Isolation of a clone containing human histone genes

Susan J.Clark, Paul A.Krieg and Julian R.E.Wells

Department of Biochemistry, The University of Adelaide, Adelaide, South Australia 5001

Received 3 February 1981

ABSTRACT

A recombinant clone containing human histone genes has been isolated. The clone, λ HH-01, was selected from a genomic library using chicken histone cDNA and a cloned fragment containing chicken histone genes as probes. Sub-clones from λ HH-01 have been mapped and coding regions located with cDNA. The human H3 gene has been identified by DNA sequence analysis.

INTRODUCTION

The organisation of histone genes has been studied extensively in several invertebrate systems, notably sea urchin species (1) and *Drosophila* (2). In these cases a repeating unit contains each of the five histone genes interspersed with A-T-rich 'spacer' DNA. In yeast (3), the two copies of linked H2A-H2B genes are widely separated from each other and from H3 and H4 genes.

For vertebrates the information is limited. For the chicken system, despite our initial indications of a simple repeat (4) more recent data from a recombinant histone gene clone indicate that this is not the case (5,6). In *Xenopus*, there is a suggestion that histone genes may occur in a simple repeat (7).

In contrast to the lack of data on vertebrate histone genes, there is much known about histone proteins and their interaction with DNA (8). Extensive studies have suggested co-ordination of histone and DNA synthesis in S-phase of the cell cycle (9,10,11 for reviews). This view has now been challenged by Groppi and Coffino (12) who find that histones

are synthesised at equivalent rates in G1 and S phases in two mammalian cell lines. The analysis of histone gene transcription in mammalian cells (13), and its relationship to the cell cycle (14), will clearly benefit from the use of mammalian histone gene probes. Here we report the isolation and partial characterisation of a recombinant clone containing human histone genes.

MATERIALS AND METHODS

Selection of Human Histone Gene Recombinant

The human "library" was a generous gift of Dr. T. Maniatis (California Institute of Technology) and was constructed from a partial HaeIII and AluI digest of human DNA, ligated to EcoRI linkers and inserted into λ Charon4A as described (15).

10^5 Plaques were screened by the Benton and Davis procedure (16). Those plaques which were positive for enriched chicken histone cDNA (4) were re-screened with a cloned fragment from the chicken histone clone λ CH-01 (5). The 3.3 Kb insert in pBR325 contains chicken H2A and H2B gene sequences (5,6). One recombinant, λ HH-01, was chosen for further characterisation.

Sub-clones of λ HH-01

To enable ready amplification of required portions of λ HH-01, fragments were sub-cloned into plasmid vectors by standard methodology. Appropriate restriction enzyme fragments (see Fig. 1 for details) were ligated to DNA from pBR322 or pBR325. The plasmid DNA was linearised with restriction enzyme(s) and de-phosphorylated prior to the ligation reaction (17). After transformation of *E. coli* K12, strain E392, sub-clones were selected by antibiotic sensitivity, and by colony hybridisation (18) with chicken histone cDNA.

All experiments involving recombinants and viable organisms were carried out in a certified C₃ laboratory in Adelaide as recommended by ASCORD (Australian Academy of Science Committee on Recombinant DNA).

RESULTS AND DISCUSSION

The recombinant clone λ HH-01 was selected from a library using two probes, chicken histone cDNA and a 3.3 Kb cloned fragment containing chicken H2A and H2B genes (5,6). By inference it is likely that λ HH-01 contains H2A and/or H2B genes. The purpose of this report is to show a map of the insert DNA, to indicate coding regions with cDNA and finally to show by DNA sequencing that histone genes are present in the human recombinant clone.

Mapping data for the 14.5 Kb insert is shown in Fig. 1. The expanded regions are emphasised because they were subsequently shown to contain coding regions and sub-cloned (see

FIG. 1

Restriction endonuclease map of λ HH-01 insert DNA. Map distances are shown in kilobase pairs. Restriction sites are shown as E = EcoRI, H = HindIII, Sm = SmaI, Sl = SalI, Sc = SacII, P = PvuII and B = BamI. E^L denotes the boundaries at the insert which was constructed with EcoRI linkers.

Fig. 2). The lack of a repeating restriction pattern in λ HH-01 suggests that human histone genes may not be organised in a simple repeating unit of the size found in the invertebrate systems (1,2) or in *Xenopus* (7).

Digestion of λ HH-01 DNA with EcoRI and HindIII results in four major insert fragments of 1.6, 2.6, 4.3 and 4.8 Kb (Fig. 1). As shown in Fig. 2(A) when these fragments are separated by gel electrophoresis, transferred to nitrocellulose (19) and probed with chicken histone cDNA (4), the three larger insert segments hybridise.

To investigate the organisation of the λ HH-01 clone in more detail these three coding fragments indicated in Fig. 2(A) were recloned into plasmid vectors to give the sub-clones; pHu2.6H, pHu4.3EH and pHu4.8E.

FIG. 2

Detection of histone coding sequences in λ HH-01. DNA from digests of λ HH-01 and sub-clones derived from it electrophoresed on 1% agarose gels was detected with ethidium bromide and transferred to nitrocellulose filters (18). The size of the bands was calculated from the migration of markers run in the same gel. Autoradiograms of the filters hybridised with chicken histone cDNA are shown alongside the respective gel tracks.

Track (A) EcoRI HindIII double digest of λ HH-01.
Track (B) SacII HindIII double digest of pHu2.6H.
Track (C) PvuII EcoRI double digest of pHu4.3EH.
Track (D) SmaI EcoRI double digest of pHu4.8E,

Digests of each sub-clone were probed with histone cDNA and the results are shown in Fig. 2(B,C,D). In pHu2.6H, cDNA hybridises to the HindIII/SacII fragment 0.57 Kb long and to the SacII fragment of 1.4 Kb. In pHu4.3EH, cDNA hybridises to the 0.6 and 0.9 Kb PvuII fragments. In pHu4.8E, cDNA hybridises to the SmaI fragments 1.4 Kb and 1.8 Kb long. The relative positions of these fragments containing coding regions is shown with respect to EcoRI and HindIII sites in λ HH-01 (Fig. 3). The data indicate a minimum of three coding regions in λ HH-01, and a minimum non-transcribed spacer region of 3.5 Kb. A large spacer region is also seen in a chicken genomic histone gene clone (5).

Of the regions shown to contain coding sequences by cDNA hybridisation (Fig. 2), the most convenient for DNA sequencing studies was the 0.57 Kb HindIII/SacII fragment from pHu2.6H (Fig. 2(B)). This band was isolated, dephosphorylated, kinased with γ - ^{32}P -ATP, and subjected to the chemical degradation reactions of Maxam and Gilbert (20). A partial sequence of this

FIG. 3

Organisation of coding regions in human histone gene sub-clones. The EcoRI/HindIII restriction map of λ HH-01 insert of 14.5 Kb is shown with the sub-clones pHu2.6H, pHu4.3EH, pHu4.8E marked. The boxed regions correspond to the restriction fragments shown in Fig. 2(B,C,D) which hybridise to chicken histone cDNA. Divisions within these boxes indicate fragments which may represent either a single coding region spanning a restriction site or a region containing more than one histone gene.

region is shown in Fig. 4. It includes 160 bases to the 5' side and 78 bases to the 3' side of the initiation codon. The coding sequence identifies this region as human histone H3. The twenty-six amino acids predicted from the sequence are identical to those of H3 histone from sea urchin (21) chicken (22) and calf (23) in agreement with the highly conserved nature of the protein.

FIG. 4.

DNA sequence from the HindIII/SacII 0.57 Kb fragment of pHu2.6H. This predicts the amino acid sequence for human histone H3. The nucleotide sequence for sea urchin (*Psammechinus miliaris*) (21), is shown below the λ HH-01 sequence. Nucleotides in the 5' flanking sequences are indicated by negative integers relative to the initiation codon and the amino acids of the structural gene are numbered in italics. Boxed regions indicate conserved 5' prelude sequences and stars indicate differences between human and sea urchin sequences within the coding region.

In the 5' flanking region there are sequence motifs related to conserved sequences found in sea urchin histone genes and other polymerase II genes (21). These include a pentameric sequence GTTTC, closely followed by a "Hogness box", TATAAATAA and a possible "cap site" CTACTTT (Fig. 4). It is interesting to note that the cap site shows closer homology to vertebrate polymerase II genes than to the equivalent sea urchin histone 5' region (21).

A comparison of codon usage for the first 26 amino acids of human histone H3 and *P. miliaris* H3 (21) is also shown in Fig. 4. Despite the absolute conservation of amino acid sequence, there are third base changes in 19 of the 26 codons with strong preference for G as third base residue in the human gene. Such divergence, if maintained in remaining regions of the gene, would place restrictions on the use of sea urchin H3 gene probes to locate human H3 genes.

The clone λ HH-01 has been shown to contain extensive histone coding regions interspersed with non-coding spacer regions. Within the coding region the histone gene for H3 has been identified by sequence analysis. Current work is directed at identifying the other individual histone genes contained in the clone.

ACKNOWLEDGMENTS

We thank J. Dinan and S. Bruschi for technical assistance and R. Harvey for helpful discussion. We are indebted to T. Maniatis for the gift of the human library.

REFERENCES

1. Kedes, L.H. (1979). *Ann. Rev. Biochem.* **48**, 837-870.
 2. Lifton, R.P., Goldberg, M.L., Karp, R.W. and Hogness, D.S. (1977). *Cold Spring Harbour Symposia on Quantitative Biology* **42**, 1047-1052.
 3. Hereford, L., Fahrner, K., Woolford, J., Rosbach, M. and Kabach, D.B. (1979). *Cell* **18**, 1261-1271.
 4. Crawford, R.J., Krieg, P.A., Harvey, R.P., Hewish, D.A. and Wells, J.R.E. (1979). *Nature* **279**, 132-136.
 5. Harvey, R.P. and Wells, J.R.E. (1979). *Nucleic Acids Res.* **7**, 1787-1798.
 6. Harvey, R.P., Krieg, P.A., Coles, L.S. and Wells, J.R.E. (1981). In preparation.
-

7. Moorman, A.F.M., de Laaf, R.T.M., Destreé, O.H.J., Telford, J. and Birnstiel, M.L. (1980). *Gene* **10**, 185-193.
8. Kornberg, R.D. (1977). *Ann. Rev. Biochem.* **46**, 931-954.
9. Borun, T.W. (1975). In *Results and Problems in Cell Differentiation 7, Cell Cycle and Cell Differentiation*, J. Reinert and H. Holtzer, eds. (New York: Springer-Verlag) pp. 249-290.
10. Elgin, S.C.R. and Weintraub, H. (1975). *Ann. Rev. Biochem.* **44**, 725-774.
11. Pederson, T. (1976). In *Protein Synthesis 2*, E.H. McConkey, ed. (New York: Marcel Dekker) pp. 69-123.
12. Groppi, V.E. and Coffino, P. (1980). *Cell* **21**, 195-204.
13. Melli, M., Spinelli, G., Wyssling, H. and Arnold, E. (1977). *Cell* **11**, 651-661.
14. Stein, G., Stein, J., Shephard, E., Park, W. and Philipps, I. (1977). *Biochem. Biophys. Res. Commun.* **77**, 245-252.
15. Maniatis, T., Hardison, R.C., Lacy, E., Laver, J., O'Connell, C., Quon, D., Sim, G.K. and Efstradiatis, A. (1978). *Cell* **15**, 687-701.
16. Benton, W.D. and Davis, R.W. (1977). *Science* **196**, 180-182.
17. Ulrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W.J. and Goodman, H.M. (1977). *Science* **196**, 1313-1319.
18. Grunstein, M.L. and Hogness, D.S. (1975). *Proc. Natl. Acad. Sci. USA* **72**, 3961-3965.
19. Southern, E.M. (1975). *J. Mol. Biol.* **98**, 503-517.
20. Maxam, A.M. and Gilbert, W. (1979). *Methods in Enzymology*, Grossman, L. and Moldave, K., eds. Vol. 65, part I, pp. 499-560, Academic Press, N.Y.
21. Busslinger, M., Portmann, R., Irminger, J.C. and Birnstiel, M.L. (1980). *Nucleic Acids Res.* **8**, 957-977.
22. Brandt, W.F. and van Holt, C. (1974). *Eur. J. Biochem.* **46**, 407-409.
23. DeLange, R.J., Hooper, J.A. and Smith, E.L. (1973). *J. Biol. Chem.* **248**, 3261-3274.