Volume 13 Number 19 1985 Nucleic Acids Research

Investigation of the tertiary folding of *Escherichia coli* 16S RNA by *in situ* intra-RNA cross-linking within 30S ribosomal subunits

Johannes Atmadja and Richard Brimacombe
Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, D-1000 Berlin-Dahlem, FRG, and

Helmut Blöcker and Ronald Frank
Gesellschaft für Biotechnologische Forschung, Abteilung DNA-Synthese, D-3000 Braunschweig-Stöckheim, FRG

Received 16 July 1985; Accepted 6 September 1985

ABSTRACT

Intra-RNA cross-links were introduced into E. coli 30S ribosomal subunits by mild ultraviolet irradiation. The subunits were partially digested with cobra venom nuclease, followed in some cases by a second partial digestion with ribonuclease H in the presence of the hexanucleotide d-(CTTCCC). The cross-linked RNA complexes were separated by two-dimensional gel electrophoresis and the sites of cross-linking analysed by our published procedures. Tertiary structural cross-links in the 16S RNA were identified between positions 31 and 48, between oligonucleotides 1090-1094 and 1161-1164, and between oligonucleotides 1125-1127 and 1280-1281. The first of these imposes a rigid constraint on the relative orientations of helices 3 and 4 of the 16S secondary structure. A further tertiary cross-link (which could not be precisely localised) was found between regions 1-72 and 1020-1095, and secondary structural cross-links were identified between positions 497 and 545-548, and positions 1238-1240 and 1298.

INTRODUCTION

The application of intra-RNA cross-linking techniques is the most direct and obvious approach for probing the tertiary structure of a large RNA molecule, and in the case of 16S and 23S ribosomal RNA several research groups have been involved in this type of study (1-9). For technical reasons, many of these investigations (5-9) have been confined to an analysis of cross-links induced within isolated RNA. In contrast, we have always been of the firm opinion that studies on the tertiary structure of ribosomal RNA can only be meaningful if the cross-links are induced in situ in the intact ribosomal subunits or 70S ribosomes (1-4).

In any cross-linking experiment of this type, the first step in the identification of the sites of cross-linking is to subject the cross-linked subunits (or subsequently isolated RNA) to

a partial nuclease digestion procedure, in order to obtain cross-linked RNA complexes of a suitable size for analysis. However, it has already become clear that the nature of the partial digestion procedure has an enormous influence on the spectrum of cross-links observed. In our earlier experiments with 30S and 50S subunits, using ribonuclease T_1 to generate fragments of the cross-linked RNA, we were only able to find cross-links that were located within the secondary structures of 16S (1) or 23S (2) RNA. These experiments were made with ribosomal subunits that had been cross-linked by mild ultraviolet irradiation. The first tertiary cross-link (i.e. a cross-link between RNA regions that are remote from each other in both the primary and secondary structures) was found in 23S RNA from subunits cross-linked with bis-(2-chloroethyl)-methylamine, and in this case (3) the partial digestions were made with cobra venom nuclease (10) coupled with a second nuclease digestion. Subsequently, using these same digestion procedures, we were able to identify a series of tertiary structural cross-links in 23S RNA (4), which had been induced by irradiation of 50S subunits under the same conditions as in our earlier studies (2).

In this paper we describe a similar series of experiments with the E. coli 30S subunit. The subunits were cross-linked by mild irradiation, and digested by treatment with cobra venom nuclease. In some cases the venom digestion was followed by a further digestion with ribonuclease H in the presence of the hexanucleotide d-(CTTCCC) (cf. 11, 12,4). Cross-linked RNA complexes were isolated from these digests by two-dimensional gel electrophoresis and subjected to oligonucleotide analysis by our established methodology (3,4), and we were able to identify a total of four tertiary structural and two secondary structural cross-links. The significance of these results with regard to the three-dimensional organisation of the 16S RNA within the 30S subunit is discussed.

MATERIALS AND METHODS

Preparation, irradiation and partial digestion of 30S subunits. ^{32}P-labelled 30S ribosomal subunits from E. coli strain MRE 600 were prepared by the rapid method of Stiege et al (3). The sub-

units (ca. 4 A_{260} units, 1.5 x 10^9 counts/min) were irradiated for 6 min at a concentration of 5 A_{260} units/ml in 25 mM triethanolamine-HCl pH 7.8, 50 mM KCl, 5 mM $MgCl_2$, 6 mM 2-mercaptoethanol, under the conditions already described (13). The $MgCl_2$ concentration was then raised to 10 mM and that of the KCl to 300 mM, and a suitable quantity of cobra (Naja naja oxiana) venom nuclease (10) was added (cf. 3,4). After incubation at 37° for 1 hr, the reaction was stopped by addition of EDTA, followed by ethanol precipitation and proteinase K treatment, as already published (3), with the exception that the proteinase K digestion was carried out in the presence of 400 mM NaCl in addition to 0.5% sodium dodecyl sulphate.

In some cases the samples were subjected at this stage to a digestion with ribonuclease H. For this purpose the proteinase K-treated reaction mixtures were extracted overnight at 4° with phenol, and RNA was isolated from the aqueous phase by ethanol precipitation. After washing the pellet with 80% ethanol and drying, the RNA (ca. 1.2 A_{260} units) was taken up in 10 mM Tris-HCl pH 7.8 (16 µl). This solution was adjusted to 40 mM Tris-HCl pH 7.8, 50 mM KCl, 10 mM $MgCl_2$, 0.1 mM EDTA, 0.1 mM dithiothreitol, 1% Triton X-100, in a total volume of 80 µl, containing 0.4 A_{260} units of d-(CTTCCC) and a suitable quantity of ribonuclease H (cf. 4,14). Incubation was for 16 hr at 4°, after which the reaction was stopped by addition of EDTA, followed by ethanol precipitation and proteinase K treatment as above. The oligonucleotides were synthesized in solution following established phosphotriester procedures (15), and were purified by ion exchange chromatography under denaturing conditions (16).

Two-dimensional gel electrophoresis and oligonucleotide analysis of cross-linked complexes. Cross-linked RNA samples, after digestion with venom nuclease alone or with venom nuclease and ribonuclease H, were subjected to a final ethanol precipitation in order to concentrate them and remove salt, and were then applied to the two-dimensional electrophoresis system of ref. 3. As before (3) two second dimension gels were made for each first dimension gel strip, the upper half of the latter being run on a 10% second dimension gel, and the lower half on a 20% gel. After electrophoresis, the cross-linked complexes were located

on the gels by autoradiography, and were extracted in the usual manner (17). In one case (complex E, see results), the isolated complex was further digested with ribonuclease T_1 in the presence of 50 µg of unlabelled carrier tRNA, followed by separation on a second two-dimensional gel, exactly as described previously (3).

Isolated cross-linked RNA complexes were digested completely with either ribonuclease T_1 or ribonuclease A as before (3,4), using the "minifingerprint" system of Volckaert and Fiers (18) to separate the oligonucleotide products. (In some later experiments the ribonuclease T_1 fingerprints were run in buffers that were 20% more concentrated in both chromatographic dimensions than those of ref. 18). Secondary digestions of the oligonucleotides from the fingerprints were also made as before (3,4), using the "double-digestion" system (18), and in some cases tertiary digestions were made, as described (4). Oligonucleotide data were fitted to the 16S RNA sequence of Brosius et al (19).

RESULTS

Examples of autoradiographs of two-dimensional polyacrylamide gels obtained from irradiated ^{32}P-labelled 30S ribosomal subunits after partial nuclease digestion are shown in Fig. 1. The left-hand panel shows a gel corresponding to the upper region of the first dimension gel strip (cf. ref. 3) and the sample applied was from irradiated subunits digested with cobra venom nuclease alone (see Materials and Methods). It has previously been demonstrated (3,4) that spots running above the "diagonal" of free RNA fragments in this system correspond to the intra-RNA cross-linked complexes. The lettered complexes are those which led to clearcut and reproducible cross-link site analyses, and they are lettered according to their position in the 16S RNA sequence (cf. Fig. 4, below). It can be seen that the pattern of cross-linked complexes in Fig. 1 (left panel) is dominated by the series of spots marked "C", and the subsequent oligonucleotide analyses showed that all these complexes contained the same cross-link, but with the RNA fragments involved being of different lengths in each case. This particular gel pattern was extremely reproducible, although the relative intensities of some

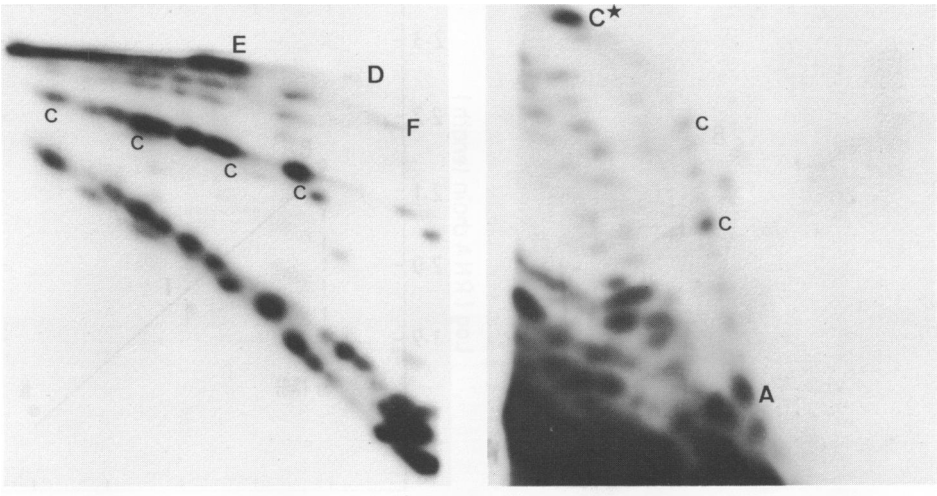

Figure 1: Separation of ^{32}P-labelled cross-linked RNA complexes on two-dimensional gels. The left-hand panel is a cobra venom digest on a 10% gel, corresponding to the upper region of the first-dimension gel strip (cf. ref.3). The right-hand panel is a combined cobra venom nuclease and ribonuclease H digest (see Materials and Methods) on a 20% gel, corresponding to the lower half of the first-dimension gel strip. Directions of electrophoresis are from left to right (first dimension) and top to bottom (second dimension). The lettered spots are the cross-linked complexes discussed in the text.

of the spots tended to vary. For instance, in the particular example shown, spot "E" is relatively strong whereas spots "D" and "F" are relatively weak.

Although it is impossible to make an accurate estimate of the yield of a cross-link in this system, as a result of the selectivity of any particular partial digestion condition (cf. Introduction), it is clear that the yield of cross-link C is extremely high, with the result that complexes were observed containing multiple cross-links (i.e. cross-link C together with a further unidentified cross-link). These can be seen in the left panel of Fig. 1, running as a double row of spots in between complex E and the strong line of spots marked C. These multiple cross-links, as well as the smeared complexes which only just enter the second dimension gel (to the left of complex E) were not analysed in detail.

The right-hand panel of Fig. 1 shows a gel corresponding to

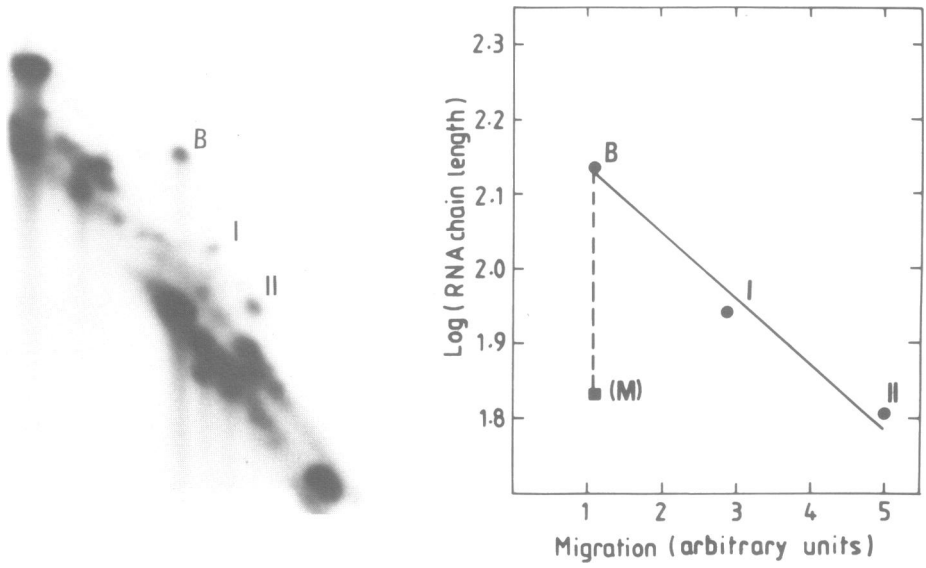

Figure 2: Left panel, another example of a combined cobra venom nuclease and ribonuclease H digest, as in the right-hand panel of Fig. 1. Spot B is a cross-linked complex, spots I and II are simple RNA fragments. The right-hand panel is a plot of mobility in the first gel dimension versus log-(chain length) for spots B, I and II. The point marked M indicates where spot B would lie if it were a fortuitous mixture of its two RNA components (see text).

part of the lower region of a different first dimension gel strip, and in this case the sample applied was from subunits digested with both cobra venom nuclease and ribonuclease H together with the hexanucleotide d-(CTTCCC) (see Materials and Methods). The reason for choosing this particular hexanucleotide for studies on 23S RNA was discussed in detail in ref. 4, although in the latter case the ribonuclease H reaction was used as a second digestion procedure for the individual complexes after isolation on the two-dimensional gels. Here, with 16S RNA, the same hexanucleotide gave the most extensive digestion among several tested. However, in these experiments the ribonuclease H digestion was made not with the isolated complexes but with the whole reaction mixture immediately after the venom nuclease digestion, the rationale being that a series of related venom digestion products (such as the fragments containing cross-link C (Fig. 1, left panel)) might then give rise to the same or similar

smaller digestion products, with a corresponding improvement in yield of the latter.

This strategy proved to be partially successful, and in the example shown in the right-hand panel of Fig. 1 a number of strong spots corresponding to cross-linked complexes can be seen, most notably those marked "A" and "C*", C* being the version of complex C whose analysis will be described in detail below (cf. Fig. 3). On the other hand, the pattern of fragments tended to be variable with this procedure, and again tended to be dominated by various versions of the ubiquitous complex C. In order to illustrate this variability, a further example of a combined venom nuclease and ribonuclease H digestion is shown in Fig. 2. Again the gel shown is from the lower region of the first dimension gel strip, but in this instance the ribonuclease H digestion led to much less hydrolysis than in the example shown in Fig. 1 (right panel), and enabled us to isolate and subsequently make a very clear analysis of complex "B". (The spots marked I and II in Fig. 2 are simple continuous RNA fragments, but their analysis was important in establishing the nature of complex B (see below)).

Many cross-linked complexes from a number of two-dimensional gels such as those depicted in Figs. 1 and 2 were subjected to oligonucleotide analysis by digestion with ribonucleases A or T_1, as outlined in Materials and Methods. The fingerprint data obtained led to the identification of six cross-links (A - F), and in the following section one example of an RNA complex corresponding to each of these cross-links will be described in detail. Examples of the fingerprints concerned are shown in Fig. 3, and the distribution of the cross-links in the 16S RNA sequence (19) is given in Fig. 4. Fig. 4 also indicates the sequence region in the vicinity of the cross-linked sites in each case, so that the reader can follow the arguments involved in the various site determinations.

Complex A: Complexes containing this cross-link were observed on gels both with and without ribonuclease H digestion, the best example being from the gel shown in Fig. 1 (right panel, with ribonuclease H). The ribonuclease T_1 fingerprint of the latter is given in Fig. 3, and covers the sequence region from

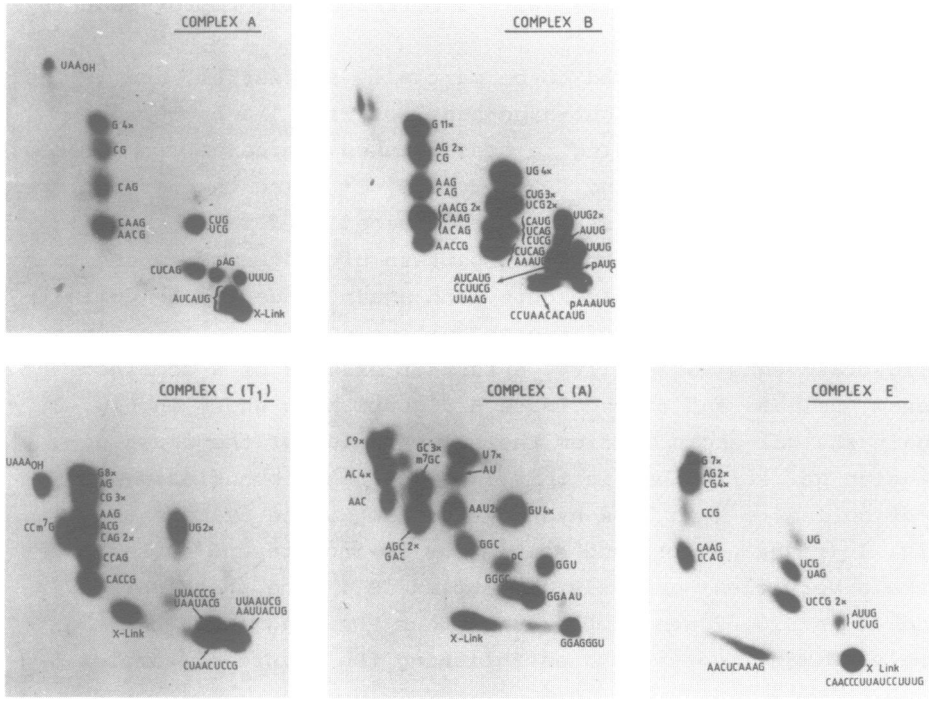

Figure 3: Examples of fingerprints of cross-linked complexes on polyethyleneimine plates (18). The first dimension is from right to left, the second from bottom to top. The identities and molarities of the oligonucleotides are given, "X-link" denoting the cross-linked oligonucleotide complex in each case. The top row shows T_1-fingerprints of Complexes A and B; the bottom row shows both a T_1- and A-fingerprint of Complex C, and a T_1-fingerprint of Complex E (after a second partial digestion procedure, see text).

positions 10-72 of the 16S RNA (Fig. 4). The characteristic oligonucleotides defining this region (UUUG, AUCAUG, CUCAG, etc) can be seen in the fingerprint. The 5'-end was located by the absence of AAG and AG (positions 7-11) from the T_1-fingerprint, and the presence of UUUG (positions 12-15). This was confirmed by the absence of GAAGAGU (positions 6-12) from the ribonuclease A fingerprint, and indicates a probable ribonuclease H site (positions 6-9) with a complementarity to the first four bases of CTTCCC. A spot corresponding to pAG (positions 10-11) was found in the T_1-fingerprint (Fig. 3). The 3'-end of the fragment was delineated by the absence of UAACAG (70-75) and the presence of UAA_{OH} (70-72) in the T_1-fingerprint. This terminus corre-

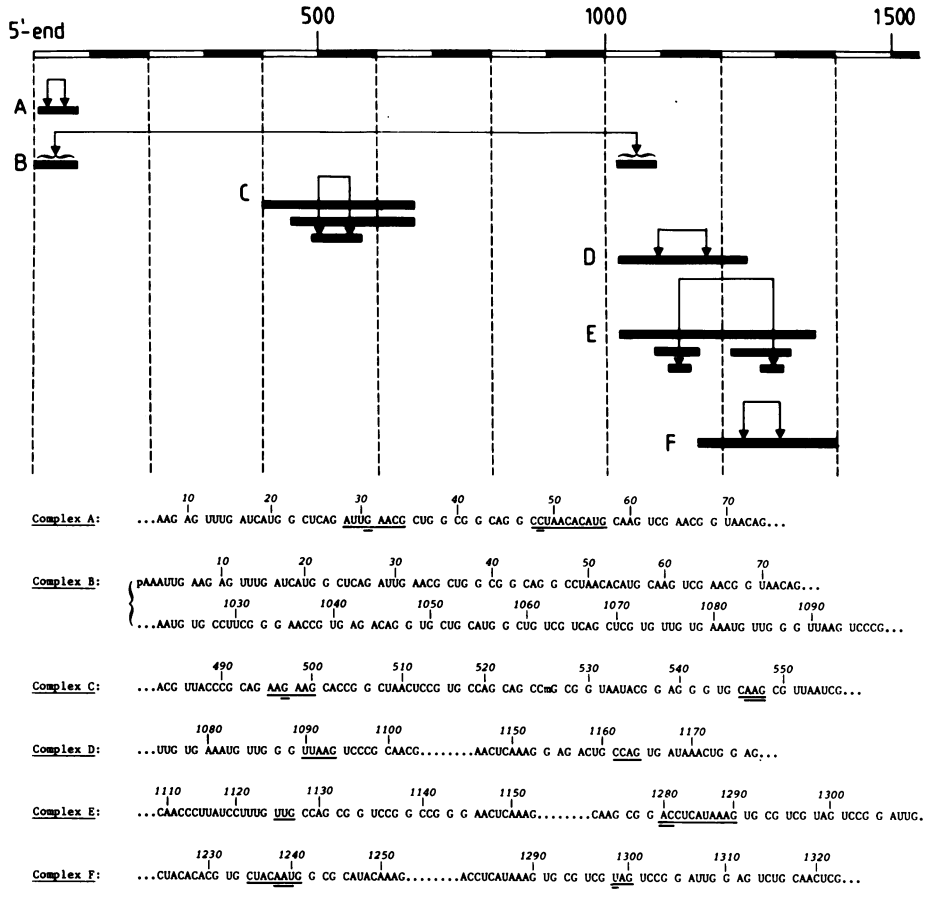

Figure 4: Location of the RNA regions containing the cross-linked complexes in 16S RNA (cf. Figs. 1-3). The sequence is numbered from the 5'-end. For clarity, in the upper part of the diagram, only selected examples from each family of complexes are shown; the thick bars denote the sequence region concerned, the thin lines with arrowheads indicating the cross-links in each case. In Complex B, the cross-link site could not be localised. The lower part of the Figure shows the details of the 16S sequence in the vicinity of each cross-link site, the sequence being divided into the ribonuclease T_1 oligonucleotides. The cross-linked T_1-oligonucleotides are underlined, with double underlining indicating the precise cross-link site within the T_1-oligonucleotide, insofar as this could be determined.

sponds to the established cobra venom cutting point at position 72 (20).

The cross-link site was indicated by the absence of AUUG (28-31) and CCUAACACAUG (47-57) from the T_1-fingerprint (Fig. 3),

and furthermore only one mole of AACG (occurring at positions 32-35 and 65-68) was found. Instead of these oligonucleotides a new spot corresponding to the cross-linked oligonucleotides was observed at the origin of the T_1-fingerprint. Unless one allows the highly unlikely possibility of a "triple" cross-link involving the AACG at positions 65-68, the only reasonable interpretation of this result is that indicated in Fig. 4, in which the cross-linked complex in the T_1-fingerprint is AUUGAACG (28-35) linked to CCUAACACAUG (47-57). The fact that G (31) was not hydrolysed by ribonuclease T_1 already hints that this nucleotide is one component of the cross-link. Secondary digestion of the cross-linked spot at the origin of the T_1-fingerprint with ribonuclease A liberated AAC, AC, 2G, 2AU, 2U, C and an undigested residue; this is consistent with a cross-link from G (31) to C (47 or 48), which would give precisely these products, leaving GAAC (31-34) linked to C (47 or 48) as the undigested residue. Analysis of the ribonuclease A fingerprint showed the presence of the oligonucleotides AGAU (26-29), AGGC (44-47) and AAC, which only occurs once in this region, at positions 50-52 (cf. Fig. 4). These data are also consistent with the above cross-link assignment, and the presence of AGGC proves further that C (47) could not be involved in the cross-link, which must therefore be between G (31) and C (48). The cross-linked oligonucleotide should appear on the A-fingerprint as GAAC (31-34) linked to C (48), and an extra spot was indeed observed whose mobility was consistent with this composition.

Complex B: This complex was found in the combined venom plus ribonuclease H digest, as illustrated in Fig. 2, and the T_1-fingerprint of the complex is shown in Fig. 3. Examination of the latter indicates that two distinct regions of the 16S RNA are involved, namely positions 1-72 (similar to the region encompassed by complex A, above), and 1022-1095 (approx.). The 5'-end of the first of these regions was defined by the presence of the 5'-oligonucleotide of 16S RNA (positions 1-6), and the 3'-end by the presence of two moles of AACG (cf. complex A) and the absence of UAACAG (70-75) on the T_1-fingerprint. GGU (68-70) was present in the A-fingerprint, and thus the same venom nuclease cut at position 72 is involved, as was the case with

complex A. The 5'-end of the second RNA region was localised to position 1022 by the presence of a spot corresponding to pAUG in the T_1-fingerprint, which is again an established venom nuclease site (20). The 3'-end of this region is approximately at position 1095, as indicated by the presence of UUAAG (1090-1094) and absence of UCCCG (1095-1099) from the T_1-fingerprint, and the presence of AAGU (1092-1095) in the A-fingerprint. The latter position is most likely a ribonuclease H cut, since GGGUUAAG (1087-1094) is complementary to CTTCCC, if the two U-residues are looped out.

However, in complex B no missing oligonucleotides or extra spot corresponding to the cross-link could be observed, and this raises the possibility that the two RNA fragments (both 72-73 bases long) were simply running together fortuitously in the same spot on the gel (Fig. 2). This possibility was eliminated by analysing the two fragments marked I and II (Fig. 2), which were simple continuous RNA fragments with chain lengths of 88 and 64 residues, respectively; a plot of mobility in the first gel dimension against log (chain-length) shows clearly that complex B cannot conceivably be a fortuitous mixture (Fig. 2, right panel). It follows that the cross-link must involve "uncharacteristic" areas of both sequence regions (e.g. positions 36-42 and 1058-1064), where the absence of a short oligonucleotide from each region would be difficult to detect in either the A- or T_1-fingerprint. The cross-linked pair could also in this case easily be missed in the rather complex region of the T_1-fingerprint near the origin (Fig. 3). Despite our failure to pinpoint the cross-link site, complex B is included for reasons which will be apparent from the Discussion section (see below).

Complex C: A large number of fragments containing cross-link C were found on the two-dimensional gels (cf. Fig. 1), but the actual site of the cross-link could only be determined from analyses of the shorter fragments, of which C* (Fig. 1, right panel) was the most useful. Ribonuclease A- and T_1-fingerprints of this complex are included in Fig. 3, the T_1-fingerprint showing clearly the modified oligonucleotide CCmG (525-527), which is very characteristic for this region of the 16S RNA. The 5'-end of the complex lay within the oligonucleotide CUCAUUG (475-481), as

shown by the presence of AU in the secondary digest of the spot at the origin of the T_1-fingerprint (not indicated in Fig. 3); this would correspond to one of the known cobra venom sites at position 475 or 476 (20). The 3'-end of the complex was at position 574, since $UAAA_{OH}$ (571-574) was present on the T_1-fingerprint, whereas UAAAG (571-575) was absent. A number of possible partial complementarities to CTTCCC occur at this point, as potential ribonuclease H cutting sites. The cross-link site was characterised by an anomalous spot on the T_1-fingerprint (Fig.3), and the absence of CAAG (545-548) together with only a very weak spot corresponding to AAG, although the latter should occur twice in this region (positions 495-500, Fig. 4). The anomalous spot liberated C and an undigestible oligonucleotide upon secondary digestion with ribonuclease A. The A-fingerprint (Fig. 3) showed correspondingly the absence of AAGC (546-549), and the cross-linked oligonucleotide appeared as a rather smeared spot whose position was close to that of the normal ribonuclease A product AGAAGAAGC (493-501). However this spot liberated AG, C and an undigested residue, but no AAG, on secondary digestion with ribonuclease T_1, and it follows that the oligonucleotide is in fact AGAAGAAGC linked to AAGC, with the actual site of cross-linking being from G (497) to AAG (546-548). In other analyses some AAG was liberated during the secondary digestion of this oligonucleotide, suggesting a possible heterogeneity at the cross-link site, involving one or other of the four A-residues surrounding G (497).

Complex D: This complex (Fig. 1, left panel) always appeared in rather long RNA fragments, and the yield was not sufficiently high to allow a further partial digestion (cf. ref. 4) to be attempted. Typically the RNA fragment concerned covered the region between positions 1020 to 1260 (Fig. 4), although the positions of the 5'- and 3'-ends could not be precisely determined in such a long RNA fragment. The cross-link site was characterised by the absence of UUAAG (1090-1094) from the T_1-fingerprint (not shown in Fig. 3), and the presence of only one mole of CCAG, which should occur twice in this region (at positions 1128-1131 and 1161-1164). Analysis of the A-fingerprint showed that the latter (1161-1164) was the missing CCAG (Fig. 4), since

AGU (1163-1165) was reproducibly entirely absent. (There is a potential cobra venom cutting point at position 1162 (20), and the absence of CCAG and AGU could be explained if the venom nuclease had cut within the cross-linked complex at this position; such a cut would however have led to the appearance of pAG in the T_1-fingerprint, but this oligonucleotide was never observed). The fingerprints and secondary digestion products of complex D were rather complicated in the region near the origin where the cross-linked oligonucleotide (UUAAG linked to CCAG) would be expected to appear, and as a result we were not able to demonstrate the presence of this cross-linked oligonucleotide with certainty. We therefore regard the cross-link site assignment in this case as tentative.

Complex E: As with cross-link D, cross-link E (Fig. 1, left panel) occurred in very long RNA fragments, covering positions ca. 1030 to 1360. Again, due to the length of the fragments, the 5'- and 3'-ends could not be precisely determined, but one component of the cross-link site was already clearly shown by the absence of ACCUCAUAAAG (1280-1290) from the T_1-fingerprint. Unlike complex D, complex E was always obtained in high yield, which enabled us to carry out second partial digestions using ribonuclease T_1, as described in Materials and Methods (cf. ref. 3). These second partial digestions gave several shorter versions of the cross-linked complex (cf. Fig. 4, upper panel), and the T_1-fingerprint of the shortest is shown in Fig. 3. The RNA in the latter fingerprint arises from two distinct regions, the first covering approximately positions 1109-1153 and the second positions 1273-1308. As before (3,4) the 5'- and 3'-ends of the fragments from such partial T_1-digests are usually heterogeneous and therefore difficult to define precisely; (for this reason the spot on the fingerprint (Fig. 3) corresponding to AUUG (1306-1309) and UCUG (1313-1316) is relatively weak). Nevertheless the location of the cross-link site was completely unambiguous, since UUG (1125-1127) and ACCUCAUAAAG (1280-1290) were reproducibly absent from the T_1-fingerprint. The cross-linked product ran together with CAACCCUUAUCCUUUG (1109-1124) at the origin of the fingerprint (Fig. 3), and secondary digestion with ribonuclease A showed the presence of C, U, G,

AU, AAC and AAAG in the expected molar ratios, together with an undigested residue. AC was however clearly absent, which indicates that the cross-link is between UUG (1125-1127) and AC (1280-1281); the latter assignment was further confirmed by the absence of GGAC (1278-1281) from the A-fingerprint of the original RNA complex.

Complex F: Complex F (Fig. 1, left panel) comes from a region of the 16S RNA covering positions ca. 1160-1400. Again, the 5'- and 3'-ends were not precisely determined, but the position of the cross-link site was clear from the T_1-fingerprint, since CUACAAUG (1234-1241) and UAG (1298-1300) were both reproducibly absent, UAG occurring only once in the sequence region concerned. Although the fingerprint was quite complex (and is therefore not included in Fig. 3), a spot corresponding to the cross-linked oligonucleotide was observed, which yielded C, U, G, AC and AG and an undigested residue upon secondary digestion with ribonuclease A, but no AAU. Thus the cross-link site is between AAU (1238-1240) and U (1298). The A-fingerprint gives no additional information in this instance.

DISCUSSION

As before (3,4), we have described the determination of each cross-link site in some detail, because with a view to three-dimensional model-building studies it is important to define the precision of each determination and to know where there are possible ambiguities. The positions of all the cross-links A-F in the secondary structure of 16S RNA (21) are indicated in Fig. 5. The Figure also includes the cross-links that we previously reproducibly identified (1) in irradiated 30S subunits, and the two data sets serve to underscore the point made in the Introduction, namely that the choice of partial digestion conditions has a decisive influence on the spectrum of cross-links observed. None of the previous cross-links (1) was seen in the present series of experiments, although it is likely that some of them were present in the longer fragments containing multiple cross-links (cross-link C plus a further cross-link; see Results and Fig. 1, left panel), which we did not analyse in detail.

Of the cross-links described here, C and F are secondary

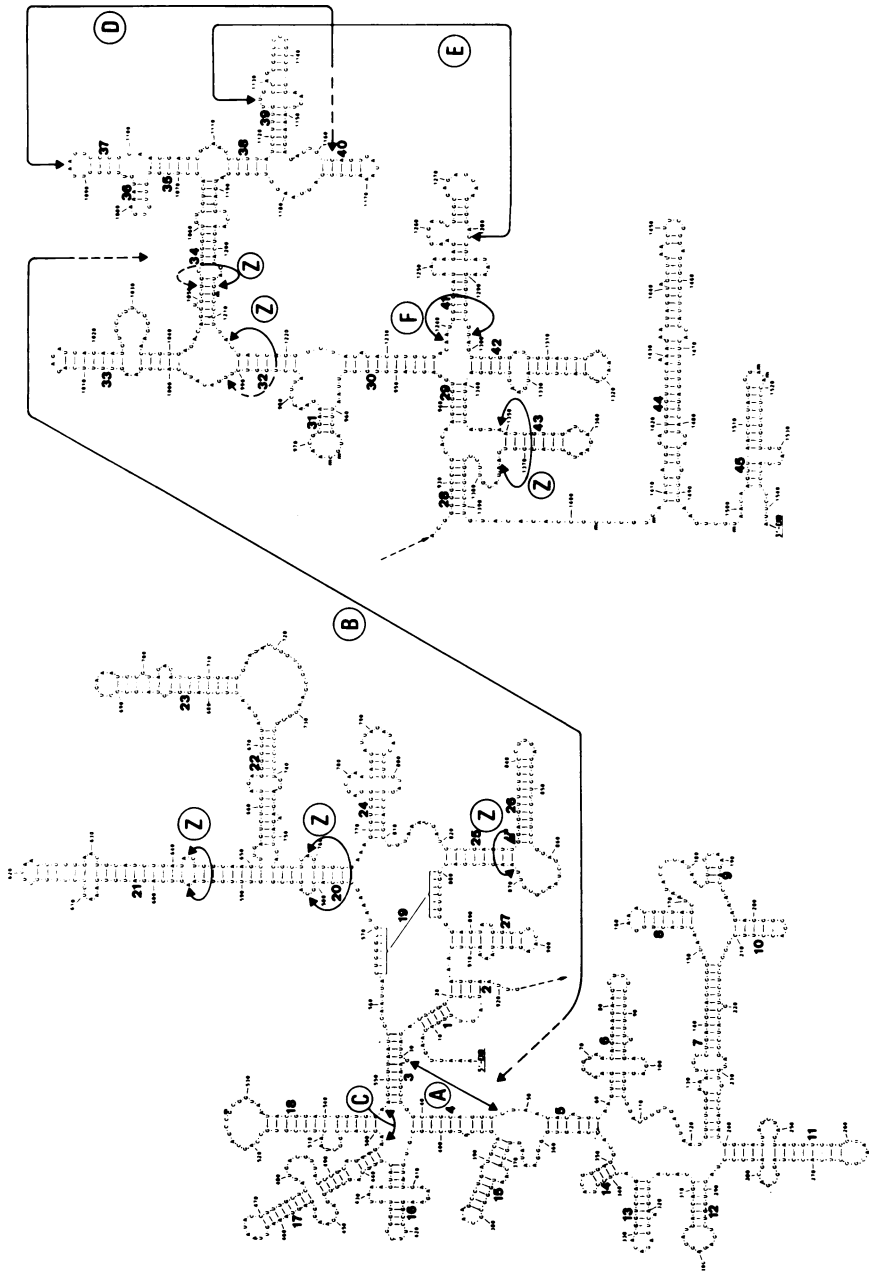

Figure 5: Location of the cross-links in the secondary structure of 16S RNA (21). Cross-links A to F are indicated by the appropriate letter, whereas previously published cross-links (1) are denoted by the letter Z. Dashed lines indicate uncertainties in the determination of the cross-link site concerned (cf. text and Fig. 4).

structural cross-links (Fig. 5). Cross-link C is noteworthy for the very high yield in which it is formed, and also for the fact that both components of the cross-link site involve AAG. It is thus a purine-purine cross-link, and a very similar situation was observed in the 50S subunit (4), in which AAG (602-604) and AAG (654-656) of the 23S RNA were linked together by ultraviolet irradiation. Further the latter cross-link was also a secondary structural cross-link in the 23S RNA as well as being the most abundant cross-link found in the 50S subunit. Cross-link F (Fig. 5) is noteworthy for the fact that it involves positions 1238-1240 of the 16S RNA, and these are the positions that have also been found cross-linked to protein S7 both by irradiation of 30S subunits (22) and by reaction with 2-iminothiolane (23). This seems therefore to be a "hot spot" in the 16S RNA.

The most interesting cross-links are however the tertiary cross-links A, B, D and E (Fig. 5). Cross-link A is a rather short-range link, and it is not immediately obvious from the secondary structure whether the two components of the cross-link site (G-31 and C-48) can be brought into close proximity to one another. Accordingly, we have built a detailed model of helices 3 and 4, in which the two helices are folded back so that they are antiparallel, and then tilted so that their axes are at roughly 45° to one another; this does indeed readily bring G-31 and C-48 into contact, with the looped-out A-residue (397) in helix 4 pointing towards the major groove of helix 3. The cross-link thus imposes a very precise restriction on the relative orientations of helices 3 and 4, and it is clear from the central positions occupied by these helices in the secondary structure that this restriction will be of key importance in any three-dimensional model of the 16S RNA.

Cross-link B, although we were unable to localise it within its two RNA components, is also very relevant to the three-dimensional topography of the 16S RNA, as it indicates that regions from helices 1-6 and 33-37 must come into contact within the 30S particle. It is these two regions which were previously proposed to be involved in intramolecular "switching" in the RNA (24), and a more recent examination (25) has shown that parts of the proposed switch structures appear to be phylogenetically con-

served. These alternative base pairings or switches involve bases 35-37 with 1067-1069 (tentative) and bases 388-395 with 1058-1066 (25). Cross-link B shows that these sequence regions are at least neighbours (cf. Fig. 5), and is the first intra-RNA cross-link to be found *in situ* in the 30S subunit linking the "head" region (bases 926-1391, see ref. 26 for review) with the "body" of the subunit.

Cross-links D and E both lie within the "head" region, and, particularly cross-link E, provide important restrictions on the possible orientations of helices 30 through 41. Combined with the known sites of cross-linking to protein S7 in helices 41 (22) and 43 (23), the data begin to give a picture of the detailed three-dimensional arrangement of the RNA in the head of the 30S subunit (see ref. 25 for review). A corresponding series of intra-RNA cross-links induced by treatment of 30S subunits with *bis*-(2-chloroethyl)-methylamine places similar restrictions on the topographical arrangement of the RNA in the "body" region (J. Atmadja, W. Stiege and R. Brimacombe, unpublished results).

Once again, as in the Introduction, we emphasize that all of the cross-links described here were induced *in situ* within the ribosomal subunits, and should be contrasted with the work of other research groups, who are currently attempting to investigate the tertiary structure of the ribosomal RNA by cross-linking of isolated RNA (5,8,9; see also refs. 25 and 26 for a discussion of this point). The latter studies have identified a series of tertiary cross-links, but there is as yet no way of telling which of these are relevant to the structure of the 16S RNA *in situ* in the subunit, and which are isolational artefacts. In any cross-linking study, the possibility of specific artefacts can of course never be entirely excluded, but in the case of the ultraviolet-induced cross-linking this possibility can be reduced still further by performing the irradiation on growing *E. coli* cells as opposed to ribosomal subunits. A series of such experiments has been made in our laboratory (W. Stiege, J. Atmadja and R. Brimacombe, manuscript in preparation), and the cross-links so far found by the "*in vivo*" cross-linking method agree precisely with those described here, as well as with those already reported for the 50S subunit (4).

ACKNOWLEDGEMENTS

We thank M. Becker, W. Heikens, J. Hollatz and H. Krause for their help in the synthetic work, and Dr. H.G. Wittmann for his critical reading of the manuscript.

REFERENCES

1 Zwieb, C. and Brimacombe, R. (1980) Nucleic Acids Res. 8, 2397-2411
2 Glotz, C., Zwieb, C., Brimacombe, R., Edwards, K. and Kössel, H. (1981) Nucleic Acids Res. 9, 3287-3306.
3 Stiege, W., Zwieb, C. and Brimacombe, R. (1982) Nucleic Acids Res. 10, 7211-7229
4 Stiege, W., Glotz, C. and Brimacombe, R. (1983) Nucleic Acids Res. 11, 1687-1706
5 Wollenzien, P., Hearst, J.E., Thompson, P. and Cantor, C.R. (1979) J. Mol. Biol. 135, 255-269
6 Turner, S., Thompson, J.F., Hearst, J.E. and Noller, H.F. (1982) Nucleic Acids Res. 10, 2839-2849
7 Turner, S. and Noller, H.F. (1983) Biochemistry 22, 4159-4164
8 Thompson, J.F. and Hearst, J.E. (1983) Cell 32, 1355-1365
9 Expert-Bezançon, A., Milet, M. and Carbon, P. (1983) Eur. J. Biochem. 136, 267-274
10 Vassilenko, S.K. and Ryte, V.C. (1975) Biokhimiya 40, 578-583
11 Donis-Keller, H. (1979) Nucleic Acids Res. 7, 179-192
12 Mankin, A.S., Skripkin, E.A., Chichkova, N.V., Kopylov, A.M. and Bogdanov, A.A. (1981) FEBS Lett. 131, 253-256
13 Möller, K. and Brimacombe, R. (1975) Mol. Gen. Genet. 141, 343-355
14 Darlix, J.L. (1975) Eur. J. Biochem. 51, 369-376
15 Chattopadhyaya, J.B. and Reese, C.B. (1980) Nucleic Acids Res. 8, 2039-2053
16 Köster, H., Blöcker, H., Frank, R., Geussenhainer, S. and Kaiser, W. (1978) Liebigs Ann. Chem., 839-853
17 Yuki, A. and Brimacombe, R. (1975) Eur. J. Biochem. 56, 23-34
18 Volckaert, G. and Fiers, W. (1977) Analyt. Biochem. 83, 228-239
19 Brosius, J., Palmer, M.L., Kennedy, P.J. and Noller, H.F. (1978) Proc. Natl. Acad. Sci. USA 75, 4801-4805
20 Vassilenko, S.K., Carbon, P., Ebel, J.P. and Ehresmann, C. (1981) J. Mol. Biol. 152, 699-721
21 Maly, P. and Brimacombe, R. (1983) Nucleic Acids Res. 11, 7263-7286
22 Zwieb, C. and Brimacombe, R. (1979) Nucleic Acids Res. 6, 1775-1790
23 Wower, I. and Brimacombe, R. (1983) Nucleic Acids Res. 11, 1419-1437
24 Glotz, C. and Brimacombe, R. (1980) Nucleic Acids Res. 8, 2377-2395
25 Brimacombe, R., Atmadja, J., Kyriatsoulis, A. and Stiege, W. (1986) Texas Ribosome Conference, Springer-Verlag, in press
26 Brimacombe, R., Maly, P. and Zwieb, C. (1983) Progr. Nucleic Acid Res. 28, 1-48