A liquid chromatography/electrospray mass spectrometric study on the post-transcriptional modification of tRNA

Hisaaki Taniguchi* and Nobuhiro Hayashi

Division of Biomedical Polymer Science, Institute for Comprehensive Medical Science, Fujita Health University, Toyoake, Aichi 470-1192, Japan

Received October 31, 1997; Revised and Accepted February 3, 1998

ABSTRACT

Liquid chromatography/electrospray mass spectrometry is one of the rapidly developing techniques with which mass of large hydrophilic polymers such as proteins and nucleic acids can be determined precisely. The technique was applied to studies on the modifications of tRNAs. Various tRNA species purified from *Escherichia coli* were directly injected into a capillary reversed-phase column and the desalted and concentrated tRNAs were analyzed on-line with an electrospray mass spectrometer. In some cases, small but significant differences were noted between the theoretical and observed molecular masses, suggesting that there exist still unknown modifications. Under high resolution measurements, multiple peaks corresponding to species modified to a varying extent were resolved. To study the structures in detail, the isolated tRNA species were digested with ribonuclease T1, and the resulting mixture of fragments were analyzed by the same liquid chromatography/mass spectrometry. In this way, most of the fragments were easily identified solely from their masses, and the positions where the expected and real structures differ were revealed. The results obtained showed the presence of micro-heterogeneity among tRNAs and demonstrated at the same time the power of the hyphenated technique for the structural analysis on nucleic acids.

INTRODUCTION

Recent development in mass spectrometry has made it possible to ionize large hydrophilic polymers such as proteins and determine their molecular mass with high precision and resolution (1–3). The precision obtained with a conventional quadrupole mass spectrometer coupled with electrospray ionization is often better than ±0.01%, so that even a slight deviation from the theoretical mass based on the gene sequence can be easily detected. We have previously shown that the application of the method to studies of protein phosphorylation and those of protein acylation is very successful (4–6). Peaks of proteins phosphorylated to different degrees were resolved, and the presence of *in vivo* phosphorylated species could be demonstrated (7,8). Capillary high performance liquid chromatography connected on-line to the electrospray mass spectrometer (LC/MS) was further used to identify various *in vivo* phosphorylation sites (6,8,9).

However, the application of the mass spectrometry to another important biopolymer, nucleic acids, is still in its infancy (for review see reference 10). This is mainly due to the lack of suitable solvents (for the electrospray ionization) or matrices (for matrix-assisted laser desorption ionization) which can be used to ionize large polynucleotides efficiently. Furthermore, the negatively-charged phosphate backbones have high affinities for non-volatile cations such as Na⁺ and K⁺. This necessitates either the extensive removal of the ions (11) or the use of elaborate machinery such as a Fourier transform mass spectrometer (12). Recent studies from several laboratories overcame these problems by choosing suitable solvents for the ionization (13,14). The addition of organic bases was also found to be efficient in suppressing the Na⁺ and K⁺ adduct ions (15). Furthermore, the direct on-line LC/MS analysis of oligonucleotides has been reported (16).

In the present report, we describe the application of the LC/MS to structural studies of tRNA. Choice of suitable solvents made it possible to observe only molecular peaks without any adduct peaks. Mixtures of polynucleotide fragments produced by endonuclease digestion could be separated by the column chromatography, and their masses were immediately determined with the mass spectrometer. Time consuming and often troublesome off-line fractionation and sample preparations used in the previous studies (11,13) became unnecessary. The results thus obtained revealed micro-heterogeneity of tRNA and demonstrated the usefulness of the LC/MS analysis to elucidate the structures of nucleic acids including the various post-transcriptional modifications.

MATERIALS AND METHODS

Materials
tRNA⁰Phe purified from brewers’ yeast was purchased from Sigma. Crude *Escherichia coli* tRNA mixture was obtained from *E.coli* A 19 cells at the late log phase, and tRNA⁰Phe was purified.

*To whom correspondence should be addressed. Tel: +81 562 93 9381; Fax: +81 562 93 8832; Email: htanigut@fujita-hu.ac.jp
from the mixture as described (17,18). tRNA\textsubscript{Glu} was purified from the same source using conventional chromatography as described previously (19). A mixture of two Met-specific tRNAs, tRNA\textsubscript{Met} and tRNA\textsubscript{Met}, was purified from \textit{E.coli} (20). The \textit{in vitro} transcript of \textit{E.coli} tRNA\textsubscript{Lys} was made with T7 RNA polymerase (19,21). Digestion with ribonuclease T1 (Sankyo, Tokyo, Japan) was carried out as described previously (19).

**RESULTS**

**LC/MS analysis**

A capillary high performance liquid chromatography column (0.3 × 15 cm) packed with a polymer-based reversed-phase material (PerSeptive Biosystems, R2/H perfusion chromatography material) was connected on-line to the electrospray interface of a quadrupole mass spectrometer (PE Sciex API-III) as described previously (6,8). The mass spectrometer was operated in negative mode with the ion source voltage set at –3500 V. The column was eluted with a linear gradient of water-acetonitrile containing 0.3 mM tributylamine-acetate (pH 5.5). The mass values of neutral molecules are used throughout the manuscript.

**LC/MS analysis of yeast tRNA\textsubscript{Phe}**

We have already reported that the addition of organic amine such as tripropylamine or tributylamine as a counter ion during the reversed-phase chromatography not only affected the chromatographic behaviour but also the efficiency of the ionization of oligonucleotides (22). With increasing chain length of the substituent group of the amine, both DNA and RNA bind more strongly to the reversed-phase column and the retention time increases accordingly. At the same time, the ionization efficiency increases with a concomitant decrease in the Na\textsuperscript{+} and K\textsuperscript{+} adduct peaks. Triethylamine, which has been used in the previous study, was found to be less effective both in the adduct suppression and ionization efficiency (22). In the presence of 0.3 mM tributylamine, almost only molecular peaks were observed even with a large polynucleotide such as tRNA\textsubscript{Phe} (Fig. 1a). The molecular mass of the major species was calculated to be 24 953.4 ± 3.5 Da from the multiply-charged ions. This is very close to the theoretical mass based on the expected structure including various post-transcriptional modifications (24 951.3 Da). This clearly indicates the lack of any adduct ions that should shift the apparent mass peaks. Due to the effective desalting and concentration by the on-line capillary reversed-phase column chromatography, the rigorous removal of salts by pretreatment (11,13) was unnecessary.

As can be seen in the spectrum (Fig. 1a), two series of the peaks belonging to minor species were also observed. The presence of two minor components with molecular masses smaller than that of the major component was clearly seen in the mass spectrum obtained by deconvolution (Fig. 1b). One component with a molecular mass of ~24 623 Da may be a truncated form of the tRNA\textsubscript{Phe}. The mass difference between the major peak and this minor peak (~300 Da) suggests that it is truncated at the 3′-end, since the omission of one ‘A’ would cause a mass decrease of 329.2 Da. On the other hand, the identity of the middle peak is not clear. While the patterns of the distribution of the multiply-charged ions were similar between the major component and the truncated form, the ions belonging to the middle component showed a quite different distribution pattern (Fig. 1a). Since the distribution of the multiply-charged ions reflects the accessibility of the solvent and hence the three-dimensional structure of the sample molecules, this suggests that the third component is a contaminating tRNA species having a structure quite different from those of the other two species.

**LC/MS analysis of \textit{E.coli} tRNA\textsubscript{Lys}**

Several tRNA species purified from \textit{E.coli} were then subjected to the same LC/MS analysis. One example is shown in Figure 2a. Deconvolution of a typical electrospray mass spectrum showing multiply-charged ions (inset) gave a mass spectrum of tRNA\textsubscript{Lys} with a sharp single peak. A small peak of a minor species well resolved from the major peak was also observed. However, the molecular mass calculated for the main peak (24 796 Da) is significantly larger than the theoretical one (24 781 Da) (23–26). Since the reliability and the precision is an important issue, we then measured an \textit{in vitro} transcript of \textit{E.coli} tRNA\textsubscript{Lys} under the same conditions. The \textit{in vitro} transcript should have the same structure as the previous sample except for the post-transcriptional modifications. As shown in Figure 2b, one major species together with at least two minor species with higher molecular masses were observed both in the original mass spectrum (inset) and in the deconvoluted mass spectrum. The mass of the major species was calculated from the multiply-charged ions to be 24 444.2 ± 3.3 Da, which is within experimental error of the theoretical mass (24 441.6 Da). The minor species contain probably one or two more additional bases. This demonstrated the accuracy of the mass determination of the present method and suggested that the
Figure 2. Mass spectra of E. coli tRNA\textsuperscript{Lys} purified from E. coli A19 cells by successive column chromatography was directly injected into the LC/MS apparatus (a). The original electrospray mass spectrum with multiply-charged ion series (inset) was deconvoluted to give a mass spectrum. An in vitro transcript of tRNA\textsuperscript{Lys} was made with T7 RNA polymerase (21) and analyzed by the LC/MS (b). A part of the original mass spectrum was expanded to show the presence of three species (inset). Peaks belonging to each species are labeled as a, b and c.

Figure 3. Mass spectra of E. coli tRNA\textsuperscript{Met} and tRNA\textsuperscript{Glu}. A mixture of two Met-specific tRNAs, tRNA\textsuperscript{fMet} and tRNA\textsuperscript{Met}, was purified from E. coli (20) and analyzed by the same LC/MS technique (a). tRNA\textsuperscript{Glu} purified from E. coli A19 cells by ion exchange column chromatography was directly analyzed by the LC/MS (b). The preparation used has previously shown to contain tRNA\textsuperscript{Glu} by sequencing (19). Two settings of the resolution of the mass spectrometer were used. Under normal low resolution conditions, a shoulder on the high mass side was observed (dotted line). More than four peaks were resolved under high resolution conditions (solid line). In the latter experiment, the mass resolution of the quadrupole mass analyzer was increased and the mass step of the data acquisition was decreased from 0.5 to 0.1 a.m.u. The major species is larger than the species corresponding to the known structure (24 528.9 Da) by 13 Da (inset).

difference between the expected and real structure of tRNA\textsuperscript{Lys} purified from E. coli was not due to the artifacts associated with the measurement itself but to the presence of unknown modifications. Furthermore, it should be noted that the accuracy and the resolution obtained in the present study are good enough to sequence the polynucleotides, although C and U nucleotides in RNA cannot be distinguished due to their small mass difference. The difference (330 Da) between the major peak (24 444 Da) and the contaminating peak (24 774 Da) is very close to the mass difference expected with the addition of an A (329.2 Da). For comparison, the addition of a G, C or U would give a mass difference of 345.2, 305.2 or 306.2 Da, respectively.

**LC/MS analysis of E. coli tRNA\textsuperscript{Met}**

Another example is tRNA\textsuperscript{Met}. Since the preparation used contained tRNAs for both formyl Met (initiator) and Met (elongator), this is an interesting case to demonstrate the power of the method to measure complex mixtures. At least three peaks were resolved as shown in Figure 3a. The major peak had a molecular mass of 24 910 Da, while the minor peak of higher mass was 25 132 Da. These two peaks should correspond to the initiator tRNA\textsuperscript{Met} (theoretical mass, 24 926.1 Da) and the elongator tRNA\textsuperscript{Met} (theoretical mass, 25 144.3 Da), although the observed masses were slightly but significantly lower than the theoretical ones. The peak areas of the two peaks in the deconvoluted spectra correspond well to the amounts of the two species in the preparation (20), suggesting that the relative quantification of the various tRNA species by the present method is possible. The origin of the third component of 24 765 Da is not clear; the deletion of a single nucleotide would not account for the differences from the two other peaks.

**LC/MS analysis of E. coli tRNA\textsuperscript{Glu}**

The next tRNA species subjected to the mass analysis was tRNA\textsuperscript{Glu} (Fig. 3b, dotted line). Only one single peak without any minor contaminants was observed, indicating an apparent homogeneity, corresponding to a single band observed in gel electrophoresis (data not shown). However, the molecular mass calculated (24 540 Da) was again significantly larger than the theoretical one (24 528.9 Da). Furthermore, a shoulder at the high mass side was noted, and the peak gave a broader impression than those obtained with other tRNA species (for example see Fig. 2). To resolve the underlying components, the same sample was subjected to the LC/MS under high resolution measurement conditions, where the mass resolution of the quadrupole mass analyzer was increased (Fig. 3b, solid line). The shoulder on the higher mass side was now resolved into at least two peaks with a separation of ~14 Da (Fig. 3b, inset), while the major peak is also split into at least two peaks. The mass of the shoulder (24 527 Da) corresponds very well to the theoretical mass of the known
Figure 4. A total ion chromatogram of tRNA\textsuperscript{Glu} RNase T1 digest and the assignments obtained by the LC/MS analysis. tRNA\textsuperscript{Glu} purified from \textit{E.coli} was digested with RNase T1 was directly injected into the LC/MS apparatus as described under Materials and Methods. The intensity of the total ion current from the detector was plotted against the retention time (a). Most of the fragments were identified solely from the masses. The identified fragments are indicated by arrows under the sequence of tRNA\textsubscript{2 Glu} (b).

structure (24 528.9 Da). The mass difference between the major peak and the shoulder was \(\sim 13\) Da. Note that the mass difference caused by methylation, a common base modification, is 14 Da, suggesting that the major species observed contains one more methylated base that has been left unnoticed. In addition, the tRNA\textsuperscript{Glu} preparation seems to contain a number of species having molecular masses of small differences. Other possibilities such as the presence of contaminated tRNA species can be ruled out, since the LC/MS analysis of the RNase T1 digests excluded the presence of other structurally-unrelated tRNA species in the preparation used as will be described below.

**LC/MS analysis of \textit{E.coli} tRNA\textsuperscript{Glu} RNase T1 digests**

To elucidate the cause of the micro-heterogeneity observed, the same LC/MS analysis was conducted with RNase digests of tRNA\textsuperscript{Glu}. In this case, the purified tRNA\textsuperscript{Glu} was first digested with RNase T1, which cleaves RNA specifically at the 3' side of G, and the resulting mixture of polynucleotides was directly injected into the reversed-phase column. As shown in the total ion chromatogram (Fig. 4a), the polynucleotide fragments were separated during the reversed-phase chromatography, and the eluted peaks were analyzed by mass spectrometry. From the known sequence, theoretical masses of fragments can be easily calculated, and they were compared with those obtained experimentally. Most of the fragments were unequivocally identified solely from their masses (Table 1). The fragments thus identified cover most of the whole structure except for short segments containing successive Gs (Fig. 4b). Some deviations from the known structure were, as expected, observed. One fragment that was tentatively assigned as fragment G2 showed the most notable heterogeneity (Fig. 5a). The peak having a mass close to the theoretical one (2808.8 Da) was a minor component, and three peaks of higher masses were observed. Another fragment was tentatively assigned as G3, but the mass obtained was larger than the theoretical one by 6 Da. These results suggest that the micro-heterogeneity and the discrepancy between the observed and theoretical masses of tRNA\textsuperscript{Glu} originate mainly from the regions close to the 5'-end. The hitherto unnoticed modifications near the 5'-end may be functionally very important, since the part of the molecule, i.e., the acceptor stem, is involved in the recognition of tRNA species by specific amino-acyl tRNA synthase (27,28).

| Table 1. Assignment of \textit{E.coli} tRNA\textsuperscript{Glu} RNase T1 digest |
|-----------------|-----------------|-----------------|
| Mass of M (Da) | Theoretical     | Observed        |
| G2              | 2807.7          | 2808.8/2819.6*  |
| G3              | 1610.0          | 1616.3*         |
| G6              | 1608.0          | 1607.9          |
| G8              | 1937.2          | 1937.4          |
| G6–8            | 3872.4          | 3871.1          |
| G5–11 or G6–12  | 8405.1          | 8412.9          |
| G9              | 3210.0          | 3210.9          |
| G13             | 1962.2          | 1961.2          |
| G17             | 1294.8          | 1295.3          |
| G18             | 3183.9          | 3182.7          |
| G18–22          | 5199.1          | 5198.5          |
| G20–23          | 2547.7          | 2541.6*         |

*Fragment where the expected and observed masses differ significantly.

**Post-transcriptional modifications in the anticodon loop of \textit{E.coli} tRNA\textsuperscript{Met}**

Other tRNA species were also subjected to the same LC/MS analysis. One interesting observation was the heterogeneity of a fragment containing the anticodon loop of tRNA\textsuperscript{Met}. Although the sample used was a mixture of at least three species (Fig. 3a), most of the fragments produced by the RNase T1 digestion were assigned to either to tRNA\textsuperscript{Met} or to tRNA\textsuperscript{Met} (data not shown). Since no other fragments were observed in significant amounts, the third component of 24 765 Da (Fig. 3a) is probably closely related to one of the two Met specific tRNAs. Among various fragments thus assigned, one fragment corresponding to the anticodon loop of tRNA\textsuperscript{Met} showed heterogeneity (Fig. 5b). In addition to the major species of 5273.7 Da, which is in good agreement with the theoretical mass (5276.2 Da), a minor peak of 5231.4 Da was observed. Since the mass difference between the two peaks, 42.3 Da, is very close to that associated with acetylation (42.04 Da), we conclude that the two peaks arise from species modified differently at the first base of the anticodon. It is less likely...
the eluent that helps the ionization process. Second, the formation of the Na\textsuperscript{+} or K\textsuperscript{+} adduct peaks is well suppressed with a concomitant increase of molecular peaks. The substitution of the tributylamine with less hydrophobic organic amines such as triethylamine, which has been used in the previous study (16), produced mass spectra with many adduct ion peaks too complex to be analysed satisfactorily.

Table 2. Theoretical and observed molecular masses of various tRNA species

<table>
<thead>
<tr>
<th>tRNA</th>
<th>Mass(^a) (Da)</th>
<th>Theoretical</th>
<th>Observed</th>
<th>(\Delta)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E.coli tRNA\textsuperscript{Lys}\textsuperscript{transcript}</td>
<td>24 781.0</td>
<td>24 796.4 ± 1.9</td>
<td>+15.4 ± 1.9</td>
<td></td>
</tr>
<tr>
<td>E.coli tRNA\textsuperscript{Lys}\textsuperscript{transcript}</td>
<td>24 441.6</td>
<td>24 444.2 ± 3.3</td>
<td>+2.6 ± 3.3</td>
<td></td>
</tr>
<tr>
<td>(in vitro transcript)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E.coli tRNA\textsuperscript{Glu}</td>
<td>24 528.9</td>
<td>24 540.6 ± 4.7</td>
<td>−11 ± 4.7</td>
<td></td>
</tr>
<tr>
<td>E.coli tRNA\textsuperscript{Met}</td>
<td>24 926.1</td>
<td>24 910.3 ± 4.8</td>
<td>−15.8 ± 4.8</td>
<td></td>
</tr>
<tr>
<td>E.coli tRNA\textsuperscript{Met}</td>
<td>25 144.3</td>
<td>25 131.7 ± 7.1</td>
<td>+12.6 ± 7.1</td>
<td></td>
</tr>
<tr>
<td>yeast tRNA\textsuperscript{Phe}</td>
<td>24 951.3</td>
<td>24 953.4 ± 3.5</td>
<td>+2.1 ± 3.5</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Mean ± S.D.

The results of the mass measurements indicate that the structures of some of the tRNA species purified from *E. coli* differ from the expected structure (Table 2). Here, the precision of the measurement is clearly not the issue, as the *in vitro* transcript of tRNA\textsuperscript{Lys} as well as yeast tRNA\textsuperscript{Phe} showed masses very close to the theoretical values. Furthermore, the mass spectrum of tRNA\textsuperscript{Glu} obtained at high resolution indicates that the tRNA contains not a single species but a group of species with small mass differences. Mass differences between some of the species are ~14–15 Da, suggesting that they are due to methylation (causing a mass difference of 14 Da), a common post-transcriptional modification of bases. Therefore, the tRNA\textsuperscript{Glu} species corresponding to the expected structure is probably the minor species, and the majority contains various unknown modifications to different degrees. The observed mass of tRNA\textsuperscript{Lys} and that of tRNA\textsuperscript{Met} are either larger or smaller than the theoretical mass values by ~14 Da. This suggests again that there is an unknown methylation site, or the known methylation is only partial. Although the extent of several modifications has been reported to be only partial and the presence of micro-heterogeneity has been documented, the extent of partial modifications could not be well determined experimentally. The mass spectrometry can ‘visualise’ this micro-heterogeneity. The LC/MS analysis on the ribonuclease digests described in the present report should help the determination and the identification of the unknown modifications. McCloskey and co-workers have, in fact, shown that the deacylation occurred during the mild treatments employed in the present study. Since the base modification of the anticodon plays important roles in the codon–anticodon recognition (wobbling) and the identity of tRNAs (recognition by specific amino-acyl tRNA synthase), the present method should give a means to analyze the structure–function relationship.

**DISCUSSION**

The LC/MS analysis that has been previously used for studies of the post-translational modifications of proteins is now successfully applied to the structural studies of tRNA. The use of the on-line reversed-phase column chromatography allows a direct analysis of samples containing various salts and additives without any pretreatment. Previously, extensive desalting has been necessary to obtain electrospray mass spectra of reasonable quality, even under the same measuring conditions. The advantage of the addition of the organic amine with long substituents is twofold. First, the binding of hydrophilic polynucleotides to the reversed-phase column becomes stronger with increasing chain length of the substituents. This results not only in better chromatographic separation but also in the increase of organic solvent concentration in the eluent that helps the ionization process. Second, the formation of the Na\textsuperscript{+} or K\textsuperscript{+} adduct peaks is well suppressed with a concomitant increase of molecular peaks. The substitution of the tributylamine with less hydrophobic organic amines such as triethylamine, which has been used in the previous study (16), produced mass spectra with many adduct ion peaks too complex to be analysed satisfactorily.

**Figure 5.** Mass spectra of RNase T1 fragments showing heterogeneity. tRNA\textsuperscript{Glu} was digested with RNase T1 and analyzed by the LC/MS. A part of the scans containing the heterogeneous fragment G2 was shown (a). A deconvoluted spectrum showing the heterogeneity of a fragment containing the anticodon loop of tRNA\textsuperscript{Met} was observed during the LC/MS analysis of the RNase T1 digest of tRNA\textsuperscript{Met}-tRNA\textsuperscript{Met} mixture (b). The position of an N\textsuperscript{4}-acetyl C in the anticodon of tRNA\textsuperscript{Met} is indicated by an arrowhead.
in the samples. Therefore, peak relative intensities observed in the deconvoluted spectra give at least rough, in some instances quantitative, estimates of the relative amounts of the species present in the samples. A notable example in the present study is the heterogeneity of the fragment containing the anticodon loop of tRNA\textsuperscript{Met}. The presence of a modified base in the anticodon may regulate the activities of the tRNA, and it would be of interest to analyse the dynamic regulation of the modification. The present method should allow a rapid, relative quantification of the modifications under various cellular conditions such as heat shock.

ACKNOWLEDGMENTS

This work was supported in part by Grants-in-Aid from the Fujita Health University, a Grant-in-Aid for Developmental Scientific Research (07558219) and Grants-in-Aid for Scientific Research on Priority Areas (08283220 and 09272222) from the Ministry of Education, Science, Sports and Culture, Japan.

REFERENCES