

Positioning of subdomain III_d and apical loop of domain II of the hepatitis C IRES on the human 40S ribosome

Elena Babaylova¹, Dmitri Graifer¹, Alexey Malygin¹, Joachim Stahl²,
Ivan Shatsky³ and Galina Karpova^{1,*}

¹Institute of Chemical Biology and Fundamental Medicine, Siberian Branch of the Russian Academy of Sciences, Novosibirsk, 630090, Russia, ²Max-Delbrück-Centrum für Molekulare Medizin (MDC), Berlin-Buch, 13092 Berlin, Germany and ³Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow, 119899, Russia

Received September 11, 2008; Revised November 19, 2008; Accepted December 9, 2008

ABSTRACT

The 5'-untranslated region of the hepatitis C virus (HCV) RNA contains a highly structured motif called IRES (Internal Ribosome Entry Site) responsible for the cap-independent initiation of the viral RNA translation. At first, the IRES binds to the 40S subunit without any initiation factors so that the initiation AUG codon falls into the P site. Here using an original site-directed cross-linking strategy, we identified 40S subunit components neighboring subdomain III_d, which is critical for HCV IRES binding to the subunit, and apical loop of domain II, which was suggested to contact the 40S subunit from data on cryo-electron microscopy of ribosomal complexes containing the HCV IRES. HCV IRES derivatives that bear a photoactivatable group at nucleotide A275 or at G263 in subdomain III_d cross-link to ribosomal proteins S3a, S14 and S16, and HCV IRES derivatized at the C83 in the apex of domain II cross-link to proteins S14 and S16.

INTRODUCTION

Hepatitis C virus (HCV) is one of the most dangerous human pathogens. It infects about 180 million people worldwide (1) and can lead to cirrhosis, hepatocellular carcinoma and other damages of the liver (2,3). HCV is a flavivirus whose genome consists of one positive 9600 nucleotide-long RNA strand lacking a cap at the 5'-terminus (4). One of the key steps of the HCV life cycle in human host cells is initiation of viral RNA translation. It is mediated by a specific highly structured motif comprising nucleotides 40–372 of the viral RNA;

it includes the 5' untranslated region and the adjacent part of the coding sequence, and is known as IRES (Internal Ribosomal Entry Site) (5–7) (Figure 1). The HCV IRES is responsible for cap-independent internal initiation of translation; it is an alternative to the conventional mechanism generally used by capped eukaryotic cellular mRNAs involving a number of initiation factors that provide association of the mRNA to the cap, leading 40S ribosome binding and scanning to the initiation codon (6). The organization of domains and several RNA motifs in the domains are conserved among different strains of HCV and related viruses of the *Flaviridae* family (6,8,9). The HCV IRES folds under physiological ionic conditions into a defined 3D structure whose integrity is essential for efficient IRES-mediated translation (10); spatial structures of the HCV IRES domains have already been resolved by X-ray crystallography and NMR (11).

The sequence of events leading to the formation of 80S initiation complex with IRES competent for the start of viral RNA translation is well known (12–14). The functional roles of individual IRES domains and subdomains have been studied using mutant IRESes (6,12–16), cross-linking with chemically reactive derivatives of oligonucleotides complementary to target RNA sequences (17), and enzymatic and chemical footprinting (15,16,18). According to the generally accepted model, in the first step of the assembly of the 80S initiation complex, the HCV IRES forms a stable binary complex with the 40S subunit where the initiation codon is positioned on the P site without any initiation factors. Binary complex formation crucially depends on the basal portion of domain III (mainly hairpins III_d and III_e) and on the four-way junction III_{abc} (Figure 1). Thereafter, eIF3 and the ternary complex (eIF2•Met-tRNA_i^{Met}•GTP) assemble to form a 48S preinitiation complex; binding of eIF3 is mediated by domain III_b and four-way junction III_{abc}.

*To whom correspondence should be addressed. Tel: +7 383 335 62 29; Fax: +7 383 333 36 77; Email: karpova@niboch.nsc.ru

target for cross-linking (19). Other investigators using HCV IRES with uridines randomly substituted by 4-thiouridines reported cross-linked proteins S2, S3, S10, S15, S16/S18 and S27 (20). These 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. Based on the cross-linking results and the data on inability of 40S subunits from lower eukaryotes to bind HCV IRES, it was suggested that the IRES-binding site on the 40S subunit is formed by protein sequences specific of higher eukaryotes rather than by the 18S rRNA that is fairly similar between yeast and human (78% identity) [(20) and refs therein].

Cryo-electron microscopic (cryo-EM) visualization of HCV IRES in its complexes with the 40S subunit (21) and the 80S ribosome (22) provided a general idea of the positioning of the IRES on the ribosome; two ribosomal proteins, namely, S5 on the head of the subunit interface and S14 on the platform were mapped close to the apex of domain II of the IRES (22). However, domain II in contrast to subdomains IIIId/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.

Highly sensitive tools to investigate the molecular environments of RNA ligands bound to the ribosome are RNA derivatives bearing a perfluorophenyl azide cross-linker at specific locations. Using a set of short mRNA analogues bearing such groups at the specific locations, the structure of the mRNA-binding channel of the human 80S ribosome has been studied in detail (23–26). The data on ribosomal components neighboring mRNA nucleotides in positions –9 to +12 with respect to the first nucleotide of the P site codon were recently confirmed in a study of 48S initiation complexes using mRNAs bearing 4-thiouridines capable of forming ‘zero-length’ cross-links (27). This additionally demonstrated that perfluorophenyl azide-modified RNAs provide correct and adequate information on the structures of binding sites of various RNA ligands bound to the mammalian ribosome. Recently, we suggested a novel strategy that makes it possible to selective introduce perfluorophenyl azide groups into long structured RNAs (28). This strategy is 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. After hydrolysis of the phosphoramidate bond in the covalent adduct formed, aryl azide moiety is selectively introduced at the liberated benzylamine group. Using HCV IRES derivative bearing the cross-linker at A296 in hairpin IIIe for a site-directed cross-linking study on human ribosomes, we determined the ribosomal components neighboring this keystone subdomain in the binary complex of the IRES with the 40S subunit. This approach makes it possible to monitor the molecular environment of each of structural elements of the IRES on the ribosome at every step of translation initiation of the viral RNA.

In this study, applying the strategy based on site-specific modification, we obtained HCV IRES derivatives bearing a perfluorophenyl azide group at various nucleotides in hairpin IIIId or in the apex of domain II (Figure 1), and with the use of these derivatives determined the ribosomal components neighboring these nucleotides in the binary complex of HCV IRES with the human 40S ribosomal subunit. The results obtained show which ribosomal components neighbor the structural elements of HCV IRES critical for its binding to the 40S subunit and which components surround the apex of domain II that is located close to the ribosomal surface based on cryo-EM data.

MATERIALS AND METHODS

Acrylamide, *N,N'*-methylene-*bis*-acrylamide and urea were purchased from AppliChem, ribonucleoside triphosphates were from Sigma, Sephadex G-75 from Pharmacia and T7 RNA polymerase from NEB. DNA oligomers complementary to definite sequences of the HCV IRES were synthesized in the Laboratory of Medical Chemistry (Institute of Chemical Biology and Fundamental Medicine SB RAS), [α -³²P]GTP (1 mCi/nmol), 4-(*N*-2-chloroethyl-*N*-methylamino)benzylamine (ClRCH₂NH₂) and *N*-oxysuccinimide ester of 4-azidotetrafluorobenzoic acid were synthesized in the Laboratory of Biotechnology of this Institute.

Preparation of labeled HCV IRES RNA

The fragment corresponding to HCV RNA nucleotides 40–372 was obtained by *in vitro* transcription of plasmid pXL40-372.NS (29) linearized by BamHI (‘Fermentas’, Lietuva) as described (17) using [α -³²P]GTP. It was isolated and examined as described (28). The specific radioactivity of the RNA transcript was typically about 25 000 cpm/pmol.

Introduction of perfluorophenyl azide groups into the IRES

4-(*N*-2-chloroethyl-*N*-methylamino)benzyl-5'-phosphoramides of oligodeoxyribonucleotides were synthesized and purified as described (17). Site-specific modification of the HCV IRES was carried out as described (28) starting with 300 pmol of the RNA transcript. Isolation of the RNA modified with deoxy-oligomers derivatives, determination of the modified RNA nucleotides, hydrolysis of the phosphoramidate bond in the deoxyribonucleotide derivatives attached to the RNA, selective introduction of perfluorophenyl azides at the aliphatic amine groups liberated after hydrolysis and purification of the photoactivatable derivatives of HCV IRES were carried out as described (28).

Ribosomes, ribosomal complexes and cross-linking procedures

40S ribosomal subunits with intact rRNA were isolated from unfrozen human placenta as indicated (30). Prior to use, the subunits were re-activated by incubation in binding buffer A (20 mM Tris-HCl, pH 7.5, 100 mM KCl, 2.5 mM MgCl₂ and 0.25 mM spermidine) at 37°C

for 10 min. Binary complexes of HCV IRES and 40S subunits were obtained by incubating the subunits (2.0×10^{-6} M if not specified otherwise) with IRES (0.5×10^{-6} M if not specified otherwise) in buffer A at 37°C for 10 min. The extent of binding of the 5'-³²P-labeled RNAs to 40S subunits was examined by the nitrocellulose filtration assay using alkali pretreated filters as described (31). In the experiments on the determination of cross-linked ribosomal proteins, the binary complexes were formed starting from 10–20 pmol of the derivatized HCV IRES. UV-irradiation of the complexes was carried out according to (25). The distribution of the cross-linked labeled IRES between the rRNA and ribosomal proteins was examined by centrifugation in a 5–20% sucrose density gradient in the presence of SDS and EDTA as described (32).

Isolation and identification of ribosomal proteins cross-linked to HCV IRES derivatives

To isolate ribosomal proteins, the irradiated complexes were treated under conditions for dissociation of the subunits into the rRNA and proteins and exhaustively hydrolyzed with RNases A and T1 according to (28). The proteins were separated from oligoribonucleotides by gel-filtration and analyzed by 1D and 2D PAGE as described (28). To analyze cross-linked proteins by immunoprecipitation, we used rabbit antibodies against proteins S2 and S3, goat antibodies against proteins S3a and S13/S16 and sheep antibodies against protein S14 obtained as described (33). These antibodies did not cross-react with other 40S proteins. Irradiated binary complexes of the ³²P-labeled IRES derivatives and 40S subunits were supplied with 2-mercaptoethanol, SDS and EDTA to final concentrations of 0.2%, 0.1% and 10 mM, respectively, together with 100 pmol of unlabeled IRES. The mixture was incubated at 37°C for 10 min to dissociate the subunits into rRNA and proteins and then diluted with TBS (20 mM Tris-HCl, pH 7.5, 150 mM NaCl and 0.4% Triton X-100). The antibodies were immobilized on protein-G-Sepharose (Fluka) by stirring 20 µl of the sorbent suspension with 5–20 µl of a solution containing 50 mM Tris-HCl, pH 9.0, 4 M urea and 0.3 mg/ml antibodies for 2 h at 0°C. The supernatant was discarded and the sorbent was washed three times with 100 µl TBS. To prevent unspecific binding of RNAs and cross-linked proteins to protein-G-Sepharose, the washed sorbent with immobilized antibodies was supplied with 20 µl of a solution containing 1 A₂₆₀ unit of total *Escherichia coli* tRNA and bovine serum albumin (final concentration of 0.5 mg/ml); the mixture was incubated 1 h at 0°C with stirring. The supernatant was discarded and the sorbent was mixed with a solution containing cross-linked proteins and incubated for 3 h at 0°C. The sorbent with bound cross-linked proteins was pelleted by low-speed centrifugation, washed five times with TBS buffer (100 µl each time) and their radioactivity was measured by Cherenkoff counting. The values obtained with antibodies in a control experiment with unmodified IRES were subtracted from the values obtained with IRES derivatives.

RESULTS

HCV IRES derivatives

To introduce photoactivatable perfluorophenyl azide groups into hairpin IIIId or the apex of domain II of the IRES (Figure 1), an approach based on the site-specific alkylation of RNA with phosphoramidate derivatives of deoxyribonucleotides described previously (28) and schematically shown in Figure 2a was applied. To increase the yield of IRES alkylation with the derivatives of deoxyoligomers complementary to sequences 259–276 and 62–81, helper oligomers were used that facilitate binding of oligomers bearing the alkylating group to the structured RNA [see (34) and Refs. therein] (Figures 1 and 2b). Helpers were not used in the experiments with the derivative of oligomer complementary to sequence 248–267 since the yield of HCV IRES alkylation was sufficiently high. The covalent adducts resulting from IRES alkylation with the oligonucleotide derivatives were separated from unmodified RNA by denaturing PAGE on an 8% gel (the respective electrophoregrams are not shown). Generally, about 50% of the RNA was converted into the covalent adduct in the course of alkylation.

To identify modified IRES nucleotides, reverse transcription was used that makes it possible to detect modified nucleotides by stop or pause of primer extension. The modification site is generally assumed to be the nucleotide 5' of the primer extension stop site. Therefore, stops at A276 (Figure 3a, left panel, lane 1), U264 (left panel, lane 2) and C84 (right panel, lane 3) correspond to modified A275, G263 and C83, respectively (Figures 1 and 2b). With HCV IRES alkylated with the derivative of oligomer complementary to sequence 248–267, a weaker stop at G263 was also observed (Figure 3a, left panel, lane 2). This signal could not be assigned to the primer extension stop caused by the modification of U262 since uridine is unable to react with aromatic 2-chloroethyl amines at nearly neutral pH (35). Thus, the signal at nucleotide in position 263 was assigned to a pause of reverse transcription that could occur at guanine alkylated on the N7 atom, which is not involved in Watson-Crick base pairing. One could expect that in the heteroduplex of the HCV IRES with the derivative of oligomer complementary to sequence 248–267 the alkylating group is located closer to nucleotides G265–267 rather than to G263 (Figure 1). The unusual target of alkylation in this case might be due to peculiarities of the spatial structure of the IRES and also to the low reactivity of guanines in oligoG fragments towards alkylation with aromatic 2-chloroethyl amines (36). Weaker stops with the derivative of oligomer complementary to sequence 62–81 probably correspond to minor alkylation sites adjacent to C83 (taking into account the discussion above, these are G82 and A81).

To introduce a photoactivatable moiety on modified RNA nucleotides, the phosphoramidate bond in the covalent adducts obtained from ³²P-labeled IRES was hydrolyzed under mild acidic conditions (pH 4.1) that resulted in the release of the aliphatic amino group –RCH₂NH₂ linked to the RNA. The modified IRES RNAs were purified by denaturing 8% PAGE and the –RCH₂NH₂ moieties were selectively derivatized by treatment with the

Binding and cross-linking of photoactivatable HCV IRES derivatives to 40S subunits

Filter retention assays revealed that the affinities of HCV IRES derivatives bearing the cross-linker at A275, G263 or C83 to 40S subunits were similar to each other and did not significantly differ from that of unmodified IRES (Figure 3b) as was shown previously with HCV IRES similarly derivatized at the nucleotides A296 and G87 (28). For the cross-linking experiments, saturating concentrations of 40S subunits were used. To obtain the cross-links, the binary complexes of the IRES derivatives and the 40S subunits were irradiated with mild UV light (>280 nm). Subsequent analyses of the distribution of the radioactive label between the rRNA and the ribosomal proteins after centrifugation in sucrose density gradients containing SDS and EDTA did not reveal radioactivity in the fractions containing 18S rRNA (the respective sedimentation profiles are not presented). Radioactivity was found only in the top fractions containing proteins and IRES not cross-linked to ribosomal components. Therefore, none of the HCV IRES derivatives cross-linked to the rRNA.

Identification of ribosomal proteins cross-linked to the photoactivatable HCV IRES derivatives

Ribosomal proteins cross-linked to derivatives of the IRES were first resolved by 1D SDS-PAGE. Prior to analysis, the RNA cross-linked to proteins was exhaustively hydrolyzed with RNases A and T1; ribosomal proteins were separated from labeled oligoribonucleotides resulting from hydrolysis by gel-filtration under strongly denaturing conditions (6 M guanidine chloride) to decrease possible effects related to unspecific binding of

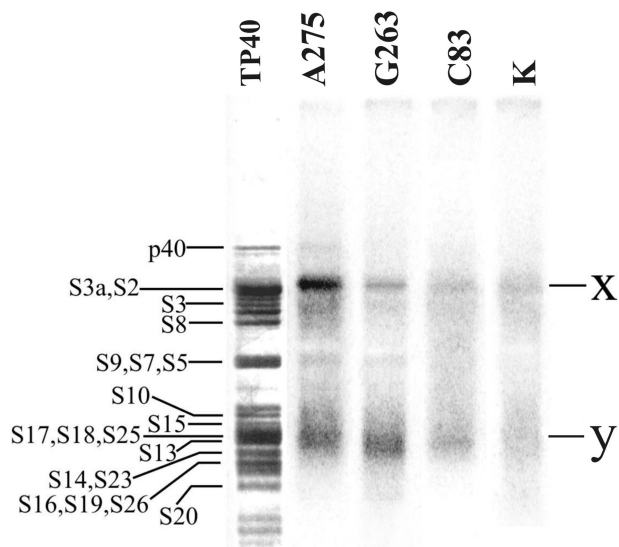


Figure 4. Autoradiogram of the 40S ribosomal proteins cross-linked to the ^{32}P -labeled derivatives of HCV IRES by 1D PAGE in the presence of SDS. Nucleotides bearing the cross-linker are indicated. Lane K, proteins isolated from an irradiated control complex of unmodified HCV IRES with 40S subunits. Lane TP40, silver stained gel; positions of the 40S ribosomal proteins are indicated (43,44).

oligoribonucleotides to the proteins. The results presented in Figure 4 show that all IRES derivatives cross-linked mainly to two protein groups corresponding to the bands 'x' in the upper and 'y' in the lower parts of the gel, the intensities of the bands depending on the location of the cross-linker in the IRES. With the IRES derivatized at C83, the yield of cross-linking was significantly lower than with the two other derivatives (Figure 4). It should be also noted that in the lane corresponding to the control RNA (lane K), weak diffuse bands appeared at positions similar to those of 'x' and 'y'. These bands clearly originated from a small portion of unspecific complexes of ribosomal proteins with labeled oligoribonucleotides resulting from hydrolysis of the IRES and remaining in the protein fraction after gel-filtration (see above), and not separated from the proteins even under conditions of SDS-PAGE.

To identify candidate cross-linked proteins, we took into account our previous data (25,26) allowing evaluation of the effect of oligoribonucleotides cross-linked to ribosomal proteins based on their electrophoretic mobilities. Moreover, exhaustive hydrolysis of HCV IRES with RNases A and T1 should produce fragments ApA^*pApGp , G^*pUp and $\text{C}^*\text{pCp/C}^*\text{p}$ (asterisks indicate modified nucleotides) cross-linked to proteins in the case of the HCV IRES derivatized at the A275, G263 and C83, respectively (Figures 1 and 2b). These fragments somewhat decrease the electrophoretic mobility of the cross-linked proteins (the greater the molecular mass of the protein, the lower the effect of the cross-linked oligomer). Considering all these factors, one could suggest that bands 'x' might correspond to cross-linked proteins S2/S3a and less probably to protein S6, and bands 'y' to cross-linked proteins S14/S16/S19/S23/S26. With the IRES derivatized at A275 and at G263, a minor band in the middle of the electrophoregram was also detected; its position with respect to proteins S9, S7 and S5 was very similar to that of cross-linked S5 in the experiments with the IRES derivatized at A296 (28). With IRES modified at A275 an additional weak band could be seen at position corresponding to a single protein p40, which was shown earlier to cross-link to A296 in subdomain IIIe (28).

Cross-linked protein(s) were more precisely identified resolving them after exhaustive hydrolysis of HCV IRES with RNases A and T1 by 'basic-SDS' 2D PAGE (separation in the first dimension was at pH 8.3, and in the second dimension at pH 6.75 in the presence of SDS), the results are presented in Figure 5. Cross-linked proteins were identified considering that in the electrophoretic system used, oligoribonucleotide fragments cross-linked to proteins decrease their mobilities in the first dimension but barely affect their migration in the second dimension. Consequently, the radioactive spot of the cross-linked protein should be somewhat shifted to the left of the corresponding stained spot of the unmodified protein. The extents of the shifts were evaluated based on data of our previous studies on cross-linking of short mRNA analogues to ribosomes (25,26). The results showed that HCV IRES derivatized at the A275 and the G263 cross-linked to the same three groups of proteins, namely, to S3/S3a (less probably, S2), S10/S15/S14 and S13/S16 (much less

probably, S25), the distribution of cross-links between these groups depended on the location of the cross-linker in the HCV IRES (Figure 5a and b). With C83-modified IRES, only two spots corresponding to cross-linked proteins S10/S15/S14 and S13/S16 were detected (Figure 5c). Given the results of the 1D gels (Figure 4), more exact identification of the cross-linked proteins could be performed. In the group S13/S16/S25, only protein S16 was actually cross-linked since S13 and S25 were not among the candidate cross-linked proteins (see above). For the same reason, proteins S10 and S15 could be excluded from the group S10/S15/S14 and the only cross-linked protein

was S14. Similarly, protein S3a could be identified as the only cross-linked protein in the group S3/S3a. Thus, combining the data of 1D and 2D PAGE made it possible to conclude that all HCV IRES derivatives used here cross-linked mainly to proteins S14 and S16, and the derivatives bearing the modifying group at A275 or G263 also cross-linked to protein S3a (Figure 5d).

The identity of the cross-linked proteins S3a, S14 and S16 was confirmed by immunoprecipitation using specific antibodies. In parallel, immunoprecipitation with antibodies against proteins S2 and S3 whose modification was practically excluded on the basis of the comparison of the results of 1D and 2D PAGE analysis was carried out as control. The results shown in Figure 6 are in a good accordance with the data of 1D and 2D PAGE separations discussed above (Figure 5) and confirm cross-linking of the HCV IRES derivatives to ribosomal proteins S3a, S14 and S16.

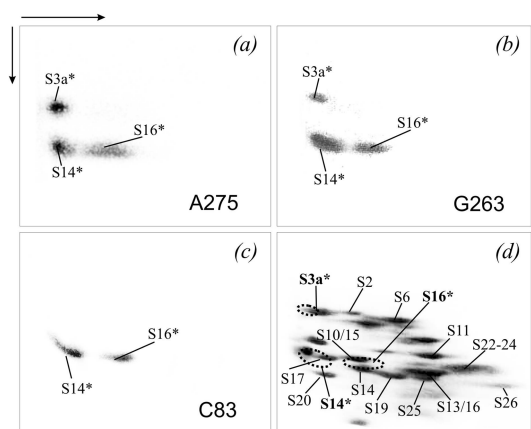


Figure 5. Analysis by 2D PAGE of proteins cross-linked to ³²P-labeled HCV IRES derivatives in the binary complexes with the 40S subunits. (a–c), The autoradiograms correspond to the experiments with HCV IRES derivatives containing the cross-linker at nucleotide A275, G263 or C83 (marked in the respective panels). (d), Coomassie stained gel corresponding to (b) (as example); the positions of the proteins are indicated (44, 45). The locations of the radioactive spots corresponding to the cross-linked proteins are indicated on the stained gel by dotted lines. The cross-linked proteins are highlighted by an asterisk (*), and are also in bold in (d).

DISCUSSION

Choice of sequences within the HCV IRES for the site-specific introduction of cross-linker and binding properties of the derivatized IRES

The criteria for the choice of derivatives of oligonucleotides for introduction of cross-linker into the HCV IRES were based on our earlier data on alkylation of the IRES with the derivatives of oligodeoxyribonucleotides complementary to its specific sequences and testing abilities of the resulting covalent adducts to bind to the 40S subunits (17). The oligos were 20-mers that was enough to enable them to form relatively stable heteroduplexes specifically with the target sequences, and contained an alkylating group at the 5'-termini. This group generally modifies RNA nucleotides near the 5'-end of the oligomer bound to the RNA, the more accessible is target RNA sequence, the higher is yield of the covalent adduct (34). To increase the

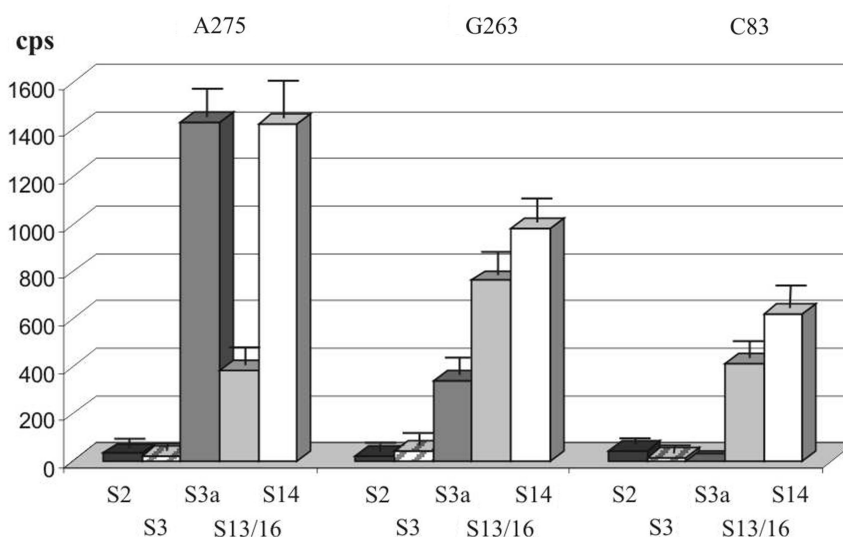


Figure 6. Analysis of 40S proteins cross-linked to ³²P-labeled HCV IRES derivatives (nucleotides bearing the cross-linker are indicated at the top) by immunoprecipitation using antibodies against mammalian 40S ribosomal proteins (indicated at the bottom).

yields of the covalent adducts, we used helper oligomers to disrupt the IRES secondary structure in the target region and thus to facilitate formation of the heteroduplexes with alkylating derivatives. We found that affinity of 40S subunits for the HCV IRES covalent adducts bearing deoxy oligomers complementary to sequences in hairpin III_d and III_e was drastically lower than that for the unmodified IRES (17). In contrast, attachment of a deoxy oligomer complementary to the apical loop of domain II barely affected the affinity of the IRES for the 40S subunits (17). These results confirmed earlier data (12–16,18) that HCV IRES subdomains III_d and III_e are the main determinants for binding to 40S subunits and that domain II does not contribute to the binding. Based on all these, we chose target RNA sequences for site-specific introduction of the cross-linker (Figure 1). Two of them were in hairpin III_d [hairpin III_e has been studied in our previous report (28)] and third one was mainly in the stem of domain II to introduce cross-linker to the middle of the apical loop [since IRES derivative with a photoactivatable group on the G87 used in our previous study (28) did not cross-link to the 40S subunit].

Binding experiments showed that introduction of an aryl azide cross-linker at nucleotides C83, G263 and A275 almost did not affect affinity of the HCV IRES to the 40S subunits (Figure 3b). This implies that HCV IRES modified at these nucleotides retained its functional activity in the binding to the 40S subunit. This conclusion could be done on the IRES modified at nucleotides G263 or A275 in subdomain III_d critical for the binding to the 40S subunit as well as on the IRES modified at the C83 in domain II, which is not essential for the binding since apical hairpins of domain II and subdomain III_d adopt similar folds (37). Besides, apical hairpin of domain II does not interact with other subdomains of the IRES (37) so modification in the loop should not affect the IRES spatial structure.

Ribosomal components neighboring the HCV IRES on the 40S subunit

Two proteins cross-linked to HCV IRES derivatives, S3a and S16, were found previously to also cross-link using the derivative bearing perfluorophenyl azide at A296 in subdomain III_e (28). Besides, this latter derivative as mentioned above, cross-linked to proteins S5 and p40 but much more strongly than the HCV IRES derivatives used here that bore the modifying group in subdomain III_d suggesting that these proteins are closer to hairpin III_e than to III_d. On the other hand, cross-linking to protein S14 was not found with HCV IRES bearing the modifying group in hairpin III_e (28), which implies that this protein likely neighbors subdomain III_d and the apical loop of domain II but not subdomain III_e. We failed to find any cross-links of HCV IRES derivatized at the G87 (in the upper part of the stem of domain II) with the 40S subunit in the binary complex (28), in contrast to the derivative bearing the same cross-linker at C83 (in the apical loop of domain II) used here (Figure 1). This is a notable result since a distance between the C83 and the G87 is within the length of the cross-linker (14 Å) used in

this study (28). Therefore, we have a basis to suggest that a distance between the C83 and the 40S subunit does fall within the range of cross-linker action while the distance from the G87 to the subunit does not. This means that the apical loop of domain II in the binary complex is located closer to the 40S subunit than the stem of this domain.

The results obtained in the present study and in our previous report (28) significantly differ from data of two earlier reports on the determination of the 40S subunit components neighboring the IRES as a whole during initiation of the viral RNA translation discussed in the Introduction (19,20). In the article by Otto *et al.* (20), proteins S2, S3, S10, S15, S16/S18 and S27 were reported to be cross-linked, of which S3, S2 and S27 were the major targets. It is now clear that a significant part of this set of proteins, namely, proteins S2, S3 and S15, belongs to the environment of the mRNA at the decoding site and downstream of it. This is based on data on cross-linking of these proteins in 80S ribosomal complexes to unstructured mRNA analogues with perfluorophenyl azide-modified nucleotides in positions +4 to +12 (25,26) or with 4-thiouridines in positions +4 to +7 (38) with respect to the first nucleotide of the P site codon. Recently these results were confirmed with 4-thiouridine-containing mRNAs in the 48S/80S initiation complexes, in which the modified nucleotide in position +11 cross-linked to proteins S3 and S2, and in positions +4/+5 cross-linked to protein S15 (27). On the cryo-EM images of the 40S subunit (39), locations of proteins S2 and S3 are similar to those of their prokaryotic counterparts S5p and S3p in the 30S subunit that form the mRNA entry site, and S15 lies on the intersubunit side of the small subunit. All these point to unlikely neighborhood of the mentioned proteins and the 5'-untranslated part of HCV IRES. Thus, we have all reasons to suggest that proteins S2, S3 and S15 found in the report (20) cross-linked not to the 5'-untranslated region of the HCV IRES but to the part of the coding sequence domain IV according to the structure presented in (8) that could contain 4-thiouridines in positions +2, +10 and +13. Consequently, only proteins S10, S16/S18 and S27 can be considered as cross-linked to the 5'-untranslated region of the HCV RNA in the binary complex. This set overlaps the set found here only by protein S16. One possible explanation for this discrepancy is that proteins S10 and S27 could be cross-linked to 4-thiouridines located in IRES positions far from C83, G263, A275 or A296 that bore perfluorophenyl azide cross-linkers. As for the data presented in (19), they cannot be compared directly with our results because in the former experiments, mixtures of HCV IRES and HeLa cytoplasmic extracts but not the purified binary complexes were used for irradiation. This did not make it possible to specify which steps of HCV RNA translation were monitored.

Our results reported here and previously (28) provide a view of the molecular environment of subdomains III_d/e, the main determinants for HCV IRES binding to the 40S subunit, and of the apical loop of domain II, which was suggested to locate close to the head of the subunit near protein S5 and the mRNA exit site on the basis of cryo-EM data (22). According to our results, protein S14

neighbors both sides of the stem of subdomain III_d and the apical loop of domain II, protein S16 is located close to both hairpins III_d/e and to the apical loop of domain II, protein S3a neighbors hairpins III_d/e, and proteins S5 and p40 are mainly close to subdomain III_e. Remarkably, no cross-links to the 18S rRNA were detected confirming the earlier suggestion that ribosomal proteins but not the 18S rRNA play the main role in binding of the HCV IRES to the 40S subunit (20).

The sets of ribosomal proteins cross-linked to perfluorophenyl azide-derivatized HCV IRES partially overlapped the sets of proteins cross-linked in the 80S ribosomal complexes to mRNA analogues that bore similar cross-linkers at nucleotides in positions -9 to $+12$ (25,26). Only protein S14 cross-linked both to HCV IRES derivatives and to mRNA analogues bearing cross-linkers in positions -6 to -9 (25). Protein S14 was also cross-linked to mRNAs bearing 4-thiouridine in the mentioned positions; besides, protein S5 was cross-linked to 4-thiouridines in positions $-3/-4$ (27). Proteins S5 and S14 are homologous to prokaryotic S7p and S11p, respectively that are known as constituents of the 30S ribosomal exit site of the mRNA (40). Thus, the HCV IRES site has minor overlap with the mRNA-binding channel and is located near the mRNA exit site.

Location of HCV IRES on the 40S subunit

We applied our cross-linking results to the cryo-EM model of the mammalian 40S ribosomal subunit (39) (Figure 7a) and compared the location of the IRES cross-linked to the proteins with its position on the 40S subunit in the binary complex deduced from cryo-EM data (21) (Figure 7b). Cryo-EM data on the 80S ribosomal complex with HCV IRES (22) were not considered since this complex reflects other steps of initiation of viral RNA translation. It should be noted that comparison of our results with cryo-EM-derived data has its limitations since only positions of the conserved parts of ribosomal proteins could be mapped on the models derived from cryo-EM data, while the IRES likely interacts with higher eukaryote-specific sequences of the proteins (20) (mentioned in the Introduction section). Nevertheless, the cross-linking results as a whole are in good agreement with the data obtained by cryo-EM reconstitution of the binary complex of HCV IRES with the 40S subunit (compare Figure 7a and b). Clearly, all the cross-linked proteins are located mainly on the solvent side of the 40S subunit opposite the beak around the cleft between the head and the body, near a site homologous to the mRNA exit site in the 30S subunit and overlapping the E site. The only pronounced dissimilarity between our cross-linking results and the cryo-EM data is that protein S5 appears to neighbor subdomain III_e (28) rather than domain II (21). This in fact is not a real discrepancy because the absence of cross-link cannot indicate lack of interaction [for a detailed discussion, see (28)], and our results do not exclude possible interaction of protein S5 with domain II of HCV IRES.

Protein S3a neighboring the HCV IRES according to our data has no prokaryotic counterparts, and therefore

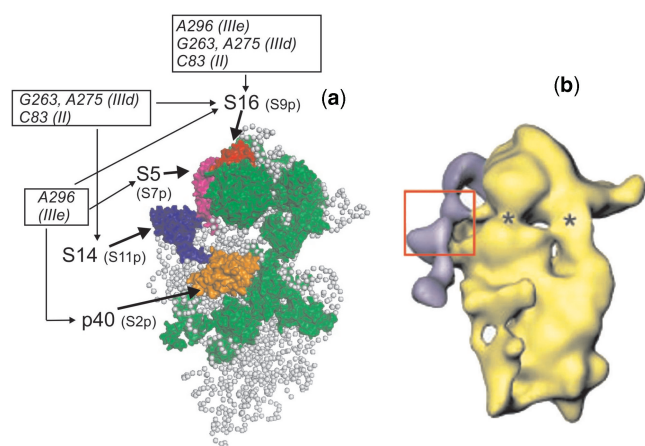


Figure 7. Structural model of the mammalian 40S ribosomal subunit (a) and of the complex of the 40S subunit with HCV IRES (b) obtained from cryo-EM data. View from the solvent side. (a) Model of the 40S subunit of the elongating ribosome [adapted from (39), PDB accession number 2ZKQ]. Locations of ribosomal proteins cross-linked to HCV IRES nucleotides belonging to domain II and subdomains III_d (this study) and III_e (17) are shown (p40 in orange, S5 in purple, S14 in dark blue and S16 in red), prokaryotic homologues of these proteins are indicated in brackets. All other proteins that have prokaryotic counterparts are shown in green and proteins specific for eukaryotes are not shown; gray balls represent the 18S rRNA. (b) Surface representation of the 40S ribosomal subunit in complex with the HCV IRES [adapted from (21)]. The cryo-EM map is in yellow. The difference map corresponding to the HCV IRES is superimposed and presented in purple. The suggested location of hairpins III_d/e is boxed. Asterisks show sites of conformational changes in the 40S subunit caused by HCV IRES binding.

its position on the eukaryotic 40S subunit remains unknown. It is conceivable that this protein is located in the same region of the subunit as p40, S5, S14 and S16, namely on the solvent side of the head far from the beak. On the cryo-EM map of the mammalian 40S ribosomal subunit, density corresponding to eukaryote-specific proteins appears close to proteins S14, S5 and p40 (39). A protein density located at the platform near proteins S14 and p40 most likely conforms position of S3a since earlier immunoelectron microscopy studies showed that antibodies against S3a bound at the mentioned region (41).

Thus, our cross-linking results provide significant novel information on the organization of the HCV IRES-binding site on the 40S ribosomal subunit in the binary complex, which forms during the first step of translation initiation of the viral RNA, highlighting the key role played by the ribosomal proteins p40, S3a, S5, S14 and S16 in the organization of this site. The binding sites of various IRESes on the 40S subunit could be significantly different. For instance, in contrast to the HCV IRES, Cricket paralysis virus (CrPV) IRES is located entirely in the mRNA-binding cleft, reaching into the P and A sites (42). Nevertheless, both IRES-binding sites share some extent of similarity: both CrPV IRES and HCV IRES neighbor protein S5 and moreover, the CrPV IRES is located close to S16 as seen upon comparing the models presented in (21,39,42). It is reasonable to suggest also that proteins play a key role in the formation of the 40S ribosomal-binding site not only for HCV IRES but

also for other IRESes. However, we cannot extend this suggestion to the later steps of initiation of translation of viral RNAs since in the 80S ribosomal complex with HCV IRES, close proximity of the IRES with 18S rRNA helices was detected (22).

ACKNOWLEDGEMENTS

We gratefully thank Anne-Lise Haenni for critical reading of this manuscript.

FUNDING

This work was supported by the Russian Foundation for Basic Research (grant #08-04-00508 to G.K.) and the Presidium of Russian Academy of Sciences (Program on Molecular and Cell Biology). The Open Access publication charges were waived by Oxford University Press.

Conflict of interest statement. None declared.

REFERENCES

- Brocard, M., Paulous, S., Komarova, A.V., Deveaux, V. and Kean, K.M. (2007) Evidence that PTB does not stimulate HCV IRES-driven translation. *Virus Genes*, **35**, 5–15.
- Wasley, A. and Alter, M.J. (2000) Epidemiology of hepatitis C: geographic differences and temporal trends. *Semin. Liver Dis.*, **20**, 1–16.
- Pawlotsky, J.M. (2004) Pathophysiology of hepatitis C virus infection and related liver disease. *Trends Microbiol.*, **12**, 96–102.
- Rosenberg, S. (2001) Recent advances in the molecular biology of hepatitis C virus. *J. Mol. Biol.*, **313**, 451–464.
- Tsukiyama-Kohara, K., Iizuka, N., Kohara, M. and Nomoto, A. (1992) Internal ribosome entry site within hepatitis C virus RNA. *J. Virol.*, **66**, 1476–1483.
- Pestova, T.V., Shatsky, I.N., Fletcher, S.P., Jackson, R.J. and Hellen, C.U.T. (1998) A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev.*, **12**, 67–83.
- Hellen, C.U.T. and Sarnow, P. (2001) Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev.*, **15**, 1593–1612.
- Honda, M., Beard, M.R., Ping, L.H. and Lemon, S.M. (1999) A phylogenetically conserved stem-loop structure at the 5' border of the internal ribosome entry site of hepatitis C virus is required for cap-independent viral translation. *J. Virol.*, **73**, 1165–1174.
- Pestova, T.V. and Hellen, C.U. (1999) Internal initiation of translation of bovine viral diarrhoea virus RNA. *Virology*, **258**, 249–256.
- Kieft, J.S., Zhou, K., Jubin, R., Murray, M.G., Lau, J.Y. and Doudna, J.A. (1999) The hepatitis C virus internal ribosome entry site adopts an ion-dependent tertiary fold. *J. Mol. Biol.*, **292**, 513–529.
- Pisarev, A.V., Shirokikh, N.E. and Hellen, C.U.T. (2005) Translation initiation by factor-independent binding of eukaryotic ribosomes to internal ribosomal entry sites. *C.R. Biologies*, **328**, 589–605.
- Otto, G.A. and Puglisi, J.D. (2004) The pathway of HCV IRES-mediated translation initiation. *Cell*, **119**, 369–380.
- Ji, H., Fraser, C.S., Yu, Y., Leary, J. and Doudna, J.A. (2004) Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA. *Proc. Natl Acad. Sci. USA*, **101**, 16990–16995.
- Locker, N., Easton, L.E. and Lukavsky, P.J. (2007) HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly. *EMBO J.*, **26**, 795–805.
- Kolupaeva, V.G., Pestova, T.V. and Hellen, C.U. (2000) An enzymatic footprinting analysis of the interaction of 40S ribosomal subunits with the internal ribosomal entry site of hepatitis C virus. *J. Virol.*, **74**, 6242–6250.
- Kieft, J.S., Zhou, K., Jubin, R. and Doudna, J.A. (2001) Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA*, **7**, 194–206.
- Malygin, A.A., Graifer, D.M., Laletina, E.S., Shatsky, I.N. and Karpova, G.G. (2003) An approach to identify the functionally important RNA sites by complementary addressed modification. *Molecular Biology*, **37**, 873–879 (translated from *Molekulyarnaya Biologiya* **37**, 1027–1034).
- Lytle, J.R., Wu, L. and Robertson, H.D. (2002) Domains on the hepatitis C virus internal ribosome entry site for 40S subunit binding. *RNA*, **8**, 1045–1055.
- Fukushi, S., Okada, M., Stahl, J., Kageyama, T., Hoshino, F.B. and Katayama, K. (2001) Ribosomal protein S5 interacts with the internal ribosomal entry site of hepatitis C virus. *J. Biol. Chem.*, **276**, 20824–20826.
- Otto, G.A., Lukavsky, P.J., Lancaster, A.M., Sarnow, P. and Puglisi, J.D. (2002) Ribosomal proteins mediate the hepatitis C virus IRES-HeLa 40S interaction. *RNA*, **8**, 913–923.
- Spahn, C.M.T., Kieft, J.S., Grassucci, R.A., Penczek, P.A., Zhou, K., Doudna, J.A. and Frank, J. (2001) Hepatitis C virus IRES RNA-induced changes in the conformation of the 40S ribosomal subunit. *Science*, **291**, 1959–1962.
- Boehringer, D., Thermann, R., Ostareck-Lederer, A., Lewis, J.D. and Stark, H. (2005) Structure of the hepatitis C virus IRES bound to the human 80S ribosome: remodeling of the HCV IRES. *Structure*, **13**, 1695–1706.
- Demeshkina, N., Repkova, M., Ven'yaminova, A., Graifer, D. and Karpova, G. (2000) Nucleotides of 18S rRNA surrounding mRNA codons at the human ribosomal A, P and E sites, respectively: a cross-linking study with mRNA analogues carrying aryl azide group at either the uracil or the guanine residue. *RNA*, **6**, 1727–1736.
- Demeshkina, N., Laletina, E., Meschaninova, M., Ven'yaminova, A., Graifer, D. and Karpova, G. (2003) Positioning of mRNA codons with respect to 18S rRNA at the P and E sites of human ribosome. *Biochim. Biophys. Acta*, **1627**, 39–46.
- Graifer, D., Molotkov, M., Styazhkina, V., Demeshkina, N., Bulygin, K., Eremina, A., Ivanov, A., Laletina, E., Ven'yaminova, A. and Karpova, G. (2004) Variable and conserved elements of human ribosomes surrounding the mRNA at the decoding and upstream sites. *Nucleic Acids Res.*, **32**, 3282–3293.
- Molotkov, M.V., Graifer, D.M., Popugaeva, E.A., Bulygin, K.N., Meschaninova, M.I., Ven'yaminova, A.G. and Karpova, G.G. (2006) mRNA 3' of the A site bound codon is located close to protein S3 on the human 80S ribosome. *RNA Biology*, **3**, 122–129.
- Pisarev, A.V., Kolupaeva, V.G., Yusupov, M.M., Hellen, C.U.T. and Pestova, T.V. (2008) Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes. *EMBO J.*, **27**, 1609–1621.
- Laletina, E., Graifer, D., Malygin, A., Ivanov, A., Shatsky, I. and Karpova, G. (2006) Proteins surrounding hairpin IIIe of the hepatitis C virus internal ribosome entry site on the human 40S ribosomal subunit. *Nucleic Acids Res.*, **34**, 2027–2036.
- Reinolds, J.E., Kaminsky, A., Kettinen, H.J., Grace, K., Clarke, B.E., Carroll, A.R., Rowlands, D.J. and Jackson, R.J. (1995) Unique features of internal initiation of hepatitis C virus RNA translation. *Embo J.*, **14**, 6010–6020.
- Matasova, N.B., Myltseva, S.V., Zenkova, M.A., Graifer, D.M., Vladimirov, S.N. and Karpova, G.G. (1991) Isolation of ribosomal subunits containing intact rRNA from human placenta. Estimation of functional activity of 80S ribosomes. *Analyt. Biochem.*, **198**, 219–223.
- Graifer, D.M., Malygin, A.A., Matasova, N.B., Mundus, D.A., Zenkova, M.A. and Karpova, G.G. (1997) Studying functional significance of the sequence 980-1061 in the central domain of human 18S rRNA using complementary DNA probes. *Biochim. Biophys. Acta*, **1350**, 335–344.
- Malygin, A.A., Graifer, D.M., Bulygin, K.N., Zenkova, M.A., Yamkovoy, V.I., Stahl, J. and Karpova, G.G. (1994) Arrangement of mRNA at the decoding site of human ribosomes. 18S rRNA nucleotides and ribosomal proteins cross-linked to oligouridylylate derivatives with alkylating groups at either the 3' or the 5'-termini. *Eur. J. Biochem.*, **226**, 715–723.
- Noll, F., Theise, H. and Bielka, H. (1974) Studies on proteins of animal ribosomes. XVIII. Reaction of ribosomes and ribosomal

- proteins with antibodies against ribosomal proteins. *Acta Biol. Med. Germ.*, **33**, 547–553.
34. Bulygin, K., Malygin, A., Karpova, G. and Westermann, P. (1998) Site-specific modification of 4.5S RNA apical domain by complementary oligodeoxynucleotides carrying an alkylating group. *Eur. J. Biochem.*, **251**, 175–180.
 35. Ross, W.C.J. (1962) *Biological Alkylating Agents*, Butterworths and Co., Ltd., London, pp. 51–63.
 36. Vlassov, V.V. and Skobel'syna, L.M. (1978) Studying macrostructure of tRNA^{Phe} (*E.coli*) by chemical modifications. *Bioorgan. Khim.*, **4**, 550–561.
 37. Lukavsky, P.J., Kim, I., Otto, G.A. and Puglisi, J.D. (2003) Structure of HCV IRES domain II determined by NMR. *Nat. Struct. Biol.*, **10**, 1033–1038.
 38. Bulygin, K., Chavatte, L., Frolova, L., Karpova, G. and Favre, A. (2005) The first position of a codon placed in the A site of the human 80S ribosome contacts nucleotide C1696 of the 18S rRNA as well as proteins S2, S3, S3a, S30 and S15. *Biochemistry*, **44**, 2153–2162.
 39. Chandramouli, P., Topf, M., Menetret, J.-F., Eswar, N., Cannone, J.J., Gutell, R.R., Sali, A. and Akey, C.W. (2008) Structure of the Mammalian 80S Ribosome at 8.7 Å Resolution. *Structure*, **16**, 535–548.
 40. Yusupova, G.Zh., Yusupov, M.M., Cate, J.H.D. and Noller, H.F. (2001) The path of messenger RNA through the ribosome. *Cell*, **106**, 233–241.
 41. Lutsch, G., Stahl, J., Kargel, H.-J., Noll, F. and Bielka, H. (1990) Immunoelectron microscopic studies on the location of ribosomal proteins on the surface of the 40S ribosomal subunit from rat liver. *Eur. J. Cell. Biol.*, **51**, 140–150.
 42. Spahn, C.M., Jan, E., Mulder, A., Grassucci, R.A., Sarnow, P. and Frank, J. (2004) Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. *Cell*, **118**, 465–475.
 43. Collatz, E., Ulbrich, N., Tsurugi, K., Lightfoot, H.N., MacKinlay, W., Lin, A. and Wool, I. (1977) Isolation of eukaryotic ribosomal proteins. *J. Biol. Chem.*, **252**, 9071–9080.
 44. Malygin, A.A., Shauro, D.D. and Karpova, G.G. (2000) Proteins S7, S10, S16 and S19 of the human 40S ribosomal subunit are most resistant to dissociation by salt. *Biochim. Biophys. Acta*, **1494**, 213–216.
 45. Madjar, J.-J., Arpin, M., Buisson, M. and Reboud, J.-P. (1979) Spot position of rat liver ribosomal proteins by four different two-dimensional electrophoresis in polyacrylamide gel. *Mol. Gen. Genet.*, **171**, 121–134.