Nucleotide sequence of two mouse histone H4 genes

Verena S.Meier, Ruth Böhni and Daniel Schümperli*

Institut für Molekularbiologie II der Universität Zürich, Hönggerberg, 8093 Zürich, Switzerland
Submitted December 13, 1988 Accession nos X13235, X13236

Two histone H4 genes (Fig. 1) were isolated from a BALB.A2G-Mx mouse genomic DNA library in lambda EMBL3 (Hug et al. (1988) Mol.Cell.Biol. **8**, 3065-3079). The gene on clone 12 has a 5' noncoding region of only 9 nucleotides. The stem of the 3'-terminal palindrome of the clone 53 H4 gene has the unusual sequence GGTCCT. In both genes (as well as in all other published vertebrate H4 genes), the promoter contains a recognition sequence for transcription factor SP1 (Dynan and Tjian (1983) Cell **35**, 79-87) immediately upstream from a very extensively conserved H4-specific box (Clerc et al. (1983) Nucl. Acids Res. **11**, 8641-8657). Clone 53, but not clone 12, additionally contains one copy each of H3 and H2A genes (H2B and H1 genes have not yet been screened for).

Clone 12 (acc.no. X13235)

```
        10         20         30         40         50         60         70         80         90        100
GGGAATGAAGTTTTTCTGCAGGATTTAAATGTGGTCTTTAAGAGACACCGCATGCAAAGAATAGCTGGGGCTTGCTAGCCAATGAAAACATTCAGATTCC
       110        120        130        140        150        160        170        180        190        200
AATGACGCATCCTTTTTTCTCCACCCCCTTCCAAGACCCGGATTCGGAAACCCCGCCTAACGCTCTAGTTTTCAACCAGGTCCGCAGAAGGCCTATTTAA
       210        220        230        240        250        260        270        280        290        300
AGGGACGATTGCTGTCTCCCTGCTGTCATAACCATGTCTGGACGTGGCAAGGGTGGTAAAGGCCTTGGGAAAGGCGGCGCTAAGCGCCACCGTAAGGTTC
       310        320        330        340        350        360        370        380        390        400
TCCGCGATAACATCCAGGGCATCACCAAGCCTGCCATCCGCCGCCTGGCCCGGCGCGGGGGAGTGAAGCGCATCTCCGGCCTCATCTACGAGGAGACCCG
       410        420        430        440        450        460        470        480        490        500
CGGTGTGCTGAAGGTGTTCCTGGAGAACGTGATCCGCGACGCCGTCACCTACACGGAGCACGCCAAGCGCAAGACCGTCACCGCCATGGACGTGGTCTAC
       510        520        530        540        550        560        570        580        590        600
GCGCTCAAGCGCCAGGGCCGCACTCTCTACGGATTCGGCGGTTAATCGACTAACAAACGATTTTCCACTGTCAACAAAAGGCCCTTTTCAGGGCCACCCA
       610        620        630
CAAATTCCTAGAAGGAGTTGTTCACTTACCGAAGCTT
```

Clone 53 (acc.no. X13236)

```
        10         20         30         40         50         60         70         80         90        100
CCTAAATACTTTACAGGGGTATTTCTTTTAATCCTCCAGGTACTGGTCCAATAAAAAATGTAACTTGCTAAAATCCTTATAGAATAATATAACATATAAT
       110        120        130        140        150        160        170        180        190        200
AAAAGCAGTATACTCAAATCTGGGGAAACCTGAGAGACTCCATGTCTGCATGTGCCCTATACAGTACTCCAGAGTCACCGCCCTGAAATGGGCTGAGCAA
       210        220        230        240        250        260        270        280        290        300
AAGAAGCCTTTGGTCCCCGCCCACTTGAAGTTCTCAACCAGGTCCGATTAGAGCCTATAATAACACAGCGAATTCTTAGTTCACCAGCTTCCTACTGAAA
       310        320        330        340        350        360        370        380        390        400
GCTTGTAAGAATTTAACTATGTCTGGTCGCGGCAAGGGAGGAAAAGGCCTGGGCAAAGGTGGCGCTAAGCGCCATCGCAAAGTGCTGCGCGACAACATCC
       410        420        430        440        450        460        470        480        490        500
AGGGCATCACGAAGCCCGCCATCCGCCGCCTGGCCCGGCGCGGAGGAGTGAAGCGCATCTCCGGCCTCATCTACGAGGAGACCCGCGGTGTGCTGAAGGT
       510        520        530        540        550        560        570        580        590        600
GTTCCTGGAGAACGTGATCCGCGACGCCGTCACCTACACGGAGCACGCCAAGCGCAAGACCGTCACCGCCATGGACGTGGTCTACCGGCTCAARCGCCAG
       610        620        630        640        650        660        670        680        690        700
GGCCGCACCCTCTACGGATTCGGGGGTTAAGGGACTAGAGCGATTTTCTTAACTGTTACAAAAGGTCCTTTTCAGGACCACTCACTGAATTCCTTAAAAA
       710        720        730        740        750        760        770        780        790        800
GCTGTGCATTTTGCTCCCTCACCCTTCCCCCACACTGGCAGCCGCCCTCACTAATAGCTTCAGCTTTAACTCAGGCTGTGGATAGTCAGCATAGTTGAGA
       810        820        830        840        850        860        870        880        890        900
TTCCCCACCCCCACCCCCACTCCTCCAGGCTGAATTTGYGCACAAAGTTAGTACAGCCTCACCTTGTACAAATTGTTAAAAGTAAGACACTTCAAGGCTC
```

Fig. 1. Sequence of two new mouse H4 genes. Important sequence elements (underlined) are: SP1 recognition sequence, H4-specific box, TATA-box and transcription start site (*), initiation codon, termination codon, hairpin element and spacer element of the 3' processing signal.

This work was supported by Swiss National Science Foundation grant No. 3.036.87 and by the State of Zürich.

*To whom correspondence should be addressed