

REGULAR ARTICLE

Extending ribosomal protein identifications to unsequenced bacterial strains using matrix-assisted laser desorption/ionization mass spectrometry

Moo-Jin Suh¹, Daisy-Malloy Hamburg¹, Steven T. Gregory², Albert E. Dahlberg² and Patrick A. Limbach¹

¹ Rieveschl Laboratories for Mass Spectrometry, Department of Chemistry, University of Cincinnati, Cincinnati, OH, USA

² Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, RI, USA

A protocol has been developed that allows protein identifications using available DNA-based or protein sequences from a reference strain of a bacterial species to be extended to bacterial strains for which no prior DNA-based or protein sequence information exists. The protocol is predicated on careful isolation of a specific sub-cellular group of proteins. In this study, ribosomal proteins were chosen due to their high relative abundance and similarity in copy number per cell. After isolation of ribosomal proteins, MALDI-MS is used to acquire accurate protein molecular weights. An iterative comparison of reference protein molecular weights and identities is made to the resulting data, allowing for the straightforward identification of ribosomal proteins from any non-reference strains. This approach can reveal differences between proteins at the amino acid or post-translational level. The protocol was developed, validated and applied to ribosomal proteins from three strains of the extreme thermophile *Thermus thermophilus*. This approach revealed that nearly 60% of the ribosomal proteins from all three strains are identical. The extension of protein identification to additional bacterial strains can be useful in phylogenetic studies as well as in biomarker identification.

Received: December 31, 2004

Revised: March 3, 2005

Accepted: April 18, 2005

Keywords:

MALDI-TOF MS / Posttranslational modification / Ribosome / Subcellular proteomics / *Thermus thermophilus* / Top-down Proteomics

1 Introduction

MS is presently one of the most powerful techniques for investigating proteins and their post-translational modifications [1]. The development of proteomics has benefited from the application of “bottom-up” (proteolytic fragment identi-

fication) [2] and “top-down” (intact protein mass measurement) [3] strategies coupled to MS, allowing for large-scale analysis of proteins present in an organism, tissue, or cell under a given set of physiological conditions [1]. A key component for the successful application of these strategies in proteomics is the presence of DNA or protein databases, which allow experimental data to be characterized based upon the closeness of fit to protein sequences available in these databases. However, there are often investigations requiring knowledge of protein sequences from species or strains that are not publicly available.

The ribosome, found in all organisms, is the subcellular organelle that performs the activity of protein synthesis. In addition to recent crystal structures [4–6], mass spectromet-

Correspondence: Dr. Patrick A. Limbach, Rieveschl Laboratories for Mass Spectrometry, Department of Chemistry, PO Box 210172, University of Cincinnati, Cincinnati, OH 45221, USA

E-mail: pat.limbach@uc.edu

Fax: +1-513-556-9239

Abbreviation: SA, sinapinic acid

ric characterization of ribosomal proteins [7–9] has drawn increasing attention in the area of subcellular proteomics [10]. MALDI-TOF MS has been a popular choice for protein identification from protein mixtures with moderate complexity due to its rapid speed and readily interpretable data. MALDI-MS was used to observe the ribosomal proteins from *Escherichia coli* ribosomes [11] and to identify the mutations present in ribosomal proteins relating to antibiotic resistance [12]. Pineda *et al.* [13] used MALDI-MS for identification of intact microorganisms based on biomarker masses derived from ribosomal proteins.

In a previous work, we have optimized the MALDI-MS conditions for analyzing *E. coli* ribosomes [14]. From a single analysis, 51 out of 56 ribosomal proteins were observed at a mass accuracy of ± 1 Da, allowing for the identification of acetylation, methylation, and other modifications. Here, we describe the extension of that work into a strategy that allows protein identifications to be extended to mixtures of proteins for which no genomic information is available. In particular, bacterial ribosomal proteins from a reference strain serve as the primary database to which MALDI-MS experimental data from ribosomal proteins arising from unsequenced strains are compared. An iterative matching strategy allows one to readily identify the protein components and determine specific differences arising in ribosomal proteins from the various strains investigated.

This approach was developed and validated using ribosomes from the extreme thermophile *Thermus thermophilus*, chosen in part because of the available crystal structures of ribosomes from this organism [4–6] and the two available genome sequences ([15]; GenBank AE017221). Initial development utilized the two sequenced strains HB27 and HB8, isolated from terrestrial hot springs in Japan [16] and then applied to the as yet unsequenced *T. thermophilus* strain IB21, isolated from a geographically distant submarine hot spring in Iceland [17]. Nearly 60% of the ribosomal proteins were found to be identical among the three strains and over 94% of the expected ribosomal proteins were identified by this approach.

2 Materials and methods

2.1 Materials

TFA, buffer reagents and protein calibration kit were obtained from Sigma (St. Louis, MO, USA). Sinapinic acid (SA) was obtained from Fluka (Milwaukee, WI, USA). Sucrose (DNase and RNase-free grade) was purchased from Acros Organics (Fairlawn, NJ, USA). Acids and organic solvents were HPLC grade or better.

2.2 Ribosome preparations

Conditions for culturing *T. thermophilus* HB8 (ATCC27634) and IB21 (ATTC43815) and isolating ribosomes have been described elsewhere [18]. Aliquots of intact 70S ribosomes at

a concentration of 22 mg/mL in a buffer containing 10 mM HEPES-KOH at pH 7.6, 10 mM MgCl₂, 50 mM NH₄Cl, and 5 mM β -mercaptoethanol were stored at -80°C . Intact 70S ribosome solutions were diluted prior to MALDI-MS analysis.

The 30S and 50S ribosomal subunits were isolated through a 0–45% sucrose gradient. Sucrose gradients were formed by layering different concentrations of sucrose in dissociation buffer (10 mM Tris-HCl pH 7.6, 1 mM MgCl₂, 60 mM NH₄Cl, and 4 mM β -mercaptoethanol) followed by diffusion. Approximately 30 A₂₆₀ units of intact 70S ribosomes were applied on the top of the gradient and the gradients were centrifuged in an SW28 rotor in a Beckman ultracentrifuge at 4°C for 17 h at 19 000 rpm. Fractions of 1.1–1.2 mL were collected manually and the absorbance at 260 nm was measured to determine the location of the 50S and 30S subunits.

Ribosomal subunits were precipitated by addition of three volumes of cold ethanol to each fraction, overnight incubation at -20°C followed by centrifugation at 13 000 rpm for 10 min. The pellets were resuspended in 10 mM Tris-HCl pH 7.6, 10 mM MgCl₂, 50 mM NH₄Cl, 0.25 mM EDTA, 4 mM β -mercaptoethanol, and aliquots were further analyzed.

2.3 MALDI-MS analysis

All MALDI-TOF MS experiments were done on a Bruker Reflex IV reflectron MALDI-TOF mass spectrometer (Bruker Daltonics, Billerica, MA, USA) equipped with a nitrogen laser as previously described [14]. Protein mass spectra were obtained in the positive ion mode at an acceleration voltage of 20 kV, extraction plate voltage of 17.1 kV and lens voltage of 10.1 kV by accumulating 300 laser shots. All samples were analyzed under identical instrumental parameters.

For all protein analyses, saturated SA in 33% aqueous ACN/0.1% TFA was used as the matrix. One microliter of sample prepared by mixing 1 μL of acidified ribosomal solution (around 2 pmol ribosome) with 9 μL of matrix was loaded into a MALDI target and allowed to air dry. Calibration of protein mass spectra was done using *E. coli* MRE 600 ribosomal proteins as external calibrants and then recalibrated internally with well-resolved ribosomal proteins from *T. thermophilus*.

2.4 Data analysis

The amino acid sequences of *T. thermophilus* HB27 [15] and HB8 ribosomal proteins were obtained from GenBank (<http://www.ncbi.nlm.nih.gov/genomes>) with accession numbers AE017221 and AP008226 (chromosome), respectively. The theoretical molecular weights of these proteins were calculated using the SequenceEditor software provided by the MALDI manufacturer.

3 Results and discussion

3.1 General approach

Although approaches for protein identification based upon bottom-up, top-down or a combination of the two methods are effective, such approaches are predicated on the availability of database information for the proteins of interest. In circumstances where a small, rigorously isolated subset of proteins is to be investigated, such approaches can be time-consuming or costly when extended to bacterial strains for which no sequence information exists. For example, the analysis of ribosomal proteins by a bottom-up approach required multiple 1-D or 2-D LC-MS/MS analyses to achieve identification of up to 95% of the proteins present [9]. Even in those cases, additional top-down experiments were necessary to confirm possible post-translational modifications. In circumstances where a sub-population of proteins can be fractionated, we reasoned that a direct approach based upon accurate molecular weight measurements that could be compared to reference molecular weights might serve as a more efficient approach for protein identification requiring minimal sample and being extremely rapid when implemented using MALDI-TOF MS.

The approach developed to readily confirm assigned protein identities and to distinguish possible sequence variations is provided in the flowchart of Fig. 1. Starting with proteins assigned from annotated sequenced DNA, MS data arising from the analysis of intact proteins is sequentially evaluated to classify proteins. Proteins can be classified into four categories: (1) Proteins yielding identical molecular weights to DNA-derived sequences; (2) Proteins yielding molecular weights corresponding to N-terminal methionine loss; (3) Proteins yielding molecular weights consistent with conserved post-translational modifications; and (4) Proteins that cannot be assigned directly which may reflect differences at the primary sequence or post-translational level.

To demonstrate the applicability of this approach, ribosomal proteins from *T. thermophilus* were evaluated. *T. thermophilus* HB27 was chosen as the reference strain from which starting molecular weight and protein assignments were based. *T. thermophilus* HB8 was then used as the source of experimental data for developing the analytical approach. Although the genome of *T. thermophilus* HB8 is currently unpublished, it has been deposited in GenBank, thereby providing a reasonable model system to gauge the effectiveness of the approach. Theoretical molecular weights of ribosomal proteins from *T. thermophilus* HB 27 were first calculated based on DNA-derived sequences. The 54 annotated

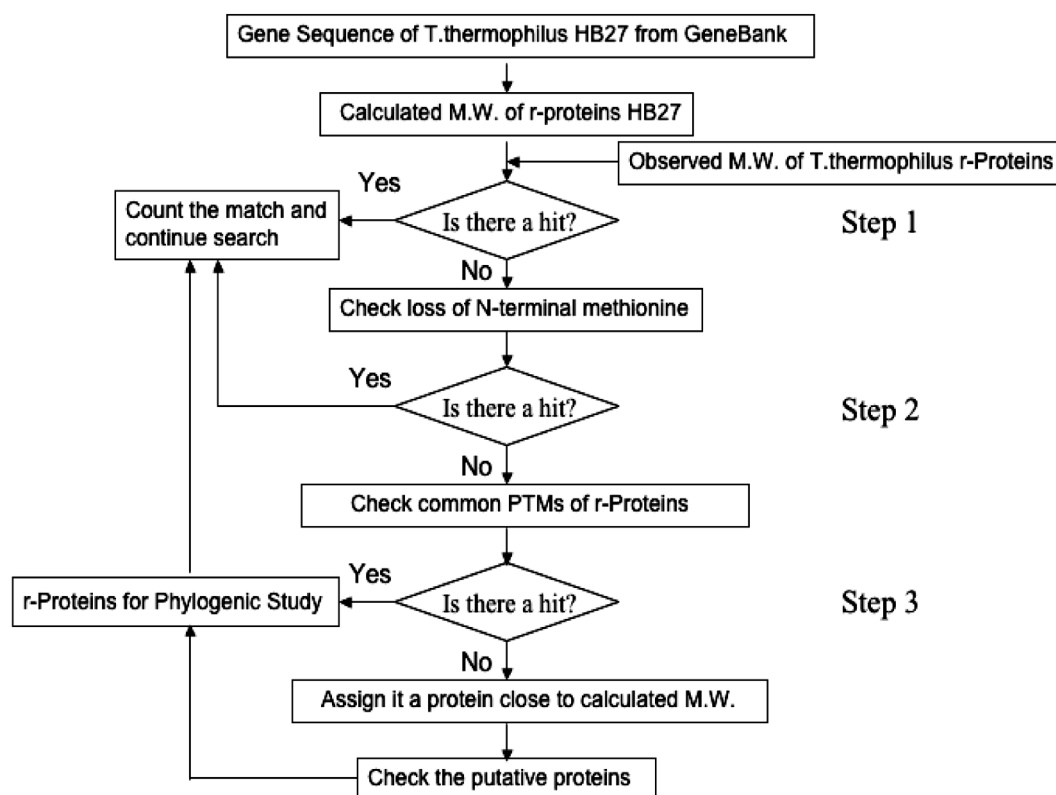


Figure 1. Flowchart outlining procedure used to assign MALDI-MS data to proteins from strains not present in any protein or gene database.

gene sequences of ribosomal proteins available for *T. thermophilus* HB27 yielded the theoretical molecular weights summarized in Table 1.

Experimental data were obtained from *T. thermophilus* HB8 using the previously described MALDI-MS approach [14]. Figure 2 contains representative mass spectra from intact 70S ribosomes (Fig. 2a) as well as each ribosomal subunit (Fig. 2b and c). A total of 54 peaks were tentatively detected from 70S ribosomes, with 34 arising from the large subunit and 20 arising from the small subunit as determined by comparing mass values obtained in Fig. 2a with those from Fig. 2b and c, respectively. Only mass values assigned to a particular subunit that also appear in the mass spectral data for intact 70S ribosomes were evaluated as described below.

3.2 Category 1: Protein molecular weight matches

The next step in the procedure is to assign proteins from the experimentally characterized strain whose measured molecular weights match those calculated from the reference strain. Table 2 lists 13 ribosomal proteins that can be assigned in this manner. Ten large subunit ribosomal proteins (L36, L34, L29, L23, L21, L24, L22, L17, L15 and L25) and three small subunit proteins (S16, S6 and S8) for the HB8 strain are assumed to have the same amino acid composition and primary sequence as those proteins in the HB27 strain. As only 24% of the proteins from the HB8 strain can be assigned on the basis of molecular weight matches to proteins from the HB27 strain, additional evaluation of the data is required.

3.3 Category 2: N-terminal methionine cleavage

N-terminal methionine cleavage is the most common post-translational modification found in prokaryotes [19, 20]. This modification, which is due to methionine aminopeptidase [21], can generally be predicted based upon the N-terminal sequence of the protein as the activity of methionine aminopeptidase is thought to be affected by the size of the amino acid in the penultimate N-terminal sequence location. N-terminal methionine cleavage occurs more than 80% of the time when Gly, Ala, Pro, Ser, and Thr are in the penultimate position whereas cleavage rarely occurs when Arg, Asn, Ile, Leu, Lys, Phe, Trp, and Tyr are in the penultimate position [22]. In rare cases, such as when Val is in the penultimate position, the N-terminal methionine can be retained or partially cleaved depending upon the size of the amino acid residue located in the antepenultimate position [19].

As N-terminal methionine cleavage is easily predicted for prokaryotic organisms, the second step in the protocol involves comparing unassigned molecular weight values to DNA-derived sequences calculated based upon N-terminal methionine loss. Calculated molecular weight values for the HB27 strain ribosomal proteins that are predicted to undergo loss of methionine are also included in Table 1.

Table 1. Calculated molecular weights of *T. thermophilus* HB27 ribosomal proteins

Protein	Cal. MW (Da)	A.A. in position 2	Met loss probable?	Cal. MW-Met ^{a)} (Da)
L1	24 825.8	Pro	Y	24 694.6
L1E	22 708.4	Tyr	N	
L2	30 468.6	Ala	Y	30 337.4
L3	22 408.3	Lys	N	
L5	21 043.8	Pro	Y	20 912.6
L6	19 532.0	Ser	Y	19 400.8
L7/L12	13 067.3	Ala	Y	12 936.1
L9	16 411.3	Lys	N	
L10	18 565.9	Pro	Y	18 434.7
L11	15 505.3	Lys	N	
L13	18 496.1	Val	Y or N	18 364.9
L14	13 288.7	Ile	N	
L15	16 281.2	Lys	N	
L16	15 947.0	Leu	N	
L17	13 715.1	Arg	N	
L18	12 611.9	Ala	Y	12 480.7
L19	17 180.9	Asn	N	
L20	13 743.3	Pro	Y	13 612.1
L21	11 047.2	Phe	N	
L22	12 780.1	Glu	N	
L23	10 736.9	Lys	N	
L24	12 056.6	Arg	N	
L25	23 204.8	Glu	N	
L27	9 480.1	Ala	Y	9 348.9
L28	11 006.2	Ser	Y	10 875.0
L29	8 650.3	Lys	N	
L30	6 785.2	Pro	Y	6 654.0
L31	11 253.0	Pro	Y	11 121.8
L32	6 717.1	Ala	Y	6 585.9
L33	6 615.9	Ala	Y	6 484.7
L35	7 127.7	Lys	N	
L36	4 421.4	Lys	N	
S1	59 985.0	Glu	N	
S2	29 276.9	Pro	Y	29 145.7
S3	26 701.2	Gly	Y	26 567.0
S4	24 324.5	Gly	Y	24 193.3
S5	17 557.6	Pro	Y	17 426.4
S6	11 972.9	Arg	N	
S7	18 016.1	Ala	Y	17 884.9
S8	15 837.6	Leu	N	
S9	14 401.7	Glu	N	
S10	11 930.0	Pro	Y	11 798.8
S11	13 712.9	Ala	Y	13 581.7
S12	14 599.4	Ala	Y	14 468.2
S13	14 304.9	Ala	Y	14 173.7
S14	7 139.7	Ala	Y	7 008.5
S15	10 554.4	Pro	Y	10 423.2
S16	10 387.0	Val	Y or N	10 255.8
S17	12 296.7	Pro	Y	12 165.5
S18	10 231.3	Ser	Y	10 100.1
S19	10 581.5	Pro	Y	10 450.3
S20	11 689.1	Ala	Y	11 557.9
Thx	3 337.0	Gly	Y	3 205.8

a) For loss of N-terminal methionine 131.2 Da were subtracted from the protein molecular weights.

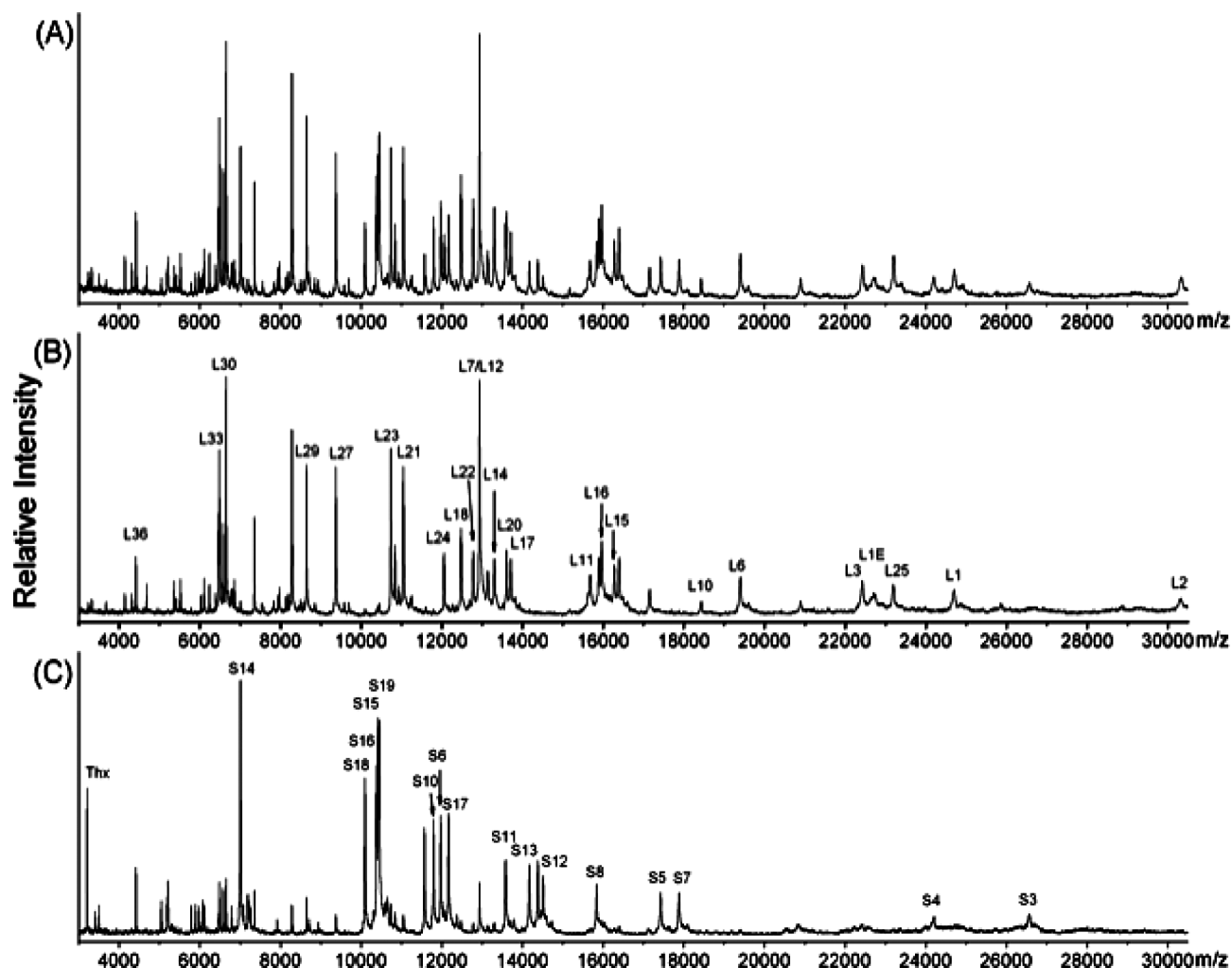


Figure 2. MALDI mass spectra of ribosome or ribosomal subunits from *T. thermophilus* HB8. (a) 70S ribosomal proteins (b) 50S (large subunit) ribosomal proteins, and (c) 30S (small subunit) ribosomal proteins. Data acquired in positive ion mode.

Table 2. The 13 ribosomal proteins from *T. thermophilus* HB8 assigned based upon direct correspondence between molecular weights predicted from the HB27 genome sequence and experimentally measured value

Gene HB27	HB27 [M+H] ⁺	HB8 [M+H] ⁺
L36	4 422	4 422
L34	6 109	6 110
L29	8 651	8 652
S16	10 388	10 387
L23	10 738	10 737
L21	11 048	11 048
S6	11 974	11 973
L24	12 058	12 057
L22	12 781	12 781
L17	13 716	13 717
S8	15 839	15 839
L15	16 282	16 284
L25	23 206	23 207

Table 3 contains experimental molecular weights for the HB8 strain ribosomal proteins which match those listed in Table 1 assuming methionine loss. Twenty-two additional proteins for the HB8 strain could be identified in this fashion. Of these, 8 proteins can be assigned to the large ribosomal subunit (L33, L30, L18, L7/L12, L20, L10, L6 and L2) and 13 assigned to the small ribosomal subunit (Thx, S14, S18, S15, S19, S10, S17, S11, S13, S5, S7, S4 and S3).

Eleven ribosomal proteins predicted to undergo N-terminal methionine cleavage based upon the HB27 DNA-derived sequence (L32, L31, L28, L27, L13, L5, L1, S20, S16, S12 and S2) were not observed as truncated proteins from the HB8 strain. Of these, ribosomal protein S16 was already assigned as a Category 1 protein with molecular weight matching the predicted one from the HB27 DNA-derived sequence. As S16 from the HB27 sequence has Val in the penultimate position and Lys in the antepenultimate position, N-terminal methionine cleavage would not be probable.

Table 3. The 22 ribosomal proteins from *T. thermophilus* HB8 assigned based upon direct correspondence between predicted molecular weight based upon HB27 sequence with loss of N-terminal methionine and experimentally measured value

Gene HB27	HB27 [M+H] ⁺	HB8 [M+H] ⁺	Δ mass
Thx	3 338	3 206	132
L33	6 617	6 486	131
L30	6 786	6 655	131
S14	7 140	7 010	130
S18	10 232	10 101	131
S15	10 555	10 424	131
S19	10 583	10 451	132
S10	11 931	11 799	132
S17	12 298	12 167	131
L18	12 613	12 481	132
L7/L12	13 068	12 937	131
S11	13 714	13 583	131
L20	13 744	13 612	132
S13	14 306	14 174	132
S5	17 559	17 428	131
S7	18 018	17 886	132
L10	18 568	18 436	132
L6	19 533	19 401	132
S4	24 326	24 192	134
S3	26 702	26 570	132
L2	30 470	30 337	133

After the second step of this procedure a total of 35 ribosomal proteins (65%) from the HB8 strain can be assigned based upon information from the HB27 strain genome.

3.4 Category 3: Common or conserved post-translational modifications

While the procedure outlined above identifies approximately 65% of the HB8 counterparts to HB27 ribosomal proteins, 19 ribosomal proteins arising from the annotation of the HB27 genome remain unassigned in the HB8 experimental data and are listed in Table 4. Thus, the next step undertaken was to consider commonly observed post-translational modifications, especially those known or suspected to be highly conserved within this family of proteins. Among ribosomal proteins, post-translational modification of L3, L11, S12 and L7/L12 are well-documented [9, 18, 23–27]. The commonly observed post-translation modifications of ribosomal proteins in prokaryotic organisms are methylation, acetylation, and β -methylthiolation.

Ribosomal protein L3 in *E. coli* is modified with an N⁵-methylglutamine at position 150 [23, 24] by the prnB methyltransferase [28]. This same protein from *Rhodospseudomonas palustris* was identified as having a single site of methylation at either K155 or K158 [9]. The mass predicted from the HB27 genome sequence was 22 408 Da, and assuming a single methylation the molecular ion should be detected at m/z 22 423. The observed m/z value closest to

this mass is 22 437, which is 14 Da higher than expected. Assuming the monomethylation is conserved, this mass difference could arise from differences in the amino acid sequence or additional methylations of L3. Although this modification must be validated by additional experiments, based upon the information derived from this study plus prior findings, the peak detected at 22 437 was assigned to L3 with an N⁵-methylglutamine acid post-translational modification.

Ribosomal protein L11 belongs to the class of conserved proteins from prokaryotic and eukaryotic organisms [29]. In *E. coli* ribosomal protein L11 is trimethylated at three amino acid positions by a single enzyme, L11 methyltransferase [25]. Evaluating the experimental data from Fig. 2a and b, the observed m/z value closest to the mass of L11 predicted from the HB27 genome sequence was 15 678. This value is 172 Da higher than the calculated value, and assuming methylations of ribosomal protein L11 from the HB8 strain, this mass difference corresponds to 12 methyl groups. Methylation of ribosomal protein L11 from *T. thermophilus* HB8 was reported by HPLC separation and sequencing although the extent of methylation of L11 was unknown [26]. Further, Dahlberg and co-workers [18] have recently shown that mutant HB8 strains lacking the *prmA* gene encoding the L11 methyltransferase results in a detected molecular weight for ribosomal protein L11 that agrees with that calculated from the DNA-derived sequence of HB27. In addition, ribosomal protein L11 from *R. palustris* was identified having multiple methylations although it was ambiguous whether the modification was 12 methyl groups or nine methyl groups and an acetyl group [9]. Based upon the information derived from this study plus prior findings, the peak detected at 15 678 was assigned to methylated L11.

Ribosomal protein S12, which plays a role in the maintenance of translational accuracy, is also a phylogenetically conserved ribosomal protein [30]. S12 from *E. coli* was identified as being post-translationally modified with a 3-methylthioaspartic acid at D88 by use of MALDI-TOF MS in combination with post-source decay sequencing [27]. S12 from *R. palustris* was also identified and the 3-methylthioaspartic acid at D88 localized with “top-down” and “bottom-up” approaches coupled to FT-ICR MS [9]. From the experimental data generated in this study, no ion corresponding to the unmodified ribosomal protein S12 at m/z 14 600 was detected. However, an ion at m/z 14 515 was found, and this ion would correspond to S12 with N-terminal methionine cleavage and a 3-methylthioaspartic acid. Although this modification must be validated by additional experiments, based upon the information derived from this study plus prior findings, the peak detected at 14 516 was assigned to S12 with a 3-methylthioaspartic acid post-translational modification.

It is known that ribosomal protein L7/L12 exists in two isoforms in *E. coli* (with L7 being an N-terminally acetylated form of L12) and other prokaryotic organisms [31, 32]. In addition to the ion detected at m/z 12 937, which was assigned to N-terminal methionine cleavage (Table 3), two

Table 4. Ribosomal proteins from *T. thermophilus* HB8 that cannot be assigned based upon direct correspondence between predicted molecular weight based upon HB27 sequence (with or without loss of N-terminal methionine) and experimentally measured value. Putative protein assignments are listed along with possible post-translational modifications (see Sections 3.4 and 3.5 for details)

Gene HB27	HB27 [M+H] ⁺	HB8 [M+H] ⁺	Δ mass	Modification/Notes
L32	6 718	6 575	−143	Met loss (6587): −12 Δ mass
L35	7 129	7 354	+225	
–	–	8 287		Unassigned/ L31 (HB8)
L27	9 481	9 378	−103	Met loss (9350): +28 Δ mass
L28	11 007	10 848	−159	Met loss (10 876): −28 Δ mass
L31	11 254	not detected		Met loss (11 123)
S20	11 690	11 572	−118	Met loss (11 559): +13 Δ mass
L7/L12	13 068	12 964	−104	Met loss (12 937): +27 Δ mass
L7/L12	13 068	12 978	−90	Met loss (12 937): +41 Δ mass
L14	13 290	13 304	+14	
S9	14 403	14 384	−19	
S12	14 600	14 516	−84	Met loss + β -methylthiolation
L11	15 506	15 678	+172	12 methylations
–	–	15 896		Unassigned/ L13 (HB8)
L16	15 948	15 962	+14	
L9	16 412	16 397	−15	
L19	17 182	17 152	−30	
L13	18 497	not detected		Met loss (18 366)
L5	21 046	20 900	−146	Met loss (20 914): −14 Δ mass
L3	22 409	22 437	+28	Methylation + 14 Da
L1e	22 709	22 724	+15	
L1	24 827	24 698	−129	Met loss (24 696)
S2	29 278	not detected		
S1	59 986	not detected		

additional peaks are detected with mass values 27.3 and 41.7 Da higher than the assigned mass. These ions likely correspond to dimethylated and trimethylated or acetylated modifications to L7/L12, respectively. Although insufficient mass accuracy is available to distinguish trimethylation (42.09 Da) from acetylation (42.04 Da), the ion detected at m/z 12 978 is consistent with the existence of the N-terminal acetylation of L12 to form L7, and this additional peak is assigned to the L7/L12 complex. Interestingly, the ion detected at m/z 12 964, which is consistent with dimethylation or formylation, was only detected when ribosome subunits were analyzed in diluted form due to overlaps caused by the more abundant ions representing L7 and L12. Thus, this modified protein is likely to be present in very low abundances within the ribosome, which might limit its identification via conventional protein analysis methods.

3.5 Category 4: Unmatched peaks

After completion of steps 1 through 3 of this procedure, 38 of the 54 ribosomal proteins from *T. thermophilus* HB8 can be assigned, and two additional m/z values can be assigned as post-translationally modified isoforms of L7/L12. Tentative assignments of the remaining ribosomal proteins were made as follows. First, any proteins expected to undergo N-termi-

nal methionine cleavage were matched to the nearest m/z values. As a result, ribosomal proteins L32, L27, L28, L5, L1 and S20 were tentatively identified and the resulting mass differences between the HB27 derived molecular weights and the experimentally measured molecular weights are listed in Table 4. The mass differences range from +28 to −28 Da and could be due to either post-translational modifications or differences between the HB27 and HB8 sequences.

The remaining m/z values were then tentatively assigned based upon molecular weight comparisons to HB27 ribosomal proteins. In this fashion, ribosomal proteins L35, L14, L16, L9, L19, L1e (equivalent to L4) and S9 were identified and the resulting mass differences between the HB27 derived molecular weights and the experimentally measured molecular weights are listed in Table 4. With the exception of L35 (Δ mass = 225 Da), all other experimentally measured molecular weights differed from that calculated from HB27 from +57 to −30 Da and could be due to post-translational modifications or sequence differences. Two anomalous m/z values from the large subunit were obtained from the mass spectral data, 8287 and 15 896, and could not be assigned to any ribosomal protein by the procedure described. Finally, no reasonable m/z values corresponding to ribosomal proteins L31, L13, S2 or S1 were detected in the data shown in Fig. 2.

By the described procedure, assignments of 50 out of 54 ribosomal proteins (93%) from *T. thermophilus* HB8 could be made using the gene sequences of *T. thermophilus* HB27 ribosomal proteins.

3.6 Validation of procedure with *T. thermophilus* HB8 genome annotations

To validate the effectiveness of this strategy for assigning proteins based upon direct analysis of protein molecular weight, theoretical molecular weights of the ribosomal proteins from the annotated *T. thermophilus* HB8 genome were calculated and are listed in Table 5. All of the ribosomal protein assignments in Tables 2 and 3 are verified to be correct, as no differences were found between the HB27 and HB8 sequences for these proteins.

Significantly, the tentative assignments of ribosomal proteins in Table 4 are validated to be correct and the original unassigned m/z values from the large subunit proteins can now be assigned. The experimental mass differences between MALDI-MS data and HB27 sequence predictions are found to match exactly to mass differences expected on the basis of sequence heterogeneity between the two strains for L32, L27, L28, L14, L16, L9, L19, L5, L1, S20 and S9. The deposited HB27 sequence for L35 has a truncated N terminus as compared to the deposited HB8 sequence, and the experimental data obtained here yield a molecular weight for L35 consistent with the HB8 sequence, suggesting misassignment of the initiator methionine in the HB27 annotation. The two m/z values originally unassigned, 8287 and 15 896, can now be assigned to L31 and L13, respectively. In both cases, the HB8 sequences have a truncated N terminus as compared to the deposited HB27 sequences, again suggesting misassignment of the initiator methionine in the HB27 annotation.

Only two ribosomal proteins, S12 and L4 (designated as L1e in the HB27 annotation), were found to yield experimental molecular weights which differed from the expected ones based upon the HB8 genome annotation. S12 was assigned as a post-translationally modified protein with HB27 predicted sequence in accordance with the HB8 experimental data. However, the predicted HB8 sequence has a valine residue inserted after the N-terminal methionine that is absent in the predicted HB27 sequence. Based upon the experimental data obtained in this study, this inserted valine is likely in error. Ribosomal protein L1e (HB27 annotation) or L4 (HB8 annotation) differ in deposited sequences by only an N-terminal truncation. The calculated molecular weight for L4 is 23 235 Da but no ion was detected in that m/z range. The calculated molecular weight for L1e is 22 709 Da and an ion at m/z 22 724 ($\Delta m = 15$) is tentatively assigned as L1e. If correct, this result suggests the sequence of L1e is incorrect or this protein is post-translationally modified. Further studies will be required to verify this assignment.

Based upon the strategy outlined above, protein identifications could be extended from the *T. thermophilus* HB27 annotated sequences to 50 out of 54 ribosomal proteins for HB8. These assignments were validated and extended (in the case of L31 and L13) by comparison to *T. thermophilus* HB8 annotated sequences, resulting in the identification of 52 out of 54 ribosomal proteins. The only two unassigned proteins, S2 and S1, could not be completed due to an absence of experimental data.

3.7 Application of procedure to *T. thermophilus* IB21 strain

While the results obtained above were encouraging, the effectiveness of this strategy is best demonstrated by its extension to ribosomal proteins isolated from a strain of *T. thermophilus* for which no genome (or protein) sequence exists. *T. thermophilus* IB21 was originally isolated from a submarine hot spring in Iceland [17] and is thus geographically and ecologically isolated from strains HB27 and HB8 found in terrestrial hot springs in Japan [16]. Assessing the extent to which such geographic isolation is reflected in ribosomal protein variation is an ideal application of this approach. Thus, ribosomal proteins from the *T. thermophilus* IB21 strain were isolated and analyzed by MALDI-MS. Figure 3 shows a representative MALDI mass spectrum of ribosomal proteins from *T. thermophilus* IB21. Following the procedure described above, the observed mass spectral peaks were systematically identified with assignments listed in Table 6.

A total of 57 distinct m/z values were identified from the MALDI-MS data. Of these, 14 m/z values correspond directly to the DNA-derived HB27 proteins (category 1), 19 can be identified as arising solely from N-terminal methionine cleavage (category 2), six are due to the conserved post-translational modifications to S12, L3, L11 and L7/L12 (category 3), 12 were identified based upon closeness of the experimental data to that calculated from HB27 sequences (category 4), two m/z values (8286 and 15 897) were assigned based upon HB8-derived assignments, and four m/z values (4376, 4957, 25 696 and 25 912) remained unassigned. From these assignments, 51 out of 54 ribosomal proteins (94%) could be identified with only ribosomal proteins S3, S2 and S1 not being detected and assigned from the MALDI-MS data. Thus, application of this procedure to multiple strains of *T. thermophilus* yields approximately the same number of assigned proteins, and this procedure should prove generally applicable to other strains of *T. thermophilus* or other organisms for which proteins from a particular strain have been sequenced.

3.8 Comparison of *T. thermophilus* strains

This approach also allows the examination of the level of conservation among different strains of the same organism. Figure 4 illustrates the comparison of calculated protein

Table 5. Calculated molecular weights of *T. thermophilus* HB8 ribosomal proteins based upon translation from genome annotation. Differences between HB8 and HB27 translated sequences are noted where appropriate

Protein	Cal. MW (Da)	A.A. in position 2	Met loss possible	Cal. MW- Met ^{a)} (Da)	Δ HB27 to HB8
L1	24 830.9	Pro	Y	24 699.7	I→V(20), H→R(28)
L2	30 468.6	Ala	Y	30 337.4	
L3	22 408.3	Lys	N		
L4	23 235.1	Lys	N		ΔC-terminus (HB27)
L5	21 029.8	Pro	Y	20 898.6	L→V(5)
L6	19 532.0	Ser	Y	19 400.8	
L7/L12	13 067.3	Ala	Y	12 936.1	
L9	16 397.3	Lys	N		I→V(107)
L10	18 565.9	Pro	Y	18 434.7	
L11	15 505.3	Lys	N		
L13	15 895.0	Lys	N		ΔC-terminus (HB8)
L14	13 302.7	Ile	N		V→I(69)
L15	16 281.2	Lys	N		
L16	15 963.0	Leu	N		F→Y(32)
L17	13 715.1	Arg	N		
L18	12 611.9	Ala	Y	12 480.7	
L19	17 151.9	Asn	N		K→Q(123); V→A(135); E→K(142);
L20	13 743.3	Pro	Y	13 612.1	
L21	11 047.2	Phe	N		
L22	12 780.1	Glu	N		
L23	10 736.9	Lys	N		
L24	12 056.6	Arg	N		
L25	23 204.8	Glu	N		
L27	9 508.1	Ala	Y	9 376.9	K→R(11)
L28	10 978.2	Ser	Y	10 847.0	R→K(81)
L29	8 650.3	Lys	N		
L30	6 785.2	Pro	Y	6 654.0	
L31	8 285.6	Lys	N		ΔC-terminus (HB8)
L32	6 705.1	Ala	Y	6 573.9	I→T(32)
L33	6 615.9	Ala	Y	6 484.7	
L34	6 109.4	Lys	N		
L35	7 484.2	Pro	Y	7 353.0	ΔC-terminus (HB27)
L36	4 421.4	Lys	N		
S1	59 971.0	Glu	N		V→I(140)
S2	29 276.9	Pro	Y	29 145.7	
S3	26 701.2	Gly	Y	26 570.0	
S4	24 324.5	Gly	Y	24 193.3	
S5	17 557.6	Pro	Y	17 426.4	
S6	11 972.9	Arg	N		
S7	18 016.1	Ala	Y	17 884.9	
S8	15 837.6	Leu	N		
S9	14 382.7	Glu	N		R→H(57)
S10	11 930.0	Pro	Y	11 798.8	
S11	13 712.9	Ala	Y	13 581.7	
S12	14 882.8	Val	Y or N	14 751.6	inserted V(2) (HB8)
S13	14 304.9	Ala	Y	14 173.7	
S14	7 139.7	Ala	Y	7 008.5	
S15	10 554.4	Pro	Y	10 423.2	
S16	10 387.0	Val	Y or N	10 255.8	
S17	12 297.7	Pro	Y	12 166.5	
S18	10 231.3	Ser	Y	10 100.1	
S19	10 581.5	Pro	Y	10 450.3	
S20	11 703.1	Ala	Y	11 571.9	V→I(40)
Thx	3 337.0	Gly	Y	3 205.8	

a) For loss of N-terminal methionine 131.2 Da were subtracted from the protein molecular weights.

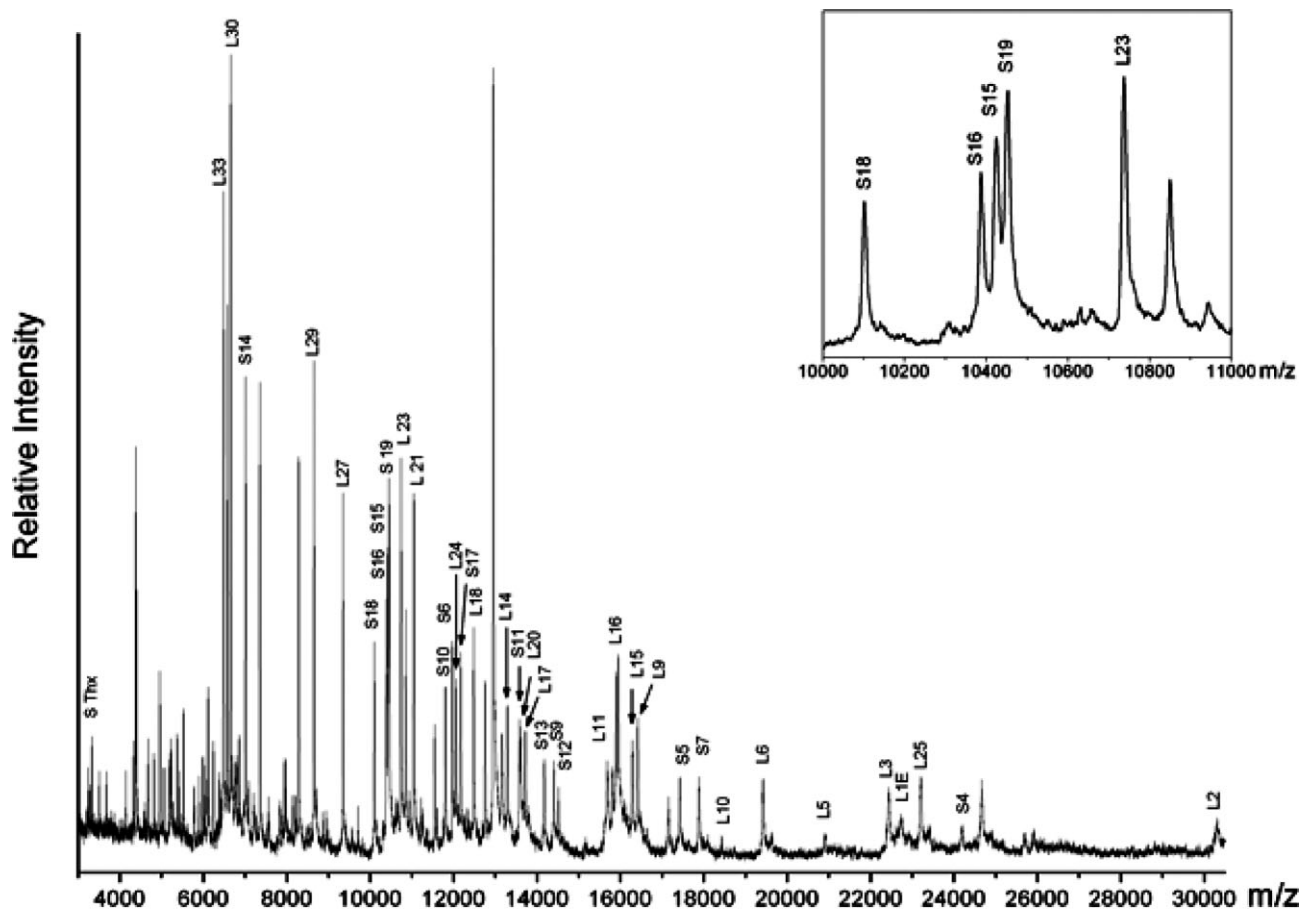


Figure 3. MALDI mass spectrum of ribosomal proteins from intact 70S ribosomes of *T. thermophilus* IB21.

Table 6. Assignment of ribosomal proteins from *T. thermophilus* IB21 based upon MALDI-MS data obtained in Fig. 3 and HB27 translated sequences^{a)}

Exp. <i>m/z</i>	Assignment	Δ mass ^{b)}	Category	Modifications/Notes
3 206	Thx ^{a)}	−132	2	Met loss
4 376	Not assigned			
4 395	L36	−27	4	
4 957	Not assigned			
6 111	L34 ^{a)}	2	1	
6 485	L33 ^{a)}	−132	2	Met loss
6 575	L32 ^{a)}	−143	4	Met loss (6587): −12 Δ mass
6 655	L30 ^{a)}	−131	2	Met loss
7 010	S14 ^{a)}	−130	2	Met loss
7 355	L35 ^{a)}	226	4	
8 286	L31 ^{a)}		4	From HB8 assignment
8 651	L29 ^{a)}	0	1	
9 350	L27	−131	2	Met loss
10 101	S18 ^{a)}	−131	2	Met loss
10 387	S16 ^{a)}	−1	1	
10 424	S15 ^{a)}	−131	2	Met loss
10 452	S19 ^{a)}	−131	2	Met loss
10 737	L23 ^{a)}	−1	1	
10 849	L28 ^{a)}	−158	4	Met loss (10876): −27 Δ mass
11 048	L21 ^{a)}	0	1	

Table 6. Continued

Exp. <i>m/z</i>	Assignment	Δ mass ^{b)}	Category	Modifications/Notes
11 545	S20	−145	4	Met loss (11559): −14 Δ mass
11 800	S10 ^{a)}	−131	2	Met loss
11 974	S6 ^{a)}	0	1	
12 059	L24 ^{a)}	1	1	
12 166	S17 ^{a)}	−132	2	Met loss
12 481	L18 ^{a)}	−132	2	Met loss
12 753	L22	−28	4	
12 952	L7/L12	−116	3	Met loss (12937): +15 Δ mass
12 979	L7/L12 ^{a)}	−89	3	Met loss (12937): +42 Δ mass
13 157	L7/L12	89	3	
13 290	L14	0	1	
13 583	S11 ^{a)*}	−131	2	Met loss
13 613	L20 ^{a)}	−131	2	Met loss
13 718	L17 ^{a)}	2	1	
14 175	S13 ^{a)}	−131	2	Met loss
14 403	S9	0	1	
14 516	S12 ^{a)}	−84	3	Met loss (14469) + β -methylthiolation
15 677	L11 ^{a)}	171	3	Methylation
15 811	S8	−28	4	
15 897	L13 ^{a)}		4	From HB8 assignment
15 948	L16	0	1	
16 283	L15 ^{a)}	1	1	
16 413	L9	1	1	
17 153	L19 ^{a)}	−29	4	
17 427	S5 ^{a)}	−132	2	Met loss
17 885	S7 ^{a)}	−133	2	Met loss
18 435	L10 ^{a)}	−133	2	Met loss
19 415	L6	−119	4	Met loss (19401): +14 Δ mass
20 915	L5	−131	2	Met loss
22 437	L3 ^{a)}	28	3	Methylation + 14 Da
22 722	L1e ^{a)}	13	4	
23 208	L25 ^{a)}	2	1	
24 199	S4	−127	4	Met loss (24194): +5 Δ mass
24 666	L1	−162	4	Met loss (24696): −30 Δ mass
25 696	Not assigned			
25 912	Not assigned			
	S3			Not detected
	S2 ^{a)}			Not detected
30 336	L2 ^{a)}	−134	2	Met loss
	S1 ^{a)}			Not detected

a) Denote protein assignments for the IB21 strain which are identical to experimental *m/z* values and assignments for the HB8 strain. A total of 57 distinct *m/z* values were identified from the MALDI-MS data.

b) Δ mass = HB27 calculated mass – experimental mass.

molecular ions from HB27 gene information (Fig. 4a) to the experimental data obtained from HB8 (Fig. 4b) and IB21 (Fig. 4c) strains. As is readily observed in this figure, L22 obtained from HB8 agrees with the value calculated for HB27 yet the same protein in the IB21 strain is found to have a 28-Da mass decrease. In addition, L14 obtained from IB21 agrees with the value calculated for HB27 yet is observed at a mass 14 Da higher in the HB8 strain. Moreover, ribosomal protein S20 is detected at different *m/z* values for the HB8 and IB21 strains and neither value matches that predicted from the HB27 strain.

As noted in Table 6, 33 ribosomal proteins were found to yield identical molecular weights from the HB8, HB27 and IB21 strains. Fourteen ribosomal proteins (L36, L27, L22, L16, L14, L7/L12, L9, L6, L5, L1, S20, S9, S8 and S4) were found to yield different molecular weights between the HB8 and IB21 strains. Ribosomal proteins L36, L22, L6, L7/L12, L1 and S8 were found to have identical molecular weights for the HB8 and HB27 strains yet were detected at different *m/z* values in the IB21 strain. Conversely, ribosomal proteins L27, L16, L14, L9, L5 and S9 were found to have the same molecular weights for the HB27 and IB21 strains yet were

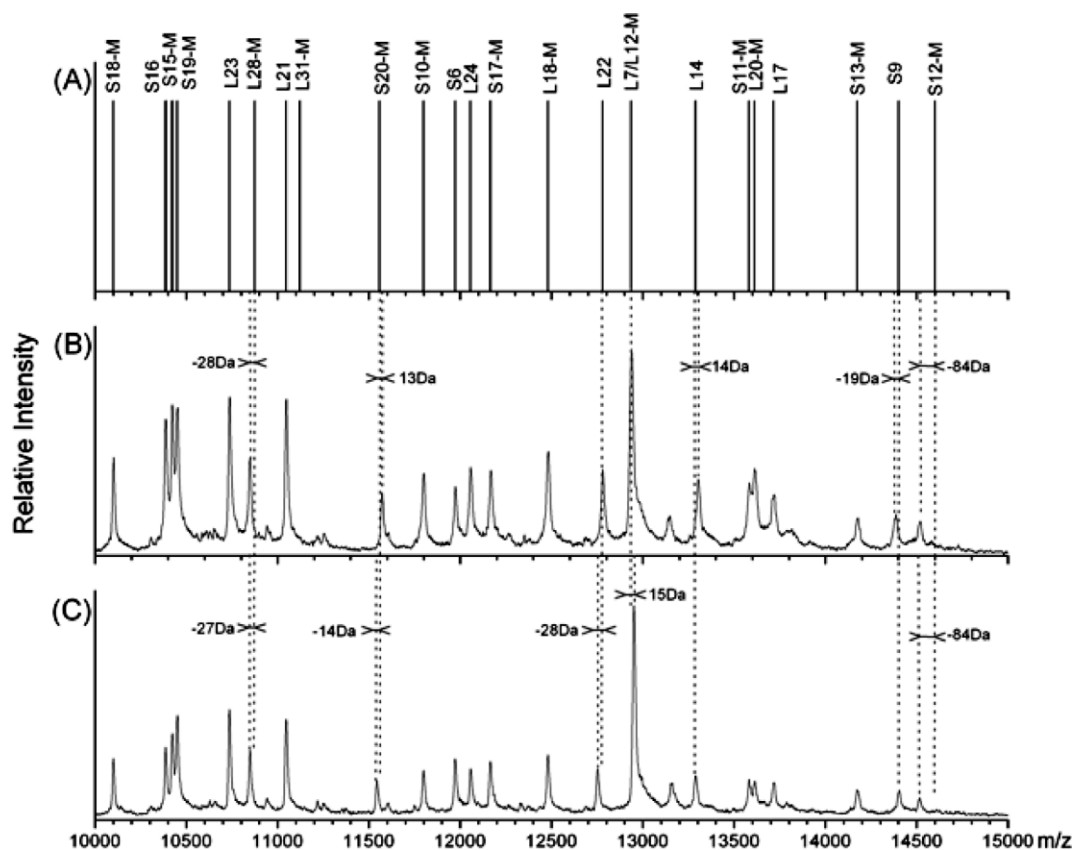


Figure 4. Comparison of MALDI mass spectral data to that predicted from DNA-derived sequences. The plots represent an expansion of the mass range between m/z values of 10 000 and 15 000. (a) Theoretical peaks predicted for HB27 strain. (b) Experimental data for HB8 strain. (c) Experimental data for IB21 strain. "M" represents methionine.

detected at different m/z values in the HB8 strain. S20 was the only ribosomal protein found to yield a different molecular weight for all three strains. Thus, S20 would be an example of a protein that could serve as a biomarker to distinguish among these strains of *T. thermophilus*.

Two ribosomal proteins (S2 and S1) were not detected in either strain. It has previously been reported that detection of all ribosomal proteins by MALDI-MS is difficult due to a combination of ion suppression and sample losses during the preparation step. However, during similar analyses of *E. coli* ribosomal proteins, only S1 was unable to be detected by MALDI-MS [14]. Thus, further investigations are necessary to characterize the two proteins not detected in this work.

3.9 General applicability

Although the approach developed here focused on ribosomal proteins from *T. thermophilus*, it should be generally applicable to ribosomal proteins from any organism for which a sequenced, reference strain exists. For example, there are known to be 17 different strains of *Helicobacter pylori*. Genome information is available for *H. pylori* 26 695 ([33], GenBank

accession number AE000511) and J99 ([34], GenBank accession number AE001439). To examine whether the general approach outlined here can be extended to other organisms, the expected molecular weights for ribosomal proteins of *H. pylori* 26 695 and J99 are presented in Table 7. As with *T. thermophilus*, the significant conservation among ribosomal proteins within a specific organism allows assigning the ribosomal proteins of an unsequenced strain from values calculated from a reference, sequenced strain. In this fashion, the approach outlined here should prove particularly useful in microbiology and other areas where specific subsets of proteins can be rigorously isolated and analyzed by MALDI-TOF MS.

4 Concluding remarks

A straightforward procedure for assigning prokaryotic proteins based upon MALDI-MS data and DNA-derived protein sequences has been presented. This procedure allows currently available protein database information specific for a particular strain of an organism to be extended to strains for which no database information exists. As illustrated with

Table 7. Calculated ribosomal protein molecular weights from two strains of *H. pylori* (26 695 and J99) demonstrating general similarity among ribosomal proteins from within a particular species^{a)}

Protein	<i>H. pylori</i> 26 695	<i>H. pylori</i> J99	Δ mass
L1	25 266.7	25 360.9	−94.2
L10	18 605.8	18 572.8	33
L11	15 329.2	15 329.2	0
L13	16 112.1	16 142.1	−30
L14	13 279.8	13 285.9	−6.1
L15*	14 905.5	14 634.1	271.4
L16	16 046.1	16 062.1	−16
L17	13 364.5	13 392.5	−28
L18	13 594.9	13 548.9	46
L19	13 600	13 616	−16
L2	30 271.4	30 271.4	0
L20	14 031.7	14 058.8	−27.1
L21	11 808.1	11 837.2	−29.1
L22	13 076.5	13 048.5	28
L23	10 454.4	10 454.4	0
L24	7 905.6	7 914.6	−9
L25	19 879.6	19 911.7	−32.1
L27	9 778.2	9 676.1	102.1
L28	6 929.3	6 929.3	0
L29	7 683	7 752.1	−69.1
L3	21 191.8	21 217.8	−26
L31	7 652	7 652	0
L32	5 660.7	5 646.7	14
L33	6 066.2	6 066.2	0
L34	5 246.4	5 246.4	0
L35	7 259.8	7 259.8	0
L36	4 307.4	4 321.5	−14.1
L4	24 023.7	23 989.7	34
L5	20 228.2	20 264.2	−36
L6	19 486.6	19 565.7	−79.1
L7/L12	13 313.5	13 314.5	−1
L9*	16 591.6	16 446.3	145.3
S1	62 827.1	62 528.8	298.3
S10	11 917.9	11 917.9	0
S11	14 029.5	14 015.5	14
S12	15 106	15 106	0
S13	13 575	13 480.9	94.1
S14	7 042.6	7 056.6	−14
S15	10 390.4	10 391.4	−1
S16	8 971.6	8 985.6	−14
S17	9 938.9	9 938.9	0
S18	10 448.4	10 448.4	0
S19	10 687.6	10 673.6	14
S2	30 688	30 649.9	38.1
S20	10 196.2	10 196.2	0
S21	8 613.2	8 614.2	−1
S3	26 380.2	26 496.4	−116.2
S4	23 964.8	24 004.9	−40.1
S5	16 358.4	16 358.4	0
S6	16 971.5	17 042.6	−71.1
S7	17 971.1	17 971.1	0
S8	15 183.7	15 183.7	0
S9	14 508.2	14 508.2	0

a) As with ribosomal proteins from *T. thermophilus*, calculated differences between strains are typically less than 40 Da.

* Represents difference in annotated initiator methionine codon from the DNA sequence.

T. thermophilus, more than 90% of the analyzed proteins can be assigned by this procedure. And while around 60% of the ribosomal proteins for the HB8 and IB21 strains were found to yield molecular weights consistent with the published protein sequences for the HB27 strain, 20% of the investigated proteins yielded molecular weights which differed among the three strains.

Although the MALDI-MS approach provides a rapid and sensitive route for extending protein assignments, the approach as described cannot determine the exact chemical constituents that lead to mass differences between measured and database values. For these discrepancies, either a negative or a positive mass shift is observed. As illustrated with the HB8 experimental data, negative mass shifts are most likely due to differences in the primary amino acid sequence of the proteins. This may be caused by a polymorphism leading to the substitution of a single amino acid in the corresponding protein sequence, or truncation of the N terminus, possibly as a result of misassignment of the initiating methionine. Positive mass shifts are either due to differences in the primary amino acid sequence of the proteins or due to differential post-translational modifications. While the data generated by MALDI-MS can neither distinguish between these two possibilities nor provide information on the specific site of amino acid substitution or post-translational modification, these data do allow one to focus further efforts on only those proteins suspected of differing amongst the various strains. Further, this approach allows to readily identify proteins that could serve as potential biomarkers for establishing strain identity from unknown samples.

Financial support of this work was provided by the National Institutes of Health (GM58843 to PAL and GM19756 to AED) and the University of Cincinnati. M.-J. Suh would like to acknowledge support from the URC Graduate Student Summer Fellowship Program at the University of Cincinnati. We would like to acknowledge Jill Thompson for helpful suggestions for ribosomal subunit purification and Jennifer Carr for preparation of the IB21 strain.

5 References

- [1] Aebersold, R. and Mann, M., *Nature* 2003, 422, 198–207.
- [2] Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., Garvik, B. M., Yates, J. R., III, *Nat. Biotechnol.* 1999, 17, 676–682.
- [3] Kelleher, N. L., Lin, H. Y., Valaskovic, G. A., Aaserud, D. J., Fridriksson, E. K., McLafferty, F. W., *J. Am. Chem. Soc.* 1999, 121, 806–882.
- [4] Schlutzenzen, F., Tocilj, A., Zaravich, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A. *et al.*, *Cell* 2000, 102, 615–623.
- [5] Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Jr, Morgan-Warren, R. J., Carter, A. P., Vonnrhein, C., Hartsch, T., Ramakrishnan, V., *Nature* 2000, 407, 327–339.

- [6] Yusupov, M. M., Yusupova, G. Z., Baucom, A., Lieberman, K., Earnest, T. N., Cate, J. H. D., Noller, H. F., *Science* 2001, **292**, 883–896.
- [7] Benjamin, D. R., Robinson, C.V., Hendrick, J.P., Hartl, F.U., Dobson, C. M., *Proc. Natl. Acad. Sci. USA* 1998, **95**, 7391–7395.
- [8] Moini, M., Huang, H., *Electrophoresis* 2004, **25**, 1981–1987.
- [9] Strader, M. B., Verberkmoes, N. C., Tabb, D. L., Connelly, H. M., Barton, J. W., Bruce, B. D., Pelletier, D. A. *et al.*, *J. Proteome Res.* 2004, **3**, 965–978.
- [10] Dreger, M., *Mass Spectrom. Rev.* 2003, **22**, 27–56.
- [11] Arnold, R. J., Reilly, J. P., *Anal. Biochem.* 1999, **269**, 105–112.
- [12] Wilcox, S. K., Cavey, G.S., Pearson, J. D., *Antimicrob. Agents Chemother.* 2001, **45**, 3046–3055.
- [13] Pineda, F. J., Antoine, M. D., Demirev, P. A., Feldman, A. B., Jackman, J., Longenecker, M., Lin, J. S., *Anal. Chem.* 2003, **75**, 3817–3822.
- [14] Suh, M.-J., Limbach, P. A., *Eur. J. Mass Spectrom.* 2004, **10**, 89–99.
- [15] Henne, A., Brüggemann, H., Raasch, C., Wiezer, A., Hartsch, T., Liesegang, H., Johann, A. *et al.*, *Nat. Biotechnol.* 2004, **22**, 547–553.
- [16] Oshima, T., Imahori, K., *Int. J. Syst. Bacteriol.* 1974, **24**, 102–112.
- [17] Kristjansson, J. K., Hreggvidsson, G. O., Alfredsson, G. A., *Appl. Environ. Microbiol.* 1986, **52**, 1313–1316.
- [18] Cameron, D. M., Gregory, S. T., Thompson, J., Suh, M.-J., Limbach, P. A., Dahlberg, A. E., *J. Bacteriol.* 2004, **186**, 5819–5825.
- [19] Polevoda, B., Sherman, F., *J. Biol. Chem.* 2000, **275**, 36479–36482.
- [20] Demirev, P. A., Lin, J. S., Pineda, F. J., Fenselau, C., *Anal. Chem.* 2001, **73**, 4566–4573.
- [21] Gonzales, T., Robert-Baudouy, J., *FEMS Microbiol. Rev.* 1996, **18**, 319–344.
- [22] Hirel, P. H., Schmitter, M. J., *Proc. Natl. Acad. Sci. USA* 1989, **86**, 8247–8251.
- [23] Lhoest, J., Colson, C., *Mol. Gen. Genet.* 1977, **154**, 175–180.
- [24] Muranova, T. A., Muranov, A. V., Markova, L. F., Ovchinnikov, Y. A., *FEBS Lett.* 1978, **96**, 301–305.
- [25] Dognin, M. J., Wittmann-Liebold, B., *Eur. J. Biochem.* 1980, **112**, 131–151.
- [26] Triantafyllidou, D., Simitsopoulou, M., Franceschi, F., Choli-Papadopoulou, T., *J. Protein Chem.* 1999, **18**, 215–223.
- [27] Kowalak, J. A., Walsh, K. A., *Protein Sci.* 1996, **5**, 1625–1632.
- [28] Colson, C., Lhoest, J. and Urlings, C., *Mol. Gen. Genet.* 1979, **169**, 245–250.
- [29] Tompson, J., Musters, W., Cundliffe, E., Dahlberg, A. E., *EMBO J.* 1993, **12**, 1499–1504.
- [30] Kurland, C. G., Jorgensen, F., Richter, A., Ehrenberg, M., Bilgin, N., Rojas, A. M., *Through the Accuracy Window*, American Society for Microbiology, Washington, D.C. 1990, pp. 513–526.
- [31] Wittmann, H. G., Littlechild, J. A., Wittmann-Liebold, B., *Ribosomes*, Univ. Park Press, Baltimore 1980, pp. 51–88.
- [32] Amaro, A. M., Jerez, C. A., *J. Bacteriol.* 1984, **158**, 84–93.
- [33] Tomb, J.-F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D., Ketchum, K. A. *et al.*, *Nature* 1997, **388**, 539–547.
- [34] Alm, R. A., Ling, L.-S. L., Moir, D. T., King, B. L., Brown, E. D., Doig, P. C., Smith, D. R. *et al.*, *Nature* 1999, **397**, 176–180.