Nucleic Acids Research, 1998, Vol. 26, No. 4 887–895

The path of the growing peptide chain through the 23S rRNA in the 50S ribosomal subunit; a comparative cross-linking study with three different peptide families

Kyoung Moo Choi and Richard Brimacombe*

Max-Planck-Institut für Molekulare Genetik, AG-Ribosomen, Ihnestrasse 73, 14195 Berlin, Germany

Received December 1, 1997; Revised and Accepted January 7, 1998

ABSTRACT

As part of a programme to investigate the path of the nascent peptide through the large ribosomal subunit, peptides of different lengths (up to 30 amino acids), corresponding to the signal peptide sequence and N-terminal region of the *Escherichia coli* ompA protein, were synthesized *in situ* on *E.coli* ribosomes. The peptides each carried a diazirine moiety attached to their N-terminus which, after peptide synthesis, was photoactivated so as to induce cross-links to the 23S rRNA. The results showed that, with increasing length, the peptides became progressively cross-linked to sites in Domains V, II, III and I of the 23S rRNA, in a similar manner to that previously observed with a family of peptides derived from the tetracycline resistance gene. However, the cross-links to Domain III appeared at a shorter peptide length (12 aa) in the case of the ompA sequence, and an additional cross-link in Domain II (localized to nt 780–835) was also observed from this peptide. As with the tetracycline resistance sequence, peptides of all lengths were still able to form cross-links from their N-termini to the peptidyl transferase centre in Domain V. A further set of peptides (30 or 50 aa long), derived from mutants of the bacteriophage T4 gene 60 sequence, did not show the cross-links to Domain III, but their N-termini were nevertheless cross-linked to Domain I and to the sites in Domains II and V. The ability of relatively long peptides to fold back towards the peptidyl transferase centre thus appears to be a general phenomenon.

INTRODUCTION

Although the primary role of the ribosome is to translate mRNA sequences into peptides, a number of recent studies have in combination provided a clear demonstration that in addition the ribosome itself plays a decisive role in the folding of the newly-synthesized peptide chains into their active conformations as proteins. However, the details of this process, as well as the related question of how the nascent peptide chain threads its way through the large ribosomal subunit, are still far from clear. With regard to the folding process, it has been shown that *Escherichia coli* ribosomes, their 50S subunits or even native 23S rRNA can promote refolding of denatured proteins (1,2). If the 23S rRNA was denatured by heat treatment (2) or EDTA (1), then stimulation of the refolding process was lost. In another study (3), the refolding was inhibited by addition of specific oligonucleotides complementary to sequences within the peptidyl transferase area of Domain V of the 23S rRNA. The stimulatory effect of ribosomes on the refolding process could be enhanced (2) by addition of the chaperonins GroEL and GroES, but in this context it has recently been calculated that under normal stress-free conditions these chaperonins are only present in the cell in sufficient amounts to facilitate the folding of a very small proportion (<5%) of the total cellular protein (4). Furthermore, it has been shown (5) that the refolding of 'artificially' denatured proteins takes place considerably more slowly than the *in situ* co-translational folding process. The co-translational folding can lead to the acquisition of full enzymatic activity in the newly-synthesized protein while it is still attached to the ribosome, provided that the protein sequence is artificially extended at its C-terminus so as to enable the complete natural protein sequence to grow clear of the 50S subunit; this has been demonstrated in the cases of rhodanese (6) and firefly luciferase (7) carrying C-terminal extensions of at least 23 and 26 aa, respectively.

The lengths of these C-terminal extensions correlate with older studies (8,9) concerning the question of the path of the nascent peptide, which indicated that 30–40 aa of the latter are shielded from attack by proteolytic enzymes and are thus 'buried' within the ribosome. The shielded amino acids are presumed to lie on a pathway connecting the peptidyl transferase centre with the protein exit site on the 50S subunit; the former is located at the subunit interface (10), whereas the latter is on the solvent side of the 50S subunit (11). There is still considerable controversy concerning the question as to whether the nascent peptide path is better described as a 'channel' (12) or a 'tunnel' (13–15), and in fact recent electron microscopic reconstructions of the *E.coli* ribosome (16,17) indicate that the large subunit has a porous structure with several channels or tunnels which could be possible candidates for this function. From the foregoing discussion it is clear that the channel or tunnel must be wide enough to accommodate an at least partially folded protein domain.

In order to gain some insight into the co-translational folding process and to identify points on the 23S rRNA that are located along the nascent peptide pathway, we recently undertook a systematic cross-linking study (18,19), using peptides synthesized *in situ* on the ribosome. Each peptide carried a photoreactive affinity group attached either to the α-amino group of its N-terminal methionine or the ε-amino group of an adjacent lysine

*To whom correspondence should be addressed. Tel: +49 30 8413 1592; Fax: +49 30 8413 1690

residue, prepared by derivatizing the methionine or lysine with the N-hydroxysuccinimide ester of 4-(3-trifluoromethyl diazirino)-benzoic acid ('TDB'; 20,21). Peptides of up to 33 aa in length were analysed, and the results (18,19) showed that the N-terminus of the nascent peptide makes a series of dramatic leaps within the secondary structure of the 23S rRNA, first contacting the peptidyl transferase centre in Domain V of the latter and then, with increasing peptide lengths, successive sites in Domains IV, II, III and finally I. A further observation was that the N-termini of even the longest peptides were still able to 'reach back' and become cross-linked to the sites in the peptidyl transferase centre.

All of the longer peptides studied (19) had amino acid sequences corresponding to the N-terminal region of the tetracycline resistance gene (22,23). This raises the obvious question as to whether the results obtained were of general significance, or whether they were specific for the particular family of peptide sequences used. Here we report a similar cross-linking study, made with two completely unrelated families of peptides. The first set of sequences correspond to the N-terminal region of the pro-ompA protein (24), which is a secretory precursor of ompA carrying a signal peptide sequence. The latter is expected to adopt an α-helical configuration (25), and the N-terminal ompA peptides are thus very different from the tetracycline resistance peptides (19) both in terms of composition, secondary structure and biological function. The second family of sequences are related to the bacteriophage T4 gene 60 (26,27), whose mRNA has the unusual property of containing a 50 nt long untranslated or 'bypassed' region. The efficiency of the bypassing process is known to be closely linked to the amino acid sequence of the nascent peptide chain already synthesized (26). In comparison with the previous data (18,19), the cross-links observed with these two families of peptides show both striking similarities and significant differences.

MATERIALS AND METHODS

Materials

Synthetic oligodeoxynucleotides were purchased from T.I.B. (Berlin), *E.coli* $tRNA_f^{Met}$ from Sigma and bulk *E.coli* tRNA from Boehringer (Mannheim). ^{35}S- or ^{3}H-labelled methionine, and ^{35}S-labelled cysteine, were obtained from Amersham-Buechler (Braunschweig). The N-hydroxysuccinimide ester of the cross-linking reagent TDB was a generous gift from Dr Dmitry Bochkariov (University of California), and the constructs carrying sequences derived from the T4 gene 60 sequence in plasmid pBR322 were kindly provided by Dr John Atkins (University of Utah). Initiation factors were a kind gift from Dr Claudio Gualerzi (University of Camerino).

Preparation of mRNA

The sequences of the mRNA constructs and/or those of the corresponding *in situ* synthesized peptides used in this study are listed in Figure 1 (see Results). In the case of the ompA peptides (Fig. 1A), the mRNAs were transcribed as in (18,19) from appropriate synthetic oligodeoxynucleotides complementary to the mRNA sequence concerned and carrying the T7 promoter sequence at their 3′-ends; T7 transcription was carried out by the method of Lowary *et al.*, (28) in the presence of the short (18 base) oligodeoxynucleotide template complementary to the T7 promoter sequence. In the case of the longer gene 60 peptides (Fig. 1B), the DNA sequences for transcription were prepared by PCR from the appropriate plasmid constructs, using standard procedures (29). The PCR-amplified DNA samples were used directly for T7 transcription, the oligodeoxynucleotides for the PCR reactions being listed in Table 1. [In the case of the 30–31 aa peptide (Fig. 1B), the plasmid used for the PCR reaction was in fact construct 'bb17' (26), for reasons which are not relevant to this study; since the nucleotide sequence in the transcribed region here is identical to that of the construct BA3 (26) used to prepare the corresponding 49–50 aa peptide (Fig. 1B), we refer to this peptide for simplicity as 'BA3, 30–31 aa']. In control experiments with the gene 60 BA3 49–50 aa peptide, the penultimate cysteine codon was replaced by the codon for aspartic acid (see Table 1 and Results). In all cases (with the ompA or gene 60 peptides), the transcribed mRNA was purified by gel electrophoresis as previously described (20).

Biosynthesis of peptides

Initiator $tRNA_f^{Met}$ or bulk tRNA were charged with the appropriate amino acids using tRNA-free S-150 enzymes (31). The initiator Met-$tRNA_f^{Met}$ was labelled with [^{3}H]methionine to a specific activity of ~250 c.p.m./pmol, and was derivatized at the α-NH_2 group with the N-hydroxysuccinimide ester of the diazirine reagent TDB as before (18,19). The amino acid mixture used for charging the bulk elongator tRNAs was supplemented with [^{35}S]methionine or [^{35}S]cysteine (~6×10^5 c.p.m./pmol) as appropriate, for labelling of the penultimate residue in each peptide (cf. Fig. 1). Reaction mixtures for peptide synthesis were prepared as before (18,31), using 70S tight couple ribosomes from *E.coli* together with the appropriate T7-transcribed mRNA (see above). After the *in situ* peptide synthesis reaction, the diazirine moiety was activated by irradiation at 350 nm, as previously described (21). In control experiments, the irradiation was carried out prior to peptide synthesis.

Isolation of cross-linked products and analysis of sites of cross-linking to 23S rRNA

23S rRNA cross-linked to radioactive peptidyl-tRNA was isolated by a series of three sucrose gradients, as previously described (30); the first gradient (at high magnesium concentration) separates the 70S ribosomal complexes from unbound tRNA etc., the second (at low magnesium concentration) separates the 50S and 30S subunits and the third (in the presence of SDS) separates 23S rRNA (containing cross-linked peptidyl-tRNA) from ribosomal proteins (or proteins cross-linked to peptidyl-tRNA). Localization of the cross-link sites in the 23S rRNA was performed by ribonuclease H digestion in the presence of complementary oligodeoxy-nucleotides and by primer extension analysis, using our standard procedures (32,33), as in (18,19).

RESULTS

As in our previous experiments (19), the peptides to be studied (Fig. 1) were synthesized *in situ* on *E.coli* 70S ribosomes, and were encoded by suitable mRNA constructs prepared by T7 transcription (see Materials and Methods). In the case of the ompA family of sequences, four peptides were investigated with lengths from 6–7 up to 30–31 aa. The peptides thus cover the whole of the signal sequence, with the two longer peptides extending into the sequence of the mature ompA protein (Fig. 1A). In the gene 60 family (Fig. 1B), two peptide sequences

related to the wild-type construct BA3 (26) were studied, with lengths of 30–31 and 49–50 aa, respectively. The shorter of these two sequences was chosen so as to be able to make a direct comparison with the corresponding 30–31 aa peptide from the ompA family, whereas the longer BA3 peptide extends to the beginning of the coding gap sequence (26), which is 50 nt long and starts at the 51st codon. The construct BA3 shows a 98% 'bypass' of the coding gap (26), and for comparison we took a similar peptide from construct BH3, where the bypass efficiency is drastically reduced to 0.7% (26); the crucial sequences in the central region of the peptide chains which control the bypass event (26) are underlined in Figure 1B. These peptides (Fig. 1) are thus very different in properties, both in relation to each other and to the series of tetracycline resistance peptides previously studied (19). Accordingly, if significant variations in the path of the nascent peptide through the 50S subunit do exist, one might expect them to be revealed in the corresponding patterns of cross-linking to the 23S rRNA.

In every peptide (Fig. 1), the photoreactive diazirine moiety (TDB) was attached to the α-amino group of the N-terminal methionine. Our previous experiments (18,19) had shown that the cross-linking patterns were essentially the same, regardless as to whether this photo-label was on the N-terminal methionine or on the ε-amino group of an immediately adjacent lysine residue. The N-terminal methionine residue was, in addition, radioactively labelled at low specific activity with tritium, so as to be able to monitor the charging of the initiator $tRNA_f^{Met}$ and the efficiency of the derivatization with TDB. In order to visualize the peptides after the cross-linking reactions, a high specific activity ^{35}S-label in either cysteine or methionine was introduced at the penultimate position of each peptide (Fig. 1). The penultimate position was chosen so as to safeguard against possible 'ragged ends' in the T7 transcripts, which could result in a failure of the amino acid at the extreme C-terminus of the peptide to be encoded. (The lengths of the peptides are given as '6–7 aa', etc., accordingly). The insertion of a ^{35}S-label at this position in the peptide sequences ensures that only the fully-translated peptides are considered in the subsequent cross-linking analyses; shorter peptide products do not contain the label and are thus simply not visible. It should however be noted that the gene 60 peptides contain internal methionine residues (Fig. 1B) which cannot be formally excluded as possible alternative initiation sites, and the BH3 peptide carries an internal cysteine residue which also becomes radioactively labelled.

After peptide synthesis and photoactivation of the TDB, the peptides cross-linked to rRNA were isolated by a series of three sucrose gradients, as summarized in Materials and Methods. Analysis of the ^{35}S-radioactivity associated with the 70S ribosomes in the first of these gradients (at high magnesium concentration) indicated that ~25% of the ribosomes had participated in synthesis of the full-length peptides (cf. 19). The analysis of the second sucrose gradients (at low magnesium concentration) already revealed an important difference between the tetracycline resistance peptide family (19) and the ompA and gene 60 families. Namely, the longer ompA peptides and all of the gene 60 peptides (Fig. 1) showed significant levels of cross-linking to the 30S subunit. For reasons of space, the analysis of this 30S cross-linking will be presented elsewhere (K.M.Choi, J.F.Atkins, R.Gesteland and R.Brimacombe, submitted for publication). Here we are concerned only with a comparison of the cross-links to 23S rRNA within the 50S subunit between the three peptide families and the sucrose gradients showed that [as before (19)] ~2–4% of the 70S ribosomes carried a cross-linked peptidyl-tRNA, a value which is typical for the diazirine reagents (cf. 19,33). The cross-linked radioactivity in the 50S subunit was approximately equally divided between the 23S rRNA and protein moieties, as evidenced by the third (SDS-containing) sucrose gradients (data not shown). Control experiments in which the TDB-derivatized methionyl-tRNA had been pre-irradiated before being added to the protein-synthesizing reaction mixtures showed only insignificant amounts of radioactivity in the 23S rRNA fractions, indicating that the cross-linking reaction is dependent on the TDB-derivative. A further control was made in the case of the BA3 49–50 aa peptide (Fig. 1) by substituting the cysteine codon in the mRNA at the 49th position by the natural aspartic acid codon (26). Here again, no significant radioactivity was associated with the 23S rRNA fractions, demonstrating that the incorporation of ^{35}S-label was dependent on the presence of cysteine at its expected position in the peptide sequence.

In order to localize the sites of cross-linking in the rRNA, the cross-linked 23S rRNA–peptide complexes were first 'scanned' with ribonuclease H in the presence of a standard set of pairs of oligodeoxynucleotides complementary to sites spread along the whole length of the 23S rRNA (cf. 19; see Fig. 2). Regions of the 23S rRNA, shown in this way to contain cross-link sites, were subjected to further ribonuclease H digestions in the presence of different pairs of oligodeoxynucleotides, so as to narrow down the cross-linked regions as far as possible (Fig. 3). As before (19), final determinations of the precise cross-links sites within the rRNA regions identified by ribonuclease H digestion could be made by the primer extension method (34).

Table 1. Sequences of oligodeoxynucleotides for PCR amplification of gene 60 sequences

DNA template to be amplified[a]	Oligodeoxynucleotide sequence (5′–3′)
BA3, 30–31 aa; BA3 49–50 aa (3′-side)	TAATACGACTCACTATAGGGAGAAAAA**ATG**AAAAGCTTAATGAAATTT[b]
BH3, 46–47 aa (3′-side)	TAATACGACTCACTATAGGGAGAAAAA**ATG**AAAAGCTTTGTAAAAATT[b]
BA3, 49–50 aa; BH3 46–47 aa (5′-side)	TCC**ACA**GTGATCTGCGTCTGTCATAAT[c]
BA3, 49–50 aa (5′-side, control)	TCC**ATC**GTGATCTGCGTCTGT[d]
BA3, 30–31 aa (5′-side)	TCC**ACA**AGAACGACGAACATTGTTCTG[c]

[a]Cf. Figure 1B.
[b]The T7 promoter sequence is underlined, the ATG start codon in heavy type.
[c]The ACA anticodon for cysteine is in heavy type.
[d]The ATC anticodon for aspartic acid is in heavy type. See text for explanation.

Since the object of this study is to compare the cross-links to 23S rRNA from the ompA and gene 60 peptide families with those from the tetracycline resistance family, we give a short synopsis of the latter data at this point (Table 2), before going on to describe the new data with the ompA and gene 60 families. Starting from the 5′-end of the 23S rRNA, the cross-link to nt 91 (Table 2) was only observed with the longest (29–33 aa) tetracycline peptide. The second cross-link site (at nt 750) first made its appearance with the 6 aa peptide, and thereafter was observed with all the longer peptides. The neighbouring cross-link (No. 3) in the 780–835 region is specific for the ompA family (see below) and was not found with any of the tetracycline peptides; the region is however included in Table 2 for easier comparison with the subsequent Tables. Cross-links Nos 4 and 5 were both observed in the tetracycline series with the two longest peptides (25–26 and 29–33 aa) only. Cross-link 5 was to nt 1614 of the 23S rRNA, whereas cross-link 4 could only be localized to the 1305–1350 region. The next cross-link (to nt 1781) was only seen with the 4 aa peptide (Table 2). In contrast, cross-link No. 7 (to nt 2062) was found in every peptide examined. Cross-links 8 and 9 (to nt 2506 and 2585, respectively) were observed somewhat erratically in the different peptides, and the final cross-link (to nt 2609) was, like the cross-link to nt 2062, found with all of the peptides, with the exception in this case of the '1 aa' peptide, which represents simply the cross-linking from TDB-methionyl-tRNA itself.

Cross-links to 23S rRNA from the ompA peptides

Typical ribonuclease H 'scans' of the cross-links from the four ompA peptides (Fig. 1A) is illustrated in Figure 2A–D and the final locations of the cross-link sites are summarized in Table 3. The shortest ompA peptide studied (6–7 aa) shows the appearance of cross-linked products in lanes 3, 8, 10 and 11 of Figure 2A. As before (18,19), the cross-linked complexes released by ribonuclease H digestion appear as doublet bands on the gels; in each doublet the slower-moving band represents the peptidyl-tRNA cross-linked to the 23S rRNA fragment, whereas in the faster-moving band the tRNA moiety has been lost due to the relatively harsh conditions (55°C) of the ribonuclease H incubation. The band in lane 3 (Fig. 2A) indicates the presence of a cross-link between approximately nt 531 and 779 in the 23S rRNA (see legend to Fig. 2); that in lane 8 a cross-link between nt 1966 and 2235 and that in lane 11 between nt 2511 and 2623. The 395 nt band in lane 10 represents the same cross-link as that seen in lane 11, now appearing in the 23S rRNA region between nt 2511 and the 3′-terminus of the molecule (at nt 2904). Further ribonuclease H digests (not shown, but cf. Fig. 3) enabled these cross-link sites

A. OmpA Sequences:

1. Pro-ompA: M K K T A I A I A V A L A G F A T V A Q A A P K D N T W Y T

2. 6 - 7 aa: GGGAGAAAAAAUGAAAAAGACAGCUAUGGCG
dM K K T A M* A

3. 12-13 aa: GGGAGAAAAAAUGAAAAAGACAGCUAUCGCGAUUGCAGUGGCAAUGGCU
dM K K T A I A I A V A M* A

4. 24-25 aa: GGGAGAAAAAAUGAAAAAGACAGCUAUCGCGAUUGCAGUGGCACUGGCUGGUUUCGCUACCGUAGCGCAGGCCGCUCCGAUGGAU
dM K K T A I A I A V A L A G F A T V A Q A A P M* D

5. 30-31 aa: GGGAGAAAAAAUGAAAAAGACAGCUAUCGCGAUUGCAGUGGCACUGGCUGGUUUCGCUACCGUAGCGCAGGCCGCUCCGAAAGAUAACACCUGGUACAUGACU
dM K K T A I A I A V A L A G F A T V A Q A A P K D N T W Y M* T

B. Gene 60 Sequences:

1. BA3, 30-31 aa: dM K S L M K F V K I D S S S V D M K K Y K L Q N N V R R S Ċ* G

2. BA3, 49-50 aa: dM K S L M K F V K I D S S S V D M K K Y K L Q N N V R R S I K S S S K I Y A N V A I M T D A D H Ċ* G
(D)

3. BH3, 46-47 aa: dM K S F V K I D S S S V D M K N I N Ċ* R T M F V V L L N P L Q K I Y A N V A I M T D A D H Ċ* G
(D)

Figure 1. mRNA and peptide sequences used in this study. (**A**) Sequences related to the pro-ompA protein (24). The first line gives the pro-ompA peptide sequence, with the N-terminal sequence of the mature protein underlined. The subsequent lines show in each case the transcribed mRNA sequence (in small type), with the corresponding peptide sequence underneath. 'aa' stands for amino acids, 'dM' denotes diazirine-modified methionine and 'M*', ^{35}S-labelled methionine. (**B**) Sequences related to Gene 60. Each line gives the translated peptide sequence, together with the name of the mutant (BA3, BH3) from which the sequence is derived (26). The altered sequence in the centre of the two longer peptides is underlined. 'dM' denotes modified methionine as in (A), and 'C', with an asterisk above, denotes ^{35}S-labelled cysteine, which was substituted for the aspartic acid residues ('D') at the C-terminus of the longer peptides. See text for further explanation.

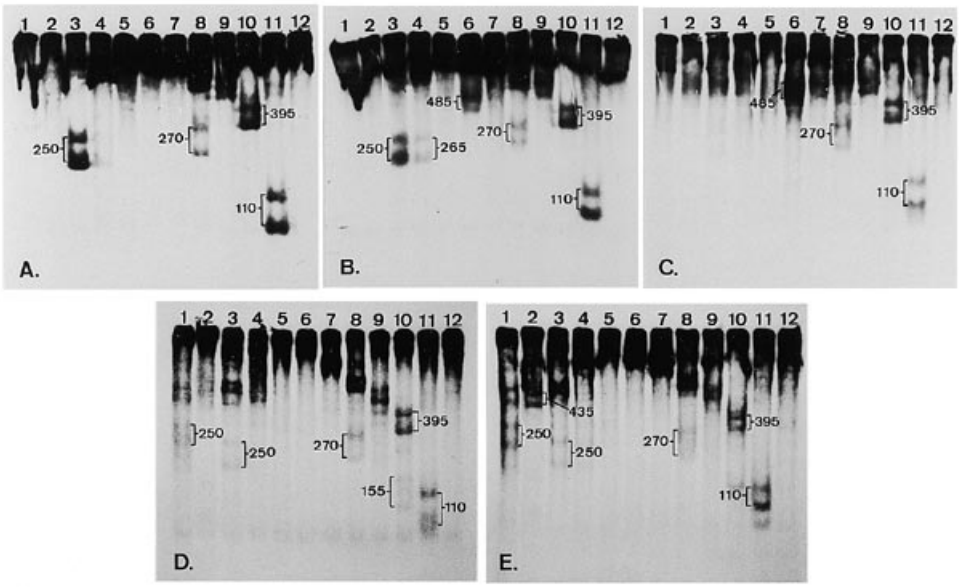

Figure 2. Autoradiograms of ribonuclease H digests of ^{35}S-labelled peptidyl-tRNA–23S rRNA cross-linked complexes on 5% polyacrylamide gels. In each gel the digests were performed in the presence of pairs of oligodeoxynucleotides centred on the following positions in the 23S rRNA sequence (the number in parentheses gives in each case the approximate size of the ribonuclease H fragment expected to be released between the two oligodeoxynucleotides): lane 1, 249/437 (190 nt); lane 2, 437/531 (95 nt); lane 3, 531/779 (250 nt); lane 4, 779/1044 (265 nt); lane 5, 1044/1258 (215 nt); lane 6, 1258/1741 (485 nt); lane 7, 1741/1966 (225 nt); lane 8, 1966/2235 (270 nt); lane 9, 2235/2358 (125 nt); lane 10, 2358/2511 (155 nt); lane 11, 2511/2623 (110 nt); lane 12, 2623/2822 (200 nt). Radioactive doublet bands (see text) corresponding to these fragments are marked accordingly. (**A**) Cross-links from ompA 6–7 aa (Fig. 1A); (**B**) ompA 12–13 aa; (**C**) ompA 24–25 aa; (**D**) ompA 30–31 aa; (**E**) cross-links from BA3 49–50 aa (Fig. 1B).

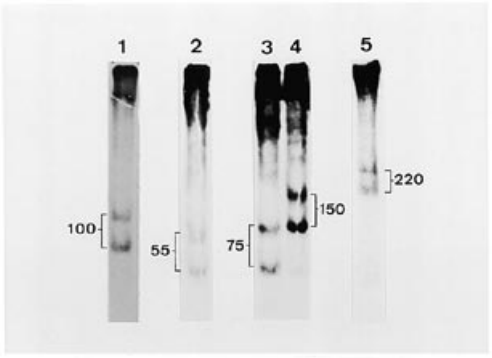

Figure 3. Examples of autoradiograms of further ribonuclease H digestions to narrow down the cross-linked 23S rRNA regions, on 5% gels as in Figure 2. The sizes of the radioactive doublet bands released are marked as in the latter Figure. Lane 1: the cross-linked ompA 30–31 aa peptide, digested with oligodeoxynucleotides centred on 23S rRNA positions 101/268 (releasing an internal fragment of ~165 nt). Lane 2: the ompA 12–13 aa peptide, with oligodeoxynucleotides at positions 779/833 (55 nt). Lane 3: ompA 12–13 aa, positions 1258/1335 (75 nt). Lane 4: ompA 12–13 aa, positions 1593/1741 (150 nt). Lane 5: ompA 30–31 aa, digested with three oligodeoxynucleotides at positions 2510/2602/2822 (90 or 220 nt, respectively). See text for further explanation.

to be narrowed down to the same regions (rounded off to the nearest 5 nt) as those found with the tetracycline resistance peptides, namely to nt 730–760 (the cross-link from lane 3), nt 2055–2090 (that from lane 8) and to two cross-links (from lane 11) between nt 2570–2605 and 2605–2625, respectively (Table 3). Furthermore, the subsequent primer extension analyses (not shown, but cf. 19) demonstrated that the cross-link sites within these regions were identical to those previously found (nt 750, 2062, 2585 and 2609; Tables 2 and 3). Thus, with the exception of the cross-link to nt 2506 (Table 2), the cross-links from the 6–7 aa ompA peptide are exactly the same as those from the 6 aa tetracycline peptide (compare Tables 2 and 3).

In contrast, significant differences are apparent between the respective cross-linking patterns for the 13–15 aa tetracycline peptide (Table 2) and the 12–13 aa ompA peptide (Table 3, Fig. 2B). It can be seen from Figure 2B that the cross-link bands in lanes 3, 8, 10 and 11 are the same as those in Figure 2A, and the further analyses confirmed that the cross-link sites were indeed the same. However, bands are also visible in lanes 4 and 6 of Figure 2B, corresponding to cross-link sites between nt 779 and 1044 (lane 4) and 1258 and 1741 (lane 6). The further ribonuclease H digestions enabled the first of these sites to be narrowed down to the region between approximately nt 780–835 (Fig. 3, lane 2), whereas the second resolved into two sites between nt 1305–1350 and 1590–1625, respectively; Figure 3 (lanes 3 and 4) shows typical digestions that were part of this ‘narrowing down’ process, where the cross-linked region between nt 1258 and 1741 has been ‘split’ into two sites between approximately nt 1258–1336 and 1593–1741. In the subsequent primer extension analyses, a reverse transcriptase stop signal was only observed in one of the three cross-linked regions, namely at nt 1614 (cross-link No. 5, Table 3). For cross-links Nos 3 and 4, no stop signal could be found, despite repeated attempts. This reproduces our previous finding with the two longest tetracycline peptides (19) in the case of the 1305–1350 region; the limitations of the primer extension method for the analysis of cross-link sites have been discussed at length elsewhere (19,35).

Table 2. Summary of 23S rRNA cross-links from the tetracycline resistance gene peptides

No.	Localization by RNAse H	Primer extension site	Peptide length 1	4	6	9	13–15	25–26	29–33
1	60–100	91	–	–	–	–	–	–	+
2	730–760	750	–	–	+	+	+	+	+
3	(780–835)	(nd)	–	–	–	–	–	–	–
4	1305–1350	nd	–	–	–	–	–	+	+
5	1590–1625	1614	–	–	–	–	–	+	+
6	1770–1810	1781	–	+	–	–	–	–	–
7	2055–2090	2062	+	+	+	+	+	+	+
8	2460–2510	2506	+	(+)	(+)	–	(+)	–	(+)
9	2570–2605	2585	+	(+)	(+)	(+)	+	+	+
10	2605–2625	2609	–	+	+	+	+	+	+

The Table is (in modified format) from (19). '+' indicates that a cross-link was reproducibly present, '(+)' that it was only weakly or occasionally present and '–' that the cross-link was reproducibly absent. The RNAse H column shows the shortest RNA regions found to encompass the cross-link site by ribonuclease H digestion, and the primer extension column indicates the precise site of cross-linking; 'nd' means that no primer extension stop signal was observed within the region identified by ribonuclease H digestion. Cross-link No. 3 (in parentheses) was not found with this family of peptides.

Table 3. Summary of 23S rRNA cross-links from the ompA gene peptides

No.	Localization by RNAse H	Primer extension site	Peptide length 6–7	12–13	24–25	30–31
1	60–100	91	–	–	–	+
2	730–760	750	+	+	–	(+)
3	780–835	nd	–	+	–	(+)
4	1305–1350	nd	–	+	+	–
5	1590–1625	1614	–	+	+	–
6	(1770–1810)	(1781)	–	–	–	–
7	2055–2090	2062	+	+	+	+
8	2460–2510	2506	–	–	–	(+)
9	2570–2605	2585	+	+	–	–
10	2605–2625	2609	+	+	+	+

See legend to Table 2 for explanation. Cross-link No. 6 (in parentheses) was not found with this family of peptides. The peptide sequences are listed in Figure 1A.

With the ompA 24–25 aa peptide (Fig. 2C), the pattern is again different, in that the cross-links in lanes 3 and 4 have now disappeared. Furthermore, ribonuclease H digestions showed that the cross-link in lane 11 only corresponded to the cross-link site at nt 2609, with the site at nt 2585 (see Table 3) now being absent (data not shown, but cf. Fig. 3, lane 5). The cross-link at nt 2062 (corresponding to the band in lane 8) and those in the 1305–1350 and 1590–1625 regions (corresponding to the band in lane 6) were however all still present. The cross-linking patterns from the longest ompA peptide (30–31 aa; Fig. 2D) tended to be less well-defined, but nevertheless showed distinctive features. The most important of these is the reproducible appearance of a cross-link band in lane 1 of the gel, where the ribonuclease H cuts were at approximately nt 249 and 437. The mobility of the cross-linked band is somewhat anomalous here (as noted in reference 19 for the corresponding 29–33 aa tetracycline peptide), but suggests a cross-link site within the 5′ 250 nt of the 23S rRNA rather than between positions 249 and 437. The slower-moving band in this gel lane of Figure 2D (~435 nt) arises from an incomplete ribonuclease H scission at position 249. Further ribonuclease H digestions confirmed the location of the cross-link site, as exemplified by the digest of Figure 3 (lane 1), where the oligodeoxynucleotides gave cuts at approximately nt 101 and 268; the observed radioactive band (~100 nt) corresponds to a cross-link site within the 5′ 100 nt region, which subsequently proved to be identical to the site at nt 91 (cross-link No. 1, Table 3) found previously with the longest tetracycline peptide (19; Table 2). Faint bands were sometimes seen with the 30–31 aa peptide in gel lanes 3 and 4, the former but not the latter being visible in the example shown in Figure 2D, corresponding to cross-links 2 and 3 of Table 3, respectively. However, cross-links 4 and 5, which should appear in lane 6, were now reproducibly absent. Cross-link 7 (lane 8) was always present, and a site corresponding to cross-link 8 (narrowed down to nt 2460–2510) was sometimes seen in lane 10 (the 155 nt band in Fig. 2D). Finally, as with the 24–25 aa peptide, the cross-link in lane 11 corresponded only to the site at nt 2609, but not that at nt 2585, as evidenced by the ribonuclease H digest in Figure 3 (lane 5); this digest shows a 220 nt band but not a 90 nt band (see legend to Fig. 3).

Cross-links to 23S rRNA from the gene 60 peptides

The data for the gene 60 peptides can be described very simply, because the cross-linking patterns for all three gene 60 peptides

(Fig. 1B) were essentially identical to that obtained with the 30–31 aa ompA peptide (Fig. 2D). The results are summarized in Table 4, and an example of the ribonuclease H scan gel is shown in Figure 2E, for the BA3 49–50 aa peptide. Cross-link No. 1 near to the 5′-end of the 23S rRNA is prominently visible in lane 1 of Figure 2E, and the same cross-link appears in lane 2 as the 435 nt band representing the region between the first ribonuclease H cut (at approximately nt 437) and the 5′-terminus. Cross-link No. 2 at nt 750 (seen as the 250 nt band in lane 3) was observed with varying intensities among the gene 60 peptides (cf. Table 4) and cross-link No. 3 (lane 4) was also sometimes present, although not in the example shown (Fig. 2E). The two cross-links (Nos. 4 and 5) in the 1305–1350 and 1590–1625 regions were reproducibly absent (lane 6) and No. 7 (at nt 2062) was weakly but reproducibly present (lane 8). As with the 30–31 aa ompA peptide, cross-link No. 8 (at nt 2506) was occasionally observed (lane 10; not visible in Fig. 2E), and the final cross-link in lane 11 was that to nt 2609 but not to nt 2585 (cf. Fig. 3, lane 5 and see Table 4). As already noted above, the gene 60 peptides contain internal methionine residues which could represent alternative initiation sites, with the result that the *in situ* synthesized peptides could in principle contain shorter products than those expected (cf. Fig. 1B). However, the cross-linking patterns obtained (e.g. the consistent appearance of the cross-link site at nt 91 of the 23S rRNA with all peptides of 30 aa or longer) indicate that it is the full-length peptides which play the most significant role in the results. Most important is the fact that no differences were found between the 46 and 47 aa peptide from construct BH3 [with low ribosomal bypass efficiency (26)] and the 49–50 aa peptide from BA3 (with high bypass efficiency).

DISCUSSION

Cross-linking studies of the type described here are time-consuming, and therefore, only a very limited number of different peptides can be investigated. The families that we have examined, namely peptides from the tetracycline resistance gene (19), the ompA protein gene and the bacteriophage T4 gene 60, were chosen so as to be widely different from each other in terms of sequence, secondary structure and biological properties. The data from all three peptide families (Table 2) are summarized in Figure 4, which shows how the nascent peptide moves through the various domains of the 23S rRNA (cf. 19).

The shortest possible ‘peptide’, namely the diazirine-derivatized methionyl-tRNA itself (Table 2) is cross-linked to nt 2062, 2506 and 2585 in the peptidyl transferase ring (36; Fig. 4). Cross-links from other photoreactive aminoacyl-tRNA derivatives to this area of the 23S rRNA have been reported, namely from a benzophenone-derivatized aminoacyl-tRNA to nts 2451, 2506 and 2585 (37) and from an azido-hippuric acid derivative to nt 2439 (32). Our additional cross-link to the peptidyl transferase ring at nt 2609 (Table 2) was observed with the peptides of 4 aa or longer, but no diazirine cross-linking was ever found to nt 2439 or 2451. This is more likely to reflect the chemistry and structural affinities of the various cross-linking reagents, rather than any special individual properties of the ribosomal complexes being cross-linked. We use the diazirine derivatives in our experiments because of their high reactivity, and because they are activated at 350 nm, a wavelength at which the ribosome is ‘transparent’.

The cross-link from the tetra-peptide in the tetracycline resistance family to nt 1781 (Table 2) was not re-examined in the current series of experiments, where the shortest sequence investigated was the 6–7 aa ompA peptide (Table 3). As previously noted (19), nt 1781 is in the immediate neighbourhood of nt 2609, as evidenced by the fact that precisely these 2 nt were found to be partners in a UV-induced intra-RNA cross-link within the 23S rRNA (35). A similar UV-induced cross-link between helices 35 and 73 is also well-established (35), nt 750 in helix 35 being the next point on the peptide path (Fig. 4). The cross-link to nt 750 appeared with the hexapeptides in both the ompA and tetracycline series (Tables 2 and 3), but, whereas this cross-link persisted with all the longer tetracycline peptides, it disappeared (or was very weak) with the ompA 24–25 and 30–31 aa peptides. The 12 aa ompA peptide showed the new cross-link site in the 780–835 area (Fig. 4), and the cross-links to the 1305–1350 area and to nt 1614 also appeared in the 12 aa ompA peptide (Table 3). An independent topographical connection between the two areas (750 and 780–835; 1305–1350 and 1614) is provided by the intra-RNA cross-link (35) between helices 33 and 54. In the tetracycline series the cross-links to nt 1305–1350 and 1614 were first seen in the 25–26 aa rather than the 13–15 aa peptide (Table 2), and whereas in the latter series these two cross-links were also present in the longest (29–33 aa) peptide, they had disappeared again in the corresponding (30–31 aa) ompA peptide (Table 3). They were also not seen in any of the gene 60 family of peptides (Table 4), although here the cross-links at nt 750 and 780–835 were occasionally weakly present.

Table 4. Summary of 23S rRNA cross-links from the gene 60 peptides

No.	Localization by RNAse H	Primer extension site	Peptide length 30–31 (BA3)	46–47 (BH3)	49–50 (BA3)
1	60–100	91	(+)	(+)	+
2	730–760	750	–	+	(+)
3	780–835	nd	(+)	(+)	(+)
4	(1305–1350)	(nd)	–	–	–
5	(1590–1625)	(1614)	–	–	–
6	(1770–1810)	(1781)	–	–	–
7	2055–2090	2062	+	(+)	(+)
8	2460–2510	2506	(+)	(+)	(+)
9	(2570–2605)	(2585)	–	–	–
10	2605–2625	2609	+	+	+

See legend to Table 2 for explanation. Cross-links Nos 4, 5, 6 and 9 (in parentheses) were not found with this family of peptides. The peptide sequences are listed in Figure 1B.

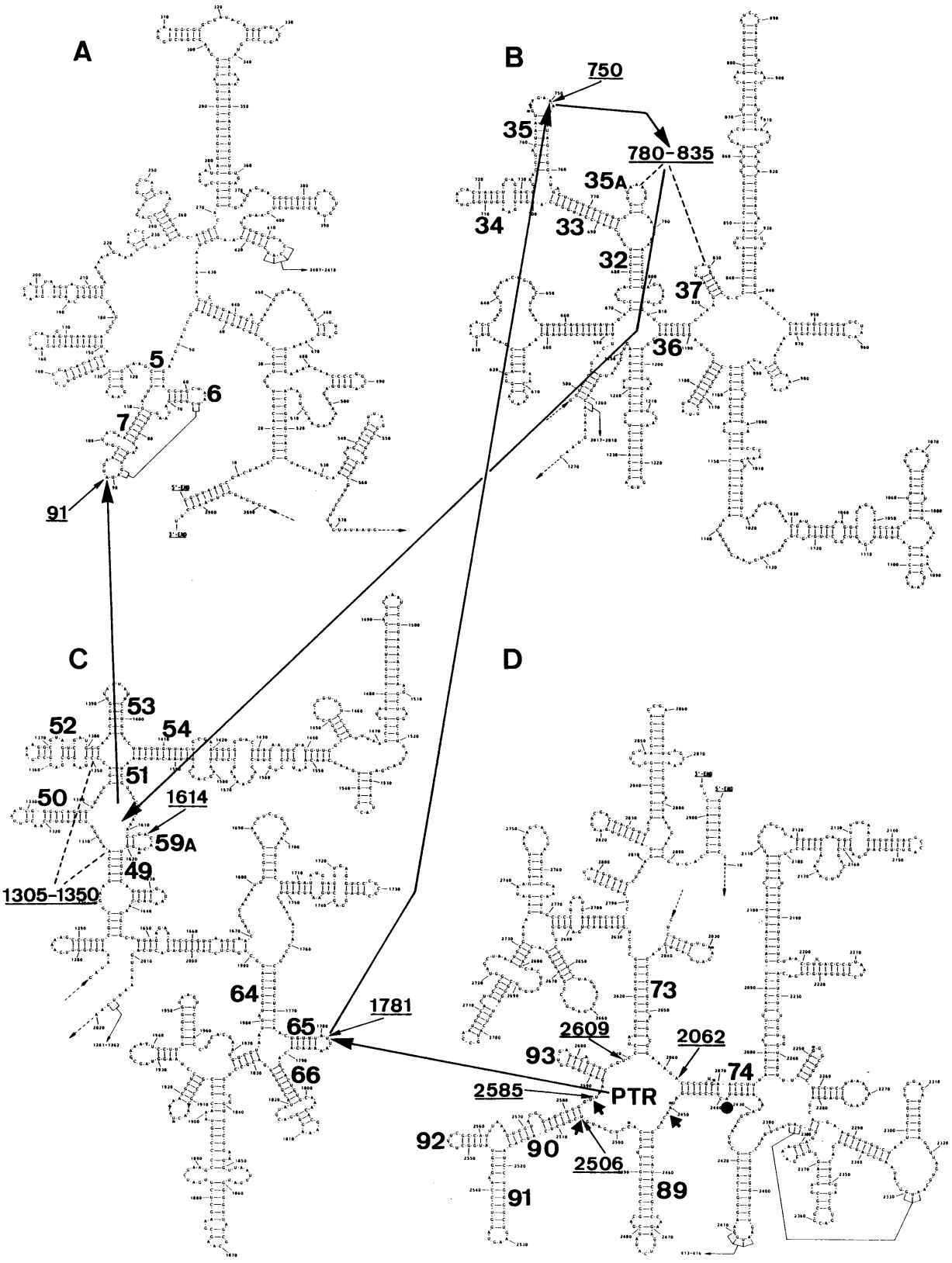

Figure 4. The path of the nascent peptide chain through the 23S rRNA (19). The complete secondary structure of the 23S rRNA is shown, with helices in the neighbourhood of the cross-link sites (cf. Tables 2–4) numbered as in (35). The cross-linked nucleotides (or regions) from the latter Tables are underlined, and the overall path of the peptide is indicated by the heavy arrowed lines. 'PTR' stands for peptidyl transferase ring (36) and the 'blunt' arrows in this area are cross-link sites from aminoacyl-tRNA (37). The black circle under helix 74 is also a cross-link site from aminoacyl-tRNA (32) (see text).

All of the peptides of 30 aa or longer exhibited the cross-link to nt 91 (Tables 2–4) at the 5′-end of the 23S rRNA. Again as previously noted (19), the nt 1614 and 1305–1350 cross-links are close to the binding site of protein L23 in helices 51–54 (38), whereas the nt 91 cross-link is close to an RNA–protein cross-link site for protein L23 in helix 6 (35), so that these two areas should be topographically close. Protein L29 similarly has a cross-link site in helix 7 (35), and both L23 and L29 have been mapped by immuno-electron microscopy (39) at locations on the solvent side of the 50S subunit near to the presumptive exit site of the nascent peptide chain (40). Notably, no further new cross-links were observed in the longer (46–47 and 49–50 aa) peptides in the gene 60 series (Table 4), and we draw the conclusion that the cross-link at nt 91 must be at or near the peptide exit site on the 50S subunit. It is also worth noting that nt 91 lies within the sequence corresponding to 5.8S rRNA in eukaryotes, thus raising the possibility that the 5.8S molecule could form part of the exit route for the nascent peptide.

It is intriguing that the most persistent cross-links in all the peptide families (Tables 2–4) are those to nt 91 on the one hand, and those to nt 2062 and 2609 in the peptidyl transferase ring (Fig. 4) on the other. The more variably-occurring cross-links (apart from those at nt 2506 and 2585) are at intermediate positions along the peptide path. For co-translational folding of the nascent peptide to take place within the ribosome (6,7), one would expect that the peptide channel or tunnel must be sufficiently wide to accommodate partially or completely folded protein domains, and that the protein chain must be highly flexible within this channel or tunnel. In particular, it is logical that even the longest peptides can still become cross-linked to sites in the peptidyl transferase ring, because the dynamically folding peptide needs to be able to ‘scan’ the latest adducts to the nascent chain. Interactions between the nascent peptide and the rRNA at the peptidyl transferase centre have already been proposed by other authors (41). Different peptide sequences will show different folding properties, accounting for the variability of the cross-linking patterns with respect to peptide length in the intermediate positions. On the other hand, the fact that the same sites on the 23S rRNA are involved in the cross-links with all the peptide families suggests that certain nucleotides are particularly exposed, or are ‘hot spots’ for the diazirine cross-linking reagent. At all events, our results clearly contradict the concept of a growing peptide that is rigidly constrained into a narrow channel or tunnel. Finally, the identical cross-linking patterns observed from the N-terminal of the two longer gene 60 peptides (BH3, 46–47 aa and BA3, 49–50 aa; Table 4) suggest that the level of coding-gap bypass efficiency [which is high for the construct BA3 and low for BH3 (26)] is not mediated by any gross variation in the path of the nascent peptide within the 50S subunit. As mentioned in the Results section, the gene 60 peptides, concomitantly showed significant levels of cross-linking to the 30S subunit (Choi, Atkins, Gesteland and Brimacombe, manuscript submitted); this is indicative of an even higher degree of flexibility, and a transient interaction of the nascent peptide with the 30S subunit could provide a clue to the mechanism of the coding gap bypass phenomenon.

ACKNOWLEDGEMENT

This work was financially supported by a grant from the Deutsche Forschungsgemeinschaft (Br 632/3-3).

REFERENCES

1 Das, B., Chattopadhyay, S., Bera, A. K. and Dasgupta, C. (1996) *Eur. J. Biochem.*, **235**, 613–621.
2 Kudlicki, W., Coffman, A., Kramer, G. and Hardesty, B. (1997) *Folding & Design*, **2**, 101–108.
3 Chattopadhyay, S., Das, B. and Dasgupta, C. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 8284–8287.
4 Lorrimer, G. H. (1996) *FASEB J.*, **10**, 5–9.
5 Fedorov, A. N. and Baldwin, T. O. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 1227–1231.
6 Kudlicki, W., Chirgwin, J., Kramer, G. and Hardesty, B. (1995) *Biochemistry*, **34**, 14284–14287.
7 Makeyev, E. V., Kolb, V. A. and Spirin, A. S. (1996) *FEBS Lett.*, **378**, 166–170.
8 Malkin, L. I. and Rich, A. (1967) *J. Mol. Biol.*, **26**, 329–346.
9 Blobel, G. and Sabatini, D. D. (1970) *J. Cell Biol.*, **45**, 130–145.
10 Stöffler, G. and Stöffler-Meilicke, M. (1986) In Hardesty, B. and Kramer, G. (eds) *Structure, Function and Genetics of Ribosomes*, Springer Verlag, New York, NY. pp. 28–46.
11 Bernabeu, C. and Lake, J. A. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 3111–3115.
12 Ryabova, L. A., Selivanova, O. M., Baranov, V. I., Vasiliev, V. D. and Spirin, A. S. (1988) *FEBS Lett.*, **226**, 255–260.
13 Milligan, R. A. and Unwin, P. N. T. (1986) *Nature*, **319**, 693–695.
14 Yonath, A., Leonard, K. R. and Wittmann, H. G. (1987) *Science*, **236**, 813–816.
15 Crowley, K. S., Reinhart, G. D. and Johnson, A. E. (1993) *Cell*, **73**, 1101–1115.
16 Stark, H., Mueller, F., Orlova, E. V., Schatz, M., Dube, P., Erdemir, T., Zemlin, F., Brimacombe, R. and van Heel, M. (1995) *Structure*, **3**, 815–821.
17 Stark, H., Orlova, E. V., Rinke-Appel, J., Jünke, N., Mueller, F., Rodnina, M., Wintermeyer, W., Brimacombe, R. and van Heel, M. (1997) *Cell*, **88**, 19–28.
18 Stade, K., Riens, S., Bochkariov, D. and Brimacombe, R. (1994) *Nucleic Acids Res.*, **22**, 1394–1399.
19 Stade, K., Jünke, N. and Brimacombe, R. (1995) *Nucleic Acids Res.*, **23**, 2371–2380.
20 Nassal, M. (1983) *Liebigs Ann. Chem.*, **1983**, 1510–1523.
21 Bochkariov, D. and Kogon, A. A. (1992) *Anal. Biochem.*, **204**, 90–95.
22 Sutcliffe, J. G. (1979) *Cold Spring Harbor Symp. Quant. Biol.*, **43**, 77–90.
23 Peden, K. W. (1983) *Gene*, **22**, 277–280.
24 Movva, N. R., Nakamura, K. and Inouye, M. (1980) *J. Biol. Chem.*, **255**, 27–29.
25 Emr, S. C. and Silhavy, T. J. (1982) *J. Cell Biol.*, **95**, 689–696.
26 Weiss, R. B., Huang, W. M. and Dunn, D. M. (1990) *Cell*, **62**, 117–126.
27 Gesteland, R. F., Weiss, R. B. and Atkins, J. F. (1992) *Science*, **257**, 1640–1641.
28 Milligan, J. F., Groebe, D. R., Witherell, G. W. and Uhlenbeck, O. C. (1987) *Nucleic Acids Res.*, **15**, 8783–8798.
29 McPherson, M. J., Quirke, P. and Taylor, G. R. (1991) In Rickwood, D. and Hames, B. D. (eds) *PCR; A Practical Approach.* Oxford University Press, pp. 1–253.
30 Stade, K., Rinke-Appel, J. and Brimacombe, R. (1989) *Nucleic Acids Res.*, **17**, 9889–9908.
31 Rheinberger, H. J., Geigenmüller, U., Wedde, M. and Nierhaus, K. H. (1988) *Methods Enzymol.*, **164**, 658–670.
32 Mitchell, P., Stade, K., Osswald, M. and Brimacombe, R. (1993) *Nucleic Acids Res.*, **21**, 887–896.
33 Döring, T., Mitchell, P., Osswald, M., Bochkariov, D. and Brimacombe, R. (1994) *EMBO J.*, **13**, 2677–2685.
34 Moazed, D., Stern, S. and Noller, H. F. (1986) *J. Mol. Biol.*, **187**, 399–416.
35 Brimacombe, R. (1995) *Eur. J. Biochem.*, **230**, 365–383.
36 Vester, B. and Garrett, R. A. (1988) *EMBO J.*, **7**, 3577–3587.
37 Steiner, G., Kuechler, E. and Barta, A. (1988) *EMBO J.*, **7**, 3949–3955.
38 Vester, B. and Garrett, R. A. (1984) *J. Mol. Biol.*, **179**, 431–452.
39 Stöffler, G. and Stöffler-Meilicke, M. (1986) In Hardesty, B. and Kramer, G. (eds) *Structure, Function and Genetics of Ribosomes.* Springer Verlag, New York, pp. 28–46.
40 Bernabeu, C. and Lake, J. A. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 3111–3115.
41 Harrod, R. and Lovett, P. S. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 8650–8654.