cDNA sequences of chicken nucleolin/C23 and NO38/B23, two major nucleolar proteins

G.Maridor and E.A.Nigg
Swiss Institute for Experimental Cancer Research (ISREC), CH-1066 Epalinges, Switzerland

Submitted January 31, 1990 EMBL accession nos X17199, X17200

Nucleolin and NO38 are abundant nucleolar proteins implicated in the transcription and processing of ribosomal RNA, and in the packaging and transport of ribosomal proteins or preribosomal particles (for references see 1–5). cDNAs encoding the chicken homologs of these proteins were cloned (by immunoscreening; 6) from λgt11 cDNA expression libraries prepared from poly(A)$^+$ RNA of 10 day old chicken embryos (7). DNA from recombinant phages was prepared using lambdasorb (Promega), and preliminary sequence determination was carried out directly on λ phages, using Klenow polymerase and primers obtained from Biolabs. In the case of NO38, the complete coding information was contained within a single 1.1 kb *Eco*RI fragment, which was subcloned into the *Eco*RI site of the pGEM-3Zf(−) vector (Promega). The nucleolin insert contained four internal *Eco*RI sites, and, therefore, a 2.5 kb fragment was excised from the recombinant λgt11 phage using *Dra*II. The resulting fragment was filled up with Klenow polymerase and blunt-end ligated into the *Sma*I site of pGEM-3Zf (−). Nested deletion templates for sequence analysis were then generated by exonuclease III digestion (8), using *Xba*I and *Sph*I for plasmid linearization. Sequencing of double-stranded plasmid DNA (9) was carried out (in both orientations) by the dideoxy method (10) using T7 polymerase (Sequenase, United States Biochemicals).

ACKNOWLEDGEMENTS

We are grateful to Dr B.Vennström (Stockholm) for a kind gift of a chicken λgt11 library. This work was supported by the Swiss National Science Foundation (3.316 and 3.431) and the Swiss Cancer League (371.87.1).

REFERENCES

1. Jordan,G. (1987) *Nature* **329**, 489–490.
2. Lapeyre,B., Bourbon,H. and Amalric,F. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 1472–1476.
3. Schmidt-Zachmann,M.S. and Franke,W.W. (1988) *Chromosoma* **96**, 417–426.
4. Chang,J.-H. and Olson,M.O.J. (1989) *J. Biol. Chem.* **264**, 11732–11737.
5. Borer,R.A., Lehner,C.F., Eppenberger,H.M. and Nigg,E.A. (1989) *Cell* **56**, 379–390.
6. Young,R.A. and Davis,R.W. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1194–1198.
7. Sap,J., Munoz,A., Damm,K., Goldberg,Y., Ghysdael,J., Leutz,A., Beug,H. and Vennström,B. (1986) *Nature* **324**, 635–640.
8. Henikoff,S. (1984) *Gene* **28**, 351–359.
9. Chen,E.Y. and Seeburg,P.H. (1985) *DNA* **4**, 165–170.
10. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.

```
   1 GGCCCGCTCCGTTCCCCGCCGCCACCCCCCCAGGGACGCGCATTCCCACCCCCCGAGGCAGTTCGGCCGCCGCCCCAGCGCCGCCAGACACCGCCGCCATC
 101 ATGGTGAAGCTCGCCAAGACTCCCAAGAATCAAATGAAACAGAAAAAAATGGCGCCTCCCCCCAAAAAGGTCGAGGAAAGCGAAGAGGAAGAGTCCTCCG
 201 ACTTAGAGGAAAGCAGCGGGGAAGAGGTGATGGTCCCTCCAAAGAAACAACAAAAAGCAGCAGTTACACCAGCCAAGAAGGCTGCTACCCCTGCAAAGAA
 301 GGCTGCTACTCCTGCAAAAAAGGCAGTGACACCAGCCAAGAAGGCTGTGGCTACTCCAGCTAAAAAGGCTGTTGCTCCATCCCCCAAAAAGGCTGCTGTC
 401 GTAGGCAAAGGGGCAAAAAATGGCAAGAATGCCAAAAAGGAAGAGAGCGAGGAGGAAGATGAAGACGATGAAGATGATGAGGAAGATGAGGATGAAGAAG
 501 AGGAGTCTGATGAGGAAGAGGAACCAGCAGTGCCTGTGAAGCCTGCAGCCAAAAAGTCCGCAGCAGCTGTACCAGCCAAAAAGCCTGCAGTTGTGCCAGC
 601 AAAGCAAGAATCTGAGGAGGAGGAGGAGGAAGATGATGAAGAGGAGGACGAGGAAGATGACGAGTCTGAAGATGAGGCCATGGACACAACTCCTGCTCCT
 701 GTGAAGAAACCTACTCCAGCAAAGGCTACACCAGCTAAAGCCAAGGCGGAGTCTGAAGATGAGGAAGATGAGGAGGATGAGGATGAAGATGAGGAGGATG
 801 AAGATGATGAAGAGGAAGATGAGGAAGAAAGTGAGGATGAAAAACCTGTCAAGGAAGCACCTGGAAAAAGGAAAAAAGAAATGGCCAATAAGAGTGCACC
 901 AGAGGCCAAGAAAAAGAAAACAGAAACACCTGCTTCAGCTTTCTCACTCTTCGTGAAAAACTTGACCCCCACCAAGGACTATGAAGAACTGAGGACTGCC
1001 ATCAAAGAATTCTTTGGCAAGAAAAATCTCCAAGTCTCAGAAGTCAGAATCGGTTCTTCCAAGCGGTTTGGCTATGTGGACTTCTTATCTGCTGAAGATA
1101 TGGATAAAGCTCTTCAACTGAATGGAAAGAAGCTGATGGGTTTGGAAATCAAACTGGAAAAAGCAAAGAGCAAAGAAAGTCTTAAAGAAAATAAGAAAGA
1201 GAGAGATGCCAGAACACTGTTTGTGAAGAATCTGCCCTACCGTGTAACTGAAGATGAAATGAAAAAATGTATTTGAAAACGCTTTAGAAGTCCGACTAGTA
1301 CTTAACAAGGAAGGGAGCAGCAAAGGGATGGCCTACATTGAATTCAAAACAGAAGCTGAGGCAGAAAAAGCACTGGAGGAGAAGCAAGGCACAGAGGTTG
1401 ACGGTCGTGCCATGGTCATTGACTACACCGGTGAGAAGAGCCAACAAGAAAGTCAGAAAGGAGGTGGAGAGAGGGAGTCAAAGACTCTGATTGTGAACAA
1501 CCTGTCATATGCTGCCTCAGAAGAAACTCTCCAGGAACTGTTTAAAAAAGCGACTTCCATCAAGATGCCACAGAACAACCAGGGCAGGCCTAAAGGGTAT
1601 GCATTTGTAGAATTTCCCACAGCCGAGGATGCCAAAGAGGCATTGAATTCCTGTAACAACACAGAAATTGAAGGCAGAGCAATCAGGCTGGAATTCAGTT
1701 CACCATCGTGGCAGAAAGGGAACATGAATGCAAGAGGAGGATTTAACCAACAAAGCAAAACATTGTTTGTCAGAGGCCTTTCTGAGGACACAACGGAGGA
1801 GACGCTAAGAGAATCATTTGAAGGCTCTATAAGTGCTAGAATAGTCACAGATAGAGACACTGGATCTTCTAAAGGGTTTGGTTTTGTGGACTTCAGCTCC
1901 CCAGAAGATGCCAAAGCAGCTAAAGAAGCTATGGAGGATGGAGAGATAGACGGAAACAAAGTGACCCTTGATTTTGCCAAACCAAAGGGTGAATTTCAGC
2001 GTGGCGGCGGATTTGGTGGTGGATTTGGTGGTCGTGGTGGCAGAGGAGGCAGAGGAGGAGGAAGAGGTGGATTTGGTGGCAGAGGTGGTGGCAGAGGGTT
2101 TGGAGGTAGAGGAGGTGGCTTCCGAGGAGGCAGAGGAGGAGGTGGAGATCACAAGCCACAAGGGAAGAAGATCAAGTTTGAATAAACTTCCCTTTTCCTT
2201 TCCCTTCTCTTGATCTCTGAGACTATCTGAAAGGACTCCAGGGGTTTTTATTCTCCTTTTATCTTGGTGGAGCCTTCGGAGGACATTCCAGGGTGCAAAG
2301 GAGTACTGTGAGCCACTTGGAAGAAAGAATTTTCATTTCAAGGAAGAAAAGCAAGCTATTTTGTCTGCTTGCTTATTCAAGTGAACTTTTTAACAGAAAT
2401 GAGCGATAGCTGTGGGAATTTACCCTTTGTCTGTGAGTTGTCTTGAATTTATAATGTTTTACCCATGTACAAAAATATTTTTTCCTATAAACTTCTGTA
2501 GCATTTTTTTGCTGTAAAAATGCAGAATGTTTTATCATTTGTGCTTCAGCACCCTGTCTTGGACAGATTAAAGGAC  2576
```

Figure 1. cDNA sequence of chicken nucleolin. The cDNA sequence shown contains 2576 nucleotides; the translation start (ATG) and stop (TAA) codons are underlined. These results complete a preliminary report of a partial (C-terminal) sequence for chicken nucleolin (5), and they correct two mistakes in that sequence, namely a GC for CG substitution at positions 1355/56 (resulting in codon 419 coding for Ala instead of Arg), and an A for C substitution at position 1659 (resulting in codon 520 coding for Asn instead of Thr).

```
   1 GAATTCCGGTCCGCAGCGCCGCATCCCGTACAGCTCCTCCCGCGCGAGCAGGCCGAGAAGATGGAGGACAGCGCCATGGACATGGAGAGCATGGGCCCGC
 101 TGCGCCCGCAGACCTTCCTCTTCGGCTGCGAGCTTAAAGCAGAGAAGGAATATCAGTTCAAAGTAGATGATGAGGAAAACGAACATCAGCTGTCCTTGAG
 201 AACGGTTACATTAGGGGCTGGAGCCAAAGACGAATTACATGTTGTAGAAGCAGAAGCACTGGACTACGAAGGCAACCCAACTAAAGTAGTACTGGCATCT
 301 CTGAAAATGTCTGTGCAGCCTACAGTTTCACTAGGTGGATTTGAGATCACACCACCATTTGTCTTGAGGTTAAAATGTGGTTCGGGGCCTGTTTATGTCA
 401 GTGGTCAGCATCTTGTAGCATTAGAGGAAGAGCCAGAATCAGAGGATGAGGAGGAGGATACAAAAATAGGGAATGCTTCAACAAAGAGACCAGCAAGTGG
 501 AGGAGGAGCTAAAACACCACAGAAAAAACCAAAATTATCAGAAGATGATGAGGACGATGATGAAGATGAGGATGATGAGGATGATGAAGACGACTTG
 601 GATGATGATGAGGAGGAGATTAAAACACCAATGAAGAAACCTGCCCGCGAGCCTGCAGGAAAAAATATGCAGAAAGCAAAGCAAAATGGAAAAGACTCAA
 701 AGCCGTCCACACCAGCATCTAAAACAAAAACTCCAGATTCCAAGAAGGACAAATCTCTAACTCCAAAAACACCGAAAGTTCCTCTGTCGTTAGAGGAGAT
 801 CAAAGCAAAAATGCAGGCCTCCGTAGACAAGGGTTGTTCCCTTCCTAAGCTGGAGCCCAAATTTGCCAACTATGTTAAGAATTGCTTCAGGACGGAGGAC
 901 CAAAAGGTCATTCAAGCTCTCTGGCAGTGGAGACAGACTCTGTAAGAGAACAATTTAAACAGTTTGTTAAAGTCTGCAGTCTTACTTCTGTAACCATTAT
1001 TTGGCTGTTCTTTTTACAAATGCTGAAAGAGCTTTCCCTACAGTGTCTGATAAATGTCATCCAGATTACCTTGCCAAGAATGTGTTGTCCAAAATGCCTG
1101 TTTAGTTTTTAAAGATGGGACTCCACCCCGAGAATTC  1137
```

Figure 2. cDNA sequence of chicken NO38. The cDNA sequence shown contains 1137 nucleotides; the translation start (ATG) and stop (TAA) codons are underlined. These results complete a preliminary report of short partial (N- and C-terminal) sequences for chicken NO38 (5).