

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 lamdasorb (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 EcoRI fragment, which was subcloned into the EcoRI site of the pGEM-3Zf(-) vector (Promega). The nucleolin insert contained four internal EcoRI sites, and, therefore, a 2.5 kb fragment was excised from the recombinant λgt11 phage using DraII. The resulting fragment was filled up with Klenow polymerase and blunt-end ligated into the SmaI site of pGEM-3Zf(-). Nested deletion templates for sequence analysis were then generated by exonuclease III digestion (8), using XbaI and SphI 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

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

```

1  GGCCCGCTCCGTTCCCGCCGGCCACCCCCAGGGAGCGCATTCCACCCCGCGAGGCGTTCGGCCGGCCCGCCAGACACCCCGCCATC
101  ATGTCGAAGCTCCGCAAGACTCCDAAAGATCAAAATGAACAGAAAAAATGGCCGCTCCGCCAAAAGGTCGAGGAAAGCGAAGAGAGACTTCCTCCG
201  ACTTAGCAGAAAGCAGCGGGGAGAGAGGTTGATGTCCTCCCAAGAAACAACAATAAGCCAGCTTACCCAGCCAGAGAGGCTCCAGCCCTCGCAAGAA
301  GCGTCTACTCTCCGAAAAAGCCAGTCAACAGCCAGCAAGAGGCTTGGCTACTCCAGCTAAAGAGCCCTTTCCCAATCCCGCAAAAGGCTGGCTTC
401  GTTAGCCAAAGGGCCAAAAATGGCAAGATGCCAAAAGGAGAGAGCCAGAGAGAGATGAAGCACTGAAGATCAATGAGGAAATGAGGATGAGAGAG
501  AGAGTCTGATGAGCAGAGGAAACAGCAGGATGCTCTGCAAGCCCTCGAGCCAAAAGTCCGAGCAGCTGTACAGCCAAAAGGCTCCAGTCTGTCAGC
601  AAAGCAGAAATCTAGGGAGAGGAGGAGCAAGATGATGAAGAGGAGGAGGAGGAGATGAAGGATCGAAGATGAGGCCATGGACACAACTCTGCTCTCT
701  GTGAGAAACCTACTCTCCGCAAGGCTTCAACAGCTAAAGCCAGGCGAGCTCTGAGATGAGGAGATGAGGAGGATGAGGATGAGGATGAGGAGGATG
801  AAGTCTGATGAGGAGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
901  AGAGCCAGAAAGAGAAACAGAAACAGCTCTTTCAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
1001  ATCAAGAAATCTCTTGGCAAGAAAAATCTCCAAAGCTCAGAAAGTCAAGATGGGTTCTTCCCAAGCGGTTGGCTATCTGGACTTCTATCTGCTGAAGATA
1101  TGGATAAGCTCTCACTGAAATGGAAGAAAGCTGATGGGTTTGGAAATCAAACTGGAAAAAGCAAGAGCAAGAAAGCTTTTAAAGAAAATTAAGAAAGA
1201  GGAGATGCCAGAACTCTTGTGAGAAATCTGCGCTACCGTCTAACTGAAATGAACTGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT
1301  CTTAAGCAGAAAGGAGCAGAGCAGAGGATGGCTGATGCAATTTGAAACAGAAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
1401  ACCGCTCTGCACTGCTTACTACAGCTTGAAGAGGCAACAGAAAGCTGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
1501  CCTGCTCATATCTGCTCCAGAGAAATCTCCAGAACTGTTTAAAAAGGAGGACTTCCATCAGATGCCACAGAAACACCAGGCGAGGCTCAAAAGGATAT
1601  GCATTTGTAGAAATTCGCAAGCGGAGGATGCCAAGAGGCAATGAAATCTCTTAAACACAGAAATTAAGGCGAGGAGCAATCAGGCTGGATTTACCT
1701  CAGCATCTGCGCAGAAAGGCAAGCAATGAAATGCAAGAGGAGGATTTAACCAACAAGCAAAACATTTTCTCAGAGGCTTCTGAGGACACAAAGCAGGA
1801  GAGCTTAGAGAAATCAATTTGAGGCTGATAGGCTGATAGGCTAGAAATGTCACAGATGAGAGAGCTGGAATTTCTTAAAGGCTTTGGTTTCTGACTCAGCTCC
1901  CCAGAAAGCTCCCAAGAGCTTAAAGAACTTGCAGGATGGAGGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
2001  GTGCGCGGCAATTTGGCTGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
2101  TCGAGGATAGAGGAGGCTGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
2201  TCCCTTCTCTGATCTCTGACATCTGACAGATCTGAAAGAGACTCCAGGCTTTTATCTCTCTTTATCTCTCTGAGGCTTCGGAGGACTTCCAGGCTTC
2301  GAGTCTGTAGAGCACTGGAGAAAGAAATTTTCAATTCAGAGGAGAAAGCAAGCTATTTTCTCTGCTGCTTATTCAGAGTCAATTTTAAACAGAAAT
2401  GAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
2501  GCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT

```

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 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  GAATTCGGCTCCGAGGCGCCGATCCCGTACAGCTCTCCGCGCCAGGCGGAGAGATGAGGACAGCGCCATGGAATGAGGAGCAATGGCCCCC
101  TGCCGCCAGGACTTCTCTCTGCGCTGCGAGCTTAAGCAGAGGAGAAATACAGTCTCAAGTATGATGATGAGGAAAAAGCAACATAAGCTCTCTCTGAG
201  AAGCTTTACAATTAGGGCTTGGAGCAAGAGCAATTAACAATGTTAGAACAGAGCACTGGACTCAGAGGCAACCCAACTAAAGTGTACTGCTGCTCTCT
301  CTGAATAATCTCTCTGAGCCTTACGTTTCACTAGGTTGATTTGAGATGACACCAACCTATTTCTTAAAGGCTTTGGTTTCTGACTCAGCTCC
401  GTGCTCAGCTCTCTGATCAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
501  AGGAGGACTTAAACACACAGAAAAAATAAATTCAGAAATGATGAGGAGGATGATGAGGATGAGGATGATGAGGATGATGAGGATGATGAGGATGATGAGG
601  GATGATGATGAGGAGGATGATGAGGATGATGAGGATGATGAGGATGATGAGGATGATGAGGATGATGAGGATGATGAGGATGATGAGGATGATGAGGAT
701  AGCCCTCCACACAGCACTTAAACAAAACTCCAGATTTCCAGAAAGGACAAAACTCTAACTCCAAAACACCGAAAGTTCTCTCTCTCTCTCTCTCTCT
801  CAAAGCAAAAATGCAAGGCTCTCTTGAACAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
901  CAAGAAGTCTTCAAGCTCTCTGAGCTGAGCAGACTCTCTGAGGAGCAAAATTAACAGTTTGTAAAAGTCTGAGCTTCACTCTCTGATACCTTAT
1001  TTGCTCTCTCTTTTAAATAAGTCAAGAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
1101  TTAGTTTTTAAAGATGGCACTCCAGGAGGATTC

```

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).