Review

Structural and functional topography of the human ribosome

Dmitri Graifer and Galina Karpova

Laboratory of Ribosome Structure and Functions, Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, Novosibirsk 630090, Russia

*Correspondence address. Tel: +7-383-363-5140; Fax: +7-383-363-5153; E-mail: karpova@niboch.nsc.ru

This review covers data on the structural organization of functional sites in the human ribosome, namely, the messenger RNA binding center, the binding site of the hepatitis C virus RNA internal ribosome entry site, and the peptidyl transferase center. The data summarized here have been obtained primarily by means of a site-directed cross-linking approach with application of the analogs of the respective ribosomal ligands bearing cross-linkers at the designed positions. These data are discussed taking into consideration available structural data on ribosomes from various kingdoms obtained with the use of cryo-electron microscopy, X-ray crystallography, and other approaches.

Keywords human ribosome; photoaffinity cross-linking; mRNA and tRNA binding site; hepatitis C IRES binding site; ribosomal component

Received: September 22, 2011 Accepted: November 6, 2011

Introduction

Ribosomes are cellular organelles that translate the genetic information incoming as messenger RNA (mRNA) into the polypeptide chains of proteins. This function is common to the ribosomes of all organisms, from eubacteria to humans. During translation, the ribosome interacts with two types of specific RNA ligands, the mRNA and transfer RNA (tRNA). Each species of tRNA carries a specific amino acid residue for the synthesis of a polypeptide chain. Insertion of the required amino acid residue into the growing polypeptide chain at the right site is achieved due to the recognition of an mRNA codon at the ribosomal aminoacyl (A) site by the anticodon of the cognate aminoacyl-tRNA via base pairing. This recognition leads to the binding of the cognate aminoacyl-tRNA at the A site. Then the ribosome catalyses formation of the peptide bond between the amino acid residue of the aminoacyl-tRNA at the A site and the polypeptide moiety of the peptidyl-tRNA bound at the neighboring peptidyl (P) site. Each tRNA is specific to only one amino acid; the correspondence between mRNA codons and amino acid residues is governed by a genetic code that is practically universal for cytoplasmic ribosomes from all kingdoms.

The ribosome is a very complicated molecular machine consisting of two subunits, small and large ones; each contains ribosomal RNA (rRNA) and several dozen ribosomal proteins. The structures of ribosomes from prokaryotes and eukaryotes share significant similarity. The secondary structures of rRNAs possess conserved cores common to all kingdoms; in eukaryotes rRNAs are longer than in prokaryotes and the conserved cores are interrupted in several sites by variable regions named expansion segments [1] (http://www.rna.ccbb.utexas.edu). Eukaryotic ribosomal subunits contain a complete set of ribosomal proteins homologous to their eubacterial counterparts plus a number of additional proteins (about one-third of the total number). The locations of the main functional sites on the subunits are in general similar for ribosomes from all kingdoms. The small subunits carry an mRNA binding center, including the decoding site where aminoacyl-tRNA cognate to the mRNA codon at the A site is selected, and the large subunits contain the peptidyl transferase center (PTC) responsible for catalysis of peptide bond formation. Three tRNA binding sites, namely, A, P, and E (E is exit site for discharged tRNA where it binds before leaving the ribosome) are formed by the interfacial areas of both subunits.

Eukaryotic ribosomes are able to translate not only cellular mRNAs but also viral RNAs, but the mechanisms of translation initiation of cellular and viral RNAs are generally different. Cellular mRNAs contain a special 7-methylguanosine cap structure at their 5'-termini that is recognized by initiation factor eIF4E, which, together with other initiation factors, promotes the binding of 40S subunits complexed with initiator Met-tRNA\textsubscript{Met} and eIFs (i.e. the 43S complex) to the 5'-terminus of mRNA. The subsequent movement of a 40S subunit along the mRNA (scanning) leads to positioning of the AUG codon at the P site, where it interacts with the initiator tRNA [2,3]. An
alternative mechanism is utilized by certain uncapped viral RNAs that contain in their 5′-untranslated region (UTR) highly structured regions, called IRESs (internal ribosome entry sites) that can bind to 40S subunits such that the initiation AUG codon is placed immediately at the P site region [4–6]. In particular, IRES-mediated translation initiation is inherent to the RNA of hepatitis C virus (HCV), one of the most dangerous human pathogens. One of the key steps of the viral life cycle in human host cells is the initiation of viral RNA translation, which implies that ribosomes are key players in progression of the viral infection in the cell.

Clearly, knowledge of the ribosome structure and an understanding of the molecular mechanisms of protein synthesis and its regulation are of basic importance for life sciences. The structure of the prokaryotic ribosome has been extensively studied by various biochemical approaches [7,8]. The accumulated data enabled the later X-ray crystallographic images to be interpreted to reveal the ribosome’s structure at the atomic level [9–12]. However, eukaryotic ribosomes, especially mammalian ones, despite their general similarities to the prokaryotic ribosomes, are much more complicated and less studied. This can be explained, in particular, by the fact that several approaches that have been fruitfully used for studying eubacterial ribosomes are not applicable yet to mammalian ribosomes, e.g. approaches based on the reconstitution of active ribosomal subunits from the total ribosomal protein and rRNA because there is still no methodology for the reconstitution of eukaryotic ribosomal subunits in vitro. One of these approaches includes site-directed mutagenesis to determine the roles of particular nucleotides of rRNA and amino acid residues of proteins or site-directed introduction of cross-linking groups into proteins to identify their inter- or intra-subunit contacts. Cryo-electron microscopy (cryo-EM) has been applied widely to eukaryotic ribosomes in the last decade [13–16] but it has not yet provided direct visualization of mRNA codons or the acceptor terminus of aminoacyl-tRNA on the ribosome, although the locations of eukaryote-specific ribosomal proteins on the subunits have been reported recently [17]. X-ray crystallography has been applied so far only to lower eukaryotic ribosomes free from mRNA and tRNAs [18,19]. The available data on ribosomal structure remain, to a significant extent, contradictory, especially in the localization of ribosomal proteins that have no eubacterial counterparts. In particular, the structures and locations of proteins S4e, S17e, S24e, S21e, S26e, and S30e are strongly or even completely different in X-ray crystallographic [19] and cryo-EM [17] studies. Moreover, one can find striking differences between the locations of several ribosomal proteins on the 40S subunit in the structures obtained by X-ray crystallography in different groups [18,19]. In particular, the sets of ribosomal proteins that comprise the subunit beak differ in these studies.

Direct data on the structural and functional organization of eukaryotic, in particular, human, ribosomes have been obtained with the use of a site-directed cross-linking (affinity labeling) approach that had been applied successfully to eubacterial ribosomes [20–24]. The approach is based on the use of chemically active analogs of various ribosomal ligands (mRNAs, tRNAs, and so on) that bear reactive groups (cross-linkers) to reveal the ribosomal components contacting these ligands or located near their binding sites. The cross-linkers are shown not to interfere with the formation of the specific complexes of the analogs with the ribosome. After formation of the cross-links within the complex and identification of the cross-linked ribosomal components, conclusions concerning the structure of the ribosomal ligand binding centers can be drawn. The application of affinity labeling approaches to study ribosomes was pioneered as early as ~40 years ago with chemically reactive derivatives of tRNA [25] and mRNA [26].

In the present review, we summarize experimental data on studying the structural and functional topography of human ribosomes that were obtained with the use of chemically reactive analogs of mRNA, HCV IRES, and tRNA bearing a cross-linker at designed locations. A comparison of the cross-linking results with cryo-EM and X-ray data on the structures of eukaryotic ribosomes made it possible to draw conclusions concerning the structural organization of functional sites in human ribosomes and to reveal similarities and differences in the organization of these sites in mammalian and eubacterial ribosomes.

Reactive Analogs of Ligands Used for Affinity Labeling of Human Ribosomes

Photoactivatable mRNA analogs

In the majority of studies, derivatives of relatively short (6–12 nt long) oligoribonucleotides were used as mRNA analogs. These derivatives bore a p-azidotetrafluorobenzoyl (ATB) group at the C5 of the selected uridine or the N7 of the selected guanosine, or the C8 of the selected adenosine, or at the 5′-terminal phosphate (Fig. 1). To introduce the ATB group in oligoribonucleotides, a general approach was used based on the initial insertion of a spacer containing an aliphatic amine group at the designed nucleotide residue with subsequent selective benzylation of this amino group with an N-hydroxysuccinimide ester of p-azidotetrafluorobenzoic acid [27–29]. Application of this type of mRNA analogs has two major advantages. First, the photoactivatable group used provides a relatively high yield of cross-links upon short (~1 min) irradiation with mild ultraviolet (UV) light (λ > 290 nm). Perfluoro-aromatic azides excited with UV light lose a nitrogen molecule and turn into the biradical nitrene, which is an active intermediate and can give a high yield of cross-links with both nucleic acids and proteins.
Second, short mRNA analogs can be exactly positioned on the ribosome by the addition of tRNA cognate to the selected mRNA codon (see below).

In addition to aryl azide derivatives of oligoribonucleotides that provide information on the mRNA environment on the ribosome, longer mRNA analogs (42–45-mers) containing single 4-thiouridine (sU) at designed positions were used. This type of mRNA analogs can form ‘zero-length’ cross-links reflecting direct contacts of mRNA with the ribosome. Such analogs had been widely applied earlier for studying mRNA binding center of eubacterial ribosomes [8]; sU-containing mRNA analogs can be easily prepared by in vitro T7 transcription from synthetic DNA templates [33,34].

Photoactivatable HCV IRES derivatives

An HCV RNA fragment corresponding to nucleotides 40–372 was used as an IRES element. This fragment was obtained by in vitro transcription of an appropriate plasmid using either 32P-labeled or biotinylated nucleoside triphosphates [35,36]. To introduce ATB groups at the selected HCV IRES positions, the strategy was based on the site-specific alkylation of an RNA with [4-(N-2-chloroethyl-N-methylamino)benzyl]phosphoramidate derivatives of oligodeoxyribonucleotides complementary to a sequence adjacent to the target site. An approach for site-specific modification of nucleic acids was suggested by Prof. Grineva and initially was called as ‘complementary addressed modification’ [37]. Site-specific alkylation occurs at any RNA nucleotide (except for U) that is located close to the modified terminus of the deoxy-oligomer derivative bound in the complementary complex with the target RNA sequence, leading to the formation of a covalent adduct [38]. After hydrolysis of the phosphoramidate bond in the covalent adduct, an aryl azide moiety is selectively introduced at the liberated aliphatic amine group (Fig. 2). The nucleotides alkylated with oligonucleotide derivatives could be determined readily using the property of reverse transcriptases to make stops or pauses at the modified nucleotides (see below). Using this strategy, a set of five HCV IRES derivatives bearing the cross-linker in the domain II and subdomains IIId and IIIe were obtained (Fig. 3).

tRNA analogs

Two types of reactive tRNA analogs were applied to study the molecular environment of the 3′-terminus of tRNA on the ribosome: analogs bearing sU as an additional 3′-terminal nucleotide [39], and an analog with an oxidized 3′-terminal ribose [40]. The sU-containing derivative was obtained from yeast tRNAAsp transcript by ligation with [32P]ps4Up and subsequent dephosphorylation of the product [41].

Model Complexes of Ribosomes with mRNA and tRNA Analogs

Cross-linking of mRNA and tRNA analogs with ribosomes was investigated in model complexes obtained without translation factors at 20°C and Mg2+ concentrations of 10–13 mM (in several experiments, lower Mg2+ concentrations and the polyamines spermine and spermidine were used whereever specified). The scheme for complex formation was based on two main observations. First, it was generally accepted that tRNAs in all forms have higher affinities to the ribosomal P site in the absence of translation factors [42–44]. The second was that the binding of short (up to 12 nt long) mRNA analogs to 80S ribosomes depends almost completely on the presence of the tRNA cognate to one of its codons (i.e. is not detected without tRNA [27,45–47]). Therefore, in the complex with the 80S ribosome, an mRNA analog is fixed on the ribosome in a designed position by interaction with the tRNA anticodon at the P site. To obtain complexes where, in addition to the P site, either the A or the E site was occupied with tRNA molecule too, tRNA cognate to the codon present at the respective site was bound to the ternary complex of 80S ribosomes with the mRNA analog and tRNA at the P site.
Generally, *Escherichia coli* tRNAs were used instead of the human species since they are more widely available and for tRNA<sup>Phe</sup> it had been shown that the qualitative and quantitative binding properties of human and *E. coli* tRNAs are practically the same [42]. To obtain 80S pre-termination complexes, mRNA analogs containing a stop codon with a cross-linker targeted to the A site were used; the addition of eRF1 to these complexes led to the formation of the termination complexes [48–51]. Examples of model complexes of mRNA analogs with 80S ribosomes and tRNAs are presented in Fig. 4. It is seen that the cross-linker could be placed in a wide spectrum of positions in the mRNA analog, i.e. in the P site-bound codon, or 5′ of it, and the A site-bound codon, or 3′ of it. To obtain comprehensive information about the arrangement of mRNA binding centers in the human ribosome, we used a large set of mRNA analogs that varied in their lengths, sequences, and positions of the cross-linker.

It is not so easy to achieve unambiguous tRNA-directed positioning on the 80S ribosome with mRNA analogs containing several dozen nucleotides, since polynucleotides often have a large intrinsic affinity for ribosomes in the absence of tRNA. In these cases, the presence of tRNA does not lead to a change in the location of the mRNA.

---

**Figure 2** Scheme of the site-specific introduction of a photoactivatable group into selected RNA sites. It was based on the complementary-addressed alkylation of the RNA with [4-(N-2-chloroethyl-N-methylamino)benzyl]-phosphoramides of oligodeoxyribonucleotides.

**Figure 3** Structural organization of the HCV IRES [6]. The initiation AUG codon is underlined. RNA sequences complementary to deoxy-oligomers used for site-specific modification of the HCV IRES are marked with thick lines with arrows indicating the 5′-phosphates derivatized with alkylating groups. Nucleotides of the HCV IRES modified with the oligonucleotide derivatives are shaded.

**Figure 4** Examples of model 80S ribosomal complexes with mRNA analogs consisting of two to four codons. mRNA positions, in which cross-linkers have been introduced, are marked (position +1 corresponds to the first nucleotide of the P site-bound codon). A, P, and E are the tRNA binding sites.
analog, shown by unaltered binding levels and cross-linking patterns of these analogs in the ribosomal complexes [52,53]. To overcome this difficulty, special adenine-rich mRNA analogs were designed whose affinities to 80S ribosomes were relatively low [54,55]. With these analogs, the addition of tRNA cognate to the selected codon resulted in an increase in the level of mRNA binding and the appearance of new cross-links that could be unambiguously assigned to the complex where mRNA analog was fixed on the ribosome by codon–anticodon interaction at the P site. The sequences of longer mRNA analogs containing s4U residues at various positions and the types of complexes that they form with 80S ribosomes are presented in Fig. 5.

The cross-linking of HCV IRES derivatives was studied in their binary complexes with 40S subunits obtained at 2.5 mM Mg2+ concentration. Under these conditions, the HCV IRES forms a stable binary complex, in contrast to all other RNAs; formation of this complex is the first step of the cap-independent initiation of HCV RNA translation [5,6]. Introduction of ATB groups at specific positions in domain II and subdomains IIIId/e of HCV IRES did not interfere with its ability to bind to 40S subunits [36].

Model complexes of 80S ribosomes with tRNAAsp analogs were obtained as described above, and a 12-mer oligoribonucleotides bearing an Asp codon were used as mRNA analogs (Fig. 6).

Cross-linking Procedure and Identification of the Cross-linked Ribosomal Components

All mRNA analogs were 5'-32P-labeled prior to use in cross-linking experiments; in some experiments, s4U-containing analogs were labeled at the internal phosphate 3' to the s4U residue. In the latter case, labeled analog was obtained by ligation of two ‘halves’, the 3'-half being 5' -end labeled [34]. To obtain cross-links with photoactivatable mRNA and tRNA analogs, the respective ribosomal complexes were irradiated with mild UV light; the conditions of irradiation were somewhat different between mRNA analogs bearing aryl azides [47] and s4U [54]. A tRNAAsp analog with an oxidized 3'-terminal ribose was cross-linked to ribosomal proteins by incubation of the ribosomal complex with sodium cyanoborohydride. With short photoactivatable mRNA analogs and with tRNA analogs (bound to 80S ribosomes in the presence of short mRNA analogs), the distribution of the cross-linked 32P-labeled analog between the ribosomal subunits can be examined easily by centrifugation in a sucrose density gradient (10%–30%) under conditions, in which 80S ribosomes are dissociated into the subunits (at 3 mM Mg2+ concentration and 300–500 mM KCl) [27,28,46,47,49]. Under these conditions, complexes dissociate, so radioactivity in the subunit fractions indicates covalent attachment of the labeled mRNA or tRNA analog. Generally, mRNA analogs cross-linked only to the 40S subunits, and tRNA analogs to the 60S subunits. The distribution of the cross-links between the ribosomal subunits can be examined easily by centrifugation in a sucrose density gradient (10%–30%) under conditions, in which 80S ribosomes are dissociated into the subunits (at 3 mM Mg2+ concentration and 300–500 mM KCl) [27,28,46,47,49]. Under these conditions, complexes dissociate, so radioactivity in the subunit fractions indicates covalent attachment of the labeled mRNA or tRNA analog. Generally, mRNA analogs cross-linked only to the 40S subunits, and tRNA analogs to the 60S subunits. The distribution of the cross-links between the rRNA and ribosomal proteins in the isolated 40S subunits could be examined easily, after denaturing the cross-linked subunits by incubation in sodium dodecyl sulfate (SDS) and ethylenediaminetetraacetic acid (EDTA), by either polyacrylamide gel electrophoresis (PAGE) [47,56] or centrifugation in a sucrose density gradient in the presence of SDS and EDTA [45,46,57].

Determination of cross-linking sites in rRNA

The methodology that was applied for identifying the nucleotides cross-linked in rRNA is based on a two-step procedure. At first, the rRNA sequences containing sites of cross-linking of 32P-labeled mRNA or tRNA analogs are determined using hydrolysis of cross-linked rRNA with
RNase H in the presence of deoxy-20-mers complementary to various rRNA sequences, with subsequent separation of the resulting RNA fragments by PAGE [46,47,49,57]. The exact identification of the cross-linked rRNA nucleotides is performed by primer extension using the cross-linked rRNA as a template; primers, typically deoxy-20-mers, are chosen based on the results of experiments with RNase H. The cross-linking site is generally assumed to be the nucleotide 5′ to the primer extension stopping (pause) site whose position is determined by PAGE analysis of the primer extension products in a denaturing gel, with a parallel sequencing reaction on unmodified rRNA. Control experiments were done with RNA isolated from unmodified subunits that passed through the same procedures as the modified ones, but without chemically reactive ribosomal ligands.

Identification of cross-linked ribosomal proteins

With mRNA analogs, total 40S ribosomal protein was isolated from the modified subunits, and labeled proteins were identified using one- and two-dimensional PAGE in various electrophoretic systems with subsequent autoradiography of the gels [28,45,47,54,56]. The radioautograms were superimposed on the respective stained gels and the positions of the radioactive bands with respect to those of unmodified proteins were determined. In those cases when it was possible without loss of the label, prior to electrophoresis, mRNA analogs cross-linked to proteins were hydrolysed with RNase A to reduce the effect of oligonucleotide on the electrophoretic mobility of the protein. In some cases, the identity of a cross-linked protein was confirmed by immunoblotting [47], mass spectrometry [54], or other approaches [47].

To identify ribosomal proteins cross-linked to tRNA analog oxidized at the 3′-terminal ribose, cross-linked protein was isolated by one-dimensional PAGE, and the appropriate piece was excised from the gel, treated with RNase T1 and proteinase K, and the resulting peptides were determined by mass spectrometry to identify the cross-linked protein [40].

Ribosomal proteins cross-linked to HCV IRES derivatives were identified by one-dimensional and two-dimensional PAGE after exhaustive hydrolysis of the cross-linked IRES with RNases A and T1 [35,36]; the identities of the cross-linked proteins were then confirmed by immunoprecipitation with the use of rabbit antibodies specific to the candidate proteins [36].

Arrangement of mRNA Binding Site of the Mammalian Ribosome

The application of a large set of mRNA analogs varying in their sequence, the nature of cross-linker and the site of its attachment allowed us to obtain detailed information about the molecular environment of mRNA nucleotides in positions from −9 to +12 with respect to the first nucleotide of the P site codon on the 80S ribosome (Tables 1 and 2 and Fig. 7). This made it possible to find universal features of the mRNA binding site in all kingdoms and to identify peculiarities specific to eukaryotic (or even to mammalian) ribosomes. Cross-linking data summarized in Table 1 show that cross-linking of mRNA analogs to the 18S rRNA occurs mainly when modified nucleotides are located at the A or the P site (in positions from +6 to +1); cross-linking to ribosomal proteins is observed as well. The yield of the cross-links with ribosomal proteins is similar to that with the 18S rRNA, while in eubacterial 30S subunits A and P site codons of mRNA contact only rRNA. mRNA regions 5′ of the P site codon or 3′ of the A site codon cross-link presumably to ribosomal proteins. Moreover, those mRNA nucleotides that are the most distant from the A and the E sites (both downstream and upstream ones) neighbor practically only ribosomal proteins. All these indicate that proteins play more significant roles in the formation of the ribosomal mRNA binding site in eukaryotes than in eubacteria.

Nucleotides of 18S rRNA at the mRNA binding site

Cross-linking of short ATB-derivatized mRNA analogs to the 18S rRNA was completely position specific, i.e. it was not observed in binary mixtures of the analogs with 80S ribosomes. All cross-linked 18S rRNA nucleotides fall into the conserved core of the small subunit rRNA secondary structure (Fig. 8), which makes it possible to find corresponding nucleotide positions in the eubacterial 16S rRNA. It shows that positions of cross-linked 18S nucleotides almost completely coincide with those of 16S rRNA nucleotides that interact with or neighbor the respective mRNA positions known from X-ray data on complexes of mRNA with eubacterial ribosomes [9–11]. This is clearly seen when these 16S rRNA nucleotides are marked on the atomic model of the eubacterial 30S ribosomal subunit (Fig. 9). Thus, our studies experimentally confirmed the widely accepted idea on the conserved rRNA ‘core’ of the ribosome and showed that the small subunit rRNA core is involved mainly in the binding of mRNA codons interacting with tRNAs at the A and the P sites. It should be noted that 18S rRNA nucleotides cross-linked to s4U and ATB-derivatized nucleotides placed at the same mRNA positions in some cases do not exactly coincide (Table 1). However, in the 30S subunit spatial structure, 16S rRNA positions corresponding to these 18S rRNA nucleotides are very close (within 10–15 Å) to each other. Thus, cross-links of perfluorophenyl azide-modified mRNA analogs to the 18S rRNA actually reflect the immediate neighborhood of the derivatized mRNA nucleotides in the eukaryotic
ribosomal complexes. Moreover, molecular environment of the A site codon practically did not depend on the codon type (sense or stop codon) and on the presence of termination factor eRF1 [49, 63]. Evidently, the application of different types of cross-linkers provided correct and comprehensive information on the molecular environment of mRNA on the ribosome. Remarkably, data very similar to those presented in Table 1 have been obtained in Pestova’s group, where a panel of mRNA-containing s^4U at positions from -26 to +11 relatively to the first nucleotide of the AUG codon were used for cross-linking in rabbit 48S and 80S initiation complexes assembled with the use of a set of the appropriate initiation factors [64, 65]. Thus, the conserved core of mammalian 18S rRNA accommodates mRNA very similarly during all steps of translation.

### Protein environment of mRNA region 3’ of the A site codon

Similarities in protein environments of mRNA on eubacterial and mammalian ribosomes concern presumably mRNA region 3’ of the A site codon that neighbors mainly ribosomal protein (rp) S3e (major) and rpS2e (Table 2, Figs. 7 and 9), which is confirmed by the data obtained with s^4U-containing mRNA analogs bound in 48S and 80S initiation complexes [65]. RpS3e and rpS2e are homologous to

### Table 1 Summary of mRNA analogs cross-links with the 18S rRNA in human 80S ribosomal complexes

<table>
<thead>
<tr>
<th>Position of nucleotide bearing cross-linker with respect to the first nucleotide of the P site codon</th>
<th>Nature of the modified nucleotide and the cross-linker</th>
<th>Share of cross-linked 18S rRNA (in % of the total amount of cross-linked 40S components)</th>
<th>Cross-linked 18S rRNA nucleotides (corresponding nucleotides of <em>Escherichia coli</em> 16S rRNA are given in brackets)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>-9</td>
<td>U, G (ATB)</td>
<td>&lt;2</td>
<td>–</td>
<td>[47]</td>
</tr>
<tr>
<td>-6</td>
<td>U, G (ATB)</td>
<td>&lt;3</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>-4</td>
<td>U, G (ATB)</td>
<td>&lt;3</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>-3</td>
<td>5’p (ATB)</td>
<td>6–8</td>
<td>G1207 (G926)</td>
<td>[57]</td>
</tr>
<tr>
<td>-3</td>
<td>U (ATB)</td>
<td>7</td>
<td>G961 (G693)</td>
<td>[46]</td>
</tr>
<tr>
<td>-3</td>
<td>G (ATB)</td>
<td>n.d.</td>
<td>G1207</td>
<td>[58]</td>
</tr>
<tr>
<td>-2</td>
<td>U (ATB)</td>
<td>n.d.</td>
<td>G961</td>
<td>–</td>
</tr>
<tr>
<td>-1</td>
<td>U (ATB)</td>
<td>n.d.</td>
<td>G1207</td>
<td>–</td>
</tr>
<tr>
<td>-3, -2, -1</td>
<td>s^4U</td>
<td>n.d.</td>
<td></td>
<td>Fragment 1699–1704 (1398–1403) [55]</td>
</tr>
<tr>
<td>+1</td>
<td>5’p (ATB)</td>
<td>40–70</td>
<td></td>
<td>G1207 [57]</td>
</tr>
<tr>
<td>+1</td>
<td>U (ATB)</td>
<td>50</td>
<td></td>
<td>G1207 [46]</td>
</tr>
<tr>
<td>+1</td>
<td>G (ATB)</td>
<td>30</td>
<td></td>
<td>G1702 (G1401) [47]</td>
</tr>
<tr>
<td>+2, +3</td>
<td>U (ATB)</td>
<td>n.d.</td>
<td></td>
<td>G1207 [58]</td>
</tr>
<tr>
<td>+4</td>
<td>G (ATB)</td>
<td>70</td>
<td>A1823–1825 (1491–1493)</td>
<td>[46]</td>
</tr>
<tr>
<td>+4</td>
<td>U (ATB)</td>
<td>n.d.</td>
<td>A1823–1825</td>
<td>[47, 49]</td>
</tr>
<tr>
<td>+4</td>
<td>s^4U</td>
<td>&lt;30</td>
<td>C1696 (C1395)</td>
<td>[54]</td>
</tr>
<tr>
<td>+5</td>
<td>U (ATB)</td>
<td>30</td>
<td>A1823–1825</td>
<td>[47]</td>
</tr>
<tr>
<td>+5</td>
<td>G (ATB)</td>
<td>n.d.</td>
<td>A1823–1825, G626 (G530)</td>
<td>–</td>
</tr>
<tr>
<td>+6</td>
<td>U (ATB)</td>
<td>30</td>
<td>A1823–A1825</td>
<td>[58]</td>
</tr>
<tr>
<td>+5, +6</td>
<td>s^4U</td>
<td>n.d.</td>
<td>C1696 (1395), C1698 (1397), fragment 1820–1825</td>
<td>[55]</td>
</tr>
<tr>
<td>+7</td>
<td>G/U (ATB)</td>
<td>20/8</td>
<td>A1823–A1825, region 605–630 (509–534), C1698</td>
<td>[59]</td>
</tr>
<tr>
<td>+9</td>
<td>G/U (ATB)</td>
<td>10/7</td>
<td>A1823–A1825, region 605–630, C1698</td>
<td>–</td>
</tr>
<tr>
<td>+12</td>
<td>G/U (ATB)</td>
<td>2</td>
<td>A1823–A1825, region 605–630, C1698</td>
<td>–</td>
</tr>
</tbody>
</table>

Major cross-links are given in bold; n.d., not determined.
prokaryotic rpS3p and rpS5p, respectively, which are known as components of the ‘ribosomal mRNA entry site’ [66]. Earlier, rpS3p and rpS5p were found to cross-link to mRNA analogs with s^{4}U residues in positions +8 and +11 [67]. According to X-ray data, these proteins interact with phosphates of mRNA nucleotides in positions +11 to +15 [68]; in particular, residues Arg131, Arg132, Lys135, and Arg164 in the rpS3p contact mRNA. Alignment of amino acid sequences of Thermus thermophilus rpS3p and human rpS3e shows that three of these residues (Arg131, Arg132, and Lys135) are located in the conserved parts of the proteins [69]. A rough mapping of the mRNA analogs cross-linking sites on the rpS3e using CNBr-induced cleavage of the cross-linked protein with the subsequent identification of the labeled peptides reveals that these sites are actually located in the N-terminal half of the protein (amino acid 2–127) containing all three above-mentioned conserved amino acid residues [69]. This indicates that arrangements of mRNA regions 3′ of the A site codon on the eubacterial and mammalian ribosomes share significant extent of similarity. However, an X-ray crystallographic study of yeast 80S ribosome [18] showed that arrangement of the ribosomal area corresponding to the region of mRNA entry site was substantially different in 70S and 80S ribosomes. In particular, distinctions have been found in protein–rRNA and protein–protein interactions involving rpS3e and rpS2e. Thus, similarities in the organization of binding sites of the mRNA

<table>
<thead>
<tr>
<th>Position of nucleotide bearing cross-linker with respect to the first nucleotide of the P site codon</th>
<th>Nature of the modified nucleotide and the cross-linker</th>
<th>Share of cross-linked ribosomal proteins (in % of the total amount of cross-linked 40S components)</th>
<th>Cross-linked ribosomal proteins</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>−9</td>
<td>U (ATB)</td>
<td>&gt;98</td>
<td>S26</td>
<td>[47]</td>
</tr>
<tr>
<td>−9</td>
<td>G (ATB)</td>
<td>&gt;98</td>
<td>S26, S14</td>
<td></td>
</tr>
<tr>
<td>−6</td>
<td>U (ATB)</td>
<td>&gt;97</td>
<td>S26</td>
<td></td>
</tr>
<tr>
<td>−6</td>
<td>G (ATB)</td>
<td>&gt;97</td>
<td>S26, S14</td>
<td></td>
</tr>
<tr>
<td>−4</td>
<td>U, G (ATB)</td>
<td>&gt;97</td>
<td>S26, S3</td>
<td></td>
</tr>
<tr>
<td>−3</td>
<td>U (ATB)</td>
<td>93</td>
<td>S3, S26, S2, S5/S7, S14</td>
<td>[60]</td>
</tr>
<tr>
<td>+1</td>
<td>U (ATB)</td>
<td>50</td>
<td>S3, S26, S2, S14, S23</td>
<td>[60]</td>
</tr>
<tr>
<td>+4</td>
<td>G (ATB)</td>
<td>30</td>
<td>S15, S2, S30</td>
<td>[28]</td>
</tr>
<tr>
<td>+4</td>
<td>U (ATB)</td>
<td>50</td>
<td>S15, S3</td>
<td>[47]</td>
</tr>
<tr>
<td>+5</td>
<td></td>
<td>70</td>
<td>S15, S3</td>
<td></td>
</tr>
<tr>
<td>+6</td>
<td></td>
<td>70</td>
<td>S15, S3</td>
<td></td>
</tr>
<tr>
<td>+4</td>
<td>s^{4}U</td>
<td>&gt;70</td>
<td>S15, S3, S3a, S2, S30</td>
<td>[54]</td>
</tr>
<tr>
<td>+5, +6</td>
<td>s^{4}U</td>
<td>n.d.</td>
<td>S15, S2/S3/S3a, S30</td>
<td></td>
</tr>
<tr>
<td>+7</td>
<td>G/U (ATB)</td>
<td>80/92</td>
<td>S3, S15, S2, S30</td>
<td>[56]</td>
</tr>
<tr>
<td>+9</td>
<td>G/U (ATB)</td>
<td>90/93</td>
<td>S3, S2, S15, S30</td>
<td></td>
</tr>
<tr>
<td>+12</td>
<td>G/U (ATB)</td>
<td>98</td>
<td>S3, S2</td>
<td></td>
</tr>
</tbody>
</table>

Major cross-links are given in bold; n.d., not determined.

**Figure 7 Schematic representation of the protein neighborhood of mRNA on the 80S ribosome** Solid lines correspond to the major neighbors; eubacterial counterparts of human ribosomal proteins are given in brackets.
region 3' of the A site codon on 70S and 80S ribosomes are limited, and further more detailed mapping of mRNA cross-linking sites on rpS3e and rpS2e could make it possible to reveal differences in mRNA binding tracks in the 70S and 80S ribosomal entry site regions.

It should be noted that cross-linking of mRNA analogs to rpS3e and/or rpS2e was detected also in the majority of other complexes where nucleotide bearing the cross-linker was in positions from −4 to +6 (Table 2). Besides, with all mRNA analogs used, cross-linking to these proteins was observed in binary complexes of the analogs with 80S ribosomes that were hardly detectable by nitrocellulose filtration technique due to their lability. Evidently, this cross-linking occurred because of rapid turnover of mRNA analogs in these complexes leading to relatively high yield of cross-linking in spite of low stability of the complexes. Therefore, rpS3e and rpS2e were not assigned to the environment of mRNA nucleotides in positions from −4 to +6. Evidently, all mRNA analogs have a pronounced affinity to the 80S ribosomal entry site, where they bind in a labile manner in the absence of tRNA.

Figure 8 Positions of cross-linked nucleotides in the 18S rRNA secondary structure It was taken from http://www.rna.ccbb.utexas.edu.

The results obtained with ATB-derivatized and s4U-containing mRNA analogs are shown with solid and dotted lines, respectively. Positions of nucleotides bearing cross-linkers with respect to the first nucleotide of the P site codon are boxed. Helices of the 18S rRNA containing cross-linked nucleotides are numbered according to [61].

Figure 9 Path of mRNA on the eubacterial 30S subunit as revealed from the X-ray study of a complex of eubacterial 70S ribosome with mRNA and tRNAs [62]. Locations of 16S rRNA nucleotides whose positions in the secondary structure correspond to those 18S rRNA nucleotides that cross-linked to mRNA analogs in the human ribosome (are given in brackets) are marked with yellow circles. The mRNA positions are presented in red boxes. Locations of several 30S proteins homologous to cross-linked human 40S proteins are shown.

Protein environment of mRNA at the decoding site

The most distinctive feature of the protein environment of the A site codon on the 80S ribosome is rpS15e. This protein cross-links to ATB-derivatized nucleotides of mRNA analogs in positions +4 to +7 (the yield of cross-linking decreases significantly when the modified nucleotide moves from position +4 to position +6 [47]; cross-links from position +7 are minor, and from downstream positions are hardly detectable [56]). mRNA analogs bearing s4U cross-linked to rpS15e when the photoactivatable nucleotide was in positions +4 and +5 in complexes representing elongation [54] and initiation [64] steps of translation. All these data imply that rpS15 contacts the first and second nucleotides of the A site codon and is relatively close to mRNA nucleotides in positions +6 and +7. RpS15e is homologous to eubacterial rpS19p that is located mainly on the 30S subunit head away from the mRNA track, but has a C-terminal tail extending toward the decoding site [68,70]. It was suggested that the tail of mammalian rpS15e comes closer to the decoding site than that of its eubacterial counterpart rpS19p [47]. RpS15e has been mapped on the head of 40S subunit in several cryo-EM studies based on docking of known atomic structure of eubacterial rpS19p into the 40S density map [15–17], but the unstructured C-terminal tail that has almost no sequence homology with the tail of eubacterial rpS19p [71], could not be mapped this way. Moreover, the main part of the...
unstructured C-terminal tail of rpS15e is lacking in the 40S subunit atomic structure obtained by X-ray crystallography [19]. To determine peptides of rpS15e cross-linked to labeled mRNA analogs in the 80S ribosomal complexes, cross-linked labeled rpS15 was isolated from the irradiated ribosomal complexes by SDS–PAGE [71]. Radioactive bands corresponding to cross-linked rpS15e were excised from the dried gels and treated with various specific proteolytic agents with subsequent electrophoretic separation of the resulting modified oligopeptides. Then these oligopeptides were identified based on comparison of the electrophoretic data with the map of rpS15e cleavages. Finally, the cross-link was mapped to the C-terminal rpS15e decapeptide 131–140 specific to eukaryotes and archaea. The same results were obtained with an mRNA analog bearing ATB-derivatized stop codon targeted to the A site in the presence of termination factor eRF1 [71]. It was suggested that the C-terminal tail of human rpS15e, in contrast to the tail of eubacterial rpS19p, is able to reach the ribosomal decoding site and interact with the mRNA due to its larger length and specific location in the 40S subunit. A variant of location of the C-terminal sequence of human rpS15e with respect to the mRNA based on similar positioning of the structured parts of rpS15e and rpS19p in the small ribosomal subunits [15–17] has been proposed (Fig. 10). In the atomic model of the 40S subunit [19], the last solved residue, an amino acid in position 130, is ≏30 Å away from the of the putative site for the A site codon of mRNA (deduced on the basis of similar positioning of the A site codon with respect to the conserved rRNA decoding site in prokaryotic and eukaryotic ribosomes). These are in a good agreement with our results and with the proposed variant of location of the C-terminal tail of rpS15e in the 40S subunit (see above).

**Protein environment of mRNA region 5′ of the P site codon**

It is seen that one of the main components of the mRNA environment upstream of the P site codon is rpS26e that is almost the only neighbor of mRNA nucleotides in positions −4 to −9. These data are mostly confirmed by later results obtained with the above-mentioned rabbit 48S and 80S initiation complexes [65], which indicates that rpS26e is a key-stone component of the ribosomal binding site of mRNA region 5′ of the E site codon, at least at the initiation and elongation of translation. This protein lacks eubacterial counterparts, and therefore, molecular environment of mRNA 5′ of the E site codon on the mammalian ribosome is strikingly different from that on the eubacterial ribosome. It is worth mentioning here that labeling of this protein has been reported in early studies of the mammalian ribosomal mRNA binding site with the use of derivatives of oligo U bearing an alkylating group at the 5′-terminal phosphate bound to human and rat ribosomes in the presence of tRNA<sub>Phe</sub> [45,72,73]. In these reports, positions of the derivatized mRNA nucleotides on the ribosome remained indefinite; nevertheless, it is evident that they could be either in the first position of the P site codon or 5′ of it. RpS26e has been mapped on 40S subunits from lower eukaryotes initially by cryo-EM [17] and then more precisely by X-ray crystallography [19] in a region of the platform side facing the head near the mRNA exit site region. The X-ray study showed that, in contrast to the flexible 3′ end of eubacterial 16S rRNA, the 3′ end of the eukaryotic 18S rRNA appears to be locked in place by tight interactions with rpS26e [19] and therefore it is not accessible for contacts with mRNA region 5′ of the E site codon. The latter is a good explanation for hardly detectable cross-linking to the 18S mRNA from the respective mRNA positions. Based on the comparison of our results with the mentioned X-ray data and the results reported previously [65], one can suggest that rpS26e directly interacts with mRNA nucleotides in positions −6 to

---

**Figure 10** A proposed variant of the location of the C-terminal sequence of human rpS15e with respect to the mRNA. The scheme is based on similar locations of the structured parts of rpS15e and rpS19p in the small ribosomal subunits. Atomic model of the 70S ribosomal complex was taken from [70] (PDB ID code 1IBM). All components of the complex except rpS19p and the mRNA were removed from the model; a minimal distance between these two molecules is shown. The C-terminal sequence 120–145 of human rpS15e (presented in dark red) was built arbitrarily by sequential attachment of the respective amino acid residues taken from the model 1IBM, and then attached virtually to the structured part of rpS19p (presented in purple) to fit the cross-linking results. Boxed tetrapeptide 137-HSSR-140 is the most probable candidate for interaction with the mRNA codon.
–9. Apparently, weaker cross-linking of ATB-derivatized uridines in positions –4 [47] and +1 [60] to rpS26e was detectable because the ATB group used was attached to uridines via a flexible linker enabling this group to reach targets located close to the mRNA but not making direct contacts with it. With the same approach as described above for mapping of cross-link on the rpS15e, it was found that cross-linking site of mRNA analogs bearing ATB group in positions 26 and 29 on the rpS26e was located in the fragment 60–71. Alignment of amino acid sequences of eukaryotic (from lower fungal to mammalian) and archaeal proteins from the rpS26e family showed that this fragment contains motif 62-YxxPKxYxK-70 conserved in eukaryotic but not in archaeal rpS26e.

Possible roles of peptides 131–140 of rpS15e and 60–71 of rpS26e in translation
Specific features of the arrangement of mRNA binding site on the mammalian ribosome found by site-directed cross-linking likely relate to more complex and multistep regulation of protein synthesis in eukaryotes. Analysis of X-ray structure of the *Tetrahymena thermophila* 40S subunit [19] showed that eukaryote-specific motif YxxPKxYxK of rpS26e is not implicated in the intra-ribosomal interactions, implying its involvement in the translation process in a eukaryote-specific manner, e.g. in binding of eukaryote-specific ligands. One of these ligands could be eukaryotic initiation factor eIF3 that is involved in recruitment of the mRNA to the 40S subunit and has no counterparts in bacteria and archaea [2]. This suggestion is based on the data on overlapping locations of rpS26e [19] and eIF3 [74] on the 40S subunit (Fig. 11) and the data on contacts of mRNA nucleotides in position −8 and upstream of it with eIF3 [65]. The role of the mentioned rpS26e motif could extend beyond the initiation step by its possible contribution to the maintenance of the mRNA path from the region of codon–anticodon interactions to the exit site during translation. Finally, the YxxPKxYxK motif could be a potential target for the action of some kind of regulatory factors modulating ribosomal function during translation since it is partly accessible at the surface of the 40S subunit solvent side [19].

Several hypotheses have been suggested to explain possible roles of the C-terminal eukaryote/archaeal-specific decapeptide of rpS15e found at the decoding site of human ribosomes [71]. Remarkably, this decapeptide was not found in the Protein Data Bank in any other proteins than rpS15e; only short fragments of this sequence (not longer than penta- or hexapeptide motifs) were traced in several protein families. The data point to an important role of this part of the unstructured C-terminal tail of rpS15e in the translation process in eukaryotes and probably in archaea. During initiation, as suggested in the previous report [65], it could be important for the selection of initiation codon and/or for the stabilization of 48S complexes since rpS15e contacts mRNA positions +4 and +5 in these complexes. During elongation, it possibly contributes to the fixation of mRNA codon at the decoding site that, in turn, could provide higher accuracy of translation and less probability of a frameshift. During termination, the decapeptide possibly interacts with polypeptide chain release factor eRF1 and thus mediates stop codon decoding. This assumption is based on the data on shielding of rpS15e from contacts with mRNA by eRF1 [75]. Finally, the C-terminal

Figure 11  Comparison of locations of eIF3 and ribosomal proteins on the 40S subunit  (A) Cryo-EM image of the complex of eIF3 with the rabbit 40S subunit (adapted from [74]). (B) X-ray structure of the 40S subunit from *T. thermophila* where ribosomal proteins are shown [19], an alternative designation of rpS1e is given in brackets. Locations of rpSA and rpS26e neighboring eIF3 on the cryo-EM image (A) are shown by homology with positioning of their counterparts on the X-ray structure (B).
decapetide of rpS15e may be involved in the regulation of translation in a manner specific for eukaryotes.

**HCV IRES Binding Site on the 40S Subunit**

The sequence of events leading to the formation of the HCV IRES-containing 80S initiation complex competent for the start of the viral RNA translation is well known [76–78]. The functional roles of individual IRES domains and subdomains (Fig. 3) have been studied using mutant IRESes [5,76–80], cross-linking with chemically reactive derivatives of oligonucleotides complementary to target RNA sequences [81], and enzymatic and chemical footprinting [79,80,82]. According to the generally accepted idea, at the first step of the assembly of the 80S initiation complex, HCV IRES forms a stable binary complex with the 40S subunit, where the initiation codon is positioned near the P site without any initiation factors. The binary complex formation crucially depends on the basal portion of domain III (mainly hairpins IIId and IIIe) and on the four-way junction IIIabc (Fig. 3). Thereafter, elf3 and the ternary complex (elf2-Met-tRNA_{Met}•GTP) assemble to form the 48S preinitiation complex; binding of elf3 is mediated by domain IIIb and four-way junction IIIabc. Binding of elf5 to this complex promotes hydrolysis of elf2-bound GTP and the subsequent release of elf2•GDP, these are modulated by HCV IRES domain II [78]. Finally, the release of the remaining factor elf3 requires elf5B and GTP hydrolysis and occurs during joining of the 60S subunit to form the 80S initiation complex. When elf2 is inactivated by phosphorylation under stress conditions, HCV IRES can direct translation without elf2 and elf5 using a bacterial-like pathway requiring as initiation factors only elf5B (an analog of bacterial IF2) and elf3 [83].

Knowledge on the ribosomal components responsible for binding of HCV IRES to 40S subunits is important to understand the mechanism of translation initiation of the viral RNA and could be essential for the development of novel efficient anti-HCV therapies. However, at the beginning of our studies, the structural organization of the IRES binding site on the 40S subunit was less studied than functional roles of its specific domains, and the resulting data were rather contradictory. So, direct UV-induced cross-linking of HCV IRES to the 40S subunit revealed rpS5e as the only target for cross-linking [84]. Other investigators using HCV IRES with uridines randomly substituted with 4-thiouridines reported a set of cross-linked proteins (rpS2e, rpS3e, rpS10e, rpS15e, rpS16e/S18e, and rpS27e), of which rpS3e, rpS2e, and rpS27e were the major targets [85]. Finally, with the use of yeast genetics it was shown that rpS25e is required for translation initiation of HCV IRES [as well as of cricket paralysis virus (CrPV)] while the lack of this protein in yeast strains does not lead to serious defects in global translation, read-through, ribosome biogenesis, and programmed ribosomal frameshifting [86]. The above-mentioned studies highlighted a set of proteins that could surround the 5′-terminal part of HCV RNA, but did not provide information on the specific ribosomal proteins neighboring particular RNA nucleotides. Cryo-EM visualization of HCV IRES in its complexes with the 40S subunit [87] and the 80S ribosome [88] provided a general idea of the IRES positioning on the ribosome. So, two ribosomal proteins, namely, rpS5e on the head of the subunit interface and rpS14e on the platform were mapped close to the apex of subdomain IIb of the IRES [88]; recent cryo-EM data on positioning of CrPV IRES [89] make it possible to propose that N-terminal tail of rpS25e extends toward rpS5e. However, domain II, in contrast to subdomains IIId/e and the four-way junction IIIabc, is not essential for binding of HCV IRES to 40S subunit (see above). Thus, ribosomal proteins that could contact keystone structural elements of HCV IRES leading to its high affinity for the 40S subunit remained unknown.

With the use of HCV IRES derivatives bearing ATB group at the selected nucleotides, it was shown that ribosomal proteins but not the 18S rRNA take part in the formation of HCV IRES binding site on the 40S subunit. The summary on our cross-linking results is presented in Table 3 and Fig. 12. These results are in good agreement with location of HCV IRES on the 40S subunit determined by cryo-EM at a low resolution [87]. Seen from Fig. 12, all cross-linked proteins are located mainly on the solvent side of the 40S subunit opposite to the beak around the cleft between the head and the body, at an area homologous to the region of mRNA exit site in the 30S subunit and overlapping with the E site. The only pronounced dissimilarity between our cross-linking results and the cryo-EM data is that rpS5e appears to neighbor subdomain IIie [35] rather than domain II [87]. This in fact is not a real discrepancy because the absence of cross-link can not indicate lack of neighborhood or even interaction of the IRES with rpS5e since the protein could not have a suitable target for cross-linking or ATB group could be oriented unfavorably with respect to the protein to obtain a cross-link. The same reasons may explain the lack of cross-linking of HCV IRES derivatives to rpS25e.

The sets of 40S ribosomal proteins cross-linked to s4'U-containing HCV IRES [85] and to ATB-derivatized IRESes (Table 3) are significantly different. It is clear now that a significant part of the set of proteins found previously [85], namely rpS2e, rpS3e, and rpS15e, belongs to the environment of the mRNA at the decoding site and downstream of it (Table 2, Fig. 7, and the respective discussion in the text). These proteins most probably did not cross-link to the 5′-UTR of the HCV IRES but did to the part of the coding sequence that could contain 4-thiouridines in
positions +2, +10, and +13 (Fig. 2). Consequently, only proteins rpS10e, rpS16e/S18e, and rpS27e may be considered as cross-linked to the 5′-UTR of the HCV RNA in the binary complex. This set overlaps with the set found in [35,36] only by protein rpS16e. One possible explanation for this discrepancy is that proteins rpS10e and rpS27e are cross-linked to 4-thiouridines located in IRES positions far from nucleotides C83, G263, A275, or A296 that bore per-fluorophenyl azides. Finally, identification of one of cross-linked proteins as rpS10e [85] could be mistaken since, according to the recent X-ray data on the 40S subunit structure [19], this protein is located at the beak region, which is incompatible with all the above-mentioned data on the location of HCV IRES on the subunit.

Our cross-linking results provided significant novel information on the organization of the HCV IRES binding site on the 40S ribosomal subunit at the first step of translation initiation of the viral RNA, highlighting the key role played by the ribosomal proteins SA (p40), S3a, S5e, S14e, and S16e in the organization of this site. It is worth mentioning here that location of rpS3a on the 40S subunit that had been predicted based on the cross-linking results [36] has been exactly confirmed by later X-ray crystallographic data [19] (Fig. 12). Remarkably, HCV IRES binding site has only a minor overlap with the mRNA binding track in the region of the mRNA exit site (Fig. 12). These data can be used for creation of new types of antiviral drugs targeted to proteins involved in the formation of only the IRES binding site. It should be noted here that binding sites of various IRESes on the 40S subunit could be significantly different. For instance, in contrast to the HCV IRES, CrPV IRES is located entirely in the mRNA binding cleft, reaching into the P and the A sites [90]. Nevertheless, these IRES binding sites share some extent of similarity, in particular, both CrPV IRES and HCV IRES neighbor rpS5e and moreover, the CrPV IRES is located close to rpS16e as one could see upon

### Table 3

<table>
<thead>
<tr>
<th>Position and nature of nucleotide bearing cross-linker</th>
<th>HCV domain (subdomain)</th>
<th>Cross-linked ribosomal proteins</th>
<th>Eubacterial counterparts of cross-linked ribosomal proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>G263</td>
<td>IIId</td>
<td>S3a</td>
<td>No counterpart</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S14e</td>
<td>S11p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S16e</td>
<td>S9p</td>
</tr>
<tr>
<td>A275</td>
<td></td>
<td>S3a</td>
<td>No counterpart</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S14e</td>
<td>S11p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S16e</td>
<td>S9p</td>
</tr>
<tr>
<td>A296</td>
<td>IIe</td>
<td>S3a</td>
<td>No counterpart</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S14e</td>
<td>S11p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S16e</td>
<td>S9p</td>
</tr>
<tr>
<td>C83</td>
<td>Apical loop of domain II</td>
<td>S14e</td>
<td>S11p</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S16e</td>
<td>S9p</td>
</tr>
<tr>
<td>G87</td>
<td>Near the apical loop of domain II</td>
<td>Not detected</td>
<td></td>
</tr>
</tbody>
</table>

Major cross-links are given in bold.
Roles of rpSA (p40) in the 40S Structure and Translation of Canonical and IRES-containing mRNAs

Notably that rpSA (mentioned as ribosomal protein p40 in [35,91] and as precursor of the laminine-binding protein in other works, e.g. [92] and references therein), one of proteins neighboring HCV IRES (see above), is not a constant ribosomal component. It is known that binding of rpSA to 40S subunits accompanies formation of polysomes during active cell growth; in contrast, dissociation of polysomes leads to a partial loss of rpSA from the 40S subunits [91]. In accordance with these, it was found that 40S subunits isolated from human placenta varied in the extent of the rpSA loss, and subunits lacking this protein could directly bind recombinant human rpSA [93]. Remarkably, affinity of HCV IRES to 40S subunits strongly depends on their rpSA content [94]. It was shown that monoclonal antibodies against C-terminal domain (CTD) of rpSA blocked HCV IRES binding to 40S subunits, but had no effect on the binding of Phe-tRNA to the poly(U)-programmed ribosomes [94]. To study effect of the antibodies on the translation of various mRNAs in a cell-free system, two mRNAs were used: one carried 5′-UTR of actin mRNA, another contained HCV IRES. It was found that binding of the antibodies to their epitopes in 40S subunits inhibited mRNA translation independently on the nature of its 5′-UTR, but the effect was larger with the IRES-containing mRNA [94]. The results obtained with mRNA-containing 5′-UTR of actin mRNA suggest that rpSA has a regulatory function in the process of translation of cellular mRNAs. The protein, when bound to the 40S subunit, could enhance or coordinate its interaction with initiation factor eIF3, whose binding site is close to site of the conserved part of rpSA (Fig. 11), and thus activate translation initiation. In contrast, dissociation of rpSA could lead to decrease of the eIF3 affinity to 40S subunits, resulting in the reduction of the total translation efficacy. This implies that the conserved part of rpSA on the 40S subunit is involved in the interaction with the factor that, in turn, is likely implicated in binding with rpS26e (see above). All these are in a good accordance with close locations of rpS26e and the conserved part of rpSA on the 40S subunit [19] (Fig. 11).

Multiple sequence alignment of ribosomal proteins belonging to the rpS2p family that includes rpSA revealed that the eukaryotic proteins had C-terminal extensions, whose lengths and sequences are strongly variable; higher eukaryotic proteins have longer C-terminal tails [95]. In the model of the lower eukaryotic ribosomal 40S subunit [19], the CTD of rpS0 (a counterpart of higher eukaryotic rpSA) is located closer to other proteins of the HCV IRES binding site than the main conserved part of the protein (Fig. 12). Likely, much longer CTD of human rpSA comes just to the HCV IRES binding site, and it is reasonable to suggest that it is involved in the IRES binding while the conserved part of rpSA apparently takes part in the regulation of eIF3-dependent initiation of cellular mRNAs translation (see above). With the use of a set of truncated forms of the recombinant human rpSA, it has been found that CTD of rpSA is responsible for binding of the protein to the 40S subunit, and fragment 236–262 is the key player in this process [95]. Hydroxyl-radical footprinting showed that helix 40 of the 18S rRNA is involved in the CTD binding in the 40S subunit [95]. Remarkably, the CTD of rpSA has another function operating as a receptor in the binding of extracellular ligands (envelope proteins of several viruses and prion proteins) by Laminine receptor, which is a dimer of rpSA. However, this process involves CTD regions that do not overlap with the region responsible for the binding to 40S subunit [92].

Arrangement of tRNA Acceptor End on the Mammalian Ribosome

First, the positioning of 3′-terminal of tRNA in the 80S ribosomal A and P sites was studied using a tRNA^Asp analog that bore s^4U attached to the 3′-terminal adenosine (s^4U77–tRNA^Asp) [39]. This analog cross-linked exclusively to the 28S rRNA when bound at the A or the P site. All cross-linked 28S rRNA nucleotides were found in the fragment 4302–4540 (Table 4) belonging to the highly conserved domain V that in prokaryotic ribosomes is involved in the PTC formation [1,9,11,96]. However, the direct comparison of cross-linking results obtained with the use of s^4U77–tRNA^Asp with available data on the respective prokaryotic ribosomal complexes would be incorrect since this tRNA analog contained an additional nucleotide at the 3′-terminus. The environment of the s^4U77 on the 50S subunit was predicted [39] by means of a molecular modeling approach using the available X-ray structural data on 50S subunits complexed with a substrate mimicking 3′-terminal part of tRNA [97] and taking into account probable thermal fluctuations of the CCA end. These made it possible to obtain information concerning nucleotides of 23S rRNA that could appear within 6–7 Å of s^4U77 and could be candidate targets for cross-linking. Few of them
were indeed found to cross-link to the 28S rRNA (Table 4). It is also essential to note that, according to molecular modelling data, two of the cross-linked 28S rRNA nucleotides determined [39], U4502 and especially U4469 (2596 and 2563, respectively, in E. coli 23S rRNA), were located too far from s4U77 to be cross-linked. Therefore, the cross-linking data suggest that the mentioned 28S rRNA nucleotides (at least U4469) are located much closer to s4U77 than the corresponding 23S rRNA nucleotides in the model prokaryotic complex used for the calculations. Dissimilarities between the set of the cross-linked 28S rRNA nucleotides and the set of candidate targets may result from allosteric effects of ribosomal proteins interacting with domain V on the mutual arrangement of the CCA end of tRNA and the 28S rRNA even if these proteins are located far from the PTC region in the mammalian complex. Such effects of ribosomal proteins on the peptidyl transferase activity of yeast ribosomes were discussed in detail [98].

The data obtained with s4U77–tRNAAsp somewhat modified the generally accepted idea that the structure of the PTC region is the same in all kingdoms based on the strong conservation of the secondary structure and a substantial part of the primary structure of domain V of 23S-like rRNAs. Thus, the arrangements of the prokaryotic and the mammalian PTC regions do share significant similarity, but also have pronounced differences that seem to relate to the more complex structure of the mammalian ribosome and different dynamic properties of ribosomes from various kingdoms.

Data complementary to the results obtained with s4U77–tRNAAsp were reported with tRNAAsp analog whose 3’-terminal ribose was oxidized (tRNAox) [40]. The application of this analog targeted to proteins made it possible to find that 3’-terminus of tRNA bound at the 80S ribosomal P site was located closely to 60S rpL36a-like (rpL36AL). Thus, for the first time it has been demonstrated that a ribosomal protein is located at a zero distance to the CCA 3’-terminal adenosine of P site-bound tRNA.

Table 4 Summary of s4U77–tRNAAsp cross-links with the 28S rRNA in human 80S ribosomal complexes [39]

<table>
<thead>
<tr>
<th>P site</th>
<th>Human 28S rRNA numbering</th>
<th>Deinococcus radiodurans numbering</th>
<th>E. coli numbering</th>
<th>Human 28S rRNA numbering</th>
<th>D. radiodurans numbering</th>
<th>E. coli numbering</th>
</tr>
</thead>
<tbody>
<tr>
<td>U4461</td>
<td>2534</td>
<td>2555</td>
<td></td>
<td>U4461</td>
<td>2535</td>
<td>2555</td>
</tr>
<tr>
<td>U4462</td>
<td>2535</td>
<td>2556</td>
<td></td>
<td>U4462</td>
<td>2535</td>
<td>2556</td>
</tr>
<tr>
<td>U4502</td>
<td>2575</td>
<td>2596</td>
<td></td>
<td>C4507</td>
<td>2580</td>
<td>2601</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>U4469</td>
<td>2542</td>
<td>2563</td>
</tr>
</tbody>
</table>

Major cross-links are given in bold.

RpL36AL, which is strongly conserved in eukaryotes, belongs to the L44e family of ribosomal proteins, a representative of which is rpL44e in Haloarcula marismortui. These results are not in agreement with the structure of the archaeal 50S ribosomal subunit lacking protein electron density in a radius of ~18Å of the PTC, where L44e is located at the E site apart from the P site [99]. The mentioned structure is a basis of the current view that the PTC works as a ribozyme and that ribosomal proteins are not involved in catalysis of peptide bond formation. The disagreements between cross-linking results with mammalian 80S ribosomes and X-ray data on prokaryotic ribosomal complexes were discussed previously [40], and one of the most probable explanations presented there concerned well-known flexibility of the CCA end of tRNA. This flexibility could allow 3’-end of the tRNAox positioned at the P site to migrate from the P- to the E site where it could interact with nearby proteins.

**Conclusions**

Application of a site-directed cross-linking approach made it possible for the first time to reveal features of the structural organization of functional centers of the human ribosome and to evaluate extent of their similarity with the respective sites of prokaryotic ribosomes. It was found that the conserved core of small subunit rRNA almost identical in all kingdoms is involved in the structural organization of 80S ribosomal binding sites of mRNA codons that interact with tRNA anticodons at the A and the P sites, while ribosomal proteins surround primarily mRNA parts upstream and downstream of these codons. Moreover, it turns out that the protein environment of mRNA on the human ribosome has a number of features that have no homology in the eubacterial ribosome. In particular, eukaryote-specific and eukaryote/archaea-specific peptides of ribosomal proteins are key players in accommodation of mRNA region 5’ of the codons interacting with tRNAs and in the formation of the decoding site, respectively. The application of
short mRNA analogs bearing cross-linkers at the designed positions goes beyond studying the structure of the 80S ribosomal mRNA binding site. So, recently these analogs were successfully used to determine peptides of polypeptide chain release factor eRF1 neighboring stop codons in human 80S termination complexes [50,51] and the data obtained provide a new insight into the mechanism of stop signal decoding [51] that remained largely unknown due to the lack of cryo-EM and X-ray crystallographic data on 80S termination complexes. As for the PTC that is generally assumed to be strongly conserved and composed only of the rRNA, features distinguishing it from the respective prokaryotic center were also detected; the most essential one is the involvement of a ribosomal protein in the accommodation of 3′-terminal nucleotide of tRNA at the P site. Finally, application of a site-directed cross-linking made it possible to determine 40S ribosomal components implicated in the formation of the hepatitis C IRES binding site, a functional center that is specific to the mammalian ribosome. Involvement of ribosomal proteins but not rRNA in its formation implies participation of eukaryote-specific proteins or peptides in the HCV IRES binding. Subsequent studies on the contributions of rpSA into the ribosomal structure and the IRES binding shed light on the particular roles of the CTD of rpSA specific to higher eukaryotes in binding of the protein to the 40S subunit and in binding of the subunit with the HCV IRES. In general, data obtained by a site-directed cross-linking approach are in a good agreement with available X-ray and cryo-EM models. However, the detailed comparison is difficult because complexes of higher eukaryotic ribosomes with mRNA, tRNA, and HCV IRES are not studied so far by X-ray crystallography. As for cryo-EM, today this approach does not allow particular mRNA codons, CCA ends of tRNAs, and eukaryote-specific peptides of proteins to be directly visualized. Evidently, cross-linking data provide a biochemical basis for future creating of atomic models of higher eukaryotic ribosomes and their functional sites with application of X-ray and cryo-EM approaches, and for further studying in detail molecular mechanisms of the work of human translational machinery operating with cellular and viral mRNAs. The next frontier of the research is to understand functional roles of specific structural elements of higher eukaryotic ribosomal functional sites in translation and its regulation.

Acknowledgement

We gratefully thank Ian Eperon for critical reading of this manuscript.

Funding

This work was supported by the grants from Russian Foundation for Basic Research (including the current grant 11-04-00597 to G.K.) and the Program ‘Molecular and Cell Biology’ of the Presidium of the Russian Academy of Sciences (to G.K.).

References


Structural and functional topography of the human ribosome


