Volume 15 Number 8 1987 Nucleic Acids Research

RNA-protein cross-linking in *Escherichia coli* 30S ribosomal subunits; determination of sites on 16S RNA that are cross-linked to proteins S3, S4, S5, S7, S8, S9, S11, S13, S19 and S21 by treatment with methyl p-azidophenyl acetimidate

Monika Osswald, Barbara Greuer, Richard Brimacombe*, Georg Stöffler[1], Hans Bäumert[2] and Hugo Fasold[2]

Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, D-1000 Berlin-Dahlem, FRG, [1]Institut für Mikrobiologie, Medizinische Fakultät, Universität Innsbruck, A-6020 Innsbruck, Austria and [2]Institut für Biochemie der Johann Wolfgang Goethe-Universität, D-6000 Frankfurt-Main, FRG

Received March 6, 1987; Accepted April 3, 1987

SUMMARY

RNA-protein cross-links were introduced into E. coli 30S ribosomal subunits by treatment with methyl p-azidophenyl acetimidate. After partial nuclease digestion of the RNA moiety, a number of cross-linked RNA-protein complexes were isolated by a new three-step procedure. Protein and RNA analysis of the individual complexes gave the following results: Proteins S3, S4, S5 and S8 are cross-linked to the 5'-terminal tetranucleotide of 16S RNA. S5 is also cross-linked to the 16S RNA within an oligonucleotide encompassing positions 559-561. Proteins S11, S9, S19 and S7 are cross-linked to 16S RNA within oligonucleotides encompassing positions 702-705, 1130-1131, 1223-1231 and 1238-1240, respectively. Protein S13 is cross-linked to an oligonucleotide encompassing positions 1337-1338, and is also involved in an anomalous cross-link within positions 189-191. Protein S21 is cross-linked to the 3'-terminal dodecanucleotide of the 16S RNA.

INTRODUCTION

For a number of years we have been engaged in the development of RNA-protein cross-linking techniques, as a tool for the investigation of topographical neighbourhoods between protein and RNA in the Escherichia coli ribosome (1-5). Our objective is to determine a network of cross-link sites on the ribosomal RNA corresponding to the various individual ribosomal proteins, and hence to build up a detailed picture of the distribution of the proteins on the RNA in situ in the ribosomal subunits. We have investigated a number of bifunctional cross-linking reagents for this purpose, including nitrogen mustard derivatives (1), methyl p-azidophenyl acetimidate (2,6) and 2-iminothiolane (3-5). In the case of the latter reagent, eight RNA-protein cross-link sites have been precisely identified on the 16S RNA (4,5), and six on the 23S RNA (3).

In our studies, the cross-links are introduced into the intact ribosomal subunits by treatment with the reagent, and this is followed by a partial nuclease digestion procedure in order to generate RNA fragments of a suitable size for the subsequent analysis of the individual cross-linked RNA-protein products. The spectrum of digestion products obtained is unavoidably very complex, and furthermore there is usually a strong tendency for aggregation to occur. In the past this has had the result that large amounts of potentially important cross-linking data have been lost during the course of the separation procedures. Recently however we have developed a new procedure (7), which has to a large extent overcome this problem. The RNA-protein complexes are isolated in three steps, involving (a) removal of non-cross-linked protein, (b) removal of non-cross-linked RNA fragments and (c) separation of the individual cross-linked products on a two-dimensional polyacrylamide gel system containing high salt concentrations to prevent aggregation. A further advance has been the development of a new and rapid immunological method (8) for identifying the proteins in the isolated cross-linked complexes. The application of these procedures has enabled us to increase the catalogue of identified RNA-protein cross-link sites on the 16S RNA from eight sites (4,5) to twenty-eight, within the space of a few months.

This paper is concerned with the documentation of a series of cross-link sites obtained using the new procedure with the reagent methyl p-azidophenyl acetimidate ("APAI", 2,6). A corresponding series of cross-links obtained with bis-(2-chloroethyl)-methylamine ("nitrogen mustard", cf. 1,9) is dealt with in a parallel manuscript (10). APAI is an imidoester, which reacts with lysine groups of the ribosomal proteins in aqueous solution at neutral pH. Subsequent photoactivation of the azide group leads to a cross-linking reaction, whereby 5 - 8% of the total protein becomes attached to the ribosomal RNA (2). We describe the identification of twelve APAI-induced RNA-protein cross-link sites on the *E. coli* 16S RNA, involving ten out of the twenty-one 30S ribosomal proteins.

MATERIALS AND METHODS

Preparation, cross-linking and partial digestion of ribosomal subunits: ^{32}P-labelled 30S subunits from E. coli strain MRE 600 (ca. 5 A_{260} units, 1.5 x 10^9 counts/min total) were prepared by the method of Stiege et al (11). The subunits, in 0.5 ml of 25 mM triethanolamine-HCl pH 7.8, 50 mM KCl, 5 mM magnesium acetate, 6 mM 2-mercaptoethanol, were treated in the dark with an equal volume of APAI (2 mg/ml, 9 mM) freshly dissolved in 100 mM triethanolamine-HCl pH 9, 50 mM KCl and 5 mM magnesium acetate (cf. ref. 2). This gives a solution with a final pH of 8.5. After incubating at 30° for 30 min, excess reagent was removed by ethanol precipitation, and the subunits were resuspended in 25 mM triethanolamine-HCl pH 7.8, 50 mM KCl, 5 mM magnesium acetate to a concentration of 5 A_{260} units/ml. This solution was subjected to ultraviolet irradiation for 2 min under the conditions previously described (2,7). In some experiments the magnesium concentration was reduced to 0.5 mM during both steps of the cross-linking procedure. After irradiation, the subunits were again precipitated with ethanol, and redissolved in a small volume (50 - 100 µl) of 10 mM Tris-HCl pH 7.8, 20 mM NH_4Cl, 1 mM magnesium acetate. The solution was incubated with a suitable quantitiy of cobra (Naja naja oxiana) venom nuclease (12) for 1 hr at 37° (7,9).

Separation of RNA-protein cross-linked complexes: Full details of this method are given in ref. 7; the following is a brief resumé of the principles involved. In order to remove non-cross-linked protein, the cross-linked and nuclease-digested subunits were subjected to electrophoresis through a 10 - 34% sucrose gradient in the presence of 7 M urea, 0.17% Triton-X 100 (Merck, Darmstadt), 170 mM tri-sodium citrate, 2.5 mM EDTA, 6 mM 2-mercaptoethanol and 50 mM Tris-citric acid buffer pH 8. The sucrose gradients were prepared in glass tubes (19 cm x 0.5 cm), closed at the bottom end with dialysis membrane. After electrophoresis at 130 volts overnight, the dialysis membrane was pierced and the gradient fractionated. Fractions containing RNA and RNA-protein complexes were diluted 20-fold with a buffer containing 500 mM NaCl, 1 mM EDTA, 6 mM 2-mercaptoethanol and 10 mM Tris-citric acid pH 8. This solution was warmed to 45° and passed slowly over a Whatman GF/C glass fibre filter (5, and cf. ref.

13), to separate the RNA-protein complexes from non-cross-linked RNA fragments. After washing the filter with the same buffer, the RNA-protein complexes (which adhere to the filter) were eluted by washing the filter with successive 50 µl aliquots of a buffer containing 0.1% SDS, 170 mM tri-sodium citrate, 1 mM EDTA, 5 M urea, 100 mM 2-mercaptoethanol and 10 mM Tris-citric acid pH 8.

Fractions containing the most radioactivity were combined (ca. 20 x 10^6 counts/min) and loaded into a single well of a 6 - 12% polyacrylamide gradient slab gel (40 cm long) containing 0.1% SDS, 7 M urea and 500 mM Na^+ ions (see ref. 7 for details of gel recipe). After electrophoresis for 48 hr at 150 volts, with a change of reservoir buffer after 24 hr, the gel strip containing the sample was cut into 0.17 cm slices, and each slice was eluted with 70 µl of a buffer consisting of 300 mM sodium acetate, 1% CHAPS (3-(3-cholamidopropyl)-dimethylammonio-3-propane sulphonate, Serva, Heidelberg), 0.1% SDS, 1 mM EDTA, 6 mM 2-mercaptoethanol, 6 M urea and 25 mM Tris-acetate buffer pH 8. Each eluate was loaded into a well of the second dimension gel (7), which consisted of 6% polyacrylamide in the presence of 500 mM sodium acetate, 6 M urea and 1% CHAPS as principal ingredients. The second dimension gels (40 cm long) were run for 3 days at 150 volts, with a change of reservoir buffer every 24 hr, and were then subjected to autoradiography for ca. 4 hr. Radioactive spots of interest were cut from the gels, and extracted overnight with 1% SDS, 1 mM EDTA, 150 mM sodium acetate, 6 mM 2-mercaptoethanol and 10 mM Tris-HCl pH 7.8. The RNA-protein complexes were precipitated from these extracts with ethanol, and finally taken up in a small volume (20 - 40 µl) of 0.1% SDS, 10 mM Tris-HCl pH 7.8. Yields of the individual complexes at this stage varied from 5,000 to 50,000 counts/min of ^{32}P-RNA.

Analysis of cross-linked complexes: aliquots of each isolated RNA-protein complex were subjected to protein analysis by the immunological method described in ref. 8. For this purpose, antibodies to the ribosomal proteins to be tested were "spotted" onto a nitrocellulose sheet, and after a suitable washing procedure, equal aliquots of the RNA-protein complex (containing ca. 200 counts/min per aliquot) were pipetted onto each antibody spot.

After a further wash, the nitrocellulose sheet was dried and subjected to autoradiography overnight. RNA radioactivity was only retained in those spots containing the antibody cognate to the protein in the RNA-protein complex. Further aliquots of each RNA-protein complex were subjected to proteinase K digestion followed by phenol extraction as described in ref. 7, in order to remove the protein moiety as far a possible. The residual oligopeptide-RNA complexes were precipitated with ethanol and then totally digested with ribonuclease T_1 or ribonuclease A for "fingerprint" analysis by polyethyleneimine cellulose thin-layer chromatography (14), according to our usual procedure (e.g. refs. 11,15). Each oligonucleotide spot from the fingerprints was subjected to a secondary digestion with ribonuclease A (in the case of oligonucleotides from T_1-fingerprints) or ribonuclease T_1 (for oligonucleotides from A-fingerprints), and the RNA sequences and positions of the cross-link sites were deduced by comparing the oligonucleotide data with the 16S RNA sequence of Brosius et al (16).

RESULTS AND DISCUSSION

30S ribosomal subunits, uniformly labelled with ^{32}P, were cross-linked with APAI (2), and subjected to partial digestion with cobra venom nuclease (12), as described in Materials and Methods. The venom nuclease gives a well-defined pattern of partial digestion products with ribosomal subunits (17), and has been routinely used for some time in our RNA-RNA cross-linking studies (e.g. refs. 9,11). This is the first time, however, that we have applied this enzyme in RNA-protein cross-linking experiments, where in the past we have used either ribonuclease T_1 (4,5) or ribonuclease H (5).

As outlined in the Introduction, the isolation of the RNA-protein complexes from the partial nuclease digest proceeds in three stages (7), examples of which are illustrated in Figs. 1 and 2. The first stage (Fig. 1a) is the removal of non-cross-linked protein by electrophoresis through a sucrose gradient in the presence of a non-ionic detergent (Triton X-100). Preliminary experiments with subunits containing labelled protein showed that the non-cross-linked proteins remain at the top of these gradi-

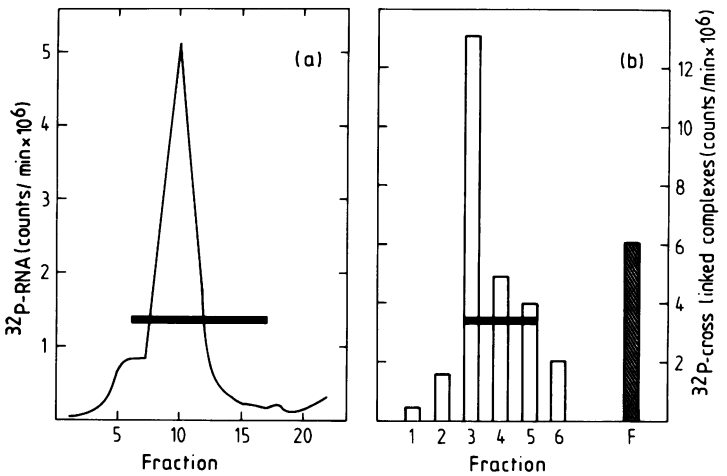

Figure 1: (a) Removal of non-cross-linked proteins from the partially digested cross-linked 30S subunits. The figure is a sucrose gradient electrophoresis profile, with the direction of electrophoresis from right to left. Non-cross-linked protein remains at the top of the gradient (right), as shown by experiments with labelled proteins (7). The bar indicates the fractions pooled for glass-fibre filtration. (b) Step-wise elution of RNA-protein complexes from the glass-fibre filter. The vertical open bars indicate the radioactivity in the individual eluate fractions, the cross-hatched bar (F) denoting the residual radioactivity on the eluted filter. The horizontal bar shows the fractions pooled for gel electrophoresis.

ents (7), whereas the RNA fragments move into the gradient at a rate which is independent of their chain-length. The RNA-protein complexes form a broad smear across the gradient, as indicated by the bar in Fig. 1a.

Removal of the non-cross-linked protein is important, as free proteins have a detrimental effect on the second stage of the separation, namely the removal of non-cross-linked RNA fragments by glass-fibre filtration. The pooled fractions from the sucrose gradient electrophoresis are diluted with buffer and passed over a glass-fibre filter, which retains the cross-linked RNA-protein complexes, and allows most of the free RNA fragments to pass through (5,13). The cross-linked complexes are then eluted from the filter, as illustrated in Fig. 1b. Typically, most of the radioactivity (75%) is recovered in two to three fractions (100 - 150 µl), although, as can be seen from Fig. 1b, about 25% remains irreversibly bound to the filter.

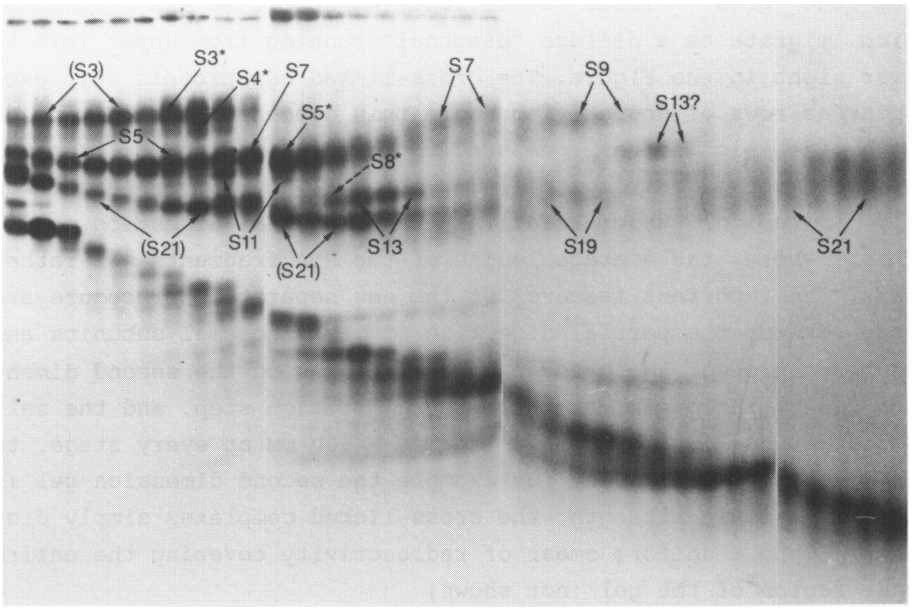

Figure 2: "Two-dimensional" polyacrylamide gel separation of RNA-protein cross-linked complexes. Direction of the first dimension is from left to right, and that of the second dimension from top to bottom. Each second dimension gel has 12 slots, containing the individual eluates from the first-dimension gel slices; four such gels are combined to give the complete pattern. The rows of RNA-protein complexes lie above the "diagonal" of free RNA, and are marked with arrows giving the identity of the protein found in each row of spots. Asterisks denote cross-links to the 5'-terminal region of the 16S RNA, the dotted line (protein S8) indicating the usual position of a complex not seen in this particular gel. Brackets denote cross-linked complexes not discussed in this paper.

The cross-linked RNA-protein complexes eluted from the glass filter are separated by "two-dimensional" gel electrophoresis (Fig. 2). In this gel system (7), the first dimension is a high-percentage polyacrylamide gel run in the presence of SDS, whereas the second dimension is a low-percentage gel in the presence of the zwitter-ionic detergent "CHAPS". The transition from high to low acrylamide concentration necessitates an elution step between the two gel dimensions (see Materials and Methods and ref. 7), so that the result (Fig. 2) is a "pseudo-two-dimensional gel", in which each slot of the second dimension gel contains the eluate from a 0.17 cm slice of the first dimension gel.

It can be seen from the gel pattern of Fig. 2 that the eluate

from the glass filter still contains some free RNA fragments, which migrate as a diffuse "diagonal" running from upper left to lower right in the Figure. The cross-linked RNA-protein complexes appear as rows of spots lying above this diagonal, thus giving an "inverted" picture as compared to that found with our original two-dimensional gel system (e.g. ref. 4). The latter system was only successful under relatively harsh partial digestion conditions, where the average length of the RNA fragments was rather short. The important features of the new separation procedure are that, between the partial digestion of the ribosomal subunits and the separation of the cross-linked complexes on the second dimension gel, there is no ethanol precipitation step, and the salt concentration is maintained at 300 - 500 mM at every stage, to minimize aggregation. If for example the second dimension gel is run at low ionic strength, the cross-linked complexes simply disappear into a uniform smear of radioactivity covering the entire upper region of the gel (not shown).

The RNA-protein complexes are denoted in Fig. 2 according to the protein subsequently identified in the individual gel bands. The gel patterns were remarkably reproducible, although occasional variations were observed; the complex containing protein S8 (indicated by a dotted line in Fig. 2) was for instance not found in this particular gel. Since the RNA is uniformly labelled, the distribution of radioactivity in the gel is in one sense deceptive, because the most prominent bands are those containing the longest RNA fragments. In some cases it was not possible to locate the cross-link site or sites within these longer fragments, and the complexes concerned (containing proteins S3 and S21) are indicated by brackets in Fig. 2. The more useful complexes were those containing shorter RNA fragments, which lie farther to the right in the gel. Despite their diffuse appearance in the gel, these complexes yielded in general the most information.

Radioactive spots containing the RNA-protein complexes were cut out from the gels (Fig. 2) and the complexes eluted. Usually the corresponding gel bands from two or three adjacent slots of the second dimension gel were combined, insofar as this did not involve the risk of cross-contamination with another cross-linked

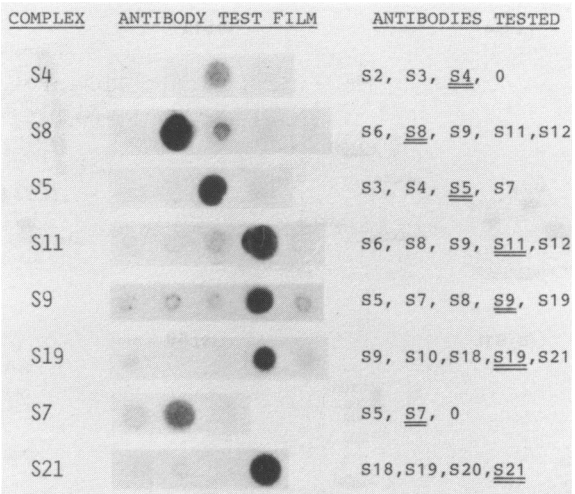

Figure 3: Immunological identification of proteins in the cross-linked complexes (8). The key on the right shows the antibodies tested against each complex, "O" denoting controls minus antibody. The proteins giving a positive reaction are underlined, as can be seen by comparison with the position of the radioactive spots on the "antibody test film" strip.

species. Each isolated complex was subjected to the immunological protein identification procedure (see Materials and Methods and ref. 8), as well as to oligonucleotide analysis in order to deduce the positions of the cross-link sites. The results of a large number of such analyses are summarized in the following sections. Fig. 3 shows examples of the protein identifications, and Fig. 4 illustrates some typical oligonucleotide fingerprints. In Fig. 5 the locations in the 16S RNA of the principal RNA fragments corresponding to each protein are given, and Fig. 6 shows the details of the RNA sequence in the vicinity of the deduced cross-link sites. In each Figure, as in the following discussion, the cross-linked complexes are described in the order in which they occur on the 16S RNA, in a 5'- to 3'-direction, and are named according to the protein which they contain. Twelve cross-link sites are described, consisting of five cross-links to the extreme 5'- or 3'-termini of the 16S RNA (involving proteins S3, S4, S5, S8 and S21), and seven cross-links at internal points in the RNA sequence, involving proteins S5, S7, S9, S11, S13 and S19. In the case of protein S13, two cross-link sites were iden-

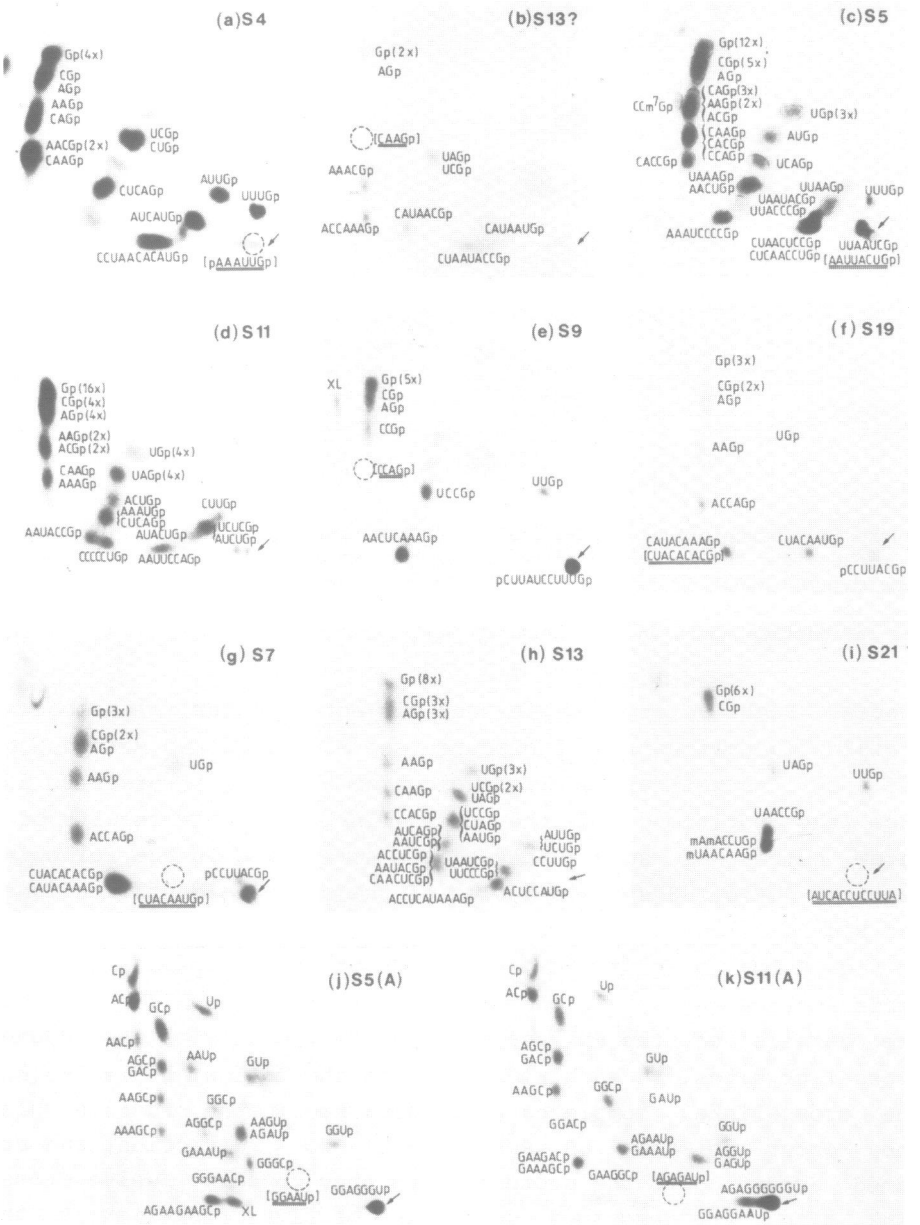

Figure 4: Examples of fingerprint analyses (14) of the cross-linked complexes. (a) to (i) are ribonuclease T_1 fingerprints, (j) and (k) are ribonuclease A fingerprints. The first dimension runs from right to left, and

the second dimension from bottom to top, the arrows indicating the sample application points. Identities and molarities of the oligonucleotides are shown (molarities not being given for the ribonuclease A fingerprints, as these were not routinely quantified). Underlined oligonucleotides in square brackets, together with the dotted circles, indicate the identities and expected positions of the missing (i.e. cross-linked) sequences. "XL" denotes an oligopeptide-oligonucleotide cross-linked species. (The oligonucleotide spots in some of the fingerprints are ringed, as a result of being photographed subsequent to the routine processing procedure for secondary digestion and analysis.)

tified, one of which is anomalous and is accordingly designated as S13? (see below).

Two points should be mentioned in connection with the interpretation of the oligonucleotide data (Figs. 4 and 6). First, the 5'- and 3'-ends of the venom nuclease-generated RNA fragments were sometimes heterogeneous in the complexes described below (and hence difficult to define precisely), as we have observed previously (e.g. refs. 4, 5, 9). In such cases the RNA sequences are discussed in terms of the first or last observed characteristic ribonuclease T_1 oligonucleotide at each end of the fragment. Secondly, the positions of the cross-links were as usual inferred from the absence of a characteristic oligonucleotide from its expected position on either the ribonuclease T_1 or ribonuclease A

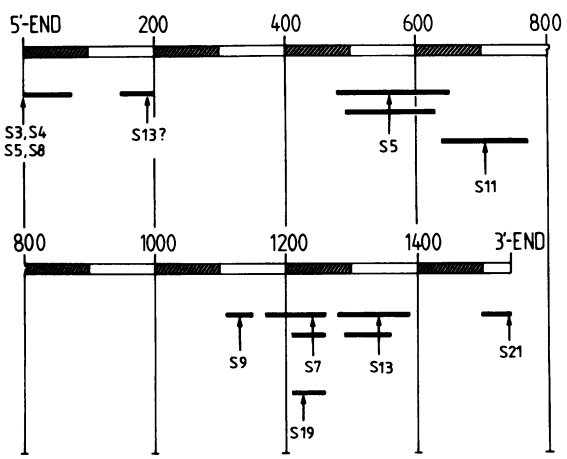

Figure 5: Locations of the cross-linked complexes in the 16S RNA sequence. The horizontal bars indicate the sequence positions of the RNA fragments in the complexes, with the arrow and protein number denoting the positions of the cross-link sites to the various proteins within these RNA fragments.

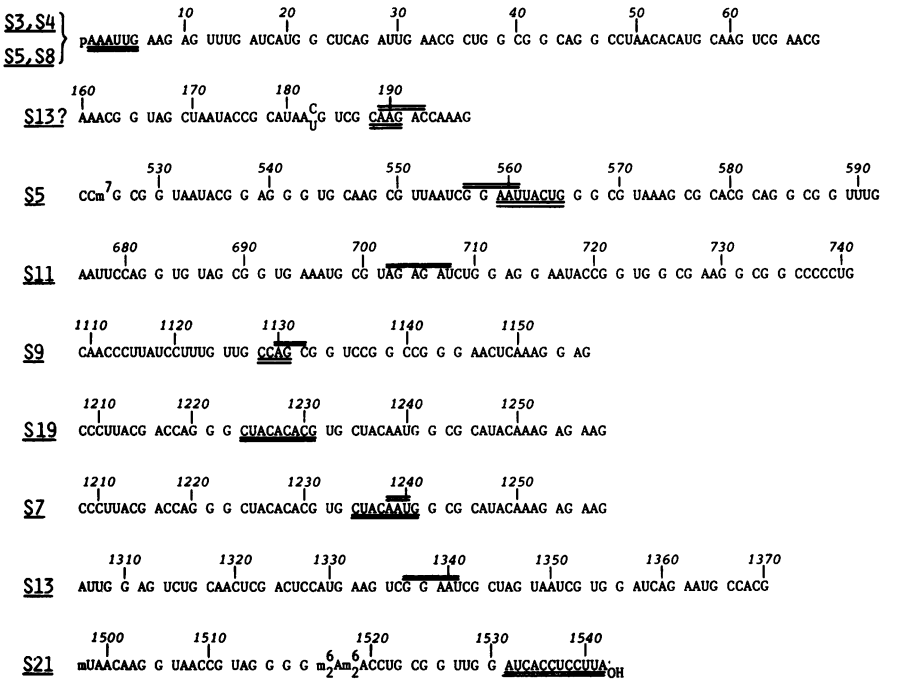

Figure 6: Details of the RNA sequence (16) in the vicinity of each cross-link site. The sequence is divided into ribonuclease T_1 oligonucleotides, to facilitate comparison with Fig. 4. Missing (i.e. cross-linked) T_1-oligonucleotides are underlined, and the corresponding missing ribonuclease A oligonucleotides are marked by lines above the sequence concerned.

fingerprint, or from absences on both fingerprints. Since the cross-linked complexes were treated with proteinase K prior to the fingerprint analysis (see Materials and Methods), one would expect the residual cross-linked product to appear as an oligonucleotide-oligopeptide spot on the fingerprint. In practice, however, such cross-linked spots could seldom be analysed, because the proteinase K digestion is usually heterogeneous, leading to a number of peptide-oligonucleotide products, which appear as fast-running and ill-defined spots on the fingerprint (cf. ref. 5). In other cases these cross-linked products are insoluble, and are not transferred efficiently to the polyethyleneimine chromatography plates.

Before describing the individual complexes, it should also be noted that (as mentioned in Materials and Methods) the magnesium

concentration during the cross-linking reaction was reduced from 5 mM to 0.5 mM in some experiments. No significant differences in the gel patterns (cf. Fig. 2) were observed at the lower magnesium concentration, and the subsequent protein and RNA analyses were also precisely comparable to those obtained at the higher magnesium concentration. This is an important point, as it shows that the cross-linking method does not distinguish between the active and inactive conformations of the 30S ribosomal subunit (18), and thus supports the view (19) that the active-inactive transition is a relatively small local effect, rather than a general unfolding of the subunits. Further, from the point of view of our model-building studies (20), the result has the consequence that cross-linking data obtained with active subunits can be directly correlated with the protein distribution derived by neutron scattering studies (21), although these latter data were obtained at low (0.5 mM) magnesium concentration.

Crosslinks to the 5'-terminus of the 16S RNA: Four proteins, S3, S4, S5 and S8, were found cross-linked to the 5'-terminal region of the 16S RNA. The corresponding four complexes are marked with an asterisk in Fig. 2, and examples of the immunological protein identification (for the S4 and S8 complexes) are shown in Fig. 3. In each case the ribonuclease T_1 fingerprints were identical (that from the S4 complex being illustrated in Fig. 4a), and corresponded to the 5'-terminal 72 nucleotides of the RNA (Fig. 5). This is a well-established venom nuclease fragment (17), and characteristic oligonucleotides from this region can be seen in the fingerprint (compare Figs. 4a and 6). The 5'-terminal hexanucleotide pAAAUUGp (usually present in the venom nuclease fragment) was however missing from its expected position on the fingerprint. This oligonucleotide is therefore the candidate for the cross-link site. In a number of experiments, spots corresponding to the cross-linked oligopeptide-oligonucleotide were observed (not visible in Fig. 4a), which in each case yielded Up, Gp and an undigested residue upon secondary digestion with ribonuclease A. Since all the other ribonuclease T_1 oligonucleotides which could have given this combination of products are accounted for in the fingerprint (cf. Figs. 4a and 6), this confirms that the 5'-terminal hexanucleotide contains the cross-link, and indi-

cates that the actual site of cross-linking is within the terminal tetranucleotide AAAU.

Complex S13?: As already mentioned above, this complex is designated with a query, because, although the identification of the cross-link site itself was unequivocal, the RNA region found was anomalous with respect to our model of the 30S subunit (20). The immunological reaction with anti-S13 was only weak (not shown in Fig. 3), and the identification of S13 was based on a protein gel analysis of the cross-linked complex after total digestion with ribonucleases A and T_1, as in ref. 4. This analysis was made in the two-dimensional system of ref. 22, and gave a radioactive spot which corresponded to a residual S13-oligonucleotide complex, running close to the position of protein S13 (data not shown, but cf. the protein analyses in ref. 4). The ribonuclease T_1 fingerprint (Fig. 4b) was very reproducible and covered the region from positions 160-198 (Fig. 6), or occasionally from 166-198 (encompassing the U/C heterogeneity at position 183 (23)), but the oligonucleotide CAAGp (pos. 188-191) was always absent (Fig. 4b). Further, AAGACp (pos. 189-193) was missing from the corresponding ribonuclease A fingerprint, and we therefore conclude that the cross-link site lies within the sequence AAG (pos. 189-191, Fig. 6).

Neutron scattering data (21) have shown that protein S13 is located on the "head" of the 30S subunit, whereas nucleotides 189-191 can only be placed at the base of the subunit (20). This anomaly could possibly be accounted for by an inter-subunit cross-link in a "head-to tail" 30S dimer in our experiments. No dimers were detected in the cross-linked subunits, however, as judged by their behaviour in sucrose gradients (data not shown), although a small amount of dimer formation may have passed undetected. An alternative explanation is that protein S13 is only loosely bound to the 30S subunit, and can "wander" over the subunit surface. Protein-protein cross-linking studies have shown that S13 becomes cross-linked to many more proteins than its position on the 30S subunit (as determined by neutron scattering) should allow (see ref. 21 for discussion), and the cross-link site to RNA positions 189-191 is discussed further in ref. 20.

Complex S5: Various versions of the S5 complex, with attached

RNA fragments of different lengths, were found interspersed with S7 complexes (Fig. 2). Cross-contamination with the S7 complexes was quite common, but clean complexes which only contained S5 were also found (see the protein analysis in Fig. 3). In these complexes the RNA sequence involved covered the region from positions ca. 480 to 650 (Fig. 5), and the ribonuclease T_1 fingerprint data (Fig. 4c) showed that the oligonucleotide AAUUACUGp (pos. 559-566) was always missing. GGAAUp (pos. 557-561) was missing from the correponding ribonuclease A fingerprint (Fig. 4j), and the latter fingerprint also showed a cross-link spot ("XL") which gave Gp and an undigested residue upon secondary digestion with ribonuclease T_1. The cross-link site therefore lies within the sequence AAU (pos. 559-561, Fig. 6).

Complex S11: Both protein identification (Fig. 3) and ribonuclease T_1 fingerprint (Fig. 4d) were clearcut and reproducible for this complex, the RNA arising from positions ca. 640 to 770 of the 16S RNA sequence (Fig. 5). No cross-link site could be discerned from the ribonuclease T_1 fingerprint (i.e. no characteristic oligonucleotide was absent), but the oligonucleotide AGAGAUp (pos. 702-707) was reproducibly missing from the corresponding ribonuclease A fingerprint (Fig. 4k). Since AUCUGp (pos. 706-710) was present in the T_1-fingerprint, we conclude that the cross-link site lies within the sequence AGAG (pos. 702-705, Fig. 6).

Complex S9: This complex was exceptionally well-defined. The protein analysis (Fig. 3) was unequivocal, and the RNA fragment covered the sequence region from positions 1114 to 1156. The 5'-end was defined by the oligonucleotide pCUUAUCCUUUGp (Fig. 4e) indicating a venom nuclease cut at position 1113-1114, and the 3'-end was defined by the presence in the ribonuclease T_1 fingerprint of AGp, which occurs at positions 1155-1156, following the characteristic oligonucleotide AACUCAAAGp (Figs. 4e and 6). The oligonucleotide CCAGp (pos. 1128-1131) was missing from the T_1-fingerprint, which instead showed a cross-link spot ("XL" in Fig. 4e), giving Cp and an undigested residue upon secondary digestion with ribonuclease A. Furthermore the oligonucleotide AGCp (pos. 1130-1132) was absent from the corresponding ribonuclease A

fingerprint. The cross-link site is thus established within the sequence AG at positions 1130-1131 (Fig. 6).

Complex S19: The RNA fragment in this complex was of a similar length to that in the S9 complex, in this case covering the sequence region from position 1209 (as indicated by the oligonucleotide pCCUUACGp, Fig. 4f) to position 1257 (as indicated by the presence in the T_1-fingerprint of AAGp, which follows the characteristic oligonucleotide CAUACAAAGp (Fig. 6)). The protein analysis was unambiguous (Fig. 3), and the cross-link site was within the oligonucleotide CUACACACGp (pos. 1223-1231), since this oligonucleotide was always absent from the T_1-fingerprint (Fig. 4f). Because of the repeated AC sequence within this oligonucleotide, the ribonuclease A fingerprint data did not allow the position of the cross-link site to be narrowed down any further.

Complex S7: This complex appeared in several positions on the two-dimensional gel (Fig. 2), interspersed with S5 complexes, as noted above. The shortest version of the S7 complex contained the identical RNA fragment to the S19 complex just discussed. In this case however the missing oligonucleotide was CUACAAUGp (pos. 1234-1241, Fig. 4g), and since the corresponding ribonuclease A fingerprint contained no AAUp, we conclude that the cross-link site lies within the sequence AAU (pos. 1238-1240, Fig. 6). The protein analysis was clean (Fig. 3), and the cross-link site is thus identical to that found for protein S7 using 2-iminothiolane as cross-linking reagent (4).

Complex S13: As with complex S13? (above), the immunological reaction with anti-S13 was weak, and the protein identification was based on a two-dimensional gel analysis. In this case the RNA covered the region between positions ca. 1280 and 1390 of the 16S RNA (Fig. 5), but no cross-link site could be discerned from the ribonuclease T_1 fingerprint (Fig. 4h). However, in the corresponding ribonuclease A fingerprint, the oligonucleotide GGAAUp (pos. 1337-1341) was always missing, although the neighbouring isomeric oligonucleotide (GAAGUp, pos. 1331-1335) was present at the same position on the fingerprint (not shown in Fig. 4). Since AAUCGp (pos. 1339-1343) was present in the T_1-fingerprint, the cross-link site must lie within the sequence GG (pos. 1337-1338, Fig. 6). In contrast to complex S13?, this position is compatible

with the location of protein S13 in the 30S subunit (20).

Complex S21: The complex containing S21 appeared as a broad smear on the right-hand side of the two-dimensional gels (Fig. 2). The protein anlysis was very clean (Fig. 3), and the RNA arose from the extreme 3'-terminus of the 16S RNA, covering positions 1498 onwards. Slightly shorter or longer vesions of the complex were also observed. The 3'-terminal oligonucleotide AUCACCUCCUUA (pos. 1531-1542) was never found in its usual position near the origin of the T_1-fingerprint (Fig. 4i), and occasionally a spot corresponding to the cross-link was seen at the left-hand edge of the fingerprint, which liberated AUp, ACp, Cp and Up upon secondary digestion with ribonuclease A. The cross-link thus lies within the 3'-terminal dodecanucleotide of the 16S RNA (Fig. 6), as was previously observed in our experiments with 2-iminothiolane (5). The ribonuclease A fingerprint data did not allow the cross-link site to be narrowed down any further.

CONCLUSIONS

We have described here the identification of twelve APAI-induced RNA-protein cross-link sites on the 16S RNA. The localizations of the sites are to within two or three nucleotides, with the exceptions of the cross-links to proteins S19 and S21, where the localization is to a nona- and a dodecanucleotide, respectively. These localizations are deduced on the basis of the minimum common sequence that is missing from the ribonuclease A and T_1 fingerprints in each case (Fig. 6). It should be noted however that the chemistry of the photolytic reaction between APAI and the RNA is not know, and may vary from one cross-link to another. If the cross-linked base becomes resistant to hydrolysis by ribonuclease T_1 or A, then this could lead to situations where the actual cross-link site is one base further to the 5'-side of the deduced oligonucleotide (e.g. to position G-558 in the S5 complex, Fig. 6). Despite this reservation, the localizations are still amply accurate enough for our purpose, as defined in the Introduction.

The locations of all the APAI cross-links in the secondary structure of 16S RNA are shown in Fig. 7, (the structure being a slightly modified version of our previous model (24), as dis-

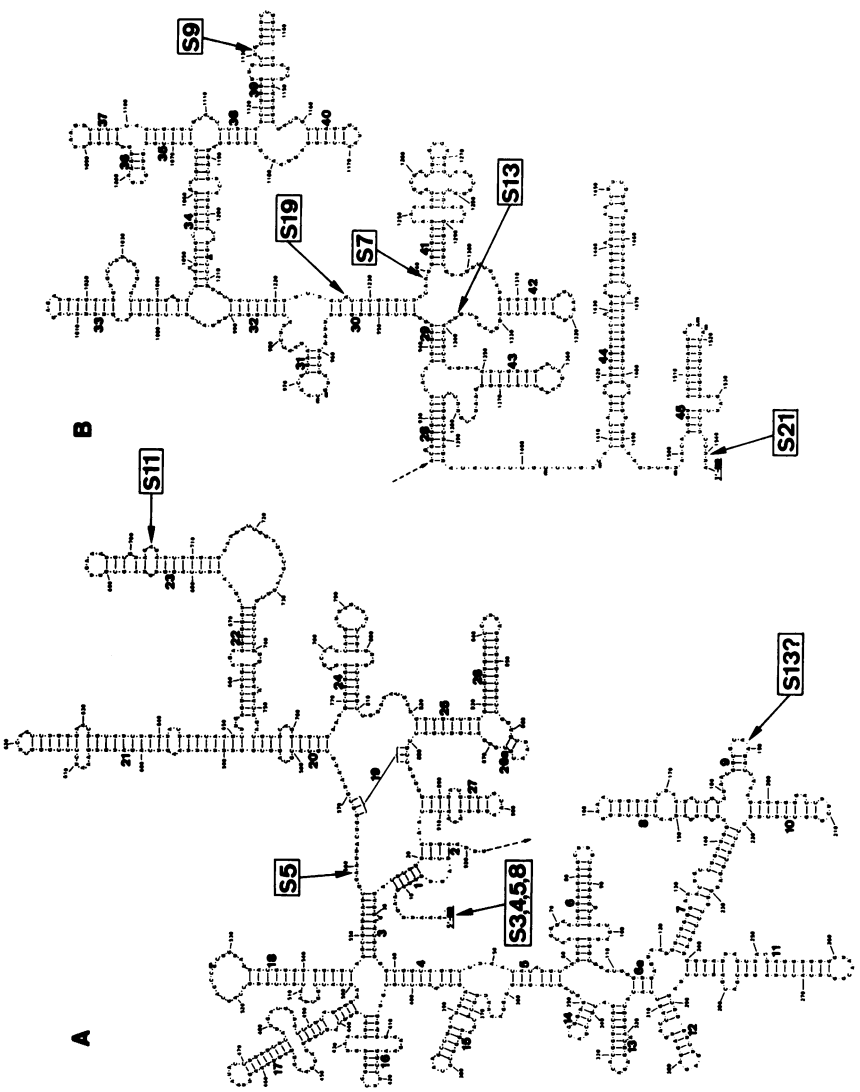

Figure 7: Locations of the APAI-induced cross-link sites in the secondary structure of 16S RNA (cf. ref. 24). The arrows representing the cross-links are drawn to the central residue of the oligonucleotide implicated in each cross-link (cf. Fig. 6).

cussed in ref. 20). The APAI cross-linking reaction seems to prefer single-stranded regions of the RNA, although there may be other sites which are discriminated against by the cobra venom nuclease digestion; as we have often noted (e.g. refs. 5,9,11), the spectrum of observed cross-links is highly dependent on the nature of the partial digestion procedure. The cross-links to proteins S5, S11, S9, S19 and S13 all represent entirely new neighbourhoods for ribosomal proteins in the 16S RNA. In combination with our previous data, as well as with a comparable set of cross-link sites obtained with nitrogen mustard (10), these cross-link sites have played a central role in the construction of a detailed model for the three-dimensional folding of the 16S RNA in the 30S subunit. With the notable exception of cross-link S13?, all of the cross-link sites are fully compatible with this model, which is presented and discussed in detail in ref. 20.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. H.G. Wittmann for his support and for helpful discussions.

*To whom correspondence should be addressed

REFEENCES

1. Ulmer, E., Meinke, M., Ross, A., Fink, G. and Brimacombe, R. (1978) Mol. Gen. Genet. 160, 183-193.
2. Rinke, J., Meinke, M., Brimacombe, R., Fink, G., Rommel, W. and Fasold, H. (1980) J. Mol. Biol. 137, 301-314.
3. Wower, I., Wower, J., Meinke, M., and Brimacombe, R. (1981) Nucleic Acids Res. 9, 4285-4302.
4. Wower, I. and Brimacombe, R. (1983) Nucleic Acids Res. 11, 1419-1437.
5. Kyriatsoulis, A., Maly, P., Greuer, B., Brimacombe, R., Stöffler, G., Frank, R. and Blöcker, H. (1986) Nucleic Acids Res. 14, 1171-1186.
6. Fink, G., Fasold, H., Rommel, W. and Brimacombe, R. (1980) Analyt. Biochem. 108, 394-401.
7. Brimacombe, R., Stiege, W., Kyriatsoulis, A. and Maly, P. (1987) Methods in Enzymology, in press.
8. Guller, H., Brimacombe, R., Stöffler-Meilicke, M. and Stöffler, G. (1987), submitted for publication.
9. Atmadja, J., Stiege, W., Zobawa, M., Greuer, B., Oβwald, M. and Brimacombe, R. (1986) Nucleic Acids Res. 14, 659-673.
10. Greuer, B., Oβwald, M., Brimacombe, R. and Stöffler, G. (1987), submitted for publication.
11. Stiege, W., Zwieb, C. and Brimacombe, R. (1982) Nucleic Acids Res. 10, 7211-7229.

12. Vassilenko, S.K. and Ryte, V.C. (1975) Biokhimiya 40, 578-583.
13. Thomas, C.A., Saigo, K., McCleod, E. and Ito, J. (1979) Analyt. Biochem. 93, 158-166.
14. Volckaert, G. and Fiers, W. (1977) Analyt. Biochem. 83, 228-239.
15. Maly, P., Rinke, J., Ulmer, E., Zwieb, C. and Brimacombe, R. (1980) Biochemistry 9, 4179-4188.
16. Brosius, J., Palmer, M.L., Kennedy, P.J. and Noller, H.F. (1978) Proc. Nat. Acad. Sci. USA 75, 4801-4805.
17. Vassilenko, S.K., Carbon, P., Ebel, J.P. and Ehresmann, C. (1981) J. Mol. Biol. 152, 699-721.
18. Zamir, A., Miskin, R. and Elson, D. (1971) J. Mol. Biol. 60, 347-364.
19. Moazed, D., Van Stolk, B.J., Douthwaite, S. and Noller, H.F. (1986) J. Mol. Biol. 191, 483-493.
20. Brimacombe, R., Atmadja, J. and Stiege, W. (1987), submitted for publication.
21. Moore, P.B., Capel, M., Kjeldgaard, M. and Engelman, D.M. (1986) in "Structure, Function and Genetics of Ribosomes", Hardesty, B. and Kramer, G., Eds., pp. 87-100, Springer-Verlag, New York.
22. Geyl, D., Böck, A. and Isono, K. (1981) Mol. Gen. Genet. 181, 309-312.
23. Carbon, P., Ehresmann, C., Ehresmann, B. and Ebel, J.P. (1979) Eur. J. Biochem. 100, 399-410.
24. Maly, P. and Brimacombe, R. (1983) Nucleic Acids Res. 11, 7263-7286