RESEARCH ARTICLE

Variation of the chemical reactivity of *Thermus thermophilus* HB8 ribosomal proteins as a function of pH

William E. Running and James P. Reilly

Department of Chemistry, Indiana University, Bloomington, IN, USA

Received: May 31, 2010
Revised: July 28, 2010
Accepted: August 4, 2010

Ribosomes occupy a central position in cellular metabolism, converting stored genetic information into active cellular machinery. Ribosomal proteins modulate both the intrinsic function of the ribosome and its interaction with other cellular complexes, such as chaperonins or the signal recognition particle. Chemical modification of proteins combined with mass spectrometric detection of the extent and position of covalent modifications is a rapid, sensitive method for the study of protein structure and flexibility. By altering the pH of the solution, we have induced non-denaturing changes in the structure of bacterial ribosomal proteins and detected these conformational changes by covalent labeling. Changes in ribosomal protein modification across a pH range from 6.6 to 8.3 are unique to each protein, and correlate with their structural environment in the ribosome. Lysine residues whose extent of modification increases as a function of increasing pH are on the surface of proteins, but in close proximity either to glutamate and aspartate residues, or to rRNA backbone phosphates. Increasing pH disrupts tertiary and quaternary interactions mediated by hydrogen bonding or ionic interactions, and regions of protein structure whose conformations are sensitive to these changes are of potential importance in modulating the flexibility of the ribosome or its interaction with other cellular complexes.

Keywords:
Bacterial ribosomes / Chemical labeling / MS / pH dependence / Protein chemistry / Technology

1 Introduction

Ribosomes are protein–nucleic acid complexes whose central role is the synthesis of cellular proteins. The bacterial ribosome is composed of in excess of 50 distinct proteins and three large ribonucleic acid polymers (rRNA). These components are distributed between a small subunit (0.8 MDa molecular weight, with a sedimentation coefficient of 30S) and a large subunit (1.5 MDa molecular weight, with a sedimentation coefficient of 50S). The small subunit translates mRNAs into proteins using aminoacylated transfer RNAs, a process mediated primarily by the ribosomal RNA. The large subunit is a ribozyme containing the catalytic machinery for forming new peptide bonds [1]. The bacterial ribosome is also of practical interest since it is one of the four primary targets for antibiotic action [2]. Because antibiotic resistance is an inevitable consequence of the use of antibiotics, continued study of this primary target complex is necessary to develop new antibacterial drugs [2, 3]. These considerations, as well as the challenge of solving the crystal structure of such a large complex, motivated work that ultimately led to Yonath, Steitz and Ramakrishnan's receipt of the 2009 Nobel Prize in Chemistry for the development of cryocrystallographic techniques and the solution of the crystal structures of a large subunit from a halophilic archaeon, and large and small subunit structures from a pair of closely related bacteria [4].

X-ray crystallography and nuclear magnetic resonance spectroscopy have been the preeminent techniques for the determination of atomic resolution biomolecular structures. Their limitations include the requirement of milligrams of

Correspondence: Professor James P. Reilly, Department of Chemistry, Indiana University, 800 E. Kirkwood Ave., Bloomington, IN 47408, USA
E-mail: reilly@indiana.edu
Fax: +1-812-855-8300

Abbreviation: SMTA, *S*-methylthioacetimidate

protein, long periods of method development (*e.g.* screening crystallization conditions), destructive data collection procedures (*e.g.* sample denaturation or damage by X-rays), and long times for data reduction and analysis [5, 6]. By contrast, MS-based methods use several orders of magnitude less material *per* analysis, and shorter experiment times [7]. Coupled LC-MS allows proteins or large complexes to be analyzed from near-native solution conditions. Direct electrospray of bacterial ribosomes revealed species-specific differences in the stoichiometry of the ribosomal stalk protein L7/L12 [8–9] (The alternate names indicate the presence or absence of N-terminal acetylation in an otherwise identical polypeptide). Analysis of samples in solution also enables mass spectrometric determination of the structural effects of changes in biochemically interesting solution parameters such as temperature, pH and ionic strength, and composition to be studied.

A number of mass spectrometric techniques exist for capturing biomolecular structural data. These include hydrogen–deuterium exchange [10–12], and covalent modification [13–23]. Covalent modification techniques can either decrease the masses of proteins, as is seen with limited proteolysis of native structures to remove conformationally flexible ''loops and fringes'' of a protein [13–16], or increase protein masses, as in selective modification of specific amino acid side chains [17–24]. In contrast to hydrogen–deuterium exchange, covalent modifications are usually not labile, which simplifies the handling and analysis of the labeled proteins. An additional advantage of covalent modification procedures is that hydrolytic enzymes and side chain modifying reagents have well-characterized residue selectivities that allow additional sequence information to be derived from mass spectrometric analysis of modified proteins or peptides. Residue selective chemical modification allows researchers to take advantage of an extensive arsenal of protein modification reagents [25].

Over the last several years, we have employed the thioimidate modifying reagent *S*-methylthioacetimidate (SMTA) to probe the structure of both isolated proteins [26, 27] and bacterial ribosomes [28–31]. In the studies of bacterial ribosome structure, the experimental labeling pattern for each ribosomal protein was examined by top-down and bottom-up proteomic methods and compared with predictions based on the inspection of subunit or whole ribosome crystal structures. The observed average extent of modification for whole proteins, and the modification or protection of specific lysine residues in tryptic peptides, showed excellent agreement with the reactivity predicted by the examination of crystal structures. The only disagreement between experimental results and predictions based on available crystal structures were due either to PTM of amino groups in ribosomal proteins, which eliminated their reactivity with SMTA, or flexible loops in the proteins' sequences whose disordered structure made them effectively invisible in the crystals. The purpose of the research presented here was to expand the range of conditions for obtaining structural data from macromolecular complexes outside of our usual conditions (5 mM $MgCl_2$, 225 mM Tris, pH 8.6), and outside the conditions used for the crystallization of the ribosome (0.2 M KSCN, 0.1 M Tris acetate, pH 7) [32]. The most easily altered solution parameters are pH, ionic strength, and ionic composition. The variation of pH was the most appropriate choice, as it was expected to change the hydrogen bonding environment of lysine residues. Ribosomes from *Thermus thermophilus* HB8 are an excellent model system because of the availability of high-resolution crystal structures of its 30S [33–35] and 70S [32, 36] subunits.

2 Materials and methods

2.1 Chemicals and reagents

Water was purified by a Barnstead Thermolyne Nanopure system. Ammonium chloride, 2-mercaptoethanol, and methylamine as a 40% v/v solution in water were supplied by Aldrich. HPLC grade ACN, methanol, TFA, and formic acid were products of J. T. Baker. Peptone and yeast extract for bacterial growth media were provided by DIFCO. Reagent grade sodium chloride and EDTA were obtained from Mallinckrodt. 2-Amino-2-(hydroxymethyl)-1,3-propanediol (TRIZMA, Tris-free base), horse heart myoglobin, HEPES (free acid), leucine enkephalin, MES (free acid), sucrose, and proteomics-grade alkylated porcine trypsin were purchased from Sigma.

SMTA was synthesized according to Thumm's method as described by Beardsley [37, 38]. SMTA modifies solvent-exposed lysine residues according to Scheme 1, increasing the mass of peptides or proteins by 41.05 Da for each modified residue. The small size of the amidino group and its preservation of a positive charge appear to produce minimal disruption in the structure of modified proteins [39, 40]. The excellent agreement between the predicted and experimental extent of amidination of native ribosomal proteins observed in our past experiments indicates that native quaternary structures are preserved as well [28–31].

Scheme 1. SMTA reaction (R = protein or peptide).

2.2 Bacterial growth

Lyophilized *T. thermophilus* HB8 cells obtained from the American Type Culture Collection (ATCC 27634) were rehydrated and cultured in a shaker bath at 70°C according to ATCC instructions. Starter cultures from this growth, containing 15% glycerol by volume, were frozen at −80°C until use. A growth chamber containing 500 mL of ATCC ♯697 medium, supplemented with Castenholtz salts and Nitsch's trace element mixture was temperature equilibrated at 70°C in an incubator cabinet. Growth was initiated by inoculation with one 10-mL starter culture. The cultures were continuously sparged with house air that was filtered through 1 L of sterile water. In addition to removing particulate materials and contaminating pump oil from the air stream, this filter also humidified the sparging air and reduced evaporative loss from the cultures. Culture volumes were maintained at 500 mL by periodic addition of sterile water. Under these conditions, *T. thermophilus* HB8 required about 60 h of growth to reach stationary phase, at which point the cells were harvested. Typically, cells from three separate inoculations were pooled. The cells were yellow–orange and cohesive, but did not stick to the walls of the reactor or centrifuge bottles.

2.3 Sample preparation

A procedure based on Spedding's preparation of bacterial ribosomes [41], previously employed for the preparation of ribosomes from *Caulobacter crescentus* [28], *Escherichia coli* [29], *Deinococcus radiodurans* R1 [30], and *Bacillus subtilis* [31], was employed. Protein concentrations were estimated using the Bradford dye binding assay with bovine serum albumin as a standard [42]. Ribosome preparations had an apparent protein concentration of 11.5 mg/mL.

To characterize the unmodified ribosomal proteome of *T. thermophilus* HB8, ribosomal RNA was removed using Hardy's acetic acid extraction procedure (2 vol. glacial acetic acid and 0.1 vol. 1 M $MgCl_2$, incubated at room temperature for 10 min), followed by centrifugation for 10 min at 14 100 × *g* in an Eppendorf Microfuge (Eppendorf North America, Westbury, NY, USA) [43]. The protein-containing supernatant of an acetic acid extract had a concentration of 3.8 mg/mL and was analyzed directly by two-dimensional chromatography as described below.

In addition to the supernatant of acetic acid extracts, two other samples were generated by reacting either native ribosomes or a solution of denatured proteins from disassembled ribosomes with SMTA. When native ribosomes are reacted with SMTA, surface accessible lysine residues that are not involved in tertiary structural interactions or quaternary interactions with other ribosomal proteins or rRNA are preferentially modified according to Scheme 1 [28–31]. The procedure for modification of native *T. thermophilus* HB8 ribosomes was identical to the procedure employed to modify *D. radiodurans* R1 ribosomes [30]. Equal volumes of ribosome suspension and 200 mM SMTA in 250 mM Tris-free base were mixed and incubated for 1 h at room temperature. The reaction was terminated by addition of $MgCl_2$ and acidification of the reaction mixture with glacial acetic acid to precipitate rRNA as described above. These native amidinated samples were analyzed by two-dimensional chromatography as described below.

To study the effect of pH on the reactivity of proteins in native ribosomes, ribosomes were reacted with SMTA as described above in solutions containing different buffer species. SMTA is isolated as the hydroiodide salt of the amidino group and it releases one equivalent of a strong mineral acid when dissolved, dramatically lowering the pH of the solution: the pH of 250 mM Tris-free base is 10.6, but the pH of the standard native ribosome modification is 8.6. Buffer solutions were prepared so that the addition of SMTA would result in a reaction pH close to the buffer species' pK_a. Solutions of MES and HEPES were prepared by dissolving the free acid form of the buffer in water and adjusting the pH with concentrated ammonium hydroxide. Buffers containing tris were prepared by adjusting a solution of the free base form with concentrated HCl. To modify native ribosomes at a controlled pH, an aliquot of ribosome suspension was mixed with an equal volume of 200 mM SMTA in 2 M MES (pH 6.6), 2 M HEPES (pH 7.5), or 2 M Tris (pH 8.3). The final concentration of each of the Good's buffers in the final solution was 1 M, and the concentration of Mg^{2+} ion was 5 mM. Reactions were incubated for 1 h at room temperature and the reaction was terminated by the precipitation of rRNA with glacial acetic acid and $MgCl_2$. The pH values noted above in parentheses were measured from the reactions after 1 h of incubation at room temperature, prior to the addition of glacial acetic acid and $MgCl_2$, using an IQ400 ion selective field effect transistor (ISFET, Hach Company, Loveland, CO, USA) electrode interfaced to a Handspring Visor handheld computer (Palm, Sunnyvale, CA, USA).

To modify denatured ribosomal proteins from disassembled ribosomes, proteins were removed from the acetic acid extract supernatant by acetone precipitation [44]. A 100-μL aliquot of acetic acid extract was chilled on ice, and then 5 vol. ice-cold acetone were added. The mixture was allowed to stand on ice for 1 h, and then the precipitated proteins were separated from the supernatant by a brief spin (*ca.* 1 min) at 1000 × *g*. The aqueous acetone supernatant was removed by aspiration and the precipitated proteins were resuspended in 50 μL of aqueous ACN (30/70 water/ACN v/v) containing 6 M urea and 25 mM ammonium bicarbonate. When the protein was fully redissolved, an equal volume of 200 mM SMTA in 250 mM Tris-free base was added and the reaction was incubated for 1 h at room temperature. The reaction was stopped by the addition of 10 μL of glacial acetic acid, and the denatured, disassembled protein sample was analyzed by 2-D LC.

2.4 Protein chromatography

A previously described automated 2-D-LC system was used to fractionate proteins from acetic acid extracts of ribosomal proteins before or after their modification [45]. The first dimension of this separation uses strong cation exchange chromatography (SCX, Toso-Haas SP-NPR SCX column, 4.6 mm × 35 mm, Tosoh Bioscience, Montgomeryville, PA, USA) developed with the gradient shown in Supporting Information Table 1, where mobile phase A is 20 mM acetic acid and 6 M urea, adjusted to pH 5.1 with methylamine, and mobile phase B is the same except for the addition of 500 mM NaCl. The gradient was supplied by a Waters Alliance 2695 chromatograph. Proteins eluted from the SCX dimension were trapped on short C4 columns (Thermo BioBasic C4 Javelin, 1 mm × 20 mm). Excess salt and urea were washed away to waste and the trapping columns were then serially connected in line to a longer C4 column (Thermo BioBasic Pioneer C4, 1 mm × 100 mm) and developed with the reversed-phase gradient shown in Supporting Information Table 2, where mobile phases A and B are water and ACN, each containing 0.1% TFA and 0.2% formic acid. The gradient was supplied by a Waters Alliance 2795 chromatograph.

The output of the second dimension of the automated 2-D LC system can either be developed directly into a mass spectrometer to generate whole protein data or fractionated for subsequent digest and peptide analysis. For whole protein analysis of 2-D LC experiments, the contents of the trapping columns were developed directly into a Waters Micromass QTOF Micro mass spectrometer, and total ion chromatograms corresponding to a time window from 15 to 55 min were collected. The 50 μL/min effluent was split down to 5 μL/min. Raw spectra were deconvoluted using either MaxEnt 1 from Waters Micromass or Bioanalyte ProTrawler/Regatta (BioAnalyte, Portland, ME, USA). Experimental deconvoluted masses are typically within 1.5 Da of values calculated from protein sequences.

More accurate masses were measured using a Thermo LTQ-FT Ultra hybrid ion trap/FT-ICR MS (Thermo Fisher, Bremen, Germany) equipped with a nano-ESI source. Aliquots of acetic acid extracted ribosomal proteins were concentrated and desalted on homemade trapping columns fabricated in 150 μm id fused-silica capillaries using a modification of the method of Wang *et al.* [46], and chromatographed on fused-silica nano-ESI tips (75 μm × 15 cm) prepared with a Sutter Instruments P-2000 Micropipette Puller (Sutter Instruments, Novato, CA, USA). Both traps and nanospray tips were packed with Phenomenex Jupiter C4 beads. The gradient shown in Supporting Information Table 3 was delivered by a Dionex Ultimate-3000 chromatograph (Dionex Corporation, Sunnyvale, CA, USA). Spectra were collected across a window of 300–2000 Th at a resolution of 100 000 and stored as centroided spectra. Raw spectra were extracted by summing across chromatographic peaks and deconvoluted using Xtract (Thermo Fisher, San Jose, CA, USA). The reported isotopically resolved masses were obtained directly from the deconvoluted spectra.

2.5 Peptide analysis

Protein containing fractions from the 2-D LC protein separation were collected for tryptic digestion by injecting 16 μL of 80% aqueous isopropyl alcohol and isocratically eluting with 90% mobile phase B. Dried samples were resuspended in 20 mM ammonium bicarbonate and digested with alkylated porcine trypsin, incubated 12–16 h at 37°C, and acidified with 0.2% formic acid in water for LC-MS/MS analysis. Peptide separations were performed using a Dionex Ultimate 3000 chromatography system. Peptide digests were concentrated and desalted on homemade trapping columns fabricated in 150 μm id fused-silica capillaries and analyzed on packed 75 μm id nanospray tips. Traps and nanospray tips were packed with Michrom Magic C18 beads (5 μm diameter, 100 Å pore size, Michrom Bioresources, Auburn, CA, USA). The gradient used is shown in Supporting Information Table 4. LC-MSMS of tryptic digests of ribosomal proteins was accomplished using a hybrid Thermo LTQ-FT. Instrument methods measured peptide precursor masses in the FT cell while the five most intense masses were collisionally fragmented in the linear ion trap. Total ion chromatograms were converted to MASCOT Generic Format using TurboRAW2MGF [47], and peptides were identified by MASCOT searches against the *T. thermophilus* HB8 proteome. Additional assignments in MS/MS spectra were generated by entering peptide sequences into UCSF Prospector MS-Product module (prospector.ucsf.edu). Amidinated lysine residues were represented as a user-defined residue and the maximum charge state was set to +3 or +4.

2.6 Bioinformatics

The genome and proteome of *T. thermophilus* HB8 were downloaded from The Institute for Genome Research's Comprehensive Microbial Resource (http://www.tigr.org, now the J. Craig Venter Institute). Because of the error in the assignment of ribosomal protein S12's start site discussed below, the sequence of S12 was downloaded from the Swiss-Prot database (entry Q5SHN3, www.expasy.org). Predictions of the extent of labeling of proteins in the native ribosome were assessed using the structures of the 50S and 30S subunits from the published 2.8 Å crystal structures of the 70S ribosomal particle of *T. thermophilus* HB8 (50S PDB file 2J01 and 2J03; 30S PDB files 2J00 and 2J02) [32]. Structures were visualized using PyMOL v. 0.99 (DeLano Scientific, www.pymol.org). The maximum extent of modification of each protein in a native ribosome was determined as detailed earlier [28–31]. Protein sequence homologies were evaluated

using blast2p at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/blast/).

3 Results

3.1 Whole proteins

In a recent paper, Suh *et al*. measured the masses of 52 of the 54 predicted ribosomal proteins of *T. thermophilus* HB8 using MALDI-TOF, failing only to observe small subunit proteins S1 and S2 [48]. During the course of studying the pH dependence of SMTA reactivity of these proteins in native ribosomes, we have measured the mass of all 53 commonly observed proteins with an ESI QTOF, failing only to observe the largest small subunit protein, S1. These initial identifications are corroborated with ESI FT-ICR measurements, shown in Table 1. Our protein assignments are further supported by MS/MS analysis of peptides isolated from trapping columns that contain proteins with the listed masses. It is also noteworthy that the digestion of several whole protein fractions yielded peptides comprising 77% sequence coverage of small subunit protein S1. However, typical of previous attempts to observe intact protein S1 directly in a bacterial ribosome mass spectrometric study, we were unable to find a mass corresponding to the 59 970.4 Da mass predicted by the proteome sequence in our whole protein analysis [28–31, 48]. In mesophilic bacteria such as *E. coli*, S1 associates most strongly with actively translating polysomes [49], and is usually not observed in ribosomes prepared according to Spedding's procedure, leading some to speculate that it is not a genuine ribosomal protein [50]. The presence of this protein in *T. thermophilus* HB8 ribosome samples may be due to the greater structural resilience of this thermophilic bacterium's proteins and their complexes.

An FT-ICR MS instrument was used to reinforce these identifications by improving the accuracy of the whole protein mass measurements. Because this instrument routinely resolves isotope structure in mass spectra, some protein mass values appear to differ from the isotopically averaged numbers measured with the TOF spectrometer. However, our assignments are consistent. The fifth through seventh columns of Table 1 list calculated isotopically resolved masses, the corresponding experimental values, and the absolute value of the differences between calculated and experimental masses in parts *per* million. The high resolving power of an FTMS instrument ($R = 100\,000$ for these experiments) requires additional calculation to arrive at a theoretical mass predicted from a protein's sequence for comparison with experimental data. The masses in the fifth column of Table 1 are calculated from a molecular formula derived from the protein's sequence using the ProtParam tools provided by ExPASy. The molecular formula is modified to include sequencing errors, PTMs, and the formation of non-native disulfide bonds due to oxidation during sample handling. Because the intensity of the monoisotopic peak of a protein with a mass over 10 kDa will be insignificant [51], series of isotopic masses are calculated by adding multiples of the mass increment between isotopic mass peaks, 1.00235 Da [52]. The experimental masses in the sixth column of Table 1 correspond to the most intense isotopic mass in the deconvoluted spectrum for each protein [53] For example, based on its sequence in the *T. thermophilus* HB8 proteome, ribosomal protein L18 has a molecular formula of $C_{561}H_{939}N_{177}O_{151}S_1$. Subtraction of a methionine residue, C_5H_9NOS, results in a calculated monoisotopic mass of 12 473.05 Da for the uncharged protein. The mass of the seventh isotopic peak above the monoisotopic mass of this protein is calculated by adding 7×1.00235 Da to this value to give the tabulated mass of 12 480.07 Da. The most intense mass observed in the experimental isotopic peak distribution, 12 480.07 Da, agrees with the calculated value to within 1 ppm. Accurate mass measurements were collected for all *T. thermophilus* HB8 ribosomal proteins except protein THX, a small subunit protein unique to *T. thermophilus*, which was not observed in chromatograms detected with the FT instrument. These mass measurements show low relative errors (typically 1–4 ppm) and are consistent with modifications that we have proposed for each protein. The somewhat lower accuracy of several mass measurements is probably due to the lower intensity of some proteins' signals that are the result of poor chromatographic separation.

Table 1 also lists the total number of amidination sites (*i.e.* the number of lysine residues plus one for the amino terminus), and the extent of modification observed for protein samples derived from denatured, disassembled ribosomes. We have previously observed that exhaustive amidination of denatured, disassembled ribosomes helps to corroborate protein identifications, proposed PTMs, or possible sequence errors, as the amidination mass shift allows one to count the number of unmodified lysine residues in the protein in a very straightforward manner [30]. Two examples of the application of this technique to *T. thermophilus* HB8 ribosomal proteins L3 and L11 are presented in Fig. 1. Deconvoluted whole protein spectra for unmodified and modified ribosomal protein L3 are shown in Fig. 1A and B, respectively. Comparison of these two panels demonstrates unambiguously that this protein has 19 free amino groups as predicted from the proteomic sequence. Spectra for unmodified and SMTA-modified ribosomal protein L11 are shown in Fig. 1C and D. The mass of the most intense peak in the spectrum corresponds to the addition of amidino groups to 10 of protein L11's 14 modifiable amino groups, a result that is discussed below.

3.2 pH variation of native amidination

To determine the effect of pH changes on the reactivity of the surface-exposed lysine residues of proteins in native ribosomes, we reacted ribosomes with SMTA in three

Table 1. Observed ribosomal proteins *Thermus thermophilus* HB8

Protein	Isotopically averaged masses			FT-ICR determined masses			Modifications[g]	Extent of amidination[a]			% Sequence coverage[b]
	Calc.	Obs.	Δm[c]	Calc.[d]	Obs.[e]	Ppm[f]		Maximum	Predicted	Disassembled	
L1	24 830.6	24 702.2	−128.4	24 699.34	24 699.31 – *15*	1	-Met	21	15	21	51
L2	30 468.2	30 340.4	−127.8	30 338.67	30 338.71 – *20*	2	-Met	26	14	26	62
L3	22 408.1	22 438.0	29.9	22 434.19	22 434.21 – *12*	1	+2(CH_2)	19	10	19	50
L4 (L1e)	23 234.8	23 237.9	3.1	23 230.56	23 230.60 – *10*	2		15	10	15	70
L5	21 029.6	20 899.8	−129.8	20 891.38	20 891.45 – *6*	4	-Met	13	8	13	69
L6	19 531.8	19 402.6	−129.2	19 399.88	19 399.81 – *11*	3	-Met	17	11	17	78
L7/L12	13 067.2	12 937.1	−130.1	12 935.23	12 935.24 – *7*	1	-Met	18	n.p.	18	40
L9	16 397.2	16 398.7	1.5	16 396.19	16 396.19 – *9*	<1		14	14	14	63
L10	18 565.7	18 435.2	−130.5	18 431.16	18 431.09 – *8*	3	-Met	14	n.p.	14	20
L11	15 505.1	15 674.4	169.3	15 672.48	15 672.48 – *9*	<1	+12(CH_2)	14	n.p.	10	39
L13	15 894.8	15 896.4	1.6	15 894.83	15 894.75 – *10*	5		18	14	18	68
L14	13 302.6	13 303.7	1.1	13 302.31	13 302.31 – *8*	<1		12	8	12	82
L15	16 281.0	16 282.1	1.1	16 281.08	16 280.86 – *10*	13		19	10	19	55
L16	15 962.8	15 964.1	1.3	15 961.63	15 961.64 – *9*	1		15	7	15	55
L17	13 715.0	13 716.0	1.0	13 714.76	13 714.75 – *8*	1		10	4	10	61
L18	12 611.8	12 481.1	−130.7	12 480.07	12 480.07 – *7*	<1	-Met	12	9	12	57
L19	17 151.7	17 153.0	1.3	17 149.43	17 149.43 – *8*	<1		12	9	12	65
L20	13 743.1	13 611.7	−131.4	13 610.67	13 610.57 – *7*	7	-Met	15	4	15	54
L21	11 047.1	11 048.1	1.0	11 046.32	11 046.33 – *6*	<1		14	8	14	77
L22	12 780.0	12 781.3	1.3	12 779.16	12 779.12 – *7*	3		12	7	12	62
L23	10 736.8	10 737.4	0.6	10 736.11	10 736.13 – *6*	2		17	12	17	64
L24	12 056.5	12 056.0	−0.5	12 051.74	12 051.74 – *7*	<1		21	15	21	70
L25	23 204.5	23 207.0	2.5	23 195.27	23 195.36 – *5*	4		13	9	13	54
L27	9508.0	9377.6	−130.4	9376.16	9376.16 – *5*	<1	-Met	7	5	7	9
L28	10 978.1	10 847.7	−130.4	10 846.31	10 846.30 – *6*	<1	-Met	16	9	16	62
L29	8650.2	8650.8	0.6	8648.92	8648.92 – *4*	1		10	6	10	68
L30	6785.1	6654.2	−130.9	6652.99	6652.89 – *3*	<1	-Met	8	6	8	83
L31	8285.5	8285.6	0.1	8281.12	8281.13 – *5*	1		6	5	6	0
L32	6705.0	6570.7	−134.3	6569.44	6569.44 – *4*	<1	-Met	9	2	9	0
L33	6615.8	6482.6	−133.2	6480.44	6480.44 – *4*	<1	-Met	9	5	9	0
L34	6109.3	6107.2	−2.1	6108.58	6108.57 – *3*	1		8	4	8	0
L35	7484.1	7353.1	−131.0	7352.29	7352.30 – *4*	<1	-Met	16	5	16	67
L36	4421.3	4421.3	0.0	4418.41	4418.41 – *2*	<1		7	n.p.	7	0
S2	29 276.6	29 148.6	−128.0	29 143.46	29 143.35 – *16*	4	-Met	15	14	15	46
S3	26 700.9	26 572.3	−128.6	26 563.74	26 563.61 – *10*	5	-Met	17	8	17	55
S4	24 324.3	24 193.3	−131.0	24 195.06	24 195.18 – *21*	5	-Met	15	11	15	55
S5	17 557.4	17 427.8	−129.6	17 425.44	17 425.41 – *10*	2	-Met	9	6	9	48
S6	11 972.8	11 973.6	0.8	11 971.33	11 971.33 – *6*	<1		5	5	5	89
S7	18 015.9	17 886.6	−129.3	17 885.57	17 885.55 – *12*	1	-Met	13	10	13	50
S8	15 837.5	15 838.8	1.3	15 836.73	15 836.72 – *9*	1		10	6	10	77
S9	14 382.5	14 383.7	1.2	14 382.80	14 382.81 – *9*	<1		12	6	12	68
S10	11 929.9	11 799.6	−130.3	11 925.61	11 925.45 – *3*	13		9	5	9	74

Table 1. Continued

Protein	Isotopically averaged masses			FT-ICR determined masses			Modifications[g]	Extent of amidination[a]			% Sequence coverage[b]
	Calc.	Obs.	Δm[c]	Calc.[d]	Obs.[e]	Ppm[f]		Maximum	Predicted	Disassembled	
S11	13 712.8	13 584.4	−128.4	13 582.29	13 582.26 – *9*	2	-Met	16	7	16	51
S12	14 599.2	14 514.9	−84.3	14 513.29	14 513.23 – *8*	4	-Met, $+SCH_3$ [h]	22	14	22	37
S13	14 304.7	14 174.8	−129.9	14 172.99	14 172.95 – *8*	2	-Met	14	11	14	62
S14	7139.7	7008.5	−131.2	7003.89	7003.90 – *4*	2	-Met.	8	5	8	0
S15	10 554.3	10 423.9	−130.4	10 422.76	10 422.76 – *6*	<1	-Met	8	6	8	62
S16	10 386.9	10 387.4	0.5	10 386.60	10 386.50 – *6*	10	Also modified[i]	8	5	8	46
S17	12 297.6	12 167.2	−130.4	12 169.88	12 169.84 – *11*	3	-Met	16	9	16	62
S18	10 231.2	10 101.3	−129.9	10 101.96	10 101.83 – *8*	13	-Met	15	9	15	45
S19	10 581.4	10 450.9	−130.5	10 448.72	10 448.71 – *5*	1	-Met	14	9	14	68
S20	11 703.0	11 572.5	−130.5	11 573.83	11 573.80 – *9*	2	-Met	20	10	20	53
THX	3337.0	3206.5	−130.5		No appearance		-Met	7	3	7	0

a) Maximum extent of amidination (number of lysines plus 1 amino terminus), number of accessible lysines predicted by inspection of crystal structures, and observed extent of modification under denaturing conditions. Predicted values of ''n.p.'' indicate that the protein is not present in the crystal structure.

b) Maximum percent sequence coverage seen in LC-MS/MS separations of tryptic digests of native amidinated samples. Smaller proteins provide too many small tryptic peptides to be observed, and are annotated with ''0%.''

c) Observed experimental mass minus calculated isotopically averaged mass based on the proteome sequence.

d) Calculated isotopic mass that takes into account all PTMs, including the possibility of non-native disulfide bonds.

e) Most intense isotopic peak. The hyphenated, italicized suffix indicates the specific isotopic mass isotopomer identified (see text).

f) Absolute value of the parts-per-million difference between calculated and experimental masses. ''<1'' indicates a value less than 1 or greater than −1.

g) Modifications: ''-Met'': N-terminal methionine removed, ''$+n(CH_2)$'': addition of *n* methyl groups.

h) S12 homologs in *E. coli*, *R. palustris, D. radiodurans,* and *B. subtilis* are β-thiomethylated (+46.1 Da) at an aspartate homologous to D88 in *E. coli*. Significant amounts of unmodified *T. thermophilus* S12 are also observed in our experiments.

i) Second amino acid in the sequence is V. Trace amounts of ''-Met'' S16 are also observed.

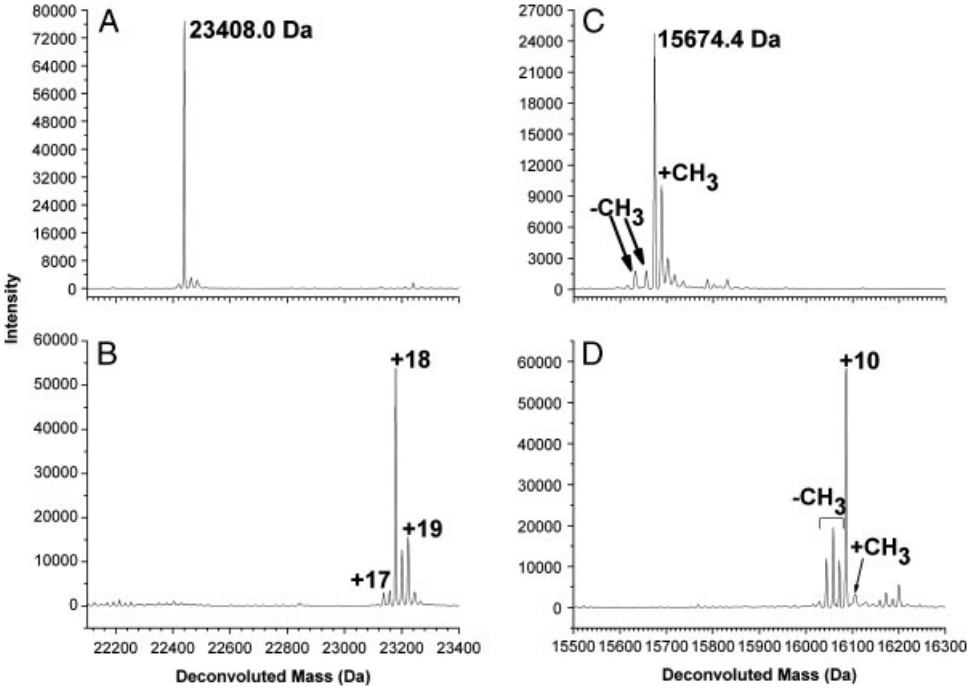

Figure 1. Deconvoluted mass spectra of large subunit proteins L3 and L11. (A) Deconvoluted spectrum of unmodified protein L3. The label indicates the experimental mass of the protein, including PTMs. (B) Deconvoluted spectrum of L3 from a denatured, disassembled ribosomal protein sample after reaction with SMTA. (C) Deconvoluted spectrum of protein L11. The label indicates the experimental mass of the protein, including PTMs. ''$-CH_3$'' and ''$+CH_3$'' labels indicate forms of L11 that are under- or overmethylated with respect to the major form that contains 12 methylations. (D) Deconvoluted spectrum of L11 from denatured, disassembled ribosomal protein sample after reaction with SMTA. In spectra (B) and (D), numerical labels indicate the number of amidino groups added to each protein.

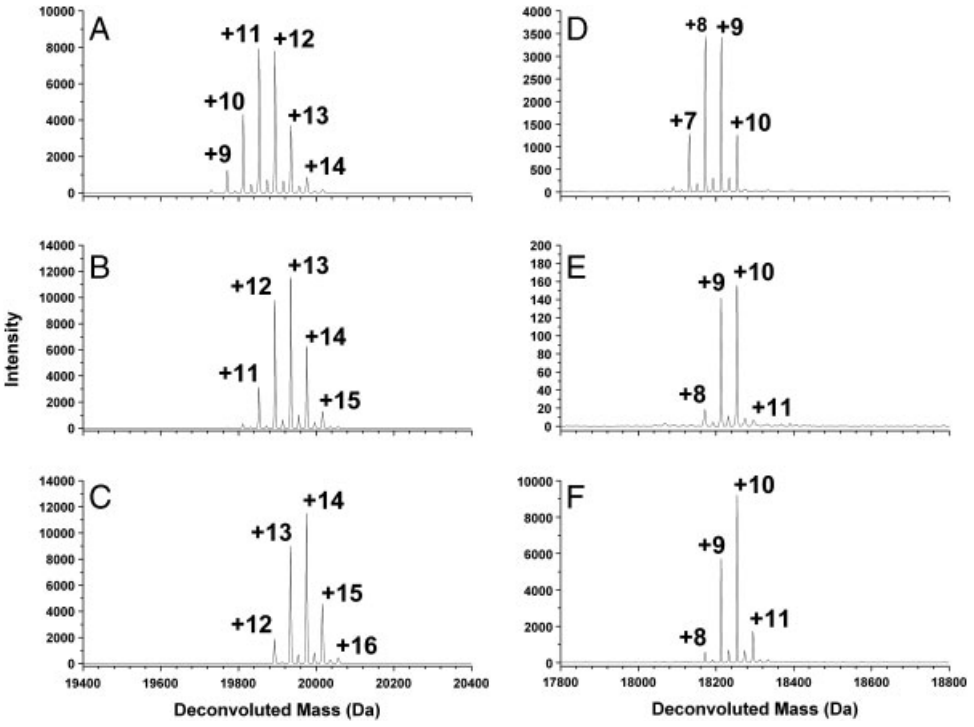

Figure 2. Changes in the average extent of labeling as a function of pH. (A–C) Large subunit protein L6 labeled with SMTA at pH 6.6 (A), pH 7.5 (B), and pH 8.3 (C). (D–F) Small subunit protein S7 at the same three pH values: pH 6.6 (D), pH 7.5 (E), and pH 8.3 (F). Numerical labels indicate the number of amidino groups associated with the observed mass shift.

different buffers: MES (pH 6.6), HEPES (pH 7.5), and Tris (pH 8.3), under non-denaturing conditions. Figure 2 shows examples of the MS data collected for whole ribosomal proteins extracted from ribosomes reacted with SMTA at each pH. The first column, panels A–C, shows the change in the extent of labeling for large subunit protein L6 at each pH. Analogous data for small subunit protein S7 are shown in panels D–F. Both series show an increase in the extent of amidination for each protein, as indicated by the labels attached to each peak. Each protein's extent of modification does not reach the maximum value predicted from their sequences but does approach the value predicted from inspection of the ribosome crystal structure to determine surface accessible lysine residues at the highest

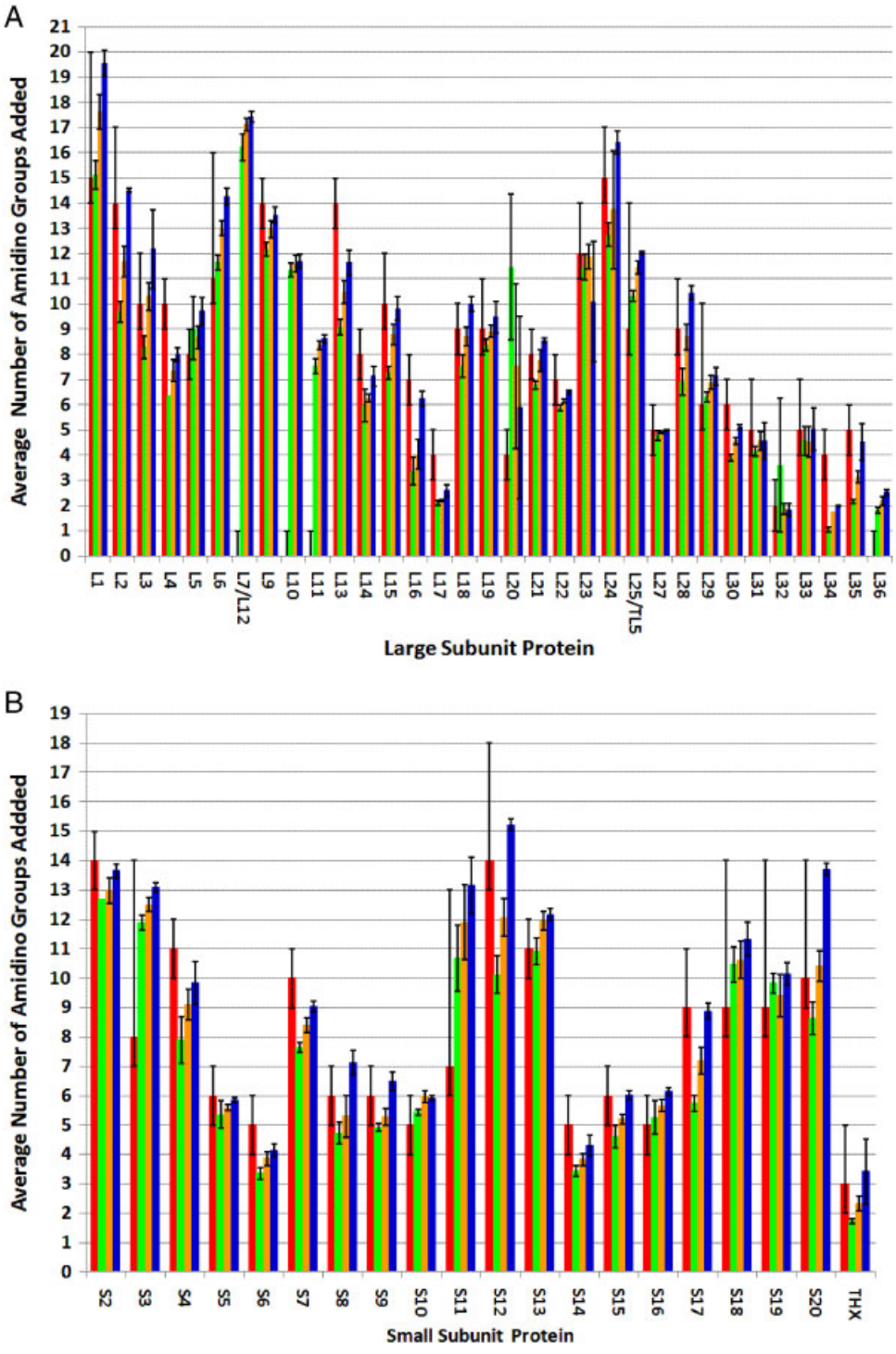

Figure 3. Weighted average extent of lysine labeling of *T. thermophilus* HB8 ribosomal proteins as a function of pH. (A) Large subunit proteins. (B) Small subunit proteins. In each cluster of data, the red bar represents the extent of labeling predicted from examination of the crystal structure. Green bar: MES buffer, pH 6.6. Orange bar: HEPES buffer, pH 7.5. Blue bar: Tris buffer, pH 8.3. Error bars are the result of three independent measurements or as described in the text.

experimental pH value, 8.3. This behavior indicates that native protein–rRNA interactions are preserved during the labeling reaction at each pH and suggests that the changes in protein reactivity are due to pH-dependant changes in the environment of surface-exposed lysines. By using the integrated area intensity and the number of amidino group additions represented by each peak in spectra like those in Fig. 1, average extents of modification for each protein in the sample can be calculated for pH of each solution studied. Extent of modification data for amidination of ribosomal proteins under native conditions is depicted as histograms in Fig. 3A and B. These values were derived from data like those shown in Fig. 2 by taking intensity weighted averages of the number of amidino groups added to each protein. Figure 3A summarizes the pH variation of SMTA reactivity for large subunit proteins, and Fig. 3B shows analogous data for small subunit proteins. In Fig. 3, the left-most value in the cluster of data for each protein is the predicted extent of labeling, derived from inspection of the crystal structures of *T. thermophilus* HB8's 50S and 30S subunits, represented with a red bar. Experimental data follow from left to right for pH 6.6 (green bars) through pH 7.5 (orange bars) to pH 8.3 (blue bars). Error bars shown for experimental values are the average of at least three determinations, while the error bars attached to the predicted reactivity assume a ± 1 error in counting lysines. In some cases there is an additional positive error due to crystallographic disorder as exemplified by ribosomal proteins L1 and S12. Ribosomal protein L1's structure in PDB file 2J01 lacks the first 17 and last 4 amino acid residues in the protein's sequence and a short internal loop containing residues 29–34, including lysine residues 3, 6, 14, and 31. This absence is reflected by increasing the positive limit of the error bar on the predicted extent of modification to five residues. The crystal structure of small subunit protein S12 is likewise missing four residues from its N-terminal region and 13 residues from the C-terminus, including lysine residues 130, 134, and 135, resulting in an increase by four residues on the positive error bar. No predictions for the extent of modification are shown for proteins L7/L12, L10, L11, and L36 since these proteins are missing from the crystal structure.

Most of the proteins in the sample show the same general trend in reactivity: lower extent of modification at lower pH, increasing to the extent of modification predicted by inspection of the crystal structure. The observed changes in protein reactivity are unquestionably due to the combined effects of pH on the mechanism of SMTA reaction with amino groups and on the conformation of the proteins being modified. However, the reactivity of ribosomal protein L7/L12 as a function of pH serves as an internal control that allows us to differentiate between the contributions of these two sources of variation. Protein L7/L12 is one of the few proteins in the ribosome that interacts only with other proteins, rather than with rRNA. Dimers of L7/L12 interact with ribosomal protein L10 to form a structurally resilient pentamer or heptamer with protein L10 *via* L7/L12's N-terminal region (the details of the interaction depend on the bacterial species [9]). Protein L7/L12 is extremely flexible, as evidenced by its absence from crystal structures of whole ribosomes [32, 36], and has been hypothesized to serve as a docking clamp for elongation factors during the translation cycle [54]. As the pH of the solution is raised from 6.6 to 8.3, the average extent of modification of L7/L12 increases from 16.2 ± 0.5 to 17.4 ± 0.2 amidino groups added out of a maximum of 18 possible amidinations. These data indicate that the maximum possible variation in protein labeling due to changes in the rate or mechanism of amidination as a function of pH is ± 1 amidino group, which is comparable to our error in predicting lysine reactivity by examining crystal structure.

Most of the proteins in Fig. 3 show only minor variation in their extents of labeling under native conditions as the pH increases from 6.6 to 8.3. Out of 53 proteins, 29 show an average change their reactivity of only 1 lysine residue across this pH range. This is an important point because it indicates that overall, native structure is being preserved in solution. If changes in solution composition or ionic strength were disrupting the structure of the ribosomes, one would expect more dramatic changes in the number of amidino groups added across the entire range of ribosomal proteins. *T. thermophilus* HB8 ribosomal proteins can be divided into three groups according to the difference between their extent of labeling at pH 8.3 and 6.6. A total of 29 proteins, including large subunit proteins L5, L7/L12, L11, and L27, show minimal alterations in their average extent of labeling in the pH range 6.6–8.3, with an average increase in the extent of modification of one lysine residue as the pH of the solution increases. Proteins that show labeling behavior of this type are either highly solvent exposed or deeply buried in the ribosomal RNA. The proteins of the ribosomal stalk, L7/L12, provide examples of highly solvent accessible proteins. Because most of this protein's lysine residues are available for reaction with SMTA, protein L7/L12 shows nearly constant reactivity as a function of pH, with an increase of only 1.2 lysine residues reacted as the pH is increased. Ribosomal protein L27 provides an example of a protein that is deeply buried in the 23S rRNA structure. Because it is inaccessible to SMTA, this protein's extent of reactivity increases by only 0.2 lysine residues as the pH is increased from 6.6 to 8.3. A second group of 15 proteins, including proteins L1, L2, L3, L6, L16, S8, and S12, shows a broader variation of reactivity across the pH range. These proteins have larger numbers of solvent accessible lysine residues (shown in Table 1 in the predicted column). Examples discussed in detail below provide a molecular explanation for the pH-dependant variation of the SMTA reactivity of this second group of proteins. Finally, eight proteins, including L4, L13, L20, and S4, show anomalous behavior that does not fit into either previously described category. These proteins show anomalous changes in reactivity with SMTA as a function of pH.

4 Discussion

4.1 Whole proteins

A recent paper by Suh *et al.* produced a nearly complete list of the ribosomal proteins of *T. thermophilus* HB8 [48]. Our results in Table 1 of this paper extend Suh *et al.*'s results by adding a whole protein mass for ribosomal protein S2. In addition, our identifications are supported by MASCOT searches against LC-MS/MS data for tryptic peptides or highly accurate mass measurements with an FT-ICR. The FT instrument was particularly effective for corroborating the identification of the small ribosomal proteins, such as L29–L36 and S14. These small proteins have larger than average percentage of lysine and arginine residues, ranging from 18% in L31 to 45% in L34 *versus* an average occurrence of K+R of 11% [55]. This high abundance of tryptic cleavage sites results in numerous single amino acids, di- and tri-peptides, and low sequence coverage for these proteins. In the past, we have used C-terminal sequencing with a mixture of fungal carboxypeptidases to confirm assignments of small ribosomal proteins [30]. This procedure works well, but is time consuming because it requires empirical adjustment of the carboxypeptidase concentration and the chromatographic assay of samples from fixed time points during the digestion. Nanospray chromatography and FT-ICR analysis require less sample handling and produce accurate identifications from a single separation. Only two identifications in Table 1 are in conflict with predictions based on the proteome retrieved from The Institute for Genome Research (TiGR).

Ribosomal protein S12 undergoes a unique PTM, the formation of a secondary methylthioether at the beta position of an aspartate residue, D88 in *E. coli* [56], D92 in the proteomic sequence of *T. thermophilus* HB8. The 14 514.9 Da mass reported in Table 1, as well as the 14 516 Da mass assigned as S12 by Suh *et al.*, agrees with the mass predicted from the sequence in Swiss-Prot record Q5SHN3 rather than the sequence presented as ORF TTHA1697 of the proteome retrieved from TiGR. Our chromatograms regularly show nearly equivalent amount of S12 with and without the PTM. The presence of two forms of this protein may be an artifact of the *T. thermophilus* HB8 cultures having been grown into stationary phase. The mass predicted from ORF TTHA1697 after accounting for PTMs, 14 797.5 Da, was absent from any whole protein experiments. However, a mass of 14 514.9 Da was observed in fractions that yielded 46% sequence coverage of S12. The discrepancy is due to incorrect assignment of the protein's start codon. The genomic sequence assigns a valine codon (GTG) as the start codon, while the correct choice appears to be the leucine codon (CTG) nine nucleotides further into the sequence. Both valine and leucine codons are used infrequently by bacteria as alternative start codons, although their frequency of use is far below that of the canonical methionine start codon ATG [57]. Reassignment of the start codon results in a predicted mass of 14 599.2 Da, or 14 514.1 Da after accounting for the removal of methionine and methylthiolation at position D89. Although there are no peptides in our tryptic digest samples containing S12's modified D89, highly accurate mass measurements with an ICR FT-MS instrument corroborate the identification of this mass as protein S12. The most intense isotopic mass peak observed for this mass was the eighth, with a mass of 14 513.29 Da, 4 ppm lighter than the mass calculated for S12 with the PTMs of methionine removal and methylthiolation.

The observed mass of small subunit protein S16 indicates that the N-terminal methionine is retained in the protein as isolated from *T. thermophilus* HB8 cells. This is noteworthy because the second residue in the sequence is a valine residue, and valine residues have a small enough steric bulk that proteins with a penultimate valine can serve as substrates for methionine aminopeptidase. Retention of protein S16's N-terminal methionine can be explained by the substrate specificity of methionine aminopeptidase. Valine-containing peptides are among the slowest substrates for this enzyme, showing 60% cleavage efficiency *in vivo* relative to peptides containing glycine, alanine, serine, threonine, and proline in the second position [58]. Although the mass observed for the sixth isotopic mass peak above the monoisotopic mass of S16 with FT-MS is 10 ppm lower than the predicted mass, 10 386.50 Da *versus* 10 386.60 Da, this accuracy is more than sufficient to corroborate the retention of the N-terminal methionine.

The utility of combining whole protein mass measurements with SMTA labeling of denatured proteins from disassembled ribosomes is illustrated by the results displayed in Fig. 1A and B. Figure 1A displays the deconvoluted spectrum for large subunit protein L3 from an unmodified ribosomal protein extract. The experimental mass for this protein, 22 438.0 Da, is 29.9 Da larger than the mass predicted from the proteomic sequence. The most straightforward explanation for this mass difference is PTM of the protein by two methylations. Also shown in Table 1 is the mass obtained for protein L3 with ICR FT-MS, 22 434.21 Da. This measurement is accurate to within less than 1 ppm of the theoretical mass calculated by assuming two methylations and provides additional support for the proposed PTM. Tryptic digests of L3-containing fractions did not reveal any methylated peptides. However, examination of the SMTA labeling pattern in Fig. 1B provides information on the chemical nature of the PTMs. Protein L3's sequence predicts 19 amidine reactive sites, 18 lysine residues, and the protein amino terminus. Figure 1B shows that the protein with a mass of 22 438.0 Da has 19 modifiable amino groups, since the largest mass in the spectrum derived from a sample of denatured, disassembled proteins corresponds to the addition of 19 amidino groups, supporting the identification of protein L3 and the conclusion that the PTM applied to protein L3 is two separate methylations rather than a dimethylation, which would render the residue unreactive to SMTA, or monomethylation of the protein's N-terminal amino group. Liu and Reilly

have shown that monomethylated lysine residues can be amidinated, while monomethylated amino termini cannot [29]. This observation rules out the possibility that either methylation is applied to the N-terminus of *T. thermophilus* HB8 L3. The canonical PTM of protein L3 observed in *E. coli*, γ-N-methylation at glutamine 150 [59], would not interfere with amidination either. When the protein L3 sequences from both organisms are compared, *E. coli*'s Q150 aligns with *T. thermophilus* HB8's K145. Despite the substitution of a functionally homologous residue for Q150, K145 is still a potential substrate for post-translational methylation. The low intensity of the fully amidinated, +19, peak in the spectrum of Fig. 1B is due to the structural resiliency of proteins isolated from an organism with an optimal growth temperature of 70°C. Even under the conditions used to modify the proteins shown in Fig. 1B, some residual secondary or tertiary structure impedes the reaction of SMTA at one site in the protein, resulting in less than total modification. In past results using the ribosomal proteins from organisms with less extreme optimal growth temperatures, 20–37°C, it has been typical to observe some evidence of less than complete SMTA derivatization with proteins such as L7/L12 that are noted for retaining significant structure even under strongly denaturing conditions such as our SCX separation conditions [60].

The spectra of unmodified and disassembled, denatured protein L11 in Fig. 1C and D provide another example of the use of SMTA modification to identify a protein whose PTMs cause its experimental mass to differ from the predicted value. Ribosomal protein L11 is modified by the addition of ten amidino groups under denaturing conditions, as shown in Fig. 1D, consistent with the commonly observed permethylation of bacterial L11 homologs at the amino terminus and two or more lysine residues [28–31, 61]. The deconvoluted spectra of L11 in Fig. 1C also shows some heterogeneity in the extent of methylation, with noticeable intensity for forms of the protein lacking 1, 2, or 3 methylations, and a form with an extra methylation. Although such heterogeneity has been observed before in *C. crescentus* L11 methylation as a function of growth conditions (Running, W. E., Reilly, J. P., unpublished observations), there is probably no significance to the observation, as both *E. coli* and *T. thermophilus* cells with non-functional L11 *prmA* methylase enzymes show no growth defects [62, 63].

4.2 pH-dependant labeling

Past experience in our lab with SMTA [28–31] along with the present results demonstrate that SMTA reactivity with surface accessible lysine residues is practically complete, even at a pH value as low as 6.6, which represents a hydrogen ion concentration four orders of magnitude higher than lysine's pK_a of 10.5 [1]. At neutral pH (7.0), only 0.1% of lysine amino groups are in the more nucleophilic unionized form. However, reaction with SMTA traps unionized lysine residues and pulls the ionization equilibrium to the right. Lysine reactivity is primarily limited by structural factors. The observation that changes in ribosomal protein reactivity as a function of pH are not uniform supports the conclusion that the results summarized in Fig. 3 are not due to the simple denaturation of the ribosomal particle, but instead reflect functionally significant alterations in the flexibility of specific proteins. The reactivity of some proteins increases only by the modification of a single additional lysine residue as the pH increases from 6.6 to 8.3. As mentioned above in the discussion of the use of L7/L12 reactivity as an internal control, this small amount of variation may be due either to changes in the rate of reaction of SMTA with amino groups or to small changes in protein conformation. The reactivity of some proteins increases by three lysine residues or more, with L1, L2, S12, and S20 exposing as many as five additional lysine residues at the higher pH. If the changes in ribosomal protein labeling were due solely to pH dependence on the rate of reaction of lysine residues with SMTA, or on an unfavorable equilibrium for amidino-lysine product formation, one might expect order of magnitude differences in the extent of modification as a function of pH. More significantly, the effects of reaction rates or equilibrium position would be uniform across the entire population of proteins, rather than being concentrated in a small group of proteins. The ionic strength of the reaction solutions could also cause changes in protein reactivity. The SMTA is prepared as the hydroiodide salt of the thioamide, and this necessitates buffer concentrations higher than those normally recommended for biochemical procedures (*e.g.* 1 M *versus* 50–100 mM) to maintain constant pH in the modification reaction. However, all proteins' average extents of modification at the highest pH value studied are within error limits of the maximum extents of lysine modification predicted by inspection of the crystal structures for *T. thermophilus* HB8 ribosomes. This observation makes complete dissociation or denaturation of the proteins under our native modification conditions an unlikely explanation for the observed changes in reactivity, except in the case of proteins such as L20 and L34.

Explaining the observed effects of pH changes on the SMTA reactivity of ribosomal proteins requires a consideration of which amino acid residues might be responsible for pH-linked conformational changes. Several amino acids have side chains with ionizable protons: aspartic acid, glutamic acid, histidine, cysteine, lysine, tyrosine, and arginine [64, 65]. However, only aspartic and glutamic acid, histidine and cysteine have pK_a values that are near the range over which our modification experiments were performed and so disruption of structural interactions involving these residues are probably responsible for the observed effects. Two possibilities for the type of structural feature that would give rise to the labeling behavior we observe are D, E, H, or C residues involved in networks of hydrogen bonds and salt bridges that constrain lysine-containing flexible loops to a particular conformation, or direct binding interactions between those residues and metal cations that play roles in rRNA structure. For exam-

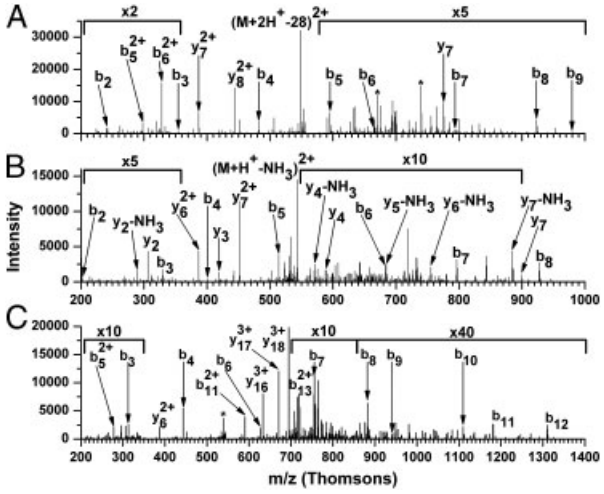

Figure 4. Spectra of amidinated peptides from ribosomal protein S20 reacted with SMTA at pH 6.6. (A) Residues 39–48 (K)K^{Am}AIQLAQEGK(A). (B) Residues 49–57 (K) AEEALK^{Am}IMR(K). (C) Residues 39–57 (K)KAIQLAQEGK^{Am}AEEALK^{Am}IMR(K). Peaks marked with an asterisk (*) are internal sequence ions consistent with the sequence and labeling pattern of each peptide. (A and B) have an identical *m/z* scale distinct from (C).

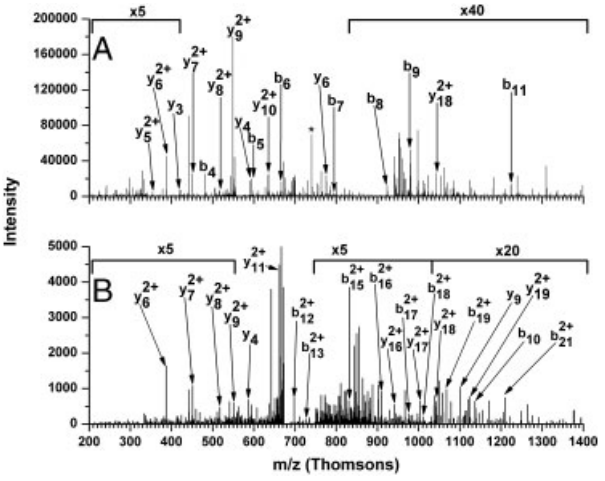

Figure 5. Spectra of amidinated peptides from ribosomal protein S20 reacted with SMTA at pH 7.5. (A) Residues 39–57, (K) K^{Am}AIQLAQEGK^{Am}AEEALK^{Am}IMR(A). (B) Residues 35–57, (K)TLSK^{Am}K^{Am}AIQLAQEGK^{Am}AEEALK^{Am}IMR(A). Peaks marked with an asterisk (*) are internal sequence ions consistent with the sequence and labeling pattern of each peptide.

ple, loss of a proton from a network of hydrogen bonds could generate a full negative charge on a carboxylate. The resulting charge–charge repulsions between adjacent carboxylates, or between the carboxylate and rRNA backbone phosphates, could push the affected portion of the protein into a conformation that exposes a previously inaccessible lysine. There are precedents for both of these roles in monomeric protein structures [66].

Analysis of the tryptic digests of protein samples stripped from trapping columns provides MS/MS spectra demonstrating the molecular details underlying the changes in the SMTA reactivity of ribosomal proteins. Figure 4A–C shows MS/MS spectra for peptides containing residues T35 to R57 from a digest of S20 modified at pH 6.6. Comparison of the *m/z* values of the b_8 and b_9 ions between Fig. 4A and C localizes the amidino group to lysine 39, as the mass-to-charge ratio of these fragments is 41 Th higher in Fig. 4A. The doubly charged b_5 ion in Fig. 4A and C shows an equally consistent 20 Th mass shift in the upper panel. These three spectra demonstrate differential modification at K39 and constant modification at K54. At pH 7.5 or 8.3, these smaller peptides are not present and have been replaced by both a triply amidinated peptide encompassing residues 49–57, with amidinations at K39, K48, and K54, and a quadruply amidinated peptide containing residues 35–57 in which all four lysine residues have been amidinated. Spectra for these more heavily modified peptides are shown in Fig. 5A and B. Again, observation of a 41 Th increase of the *m/z* values of b_8 and b_9 ions between Figs. 4C and 5A demonstrates the presence of the amidino group at K39. Together, the spectra of Figs. 4 and 5 demonstrate the pH-dependant differential labeling of residues K38, K39, and K48 and account for three of the five lysine residues that become susceptible to labeling when solution pH is increased from 6.6 to 8.3. The position of this peptide in the intact 30S subunit is shown in Fig. 6. The 16S rRNA is depicted as a blue solvent accessible or Connolly surface, ribosomal proteins are shown in light blue, and the bulk of S20 is depicted in white. Residues 39–57 are shown in yellow, with red lysines and green glutamic acid residues. Inspection of the crystal structure makes it clear that the reactivity of lysine 54 is due to its unhindered exposure to the solvent. The differential reactivity of residues 38, 39, and 48 is due to their proximity to anionic groups: lysine 38 forms a salt bridge with the phosphate of 16S rRNA residue C1440, lysine 39 forms a salt bridge with the phosphate of G1456, and lysine 48 is surrounded by glutamic acid residues 46 and 51, forming a salt bridge with E46's carboxylate. Lysine 39 and 48 are differentially modified at pH 6.6, suggesting that these interactions are more labile than the K38-C1440 interaction, whose strength is demonstrated by the observation of peptides starting with residue 39 even at pH 8.3 (note that amidination prevents tryptic cleavage at lysine residues).

Ribosomal protein L2 presents a case similar to S20 in which alterations of the pH of the solution appear to modify protein–rRNA contacts. Figure 7A and B compares MS/MS spectra of singly and doubly amidinated peptides containing residues 70–88 of protein L2. The doubly charged y_{15} ion has an *m/z* value consistent with amidination on K78, while the doubly charged y_{17}, y_{18}, b_{15}, and b_{16} ions increase by 20 Th from Fig. 7A and B, indicating amidination at K72 when the pH increases from 6.6 to 8.3. Similarly, Fig. 8A and B shows that K102 is protected at pH 6.6 but reactive at pH 8.3. The

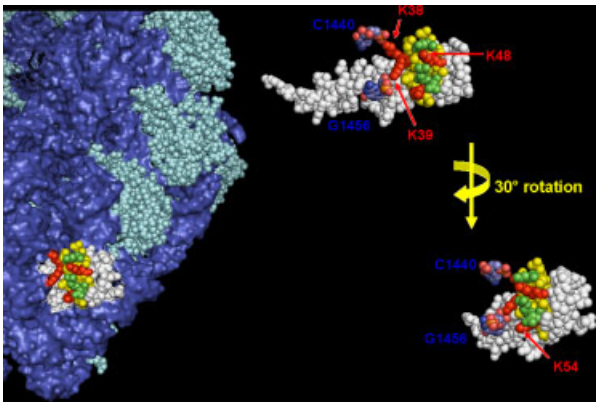

Figure 6. The differentially modified region of protein S20. 16S rRNA is shown in blue, other proteins in light blue. Protein S20 is shown in white, and the peptide containing residues 39–57 is shown in yellow. Lysine residues 54, 39, 38, and 48 are visible in red, and glutamate residues 50, 51, and 46 are shown in green. To the right, protein S20 is shown in isolation in the same orientation (top) and rotated 180 degrees (bottom). Also shown to the right are the interactions of lysine 38 with C1440 and lysine 39 with G1456. Nucleotides are shown with carbon light blue, nitrogen blue, phosphorous orange, and oxygen red.

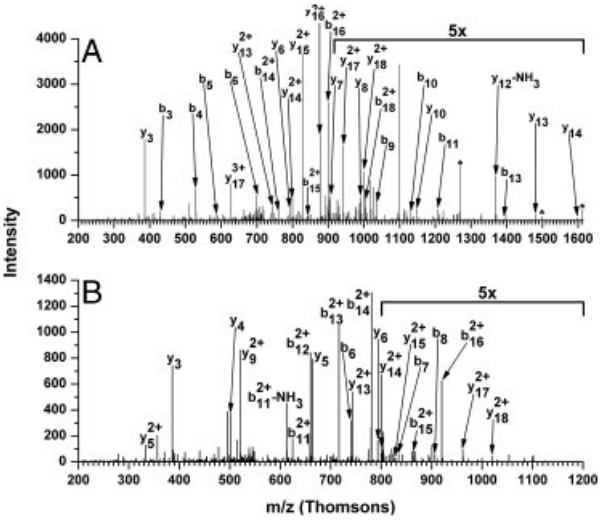

Figure 7. Singly and doubly amidinated versions of protein L2 peptide containing residues 70–88, (R)WDK^{72}VGIPAK^{78}VAAIEYDPNR(S). (A) Singly amidinated at position K78. (B) Doubly amidinated at positions K72 and K78.

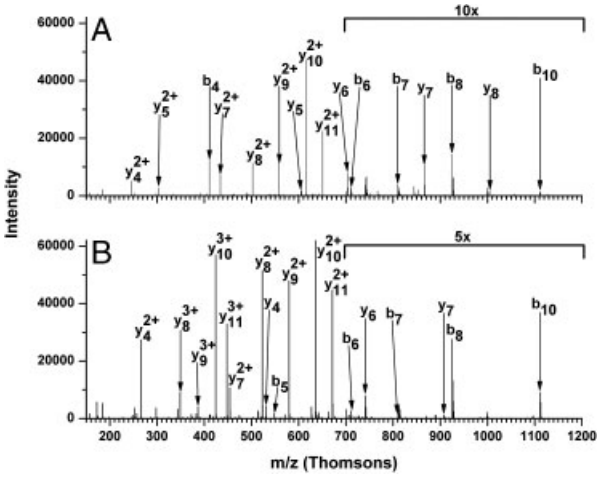

Figure 8. Differential reactivity of lysine 102 from protein L2 in a peptide containing residues 92–103, (R)IALLHYVDGEK^{102}R(Y). (A) Peptide from a tryptic digest of ribosomes modified at pH 6.6. (B) Peptide from a digest of ribosomes modified at pH 8.3.

41-Th shift of the y_6 and y_7 ions, as well as the 20-Th shift of the doubly charged y_4 ions, is diagnostic of the position of the modification. Ribosomal protein L2's interaction with the 23S rRNA is shown in Fig. 9 using the same color scheme as Fig. 6. These three lysine residues interact with a loop of rRNA in the region from G1491 to C1502 although only K102 is close enough to form direct hydrogen bonds with an rRNA residue, C1501. The ε-amino group of K72 is 8.0 Å away from the phosphate of G1491 and 6.4 Å away from the phosphate of G1492. However, the amino group of K72 does appear to interact with a network of carboxylates that include D71, D99, and E101. Arginine 103 also appears to be a participant in this cup-like network of interactions. Conversely, despite an apparent close proximity to the rRNA surface, K78's ε-amino is 4.4 and 5.1 Å from the backbone oxygens of C1502. Our data are not complete enough to propose a detailed mechanism for the changes in protein L2's reactivity at these positions. We hypothesize that despite the flexibility of the lysine side chain, residue K78 is always reactive across the pH range studied because it is not close enough to the rRNA surface. On the other hand, as the pH of the solution increases from 6.6 to 8.3, the network of interactions surrounding K72 is disrupted, increasing this residue's reactivity with SMTA. At higher pH, the buttress of rRNA comprising nucleotides 1490–1504 moves away from the surface of L2 due to charge–charge interactions with D71, D99, and E101, and K102 becomes fully reactive.

Peptides obtained from the digest of protein fractions containing L1 include peptides covering this protein's sequence from residues 10–47. Following modification at a pH of 6.6, L1 digests include a peptide containing residues 10–28, amidinated at position K14, and a peptide containing residues 32–47, amidinated at position K37. Spectra are shown in Fig. 10A and B. At pH 7.5 or 8.3, the SMTA reactivity of this region of the protein has changed, as demonstrated by the appearance of peptides containing residues 10–28, with amidino groups at positions K14 and K19, and residues 29–47, with amidinations at K31 and K37. Figure 11A and B shows MS/MS spectra demonstrating the increased reactivity of lysines 19 and 31. Comparison of Figs. 10A and 11A shows a 41-Th shift of the b_{10} ion that identifies the amidinated lysine as K19. The 381-Th increase

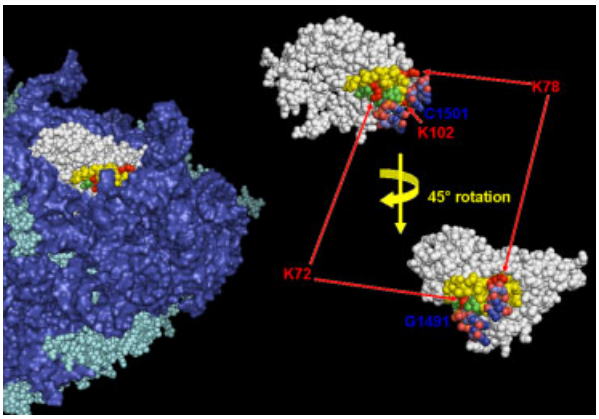

Figure 9. Ribosomal protein L2 *in situ*. The region containing the peptides discussed in the text is depicted in yellow, lysine residues are in red, and additional residues (D71, D99, E101, R103) are in green. The right hand portion of the figure shows L2 in isolation in the same orientation as the left hand side. The right hand, bottom portion of the figure shows L2 after a 45° rotation around a vertical axis in the plane of the paper. Nucleotide sequences containing G1491 and G1492 and C1501 are shown as in Fig. 6.

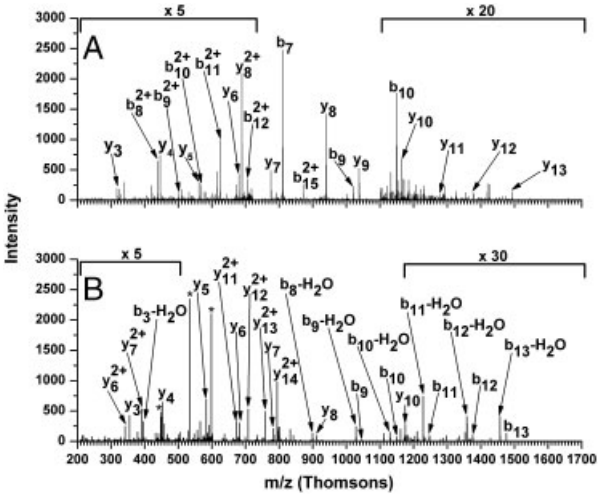

Figure 10. Spectra of peptides from tryptic digests of protein L1 labeled at pH 6.6. (A) Residues 10–28, (R)ALLEK^{Am}VDPNKVYTIDEAAR(L). (B) Residues 32–47, (K)ELATAK^{Am}FDETVEVHAK(L).

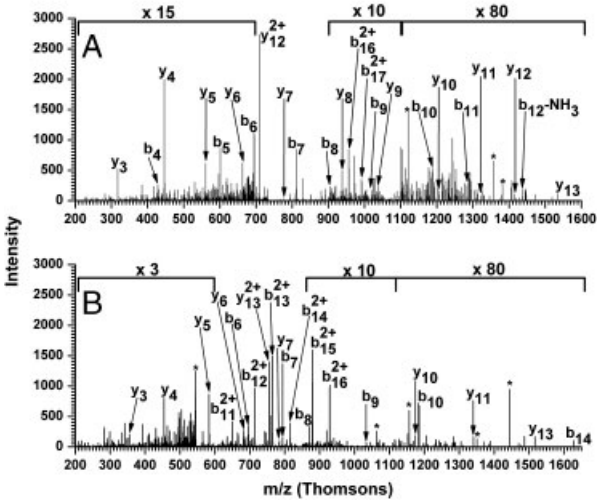

Figure 11. Spectra for tryptic peptides of protein L1 labeled at pH 7.5. (A) Residues 10–28, (R)ALLEK^{Am}VDPNK^{Am}VYTIDEAAR(L). (B) Residues 29–47, (R)LVK^{Am}ELATAK^{Am}FDETVEVHAK(L). Peaks marked with an asterisk (*) are internal sequence ions consistent with the sequence and labeling pattern of each peptide.

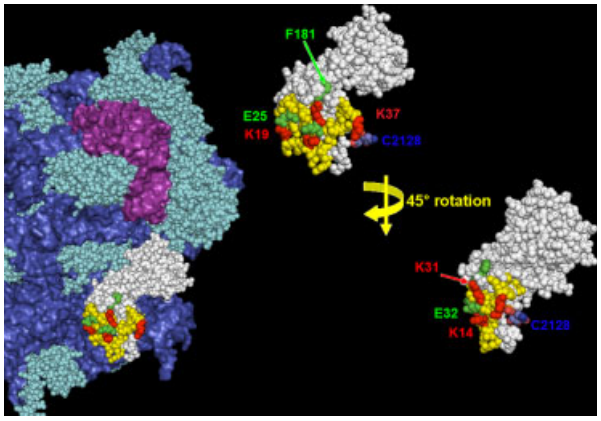

Figure 12. The differentially modified region of protein L1. The color scheme is identical to that used in Fig. 6, with the addition of the 5S rRNA, labeled in purple. The right-hand portion of the figure shows L1 with no rRNA except nucleotide C2128 present, in either the same orientation as the left-hand side of the figure (top) or rotated 45° around a vertical axis in the plane of the figure (bottom).

of the precursor ion m/z from Fig. 10B to 11B, combined with identical m/z values for y_3–y_7 and y_{10} for the spectra, are consistent with the addition of an LVK tripeptide with an amidinated lysine residue. These peptides report on the behavior of what appears to be an especially flexible portion of the protein. In the original crystal structure of the *T. thermophilus* HB8 70S ribosomal particle [32], the N-terminal region of L1 from residue P2 to K18 is too disordered to provide a structure, as is a loop encompassing residues L29 to A34. In a more recent structure of *T. thermophilus* HB8 70S ribosomes co-crystallized with the initiation factor EF-P [36], a substantial portion of the N-terminus and the entire loop from L29 to A34 is visible. Figure 12 shows the region of the 50S subunit containing L1, generated from PDB file 3HUX. In this structure, the amino group of K14 is too far from the oxygens of E32 for strong hydrogen bonding interactions, which correlates well with the observed reactivity of K14. The analogous distances

from the amino group of K19 to the oxygen atoms of E25 are 4.0 and 3.6 Å, close enough given the flexibility of the lysine residue's side chain to form an interaction that prevents SMTA modification of this residue at lower pH. The other two observed lysine residues in this region show more puzzling behavior. Lysine 31 appears to be freely accessible. The only adjacent residue that could interact is F181, whose backbone carbonyl oxygen is within 2.5 Å of K31's amino group in the structure reported by PDB file 3HUX. Protection of K31 at low pH indicates that the strength of this hydrogen bond is enhanced by the hydrophobic environment contributed by P182 and P183, which wrap part of the way around the lysine. Finally, K37 appears to be in very close proximity to nucleotide C2128 of the 23S rRNA. Closer inspection of the structures shown in the lower right-hand corner of Fig. 12 demonstrates that the lysine residue drapes its aliphatic chain over this cytosine residue's backbone phosphate rather than interacting directly. This puzzling lack of a salt bridge is supported by the observation that K37 is amidinated at all three pH values studied.

4.3 Conclusions

The results presented in this report demonstrate the utility of SMTA as a labeling reagent across a physiologically relevant pH range. The robust structure of the thermophilic *T. thermophilus* HB8 ribosome makes it an ideal experimental sample, since denaturation is less likely to be the cause for the observed changes in reactivity. The availability of a high-resolution crystal structure with observed side chain conformations allows us to interpret pH-induced changes in proteins' SMTA labeling patterns in terms of alterations in side chain hydrogen bonding or salt bridge interactions. It is significant that one of our examples is a surface loop of large subunit protein L1 that is disordered in the isolated ribosome structure. Because our data were collected using isolated ribosomes rather than actively translating complexes, we cannot propose a specific role for these regions of L1. However, the modulation of specific lysine residues' reactivity by such mild, non-denaturing alterations of solution conditions suggests that we are detecting locally flexible regions of proteins whose conformational changes could be of functional significance. This initial use of SMTA labeling at pH values other than the one resulting from the neutralization of Tris-free base by SMTA will ultimately be extended to study the effects of pH and ionic strength on macromolecular complexes that have not been crystallized. In addition to extending the scope of SMTA labeling as a structural probe, our labeling data can provide structural constraints for computational chemists who seek to model the effects of solution parameters such as pH and salt concentration on the structure of biological machines like the ribosome [67]. While pursuing an initial characterization of the ribosomal proteome of *T. thermophilus* HB8, we have also confirmed the identifications made by Suh *et al.* [48] and extended their observations to include ribosomal protein S2, and more accurate measurements of all protein masses. The integration of whole protein mass measurements of unmodified proteins and proteins from denatured, disassembled ribosomes has been shown to be a useful procedure to corroborate protein identifications, PTMs, and sequence errors.

Most of *T. thermophilus* HB8's ribosomal proteins show minimal variation in their SMTA reactivity as the pH is changed from 6.6 to 8.3. These proteins appear to be highly solvent exposed or deeply buried in the ribosome structure, and therefore their reactivity is not influenced by changes in solution composition. A smaller group of proteins have lysine residues in local environments that are flexible enough that their reactivity with SMTA is increased by decreases in hydrogen ion concentrations. Correlation of the labeling patterns of representative examples of such proteins with LC-MS/MS identification of labeling positions shows that these alterations occur in flexible portions of the proteins' sequence, regions whose conformational variability may be of some functional importance. A contributing factor to this pH-induced conformational flexibility appears to be lysine proximity to glutamic and aspartic acid residues. Increasing the pH of the solution ionizes the carboxylates, and charge–charge repulsions force the local protein structure to become more open, increasing the SMTA reactivity of differentially modified lysine residues as seen in the examples presented for proteins L1 and S20. Similar repulsive interactions between protein carboxylates and rRNA backbone phosphates also contribute to these structural transitions, as seen in the example of protein L2.

One observation that informs our pH dependence experiments is that an isolated ribosome is likely to provide a less than ideal context for interpretation of the results of our labeling experiments. Ribosomal proteins control the flexibility of the ribosomal RNA, preventing it from falling into inactive conformations during assembly or any of the segments of the translation cycle. Each of these processes involves numerous accessory proteins: initiation factors, elongation factors, termination factors, and the like. The observation of some amount of flexibility in a particular region of the ribosome, or in a particular domain of one of its constituent proteins, is not significant without any indication of what other cellular proteins or complexes the flexible regions normally interact with. Labeling patterns obtained with ribosomes interacting with cofactors or substrates may be more relevant to the functional importance of conformational changes in the ribosome.

W. E. R. would like to thank Ellen Quardokus of the Indiana University Biology Department for her advice and assistance culturing Thermus thermophilus HB8. This work was supported in part by the Indiana METACyte Initiative of Indiana University, funded in part through a major grant from the Lilly

Endowment, Inc., and by the National Science Foundation Grants CHE-1012855 and CHE-0832651

The authors have declared no conflict of interest.

5 References

[1] Nelson, D. L., Cox, M. M., *Lehninger Principles of Biochemistry*, 3rd. ed., Worth Publishers, New York 2000.

[2] Walsh, C. T., *Antibiotics: Actions, Origins, Resistance*, ASM Press, Washington, DC 2003.

[3] Livermore, D. M., Bacterial resistance: origins, epidemiology, and impact. *Clin. Infect. Dis.* 2003, *36*, S11–S23.

[4] Ehrenberg, M., Structure and function of the ribosome. Scientific background on the Nobel prize in chemistry 2009. Accessed February 15, 2009 at http://nobelprize.org/nobel_prizes/chemistry/laureates/2009/sci.html.

[5] Drenth, J., *Principles of Protein X-ray Crystallography*, Springer-Verlag, New York 1994.

[6] Reed, D. G., *Protein NMR Techniques, Methods in Molecular Biology v. 60*, Humana Press, Totowa, NJ 1997.

[7] Akashi, S., Sirouzu, M., Terada, T., Ito, Y., Yokoyama, S., Takio, K., Characterization of the structural difference between active and inactive forms of the Ras protein by chemical modification followed by mass spectrometric peptide mapping. *Anal. Biochem.* 1997, *248*, 15–25.

[8] Benjamin, D. R., Robinson, C. V., Hendrick, J. P., Hartl, F. U., Dobson, C. M., Mass spectrometry of ribosomes and ribosomal subunits. *Proc. Natl. Acad. Sci. USA* 1998, *95*, 7391–7395.

[9] Ilag, L. L., Videler, H., McKay, A. R., Sobott, F. *et al.*, Heptameric (L12)6/L10 rather than canonical pentameric complexes are found by tandem MS of intact ribosomes from thermophilic bacteria. *Proc. Natl. Acad. Sci. USA* 2005, *102*, 8192–8197.

[10] Englander, S. W., Mayne, L., Sosnick, T. R., Hydrogen exchange: the modern legacy of Linderstrøm-Lang. *Protein Sci.* 1997, *6*, 1101–1109.

[11] Ehring, H., Hydrogen exchange/electrospray ionization mass spectrometry studies of structural features of proteins and protein/protein interactions. *Anal. Biochem.* 1999, *267*, 252–259.

[12] Smith, D. L., Deng, Y., Zhang, Z., Probing the non-covalent structure of proteins by amide hydrogen exchange and mass spectrometry. *J. Mass Spectrom.* 1997, *32*, 135–146.

[13] Neurath, H., in Jaenicke, R. (Ed.), *Protein Folding*, Elsevier, New York 1980, pp. 501–523.

[14] Hubbard, S. J., The structural aspects of limited proteolysis of native proteins. *Biochim. Biophys. Acta* 1998, *1382*, 191–206.

[15] Suh, M.-J., Pourshahian, S., Limbach, P. A., Developing limited proteolysis and mass spectrometry for the characterization of ribosome topology. *J. Am. Soc. Mass Spectrom.* 2007, *18*, 1304–1317.

[16] Hamburg, D.-M., Suh, M.-J., Limbach, P. A., Limited proteolysis analysis of the ribosome is affected by subunit association. *Biopolymers* 2009, *91*, 410–422.

[17] Glocker, M. O., Borchers, C., Fiedler, W., Suckau, D., Przybylski, M., Molecular characterization of surface topology in protein tertiary structures by amino-acylation and mass spectrometric peptides mapping. *Bioconjugate Chem.* 1994, *5*, 583–590.

[18] Glocker, M. O., Nock, S., Sprinzl, M., Przybylski, M., Characterization of surface topology and binding area in complexes of the elongation factor proteins EF-Ts and EF-Tu-GDP from *Thermus thermophilus*: a study by protein chemical modification and mass spectrometry. *Chem. Eur. J.* 1998, *4*, 707–715.

[19] Fiedler, W., Borchers, C., Macht, M., Deininger, S.-O., Przybylski, M., Molecular characterization of a conformational epitope of hen egg white lysozyme by differential chemical modification of immune complexes and mass spectrometric peptide mapping. *Bioconjugate Chem.* 1998, *9*, 236–241.

[20] Mendoza, V. L., Vachet, R. M., Protein surface mapping using diethylpyrocarbonate with mass spectrometric detection. *Anal. Chem.* 2008, *80*, 2895–2904.

[21] Guan, J.-Q., Chance, M. R., Structural proteomics of macromolecular assemblies using oxidative footprinting and mass spectromtery. *Trends Biochem. Sci.* 2005, *30*, 583–592.

[22] Sharp, J. S., Becker, J. M., Hettich, R. L., Analysis of protein solvent accessible surfaces by photochemical oxidation and mass spectromtery. *Anal. Chem.* 2004, *76*, 672–683.

[23] Hambly, D. M., Gross, M. L., Laser flash photolysis of hydrogen peroxide to oxidize protein solvent-accessible residues on the microsecond timescale. *J. Am. Soc. Mass Spectrom.* 2005, *16*, 2057–2063.

[24] Tong, X., Wren, J. C., Konermann, L., Effects of protein concentration on the extent of γ-ray-mediated oxidative labeling studied by electrospray mass spectrometry. *Anal. Chem.* 2007, *79*, 6376–6382.

[25] Lundblad, R. L., *Chemical Reagents for Protein Modification*, 3rd Edn, CRC Press, Boca Raton, FL 2005.

[26] Janecki, D. J., Beardsley, R. L., Reilly, J. P., Probing protein tertiary structure with amidination. *Anal. Chem.* 2005, *77*, 7274–7281.

[27] Liu, X., Broshears, W. C., Reilly, J. P., Probing the structure and activity of trypsin with amidination. *Anal. Biochem.* 2007, *367*, 13–19.

[28] 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*, 2935–2946.

[29] Liu, X., Reilly, J. P., Correlating the chemical modification of *Escherichia coli* ribosomal proteins with crystal structure data. *J. Proteome Res.* 2009, *8*, 4466–4478.

[30] Running, W. E., Reilly, J. P., Ribosomal proteins of *Deinococcus radiodurans*: their solvent acessibility and reactivity. *J. Proteome Res.* 2009, *8*, 1228–1246.

[31] Lauber, M. A., Running, W. E., Reilly, J. P., *B. subtilis* ribosomal proteins: structural homology and post-translational modifications. *J. Proteome Res.* 2009, *8*, 4193–4206.

[32] Selmer, M., Dunham, C. M., Murphy IV, F. V., Weixlbaumer, A. *et al.*, Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* 2006, *313*, 1935–1942.

[33] Wimberly, B. T., Brodersen, D. E., Clemons, W. M. Jr., Morgan-Warren, R. J. *et al.*, Structure of the 30S ribosomal subunit. *Nature* 2000, *407*, 327–339.

[34] Ogle, J. M., Brodersen, D. E., Clemons, W. M. Jr., Tarry, M. J. *et al.*, Recognition of cognate Transfer RNA by the 30S ribosomal subunit. *Science* 2001, *292*, 897–902.

[35] Ogle, J. M., Murphy IV, F. V., Tarry, M. J., Ramakrishnan, V., Selection of tRNA by the ribosome requires a transition from an open to a closed form. *Cell* 2002, *111*, 721–732.

[36] Blaha, G., Stanley, R. E., Steitz, T. A., Formation of the first peptide bond: the structure of EF-P bound to the 70S ribosome. *Science* 2009, *325*, 966–970.

[37] Thumm, M., Hoenes, J., Pfleiderer, G., *S*-Methylthioacetimidate is a new reagent for the amidination of proteins at low pH. *Biochim. Biophys. Acta* 1987, *923*, 263–267.

[38] Beardsley, R. L., Reilly, J. P., Quantitation using enhanced signal tags: a technique for comparative proteomics. *J. Proteome Res.* 2003, *2*, 15–21.

[39] Reynolds, J. H., Acetimidation of bovine pancreatic ribonuclease A. *Biochemistry* 1968, *7*, 3131–3135.

[40] Fazili, K. M., Mir, M. M., Qasim, M. A., Changes in chemical stability upon chemical modification of lysine residues in bovine serum albumin by different reagents. *Biochem. Mol. Biol. Int.* 1993, *31*, 807–816.

[41] Spedding, G., in Rickwood, D., Hames, B. D. (Eds.). *Ribosomes and Protein Synthesis: A Practical Approach*, Oxford Press, New York 1990, pp. 1–29.

[42] Read, S. M., Northcote, D. H., Minimization of variation in response to different proteins of the coomassie blue G dye-binding assay for proteins. *Anal. Biochem.* 1981, *116*, 53–64.

[43] Hardy, S. J. S., Kurland, C. G., Voynow, P., More, G., Ribosomal proteins of *E. coli*. I purification of the 30S ribosomal proteins. *Biochemistry* 1969, *8*, 2897–2905.

[44] Barritault, D., Expert-Bezancon, A., Guerin, M. F., Hayes, D., The use of acetone precipitation in the isolation of ribosomal proteins. *Eur. J. Biochem.* 1976, *63*, 131–135.

[45] 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.

[46] Wang, L.-C., Okitsu, C. Y., Kochounian, H., Rodriguez, A. *et al.*, A simple and inexpensive on-column frit fabrication method for fused silica capillaries for increased capacity and versatility in LC-MS/MS applications. *Proteomics* 2008, *8*, 1758–1761.

[47] Mann, B., Madera, M., Sheng, Q., Tang, H. *et al.*, ProteinQuant Suite: a bundle of automated software tools for label-free quantitative proteomics. *Rapid Commun. Mass Spectrom.* 2008, *22*, 3823–3834.

[48] Suh, M.-J., Hamburg, D.-M., Gregory, S. T., Dahlberg, A. E., Limbach, P. A., Extending ribosomal protein identifications to unsequenced bacterial strains using matrix assisted laser desorption/ionization mass spectrometry. *Proteomics* 2005, *5*, 4818–4831.

[49] Subramanian, A.-R., Structure and functions of ribosomal protein S1. *Prog. Nucl. Acids Res. Mol. Biol.* 1983, *28*, 101–142.

[50] Ramakrishnan, V., White, S. W., Ribosomal protein structures: insights into the architecture, machinery and evolution of the ribosome. *Trends Biochem. Sci.* 1998, *23*, 208–212.

[51] McLafferty, F. W., High-resolution tandem FT mass spectrometry above 10 kDa. *Acc. Chem. Res.* 1994, *27*, 379–386.

[52] Horn, D. M., Zubarev, R. A., McLafferty, F. W., Automated reduction and interpretation of high resolution electrospray mass spectra of large molecules. *J. Am. Soc. Mass Spectrom.* 2000, *11*, 320–332.

[53] Zabrouskov, V., Senko, M. W., Du, Y., Leduc, R. D., Kelleher, N. L., New and automated MS^n approaches for top-down identification of modified proteins. *J. Am. Soc. Mass Spectrom.* 2005, *16*, 2027–2038.

[54] Diaconu, M., Kothe, U., Schlünzen, F. *et al.*, Structural basis for the function of the ribosomal L7/L12 stalk in factor binding and GTPase activation. *Cell* 2005, *121*, 991–1004.

[55] Creighton, T. E., *Proteins: Structures and Molecular Properties,* 2nd Edn, W. H. Freeman, New York 1993.

[56] Kowalak, J. A., Walsh, K. A., β-Methylthioaspartic acid: identification of a novel posttranslational modification in ribosomal protein S12 from *Escherichia coli*. *Protein Sci.* 1996, *5*, 1625–1632.

[57] Miller, J. H., GUG and UUG are initiation codons *in vivo*, *Cell* 1974, *1*, 73–76.

[58] Frottin, F., Martinez, A., Peynott, P., Mitra, S. *et al.*, The proteomics of N-terminal methionine cleavage. *Mol. Cell. Proteomics* 2006, *5*, 2336–2349.

[59] Muranova, T. A., Muranov, A. V., Markova, L. F., Ovchinnikov, Y. A., The primary structure of ribosomal protein L3 from *Escherichia coli* 70S ribosomes. *FEBS Lett.* 1978, *96*, 301–305.

[60] Diedrich, G., Burkhardt, N., Nierhaus, K. H., Large scale isolation of proteins of the large subunit from *Escherichia coli* ribosomes. *Protein Exp. Purif.* 1997, *10*, 42–50.

[61] Strader, M. B., VerBerkmoes, N. C., Tabb, D. L., Connelly, H. M. *et al.*, Characterization of the 70S ribosome from *Rhodopseudomonas palustris* using an integrated ''Top Down'' and ''Bottom Up'' mass spectrometric approach. *J. Proteome Res.* 2004, *3*, 965–978.

[62] Vanet, A., Plumbridge, J. A., Guerin, M.-F., Alix, J.-H., Ribosomal protein methylation in *Escherichia coli*: the gene prmA, encoding the ribosomal protein L11 methylase, is dispensable. *Mol. Microbiol.* 1994, *14*, 947–958.

[63] Cameron, D. M., Gregory, S. T., Thompson, J., Suh, M.-J. *et al.*, *Thermus thermophilus* L11 methyltransferase, prmA,

is dispensable for growth and preferentially modifies free ribosomal protein L11 prior to ribosome assembly. *J. Bacteriol.* 2004, *186*, 5819–5825.

[64] Forsyth, W. R., Antosiewicz, J. M., Robertson, A. D., Empirical relationships between protein structure and carboxyl pK_a values in proteins. *Proteins* 2002, *48*, 388–403.

[65] Grimsley, G. R., Scholtz, J. M., Pace, C. N., A summary of the measured pK values of the ionizable groups in folded proteins. *Protein Sci.* 2009, *18*, 247–251.

[66] Qin, B. Y., Bewley, M. C., Creamer, L. K., Baker, H. M. *et al.*, Structural basis of the tanford transition of bovine β-lactoglobulin. *Biochemistry* 1998, *37*, 14014–14023.

[67] Pankavic, S., Miao, Y., Ortoleva, J., Shreif, Z., Ortoleva, P., Stochastic dynamics of bionanosystems: multiscale analysis and specialized ensembles. *J. Chem. Phys.* 2008, *128*, 234908-1–234908-13.

