REVIEW ARTICLE

Structure and function of ribosomal RNA

Richard BRIMACOMBE and Wolfgang STIEGE
Max-Planck-Institut für Molekulare Genetik, Abteilung Wittmann, Berlin-Dahlem,
Germany

Introduction

The last few years have seen a considerable advance in our understanding of ribosomal RNA (rRNA). The large rRNA molecules have well-defined secondary structures that have been strongly conserved across the evolutionary spectrum, and there is an increasing body of evidence that the rRNA plays key roles in both assembly and function of the ribosomal particles. In order to correlate these structural and functional properties it is clear that one of the central objectives is a detailed determination of the three-dimensional organization of the rRNA *in situ* in the ribosome, and the purpose of this article is to review some of the most recent developments in this area. We begin with a summary of the available primary structural data for rRNA molecules, and follow with a brief description of how these data have been used to derive secondary structure models. The main part of the article is concerned with the three-dimensional aspects, namely the folding of the rRNA secondary structures into a third dimension, the interaction with ribosomal proteins, and the location of functionally significant sites. The reader is referred to other reviews for a more detailed treatment of certain subjects, and in general our citation of the literature is restricted to the most recent data on a particular topic.

Primary structure

In every ribosome the bulk of the ribosomal RNA consists of two large molecules, one in each ribosomal subunit, and a list of the species for which complete nucleotide sequences are available is given in Table 1. A few of these sequences were obtained by direct sequencing of the rRNA, but in most cases the corresponding rDNA sequence was determined. In this context it should be noted that a number of genes for rRNA contain introns (reviewed by Clark *et al.*, 1984; Noller *et al.*, 1981). It can be seen from Table 1 that the size of the rRNA molecules varies considerably, from 9S (approx. 640 nucleotides) in the small subunit up to 18S (approx. 1870 nucleotides), and from 12S

(approx. 1230 nucleotides) up to 28S (4800 nucleotides) in the large subunit. The mitochondrial rRNA molecules are the most variable in length, whereas those of chloroplasts and bacteria are relatively constant (16S and 23S in the small and large subunits, respectively), as are also the molecules from eukaryotic cytoplasmic ribosomes (17–18S and 26–28S, respectively).

With the exception of some of the smaller mitochondrial ribosomes, the large ribosomal subunit always contains a 5S rRNA molecule (120 nucleotides long), and 5S sequences from 175 species have so far been determined (reviewed by Erdmann *et al.*, 1984). In addition other small RNA species, namely 5.8S, 4.5S and 2S rRNA, occur in the large subunit, but sequence comparison studies (reviewed by Brimacombe *et al.*, 1983; Noller, 1984) have shown that these small molecules have clear counterparts in bacterial 23S rRNA. 5.8S RNA (see Erdmann *et al.*, 1984, for list of sequences) is present in the large subunit of eukaryotic cytoplasmic ribosomes and corresponds to the 5'-terminal region of 23S RNA (Nazar, 1980), and 2S RNA, which is present in the *Drosophila melanogaster* large ribosomal subunit, is a sub-fragment of 5.8S RNA (Pavlakis *et al.*, 1979). 4.5S RNA, on the other hand, occurs in chloroplast ribosomes, and corresponds to the 3'-terminal region of 23S RNA (MacKay, 1981). Thus, these small RNA species are the result of extra post-transcriptional processing events in the rRNA molecules, and there are other examples, such as the small subunit rRNA from *Paramecium* (see Table 1), or the large subunit rRNA from *D. melanogaster* (not yet sequenced), which occurs in two halves (Glover, 1981). The most extreme example is the ribosome of *Crithidia fasciculata*, whose large subunit contains no less than four extra small RNA species in addition to 5S and 5.8S RNA (Schnare *et al.*, 1983), although it has yet to be established whether all of these have 23S RNA counterparts. In contrast to this type of extra processing event, it has recently been proposed that some fungal mitochondrial ribosomes (which, as noted above, lack 5S RNA) may have actually incorporated a 5S-like sequence into their large subunit RNA molecules (Thurlow *et al.*, 1984).

Table 1. *Known sequences of rRNA or rDNA molecules*

The *s* values given correspond to the principal component of the mature rRNA molecule concerned. Some ribosomal subunits contain additional small RNA molecules as follows. ^aA short fragment arising from the 3'-end of the small subunit RNA gene; ^ba 5.8S or 5.8S-like RNA fragment from the 5'-end of the large subunit gene; ^ca 4.5S RNA fragment from the 3'-end of the large subunit gene (see the text). – Indicates that the sequence concerned is not available.

	Small subunit		Large subunit	
	<i>s</i> value	Reference	<i>s</i> value	Reference
Mitochondria				
<i>Trypanosoma brucei</i>	9S	Eperon <i>et al.</i> (1983)	12S	Eperon <i>et al.</i> (1983)
Mouse	12S	Van Etten <i>et al.</i> (1980)	16S	Van Etten <i>et al.</i> (1980)
Rat	12S	Kobayashi <i>et al.</i> (1981)	16S	Saccone <i>et al.</i> (1981)
Human	12S	Eperon <i>et al.</i> (1980)	16S	Eperon <i>et al.</i> (1980)
Bovine	12S	Anderson <i>et al.</i> (1982)	16S	Anderson <i>et al.</i> (1982)
<i>Paramecium primaurelia</i>	13S ^a	Seilhammer <i>et al.</i> (1984a)	20S ^b	Seilhammer & Cummings (1981)
<i>Paramecium tetraurelia</i>	13S ^a	Seilhammer <i>et al.</i> (1984a)	20S ^b	Seilhammer <i>et al.</i> (1984b)
<i>Saccharomyces cerevisiae</i>	15S	Sor & Fukuhara (1980)	21S	Sor & Fukuhara (1983)
<i>Aspergillus nidulans</i>	16S	Köchel & Küntzel (1981)	23S	Köchel & Küntzel (1982)
Maize	18S	Chao <i>et al.</i> (1984)	26S	Dale <i>et al.</i> (1984)
Wheat	18S	Spencer <i>et al.</i> (1984)	–	–
Chloroplasts				
Maize	16S	Schwarz & Kössel (1980)	23S ^c	Edwards & Kössel (1981)
Tobacco	16S	Tohdoh & Sugiura (1982)	23S ^c	Takaiwa & Sugiura (1982)
<i>Euglena gracilis</i>	16S	Graf <i>et al.</i> (1982)	–	–
<i>Chlamydomonas reinhardtii</i>	16S	Dron <i>et al.</i> (1982)	–	–
Eubacteria				
<i>Escherichia coli</i>	16S	Brosius <i>et al.</i> (1978), Carbon <i>et al.</i> (1978)	23S	Brosius <i>et al.</i> (1980), Branlant <i>et al.</i> (1981)
<i>Anacystis nidulans</i>	16S	Tomioka & Sugiura (1983)	23S	Kumano <i>et al.</i> (1983), Douglas & Doolittle (1984)
<i>Proteus vulgaris</i>	16S	Carbon <i>et al.</i> (1981)	–	–
<i>Bacillus brevis</i>	16S	Kop <i>et al.</i> (1984a)	–	–
<i>Bacillus stearothermophilus</i>	–	–	23S	Kop <i>et al.</i> (1984b)
<i>Mycoplasma capricolum</i>	16S	Iwami <i>et al.</i> (1984)	–	–
Archaeobacteria				
<i>Halobacterium volcanii</i>	16S	Gupta <i>et al.</i> (1983)	–	–
<i>Halococcus morrhua</i>	16S	Leffers & Garrett (1984)	–	–
Eukaryotic cytoplasm				
<i>Dictyostelium discoideum</i>	17S	McCarroll <i>et al.</i> (1983)	–	–
Rice	17S	Takaiwa <i>et al.</i> (1984)	–	–
Maize	17S	Messing <i>et al.</i> (1984)	–	–
<i>Saccharomyces cerevisiae</i>	18S	Rubtsov <i>et al.</i> (1980)	26S ^b	Georgiev <i>et al.</i> (1981)
<i>Saccharomyces carlsbergensis</i>	–	–	26S ^b	Veldman <i>et al.</i> (1981)
<i>Physarum polycephalum</i>	–	–	26S ^b	Otsuka <i>et al.</i> (1983)
<i>Artemia salina</i>	18S	Nelles <i>et al.</i> (1984)	–	–
<i>Xenopus laevis</i>	18S	Salim & Maden (1981)	28S ^b	Ware <i>et al.</i> (1983)
Mouse	18S	Raynal <i>et al.</i> (1984)	28S ^b	Hassouna <i>et al.</i> (1984)
Rat	18S	Torczynski <i>et al.</i> (1983), Chan <i>et al.</i> (1984)	28S ^b	Hadjiolov <i>et al.</i> (1984), Chan <i>et al.</i> (1983)
Rabbit	18S	Connaughton <i>et al.</i> (1984)	–	–

There are a small number of modified nucleotides in ribosomal RNA. In *E. coli*, the 16S rRNA contains nine methylated bases (Carbon *et al.*, 1979) and the 23S rRNA ten methylated bases and three pseudouridine residues (Branlant *et al.*, 1981).

Modified nucleotides are more common in eukaryotic rRNA molecules, the additional modifications consisting mostly of 2'-*O*-methylated ribose moieties and pseudouridines (reviewed by Brimacombe *et al.*, 1983).

Secondary structure

The derivation of secondary structure models for the rRNA molecules has attracted a great deal of attention during the last 4 or 5 years, and these structures form the basis for many of the three-dimensional studies which will be discussed below. However, since the secondary structure is a topic that has been reviewed several times in detail lately, we will only give a brief resumé here.

The main effort has concentrated on the *E. coli* 16S and 23S rRNA, and three essentially similar secondary structure models were proposed for both molecules (16S RNA: Noller & Woese, 1981; Stiegler *et al.*, 1981a; Zwieb *et al.*, 1981; and 23S RNA: Glotz *et al.*, 1981; Noller *et al.*, 1981; Branlant *et al.*, 1981). In each case the structures were derived by a two-track approach, involving firstly the collection of experimental data from the *E. coli* RNA and secondly phylogenetic comparisons with ribosomal RNA sequences from other organisms (reviewed by Brimacombe *et al.*, 1983; Woese *et al.*, 1983; Noller, 1984). The principle of the phylogenetic approach (Fox & Woese, 1975) is simply that, if two similar but not identical sequences are compared, then base changes in one strand of a putative double-helical region must be compensated by corresponding base changes in the opposing strand, in order to preserve the base-pairing pattern (see Brimacombe, 1984, for a short review). With the number of rRNA sequences now available (see Table 1) this approach has proved a very powerful method for refining and extending the secondary structures, and as a result the latest models (Maly & Brimacombe, 1983; Noller, 1984) have converged to the point where they are identical in all but a few areas.

Our secondary structure models for the 16S and 23S rRNA from *E. coli* (slightly modified from Maly & Brimacombe, 1983) are shown in Figs. 1 and 2, respectively. The principal difference to the model of Noller (1984) in the case of 16S RNA is in the region of the right-hand six base pairs of helix 7 together with helix 10 (Fig. 1), and there is also a difference of opinion as to whether helix 19 exists or not. Otherwise, apart from minor variations in the detailed base-pairing of some half a dozen of the loops, the models are identical. In 23S RNA (Fig. 2) there are differences in the regions of helix 3/helix 18, helix 23, helix 44, and helices 47 and 49. Helices 5, 28 and 37 are not in Noller's model (Noller, 1984), but otherwise, as with 16S RNA, there are only minor variations in the base-pairing in a handful of the secondary structural loops. Our model (Fig. 2) shows two possible versions of helix 20, but we are now of the opinion that the interaction between bases 579–585 and 1255–1261 (Noller, 1984) is the correct one. Some of the

principal long-range interactions in the RNA models (base-pairings between regions that are remote from one another in the primary sequence; cf. Figs. 1 and 2) have recently been visualized by electron microscopy of partially denatured 16S or 23S RNA (Klein *et al.*, 1983, 1984).

The phylogenetic approach has shown that a number of non-Watson-Crick base-pairings appear to be able to participate in helix formation. Apart from the well-known G–U pairing, G–A pairs are often seen, and a recent n.m.r. study (Kan *et al.*, 1983) has shown that G–A pairs can indeed exist in double helices. Pyrimidine–pyrimidine 'mis-matches' as well as A–C pairs are also sometimes found in the rRNA molecules. However, the most important finding to emerge from the phylogenetic studies is the now firmly-established concept that the secondary structures of the rRNA have been conserved to a remarkable extent throughout evolution, regardless of the length of the particular rRNA molecule concerned. As was already mentioned in the preceding section (see Table 1), the lengths of the rRNA sequences from different species vary by a factor of three in both subunits. If one regards the bacterial 16S and 23S RNA molecules as the 'norm', then these differences in length have to be accommodated by large deletions when going 'down' to the small mitochondrial rRNA, or by corresponding insertions when going 'up' to the large eukaryotic rRNA. The deletions and insertions are found in specific regions of the secondary structures and are clearly distinguishable from the rest of the structures, which constitute a conserved 'core' present in every species of ribosomes (reviewed by Brimacombe *et al.*, 1983; Brimacombe, 1984). The secondary structural core is distributed along the whole length of the rRNA molecules, and there are numerous stretches of primary sequence of up to about 15 bases in length which have been almost universally conserved. One of the most well-studied regions is the 3'-terminal loop of the small subunit rRNA (Van Knippenberg *et al.*, 1984; cf. Fig. 1), although it must be emphasized that this loop is by no means exceptional, many other regions of both rRNA molecules showing similar patterns of conservation in both primary and secondary structure.

In addition to the structures for 16S and 23S RNA referred to above, detailed secondary structure models based on sequence comparisons have been proposed for all of the major classes of rRNA (cf. Table 1), namely—in the case of the small subunit—for mammalian mitochondrial 12S RNA (e.g. Mankin & Kopylov, 1981; Stiegler *et al.*, 1981b; Zwieb *et al.*, 1981) and for eukaryotic 18S RNA (e.g. Mankin *et al.*, 1981; Olsen *et al.*, 1983; Atmadja *et al.*, 1984; Chan *et al.*, 1984). The

sequence data and secondary structures have also been used for the examination of evolutionary relationships between species (Küntzel & Köchel, 1981; Gray *et al.*, 1984). In the case of the large subunit, models have been published for mammalian mitochondrial 16S RNA (e.g. Glotz *et al.*, 1981; Branlant *et al.*, 1981) as well as for eukaryotic 26S RNA (Veldman *et al.*, 1981) and 28S RNA (e.g. Clark *et al.*, 1984; Michot *et al.*, 1984; Hadjiolov *et al.*, 1984). The equivalence of eukaryotic 5.8S RNA to the 5'-terminal region of prokaryotic 23S RNA has already been mentioned in the previous section, and the interaction between 5.8S RNA and the 5'-region of the 26S or 28S RNA is now well-documented (Vaughn *et al.*, 1984; Walker *et al.*, 1983; see Brimacombe *et al.*, 1983, for review). Evidence is also accumulating for an interaction of 5.8S RNA with the 3'-end of the 26S or 28S RNA (Kelly & Cox, 1981; Georgiev *et al.*, 1984), but the precise nature of this putative interaction is not yet clear; although there is strong evidence that the extreme 5'- and 3'-ends of prokaryotic 23S RNA interact with one another (helix 1, Fig. 2), this helix does not have a direct counterpart in 5.8S and 28S RNA, since the 5'-terminal residue of 5.8S RNA corresponds to the 12th base from the 5'-end of 23S RNA, which lies outside helix 1.

In the case of 5S RNA, the large number of sequences available (Erdmann *et al.*, 1984) has led to the publication of many phylogenetically derived secondary structure models (e.g. Stahl *et al.*, 1981; Studnicka *et al.*, 1981; Pieler & Erdmann, 1982; Küntzel *et al.*, 1983; Trifonov & Bolshoi, 1983). These models are all very similar, and it is doubtful whether the determination of yet more sequences will contribute anything further of significance to the secondary structure. As has been pointed out (Brimacombe, 1984) this law of diminishing returns is already beginning to operate for the 16S and 23S RNA molecules, and here the method recently described by Qu *et al.* (1983) in which a DNA primer in a conserved sequence region from one species is used to sequence the

adjacent (non-conserved) regions in several heterologous RNA molecules by reverse transcription should prove useful; this method allows the rapid collection of partial sequence data in regions of interest from related species.

A more serious limitation of the sequence comparison approach is that the universally conserved stretches of primary sequence in the 16S or 23S RNA (which are presumably of particular functional importance) cannot be placed in base-paired structures by this method, since there are *a priori* never any compensating base changes in such conserved stretches. For the same reason it is unlikely that sequence comparisons *per se* will be of much help in determining the tertiary folding of the RNA, since tertiary interactions would not in general be expected to span more than three or four base pairs at the most, and the chances of finding convincing sets of compensating base changes are therefore low. So far only one such phylogenetically consistent tertiary interaction has been observed, namely helix 2 in the 16S RNA (Fig. 1), which involves base-pairing of a region of the RNA (bases 915–918) to the loop end of another helix (helix 1). Model-building studies (P. Maly & R. Brimacombe, unpublished results) show that it is stereochemically possible for helix 2 to co-exist with helix 1.

Tertiary structure and inter-RNA contacts

So far tRNA is the only RNA molecule for which a precise three-dimensional structure has been determined, by means of X-ray crystallography (e.g. Ladner *et al.*, 1975). In the case of the ribosomal RNAs, 5S RNA from *Thermus thermophilus* is the only intact molecule to have been crystallized (Morikawa *et al.*, 1982), but the crystals were not of sufficiently high quality to allow a structural analysis to be undertaken. More recently a fragment of *E. coli* 5S RNA has been crystallized, as well as a complex of this fragment with ribosomal protein L25 (Abdel-Meguid *et al.*, 1983), and these crystals appear to be more

Fig. 1. Secondary structure of *E. coli* 16S rRNA, including sites of intra-RNA and RNA-protein cross-linking, and protein binding sites

The structure is divided into two domains and the helices numbered as in Maly & Brimacombe (1983). The arrows connected by lines denote intra-RNA cross-links. Broken lines indicate an uncertainty in the cross-link site analysis, or that the cross-links were formed in 'naked' RNA. The intra-RNA cross-links are from the following sources: a, Zwieb & Brimacombe (1980); b, J. Atmadja & R. Brimacombe, unpublished results; c, Expert-Bezançon *et al.* (1983); d, Thompson & Hearst (1983a); e, Turner *et al.* (1982); f, Prince *et al.* (1982), tRNA. RNA-protein cross-links are indicated by an arrow to the corresponding protein (boxed), and are from: g, Zwieb & Brimacombe (1979); h, Wower & Brimacombe (1983); i, Chiaruttini *et al.* (1982); j, Czernilofsky *et al.* (1975); k, A. Kyriatsoulis, I. Wower & R. Brimacombe, unpublished results. The precision of determination of the individual cross-link sites (both intra-RNA and RNA-protein) varies from one case to another (cf. the original literature). Protein binding sites, enclosed by dotted lines, are from: l, Thurlow *et al.* (1983).

promising. For the large rRNA molecules, although considerable progress has been made in the crystallization of ribosomal subunits (e.g. Arad *et al.*, 1984; Yonath *et al.*, 1984), it will obviously be a long time before structural resolution at the atomic level can even be contemplated.

At the nucleotide level, Peattie & Gilbert (1980) showed that a chemical modification analysis could be used to monitor the denaturation of both tertiary and secondary structure in tRNA, specific residues becoming progressively accessible to the modifying reagents as the tertiary and then the secondary structure melts out. A similar approach, using both chemical modification and enzymic cleavage, has been applied by many authors to the study of the structure of 5S, 4.5S and 5.8S rRNA, both free in solution or *in situ* in the ribosome [e.g. McDougall & Nazar, 1983; Rabin *et al.*, 1983; Silberklang *et al.*, 1983; Miura *et al.*, 1983; Pieler *et al.*, 1984; Göringer *et al.*, 1984a (5S RNA); Kumagai *et al.*, 1983 (4.5S RNA); Liu *et al.*, 1983 (5.8S RNA)]. For the large rRNA molecules, chemical modification and enzymic cleavage were used to help derive the secondary structure models (see preceding section, and Brimacombe *et al.*, 1983, for review), and more recently very detailed studies have been made of the 3'-terminal domains of both small and large subunit rRNA (Douthwaite *et al.*, 1983; Garrett *et al.*, 1984). [The 3'-terminal domain of 16S RNA has in addition been studied by n.m.r. and microcalorimetry (Heus *et al.*, 1983a,b).] The effects of tRNA binding or subunit association on the accessibility of individual residues in the RNA have also been reported (e.g. Meier & Wagner, 1984, 1985). These experiments have been confined to a study of the terminal regions of the rRNA molecules concerned, since they rely on end-labelling techniques and sequencing gels for the analysis of the modified or cleaved RNA sites and in consequence are limited by the resolving power of the sequencing gels (approx. 150–200 nucleotides). The method of Qu *et al.* (1983) using a DNA primer and reverse transcription offers the possibility of gathering information from the whole length of the RNA molecule, and the whole molecule can of course also be examined if uniformly labelled RNA is used, as in the extensive studies by Noller and his colleagues on the modification of rRNA by kethoxal *in situ* in ribosomes or ribosomal subunits in various functional states (e.g. Brow & Noller, 1983; Hogan *et al.*, 1984).

All of these approaches suffer from the disadvantage that, although they yield precise information on the environment of individual nucleotides within the secondary or tertiary structures, they do not enable a tertiary structure to be deduced. In other words the interpretation of a modification

analysis in terms of a known crystal structure, as made by Peattie & Gilbert (1980), does not work in reverse. Another point to be made here is that the results from this type of study concerning the implied shielding of certain rRNA residues in different functional states (e.g. in the presence of tRNA, or in 70S ribosomes as opposed to isolated subunits) must be interpreted with caution. This has recently been made very clear by the experiments of Meier & Wagner (1984, 1985) who found that some of the guanine residues in *E. coli* 16S RNA that were protected from kethoxal modification in 70S ribosomes or polysomes (Brow & Noller, 1983) actually showed an enhanced reactivity to dimethyl sulphate. Since dimethyl sulphate reacts with N-7 of guanine, whereas kethoxal binds to the opposite side of the purine ring, this result shows that the observed effects are due to a rotation of the guanine residue concerned, rather than to a 'bulk shielding' by another ribosomal component. The real significance of this type of data will only become clear when more is known about the tertiary structure from other approaches.

More direct, albeit often less precise, information on the tertiary folding of the rRNA can be obtained by intramolecular cross-linking. Intramolecular cross-links fall into two classes, namely those that are 'within' secondary structural elements (and which therefore help to define the secondary structure) and those that lie 'outside' the secondary structure (and which therefore reveal the tertiary folding). Both types of cross-link have been identified in *E. coli* rRNA. In 5S RNA, secondary structural cross-links in the stem region were found (Wagner & Garrett, 1978; Rabin & Crothers, 1979), as well as a tertiary cross-link between residues G-41 and G-72 (Hancock & Wagner, 1982). A further cross-link, between residues G-69 and G-108, was identified in 5S RNA cross-linked *in situ* in 50S subunits (Stiege *et al.*, 1982).

In the 16S and 23S rRNA psoralen derivatives have been used to identify both types of cross-link. The early experiments (Wollenzien *et al.*, 1979) relied on electron microscopy to identify the positions of tertiary cross-links in the 16S RNA, but this method is not sufficiently accurate to allow an unambiguous localization of the cross-links in the sequence. Subsequently, gel electrophoretic separation of partially digested cross-linked RNA combined with photo-reversal of the psoralen reaction enabled some secondary structural cross-links to be detected in 16S (Turner *et al.*, 1982) and 23S (Turner & Noller, 1983) RNA, as well as a series of both secondary and tertiary cross-links in 16S RNA (Thompson & Hearst, 1983a). Following photoreversal the precise sites of psoralen reaction cannot in general be directly determined, and these cross-link assignments were therefore made on the

basis of the known preference of psoralen for U residues (Bachelierie *et al.*, 1981).

Psoralen has the disadvantage that it does not react appreciably with active ribosomal subunits (Thammana *et al.*, 1979), and all of the experiments just quoted were made on isolated 16S or 23S RNA. A tertiary cross-link in 16S RNA identified by Expert-Bezançon *et al.* (1983) was also formed in isolated RNA, in this case using a glyoxal derivative as reversible cross-linker. In such experiments there is a real danger that, although secondary structural features are likely to 'snap back' into their correct configurations after a denaturing RNA isolation procedure, the tertiary structure may be very different to that in the native ribosome. It should be remembered in this context that rRNA never exists as a separate entity in the cell, so that even a careful 'renaturing' incubation of the isolated RNA may not be very meaningful. Chu *et al.* (1983) have attempted to circumvent this problem by allowing the psoralen to intercalate into inactive subunits, which are then re-activated by dialysis prior to the irradiative cross-linking step. Although a reasonable level of psoralen incorporation can be achieved by this procedure, our own experience is that, after phenol extraction or gel electrophoresis, only very small amounts remain bound to the RNA (M. Kosack & R. Brimacombe, unpublished results). This exacerbates an ever-present problem in any cross-linking or chemical modification study, namely that one can never be certain that the observed cross-link or modified site does not represent a minor or irrelevant fraction of the ribosome population.

In our own experiments, irreversible cross-links are introduced into RNA within intact ribosomal subunits, either by direct u.v. irradiation or by reaction with nitrogen mustard. After partial digestion, the cross-linked products are isolated by two-dimensional electrophoresis and the cross-link sites determined by classical fingerprinting techniques. The early studies (Zwieb & Brimacombe, 1980; Glotz *et al.*, 1981) only led to the identification of secondary structural cross-links, but, by using different partial digestion conditions, a number of tertiary as well as secondary intra-RNA cross-links have now been localized in both 50S (Stiege *et al.*, 1982, 1983) and in 30S (J. Atmadja & R. Brimacombe, unpublished results) subunits. Very recently we have demonstrated that in the

case of the u.v.-induced cross-linking reaction the same cross-links are formed *in vivo* in 70S ribosomes when growing *E. coli* cells are irradiated (W. Stiege & R. Brimacombe, unpublished results). The locations of all the cross-links described in this section are included in Figs. 1 and 2.

tRNA has also been used as a target for cross-linking studies. In one series of experiments, Prince *et al.* (1982) showed that the hypermodified uridine at the 5'-anticodon position of tRNA^{Val} or tRNA^{Ser} is cross-linked to C-1400 in the *E. coli* 16S RNA sequence. Furthermore a precisely analogous cross-link has been identified in yeast 18S RNA (Ofengand *et al.*, 1982), and the same authors have described an elegant method for the localization of such cross-link sites by a combined end-labelling and cleavage procedure (Ehresmann & Ofengand, 1984). In another series of experiments, Barta *et al.* (1984), using reverse transcriptase to identify the cross-linked nucleotides, identified U-2584 and U-2585 in the *E. coli* 23S RNA as the sites of reaction with a Phe-tRNA molecule carrying an affinity label on the amino acid residue. In both sets of experiments, the tRNA was bound to the ribosomal P-site, and in both cases the cross-links are to highly conserved regions in the 16S and 23S RNA, respectively. Thus the locations of 'both ends' of a tRNA molecule at the P-site are now precisely pin-pointed in the ribosomal RNA, and these positions are also included in Figs. 1 and 2. A cross-link reported in older experiments between a poly(U) analogue and 16S RNA (Wagner *et al.*, 1976) is also relevant in this context.

Finally in this section it should be noted that interactions between isolated 16S and 23S RNA (Burma *et al.*, 1983) and between 5S and 18S RNA (Azad, 1979; Kelly & Cox, 1982) have been reported. 16S and 23S RNA have also been cross-linked together within 70S ribosomes (Zwieb *et al.*, 1978). So far, however, neither a precise cross-link site nor a phylogenetically conserved base-paired interaction have been identified in these cases.

RNA-protein interaction

The ribosomal proteins can be divided into two classes, namely those which are able to form specific complexes with isolated rRNA and those which are not. For the proteins in the first class, 'binding sites' on the RNA have been determined

Fig. 2. Secondary structure of *E. coli* 23S RNA, showing sites of intra-RNA and RNA-protein cross-linking, and protein binding sites (cf. Fig. 1)

The structure is divided into four domains as in Maly & Brimacombe (1983). Intra-RNA cross-links (cf. Fig. 1) are from: m, Stiege *et al.* (1983); n, Glotz *et al.* (1981); o, Stiege *et al.* (1982); p, Turner & Noller (1983); q, Barta *et al.* (1984), tRNA. RNA-protein cross-links are from: r, Wower *et al.* (1981); s, Maly *et al.* (1980). Protein binding sites are from: t, Brimacombe *et al.* (1983) (review); u, Schmidt *et al.* (1981); v, Vester & Garrett (1984). Other references are as in the legend to Fig. 1.

in a number of cases by partial digestion of the RNA-protein complex followed by analysis of the protein-protected RNA fragments. In *E. coli*, such binding sites have been reported for proteins L18 and L25 on 5S RNA, for S4, S7, S8, S15, S17 and S20 on 16S RNA, and for L1, L11, L23 and L24 on 23S RNA (see Zimmermann, 1980; Brimacombe *et al.*, 1983, for reviews). Detailed descriptions of some of these binding sites have been made, using the same type of nuclease or chemical accessibility approach as that described in the foregoing section for probing the tertiary structure of the RNA. Thus the sites of L18 and L25 on 5S RNA are now well documented (Garrett *et al.*, 1981), and ribonuclease α -sarcin has been recently used to 'footprint' these sites as well as that of L5 in the presence of L18 (Huber & Wool, 1984). A corresponding protein binding site on eukaryotic 5S RNA has been studied in a similar way (Nazar & Wildeman, 1983). The *E. coli* L25-5S RNA complex has been examined by n.m.r. (Kime & Moore, 1984), and the crystallization of an L25-5S RNA fragment complex was already noted in the preceding section.

In 16S RNA, a careful study has been made of the binding site for protein S8 (Thurlow *et al.*, 1983), as well as of complexes containing S8 together with S15, S6 and S18 (Gregory *et al.*, 1984), and the position of the S8 site on the RNA is indicated in Fig. 1. The interaction of initiation factor IF3 with the 3'-terminal loop of 16S RNA has also been analysed (Wickstrom, 1983); here it should be noted that an older report involving protein S1 and this same RNA region now appears to represent a non-specific interaction (reviewed by Subramanian, 1983). In 23S RNA, protein L1 has a well-defined and highly-conserved binding site (reviewed by Brimacombe *et al.*, 1983), and the binding site of L23 has recently been described in detail (Vester & Garrett, 1984). Protein L11 has a very small binding site on the 23S RNA (Schmidt *et al.*, 1981), and the 'L8 complex' (consisting of protein L10 and two dimers of protein L7/L12) binds to the same region (Beauclerk *et al.*, 1984). The positions of the sites for the single proteins are shown in Fig. 2, and a region of 23S RNA identified in a ribonucleoprotein fragment together with 5S RNA and proteins L5, L18 and L25 (Branlant *et al.*, 1981) is also indicated. Other proteins or groups of proteins, such as S4, S20, L24 or S7/S9/S10/S13/S19 (see Brimacombe *et al.*, 1983, for review) have been found associated with rather large regions of the 16S or 23S RNA (not shown in Figs. 1 and 2), and it is clear that the locations of potential nuclease sites within the tertiary structure of the RNA, as well as in regions 'protected' by the protein, influence the size and nature of the binding site identified for any given

protein by this approach (Brimacombe *et al.*, 1983).

The alternative approach to the study of RNA-protein interactions is the application of RNA-protein cross-linking techniques. Cross-linking is a purely topographical probe, and the existence of a cross-link between two ribosomal components does not necessarily imply a concomitant 'binding' or physical interaction. Nonetheless, as with the intra-RNA cross-linking described in the preceding section, an RNA-protein cross-linking analysis has the potential of providing precisely the type of information needed to integrate the RNA structure into the ribosomal particles (see the following section), as well as giving data on those proteins that are unable to bind to the rRNA. The methods and reagents that have been applied to RNA-protein cross-linking in the ribosome, as well as the problems associated with this approach, were reviewed in detail recently (Brimacombe *et al.*, 1983), and so will only be treated briefly here. Since our above-mentioned review, Brewer & Noller (1983) have described a new cross-linking reagent which is a bis-kethoxal derivative, and the use of γ -irradiation for inducing RNA-protein cross-links has also been reported (Cazillis *et al.*, 1984). However, since it is already clear that most if not all of the ribosomal proteins are capable of being cross-linked to their cognate RNA molecules *in situ* in the ribosomal subunits by one reagent or another, the question of determining the sites of cross-linking to individual proteins is now the dominant problem. A number of such cross-linking sites to various proteins on *E. coli* 23S RNA (Wower *et al.*, 1981) or 16S RNA (Wower & Brimacombe, 1983) have been determined by using 2-iminothiolane as the cross-linking agent, and these results are included in Figs. 1 and 2, together with some more recent assignments obtained after cross-linking with 2-iminothiolane or nitrogen mustard (A. Kyriatsoulis, I. Wower & R. Brimacombe, unpublished results). Fig. 1 also includes a site of cross-linking to protein S12 (Chiaruttini *et al.*, 1982), as well as older data concerning proteins cross-linked to the 3'-terminus of the 16S RNA (see Brimacombe *et al.*, 1983).

In these experiments the cross-linking reaction is carried out with intact ribosomal subunits, and involves at some stage a partial digestion procedure and the isolation of a cross-linked protein-RNA fragment. Since the mixture of products obtained by such a procedure is exceedingly complex, we do not consider a cross-link site to be established unless the presence of the protein concerned on the RNA fragment analysed has been positively demonstrated. For this reason, the recently-described cross-link of elongation factor EF-G to residue 1067 of the 23S RNA (Sköld,

1983) is not included in Fig. 2, although this result would fit well to other data concerning this region of the RNA (Garrett, 1983). EF-G has been cross-linked to the 3'-region of 23S RNA (Bochkareva & Girshovich, 1984), but the cross-link has not yet been localized. This latter result, together with other cross-link assignments where the RNA site could not be pin-pointed to within about ten nucleotides, are also not included in Figs. 1 and 2. Chiam & Wagner (1983) have demonstrated a number of RNA-protein cross-links across the ribosomal subunit interface, but again no precise sites have so far been reported.

Packing the rRNA into the ribosomal subunits

Until quite recently almost nothing was known about the topography of the rRNA within the ribosomal subunits, although data concerning the corresponding arrangement of the ribosomal proteins in *E. coli* have been accumulating for a number of years (see Wittmann, 1983, for review). The most important techniques for studying protein topography include the identification of protein pairs that can be cross-linked together (e.g. Lambert *et al.*, 1983), the measurement of inter-protein distances and protein shapes by neutron scattering (e.g. Ramakrishnan *et al.*, 1984; Nierhaus *et al.*, 1983), the measurement of distances between specific amino acids on different proteins carrying fluorescent probes (e.g. Steinhäuser *et al.*, 1983), and the localization of antigenic sites in individual proteins on the subunit surfaces by immune electron microscopy (e.g. Stöffler-Meilicke *et al.*, 1983; Winkelman & Kahan, 1983; Breitenreuter *et al.*, 1984). The experimental error associated with the physical measurements is relatively large, but nonetheless there is a good consensus of agreement between the various sets of data.

In all of these methods the proteins may be considered as amorphous spheres or ellipsoids, and thus, although the amino acid sequences of all the *E. coli* ribosomal proteins are known (Wittmann, 1982), the topographical studies have in general been conducted without reference to these sequences. Only in a few cases have the precise amino acid residues involved in protein-protein cross-links been localized (e.g. Allen *et al.*, 1979). The amino acid sequence data will be able to be incorporated at a later stage, when (for example) more X-ray crystallographic data on the individual proteins become available. In contrast, since there is only one large rRNA molecule per ribosomal subunit, the RNA topography problem cannot be considered without reference to the sequence, and from the outset the question is how to package the 16S and 23S rRNA sequences into the known dimensions of the 30S and 50S subunits and how to

intersperse the sequences correctly in relation to the known distribution of the ribosomal proteins, as just outlined above. To this end, the data described in the preceding sections, concerning secondary structure of the rRNA, intra-RNA and RNA-protein cross-linking, and RNA-protein binding sites (Figs. 1 and 2), are all obviously directly relevant. In addition, the distances between the 3'-ends of 5S, 16S and 23S RNA in 70S ribosomes are known from energy transfer measurements with fluorescent probes (Odom *et al.*, 1980), and the distances between the 3'-end of 16S RNA and probes attached to proteins S1 and S21 have similarly been determined in various functional states (e.g. Odom *et al.*, 1984a). In this context the two highly specific cross-links to tRNA already discussed (see section on inter-RNA interaction) represent in effect a 'distance measurement' between residues 1400 in 16S RNA and residues 2584/5 in the 23S RNA (cf. Figs. 1 and 2), since tRNA is a rather rigid molecule. tRNA has also been linked to protein S19 via an affinity label attached to the U-8 position (Lin *et al.*, 1984), although in this latter instance the tRNA was located at the A-site, not the P-site.

There is now a good consensus of agreement regarding the detailed three-dimensional shapes of the 30S and 50S subunits as determined by electron microscopy (see Lake, 1983, for review). Three-dimensional reconstructions have been made [e.g. Verschoor *et al.*, 1984 (30S)], and the question of the handedness of the structures has also been resolved with reasonable certainty [e.g. Vasiliev *et al.*, 1983 (50S)]. The shape of the rRNA within the subunits is beginning to emerge from electron microscopic studies both by negative staining [Knauer *et al.*, 1983 (30S); Oettl *et al.*, 1983 (50S)] and by electron spectroscopic imaging [Korn *et al.*, 1983 (30S)]. Physical techniques have been used to demonstrate that only six proteins need to bind for the 16S RNA to acquire the compactness of a 30S particle (Serdyuk *et al.*, 1983).

The well-defined shapes of the ribosomal subunits have enabled a number of sites on the rRNA to be mapped by immune electron microscopy. These sites include the 3'-ends of 5S and 23S RNA in the *E. coli* 50S subunit, the 5'- and 3'-ends of 16S RNA in the 30S subunit, as well as the location in the 30S of the two adjacent dimethyladenosine residues (positions 1518–1519, Fig. 1) in 16S RNA (see Brimacombe *et al.*, 1983, for review and references). More recently the location of the 7-methyl guanine residue in 16S RNA (position 527, Fig. 1) has been localized in the 30S subunit (Trempe *et al.*, 1982), and the 5'-end of 5S RNA in the 50S subunit has been placed (and the position of the 3'-end of the 23S RNA confirmed) using an RNA processing-deficient mutant (Clark & Lake, 1984).

The 5'- and 3'-ends of the 5S RNA are, as expected from the secondary structure, at the same position in the 50S subunit, and the secondary structure of the 23S RNA (Fig. 2) predicts that the 5'- and 3'-ends of the molecule should also have coincident locations, in contrast with the 5'- and 3'-ends of 16S RNA (Fig. 1). The hypermodified base in tRNA cross-linked to C-1400 of the 16S RNA (Fig. 1) has been located on the 30S particle (Gornicki *et al.*, 1984), and haptenized poly(U) also maps in the same region (Evstafieva *et al.*, 1983). The latter result is of particular interest, since the 3'- and 5'-ends of the message were found at the same location, suggesting that the message makes a 'U-turn' in the ribosome. This at first sight surprising result is in fact reasonable, since mRNA itself has a defined secondary and tertiary structure, and the 'U-turn' would enable the message to be read with far less disruption to its structure than in the 'threaded-through' concept that is usually assumed in sketches of the functioning ribosome.

All of the above data, together with locations of RNA regions which can be inferred from their proximity to known protein positions in RNA-protein cross-linking or binding studies (see previous section and Figs. 1 and 2), are illustrated in Fig. 3. The reader is also referred to Spirin (1983) for a model of the location of tRNA on the ribosome.

Function of rRNA

The high degree of structural conservation in rRNA strongly implies that the RNA is directly involved in the ribosomal function, and it is indeed likely (Crick, 1968) that the primitive ribosome was an RNA molecule. The involvement of the 3'-terminus of the 16S rRNA in mRNA binding (Shine & Dalgarno, 1974; Steitz & Jakes, 1975) is now well-established, and resistance to several antibiotics has been traced to point mutations in 16S or 23S RNA (e.g. Garrett, 1983; Sigmund *et al.*, 1984). The mutations in 23S RNA are implicated in peptidyl transfer, and as already mentioned above the same region of the 23S RNA structure has been identified in a cross-link to a peptidyl-tRNA affinity analogue (Barta *et al.*, 1984). This latter RNA region and the corresponding region in 16S RNA that is cross-linked to the anticodon of tRNA (Prince *et al.*, 1982; cf. Figs. 1 and 2) are both part of the highly conserved secondary structural core (see the section on secondary structure), and preliminary model-building studies in our laboratory indicate that the conserved core is concentrated in the interface regions of both subunits.

Function, however, implies movement, and evidence for specific conformational changes within the ribosomal particles is beginning to

accumulate. For example, 5.8S RNA appears to have two conformers *in situ* in the ribosome (Lo *et al.*, 1984). The distance between the 3'-end of 16S RNA and a fluorescent label on protein S21 increases when the 30S binds to the 50S subunit (Odom *et al.*, 1984b), and the accessibility of the 3'-end of 16S RNA to oligonucleotide binding is altered in active versus inactive subunits (Van Duin *et al.*, 1984). Specific conformational 'switches' (a switch being defined as the existence of two mutually exclusive secondary structural elements) have been proposed by various authors for 16S and 23S RNA (e.g. Glotz & Brimacombe, 1980; Glotz *et al.*, 1981; Thompson & Hearst, 1983a; Atmadja *et al.*, 1984) and also for 5S RNA (Trifonov & Bolshoi, 1983; De Wachter *et al.*, 1984).

The most likely stage in the protein biosynthetic process where the rRNA might need to undergo radical conformational changes is during the translocation process (Glotz & Brimacombe, 1980; Thompson & Hearst, 1983b), since in this step the entire mRNA-tRNA-peptide complex must be physically shifted by one codon length across the ribosome, and this could be achieved by a cycle of switches in the ribosomal RNA. [In this context the reader is also referred to the work of Nierhaus & Rheinberger (1984), who have found strong evidence for the existence of a third tRNA binding site on the ribosome, in addition to the classical A- and P-sites.] A very promising way to test hypothetical RNA functions or switches is to use rRNA molecules with artificially altered sequences. The altered sequences can be obtained either by manipulation *in vitro* or by site-directed mutagenesis. An example of the former is the early work of Zagorska *et al.* (1980) who showed the effect on 30S subunit function of removing 160 nucleotides from the 3'-end of 16S RNA. More recently it has been elegantly shown that 5S RNA can function normally in the absence of the conserved GAAC sequence at positions 44-47 (Zagorska *et al.*, 1984), thereby discounting an earlier proposal that this sequence is involved in binding the GTTC sequence of tRNA to the ribosome. In the case of site-directed mutagenesis, rRNA molecules have been constructed with modifications in the central region (bases 600-900) of 16S RNA (Stark *et al.*, 1984; Zwieb & Dahlberg, 1984a), in the 1760-region of 23S RNA (Zwieb & Dahlberg, 1984b), and in 5S RNA (Göringer *et al.*, 1984b). This approach is still in its infancy, and at the moment is limited both by the lack of sensitivity of available functional assays as well as by simply not knowing where in the rRNA structure to place the mutations; the 4500 bases in 16S and 23S RNA offer a rather wide range of choice. Nevertheless, as information on the three-

50S Subunit

FRONT

REAR

Labels for FRONT view:

- C1400 (tRNA)
- 1240X {1377-78X}
- S7
- 3'-end, 1542
- 723-24X {1542X}
- S21
- m⁶A, 1518-19
- S17
- 1323X
- 629-33X
- (S8)
- (S12)
- 5'-end
- 580-650B {629-33X {651-54X}}
- m⁷G, 527

Labels for REAR view:

- 1240X {1377-78X}
- S7
- 723-24X {1542X}
- S21
- 1323X
- (S12)
- S17
- 629-33X
- (S8)
- 580-650B {629-33X {651-54X}}
- 5'-end

The RNA sites were either directly localized by immune electron microscopy (see the text for references), or deduced from their proximity to a protein antigenic site on the subunit surface (Stöffler-Meilicke *et al.*, 1983; Breitenreuter *et al.*, 1984; Lake, 1983). In the latter cases, the RNA regions are distinguished by the suffix 'B' (deduced from a protein binding site) or 'X' (deduced from an RNA-protein cross-link site; cf. Figs. 1 and 2). The numbers give the position in the rRNA sequence (from the 5'-end). The approximate locations of proteins S8 and S12 are inferred from the neutron data of Ramakrishnan *et al.* (1984).

The authors are grateful to Dr. H. G. Wittmann for his critical reading of the manuscript.

Abdel-Meguid, S. S., Moore, P. B. & Steitz, T. A. (1983) *J. Mol. Biol.* **171**, 207-215
Allen, G., Capasso, R. & Gualerzi, C. (1979) *J. Biol. Chem.* **254**, 9800-9806

- Anderson, S., de Bruijn, M. H. L., Coulson, A. R., Eperon, I. C., Sanger, F. & Young, I. G. (1982) *J. Mol. Biol.* **156**, 683–717
- Arad, T., Leonard, K., Wittmann, H. G. & Yonath, A. (1984) *EMBO J.* **3**, 127–131
- Atmadja, J., Brimacombe, R. & Maden, B. E. H. (1984) *Nucleic Acids Res.* **12**, 2649–2667
- Azad, A. A. (1979) *Nucleic Acids Res.* **7**, 1913–1929
- Bachelierie, J. P., Thompson, J. F., Wegnez, M. R. & Hearst, J. E. (1981) *Nucleic Acids Res.* **9**, 2207–2222
- Barta, A., Steiner, G., Brosius, J., Noller, H. F. & Kuechler, E. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3607–3611
- Beaclarck, A. A. D., Cundliffe, E. & Dijk, J. (1984) *J. Biol. Chem.* **259**, 6559–6563
- Bochkareva, E. S. & Girshovich, A. S. (1984) *FEBS Lett.* **171**, 202–206
- Branlant, C., Krol, A., Machatt, M. A., Pouyet, J., Ebel, J. P., Edwards, K. & Kössel, H. (1981) *Nucleic Acids Res.* **9**, 4303–4324
- Breitenreuter, G., Lotti, M., Stöffler-Meilicke, M. & Stöffler, G. (1984) *Mol. Gen. Genet.* **197**, 189–195
- Brewer, L. A. & Noller, H. F. (1983) *Biochemistry* **22**, 4310–4315
- Brimacombe, R. (1984) *Trends Biochem. Sci.* **9**, 273–277
- Brimacombe, R., Maly, P. & Zwieb, C. (1983) *Prog. Nucleic Acid Res.* **28**, 1–48
- Brosius, J., Palmer, M. L., Kennedy, P. J. & Noller, H. F. (1978) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 201–204
- Brosius, J., Dull, T. J. & Noller, H. F. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 201–204
- Brow, D. A. & Noller, H. F. (1983) *J. Mol. Biol.* **163**, 27–46
- Burma, D. P., Nag, B. & Tewari, D. S. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4875–4878
- Carbon, P., Ehresmann, C., Ehresmann, B. & Ebel, J. P. (1978) *FEBS Lett.* **94**, 152–156
- Carbon, P., Ehresmann, C., Ehresmann, B. & Ebel, J. P. (1979) *Eur. J. Biochem.* **100**, 399–410
- Carbon, P., Ebel, J. P. & Ehresmann, C. (1981) *Nucleic Acids Res.* **9**, 2325–2333
- Cazillis, M., Giocanti, N., Houssais, J. F. & Ekert, B. (1984) *Eur. J. Biochem.* **139**, 439–445
- Chan, Y. L., Olvera, J. & Wool, I. G. (1983) *Nucleic Acids Res.* **11**, 7819–7831
- Chan, Y. L., Gutell, R., Noller, H. F. & Wool, I. G. (1984) *J. Biol. Chem.* **259**, 224–230
- Chao, S., Sederoff, R. & Levings, C. S. (1984) *Nucleic Acids Res.* **12**, 6629–6644
- Chiam, C. L. & Wagner, R. (1983) *Biochemistry* **22**, 1193–1200
- Chiaruttini, C., Expert-Bezançon, A., Hayes, D. & Ehresmann, B. (1982) *Nucleic Acids Res.* **10**, 7657–7676
- Chu, Y. G., Wollenzien, P. L. & Cantor, C. R. (1983) *J. Biomol. Struct. Dynamics* **1**, 647–656
- Clark, M. W. & Lake, J. A. (1984) *J. Bacteriol.* **157**, 971–974
- Clark, C. G., Tague, B. W., Ware, V. C. & Gerbi, S. A. (1984) *Nucleic Acids Res.* **12**, 6197–6220
- Connaughton, J. F., Rairkar, A., Lockard, R. E. & Kumar, A. (1984) *Nucleic Acids Res.* **12**, 4731–4745
- Crick, F. H. C. (1968) *J. Mol. Biol.* **38**, 367–379
- Czernilofsky, A. P., Kurland, C. G. & Stöffler, G. (1975) *FEBS Lett.* **58**, 281–284
- Dale, R. M. K., Mendu, N., Ginsburg, H. & Kridl, J. C. (1984) *Plasmid* **11**, 141–150
- De Wachter, R., Chen, M. W. & Vandenberghe, A. (1984) *Eur. J. Biochem.* **143**, 175–182
- Douglas, S. E. & Doolittle, W. F. (1984) *Nucleic Acids Res.* **12**, 3373–3386
- Douthwaite, S., Christensen, A. & Garrett, R. A. (1983) *J. Mol. Biol.* **169**, 249–279
- Dron, M., Rahire, M. & Rochaix, J. D. (1982) *Nucleic Acids Res.* **10**, 7609–7620
- Edwards, K. & Kössel, H. (1981) *Nucleic Acids Res.* **9**, 2853–2869
- Ehresmann, C. & Ofengand, J. (1984) *Biochemistry* **23**, 438–445
- Eperon, I. C., Anderson, S. & Nierlich, D. P. (1980) *Nature (London)* **286**, 460–467
- Eperon, I. C., Janssen, J. W. G., Hoeijmakers, J. H. J. & Borst, P. (1983) *Nucleic Acids Res.* **11**, 105–125
- Erdmann, V. A., Wolters, J., Huysmans, E., Vandenberghe, A. & de Wachter, R. (1984) *Nucleic Acids Res.* **12**, r133–r166
- Evstafieva, A. G., Shatsky, I. N., Bogdanov, A. A., Semenov, Y. P. & Vasiliev, V. D. (1983) *EMBO J.* **2**, 799–804
- Expert-Bezançon, A., Milet, M. & Carbon, P. (1983) *Eur. J. Biochem.* **136**, 267–274
- Fox, G. & Woese, C. R. (1975) *Nature (London)* **256**, 505–507
- Garrett, R. A. (1983) *Trends Biochem. Sci.* **8**, 189–190
- Garrett, R. A., Douthwaite, S. & Noller, H. F. (1981) *Trends Biochem. Sci.* **6**, 137–139
- Garrett, R. A., Christensen, A. & Douthwaite, S. (1984) *J. Mol. Biol.* **179**, 689–712
- Georgiev, O. I., Nikolaev, N., Hadjiolov, A. A., Skryabin, K. G., Zakharyev, V. M. & Bayev, A. A. (1981) *Nucleic Acids Res.* **9**, 6953–6958
- Georgiev, O. I., Dudov, K. P., Hadjiolov, A. A. & Skryabin, K. G. (1984) *Folia Biol. (Prague)* **30**, 1–14
- Glitz, C. & Brimacombe, R. (1980) *Nucleic Acids Res.* **8**, 2377–2395
- Glitz, C., Zwieb, C., Brimacombe, R., Edwards, K. & Kössel, H. (1981) *Nucleic Acids Res.* **9**, 3287–3306
- Glover, D. M. (1981) *Cell* **26**, 297–298
- Göringer, H. U., Szymkowiak, C. & Wagner, R. (1984a) *Eur. J. Biochem.* **144**, 25–34
- Göringer, U., Wagner, R., Jacob, W. F., Dahlberg, A. E. & Zwieb, C. (1984b) *Nucleic Acids Res.* **12**, 6935–6950
- Gornicki, P., Nurse, K., Hellmann, W., Boublik, M. & Ofengand, J. (1984) *J. Biol. Chem.* **259**, 10493–10498
- Graf, L., Roux, E., Stutz, E. & Kössel, H. (1982) *Nucleic Acids Res.* **10**, 6369–6381
- Gray, M. W., Sankoff, D. & Cedergren, R. J. (1984) *Nucleic Acids Res.* **12**, 5837–5852
- Gregory, R. J., Zeller, M. L., Thurlow, D. L., Gourse, R. L., Stark, M. J. R., Dahlberg, A. E. & Zimmermann, R. A. (1984) *J. Mol. Biol.* **178**, 287–302
- Gupta, R., Lanter, J. M. & Woese, C. R. (1983) *Science* **221**, 656–659
- Hadjiolov, A. A., Georgiev, O. I., Nosikov, V. V. & Yavachev, L. P. (1984) *Nucleic Acids Res.* **12**, 3677–3693

- Hancock, J. & Wagner, R. (1982) *Nucleic Acids Res.* **10**, 1257-1269
- Hassouna, N., Michot, B. & Bachellerie, J. P. (1984) *Nucleic Acids Res.* **12**, 3563-3583
- Heus, H. A., van Kimmenade, J. M., van Knippenberg, P. H., Haasnoot, C. A., de Bruin, S. H. & Hilbers, C. W. (1983a) *J. Mol. Biol.* **170**, 939-956
- Heus, H. A., van Kimmenade, J. M., van Knippenberg, P. H. & Hinz, H. J. (1983b) *Nucleic Acids Res.* **11**, 203-210
- Hogan, J. J., Gutell, R. R. & Noller, H. F. (1984) *Biochemistry* **23**, 3322-3330
- Huber, P. W. & Wool, I. G. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 322-326
- Iwami, M., Muto, A., Yamao, F. & Osawa, S. (1984) *Mol. Gen. Genet.* **196**, 317-322
- Kan, L. S., Chandrasegaran, S., Pulford, S. M. & Miller, P. S. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4263-4265
- Kelly, J. M. & Cox, R. A. (1981) *Nucleic Acids Res.* **9**, 1111-1121
- Kelly, J. M. & Cox, R. A. (1982) *Nucleic Acids Res.* **10**, 6733-6745
- Kime, M. J. & Moore, P. B. (1984) *Biochemistry* **23**, 1688-1695
- Klein, B. K., King, T. C. & Schlessinger, D. (1983) *J. Mol. Biol.* **168**, 809-830
- Klein, B. K., Forman, P., Shiomi, Y. & Schlessinger, D. (1984) *Biochemistry* **23**, 3927-3933
- Knauer, V., Hegerl, R. & Hoppe, W. (1983) *J. Mol. Biol.* **163**, 409-430
- Kobayashi, M., Seki, T., Katsuyuki, Y. & Koike, K. (1981) *Gene* **16**, 297-307
- Köchel, H. G. & Küntzel, H. (1981) *Nucleic Acids Res.* **9**, 5689-5696
- Köchel, H. G. & Küntzel, H. (1982) *Nucleic Acids Res.* **10**, 4795-4801
- Kop, J., Kopylov, A. M., Magrum, L., Siegel, R., Gupta, R., Woese, C. R. & Noller, H. F. (1984a) *J. Biol. Chem.* **259**, 15287-15319
- Kop, J., Wheaton, V., Gupta, R., Woese, C. R. & Noller, H. F. (1984b) *DNA* **3**, 347-357
- Korn, A. P., Spitnik-Elson, P., Elson, D. & Ottensmeyer, F. P. (1983) *Eur. J. Cell Biol.* **31**, 334-340
- Kumagai, I., Bartsch, M., Subramanian, A. R. & Erdmann, V. A. (1983) *Nucleic Acids Res.* **11**, 961-970
- Kumano, M., Tomioka, N. & Sugiura, M. (1983) *Gene* **24**, 219-225
- Küntzel, H. & Köchel, H. G. (1981) *Nature (London)* **293**, 751-755
- Küntzel, H., Piechulla, B. & Hahn, U. (1983) *Nucleic Acids Res.* **11**, 893-900
- Ladner, J. E., Jack, A., Robertus, J. D., Brown, R. S., Rhodes, D., Clark, B. F. C. & Klug, A. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4414-4418
- Lake, J. A. (1983) *Prog. Nucleic Acid Res.* **30**, 163-194
- Lambert, J. M., Boileau, G., Cover, J. A. & Traut, R. R. (1983) *Biochemistry* **22**, 3913-3920
- Leffers, H. & Garrett, R. A. (1984) *EMBO J.* **3**, 1613-1619
- Lin, F. L., Kahan, L. & Ofengand, J. (1984) *J. Mol. Biol.* **172**, 77-86
- Liu, W., Lo, A. C. & Nazar, R. N. (1983) *J. Mol. Biol.* **171**, 217-224
- Lo, A. C., Liu, W. & Nazar, R. N. (1984) *Eur. J. Biochem.* **141**, 549-553
- MacKay, R. M. (1981) *FEBS Lett.* **123**, 17-18
- Maly, P. & Brimacombe, R. (1983) *Nucleic Acids Res.* **11**, 7263-7286
- Maly, P., Rinke, J., Ulmer, E., Zwieb, C. & Brimacombe, R. (1980) *Biochemistry* **19**, 4179-4188
- Mankin, A. S. & Kopylov, A. M. (1981) *Biochem. Int.* **3**, 587-593
- Mankin, A. S., Kopylov, A. M. & Bogdanov, A. A. (1981) *FEBS Lett.* **134**, 11-14
- McCarroll, R., Olsen, G. J., Stahl, Y. D., Woese, C. R. & Sogin, M. L. (1983) *Biochemistry* **22**, 5858-5867
- McDougall, J. & Nazar, R. N. (1983) *J. Biol. Chem.* **258**, 5256-5259
- Meier, N. & Wagner, R. (1984) *Nucleic Acids Res.* **12**, 1473-1487
- Meier, N. & Wagner, R. (1985) *Eur. J. Biochem.* **146**, 83-87
- Messing, J., Carlson, J., Hagen, G., Rubenstein, I. & Oleson, A. (1984) *DNA* **3**, 31-40
- Michot, B., Hassouna, N. & Bachellerie, J. P. (1984) *Nucleic Acids Res.* **12**, 4259-4279
- Miura, K., Tsuda, S., Iwano, T., Ueda, T., Harada, F. & Kato, N. (1983) *Biophys. Acta* **739**, 181-189
- Morikawa, K., Kawakami, M. & Takemura, S. (1982) *FEBS Lett.* **145**, 194-196
- Nazar, R. N. (1980) *FEBS Lett.* **119**, 212-214
- Nazar, R. N. & Wildeman, A. G. (1983) *Nucleic Acids Res.* **11**, 3155-3168
- Nelles, L., Fang, B. L., Volckaert, G., Vandenberghe, A. & de Wachter, R. (1984) *Nucleic Acids Res.* **12**, 8749-8768
- Nierhaus, K. H. & Rheinberger, H. J. (1984) *Trends Biochem. Sci.* **9**, 428-432
- Nierhaus, K. H., Lietzke, R., May, R. P., Nowotny, V., Schulze, H., Simpson, K., Wurmbach, P. & Stuhmann, H. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 2889-2893
- Noller, H. F. (1984) *Annu. Rev. Biochem.* **53**, 119-162
- Noller, H. F. & Woese, C. R. (1981) *Science* **212**, 403-411
- Noller, H. F., Kop, J., Wheaton, V., Brosius, J., Gutell, R. R., Kopylov, A. M., Dohme, F., Herr, W., Stahl, D. A., Gupta, R. & Woese, C. R. (1981) *Nucleic Acids Res.* **9**, 6167-6189
- Odom, O. W., Robbins, D. J., Lynch, J., Dottavio-Martin, D., Kramer, G. & Hardesty, B. (1980) *Biochemistry* **19**, 5947-5954
- Odom, O. W., Dabbs, E. R., Dionne, C., Müller, M. & Hardesty, B. (1984a) *Eur. J. Biochem.* **142**, 216-267
- Odom, O. W., Deng, H. Y., Dabbs, E. R. & Hardesty, B. (1984b) *Biochemistry* **23**, 5069-5076
- Oettl, H., Hegerl, R. & Hoppe, W. (1983) *J. Mol. Biol.* **163**, 431-450
- Ofengand, J., Gornicki, P., Chakraborty, K. & Nurse, K. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2817-2821
- Olsen, G. J., McCarroll, R. & Sogin, M. L. (1983) *Nucleic Acids Res.* **11**, 8037-8049
- Otsuka, T., Nomiyama, H., Yoshida, H., Kukita, T., Kuhara, S. & Sakaki, Y. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3163-3167
- Pavakis, G. N., Jordan, B. R., Wurst, R. M. & Vournakis, J. N. (1979) *Nucleic Acids Res.* **7**, 2213-2238

- Peattie, D. A. & Gilbert, W. (1980) *Proc. Natl. Acad. Sci. U.S.A.* **77**, 4679-4682
- Pieler, T. & Erdmann, V. A. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 4599-4603
- Pieler, T., Schreiber, A. & Erdmann, V. A. (1984) *Nucleic Acids Res.* **12**, 3115-3126
- Prince, J. B., Taylor, B. H., Thurlow, D. L., Ofengand, J. & Zimmermann, R. A. (1982) *Proc. Natl. Acad. Sci. U.S.A.* **79**, 5450-5454
- Qu, H. L., Michot, B. & Bachellerie, J. P. (1983) *Nucleic Acids Res.* **11**, 5903-5920
- Rabin, D. & Crothers, D. M. (1979) *Nucleic Acids Res.* **7**, 689-703
- Rabin, D., Kao, T. & Crothers, D. M. (1983) *J. Biol. Chem.* **258**, 10813-10816
- Ramakrishnan, V., Capel, M., Kjeldgaard, M., Engelmann, D. M. & Moore, P. B. (1984) *J. Mol. Biol.* **174**, 265-284
- Raynal, F., Michot, B. & Bachellerie, J. P. (1984) *FEBS Lett.* **167**, 263-268
- Rubtsov, P. M., Musakhanov, M. N., Zakharyev, V. M., Krayev, A. S., Skryabin, K. G. & Bayev, A. A. (1980) *Nucleic Acids Res.* **8**, 5779-5794
- Saccone, C., Cantatore, P., Gadaleta, G., Gallerani, R., Lanave, C., Pepe, G. & Kroon, A. M. (1981) *Nucleic Acids Res.* **9**, 4139-4148
- Salim, M. & Maden, B. E. H. (1981) *Nature (London)* **291**, 205-208
- Schmidt, F. J., Thompson, J., Lee, K., Dijk, J. & Cundliffe, E. (1981) *J. Biol. Chem.* **256**, 12301-12305
- Schnare, M. N., Spencer, D. F. & Gray, M. W. (1983) *Can. J. Biochem.* **61**, 38-45
- Schwarz, Z. & Kössel, H. (1980) *Nature (London)* **283**, 739-742
- Seilhammer, J. & Cummings, D. J. (1981) *Nucleic Acids Res.* **9**, 6391-6406
- Seilhammer, J. J., Olsen, G. J. & Cummings, D. J. (1984a) *J. Biol. Chem.* **259**, 5167-5172
- Seilhammer, J. J., Gutell, R. R. & Cummings, D. J. (1984b) *J. Biol. Chem.* **259**, 5173-5181
- Serdnyuk, I. N., Agalarov, S. C., Sedelnikova, S. E., Spirin, A. S. & May, R. P. (1983) *J. Mol. Biol.* **169**, 409-425
- Shine, J. & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. U.S.A.* **71**, 1342-1346
- Sigmund, C. D., Ettagebi, M. & Morgan, E. A. (1984) *Nucleic Acids Res.* **12**, 4653-4663
- Silberklang, M., RajBhandary, U. L., Lück, A. & Erdmann, V. A. (1983) *Nucleic Acids Res.* **11**, 605-617
- Sköld, S. E. (1983) *Nucleic Acids Res.* **11**, 4923-4932
- Sor, F. & Fukuhara, H. (1980) *C.R. Acad. Sci. Paris Ser. D.* **291**, 933-936
- Sor, F. & Fukuhara, H. (1983) *Nucleic Acids Res.* **11**, 339-348
- Spencer, D. F., Schnare, M. N. & Gray, M. W. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 493-497
- Spirin, A. S. (1983) *FEBS Lett.* **156**, 217-221
- Stahl, D. A., Luehrsen, K. R., Woese, C. R. & Pace, N. R. (1981) *Nucleic Acids Res.* **9**, 6129-6137
- Stark, M. J., Gregory, R. J., Gourse, R. L., Thurlow, D. L., Zwieb, C., Zimmermann, R. A. & Dahlberg, A. E. (1984) *J. Mol. Biol.* **178**, 303-322
- Steinhäuser, K. G., Woolley, P., Dijk, J. & Epe, B. (1983) *Eur. J. Biochem.* **137**, 337-345
- Steitz, J. A. & Jakes, K. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 4734-4738
- Stiege, W., Zwieb, C. & Brimacombe, R. (1982) *Nucleic Acids Res.* **10**, 7211-7229
- Stiege, W., Glotz, C. & Brimacombe, R. (1983) *Nucleic Acids Res.* **11**, 1687-1706
- Stiegler, P., Carbon, P., Zuker, M., Ebel, J. P. & Ehresmann, C. (1981a) *Nucleic Acids Res.* **9**, 2153-2172
- Stiegler, P., Carbon, P., Ebel, J. P. & Ehresmann, C. (1981b) *Eur. J. Biochem.* **120**, 487-495
- Stöffler-Meilicke, M., Noah, M. & Stöffler, G. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80**, 6780-6784
- Studnicka, G. M., Eiserling, F. A. & Lake, J. A. (1981) *Nucleic Acids Res.* **9**, 1885-1904
- Subramanian, A. R. (1983) *Prog. Nucleic Acid Res.* **28**, 101-142
- Takaiwa, F. & Sugiura, M. (1982) *Eur. J. Biochem.* **124**, 13-19
- Takaiwa, F., Oono, K. & Sugiura, M. (1984) *Nucleic Acids Res.* **12**, 5441-5448
- Thammana, P., Cantor, C. R., Wollenzien, P. L. & Hearst, J. E. (1979) *J. Mol. Biol.* **135**, 271-283
- Thompson, J. F. & Hearst, J. E. (1983a) *Cell* **32**, 1355-1365
- Thompson, J. F. & Hearst, J. E. (1983b) *Cell* **33**, 19-24
- Thurlow, D. L., Ehresmann, C. & Ehresmann, B. (1983) *Nucleic Acids Res.* **11**, 6787-6802
- Thurlow, D. L., Mason, T. L. & Zimmermann, R. A. (1984) *FEBS Lett.* **173**, 277-282
- Tohdoh, N. & Sugiura, M. (1982) *Gene* **17**, 213-218
- Tomioka, N. & Sugiura, M. (1983) *Mol. Gen. Genet.* **191**, 46-50
- Torczyński, R., Bollon, A. P. & Fuke, M. (1983) *Nucleic Acids Res.* **11**, 4879-4890
- Trempe, M. R., Ohgi, K. & Glitz, D. G. (1982) *J. Biol. Chem.* **257**, 9822-9829
- Trifonov, E. N. & Bolshoi, G. (1983) *J. Mol. Biol.* **169**, 1-13
- Turner, S. & Noller, H. F. (1983) *Biochemistry* **22**, 4159-4164
- Turner, S., Thompson, J. F., Hearst, J. E. & Noller, H. F. (1982) *Nucleic Acids Res.* **10**, 2839-2849
- Van Duin, J., Ravensbergen, C. J. C. & Doornbos, J. (1984) *Nucleic Acids Res.* **12**, 5079-5086
- Van Etten, R. A., Walberg, M. W. & Clayton, D. A. (1980) *Cell* **22**, 157-170
- Van Knippenberg, P. H., van Kimmenade, J. M. A. & Heus, H. A. (1984) *Nucleic Acids Res.* **12**, 2595-2604
- Vasiliev, V. D., Selivanova, O. M. & Ryazantsev, S. N. (1983) *J. Mol. Biol.* **171**, 561-569
- Vaughn, J. C., Sperbeck, S. J., Ramsey, W. J. & Lawrence, C. B. (1984) *Nucleic Acids Res.* **12**, 7479-7502
- Veldman, G. M., Klootwijk, J., de Regt, V. C. H. F., Planta, R. J., Branlant, C., Krol, A. & Ebel, J. P. (1981) *Nucleic Acids Res.* **9**, 6935-6952
- Verschoor, A., Frank, J., Radermacher, M., Wagenknecht, T. & Boublik, M. (1984) *J. Mol. Biol.* **178**, 677-698
- Vester, B. & Garrett, R. A. (1984) *J. Mol. Biol.* **179**, 431-452
- Wagner, R. & Garrett, R. A. (1978) *Nucleic Acids Res.* **5**, 4065-4075

- Wagner, R., Gassen, H. G., Ehresmann, C., Stiegler, P. & Ebel, J. P. (1976) *FEBS Lett.* **67**, 312–315
- Walker, T. A., Endo, Y., Wheat, W. H., Wool, I. G. & Pace, N. R. (1983) *J. Biol. Chem.* **258**, 333–338
- Ware, V. C., Tague, B. W., Clark, C. G., Gourse, R. L., Brand, R. C. & Gerbi, S. A. (1983) *Nucleic Acids Res.* **11**, 7795–7817
- Wickstrom, E. (1983) *Nucleic Acids Res.* **11**, 2035–2052
- Winkelman, D. A. & Kahan, L. (1983) *J. Mol. Biol.* **165**, 357–374
- Wittmann, H. G. (1982) *Annu. Rev. Biochem.* **51**, 155–183
- Wittmann, H. G. (1983) *Annu. Rev. Biochem.* **52**, 35–65
- Woese, C. R., Gutell, R., Gupta, R. & Noller, H. F. (1983) *Microbiol. Rev.* **47**, 621–669
- Wollenzien, P., Hearst, J. E., Thammana, P. & Cantor, C. R. (1979) *J. Mol. Biol.* **135**, 255–269
- Wower, I. & Brimacombe, R. (1983) *Nucleic Acids Res.* **11**, 1419–1437
- Wower, I., Wower, J., Meinke, M. & Brimacombe, R. (1981) *Nucleic Acids Res.* **9**, 4285–4302
- Yonath, A., Bartunik, H. D., Bartels, K. S. & Wittmann, H. G. (1984) *J. Mol. Biol.* **177**, 201–206
- Zagorska, L., Szkopinska, A., Klita, S. & Szafranski, P. (1980) *Biochem. Biophys. Res. Commun.* **95**, 1152–1159
- Zagorska, L., van Duin, J., Noller, H. F., Pace, B., Johnson, K. D. & Pace, N. R. (1984) *J. Biol. Chem.* **259**, 2798–2802
- Zimmermann, R. A. (1980) in *Ribosomes* (Chambliss, G., Craven, G. R., Davies, J., Davis, K., Kahan, L. & Nomura, M., eds.), pp. 135–169, University Park Press, Baltimore
- Zwieb, C. & Brimacombe, R. (1979) *Nucleic Acids Res.* **6**, 1775–1790
- Zwieb, C. & Brimacombe, R. (1980) *Nucleic Acids Res.* **8**, 2397–2411
- Zwieb, C. & Dahlberg, A. E. (1984a) *Nucleic Acids Res.* **12**, 4361–4375
- Zwieb, C. & Dahlberg, A. E. (1984b) *Nucleic Acids Res.* **12**, 7135–7152
- Zwieb, C., Ross, A., Rinke, J., Meinke, M. & Brimacombe, R. (1978) *Nucleic Acids Res.* **5**, 2705–2720
- Zwieb, C., Glotz, C. & Brimacombe, R. (1981) *Nucleic Acids Res.* **9**, 3621–3640