STOICHIOMETRY OF CHROMATIN PROTEINS

W. T. Garrard, W. R. Pearson, S. K. Wake, and James Bonner

Division of Biology
California Institute of Technology
Pasadena, California 91109

Received February 22,1974

SUMMARY: The stoichiometry of rat liver chromatin proteins has been estimated by quantitative disc electrophoresis. There are about 1.3 x 10^8 histone and 2.4 x 10^7 nonhistone polypeptide molecules per haploid genome, an average of one such molecule for every 21 and 117 base pairs of DNA, respectively. Nonhistone proteins have been separated into 115 polypeptide size classes. The total number of molecules represented in distinct size classes ranges from 4.2 x 10^3 (limits of sensitivity) to 1.7 x 10^6 per haploid genome, with a mean of 2.1 x 10^5.

In an effort to understand how two major components of chromatin, the histone and nonhistone proteins, might contribute to chromatin structure and function, we have estimated the number of polypeptide copies of these species per haploid genome.

METHODS: Previously published techniques, with minor modifications (1), were used for preparation of sheared rat liver chromatin (2), chemical analyses (2), and isolation of chromatin proteins (3). The chemical composition of chromatin was 1.00:1.10 ± 0.03:0.60 ± 0.07:0.03 ± 0.005 (DNA:acid soluble protein:nonhistone protein:RNA). Histone (4) and nonhistone (5) gels were stained with amido black and Coomassie brilliant blue R-250, respectively, destained, and scanned at 600 nm with a Gilford 2000 spectrophotometer equipped with a 0.05 mm slit (1). The amount of sample applied to gels was adjusted so that densitometry of stained bands was within the linear portion of absorbance measurements. A computer program was used to fit Gaussian curves to visualized bands. The relative mass proportions of individual bands were determined on the assumption that mass was proportional to the amount of stain bound. This assumption holds true for the histones (6), while deviations of less than twofold can be expected for the nonhistone chromatin proteins (7). The rat haploid genome size was assumed to be 1.8 x 10^{12} daltons or 2.8 x 10^9 base pairs (8). Stoichiometry was estimated using the following relationship: number of molecules of a given component per haploid genome = (ng given component per mg DNA) (molecular weight component)$^{-1}$ (1.8 x 10^6 mg DNA per nmole genome).

RESULTS AND DISCUSSION: Table I summarizes calculations of stoichiometry for the five major histone species of rat liver chromatin. There are approximately 1.3 to 1.4 x 10^8 histone molecules per haploid genome, an average of one molecule for every 21 base pairs of DNA. This value closely agrees with an estimate made by Phillips for the histones of calf thymus chromatin (11).

TABLE I

STOICHIOMETRY OF THE HISTONES OF RAT LIVER CHROMATIN

A. Published Properties (9,10)

Histone Species	Molecular Weight x 10^{-3}	Basic Amino Acids Per Molecule[a] N-Terminal Portion	Middle Portion	C-Terminal Portion	Total
I	21.0	12	10	43	65
IIb1	14.0	14	7	9	30
IIb2	13.8	15	6	10	31
III	15.3	18	-	15	33
IV	11.3	17	-	10	27

B. Rat Liver Chromatin Composition

Histone Species	% Total Mass	μg Protein[b] / mg DNA	n moles / mg DNA	Molecules x 10^{-6}[c] / Haploid Genome	Basic Groups x 10^{-6} / Haploid Genome Most Basic Half	Total
I	10.5	110	5.2	9.4	403	610
IIb1	14.9	157	11.2	20.2	282	605
IIb2	25.7	270	19.6	35.2	527	1090
III	29.7	312	20.4	36.8	657	1204
IV	19.2	202	17.9	32.2	550	870
			Totals	133.8	2419	4379

---

[a] Basic amino acids include arginine, lysine, and histidine and their derivatives.

[b] The mass ratio of these histones per mg chromatin DNA is 1.05 mg.

[c] The rat haploid genome contains 2.8 x 10^9 base pairs or 5.6 x 10^9 phosphate groups, and has a molecular weight of 1.8 x 10^{12} (8).

Table I also shows the net number of basic amino acids contributed by the major histone species. If all such amino acids could interact electrostatically with DNA phosphate groups (e.g., no modifications or steric limitations) then about 78% of the DNA phosphates would be neutralized. However, reports in the literature on the ability of chromatin to bind metals (12), dyes (13), and polylysine (14) suggest that only about 50-60% of such phosphates are neutralized. Ignoring possible contributions by nonhistone proteins, this implies that histones are bound to chromatin such that most (64-77%) but not all of their basic amino acids are involved in this type of electrostatic interaction. This numerology sets limits on what portions of each histone molecule might be bound to DNA. Interaction of only the more basic portion of each histone molecule would neutralize too few (43%) of the DNA phosphates (Table I). If in addition, the less basic C-terminal portions of histones IIb1 and IIb2 were to interact with DNA, but not the basic groups in the hydrophobic centers of these molecules, then the observed charge neutralization in chromatin would be met (53%). Model building studies on the histone IV-DNA complex at criteria similar to those dictated by our calculations have shown that certain non-basic portions of the molecule would have to loop out to allow enough space for the basic-portions to mediate electrostatic interactions (11). It seems clear that the above estimates illustrate that the majority of chromatin DNA must be covered with closely neighboring histone molecules, a conclusion reached earlier by Phillips (11).

Unlike the histones, the nonhistone chromatin proteins of rat liver are highly heterogeneous (Fig. 1). This heterogeneity is not due to gross contamination by extrachromosomal proteins; results of a radioactive mixing experiment suggest that over 86% of such protein, as prepared by our method, originates from chromosomal material (1). Furthermore, gel patterns of different preparations of nonhistone proteins are nearly identical (1). A priori, this complex set of polypeptides can be supposed to include structural proteins, enzymes of chromosomal metabolism, and regulatory elements.

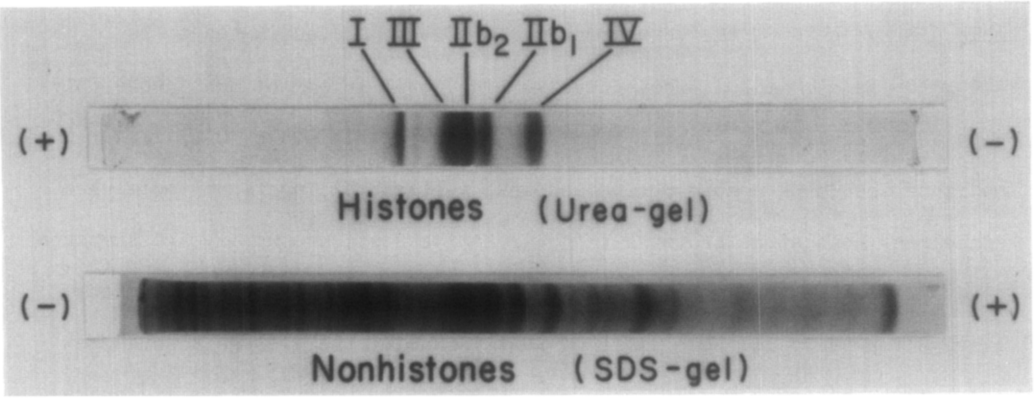

FIGURE 1. Acrylamide gel patterns of rat liver chromatin proteins. Samples of 25 μg histone and 100 μg nonhistone protein were separated on 7 and 10 cm gels, respectively. Migration was from left to right in both cases.

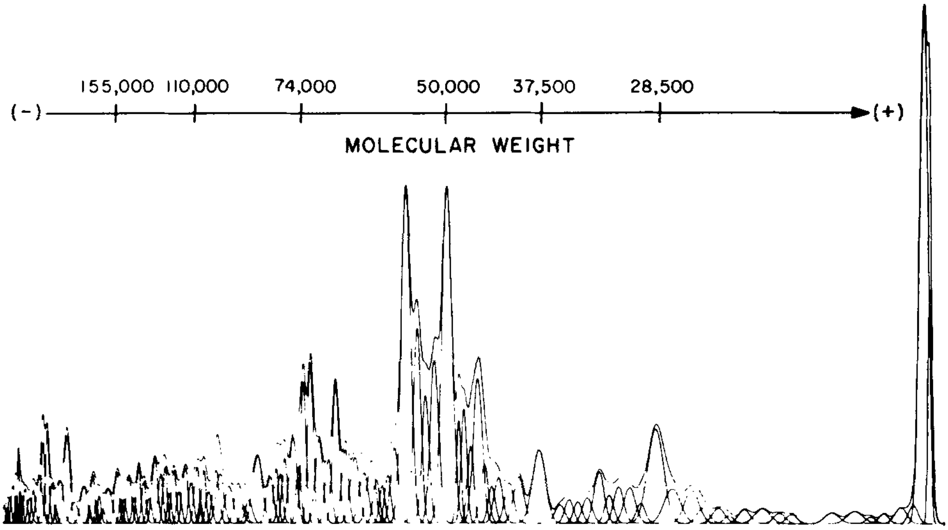

FIGURE 2. Composite gel scan of rat liver nonhistone chromatin polypeptides. Each of the 115 visibly identifiable bands was fit to a Gaussian curve. Samples of 50 and 200 μg protein were separated on 22 sodium dodecyl sulfate gels. Profiles are within the linearity of absorbance measurements. Molecular weights were assigned from a standard curve (1).

Figure 2 shows a gel scan of nonhistone polypeptides in which each of the 115 visibly identifiable bands has been fit with a Gaussian curve. At present, the molecular complexities of these bands is unknown; some could be composed of unique polypeptides, or families of molecules which are evolutionally related, or mixtures of components. The number- and weight-average molecular weights of the total nonhistone mixture are estimated to be 4.5×10^4 and 5.0×10^4, respectively. The total number of nonhistone polypeptide molecules per haploid genome is approximately 2.4×10^7. This estimate should be considered an absolute maximum for functional molecules since many nonhistone proteins in their native state presumably are composed of subunits. At this maximum, there would be one nonhistone molecule on average for every 117 base pairs of DNA. Assuming the average nonhistone molecule to be roughly spherical, with a molecular weight of 5×10^4 and a density of 1.3 g per cm^3, such a molecule would have a diameter of approximately 50 Å, and thus could cover about 15 base pairs of DNA.

The average number of molecules per haploid genome represented in an individual nonhistone band is estimated to be 2.1×10^5 ($2.4 \times 10^7/115$), and Figure 3 summarizes detailed calculations for 98 of the Gaussian components (bands of $>2 \times 10^5$ and $<1.8 \times 10^4$ M.W. are not included because their size could not be assigned precisely). The average frequency for a component of this subset is 1.8×10^5 molecules per haploid genome, with a range from

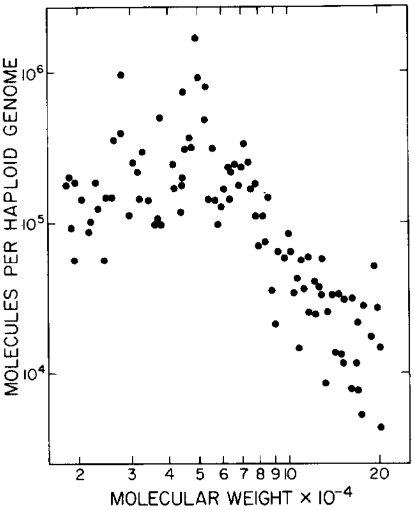

FIGURE 3. Frequences of molecules of rat liver nonhistone polypeptide size classes. Points are derived from Gaussian curves shown in Figure 2. Note that data are plotted on a log-log scale.

4.2×10^3 (limits of sensitivity) to 1.7×10^6, or one molecule of these latter classes on average for every 6.7×10^5 to 1.6×10^3 base pairs of DNA. While the molecular complexities of the nonhistone components resolved by this study are unknown, it seems clear that at least several bands are composed of individual species. Fold-purification data for RNA polymerase (15), and chromatin-bound protease (16) suggest that these polypeptides are present at 5 to 10×10^3, and 43×10^3 copies per haploid genome, respectively; proteins present near these frequencies at the appropriate molecular weights can be found in Figure 3.

Nonhistone proteins presumably are bound to chromatin by a variety of means (interactions with DNA, histones, themselves, and RNA). Since specificity of interaction of certain species might be determined by the repetitive DNA sequences of chromatin, it is of interest to compare our estimates with the sequence composition of the rat genome. About half of the Gaussian components are present in numbers high enough ($>10^5$ molecules per haploid genome) to interact with the 10% of the nuclear DNA which is highly repetitious (17). In contrast, bands with frequencies similar to those of moderately repetitive DNA (20% of rat DNA, ref. 17) are present near the detection limit of this study (4×10^3) and constitute a trace of the total. Thus, our estimates of the frequencies of nonhistones do not correlate well with DNA sequence composition. Some of the bands, particularly those of high frequency, may be composed of a heterogeneous population of similar polypeptides. These could provide proteins in number similar to moderately repetitive DNA sequence frequencies for tens to hundreds of different sequence families. With regard to models of gene regulation (18), this may be the simplest interpretation of the reported differences in the electrophoretic patterns of major nonhistones between tissues (19), and during development (20) and the cell cycle (21). On the other hand, certain regulatory proteins may need to be at high concentrations in order to saturate their specific sites amongst an enormous amount of competitor DNA (Riggs and Linn, personal communication). Current techniques of protein separation and analysis are available to afford information pertinent to these possibilities.

In conclusion, the estimates of this study reveal that the majority of chromatin DNA is covered with closely spaced histone molecules, which exceed the nonhistones in number by at least 5-fold. The frequencies of nonhistone molecules present in distinct size classes raises an important question regarding models of chromatin structure, DNA sequence composition, and gene regulation: What is the absolute limit of heterogeneity of the nonhistone proteins? Finally, it should be noted that chromatin fractionation results suggest that transcribed regions are enriched in certain nonhistones, and impoverished in histones (17). Therefore, the packing of these proteins within various genomal loci is clearly more crowded than our calculations suggest.

ACKNOWLEDGEMENTS. This work was supported in part by U.S. Public Health Service Grant GM 13762, by Damon Runyon Memorial Fund for Cancer Research Grant DRF-755 (W. T. G.), by U. S. Public Health Service Training Grant GM-86 (W. R. P.), and by The California Foundation for Biochemical Research (S. K. W.)

REFERENCES

1. Garrard, W. T. and Bonner, J., submitted to J. Biol. Chem.
2. Bonner, J., Chalkley, G. R., Dahmus, M., Fambrough, D., Fujimura, F., Huang, R. C., Huberman, J., Jensen, R., Marushige, K., Ohlenbush, H., Olivera, B. M., and Widholm, J. (1968) Methods Enzymol. 12B, 3-65.
3. Elgin, S. C. R. and Bonner, J. (1970) Biochemistry 9, 4440-4447.
4. Panyim, S. and Chalkley, R. (1969) Arch. Biochem. Biophys. 130, 337-346.
5. King, J. and Laemmli, V. K. (1971) J. Mol. Biol. 62, 465-477.
6. Fambrough, D. M., Fujimura, F., and Bonner, J. (1968) Biochemistry 7, 575-584.
7. Fazekas de St. Groth, S., Webster, R. G., and Datyner, A. (1963) Biochim. Biophys. Acta 71, 377-391.
8. Britten, R. J. and Davidson, E. H. (1971) Quart. Rev. Biol. 46, 111-138.
9. Bustin, M., Rall, S. C., Stellwagen, R. H., and Cole, R. D. (1969) Science 163, 391-393.
10. Elgin, S. C. R. and Bonner, J. (1973) The Biochemistry of Gene Expression in Higher Organisms, pp. 142-163, Australia and New Zealand Book Co., Pty. Ltd., Artarmon, Australia.
11. Phillips, D. M. P. (1971) Histones and Nucleohistones, pp. 47-83, Plenum Press, London, England.
12. Schmidt, G., Cashion, P. J., Shigetaka, S., Joseph, J. P., Demarco, P., and Cohen, M. B. (1972) Arch. Biochem. Biophys. 149, 513-527.

13. Itzhaki, R. F. (1971) Biochem. J. 122, 583-592.
14. Itzhaki, R. F. and Cooper, H. K. (1973) J. Mol. Biol. 75, 119-128.
15. Weaver, R. F., Blatti, S. P., and Rutter, W. J. (1971) Proc. Nat. Acad. Sci. USA 68, 2994-2999.
16. Chong, M.-T., Garrard, W. T., and Bonner, J. (In preparation).
17. Bonner, J., Garrard, W. T., Gottesfeld, J., Holmes, D. S., Sevall, J. S., and Wilkes, M. (1973) Cold Spring Harbor Symp. Quant. Biol. 38, In press.
18. Britten, R. J. and Davidson, E. H. (1969) Science 165, 349-357.
19. Wu, F. C., Elgin, S. C. R., and Hood, L. E. (1973) Biochemistry 12, 2792-2797.
20. Vidali, G., Boffa, L. C., Littau, V. C., Allfrey, K. M., and Allfrey, V. G. (1973) J. Biol. Chem. 248, 4065-4068.
21. Bhorjee, J. S. and Pederson, T. (1972) Proc. Nat. Acad. Sci. USA 69, 3345-3349.