Relationship between Protein and DNA Structure in Calf Thymus Chromatin. II. Conformational Aspects†

Sue Hanlon,* Roger S. Johnson, and Aurelia Chan

ABSTRACT: Analysis of the circular dichroism properties, below 240 nm, of the residual complexes of calf thymus chromatin described in paper I (Hanlon, S., Johnson, R. S., and Chan, A. (1974), *Biochemistry 13*, 3963) yields conformational information about the proteins which are dissociated by the reagents employed as well as those which remain in complex at a given value of the protein content, W (gram of protein/gram of DNA). As the various compositional classes of proteins are dissociated by NaCl and by sodium deoxycholate, it appears that three different conformational classes are removed. When the W value of intact chromatin is reduced from 1.7 to 0.85 g/g, a class of proteins consisting of 13% α helix, 21% β structure, and 66% random coil forms is removed. Between a W of 0.85 and 0.33 g/g, a class of moderately high helix content proteins (60% α and 40% random coil) is dissociated. Removal of this latter class parallels almost exactly the loss of C character with linear increases in the per cent B conformation present in the DNA constituent. Below a W of 0.33, where the per cent B character is maximal (100%), the proteins removed (which are essentially nonhistone proteins) have negligible helix content. As was observed for the DNA conformational changes (paper I), there is no statistical difference in the conformational characteristics of the proteins removed by NaCl and by sodium deoxycholate at W values above 0.3. Thermal melting experiments reveal that although the melting profiles of the residual complexes are more complex and less readily resolvable into the separate transitions seen in the intact controls, the correlation between the fraction of bases melting out in the high melting transition with the fraction of bases in the C conformation, originally observed for the intact controls, is also seen in the case of the residual complexes prepared by any of the dissociation procedures. These data taken together with the observations presented in paper I have added additional details to our original model for calf thymus chromatin (Hanlon, S., Johnson, R. S., Wolf, B., and Chan, A. (1972), *Proc. Nat. Acad. Sci. U. S. 69*, 3263). In the current version, the DNA bases in the C conformation are maintained by a structural package of histones of high average helix content, more or less embedded in the grooves of the DNA helix. Remaining regions of the DNA genome are either directly exposed to solvent or else in complex with more structurally disorganized histones and nonhistone proteins.

In the preceding paper of this series (part I), we posed the question of the existence of a unique class of proteins responsible for the maintenance of the C conformation of the DNA constituent of calf thymus chromatin. In that study, we attempted to answer the question from the standpoint of composition. The results demonstrated that although histone proteins were presumably responsible for the stabilization of that fraction of base pairs in the C conformation, more than one histone compositional class was involved.

It was also apparent, however, that the full complement of histone proteins originally present was not required to maintain the given C character present in the intact controls. An appreciable fraction, approximately 50% of the total weight of histone, could be removed without a significant change in the fraction of base pairs in the C form. Furthermore, there were no clear-cut distinctions in gross composition between these two sets of histones. These facts would imply that either the C character was maintained in intact chromatin by a cooperative action of all histones in complex, or, alternatively, there existed two or more classes of histone proteins whose differences were based on properties other than composition.

In this present study, we wish to explore the latter possibility from the standpoint of conformational differences in the histones associated with the DNA of chromatin. In addition, we also present data which rule out the possibility of a cooperative influence of histone binding on C structure in chromatin.

Experimental Section

Experimental methods, equipment, and precautions observed in this work are identical, except as described below, with that given in paper I, as the data were derived from the same experiments.

The thermal melting profiles were obtained with a Cary Model 14 CMR recording spectrophotometer, equipped with thermostated adaptors. Temperature was maintained by a Haake circulating water bath and monitored by a bridge and thermistor probe assembly manufactured by Yellow Springs Instrument Co. The probe was placed directly in the sample solution, out of the light path, contained in the 1-cm matched quartz cuvet in the sample compartment. The reference cell, maintained at the same temperature as the sample cell, contained all components of the sample solution except the macromolecular solute.

The melting profiles were obtained over the temperature range of 15–97°. Absorbance was monitored at 259 nm continuously and full spectra were generally obtained at temperature intervals of 2–4° in regions of rapidly changing absorbance. The maximum value at 259 nm was recorded at smaller intervals (*ca.* 0.2–0.4°). Five to fifteen minutes were generally allotted for equilibration at the temperature at which full spectra were taken.

The fractions of α, β, and random coil structures of the protein constituents of the chromatin complexes were calculated

† From the Department of Biological Chemistry, College of Medicine, University of Illinois, Chicago, Illinois 60612. *Received November 21, 1973.* This work was supported by a grant from the National Science Foundation (GB 24550). Taken in part from the Ph.D. Thesis of R. S. J.

from the values of $[\theta]^{obsd}$ below 240 nm of the CD[1] spectra of the complexes, as described in the Results section. For these calculations, reference spectra of Greenfield and Fasman (1969) as well as Chen *et al.* (1972) were employed. The latter spectra permit an unambiguous estimate of the fraction of α structures, f_α, present from the value of $[\theta]$ at 222 nm. All other wavelength calculations using these reference spectra, as well as those employing the reference spectra taken from the data of Greenfield and Fasman (1969), required the solution of simultaneous equations. For the latter calculations, we used the linear least-mean-squares averaging procedure described in paper I. In the present instance, the appropriate expression for the sum of the squares of the residuals is

$$R = \sum_{\lambda_i} \{[\theta]_{\lambda_i} - f_\alpha[\theta]_{\lambda_i}^{\alpha} - f_R[\theta]_{\lambda_i}^{R} - (1 - f_\alpha - f_R)[\theta]_{\lambda_i}^{\beta}\}^2 \quad (1)$$

thus forcing the fractions of all components to sum to 1. $[\theta]_{\lambda_i}$ represents one of two protein ellipticity functions described in the Results section. f_α and f_R are fractions of peptides in the α helix and random coil, respectively. $[\theta]^{\alpha}$, $[\theta]^{R}$, and $[\theta]^{\beta}$ are the reference ellipticities of the α helix, β sheet, and random coil, taken from the cited literature sources. The expression for R was then differentiated with respect to f_α and f_R, and the simultaneous equations so obtained were solved for the two coefficients, f_α and f_R. The fraction, f_β, of peptide residues in β structures was obtained from the relationship, $1 = f_\alpha + f_\beta + f_R$.

Results and Discussion

As noted in the previous paper (I), the CD spectra of the residual complexes obtained by removal of protein from intact chromatin transform markedly upon reduction of the protein content below values of 1 g of protein/g of DNA (W). The changes in the positive band have been interpreted in terms of a conversion of base pairs in the C conformation to the B conformation. Changes were also observed in the negative band below 240 nm. It can be demonstrated that these changes are greater than can be accounted for on the basis of a simple reduction of the protein content, *per se*. Transformations in the values of $[\theta]_{222}$ are particularly marked, which suggests that either a loss of helical structure in the protein constituents left in the complex occurs, or, alternatively, there is a marked difference between the helix content of the proteins which are removed and those which remain in the complex as W decreases.

In order to extract information of a quantitative nature about the conformational properties or changes of the protein constituents of these complexes, one must make the same crucial assumptions which have been employed in the analysis of the conformational contributions of the DNA constituent (Hanlon *et al.*, 1972; Johnson *et al.*, 1972). These are, briefly, (1) the secondary structures found in fiber form are reasonable approximations of the forms found in complex under the present conditions, (2) no new forms are generated under the present conditions, and (3) the reference spectra of these various secondary structures obtained by others under different conditions of solvent are suitable for use in the analysis of the conformations in complex. In addition, we must also assume that the base and the peptide chromophores do not optically interact with each other, thereby grossly distorting their respective CD signals in the wavelength regions where their absorption bands overlap.

These assumptions permit us to correct the observed ellipticities, $[\theta]_\lambda^{obsd}$, of the intact and residual complexes for the contribution of the mixed conformation of the DNA base pairs, $([\theta]_\lambda^{B,C})$. The difference, $([\theta]_\lambda^{obsd} - [\theta]_\lambda^{B,C})$, which has been designated $[\theta]_\lambda'$, thus contains the protein contribution to the spectrum. This contribution, however, is on a molar nucleotide basis, whereas conformational analysis of the spectrum requires it to be on a molar peptide basis, $[\theta]_\lambda^{prot.,av}$. The conversion to the latter basis is effected by a normalization factor, and the relationship between $[\theta]_\lambda'$ and $[\theta]_\lambda^{prot.,av}$ is

$$[\theta]_\lambda' = (M_N/M_P)\, W[\theta]_\lambda^{prot.,av} \quad (2)$$

where M_N and M_P are the nucleotide and peptide residue weights, respectively.

If there are discrete conformational classes present among these proteins in any given complex, then

$$W[\theta]_\lambda^{prot.,av} = \sum_i w_i[\theta]_\lambda^i \quad (3)$$

where w_i is the individual weight ratio of the ith conformationally different class whose ellipticity at wavelength λ is $[\theta]_\lambda^i$. Substituting this expression into eq 2, we obtain

$$[\theta]_\lambda' = (M_N/M_P)\sum_i w_i[\theta]_\lambda^i \quad (4)$$

If these i classes are removed sequentially, one class at a time, then it can be shown that

$$\frac{d[\theta]_\lambda'}{dW} = \left(\frac{M_N}{M_P}\right)\frac{\partial \Sigma w_i[\theta]_\lambda^i}{\partial w_i} = \left(\frac{M_N}{M_P}\right)\left[\frac{\partial(w_i[\theta]_\lambda^i)}{\partial w_i}\right]_{w_{j\neq i}} \quad (5)$$

and

$$\frac{d[\theta]_\lambda'}{dW} = \frac{M_N}{M_P}\left\{[\theta]_\lambda^i + w_i\left(\frac{\partial[\theta]_\lambda^i}{\partial w_i}\right)_{w_{j\neq i}} + \sum_{j\neq i} w_j\left(\frac{\partial[\theta]_\lambda^j}{\partial w_i}\right)_{w_{j\neq i}}\right\} \quad (6)$$

where the index i refers to the specific class which is being removed. This expression can be simplified if we make the further assumption—which is amenable to experimental verification—that the partial derivatives $(\partial[\theta]_\lambda^{i\text{ or }j}/\partial w_i)_{w_{j\neq i}}$ are 0. This amounts to the assumption that either conformational changes in all classes of proteins do not occur as the ith class is removed, or, alternatively, that conformational changes which do ensue are compensatory. This first is perhaps a risky assumption in the present instance, and the latter seems somewhat improbable. Both, however, can be verified by subsequent calculations based on results of this analysis.

With this assumption

$$d[\theta]_\lambda'/dW = (M_N/M_P)[\theta]_\lambda^i \quad (7)$$

If these i conformational classes are removed together but independently of one another, $[\theta]_\lambda'$ is replaced in eq 7 by a weighted average of the ellipticities of the several classes which are being simultaneously dissociated.

When these conformational classes are quantitatively significant, a plot of $[\theta]_\lambda'$ *vs.* W should show several linear regions. The *slopes* should yield values of $[\theta]^i$ for the individual classes which are being *dissociated* as the W values decrease. In contrast, the $[\theta]$ value at a given W should reflect the weighted average, $[\theta]_\lambda^{prot.,av}$, of the proteins *remaining* in complex. These values of $[\theta]^i$ and $[\theta]_\lambda^{prot.,av}$ can then be translated into confor-

[1] Abbreviations used are: CD, circular dichroism; NaEDTA, the sodium salt of ethylenediaminetetraaceticacid; TNH, purified calf thymus chromatin; T_m, the midpoint of thermal melting transitions.

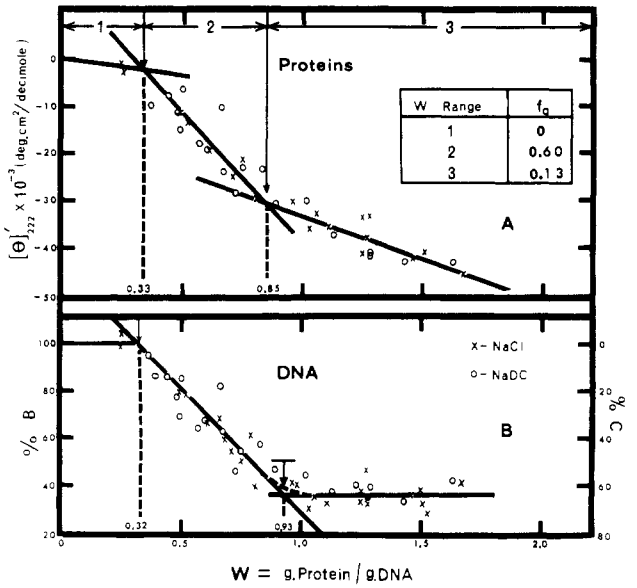

FIGURE 1: Correlation of secondary structures of the protein and DNA constituents of calf thymus chromatin with protein content. (A) The protein contribution to the CD spectrum at 222 nm, $[\theta]_{222}'$, is plotted on the ordinate against the protein content, W, on the abscissa for the residual complexes prepared by the NaCl (×) and the sodium deoxycholate (○) dissociation procedures. The inset gives the results of the analysis of the average helix content of the proteins removed in the indicated ranges of W values (see text). (B) The per cent B and the per cent C character are plotted on the left- and right-hand ordinates, respectively, against the protein content, W, for the same residual complexes described above in A. The lines through the points in both A and B represent the results of the linear regression analyses of the data from both sets of experiments. The W values within the borders of the figures are the intersection points calculated as described in the text. The horizontal bar over the arrow in B represents the range of those W values of the intersection falling within the 95% zone of confidence for the regression characteristics of data at $W \geq 1.00$.

TABLE I: Characteristics of the Linear Regression Analysis of the Dependence of the Protein Contribution at 222 nm to the Circular Dichroism Spectrum on Protein Content of Calf Thymus Chromatin.

Characteristic of Function	Dissociation Procedure: NaCl	Na Deoxycholate	Combined
$0.25 \leq W \leq 1.00$		$W = m[\theta]_{222}' + b$	
$m \times 10^3$	−0.0211	−0.0163	−0.0178
b	0.236	0.326	0.302
r	−0.944	−0.840	−0.879
t	−6.40	−5.13	−7.84
Significance	Yes, $P < 0.005$	Yes, $P < 0.001$	Yes, $P < 0.001$
$F_{slope} = 0.848$ (N.S.); $F_{intercept} = 0.138$ (N.S.)			
$1.00 \leq W \leq 2.25$		$[\theta]_{222}' = mW + b$	
$m \times 10^{-3}$	−16.65	−19.45	−17.80
$b \times 10^{-3}$	−16.49	−14.60	−15.65
r	−0.793	−0.794	−0.779
t	−3.69	−2.61	−4.65
Significance	Yes, $P < 0.01$	Yes, $P \leq 0.06$	Yes, $P < 0.01$
$F_{slope} = 0.118$ (N.S.); $F_{intercept} = 1.39$ (N.S.)			

mational information, using the reference spectra in the literature, as described in the Experimental Section.

Analysis of our data in this fashion yields the results shown in the upper part of Figure 1 and in Tables I, II, III, and IV. In Figure 1A we have plotted $[\theta]'$ at 222 nm against the W values for the residual complexes obtained from the NaCl (×) and the sodium deoxycholate (○) dissociation procedures. As was true for the per cent B values (reproduced in the bottom portion of the same figure), the dependence of $[\theta]'$ on W appears to be linear. The characteristics of the linear regression analyses are given in Table I. A covariance analysis reveals that there is no essential difference in the slopes and intercepts of the set of straight lines generated by the data from the two separate dissociation experiments above $W = 0.25$. Below $W = 0.25$, we have only two data points and 0 for the complexes prepared by the NaCl method. The line drawn in Figure 1A represents that determined by 0 and the average value of $[\theta]_{222}'$ of the two complexes of approximately equal protein content.

The intersection points of these three lines were determined algebraically using the results of the linear regression analyses for the combined data presented in Table I and the assumed linear characteristics of the NaCl complexes below a W of 0.25. These intersection points, shown in Figure 1A, corresponded to W values of 0.33 and 0.85 for the data at 222 nm.

The behavior of $[\theta]_\lambda'$ at other wavelengths (240, 230, 225, and 220 nm) as a function of W was similar to that described above for 222 nm. Values of $[\theta]_\lambda^i$ obtained from the slopes of the combined data at these latter wavelengths are given in Table II.

Operating on the basis that our assumptions and interpretations were correct, we calculated the fractions of α helix (f_α), β structure (f_β), and random coil (f_R) from the values of $[\theta]^i$ presented in Table II. These fractions were calculated from both the data at a single wavelength, 222 nm, as well as the data from the several wavelengths given in Table II by the least-mean-squares averaging procedure described in the Experimental Section. The values from the two sets of reference spectra used (*i.e.*, Chen *et al.* (1972); Greenfield and Fasman (1969)) are presented in Table III. With the exception of the entries for $0.25 \leq W \leq 1.00$ obtained by the least-mean-squares averaging procedure, the agreement between values calculated by the several procedures and different reference spectra is reasonably good. In general, however, the better fits of the data are observed with the reference spectra of Chen *et al.* (1972). We thus feel that the values of f_α, f_β and f_R obtained in the latter set of calculations are more reliable.

Above a W value of 1, the magnitude of the various fractions calculated from the several sets of wavelengths as well as the value of $[\theta]^i$ at 222 nm are in reasonable agreement. The average values, based on the two sets of reference spectra, of the helix content, β structure, and random coil are 0.13, 0.21, and 0.66, respectively. The deviation of the value of f_α and f_R calculated from the value of $[\theta]_{222}^i$ using the Greenfield and Fasman limit spectra can reasonably be explained on the basis that the assumption that only α helix and random coil structures were present is in error. Below a W value of 1, however, the values of f_β calculated from the least-mean-squares averaging procedure are negative, implying that the data cannot reasonably be fitted by the sets of reference spectra available. We feel that this is attributable to experimental error due to a high helix content rather than a failure of the assumptions which we have previously outlined concerning the presence of standard conformations of peptide structures. This is supported by the fact that if the data for $[\theta]'$ are used directly to calculate values of $[\theta]^{prot.,av}$ at various W values and these, in turn, are used to

TABLE II: Molecular Ellipticities of the Protein Fractions *Dissociated* from Calf Thymus Chromatin by NaCl and Sodium Deoxycholate.

	$[\theta]_\lambda{}^i \times 10^{-3}$ (deg cm²/dmol)[b]				
W Range	240 nm	230 nm	225 nm	222 nm	220 nm
$0 \leq W \leq 0.25$[a]	−0.85[a]	−1.94[a]	−2.28[a]	−2.86[a]	−4.86[a]
$0.25 \leq W \leq 1.00$	−3.16	−14.69	−19.78	−19.90	−18.27
$1.00 \leq W \leq 2.25$	−0.61	−4.61	−6.18	−6.32	−6.79

[a] All values repeated in this row of figures are appropriate only for the NaCl experiments. [b] Calculated by eq 7 in text and the slopes of $[\theta]_\lambda'$ *vs.* W. M_N was taken as 310 and M_P as 110.

TABLE III: Fractional Distribution of Various Secondary Structures in Proteins *Dissociated* from Calf Thymus Chromatin by NaCl and Sodium Deoxycholate.

		Least-Mean-Squares Anal. of $[\theta]_\lambda{}^i$			$[\theta]_{222}{}^i$ Anal.	
W Range	Ref Spectra	f_α	f_β	f_R	f_α	$f_{(\beta+R)}$ or f_R
$0 \leq W \leq 0.25$	Chen *et al.* (1972)	0.02	0.14	0.84	0.01	0.99
	Greenfield and Fasman (1969)	0.06	0.28	0.66	0.17	0.83
$0.25 \leq W \leq 1.00$	Chen *et al.* (1972)	0.60	−0.27	0.67	0.60	0.40
	Greenfield and Fasman (1969)	0.87	−0.58	0.71	0.60	0.40
$1.00 \leq W \leq 2.25$	Chen *et al.* (1972)	0.15	0.11	0.74	0.12	0.88
	Greenfield and Fasman (1969)	0.12	0.32	0.56	0.26	0.74

TABLE IV: Fractional Distribution of Various Secondary Structures in Proteins *Remaining* in the Residual Complexes of Calf Thymus Chromatin.

		Least-Mean-Squares Anal. of $[\theta]_\lambda{}^{prot.,av}$			$[\theta]_{222}{}^{prot.,av}$ Anal.	
W	Ref Spectra	f_α	f_β	f_R	f_α	$f_{(\beta+R)}$ or f_R
	Combined Data from NaCl and Na Deoxycholate Experiments					
0.40	Chen *et al.* (1972)	0.08	0.18	0.74	0.08	0.92
	Greenfield and Fasman (1969)	0.04	0.40	0.56	0.23	0.77
0.60	Chen *et al.* (1972)	0.25	0.03	0.72	0.25	0.75
	Greenfield and Fasman (1969)	0.32	0.07	0.61	0.35	0.65
0.80	Chen *et al.* (1972)	0.34	−0.05	0.71	0.33	0.67
	Greenfield and Fasman (1969)	0.45	−0.09	0.64	0.41	0.59
1.00	Chen *et al.* (1972)	0.32	−0.01	0.70	0.32	0.68
	Greenfield and Fasman (1969)	0.41	−0.03	0.62	0.40	0.60
1.20	Chen *et al.* (1972)	0.29	0.00	0.71	0.29	0.71
	Greenfield and Fasman (1969)	0.37	0.01	0.62	0.38	0.62
1.40	Chen *et al.* (1972)	0.27	0.02	0.71	0.26	0.74
	Greenfield and Fasman (1969)	0.34	0.05	0.61	0.36	0.64
	Residual Complex from Ca–Urea Experiment					
0.50	Chen *et al.* (1972)	0.27	0	0.73	0.27	0.73
	Greenfield and Fasman (1969)	0.37	0	0.63	0.36	0.64

calculate the fraction of α, β, and random coil structures left in complex, these fractional values are reasonable in all W ranges. This point is demonstrated by the data shown in Table IV, which gives the fractions of the various peptide structures based on the values of $[\theta]_\lambda'$ at several W values taken from the linear least-mean-squares regression analysis of the combined data. If structures other than α, β and random coils were actually present in the complex to any significant extent, one would expect the values of $[\theta]^{prot.,av}$ to reflect this fact as well. Hence, one would expect f_α, f_β or f_R to show anomalies such as highly negative values.

It is quite apparent from the values of $[\theta]^i$ in the W range between 0.25 and 1.00 that the proteins being dissociated contain a large fraction of their residues in helical form. If it is assumed that only α helix and random coil forms exist, then an approximate value of f_α can be calculated at every wavelength. Except for the data at 240 nm, the values are reasonably consistent within ±0.05 and average about 0.6 for both sets of reference spectra. This average is in excellent agreement with the value calculated from the data at 222 nm, using the reference

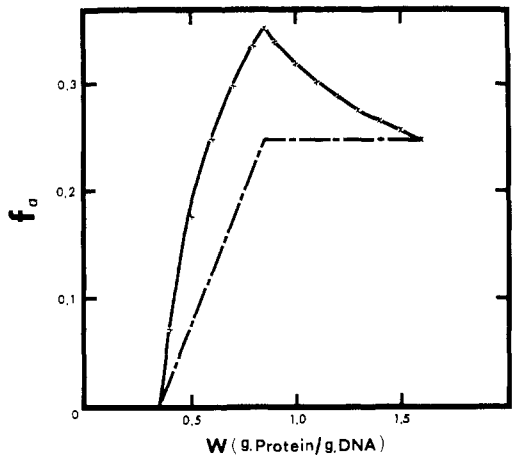

FIGURE 2: Predicted and observed dependence of the helix content of the proteins left in complex on protein content, W, of the residual complexes of calf thymus chromatin. The fraction of residues in the α helical structure, f_α, based on CD measurements at 222 nm and the reference spectra of Chen *et al.* (1972), are plotted against the W values of the residual complexes. The crosses (×) represent the observed data based on the least-mean-squares regression values of $[\theta]_{222}'$ of the combined data from the NaCl and the sodium deoxycholate experiments at a given W value. The solid line represents the predicted values based on eq 8 in the text. The dashed line represents the predicted behavior for a loss of α helix character in the proteins left in complex after the protein content of the complexes has been reduced to 0.85 g of protein/g of DNA.

spectra of Chen *et al.* (1972). It should be noted from the data presented in the lower part of Figure 1 that the major loss of C character in the DNA component occurs in this same range of W in which the high helix content proteins are dissociating.

As was shown in paper I, the W range encompassing the dissociation of arginine-rich histones overlaps, to a certain extent, in the two dissociation procedures. Thus, it might be argued that the high helix content proteins which were being removed between W values of 0.9 and 0.3 were these histones, and this process was only coincidental with the increase in the B character of the DNA. This possibility may be ruled out, however, by the fact that the helix content of the proteins of the complexes prepared by the Ca–urea procedure is only moderate, amounting to *ca.* 30%, regardless of whether the Chen *et al.* (1972) or the Greenfield and Fasman (1969) reference spectra are used. The last entry in Table IV gives the actual values for a typical complex of this type. If these arginine-rich proteins were indeed of high helix character when complexed alone to DNA, one would have expected this value of f_α to be 0.6.

The data for the W range below 0.25 are nonexistent. Using the last two points from the NaCl experiments to represent an upper limit of the helix content for those proteins dissociating below this limit, one obtains an f_α of *ca.* 0. It is clear that the helix content of these proteins, which are nonhistone proteins, must be negligible.

As stated earlier, this conformational anaylsis is based on the assumption that no change of structure occurs in the members of a given class of proteins in a complex as members of that class are removed. If this were not the case, one would anticipate that the linear regression of $[\theta]_\lambda'$ on W for W values greater than 1 should exhibit an intercept of 0 at $W = 0$. As is apparent from Figure 1 and the data shown on Table I, this is definitely not the case.

This assumption can also be verified in a more positive fashion by calculating directly the values of f_α, f_β, and f_R of the proteins remaining in complex, from the experimental data, $[\theta]_\lambda^{\text{prot.,av}}$, as was done in Table IV and comparing these values to the predicted values, based on the fact that in a complex of weight ratio, W, the predicted value of the fractional amount of any one of these conformations (α, β, or random coil structure) should be a weighted average of the contributions of each of the discrete fractions remaining in the given complex. For the α helix content, this would amount to a predicted value of

$$f_\alpha = \frac{\Sigma w_i f_{\alpha i}}{W} = \frac{w_1(0)}{W} + \frac{w_2(0.60)}{W} + \frac{w_3(0.13)}{W} \quad (8)$$

where the subscripts refer to the three W ranges in which the different classes are removed, and w_1, w_2, and w_3 are the weights of protein/gram of DNA left in each zone.

A comparison of these two values of f_α, the predicted and the observed, as a function of W is shown in Figure 2. The predicted values, calculated from eq 8, are shown as a continuous solid line. For these calculations, the f_α values for the various protein dissociation ranges were taken from the f_α values calculated from $[\theta]_\lambda^i$ data using the reference spectra of Chen *et al.* (1972). The points, which fall precisely on this curve, are the observed values of f_α calculated from the values of $[\theta]^{\text{prot.,av}}$ at 222 nm, using the reference curves of Chen *et al.* (1972). The dashed line represents the expected dependence in f_α if a loss of helix content in the remaining proteins ensued after the W value was reduced to 0.85. The failure of the observed values of f_α to conform to this latter curve demonstrates the validity of the assumption that marked conformational changes in protein structure either do not occur at all in the proteins remaining in complex as neighboring proteins are dissociated, or alternatively, occur in a compensatory fashion.

In previous experiments, we have noted increases in the fractions of base pairs in the C conformation in the presence of divalent ions such as Ca^{2+} or Mg^{2+} and, correspondingly, decreases upon the addition of NaEDTA.[1] In the light of conformational correlations between protein and DNA structure revealed by these experiments, it was of interest to ascertain whether these agents also had an effect on the protein spectra. Such effects were not observed, however. The values of $[\theta]^i$ and $[\theta]^{\text{prot.,av}}$ were unaffected by the presence of either of these divalent ions or the chelator, NaEDTA. One may thus conclude that changes in the C character of the DNA constituent of these complexes which ensue upon the addition of these agents are mainly attributable to the direct interaction of the divalent ions with DNA, rather than indirect effects on the conformation of associated proteins.

These analyses rest, of course, on a number of assumptions which may introduce errors, or, in some cases, make this approach invalid. Errors in the fractional amounts of various secondary structures of the protein constituent may arise because of inaccuracies in the reference spectra employed. As long as these reference spectra do not change dramatically, during the reduction of the protein content, such errors should not seriously jeopardize the major conclusion that the helix content of the proteins removed in the W range in which the C structure of the DNA constituent is being lost is substantially higher than that of the proteins removed in W ranges in which the C character is independent of W.

There is one possibility, however, which would seriously affect this analysis and, in fact, would invalidate it in its present form. This is the possibility that a tertiary structure exists in native chromatin in which the asymmetric secondary structures of the proteins are themselves arranged in an asymmetric—or nonrandom—array in such a manner as to contribute an additional band to the CD spectrum. Upon disruption of this tertiary structure, by removal of the key contributing proteins, this

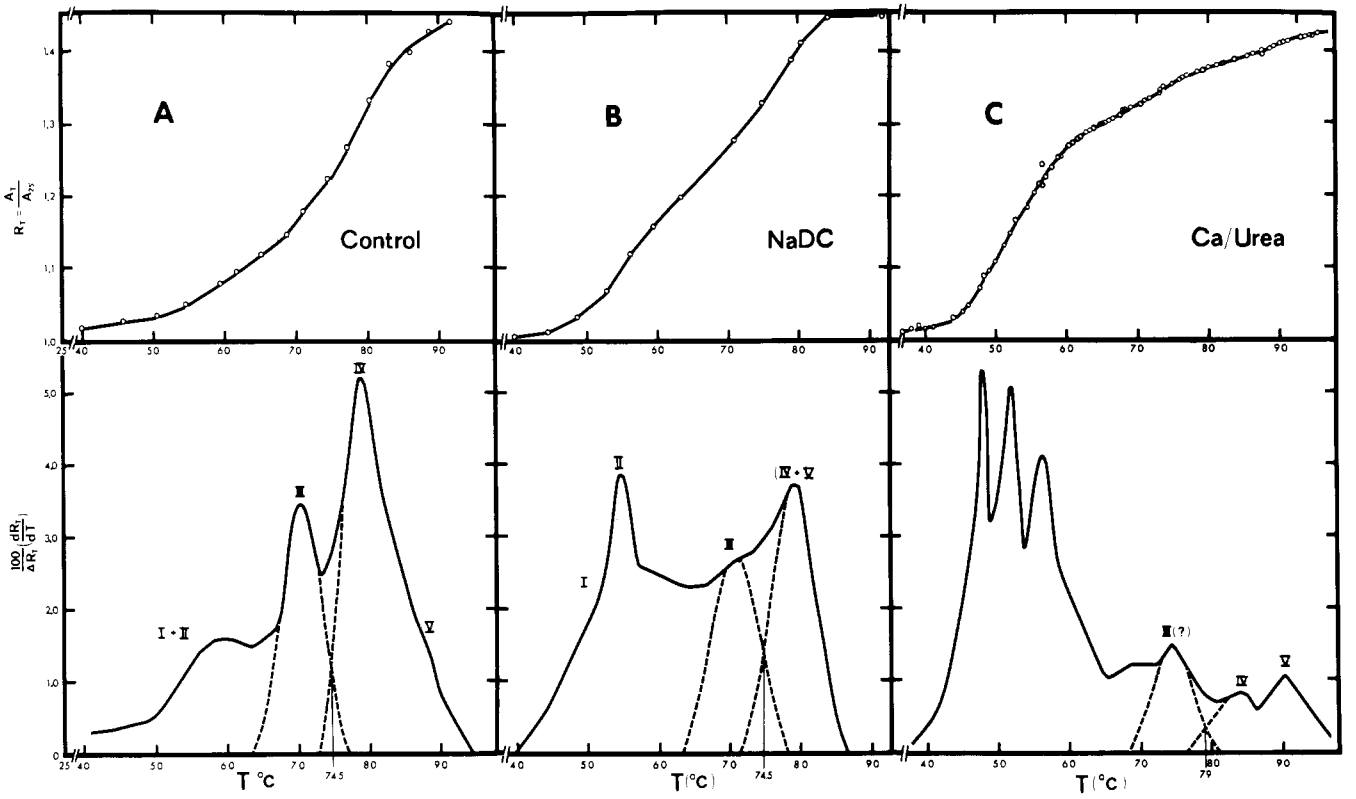

FIGURE 3: Thermal melting profiles of intact and residual complexes of calf thymus chromatin at low ionic strengths. In the upper part of these figures, the integral profiles are plotted in terms of the relative absorbance, R_T, as a function of temperature, T (°C). The lower set of figures represents the normalized derivative patterns, calculated from the smoothed integral profiles. The Roman numerals represent the various thermal transitions. Transitions III and IV have been resolved into two gaussians, as described in the text. The temperature at which they intersect, T_C, is indicated by a thin vertical line: (A) intact TNH (control), with a W of 1.3 g/g, in 2.5×10^{-3} M Tris-2.5×10^{-4} M NaEDTA (pH 7.5); (B) residual TNH prepared by the sodium deoxycholate dissociation procedure, with a W of 0.60 g/g in 2.5×10^{-3} M Tris-2.5×10^{-4} M NaEDTA (pH 7.5); (C) residual TNH prepared by the Ca-urea procedure with a W of 0.45 g/g, in 1×10^{-3} M NaCl-2.5×10^{-4} M NaEDTA (pH 7.5).

band would correspondingly be lost. If such a band were negative and had the shape of an α helix spectrum between 220 and 240 nm, then the same patterns of $[\theta]_\lambda'$ *vs.* W and the same self-consistency in the calculations of the fractions of various secondary structures by the procedures described would be observed.

Although this is not a possibility which we can unequivocally exclude at the present time, we, nevertheless, feel that it is an unlikely explanation for the observed changes in $[\theta]_\lambda'$ *vs.* W. If such were the case it is unreasonable that only the peptide spectrum would be affected without an accompanying effect in the DNA spectrum. Yet we do not find any unaccountable distortions of the DNA spectra. Until some of the studies currently under way in this laboratory are completed, we prefer to interpret our data along the simpler line of differences in secondary structure of classes or proteins. Tertiary structure may indeed be lost in the W range between 0.25 and 1.0 g/g, but we do not think that this is directly affecting the CD spectrum, either in the peptide or the nucleotide regions of the spectrum.[2]

Thermal Melting Profiles. We have previously noted that the fraction of bases in the C conformation matched the fraction of bases melting out in the high melting transitions of the thermal melting profiles of intact calf thymus chromatin. We therefore undertook a set of thermal melting experiments to ascertain whether this same correlation held for the residual complexes.

Typical thermal melting profiles (in both the integral and the derivative forms) for the several residual complexes and a control are shown in Figures 3 and 4. The integral curves have been plotted as R_T, the ratio of the absorbance of the maximum (259 nm) at temperature, T, to the absorbance at 25° (corrected for the slight volume dilution due to thermal expansion) against T. The derivative patterns have been plotted in a normalized form, $[(100/\Delta R_T)(dR_T/dT)]$, where ΔR_T is the total change in R_T between 25 and 97°.

The several transitions seen in the intact control have been numbered in conventional fashion, using the Roman numeral designation. A typical pattern for intact TNH is shown in Figure 3A. In our preparations, we usually see a shoulder at *ca.* 90°, and have correspondingly assigned the number V to this transition, which is generally not reported as a separate transition by other workers. Although we differ in some details, we are in general agreement with the assignments of these various transitions made by others (Shih and Bonner, 1970; Ansevin and Brown, 1971; Ansevin *et al.*, 1971; Li and Bonner, 1971; Li *et al.*, 1973). The areas under transitions III, IV, and, in our preparations, V, generally decrease with decreasing histone content of the complex whereas the areas associated with transitions I and II increase. Thus, transitions III, IV, and V have been assigned to the melting of protein covered DNA segments, whereas the two lower transitions are attributed to the

[2] It might be noted that simple random aggregation alone can also contribute extraneous bands to a CD spectrum due to light scattering and other optical effects. Previous studies (Johnson *et al.*, 1972) have demonstrated the absence of such artifacts in the spectrum of intact chromatin isolated by a variety of methods. (Chromatin preparations subjected to urea concentrations known to permanently disrupt aggregates showed no significant differences in their spectral properties after urea removal when compared to control solutions of aggregated preparations which had not been subjected to these urea concentrations.)

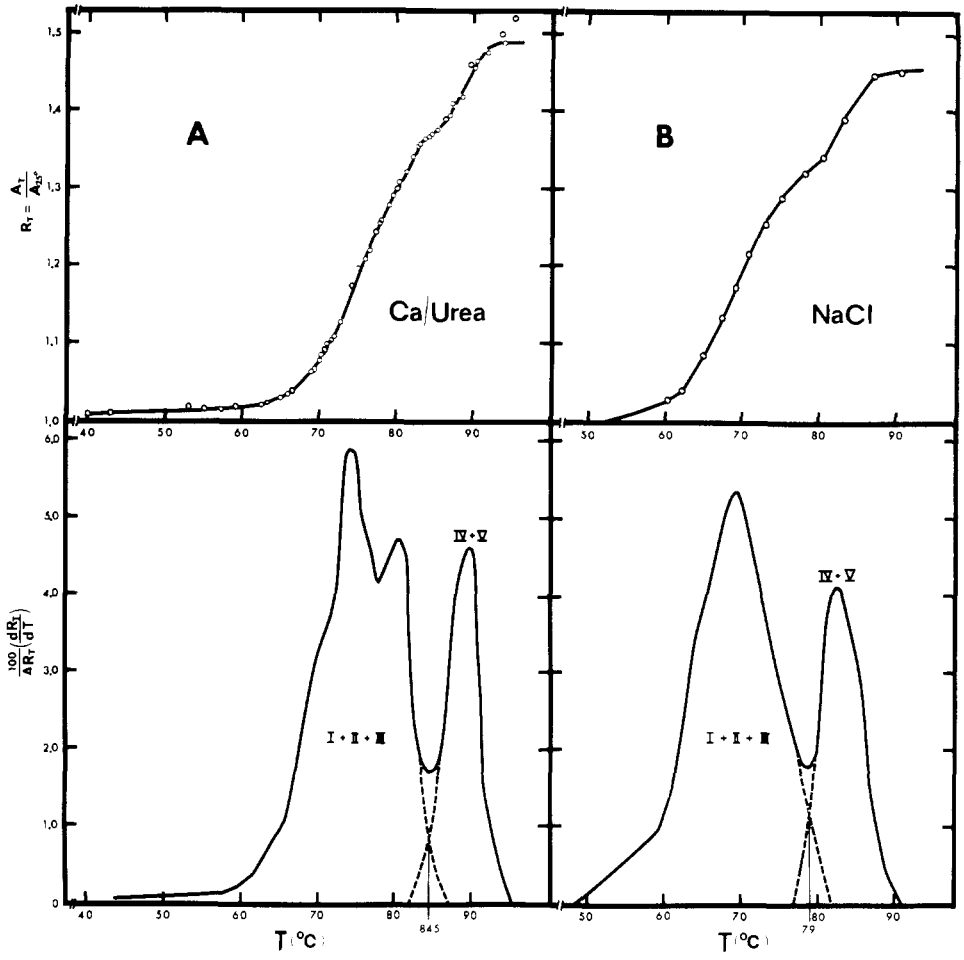

FIGURE 4: Thermal melting profiles of residual complexes of calf thymus chromatin at higher salt concentrations. The format of this figure is the same as that described for Figure 3: (A) residual TNH prepared by the Ca-urea dissociation procedure, with a W of 0.50, in 2×10^{-2} M NaCl-1×10^{-2} M Tris-2.5×10^{-4} M NaEDTA (pH 7.5); (B) residual TNH prepared by the NaCl dissociation procedure, with a W of 0.61, in 1×10^{-2} M NaCl-2.5×10^{-3} M Tris (pH 7.5).

melting of protein-free DNA regions. Transition I has a T_m which corresponds to the T_m of linear protein-free DNA (of the same base composition) and hence has been assigned to the melting of protein-free regions of DNA at the ends of chromatin molecules. By default transition II has been assigned to loops of protein-free DNA stabilized at both ends by more thermally stable histone associated portions of DNA which do not melt out at that T_m.

In our experiments, it is sometimes difficult to resolve transitions I, II, and III from each other. These low melting transitions, however, are always clearly differentiable in intact chromatin from the high melting transitions, IV and V, and we have thus had no difficulty in calculating the fraction of bases melting out in these higher transitions. This fraction, as indicated previously, matches almost exactly in every control preparation examined the fraction of bases in the C conformation, as measured by the CD spectrum.

In contrast to the intact parent chromatin preparations, however, the residual complexes prepared by sodium deoxycholate or by the Ca-urea treatment exhibit poorly resolved transitions and/or additional complexities not seen in the parent chromatin profile. These points are demonstrated in Figures 3B and C, which show the profiles of a sdium deoxycholate complex and a Ca-urea complex, respectively. In the pattern for the sodium deoxycholate complex (Figure 3B), transitions III and IV are completely unresolved. In order to extract the fraction of bases melting out in the higher melting transitions, one must resort to curve fitting. For this purpose, we have assumed that the derivatives of transitions III and IV are gaussian in shape about their T_m values, and the position of these T_m values is the same as that observed in the intact control in the same solvent (71 and 79°, respectively). The temperature, T_C, is then located at which the gaussians associated with the two transitions make an equal contribution to the derivative melting profile. The fraction of bases melting out in transition IV and any higher transition is then taken as the fraction of bases melting out in the integral curve above the temperature T_C. Correspondingly, the fraction of bases melting out in transitions I, II, and III is that fraction of bases melting out below T_C.

The value of T_C for this particular solution is 74.5°. A similar analysis of the control as well as other preparations in this solvent also gave the same value, as is demonstrated by the data shown in Table V. The presence or absence of NaEDTA has no marked influence on this position of T_C, although it shifts to a higher temperature with marked increases in ionic strength, as might be expected. (The arginine-rich complexes produced by Ca-urea also exhibit an elevated T_c which is undoubtedly due to an asymmetry in the base composition of the histone covered base pairs in these complexes (Clark and Felsenfeld, 1971; Marmur and Doty, 1962)). Regardless of solvent, the fraction of bases melting out in transitions above T_C agrees very well with the fraction of bases in the C conformation, as Table V shows.

TABLE V: Correlation of Melting Profile and Circular Dichroism Characteristics of Residual Complexes and Controls of Calf Thymus Chromatin.

Dissoc. Reagent	W (g of Prot/ g of DNA)	Solvent (pH 7–7.5)	T_c (°C)	$100f_{T \leq T_c}$	% B (CD)	$100f_{T \geq T_c}$	% C (CD)
None (control)	1.43	2.5×10^{-3} M Tris	75	33	33	67	67
0.03 M NaDC[a]	0.83	2.5×10^{-3} M Tris	~75	50	57	50	43
0.05 M NaDC	0.49	2.5×10^{-3} M Tris	~75	56	66	44	34
1.0 M NaCl	0.61	1×10^{-2} M NaCl–2×10^{-4} M Tris	79	72	66	28	34
1.0 M NaCl	0.61	5×10^{-2} M NaCl–1×10^{-2} M Tris	84	68	66	32	34
None (control)	1.29	2.5×10^{-3} M Tris–2.5×10^{-4} M NaEDTA	74.5	51	52	49	48
0.03 M NaDC	0.75	2.5×10^{-3} M Tris–2.5×10^{-4} M NaEDTA	74.5	67	66	33	34
0.05 M NaDC	0.60	2.5×10^{-3} M Tris–2.5×10^{-4} M NaEDTA	74.5	72	75	28	25
0.07 M NaDC	0.44	2.5×10^{-3} M NaCl–2.5×10^{-3} M Tris–2.5×10^{-4} M NaEDTA	79	88	86	12	14
1 mM $CaCl_2$, 5 M urea	0.50	2×10^{-2} M NaCl–1×10^{-2} M Tris–2.5×10^{-4} M NaEDTA	84.5	76	77	24	23
1 mM $CaCl_2$, 5 M urea	0.45	1×10^{-3} M NaCl–2.5×10^{-4} M NaEDTA	79.5	85	86	15	14

[a] Sodium deoxycholate.

In many of the profiles of the residual complexes prepared by these three dissociation procedures, there frequently appear transitions not seen in the parent intact control. This is particularly evident in the case of the Ca–urea complexes, one of whose profiles is shown in Figure 3C. Part of this complexity is undoubtedly attributable to the fact that the G + C content of the protein covered regions is markedly enriched, as reported by Clark and Felsenfeld (1971). This complexity creates problems in the identification of the several transitions. We assume that the transition at 90° in this profile corresponds to transition V normally found in the intact controls. We also assume that the transition at 84.5° is that portion of transition IV in the intact controls due to the association of the arginine-rich proteins with the higher than average G + C content regions of DNA.

The identification of transition III is more difficult since there appears to be two transitions in the vicinity of the original transition III found in intact chromatin. It is again reasonable to assume, however, that he transition at *ca.* 75° reflects that portion of bases melting out in transition III with the shift in T_m due to the higher than average G + C content of these histone covered stretches of DNA. The suggestion of a transition at *ca.* 69° may simply be due to the extraneous protein (other than the arginine-rich histones) found in these complexes.

On the other hand, it is possible that this doublet character of the transitions of the Ca–urea complexes reflects an effect which was first invoked by Li and Bonner (1971), to explain the presence of transitions III and IV in intact chromatin. According to these authors, transitions III and IV reflect the melting of two sets of DNA bases linked by a common histone molecule. The higher transition (IV) reflects the melting of that portion of the bases associated with the more basic end of the histone whereas transition III corresponds to the melting behavior of the bases associated with the less basic half of the same histone. Although we do not agree with this interpretation for the given transitions in intact chromatin (Hanlon *et al.*, 1972; Johnson *et al.*, 1972) there is the possibility that the doublet nature of the transitions seen in Figure 3C is due to this effect operating within each transition. That is, the transitions at 85 and 90° could be due to the association of a single histone molecule, or class of molecules, involved in maintaining C structure of DNA, with the higher transition being due to the association of the more basic portions of the polypeptide chain of this given histone. The two transitions seen at *ca.* 69 and 75° could be due to a similar association of another type or class of histone molecules, associated with regions of DNA in the B conformation.

This interpretation is supported in part by the fact that at higher salt concentrations, this doubling of the transitions above 80°, at least, disappears, and the combined transition shifts to 90°. This is demonstrated by the profile shown in Figure 4A, which shows the melting behavior of the Ca–urea complex at a higher NaCl concentration (0.025 M NaCl). Figure 4B shows the melting behavior of a NaCl complex at a comparable W value at a somewhat higher ionic strength. The high melting transitions, in both cases, are much more easily separated from the transitions of lower T_m. The fraction of bases melting out in these last transitions still matches very closely the per cent C character, as is seen from the entries in Table V.

Conclusion

Both the thermal melting data and the CD spectra confirm the fact that discrete classes of proteins which differ in conformation and ability to thermally stabilize DNA are associated with the two different conformational regions of DNA, both in intact chromatin as well as the residual complexes. The pro-

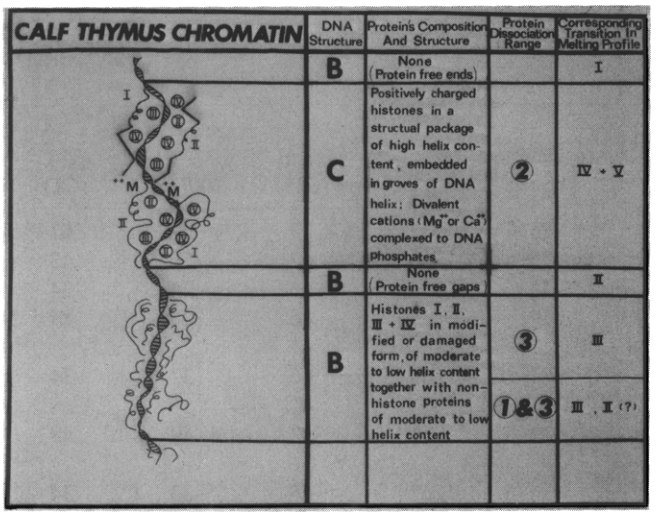

FIGURE 5: Schematic model of intact calf thymus chromatin. The circled Roman numerals in the schematic model to the left are meant to represent high helix content histones of the designated compositional classes. The structural package of histone proteins of average helix content of *ca.* 60% is completed by the addition of external histones of a more flexible character and moderate helix content. This package effectively excludes water and other components of the solvent from the vicinity of the DNA helix which is in the C conformation. The remaining portions of the DNA genome, which are in the B conformation, are either protein free or else associated with more structurally disorganized histone and nonhistone proteins of moderately low or zero helix content. The circled Arabic numbers in the second to last column refer to the ranges in *W* shown in Figure 1. The Roman numerals in the last column at the right are the transitions in the melting profiles, shown in Figure 3A, in which the DNA bases in various regions of this complex are assumed to melt.

teins which are uniquely associated with the C conformational regions are of high helix content. We suggest that the ability to fold in this fashion permits these histones to fit effectively in one or both of the grooves of the DNA structure, thus making for maximal displacement of water, maximal protection, and maximal efficiency in the neutralizing and shielding of the negatively charged phosphates of the DNA backbone. Thus, these regions are the most thermally stable ones, as is evident from the correlations with the melting profile data. Previous studies have also shown that these regions are also protected from the effects of nuclease digestion (Hanlon *et al.*, 1972; Johnson *et al.*, 1972).

We have summarized the results of papers I and II of this series as well as experiments described in previous communications from our laboratory and those of others (Richards and Pardon, 1970; Smart and Bonner, 1971b,c; Rill and Van Holde, 1974) in the form of a schematic model for calf thymus chromatin presented in Figure 5. In this model (a revised version of the earlier structure (Hanlon *et al.*, 1972; Johnson *et al.*, 1972)) the histones associated with the C conformational regions of DNA form a unique structural package of a high average degree of helicity, in the neighborhood of 60%. The bulk of this package is probably embedded in the grooves of the DNA helix. These histones are also associated with more flexible histones probably lying on the exterior. The latter play a supplementary role in terms of providing additional charge neutralization of the DNA phosphates and may also facilitate or participate in macromolecular interactions. The removal of a portion of these more flexible histones in the presence of the more tightly bound high helix content proteins does not markedly alter the C character of the complex.

One of the flexible extended histones in this model has been portrayed as histone I as it can be removed in its entirety in the NaCl experiments without a noticeable increase in the B character of the residual complexes. In addition, experiments in the literature, summarized by Smart and Bonner (1971b,c), suggest that this histone fraction is on the exterior of the chromatin molecule and facilitates intermolecular interactions. On the basis of the results in sodium deoxycholate (Smart and Bonner, 1971b,c) subfractions of histones II may play a similar role. The high helix content proteins are presumably among the arginine rich (III + IV) and the slightly lysine rich (II).

There is a considerable amount of data in the literature to suggest that chromatin is more compact than the DNA which is derived from it (Zubay and Doty, 1959; Ohba, 1966; Wilhelm *et al.*, 1970; Rill and Van Holde, 1974). This compactness has been attributed (Smart and Bonner, 1971b; Rill and Van Holde, 1974) to the fact that the complex assumes to a *partial* extent in solution some variation of the supercoiled form seen in fiber form (Pardon *et al.*, 1967; Pardon and Wilkins, 1972). In this regard, it is interesting that in the experiments of Smart and Bonner (1971b) the enhanced linear dichroism observed for pea bud chromatin (which can be explained by supercoiling) is los over, roughly, the same range of histone content where we find appreciable losses in the C strucutre. Garrett (1971) also found that the characteristic reflections attributed to the supercoiled form were lost by the residual complexes prepared by a NaCl dissociation over the same range of NaCl concentrations corresponding to that which, on the average, results in the destruction of the C character in our studies. These facts suggest that there is a correlation between tertiary structure and the secondary C structure of the DNA helix. Indeed, we have reason to believe that the C form lends itself to a more compact tertiary folding of DNA (B. Wolf and S. Hanlon, manuscript in preparation) and hence we have attempted to summarize these speculations in the model by a suggested tertiary fold in these C regions. *We wish to emphasize, however, that the lowering of the positive CD band can be accounted for solely by the C structure* of the DNA constituent and not to any tertiary structure of the nucleoprotein complex.

In addition to the external histones another portion of the low helix content proteins must be complexed directly to DNA. (We have represented this portion in the model shown in Figure 5 by a set of schematic disorganized structures attached to regions of DNA in the B form.) This conclusion is based on the results of the melting experiments. In these experiments, as previously noted, we have found that the fraction of bases in the B conformation corresponds to the fraction of bases melting out in transitions I, II, and III, both in intact chromatin and the residual complexes. These regions of DNA also appear to be relatively accessible to solvent components, such as nucleases (Hanlon *et al.*, 1972; Johnson *et al.*, 1972). In native intact chromatin isolated by the methods described, approximately 20% of the bases melt out in the histone associated transition III (Johnson *et al.*, 1972). An examination of the literature reveals that this fraction of bases melting out in transition III can be increased, at the expense of the fraction melting out in the higher T_m transitions, IV and V, by proteolytic damage (Ansevin *et al.*, 1971) as well as the presence of urea, a protein denaturant (Ansevin *et al.*, 1971; Subirana, 1973). These facts taken together suggest that the histones in association with the DNA bases melting out in transition III are structurally disorganized.

It might be noted that in both the reported data in the litera-

ture and our own results, it is quite clear that there is no *increase* in the fraction of bases melting out in transition III as histones are removed from intact chromatin. This is demonstrated in the data presented by Li and Bonner (1971) and Li *et al.* (1973) for the residual complexes of pea bud and calf thymus chromatin prepared by the NaCl dissociation procedure. The fraction of bases in transition III in their experiments undergoes a very definite decrease with decreasing histone content. In the case of the residual complexes of calf thymus chromatin which we have prepared by the sodium deoxycholate procedure of Smart and Bonner (1971a), this decrease is harder to demonstrate as the melting profiles are less readily resolvable. It is certain, however, that the fraction of bases melting out in the temperature range in which transition III is normally found in the intact control does not increase in the residual complexes. In the light of our interpretation of the nature of transition III, these data support the contention that the removal of histones competent to maintain the C character of intact chromatin does not result in the unfolding of the remaining members of this class. If the latter were actually the case, one would expect to see an enrichment of transition III with decreasing protein content. This is not observed.

It is possible that this group of "inept" histones associated with transition III has been derived from the structural package of histones maintaining the C structure of DNA by inadvertent proteolytic damage incurred during the course of the isolation. Indeed, the disc gel electrophoresis patterns confirm that there is proteolytic damage to histones. Alternatively, this class could also have arisen as a result of cellular mediated modifications such as phosphorylation, acetylation, or directed proteolysis, which could serve as mechanisms of derepression. The disorganization of histone structure would then serve to expose the bases to the cellular environment.

In any event, it is clear that the histone proteins involved in maintaining the fraction of bases in C structure in nucleohistone are of a conformational character different from those histones associated with the DNA in the B form in complex. We have ascribed this conformational difference to a difference in the helix content of the histones as this is the interpretation which best accounts for our data.

Acknowledgments

We wish to thank Dr. Barry Wolf and Dr. Stephen Berman for helpful discussions and insights.

References

Ansevin, A. T., and Brown, B. S. (1971), *Biochemistry 10,* 1133.

Ansevin, A. T., Hnilica, L. S., Spelsberg, T. C., and Kehm, S. L. (1971), *Biochemistry 10,* 4793.

Chen, Y. H., Yang, J. T., and Martinez, H. M. (1972), *Biochemistry 11,* 4120.

Clark, R. J., and Felsenfeld, G. (1971), *Nature (London) New Biol. 240,* 226.

Garrett, R. A. (1971), *Biochemistry 12,* 2227.

Greenfield, N., and Fasman, G. D. (1969), *Biochemistry 8,* 4108.

Hanlon, S., Johnson, R. S., Wolf, B., and Chan, A. (1972), *Proc. Nat. Acad. Sci. U. S. 69,* 3263.

Johnson, R. S., Chan, A., and Hanlon, S. (1972), *Biochemistry 11,* 4347.

Li, H. J., and Bonner, J. (1971), *Biochemistry 10,* 1461.

Li, H. J., Chang, C., and Weiskopf, M. (1973), *Biochemistry 12,* 1763.

Marmur, J., and Doty, P. (1962), *J. Mol. Biol. 5,* 109.

Ohba, Y. (1966), *Biochim. Biophys. Acta 123,* 76.

Pardon, J. F., and Wilkins, M. H. F. (1972), *J. Mol. Biol. 68,* 115.

Pardon, J. F., Wilkins, M. H. F., and Richards, B. M. (1967) *Nature (London) 215,* 508.

Richards, B. M., and Pardon, J. F. (1970), *Exp. Cell Res. 62,* 184.

Rill, R., and Van Holde, K. E. (1974), *J. Mol. Biol. 83,* 459.

Shih, T. Y., and Bonner, J. (1970), *J. Mol. Biol. 48,* 469.

Smart, J. E., and Bonner, J. (1971a), *J. Mol. Biol. 58,* 651.

Smart, J. E., and Bonner, J. (1971b), *J. Mol. Biol. 58,* 661.

Smart, J. E., and Bonner, J. (1971c), *J. Mol. Biol. 58,* 675.

Subirana, J. A. (1973), *J. Mol. Biol. 74,* 363.

Wilhelm, F. X., Champagne, M. H., and Daune, M. P. (1970), *Eur. J. Biochem. 15,* 321.

Zubay, G., and Doty, P. (1959), *J. Mol. Biol. 1,* 1.