# Structural Analysis of a Prokaryotic Ribosome Using a Novel Amidinating Cross-Linker and Mass Spectrometry

Matthew A. Lauber and James P. Reilly*

Department of Chemistry, Indiana University, Bloomington, Indiana 47405, United States

**S** *Supporting Information*

**ABSTRACT:** The structure of the *Escherichia coli* ribosome, a 2.5 MDa ribonucleoprotein complex containing more than 50 proteins, was probed using the novel amidinating cross-linker diethyl suberthioimidate (DEST) and mass spectrometry. Peptide cross-links derived from this complex structure were identified at high confidence (FDR 0.8%) from precursor mass measurements and collision-induced dissociation (CID) fragmentation spectra. The acquired cross-linking data were found to be in excellent agreement with the crystal structure of the *E. coli* ribosome. DEST cross-links are particularly amenable to strong cation exchange (SCX) chromatography, facilitating a large-scale analysis. SCX enrichment and fractionation were shown to increase the number of cross-link spectra matches in our analysis 10-fold. Evidence is presented that these techniques can be used to study complex interactomes.



**KEYWORDS:** Protein cross-linking, mass spectrometry, amidination, diethyl suberthioimidate, DEST, ribosome, *E. coli*
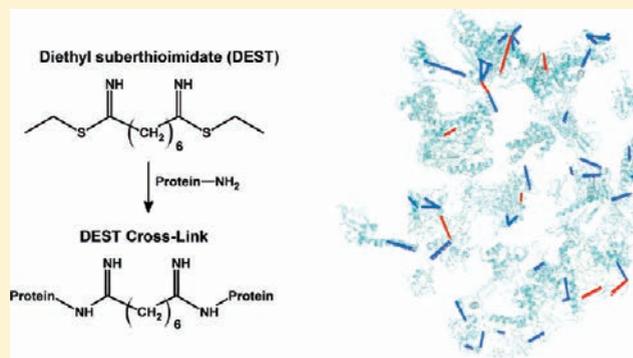
## INTRODUCTION

Protein—protein interactions are required for most biological functions. The abilities of RNA polymerase to transcribe DNA,[1] the spliceosome to process gene transcripts,[2] and the ribosome to translate mRNA,[3] in particular, require the orchestrated interactions of many proteins. Unfortunately, the quaternary structures that enable these interactions have been challenging to study. For example, although X-ray crystallography can provide atomic resolution, it is often very difficult to produce and crystallize large macromolecular complexes due to their dynamic nature.[4] It has therefore become apparent that complementary methods are needed to best elucidate the structures and interactions of macromolecular complexes.[4]

Chemical cross-linking can be applied to study quaternary structures and thus aid the development of hybrid structure models. It can capture dynamic interactions without being markedly hindered by sample structural heterogeneity. Modern mass spectrometers have even made it possible to identify cross-links and the exact residues involved in linkages with reasonably high confidence.[5,6] In most cross-linking analyses, protein functional groups are targeted for derivatization with a molecule that contains two reactive groups separated by a spacer arm of known length. Only functional groups closer than the length of the spacer arm are capable of being linked. Identification of the residues involved in a cross-link thereby provides distance constraints for structural modeling.

Despite these advantages, the potential of cross-linking studies can be limited by several analytical challenges. Detection of cross-linked peptides in the proteolytic digests of derivatized proteins is often impeded by the combination of their low stoichiometric yield and the presence of other peptide species. Enrichment techniques have accordingly become an intense focus for researchers in the field.[7−9] Due to the combinatorial nature of cross-linking reactions, mixtures containing cross-linked peptides can be extremely heterogeneous. Their fractionation has therefore been necessary in the few large-scale cross-linking studies conducted to date.[10,11] Lastly, identification of cross-links typically requires mass accuracies on the order of a few parts per million if large sequence databases are to be searched.

To facilitate cross-linking studies, we previously developed the novel amidinating cross-linker, diethyl suberthioimidate (DEST).[12] This reagent is a water-soluble, eight-atom (11 Å) spacer arm cross-linker, similar to commonly used and commercially available cross-linkers, such as bissulfosuccinimidyl suberate (BS3). Unlike its analogs, however, DEST modifies primary amines at physiological pH without sacrificing their native basicity. As a result, the use of DEST does not perturb the electrostatic properties of a protein and is thus unlikely to disrupt native protein structure. The fact that this reagent preserves the basicity of targeted amines upon modification also means that the ionization efficiency of the residue it modifies is not adversely affected and that the cross-links it forms can be easily separated from other components of tryptic digests using strong

cation exchange (SCX) chromatography. We have previously employed this cross-linker to observe intraprotein cross-linking of cytochrome $c$[12] and have subsequently become interested in applying it to study very complex structures, as a step toward understanding their dynamic interactomes.

One such complex of interest is the *Escherichia coli* ribosome, the cellular machine responsible for translating mRNA into protein. It is a 2.5 MDa ribonucleoprotein complex composed of 53 unique proteins and is one of a few large macromolecular complexes with a structure, albeit incomplete, determined by X-ray crystallography. The *E. coli* ribosome, as well as other prokaryotic ribosomes, is thus a rather unique system in which to investigate the feasibility of large-scale mass spectrometric studies. In previous work, we have demonstrated the effectiveness of monofunctional thiomidates as chemical labeling reagents by probing the tertiary and quaternary structures of the proteins in prokaryotic ribosomes and demonstrating that the labeling of proteins correlates extremely well with predictions derived from crystal structures.[13−17] The study of ribosome cross-linking is far more challenging. Although there are only 678 unique modifiable amines in the sample, the combinatorial nature of cross-linking means there are about 230 000 potential modification products with which to be concerned.

In the present work, we have probed the structure of the *E. coli* ribosome using the DEST cross-linker and mass spectrometry. This is one of the most complex structures to which cross-linking combined with mass spectrometry has been applied. Peptide cross-links derived from this structure were identified with high confidence from single precursor mass measurements and collision-induced dissociation (CID) fragmentation spectra. Data could be directly compared with the crystal structure of the *E. coli* ribosome. This study also demonstrates the compatibility of DEST cross-linking with cation exchange chromatography. Finally, evidence of ribosome−ribosome interactions in the data suggests that these techniques constitute a viable method for studying complex interactomes.

## ■ EXPERIMENTAL PROCEDURES

### Materials

Acetonitrile (ACN), glacial acetic acid, hydrochloric acid, sodium hydroxide, trifluoroacetic acid (TFA), and water were purchased from EMD Chemicals (Gibbstown, NJ). Anhydrous diethyl ether was obtained from Fisher (Fair Lawn, NJ). Proteomics grade trypsin (T-6567), magnesium acetate tetrahydrate, Trizma base, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), phenylmethylsulfonyl fluoride (PMSF), and sucrose were purchased from Sigma (St. Louis, MO). Ammonium chloride, formic acid (FA), magnesium chloride hexahydrate, 2-mercaptoethanol, and suberonitrile were obtained from Aldrich (Milwaukee, WI). Calcium chloride dihydrate, dichloromethane, (ethylenedinitrilo)tetraacetic acid (EDTA), sodium chloride, sodium phosphate monobasic, and type 3A molecular sieves were purchased from Mallinckrodt Baker (Phillipsburg, NJ). Ethanethiol was obtained from Acros Organics (Pittsburgh, PA). Hydrogen chloride (technical grade) was purchased from Matheson (Cucamonga, CA). Bacto-tryptone and bacto-yeast extract were provided by Becton, Dickinson and Company (Sparks, MD). Complete EDTA-free Protease Inhibitor ("Mini") Tablets were supplied by Roche Applied Science (Indianapolis, IN). Rapigest SF was purchased from Waters (Milford, MA).

### Synthesis of Diethyl Suberthioimidate (DEST)

As described previously,[12] diethyl suberthioimidate (DEST) was prepared from ethanethiol and suberonitrile via the Pinner synthesis (Supporting Information, Scheme 1).[18] Suberonitrile (9 mmol) in anhydrous dichloromethane (1:3 v/v) was added to ice-cold ethanethiol (90 mmol). The reaction mixture, while being kept on ice and constantly stirred, was sparged with hydrogen chloride gas for 1 h and subsequently kept at 4 °C for an additional 16 h. Anhydrous diethyl ether was then added to aid precipitation, and the mixture was stored at −20 °C until a solid had formed. This solid, after being washed several times with anhydrous diethyl ether, was stored in a vacuum desiccator at 4 °C until needed.

### Preparation of *E. coli* Ribosomes

*E. coli* K12 cells were grown for the preparation of ribosomes. Starter cultures were grown overnight at 37 °C while being aerated at 180 rpm. Ribosomes were isolated from cells as previously described by Spedding.[19,20]

### Cross-Linking Conditions and Sample Preparation

Solutions of ribosomes were exchanged into phosphate-buffered saline (20 mM sodium phosphate/150 mM sodium chloride, pH 7) containing 10 mM magnesium acetate using Amicon Ultra 10K centrifugal filter devices (Millipore, Eschborn, Germany) prior to cross-linking. Reactions for two biological replicates, each with two technical replicates, were carried out at a total protein concentration of 0.5 mg/mL (an approximately 0.7 $\mu$M ribosome solution assuming there is 760 000 g of protein per mole of ribosomes) and DEST concentration of 2.5 mM. The protein concentration of the ribosome solutions was estimated by means of a Bradford assay, in which bovine serum albumin was used as a standard. After proceeding at room temperature for 12 h, the reactions were quenched by addition of 0.5 M Tris to a final concentration of 50 mM.

These are conditions that allow DEST to react until it is near fully hydrolyzed; these are also conditions that yield only partial modification of proteins.[12] Shorter reaction times may be necessary for studying less stable biological systems. To precipitate the rRNA, glacial acetic acid and 1 M magnesium chloride were added to the reaction mixture such that the final solution contained 3:6:1 (v/v/v) reaction mixture/glacial acetic acid/1 M magnesium chloride. The samples were vortexed and allowed to remain at room temperature for 40 min before the rRNA precipitate was separated by centrifugation at 14 100$g$ for 20 min. Supernatant from each reaction mixture, containing soluble cross-linked and unmodified ribosomal proteins, was desalted and cleared of hydrolyzed reagent by exchange into 50 mM Tris using Amicon Ultra 3K centrifugal filter devices (Millipore, Eschborn, Germany). The concentrate was then dried and stored at −20 °C until later sample preparation for proteolytic digestion.

### Proteolytic Digestion

Dried aliquots of DEST-modified proteins (40 $\mu$g) and proteomics grade trypsin (2 $\mu$g) were reconstituted in Rapigest (Waters, Milford, MA) containing solution, such that the digestion was carried out in 18 $\mu$L of 100 mM Tris/10 mM calcium chloride (pH 8) and 0.2% (w/v) Rapigest. 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 1%. To hydrolyze Rapigest, the quenched digests were incubated for

30 min at 37 °C. Insoluble Rapigest byproduct was cleared from the samples by centrifuging at 13 000g.

## SCX Enrichment and Fractionation of DEST Cross-Links

Strong cation exchange (SCX) chromatography was used in two ways to simplify proteolytic digests of cross-linked proteins prior to their analysis by nanoLC-MS/MS. Interpeptide cross-links were enriched from non-cross-linked peptide (linear peptide) species via a low pH separation and then fractionated via an intermediate pH separation (Scheme 2).

Interpeptide cross-links were enriched using SCX chromatography by exploiting the fact that they contain more basic functional groups, and thus positive charges at low pH, than other species in a proteolytic digest. We have previously demonstrated the utility of this approach.[12] As before, tryptic digest of the cross-linked sample was loaded on an SCX column (TSKgel SP-NPR, 4.6 mm × 35 mm, Tosoh Bioscience, Montgomeryville, PA) equilibrated with 0.1% TFA in water (pH 2). Most non-cross-linked (linear peptide) species were eluted from the column with 300 mM NaCl mobile phase. Interpeptide cross-links were then eluted from the SCX column onto a C18 trapping column (Thermo Hypersil-Keystone Javelin, 1.0 mm × 20 mm, Bellefonte, PA) with mobile phase containing 1000 mM NaCl. After being desalted, the contents of the C18 trapping column were eluted with organic mobile phase. This enriched fraction was dried under vacuum and resuspended in 0.1% TFA for subsequent LC-MS/MS analysis.

The amount of sample remaining after the 300 mM salt wash step of the enrichment was estimated via a micro-BCA assay.[21,22] The $A_{562}$ was measured for the sample before it was loaded onto the SCX column as well as after it was processed via the above procedure. Based on two biological replicates, the amount of sample remaining after the salt wash procedure was estimated to be 15% of the original sample mass.

DEST interpeptide cross-links were also subjected to fractionation using SCX chromatography. This chromatography was directly integrated with the enrichment of DEST interpeptide cross-links. After the 300 mM salt wash of the enrichment, the mobile phases were changed to a 20 mM sodium acetate, pH 5, buffer, and a 100 min gradient from 0 to 300 or 400 mM NaCl, followed with a 10 min isocratic hold of 1000 mM NaCl, was implemented. DEST interpeptide cross-links eluting from the SCX column were collected sequentially onto 10 C18 trapping columns (Thermo Hypersil-Keystone Javelin, 1.0 mm × 20 mm, Bellefonte, PA). An apparatus similar to one previously described was used to switch the flow path from one trapping column to the next every 10 min.[23] After being desalted with 0.1% TFA in water at 0.3 mL/min for 20 min, the contents of the C18 trapping columns were sequentially eluted with a 10 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. These fractions of the SCX were then resuspended in 0.1% TFA in water for subsequent LC-MS/MS analysis.

## Capillary LC-ESI-MS/MS

Capillary LC-ESI-MS/MS was performed using an IntegraFrit capillary trapping column packed with 1.5 cm of C18 (150 μm × 11 cm, New Objective, Woburn, MA; Magic C18, 5 μm, 200 Å, Michrom BioResources, Auburn, CA), a capillary analytical column packed with 15 cm of C18 (75 μm × 15 cm, Magic C18, 5 μm, 100 Å, Michrom BioResources, Auburn, CA), an LTQ Orbitrap XL mass spectrometer (Thermo Electron,

Bremen, Germany), and a Dionex chromatography system (Ultimate 3000, Dionex, Sunnyvale, CA). In each experiment, approximately 1 μg of tryptic digest was injected onto a trapping column to remove salts and contaminants by flushing for 10 min with mobile phase A (0.1% FA in 97:3 water/ACN) at a flow rate of 10 μL/min. The flow rate was then reduced to 0.3 μL/min, effluent from the trapping column was directed to the capillary LC column, and a 100-min gradient between 0% and 35% mobile phase B (0.1% FA in ACN) was implemented. Eluting peptides were electrosprayed into an LTQ-Orbitrap mass spectrometer operating in data-dependent mode to acquire a full MS scan (300−1800 $m/z$) and subsequent MS/MS scans of the three most intense precursor ions. The AGC target value was set to $5 \times 10^5$ for MS scans and $2 \times 10^5$ for MS/MS scans. CID of the precursors occurred in the LTQ at 35% normalized collision energy. Isolation width was set to 2 $m/z$ and monoisotopic precursor selection was enabled. Both the MS and MS/MS scans were acquired in the orbitrap with resolution set to 30 000 and 7500, respectively. Dynamic exclusion was employed with the following settings: a 90 s exclusion duration time, maximum exclusion list of 500, and one repeat count. In addition, charge state rejection was enabled for 1+, 2+, and unassigned charge states, except when it was of interest to determine the changes in sample composition due to SCX enrichment. Because cross-linked peptides formed in these experiments tend to have a higher charge state upon ESI than non-cross-linked (linear peptide) species, rejection of low charge states from MS/MS acquisition provides a useful bias for detecting cross-links. MS/MS spectra were subjected to data reduction using Mascot Distiller (version 2.3.2.0), such that precursor masses were redetermined via interpretation of isotopic distributions and MS/MS spectra were deisotoped. Processed data were saved as .dta files then merged into .mgf files with syntax appropriate for database searching.[11]

## xQuest Database Searching

MS/MS data contained in merged .mgf files were subsequently searched against the sequences of constitutive proteins from the E. coli 30S and 50S ribosomal subunits for interpeptide cross-links using the web interface of xQuest (version 2.721).[11] This is an algorithm that can be used to search for theoretical cross-links that have masses matching measured precursor masses and to subsequently assign the fragment masses of MS/MS spectra. In its interpretation of MS/MS spectra, xQuest assumes that a cross-link precursor will fragment at only one peptide bond. For our searches, xQuest was used with default settings, except that MS tolerance was set to 5 ppm, the MS/MS tolerance was set to 0.01 $m/z$, variable modification was set for methionine oxidation (15.9949), no fixed modifications were used, and mass shifts for cross-linking products were manually set to 136.10005 for the "xlink mass-shift" and 154.11061 (ADE), 199.10308 (TEDE), and 257.17394 (Tris dead-end) for the "monolink mass-shifts". Matches from all searches were required to have precursor mass errors ≤4.3 ppm and ≥15% of the ion current in a given MS/MS spectrum assigned as b- and y-type ions. This value was found acceptable, since xQuest does not include ammonia and water loss fragment ions in its scoring. MS/MS matches were also required to have normalized cross-correlation scores ≥0.052 for both ion series containing and not containing the cross-link mass shift.

xQuest matches for intraprotein and interprotein cross-links were required to meet different scoring thresholds.

Intraprotein cross-link matches were required to have xQuest scores $\geq 25.5$. Furthermore, a number of intraprotein cross-link matches were manually inspected to determine whether they indeed provided unambiguous identifications. This was required for special cases in which the detected cross-links were formed from two peptides, which together comprised a contiguous sequence from the protein. 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 types of matches were manually inspected and disregarded if the fragmentation did not unambiguously indicate the cross-link structure.

Interprotein cross-link matches were required to have xQuest scores $\geq 27.3$ and were required to have xQuest scores at least 2 units greater than their corresponding second rank matches. Moreover, interprotein cross-link matches were disregarded unless there were three unique fragments assigned on each peptide chain. Spectra matches were assessed for this characteristic after assigning neutral losses and second isotopic peaks. Fragments that could be ambiguously matched to the sequences of either peptide chain in the cross-link were not included in this count. The cross-link matches were, in addition, required to have at least eight of these sequence-specific fragments assigned.

Finally, spectra that corresponded to cross-links, either intraprotein or interprotein, with multiple possible linkage patterns were manually inspected. Exact linkages were only proposed when the two cross-linked residues could be unequivocally defined by the observed fragmentation.

### False Discovery Rate Analysis

The false discovery rate (FDR) of this analysis was first estimated by searching files that produced matches against randomized sequences of the *E. coli* 30S and 50S ribosomal proteins. This yielded 2 intraprotein and 0 interprotein (non-intersubunit) decoy matches, indicating that the FDR for our identified cross-link spectra matches, whether intraprotein or interprotein, was less than 2%. It was clear, however, that this was not an accurate estimate of the FDR. When there are only a limited number of identified spectra matches, a decoy database that is of equal size to a target database is inadequate for accurately defining a FDR. A solution to this problem is to use a decoy database that is larger than the target database and to scale the number of decoy matches based on the difference in size between the decoy and target search spaces.[24] We constructed a large decoy database that retained the characteristics of the *E. coli* ribosomal proteins by concatenating and subsequently randomizing a database that contained two copies of each protein sequence. Searches against this decoy database yielded 501 and 5 decoy matches meeting the criteria for confident intraprotein and interprotein matches, respectively. FDRs for both the identification of intraprotein and interprotein cross-link matches were calculated in two steps. The first step was to scale the number of decoy matches by how many times larger the decoy search space was than the search space needed to identify either intraprotein or interprotein cross-links. The decoy search space contained 8 555 316 theoretical cross-linked peptides, while the search space needed to identify intraprotein cross-links contained 34 170 theoretical cross-linked peptides. Thus, the 501 decoy matches corresponding to intraprotein cross-link matches were scaled to 2.0. Likewise, the search space needed to identify interprotein (non-intersubunit) cross-links contained 801 888 theoretical cross-linked peptides, so the 5 decoy matches corresponding to the

interprotein cross-link matches were scaled to 0.5. The FDRs were then obtained by dividing the scaled number of decoy matches by the number of matches produced when searching the experimental data against the ribosomal protein sequences.

### Mascot Database Searching

MS/MS data contained in the .mgf files were searched for peptide species other than interpeptide cross-links (namely, linear peptide species) using Mascot (version 2.2.0) with a database containing the sequences of *E. coli* ribosomal proteins. The MS and MS/MS tolerances for these searches were 5 ppm and 0.01 $m/z$, respectively. Numerous variable modifications for protein N-termini and lysine residues were included in these searches, such as the addition of amide dead ends (+154.11061 Da), thioester dead ends (+199.10308), Tris dead ends (+257.17394), and intrapeptide cross-links (+136.10005 Da). Other settings included a variable modification for methionine oxidation and an allowed number of missed tryptic cleavages equal to 6. Ion matches with scores less than 17 ($p > 0.005$) were ignored.

### Protein Sequence Data

The *E. coli* K12 proteome was obtained from The J. Craig Venter Institute (GenBank Accession. Version U00096.2).

### Crystal Structure of the *E. coli* 70S Ribosome

Distances between α-carbons of lysine residues in the *E. coli* ribosome were determined using a program developed in-house and the coordinates from PDB files 2AW4 and 2AVY.[25] PyMOL v. 0.99 (DeLano Scientific, www.pymol.org) was employed for the visualization and manipulation of crystal structures.[26]

## ■ RESULTS AND DISCUSSION

### Cross-Linking of the *E. coli* Ribosome

The *E. coli* ribosome was subjected to cross-linking with DEST, a novel bifunctional thioimidate reagent with an 11 Å spacer arm. Ribosomes prepared at less than intracellular concentrations (ca. 1 $\mu$M versus 10 $\mu$M)[27,28] were modified in a physiologically relevant buffer with DEST present at a 5:1 reagent to protein amine ratio. DEST modifies the primary amines of proteins, thereby introducing a set of amidine-linked reaction products, including cross-links and so-called "dead ends". As noted in Scheme 1, cross-links between two amines introduce 136 Da mass shifts. Dead ends, which form when one end of the reagent reacts with an amine while the other end is hydrolyzed, introduce mass shifts of either 154 or 199 Da, depending on whether hydrolysis eliminates the thiol to form an amide (ADE) or eliminates ammonia to form a thioester (TEDE).

To estimate the number of structurally possible DEST cross-links supported by the *E. coli* ribosome, we investigated the distances between modifiable residues in the previously reported crystal structure.[25] A specific focus was made on lysine residues and not protein N-termini, given that lysine residues significantly out number protein N-termini and that the N-termini of several ribosomal proteins (L11, L16, L13, S5, S11, and S18) are not available for amidination due to post-translational methylation or acetylation.[15,20] The *E. coli* ribosome contains 631 unique, modifiable lysine residues; 548 of these are modeled in the crystal structure. We calculated the distances between α-carbons of these residues and counted the number of values that were less than the maximum α-carbon to α-carbon cross-linking distance that DEST is capable of bridging. We defined this to be 24 Å, the

### Scheme 1. Major Reaction Products of DEST Cross-Linking
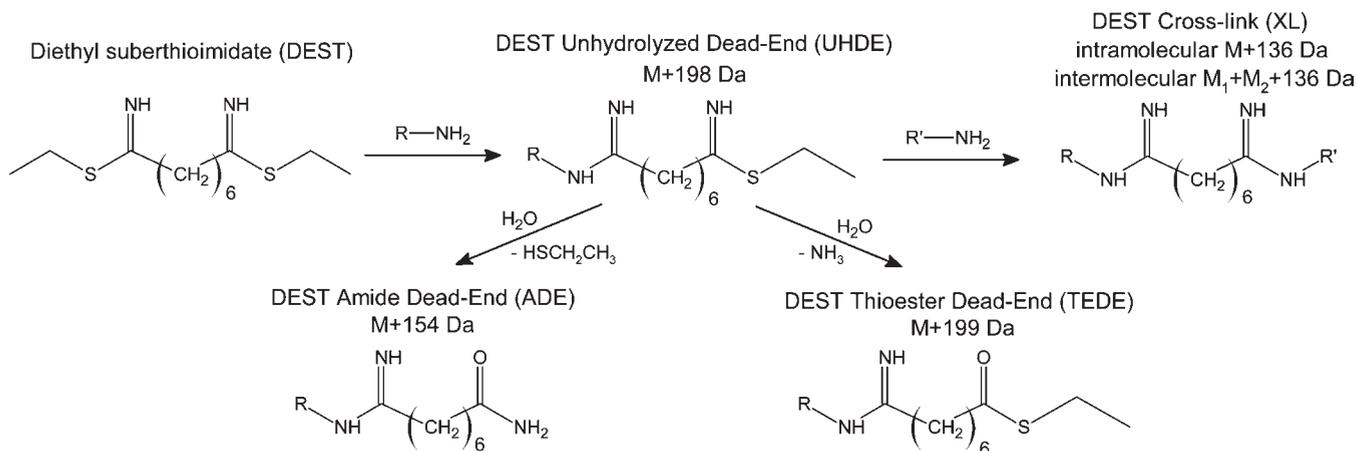


### Scheme 2. SCX Enrichment and Fractionation of DEST Interpeptide Cross-Links
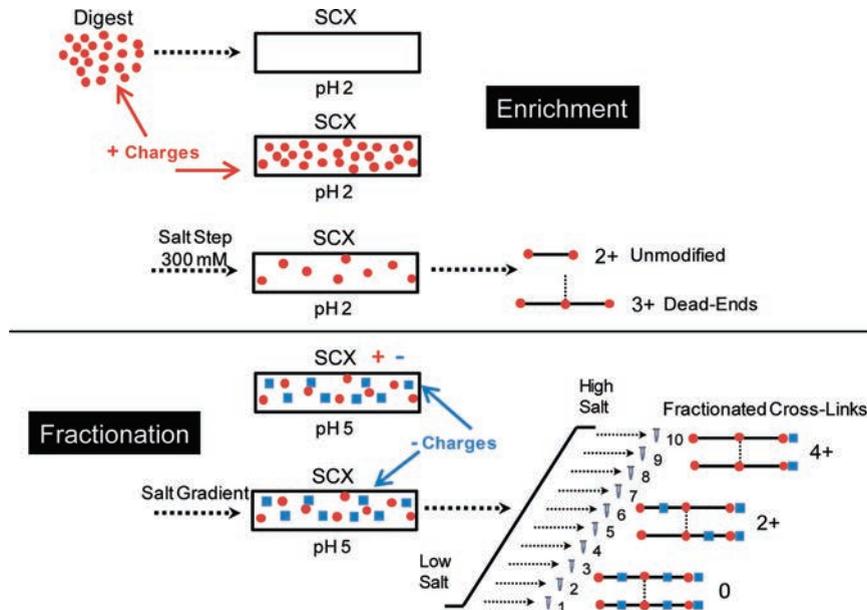


sum of DEST's spacer arm length and the length of two lysine side chains. This led to the estimate that there were at least 2200 lysine pairs within linkable distance in the ribosome sample. Of course, not all of the amines of lysine residues in the ribosome are solvent-accessible. In fact, judging by manual interpretation of the crystal structure, only 396 are fully accessible.[15] However, these fully accessible lysine residues are still capable of producing 1100 cross-links. A proteolytic digest of cross-linked ribosomal proteins was therefore expected to be overwhelmingly complex, because it would contain many peptides and many cross-links.

### Strategies for the Enrichment and Fractionation of Cross-Links

To reduce the complexity of such a digest and facilitate the detection of cross-links, proteolyzed samples were subjected to SCX chromatography. The fact that interpeptide cross-links formed by succinimidyl ester reagents have a considerable number

of basic functional groups has already been exploited by using SCX to partially separate cross-links from unmodified and dead-end-modified peptides that happen to have fewer basic functional groups.[10,11] DEST interpeptide cross-links should be better suited to such physical enrichment, because they contain two additional basic functional groups. DEST cross-links, if produced by tryptic proteolysis, contain at least six such groups, while the undesired unmodified and dead-end-modified peptides of the digest tend to contain only two or three. The technique, illustrated in Scheme 2, was simple to implement.[12] A tryptic digest of DEST-modified ribosomal proteins was loaded onto an SCX column at low pH (pH 2), to ensure that all ionizable groups were protonated, and then washed with 300 mM sodium chloride. This salt concentration cleared the sample of low charge species without causing significant loss of the more highly charged interpeptide cross-links. To further reduce the complexity of the digest before it was analyzed, we combined this low pH
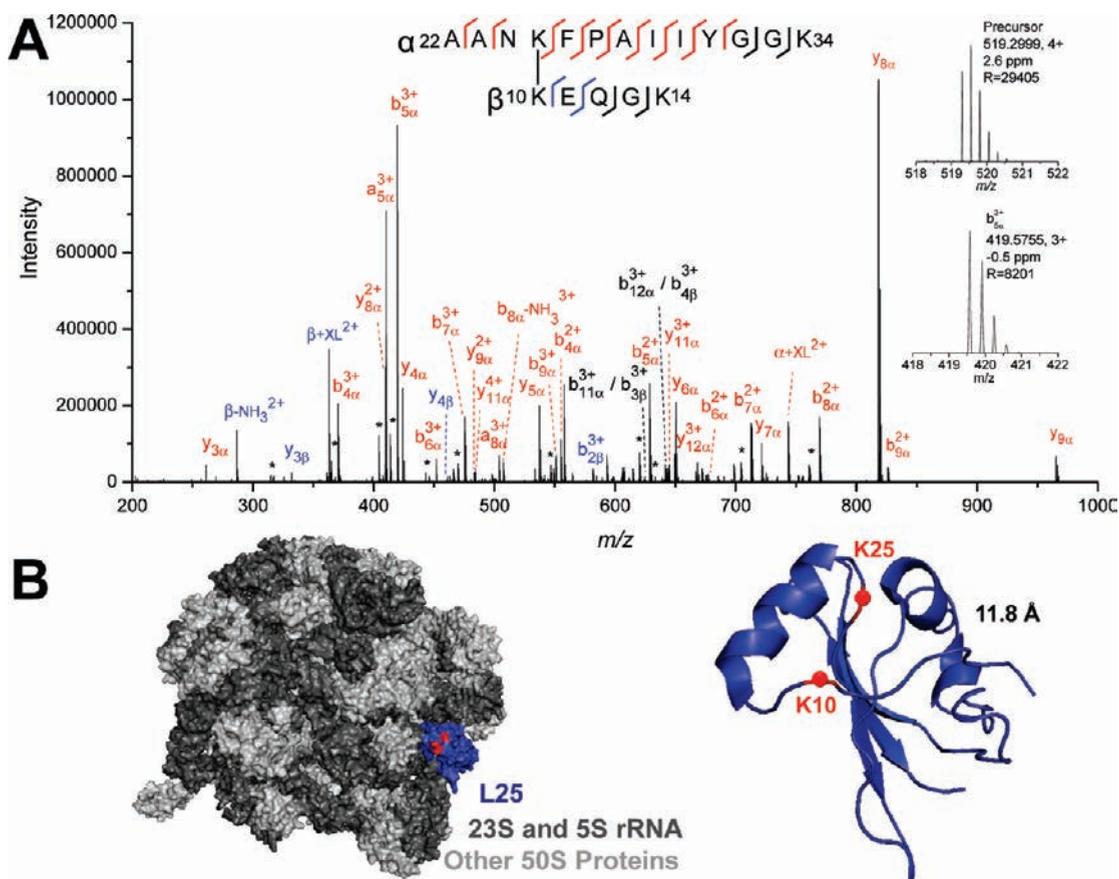
**Figure 1.** Intraprotein cross-link between K10 and K25 of ribosomal protein L25. (A) Orbitrap MS/MS spectrum of AANKFPAIIYGGK cross-linked to KEQGK. Peak assignments are shown in red if they correspond to the $\alpha$ peptide, blue if they correspond to the $\beta$ peptide, and black if they could correspond to either peptide chain. Peaks corresponding to fragmentation of the amidine bond of the cross-link, $\alpha$+XL and $\beta$+XL, have been manually assigned. These fragmentation pathways, shown in Supporting Information Figure 1, have been discussed previously.[12] The $\beta$-$NH_3^{2+}$ ion does not contain the cross-linker. Fragment ions have been assigned according to the nomenclature proposed by Schilling and co-workers.[39] Peaks marked with asterisks correspond to neutral losses of other assignments. (B) Structural context of the linkage with respect to the 50S subunit (PDB 2AW4). The residues involved in the identified linkage are marked in red.

enrichment with subsequent intermediate pH fractionation. At intermediate pHs, carboxylic acids become deprotonated. Since DEST cross-links, and any peptide species for that matter, contain differing numbers of such groups, their net charges at intermediate pH will differ, thus enabling productive fractionation. While the sample enriched for DEST interpeptide cross-links was still adsorbed to the SCX column, the mobile phase pH was shifted and a salt gradient was implemented. Ten fractions of the eluting DEST interpeptide cross-links were collected. This sample preparation was completed for two biological replicates, each with two technical replicates. Because we were interested in investigating the value of this sample handling process, the effects of each of the chromatographic steps were traced for one of the biological replicates by preparing three additional types of samples, one in which no SCX chromatography was completed, one in which only SCX enrichment was completed, and one in which only SCX fractionation was completed.

## Mass Spectrometric Analysis of Cross-Linked Peptides

Peptide mixtures produced by the differing sample preparations were analyzed via nanoLC-MS/MS with an LTQ-Orbitrap high-resolution hybrid mass spectrometer using a strategy in which both precursor and fragmentation mass spectra were

acquired in the Orbitrap. These experiments produced accurate precursor and fragment ion masses that were searched by the algorithm xQuest for matches to theoretical lysine to lysine interpeptide cross-links derived from the sequences of the ribosomal proteins. Score thresholds for cross-link spectra matches were established through decoy database searching, where the goal was to obtain a data set of identified linkages with an FDR <5%. We found that intraprotein cross-link matches could be accepted with lower xQuest scores than interprotein cross-links matches. This is easily explained by the fact that the size of the search space needed to identify intraprotein cross-links is much smaller than the search space required to identify interprotein cross-links. Consequently, there is a lower likelihood of an intraprotein cross-link match being a false positive. For intraprotein cross-link matches, use of an xQuest score threshold was sufficient for meeting our FDR goal. This was not the case for interprotein cross-link matches, which were found to require manual validation to guarantee that they contained fragmentation coverage across both peptides of the cross-links.

An example of an intraprotein cross-link spectrum match conforming to these criteria is shown in Figure 1. This particular match involves a cross-link between two peptides derived from ribosomal protein L25, AANKFPAIIYGGK and KEQGK.
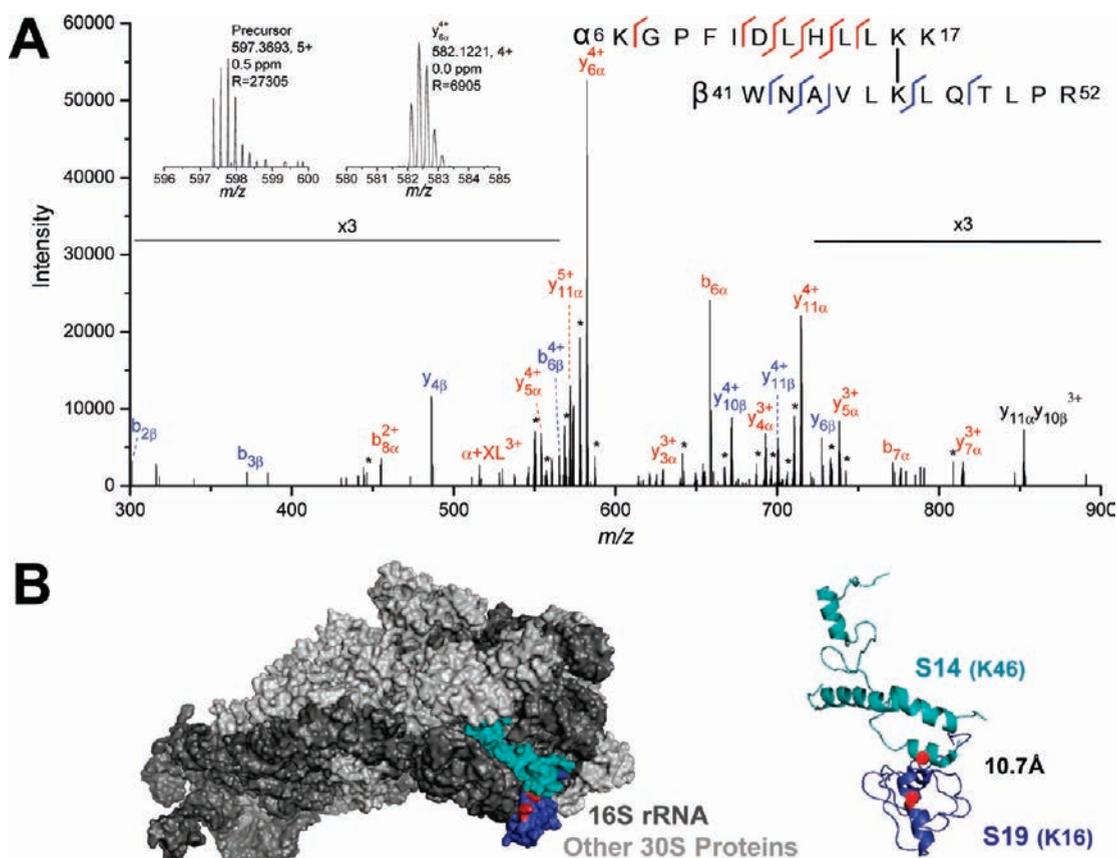
**Figure 2.** Interprotein cross-link between K16 of ribosomal protein S19 and K46 of ribosomal protein S14. (A) Orbitrap MS/MS spectrum of KGPFIDLHLLKK cross-linked to WNAVLKLQTLPR. Peak assignments are shown in red if they correspond to the α peptide and blue if they correspond to the β peptide. A peak corresponding to fragmentation of the amidine bond of the cross-link, α+XL, has been manually assigned. This fragmentation pathway, shown in Supporting Information Figure 1, has been discussed previously.[12] Fragment ions have been assigned according to the nomenclature proposed by Schilling and co-workers.[39] Peaks marked with asterisks correspond to neutral losses of other assignments. (B) Structural context of the linkage with respect to the 30S subunit (PDB 2AVY). The residues involved in the identified linkage are marked in red.

Unlike the fragmentation spectrum of a single peptide, this spectrum contains ions resulting from cleavage of both peptide chains. To illustrate this, peak assignments are shown in red if they correspond to fragmentation on the α peptide, and blue if they correspond to the β peptide. Some fragment ions were observed that could correspond to either peptide chain, given that they share a C-terminal sequence of GK. This type of ambiguous ion assignment, such as the peak labeled $b_{11\alpha}^{3+}/b_{3\beta}^{3+}$, is shown in black. As can be seen, the masses observed for precursor and fragment ions are in excellent agreement with theoretical masses. The majority (>90%) of the fragments ions assigned in this CID spectrum were in fact found to match theoretical masses to within 3 ppm. This impressive mass accuracy allowed xQuest database searching to be conducted with narrow mass tolerances, ensuring a small likelihood of obtaining false positives. Intraprotein cross-link matches, such as this, were identified at an FDR of 0.8% and, as noted earlier, did not require manual validation.

In contrast, interprotein cross-link matches required additional considerations to obtain an equally low FDR value. Figure 2 illustrates an interprotein cross-link spectrum. The assignments made to the spectrum demonstrate that this is a cross-link between the S19 peptide KGPFIDLHLLKK and the S14 peptide WNA-VLKLQTLPR. Fragmentation of this precursor yielded a significant number of informative ions; a total of 16 unique frag-

ment assignments could be attributed to one specific cross-link structure, notably one in which the linkage between the peptides involved the 11th residue from the N-terminus of the S19 (α) peptide. It is important to note that this spectrum does not contain any assignments eliminating the possibility of the linkage involving the C-terminal lysine residue of the S19 peptide. However, this residue is assumed to be unmodified, since amidation is known to block tryptic cleavage.[14,29] In this specific spectrum, nine assignments could be attributed to the α peptide sequence, and seven assignments could be attributed to the β peptide sequence. The sequences of both the α and β peptides were thus confidently matched. This was found to be a very strong indicator of a true positive. While analyzing decoy database searches, we often observed false positive matches in which the assigned fragmentation on the peptides was limited or biased to only one of the peptide chains. We consequently manually validated our interprotein cross-link matches to ensure that at least three unique fragments were assigned to each peptide chain and that there were a minimum of eight such fragments assigned per cross-link structure. In doing this, we reduced the FDR for interprotein cross-link matches from approximately 10% to 0.8%.

From our cross-linking analysis of the *E. coli* ribosome, we identified 325 cross-link spectra matches that were associated with 71 different peptide linkages. Tables 1 and 2 summarize

**Table 1. Intraprotein Cross-Link Matches from the 30S Subunit and the 50S Subunit[a]**

| 30S protein | topology[b] | distance | xQuest matches | 50S protein | topology[b] | distance | xQuest matches |
|---|---|---|---|---|---|---|---|
| S3 | K44−K85 | 13.0 | 1 | L1 | K5−K13 | [c] | 4 |
| **S3** | **K48−K85** | 16.2 | 8 | **L1** | **K53−K166** | [c] | 4 |
| S3 | K78−K85 | 11.1 | 1 | **L1** | **K53−K204** | [c] | 2 |
| **S3** | **K107−K146** | 14.4 | 19 | L3 | K116−K159/K160 | 19.6/16.1 | 1 |
| S3 | K107−K225 | [c] | 1 | L4 | K99−K106 | 14.9 | 1 |
| **S4** | **K82−K184** | 14.7 | 11 | L4 | K123−K137 | 7.5 | 1 |
| **S4** | **K150−K155** | 6.7 | 7 | L4 | K132−K166 | 13.7 | 1 |
| **S4** | **K155−K166** | 15.9 | 11 | **L9** | **K42−K57** | 25.0 | 16 |
| S4 | K155−K176 | 9.6 | 1 | L10 | K36−K104 | [c] | 1 |
| **S4** | **K166−K182** | 13.0 | 10 | **L14** | **K44−K54** | 5.6 | 11 |
| S6 | K35−K106 | [c] | 1 | **L14** | **K51−K114** | 19.5 | 13 |
| S6 | K56−K104 | [c] | 1 | **L14** | **K59−K111** | 16.3 | 4 |
| S6 | K56−K106 | [c] | 2 | L15 | K63−K70 | 21.6 | 1 |
| **S7** | **K148−K170** | [c] | 10 | **L16** | **K58−K118** | 21.1 | 3 |
| S9 | K99−K114 | 36.6 | 1 | L17 | K78−K121 | 25.0 | 1 |
| S14 | K18/K22−K46 | 19.3/18.0 | 1 | L18 | K63−K76 | 14.9 | 1 |
| S14 | K18−K46 | 19.3 | 2 | **L19** | **K62−K86** | 15.8 | 4 |
| **S14** | **K22−K46** | 18.0 | 6 | **L20** | **K77−K84** | 11.5 | 6 |
| S14 | K75−K82 | 12.4 | 1 | **L22** | **K6−K73** | 10.8 | 6 |
| S17 | K29−K38 | 9.2 | 1 | L22 | K16−K73 | 17.8 | 1 |
| **S19** | **K16−K20** | 6.0 | 9 | **L24** | **K16−K42** | 16.1 | 2 |
| S19 | K20−K28 | 12.3 | 2 | L24 | K18/K20−K42 | 19.5/19.8 | 2 |
| **S20** | **K33−K48** | 9.4 | 6 | **L25** | **K10−K25** | 11.8 | 36 |
| **S20** | **K68−K75** | 12.8 | 5 | L25 | K10−K68 | 17.8 | 1 |
| total | | | 118 | **L28** | **K61−K76** | [c] | 6 |
| | | | | L29 | Nterm/K2−K9 | 17.5/13.5 | 1 |
| | | | | L29 | K4−K9 | 11.8 | 3 |
| | | | | L33 | K9−K49 | 15.0 | 1 |
| | | | | **L33** | **K9−K52** | 12.1 | 10 |
| | | | | total | | | 144 |

[a] Bold indicates that the linkage was observed in both biological replicates. [b] Topologies shown with a slash separating residues indicate an ambiguous linkage. The distances for either linkage are provided. [c] Not applicable; one or both of the residues involved in the linkage are not present in the *E. coli* ribosome crystal structure.

intraprotein and interprotein linkages, along with the number of spectra matches supporting each identification. There is clearly a high degree of redundancy as well as variability with which linkages are represented in our data. For example, the linkage between K10 and K25 of ribosomal protein L25 was identified nearly 40 times more frequently than the linkage between K10 and K68 in the same protein. There are two primary factors that likely account for this observation. Differing rates of cross-link formation among linkages is one. Cross-links are obviously more likely to form when residue reactivities/solvent accessibilities, interresidue lengths, side chain orientations, and interresidue steric bulk are favorable for the reaction to occur. In the case of the noted L25 intraprotein linkages, K10 and K25 are closer than K10 and K68. These considerations alone are not likely to fully explain the variability with which linkages are represented in the data. A bias due to the methods used to identify the cross-links is certainly another factor. Differences in the ionization efficiencies of the linked peptides and the fragmentation behavior of the cross-links, for example, should introduce an additional effect. The most significant consequence of this variability and corresponding redundancy is that the low FDR of 0.8% for spectra matches increases to 3% for linkage identifications.

## Comparison of Cross-Linking and X-ray Crystallography Data

Of the 71 identified linkages, 52 were intraprotein and 19 were interprotein. None of the latter, however, connected the large and small ribosomal subunits. Crystal structure data indicated that only eight pairs of lysines were close enough for DEST to bridge the 30S and 50S subunits, so this was relatively unsurprising. It is further noteworthy that only 20% of the identified linkages were interprotein, indicating a heavy bias toward the identification of intraprotein linkages. Nevertheless, our identifications mirror the distribution of intra- and interprotein linkages deemed structurally possible in the crystal structure. Only about 20% of the lysine pairs bridging distances less than 24 Å were interprotein, while 80% were intraprotein. This highlights the nature of the ribosome, a macromolecular complex composed of two-thirds ribonucleic acid and one-third protein by mass.[30] Many ribosomal proteins are spatially separated due to the presence of RNA, limiting the number of interprotein cross-links that can form.

Distances between the lysine residues in the observed linkages, as calculated from the crystal structure of the *E. coli* ribosome, are listed in Tables 1 and 2. In our results, there are 240 cross-link

**Table 2. Interprotein Cross-Link Matches**[a]

| | protein 1 | protein 2 | topology[b] | distance ($C_\alpha-C_\alpha$, Å) | xQuest matches |
|---|---|---|---|---|---|
| **30S** | S2 | S8 | K25/K27−K68 | 16.1/19.1 | 2 |
| | S3 | S14 | K85−K6 | 30.0 | 1 |
| | S5 | S8 | K158−K40 | 12.2 | 1 |
| | **S6** | **S18** | **K106−K29** | [c] | 15 |
| | **S14** | **S19** | **K46−K16** | 10.7 | 3 |
| | S18 | S21 | K29−K4 | 16.3 | 2 |
| **50S** | L2 | L12 | K67−K81 | [c] | 1 |
| | L6 | L12 | K28−K70 | [c] | 2 |
| | L6 | L12 | K85−K59 | [c] | 2 |
| | L6 | L12 | K85−K81 | [c] | 1 |
| | L9 | L10 | K57−K143 | [c] | 1 |
| | **L9** | **L28** | **K42−K43** | [c] | 10 |
| | **L9** | **L28** | **K42−K61** | [c] | 5 |
| | **L9** | **L28** | **K57−K43** | [c] | 7 |
| | L11 | L12 | K99−K70 | [c] | 2 |
| | **L16** | **L25** | **K127−K83** | 9.3 | 4 |
| | L17 | L32 | K121−K31 | 17.6 | 1 |
| | **L18** | **L27** | **K17−K61** | 19.9 | 2 |
| | L22 | L32 | K28−K36 | 19.4 | 1 |
| total | | | | | 63 |

[a] Bold indicates that the linkage was observed in both biological replicates. [b] Topologies shown with a slash separating residues indicate an ambiguous linkage. The distances for either linkage are provided. [c] Not applicable, one or both of the residues involved in the linkage are not present in the *E. coli* ribosome crystal structure.



**Figure 3.** Structural context of cross-linking data. Identified linkages involving residues present in the crystal structure of the *E. coli* ribosome (PDB 2AVY and 2AW4)[25] are shown. Ribosomal proteins are displayed in cyan, intraprotein linkages in blue, and interprotein linkages in red. For clarity, rRNA are not shown.

spectra matches in which linked residues are both unambiguously identified and present in the crystal structure. The structural context of these data, and the 46 linkages they indicate, is displayed in Figure 3. In addition, Figure 4 displays the distribution of 240 spectral matches as a function of the distance between the linked lysine residues. Note that all but 19 of the matches corresponded to inter-residue distances less than or equal to 24 Å, the maximum α-carbon to α-carbon cross-linking distance for DEST. Furthermore, all but two corresponded to interresidue distances less than or equal to 25 Å, a distance that still demonstrates agreement given the anticipated error in the crystal structure. An estimate for the coordinate error is not provided in the PDB entry of the *E. coli* ribosome structure. Such a value is, nevertheless, provided for the crystal structure of the *T. thermophilus* ribosome (PDB 2J00 and 2J01).[31] In this crystal structure of similar resolution, the estimated coordinate error is on the order of 1 Å. From our interpretation, then, 99% of the cross-link spectra matches are consistent with the crystal structure of the *E. coli* ribosome. DEST cross-linking thus provides reliable structural information about macromolecular structures. One of the two cross-link spectra matches that corresponded to an interresidue distance greater than 25 Å indicates a linkage between ribosomal proteins S3 and S14. These two proteins reside at adjacent positions in the crystal structure. Still, the residues involved in the linkage are separated by 30.0 Å (Supporting Information Figure 2). Conformational flexibility or alternative conformations of these proteins in solution may account for the observation of this linkage. The second spectrum match that is inconsistent with the crystal structure defines an intraprotein linkage in ribosomal protein S9 between K99 and K114. These
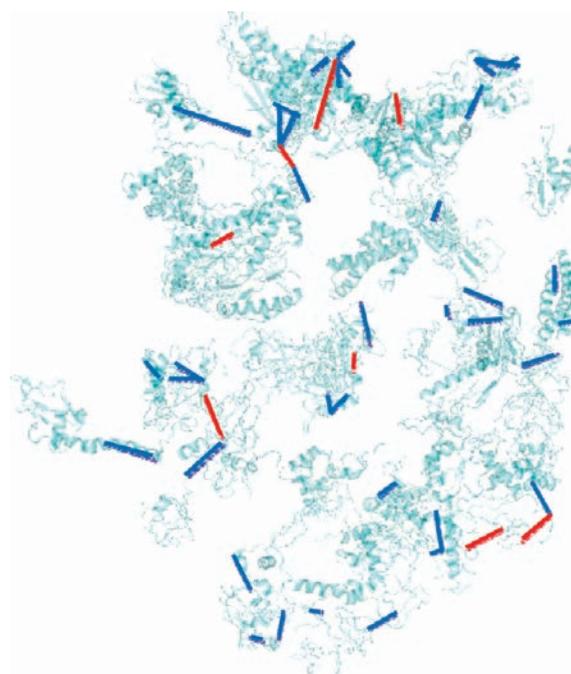
residues are separated by 36.6 Å (Supporting Information Figure 3). A discrepancy as large as this is not likely to be accounted for by conformational flexibility of the protein. It is more likely that sample preparation or storage yielded a population of ribosomes with a structural anomaly, such as damaged RNA. Lysine 114 is part of an unstructured domain anchored by RNA; damage to or cleavage of this RNA, by endogenous RNase, for example, would release this protein domain, allowing K114 to come within a linkable distance of K99. Alternatively, this linkage may simply be a false positive. The FDR determined for this analysis indicates that it is highly probable for our data to contain at least one false positive.

About one-fourth of the linkages identified in this study involved residues not present in the crystal structure of the *E. coli* ribosome. The majority of these linkages are consistent with other structural studies of prokaryotic ribosomes. For instance, although the location of L28 is not properly defined in the *E. coli* crystal structure, three different linkages from our data demonstrate that it must make extensive contacts with ribosomal protein L9. The electron density of L28 was properly refined in the *T. thermophilus* ribosome crystal structure,[31] and indeed from this structure, it can be seen that L28 directly binds to L9. Nevertheless, a couple of these linkages could not be readily explained given what is currently known about ribosome structure. These results are the subject of further discussion below.

## Cross-Linking of the Stalk Complex

Cross-linking of the proteins that constitute the stalk complex was particularly intriguing, since this prominent feature of the ribosome has not yet been fully refined in any crystal structure. The stalk complex, formed by one copy of L10 and four copies of L12, has, nonetheless, been structurally elucidated in a hybrid model constructed from the combined use of
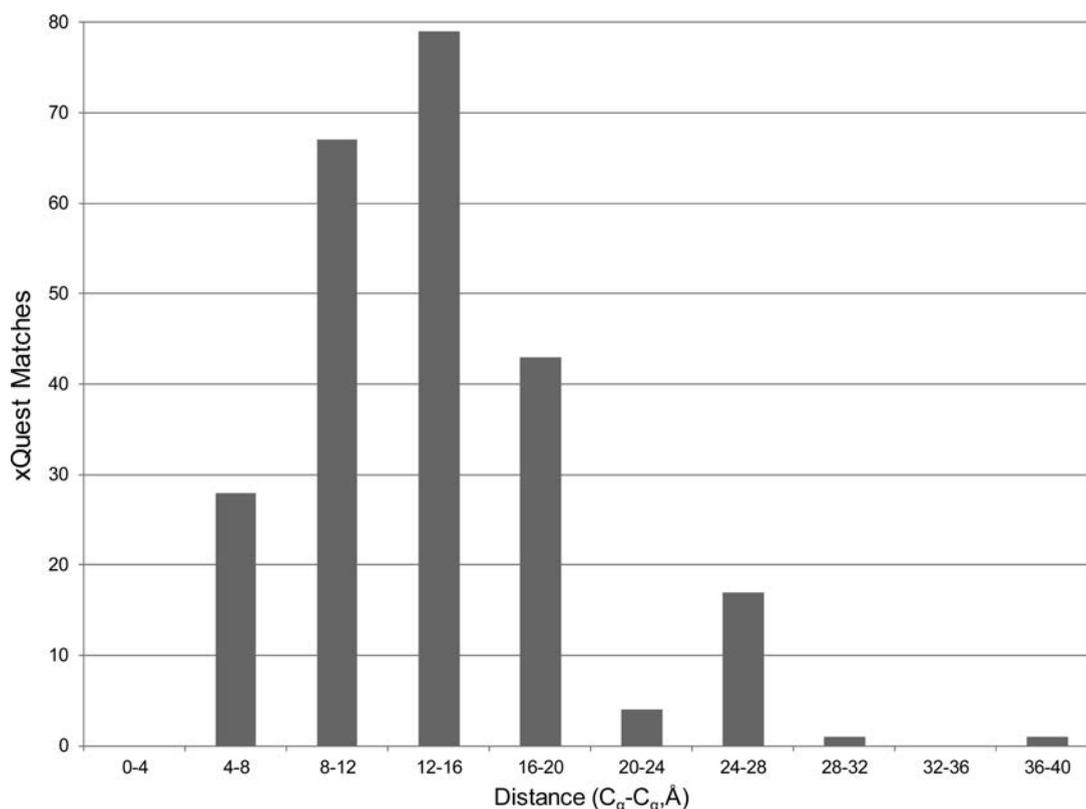
**Figure 4.** xQuest cross-link matches as a function of distance between lysine residues in the identified linkages.

electron microscopy and X-ray crystallography.[32] This model demonstrates that the base of the complex is situated adjacent to ribosomal proteins L6 and L11; our results capture this structural feature of the ribosome, as we identified four different linkages between these two proteins and ribosomal protein L12.

Two linkages involving the proteins of the stalk complex (L10 and L12) that cannot be explained by the cryoEM/crystal structure model mentioned above were identified. These linkages correspond to K67 of L2 cross-linked to K81 of L12 and K57 of L9 cross-linked to K143 of L10. Interestingly, L2 and L9 reside on the opposite side of the ribosome from the stalk complex. Figure 5A displays the crystal structure of the *E. coli* ribosome and demonstrates the position of the stalk complex with respect to the proteins to which it was found cross-linked. In Figure 5B, the structure has been rotated 90° in either direction, making the residues involved in these linkages visible. Since we estimate the distance across the ribosome to be at least 50 Å greater than the span of any of the proteins[25,32,33] and the cross-linker, these linkages are believed to correspond to interactions between ribosomes. A model for these interactions is shown in Figure 6, where the stalk is depicted as a large red circle. Our model shows that the stalk protein complex may, in addition to other functions, serve as a hub point between ribosomes and interact with ribosomal proteins L9 and L2 on an adjacent ribosome. Because the reaction conditions used here for cross-linking ribosomes have not been found to yield nonspecific cross-linking of monomeric proteins,[12] we believe these ribosome—ribosome interactions are of functional significance.

Cross-linking between the stalk complex and the L2 region on the opposite face of the *E. coli* ribosome has been detected before,

albeit via different methods. Dey and co-workers previously incorporated a photoreactive group into the structure of L12 as a means to conduct site-directed cross-linking and to identify the proteins with which the stalk complex interacts.[34] Their interpretation for L12 being cross-linked to L2, made at a time when high-resolution structures did not exist, was that an elongated structure of L12 might extend across the body of the 50S subunit and that this interacts with L2. In light of more recent work modeling the ribosome and the stalk complex,[32] this seems unlikely. Even a fully extended L12 protein could not span the distance between the base of the stalk complex and ribosomal protein L2. Their results are consistent with our suggestion that ribosome—ribosome interactions are occurring.

A recent analysis of *E. coli* polysomes, clusters of ribosomes on an mRNA, by cryoelectron tomography is consistent with our ribosome—ribosome interaction hypothesis.[35] In this study, adjacent ribosomes were found to adopt preferential configurations, in which 30S subunits aligned with orientations that differed by 180°. Although the accuracy of docking the ribosome crystal structure[25] into their averaged 3D density of adjacent ribosomes was limited, it was suggested that in the most common configuration contact between ribosomes extends from the L1 arm of one ribosome to a region near protein S4 in an adjacent ribosome. Furthermore, the docking showed that in this configuration ribosomal protein L9 extends out from one ribosome toward the large subunit of another. Our identification of a linkage between ribosomal protein L9 and the stalk complex is consistent with this preferred configuration of ribosome—ribosome interaction. However, it is uncertain whether the ribosome—ribosome interactions identified in our analysis correspond to polysomes or free ribosomes. Our preparation
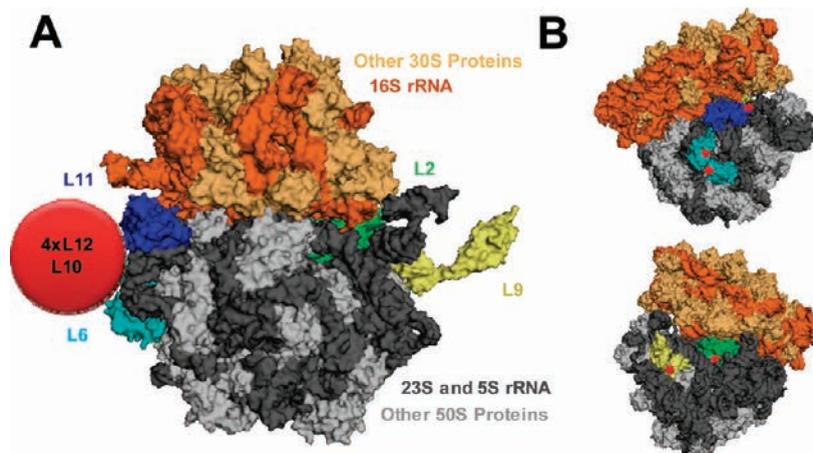
**Figure 5.** Cross-linking of the stalk complex. (A) The crystal structure of the *E. coli* ribosome (PDB 2AVY and 2AW4).[25] The stalk complex, positioned at its canonical binding site, is represented by a large red circle. The proteins found to cross-link to the stalk complex, L2, L6, L9, and L11, are labeled. (B) Rotations of the *E. coli* ribosome. The L6/L11 (top) and L2/L9 faces (bottom) of the ribosome are shown. Residues involved in the linkages with the stalk complex are highlighted in red.
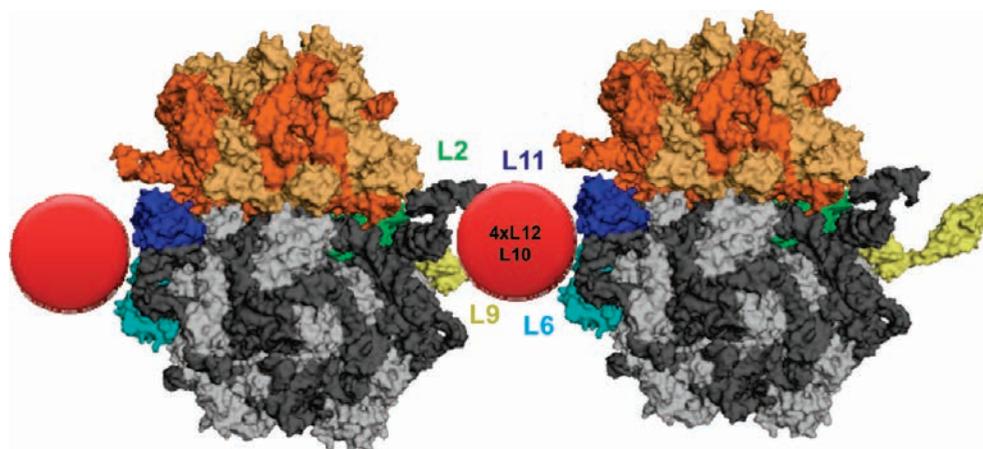


**Figure 6.** Model of ribosome—ribosome interactions based on DEST cross-linking data.

of ribosomes, while expected to contain primarily free ribosomes, may contain some polysomes. Further study involving purification of free ribosomes and polysomes is needed to distinguish the origin of our identifications.

Ribosome—ribosome interactions may be of functional significance. Previous studies have indicated that L9, one of the proteins mentioned above, plays a role in enhancing the fidelity of translation by minimizing ribosome slippage.[36,37] Among the hypotheses for how L9 is able to do so, it has been suggested that L9 may function as a strut[38] to couple ribosomes within a polysome.[36] This coupling could prevent the downstream ribosome from sliding relative to the mRNA. Our results lend credence to this hypothesis by demonstrating the involvement of L9 in ribosome—ribosome interactions. Further study is needed to better characterize ribosome—ribosome interactions and their functional significance.

### Amenability of DEST Cross-Links to SCX Enrichment and Fractionation

It was of interest to assess the extent to which SCX chromatography improved the analysis of DEST cross-links. The number of spectra matches identified per LC-MS/MS experiment has

been cataloged as a function of sample preparation for one biological replicate in Figure 7. Without prior sample preparation, an average of only 5.5 cross-link spectra matches was obtained per LC-MS/MS experiment. After SCX enrichment, however, the average number of cross-link spectra matches per analysis was increased 2.5-fold to 13.5. This can be attributed to the fact that the enrichment physically altered the composition of the digest, making it more abundant in DEST interpeptide cross-links and less abundant in linear peptide species. To support that this was the case, the above samples were analyzed by LC-MS/MS experiments in which, unlike before, there was no charge rejection to bias detection against linear peptides, and the resulting data were searched for both interpeptide cross-links and linear peptide species. As anticipated, the ratio of interpeptide cross-link spectra matches to linear peptide spectra matches was found to increase from 0.018 to 0.34, a 20-fold change, due to the enrichment. It is worth noting that these ratios of spectra matches are not expected to accurately translate to the percent composition of the sample, because cross-linked peptides are significantly more challenging to identify with confidence than linear peptides.
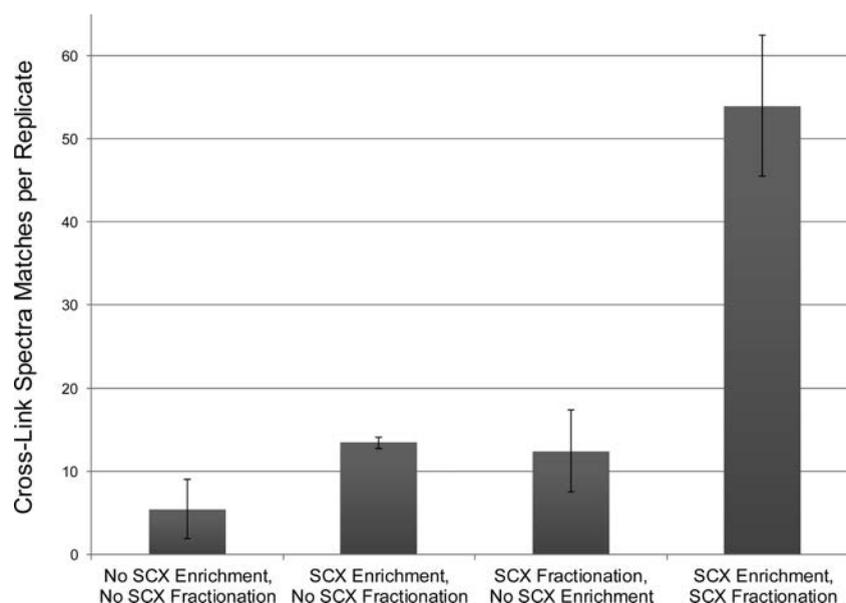
**Figure 7.** Cross-link spectra matches as a function of sample preparation.

As explained previously, enrichment of DEST interpeptide cross-links using SCX integrates easily with their fractionation (Scheme 2). When DEST interpeptide cross-links were both enriched and fractionated by SCX, the average number of cross-link spectra matches per analysis increased to 54, a 10-fold increase over the sample analyzed without prior sample preparation. For comparison sake only, we additionally prepared a digest sample that was subjected to fractionation without prior enrichment. This sample preparation yielded an average of 12.5 matches per analysis. Clearly, the effects of the enrichment and fractionation are multiplicative, validating the use of this enrichment technique.

## ■ CONCLUSION

We have probed the structure of the *E. coli* ribosome, a 2.5 MDa ribonucleoprotein complex containing more than 50 proteins, using the novel amidinating cross-linker DEST and mass spectrometry. This study demonstrates that peptide cross-links derived from very complex structures can be identified at high confidence from single precursor mass measurements and CID fragmentation spectra. Employing the excellent mass accuracy of the LTQ Orbitrap, we identified 325 cross-link spectra at a FDR of 0.8%. These identifications led to the assignment of 71 unique linkages with an FDR of 3%. Our results further show that DEST cross-linking yields reliable information about macromolecular structure, as the acquired cross-linking data are in excellent agreement with the crystal structure of the *E. coli* ribosome. Our results also demonstrate that the amenability of DEST cross-links to SCX chromatography is very useful in a large-scale analysis. This characteristic was, in fact, shown to increase the number of cross-link spectra matches in our analysis 10-fold. Finally, given that we were able to identify ribosome—ribosome interactions of potentially significant function, our structural analysis of the *E. coli* ribosome demonstrates that this technique constitutes a viable method for studying complex interactomes.

## ■ ASSOCIATED CONTENT

### Ⓢ Supporting Information

Scheme for the synthesis of DEST, tables of xQuest MS/MS matches for interpeptide cross-links, figure indicating the fragmentation of the amidine bond of a DEST cross-link, figure demonstrating the structural context of an interprotein linkage between ribosomal proteins S3 and S14, and figure demonstrating the structural context of an intraprotein linkage between K99 and K114 in ribosomal protein S9. This material is available free of charge via the Internet at http://pubs.acs.org.

## ■ AUTHOR INFORMATION

### Corresponding Author
*E-mail: reilly@indiana.edu. Phone: (812)-855-1980. Fax: (812)855-8300.

## ■ ACKNOWLEDGMENT

## ■ REFERENCES

(1) Selth, L. A.; Sigurdsson, S.; Svejstrup, J. Q. Transcript elongation by RNA polymerase II. *Annu. Rev. Biochem.* **2010**, *79*, 271–293.

(2) Valadkhan, S. The spliceosome: Caught in a web of shifting interactions. *Curr. Opin. Struct. Biol.* **2007**, *17* (3), 310–315.

(3) Rodnina, M. V.; Wintermeyer, W. Recent mechanistic insights into eukaryotic ribosomes. *Curr. Opin. Cell Biol.* **2009**, *21* (3), 435–443.

(4) Russell, R. B.; Alber, F.; Aloy, P.; Davis, F. P.; Korkin, D.; Pichaud, M.; Topf, M.; Sali, A. A structural perspective on protein-protein interactions. *Curr. Opin. Struct. Biol.* **2004**, *14* (3), 313–324.

(5) Singh, P.; Panchaud, A.; Goodlett, D. R. Chemical cross-linking and mass spectrometry as a low-resolution protein structure determination technique. *Anal. Chem.* **2010**, *82* (7), 2636–2642.

(6) Leitner, A.; Walzthoeni, T.; Kahraman, A.; Herzog, F.; Rinner, O.; Beck, M.; Aebersold, R. Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol. Cell. Proteomics* **2010**, *9* (8), 1634–1649.

(7) Chowdhury, S. M.; Du, X.; Tolic, N.; Wu, S.; Moore, R. J.; Mayer, M. U.; Smith, R. D.; Adkins, J. N. Identification of cross-linked peptides after click-based enrichment using sequential collision-induced dissociation and electron transfer dissociation tandem mass spectrometry. *Anal. Chem.* **2009**, *81* (13), 5524–5532.

(8) Yan, F.; Che, F.-Y.; Rykunov, D.; Nieves, E.; Fiser, A.; Weiss, L. M.; Hogue Angeletti, R. Nonprotein based enrichment method to analyze peptide cross-linking in protein complexes. *Anal. Chem.* **2009**, *81* (17), 7149–7159.

(9) Yang, L.; Tang, X.; Weisbrod, C. R.; Munske, G. R.; Eng, J. K.; von Haller, P. D.; Kaiser, N. K.; Bruce, J. E. A photocleavable and mass spectrometry identifiable cross-linker for protein interaction studies. *Anal. Chem.* **2010**, *82* (9), 3556–3566.

(10) Chen, Z.; Jawhari, A.; Fischer, L.; Buchen, C.; Tahir, S.; Kamenski, T.; Rasmussen, M.; Lariviere, L.; Bukowski-Wills, J.-C.; Nilges, M.; Cramer, P.; Rappsilber, J. Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* **2010**, *29* (4), 717–726.

(11) Rinner, O.; Seebacher, J.; Walzthoeni, T.; Mueller Lukas, N.; Beck, M.; Schmidt, A.; Mueller, M.; Aebersold, R. Identification of cross-linked peptides from large sequence databases. *Nat. Methods* **2008**, *5* (4), 315–318.

(12) Lauber, M. A.; Reilly, J. P. Novel amidinating cross-linker for facilitating analyses of protein structures and interactions. *Anal. Chem.* **2010**, *82* (18), 7736–7743.

(13) Running, W. E.; Reilly, J. P. Variation of the chemical reactivity of *Thermus thermophilus* HB8 ribosomal proteins as a function of pH. *Proteomics* **2010**, *10* (20), 3669–3687.

(14) Running, W. E.; Reilly, J. P. Ribosomal proteins of *Deinococcus radiodurans*: Their solvent accessibility and reactivity. *J. Proteome Res.* **2009**, *8* (3), 1228–1246.

(15) Liu, X.; Reilly, J. P. Correlating the chemical modification of *Escherichia coli* ribosomal proteins with crystal structure data. *J. Proteome Res.* **2009**, *8* (10), 4466–4478.

(16) Lauber, M. A.; Running, W. E.; Reilly, J. P. *B. subtilis* ribosomal proteins: Structural homology and post-translational modifications. *J. Proteome Res.* **2009**, *8* (9), 4193–4206.

(17) Beardsley, R. L.; Running, W. E.; Reilly, J. P. Probing the structure of the *Caulobacter crescentus* ribosome with chemical labeling and mass spectrometry. *J. Proteome Res.* **2006**, *5* (11), 2935–2946.

(18) Roger, R.; Neilson, D. G. The chemistry of imidates. *Chem. Rev.* **1961**, *61*, 179–211.

(19) Spedding, G. In *Ribosomes and Protein Synthesis: A Pratical Approach*; Rickwood, D., Hames, B. D., Eds.; Oxford Press: New York, 1990; pp 1−29.

(20) Arnold, R. J.; Reilly, J. P. Observation of *Escherichia coli* ribosomal proteins and their posttranslational modifications by mass spectrometry. *Anal. Biochem.* **1999**, *269*, 105–112.

(21) Olson Bradley, J. S. C.; Markwell, J. Assays for determination of protein concentration. In *Current Protocols in Protein Science*; Coligan, J. E., Ed.; John Wiley and Sons, Inc.: Brooklyn, NY, **2007**, Chapter 3, Unit 3.4.

(22) Kapoor, K. N.; Barry, D. T.; Rees, R. C.; Dodi, I. A.; McArdle, S. E. B.; Creaser, C. S.; Bonner, P. L. R. Estimation of peptide concentration by a modified bicinchoninic acid assay. *Anal. Biochem.* **2009**, *393* (1), 138–140.

(23) Karty, J. A.; Running, W. E.; Reilly, J. P. Two dimensional liquid phase separations of proteins using online fractionation and concentration between chromatographic dimensions. *J. Chromatogr. B.* **2007**, *847*, 103–113.

(24) Tang, W. H.; Shilov, I. V.; Seymour, S. L. Nonlinear fitting method for determining local false discovery rates from decoy database searches. *J. Proteome Res.* **2008**, *7* (9), 3661–3667.

(25) Schuwirth, B. S.; Borovinskaya, M. A.; Hau, C. W.; Zhang, W.; Vila-Sanjurjo, A.; Holton, J. M.; Cate, J. H. D. Structures of the bacterial ribosome at 3.5 A resolution. *Science* **2005**, *310* (5749), 827–34.

(26) DeLano, W. L. *The PyMOL Molecular Graphics System*; DeLano Scientific: Palo Alto, CA, USA, 2002.

(27) Umekage, S.; Ueda, T. Spermidine inhibits transient and stable ribosome subunit dissociation. *FEBS Lett.* **2006**, *580* (5), 1222–1226.

(28) Davis, B. D. Role of subunits in the ribosome cycle. *Nature* **1971**, *231* (5299), 153–7.

(29) Liu, X.; Broshears, W. C.; Reilly, J. P. Probing the structure and activity of trypsin with amidination. *Anal. Biochem.* **2007**, *367* (1), 13–19.

(30) Voet, D.; Voet, J. G. In *Biochemistry*, 3rd ed.; J. Wiley & Sons: Hoboken, NJ, 2004; p 67.

(31) Selmer, M.; Dunham, C. M.; Murphy, F. V.; Weixlbaumer, A.; Petry, S.; Kelley, A. C.; Weir, J. R.; Ramakrishnan, V. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **2006**, *313*, 1935–1942.

(32) Diaconu, M.; Kothe, U.; Schluenzen, F.; Fischer, N.; Harms, J. M.; Tonevitsky, A. G.; Stark, H.; Rodnina, M. V.; Wahl, M. C. Structural basis for the function of the ribosomal L7/12 stalk in factor binding and GTPase activation. *Cell* **2005**, *121* (7), 991–1004.

(33) Bocharov, E. V.; Sobol, A. G.; Pavlov, K. V.; Korzhnev, D. M.; Jaravine, V. A.; Gudkov, A. T.; Arseniev, A. S. From structure and dynamics of protein L7/L12 to molecular switching in ribosome. *J. Biol. Chem.* **2004**, *279* (17), 17697–17706.

(34) Dey, D.; Bochkariov, D. E.; Jokhadze, G. G.; Traut, R. R. Crosslinking of selected residues in the N- and C-terminal domains of *Escherichia coli* protein L7/L12 to other ribosomal proteins and the effect of elongation factor Tu. *J. Biol. Chem.* **1998**, *273* (3), 1670–1676.

(35) Brandt, F.; Etchells, S. A.; Ortiz, J. O.; Elcock, A. H.; Hartl, F. U.; Baumeister, W. The native 3D organization of bacterial polysomes. *Cell* **2009**, *136* (2), 261–271.

(36) Herr, A. J.; Nelson, C. C.; Wills, N. M.; Gestland, R. F.; Atkins, J. F. Analysis of the roles of tRNA structure, ribosomal protein L9, and the bacteriophage T4 gene 60 bypassing signals during ribosome slippage on mRNA. *J. Mol. Biol.* **2001**, *309* (5), 1029–1048.

(37) Adamski, F. M.; Atkins, J. F.; Gesteland, R. F. Ribosomal protein L9 interactions with 23 S rRNA: The use of a translational bypass assay to study the effect of amino acid substitutions. *J. Mol. Biol.* **1996**, *261* (3), 357–371.

(38) Hoffman, D. W.; Davies, C.; Gerchman, S. E.; Kycia, J. H.; Porter, S. J.; White, S. W.; Ramakrishnan, V. Crystal structure of prokaryotic ribosomal protein L9: a bi-lobed RNA-binding protein. *EMBO J.* **1994**, *13* (1), 205–212.

(39) Schilling, B.; Row, R. H.; Gibson, B. W.; Guo, X.; Young, M. M. MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides. *J. Am. Soc. Mass Spectrom.* **2003**, *14* (8), 834–850.