Selective isolation and detailed analysis of intra-RNA cross-links induced in the large ribosomal subunit of *E. coli*: a model for the tertiary structure of the tRNA binding domain in 23S RNA

Philip Mitchell, Monika Osswald, Dierk Schueler and Richard Brimacombe*
Max-Planck Institut fuer Molekulare Genetik, Abt. Wittmann, Berlin-Dahlem, FRG

Received June 15, 1990; Accepted July 10, 1990

ABSTRACT

Intramolecular RNA cross-links were induced within the large ribosomal subunit of *E. coli* by mild ultraviolet irradiation. Regions of the 23S RNA previously implicated in interactions with ribosomal-bound tRNA were then specifically excised by addressed cleavage using ribonuclease H, in conjunction with synthetic complementary decadeoxyribonucleotides. Individual cross-linked fragments within these regions released by such 'directed digests' were isolated by two-dimensional gel electrophoresis and the sites involved in the cross-links determined using classical oligonucleotide analysis techniques. Using this approach, seven 'new' cross-links could be precisely localised, between positions 1782 and 2608 – 2609, 1940 and 2554, 1941 – 1942 and 1964 – 1965, 1955 and 2552 – 2553, 2145 – 2146 and 2202, 2518 – 2519 and 2544 – 2545, and between positions 2790 – 2791 and 2892 – 2895 in the 23S RNA sequence. These data, in conjunction with data from RNA-protein cross-linking studies carried out in our laboratory, were used to define a model for the tertiary organisation of the tRNA binding domain of 23S RNA *'in situ'*, in which the specific nucleotides associated with tRNA binding in the 'A' and 'P' sites are clustered at the base of the 'central protuberance' of the 50S subunit.

INTRODUCTION

A number of specific nucleotides within the 16S and 23S rRNAs of *E. coli* have been associated with the binding of tRNA, largely through photoaffinity cross-linking (1–4) and chemical footprinting studies (5,6). The availability of models for the three-dimensional organisation of 16S RNA *'in situ'* in the 30S subunit (7,8) have allowed those nucleotides in the 16S RNA to be mapped within the architecture of the small subunit. Although well separated in the primary sequence and secondary structure of the 16S RNA, these nucleotides are localised in two clusters in the three-dimensional model, one on each side of the 'head' of the 30S particle (9). On the other hand, the lack of knowledge concerning the detailed three-dimensional arrangement of the 23S RNA has meant that nucleotides within the RNA of the large subunit identified as being proximal to tRNA sites can only be discussed with respect to the phylogenetically established secondary structure models (10,11). The majority of these latter sites are clustered in a region of the secondary structure associated with the peptidyl transferase activity through biochemical and genetic studies (12), although a number of other nucleotides that have been implicated in tRNA binding (*e.g.* 4,6) lie in distinct regions of the secondary structure. Clearly, a better understanding of the structural organisation of the 23S RNA is needed to gain further insight into the interaction of tRNA with the ribosome and to provide a basis for more detailed mechanistic studies.

The precise analysis of intra-RNA cross-links generated *'in situ'* is the most direct method for obtaining information concerning the topographical folding of rRNA within the ribosome. In previous studies, we have reported the analyses of a number of intra-RNA cross-linked complexes within 23S RNA generated through partial digestion with the double strand-specific cobra venom nuclease (CSV) (13,14). Recently, we have developed a more suitable approach for the isolation of individual intra-RNA cross-links by addressed cleavage using ribonuclease H. This enzyme is known to induce specific cuts in RNA molecules where the RNA is involved in a hybrid helix with DNA (15). Thus, by carefully designing deoxyribonucleotide templates and hybridising them with cross-linked RNA under controlled conditions, ribonuclease H partial digests can be tailored so as to excise individual cross-links in regions of the RNA molecule chosen at will. Because only a very few cuts are introduced into the RNA, the cross-linked complexes can be isolated with high purity in a semi-quantitative manner, using our established gel separation techniques. Moreover, this method can be efficiently used to probe selected areas of the molecule for additional cross-linking data.

In the series of experiments reported here, we have utilised this 'directed digest' approach, combined with additional developments in our methods of oligonucleotide analysis, to investigate the tertiary structure of 23S RNA regions implicated in the binding of tRNA; these include the 'peptidyl transferase

* To whom correspondence should be addressed

ring' (12) and the binding domains of ribosomal proteins L1 (16) and L2 (17). Using this approach, seven new cross-links could be precisely localised, including three long range 'tertiary' cross-links between regions widely separated in the secondary structure.

These new data, along with previously established intra-RNA and RNA-protein cross-links (reviewed in ref. 11), have been used to derive a three-dimensional model for the organisation of *E. coli* 23S RNA '*in situ*' in the 50S subunit (see ref. 18 for a preliminary description). Here we present a detailed arrangement for helices 64–93 of the 23S RNA secondary structure (11), which are involved in tRNA binding, and which are constrained by the cross-linking data into a distinct domain in the tertiary structure.

MATERIALS AND METHODS

The preparation of ribonuclease H and DNA oligonucleotide templates, and the isolation of uniformally 32P-labelled 23S RNA from '*in situ*' cross-linked 50S subunits were carried out as described in (19). Ribonuclease H digests of the cross-linked RNA were performed as in (19) with the following modifications: the RNA was preincubated with the appropriate decadeoxyribonucleotide templates at 55°C for 10 min., after which the mixture was made 1 mM in MgOAc, 0.1 mM in DTT, and a suitable quantity of ribonuclease H was added. Incubation was continued for a further 30 min. at 55°C. This modification has the dual advantage of a higher specificity of the digest and an increased accessibility of the target sites in the RNA. The digestion was stopped by the addition of 1/10 vol. of 1% SDS, 20 mM EDTA-KOH, pH 7.0. The digest was treated with 1 mg/ml proteinase K (55°C, 30 min.) and the RNA fragments recovered by phenol extraction and ethanol precipitation (19). The two dimensional gel system used to resolve the partial digests was a modification of the one given in (13): a 5% gel was used in the first dimension, from which a single gel strip was excised and directly polymerised (without prior washing) into a 12% second dimension gel. Cross-linked complexes were recovered from the gels and digested totally with ribonuclease T_1 or A, as in (13). The oligonucleotides released were separated by two dimensional chromatography on polyethyleneimine (PEI) cellulose plates, using the system of Volckaert and Fiers (20) or the 'alternative' chromatographic procedure (56% formic acid, 5 M urea; 7.6% formic acid/pyridine pH 4.3, 4 M urea (19)). Secondary digestions and analyses of oligonucleotides were performed as in (19).

In some instances, ribonuclease T_1 end-products were further characterised by digestion with ribonuclease U_2. These digestions were performed with 5 units of ribonuclease U_2 (Sankyo, Tokyo) in 10μl of 10 mM NH_4OAc, pH 4.5, 1 mM EDTA, with 5 μg carrier tRNA for 30 min. at 37°C, followed by slow warming (15 min.) to 60°C. The digests were lyophilised, taken up in 5 μl 0.1% SDS containing a little xylene cyanol, and resolved using the standard fingerprint system (20). After recovery from the PEI cellulose plates, oligonucleotides were digested with ribonuclease T_2 (Sankyo, Tokyo) as in (21) and the resulting 3'-nucleotide phosphates resolved on PEI cellulose plates using the system of Bernardi (22). In all cases, the oligonucleotide data were fitted to the 23S RNA sequence of Brosius *et al.* (23).

A 'wire and tube' model for the '*in situ*' structure of 23S RNA was constructed as described in (7). Coordinates of RNA helices measured from the model were fed into a VAX 8000 computer graphics system, again as described in (7).

RESULTS AND DISCUSSION

Directed digests of cross-linked rRNA with ribonuclease H

Fig. 1 shows a two dimensional gel separation of fragments generated by addressed cleavage using ribonuclease H. In this example, the digest was performed using 4 decadeoxyribonucleotide templates complementary to positions 1808–1817,

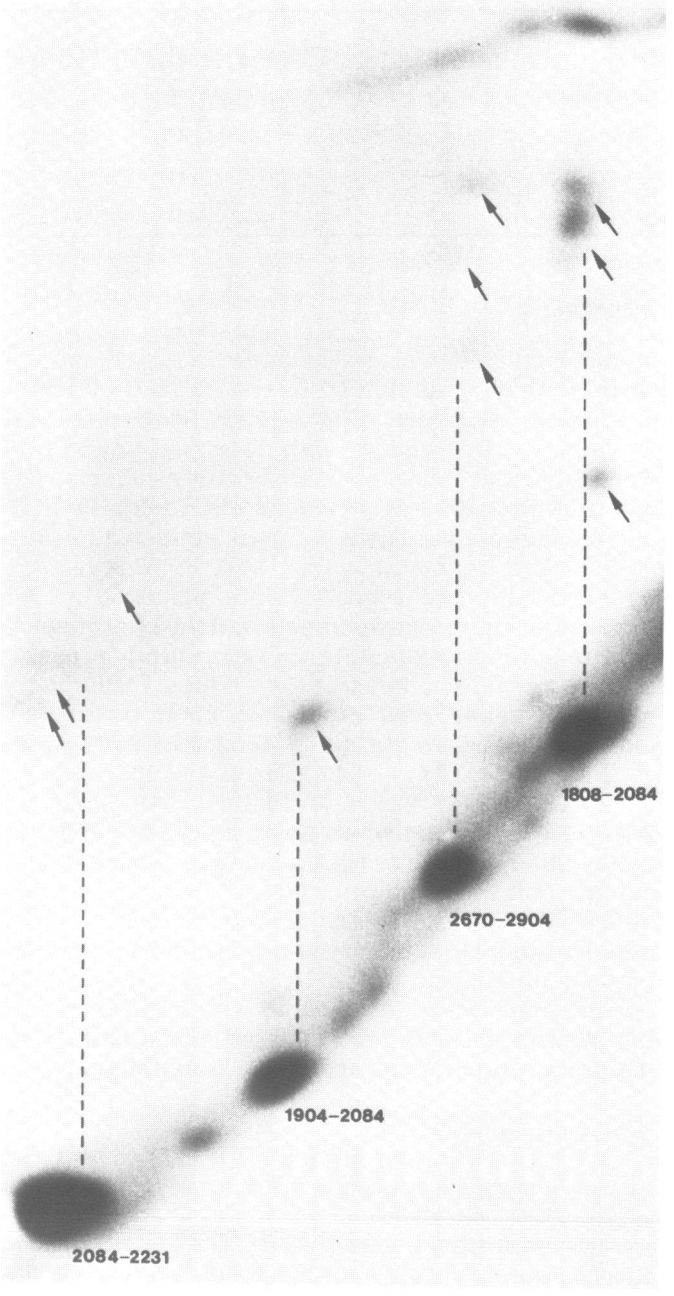


Figure 1. Two dimensional gel electrophoresis of cross-linked 23S RNA, following addressed cleavage with ribonuclease H. The digest was performed using decadeoxyribonucleotides complementary to positions 1808–1817, 1904–1913, 2084–2093 and 2231–2240 in the 23S RNA sequence (numbered from the 5'-end). Separation of the fragments was from right to left in the first dimension and from top to bottom in the second dimension. The specifically excised single stranded fragments, lying in a 'diagonal' on the gel, are labelled according to their positions in the 23S RNA sequence. Cross-linked complexes (shown with arrows) are retarded in the second dimensional separation, appearing as spots above the corresponding fragments in the diagonal (connected by the broken lines).

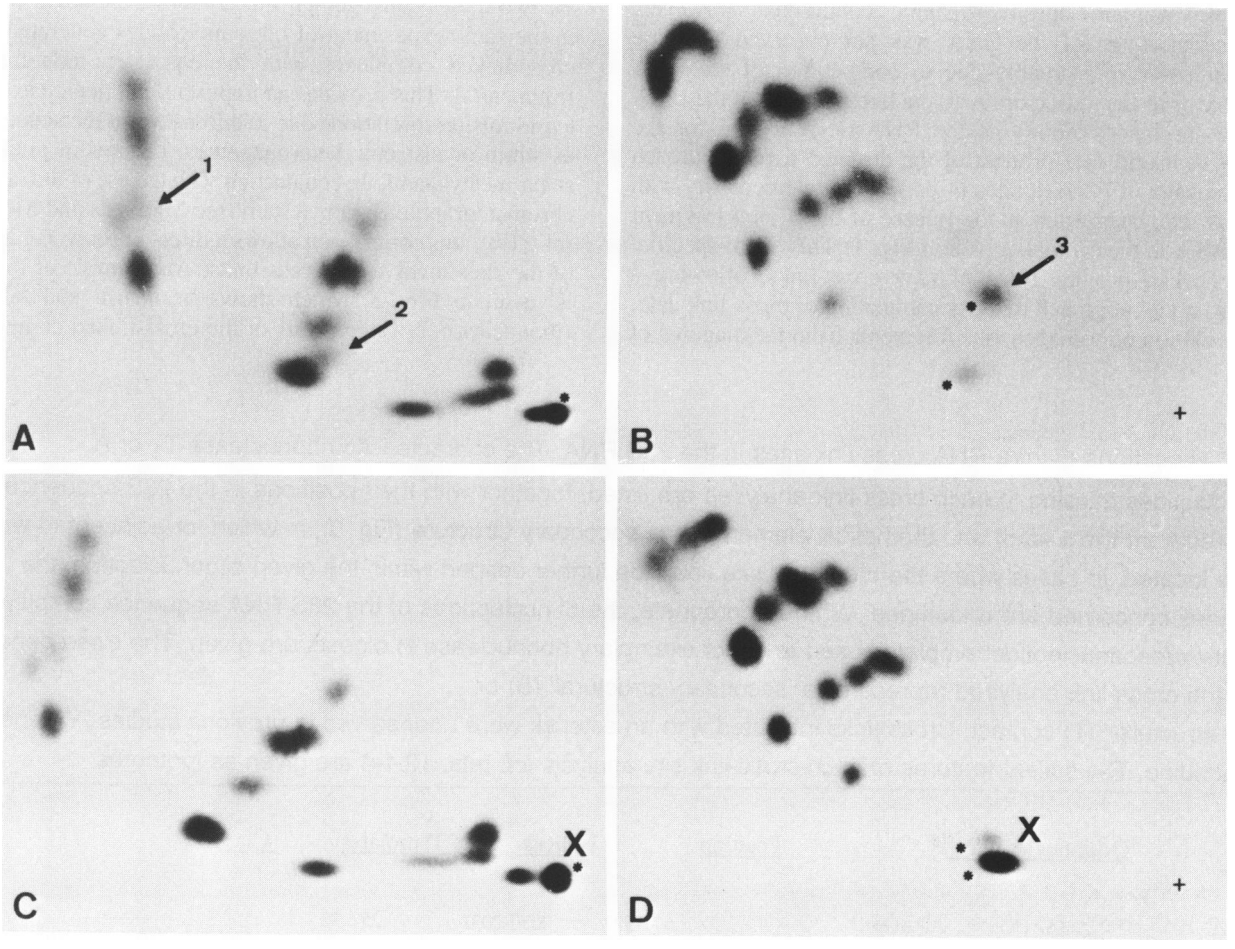


Figure 2. Ribonuclease T_1 fingerprints of a non-cross-linked RNA fragment (panels A and B) and the corresponding cross-linked complex (panels C and D) from positions *ca.* 1911−2054 in the 23S RNA sequence. Panels A and C show the 'standard' chromatographic system, panels B and D the 'alternative' system (see Materials and Methods). Direction of chromatography is from right to left in the first dimension and from bottom to top in the second dimension. Sample application points in panels B and D are denoted by '+'. The asterisks at the sample application points in panels A and C indicate oligonucleotides which resolve into pairs of spots (also asterisked) in panels B and D, respectively. Oligonucleotides 1, 2 and 3 from the non-cross-linked fragment (indicated by arrows in panels A and B) are 'missing' in the cross-linked samples, where instead the cross-linked oligonucleotide 'X' can be seen. (Oligonucleotide 1 co-elutes with an isomeric oligonucleotide, also present in the fragment; hence the spot is still present, albeit weaker, in panel C).

1904−1913, 2084−2093 and 2231−2240 in the 23S RNA sequence (*cf.* Fig. 3 below) The gel has the characteristic features (*cf.* refs. 13,14) of a 'diagonal' of non-cross-linked fragments, above which are spots (indicated with arrows) corresponding to the cross-linked complexes. In all cases, the gel patterns for a particular digest were highly reproducible. The fragment pattern is very simple, in contrast to the patterns obtained from CSV nuclease digests (Fig. 1, ref. 13), or ribonuclease H partial digests using sets of deoxyoligonucleotides complementary to sequences at approximately 50 nucleotide intervals along the complete length of the 23S RNA (Fig. 2, ref. 19). The diagonal is dominated by fragments generated by specific cleavage at the targeted sites, with the noteable exception in this case of a single strong spot corresponding to the last *ca.* 230 nucleotides at the 3′-end of the molecule (positions 2670−2904). Moreover, the pattern of cross-linked complexes is equally simplified. The cross-links are present for the most part as discrete spots running directly above the corresponding single-stranded fragment in the diagonal, and are free of contamination from complexes within other regions of the molecule.

In this particular example, we investigated the two distinct regions in the 23S RNA secondary structure previously associated with protein L1 through binding studies (16) and our own RNA-protein cross-linking experiments (reviewed in ref. 11 and *cf.* Fig. 3, below). Although cross-links were found within each of the excised L1-regions (and in the resulting fragment between the two regions), no intra-RNA cross-links connecting these two distinct portions of the molecule were observed. Such cross-links would be expected to run at positions in the first dimensional gel corresponding to the combined molecular weight of a pair of the excised fragments. However, using other similar 'directed digests', long range tertiary cross-links of this type were indeed found (cross-links 3, 4 and 6, see Table 1 below).

Most of the individual cross-links identified in this series of experiments were isolated within a set of overlapping fragments by using varying combinations of deoxyoligonucleotides. This flexibility of the directed digest approach permits the location of undefined cross-link sites to be gradually narrowed down until they can be unequivocally assigned. In this manner, the digests can be tailored to avoid complications in the oligonucleotide analyses, by avoiding regions containing non-characteristic oligonucleotides or isomeric oligonucleotides which co-elute

under the standard chromatographic conditions. However, addressed cleavage of the RNA was not observed in some particular cases, presumably due to competition of the RNA secondary or tertiary structure with the hybridisation of the DNA template, or inaccessibility of the RNA-DNA hybrid for the enzyme. A useful side product of the directed digest approach is the cleavage of RNA at sites of partial complementarity with the DNA templates, such as the release of the 3'-end fragment of 23S RNA in the example given in Fig. 1. This 'semi-specific' hydrolysis can provide hints of new cross-links, allowing a 'stepping stone' approach to the accumulation of cross-link data.

The isolation of homogeneous fragments from the diagonal of the two-dimensional gel (*cf.* Fig. 1) allows the direct comparison in the same experiment of oligonucleotides generated from the cross-linked complexes with the equivalent non-cross-linked fragment(s). This provides an important internal control that can avoid misinterpretations due to anomalies in the sequence, such as strain or cistronic heterogeneities, or sites of nucleoside or sugar methylations. In conjunction with the use of the 'alternative' chromatographic elution system (see Materials and Methods and ref. (19)), this comparison allows a direct unequivocal assignment of the sites involved in cross-links. An example of this method is given in Fig. 2, which shows 'standard' and 'alternative' ribonuclease T_1 fingerprints of the cross-linked complex 5 (see

Table 1: Locations of intra-RNA cross-link sites in the 23S RNA. The characteristic ribonuclease T_1 or A oligonucleotides missing in each cross-link analysed are listed, together with their positions in the 23S sequence (numbered from the 5'-end) and the helical element in the secondary structure (Fig. 3), in which, or adjacent to which, they are located. In cases where the cross-link site could be further defined within the given oligonucleotide, the nucleotides concerned are underlined. Where appropriate, the 5'-nucleotides of the 23S RNA sequence complementary to the deoxydecanucleotide templates used to direct exemplary ribonuclease H digests are given. The class denotes wether the cross-link analysed represents a 'secondary structural' (S) or 'tertiary structural' (T) contact. Cross-links indicated with an asterisk were unobserved in previous studies. Y denotes psuedouridine. The salient features of each cross-link site analysis (*cf.* refs. 13,14) are given as footnotes.

No.	Oligonucleotides	Position	Helices	Templates	Class	
1	UAUAAUGp : A^mAGp	571-577 :	2030-2032	25 : 72	-----	T
2	ACUAAUG^{m}YTGp : UCCCUAUCUGp	739-748 :	2609-2618	35 : 73	-----	T
3	UUUAUUAAAAACACAGp : GGUp	1777-1792 :	2607-2609	65 : 93	1703, 1767, 2568, 2619	T*
4	AAATUCCUUGp : U^mGUp	1936-1945 :	2552-2554	70 : 92	1904, 1957, 2506, 2568	T*
5	AAATUCCUUGp : ACCUG.CACGp	1936-1945 :	1960-1968	70 : 71	1904, 2050	T*
6	AAGUp : U^mGUp	1952-1955 :	2552-2554	71 : 92	1904, 1957, 2506, 2568	T*
7	CCAGp : GGUp	2145-2148 :	2201-2203	78 : 79	2084, 2231	T*
8	UCUCCUCCUAAAGp : CUCAACGp	2257-2269 :	2422-2428	80 : 88	-----	T
9	CUCAUCACAUCCUGp : AAGGGUp	2510-2523 :	2541-2546	91 : 91	2438, 2568	S*
10	UUCUCCCUGp : GAGGCp	2783-2791 :	2892-2896	98 : 1	1808, 1904, 2084, 2231	T*

Cross-links 1 and 2 were previously reported (14) but could be further localised by isolation of the cross-linked residue using the 'alternative' fingerprint system. **1**: Secondary analysis with ribonuclease A of the cross-linked T_1 oligonucleotide gave Gp, AUp and AAUp but no Up. **2**: Secondary analysis of the cross-linked T_1 oligonucleotide gave ACp, Up, AAUp, and G^mYp but the AUp was either weak or absent. Ribonuclease U_2 digestion of the cross-linked residue gave UCUGp. The 5'-site was given by the absence of the overlapping ribonuclease A oligonucleotide GAAAAAUp (positions 748–754). **3**: UUUAUUAAAAACACAGp was clearly absent from the 'alternative' ribonuclease T_1 fingerprint (*cf.* Fig. 2D). Secondary analysis gave Up, AUp, ACp, AGp with a weak AAAAACp. GGUp was missing in the ribonuclease A fingerprint analyses, whereas the overlapping T_1 oligonucleotide UUCGp (2604–2607) was present. The T_1 oligonucleotide UCCCUAUCUGp (2609–2618) was sometimes present and sometimes absent, indicating a microheterogeneity of the 3'-end of the cross-link site between positions 2608–2609. **4**: AAATUCCUUGp was clearly absent from the 'alternative' ribonuclease T_1 fingerprint. CU^mGp was present in the ribonuclease T_1 fingerprint but U^mGUp was absent from the ribonuclease A fingerprint, hence the 3'-site was localised to U2554. The 5'-site is tentatively allocated to U_{1940} due to the submolar AAATp spot in the secondary analysis. **5**: AAATUCCUUGp, ACCUGp and CACGp were absent from the ribonuclease T_1 fingerprint analyses (see Fig. 2). Secondary analysis with ribonuclease A of the cross-linked complex gave AAATp, Up, Gp, ACp and a weak Cp spot. **6**: AAGUp, U^mGUp, and the overlapping oligonucleotides UUCCGp and CU^mGp (positions 1955–1959 and 2551–2553 respectively) were clearly absent from the ribonuclease A and T_1 fingerprints respectively. **7**: CCAGp and GGUp were not present in the ribonuclease T_1 and A fingerprints respectively. Position 2202 is heterogenous, with G not U as in Fig. 3. The cross-linked ribonuclease T_1 residue gave AGp on secondary analysis. **8**: CUCAACGp and UCUCCUCCUAAAGp were clearly absent from the ribonuclease T_1 fingerprints. The cross-linked ribonuclease T_1 residue resolved into two spots by 'alternative' fingerprint system which gave identical secondary analyses with strong Cp and Up spots and AAAGp but no AACp or Gp. **9**: CUCAUCACAUCCGp and AAGGGUp were absent from the ribonuclease T_1 and A fingerprints respectively. UAUGp (positions 2546–2549) was present, as was UCCCAAGp (positions 2537–2543). Secondary analysis of the cross-linked ribonuclease T_1 residue gave submolar AUp, and digestion with ribonuclease U_2 did not generate UCCUGp. **10**: This cross-link was isolated fortuitously, due to hydrolysis of the RNA by ribonuclease H at a site of partial complementarity with the templates indicated (see text for discussion). UUCUCCCUGp and GAGGCp were clearly absent from the ribonuclease T_1 and A fingerprint analyses respectively. Secondary analysis of the cross-linked ribonuclease T_1 residue yielded strong Up and Cp spots and a cross-linked residue, but no Gp.

Table 1 below) and the corresponding non-cross-linked RNA fragment between positions *ca.* 1911–2054 of the 23S RNA sequence. Three oligonucleotides which are present in the control fragment (indicated by arrows) but which are absent from the cross-linked complex, constitute the cross-linked ribonuclease T_1 residue. This gives an additional characteristic spot (X) in the fingerprints of the cross-linked sample, comprising the adjacent oligonucleotides ACCUGp (spot 2) and CACGp (spot 1) covalently linked to AAATUCCUUGp (spot 3) (*cf.* Table 1).

We have previously reported the use of immobilised cDNA clones complementary to sequences within the 16S RNA to hybridise specific populations of RNA fragments after a partial digest with CSV nuclease (24). In this manner, the complexity of the gel patterns from CSV digests could be markedly reduced, and the technique proved a successful method to isolate cross-linked fragments located in selected regions of the RNA. However, this approach cannot overcome the inherently heterogeneous nature of the partial digestion with CSV and, moreover, it is not suitable for the analysis of low-yield cross-links, due to the losses incurred during the hybridisation step. In contrast, directed digests with ribonuclease H are semi-quantitative and the product fragments can be tailored to suit the particular requirements of the experiment.

Localisation of cross-links within the 23S RNA secondary structure

The sites of intramolecular cross-links within the 23S RNA which were localised by a large number of directed digests in these studies are summarised in Table 1, and are superimposed on the updated version of our model for the secondary structure of 23S RNA (11) in Fig. 3. Also included are three tertiary cross-links (numbered 1, 2 and 8 in Table 1) which were isolated by ribonuclease H digests using sets of deoxyoligonucleotides complementary to sites along the whole length of the 23S RNA sequence (*cf.* ref. 19); these were reported in our earlier studies, but have now been more precisely localised by oligonucleotide fingerprinting using the 'alternative' chromatographic elution system (*cf.* Fig. 2) and, in the case of cross-link 2, ribonuclease U_2 digestion of the cross-linked ribonuclease T_1 residue. Of the seven new cross-links, six represent tertiary contacts in the folding of the rRNA and thus can be used to impose spatial constraints on the phylogenetically established secondary structure of 23S RNA in model building studies.

The most important topographical data are provided by cross-links 3, 4 and 6, which interconnect specific sites in the L2 binding domain (17) with sites in the vicinity of the 'peptidyl transferase ring', at the junction of helices 73, 74, 89, 90 and 93 (see Fig. 3). The analysis of cross-link 3 (see Table 1) is similar to that of a tertiary cross-link tentatively reported earlier (cross-link H, ref. 14), but in the latter case the 3'-site of the cross-link was localised to positions 2584–2588. These nucleotides lie adjacent to positions 2607–2609 in the secondary structure at the base of helix 93 (see Fig. 3), and thus cross-link 3 is not incompatible with this earlier observation. However, a cross-link involving the previously reported site was not observed in these experiments. Cross-link 6 connects helices 71 and 92 of the secondary structure and was preliminarily localised through ribonuclease H digests using sets of DNA templates (19). Directed digests to excise this cross-link revealed two contacts between the same fragments (cross-links 4 and 6) which, taken together with cross-link 5, demonstrate a very close juxtaposition of the loop ends of helices 70, 71 and 92.

A three-dimensional model of the tRNA binding domain of 23S RNA

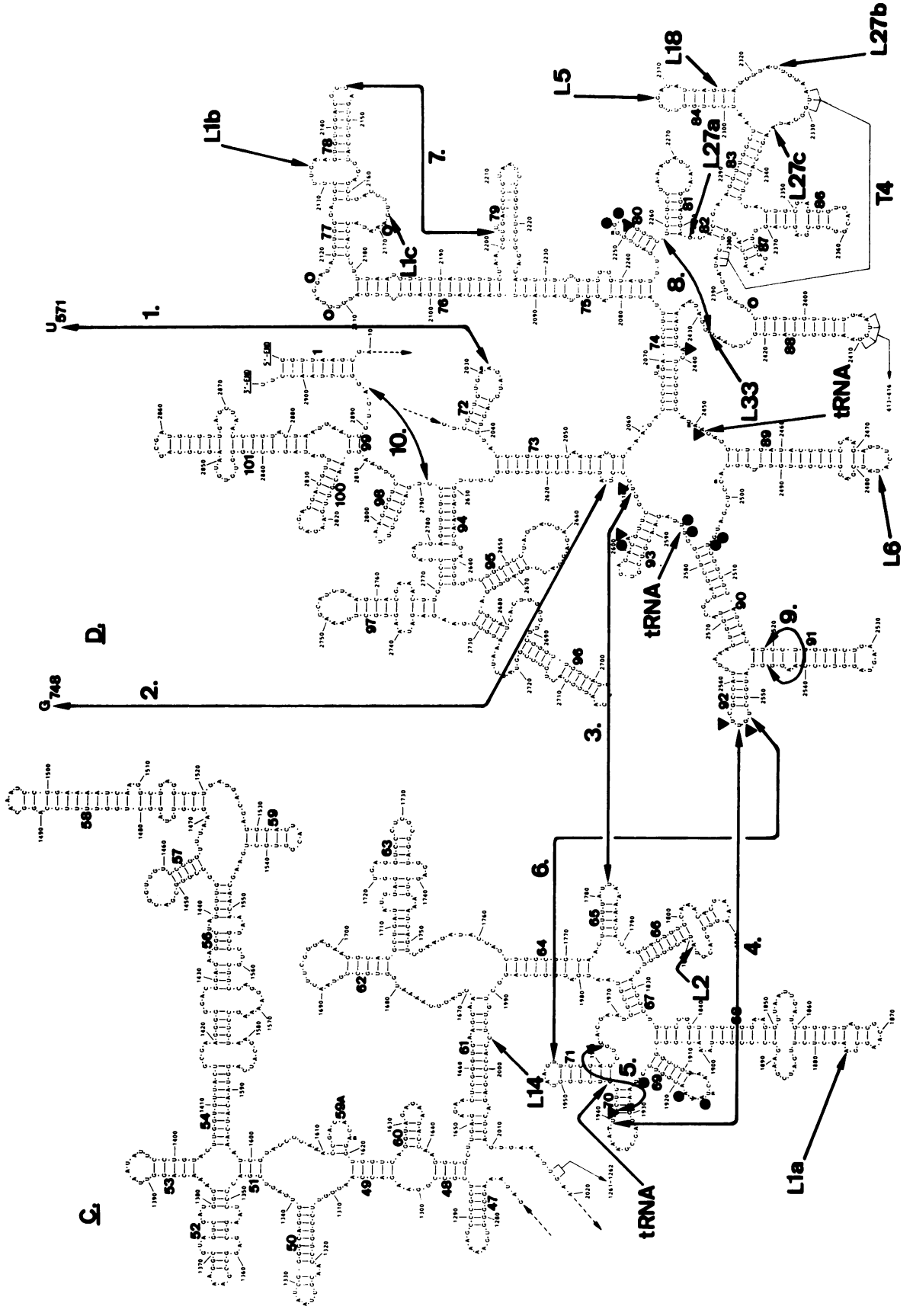
In addition to the locations of intra-RNA cross-links analysed in this study, Fig. 3 also shows topographical data obtained from RNA-protein cross-linking studies carried out in our laboratory (reviewed in ref. 11,18). These two independent sets of cross-linking data define points of topographical neighborhood 'in situ' and were used to impose constraints upon the folding of the secondary structure of 23S RNA in three-dimensional model building studies.

RNA-protein cross-linking data provide the link between the tertiary structure of the rRNA and the spatial distribution of the proteins within the ribosomal subunit. Both recent models for the tertiary structure of 16S RNA '*in situ*' in the 30S subunit (7,8) relied upon the known relative positions of the ribosomal proteins as defined by low angle neutron scattering studies (25) to correlate the tertiary structure of the RNA with the spatial arrangement of the proteins. At the moment, a corresponding map is not available for the 50S subunit. However, a model has been proposed for the arrangement of the proteins within the 50S subunit (26), based on immuno electron microscopy (IEM) data and the immunological analysis of a large number of inter-protein cross-links. The high degree of compatibility of the IEM model with the neutron scattering map in the case of the 30S subunit (see *e.g.* ref. 11) suggests that the IEM-derived protein arrangement of ref. 26 for the 50S subunit is likely to be reliable.

Accordingly, we have incorporated our complete intra-RNA and RNA-protein cross-link data into a model for the three-dimensional structure of 23S RNA in the large ribosomal subunit, using the protein map (26) to align the arrangement of the RNA with the spatial distribution of the proteins. The tRNA binding domain of the 23S RNA involves sites in the region comprising helices 64 to 93 (Fig. 3), and the RNA-protein cross-linking data in this portion of the molecule define sites in close topographical neighborhood with proteins L1, L2, L5, L6, L14, L18, L27 and L33 (see Fig. 3). With the exception of L33, epitopes to all of these proteins have been located on the ribosomal subunit by IEM studies. However, cross-link data pairing L33 with both L1 and L27 allows it to be placed between the latter two proteins, close to the base of the 'L1 protuberance' (26).

The computer version of the 23S RNA model is presented in Fig. 4. The view given is of the interface side of the 50S subunit. The overall shape of the large subunit observed in electron micrographs (including the 'stalk', 'central protuberance' and 'L1 protuberance') is clearly recognisable in the model (*cf.* ref. 26). The tertiary structure of the 23S RNA is presented as series of cylindrical elements corresponding to the individual helices of the secondary structure (*cf.* ref. 7). No information is available for the detailed arrangement of single stranded regions of the molecule, or for the effects of loops or bulges upon individual helices, and therefore they are not indicated in the computer model, although present in the 'wire and tube' model (*cf.* ref. 7).

This preliminary model provides a framework within which to incorporate steadily incoming new cross-link data. Although the general folding pattern of the RNA in the large subunit has become clear, sufficient data are not available at this stage to present a detailed model of the complete 23S RNA. In contrast, the high degree of spatial constraints imposed on the organisation of helices 64–93 of the secondary structure define a distinct self-contained domain in the three-dimensional model of the 23S RNA, positioned in the upper portion of the large subunit on the interface side. Helices constituting this detailed region of the


L1b
L1c
L5
L18
L27b
L27a
L27c
T4
L33
tRNA
L6
L2
L14
L1a
tRNA
tRNA
U571
G748
1.
2.
3.
4.
5.
6.
7.
8.
9.
10.
D.
C.

model are denoted by heavier shading in Fig. 4 and numbered with respect to Fig. 3. The other regions of the structure are also included in Fig. 4 to illustrate the orientation of the tRNA binding domain of 23S RNA within the complete molecule: this is established through a number of long range connections (see ref. 18 for discussion), *e.g.* intra-RNA cross-links 1 and 2 and a tertiary interaction between positions 413–416 and 2407–2410 in the 23S RNA sequence (27) (see Fig. 3). These remaining helices are however left unnumbered to emphasize the preliminary nature of these regions of the model.

Figure 4a shows the locations of RNA-protein cross-link sites within the tRNA binding domain in the three-dimensional model. We have previously reported cross-links to L1 between helices 77 and 78 (11) within the known binding region of the protein (16), and have since been able to precisely assign an additional L1 cross-link site in helix 68 to positions 1876–1878 (M. Osswald and R. Brimacombe, unpublished results), providing evidence that the orientation of this latter helix is directed towards the 'L1 protuberance' in the ribosomal subunit.

As noted above, the most important intra-RNA cross-link data for defining this region of the model are provided by the three long range intra-RNA cross-links 3, 4 and 6. Cross-link 3 connects the loop-end of helix 65 (extending from the right hand side of helix 65 in Fig. 4) to the single stranded region at the base of helix 93 (helix 93 is barely visible in Fig.4a, being concealed by helix 92). Cross-links 4 and 6 connect helices 70, 71 and 92, and are located at the front of the model (as viewed in Fig. 4), at the base of the 'central protuberance'. Additional data are provided for the arrangement of all the remaining independent helical systems within this region of the 23S RNA. Thus, cross-link 8 folds the helix system 76–79 tightly back upon itself to form a compact structure which, in addition to helix 68 as already noted, constitutes the RNA moiety of the 'L1 protuberance' (helix 79 is hidden from view in Fig. 4, lying behind helix 76). The sub-domain consisting of helices 80–88 is located in the 'central protuberance' of the 50S subunit. This region is constrained by the three cross-link sites to L27, the cross-links to L5 and L18, the phylogenetically established tertiary interaction (27) between positions 2328–2330 and 2385–2387 (labelled **T4** in Figs. 3,4), and cross-link 8, which spans the 'ring' connecting helices 74, 75, 80, 81, 82 and 88 (see Fig. 3). The orientations of helices 74 and 89 are governed by defined RNA-protein cross-link sites to L33 and L6 respectively (*cf.* ref. 26). Similarly, helices 61 and 66 contain cross-links to proteins L14 and L2 respectively.

Functional significance of the model

The site of peptide bond formation has been localised to the highly conserved 'peptidyl transferase ring' connecting helices 74, 75, 89, 90 and 93 in the secondary structure of 23S RNA (see Fig. 3), largely on the basis of the genetic analysis of organisms resistant to chloramphenicol and anisomycin (both of which are antibiotic inhibitors of peptidyl transfer) (see ref. 12). Furthermore, a photoreactive analogue of puromycin was specifically cross-linked to 23S RNA at positions G_{2502} and U_{2504} within the 'peptidyl transferase ring' (28). Chemical footprinting studies (29) have shown that chloramphenicol, erythromycin, carbomycin and vernamycin B (which compete for the same binding site on the ribosome (30) although they inhibit protein synthesis by different mechanisms) have overlapping binding sites in this region of the RNA secondary structure. Vernamycin B also gave a strong footprint at position A_{752}, thus providing functional data which correlates well with the intra-RNA cross-link between positions 748 and 2613–2614 (number 2, Table 1) observed by us. This cross-link serves to bring the loop end of helix 35 (11) into the vicinity of the 'peptidyl transferase ring' in our model.

More directly pertaining to the binding site of tRNA, nucleotides 2451/2452 and 2584/2585 within the 'peptidyl transferase ring' were the sites of cross-linking to a photoreactive aminoacyl affinty analogue of charged $tRNA^{Phe}$, bound in the 'P' and 'A' sites respectively (3) (see Fig. 3). In other studies (4), an azidoadenosine nucleotide incorporated at position A_{76} of $tRNA^{Phe}$ was cross-linked to position G_{1945} in the 23S RNA, located at the junction between helices 70 and 71 (Fig. 3). In the same experiment, the tRNA was also cross-linked to L27, one of the most frequently labelled proteins in studies using probes directed at the peptidyl transferase center. Thus, G_{1945}, the sites in 23S RNA cross-linked to L27 and the 'peptidyl transferase ring' would be expected to be adjacent in the three-dimensional structure. Indeed, these regions of the 23S RNA are in very close proximity in our model, lying at the base of the 'central protuberance' of the 50S subunit (see Fig. 4).

In quite different studies from the photo-affinity labelling experiments described above, proteins L2, L3, L4 and L16 (in addition to the 23S RNA) have been shown to be essential for peptidyl transferase activity by partial reconstitution experiments '*in vitro*' (*e.g.* 31,32). Together with L27, these proteins are clustered in the IEM protein map of Walleczek *et al.* (26) towards the geometric center of the 50S subunit. This supports the location of the peptidyl transferase region directly beneath the 'central protuberance' of the large subunit, in accordance with our model (Fig. 4), although the position of the peptidyl transferase center is often interpreted as being more towards the 'L1 protuberance' (*e.g.* ref. 33).

The peptidyl transferase center defines the location of the CCA end of tRNA in the 'A' or 'P' sites on the large subunit. Using chemical footprinting techniques, Moazed and Noller (6) have recently identified a comprehensive catalogue of sites in the 23S RNA which are responsive to tRNA binding in the ribosomal 'A', 'P' and 'E' sites. As noted above, several of these sites implicate additional regions of the 23S RNA secondary structure in the binding of ribosomal bound tRNA. The locations of these tRNA footprinting data within the secondary structure of 23S RNA are given in Fig. 3, and their distribution in the three-dimensional model is shown in Fig. 4b.

The 'P' site footprints are clearly brought together into a tight group in the three-dimensional structure of 23S RNA, lying at the base of the 'central protuberance' on the interface side of the large subunit. Those sites protected explicitly by tRNA bound

Figure 3. The secondary structure of positions 1271–2904 of the *E. coli* 23S RNA (regions C and D, ref. 11), showing the locations of intra-RNA and RNA-protein cross-link data used to define the model for the tRNA binding domain. Helices are numbered according to reference 11. Intra-RNA cross-links are denoted by double headed arrows and numbered as in Table 1. RNA-protein cross-link sites are denoted by arrows and labelled with the respective ribosomal protein. Where more than one site is given (proteins L1 and L27) these are distinguished by the suffixes a, b and c. The location of tRNA cross-links or affinity labelling sites (2–4) and chemical footprints (6) within the 23S RNA are also given: sites involved in cross-linking (4) or affinity labelling with analogues of aminoacyl tRNA (2,3) are indicated by arrows and labelled tRNA, whereas the chemical footprinting data (6) are defined according to 'A', 'P' and 'E' sites (▼, ● and ○, respectively). **T4** indicates the phylogenetically established tertiary interaction between positions 2328–2330 and 2385–2387 (27). See text for further description.

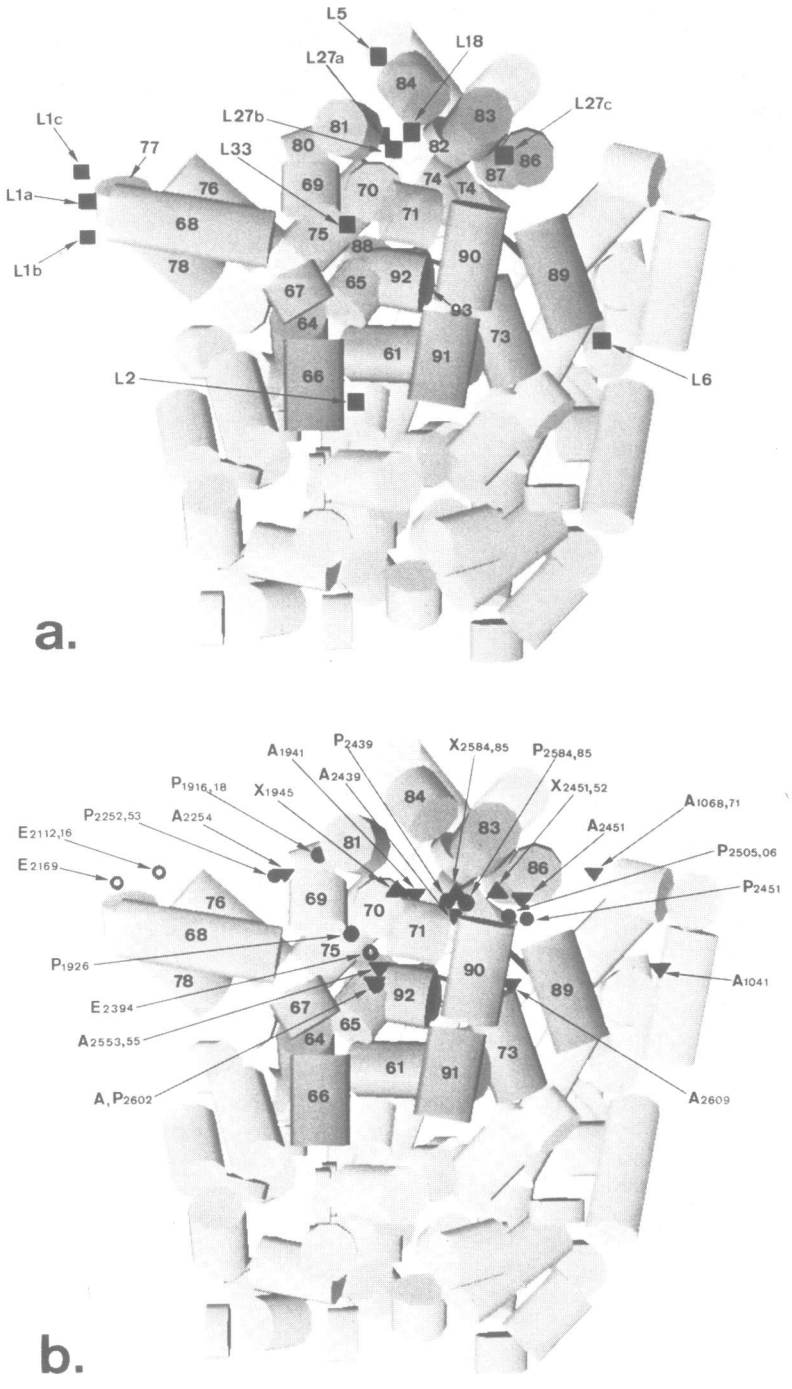


Figure 4. Computer graphics model of the tertiary structure of 23S RNA. The view given is of the interface side of the large subunit. Helices constituting the tRNA binding domain are denoted by heavier shading and numbered as in Fig. 3. **T4** denotes a phylogenetically established tertiary interaction, shown in Fig. 3. *Panel a* shows the locations of RNA-protein cross-link sites in the tertiary model: these are indicated by squares and labelled with the respective ribosomal protein, as in Fig. 3 (the site involved in a cross-link to L14 in helix 61 is not visible, being obscured in this view of the model by helix 91). *Panel b* shows the distribution of topographical data relating to tRNA in the model. The tRNA footprint sites (6) are indicated using the same symbols as in Fig. 3, and labelled according to the corresponding ribosomal binding site ('A', 'P' or 'E' sites, respectively). Their positions in the 23S RNA sequence are given, numbered from the 5'-end. Sites involved in cross-linking or affinity labelling with tRNA are denoted by ▲. These sites are labelled X and their positions in the 23S RNA sequence are likewise indicated.

in the 'A' site largely overlap the 'P' site footprints, with the exceptions of positions G_{1041}, G_{1068} and G_{1071}, which were shown to be dependent upon the presence of ribosomal bound EF-Tu. Nucleotides protected by tRNA bound in the 'E' site are less numerous and are located towards the 'L1 protuberance'. The distribution pattern of the footprint data on the model is in agreement with the proposed direction of movement of the tRNA molecule (34) from 'A' to 'P' to 'E' sites across the interface of the large subunit, from right to left with respect to the view shown in Fig. 4. The locations of 'E' site protections on the upper side of the L1 protuberance may indicate that the path of the tRNA leaving the ribosome proceeds over the top of this structure.

The cluster of sites beneath the 'central protuberance' of the 50S subunit would approximately accommodate a tRNA molecule (bound in either the 'A' or 'P' sites) positioned with its 3'-end at the 'peptidyl transferase ring' and the anticodon loop directed towards the 'L1 protuberance'. This orientation is in agreement with the observed cross-link between the anticodon loop of a photoreactive derivative of yeast $tRNA^{Met}$ and L1 for 'P' site bound tRNA (35). In this position, the tRNA anticodon loop could come into contact with the decoding site in the cleft of the 30S subunit and is suitably positioned to satisfy the cross-links observed to S19, L5 and L27 with the 'elbow' of tRNA (36). Further support for such an alignment is given by cross-links to L33 and a region of the 23S RNA between helices 81–86 using an affinity label positioned on the 'D' loop of lupin $tRNA^{Met}$ bound in the 'P' site (P. Gornicki, unpublished results). As noted above, L33 is located between L1 and L27, the latter protein (together with L5) being at the base of the 'central protuberance' of the large subunit in the 50S protein map (see Fig. 1, Ref. 26).

Moazed and Noller (6) reported the majority of footprint sites to be dependent on the structural integrity of the 3'-end of the tRNA. Clearly, the number of sites implicated in tRNA interaction and their distribution in the model suggest they could not be all in direct contact with the CCA end of the tRNA. Unfortunately, data from chemical footprinting studies do not allow one to distinguish between 'true' protections and alterations in reactivity due to allosteric effects. However, it is of note that of those sites reported which were not affected by removal of the 3'-terminus and which are thus possibly in direct contact with other portions of the tRNA molecule, only C_{2254} (which is afforded protection by 'A' site bound tRNA) is outside the regions that are implicitly associated with the 'peptidyl transferase ring' by cross-links 4 and 6 and the tRNA-23S RNA cross-link at position 1945 (see Fig. 3). C_{2254} is at the loop end of helix 80 and positioned near to the 'L1 ridge' in the three-dimensional model (Fig. 4) through the locations of cross-link 8 and the cross-link site to L33 (see Fig. 3), in agreement with the suggested orientation of the tRNA on the 50S subunit.

The striking phylogenetic conservation of certain helical elements and specific nucleotides within the rRNA secondary structure (27) probably reflects the molecular basis of translation at the level of a series of RNA-RNA interactions. The location of conserved helices in the 23S RNA on the interface surface of the 50S subunit in the model (18) is thus in keeping with the direct involvement of rRNA in the processes of protein synthesis. Indeed, many of the conserved helices are interposed in the secondary and tertiary structure, forming a group constituting the core of the tRNA binding domain. This group is composed of helices 61, 64, 66, 67, 69–75, 80, 81, 89, 90, 92 and 93 (see Figs. 3 and 4).

The model of the tRNA binding domain of 23S RNA presented here provides a first insight into the three-dimensional structure of a functionally significant region of the RNA of the large subunit. In conjunction with models for the organisation of 16S RNA in the small subunit (7,8), this affords a basis for beginning to understand molecular interactions between rRNA and ribosomal ligands during the translational process.

REFERENCES

1. Prince, J.B., Taylor, B.H., Thurlow, D.L., Ofengand, J. and Zimmermann, R.A. (1982) Proc. Natl. Acad. Sci. USA, **79**, 5450–5454.
2. Barta, A., Steiner, G., Brosius, J., Noller, H.F. and Kuechler, E. (1984) Proc. Natl. Acad. Sci. USA, **81**, 3607–3611.
3. Steiner, G., Kuechler, E. and Barta, A. (1988) EMBO J., **7**, 3949–3955.
4. Wower, J., Hixson, S.S. and Zimmermann, R.A. (1989) Proc. Natl. Acad. Sci. USA, **86**, 5232–5236.
5. Moazed, D. and Noller, H.F. (1990) J. Mol. Biol., **211**, 135–145.
6. Moazed, D. and Noller, H.F. (1989) Cell, **57**, 585–597.
7. Brimacombe, R., Atmadja, J., Stiege, W. and Schueler, D. (1988) J. Mol. Biol., **199**, 115–136.
8. Stern, S., Weiser, B. and Noller, H.F. (1988) J. Mol. Biol., **204**, 447–481.
9. Stiege, W., Stade, K., Schueler, D. and Brimacombe, R. (1988) Nucleic Acids Res., **16**, 2369–2388.
10. Noller, H.F. (1984) Ann. Rev. Biochem., **53**, 119–162.
11. Brimacombe, R., Greuer, B., Mitchell, P., Osswald, M., Rinke-Appel, J., Schueler, D. and Stade, K. (1990) In Hill, W. *et al.* (ed.), The Structure, Function and Evolution of Ribosomes. ASM Press, Washington. in the press.
12. Vester, B. and Garrett, R.A. (1988) EMBO J., **7**, 3577–3587.
13. Stiege, W., Zwieb, C. and Brimacombe, R. (1982) Nucleic Acids Res., **10**, 7211–7229.
14. Stiege, W., Glotz, C. and Brimacombe, R. (1983) Nucleic Acids Res., **11**, 1687–1706.
15. Donis-Keller, H. (1979) Nucleic Acids Res., **7**, 179–192.
16. Branlant, C., Krol, A., Sri Widada, J., Ebel, J.P., Sloof, P. and Garrett, R.A. (1976) Eur. J. Biochem., **70**, 457–469.
17. Beauclerk. A.A.D. and Cundliffe, E. (1988) EMBO J., **7**, 3589–3594.
18. Brimacombe, R., Gornicki, P., Greuer, B., Mitchell, P., Osswald, M., Rinke-Appel, J., Schueler, D. and Stade, K. (1990) Biochem. Biophys. Acta, in the press.
19. Brimacombe, R., Greuer, B., Gulle, H., Kosack, M., Mitchell, P., Osswald, M., Stade, K. and Stiege, W. (1990) In Spedding, G. (ed.), Ribosomes and Protein Synthesis; a Practical Approach. IRL Press, Oxford, in the press.
20. Volckaert, G. and Fiers, W. (1977) Anal. Biochem., **83**, 228–239.
21. Zwieb, C. and Brimacombe, R. (1979) Nucleic Acids Res., **6**, 1775–1790.
22. Bernardi, A. (1974) Anal. Biochem., **59**, 501–507.
23. Brosius, J., Dull, T.J. and Noller, H.F. (1980) Proc. Natl. Acad. Sci. USA, **77**, 201–204.
24. Stiege, W., Kosack, M., Stade, K. and Brimacombe, R. (1988) Nucleic Acids Res., **16**, 4315–4329.
25. Capel, M.S., Kjeldgaard, M., Engelman, D.M. and Moore, P.B. (1988) J. Mol. Biol., **200**, 65–87.
26. Walleczek, J., Schueler, D. Stoeffler-Meilicke, M., Brimacombe, R. and Stoeffler, G. (1988) EMBO J., **7**, 3571–3576.
27. Gutell, R.R. and Fox, G.E. (1988) Nucleic Acids Res., **16**, r175–r269.
28. Hall, C.C., Johnson, D. and Cooperman, B.S. (1988) Biochem., **27**, 3983–3990.
29. Moazed, D. and Noller, H.F. (1987) Biochimie, **69**, 879–884.
30. Fernandez-Munoz, R., Monro, R.E., Torres-Pinedo, R. and Vazquez, D. (1971) Eur. J. Biochem., **23**, 185–193.
31. Hampl, H, Schulze, H. and Nierhaus, K.H. (1981) J. Biol. Chem., **256**, 2284–2288.
32. Auron, P.E. and Fahnestock, S.R. (1981) J. Biol. Chem., **256**, 10105–10110.
33. Stoeffler, G. and Stoeffler-Meilicke, M. (1984) Ann. Rev. Bioeng., **13**, 303–330.
34. Noller, H.F., Moazed, D., Stern, S., Powers, T., Allen, P.A., Robertson, J.M., Weiser, B. and Triman, K. (1990) In Hill, W. *et al.* (eds.), The Structure, Function and Evolution of Ribosomes. ASM Press, Washington, in the press.
35. Podkowinski, J. and Gornicki, P. (1989) Nucleic Acids Res., **17**, 8767–8782.
36. Ofengand, J., Ciesiolska, J., Denman, R. and Nurse, K. (1986) In Hardesty, B. and Kramer, G. (ed.), Structure, Function, and Genetics of Ribosomes. Springer Verlag, New York and Heidelberg, pp. 473–494.