Balhorn, R., Balhorn, M., Morris, H. P., and Chalkley, R. (1972), *Cancer Res. 32,* 1775–1784.
Butler, P. J. G., Harris, J. I., Hartley, B. S., and Leberman, R. (1969), *Biochem. J. 112,* 679–689.
De Lange, R. J., Fambrough, D. M., Smith, E. L., and Bonner, J. (1969), *J. Biol. Chem. 244,* 319–334.
Desai, L. S., and Foley, G. E. (1970), *Biochem. J. 119,* 165–170.
Greenaway, P. J., and Murray, K. (1971), *Nature (London), New Biol. 229,* 233–238.
Johns, E. W. (1967), *Biochem. J. 105,* 611–614.
Kinkade, J. M., Jr., and Cole, R. D. (1966), *J. Biol. Chem. 241,* 5790–5797.
Laine, B. (1974), Thèse de Doctorat du 3e Cycle en Biochimie, Université de Lille I.
Laine, B., Sautière, P., and Biserte, G. (1973), *Proc. Int. Congr. Biochem., 9th,* 110.
Luck, J. M., Rasmussen, P. S., Satake, K., and Tsvetikov, A. N. (1958), *J. Biol. Chem. 233,* 1407–1414.
Panyim, S., Bilek, D., and Chalkley, R. (1971), *J. Biol. Chem. 246,* 4206–4215.
Panyim, S., and Chalkley, R. (1969), *Arch. Biochem. Biophys. 130,* 337–346.
Patthy, L., and Smith, E. L. (1975), *J. Biol. Chem. 250,* 1919–1920.
Patthy, L., Smith, E. L., and Johnson, J. (1973), *J. Biol. Chem. 248,* 6834–6840.
Sautière, P., Moschetto, Y., Dautrevaux, M., and Biserte, G. (1970), *Eur. J. Biochem. 12,* 222–226.
Sautière, P., Tyrou, D., Laine, B., Mizon, J., Ruffin, P., and Biserte, G. (1974), *Eur. J. Biochem. 41,* 563–576.
Sautière, P., Tyrou, D., Moschetto, Y., and Biserte, G. (1971), *Biochimie 53,* 479–483.
Sautière, P., Wouters-Tyrou, D., Laine, B., and Biserte, G. (1975), *Ciba Found. Symp. 28,* 77–93.
Shlyapnikov, S. V., Arutyunyan, A. A., Kurochkin, S. N., Memelova, L. V., Nesterova, M. V., Sashchenko, L. P., and Severin, E. S. (1975), *FEBS Lett. 53,* 316–319.
Wilson, R. K., Starbuck, W. C., Taylor, C. W., Jordan, J., and Busch, H. (1970), *Cancer Res. 30,* 2942–2951.

# Physical Studies of the Nonhistone Chromosomal Proteins HMG-1 and HMG-2†

C. Baker,‡ I. Isenberg,* G. H. Goodwin, and E. W. Johns

ABSTRACT: The nonhistone chromosomal proteins, HMG-1 and HMG-2, have a folded conformation, with a high $\alpha$-helical content, over a wide pH range. At high and low pH values, the molecules unfold. Both molecules contain cysteine and tryptophan. The tryptophans appear to be buried in the folded form. HMG-1 shows aggregation at pH 5.7, as does HMG-2 at pH 9.0. The folded form is insensitive to high concentrations of salt, suggesting that charge-charge interaction plays no role in stabilizing the tertiary structure.

Early work on nonhistones was hampered by their insolubility. However, it was found (Goodwin et al., 1973) that differential $Cl_3CCOOH$ precipitation could separate the nonhistones extracted from chromatin with 0.35 M NaCl into low and high molecular weight fractions. The low molecular weight fraction is freely soluble. Chromatography of this fraction, named the high mobility group, or HMG[1] proteins, on a carboxymethyl-cellulose column at pH 9.0 (Goodwin and Johns, 1973) yields two pure fractions called HMG-1 and HMG-2. These are the same proteins found as contaminants in F1 isolated by perchloric acid extraction (Johns 1964a, 1964b).

HMG-1 and HMG-2 are present in about $10^5$ to $10^6$ copies per nucleus (Johns et al., 1975; Walker et al., submitted) making it unlikely that they are specific gene activators. HMG-1 has been shown to complex with F1 and bind nonspecifically to DNA (Shooter et al., 1974; Goodwin et al., 1975). HMG-1 stimulates chromatin template activity (Johns et al., 1975).

HMG-1 and HMG-2 contain approximately 25% basic and 30% acidic residues which appear to be asymmetrically distributed. Preliminary results on the sequencing of the N-terminal residues indicate that HMG-1 and HMG-2 are similar, but contain some amino acid substitutions (Johns et al., 1975)

In this work we present a number of fundamental properties of HMG-1 and HMG-2. We report that they each contain cysteine and tryptophan. We give molar extinction coefficients for each protein. We describe a number of physical properties of HMG-1 and HMG-2, and present new data on the electrophoretic mobility of these proteins, and describe the results of CD fluorescence anisotropy, and turbidity measurements. We show that HMG-1 is an acidic and HMG-2 a basic protein. Among other things, we show that these proteins can exist in a highly folded conformation which has a high $\alpha$-helical content. The conformation is very insensitive to high salt concentration. We show that

† From the Department of Biochemistry and Biophysics (C.B. and I.I.), Oregon State University, Corvallis, Oregon 97331, and the Chester Beatty Research Institute (G.H.G. and E.W.J.), Institute of Cancer Research, Royal Cancer Hospital London, England. *Received July 22, 1975.* The work of C. Baker and I. Isenberg was supported by U.S. Public Health Service Grant CA-10872, that of G. H. Goodwin and E. W. Johns by the Medical Research Council and the Cancer Research Campaign.

‡ Present address: The Rockefeller University, New York, New York.

[1] Abbreviation used: HMG, high mobility group.

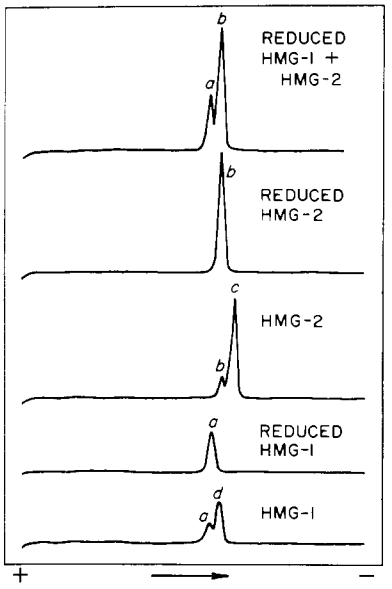


FIGURE 1: Densitometer tracings using the 10% urea gels of Orrick et al. (1973). Samples were dissolved in 10 M urea and reduced with β-mercaptoethanol.

Table I: Molar Extinction Coefficient.[a]

| | By Nitrogen Determination | By Amino Acid Analysis |
|---|---|---|
| HMG-1 | 21.9 ± 1.6[b] (5)[c] | 20.9 ± 1.6 (5) |
| HMG-2 | 20.5 ± 1.3 (3) | 20.0 ± 2.0 (3) |

[a] At 280 nm, in units of $10^3$ l. $mol^{-1}$ $cm^{-1}$. [b] Standard deviation. [c] Number of Analyses.

the folded conformation exists over a wide pH range but over part of the range the proteins aggregate. We therefore have defined the conditions that should be used in future physical studies.

## Experimental Section

Calf-thymus HMG-1 and HMG-2 were prepared as previously described (fractions A and C, respectively, of Goodwin and Johns, 1973). HMG-1 and HMG-2 were electrophoresed on 10% urea gels for 5 h (Orrick et al., 1973). Samples were 0.5 mg/ml in 10 M urea.

To obtain the molar extinction coefficients of HMG-1 and HMG-2, proteins were dissolved in 0.01 N HCl at a concentration of about 1 mg/ml. Protein concentrations were determined by two methods: (1) nitrogen analysis using the method of Jaenicke (1974) [Assuming that 14% of the acidic amino acids are amidated (Walker et al., submitted), the protein concentrations were then calculated using data from amino acid analyses.] (2) Amino acid analysis. To 1 ml of the protein solution, about 0.25 μmol of norleucine was added as an internal standard. Hydrolysis was carried out for 24 h in 6 N HCl at 110 °C. After evaporating to dryness and redissolving in 0.01 N HCl, an aliquot was analyzed using a JEOL JLC-6AH amino acid analyzer. The original protein concentration was determined by the summation of the individual amino acids. The molar extinction coefficients, obtained by these methods, are given in Table I. The molecular weight of HMG-1 was taken to be 26 500 (Shooter et al., 1974) and that for HMG-2 to be 26 000 (Goodwin et al., 1975). In this paper the values 20 900 and 20 000 l. $mol^{-1}$ $cm^{-1}$ will be used for HMG-1 and HMG-2, respectively.

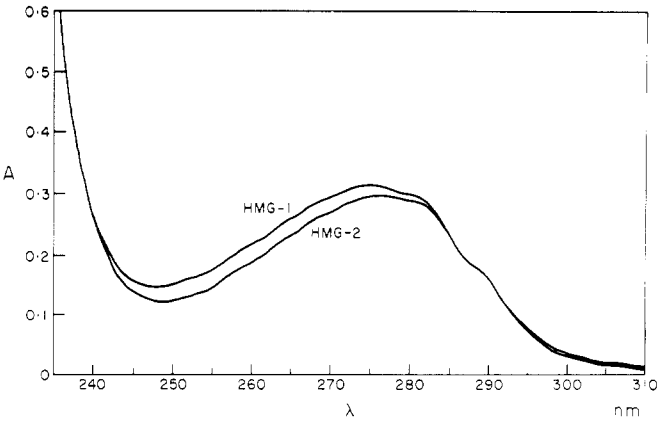


FIGURE 2: Uv absorbance spectra of equimolar concentrations, $1.44 \times 10^{-5}$ M, of HMG-1 and HMG-2 in 0.01 N HCl and $10^{-4}$ M dithiothreitol.

The tryptophan content was estimated from the experimental extinction coefficients and the tyrosine content determined by amino acid analysis. The molar extinction coefficients for tyrosine and tryptophan were taken from Edelhoch (1967). The presence of tryptophan was confirmed by fluorometry.

Protein stock solutions were prepared as follows: 5–8 mg of protein and an equal weight of dithiothreitol (Sigma) were dissolved in 0.2–0.3 ml of glass distilled water, and 0.4–0.5 ml of 0.1 M Tris, 9 M urea (pH 9.0), was added. The sample vial was flushed with argon and incubated at 40 °C for 1 h. The solution was then desalted on a G-25 Fine Sephadex column (1.2 × 26 cm) equilibrated with 0.01 M HCl and 1 mM dithiothreitol. Fractions were pooled to make a stock solution of known concentration, which was kept refrigerated until used. Electrophoresis of stock solutions showed that the protein was free of inter- and intramolecular disulfide bonds. Stock solutions were diluted into either glass-distilled water or dilute NaOH for pH titrations or into the appropriate buffers for NaCl titrations. In the latter case the sample was then titrated to the desired pH with NaOH or HCl before titrating with NaCl.

Fluorescence anisotropy $r = (I_{\parallel} - I_{\perp})/(I_{\parallel} + 2I_{\perp})$ was measured on a computer-controlled polarization spectrometer (Ayers et al., 1974). All measurements were at 20 °C. Samples were excited at 295 nm and emission was measured at 340 nm. At these wavelengths only tryptophan is excited, eliminating the complication of depolarization by tyrosine to tryptophan energy transfer.

Light scattering was measured turbidometrically on a Cary 14 spectrophotometer as $A_{360}$ values.

Circular dichroism (CD) was measured on a Jasco Model J-10 CD recorder at 20 °C. Data are reported as $\Delta\epsilon$ in units of l. $cm^{-1}$ $mol^{-1}$ of residue. The amino acid analyses were used to estimate the number of residues per 26 500 daltons for HMG-1 and the number per 26 000 daltons for HMG-2. We obtained 230 and 232 residues per molecule of HMG-1 and HMG-2, respectively. CD spectra were analyzed for the amount of α-helix, β-sheet, and random coil by a new method (Baker and Isenberg, 1976).

## Results

*Electrophoresis.* Electrophoresis of HMG-1 and HMG-2 on 10% urea gels gives two closely spaced bands for each protein (Figure 1). Reduction with β-mercaptoethanol be-

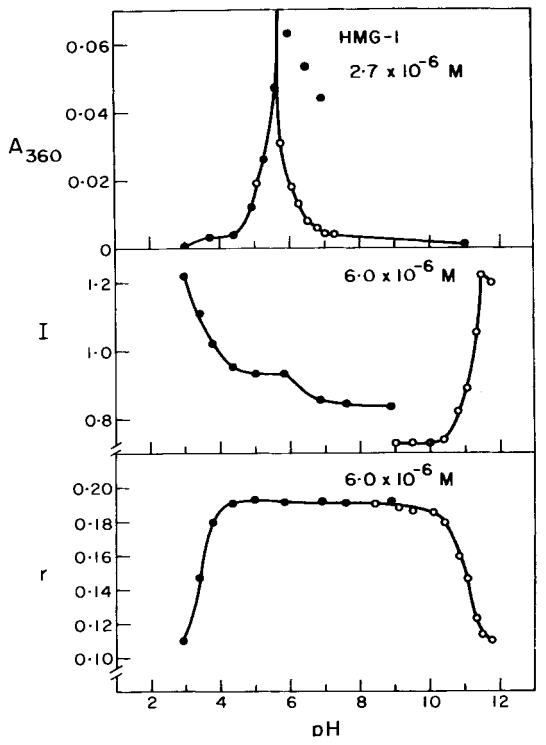


FIGURE 3: Titration of reduced HMG-1 with NaOH starting with HMG-1 in 1.2 mM HCl, 0.12 mM dithiothreitol (●); and with HCl starting with a similar solution which had been raised instantaneously to pH 12 with NaOH (○).

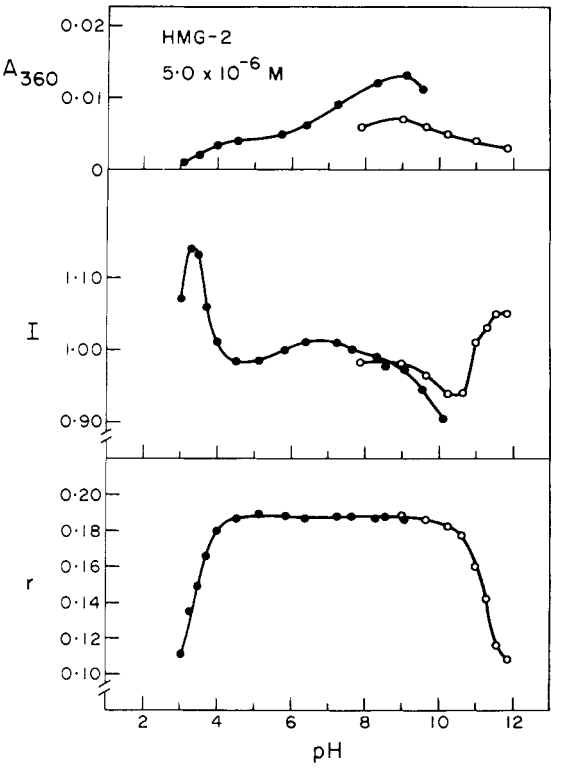


FIGURE 4: Titration of reduced HMG-2 with NaOH starting with HMG-2 in 0.8 mM HCl and 0.08 mM dithiothreitol (●); and with HCl starting with HMG-2 at pH 12 (○).

fore electrophoresis gives only the slower moving band for both proteins. Co-electrophoresis of reduced HMG-1 and HMG-2 gives two bands. Amino acid analysis gives 2 cysteines per protein molecule for both HMG-1 and HMG-2. Apparently in the oxidized form, an internal disulfide bridge prevents the molecule from totally relaxing when denatured by urea, thus giving a lower frictional coefficient for the oxidized form.

*Tyrosine and Tryptophan Content.* Figure 2 shows the uv absorbance spectra of equal concentrations of HMG-1 and HMG-2 in 0.01 M HCl. Using molar extinction coefficients at 280 nm of 20 900 l. $mol^{-1}$ $cm^{-1}$ for HMG-1 and 20 000 l. $mol^{-1}$ $cm^{-1}$ for HMG-2, plus the molar extinction for the two spectral coefficients of tyrosine and tryptophan (Edelhoch, 1967), we estimate that there are 7 tyrosines and 2 tryptophans per molecule of HMG-1 and 6 tyrosines and 2 tryptophans per molecule of HMG-2.

*Fluorescence and Light Scattering: Effects of Salt and pH Changes.* The fluorescence anisotropy of HMG-1 and HMG-2 show similar pH dependences (Figures 3 and 4). Either protonation of the acidic residues or deprotonation of the basic residues denatures the protein. There is a broad plateau region between pH 4.5 and pH 10 where the anisotropy remains nearly constant. The fluorescence intensity is quenched by folding, but the exact pH dependence differs for each (Figures 3 and 4). The turbidity of HMG-1 is very different from that of HMG-2 (Figures 3 and 4). HMG-1 gives a large scatter peak at pH 5.7. This is near the estimated pI so this is probably a nonspecific aggregation due to net charge neutralization. HMG-2 also shows a scatter peak near its pI (9.0) but it is much lower.

Once the aggregated state is formed, it remains fairly stable to pH changes so care must be exercised as to the path taken to reach a particular pH.

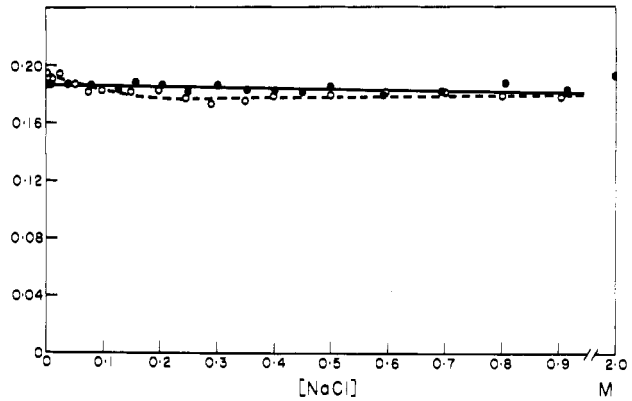


FIGURE 5: NaCl titration of 6.2 × $10^{-6}$ M HMG-1 in 10 mM Tris, pH 8 (●). NaCl titration of 5.0 × $10^{-6}$ M HMG-2 in 10 mM cacodylic acid, pH 6.25 (○).

One should note that the anisotropy is not significantly increased by aggregation. Therefore the relatively high anisotropy of the plateau is not a result of intermolecular interactions. Secondary structure alone will not immobilize the tryptophan side groups either. Therefore it can be concluded that intramolecular folding occurs such that the tryptophans are buried inside the molecule.

The addition of NaCl to HMG-1 at a nonaggregating pH decreases the anisotropy only slightly. Most of the change occurs by 0.1 M NaCl, at which point the anisotropy has leveled off at 0.178. Addition of NaCl to HMG-2 produces no significant change in anisotropy. There is not even the small decrease seen for HMG-1 at low salt concentrations (Figure 5).

*Circular Dichroism: Salt and pH Effects.* Figures 6 and 7 show the CD spectra of HMG-1 and HMG-2 under vari-

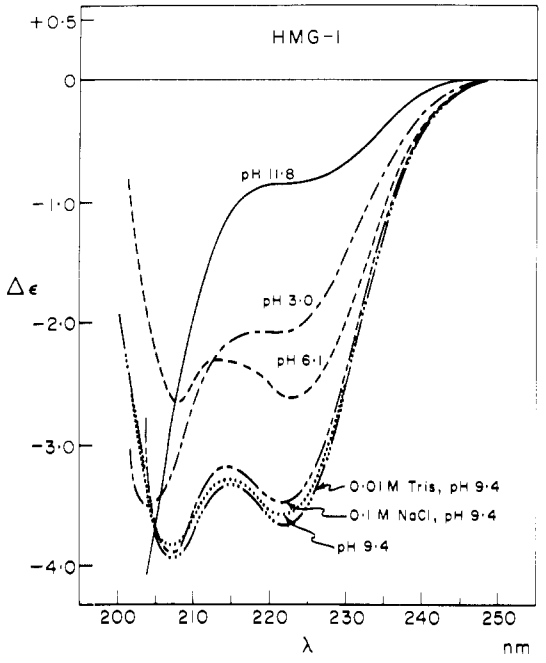


FIGURE 6: CD of $6.0 \times 10^{-6}$ M HMG-1, $10^{-4}$ M dithiothreitol at the pH's indicated. The NaCl sample also contained 10 mM Tris. $\Delta\epsilon$ in units of l. $cm^{-1}$ $mol^{-1}$ of residue.

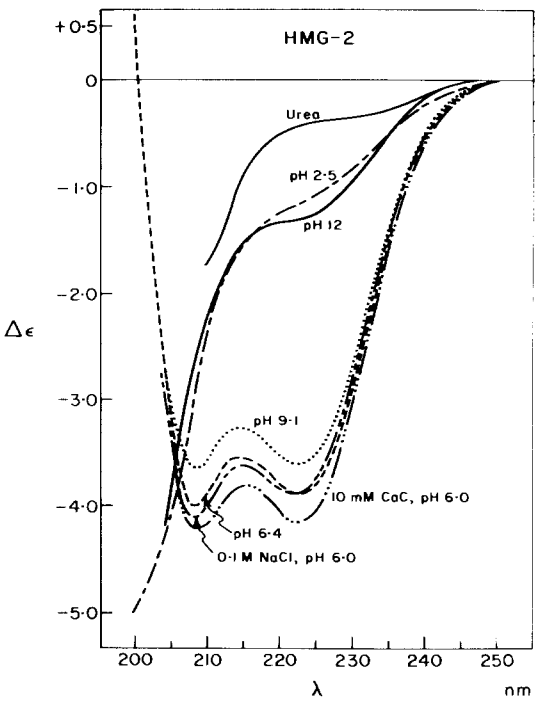


FIGURE 7: CD of $4.8 \times 10^{-6}$ M HMG-2, $10^{-4}$ M dithiothreitol. pH's were adjusted with HCl and NaOH. The NaCl sample also contained 10 mM cacodylic acid. $\Delta\epsilon$ in units of l. $cm^{-1}$ $mol^{-1}$ of residue.

ous conditions. As discussed below, the CD of HMG-1 demonstrates the existence of an $\alpha$-helical content of 41%. Addition of 0.1 M NaCl changes the CD only slightly. At pH 6.1 HMG-1 is highly aggregated. The CD spectrum is similar in shape but is decreased in magnitude. This decrease may be a scattering artifact. At pH 3.0 the anisotropy is near its minimum, but the CD shows that residual structure exists. At pH 11.8 the $\alpha$-helical contribution to the CD of HMG-1 is less than that found at pH 3, and going to pH 13 does not change the CD further.

The CD of HMG-2 shows that it has 50% $\alpha$-helical content in the plateau region. The CD at pH 9.1 is less than at pH 6.4, and this may also be a scattering artifact. Addition of 0.1 M NaCl at pH 6.4 produces no change in the CD of HMG-2. The CD of HMG-2 at pH 12 is nearly identical with that at pH 2.5. At these pH values, the protein is largely unfolded, but the CD in 5 M urea at pH 2.4 indicates that extreme pH's alone will not totally unfold the protein.

The CD of HMG-1 and HMG-2 were analyzed by a newly developed method (Baker and Isenberg, 1976). This method provides a number of tests for the choice of reference spectra. One of the tests is that the sum of the various spectral components should add to unity; another is that the results should be independent of the spectral range used in the analysis. With these criteria a number of reference spectra could be rejected as unsuitable for HMG-1 and HMG-2. For example, use of the poly(L-lysine) CD spectra for $\alpha$, $\beta$, and random reference spectra to analyze the CD of HMG-1 gave a sum of only 0.69 instead of unity. In contrast to this, use of the Chen et al. (1974) spectra gave a more satisfactory sum of 1.04 and an analysis of 0.41, 0.04, and 0.59 for $\alpha$, $\beta$ and random coil components, respectively. For HMG-2, we obtained 0.50, −0.017, and 0.49 for the $\alpha$, $\beta$, and random coil components, respectively. The small negative component for the $\beta$-sheet content is, of course, not meaningful; within experimental error, the $\beta$-sheet content is zero.

## Discussion

HMG-1 and HMG-2 are similar proteins which differ mainly in molecular weight, pI, and degree of aggregation near the pI. It is interesting to note that HMG-1 is an acidic, and HMG-2 is a basic protein.

HMG-1 and HMG-2 each have two cysteines and two tryptophans, results which have hitherto not been noted. The tryptophans must have similar distributions and similar conformations in the folded state since the tryptophan anisotropy has the same pH dependency for each. HMG-1 and HMG-2 exist in a highly structured conformation, rich in $\alpha$-helix over a broad pH range. This conformation is insensitive to pH and salt changes.

It is not possible to give a direct functional test for the native conformation of HMG-1 and HMG-2 since, at the present time, the functions of these proteins are not known. Nevertheless, it is possible to give the same type of indirect criteria that have been found useful in the study of histones (Van Holde and Isenberg, 1975; Isenberg, 1976). Extensive experience in studies of enzymes has shown that functional activity is associated with a highly structured compact conformation. In lieu of a more direct criterion, we may, at least tentatively, adopt such a test for HMG-1 and HMG-2, although it is possible that small parts of the molecule could be in a nonnative form even though most of the protein is highly structured. In this sense we state that the neutral, folded forms of HMG-1 and HMG-2 are either in the native state, or close to it.

Sequence data so far obtained (Johns et al., 1975) indicate that the N-terminal sequences of both HMG-1 and HMG-2 are highly basic, and that for HMG-1 both hydrophobic and acidic regions exist. Since the folded conformation can be disrupted by going to high or low pH, but not by high salt, it is possible that charge–charge interactions do not play an important role in stabilizing the folded state. If so, the charges would be asymmetrically distributed, not

only in the sequence, but also in the folded form. We may then speculate that acidic and basic regions would be available to interact with DNA, and proteins such as histones.

References

Ayers, W. A., Small, E. W., and Isenberg, I. (1974), *Anal. Biochem. 58,* 361.
Baker, C. C., and Isenberg, I. (1976), *Biochemistry 15,* 629.
Chen, Y., Yang, J. T., and Chau, K. H. (1974), *Biochemistry 13,* 3350.
Edelhoch, H. (1967), *Biochemistry 6,* 1948–1954.
Goodwin, G. H., and Johns, E. W. (1973), *Eur. J. Biochem. 40,* 215–219.
Goodwin, G. H., Sanders, C., and Johns, E. W. (1973), *Eur. J. Biochem. 38,* 14–19.
Goodwin, G. H., Shooter, K. V., and Johns, E. W. (1975), *Eur. J. Biochem. 54,* 427.
Greenfield, N., and Fasman, G. D. (1969), *Biochemistry 8,* 4108.
Isenberg, I., (1976), in Search and Development—A Volume Dedicated to Albert Szent-Gyorgyi, Kaminer, B., Ed., New York, N.Y., Academic Press (in press).
Jaenicke, L. (1974), *Anal. Biochem. 61,* 623.
Johns, E. W. (1964a), *Biochem. J. 92,* 55–59.
Johns, E. W. (1964b), The Nucleohistones: Proceedings of the first World conference on histone biology and chemistry, Bonner, J., and Tso, P., Ed., San Francisco, Holden-Day, pp 52–58.
Johns, E. W., Goodwin, G. H., Walker, J. M., and Sanders, C. (1975), *Ciba Found. Symp. 28,* 95–112.
Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951), *J. Biol. Chem. 193,* 265.
Orrick, L. R., Olson, M. O. J., and Busch, H. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 1316–1320.
Shooter, K. V., Goodwin, G. H., and Johns, E. W. (1974), *Eur. J. Biochem. 47,* 263–270.
Van Holde, K. E., and Isenberg, I. (1975), *Acc. Chem. Res. 8,* 327.
Walker, J. M., Goodwin, G. H., and Johns, E. W., submitted for publication.

# Synthesis and Properties of Carbonylbis(methionyl)insulin, a Proinsulin Analogue Which Is Convertible to Insulin by Cyanogen Bromide Cleavage†

Wolf-Dieter Busse‡ and Frederick H. Carpenter*

ABSTRACT: The preparation and use of carbonylbis(L-methionine *p*-nitrophenyl ester) as a reversible cross-linking reagent for insulin are described. The reaction of 1 equiv of reagent with zinc insulin in dimethylformamide in the presence of triethylamine yields as one of the products $N^{\alpha A1}$,$N^{\epsilon B29}$-carbonylbis(methionyl)insulin, (CBM-insulin). The CBM-insulin was characterized by end group analysis and by the products formed on tryptic and chymotryptic cleavage. It possessed 91% of the immunological and 6.5% of the hormonal activity of insulin. Treatment of CBM-insulin with cyanogen bromide (CNBr) in 70% formic acid for 1 h resulted in nearly complete removal of the methionine bridge to yield insulin. A small amount of a side product was removed on DEAE-cellulose at pH 7.2 to give an overall recovery of insulin of 70–80%. Oxidative sulfitolyses of CBM-insulin gave the hexa(*S*-sulfonate) which was reduced with dithiothreitol to yield reduced CBM-insulin. The latter compound, containing 6 sulfhydryls, exhibited a pH-dependent circular dichroic spectrum. The form at pH 10 exhibited a spectrum typical of random coil which was converted to a form at pH 7.8 which was characterized by a negative extremum at 213 nm. The change in the spectrum at 213 nm with pH was characterized by an apparent p$K_a$ of 8.5. Studies on the reoxidation of reduced CBM-insulin were performed at pH values between 7.8 and 10 and at protein concentrations of 0.01–1 mg/ml. The best yields (ca. 85%) of the correctly paired disulfide bonds were obtained in reoxidations at pH 9.5–10 at protein concentration of 0.01–0.1 mg/ml. CBM-insulin, which had been isolated from reoxidation at high pH of the reduced CBM-insulin, was cleaved by CNBr to yield a fully active insulin in an overall yield of 60% from the reduced CBM-insulin.

The present paper elaborates on preliminary communications (Busse and Carpenter, 1974; Busse et al., 1974; Busse, 1975) on the synthesis and application of a new bifunctional "methionine" reagent, carbonylbis(methionine *p*-nitrophenyl ester), which may be used for intra- and intermolecular cross-linking of the amino groups of insulin. From the reaction of insulin (**1**) with 1 equiv of CO(Met-ONp)$_2$,[1] an in-

† From the Department of Biochemistry, University of California, Berkeley, California 94720. *Received October 16, 1975.* This study was supported in part by Grants AM 00608 and EY 00813 from the National Institutes of Health.

‡ Recipient of a Travel Award from the Deutsches Forschungsgemeinschaft. W.-D.B. was on leave from the Deutsches Wollforschungsinstitut, Aachen, Germany, from March 1, 1973 to October 31, 1974.

[1] The abbreviations for amino acid derivatives and peptides are those proposed by the IUPAC-IUB Commission on Biochemical Nomenclature (1972): Boc, *tert*-butyloxycarbonyl; dansyl, 5-dimethylaminonaphthalene-1-sulfonyl; DCC, *N,N'*-dicyclohexylcarbodiimide; ONp, *p*-nitrophenyl. In addition, the following abbreviations are used: CBM-, carbonylbis(L-methionyl); CBM-insulin, insulin which is cross-linked by the CBM residue between the $\alpha$-amino group of glycine A1 and the $\epsilon$-amino group of lysine B29; CD, circular dichroism; CO(Met-ONp)$_2$, carbonylbis(L-methionine *p*-nitrophenyl ester); CNBr, cyanogen bromide.