because the breadth of the bands appears to decrease with digestion, the average value of 17 is probably an underestimate.

Table 2. Nucleosomal DNA sizes

	Chromatin	Mono	Di	Tri	Unit	Spacer*	Spacer difference
A†	New	168	387	563	198	58	
	Total	164	397	579	208	68	10
B†	New	151	349	528	188	48	
	Total	150	365	548	199	59	11
‡	New	144	331		187	47	
	Total	148	358		210	70	23
C†	New	140	323	527	194	54	
	Total	139	348	585	223	83	29
‡	New	138	320		182	42	
	Total	140	336		195	55	13
	Average						17 ± 8

* Assuming a core size of 140 base pairs.
† Fig. 3 panel number.
‡ Gel not shown; enzyme concentration as for preceding gel.

DISCUSSION

Much evidence has been accumulated to indicate that nucleosomes from different species and cell types have different repeat lengths (5–11). A correlation between nucleosome size and transcriptional activity has been suggested (9–10). As has been pointed out by Thomas and Thompson (11), however, it would appear unlikely that two different cells from the same organism would have a significant difference in their average repeat size because of differences in the small fraction of total DNA coding for structural genes. Some reports of the presence of transcribed genes in nucleosome structures have included data that indicate that the spacing for transcribed and nontranscribed DNA is the same. Two of these methods—comparison of unlabeled DNA in parallel gels (17) and comparison of unlabeled total DNA and labeled, hybridized probe for the examined sequence (15, 16, 18)—are severely limited in their accuracy due to difficulties in aligning and scaling the different sets of data. The third method—double-labeling—has been used to show that 60- to 120-min labeled ribosomal DNA and long-term labeled total DNA from *Tetrahymena* macronuclei nucleosomes have the same size (13, 14). No direct evidence of a relationship between nucleosome spacing and transcriptional activity has been reported.

As we have shown above, newly replicated nucleosomal DNA from mouse Friend cells is different in size from total nucleo-

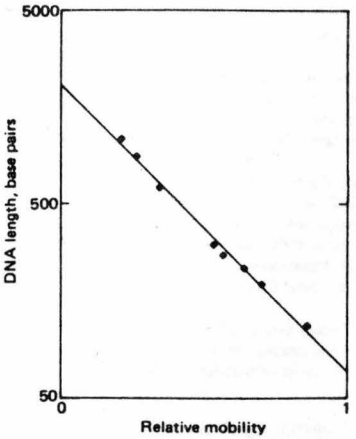


FIG. 4. Molecular weight standardization curve for 4% polyacrylamide gels. DNA length L (base pairs) is plotted semilogarithmically versus mobility relative to bromophenol blue for ϕX174 replicative form DNA cut with *Hae* III [♦, lengths determined from the DNA sequence (24)]. The line represents the least-squares fit log $L = 3.308 - 1.473\ R_F$ (correlation coefficient $r = -0.9963$, % error = 1.06) for the ϕX174 markers, whose lengths are 118, 194, 234, 271 and 278, 310, 606, 872, and 1078 base pairs.

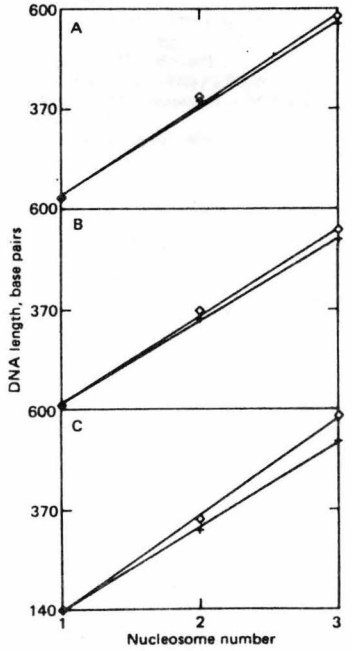


FIG. 5. Determination of nucleosome spacing. DNA length (base pairs, determined from the center of fitted Gaussian curves) is plotted versus nucleosome multimer number (n) for the gels from Fig. 3. +, ^{3}H; ◇, ^{14}C. The lines are least-squares fits $L = a + bn$, in which a is the sum of the lengths of the ends on each multimer minus the spacer length and b is the unit nucleosome size (core length plus the spacer length) (see Table 2).

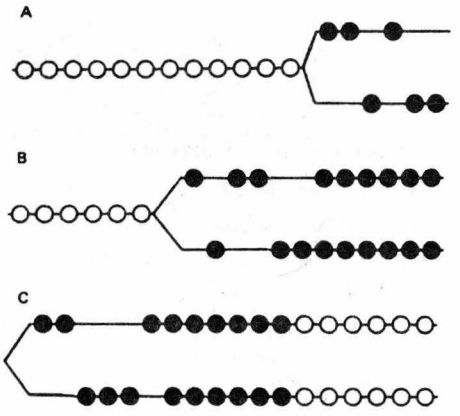


FIG. 6. Changes in nucleosome arrangement during chromatin replication. All circles represent 140-base-pair nucleosome cores. ○, nucleosomes with normal (200-base-pair) spacing; ●, nucleosomes with shortened (180-base-pair) spacing. The number of nucleosomes depicted in each configuration is arbitrary. In reality, the number of nucleosomes (and hence the length of time) between the passage of the replication fork and regaining of normal spacing is certainly much larger. (*A*) Chromatin after replication has been initiated. Compression of nucleosomes has occurred to allow for replication complex. A random dispersive model is shown, but semiconservative and conservative models are not ruled out. Current data are conflicting in this regard, but the proposed compression during replication is independent of the exact dispersion mechanism. Nuclease digestion of chromatin labeled for short times (1–30 sec, depending on the rate of replication for the specific cell type) yields an increased rate of production of both acid-soluble material and mononucleosomes. (*B*) New nucleosome cores cover the free DNA, but shortened spacing remains intact. Additional free DNA generated by compression is present at the fork. Longer pulses (5–10 min) detect no difference in production of acid-soluble material, decreased oligonucleosome production, and shortened nucleosome spacing. (*C*) Normal spacing is slowly regained by nucleosome sliding (20–30 min). The amount of free DNA at the fork is thus kept constant.

somal DNA. This might suggest a relationship between nucleosome spacing and chromatin replication rate, because the general correlation between shortened nucleosomal spacings and high transcriptional activity [reviewed by Kornberg (1)] can also be made between shortened spacing and short generation times. However, there is no difference in spacing between Chinese hamster ovary cell nuclei and mitotic chromosomes (17), or between confluent and exponentially growing C6 rat glial tumor cells (10). These experiments used unlabeled DNA in parallel gels, and hence do not have the sensitivity of our experiments. However, the conclusion that nucleosome spacing is not significantly related to rate of chromatin replication is probably justified, because the fraction of chromatin in newly replicated form is too small to account for spacing differences in total DNA from cells with different generation times. Instead, a gradual increase in nucleosome spacing with time is possible (perhaps to compensate for small amounts of histone degradation or as a result of changes in the level of histone modification).

Some laboratories have reported that chromatin pulse-labeled for from 0.5 to 10 min produces acid-soluble material at an increased rate (relative to total chromatin) when digested with staphylococcal nuclease (26–28). These studies have generally used shorter pulses than we have, and those that have used longer pulses have indicated that the difference disappears (26,

28). Considering the variation in chromatin replication rate among different cell lines, it would appear that the published data are consistent with an increased sensitivity during and immediately following DNA replication and normal sensitivity once histone cores are deposited on the free DNA. The short-lived acid-soluble difference is followed by a longer-term decrease in production of nucleosomes by staphylococcal nuclease, due to smaller nucleosome spacers (Fig. 6).

In any case, our results indicate that elongation can take place after histone deposition (perhaps caused by the binding of histone H1 and/or some nonhistone protein). This suggests that the interaction between core histones and DNA is not strong enough to prevent nucleosome "sliding," a conclusion that may be significant for models of chromatin replication and transcription.

After this manuscript was prepared, Levy and Jakob (29) and Seale (30) reported similar results for sea urchin embryos labeled for 7 sec and HeLa cell nuclei labeled *in vitro* for 20 min, respectively.

ϕX174 DNA cleaved with *Hae* III was the gift of Richard Parker. We thank Thomas Sargent for labeling the ϕX174 DNA and for helpful discussion of this manuscript. This work was supported by U.S. Public Health Service Grant GM-13762. R.F.M. is a National Institutes of Health predoctoral trainee. R.B.W. is the holder of a Medical Research Council of Canada Centennial fellowship.

1. Kornberg, R. D. (1977) *Annu. Rev. Biochem.* **46**, 931–954.
2. Whitlock, J. P., Jr. & Simpson, R. T. (1976) *Biochemistry* **15**, 3307–3314.
3. Bakayev, V. V., Bakayeva, T. G. & Varshavsky, A. J. (1977) *Cell* **11**, 619–629.
4. Noll, M. & Kornberg, R. D. (1977) *J. Mol. Biol.* **109**, 393–404.
5. Noll, M. (1976) *Cell* **8**, 349–355.
6. Morris, N. R. (1976) *Cell* **8**, 357–363.
7. Thomas, J. O. & Furber, V. (1976) *FEBS Lett.* **66**, 274–280.
8. Spadafora, C., Bellard, M., Compton, J. L. & Chambon, P. (1976) *FEBS Lett.* **69**, 281–285.
9. Morris, N. R. (1976) *Cell* **9**, 627–632.
10. Compton, J. L., Bellard, M. & Chambon, P. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 4382–4386.
11. Thomas, J. O. & Thompson, R. J. (1977) *Cell* **10**, 633–640.
12. Lohr, D., Corden, J., Tatchell, K., Kovacic, R. T. & van Holde, K. E. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 79–83.
13. Piper, P. W., Celis, J., Kaltoft, K., Lerr, J. C., Nielsen, O. F. & Westergaard, O. (1976) *Nucleic Acids Res.* **3**, 493–505.
14. Mathis, D. J. & Gorovsky, M. A. (1976) *Biochemistry* **15**, 750–755.
15. Bokhon'ko, A. & Reeder, R. H. (1976) *Biochem. Biophys. Res. Commun.* **70**, 146–152.
16. Bostock, C. J., Christie, S. & Hatch, F. T. (1976) *Nature (London)* **262**, 516–519.
17. Compton, J. L., Hancock, R., Oudet, P. & Chambon, P. (1976) *Eur. J. Biochem.* **70**, 555–568.
18. Gottesfeld, J. M. & Melton, D. A. (1978) *Nature (London)* **273**, 317–319.
19. Dube, S. K., Gaedicke, G., Kluge, N., Weimann, B. J., Melderis, H., Steinheider, G., Crozier, T. & Ostertag, W. (1974) in *Proceedings of the 4th Symposium: Princess Takamatsu Cancer Research Fund, Differentiation and Control of Malignancy in Tumor Cells* (Univ. of Tokyo Press, Tokyo), pp. 99–136.
20. Marmur, J. (1961) *J. Mol. Biol.* **3**, 208–218.
21. Maniatis, T., Jeffrey, A. & van deSande, H. (1975) *Biochemistry* **14**, 3787–3794.
22. Berkner, K. L. & Folk, W. R. (1977) *J. Biol. Chem.* **252**, 3176–3184.
23. Pearson, W. R., Davidson, E. H. & Britten, R. J. (1977) *Nucleic Acids Res.* **4**, 1727–1737.
24. Wallace, R. B., Sargent, T. D., Murphy, R. F. & Bonner, J. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3244–3248.
25. Sanger, F., Air, G. M., Barrel, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C. A., III, Slocombe, P. M. & Smith, M. (1977) *Nature (London)* **265**, 687–695.
26. Seale, R. L. (1975) *Nature (London)* **255**, 247–249.
27. Hildebrand, C. E. & Walters, R. A. (1976) *Biochem. Biophys. Res. Commun.* **73**, 157–163.
28. Burgoyne, L. A., Mobbs, J. D. & Marshall, A. J. (1976) *Nucleic Acids Res.* **3**, 3293–3304.
29. Levy, A. & Jakob, K. M. (1978) *Cell* **14**, 259–267.
30. Seale, R. L. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 2717–2721.

PROPERTIES OF NEWLY-REPLICATED CHROMATIN: ANALYSIS USING SHALLOW METRIZAMIDE GRADIENTS†

Robert F. Murphy*, R. Bruce Wallace‡ and James Bonner

Running title: Properties of newly-replicated chromatin

†From the Division of Biology, California Institute of Technology, Pasadena, CA 91125. This work was supported by U.S. Public Health Service Grant GM-13762.

‡Present address: Division of Biology, City of Hope National Medical Center, 1500 East Duarte Road, Duarte, CA 91010.

Abstract

The properties of chromatin containing newly synthesized DNA and protein are investigated. Nuclease digestion rates suggest that histones are segregated conservatively at the replication fork. A fraction of soluble nucleosomes enriched in newly-replicated DNA has been isolated by means of its increased density in metrizamide relative to bulk chromatin. This fraction is shown to be packaged into nucleosomes, but at an interval of approximately 160 base-pairs, as opposed to the 200 base-pair repeat seen for total chromatin. Evidence is presented that the density difference is due to this altered spacing. While pulse-labeled DNA is present in this dense fraction, nucleosomes labeled with short pulses of arginine or acetate are of normal density. The data presented are consistent with the conclusion that newly replicated DNA is associated with pre-existing histones in a short-lived, compact structure, while newly-synthesized histones are deposited at normal spacing some distance from the replication fork.

The events involved in the replication of eukaryotic chromatin have been analyzed by fixation of chromatin with formaldehyde (e.g., Fakan et al., 1972; Jackson et al., 1975; Seale, 1976a; Hancock, 1977) and inhibition of protein synthesis with cycloheximide (e.g., Seale & Simpson, 1975; Weintraub, 1976). Some of these investigations have yielded conflicting results. In addition, possible disruption of chromatin structure resulting from the methods of sample preparation and analysis used has led to objections to the results of previous experiments. In this paper we describe the determination of some of the properties of replicating chromatin using metrizamide density gradient centrifugation of native chromatin.

Materials and Methods

Cell culture and labeling. Friend leukemia cells, uninducible clone F4+, were grown in suspension culture as described previously (Keppel et al., 1977). The generation time under our conditions was 15.2 hr. For various experiments, cells were labeled with ^{14}C-thymidine (61 Ci/mmol, Amersham) at 4 nCi/ml, ^{14}C-L-arginine (312 mCi/mmol, Schwartz/Mann) at 50 nCi/ml, ^{3}H-thymidine (22 Ci/mmol, Amersham) at 40 uCi-min/ml, ^{3}H-L-arginine (8.8 Ci/mmol, Amersham) at 250 uCi-min/ml, and ^{3}H-acetic acid (2 Ci/mmol, Moravek Biochemicals) at 300 uCi-min/ml (1 Ci=3.7 x 10^{10} becquerels). All samples for scintillation counting were brought to 0.1% SDS in a final volume of 0.4 ml and counted in 3 ml of Aquasol-2 (New England Nuclear). Under these conditions, the spillover from the ^{3}H window to the ^{14}C window was 4%, and the ^{14}C to ^{3}H spillover was 11%. No effect on the counting efficiencies or spillover ratios was seen from 0-50% metrizamide. The presence of the SDS eliminated almost all self-quenching of large molecules.

Preparation of nuclei and nucleosomes. Nucleosomes were prepared from purified nuclei as described previously (Murphy et al., 1978), with the exception that micrococcal nuclease was purchased from Worthington (1 unit produces 1 A_{260} unit of acid-soluble material from DNA in 30 min at pH 8.0, 37^{o}C). To facilitate comparison of different experiments, digestions were carried out for 5 min at 37^{o}C at specified ratios of nuclease units to number of nuclei.

Metrizamide gradient centrifugation. Metrizamide was obtained from Nyegaard and Co. A/S, Oslo, Norway. All metrizamide solutions contained 2.5 mM Tris pH 8.0/ 2.5 mM ethyleneglycol-bis-(-aminoethyl ether)N,N'-tetraacetic acid (EGTA). 2.3 ml of sample in 30% (w/v) metrizamide was layered over 1.7 ml of 60% (w/v) metrizamide in a ½ x 2 in cellulose nitrate centrifuge tube, and overlayered with 1 ml of Tris/EGTA. Gradients were run at 10,000 rpm for 16 hr at 4°C in the Sorvall TV-865 rotor, and fractions collected dropwise from the bottom. Density was determined from the refractive index at 25°C (Birnie, 1978). DNA samples for electrophoresis were prepared by proteinase K (EM Laboratories) digestion of gradient fractions.

Polyacrylamide gel electrophoresis and fluorography. 4% polyacrylamide slab gels containing 89 mM Tris/89 mM Boric acid/ 2.5 mM EDTA (pH 8.3) were prepared as described by Maniatis et al. (1975). ^{3}H radioactivity was visualized by fluorography (Bonner & Laskey, 1974; Laskey & Mills, 1975). 2.5% polyacrylamide/ 0.5% agarose Tris/Borate/EDTA tube gels were prepared similarly. Tube gels (0.6 x 15 cm) were run for 3 hr at 100V and fractionated into 2 mm slices (Aliquogel fractionator, Gilson). Gel fractions were counted as described above. Counting of gels polymerized with various sizes of ^{3}H and ^{14}C DNA showed no quenching by the polyacrylamide fragments, and minimal dependence of the efficiency and spillover on DNA length.

Results

Nuclease sensitivity of pulse-labeled chromatin. In a previous paper (Murphy et al., 1978), we demonstrated that chromatin pulse-labeled for 10 min was digested to nucleosomes more slowly than bulk chromatin, using the criteria of solubility in Tris/EGTA. We attributed this to the shortened nucleosome spacing of newly-replicated chromatin (Murphy et al., 1978; Levy & Jakob, 1978; Seale, 1978b). While an increased rate of digestion to acid-soluble DNA by micrococcal nuclease had been reported for pulse-labeled chromatin (Hildebrand & Walters, 1976; Seale, 1976b), we did not observe this difference for our 10 min labeled material. Because of possible differences in replication rates between different cell lines, we have determined the relative rates of production of acid-soluble and Tris/EGTA soluble material for cells labeled for 1-30 min. Figure 1 shows that digestion of newly-replicated chromatin results in an increased rate of production of acid-soluble material, but a decreased rate of production of soluble nucleosomes. In view of the smaller spacing and decreased nucleosome production of newly-replicated chromatin, we attempted to fractionate unfixed nucleosomes using metrizamide gradients.

Buoyant density of pulse-labeled chromatin. Cells were labeled with ^{14}C-thymidine for 24 hr, and then with ^{3}H-thymidine for 1, 10 or 100 min. After brief digestion with micrococcal nuclease, the total nucleosomal fraction was centrifuged in metrizamide. Figure 2 shows that the 1 min

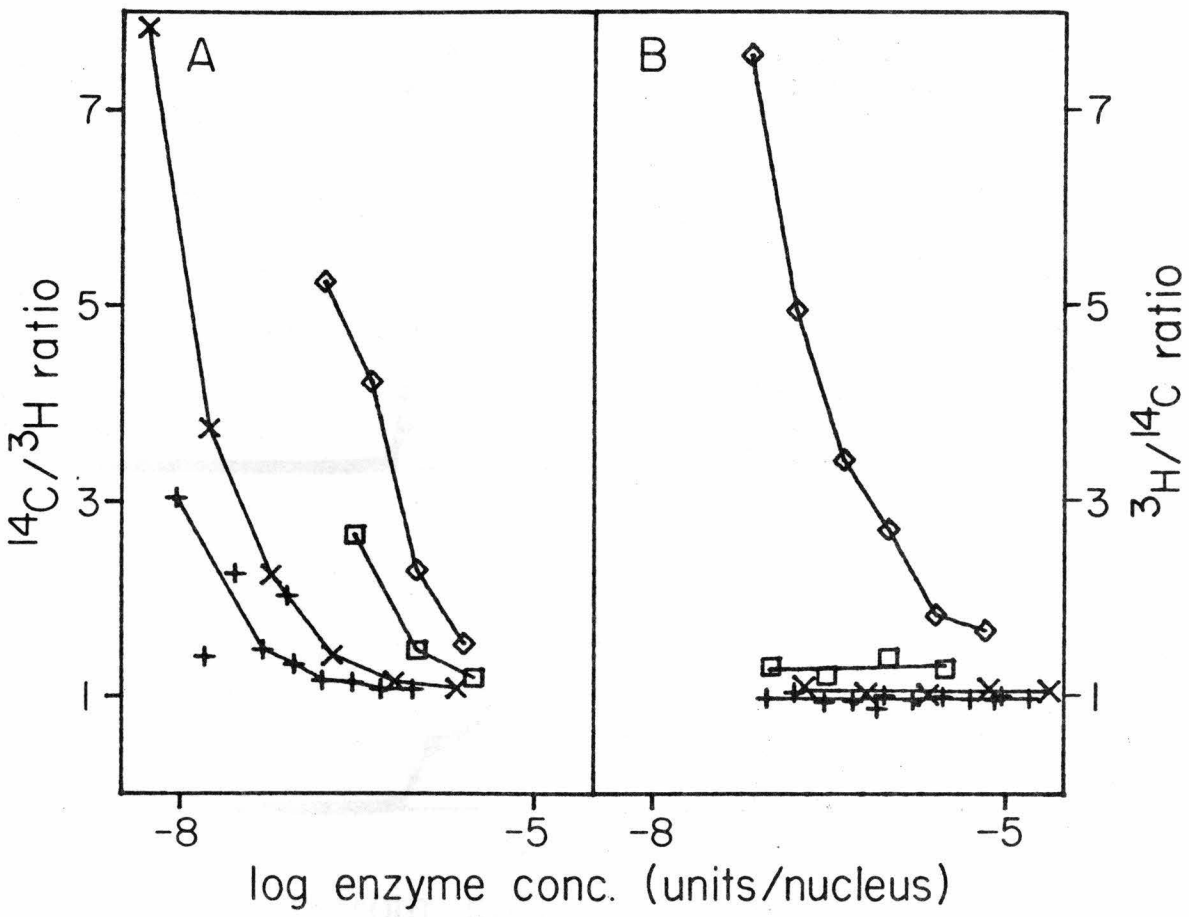


Figure 1. Digestion of pulse-labeled chromatin with micrococcal nuclease. Cells were labeled for 24 hr with ^{14}C-thymidine, followed by 1 min (◇), 5 min (□), 10 min (X), or 30 min (+) with ^{3}H-thymidine. A) $^{14}C/^{3}H$ ratios for Tris/NaCl/Mg soluble fraction (equivalent to acid soluble). B) $^{3}H/^{14}C$ ratios for Tris/EGTA soluble nucleosomes.

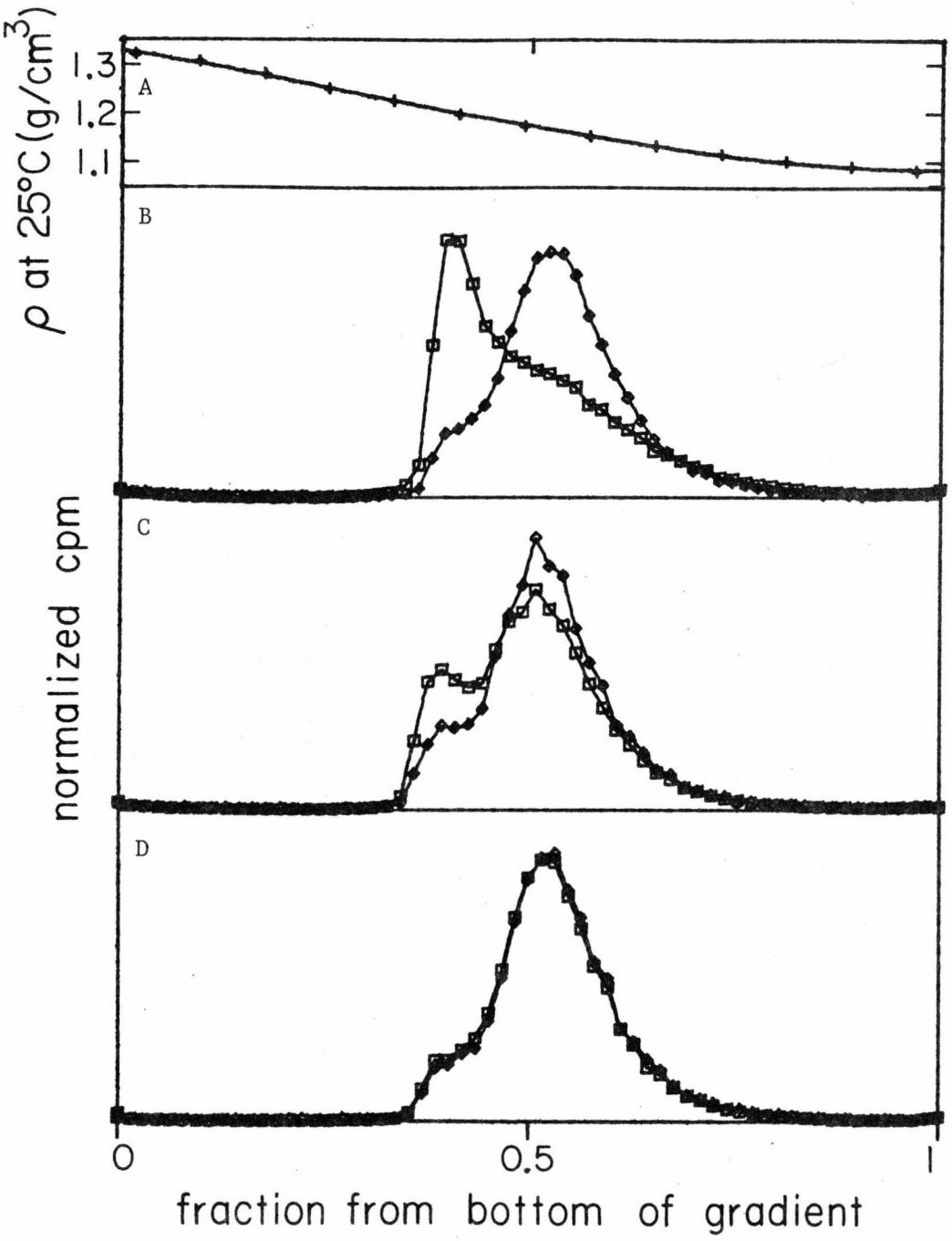


Figure 2. Separation of newly-replicated chromatin from bulk chromatin in metrizamide gradients. Cells were labeled for 24 hr with ^{14}C-thymidine and then for 1 min (B), 10 min (C) or 100 min (D) with ^{3}H-thymidine. After digestion with micrococcal nuclease at 5.33 x 10^{7} units/nucleus and removal of acid-soluble material (2.4% ^{14}C; 8.8, 2.4, 2.4% ^{3}H), the total nucleosomal fraction was centrifuged in metrizamide. The data in B-D are plotted so that the total area of each panel represents seven times the number of cpm in the gradient. + density. ◊ ^{14}C cpm. □ ^{3}H cpm. (See Table I.)

Table I. Density of chromatin fractions in metrizamide gradients of Figure 2[a]

	Heavy			Light	
	^{3}H	%T[b]	^{14}C	%T	^{3}H	^{14}C
1 min	1.2034	28.2	1.2070	2.9	1.1712	1.1664
10 min	1.2106	10.4	1.2107	6.0	1.1730	1.1701
100 min	1.2097	3.6	1.2093	3.4	1.1678	1.1669
Mean	1.2079	----	1.2090	4.1	1.1707	1.1678
SD	0.0039	----	0.0019	1.7	0.0026	0.0020

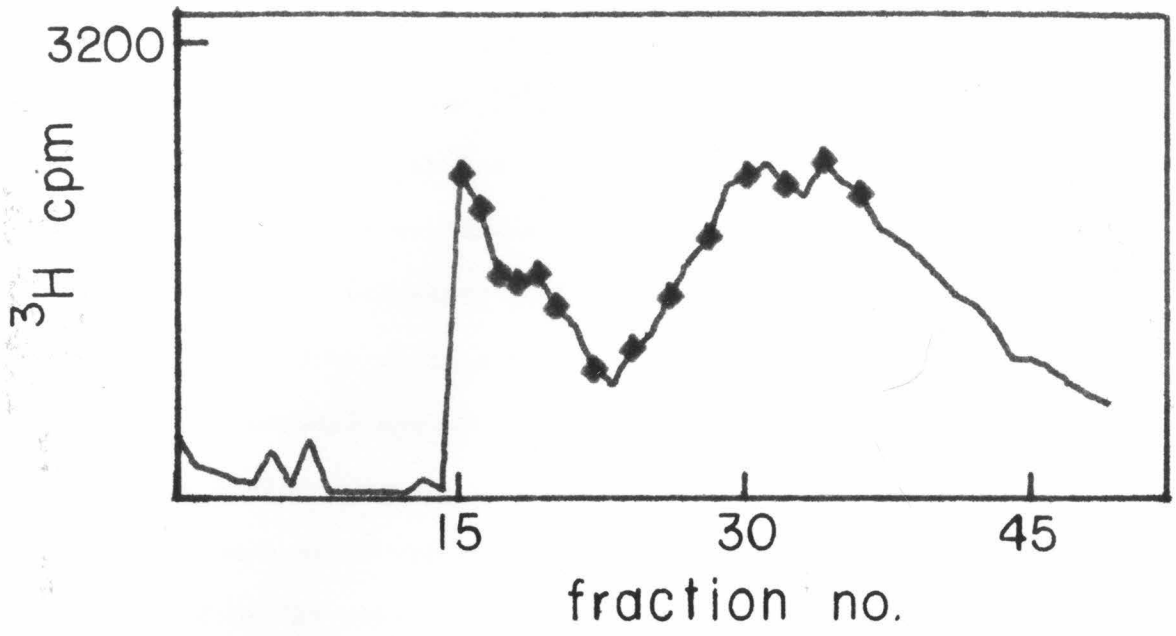
[a]Peak densities and percentages were determined using a non-linear least-squares fitting program (Murphy et al., 1979) to fit two Gaussian curves to the data.
[b]Percent of total cpm in gradient.

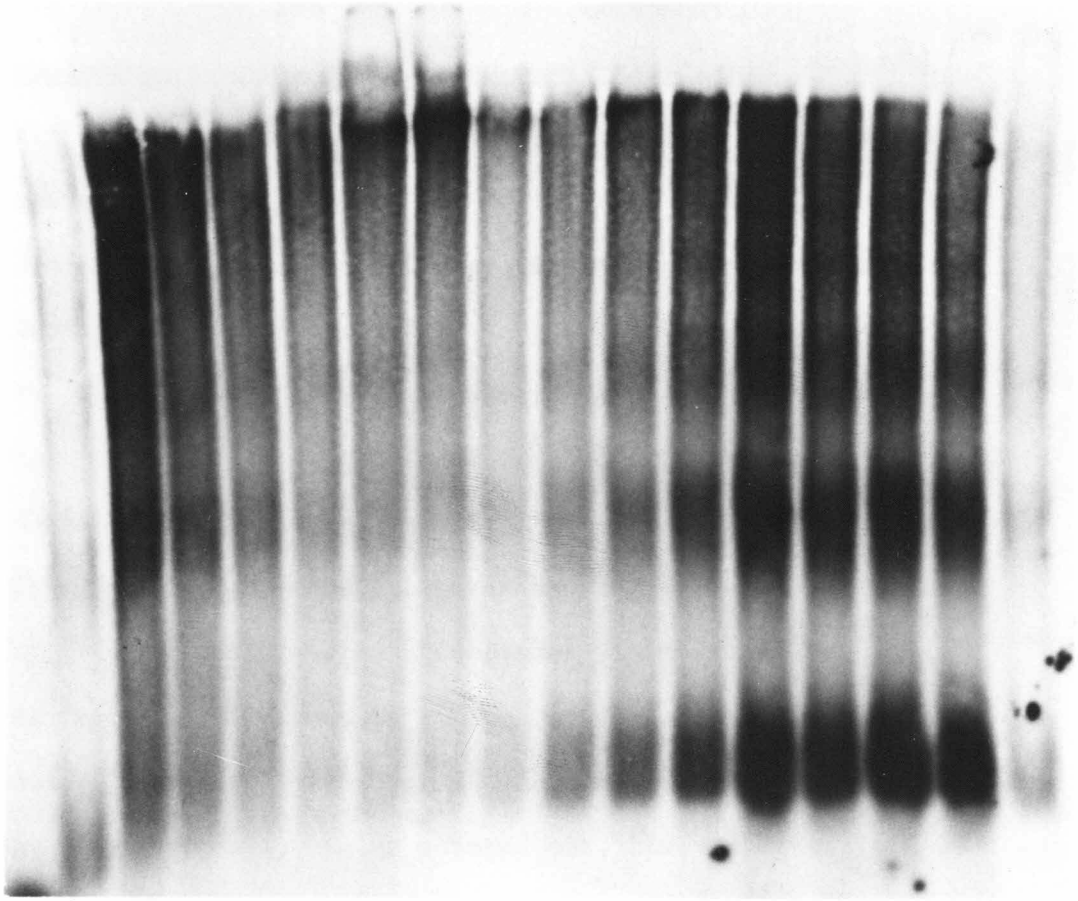
labeled nucleosomes are more dense than the long-term labeled nucleosomes, and that this density difference disappears on longer labeling. Table I gives the densities of the two chromatin fractions, and shows that bulk chromatin has a dense shoulder constituting approximately 5% of the total DNA. Birnie et al. (1973) found the density of native mouse DNA to be 1.118 g/cm^3, and the density of purified proteins to vary from 1.24 to 1.29 g/cm^3. The density of nuclear DNA (proteinase K treated) in metrizamide is 1.12 g/cm^3, and that of nucleosomal protein (DNAse I treated) is 1.23 g/cm^3 (data not shown). Thus, it is likely that the increased density of newly-replicated chromatin is due to an increased protein/DNA ratio, as would be expected due to its decreased nucleosomal spacing.

To confirm that the material in the dense peak is present in nucleosomes, the material from that peak was isolated and run on isokinetic sucrose gradients (Noll, 1967). The ^3H labeled material showed the typical nucleosomal pattern (data not shown). Figure 3 also shows that the DNA of this fraction has nucleosomal spacing. Since there is more newly-replicated DNA in the mononucleosome bands of the light fractions than in the corresponding bands of the denser fractions, the separation may be at least partly due to a decreased rate of production of low multimers of newly-replicated chromatin.

Removal of density difference by extensive nuclease digestion. If the density difference were totally due to the presence of newly-replicated DNA in higher nucleosome multimers,

Figure 3 (next page). Nucleosome distribution of metrizamide gradient fractions. Nuclei labeled for 3 min with ^{3}H-thymidine were digested at 2 x 10^{6} units/nucleus (18% acid-soluble), and the total nucleosomal fraction was centrifuged in metrizamide. A) ^{3}H distribution across gradient. B) 13 day fluorogram of a 4% polyacrylamide slab gel. The outermost slots contain DNA from an aliquot of the nucleosomes before centrifugation. The other slots (from left to right) contain DNA from the fractions marked by diamonds in A.


3200
^{3}H cpm
15
30
45
fraction no.



it would disappear if soluble nucleosomes (i.e., low multimers) were isolated and run. That this is not the case is demonstrated by the fact that 1 min labeled soluble nucleosomes are more dense than long-term labeled nucleosomes (Figure 4). However, the density difference disappears as the chromatin is digested to monomers, as would be predicted if the higher density were due to decreased nucleosomal spacing.

Figure 4 also suggests that there is a shift in density of bulk nucleosomes as nuclease digestion proceeds. This is further demonstrated by the data from a number of experiments summarized in Table II.

Nucleosome spacing of isolated newly-replicated chromatin. The data presented above show that it is possible to isolate a fraction of nucleosomes which are enriched in newly-replicated DNA. To confirm that indeed the nucleosome spacing of this fraction is significantly smaller than that of bulk chromatin from the same digest, DNA from these fractions was isolated and analyzed on polyacrylamide gels (Figure 5). Estimating molecular weight from relative mobility using ^{32}P-labeled PBR322 DNA digested with the restriction endonuclease HinF1, the repeat spacing of the 1 min labeled nucleosomes is about 160 bp, while that of the bulk chromatin is 200 bp. This is smaller than our previous estimate of 180 bp for 10 min labeled nucleosomes (Murphy et al., 1978). As we pointed out, however, this is due to the reversion to normal spacing which occurs in the 10 min labeling period.

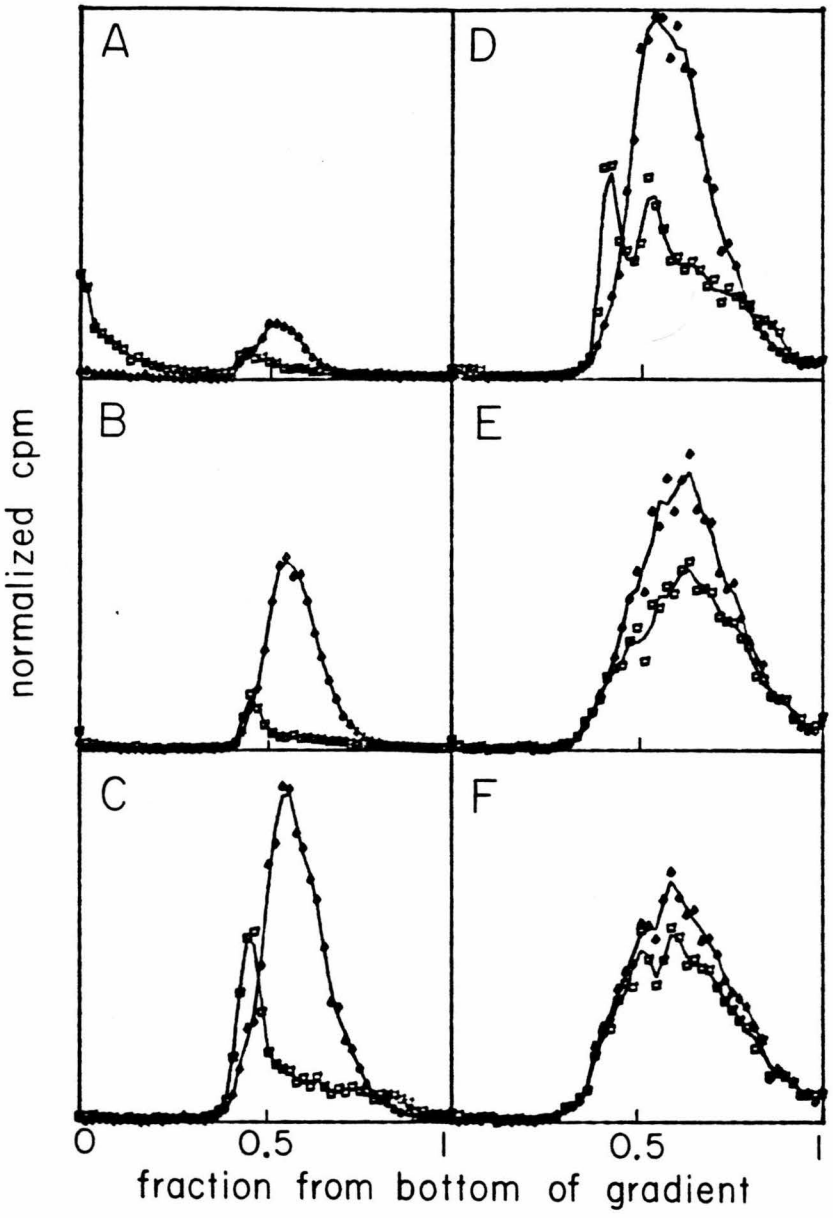


Figure 4. Effect of extent of digestion on separation of newly-replicated chromatin from bulk chromatin. Tris/EGTA-soluble nucleosomes from cells labeled for 24 hr with ^{14}C-thymidine (◇) followed by 1 min with ^{3}H-thymidine (□) were centrifuged in metrizamide. The total area of each panel is six times the total nuclear cpm. Digestions (A-F) were from 1.67×10^{8} to 6.67×10^{6} units/nucleus. ^{14}C acid-soluble increases from 0.6% to 23%, and ^{3}H acid-soluble from 4% to 39%. The lines through the points are 5 point quadratic smooths of the data (Savitzky & Golay, 1964).

Table II. Buoyant density of nucleosomes as a function of extent of digestion[a]

Enzyme Range[b]	Thymidine N[c]	Thymidine Mean	Thymidine SD	Arginine N	Arginine Mean	Arginine SD
-8 to -7.1	5	1.1577	0.0024	3	1.1580	0.0078
-7 to -6.1	7	1.1517	0.0059	7	1.1547	0.0068
-6 to -5.1	5	1.1363	0.0049	2	1.1417	0.0007

[a]Peak densities were determined by fitting a Gaussian curve to data for 24 hr labeled material.
[b]Log of enzyme concentration (units/nucleus); lower and upper limits for grouped data.
[c]Number of gradients.

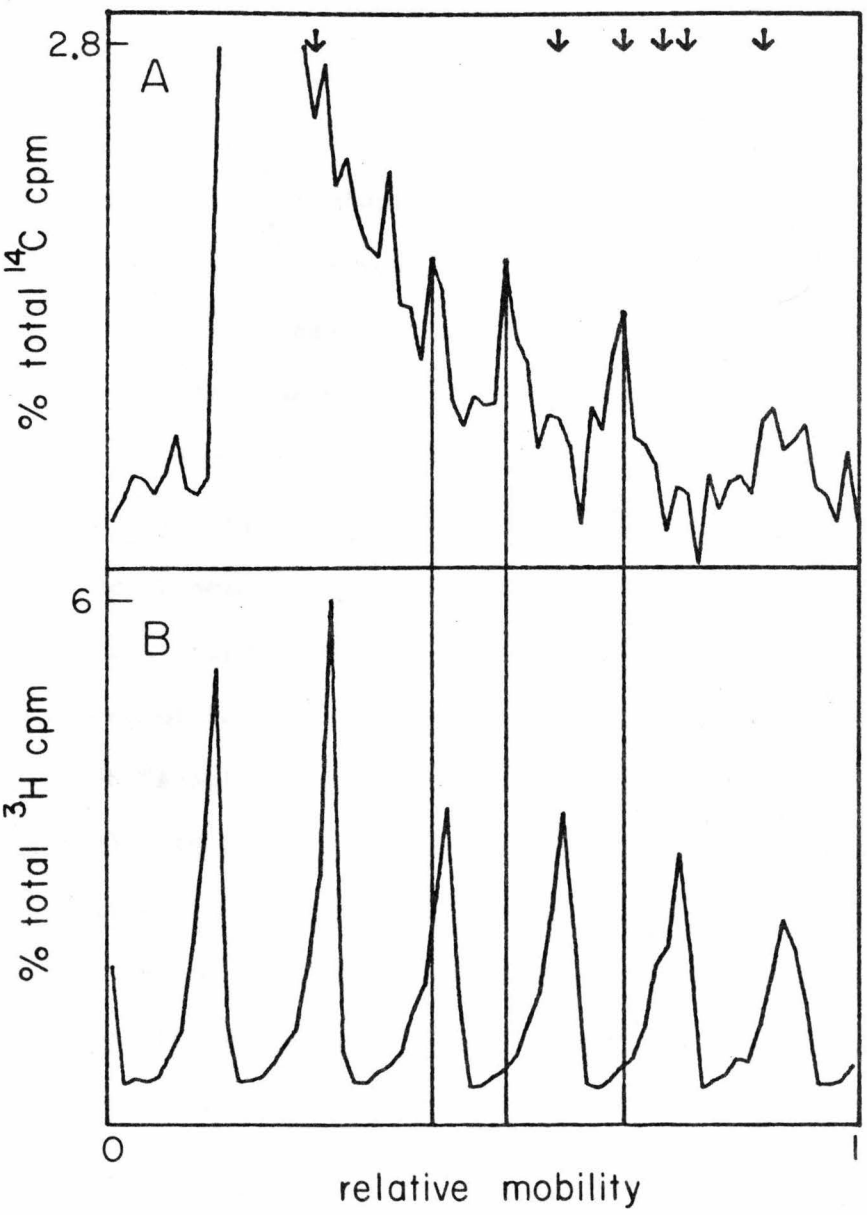


Figure 5. Nucleosomal DNA spacing of heavy fraction from metrizamide gradients. 2.5% polyacrylamide/0.5% agarose gels. Samples were from the gradient shown in Figure 4C. A) ^{14}C cpm for Tris/EGTA soluble fraction before centrifugation. B) ^{3}H cpm for heavy peak from gradient (fractions 25&26). Mobility is relative to bromphenol blue. The arrows mark the positions of ^{32}P-labeled PBR322 DNA cut with HinF1, and correspond to DNA sizes of 1631, 516/506, 396, 344, 298, and 221/220 base-pairs.

Buoyant density of newly-deposited nucleosomal protein.

Some investigators have suggested that newly-synthesized chromosomal protein is deposited on unreplicated DNA (Jackson et al., 1976; Seale, 1976a; Hancock, 1977; Hancock, 1978). There has been some uncertainty about this conclusion, since CsCl gradients of formaldehyde-fixed chromatin were used. To determine the degree of association of newly-synthesized histone with newly-replicated DNA using unfixed chromatin, cells were labeled with ^{14}C-L-arginine for 24 hr, and then with ^{3}H-L-arginine for 5 min. Tris/EGTA soluble nucleosomes from these cells were run on metrizamide gradients after digestion to varying extents. Figure 6B shows one of these gradients, along with a gradient of 24 hr ^{14}C-thymidine/ 5 min ^{3}H-thymidine labeled Tris/EGTA soluble nucleosomes (at the same nuclease concentration). In no case were ^{3}H-arginine counts detected in the dense region. Indeed, the ^{14}C and ^{3}H counts were nearly identical, as can be seen in Figure 6. Similar results were obtained for cells labeled with arginine for 10 or 100 min, and for cells labeled with acetate for 1, 10 or 100 min.

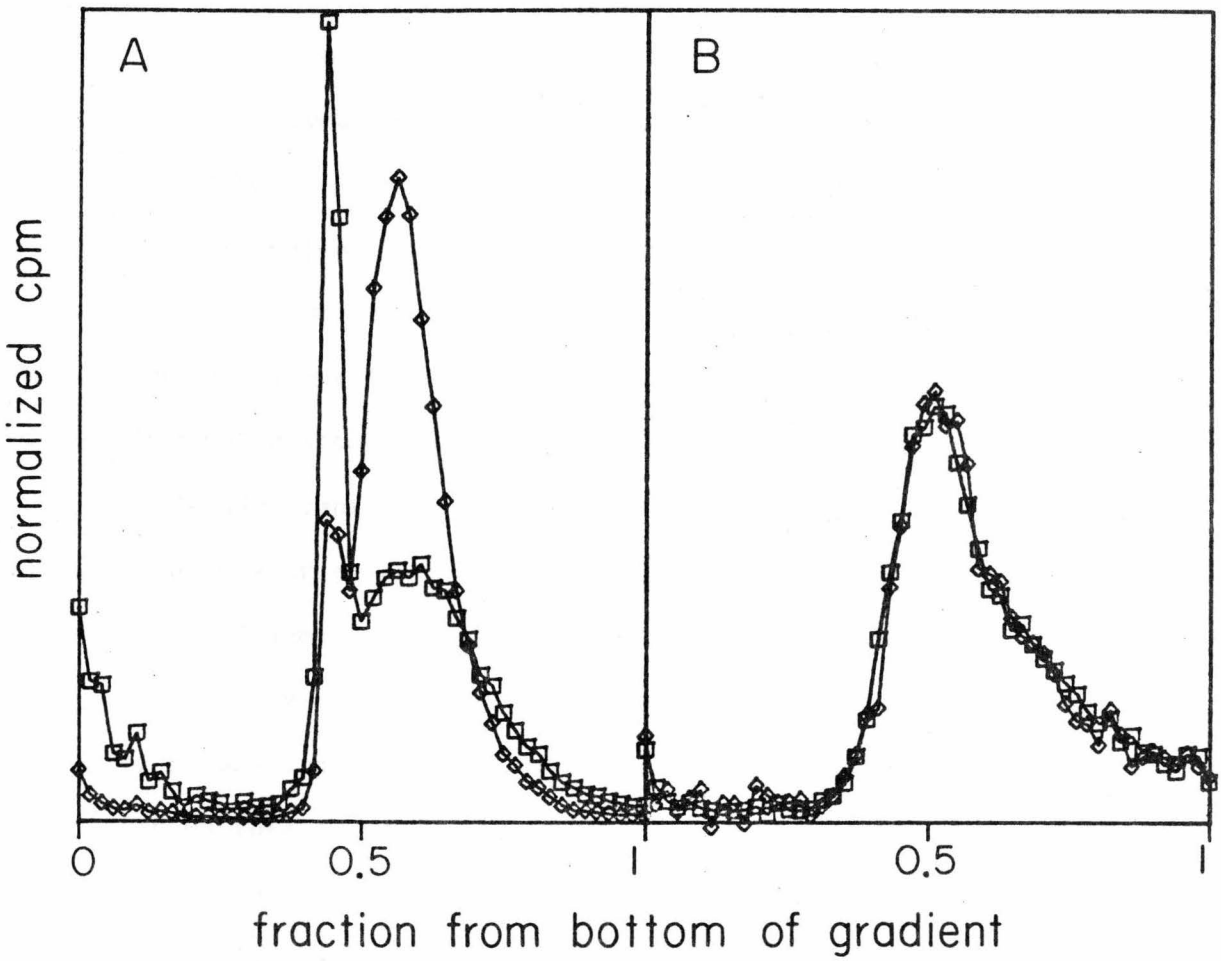


Figure 6. Comparison of density distribution of thymidine and arginine labeled Tris/EGTA soluble nucleosomes. Digestion was at 1.0 x 10^7 units/nucleus. A) 24 hr ^{14}C-thymidine (◇ ; 2.8% acid-soluble)/ 5 min ^{3}H-thymidine (□ ; 3.7% acid-soluble). B) 24 hr ^{14}C-L-arginine (◇)/ 5 min ^{3}H-L-arginine (□). The total area of each panel is 6.6 times the total cpm in each gradient.

Discussion

Nuclease sensitivity and histone segregation mechanism.

Chromatin pulse-labeled with thymidine for short periods of time has been shown to be digested to acid-soluble material more quickly than bulk chromatin by micrococcal (staphylococcal) nuclease (Hildebrand & Walters, 1976; Seale, 1976b; Levy & Jakob, 1978) and DNase I (Seale, 1975; Burgoyne et al., 1976). This increased production of acid-soluble material has also been demonstrated for chromatin from cells which have been exposed to cycloheximide (Seale & Simpson, 1975; Weintraub, 1976). While Seale (1976b) suggested that newly-replicated chromatin was digested to nucleosomes faster than total chromatin, Marshall & Burgoyne (1976) demonstrated a slightly decreased rate of production of nucleosomes for 10 min labeled chromatin using sucrose gradients. We have confirmed this decreased rate of production of nucleosomes by measuring solubility in Tris/EGTA, which is proportional to fragment length (Murphy et al., 1978 and Figure 1).

These seemingly contradictory results can best be explained by assuming a conservative mechanism of histone segregation at the replication fork. Such a mechanism has been demonstrated using cycloheximide treated cells (Weintraub, 1976; Riley & Weintraub, 1979) and by histone-histone cross-linking (Leffak et al., 1977). The fact that other investigators (e.g., Jackson et al., 1975) have presented evidence for a random mechanism may be a result of problems involved in chromatin fixation, or

difficulties in correcting for the presence of non-histone protein.

Metrizamide fractionation of chromatin. Because metrizamide is an inert, non-ionic density gradient material, it has been used to examine chromatin properties (e.g., Birnie et al., 1973). Rickwood et al. (1974) demonstrated the separation of chromatin into two fractions as a function of the extent of mechanical shearing, and took this as an indication of the presence of protein-rich and protein-poor regions. The data of Figure 4 and Table II suggest that the separation is instead due to the decreased density of low nucleosome multimers. This shift may well be due to the removal of non-histone proteins and/or histone H1 during digestion (Noll & Kornberg, 1977).

Buoyant density of pulse-labeled chromatin. In CsCl gradients of formaldehyde-fixed chromatin, a decreased density has been reported for pulse-labeled chromatin (Fakan et al., 1972; Seale & Simpson, 1975; Jackson et al., 1976), cycloheximide-treated chromatin (Seale & Simpson, 1975), and in vitro labeled chromatin (Seale, 1978a). This decreased density in CsCl corresponds to the increased density in metrizamide demonstrated above, since protein is less dense than DNA in CsCl, but more dense than DNA in metrizamide. Both increased (Levy et al., 1975) and decreased (Burke & Pearson, 1976) densities in metrizamide have been reported for pulse-labeled chromatin which had been mechanically sheared. As Noll et al. (1975) have pointed out, there is significant disruption of nucleosome

structure during mechanical shearing. In addition, the results presented above demonstrate the importance of controlling the extent of shearing during sample preparation. These facts make interpretation of the previous reports difficult.

Hancock (1974) has shown that the preparation of nuclei and chromatin by methods very similar to those used in these experiments do not allow exchange of histones between deoxyribonucleoprotein molecules. The lack of histone exchange, the fact that nuclease digestion was used to prepare nucleosome samples, and the use of the non-ionic density gradient material make it very unlikely that the results presented above are due to distortion or disruption of the chromatin structure.

Deposition of newly-synthesized histone. Having established a method for isolating native nucleosomes containing newly-replicated DNA, it was then possible to examine the distribution of newly-synthesized histones. Figure 6 shows that no significant deposition of this histone occurs on new DNA during the lifetime of the closely packed nucleosomal structure. This is in agreement with the results reported for fixed chromatin (Jackson et al., 1976; Seale, 1976a; Hancock, 1977; Hancock, 1978). Our data do not rule out the possibility that new histones are deposited on the unpackaged new DNA strand at normal spacing. However, previous results with fixed chromatin would appear to make this unlikely.

Acknowledgments

We thank Thomas D. Sargent for the gift of the ^{32}P-labeled PBR322 markers, and for helpful discussion of this manuscript.

References

Birnie, G.D. (1978) Methods in Cell Biology 18, 23-29.

Birnie, G.D., Rickwood, D., & Hell, A. (1973) Biochim. Biophys. Acta 331, 283-294.

Bonner, W.M., & Laskey, R.A. (1974) Eur. J. Biochem. 46, 83-88.

Burgoyne, L.A., Mobbs, J.D., & Marshall, A.J. (1976) Nucl. Acids Res. 3, 3293-3304.

Burke, J.F., & Pearson, C.K. (1976) Biochem. Soc. Trans. 4, 787-789.

Fakan, S., Turner, G.N., Pagano, J.S., & Hancock, R. (1972) Proc. Natl. Acad. Sci. USA 69, 2300-2305.

Hancock, R. (1974) J. Mol. Biol. 86, 649-663.

Hancock, R. (1977) Methods in Cell Biology 16, 459-466.

Hancock, R. (1978) Proc. Natl. Acad. Sci. USA 75, 2130-2134.

Hildebrand, C.E., & Walters, R.A. (1976) Biochem. Biophys. Res. Comm. 73, 157-163.

Jackson, V., Granner, D.K., & Chalkley, R. (1975) Proc. Natl. Acad. Sci. USA 72, 4440-4444.

Jackson, V., Granner, D., & Chalkley, R. (1976) Proc. Natl. Acad. Sci. USA 73, 2266-2269.

Keppel, F., Allet, B., & Eisen, H. (1977) Proc. Natl. Acad. Sci. USA 74, 653-656.

Laskey, R.A., & Mills, A.D. (1975) Eur. J. Biochem. 56, 335-341.

Leffak, I.M., Grainger, R., & Weintraub, H. (1977) Cell 12, 837-845.

Levy, A., & Jakob, K.M. (1978) Cell 14, 259-267.

Levy, A., Jakob, K.M., & Moav, B. (1975) Nucl. Acids Res. 2, 2299-2303.

Maniatis, T., Jeffrey, A., & van deSande, H. (1975) Biochemistry 14, 3787-3794.

Marshall, A.J., & Burgoyne, L.A. (1976) Nucl. Acids Res. 3, 1101-1110.

Murphy, R.F., Wallace, R.B., & Bonner, J. (1978) Proc. Natl. Acad. Sci. USA 75, 5903-5907.

Murphy, R.F., Pearson, W.R., & Bonner, J. (1979) Nucl. Acids Res., in press.

Noll, M. (1967) Nature 215, 360-363.

Noll, M., & Kornberg, R.D. (1977) J. Mol. Biol. 109, 393-404.

Noll, M., Thomas, J.O., & Kornberg, R.D. (1975) Science 187, 1203-1206.

Rickwood, D., Hell, A., Malcolm, S., Birnie, G.D., MacGillivray, A.J., & Paul, J. (1974) Biochim, Biophys. Acta 353, 353-361.

Riley, D., & Weintraub, H. (1979) Proc. Natl. Acad. Sci. USA 76, 328-332.

Savitzky, A., & Golay, M.J.E. (1964) Anal. Chem. 36, 1627-1639.

Seale, R.L. (1975) Nature 255, 247-249.

Seale, R.L. (1976a) Proc. Natl. Acad. Sci. USA 73, 2270-2274.

Seale, R.L. (1976b) Cell 9, 423-429.

Seale, R.L. (1978a) Cold Spring Harbor Symp. Quant. Biol. 42, 433-438.

Seale, R.L. (1978b) Proc. Natl. Acad. Sci. USA 75, 2717-2721.

Seale, R.L. & Simpson, R.T. (1975) J. Mol. Biol. 94, 479-501.

Weintraub, H. (1976) Cell 9, 419-422.

PART V

Chapter 6. Computer programs for analysis of nucleic acid hybridization, thermal denaturation, and gel electrophoresis data

COMPUTER PROGRAMS FOR ANALYSIS OF NUCLEIC ACID HYBRIDIZATION, THERMAL DENATURATION, AND GEL ELECTROPHORESIS DATA

Robert F. Murphy, William R. Pearson* and James Bonner

Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA

*Department of Microbiology, Johns Hopkins Medical School, Baltimore, MD 21205, USA

ABSTRACT

Computer programs for the analysis of data from techniques frequently used in nucleic acids research are described. In addition to calculating non-linear least-squares solutions to equations describing these systems, the programs allow for data editing, normalization, plotting and storage, and are flexible and simple to use. Typical applications of the programs are described.

INTRODUCTION

The increasing complexity and volume of data being generated in biochemical and biophysical experimentation, and the proliferation of mini-computer systems, has created a need for portable, interactive data storage and analysis programs. This paper describes non-linear least-squares fitting programs which have been used for the analysis of data from gel electrophoresis (1-4), DNA-DNA and RNA-DNA hybridization (5-7), and DNA, chromatin and protein-DNA thermal denaturation (8-10). These programs have been implemented on a PDP-11 mini-computer system, and are written in FORTRAN for ease of transfer to other computer systems. The programs require approximately 16,000 16-bit words of memory and a disk mass storage device (such as flexible disk). Mini-computers capable of running this data analysis system are currently available for less than $5,000.

In this paper we will first describe three different programs in the package, and then discuss the reliability and significance of the parameter estimates calculated by the programs.

PROGRAM DESCRIPTIONS

The three programs in the least-squares fitting package are COTFIT, for analysis of nucleic acid hybridization and denaturation data, GELFIT, for determination of the positions and areas of Gaussian curves fit to gel electrophoresis data, and MELSMR, for removal of noise from closely-spaced data. These programs use a common data file format, which allows interaction of the various programs for uses other than those described above. For example, CsCl density gradient data could be entered with COTFIT, smoothed with MELSMR, and then analyzed with GELFIT.

COTFIT

COTFIT is a non-linear least-squares fitting program developed from the NNNBAT program described by Pearson, Davidson and Britten (11), which was in turn based on the FINGER program of Britten, Graham and Neufeld (12). The program accepts English commands to control data entry and fitting. Data may be read from disk files or entered at the terminal. In addition to least-squares fitting, COTFIT provides general facilities for entering and editing data, and offers a variety of options for plotting and printing the curves calculated from the data.

Initial parameter estimates are improved by a modification of the method of Marquardt (13). Parameter values which would produce undefined function values are detected without causing arithmetic errors. This enables the fitting routine to try a broad range of possible parameter values without causing program termination.

As an additional option, files containing function values (with or without specified errors in the parameters) over a given interval can be generated. This feature is useful for displaying the functions under various conditions.

The functions described by Pearson, Davidson and Britten (11) have been modified, and a new function has been added. The NNNBAT function names and the corresponding COTFIT names are listed below.

N	COTFIT	NNNBAT	DESCRIPTION
1	FINGER	FINGER	second order DNA renaturation.
2	DIGEST	WHATOR	variable order renaturation.
3	DRIVEN	NUFORM	tracer/driver reaction with different nucleation rates.
4	EXCESS	EXCESS	first order renaturation.
5	MELTFN	------	thermal denaturation function.

The MELTFN function is

$$H(T) = I + \sum_{i=1}^{m} Fi\{0.5 + \frac{1}{\sqrt{\pi}} \int_{Tmi}^{T} \exp\left[\frac{-(x-Tmi)^2}{2Si^2}\right]dx\} \quad (1)$$

where H is the hyperchromicity at temperature T, I is the initial hyperchromicity, m is the number of components, Fi, Tmi, and Si are the hyperchromicity, transition midpoint and transition width (equivalent to the standard deviation of a Gaussian curve) of component i. The function is a normalized form of the error function, which is the integral of the Gaussian distribution. The program uses a polynomial approximation to this integral (14). The addition of this function allows direct fitting of thermal denaturation data, as opposed to the more common use of the Gaussian function to fit derivatized data. Although the two methods are theoretically identical, in practice the new approach eliminates errors introduced by the derivatization and smoothing process, and requires significantly fewer data points to determine component parameters. This is especially useful for analyzing DNA melts assayed by hydroxyapatite or filter binding (R.B. Wallace, G. Schaeffer, T. Hirose, K. Itakura, R.F. Murphy and J. Bonner, submitted to Nucleic Acids Res).

GELFIT

The GELFIT program uses a set of commands which is compatible with that of the COTFIT program, and fits Gaussian curves to integer data with evenly-incremented X values. The program can calculate the molecular weight of the species in a given band using a polynomial approximation to standards data specified by the user. The simplicity of the function being fitted allows the program to make initial estimates of the number and position of the bands in a given gel, and then proceed with fitting. The restriction to Y-only integer data was added to minimize data storage space and computation time in view of the

large numbers of data points commonly collected during gel scanning in our laboratory. The fitting method is similar to that used by COTFIT. In order to reduce program size, a band matrix (15) is used in place of the symmetric triangular matrix COTFIT uses. The number of adjacent curves whose values are allowed to affect an individual curve's parameter estimates can be adjusted at run-time by the user.

MELSMR

The MELSMR program smooths and/or derivatizes data having a constant X increment between points using the method of Savitzky and Golay (16-18). The degree to which the data are smoothed is controlled by the user. Digital data which have been collected from instruments such as spectrophotometers frequently contain fluctuations in the less significant digits. MELSMR provides a means of reducing or eliminating this noise.

RESULTS AND DISCUSSION

While these programs can significantly shorten the time required to analyze gel and melt data, and are essential for accurate measurement of nucleic acid hybridization rates and component amounts, the significance of the calculated parameter estimates must not be overestimated. This section addresses three issues encountered in fitting nucleic acid data with least squares programs: 1) The significance of the exponent in S1 nuclease-assayed reassociation data analysis; 2) the effect of using equation (1) to analyze melt data; and 3) the reproducibility of melt data parameter estimates.

Analysis of nuclease-assayed reassociation data

Morrow (19) and Smith, Britten and Davidson (20) have analyzed the kinetics of DNA-DNA reassociation assayed by the single-strand-specific S1 nuclease of *Aspergillus oryzae*. They concluded that data from S1 nuclease assayed renaturations were best fit using the equation

$$\frac{S}{Co} = (1 + kCot)^{-n} \qquad (2)$$

where S is the S1 nuclease sensitive (single stranded) DNA NT concentration, Co is the total DNA NT concentration, t is time, k

Table 1. Parameters for equation (3) for data of Sala-Trepat et al (6)

DRIVER*	c-DNA+	N++	U	F	K	n	%E+++
liver	RSA	16	-0.0482	1.0490	0.001104	0.440	3.021
			0.1154	0.8782	0.000466	0.965	2.257
kidney	RSA	14	-0.0389	1.0414	0.001160	0.440	3.760
			0.1240	0.8543	0.000472	0.976	3.033
hepatoma	RSA	17	-0.0713	1.0644	0.001109	0.440	2.404
			0.0885	0.8980	0.000506	0.903	1.748
liver	AFP	14	-0.0543	1.0510	0.001344	0.440	2.204
			0.0628	0.9280	0.000736	0.745	1.718
kidney	AFP	12	-0.0643	1.0542	0.001317	0.440	3.370
			0.0925	0.8884	0.000498	1.000	2.128
hepatoma	AFP	17	-0.0663	1.0593	0.001420	0.440	2.932
			0.0742	0.9110	0.000673	0.846	2.422

* sheared to 300-400 nucleotides. The first values are with n fixed, the second with n allowed to vary.
\+ 1000-2200 nucleotides (7). RSA=rat serum albumin. AFP=alpha feto-protein.
++ N = number of points.
+++ %E = root mean square error (RMS) divided by data mean.

is the rate constant which would be observed if the reaction were assayed on HAP, and n was found to be 0.44 (19) or 0.453 (20) for driver and tracer DNAs of the same size. The deviation from second order kinetics indicated by a value of n less than 1 was attributed to the lowered reactability of the single strand regions of partial duplexes relative to free single strands. Smith, Britten and Davidson (20) observed that while no simple physical meaning can be associated with the exponent n, equation (2) is useful for data reduction.

For incomplete reactions, or reactions consisting of multiple components, COTFIT uses the DIGEST function

$$F(t) = U + \sum_{i=1}^{m} Fi(1 + KiCot)^{-n} \qquad (3)$$

where U is the fraction unreacted (single-stranded) at infinite time, m is the number of components, and Fi and Ki are the fraction and rate for component i. Table 1 contains the best-fit parameters of this function for the data of Sala-Trepat et al (6)

Table 2. Comparison of melt fitting methods using data from Wallace et al (8)

Difference	Mean	StdDev	%Mean relative to A
|F(B) - F(A)|	0.0167	0.0285	42.0
|Tm(B) - Tm(A)|	1.1	2.9	1.99
|S(B) - S(A)|	2.1	1.1	33.0
Parameter	**Mean**	**StdDev**	**Maximum**
S(A)	6.2	2.4	11.8
S(B)	4.4	1.6	7.7
%E(A)	5.23	1.32	7.41
%E(B)	1.22	0.41	1.92

Initial absorbances were 0.9-1.3 A260 units at 25°C. Data were collected every 0.4° while heating at 0.25°/min.
A. Parameters for Gaussian curves fit to 15-point cubic-quartic first derivative of absorbance data (using GELFIT).
B. Parameters for equation (1) fit to the absorbance data normalized to fraction hyperchromicity (using COTFIT).

when the exponent n is fixed at 0.44 or allowed to vary. The unfixed exponents vary from 0.745 to 1. This shift toward second-order kinetics may be due either to the different driver and tracer lengths used, or to the interrupted nature of the albumin and AFP genes in rat DNA (T.D. Sargent, J.R. Wu, J. Sala-Trepat, R.B. Wallace, T. Reyes, and J. Bonner, manuscript in preparation). In either case, the results demonstrate the need for careful determination of the exponent for individual S1 assayed experiments, since the calculated K values may vary by greater than 2.5 fold. Values of n significantly different from the expected value may indicate inaccurate tracer and driver length determinations or other systematic error.

Analysis of DNA and chromatin thermal denaturation data

Since many estimates of chromatin and DNA melting component parameters have been made by fitting Gaussian curves to derivatized data, we have compared this method to the use of the COTFIT MELTFN function. Data from our previously published melts of various chromatin and nucleosome samples (8), which had been derivatized with MELSMR and fit with GELFIT, were re-fit using COTFIT. As Table 2 shows, the two methods yield similar results, the Tm's differing by an average of only 2%. However, COTFIT yields an average error almost five times lower than that

Table 3. Rat liver chromatin melting transitions

	Sample 1	Sample 2	Sample 3	Mean	StdDev	%StdDev
I	0.0068	0.0033	0.0047	0.0049	0.0018	36
F1	0.0989	0.1034	0.1076	0.1033	0.0044	4
Tm1	65.5	60.7	56.4	60.9	4.5	7
S1	9.5	7.9	6.2	7.9	1.6	21
F2	0.0789	0.0978	0.0508	0.0758	0.0237	31
Tm2	76.9	76.8	69.1	74.3	4.5	6
S2	5.8	5.2	3.7	4.9	1.1	23
F3	0.0742	0.0811	0.0652	0.0735	0.0080	11
Tm3	77.9	77.4	76.4	77.2	0.8	1
S3	6.0	9.7	8.4	8.0	1.9	24
F4	0.0660	0.0431	0.1188	0.0760	0.0389	51
Tm4	82.4	82.6	79.8	81.6	1.6	2
S4	2.1	1.9	4.2	2.7	1.3	46
%E	1.67	1.09	1.62	1.46	0.30	22

Melts were performed in 0.25 mM EDTA pH 8. Initial absorbances were 0.7-1.5 A260 units at 25°C. Data were collected from 25-95°C every 0.4° while heating at 0.25°/min, and fit to equation (1) using COTFIT. The average percent standard deviations in F, Tm, and S are 24, 4, and 29, respectively.

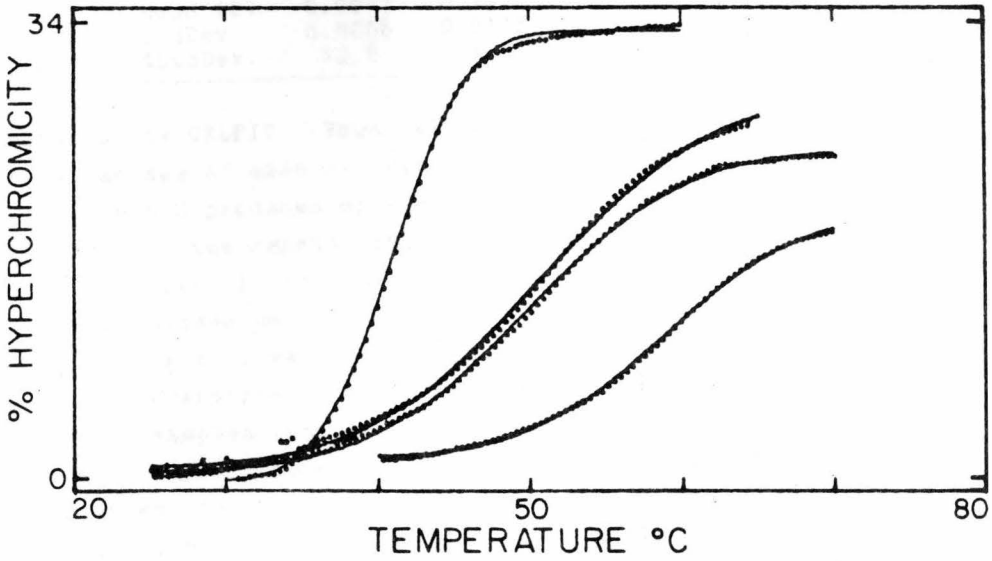


Figure 1. Fitting of DNA melting transitions using MELTFN. From left to right are rat liver DNA and oligomers A (every 4th point shown), C (every other point shown), and B (every other point shown) (see Table 4).

Table 4. DNA melting transitions

	I	F1	Tm1	S1	%E
I. Unsheared Rat liver DNA. 1.1-1.6 A260 units/ml, 0.25 °C/min.					
A. 0.25 mM EDTA pH 8 0.4 °C/pt, 30-60 °C.					
Mean (5)	-0.0011	0.3369	43.6	3.86	1.88
StdDev	0.0040	0.0108	3.9	0.47	0.81
%StdDev	378	3.2	8.9	13.1	43.4
B. 0.01 x SSC pH 7 0.4 °C/pt, 35-75 °C.					
Mean (2)	0.0064	0.3554	58.2	6.09	1.83
StdDev	0.0008	0.0069	0.1	0.28	0.01
%StdDev	12.6	2.0	0.2	4.6	0.4
II. Synthetic oligonucleotides. 1 M NaCl 0.01 M PB, 0.5 °C/min.					
A. CCGAATTCGG / GGCTTAAGCC 0.77-0.78 A260 units/ml, 0.1 °C/pt, 25-65 °C.					
Mean (2)	0.0049	0.2808	50.3	9.28	3.18
StdDev	0.0019	0.0223	0.4	0.49	0.39
%StdDev	38.1	7.9	0.8	5.2	12
B. GGATCACCGCC / CCTAGTGGCGG 0.72-0.77 A260 units/ml, 0.2 °C/pt, 40-70 °C.					
Mean (3)	0.0151	0.1817	58.6	7.26	2.06
StdDev	0.0012	0.0050	0.2	0.40	0.56
%StdDev	7.9	2.7	0.3	5.5	27
C. CATGAATTCATG / GTACTTAAGTAC 0.55-0.57 A260 units/ml, 0.2 °C/pt, 25-70 °C.					
Mean (2)	0.0047	0.2396	49.7	8.34	2.29
StdDev	0.0006	0.0118	0.3	0.32	0.29
%StdDev	12.8	4.9	0.6	3.8	13

produced by GELFIT. Thus, although the MELSMR/GELFIT method has the advantage of ease of visual interpretation of plots, the COTFIT method produces more accurate parameter estimates.

To test the reproducibility of this function, data from melts of rat liver chromatin, rat liver DNA, and synthetic oligonucleotides were fit using COTFIT (Tables 3 and 4 and Figure 1). The rat liver chromatin and DNA melts are of different sample preparations, and the oligonucleotide melts are of different samples from the same preparation. All melts were from separate runs. Some fluctuations in Tm resulting from differences in buffer concentration can be seen for the 0.25 mM EDTA melts, but the standard deviation in Tm is still less than 10%. The deviations for the other melts are much smaller.

The closeness of the data and fit is in accordance with the shape predicted by theoretical treatments of nucleic acid melting

(21-24), although there is no immediate correlation between the parameters of equation (1) and the physical parameters of the system. The agreement of the rat DNA melt data with equation (1) is probably due to the variation in nucleotide composition of rat DNA, and a resulting combination of a Gaussian distribution of small transitions (25). Deviations from the fitted curve are more apparent for the oligonucleotide melts, as might be expected. In light of the ease of estimation of the parameters of equation (1), our results lend support to its use in fitting nucleic acid thermal denaturation data, especially for comparative purposes.

CONCLUSIONS

We have described a set of flexible, interactive programs for the analysis and storage of biochemical data. The ability of the programs to accept English commands and prompt the operator for needed information allows even an inexperienced computer user to analyze a reassociation curve or gel profile in under an hour. In addition to the analysis of data from nucleic acid hybridization, thermal denaturation, and gel electrophoresis (1-10), the programs may be used for a number of other applications, such as resolution of components in velocity and equilibrium density gradients and the determination of rate constants for enzyme reactions. The programs also provide a framework for the development of other data analysis systems.

The programs described in this paper are available from the authors on a variety of machine-readable media.

ACKNOWLEDGMENTS

We thank T.D. Sargent, J.W. Posakony and R.B. Wallace for helpful discussion and critical review of this manuscript. This research was supported by U.S. Public Health Service Grants GM-13762 and GM-00086.

REFERENCES

1. Garrard, W.T., Pearson, W.R., Wake, S.K., and Bonner, J. (1974) Biochem. Biophys. Res. Comm. 58, 50-57.

2. Murphy, R.F., and Bonner, J. (1975) Biochim. Biophys. Acta 405,62-66.

3. Gottesfeld, J.M., Murphy, R.F., and Bonner, J. (1975) Proc. Nat. Acad. Sci. USA 72, 4404-4408.

4. Murphy, R.F., Wallace, R.B., and Bonner, J. (1978) Proc. Nat. Acad. Sci. USA 75, 5903-5907.

5. Wallace, R.B., Dube, S.K., and Bonner, J. (1977) Science 198, 1166-1168.

6. Sala-Trepat, J.M., Sargent, T.D., Sell, S., and Bonner, J. (1979) Proc. Nat. Acad. Sci. USA 76, 695-699.

7. Sala-Trepat, J.M., Dever, J., Sargent, T.D., Thomas, K., Sell, S., and Bonner, J. (1979) Biochemistry, in press.

8. Wallace, R.B., Sargent, T.D., Murphy, R.F., and Bonner, J. (1977) Proc. Nat. Acad. Sci. USA 74, 3244-3248.

9. Bonner, J., Wallace, R.B., Sargent, T.D., Murphy, R.F., and Dube, S.K. (1978) Cold Spring Harbor Symp. Quant. Biol. 42, 851-857.

10. Bakke, A.C., Wu, J.R., and Bonner, J. (1978) Proc. Nat. Acad. Sci. USA 75, 705-709.

11. Pearson, W.R., Davidson, E.H., and Britten, R.J. (1977) Nucleic Acids Res. 4, 1727-1737.

12. Britten, R.J., Graham, D.E., and Neufeld. B.R. (1974) in Methods in Enzymology (L. Grossman and K. Moldave, Eds.), Vol. 29E, pp. 363-418, Academic Press, New York.

13. Marquardt, D.W. (1963) J. Soc. Indust. Appl. Math. 11, 431-441.

14. Gautschi, W. (1972) in Handbook of Mathematical Functions (M. Abramowitz and I.A. Stegun, Eds.), pp. 295-329, Dover Publications, New York.

15. Martin, R.S., and Wilkinson, J.H. (1971) in Linear Algebra (J.H. Wilkinson and C. Reinsch, Eds.), pp. 50-56, Springer-Verlag, New York.

16. Savitzky, A., and Golay, M.J.E. (1964) Anal. Chem. 36, 1627-1639.

17. Steinier, J., Termonia, Y., and Deltour, J. (1972) Anal. Chem. 44, 1906-1909.

18. Madden, H.H. (1978) Anal. Chem. 50, 1383-1386.

19. Morrow, J. (1974) Ph. D. Thesis, Stanford University.

20. Smith, M.J., Britten, R.J., and Davidson, E.H. (1975) Proc. Nat. Acad. Sci. USA 72, 4805-4809.

21. Zimm, B.H. (1960) J. Chem. Phys. 33, 1349-1356.

22. Applequist, J., and Damle, V. (1963) J. Chem. Phys. 39, 2719-2721.

23. Applequist, J., and Damle, V. (1965) J. Am. Chem. Soc. 87, 1450-1458.

24. Abzel, M.Ya. (1973) Phys. Rev. Lett. 31, 589-593.

25. Abzel, M.Ya. (1979) Proc. Nat. Acad. Sci. USA 76, 101-105.

APPENDIX A

GELFIT, a non-linear, least-squares Gaussian curve fitting program

The following program has been used for determination of peak positions and areas for polyacrylamide gel scans (Chapters 1, 2, and 4), thermal denaturation derivative profiles (Chapter 3), and density gradient distributions (Chapter 5). In this listing, input/output related routines, some of which may be computer system specific, have been omitted. Most of these are part of a library of subroutines for handling of standardized data files. The data file format used allows access by all programs in the system to files created by any program.

```
C-----------------------------------------------------------------------
C GELFIT.FOR
C-----------------------------------------------------------------------
      PROGRAM GELFIT
C V04-10        28-JUN-79
C
C R.F.MURPHY    CREATED 23-AUG-76
C DIVISION OF BIOLOGY
C CALIFORNIA INSTITUTE OF TECHNOLOGY
C PASADENA, CA 91125
C (213) 795-6811 EXT.1915
C
C FITS GAUSSIAN COMPONENTS TO VARIABLE FORMAT INTEGER DATA FILES
C USING BAND MATRICES TO MINIMIZE STORAGE REQUIREMENTS.
C
C   USERS OF THIS PROGRAM SHOULD CITE:
C       Murphy, R.F., Pearson, W.R., &Bonner, J. (1979)
C       "Computer programs for analysis of nucleic acid
C       hybridization, thermal denaturation, and gel
C       electrophoresis data"
C       Submitted to Nucleic Acids Research.
C
C MODULES USED
C
C       GELFIT  V04-10  MAIN PROGRAM
C       GELCMD  V02-05  INTERPRETS COMMANDS
C       PLTLIM  V00-01  GET DEFAULT PLOT LIMITS
C       COMNUM  V00-04  GET COMMANDS USING GTLIN (SYSLIB)
C       COMIND  V00-03  USED BY GELCMD
C       IMXMIN  V01-02  DETERMINES MAX & MIN OF INTEGER ARRAY
C       HELP    V04     PRINTS A HELP FILE FROM DEVICE SY:
C       SORTG   V00     SORTS CURVES
C       FTGAUS  V04-04  MAIN FITTING ROUTINE
C       GELFUN  V00     CALCULATES GAUSSIAN FIT VALUE
C       IBANDT  V02-05  FORMS DECOMPOSITION BAND MATRIX
C       BANSOL  V02-05  GENERATES PARAMETER CHANGE VECTORS
C       GELLST  V02     TYPES OUT NEW FIT AND MW'S
C       GELPLT  V01-01  PLOT DATA AND FIT ON TERMINAL
C       TTPLOT  V03-01  USED BY GELPLT
C       GUESSG  V00-04  MAKE STARTING PARAMETER GUESSES
C       GENERD  V01     GENERATE DATA FROM PARAMETERS
C       GAUS    V00     GENERATES GAUSSIAN RANDOM NUMBERS
C
C BIODAS LIBRARY SUBROUTINES
C
C       REMARK  V02-05  PRINTS MESSAGES ON UNIT LWD
C       ERROR   V00     PRINTS FORMATTED ERROR MESSAGES
C       OPEN    V01     CONNECTS I/O DEVICES TO LOGICAL UNITS
C       IRDY0   V01-03  READS A BIODAS FORMAT DATA FILE
C       IFMTYP  V01-01  DETERMINES VARIABLE FORMAT DATA TYPE
C       LRDY1   V01     READS REAL DATA, CONVERTS TO INTEGER
C       IRDP    V01     READS PARAMETERS ONLY (STANDARDS FILE)
C       NTITLE  V01-01  UPDATE FORMATTED TITLE FOR OUTPUT FILE
C
C CONNECTION OF FILES TO LOGICAL UNITS IS HANDLED BY THE
C SUBROUTINE "OPEN", AND IS DONE ONLY IN THIS MODULE
C AND IN "HELP."
C
C THIS PROGRAM WILL RUN "AS IS" (IF OVERLAID) ON ANY PDP-11
C WITH AT LEAST 28K WORDS MEMORY RUNNING UNDER RT-11 V03B.
C (THE RT-11 VERSION OCCUPIES 15.5K WORDS, EXCLUDING BUFFERS)
C TO INCREASE THE NUMBER OF CURVES, CHANGE THE DIMENSIONS
C ON P AND FIXARR, CHANGE THE DATA STATEMENT FOR MAXPAR,
C AND CHANGE THE DIMENSIONS ON PARTDX,PARTDY,DELTA,OLDP AND IMAP
C IN MODULE "FTGAUS" (OR REMOVE THE COMMENT INDICATOR
C FROM THE COMMONS BELOW AND CHANGE THE DIMENSIONS; SEE
C "FTGAUS" FOR EXPLANATION OF EFFECT.)
C TO INCREASE THE NUMBER OF NEIGHBORS, CHANGE THE DIMENSION
C FOR WRKSPC, AND CHANGE MAXBND.
C TO INCREASE THE NUMBER OF POINTS, CHANGE THE DIMENSION
C ON IY, AND CHANGE MAXPTS.
C
C TO SAVE CORE, AN OVERLAID PROGRAM CAN BE GENERATED.
C OVERLAY REGION        MODULES
C ROOT          GELFIT,REMARK,ERROR
C #1            GELCMD
C #1            IRDY0,IFMTYP,IRDP,NTITLE,IWDY0
C #1            FTGAUS,IBANDT,BANSOL
C #1            OPEN,HELP,SORTG,GELLST
C #1            GELPLT,TTPLOT
C #1            GUESSG,GENERD,GAUS
C #2            COMNUM
C #2            PLTLIM
C #2            IMXMIN
C #3            COMIND
C #3            GELFUN
C
      COMMON/V0410/V0410
         INTEGER V0410
      COMMON/PRGNAM/PRGNAM(16)
         BYTE PRGNAM
      COMMON/FILCOM/OUTFIL(16),INPFIL(16),LSTFIL(16),
     *PLTFIL(16),STDFIL(16)
         BYTE OUTFIL,INPFIL,LSTFIL,PLTFIL,STDFIL
      COMMON/LOGDEV/LRD,LWD,LFD
         INTEGER LRD,LWD,LFD
      COMMON/DATA/TITLE(80),MINMAX(4),NOPTS,MAXPTS,IY(600)
         BYTE TITLE
         INTEGER NOPTS,MAXPTS,IY
         REAL MINMAX
      COMMON/CURVES/NOCURV,NOPAR,MAXPAR,P(90)
         INTEGER NOCURV,NOPAR,MAXPAR
         REAL P
      COMMON/FITPAR/NIT,NCIT,RMSQU,DELQU,RATMAX,NEIGHB,NFPAR,
     *  MAXBND,XDEL,NFREE,QCOND,RMS,CHISQ,DELMX
         INTEGER NIT,NCIT,NEIGHB,NFPAR,MAXBND,NFREE,QCOND
         REAL RMSQU,DELQU,RATMAX,XDEL,RMS,CHISQ,DELMX
      COMMON/PARCAL/PARCAL
         LOGICAL PARCAL
      COMMON/FIXCOM/FIXED,UNFIXD,FIXARR(30)
         BYTE FIXED,UNFIXD,FIXARR
      COMMON/RANCOM/IRANV1,IRANV2
         INTEGER IRANV1,IRANV2
      COMMON/PLOTCM/XPMIN,XPMAX,YPMIN,YPMAX,NXPLOT,
     *  NYPLOT,NYPERX,LOGFLG,CRTPLT
         LOGICAL*1 LOGFLG,CRTPLT
         INTEGER NXPLOT,NYPLOT,NYPERX
         REAL XPMIN,XPMAX,YPMIN,YPMAX
      COMMON/STDCOM/NOSTD,MAXSTD,LOGSTD,STD(8)
         INTEGER NOSTD,MAXSTD
```

```
          LOGICAL LOGSTD
          REAL STD
C!      COMMON/PARTDX/PARTDX(90)
C!         REAL PARTDX
C!      COMMON/PARTDY/PARTDY(90)
C!         REAL PARTDY
C!      COMMON/DELTA/DELTA(90)
C!         REAL DELTA
C!      COMMON/OLDP/OLDP(90)
C!         REAL OLDP
C!      COMMON/IMAP/IMAP(90)
C!         INTEGER IMAP
        COMMON/GUESCM/YFILT,ZFILT
           REAL YFILT,ZFILT
        COMMON/VFRMAT/VFRMAT(10)
           INTEGER VFRMAT
        COMMON/DEBUG/DEBUG
           INTEGER DEBUG
C CHARACTER ARRAYS
        BYTE HLPTMP(80)
        EQUIVALENCE (HLPTMP(1),WRKSPC(1))
C NUMERICAL ARRAYS
        REAL WRKSPC(900)
C
        INTEGER COMAND
C
C OPEN RETURNS -1 IF FILE OPENED, OTS ERROR # OTHERWISE.
        INTEGER OPEN
        EXTERNAL GELFUN
C
        DATA PRGNAM     /'G','E','L','F','I','T',
        *' ','V','0','4','-','1','0',' ',',',' '/
        DATA LRD,LWD,LFD        /5,7,2/
        DATA MAXPTS     /600/
        DATA MAXPAR     /90/
        DATA MAXBND     /450/
        DATA MAXSTD     /8/
        DATA FIXED,UNFIXD       /'*',' '/
C
C INITIAL VALUE DATA STATEMENTS
C
        DATA NOPTS, NOPAR, DEBUG        /0,0,0/
        DATA RMSQU,DELQU,RATMAX /0.,0.,0./
        DATA OUTFIL     /'D','K',':','D','A','T','A','.',
        *       'D','A','T',' ',' ',' ',' ',' ',' '/
        DATA LSTFIL     /'L','P',':','D','A','T','A','.',
        *       'L','S','T',' ',' ',' ',' ',' ',' '/
        DATA PLTFIL     /'L','P',':','D','A','T','A','.',
        *       'P','L','T',' ',' ',' ',' ',' ',' '/
C SETTING THESE VARIABLES TO ZERO INITIALZES THE RANDOM
C NUMBER GENERATOR.
        DATA IRANV1,IRANV2      /0,0/
C
C
C PRINT PROGRAM NAME AND VERSION NUMBER.
        CALL REMARK(PRGNAM)
        COMAND = 0
C MAIN GELFIT LOOP START. GET COMMAND.
5       CALL GELCMD(COMAND)
        IF(COMAND.EQ.26) GOTO 260
        IF(COMAND-23) 6,230,240
6       GOTO (10,20,30,40,50,60) COMAND
C
C (01)  HELP
C
10      CALL HELP(HLPTMP)
        GOTO 5
C
C (02)  INPUT
C
20      IF(OPEN(LFD,INPFIL,'DAT','RDO','NC').GE.0) GOTO 5
        CALL IRDYO(LFD,TITLE,NOPTS,NOPAR,MINMAX,
        *VFRMAT,IY,P,MAXPTS,MAXPAR)
        CLOSE(UNIT=LFD)
        NOCURV = NOPAR/3
        IF(NOPTS.GT.1) XDEL = (MINMAX(2)-MINMAX(1))/(NOPTS-1)
        NEIGHB = -1
        RATMAX = 0.
        RMSQU = 0.
        DELQU = 0.0001
        IF(NOPAR.LE.0) GOTO 5
        DO 12 I=1,NOPAR
           FIXARR(I) = UNFIXD
12      CONTINUE
        CALL SORTG(P,NOPAR,0)    ! 0 FOR ORDERING BY XMAX
        GOTO 5
C
C (03)  STANDARDS
C
30      IF(OPEN(LFD,STDFIL,'DAT','RDO','NC').GE.0) GOTO 5
        CALL IRDP(LFD,WRKSPC,NOSTD,STD,MAXSTD)
        CLOSE(UNIT=LFD)
        GOTO 5
C
C (04)  CLOSE
C
40      CALL SORTG(P,NOPAR,1)    ! 1 FOR ORDERING FOR OUTPUT
        CALL IMXMIN(IY,NOPTS,IYMAX,IYMIN)
        MINMAX(3) = FLOAT(IYMIN)
        MINMAX(4) = FLOAT(IYMAX)
        IF(OPEN(LFD,OUTFIL,'DAT','NEW','NC').GE.0) GOTO 5
        CALL NTITLE(TITLE,OUTFIL,PRGNAM,0)
        CALL IWDYO(LFD,TITLE,NOPTS,NOPAR,MINMAX,VFRMAT,IY,P)
        CLOSE(UNIT=LFD)
        GOTO 5
C
C (05)  FIT
C
50      CALL FTGAUS(WRKSPC)
        CALL SORTG(P,NOPAR,0)
C DON'T UPDATE REVISION NUMBER UNLESS SOME ITERATIONS WERE DONE
        I = 4
        IF(NIT.LE.0) I = 0
        CALL NTITLE(TITLE,OUTFIL,PRGNAM,I)
        IF(OPEN(LFD,LSTFIL,'LST','NEW','CC').GE.0) GOTO 5
        CALL GELLST
        CLOSE(UNIT=LFD)
        GOTO 5
C
C (06)  PLOT
```

```
C
60      CALL NTITLE(TITLE,OUTFIL,PRGNAM,0)
        IF(OPEN(LFD,PLTFIL,'PLT','NEW','CC'),GE,0) GOTO 5
        CALL GELPLT(WRKSPC,MAXBND*2)
        CLOSE(UNIT=LFD)
        GOTO 5
C
C (23)  GUESS
C
230     CALL GUESSG(IY,NOPTS,MINMAX(1),XDEL,INT(YFILT),ZFILT,
        *P,NOPAR,MAXPAR)
        DO 232 I=1,NOPAR
           FIXARR(I) = UNFIXD
232     CONTINUE
        NOCURV = NOPAR/3
        GOTO 5
C
C (24)  GENERATE
C
C GELCMD RETURNS THE NUMBER OF POINTS TO GENERATE IN 'NOPTS',
C THE START AND STOP X VALUES IN 'MINMAX(1)' AND 'MINMAX(2)',
C AND THE RELATIVE STANDARD DEVIATION FOR ALL UNFIXED PARAMETERS
C IN 'MINMAX(3)' (IF ZERO, THE EXACT FIT IS GENERATED).
C
C PUT DESIRED STANDARD DEVIATIONS INTO WORKING ARRAY.
240     DO 242 I=1,NOPAR
           WRKSPC(I) = 0.
           IF(FIXARR(I).NE.FIXED) WRKSPC(I) = MINMAX(3)*P(I)
242     CONTINUE
C THE .FALSE. INDICATES NO 'LOGX' GENERATION.
C THE TWO .TRUE.'S INDICATE INTEGER Y ONLY DATA GENERATION.
C WRKSPC(NOPAR+1) IS USED FOR STORAGE OF OLD PARAMETERS
C WHEN NEW PARAMETERS ARE GENERATED FOR EACH DATA POINT.
        CALL GENERD(MINMAX(1),MINMAX(2),,FALSE.,NOPAR,P,WRKSPC,
        *WRKSPC(NOPAR+1),GELFUN,NOPTS,,,,IY,,TRUE.,,TRUE.)
        XDEL = (MINMAX(2)-MINMAX(1))/(NOPTS-1)
        GOTO 5
C
C (26)  EXIT
C
260     CALL EXIT
        END

C-------------------------------------------------------------
C GELCMD.FOR
C-------------------------------------------------------------
        SUBROUTINE GELCMD(COMAND)
C V02-05        28-JUN-79
C
C R.F.MURPHY    CREATED 15-MAY-78
C
C GETS AND INTERPRETS COTFIT COMMAND STRINGS,
C RETURNING AN INTEGER COMMAND NUMBER FOR
C THOSE COMMANDS WHICH NEED PROCESSING BY THE
C MAIN PROGRAM.
C
C COMMAND OUTLINE
C
C               GELFIT                  COTFIT
C
```

```
C # COMMAND     OPTIONS ARGS            OPTIONS ARGS
C
C 01 Help       Long                    Long
C 02 Input      Add     FILE            Add,Old FILE
C 03 Standards  NOLOG   FILE                  R
C 04 Close                              LIst,PLot
C 05 FIT                2I                    2I
C 06 PLot       Crt     4R,3I           Crt,LP,LOG,NOLOG 4R,3I
C 07 Output             FILE            LIst,PLot FILE
C 08 New
C 09 Add                3R                    13R
C 10 Delete             2I                    2I
C 11 Modify             2I                    2I
C 12 Type               2I                    2I
C 13 FIX                2I                    2I
C 14 Unfix              2I                    2I
C 15 POints
C 16 Kill               2I                    2I
C 17 List               2I                    2I
C 18 Quit               2R                    2R
C 19 Bands              I               ----------
C 20 Width              R               ----------
C 21 Xscale             2R                    2R
C 22 Yscale             2R                    2R
C 23 GUess              2R              ----------
C 24 GEnerate           I,3R            LOG,NOLOG I,3R
C 25 DEBug              I                     I
c 26 Exit
C
C INPUTS (MARKED >) AND OUTPUTS (MARKED <)
C
C  COMAND          <ARG>       >LAST COMMAND <NEW COMMAND #
C                              >0:OK; =0:BAD COM; =-1:AMB COM
C                              =-2:BAD SWITCH; =-3:AMB SWITCH
C  LRD             /LOGDEV/    >LOGICAL READ DEVICE
C  LWD             /LOGDEV/    >LOGICAL WRITE DEVICE
C  OUTFIL(16)      /FILCOM/    ><OUTPUT FILE NAME
C  INPFIL(16)      /FILCOM/    ><INPUT FILE NAME
C  LSTFIL(16)      /FILCOM/    ><LISTING FILE NAME
C  PLTFIL(16)      /FILCOM/    ><PLOT FILE NAME
C  STDFIL(16)      /FILCOM/    ><STANDARDS FILE NAME
C  TITLE(80)       /DATA/      <NEW TITLE AFTER POINTS
C  XMIN            /DATA/      >
C  XMAX            /DATA/      >
C  YMIN            /DATA/      >
C  YMAX            /DATA/      >
C  NOPTS           /DATA/      ><# OF DATA POINTS
C  MAXPTS          /DATA/      >MAX VAL FOR NOPTS
C  IY(MAXPTS)      /DATA/      ><Y DATA VALUES
C  VFRMAT          /VFRMAT/    <OUTPUT DATA FILE FORMAT
C  NOCURV          /CURVES/    ><# OF FIT CURVES (NOPAR/3)
C  NOPAR           /CURVES/    ><# OF FIT PARAMETERS
C  MAXPAR          /CURVES/    >MAX VAL FOR NOPAR
C  P(MAXPAR)       /CURVES/    ><PARAMETERS
C  NIT             /FITPAR/    ><# ITS FOR REFINE
C  NCIT            /FITPAR/    ><# CITS FOR REFINE
C  RMSQU           /FITPAR/    ><RMS QUIT VALUE
C  DELQU           /FITPAR/    ><DELMX QUIT VALUE
C  RATMAX          /FITPAR/    ><MAX RATIO OF HEIGHT TO HWIDTH
C  NEIGHB          /FITPAR/    ><# OF NEIGHBORING BANDS IN FIT
```

```
C  NFPAR          /FITPAR/  ><NUMBER OF FREE PARAMETERS
C  MAXBND         /FITPAR/  ><MAX PROD OF NFPAR AND NOPAR
C  XDEL           /FITPAR/  ><X CHANGE BETWEEN POINTS
C  FIXED          /FIXCOM/  >ASCII VALUE TO SHOW PARAM IS FIXED
C  UNFIXD         /FIXCOM/  >' ' ' ' ' ' NOT FIXED
C  FIXARR(MAXPAR)/FIXCOM/   ><FLAGS FOR FIXING PARAMS
C  XPMIN,XPMAX    /PLOTCM/  <
C  YPMIN,YPMAX    /PLOTCM/  <
C  NXPLOT         /PLOTCM/  <# CHARS PER LINE IN PLOT
C  NYPLOT         /PLOTCM/  <# LINES PER PLOT
C  NYPERX         /PLOTCM/  <# Y DOTS PER X (FOR FIT PLOT)
C  LOGFLG         /PLOTCM/  <.TRUE. MEANS PLOT Y VS. LOG(X)
C  CRTPLT         /PLOTCM/  <.TRUE. MEANS CONDENSED PLOT
C  NOSTD          /STDCOM/  <ORDER OF STANDARDS FIT
C  LOGSTD         /STDCOM/  <FIT IS TO LOG MW
C  DEBUG          /DEBUG/   ><DEBUGGING FLAG
C
      INTEGER COMAND
      COMMON/LOGDEV/LRD,LWD,LFD
         INTEGER LRD,LWD,LFD
      COMMON/FILCOM/OUTFIL(16),INFILE(16),LSTFIL(16),
     *   PLTFIL(16),STDFIL(16)
         BYTE OUTFIL,INFILE,LSTFIL,PLTFIL,STDFIL
      COMMON/DATA/TITLE(80),XMIN,XMAX,YMIN,YMAX,
     *   NOPTS,MAXPTS,IY(100)
         BYTE TITLE
         INTEGER NOPTS,MAXPTS,IY
         REAL XMIN,XMAX,YMIN,YMAX
      COMMON/VFRMAT/VFRMAT(10)
         INTEGER VFRMAT
      COMMON/CURVES/NOCURV,NOPAR,MAXPAR,P(60)
         INTEGER NOCURV,NOPAR,MAXPAR
         REAL P
      COMMON/FITPAR/NIT,NCIT,RMSQU,DELQU,RATMAX,NEIGHB,
     *  NFPAR,MAXBND,XDEL
         INTEGER NIT,NCIT,NEIGHB,NFPAR,MAXBND
         REAL RMSQU,DELQU,RATMAX,XDEL
      COMMON/FIXCOM/FIXED,UNFIXD,FIXARR(1)
         BYTE FIXED,UNFIXD,FIXARR
      COMMON/PLOTCM/XPMIN,XPMAX,YPMIN,YPMAX,NXPLOT,
     *  NYPLOT,NYPERX,LOGFLG,CRTPLT
         LOGICAL*1 LOGFLG,CRTPLT
         INTEGER NXPLOT,NYPLOT,NYPERX
         REAL XPMIN,XPMAX,YPMIN,YPMAX
      COMMON/STDCOM/NOSTD,MAXSTD,LOGSTD
         INTEGER NOSTD,MAXSTD
         LOGICAL LOGSTD
      COMMON/GUESCM/YFILT,ZFILT
         REAL YFILT,ZFILT
      COMMON/DEBUG/DEBUG
         INTEGER DEBUG
C WHEN 'ADDFLG' IS .TRUE., DATA IS BEING ADDED TO PREVIOUS VALUES
      LOGICAL*1 ADDFLG
C WHEN 'LPTPLT' IS .TRUE., LINE-PRINTER DEFAULTS ARE SELECTED.
      LOGICAL*1 LPTPLT
C THIS ARRAY CONTAINS THE DEFAULT FILE FORMAT.
      INTEGER IFRMAT(10),MAXIFR
C THESE ARRAYS AND VARIABLES CONTAIN DATA FOR THE COMNUM SUBROUTI
      BYTE COMSTR(128),SWISTR(31),CSTRNG(81)
      INTEGER COMPTR(2,26),SWIPTR(2,9),SWIFLG(3,11)
      INTEGER MAXCOM,MAXSWI,MAXFLG,LSTRNG,IPTR,LARGS
C 'ARGTYP' <0 IF COMMAND TAKES INTEGER ARGS, >0 IF REAL ARGS.
C 'REQDAT' .TRUE. IF COMMAND REQUIRES PREVIOUS DATA POINTS.
C 'REQPAR' .TRUE. IF COMMAND REQUIRES PREVIOUS PARAMS.
      BYTE ARGTYP(26)
      LOGICAL*1 REQDAT(26),REQPAR(26)
C
      INTEGER I,J,K,IARG1,IARG2
      REAL RARG1, RARG2
      REAL OMN,OMX,XT
C PROCEDURE TYPE DECLARATION.
      INTEGER COMNUM
C
      DATA MAXCOM     /26/
      DATA MAXSWI     /9/
      DATA MAXFLG     /11/
      DATA COMSTR/'H','E','L','P','I','N','P','U','T','S',
     *'T','A','N','D','A','R','D','S','C','L',
     *'O','S','E','F','I','T','P','L','O','T',
     *'O','U','T','P','U','T','N','E','W','A',
     *'D','D','D','E','L','E','T','E','M','O',
     *'D','I','F','Y','T','Y','P','E','F','I',
     *'X','U','N','F','I','X','P','O','I','N',
     *'T','S','K','I','L','L','L','I','S','T',
     *'Q','U','I','T','B','A','N','D','S','W',
     *'I','D','T','H','X','S','C','A','L','E',
     *'Y','S','C','A','L','E','G','U','E','S',
     *'S','G','E','N','E','R','A','T','E','D',
     *'E','B','U','G','E','X','I','T'/
      DATA COMPTR/4,1,9,5,18,10,23,19,26,26,
     *   30,28,36,31,39,37,42,40,48,43,
     *   54,49,58,55,61,61,66,62,72,68,
     *   76,73,80,77,84,81,89,85,94,90,
     *   100,95,106,101,111,108,119,113,124,122,
     *   128,125/
      DATA SWISTR/'O','L','D','A','D','D','L','O','N','G',
     *'C','R','T','L','O','G','N','O','L','O',
     *'G','L','P','L','I','S','T','P','L','O','T'/
      DATA SWIPTR/3,1,6,4,10,10,13,11,16,16,
     *   21,21,23,23,27,25,31,29/
C SWITCHES WHICH ARE ONLY USED BY COTFIT
C HAVE THE COMMAND # SET TO 0.
      DATA SWIFLG     /1,0,0,
     *   2,2,0,
     *   3,1,0,
     *   4,6,0,
     *   5,0,0,
     *   6,3,0,
     *   5,0,0,
     *   6,0,0,
     *   7,6,0,
     *   8,7,0,
     *   9,7,0/
      DATA ARGTYP     /0,0,0,0,-1,0,0,0,0,-1,
     *   -1,-1,-1,-1,0,-1,-1,1,-1,0,
     *   1,1,1,0,-1,0/
      DATA REQDAT     /0,0,0,1,1,1,0,0,0,0,
     *   0,0,0,0,0,1,1,0,0,0,
     *   1,1,1,0,0,0/
      DATA REQPAR     /0,0,0,1,1,0,0,0,1,1,
```

```
     *   1,1,1,1,0,0,0,0,0,0,
     *   0,0,0,1,0,0/
      DATA IFRMAT     /'(1','0I','8)',',  ',' ',
     *'  ',',  ',',  ',',  ',',  '/
      DATA MAXIFR     /10/
C
C IF LAST COMAND WAS Input DO AN IMPLICIT New
      IF(COMAND.EQ.2.AND.NOCURV.LE.0.AND.NOPTS.GT.0) GOTO 80
C RETURN HERE TO GET NEW COMMAND.
C CONVERT COMMAND AND ASSOCIATED SWITCHES INTO COMMAND NUMBER
C AND FLAGS IN SWIFLG.
1     CALL RCTRLO
      LSTRNG = 0
      COMAND = COMNUM(CSTRNG,LSTRNG,COMSTR,COMPTR,MAXCOM,
     *SWISTR,SWIPTR,MAXSWI,SWIFLG,MAXFLG,IPTR)
C CHECK FOR BAD COMMAND OR BAD OPTION.
      IF(COMAND.GT.0) GOTO 4
      GOTO (530,520,510,500) COMAND+4
C IPTR --> START OF ARGUMENTS.
C MAKE LARGS = LENGTH OF ARGUMENTS STRING.
4     LARGS = LSTRNG - IPTR + 1
C GET INTEGER ARGS IF NEEDED.
      IF(ARGTYP(COMAND)) 5,9,7
5     IF(LARGS.LE.0) GOTO 6
      DECODE(LARGS,2000,CSTRNG(IPTR),ERR=500) IARG1,IARG2
      GOTO 9
6     IARG1 = 0
      IARG2 = 0
      GOTO 9
C GET REAL ARGS IF NEEDED.
7     IF(LARGS.LE.0) GOTO 8
      DECODE(LARGS,2010,CSTRNG(IPTR),ERR=500) RARG1,RARG2
      GOTO 9
8     RARG1 = 0.
      RARG2 = 0.
C MAKE SURE REQUIRED DATA AND CURVES EXIST.
9     IF(REQDAT(COMAND).AND.NOPTS.LE.0) GOTO 540
      IF(REQPAR(COMAND).AND.NOPAR.LE.0) GOTO 550
      GOTO (10,20,30,40,50,60,70,80,90,100,100,100,130,130,
     *150,160,160,180,190,200,210,210,230,240,250,260) COMAND
C
C (01)  H(ELP) (/LONG) - PRINT COMMAND SUMMARY
C
10    IF(SWIFLG(3,3).NE.0) RETURN
      WRITE(LWD,1040)
      GOTO 1
C
C (02)  I(NPUT)(/ADD) (INFILE.EXT) - GET INPUT FILE
C
C CHECK FOR /ADD.
20    ADDFLG = (SWIFLG(3,2).NE.0)
C SEE IF A FILENAME WAS SPECIFIED.
      IF(LARGS.GT.0) GOTO 21
C NO ARG SPECIFIED, PROMPT.
      WRITE(LWD,1010)
      READ(LRD,2020) INFILE
      GOTO 22
C USE ARGUMENT SPECIFIED.
21    DECODE(LARGS,2020,CSTRNG(IPTR)) INFILE
C DEFAULT OUTPUT FILE IS THE INPUT FILE.
C DON'T CHANGE OUTPUT FILE IF ADDING.
22    IF(ADDFLG) GOTO 24
      DO 23 I=1,16
         OUTFIL(I) = INFILE(I)
23    CONTINUE
C GET DEFAULTS FOR LIST AND PLOT FILES
      CALL SETFIL
C NOT ADDING, DELETE OLD POINTS.
      NOPTS = 0
C RETURN TO MAIN PROGRAM TO READ INPUT FILE.
24    RETURN
C
C (03)  S(TANDARDS)(/NOLOG) (STDFIL.EXT) - GET STANDARDS FILE
C
30    IF(LARGS.GT.0) GOTO 31
      WRITE(LWD,1050)
      READ(LRD,2020) STDFIL
      IF(STDFIL(1).NE.' ') GOTO 32
C NO FILE SPEC. CANCEL PREVIOUS STANDARDS.
      NOSTD = 0
      GOTO 1
31    DECODE(LARGS,2020,CSTRNG(IPTR)) STDFIL
C FITS ARE TO LOG MW UNLESS /NOLOG IS SPECIFIED.
32    LOGSTD = (SWIFLG(3,6).EQ.0)
      RETURN
C
C (04)  C(LOSE) - RETURN TO WRITE OUTPUT FILE
C
40    RETURN
C
C (05)  FIT (NIT)(,NCIT) - RETURN FOR FITTING
C
50    NIT = IARG1
C LET FIT IMPLY FIT 10.
      IF(LARGS.LE.0) NIT = 10
      NCIT = IARG2
C LET FIT? IMPLY FIT?,30
      IF(NCIT.LE.0) NCIT = 30
      IF(NEIGHB.LT.0) NEIGHB=MINO(NOCURV-1,(MAXBND/NOPAR)/3-1)
      NFPAR = (NEIGHB+1)*3
      IF(NFPAR.GT.NOPAR) NFPAR = NOPAR
      IF((NFPAR*NOPAR).LE.MAXBND) RETURN
      WRITE(LWD,1060) NOPAR,NFPAR,MAXBND
      GOTO 1
C
C (06)  PL(OT) (/CRT) - PLOT POINTS & FIT ON TERMINAL
C
60    CALL IMXMIN(IY,NOPTS,IYMAX,IYMIN)
      CRTPLT = (SWIFLG(3,4).NE.0)
      LPTPLT = (SWIFLG(3,9).NE.0)
      IF(LARGS.LE.0) GOTO 62
      DECODE(LARGS,2040,CSTRNG(IPTR),ERR=500) XPMIN,XPMAX,
     *YPMIN,YPMAX,NXPLOT,NYPLOT,NYPERX
      CRTPLT = (.NOT.LPTPLT)
      GOTO 66
62    CALL PLTLIM(XMIN,XMAX,XPMIN,XPMAX)
      YMIN = FLOAT(IYMIN)
      YMAX = FLOAT(IYMAX)
      CALL PLTLIM(YMIN,YMAX,YPMIN,YPMAX)
      NYPERX = 5
```

```
      IF(.NOT.LPTPLT) GOTO 64
C USE LINE-PRINTER DEFAULTS
      NXPLOT = 101
      NYPLOT = 51
      GOTO 66
64    IF(.NOT.CRTPLT) GOTO 65
C USE CRT DEFAULTS AND USE CONDENSED PLOTTING
      NXPLOT = 68
      NYPLOT = 16
      GOTO 66
C USE TELETYPE DEFAULTS AND SET CONDENSED PLOTTING FLAG
65    NXPLOT = 61
      NYPLOT = 41
      CRTPLT = .TRUE.
66    RETURN
C
C (07)  O(UTPUT)(/LIST)(/PLOT)(OUTFIL.EXT)-SET OUTPUT FILENAMES
C
70    IF(SWIFLG(3,10).NE.0) GOTO 72
      IF(SWIFLG(3,11).NE.0) GOTO 74
      IF(LARGS.LE.0) GOTO 71
      DECODE(LARGS,2020,CSTRNG(IPTR)) OUTFIL
      GOTO 1
71    WRITE(LWD,1030)
      READ(LRD,2020) OUTFIL
C SETTING OUTPUT FILENAME ALSO SETS PLOT AND LIST FILENAMES.
      CALL SETFIL
      GOTO 1
72    IF(LARGS.LE.0) GOTO 73
      DECODE(LARGS,2020,CSTRNG(IPTR)) LSTFIL
      GOTO 1
73    WRITE(LWD,1240)
      READ(LRD,2020) LSTFIL
      GOTO 1
74    IF(LARGS.LE.0) GOTO 75
      DECODE(LARGS,2020,CSTRNG(IPTR)) PLTFIL
      GOTO 1
75    WRITE(LWD,1250)
      READ(LRD,2020) PLTFIL
      GOTO 1
C
C (08)  N(EW) - GET NEW CURVES
C
80    NOPAR = 0
C MAKE SURE ADD COMMAND READS FROM TELETYPE
      LARGS = 0
C
C (09)  A(DD) - ADD TO PREVIOUS PARAMETERS
C
90    IF(NOPAR.GE.MAXPAR) GOTO 560
      IF(LARGS.GT.0) GOTO 93
      J = NOPAR
91    WRITE(LWD,1070)
92    READ(LRD,2010,ERR=91,END=94) (P(I),I=J+1,J+3)
      J = J + 3
      IF(J.LT.MAXPAR) GOTO 92
      GOTO 94
93    DECODE(LARGS,2010,CSTRNG(IPTR),ERR=500)
      * (P(I),I=NOPAR+1,NOPAR+3)
      J = NOPAR + 3
94    DO 95 I=NOPAR+1,J
        FIXARR(I) = UNFIXD
95    CONTINUE
      NOPAR = J
      NOCURV = J/3
      NEIGHB = -1
      GOTO 1
C
C     MAKE SURE DELETE, MODIFY AND TYPE
C     ARGUMENTS ARE WITHIN LIMITS
C
C LET 'COMAND N,<CR>' IMPLY 'COMAND N,NOCURV<CR>'
100   IF(CSTRNG(LSTRNG).NE.',') GOTO 101
      IARG2 = NOCURV
      GOTO 102
C LET 'COMAND N<CR>' IMPLY 'COMAND N,N<CR>'
101   IF(IARG2.EQ.0) IARG2 = IARG1
C LET 'COMAND<CR>' IMPLY 'COMAND 1,NOCURV<CR>'
      IF(LARGS.GT.0) GOTO 102
      IARG1 = 1
      IARG2 = NOCURV
102   IF(IARG1.LE.0.OR.IARG1.GT.NOCURV.OR.IARG2.LT.IARG1
      *.OR.IARG2.GT.NOCURV) GOTO 600
C CONVERT CURVE NUMBERS TO PARAMETER NUMBERS
      IARG1 = IARG1*3 - 2
      IARG2 = IARG2*3
      IF(COMAND.NE.10) GOTO 110
C
C (10)  D(ELETE) (I)(,J) - DELETE CURVES
C
      I = IARG2 - IARG1 + 1
      NOPAR = NOPAR - I
      NEIGHB = -1
103   IF(IARG1.GT.NOPAR) GOTO 104
        P(IARG1) = P(IARG1+I)
        FIXARR(IARG1) = FIXARR(IARG1+I)
        IARG1 = IARG1 + 1
        GOTO 103
104   NOCURV = NOPAR/3
      GOTO 1
C
C (11)  M(ODIFY) (I)(,J) - CHANGE CURVES
C (12)  T(YPE) (I)(,J) - TYPE CURVES
C
110   WRITE(LWD,1080)
111     WRITE(LWD,1090) IARG1/3+1,
      *   (P(I),FIXARR(I),I=IARG1,IARG1+2)
C IF TYPING, DON'T GET NEW VALUES.
        IF(COMAND.EQ.12) GOTO 113
112     WRITE(LWD,1100)
        READ(LRD,2010,ERR=114) (P(I),I=IARG1,IARG1+2)
113     IARG1 = IARG1 + 3
        IF(IARG1.GT.IARG2) GOTO 1
        GOTO 111
114   CALL ERROR('W','input error - repeat last curve.')
      GOTO 112
C
C (13)  FIX (I)(,J) - HOLD CURVE I PARAMETER J CONSTANT
C (14)  U(NFIX) (I)(,J) - ALLOW CURVE I PARAMETER J TO VARY
C
```

```
130     IF(IARG1.LE.0.OR.IARG1.GT.NOCURV) GOTO 570
        IF(IARG2.LT.0.OR.IARG2.GT.3) GOTO 600
        IARG1 = (IARG1-1)*3 + IARG2
        IF(IARG2.EQ.0) GOTO 131
        IARG2 = IARG1
        GOTO 132
C 2ND ARG=0 - FIX ENTIRE CURVE.
131     IARG1 = IARG1 + 1
        IARG2 = IARG1 + 2
132     K = FIXED
        IF(COMAND.EQ.14) K = UNFIXD
        DO 133 I =IARG1,IARG2
           FIXARR(I) = K
133     CONTINUE
        GOTO 1
C
C (15)  PO(INTS) - ADD TO PREVIOUS DATA POINTS
C
C IF NO DATA YET, DO A TT: INPUT COMMAND
C MAKE SURE THERE'S ROOM FOR MORE POINTS.
150     IF(NOPTS.GE.MAXPTS) GOTO 580
        IF(NOPTS.GT.0) GOTO 154
C IF NO POINTS ALREADY, GET TITLE, ETC.
        NOPTS = 0
        NOPAR = 0
        DO 152 I=1,43
           TITLE(I) = ' '
152     CONTINUE
        WRITE(LWD,1020)
        READ(LRD,2020) (TITLE(I),I=44,80)
        DO 153 I=1,MAXIFR
           VFRMAT(I) = IFRMAT(I)
153     CONTINUE
        WRITE(LWD,1110)
        READ(LRD,2010) XMIN,XDEL
        IF(XDEL.EQ.0.) XDEL = 1.
154     WRITE(LWD,1120)
155        READ(LRD,2000,ERR=156,END=158) IY(NOPTS+1)
           NOPTS = NOPTS + 1
           IF(NOPTS.GE.MAXPTS) GOTO 580
           GOTO 155
156        CALL ERROR('W','input error - repeat last point.')
           GOTO 155
158     IF(NOPTS.LE.0) GOTO 540
        XMAX = XMIN + (NOPTS-1)*XDEL
C DO AN IMPLICIT NEW
        IF(NOPAR.LE.0) GOTO 80
        GOTO 1
C
C       MAKE SURE KILL & LIST ARGUMENTS ARE WITHIN LIMITS
C
C LET 'COMAND N,<CR>' IMPLY 'COMAND N,NOPTS<CR>'
160     IF(CSTRNG(LSTRNG).NE.',') GOTO 161
        IARG2 = NOPTS
        GOTO 162
C LET 'COMAND N<CR>' IMPLY 'COMAND N,N<CR>'
161     IF(IARG2.EQ.0) IARG2 = IARG1
C LET 'COMAND<CR>' IMPLY 'COMAND 1,NOPTS<CR>'
        IF(LARGS.GT.0) GOTO 162
        IARG1 = 1
        IARG2 = NOPTS
162     IF(IARG1.LE.0.OR.IARG1.GT.NOPTS.OR.IARG2.LT.IARG1
        *.OR.IARG2.GT.NOPTS) GOTO 590
        IF(COMAND.NE.16) GOTO 170
C
C (16)  K(ILL) (M)(,N) - DELETE POINTS
C
        I = IARG2 - IARG1 + 1
        NOPTS = NOPTS - I
        XMAX = XMIN + (NOPTS-1)*XDEL
163        IF(IARG1.GT.NOPTS) GOTO 1
           IY(IARG1) = IY(IARG1+I)
           IARG1 = IARG1 + 1
           GOTO 163
C
C (17)  L(IST) (M)(,N) - LIST POINTS
C
170     WRITE(LWD,1130) IARG1,IARG2
        WRITE(LWD,VFRMAT) (IY(I),I=IARG1,IARG2)
        GOTO 1
C
C (18)  Q(UIT) (RMSQU)(,DELQU) - SET QUIT CRITERIA
C
180     IF(LARGS.GT.0) GOTO 181
        WRITE(LWD,1190)
        READ(LRD,2010) RMSQU
        WRITE(LWD,1200)
        READ(LRD,2010) DELQU
        GOTO 1
181     RMSQU = RARG1
        DELQU = RARG2
        GOTO 1
C
C (19)  B(ANDS) (I) - SET # OF NEIGHBORING BANDS IN FIT
C
190     IF(IARG1.LT.0.OR.IARG1.GE.NOCURV) GOTO 570
        NEIGHB = IARG1
        GOTO 1
C
C (20)  W(IDTH) (A) - SET MAX RELATIVE WIDTH
C
200     IF(LARGS.GT.0) GOTO 201
        WRITE(LWD,1140)
        READ(LRD,2010) RATMAX
        GOTO 1
201     RATMAX = RARG1
        GOTO 1
C
C GET FACTORS FOR XSCALE AND YSCALE.
C
210     IF(LARGS.GT.0) GOTO 211
        WRITE(LWD,1230)
        READ(LRD,2010) RARG1,RARG2
211     IF(COMAND.EQ.22) GOTO 220
C
C (21)  X(SCALE) - SCALE X DATA
C
        XMIN = (XMIN+RARG1)*RARG2
        XMAX = (XMAX+RARG1)*RARG2
        XDEL = XDEL*RARG2
```

```
        DO 212 I=1,NOPAR,3
           P(I) = (P(I)+RARG1)*RARG2
           P(I+2) = P(I+2)*RARG2
212     CONTINUE
        RATMAX = RATMAX*RARG2
        GOTO 1
C
C (22)  Y(SCALE) A,B - SCALE Y DATA
C
220     YMIN = (YMIN+RARG1)*RARG2
        YMAX = (YMAX+RARG1)*RARG2
        DO 221 I=2,NOPAR,3
           P(I) = (P(I)+RARG1)*RARG2
221     CONTINUE
        DO 222 I=1,NOPTS
           IY(I) = INT( (FLOAT(IY(I))+RARG1)*RARG2 )
222     CONTINUE
        RATMAX = RATMAX/RARG2
        GOTO 1
C
C (23)  GU(ESS) YFILTER,ZFILTER - GENERATE PARAMETER GUESSES
C
230     IF(LARGS.LE.0) GOTO 232
        YFILT = RARG1
        ZFILT = RARG2
        GOTO 234
C NO VALUES SPECIFIED. USE DEFAULTS.
232     CALL IMXMIN(IY,NOPTS,IYMAX,IYMIN)
        YFILT = 0.05*IYMAX
        ZFILT = 0.001*(IYMAX-IYMIN)/XDEL
234     RETURN
C
C (24)  GE(NERATE) NOPTS,XMIN,XMAX,E - GENERATE DATA FROM PARAMS
C
240     IF(LARGS.LE.0) GOTO 241
        DECODE(LARGS,2050,CSTRNG(IPTR),ERR=500)
        * NOPTS,XMIN,XMAX,YMIN
        IF(NOPTS.LE.1) NOPTS = MAXPTS
        GOTO 242
241     NOPTS = MAXPTS
        XMIN = 0.
        XMAX = 1.
        YMIN = 0.
C MAKE SURE A VALID OUTPUT FORMAT EXISTS.
242     DO 243 I=1,MAXIFR
           VFRMAT(I) = IFRMAT(I)
243     CONTINUE
C COPY COMMAND INTO TITLE.
        DO 244 I=44,80
           TITLE(I) = ' '
244     CONTINUE
        DO 245 I=1,MIN0(LSTRNG,33)
           TITLE(I+43) = CSTRNG(I)
245     CONTINUE
        RETURN
C
C (25)  DEB(UG) I - SET DEBUGGING FLAG
C
250     DEBUG = IARG1
        GOTO 1
C
C (26)  E(XIT) - RETURN TO MAIN PROGRAM FOR EXIT
C
260     RETURN
        GOTO 1
C
C       ERROR MESSAGES
C
500     CALL ERROR('F','illegal command - type H<cr> for help.')
        GOTO 1
510     CALL ERROR('F','ambiguous command.')
        GOTO 1
520     CALL ERROR('F','illegal option.')
        GOTO 1
530     CALL ERROR('F','ambiguous option.')
        GOTO 1
540     CALL ERROR('F','no points.')
        GOTO 1
550     CALL ERROR('F','no curves.')
        GOTO 1
560     CALL ERROR('F','no more curves.')
        GOTO 1
570     CALL ERROR('F','bad curve number.')
        GOTO 1
580     CALL ERROR('F','no more points.')
        GOTO 1
590     CALL ERROR('F','bad point number.')
        GOTO 1
600     CALL ERROR('F','bad parameter number.')
        GOTO 1
C
C       OUTPUT FORMATS
C
1010    FORMAT('$INPUT FILE: ')
1020    FORMAT('$TITLE: ')
1030    FORMAT('$OUTPUT FILE: ')
1040    FORMAT(' GELFIT Command Summary:'/
        *'0The minimum required letters of each command or
        * option'/' are shown. All commands are terminated
        * by the return key.'/' <list> indicates a set of
        * optional arguments.'/
        *'0A <list>',T15,'ADD more curves'/
        *' B k',T15,'fit k BANDS wide'/
        *' C',T15,'CLOSE output file'/
        *' D i,J',T15,'DELETE curves i-J'/
        *' FIT k,l',T15,'FIT k its with l cits'/
        *' FIX i,J',T15,'FIX curve i param J'/
        *' GE i,x1,x2,e',T15,'GENERATE data from params'/
        *' GU yfil,zfil',T15,'GUESS curve parameters'/
        *' H',T15,'type this HELP summary'/
        *'   /L',T15,'type SY:GELFIT.HLP'/
        *' I infile',T15,'set INPUT file'/
        *'   /A',T15,'ADD to previous data'/
        *' K m,n',T15,'KILL points m-n'/
        *' L m,n',T15,'LIST points m-n'/
        *' M i,J',T15,'MODIFY curves i-J'/
        *' N',T15,'set NEW curves'/
        *' O outfil',T15,'set OUTPUT filename'/
        *'   /LI',T20,'set LIST filename'/
        *'   /PL',T20,'set PLOT filename'/
```

```
     *' PL <list>',T15,'PLOT data and fit'/
     *'   /C',T15,'plot in CRT format'/
     *'   /LP',T15,'plot in LP format'/
     *' PO',T15,'add more POINTS'/
     *' Q a,b',T15,'set QUIT criteria'/
     *' S stdfil',T15,'get STANDARDS'/
     *' T i,j',T15,'TYPE curves i-j'/
     *' U i,j',T15,'UNFIX curve i param j'/
     *' W a',T15,'set max relative WIDTH'/
     *' X a,b',T15,'XSCALE (x+a)*b'/
     *' Y a,b',T15,'YSCALE (y+a)*b'/)
1050 FORMAT('$STANDARDS FILE: ')
1060 FORMAT('ORND MAT TOO LRG'/' #P=',I6,' #N=',I6,
     *' MAXBND=',I6/)
1070 FORMAT('$ENTER CURVES - X,Y,H<CR> (^Z=END): ')
1080 FORMAT(' ...N.....X.....,.....Y.....,H.....')
1090 FORMAT(I5,3(G12.5,A1,1X))
1100 FORMAT('$NEW VALUES: ')
1110 FORMAT('$XMIN,XINC: ')
1120 FORMAT(' ENTER DATA, ONE INTEGER PER LINE (^Z=END)')
1130 FORMAT('ODATA POINTS',I4,' TO ',I4,':')
1140 FORMAT('$MAX.REL.HALFWIDTH: ')
1190 FORMAT('$RMS QUIT: ')
1200 FORMAT('$MAX.CHANGE QUIT: ')
1230 FORMAT('$FACTORS: ADDITIVE, MULTIPLICATIVE: ')
1240 FORMAT('$LISTING FILE: ')
1250 FORMAT('$PLOT FILE: ')
C
C       INPUT FORMATS
C
2000 FORMAT(2I6)
2010 FORMAT(3F12.0)
2020 FORMAT(80A1)
2040 FORMAT(4F8.0,3I4)
2050 FORMAT(I4,3F8.0)
       END
       SUBROUTINE SETFIL
C
C CREATE DEFAULT PLOT AND LIST FILENAMES
C
       COMMON/FILCOM/OUTFIL(16),INPFIL(16),LSTFIL(16),
       *PLTFIL(16)
          BYTE OUTFIL,INPFIL,LSTFIL,PLTFIL
C DEFAULT PLOT AND LIST FILES ARE LP:XXXXXX
C FIND END OF DEVICE NAME, IF SPECIFIED
       DO 27 I=1,4
          IF(INPFIL(I).EQ.':') GOTO 28
27     CONTINUE
       I = 0
28     LSTFIL(1) = 'L'
       LSTFIL(2) = 'P'
       LSTFIL(3) = ':'
       DO 29 J=1,6
          IF(INPFIL(I+J).EQ.',') GOTO 30
          LSTFIL(J+3) = INPFIL(I+J)
29     CONTINUE
30     DO 31 I=J+3,16
          LSTFIL(I) = ' '
31     CONTINUE
       DO 32 I=1,16
          PLTFIL(I) = LSTFIL(I)
32     CONTINUE
       RETURN
       END

C-------------------------------------------------------------
C COMNUM.FOR
C-------------------------------------------------------------
       INTEGER FUNCTION COMNUM(CSTRNG,LSTRNG,COMSTR,COMPTR,
       *MAXCOM,SWISTR,SWIPTR,MAXSWI,SWIFLG,MAXFLG,IPTR)
C V00-04
C
C R.F.MURPHY     07-NOV-78
C
C CONVERTS A COMMAND STRING AND ASSOCIATED SWITCHES TO A
C COMMAND NUMBER AND SET FLAGS IN THE SWITCH FLAG ARRAY.
C
C INPUT ARRAY FORMATS
C
C  CSTRNG(LSTRNG)  INPUT COMMAND STRING
C  COMSTR()        CONSECUTIVE ASCII NAMES OF COMMANDS.
C  COMPTR(2,1)     FIRST ELEMENT IS POINTER TO END OF
C                  FULL COMMAND NAME, SECOND ELEMENT IS
C                  POINTER TO END OF UNIQUE COMMAND NAME.
C  SWISTR()        SIMILAR TO COMSTR FOR SWITCHES.
C  SWIPTR(2,1)     SIMILAR TO COMPTR FOR SWITCHES.
C  SWIFLG(3,1)     (1,N) IS SWITCH #.
C                  (2,N) IS # OF COMMAND IT IS ALLOWED ON.
C                  (3,N) IS SET TO # OF TIMES SEEN.
C
       BYTE CSTRNG(LSTRNG),COMSTR(1),SWISTR(1)
       INTEGER LSTRNG,COMPTR(2,MAXCOM),MAXCOM,SWIPTR(2,MAXSWI),
       *MAXSWI,SWIFLG(3,MAXFLG),MAXFLG,IPTR
       INTEGER I,J,SWINUM
       INTEGER COMIND
C
       BYTE PROMPT(3)
C
       DATA PROMPT     /10,'$',128/
C
       IF(LSTRNG.GT.0) GOTO 3
       CALL GTLIN(CSTRNG,PROMPT)
       DO 1 LSTRNG = 1,81
          IF(CSTRNG(LSTRNG).EQ.0) GOTO 2
1      CONTINUE
2      LSTRNG = LSTRNG - 1
       GOTO 4
3      CSTRNG(LSTRNG+1) = 0
C CLEAR SWITCH FLAGS.
4      I = 1
5      IF(I.GT.MAXFLG) GOTO 6
          SWIFLG(3,I) = 0
          I = I + 1
          GOTO 5
C FIND COMMAND NAME IN COMSTR ARRAY.
6      IPTR = 1
       COMNUM = COMIND(CSTRNG,COMSTR,COMPTR,MAXCOM,IPTR)
       IF(COMNUM.LE.0) GOTO 50
C VALID COMMAND FOUND, CHECK FOR SWITCHES.
C IPTR HAS BEEN SET TO THE FIRST CHAR AFTER THE COMMAND NAME.
```

```
10      IF(CSTRNG(IPTR).NE.'/') GOTO 40
C MAKE SURE SWITCHES ARE ALLOWED.
        IF(MAXSWI.LE.0) GOTO 30
C SKIP THE /.
        IPTR = IPTR + 1
        SWINUM = COMIND(CSTRNG,SWISTR,SWIPTR,MAXSWI,IPTR)
        IF(SWINUM) 35,30,15
C VALID SWITCH FOUND, SET FLAG AND CHECK FOR MORE.
15      DO 20 J=1,MAXFLG
           IF(SWIFLG(1,J).NE.SWINUM.OR.SWIFLG(2,J).NE.COMNUM)
        *    GOTO 20
           SWIFLG(3,J) = SWIFLG(3,J) + 1
           GOTO 10
20      CONTINUE
C NOTHING SHOULD PASS THROUGH THIS DO LOOP.
C
C BAD SWITCH
30      COMNUM = -2
        GOTO 50
C AMBIGUOUS SWITCH.
35      COMNUM = -3
        GOTO 50
C ADVANCE POINTER TO START OF ARGUMENTS, IF ANY.
40      IF(CSTRNG(IPTR).NE.' ') GOTO 50
        IPTR = IPTR + 1
        GOTO 40
50      RETURN
        END

C-------------------------------------------------------------
C COMIND.FOR
C-------------------------------------------------------------
        INTEGER FUNCTION COMIND(STRING,STRNGS,STRPTR,
        *MAXIND,IPTR)
C V00-03
C
C R.F.MURPHY    07-NOV-78
C
C FINDS A STRING IN A LARGER STRING ARRAY AND
C RETURNS THE STRING NUMBER
C       SEE COMNUM.FOR FOR ARGUMENT FORMAT.
C
        BYTE STRING(1),STRNGS(1)
        INTEGER STRPTR(2,MAXIND),MAXIND,IPTR
C
        DO 20 COMIND = 1,MAXIND
           I = IPTR
           J = 1
C BEGINNING OF STRING N IS 1 PLUS END OF STRING N-1.
           IF(COMIND.GT.1) J = 1 + STRPTR(1,COMIND-1)
           K = STRPTR(1,COMIND)
10         IF(STRING(I).GT.'Z'.OR.STRING(I).LT.'A') GOTO 30
           IF(J.GT.K) GOTO 25
           IF(STRING(I).NE.STRNGS(J)) GOTO 20
           I = I + 1
           J = J + 1
           GOTO 10
20      CONTINUE
C NO MATCH.
25      COMIND = 0
        GOTO 40
C MATCH SO FAR, SEE IF ENOUGH CHARS WERE SPECIFIED.
30      IF(J.GT.STRPTR(2,COMIND)) GOTO 40
C NOT ENOUGH MATCHED. IF NONE MATCHED, COMMAND IS ILLEGAL.
        IF(J.EQ.1) GOTO 25
C SOME MATCHED, BUT COMMAND IS AMBIGUOUS.
        COMIND = -1
C SAVE END OF MATCHED STRING.
40      IPTR = I
        RETURN
        END

C
C PLTLIM.FOR
C
        SUBROUTINE PLTLIM(MIN,MAX,PLTMIN,PLTMAX)
C V00-01        01-FEB-79
C
C R.F.MURPHY    CREATED 27-DEC-78
C
C CALCULATES DEFAULTPLOT LIMITS FROM ACTUAL
C DATA MIN AND MAX.
C
        REAL MIN,MAX,PLTMIN,PLTMAX
        INTEGER POWER
        REAL T,AT
C
        T = AMAX1(ABS(MIN),ABS(MAX))
        POWER = 0
        IF(T.EQ.0) GOTO 10
C ALLOWABLE LIMITS ARE A.B*10**LOG(X),
C WHERE A AND B ARE SINGLE DIGITS.
        T = ALOG10(T)
        AT = AINT(T)
        POWER = INT(AT) - 1
C ROUND POWER DOWN FOR NEGATIVE LOGS.
        IF(AT.GT.T) POWER = POWER - 1
C ROUND DOWN FOR MINIMUM.
10      T = MIN/10.**POWER
        AT = AINT(T)
        IF(AT.GT.T) AT = AT -1.
        PLTMIN = AT*10.**POWER
C ROUND UP FOR MAXIMUM.
        T = MAX/10.**POWER
        AT = AINT(T)
        IF(AT.LT.T) AT = AT + 1.
        PLTMAX = AT*10.**POWER
        RETURN
        END

C-------------------------------------------------------------
C GUESSG.FOR
C-------------------------------------------------------------
        SUBROUTINE GUESSG(IY,NOPTS,XMIN,XDEL,IYFILT,
        *ZFILT,P,NOPAR,MAXPAR)
C V00-04        28-JUN-79
C
C R.F.MURPHY    CREATED 29-DEC-78
C
C MAKES PARAMETER GUESSES FOR GELFIT. A CURVE IS CREATED
```

```
C BETWEEN LOCAL MAXIMA AND MINIMA OF THE
C FIRST DERIVATIVE (INFLECTION POINTS).
C THESE MAXIMA AND MINIMA MUST BE GREATER THAN
C THE Z FILTER, AND THE Y MAXIMUM IN
C THAT INTERVAL MUST BE GREATER THAN THE Y FILTER.
C REASONABLE VALUES FOR THE FILTERS ARE:
C       IYFILT = 0.05*IYMAX
C       ZFILT = 0.001*(IYMAX-IYMIN)/XDEL
C
        INTEGER NOPTS, IY(NOPTS), NOPAR, MAXPAR, IYFILT
        REAL XMIN, XDEL, ZFILT, P(MAXPAR)
C
        LOGICAL*1 UP,LFOUND,RFOUND,INFTST
        REAL YLEFT,YRIGHT,ZLEFT,ZRIGHT,ZLINF,ZRINF
C
C SMOOTH THE Y DATA USING A SIMPLE POLYNOMIAL SMOOTH.
C SET UP INITIAL VALUES.
        NOPAR = 0
        UP = .FALSE.
        LFOUND = .FALSE.
        YLEFT = (3.*IY(1)+IY(2))/4.
        YRIGHT = (IY(1)+2.*IY(2)+IY(3))/4.
        ZLEFT = (YRIGHT-YLEFT)/XDEL
C REQUIRE INFLECTION POINT TO BE GREATER THAN LEFT BOUNDARY.
        ZLINF = ZLEFT
        YLEFT = YRIGHT
C
C START CHECK LOOP.
C
C 'UP' IS A FLAG FOR THE PREVIOUS CONDITION ZRIGHT>ZLEFT
C INFLECTION POINT CRITERIA:
C
C       LEFT) UP & (ZRIGHT-ZLINF)<ZFILT
C
C       RIGHT) .NOT.UP & (ZRIGHT-ZRINF)>ZFILT
C
C CURVE CRITERIA:
C       IY(IMAXP)>IYFILT,
C               WHERE IMAXP IS BETWEEN INFLECTION POINTS.
C
        DO 50 I=2,NOPTS-2
           YRIGHT = (IY(I)+2.*IY(I+1)+IY(I+2))/4.
           ZRIGHT = (YRIGHT-YLEFT)/XDEL
C SEE IF LEFT INF POINT FOUND ALREADY.
           IF(LFOUND) GOTO 10
C LEFT INF POINT REQUIRES UP; DON'T CHANGE ZLINF UNTIL TRUE.
           IF(.NOT.UP) GOTO 40
           IF(ZLEFT.GT.ZLINF) ZLINF = ZLEFT
           IF( (ZRIGHT-ZLINF) .GE.ZFILT) GOTO 40
C FOUND LEFT INF POINT. INIT POINTERS.
           LFOUND = .TRUE.
           RFOUND = .FALSE.
           ZRINF = ZRIGHT
           ILEFT = I-1
           IMAXP = I
           GOTO 40
C ALREADY FOUND LEFT. UPDATE MAX POINTER.
10         IF(IY(I).GT.IY(IMAXP)) IMAXP = I
C CHECK RIGHT INF CRITERIA.
           INFTST = (.NOT.UP.AND.(ZRIGHT-ZLEFT).GT.ZFILT)
C UPDATE LOWEST Z SEEN.
           IF(ZRIGHT.LT.ZRINF) ZRINF = ZRIGHT
C SEE IF PREVIOUS POINT PASSED CRITERIA.
           IF(RFOUND) GOTO 20
           RFOUND = INFTST
           GOTO 40
C LAST POINT PASSED. IF THIS DOES ALSO, CONTINUE SEARCH.
20         IF(INFTST) GOTO 40
C REGION PASSES Z CRITERIA. CHECK Y FILTER.
           IF(IY(IMAXP).LE.IYFILT) GOTO 30
           J = NOPAR + 3
           IF(J.GT.MAXPAR) GOTO 60
           NOPAR = J
           P(NOPAR-2) = XMIN + (IMAXP-1)*XDEL
           P(NOPAR-1) = FLOAT(IY(IMAXP))
           P(NOPAR) = (I-ILEFT)*XDEL/4.
30         LFOUND = .FALSE.
40         UP = (ZRIGHT.GT.ZLEFT)
           YLEFT = YRIGHT
           ZLEFT = ZRIGHT
50      CONTINUE
        GOTO 70
60      CALL ERROR('W','param array full - curve(s) lost.')
70      RETURN
        END

C-------------------------------------------------------------
C GENERD.FOR
C-------------------------------------------------------------
        SUBROUTINE GENERD(X1,X2,LOGFLG,NOPAR,P,ERRP,OLDP,
        *FUNCTN,NOPTS,X,Y,IX,IY,INTFLG,YONFLG)
C V01            29-MAY-79
C
C R.F.MURPHY     CREATED 29-MAY-79
C
C GENERATE DATA FROM FUNCTION PARAMETERS, ALLOWING
C PARAMETERS TO VARY IN NORMAL DISTRIBUTIONS WITH
C SPECIFIED STANDARD DEVIATIONS. DATA IS EVENLY
C X-INCREMENTED UNLESS LOGFLG IS .TRUE., IT IS
C REAL UNLESS INTFLG IS .TRUE., AND IT IS
C X&Y UNLESS YONFLG IS .TRUE.
C OLDP NEED ONLY BE SUPPLIED IF AT LEAST ONE ERRP IS NON-ZERO.
C ONLY THE NECESSARY SUBSET OF X,Y,IX,IY (DETERMINED BY
C THE INTFLG AND YONFLG SETTINGS) NEED BE SUPPLIED.
C
        INTEGER NOPAR, NOPTS
        LOGICAL LOGFLG,INTFLG,YONFLG
        INTEGER IX(NOPTS),IY(NOPTS)
        REAL X1,X2,FUNCTN
        REAL P(NOPAR),ERRP(NOPAR),OLDP(NOPAR),X(NOPTS),Y(NOPTS)
C
        INTEGER I,J
        REAL XT,XDEL,YT
C
C MAKE SURE ARGUMENTS ARE VALID.
        IF(NOPTS.LE.1.OR.NOPAR.LE.0) GOTO 50
C SAVE ORIGINAL PARAMETERS.
5       DO 10 I=1,NOPAR
           IF(ERRP(I).NE.0.) OLDP(I) = P(I)
10      CONTINUE
```

```
        XDEL = 0.
        IF(NOPTS.GT.1) XDEL = (X2-X1)/(NOPTS-1)
        XT = X1
        IF(.NOT.LOGFLG) GOTO 15
        XDEL = 10.**XDEL
        XT = 10.**XT
C CALCULATE NEW PARAMETERS FOR EACH POINT,
C AND EVALUATE THE FUNCTION.
15      DO 30 I=1,NOPTS
           DO 20 J=1,NOPAR
              IF(ERRP(J).EQ.0.) GOTO 20
              P(J) = GAUS(OLDP(J),ERRP(J))
20         CONTINUE
           IF(YONFLG) GOTO 24
           IF(INTFLG) GOTO 22
           X(I) = XT
           GOTO 24
22         IX(I) = INT(XT)
24         YT = FUNCTN(XT)
           IF(INTFLG) GOTO 26
           Y(I) = YT
           GOTO 28
26         IY(I) = INT(YT)
28         IF(LOGFLG) GOTO 29
           XT = XT + XDEL
           GOTO 30
29         XT = XT*XDEL
30      CONTINUE
C RESTORE ORIGINAL PARAMETERS.
        DO 40 I=1,NOPAR
           IF(ERRP(I).NE.0.) P(I) = OLDP(I)
40      CONTINUE
50      RETURN
        END

C-------------------------------------------------------------
C FTGAUS.FOR
C-------------------------------------------------------------
        SUBROUTINE FTGAUS(BANDPP)
C V04-04        28-JUN-79
C
C R.F.MURPHY    CREATED 1-JUL-77
C
C THIS SUBROUTINE FITS GAUSSIAN COMPONENTS USING A
C MODIFIED DAMP LEAST SQUARES ALGORITHM DESIGNED BY MARQUARDT.
C J. SOC. IND. APPL. MATH. (1965) 11:431-441.
C RETURNS CHISQ AND DEGREES OF FREEDOM
C FOR SIGNIFICANCE CALCULATION.
C
C INPUTS
C
C  LWD             /LOGDEV/  LOGICAL OUTPUT DEVICE
C  XMIN            /DATA/    X VALUE ASSOC. WITH IY(1)
C  XMAX            /DATA/    X VALUE ASSOC. WITH IY(NOPTS)
C  NOPTS           /DATA/    # OF DATA POINTS
C  IY(NOPTS)       /DATA/    ARRAY OF Y DATA VALUES
C  NOCURV          /CURVES/  NOPAR/3
C  NOPAR           /CURVES/  # OF FIT PARAMETERS
C  P(NOPAR)        /CURVES/  PARAMETER ARRAY
C  NIT             /FITPAR/  # OF ITERATIONS TO TRY
C  NCIT            /FITPAR/  # OF CORRECTIVE TRIES PER ITERATION
C  RMSQU           /FITPAR/  QUIT WHEN RMS REACHES THIS VALUE
C  DELQU           /FITPAR/  QUIT WHEN MAX CHANGE REACHES THIS
C  RATMAX          /FITPAR/  MAX RATIO OF HW TO YMAX
C  NFPAR           /FITPAR/  # PARAM WIDTH OF BAND MATRIX
C  XDEL            /FITPAR/  CHANGE IN X BETWEEN IY(I) & IY(I+1)
C  FIXED           /FIXCOM/  CHARACTER FOR FIXED PARAM
C  FIXARR(NOPAR)   /FIXCOM/  INDICATES WHICH PARAMS TO FIX
C  GELFUN          <GLOBAL>  CALCULATES GAUSSIAN FUNCTION
C   DEBUG          /DEBUG/   DEBUGGING FLAG
C
C WORKING ARRAYS AND VARIABLES
C       (COMMONS ARE DIMENSIONED HERE FOR 30 CURVES
C       SO THAT THEY ARE INCLUDED IN WHICHEVER
C       OVERLAY THIS MODULE IS PLACED IN;
C       COMMONS MAY BE EXTENDED BY MAIN PROGRAM
C       IN WHICH CASE THEY WILL BE IN THE ROOT SEGMENT)
C
C  BANDPP(NDIFF*NFPAR*2)  <ARG>  PARTIAL OF EACH PARAM TO
C                                EACH PARAM IN LOWER HALF;
C                                SPACE FOR DECOMPOSITION
C                                OF BANDPP IN UPPER HALF
C  PARCAL          /PARCAL/  GELFUN FLAG TO CALCULATE PARTDX
C  DELMX           /FITPAR/  MAX REL CHANGE IN ANY PARAMETER
C  PARTDX(NOPAR)   /PARTDX/  PARTIAL DERIV WITH RESPECT TO X
C  PARTDY(NDIFF)   /PARTDY/  SCALED PARTIAL DERIVATIVES
C  DELTA(NDIFF)    /DELTA/   AMOUNT TO CHANGE PARAMS BY
C  OLDP(NOPAR)     /OLDP/    STORAGE FOR PREVIOUS PARAMS
C  IMAP(NDIFF)     /IMAP/    MAPS ORIGINAL TO VARIED PARAM#'S
C
C MODULES USED
C
C  IBANDT       DECOMPOSE BANDPP MATRIX
C  BANSOL       SOLVE DECOMPOSITION FOR CHANGE VECTOR.
C
C OUTPUTS
C
C  P(NOPAR)        /CURVES/  NEW PARAMETER VALUES
C  NFREE           /FITPAR/  # OF DEGREES OF FREEDOM
C  QCOND           /FITPAR/  CONDITION # WHICH CAUSED QUIT
C                               0 = NIT DONE
C                               1 = NCIT EXCEEDED
C                               2 = DELQU ATTAINED
C                               3 = RMSQU ATTAINED
C                               4 = EQUATIONS SINGULAR
C                               5 = RATMAX FAILURE
C  RMS             /FITPAR/  ROOT MEAN SQUARE ERROR OF FIT
C  CHISQ           /FITPAR/  "GOODNESS OF FIT"
C
C METHOD:
C
C 1.    STORE PARAMETERS.
C 2.    CALCULATE PARTDX,PARTDY,BANDPP,DIFFSQ.
C       CALCULATE RMS, CHISQ.
C       OUTPUT RMS, CHISQ IF ANY ITS DONE.
C 3.    CHECK TERMINATION. IF DONE, RETURN.
C 4.    SCALE PARTDY AND BANDPP.
C 5.    ADD CORRECTION FACTOR (INITIALLY 0.01) TO
C       COLUMN NFPAR OF BANDPP.
C       DECOMPOSE BANDPP AND SOLVE FOR DELTA.
```

```
C       SCALE DELTA.
C 6.    ADD XKGAM*DELTA TO P.
C       CALCULATE DIFFSQ.
C       IF FIT WORSE: CHECK COSINE CRITERION:
C       *               GT, DECREASE XKGAM, GOTO 6.
C       *               LE, INCREASE FACTOR, GOTO 5.
C       FIT BETTER, CHECK HALFWIDTHS
C       *               BAD, CORRECT AND GOTO 2.
C       FIT OK, DECREASE CORRECTION FACTOR.
C       GOTO 1.
        REAL FUNCTN
        EXTERNAL FUNCTN
C
        REAL BANDPP(1)
        COMMON/LOGDEV/LRD,LWD
           INTEGER LRD,LWD
        COMMON/DATA/TITLE(80),XMIN,XMAX,YMIN,YMAX,
        *NOPTS,MAXPTS,IY(1)
           BYTE TITLE
           INTEGER NOPTS,MAXPTS,IY
           REAL XMIN,XMAX,YMIN,YMAX
        COMMON/CURVES/NOCURV,NOPAR,MAXPAR,P(1)
           INTEGER NOCURV,NOPAR,MAXPAR
           REAL P
        COMMON/FITPAR/NIT,NCIT,RMSQU,DELQU,RATMAX,NEIGHB,NFPAR,
        *  MAXBND,XDEL,NFREE,QCOND,RMS,CHISQ,DELMX
           INTEGER NIT,NCIT,NEIGHB,NFPAR,MAXBND,NFREE,QCOND
           REAL RMSQU,DELQU,RATMAX,XDEL,RMS,CHISQ,DELMX
        COMMON/PARCAL/PARCAL
           LOGICAL PARCAL
        COMMON/FIXCOM/FIXED,UNFIXD,FIXARR(1)
           BYTE FIXED,UNFIXD,FIXARR
        COMMON/DEBUG/DEBUG
           INTEGER DEBUG
        COMMON/PARTDX/PARTDX(90)
           REAL PARTDX
        COMMON/PARTDY/PARTDY(90)
           REAL PARTDY
        COMMON/DELTA/DELTA(90)
           REAL DELTA
        COMMON/OLDP/OLDP(90)
           REAL OLDP
        COMMON/IMAP/IMAP(90)
           INTEGER IMAP
C
        LOGICAL*1 ALLOK
C ITC IS THE CORRECTION ITERATION COUNTER.
C ITNO CONTAINS THE CURRENT ITERATION NUMBER.
C I-->M ARE TEMPORARY VARIABLES.
        INTEGER I,IPMAX,ITC,ITNO,J,K,L,M
C THESE VARIABLES HAVE SPECIFIC VALUES DURING THE FITTING.
        INTEGER NDIFF,NDIFF1,NFPAR1,NPPSIZ,NPPSZ1
C MISCELLANEOUS VARIABLES.
        INTEGER IYFUNV,IYDATV
        REAL DIFFN,DIFFSQ,DIFF,DIFF2,FLAM,XKGAM,XLAM,
        *GAMNUM,TEMP,TEMPSQ,DEL,DELSQ,STIME,X
C PROCEDURE TYPE DECLARATIONS.
        INTEGER IBANDT
        REAL GELFUN
C***
        REAL SECNDS
C
C CREATE MAP ARRAY TO POINT TO UNFIXED PARAMETERS.
        NDIFF = 0
        I = 1
D       GOTO 1
        IF(DEBUG.NE.0)
        * CALL ERROR('W','Debussing code not compiled.')
1       IF(I.GT.NOPAR) GOTO 3
           IF(FIXARR(I).EQ.FIXED) GOTO 2
           NDIFF = NDIFF + 1
           IMAP(NDIFF) = I
2          I = I + 1
           GOTO 1
3       IF(NDIFF.GT.0) GOTO 4
C ALL PARAMETERS ARE FIXED. DON'T DO ANY ITERATIONS.
        NIT = 0
        NDIFF = 1
        IMAP(1) = 1

C INITIALIZE VARIABLES.
4       NFREE = NOPTS - NDIFF
        NDENOM = MAX0(1,NFREE)
        IF(NFREE.LE.0.)
        * CALL ERROR('W','solution underdetermined.')
        IF(NFPAR.GT.NDIFF) NFPAR = NDIFF
        NFPAR1 = NFPAR - 1
        NDIFF1 = NDIFF - 1
        NPPSIZ = NDIFF*NFPAR
        NPPSZ1 = NPPSIZ + 1
        XLAM = 0.01
        QCOND = 0
        ITC = 0
        ITNO = 0
        ALLOK = .TRUE.
        IF(NIT.GT.0) WRITE(LWD,1000)
C MAIN ITERATION LOOP.
C SAVE PREVIOUS PARAMETER VALUES.
5       DO 10 I=1,NOPAR
           OLDP(I) = P(I)
10      CONTINUE
C RETURN HERE IF RATMAX FAILS.
C CALCULATE PARTIAL DERIVATIVE MATRICES.
C BANDPP IS ACTUALLY BANDPP(NDIFF,NFPAR), WHERE BANDPP(I,J)
C IS THE PRODUCT OF THE PARTIALS OF PARAMS I AND I+NFPAR-J.
C THUS, THE SQUARES OF THE PARTIALS OF THE PARAMS ARE IN
C COLUMN NFPAR.
C
C IF PARCAL=.TRUE., 'GELFUN' WILL CALCULATE PARTDX.
15      PARCAL = (ITNO.LT.NIT.AND.ITC.LE.NCIT)
        IF(.NOT.PARCAL) GOTO 35
        DO 20 I=1,NDIFF
           PARTDY(I) = 0.
20      CONTINUE
        DO 30 I=1,NPPSIZ
           BANDPP(I) = 0.
30      CONTINUE
35      DIFFSQ = 0.
        CHISQ = 0.
        X = XMIN
```

```
      DO 60 I=1,NOPTS
C ROUND OFF FUNCTION VALUE TO ACCURACY OF DATA;
C MINIMIZES EFFECT OF SMALL DIFFERENCES ON CHISQ AND RMS.
         IYFUNV = INT(GELFUN(X)+.5)
         IYDATV = IY(I)
         DIFF = IYDATV - IYFUNV
         DIFF2 = DIFF*DIFF
         DIFFSQ = DIFFSQ + DIFF2
C MINIMIZE ERROR DUE TO POINTS WHERE Y(X) = 0.
         IF(IYFUNV.EQ.0) IYFUNV = IYDATV
C DIFF2 = 0. IF IYFUNV = 0.
         IF(IYFUNV.NE.0) CHISQ = CHISQ + DIFF2/IYFUNV
         X = X + XDEL
         IF(.NOT.PARCAL) GOTO 60
         DO 50 J=1,NDIFF
            PARTDY(J) = PARTDY(J) + DIFF*PARTDX(IMAP(J))
            K = NPPSIZ - NDIFF + J
            L = MINO(J+NFPAR1,NDIFF)
            DO 40 M=J,L
               BANDPP(K)=
     *         BANDPP(K)+PARTDX(IMAP(J))*PARTDX(IMAP(M))
               K = K - NDIFF1
40          CONTINUE
50       CONTINUE
60    CONTINUE
D     IF(DEBUG.EQ.0) GOTO 59
D     CALL REMARK('AFTER 60')
D     CALL REMARK('PARTDY')
D     WRITE(LWD,3000) (PARTDY(I),I=1,NDIFF)
D     CALL RCTRLO
D     CALL REMARK('BANDPP')
D     WRITE(LWD,3000) (BANDPP(I),I=1,NPPSIZ)
D     CALL RCTRLO
D59   CONTINUE
C CALCULATE ERROR VALUES AND CHECK TERMINATION CONDITIONS.
      RMS = SQRT(DIFFSQ/NOPTS)
      IF(ITNO.EQ.0) GOTO 62
C***+
      STIME = SECNDS(STIME)
      IF(ITC.GT.NCIT.OR..NOT.ALLOK) GOTO 61
      WRITE(LWD,1010) STIME,RMS,DELMX,IPMAX,P(IPMAX),ITC
C     WRITE(LWD,1010) RMS,DELMX,IPMAX,P(IPMAX),ITC
      ITC = 0
      GOTO 62
61    WRITE(LWD,1010) STIME
C REMEMBER TIME EACH ITERATION IS STARTED.
62    STIME = SECNDS(0.)
C***-
      ITNO = ITNO + 1
      IF(ITNO.LE.NIT) GOTO 63
      GOTO 400
63    IF(ITC.LE.NCIT) GOTO 64
      QCOND = 1
      IF(.NOT.ALLOK) QCOND = 5
      GOTO 400
64    IF(DELMX.GT.DELQU.OR.ITNO.EQ.1) GOTO 66
      QCOND = 2
      GOTO 400
66    IF(RMS.GT.RMSQU) GOTO 68
      QCOND = 3
      GOTO 400
C FILL PARTDX WITH THE ABSOLUTE VALUE OF THE PARTIAL OF
C EACH UNFIXED PARAM. SCALE MATRICES BY PARTDX.
68    DO 80 I=1,NDIFF
         K = NPPSIZ - NDIFF + I
         PARTDX(I) = SQRT(BANDPP(K))
         IF(PARTDX(I).NE.0.) PARTDY(I) = PARTDY(I)/PARTDX(I)
         L = MAX0(1,I-NFPAR1)
         DO 70 J=I,L,-1
            X = PARTDX(I)*PARTDX(J)
            IF(X.NE.0.) BANDPP(K) = BANDPP(K)/X
            K = K - NDIFF
70       CONTINUE
80    CONTINUE
D     IF(DEBUG.EQ.0) GOTO 82
D     CALL REMARK('AFTER 80')
D     CALL REMARK('PARTDX')
D     WRITE(LWD,3000) (PARTDX(I),I=1,NDIFF)
D     CALL RCTRLO
D     CALL REMARK('BANDPP')
D     WRITE(LWD,3000) (BANDPP(I),I=1,NPPSIZ)
D     CALL RCTRLO
D82   CONTINUE
C INITIALIZE CORRECTION VARIABLES.
      FLAM = 0.
      XKGAM = 1.
C CORRECTIVE LOOP.
C ADD FACTOR TO COLUMN NFPAR.
100      FLAM = XLAM - FLAM
         DO 110 I=NPPSIZ-NDIFF1,NPPSIZ
            BANDPP(I) = FLAM + BANDPP(I)
110      CONTINUE
C DECOMPOSE BANDPP.
         IF(IBANDT(BANDPP,BANDPP(NPPSZ1),
     *   NDIFF,NFPAR).NE.0) GOTO 130
C EQUATIONS ARE SINGULAR. RESTORE OLD PARAMS AND FORCE RETURN.
            QCOND = 4
            DO 120 I=1,NOPAR
               P(I) = OLDP(I)
120         CONTINUE
            ITNO = NIT + 1
            GOTO 5
C SOLVE FOR DELTA (PARAMETER CHANGE VECTOR).
130      CALL BANSOL(BANDPP(NPPSZ1),NDIFF,NFPAR,DELTA,PARTDY)
D     IF(DEBUG.EQ.0) GOTO 132
D     CALL REMARK('AFTER 130')
D     CALL REMARK('PARTDX')
D     WRITE(LWD,3000) (PARTDX(I),I=1,NDIFF)
D     CALL RCTRLO
D     CALL REMARK('DELTA')
D     WRITE(LWD,3000) (DELTA(I),I=1,NDIFF)
D     CALL RCTRLO
D132  CONTINUE
C SCALE THE DELTAS.
         DO 140 I=1,NDIFF
            IF(PARTDX(I).NE.0.) DELTA(I)=DELTA(I)/PARTDX(I)
140      CONTINUE
C CORRECTIVE LOOP FOR FITS THAT MEET COSINE CRITERION.
200         DO 210 I=1,NDIFF
               P(IMAP(I)) = P(IMAP(I)) + XKGAM*DELTA(I)
```

```
210           CONTINUE
D       IF(DEBUG.EQ.0) GOTO 212
D       CALL REMARK('AFTER 210')
D       CALL REMARK('P')
D       WRITE(LWD,3000) (P(I),I=1,NOPAR)
D       CALL RCTRLO
D212    CONTINUE
C CALCULATE ERROR OF NEW FIT.
              DIFFN = 0.
              PARCAL = .FALSE.
              X = XMIN
              DO 220 I=1,NOPTS
                 DIFF = IY(I) - INT(GELFUN(X)+.5)
                 DIFFN = DIFFN + DIFF*DIFF
                 X = X + XDEL
220           CONTINUE
C IF FIT IS BETTER, JUMP OUT OF CORRECTIVE LOOP.
              IF(DIFFN.LT.DIFFSQ) GOTO 300
C FIT IS WORSE. RESTORE OLD PARAMS AND CHECK COSINE CRITERION.
D       IF(DEBUG.NE.0) CALL REMARK('FIT WORSE')
              DO 230 I=1,NOPAR
                 P(I) = OLDP(I)
230           CONTINUE
              ITC = ITC + 1
              IF(ITC.GT.NCIT) GOTO 5
C CALCULATE VALUES FOR COSINE CORRECTION.
              GAMNUM = 0.
              TEMPSQ = 0.
              DELSQ = 0.
              DO 240 I=1,NDIFF
                 TEMP = PARTDY(I)
                 DEL = DELTA(I)
                 GAMNUM = GAMNUM + TEMP*DEL
                 TEMPSQ = TEMPSQ + TEMP*TEMP
                 DELSQ = DELSQ + DEL*DEL
240           CONTINUE
C THAT'S COS(PI/4).
              TEMP = SQRT(TEMPSQ*DELSQ) * 0.7071067812
              IF(GAMNUM.LE.TEMP) GOTO 250
C COSINE CRITERION MET. DECREASE AMOUNT ADDED TO PARAMS.
              XKGAM = XKGAM/2.
              GOTO 200
C CRITERION NOT MET. INCREASE AMOUNT ADDED TO BANDPP.
250         XLAM = 10.*XLAM
            GOTO 100
C FIT IS BETTER. CHECK HALFWIDTH RATIOS.
300     IF(RATMAX.EQ.0.) GOTO 308
D       IF(DEBUG.NE.0)
D       * CALL REMARK('FIT BETTER - CHECKING RATMAX')
        ALLOK = .TRUE.
        DO 302 I=3,NOPAR,3
           TEMP = P(I-1)
           IF(TEMP.EQ.0.) GOTO 302
           IF(ABS(P(I)/TEMP).LE.RATMAX) GOTO 302
           ALLOK = .FALSE.
C RATMAX FAILURE. NARROW FAILING CURVE TO YMAX*RATMAX*.8.
C WIDEN PRECEDING CURVE, EITHER HW*2 OR YMAX*RATMAX*.8.
           P(I) = RATMAX*TEMP*.8
           IF(I.GT.3) P(I-3)=AMIN1(P(I-3)*2.,P(I-4)*RATMAX*.8)
302     CONTINUE
        IF(ALLOK) GOTO 308
C AT LEAST ONE HW TOO BIG.
        ITC = ITC + 1
        GOTO 15
C ACCEPTABLE FIT. DECREASE AMOUNT ADDED TO DIAGONAL OF BANDPP,
C CALCULATE THE MAXIMUM RELATIVE CHANGE IN A PARAM,
C AND GO BACK FOR NEXT ITERATION.
308     XLAM = XLAM/10.
D       IF(DEBUG.NE.0) CALL REMARK('FIT ACCEPTABLE')
        DELMX = 0.
        IPMAX = 1
        DO 310 I=1,NOPAR
           DEL = ABS(P(I)-OLDP(I))
           IF(OLDP(I).NE.0.) DEL = ABS(DEL/OLDP(I))
           IF(DEL.LE.DELMX) GOTO 310
           DELMX = DEL
           IPMAX = I
310     CONTINUE
        GOTO 5
C FITTING TERMINATED.
C MAKE SURE ALL HALF-WIDTHS ARE POSITIVE.
400     DO 410 I=3,NOPAR,3
           P(I) = ABS(P(I))
410     CONTINUE
500     RETURN
C***+
1000    FORMAT(//1X,'ET(SEC)',5X,'RMS',7X,'MAX.REL.CHANGE IN P#',
        *'  NEW VALUE',4X,'CIT#')
C1000   FORMAT(//5X,'RMS',7X,'MAX.REL.CHANGE IN P#',
C       *'  NEW VALUE',4X,'CIT#')
1010    FORMAT(2X,F5.0:2X,2G14.7,'  (',I2,')',G14.7,I4)
C1010   FORMAT(1X,2G14.7,'  (',I2,'),G14.7,I4)
C***-
D3000   FORMAT(9(1X,G12.5))
        END

C-------------------------------------------------------------
C GELFUN.FOR
C-------------------------------------------------------------
        REAL FUNCTION GELFUN(X)
C V00
C
C R.F.MURPHY     CREATED 17-DEC-78
C
C FOR USE WITH GELFIT.
C RETURNS VALUE OF FIT AT X.
C
C INPUTS
C
C  X               <ARG>      VALUE TO EVALUTE FUNCTION AT
C  NCRVS          /CURVES/   NUMBER OF CURVES IN FIT
C  P(3,NCRVS)     /CURVES/   FIT PARAMETERS
C  PARCAL          /PARCAL/   IF .TRUE, DO PARTDX CALCULATION
C
C OUTPUTS
C
C  GELFUN          <RESULT>   VALUE OF FUNCTION AT X
C  PARTDX(3,NCRVS) /PARTDX/   PARTIAL DERIVATIVE OF EACH PARAM
C
        REAL X
```

```
C
      COMMON/CURVES/NCRVS,NP,MAXPAR,P(3,1)
         INTEGER NCRVS,NP,MAXPAR
         REAL P
      COMMON/PARCAL/PARCAL
         LOGICAL PARCAL
      COMMON/PARTDX/PARTDX(3,1)
         REAL PARTDX
C
      INTEGER I,J
      REAL AMP,ARG,DIDSK,EXARG,XARG
C
      GELFUN = 0.
      DO 20 J = 1,NCRVS
         IF(P(3,J).EQ.0.) GOTO 10
         ARG = (X-P(1,J))/P(3,J)
         XARG = ARG*ARG
         IF(XARG.GT.15.) GOTO 10
         EXARG = EXP(-XARG)
         GELFUN = GELFUN + P(2,J)*EXARG
         IF(.NOT.PARCAL) GOTO 20
         PARTDX(2,J) = EXARG
         AMP = 2.*P(2,J)/P(3,J)
         DIDSK = AMP*ARG*EXARG
         PARTDX(1,J) = DIDSK
         PARTDX(3,J) = ARG*DIDSK
         GOTO 20
10       IF(.NOT.PARCAL) GOTO 20
         DO 15 I = 1,3
            PARTDX(I,J) = 0.
15       CONTINUE
20    CONTINUE
      RETURN
      END

C-------------------------------------------------------------
C IBANDT.FOR
C-------------------------------------------------------------
      INTEGER FUNCTION IBANDT(A,XL,N,M)
C V02-05
C
C R.F.MURPHY     21-JUN-77
C
C FORMS LL* DECOMPOSITION OF A SYMMETRIC
C POSITIVE BAND MATRIX N*M
C   SEE BANSOL.FOR FOR REFERENCE.
C
      REAL A(N,M),XL(N,M)
      INTEGER N,M
      INTEGER I,J,K,P,R,S
      DOUBLE PRECISION Y
      DO 200 I=1,N
      P=1
      IF (I.LT.M) P=M-I+1
      R=I-M+P
      DO 100 J=P,M
      S=J-1
      Q=M-J+P
      Y=A(I,J)
      IF (P.GT.S) GOTO 30
      DO 20 K=P,S
      Y=Y-XL(I,K)*XL(R,Q)
20    Q=Q+1
30    IF (J.NE.M) GOTO 50
      IF (Y.LE.0.) GOTO 210
      XL(I,J)=1./DSQRT(Y)
      GOTO 90
50    XL(I,J)=Y*XL(R,M)
90    R=R+1
100   CONTINUE
200   CONTINUE
      IBANDT = -1
      RETURN
210   IBANDT = 0
      RETURN
      END

C-------------------------------------------------------------
C BANSOL.FOR
C-------------------------------------------------------------
      SUBROUTINE BANSOL(XL,N,M,X,B)
C V02-05
C
C W.R.PEARSON, R.F.MURPHY       22-AUG-76
C
C MODIFICATION HISTORY:
C       22-AUG-76       ADAPTED FROM WILKINSON AND REINSCH
C                       LINEAR ALGEBRA 1971 SPRINGER VERLAG
C                       CONTRIBUTION I/4
C                       BY WILLIAM PEARSON
C       18-JUN-77       CONVERTED FOR USE WITH GELFIT
C                       BY R.F.MURPHY
C
C SOLVES THE EQUATIONS AX=B USING
C IBANDT DECOMPOSITION.  ACTUALLY
C SOLVES LY=B,UX=Y
C
      DIMENSION XL(N,M),X(1),B(1)
      INTEGER I,J,K,P,Q,S
      DOUBLE PRECISION Y
      S=M-1
C
C       SOLUTION OF LY=B
C
      DO 50 I=1,N
      P=1
      IF (I.LT.M) P=M-I+1
      Q=I
      Y=B(I)
      IF (S.LT.P) GOTO 30
      DO 20 K=S,P,-1
      Q=Q-1
   20 Y=Y-XL(I,K)*X(Q)
   30 X(I)=Y*XL(I,M)
   50 CONTINUE
D       WRITE(5,55) (X(I),I=1,N)
D  55 FORMAT(1X,8E10.3)
C
C       SOLUTION OF UX=Y
C
```

```
      DO 100 I=N,1,-1
      P=1
      IF ((N-I).LT.M) P=M-N+I
      Y=X(I)
      Q=I
      IF (S.LT.P) GOTO 80
      DO 70 K=S,P,-1
      Q=Q+1
   70 Y=Y-XL(Q,K)*X(Q)
   80 X(I)=Y*XL(I,M)
  100 CONTINUE
      RETURN
      END

C-------------------------------------------------------------------
C GELLST.FOR
C-------------------------------------------------------------------
      SUBROUTINE GELLST
C V02          01-JUN-79
C
C R.F.MURPHY    CREATED 13-DEC-78
C
C TYPES THE CURRENT FIT AND ERROR ESTIMATES FOR GELFIT.
C
      COMMON/LOGDEV/LRD,LWD,LFD
         INTEGER LRD,LWD,LFD
      COMMON/FILCOM/OUTFIL(16),INPFIL(16),LSTFIL(16),
     *PLTFIL(16),STDFIL(16)
         BYTE OUTFIL,INPFIL,LSTFIL,PLTFIL,STDFIL
      COMMON/DATA/TITLE(80),MINMAX(4),NOPTS,MAXPTS,IY(1)
         BYTE TITLE
         INTEGER NOPTS,MAXPTS,IY
         REAL MINMAX
      COMMON/CURVES/NCRVS,NPAR,MXPAR,P(3,1)
         INTEGER NCRVS,NPAR,MXPAR
         REAL P
      COMMON/FITPAR/NIT,NCIT,RMSQU,DELQU,RATMAX,NEIGHB,NFPAR,
     *  MAXBND,XDEL,NFREE,QCOND,RMS,CHISQ,DELMX
         INTEGER NIT,NCIT,NEIGHB,NFPAR,MAXBND,NFREE,QCOND
         REAL RMSQU,DELQU,RATMAX,XDEL,RMS,CHISQ,DELMX
      COMMON/FIXCOM/FIXED,UNFIXD,FIXARR(3,1)
         BYTE FIXED,UNFIXD,FIXARR
      COMMON/STDCOM/NOSTD,MAXSTD,LOGSTD,STD(1)
         INTEGER NOSTD,MAXSTD
         LOGICAL LOGSTD
         REAL STD
      COMMON/WRKSPC/I,J,FRACT,TOTP,TOTD,YMEAN,
     *   PERROR,PDIFF,MW,XT,SUM(1)
         INTEGER I,J
         REAL FRACT,TOTP,TOTD,YMEAN,PERROR,PDIFF,MW,XT,SUM
C
      WRITE(LFD,1010) TITLE,NOPTS,NPAR,NCRVS,NFPAR
      IF(NOSTD.LE.0) GOTO 2
      WRITE(LFD,1012) STDFIL
2     WRITE(LFD,1014)
      IF(NOSTD.GT.0) GOTO 5
      WRITE(LFD,1016)
      GOTO 10
5     WRITE(LFD,1018)
10    TOTP = 0.
      DO 15 I=1,NCRVS
         SUM(I)=2.*P(2,I)*P(3,I)
         TOTP = TOTP + SUM(I)
15    CONTINUE
      FRACT = 0.
      DO 20 I=1,NCRVS
         IF(TOTP.NE.0.) FRACT=SUM(I)/TOTP
         WRITE(LFD,1020)
     *    I,(P(J,I),FIXARR(J,I),J=1,3),SUM(I),FRACT
         IF(NOSTD.GT.0) GOTO 17
C NO STANDARDS IN USE, PUT OUT END OF LINE.
         WRITE(LFD,1016)
         GOTO 20
C CALCULATE AND OUTPUT MOLECULAR WEIGHT.
17       MW = 0.
         XT = 1.
         DO 18 J=1,NOSTD
            MW = MW + STD(J)*XT
            XT = XT*P(1,I)
18       CONTINUE
         IF(LOGSTD) MW = 10.**MW
         WRITE(LFD,1022) MW
20    CONTINUE
C FIND MEAN OF DATA.
      TOTD = 0.
      DO 30 I=1,NOPTS
         TOTD = TOTD + FLOAT(IY(I))
30    CONTINUE
      YMEAN = TOTD/NOPTS
C CORRECT TOTAL AREA FOR BOUNDARIES AND X INCREMENT.
      TOTD = (TOTD - FLOAT( (IY(1)+IY(NOPTS))/2) ) * XDEL
      PERROR = 100.*RMS/YMEAN
      PDIFF = 100.*ABS((TOTP-TOTD)/TOTD)
      WRITE(LFD,1030)
     *TOTP,TOTD,PDIFF,RMS,CHISQ,NFREE,YMEAN,PERROR
      IF(RATMAX.NE.0.) WRITE(LFD,1032) RATMAX
      IF(QCOND.EQ.0) GOTO 200
      WRITE(LFD,1040)
      GOTO(110,120,130,140,150) QCOND
110   WRITE(LFD,1050) NCIT
      GOTO 200
120   WRITE(LFD,1060) DELQU
      GOTO 200
130   WRITE(LFD,1070) RMSQU
      GOTO 200
140   WRITE(LFD,1080)
      GOTO 200
150   WRITE(LFD,1090) RATMAX
200   RETURN
1010  FORMAT('1'///1X,80A1,//1X,I5,' points',I3,' parameters',
     * I3,' curves,  Band matrix',I3,' parameters wide'/)
1012  FORMAT(' Using MW standards file ',16A1/)
1014  FORMAT(7X,'Xmax',10x,'Ymax',9x,'HalfW',10x,'Area',
     *6X,'Fraction'$)
1016  FORMAT('+',1X/)
1018  FORMAT('+',4X,'Mol.Weight'/)
1020  FORMAT(I4,3(G12.5,A1,1X),2G12.5$)
1022  FORMAT('+',G12.5/)
1030  FORMAT('0Area(fit) =',G15.8,' Area(data) =',
     * G15.8,' %diff =',F7.3/'0RMS =',G14.7,8X,
```

```
     * 'CHISQ =',G14.7, 7X,'NDF =',I4/
     * '0Ymean =',G15.8,5X,'%E(RMS/Ymean) =',F7.3//)
1032 FORMAT(' Maximum ratio of HW to YMAX set to ',g14.7//)
1040 FORMAT('$Fitting terminated with ')
1050 FORMAT('+limit of',I4,' CITs surpassed'///)
1060 FORMAT('+Max.Rel.Change <= ',G11.4///)
1070 FORMAT('+RMS <= ',G11.4///)
1080 FORMAT('+equations singular'///)
1090 FORMAT('+maximum HW/Ymax ratio > ',G11.4///)
     END

C-------------------------------------------------------------
C GELPLT.FOR
C-------------------------------------------------------------
     SUBROUTINE GELPLT(CORE,MAXCOR)
C V01-01        28-JUN-79
C
C R.F.MURPHY    CREATED 17-DEC-78
C
C PLOT DATA AND FIT FOR GELFIT USING TTPLOT.
C
     INTEGER MAXCOR
     REAL CORE(MAXCOR)
     COMMON/LOGDEV/LRD,LWD,LFD
        INTEGER LRD,LWD,LFD
     COMMON/DATA/TITLE(80),XMIN,XMAX,YMIN,YMAX,NOPTS
        BYTE TITLE
        INTEGER NOPTS
        REAL XMIN,XMAX,YMIN,YMAX
     COMMON/CURVES/NOCURV,NOPAR
        INTEGER NOCURV,NOPAR
     COMMON/PARCAL/PARCAL
        LOGICAL PARCAL
     COMMON/PLOTCM/XPMIN,XPMAX,YPMIN,YPMAX,NXPLOT,
     *  NYPLOT,NYPERX,LOGPLT,CRTPLT
        LOGICAL*1 LOGPLT,CRTPLT
        INTEGER NXPLOT,NYPLOT,NYPERX
        REAL XPMIN,XPMAX,YPMIN,YPMAX
C
     LOGICAL*1 FPFLAG
     EXTERNAL XFUNCT,YFUNCT,GELFUN
C
     IF(NXPLOT/4.LT.MAXCOR) GOTO 1
     CALL ERROR('F','not enough core for plot.')
     RETURN
1    IF(CRTPLT) GOTO 41
     WRITE(LFD,1000) TITLE
     GOTO 42
41   WRITE(LFD,1005) TITLE
C INDICATE JUST DATA PLOT.
42   IF(NOPAR.LE.0) NYPERX = 0
     PARCAL = .FALSE.
     CALL TTPLOT(LFD,XFUNCT,YFUNCT,NOPTS,GELFUN,
     *CRTPLT,XPMIN,XPMAX,YPMIN,YPMAX,NXPLOT,NYPLOT,
     *NYPERX,CORE,4*MAXCOR)
     RETURN
1000 FORMAT('1'///1X,80A1//)
1005 FORMAT('1',80A1/)
     END
     REAL FUNCTION XFUNCT(I)
C
     COMMON/DATA/TITLE(80),XMIN
        BYTE TITLE
        REAL XMIN
     COMMON/FITPAR/JUNK(11),XDEL
        INTEGER JUNK
        REAL XDEL
C
     XFUNCT = XMIN + (I-1)*XDEL
     RETURN
     END
     REAL FUNCTION YFUNCT(I)
C
     COMMON/DATA/TITLE(80),MINMAX(4),NOPTS,MAXPTS,IY(1)
        BYTE TITLE
        INTEGER NOPTS,MAXPTS
        REAL MINMAX
C
     YFUNCT = FLOAT(IY(I))
     RETURN
     END

C-------------------------------------------------------------
C TTPLOT.FOR
C-------------------------------------------------------------
     SUBROUTINE TTPLOT(LLD,X,Y,NOPTS,FUNCTN,CRTPLT,
     *XMIN,XMAX,YMIN,YMAX,NX,NY,NYPERX,WRK,MAXWRK)
C V03-01        28-JUN-79
C
C R.F.MURPHY    CREATED 31-OCT-78
C
C PLOTS X&Y DATA FUNCTIONS AND/OR A FIT FUNCTION ON A TERMINAL.
C
C INPUTS
C
C       LLD     LOGICAL UNIT TO SEND PLOT TO
C       X       X DATA FUNCTION
C       Y       Y DATA FUNCTION
C       NOPTS   # OF DATA POINTS
C       FUNCTN  FIT FUNCTION
C       CRTPLT  FLAG (.TRUE. IMPLIES USE CONDENSED FORMAT)
C       XMIN,XMAX,YMIN,YMAX     FIT PLOT LIMITS
C       NX      # OF COLUMNS IN PLOT
C       NY      # OF ROWS IN PLOT
C       NYPERX  # OF FUNCTION VALUES TO PLOT PER X
C       WRK     TEMPORARY WORK SPACE
C       MAXWRK  # OF DOUBLE-WORDS OF SPACE IN WRK
C
C OUTPUTS
C
C       WRK(1,1)        # OF DATA VALUES OFF PLOT SURFACE
C
     LOGICAL*1 CRTPLT
     BYTE WRK(NX,NY)
     INTEGER LLD,NOPTS,NX,NY,NYPERX,MAXWRK
     REAL X,Y,FUNCTN,XMIN,XMAX,YMIN,YMAX
C
     BYTE ERR
     INTEGER I,IX,IY
     REAL XINC,XINC2,YINC,XT,YT
```

```
      INTEGER PLTPNT
C
C DEFINE PLOT POINTER FUNCTION.
      PLTPNT(A1,A2,A3) = INT( 1.5 + (A1-A2)/A3 )
C
C MAKE SURE THERE'S SOMETHING TO PLOT.
      ERR = 0
      IF(NOPTS.LE.0.AND.NYPERX.LE.0) GOTO 990
C CALCULATE X & Y PLOT INCREMENTS
      XINC = (XMAX-XMIN)/(NX-1)
      IF(NYPERX.GT.0) XINC2 = XINC/NYPERX
      YINC = (YMAX-YMIN)/(NY-1)
      MAXNY = MAXWRK/NX
      IF(MAXNY.LE.0) GOTO 990
C PUT OUT TOP X AXIS.
      CALL XAXIS(LLD,XMIN,XINC,NX,CRTPLT,'T')
C INIT FOR PLOT OUTPUT.
C GENERATE AS MANY LINES AT ONCE AS POSSIBLE.
      NY1 = 0
      NY2 = MINO(NY,MAXNY)
C FILL PAGE IMAGE ARRAY WITH BLANKS
4     DO 10 IX=1,NX
         DO 5 IY=1,NY2-NY1
            WRK(IX,IY) = ' '
5        CONTINUE
10    CONTINUE
C PUT FUNCTION IMAGE INTO ARRAY
      IF(NYPERX.LE.0) GOTO 22
      XT = XMIN
      DO 20 IX=1,NX
         DO 15 I=1,NYPERX
            YT = FUNCTN(XT)
            XT = XT + XINC2
            IY = PLTPNT(YMAX,YT,YINC)
            IF(IY.LE.NY1.OR.IY.GT.NY2) GOTO 15
            IY = IY - NY1
            WRK(IX,IY) = '*'
15       CONTINUE
20    CONTINUE
C PLOT DATA OVER FUNCTION, USING NUMBERS TO
C INDICATE MULTIPLE POINTS.
22    IF(NOPTS.LE.0) GOTO 41
      DO 40 I=1,NOPTS
         IX = PLTPNT(X(I),XMIN,XINC)
         IF(IX.LE.0.OR.IX.GT.NX) GOTO 38
         IY = PLTPNT(YMAX,Y(I),YINC)
         IF(IY.LE.NY1.OR.IY.GT.NY2) GOTO 38
         IY = IY - NY1
         IF(WRK(IX,IY).EQ.'*'.OR.WRK(IX,IY).EQ.' ') GOTO 30
C POINT IS ALREADY THERE. BUMP COUNTER.
         IF(WRK(IX,IY).GE.'0'.OR.WRK(IX,IY).LT.'9') GOTO 24
         WRK(IX,IY) = WRK(IX,IY) + 1
         GOTO 40
C USE + FOR MORE THAN TEN SUPERIMPOSED POINTS
24       WRK(IX,IY) = '+'
         GOTO 40
C NO POINT YET. START WITH 0.
30       WRK(IX,IY) = '0'
         GOTO 40
38       ERR = ERR + 1
40    CONTINUE
C PUT OUT PAGE IMAGE AND Y AXES
41    YT = YMAX - YINC*NY1
      DO 70 IY=1,NY2-NY1
         IF(MOD(IY+NY1,10).NE.1) GOTO 50
         IF(CRTPLT) GOTO 42
         WRITE(LLD,1020) YT
         GOTO 43
42       WRITE(LLD,1040) YT
43       WRITE(LLD,1022) (WRK(IX,IY),IX=1,NX)
         IF(CRTPLT) GOTO 44
         WRITE(LLD,1024) YT
         GOTO 60
44       WRITE(LLD,1044)
         GOTO 60
50       IF(CRTPLT) GOTO 51
         WRITE(LLD,1030)
         GOTO 52
51       WRITE(LLD,1050)
52       WRITE(LLD,1022) (WRK(IX,IY),IX=1,NX)
         WRITE(LLD,1034)
60       YT = YT - YINC
70    CONTINUE
      IF(NY2.GE.NY) GOTO 72
      NY1 = NY2
      NY2 = MINO(NY,NY2+MAXNY)
      GOTO 4
C PUT OUT BOTTOM XAXIS.
72    CALL XAXIS(LLD,XMIN,XINC,NX,CRTPLT,'B')
990   WRK(1,1) = ERR
      RETURN
C
C OUTPUT FORMATS
C
C FORMAT #'S ENDING IN A 0 START AN OUTPUT LINE,
C THOSE ENDING IN A 2 CONTINUE IT, AND THOSE
C ENDING IN A 4 END IT.
C
1020  FORMAT('+',E12.5,' -'$)
1022  FORMAT('+',A1$)
1024  FORMAT('+- ',E12.5/)
1030  FORMAT('+',13X,'|'$)
1034  FORMAT('+|'/)
1040  FORMAT('+',E9.2,'-'$)
1044  FORMAT('+-'/)
1050  FORMAT('+',9X,'|'$)
      END
      SUBROUTINE XAXIS(LLD,XMIN,XINC,NX,CRTPLT,TOPBOT)
C
      LOGICAL*1 CRTPLT
      BYTE TOPBOT
      INTEGER LLD,NX
      REAL XMIN,XINC
      INTEGER I,IT,J
      REAL XT
C
      IF(TOPBOT.EQ.'B') GOTO 30
10    IF(CRTPLT) GOTO 12
      WRITE(LLD,1000)
      GOTO 14
```

```
12      WRITE(LLD,1060)
14      XT = XMIN
        DO 20 I=1,NX-20,20
           WRITE(LLD,1002) XT
           XT = XT + 20.*XINC
20      CONTINUE
        WRITE(LLD,1004) XT
        IF(TOPBOT.EQ.'B') RETURN
30      IF(CRTPLT) GOTO 32
        WRITE(LLD,1010)
        GOTO 34
32      WRITE(LLD,1070)
34      DO 40 I=1,NX
           IT = I
           J = '-'
           IF(MOD(IT,10).EQ.1) J = '|'
           WRITE(LLD,1012) J
40      CONTINUE
        WRITE(LLD,1014)
        IF(TOPBOT.EQ.'T') RETURN
        GOTO 10
C
C OUTPUT FORMATS
C
1000    FORMAT('+',7X$)
1002    FORMAT('+',E12.5,8X$)
1004    FORMAT('+',E12.5/)
1010    FORMAT('+',14X$)
1012    FORMAT('+',A1$)
1014    FORMAT('+'/)
1060    FORMAT('+',3X$)
1070    FORMAT('+',10X$)
        END

C-------------------------------------------------------------
C SORTG.FOR
C-------------------------------------------------------------
        SUBROUTINE SORTG(A,NROW,PFLAG)
C V00
C
C R.F.MURPHY    17-JUN-78
C
C SORT A LINEAR ARRAY AS IF IT WHERE DIMENSIONED
C NROW*3 SO THAT THE FIRST ROW IS EITHER IN INCREASING ORDER
C OR IN ORDER OF 'XMIN' (XMAX-3*HW)
C
        INTEGER NROW, PFLAG
        REAL A(1)
        LOGICAL*1 DONE
C
        DATA NCOL/3/
C
        IF(NROW.LE.NCOL) GOTO 30
        NCOL1=NCOL-1
        DO 20 J=1,NROW-NCOL,NCOL
        L=NROW-J-NCOL1
        DONE=.TRUE.
        DO 10 K=1,L,NCOL
        IF(PFLAG) GOTO 1
        IF(A(K).LE.A(K+NCOL)) GOTO 10
        GOTO 2
1       IF((A(K)-3.*A(K+2)).LE.
        *(A(K+NCOL)-3.*A(K+NCOL+2))) GOTO 10
2       DO 5 I=0,NCOL1
        I1=K+I
        ATEM=A(I1)
        A(I1)=A(I1+NCOL)
        A(I1+NCOL)=ATEM
5       CONTINUE
        DONE=.FALSE.
10      CONTINUE
        IF(DONE) GOTO 30
20      CONTINUE
30      RETURN
        END
```

APPENDIX B

GELFIT fitting session example

In the example which follows, the operator first requested a command summary. If more information had been needed about a particular command, the user could have typed HELP/LONG to obtain it. The file GRR209.DAT was then read in for use in the standards calculation, and the LOG molecular weight calculation suppressed. (Normally, standards data are expressed as log molecular weight versus relative mobility.) The data for a metrizamide gradient (Chapter 5, Figure 4B) were then read from a file which had been created from the output of a scintillation counter by a data collection program. The program was then requested to estimate peak positions (the GUESS command), with the constraint that each peak be greater than 1000 counts per minute. The guesses were refined by fitting, with the maximum number of iterations set to 20. As the fitting proceeded, the values of important parameters were output. A listing of the final parameter values was printed, which includes the calculated "Mol.Weight" (in this case the peak density). A rough plot was generated on the terminal, the fit values stored in a new copy of the input file, and the program exited. The entire process took a little over a minute.

```
 .DATE
 ----
2-Jul-79

.TIME
 ----
16:30:18

.RU GELFIT
 ---------
GELFIT V04-10B
$HELP
 ----

GELFIT Command Summary:

The minimum required letters of each command or option
are shown. All commands are terminated by the return key.
<list> indicates a set of optional arguments.

A <list>      ADD more curves
B k           fit k BANDS wide
C             CLOSE output file
D i,j         DELETE curves i-j
FIT k,l       FIT k its with l cits
FIX i,j       FIX curve i param j
GE i,x1,x2,e GENERATE data from params
GU yfil,zfil GUESS curve parameters
H             type this HELP summary
  /L          type SY:GELFIT.HLP
I infile      set INPUT file
  /A          ADD to previous data
K m,n         KILL points m-n
L m,n         LIST points m-n
M i,j         MODIFY curves i-j
N             set NEW curves
O outfil      set OUTPUT filename
  /LI              set LIST filename
  /PL              set PLOT filename
PL <list>     PLOT data and fit
  /C          plot in CRT format
  /LP         plot in LP format
PO            add more POINTS
Q a,b         set QUIT criteria
S stdfil      set STANDARDS
T i,j         TYPE curves i-j
U i,j         UNFIX curve i param j
W a           set max relative WIDTH
X a,b         XSCALE (x+a)*b
Y a,b         YSCALE (y+a)*b
$STANDARDS/NOLOG GRR209
 ----------------------
$INP GRH202
 ----------
$GUESS 1000
 ----------
$FIT 20
```

ET(SEC)	RMS	MAX.REL.CHANGE	IN	P#	NEW VALUE	CIT#
1.	447.9610	0.9626734	(	6)	0.4717965E-01	0
2.	370.2311	1.174351	(	6)	0.1025851	0
1.	321.2720	1.110425	(	6)	0.2164982	0
2.	286.9051	0.1601599	(	2)	3801.431	0
2.	286.1761	0.2930580E-01	(	5)	1010.346	0
2.	286.1088	0.9214476E-02	(	6)	0.2216258	0
1.	286.0698	0.4176276E-02	(	6)	0.2207003	0
7.	286.0692	0.8147353E-03	(	6)	0.2205205	9
2.	286.0439	0.6298451E-03	(	6)	0.2203816	1
3.	286.0317	0.8411340E-04	(	6)	0.2203630	2

GRH202 $03 GELFIT V04-10B 02-JUL-79 PRF62.2-2 1'THY (SP.04,.112)

53 points 6 parameters 2 curves, Band matrix 6 parameters wide

Using MW standards file GRR209.DAT

	Xmax	Ymax	HalfW	Area	Fraction	Mol.Weight
1	0.46609	3785.3	0.29404E-01	222.61	0.33224	1.1775
2	0.59149	1015.2	0.22036	447.41	0.66776	1.1400

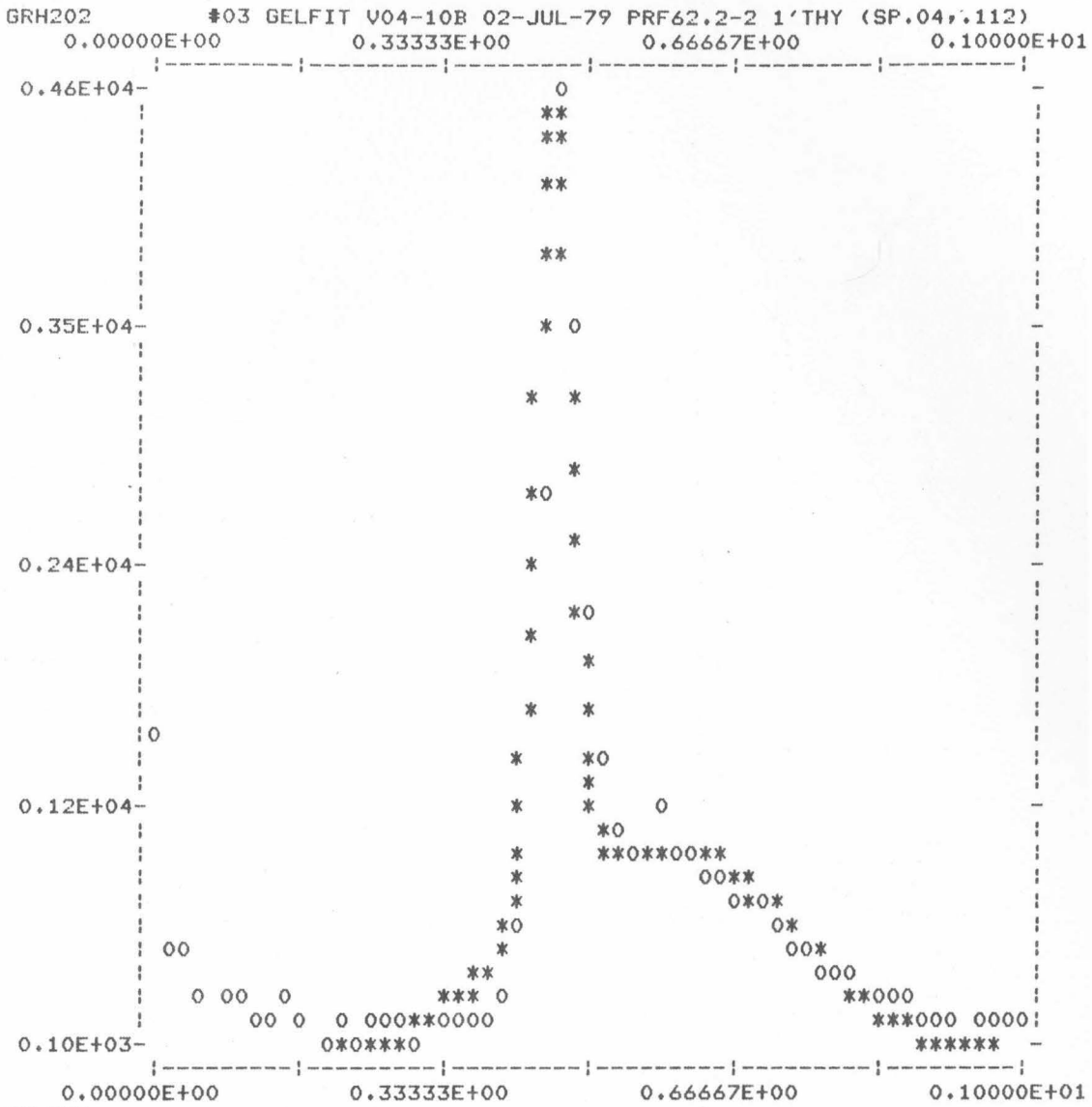
Area(fit) = 670.01251 Area(data) = 675.00000 %diff = 0.739

RMS = 286.0317 CHISQ = 2882142. NDF = 47

Ymean = 678.30188 %E(RMS/Ymean) = 42.169

Fitting terminated with Max.Rel.Change <= 0.1000E-03

$PLOT




```
$CLOSE
 -----
$EXIT
 ----

.TIME
 ----
16:31:19
```