PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Aspects of the Regulation of Histone Genes

M. L. Birnstiel, A. Kressmann, W. Schaffner, R. Portmann and M. Busslinger

Phil. Trans. R. Soc. Lond. B 1978 **283**, 319-324
doi: 10.1098/rstb.1978.0031

Phil. Trans. R. Soc. Lond. B. **283**, 319–324 (1978) [319]
Printed in Great Britain

Aspects of the regulation of histone genes

By M. L. Birnstiel, A. Kressmann, W. Schaffner, R. Portmann and M. Busslinger
Institut für Molekularbiologie II der Universität Zürich, Winterthurerstrasse 266*A*, 8057 *Zürich, Switzerland*

[Plate 1]

Sequencing of cloned histone DNA of the sea urchin *Psammechinus miliaris* has confirmed the map of the histone genes obtained earlier by rather less refined techniques. Sequencing of spacer has revealed that it is unlikely to code for protein. Some interesting sequences in the prelude regions to the structural genes have been found. The technique of injecting DNA into the germinal vesicle of the *Xenopus* oocyte has been greatly simplified, so that now many of the parameters governing the transcription of the injected genes can be investigated. Some mRNA-like molecules appear when circular histone DNA is inserted into the oocyte nucleus. We are cautiously optimistic that the technique can be further developed and will provide a useful tool for the study of the molecular mechanisms governing the expression of structural genes coding for proteins.

Introduction

Those biologists who have adopted a biochemical and quantitative approach to the study of development have for many years endeavoured to relate differentiation to changes in the protein pattern. Among the most interesting molecular markers for development are those gene products which derive from a series of similar structural genes such as those for the haemoglobins and for isozymes, which may be used to define states of differentiation. Nowadays, we should like to know what is the nature of the regulatory sequences adjacent to the structural genes which allow them to be expressed in this specific way.

Histone variants in the sea urchins

Now, since it has been established that histone variants exist, at least in some species, the histones may also be considered as a family of related proteins carrying out similar functions in chromatin structure. In the sea urchin, where the histone genes are highly redundant, the polymorphism of the histone genes has been exploited during development in that specific histone variants are often associated with specific tissues. These histone variants may well represent 'flaws' in a crystal to the crystallographers, but for us they may offer a real opportunity to learn more about development.

A case in point is the change in the H2B histone component during early development in the sea urchin. As shown by Cohen, Newrock & Zweidler (1975) the cleaving egg synthesizes first α-type H2B, followed by the production of γ- and δ-type H2B in the gastrula. It has been further shown by amino acid sequence analysis that gastrula H2B differs again from sperm H2B, both in the number of amino acid moieties and in sequence. Thus sperm H2B contains

143 amino acids compared to the 125 in embryonic H2B and there are frequent amino acid substitutions, especially at the N-terminal end (Strickland *et al.* 1977).

In our own sequencing work on cloned histone DNA (Clarkson *et al.* 1976) we have found that the H2A of the commonest type of histone clone differs in at least 10 amino acid positions from the known H2A protein sequence of gonads from the same species (Schaffner *et al.* 1978, in the press). The considerable divergence in protein sequence exhibited by these variants in a single species must mean that the genes for the histone variants must have been separated in evolution for quite some time from the so-called standard histones.

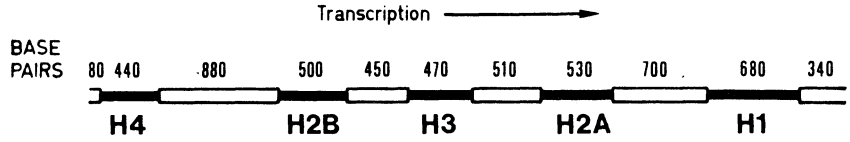


FIGURE 1. Structural map of the histone gene repeat unit cloned in a λ-vector. The distribution of spacers and genes is shown in a schematic way. The respective lengths of the various DNA segments have been noted in base pairs.

CLONING OF HISTONE DNA OF THE SEA URCHIN

The simplest explanation that can be offered for the large redundancy of histone genes is that it arises from the need for very rapid synthesis of histone protein during cleavage in the sea urchin. If so, cloned histone DNA should consist predominantly of one major type of histone genes, namely those coding for the sequence produced during cleavage. In collaboration with Clarkson *et al.* (1976) we have produced 68 independent histone DNA clones, each clone consisting of a DNA fragment about 6000 base pairs long which may be assumed to contain a set of the five different histone genes. Restriction analysis has shown that about three-quarters of the clones tested show an identical restriction pattern whereas the remaining quarter show considerable differences from the majority class and in addition differ among themselves (unpublished result). These histone clones then are clearly divergent in sequence from the majority class and are good candidates for harbouring histone variant genes. This idea is further strengthened by the fact that while the majority class of cloned DNA hybridizes with good base fidelity to histone mRNA from cleaving embryos, the minority class contains clones which hybridize only poorly to this same cleavage histone mRNA and therefore must also be divergent in their gene sequences. Since some histone variants show tissue specificity in the sea urchin it is clear that they may be used as markers to define differentiated states of at least some of these tissues. Thus histones may represent a case similar to that of the 'isozymes'. In order to understand how the regulation of these histone genes is encoded in the DNA we must first compare the histone genes with one another to determine what forms a common theme among all histone genes and what is associated with tissue specific expression.

MOLECULAR STRUCTURE OF THE HISTONE DNA REPEAT UNIT

From our previous work (Schaffner, Gross, Telford & Birnstiel 1976) it is known that the coding sequence for the histones in the sea urchin *Psammechinus* are clustered together in a repeat unit of 6×10^3 bases in the order shown in figure 1.

All coding sequences lie on the same DNA strand and transcription is known to proceed from left to right (Gross, Probst, Schaffner & Birnstiel 1976). The distribution and length of

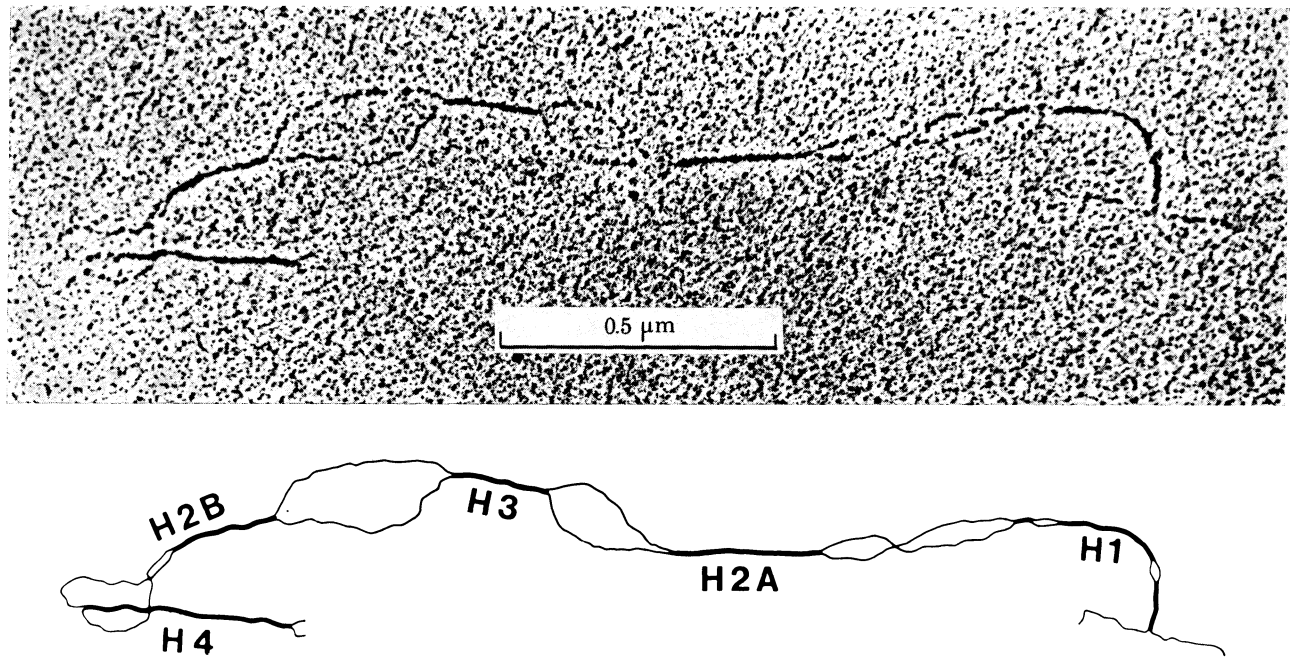


FIGURE 2. Electron micrograph of a partly denatured histone DNA molecule reclaimed from a λ recombinant. The spacers intervening between the mRNA-coding segments are AT-rich and hence melt at low temperature. The assignment of the double-stranded regions to specific mRNAs was made on the basis of an extensive biochemical analysis. The identity, map position and polarity of the coding sequences have been confirmed unequivocally by sequencing of the histone DNA.

(*Facing p.* 320)

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

the spacers intervening between coding sequences can be best demonstrated by partial denaturation mapping of histone DNA (cf. figure 2, plate 1) reclaimed from λ recombinants (Portmann, Schaffner & Birnstiel 1976). This technique exploits the fact that spacer is AT-rich and melts at lower temperature than the high GC coding sequences (figure 2).

We are now in the process of sequencing parts of the histone repeat unit of the majority type and about one third of the total sequence is known at present. The techniques which were used, i.e. those of Maxam & Gilbert (1977) and Sanger & Coulson (1976), rely on the availability of restriction sites with defined topologies, and over 100 such sites have been accurately mapped (unpublished results). We have sequenced sections (Birnstiel, Schaffner & Smith 1977*a*), some quite large, from all five genes and this has enabled us to provide irrefutable evidence that the identification, the polarity and the map positions of the histone genes obtained in our previous experiments are correct.

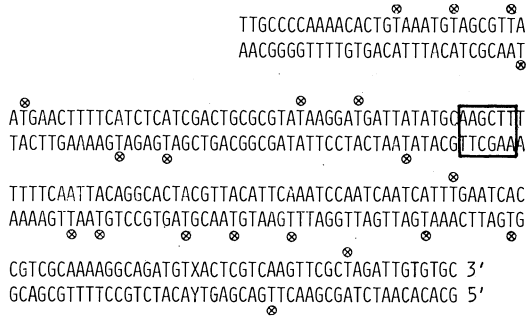


FIGURE 3. Partial sequence of the *Hin*dIII spacer. Starting from the *Hin*dIII sites at the ends of the histone DNA molecules, the DNA sequence of part of this spacer intervening between H1 and H4 was determined. There are no short repeats in this spacer region. If transcribed into RNA in either direction, there would be many stop codons (indicated with ⊗) in all three reading frames which would terminate translation; hence, this spacer segment would not be expected to code for any proteins. The boxed sequence is the *Hin*dIII restriction site.

In addition to the information obtained about structural genes the DNA sequencing enables us to ask three questions about spacer DNA:

(*a*) Are the spacer sequences between genes different from one another?

(*b*) Do they contain short repeat sequences as other spacers from other genes do?

(*c*) Could they possibly code for a protein?

From the data we have so far obtained on the spacer regions intervening between the H1 and the H4 and between the H4 and the H3 genes it is clear that these spacers are different from one another, that they contain no short repeats. If transcribed in either direction they would contain such a great number of stop signals in all reading frames that no protein of reasonable length would be produced (figure 3).

The prime object of sequencing as far as we are concerned is the study of regulation and from this standpoint the most interesting regions are the interphases between spacers and genes where we would expect to find the signals for regulation of transcription and translation. Such regulatory sequences would be expected to include the leading and the trailing sequence of the mRNAs. Furthermore, they would either contain five promotors and/or termination signals if there were monocistronic transcription of the genes, or one promotor and terminator plus many maturation signals for cleavage of the precursor RNA if there were a polycistronic transcriptional unit. Since such signals would form a recurrent theme within the DNA we would expect to find repeated similar sequences occurring 'upstream' from the structural genes.

A comparison of the sequences adjacent to the initiator codons of the H1 and the H2A genes clearly shows extensive sequence homology (figure 4).

Similarly, the sequences lying approximately 100 nucleotides further 'upstream' from the initiator codon of H4 and H2A also show sequence homology (Schaffner *et al.* 1978, in the press).

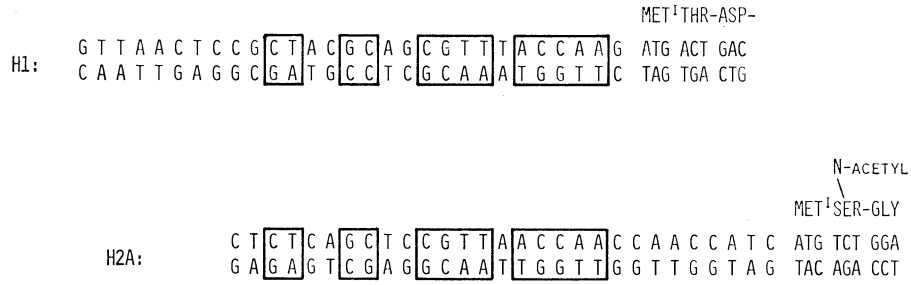


FIGURE 4. Prelude sequences to the H1 and H2A genes. Upstream from the initiation codon of the H1 and the H2A structural genes sequences exhibiting great similarity have been noted. Such sequences, if transcribed, could represent the ribosomal attachment sites for the H1 and the H2A mRNAs.

INVESTIGATION OF CLONED HISTONE DNA BY A NOVEL OOCYTE NUCLEUS INJECTION TECHNIQUE

It is of course very tempting to speculate that the sequences in these prelude regions adjacent to the initiator codons represent ribosome attachment signals. However, this can only be substantiated by experiments of the type in which we destroy or alter the sequence by some means and see what effect if any this has on transcription and translation.

It would also be extremely difficult at the moment to guess at the function of the rather conserved sequence further 'upstream' from the coding sequence. However, it is quite clear in the present area of genetical engineering that in order to understand these sequences we must look at their expression in a cell nucleus into which the histone DNA has been transported either by a viral vector or by mechanical injection.

In our own nuclear injection experiments we were of course spurred on by the successes of Gurdon's group (see this Symposium) which showed that the DNA molecules such as ϕX, plasmids, animal viruses and *Drosophila* histone DNA were extensively transcribed after injection into the oocyte nucleus.

The problem facing would-be egg-nucleus injectors is to locate and inject the DNA into the nucleus of an oocyte which is opaque and full of yolk. The breakthrough in this respect came for us about 9 months ago when we developed a technique for egg nucleus injection (Kressman *et al.* 1977). When a *Xenopus* oocyte is centrifuged at low speed the nucleus floats up towards the surface of the oocyte and in doing so displaces pigment granules at the surface of the oocyte. What is then seen under the microscope is a dark ring of pigment with a lighter area inside which delimits the boundaries of the nucleus lying immediately underneath. If one pierces the lighter area just under the surface of the oocyte one can be sure that one has entered the nucleus. After centrifugation oocytes are apparently unimpaired and continue to synthesize RNA in a linear fashion for days.

Oocyte nucleus injection techniques would be of a special value if it were possible to take any cloned gene and transcribe and translate it in a specific way inside the living oocyte. It is

clear that in the case of the 5*S* DNA of the *Xenopus* species the results are impressive (D. D. Brown & J. B. Gurdon, unpublished results; see also this symposium) and there is specificity of transcription. We have been studying the transcription of cloned sea urchin histone DNA in the *Xenopus* oocyte in the full knowledge that we are now crossing the species boundary as well as studying a gene which is presumably transcribed by a polymerase much less abundant than the polymerase C thought to transcribe the 5*S* DNA.

The first question one can ask is: is the histone DNA transcribed at all? To answer this question, RNA synthesized in the presence and absence of histone DNA is hybridized to cloned histone DNA. The answer obtained is identical to that obtained by the Gurdon group for a variety of heterologous DNAs, i.e. about 5% of all cellular transcripts can back-hybridize to the template DNA. There is also a stimulation of total cellular RNA synthesis by a factor of about 2 as a consequence of DNA injection.

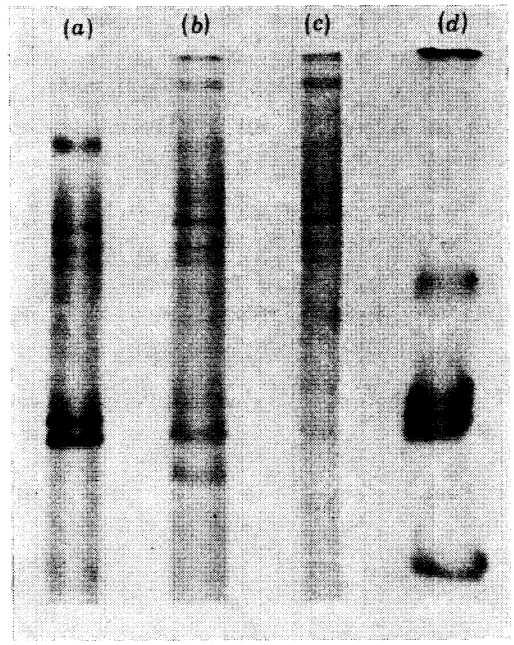


FIGURE 5. RNA molecules synthesized on sea urchin histone DNA injected into the oocyte nucleus of *Xenopus*. A fluorograph of an electroferogram of labelled RNA molecules synthesized by the oocyte nucleus is shown (*a*) after injection of histone DNA circularized by ligation, and (*b*) after injection of linear histone DNA molecules; (*c*) pattern obtained after injection of label only (no DNA); (*d*) pattern of histone mRNA obtained from pulse labelled sea urchin embryos.

The second question one can ask is: are there histone mRNA-like molecules formed in the injected oocytes? For this question we have injected linear histone DNA molecules as well as histone DNA circularized by ligation and compared the RNA electrophoresis pattern to that of the sea urchin 9*S* histone mRNA on the one hand and to the RNA pattern of the oocyte injected with Simian virus 40 (SV40) DNA or just with water.

The RNA produced in the presence of circular histone DNA contains two bands co-migrating with the messengers H2A and H2B; however, no RNAs coincident with H4, H3 or H1 can be detected (Kressman *et al.* 1977). Linear molecules also give rise to two types of bands, one co-migrating with H2A and one band which moves faster than the standard H2A band (figure 5).

From these experiments it is evident that while some mRNA-like molecules are made, not all the five histone messengers appear. In addition it is also clear that the configuration of the DNA itself, whether linear or circular, influences the type of product which is made. We know from many experiments that the picture that we present here is not the full story, especially in the case of the linear DNAs which we have investigated rather thoroughly. In the case of linears we know that there is also a large amount of low molecular weight RNA formed of the 4*S* type. This low molecular weight RNA must have arisen from rather unspecific transcription of histone DNA since it back-hybridizes to both sense and anti-sense strand of the histone DNA with relatively little discrimination. So we further conclude that when working with linears there is some unspecific background arising from transcription of heterologous DNAs.

References (Birnstiel *et al.*)

Birnstiel, M. L., Schaffner, W. & Smith, H. O. 1977*a* *Nature, Lond.* **266**, 603–607.

Kressman, A., Clarkson, S. G., Telford, J. & Birnstiel, M. L. 1977 *Cold Spring Harb. Symp. quant. Biol.* (In the press.)

Clarkson, S. G., Smith, H. O., Schaffner, W., Gross, K. W. & Birnstiel, M. L. 1976 *Nucl. Acids Res.* **3**, 2617–2631.

Cohen, L., Newrock, K. & Zweidler, A. 1975 *Science, N.Y.* **190**, 994–997.

Gross, K., Probst, E., Schaffner, W. & Birnstiel, M. 1976 *Cell* **8**, 455–469.

Maxam, A. M. & Gilbert, W. 1977 *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564.

Portmann, R., Schaffner, W. & Birnstiel, M. 1976 *Nature, Lond.* **264**, 31–34.

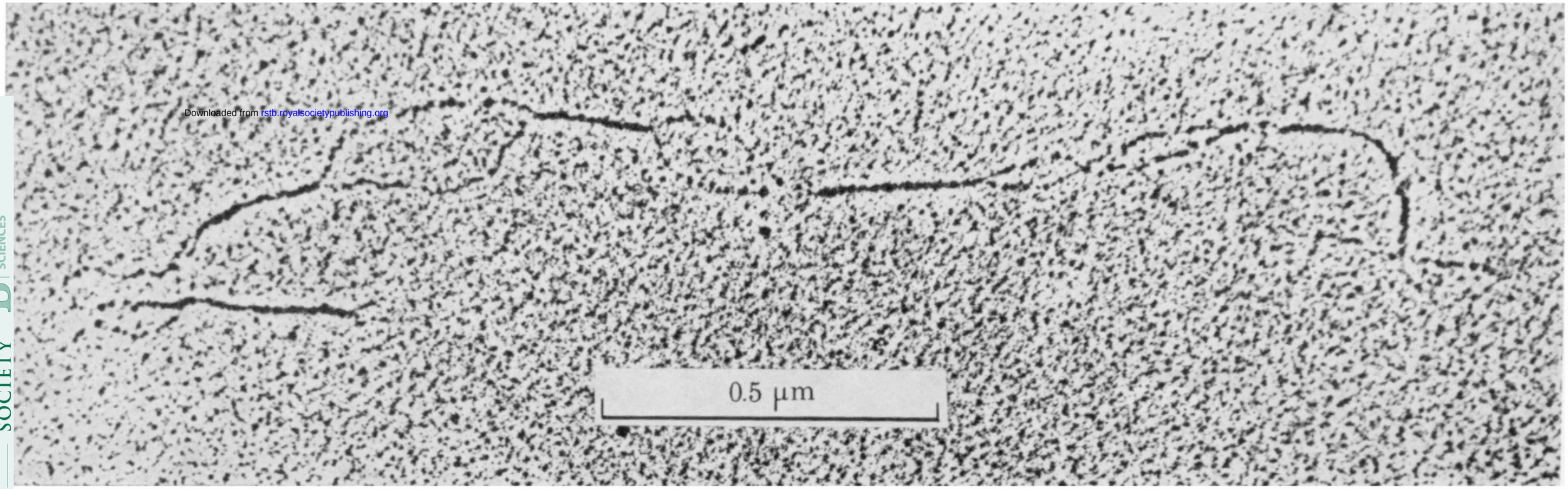
Sanger, F. & Coulson, A. R. 1975 *J. molec. Biol.* **94**, 441–448.

Schaffner, W., Gross, K., Telford, J. & Birnstiel, M. 1976 *Cell* **8**, 471–478.

Schaffner, W., Smith, H. O. & Birnstiel, M. L. 1978 *Cell*. (In the press.)

Strickland, M., Strickland, N., Brandt, W. F. & von Holt, C. 1977 *Eur. J. Biochem.* **77**, 263–275.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES



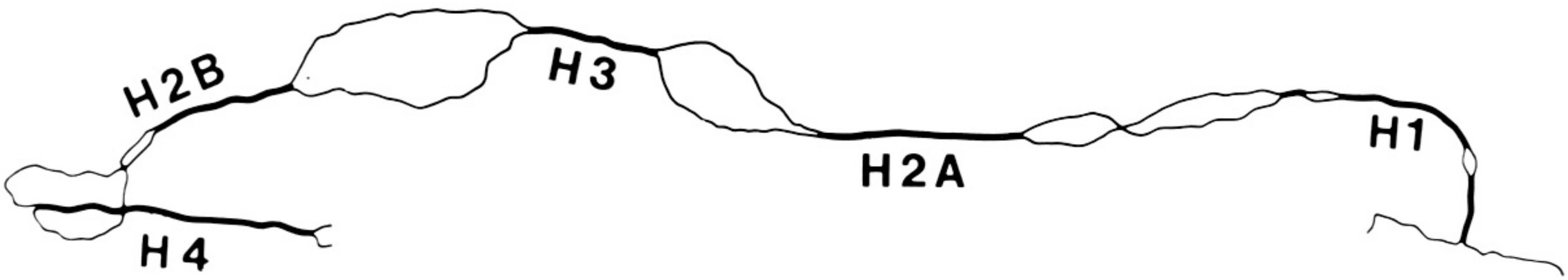


FIGURE 2. Electron micrograph of a partly denatured histone DNA molecule reclaimed from a λ recombinant. The spacers intervening between the mRNA-coding segments are AT-rich and hence melt at low temperature. The assignment of the double-stranded regions to specific mRNAs was made on the basis of an extensive biochemical analysis. The identity, map position and polarity of the coding sequences have been confirmed unequivocally by sequencing of the histone DNA.

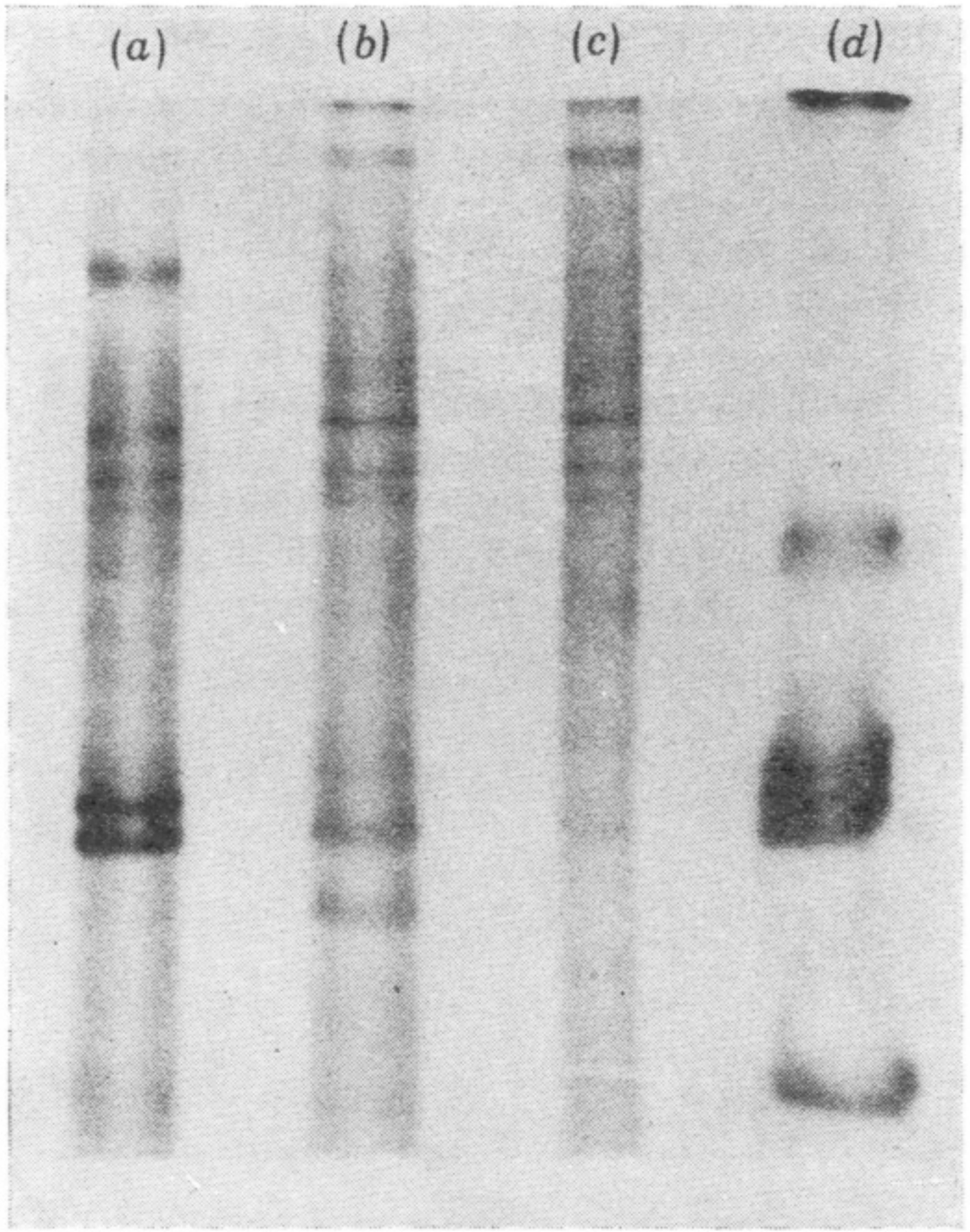


FIGURE 5. RNA molecules synthesized on sea urchin histone DNA injected into the oocyte nucleus of *Xenopus*. A fluorograph of an electroferogram of labelled RNA molecules synthesized by the oocyte nucleus is shown (*a*) after injection of histone DNA circularized by ligation, and (*b*) after injection of linear histone DNA molecules; (*c*) pattern obtained after injection of label only (no DNA); (*d*) pattern of histone mRNA obtained from pulse labelled sea urchin embryos.