**Dictyostelium discoideum** contains a single-copy gene encoding a unique subtype of histone H1

( Genomic clone; Southern blot analysis; globular domain )

Loren J. Hauser, Madhu S. Dhar and Donald E. Olins

*University of Tennessee, Graduate School of Biomedical Sciences, Biology Division, Oak Ridge National Laboratory, Oak Ridge, TN, 37831-8077, USA*

Received by A.-M. Skalka: 21 June 1994; Revised/Accepted: 17 October/25 October 1994; Received at publishers: 8 December 1994

**SUMMARY**

A *Dictyostelium discoideum* genomic library was screened using a degenerate oligodeoxyribonucleotide derived from the peptide, GPKAPT, obtained from the N terminus of purified histone H1. Two identical H1 clones were isolated. Comparative sequence data reveal a typical H1 three-domain structure with considerable homology to the globular domain of higher eukaryotic H1 histones, especially to plant H1 histones. Southern blot analysis shows that this gene is probably a single-copy gene, and suggests that any other H1 gene(s), if present, must be very different in sequence. Amino acid (aa) sequence comparison of the globular core of *D. discoideum* H1 to the consensus globular core reveals the absence of a 6-aa motif, GXGXXG, from *D. discoideum*. This motif matches the consensus for a putative nucleotide-binding loop, which is also absent in plant H1 histones like *Arabidopsis thaliana*, pea and wheat.

**INTRODUCTION**

Histone H1 is believed to be the primary protein responsible for the formation and stabilization of the 30-nm chromatin fiber seen in eukaryotes (Van Holde, 1989). The H1 histones of multicellular eukaryotes contain three domains: a short unstructured N-terminal tail, a central globular core and a long basic unstructured C-terminal tail (Cole, 1987). Current structural models postulate that the globular domain binds to the nucleosome, while the two tails bind to the linker DNA between the nucleosomes (Van Holde, 1989).

Histone H1 does not appear to be present in *Saccharomyces cerevisiae* (Grunstein, 1990), limiting the use of yeast as a model system to study higher-order chromatin structure and H1 function. *D. discoideum* (the cellular slime mold), was shown to possess histone H1 (Parish and Schmidlin, 1985). This organism is haploid during vegetative growth, contains a small genome of $5 \times 10^7$ bp (Nellen et al., 1987), has clear mitotic chromosomes (Olins and Olins, 1993) and proceeds through a well-defined and synchronous multicellular developmental pathway (Nellen et al., 1987). It is readily transformed and has a high frequency of homologous recombination (Knecht and Kessen, 1990). Hence, these properties make *D. discoideum* a good model system to study eukaryotic higher-order chromatin structure and function. As a step towards this goal, we report the cloning and sequence analysis of the gene for histone H1 from *D. discoideum*. Our data show a single subtype of histone H1 present as a single copy gene, in sharp contrast to other eukaryotes.
where multiple subtypes and multiple copies of histone H1 are common (Cole, 1987).

EXPERIMENTAL AND DISCUSSION

(a) Isolation of histone H1 from *D. discoideum*

Nuclei were isolated (Charlesworth and Parish, 1975), extracted with 5% HClO₄ and H1 recovered by precipitation with acidified acetone (Parish and Schmidlin, 1985). Crude H1 was redissolved in 0.1% TFA and further purified by reverse-phase HPLC using a C₄ column and a water-acetonitrile gradient in 0.1% TFA. The H1 was identified as a single well-separated peak (data not shown). The purified histone H1 was subsequently chromatographed on a COO⁻ HPLC (MA7C, Bio-Rad) column with a guanidine-HCl gradient in 100 mM Na-phosphate pH 6.8. A single H1 peak was observed. The N terminus of the reverse-phase-HPLC-purified protein was sequenced three times by two different facilities and was determined to be GPKAPTTPTKKAAATK.

(b) Isolation of H1 genomic clones

A degenerate oligo, 5'-GGTCCAAAARGCDCCAAC (R = A or G; D = A, C or T), was designed based on the sequence GPKAPT, in conjunction with the codon preference of *D. discoideum* (Warrick and Spudich, 1988). A *D. discoideum* genomic library (gift of Dr. R. Firtel, UCSD) in a modified pAT153 (Dynes and Firtel, 1989) was screened according to Maniatis et al. (1982), using the oligo described above. Positive clones obtained after three rounds of screening were sequenced using the above mentioned oligo. Subsequently, they were mapped using a number of restriction enzymes like EcoRI, BamHI, etc., commonly found in the multiple cloning site and some enzymes like *RsaI*, which have a 4-bp recognition site. Two identical histone H1 genomic clones, pDH1-23 and pDH1-34, with an insert size of about 4.8 kb were obtained. The restriction map of pDH1-23 is shown in Fig. 1.

(c) Southern blot analysis

In order to determine whether there are multiple copies of the same or other related H1 subtypes, the insert in pDH1-23 was hybridized (Maniatis et al., 1982) to four genomic blots of *D. discoideum*. Subsequent to low-stringency hybridization (50°C, 6 × SSC), each blot was washed at a different stringency; 0.1 × SSC, 0.3 × SSC, 1 × SSC or 3 × SSC at 50°C (Fig. 2). The absence of any other hybridizing restriction fragments, other than those in the cloned gene imply the existence of only one single-copy H1 gene.

(d) Sequence analysis

Sequence analysis of the genomic clone was carried out using the GCG software package (Genetics Computer Group, 1991). Partial sequence analysis of pDH1-23, revealed an ORF coding for 180 aa and about 149 bp of the 5' and the 3' flanking regions (GenBank accession No. L33457). The coding region comprised the complete aa sequence of *Dictyostelium* H1. Codon frequency of the deduced sequence is in close agreement with published data (Warrick and Spudich, 1988).

The predicted aa sequence reveals a typical H1 three-domain structure (Cole, 1987), with considerable homology to the globular domain of higher eukaryotic H1 histones, especially to plant H1 histones (Fig. 3). The aa sequence comparison to the consensus globular core shows that the *Dictyostelium* H1 core is 5 aa shorter. A stretch of 6 aa, GTGASG (underlined in Fig. 3), is replaced by a single A. This shorter globular core is also present in plant H1 histones, like *Arabidopsis*, pea and wheat (Wells and McBride, 1989; Gantt and Lenvik, 1991; Wells and Brown, 1991). This motif matches the consensus (GXGXXG) for a putative nt binding loop (Nilsson et al., 1992), and has been shown to be in a loop

---

![Fig. 1. Restriction map of pDH1-23. The total length of the insert is about 4.8 kb and the coding region is 540 bp. Restriction enzymes that do not cut pDH1-23 are EcoRV, ClaI, HindIII, PstI, KpnI, Sall, SacI, HpaII and TaqI. The polylinker sites shown were inserted into pAT153 from pSP73 (Promega, Madison, WI, USA).](image-url)
null mutants of H1 using gene disruption, (ii) partial deletion mutants by removing all or a part of each of the three domains of H1, (iii) mutants with the addition of the nt binding loop and (iv) site-directed mutants of known phosphorylation sites.

ACKNOWLEDGEMENTS

We thank C. Smith (University of Tennessee Health Science Center) and Claude Stringer (Biology Division, National Institute of Health, Bethesda, MD) for help with the experiments and the manuscript. We thank C. Smith (University of Tennessee Health Science Center) and Claude Stringer (Biology Division, National Institute of Health, Bethesda, MD) for help with the experiments and the manuscript. We also thank David B. S. Robinson and Oleg Ptitsyn for help with the experiments and the manuscript. We acknowledge the use of the National Resource for Eukaryotic Genome Analysis, a National Institutes of Health Resource Center, operated by the University of Tennessee Health Science Center and funded by a contract from the National Institutes of Health. We also acknowledge the use of the National Resource for Eukaryotic Genome Analysis, a National Institutes of Health Resource Center, operated by the University of Tennessee Health Science Center and funded by a contract from the National Institutes of Health. We are grateful to Dr. C. Smith (University of Tennessee Health Science Center) and Claude Stringer (Biology Division, National Institute of Health, Bethesda, MD) for help with the experiments and the manuscript. We also acknowledge the use of the National Resource for Eukaryotic Genome Analysis, a National Institutes of Health Resource Center, operated by the University of Tennessee Health Science Center and funded by a contract from the National Institutes of Health.
Oak Ridge National Laboratory) for performing the protein sequencing; and Drs. Ada L. Olins and Richard Mural for critical reading of the manuscript. Research was sponsored by the US Department of Energy, under contract DE-AC05-84OR21400 with Martin Marietta Energy Systems, Inc. and NSF grant MCB9117123 to D.E.O.

REFERENCES


