Mass spectrometric analysis of the human 40S ribosomal subunit: Native and HCV IRES-bound complexes

YONGHAO YU,[1,4] HONG JI,[1,2,3] JENNIFER A. DOUDNA,[1,2,3] AND JULIE A. LEARY[1,4]

[1]Department of Chemistry, [2]Department of Molecular and Cell Biology, and [3]Howard Hughes Medical Institute, University of California, Berkeley, California 94720, USA
[4]Genome Center, Departments of Chemistry and Molecular and Cellular Biology, University of California, Davis, California 95616, USA

(RECEIVED December 16, 2004; FINAL REVISION February 28, 2005; ACCEPTED March 9, 2005)

Abstract

Hepatitis C virus uses an internal ribosome entry site (IRES) in the viral RNA to directly recruit human 40S ribosome subunits during cap-independent translation initiation. Although IRES-mediated translation initiation is not subject to many of the regulatory mechanisms that control cap-dependent translation initiation, it is unknown whether other noncanonical protein factors are involved in this process. Thus, a global protein composition analysis of native and IRES-bound 40S ribosomal complexes has been conducted to facilitate an understanding of the IRES ribosome recruitment mechanism. A combined top-down and bottom-up mass spectrometry approach was used to identify both the proteins and their posttranslational modifications (PTMs) in the native 40S subunit and the IRES recruited translation initiation complex. Thirty-one out of a possible 32 ribosomal proteins were identified by combining top-down and bottom-up mass spectrometry techniques. Proteins were found to contain PTMs, including loss of methionine, acetylation, methylation, and disulfide bond formation. In addition to the 40S ribosomal proteins, RACK1 was consistently identified in the 40S fraction, indicating that this protein is associated with the 40S subunit. Similar methodology was then applied to the hepatitis C virus IRES-bound 40S complex. Two 40S ribosomal proteins, RS25 and RS29, were found to contain different PTMs than those in the native 40S subunit. In addition, RACK1, eukaryotic initiation factor 3 proteins and nucleolin were identified in the IRES-mediated translation initiation complex.

Keywords: protein mapping; ribosome; mass spectrometry

Macromolecular protein complexes are involved in numerous biological processes. In many cases, the composition and dynamics of these complexes differ in various physiological environments (Pugh 1996; Pestova et al. 2001; Will and Luhrmann 2001; Madison-Antenucci et al. 2002). The yeast two-hybrid method is one of the widely used methods to study protein–protein interactions within these protein mixtures. One limitation of this method is that it generally focuses on the binary interactions between one protein target and other proteins at a time (Phizicky and Fields 1995). Alternatively, the purified intact protein complexes can be fractionated using two-dimensional gel electrophoresis followed by in-gel proteolysis and protein identification. However, this method is slow, lacks reproducibility, and is difficult to automate (Shevchenko et al. 1996; Hanash 2000; Lopez et al. 2000).

Protein identification using liquid chromatography combined with mass spectrometry has become a method

Reprint requests to: Julie A. Leary, Department of Chemistry, Division of Molecular and Cellular Biology, University of California, Davis, CA 95616, USA; e-mail: jaleary@ucdavis.edu; fax: (530) 754-9658.

Article published online ahead of print. Article and publication date are at http://www.proteinscience.org/cgi/doi/10.1110/ps.041293005.

of choice to study the composition of such complexes because of the efficiency of comprehensive polypeptide analysis. In the "Shotgun," or bottom-up approach (Hunt et al. 1986; Link et al. 1999; Zhou et al. 2002; Ranish et al. 2003), the protein mixture is first digested using various proteases. The resulting peptide mixture is fractionated using liquid chromatography and subjected to MS and MS/MS. The acquired mass spectra are then used to search against translated genomic databases to identify the proteins. However, this method is sometimes limited because the peptides containing the post-translational modifications can be lost during the separation step and thus are not detected (MacCoss et al. 2002; VerBerkmoes et al. 2002). In addition to sample losses, PTMs are difficult to determine with the bottom-up approach because there are digestion limitations, issues related to efficient LC-MS/MS detection, and in some cases, the search algorithms lacking the appropriate search parameters. Alternatively, the proteins in the mixtures can be identified using a top-down method (Lee et al. 2002; Ge et al. 2003; Kelleher 2004) in which intact proteins are fractionated by liquid chromatography and identified using accurate mass measurement and/or tandem mass spectrometry. Due to the greater variation of the molecular properties of the intact proteins, it is more difficult to standardize the experimental conditions conducive to separation and mass spectrometry (Chong et al. 2001; Du et al. 2004). Thus, a comprehensive method, which combines both bottom-up and top-down approaches, ensures that the information necessary for both protein identification and determination of PTMs can be accurately determined (VerBerkmoes et al. 2002; Strader et al. 2004).

Initiation of eukaryotic protein synthesis requires the recruitment of the 40S ribosomal subunit to the initiation codon in a messenger RNA. In a cap-dependent mechanism, the protein translation begins with the interaction between the eukaryotic translation initiation factor (eIF) protein and the 7-methyl guanosine cap at the free 5′ end of the mRNA. Subsequently, the 40S subunit binds at or near the cap and scans to the first AUG codon. In contrast to the majority of the cellular mRNAs, many viral RNAs contain a highly structured sequence, called internal ribosome entry site (IRES), located in the 5′ end of the noncoding region. IRES binds directly to the 40S ribosomal subunit and initiates protein translation in the absence of many canonical eIFs (Pestova et al. 2001; Ji et al. 2004). In order to better understand this cap-independent translation initiation mechanism, the 40S complex from human HeLa cell lysate was purified using hepatitis C virus (HCV) IRES as the affinity tag and analyzed using bottom-up and top-down mass spectrometry analysis (Lu et al. 2004). This was compared with the proteins identified in the native 40S subunit. Only small differences in the composition of 40S proteins was found between native and IRES-bound complexes, suggesting that the IRES RNA does not selectively target a subset of ribosomes during translation initiation. However, the presence of nucleolin in IRES-containing complexes suggests a possible role of this nucleolar protein in IRES-mediated translation initiation. RACK1 was also unambiguously identified.

Results and Discussion

Top-down analysis of the native 40S subunit

After lyophilization, each fraction collected off line was analyzed by FTICR-MS. A typical example of the mass spectrum acquired for one of the fractions is shown in Figure 1. Half of the lyophilized fraction was resuspended in 79:20:1 AcN:H_2O:formic acid, mixed with myoglobin and analyzed by FTICR-MS. The remaining sample was saved for MS/MS analysis as needed. Two charge state envelopes were observed (Fig. 1A). After deconvolution and calibration (Fig. 1B), one of the charge state envelopes was assigned to myoglobin and the other corresponded to a protein with a molecular weight of 26,598.471 (Fig. 1C). This mass did not match any of the native 40S proteins suggesting the presence of a possible PTM. After searching the 40S ribosomal protein database, and considering various PTMs including loss of methionine, acetylation, methylation, phosphorylation and disulfide bond, this ion was found to match the RS3 protein minus a methionine residue and addition of an acetyl group (theoretical mass of 26,598.442). The mass error for this assignment is 1.06 ppm.

To localize the position of the acetylation and determine the site of methionine loss, the remaining half of the original fraction was resuspended in 100 mM NH_4HCO_3 solution, digested with trypsin, and the resulting peptide mixture analyzed by FT-ICR MS. The N-terminal peptide T1 (NH_2-MAVQISK) was absent in the digest; however, an ion at m/z 687.41 was clearly present (Fig. 2A). The mass of this ion corresponded to the acetylated form of T1, which was also missing the methionine residue. To confirm this and further localize the modified amino acid, this ion was collisionally activated using SORI. The resulting MS/MS spectrum (Fig. 2B) clearly indicated that T1 was acetylated at the newly formed N-terminal alanine. Fractions collected offline were treated in a similar manner such that the site of modification was determined.

A total of 22 ribosomal proteins (69% of all the RS proteins) were unambiguously identified using the top-down approach (Table 1A). Each protein contained a modification including loss of methionine, acetylation, methylation, and disulfide bond formation. Loss of methionine is a common PTM and was observed on most of the identified 40S proteins and always from the N terminus. Using a similar strategy as shown for RS3,

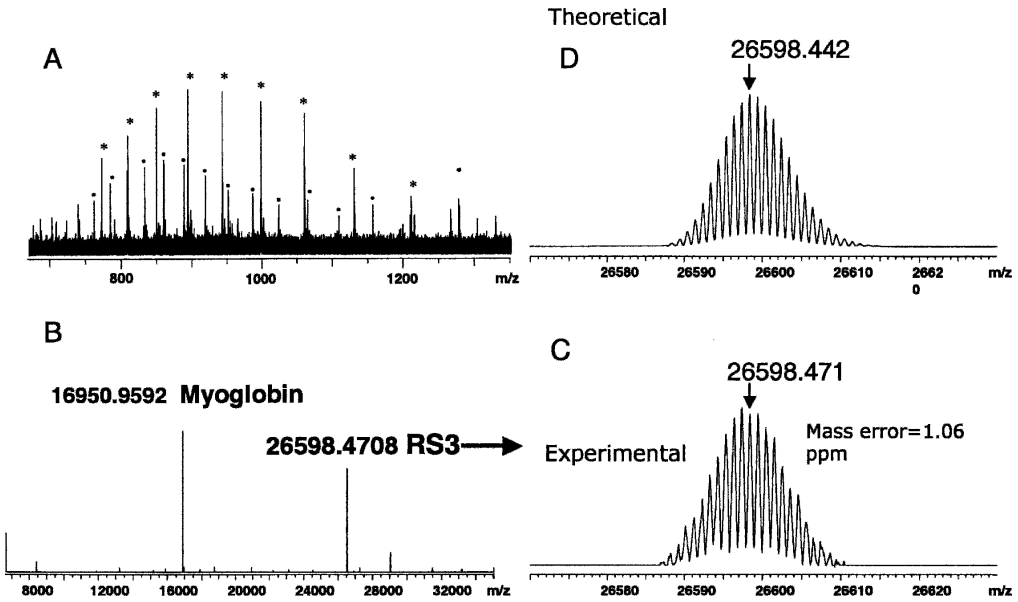

Figure 1. Top-down identification of the RS3 protein (*A*) FT-ICR MS spectra of the RPLC fraction. *, myoglobin; •, 40S ribosomal protein. (*B*) Deconvoluted mass spectrum. (*C*) Enlargement of ~26,600 Da region showing isotopic distribution. (*D*) Theoretical isotopic distribution of RS3 with loss of methionine and acetylation.

the following proteins were found to contain N-terminal acetylation: RS15, RS18, RS20, and RS28.

Although trimethylation and acetylation has a close mass difference, the yeast homologs of the MS identified acetylated ribosomal proteins have also been found to be N-terminally acetylated in N-terminal acetyltransferase mutation studies (Arnold et al. 1999; Lee et al. 2002). This strongly supported the identification of acetylation for these ribosomal proteins. However, these data are not 100% conclusive for acetylation, and thus trimethylation cannot be absolutely excluded.

In addition to acetylation, methylation was observed on three human 40S ribosomal proteins. RS10 contains one methylation site, in addition to the N-terminal

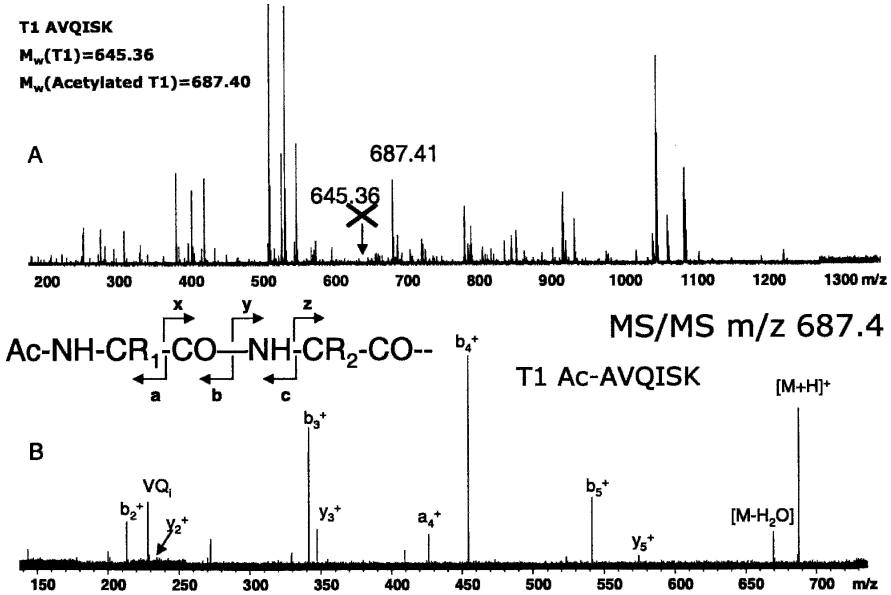

Figure 2. (*A*) MS spectrum of the trypsin digest of RS3. The observation of the ion at *m*/*z* 687.41 indicated that the T1 peptide was modified with both acetylation and loss of methionine. (The calculated MW of the ions corresponded to the protonated form.) (*B*) MS/MS spectrum of the ion at *m*/*z* 687.4. The fragment ions confirmed that the acetylation site was on the N terminus of the T1 peptide.

Table 1. *Identified proteins in the native 40S subunits (1A) and the IRES bound complex (1B)*

Name	Calculated mass	Observed mass	Δm (ppm)	PTMs
A. Proteins identified in the native 40S subunit				
RS10	18,952.9215	18,952.9738	2.8	(+Me+Ac)
RS12	14,422.4921	14,422.4932	0.1	(−Met+2Me)
RS13	17,090.6539	17,090.6553	0.1	(−Met)
RS14	16,140.5153	16,140.4440	4.4	(−Met)
RS15	16,950.1690	16,950.2536	5.0	(−Met+Ac)
RS16	16,313.0248	16,312.9994	1.6	(−Met)
RS17	15,418.3733	15,418.3531	1.3	(−Met)
RS18	17,628.8577	17,628.8398	1.0	(−Met+Ac)
RS24	15,464.4655	15,464.4521	0.9	(+Ac)
RS30	6647.8274	6647.8071	3.1	(−Met)
RS5	22,787.0589	22,787.1386	3.5	(−Met+Ac)
RS19	15,928.5187	15,928.5118	0.4	(−Met)
RS15a	14,707.9501	14,707.9392	0.7	(−Met)
RS20	13,283.2996	13,283.2977	0.1	(−Met+Ac)
RS21	9152.5917	9152.5843	0.8	(+Ac)
RS25	13,638.7116	13,638.8055	6.9	(−Met+2Me)
RS28	7882.2223	7882.2098	1.6	(+Ac)
RS29	6541.2411	6541.2263	2.3	(−Met),(S-S)x2
RS3	26,598.4427	26,598.4708	1.1	(−Met+Ac)
RS7	22,168.3046	22,168.3275	1.0	(+Ac)
RS9	22,459.5509	22,459.5431	0.3	(−Met)
RSP4	32,764.4569	32,764.5798	3.8	(−Met+Ac)
RS11[a]	—	—	—	NF
RS6[a]	—	—	—	NF
RS8[a]	—	—	—	NF
RS4[a]	—	—	—	NF
RS26[a]	—	—	—	NF
RS3a[a]	—	—	—	NF
RS2[a]	—	—	—	NF
RS23[a]	—	—	—	NF
RS27a[a]	—	—	—	NF
RACK1[a]	—	—	—	NF
B. Proteins identified in the IRES-bound complexes				
RS10	18,952.9215	18,952.9196	0.1	(+Me+Ac)
RS12[a]	—	—	—	NF
RS13	17,090.6539	17,090.6009	3.1	(−Met)
RS14	16,140.5153	16,140.5931	4.8	(−Met)
RS15[a]	—	—	—	NF
RS16	16,313.0248	16,313.0805	3.4	(−Met)
RS17	15,418.3733	15,418.3843	0.7	(−Met)
RS18	17,628.8577	17,628.7590	5.6	(−Met+Ac)
RS24	15,464.4655	15,464.4695	0.3	(+Ac)
RS30[a]	—	—	—	NF
RS5[a]	—	—	—	NF
RS19	15,928.5187	15,928.5312	1.1	(−Met)
RS15[a]	14,707.9501	14,707.9445	0.4	(−Met)
RS20	13,283.2996	13,283.3289	2.2	(−Met+Ac)
RS21	9152.5917	9152.6253	3.7	(+Ac)
RS25	13,638.7116	13,638.8100	7.2	(−Met+Me)&(−Met+2Me)
RS28	7882.2223	7882.2087	1.7	(+Ac)
RS29	6545.2723	6545.2730	0.1	(−Met)
RS3[a]	—	—	—	NF
RS7	22,168.3046	22,168.2861	0.8	(+Ac)
RS9	22,459.5509	22,459.4683	3.7	(−Met)
RSP4[a]	—	—	—	NF
RS11	18,340.9920	18,340.9939	1.1	(−Met+Ac)
RS6	28,680.0313	28,680.1193	3.1	No
RS8	24,073.1600	24,073.2000	1.7	(−Met)
RS4	29,466.0609	29,466.0893	1.0	(−Met)

(*continued*)

Table 1. *Continued*

Name	Calculated mass	Observed mass	Δm (ppm)	PTMs
RS26	12,882.0480	12,881.9986	3.8	(−Met),(S-S)
RS3a	29,813.7465	29,813.7404	0.2	(−Met)
RS2[a]	—	—	—	NF
RS23[a]	—	—	—	NF
RS27[a]	—	—	—	NF
RS27[a]	—	—	—	NF
RACK1[a]	—	—	—	NF
Nucleolin[a]	—	—	—	NF
eIF3a[a]	—	—	—	NF
eIF3c[a]	—	—	—	NF
eIF3i[a]	—	—	—	NF
eIF3b[a]	—	—	—	NF
eIF3e[a]	—	—	—	NF
eIF3f[a]	—	—	—	NF
eIF3d[a]	—	—	—	NF
eIF3h[a]	—	—	—	NF
eIF3j[a]	—	—	—	NF
eIF3 subunit 6 interacting protein[a]	—	—	—	NF

[a] Proteins that were identified only by the bottom-up method. The other proteins were identified by both the top-down and the bottom-up method. The data presented are from single mass measurements. NF: no PTMs were found for the proteins using the bottom-up method because the modified peptides were missing.

acetylation, while both RS12 and RS25 were found to be dimethylated and these modifications were stoichiometric. Several yeast and human ribosomal proteins have also been found to be methylated (Lhoest et al. 1984; Odintsova et al. 2003).

Disulfide formation in one of the human 40S ribosomal proteins was also detected. RS29 was found to have a mass 4 Da lower than the predicted mass and, based on the high resolution mass measurement, was assigned to contain two disulfide bonds. Although its function remains unclear, a previous study by Forbes et al. (2004) showed that *Methanococcus jannaschii* 50S ribosomal protein L34e also contained a PTM consisting of two disulfide bonds.

Not all of the ribosomal proteins could be identified based solely on the top-down approach. For example, assignment of the protein with the measured mass of 9413.8471 Da was complicated by the fact that it was a possible match for either RS27 with di-acetylation (1.74 ppm) or RS27a with the formation of two disulfide bonds (2.26 ppm). Similarly, several ribosomal proteins of higher molecular weight were not identified by the top-down method, and it is possible that they were irreversibly adsorbed within the porous resin during the reverse-phase separation step and thus were not detected (Lubman et al. 2002).

Bottom-up analysis of the native 40S subunit

Bottom-up identification of 40S ribosomal proteins was performed on an LC-Q-TOF Micro mass spectrometer. The Q-TOF instrument was equipped with a dual electrospray source, which alternated between the sample channel and the internal standard channel, and the instrument was set to switch between MS and MS/MS. Real-time calibration of the MS as well as MS/MS spectra ensures high mass measurement accuracy of both the precursor ions (average mass error < 20 ppm) and the fragment ions (average mass error < 0.1 Da).

One typical LC-MS/MS experiment identified 31 out of 32 (97% protein coverage) possible 40S ribosomal proteins (Table 1). Twenty-nine of the 40S ribosomal proteins were identified using three or more peptides. The only redundant hit was the RS4 protein. However, since the 40S subunit was extracted from HeLa cells, RS4 X was identified as the only possible isoform. The missing protein RS27 (MW = 9.3 kDa) has 15 basic amino acids (Arg and Lys) in its 83 amino acid sequence. It is possible that the smaller tryptic peptides were lost in the separation step and eluded identification. Whereas the top-down approach was essential for identifying the presence of modifications, the bottom-up approach gave better coverage even though most of the peptides containing the modifications were missing.

In addition to the 40S ribosomal proteins, guanine nucleotide-binding protein beta subunit-like protein (RACK1) was identified in the 40S fraction with high confidence (15 identified peptides with 57% sequence coverage). RACK1 is a cytoplasmic protein and functions as the receptor for activated protein kinase C (McCahill et al. 2002). Immunofluorescence studies showed that RACK1 harbors a significant amount of activated protein kinase C and shuttles between different intracellular sites

(Angenstein et al. 2002). A previous report by Shor et al. (2003) showed that a number of ribosomal proteins coprecipitated with the immunopurified RACK1. The consistent identification of RACK1 in the 40S fraction provides direct evidence that RACK1 forms a stable complex with the 40S subunit (Link et al. 1999; Sengupta et al. 2004). While its function has yet to be fully elucidated, studies have indicated that RACK1 might affect joining of 40S subunit and 60S subunits. For example, a *Saccharomyces cerevisiae* strain lacking Cpc2, a RACK1 ortholog, shows an increased number of half-mer polyribosomes (the 43S preinitiation complexes composed of mRNA, the 40S subunit and initiation factors) relative to a wild-type strain (Chantrel et al. 1998). Since RACK1 binds to activated protein kinase C, it might also be related to the phosphorylation of its associated proteins and act as a modulator in the protein kinase C signaling pathway (McCahill et al. 2002).

Top-down analysis of the HCV IRES bound translation initiation complex

The top-down method was used to analyze purified IRES-bound 40S ribosomal complexes as described above. A total of 22 40S ribosomal proteins were identified with good mass measurement accuracy. Only two of the 40S ribosomal proteins, RS29 and RS25, in the IRES pull-down complex were found to contain different post-translational modifications compared to the native 40S subunit (Table 1B). Two 40S ribosomal proteins, RS28 and RS29, were identified in the LC fraction at 44 min separation of the IRES pulldown–40S complex (Fig. 3). RS28 was found to be N-terminally acetylated, as found in the native 40S subunit. However, RS29 in the native sample contained two disulfide bonds, whereas it did not contain this modification in the IRES recruited sample. Another protein found to contain a different PTM in the IRES pulldown 40S subunit was RS25. RS25 was identified as being fully dimethylated in the native 40S subunit, but was present in both the mono- and dimethylated forms in the pulldown complex (Fig. 4). Six of the 22 40S proteins identified in the pulldown sample, including RS11, RS4, RS6, RS8, RS26, and RS3A, were not identified in the top-down analysis of the native 40S subunit; however, they were detected using the bottom-up approach. Whether these six proteins have different PTMs from the native 40S complex is unknown, as identification of the PTMs using the bottom-up method was incomplete. Thus, the combination of the top-down and bottom-up approaches was essential in this particular study.

Bottom-up analysis of the HCV IRES bound translation initiation complex

The IRES bound 40S subunit was digested with trypsin and the resulting tryptic peptide mixture was also analyzed by LC-MS/MS. The identified proteins included the expected 40S ribosomal proteins, the eukaryotic translation initiation factor 3 (eIF3) subunits and the protein factors associated with the IRES/40S complex. Using the bottom-up approach, 32 of the 40S ribosomal proteins were identified in the IRES pulldown 40S subunit. Twenty-eight of the 40S ribosomal proteins were identified using three or more peptides. However, post-translational modifications were again only identified for a small percentage of the proteins; i.e., RSP4 and RS18. This is most probably due to the fact that the proteins were modified on the N terminus, and these peptides were often missing in the digest analysis.

Of all the eukaryotic initiation factors, eIF3 is the most structurally complicated and has at least 12 component proteins (Hellen and Sarnow 2001; Morris-Desbois et al. 2001). Although the IRES can directly recruit the 40S subunit ($K_d = 1.9 \pm 0.4$ nM), it also binds to eIF3 with a somewhat lower affinity ($K_d = 35 \pm 5$ nM) (Kieft et al. 2001), and it has been

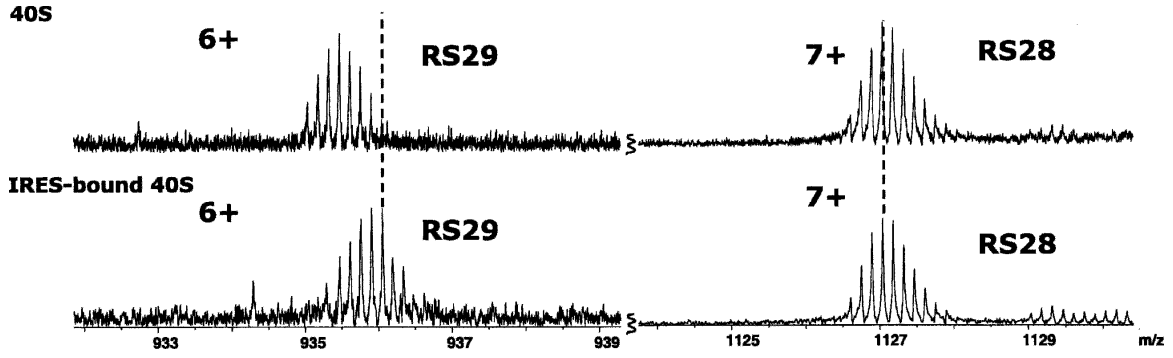

Figure 3. Identifications of RS28 and RS29 in the native 40S subunit and the IRES/40S complex using the top-down approach. The corresponding charge states are indicated on each expanded scale.

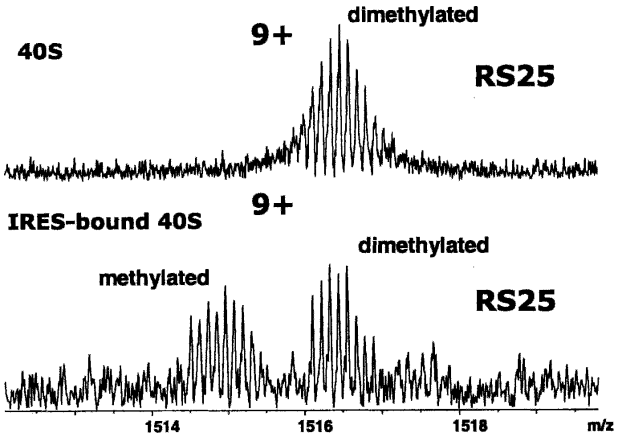

Figure 4. Expanded FT-ICR mass spectra of the RS25 protein in the native and IRES/40S complex. Clear difference is observed in the peak profiles showing differences in the modifications.

shown that four eIF3 subunits may interact directly with the IRES (Sizova et al. 1998). Here, all of the eIF3 proteins, except for eIF3g and eIF3k, were identified in the IRES bound complexes (Table 1B), which clearly indicated that it was associated with the IRES mediated translation initiation complex. However, the role that eIF3 plays in the IRES mediated translation is not fully understood. One possibility is that eIF3 may help the IRES to bind to the 40S subunits rather than 80S ribosomes. Alternatively, eIF3 may be needed to recruit other protein factors required for this cap-independent initiation mechanism (Kieft et al. 2001).

Two additional proteins, RACK1 and nucleolin, were also identified in the IRES-associated 40S sample. Because RACK1 was previously found in the native 40S subunit, its presence was not unexpected. Nucleolin, however, was only observed in the 40S/IRES complex. Nucleolin resides primarily in the nucleolus and is involved in ribosome assembly, rDNA transcription and rRNA maturation (Ginistry et al. 1999). However, a previous study by Waggoner and Sarnow (1998) showed that after cells were infected with poliovirus, nucleolin translocated from the nucleolus and accumulated in the cytoplasm. The authors further showed that this relocalization event was not a consequence of inhibition of host cell translation or of inhibition of host cell transcription by the poliovirus infection. Based on these two findings, the authors raise the possibility that nucleolin may be trapped in the cytoplasm by ligands generated by the virus. Later, Izumi et al. (2001) reported that the poliovirus IRES-mediated translation was stimulated when cells were cotransfected with a plasmid encoding the full-length nucleolin. Our identification of nucleolin in the HCV IRES-bound 40S subunit shows that nucleolin is physically associated with the HCV IRES-mediated translation initiation complex. Experiments indicating the presence of nucleolin were repeated several times for both the native and the IRES-bound complexes using different preparations. Nucleolin was never observed in the native 40S sample.

Conclusions

Bottom-up and top-down mass spectrometry approaches were used as complementary methods to identify proteins from native human 40S ribosomal subunits as well as the IRES-bound complex. Although the bottom-up approach provides excellent protein coverage, in most cases it missed in the identification of PTMs. In contrast, intact protein masses measured by the top-down method readily provide information about modifications, but some proteins were not detected. Combining the top-down and bottom-up methods provides both maximum coverage and PTM identification.

All the proteins identified by the top-down method were found to contain post-translational modifications, including methionine truncation, acetylation, methylation, and disulfide bonds. RACK1 was also detected, providing direct evidence for its stable association with the 40S subunit. In addition, the presence of nucleolin in IRES-bound 40S complexes but not native 40S subunit samples suggests its possible function during HCV IRES-mediated cap-independent translation initiation.

Materials and methods

Materials

The reagents used in this study, including formic acid, $CaCl_2$, $Mg(OAc)_2$, β-mercaptoethanol, urea, iodoacetamide, acetonitrile, myoglobin from equine heart, Tris, sucrose, EDTA, $MgCl_2$, DTT, puromycin, KCl, angiotensin II, and leucine enkephalin were purchased from Sigma. Glacial acetic acid and trifluoroacetic acid were purchased from EMD and J. T. Baker, respectively. Sequencing grade trypsin was purchased from Promega. Coomassie plus protein assay reagent was purchased from Pierce.

Top-down identification of 40S ribosomal proteins

The 40S ribosomal proteins were prepared by using an acid extraction method (Hardy et al. 1969). One volume of the 40S solution, which contained 500 μg of ribosomal proteins (~100 μL), was mixed with 0.25 volume of 1 M $Mg(OAc)_2$ and 2 volumes of glacial acetic acid. The solution was incubated on ice for 1 h and centrifuged at 12,000 rpm for 20 min. The pellet was resuspended and extracted by adding an adequate volume of 0.1 M $Mg(OAc)_2$ and glacial acetic acid. Formic acid was used in a different sample preparation in order to verify that the PTMs observed were not due to sample handling. The solution was vortexed for 10 min and centrifuged again. The suspensions were combined and injected directly onto either a 3×150-mm

Polaris C18 column (Varian) or a 4.6 × 250-mm 214TP54 C4 column (Vydac). Reverse-phase separation was performed by HPLC (Waters) at a flow rate of 0.5 mL/min using solvent A (0.1% TFA in water) and solvent B (0.1% TFA in acetonitrile). Proteins were eluted using the following gradient: (1) 0–3 min 10% B, (2) 3–33 min 10–30% B, (3) 33–37 min 30–37% B, (4) 73–103 min 37–50% B, (5) 103–113 min 50–80% B, or (6) 113–123 min 80–95% B. Chromatograms were recorded using UV detection at both 215 nm and 280 nm. One hundred twenty fractions were collected, lyophilized using a speed-vac (Eppendorf), and stored at −80°C for further analysis by Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS).

Mass spectra were acquired on a Bruker 7-tesla FT-ICR MS equipped with an Apollo (Bruker) electrospray ion source. Lyophilized samples were resuspended in 79:20:1 acetonitrile:water: formic acid. For internal calibration, equine myoglobin was added into the solution to a suitable concentration. Samples were infused into the mass spectrometer at 1 μL/min. Desolvation of the ions was performed using a ~100°C drying gas and ~150 V capillary exit voltage. Ions were externally accumulated in a radio frequency-only hexapole for 0.5–2 sec, and two to eight ion packets were transferred into the ICR cell for mass analysis. All samples were collected using gated trapping. Ions were detected using chirp excitation and broadband data acquisition using an average of 16–160 time domain transients containing 1024 k data points. The free induction decay (FID) spectra were zero filled, Gaussian-multiplied, and Fourier transformed. A resolution of 40,000 was routinely achieved. Calibrations were performed against equine myoglobin using the most abundant ions from four charge states. All the data were acquired and processed using Xmass v 6.0.0 (Bruker).

Tandem mass spectrometry experiments were performed by first isolating the ions of interest using correlated harmonic excitation field (CHEF) isolation sweep. Argon was then pulsed into the cell to a pressure of $\sim 10^{-6}$ mbar and the ions were collisionally activated for 3 sec using sustained off-resonance irradiation (SORI) at 1000 Hz above the ion's cyclotron frequency.

Bottom-up identification of 40S ribosomal proteins

One hundred micrograms of 40S ribosomal proteins were denatured in buffer A (6 M urea, 2 mM β-mercaptoethanol and 200 mM NH_4HCO_3, pH 7.5) at 50°C for 20 min. Proteins were then alkylated by adding iodoacetamide to a final concentration of 50 mM, followed by incubation in the dark for 20 min. The urea concentration was then diluted to 2 M and 1 mM $CaCl_2$ was added. Trypsin was added into the solution to a final substrate-to-enzyme ratio (w/w) of 50:1 and digestion was performed overnight at 37°C. The peptide mixture was desalted using an Oasis RP SPE cartridge (Waters), lyophilized, resuspended in water, and filtered with a 0.2-μm filter (Fisher). All of the tryptic peptide mixture was then subjected to LC-MS/MS (Waters) analysis. Reverse phase separation was performed at 1.5 μL/min without split on a Vydac 218MS C18 column (0.3 × 150 mm) in a 180 min gradient of 5–45% D (solvent C: 97/340 water/acetonitrile with 0.1% formic acid; solvent D: 97/3 acetonitrile/water with 0.1% formic acid).

MS and MS/MS experiments were performed on a Q-TOF micro mass spectrometer (Waters). A dual electrospray source alternated between the sample emitter and the calibrant emitter. A peptide mixture containing leucine enkephalin and angiotensin was used as the calibrant solution. The pseudo-internal calibration corrected for the effect of temperature change on the mass measurement accuracy during the entire LC-MS/MS experiment. The three most abundant multiply charged ions (+2, +3, +4) in the survey scans were subjected to tandem MS experiments. Argon was used as the collision gas. The collision energy profile applied was previously optimized for ions with different charge states (+2, +3, +4) and mass-to-charge regions. Each multiply charged precursor ion was calibrated against the doubly charged angiotensin ion (m/z 523.7751) and each fragment ion was calibrated against the singly charged leucine enkephalin ion (m/z 556. 2771).

The calibrated MS and MS/MS spectra were then used to search against all the entries of *Homo sapiens* in the SwissProt database using both Mascot and ProteinLynx (Waters). Several criteria were used to ensure the correct identifications of the 40S ribosomal proteins. First, 20 ppm and 0.1 Da were set as the maximum mass errors allowed for the precursors ions and the fragment ions, respectively, in identifying the protein hits. Second, if two or less peptides were used to identify a protein hit, manual inspection was performed to verify the hits.

Purification of the human 40S subunit and IRES-bound 40S complex

Samples of 40S ribosomal subunits were isolated from HeLa cell (National Cell Culture Center) lysate essentially using the method as described (Pestova et al. 1996). Briefly, the KCl concentration in HeLa cell lysate was adjusted to 0.5 M and 20 mL of this lysate was loaded onto a 45-mL sucrose cushion (20 mM Tris, pH 7.5, 0.5 M KCl, 1 M sucrose, 1 mM EDTA, 6 mM $MgCl_2$, and 2 mM DTT) and centrifuged in a Beckmann ultracentrifugation Ti 45 rotor at 4°C at 40,000 rpm for 16 h. The 0.5 M ribosomal salt wash (0.5 M RSW) pellet was resuspended in buffer B (20 mM Tris, pH 7.5, 100 mM KCl, 2.5 mM $MgCl_2$, 2 mM DTT) at a concentration of 100 OD/mL at 260 nm. Approximately 2.3 mM puromycin in buffer A was added to dissociate 80S ribosomes into 60S and 40S subunits. The ribosome suspension was incubated on ice for 10 min, and then at 37°C for 10 min. This was followed by the addition of KCl to a final concentration of 0.5 M. The ribosome suspension was then layered onto 10–50% sucrose gradient in buffer C (20 mM Tris, pH 7.5, 0.5 M KCl, 2.5 mM $MgCl_2$, 2 mM DTT) and centrifuged at 23,000 rpm in a Beckmann SW28 rotor at 4°C for 16 h. The 40S subunits were recovered by fractionating the gradient and monitoring the absorbance at 260 nm. Purification of IRES-bound translation initiation complexes was performed as described elsewhere (Ji et al. 2004). As phosphatase inhibitors were not used during preparation of the HeLa cell lysate, no information was obtained regarding possible phosphorylation sites.

Acknowledgments

We thank Waters Corp. for providing the Q-TOF Micro mass spectrometer.

References

Angenstein, F., Evans, A.M., Settlage, R.E., Moran, S.T., Ling, S., Klintsova, A.Y., Shabanowitz, J., Hunt, D.F., and Greenough, W.T. 2002. A receptor for activated C kinase is part of messenger ribonucleoprotein complexes associated with polyA-mRNAs in neurons. *J. Neurosci.* **22:** 8827–8837.

Arnold, R.J., Polevoda, B., Reilly, J.P., and Sherman, F. 1999. The action of N-terminal acetyltransferases on yeast ribosomal proteins. *J. Biol. Chem.* **274:** 37035–37040.

Chantrel, Y., Gaisne, M., Lions, C., and Verdiere, J. 1998. The transcriptional regulator Hap1p (Cyp1p) is essential for anaerobic or heme-deficient growth of *Saccharomyces cerevisiae*: Genetic and molecular characterization of an extragenic suppressor that encodes a WD repeat protein. *Genetics* **148:** 559–569.

Chong, B.E., Hamler, R.L., Lubman, D.M., Ethier, S.P., Rosenspire, A.J., and Miller, F.R. 2001. Differential screening and mass mapping of proteins from premalignant and cancer cell lines using nonporous reversed-phase HPLC coupled with mass spectrometric analysis. *Anal. Chem*. **73:** 1219–1227.

Du, Y., Meng, F., Patrie, S.M., Miller, L.M., and Kelleher, N.L. 2004. Improved molecular weight-based processing of intact proteins for interrogation by quadrupole-enhanced FT MS/MS. *J. Proteome Res*. **3:** 801–806.

Forbes, A.J., Patrie, S.M., Tylor, G.K., Kim, Y., Jiang, L., and Kelleher, N.L. 2004. Targeted analysis and discovery of posttranslational modifications in proteins from *Methanogenic archaea* by top-down MS. *Proc. Natl. Acad. Sci*. **101:** 2678–2683.

Ge, Y., ElNaggar, M., Sze, S.K., Oh, H.B., Begley, T.P., McLafferty, F.W., Boshoff, H., and Barry III, C.E. 2003. Top down characterization of secreted proteins from *Mycobacterium tuberculosis* by electron capture dissociation mass spectrometry. *J. Am. Soc. Mass Spectrom*. **14:** 253–261.

Ginistry, H., Sicard, H., and Bouvet, P. 1999. Structure and functions of nucleolin. *J. Cell Sci*. **112:** 761–772.

Hanash, S.M. 2000. Biomedical applications of two-dimensional electrophoresis using immobilized pH gradients: Current status. *Electrophoresis* **21:** 1202–1209.

Hardy, S.J.S., Kurland, C.G., Voynow, P., and Mora, G. 1969. Ribosomal proteins of *Escherichia coli*. I. Purification of the 30 S ribosomal proteins. *Biochemistry* **8:** 2897–2905.

Hellen, C.T. and Sarnow, P. 2001. Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes Dev*. **15:** 1593–1612.

Hunt, F.D., Yates III, J.R., Shabanowitz, J., Winston, S., and Hauer, C.R. 1986. Protein sequencing by tandem mass spectrometry. *Proc. Natl. Acad. Sci*. **83:** 6233–6237.

Izumi, R.E., Valdez, B., Banerjee, R., Srivastava, M., and Dasgupta, A. 2001. Nucleolin stimulates viral internal ribosome entry site-mediated translation. *Virus Res*. **76:** 17–29.

Ji, H., Fraser, C.S., Yu, Y., Leary, J.A., and Doudna, J.A. 2004. Coordinated assembly of human ribosomes by the hepatitis C virus IRES RNA. *Proc. Natl. Acad. Sci*. **101:** 16990–16995.

Kelleher, N.L. 2004. Top-down proteomics. *Anal. Chem*. **76:** 196A–203A.

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.

Lee, S., Berger, S.J., Martinovicá, S., Paša-Tolicá, L., Anderson, G.A., Shen, Y., Zhao, R., and Smith, R.D. 2002. Direct mass spectrometric analysis of intact proteins of the yeast large ribosomal subunit using capillary LC/FTICR. *Proc. Natl. Acad. Sci*. **99:** 5942–5947.

Lhoest, J., Lobet, Y., Costers, E., and Colson, C. 1984. Methylated proteins and amino-acids in the ribosomes of *Saccharomyces-cerevisiae*. *Eur. J. Biochem*. **141:** 585–590.

Link, A.J., Eng, J., Schieltz, D.M., Carmack, E., Mize, G.J., Morris, D.R., Garvik, B.M., and Yates III, J.R. 1999. Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol*. **17:** 676–682.

Lopez, M.F., Kristal, B.S., Chernokalskaya, E., Lazarev, A., Shestopalov, A.I., Bogdanova, A., and Robinson, M. 2000. High-throughput profiling of the mitochondrial proteome using affinity fractionation and automation. *Electrophoresis* **21:** 3427–3440.

Lu, H., Li, W., Noble, W.S., Payan, D., and Anderson, D.C. 2004. Riboproteomics of the hepatitis C virus internal ribosomal entry site. *J. Proteome Res*. **3:** 949–957.

Lubman, D.M., Kachman, M.T., Wang, H., Gong, S., Yan, F., Hamler, R.L., O'Neil, K.A., Zhu, K., Buchanan, N.S., and Barder, T.J. 2002. Two-dimensional liquid separations-mass mapping of proteins from human cancer cell lysates. *J. Chromatogr. B* **782:** 183–196.

MacCoss, M.J., McDonald, W.H., Saraf, A., Sadygov, R., Clark, J.M., Tasto, J.J., Gould, K.L., Wolters, D., Washburn, M., Weiss, A., et al. 2002. Shotgun identification of protein modifications from protein complexes and lens tissue. *Proc. Natl. Acad. Sci*. **99:** 7900–7905.

Madison-Antenucci, S., Grams, J., and Hajduk, S.L. 2002. Editing machines: The complexities of trypanosome RNA editing. *Cell* **108:** 435–438.

McCahill, A., Warwicker, J., Bolger, G.B., Houslay, M.D., and Yarwood, S.J. 2002. The RACK1 scaffold protein: A dynamic cog in cell response mechanisms. *Mol. Pharmocol*. **62:** 1261–1273.

Morris-Desbois, C., Réty, S., Ferro, M., Garin, J., and Jalinot, P. 2001. The human protein HSPC021 interacts with Int-6 and is associated with eukaryotic translation initiation factor 3. *J. Biol. Chem*. **276:** 45988–45995.

Odintsova, T.I., Mueller, E., Ivanov, A.V., Egorov, T.A., Bienert, R., Vladimirov, S.N., Kostka, S., Otto, A., Wittmann-Liebold, B., and Karpova, G.G. 2003. Characterization and analysis of posttranslational modifications of the human large cytoplasmic ribosomal subunit proteins by mass spectrometry and edman sequencing. *J. Protein Chem*. **22:** 249–258.

Pestova, T.V., Hellen, C.U., and Shatsky, I.N. 1996. Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Mol. Cell Biol*. **16:** 6859–6869.

Pestova, T.V., Kolupaeva, V.G., Lomakin, I.B., Pilipenko, E.V., Shatsky, I.N., Agol, V.I., and Hellen, C.U.T. 2001. Molecular mechanisms of translation initiation in eukaryotes. *Proc. Natl. Acad. Sci*. **98:** 7029–7036.

Phizicky, E.M. and Fields, S. 1995. Protein–protein interactions: Methods for detection and analysis. *Microbiol. Rev*. **59:** 94–123.

Pugh, B.F. 1996. Mechanisms of transcription complex assembly. *Curr. Opin. Cell. Biol*. **8:** 303–311.

Ranish, J.A., Yi, E.C., Leslie, D.M., Purvine, S.O., Goodlett, D.R., Eng, J., and Aebersold, R. 2003. The study of macromolecular complexes by quantitative proteomics. *Nat. Genet*. **33:** 349–355.

Sengupta, J., Nilsson, J., Gursky, R., Spahn, C.M.T., Nissen, P., and Frank, J. 2004. Identification of the versatile scaffold protein RACK1 on the eukaryotic ribosome by cryo-EM. *Nat. Struct. Mol. Biol*. **11:** 957–962.

Shevchenko, A., Jensen, O.N., Podtelejnikov, A.V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Boucherie, H., and Mann, M. 1996. Linking genome and proteome by mass spectrometry: Large-scale identification of yeast proteins from two dimensional gels. *Proc. Natl. Acad. Sci*. **90:** 14440–14445.

Shor, B., Calaycay, J., Rushbrook, J., and McLeod, M. 2003. Cpc2/RACK1 is a ribosome-associated protein that promotes efficient translation in *Schizosaccharomyces pombe*. *J. Biol. Chem*. **278:** 49119–49128.

Sizova, D.V., Kolupaeva, V.G., Pestova, T.V., Shatsky, I.N., and Hellen, C.U.T. 1998. Specific interaction of eukaryotic translation initiation factor 3 with the 5′ nontranslated regions of hepatitis C virus and classical swine fever virus RNAs. *J. Virol*. **72:** 4775–4782.

Strader, M.B., VerBerkmoes, N.C., Tabb, D.L., Connelly, H.M., Barton, J.W., Bruce, B.D., Pelletier, D.A., Davison, B.H., Hettich, R.L., Larimer, F.W., et al. 2004. Characterization of the 70S ribosome from *Rhodopseudomonas palustris* using an integrated "top-down" and "bottom-up" mass spectrometric approach. *J. Proteome Res*. **3:** 965–978.

VerBerkmoes, N.C., Bundy, J.L., Hauser, L., Asano, K.G., Razumovskaya, J., Larimer, F., Hettich, R.L., and Stephenson Jr., J.L. 2002. Integrating "top-down" and "bottom-up" mass spectrometric approaches for proteomic analysis of *Shewanella oneidensis*. *J. Proteome Res*. **1:** 239–252.

Waggoner, S. and Sarnow, P. 1998. Viral ribonucleoprotein complex formation and nucleolar-cytoplasmic relocalization of nucleolin in poliovirus-infected cells. *J. Virol*. **72:** 6699–6709.

Will, C.L. and Luhrmann, R. 2001. Spliceosomal UsnRNP biogenesis, structure and function. *Curr. Opin. Cell. Biol*. **13:** 290–301.

Zhou, Z., Licklider, L.L., Gygi, S.P., and Reed, R. 2002. Comprehensive proteomic analysis of the human spliceosome. *Nature* **419:** 182–185.