RNA-protein cross-linking in *Escherichia coli* ribosomal subunits: localization of sites on 16S RNA which are cross-linked to proteins S17 and S21 by treatment with 2-iminothiolane

Apostolos Kyriatsoulis*, Peter Maly†, Barbara Greuer, Richard Brimacombe§, Georg Stöffler¹, Ronald Frank² and Helmut Blöcker²

Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, D-1000 Berlin-Dahlem, FRG, ¹Institut für Mikrobiologie, Medizinische Fakultät, Universität Innsbruck, Fritz-Pregl-Str. 3, A-6020, Innsbruck, Austria and ²Gesellschaft für Biotechnologische Forschung, Abteilung DNA-Synthese D-3000 Braunschweig-Stöckheim, FRG

Received 29 November 1985; Accepted 14 January 1986

SUMMARY

Treatment of *E. coli* ribosomal subunits with 2-iminothiolane coupled with mild ultraviolet irradiation leads to the formation of a large number of RNA-protein cross-links. In the case of the 30S subunit, a number of sites on 16S RNA that are cross-linked to proteins S7 and S8 by this procedure have already been identified (see ref. 6). Here, by using new or modified techniques for the partial digestion of the RNA and the subsequent isolation of the cross-linked RNA-protein complexes, three new iminothiolane cross-links have been localized: Protein S17 is cross-linked to the 16S RNA within an oligonucleotide encompassing positions 629-633, and protein S21 is cross-linked to two sites within oligonucleotides encompassing positions 723-724 and positions 1531-1542 (the 3'-end of the 16S RNA).

INTRODUCTION

Although a great deal is known about the topological arrangement within the *Escherichia coli* ribosome of the ribosomal proteins on the one hand and the ribosomal RNA on the other (see refs. 1-3 for review), there is still relatively little information available concerning the contacts between protein and RNA. The classical "binding site" approach (see ref. 4 for review) is severely limited by the fact that only a few ribosomal proteins are capable of binding specifically to the ribosomal RNA, and still fewer yield "binding sites" of a size which is useful for detailed topographical studies. The alternative approach - the application of RNA-protein cross-linking techniques - has also made disappointingly slow progress, in this case largely as a result of the technical problems involved in the separation and analysis of the cross-linked products. In our previous cross-linking studies (5,6), RNA-protein cross-links were induced in the ribosomal subunits by treatment with 2-iminothiolane (7).
followed by a brief ultraviolet irradiation. After removal of non-cross-linked protein by sucrose gradient centrifugation, the RNA was fragmented by partial nuclease digestion, and the cross-linked RNA-protein complexes were separated in a specially developed two-dimensional gel electrophoresis system (8). The sites of cross-linking on the RNA were then determined by classical fingerprint analysis. Using this method we were able to identify six RNA-protein cross-link sites in the 50S ribosomal subunit from a single cross-linking mixture (5), and correspondingly five sites in the 30S subunit (6).

These experiments were the first in which a simultaneous analysis of several protein cross-linking sites on the RNA could be made, but at the same time it was clear that in both cases (5, 6) many more cross-links were present which had escaped identification, either as a result of the selective nature of the partial digestion conditions, or due to insufficient resolution on the two-dimensional gels. In this paper we describe a further series of experiments with the 30S subunit, in which we have been able to locate three of these "missing" iminothiolane cross-links on the 16S RNA. One of the cross-links is to protein S17, and the other two are to protein S21. The experiments reported here represent part of an intensive search for improved methodology for the isolation and analysis of the cross-linked RNA-protein complexes.

MATERIALS AND METHODS

Preparation of cross-linked 30S subunits, and isolation of cross-linked RNA-protein complexes. $^{32}$P-labelled 30S subunits from E. coli strain MRE 600 (ca. 5 - 8 A$_{260}$ units, 1 - 2 x 10$^9$ counts/min total) were prepared by the method of Stiege et al (9). Cross-linking with 2-iminothiolane (20 mM) was carried out as described previously (5), with the exception that the subsequent ultraviolet irradiation procedure was reduced from 3 to 2 min. Non-cross-linked protein was removed by centrifugation on sucrose gradients containing dodecyl sulphate as before (5), and the RNA plus cross-linked RNA-protein complexes isolated from these gradients was submitted to one of two partial digestion and separation procedures, as follows:
Procedure (a): The samples were digested with ribonuclease T₁ (1.5 - 4.0 enzyme units per A₂₆₀ unit RNA) exactly as before (6), and were applied to the Triton X-100/dodecyl sulphate two-dimensional gel system previously described (5,8), with the exceptions that the first dimension gel strips were not rinsed before polymerising into the slab gel used for the second dimension, and the second dimension dodecyl sulphate gels (13% polyacrylamide) included a 6% spacer gel (10). The first dimension gel strips were polymerised into this spacer gel.

Procedure (b): The samples were digested with ribonuclease H (cf. 11-13) in the presence of the deoxyoligonucleotides d-(CGTTGC) and d-(CCGCCT), in a buffer containing 40 mM Tris-HCl pH 7.8, 50 mM KCl, 10 mM MgCl₂, 0.1 mM EDTA, 0.1 mM dithiothreitol and 0.2% Triton X-100 (13, cf. 11). The reaction mixtures (typically ca 350 µl) contained 15 A₂₆₀ units/ml RNA plus cross-linked complexes and 5 A₂₆₀ units/ml of each hexanucleotide, and were incubated at 4° for 16 hr (13). The oligonucleotides were synthesized in solution following established triester procedures (14), and were purified by ion exchange chromatography under denaturing conditions (15).

The ribonuclease H digestion was stopped by addition of EDTA to a final concentration of 10 mM, and the hydrolysate was diluted 10-fold with a buffer containing 10 mM Tris-HCl pH 7.8, 1 mM EDTA, 300 mM NaCl. The solution was warmed to 60° and passed slowly through a glass filter (Whatman GF-C, 25 mm dia.), which was then washed with the same buffer. The filtration process has the property of separating free RNA from RNA-protein complexes (cf. 16). The free RNA fragments (representing usually over 95% of the radioactive material) could if required be recovered from the filtrate by ethanol precipitation, whereas the RNA-protein complexes were recovered by eluting the filter with 5 successive aliquots (75 µl) of a buffer containing 10 mM Tris-HCl pH 7.8, 1% sodium dodecyl sulphate, 0.1% 2-mercaptoethanol. The combined eluates were made 100 mM in sodium acetate and the RNA-protein complexes precipitated with ethanol to concentrate them. The precipitate was redissolved in 40 µl of 50 mM Tris-citrate pH 8.8, 2.5 mM EDTA, 2% sodium dodecyl sulphate and 5% 2-mercaptoethanol, and after incubating for 5 min at 60° the
sample was loaded onto a 4 - 12% polyacrylamide gradient gel containing 7 M urea and dodecyl sulphate (gel recipe as in ref. 9, except that the dodecyl sulphate concentration was raised to 0.7%). The gel (40 cm long) was run until a bromophenol blue marker had travelled ca 24 cm, and then 1 cm wide x 12 cm long strips were cut out (from the upper and lower portions of the gel, respectively) and polymerised onto 7% second dimension gels in the system of Maxam and Gilbert (17). Before polymerising, the gel strips were washed for 5 min at 40° with reservoir buffer (17,9) containing 5% 2-mercaptoethanol. The gels were run until a bromophenol blue marker had travelled ca 10 cm.

Analysis of isolated RNA-protein cross-linked complexes. The 32P-labelled cross-linked complexes were located on the two-dimensional gels from procedure (a) and (b) (above) by autoradiography, and were extracted as previously described (8). Radioactivity at this stage was of the order of 5 - 10 x 10^3 counts/min per complex, and each sample was divided into two parts, for RNA and protein analysis, respectively. For the RNA analysis, each sample was fully digested with ribonuclease T1, followed by incubation with proteinase K, exactly as described (6), and the oligonucleotides released were analysed by "fingerprinting" on polyethyleneimine-cellulose plates (18). Secondary digestions of individual oligonucleotides from the fingerprints were made with ribonuclease A, again as described (8), and the data were fitted to the 16S RNA sequence of Brosius et al (19).

For the protein analysis, each sample was examined either by two-dimensional gel electrophoresis in the system of Mets and Bogorad (20) or by immuno double diffusion (Ouchterlony) using antibodies to individual proteins from the 30S subunit (21). In the former case the samples were mixed with 1 A260 unit of unlabelled 30S subunits and fully digested with ribonucleases A and T1 before application to the two-dimensional gel (20), as in ref. 6. After running, the gels were both stained and autoradiographed. In the latter case, a suitable amount of total protein from the 30S subunit in 3 mM Tris-HCl pH 7.8 was added to the samples in either 8 M urea or 0.1% sodium dodecyl sulphate, and the immuno double diffusion was carried out as described (22,5). After five days of diffusion at 4° the plates were washed for
RESULTS AND DISCUSSION

Isolation of cross-linked RNA-protein complexes. As indicated in the Introduction, the choice of partial nuclease digestion conditions and separation system are the two most crucial factors in the analysis of the complex mixture of cross-linked products generated by treatment of ribosomal subunits with a bifunctional reagent. Figure 1 shows schematically the separation of cross-linked RNA-protein complexes which we obtained by two-dimensional gel electrophoresis in our previous experiments (6) with iminothiolane-treated 30S subunits from E. coli. These experiments, in which the partial digestion of the cross-linked material was made with ribonuclease T₁, led to the identification of three cross-link sites for protein S8 and two for S7 (6), but a number of other complexes were also observed on the gels for which an unequivocal analysis was not obtained. In particular, two groups of complexes were reproducibly seen corresponding to proteins S19/S20/S21 and to S15/S16/S17 (Fig. 1), and the elucidation of these complexes is the subject of this paper.

The S15/S16/S17 complex will be considered first, as it proved possible to resolve this complex simply by making some minor modifications to the two-dimensional gel system, as described in Materials and Methods under "procedure (a)". In this modified system the pattern of spots was essentially the same as that in Fig. 1 (cf. ref. 6, data not shown), but the resolution was better in the "S15/S16/S17 region" (although at the same time it was less good in other regions of the gel). Analysis of these gels enabled a reproducible identification of the S15/S16/S17 complex to be made, and it proved to separate into two components; one of the components contained protein S17, whereas the other was an entirely unrelated complex containing protein S21. (It can be seen from Fig. 1 that the S15/S16/S17 complex lies effectively on the same "reverse diagonal" (cf. refs. 5,6,8) as the S19/S20/S21 complex, so the finding of complexes with S17 and S21 at almost the same position in the gel is reasonable if the attached RNA fragments are of similar length). The analysis of
both of these S17 and S21 cross-linked complexes is described in the appropriate section below, the S21 complex being referred to as "S21a".

In the case of the S19/S20/S21 complex, a clear cross-link site at positions 723-724 of the 16S RNA was previously observed (6), but the yield of the complex — under the ribonuclease T₁ partial digestion conditions used — was always too low to enable us to identify the protein unambiguously. An entirely different approach was therefore applied, which led to the identification of a cross-link site for protein S21 corresponding to the S19/S20/S21 complex of Fig. 1. This complex is referred to in the sections below as "S21b", and the experimental approach is described in Materials and Methods under "procedure (b)". Here, instead of making the partial digestions with ribonuclease T₁ as above, ribonuclease H together with defined hexadeoxynucleotides (12,13) was used, in order to generate a different spectrum of RNA fragments containing the cross-linked complexes. Oligodeoxynucleotides for this purpose were chosen on the basis of a computer search of the single-stranded regions of the 16S RNA secondary structure (23) as described in ref. 13, and, out of a number of oligonucleotides tested, the combination d-(CCGCCT)
and d-(CGTTGC) was found to yield a suitable distribution of ribonuclease H cuts in the 16S RNA. The digestions were carried out at 4° (cf. 13), and Triton X-100 was included in the reaction mixtures to improve the solubility of the fragmented RNA-protein complexes (see Materials and Methods). The digestion was dependent on the presence of both enzyme and deoxyoligonucleotides.

The ribonuclease H digests contain some relatively large RNA fragments, which disturb the separation of the cross-linked complexes in our two-dimensional gel system (cf. Fig. 1), and therefore the next step in the procedure was to remove most of the free RNA by the use of glass filters. Such filters have been shown to separate DNA from DNA-protein complexes (16), and preliminary experiments with double-labelled cross-linked mixtures showed that the filters behaved in the same way with RNA, i.e. free 3H-labelled RNA fragments passed through the filters, whereas [3H]RNA-[14C]protein cross-linked complexes were quantitatively retained. Non-specific adsorption of free RNA was reduced by conducting the filtration at 60° as described in Materials and Methods, and the complexes could subsequently be eluted from the filters in 80 - 90% yield with dodecyl sulphate. The procedure served to enrich the cross-linked RNA-protein complexes by about 20-fold, and the correspondingly enriched 32P-labelled complexes were separated by gel electrophoresis in a different system to that of Fig. 1, as outlined in the following paragraph.

Figure 2 (left) shows a typical separation on a 4-12% polyacrylamide gradient gel (see Materials and Methods, procedure "b") of the filtrate fraction (free RNA) and eluate fraction (RNA-protein complexes) from a 32P-labelled cross-linked 30S preparation after the glass filtration step. In preliminary experiments the distribution of proteins in the cross-linked complexes was monitored by soaking a gel strip from the eluate fraction (cf. Fig. 2, left) in a solution containing ribonucleases A and T1, and then polymerising the strip onto the top of a Laemmli-Favre (10) dodecyl sulphate gel; the released 32P-labelled oligonucleotide-protein complexes were separated in this gel and subsequently visualized by autoradiography. The results (data not shown) indicated that complexes corresponding to most
if not all of the 30S proteins were present in the eluate fraction, and this is the first time that we have observed in a partial nuclease digest a representative spectrum of the proteins cross-linked to RNA by a bi-functional reagent. However, the pat-
tern of products in the partial digests was extremely complex, and so far we have only been able to purify the smallest and fastest-running components of the mixture, which proved to contain complex S21b. This purification was achieved by a second gel electrophoretic step in which the gel strip containing the eluate fraction (Fig. 2, left) was polymerised onto a Maxam-Gilbert gel (17), as described in Materials and Methods. In this gel system fragments of free RNA still remaining after the glass filtration procedure run into the gel, whereas protein-RNA complexes remain at the origin. A typical separation is shown in Fig. 2 (right), in which a number of radioactive spots corresponding to cross-linked complexes can be seen; however, only the spots marked "S21b" led to unequivocal analyses. These complexes were extracted from the gel, and, as with the S17 and S21a complexes (above), were subjected to both protein identification and RNA fingerprinting.

**Analysis of isolated RNA-protein cross-linked complexes.** Ribonuclease T1 fingerprints of the complexes (S17, S21a and S21b) on polyethyleneimine cellulose thin-layer plates (18) are shown in Fig. 3, and examples of the corresponding protein analyses are illustrated in Fig. 4. The location of the cross-link sites
Figure 4: Protein analysis of cross-linked complexes. Left: A two-dimensional gel (cf. ref. 6) of complex S21b, after total digestion with ribonucleases A and T1 (see Materials and Methods). The Figure is an autoradiogram of the gel, with the positions of the principal stained protein spots indicated by the circles. The arrow denotes the direction in which the cross-linked radioactive complex is displaced from the corresponding free protein. Direction of electrophoresis is from left to right (first dimension) and top to bottom (second dimension). Right: Autoradiograms of immuno double diffusion tests made with the cross-linked complexes. In each case the centre well contained the complex, and selected antibodies to the individual proteins were placed in the peripheral wells as indicated. The proteins in the complexes in the top row were identified as S17 and S21, respectively, the latter from complex S21a. The complexes in the bottom row show respectively a mixture of S18 and S21 (from complex S21a, see text), and S21 alone (from complex S21b). The complexes, particularly those containing S21, tended to aggregate in the sample wells.

in the 16S RNA are summarized in Fig. 5.

Complexes S17 and S21a were barely separable from one another on the two-dimensional gels (see foregoing section), and as a result the isolated complexes were usually cross-contaminated to a slight extent. This cross-contamination can be seen in the respective fingerprints (Fig. 3), but it was nevertheless possible to make clearcut and reproducible analyses of the cross-linked oligonucleotides. In the case of the S17 complex, the RNA component extended from bases 593 to 645 of the 16S RNA (cf. Figs. 3 and 5), but the oligonucleotide AACUG (positions 629-633) was only present in trace amounts (Fig. 3) and therefore represents the site of cross-linking to the protein (cf. 5,6). The fact that small amounts of AACUG were found at all suggests that
the cross-link site may be heterogeneous, and that a further unidentified minor cross-link site may be present. The residual protein-oligonucleotide complex appears on the fingerprint as a pair of fast-moving oligopeptide-oligonucleotide fragments (arrowed in Fig. 3), as a consequence of the proteinase K treatment which routinely followed the ribonuclease T₁ digestion for the fingerprint analyses (see Materials and Methods). These fragments were subjected to secondary digestion with ribonuclease A, but in this particular case the secondary digestion did not yield any additional information regarding the cross-link site (in contrast to the S21a complex, below). In addition to the fingerprint analysis, aliquots of the complex were examined by immunodouble diffusion using antibodies to selected proteins from the 30S subunit. The subsequent autoradiography of the immunodiffusograms (Fig. 4) led to an unambiguous identification of the protein component as S17. Thus, protein S17 is cross-linked within
the oligonucleotide spanning positions 629-633.

The RNA component of complex S21a arose from the extreme 3'-end of the 16S RNA (cf. Figs. 3 and 5), covering positions 1490 (or sometimes 1498) to 1542. The 3'-terminal oligonucleotide AUCA\textsubscript{OH}CUCUAU (positions 1531-1542) did not appear in its usual position on the fingerprint, but ribonuclease A digestions of the fast-running oligopeptide-oligonucleotide fragments (arrowed in Fig. 3) showed in each case the presence of AU, AC, C and U, which indicates that the cross-link site was within this terminal oligonucleotide, as illustrated in Fig. 5. The immuno double diffusion test (Fig. 4) allowed an unambiguous identification of the protein component of the complex as S21, although in some samples S18 was seen in addition. However, since the latter protein was never found without S21 being present as well, we do not consider it proven that S18 is also cross-linked to the 3'-terminus. The finding of S21 at the 3'-terminus of the RNA is consistent with the earlier finding of Czernilofsky et al (24), using periodate oxidation of the 16S RNA.

In our previous analysis (6) of the S19/S20/S21 complex (Fig. 1), we found RNA fragments extending over positions 704-733, with UG at positions 723-724 as the cross-link site (cf. Fig. 5). Complex S21b, being the product of a ribonuclease H as opposed to a ribonuclease T\textsubscript{1} partial digestion, contained a different set of RNA fragments attached to the protein. The ribonuclease T\textsubscript{1} fingerprints showed that these extended over positions 689 to ca. 727 (cf. Fig. 3), with longer and shorter fragments being found as indicated in Fig. 5. At the 3'-end of this region there is a potential ribonuclease H cutting site due to perfect complementarity between AGGCCG (729-734) and the oligodeoxynucleotide d-(CCGCCT). The preferred cutting points for ribonuclease H are at the ends of the hybrid helices (12), and a cut at position 728-729 would be consistent with the observation that AAG (728-730) was never found in the fingerprints of complex S21b. In contrast, at the 5'-end of the complexes (positions 689-700) only partial complementarities exist with respect to either d-(CCGCCT) or d-(CGTTGC), and the occurrence of several "weak" ribonuclease H cuts in this region would account for the observed heterogeneity at the 5'-end of the
complexes. The fingerprints of the S21b complexes were extremely clean, which is important for the following argument.

In the longer S21b complexes (such as that shown in Fig. 3), CG (positions 689-690, 699-700 or 726-727) and UG (692-693 or 723-724) were both found, but in the shortest complex (positions 701-727, Fig. 5) no UG was observed at all, although CG was still present. The latter could not be the CG at position 699-700, since the (heterogeneous) 5'-end of the complex was represented by the oligonucleotides pCG (699-700) and pUAG (701-703); (ribonuclease H gives hydrolysis products with 5'-phosphate groups). Thus the CG observed must have been that at position 726-727, and since no UG was found, the UG at position 723-724 corresponds as before (6) to the cross-link site. No fast-running oligopeptide-oligonucleotide fragments from the cross-link site were seen on the fingerprints (cf. Fig. 3), which may be due to the rather insoluble nature of the residual protein-oligonucleotide complex in this instance. This insolubility was also reflected in the protein analysis of the complex; Fig. 4 (left) shows a two-dimensional protein gel from complex S21b, in which a rather smeared radioactive band can be seen corresponding to the protein-oligonucleotide residue. This band appears to emanate from proteins S19/S20/S21 as observed previously (6), and the subsequent immuno diffusion test (also shown in Fig. 4) reproducibly gave a positive reaction only with protein S21. Thus protein S21 is cross-linked both to positions 723-724, and to the 3'-terminal oligonucleotide of the 16S RNA.

CONCLUSIONS

It is clear from the results we have obtained so far (5,6, and this paper) that an ideal general method for the analysis of RNA-protein cross-link sites in ribosomal subunits has yet to be discovered. Whereas the partial digestion procedure with ribonuclease T₁ (5,6) leads to a relatively simple pattern of products (cf. Fig. 1), many cross-links are lost as a result of the selectivity of the enzyme. With ribonuclease H on the other hand, this loss is avoided, but the pattern of cross-linked hydrolysis products then becomes very complex (cf. Fig. 2). The complexity is at least partly due to the occurrence of imperfectly base-
paired hybrids between the RNA and the deoxyoligonucleotides at the temperature used (4°), which causes heterogeneous cutting by the ribonuclease H. The specificity can be improved by raising the temperature, but then an impractically large number of different oligonucleotides would have to be included in each reaction in order to generate RNA fragments of a suitably short length for the subsequent analysis. It seems likely that in the future we will have to continue to use a variety of different analytical approaches, as we have done here, in the expectation that each one will lead to the identification of a small sub-set of the cross-link sites generated by a particular cross-linking agent. Further development of the techniques described in this paper has already led to the preliminary identification of another group of cross-link sites involving proteins S3, S10, S17 and S18, which will be reported in due course.

Despite the technical problems, the catalogue of known RNA-protein cross-link sites in the 30S subunit is beginning to accumulate. The iminothiolane-induced cross-links now include three sites for protein S8 (within oligonucleotides 629-633, 651-654 and 593-597), two for S7 (1238-1240 and 1377-1388), two for S21 (723-724 and 1531-1542), and one for S17 (629-633). The new data show two striking features. First, the cross-link site to protein S17 (which has a "binding site" in the 5'-terminal region of the 16S RNA (4)) is to precisely the same oligonucleotide (629-633) as one of the S8 cross-links. This result is in good agreement with the recent finding by neutron scattering studies (25) as well as immuno electron microscopy (3) that S8 and S17 are very close neighbours in the 30S subunit. The second feature of interest is the finding of the cross-link site to protein S21 in the central region of the 16S RNA (positions 723-724). While the corresponding cross-link of S21 to the 3'-terminus of the RNA is consistent with a number of other findings (summarised in ref. 26), the cross-link to positions 723-724 represents an entirely new neighbourhood for protein S21 in the ribosome. This type of result is of obvious importance for model-building studies, and our detailed model of the 30S subunit (described as a preliminary report in ref. 27) will be presented elsewhere.
ACKNOWLEDGEMENTS

We thank M. Becker, W. Heikens, J. Hollatz and H. Krause for their help in the synthesis of oligodeoxynucleotides, Dr. M. Stöffler-Meilicke and R. Hasenbank for their help with the immuno-assays, and Dr. H. G. Wittmann for his continued support.

Present addresses: *I. Medizinische Klinik der Johannes Gutenberg Universität, Langenwitstr. 1, D-6500 Mainz, FRG and *Biochemisches Institut der Universität Zürich, Winterthurerstr. 190, CH-8057 Zürich, Switzerland

§To whom correspondence should be addressed

REFERENCES