Development of
SI-traceable and High Accurate Method
for Quantitative Analysis of DNA

January 2017

Sachie KATTO

Development of SI-traceable and High Accurate Method for Quantitative Analysis of DNA

A Dissertation Submitted to
the Graduate School of Life and Environmental Sciences,
the University of Tsukuba
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Biotechnology
(Doctoral Program in Bioindustrial Sciences)

Sachie KATTO

Contents

Abstract

In the bioindustry, DNA quantification is a widely used and essential routine process. Such routine DNA measurements could benefit greatly from more accurate and reliable DNA measurements; however, neither an appropriate DNA reference material nor a suitable DNA quantification method is currently available. In this study, I developed SI-traceable and accurate methods for qualitative analysis of DNA, which may allow the characterization of DNA certified reference material (CRM).

In Chapter 2, I have described a highly sensitive method for the analysis of deoxynucleotide monophosphates (dNMPs), which involves the use of liquid chromatography/mass spectrometry (LC/MS) and a new metal-free column. The new column solves the problem of the phosphate group in dNMPs interacting with the metal of the device or column. After optimization of the analytical conditions, the limits of detection (LODs) of dNMPs ranged from 5.4 ng/g to 6.3 ng/g. These values were 10 times lower than the LODs of previous methods. I applied the method to determine the base composition and quantification of a 20-mer oligonucleotide. Despite using only a very small sample of 14.5 ng, I was able to determine the base composition, and the result was consistent with theoretical values. I was also able to quantify the mass fraction of oligonucleotide at 8.2% expanded uncertainty ($k = 2$).

In Chapter 3, the development of liquid chromatography-isotope dilution mass spectrometry (LC-IDMS) with formic acid hydrolysis for the accurate quantification of λDNA was described. The over-decomposition of nucleobases in formic acid hydrolysis was restricted by optimizing the reaction temperature and reaction time and was accurately corrected by using dNMPs and isotope-labeled dNMPs as the calibrator and internal standard, respectively. The present method can quantify λDNA with an expanded uncertainty of 3.4% using 300 ng of λDNA. The analytical

results obtained with the present method were validated by comparison with the results of phosphate-base quantification by inductively coupled plasma-mass spectrometry (ICP-MS). The results showed good agreement between the two methods. I conclude that the formic acid hydrolysis/LC-IDMS method can quantify λDNA accurately and shows promise as the primary method for the certification of DNA as reference material.

In Chapter 4, I have established a method for the separation and quantification of DNA fragments in mixed DNA samples, using high performance liquid chromatography (HPLC) with an anion-exchange column. Using an NaCl concentration gradient, DNA fragments in mixed DNA samples were separated well. A calibration curve from 0.05 to 12.4 ng/μL was obtained, which showed high linearity and a correlation coefficient of 0.9999. The LOD for S/N=3 was 0.02 ng/μL and the limit of quantification (LOQ) for S/N=10 was 0.06 ng/μL. The relative standard deviation (RSD) was less than 2% for measurement of peak area repeatability. The recovery of approximately 1 ng/μL of a specific DNA spiked in a mixed DNA sample was 99.9 ± 3.6%. The method was able to measure the degradation rate of 600 bp DNA with a variation of approximately 1%.

From these results, obtained by a combination of developed methods, the sample DNA can be quantified with SI-traceability and high accuracy. The use of a combination of methods has the advantage of balancing the strengths and weaknesses of each method. This would make possible the development of DNA CRM and allow reliable DNA quantification.

Chapter 1 General introduction

DNA quantification is an essential process that is widely carried out in bioindustrial science. In this regard, highly reliable DNA quantification is important for the determination of the legislative threshold for genetically modified organisms (GMOs) [1]. In this chapter, an overview of current DNA quantification methods, general DNA quantification methods, and DNA and DNA-related compound quantification methods using high-performance liquid chromatography (HPLC) and liquid chromatography-mass spectrometry (LC-MS) is provided, and the objectives are discussed.

1.1 Fields engaged in DNA quantification and importance of accurate DNA quantification

In many fields related to bioindustrial science, including food analysis and medicine, DNA quantification is carried out as routine work. In food analysis, GMO and microbial contamination of food is quantified by DNA measurement. Because the quantification of GMO contamination is important to meet legislative requirements by country [1], an accurate GMO quantification method is required. In medicine, because nucleic acid drugs with DNA or RNA as the main drug component have been developed, stringent quality control of nucleic acid drugs is also required [2, 3] and the number of genetic tests has been increasing year by year. To confirm the reliability and validity of a genetic test, the same and/or comparable results should be obtained regardless of when or where the genetic test is done, and therefore, the results of the genetic test should be traceable to International System (SI) units.

Obviously, accurate DNA quantification is a must in bioindustrial science. However, neither an appropriate DNA reference material nor a suitable method for DNA quantitative measurement is available.

1.2 General DNA quantification methods

There are three methods for general DNA quantification and their advantages and disadvantages are discussed as follows.

One method is ultraviolet (UV) absorbance measurement [4, 5]. Because of its low cost and simple principle, i.e., the Beer-Lambert law, UV absorbance measurement at 260 nm, which is the local absorption maximum of DNA, is carried out for the determination of DNA concentration. However, the conversion factors for this method are not SI-traceable and the variation of the quantitative values obtained by this method is very large.

The other methods are the PicoGreen assay and the real-time quantitative polymerase chain reaction (qPCR) in combination with a fluorescent dye [5, 6]. Because those methods use fluorescence for detection, solutions having low DNA concentrations can be evaluated. However, the PicoGreen assay and qPCR require an external calibrator that has the same sequence as the analyte DNA sequence for SI-traceable and accurate DNA quantification.

Because of problems arising from the use of conversion factors in UV absorbance measurement and the dependence of the PicoGreen assay and qPCR on an external calibrator, those methods fail to accurately quantify DNA. Therefore, an SI-traceable and accurate method for the characterization of DNA certified reference material (CRM) for effective comparison of quantified values, high quality control of DNA quantification, and method validation is required.

1.3 Quantification of DNA and DNA-related compounds using analytical instruments

Such analytical instruments as HPLC and LC-MS are widely used for the quantification of compounds with SI traceability and high accuracy. These analytical instruments are also expected to be powerful tools for the quantification of DNA and DNA-related compounds with SI traceability

and high accuracy. A summary of those methods is provided below.

DNA is produced by polymerizing deoxyribonucleotide monophosphate (dNMPs), which are DNA monomers, and the molecular weights of produced DNAs vary depending on the degree of polymerization, i.e., from a few nucleotides rectitude as seen in a single-stranded oligonucleotide to several million base pairs (bp) of double-stranded eukaryotic genomes. Methods for measuring DNA include capillary electrophoresis (CE) [7] and HPLC with a size-exclusion column [8, 9], an ion-pair reversed-phase column, [10, 11] or an anion-exchange column [10, 12-16]. Because CE-based DNA quantification methods have high resolution but yield a large variation in the results due to the small injection volume, and HPLC-based methods have high accuracy but poor resolution, DNA quantification has remained a difficult task. A method for quantifying DNA with high resolution and high accuracy is required.

On the other hand, as a different approach to accurate DNA quantification, a DNA quantification method that analyzes DNA monomers, such as dNMPs or deoxyribonucleoside (dNs), by HPLC or LC-MS was reported [3, 15, 17-22]. Those DNA monomers were generated by enzymatic digestion. Because that method can quantify DNA indirectly through the quantification of dNMPs or dNs, a DNA standard is not required. Furthermore, that method can quantify DNA with SI traceability and high accuracy by using both an appropriate single-nucleotide standard and LC-isotope dilution mass spectrometry (LC-IDMS) [15, 17-20]. For accurate DNA quantification by LC-IDMS, complete DNA digestion and control of over-decomposition are required. However, when enzymatic digestion is used with LC-IDMS for DNA quantification, a large amount of DNA sample, more than 2 μg, is required because dNMPs are difficult to analyze with high sensitivity by LC-MS.

Thus, the analysis of DNA and DNA-related compounds by HPLC and LC-MS is not easy but offers such promising features as SI traceability and high accuracy.

1.4 Objectives

As mentioned above, although accurate DNA quantification is rapidly gaining importance, neither DNA CRM nor a suitable method for DNA quantification is available. In this research work, I was aiming to develop SI-traceable and high-accuracy methods for the quantitative analysis of DNA. As those methods can be used for the characterization of DNA CRM, they will be applicable to all types of DNA quantification.

Chapter 2 Metal-free columns for dNMP determination by LC-MS and application to oligonucleotides

2.1 Introduction

A growing need to quantify dNMPs and oligonucleotides has emerged in recent years. The analysis of dNMPs is necessary for the identification and quality control of nucleic acid drugs [2, 3] and the determination of DNA methylation [23]. dNMPs are DNA monomers, and they have been used as analytes for the accurate quantification of oligonucleotides and DNA [17-21]. HPLC is the instrument of choice in many of those studies. HPLC is widely employed for the separation and quantification of small molecules with high precision and accuracy. Many detection techniques are used in the analysis of dNMPs by HPLC, including ultraviolet-visible (UV-Vis) spectrometry [3], mass spectrometry (MS) [15, 17-20, 22], and inductively coupled plasma MS (ICP-MS) [24-27]. Among those techniques, ICP-MS has the highest sensitivity for dNMPs [26, 27]. Because ICP-MS is capable of high-sensitivity elemental analysis, the high-sensitivity analysis of dNMPs can be achieved by detecting phosphorus in dNMPs.

dNMPs are hydrophilic compounds having a nucleic acid base, a deoxyribose, and a phosphate group. In the analysis of dNMPs using a conventional octa decyl silyl (ODS) column, peak tailing attributed to strong chelate complexation between the phosphate group and an active metal ion, such as Fe(III), on the surface of LC and the LC column was noted [25, 28]. Therefore, for accurate DNA quantification by LC, approximately 2 to 100 μg of DNA sample was required [15, 17-20, 22]. However, as the amount of DNA sample is limited, a highly sensitive analytical method for dNMPs, which would enable accurate quantification with a small amount of DNA, is desired.

In this work, I examined a new metal-free column for LC-MS with an eye to developing highly sensitive analytical methods for dNMPs. As the chromatographic tube and frit in conventional

ODS columns are made of stainless steel, active metal ions are eluted from them. In contrast, the chromatographic tube and frit in the newly developed column are made of metal-free material, such as glass-lined stainless steel and polyethylene, and therefore, metal ion elution from them is reduced [29, 30]. By using this column, good peak shape was obtained for dNMPs. MS was used for detection, and stable isotope labeled dNMPs were used as the internal standard (IS). The stable isotope labeled compounds had the same chemical characteristics as the analyte and were therefore an ideal IS for the correction of injection or ionization efficiency. I also applied this method to the determination of oligonucleotide base composition.

2.2 Experimental

2.2.1 Reagents and samples

2'-Deoxyadenosine 5'-monophosphate disodium salt (dAMP), 2'-deoxycytidine 5'-monophosphate disodium salt (dCMP), 2'-deoxyguanosine 5'-monophosphate disodium salt (dGMP), and 2'-deoxythymidine 5'-monophosphate disodium salt (dTMP) were purchased from MP Biomedicals (CA, USA). The dNMPs were dissolved in water and their mass fractions were quantified by ^{1}H NMR measurement. ^{13}C- and ^{15}N-labeled dAMP, dCMP, and dTMP were purchased from Spectra Stable Isotopes (MD, USA), and ^{13}C- and ^{15}N-labeled dGMP was purchased from ISOTEC International (GA, USA). Ammonium acetate, ammonia for pH adjustment, and LC-MS grade methanol were purchased from Wako (Osaka, Japan). Phosphodiesterase I (PD1) was purchased from Sigma-Aldrich Japan (Tokyo, Japan). Synthesized 20 mer oligonucleotide (5'-TAATACGACTCACTATAGGG-3') was used as sample. Lyophilized 20 mer oligonucleotide was dissolved in water and concentration was roughly determined by UV spectrometry. Water used for preparation was purified water from Milli-Q Advantage (Millipore, Tokyo, Japan).

2.2.2 Equipment and measurement

The HPLC system used was the LC-10A series (Shimadzu, Kyoto, Japan), and a 0.13 mm i.d. polyether ether ketone tube was adopted to reduce dead volume. The LCMS-8030 triple quadrupole mass spectrometer (Shimadzu) was used as the detector. Separation of dNMPs was achieved with a metal-free column (L-column 2 ODS, 2.0 mm i.d. × 150 mm, 3 μm, metal-free, Chemicals Evaluation and Research Institute (CERI), Saitama, Japan) [29, 30]. To evaluate the effect of metals, a conventional ODS column that used metal frit (L-column 2 ODS, 2.0 mm i.d. × 150 mm, 3 μm, CERI) was adopted. To optimize the analytical conditions, methanol concentration and pH of the aqueous mobile phase were determined. The optimum conditions were 1% methanol and 20 mM ammonium acetate (pH 8.0) at the flow rate of 0.2 mL/min. Injection volume was 10 μL and column bath temperature was 40^{o}C. MS conditions were electrospray ionization and selective ion monitoring in the negative ion mode. Flow rate of nebulizing gas was 3 L/min, desolvation line (DL) temperature was 250^{o}C, heat block temperature was 400 ^{o}C, and flow rate of drying gas was 15 L/min. Ions monitored for dAMP, dCMP, dGMP, dTMP, isotope-labeled dAMP, isotope-labeled dCMP, isotope-labeled dGMP, and isotope-labeled dTMP were 330.1, 306.0, 346.1, 321.0, 345.1, 318.1, 361.1, and 333.1, respectively. The UV spectrometer used for the quantification of 20 mer oligonucleotide was a BioSpec-nano (Shimadzu).

2.2.3 dNMP standards for calibration curve preparation

Mixed dNMP standards of six different mass fractions were used to prepare the calibration curve. The mass fractions of the four dNMP in each standard solution were 5 ng/g, 10 ng/g, 50 ng/g, 100 ng/g, 450 ng/g, and 900 ng/g, respectively. The standard solutions also contained isotope-labeled dNMPs at the mass fraction of 900 ng/g dNMPs. In all cases, samples were prepared gravimetrically.

2.2.4 Determination of base composition and quantification of 20 mer oligonucleotide by enzymatic digestion

The calibration blend and the sample blend were prepared. The calibration blend was a mixture of dNMPs and isotope-labeled dNMPs at the same concentration. The concentrations of dNMPs and isotope-labeled dNMPs were matched with the estimated concentration of individual dNMPs in the digested 20 mer oligonucleotide. The sample blend was a mixture of approximately 2 ng/μL 20 mer oligonucleotide solution and isotope-labeled dNMPs, the concentrations of which were the same as those of the calibration blend. Ten μL of the calibration blend or the sample blend was incubated at 37^{o}C for 2 h in a 200 μL sample tube after the addition of 10 μL of 2 mU/mL PD1 solution. The digestion was stopped by heat shock at 95 ^{o}C for 20 min, to deactivate PD1. To remove PD1 from the digested sample, ultra-filtration was carried out as follows. A Nanosep 30K Omega device (Nihon Pall Ltd., Tokyo, Japan) was conditioned with 500 μL of water and centrifugation was carried out at 14,000 × g for 10 min at room temperature twice. The digested sample was diluted with 30 μL of water and loaded on a Nanosep 30K Omega cartridge. The cartridge was centrifuged at 14,000 × g for 10 min at room temperature. The filtrate was used as LC-MS sample for the quantification of dNMPs. Enzymatic digestion was conducted in triplicate, and each sample was analyzed four times by LC-MS.

2.3 Results and Discussion

2.3.1 Optimization of analytical conditions for dNMPs using metal-free column

For the analysis of dNMPs on the metal-free column, the analytical conditions were optimized with 900 ng/g dNMP standard mixture. Because dNMPs having both phosphate and amine groups are usually ionized in aqueous solution, they are not likely to be retained on an ODS column. In this research, the concentration of methanol and the pH of the aqueous mobile phase

were optimized. To estimate the effect of metal in the column, dNMP standard was analyzed on both metal-free column and conventional column under the same LC conditions. The conventional column used the same material as the metal-free column, except that a metal frit was used. In previous studies that analyzed dNMPs for DNA quantification [17, 19], isocratic conditions with a very low concentration of methanol were used. After performing experiments using several concentrations of methanol, we found that good separation was achieved with 1% methanol.

Because MS was used as the detector, acetic acid buffer was used as the aqueous mobile phase. We examined three kinds of aqueous mobile phase: 0.1% acetate at pH 3.3, 20 mM ammonium acetate buffer at pH 5.9, and 20 mM ammonium acetate buffer at pH 8.0. The optimum mobile phase was selected on the basis of the signal-to-noise ratio (S/N) and the peak shape of 900 ng/g dNMPs. As the index of peak shape, tailing factor (T_f) was used. T_f was calculated using equation (1) [18].

$$T_f = \frac{W_{0.05}}{2d} \quad (1)$$

Here, $W_{0.05}$ is the peak width at 5% peak height and d is the former peak width (from peak start to peak tip) at 5% peak height. $T_f < 1$ means peak fronting, $1 < T_f < 1.2$ indicates good peak shape, and $1.2 < T_f$ means peak tailing. When 0.1% acetate at pH 3.3 was used as the mobile phase, S/N of dNMP analysis ranged from 220 to 880, and the highest sensitivity was observed among the three kinds of buffer used in this study. However, T_f of each dNMP ranged from 2.3 to 3.2, and peak tailing was observed. When 20 mM ammonium acetate buffer at pH 5.9 was used, S/N of dNMP analysis ranged from 100 to 260, and the lowest sensitivity was observed. T_f ranged from 2.1 to 3.5, and peak tailing was observed. When 20 mM ammonium acetate buffer at pH 8.0 was used, S/N and T_f were from 320 to 680 and from 1.3 to 1.8, respectively. Therefore, dNMPs showed good peak shape with 20 mM ammonium acetate at pH 8.0. Zhang et al. [28] reported that compounds containing a phosphate group showed peak tailing when a conventional reversed-phase column and

an aqueous mobile phase at pH 3 to 7 were used. The peak tailing was due to the interaction between the deprotonated phosphate groups and the protonated silanol groups. Our result is consistent with Zhang et al. report, and interaction between the phosphate group and the silanol group occurred on the metal-free column.

On the other hand, to estimate the effect of metals, 900 ng/g dNMP standard was analyzed using both the metal-free column and the conventional column under the same LC conditions of 20 mM ammonium acetate at pH 8.0 with 1% methanol as mobile phase. When the conventional ODS column was used, S/N of dNMP analysis ranged from 70 to 110, and T_f ranged from 2.3 to 3.3. These S/N values were approximately one-third to one-eighth of those when the metal-free column was used, and significant peak tailing was observed. The results suggest that the effect of the interaction between metals in the conventional ODS column and the phosphate group was not negligible even under the basic condition. Therefore, a basic mobile phase and a metal-free column were deemed the better choice when the analyte contained a phosphate group, such as dNMPs.

From the results, the metal-free column and 20 mM ammonium acetate buffer at pH 8.0 with 1% methanol were chosen as the optimum LC conditions because S/N was sufficient for the quantification of dNMPs and good peak shape was obtained. The chromatogram obtained by analyzing 900 ng/g dNMP mixture under the optimum conditions is shown in Fig. 2-1. In Fig. 2-1, dGMP and dTMP were eluted at the same time; however, those compounds could be separated by MS because of their different m/z values, namely, 346.1 and 321.0, respectively.

2.3.2 Method validation

To validate this method, the linear range of the calibration curve, the limit of detection (LOD, S/N = 3), and the limit of quantification (LOQ, S/N = 10) were calculated. The results are shown in Table 2-1. A calibration curve with high linearity was obtained for each dNMP. The LODs

ranged from 5.4 ng/g to 6.3 ng/g, and were ten times lower than the LODs obtained by HPLC-ICP-MS and μHPLC-ICP-MS [26, 27], indicating that the developed method with conventional MS detection has high sensitivity. The repeatability of this method was evaluated by measuring the peak area ratios of the six dNMP standards and expressed as relative standard deviation (RSD, n = 4). At the concentration of 50 ng/g, which was ten times higher than LOD, RSDs ranged from 2.5% to 4%. At concentrations higher than 450 ng/g, which were approximately one hundred times higher than LOD, RSDs were less than 1%. Therefore, using the developed method, dNMPs at concentrations as low as approximately 50 ng/g could be analyzed with high sensitivity and high repeatability of less than 5%.

2.3.3 Application to analysis of 20 mer oligonucleotide

Analysis of dNMPs from digested 20 mer oligonucleotide

The developed method was used to analyze dNMPs from an enzymatically digested 20 mer oligonucleotide. The obtained mass fractions of dNMPs from the digested 20 mer oligonucleotide are shown in Table 2-2(a). The mass fractions of dNMPs in the digested sample were from 306 ng/g to 545 ng/g. When dNMPs in the digested sample were measured by LC-MS, the peak area ratios of the sample blend and the calibration blend varied from 2.1% to 6.9%. The reason for the large variation was that the actual mass fractions of dNMPs in the digested sample measured by LC-MS were approximately 50 ng/g or less than 50 ng/g because the sample was diluted. Because the repeatability of 50 ng/g dNMP standard was approximately 5%, the enzymatically digested sample could be analyzed by this method with the same variation as the dNMP standard. From the results, we estimated the uncertainty according to the GUM guide (Guide to the Expression of Uncertainty in Measurement, ISO) [19]. The uncertainties of this measurement were calculated by combining the relative standard uncertainties associated with (1) the mass ratio of sample blends and

calibration blends (weighting), (2) the mass fraction of the standard dNMP solution, (3) the variation arising from LC-MS measurements of sample blends and calibration blends, and (4) the variation arising from enzymatic digestion. The estimated uncertainties are listed in Table 2-3. The expanded uncertainties (k = 2) of dNMP quantification were from 3.0% to 7.2%.

Determination of base composition of 20 mer oligonucleotide

The base composition of 20 mer oligonucleotide was calculated from the mass fractions of dNMPs, and found to be dAMP: dCMP: dGMP: dTMP = 1.73: 1.05: 0.95: 1. The theoretical base composition of 20 mer oligonucleotide, which was calculated from the sequence, was dAMP: dCMP: dGMP: dTMP = 1.75: 1: 1: 1. (T of 5' terminus of 20 mer oligonucleotide did not have a phosphate group for the characteristic of synthesis, and thymidine was generated when 20 mer oligonucleotide was digested. The number of dTMPs was one less than the sequence.) Therefore, the base composition calculated by this method was consistent with the theoretical base composition within 5%.

Quantification of 20 mer oligonucleotide

The mass fractions of 20 mer oligonucleotide were calculated from the mass fraction of each dNMP, and are shown in Table 2-2(b). The results were consistent within the range of expanded uncertainties (k = 2). By using those mass fractions, the final mass fraction and uncertainty of 20 mer oligonucleotide were calculated with reference to the report of Kinumi et al. [19]. The final mass fraction of 20 mer oligonucleotide quantified by this method was 1.45 μg/g ± 0.12 μg/g (mass fraction ± expanded uncertainty (k = 2)). The sample amount used was calculated to be approximately 14.5 ng for one enzymatic digestion. Therefore, by using this method, a very small amount of sample, 14.5 ng, is sufficient for the quantification of 20 mer oligonucleotide with an expanded uncertainty of 8.2% (k = 2). Because this uncertainty was equal to those of previous

reports that quantified oligonucleotide or DNA by LC-IDMS using dNMPs [19, 20], this method could quantify oligonucleotide with equivalent uncertainty using 1/20 to 1/1,000 of the sample amount.

2.4 Conclusions

An analytical method that uses LC-MS and a metal-free column for the analysis of dNMPs with high sensitivity was developed. By using the metal-free column, the interaction between the phosphate group of dNMPs and metal in the column was reduced. Optimization of the analytical conditions improved the peak shape and the sensitivity of LC-MS detection: 5.4 ng/g to 6.3 ng/g dNMPs could be detected by LC-MS, and those values were ten times lower than the LODs of previous reports. Despite use of a very small amount of sample, 14.5 ng, the base composition was obtained, and the results were consistent with theoretical values. The mass fraction of oligonucleotide was quantified with an expanded uncertainty of 8.2% (k = 2). This method shows promise for use in the quality control of nuclear acid drugs, which requires determination of the base composition and oligonucleotide quantification.

Table 2-1 The linear range, LOD and LOQ of dNMPs by this method.

	Linear range (ng/g)	LOD (ng/g)	LOQ (ng/g)
dAMP	795	6.3	21
dCMP	930	5.8	19
dGMP	888	5.6	19
dTMP	908	5.4	18

Table 2-2 Results of dNMP analysis of digested 20 mer oligonucleotide. (a) Mass fraction of dNMPs; (b) calculated Mass fraction of 20 mer oligonucleotide.

(a)	dNMPs		(b)	20 mer oligonucleotide	
	Mass fraction (ng/g)	Expanded uncertainty (ng/g, *k*=2)		Mass fraction (μg/g)	Expanded uncertainty (μg/g, *k*=2)
dAMP	545	16		1.44	0.04
dCMP	306	20		1.52	0.10
dGMP	315	23		1.39	0.10
dTMP	306	15		1.45	0.07

Table 2-3 Uncertainty components of 20 mer oligonucleotide quantification obtained from each dNMPs analysis. Each value is shown as relative standard uncertainty (%).

	Relative standard uncertainty (%)			
	dAMP	dCMP	dGMP	dTMP
Peak area ratio of sample blend	0.84	1.49	1.08	1.52
Peak area ratio of calibration blend	1.11	2.82	3.34	1.85
Enzymatic digestion	0.49	0.00	0.72	0.00
Mass fraction of the standard dNMPs solution	0.23	0.85	0.40	0.37
Mass ratio of sample blend	0.09	0.09	0.09	0.09
Mass ratio of calibration blend	0.01	0.01	0.01	0.01
Combined standard uncertainty	1.5	3.3	3.6	2.4

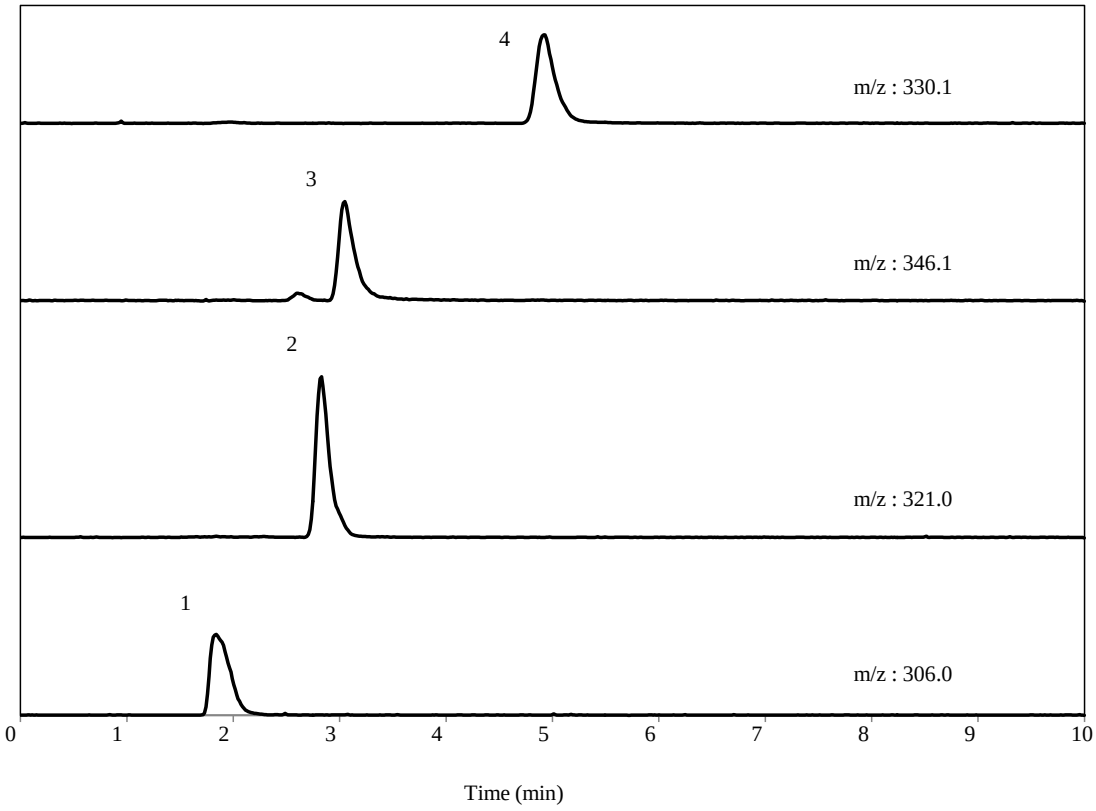

Fig.2-1 The LC-MS chromatogram of 900 ng/g dNMP mixture. 1: dCMP, 2: dTMP, 3: dGMP, 4: dAMP.

Chapter 3 Formic acid hydrolysis/liquid chromatography-isotope dilution mass spectrometry: an accurate method for large DNA quantification

3.1 Introduction

DNA quantification is carried out in many fields, including food analysis, medicine, and pharmacy. In food analysis, GMO or microbial contamination of food is quantified by DNA measurement. Because the detection of GMO contamination is important to meet legislative requirements by country [1], a method for the accurate quantification of GMOs is needed. The detection of disease-causing microorganisms in medicine or the quality control of nucleic acid drugs in pharmacy is carried out as routine work by measuring DNA. The DNA size of interest ranges from a few nucleotide bp as seen in single-stranded oligonucleotides to millions of bp typical of double-stranded eukaryotic genomes [2]. Routine DNA measurement could benefit greatly from more accurate and reliable DNA measurement; however, neither an appropriate DNA reference material nor a suitable method for DNA quantitative measurement is available.

For the determination of DNA concentration, UV absorbance and fluorescence measurements are generally conducted [4, 5]. The measurement of UV absorbance at 260 nm, which is the local absorption maximum of DNA, is commonly carried out for the determination of DNA concentration; however, the conversion factors for this method are not SI-traceable. Other popular methods include the PicoGreen assay and real-time qPCR in combination with a fluorescent dye [5, 6]. However, because of problems arising from the use of conversion factors in UV absorbance measurement and the dependence of the PicoGreen assay and qPCR on external calibrators, these methods fail to accurately quantify DNA [20]. Therefore, an accurate and SI-traceable method for DNA CRM characterization, which will enable effective comparison of quantified values, quality

control in routine work, and method validation, is required.

Five methods having potential SI-traceability include direct counting [31], digital PCR (dPCR) [32], quantitative nuclear magnetic resonance (qNMR) [33], ICP-optical emission spectrometry (ICP-OES) [4, 34, 35], and LC-IDMS [17-19, 36]. The direct counting method and dPCR address different measurands, such as countable particle and amplifiable targets, respectively, rather than total DNA. qNMR lacks sufficient sensitivity for total DNA measurement. ICP-OES and LC-IDMS are the methods of choice for the accurate quantification of total DNA. ICP-OES is based on the measurement of DNA phosphorus (P) content. LC-IDMS, on the other hand, measures DNA compounds, such as dNMPs or dNs, after the enzymatic digestion of DNA [17-21, 36]. Because LC-IDMS can quantify DNA indirectly through the quantification of dNMPs or dNs, a DNA standard having an accurate value, such as concentration or mass fraction, is not required. Furthermore, LC-IDMS is fully traceable to SI units by using an appropriate single-nucleotide standard.

For the accurate quantification of DNA by LC-IDMS, complete digestion and control of over-decomposition of DNA compounds are required. As regards DNA digestion, two methods have been reported. One is enzymatic digestion, which has already been mentioned earlier, and the other is acid hydrolysis. The former can control the over-decomposition of dNMPs or dNs by optimizing the reaction conditions. However, large DNAs having more than 10,000 bp might not be digested completely, as was shown by Dong et al. [20], who applied enzymatic digestion to λDNA quantification but had to rely on ultrasonication with a special instrument prior to the enzymatic digestion. Acid hydrolysis could hydrolyze DNA into nucleobases completely, but the nucleobases were over-decomposed by hydrochloric acid or formic acid [2, 37-40]. In hydrochloric acid hydrolysis, accurate measurement was difficult due to the influence of over-decomposed compounds. In formic acid hydrolysis, cytosine was deaminated to yield uracil [2]. If the over-decomposition of

nucleobases could be well controlled, acid hydrolysis would be a promising tool for the digestion of DNA, particularly large DNA. Furthermore, uncertainty arising from the acid hydrolysis of DNA would be reduced because the DNA digestion step would be less complicated than the enzymatic quantification with ultrasonication.

A method for the accurate quantification of large DNA was developed, which involves acid hydrolysis and LC-IDMS. For DNA digestion, formic acid was chosen because the reaction with formic acid was milder than that with hydrochloric acid. As sample, we used λDNA, which is the genomic DNA of bacteriophage lambda, and dsDNA having 48,502 bp. By using dNMPs or isotope-labeled dNMPs (LdNMPs) as the calibrator or the internal standard, hydrolyzing dNMPs and LdNMPs into nucleobases, and optimizing the reaction conditions, the accurate quantification of nucleobases was achieved. Phosphorus was quantified in the same λDNA sample by ICP-MS and the results of the developed method were compared with those of ICP-MS.

3.2 Experimental

Water used for preparation and dilution was purified water from Milli-Q Advantage (Millipore, Tokyo, Japan).

3.2.1 λDNA solution

λDNA solution was purchased from Nippon Gene (Tokyo, Japan). The concentration of λDNA solution was roughly estimated to be 670 ng/μL with a UV spectrometer (BioSpec-nano, Shimadzu). λDNA quality was evaluated by gel electrophoresis using 0.3% agarose H (Wako, Osaka, Japan). Because a single band measuring approximately 50,000 bp was observed, λDNA was considered intact. λDNA solution was diluted to 10 ng/μL with water and used as sample.

3.2.2 Reagents and samples

dAMP, dCMP, dGMP, and dTMP were purchased from MP Biomedicals (CA, USA). ^{13}C- and ^{15}N-labeled dAMP, dCMP, and dTMP were purchased from Spectra Stable Isotopes (MD, USA), and ^{13}C- and ^{15}N-labeled dGMP was purchased from ISOTEC International (GA, USA). LC-MS grade acetic acid, LC-MS grade methanol, special grade formic acid, and 3-(trimethylsilyl)1-propanesulfonic acid-d_6 sodium salt (DSS-d_6) reference material were purchased from Wako. Potassium hydrogen phthalate (PHP, NMIJ CRM 3001-b) was from the National Metrology Institute of Japan (NMIJ, Ibaraki, Japan). Ultrapure grade nitric acid, phosphate ion standard solution (JCSS), and cobalt (Co) standard solution (JCSS) were purchased from Kanto Chemical (Tokyo, Japan).

3.2.3 Equipment and measurements

LC-MS/MS

For nucleobase measurement, an LC-10A series HPLC (Shimadzu, Kyoto, Japan) and an LCMS-8030 triple quadrupole mass spectrometer (Shimadzu) were used. Nucleobase separation was carried on a Kinetex XB-C18 (4.6 mm × 250 mm, 5 μm, Phenomenex, CA, USA) at the flow rate 0.5 mL/min. The mobile phase for nucleobase separation was 8% methanol with 0.1% acetic acid in water. MS conditions were electrospray ionization (ESI) and multiple reaction monitoring (MRM) in the positive ion mode. Flow rate of nebulizing gas was 3 L/min, desolvation line (DL) temperature was 250^{o}C, heat block temperature was 400^{o}C, and flow rate of drying gas was 15 L/min. Ions monitored for adenine (A), cytosine (C), guanine (G), thymine (T), ^{13}C- and ^{15}N-labeled A, ^{13}C- and ^{15}N-labeled C, ^{13}C- and ^{15}N-labeled G, and ^{13}C- and ^{15}N-labeled T, and collision energies are shown in Table 3-1.

ICP-MS

For P measurement, ICP-MS (ELEMENT XR, Thermo Fisher Scientific, Tokyo, Japan) was used. Incident Rf power was 1.5 kW and reflected power was <1 W. The flow rates of argon as outer gas, intermediate gas, carrier gas, and make-up gas were 15 L/min, 0.8 L/min, 1.0 L/min, and 0.2 L/min, respectively.

NMR

All ^{1}H NMR spectra were measured with a VNS 600 spectrometer (Agilent Technologies, CA, USA) operating at 599.90 MHz with ^{1}H resonance frequency. A typical set of ^{1}H NMR experimental parameters were as follows: 59523.8 Hz (99.2 ppm) spectral width, 4.0 s acquisition time, 13.0 μs (90°) pulse width, 60 s relaxation delay, and 32 transients acquired. Data processing was performed using MestReNova ver. 9.0.1. All signals were integrated without including ^{13}C satellite signals [41].

3.2.4 DNA quantification

Mass fraction measurement of each dNMP standard solution

Each dNMP standard solution was prepared by dissolving in water and the mass fraction of each dNMP standard solution was determined by qNMR. Briefly, the solvent of weighed dNMP standard solution was removed by evaporation and replaced by DSS-d_6 in D_2O solution, an internal standard solution for qNMR. The mass fraction of DSS-d_6 in D_2O solution was also determined by qNMR using PHP (NMIJ CRM 3001-b, 99.991% ± 0.014% (certified value ± expanded uncertainty)) as reference material [42]. For the evaluation of the mass fraction of DSS-d_6 in D_2O solution, PHP was dissolved in weighed DSS-d_6 in D_2O solution. The mass fractions were estimated from the ^{1}H NMR signals, and the signals that overlapped with impurity signals were corrected on the basis of the amount of impurities determined by HPLC. The estimated mass fractions and the uncertainties of dAMP, dCMP, dGMP, and dTMP solution and DSS-d_6 in D_2O

solution were 7878 μg/g ± 37 μg/g, 8747 μg/g ± 44 μg/g, 9797 μg/g ± 78 μg/g, 8888 μg/g ± 18 μg/g, and 796.5 μg/g ± 0.9 μg/g, respectively. The main uncertainty components of dNMP standard solution as estimated by qNMR were the variation arising from repeatability and the variation arising from the intermolecular signals of dNMP and the mass fraction of DSS-d_6 in D_2O solution.

Preparation of mixed dNMP standard stock solution and LdNMP internal standards

A mixed dNMP standard stock soluton containing around 3.5 μg/g of each dNMP was prepared gravimetrically. This mass fraction corresponded to the number of nucleobases in the λDNA sample assuming complete hydrolysis of the λDNA into nucleobases. The molecular weights of dAMP, dCMP, dGMP, and dTMP were 331.2, 307.2, 347.2, and 322.2, respectively. The LdNMPs were dissolved in 10 mM tris(hydroxymethyl)aminomethane (Tris)-HCl buffer (pH 8.1). A mixed LdNMP stock solution was also prepared at the mass fraction as that of the mixed dNMP stock standard. The mixed LdNMP stock solution was subjected to LC-MS under the conditions indicated in the previous report [10], and the four LdNMPs were detected at the detection limit of 6.3 ng/g. Before IDMS, the standard curve of each dNMP was prepared by using five kinds of calibration blend containing different mass fractions of dNMPs (from 1.4 μg/g to 1.8 μg/g) but the same mass fraction of LdNMP, and the slopes, intercepts, and correlation coefficients of the standard curves are listed in Table 3-2.

Preparation of sample blend and calibration blend solutions

Sample blend and calibration blend were prepared by mixing equal volumes of the LdNMP stock solution and the 10 ng/μL λDNA sample or the mixed dNMP standard stock solution. Briefly, sample blend was a mixture of 400 μL of λDNA sample and 400 μL of mixed LdNMP stock solution. Calibration blend was a mixture of 200 μL of mixed dNMP standard stock solution and 200 μL of mixed LdNMP stock solution, and the mass fraction of dNMPs were around 1.5 μg/g. The

solutions were weighed on a balance.

Acid hydrolysis

Fifty μL of the sample blend or the calibration blend and 200 μL of 88% formic acid were added into a crimp-top vial (Agilent, Tokyo, Japan), and the vial was sealed with aluminum seal and heated. After heating, solvent, including formic acid, was evaporated under nitrogen flow. Then, 250 μL of water was added and the solution was subject to analysis. To determine the optimum conditions for acid hydrolysis, different reaction temperatures (130°C, 150°C, and 170°C) and reaction times (0 h, 2 h, 4 h, 6 h, 8 h, and 24 h) were examined. Acid hydrolysis was conducted six times, and each sample was analyzed six times by LC-MS/MS.

Quantification of nucleobase and calculation of λDNA mass fraction

Equation 2 was used to calculate each nucleobase amount of substance content.

$$C_{\mathrm{nucleobase,i}} = C_{\mathrm{dNMP,i}} \times \frac{R_{\mathrm{sample}}}{R_{\mathrm{calibration}}} \times \frac{MR_{\mathrm{calibration}}}{MR_{\mathrm{sample}}} \times D \qquad (2)$$

Here, $C_{\mathrm{nucleobase}}$ is nucleobase amount of substance content, C_{dNMP} is the amount of substance content of each dNMP in the mixed dNMP stock standard, R_{sample} is the peak area ratio of natural nucleobase and labeled nucleobase in sample blend, $R_{\mathrm{calibration}}$ is the peak area ratio of natural nucleobase and labeled nucleobase in calibration blend, MR_{sample} is the mass ratio of λDNA sample and mixed LdNMP stock solution in sample blend, $MR_{\mathrm{calibration}}$ is the mass ratio of mixed dNMP stock standard and mixed LdNMP stock solution in calibration blend, and D is the acid hydrolysis factor. The index "i" indicates the four nucleobases A, C, G, and T.

λDNA mass fraction was calculated from the amount of substance content of each nucleobase determined by using equation 3.

$$C_{\lambda\mathrm{DNA,i}} = C_{\mathrm{nucleobase,i}} \times \frac{MW_{\lambda\mathrm{DNA}}}{N_{\mathrm{nucleobase,i}}} \qquad (3)$$

Here, $C_{\lambda\mathrm{DNA}}$ is the mass fraction of λDNA, $MW_{\lambda\mathrm{DNA}}$ is the molecular weight of λDNA (3.06×10^7), and $N_{\mathrm{nucleobase,i}}$ is the number of nucleobases of the given type in one λDNA particle. The numbers of nucleobases A, C, G, and T are 24,320, 24,182, 24,182, and 24,320, respectively. The final λDNA mass fraction is the average of the four λDNA mass fraction calculated from the mass fractions of A, C, G, and T by using equation 4.

$$C_{\lambda\mathrm{DNA}} = \frac{\sum C_{\lambda\mathrm{DNA,i}}}{4} \quad (4)$$

Calculation of uncertainty

All uncertainties were calculated in accordance with GUM (Guide to the Expression of Uncertainty in Measurement, ISO) [43]. The uncertainties were calculated by using equation 5 by combining the relative standard uncertainties associated with (1) the mass ratios of sample blends (uMR_{sample}) and calibration blends ($uMR_{\mathrm{calibration}}$), (2) the mass fraction of standard dNMP solution ($uC_{\mathrm{dNMP,i}}$), (3) variation arising from LC-MS/MS measurements (uR_{sample} and $uR_{\mathrm{calibration}}$), and (4) variation arising from acid hydrolysis (uD).

$$u_{C_{\lambda\mathrm{DNA,i}}} = C_{\lambda\mathrm{DNA,i}}\sqrt{\left(\frac{u_{MR_{\mathrm{sample}}}}{MR_{\mathrm{sample}}}\right)^2 + \left(\frac{u_{MR_{\mathrm{calibration}}}}{MR_{\mathrm{calibration}}}\right)^2 + \left(\frac{u_{C_{\mathrm{dNMP,i}}}}{C_{\mathrm{dNMP,i}}}\right)^2 + \left(\frac{uR_{\mathrm{sample}}}{C_{\lambda\mathrm{DNA,i}}}\right)^2 + \left(\frac{uR_{\mathrm{calibaration}}}{C_{\lambda\mathrm{DNA,i}}}\right)^2 + \left(\frac{u_D}{C_{\lambda\mathrm{DNA,i}}}\right)^2} \quad (5)$$

uR_{sample}, $uR_{\mathrm{calibration}}$, and uD were calculated by ANOVA. Because the numbers of nucleobases in one λDNA molecule are unique, the uncertainties were negligible.

The measurement uncertainty of λDNA mass fraction was calculated by using equation 6 by combining the relative standard uncertainties associated with variations arising from each

nucleobase quantification and between nucleobase quantifications (u_{dev}),

$$u_{C_{\lambda DNA}} = \sqrt{\sum \left(\frac{u_{C_{\lambda DNA,i}}}{4}\right)^2 + (u_{dev})^2} \quad (6)$$

Here, u_{dev} is RSD of four $C_{\lambda DNA,\ i}$.

3.2.5 Validation by ICP-MS

For the quantification of P in 10 ng/μL λDNA sample, 100 μL of λDNA sample or P standard solution was introduced into a Teflon tube and this was spiked with approximately 150 ng/g Co internal standard solution. The final mass fractions of P and Co (after digestion and dilution to 5 g) were 10 ng/g and 2 ng/g, respectively. After adding 3 mL of HNO_3 into the digestion tubes, the samples were digested with UltraWAVE (Milestone General, Kanagawa, Japan). The microwave digestion conditions were heating at 230^{o}C for 20 min under 4 MPa. The digested samples were transferred into polypropylene bottles and diluted to 5 g with water. All volumes were corrected gravimetrically.

The mass fraction of λDNA was calculated from the mass fraction of P by using the number of P atoms in one λDNA molecule (97,004) and the molecular weight of λDNA (3.06×10^7). To estimate the uncertainty of ICP-MS, the EURACHEM/CITAC Guide CG 4 [24] was adopted and a spreadsheet approach was used. Because P was measured by ELEMENT XR in the high-resolution mode, the interfering ion that produced potential bias was eliminated from P. Therefore, there was no potential bias and the uncertainty of potential biases was negligible.

3.3 Results and Discussion

3.3.1 Optimization of conditions for nucleobase analysis by LC-MS/MS

In terms of chromatographic performance, the nucleobases have two advantages over dNMPs. The first advantage is that the nucleobases can be analyzed in the ESI positive ion mode

because they are basic compounds. The second advantage is that the nucleobases are more hydrophobic than dNMPs and thus can be analyzed in the presence of a higher concentration of organic solvent than dNMPs. Because a higher concentration of organic solvent leads to a higher sensitivity in the ESI positive ion mode, nucleobases can be analyzed at a higher sensitivity than dNMPs. In this study, the concentration of methanol and the pH of the aqueous mobile phase were optimized to realize good separation and sensitivity for the LC-MS/MS analysis of nucleobases. The nucleobases were not retained on the ODS column when methanol concentration was high, whereas the sensitivity of MS for nucleobases was low when methanol concentration was low. The concentration of methanol was changed from 0% to 10%, and 8% methanol under isocratic conditions was the best choice for the separation of nucleobases. In the ESI positive ion mode, a high sensitivity is usually obtained by using a low pH aqueous mobile phase. By using 0.1% acetic acid solution, S/N of 4 nmol/g A, C, G, and T was 400, 400, 420, and 100, respectively. The S/N of T was the lowest among the nucleobases. A, C, and G but not T contain an amine in their structures, which is likely ionized to a positive ion, whereas T is not likely ionized. However, because sufficient S/N was obtained for all nucleobases from the 10 ng/μL λDNA sample, 8% methanol with 0.1% acetic acid solution was chosen as the optimum mobile phase.

The chromatograms obtained by analyzing about 4 nmol/g nucleobase mixture and acid-hydrolyzed λDNA sample under the optimum conditions are shown in Fig. 3-1. A and G could not be separated completely, but those compounds could be separated by MS because they have different *m/z* values.

3.3.2 Optimization of conditions for acid hydrolysis

It has been reported that DNA is not completely hydrolyzed into nucleobases after a short reaction time, and the nucleobases are decomposed after a long reaction time [37-39]. In this study,

we optimized the conditions for formic acid hydrolysis, focusing on the reaction temperature and the reaction time for the complete digestion of λDNA and the control of the over-decomposition of nucleobases.

λDNA was hydrolyzed with formic acid for 90 min at 130°C, 150°C or 170°C. The amounts of nucleobases generated from the λDNA sample hydrolyzed at 130°C for 90 min were the smallest. When λDNA was hydrolyzed at 170°C from 90 min, uracil, which was formed by the deamination of C as reported by Fisher and Giese [40], and other compounds were detected. The amounts of nucleobases generated when λDNA was hydrolyzed at 150°C for 90 min were larger than those generated when the hydrolysis was carried out at 130°C, but the amounts of the other compounds were smaller than those generated after the hydrolysis at 170°C. From the results, 150°C was chosen as the optimum reaction temperature for formic acid hydrolysis.

To determine the optimum reaction time for the acid hydrolysis at 150°C, the reaction time was changed from 0 h to 24 h, and the peak areas of nucleobases analyzed by LC-MS/MS are shown in Fig. 3-2. The peak areas of A and G reached a maximum at 2 h and those of C and T, at 6 h. When the reaction time exceeded 8 h, the peak areas of A, C, and G decreased, but that of T did not. The results suggested that each nucleobase had a unique optimum reaction time at 150°C, and A, C, and G decomposed after more than 8 h. The amount of uracil, which is produced by the deamination of C, increased when the reaction time was longer than 2 h (data not shown). On the other hand, the relative concentrations of the nucleobases as determined by IDMS are shown in Fig. 3-3. The relative concentrations of A and G remained unchanged even after heating for more than 2 h and those of C and T became constant after heating for more than 4 h and 6 h, respectively. As a result, the relative concentration of each nucleobase remained constant even after heating for more than 6 h. By using stable isotope labeled dNMPs as the internal standard, the over-decomposition of A, C, and G was accurately corrected. In particular, the deamination of C was effectively controlled by the

internal standard because the amount of uracil increased when the reaction time exceeded 2 h but the concentration of C was constant even after 4 h. For the accurate quantification of λDNA, λDNA should be hydrolyzed into nucleobases completely. Because the concentrations of nucleobases in the sample blend at 8 h were equal to those of nucleobases corresponding to the complete hydrolysis of λDNA sample, all the nucleobases were quantitatively liberated from the DNA. Therefore, 8 h was chosen as the optimum reaction time.

3.3.3 Quantification of λDNA

Quantification of nucleobase

By using the developed LC-IDMS method, the amounts of substance contents of nucleobases generated from λDNA by acid hydrolysis at 150^{o}C for 8 h were calculated by using equation 1, and the results are shown in Table 3-3(a). The amounts of substance contents of nucleobases in the hydrolyzed λDNA sample were from 8.73 nmol/g to 9.05 nmol/g according to the number of nucleobases in a λDNA molecule. The uncertainty was estimated according to GUM [43]. The estimated uncertainties of the nucleobase quantification are listed in Table 3-4. The major uncertainty component in the quantification of A was the variation in acid hydrolysis; that in the quantification of C and T was the measurement precision of LC-MS/MS; and that in the quantification of G was the standard solution. Because S/N in the measurement of T, labeled C, and labeled T by LC-MS/MS was equal to or lower than 100, measurement variation was large in the quantification of C and T. In addition, because the uncertainty of acid hydrolysis was not detected and the uncertainty of measurement by LC-MS/MS was small in the quantification of G, the standard solution was the major uncertainty component. This, however, was not a problem as the combined standard uncertainty of G was the second smallest of the nucleobases. From the results, the nucleobases were quantified with approximately 95% expanded uncertainties of 1.7% to 2.6%.

Calculation of λDNA solution

The mass fractions of λDNA were calculated from the values of each nucleobase by using equation 2, and the results are shown in Table 3-3(b). The results were consistent within the range of expanded uncertainties (k = 2). In this method, the estimates for the individual nucleobases were fully independent event, because the variation arising from the LC-MS/MS measurement was a major uncertainty component. Therefore, by using equations 3 to 5, the mass fraction and the measurement uncertainty of λDNA sample were calculated. u_{dev} was 1.4% and was smaller than the expanded uncertainty of nucleobase quantification. The calculated mass fraction of λDNA sample quantified by this method was 11.2 μg/g ± 0.2 μg/g (mass fraction ± combined standard uncertainty). This uncertainty included all uncertainties related to λDNA quantification. Therefore, the present method can quantify λDNA with 1.5% measurement uncertainty using 300 ng of λDNA. This sample amount for quantification was 1/7 and the uncertainty was smaller than that of Dong et al.'s report [20]. The reason why a minimum amount of sample could be used was that the nucleobases could be measured with high sensitivity in the ESI positive ion mode. Dong et al. quantified λDNA by measuring dNMP [20]. Because dNMP contains a phosphate group in its structure, high-sensitivity analysis was difficult using conventional columns [36]. In the present research, nucleobases were used as analytes and high-sensitivity measurement was achieved in the ESI positive ion mode, which was why the amount of λDNA for quantification could be reduced. Another reason why the uncertainty of the quantification could be reduced was that the uncertainty arising from formic acid hydrolysis was small because formic acid hydrolysis is a simple DNA digestion method.

Together, the results confirmed that formic acid hydrolysis with LC-IDMS could be used to quantify λDNA precisely and was traceable to SI because λDNA could be quantified by

LC-IDMS with dNMP standard solution, which could be evaluated appropriately by ^{1}H NMR measurement.

3.3.4 Comparison with results of ICP-MS

To validate the developed formic acid hydrolysis with LC-IDMS, the mass fraction of λDNA obtained with the developed method was compared with that of P-base quantification obtained with ICP-MS. The mass fraction of total P and the mass fraction of λDNA calculated from total P were 1.11 μg/g ± 0.11 μg/g and 11.1 μg/g ± 1.1 μg/g (mass fraction ± expanded uncertainty (k = 2)), respectively. From the results, the mass fraction of λDNA obtained with the developed method was consistent with the results of ICP-MS within the range of expanded uncertainty.

The uncertainty of measurement bias was evaluated from the difference between the value quantified by this method and that by ICP-MS. The relative standard uncertainty was calculated by dividing the difference between those two values by √3 as 1.43%. The final uncertainty was combined both the uncertainty arising from measurement and measurement bias and the final mass fraction of λDNA was 11.2 μg/g ± 0.4 μg/g (mass fraction ± expanded uncertainty (k = 2.2)). Because the effective degree of freedom of the total λDNA quantification was 11 as calculated by using the Welch-Satterthwaite equation, the expansion factor was k = 2.2 in this study. Therefore, the present method can quantify λDNA with 3.4% expanded uncertainty using 300 ng of λDNA.

3.4 Conclusions

A new method was developed for the quantification of λDNA, which involved formic acid hydrolysis with LC-IDMS. The present method could quantify λDNA accurately and was traceable to SI by using as small an amount of sample as 300 ng. The uncertainty of the λDNA quantification was 3.4% of the expanded uncertainty (k = 2.2), and this value was smaller than that obtained by the

previous quantification method. Clearly, formic acid hydrolysis is a simple yet powerful pretreatment tool for DNA digestion. The present method could be applied to low-concentration DNA samples and the quantification of DNA CRM, which would support all types of DNA quantification.

Table 3-1 The *m/z* and collision energy of nucleobases and labeled nucleobases by LC/MS/MS.

Aanalyte	*m/z*	Collision energy (eV)
A ^{13}C- and ^{15}N-labbeled A	136.00 →119.15 146.0 → 128.0	-28
C ^{13}C- and ^{15}N-labbeled C	112.0 → 40.0 119.0 → 54.05	-37
G ^{13}C- and ^{15}N-labbeled G	152.0 → 110.0 162.0 → 117.05	-22
T ^{13}C- and ^{15}N-labbeled T	127.0 → 54.05 134.0 → 58.0	-27

Table 3-2 The slopes, intercepts and correlation coefficients of standard curves of nucleobases after formic acid hydrolysis of 5 kinds of calibration blend.

	Slope	Intercept	Correlation coefficient
A	0.922	0.075	0.9910
C	0.764	0.013	0.9924
G	0.967	0.010	0.9985
T	0.929	0.037	0.9932

Table 3-3 The results of nucleobases quantification and λDNA quantification by using developed IDMS method and ICP-MS.

(a)	Nucleobases		(b)	λDNA	
	Amount of substance content (nmol/g)	Expanded uncertainty (nmol/g, *k*=2)		Amount of substance content (pmol/g)	Expanded uncertainty (pmol/g)
A	9.05	0.23		0.373	0.010
C	8.73	0.21		0.361	0.006
G	8.89	0.17		0.368	0.007
T	8.83	0.21		0.363	0.009
Final result of LC-IDMS	-	-		0.366	0.017
ICP-MS	-	-		0.357	0.037

Table 3-4 Uncertainty components of nucleobases quantification. Each value is shown as relative uncertainty (%)

	Relative uncertainty (%)			
	A	C	G	T
Acid hydrolysis (uD)	1.162	0.000	0.000	0.973
Measurement of calibration blend ($uR_{calibaration}$)	0.216	0.675	0.382	0.612
Measurement of sample blend (uR_{sample})	0.163	0.256	0.155	0.255
Standard solution (uC_{dNMP})	0.470	0.504	0.847	0.213
Mass ratio of calibration blend ($uMR_{calibration}$)	0.004	0.004	0.004	0.004
Mass ratio of sample blend (uMR_{sample})	0.004	0.004	0.004	0.004
Combined standard uncertainty	1.28	0.88	0.94	1.20
Expanded uncertainty (k = 2)	2.57	1.76	1.88	2.39

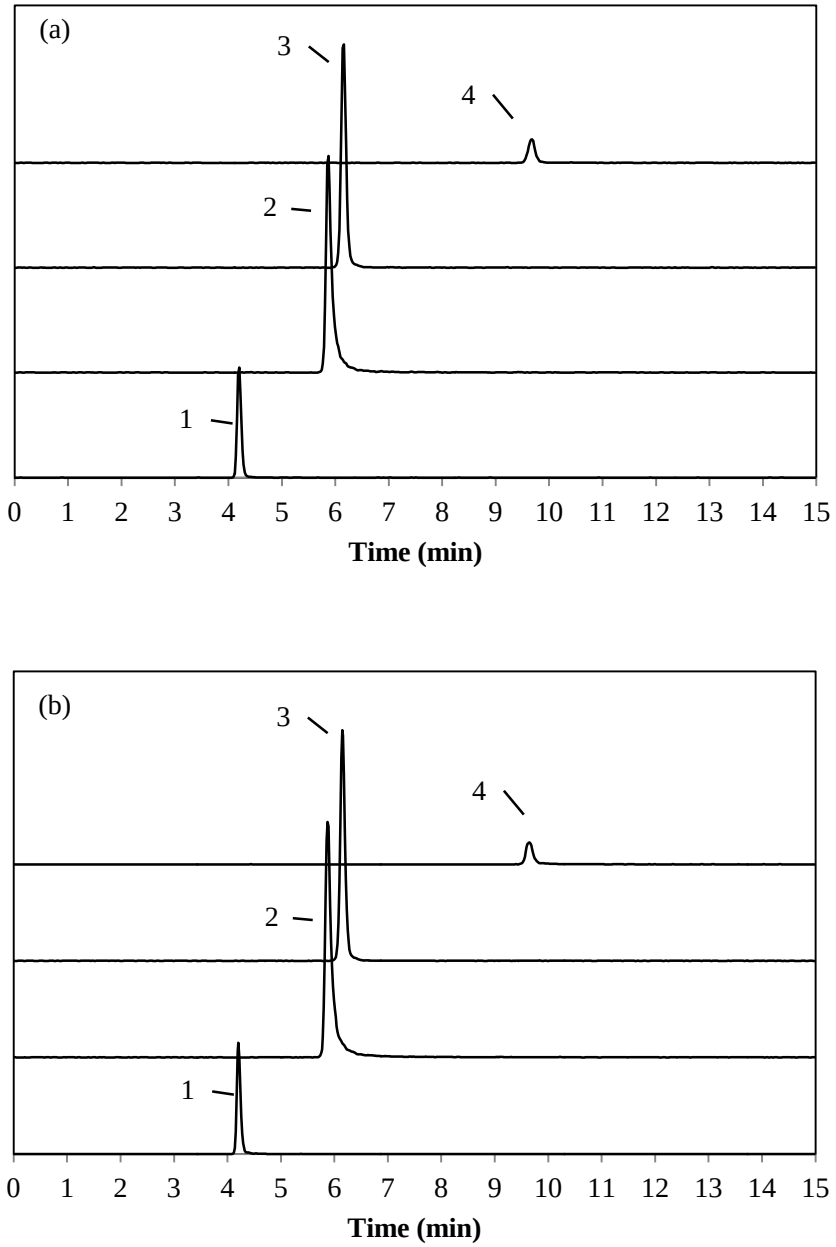

Fig. 3-1 The LC-MS chromatogram of (a) 4 nmol/g nucleobases mixture and (b) acid hydrolized λDNA sample. 1; Cytosine (*m/z*:112.0→40.0), 2; Adenine(*m/z*:136.0→119.15), 3; uanine(*m/z*:152.0→110.0), 4; Thymine(*m/z*:127.0→54.05).

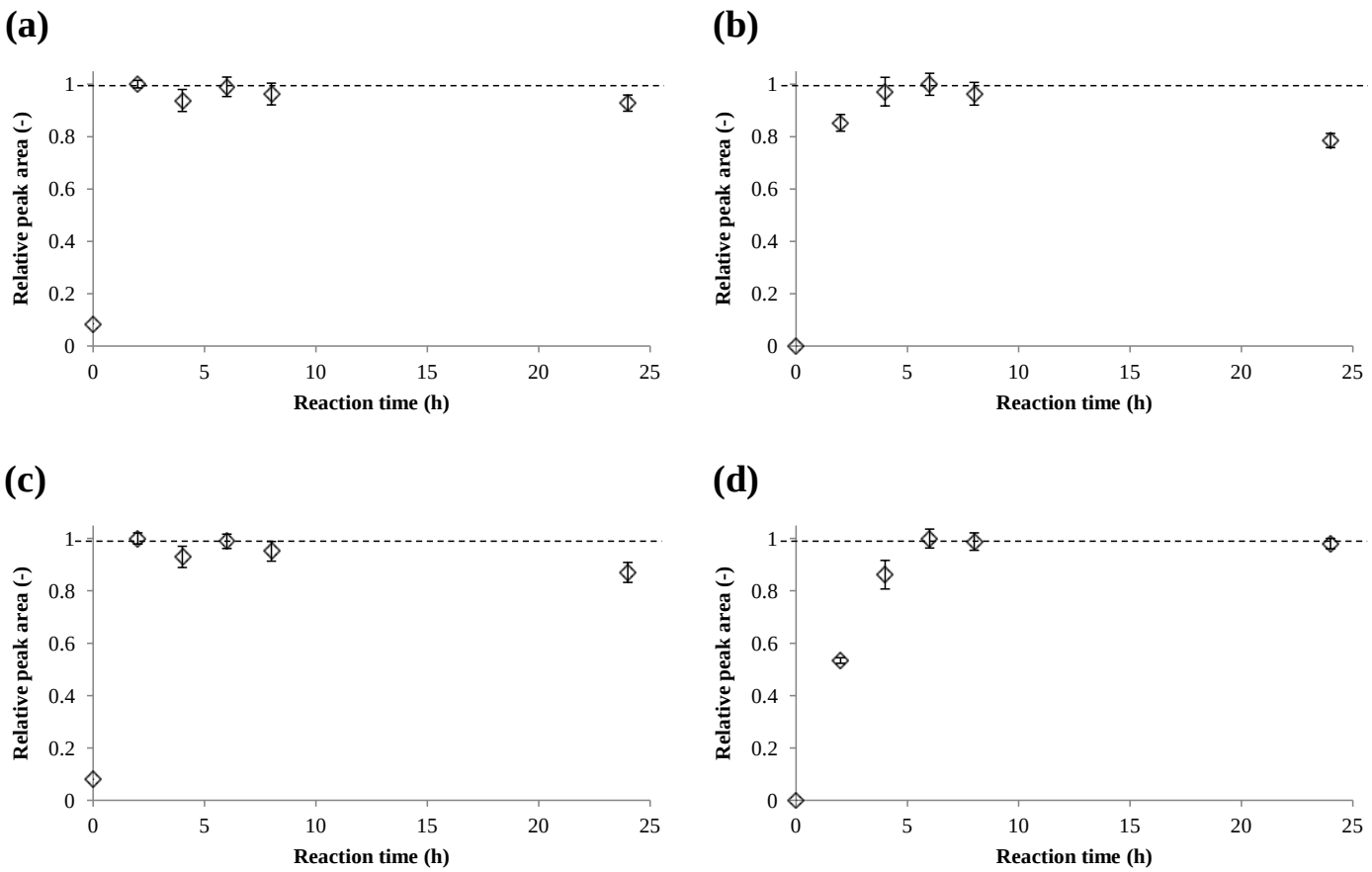

Fig. 3-2 The average and standard deviation of relative peak area (n = 4) of (a) adenine, (b) cytosine, (c) guanine and (d) thymine when the reaction time of acid hydrolysis was changed.

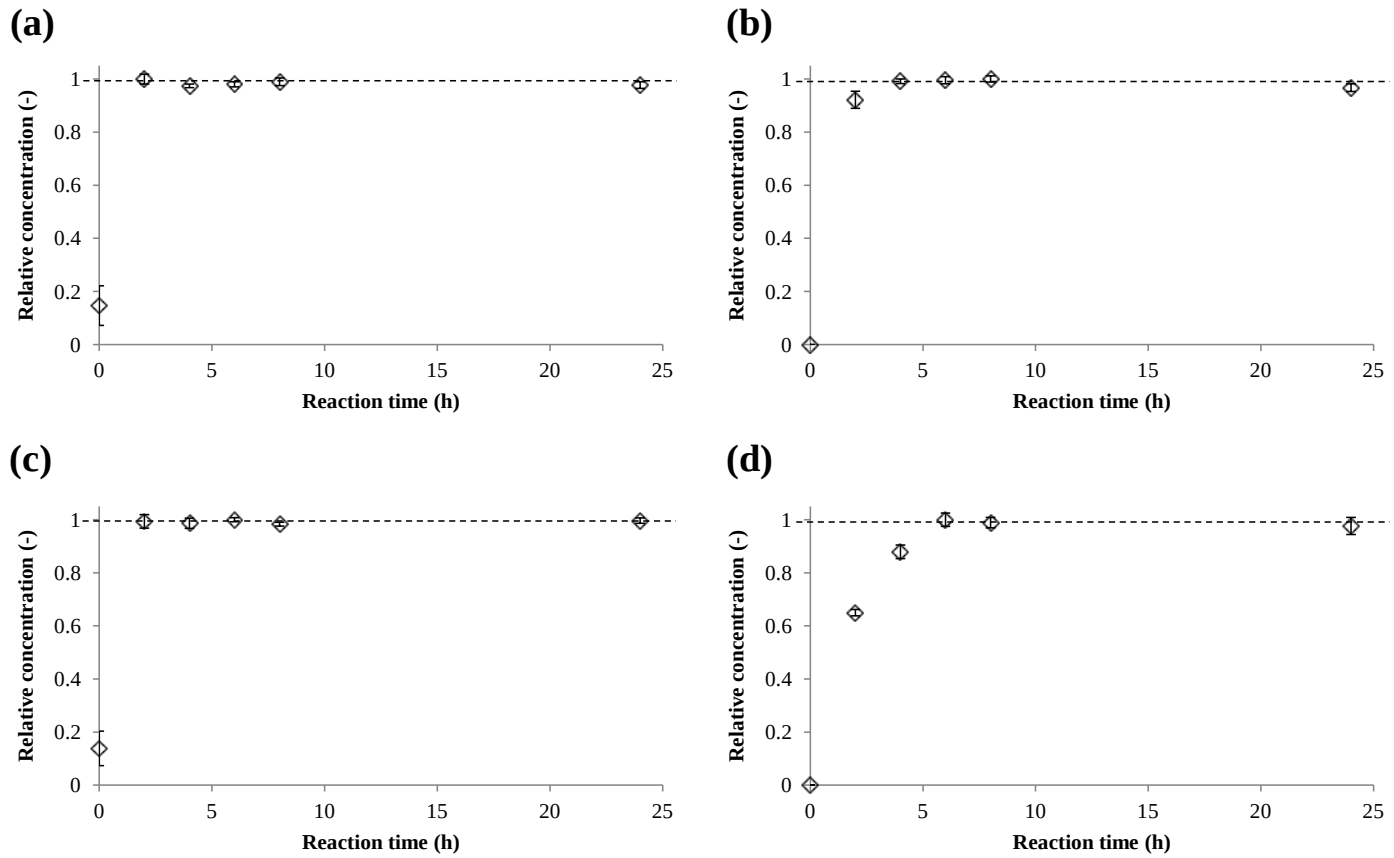

Fig. 3-3 The relative concentration and relative standard uncertainty of (a) adenine, (b) cytosine, (c) guanine and (d) thymine obtained from the same data of Fig. 3-2 when the reaction time of acid hydrolysis was changed.

Chapter 4 HPLC for separation and quantification of DNA fragments and measurement of DNA degradation

4.1 Introduction

DNA analysis and quantification is rapidly gaining importance because of the emergence of pharmaceutical products whose main components are nucleic acids, such as DNA vaccines and nucleic acid drugs. DNA quantification is also essential in detecting GMOs or microorganisms that cause disease or food poisoning. In the case of plasmid DNA, which is essential to gene therapy, it is necessary to check that the plasmid DNA in vaccines or drugs has higher than 90% purity [12]. A method that separates and quantifies original plasmid DNA and its degradation product, which has the same sequence as the original plasmid DNA, is necessary. qPCR, which analyzes double-stranded DNA, is commonly used for the detection and quantification of GMOs. However, accurate GMO quantification by qPCR is sometimes very difficult because DNA in processed food is often fragmented [1, 44, 45]. To quantify GMOs accurately in order to meet the legal requirements for GMO labeling by country, it is necessary to evaluate the degradation or size of the extracted DNA prior to qPCR. Therefore, for both plasmid DNA and double-stranded DNA, an analytical method that can separate and quantify DNA and its degradation products is necessary, so that the DNA degradation rate can be determined quantitatively and accurately.

To quantify DNA, UV absorbance or fluorescence measurement with a DNA-specific dye is performed. However, as those methods do not include a separation step, it is difficult to quantify the target DNA in a mixture of several DNAs. On the other hand, agarose gel electrophoresis, microchip electrophoresis, CE, and microcapillary hydrodynamic chromatography are used to analyze specific DNA [7, 46]. Because of their ability to separate DNA by length (molecular

weight), those techniques are used to confirm DNA of a particular length. Agarose gel electrophoresis is the classic method for DNA analysis. It is used widely because no expensive equipment is needed. However, agarose gel electrophoresis is a qualitative technique. Microchip electrophoresis and CE offer many benefits, including high resolution with a very small amount of sample, short analysis time, low cost, and low risk of contamination, but the peak area varies from 20% to 30% [7]. Therefore, the development of a high-accuracy method for DNA is required.

HPLC is widely used for the separation and quantification of compounds with high accuracy. Because the molecular weight of an oligonucleotide is larger than 1,000, few applications of HPLC for DNA analysis have been reported. In those reports, size-exclusion [8, 9], ion-pair reversed-phase [10, 11], and anion-exchange [10, 12-16] columns were used. DNA was separated according to molecular weight by using a size-exclusion column, but the resolution was poor compared with an anion-exchange column or an ion-pair reversed-phase column. HPLC with an ion-pair reversed-phase column is applicable to only small DNAs, because the increase of hydrophobicity due to the ion-pair reagent has a relatively small effect on large DNA molecules. An anion-exchange column has been used to separate plasmid DNA of different conformations [12-14] and mixed DNA samples in a molecular weight marker [10, 15, 16]. In a report on the quality control of plasmid DNA by Quaak et al. [13], plasmid DNAs having different conformations were well separated and good quantitative accuracy of 1.6% was obtained. In the quantitative analysis of double-stranded DNA by Burke et al. [15], DNA fragments in a DNA molecular weight marker were quantified by digital PCR, and another double-stranded DNA was quantified using the DNA molecular weight marker as standard. The uncertainty of this method was approximately 1% for a sample with approximately 10 ng/μL concentration.

In this study, a method by which both the main component and the degradation products of double-stranded DNA could be separated and quantitatively evaluated was examined and established.

To analyze DNA fragments ranging from a few bp to more than 10000 bp, a mixed double-stranded DNA sample consisting of 26 bp to 23130 bp DNA fragments was used. The separation conditions were examined in detail and the optimum separation conditions for both short and long DNA fragments were determined. LOD, LOQ, and measurement repeatability were evaluated by using DNA CRM, and the established method was used to evaluate the degradation rate of a 600 bp DNA fragment.

4.2 Experimental

4.2.1 Reagents

NMIJ CRM 6203-a (National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan) was used as the sample and calibrator. NMIJ CRM 6203-a is a set of double-stranded DNA solutions of 600 bp with four different sequences. In this study, D001-600-G, which is one of the four DNA solutions, was used. The certified total DNA concentration in D001-600-G was 12.4 ng/μL ± 1.1 ng/μL. DNA molecular weight markers, pUC19/Msp I digest (Wako Pure Chemical, Osaka, Japan) and λ/Hind III digest (Wako Pure Chemical, Osaka, Japan), were used as samples. The lengths of the DNA fragments in the DNA molecular weight markers are shown in Table 4-1. Tris(hydroxymethyl)aminomethane (Tris) was purchased from Sigma-Aldrich (Tokyo, Japan). Sodium chloride, hydrochloric acid, and sodium benzoate were purchased from Wako Pure Chemical (Osaka, Japan). The molecular weight marker ϕX174/Hae III digest (Life Technologies, Tokyo, Japan) was diluted 50-fold with ultrapure water and used for gel electrophoresis.

4.2.2 Sample preparation

Three types of mixed DNA samples were used to investigate the analytical conditions.

Each sample was prepared so that the concentration of each DNA fragment was almost equal to 1 ng/μL.

Sample A: D001-600-G and pUC19/Msp I were mixed in a 4 : 1 ratio.

Sample B: D001-600-G and λ/Hind III were mixed in a 5 : 1 ratio.

Sample C: D001-600-G, pUC19/Msp I, and λ/Hind III were mixed in a 20 : 5 : 4 ratio.

Sample C-blk: Sample C with ultrapure water instead of D001-600-G.

To evaluate DNA degradation rate, Sample D was used.

Sample D: approximately 10 ng/μL of 600 bp DNA solution, which was the PCR amplified product of D001-600-G.

D001-600-G calibration solutions with concentrations ranging from approximately 0.05 ng/μL to 12.4 ng/μL were prepared. Ultrapure water purified with Milli-Q Advantage (Millipore) was used to dilute all the samples. All the samples were prepared gravimetrically.

4.2.3 Equipment and measurements

Chromatography

The HPLC system was an LC-10A series (Shimadzu, Kyoto, Japan) with a UV detector. Separation was achieved with a YMC-BioPro QA-F column (4.6 mm i.d., x 100 mm, particle size 5 μm, YMC Co., Ltd., Kyoto, Japan). The mobile phase consisted of two buffers: buffer A, 20 mM Tris-HCl (pH 8.1) containing 0.5 M NaCl, and buffer B, 20 mM Tris-HCl (pH 8.1) containing 1.0 M NaCl. The isocratic mode, in which sodium chloride concentration was between 0.5 M and 1.0 M, or the gradient mode, in which sodium chloride concentration was changed from 0.7 M to 0.9 M, was used. Flow rate was 0.5 mL/min and injection volume was 10 μL. Column bath temperature was 25^{o}C and measurement wavelength was 254 nm.

Agarose gel electrophoresis

Sample separation was conducted at 50 V for 70 min with TAE buffer in agarose gel. A 5 μL volume of each sample was loaded on 0.5% Agarose S gel (for long DNA, Wako Pure Chemical, Osaka, Japan) or 3.0% Agarose 21 gel (for short DNA, Wako Pure Chemical, Osaka, Japan). After electrophoresis, the gel was stained with SYBR Gold (Life Technologies, Tokyo, Japan) diluted 10,000 times with TAE buffer for 15 min and photographed by ChemDoc (Bio-Rad, Tokyo, Japan). Fluorescence intensities were compared visually.

4.2.4 Identification of DNA peak in mixed DNA sample

After separating the DNA fragments in the molecular weight marker by agarose gel electrophoresis, the gel was cut off and purified with a QIAquick Gel Extraction Kit (Qiagen, Tokyo, Japan).

4.2.5 Evaluation of DNA degradation

I pipetted 250 μL of Sample D into each sample tube and the tubes were stored at -20^{o}C or 40^{o}C. A sample tube was taken out periodically and subjected to HPLC-UV and agarose gel electrophoresis. In the HPLC-UV chromatogram, the peak area of the 600 bp DNA fragment was measured. Relative concentration was determined by dividing the peak area of the sample measured after the temperature treatment by that of the sample measured immediately after preparation. At each time, the peak area was corrected with sodium benzoate as the external standard. Agarose gel electrophoresis was performed using 3.0% Agarose 21 gel under the conditions specified in 2.3.2.

4.3 Results and Discussions

4.3.1 Optimization of separation conditions using mixed DNA sample

The separation principle of the anion-exchange column is based on the interaction between the negatively charged groups on the analyte and the positively charged groups on the stationary phase. The elution of the analyte is carried out by changing the ionic strength of the mobile phase, and the analyte that has the smallest charge is eluted first. In DNA analysis, interaction between the negatively charged phosphate groups on the DNA backbone and the positively charged ammonium groups on the stationary phase occurs. As the number of phosphate groups increases in proportion to the molecular weight of DNA, the total charge of a larger DNA increases. The elution conditions were optimized by changing sodium chloride concentration in the mobile phase.

HPLC analyses of Samples A and B in the isocratic mode in which sodium chloride concentration was changed from 0.5 M to 1.0 M in 0.125 M increments were performed. DNA fragments in Sample A were eluted with sodium chloride at concentrations ranging from 0.625 M to 0.75 M, and all the DNA fragments in Sample A were eluted at the same time at concentrations higher than 0.75 M. DNA fragments in Sample B were also eluted with sodium chloride at concentrations ranging from 0.75 M to 0.825 M, and all the fragments in Sample B were eluted at the same time at concentrations higher than 0.825 M. Because the DNA fragments in Sample A and Sample B could not be eluted in the same conditions, it was concluded that HPLC analysis of DNA fragments having a wide range of molecular weights was difficult in the isocratic mode. Then, the gradient condition in which sodium chloride concentration was changed from 0.7 M to 0.9 M was examined. In order to optimize the gradient condition, the separation of each DNA fragment was evaluated by changing gradient steepness. Resolution was calculated from equation 7 and used as an indicator. When two peaks are separated by a baseline, the resolution is higher than 1.5.

$$\text{Resolution} = 1.176\left(\frac{t_B - t_A}{w_{hA} + w_{hB}}\right) \qquad (7)$$

Here, t_A is the retention time of analyte A, t_B is the retention time of analyte B, w_{hA} is the peak width at half height of analyte A, and w_{hB} is the peak width at half height of analyte B. Fig. 4-1 shows the dependence of resolution on gradient steepness for several pairs of DNA fragments. The steeper the gradient is, the lower the resolution is for DNA fragments equal to or shorter than 2,027 bp. On the other hand, for DNA fragments equal to or longer than 6,557 bp, the resolution increases in the gradient steepness range of 3.3 mM/min to 6.7 mM/min. Therefore, because both short and long DNA fragments can be separated well at the gradient of 6.7 mM/min, this gradient was chosen as the optimum condition. Fig. 4-2 shows the chromatogram of DNA fragments ranging from 67 bp to 23,130 bp using this optimum gradient method.

The order of elution of DNA fragments in Sample C in the optimum gradient condition was examined in detail and the result is shown in Fig. 4-2. Using an anion-exchange column, DNA fragments were generally eluted in the order of small molecular weight to large molecular weight. However, two DNA fragments were eluted in reverse order (404 bp and 489 bp). Huber reported that the elution of DNA that had high adenine (A) and thymine (T) contents in its sequence was delayed [10]. Our result is consistent with Huber's because the AT contents of 404 bp and 489 bp DNA were 56% and 44%, respectively.

In this study, several DNA fragments, such as 242 bp and 331 bp DNA fragments, 564 bp and 600 bp DNA fragments, and 2,027 bp, 2,322 bp, 4,361 bp, and 6,557 bp DNA fragments, were eluted at the same time. In all cases except the 564 bp and 600 bp DNA fragments, not baseline separation but peak tip separation was achieved by changing gradient steepness. For example, Fig. 4-3 shows the chromatogram of sample C analyzing under 20 mM/min of gradient sharpness. In the peak including 2,027 bp, 2,322 bp, 4,361 bp and 6,557 bp, the peak tip separation was occurred. On the other hand, as the degradation products would be shorter than the original DNA but have a similar sequence, it is expected that the degradation products would be eluted before the original

DNA.

4.3.2 Validation of analytical method

Linearity, LOD, LOQ

The calibration curve from 0.05 ng/μL to 12.4 ng/μL of D001-600-G was obtained with high linearity and the correlation coefficient was 0.999. LOD was 0.02 ng/μL and LOQ was 0.06 ng/μL for S/N = 3 or S/N =10, respectively. Currently, LOD for DNA analysis using an absorption spectrometer is approximately 2 ng/μL and LOD for the fluorescence method is approximately 0.1 ng/μL. LOD of our method is 20 to 30 times higher than those of the above-mentioned methods.

Evaluation of peak area repeatability

Measurements were performed in triplicate for each concentration in the calibration curve and peak area repeatability was evaluated. RSD at the concentration of 0.56 ng/μL, which was approximately 100 times higher than LOD, was 1.8%, and RSDs at 1.21 ng/μL, 6.26 ng/μL, and 12.4 ng/μL were 1.3%, 1.0%, and 1.4%, respectively. Peak area repeatability of DNA in the mixed DNA sample was also evaluated. The RSDs of 190 bp, 600 bp, and 9416 bp DNA fragments in Sample C were 1.7%, 1.0%, and 0.9%, respectively. Therefore, DNA in both single and mixed DNA samples can be analyzed with high repeatability by the present method.

Evaluation of recovery in mixed DNA sample

Sample C was measured in triplicate and the 600 bp DNA fragment in the sample was quantified. Recovery was calculated by comparing the measured value with the quantitative value. As 564 bp and 600 bp DNA fragments were eluted at the same time, the 564 bp fragment in Sample C-blk was quantified. The quantitative value of the 600 bp fragment was 1.20 ng/μL ± 0.04 ng/μL (quantitative value ± standard deviation) and the preparative value was 1.21 ng/μL ± 0.03 ng/μL.

Therefore, the recovery was 99.1% ± 3.6%.

4.3.3 Evaluation of DNA degradation

Then, this analytical method was applied to the measurement of DNA degradation. Fig. 4-4 shows the DNA degradation rates obtained by this method.

The relative concentrations of the samples stored at -20°C were constant during storage and no DNA degradation was observed. On the other hand, in the samples stored at 40°C, DNA degradation was observed and the degradation rates were 13.5% ± 0.9% after one-day storage, 41.8% ± 1.1% after one-week storage, and 67.0% ± 0.9% after two-week storage. A comparison of the chromatograms of the samples stored at -20°C and 40°C revealed that the peaks of components that were not retained on the anion-exchange column increased in number in the sample stored at 40°C. Such small molecules as deoxynucleotides, which are DNA monomers, might be generated when the 600 bp DNA fragment was degraded. In addition, high baseline noise was observed prior to the appearance of the peak of the 600 bp DNA fragment. This suggested that several small DNA fragments could be produced at concentrations lower than LOD.

Agarose gel electrophoresis, which is generally used to evaluate DNA degradation, was also performed. There was no change in the fluorescence intensity of the 600 bp DNA fragment, and no smear-like DNA was observed in the samples stored at -20°C. In contrast, in the samples stored at 40°C for one or two weeks, smear-like DNA was observed at the sides of both lower and higher molecule in the gel. The fluorescence intensity of the 600 bp DNA fragment in the sample stored for one week did not decrease, whereas that in the sample stored for two weeks decreased slightly (data not shown). Taking the results of both HPLC and agarose gel electrophoresis into consideration, more than 65% of DNA might have decomposed when the fluorescence intensity decreased in agarose gel electrophoresis.

By using HPLC-UV with an anion-exchange column, the DNA degradation rate in approximately 10 ng/μL DNA solution can be evaluated with a variation of approximately 1%. This analytical method can evaluate DNA degradation rate correctly and quantitatively.

4.4 Conclusions

An HPLC-UV method that uses an anion-exchange column to separate and quantify mixed DNA samples was developed. As a result of changing the gradient steepness of salt concentration in the mobile phase, the response to the gradient steepness in separation differed depending on the DNA fragment length. Under the optimum conditions, all DNA fragments ranging from 67 bp to 23,130 bp in the mixed DNA sample could be separated well. The calibration curve from 0.05 ng/μL to 12.4 ng/μL of D001-600-G was obtained with high linearity. LOD was 20 to 30 times higher than those of existing methods. Good measurement reproducibility of 2% or less in all the samples, and satisfactory recovery of spiked DNA fragment in the mixed DNA sample of approximately 1.2 ng/μL concentration were obtained. From these results, this analytical method can separate and quantify DNA fragments in the mixed DNA sample with high accuracy.

As a result of applying this analytical method to the evaluation of 600 bp DNA degradation, DNA degradation could be observed in the sample stored at 40^{o}C, and the degradation rate could be successfully evaluated with the precision of approximately 1%.

Taken together, the results indicate the usefulness of the HPLC-UV with anion-exchange column method for the quantitative determination of DNA degradation, which is necessary for quality control.

Table 4-1 Lengths and adenine (A) and thymine (T) contents of DNA fragments in mixed DNA sample.

	Length of DNA fragment (bp)	AT content (%)
D001-600-G	600	50
pUC19/*Msp* I	26	50
	34A	50
	34B	41
	67	46
	110	50
	111	50
	147	39
	190	50
	242	51
	331	42
	404	56
	486	44
	501	56
λ/*Hind* III	526	58
	2,027	64
	2,322	63
	4,361	55
	6,557	52
	9,416	55
	23,130	44

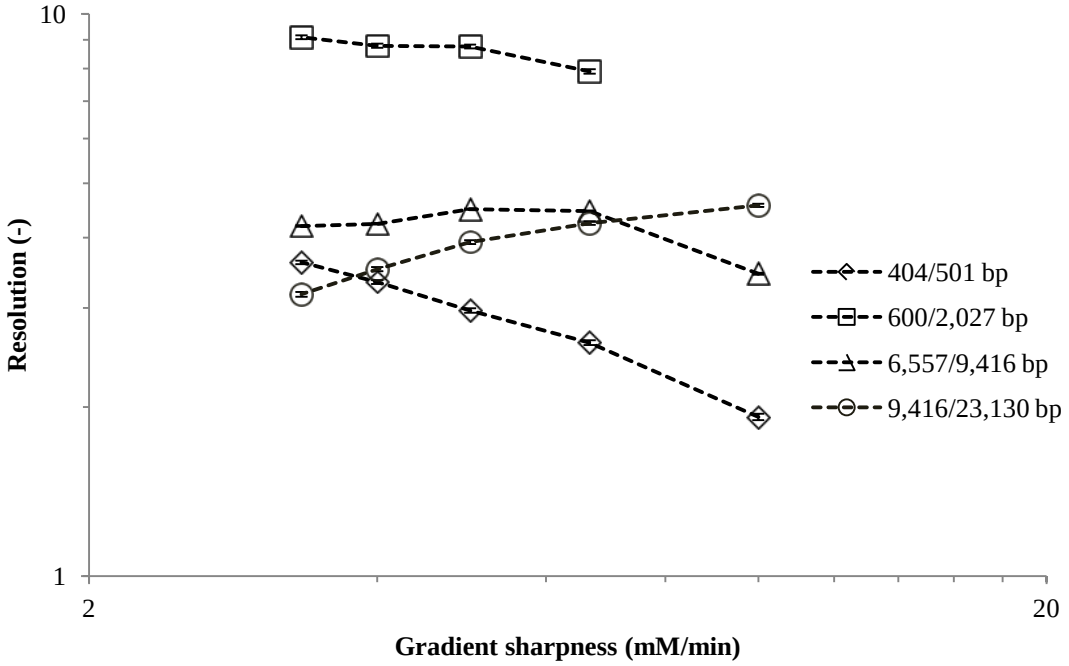

Fig. 4-1 Dependence of resolution on gradient steepness for several pairs of DNA fragments. ◇; resolution of 404 bp and 501 bp, □; resolution of 600 bp and 2,027 bp, △; resolution of 6,557 bp and 9,416 bp, and ○; resolution of 9,416 bp and 23,130 bp.

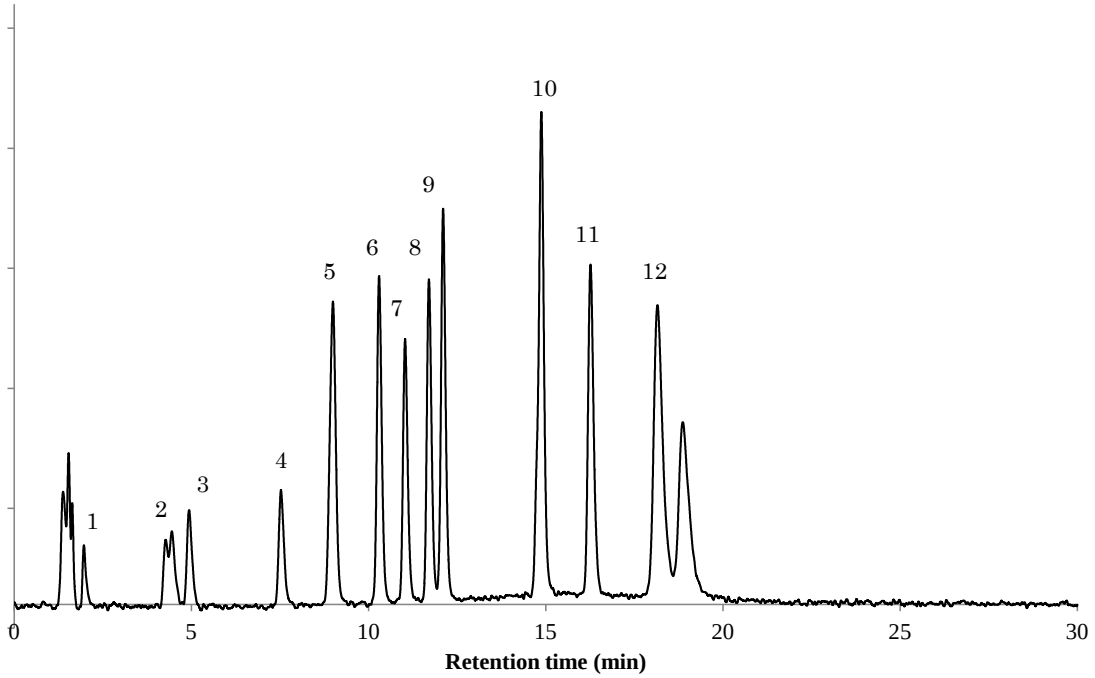

Fig. 4-2 The HPLC chromatogram of sample C analyzing under the optical condition. 1: 67 bp, 2: 110 bp, 111 bp, 3: 147 bp, 4: 190 bp, 5: 242 bp, 331 bp, 6: 489 bp, 7: 404 bp, 8: 501 bp, 9: 564 bp, 600 bp (CRM), 10: 2,027 bp, 2,322 bp, 4,361 bp, 6,557 bp, 11: 9,416 bp, 12: 23,130 bp.

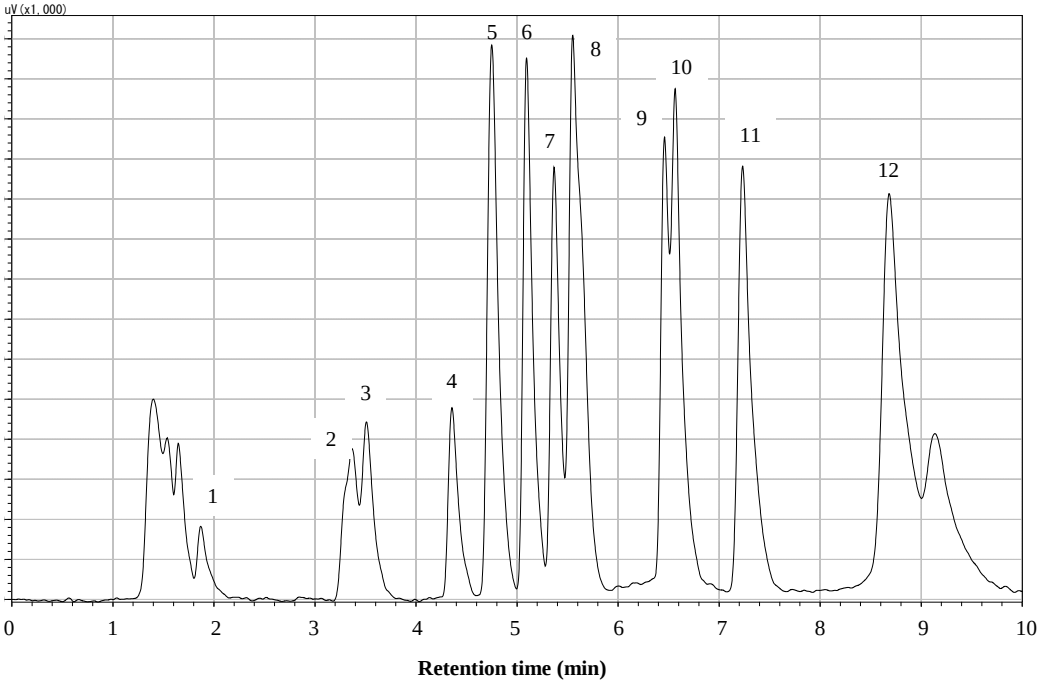

Fig. 4-3 The HPLC chromatogram of sample C analyzing under 20 mM/min of gradient sharpness. 1: 67 bp, 2: 110 bp, 111 bp, 3: 147 bp, 4: 190 bp, 5: 242 bp, 331 bp, 6: 489 bp, 7: 404 bp, 8: 501 bp, 564 bp, 600 bp (CRM), 9, 10: 2,027 bp, 2,322 bp, 4,361 bp, 6,557 bp, 11: 9,416 bp, 12: 23,130 bp.

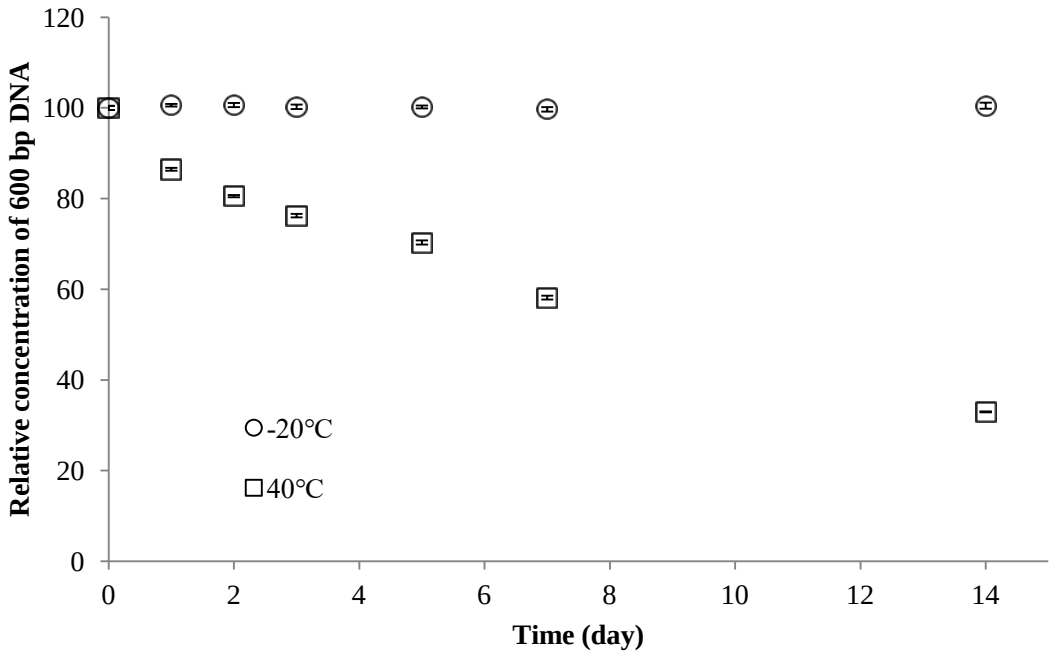

Fig. 4-4 The relative concentration and measurement variation of 600 bp DNA at -20°C and 40°C. ○: -20°C, □: 40°C.

Chapter 5 Conclusions and perspectives

5.1 Conclusions

As discussed in Chapters 2 and 3, DNA mass fraction can be determined with SI traceability and high accuracy by quantifying DNA monomers using LC-IDMS. There are two methods for DNA degradation: one is enzymatic digestion and the other is acid hydrolysis. Because those methods do not require a DNA calibrator, which has the same sequence as a DNA analyte, for DNA quantification, those methods can be applied to the characterization of DNA CRM, which will support all types of DNA quantification. On the other hand, because those methods tend to degrade DNA into dNMPs or nucleobases, the DNA sequence disappears and the obtained value is the total DNA mass fraction. To obtain the mass fraction of a particular DNA, the purification of that DNA or its separation from other DNAs in the sample is required when other DNAs are present in the sample. To this end, I have developed a HPLC-based separation method for DNA fragments, which is dependent on DNA length (molecular weight) (Chapter 4). By using HPLC with an anion-exchange column and optimizing the analytical conditions, DNA could be separated by length, thereby enabling its quantification. Using this method, DNA degradation rate was quantified with approximately 1% variation, and stability and purity studies necessary for DNA CRM development could be carried out. However, for quantification of DNA by this method the DNA calibrator which has the same sequence of DNA analyte was required.

By combining the developed methods, DNA could be quantified with SI traceability and high accuracy. The combination of the developed methods would lead to the development of DNA CRM and the improvement of the reliability of DNA quantification.

5.2 Perspectives

I would like to discuss the perspectives of DNA quantification. Over the course of my work, I could not develop a DNA quantification method coupled with sequence evaluation. DNA sequencing provides vital information about DNA and thus, the quantification of DNA having a particular sequence is very important. A possible method for quantifying DNA with sequencing evaluation is digital PCR (dPCR) [47]. Because dPCR is based on PCR, quantification depending on DNA sequence is possible. A similar method based on PCR is qPCR. However, dPCR differs from qPCR in that dPCR is an end-point measurement that does not require a DNA calibrator. In a typical dPCR experiment, a sample is randomly distributed into discrete partitions such that some contain no nucleic acid template and others contain one or more template copies [48, 49]. Two dPCR platforms are available: chip-based dPCR and droplet dPCR. In the former, dPCR is performed in small volumes and solid partitions. In the latter, dPCR is performed in partitions made up of water-in-oil emulsion droplets. In both chip-based dPCR and droplet dPCR, PCR is performed in small volumes and separate spaces [50]. Therefore, dPCR would be a possible tool for the characterization of DNA CRM, and DNA CRM characterized by dPCR has been reported [22, 49]. Many studies of DNA quantification by dPCR have been reported by National Metrology Institute (NMI) of each country, and the possibility of using dPCR for DNA quantification is rapidly becoming a reality [15, 20, 22, 48, 49, 51-54]. A DNA quantification method using dPCR would be a promising tool for DNA quantification with DNA sequence evaluation.

For DNA sequence evaluation, next-generation sequencing (NGS) has been developed over the last ten years [55, 56]. In 1977, Sanger et al. introduced an automated sequencing method [57]. Considered "first-generation" DNA sequencing technology, automated Sanger sequencing is used routinely for the sequencing of small DNA fragments. NGS is high-throughput DNA sequencing technology that is capable of sequencing large numbers of DNA sequences in a single

reaction and whose cost is lower than that of Sanger sequencing [58, 59]. Because of these features, its clinical application has been actively promoted in the United States [60, 61]. However, to ensure precision of NGS analysis, quality control using DNA CRM is required. Because NGS evaluates DNA sequences, DNA CRM is necessary to be evaluated the sequence itself precisely. NGS would be required in not only medicine but also bioindustry, and quality control of NGS by using DNA CRM would be very important.

Thus, DNA quantification is necessary in bioindustry, and strict quality control of DNA quantification is a must. I hope the SI-traceable high-accuracy DNA quantification methods that I have developed will contribute to the further development of bioindustry.

5.3 Future plan

Finally, I mentioned two future plans. One is the development of accurate DNA quantification method using dPCR. For the evaluation of accuracy of dPCR, DNA samples which have accurate concentration are needed. Because I developed SI-traceable and high accurate DNA quantification method and DNA samples could quantify with high accuracy, the accuracy of dPCR will be able to evaluate. The other is the development of DNA CRM. Because DNA CRM will support all type of DNA quantification and become improve the reliability of DNA quantification, I will develop DNA CRM which will certify the accurate concentration by combining the developed method.

Acknowledgment

Many people contributed to completion of my doctorates.

First, I am deeply grateful to Dr. Akiko Takatsu (National Institute of Advanced Industrial Science and Technology (AIST)) for constant encouragement and long term support. I am also grateful to Dr. Sin-ichiro Fujii (AIST) for helpful advices and constructive discussion. I would like to express my gratitude to Prof. Kouji Nakamura for giving me an opportunity to study at the laboratory and for supervision during the course of my doctor degree. I would like to appreciate Prof. Yingnan Yang, Assoc. Prof. Motoo Utsumi and Assoc. Prof. Kosumi Yamada for helpful advice and comments. I also appreciate Dr. Hiroshi Sakamaki (Chemicals Evaluation and Research Institute, Japan) for their kindness in giving me the metal free column and useful information.

Finally, I thank for my family for their supports and encouragements.

References

[1] T. Yoshimura, H. Kuribara, T. Matsuoka, T. Kodama, M. Iida, T. Watanabe, H. Akiyama, T. Maitani, S. Furui, A. Hino, Applicability of the quantification of genetically modified organisms to foods processed from maize and soy, Journal of Agricultural and Food Chemistry, 53 (2005) 2052-2059.

[2] H. Aygün, Determination of base composition, in: Handbook of Analysis of Oligonucleotides and Related Products, CRC Press, 2011, pp. 439-452.

[3] P.R. Brown, C.S. Robb, S.E. Geldart, Perspectives on analyses of nucleic acid constituents: the basis of genomics, Journal of Chromatography A, 965 (2002) 163-173.

[4] C.A. English, S. Merson, J.T. Keer, Use of elemental analysis to determine comparative performance of established DNA quantification methods, Analytical Chemistry, 78 (2006) 4630-4633.

[5] M.J. Holden, R.J. Haynes, S.A. Rabb, N. Satija, K. Yang, J.R. Blasic, Factors affecting quantification of total DNA by UV spectroscopy and PicoGreen fluorescence, Journal of Agricultural and Food Chemistry, 57 (2009) 7221-7226.

[6] I.M. Mackay, K.E. Arden, A. Nitsche, Real-time PCR in virology, Nucleic Acids Research, 30 (2002) 1292-1305.

[7] N. Thaitrong, H. Kim, R.F. Renzi, M.S. Bartsch, R.J. Meagher, K.D. Patel, Quality control of next-generation sequencing library through an integrative digital microfluidic platform, Electrophoresis, 33 (2012) 3506-3513.

[8] Y. Kato, M. Sasaki, T. Hashimoto, T. Murotsu, S. Fukushige, K. Matsubara, Operational variables in high-performance gel-filtration of DNA fragments and RNAs, Journal of Chromatography, 266 (1983) 341-349.

[9] H. Ellegren, T. Laas, Size-exclusion chromatography of DNA restriction fragment - Fragment length determinations and acomparison with the behavior of proteins in size-exclusion chromatography, Journal of Chromatography, 467 (1989) 217-226.

[10] C.G. Huber, Micropellicular stationary phases for high-performance liquid chromatography of double-stranded DNA, Journal of Chromatography A, 806 (1998) 3-30.

[11] L. Zhang, B. Majeed, L. Lagae, P. Peumans, C. Van Hoof, W. De Malsche, Ion-pair reversed-phase chromatography of short double-stranded deoxyribonucleic acid in silicon micro-pillar array columns: Retention model and applications, Journal of Chromatography A, 1294 (2013) 1-9.

[12] M. Diogo, J. Queiroz, D. Prazeres, Chromatography of plasmid DNA, Journal of Chromatography A, 1069 (2005) 3-22.

[13] S.G.L. Quaak, B. Nuijen, J. Haanen, J.H. Beijnen, Development and validation of an anion-exchange LC-UV method for the quantification and purity determination of the DNA plasmid pDERMATT, Journal of Pharmaceutical and Biomedical Analysis, 49 (2009) 282-288.

[14] D. Weigl, M.J. Molloy, T.M. Clayton, J. Griffith, C.R. Smith, T. Steward, B.M. Merrill, R.B. DePrince, C.S. Jone, M. Persmark, Characterization of a topologically aberrant plasmid population from pilot-scale production of clinical-grade DNA, Journal of Biotechnology, 121 (2006) 1-12.

[15] D.G. Burke, L.H. Dong, S. Bhat, M. Forbes-Smith, S. Fu, L. Pinheiro, W. Jing, K.R. Emslie, Digital polymerase chain reaction measured pUC19 marker as calibrant for HPLC measurement of DNA quantity, Analytical Chemistry, 85 (2013) 1657-1664.

[16] T. Matos, J.A. Queiroz, L. Bulow, Binding and elution behavior of small deoxyribonucleic acid fragments on a strong anion-exchanger multimodal chromatography resin, Journal of Chromatography A, 1302 (2013) 40-44.

[17] G. O'Connor, C. Dawson, A. Woolford, K.S. Webb, T. Catterick, Quantitation of

oligonucleotides by phosphodiesterase digestion followed by isotope dilution mass spectrometry: Proof of concept, Analytical Chemistry, 74 (2002) 3670-3676.

[18] C.E. Donald, P. Stokes, G. O'Connor, A.J. Woolford, A comparison of enzymatic digestion for the quantitation of an oligonucleotide by liquid chromatography–isotope dilution mass spectrometry, Journal of Chromatography B, 817 (2005) 173-182.

[19] T. Kinumi, T. Narukawa, S. Fujii, S. Eyama, M. Saeki, A. Takatsu, Quantification of an oligonucleotide containing a sequence failure product: comparison of isotope dilution mass spectrometry with other quantification methods, European Journal of Mass Spectrometry, 15 (2009) 399-407.

[20] L. Dong, C. Zang, J. Wang, L. Li, Y. Gao, L. Wu, P. Li, Lambda genomic DNA quantification using ultrasonic treatment followed by liquid chromatography-isotope dilution mass spectrometry, Analytical and Bioanalytical Chemistry, 402 (2012) 2079-2088.

[21] N.S. Hong, L.H. Shi, J.S. Jeong, I. Yang, S.K. Kim, S.R. Park, Rapid and accurate determination of deoxyribonucleoside monophosphates from DNA using micellar electrokinetic chromatography with a cationic surfactant additive, Analytical and Bioanalytical Chemistry, 400 (2011) 2131-2140.

[22] L.H. Dong, Y. Meng, J. Wang, Y.Y. Liu, Evaluation of droplet digital PCR for characterizing plasmid reference material used for quantifying ammonia oxidizers and denitrifiers, Analytical and Bioanalytical Chemistry, 406 (2014) 1701-1712.

[23] I. Yang, I.Y. Park, S.M. Jang, L.H. Shi, H.K. Ku, S.R. Park, Rapid quantification of DNA methylation through dNMP analysis following bisulfite-PCR, Nucleic Acids Research, 34 (2006) e61-e61.

[24] F. Liang, Y.L. Li, W.W. Lu, M.G. Tan, G.L. Zhang, Y. Li, Analysis of deoxynucleotide in plasmid deoxyribonucleic acid using high performance liquid chromatography-inductively coupled

plasma-mass spectrometry, Chinese Journal of Analytical Chemistry, 36 (2008) 1097-1100.

[25] S. Fujii, K. Inagaki, A. Takatsu, T. Yarita, K. Chiba, Determination of phosphorus using capillary electrophoresis and micro-high-performance liquid chromatography hyphenated with inductively coupled plasma mass spectrometry for the quantification of nucleotides, Journal of Chromatography A, 1216 (2009) 7488-7492.

[26] S. Fujii, K. Inagaki, K. Chiba, A. Takatsu, Quantification of phosphorus in DNA using capillary electrophoresis hyphenated with inductively coupled plasma mass spectrometry, Journal of Chromatography A, 1217 (2010) 7921-7925.

[27] C.L. Camp, B.L. Sharp, H.J. Reid, J. Entwisle, H. Goenaga-Infante, Analysis of mono-phosphate nucleotides as a potential method for quantification of DNA using high performance liquid chromatography-inductively coupled plasma-mass spectrometry, Analytical and Bioanalytical Chemistry, 402 (2012) 367-372.

[28] J. Zhang, Q.G. Wang, B. Kleintop, T. Raglione, Suppression of peak tailing of phosphate prodrugs in reversed-phase liquid chromatography, Journal of Pharmaceutical and Biomedical Analysis, 98 (2014) 247-252.

[29] H. Sakamaki, T. Uchida, L.W. Lim, T. Takeuchi, Evaluation of column carryover of phosphorylated peptides and fumonisins by duplicated solvent gradient method in liquid chromatography/tandem mass spectrometry, Analytical Sciences, 31 (2015) 91-97.

[30] H. Sakamaki, T. Uchida, L.W. Lim, T. Takeuchi, Evaluation of column hardware on liquid chromatography-mass spectrometry of phosphorylated compounds, Journal of Chromatography A, 1381 (2015) 125-131.

[31] L. Hyk-Min, Y. Hee-Bong, H. Nan-Sook, Y. Inchul, H. Myung-Sub, P. Sang-Ryoul, Count-based quantitation of trace level macro-DNA molecules, Metrologia, 46 (2009) 375-387.

[32] S. Bhat, K.R. Emslie, Digital polymerase chain reaction for characterisation of DNA reference

materials, Biomolecular Detection and Quantification, doi:10.1016/j.bdq.2016.1004.1001.

[33] M.J. Cavaluzzi, D.J. Kerwood, P.N. Borer, Accurate nucleic acid concentrations by nuclear magnetic resonance, Analytical Biochemistry, 308 (2002) 373-380.

[34] I. Yang, M.S. Han, Y.H. Yim, E. Hwang, S.R. Park, A strategy for establishing accurate quantitation standards of oligonucleotides: quantitation of phosphorus of DNA phosphodiester bonds using inductively coupled plasma-optical emission spectroscopy, Analytical Biochemistry, 335 (2004) 150-161.

[35] M.J. Holden, S.A. Rabb, Y.B. Tewari, M.R. Winchester, Traceable phosphorus measurements by ICP-OES and HPLC for the quantitation of DNA, Analytical Chemistry, 79 (2007) 1536-1541.

[36] S. Shibayama, H. Sakamaki, T. Yamazaki, A. Takatsu, Metal free columns for determination of deoxynucleotide monophosphate by liquid chromatography/mass spectrometry and application to oligonucleotide, Journal of Chromatography A, 1406 (2015) 210-214.

[37] O. Shimelis, R.W. Giese, Nuclease P1 digestion/high-performance liquid chromatography, a practical method for DNA quantitation, Journal of Chromatography A, 1117 (2006) 132-136.

[38] D. Eick, H.J. Fritz, W. Doerfler, Quantitative -determination of 5-methylcytosine in DNA by reverse-phase high-perfromance liquid-chromatography, Analytical Biochemistry, 135 (1983) 165-171.

[39] R. Valencia, H.N. Cong, O. Bertaux, Separation of naturally-occurring or induced methylated nucleobases of DNA by reversed-phase high-performance liquid-chromatography, Journal of Chromatography, 325 (1985) 207-220.

[40] D.H. Fisher, R.W. Giese, Determination of 5-methylcytosine in DNA by gas-chromatography electron-capture, Journal of Chromatography, 452 (1988) 51-60.

[41] T. Yamazaki, T. Ohtsuki, T. Miura, T. Suematsu, T. Horinouchi, M. Murakami, T. Saito, T. Ihara, A. Tada, M. Tahara, Y. Goda, H. Akiyama, S. Nakao, Y. Yamada, R. Koike, N. Sugimoto, Study of

sample preparation method using internal standard solution to accurate quantitative analysis with H-1 NMR spectroscopy, Bunseki Kagaku, 63 (2014) 323-329.

[42] T. Yamazaki, T. Saito, T. Miura, T. Ihara, Investigation of analysis conditions for accurate quantitative NMR analysis, Bunseki Kagaku, 61 (2012) 963-967.

[43] J. 100:2008, Evaluation of measurement data -Guide to the expression of uncertainty in measurement (GUM). http://www.bipm.org/en/publications/guides/#gum

[44] P. Corbisier, S. Trapmann, D. Gancberg, L. Hannes, P. Van Iwaarden, G. Berben, H. Schimmel, H. Emons, Quantitative determination of Roundup Ready soybean (Glycine max) extracted from highly processed flour, Analytical and Bioanalytical Chemistry, 383 (2005) 282-290.

[45] F. Moreano, U. Busch, K.H. Engel, Distortion of genetically modified organism quantification in processed foods: Influence of particle size compositions and heat-induced DNA degradation, Journal of Agricultural and Food Chemistry, 53 (2005) 9971-9979.

[46] X.Y. Wang, V. Veerappan, C. Cheng, X. Jiang, R.D. Allen, P.K. Dasgupta, S.R. Liu, Free solution hydrodynamic separation of DNA fragments from 75 to 106,000 base pairs in a single run, Journal of the American Chemical Society, 132 (2010) 40-41.

[47] B. Vogelstein, K.W. Kinzler, Digital PCR, Proceedings of the National Academy of Sciences of the United States of America, 96 (1999) 9236-9241.

[48] L.B. Pinheiro, V.A. Coleman, C.M. Hindson, J. Herrmann, B.J. Hindson, S. Bhat, K.R. Emslie, Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification, Analytical Chemistry, 84 (2012) 1003-1011.

[49] P. Corbisier, L. Pinheiro, S. Mazoua, A.M. Kortekaas, P.Y.J. Chung, T. Gerganova, G. Roebben, H. Emons, K. Emslie, DNA copy number concentration measured by digital and droplet digital quantitative PCR using certified reference materials, Analytical and Bioanalytical Chemistry, 407 (2015) 1831-1840.

[50] J.F. Huggett, C.A. Foy, V. Benes, K. Emslie, J.A. Garson, R. Haynes, J. Hellemans, M. Kubista, R. Nolan, M.W. Pfaffl, G.L. Shipley, J. Vandesompele, C.T. Wittwer, S.A. Bustin, The digital MIQE guidelines: Minimum information for publication of quantitative digital PCR experiments, Clinical Chemistry, 59 (2013) 892-902.

[51] R. Sanders, J.F. Huggett, C.A. Bushell, S. Cowen, D.J. Scott, C.A. Foy, Evaluation of digital PCR for absolute DNA quantification, Analytical Chemistry, 83 (2011) 6474-6484.

[52] S. Bhat, J.L.H. McLaughlin, K.R. Emslie, Effect of sustained elevated temperature prior to amplification on template copy number estimation using digital polymerase chain reaction, Analyst, 136 (2011) 724-732.

[53] A.S. Whale, J.F. Huggett, S. Cowen, V. Speirs, J. Shaw, S. Ellison, C.A. Foy, D.J. Scott, Comparison of microfluidic digital PCR and conventional quantitative PCR for measuring copy number variation, Nucleic Acids Research, 40 (2012) e82-e82.

[54] A.S. Whale, S. Cowen, C.A. Foy, J.F. Huggett, Methods for applying accurate digital PCR analysis on low copy DNA samples, Plos One, 8 (2013) e58177.

[55] K.V. Voelkerding, S.A. Dames, J.D. Durtschi, Next-generation sequencing: From basic research to diagnostics, Clinical Chemistry, 55 (2009) 641-641.

[56] J.M. Rizzo, M.J. Buck, Key Principles and Clinical Applications of "Next-Generation" DNA Sequencing, Cancer Prevention Research, 5 (2012) 887-900.

[57] F. Sanger, S. Nicklen, A.R. Coulson, DNA sequencing with chain-terminating inhibitors, Proceedings of the National Academy of Sciences, 74 (1977) 5463-5467.

[58] D.J. Munroe, T.J.R. Harris, Third-generation sequencing fireworks at Marco Island, Nat Biotech, 28 (2010) 426-428.

[59] E.E. Schadt, S. Turner, A. Kasarskis, A window into third-generation sequencing, Human Molecular Genetics, 19 (2010) R227-R240.

[60] A. Katsnelson, Momentum grows to make 'personalized' medicine more 'precise', Nature Medicine, 19 (2013) 249-249.

[61] J.H. Coote, M.J. Joyner, Is precision medicine the route to a healthy world?, The Lancet, 385 (2015) 1617-1617.