⌘ Author's Choice

© 2012 by The American Society for Biochemistry and Molecular Biology, Inc.
This paper is available on line at http://www.mcponline.org

Dynamics of Ribosomal Protein S1 on a Bacterial Ribosome with Cross-Linking and Mass Spectrometry*[S]

Matthew A. Lauber‡, Juri Rappsilber§, and James P. Reilly‡¶

Ribosomal protein S1 has been shown to be a significant effector of prokaryotic translation. The protein is in fact capable of efficiently initiating translation, regardless of the presence of a Shine-Dalgarno sequence in mRNA. Structural insights into this process have remained elusive, as S1 is recalcitrant to traditional techniques of structural analysis, such as x-ray crystallography. Through the application of protein cross-linking and high resolution mass spectrometry, we have detailed the ribosomal binding site of S1 and have observed evidence of its dynamics. Our results support a previous hypothesis that S1 acts as the mRNA catching arm of the prokaryotic ribosome. We also demonstrate that in solution the major domains of the 30S subunit are remarkably flexible, capable of moving 30–50Å with respect to one another. ***Molecular & Cellular Proteomics 11: 10.1074/mcp.M112.019562, 1965–1976, 2012.***

Initiation of translation is often the rate-limiting step of protein biosynthesis (1). In prokaryotes, this process is widely recognized to be directed by the Shine-Dalgarno (S.D.)[1] sequence of mRNA and its complementation with the 3′ end of 16S rRNA (2). However, binding of the S.D. sequence to the ribosome is not obligatory for initiation. Ribosomal protein S1, widely conserved in prokaryotes, (3) has been shown to efficiently initiate translation, regardless of the presence of an S.D. sequence (4, 5).

S1 is a strikingly atypical ribosomal protein, being both the largest (61 kDa) and the most acidic (pI 4.7) (6). The protein is composed of six homologous repeats each forming beta barrel domains (3) that in solution comprise a highly elongated structure spanning up to *ca.* 230 Å (7). This length is comparable to the diameter of the ribosome itself. In addition to these anomalous characteristics, S1 is also one of only two ribosomal proteins that has been attributed functional significance (6). Ribosomal protein S1, for instance, has no apparent role in the assembly of the ribosome, (2) yet is critical for translation in *E. coli* (8, 9). The functional significance of S1 is related to its most pronounced characteristic, the ability to simultaneously bind mRNA and the ribosome. Analysis of fragments produced by limited proteolysis and chemical cleavage of S1 has shown that an N-terminal fragment of S1 (residues 1–193) binds the ribosome (10) but not RNA (11). Likewise, a C-terminal fragment (res 172–557) binds RNA (12, 13) but not the ribosome (6, 10). By nature of this bi-functional structure, S1 enhances the *E. coli* ribosome's affinity for RNA ~5000 fold (14) and can directly mediate initiation of translation by binding the 5′ UTR of mRNA (4, 5). These observations have led to the hypothesis that S1 acts as a catching arm for the prokaryotic ribosome, working to bring mRNA to the proximity of the ribosome and thereby facilitate initiation (6).

Unfortunately, structural analyses capturing how S1 is able to function in this manner remain elusive. A high-resolution crystal structure of ribosome bound S1, or even free S1, does not exist, because S1 is recalcitrant to crystallography (6). Preparation of ribosomes for x-ray crystallography actually involves the deliberate removal of ribosomal protein S1 as a means to improve the reproducibility of crystallization and the quality of the ribosome crystals formed (15–17). The structure and interactions of the protein have nevertheless intrigued structural biologists for decades. However, studies completed to date have failed to convincingly demonstrate the interaction between S1 and the rest of the 30S subunit, because they were incapable of localizing the individual S1 domains (16, 18–20).

We have studied the binding of S1 to the 30S subunit by combining cross-linking with mass spectrometry. Chemical cross-linking has long been appreciated as a technique to probe protein-protein interactions (21, 22). With the advent of modern mass spectrometers, it can be very effectively employed to confidently identify the exact residues involved in linkages (23–28). In most cross-linking analyses, protein residues are targeted for covalent modification with a molecule

From the ‡Department of Chemistry, Indiana University, Bloomington, Indiana 47405; §Wellcome Trust Centre for Cell Biology, Institute of Cell Biology, The University of Edinburgh, Edinburgh EH9 3JR, UK and Institut für Biotechnologie, Technische Universität Berlin, 13353 Berlin, Germany

⌘ *Author's Choice*—Final version full access.

Received April 10, 2012, and in revised form, September 19, 2012
Published, MCP Papers in Press, October 2, 2012, DOI 10.1074/mcp.M112.019562

[1] The abbreviations used are: DEST, diethyl suberthioimidate; FA, formic acid; FDR, false discovery rate; mRNA, messenger ribonucleic acid; ppm, parts per million; rRNA, ribosomal ribonucleic acid; SCX, strong cation exchange; SD, Shine Dalgarno; SEC, size exclusion chromatography; SMTA, S-methylthioacetimidate; TCEP, tris(2-carboxyethyl)phosphine hydrochloride; UTR, untranslated region.

that contains two reactive groups separated by a spacer arm of known length. Only protein residues closer than the length of the spacer arm are capable of being linked. Identification of cross-linked residues thereby provides distance constraints for structural modeling. In this work, the novel amidinating protein cross-linker, DEST (diethyl suberthioimidate), was employed (29, 30). This amine reactive reagent, unlike commercially available reagents, preserves the native basicity of the residues it modifies while being effective at physiological pH. Use of the reagent is unlikely to perturb protein structure and the modifications it imparts are compatible with ionization for mass spectrometry. We have additionally shown that the cross-links it forms can be efficiently enriched from other components of proteolytic digests using strong cation exchange (SCX) chromatography, (30) and that DEST cross-linking of ribosomes yields structural information in excellent agreement with x-ray crystallography (29). Although DEST is an 11Å spacer arm cross-linker, it links alpha carbons up to 24Å apart because of the length and flexibility of lysine side chains. Nevertheless, this is sufficient resolution to approximate the binding positions of the 10kDa domains of S1. Furthermore, multiple cross-linking of a single domain significantly enhances the resolution with which it can be localized.

Here, through the application of protein cross-linking and high resolution mass spectrometry, we show that S1 binds to the 30S subunit near the anti-S.D. motif of the 16S rRNA, demonstrate that it is highly elongated even when bound to the ribosome, and provide evidence that its C-terminal mRNA binding region is remarkably dynamic. Our results thus indicate S1 is structurally poised, as previously hypothesized, (6) to act as the mRNA catching arm of the prokaryotic ribosome.

EXPERIMENTAL PROCEDURES

Preparation of E. coli 30S Subunits—Ribosomes were prepared from *E. coli* MRE 600 (16, 17, 31) in accord with Spedding's procedure, (32) except that the second sucrose/salt wash was omitted so as to minimize loss of ribosomal protein S1 (6). Cells were grown in Luria Bertani broth at 37 °C to mid-log phase and subsequently harvested by centrifugation. Pelleted cells were resuspended in buffer containing protease inhibitors and lysed via four passages through a French press at 11,000 psi. The lysate was then clarified and ribosomes were pelleted by centrifugation through 1.1 M sucrose. The resulting sucrose/salt washed ribosomes were stored as ~500 A_{260}/ml aliquots at −80 °C.

To dissociate and purify 30S subunits from 50S ribosomal subunits, sucrose/salt-washed ribosomes were added to a low magnesium (0.3 mM) buffer and subjected to sucrose density gradient fractionation. Fractions containing only 30S subunits were pooled, filtered and buffer exchanged into cross-linking buffer (50 mM HEPES-KOH, 100 mM KCl, 20 mM $MgCl_2$, pH 7.4 @ 20 °C) using Amicon Ultra 100K centrifugal filter devices (Millipore, Eschborn, Germany), and stored as 100 A_{260}/ml aliquots at −80 °C. Details about the preparation of 30S subunits can be found in the supplemental information.

Extraction of Ribosomal Proteins—Ribosomal proteins were extracted from 30S subunits through acetic acid precipitation of rRNA. Glacial acetic acid and 1 M $MgCl_2$ were added to samples such that the solutions contained a 3:6:1 (v/v/v) mixture of sample/glacial acetic acid/1 M $MgCl_2$. The samples were vortexed and allowed to remain at room temperature for 10 min before the rRNA precipitate was pelleted by centrifugation at 14,100 g for 10 min.

Analysis of Unmodified 30S Ribosomal Proteins—To assay the contents of our 30S subunit preparation, unmodified ribosomal proteins were analyzed by LC-MS. Ribosomal proteins were extracted from ~100 pmol of unmodified 30S subunits by the acetic acid procedure described above, loaded onto a BioBasic-4 reversed phase column (5 μm, 1 × 100 mm, Thermo Fisher Scientific, Bellefonte, PA) and eluted at 50 μl/min with a 55 min gradient between 0.1% formic acid (FA) in 90:10 water/acetonitrile (ACN) and 0.1% FA in 40:60 water/ACN. Reversed-phase effluent resulting from this separation was infused into the electrospray ionization (ESI) source of a quadrupole time-of-flight mass spectrometer (Q-TOF, Waters, MicroMass, Manchester, UK) at a flow rate of 10 μl/min. The voltage settings for the ESI needle, sample cone, and extraction cone were +3.0 kV, 35 V, and 1.5 V, respectively. Mass spectra were acquired over a range of 600–1800 *m/z.* Intact protein masses were obtained by manually summing together the raw spectra corresponding to chromatographic peaks and deconvolution using MassLynx (Waters, Milford, MA, version 4.1) and MaxEnt 1 or by automatic deconvolution using Protein Trawler software (BioAnalyte, Portland, ME).

Cross-Linking of the 30S Subunit—Aliquots of 30S subunits were diluted with cross-linking buffer (50 mM HEPES-KOH, 100 mM KCl, 20 mM $MgCl_2$, pH 7.4 @ 20 °C) to 20 A_{260}/ml (which converts to 1.4 μM)(32) and mixed in equal volume with 1.9 mM DEST, such that 30S subunits were modified at 0.7 μM in the presence of 0.95 mM DEST (Scheme S1). These conditions yielded an ~5:1 DEST to protein primary amine ratio, assuming that there are 288 such amines in this preparation of 30S subunits. After proceeding at room temperature for 6 h, reactions were quenched by adding 1 M NH_4Cl to a final concentration of 100 mM. These conditions are deliberately used to allow DEST to react until it is near fully hydrolyzed as well as yield only partial modification of proteins (30). The modified 30S subunits were subsequently cleared of hydrolyzed reagent by exchange into a post-cross-linking buffer (20 mM HEPES-NH_4OH, 100 mM NH_4Cl, 20 mM $MgCl_2$, 5 mM DTT, pH 7.4 @ 4 °C) using Amicon Ultra 100K centrifugal filter devices (Millipore, Eschborn, Germany) and repurified by sucrose density gradient fractionation to ensure only intra-30S subunit cross-linking was analyzed. Fractions containing DEST-modified 30S subunits were manually aspirated, pooled, filtered and buffer exchanged into post-cross-linking buffer using Amicon Ultra 100K centrifugal filter devices (Millipore, Eschborn, Germany), and stored as 100 A_{260}/ml (6.9 μM) aliquots at −80 °C. For one replicate analysis, 130 μl of this sample was prepared and analyzed according to the procedures below. Each replicate, as a result, required 900 pmol of 30S subunits and 270 μg of protein, assuming one mol of subunit contains 300,000 g of protein. In total, the study of S1/30S cross-linking involved two replicate analyses and ~540 μg of protein.

Size Exclusion Chromatography—Ribosomal proteins were extracted from 450 pmol of DEST-modified 30S subunits. This sample, containing ~135 μg of protein, was subsequently buffer exchanged into a low pH urea containing buffer (6% (v/v) acetic acid, 6 M urea, 100 mM NH_4 acetate, pH 3.5 @ 20 °C) using an Amicon Ultra 10K centrifugal filter device (Millipore, Eschborn, Germany). These samples were then separated by size exclusion chromatography using an Acquity UPLC BEH200 SEC column (1.7 μm, 4.6 × 150 mm, Waters, Milford, MA). It was expected that because S1 is significantly larger than other ribosomal proteins it could be enriched particularly after cross-linking using SEC. Proteins were eluted through the column with mobile phase (6% (v/v) acetic acid, 100 mM NH_4 acetate, pH 3.5 @ 20 °C) at 0.3 ml/min (Waters 2695, Milford, MA) and detected via absorbance at 280 nm (1200 series VWD SL+ UV-Vis detector, Agilent, Santa Clara, CA). Fractions corresponding to 3 to 3.8 min, 3.8 to 4.6 min, and 4.6 to 5.4 min were collected, dried under

vacuum, and stored at −20 °C until subsequent proteolytic digestion. The amount of protein in each SEC fraction was estimated based on the amount of sample injected and the A_{280} trace from the chromatogram.

Unmodified ribosomal proteins were also subjected to size exclusion chromatography to provide a frame of reference for the elution of the cross-linked proteins. Ribosomal proteins were extracted from 300 pmol of unmodified 30S subunits, prepared for injection, and chromatographed under the same conditions outlined for DEST-modified ribosomal proteins. The amount of unmodified ribosomal proteins that had been analyzed was less than the amount of cross-linked ribosomal proteins that had been analyzed. The sample of unmodified proteins was less heterogeneous and consequently produced sharper, more intense chromatographic peaks. Less sample was required to mark their elution times.

Proteolytic Digestion—Dried aliquots of SEC fractionated, DEST-modified ribosomal proteins (*ca.* 68 μg) and proteomics grade trypsin (3.4 μg) were reconstituted in 63 μl of 50 mM Tris/10 mM $CaCl_2$ (pH 8) and 0.1% (w/v) Rapigest (Waters, Milford, MA). Proteins were not reduced and alkylated. Each digest reaction was allowed to proceed at 37 °C for 24 h and subsequently quenched by adding 10% TFA to a final concentration of 0.5%. Extended digestion times and Rapigest, a surfactant, were used to minimize the number of missed cleavages. To hydrolyze Rapigest, the quenched digests were incubated for 30 min at 37 °C. By-products of the Rapigest regeant were cleared from the samples by centrifuging at 14,000 × *g*.

SCX Enrichment and Fractionation of DEST Cross-Links—Strong cation exchange (SCX) chromatography was employed to simplify proteolytic digests of cross-linked proteins prior to their analysis by nanoLC-MS/MS through 1) enriching interpeptide cross-links from noncross-linked peptide (linear peptide) species via a separation at low pH and 2) fractionating the enriched interpeptide cross-links via a separation at intermediate pH. We have previously demonstrated how DEST interpeptide cross-links can be enriched using SCX chromatography by exploiting the fact that they contain more basic residues, and thus positive charges at low pH, than other species in a proteolytic digest (29, 30). Tryptic digest of the cross-linked sample was loaded onto an SCX column (TSKgel SP-NPR, 4.6 × 35 mm, Tosoh Bioscience, Montgomeryville, PA) equilibrated with 0.1% TFA in water (pH 2). Noncross-linked (linear peptide) species were then eluted from the column with a 300 mM NaCl mobile phase. The enriched interpeptide cross-links, still adsorbed to the column, were then fractionated. The mobile phase was changed to a 20 mM sodium acetate pH 5 buffer and a gradient of NaCl was implemented (40 min ramp from 0 to 120 mM NaCl followed by a 10 min hold of 1000 mM NaCl). DEST interpeptide cross-links eluting from the SCX column were collected sequentially onto 5 C18 trapping columns (Javelin, 5 μm, 1.0 × 20 mm, Thermo Scientific, Bellefonte, PA) using a valve system (33) to switch the flow path from one trapping column to the next every 10 min. The contents of each C18 trapping column were then desalted with 0.1% TFA in water at 0.3 ml/min for 3 min and eluted with a 5 min isocratic hold (flow rate: 100 μl/min) of 5% aqueous mobile phase (0.1% TFA in water) and 95% organic mobile phase (0.1% TFA in ACN). The entire eluate from each trap was dried under vacuum and stored at −20 °C until being reconstituted in 0.1% TFA in water for subsequent LC-MS/MS analysis.

Capillary LC-ESI-MS/MS—Capillary LC-ESI-MS/MS was employed to analyze the digests of SEC-fractionated, DEST-modified ribosomal proteins that had and had not been subjected to SCX chromatography. Digests that had not been subjected to SCX chromatography were analyzed to benchmark the effect of the SCX separation and because they can occasionally yield a few unique identifications. Two replicates of each sample were prepared and analyzed. In each run, ~1 μg of sample was separated and analyzed using an 11 cm long IntegraFrit capillary trapping column packed with 1.5 cm of C18 (150 μm x 11 cm, New Objective, Woburn, MA; Magic C18, 5 μm, 200Å, Michrom BioResources, Auburn, CA), a capillary emitter column packed with 15 cm of C18 (75 μm x 15 cm, Magic C18, 5 μm, 100 Å, Michrom BioResources), an LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific, San Jose, CA), and a NanoLC-Ultra chromatography system (Eksigent, Dublin, CA). In each run, sample was injected onto the trapping column and flushed with mobile phase A (0.1% FA in 97:3 water/ACN) for 12 min at 2.5 μl/min to remove salts and contaminants. The flow rate was then reduced to 300 nL/min, the trapping column was put in line with the capillary emitter column, and a 150-min gradient between 4 and 35% mobile phase B (0.1% FA in ACN) was implemented. Eluting peptides were electrosprayed into an LTQ-Orbitrap XL mass spectrometer operating in data-dependent mode to acquire a full profile MS scan (300–1800 m/z) and centroid MS/MS scans of the three most intense precursor ions. The AGC target value was set to 7.5×10^5 for MS scans and 1×10^5 for MS/MS scans. The max fill time was set to 500 ms. Precursors were isolated with a 2 m/z window and subjected to CID in the LTQ at 35% normalized collision energy. Both the MS and MS/MS scans were acquired in the Orbitrap with resolution set to 100,000 and 7500, respectively. Dynamic exclusion was employed with the following settings: a 90 s exclusion duration, maximum exclusion list of 500, and one repeat count. In addition, charge state rejection was enabled for 1+ and 2+ charge states. MS/MS spectra were reduced to peak lists using MaxQuant (version 1.1.1.36). Default program settings were used, except that the value for top peaks per 100 Da was changed to 100 and de-isotoping was disabled.

Database Searching—MS/MS data contained in the resulting files were searched against the sequences of the *E. coli* 30S ribosomal proteins for interpeptide cross-links using the in-house program, Xi (version 1.2.315), (34) an algorithm that produces identifications based on matching experimental precursor and fragment ion masses with the theoretical precursor and fragment ion masses of cross-links generated *in silico* from a sequence database. The sequence database contained entries for the 21 *E. coli* 30S ribosomal proteins (S1-S21), without entries for possible contaminating proteins, such as trypsin and keratin. More information of the sequences of the ribosomal proteins is provided below. Searches were completed with MS tolerance set to 6 ppm, MS2 tolerance set to 10 ppm, enzyme set to trypsin, cross-linker set to DEST (mass shift: 136.10005, specificity: Lys, Protein N-term), and maximum missed cleavages set to 2 (allows for up to 2 in an entire cross-link not considering cross-linked Lys residues as cleavage sites). Protein modifications, other than DEST cross-linking, were not included in these searches. In addition, ion matching was set to include b-ions, y-ions, precursor ions, water loss ions, and ammonia loss ions. Matches from all searches were required to have at least 50% of the ion current in a given spectrum assigned, a Xi score >5, and to correspond to peptides at least 4 residues in length. We also required each match to have a score at least 1 unit higher than the score of the assigned spectrum's next best match. For instance, a match with a Xi score of 8 was considered only if the assigned spectrum's next best match was less than 7. These score thresholds were established through searching both target and decoy databases.

Moreover, many matches were manually validated to ensure that confident identifications were made. For example, intraprotein cross-link matches involving peptides from a contiguous sequence of protein were manually inspected to determine if they indeed provided unambiguous identifications. This is imperative because in this situation a potential match could also have been made to a dead-end modified peptide derived from the same sequence, because it would have the exact same mass. These matches were thus manually inspected and disregarded if the fragmentation did not indicate the

cross-link structure. That is, it was required that one or more fragment ions unique to the cross-link be observed. In addition and perhaps more importantly, interprotein cross-link matches were manually validated to ensure that a number of fragment ions had been assigned to both peptide chains of the proposed cross-link. We have previously demonstrated that interprotein, but not intraprotein, cross-link matches require this additional scrutiny (29). Because the search space needed to identify interprotein cross-links is much larger than the minimal search space needed to identify intraprotein cross-links for a particular protein, there is a greater likelihood of an interprotein cross-link match being a false positive. To this end, interprotein cross-linking matches with Xi scores >9 were disregarded unless there were 4 unique fragments matched to within 10 ppm assigned on each peptide chain. Interprotein cross-linking matches with Xi scores >5 but <9 were disregarded unless there were 4 unique fragments matched to within 3 ppm assigned on each peptide chain. Fragments ambiguously matching either peptide chain in the cross-link were not included in this count. Finally, spectra that corresponded to cross-links with multiple possible linkage patterns were manually inspected, and exact linkages were proposed only when the two cross-linked residues could be defined by the observed fragmentation. C-terminal lysine residues produced by trypsin were not considered as potential linkage sites, because amidination is known to block tryptic cleavage (35, 36).

False Discovery Rate Analysis—The false discovery rates (FDRs) of this analysis were estimated by repeating searches against combined target-decoy databases that contained both unaltered and randomized sequences of the *E. coli* 30S ribosomal proteins. Searches were completed against five such target-decoy databases, each randomized independently. Decoy searches were conducted in this manner for several reasons. Randomized sequences were employed rather than unrelated sequences to ensure that the unique amino acid composition of the ribosomal proteins was preserved. Furthermore, randomized sequences rather than reversed sequences were used so that more than one decoy database could be searched and that consequently FDRs could be more accurately defined. Lastly, the decoy databases were constructed with unaltered and randomized sequences to capture the most common type of false positive, a match in which one peptide sequence, but not both, is correctly identified. Combined target-decoy databases produce these false positives via matches in which a peptide with an unaltered sequence is proposed to be cross-linked to a peptide with a randomized sequence.

In these searches, there were two decoy matches conforming to the criteria outlined in the previous section on database searching. These, so called decoy matches, were matches that involved a randomized sequence cross-linked to either the true or randomized S1 sequences. As this was the result of searching five different sets of randomized sequences, the average number of false positives in our S1 data was estimated to be 2 divided by 5, or 0.4. The FDR was calculated by dividing this number (0.4) by the number of S1 cross-link matches, 72, produced when searching against just the unaltered ribosomal protein sequences. This resulted in an FDR for spectra matches involving S1 of ~0.6%. The calculation presented here is in accord with the generally accepted definition of an FDR, (37) except that the number of decoy matches corresponding to one analysis had been tested numerous times.

The false discovery rate for cross-links not involving S1 was determined in the same manner. Searching 5 sets of randomized sequences yielded 4 decoy matches, indicating the average number of false positives in the data set was 0.8. Given that there were 648 non-S1 cross-link matches produced when searching the unaltered ribosomal protein sequences, this resulted in an FDR for spectra matches not involving S1 of ~0.1%.

Crystal Structure Visualization—PyMOL v. 0.99 (DeLano Scientific, www.pymol.org) was employed for the visualization and manipulation of the crystal structures for the 30S subunit of the *E. coli* ribosome (PDB: 2AVY)(16, 38) and domains 4 and 6 of S1 (PDB: 2KHI and 2KHJ) (3).

Depleting S1 from 30S Subunits—Protein S1 was removed from 30S subunits by taking advantage of the fact that it strongly adsorbs to polyU Sepharose and that it can be displaced from the ribosome with 1 M NH_4Cl (11). PolyU Sepharose 4B (125 mg) was placed into 1 ml spin columns (Pierce #69725, Thermo Scientific, Rockford, IL) and hydrated with polyU buffer (10 mM HEPES NH_4OH, 20 mM $MgCl_2$, 1 M NH_4Cl, 5 mM DTT, pH 7.5 @ 20 °C). Purified 30S subunits (1.5 nmol) were then loaded onto and passed four times through the columns by spinning at 100g. Subsequently, four 500 μl volumes of polyU buffer were passed through the columns. The flow through from these separations, containing the 30S subunits depleted of S1, was collected and exchanged into cross-linking buffer (50 mM HEPES-KOH, 100 mM KCl, 20 mM $MgCl_2$, pH 7.4 @ 20 °C) using Amicon Ultra 100K centrifugal filter devices (Millipore, Eschborn, Germany). Subunits depleted of S1 were analyzed by LC-MS and subjected to cross-linking as noted above.

SMTA Labeling of Proteins—Purified 30S subunits either containing or not containing protein S1 were modified with SMTA (S-methylthioacetimidate) under conditions similar to those used in previous ribosome labeling experiments (Scheme S2) (35, 39, 40). Unlike with cross-linking, the conditions of this experiment lead to modification of all solvent accessible primary amines. The SMTA reagent is present at high excess. Reactions are routinely carried out for only one hydrolysis half-life (1 h), (41) as there is no need to consider the hydrolysis of a second intramolecular thioimidate group and its effect on subsequent procedures. Reaction of thioimidates with amines must be faster than reaction with OH groups, because even in aqueous solution, in which water is present at very high concentrations, significant levels of amidination can be achieved. This is also evidenced by the fact that amidinating reagents are highly selective for amines and do not readily modify other nucleophiles (42, 43). 30S subunits (4 μM) in cross-linking buffer (50 mM HEPES-KOH, 100 mM KCl, 20 mM $MgCl_2$, pH 7.4 @ 20 °C) were mixed in equal volume with 200 mM SMTA that had been reconstituted in 160 mM KOH. The reaction pH was measured to be ~7.4. After 1 h, the reaction was quenched by the addition of acetic acid and precipitation of RNA. Methanethiol disulfide adducts formed during the reaction (44) were thereafter reduced with 50 mM TCEP (tris(2-carboxyethyl)phosphine hydrochloride). Two replicates of each sample type were analyzed by LC-MS as described above. Intensity weighted averages for the number of modifications added to each protein were calculated using mass spectral peak lists generated by Protein Trawler software (BioAnalyte, Portland, ME).

Protein Sequence Data—The protein sequences used in this analysis for *E. coli* 30S ribosomal proteins were derived from the K12 reference genome (GenBank Accession.Version U00096.2). Two revisions, as discussed below, were made to these sequences to make them consistent with MRE 600 sequences. The Asp residue at position 126 in S2 was changed to Glu, and the sequence of S7 was shortened to end with Trp 155.

RESULTS AND DISCUSSION

Isolation and Characterization of E. coli 30S Subunits—To facilitate this analysis of S1, ribosomes were prepared with special consideration. Because the protein is weakly bound to and thus easily removed from the ribosome, (6, 45) ribosomes were purified by centrifugation with only one sucrose/salt wash step, rather than the two often utilized (32, 39, 46). In addition, the purified sucrose/salt-washed ribosomes were

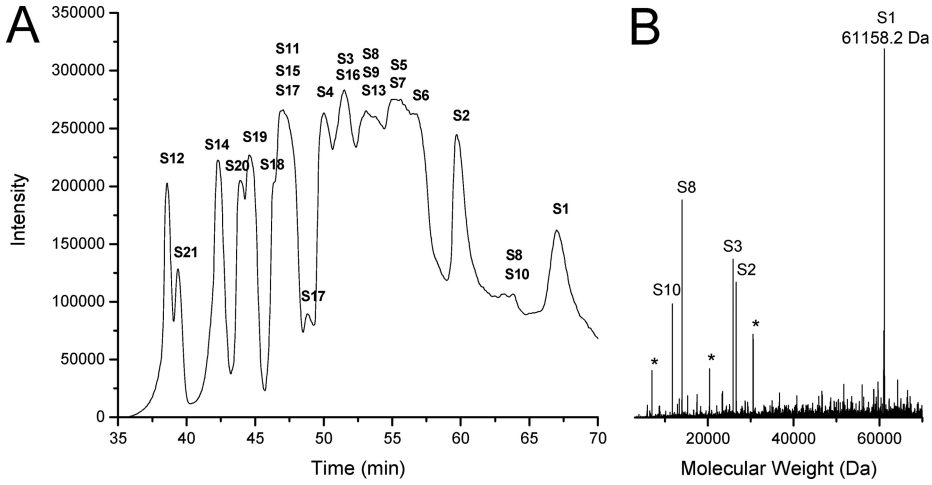

FIG. 1. **LC-MS of 30S ribosomal proteins.** *A*, Reversed phase chromatogram of proteins extracted from unmodified 30S subunits. *B*, Deconvoluted ESI mass spectrum of the species eluting between 65 and 70 min. Asterisks mark harmonic artifacts of deconvolution corresponding to the molecular weights of the proteins labeled. The labeled S1 mass is the result of an internal calibration.

dissociated and 30S subunits were isolated from 50S subunits using sucrose density gradient fractionation.

The proteins of these 30S subunits were characterized by LC-MS, as shown in Fig. 1. In this study, ribosomes were isolated from *E. coli* MRE 600, an RNase deficient strain commonly exploited to prepare homogenous ribosomes for structural analyses (16, 17, 31). The genome of this strain has not been sequenced. It has been assumed, particularly by x-ray crystallographers, (16) that ribosomal proteins from MRE 600 and the paradigmatic K12 strain share identical sequences. Indeed, the masses we measured for 19 of the 21 small subunit ribosomal proteins from MRE 600 matched those predicted using K12 sequences and canonical post-translational modifications to within the 3 Da mass accuracy of our QTOF instrument (supplemental Table S1) (39, 46–52). Two proteins, S2 and S7, were however suspected of having different sequences, because their observed masses significantly differed from those previously observed for S2 and S7 from K12 (39, 46). Analysis of sequences in the UniProt Knowledgebase (release 2012_03)(53) indicated that many *E. coli* strains, including O8 IAI1 and O9 HS, (54, 55) have S2 and S7 sequences that are both different than those from K12 and consistent with our mass measurements. These *E. coli* strains encode S2 with a Glu rather than an Asp at residue position 126, and they encode S7 in a truncated form with a sequence terminating at Trp 155.

LC-MS analysis also revealed that in contrast with previous proteomic studies of the *E. coli* ribosome, (39, 46) this preparation yielded appreciable signal for a protein whose mass matched the theoretical mass of S1 to within 2.2 Da (Fig. 1 and supplemental Table S1). Because we were particularly interested in confirming this identification, we internally calibrated the ESI TOF mass spectrum that included various charge states of co-eluting S2, S3, S8 and S10 ribosomal proteins. The resulting 61158.2 Da mass matched the S1 theoretical mass to within 0.2 Da, or 3 ppm. This mass is consistent with the K12 reference sequence and 166 of the other 167 *E. coli* S1 sequences in the UniProt Knowledgebase

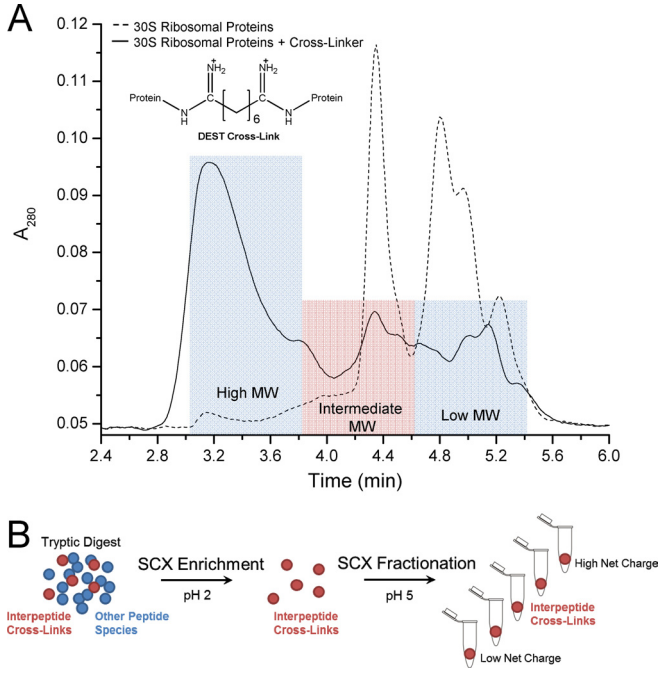

FIG. 2. **Preparation of cross-linked samples for nanoLC-ESI-MS/MS.** *A*, Size exclusion chromatogram of ribosomal proteins extracted from 450 pmol of DEST cross-linked 30S subunits (solid line) and of ribosomal proteins extracted from 300 pmol of unmodified 30S subunits (dashed line). *B*, Method for strong cation exchange enrichment and fractionation of interpeptide DEST cross-links.

(release 2012_03) (53). Ribosomal subunits prepared in this manner clearly contained protein S1 and in sufficient quantities for thorough structural analysis.

Preparation of Cross-Linked Samples—Purified 30S subunits were cross-linked at physiological pH with DEST present at an estimated 5:1 excess over protein primary amines. Proteins were thereafter extracted from the modified subunits, fractionated by size exclusion chromatography (Fig. 2*A*), and digested with trypsin. As in previous work, SCX chromatography was employed to enrich and fractionate the DEST interpeptide cross-links present in the digests by exploiting the

fact that they contain a considerable number of basic residues (Fig. 2*B*) (29, 30). Not surprisingly, these chromatographic procedures proved to be highly beneficial to this structural analysis of S1. Interprotein cross-links involving S1 were only detected in samples that had originated from the high molecular weight SEC fraction and were isolated by SCX.

Mass Spectrometric Analysis of Cross-Linked Peptides—Cross-links present in the prepared samples were analyzed using nanoLC-ESI-MS/MS with an LTQ-Orbitrap mass spectrometer. Both precursor and fragmentation mass spectra were acquired with high resolution in the Orbitrap. The resulting accurate precursor and fragment ion masses were searched for matches to theoretical interpeptide cross-links derived from sequences of the 30S ribosomal proteins. The entire workflow for analyzing the cross-linking of the 30S ribosomal subunit is provided in Scheme S3. Using the algorithm, Xi, (34) along with manual validation, we were able to identify 72 spectra matches associated with cross-linking of S1 with a false discovery rate (FDR) of ~0.6% (supplemental Tables S2 and S3). This number demonstrates that these data were obtained with very high confidence.

From these 72 spectra matches, 33 different linkages were identified, 23 corresponding to interprotein cross-linking and 10 to intraprotein cross-linking (Table I). Interprotein cross-linking between K37 of S18 and K150 of S1 was, for example, identified by means of the Orbitrap MS/MS spectrum displayed in Fig. 3. This spectrum contains ions resulting from cleavage of two different peptide chains; peak assignments shown in red correspond to fragmentation of the S18 peptide, NYITESGKIVPSR, whereas those in blue correspond to fragmentation of the S1 peptide, DTLHLEGKELEFK. In this spectrum, 14 different fragment assignments could be attributed to the S18 peptide sequence, and 15 different fragment assignments could be attributed to the S1 peptide sequence. The sequences of both peptides were thereby matched with very high confidence, particularly because all fragment masses were required to match theoretical masses to within 10 ppm (further details on spectra matching can be found under "Experimental Procedures"). That so many high confidence interprotein cross-links, such as this, were identified is noteworthy. The density of cross-linking information obtained here for S1 appears to be comparable or more extensive than that reported in recent large-scale cross-linking analyses (34, 56–59).

Cross-Linking of Protein S1—Interpretation of the identified protein linkages has led to an insightful understanding of S1. S1 interprotein cross-links outnumbered S1 intraprotein cross-links. Furthermore, the intraprotein cross-links that were detected predominately corresponded to cross-linking within a domain or between domains adjacent in sequence, such as between domains 2 and 3. This is consistent with S1 being an elongated protein, whose amino groups are more likely to cross-link to amines from other proteins on the surface of the ribosome than to other S1 amines. Interprotein cross-linking of S1 further demonstrates this point. S1 was found to cross-link to 15 different residues spanning close to 240 Å across the topography of the ribosome, from S19 near the subunit interface to S6 on the tip of the 30S platform to S3 at the mRNA entry pore (supplemental Fig. S1). These observations corroborate the conclusion from small angle x-ray scattering studies that S1 adopts a highly elongated structure (6).

These experiments most importantly provide explicit information about the localization of S1's individual domains. A model for the architecture of these domains and their characterized functions is presented in Fig. 4. Fig. 5 displays the crystal structure of the *E. coli* 30S subunit (16). As mentioned earlier, the first two domains of S1 are believed to be responsible for binding the ribosome (10, 11). The cross-linking of

TABLE I

Summary of identified linkages involving S1. (A) Intraprotein and (B) Interprotein identifications

A

Domain	Domain	Linkage	Matches
1	2	K88-K150	2
2	3	K150-K196	1
2	3	K150-K247	15
2	4	K150-K279	5
2	C-term	K150-K555	1
2	3	K158-K247	1
4	C-term	K279-K555	1
5	6	K363-K464	1
5	5	K370-K411	2
6	6	K450-K504	1
		Total	30

B

Protein 1	Protein 2	Linkage (Prot1-Prot2)	Matches
S1	S2	K14-K10	1
S1	S2	K88-K36	1
S1	S2	K363-K114	1
S1	S2	K363-K127	1
S1	S3	K279-K107	1
S1	S3	K279-K149	1
S1	S6	K150-K106	1
S1	S7	K117-K148	2
S1	S7	K150-K148	7
S1	S7	K158-K148	1
S1	S7	K229-K148	1
S1	S7	K279-K55	1
S1	S7	K279-K148	5
S1	S9	K279-K99	4
S1	S18	K150-K29	1
S1	S18	K150-K37	1
S1	S18	K272-K29	1
S1	S19	K279-K86	1
S1	S21	K150-K4	1
S1	S21	K150-K39	1
S1	S21	K247-K4	1
S1	S21	K279-K4	6
S1	S21	K279-K39	1
		Total	42

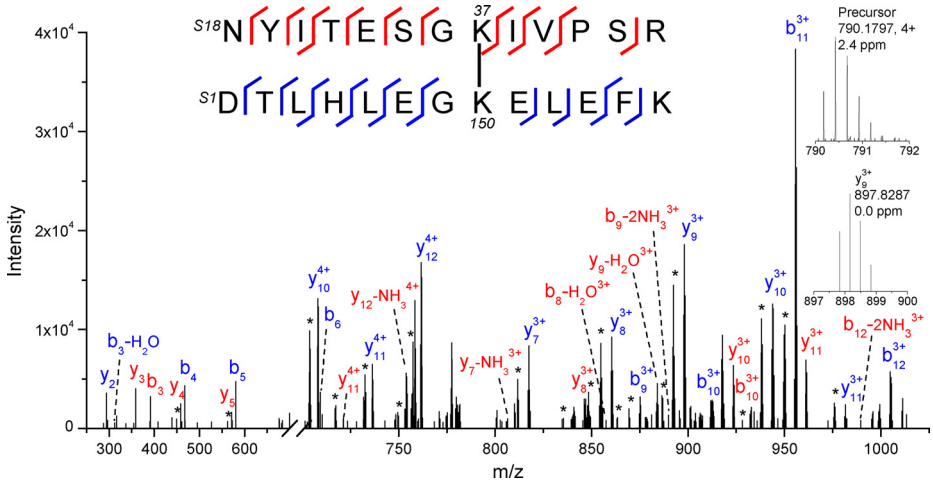

FIG. 3. **Orbitrap MS/MS of an interprotein cross-link between K37 of S18 and K150 of S1.** Peak assignments corresponding to the S18 peptide are shown in red and those corresponding to the S1 peptide are shown in blue. Peaks marked with asterisks correspond to neutral losses of other assignments.

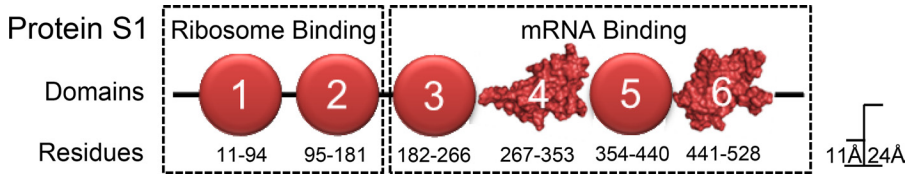

FIG. 4. **Domain architecture of protein S1.** The NMR structures of domains 4 and 6 are shown (3).

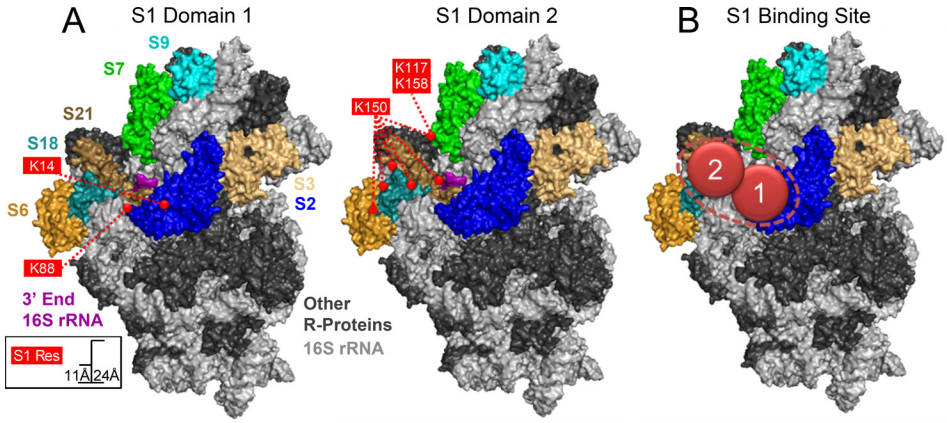

FIG. 5. **Cross-linking of the N-terminal ribosome binding region of S1 to the 30S ribosomal subunit.** *A*, Cross-linking of domains 1 and 2 involves residues localized near the 3′ end of the 16S rRNA. Residues on the 30S subunit (PDB: 2AVY)(16) that cross-linked to S1 are marked with red circles. The triangle on S6 shows the approximate position of a residue not present in the crystal structure. The S1 residues involved in these cross-links are highlighted in red. *B*, A binding site for the N-terminal ribosome binding region of S1 consistent with this cross-linking analysis.

these domains can therefore be used to elucidate the ribosomal binding site of S1. As displayed in Fig. 5*A*, domain 1 cross-linked to the N-terminal region of S2 through two different lysine residues. Domain 2, meanwhile, cross-linked to six different residues of S6, S7, S18, and S21 in an adjacent region at the platform of the 30S subunit, ~30Å away. The binding site of S1 can therefore be defined as the cleft between the platform and body/head of the 30S subunit. A model for the interaction between domains 1 and 2 of S1 and the 30S subunit that is consistent with the observed cross-linking pattern is shown in Fig. 5*B*. More detailed structural modeling requires a high resolution structure for the N-terminal ribosome binding domains of S1. Binding of S1 to this region of the 30S subunit is likely to be imperative for its ability to facilitate the initiation of translation. During initiation, mRNA binds along a groove in the 30S subunit, such that its 5′ end extends out through this cleft (60–62). This region of the 30S subunit contains the 3′ end of 16S rRNA with a sequence

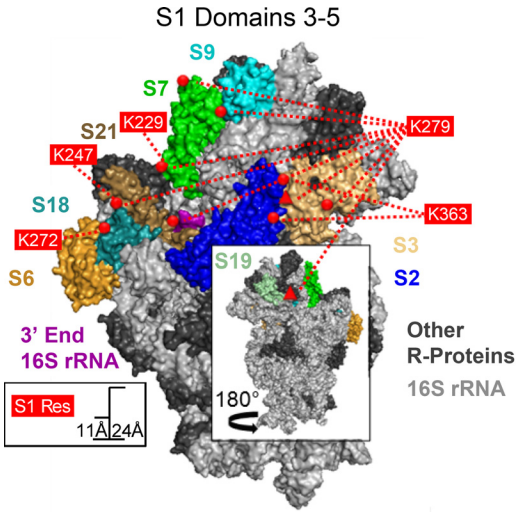

FIG. 6. **Cross-linking of the C-terminal mRNA binding region of S1 to the 30S ribosomal subunit.** Cross-linking of domains 3–5 involves residues spread far across the subunit. Residues on the 30S subunit (PDB: 2AVY)(16) that cross-linked to S1 are marked with red circles. Triangles on S3 and S19 show the approximate positions of residues not present or not solvent accessible in the crystal structure. The S1 residues involved in these cross-links are highlighted in red.

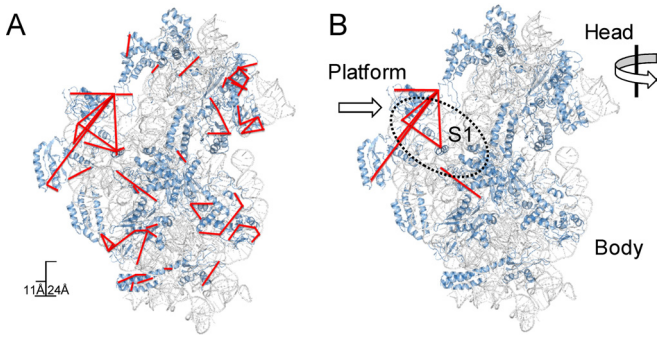

FIG. 7. **Structural context of cross-linking data.** *A*, Identified linkages involving residues present in the crystal structure of the *E. coli* 30S subunit (PDB 2AVY) are shown. Ribosomal proteins are displayed in blue, the 16S rRNA in gray, and linkages in red. *B*, Identified linkages associated with anomalously long distances. The proposed S1 ribosomal binding site is approximated with a dashed line. Movements of the 30S subunit domains inferred from the cross-link pattern are indicated with arrows.

complementary to the S.D. sequence found in many mRNA (2, 16). Our results show that S1 is also present in this location, explaining how S1, like the anti-S.D. sequence, is able to bind the 5′ UTR of mRNA and facilitate initiation (4, 5).

Protein S1 is believed to bind the mRNA through its C-terminal region (10, 12, 13). Fig. 6 displays the cross-linking of this region against the crystal structure of the 30S subunit. Interestingly, this part of S1 cross-links to a much larger region of the 30S subunit than domains 1 and 2. For example, we detected cross-linking of K279 from domain 4 to residues spanning nearly 200 Å across the 30S subunit. Such a cross-linking signature suggests that the central part of S1's mRNA binding region is capable of sweeping out a very large region in space, including to some extent the interfacial side of the subunit (see insert in Fig. 6). S1 appears then to not only be highly elongated but to also have a C-terminal region that is quite dynamic. Stochastic movement of the C-terminal region could enable S1 to capture mRNA proximal to the subunit. The dynamic nature of the C-terminal region may also play a role in supporting the binding and movement of mRNA through the ribosome during elongation.

This cross-linking study is consistent with a previous cryo-EM analysis of S1 containing 30S subunits, (18) in that the region near the 3′ end of the 16S rRNA is identified as the S1 binding site. Our results do, however, reflect a more complex behavior for S1. The physical density attributed to S1 in the cryo-EM analysis was surprisingly compact. It has since been pointed out that a significant portion of this density had been attributed to S1 by mistake. This in turn has led investigators to continue hypothesizing the protein extends out from the ribosome (63). Here, S1 is visualized as a highly elongated, as well as dynamic, protein. Moreover, our method has shed light on the localization of the individual S1 domains, which is fundamental to understanding this multi-domain protein.

Cross-Linking of Other Ribosomal Proteins—This analysis has additionally yielded structural information about ribosomal proteins other than protein S1. We identified 648 cross-link spectra matches not involving protein S1 with a FDR of ~0.1% (Tables S4 and S5). From these matches, 52 intraprotein and 25 interprotein linkages were identified (supplemental Table S6). For 545 of the 648 cross-link spectra matches, linked residues are both precisely defined by fragmentation and present in the crystal structure of the *E. coli* 30S subunit (16). The structural context of these data, and the 61 linkages they indicate, are provided in Fig. 7*A*. This depiction of the 30S subunit displays the backbones of the proteins and rRNA transparently to facilitate visualizing the α-carbon to α-carbon distances of the identified linkages. Notably, there appears to be good agreement between these cross-linking data and the crystal structure of the 30S subunit. For instance, 525 of the matches (96% of the data) correspond to inter-residue α-carbon to α-carbon distances less than or equal to 26Å, the maximum distance DEST is capable of spanning plus the coordinate error for the crystal structure (29). The remaining 20 matches indicate 7 linkages, 4 measuring between 30 and 35Å and 3 measuring 40, 46 and 76Å. Based on the FDR of this analysis, it is obviously unlikely for even a few of these identifications, let alone all of them, to have been miss-assigned. Many of the linkages were actually detected numerous times, and the cross-linked peptide mass spectra were comparable in quality to other spectra that defined the location of S1. Furthermore, it is highly curious that these anomalous matches involve linkages at or near the identified S1 binding site and that most involve one particular residue. The cross-linking pattern in this region between the

platform and head/body of the 30S subunit suggests movement of the subunit domains with respect to one another and their positions in the crystal structure. A swiveling of the head domain along with movement of the platform toward the body would be consistent with the observed cross-linking (Fig. 7B).

As protein S1 is absent from the crystal structure, we investigated whether the anomalous cross-links may have occurred as a result of S1 inducing a conformational change in the subunit. Accordingly, 30S subunits depleted of S1 were subjected to the same analysis. It has been shown that S1 can be selectively removed from ribosomes by capturing the protein on polyU Sepharose and subsequently displacing it from the ribosome with 1 M ammonium chloride (11). *E. coli* ribosomes obtained in this manner are fully competent with respect to translation, so long as S1 is added back to the system (6, 11). supplemental Fig. S2 displays the LC-MS chromatogram of ribosomal proteins extracted from 30S subunits depleted of S1. Unlike in Fig. 1, a peak for S1 was not detected, demonstrating that S1 had indeed been removed from the subunit. supplemental Fig. S2 also shows that the polyU Sepharose chromatography did not significantly affect any of the other ribosomal proteins. Cross-linking of 30S subunits depleted of S1 in this manner also yielded the anomalous cross-links near the S1 binding site. Thus, there is no indication that S1 induces a conformational change to the subunit.

The cross-links appearing to be anomalous may simply reflect differences between the cross-linking and x-ray crystallography experiments. The only crystal structure that exists contains the *E. coli* 30S subunit bound to a 50S subunit. Our experiments were performed with free 30S subunits. It is worth noting that some conformational differences exist between the crystal structures of the 50S bound 30S subunit from *E. coli* and a free 30S subunit from *T. thermophilus*. The structures in fact appear to differ by up to 10 Å in the lower region of the body domain (15, 16). More interestingly, comparison of these and other crystal structures shows that the head domain of the subunit can be captured in surprisingly different angles of rotation, even within the same crystallographic unit. The head domain has actually been found to rotate up to 12°, involving 20 Å movement (16). In addition, cryo-EM of isolated 30S subunits has demonstrated there to be pronounced structural flexibility among each major domain of the 30S subunit, including the head, platform and body. Conformers narrowed by up to 15° between the platform and head domains have been detected (64, 65).

It appears then that the observed cross-links provide evidence of structural differences between free and 50S bound 30S subunits and the remarkable structural flexibility of the subunit's platform and head domains. Our results suggest these domains are capable of moving 30–50Å with respect to one another. So, although cross-linking will never have the spatial resolution of x-ray crystallography, the approach

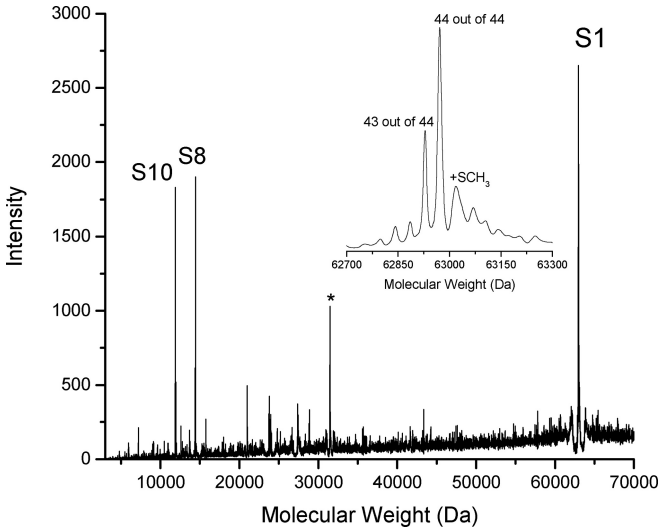

FIG. 8. **Deconvoluted ESI mass spectrum of protein S1 labeled by SMTA.** The mass region for S1 is expanded to illustrate the number of amines labeled. The $+SCH_3$ label marks S1 modified at each of its primary amines as well as methylthiolated on one of its cysteines (44). The asterisk marks a harmonic artifact resulting from deconvolution of the S1 mass.

seems to more readily capture the dynamical processes of biomolecules.

In earlier work on the 70S ribosome, we did not observe any evidence of anomalously long cross linkages between the domains of the 30S subunit (29). We cannot however draw a fair comparison between current and previous studies, given that our current methods are more sensitive and are applied to the significantly less complex 30S structure. Even so, it seems reasonable to suggest that domain movements are less restricted in the isolated rather than 50S bound 30S subunit.

Chemical Labeling of S1 and the 30S Subunit—In an attempt to derive complementary footprinting information, we employed chemical labeling with SMTA (S-methylthioacetimidate)(41) to investigate the S1–30S interaction. This monofunctional reagent reacts with protein primary amines and introduces amidino modifications corresponding to mass shifts of 41 Da. SMTA has been previously employed to probe the structures of a number of proteins and bacterial ribosomes (35, 36, 39, 40, 44, 66). In this experiment, we determined the number of modifications incorporated into proteins from 30S subunits containing and not containing S1. The results appear in supplemental Table S7 and supplemental Fig. S3. The observed extents of modification are indicative of the ribosome's structure, as previous comparisons of chemical labeling and crystal structure data have demonstrated (39). The number of protein modifications is, however, independent of the presence or absence of S1. The labeling of protein S1, itself, proved to be even more interesting. Fig. 8 displays the mass spectrum of S1 labeled with SMTA while in the presence of the 30S subunit. The expanded spectrum

shows that the majority of S1 labeled at every one of its 44 primary amino groups. Given this extent of labeling, there seemed little reason to interrogate individual residues by peptide MS/MS. Most ribosomal proteins are not fully amidinated under these modification conditions, (35, 39, 40, 66) making this result rather striking. Because S1 is particularly large, one would expect shielding of at least some amines. Lack of any shielding is another indication that S1 is highly elongated and dynamic.

The labeling of S1 is reminiscent of L12 from the large subunit stalk complex, as the majority of L12 also becomes fully amidinated on SMTA labeling (39). L12, like S1, is believed to be highly elongated and dynamic, (67) though it plays a role in the binding of GTPase translation factors, (67–71) not mRNA. It may be possible to elucidate the dynamic structures and interactions of both these proteins via labeling approaches designed to capture kinetic properties or to occur on very short (μsec) time scales, such as with pulse-chase amidination (72) or hydroxyl radical labeling (73, 74).

CONCLUSIONS

Through the application of protein cross-linking and high resolution mass spectrometry, we have detailed the ribosomal binding site of S1 and have observed evidence of its dynamics. The localization of its individual domains reveals insights into the structure and interactions of the protein. The results are consistent with the notion that the N-terminal region of S1 binds the ribosome whereas the C-terminal region binds mRNA (10–13). S1 binds to the 30S subunit near the anti-Shine-Dalgarno motif at the 3′ end of the 16S rRNA, is highly elongated even when bound to the ribosome, and has a C-terminal RNA binding region with a cross-linking signature that suggests it is remarkably dynamic. It is therefore apparent that ribosomal protein S1 is structurally poised, as previously hypothesized, to act as an mRNA catching arm in the prokaryotic ribosome (6).

Acknowledgments—We thank Salman Tahir and Lutz Fischer for their assistance with Xi, as well as Loubna Hammad and Jonathan Karty for providing access to and assistance with the LTQ-Orbitrap. In addition, we would like to thank Qi-Zhuang Ye from IUPUI for his gift of *E. coli* MRE 600.

* This work was supported by the National Science Foundation Grants CHE-1012855 and CHE-0832651 and the Wellcome Trust (Senior research fellowship 084229 and Centre core grants 077707 and 092076).

[S] This article contains supplemental Information, Schemes S1 to S4, Tables S1 to S7 and Figs. S1 to S33.

¶ To whom correspondence should be addressed: Department of Chemistry, Indiana University, Bloomington, IN 47405. Tel.: (812) 855-1980; E-mail: reilly@indiana.edu.

REFERENCES

1. Laursen, B. S., Sørensen, H. P., Mortensen, K. K., and Sperling-Petersen, H. U. (2005) Initiation of protein synthesis in bacteria. *Microbiol. Mol. Biol. Rev.* **69,** 101–123
2. Shine, J., and Dalgarno, L. (1974) 3′-Terminal sequence of Escherichia coli 16S ribosomal RNA. Complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci. U.S.A.* **71,** 1342–1346
3. Salah, P., Bisaglia, M., Aliprandi, P., Uzan, M., Sizun, C., and Bontems, F. (2009) Probing the relationship between Gram-negative and Gram-positive S1 proteins by sequence analysis. *Nucleic Acids Res.* **37,** 5578–5588
4. Boni, I. V., Isaeva, D. M., Musychenko, M. L., and Tsareva, N. V. (1991) Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. *Nucleic Acids Res.* **19,** 155–162
5. Tzareva, N. V., Makhno, V. I., and Boni, I. V. (1994) Ribosome-messenger recognition in the absence of the Shine-Dalgarno interactions. *FEBS Lett.* **337,** 189–194
6. Subramanian, A. R. (1983) Structure and functions of ribosomal protein S1. *Prog. Nucleic Acids Res. Mol. Biol.* **28,** 101–142
7. Labischinski, H., and Subramanian, A. R. (1979) Protein S1 from Escherichia coli ribosomes: an improved isolation procedure and shape determination by small-angle X-ray scattering. *Eur. J. Biochem.* **95,** 359–366
8. Kitakawa, M., and Isono, K. (1982) An amber mutation in the gene rpsA for ribosomal protein S1 in Escherichia coli. *Mol. Gen. Genet.* **185,** 445–447
9. Sørensen, M. A., Fricke, J., and Pedersen, S. (1998) Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in Escherichia coli in vivo. *J. Mol. Biol.* **280,** 561–569
10. Giorginis, S., and Subramanian, A. R. (1980) The major ribosome binding site of Escherichia coli ribosomal protein S1 is located in its N-terminal segment. *J. Mol. Biol.* **141,** 393–408
11. Subramanian, A. R., Rienhardt, P., Kimura, M., and Suryanarayana, T. (1981) Fragments of ribosomal protein S1 and its mutant form m1-S1. Localization of nucleic-acid-binding domain in the middle region of S1. *Eur. J. Biochem.* **119,** 245–249
12. Thomas, J. O., Boublik, M., Szer, W., and Subramanian, A. R. (1979) Nucleic acid binding and unfolding properties of ribosomal protein S1 and the derivatives S1-F1 and m1-S1. *Eur. J. Biochem.* **102,** 309–314
13. Suryanarayana, T., and Subramanian, A. R. (1979) Functional domains of Escherichia coli ribosomal protein S1. Formation and characterization of a fragment with ribosome-binding properties. *J. Mol. Biol.* **127,** 41–54
14. Katunin, V. I., Semenkov, Y. P., Makhno, V. I., and Kirillov, S. V. (1980) Comparative study of the interaction of polyuridylic acid with 30S subunits and 70S ribosomes of Escherichia coli. *Nucleic Acids Res.* **8,** 403–421
15. Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Jr., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) Structure of the 30S ribosomal subunit. *Nature* **407,** 327–339
16. Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. (2005) Structures of the bacterial ribosome at 3.5 A resolution. *Science* **310,** 827–834
17. Dunkle, J. A., Wang, L., Feldman, M. B., Pulk, A., Chen, V. B., Kapral, G. J., Noeske, J., Richardson, J. S., Blanchard, S. C., and Cate, J. H. D. (2011) Structures of the Bacterial Ribosome in Classical and Hybrid States of tRNA Binding. *Science* **332,** 981–984
18. Sengupta, J., Agrawal, R. K., and Frank, J. (2001) Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 11991–11996
19. Boileau, G., Sommer, A., and Traut, R. R. (1981) Identification of proteins at the binding site for protein S1 in 70S ribosomes and 30S subunits by cross-linking with 2-iminothiolane. *J. Biol. Chem.* **256,** 8222–8227
20. Agalarov, S. Ch., Kalinichenko, A. A., Kommer, A. A., and Spirin, A. S. (2006) Ribosomal protein S1 induces a conformational change of the 30S ribosomal subunit. *FEBS Lett.* **580,** 6797–6799
21. Das, M., and Fox, C. F. (1979) Chemical cross-linking in biology. *Annu. Rev. Biophys. Bioeng.* **8,** 165–193
22. Ludueña, R. F., Roach, M. C., Trcka, P. P., and Weintraub, S. (1981) N,N-Bis(alpha-iodoacetyl)-2,2′-dithiobis(ethylamine), a reversible cross-linking reagent for protein sulfhydryl groups. *Anal. Biochem.* **117,** 76–80
23. Rappsilber, J. (2011) The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* **173,** 530–540
24. Singh, P., Panchaud, A., and Goodlett, D. R. (2010) Chemical Cross-Linking and Mass Spectrometry As a Low-Resolution Protein Structure Determination Technique. *Anal. Chem.* **82,** 2636–2642
25. Leitner, A., Walzthoeni, T., Kahraman, A., Herzog, F., Rinner, O., Beck, M., and Aebersold, R. (2010) Probing native protein structures by chemical

cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell. Proteomics* **9,** 1634–1649
26. Sinz, A. (2006) Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom. Rev.* **25,** 663–682
27. Petrotchenko, E. V., and Borchers, C. H. (2010) Crosslinking combined with mass spectrometry for structural proteomics. *Mass Spectrom. Rev.* **29,** 862–876
28. Rasmussen M. I., Refsgaard J. C., Peng, L., Houen, G., and Hojrup, P. (2011) CrossWork: software-assisted identification of cross-linked peptides. *J Proteomics* **74,** 1871–1883
29. Lauber, M. A., and Reilly, J. P. (2011) Structural Analysis of a Prokaryotic Ribosome Using a Novel Amidinating Crosslinker and Mass Spectrometry. *J. Proteome Res.* **10,** 3604–3616
30. Lauber, M. A., and Reilly, J. P. (2010) Novel Amidinating Cross-Linker for Facilitating Analyses of Protein Structures and Interactions. *Anal. Chem.* **82,** 7736–7743
31. Cammack, K. A., and Wade, H. E. (1965) The sedimentation behaviour of ribonuclease-active and -inactive ribosomes from bacteria. *Biochem. J.* **96,** 671–680
32. Spedding, G. (1990). in *Ribosomes and Protein Synthesis: A Pratical Approach* (Rickwood, D., and Hames, B. D. eds.), Oxford Press, New York. pp 1–29
33. Karty, J. A., Running, W. E., and Reilly, J. P. (2007) Two dimensional liquid phase separations of proteins using online fractionation and concentration between chromatographic dimensions. *J. Chromatogr. B.* **847,** 103–113
34. Chen, Z. A., Jawhari, A., Fischer, L., Buchen, C., Tahir, S., Kamenski, T., Rasmussen, M., Lariviere, L., Bukowski-Wills, J. C., Nilges, M., Cramer, P., and Rappsilber, J. (2010) Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* **29,** 717–726
35. Running, W. E., and Reilly, J. P. (2009) Ribosomal Proteins of Deinococcus radiodurans: Their Solvent Accessibility and Reactivity. *J. Proteome Res.* **8,** 1228–1246
36. Liu, X., Broshears, W. C., and Reilly, J. P. (2007) Probing the structure and activity of trypsin with amidination. *Anal. Biochem.* **367,** 13–19
37. Choi, H., and Nesvizhskii, A. I. (2008) False Discovery Rates and Related Statistical Concepts in Mass Spectrometry-Based Proteomics. *J. Proteome Res.* **7,** 47–50
38. DeLano, W. L. (2002) The PyMOL Molecular Graphics System. DeLano Scientific, Palo Alto, CA, U.S.A.
39. Liu, X., and Reilly, J. P. (2009) Correlating the Chemical Modification of Escherichia coli Ribosomal Proteins with Crystal Structure Data. *J. Proteome Res.* **8,** 4466–4478
40. Beardsley, R. L., Running, W. E., and Reilly, J. P. (2006) Probing the Structure of the Caulobacter crescentus Ribosome with Chemical Labeling and Mass Spectrometry. *J. Proteome Res.* **5,** 2935–2946
41. Thumm, M., Hoenes, J., and Pfleiderer, G. (1987) S-Methylthioacetimidate is a new reagent for the amidination of proteins at low pH. *Biochim. Biophys. Acta* **923,** 263–267
42. Hunter, M. J., and Ludwig, M. L. (1962) The reaction of imidoesters with proteins and related small molecules. *J. Am. Chem. Soc.* **84,** 3491–3504
43. Hermanson, G. T. (2008) *Bioconjugate Techniques*, 2nd ed., Academic Press, San Diego
44. Janecki, D. J., Beardsley, R. L., and Reilly, J. P. (2005) Probing Protein Tertiary Structure with Amidination. *Anal. Chem.* **77,** 7274–7281
45. Draper, D. E., and Von Hippel, P. H. (1979) Interaction of Escherichia coli ribosomal protein S1 with ribosomes. *Proc. Natl. Acad. Sci. U.S.A.* **76,** 1040–1044
46. Arnold, R. J., and Reilly, J. P. (1999) Observation of Escherichia coli ribosomal proteins and their posttranslational modifications by mass spectrometry. *Anal. Biochem.* **269,** 105–112
47. Kowalak, J. A., and Walsh, K. A. (1996) β-Methylthio-aspartic acid: identification of a novel posttranslational modification in ribosomal protein S12 from Escherichia coli. *Protein Sci.* **5,** 1625–1632
48. Kamp, R., and Wittmann-Liebold, B. (1980) Primary structure of protein S11 from Escherichia coli ribosomes. *FEBS Lett.* **121,** 117–122
49. Wittmann-Liebold, B., and Greuer, B. (1978) The primary structure of protein S5 from the small subunit of the Escherichia coli ribosome. *FEBS Lett.* **95,** 91–98
50. Hitz, H., Schäfer, D., and Wittmann-Liebold, B. (1977) Determination of the complete amino-acid sequence of protein S6 from the wild-type and a mutant of Escherichia coli. *Eur. J. Biochem.* **75,** 497–512
51. Yaguchi, M. (1975) Primary structure of protein S18 from the small Escherichia coli ribosomal subunit. *FEBS Lett.* **59,** 217–220
52. Strader, M. B., Costantino, N., Elkins, C. A., Chen, C. Y., Patel, I., Makusky, A. J., Choy, J. S., Court, D. L., Markey, S. P., and Kowalak, J. A. (2011) A proteomic and transcriptomic approach reveals new insight into beta-methylthiolation of Escherichia coli ribosomal protein S12. *Mol. Cell. Proteomics* **10,** M110.005199
53. Apweiler, R., Martin, M. J., O'Donovan, C., Magrane, M., Alam-Faruque, Y., Antunes, R., Casanova, E. B., Bely, B., Bingley, M., Bower, L., Tursteinas, B., Chan, W. M., Chavali, G., Da Silva, A., Dimmer, E., Eberhardt, R., Fazzini, F., Fedotov, A., Garavelli, J., Garcia Castro, L., Gardner, M., Hieta, R., Huntley, R., Jacobsen, J., Legge, D., Liu, W., Luo, J., Orchard, S., Patient, S., Pichler, K., Poggioli, D., Pontikos, N., Pundir, S., Rosanoff, S., Sawford, T., Sehra, H., Turner, E., Wardell, T., Watkins, X., Corbett, M., Donnelly, M., van Rensburg, P., Goujon, M., McWilliam, H., Lopez, R., Xenarios, i., Bougueleret, L., Bridge, A., Poux, S., Redaschi, N., Argoud-Puy, G., Auchincloss, A., Axelsen, K., Baratin, D., Blatter, M.-C., Boeckmann, B., Bolleman, j., Bollondi, L., Boutet, E., Braconi Quitaje, S., Breuza, L., deCastro, E., Cerutti, L., Coudert, E., Cuche, B., Cusin, I., Doche, M., Dornevil, D., Duvaud, S., Estreicher, A., Famiglietti, L., Feuermann, M., Gehant, S., Ferro, S., Gasteiger, E., Gerritsen, V., Gos, A., Gruaz-Gumowski, N., Hinz, U., Hulo, C., Hulo, N., James, J., Jimenez, S., Jungo, F., Kappler, T., Keller, G., Lara, V., Lemercier, P., Lieberherr, D., Martin, X., Masson, P., Mooinat, M., Morgat, A., Paesano, S., Pedruzzi, I., Pilbout, S., Pozzato, M., Pruess, M., Rivoire, C., Roechert, B., Schneider, M., Sigrist, C., Sonesson, K., Staehli, S., Stanley, E., Stutz, A., Sundaram, S., Tognolli, M., Verbregue, L., Veuthey, A.-L., Wu, C. H., Arighi, C. N., Arminski, L., Barker, W. C., Chen, C., Chen, Y., Dubey, P., Huang, H., Kukreja, A., Laiho, K., Mazumder, R., McGarvey, P., Natale, D. A., Natarajan, T. G., Roberts, N. V., Suzek, B. E., Vinayaka, C. R., Wang, Q., Wang, Y., Yeh, L.-S., and Zhang, J. (2012) Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res.* **40,** D71–D75
54. Touchon, M., Hoede, C., Tenaillon, O., Barbe, V., Baeriswyl, S., Bidet, P., Bingen, E., Bonacorsi, S., Bouchier, C., Bouvet, O., Calteau, A., Chiapello, H., Clermont, O., Cruveiller, S., Danchin, A., Diard, M., Dossat, C., Karoui, M. E., Frapy, E., Garry, L., Ghigo, J. M., Gilles, A. M., Johnson, J., Le Bouguenec, C., Lescat, M., Mangenot, S., Martinez-Jéhanne, V., Matic, I., Nassif, X., Oztas, S., Petit, M. A., Pichon, C., Rouy, Z., Ruf, C. S, Schneider, D., Tourret, J., Vacherie, B., Vallenet, D., Médigue, C., Rocha, E. P., and Denamur, E. (2009) Organised genome dynamics in the Escherichia coli species results in highly diverse adaptive paths. *PLoS Genet.* **5,** e1000344
55. Rasko, D. A., Rosovitz, M. J., Myers, G. S., Mongodin, E. F., Fricke, W. F., Gajer, P., Crabtree, J., Sebaihia, M., Thomson, N. R., Chaudhuri, R., Henderson, I. R., Sperandio, V., and Ravel, J. (2008) The pangenome structure of Escherichia coli: comparative genomic analysis of E. coli commensal and pathogenic isolates. *J. Bacteriol.* **190,** 6881–6893
56. Leitner, A., Reischl, R., Walzthoeni, T., Herzog, F., Bohn, S., Forster, F., and Aebersold, R. (2012) Expanding the Chemical Cross-Linking Toolbox by the Use of Multiple Proteases and Enrichment by Size Exclusion Chromatography. *Mol. Cell. Proteomics* **11,** M111.014126
57. Kao, A., Chiu, C.-l., Vellucci, D., Yang, Y., Patel, V. R., Guan, S., Randall, A., Bald, P., Rychnovsky, S. D., and Huang, L. (2011) Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. *Mol. Cell. Proteomics* **10,** M110.002212
58. Trnka, M. J., and Burlingame, A. L. (2010) Topographic studies of the GroEL-GroES chaperonin complex by chemical cross-linking using diformyl ethynylbenzene. The power of high resolution electron transfer dissociation for determination of both peptide sequences and their attachment sites. *Mol. Cell. Proteomics* **9,** 2306–2317
59. Buncherd, H., Nessen, M. A., Nouse, N., Stelder, S. K., Roseboom, W., Dekker, H. L., Arents, J. C., Smeenk, L. E., Wanner, M. J., van Maarseveen, J. H., Yang, X., Lewis, P. J., de Koning, L. J., de Koster, C. G., and de Jong, L. (2012) Selective enrichment and identification of cross-linked peptides to study 3-D structures of protein complexes by mass spectrometry. *J. Proteomics* **75,** 2205–2215

60. Selmer, M., Dunham, C. M., Murphy, F. V.,4th, Weixlbaumer, A., Petry, S., Kelley, A. C., Weir, J. R., and Ramakrishnan, V. (2006) Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313,** 1935–1942
61. Yusupova, G. Z., Yusupov, M. M., Cate, J. H., and Noller, H. F. (2001) The path of messenger RNA through the ribosome. *Cell* **106,** 233–241
62. Yusupova, G., Jenner, L., Rees, B., Moras, D., and Yusupov, M. (2006) Structural basis for messenger RNA movement on the ribosome. *Nature* **444,** 391–394
63. Dunkle, J. A., and Cate, J. H. D. (2011) The Packing of Ribosomes in Crystals and Polysomes. in *Ribosomes: Structure, Function, and Dynamics* (Marina Rodnina, Wolfgang Wintermeyer, and Green, R. eds.), Springer-Verlag/Wien, New York. pp 65–73
64. Gabashvili, I. S., Agrawal, R. K., Grassucci, R., and Frank, J. (1999) Structure and structural variations of the Escherichia coli 30 S ribosomal subunit as revealed by three-dimensional cryo-electron microscopy. *J. Mol. Biol.* **286,** 1285–1291
65. Lata, K. R., Agrawal, R. K., Penczek, P., Grassucci, R., Zhu, J., and Frank, J. (1996) Three-dimensional reconstruction of the Escherichia coli 30 S ribosomal subunit in ice. *J. Mol. Biol.* **262,** 43–52
66. Lauber, M. A., Running, W. E., and Reilly, J. P. (2009) B. subtilis Ribosomal Proteins: Structural Homology and Post-Translational Modifications. *J. Proteome Res.* **8,** 4193–4206
67. Diaconu, M., Kothe, U., Schlünzen, F., Fischer, N., Harms, J. M., Tonevitsky, A. G., Stark, H., Rodnina, M. V., and Wahl, M. C. (2005) Structural basis for the function of the ribosomal L7/12 stalk in factor binding and GTPase activation. *Cell* **121,** 991–1004
68. Savelsbergh, A., Mohr, D., Kothe, U., Wintermeyer, W., and Rodnina, M. V. (2005) Control of phosphate release from elongation factor G by ribosomal protein L7/12. *EMBO J.* **24,** 4316–4323
69. Gao, Y. G., Selmer, M., Dunham, C. M., Weixlbaumer, A., Kelley, A. C., and Ramakrishnan, V. (2009) The Structure of the Ribosome with Elongation Factor G Trapped in the Posttranslocational State. *Science* **326,** 694–699
70. Helgstrand, M., Mandava, C. S., Mulder, F. A., Liljas, A., Sanyal, S., and Akke, M. (2007) The Ribosomal Stalk Binds to Translation Factors IF2, EF-Tu, EF-G and RF3 via a Conserved Region of the L12 C-terminal Domain. *J. Mol. Biol.* **365,** 468–479
71. Mohr, D., Wintermeyer, W., and Rodnina, M. V. (2002) GTPase Activation of Elongation Factors Tu and G on the Ribosome. *Biochemistry* **41,** 12520–12528
72. Chang, F. M., Lauber, M. A., Running, W. E., Reilly, J. P., and Giedroc, D. P. (2011) Ratiometric Pulse-Chase Amidination Mass Spectrometry as a Probe of Biomolecular Complex Formation. *Anal. Chem.* **83,** 9092–9099
73. Stocks, B. B., and Konermann, L. (2009) Structural Characterization of Short-Lived Protein Unfolding Intermediates by Laser-Induced Oxidative Labeling and Mass Spectrometry. *Anal. Chem.* **81,** 20–27
74. Konermann, L., Stocks, B. B., Pan, Y., and Tong, X. (2010) Mass Spectrometry Combined with Oxidative Labeling for Exploring Protein Structure and Folding. *Mass Spectrom. Rev.* **29,** 651–667