

Molecular cloning and expression of the functional gene encoding the M2 subunit of mouse ribonucleotide reductase: a new dominant marker gene

M.Thelander and L.Thelander

Department of Medical Biochemistry and Biophysics, University of Umeå, S-901 87 Umeå, Sweden

Communicated by P.Reichard

Mammalian ribonucleotide reductase consists of two non-identical subunits, proteins M1 and M2. M2-related DNA sequences are present on mouse chromosomes 4, 7, 12 and 13. However, M2-overproducing mouse cells show amplification of a chromosome 12-specific, single 13 kb *Hind*III fragment, which probably represents the active gene. We have isolated this fragment from parental mouse cell DNA and used it to clone and characterize the functional M2 gene. The 5770 bp transcribed M2 sequence contains ten exons separated by nine 95–917 bp introns. The 501 bp of 5' flanking DNA is G+C rich and contains TTTAAA and CCAAT sequences as well as potential Sp1 binding sites. The M2-related sequence on chromosome 13, which contains only the last six exons and several internal rearrangements, is a pseudogene. Transfection of BALB/3T3 cells with the M2 gene resulted in stable transformants with a 10-fold reduction in sensitivity to hydroxyurea, compared to control cells. This confirmed that the cloned M2 genomic DNA represents the functional gene and conclusively establishes the link between hydroxyurea resistance and M2 expression in mammalian cells. M2 genomic DNA should be a valuable dominant, selectable marker for identifying and isolating stable co-transformants.

Key words: hydroxyurea resistance/M2-gene/ribonucleotide reductase

Introduction

Ribonucleotide reductase is a highly regulated enzyme which catalyzes the first unique step in the reactions leading to DNA synthesis (Thelander and Reichard, 1979; Reichard, 1988). Enzyme activity is strongly correlated to the rate of DNA synthesis showing maximal activity in S-phase cells. In this respect it resembles other enzymes exclusively involved in DNA synthesis such as thymidine kinase (Littlefield, 1966), thymidylate synthetase (Conrad, 1971) and DNA polymerases (Fry and Loeb, 1986).

Mammalian ribonucleotide reductase is a heterodimer consisting of two non-identical subunits, proteins M1 and M2. Both subunits have been purified to homogeneity and the corresponding cDNAs have been cloned from mouse cells (Caras *et al.*, 1985; Thelander and Berg, 1986). The genes encoding proteins M1 and M2 are located on different chromosomes in both human and mouse cells and, in addition to the functional genes, both the human and mouse genomes contain three M2 pseudogenes (Yang-Feng *et al.*, 1987; Brissenden *et al.*, 1988).

Protein M1 has a mol. wt of $2 \times 84\,000$ and contains binding sites for nucleoside triphosphate allosteric effectors and ribonucleoside diphosphate substrates (Thelander *et al.*, 1980). Protein M2, with a mol. wt of $2 \times 45\,000$, contains a non-heme iron center that generates and stabilizes a tyrosyl free radical essential for activity (Thelander *et al.*, 1985). The radical is the target for a number of DNA synthesis inhibitors, e.g. hydroxyurea (Åkerblom *et al.*, 1981). Both human and rodent cells have been selected for hydroxyurea resistance and shown to overproduce protein M2 as a result of gene amplification (Thelander and Berg, 1986; Wright *et al.*, 1987).

Protein M1 has a half-life of ~20 h in proliferating cells and measurements in both cell extracts and whole cells have shown that the levels of M1 protein are almost constant throughout the cell cycle (Engström *et al.*, 1985; Mann *et al.*, 1988). In contrast, protein M2 has a half-life of only 3 h under the same conditions, and electron paramagnetic resonance measurements of the levels of active M2 protein during the cell cycle have shown strictly S-phase-correlated expression (Eriksson *et al.*, 1984; Engström *et al.*, 1985).

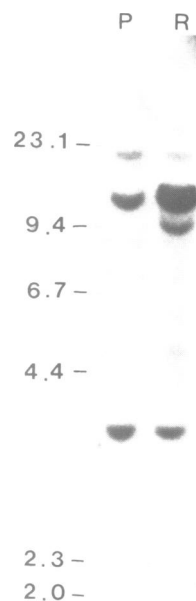


Fig. 1. Southern blot analysis of *Hind*III-digested chromosomal DNA from parental (P) and hydroxyurea-resistant mouse TA 3 cells (R) using a ^{32}P -labeled *Xho*I-generated fragment containing full-length mouse M2 cDNA as a probe (Thelander and Berg, 1986). High mol. wt DNA (~10 μg) was digested to completion with *Hind*III endonuclease and electrophoresed on a 0.8% agarose gel, transferred to nitrocellulose and hybridized as described. The DNA size markers were from *Hind*III-digested λ DNA.

Therefore, holoenzyme activity is controlled during the cell cycle by the levels of M2 protein which in turn is regulated by *de novo* synthesis and breakdown (Eriksson *et al.*, 1984). Both subunits of mammalian ribonucleotide reductase are exclusively localized to the cell cytoplasm (Engström *et al.*, 1984; Engström and Rozell, 1988).

We have cloned and sequenced the functional gene encoding mouse protein M2. This genomic clone will be useful for studying the regulation of ribonucleotide reductase during the cell cycle as well as the correlation between ribonucleotide reduction and DNA synthesis. Mouse cells transfected with the cloned M2 genomic DNA gave rise to stable transformants showing 10-fold resistance to hydroxyurea. Therefore, co-transfection with this DNA allows hydroxyurea resistance to be used as a new selectable marker.

Results

Cloning the functional gene encoding the M2 subunit of mouse ribonucleotide reductase

Somatic cell hybrid studies showed that there are four different sites for protein M2-related sequences on mouse chromosomes, located on chromosomes 4, 7, 12 and 13 (Yang-Feng *et al.*, 1987). These give rise to *HindIII* fragments of 21, 13 and 4.2 kb respectively (Figure 1). The functional M2 locus was identified by Southern blot analysis of DNA from hydroxyurea-resistant mouse TA 3 cells that overproduce protein M2 because of gene amplification. In such DNA, a single 13 kb *HindIII* fragment, showing at least a 5-fold amplification, was detected using a full-length M2 cDNA probe. Therefore this fragment most probably contains the entire M2 gene (Figure 1). The same 13 kb fragment was seen in DNA from parent type TA 3 cells and DNA isolated from mouse liver (data not shown).

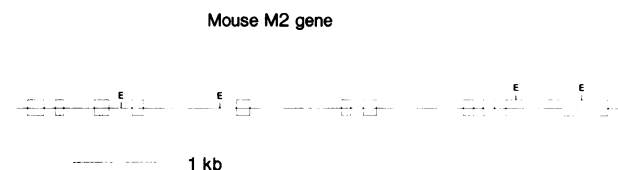


Fig. 2. Structure of the mouse M2 gene. The exons, represented by boxes, and introns are drawn to scale. *EcoRI* restriction endonuclease sites are indicated as E.

Initial attempts to isolate the active M2 gene from a Charon 4A mouse sperm genomic library were unsuccessful because fragments from the functional gene could not be distinguished from fragments originating from the various pseudogenes. Instead mouse genomic DNA was digested with *HindIII*, the fragments separated by gel electrophoresis and DNA fragments with a size of ~13 kb were recovered from the gel (Figure 1). These fragments were then cloned into the lambda vector L47, which allows direct cloning of *HindIII* fragments (Loenen and Brammar, 1980). M2 positive clones were identified by screening with an M2 cDNA probe. The 13 kb mouse DNA insert encoding protein M2 was subcloned into the plasmid vector pUC18 to give the clone pM2Hind13. Finally, this clone was analyzed by restriction endonuclease mapping and hybridization to cDNA fragments. The nucleotide sequence of the M2 gene was determined after further subcloning of the various fragments into M13 vectors.

Characterization of the mouse M2 gene

A schematic structure of the M2 gene is shown in Figure 2. There are 5770 bp from the start of the first exon (see below) to the end of the last exon. The full nucleotide sequence, which also covers 501 bp of 5' flanking DNA sequences and 369 bp of 3' flanking sequences, will be available through the EMBL Data Library (accession number X15666 M2 subunit).

The transcribed sequence is divided into ten exons separated by nine introns ranging in size from 95 to 917 bp (Table I). All introns follow the general rules and start with the sequence GT and end with AG (Mount, 1982). The active site tyrosyl free radical is encoded within exon 5.

In Figure 3 the nucleotide sequence of the 5' flanking DNA is shown. In the promoter region a TATA homology can be found at position -29 to -24 and a CCAAT sequence at position -79 to -75. In addition, two potential Sp1 binding sites in different orientations are located further upstream from the transcription start site (see below). The promoter region is highly G+C rich, containing 65% G+C in the 501 bp of 5' flanking DNA.

Transcription start site

A nearly full-length M2 cDNA was isolated from an Okayama-Berg library (Thelander and Berg, 1986). From the length of this cDNA, the M2 transcription start site was

Table I. Exon-intron splice-junction sequences in the mouse M2 gene

Exon	Position in cDNA ^a	Size of exon (bp)	Sequence of exon-intron junctions				
			5'-splice donor	intron (size)		3'-splice acceptor	
1	1-163	163	A A C A C G	g t g a g g g c g g	(114)	c c c c c t a c a g	C C C C C G
2	164-238	75	G A G C T G	g t g a g t g g c c	(304)	c a t c c a a c a g	G A A A G T
3	239-385	147	G A G G A G	g t a a t c g t t c	(229)	t a a a c t t c a g	G T G G A C
4	386-502	117	A A C T T G	g t g a g t t g g c	(917)	c t a c c a c t a g	G T G G A G
5	503-636	134	G G A A A G	g t g a g t a t c c	(904)	g g c t c t g c a g	A G A A T A
6	637-731	95	C G T A T G	g t g a g g a c c c	(127)	c t c t t g c t a g	G A G A A C
7	732-865	134	G A C G A G	g t g a g t c t g a	(858)	t c t c t t c t a g	G G T T T A
8	866-970	105	G A G C A G	g t g a g t g a c t	(95)	t g c c c t c t a g	G A G T T C
9	971-1084	114	A A C A A G	g t a a a g t g t t	(109)	a c t t t t t c a g	A T T T T C
10	1085-2113	1029	C A T G G C				

Upper-case letters signify exon sequences, and lower-case letters signify intron sequences.

^aNote that this numbering differs from the published cDNA sequence by the addition of two nucleotides to the 5' end, i.e. 1 refers to the first G in the major M2 transcript (cf. Figure 3).

out with M2 insert, excised from pM2Hind13 with *HindIII*, rather than supercoiled plasmid DNA.

From the resistant colonies, five independent clones were isolated, expanded and tested for levels of hydroxyurea resistance as compared to parent 3T3 cells. The concentrations of drug that reduced 72 h growth to 50% were 0.38, 0.42, 0.43, 0.46 and 0.58 mM respectively for the five different resistant clones. The corresponding value for the parent 3T3 cells was 0.04 mM.

Discussion

The M2-related sequence located on mouse chromosome 12 indeed represents the functional M2 gene because transfection of cells with this piece of DNA resulted in a hydroxyurea-resistant phenotype. This result conclusively establishes the link between hydroxyurea resistance and M2 expression in mammalian cells. Nothing can be said at present about the M2-related sequences present on chromosomes 4 and 7 except that they are not amplified in the hydroxyurea-resistant, M2-overproducing mouse TA 3 cells. The M2 sequence on chromosome 13 clearly must represent an inactive pseudogene since it is both truncated and differs in sequence from the active gene. In these respects it also differs from a typically processed pseudogene (Wagner, 1986).

Comparison between the 5' upstream sequence of the M2 gene with the corresponding sequences from other cell-cycle-regulated genes such as the human and mouse thymidine kinase genes (Flemington *et al.*, 1987; Seiser *et al.*, 1989) and the hamster histone H-3.2 gene (Artishevsky *et al.*, 1987) show no obvious similarities. Except for the presence of TATAAA and CCAAT elements as well as potential Sp1 binding sites, the only constant feature seems to be that the promoter region in all these genes is very G+C rich. Interestingly, the rare TATAAA variant 'TTTAAA' found in the M2 gene is identical to that found in the human thymidine kinase gene.

Further experiments are clearly needed to identify *cis*-acting regulatory DNA elements. Here the demonstrated hydroxyurea-resistant phenotype of cells stably transformed by the functional M2 genomic DNA should be useful. All five of the characterized stable mouse cell transformants showed the same 10-fold increase in hydroxyurea resistance when compared with parental cells. Therefore, the resistant phenotype in all these cases is most likely the result of the uptake of one additional functional M2 gene copy.

It is interesting that previous transfection experiments of various cell types using full-length M2 cDNA under the control of the SV40 early promoter repeatedly failed to give stable hydroxyurea-resistant transformants, although transient expression of the M2 protein in COS-7 cells was easily demonstrated (Thelander and Berg, 1986). This suggests that the genomic M2 DNA might contain sequences essential for stable M2 expression. Hydroxyurea is an inexpensive drug and can be used in low concentrations. Therefore, co-transfection with M2 genomic DNA followed by hydroxyurea selection should be a valuable method to select for stable transformants.

The primer extension experiments demonstrated that our previously isolated 'nearly full-length' M2 cDNA was in fact full-length. However, using the Okayama-Berg cloning

method, there is an inherent risk of underestimating the number of 5' terminal G residues because they cannot be distinguished from the G-tail of the linker used in the construction of the library (Okayama and Berg, 1983). This is the reason why our published M2 cDNA sequence lacked two Gs at the extreme 5' end.

It was recently demonstrated that the gene encoding the small subunit of the bacteriophage T4 ribonucleotide reductase (*nrdB*) contains an intron (Gott *et al.*, 1986; Sjöberg *et al.*, 1986). There is a high degree of similarity between this intron and eukaryotic introns, suggesting a common ancestry. Furthermore, in the intron region, the T4 protein shows a high amino acid sequence similarity to the equivalent proteins of *Escherichia coli* and mouse (Shub *et al.*, 1988). However, the T4 *nrdB* splice junction is located in the middle of exon 7 in the mouse M2 gene and therefore has not been conserved in this species.

Materials and methods

Cells

Hydroxyurea-resistant, M2-overproducing mouse mammary tumor TA 3 cells, the corresponding parental cells and BALB/3T3 cells (ATCC CCL 163) were grown as monolayer cultures in Dulbecco's modified Eagle's medium plus 10% heat-inactivated horse serum as described earlier (Thelander and Berg, 1986).

Bacterial strains, vectors and media

Mouse DNA fragments were subcloned into the plasmid vector pUC18 and transfection of *E. coli* JM109 cells (Yanisch-Perron *et al.*, 1985) was made as described by Hanahan (1983). Plasmid DNA was prepared from overnight cultures of infected bacteria grown in LB medium at 37°C in the presence of 50 µg/ml carbacilline (Astra, Sweden). The cells were gently lysed by treatment with lysozyme followed by Triton X-100 and DNA was purified by two consecutive CsCl gradient ultracentrifugations.

Genomic cloning

A mouse sperm DNA library established in the lambda Charon 4A phage vector was obtained from Dr Ken-ichi Arai, DNAX, CA (Miyatake *et al.*, 1985). For the lambda L47 phage sublibrary, DNA was purified from parent mouse TA 3 cells, digested with *HindIII* and the fragments separated by electrophoresis in a 0.5% agarose gel. DNA fragments of ~13 kb were electrophoresed into NA 45 paper (Schleicher and Schull) and eluted. L47 DNA (Amersham) was digested with *HindIII* and *SalI*, the lambda arms purified by electrophoresis in a 0.5% low melting agarose gel, eluted, ligated to the 13 kb mouse DNA *HindIII* fragments and packaged using an *in vitro* packaging kit (Amersham). The libraries were screened by plaque hybridization using an *XhoI* fragment containing the full-length mouse M2 cDNA as a probe (Thelander and Berg, 1986).

DNA sequencing

The 13 kb DNA fragment containing the M2 gene was mapped using restriction endonucleases and hybridization to DNA fragments from the various parts of the M2 cDNA. Smaller fragments were subcloned in M13 phage vectors and the DNA sequence determined by the dideoxy terminator method using Sequenase (USB) and dGTP and dTTP labeling mixes in parallel (Tabor and Richardson, 1987). Computer-assisted sequence collection and analyses were accomplished with the programs of DNASTAR, Inc.

Primer extension

A synthetic oligonucleotide 5' AGAGCATGGCGAACGAGGGC 3' (Symbicom, Umeå) corresponding to nucleotides 52–71 in the M2 cDNA (see footnote to Table I) was labeled using [γ -³²P]ATP (sp. act. 5000 C/mmol, Amersham) and T4 polynucleotide kinase. Primer extension was made as described (Calzone *et al.*, 1987) after annealing 15 fmol of oligonucleotide to poly(A) RNA from parental and hydroxyurea-resistant TA 3 cells at 50°C for 6 h. Total cellular RNA was extracted from logarithmically growing cells by the guanidium thiocyanate method (Chirgwin *et al.*, 1979) and poly(A) RNA was recovered after one cycle of adsorption and elution from an oligodeoxythymidylate [oligo(dT)]-cellulose column (Aviv and Leder, 1972).

Transfection experiments

DNA (20 µg) was used to transfect 1 × 10⁶ BALB/3T3 cells on 10 cm dishes using the calcium phosphate co-precipitation method (Graham and van der Eb, 1973) and a 20% glycerol shock (Frost and Williams, 1978) 4 h after the addition of DNA. Medium was changed following the shock and the cells were allowed to grow for two more days in non-selective medium. Cells were then passaged into a selective medium containing 0.2 mM hydroxyurea which was changed every 2–3 days. Clones became visible within 12 days and were counted after 2 weeks of selection by staining with methylene blue. Independent clones were isolated from parallel unstained dishes, expanded in the selective medium and finally tested for level of hydroxyurea resistance by measuring cell growth (Suttle and Stark, 1979) during 72 h at different concentrations of hydroxyurea.

Acknowledgements

This work was supported by grants from the Swedish Natural Science Research Council and Magn. Bergvalls Stiftelse.

References

Åkerblom, L., Ehrenberg, A., Gräslund, A., Lankinen, H., Reichard, P. and Thelander, L. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 2159–2163.
 Artishevsky, A., Wooden, S., Sharma, A., Resendez, E., Jr and Lee, A.S. (1987) *Nature*, **328**, 823–827.
 Aviv, H. and Leder, P. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1408–1412.
 Brissenden, J.E., Caras, I., Thelander, L. and Francke, U. (1988) *Exp. Cell Res.*, **174**, 302–308.
 Calzone, F.J., Britten, R.J. and Davidson, E.H. (1987) *Methods Enzymol.*, **152**, 611–632.
 Caras, I.W., Levinson, B.B., Fabry, M., Williams, S.R. and Martin, D.W., Jr (1985) *J. Biol. Chem.*, **260**, 7015–7022.
 Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) *Biochemistry*, **18**, 5294–5299.
 Conrad, A.H. (1971) *J. Biol. Chem.*, **246**, 1318–1323.
 Engström, Y. and Rozell, B. (1988) *EMBO J.*, **7**, 1615–1620.
 Engström, Y., Rozell, B., Hansson, H.-A., Stemme, S. and Thelander, L. (1984) *EMBO J.*, **3**, 863–867.
 Engström, Y., Eriksson, S., Jildevik, I., Skog, S., Thelander, L. and Tribukait, B. (1985) *J. Biol. Chem.*, **260**, 9114–9116.
 Eriksson, S., Gräslund, A., Skog, S., Thelander, L. and Tribukait, B. (1984) *J. Biol. Chem.*, **259**, 11695–11700.
 Flemington, E., Bradshaw, H.D., Jr, Traina-Dorge, V., Slagel, V. and Deininger, P.L. (1987) *Gene*, **52**, 267–277.
 Frost, E. and Williams, J. (1978) *Virology*, **91**, 39–50.
 Fry, M. and Loeb, L.A. (1986) In *Animal Cell DNA Polymerases*. CRC Press, Boca Raton, FL, pp. 13–60.
 Gott, J., Shub, D.A. and Belfort, M. (1986) *Cell*, **47**, 81–87.
 Graham, F.L. and van der Eb, A.J. (1973) *Virology*, **52**, 456–467.
 Hanahan, D. (1983) *J. Mol. Biol.*, **166**, 557–580.
 Littlefield, J.W. (1966) *Biochim. Biophys. Acta*, **114**, 398–403.
 Loenen, W.A.M. and Brammar, W.J. (1980) *Gene*, **20**, 249–259.
 Mann, G.J., Musgrove, E.A., Fox, R.M. and Thelander, L. (1988) *Cancer Res.*, **48**, 5151–5156.
 Miyatake, S., Yokota, T., Lee, F. and Arai, K. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 316–320.
 Mount, S.M. (1982) *Nucleic Acids Res.*, **10**, 459–472.
 Okayama, H. and Berg, P. (1983) *Mol. Cell. Biol.*, **3**, 280–289.
 Reichard, P. (1988) *Annu. Rev. Biochem.*, **57**, 349–374.
 Seiser, C., Knötler, M., Rudelstörfer, I., Haas, R. and Wintersberger, E. (1989) *Nucleic Acids Res.*, **17**, 185–196.
 Shub, D.A., Gott, J.M., Xu, M.-Q., Lang, B.F., Michel, F., Tomaszewski, J., Pedersen-Lane, J. and Belfort, M. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 1151–1155.
 Sjöberg, B.-M., Hahne, S., Mathews, C.Z., Mathews, C.K., Rand, K.N. and Gait, M.J. (1986) *EMBO J.*, **5**, 2031–2036.
 Suttle, D.P. and Stark, G.R. (1979) *J. Biol. Chem.*, **254**, 4602–4607.
 Tabor, S. and Richardson, C.C. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4767–4771.
 Thelander, L. and Berg, P. (1986) *Mol. Cell. Biol.*, **6**, 3433–3442.
 Thelander, L. and Reichard, P. (1979) *Annu. Rev. Biochem.*, **48**, 133–158.
 Thelander, L., Eriksson, S. and Åkerman, M. (1980) *J. Biol. Chem.*, **255**, 7426–7432.
 Thelander, M., Gräslund, A. and Thelander, L. (1985) *J. Biol. Chem.*, **260**, 2737–2741.

Wagner, M. (1986) *Trends Genet.*, **2**, 134–137.
 Wright, J.A., Alam, T.G., McClarty, G.A., Tagger, A.Y. and Thelander, L. (1987) *Somat. Cell Mol. Genet.*, **13**, 155–165.
 Yang-Feng, T.L., Barton, D.E., Thelander, L., Lewis, W.H., Srinivasan, P.R. and Francke, U. (1987) *Genomics*, **1**, 77–86.
 Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene*, **33**, 103–119.

Received on April 20, 1989