Identification of Mammalian Mitochondrial Ribosomal Proteins (MRPs) by N-Terminal Sequencing of Purified Bovine MRPs and Comparison to Data Bank Sequences: The Large Subribosomal Particle^{†,‡}

Hanns-Rüdiger Graack,§ Michael L. Bryant, and Thomas W. O'Brien*

Department of Biochemistry and Molecular Biology, College of Medicine, University of Florida, Box J-245, Gainesville. Florida 32610-0245

Received July 6, 1999; Revised Manuscript Received October 5, 1999

ABSTRACT: Bovine mitochondrial ribosomes are presented as a model system for mammalian mitochondrial ribosomes. An alternative system for identifying individual bovine mitochondrial ribosomal proteins (MRPs) by RP-HPLC is described. To identify and to characterize individual MRPs proteins were purified from bovine liver, separated by RP-HPLC, and identified by 2D PAGE techniques and immunoblotting. Molecular masses of individual MRPs were determined. Selected proteins were subjected to N-terminal amino acid sequencing. The peptide sequences obtained were used to screen different databases to identify several corresponding MRP sequences from human, mouse, rat, and yeast. Signal sequences for mitochondrial import were postulated by comparison of the bovine mature N-termini determined by amino acid sequencing with the deduced mammalian MRP sequences. Significant sequence similarities of these new MRPs to known r-proteins from other sources, e.g., *E. coli*, were detected only for two of the four MRP families presented. This finding suggests that mammalian mitochondrial ribosomes contain several novel proteins. Amino acid sequence information for all of the bovine MRPs will prove invaluable for assigning functions to their genes, which would otherwise remain unknown.

Mitochondrial ribosomal proteins (MRPs¹) are much more diverse in number, sequences, and features among species than their cytoplasmic eukaryotic or bacterial counterparts. Mammalian mitochondrial (mt) ribosomes contain many more proteins than eubacterial ribosomes, and they contain at least as many proteins as compared to cytoplasmic eukaryotic ribosomes (1-3). In yeast only a minority of MRPs known so far shows significant sequence similarities to r-proteins from other sources. Additionally, those MRPs that are similar, e.g., to $E.\ coli$ r-proteins, show strong differences in length, with insertions, and only a relatively low degree of sequence similarity (3). Thus, identification of similar MRPs from higher eukaryotes, except for the laborious classical way of gene screening, cloning, and

sequencing, suffers from the lack of reliable MRP reference sequences.

Until now, few MRPs from mammals have been characterized on the molecular level. The MRL3 gene of human and rat was identified as an overexpressed protein in Mahlavu hepatomic cells and categorized as a 60S r-protein by its sequence similarity to the E. coli EcoL3 r-protein (4). Later it was postulated to be a true MRP since its sequence is more similar to its MRP counterpart YmL9 of yeast than to the corresponding cytoplasmic yeast r-protein (5). The MRP-L12 genes of human and hamster were identified as delayed early response genes accumulating during the G₁ phase of growth-stimulated cells (6). The MRP-L12 protein is the mammalian counterpart of the E. coli EcoL7/L12 r-protein (6). The human and mouse MRP-S12 genes were identified by homology to the *Drosophila* MRP-S12 gene, encoding the mitochondrial ribosomal protein homologue of eubacterial S12 (7, 8). Recently, several other mammalian MRP sequences were identified by using N-terminal peptide information of rat MRPs as screening probes for EST and other data banks (9). In a few cases the correspondence of the identified mammalian MRPs to yeast MRPs was shown. But the sequence similarities are low, and only in one case was a correspondence shown for a mammalian and a yeast MRP, both of which are not similar to E. coli r-proteins (9).

On the other hand, mammalian MRPs are very similar to each other, as inferred from their immunologic (10) and electrophoretic properties (11) and actual amino acid sequences, when available (9). Screening of databases using N-terminal peptide information obtained from one mammal

[†] This work was supported by USPHS Grant No. GM15438 to TWO'B, by the student exchange program of the University of Tübingen, Germany, and by funds of the Schering AG, Berlin, to HRG.

 $^{^{\}ddagger}$ The N-terminal peptide sequences reported in this paper have been submitted to the MIPS Data Base with accession nos. S78750 (MRP-L2_{bov}), S78751 (MRP-L3_{bov}), S78752 (MRP-L15_{bov}), and S78753 (MRP-L31/L34_{bov}), respectively.

^{*} To whom correspondence should be addressed: Tel.: 1-(352)-392-6878; Fax: 1-(352)-392-2953; E-mail: tobrien@ufl.edu.

[§] Present address: Institute for Genetics, AG Kress, Free University of Berlin, Arnimallee 7, D-14195 Berlin, Germany.

¹ Abbreviations: MRP, mitochondrial ribosomal protein(s); Ac, acetate; AU, absorbance units; EcoLxx, *E. coli* ribosomal proteins of the large (L) ribosomal subunit. xx represents the appropriate number; EST, expressed sequence tags; kDa, kiloDalton; MM, molecular mass; mt, mitochondrial; pI, isoelectric point; RP-HPLC, reversed phase high pressure liquid chromatography; r-proteins, ribosomal proteins; SDS, sodium dodecyl sulfate; TFA, trifluoroacetic acid; 1D- and 2D-PAGE, one dimensional and two dimensional polyacrylamide gel electrophoresis, respectively.

seems to be sufficient to identify the corresponding proteins and genes of others.

We are presenting the bovine mitochondrial ribosome as a model system for mammalian mitochondrial ribosomes (12). The proteins from each subunit have been identified and characterized extensively with respect to their individuality and electrophoretic properties (1), their topographic disposition (13), RNA binding properties (14), evolutionary relationships (11), and reaction with affinity probes of ribosome functional domains (15, 16). In this paper we describe the identification of individual bovine MRPs purified by RP-HPLC as an alternative to the original 2D PAGE system (1) upon which the nomenclature for bovine MRPs is based. Also, bovine MRP peptide information is used to identify corresponding human genes which have been assigned to unknown function to date and EST sequences of corresponding proteins from human, mouse, and rat.

EXPERIMENTAL PROCEDURES

Materials. Acetonitrile and H₂O for RP-HPLC were HPLC grade by Fisher. TFA was sequence grade by Sequamat, Inc.. Urea and Tris were ultrapure grade. All other chemicals were reagent grade. SDS was recrystallized before use.

Isolation of Mitochondrial Ribosomal Proteins. Large subunits of bovine mitochondrial ribosomal proteins were isolated as described previously (1, 12). Ribosomal subunits of high purity are routinely obtained by first isolating 55S monoribosomes using lysis and sucrose gradient conditions that stabilize monoribosomes, followed by separation of the 28S- and 39S-derived subunits in a second sucrose gradient under ribosome dissociating conditions. Proteins were isolated by a modification of the acetic acid extraction procedure (17). Samples containing up to 1.2 mg of protein were stirred in 1 mL of buffer A (10 mM Tris/HCl pH 7.7, 100 mM Mg(Ac)₂) for 30 min at 4° C. Two milliliters of glacial acetic acid were added to a final concentration of 66% acetic acid. After stirring overnight at 4° C, RNA was precipitated by centrifugation at 150 000g in a Ti 65 rotor for 1 h at 4° C. The supernatant was dialyzed vs 6% acetic acid for 3.5 h, with acetic acid changes after 30, 60, and 90 min. The protein solution was used directly for HPLC injection.

HPLC Separation. A Hewlett-Packard Model 1090 A HPLC system was used with a 200 μ L injection loop and an HP 85 B personal computer with 010190-10301 B-2348 1983-3 software. Separations were performed on a Synchrom synchropak RP-C 4 column, 250×4.1 mm with $6.5 \mu m$ particle size and 300 Å pores, or subsequently on a Vydac Phenyl column, 250×4.6 mm with 6.5μ m particle size and 300 Å pores. The solvent system was as follows: solvent A: water/0.1% TFA; solvent B: acetonitrile/0.1% TFA. Fractions were collected with an LKB Redirac 2112 collector. Dialyzed protein samples (up to 7 mL) were applied to the column by multiple injections of 200 µL each at 2 min intervals with a flow rate of 0.25 mL/min and 100% solvent A. The gradients were as follows: 0% B for 10 min, from 0% to 17% B in 10 min, from 17% to 25% B in 20 min, from 25% to 28% B in 15 min, from 28% to 39% B in 140 min, from 39% to 44% in 75 min, from 44% to 60% B in 50 min, from 60% to 85% B in 40 min, and from 85% to 0% B in 10 min. The flow rates were as follows: for 195 min at 0.25 mL/min, for 75 min at 0.2 mL/min, and for 100

min at 0.25 mL/min. Samples were collected at 1 min intervals

Electrophoresis. 1D PAGE was performed according to Laemmli (18) with modifications. Samples were dried on a Savant Speed Vac concentrator and solubilized in 30 μ L of buffer F (62.5 mM Tris/HCl pH 6.8, 2% SDS, 5% β-mercaptoethanol, 10% glycerol) by boiling in a water bath for 5 min. The separator gel was 0.7% bisacrylamide and 12% or 15% acrylamide, respectively. The stacking gel was 4% acrylamide and 0.1% bisacrylamide. Electrophoresis was at room-temperature overnight at a current of 20 mA or for 5–7 h at 4° C with constant current of 40 mA. Gels were silver stained (19).

2D PAGE was performed as described (1). Electrophoresis of the first dimension gels was carried out at 0.25 mA/gel until the tracker dye entered the separation gel. The gels were removed from the glass tubes using buffer G (9 M urea, 60 mM potassium acetate, pH 4.3). Samples of individual HPLC fractions were analyzed for identification of the respective MRPs by 2D PAGE together with a low amount of total large subunit MRPs for background reference. Large subunit MRPs were isolated as described above. The proteins in 6% acetic acid were dialyzed vs 1D sample buffer H (9 M urea, 60 mM potassium acetate, 0.01% amino-ethanethiol, 100 mM β -mercaptoethanol, pH 6.7) overnight. A 10–20-fold excess (500 μ L at maximum) of proteins collected by HPLC was added to 0.3 A₂₆₀ of total large subunit MRPs (equivalent of $18.8 \,\mu g$ total mitochondrial ribosomal protein). Samples were dialyzed against 1D sample buffer H for several hours with at least two changes of buffer and used directly for the first dimension PAGE. Proteins were identified according to the standard 2D numbering system for bovine MRPs by (1).

Immunological Identification of Individual Proteins. Selected proteins were identified on dotblots or western blots (1D PAGE) of eluant fractions (20) using available monoclonal antibodies against bovine MRPs L11 and L32 and monospecific antibodies against MRPs L2, L19, and L31 (21).

N-Terminal Sequencing of Individual Proteins. Pure protein samples were applied to poly(vinylidene difluoride) (PVDF) membrane for amino acid sequence analysis using an Applied Biosystems Model 470 gas-phase protein sequenator. The resulting phenylthiohydantoin (PTH)-amino acids were analyzed with an online model 120A PTH-HPLC analyzer (Applied Biosystems) using the program 03RPTH.

Computional Analysis. Computional screening of EST and other data banks, establishing of consensus cDNA sequences, and analysis and comparison of deduced protein sequences was done as described (9). Consensus sequences were assembled from ESTs using the GCG programs "PileUp", "BestFit", and "Map" (25).

RESULTS

Purification and Characterization of Individual MRPs. The proteins of the 39S subunit of bovine mitochondrial ribosomes were separated by RP-HPLC on a synchropak RP-C 4 column. As shown previously for yeast MRPs (22, 23), a water/acetonitrile/TFA system has been used successfully for separation of ribosomal proteins. Multiple injections of dialyzed samples (up to 1180 µg proteins in 7 mL of 6% HAc) were performed to ensure retention of the proteins from

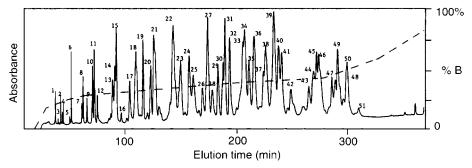


FIGURE 1: HPLC separation of 39S bovine mitochondrial ribosomal subunit proteins on a C 4 column: 1.18 mg of proteins in 5.6 mL of 6% acetic acid was injected in portions of 200 μ L each at intervals of 2 min. The gradient was as described in Experimental Procedures. The absorbance was measured at 214 nm with a total range of 0.5 AU. Elution time is given in minutes (min); the gradient extends from 0% to 100% solvent B.

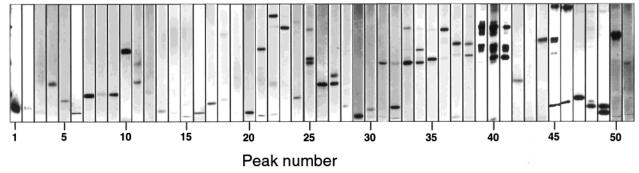


FIGURE 2: SDS PAGE of single fractions of individual peaks of bovine MRPs separated by HPLC (Figure 1). One tenth to 3/10 of the total amount of individual fractions was loaded. Fractions were collected at 1 min intervals during the HPLC separation. Numbers correspond to individual peaks of the HPLC profile (Figure 1). Protein bands correspond to the description of individual HPLC fractions in Table 1.

6% HAc. Bovine MRPs tend to precipitate upon standing in 6% or lower concentrations of HAc. After precipitation or speed vac concentration (to incipient dryness) they are not completely soluble in 66% HAc or 0.1% TFA. By collecting the efflux during multiple injections we could determine that no proteins detectable by gel electrophoresis and silver staining were washed off the column during as many as 35 injections of 200 µL each. The elution profile was not affected by multiple injections as compared to single injections of lower amounts of proteins.

The 52 proteins of the large mt ribosomal subunit were separated into 51 peaks (Figure 1), which are different by their protein composition (Figure 2). Twenty-two peaks or parts of peaks in single fractions were shown to contain only a single protein (Table 1). Other peaks contain up to six proteins of which one or two appear as major components. Proteins were identified by Western blotting of HPLC fractionated proteins (Figure 3) and/or by 2D PAGE (Figure 4). Molecular masses (MM) were determined by SDS PAGE (Table 2). A few peaks contain proteins of the neighboring peaks. Peak 30, for example, contains the proteins of peak 29 and peak 31, MRP-L19 and MRP-L32, respectively (Figure 1, Table 1). Since an overlap of two incompletely separated proteins would not form an additional peak, an aggregation form of these two proteins can be anticipated, which is not resolved completely by RP-HPLC.

Five of the purified proteins were selected for N-terminal amino acid sequencing (Table 2). MRP-L31bov and MRP-L34_{bov} turned out to be peptides of the same protein, accounting for the cross reaction of anti-L31 antibodies on L34 (11). They differ on their respective N-termini of five amino acid residues in length (Table 2), probably as a result of proteolysis during the isolation of the ribosomes (R. A.

Hessler and T. W. O'Brien, unpublished). The MRP sequences obtained were used to screen the EST and other data banks for similar peptide sequences from other sources. Several sequences of human, mouse, and rat origin were obtained. If possible, consensus cDNA sequences were assembled from EST sequences (Table 3), and deduced ORFs were characterized by comparison to the bovine peptide sequences and to each other (Figure 5). The amino acid sequencing of the mature bovine MRPs enabled us to postulate signal peptides for mt import for the deduced protein sequences of mammalian MRPs. In one case (L15) correspondence to a previously characterized ORF of unknown function was detected (Figure 5, Table 4).

MRP-L2. This 46 kDa bovine MRP is one of the largely "buried" proteins in the large subunit, based on its accessibility to radioiodination in intact subunits (13, 24). It apparently resides somewhere along the interfacial aspect of the subribosomal particle, in the vicinity of the peptidyl transferase center (15) since it is virtually inaccessible to radioiodination in the monoribosome. Several EST sequences were identified for the corresponding human and mouse MRPs. By "virtual chromosome walking" along the EST sequences complete cDNAs for MRP-L2_{human} and MRP-L2_{mouse} were assembled (Table 3). A C-terminal truncated MRP-L2_{rat} was deduced from a single EST of rat origin. The human and the mouse MRP-L2s are very similar to each other (Figure 5a) They share 79% identical amino acid residues. However, no similar r-protein from E. coli or yeast could be detected by screening the data banks. Corresponding to the mature MRP-L2_{bov} we postulate signal peptides for mt import of 27 amino acid residues for the human and mouse proteins and 29 amino acids for the rat MRP. The deduced mature MRP-L2_{human} has a calculated molecular

Table 1: Separation of 39S Bovine MRPs by HPLC ^a							
no. of peak	no. of proteins	proteins identified	main protein				
1	1						
2 3	1						
3 4	1 1						
5	1	L45	L45				
6	2	L51	L51				
7	1	L40	L40				
8	1						
9	1						
10	1						
11	2						
12	1						
13	1	I 46	T 46				
14 15	2 1	L46	L46				
16	1	L42	L42				
17	1	L42	L42				
18	2						
19	2	L41 L49	L41				
20	1	L26	L26				
21	3	L18	L18				
22	4	L6 L10	L6				
23	1	L10	L10				
24	2	T 20					
25	3	L20					
26 27	2 2	L24 L23					
28	1	L23					
29	1	L32	L32				
30	2	L32 L19	L32				
31	1	L19	L19				
32	2	L26 L19	L26				
33	4	L7 L16 L3 L15	L16				
34	4	L7 L16 L3 L15	L7				
35	3	L3 L15	L15				
36	4	L3 L15	L3				
37 38	3 3	L13L15 L13L15	L13				
36 39	4	L13L13 L2 L5	L2				
40	4	L2 L5 L11	L11				
41	4	L5 L11	L5				
42	1						
43	2	L1 L38	L38				
44	2	L1 L9	L9				
45	6	L1					
46	5	L1 L17 L28	L1				
47	1	L25	L25				
48	3	L31 L34					
49 50	2 1	L31 L34 L8	L8				
51	1	LU	L 0				

^a Protein distribution among different elution peaks. MRPs identified by PAGE and/or immunoblotting are listed. Main protein, major protein fraction of an individual peak. Protein identities follow the nomenclature system for bovine MRPs (1).

mass of 45 307 Da and a pI of 8.26. The calculated MM is in good agreement with the MM of MRP-L2_{bov} of 46 kDa as determined by SDS PAGE (Table 2). For MRP-L2_{mouse} very similar results (MM 45,602 Da; pI 8.35) were obtained. Both MRP-L2_{human} and MRP-L2_{mouse} are only slightly basic. r-Proteins are basic in general, although MRPs are not as basic on average as their cytoplasmic counterparts (*I*).

MRP-L3. The complete corresponding MRP sequences of both human and mouse were assembled from different ESTs (Table 3, Figure 5*b*). Additionally, an N-terminal truncated rat amino acid sequence was deduced from a single EST (Table 3, Figure 5*b*). The deduced MRP-L3_{human} sequence contains 379 amino acid residues. Within the assembled

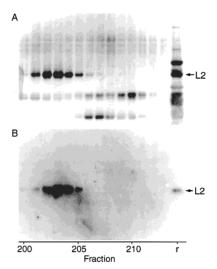


FIGURE 3: Identification of MRP-L2bov by immunoblotting. Proteins in peak 39 of Figure 1 were separated by chromatography on a Vydac phenyl column as described in Experimental Procedures. (A) SDS gel of individual fractions obtained from HPLC separation. (B) Autoradiography of immunoblot of the gel from (A) with anti-MRP-L2_{bov} antibody. Lane "r", ribosomal proteins in peak 39.

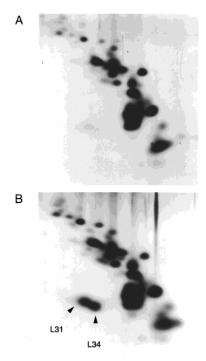


FIGURE 4: Identification by 2D PAGE of MRPs separated by HPLC. (A) 2D PAGE of proteins from $0.3 A_{260}$ of 39 S subunits of bovine mitochondrial ribosomes. First dimension: left to right. Second dimension: top to bottom. (B) 2D PAGE of proteins from $0.3 A_{260}$ of 39S subunits, as in Figure 3A, together with \sim 170 pmol of two different proteins separated by HPLC (peak 49 of Figure 1; see Figure 2 also).

consensus cDNA a poly-adenylation signal AATAAA was identified between bases 1329 and 1334. A poly(A) track begins at position 1351. Thus, we anticipate this to be the true poly-adenylation site of the mature MRP-L3_{human} mRNA. Corresponding to the mature N-terminus of the MRP-L3_{bov} protein we postulate an mt import sequence of 25 amino acid residues for the MRP-L3_{human}. The deduced mature MRP-L3_{human} has a MM of 43 540 Da, which is in good agreement with the MM of the mature MRP-L3_{bov} deter-

Table 2: N-terminal Amino Acid Sequences of Mature MRPs of Bos ${\it Taurus}^a$

protein na	me			MM	(kD)	
	10	20	30	40		
	1			i		
MRP-L2 _{bov}	AYEWGVRSTRKPEPPXLDXVYE1P				46	
MRP-L3 _{bov}	RRAAPLGPMPNEDIDVSNLERLKKY				44	
MRP-L15 _{bov}	ITGPSEPGVFQPPPKPVIVDKRGPQXRE				28	
MRP-L31 _{bov}	GEALTGAPLDNAPKEYPPKIQQLVQ				17.3	
MRP-L34 _{bov}	MRP-L34 _{bov} GAPLDNAPKEYPPKIQQLVQDIALSTLL				16.5	

^a N-terminal peptide sequences of purified bovine MRPs determined by amino acid sequencing, and molecular masses of bovine MRPs determined by second dimension SDS/urea gel electrophoresis. Amino acids are given in the single-letter code. Unidentified amino acids are given as *X*. Numbers give sequence extensions of determined peptide sequences. MM, molecular mass in kiloDaltons.

mined by SDS PAGE (Table 2). Both MMs are anticipated to be quite similar due to the strong sequence similarities between the corresponding human (and mouse) and bovine MRPs. The strong sequence conservation between MRP-L3_{human} and MRP-L3_{mouse} is shown in Figure 5b. Both proteins share 83.9% similar and 82.3% identical amino acids over their entire length of 379 amino acid residues according to the computer program "bestfit" (25). The calculated MM of the deduced mature MRP-L3_{mouse} of 43 610 Da and the calculated pI of 8.17 correspond to the respective values for MRP-L3_{human}. Within the MRP-L3_{mouse} consensus cDNA sequence a polyadenylation signal, AATAAA, was identified approximately 150 bp 3' of the translational stop codon. However, in contrast to the human ESTs many mouse ESTs derived from incompletely spliced mRNAs were identified, suggesting the existence of several introns within the MRP-L3_{mouse} genomic DNA sequence.

No r-proteins from other sources showing significant sequence similarities to the mammlian MRP-L3s were identified. However, a sequence similarity of about 40% (sequence identity of 25%) was detected between the C-terminus of MRP-L3_{human} and the Neuropolypeptide H3 (accession no. P30086) over its entire length of 187 amino acids. A stretch of 30 amino acid residues of MRP-L3_{human} shows a sequence similarity of 70% (sequence identity of 50%) to the yeast CDC25 protein. Whether these findings are functionally relevant remains to be eludicated.

MRP-L15. Several human EST sequences were identified corresponding to MRP-L15_{boy}. These sequences correspond to a previously characterized human protein termed RLX1 (accession no. P49406; (26)) which was assumed to be a putative 60S ribosomal protein by its sequence similarity to r-proteins of the EcoL19 family. However, comparison of the RLX1 to the MRP-L15bov N-terminal peptide shows a very good sequence correspondence (Figure 5c). On the other hand, the sequence similarity between the RLX1 and EcoL19 is relatively low, accounting for only 29% identity and 40% similarity, respectively, over a stretch of 95 amino acid residues (Table 4, Figure 5c). Since the RLX1 differs from the homogeneous group of the EcoL19 r-protein family so strongly, and since it does fit the MRP-L15_{bov} sequence so well, we postulate RLX1 to be the human MRP corresponding to the MRP-L15_{bov}. Since the N-terminal peptide of MRP-L15_{bov} was determined from the mature, processed

Table 3: Assembly of Identified EST Sequences: Consensus CDNA in 5' to 3' Direction^a

gene	bps nos. assembled into cDNA	EST acc. no.	total length	ORF from bps no. to no.	r-protein family
MRP-L2 _{human}	1-60	AA171584	1478	32-1294	new
	62-95	AA171584			
	90-113	R61625			
	113-180	AA179644			
	97 - 337	AA402114			
	12-539	AA486034			
	108 - 449	W73212			
	261-441r,c	AA838035			
MRP-L2 _{mouse}	1-268	W83828	1470	41 - 1306	new
	54 - 164	AA033456			
	368-505	AA103642			
	88 - 459	AA238400			
	57-560	AA170589			
	423-499	AI834889r,c			
MRP-L2 _{rat}	1-542	AA848353	542	40 - 540	new
MRP-L3 _{human}	1-112	AA907110	1360	6 - 1145	new
	107 - 371	H46322			
	56-131	AA044430			
	8-162	AA454762			
	70-504	AI187067r,c			
	146-462	AA847560r,c			
MRP-L3 _{mouse}	1-78	AI019651	1386	68-1207	new
	28-41	AA921101			
	51-207	AI196112			
	242-261	AI019651			
	213-289	AI156695			
	338-385	AI019651			
	2-20	W596031			
	105-293 209-325	AA518833 W596031			
	1-385 314-348	W57031 W76828			
	324-485	AA014376			
	262-336	W76871			
MRP-L3 _{rat}	1-421	AA891689r,c	421	*3-272	new
MRP-L12 _{mouse}	1-170	AA163676	1105	165-767	L7/L12
VIICI -L12mouse	14-25	AI931561	1105	103 707	LI/LI2
	32-459	AA240863			
	182-478	AA285498			
	350-493	W49060			
	161-214	AI837172r,c			
MRP-L12 _{rat}	1-542	AI176961r,c	542	*3-269	L7/L12
MRP-L15 _{mouse}	1-47	AA673916	1219	20-895	L19
2.20 mouse	26-110	W11186	1217	20 0,0	217
	130-553	AA673916			
	230-410	AA530552			
	151-441	AA624404			
	398-455	C78144			
	420-468	AA895618			
	506-589	C78144			

^a 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. r-protein family, affiliation of deduced ORFs to existing families of similar r-proteins. Base pair numbers (bps nos.) are the bases in individual ESTs that were assembled into the consensus sequence.

MRP, we postulate a signal peptide of 32 amino acid residues for RLX1. The mature RLX1 has a calculated molecular mass of 29 011 Da, which is in good agreement with the molecular mass of 28 kDa of the corresponding mature MRP-L15_{bov} determined by SDS PAGE (Table 2). The mature RLX1 has a calculated pI of 10.06.

The amino acid sequence of MRP-L15 $_{\rm mouse}$ was deduced from a consensus cDNA after EST assembly of 1219 bp (Table 3). The MRP-L15 $_{\rm mouse}$ protein sequence corresponds very well to the respective RLX1 and the bovine N-terminal

MRP-L2

а AYEWGVRSTRKPEPPxLDxVYEIP bovine N-terminus 11 11111 100 rat 98 mouse 98 human rat 198 mouse 198 human RTSAQNCTLATTWNRESLLLQVRGTSSTILSAKDPLPVIASREEVEATRSHVLETFYPISPTIDLQECHVYEVKDDTGFQEGYPYPHPHTLYFLEKANLR 298 mouse 298 human 398 mouse human mouse VGPVDFQPETFRKFLALYLHGVV* 421 |||| |.||||||||||||||||||||||||||||VGPVGFKPETFRKFLALYLHGAA* 421 human MRP-L3 b RRAAPLGPMPNEDIDVSNLERLKKY bovine N-terminus 111111111111111111 MAAPGGCGVFRNWEMPGLHTSAFLSRRTPPLGPMPNEDIDVSNLERLEKYRSFERYRRRAEQEARARTGGGPTGSIS*RRQIPKTKLTLGYPHPESVGPK 100 mouse 100 human ${\tt SYWSGTTFCGNFVPXWKKNGAARLRTASIPLEAVRAEWERTCGPYHKQRLAEYYGLYRDLFHGATFVPWVPLHVAYAVGEEDLIPVYHGNEVTPTEASRA}$ 200 mouse 200 human -QLAQRTFHTV 10 rat. PEVTYEADKDSLWTLLFINLDGHLLEPDAEYVHWLLTNIPSNRVAEGQETCPYLPPFPARGSGFHRFAFLLFKQDKPINFSEDTRPSPCYQLAQRTCRTF mouse 1111111:. 111111 300 human DLYKKHQEAMTPAGLAFFQCRWDDSVTHTFHQLLDMREPVFEFVRPPPYHPKQKRFPHQQPLRYLDRYRDSHEPTYGIY* rat 380 mouse 380 human MRP-L15 C ITGPSEPGVFQPPPKPVIVDKRGPQxRE bovine N-terminus 100 mouse 88 human RLX1 SNIIKOLE EcoL19 DRRKVLPIPEFYVGSILRVTTADPYASGK.TSQFLGICIKRSGNGLVATFTLRNTIEGQGVEICFELYNPRIQEIQVVKLEKRLDDNLLYLRDALPEYST mouse 187 human 1:11 111 QEQMKQDVPSFRPGDTVEVKVWVVEGSKKRLQAFEGVVÍAIRNŘGLHSAFTVŘKISNGEGVÉRVFQTHSPVVDSÍSVKRRGAVRKAKLYÝLŘERTGKAAR EcoL19 108 -RPNFNIKGIRFDLALTEEQMKEAQKWSKPWLEFDMMREYDTSKIEAALWEEIEASKKS* 58 rat 292 mouse 280 human IKERLN* EcoL19

đ

MRP-L31/34

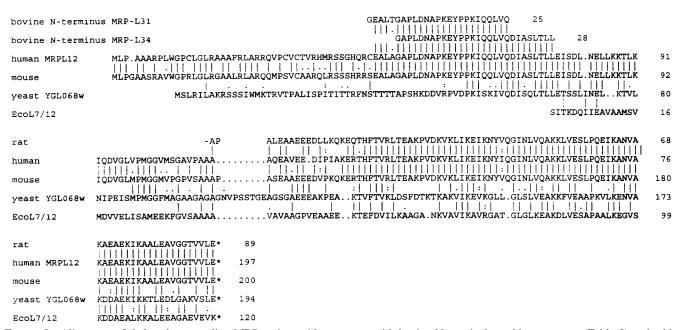


FIGURE 5: Alignment of deduced mammalian MRP amino acid sequences with bovine N-terminal peptide sequences (Table 2) and with corresponding protein sequences deduced from cDNAs of other species. Numbers give the respective amino acid position. Lines (|) mark identical amino acid positions, colons (:) mark strongly conserved amino acid residues, and points (.) mark weakly conserved amino acid residues, as calculated by the GCG program "bestfit" (25). Dashes (-) show N- or C-terminal ends of incomplete amino acid sequences, and asterisks (*) mark stop codons. Amino acid residues in lowercase letters are uncertain in their determination by amino acid sequencing (see Table 2). x marks unidentified amino acids although the positional number is valid. a, alignment of mammalian MRP-L2 protein sequences. b, alignment of mammalian MRP-L3 protein sequences. c, alignment of mammalian MRP-L15 protein sequences with the human RLX1 sequence (accession no. P49406) and the E. coli EcoL19 sequence (accession no. sw:rl19_ecoli). d, alignment of the mammalian MRP-L31/34 protein sequences with the human MRPL12 sequence (accession no. ×79864), the yeast YGL068w sequence (accession no. Z72591), and the E. coli EcoL7/L12 sequence (accession no. sw:rl7_ecoli).

Table 4: Families of Corresponding r-Proteins of Mammalian and Yeast Mitochondria and of E. Coli bovine human other E. coli r-proteins mouse yeast MRP-L2_{bov} MRP-L2_{human} MRP-L2_{mouse} none MRP-L2_{rat} none MRP-L3_{bov} $MRP-L3_{human}$ MRP-L3_{mouse} $MRP-L3_{rat}$ none none MRP-L15_{box} MRP-L15_{mouse} MRP-L15_{rat} EcoL19 RLX1 none MRP17_{hamster} EcoL7/L12 MRP-L31/34box MRPL12 MRPL12_{mouse} YGL068w MRP-L12_{rat}

peptide sequence except in the postulated signal sequence for mt import of about 44 amino acid residues, which is not well conserved between mouse and human (Figure 5c). The deduced mature MRP-L15_{mouse} has a calculated MM of 28 921 Da and a pI of 10.01; both values are very similar to the corresponding values of RLX1. Additionally, an Nterminal truncated MRP-L15_{rat} peptide was deduced from a single EST identified by virtual screening (Table 3, Figure 5c) which is extremely well conserved compared to the MRP- $L15_{\text{mouse}}$ and RLX1 sequence (Figure 5c). Further, additional ESTs showing strong sequence similarities of their deduced amino acid sequences of various lengths as compared to the RLX1 sequence were identified from Drosophila melanogaster (accession no. AA990636; 121 amino acids, 70% similarity, 49% identity), Schistosoma mansoni (accession no. T14351; 108 amino acids, 66% similarity, 50% identity), Zea mays (accession no. AA979971; 103 amino acids, 54% similarity, 37% identity), and rice (accession no. C74047; 63 amino acids; 61% similarity, 38% identity).

MRP-L31/34. This protein was identified to be the bovine counterpart of the previously characterized MRPL12 of human and MRP17 of the black-bellied hamster Cricetus cricetus (6). Both the human and the hamster MRPs are similar to the EcoL7/L12 r-protein ((6), Figure 5d). By comparison to the MRP-L31/L34_{bov} signal peptides of 44 and 46 amino acid residues for MRPL12_{human} and MRP17_{hamster} are postulated, respectively. However, both proteins have not been purified and sequenced yet. No data are available to indicate whether these proteins are modified similar to the EcoL7/L12 r-protein. The calculated MM of MRPL12 lacking a signal peptide of 44 amino acid residues is 16 497 Da. The mature MRPL12 is an acidic MRP—like its *E. coli* counterpart (27)—with a pI of 5.17. The MRPL12_{mouse} (Table 3, Figure 5d) is similar in pI (5.20) and MM (16 740 Da). However, in the fourth amino acid position of its ORF there is a frame shift which could not be resolved by EST sequence comparison. In this position of the assembled consensus cDNA a putative splice site is located as suggested by other EST sequences and splice site sequence conservation (data not shown). We therefore assembled the MRPL12_{mouse} sequence in correspondence to MRPL12_{human}. Molecular masses of MRPL12_{human} and MRPL12_{mouse} correspond well to the respective values of MRP-L31bov and MRP-L34bov of 17.3 and 16.5 kDa, respectively. The 2D map position of

MRP-L31_{bov} and MRP-L34_{bov} is consistent with their putative acidic character (Figure 3). Additionally, an N-terminal truncated ORF for MRP-L12_{rat} was deduced from a single EST identified by virtual screening (Table 3, Figure 5*d*). The deduced amino acid sequence shows strong similarities as compared to the MRPL12_{human} and the MRP-L12_{mouse} (Figure 5*d*). Further, EST deduced amino acid sequences similar to these MRPs were identified for *D. melanogaster* (accession no. AA802042; 58 amino acids, 59% similarity, 39% identity, and a second stretch of 36 amino acids; 69% similarity, 47% identity), *Caenorhabdidis elegans* (accession no. C50132; 105 amino acids, 49% similarity, 33% identity), and *Dictiostelium discoideum* (accession no. C90673; 109 amino acids, 44% similarity, 26% identity).

All these r-proteins mentioned show significant sequence similarity to the yeast ORF YGL068w (accession no. S64075). This ORF was postulated to be the yeast MRP corresponding to the EcoL7/L12 r-protein (28). Interestingly, the region of the respective signal peptides is much less conserved than sequences downstream (Figure 5*d*). Overall, YGL068w shows a sequence identity of 37% compared to the mammalian MRPs (Figure 5*d*).

Together, four different classes of mammalian MRPs have been characterized. Two of them (MRP-L2, MRP-L3) represent new classes of MRPs that have no counterparts in the yeast mitochondrial nor the E. coli ribosome. The two others are the mammalian mitochondrial counterparts of the bacterial L19 and the L7/L12 r-proteins, respectively. However, the latter MRPs are much less conserved compared to their bacterial and yeast counterparts than is common for bacterial and eukaryotic cytoplasmic r-proteins from the same family. These findings of MRP properties, first adduced from comparisons of bovine and rat r-proteins (11), are consistent with recent results regarding other mammalian MRPs (9). The mature MRP-L2_{human} and MRP-L2_{mouse} differ at 80 of their 394 amino acids, for example. This corresponds to a mutation rate of 40.6×10^{-10} substitutions per site per year, assuming a divergence time of 50 million years, surprisingly close to the value of 40×10^{-10} , predicted for mammalian MRPs as a group, based on their electrophoretic properties (11).

DISCUSSION

Four different groups of mammalian MRPs have been characterized in the present work. Although the mammalian MRPs are quite similar to each other, two of them (MRP-L2, MRP-L3) do not show any sequence similarities to r-proteins from other sources. For MRP-L2_{bov} this result was unexpected. It was anticipated that MRP-L2_{bov} would be similar (coincidentally) to E. coli L2, since both proteins are targeted by chloramphenicol affinity probes, iodoamphenicol (15) or bromoamphenicol (29). It thus appears that MRP-L2 is one of the "extra" proteins in mammalian mitochondrial ribosomes (24) that have no counterparts in bacterial or eukaryotic cytoplasmic ribosomes. MRP-L3_{boy} is also thought to reside within the peptidyl transferase domain, for it occupies a topology similar to that of MRP-L2_{bov} (13) and it is also a target of chloramphenical affinity probes (T. O. Harville and T. W. O'Brien, unpublished). Since both proteins do not have cytoplasmic counterparts, their functions are unique to mitochondrial ribosomes.

This is not the case for the mammalian MRP-L15s and MRP-L31/34s. MRP-L15 belongs to the L19 family of r-proteins (Table 4). Interestingly, by correspondence to the bovine MRP-L15 amino acid sequence determined by N-terminal amino acid sequencing, a previously characterized gene RLX1 (26) of unknown function was assigned to be a true human MRP. The RLX1 gene was surmised to belong to the L19 class of r-proteins, but no hint for an mt destination could be deduced from the pure sequence (26). Further, the future identification of more mammalian genes of unknown function as MRPs can be anticipated. In yeast, more than 50% of the MRPs are not similar to any known r-proteins of other sources (3). Identification of their genes has been achieved by analysis of mutants or N-terminal sequencing of purified individual proteins, the same method that has been applied for bovine MRPs in this paper.

The corresponding human gene for MRP-L31/34_{bov} has been identified (6). The L7/L12 family of r-proteins consists of members showing an unusually high degree of sequence similarities between each other. However, their occurrence in a 4-fold excess within the ribosome compared to the other r-proteins, their location within the large protuberance of the large ribosomal subunit, and the presence of a modified (acetylated) and an unmodified form at once in the ribosome makes this protein unique compared to other r-proteins (27). However, the two modified forms of the corresponding MRP-L31/34 detected in the mt ribosome show a different form of modification. First, the MRP-L31/34 proteins are not modified at their respective N-termini by acetylation, otherwise they would not be subjectable to N-terminal aino acid sequencing. Second, the modification they show is a specific proteolytic loss of five N-terminal amino acids from MRP-L31 to yield MRP-L34. However, MRP modification by proteolysis is known at least as far as it is detectable by the used 2D PAGE systems. In yeast mt ribosomes, four different pairs of full length and truncated MRPs are known from the large ribosomal subunit (3). Yeast MRP YmL11, although present in a single spot in the 2D PAGE system, yields two N-terminal amino acid sequences in parallel, differing for a few N-terminal amino acid residues (30). In rat mt ribosomes two proteins are present in the 2D PAGE (MRP-L22_{rat} and MRP-L24_{rat}) that yield identical N-terminal amino acid sequences by N-terminal protein sequencing (9). Thus, specific proteolytic events either during the respective purification procedure and/or during the assembly of the mt ribosomes seem to be common. However, in the case of MRP-L31/L34 we suspect the proteolytic cleavage of MRP-L31 to be a preparation artifact which can be suppressed using protease inhibitors such as NEM, TPCK, and PMSF for the purification (R. A. Hessler and T. W. O'Brien, unpublished). The 5 amino acid residues that are cleaved from MRP-L31 to yield MRP-L34 are "extra ones" relative to E. coli L7/L12, and they may be more exposed and susceptible to proteolysis during the exposure to proteases in the mitochondrial lysate while the ribosomes are being recovered by centrifugation (see Experimental Procedures). So we propose MRP-L31bov to be the true counterpart of the E. coli L7/L12 protein.

N-Terminal Signal Sequences for mt Import of Mammalian MRPs. For all of the deduced MRP sequences cleavable signal sequences for mt import were postulated. According to the relative positions of arginine residues as compared to

the postulated cleavage sites, the signal sequences presented belong to different classes (31). The MRP-L2_{rat}, MRP-L2_{mouse}, and MRP-L12_{mouse} possess import sequences of class R-2. MRP-L2_{mouse}, MRP-L2_{human}, and the mammalian MRP-L15s may be cleaved according to the R-3 mechanism (31). The MRP-L15s and MRP-L12_{mouse}, alternatively, may be imported by a two-step mechanism involving the R-10 cleavage site. MRPL12_{human} correlates to the latter class of import sequences only. The MRP-L3s belong to the R-none class (31). Such a differentiation of the classification of their putative import signal sequences has been reported for other mammalian MRPs, too (9). Interestingly, mammalian MRPs have adopted N-terminal signal sequences in cases where the corresponding yeast MRPs possess no cleavable signal $sequences \ (MRP\text{-}L28_{mammals} \ vs \ yeast \ YmL33; \ MRP\text{-}L32_{mammals}$ vs yeast YmL38 (9)). This might suggest that the shift of the respective genes from the mt DNA to the nucleus happened after the ancestors of mammals and yeast diverged during evolution.

Finally, we want to say a few words about the emerging nomenclature of mammalian MRPs. Of course, proteins and genes which have been characterized regardless of their MRP function have obtained independent names, e.g., RLX1 (26). Names of new sequences presented in this paper correspond to the established 2D PAGE system of bovine MRPs (1). Similar 2D PAGE maps of human and other mammalian MRPs have been determined in correspondence to the bovine system (T. W. O'Brien, in preparation). Thus, we have termed corresponding human, etc. MRPs according to the bovine 2D PAGE map; for example, MRP-L2bov, MRP-L2_{human}, MRP-L2_{mouse}, MRP-L2_{rat}. This nomenclature has been used also for characterization of MRPs corresponding to rat MRPs and their individual 2D PAGE system (9). Of course MRP-L31_{human} named according to the rat MRP-L31 is not identical with MRP-L31/34_{human}, corresponding to the bovine MRP-L31/34 (Table 2, Table 3). In this case, fortunately, confusion can be avoided since the human gene corresponding to MRP-L31/34bov is the MRPL12 already cloned (6). However, in the future there might arise confusion if two different human MRPs will be termed identically corresponding to the respective bovine and rat MRPs. As long as researchers working on mammlian MRPs want to name mitochondrial ribosomal proteins as "MRP-Lxx" or "MRP-Syy", respectively, they should keep in touch to avoid confusing double names. Additional proteins might be named using numbers which exceed those in the respective 2D PAGE nomenclatures. Naming of MRPs in correspondence to E. coli r-proteins is not sufficient since many MRPs are not similar in sequence to known r-protein families (3, 9, this work).

ACKNOWLEDGMENT

The authors would like to acknowledge the technical assistance of Mark Lijewski in the analysis of proteins by 2D PAGE and the assistance of Nancy Denslow in the identification of individual MRPs.

REFERENCES

Matthews, D. E., Hessler, R. A., Denslow, N. D., Edwards, J. S., and O Brien, T. W. (1982) *J. Biol. Chem.* 257, 8788

8794.

- Wool, I. G., Chan, Y.-L., and Glück, A. (1995) Biochem. Cell Biol. 73, 933-947.
- 3. Graack, H.-R., and Wittmann-Liebold, B. (1998) *Biochem. J.* 329, 433–448.
- Ou, J.-H., Yen, T. S. B., Wang, Y.-F., Kam, W. K., and Rutter, W. J. (1987) *Nucl. Acids Res.* 15, 8919

 –8934.
- Graack, H.-R., Grohmann, L., Kitakawa, M., Schäfer, K.-L., and Kruft, V. (1992) Eur. J. Biochem. 206, 373–380.
- Marty, L., and Fort, P. (1996) J. Biol. Chem. 271, 11468

 11476.
- 7. Shah, Z. H., O'Dell, K. M., Miller, S. C., An, X., and Jacobs, H. T. (1997) *Gene* 204, 55–62.
- 8. Johnson D. F., Hamon M., and Fischel-Ghodsian, N. (1998) *Genomics* 52, 363–368.
- 9. Goldschmidt-Reisin, S., Kitakawa, M., Herfurth, E., Wittmann-Liebold, B., Grohmann, L., and Graack, H.-R. (1998) *J. Biol. Chem.* 273, 34828–34836.
- Pietromonaco, S. (1986) The Evolution and Biosynthesis of the Proteins of Animal Mitochondrial Ribosomes, Ph.D. Thesis, University of Florida, Gainesville, FL.
- Pietromonaco, S., Hessler, R. A., and O'Brien, T. W. (1986)
 J. Mol. Evolution 24, 110-117.
- 12. O'Brien, T. W., and Denslow, N. D. (1996) *Methods Enzymol.* 264, 237–248.
- Denslow, N. D., and O'Brien, T. W. (1984) J. Biol. Chem. 259, 9867–9873.
- 14. Piatyszek, M., Denslow, N. D., and O'Brien, T. W. (1988) *Nucl. Acid Res.* 16, 2565–2583.
- O'Brien, T. W., Denslow, N. D., Harville, T. O., Hessler, R. A., and Matthews, D. E. (1980) in *The Organization and Expression of the Mitochondrial Genome* (Kroon, A., and Saccone, C., Eds.) pp 301–305, Elsevier/North-Holland Biomedical Press.
- Denslow, N. D., Anders, J. C., and O'Brien, T. W. (1991) J. Biol. Chem. 266, 9586–9590.
- 17. Sherton, C. C., and Wool, I. G. (1972) *J. Biol. Chem.* 247, 4460–4467.
- 18. Laemmli, U. K. (1970) Nature 227, 680-685.
- Wray, W., Boulikas, T., Wray, V., and Hancock, R. (1981)
 Anal. Biochem. 118, 197–203.
- Pietromonaco, S., Denslow, N. D., and O'Brien, T. W. (1991)
 Biochimie 73, 827–836.
- Chiles, T. C., O'Brien, T. W., and Kilberg, M. S. (1987) *Anal. Biochem.* 163, 136–142.
- 22. Graack, H.-R., Grohmann, L., and Choli, T. (1988) *FEBS Lett.* 242, 4–8.
- Kitakawa, M., Graack, H.-R., Grohmann, L., Goldschmidt-Reisin, S., Herfurth, E., Wittmann-Liebold, B., Nishimura, T., and Isono, K. (1997) Eur. J. Biochem. 245, 449–456.
- 24. O'Brien, T. W., Denslow, N. D., Faunce, W. H., Anders, J. C., Liu, J., and O'Brien, B. J. (1993) in *The Translational Apparatus: Structure, Function, Regulation, Evolution* (Nierhaus, K. H., Franceschi, F., Subramanian, A. R., Erdmann, V. A., and Wittman-Liebold, B., Eds.) pp 575–586, Plenum Publishing Corp., New York.
- GCG (1997) Wisconsin Package, Version 9.1, OpenVMS, University of Wisconsin, Madison, WI.
- 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.
- Brot, N., and Weissbach, H. (1981) Mol. Cell. Biochem. 36, 47–63.
- Mason, T. L., Pan, C., Sanchirico, M. E., and Sirum-Conolly, K. (1996) Experientia 52, 1148–1157.
- Sonenberg, N., Wilchek, M., and Zamir, A. (1973) Proc. Natl. Acad. Sci. U.S.A. 70, 1423-1426.
- Grohmann, L., Graack, H.-R., and Kitakawa, M. (1991) FEBS Lett. 284, 51–56.
- 31. Branda, S. S., and Isaya, G. (1995) *J. Biol. Chem.* 270, 27366–27373.

BI991543S