Volume 220, number 2, 283-287 FEB 04965 August 1987

Hypothesis

Similarities between a predicted secondary structure for the M1 RNA ribozyme and the tRNA binding center of 16 S rRNA from *E. coli*

Siegfried Boehm

Central Institute of Molecular Biology, Department of Molecular Biophysics, Academy of Sciences of the GDR, 1115 Berlin-Buch, GDR

Received 20 May 1987

We propose a new model for the secondary structure of the M1 RNA component of *E. coli* RNase P which is based on significant sequence homologies with parts of the *E. coli* 16 S rRNA. A large domain of the new model resembles closely the secondary structure of the tRNA binding center of 16 S rRNA. We suggest that this domain of M1 RNA when functioning as a ribozyme binds the mature part of the precursor tRNA.

M1 RNA; 16 S rRNA; Ribozyme; Secondary structure; Substrate recognition; tRNA binding

1. INTRODUCTION

RNase P from *E. coli* precisely cleaves extra nucleotides from the 5'-termini of tRNA precursors to generate the 5'-termini of mature tRNAs (for a recent review see [1]). This enzyme is composed of a RNA molecule (M1 RNA containing 377 nucleotides) and a protein component (C5 protein, 14 kDa). The M1 RNA is the catalytic subunit of RNase P and is involved in the binding of the substrate to the holoenzyme [2]. Because M1 RNA, like some other RNA molecules (reviewed, e.g., in [3]), acts as a true enzyme, it has been called 'RNA enzyme' or 'ribozyme'.

The recognition of the substrate by the M1 RNA is an unsolved problem as yet. As outlined in [1] the "substrate recognition by the M1 RNA differs from various splicing reactions in which either 'guide sequences' or 'consensus sequences' are believed to play an important role in substrate recognition [3]. Instead the recognition, if indeed governed by Watson-Crick base pairing, might rather rely on single nucleotides scattered throughout the sequence of both substrate and RNA enzyme".

An important aspect of understanding the substrate recognition and catalytic function of M1 RNA is the elucidation of its secondary and tertiary structure. The secondary structures proposed so far for the M1 RNA [1,4,5] have not provided definitive insights into possible relationships between its structure and function, e.g. they do not permit the identification of separate domains for substrate recognition and catalytic activity.

By computer assisted search for sequence homologies we have found several significant homologies in the primary structures between parts of the M1 RNA and parts of the 16 S rRNA from *E. coli*. Moreover using these data on sequence

Correspondence address: S. Boehm, Central Institute of Molecular Biology, Department of Molecular Biophysics, Academy of Sciences of the GDR, 1115 Berlin-Buch, GDR

Published by Elsevier Science Publishers B.V. (Biomedical Division)

homologies and the well established secondary structure of 16 S rRNA (reviewed in [6,7]) as a guide, the primary structure of the M1 RNA can be folded into a secondary structure so that parts of it closely resemble the tRNA binding center of 16 S rRNA.

2. RESULTS AND DISCUSSION

Fig.1 shows a base pairing scheme for the M1 RNA which is completely different for more than two-thirds of the whole molecule from a current model proposed by Altman et al. [1]. Only the double-stranded regions 1, 2 and 7 and their adjacent single-stranded parts are identical (with slight variations in helix 2) as illustrated in fig.2.

Our new model is based primarily on significant sequence homologies between parts of the M1 RNA and parts of the 16 S rRNA. The homologous sequence stretches between both RNAs which are used as a guide for the M1 RNA model are listed in table 1 and are indicated in figs 1 and 2.

The new model explains better than the model of Altman et al. the nuclease cleavage data taken from [4]. In particular, the model of the M1 RNA predicted here is more compatible with the double-strand specific cleavage sites reported in [4]. Several sites (sixteen) are located in double-stranded regions in our model (fig.1) but occur in single-stranded regions in the model of Altman et al. [1]. Thus, there is no requirement to involve tertiary base pairing interactions for these sites as assumed in [4].

The M1 RNA model proposed here contains a significantly higher amount of paired bases (about 62%) compared to the model of Altman et al. (about 49%) if the noncanonical GU and GA base pairs are included in the calculation.

An interesting feature of the proposed model are the structural repeats within the M1 RNA. The following short sequence repeats are apparent: (i) AGGGUGC (from positions 86 to 92 and 182 to 188, respectively) and (ii) GGUAAC (from positions 95 to 100 and 211 to 216, respectively).

These sequence repeats are involved in similar secondary structure motifs. In [1] it was reported that two molecules of the C5 protein interact with one molecule of M1 RNA in the holoenzyme complex. It may be assumed therefore that the noted repeated structures in the M1 RNA are each involved each in the binding of a C5 protein molecule. As shown in a schematic diagram in fig.2 the most striking feature of the suggested M1 RNA model is its similarity to rather large parts of the tRNA binding center of the 16 S rRNA.

The central part of the tRNA binding center of 16 S rRNA shown in fig.2 consists of the so-called cleft anchor [9], the adjacent decoding region around nucleotide C 1400 [10] and a 'pocket' or 'cage' which contains the helices 27–31 [11]. This 3′-domain is connected through helices 1–3 and helix 19 with more distant parts of the 16 S rRNA which are involved in tRNA binding as well (for details see [11]).

A comparison of the described central part of the tRNA binding center of 16 S rRNA with the M1 RNA structural model (fig.2) shows a surprising general analogy between large parts of both models. The M1 RNA model looks like a more or less truncated and modified variant of the described domain of 16 S rRNA. It resembles strongly the secondary structure of the 3′-domain of a minimal small subunit rRNA from kinetoplasts (see e.g. [12]).

In [11] it was reported that tRNA binding to ribosomes shields a set of highly conserved nucleotides in the 16 S rRNA from attack by chemical probes in both the presence and absence of mRNA. In fig.2 the nucleotide positions of the central part of the tRNA binding center of 16 S rRNA which are protected in the absence of

Fig.1. The predicted new secondary structure model for the *E. coli* M1 RNA. Double-stranded regions of M1 RNA are numbered from the 5′-end. Identical nucleotides within aligned homologous stretches of 16 S rRNA and M1 RNA are boxed (see text). GC and AU base pairs are represented by asterisks, GU and AG base pairs by circles. Dotted lines indicate additional (weaker) base pairs. Double-strand specific cutting sites for M1 RNA according to [4] are marked by crosses. Double crosses indicate sites which are compatible with the new model but not with the current model of Altman et al. [1]. Underlined crosses mark cleavage sites which fit Altman's model but not our model. Brackets between position 281 and 282 indicate a single nucleotide deletion in M1 RNA within the alignment of this sequence stretch with its homologous stretch in 16 S rRNA.

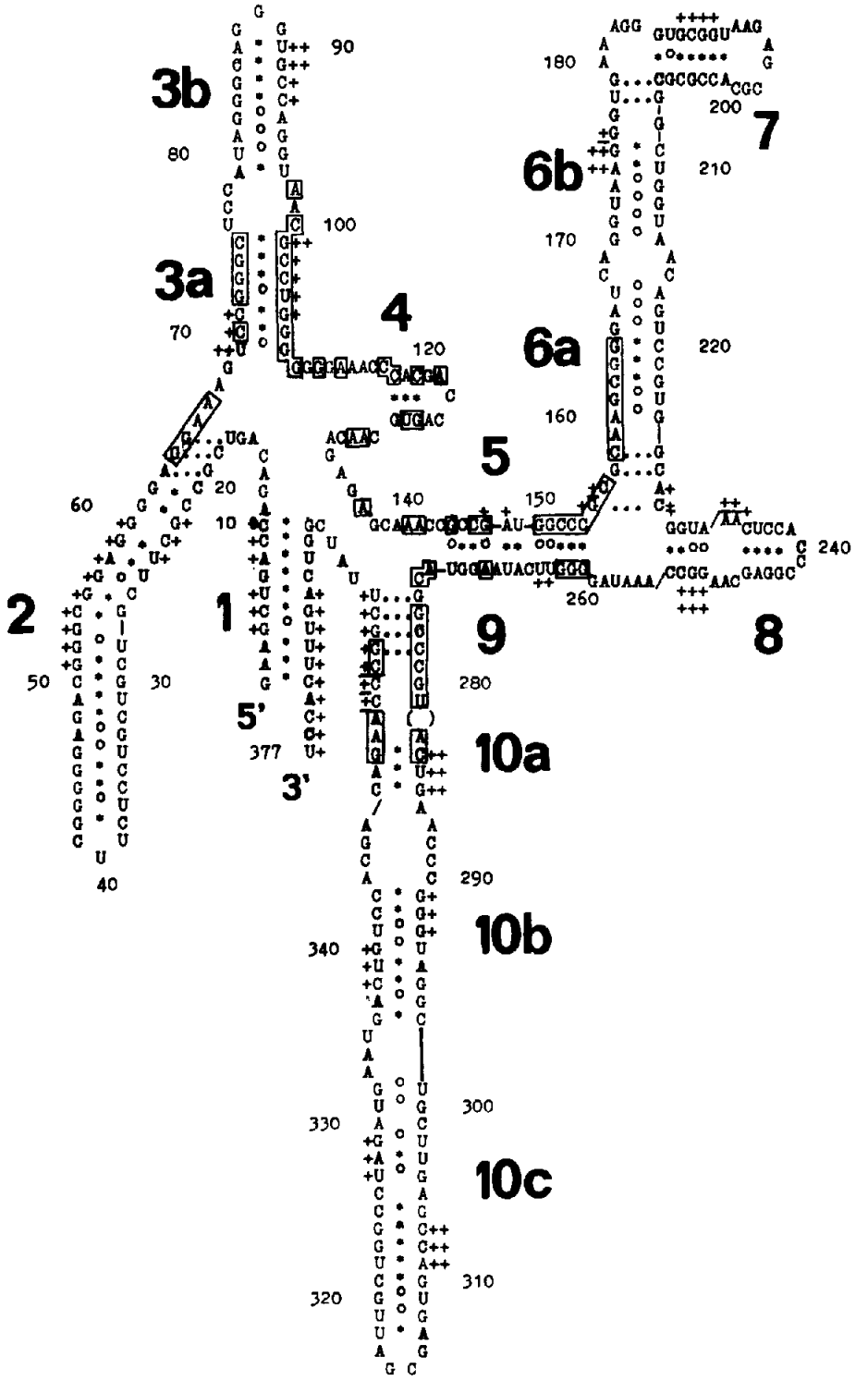
3b
3a
4
5
6b
6a
7
2
1
9
8
10a
10b
10c
5'
3'

Table 1

Homologous sequence stretches between M1 RNA and 16 S rRNA from *E. coli* used for the construction of the new M1 RNA secondary structure

Stretch no.	Nucleotide positions in		Identities/ nucleotides	Gap in alignment
	M1 RNA	16 S rRNA		
I	63– 75	557– 569	9/13	no
II	98–164	878– 942	41/66	1
III	260–283	1385–1409	14/24	1
IV	350–356	1491–1487	5/ 7	no

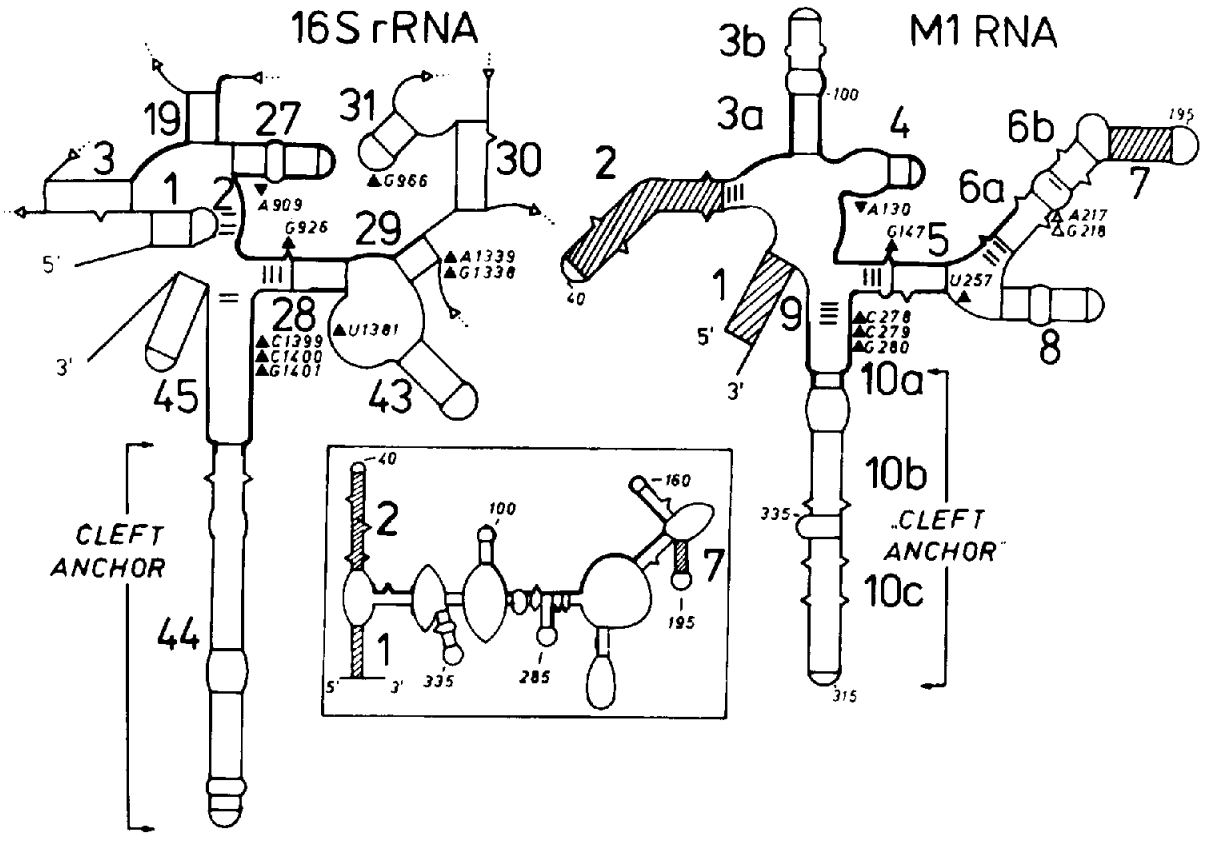

Fig.2. Schematic diagrams for the predicted M1 RNA secondary structure and the central part of the tRNA binding center of 16 S rRNA and its surroundings. The inset shows a scheme of the current model for M1 RNA proposed by Altman et al. [1]. The helices 1, 2 and 7 which are identical in the two M1 RNA models are hatched. Thick lines indicate homologous parts of the primary and/or secondary structures of M1 RNA and 16 S rRNA (see also fig.1 and table 1). Note that the drawing for 16 S rRNA is a slightly modified version of the California group picture [7,11] to show (i) schematically a possible folding of its active conformation according to the data in [8], to indicate (ii) that the base pairs of helix 2 should be regarded as tertiary interactions and to indicate (iii) more clearly the structural similarities between the M1 RNA and 16 S rRNA models. The numbering of double-stranded parts in the 16 S rRNA is according to the proposal of the Berlin group [6]. Nucleotides which are protected from attack by chemical probes through mRNA-independent tRNA binding to 16 S rRNA according to [11] are indicated by filled triangles. Corresponding nucleotide positions in the M1 RNA model are marked similarly. Other indications are as for fig.1.

mRNA by tRNA binding are indicated (see [11] for details). It is remarkable that most of these protected nucleotides are found at corresponding locations in the M1 RNA secondary structure (fig.2).

On the basis of these results we suggest that the mature part of the precursor tRNA may be bound to the central structural domain of M1 RNA which is similar to the central part of the tRNA binding center of 16 S rRNA. Indeed, deletions and mutations in the postulated tRNA binding domain affect the activity of the M1 RNA [13,14].

The M1 RNA model shown in figs 1 and 2 also exhibits structural motifs which have no obvious counterparts in the 16 S rRNA model, e.g. the helix 1 (base paired 5'- and 3'-ends of M1 RNA), the helix 2 and the domain built by helices 6–8. From deletion and point mutation experiments Altman and his group concluded recently [13,14] that in M1 RNA there is a separate catalytic center that includes nucleotides 165–255. In the model predicted here for the M1 RNA (figs 1,2) the mentioned sequence region is folded in a discrete domain which contains the helices 6–8 and their adjacent loop regions.

Thus, it appears that there are two separate domains in the M1 RNA structure, possibly influencing each other, one for the catalytic activity and binding of the 5'-extension in the tRNA precursor and another for binding of the mature part.

ACKNOWLEDGEMENTS

I would like to thank Professor T.A. Rapoport for stimulating discussions and critical reading of the manuscript. I thank also Dr B. Gross for help in the interpretation of autoradiograms from the literature and Mr C. Jakobs for drawings.

REFERENCES

[1] Altman, S., Baer, M., Guerrier-Takada, C. and Vioque, A. (1986) Trends Biochem. Sci. 11, 515–518.
[2] Guerrier-Takada, E., Gardiner, K., Marsh, T., Pace, N. and Altman, S. (1983) Cell 35, 849–857.
[3] Cech, T.R. and Bass, B.L. (1986) Annu. Rev. Biochem. 55, 599–629.
[4] Guerrier-Takada, C. and Altman, S. (1984) Biochemistry 23, 6327–6334.
[5] Reed, R.E., Baer, M.F., Guerrier-Takada, C., Donis-Keller, H. and Altman, S. (1982) Cell 30, 627–636.
[6] Brimacombe, R. and Stiege, W. (1985) Biochem. J. 229, 1–17.
[7] Gutell, R.R., Weiser, B., Woese, C.R. and Noller, H.F. (1985) Prog. Nucleic Acids Res. Mol. Biol. 32, 155–216.
[8] Moazad, D., Van Stolk, B.J., Douthwaite, S. and Noller, H.F. (1986) J. Mol. Biol. 191, 483–493.
[9] Oakes, M.I., Clark, M.W., Henderson, E. and Lake, J.A. (1986) Proc. Natl. Acad. Sci. USA 83, 275–279.
[10] Prince, J.B., Taylor, B.H., Thurlow, D.L., Ofengand, J. and Zimmermann, R.A. (1982) Proc. Natl. Acad. Sci. USA 79, 5450–5454.
[11] Moazad, D. and Noller, H.F. (1986) Cell 47, 985–994.
[12] DeLaCruz, V.F., Lake, J.A., Simpson, A.M. and Simpson, L. (1985) Proc. Natl. Acad. Sci. USA 82, 1401–1405.
[13] Guerrier-Takada, C. and Altman, S. (1986) Cell 45, 177–183.
[14] Lawrence, N.P. and Altman, S. (1986) J. Mol. Biol. 191, 163–175.