Volume 11 Number 6 1983 Nucleic Acids Research

Localisation of a series of intra-RNA cross-links in the secondary and tertiary structure of 23S RNA, induced by ultraviolet irradiation of *Escherichia coli* 50S ribosomal subunits

Wolfgang Stiege, Carola Glotz and Richard Brimacombe

Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, D-1000 Berlin-Dahlem, FRG

Received 20 January 1983; Accepted 23 February 1983

SUMMARY

Intra-RNA cross-links were introduced into *E. coli* 50S ribosomal subunits by mild ultraviolet irradiation. The subunits were partially digested with cobra venom nuclease, and the cross-linked RNA complexes were isolated by two-dimensional electrophoresis. Many of the complexes were submitted to a second partial digestion procedure. Oligonucleotide analysis of the RNA fragments obtained in this manner enabled cross-links between the following ribonuclease T1 oligonucleotides in the 23S RNA to be established: positions 292-296 and 339-350; 601-604 and 652-656; 1018-1022 and 1140-1149; 1433-1435 and 1556-1560; 1836-1839 and 1898-1903; 2832-2834 (tentative) and 2878-2885; 2849-2852 and 2865-2867 (tentative); 739-748 and 2609-2618; 571-577 and 2030-2032; 1777-1792 (tentative) and 2584-2588. The first seven of these cross-links lie within the secondary structure of the 23S RNA, whereas the last three are tertiary structural cross--links. The degree of precision of the individual determinations was variable, depending on the nucleotide sequence in the vicinity of the cross-link site concerned.

INTRODUCTION

In this laboratory we have for several years been interested in the development of intra-RNA cross-linking techniques, as a tool for investigating the three-dimensional structure of ribosomal RNA *in situ*. In 1978 we showed (1) that such cross-links could be introduced into *E. coli* ribosomal subunits either by direct ultraviolet irradiation, or by treatment with *bis*-(2-chloroethyl)-amine (cf. 2), and subsequently we were able to locate a number of the ultraviolet-induced cross-link sites in both 16S (3) and 23S (4) RNA. However, all of these cross-links were within elements of the secondary structure of the RNA (cf. also ref.5) and only recently (6) has it been possible to identify a "tertiary structural" cross-link, that is to say a cross-link between regions of the RNA which are remote from each other in both the

© IRL Press Limited, Oxford, England.

primary and secondary structures. In the latter experiments (6), the cross-links were induced in 50S subunits by treatment with bis-(2-chloroethyl)-methylamine, the subunits were digested with cobra venom nuclease (7), and the intra-RNA cross-linked complexes were then isolated by two-dimensional gel electrophoresis (cf. 3). Some of the isolated complexes were submitted to a second partial digestion procedure, followed by a further two-dimensional electrophoretic separation. Oligonucleotide analysis of the products enabled three secondary structural cross-links to be identified (two in 23S and one in 5S RNA), in addition to the tertiary structural cross-link (in 23S RNA) just mentioned.

In this paper, we describe the application of similar methodology to 50S subunits containing intra-RNA cross-links induced by mild ultraviolet irradiation. Here the two-dimensional gels obtained after digestion of the cross-linked subunits with cobra venom nuclease (cf. 6) showed very reproducible patterns ,with many well-defined cross-linked complexes appearing in high yield. The isolated complexes were subjected to further partial digestions, using either ribonuclease T_1, or cobra venom nuclease, or ribonuclease H in conjunction with oligodeoxynucleotides of defined sequence (cf. 8-10). As a result, it was possible in a number of cases to isolate progressively smaller complexes containing the cross-links, and hence to localize the cross-link sites with increasing precision in the subsequent oligonucleotide analyses. Seven secondary and three tertiary structural cross-links were identified in this way.

MATERIALS AND METHODS

Preparation, irradiation and partial digestion of 50S subunits. ^{32}P-labelled 50S ribosomal subunits from *E. coli* strain MRE 600 were prepared by the rapid method of Stiege et al (6). The subunits (ca. 10 A_{260} units, 3 x 10^9 counts/min) were irradiated for 3 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 previously described (11). 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 oxiana*) venom nuclease (7) was added (cf. 6). 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 (6).

Two-dimensional gel electrophoresis, and second partial digestion of cross-linked complexes. The venom-digested 50S subunits were separated in the two-dimensional gel system which has been published in detail (6). The cross-linked complexes were located on the gel by autoradiography, and were extracted in the usual manner (12). Aufter ethanol precipitation, selected complexes were subjected to further digestion in the presence of 50 µg of unlabelled carrier tRNA, by one of three methods: (a) Digestion with ribonuclease T_1. This was carried out as previously described (6), followed by the usual proteinase K treatment to destroy the enzyme. (b) Digestion with cobra venom nuclease. This was carried out as outlined above and in ref. 6 for the first digestion of the cross-linked 50S subunits, with the exception that the amount of enzyme used was reduced ten-fold. The reaction was stopped by the addition of EDTA, etc., as above. (c) Digestion with ribonuclease H. This was carried out in 10 µl of 40 mM Tris-HCl pH 7.8, 50 mM KCl, 10 mM $MgCl_2$, 0.1 mM EDTA, 0.1 mM dithiothreitol (cf. 13), in the presence of 0.05 A_{260} units each of d-(C-T-T-C-C-C), d-(C-G-T-T-G-C) and oligo-$(dT)_8$ together with a suitable quantity of ribonuclease H (see Results). The oligodeoxynucleotides were obtained from PL Biochemicals, Inc. Ribonuclease H was prepared by the method of Darlix (13), omitting the final purification step on DEAE-cellulose, and the enzyme was used at the minimum concentration found to induce an oligonucleotide-dependent hydrolysis. Incubation was for 16 hr at 4°, after which the enzyme was destroyed by proteinase K treatment. The digested samples were in all cases separated by a second two-dimensional gel electrophoresis procedure, exactly as described (6).

Oligonucleotide analysis. Cross-linked RNA complexes were digested completely with either ribonuclease T_1 or ribonuclease A as before (3,6), using the "mini-fingerprint" system of Volckaert and Fiers (14) to separate the products. Molarity determinations and secondary digestions of the oligonucleotides were also made as before (3,6,12), using the "double-digestion" system of

Volckaert and Fiers (14). In some cases "tertiary" digestions were carried out using ribonuclease T_2. For this purpose, material which was resistant to both ribonuclease A and T_1 treatment was eluted from the thin-layer plates and digested at pH 4.5 with ribonuclease T_2 as described (15), the products being separated on polyethyleneimine plates in the one-dimensional pyridine-acetic acid system of Bernardi (16). Oligonucleotide data were fitted to the 23S RNA sequence of Brosius et al (17).

RESULTS

A typical two-dimensional gel pattern of the RNA fragments obtained by cobra venom digestion of irradiated 50S subunits is shown in Fig. 1. The gel shows the usual features (cf. 3,6), namely a diagonal of free RNA fragments, with the cross-linked complexes appearing as spots running above this diagonal. As before (6), control experiments with non-cross-linked subunits showed no spots above the diagonal, and the pattern of cross-linked complexes was not altered if the proteinase K treatment prior to loading the gel (see Materials and Methods, and ref. 6) was omitted. Thus, the complexes do not contain cross-linked protein. The intra-RNA cross-linking reaction induced by ultraviolet irradiation of 50S subunits is already known to proceed considerably more rapidly than the concomitant RNA-protein cross-linking reaction (1), and the irradiation time of 3 minutes used in these experiments (see Materials and Methods) causes cross-linking of only 1.5% of the total 50S protein to 23S RNA (18, and cf. 11). Under these conditions, the 50S subunits retain 55% of their biological activity (18).

The pattern of principal cross-linked complexes in Fig. 1 was very reproducible, and the results described in the following sections will all be referred to in terms of this particular gel pattern. The key diagram in Fig. 1 indicates the nomenclature used for the various complexes, each individual cross-link being designated by a letter. Thus, complexes A1, A2 and A3 all contain the same cross-link, but the RNA fragments involved are of differing lengths in each case. The yields of the individual complexes extracted from the gel varied from a few tens of thousands up to several million counts/min of ^{32}P-RNA.

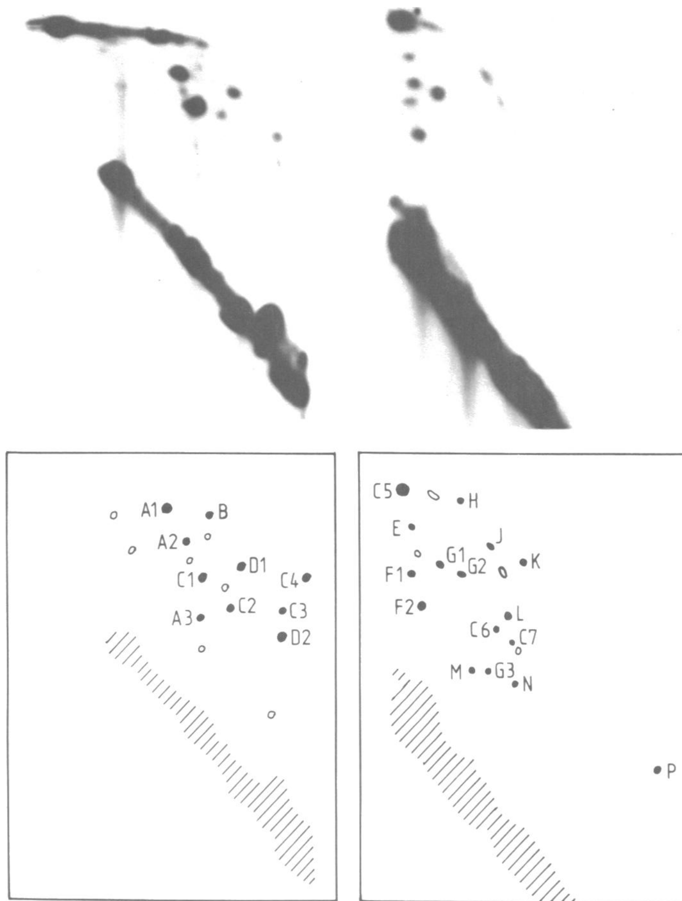

Figure 1: Separation of 32P-labelled cross-linked RNA complexes on two-dimensional gels. The left-hand gel is 10% polyacrylamide and corresponds to the upper region of the first-dimension gel strip (see ref. 6), the right-hand gel being 20%, corresponding to the lower region of the first-dimension gel strip. Directions of electrophoresis are from left to right (first dimension) and top to bottom (second dimension). The key diagrams beneath the autoradiograms indicate the nomenclature used for the cross-linked complexes.

Oligonucleotide analyses were made at this stage from aliquots of the isolated complexes, and then, wherever sufficient radioactivity was present, the remainder of each individual complex was submitted to one of the second partial digestion pro-

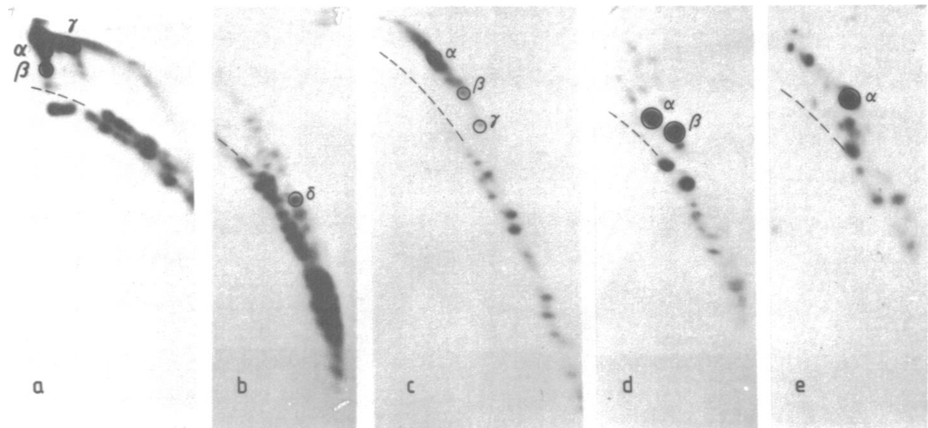

Figure 2: Two-dimensional electrophoresis of cross-linked complexes, after a second partial digestion (cf. Fig. 1). The greek letters indicate the digested complexes used for further analysis, and the dashed lines denote the continuation of the "diagonal" of free RNA fragments. a: Ribonuclease H digest of complex A2 (Fig. 1). b to e: Ribonuclease T1 partial digests of complexes A2, C5, D1 and G1 (Fig. 1), respectively.

cedures, as described in Materials and Methods. Some examples of the two-dimensional gel autoradiograms obtained after the second digestions are illustrated in Fig. 2. As before (6), these gels show a lower diagonal of non-cross-linked RNA fragments, with an upper diagonal or diagonals (e.g. Fig. 2a) corresponding to the progressively digested cross-linked complex; the upper diagonal tends to merge into the lower as the residual cross-linked complex becomes smaller. Fig. 2 also indicates which digestion products were extracted from these gels for further analysis, each being designated by a greek letter (e.g. A2-ß (Fig. 2a) is a second partial digestion product derived from fragment A2 (Fig. 1)).

Most of the second partial digestions were made with ribonuclease T_1 (cf. 6), but, since this enzyme tends to lead to a rather heterogeneous mixture of products (see e.g. Fig. 2b and c), we have also investigated the use of cobra venom nuclease or ribonuclease H for this purpose. The venom nuclease gives an apperently non-specific pattern of smeared digestion products when applied to intact ribosomal RNA as opposed to ribosomal subunits; in contrast, when the enzyme is used to digest the

isolated cross-linked complexes (which are much smaller than the intact RNA), discrete and well-defined digestion products are obtained (data not shown, but see the description of complexes C and G, below).

The application of ribonuclease H (see e.g. Fig. 2a) needs more explanation. Oligodeoxynucleotides of defined sequence have already been used as templates for the site-directed hydrolysis of ribosomal RNA by ribonuclease H (9,10), and it appears that only four complementary base-pairs are needed in the RNA-DNA hybrid duplex in order to be recognized by the enzyme (8). We have investigated the properties of the hydrolysis in more detail by making use of the fact that in the secondary structure of 23S RNA (4) there a number of stretches of four or five contiguous A-residues in looped-out regions. Thus, oligomers of thymidylic acid could be used as templates for testing the ribonuclease H hydrolysis. Trial experiments showed that, even at 4°, oligo-$(dT)_6$ was the shortest template which produced any hydrolysis, and that oligo-$(dT)_8$ was much more efficient, leading to six well-defined cuts in the 23S RNA (Glotz and Brimacombe, unpublished results). Since the cutting sites do not contain eight contiguous A- or G-residues which could pair with oligo-$(dT)_8$, the increased efficiency of the latter must be due to the stabilising effect of "dangling bases" on the RNA-DNA hybrid formation (cf. 19). In combination with the results of Donis-Keller (8), these data suggest that oligodeoxynucleotides of six bases or less must be able to form at least one or two G-C pairs in order to be an effective template for ribonuclease H. Accordingly, a computer search was made of the single-stranded regions of 23S RNA, with a view to selecting the most commonly-occurring tetra-nucleotide sequences, rich in G or C. These sequences were then extended to hexanucleotides (each hexanucleotide naturally incorporating three tetranucleotide sequences at positions 1-4, 2-5 and 3-6), and on this basis the sequences d-(C-T-T-C-C-C) and d-(C-G-T-T-G-C) were chosen as being among the most likely to generate a useful distribution of ribonuclease H cuts in the 23S RNA. These two hexanucleotides were used in combination with oligo-$(dT)_8$ as described in Materials and Methods, and indeed led to a highly specific and reproducible pattern of fragments

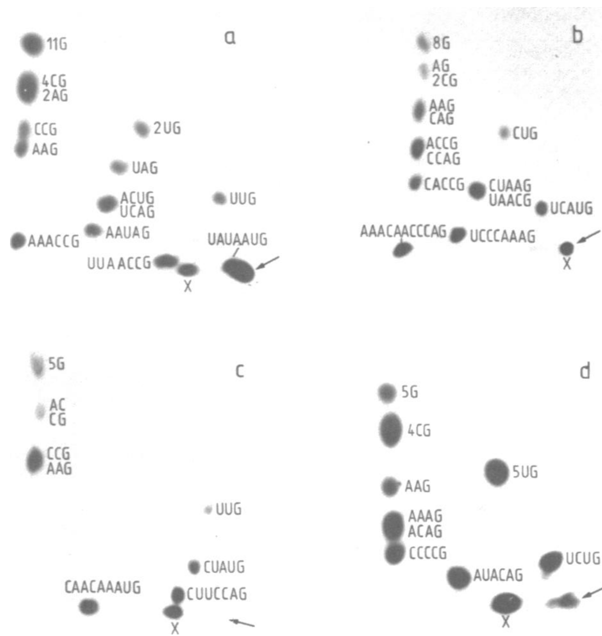

Figure 3: Examples of ribonuclease T1 fingerprints of cross-linked complexes on polyethyleneimine plates (14). The first dimension ran from right to left, the second from bottom to top, the arrows denoting the points of sample application. Identities of the oligonucleotides are shown, "X" denoting the cross-linked oligonucleotide complex in each case. a: Fingerprint of complex C6 (the spot at the origin contained U-A-C-C-U-U-U-G, A-C-U-U-A-U-A-U-U-C-U-G and U-C-U-U-A-A-C-U-G). b: Complex D1-α. c: Complex J-α. d: Complex P (cf. Figs. 1 and 2 for nomenclature).

(see Fig. 2a and the description of complexes A, B and C, below).

All the cross-linked complexes A to P (Fig. 1), together with their second partial digestion products (e.g. Fig. 2) were subjected to oligonucleotide analysis by total digestion with ribonuclease T_1 or A. Secondary, and sometimes tertiary, digestions with ribonuclease A, T_1 or T_2 were made as described in Materials and Methods, and the next section describes very briefly the results of these oligonucleotide analyses for each cross-linked complex. Fig. 3 shows some of the simpler examples from the several hundred oligonucleotide fingerprints made, and Fig. 4 summarizes the locations of the individual complexes in the 23S RNA sequence (17). Fig. 5 gives the region of the RNA

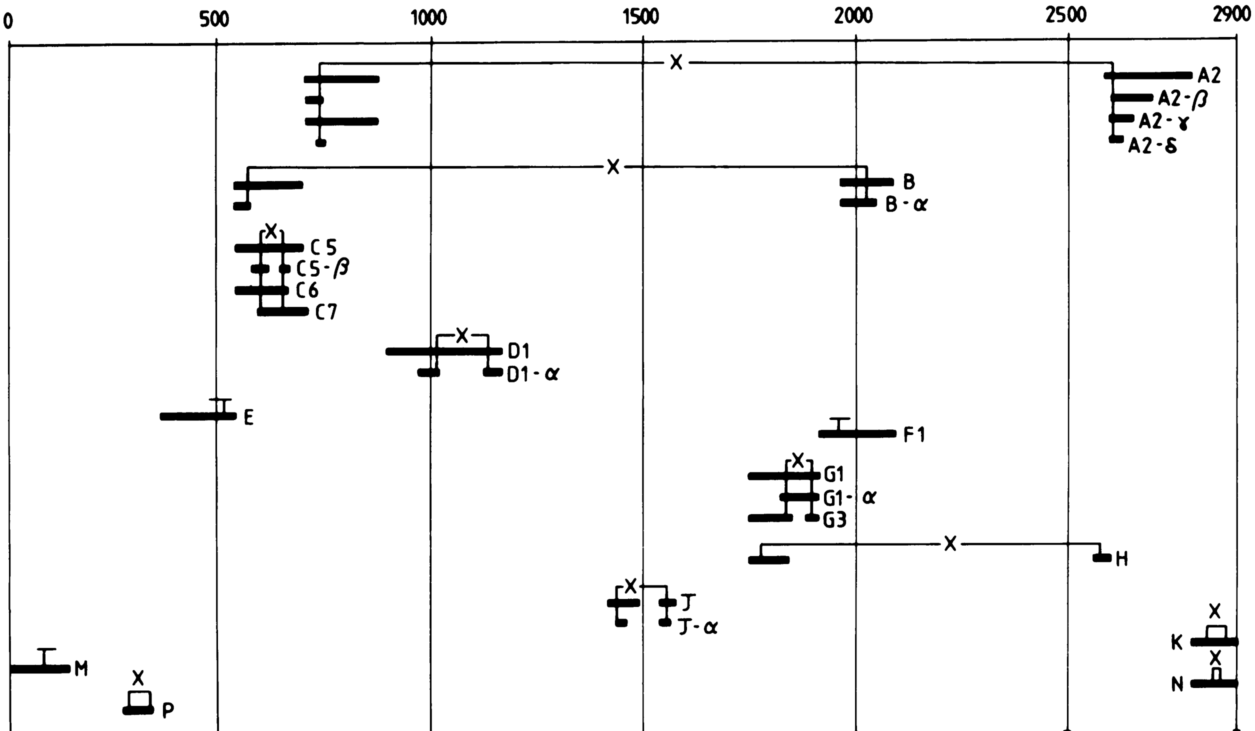

Figure 4: Location of the RNA regions containing the cross-linked complexes in 23S RNA (cf. Figs. 1 - 3). The sequence is numbered from the 5'-end. For clarity, only selected examples from each family of complexes are shown. The thick bars denote the sequence regions concerned, the thin lines marked "X" indicating the cross-link sites. In complexes E, F and M, only one component of the cross-link could be identified (see text).

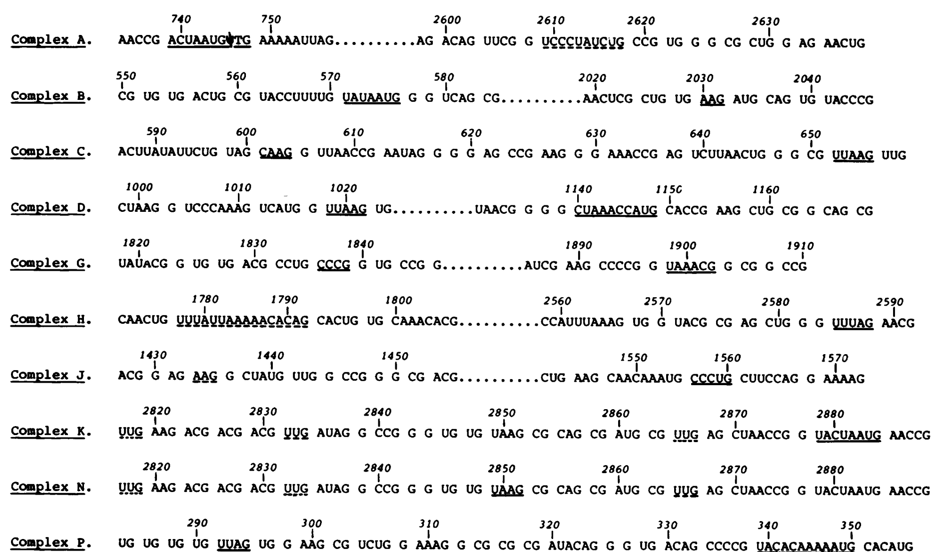

Figure 5: RNA sequences in the vicinities of the cross-link sites. The sequence is divided into the ribonuclease T1 oligonucleotides, and both components of the cross-link sites in complexes A to P are shown, with the cross-linked oligonucleotides underlined (the dotted lines indicating tentative assignments as discussed in the text). Complexes E, F, L and M are not shown.

sequence in the vicinity of each cross-link site, so that the reader can follow the arguments involved in the various site determinations.

Complex A: Three closely-related complexes (A1, A2, A3) appeared on the two-dimensional gel (Fig. 1), and of these the complex A2 was further digested by both ribonuclease T_1 and ribonuclease H (Figs. 2a and b). Analysis of the three original complexes identified the RNA regions containing the cross-link site as being bases 700-900 together with 2600-2800 (see Fig. 4). The cross-link site was localized by sequence analysis of the partial digestion products of complex A2 (examples of which are given in Fig. 4), in particular the small complex A2-δ (Fig. 2b). The 5'-component of the cross-link site lies within the oligonucleotide A-C-U-A-A-U-G-Ψ-T-G (739-748, Fig. 5). This oligonucleotide is not cleaved at position G-745, since the G-residue is modified (20). The oligonucleotide therefore remains at the origin of the ribonuclease T_1 fingerprint, and gives a very

characteristic pattern of secondary digestion products with ribonuclease A, including A-C, A-A-U and G-Ψ. These products were found in the analysis of the fingerprint from complex A2-δ, whereas the neighbouring oligonucleotides A-A-C-C-G and A-A-A-A-A-U-U-A-G (which were both present in the original complex A2, and in A2-γ) were now absent. A ribonuclease H cleavage site caused by oligo-$(dT)_8$ was observed in the latter oligonucleotide in complexes A2-ß and A2-γ (Fig. 4).

The fingerprint of complex A2-δ showed it to contain bases 2600-2630 (Fig. 5). A-G and A-A-C-U-G (2634-2638) were both absent from this complex, although the latter was present in A2-γ (Fig. 4). On the other hand, A-C-A-G (2600-2603) was present in A2-δ, although absent in A2-ß (Fig. 4). The 3'-component of the cross-link site must therefore lie between bases 2603 and 2633. Since the fingerprint of A2-δ contained U-U-C-G, U-G, C-G and C-U-G, in their normal positions, this leaves the decanucleotide U-C-C-C-U-A-U-C-U-G (2609-2618, which remains at the origin of the fingerprint) as the most likely candidate for the cross-link site. The ribonuclease A secondary digest of the oligonucleotides at the origin of the fingerprint of complex A2-δ showed an anomalous spot corresponding to the resistant cross-linked oligonucleotide, in addition to the digestion products (C, G, U, A-C, A-U, A-A-U and G-Ψ) arising from oligonucleotides 739-748 and 2609-2618. However, the single G-residues (2608, 2625, 2626, Fig. 5) and the A-G (2632-2633) cannot be excluded as alternative candidates for the 3'-component of the cross-link site.

Complex B. This complex contained RNA regions 550-700 and 1960-2090 (Fig. 4), and was subjected to further digestion with ribonuclease H. In the 5'-region of the complex, U-A-U-A-A-U-G (571-577) was missing from its normal position in the ribonuclease T_1 fingerprint of the original complex B, and the analysis of complex B-α confirmed that this oligonucleotide must contain the 5'-component of the cross-link site; B-α contained only the RNA region from bases ca. 550-580, and U-A-U-A-A-U-G was again missing. The 3'-end of this fragment was defined here by a ribonuclease H cut due to d-(C-G-T-T-G-C), which shows a five-base complementarity to residues 581-585 (Fig. 5). The 3'-region of complex B-α ran from positions 1970-2046, and although A-A-C-U-

C-G (2019-2024) and U-A-C-C-C-G (2041-2046) were both present, the complex contained no A-A-G. A-A-G should occur only once in the entire RNA region comprising complex B-α, at positions 2030-2032, and this A-A-G sequence (which has a modified A-2030 (29)) is therefore the 3'-component of the cross-link site. A cross-link between the latter trinucleotide and U-A-U-A-A-U-G would be expected to give an anomalous oligonucleotide remaining at the origin of the ribonuclease T_1 fingerprint, and this was indeed observed. The ribonuclease A digest of the material at the origin gave an anomalous spot together with A-A-U and A-U, as well as A-C, U and G (arising from the nonanucleotide at positions 562-570, which normally stays at the origin of the fingerprint).

Complex C: The cross-link contained in complex C has already been described (4), although not documented in detail. It is the most abundant cross-link which we have so far found, and appears in no less than seven variations in the gel of Fig. 1. Several of these complexes were subjected to second partial digestions by all three enzymes (ribonuclease T_1, H and venom nuclease), and in all cases the analyses showed RNA arising from the region including bases 550-670, the cross-link site being clearly characterized by the absence from the fingerprints of C-A-A-G (601-604) and U-U-A-A-G (652-656, see Figs. 3a and 5). In one case (fragment C5-ß, Fig. 2c and 4) the complex was divided into two distinct RNA fragments (cf. complexes A and B, above) by the excision of the central region (from bases 631-647). The cross-linked oligonucleotide appeared as an extra spot in the ribonuclease T_1 fingerprints (marked "X" in Fig. 3a), which gave U, C and an anomalous oligonucleotide upon digestion with ribonuclease A, suggesting that both A-A-G sequences are involved in the actual cross-link.

Complex D: This complex appeared twice in the gel of Fig. 1, and arises from bases 900-1180 of the RNA (Fig. 4). The second partial digestion of complex D1 with ribonuclease T_1 (Figs. 2d, 3b and 4) split this region into two fragments, comprising bases 990-1022 or 975-1022 together with 1132-1168. It was however already clear from the analysis of the parent complex that the cross-link site involved the oligonucleotides U-U-A-A-G (1018-

1022) and C-U-A-A-A-C-C-A-U-G (1140-1149), as both of these were missing from the fingerprint. The fingerprints of complexes D1-α and D1-β both showed as expected an anomalous oligonucleotide remaining at the origin (marked "X" in Fig. 3b), which gave A-A-A-C, A-U, C, U and G together with a ribonuclease-resistant cross-linked product upon digestion with ribonuclease A. This indicates that the actual cross-link site is from A-A-G (1020-1022) to C-U (1140-1141) or to C-1146 (see Fig. 5).

Complexes E and F: These complexes both contained continuous RNA sequences, complex E being from bases 360-540, and complex F from 1920-2090. In each case a single oligonucleotide was missing (A-A-C-C-U-G, 507-512, and C-A-C-G, 1966-1969, respectively), but in neither case was there any clear indication as to the identity of the second component of the cross-link site. Unfortunately the yields of the two complexes were not high enough to permit a second partial digestion, and they will not be discussed further here.

Complex G: Three versions of this complex were found in the gel of Fig. 1, and second partial digestions were made with both venom nuclease and ribonuclease T_1. The RNA region contained in the complex covered bases 1750-1910 (Fig. 4), with a much shorter complex (G1-α, Fig. 2e) being obtained from the second partial digestion with ribonuclease T_1. However, the most well-defined version of this cross-link was found in complex G3, which contained RNA regions 1750-1846 and 1887-1907. In all cases the oligonucleotide C-C-C-G (1836-1839) was missing from the 5'-region of the RNA, and U-A-A-A-C-G (1898-1903) from the 3'-region (Fig. 5). An anomalous spot was found in the ribonuclease T_1 fingerprint, which gave A-A-A-C, C and G on digestion with ribonuclease A, in addition to the resistant cross-linked residue. The actual cross-link must therefore be between U-1898 and one of the three C-residues (1836-1838).

Complex H: This complex appeared as a well-defined spot in the gel of Fig. 1, and was clearly composed of two distinct RNA regions, namely bases 1750-1835 and 2560-2600. The 3'-region of the RNA was not easily discernible from the ribonuclease T_1 fingerprint of the complex, since its 5'-terminus was a cobra venom nuclease cut in the middle of the characteristic oligo-

nucleotide C-C-A-U-U-U-A-A-A-G (2558-2567, Fig. 5). Fortunately however this RNA region contains two highly characteristic ribonuclease A oligonucleotides, namely A-A-A-G-U (2564-2568) and A-G-A-A-C (2587-2591), both of which were present in a ribonuclease A fingerprint of the complex. With this information it was immediately obvious that the sequence U-U-U-A-G (2584-2588) must be the 3'-component of the cross-link site, since this oligonucleotide was missing from the ribonuclease T_1 fingerprint. In the absence of second partial digestion data the 5'-component of the cross-link could not be localized within the region of bases 1750-1835, and the cross-linked oligonucleotide complex remained at the origin of the ribonuclease T_1 fingerprint, together with the large oligonucleotide U-U-U-A-U-U-A-A-A-A-A-C-A-C-A-G (1777-1792). Inspection of the fingerprint data and comparison with the secondary structure of the 23S RNA (see later) suggested that this latter oligonucleotide is in fact itself a likely candidate for the 5'-component of the cross-link, but this must be regarded as no more than a tentative assignment.

Complex J: Complex J was subjected to a second partial digestion with ribonuclease T_1, and both the original complex and the digestion products showed an interrupted RNA sequence in the region of bases 1420-1570. Location of the 3'-component of the cross-link site was simple, since the oligonucleotide C-C-C-U-G (1556-1560) was missing, although the two neighbouring oligonucleotides C-A-A-C-A-A-A-U-G and C-U-U-C-C-A-G (Fig. 5) were both present. Identification of the 5'-component was not so clearcut, since a number of smaller oligonucleotides are possible candidates. In the smallest complex (J-α, Fig. 3c), the 5'-region consisted of approximately bases 1430-1452. One copy of A-A-G was present, but no A-C-G (see Fig. 5). The A-A-G could be either the A-A-G at positions 1433-1435, or that at positions 1544-1546 in the 3'-region of the complex. C-U-G (1541-1543) was not contained in the latter region, but as already mentioned C-A-A-C-A-A-A-U-G was present. If the A-A-G (1544-1546) was the one present in the fingerprint, then A-A-G (1433-1435) is the cross-link site. Alternatively, one of the A-C-G residues flanking the RNA region (1427-1429, or 1453-1455) could be the candidate. Digestion of the cross-linked oligo-

nucleotide ("X" in Fig. 3c) with ribonuclease A liberated only C and an anomalous oligonucleotide, the latter giving A,G and a resistant product after further digestion with ribonuclease T_2. These products are consistent with the involvement of A-A-G or A-C-G in the cross-link.

Complex K: This complex covered a continuous region of RNA from bases 2800 to the 3'-end of the 23S RNA (2904), and the oligonucleotide U-A-C-U-A-A-U-G (2878-2885) was clearly absent. The other component of the cross-link site could not be unequivocally established, but it was noticed that the molarity of U-U-G (which should occur three times in this area (Fig. 5)) was too low. It is therefore plausible that one of these U-U-G residues is the other component of the cross-link. Inspection of the secondary structure (see later) indicates that U-U-G (2832-2834) is the most likely candidate. The cross-linked oligonucleotide remained at the origin of the ribonuclease T_1 fingerprint, and liberated A-C, G and U, but no A-A-U, on digestion with ribonuclease A.

Complexes L and M: In an analogous manner to complexes E and F, both these complexes contained a continuous RNA region, in which only one component of the cross-link site could be discerned. Complex L was 5S RNA, with the A-A-C-U-G sequence near to the 3'-end of the molecule missing (cf. 6). Complex M comprised bases 1-140 of the 23S RNA, with U-A-A-G (82-84) absent. As with complexes E and F, insufficient material was available for a second partial digestion.

Complex N: The RNA region involved here was precisely the same as that in complex K (bases 2800-2904). However, in this case, a different oligonucleotide was absent, namely the U-A-A-G at positions 2849-2852, but the molarity of U-U-G was again too low. Inspection of the secondary structure (see later) suggests that this time the U-U-G at positions 2865-2867 (Fig. 5) might be involved. Consistent with the involvement of U-U-G in the cross--link was the fact that the cross-linked oligonucleotide remained at the origin of the ribonuclease T_1 fingerprint, and gave A-A-G, U and an anomalous oligonucleotide upon digestion with ribonuclease A.

Complex P: The last of the complexes arose from near the

5'-end of the 23S RNA. The 5'-terminus was at base 284, as evidenced by the high U-G content (Figs. 3d, and 5), and the oligonucleotide U-U-A-G (292-295) was missing. At the 3'-end of the complex, no spot corresponding to U-A-C-A-C-A-A-A-A-A-U-G (339-350) was found, and the neighbouring C-A-C-A-U-G was also absent, although C-C-C-C-G (334-338) was present. Instead of 339-350 the spot "X" was observed (Fig. 3d), which liberated A-C, A-G, U and a resistant anomalous oligonucleotide upon digestion with ribonuclease A. This indicates clearly that the 3'-end of the complex lies within the long oligonucleotide (339-350), and that at the same time this oligonucleotide contains the 3'-component of the cross-link site.

DISCUSSION

The cross-links described in this paper provide information which is directly relevant to the detailed three-dimensional organization of the 23S RNA *in situ* in the 50S subunit. As before (6), 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 ambiguities. The ultraviolet-induced intra-RNA cross-links are more difficult to analyse precisely than those induced by nitrogen mustard treatment (6). In the latter case, the cross-links can be reasonably assumed to involve only G-residues, and the cross-linked oligonucleotides tend to have a high mobility in the fingerprint system (cf. Fig. 3) as a result of their extra positive charges (6). In contrast, the photochemical cross-links may involve any nucleotide, but since U-containing oligonucleotides are usually preferred (cf. 3,4), the cross-linked oligonucleotides tend to remain at the origin of the fingerprint, making their analysis difficult. On the other hand, the two-dimensional gel patterns obtained from the ultraviolet cross-linked subunits (cf. Fig. 1) are better-defined than those from the chemically-cross-linked subunits. The use of the second partial digestion procedure has proved to be the key to the successful determination of the ultraviolet-induced cross-links, and in those cases where one component of the cross-link site could not be located unequivo-

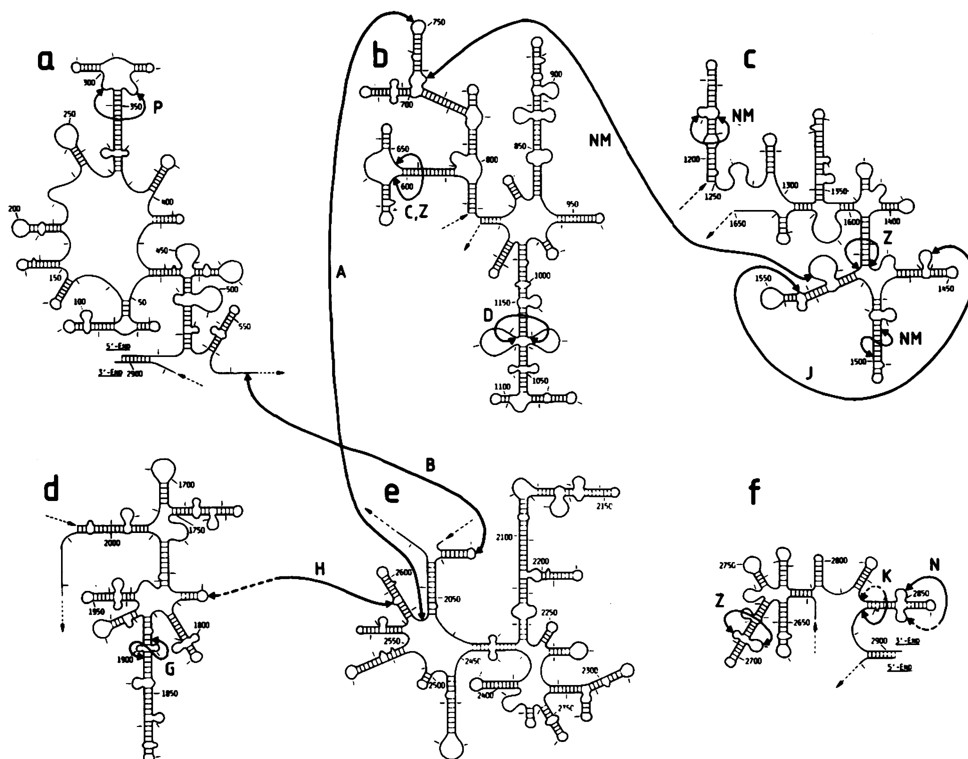

Figure 6: Location of the cross-links in the secondary structure of 23S RNA, the RNA being divided into six domains (a to f) as in ref. 4. The positions of the cross-links found in complexes A to P (cf. discussion in the text in each case) are indicated by the arrows, dotted lines indicating tentative assignments of one component of the cross-link site in complexes H, K and N. Previously identified cross-links are also included, "Z" denoting ultraviolet induced cross-links described by Zwieb et al (4), and "NM" cross-links induced by treatment with nitrogen mustard (6). The secondary structure in the region of bases 2020-2040 has altered to conform to the arrangement of Noller et al (21); this small hairpin loop was overlooked in our original model (4).

cally, the maximum possible error is limited to the size of the RNA fragment contained in the shortest complex after the second partial digestion (e.g. in complex A, the 3'-component must lie between bases 2608 and 2633). Improvements in these digestion conditions, coupled with an extra chromatographic step to separate the oligonucleotide material at the origin of the fingerprints (cf. 3), should enable the precision of the data to be

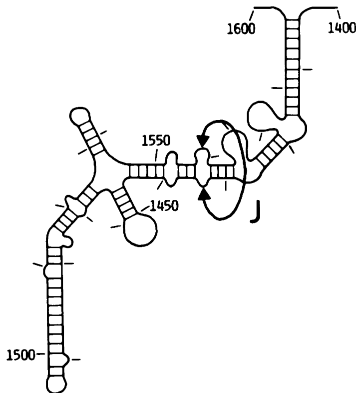

Figure 7: Location of cross-link J, according to the secondary structure model of Noller et al (21), assuming that A-A-G (1433-1435) is the 5'-component of the cross-link site (see text).

increased and further cross-links (e.g. complexes E,F,L and M) to be established.

The positions of the cross-links described here are shown in the secondary structure of 23S RNA (4) in Fig. 6, together with previously identified cross-links (4,6). As before (6), the cross-links fall into two groups, corresponding to "secondary" and "tertiary" structural cross-links. Cross-links A, B, H and possibly J are clearly tertiary structural, and the remainder lie within (and provide confirmatory evidence for) the secondary structure. Cross-link J is a special case, as this involves a region of the RNA where different secondary structures have been proposed (4,20,21). If the model of Noller et al (21) is correct here, and if the 5'-component of the cross-link site (within bases 1430-1450) is A-A-G (see above), then this cross-link falls into the pattern of the other secondary structural cross-links, as illustrated in Fig. 7. All of the latter type of cross-link (induced by ultraviolet irradiation) occur either at the end of a helical region or where the helix is interrupted. As has already been pointed out (6), this has the consequence that looped--out regions seem to prevent continuous stacking of the adjacent helical areas, since such cross-links would not be sterically possible if the base-stacking continued across the interrupted region. This is also true for cross-link G, which is at first sight analogous to the nitrogen mustard cross-link (6) at positions 1482-1501 (Fig. 6); whereas the latter occurs within an uninterrupted helix by bridging via the nitrogen mustard molecule

the existence of cross-link G (being a "zero-distance" ultraviolet cross-link) implies that the stacking must be broken at bases 1840/1900 to enable the cross-link to form. In other words, interruptions in the helices produce "bends" or "kinks" in the structure, which will be important for model-building. The tertiary structural cross-links are however still the most interesting for establishing the three-dimensional arrangement of the RNA, and enough of these are now available (Fig.6) for serious model-building studies to be started.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. V.A. Erdmann for his gift of cobra venom nuclease, to Dr. K. Isono for writing computer programmes for us, and to Dr. H.G. Wittmann for his critical reading of the manuscript.

REFERENCES

1 Zwieb, C., Ross, A., Rinke, J., Meinke, M. and Brimacombe, R. (1978) Nucleic Acids Res. 5, 2705-2720
2 Malbon, R.M. and Parish, J.H. (1971) Biochim. Biophys. Acta 246, 542-552
3 Zwieb, C. and Brimacombe, R. (1980) Nucleic Acids Res. 8, 2397-2411
4 Glotz, C., Zwieb, C., Brimacombe, R., Edwards, K. and Kössel, H. (1981) Nucleic Acids Res. 9, 3287-3306
5 Turner, S., Thompson, J.F., Hearst, J.E. and Noller, H.F. (1982) Nucleic Acids Res. 10, 2839-2849
6 Stiege, W., Zwieb, C. and Brimacombe, R. (1982) Nucleic Acids Res. 10, 7211-7229
7 Vassilenko, S.K. and Ryte, V.C. (1975) Biokhimiya 40, 578-583
8 Donis-Keller, H. (1979) Nucleic Acids Res. 7, 179-192
9 Backendorf, C., Ravensbergen, C.J.C., Van der Plas, J., Van Boom, J.H., Veeneman, G. and Van Duin, J. (1981) Nucleic Acids Res. 9, 1425-1444
10 Mankin, A.S., Skripkin, E.A., Chichkova, N.V., Kopylov, A.M. and Bogdanov, A.A. (1981) FEBS Lett. 131, 253-256
11 Möller, K. and Brimacombe, R. (1975) Mol. Gen. Genet. 141, 343-355
12 Yuki, A. and Brimacombe, R. (1975) Eur. J. Biochem. 56, 23-34
13 Darlix, J.L. (1975) Eur. J. Biochem. 51, 369-376
14 Volckaert, G. and Fiers, W. (1977) Analyt. Biochem. 83, 228-239
15 Zwieb, C. and Brimacombe, R. (1979) Nucleic Acids Res. 6, 1775-1790
16 Bernardi, A. (1974) Analyt. Biochem. 59, 501-507

17 Brosius, J., Dull, T.J. and Noller, H.F. (1980) Proc. Nat. Acad. Sci. USA 77, 201-204

18 Rinke, J., Meinke, M., Brimacombe, R., Fink, G., Rommel, W. and Fasold, H. (1980) J. Mol. Biol. 137, 301-314

19 Romaniuk, P.J. Hughes, D.W., Gregoire, R.J., Neilson, T. and Bell, R.A. (1978) J. Amer. Chem. Soc. 100, 3971-3972

20 Branlant, C., Krol, A., Machatt, M.A., Pouyet, J., Ebel, J.P., Edwards, K. and Kössel, H. (1981) Nucleic Acids Res. 9, 4303-4324

21 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