Nucleic Acids Research, 1993, Vol. 21, No. 3 649–655

Zinc finger-like motifs in rat ribosomal proteins S27 and S29

Yuen-Ling Chan, Katsuyuki Suzuki, Joe Olvera and Ira G.Wool*
Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, IL 60637, USA

Received September 24, 1992; Revised and Accepted December 12, 1992 EMBL accession nos X59375, X59051

ABSTRACT

The primary structures of the rat 40S ribosomal subunit proteins S27 and S29 were deduced from the sequences of nucleotides in recombinant cDNAs and confirmed by determination of amino acid sequences in the proteins. Ribosomal protein S27 has 83 amino acids and the molecular weight is 9,339. Hybridization of cDNA to digests of nuclear DNA suggests that there are 4–6 copies of the S27 gene; the mRNA for the protein is about 620 nucleotides in length. Ribosomal protein S29 has 55 amino acids and the molecular weight is 6,541. There are 14–17 copies of the S29 gene and its mRNA is about 500 nucleotides in length. Rat ribosomal protein S29 is related to several members of the archaebacterial and eubacterial S14 family of ribosomal proteins. S27 and S29 have zinc finger-like motifs as do other proteins from eukaryotic, archaebacterial, eubacterial, and mitochondrial ribosomes. Moreover, ribosomes and ribosomal subunits appear to contain zinc and iron as well.

INTRODUCTION

A determination of the structure of ribosomes has as a requisite the sequences of nucleotides and of amino acids in the constituent nucleic acids and proteins. An attempt is under way to acquire these sequences for mammalian (rat) ribosomes (1). The primary purpose for the accumulation of this data is the importance it has for a solution of the structure of the organelle; the structure is in turn needed for a coherent, molecular account of the function of ribosomes in protein synthesis. The information on the sequences of amino acids (the structures of the rRNAs have already been determined) may also help in understanding the evolution of ribosomes, in unraveling the function of the proteins, in defining the rules that govern the interaction of the proteins and the rRNAs, and in uncovering the amino acid sequences that direct the proteins to the nucleolus for assembly on nascent rRNA. We report here the sequences of amino acids in rat ribosomal proteins S27 and S29. Rat S29 is related to the prokaryotic family of S14 ribosomal proteins which latter can be cross-linked to puromycin (2) and, hence, are likely to be in the ribosomal A-site and involved in the binding of aminoacyl-tRNA. Rat ribosomal proteins S27 and S29 have zinc finger-like motifs and ribosomes and ribosomal subunits appear to contain the metal.

MATERIALS AND METHODS

Recombinant procedures

The recombinant DNA procedures and the methods used to determine the sequences of nucleotides in nucleic acids have been described or cited (3–5). Rat S27 and S29 were isolated from a mixture of 40S ribosomal subunit proteins by high performance liquid chromatography on a reverse phase column (Bio-Rad, Hi-Pore C4; 250 mm×4.6 mm I.D.) with a 0 to 60% acetonitrile gradient containing 0.1% trifluoroacetic acid at a flow rate of 0.8 ml/min. The sequence of the 33 amino acids at the NH_2-terminus of rat ribosomal protein S27, and the entire sequence of amino acids in S29, were determined by Edman degradation in an automated gas phase sequencer (Applied Biosystems, Model 470A). A probe for the identification of a cDNA encoding rat ribosomal protein S27 was based on the sequence SLEEEKKKH (residues 10–18 of the mature protein); the probe was a mixture of 384 different oligodeoxynucleotides each 26 bases long. Probes for the isolation of a cDNA encoding rat ribosomal protein S29 were derived from two separate sequences of amino acids in the protein: The first probe, related to the sequence HQQLYWSHP (residues 2–10), was a mixture of 512 oligodeoxynucleotides each 26 bases long; the second, encoding the sequence KYGLNMC (residues 32–38), was in a mixture of 64 oligodeoxynucleotides each 21 bases long. The oligodeoxynucleotides were synthesized on a solid support by the methoxyphosphoramidite method using an Applied Biosystems, Model 380B, DNA synthesizer (6).

The preparation of rat liver ribosomes and of ribosomal subunits

Rat liver ribosomes were prepared (7) and suspended in buffer A (20 mM Tris–HCl, pH 7.6; 500 mM KCl; 3 mM $MgCl_2$; and 20 mM β-mercaptoethanol) and dialyzed 3 h against 100 vol of buffer A. Puromycin (0.1 mM final concentration) was added to the ribosomes after dialysis and they were incubated at 37°C for 15 min. Ribosomal subunits were separated by centrifugation at 25°C in 10–30% sucrose gradients in buffer A for 16 h at

* To whom correspondence should be addressed

18,000 rpm in a Beckman SW-28 rotor (7). Subunits, 40S and 60S, were collected in buffer (20 mM Tris−HCl, pH 7.6 and 20 mM $MgCl_2$) and sedimented by centrifugation at 30,000 rpm for 16 h at 5°C in a Beckman Ti30 rotor.

Determination of the content of zinc in ribosomes and ribosomal subunits

The ribosomes or the ribosomal subunits were suspended in buffer A and 2 vol acetic acid were added. The final concentrations were: 66% acetic acid; 6.7 mM Tris−HCl, pH 7.6; 167 mM KCl; 1 mM $MgCl_2$; and 6.7 mM β-mercaptoethanol. The suspension was stirred for 1 h at 4°C and then centrifuged for 30 min at 6,500 ×*g*; the supernatant which contains the ribosomal proteins and metal ions was analyzed by atomic absorption with a Perkin-Elmer, model 306 spectrometer. The quantity of Fe, Co, Ni, Cu, Zn, and Cd in the samples was estimated from comparison with standards. The quantity of Zn was determined at 213.9 nm and compared to standards (Spex Industries) in 2% HCl. The content of metals in 66% acetic acid in 0.3× buffer A (the blank) used to suspend the ribosomes or ribosomal subunits, which was usually below the limit for detection, was subtracted from the sample value. Precautions were taken with the glassware; they were soaked in chromic acid-sulfuric acid, rinsed thoroughly with deionized water, and autoclaved.

RESULTS AND DISCUSSION

The sequences of nucleotides in recombinant cDNAs encoding rat ribosomal proteins S27 and S29

A random selection of 20,000 colonies from two rat liver cDNA libraries (3) of 20,000 and 30,000 independent transformants was screened with an oligodeoxynucleotide probe for rat ribosomal protein S27. A clone, pS27-4, was selected and the sequences of nucleotides from both strands of the cDNA and overlapping sequences for each restriction site were obtained. The cDNA insert in pS27-4 is 278 bases long with a 5′ noncoding sequence of 41 nucleotides and a single open reading frame of 237. The open reading frame encodes 79 amino acids but lacks a termination codon and a 3′ noncoding sequence. The cDNA insert in pS27-4 was made radioactive and used to screen a third rat cDNA library. The sequence of nucleotides in the cDNA insert in the longest clone, pS27-14, that gave a positive hybridization signal was determined (Fig. 1). The cDNA, apart from a long poly(A)tail, has 496 nucleotides; a 5′ noncoding sequence of 77, a single open reading frame of 255, and a 3′ noncoding region of 164. This is the longest 3′ noncoding sequence that has been found in a mammalian ribosomal protein cDNA. The overlapping sequences in pS27-4 and pS27-14 are identical and, hence, they are likely to be derived from the same gene. The open reading frame in pS27-14 starts at an AUG codon at a position that we designate +1 and ends with a termination codon (TAA) at position 253; it encodes 84 amino acids (Fig. 1). The initiation codon occurs in a context, ACC*AUG*C, that is close to the consensus sequence ACC*AUG*G (8). The hexamer AATAAA that directs posttranscriptional cleavage and polyadenylation of the 3′ end of the precursor of the mRNA (9) is at position 393−398, 21 nucleotides upstream of the start of a long poly(A) sequence.

A random selection of 80,000 cells from two cDNA libraries (see above) was screened for clones that hybridized to two oligodeoxynucleotide probes for rat ribosomal protein S29. No clone gave a positive signal with these probes. We assumed that the libraries lacked a cDNA for S29; therefore, an alternative procedure was adopted. A pair of oppositely oriented degenerate oligodeoxynucleotide primers based on the NH_2- and carboxyl-terminal amino acid sequences in S29 were synthesized and used in the polymerase chain reaction to amplify cDNAs made from rat liver poly(A)$^+$mRNA with reverse transcriptase (3). An amplified DNA of 180 base pairs was selected and cloned in a pGEM-2 vector; the cDNA insert was designated pGEM-2-S29. Transcripts prepared from pGEM-2-S29 with T7 RNA polymerase (10) were translated in a nuclease treated rabbit reticulocyte lysate and the protein that was synthesized migrated on electrophoresis in polyacrylamide gels to the same position as authentic S29. The sequence of nucleotides in both strands

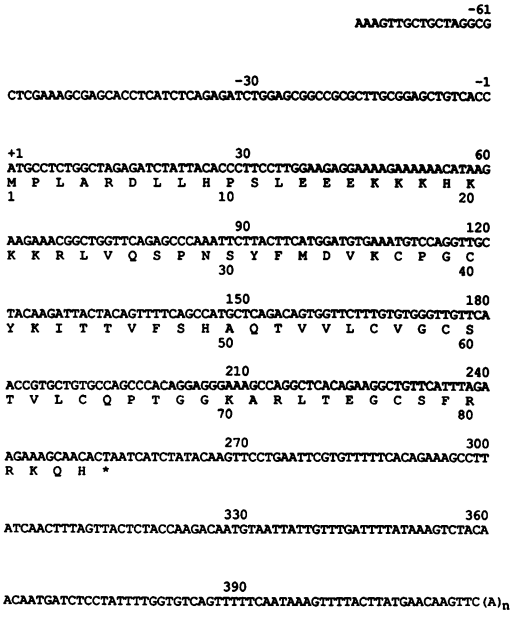

Figure 1. The sequence of nucleotides in the cDNA insert in plasmid pS27-14 and the amino acid sequence encoded in the open reading frame. The positions of the nucleotides in the cDNA insert are given above the residues; the positions of amino acids in protein S27 are designated below the residues.

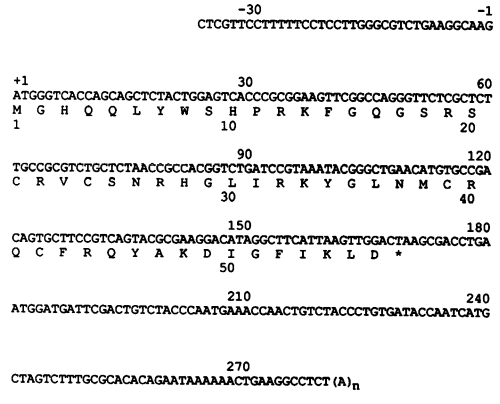

Figure 2. The sequence of nucleotides in the cDNA insert in plasmid pS29-7 and the amino acid sequence encoded in the open reading frame. The positions of the nucleotides in the cDNA insert are given above the residues; the positions of the amino acids in protein S29 are designated below the residues.

of the cDNA insert in pGEM-2-S29 was determined (Fig. 2). The DNA has a single open reading frame that encodes a protein of 56 amino acids; this sequence is congruent with that determined directly from the protein except at 5 positions which were ambiguous in the latter determination. Because of the method of preparation, pGEM-2-S29 of necessity lacked the 5′ and 3′ noncoding sequences of a S29 cDNA. For this reason a third rat cDNA library was screened using as a probe radioactive pGEM-2-S29 cDNA. Three transformants gave a positive hybridization signal and the sequences of both strands and of overlapping sequences for restriction sites for one of the clones, pS29-7, was determined by the dideoxy chain termination method.

The cDNA insert in pS29-7, apart from a long poly(A) stretch, has 317 nucleotides: a 5′ noncoding region of 37, a single open reading frame of 171, and a 3′ noncoding sequence of 109 (Fig. 2). As with pS27-14 the 3′ noncoding sequence is exceptionally long for a mammalian ribosomal protein cDNA. The open reading frame begins at an ATG codon at a position that we designate +1 and ends with a termination codon (TAA) at position 169; it encodes 56 amino acids (Fig. 2). The initiation codon occurs in a context, AAG*AUG*G, that differs little from the consensus sequence ACC*AUG*G (8). The sequence, AATAAA, that directs cleavage and polyadenylation of the 3′end of the precursor to the mature mRNA (9) is at position 260−265, 15 nucleotides upstream of the start of a long poly(A) sequence. The first 20 nucleotides of the S29 cDNA (positions −37 to −18) are, with the exception of one residue, pyrimidines. Pyrimidine sequences are present at the 5′ end of many, if not all, eukaryotic ribosomal protein mRNAs (1) and may play a role in the regulation of their translation (11).

The primary structures of rat ribosomal proteins S27 and S29

The protein specified by pS27-14 was identified from the congruence (a single residue excepted) of the sequence of 33 amino acids at the NH_2 terminus of S27, determined directly from the protein by Edman degradation, and that deduced from the sequence of nucleotides. However, the amino acid composition of the protein encoded by pS27-14 closely approximates that reported before (12) for S27Â (now designated S27a) rather than for S27 (13). This discrepancy was resolved in the following way: The protein whose NH_2-terminal amino acid sequence was determined to correspond to that of the open reading frame in pS27-14 (see above) was shown to migrate on electrophoresis in two-dimensional gels with the coordinates of S27 rather than S27a (Fig. 3). This confirmed the identity of the protein specified by pS27-14 as rat ribosomal protein S27. Therefore, in the original reports of the amino acid compositions of S27 (13) and S27a (12) the proteins must have been transposed; this no doubt resulted from the two occupying nearby positions in Kaltschmidt-Wittmann two-dimensional urea-polyacrylamide gels. The interchange of S27 and S27a had been suggested before from an analysis of the latter by the 'four-corners' method of two-dimensional polyacrylamide gel electrophoresis (14).

The molecular weight of rat ribosomal protein S27, calculated from the sequence of amino acids deduced from pS27-14, is 9,470; this is somewhat less than was estimated from the migration of the purified protein in sodium dodecyl sulfate gels (12); small, basic proteins often give anomalous molecular weights by this procedure. The NH_2-terminal methionine encoded in the S27 mRNA is removed after translation since it is not found in the amino acid sequence derived from the protein. The residue that follows the initial methionyl is prolyl which has been reported (15) to favor NH_2-terminal processing. Thus the number of residues in the mature protein is 83 and the molecular weight is 9,339.

Protein S27 has more basic residues (5 arginyl, 10 lysyl, and 4 histidyl) than acidic ones (2 aspartyl and 4 glutamyl) (Table I). As has been noted before for ribosomal proteins, the basic residues tend to be clustered; for example, there are 8 consecutive basic residues at positions 16−23. There is a prominent hydrophilic region at the NH_2 terminus; 14 of the first 22

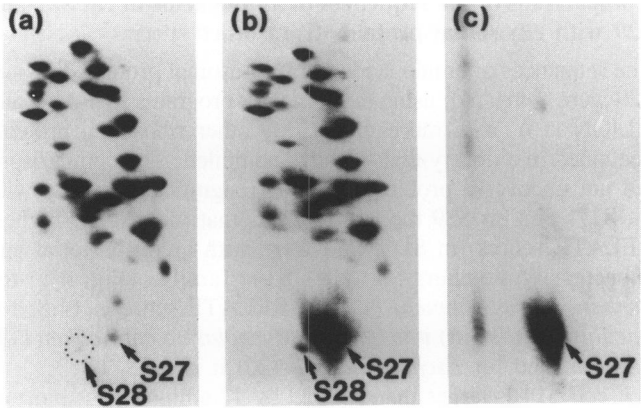

Figure 3. Identification of the protein encoded in pS27-14. In *(a)*, 80 μg of a mixture of the proteins of rat 40S ribosomal subunits was separated by two-dimensional electrophoresis (left to right in the first dimension and top to bottom in the second) in a urea-polyacrylamide gel and the proteins were stained with Coomassie blue. In *(b)*, 80 μg of a similar mixture of 40S ribosomal proteins was supplemented with the proteins obtained by extraction with 67% acetic acid from 15 μl of a micrococcal nuclease-treated rabbit reticulocyte lysate that was used to translate the pS27-14 cDNA derived transcript; the analysis of the proteins was as in *(a)*. In *(c)*, the radioactive protein synthesized in the reticulocyte lysate that was supplemented with [^{35}S]methionine employed in *(b)* was visualized by radioautography of the polyacrylamide gel.

Table 1. The amino acid compositions of rat ribosomal proteins S27 and S29

Amino acid[a]	S27 A	S27 B	S29 A	S29 B
Alanine	4	3	1	1
Arginine	5	5	7	7
Aspartic acid and asparagine	4	2 + 1	5	2 + 2
Cysteine	n.d.	6	n.d.	4
Glutamic acid and glutamine	8	4 + 4	5	0 + 5
Glycine	6	5	7	6
Histidine	4	4	4	3
Isoleucine	2	1	3	3
Leucine	7	8	4	4
Lysine	9	10	5	4
Methionine	1	2[b]	1	2[b]
Phenylalanine	3	3	3	3
Proline	5	5	2	1
Serine	6	6	4	4
Threonine	5	6	0	0
Tryptophan	n.d.	0	n.d.	1
Tyrosine	1	2	3	3
Valine	7	7	2	1
Residues		84		56

[a] The amino acid composition (in numbers of residues) was determined (A) from an analysis of an hydrolysate of purified S27 (12) or S29 (12) or inferred (B) from the sequence of nucleotides in recombinant cDNAs.
[b] The NH_2-terminal methionine is removed after translation of the mRNA.

```
RS29                                                                M
MvS14                                                               M
BsS14                          MAKKS MIA..... ..... ..... .....
McS14                          MAKKS LK...VK.QA KHPKF..... .....
EcS14                          MAKQS MKAREVKRVA LADKYFAKRA ELKAIISDVN
ScMRP2          MGNFRFPIKT KLPPGFINAR ILRDNFKRQQ FKENEILVKS .LKFIARNMN
Consensus                      mak.s ...  .K .. .....              .

RS29            GHQQLYWSHP RKFGQ..... ..GSRS...C RVCSNR..HG LIRKYGLNMC
MvS14           TKEPF....K TKYGQ..... ..GSKV...C KRCG.RKGPG IIRKYGLDLC
BsS14           .KQQR..TPK FKV.Q..... .....EYTRC ERCG.RP.HS VIRKF..KLC
McS14           .... ....N VR..... ..... .NYTRC NHCG.RP.HA VLKKFGI..C
EcS14           ASDEDRWNAV LKL.QTLPRD SSPSRQRNRC RQTG.RP.HG FLRKFGLS..
ScMRP2          LPTKLRLEAQ LKLN.ALPNY MRSTQIKNRC VDSG.H.ARF VLSDF..RLC
Consensus       .......... .k..Q        .sr...RC ..cg rp h. virkfGl.lC

RS29            RQCFRQYA.. KDI.GFIKLD
MvS14           RQCFRELA.. PKL.GFKKYD
BsS14           RICFRELAYK GQIPGVKKAS W
McS14           RLCFRKFAYE GQIPGIKKAS W
EcS14           RIKVREAAMR GQIPGLKKS
ScMRP2          RYQFRENALK GNLPGVKKGI W
Consensus       R.cfRe.A.k g.iPGvkK.. W
```

Figure 4. Comparisons of the amino acid sequences of rat S29 and the eubacterial and archaebacterial S14 family of ribosomal proteins. The alignment of the several sequences was by reiteration using the programs GAP and PRETTY from the GCG Sequence Analysis Package (40). In the consensus sequence invariant amino acids are designated by a capital letter; where there are only conservative changes (R, K, or H; I, V, or L; and D or E) or a residue is present in all but one of the proteins the designation is with a small letter. The abbreviations are: RS29, rat S29; MvS14, *Methanococcus vannielii* S14; BsS14, *Bacillus subtilis* S14; McS14, *Mycoplasma capricolum* S14; EcS14, *Escherichia coli* S14; and ScMRP2, *Saccharomyces cerevisiae* mitochondrial MRP2.

```
RS27            37   C PG    C YKITTVFSHAQTVVL C VG      C
RS29            21   C RV    C SNRHGLIRKYGLNM  C RQ      C

BsuS14          23   C ER    C GRPHSVIRKFKL    C RI      C
McS14           24   C NH    C GRPHAVLKKFGI    C RL      C
MvS14           17   C KR    C GRKGPGIIRKYGLDL C RQ      C
ScMRP2          78   C VDS     GHARFVLSDFRL    C RYQFR
EcS14           63   C RQT     GRPHGFLRKFGL      SRIKV

RL37a           39   C SF    C GKTKMKRRAVGIWH  C GS      C
RL37            18   C RR    C GSKAYHLQKST     C GK      C
HCEP52          20   C RK    C YARLHPRAVN      C RKKK    C
HCEP80          45   C PSDE  C GAGVFMASHFDRHY  C GK      C

BstL32          29   C PN    C GEWKLAHRV       C KA      C
BsuL36          11   C EK    C KVIRRKGKVMVI    C ENPK    H
EcL36           11   C RN    C KIVKRDGVIRVI    C SAEPK   H
```

Figure 5. Ribosomal proteins with zinc finger-like motifs. The abbreviations for the species and for the ribosomal proteins are: RS27, rat S27; RS29, rat S29; BsuS14, *Bacillus subtilis* S14; McS14, *Mycoplasma capricolum* S14; MvS14, *Methanococcus vannielii* S14; ScMRP2, *Saccharomyces cerevisiae* mitochondrial MRP2; EcS14, *Escherichia coli* S14; RL37a, rat L37a; RL37, rat L37; HCEP 52, human ubiquitin carboxyl-terminal extension protein 52; HCEP80, human ubiquitin carboxyl-terminal extension protein 80; BstL32, *Bacillus stearothermophilus* L32; BsuL36, *Bacillus subtilis* L36; EcL36, *Escherichia coli* L36.

residues are charged. Finally, 12 of the 19 residues at positions 46–64 are hydrophobic.

The protein specified by the reading frame in pS29-7 was identified as S29, in the first instance, from the amino acid composition (Table I) which closely approximates that derived from an hydrolysate of the protein (12). Positive identification was from the close identity of the entire sequence of amino acids determined by Edman degradation of S29 and that deduced from the sequence of nucleotides in pS29-7; the exceptions were 5 residues in the sequence of 55 amino acids from the protein that were not certain.

The molecular weight of rat ribosomal protein S29, calculated from the sequence of amino acids deduced from pS29-7, is 6,672; this is quite a bit less than was originally estimated from the migration of the purified protein in sodium dodecyl sulfate gels (12). However, when the molecular weight was redetermined using for calibration small basic proteins of known mass the value obtained was 6,500 (results not shown). The NH_2-terminal methionine encoded in the S29 mRNA is removed after translation; it is not found in the sequence obtained from the protein. The residue that follows the initial methionyl is glycyl which favors NH_2-terminal processing (15). Thus, S29 has 55 residues and the molecular weight is 6,541.

Protein S29 has more basic residues (7 arginyl, 4 lysyl, and 3 histidyl) than acidic ones (2 aspartyl) and lacks glutamic acid and threonine (Table I).

The numbers of copies of the S27 and S29 genes

The cDNA insert in pS27-4 was made radioactive and was used to probe digests made from rat liver nuclear DNA with restriction endonucleases (*Bam*HI, *Eco*RI, or *Hin*dIII) (4). The number of hybridization bands suggests that there are 4–6 copies of the S27 gene (data not shown). A similar procedure with the cDNA insert in pS29-7 gave 14–17 hybridization bands (data not shown). There are multiple copies of most, if not all, mammalian ribosomal protein genes (cf. 1 for references and discussion). However, in no instance has it been shown that more than one of the genes is functional; the presumption is that the additional copies are retroposon pseudogenes.

The sizes of the mRNAs encoding rat ribosomal proteins S27 and S29

To determine the sizes of the mRNAs coding for S27 and S29, total poly(A)$^+$mRNA from rat liver was separated by electrophoresis and screened by northern hybridization using radioactive pS27-4 and pS29-7 cDNAs. One distinct band of about 620 nucleotides was detected for S27 and one of approximately 500 nucleotides for S29 (data not shown).

Comparisons of the sequences of amino acids in rat S27 and S29 with ribosomal proteins from other species

The sequences of amino acids in rat ribosomal proteins S27 and S29 were compared, using the computer programs RELATE and ALIGN (16), with more than 1,000 other ribosomal proteins contained in a library that we have compiled. The comparisons did not uncover a protein that shared significant identity with rat S27. For rat S29 the comparisons that yielded the highest RELATE scores (in SD units) were with archaebacterial and eubacterial members of the S14 family (Fig. 4): for *Methanococcus vannielii* (17) the RELATE score is 14.8; for *Bacillus subtilis* (18) it is 6.7; for *Mycoplasma capricolum* (19) it is 6.5; and for *Escherichia coli* (20) it is 5.7.

E.coli S14 is larger than rat S29 by 42 amino acids; most of this difference is accounted for by the absence from rat S29 of the 33 amino acids at the NH_2 terminus of *E.coli* S14. *B. subtilis* S14 (60 residues) is closer in size to rat S29 (55 residues) than to *E.coli* S14 (98 residues). Despite the difference in the number of residues, *B. subtilis* and *E.coli* S14 are related in structure; the RELATE score is 8.9 SD and in an alignment of the amino acid sequences there are 27 identities in 58 possible matches; the ALIGN score is 14.2 SD. S14 is not only required for assembly of 30 S particles (21) but it may be responsible also for determining the active conformation of 16 S rRNA at the A site (22).

Zinc finger-like motifs in ribosomal proteins

Rat ribosomal proteins S27 and S29 have zinc finger-like domains (Fig. 5). The founding father of the family of proteins with zinc finger motifs is *Xenopus laevis* transcription factor IIIA (23). The consensus sequence for the domain, which occurs in tandem nine times in TFIIIA, is: -(Y,F)-X-C-X_{2-4}-C-X_3-F-X_5-L-X_2-H-$X_{3,4}$-H-X_{2-6}- (24). In many of the related proteins the histidines are replaced by a second pair of cysteines (25) as in S27 and S29 which have a domain of the form -C-X_2-C-$X_{14,15}$-C-X_2-C- at positions 37−59 and 21−42 respectively (Figs. 1, 2, and 5). The number of proteins with zinc finger motifs in one or another of its sequence variations exceeds a thousand (Zinc Finger Data Base; information courtesy of George Michaels). All have the potential to coordinate a zinc ion and almost all bind to nucleic acids, most to DNA, a few to RNA (25, 26). TFIIIA, for example, binds to both the internal control region of 5 S rDNA and to 5 S rRNA, i.e. to both the gene and its transcript (27). The binding to the internal control region of the gene is sequence specific through base-specific contacts in the major groove, whereas binding of TFIIIA to 5S rRNA is prescribed by backbone structure in the nucleic acid (28).

The secondary and tertiary structures of zinc fingers of the C_2-H_2 and C_2-C_2 types are distinct, indeed, may not be related (29). In the C_2-H_2 motif, there is a two-stranded antiparallel β-sheet that includes the two cysteines, a turn, and then an amphipathic α-helix that encompasses both histidines (30). In the C_2-C_2 type, the domain lacks a β-sheet and the amphipathic α-helix begins after the third cysteine. Nonetheless, both bind to DNA by presentation of the α-helix to the major groove. It is import that proteins having C_2-H_2-type zinc fingers can bind to DNA as monomers, TFIIIA is an example, whereas, those with C_2-C_2 motifs generally bind as dimers. Finally, the C_2-C_2 motif often occurs as a tandem repeat with the α-helix of the first repeat binding to DNA and the second providing a surface for interaction with the other protein of the dimer (29). How these motifs participate in binding to RNA, if indeed they do, is not known.

It is perhaps relevant here that the C_2-C_2 type motifs generally occur as tandem repeats in DNA binding proteins (29) but singly in ribosomal proteins. The single C_2-C_2 domain in the ribosomal proteins corresponds most closely to the first or DNA binding motif of, for example, the steroid receptor proteins (29).

In S27 and S29 the amino acid sequence between the internal cysteinyls, which we shall for convenience refer to as the linker sequence, is dominated by basic and hydrophobic residues. For S27 the sequences is, YKITTVFSHAQTVVL; of the 15 residues, 2 are basic and 7 are hydrophobic. The sequence in S29 is, SNRHGLIRKYGLNM; of the 14 residues, 4 are basic and 5 are hydrophobic. In addition, the positions of the basic and hydrophobic residues tend to be preserved (cf. later). We do not know if the linker sequences in S27 and S29 form α-helices and if they do whether they are amphipathic; we are not able to draw amphipathic wheels from the sequences without arbitrary assumptions. We note that the S27 linker sequence has tyrosine and phenylalanine and that of S29 has tyrosine.

We have searched our library of amino acid sequences for other ribosomal proteins with zinc finger-like motifs. Rat ribosomal proteins L37a (31) and L37 have domains of the same form and the linker sequences also have basic and hydrophobic residues (Fig. 5). Of the archaebacterial and eubacterial S14 family of ribosomal proteins related to rat S29 the members from *M. vannielii* (17), from *B. subtilis* (18), and from *M. capricolum* (19) have the motif (Fig. 5); whereas *E.coli* S14 (20) has at most a degenerate form of the domain. In the alignment with the others (Fig. 5) there is a cysteine in *E.coli* S14 at what would be the initial position; the other three are absent. The hydrophilic and hydrophobic character of the linker sequence is preserved in *E.coli* S14 and 8 of the 12 residues share identity with amino acids at the same positions in *B. subtilis* S14. The yeast nuclear encoded mitochondrial ribosomal protein MRP2 (32) is also related to the S14 family and it too has a degenerate form of the motif; the sequence, -CVDSGHARFVLSDFRLC-, has only two cysteinyls; they are at positions that correspond to the first and third sites in the motif (Fig. 5). Once again the linker sequence is basic (3 of 12 residues), hydrophobic (6 residues), and has two aromatic amino acids. The spin given to these findings is that *E.coli* S14 and yeast mitochondrial MRP2 once had full C_2-C_2 motifs and that parts were lost during divergent evolution; there is, of course, no direct evidence for this bias. Finally, variants of the zinc finger motif are to be found in the human ubiquitin carboxyl-terminal extension proteins CEP52 (33) and CEP80 (14); the latter is ribosomal protein S27a and the former is likely to be a large ribosomal subunit protein but its identity has not been established. Finally, we note that the ribosomal proteins with zinc finger motifs are small and that the motif constitutes an appreciable fraction of the entire sequence; for example 22 of 55 residues in S29 and 23 of 84 in S27. In addition, most of the conserved residues in related proteins, for example in the S14 set, are in the motif (cf. Fig. 4).

An approximate consensus amino acid sequence for the linker peptide can be derived. It has the form, -GxBxBxZxBBZxxZ-, where B designates a basic residue, Z a hydrophobic amino acid, and x a position with no consensus. What is most striking is the retention of the pattern of basic and hydrophobic residues.

Table 2. Metal ion content of ribosomes and ribosomal subunits determined by atomic absorption spectroscopy

Ribosome Particles	Fe	Co	Ni	Cu	Zn	Cd
			μg/ml			
40S	0.16	<0.021	<0.072	<0.012	0.29	<0.009
60S	0.82	<0.021	<0.072	<0.012	0.59	<0.009
80S	0.18	<0.021	<0.072	<0.012	0.40	<0.009
70S	0.24	<0.021	<0.072	<0.012	0.41	<0.009

Table 3. Zinc content of ribosomes and ribosomal subunits

	Ribosome Particles *mol/mol*	Zn/particle		
Expt.		I	II	III
	40S	3.2	5.0	3.7
	60S	5.8	5.8	2.7
	80S	n.d.	5.1	2.9
	70S	n.d.	3.1	1.7

n.d., not done

The presence of zinc in ribosomes and ribosomal particles

Analysis of rat liver ribosomes and ribosomal 40S and 60S subunits by atomic absorption spectroscopy indicated the presence of appreciable amounts of Zn (Table II has the raw data from an experiment). The amount of Co, Ni, Cu, and Cd is less than can be detected by the analytical method. However, ribosomal particles also contain appreciable quantities of Fe which may be

a component of the ribosome (see below), or may be a contaminant from hemoglobin or from other Fe-containing proteins in the blood or in the liver. We recognize that if Fe is a contaminant then it is possible that Zn is also. It may be significant then that *E.coli* 70S ribosomes were also found to have Fe (Table II).

What is more reassuring is the observation that at least one ribosomal protein is capable of binding iron. A protein was purified to homogeneity from rat liver, using its ability to bind iron, and characterized (34). The molecular mass (16 kDa), the metal-binding specificity, and the intracellular location (predominantly in microsomes) distinguished this protein from transferrin, lactoferrin, and ferritin. The absorption spectrum indicated the protein does not contain iron prosthetic groups such as iron-sulfur, heme, or flavins and hence is not related to cytochromes or ferrodoxins. Amino acid sequences of tryptic peptides derived from the protein (35) are identical with sequences in rat ribosomal protein P2 (36). Some 4 to 7% of the radioactive iron in hepatoma cells incubated with [^{59}Fe] transferrin was precipated with a monoclonal antibody specific for the ribosomal P proteins (P0, P1, and P2 share a common carboxyl-terminal epitope). Thus rat ribosomal protein P2 can bind iron and this capacity, perhaps shared by other ribosomal proteins, may account for our finding that ribosomes contain the metal. How P2 binds the metal is by no means obvious since the protein lacks cysteine and histidine (36); however, the protein does have acidic amino acids and coordination of iron by carboxylates does occur (37, 38).

The concentration of Zn in the samples of rat 80S ribosomes and of 60S and 40S subunits varied (Table III). For rat 80S ribosomes the stoichiometry (moles of Zn per mole of ribosomal particle) ranged from 3 to 6; for 60S subunits from 3 to 6; and for 40S subunits from 3 to 5. We cannot account for the sum of the moles in the two subunits being greater than in 80S ribosomes. We attribute the uncertainties in the calculation of the stoichiometry to the difficulty of determining with precision, either by UV absorption or chemically, the concentrations of ribosomes or of ribosomal subunits; this is primarily because these preparations are suspensions rather than solutions. We recognize that there may be a source of Zn contamination that we have overlooked. Our best estimate from the data is that there are 3 moles of Zn in each of the subunits and that the 80S ribosome has 6 moles. This would accord, but perhaps only fortuitously, with the identification of 3 proteins with zinc binding motifs in each of the ribosomal subunits. *E.coli* 70S ribosomes have a single protein, L36, with a zinc finger-like motif (a CCCH box) yet appear to have 2 moles of Zn (Table III). We caution that although rat and *E.coli* ribosomes have proteins with zinc binding motifs and may contain Zn there is no direct evidence from our experiments that the proteins bind the metal. However, it has been shown that a protein from *Entamoeba histotytica* that is related to rat S27 (there are 43 amino acid identities out of 83 possible matches between residues) will bind radioactive zinc specifically (Stanley, S.L., personal communication).

Do the zinc finger motifs participate in the binding of ribosomal proteins to rRNA?

There is no evidence that the zinc finger motifs in ribosomal proteins participate directly in binding to RNA. Indeed, it is possible that the zinc finger motif is a vestige of a former function, for example, the binding to DNA, that is fortuitously preserved in ribosomal proteins that employ other means for association with rRNA. That the zinc finger domains in the ribosomal proteins are chemical fossils is consistent with the preservation in *E.coli* S14 and in yeast MRP2 of only a part of the motif, recognizable but presumably no longer capable of chelating a metal ion and hence of forming the canonical secondary and tertiary structures. Furthermore, it is likely that the same sequences and/or structures in the *B. subtilis* and *E.coli* S14 proteins are used to bind to 16S rRNA since the binding site is highly conserved (39). If this assumption is correct then the conclusion must be that the S14 proteins do not use the structure that can be formed by the zinc finger motif in binding to rRNA since *E.coli* S14 lacks the capacity to form the structure. A reconciliation is that binding to rRNA employs the side chains of the basic and hydrophobic amino acids in the linker region of the motif in S14 rather than the finger structure *per se*.

ACKNOWLEDGMENTS

We are grateful to our former colleague Alan Lin who initiated the determination of the amino acid sequence of S29, to Andrew Davis for the metal ion determinations, to Anton Glück for assistance with the amino acid sequence comparisons and for fruitful discussions, to Veronica Paz for technical assistance, and to Arlene Timosciek for aid in the preparation of the manuscript. This research was supported by a grant GM-21769 from the National Institutes of Health.

REFERENCES

1. Wool, I.G., Endo, Y., Chan, Y.L. and Glück, A. (1990) In Hill, W.E., Dahlberg, A., Garrett, R.A., Moore, P.B., Schlessinger, D. and Warner, J.R. (eds.), The Ribosome: Structure, Function, and Evolution. Amer. Soc. Microbiol., Washington, D.C. pp. 203–214.
2. Kerlavage, A.R. and Cooperman, B.S. (1986) *Biochemistry*, **25**, 8002–8010.
3. Chan, Y.L., Lin, A., McNally, J. and Wool, I.G. (1987) *J. Biol. Chem.*, **262**, 12879–12886.
4. Chan, Y.L. and Wool, I.G. (1988) *J. Biol. Chem.*, **263**, 2891–2896.
5. Glück, A., Chan, Y.L., Lin, A. and Wool, I.G. (1989) *Eur. J. Biochem.*, **182**, 105–109.
6. Beaucage, S.L. and Caruthers, M.H. (1981) *Tetrahedron Lett.*, **22**, 1859–1862.
7. Martin, T.E. and Wool, I.G. (1969) *J. Mol. Biol.*, **43**, 151–161.
8. Kozak, M. (1986) *Cell*, **44**, 283–292.
9. Proudfoot, N.J. and Brownlee, G.G. (1976) *Nature*, **263**, 211–214.
10. Chan, Y.L., Devi, K.R.G., Olvera, J. and Wool, I.G. (1990) *Arch. Biochem. Biophys.*, **283**, 546–550.
11. Levy, S., Avni, D., Hariharan, N., Perry, R.P. and Meyuhas, O. (1991) *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 3319–3323.
12. Collatz, E., Ulbrich, N., Tsurugi, K., Lightfoot, H.N., MacKinlay, W., Lin, A. and Wool, I.G. (1977) *J. Biol. Chem.*, **252**, 9071–9080.
13. Collatz, E., Wool, I.G., Lin, A. and Stöffler, G. (1976) *J. Biol. Chem.*, **251**, 4666–4672.
14. Redman, K.L. and Rechsteiner, M. (1989) *Nature*, **338**, 438–440.
15. Flinta, C., Persson, B., Jörnvall, H. and von Heijne, G. (1986) *Eur. J. Biochem.*, **154**, 193–196.
16. Dayhoff, M.O. (1978) In Dayhoff, M.O. (ed.), Atlas of Protein Sequence and Structure. National Biomedical Research Foundation, Washington, D.C., Vol.5, Suppl. 3, pp. 1–8.
17. Auer, J., Spicker, G. and Böck, A. (1989) *J. Mol. Biol.*, **209**, 21–36.
18. Henkin, T.M., Moon, S.H., Mattheakis, L.C. and Nomura, M. (1989) *Nucleic Acids Res.*, **17**, 7469–7486.
19. Ohkubo, S., Muto, A., Kawauchi, Y., Yamao, F. and Osawa, S. (1987) *Mol. Gen. Genet.*, **210**, 314–322.
20. Yaguchi, M., Roy, C., Reithmeier, R.A.F., Wittmann-Liebold, B. and Wittmann, H.G. (1983) *FEBS Lett.*, **154**, 21–30.
21. Ramakrishnan, V., Graziano, V. and Capel, M.S. (1986) *J. Biol. Chem.*, **261**, 15049–15052.
22. Powers, T., Stern, S., Changchien, L.M. and Noller, H.F. (1988) *J. Mol. Biol.*, **201**, 697–716.

23. Miller, J., McLachlan, A.D. and Klug, A. (1985) *EMBO J.*, **4**, 1609–1614.
24. Berg, J.M. (1990) *Annu. Rev. Biophys. Biophys. Chem.*, **19**, 405–421.
25. Berg, J.M. (1986) *Science*, **232**, 485–487.
26. Klug, A. and Rhodes, D. (1987) *Trends Biochem. Sci.*, **12**, 464–469.
27. Brown, D.D. (1984) *Cell*, **37**, 359–365.
28. You, Q., Veldhoen, N., Baudin, F. and Romaniuk, P.J. (1991) *Biochemistry*, **30**, 2495–2500.
29. Schwabe, J.W.R. and Rhodes, D. (1991) *Trends Biochem. Sci.*, **16**, 291–296.
30. Pavletich, N.P. and Pabo, C.O. (1991) *Science*, **252**, 809–817.
31. Tanaka, T., Aoyama, Y., Chan, Y.L. and Wool, I.G. (1989) *Eur. J. Biochem.*, **183**, 15–18.
32. Myers, A.M., Crivellone, M.D. and Tzagoloff, A. (1987) *J. Biol. Chem.*, **262**, 3388–3397.
33. Baker, R.T. and Board, P.G. (1991) *Nucleic Acids Res.*, **19**, 1035–1040.
34. Furukawa, T., Taketani, S., Kohno, H. and Tokunaga, R. (1991) *Biochem. Biophys. Res. Comm.*, **181**, 409–415.
35. Furukawa, T., Uchiumi, T., Tokunaga, R., and Taketani, S. (1992) *Arch. Biochem. Biophys.* **298**, 182–186.
36. Wool, I.G., Chan, Y.L., Glück, A. and Suzuki, K. (1991) *Biochimie*, **73**, 861–870.
37. Armstrong, W.H., Spool, A., Papaefthymiou, G.C., Frankel, R.B. and Lippard, S.J. (1984) *J. Am. Chem. Soc.*, **106**, 3653–3667.
38. Slater, A.F.G., Swiggard, W.J., Orton, B.R., Flitter, W.D., Goldberg, D.E., Cerami, A. and Henderson, G.B. (1991) *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 325–329.
39. Stern, S., Powers, T., Changchien, L.M. and Noller, H.F. (1989) *Science*, **244**, 783–790.
40. Devereux, J., Haeberli, P. and Smithies, O. (1984) *Nucleic Acids Res.*, **12**, 387–395.