Volume 10 Number 22 1982 Nucleic Acids Research

Precise localisation of three intra-RNA cross-links in 23S RNA, and one in 5S RNA, induced by treatment of *Escherichia coli* 50S ribosomal subunits with *bis*-(2-chloroethyl)-methylamine

Wolfgang Stiege, Christian Zwieb and Richard Brimacombe

Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, Berlin-Dahlem, FRG

Received 7 September 1982; Accepted 14 October 1982

ABSTRACT

Treatment of E. coli 50S ribosomal subunits with low doses of bis-(2-chloroethyl)-methylamine ("nitrogen mustard") leads to formation of a number of intra-RNA and RNA-protein cross-links. After partial digestion of the cross-linked subunits with cobra venom nuclease, followed by destruction of the protein moiety with proteinase K, complexes containing the intra-RNA cross-links were isolated by two-dimensional gel electrophoresis. The individual complexes were subjected to oligonucleotide analysis, either directly or after a second partial digestion procedure using ribonuclease T_1, and the cross-link sites determined. In 23S RNA, the cross-links found were between bases 763 and 1567, 1210 and 1236, 1482 and 1501; in 5S RNA, base 69 was cross-linked to base 107. The significance of these cross-links in relation to the three-dimensional organization of the ribosomal RNA is discussed.

INTRODUCTION

The last two to three years have seen the appearance of several secondary structure models, derived from various types of experimental data, for both the 16S and 23S ribosomal RNA from E. coli (1-6). In the case of the 16S RNA, the models have already undergone some refinement, with the result that the latest versions of the structure from all of the research groups concerned are in very substantial agreement (1-3). In the case of the 23S RNA, three independent secondary structure models appeared for the first time during 1981 (4-6), and, although there are areas of discrepancy, here again the degree of agreement between the proposed structures is very encouraging. (The models for both 16S and 23S RNA are compared in detail in a review which will be appearing shortly (7)).

The next phase in the derivation of the structure of the ribosomal RNA is to determine how these secondary structures

0305-1048/82/1022-7211$2.00/0

(which are essentially two-dimensional models of the RNA) are folded into three dimensions in the compact ribosome. Here the most direct and obvious approach is to induce intra-RNA cross-links *in situ* in the ribosomal subunits, and to localise the sites of cross-linking. Some progress has already been made in this direction. Cantor and his colleagues (8,9) have described the partial localisation by electron microscopy of several psoralen-induced cross-links in the *E. coli* 16S RNA, and one of these cross-links has subsequently been precisely identified (10). We have described precise localisations of a number of cross-links in both 16S (11) and 23S (4) RNA, induced by mild ultraviolet irradiation of *E. coli* 30S or 50S subunits. However, all of these latter cross-links, as well as the psoralen cross-link (10) just mentioned, lie within elements of the secondary structure. As a result, although these cross-links are very useful in confirming and extending the secondary structure models, they have not so far yielded the type of information needed to develop a tertiary structure model. We have suggested (11) that this is a consequence of the partial hydrolysis conditions used in the isolation of the cross-linked RNA complexes; the hydrolyses usually involve a mild treatment with ribonuclease T_1, which favours the selection of stable secondary structure elements as digestion products, and therefore only those cross-links which occur in such stable elements will be observed.

In this paper we describe the analysis of a set of intra-RNA cross-links induced in the *E. coli* 50S subunit by *bis*-(2-chloroethyl)-methylamine ("nitrogen mustard"). This compound has been known for many years to induce cross-links in DNA (12), and is slightly more reactive than its non-methylated counterpart *bis*-(2-chloroethyl)-amine, which we have previously shown to induce both RNA-protein (13) and intra-RNA (14) cross-links in ribosomal subunits (as does the corresponding sulphur mustard compound (15)). In order to identify the sites of cross-linking, we have made use of the double-strand specific nuclease from cobra venom (16) in order to generate a set of suitable RNA fragments. This enzyme has been used very successfully (2,5) to generate specific cuts in ribosomal RNA, and of course gives a spectrum of RNA fragments which is entirely different to that

obtained with ribonuclease T_1. The cross-linked fragments are isolated by two-dimensional polyacrylamide gel electrophoresis, in a modification of the system previously described (11), and the cross-link sites are identified by oligonucleotide analysis. The mustard compounds are known to react specifically with the N-7 position of guanosine residues (17), which helps in these identifications, and we describe the localisation of three cross-links in 23S RNA and one in 5S RNA. One of the 23S RNA cross-links is a clear tertiary structural cross-link, between two remote regions of the secondary structure.

MATERIALS AND METHODS

Preparation of ribosomal subunits

100 ml cultures of E. coli MRE 600 were labelled with 20 mCi of ^{32}P-orthophosphate. After harvesting, the cells were washed with 10 mM Tris-HCl pH 7.8, 1 mM magnesium acetate, and were then disrupted by sonication in 3 ml of this buffer. 1 µl of deoxyribonuclease I (1 mg/ml) was added, and cell debris was removed by centrifugation for 15 min at 10,000 rpm. The supernatant was loaded immediately onto 10 - 40% sucrose gradients containing 0.3 mM magnesium acetate, 50 mM KCl, 6 mM 2-mercaptoethanol and 10 mM Tris-HCl pH 7.8 for separation of the ribosomal subunits (18). This rapid and simplified method considerably reduces the danger of the ^{32}P-labelled RNA (particularly the 23S RNA) becoming partially degraded during the subunit isolation procedure. After centrifugation, the 30S and 50S subunits were precipitated with ethanol, resuspended in buffer and subjected to an activation dialysis as previously described (19). Finally they were dialysed against several changes of cross-linking buffer (5 mM magnesium chloride, 50 mM KCl, 6 mM 2-mercaptoethanol, 25 mM triethanolamine-HCl pH 7.8 (cf. 13)).

Cross-linking and partial digestion of 50S subunits

^{32}P-labelled 50S subunits at a concentration of 10 A_{260} units/ml in cross-linking buffer (0.5 ml) were treated with 3 mM bis-(2-chloroethyl)-methylamine for 60 min at 37°. Unreacted reagent was destroyed by a further incubation for 15 min at 37° with 6 mM cysteamine hydrochloride, and the cross-linked sub-

units were then precipitated with ethanol and resuspended in 10 mM magnesium chloride, 300 mM KCl, 10 mM Tris-HCl pH 7.8. After pre-incubation for 10 min at 37°, a suitable quantity of cobra (Naja oxiana) venom nuclease (16) was added (see Results), and incubation was continued for 60 min at 37°. The hydrolysis was stopped by addition of EDTA to a concentration of 40 mM, followed by a brief warming (5 min) to 60°. Next, the digested subunits were again precipitated with ethanol, and protein was destroyed by incubation for 15 min at 37° with proteinase K (0.1 mg/A_{260} unit RNA) in 0.5 ml of a buffer consisting of 10 mM NaCl, 5 mM EDTA, 0.5% sodium dodecyl sulphate and 10 mM Tris-HCl pH 7.8. After a final ethanol precipitation the samples were redissolved in a small volume of first dimension gel buffer and were loaded onto two-dimensional polyacrylamide gels for separation of the cross-linked RNA fragments.

Two-dimensional gel electrophoresis

The two-dimensional gel system was a modification of that previously described (11). The first dimension was run on a 40 cm long slab gel (0.1 cm thick), consisting of a linear 3 - 15% polyacrylamide gradient (acrylamide: methylene-bis-acrylamide ratio 19:1). The gel gradient was stabilised by addition of 5% sucrose to the 15% acrylamide solution. The gel buffer contained 7 M urea, 2.5 mM EDTA, 0.1% sodium dodecyl sulphate and 50 mM Tris-HCl pH 7.8, with 0.4% dimethylaminopropionitrile and 0.03% ammonium persulphate as catalysts (cf. 20). The reservoir buffer was 2.5 mM EDTA, 0.1% dodecyl sulphate and 50 mM Trirs-citric acid pH 8.8, as before (11). After electrophoresis, the RNA fragments were located by autoradiography, and strips of 15 cm x 1 cm were cut from the upper and lower parts of the gel, respectively. These strips were soaked in second dimension gel buffer containing 2-mercaptoethanol before polymerizing into the second dimension gel (40 cm x 20 cm x 0.1 cm), exactly as previously described (11). The second dimension gel was run in the system of Maxam and Gilbert (21) as before (11), using a 10% gel for the upper strip from the first dimension gel, and a 20% gel for the lower strip.

Isolation and second partial digestion of cross-linked complexes

Intra-RNA cross-linked complexes were located by autoradio-

graphy, cut from the gel, and extracted in the presence of phenol as previously described (20). After ethanol precipitation, some of the complexes (see Results) were subjected to a second partial digestion, which was carried out in 50 µl of 10 mM magnesium chloride, 300 mM KCl, 10 mM Tris-HCl pH 7.8, together with 50 µg of unlabelled carrier tRNA. The samples were preincubated for 10 min at 37°, then 75 ng of ribonuclease T_1 was added and the mixture incubated for 30 min at 0°. Enzyme was destroyed by proteinase K treatment as above, the proteinase being added to the sample at 0°. The digested RNA was precipitated with ethanol, and loaded onto a two-dimensional gel as above, using a 20 cm long 10% polyacrylamide gel for the first dimension and a 40 cm long 20% gel for the second dimension. The cross-linked complexes were again located by autoradiography and extracted as above.

Oligonucleotide analysis

Cross-linked RNA complexes (either with or without the second partial digestion procedure) were subjected to oligonucleotide analysis by total digestion with either ribonuclease T_1 or ribonuclease A (11), followed by thin-layer chromatography on polyethyleneimine plates in the two-dimensional "minifingerprint" system of Volckaert and Fiers (22), exactly as previously described (23). The oligonucleotides were eluted and submitted to secondary digestion with ribonuclease A or T_1, and the digestion products separated on polyethyleneimine plates using the "double-digestion" system of Volckaert and Fiers (22), again as described (11,23). The data were fitted to the 23S RNA sequence of Brosius et al. (24), or the 5S RNA sequence of Brownlee et al. (25).

RESULTS AND DISCUSSION

Isolation of cross-linked complexes

The action of cobra venom nuclease on 50S ribosomal subunits generates a highly reproducible set of RNA fragments (cf. 5) ranging in size from about 40 to 400 nucleotides, and this pattern of fragments is stable over a wide range of enzyme concentrations. Accordingly, the enzyme solution (which was a generous gift from Dr. V.A. Erdmann) was used at a dilution which was in the middle of the range shown to yield the stable fragment pat-

tern in preliminary trials with unlabelled 50S subunits. Similar trials with 50S subunits treated with the nitrogen mustard cross-linking reagent showed that the pattern of RNA fragments obtained on one-dimensional gels was largely unaltered by treatment with the reagent at a concentration of 3 mM, whereas concentrations greater than 5 mM caused a progressive smearing of the hydrolysis profiles. A concentration of 3 mM was therefore used in these experiments, the cross-linking reaction with ^{32}P-labelled 50S subunits being carried out as described in Materials and Methods. As noted in the Introduction, the mustard reagents induce both intra-RNA and RNA-protein cross-links in ribosomal subunits, and analysis of the levels of RNA-protein cross-linking (P. Maly, unpublished results) showed that 3 mM bis-(2-chloroethyl)-methylamine has approximately the same effect as a 10 mM concentration of the non-methylated counterpart, which we used in our earlier work (13).

A two-dimensional polyacrylamide gel system for the isolation of intra-RNA cross-linked complexes has already been described (11). In the modified system used in this series of experiments, both gel dimensions contained a strong denaturing agent (7 M urea), the crucial factor for the separation of the cross-linked complexes being the presence of dodecyl sulphate in the first dimension and its absence in the second dimension. While we are unable to offer a convincing theoretical explanation for the separation, the resulting gel patterns (see Figs. 1 and 2) are nevertheless much more clearly defined than those previously obtained (4,11). This factor, combined with the advantage gained by using the venom nuclease as opposed to ribonuclease T_1, results in yields of cross-linked complexes which are 10- to 100-fold higher than those found under the old conditions (4,11).

The two-dimensional gel pattern obtained from venom-digested ^{32}P-labelled 50S subunits after treatment with 3 mM nitrogen mustard is shown in Fig. 1, the two halves of the diagram being the 10% and 20% second dimension gels from the upper and lower regions of the first dimension gel, respectively (see Materials and Methods). The diagonal of "free" RNA fragments can clearly be seen, together with a number of spots, corresponding to putative intra-RNA cross-linked complexes, running above the

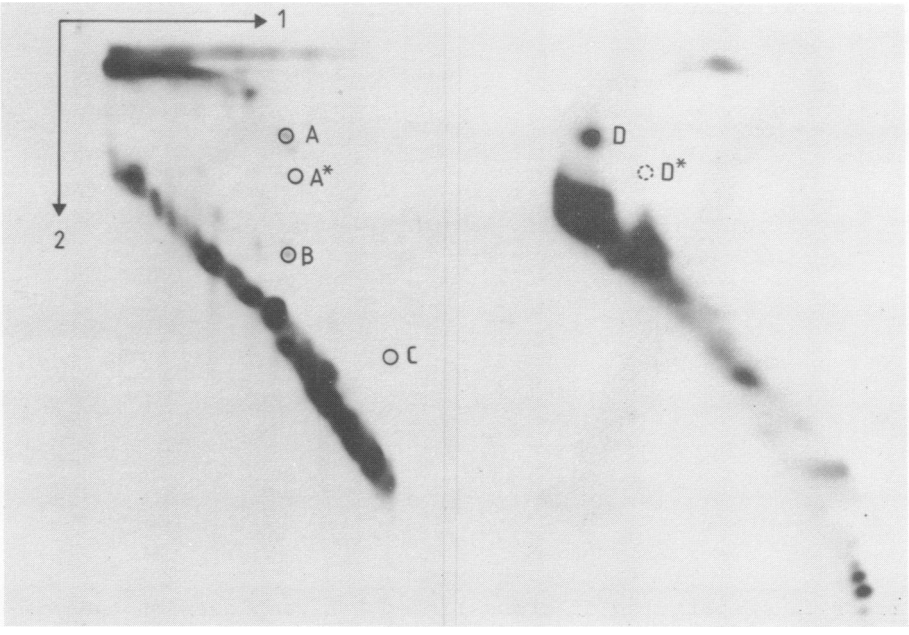

Figure 1: Separation of ^{32}P-labelled cross-linked RNA complexes on two-dimensional gels. The left-hand gel is 10% polyacrylamide and corresponds to the upper region of the first-dimension gel strip (see Materials and Methods), the right-hand gel being 20%, corresponding to the lower region of the first-dimension gel strip. Directions of electrophoresis for the first and second dimensions are indicated. The circled fragments (A to D) are those discussed in the text.

diagonal. These spots were totally absent in non-cross-linked control samples, and, moreover, their pattern was not affected if the proteinase K treatment prior to loading the gel (see Materials and Methods) was omitted. This demonstrates that the putative intra-RNA cross-link spots do not arise from RNA-protein cross-links. RNA-protein complexes, together with larger intra-RNA cross-linked fragments, are presumably contained in the strong smear of radioactivity (also entirely absent in non-cross-linked control samples) which barely enters the second dimension gel at the top left-hand corner of Fig. 1.

Although a number of other well-defined cross-linked spots can also be seen in Fig. 1, the ensuing discussion is confined to the analysis of the complexes marked A, A*, B, C, D, and D*. The other complexes will not be considered further here, since

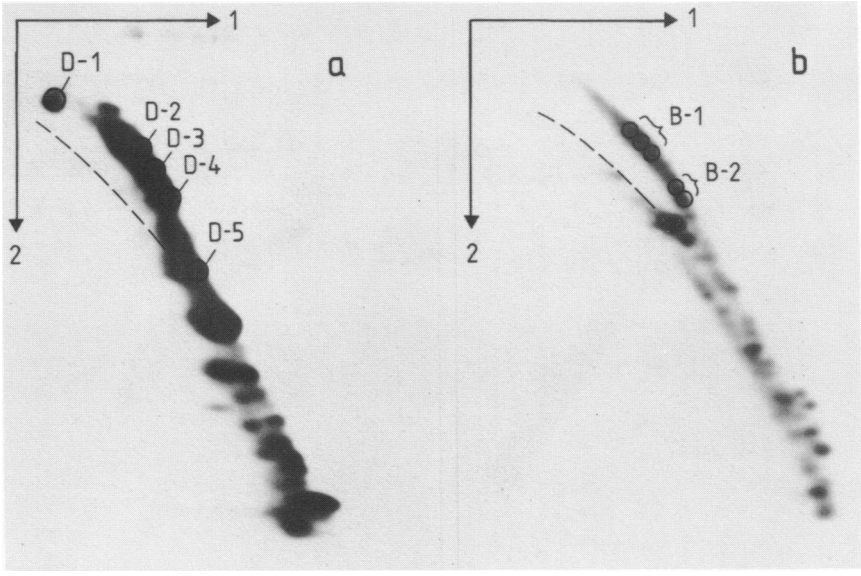

Figure 2: Separation of cross-linked RNA complexes on two-dimensional gels, after a second partial digestion. a. Gel pattern obtained from the partial digest of complex D (Fig. 1). b. Corresponding pattern from complex B (Fig. 1). Directions of electrophoresis for the first and second dimensions are indicated, and the fragments subjected to analysis are marked (see text). The dashed lines denote the continuation of the "diagonal" of free RNA fragments.

an unequivocal analysis of their cross-link sites has not so far been achieved. The strong spot A, and the closely related faint spot A*, both arise from 23S RNA, as do the spots B and C. The strong spot D on the other hand, and the weak spot D*, arise from 5S RNA. Complexes B and D were obtained in large enough amounts (0.8 - 3.5 x 10^6 counts/min) to allow a second partial digestion to be made with ribonuclease T_1, as described in Materials and Methods. The resulting two-dimensional gel patterns are illustrated in Fig. 2, which, like Fig. 1, shows a diagonal of free RNA fragments (indicated in the upper part of each Figure by the faint spots lying on the dashed lines), with the cross-linked fragments lying above this diagonal. It can be seen that as the cross-linked complexes become shorter they eventually merge into the diagonal in the lower part of each Figure. Fig. 2 indicates the complexes which were extracted from these gels for

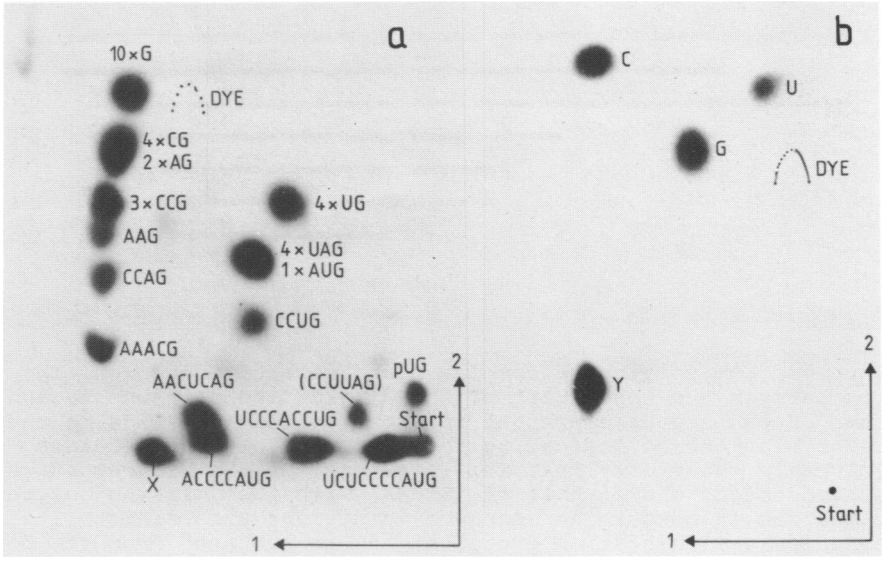

Figure 3: Oligonucleotide analysis of complex D (Fig. 1). a. Ribonuclease T_1 fingerprint of the complex, on a two-dimensional thin layer plate. The identities and molarities of the oligonucleotides are shown by each spot. b. Secondary digestion products with ribonuclease A of spot "X" (Fig. 3a), on a two-dimensional thin layer plate. Directions of chromatography for the first and second dimensions are indicated in each case, as well as the positions of a xylene cyanol dye marker. See Materials and Methods and ref. 22 for details of the systems.

analysis, denoted D-1 to D-5 (Fig. 2a) derived from complex D (Fig. 1), and B-1 and B-2 (Fig. 2b) from complex B (Fig. 1). In the latter case (Fig. 2b) two or three spots were pooled together as indicated to obtain sufficient material for the subsequent oligonucleotide analysis and cross-link site determination.

Analysis of complex D (5S RNA)

We will consider first in some detail the complexes derived from 5S RNA (D and D* in Fig. 1, and D-1 to D-5 in Fig. 2a). Oligonucleotide analysis was carried out as described in Materials and Methods, and the ribonuclease T_1 fingerprint obtained from complex D is shown in Fig. 3. The identities of the oligonucleotides were established by secondary digestion with ribonuclease A, and it can be seen from Fig. 3a that complex D gives a normal

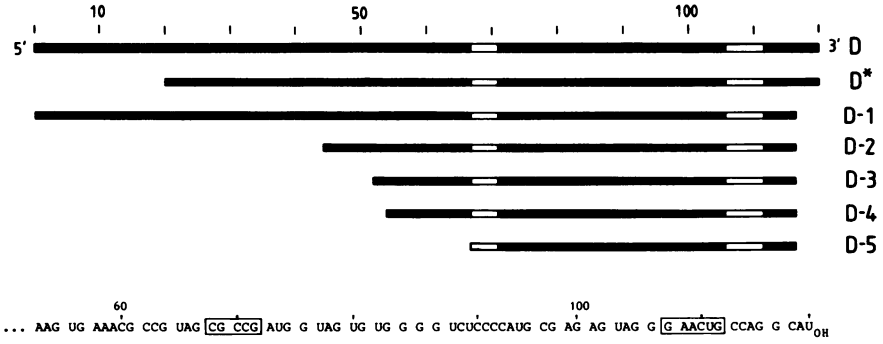

Figure 4: Composition of complex D and related complexes. The solid bars in the upper part of the diagram indicate the regions of the 5S RNA sequence present in each complex (see Figs. 1 and 2), the blank boxes indicating the "missing" oligonucleotides involved in the cross-link. The lower part of the diagram shows the sequence of the 3'-half of 5S RNA, with the "missing" oligonucleotides in boxes. The sequence is divided according to its ribonuclease T1 digestion products, and is numbered from the 5'-end with a dash at every tenth base. The cross-link is between G-69 and G-107, (see text).

5S RNA fingerprint, with the exception that A-A-C-U-G (positions 108-112) is clearly absent, and instead a new spot "X" is observed. (The 3'-terminal oligonucleotide C-A-U_{OH} runs as a smear at the top left-hand corner of the fingerprint in Fig. 3a, and was not usually analysed further. The bracketed oligonucleotide C-C-U-U-A-G arises from the known heterogeneity (25) near the 5'-end of the molecule). Digestion of the spot "X" with ribonuclease A gave the pattern of products shown in Fig. 3b, consisting of the mononucleotides C, G and U in the ratio 2:2:1, together with the anomalous spot "Y".

In the corresponding analyses of the related complexes D* and D-1 to D-5 the same spot "X" was observed in all the fingerprints, and A-A-C-U-G was always absent. The regions of the 5S RNA contained in these complexes are indicated in Fig. 4. Complex D-1 is identical to D (apart from the possible loss of the 3'-terminal C-A-U_{OH}) and complex D* has lost 20 nucleotides from the 5'-end, as a result of a cobra venom nuclease cut at this position, which has also been observed by other authors in isolated 5S RNA (26). Complexes D-2 to D-5 show a progressive loss of material from the 5'-end of the molecule, and in addi-

tion to the absence of A-A-C-U-G it was clear from the analyses of these shorter complexes that the molarity of C-C-G (which occurs 4 times in the intact 5S RNA) was unexpectedly low. For instance, complexes D-3 and D-4 had only one mole of C-C-G, although A-A-A-C-G (positions 57-61) was present, and D-5 contained no C-C-G at all. The oligonucleotide C-C-A-G (positions 113-116) was present in all the complexes. If these findings are compared with the 3'-region of the 5S RNA sequence (Fig. 4, lower half), then it becomes clear that the spot "X" (Fig. 3a) must be the sequence C-G-C-C-G (positions 68-72) cross-linked to G-A-A-C-U-G (positions 107-112), the actual cross-link being from G-69 to G-107, which renders both these G-residues resistant to enzymatic hydrolysis. Bearing in mind that reaction with the nitrogen mustard leads to the addition of a positive charge to the N-7 atom of both cross-linked G-residues (17), the mobility of spot "X" is consistent with a complex of this composition. Further, digestion of this complex with ribonuclease A should yield the mononucleotides 2C, 2G, 1U, and a resistant residue consisting of G-C cross-linked to G-A-A-C; the position and intensity of spot "Y", as well as the observed intensities of the mononucleotides C, G and U in Fig. 3b, are consistent with this conclusion. Furthermore, the "reverse" analysis of complex D (i.e. a ribonuclease A fingerprint followed by secondary digestions with ribonuclease T_1) was also consistent, in that the expected oligonucleotide A-G-G-G-A-A-C (positions 104-110) was missing. At the low level of cross-linking reagent used (3 mM), no interference in the analyses resulting from monovalently reacted mustard molecules was observed.

Analysis of 23S RNA complexes

The analyses of the cross-linked spots arising from 23S RNA (A, A*, B, B-1, B-2 and C (Figs. 1 and 2)) were made in a similar manner, and the results are summarized in Fig. 5. The top part of the Figure shows the locations in the 23S RNA of the sequences contained in the complexes, while the lower part of the Figure gives the detailed sequences in the immediate vicinity of the cross-link sites. For reasons of space, the actual fingerprints and secondary digests of the cross-linked products (cf. Fig. 3) are not shown.

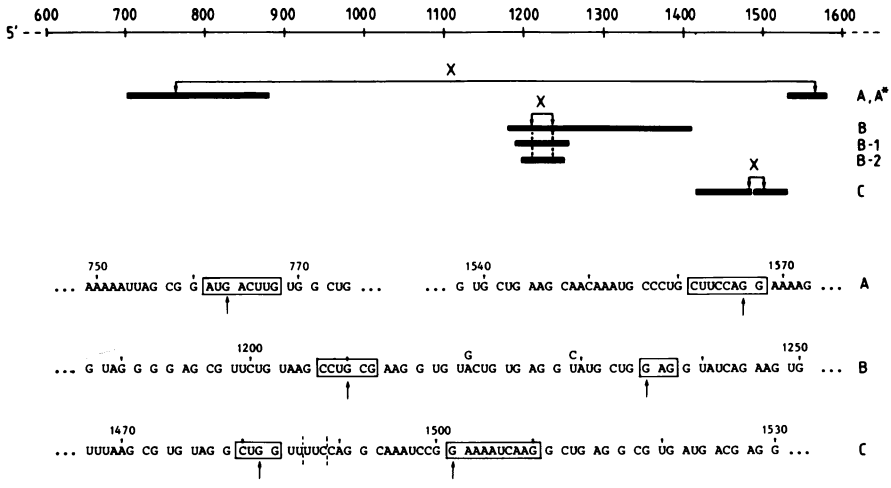

Figure 5: Compositions of complexes A, B, C and their related complexes (Figs. 1 and 2). The solid bars in the upper part of the diagram show the 23S RNA sequence regions contained in the cross-linked complexes, the arrows marked "X" denoting the positions of the cross-links in each case. The sequence is numbered from the 5'-end. The lower part of the diagram shows details of the sequences in the neighbourhood of the cross-link sites (cf. Fig. 4). The "missing" oligonucleotides involved in the cross-links are in boxes, with the sites of cross-linking indicated by arrows. In the sequence encompassing cross-link B, Brosius et al (24) found G at position 1220 and C at position 1229 (see text). In the sequence encompassing cross-link C, the dashed lines at positions 1485 and 1488 indicate the sites of cutting by cobra venom nuclease (see text and ref. 5).

Complex A showed sequences from two distinct regions of the 23S RNA, between bases 703-880 and 1531-1573 (Fig. 5), complex A* giving a virtually identical fingerprint. The 5'- and 3'-ends of these two regions are in reasonable agreement with the positions of cobra venom nuclease cuts observed by Branlant et al. (5). The shorter 3'-region of this complex (bases 1531-1573) was identified by the presence of the two rather characteristic oligonucleotides C-A-A-C-A-A-A-U-G (1547-1555) and A-A-A-A-G (1569-1573). (A-A-A-A-G also occurs 5'-adjacent to an oligonucleotide A-A-A-U-C-A-A-C-C-G at positions 216-230 in the 23S RNA (24), which would have given the same digestion products in our analysis. However, this ambiguity was resolved by the fact that A-A-A-U was found in a ribonuclease A fingerprint of complex A. A-A-A-U occurs only once in such a fingerprint of the whole

region encompassing positions 703-880 and 1531-1573, namely at positions 1551-1554, and this oligonucleotide would not be generated by a ribonuclease A digest of the 216-230 region). From molarity measurements, the oligonucleotide C-C-C-U-G (1556-1560) was present, but the adjacent C-U-U-C-C-A-G was absent, and the oligonucleotide A-G-G-A-A-A-A-G-C (1566-1574) was also absent from the corresponding ribonuclease A fingerprint. This pinpoints the 3'-component of the cross-link site at G-1567, as indicated in Fig. 5.

In the 5'-region of complex A, A-U-G (which should occur only once in the entire region comprising the complex) and the adjacent A-C-U-U-G (761-768) were both clearly absent, indicating G-763 as the other (5'-) component of the cross-link site. The cross-linked oligonucleotide on the ribonuclease T_1 fingerprint should therefore have the composition A-U-G-A-C-U-U-G linked to C-U-U-C-C-A-G-G (Fig. 5), which would be expected to remain at the origin in the fingerprint, by virtue of its high U-content. This rendered it difficult to identify positively, since two other large oligonucleotides from this sequence region also remain at the origin. Nevertheless, a ribonuclease A secondary digestion product was seen (cf. spot "Y", Fig. 3b) which had a mobility consistent with the expected resistant cross-linked residue, G-A-C linked to A-G-G.

Complex B arose from the single sequence region 1180-1410, but the 5'- and 3'-ends could not be determined precisely, and due to the length of the fragment the cross-link site remained obscure. However, analysis of the much shorter complexes B-1 and B-2 (derived from complex B, Fig. 2b) enabled the site to be deduced. The shorter fragment B-2 extended from positions 1196-1250, whereas B-1 was slightly longer in both directions (1188-1256). (The ends of each fragment were rather heterogeneous, as a result of the fact that two or three spots from the gel (Fig. 2b) were pooled to make the analysis). The region was characterized by the presence of the oligonucleotides U-U-C-U-G and U-A-A-G (1198-1206) and U-A-U-C-A-G (1240-1245), thus locating the region containing the cross-link in a well-defined hairpin loop in the 23S RNA (see Fig. 6 and refs. 4-6). However, in addition to an anomalous oligonucleotide corresponding to the

cross-link site, two other oligonucleotides were found which did not fit the published sequence of 23S RNA (24) in this region, namely (U,U,A-C)-G and (U,A-U)-G. The same anomalous products were found in this region in other unrelated experiments with non-cross-linked material (C. Glotz and R. Brimacombe, unpublished data). If one assumes a compensating G---C to A---U base-pair substitution at positions 1220 and 1229 within the hairpin loop structure (4), then oligonucleotides U-A-C-U-G and U-A-U-G would be expected at positions 1219-1223 and 1229-1232, respectively, and the data then fit the sequence perfectly. (A comparison of the E. coli and Z. mays chloroplast sequences in this region of the 23S RNA (4) shows that the sequence is not highly conserved here, and the Strasbourg group (5) has reported several heterogeneities in the E. coli sequence in this general area of the molecule.)

Making this assumption, then the observed absence of C-C-U-G (1203-1206) pinpoints G-1206 as one component of the cross-link, since, as already mentioned, the adjacent U-A-A-G (Fig. 5) was present. Inspection of the remaining sequences present in the fingerprints and the rather clear secondary structure of this region indicated that G-1236 is the only reasonable candidate for the other component of the cross-link. This would result in the cross-linked oligonucleotide in the ribonuclease T_1 fingerprint being C-C-U-G-C-G linked to G-A-G, which would yield the mononucleotides 2C, 1G, 1U, together with G-C linked to G-A-G as the resistant residue. The position of the cross-linked oligonucleotide on the fingerprint and its secondary digestion products were entirely consistent with this expectation (cf. Fig. 3b). It cannot of course be excluded that the 3'-component of the cross-link is the next G-residue (1238), but inspection of the secondary structure shows that this is unlikely.

Complex C extended from position 1416 to an established cobra venom site (5) at position 1530. However, another strong cobra venom cutting point has been reported at positions 1485 or 1488 (5), and indeed the oligonucleotide U-U-U-U-C-C-A-G (1484-1491, Fig. 5) was absent. Instead an oligonucleotide corresponding to pU-U-C-C-A-G was found, indicating that the cobra venom cut was in this case at position 1485. It follows that the

two components of the cross-link site must be one on either side of this position, in order to hold both halves of the complex C sequence region together. The 3'-side of the cross-link was readily determined, since the characteristic oligonucleotide A-A-A-A-U-C-A-A-G (1502-1510) was absent. The cross-link could therefore be either G-1501 or G-1510 (Fig. 5), but since A-A-G and not A-A-A-A-U was found in the ribonuclease A secondary digest of the cross-linked oligonucleotide, the site must be G-1501.

The 5'-component of the cross-link involves the C-U-G sequence (1480-1482), since only 1 mole of C-U-G was observed in the ribonuclease T_1 fingerprint; C-U-G occurs twice in the sequence region comprising complex C, at positions 1480-1482 and 1512-1514, and therefore if one of these two C-U-G sequences is involved in the cross-link (as indicated by the molarity) then it must be the former, in view of the cobra venom cut at position 1485 as already mentioned above. The actual cross-link site could therefore be either at G-1479 or G-1482, and must in fact be the latter, since no mononucleotide G was observed in the ribonuclease A secondary digest of the cross-linked oligonucleotide. The cross-linked product in the ribonuclease T_1 fingerprint should in this case have the composition C-U-G-G linked to G-A-A-A-A-U-C-A-A-G, yielding 2C, 1U, A-A-G, and G-G linked to G-A-A-A-A-U as secondary digestion products. This was indeed consistent with the observed ribonuclease A analysis.

Conclusions

The positions of the cross-links described above in the second and third domains of our secondary structure model for 23S RNA (4) and in the 5S model of Studnicka et al. (27) are indicated in Fig. 6. Cross-link A is certainly the most important, as it is between two G-residues which are far apart in both the primary and secondary structure. These G-residues are in single-stranded regions in all three of the recently proposed secondary structure models (4-6). The position of cross-link B must be regarded as not proven (see the above discussion), but is clearly located (in contrast to cross-link A) within the secondary structure in the 1200-1250 region. It should be noted that if the cross-link is indeed at the positions indicated, then this pre-

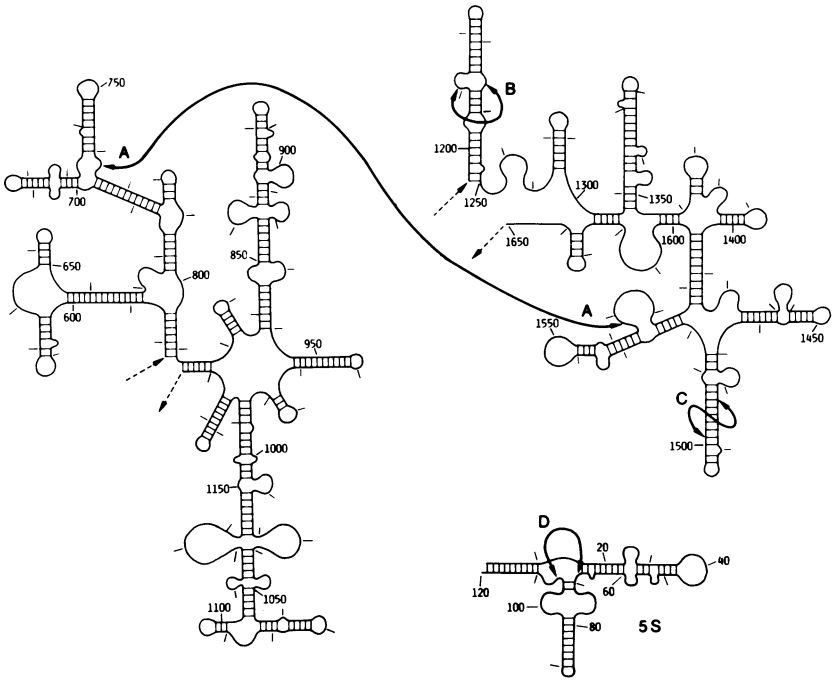

Figure 6: Location of the cross-links in the secondary structures of 23S and 5S RNA. The diagrams are schematic representations of the second and third domains of the 23S secondary structure (4) and of 5S RNA (27). The sequences are numbered from the 5'-end of each molecule, with a dash at every tenth base. The positions of the cross-links found in complexes A, B, C and D are indicated by the arrows.

cludes the possibility that the upper and central helical regions of this loop (Fig. 6) are continuously stacked, a fact which will have important consequences for future model building of interrupted helices of this nature.

Cross-link C is of yet another type, being between two strands of a double-helical region. Nitrogen mustard is known to cross-link double-helical DNA (12), and simple model-building indicates that the two G-residues indicated lie directly above one another in the helical structure. This region is a double-helix in all three of the proposed secondary structures (4-6), and a similar type of cross-link in the stem region of isolated 5S RNA was observed by Wagner and Garrett (28). Cross-link D in the 5S RNA is different to this latter cross-link, and differs

also from other established cross-links in 5S RNA (29,30). The cross-link is consistent with the model of Studnicka et al. (27), bridging the base of the hairpin loop encomapssing bases 70-106. Detailed model building may show it to be consistent with other 5S models (e.g. 31), but the important point to note here is that cross-link D was generated in situ in 50S subunits, as opposed to isolated 5S RNA.

Still not clear to us is the question of why and to what exeent the cross-linked fragments are retarded in the second dimension of our gel system (Figs. 1 and 2). An understanding of this property would have value in predicting the type of cross-link contained in a particular complex, and would help in the subsequent oligonucleotide analysis. As can be seen from the foregoing discussion, the determination of the precise cross-link sites is not straightforward, and relies on detailed argumentation from the oligonucleotide data in each individual case. It is unfortunate that end-labelling techniques cannot be applied here, since the cross-linked complexes (in particular those such as complex A) contain two 5'- and 3'-ends. Our strategy for the future is to develop better conditions for the second partial digestion procedure (cf. Fig. 2), such as the use of defined deoxynucleotides and ribonuclease H (cf. 32), in order to range in towards the cross-link site in a more controlled manner. It is clear from Fig. 1, and from similar experiments with 30S subunits as well as with both 30S and 50S subunits cross-linked by mild ultraviolet irradiation (W. Stiege, unpublished data), that a considerable number of intra-RNA cross-links still remain to be determined from these simple cross-linking systems. Some preliminary identifications of tertiary structural cross-links similar to complex A above have already been made from ultraviolet-irradiated subunits, and we are convinced that further analyses of this type will play an important role in building up a three-dimensional model of the ribosomal RNA.

ACKNOWLEDGEMENTS

We are very grateful to Dr. V.A. Erdmann for his gift of cobra venom nuclease, and to Dr. H.G. Wittmann for his continued support and encouragement.

REFERENCES

1. Noller, H.F. and Woese, C.R. (1981) Science 212, 403-411.
2. Stiegler, P., Carbon, P., Zuker, M., Ebel, J.P. and Ehresmann, C. (1981) Nucleic Acids Res. 9, 2153-2172.
3. Zwieb, C., Glotz, C. and Brimacombe, R. (1981) Nucleic Acids Res. 9, 3621-3640.
4. Glotz, C., Zwieb, C., Brimacombe, R., Edwards, K. and Kössel, H. (1981) Nucleic Acids Res. 9, 3287-3306.
5. Branlant, C., Krol, A., Machatt, M.A., Pouyet, J., Ebel, J.P., Edwards, K. and Kössel, H. (1981) Nucleic Acids Res. 9, 4303-4324.
6. Noller, H.F., Kop. J., Wheaton, V., Brosius, J., Gutell, R.R., Kopylov, A.M., Dohme, F., Herr, W., Stahl, D.A., Gupta, R. and Woese, C.R. (1981) Nucleic Acids Res. 9, 6167-6189.
7. Brimacombe, R., Maly, P. and Zwieb, C. (1983) Progr. Nucleic Acid Res., in the press.
8. Wollenzien, P., Hearst, J.E., Thammana, P. and Cantor, C.R. (1979) J. Mol. Biol. 135, 255-283.
9. Wollenzien, P. and Cantor, C.R. (1982) J. Mol. Biol., in the press.
10. Turner, S., Thompson, J.F., Hearst, J.E. and Noller, H.F. (1982) Nucleic Acids Res. 10, 2839-2849.
11. Zwieb, C. and Brimacombe, R. (1980) Nucleic Acids Res. 8, 2397-2411.
12. Geiduschek, E.P. (1961) Proc. Nat. Acad. Sci. USA 47, 950-955.
13. Ulmer, E., Meinke, M., Ross, A., Fink, G. and Brimacombe, R. (1978) Mol. Gen. Genet. 160, 183-193.
14. Zwieb, C., Ross, A., Rinke, J., Meinke, M. and Brimacombe, R. (1978) Nucleic Acids Res. 5, 2705-2720.
15. Malbon, R.M. and Parish, J.H. (1971) Biochim. Biophys. Acta 246, 542-552.
16. Vassilenko, S.K. and Ryte, V.C. (1975) Biokhimiya 40, 578-583.
17. Brookes, P.B. and Lawley, P.D. (1961) Biochem. J. 80, 496-503.
18. Morgan, J. and Brimacombe, R. (1972) Eur. J. Biochem. 29, 542-552.
19. Zwieb, C. and Brimacombe, R. (1979) Nucleic Acids Res. 6, 1775-1790.
20. Yuki, A. and Brimacombe, R. (1975) Eur. J. Biochem. 56, 23-34.
21. Maxam, A.M. and Gilbert, G. (1977) Proc. Nat. Acad. Sci. USA 74, 560-564.
22. Volckaert, G. and Fiers, W. (1977) Analyt. Biochem. 83, 228-239.
23. Maly, P., Rinke, J., Ulmer, E., Zwieb, C. and Brimacombe, R. (1980) Biochemistry 19, 4179-4188.
24. Brosius, J., Dull, T.J. and Noller, H.F. (1980) Proc. Nat. Acad. Sci. USA 77, 201-204.
25. Brownlee, G.G., Sanger, F. and Barrell, B.G. (1968) J. Mol. Biol. 34, 379-412.
26. Douthwaite, S. and Garrett, R.A. (1981) Biochemistry 20, 7301-7307.

27. Studnicka, G.M., Eiserling, F.A. and Lake, J.A. (1981) Nucleic Acids Res. 9, 1885-1904.
28. Wagner, R. and Garrett, R.A. (1978) Nucleic Acids Res. 5, 4065-4075.
29. Rabin, D. and Crothers, D.M. (1979) Nucleic Acids Res. 7, 689-703.
30. Hancock, J. and Wagner, R.A. (1982) Nucleic Acids Res. 10, 1257-1269.
31. Pieler, T. and Erdmann, V.A. (1982) Proc. Nat. Acad. Sci. USA 79, 4599-4603.
32. Donis-Keller, H. (1979) Nucleic Acids Res. 7, 179-192.