JBC Papers in Press. Published on April 5, 2000 as Manuscript M909762199

Mammalian Mitochondrial Ribosomal Proteins (4): AMINO ACID SEQUENCING, CHARACTERIZATION, AND IDENTIFICATION OF CORRESPONDING GENE SEQUENCES

running title: mammalian mitochondrial ribosomal proteins

Thomas W. O'Brien[‡]*, Jiguo Liu[‡], James E. Sylvester[¤]*, Edward B. Mougey[¤]*, Nathan Fischel-Ghodsian[¶]*, Bernd Thiede[¢], Brigitte Wittmann-Liebold[¢]*, and Hanns-Rüdiger Graack[¥]*[&]

From the [‡]Department of Biochemistry and Molecular Biology, University of Florida, Gainesville, Florida 32610-0245, [¤]Nemours Children's Clinic, Jacksonville, Florida 32207, [¶]Ahmanson Department of Pediatrics, Cedars-Sinai Medical Center and UCLA School of Medicine, Los Angeles, California 90048-1865, [¢]Max-Delbrück-Center for Molecular Medicine, Robert-Rössle-Sraße 10, D-13125 Berlin, Germany, [¥]Institute for Biology-Genetics, AG Kress, Free University of Berlin, Arnimallee 7, D-14195 Berlin, Germany.

*The Mammalian Mitochondrial Ribosomal Consortium: http://www.med.ufl.edu/biochem/tobrien/mmrc/index.htm

The peptide sequences reported in this paper have been submitted to the MIPS Data Bank with accession numbers S78761 (MRP-L18$_{bov}$, fragments), S78772 (MRP-L43$_{bov}$, fragments), S78762 (MRP-S12$_{bov}$, fragment), S78763 (MRP-S18$_{bov}$, fragment), S78766 (MRP-S28$_{bov}$, fragment).

Copyright 2000 by The American Society for Biochemistry and Molecular Biology, Inc.

[&]Corresponding author: Hanns-Rüdiger Graack, AG Kress, Institute for Biology - Genetics, Free University of Berlin, Arnimallee 7, D-14195 Berlin, Germany. Tel: +49-30-838-2629, Fax: +49-30-838-3649, Email: graack@zedat.fu-berlin.de

SUMMARY

Mitochondrial ribosomal proteins (MRPs) are required for the translation of all 13 mitochondrial encoded genes in humans. It has been speculated that mutations and polymorphisms in the human MRPs may be a primary cause of some oxidative phosphorylation disorders, or modulate the severity and tissue specificity of pathogenic mitochondrial DNA mutations. Although the sequences of most of the yeast MRPs are known, only very few mammalian, and nearly no human MRPs have been completely characterized. MRPs differ greatly in sequence, and sometimes biochemical properties, between different species, not allowing easy recognition by sequence homology. Therefore, the Mammalian Mitochondrial Ribosomal Consortium is using a direct approach of purifying individual mammalian (bovine) MRPs, determining their N-terminal and/or internal peptide sequences using different protein sequencing techniques, and using the resulting sequence information for screening EST and genomic data bases to determine human, mouse and rat homologues of the bovine proteins. Two proteins of the large and 3 proteins of the small ribosomal subunit, respectively, have been analysed in this manner. Three of them represent „new“, i. e. formerly unknown mammalian mitochondrial ribosomal protein classes. Only one out of these three different MRPs shows significant sequence similarities to known ribosomal proteins. In one case, the corresponding human genomic DNA sequences were found in the data bases, and the exon/intron structure was determined.

INTRODUCTION

Human oxidative phosphorylation disorders have over the last decade shown to be the cause of a great variety of inherited and acquired diseases, including such different clinical entities as systemic neuromuscular disorders, diabetes, aplastic anemia, deafness, and degenerative disorders by either mitochondrial or nuclear gene mutations (1,2). It has been speculated that mutations or polymorphisms in proteins involved in mitochondrial RNA processing and translation can be involved in some of these diseases, either as the primary cause, or as factors modulating the severity and/or tissue specificity of them (3).

The proper expression of the mitochondrial encoded protein genes depends on the nuclear encoded components of the mitochondrial translational system (4). In yeast, knockout mutants of nuclear encoded MRPs[1] lose their mitochondrial DNA, changing the mitochondrial genetical status to ρ^o (5), an effect also observed in mammalian nuclear mutations affecting mitochondria (6). Mammalian mitochondrial ribosomes differ significantly from other known bacterial and eukaryotic cytoplasmic ribosomes when analyzed by biochemical methods (7-10). Numbers and sizes of rRNA molecules are reduced as compared to bacterial ribosomes, and numbers of ribosomal proteins are elevated (11). However, the differences turned out to be even greater than expected when molecular biological and protein biochemical investigations began to reveal the amino acid and gene sequences of MRPs. These efforts are supported by the recent genome projects. Most of the yeast MRPs were identified by the comparison of peptide

---

[1] Abbreviations used: MRP, Mitochondrial ribosomal protein(s); 2D PAGE, two-dimensional polyacryle amide gel-electrophoresis; (k)bp, (kilo)base pair(s); MISP, mitochondrial import signal peptide; MM, molecular mass.

sequences obtained from purified mature yeast MRPs by N-terminal sequencing, with the protein sequences deduced from the yeast genome project data (reviewed in ref. 12). However, the assumption that yeast MRPs and mammalian MRPs are similar to each other as is the case for yeast and rat cytoplasmic ribosomal proteins, and that it may be possible to identify mammalian MRPs by using yeast MRP probes as screening devices is not valid. When the current project of mammalian MRP gene identification was launched in 1997 only very few mammalian MRPs were known on the molecular biological level, rather by chance than by a systematic approach. The mammalian mitochondrial homologue of the bacterial L3 ribosomal protein was identified as an overexpressed protein in Mahlavu hepato-carcinoma cells (13, 14). The mammalian homologues RPMS12 of the EcoS12 protein, a protein strongly conserved through evolution, were cloned by sequence similarities to the EcoS12 ribosomal protein and the S12 homolog of *Drosophila melanogaster* mitochondria (15, 16). L23MRP was identified by its sequence similarity to the EcoL23 ribosomal protein (17). The mammalian mitochondrial homologue of the strongly conserved EcoL7/L12 gene was cloned as a delayed-early expressed gene (18). A HSMRPS14 cDNA similar to the EcoS14 ribosomal protein was cloned (acc. no. Z99297) but not further characterized. However, none of the identified mammalian MRPs was „new“ in terms of lacking sequence similarities to known ribosomal proteins.

To characterize mammalian MRPs systematically, the Mammalian Mitochondrial Ribosomal Consortium was formed, and the initial primary experimental approach was to be based on the N-terminal sequencing of purified mature bovine MRPs. Using the obtained peptide sequence information, EST and genomic DNA data bases are screened,

and cDNA sequences are assembled *in silico*. This approach takes advantage of the existing bovine model for MRPs (7, 19), and the rapidly growing sequence data bases of human and other organisms. The obtained sequences were characterized by comparison to known ribosomal protein sequences and corresponding genomic DNA sequences were identified. Nineteen different groups of homologous mammalian MRPs have been determined so far (20-22[2]). Only 10 of them show significant sequence similarities to yeast MRPs and/or bacterial ribosomal proteins. This manuscript describes the identification of 5 mammalian MRPs.

[2] J. Anders, H.-R. Graack, J. Liu, and T. W. O'Brien, manuscript in preparation.

EXPERIMENTAL PROCEDURES

Analysis of bovine MRPs-Isolation and purification of bovine MRPs has been described (19). The numbering of the proteins follows (7). Large subunit proteins were stained with Coomassie Blue after 2D PAGE. Individual spots were excised from the gel and subjected to in-gel digestion by trypsin according to the method of Otto *et al*. (23). Resulting peptides were isolated and concentrated, purified by reverse phase HPLC and subjected to Edman sequencing (23) or individual peptides were subjected to tandem mass spectroscopic (MS/MS) analysis with a Q-TOF Micromass, Manchester, UK) equipped with a nanoflow Z-spray ion source. It should be noted that the latter method is not able to differentiate between leucine and isoleucine because of the identical MM of these two amino acids. Small subunit proteins were electroblotted from a 2D PAGE gel onto PVDF membrane, visualized by Coomassie Blue staining, and cut from the membrane. Individual spots obtained were subjected to N-terminal amino acid sequencing in an Applied Biosystems Model 494 Protein Sequencer.

Computing-Virtual screening of public EST databases was performed using the blast program of Altschul *et al*. (24) and the NCBI server. For screening of short peptides the advanced BLAST program was performed using the modified options "expected": 1000, and "other options": -e2. The analysis of the obtained sequences, the sequence comparison, and the assembly of virtual consensus cDNAs was performed as described (20) using several analytic programs of the GCG DNA analysis computer software package (25). Alignments were performed using the „bestfit" program, and multiple alignments were assembled manually. The algorithm used in the „pileup" program

classifies weak or strong „similarities“ of amino acid residues by the frequencies of exchanges in a given collection of sequences and not by chemical similarities. Therefore, classifications of amino acid exchanges termed „weak“ or „strong“ in the presented alignment figures were made manually regarding chemical similarities. For example, Ser vs. Thr are strongly conserved amino acid residues, and glutamic acid vs. glutamine represent weakly conserved amino acid residues. BLAST searches for homologous *Caenorhabdidis elegans* protein sequences were performed on the Sanger Center's server. Analysis of N-terminal peptide sequences for signal peptide properties was done using the Signal V1.1 Center for Biological Sequence Analysis server (26).

RESULTS AND DISCUSSION

Protein sequencing of individual MRPs-Proteins of the bovine small and large mitochondrial ribosomal subunits, respectively, were purified from bovine liver and sequenced (see MATERIALS AND METHODS). Seven different peptide sequences were obtained from 5 individual proteins (Table I). The MRP-$L43_{bov}$ protein turned out to be the corresponding bovine MRP to the previously characterized MRP-$L32_{rat}$ (20). The bovine MRP-L43 peptide 1 makes a perfect match in the C-terminal part of the mammalian MRP-L32s, and the bovine peptide 2 matches close to the mature N-termini of the MRP-$L32_{mammalian}$s (see ref. 20 for comparison). The MRP-$S18_{bov}$ protein corresponds to the previously published MRP-$S13_{mammalian}$ (20). However, comparison of the mature N-termini of MRP-$S18_{bov}$ and MRP-$S13_{rat}$ showed that the true N-terminus of mammalian MRP-S13s may be positioned one amino acid residue further upstream since the MRP-$S18_{bov}$ sequencing seems to be more reliable (20). Thus, we can combine the different model systems used for mammalian MRPs by the presented method (7, 20).

Characterization of gene sequences corresponding to individual mammalian MRPs-Peptide sequence information obtained was used to screen public EST data bases. By multiple screenings and comparison of ESTs from human, mouse, and rat sources consensus cDNAs were assembled electronically (Table II). Due to the identification of false positives for MRP-$S12_{human}$ by the initial EST data base screening using the MRP-$S12_{bov}$ peptide sequence as virtual screening probe, the MRP-$S12_{mouse}$ amino acid sequence (below) was used to screen for corresponding human ESTs. For all three new

MRPs (MRP-L18, MRP-S12, and MRP-S28) complete ORFs were deduced of mouse and human origin. Only partial sequences were found for the corresponding rat MRPs (Table II, FIG. 2). Additionally, two corresponding genomic clones were found in the data bases for $MRP\text{-}S12_{human}$. The gene is located in chromosomal position 16q13-21 very close to the DR-nm23 Nucleoside Diphosphate Kinase homologous gene (27). Genomic and EST sequences were assembled (Table II) and compared, permitting the exon/intron structure of the $MRP\text{-}S12_{human}$ gene to be determined (FIG. 1). The gene spans approximately 1500 bp. However, the 5' exon A containing 53 bp is not covered by the published genomic sequence: Exon B extends from bp 54-302, exon C from bp 303-342, exon D from bp 343-662, and exon E from bp 663-804. Other human MRP genes whose genomic structure and chromosomal location have been determined similarily are the $MRP\text{-}L22_{human}$ (20), located in position 22q11, and the $MRP\text{-}L5_{human}$ gene in position 21q22.1 (22). The chromosomal positions are being examined for putative linkages to known inherited diseases caused by mitochondrial disorders.

For human MRP-S28 a full length cDNA clone HSPC007 (accession no. AF070663) was identified. This clone was sequenced and an ORF of 188 amino acid residues was determined by the authors but it wasn't assigned to a specific function. Similarily, other mammalian MRP genes that had been sequenced prior to this project of systematic MRP gene characterization were identified the same way: The human RLX1 gene was sequenced (28) but the function of the deduced protein remained unknown until the annotation of this gene to be the human MRP-L15 homologue (21). The KIAA0264 cDNA (29) was shown to encode the human MRP-S4[2], the cDNA CGI-22 (acc. no. AF132956) corresponds to the $MRP\text{-}L14_{human}1$ cDNA (22) although a sequencing error in

the CGI-22 cDNA made the annotation more difficult. Except the MRP-L14$_{human}$1/CGI-22 protein, which is homologous to the EcoL2 ribosomal protein, all other genes mentioned are „new“ in respect that they have no counterparts in known ribosomal systems. These results (i) demonstrates the power of the methodology applied to determinate the identity of some ESTs as coding for MRPs, (ii) to identify unknown „new“ mammalian MRP genes and (iii) sheds some light on the numbers of „new“ genes to be expexted among the mammalian MRPs in total.

Characterization of consensus cDNAs-Assembly of consensus cDNAs from EST sequences is limited by the poor quality of EST sequences, especially in the further 5’ coding regions. Therefore, we suspect most of the assembled consensus cDNAs are incomplete at their respective 5’ ends. This phenomenon is observed generally when using ESTs for cDNA assembly (20-22). From an experience, the decreasing numbers and diminishing qualities of ESTs derived from the 5’ regions of long mRNAs pose problems for the successful assembly of cDNAs longer than about 1500 bp from ESTs of public (!) data bases. At their respective 3’ ends, the consensus cDNAs frequently show canonical poly-adenylation signals, AATAAA, and parts of the corresponding poly(A) tracks. The ORFs presented (Table II) are complete, and in the cases of MRP-L18 and MRP-S12 5’ preceeding in-frame stop codons were found 5’ to the initiation codons.

It turned out that MRP-S12$_{human}$ exists in two different splice variants (FIG. 2*c*). For MRP-S12$_{human}$1 a consensus cDNA of 847 bp was assembled. Since the N-terminus of the MRP-S12$_{human}$1 sequence was only found in the genomic DNAs, it explains why the gene

was not identified by the initial EST data base search. For MRP-$S12_{human}2$ a consensus cDNA of 662 bp was assembled (Table II). Although its ORF is open at its 5' end, it shows strong sequence similarity to both MRP-$S12_{human}1$ and MRP-$S12_{mouse}$ except for a deletion of 26 amino acid residues in length. Putatively, this is a result of alternative splicing due to the fact that the respective DNA sequences show the common features of intron sequences including canonical donor and acceptor splice sites and a poly-pyrimidine tract close to the 3' end of the putative intron (data not shown). However, among more than 20 independent EST sequences discovered for MRP-$S12_{mouse}$, no sequence corresponding to the splice variant of MRP-$S12_{human}2$ was determined. Similar to mouse, MRP-$S12_{rat}$ shows no splicing variant corresponding to MRP-$S12_{human}2$. Therefore, we can only speculate whether MRP-$S12_{human}2$ is a naturally occuring splice variant different between human and mouse, if it belongs to an incomplete, non-functional variant of MRP-$S12_{human}$, or if it is a cloning artifact.

Properties of the deduced ORFs-Comparison of the N-terminal peptides determined by amino acid sequencing with the deduced ORFs revealed the existence and extensions of the putative MISPs (FIG. 2). As it was proposed[2], mammalian MRP import may be mediated through different import signal recognition mechanisms and N-terminal processing during or after import. Also, the MISPs presented in this paper belong to different classes according to the classification of Branda and Isaya (30). MRP-L18s show MISPs of the R-none class, MRP-$S28_{mouse}$ has an R-3 MISP while MRP-$S28_{human}$ possess an R-none MISP. Mammalian mitochondrial protein import seems to be at least

as heterogeneous as in yeast (12): For human and mouse MRP-L18, we postulate MISPs of 9 amino acid residues, respectively (FIG. 2*a*). Although such short MISPs are rather unusual for mammalian imported mitochondrial proteins, the stop codons preceeding the respective initiator methionines, and the N-terminal sequence of the $\text{MRP-L18}_{\text{bovine}}$ give no other choice. The MISPs does not fit the general features of an import signal peptide, and the cleavage site obtained experimentally was not identified by the SignalP computer program (26). In contrast, mammalian MRP-S12s have no substantial cleavable MISPs; only the N-terminal methionine residues are cleaved off in the mature forms (FIG. 2*b*). Although this feature is not unusual with yeast MRPs (12), it is the first example for a mammalian MRP that is only processed on the N-terminal methionine (FIG. 2*b*). For mouse and human MRP-S28 and HSPC007, respectively, we postulate cleavable MISPs of 72 amino acid residues. Interestingly, the MISPs of these two proteins are much less conserved as compared with the conservation of the mature proteins, i. e., after cleavage of the MISPs (FIG. 2*c*). Both MISPs show the common features of mitochondrial import signal sequences such as large numbers of hydroxylated, hydrophobic and positively charged amino acid residues. However, the SignalP program identfied most probable cleavage sites between amino acid residues nos. 31 (T) and 32 (E) of mouse and human MRP-S28, respectively (26). These results do not fit the biochemical data. Both MISPs contain the N-terminal MAA motif which has been shown to be common among N-termini of mammalian MISPs for MRPs[2]. However, only one out of the presented 3 new mammalian MRP classes shows the „MAA“ motif in their respective N-termini (MRP-S28, FIG. 2*c*).

After N-terminal processing, the presented MRPs show more (pI 11.08, MRP-$S12_{mouse}$) or less (pI 7.44, MRP-$S28_{mouse}$) basic characters, as it is common for ribosomal proteins. The calculated MMs of the mature human and mouse proteins correspond very well to the MMs of the bovine MRPs (Table I) determined by SDS PAGE, showing that MRPs of mammals are closely related in their molecular properties.

Evolution of MRPs-Human and mouse MRPs deduced pairwise from the presented cDNAs are identical in length (FIG. 2). They share between 72.5 % (MRP-S28) and 87.5 % (MRP-L18) identical amino acid residues, comparable to the MRP sequence conservation previously observed among mammalian MRPs (20-22, 31). These results are in strong contrast to the sequence conservation observed between cytoplasmic ribosomal proteins of the same species. Cytoplasmic ribosomal proteins of rat and man are close to 100 % identical (32). The rate of evolutionary divergence among mammalian MRPs was estimated to be 13 times higher than that of mammalian cytoplasmic ribosomal proteins (33). Consequently, MRPs themselves are very heterogeneous among species (34). This is reflected in the low sequence identities determined between yeast and mammalian MRPs and by the experimental problem to identify mammalian MRP genes using their homologous MRP probes from yeast for screening (20-22, 31, this work). Curiously, mitochondrial ribosomes of distantly related species like yeast and mammals have similar numbers of different MRPs but only a subset of their respective MRPs is really homologous. Instead, mitochondrial ribosomes of yeast and mammals have many MRPs that are *unique* to one or the other (group of) organism(s). Two out of the 3 of the new MRPs presented in this paper have no counterparts in other ribosomal

protein systems at all (MRP-S12, MRP-S28). Only the mammalian MRP-L18 proteins showed significant sequence similarities to a known ribosomal protein from *E. coli* (EcoL24). However, the sequence conservation is poor. 30.5 % of a stretch of 87 amino acid residues of EcoL24 are identical to the MRP-$L18_{human}$ sequence. The mammalian counterparts are elongated into both N- and C-terminal directions resulting in a protein more than twice the size of the EcoL24 protein (FIG. 2*b*). These values are comparable to those that have been determined for other homologous MRPs of mammals and yeast (20-22, 31). However, only half of the mammalian MRPs known on the molecular level to date have counterparts in the yeast mitochondrial and/or *E. coli* ribosomal system at all (Table III). Interestingly, mammalian MRPs which are similar to „new" yeast MRPs that have no bacterial counterparts are very rare. MRP-L27 (mammals)/YmL27 (yeast; ref. 20) is the only example. The yeast MRPs seem to be much more distantly related to the mammalian MRPs than the cytoplasmic ribosomal proteins of the same species - rat and yeast cytoplasmic ribosomal proteins share 40 % - 80 % identical amino acid residues (32). For several of the mammalian MRPs which have been investigated so far in that matter, corresponding sequences have been identified in the *C. elegans* data bases (Table III). However, percentages of sequence identities between the worm and mammalian MRPs do not exceed 45 % and are confined to relatively short positions and match only stretches of the complete mammalian MRP sequences. Mitochondrial ribosomes must be much, much more divergent among species than the cytoplasmic ribosomes from the same species. Finally, these results raise questions why and how two different multi-molecular complexes that perform the same function - synthesis of proteins - but are separated only by two (mitochondrial) membranes evolved two different rates of

evolution. The explanation of poor DNA repair mechanisms of the mitochondria is not satisfactory since all mitochondrial and cytoplasmic ribosomal proteins are encoded by nuclear genes (in mammals). How did mitochondrial ribosomes of yeast and mammals adopt so many additional and different proteins? Did they aquire additional proteins, or did they simply not loose proteins during evolution which have been omitted by e. g. the bacteria to keep a streamlined, high-throughput but poorly regulated ribosome? We hope that the discovery of the complete set of mammalian MRPs in the future will provide some insight to these questions.

Functions of mammalian MRP-L18s-So far, the functions of mammalian MRPs have been eludicated only indirectly. Some analogous functions can be supposed if the MRP is homologous to a known ribosomal protein of, e. g., bacterial sources. MRP-$L18_{mammalian}$ is homologous to EcoL24. EcoL24 has been characterized as an early assembly protein of the large ribosomal subunit, binding directly to the 13S rRNA (35). This coincides with the characterization of the MRP-$L18_{bov}$ as a core protein of the ribosomal subunit that is found bound to the subunit core after washes with 4 M LiCl (36). However, MRP-$L18_{mammalian}$ is more than twice the size of EcoL24, raising the possibility that additional functions for this protein may be exist.

CONCLUSION

Five individual bovine MRPs have been purified and partially sequenced. Two of them, MRP-$L43_{bov}$ and MRP-$S18_{bov}$, turned out to be homologous to the recently published

mammalian MRP-L32s and MRP-S13s, respectively (20). Based on the obtained peptide information, corresponding EST and genome DNA sequence were identified *in silico* and the virtual cDNAs for human, mouse (and partially rat) were assembled. Deduced MRP sequences were characterized corresponding to the biochemical data obtained from bovine MRPs. MISPs postulated by comparison of bovine N-terminal peptide data and the deduced mammalian MRPs suggested different mechanisms for mammalian MRP import as it has been proposed for yeast MRPs (12). Only one (MRP-L18) out of three newly characterized mammalian MRPs showed significant sequence similarities to a known ribosomal protein class (EcoL24). Our data demonstrate the poor conservation of MRPs among different distantly related species such as yeast and mammals (Table III); thus raising questions concerning the mechanisms of mitochondrial ribosomal evolution compared to those of cytoplasmic ribosomes.

ACKNOWLEDGMENTS

TWO and NFG gratefully acknowledge support by NIH/NIDCD grant RO1-DC04092.

JES and EBM acknowledge support from the Nemours Research Program.

REFERENCES

1. MITOMAP. http://infinity.gen.emory.edu/mitomap.html

2. Wallace, D. C. (1999) *Science* **283**, 1482-1488

3. Fischel-Ghodsian, N. (1998) *Mol. Genet. Metab*. **65**, 97-104

4. Costanzo, M. C,. and Fox, T. D. (1990) *Annu. Rev. Genet*. **24**, 91-113

5. Myers, A. M., Pape, L. K., and Tzagoloff, A. (1985) *EMBO J*. **4**, 2087-2092

6. Grossman, L. I., and Shoubridge, E. A. (1996) *BioEssays* **18**, 983-991

7. Matthews, D. E., Hessler, R. A., Denslow, N. D., Edwards, J. S., and O'Brien, T. W. (1982) *J. Biol. Chem.* **257**, 8788-8794

8. Graack, H.-R., Grohmann, L,. and Choli, T. (1988) *FEBS Lett.* **242**, 4-8

9. O'Brien, T. W., Denslow, N. D., Anders, J. C., and Courtney, B. C. (1990) *Biochim. Biophys. Acta* **1050**, 174-178

10. Cahill, A., Baio, D. L., and Cunningham, C. C. (1995) *Anal. Biochem.* **232**, 47-55

11. Kitakawa, M., and Isono, K. (1991) *Biochimie* **73**, 813-825

12. Graack, H.-R., and Wittmann-Liebold, B. (1998) *Biochem. J.*, **329**, 433-448

13. Ou, J.-H., Yen, T. S. B., Wang, Y.-F., Kam, W. K.. and Rutter, W. J. (1987) *Nucl. Acids Res.* **15**, 8919-8934

14. Graack, H.-R., Grohmann, L., Kitakawa, M., Schäfer, K.-L. and Kruft, V. (1992) *Eur. J. Biochem.* **206**, 373-380

15. Shah, Z.H., O'Dell, K.M.C., Miller, S.C.M., An, X. and Jacobs, H.T. (1997) *Gene* **204**, 55-62

16. Johnson, D. F., Hamon, M., and Fischel-Ghodsian, N. (1998) *Genomics* **52**, 363-368

17. Tsang, P., Gilles, F., Yuan, L., Kuo, Y. H., Lupu, F., Samara, G., Moosikasuwan, J., Goye, A., Zelenetz, A. D., and Tycko, B. (1995) *Hum. Mol. Genet.* **4**, 1499-1507

18. Marty, L., and Fort, P. (1996) *J. Biol. Chem.* **271**, 11468-11476

19. O'Brien, T. W., and Denslow, N. D. (1996) *Meth. Enz.* **264**, 237-248

20. Goldschmidt-Reisin, S., Kitakawa, M., Herfurth, E., Grohmann. L., Wittmann-Liebold, B., and Graack, H.-R. (1998) *J. Biol. Chem.*, **273**, 34828-34836

21. Graack, H.-R., Bryant, M. L., and O'Brien, T. W. (1999) *Biochemistry*, **38**, 16569-16577

22. O'Brien, T. W., Fiesler, S. E., Denslow, N. D., Thiede, B., Wittmann-Liebold, B., Sylvester, J. E., Mougey, E. B., and Graack, H.-R. (1999) *J. Biol. Chem.*, **274**, 36043-36051

23. Otto, A., Thiede, B., Müller, E.-C., Scheler, C., Wittmann-Liebold, B., and Jungblut, P. (1996) *Electrophoresis* **17**, 1643-1650

24. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) *Nucl. Acids Res.* **25**, 3389-3402

25. GCG WISCONSIN PACKAGE Version 9.1-OpenVMS, September 1997

26. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) *Protein Engineering* **10**, 1-6

27. Martinez, R., Venturelli, D., Perrotti, D., Veronese, M. L., Kastury, K., Druck, T., Huebner, K., and Calabretta, B. (1997) *Cancer Res.* **57**, 1180-1187

28. Nagase, T., Miyajima, N., Tanaka, A., Sazuka, T., Seki, N., Sato, S., Tabata, S., Ishikawa, K. I., Kawarabayasi, Y., Kotani, H., and Nomura, N. (1995) *DNA Res.* **2**, 37-43

29. Nagase, T., Seki, N., Ishikawa, K., Ohira, M., Kawarabayasi, Y., Ohara, O., Tanaka, A., Kotani, H., Miyajima, N., and Nomura, N. (1996) *DNA Res.* **3**, 321-329

30. Branda, S. S., and Isaya, G. (1995) *J. Biol. Chem.* **270**, 27366-27373

31. Koc, E. C., Blackburn, K., Burkhat, W. and Spremulli, L. L. (1999) *Biochem. Biophys. Res. Comm.* **266**, 141-146

32. Wool, I. G., Chan, Y.-L., and Glück, A. (1995) *Biochem. Cell Biol*. **73**, 933-947

33. Pietromonaco , S. F., Hessler, R. A., and O'Brien, T. W. (1986) *J. Mol. Evol.* **24**, 110-117

34. Matthews, D. E., Hessler, R. A., and O'Brien, T. W. (1978) *FEBS Lett.* **86**, 76-80

35. Herold, M., and Nierhaus, K. H. (1987) *J. Biol. Chem.* **262**, 8826-8833

36. Schieber, G. L., and O'Brien, T. W. (1982) *J. Biol. Chem.* **257**, 8781-8787

FIGURE LEGENDS

FIG. 1. **Exon/intron structure of the human MRP-S12 gene.** The genomic DNA is shown as a horizontal line. Boxes mark location of cDNA sequences; black boxes represent translated cDNA sequence, open boxes show untranslated cDNA sequences. Exons are termed consecutively B, C, D, and E. Exon A is missing in this Figure. Numbers give genomic DNA positions in the corresponding clone (accession no. Z34802).

FIG. 2. **Alignment of deduced mammalian MRP amino acid sequences, and with ribosomal protein sequences of other sources.** Obtained deduced mammalian MRP amino acid sequences are aligned with bovine peptide sequences (Table I) and corresponding ribosomal protein sequences of *E. coli*. *Numbers* give the respective amino acid positions. *Lines* (|) mark identical amino acid positions, *colons* (:) mark strongly conserved amino acid residues, and *peroids* (.) mark weakly conserved amino acid residues. *Dashes* (-) show N- or C-terminal ends of incomplete amino acid sequences, and *asterisks* (*) mark stop codons. Amino acid residues in lowercase letters within the bovine peptide sequences are uncertain in their determination by amino acid sequencing (see Table I). *x* marks unidentified amino acids at particular positions in the sequence. a, alignment of mammalian MRP-L18 sequences with the *E. coli* EcoL24 sequence (accession no. Sw:rl24_ecoli). b, alignment of the mammalian MRP-S12 protein sequences. c, alignment of the mammalian MRP-S28 protein sequences.

TABLE LEGENDS

Table I. *Peptide sequences of mature bovine mitochondrial ribosomal proteins obtained by amino acid sequencing.* MM, molecular mass in kiloDaltons. Amino acids residues separated by a slash (/) and in brackets, both residues are possible. *x*, the amino acid could not be identified although the position in the respective peptide sequence is correct. Amino acid residues in lower case, determination is uncertain. (N-Term.), peptide sequence is determined from the N-terminus of the mature protein.

Table II. *Assembly of identified EST sequences: consensus cDNAs in 5' to 3' direction.* Determination of consensus cDNA sequences for deduction of mammalian MRP ORFs by assembly of EST sequences. *Asterisks* (*) label incomplete ORFs. r,c, the nucleotide sequence was determined in reverse complement orientation. Ribosomal protein family, affiliation of deduced ORFs to existing families of similar ribosomal proteins.

Table III. *Homologous ribosomal proteins from mitochondria and E. coli.* Corresponding genes, or names of contigs or accession numbers (*C. elegans*) are listed. *no*, no corresponding protein has been found in yeast and/or *E. coli*. *?*, existence of corresponding protein has not been investigated so far. References are given in parentheses. The naming of MRP-S7 is preliminary according to (37). (a), accession no. Z99297. (b), this work.

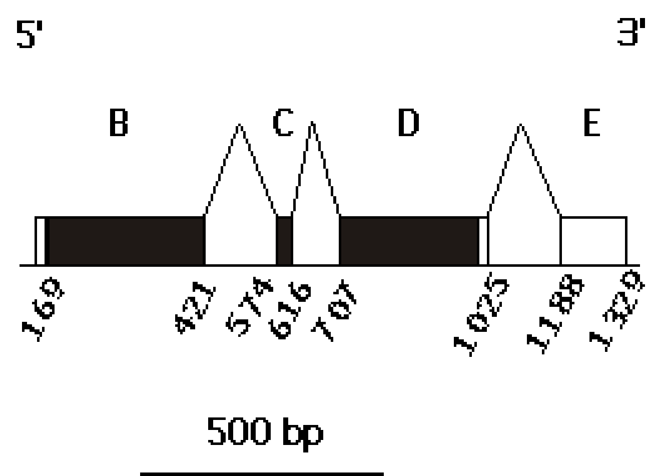

Figure 1
O'Brien et al.

a MRP-L18

```
bovine peptides   ASKVTLPQNYRYGMxQ                      rat       -GDMVEILEGKDAGKQGKVVQVVRQRNWVVLEGLNTHYRYIGR    42
                  |||.|    |||||                                   ||||||||||||||||||||||||||||||||||||||||||
mouse    MRLSALLALASKATSSPFYRYGMSRPGSIADKRKNPPWSRRRQVVVEPISDEDWHLFCGDMVEILEGKDAGKQGKVVQVVRQRNWVVLEGLNTHYRYIGR   100
         ||||||||||||.|  | |||||| |||:||||||||| ||| ||||||||||| ||||| ||||||||||||||||||.|||||||. ||||||||||:
human    MRLSALLALASKVTLPPHYRYGMSPPGSVADKRKNPPWIRRRPVVVEPISDEDWYLFCGDTVEILEGKDAGKQGKVVQVIRQRNWVVVGGLNTHYRYIGK   100
                                                                    | | :| ||| || ||| :|:.. . |:| |:|.  :
E. coli EcoL24                                              MAAKIRRDDEVIVLTGKDKGKRGKVKNVLSSGK.VIVEGINLVKKHQKP    48


bovine          GTMVPSEAPL
                |||.||||||
rat      TK..DHRGTMIPSEAPLLHHQVKLVDPVDRKPTEIQWRFTEAGERVRVSTRSGRIIPKPEFPRADGIVPETWTDGPKDTSVEDALERTYVPRLKTLEEDV   140
         ||  ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
mouse    TK..DHRGTMIASEAPLLHHQVKLVDPVDRKPTEIQWRFTEAGERVRVSTRSGRIIPKPEFPRADGIVPETWTDGPKDTSVEDALERTYVPRLKTLEEDV   198
         |   | ||||| |||||||:|||||||.|||||||.|||||||||||||||||||||||||||||||||||| |||||||||||||||||| ||||.|:|
human    TM..DYRGTMIPSEAPLLHRQVKLVDPMDRKPTEIEWRFTEAGERVRVSTRSGRIIPKPEFPRADGIVPETWIDGPKDTSVEDALERTYVPCLKTLQEEV   198
                | ..  || .   :| ... .  |   . .|| | | :||.   .   |
EcoL24   VPALNQPGGIVEKEAAIQVSNVAIFNAATGKADRVGFRF.EDGKKVRFFKSNSETIK*   104


rat      MEAMGIQETRRFKKIYWY*   158
         ||||||||||||||.|||
mouse    MEAMGIQETRRFKKVYWY*   216
         ||||||.|||::||||||
human    MEAMGIKETRKYKKVYWY*   216
```

b MRP-S12

```
bovine N-terminal peptide
         ARRKVRPRL                                             rat        -ARFPLFGLGRLVTRKSWLWQHDEPCYWRLTRV    32
         ||:||||||                                                         ||.|||||||||||||||||||||||| ||||
mouse   MARKKVRPRLIAELARRVALR.QRNQPRDSQLYALDYETLTRPHSGRRLPVRAWADVRRESASYNVLARLPLFGLGRLVTRKSWLWQHDEPCYWSLTRV    99
        ||||||||||||||||||||| | |.||||||||.|||||||| |||||||||||||||||    .| ||| |||||||||||||||||||||| ||||
human1  MARKKVRPRLIAELARRVALREQLNRPRDSQLYAVDYETLTRPFSGRRLPVRAWADVRRESRLLQLLGRLPFFGLGRLVTRKSWLWQHDEPCYWRLTRV   100
                                     |||||||||| ||||||||                          ||||||||||| |||||||||||||
human2                              -SQLYAVDYETxTRPFSGRR.........................SLGRLVTRKSWLxQHDEPCYWRLTRV    45


rat     RPDYTAQNLDHGRAWGILTFKGKSEETAREIEQVMYHDWRLVPKHEEEAFTAFTGKAEDRLNLVPYPPLLRAMILAERQKNGDTSVEEPLLNLERTRMRP   132
        |||||||||||||||||||||||||:|||||||||||||||||||||||||||| | ||||| ||||||||||||||||||||||||.||||||||||||
mouse   RPDYTAQNLDHGRAWGILTFKGKSEDTAREIEQVMYHDWRLVPKHEEEAFTAFTAKPEDRLNSVPYPPLLRAMILAERQKNGDTSVQEPLLNLERTRMRP   199
        ||||||||||||:||||||||||.|  ||||| |||||||||||||||||||||  ||| :     ||||||||:|||||||||| :||:||.:| || |
human1  RPDYTAQNLDHGKAWGILTFKGKTESEAREIEHVMYHDWRLVPKHEEEAFTAFTPAPEDSMPPCRTPPLLRAMIIAERQKNGDTSTEEPMLNVQRIRMEP   200
        ||||| |||| ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
human2  RPDYTxQNLD.GKAWGILxFKGKTESEAREIEHVMYHDWRLVPKHEEEAFTAFTPAPEDSLASVPYPPLLRAMIIAERQKNGDTSTEEPMLNVQRIRMEP   144


rat     WDYPAKQETKGRAKGTPV*   150
        ||||||||||||||||||
mouse   WDYPAKQETKGRAKGTPV*   217
        ||||||||||||||||||
human1  WDYPAKQEDKGRAKGTPV*   218
        ||||||||||||||||||
human2  WDYPAKQEDKGRAKGTPV*   162
```

Figure 2
O'Brien et al.

C **MRP-S28**

```
bovine N-terminal peptide                                                         gsPKNVESFASMl
                                                                                   ||| ||||||||
rat                                                                          -ARAE.SPKPVESFASMLRHSPLTQMGPA    27
                                                                               | | ||||||||||||||||||| |||
mouse      MAALCRSHAGTAGSRFLRALVFSKPLRNASTESGSESATHDSSAPRARSGGFASALERHSDLQRKADVRLE.SPKPVESFASMLRHSPLTQLGPA    94
           ||||||::|  | |:|||  .| :| |   ||||||| :  .  |: | | |||| ||||:| .| :     ||| ||||||||||||||| |||
HSPC007    MAALCRTRAVAAESHFLRVFLFFRPFRGVGTESGSESGSSNAKEPKTRAGVFASASERHSELLQKVEPLQKGSPKNVESFASMLRHSPLTQMGPA    95


rat        KDKLVIGRIFHIVEDDLYIDFGGKFHCVCKRPDVDGEKYQRGTRVRLRLLDLELTSRFLGGTTDTTILEADAVLLGLQESRDSKSREERLNKQ*   121
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||.  |
mouse      KDKLVIGRIFHIVEDDLYIDFGGKFHCVCKRPDVDGEKYQRGTRVRLRLLDLELTSRFLGGTTDTTILEADAVLLGLQEIRDSKSREEQPSK*   192
           ||||||||||||||.|||||||||||||||||:|||||||:||||||||||||||||||| |||||.|||.|||||.|| :||:|:||.  |
human      KDKLVIGRIFHIVENDLYIDFGGKFHCVCRRPEVDGEKYQKGTRVRLRLLDLELTSRFLGATTDTTVLEANAVLLGIQESKDSRSKEEHHEK*   193
```

Figure 2
O'Brien et al.

Protein name	Sequences	Molecular mass
MRP-$L18_{bov}$		
peptide 1	ASKVTLPQNY(R/Q)YG(M/Q)xQ (N-term.)	26.5
peptide 2	GTMVPSEAPL	
MRP-$L43_{bov}$		
peptide 1	(R/K)FDSNNVV(I/L)(I/L)EDNGNPVGTR	12.9
peptide 2	(R/K)VVDNSA(I/L)PNT	
MRP-$S12_{bov}$	ARRKVRPRL (N-term.)	27.5
MRP-$S18_{bov}$	RKTRHDPPAKSKIGxVATPP (N-term.)	20.5
MRP-$S28_{bov}$	(g/p/q)(s/r)PKNVESFASM(q/l) (N.term.)	13.0

O'Brien et al., Table 1

Assembly of identified EST sequences: consensus cDNA in 5' to 3' direction

gene	bps nos.	acc. no.	total length	ORF from no. to no.	r-protein family
MRP-L18$_{rat}$	1-600	AA892843	600	*41-514	EcoL24
MRP-L18$_{mouse}$	1-505	AA498399	985	213-860	EcoL24
	101-580	W48021			
MRP-L18$_{human}$	1-346	AL038493	960	156-803	EcoL24
	191-212	AA249773r,c			
	69-558	AI361046r,c			
	351-452	AA766104r,c			
MRP-S12$_{rat}$	1-651	AI231798	651	*2-451	new
MRP-S12$_{mouse}$	1-60	AA794791	907	49-699	new
	31-70	AI047975			
	38-162	AA059964			
	178-227	AA920808			
	212-261	AA059964			
	298-347	AI047975			
	311-360	AA059964			
	383-550	AA222183			
	92-405	AV026499			
MRP-S12$_{human}$1	7684-7920	HS371H6, r	847	23-676	new
	316-421	U80813			
	574-616				
	707-1025				
	1188-1329				

Assembly of identified EST sequences: consensus cDNA in 5' to 3' direction

MRP-S12$_{human}$2	1-208	AA323208	662	*2-487	new
	1-454	AI343550r,c			
MRP-S28$_{rat}$	1-411	AA893385r,c	411	*3-362	new
MRP-S28$_{mouse}$	1-228	AA138018	690	30-587	new
	213-474	AA036202			
	447-532	AA796533			
	88-201	AA273405			
MRP-S28$_{human}$	1-270	W56563	716	7-585	new
	91-536	AA999834r,c			

human	mouse	rat	bovine	C. elegans	yeast mito.	E. coli
large subunit						
MRL3 (13)	?	MRL3 (13)	?	?	YmL9	EcoL3
MRP-L2 (21)	MRP-L2 (21)	MRP-L2 (21)	MRP-L2 (21)	?	no	no
MRP-L3 (21)	MRP-L3 (21)	MRP-L3 (21)	MRP-L3 (21)	Y34D9.Contig160	no	no
MRP-L5 (22)	MRP-L5 (22)	MRP-L5 (22)	MRP-L5 (22)	Y62F5.ctg01193	no	no
MRP-L7 (22)	MRP-L7 (22)	MRP-L7 (22)	MRP-L7 (22)	Y92H12.Contig307	YmL10	EcoL15
MRP-L8 (20)	MRP-L8 (20)	MRP-L8 (20)	?	?	YmL11	EcoL10
MRP-L12(18)	MRP-L12 (21)	?	MRP-L31/34 (21)	C50132	YGL068w	EcoL7/L12
MRP-L14 (22)	MRP-L14 (22)	MRP-L14 (22)	MRP-L14 (22)	R08B6.ctg00286	Rml2p	EcoL2
RLX1 (21, 28)	MRP-L15 (21)	?	MRP-L15 (21)	?	no	EcoL19
MRP-L18 (b)	MRP-L18 (b)	MRP-L18 (b)	MRP-L18 (b)	C45339, U80448	no	EcoL24
MRP-L22 (20)	MRP-L22 (20)	MRP-L22 (20)	?	?	no	no
L23MRP (17)	L23mrp (17)	?	?	?	YmL41	EcoL23
MRP-L25 (20)	MRP-L25 (20)	MRP-L25 (20)	?	?	no	EcoL22
MRP-L26 (22)	MRP-L26 (22)	MRP-L26 (22)	MRP-L26 (22)	Y54E10.Contig 159	YmL8	EcoL17
MRP-L27 (20)	MRP-L27 (20)	MRP-L27 (20)	?	?	YmL27	no
MRP-L28 (20)	MRP-L28 (20)	MRP-L28 (20)	?	?	YmL33	EcoL30
MRP-L31 (20)	MRP-L31 (20)	MRP-L31 (20)	?	?	no	no
MRP-L32 (20)	MRP-L32 (20)	MRP-L32 (20)	MRP-L43 (b)	U29536	YmL38	EcoL14
small subunit						
MRP-S7 (31)	MRP-S7 (31)	MRP-S7 (31)	MRP-S7 (31)	?	MRP S7 (31)	EcoS7
MRP-S12 (b)	MRP-S12 (b)	MRP-S12 (b)	MRP-S12 (b)	Z34802	no	no
RPMS12 (15,16)	RPMS12 (15)	?	?	?	YNR036c	EcoS12
MRP-S13 (20)	MRP-S13 (20)	MRP-S13 (20)	MRP-S18 (b)	U10402	no	no
HSMRPS14 (a)	?	?	?	?	MRP2	EcoS14
MRP-S28 (b)	MRP-S28 (b)	MRP-S28 (b)	MRP-S28 (b)	Y116F11.ctg10791	no	no