

LETTER TO THE EDITOR

The cross-link from the upstream region of mRNA to ribosomal protein S7 is located in the C-terminal peptide: Experimental verification of a prediction from modeling studies

BARBARA GREUER,¹ BERND THIEDE,^{2,3} and RICHARD BRIMACOMBE¹

¹Max-Planck-Institut für Molekulare Genetik, 14195 Berlin, Germany

²Max-Delbrück-Centrum für Molekulare Medizin, 13125 Berlin, Germany

Keywords: 3D structure of 16S rRNA; decoding site of 30S subunit; ribosome structure; RNA–protein interactions

The recent rapid advances that have been made both in cryo-electron microscopy (cryo-EM) and X-ray crystallography of bacterial ribosomes or their subunits (e.g., Stark et al., 1997a, 1997b; Ban et al., 1998; Malhotra et al., 1998) have led to a correspondingly rapid advance in our understanding of the three-dimensional (3D) arrangement *in situ* of the ribosomal RNA and protein molecules. In 1997 we published a model for the 16S rRNA (Mueller & Brimacombe, 1997a), which was fitted to a cryo-EM reconstruction at 20 Å resolution of the *Escherichia coli* 70S ribosome carrying tRNAs at the ribosomal A and P sites (Stark et al., 1997a). Subsequently, on the basis of the available RNA–protein interaction data (Mueller & Brimacombe, 1997b), we were able to fit the structure of ribosomal protein S7 as determined by X-ray crystallography (Hosaka et al., 1997; Wimberly et al., 1997) into our model in such a way as to satisfy both the biochemical data and the electron density of the EM reconstruction (Tanaka et al., 1998). More recently, the 16S model has been refined to fit an EM reconstruction at 13 Å resolution (Brimacombe et al., 2000), the latter being itself a refinement of the published EM reconstruction at 18 Å (Stark et al., 1997b) of 70S ribosomes carrying an EF-Tu/tRNA ternary complex stalled with the antibiotic kirromycin. In the refined 16S model, only minor changes needed to be made in the arrangement of the rRNA region interacting with S7 and in the positioning of the protein itself.

It is known that protein S7 can be cross-linked from sites in the upstream region of mRNA, close to the P

site codon (Stade et al., 1989; Dontsova et al., 1991), as well as from sites in the anticodon loop of P site-bound tRNA (Wower et al., 1993; Döring et al., 1994). Our fitted structure (Tanaka et al., 1998) made the strong prediction that the region of S7 involved in these cross-links must be either in the β -sheet area of the protein (covering amino acids ~75–90) or at the extreme C-terminus (from amino acid ~145 within the C-terminal α -helix to the C-terminus itself at position 155). Here we demonstrate that the predominant cross-link from the upstream region of mRNA is indeed to the C-terminal region of the protein.

As in our previous studies (Dontsova et al., 1991), an mRNA analogue related to the cro-mRNA from λ -phage was prepared by T7 transcription from a suitable DNA template. The mRNA sequence was GGAAGGAGG UUGUAUGGACACCAAC·A₆G·A₆G·A₇, thus containing a strong Shine–Dalgarno sequence close to the 5' end, the cro-mRNA UUGU spacer sequence, an AUG initiator codon, and an A-rich 3' sequence; the latter was included to enable the mRNA–protein cross-linked complex to be isolated by binding to oligo(dT)-cellulose (see below). The T7 transcription was carried out using 4-thio-UTP in place of “normal” UTP, except that a small amount of ³²P-UTP was present to label the transcript. Our previous experiments (Dontsova et al., 1991) had demonstrated that the thio-U residues in the UUGU spacer sequence are effectively cross-linked to protein S7. Ribosomal complexes were prepared containing this mRNA together with tRNA_f^{Met}, and, after irradiation at 350 nm, the mRNA cross-linked to ribosomal protein was isolated by a series of sucrose gradients, first in the presence of 10 mM magnesium to purify the 70S ribosomal complex, then in the presence of 0.3 mM magnesium to separate the 30S and 50S subunits, and finally in the presence of sodium dodecyl sulfate (SDS)

Reprint requests to: Richard Brimacombe, Max-Planck-Institut für Molekulare Genetik, Ihnestrasse 73, 14195 Berlin, Germany.

³Present address: WITA GmbH, Wartheinstrasse 21, 14513 Teltow, Germany.

to separate the 16S rRNA from ribosomal proteins (Stade et al., 1989). In this last gradient, the mRNA–protein cross-linked material runs with free ribosomal protein at the top of the gradient, and was separated from the latter by binding to oligo(dT)-cellulose as just mentioned (see Materials and Methods for details).

The experimental analysis of RNA–protein cross-link sites within the protein moiety has proved to be a surprisingly difficult problem. Urlaub et al. (1997) and Thiede et al. (1998) were able to precisely localize a number of RNA–protein sites at the nucleotide/amino acid level from 30S and 50S ribosomal subunits, respectively, cross-linked with the reagent 2-iminothiolane. At the same time, however, it was evident from the number of ribosomal proteins known to be cross-linked to the rRNA by this reagent (e.g., Osswald et al., 1990) that these localizations involved only a minority of the cross-linked products, many of the cross-linked oligonucleotide–oligopeptide intermediates being lost during the analyses. It soon became clear that oligonucleotide–oligopeptides derived from the S7-mRNA cross-linked complex belonged to this latter category. Using a variety of different protease and nuclease digestion conditions, and several different high-performance liquid chromatography (HPLC) separation systems, we were consistently unable to isolate a cross-linked oligonucleotide–oligopeptide product; the cross-linked ³²P-radioactivity either remained bound to the HPLC column or else eluted as a broad smear across the entire HPLC gradient. The use of thin-layer chromatographic systems was similarly unsuccessful, and furthermore, in our hands the oligopeptide–oligonucleotide complex would not “fly” in attempts to analyze it by MALDI mass spectrometry. The problem could possibly be a consequence of the very A-rich sequence in the mRNA, but for this there was no alternative. When mRNA with a shorter A-rich 3′ sequence was tried, it would only bind to the oligo(dT)-cellulose at temperatures below 15 °C. This in turn caused precipitation of SDS together with non-cross-linked protein onto the oligo(dT)-cellulose, thereby rendering the separation step ineffective.

Accordingly, we resorted to the following simple strategy. It is well known that protein S7 has a variable C-terminal region in different strains of *E. coli* (Reinbolt et al., 1978), the protein being 23 amino acids longer in strain K than in strain B. If the mRNA cross-link to S7 does lie, as predicted, at the C-terminus of the protein, then this should be immediately apparent from a comparison of the mRNA–oligopeptides that are released by protease digestion of the cross-linked mRNA–S7 complex from the two respective *E. coli* strains. Both strains carry a lysine residue at position 147 (Reinbolt et al., 1978), and, whereas the B strain terminates at residue 154, the K strain terminates at residue 177, with an intermediate lysine residue at position 169. Thus, digestion of the protein with Lys-C should yield a C-terminal peptide of 7 amino acids in the case of the

B strain (with the sequence AFAHYRW), with the corresponding peptide from the K strain being 22 amino acids long (AFAHYRWLSLRSFSHQAGASSK). Parallel cross-linking experiments were therefore made with the two *E. coli* strains and the results are illustrated in Figure 1.

After the oligo(dT)-cellulose separation step discussed above, the cross-linked mRNA complexes from both strains were applied to a 10% polyacrylamide gel containing SDS and urea (Fig. 1A). In each case the gel shows the presence of some free mRNA, together with two bands corresponding to mRNA–protein cross-linked complexes. These bands were extracted from the gel and aliquots were subjected to immunological analysis as in Dontsova et al. (1991). Whereas the lower, weaker band contained proteins S18 and S21 (Fig. 1C, and cf. Dontsova et al., 1991), the upper, stronger band corresponded to protein S7. In the sample from *E. coli* strain K (Fig. 1A) the S7 band is more intense (perhaps indicating a somewhat modified interaction with the mRNA) and, as expected, has a slightly slower mobility in the gel than its counterpart from strain B. Aliquots of the S7-mRNA complexes were also analyzed by ribonuclease T₁ fingerprinting, to check the distribution of the cross-linking between the four thio-U residues in the spacer region of the mRNA. As already noted, the T7 transcription of the mRNA was made in the presence of small amounts of ³²P-UTP, with the result that T₁ digestion of the mRNA should yield the radioactive oligonucleotides Gp, UUGp, and UAUGp (cf. Dontsova et al., 1991). Two-dimensional thin-layer chromatography of the digests from the cross-linked S7-mRNA complexes showed that all three of these radioactive spots were present, indicating that the cross-linking is distributed among the four thio-U residues; an example is shown in Figure 1D. This result is somewhat different from that observed by Dontsova et al. (1991), where the cross-linking to S7 was predominantly (although not exclusively) from the UUG sequence.

The isolated S7-mRNA complexes were digested with Lys-C (see Materials and Methods) and again bound to oligo(dT)-cellulose so as to remove the protease and non-cross-linked peptide material. The samples were then again applied to a 10% polyacrylamide gel, as above, giving the result shown in Figure 1B. Each sample shows two bands on the gel, the weaker, upper band (“P1”) being identical for strain B and strain K, indicating that this band does not arise from the C-terminal region of the protein. We were not, however, able to establish the identity of this weaker cross-link site. In contrast, the lower and stronger band (“P2B”, “P2K”) shows a distinctly different mobility between the two strains, with P2B running slightly behind the free mRNA in Figure 1B, and P2K running considerably more slowly, consistent with the predicted amino acid content in each case (7 amino acids in P2B and 22 in P2K, as discussed above). [The proportionately small amount

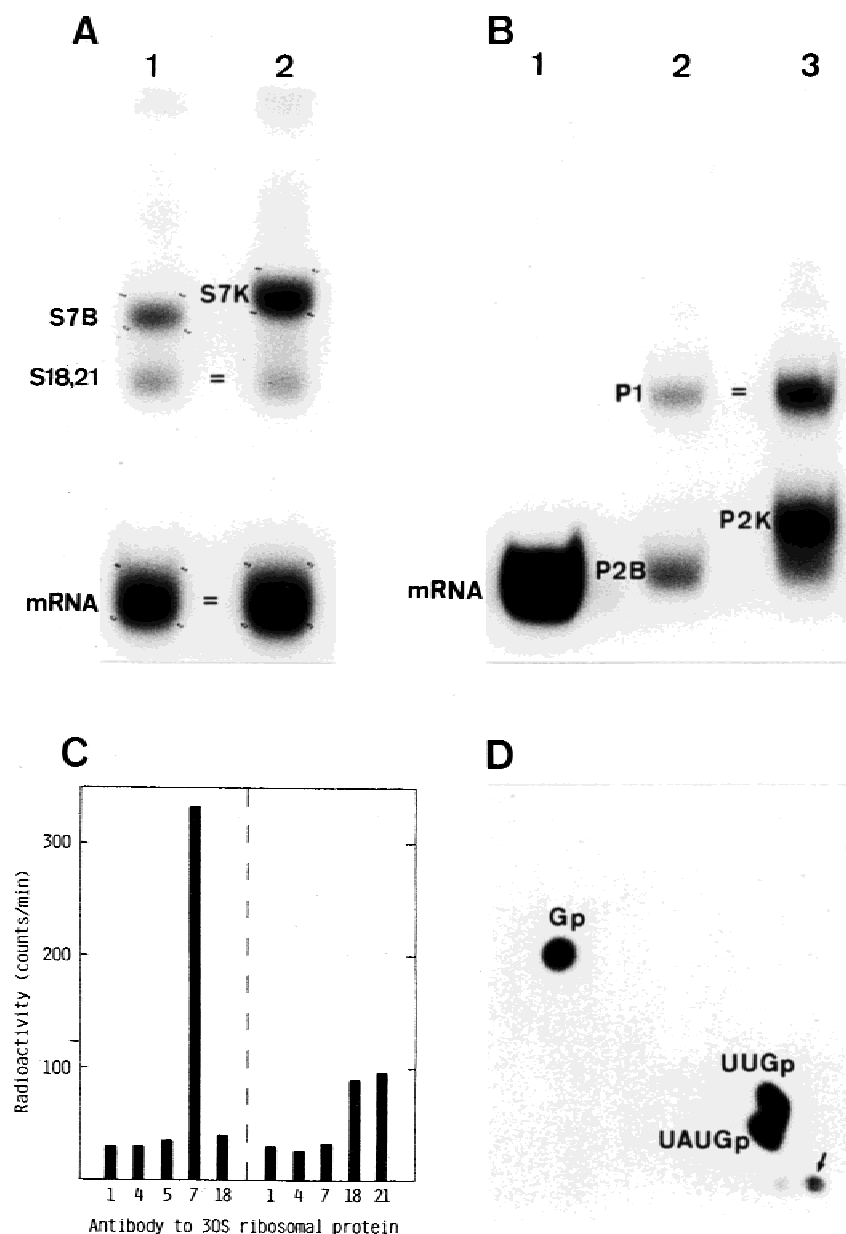


FIGURE 1. Separation and analysis of mRNA–protein cross-linked products. **A:** Autoradiogram of the separation of mRNA and mRNA–protein complexes on a 10% polyacrylamide gel. Lane 1: Sample from *E. coli* strain B. Lane 2: Sample from *E. coli* strain K. The free mRNA and cross-linked complexes containing proteins S7 (B or K), S18, and S21 are marked. **B:** Autoradiogram of the separation of S7–mRNA complexes after digestion with Lys-C on a 10% polyacrylamide gel. Lane 1: Free mRNA marker. Lane 2: Digested complex from *E. coli* strain B. Lane 3: Digested complex from *E. coli* strain K. The digestion products with the internal peptide (P1) and the C-terminal peptides (P2B, P2K) are marked. **C:** Immunological analysis of mRNA–protein complexes, using agarose-immobilized antibodies. The left half of the diagram shows analyses of the S7–mRNA complex (tested against anti-S1, S4, S5, S7, and S18), and the right half shows analyses of the S18, S21–mRNA complexes (tested against anti-S1, S4, S7, S18, and S21). **D:** Ribonuclease T₁ fingerprint of the S7–mRNA complex, by two-dimensional thin-layer chromatography. Direction of the first dimension is from right to left and that of the second from bottom to top. The arrow marks the sample application point, and the positions of the oligonucleotide digestion products are marked (see text).

of material from the K strain sample (Fig. 1B, lane 3) running with a mobility similar to that of P2B (Fig. 1B, lane 2) could arise either from a cross-contamination between the two strains or from a fortuitous partial degradation of peptide P2K.] We conclude that the principal site of cross-linking lies within the C-terminal region of protein S7. In the B strain the site must be within the sequence AFAHYRW (positions 148–154), and in the K strain the same site is likely to be involved, although we cannot, of course, exclude that here the cross-link lies within the C-terminal extension (up to position 169). We attempted to perform N-terminal sequencing of the peptide–oligonucleotide complexes isolated from gels such as that of Figure 1B, and, although the results showed a poor signal-to-noise ratio, we were

able to identify the sequence AFA at the N-terminus of both peptides P2B and P2K (data not shown).

Figure 2 shows the position of protein S7 in the refined 16S rRNA model that was fitted to the 13 Å cryo-EM reconstruction mentioned above. Figure 2A includes the minimum binding site for S7 on the 16S rRNA (Dragon & Brakier-Gingras, 1993), the Fe(II)-EDTA foot-print sites (Powers & Noller, 1995), the cross-link site to 16S rRNA from position –7 of the mRNA (Rinke-Appel et al., 1994), and the two previously identified cross-links between S7 and the 16S rRNA (Urlaub et al., 1997), as well as the S7–mRNA cross-link described here. The latter is indicated by the three highlighted U-residues in the UUGU spacer sequence of the mRNA and by the arrow pointing to the C-terminal

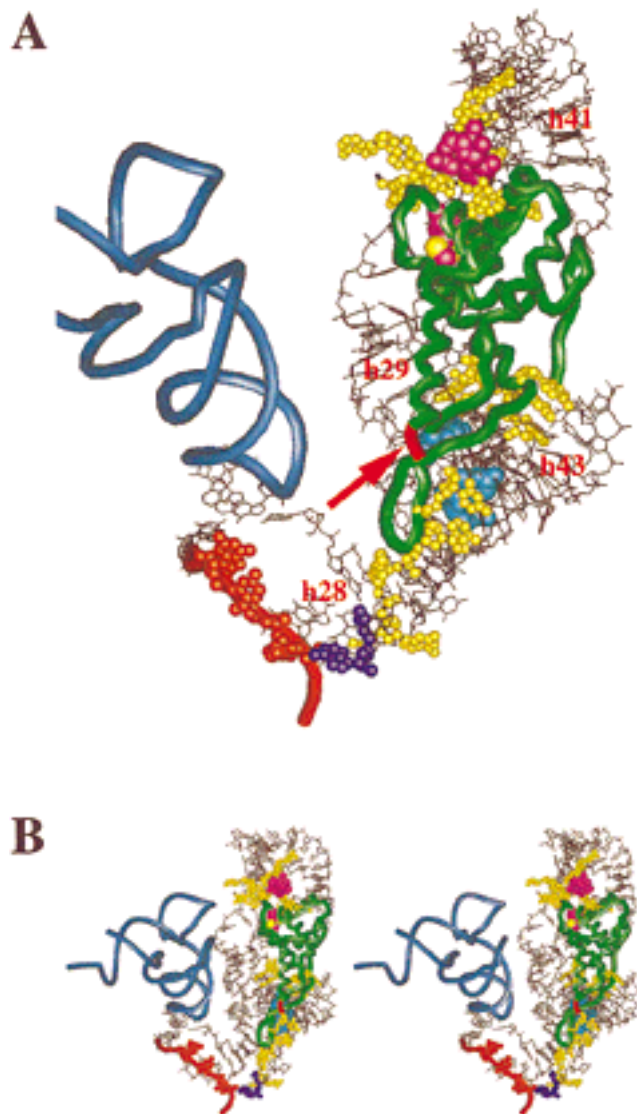


FIGURE 2. The location of protein S7 and its binding region on the 16S rRNA in relation to the positions of tRNA and mRNA. **A:** The biochemical data. Elements of the 16S rRNA comprising the minimal binding site for S7 on the 16S rRNA are shown as black wire-frame structures, together with protein S7 as a green backbone tube (as in Tanaka et al., 1998). Fe(II)-EDTA footprint sites for S7 are denoted by the yellow nucleotides and the minimum binding site for S7 has been expanded to include 7 bp of helix 41, so as to incorporate the Fe(II)-EDTA footprint sites in the latter helix. The cross-link site on 16S rRNA (position 1360) from position -7 of the mRNA is shown by the dark purple nucleotide at the lower left extremity of helix 43. Cross-links between S7 and the 16S rRNA (Met-115 to nt 1240, and Lys-75 to nt 1378) are denoted by the pink or light blue CPK amino acids and nucleotides, respectively. (The sulphur atom in Met-115 is yellow.) The mRNA (from position -6 to $+3$) is shown as a red backbone tube, with the U residues in the UUGU spacer region added as red nucleotides. The cross-link from these U residues to protein S7 is indicated by the red arrow pointing to the red C-terminal backbone residue (position 147) in the crystal structure (the eight C-terminal amino acids of the protein not being included in the latter). The P site tRNA (in part) is represented by the blue-grey backbone tube, with the codon-anticodon helix shown in wire-frame format. **B:** Stereo view of the same structure, from a different angle. See text for references.

position 147 of the crystal structure (Hosaka et al., 1997); it should be noted that the 8 amino acids at the extreme C-terminus of the protein (where the cross-link site to the mRNA is located) are disordered and therefore not included in the crystal structure (Hosaka et al., 1997). Figure 2B is a stereo view of the same structural elements, from a slightly different angle. It is perhaps surprising that the variable C-terminal region of S7 lies closest to the functionally important decoding site area. Nonetheless, with three defined cross-link sites within the protein (two to 16S rRNA and one to mRNA; Fig. 2), S7 is now the most firmly “biochemically anchored” protein in the *E. coli* 70S ribosome. Furthermore, to the best of our knowledge, the S7-mRNA cross-link described here represents the first time that an RNA-protein contact could be predicted at the peptide/oligonucleotide level and experimentally confirmed.

MATERIALS AND METHODS

Our procedures for the preparation of ^{32}P -labeled mRNA containing 4-thio uridine, binding of the mRNA to ribosomes, irradiation at 350 nm, and separation of the cross-linked products on sucrose gradients have all been described previously (Stade et al., 1989; Dontsova et al., 1992). The methods for the immunological analysis of the cross-linked proteins and for ribonuclease T_1 fingerprinting of the cross-linked mRNA have also been described (Dontsova et al., 1991). Here, ribosomal complexes from *E. coli* strain B (MRE 600) or strain K (CAN20-12E) were prepared on a larger scale (4,500 pmol 70S ribosomes, with 9,000 pmol each of mRNA and tRNA^{Met}), and after UV irradiation of the reaction mixtures the separation procedure was carried out up to the third sucrose gradient step in SDS to dissociate 16S rRNA from ribosomal protein (Stade et al., 1989; see text). The protein fraction from this last gradient, which also contains any remaining free mRNA as well as the mRNA-protein cross-linked complexes, was applied in a buffer containing 500 mM NaCl, 10 mM Tris-HCl, pH 7.8, 0.1% SDS, and 1 mM EDTA to a slurry of 50 mg oligo(dT)-cellulose (Type 7, Pharmacia) that had been pre-washed in the same buffer. After gentle shaking for 1.5 h at room temperature, the supernatant was removed and the oligo(dT)-cellulose washed three times for 15 min with the buffer. The mRNA and mRNA-protein complexes were then eluted by warming the oligo(dT)-cellulose to 55 °C for 15 min with three successive aliquots (100 μL) of a buffer containing 10 mM Tris-HCl, pH 7.8, 0.1% SDS, and 1 mM EDTA. A final elution step was made at 60 °C for 15 min with a further 100 μL of this buffer.

The eluates from the oligo(dT)-cellulose were combined, precipitated with ethanol, taken up in a small volume of 0.5% SDS, and then applied to a 10% polyacrylamide gel containing 0.1% SDS and 7 M urea. Bands corresponding to cross-linked mRNA-protein complexes (cf. Fig. 1A) were extracted from the gel in 10 mM Tris-HCl, pH 7.8, 0.1% SDS, and 1 mM EDTA and precipitated with ethanol. After removal of aliquots for immunological analysis and ribonuclease T_1 fingerprinting (see Fig. 1C,D) the S7-mRNA complex (~ 10 pmol, in 30 μL of 25 mM Tris-HCl, pH 8.5, 0.1% SDS, 1 mM EDTA) was treated with 15 μL of Lys-C (Boehringer-Mannheim; freshly

dissolved to 0.1 $\mu\text{g}/\mu\text{L}$ in 50 mM HEPES-KOH, pH 8.0, 10 mM EDTA, 5 mg/mL raffinose). The protease digestion was allowed to proceed for 16–20 h at 32 °C. The cross-linked mRNA–peptides were separated from protease and from non-cross-linked peptides by binding to oligo(dT)–cellulose as described above, and the eluted products were again applied to a 10% polyacrylamide gel (cf. Fig. 1B).

ACKNOWLEDGMENTS

We thank Dr. Knud Nierhaus for providing us with *E. coli* strain CAN20-12E, and Mr. Roman Pawlik for performing the N-terminal sequence analyses.

Received July 19, 1999; returned for revision August 9, 1999; revised manuscript received September 7, 1999

REFERENCES

- Ban N, Freeborn B, Nissen P, Penczek P, Grassucci RA, Sweet R, Frank J, Moore PB, Steitz TA. 1998. A 9 Å resolution X-ray crystallographic map of the large ribosomal subunit. *Cell* 93:1105–1115.
- Brimacombe R, Greuer B, Mueller F, Osswald M, Rinke-Appel J, Sommer I. 2000. Three-dimensional organization of the bacterial ribosome and its subunits; the transition from low-resolution models to high-resolution structures. In: Garrett RA, Douthwaite SR, Liljas A, Matheson AT, Moore PB, Noller HT, eds. *The ribosome: Structure, function, antibiotics and cellular interactions*. Washington, DC: ASM Press.
- Dontsova O, Dokudovskaya S, Kopylov A, Bogdanov A, Rinke-Appel J, Jünke N, Brimacombe R. 1992. Three widely separated positions in the 16S RNA lie in or close to the ribosomal decoding region: A site-directed cross-linking study with mRNA analogues. *EMBO J* 11:3105–3116.
- Dontsova O, Kopylov A, Brimacombe R. 1991. The location of mRNA in the ribosomal 30S initiation complex: Site-directed cross-linking of mRNA analogues carrying several photoreactive labels simultaneously on either side of the AUG start codon. *EMBO J* 10:2613–2620.
- Döring T, Mitchell P, Osswald M, Bochkariov D, Brimacombe R. 1994. The decoding region of 16S RNA: A cross-linking study of the ribosomal A, P, and E sites using tRNA derivatized at position 32 in the anticodon loop. *EMBO J* 13:2677–2685.
- Dragon F, Brakier-Gingras L. 1993. Interaction of *E. coli* ribosomal protein S7 with 16S rRNA. *Nucleic Acids Res* 21:1199–1203.
- Hosaka H, Nakagawa A, Tanaka I, Harada N, Sano K, Kimura M, Yao M, Wakatsuki S. 1997. Ribosomal protein S7: A new RNA-binding motif with structural similarities to a DNA architectural factor. *Structure* 5:1199–1208.
- Malhotra A, Penczek P, Agrawal RK, Gabashvili IS, Grassucci RA, Jünemann R, Burkhardt N, Nierhaus KH, Frank J. 1998. *E. coli* 70S ribosome at 15 Å resolution by cryo-electron microscopy: Localization of fMet-tRNA^{fMet} and fitting of L1 protein. *J Mol Biol* 280:103–116.
- Mueller F, Brimacombe R. 1997a. A new model for the three-dimensional folding of *E. coli* 16S ribosomal RNA. I. Fitting the RNA to a 3D electron microscopic map at 20 Å. *J Mol Biol* 271:524–544.
- Mueller F, Brimacombe R. 1997b. A new model for the three-dimensional folding of *E. coli* 16S ribosomal RNA. II. The RNA–protein interaction data. *J Mol Biol* 271:545–565.
- Osswald M, Greuer B, Brimacombe R. 1990. Localization of a series of RNA–protein cross-link sites in the 23S and 5S ribosomal RNA from *E. coli*, induced by treatment of 50S subunits with three different bifunctional reagents. *Nucleic Acids Res* 18:6755–6760.
- Powers T, Noller HF. 1995. Hydroxyl radical footprinting of ribosomal proteins on 16S rRNA. *RNA* 1:194–209.
- Reinbolt J, Tritsch D, Wittmann-Liebold B. 1978. The primary structure of ribosomal protein S7 from *E. coli* strains K and B. *FEBS Lett* 91:297–301.
- Rinke-Appel J, Jünke N, Brimacombe R, Lavrik I, Dokudovskaya S, Dontsova O, Bogdanov A. 1994. Contacts between 16S ribosomal RNA and mRNA, within the spacer region separating the AUG initiator codon and the Shine-Dalgarno sequence: A site-directed cross-linking study. *Nucleic Acids Res* 22:3018–3025.
- Stade K, Rinke-Appel J, Brimacombe R. 1989. Site-directed cross-linking of mRNA analogues to the *E. coli* ribosome: Identification of 30S ribosomal components that can be cross-linked to the mRNA at various points 5' with respect to the decoding site. *Nucleic Acids Res* 17:9889–9908.
- Stark H, Orlova EV, Rinke-Appel J, Jünke N, Mueller F, Rodnina M, Wintermeyer W, Brimacombe R, van Heel M. 1997a. Arrangement of tRNAs in pre- and posttranslocational ribosomes revealed by electron cryomicroscopy. *Cell* 88:19–28.
- Stark H, Rodnina MV, Rinke-Appel J, Brimacombe R, Wintermeyer W, van Heel M. 1997b. Visualization of elongation factor Tu on the *E. coli* ribosome. *Nature* 389:403–406.
- Tanaka I, Nakagawa A, Hosaka H, Wakatsuki S, Mueller F, Brimacombe R. 1998. Matching the crystallographic structure of ribosomal protein S7 to a three-dimensional model of the 16S ribosomal RNA. *RNA* 4:542–550.
- Thiede B, Urlaub H, Neubauer H, Grelle G, Wittmann-Liebold B. 1998. Precise determination of RNA–protein contact sites in the 50S ribosomal subunit of *E. coli*. *Biochem J* 334:39–42.
- Urlaub H, Thiede B, Müller EC, Brimacombe R, Wittmann-Liebold B. 1997. Identification and sequence analysis of contact sites between ribosomal proteins and rRNA in *E. coli* 30S subunits by a new approach using matrix-assisted laser desorption/ionization-mass spectrometry combined with N-terminal microsequencing. *J Biol Chem* 272:14547–14555.
- Wimberly BT, White SW, Ramakrishnan V. 1997. The structure of ribosomal protein S7 at Å reveals a β -hairpin motif that binds double-strand nucleic acids. *Structure* 5:1187–1198.
- Wower J, Sylvers LA, Rosen KV, Hixson SS, Zimmermann RA. 1993. A model for the tRNA binding sites on the *E. coli* ribosome. In: Nierhaus KH, Franceschi F, Subramanian AR, Erdmann VA, Wittmann-Liebold B, eds. *The translational apparatus: Structure, function, regulation, evolution*. New York: Plenum Press. pp 455–464.