Two new low-molecular-weight acidic proteins from calf thymus nuclei that resemble HMG (high-mobility-group) proteins 14 and 17

Peter D. CARY

*Biophysics Laboratories, Portsmouth Polytechnic, St Michael's Building, White Swan Road,
Portsmouth PO1 2DT, Hants., U.K.*

(Received 15 June 1982/Accepted 11 August 1982)

Two new, closely similar, acidic proteins were extracted and purified from calf thymus and designated AP-X and AP-Y (acidic proteins X and Y). They contain about 33% acidic residues, mostly non-amidated, and 20% lysine, but no arginine, tyrosine, histidine or tryptophan. There is a single phenylalanine residue in a molecular weight of approx. 5000. Circular dichroism and proton nuclear magnetic resonance show that they do not take up secondary or tertiary structure in free solution, as expected from the low content of hydrophobic amino acids. They appear structurally related to the high-mobility-group proteins HMG 14 and 17. Controlled extraction experiments indicate that proteins AP-X and AP-Y are at least partially located in the calf thymus nucleus.

The extraction, purification and characterization of minor chromosomal protein components such as the HMG proteins, ubiquitin, protein A-24 and histone H1⁰, has been a powerful stimulus to the study of the more unusual states of chromatin (e.g. transcriptionally active), for which such proteins may be essential requirements. In purification of HMG proteins from calf thymus by ion-exchange chromatography on CM-Sephadex we have consistently observed the presence of two components that are eluted in the run-through peak and which run on SDS/polyacrylamide gels somewhat slower than ubiquitin. They have been purified and partially characterized. Since they are acidic proteins, they have been designated AP-X (major component) and AP-Y (minor component).

Methods and results

Analytical techniques

Gel electrophoresis. SDS/0.1%-polyacrylamide gels at 15% were used as described by Laemmli (1970).

Spectroscopy. U.v. spectra were obtained on a Perkin-Elmer Coleman 572 instrument with a 1 cm-path micro-cell. C.d. spectra were obtained in a Cary 61 dichrograph and the protein concentrations were determined by amino acid analysis. The 270 MHz p.m.r. spectra were obtained in ²H₂O

on a Bruker WH270 Fourier transform spectrometer, with 5 mm tubes. Chemical shifts were measured relative to a trace of internal acetone at 2.24 p.p.m.

Amino acid analyses. These were obtained with a Rank-Hilger Chromaspek analyser. No corrections were made for hydrolytic losses.

Extraction and purification of proteins AP-X and AP-Y from calf thymus

Since the extraction of these proteins was a long and multi-step procedure, it is both described in the following paragraphs and also summarized diagrammatically in Scheme 1.

The thymus was frozen in liquid N₂ within 5 min of slaughter. Within 1 h it was submerged in 5% (w/v) HClO₄ at 4°C for sufficient time to allow removal of excess fatty tissue and external vessels. Then 500 g lots were blended in 650 ml of 5% HClO₄ at 4°C for 3 min and then spun at 2000 rev./min in a Mistral 6L centrifuge. Surface fat was poured off and 3.5 vol. of acidified acetone added. This precipitated largely histone H1. A further 2.5 vol. of acidified acetone was then added to precipitate the non-histone proteins (after standing at 4°C overnight), designated N_p. This procedure is a modification of that given by Sanders & Johns (1974).

The first purification step was that used by Goodwin & Johns (1977) in the purification of proteins HMG 1 and HMG 2 on a column of CM-Sephadex C-25 (2.6 cm × 40 cm). Typically, 4 g of non-histone protein pellet N_p was loaded in 50 ml

Abbreviations used: HMG proteins, high-mobility-group proteins; SDS sodium dodecyl sulphate.

Scheme 1. Summarized protocol for the preparation and purification of acidic proteins AP-X and AP-Y from calf thymus

of elution buffer [7.5 mM-sodium borate buffer (pH 9.0)/10 mM- β -mercaptoethanol] and the profile obtained had two main peaks and one smaller peak, as in profile I of Fig. 1. The components of the different peaks were identified by SDS/polyacrylamide-gel electrophoresis of six fractions and compared with the loaded protein as control (C): this is shown in panel (a) of Fig. 1. The first peak eluted (A) contains protein AP-X and other, presumably acidic, proteins, whereas the second peak (B) contains mostly proteins HMG 1 and HMG 2.

The concentration of protein AP-X in peak A is not great, but an alternative procedure was found more effective: if the non-histone fraction N_p was

taken up in 10 mM-HCl at 70 mg of protein/ml and then reprecipitated with 3 vol. of acidified acetone, most of the histone H1 and proteins HMG 1 and HMG 2 were removed (see Fig. 1, panel c, lanes 1 and 2). When a further 6 vol. of acidified acetone was added (to give fraction H_p), a better concentration was obtained of the faster-migrating proteins, including proteins AP-X and AP-Y. When fraction H_p was loaded on the Sephadex C-25 column in borate buffer and eluted with the same solution, then a much simpler profile resulted (profile II in Fig. 1), with a peak (S) essentially corresponding to peak A of profile I. The gel of peak S material (panel b) shows that indeed there has been a good concentration of the proteins running towards

FIG. 1. Ion-exchange chromatography on CM-Sephadex C-25 in 7.5 mM-sodium borate/10 mM- β -mercaptoethanol (pH 9.0) of non-histone precipitate N_p (profile I) and H_p (profile II)

Panel (a): SDS/polyacrylamide-gel electrophoresis of N_p and samples 1–6 from profile I. Panel (b): SDS/polyacrylamide-gel electrophoresis of peak S (sample 1) from profile II. Panel (c): SDS/polyacrylamide-gel electrophoresis of acidified-acetone-precipitated fractions from non-histone fraction N_p . Lanes 1 and 2 are loadings $\times 1$ and $\times 2$ of material precipitated with 3 vol. of acetone. Lanes 3 and 4 are loadings $\times 1$ and $\times 2$ of the material subsequently precipitated with 6 vol. of acetone (H_p). Abbreviations: H1, histone H1; UB, ubiquitin.

the bottom of the gel (including AP-X and AP-Y). Note that protein AP-X normally runs with a sharp front and a broad trailer in this gel system.

Material from peak S (profile II) or from peak A (profile I) was then loaded on a Sephadex G-75 column (2.7 cm $\times$ 140 cm) in 10 mM-HCl and eluted with the same solution. The A_{240} profile obtained is III in Fig. 2 and the main peak X contains largely proteins AP-X, AP-Y and ubiquitin, as seen from gel electrophoresis of fractions labelled 1–5 (Panel *a* in Fig. 2). In order to separate proteins AP-X and AP-Y from ubiquitin, the total protein in peak X was loaded on a Sephadex G-50 column (1.3 cm $\times$ 100 cm) in 10 mM-HCl and eluted with this solution. The elution obtained is profile IV in Fig. 2, and

electrophoresis of samples 1–9 shows that, whereas samples 1–5 have been freed of ubiquitin, samples 6–9 have not. To maximize the yield of proteins AP-X and AP-Y, these fractions 6–9 were collected from several runs of the Sephadex G-50 column, pooled, and re-run on Sephadex G-50. This recycle procedure was repeated three times in all. To obtain a complete separation of proteins AP-X and AP-Y, a column (1.1 cm $\times$ 32 cm) of CM-Sephadex C-25 was again used, but this time in 10 mM-potassium phosphate buffer, pH 5.9, with elution with a step gradient of 0, 50 mM-, 150 mM-, 300 mM- and 1000 mM-NaCl. The A_{240} profile is shown in Fig. 3; sample I is almost pure protein AP-X, but all subsequent fractions are still mixtures (gel panel *a*).

Fig. 2. Further chromatography of material obtained from the ion-exchange chromatography shown in Fig. 1. Elution profile III (●) is peak S of Fig. 1 chromatographed on Sephadex G-75 in 10 mM-HCl. Samples 1–5 are analysed on an SDS/polyacrylamide gel in panel (a). Elution profile IV (○) is total peak X of profile III rechromatographed on Sephadex G-50 in 10 mM-HCl. Samples 1–9 are analysed on an SDS/polyacrylamide gel in panel (b). Abbreviation: UB, ubiquitin.

Fig. 3. Ion-exchange chromatography on CM-Sephadex G-25 of protein AP-X/AP-Y mixture obtained from the Sephadex G-50 chromatography shown in Fig. 2

Elution was with a step gradient of 0, 50, 150, 300 and 1000 mM-NaCl. Panel (a) shows an SDS/polyacrylamide gel of samples 1–6, and panel (b) shows a SDS/polyacrylamide gel of pure samples of proteins AP-X and AP-Y compared with standards of cytochrome *c* (CC), bovine pancreatic trypsin inhibitor (TI) and the C-terminal CNBr fragment of calf thymus histone H2B (H2Bc) comprising residues 63–125.

A recycle procedure was again employed in which all mixed fractions were pooled, freeze-dried, desalted on Sephadex G-25, and re-run on the Sephadex C-25 with the same gradient. In this way the amount of pure protein AP-X obtained was increased, and gradually a sample of pure protein AP-Y emerged (albeit in much lower yield). As a demonstration of purity and to estimate molecular weight, the final pure samples of proteins AP-X and AP-Y were compared by gel electrophoresis with cytochrome *c*, bovine pancreatic trypsin inhibitor and the C-terminal half of histone H2B (H2Bc, residues 63–125): see panel (b) in Fig. 3. The proteins AP-X and AP-Y run as single bands of width similar to that of the standards. There is thus no evidence of any microheterogeneity in this system.

Characterization of proteins AP-X and AP-Y

Amino acid analysis. Table 1 shows results for both proteins AP-X and AP-Y. They are clearly very similar proteins with a high proportion of acidic residues (approx. 33%), exceeding that of basic residues (approx. 20%), and explaining why they are eluted early from the CM-Sephadex columns. The two proteins differ in that protein AP-X has more glutamic acid than protein AP-Y, whereas protein AP-Y has more aspartic acid than protein AP-X. The number of residues of each type has been calculated on the basis of a single phenylalanine residue. This implies a total of about 44 residues in protein AP-X and 42 residues in protein AP-Y, and that there is no arginine, histidine or tyrosine in either (as unequivocally demonstrated from n.m.r.

Table 1. *Amino acid analysis of proteins AP-X and AP-Y and chymotryptic fragments of AP-X, designated AP-X1 and AP-X2*

For details see the text. No corrections were applied for hydrolytic losses.

Amino acid	AP-X		AP-Y		APX-1		APX-2	
	(mol %)	(Phe = 1)	(mol %)	(Phe = 1)	(mol %)	(Leu = 1)	(mol %)	(Leu = 1)
Met	1.2	0	0.5	0	0	0	1.3	0
Asp	9.9	4	11.1	5	4.4	1	13.4	3
Thr	7.2	3	10.2	4	8.1	2	1.3	0
Ser	8.7	4	5.0	2	7.3	2	8.6	2
Glu	24.0	11	19.8	9	33.3	9	15.2	3
Pro	7.4	3	7.3	3	10.2	3	10.8	2
Gly	3.4	2	4.2	2	5.4	2	4.1	1
Ala	5.8	3	7.5	3	6.3	2	11.1	2
Val	0.7	0	1.0	0	1.0	0	1.4	0
Ile	4.2	2	4.1	2	3.6	1	6.9	1
Leu	4.9	2	6.4	3	4.0	1	4.5	1
Tyr	0.2	0	0.3	0	Trace	0	Trace	0
Phe	2.3	1	2.3	1	Trace	0	6.1	1
His	0.2	0	0.5	0	Trace	0	Trace	0
Lys	19.7	9	19.0	8	16.5	5	15.3	3
Arg	0.4	0	0.8	0	Trace	0	Trace	0
Cys	0.2	0	0.2	0	Trace	0	Trace	0
Total residues ...		44		42		28		19

spectra; see below). The u.v. spectra of both proteins AP-X and AP-Y also show the absence of tyrosine and additionally indicate that no tryptophan is present. If only one phenylalanine residue were present, then only two fragments would be expected from cleavage with α -chymotrypsin. Protein AP-X was therefore treated with α -chymotrypsin in 50 mM-potassium phosphate buffer at pH 8.0, at an enzyme/substrate ratio of 1:200 (w/w).

The products were separated on a column (0.7 cm $\times$ 130 cm) of Sephadex G-25 in 10 mM-HCl and two fragments were obtained, designated peptides AP-X1 and AP-X2. Their analyses are also given in Table 1, and the number of residues in each fragment was calculated on the basis of a single leucine residue in each. This implies 28 residues in the larger fragment AP-X1 and 19 in the smaller, AP-X2, i.e. a total of 47 amino acids, which is close to that calculated for intact protein AP-X. Since peptide AP-X2 contains the phenylalanine residue, it is the *N*-terminal fragment. The analyses of protein AP-X and peptide AP-X2 suggest the possible presence of methionine, and this is supported by the observation of a sharp resonance at 2.13 p.p.m. in the n.m.r. spectrum of both protein AP-X and the peptide AP-X2 (peak B, Fig. 4). Protein AP-X was therefore treated with CNBr in 70% (v/v) formic acid and the products were run on an SDS/polyacrylamide gel. No apparent cleavage was observed, however, so if methionine is present in protein AP-X it must be very close to the *N*-terminus.

Amidation. Inspection of the n.m.r. spectra in Fig. 4 shows that proteins AP-X and AP-Y are amidated to only a low extent at glutamic acid residues. At pH 3, the resonance of the γ -CH₂ protons of glutamic acid is at 2.5 p.p.m., and that for those of glutamine is at 2.4 p.p.m. It is clear that for both proteins the number of glutamine residues is much less than glutamic acid, particularly for protein AP-X. As regards aspartic acid, the existence of amidation is clearly seen in protein AP-Y from the peak at 2.85 p.p.m., which can be assigned to β -CH₂ groups of asparagine. The proportion of asparagine in protein AP-X is less than in protein AP-Y and is entirely in peptide AP-X2, the *N*-terminal fragment.

Nuclear magnetic resonance. Proton spectra serve a double purpose in this work, firstly to study conformation and secondly to obtain more information on chain length. The latter will be considered first. The estimates in Table 1 are based on a single phenylalanine residue in protein AP-X. This is supported by the observation of a very sharp multiplet at approx 7.3 p.p.m. in protein AP-X. If two phenylalanine residues are present, small chemical-shift differences between the two residues normally indicate this. Furthermore, the β -CH₂ of the phenylalanine can be observed in protein AP-X2 as a sharp multiplet (peak C at 3.3 p.p.m.) and in the parent protein AP-X as a somewhat broader peak to the low-field side of the lysine ϵ -CH₂ peak. Since the phenylalanine is expected to be *C*-terminal in peptide AP-X2, the fragment was titrated from pH 3.2 to 7.0. Peak C was found to move upfield from 3.28 to

Fig. 4. 270 MHz proton spectra at pH 3.0 of proteins AP-X, AP-Y and the two chymotryptic fragments of AP-X, designated peptide AP-X1 (C-terminal) and peptide AP-X2 (N-terminal)

The bottom spectrum is an 'equimolar' addition of the spectra of peptides AP-X1 and AP-X2 for comparison with intact protein AP-X. Peaks: A, α -amino *N*-acetyl-CH₃; B, methionine CH₃; C, phenylalanine β -CH₂; D, α -CH from *N*-terminal residue of peptide AP-X1.

3.20 p.p.m. over this range, as expected if it were C-terminal. In this titration range no change in peak shape was observed, and there was no sign of a β -CH₂ group from another phenylalanine, so we conclude that the numbers of residues given in Table 1 are essentially correct. It is of course possible that three fragments could have been generated by chymotryptic cleavage and a small fragment not seen in gel electrophoresis could have been lost in separations. The n.m.r. spectra of peptides AP-X1 and AP-X2 were therefore summed and the resultant found to be close to that of intact protein AP-X, i.e. no polypeptide chain appears to have been lost. The only clear differences are peak C (as mentioned

above) and peak D, which was observed to change chemical shift on titration and could therefore be due to the *N*-terminal residue of the C-terminal peptide fragment AP-X1.

A second important feature revealed by the n.m.r. spectra is a sharp resonance in AP-X and in the *N*-terminal peptide fragment AP-X2 at 2.10 p.p.m. (peak A, Fig. 4). This is the precise chemical shift observed for the methyl group of *N*-acetylserine in histone H4, and we conclude that protein AP-X is probably acetylated at its *N*-terminus. The chemical shift is not correct for *N* ^{ϵ} -acetyl-lysine (1.99 p.p.m.). Protein AP-Y has a sharp resonance in this region of the spectrum, but this is at 2.05 p.p.m. and so is

neither *N*-terminal acetyl nor *N*^ε-acetyl-lysine. It could be an impurity.

Conformation

If a solution of proteins AP-X or AP-Y is adjusted to neutral pH and NaCl added to 1 M, no change is observed in the n.m.r. spectrum except ionization-induced changes in the glutamic acid and aspartic acid resonances. In particular, no ring-current-shifted methyl-group peaks appear upfield of 1 p.p.m., and no perturbations of the single aromatic resonance are observed. This implies that no tertiary-structure folding occurs.

Circular dichroism. When measured in 100 mM-potassium phosphate buffer at pH 6.5, both proteins AP-X and AP-Y exhibit an ellipticity at 222 nm of about $-3400 \text{ deg} \cdot \text{cm}^2 \cdot \text{dmol}^{-1}$. On the basis of a random coil ellipticity of -1000° (Moss *et al.*, 1976) and an ellipticity of -30000° for 100% helix at this wavelength (Chen *et al.*, 1974), this indicates about 8% helix. Curve fitting as in Moss *et al.* (1976), indicates 9.5% α -helix. These results, taken at face value, mean four helical residues, the absolute minimum possible for a single turn of helix. There was no evidence of significant β -sheet structure. Bearing in mind that there is no evidence of tertiary structure formation, we conclude that proteins AP-X and AP-Y are in fact disordered proteins in free solution, in the way previously demonstrated for proteins HMG 14 and HMG 17 (Abercrombie *et al.*, 1978; Cary *et al.*, 1980).

Are proteins AP-X and AP-Y nuclear or cytoplasmic?

If a pure preparation of calf thymus nuclei is obtained by repeated washing of nuclei in a relatively high-ionic-strength medium containing only a low concentration of sucrose (10 mM-MgCl₂, 50 mM-NaHSO₃, 10 mM-Tris/HCl, 0.25 M-sucrose), then 5% HClO₄ is able to extract only histone H1 and very low amounts of protein HMG 1. The remainder of the non-histone proteins are found in the washing solution. A low-ionic-strength/high-sucrose medium was therefore used to try to retain non-histone proteins in the nucleus. Thymus was blended for 3 min in 4 mM-MgCl₂/2.2 M-sucrose/1 mM-phenylmethanesulphonyl fluoride at 4°C (solution A). The fragmented suspension containing intact nuclei was twice passed through cheesecloth, layered on top of solution A and centrifuged in a swing-out rotor at 40000g for 1 h at 4°C. The nuclear pellet was resuspended in solution A and once again centrifuged through solution A as above. Observation of the final pellet in the light microscope indicated that a very pure nuclear preparation had been obtained. These nuclei were extracted with 5% HClO₄ and the proteins obtained compared with those obtained by 5%-HClO₄ extraction of whole cells (as described

above for protein AP-X preparation). Fig. 5 shows a gel giving this comparison. Since proteins AP-X and AP-Y are present in relatively low amounts, two loadings are used for nuclei (lanes 1 and 2) and two for whole cells (lanes 4 and 5). The massive loadings in lanes 1 and 4 result in distorted band shapes, but comparison with the adjacent lanes 2 and 5 permits identification of bands. Protein AP-X can readily be identified in lane 5, from the standards in lanes 3 and 7, and also in lane 2. The distorted bands in lanes 1 and 4 run very slightly slower than in lanes 2 and 5, but those corresponding to AP-X can nevertheless be identified. Protein AP-Y is present in much lower amounts than protein AP-X, and it is not possible to draw any firm conclusions on its cellular distribution. The gel in Fig. 5 shows that, under the conditions of extraction used, some protein AP-X is found in the nucleus. Since the band corresponding to protein AP-X is much more intense in the whole-cell extract than in the nuclear extract, it could be that AP-X is also located in the cytoplasm. However, the yield from nuclei may be much decreased by leakage during their preparation, and so it could also be that protein AP-X is entirely nuclear in origin, i.e. no definitive statement on its cytoplasmic presence can be made. In this respect it is very much like ubiquitin, which is also readily identified in Fig. 5 and is likewise best prepared from whole cells (Walker *et al.*, 1978). Ubiquitin is also present in the nucleus (lane 1) like protein AP-X, but in a somewhat lower amount (lane 2). The con-

Fig. 5. Comparison of 5%-HClO₄ extracts of calf thymus whole cells and nuclei (washed with a low-ionic-strength medium)

Lanes 1 and 2 are from nuclei, loadings $\times 3$ and $\times 1$, and lanes 4 and 5 are similarly from cells. Abbreviations: UB, ubiquitin; H1, histone H1.

clusion is that protein AP-X is located, at least partially, in the calf thymus nucleus.

Discussion

Protein AP-X is present in calf thymus cells in an amount somewhat below that of the HMG proteins (see Fig. 5), whereas the amount of protein AP-Y is only 10–20% that of protein AP-X. The extraction procedures used here suggest them to be located in both nucleus and cytoplasm, but since non-histone nuclear proteins readily leak out from the nucleus during purification of nuclei, it is possible that they are solely nuclear proteins. They both consist of about 45 residues, of which approx. 34% are acidic (largely non-amidated) and approx. 20% basic (lysine). They are thus acidic proteins. Could they be degradation products of other proteins? Inspection of the sequences of the histones, the HMG proteins 1, 2, 14, 17, T and protein H6 shows that nowhere is there a run of some 44 amino acids with the composition determined for proteins AP-X or AP-Y. Since the amount of protein AP-X is at least 10% that of the HMG proteins, one would expect any precursor to be present in quantities typical for the HMG proteins (or above). Since proteins AP-X and AP-Y have been prepared from very fresh thymus, it is highly likely that they are genuine proteins. It is just possible that protein AP-Y is a degradation product of protein AP-X, for example by loss of *N*-terminal acetyl methionine, but the increased amounts of threonine and aspartic acid in protein AP-Y make this very unlikely.

Both proteins AP-X and AP-Y do not seem able to take up secondary or tertiary structure in free solution, as expected from the low content of hydrophobic residues. In this they resemble proteins HMG 14 and HMG 17. Similarity to proteins HMG 14 and HMG 17 lies also in the very high content of

acidic and basic residues and the low molecular weight, though whether there is any true homology must await sequence determination. The composition of the *N*-terminal 40% of protein AP-X (i.e. peptide AP-X2) and the *C*-terminal 60% (i.e. peptide AP-X1) indicate a fairly even distribution of acidic and basic residues in protein AP-X, unlike proteins HMG 14 and HMG 17, which have somewhat asymmetric sequences. Proteins AP-X and AP-Y appear to be *N*-terminally acetylated, which is a feature of certain histones (H4, H2A and H1) and not of the HMG proteins. Thus, although we are reluctant at present to designate proteins AP-X and AP-Y as 'high mobility group' (HMG; Goodwin *et al.*, 1973), they do appear similar to proteins HMG 14 and HMG 17. Since they have no conformation of their own, they must acquire one *in vivo* by binding to other components, either protein or nucleic acid.

References

- Abercrombie, B. D., Kneale, G. G., Crane-Robinson, C., Bradbury, E. M., Goodwin, G. H., Walker, J. M. & Johns, E. W. (1978) *Eur. J. Biochem.* **84**, 173–177
- Cary, P. D., King, D. S., Crane-Robinson, C., Bradbury, E. M., Rabbani, A., Goodwin, G. H. & Johns, E. W. (1980) *Eur. J. Biochem.* **112**, 577–588
- Chen, Y. H., Yang, J. T. & Chau, K. H. (1974) *Biochemistry* **13**, 3350–3359
- Goodwin, G. H., Sanders, C. & Johns, E. W. (1973) *Eur. J. Biochem.* **38**, 14–19
- Goodwin, G. H. & Johns, E. W. (1977) *Methods Cell Biol.* **6**, 257–267
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- Moss, T., Cary, P. D., Abercrombie, B. D., Crane-Robinson, C. & Bradbury, E. M. (1976) *Eur. J. Biochem.* **71**, 337–350
- Sanders, C. & Johns, E. W. (1974) *Biochem. Soc. Trans.* **2**, 547–550
- Walker, J. M., Goodwin, G. H. & Johns, E. W. (1978) *FEBS Lett.* **90**, 327–330