4936–4940 Nucleic Acids Research, 1993, Vol. 21, No. 21

Ribosomal protein L25 from *Trypanosoma brucei*: phylogeny and molecular co-evolution of an rRNA-binding protein and its rRNA binding site

Stan Metzenberg, Christine Joblet, Philippe Verspieren and Nina Agabian
Intercampus Program in Molecular Parasitology, University of California–San Francisco, San Francisco, CA 94143-1204, USA

Received July 14, 1993; Revised and Accepted September 17, 1993 GenBank accession no. L21172

ABSTRACT

The gene encoding ribosomal protein L25, a primary rRNA-binding protein, was isolated from the protozoan parasite *Trypanosoma brucei*. Hybridization studies indicate that multiple copies of the gene are present per *T.brucei* haploid genome. The C-terminal domain of L25 protein from *T.brucei* is strikingly similar to L23a protein from rat, L25 proteins from fungal species, and L23 proteins from eubacteria, archaebacteria, and chloroplasts. A phylogenetic analysis of L23/25 proteins and the putative binding sites on their respective LSU-rRNAs (large subunit rRNAs) provides a rare opportunity to study molecular co-evolution between an RNA molecule and the protein that binds to it.

INTRODUCTION

The ribosomal protein L25 is highly conserved, binds tightly to the LSU-rRNA (28S/23S large subunit rRNA), and may be situated near the peptidyl transferase site in the ribosome[1–3]. It is a 'primary rRNA-binding protein', in that it is one of the first to be assembled on the rRNA during ribosome biosynthesis, and its binding does not depend on other protein factors[4]. Identifying the homologs of prokaryotic ribosomal proteins in eukaryotes is often problematic, but L25 is one of the rare examples where the conservation of protein sequence has been sufficient to allow cross-identification of ribosomal protein homologs in a variety of organisms (in eubacteria and chloroplasts the L25 homologs are named L23; homologs of the protein will be denoted L23/25 in this paper)[5, 6]. On the LSU-rRNA, the L23/25 binding site has been mapped in eukaryotes and prokaryotes, and is also highly conserved[7, 8]. The different L23/25 homologs, and their respective rRNA binding sites, therefore present an excellent opportunity for the study of molecular co-evolution[9] between interacting macromolecules.

The gene encoding ribosomal protein L23/25, whose homolog in the kinetoplastid species *Trypanosoma brucei* is described in this manuscript, has been isolated and sequenced from at least 15 other eukaryotes, prokaryotes, or chloroplasts. In addition to segments that are well conserved in all L23/25 proteins, the eukaryotic homologs encode an N-terminal sequence directing localization to the nucleus[10, 11]. The structural domain required for binding of the L23/25 protein to the LSU-rRNA has been determined for the L25 protein of *Saccharomyces cerevisiae*, and consists of approximately 80 amino acids near the C-terminus[12]. Remarkably, the RNA-binding domains of L23/25 proteins are sufficiently conserved that the L25 protein from *Saccharomyces cerevisiae* will bind to the *Escherichia coli* 23S rRNA; conversely the *E.coli* L23 protein will bind to the *S.cerevisiae* 26S rRNA[8].

During an analysis of poly A^+ mRNA-binding proteins in *T.brucei*, a cDNA clone encoding ribosomal protein L25 was isolated fortuitously. The identification of the clone as a trypanosomal homolog of L25 was achieved by comparison of the predicted protein sequence with the NBRF-PIR database. Comparisons between the predicted sequence of the *T.brucei* L25 protein and its eukaryotic homologs identified a domain which was likely to function as a nuclear localization sequence at the N-terminus of the protein, and an rRNA-binding domain near the C-terminus. Additional phylogenetic comparison of the binding sites in the LSU-rRNA from many of these organisms is consistent with molecular co-evolution of the protein and its rRNA target.

MATERIALS AND METHODS

Bloodstream-stage cells of *Trypanosoma brucei* (subspecies brucei), strain 1.1, were propagated by serial passage in mice and rats, while procyclic-stage cells were propagated in BSM medium supplemented with 10% fetal bovine serum and penicillin (250 u/ml).

Antisera to RNA-binding proteins were prepared by passing a cytoplasmic extract of *T.brucei* (supernatant of 10,000×g centrifugation of trypanosomes lysed with NP-40) over a poly(A) sepharose column, removing loosely bound proteins by washing the column with buffer A (0.1 M NaCl, 0.03 M Tris-Cl pH 7.5, 0.015 M 2-mercaptoethanol and 0.001 M EDTA) supplemented with 10 μg/ml poly(C), and eluting the remaining proteins with buffer A supplemented with 7 M guanidine HCl. The eluted proteins were dialyzed against buffer A, concentrated and mixed with Freund's incomplete adjuvant, and injected as an immunogen into Balb/c mice. The mice were boosted with the same immunogen, mixed with Freund's complete adjuvant, at two intervals of 3 weeks, and then bled for antibodies. A cDNA

library, cloned in the vector λgt11 (Promega, Inc.), was screened by the method described in the Molecular Cloning Manual[13].

For the purposes of Southern blot analysis, DNA was prepared from bloodstream-stage cells[14], and dialyzed against TE (0.01 M Tris-Cl, 0.001 M EDTA, pH 7.5) at 4°C. Restriction endonucleases were purchased from New England Biolabs Inc. (Beverley, MA), and endonuclease digestions were performed under the recommended salt conditions for each enzyme. pBSII was obtained from Stratagene, Inc. (LaJolla, CA). Electrophoresis and Southern blot analyses were performed, as described in the Molecular Cloning Manual[13].

DNA sequence was obtained using an ABI-370 DNA sequencer apparatus, and reagents from Applied Biosystems Inc. (Foster City, CA). The nucleotide sequence of *T.brucei* L25 has been assigned the GenBank accession number: L21172. Phylogenetic analyses of RNA and protein sequences were performed using the Doolittle program[15−17] of EuGene (Molecular Biology Information Resource, Baylor College, Houston, TX). In this program, a multiple protein sequence alignment was constructed by the Needleman and Wunsch algorithm, which maximizes similarities. Starting with the pair of species with the highest similarity score, the program inserted neutral elements to fill gaps, preserving the gap structures during the progressive alignment of additional species (the rRNA sequences were pre-aligned by the Ribosomal Database Project, as described later). From the resulting percentage difference matrix, the branching order and branch lengths were calculated for tree construction. For the tree of protein sequences, the branch lengths were derived from −ln(S), where S is the amino acid similarity between a protein species and its hypothetical ancestral species. In the case of the nucleic acid sequence tree, the distances represent the number of simple base replacements separating the species from its hypothetical ancestor. To analyze coincident variation between protein and RNA sequence, we constructed a computer program that compares the columns of aligned sequences for any similarity in the patterns of variation between species. In the algorithm of this program, the pattern of variation was determined at each column of aligned residues (either RNA or amino acid), and patterns matching in both the protein domain and rRNA site were tabulated. For these analyses, sequence alignments of L23/25 protein were developed using the Doolittle program, while LSU-rRNA sequences and alignments were derived from the Ribosomal Database Project[18] on the anonymous ftp server at Argonne National Laboratory in Argonne, Illinois (Ver 2.1, May 1993). From an alignment with parameters of SEQUENCE_POSITIONS = *E.coli*: [1280− 1430,1570−1635], the following columns were included: [15−96, 187−239, 324−327, 396−408, 453−461, 482−492, 539−574]. These columns of the alignment were chosen to satisfy two criteria; they correspond to regions of the *S.cerevisiae* LSU-rRNA protected by L25 binding from T_1-nuclease attack, and at least 3 of the 11 species have overlapping sequence.

RESULTS AND DISCUSSION

A cDNA clone, λ_{613}, was isolated from a λ library by antibody screening, as described in the Materials and Methods. This clone consisted of (from 5′ to 3′): 16 nt. of the 3′ end of the *T.brucei* spliced leader[19], 28 nt. of 5′ untranslated sequence, an open reading frame of 492 nt. and a 3′ untranslated sequence of 160 nt. Comparison of the predicted protein sequence encoded by the λ_{613} clone with the NBRF-PIR database, indicates that the λ_{613} open reading frame encodes a homolog of the ribosomal protein denoted L23a or L25 in eukaryotes, and L23 in most prokaryotes. As shown in the sequence alignment in Figure 1, the similarity is highest between the *T.brucei* homolog and the rat L23a sequence[20] (46% identity over 148 amino acids) and lower between the *T.brucei* sequence and its L23 homolog in *E.coli*[21] (30% identity over 82 amino acids).

To determine the number of copies of the L25 gene in the *T.brucei* genome, agarose gel electrophoresis and Southern blot analyses were performed on bloodstream-stage trypanosome DNA, as shown in Figure 2. Digestion of *T.brucei* DNA with restriction endonucleases Bam HI, Eco RI, Kpn I or Hind III (lanes 3−6, respectively of Figure 2) resulted in a single band hybridizing to a plasmid clone of the λ_{613} insert. A reconstruction experiment with 20 pg and 2 pg (lanes 1−2 of Figure 2) of digested pBSII vector DNA indicates that each trypanosome genome equivalent contains approximately 2−4 copies of the L25 gene. Digestion of *T.brucei* DNA with Stu I results in three principal bands (lane 7, Figure 2). A 0.24 kbp fragment spanning two internal sites in the L25 gene was too small to appear on this gel. The remaining bands are therefore consistent with at least two copies of the L25 gene per genome. Additional mapping and sequence analysis of a genomic clone confirmed that two tandem copies of the L25 gene are present per haploid genome[22]. Although rehybridization of the blot with a *T.brucei* β-tubulin probe showed no evidence of partial digestion products (data not shown), the faint hybridizing band at 1.1 kbp may indicate incomplete digestion with Stu I, or alternatively a restriction site polymorphism in a fraction of the genomic copies. Pulse-field gel and Southern blot analysis of undigested *T.brucei* DNA indicates that the L25 gene copies are all contained either on a single chromosome with a size range of 2−5 Mbp, or on multiple chromosomes with indistinguishable electrophoretic mobilities (data not shown). Further study of the genomic structure of the L25 locus is currently in progress.

The minimal subset of L23/25 amino acid sequences responsible for rRNA binding has been localized in the *S.cerevisiae* homolog[10, 12]. This region is highly conserved in *T.brucei* and other eukaryotic L23/25 proteins, as shown by the alignment (positions 84−161) in Figure 1, and the phylogenetic comparison in Figure 3A. As shown in Figure 1, this binding domain is highly similar between rat and trypanosomes; when conservative amino acid substitutions are included, the sequences are 85% similar over 65 residues. The prokaryotic L23/25 proteins also share sequence features with their eukaryotic homologs, particularly at positions 114−118 (KTEIK motif), 132−135 (VNTL motif) and 155−161 (KKAYIRL motif) (see Figure 1). The archaebacterial and eukaryotic L23/25 proteins can be distinguished from eubacterial and chloroplastid homologs, however, by the absence of 13 amino acid residues (at positions 142−154 in Figure 1) in the C-terminal region of the protein[23]. This segment of the eubacterial and chloroplastid L23/25 homolog is variable, but usually contains numerous basic residues. The L23/25 proteins of Gram-positive bacteria and chloroplasts have an additional amino acid deletion (position 101 in Figure 1) in comparison with Gram-negative bacteria, archaebacteria and eukaryotes, while *S.cerevisiae* has a unique deletion at positions 128−132. All prokaryotic L23 proteins lack an amino-terminal extension, which in eukaryotes varies in length between 61 amino acids in yeast, and 83 amino acids in *T.brucei*. The N-terminus of the *S.cerevisiae* L25 protein (Figure 1, positions 23−63) promotes import into the nucleus

```
                  10         20         30         40         50         60
                   |          |          |          |          |          |
T.brucei  MPAKAASAAASKKNSAPKSAVSKKVAKKGAPAAAAKPTKVVKVTKRKAYTRPQFRRPHTY
                   : :              .  :.. :  :    :    :.:. : : :::: :
R.norvegi.--------MAPKAKKEAPAPPKAEAKAKALKAKKAVLKGVHSHKKKKIRTSPTFRRPKTL
S.cerevis.----------------------MAPSAKATAAKKAVVKGTNGKKALKVRTSATFRLPKTL
H.jadinii ----------------------MAPSTKAASAKKAVVKGSNGSKALKVRTSTTFRLPKTL
M.vanniel.------------------------------------------------------------
H.marismo.------------------------------------------------------------
E.coli    ------------------------------------------------------------
Y.pseudot.------------------------------------------------------------
B.stearot.------------------------------------------------------------
M.caprico.------------------------------------------------------------
CZ.mays   ------------------------------------------------------------
CO.sativ. ------------------------------------------------------------
CN.tabac. ------------------------------------------------------------
CM.polym. ------------------------------------------------------------

                  70         80         90         100        110        120
                   |          |          |  *       |    *     |          |
T.brucei  RRPATVKPSSNVSAIKNKWDAFRIIRYPLTTDKAMKKIEENNTLTFIVDSRANKTEIKKA
          :     :         .:: : . ::..::::. :::::::.:::: :::: .:::  :: :
R.norvegi.RLRRQPKYPRKSAPRRNKLDHYAIIKFPLTTESAMKKIEDNNTLVFIVDVKANKHQIKQA
S.cerevis.KLARAPKYASKAVPHYNRLDSYKVIEQPITSETAMKKVEDGNILVFQVSMKANKYQIKKA
H.jadinii KLTRAPKYARKAVPHYQRLDNYKVIVAPIASETAMKKVEDGNTLVFQVDIKANKHQIKQA
M.vanniel.------------------MDAFDVIKTPIVSEKTMKLIEEENRLVFYVERKATKEDIKEA
H.marismo.------------------MSWDVIKHPHVTEKAMNDMDFQNKLQFAVDDRASKGEVADA
E.coli    ------------MLAIREERLLKVLRAPHVSEKASTAMEKSNTIVLKVAKDATKAEIKAA
Y.pseudot.--------------MIREERLLKVLRSPHVSEKASAAMEKNNTIVLKVAKDATKAEIKAA
B.stearot.------------------MKDPRDIIKRPIITENTMNLIGQ-KKYTFEVDVKANKTEVKDA
M.caprico.------------------MHITEVLKKPVLTEKSFAGHKD-NVYTFLVDKKANKVQIKKT
CZ.mays   ----------------------MDGIKYAVFTEKSLRLLGK-NQYTFNVESGFTKTEIKHW
CO.sativ. ----------------------MDGIKYAVFTEKSLRLLGK-NQYTFNVESGFTKTEIKHW
CN.tabac. ----------------------MDGIKYAVFTDKSIRLLGK-NQYTSNVESGSTRTEIKHW
CM.polym. ----------------------MNQVKYPVLTEKTIRLLEK-NQYSFDVNIDSNKTQIKKW

                  130        140        150        160        170
                 *    |   *    |          |          |   *      |
T.brucei  IRKLYQVKTVKVNTLIRPDGL-------------KKAYIRLSASYMGLV
          ..::: .  .:::::::::::               :::::.::   : .:
R.norvegi.VKKLYDIDVAKVNTLIRPDGE-------------KKAYVRLAPDYDALDVANKIGII
S.cerevis.VKELYEV-----NILVRPNGT-------------KKAYVRLTADYDALDIANRIGYI
H.jadinii VKDLYEVDVLAVNTLIRPNGT-------------KKAYVRLTADHDALDIANKIGYI
M.vanniel.IKQLFNAEVAEVNTNITPKGQ-------------KKAYIKLKDEYNAGEVAASLGIY
H.marismo.VEEQYDVTVEQVNTQNTMDGE-------------KKAVVRLSEDDDAQEVASRIGVF
E.coli    VQKLFEVEVEVVNTLVVKGKVKRHGQRIGRRSDWKKAYVTLKEGQNLDFVGGAE
Y.pseudot.VQKLFEVEVEDVNTLLVKGKSKRHGQRVGRRSDWKKAYVTLKEGQNLDFIGGAE
B.stearot.VEKIFGVKVEKVNIMNYKGKFKRVGRYSGYTNRRKKAIVTLTPDSKEIELFEV
M.caprico.FEEIFEVKVESVRTINYDAKEKRLGKYVGKKPSYKKAIITLKEGQKLDVLSDL
CZ.mays   VELFFGVKVVAVNSHRLPGKGRRMGPILGHTMHYRRMIITLQPGYSIPLLDRETN
CO.sativ. VELFFGVKVVAVNSHRLPGKGRRMGPILGHTMHYRRMIITLQPGYSIPLLDREKN
CN.tabac. VELFFGVKVIAMNSHRLPGKSRRMGPIMGHTMHYRRMIITLQPGYSIPPLRKKRT
CM.polym. IELFFNVKVISVNSHRLPKKKKKIGTTTGYTVRYKRMIIKLQSGYSIPLFSNK
```

Figure 1. Amino acid sequences of L23/25 homologs. Species names are abbreviated at the left, and are respectively: *Trypanosoma brucei*, *Rattus norvegicus*[20], *Saccharomyces cerevisiae*[29], *Hansenula jadinii*[30], *Methanococcus vannielii*[31], *Halobacterium marismortui*[23], *Escherichia coli*[21], *Yersinia pseudotuberculosis*[32], *Bacillus stearothermophilus*[33], *Mycoplasma capricolum*[34]; and chloroplasts from *Zea mays*[35], *Oryza sativa*[36], *Nicotinium tabacum*[37], and *Marchantia polymorpha*[38]. Position numbers in the alignment are indicated above the sequence. Positions 1–80 are unaligned, whereas gaps (hyphens) are introduced in the remaining sequence to maximize alignment. Asterisks refer to points of coincident variation, referred to in Figure 3C. Between the *R.norvegicus* and *T.brucei* sequences, identical residues are indicated by a colon, and conservative substitutions are indicated by a period.

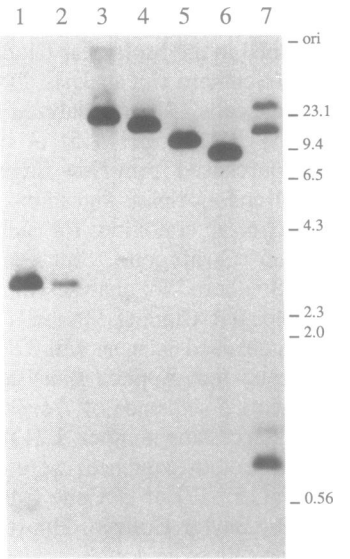

Figure 2. Southern blot of *T.brucei* DNA, probed with cloned *T.brucei* L25 cDNA. 2.5 μg of *T.brucei* DNA was digested with the restriction enzymes Bam HI (lane 3), Eco RI (lane 4), Kpn I (lane 5), Hind III (lane 6) or Stu I (lane 7). A reconstruction with 20pg (lane 1) or 2pg (lane 2) of linearized vector (pBSII digested with EcoRI) is included as a hybridization control. DNA markers (λ DNA restricted with Hind III) are shown at right (ori. = gel origin).

by the presence of sequence motifs KKAVVK and KKALK (Figure 1, positions 33 and 43, respectively)[10, 11]. The rat L23 has similar motifs at its N-terminus (KKAVLK and AKALK at Figure 1 positions 33 and 27, respectively), as does the *T.brucei* L25 (KKVAKK at position 23). These short motifs are likely to also be nuclear localization sequences in rat and *T.brucei*.

Since the analyses presented in this paper, and in previous studies[5, 6, 20, 23, 24], indicate that the L23/25 homolog is partially conserved in a wide variety of organisms, an evolutionary comparison of the LSU-rRNA sequences was also performed, to determine whether their putative L23/25 binding sites are similarly conserved (see Figure 3B). The segments chosen for analysis can be aligned with the L25 binding site in the *S.cerevisiae* LSU-rRNA, which was mapped by T_1-nuclease protection studies[2]. Evolutionary comparisons were limited to regions where LSU-rRNA sequences from at least three species could be aligned, thus eliminating several large expansion segments from the alignment. The lengths of the phylogenetic tree branches represent distance measures (statistical similarities) between the pre-aligned sequences. A comparison of the trees in Figures 3A and 3B indicates that the binding site within the rRNA varies significantly among eukaryotes, in comparison to the cognate L23/25 domains which are relatively similar. The branching orders of the rRNA and L23/25 sequences are nearly identical in the two comparisons. The only exceptions to this are the slightly different arrangements of *R. norvegicus*, which is grouped either with fungi in the rRNA tree, or with trypanosomes in the protein tree; and *M.vannielli*, which is grouped with *H.marismortui* in the rRNA tree but not grouped in the protein tree. The slight difference in the phylogeny of *N.tabacum* chloroplasts may not be statistically significant, given the short branch lengths separating these nodes of the trees. Other than these few differences, the similarity in branching order is consistent with co-evolution of the protein domain and rRNA binding site.

If the L23/25 domains and their respective rRNA binding sites have evolved in concert, the nucleotide/amino acid interactions may be different in many of the species indicated in Figure 3. As an example, *S.cerevisiae* L25 and *E.coli* L23 each bind the LSU-rRNA of the other species, and RNase-T1 protected fragments are similar in both homologous and heterologous binding complexes. However, both *S.cerevisiae* L25 and *E.coli* L23 fail to bind the LSU-rRNA of mice[8]. This incompatibility has been ascribed to the differing expansion segments in the rRNA of each species, or to the non-conservation of specific residues.

In order to extend our understanding of non-conserved elements of the L23/25 domain and rRNA binding site, we designed a computer program to search for points of coincident variation

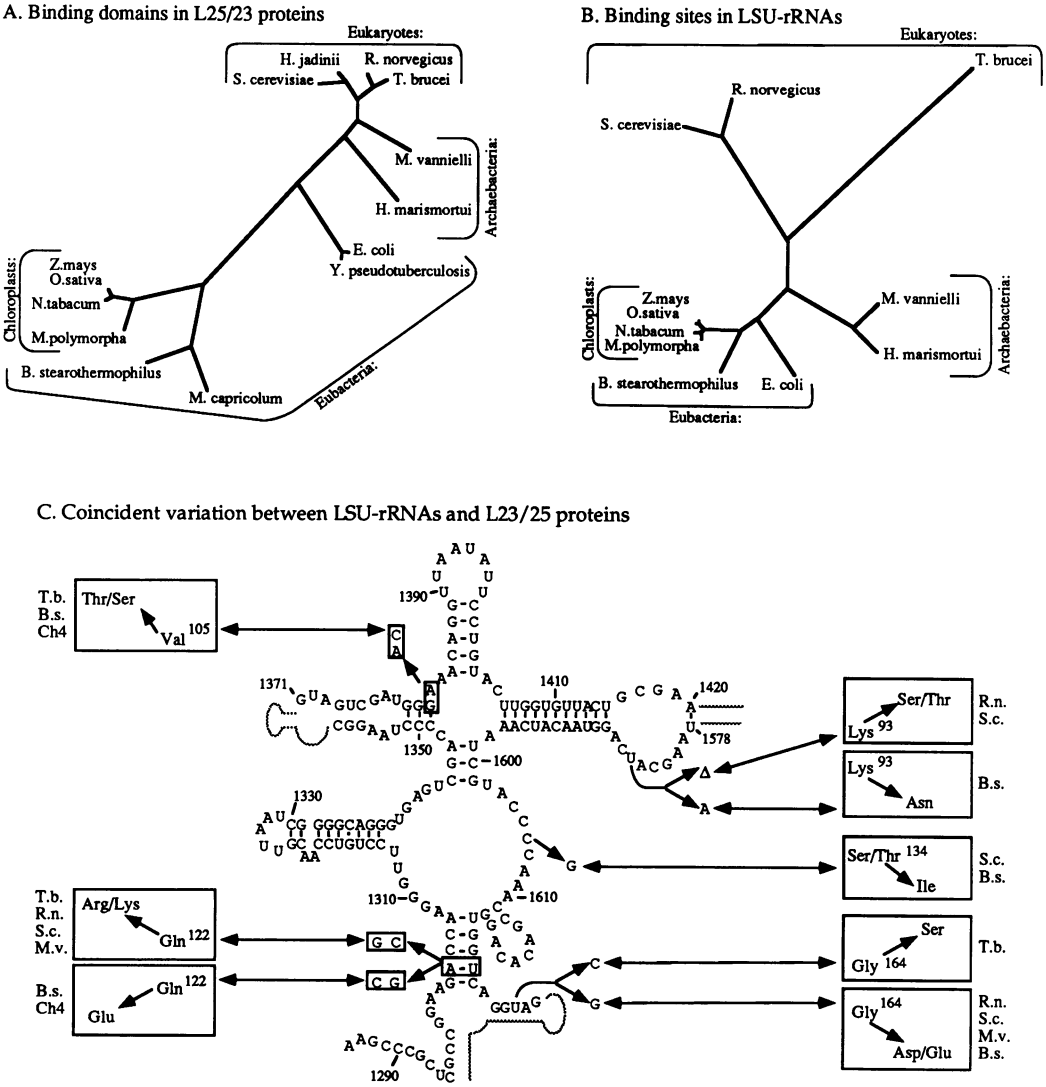

Figure 3. A) Phylogenetic tree of the L23/25 binding domain. The alignment shown in Figure 1 (positions 84−161) was used as the basis for the distance tree, as described in Materials and Methods. Species names are indicated in the legend to Figure 1. **B)** Phylogenetic tree of the L23/25 binding site in LSU-rRNA. The alignment used as a basis for the tree was derived from the Ribosome Database Project (see Materials and Methods). **C)** Position of coincident variations between LSU-rRNA and L23/25 proteins are indicated on the *E.coli* LSU-rRNA sequence. Single arrowheads indicate a mutation from the *E.coli* sequence coincident with (double arrowheads) a mutation in the L23/25 sequence (large boxes). Species with each respective mutation are tabulated by their initials, as indicated immediately outside of the large boxes (note that Ch4 = all four chloroplast species). Amino acid numbers refer to position numbers in Figure 1. Numbers on the LSU-rRNA refer to the *E.coli* sequence numbers. Stippled lines indicate extensions of sequence that lie outside the L23/25 binding site. Protein and rRNA sequences from the extreme halophile *H.marismortui* were not included in this analysis, however, due to the unusual acidity and salt-resistance of its ribosomal proteins[39].

between protein and LSU-rRNA. Coincident variation is defined as follows: For any nucleotide in the LSU-rRNA and for any amino acid residue in the L23/25 protein, if a mutation of the specified nucleotide between any two species is always accompanied by a non-conservative change in the specified amino acid residue (and vice-versa), the nucleotide and amino acid are said to vary coincidentally. Such cases of coincident variation may arise by chance, but may also include pairs of residues where a biologically significant interaction between rRNA and protein requires that a mutation in one of the macromolecules be accompanied by a compensatory change in the other.

A coincident variation analysis of the species indicated in Figure 3B are shown in Figure 3C, with mutations indicated on the *E.coli* LSU-rRNA. Amino acid residue numbers refer to the alignment shown in Figure 1, and the amino acids referred to in Figure 3C are highlighted in Figure 1 with an asterisk. The coincident variations fall into two classes: either the gain or loss of an acidic or basic amino acid side group, or a variation between aliphatic and aliphatic hydroxyl amino acids. As shown in Figure 3C, mutation of the ^{1304}A-^{1624}U pair to G-C or C-G coincides with a mutation of ^{122}Gln to a Arg/Lys or Glu residue respectively; mutation of ^{1629}U to G or C coincides with a mutation of ^{164}Gly to an Asp/Glu or Ser residue, respectively; mutation of ^{1584}U to an A, or its deletion (Δ), coincides with a mutation of ^{93}Lys to an Asn or a Ser/Thr residue respectively; mutation of ^{1382}G^{1383}A to AC coincides with a mutation of ^{105}Val to Ser/Thr; and mutation of ^{1606}C to G coincides with a mutation of ^{134}Ser to Ile. Coincident variations that fall along well established evolutionary boundaries (such as eukaryote/non-eukaryote, eubacteria/non-eubacteria and chloroplast/non-

chloroplast) were detected in this analysis, but are not shown. The coincident variations indicated in Figure 3C cannot be ascribed to a single evolutionary change (such as the distinction between eukaryotes and non-eukaryotes), and may therefore be significant as examples of molecular co-evolution.

One interesting question is how compensatory mutations can be established in functionally important residues, without strong selection against the initial (non-compensated) mutation. In the case of this ribosomal protein−rRNA interaction, the establishment of compensatory mutations may be facilitated by the fact that the rRNA is encoded by a multigene family [25−27]. For example, a rRNA mutation affecting the L23/25 binding function might be detrimental, were it not buffered by wild-type rRNA gene copies. Compensatory mutations in the L23/25 gene that restore binding function to the mutated rRNA could then develop in a background that was adaptively neutral, thus fixing the original rRNA mutation in the population. Conversely, molecular drive within the rDNA family may first amplify the mutant rRNA gene to the point where individual fitness is affected, analogous to the model proposed by Dover for rDNA and pol I genes[25].

Although it is premature to state that the coincident variations represent direct or compensatory linkages between protein and rRNA, they may serve as a basis for experimental dissection of the binding site. Complementation analysis between mutants of L23/25 protein and LSU-rRNA (which to date have been studied individually[12, 28]) may be useful in leading to an understanding of this conserved macromolecular interaction. Only such experimental dissection can establish that coincident variations seen in evolution are actually compensatory.

ACKNOWLEDGEMENTS

We thank Drs. Aïda Metzenberg, Robert Metzenberg, Jane Gitschier, Russell Doolittle, Gabriel A. Dover, Mitchell L. Sogin and members of the Agabian laboratory for helpful advice, discussions and/or critical review of the manuscript. S.M. is supported by a NRSA fellowship (GM13842). This work is supported by grants to N.A. from the John D. and Catherine T. MacArthur Foundation, and National Institutes of Health grant AI21975. N.A. is a Burroughs Wellcome Scholar of Molecular Parasitology.

REFERENCES

1. Vester, B. and Garrett, R.A. (1984) J. Mol. Biol. 179, 431−452.
2. El-Baradi, T.T.A.L., Raué, H.A., de Regt, V.C., Verbree, E.C., and Planta, R.J. (1985) EMBO J. 4, 2101−7.
3. Weitzmann, C.J. and Cooperman, B.S. (1990) Biochemistry 29, 3458−65.
4. El-Baradi, T.T.A.L., Raué, H.A., de Regt, C.H., and Planta, R.J. (1984) Eur J Biochem 144, 393−400.
5. Koepke, A.K. and Wittmann, L.B. (1989) Can. J. Microbiol. 35, 11−20.
6. Raue, H.A., Otaka, E., and Suzuki, K. (1989) J. Mol. Evol. 28, 418−26.
7. Egebjerg, J., Christiansen, J., and Garrett, R.A. (1991) J. Mol. Biol. 222, 251−64.
8. El-Baradi, T.T.A.L., de Regt, V.D.H.F., Planta, R.J., Nierhaus, K.H., and Raue, H.A. (1987) Biochimie 69, 939−948.
9. Dover, G.A. (1992) Bioessays 14, 281−7.
10. Rutgers, C.A., Schaap, P.J., van 't Riet, J., Woldringh, C.L., and Raue, H.A. (1990) Biochimica et Biophysica Acta 1050, 74−9.
11. Schaap, P.J., van 't Riet, J., Woldringh, C.L., and Raue, H.A. (1991) J. Mol. Biol. 221, 225−37.
12. Rutgers, C.A., Rientjes, J.M., van 't Riet, J., and Raue, H.A. (1991) J. Mol. Biol. 218, 375−85.
13. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
14. Milhausen, M., Nelson, R.G., Parsons, M., Newport, G., Stuart, K., and Agabian, N. (1983) Mol. Biochem. Parasitol. 9, 241−254.
15. Altschul, S.F. and Erickson, B.W. (1986) Bull. Math. Biol. 48, 603−616.
16. Klotz, L.C. and Blanken, R.L. (1981) J. Theor. Biol. 91, 261−272.
17. Feng, D. and Doolittle, R.F. (1987) J. Mol. Evol. 25, 351−360.
18. Olsen, G.J., Overbeek, R., Larson, N., Marsh, T.L., McCaughey, M.J., Maciukenas, M.A., Kuan, W.M., Macke, T.J., Xing, Y., and Woese, C.R. (1992) Nucl. Acids Res. 20 (Supplement), 2199−2200.
19. Agabian, N. (1990) Cell 61, 1157−1160.
20. Suzuki, K. and Wool, I.G. (1993) J. Biol. Chem. 268, 2755−61.
21. Wittman-Liebold, B. and Greuer, B. (1979) Febs Lett. 108, 69−74.
22. Joblet, C., Verspieren, P., Metzenberg, S., and Agabian, N. (Manuscript in preparation)
23. Hatakeyama, T. and Kimura, M. (1988) Eur. J. Biochem. 172, 703−711.
24. Fearon, K. and Mason, T.L. (1992) J. Biol. Chem. 267, 5162−70.
25. Dover, G.A. and Flavell, R.B. (1984) Cell 38, 622−623.
26. Hancock, J.M., Tautz, D., and Dover, G.A. (1988) Mol. Biol. Evol 5, 393−414.
27. Hancock, J.M. and Dover, G.A. (1990) Nucl. Acids Res. 18, 5949−5954.
28. Kooi, E.A., Rutgers, C.A., Mulder, A., van 't Riet, J., Venema, J., and Raue, H.A. (1993) Proc. Natl. Acad. Sci. USA 90, 213−6.
29. Leer, R.J., van Raamsdonk-Duin, M.M.C., Hagendoorn, M.J.M., Mager, W.H., and Planta, R.J. (1984) Nucl. Acids Res. 12, 6685−6700.
30. Woudt, L.P., Mager, W.H., Beek, J.G., Wassenaar, G.M., and Planta, R.J. (1987) Curr. Genet. 12, 193−198.
31. Koepke, A.K.E. and Wittmann-Liebold, B. (1988) FEBS Lett. 239, 313−318.
32. Gross, U., Chen, J.H., Kono, D.H., Lobo, J.G., and Yu, D.T.Y. (1989) Nucl. Acids Res. 27, 3601−3602.
33. Kimura, M., Kimura, J., and Ashman, K. (1985) Eur. J. Biochem. 150, 491−497.
34. Ohkubo, S., Muto, A., Kawauchi, Y., Yamao, F., and Osawa, S. (1987) Mol. Gen. Genet. 210, 314−322.
35. McLaughlin, W.E. and Larrinua, I.M. (1988) Nucl. Acids Res. 16, 8183.
36. Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori, M., Kondo, C., Honji, Y., Sun, C.R., Meng, B.Y., Li, Y.Q., Kanno, A., Nishizawa, Y., Hirai, A., Shinozaki, K., and Sugiura, M. (1989) Mol. Gen. Genet. 217, 185−194.
37. Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matsubayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohto, C., Torazawa, K., Meng, B.Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., and Sugiura, M. (1986) EMBO J. 5, 2043−2049.
38. Ohyama, K., Fukuzawa, H., Kohchi, T., Shirai, H., Sano, T., Sano, S., Umesono, K., Shiki, Y., Takeuchi, M., Chang, Z., Aota, S., Inokuchi, H., and Ozeki, H. (1986) Nature 322, 572−574.
39. Kimura, M., Arndt, E., Hatakeyama, T., Hatakeyama, T., and Kimura, J. (1989) Can. J. Microbiol. 35, 195−9.