RECOGNITION OF RNA BY RIBOSOMAL PROTEIN S1: INTERACTION OF S1 WITH 23 S rRNA OF *ESCHERICHIA COLI*

Alain KROL, Christiane BRANLANT, and Jean-Pierre EBEL

Institut de Biologie Moléculaire et Cellulaire du CNRS, Université Louis Pasteur, 15 rue Descartes, 67084 Strasbourg Cedex, France

and

Louis Peter VISENTIN*

The Division of Biological Sciences, The National Research Council of Canada, Ottawa, Ontario, K1A OR6, Canada

Received 27 June 1977

1. Introduction

Specific interaction of proteins and nucleic acids is both a structural prerequisite for special cell organelles such as ribosomes and chromatin, as well as the chemical basis for the control of cellular information flow. This form of biological 'cybernetics' is best exemplified by operon–repressor interactions [1,2], tRNA–tRNA synthetase recognition [3] and more recently virion-RNA–protein complexes [4] and mRNA discrimination by the ribosome [5,6]. The latter is a complex system in *Escherichia coli* cells involving a region of the 3′-terminus of 16 S rRNA and a ribosomal protein, S1 [7].

Ribosomal protein S1 has been shown to be necessary for the correct translation of synthetic and natural mRNA [8], the replication of Qβ RNA as a subunit of Qβ replicase [9], and in functions involving ribosome subunit interactions [10]. The multiplicity and complexity of these reactions suggest very specific recognition properties for S1, a conclusion supported by recent biophysical studies indicating that S1 can discriminate between double and single stranded synthetic polynucleotides [11,12]. However, it is not known whether the specific recognition of nucleic acids by S1 is due to some specific structural arrangement of nucleotide bases like a palindrome, cruciform [13], a particular nucleotide sequence or pattern therein.

We report herein that S1 forms a specific complex with 23 S rRNA of *E. coli* which, when treated with T_1 ribonuclease yields a ribonucleoprotein particle (RNP) containing a single stranded RNA fragment, 17 nucleotides long, derived from the L1 binding site on 23 S rRNA. The properties of the complex formation suggest sequence recognition by S1, a sequence, which bears remarkable resemblance to the S1 binding site on Qβ RNA and the 3′-terminus of 16 S rRNA. We discuss these findings in relation to RNA 'pattern recognition' by S1 and S1 function. The type of disjunctive linear sequence of RNA that is common to S1 binding regions may be the basic unit of the S1 recognition and/or interaction.

2. Materials and methods

2.1. *Materials*

Preparation of ^{32}P-labelled 23 S rRNA: ^{32}P-labelled 23 S rRNA was prepared from ribosomes of *E. coli* MRE 600 as previously described [14]. The 23 S rRNA was deproteinized either by three successive

* A CNRS-NRCC Exchange fellow 1976-77

phenol extractions in TM4 buffer (10 mM Tris–HCl (pH 7.5), 0.1 mM $MgCl_2$) or by three successive extractions in phenol-0.5% Na dodecylsulphate in water. Traces of the remaining phenol or SDS were eliminated by centrifuging the 23 S rRNA through a 10% sucrose cushion in TM3 buffer (10 mM Tris–HCl (pH 7.5), 1 mM $MgCl_2$).

Preparation of *E. coli* S1: S1 was prepared from the ribosomes of *E. coli* MRE 600 as previously described [15] and its purity established by electrophoresis, amino acid composition, amino-terminal sequence (Met–Thr–Glu–Ser–Phe–), and its immunological cross-reactivity against anti-sera prepared against S1 and/or subunit 1 of Qβ replicase [15], which was generously supplied by Dr Albert Wahba of the Laboratory of Molecular Biology, Sherbrooke University Medical School, Sherbrooke, P.Q. Canada.

2.2. *Methods*

Preparation of the protein S1–23 S rRNA complexes: S1–23 S [^{32}P]RNA complexes were prepared as described earlier [16] in TMK reconstitution buffer (10 mM Tris–HCl (pH 7.5), 20 mM $MgCl_2$, 350 mM KCl, 6 mM β-mercaptoethanol). Various molar ratios of S1 – RNA, from 0.5 : 1 to 30 : 1 were used to form the S1–23 S rRNA complexes starting from an S1 stock solution containing 10 mg/ml of S1 in 10 mM Tris buffer (pH 7.5) and 6 mM β-mercaptoethanol. The temperature of incubation was either 0°C or 42°C for 1 h. Variation in K^+ and or Mg^{2+} were assessed in similar systems.

The complexes were digested with T_1 ribonuclease (Sankyo, Japan) at a concentration of 1 mg/ml at enzyme–complex ratios varying from 1 : 10 down to 1 : 5 (w/w). The RNP fragments were subsequently fractionated electrophoretically in 8% polyacrylamide gels containing Mg^{2+} [13]. The gel slices containing the S1–RNA complexes were either dissociated in polyacrylamide gels containing 8M urea or eluted directly by electrophoresis methods [16]. The RNA contained therein was digested with T_1 or pancreatic ribonuclease and fingerprinted [17,18].

Complexes with Tobacco mosaic virus Omega fragment and S1 were carried out as for the 23 S rRNA except that the 5′-terminal fragment was purified previously and was a kind gift of Dr Ken Richards of the Institut de Biologie Moléculaire du CNRS, Virology Group, Strasbourg, France.

3. Results and discussion

The fractionation of the products arising from the enzymatic digestion of the 23 S rRNA–S1 complex is shown in fig.1a. The arrow indicates the band corresponding to the complex which is absent when the 23 S rRNA is incubated in the absence of S1. As has been reported previously for the binding of Qβ RNA by S1, saturation binding is not achieved even at molar ratios of S1–RNA of 10 : 1 or 15 : 1 [19]. However, on subsequent analysis of the RNP formed, whatever the concentration of S1, yielded a single, unique species of rRNA indicative of the binding of a specific nucleotide sequence rather than a non-specific random interaction. Binding of 23 S rRNA by S1 is independent of temperature and occurs at 0°C or 42°C with no difference in specificity. In a similar fashion a systematic analysis of the addition of K^+ and Mg^{2+} indicated that complex formation

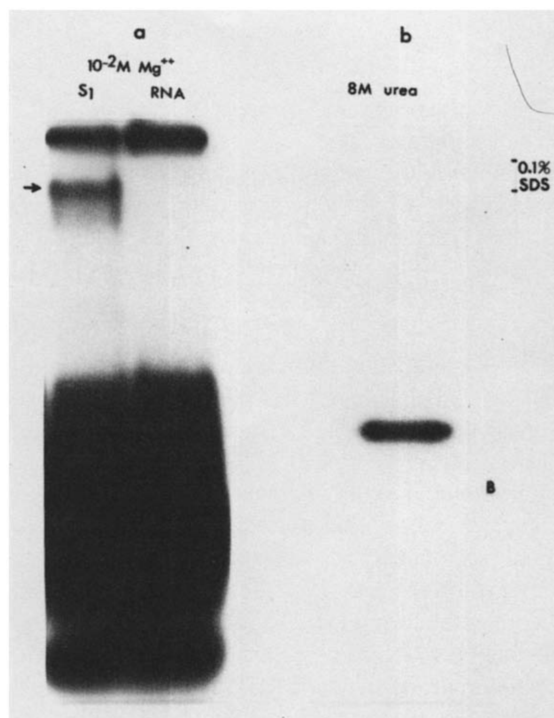

Fig.1. (a) Fractionation of the products of hydrolysis of 23 S rRNA–protein S1 complex. (b) Fractionation of the products resulting from the dissociation of the RNP shown in figure 1a by electrophoresis in a polyacrylamide gel slab containing 8 M urea and SDS (see Materials and methods).

between S1 and 23 S rRNA is optimal and unique under conditions specified for reconstitution at K^+ concentrations between 0.3–0.5 M [24]. In this range and under optimal digest conditions (1 : 5), a single RNA fragment is generated with no contaminating background. However, S1 was found to interact with 23 S rRNA in the absence of Mg^{2+} which indicates that a defined tertiary structure is not necessary for S1 recognition of RNA. These data are supported by earlier studies which indicated a Mg^{2+} independence for S1 binding to Qβ RNA [19] and in part, by biophysical studies that show that S1 has a high affinity for single stranded synthetic polynucleotides [11,12,20] as well as a single stranded region at the 3′-terminus of 16 S rRNA [7].

The S1–23 S RNP as shown in fig.1a was subsequently dissociated by electrophoresis in acrylamide gels containing 8 M urea (fig.1b) and the single unique fragment was further analyzed by fingerprinting the T_1. (fig.2) and pancreatic RNAase digests. As can be seen in fig.2, a single oligonucleotide is present in the T_1 ribonuclease digest with the characteristic mobility of T_1 oligonucleotide 2 from the 13 S fragment and T_1 oligonucleotide 3 of the 18 S fragment [21]. The fingerprint of the pancreatic digest reveals that this oligonucleotide contains AAAUp, AAUp, 2ACp, and several Up and Cp. Therefore it can be identified as oligonucleotide 3 of the 18 S fragment: AAAUACCCUUUACAAUG which is only present

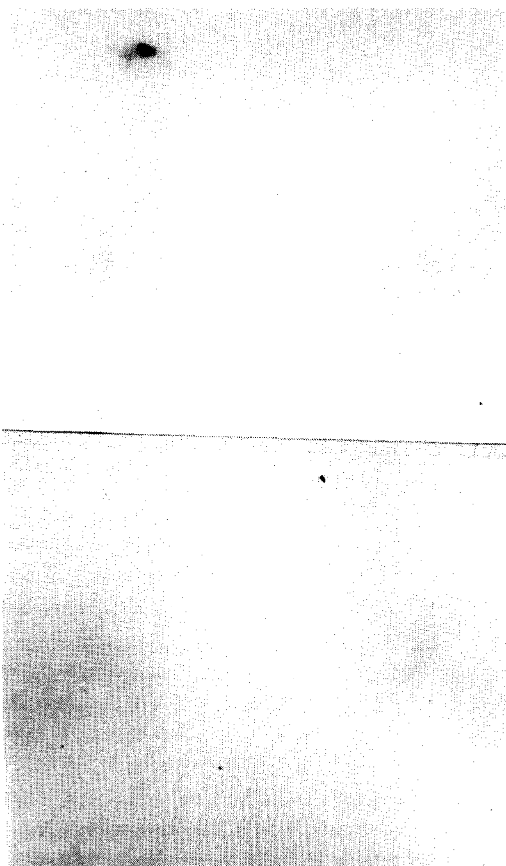

Fig.2. Fingerprint of a T_1 + alkaline phosphatase digest of fragment shown in fig.1b.

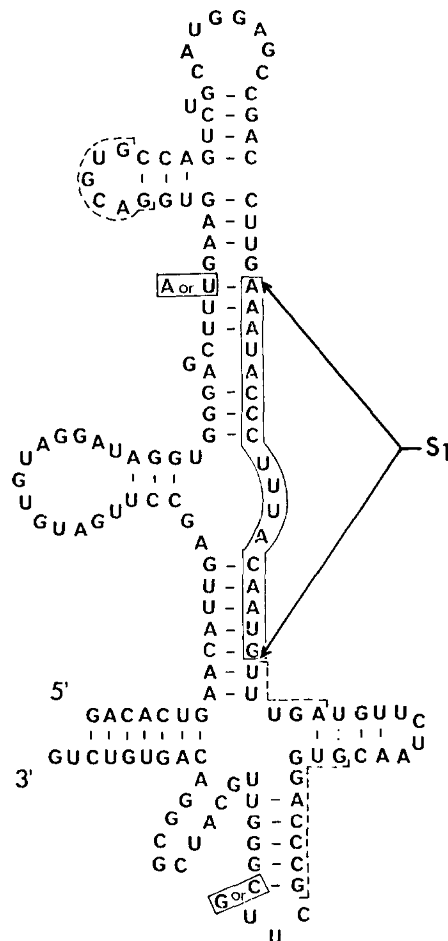

Fig.3. Secondary structure of the L1 binding site on 23 S RNA showing the binding site of protein S1 [22].

once in the 23 S RNA [21] in the binding site for protein L1 [22]. When digest conditions were less drastic the entire rRNA binding site for protein L1 was found associated with protein S1. The protection of this region (against ribonuclease action) was poorer than that exhibited by L1, since upon dissociation in an 8 M urea gel, the ribonucleoprotein releases only very small fragments. The oligonucleotide that interacts with protein S1 is supposed to be involved in a base pairing according to the model proposed for the L1 binding site (fig.3). It is conceivable that S1 destabilizes this base pairing and therefore the secondary structure of the whole region, which would explain the high degree of degradation by T_1 ribonuclease.

However, the finding that a restricted region of the 23 S rRNA constituted an S1 binding site was remarkable. This would indicate that S1 can recognize, select, and interact with a specific sequence, 17 nucleotides long, even though it is interfaced with a 23 S rRNA molecule encompassing 3000 nucleotides and having alternative purine or pyrimidine-rich stretches that could serve as binding sites.

In fact, table 1 indicates the frequency of other sequence patterns containing pyrimidines that could potentially serve as S1 binding sites, in 23 S rRNA but in fact, do not. This interaction with a limited sequence of rRNA is contrasted to the interactions of other ribosomal proteins with rRNA where multisite attachment appears to be the rule [14,16]. This reflects the unique structural [23] and functional properties of S1 [5] and its dynamic mode of action during protein synthesis as compared to the more rigid structural role of other ribosomal proteins that are involved in site specific complex formation at the early stages of ribosome assembly [24].

Table 1
Frequency of sequence patterns present in the recognition site for protein S1 in 23 S T_1 – oligonucleotides containing more than one uridine

Sequence pattern	Frequency[a]
(1) Pyrimidine runs greater than 4	22
(2) AC. . .	7
(3) Pyrimidines. . .UUA pyrimidines	4
(4) AC. . .A pyrimidines	2
(5) AU. .AC. . pyrimidines	1
(6) AU. .AC.	1

[a] According to [21]

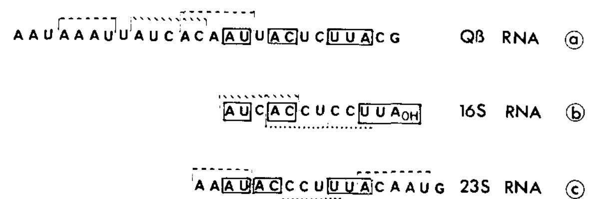

Fig.4. Comparison of S1 binding sites on Qβ RNA (a) [19], 16 S RNA (b) [7] and 23 S RNA (c). Dashed lines encompass areas of homology between Qβ RNA and 23 S RNA of binding sites. Dotted lines encompass areas of homology between 23 S rRNA and 16 S rRNA S1 binding sites. Slanted-dashed lines dilineate Qβ RNA and 16 S RNA homologies in S1 binding sites. The enclosed areas are the patterns that occur in all three S1 binding sites on Qβ RNA, 16 S RNA and 23 S RNA.

We compared the sequence of this S1 binding site on 23 S rRNA with the sequences of S1 binding sites on Qβ RNA and 16 S rRNA. This comparison is shown in fig.4. We examined these sequences for homologies and it is obvious from the data displayed that several restricted tetranucleotide homologies can be found. However, only the enclosed sequences shown in fig.4 are found in all three types of RNA. A linear periodicity for the common sequences is evident and each group is separated by C and/or U spacers of varying frequency. Whether this type of periodic clustering of unique sequences is what S1 really recognizes is not yet certain and alternative explanations are possible. However, we present this segmented clustering of sequences as a model, recognition or interaction site for S1, a model, unlike the palindrome or cruciform shapes recognized by other functional proteins [13], that is based on a linear pattern.

Another sequence that could be considered as a possible interaction site for S1 is a repetitive cluster which appears in Qβ RNA and 23 S rRNA, that is UUAC . . . . UUAC. As a preliminary test of this proposition we made complexes of S1 with a unique segment of tobacco mosaic virus RNA (the Omega region at the 5′-terminus) which is approximately 70 nucleotides, completely lacking in G and contain-

ing a repetitive sequence of UUAC with spacers (Ken Richards, et al., unpublished results).

S1 binds this oligonucleotide almost quantitatively (Krol et al. unpublished results) and the site fits our model of pattern recognition. Additionally we have tested the binding of S1 to 16 S RNA and have found that S1 binds the native 16 S molecule. It is known that S1 binds the 3′-terminus of 16 S rRNA [7] and it would be of interest to determine whether S1 binds other interior regions known to contain the UUAC sequence preceded by AC and/or AU [25]. However, in the 23 S rRNA, the UUAC sequence does not in itself, appear to be sufficient to act as an interaction site as this sequence occurs in other 23 S rRNA regions [21]. Therefore it would appear that a pattern including this sequence is required for specific interaction.

As to the in vivo situation regarding the binding of S1 to 23 S rRNA we cannot conclude that S1 interacts with the 23 S rRNA in the intact 50 S subunit. Recent data indicates that S1 will only bind to the 50 S subunit at a 50 S : S1 ratio of 0.15 [23]. This does not exclude the possibility that S1 interacts temporally on both subunits during initiation. It should be noted that L1 protein is accessible on the surface of the ribosome and the L1 RNA binding site is expected to be at the subunit interface [26]. Similarly, it has been shown that S1 stimulates the binding of tRNA by the 50 S subunit [10] suggesting that a dynamic interaction is possible. This dynamic aspect of S1 interaction has been shown for 30 S–S1 complex [23] and certainly would be a prerequisite for a protein so important in the fixation of mRNA and simultaneously in Qβ RNA replication [5,6] and possibly even transcription [27].

As S1 has an unusually elongated shape ([23], and Vachette personal communication) and an unusually short and characteristic binding site it may prove to be most amenable to biophysical and chemical studies of the protein binding sites involved in RNA recognition. Certainly it should be useful in the analysis of a model for interaction and recognition which proposes stacking interactions between a linear RNA sequence and aromatic functional groups of proteins [28].

Acknowledgments

This work was supported by the Délégation Générale à la Recherche Scientifique et Technique, by the Commissariat à l'Energie Atomique and the National Research Council of Canada.

The authors thank A. Cosset for her skilled technical assistance, and to Dr S. Hasnain, Dr K. Johnson, Mr Wally Rowsome and Mr D. Griffith for the preparations of S1.

References

[1] Gilbert, W. and Maxam, A. (1973) Proc. Natl. Acad. Sci. USA 70, 3581–3584.
[2] Bahl, C. P., Wu, R., Stawinsky, J. and Narang, S. A. (1977) Proc. Natl. Acad. Sci. USA 74, 966–970.
[3] Rich, A. (1974) Biochimie 56, 1441–1449.
[4] Flanegan, J. B., Pettersson, R. F., Ambros, V., Hewlett, M. J. and Baltimore, D. (1977) Proc. Natl. Acad. Sci. USA 74, 961–965.
[5] Steitz, J. A., Wahba, A. J., Laughrea, M. and Moore, P. B. (1977) Nucl. Acid. Res. 4, 1–14.
[6] Sobura, J. E., Chowdry, M. R., Hawley, D. A. and Wahba, A. J. (1977) Nucl. Acid Res. 4, 17–29.
[7] Dahlberg, A. E. and Dahlberg, J. E. (1975) Proc. Natl. Acad. Sci. USA 721, 2940–2944.
[8] Van Dieijen, G., Van der Laken, C. J., Van Knippenberg, P. H. and Van Duin, J. (1975) J. Mol. Biol. 93, 351–366.
[9] Wahba, A. J., Miller, M. J., Niveleau, A., Landers, T. A., Carmichael, G. G., Weber, K., Hawley, D. A. and Slobin, L. I. (1974) J. Biol. Chem. 249, 3314–3316.
[10] Thomas, G., Sweeney, R., Chang, C. and Noller, H. F. (1975) J. Mol. Biol. 95, 91–102.
[11] Bear, D. G., Ng, R., Van Derveer, D., Johnson, N. P., Thomas, G., Schleich, T. and Noller, H. (1976) Proc. Natl. Acad. Sci. USA 73, 1874–1878.
[12] Draper, D. E. and Von Hippel, P. H. (1976) ICN-UCLA Symposium on Molecular and Cellular Biology 5, in press.
[13] Helene, C. (1977) La Recherche 8, 122–132.
[14] Branlant, C., Sri Widada, J., Krol, A. and Ebel, J. P. (1977) Eur. J. Biochem. 74, 155.
[15] Visentin, L. P., Hasnain, S., Johnson, K., Griffith, D., Gallin, W. and Wahba, A. (1977) FEBS Lett. 79, 258.
[16] Branlant, C., Krol, A., Sri Widada, J., Fellner, P. and Crichton, R. (1973) FEBS Lett., 35, 265–272.
[17] Sanger, F., Brownlee, G. G. and Barrell, B. G. (1965) J. Mol. Biol. 13, 373–398.
[18] Brownlee, G. G. and Sanger, F. (1967) J. Mol. Biol. 23, 337–353.
[19] Senear, A. W. and Steitz, J. A. (1976) J. Biol. Chem. 251, 1902–1912.
[20] Szer, W., Hermoso, J. M. and Boublik, M. (1976) Biochem. Biophys. Research Comm. 70, 957–964.
[21] Branlant, C. and Ebel, J. P. (1977) J. Mol. Biol. 3, 215–256.

[22] Branlant, C., Korobko, V. and Ebel, J. P. (1976) Eur. J. Biochem. 70, 471–482.
[23] Laughrea, M. and Moore, P. B. (1977) J. Mol. Biol. 112, 399–421.
[24] Mizushima, S. and Nomura, M. (1970) Nature 226, 1214–1218.
[25] Ehresmann, C., Stiegler, P., Fellner, P. and Ebel, J. P. (1975) Biochimie 57, 711–748.
[26] Stöffler, G. (1974) in: Ribosomes (Nomura, M., Tissieres, A. and Lengyel, S. eds) pp. 615–667 Cold Spring Harbor.
[27] Van Dieijen, G., Van Knippenberg, P. H., Van Duin, J., Koekman, B. and Pouwels, P. H. (1977) Molec. Gen. Genet. 153, 75–80.
[28] Brun, F., Toulme, J. J. and Helene, C. (1975) Biochemistry 14, 558–563.