Doctoral Dissertation

Academic Year 2020

# Development of comprehensive dipeptide analysis method and its application to cancer study

Graduate School of Media and Governance

Keio University

Hitoshi Ozawa


Keio University
1858
CALAMVS GLADIO FORTIOR

# Development of comprehensive dipeptide analysis method and its application to cancer study

Hitoshi Ozawa

A dissertation for the degree of *Doctor of Philosophy*
in the Graduate School of Media and Governance
Systems Biology Program

KEIO UNIVERSITY

2020

# Thesis abstract

## Development of comprehensive dipeptide analysis method and its application to cancer study

Dipeptides, two amino acids linked by a peptide bond, have attracted much attention not only as functional substances and biomaterials but also as biomarkers that can diagnose diseases. Thus, comprehensive analysis method for dipeptides has recently increased. However, such a technique has yet to be developed. Due to existence of many structural isomers, accomplishment of comprehensive dipeptide analysis has been difficult. In this study, I developed methods for the comprehensive analysis of dipeptides, which include separation of structural isomers of dipeptides using liquid chromatography tandem mass spectrometry and capillary electrophoresis tandem mass spectrometry. Subsequently, I applied the methods to cancer study, carrying out cancer-specific profiling of dipeptides in hepatocellular carcinoma.

The first three chapters of this doctoral dissertation open with details on the research background and strategy, analytical and data processing methods used in the study. The fourth chapter describes the development of comprehensive methods for the analysis of dipeptides using liquid chromatography tandem mass spectrometry and capillary electrophoresis tandem mass spectrometry. These methods enabled simultaneous quantitation of 335 types of dipeptides for the first time. The fifth chapter demonstrates the application of the proposed methods to profiling of tumor tissues and surrounding non-tumor tissues obtained from patients with liver cancer and revealed the characteristic dipeptide profiles before and after the onset of a tumor in hepatitis-derived hepatocellular carcinoma. The sixth chapter concludes this study. These novel approaches of my doctoral dissertation will be of great help to elucidate characteristics of dipeptides in various types of cancer.

Keywords: Hepatocellular carcinoma, Hepatitis, Dipeptide, Liquid chromatography tandem mass spectrometry, Capillary electrophoresis tandem mass spectrometry

# 主論文要旨

## ジペプチドの一斉分析法の開発と癌研究への応用

2 つのアミノ酸がペプチド結合でつながったジペプチド類は、機能性物質や生体材料として注目されるばかりでなく、疾病を診断する有用なバイオマーカーとしてもその価値が見出されており、近年、各種のジペプチドを測定する重要性が高まっている。しかし、これまでの研究では、一部のジペプチドの測定法は開発されてきたものの、ジペプチド類を網羅的に一斉分析できる測定法は開発されてこなかった。ジペプチドには、アミノ酸の結合順序が異なる構造異性体が数多く含まれており、これが、ジペプチド類の一斉分析を極めて難しくしていたからである。そこで、本研究では、液体クロマトグラフィータンデム質量分析計とキャピラリー電気泳動タンデム質量分析計を用いて、ジペプチドの構造異性体の分離を含めた、ジペプチド類の一斉分析法の開発を行った。続いて、本手法を癌研究に応用し、肝細胞癌の癌特異的ジペプチド類のプロファイリングを行った。

本博士論文では、本研究における研究背景、研究方針、分析方法およびデータ処理方法を第 1 章から第 3 章で説明する。第 4 章では、世界で初めて３３５種類のジペプチドの同時定量を可能にした液体クロマトグラフィータンデム質量分析計とキャピラリー電気泳動タンデム質量分析計を用いたジペプチドの一斉分析法を開発した。さらに、第 5 章では、本手法を用いて、肝細胞癌の腫瘍組織、周囲の非腫瘍組織のジペプチドプロファイルの違いと、肝炎由来の肝細胞癌における腫瘍発症前後の特徴的なジペプチドプロファイルを明らかにし、第 6 章で結論を述べる。本博士論文における新しいアプローチは、様々な癌の特徴を解明する上で非常に有用であることが期待できる。

キーワード：肝細胞癌、肝炎、ジペプチド、液体クロマトグラフィータンデム質量分析法、キャピラリー電気泳動タンデム質量分析法

# Table of contents

## *Table of contents*

# List of tables

# List of figures

# Chapter 1

## Research background and research plan

One of the important approach to understanding the characteristics of cancer is metabolome analysis, which can reveal the metabolomic profile *in vivo*. Metabolome analysis has been applied to various tumor tissues such as gastric cancer [1], liver cancer [2,3], prostate cancer [4,5], breast cancer [6], oral cancer [7], and lung cancer [4,8]. Additionally, several studies have reported the discovery of potential serum biomarkers of hepatocellular carcinoma (HCC) by metabolome analysis using gas chromatography mass spectrometry (GC/MS) and liquid chromatography mass spectrometry (LC-MS) [9-12].

As post-amino acids, dipeptides are highly diverse with different physical and functional properties from amino acids [13]. In recent years, dipeptides have attracted attention as functional biomaterials, particularly as potential disease biomarkers [14]. Dipeptides comprise two amino acids linked by a peptide bond and are therefore highly diverse, e.g., 400 dipeptides can be constructed from 20 proteinogenic α-amino acids alone. For instance, carnosine and anserine contain an imidazole group derived from histidine and can remove reactive oxygen, and thus play a role as endogenous

antioxidants [15,16]. In addition, leucine-histidine suppresses microglia activity, reduces pro-inflammatory cytokine production, and ameliorates depression and depression-related emotional disturbances [17]. It has also been reported that dipeptides consisting of aromatic amino acids and leucine, such as Tyr-Leu, Phe-Leu, and Trp-Leu, have anxiolytic-like activity in mice [18,19]. Moreover, other dipeptides such as Val-Tyr (antihypertensive effect) [20] and Tyr-Arg (analgesic activity) [21] also exhibit bioactivity, and some diseases are known to result in dipeptide profile changes as exemplified by those in non-small-cell lung cancer tissues [22]. The physiological activity of artificially synthesized dipeptides has also been reported [23,24].

Some dipeptides have also been used as biomarkers of disease. For example, prolyl-4-hydroxyproline, a dipeptide produced when collagen is degraded, is used as a urinary biomarker of bone resorption [25,26]. Furthermore, our metabolome research group found that the concentration of γ-glutamyl dipeptides in the serum fluctuates in nine types of liver diseases such as hepatitis, cirrhosis, and HCC, indication the potential of these biomarkers for liver disease screening [27,28]. Moreover, dipeptides are also found in fermented foods such as soy sauce [29], wine [30,31], and Japanese sake [32] and have therefore attracted attention as new functional food ingredients.

Thus, analyzing dipeptides in a biological sample may lead to the discovery of new

functional components and various disease biomarkers. According to previous study, several dipeptide analysis methods have been reported. For instance, derivatization with derivatization reagent (6-aminoquinolyl-*N*-hydroxysuccinimidyl carbamate [32] or 2,4,6-trinitrobenzenesulfonate [33]) followed by reversed-phase LC-MS analysis. These approaches are suitable for the analysis of dipeptides poorly retained on reversed-phase columns and allows the detection sensitivity to be improved by derivatization. However, the efficiency of derivatization may not be the same for all dipeptides, which causes uncertainly in quantitation. Ion-pair reversed-phase LC-MS can also be used for dipeptide analysis [30] but suffers from possible interference due to ion-pairing reagents and the resulting matrix effects. As a technique well suited for the separation of polar compounds, hydrophilic interaction chromatography has been applied to dipeptide analysis [14,34]. However, methods developed so far have only been applied to certain dipeptides and may not be suitable for all dipeptides.

Although establishment of comprehensive analytical method for dipeptide has become very important because of diversity of dipeptide, few comprehensive dipeptide analyses have been performed. One main reason for this is that all dipeptides except those composed of the same amino acid, have structural isomers with opposite amino acid binding orders. Because these isomers have the same molecular weight, it is difficult to

distinguish them by mass spectrometry. Moreover, the database search strategy using the tandem mass spectra of peptides is commonly used as a conventional proteomic approach, general search algorithms are often limited to peptides containing seven or more amino acids [35]. Because short peptides such as dipeptides often do not generate unique fragments derived from amino acid sequences, greater certainty can be achieved by identifying individual dipeptide fragments using standard reagents.

To overcome these difficulties, we developed a comprehensive dipeptide analytical method using liquid chromatography tandem mass spectrometry (LC-MS/MS) and capillary electrophoresis tandem mass spectrometry (CE-MS/MS) [13]. LC-MS can separate a wide range of compounds and have been used for metabolomics study [9-12]. CE-MS is suitable for the analysis of polar compounds, as exemplified by our metabolome research group's work for the comprehensive and quantitative analysis of charged metabolites in biological samples [36-39].

In this study, the analytical conditions including capillary length and background electrolyte (BGE) pH in CE-MS/MS and column type in LC-MS/MS are optimized. In addition, method validations including repeatability, linearity, sensitivity and spike recovery test are performed in order to confirm the precision of quantitation.

As shown in Figure 1, we achieve the quantitation of 335 dipeptides into 361 kinds

dipeptides composed by two amino acids except dipeptides including cysteine using two methods. It is possible to dramatically increase the number of quantifiable dipeptides by combining two analytical methods with different separation mechanisms. Although the developed method is limited to 335 dipeptides, it can distinguish dipeptide isomers and enables absolute quantitation, with highly sensitive analysis achieved via multiple reaction monitoring (MRM).

I applied the method to diet experiment of mouse, succeeding the elucidation of dipeptide profile difference between liver of mouse fed normal diet or high-fat diet, showing that our method can apply to dipeptide analysis in biological samples.

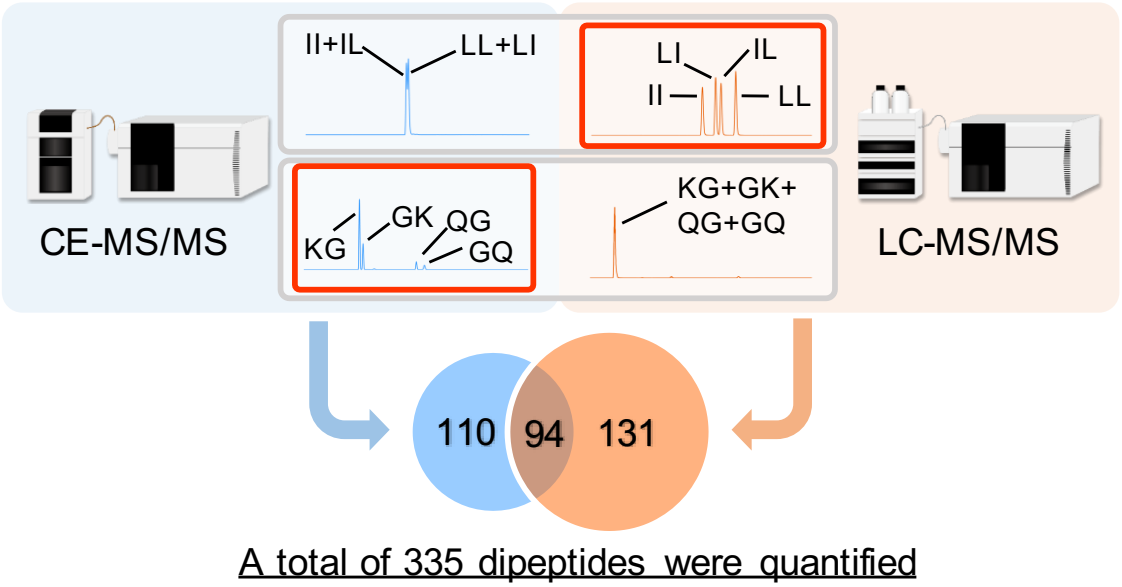



Figure 1. Development of comprehensive dipeptide analysis method using LC-MS/MS and CE-MS/MS. We achieve the quantitation of 335 dipeptides into 361 kinds dipeptides composed by two amino acids except dipeptides including cysteine and it is possible to dramatically increase the number of quantifiable dipeptides by combining two analytical methods with different separation mechanisms.

Furthermore, I applied the method to compare the dipeptide profiles of tumors and surrounding non-tumor tissues of patients with liver cancer [40]. HCC is the third leading cause of cancer-related deaths worldwide [41]. A major cause of HCC is chronic hepatitis caused by hepatitis virus infection. Infection with hepatitis B virus (HBV) or C virus (HCV) causes hepatitis, and long-term destruction and regeneration of hepatocytes leads to cirrhosis and finally HCC [42]. The most characteristic feature of HCC is that tumor growth is rapid and initial symptoms are unlikely to be detected. When a tumor is found, it has often spread to other organs [43]. Hence, elucidation of cancer-specific feature of HCC pathogenesis such as hepatitis is of great help to cancer treatment. In this study, the characteristics of the amino acids constituting the dipeptide detected in the tissues were also examined. Furthermore, the dipeptide profiles in HCC with different etiologies were compared. It was found that the dipeptide profiles in non-tumor and tumor tissues differed, and hepatitis-derived cancer has a characteristic dipeptide profile before and after tumor onset. The details of experimental results are going to be discussed in other chapters and the outline of HCC study is shown in Figure 2.

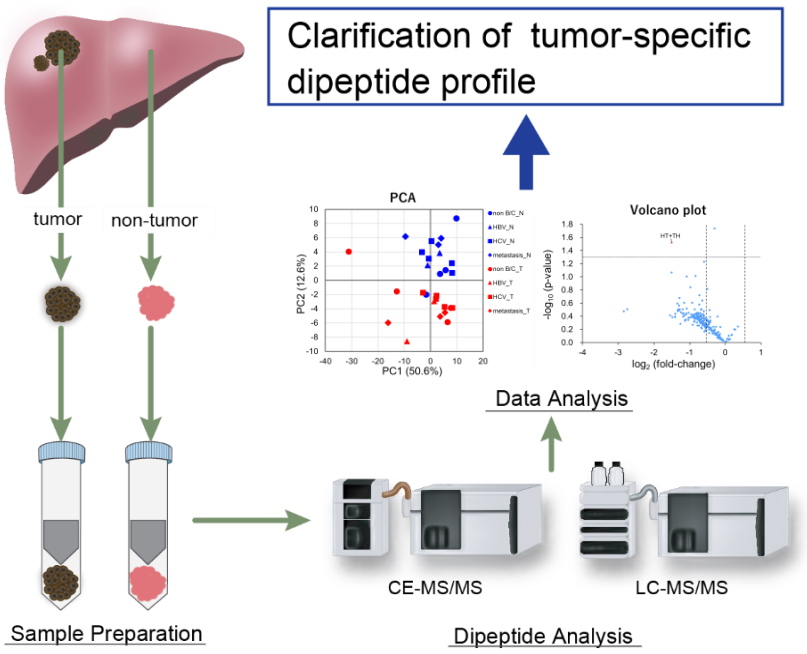



Figure 2. The outline of hepatocellular carcinoma study using developed comprehensive dipeptide analysis method. It was found that the dipeptide profiles in non-tumor and tumor tissues differed, and hepatitis-derived cancer has a characteristic dipeptide profile before and after tumor onset.

# Chapter 2

## Analytical method

### 2.1 Chemicals

Methionine sulfone (internal standard for CE-MS/MS) was purchased from Alfa Aesar (Ward Hill, MA, USA), and α-dipeptides (361 individual compounds) except for those containing cysteine residues were obtained from Sigma-Aldrich (St. Louis, MO, USA), BACHEM (Bubendorf, Switzerland), or MP Biomedicals (Santa Ana, CA, USA). Commercially unavailable dipeptides and Phe-Gly-$^{13}C_9$-$^{15}N_1$ (internal standard for LC-MS/MS) were synthesized by BEX Co. Ltd. (Tokyo, Japan). Methanol (LC-MS grade) and chloroform (reagent grade) were purchased from Wako Pure Chemical Industries Ltd. (Osaka, Japan). Individual stock solutions (10 or 100 mM) were prepared in Milli-Q water (Millipore, Billerica, MA, USA), 0.1 M HCl, 0.1 M NaOH, or methanol. The working standard solutions were prepared by diluting the stock solutions with Milli-Q water immediately before analysis.

## 2.2 CE-MS/MS conditions

CE-MS/MS analysis was performed using an Agilent 1600 CE system (Agilent Technologies, Waldbronn, Germany), an Agilent 6490 triple quadrupole mass spectrometer (Agilent Technologies, Palo Alto, CA, USA), an Agilent 1200 series isocratic high performance liquid chromatography (HPLC) pump, and an Agilent G1607A CE-ESI-MS sprayer kit. The CE and MS instruments were coupled using an electrospray ionization (ESI) interface. The CE system was controlled by Agilent CE Chem Station software, while MS/MS acquisition and data analysis were performed using Agilent MassHunter software (version B.06.00).

Dipeptide separation was carried out on a fused-silica capillary (50 μm I.D., 360 μm O.D., 135 cm total length, Polymicro Technologies, Phoenix, AZ, USA) using aqueous acetic acid (200 mM, pH 2.74) as a BGE, Prior to first use, a new capillary was rinsed with the BGE for 20 min. Equilibration by 4 min of flushing with the BGE was performed before each run. The sample solution was injected at 5 kPa for 15 s (～15 nL), and a positive voltage of 30 kV was applied. The capillary was maintained at 20 ℃, and the sheath liquid (methanol/water = 1:1, v/v) was delivered at 10 μL/min.

ESI-MS/MS analysis was conducted in positive-ion mode using the following source parameters: dry gas temperature = 280 ℃, dry gas flow rate = 11 L/min, nebulizer

pressure = 10 psi, capillary voltage = 4 kV, fragmentor voltage = 380 V, cell accelerator voltage = 7 V, high- and low-pressure radio-frequency voltage of ion funnel = 150 and 60 V, respectively, dwell time = 5 ms.

## 2.3 LC-MS/MS conditions

LC-MS/MS analysis was carried out using an Agilent 1290 Infinity HPLC system coupled to an Agilent 6490 triple quadrupole mass spectrometer with an Agilent jet stream (AJS) ESI interface. System control and data acquisition analysis were performed using Agilent MassHunter software.

Dipeptides were separated on an Acquity UPLC high strength silica (HSS) penta fluoro phenyl (PFP) column (Acquity PFP, 2.1 × 150 mm, 1.8 μm; Waters). The mobile phase comprised 0.1 vol % aqueous formic acid (A) and 0.1 vol % formic acid in 95 vol % aqueous acetonitrile (B). The flow rate equaled 0.2 mL/min, and the following linear gradient was used: 0-3 min, 1% B; 3-30 min, 1 to 50% B; 30-30.1 min, 50 to 99% B; 30.1-35 min, 99% B; 35-35/1 min 99 to 1% B, followed by equilibration with 1% B for 15 min. The injection volume was 1 μL, and the column temperature was maintained at 45 °C. During column evaluation, five other columns, namely XBridge octadecyl silyl (C18) (XBridge C18, 2.1 × 150 mm, 3.5 μm; Waters), Luna Omega Polar C18 (Luna

C18, 2.1 × 150 mm, 1.6 μm; Phenomenex), Discovery HS F5 (Discovery PFP, 2.1 × 150 mm, 3 μm; Sigma-Aldrich), Capcell Core PFP (Capcell PFP, 2.1 × 150 mm, 2.7 μm; Shiseido), and Luna PFP (2) (Luna PFP, 2.0 × 150 mm, 3 μm; Phenomenex) were tested.

AJS-ESI-MS/MS analysis was performed in positive-ion mode using the following source parameters: dry gas temperature = 280 ℃, dry gas flow rate = 12 L/min, nebulizer pressure = 30 psi, sheath gas temperature = 380 ℃, sheath gas flow rate = 12 L/min, capillary voltage = 3.5 kV, nebulizer voltage = 2.0 kV, fragmentor voltage = 380 V, cell accelerator voltage = 7 V, high- and low-pressure radio frequency voltage of ion funnel = 150 and 60 V, respectively, dwell time = 2 ms.

## 2.4 Animal experiments

Male C57BL/6 wild-type mice (aged 8-10 weeks) obtained from SLC (Hamamatsu, Japan) were maintained under temperature- and humidity-controlled specific pathogen-free conditions on a 12 h dark/12 h light cycle and divided into two groups (*N* = 8) depending on whether they were fed a normal diet (CE-2, CLEA, Tokyo, Japan) or a high-fat diet (D 12451, fat 45%, research diet (New Brunswick, NJ, USA)). After 1 month, a significant difference in body weight was observed between the two groups

(Figure 3). Finally, 2 months after experiment initiation, all mice were euthanized after 6 h of fasting to obtain liver samples. All animal experiments were approved by the Animal Care and Use Committee of the Nagoya University Graduate School of Medicine and were conducted in Nagoya University Graduate School of Medicine.

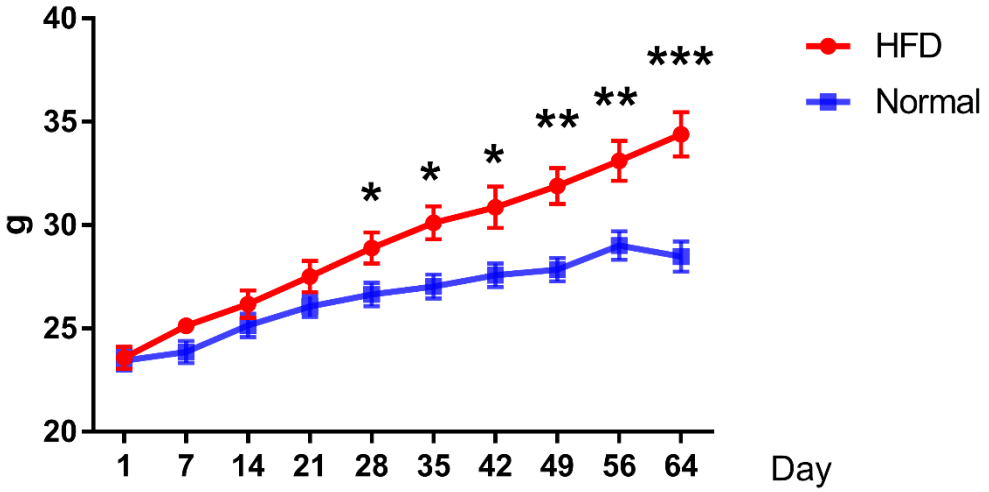



Figure 3. The relationship between breeding days and body weight in each diet. Male C57BL/6 wild-type mice (aged 8-10 weeks) obtained from SLC (Hamamatsu, Japan) were maintained under temperature- and humidity-controlled specific pathogen-free conditions on a 12 h dark/12 h light cycle and divided into two groups ($N = 8$). Normal, normal diet (CE-2, CLEA, Tokyo, Japan); HFD, high fat diet (D 12451, fat 45%, research diet (New Brunswick, NJ, USA)). Student's *t*-test was used for statistical analysis. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

## 2.5 Sample preparation

In standard sample preparation, a liver sample (～50 mg) was homogenized in methanol (500 μL) containing internal standards (20 μM methionine sulfone and 50 μM Phe-Gly-$^{13}C_9$-$^{15}N_1$) using a Shake Master NEO instrument (Bio Medical Science, Tokyo, Japan). Subsequently, chloroform (500 μL) and Milli-Q water (200 μL) were added, and the solution was centrifuged at 4,600 *g* for 15 min at 4 ℃. The upper aqueous layer (300 μL) was centrifugally filtered through a 5 kDa cutoff filter (Human Metabolome Technologies, Tsuruoka, Japan) to remove proteins, and the filtrate was lyophilized and dissolved in Milli-Q water (25 μL) immediately before CE-MS/MS analysis. The obtained sample was further diluted 5-fold using Milli-Q water to mitigate the matrix effect and subjected to LC-MS/MS analysis.

Matrix effect investigation in LC-MS/MS was performed using chicken liver, and both 20 selected dipeptides (at a concentration of 5 μM) and internal standard (Phe-Gly-$^{13}C_9$-$^{15}N_1$, at a concentration of 50 μM) were added after sample preparation. The sample pretreated with 50 mg chicken liver was sequentially diluted with Milli-Q water to prepare samples containing tissue corresponding to 25, 10 and 5 mg.

The recovery test was performed using chicken liver samples (~100 mg) spiked with 40 selected dipeptides (20 for CE-MS/MS at a concentration of 10 μM and 20 for

LC-MS/MS at a concentration of 5 μM) either before or after sample preparation. In post-addition test, internal standard was also added after sample preparation. In pre-addition test, internal standards were prepared to become 20 μM (for methionine sulfone), 50 μM (for Phe-Gly-$^{13}C_9$-$^{15}N_1$). In post-addition test, internal standards were prepared to become 100 μM (for methionine sulfone), 250 μM (for Phe-Gly-$^{13}C_9$-$^{15}N_1$) for CE-MS/MS and 20 μM (for methionine sulfone), 50 μM (for Phe-Gly-$^{13}C_9$-$^{15}N_1$) for LC-MS/MS.

# Chapter 3

## Data processing method

### 3.1 Quantification and statistical analysis

#### 3.1.1 Internal standard method

The number of moles (nmol/g) of each metabolite was quantified as follows. The ratio between concentrations of metabolite and internal standard (IS) is proportional to the ratio between areas of metabolite and IS (equation (1)). Methionine sulfone and Phe-Gly-$^{13}C_9$-$^{15}N_1$ were used as IS for CE-MS/MS and LC-MS/MS, respectively.

$$\frac{C}{C_{\mathrm{IS}}} = k \cdot \frac{X}{X_{\mathrm{IS}}} \tag{1}$$

Where $C$ (nmol/L) and $C_{\mathrm{IS}}$ indicate concentrations of metabolite and IS in suspension after homogenization, respectively, $X$ is the metabolite area, $X_{\mathrm{IS}}$ is the IS area, and $k$ is a constant. For sample preparation, Equation (1) is also valid for the standard (STD) and equation (2) is derived from equation (1).

$$k = \frac{C}{C_{\mathrm{IS}}} \cdot \frac{X_{\mathrm{IS}}}{X} = \frac{C(\mathrm{STD})}{C_{\mathrm{IS}}(\mathrm{STD})} \cdot \frac{X_{\mathrm{IS}}(\mathrm{STD})}{X(\mathrm{STD})}$$

$$\Leftrightarrow C = \frac{X}{X_{\mathrm{IS}}} \cdot \frac{X_{\mathrm{IS}}(\mathrm{STD})}{X(\mathrm{STD})} \cdot C(\mathrm{STD}) \cdot \frac{C_{\mathrm{IS}}}{C_{\mathrm{IS}}(\mathrm{STD})} \tag{2}$$

The number of moles (nmol/g) of each metabolite can be calculated using equation (3).

$$x = \frac{C\theta}{m} \tag{3}$$

Where $x$ is the number of moles (nmol/g) of metabolite, $\theta$ is the volume of methanol after homogenization (L), and $m$ is tissue weight (g).

## 3.1.2 Principal component analysis with auto scaling

If measurement data are represented as $x_{ij}$ (i$^{th}$ specimen, j$^{th}$ metabolite), mean $\bar{x}_j$ (j$^{th}$ metabolite), variance $s_{jj}$ (j$^{th}$ metabolite), and covariance $s_{jk}$ (j$^{th}$, k$^{th}$ metabolite) are given as equations (4)–(6).

$$\bar{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij} \tag{4}$$

$$s_{jj} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)^2 \tag{5}$$

$$s_{jk} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)\left(x_{ik} - \bar{x}_k\right) \tag{6}$$

Where $n$ is the number of specimens. Standardized data $X_{ij}$ (i$^{th}$ specimen, j$^{th}$ metabolite), mean $\bar{X}_j$ (j$^{th}$ metabolite), variance $s^*_{jj}$ (j$^{th}$ metabolite), and covariance $s^*_{jk}$ (j$^{th}$, k$^{th}$ metabolite) are derived in equations (7)–(11).

$$X_{ij} = \frac{x_{ij} - \bar{x}_j}{\sqrt{s_{jj}}} \tag{7}$$

$$\bar{X}_j = \frac{1}{n}\sum_{i=1}^{n} X_{ij} = \frac{\bar{x}_j - \bar{x}_j}{\sqrt{s_{jj}}} = 0 \tag{8}$$

$$s_{jj}^{*} = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_{ij} - \bar{X}_j\right)^2 = \frac{1}{n-1}\sum_{i=1}^{n} {X_{ij}}^2$$

$$= \frac{1}{s_{jj}} \cdot \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)^2 = \frac{1}{s_{jj}} \cdot s_{jj} = 1 \tag{9}$$

$$s_{jk}^{*} = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_{ij} - \bar{X}_j\right)\left(X_{ik} - \bar{X}_k\right) = \frac{1}{n-1}\sum_{i=1}^{n} X_{ij}X_{ik}$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{x_{ij} - \bar{x}_j}{\sqrt{s_{jj}}}\right)\left(\frac{x_{ik} - \bar{x}_k}{\sqrt{s_{kk}}}\right)$$

$$= \frac{1}{\sqrt{s_{jj}}\sqrt{s_{kk}}} \cdot \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)\left(x_{ik} - \bar{x}_k\right)$$

$$= \frac{s_{jk}}{\sqrt{s_{jj}}\sqrt{s_{kk}}} = r_{jk} = r_{kj} \tag{10}$$

$$r_{jk} = \frac{s_{jk}}{\sqrt{s_{jj}}\sqrt{s_{kk}}} \tag{11}$$

Where $r_{jk}$ (j$^{th}$, k$^{th}$ metabolite) in equation (11) is the correlation coefficient. Principal component coordinate $Y_i$ (i$^{th}$ specimen) is defined in equation (12) and mean $\bar{Y}$ is derived in equation (13).

$$Y_i = \sum_{j=1}^{m} p_j X_{ij} \tag{12}$$

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m} p_j X_{ij}$$

$$= \sum_{j=1}^{m}\left(p_j \cdot \frac{1}{n}\sum_{i=1}^{n} X_{ij}\right) = \sum_{j=1}^{m} p_j \bar{X}_j = 0 \tag{13}$$

Where $m$ is the number of metabolites and $p_j$ (j[th] metabolite) is a constant. Therefore, the variance $V(Y)$ of the principal component coordinate is derived in equation (14).

$$V(Y) = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = \frac{1}{n-1}\sum_{i=1}^{n}{Y_i}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(\sum_{j=1}^{m} p_j X_{ij}\right)^2$$
$$= \sum_{j=1}^{m}\left({p_j}^2 \cdot \frac{1}{n-1}\sum_{i=1}^{n}{X_{ij}}^2\right) + \sum_{j=1}^{m}\sum_{\substack{k=1 \\ j\neq k}}^{m}\left(p_j p_k \cdot \frac{1}{n-1}\sum_{i=1}^{n} X_{ij}X_{ik}\right)$$
$$= \sum_{j=1}^{m}{p_j}^2 s_{jj}^* + \sum_{j=1}^{m}\sum_{\substack{k=1 \\ j\neq k}}^{m} p_j p_k s_{jk}^* = \sum_{j=1}^{m}{p_j}^2 + \sum_{j=1}^{m}\sum_{\substack{k=1 \\ j\neq k}}^{m} p_j p_k r_{jk} \quad (14)$$

Then, standardization of sum of square of $p_j$ is assumed as equation (15). Principal component coordinate is selected by maximizing variance. Hence, the Lagrange multipliers method can be applied as equation (16) by using equations (14) and (15).

$$\sum_{j=1}^{m}{p_j}^2 = 1 \quad (15)$$

$$F(\{p_i\}_{i=1}^{m}, \lambda) = \sum_{j=1}^{m}{p_j}^2 + \sum_{j=1}^{m}\sum_{\substack{k=1 \\ j\neq k}}^{m} p_j p_k r_{jk} - \lambda\left(\sum_{j=1}^{m}{p_j}^2 - 1\right) \quad (16)$$

Where $F$ is Lagrangian and $\lambda$ is the Lagrange multiplier. Equations (17) and (18) are derived by solving equation (16).

$$\frac{\partial F}{\partial p_i} = 2p_i + 2\sum_{\substack{j=1 \\ i\neq j}}^{m} p_j r_{ij} - 2\lambda p_i = p_i + \sum_{\substack{j=1 \\ i\neq j}}^{m} p_j r_{ij} - \lambda p_i = 0$$
$$\Leftrightarrow p_i + \sum_{\substack{j=1 \\ i\neq j}}^{m} r_{ij} p_j = \lambda p_i \quad (17)$$

$$\frac{\partial F}{\partial \lambda} = \sum_{j=1}^{m} {p_j}^2 - 1 = 0 \quad (18)$$

Furthermore, equation (17) can express matrix formation of the eigenvalue problem (equations (19)–(21)).

$$D\boldsymbol{p}^t = \lambda \boldsymbol{p}^t \quad (19)$$

$$D = \begin{bmatrix} 1 & \cdots & r_{1m} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & 1 \end{bmatrix} \quad (20)$$

$$\boldsymbol{p} = [p_1 \quad p_2 \quad \cdots \quad p_m] \quad (21)$$

Where $D$ is the correlation matrix, line vector $\boldsymbol{p}$ represents eigenvector, and Lagrange multiplier $\lambda$ indicates eigenvalue. The eigenvalue problem that is specific for principal component of eigenvector and eigenvalue is displayed in equations (22) and (23) ($i^{th}$ principal component).

$$D{\boldsymbol{p}_i}^t = \lambda_i {\boldsymbol{p}_i}^t \quad (22)$$

$$\boldsymbol{p}_i = [p_{1i} \quad p_{2i} \quad \cdots \quad p_{mi}] \quad (23)$$

Factor loading $q_{ji}$ ($i^{th}$ principal component, $j^{th}$ metabolite) in the loading plot is defined in equation (24).

$$q_{ji} = \sqrt{\lambda_i} p_{ji} \quad (24)$$

## 3.1.3 $T^2$ statistics

Principal component coordinate $Y_{ik}$ ($i^{th}$ specimen, $k^{th}$ principal component) can be redefined in equation (25).

$$Y_{ik} = \sum_{j=1}^{m} p_{jk} X_{ij} \tag{25}$$

Then, the Mahalanobis distance is defined in equation (26).

$$M_i{}^2 = \sum_{j=1}^{A} \frac{Y_{ij}{}^2}{\lambda_j}$$

$$= [Y_{i1} \quad Y_{i2} \quad \dots \quad Y_{iA}] \begin{bmatrix} 1/\lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1/\lambda_A \end{bmatrix} [Y_{i1} \quad Y_{i2} \quad \dots \quad Y_{iA}]^t \tag{26}$$

Where $M_i$ ($i^{th}$ specimen) represents the Mahalanobis distance and $A$ is the number of principal components. The value of $A$ is usually 2. Equation (25) can be expressed as a matrix in equations (27)–(29).

$$[Y_{i1} \quad Y_{i2} \quad \dots \quad Y_{iA}] = [X_{i1} \quad X_{i2} \quad \dots \quad X_{im}] \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1A} \\ p_{21} & p_{22} & \cdots & p_{2A} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mA} \end{bmatrix}$$

$$= \boldsymbol{X}_{ci} P_A \tag{27}$$

$$\boldsymbol{X}_{ci} = [X_{i1} \quad X_{i2} \quad \dots \quad X_{im}] \tag{28}$$

$$P_A = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1A} \\ p_{21} & p_{22} & \cdots & p_{2A} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mA} \end{bmatrix} \tag{29}$$

And the diagonal matrix in equation (26) can be expanded as equations (30) and (31).

$$\begin{bmatrix} 1/\lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1/\lambda_A \end{bmatrix} = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_A \end{bmatrix}^{-1} = L^{-1} \tag{30}$$

$$L = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_A \end{bmatrix} \tag{31}$$

The Mahalanobis distance is given as equation (32).

$$M_i{}^2 = \boldsymbol{X}_{ci} P_A L^{-1} (\boldsymbol{X}_{ci} P_A)^t = \boldsymbol{X}_{ci} P_A L^{-1} P_A{}^t \boldsymbol{X}_{ci}{}^t \tag{32}$$

The $T^2$ statistic is equivalent to the square of the Mahalanobis distance. Therefore, the $T^2$ statistic is given as equation (33).

$$T_i{}^2 = M_i{}^2 = \boldsymbol{X}_{ci} P_A L^{-1} P_A{}^t \boldsymbol{X}_{ci}{}^t \tag{33}$$

Moreover, if $A = m$, equation (34) is obtained.

$$DP_A = P_A L \Leftrightarrow L = P_A{}^{-1} D P_A \Leftrightarrow L^{-1} = P_A{}^{-1} D^{-1} P_A \tag{34}$$

Then, $P_A$ becomes an orthogonal matrix (equation (35)) because $P_A$ can diagonalize the symmetric matrix $D$.

$$P_A{}^{-1} = P_A{}^t \tag{35}$$

Therefore, if $A = m$, the Mahalanobis distance is given as equation (36).

$$M_i{}^2 = \boldsymbol{X}_{ci} P_A P_A{}^t D^{-1} P_A P_A{}^t \boldsymbol{X}_{ci}{}^t = \boldsymbol{X}_{ci} D^{-1} \boldsymbol{X}_{ci}{}^t \tag{36}$$

As shown in equation (36), it is understood that the Mahalanobis distance includes the correlation between variates in the Euclidean distance.

## 3.1.4 Control limit

The control limit was evaluated using equation (37).

$$\mathrm{CL}_{1-\alpha} = \frac{(n-1)^2}{n} x_{1-\alpha}^{\beta}\left[\frac{A}{2}, \frac{n-A-1}{2}\right] \quad (37)$$

Where $n$ is the number of specimens, $A$ is the number of principal components, $\alpha$ is the significance level, and $x_{1-\alpha}^{\beta}$ is the $(1-\alpha)$-quantile of the beta distribution of parameter $[A/2, (n-A-1)/2]$. Upper control limit (UCL) is defined as $\alpha = 0.01$. In the examination of outliers using Hotelling's $T^2$ test, the $T^2$ statistic (equation (33)) and upper control limit were used to evaluate the results of PCA with auto scaling [44].

## 3.2 Instrumental detection limit

For linearity evaluation, standard solutions were prepared at concentrations of 0.1, 0.5, 1, 5, and 10 μM, each containing methionine sulfone (200 μM) as the IS for CE-MS/MS. For LC-MS/MS, 0.0192, 0.096, 0.48, 2.4, and 12 μM standard solutions were prepared, each containing Phe-Gly-$^{13}C_9$-$^{15}N_1$ (10 μM) as an IS.

For limit of detection (LOD) calculations, the instrumental detection limit ($C_{\mathrm{IDL}}$) was determined using the values (n = 6, α = 0.01, α is the significant level) obtained for 0.001, 0.01, and 0.1μM standard solutions (equation (38)).

$$C_{\mathrm{IDL}} = \frac{t_{1-\alpha} S_{\mathrm{STD}} C_{\mathrm{STD}}}{\bar{X}_{\mathrm{STD}}} \quad (38)$$

Where $S_{\mathrm{STD}}$ is the standard deviation of the peak area of the dipeptide in the standard solution, $C_{\mathrm{STD}}$ is the concentration of the standard solution, $\bar{X}_{\mathrm{STD}}$ is the average peak area of the dipeptide in the standard solution, and $t$ is the statistic of the *t*-test. Equation (38) can be derived as follows. A schematic figure of IDL is shown in Figure 4.

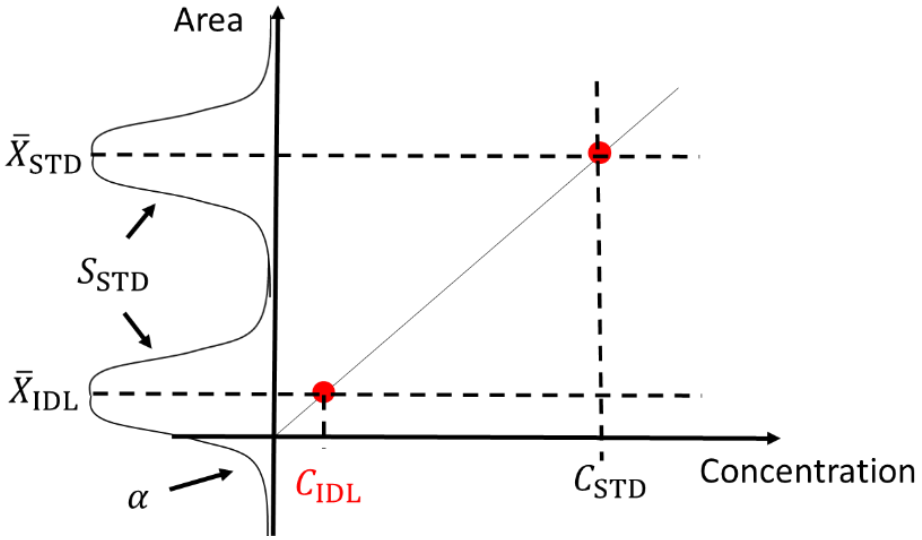



Figure 4. The schematic figure of instrumental detection limit ($C_{IDL}$) [45]. $\bar{X}_{IDL}$ is the peak area of limit of detection. $S_{STD}$ is the standard deviation of the peak area of the dipeptide in the standard solution. $C_{STD}$ is the concentration of the standard solution. $\bar{X}_{STD}$ is the average peak area of the dipeptide in the standard solution. α is the significant level.

Where $\bar{X}_{\mathrm{IDL}}$ is the peak area of limit of detection. Null hypothesis $H_0$ and alternative hypothesis $H_1$ are set as $\mu = 0$ and $\mu > 0$ (where $\mu$ is the population mean of peak area). Then, the rejection region of $H_0$ is determined as equation (39) (for IDL, standard deviation is conventionally used instead of standard error).

$$t_{1-\alpha} < \frac{\bar{X} - \mu}{S_{\mathrm{STD}}} = \frac{\bar{X}}{S_{\mathrm{STD}}} \tag{39}$$

Therefore, $\bar{X}_{\mathrm{IDL}}$ is given as equation (40).

$$\bar{X}_{\mathrm{IDL}} = t_{1-\alpha} S_{\mathrm{STD}} \tag{40}$$

About $3S_{\mathrm{STD}}$ is conventionally used as $\bar{X}_{\mathrm{IDL}}$. According to Figure 4, the slope of the calibration curve is $\bar{X}_{\mathrm{STD}} \,/\, C_{\mathrm{STD}}$. Equation (38) is derived from equation (41).

$$\bar{X}_{\mathrm{IDL}} = t_{1-\alpha} S_{\mathrm{STD}} = \frac{\bar{X}_{\mathrm{STD}}}{C_{\mathrm{STD}}} C_{\mathrm{IDL}} \;\Leftrightarrow\; C_{\mathrm{IDL}} = \frac{t_{1-\alpha} S_{\mathrm{STD}} C_{\mathrm{STD}}}{\bar{X}_{\mathrm{STD}}} \tag{41}$$

## 3.3 Resolution

The resolution ($R_{\mathrm{s}}$) for the dipeptide isomer was calculated using equation (42), which is defined by the general rules of JIS for high performance liquid chromatography.

$$R_{\mathrm{s}} = \frac{1.18(t_2 - t_1)}{W_{0.5,1} + W_{0.5,2}} \tag{42}$$

Where $t_1$ and $t_2$ are migration and retention times, respectively, of the dipeptide isomers, and $W_{0.5}$ refers to temporal peak width at half height.

## 3.4 Software of data processing

Acquired data were analyzed using the MassHunter software (version B.06.00, Agilent Technologies). Principal component analysis (PCA) was performed using the JMP (version 14.0.0) or SIMCA software (version 13, Umetrics AB, Umeå, Sweden).

# Chapter 4

## Development of comprehensive dipeptide analysis method

### 4.1 Optimization of the MRM transition settings for dipeptides

For each dipeptide, optimization was performed using the optimizer function of the Mass Hunter software. The working solutions of individual dipeptides and internal standards (at a concentration of 10 or 100 μM) were prepared by diluting the stock solutions with Milli-Q water, and 1 μL samples were flow-injected into the mass spectrometer. The mobile phase corresponded to 0.1 vol % formic acid in 50 vol % aqueous acetonitrile, and the flow rate equaled 0.2 mL/min. Other instrumental conditions were identical to those of LC-MS/MS described in section 2.3. MRM transition settings including precursor ion, product ion, and collision energy are optimized (Table A1). Herein, the amino acids making up the dipeptides were expressed in one-letter code.

## 4.2 Optimization of CE-MS/MS conditions

In CE-MS, the migration velocity of a given compound depends on its charge and ionic radius. In the case of isomers, these values are generally almost identical, which complicates separation. Inspired by our previous work, which showed that the strict adjustment of the BGE pH allows the separation of γ-glutamyl peptides [27], I probed the effect of BGE pH on the separation behavior (migration time and resolution) of dipeptide isomers. The results for three dipeptide isomers that exhibit characteristic separation behavior are shown in Figure 5A, B. In the case of 1 M formic acid (pH 1.58), which was used for routine cationic metabolite analysis [46], only the QD isomer was separated. On the other hand, when the BGE was changed to acetic acid (100-2000 mM) and its pH was adjusted by the addition of aqueous ammonia (5.6 wt %), all tested isomers were separated between pH 2.33 and 2.9. However, some isomers could not be separated at pH values above 3.5. Considering the resolution and migration time, I determined the optimal BGE pH as 2.74 (Figure 6). The pH dependence of resolution is attributable to the acid dissociation constant of carboxylic acid groups of amino acids because pK1 equals pH when one half of the carboxylic acid groups is protonated and migration time depends on the charge and hydrodynamic radius of the amino acid. pK1 values for Q, D, V, K, and G are 2.17, 1.88, 2.32, 2.20, and 2.35, respectively [47].

Therefore, the larger resolution of QD-DQ compared with that of VD-DV and KG-GK could be due to differences in the pK1 of the structural isomer and D could contribute to large resolutions of QD-DQ and VD-DV at a low pH because of differences in the low pK2 of its carboxyl group of side chain of the structural isomer due to inductive effect. However, the 26 types of dipeptides, for which the structural isomers cannot be separated, could be quantitated at pH values ranging from 9 to 11 because these amino acids have pKa values ranging from 9 to 11, and the difference between the pK2 or pK3 values of the structural isomer is larger than those of pK1 values. pK2 or pK3 values of Q, D, V, K, and G are 9.13, 9.60, 9.62, 8.90, and 9.78, respectively [47]. However, the alkaline solution of BGE needs to be investigated further because if the pH of BGE rises, enough separation of the structural isomer cannot be achieved as the electroosmotic flow becomes faster. As the efficiency of dipeptide isomer separation can be enhanced by increasing the capillary length, I investigated the effect of this parameter on the migration time and resolution (Figure 5C, D). At a capillary length of 100 cm, the migration time was short, but the separation was insufficient. On the other hand, when the capillary length exceeded 140 cm, the resolution did not increase but the migration time did. Therefore, I set the capillary length to 135 cm (Figure 7). A similar trend was observed for other dipeptides analyzed by CE-MS/MS.

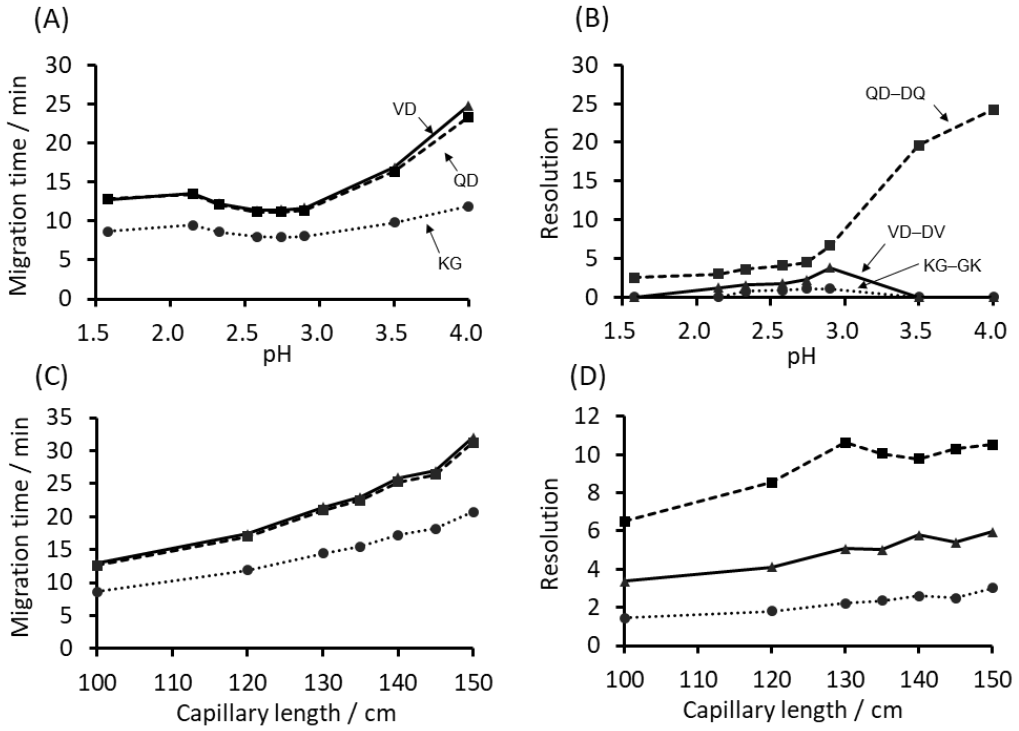



Figure 5. Optimization of CE-MS/MS conditions. Effect of pH on (A) migration time and (B) peak resolution. Effect of capillary length on (C) migration time and (D) peak resolution. The resolution ($R_s$) for dipeptide isomer was calculated using equation (42) which is defined as JIS high performance liquid chromatography common rule. Considering the resolution and migration time, we determined the optimal BGE pH as 2.74 and set the capillary length to 135 cm.

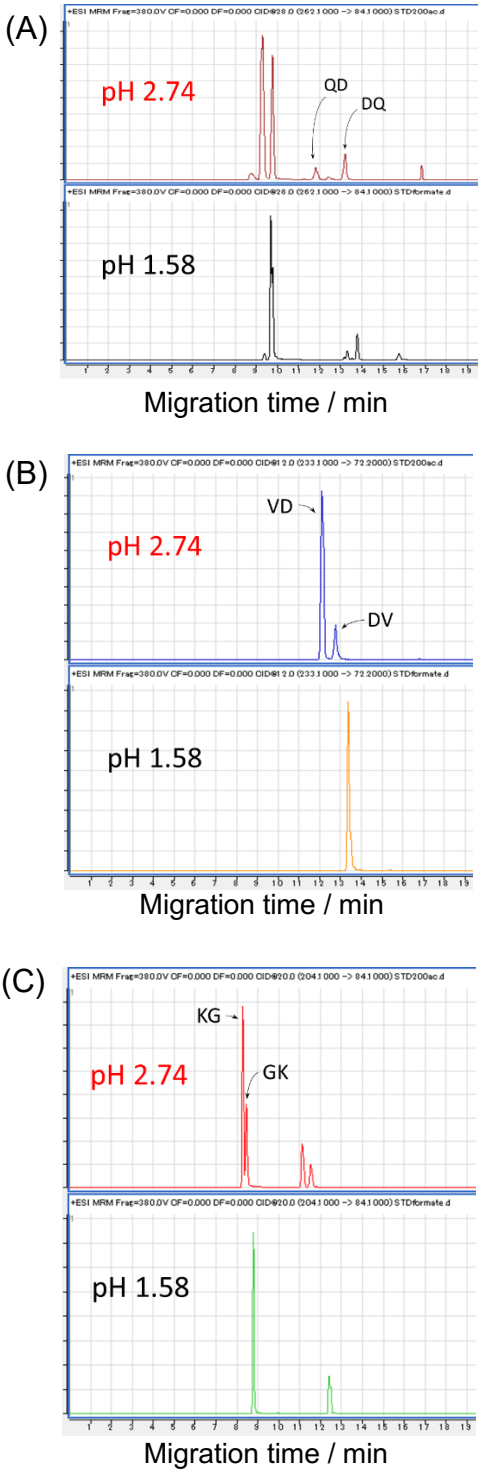



Figure 6. Separation of structural isomers of dipeptides by controlling pH. Capillary length is 100 cm. Electropherogram of QD-DQ (A), VD-DV (B) and KG-GK (C).

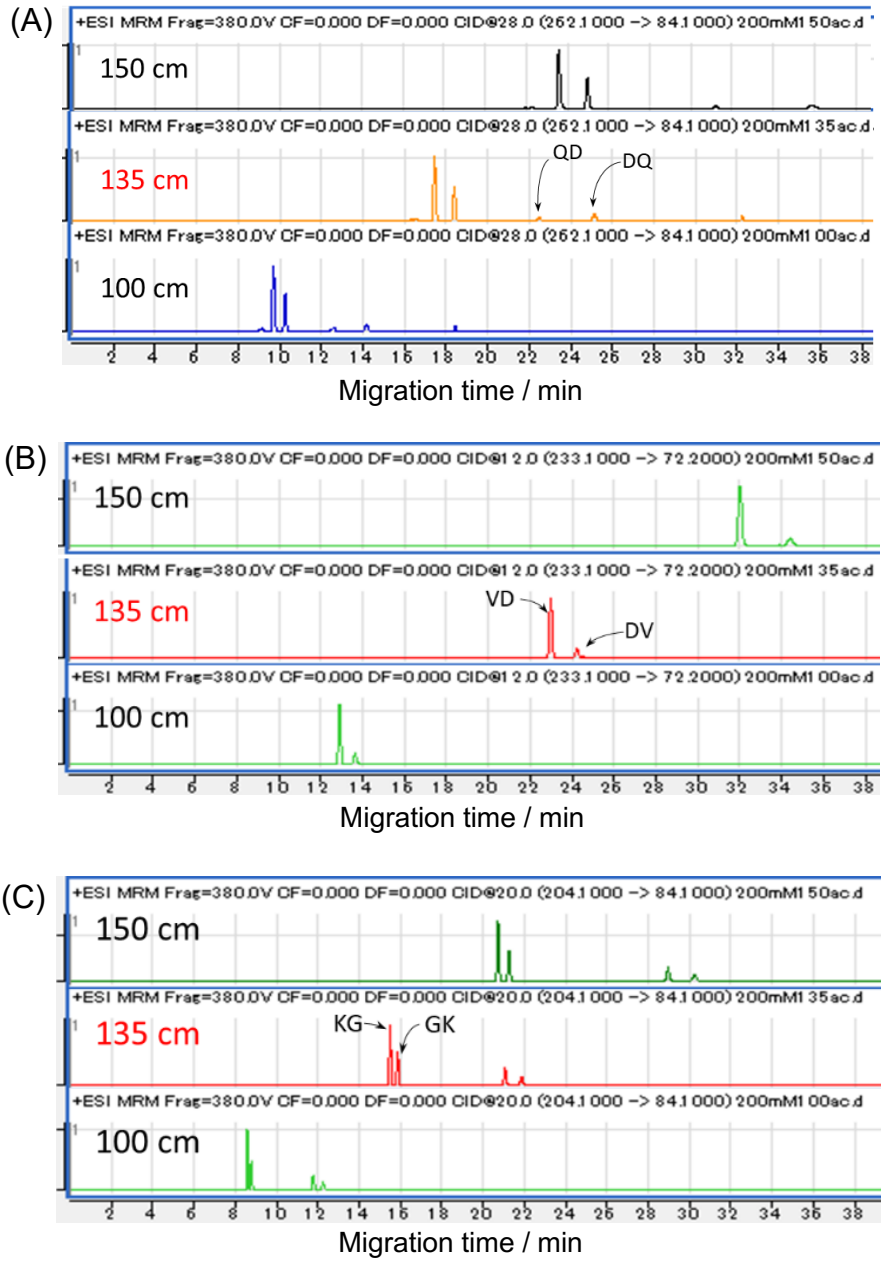



Figure 7. Separation of structural isomers of dipeptides by controlling capillary length. pH is 2.74. Electropherogram of QD-DQ (A), VD-DV (B) and KG-GK (C).

## 4.3 Optimization of LC-MS/MS conditions

Column type is the most important factor affecting the separation of dipeptides and other analytes by LC-MS. Herein, I evaluated six columns under the condition of constant eluent composition and gradient. Three dipeptide isomers representing typical trends were selected as test compounds, with the obtained resolutions shown in Figure 8. FG isomers were well separated on all tested columns. In view of the aromatic nature of phenylalanine, dipeptides containing this amino acid were well retained on both C18 and PFP columns. On the contrary, PR and LI isomers could only be separated on some columns. Specifically, PR isomers could be separated by all columns except for one PFP column, with the separation efficiency of PFP columns exceeding that of C18 columns. In line with the fact that PFP columns can separate polar compounds well [48,49], I observed that hydrophilic dipeptides were better retained and separated on these columns. LI isomers could be separated by four columns (two C18 and two PFP ones), but the achieved separation efficiency was smaller than in the case of other dipeptides. Taking these results into account, the Acquity column was found to separate the largest number of dipeptide isomers among the tested columns. When other dipeptide isomers were investigated, a similar trend was observed. Hence, the Acquity column was selected for further experiment.

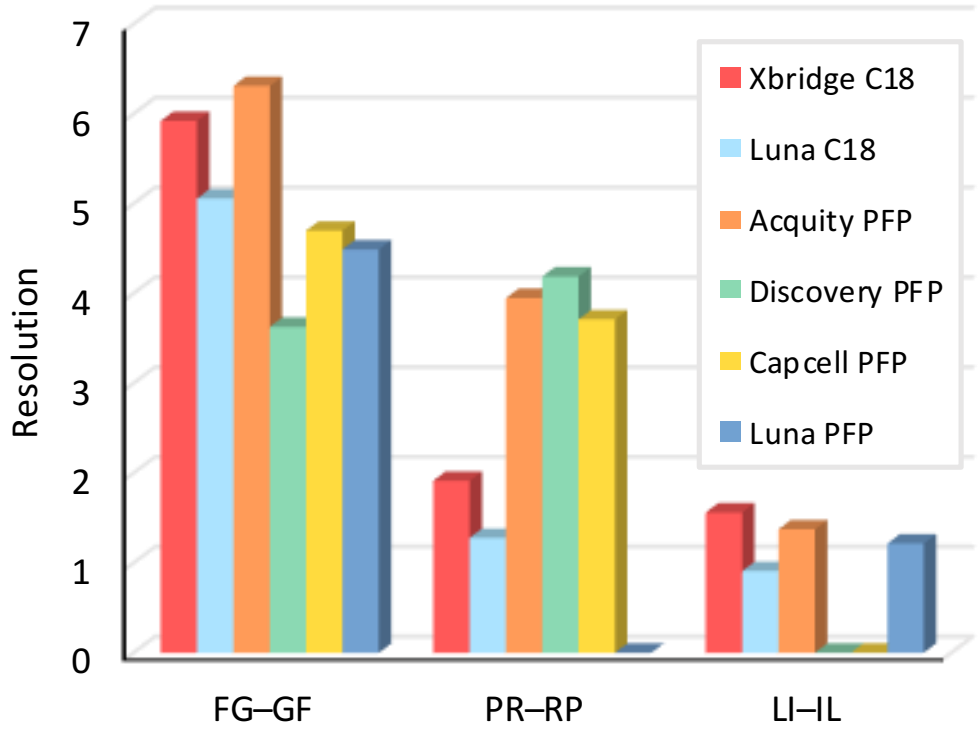



Figure 8. Separation of three dipeptide isomers on six different octadecylsilyl (C18) or pentafluorophenyl (PFP) columns. The resolution ($R_s$) for dipeptide isomer was calculated using equation (42) which is defined as JIS high performance liquid chromatography common rule. The Acquity column was found to separate the largest number of dipeptide isomers among the tested columns.

## 4.4 Method validation

### 4.4.1 Separation of structural isomer

Under the optimized CE-MS/MS and LC-MS/MS conditions, isomer separation was verified for 361 dipeptides comprising 19 L-amino acids excluding cysteine, as this amino acid is unstable and easily forms disulfide bonds with other thiol-containing compounds during pretreatment. Figure 9 shows an example of dipeptide isomer analysis by CE-MS/MS and LC-MS/MS. Notably, II, LI, II, IL and LL are all structural isomers with the same MRM transition. In CE-MS/MS, two isomers featured overlapped peaks (Figure 9A), whereas all four isomers were baseline-separated in LC-MS/MS (Figure 9B). On the other hand, in the cases of KG, GK, QG and GQ, all four dipeptides were coeluted in LC-MS/MS (Figure 9D) but were completely separated in CE-MS/MS (Figure 9C). Therefore, I concluded that an appropriate analytical method has to be selected for each dipeptide isomer. Figure 9E presents an overview of the separation of 361 dipeptides, revealing that 335 dipeptides could be separated from their structural isomers by CE-MS/MS and/or LC-MS/MS, whereas 26 dipeptides could be detected but could not be separated from their isomers. For the latter dipeptides, separation may possibly be achieved by changing the column and the eluent.

Herein, the standard solution for either CE-MS/MS or LC-MS/MS analysis contained

the two isomers of each dipeptide in equal amounts, and each dipeptide was quantified as the sum of its two isomers. Finally, 128 and 233 dipeptides were analyzed by CE-MS/MS and LC-MS/MS respectively. MRM electropherogram and chromatogram for a standard mixture of 361 dipeptides are shown in Figure A1.

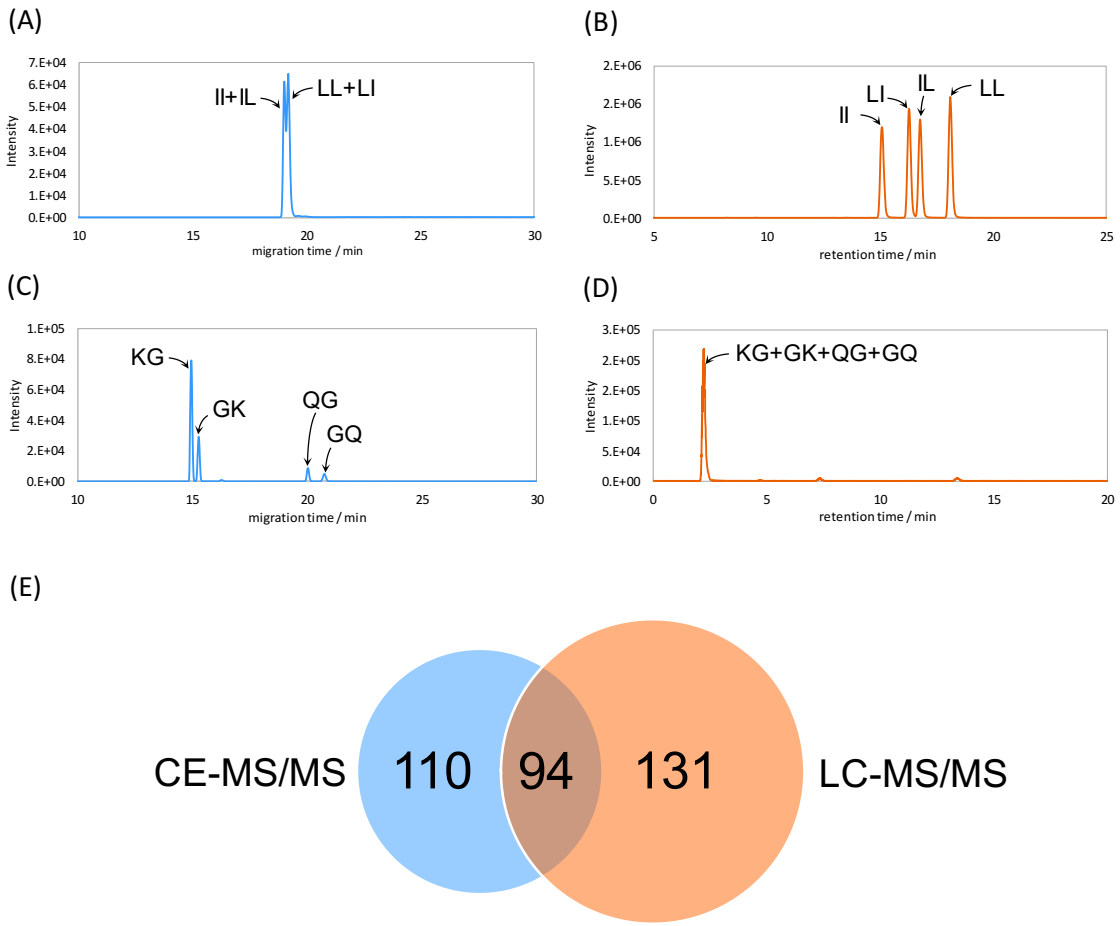



Figure 9. Example of dipeptide isomer separation by CE-MS/MS (A and C) and LC-MS/MS (B and D). (E) Venn diagram showing the number of dipeptides that can be quantified alone for each method. As a result, 94 dipeptides could be separated from their isomers by either method. The concentration of each dipeptide was 10 μM.

## 4.4.2 Repeatability, linearity and sensitivity

The results of repeatability, linearity, and sensitivity are presented in Table 1 and Table 2. The repeatability test was carried out by injecting the standard solution (20 μM for CE-MS/MS, 15 μM for LC-MS/MS) of 40 selected dipeptides six times to determine their migration/retention times and peak areas. The relative standard deviations (RSDs) of the intraday measurements ranged from 0.049 to 3.2% for migration or retention time and from 1.1 to 28.2% for peak area. Further, the RSDs of the interday measurements (3 days) ranged from 0.055 to 1.7% for migration or retention time and from 3.4 to 13.4% for peak area. The calibration curves of the tested dipeptides showed linearity over a wide concentration range (0.1-10 μM for CE-MS/MS, 0.0192-12 μM for LC-MS/MS), with the corresponding correlation coefficients lying between 0.981 and 0.999. The LODs, which were determined as described in section 3.3, varied from 2.9 to 83.1 nM for CE-MS/MS and from 0.088 to 3.2 nM for LC-MS/MS.

Table 1. Repeatability of the developed method

| Dipeptide | Q1 (*m/z*) | Q3 (*m/z*) | Repeatability (RSD, %) | | | |
|---|---|---|---|---|---|---|
| | | | Intraday (n = 6) | | Interday (3 days, n = 18) | |
| | | | Migration / retention time | Peak area | Migration / retention time | Peak area |
| **CE-MS/MS** | | | | | | |
| RR | 166.1 | 70.1 | 1.6 | 5.6 | 1.0 | 5.3 |
| SH | 243.1 | 110 | 2.2 | 13.3 | 1.2 | 7.5 |
| HS | 243.1 | 110 | 2.2 | 13.7 | 1.2 | 9.6 |
| RD | 290.1 | 70.1 | 2.2 | 11.7 | 1.2 | 6.9 |
| HQ | 284.1 | 110.1 | 2.3 | 28.2 | 1.3 | 13.4 |
| HD | 271.1 | 110.1 | 2.4 | 9.6 | 1.3 | 8.5 |
| AG | 147.1 | 147.1 | 1.6 | 8.3 | 1.0 | 6.7 |
| SA | 177.1 | 60.2 | 2.7 | 11.6 | 1.4 | 9.9 |
| TA | 191.1 | 74.1 | 2.7 | 7.7 | 1.5 | 8.6 |
| EA | 219.1 | 201.1 | 2.8 | 9.0 | 1.5 | 9.5 |
| AN | 204.1 | 133.1 | 2.8 | 16.1 | 1.5 | 9.7 |
| GN | 190.1 | 173.1 | 2.9 | 12.0 | 1.5 | 9.5 |
| NS | 220.1 | 140 | 2.9 | 12.8 | 1.5 | 9.4 |
| TS | 207.1 | 74.1 | 2.9 | 13.0 | 1.5 | 8.2 |
| GE | 205.1 | 84.2 | 2.9 | 9.5 | 1.5 | 7.9 |
| SE | 235.1 | 60.1 | 3.0 | 13.9 | 1.5 | 10.6 |
| WQ | 333.2 | 159 | 3.0 | 14.0 | 1.6 | 10.0 |
| GD | 191.1 | 134.1 | 3.1 | 10.6 | 1.6 | 9.3 |
| SD | 221.1 | 60.1 | 3.1 | 9.9 | 1.6 | 10.5 |
| DD | 249.1 | 133.9 | 3.2 | 9.9 | 1.7 | 6.8 |
| **LC-MS/MS** | | | | | | |
| MS | 237.1 | 104 | 0.28 | 3.9 | 0.20 | 6.6 |
| VQ | 246.1 | 72.2 | 0.15 | 3.7 | 0.21 | 5.9 |
| GP | 173.1 | 116.1 | 0.15 | 2.6 | 0.16 | 4.9 |
| VG | 175.1 | 55.1 | 0.29 | 8.4 | 0.23 | 9.4 |
| VA | 189.1 | 72.2 | 0.060 | 7.7 | 0.17 | 5.7 |
| MQ | 278.1 | 147.1 | 0.17 | 3.0 | 0.29 | 9.4 |
| AP | 187.1 | 116.1 | 0.27 | 6.0 | 1.1 | 9.1 |
| GV | 175.1 | 129.1 | 0.17 | 6.5 | 0.39 | 8.0 |
| LT | 233.1 | 86 | 0.21 | 7.1 | 0.74 | 8.1 |
| IA | 203.1 | 86.1 | 0.18 | 3.3 | 0.20 | 5.9 |
| SI | 219.1 | 60 | 0.10 | 5.4 | 0.087 | 7.6 |
| NI | 246.2 | 229 | 0.077 | 5.0 | 0.087 | 5.8 |
| MV | 249.1 | 104.1 | 0.058 | 1.8 | 0.075 | 4.2 |
| PI | 229.2 | 70.1 | 0.11 | 5.2 | 0.10 | 6.7 |
| GF | 223.1 | 120.1 | 0.071 | 4.3 | 0.071 | 6.9 |
| IY | 295.2 | 86.2 | 0.049 | 1.1 | 0.061 | 4.9 |
| VF | 265.2 | 72.1 | 0.064 | 3.3 | 0.055 | 3.4 |
| IL | 245.2 | 86.1 | 0.10 | 3.9 | 0.067 | 4.8 |
| LF | 279.2 | 86.2 | 0.12 | 2.7 | 0.077 | 5.4 |
| WF | 352.2 | 335 | 0.10 | 2.2 | 0.062 | 6.1 |

Table 2**.** Linearity, and sensitivity of the developed method

| Dipeptide | Q1 (*m/z*) | Q3 (*m/z*) | Calibration curve | LOD (nM) |
|---|---|---|---|---|
| | | | correlation coefficient | |
| **CE-MS/MS** | | | | |
| RR | 166.1 | 70.1 | 0.996 | 83 |
| SH | 243.1 | 110 | 0.998 | 73 |
| HS | 243.1 | 110 | 0.999 | 30 |
| RD | 290.1 | 70.1 | 0.999 | 2.9 |
| HQ | 284.1 | 110.1 | 0.981 | 8.1 |
| HD | 271.1 | 110.1 | 0.999 | 72 |
| AG | 147.1 | 147.1 | 0.999 | 6.0 |
| SA | 177.1 | 60.2 | 0.999 | 55 |
| TA | 191.1 | 74.1 | 0.996 | 55 |
| EA | 219.1 | 201.1 | 0.998 | 6.3 |
| AN | 204.1 | 133.1 | 0.999 | 4.2 |
| GN | 190.1 | 173.1 | 0.999 | 6.4 |
| NS | 220.1 | 140 | 0.999 | 5.9 |
| TS | 207.1 | 74.1 | 0.999 | 79 |
| GE | 205.1 | 84.2 | 0.995 | 4.7 |
| SE | 235.1 | 60.1 | 0.999 | 3.1 |
| WQ | 333.2 | 159 | 0.998 | 6.1 |
| GD | 191.1 | 134.1 | 0.999 | 3.3 |
| SD | 221.1 | 60.1 | 0.999 | 3.4 |
| DD | 249.1 | 133.9 | 0.993 | 3.3 |
| **LC-MS/MS** | | | | |
| MS | 237.1 | 104 | 0.996 | 0.74 |
| VQ | 246.1 | 72.2 | 0.999 | 0.62 |
| GP | 173.1 | 116.1 | 0.999 | 0.55 |
| VG | 175.1 | 55.1 | 0.998 | 3.2 |
| VA | 189.1 | 72.2 | 0.999 | 0.52 |
| MQ | 278.1 | 147.1 | 0.998 | 0.66 |
| AP | 187.1 | 116.1 | 0.995 | 0.39 |
| GV | 175.1 | 129.1 | 0.999 | 3.1 |
| LT | 233.1 | 86 | 0.999 | 0.41 |
| IA | 203.1 | 86.1 | 0.999 | 0.19 |
| SI | 219.1 | 60 | 0.998 | 0.91 |
| NI | 246.2 | 229 | 0.999 | 0.60 |
| MV | 249.1 | 104.1 | 0.998 | 0.28 |
| PI | 229.2 | 70.1 | 0.999 | 0.12 |
| GF | 223.1 | 120.1 | 0.999 | 0.54 |
| IY | 295.2 | 86.2 | 0.999 | 0.31 |
| VF | 265.2 | 72.1 | 0.999 | 0.097 |
| IL | 245.2 | 86.1 | 0.999 | 0.13 |
| LF | 279.2 | 86.2 | 0.999 | 0.088 |
| WF | 352.2 | 335 | 0.999 | 0.37 |

## 4.4.3 Matrix effect investigation in LC-MS/MS

Matrix effect is defined as the effect of co-eluting sample matrix components on the ionization of the target compounds and causes ion suppression and enhancement. I investigated the matrix effect in LC-MS/MS measurement with different liver weight (5, 10, 25 and 50 mg). Recovery rates of 50 mg case and 5 mg case of recovery rate are shown in Figure 10. Figure 10A, 10B show recovery rates are improved in all dipeptides as liver weight becomes small and the results of MS, VG, VA, IA, NI in 50 mg are affected by ion suppression and the result of VQ in 50 mg is affected by ion enhancement. As shown in Figure 10C, the matrix effects of these 6 dipeptides (MS, VG, VA, IA, NI, VQ) are mitigated by dilution and appropriate recovery rate is obtained in 5 or 10 mg.

Therefore, these results indicate matrix effect could affect LC-MS/MS measurement in dipeptide analysis and I diluted sample subjected to LC-MS/MS analysis 5 times in sample preparation to mitigate the matrix effect.

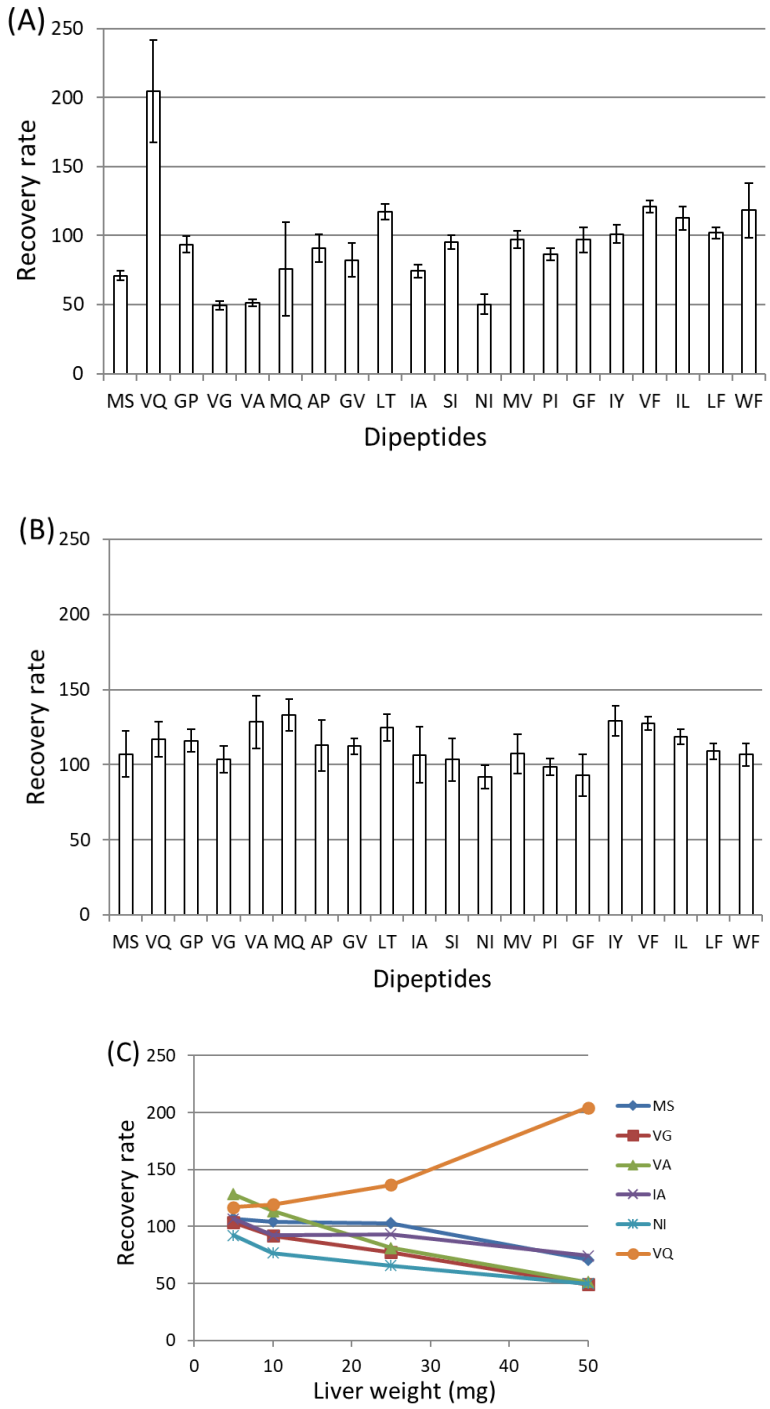



Figure 10. Comparison of matrix effects between 50 mg (A) and 5 mg (B). Dipeptides are arranged in order of retention time. (C) Liver sample weight dependence in 6 dipeptides (MS, VG, VA, IA, NI,VQ) of recovery rate.

## 4.4.4 Spike and recovery test

As mentioned above (section 2.5), the recovery test was performed using chicken liver samples spiked with 40 selected dipeptides (20 for CE-MS/MS at a concentration of 10 μM and 20 for LC-MS/MS at a concentration of 5 μM) either before or after sample preparation. Although most dipeptides showed good recovery (70-135%), several featured poor recoveries (Table 3). The recovery of RR decreased in the case of pre-addition (45.4 ± 4.9%) but not in the case of post-addition (128.2 ± 7.4%), which indicated that this dipeptide decomposed during sample preparation. In contrast, the recoveries of MQ and DD were less than 70% in both cases (MQ: 52.3 ± 3.2 and 60.1 ± 2.0% and DD: 54.3 ± 14.8 and 66.9 ± 8.8% for pre- and post-addition, respectively).

The instability of several amino acids has been widely reported. For example, glutamine and asparagine are easily hydrolyzed and converted to glutamic acid and aspartic acid, respectively, while sulfur-containing amino acids such as methionine, tyrosine, and cysteine are unstable in aqueous solution and are easily oxidized to form disulfides [50].

Table 3. Recoveries of dipeptides from spiked liver samples

| Dipeptide | Pre-addition recovery | Post-addition recovery |
|---|---|---|
| **CE-MS/MS** | | |
| RR | 45.4 ± 4.9 | 128.2 ± 7.4 |
| SH | 94.3 ± 8.9 | 91.5 ± 8.7 |
| HS | 112.3 ± 2.3 | 101.9 ± 6.1 |
| RD | 114.7 ± 1.0 | 110.6 ± 12.4 |
| HQ | 128.1 ± 2.8 | 95.8 ± 9.1 |
| HD | 132.3 ± 2.7 | 116.4 ± 7.5 |
| AG | 98.6 ± 7.2 | 72.3 ± 4.5 |
| SA | 133.9 ± 6.0 | 107.8 ± 2.3 |
| TA | 111.9 ± 12.5 | 110.5 ± 13.9 |
| EA | 134.5 ± 4.6 | 111.7 ± 4.0 |
| AN | 124.8 ± 2.6 | 106.0 ± 8.0 |
| GN | 102.4 ± 7.0 | 99.6 ± 18.0 |
| NS | 111.5 ± 2.0 | 95.9 ± 3.8 |
| TS | 109.9 ± 3.0 | 95.9 ± 4.6 |
| GE | 113.1 ± 3.1 | 98.7 ± 6.9 |
| SE | 110.0 ± 8.0 | 100.6 ± 5.7 |
| WQ | 119.3 ± 8.0 | 104.2 ± 8.5 |
| GD | 112.5 ± 7.8 | 103.5 ± 5.5 |
| SD | 102.9 ± 9.4 | 89.0 ± 3.6 |
| DD | 54.3 ± 14.8 | 66.9 ± 8.8 |
| **LC-MS/MS** | | |
| MS | 81.2 ± 3.3 | 78.7 ± 2.2 |
| VQ | 101.8 ± 4.3 | 91.0 ± 2.1 |
| GP | 110.4 ± 6.7 | 104.4 ± 6.3 |
| VG | 104.1 ± 3.0 | 99.7 ± 10.7 |
| VA | 87.5 ± 5.9 | 82.1 ± 3.5 |
| MQ | 52.3 ± 3.2 | 60.1 ± 2.0 |
| AP | 85.5 ± 3.5 | 80.8 ± 4.5 |
| GV | 100.0 ± 3.1 | 92.0 ± 1.4 |
| LT | 96.7 ± 2.6 | 95.2 ± 0.9 |
| IA | 105.4 ± 5.8 | 97.6 ± 3.1 |
| SI | 89.9 ± 7.5 | 99.2 ± 2.0 |
| NI | 87.2 ± 8.7 | 83.1 ± 2.5 |
| MV | 92.7 ± 1.6 | 105.0 ± 2.6 |
| PI | 99.7 ± 7.8 | 99.5 ± 2.2 |
| GF | 118.0 ± 5.7 | 110.4 ± 3.1 |
| IY | 117.6 ± 5.9 | 109.0 ± 2.6 |
| VF | 110.4 ± 6.1 | 104.8 ± 1.5 |
| IL | 107.6 ± 4.0 | 102.4 ± 2.9 |
| LF | 113.1 ± 6.1 | 110.2 ± 3.8 |
| WF | 112.8 ± 6.9 | 114.0 ± 4.0 |

In fact, I observed that the recoveries of di- and tripeptides containing cysteine and homocysteine from serum and liver are extremely poor [27]. Pyroglutamate is spontaneously formed from glutamine and glutamic acid, and the cyclization of N-terminal glutamine to pyroglutamate has been observed for recombinant monoclonal antibodies [51]. Thus, the low recoveries of some dipeptides studied herein were ascribed to the presence of the abovementioned amino acids and could presumably be increased through the use of isotopically labeled dipeptides.

## 4.5 Diet experiment of mouse

The developed method was applied to the quantitation of dipeptides in the liver of mice fed either a normal or a high-fat diet, and 96 and 164 dipeptides (24 were overlapped) were determined by CE-MS/MS and LC-MS/MS, respectively. The content of each dipeptide in 1 g of liver tissue was determined by normalizing the peak area of each dipeptide with respect to the area of the internal standard and by using standard calibration curves. The differences in the dipeptide profiles obtained for the two diets were revealed by PCA (Figure 11). In the PCA score plot, mice fed a high-fat diet were adequately separated from those fed a normal diet for the first two principal components, which suggests that diet affected the dipeptide profile of mouse liver.

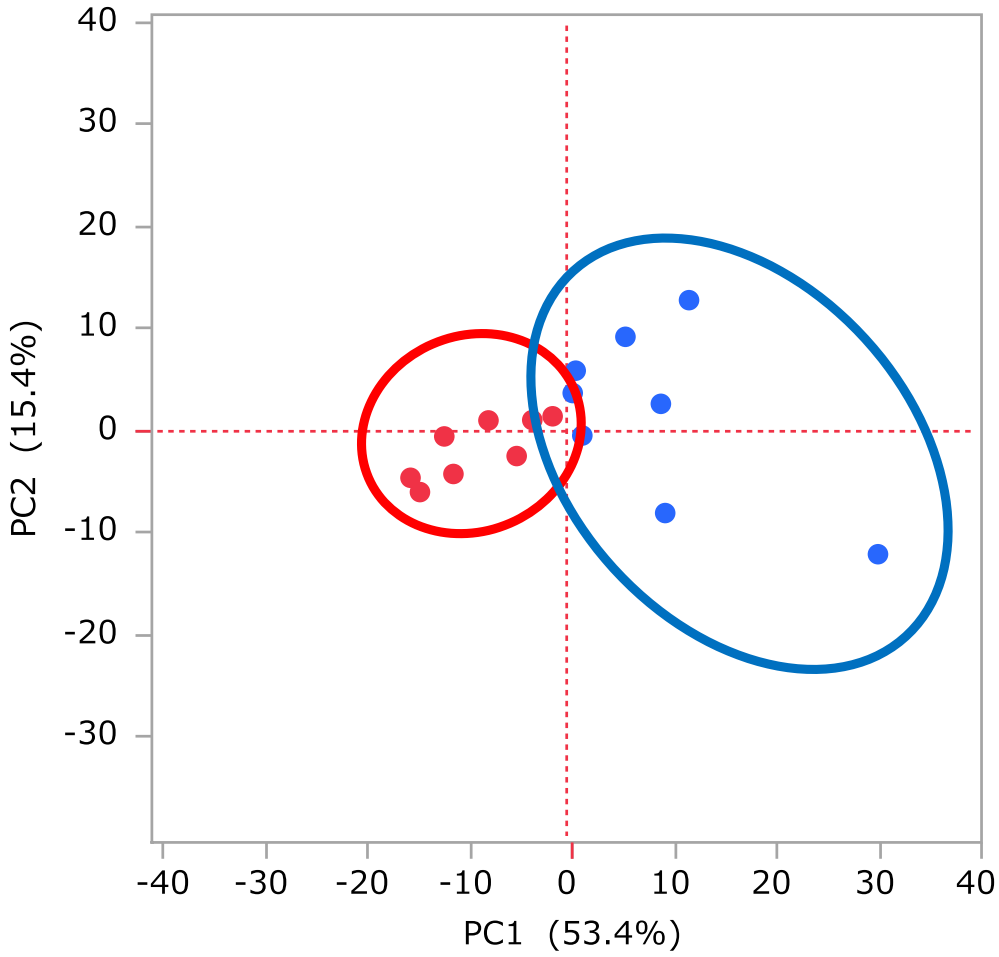



Figure 11. PCA score plot of dipeptide profiles obtained for the livers of mice fed normal (blue symbols) and high-fat (red symbols) diets. Mice fed a high-fat diet were adequately separated from those fed a normal diet for the first two principal components, which suggests that diet affected the dipeptide profile of mouse liver.

## *4. Development of comprehensive dipeptide analysis method*

Next, volcano plots were used to determine the dipeptides for which the contents significantly changed upon going from a normal diet to a high-fat one (Figure 12). The shift to a high-fat diet was shown to decrease the levels of 29 dipeptides ($p < 0.05$, fold-change < 0.5) and increase the levels of none. It is worth mentioning that the levels of many metabolites have been positively or negatively correlated with obesity, e.g., high-fat diets are known to increase the levels of lipid metabolites while decreasing those of lipid metabolism intermediates through abnormal lipid and energy metabolism [52]. Carnosine and anserine belong to the family of histidyl dipeptides and are abundant in the skeletal muscles of vertebrates [53]. The contents of these dipeptides, which scavenge reactive oxygen and carbonyl species, are known to be decreased in the liver of obese and diabetic rodents [54,55]. In the present study, the levels of SH, HA, HN, and HG were significantly decreased in the liver of mice on a high-fat diet. Although I have not measured the levels of carnosine and anserine, the decrease in the levels of the above four histidyl dipeptides was ascribed to their consumption for the removal of reactive oxygen and carbonyl species generated by the administration of a high-fat diet. Although the functions of the other 25 dipeptides (29 – 4 = 25) are not well understood, some of them may be associated with obesity and used as biomarkers of obesity-related diseases.

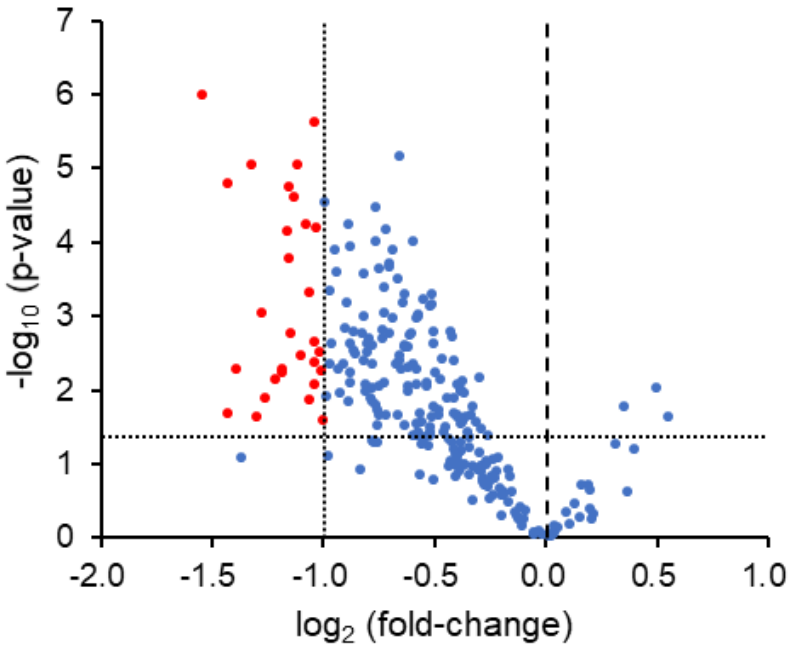



Figure 12**.** Volcano plot of differential dipeptides for different diets. For each dipeptide, $-\log_{10}$ (*p*-value) was plotted vs. $\log_2$ (fold-change). The shift to a high-fat diet significantly decreased the levels of 29 dipeptides in mouse liver ($p < 0.05$ and fold-change < 0.5). Student's *t*-test was used for statistical analysis.

# Chapter 5

## Elucidation of cancer-specific profile of liver cancer

### 5.1 Clinical characteristics of patients

This study was approved by the Tokushima University Hospital Ethics Committee (Approved no. 1815), and the corresponding regulatory agencies and all experiments were carried out in accordance with approved guidelines. All patients involved in the study signed an informed consent form.

Tumor and surrounding non-tumor liver tissues were surgically resected from 13 patients with HCC and 3 patients with metastatic liver cancer. The resection of these liver tissues is conducted in Tokushima University Hospital. The clinical information of the patients is listed in Table 4.

## *5. Elucidation of cancer-specific profile of liver cancer*

Table 4. Clinical characteristics of patients in this study

| ID | Age | Sex | BMI | Type | Virus | Stage |
|---|---|---|---|---|---|---|
| 1 | 68 | Female | 31.9 | HCC | Non B/C | III |
| 2 | 73 | Female | 23.1 | HCC | Non B/C | II |
| 3 | 73 | Male | 17.3 | HCC | Non B/C | II |
| 4 | 72 | Male | 24.7 | HCC | Non B/C | II |
| 5 | 63 | Male | 18.2 | HCC | Non B/C | II |
| 6 | 67 | Male | -* | HCC | Non B/C | III |
| 7 | 44 | Male | 22.3 | HCC | HBV | II |
| 8 | 65 | Male | 19.0 | HCC | HBV | III |
| 9 | 61 | Male | 21.8 | HCC | HCV | II |
| 10 | 57 | Male | 21.5 | HCC | HCV | I |
| 11 | 78 | Female | 20.5 | HCC | HCV | III |
| 12 | 75 | Male | 25.4 | HCC | HCV | II |
| 13 | 60 | Male | 21.8 | HCC | HCV | IVB |
| 14 | 78 | Female | 22.1 | MLC | Non B/C | |
| 15 | 68 | Female | 23.3 | MLC | Non B/C | |
| 16 | 80 | Male | 26.7 | MLC | Non B/C | |

BMI, body mass index; HCC, hepatocellular carcinoma; MLC, metastatic liver cancer; Non B/C, non B and non C hepatitis; HBV, hepatitis B virus; HCV, hepatitis C virus, *, missing value

### *5. Elucidation of cancer-specific profile of liver cancer*

All liver cancer cases were the first instance of cancer, and the patients had no treatment history prior to surgery. The mean ± standard deviation of the patients' age and body mass index were 67.6 ± 9.3 years and 22.6 ± 3.6, respectively. The HCC group contained 6 non B/C samples, 2 HBV samples, and 5 HCV samples. The stages of cancer in HCC varied from I to IVB.

In this study, 140 and 96 dipeptides were detected in the liver by LC-MS/MS and CE-MS/MS, respectively. Among the total of 236 dipeptides detected in both methods, the peak of 14 dipeptides could not be distinguished by MRM transition and retention time. The amount of each dipeptide was standardized to nmol/g liver tissue, and subsequent analysis was performed using this value.

## 5.2 Outlier analysis

First, PCA was performed to identify trends in the dipeptide profiles of all samples measured in this study (Figure 13).

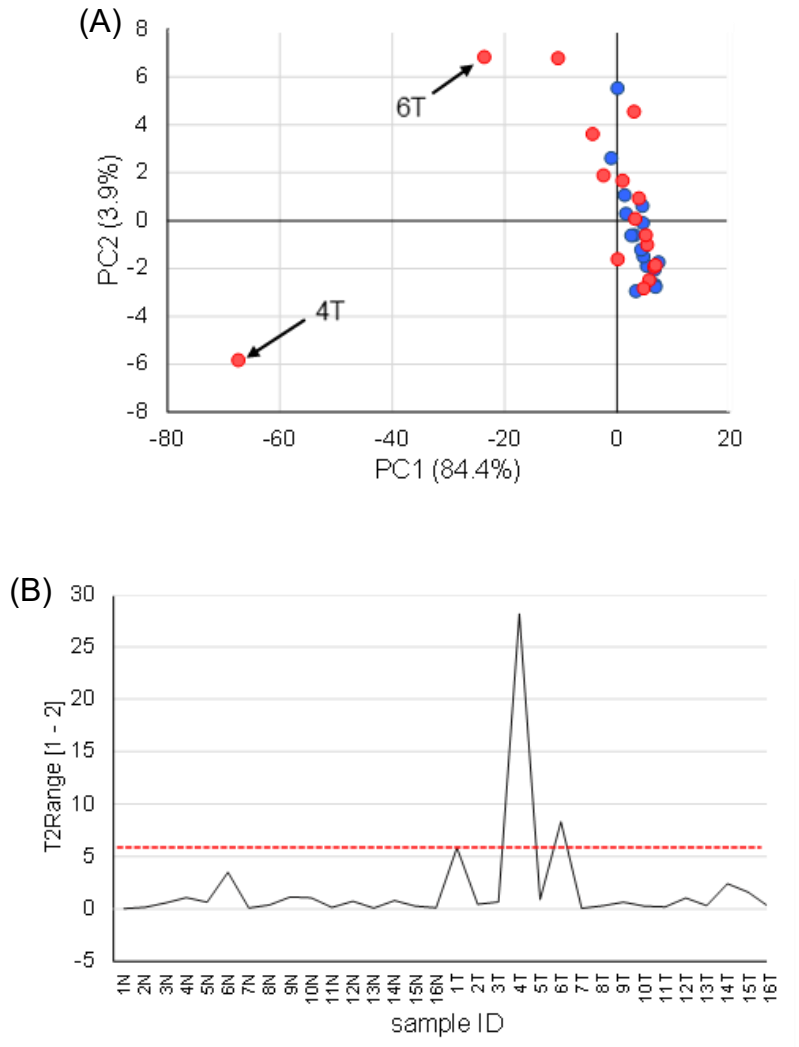



Figure 13. Outlier analysis based on quantified dipeptides obtained from patients with liver cancer. (A) Principal component analysis score plots using the auto-scaled dipeptide data of paired non-tumor (blue) and tumor (red) tissues. The contribution ratios were 84.4% and 3.9% for PC1 and PC2, respectively. (B) Hotelling's $T^2$ range plot of all samples (equation (33)). The red dashed line indicates the upper control limit (equation (37), $\alpha$= 0.01).

The results of the PCA score plots (Figure 13A) suggested that two tumor tissues were supposed outliers. Therefore, outlier analysis was carried out to investigate whether outliers were included in the measured samples. Figure 13B shows the results of outlier analysis using Hotelling's $T^2$ statistics in PCA. As Hotelling's $T^2$ statistics of two tumor tissues (No. 4T and 6T) exceeded the upper control limit at a significance level of 0.01, these samples were considered as outliers and excluded from subsequent analysis.

## 5.3 Principal component analysis of all non-tumor and tumor tissues

Next, PCA based on Pareto scaling was performed using the remaining 14 samples (each sample contained non-tumor and tumor tissues, respectively), excluding samples showing outliers (Figure 14).

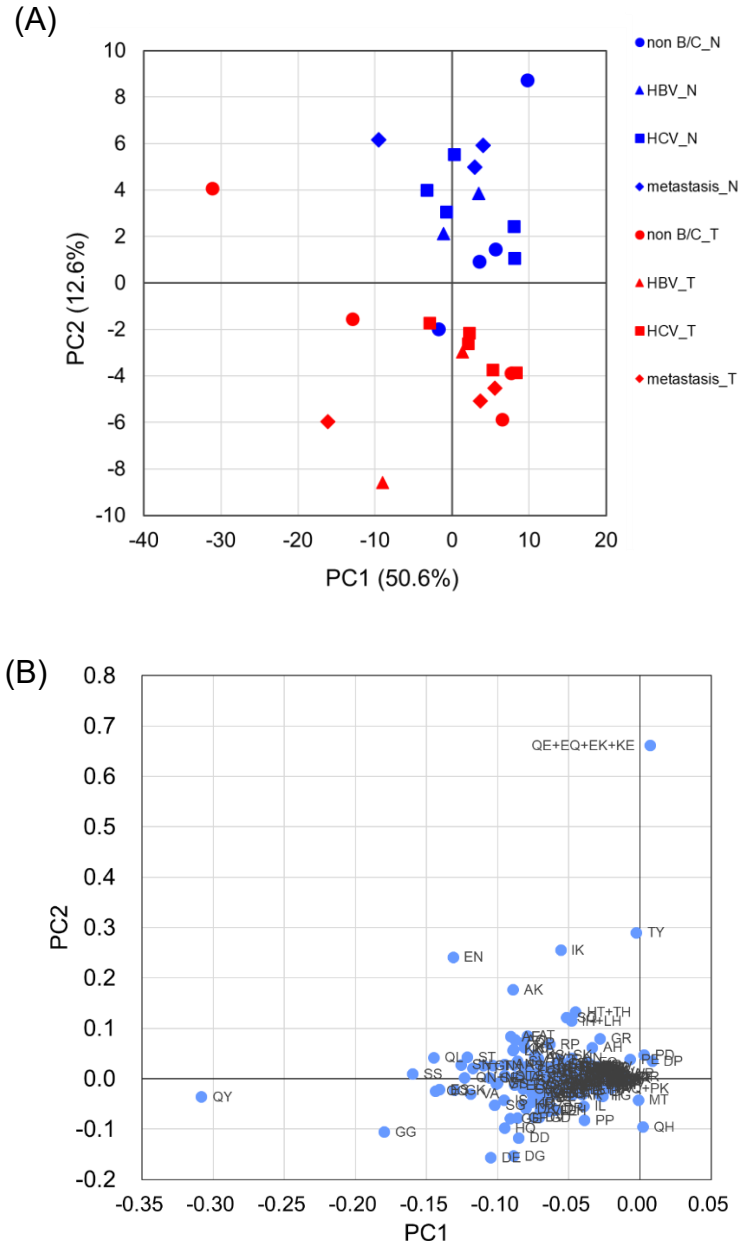



Figure 14. Principal component analysis using the Pareto-scaled dipeptide data. (A) PCA score plots of paired non-tumor (blue) and tumor (red) tissues in different type of liver cancers. The contribution ratios were 50.6% and 12.6% for PC1 and PC2, respectively. (B) PCA loading plots of dipeptides on the first two principal components.

In the PCA score plot (Figure 14A), non-tumor and tumor tissues were sufficiently separated mainly by principal component 2. In addition, QE (Gln-Glu) + EQ (Glu-Gln) + EK (Glu-Lys) + KE (Lys-Glu) (overlapped peak) and TY (The-Tyr), IK (Ile-Lys), and EN (Glu-Asn) showed relatively large values in principal component 2 of the loading plot (Figure 14B), suggesting that these dipeptides contributed to the separation of non-tumor and tumor tissues.

## 5.4 Characteristics of dipeptides detected in liver tissue

To determine the characteristics of the dipeptides detected in each sample, grouping was performed for each amino acid constituting the N-terminus (Figure 15A) and C-terminus (Figure 15B).

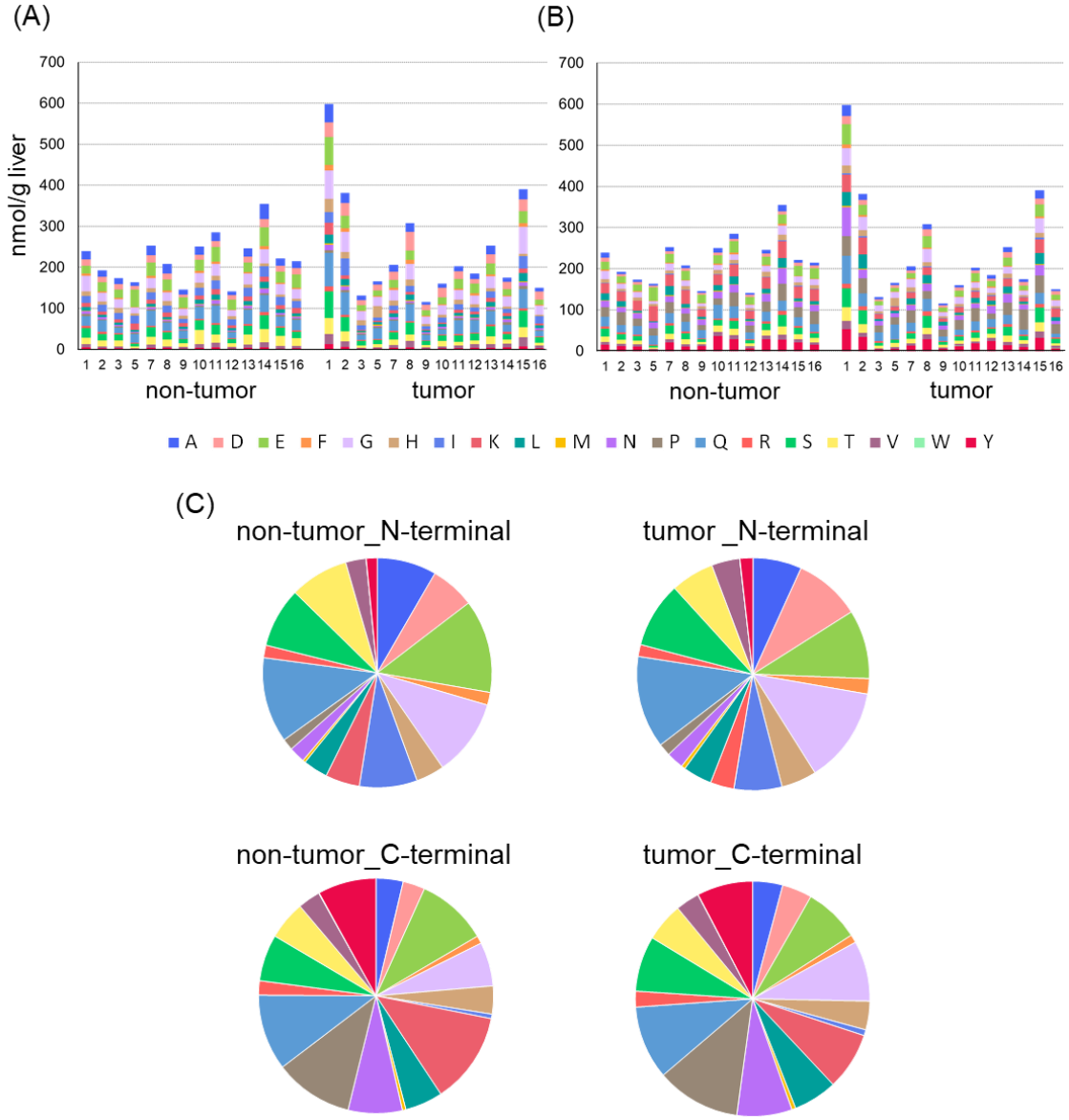



Figure 15. Amino acid composition of N-terminus (A) and C-terminus (B) in dipeptide detected from each liver tissue. The columns represent the number of moles (nmol/g liver). (C) Difference in average composition ratio of N-terminal and C-terminal amino acids in each tissue.

Relatively high accumulation of dipeptide was observed in the non-tumor tissue of sample ID14 and tumor tissues of samples ID1, 2, and 15, but no obvious difference was observed in the amino acid composition of these samples compared to the other samples. This trend was similar between non-tumor and tumor tissues. The amino acid compositions at the C- and N-termini significantly differed (Figure 15C). For example, dipeptides containing alanine (A), aspartic acid (D), and isoleucine (I) were predominant at the N-terminus, whereas dipeptides containing lysine (K), asparagine (N), proline (P), and tyrosine (Y) were increased at the C-terminus.

## 5.5 Dipeptide profile with and without hepatitis in hepatocellular carcinoma

The dipeptide profiles of non-hepatitis- and hepatitis-derived HCC were compared. Volcano plots were prepared using the quantitated values of dipeptides detected in non-hepatitis (Non B/C, n = 4) and hepatitis (HBV or HCV n = 7) samples form tumor (Figure 16A) and non-tumor tissues (Figure 16B) of HCC, respectively.

In tumor tissues, only the amount of HT+TH was significantly decreased ($p < 0.05$, fold-change < 0.67) in 6 dipeptides in hepatitis samples, including VI, IY, IE, TI, VN, and VY, was observed in non-tumor tissues. According to the results of the volcano plot, the number of dipeptides that was increased in hepatitis samples was slightly higher in

nontumor tissues, whereas this number was mostly decreased in tumor tissues.

Finally, I investigated the relationship of the dipeptide abundance between non-tumor and tumor tissues with and without hepatitis. Figure 16C shows the amounts of dipeptides in both tissues; a significant different was observed between tissues. One of the dipeptides that showed significance in tumor tissue showed no significance in non-tumor tissue, whereas six of the dipeptides that showed significance in non-tumor tissue showed no significance in tumor tissue.

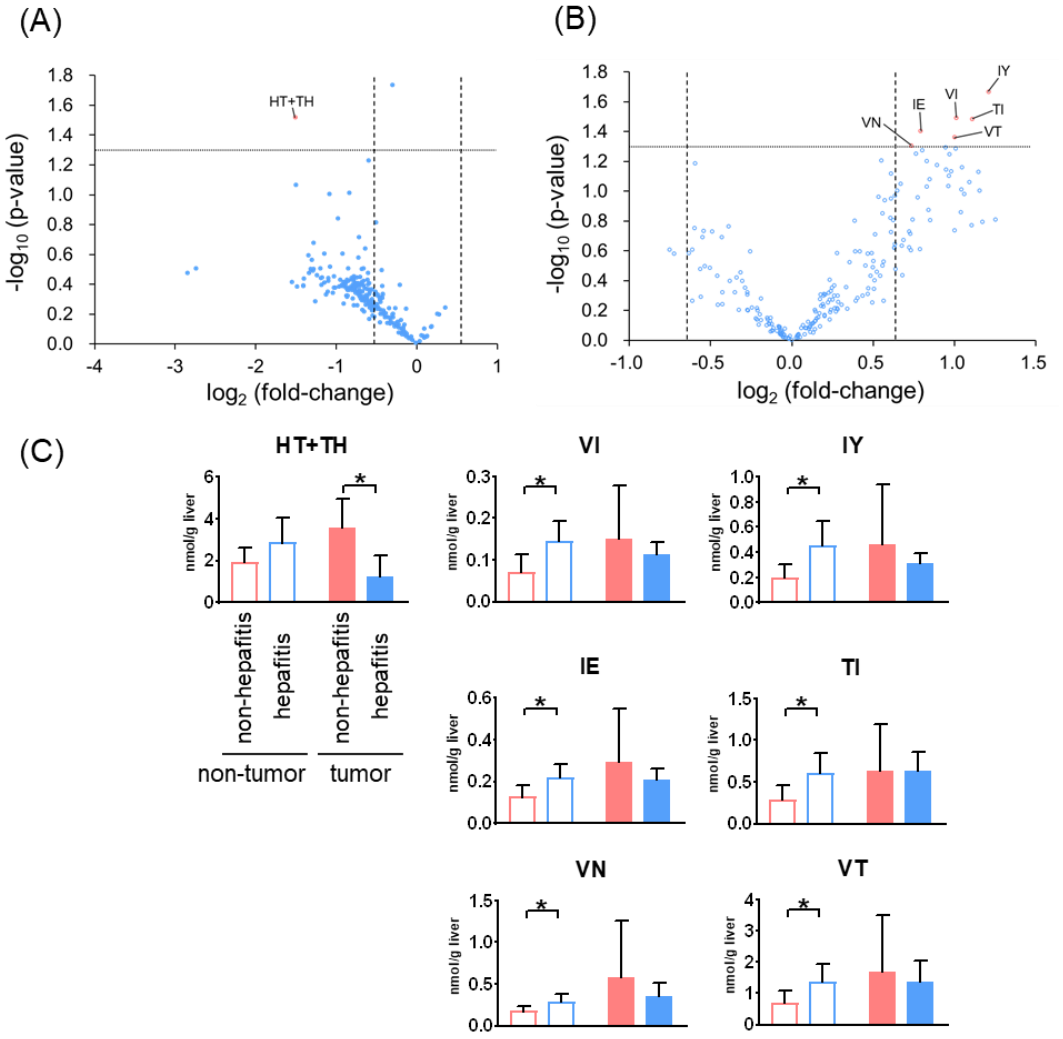



Figure 16. Volcano plot of differential dipeptides in tumor (A) and non-tumor (B) tissues with and without hepatitis in hepatocellular carcinoma. For each dipeptide, the -$\log_{10}$($p$-value) was plotted vs $\log_2$(fold-change). (C) Dipeptides showing a significant difference in the volcano plot ($p < 0.05$ and fold-change < 0.67 or > 1.5). The columns represent the average number of moles (nmol/g tissue) and error bars indicate the standard deviation. *$p < 0.05$. The Mann-Whitney $U$-test was used for statistical analysis.

## 5.6 Discussion

In this study, I performed comprehensive dipeptide analysis of paired tumors and surrounding non-tumor liver tissues obtained from 13 patients with HCC and 3 patients with metastatic liver cancer. I successfully quantified 236 dipeptides (14 overlapped) with our previously developed method by LC-MS/MS and CE-MS/MS [13]. This is the largest number of dipeptides detected in the liver tissue to date.

Unlike analysis of body fluids such as blood and urine, it is necessary to consider variability between samples when performing metabolome analysis of tissues. Particularly, tumor tissues exhibit cancer heterogeneity, and the amounts of metabolites and dipeptides may greatly differ depending on the sample location. Therefore, effectively removing outliers is useful for searching for effective biomarkers in subsequent analysis. In this study, PCA using auto-scaled dipeptide data revealed that the score plots of samples 4T and 6T tended to deviate from the score plots of the other samples. Therefore, in outlier analysis with Hotelling's $T^2$ statistics, these two samples exceeded the upper control limit at a significance level of 0.01, and were therefore excluded from subsequent analysis.

PCA was performed again on the remaining 14 non-tumor and tumor tissues using Pareto-scaled dipeptide data, which showed that non-tumor and tumor tissues were

separated for principal component 2. In contrast, no significant difference was observed depending on the factors causing liver cancer. These results demonstrate that the difference in the dipeptide profile in the liver depends on whether it is a non-tumor- or tumor-derived tissue and does not depend on the factors causing cancer. This finding is similar to the serum metabolic profiles obtained by LC-MS from two HCC cohorts infected with HBV or HCV [56].

The characteristics of the detected dipeptides were also examined. Comparison of the total amount of dipeptides in non-tumor and tumor tissues in each specimen showed that 7 of 14 specimens showed an increase and the others showed a decrease in the amounts of dipeptides in the tumor tissue. When amino acids in various tumor tissues were measured by metabolome analysis [4,57], the samples showed different tendencies, possibly because of the heterogeneity of cancer. In addition, no significant difference was found in the amino acid composition between non-tumor and tumor tissues, whereas a significant difference was observed at the N- and C-termini. Studies have reported that γ-glutamylcysteine synthetase [28] and non-ribosomal peptide synthetase [58] are dipeptide synthetases; however, proteolysis is common during dipeptide generation. Therefore, substrate specificity of the peptidase could be the main cause for this. For example, chymotrypsin, which is a serine protease, specifically cleaves the

C-terminal of aromatic amino acids. In addition, elastase, which has a shallow bottom in the substrate binding pocket, specifically cleaves the peptide bond at the C-terminus of amino acids with small side chains. However, various factors require further analysis, such as where the dipeptides detected in the liver tissue are produced. Bioinformatics-based proteome analysis could help reveal the protease involved.

Although there are several routes to HCC development, HBV and HCV infections are the most important risk factors [59]. Therefore, investigating the differences in dipeptide profiles depending on the presence or absence of hepatitis is important for understanding the process of HCC development. In addition, determining the differences in dipeptide profiles will enable identification of biomarker candidates for future evaluation using biofluids. Nevertheless, there are limited numbers of studies searching for biomarkers by measuring dipeptides in body fluids. Recently, the dipeptides, such as hydroxyproline-Leu, EW and FF, have been selected as promising predictive biomarkers for the diagnosis of epithelial ovarian cancer [60]. In this study, analysis of dipeptides with and without hepatitis revealed that 7 dipeptides were significantly changed ($p < 0.05$, fold-change $< 0.67$ or $> 1.5$). Among the tumor tissues, HT+TH was significantly decreased in hepatitis samples, and most other dipeptides also tended to decrease. In contrast, in non-tumor tissues, VI, IY, IE, TI, VN, and VT were

significantly increased in hepatitis samples, and many other dipeptides also tended to increase. For the 7 dipeptides showing significant differences, the amounts in corresponding tissues were also examined, but none of the dipeptides showed a common change between tissues. Thus, hepatitis-derived HCC showed a characteristic tendency, with the amount of dipeptide increased in the surrounding non-tumor tissue but decreased in the tumor tissue. This suggests that the change in the dipeptide profile due to the presence or absence of hepatitis already occurred before tumor formation and was maintained throughout the production of different dipeptides and metabolic mechanisms even after tumor generation.

# Chapter 6

## Concluding remarks

In my present study, I approached understanding the characteristics of liver cancer from metabolome analysis, which can reveal the metabolite profile *in vivo*. Dipeptide, two amino acids linked by a peptide bond, are important metabolites in metabolome analysis and are recently attracting attention as functional biomaterials as well as biomarkers of disease. The analysis of dipeptides in tumor tissue could lead to the discovery of new metabolic mechanisms and biomarkers in cancer.

Although previous studies have analyzed some dipeptides there has been a lack of a comprehensive analysis. The reason being that all dipeptides, except those composed of the same amino acid, have structural isomers with opposite amino acid binding orders, making it difficult to separate the many structural isomers by a single analytical method. Therefore we worked on the development of comprehensive method for analyzing dipeptides.

An analytical platform for the comprehensive analysis of dipeptides was developed, featuring two methods with different separation principles. The designed technique was used to analyze 110 and 225 dipeptides by CE-MS/MS and LC-MS/MS, respectively,

and was finally employed to determine 335 dipeptides (and separate them from their structural isomer). Good repeatability, linearity, and recovery were observed, and the instrumental detection limit of the tested dipeptides (<83.1 nM) and acceptable recoveries (70-135%) suggested that the developed method is acceptable for dipeptide profiling. The presented approach was applied to quantitatively determine dipeptide level changes in mouse liver due to a shift from a normal diet to a high-fat one. Our technique using multiple different separation principles could apply other overlapped peaks of other metabolites and our method can be used in combination with other dipeptide measurement techniques, such as dipeptide included β-amino acid and γ-glutamyl dipeptide measurement, and other sophisticated techniques for measuring specific dipeptides, such as derivatization to elucidate various dipeptide profiles and detect important dipeptides. Our method was introduced in Yamagata newspaper (2020/7/4).

Furthermore, I applied our method to liver cancers including HCC because the analyzing dipeptides in HCC could lead to new discovery of cancer-specific feature of HCC. HCC is the third leading cause of cancer-related deaths worldwide and a major cause of HCC is chronic hepatitis caused by hepatitis virus infection. Hence, elucidation of cancer-specific feature of HCC pathogenesis such as hepatitis is of very importance

for cancer treatment. I detected 236 dipeptides in liver cancer tissue using a comprehensive dipeptide analytical method involving CE-MS/MS and LC-MS/MS. Similar to previously reported metabolite results, although the dipeptide profiles between non-tumor and tumor tissues differed, no clear difference was observed in the etiological comparison. I also found that the N- and C-terminal amino acid compositions of the detected dipeptides significantly differed in both tissues, suggesting the substrate specificity of enzyme proteins, such as serine protease and these results could be related to activity of the zymogen of the pancreas. Comparison of the dipeptide profiles depending on the presence or absence of hepatitis suggested that hepatitis-derived cancer has a characteristic dipeptide profile before and after tumor onset. Studies on dipeptide analysis in disease are limited because our method has been recently developed. Detailed investigation of liver cancer, such as HCC, and other liver diseases, such as cirrhosis and nonalcoholic steatohepatitis, as well as other diseases is warranted in the future to elucidate the role of dipeptides in various diseases.

In this study, we developed an important platform for dipeptide analysis. This method enables detection of many types of dipeptides precisely without the need for derivatization. Moreover, our method can be applied to various samples and is expected to facilitate the search for new dipeptide applications, such as functional components of

foods and disease biomarkers. Our method has great application in various fields, such as agriculture and medicine. Dipeptides have attracted attention as post-amino acids with physical properties and functions different from those of amino acids. The liver cancer study, which used our method, provides insights into HCC pathogenesis and may help identify novel biomarkers for diagnosis. Moreover, these novel approaches will help elucidate characteristics of various cancers. My doctoral dissertation will open new doors in the search for novel cancer profiles, biomarkers, therapeutic targets, and functional components of foods.

# Acknowledgements

This work was supported by many people. My chief examiner Masaru Tomita and my sub-chief examiner Tomoyoshi Soga advised me about this work and offered me a good environment to study. My sub-chief examiner Akio Kanai advised me about this thesis. Department of Nephrology, Nagoya University Graduate School of Medicine provided livers of mice on normal and high-fat diets. Department of Clinical Nutrition and Food Management, Graduate School of Biomedical Sciences, Tokushima University Graduate School provided livers of patients with cancers. My sub-chief examiner Akiyoshi Hirayama taught me experimental procedures and data processing. Research staff Takamasa Ishikawa and technical staff of the metabolome research group helped me measure biological samples. I would like to thank everyone involved in this work and also thank my parents.

# References

[1] Wu, H.; Xue, R.; Tang, Z.; Deng, C.; Liu, T.; Zeng, H.; Sun, Y.; Shen, X. Metabolomic investigation of gastric cancer tissue using gas chromatography/mass spectrometry. *Anal. Bioanal. Chem.* **2010**, *396*, 1385-1395.

[2] Huang, Q.; Tan, Y.; Yin, P.; Ye, G.; Gao, P.; Lu, X.; Wang, H.; Xu, G. Metabolic characterization of hepatocellular carcinoma using nontargeted tissue metabolomics. *Cancer Res.* **2013**, *73*, 4992-5002.

[3] Ferrarini, A.; Di Poto, C.; He, S.; Tu, C.; Varghese, R. S.; Kara Balla, A.; Jayatilake, M.; Li, Z.; Ghaffari, K.; Fan, Z., et al. Metabolomic Analysis of Liver Tissues for Characterization of Hepatocellular Carcinoma. *J. Proteome Res.* **2019**, *18*, 3067-3076.

[4] Kami, K.; Fujimori, T.; Sato, H.; Sato, M.; Yamamoto, H.; Ohashi, Y.; Sugiyama, N.; Ishihama, Y.; Onozuka, H.; Ochiai, A., et al. Metabolomic profiling of lung and prostate tumor tissues by capillary electrophoresis time-of-flight mass spectrometry. *Metabolomics* **2013**, *9*, 444-453.

[5] Sreekumar, A.; Poisson, L. M.; Rajendiran, T. M.; Khan, A. P.; Cao, Q.; Yu, J.; Laxman, B.; Mehra, R.; Lonigro, R. J.; Li, Y., et al. Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature* **2009**, *457*, 910-914.

[6] Budczies, J.; Denkert, C.; Muller, B. M.; Brockmoller, S. F.; Klauschen, F.; Gyorffy, B.; Dietel, M.; Richter-Ehrenstein, C.; Marten, U.; Salek, R. M., et al. Remodeling of central metabolism in invasive breast cancer compared to normal breast tissue – a GC-TOFMS based metabolomics study. *BMC Genomics* **2012**, *13* 334.

[7] Ishikawa, S.; Sugimoto, M.; Kitabatake, K.; Sugano, A.; Nakamura, M.; Kaneko, M.; Ota, S.; Hiwatari, K.; Enomoto, A.; Soga, T., et al. Identification of salivary metabolomic biomarkers for oral cancer screening. *Sci. Rep.* **2016**, *6*, 31520.

[8] Hori, S.; Nishiumi, S.; Kobayashi, K.; Shinohara, M.; Hatakeyama, Y.; Kotani, Y.; Hatano, N.; Maniwa, Y.; Nishio, W.; Bamba, T., et al. A metabolomic approach to lung cancer. *Lung Cancer* **2011**, *74*, 284-292.

[9] Baniasadi, H.; Gowda, G. A.; Gu, H.; Zeng, A.; Zhuang, S.; Skill, N.; Maluccio, M.; Raftery, D. Targeted metabolic profiling of hepatocellular carcinoma and hepatitis C using LC-MS/MS. *Electrophoresis* **2013**, *34*, 2910-2917.

[10] Chen, F.; Xue, J.; Zhou, L.; Wu, S.; Chen, Z. Identification of serum biomarkers of hepatocarcinoma through liquid chromatography/mass spectrometry-based metabolomic method. *Anal. Bioanal. Chem.* **2011**, *401*, 1899-1904.

[11] Gao, R.; Cheng, J.; Fan, C.; Shi, X.; Cao, Y.; Sun, B.; Ding, H.; Hu, C.; Dong, F.; Yan, X. Serum Metabolomics to Identify the Liver Disease-Specific Biomarkers for

Progression of Hepatitis to Hepatocellular Carcinoma. *Sci. Rep.* **2015**, *5*, 18175.

[12] Han, J.; Qin, W. X.; Li, Z. L.; Xu, A. J.; Xing, H.; Wu, H.; Zhang, H.; Wang, M. D.; Li, C.; Liang, L., et al. Tissue and serum metabolite profiling reveals potential biomarkers of human hepatocellular carcinoma. *Clin. Chim. Acta* **2019**, *488*, 68-75.

[13] Ozawa, H.; Hirayama, A.; Ishikawa, T.; Kudo, R.; Maruyama, M.; Shoji, F.; Doke, T.; Ishimoto, T.; Maruyama, S.; Soga, T., et al. Comprehensive Dipeptide Profiling and Quantitation by Capillary Electrophoresis and Liquid Chromatography Coupled with Tandem Mass Spectrometry. *Anal. Chem.* **2020**, *92*, 9799-9806.

[14] Tang, Y.; Li, R.; Lin, G.; Li, L. PEP search in MyCompoundID: Detection and Identification of Dipeptides and Tripeptides Using Dimethyl Labeling and Hydrophilic Interaction Liquid Chromatography Tandem Mass Spectrometry. *Anal. Chem.* **2014**, *86*, 3568-3574.

[15] Fonteh, A. N.; Harrington, R. J.; Tsai, A.; Liao, P.; Harrington, M. G. Free amino acid and dipeptide changes in the body fluids form Alzheimer's disease subjects. *Amino Acids* **2007**, *32*, 213-224.

[16] Saoi, M.; Percival, M.; Nemr, C.; Li, A.; Gibala, M.; Britz-McKibbin, P. Characterization of the Human Skeletal muscle Metabolome for Elucidation the Mechanisms of Bicarbonate Ingestion on Strenuous Interval Exercise. *Anal. Chem.*

**2019**, *91*, 4709-4718.

[17] Ano, Y.; Kita, M.; Kitaoka, S.; Furuyashiki, T. Leucine-Histidine Dipeptide Attenuates Microglial Activation and Emotional Disturbances Induced by Brain Inflammation and Repeated Social Defeat Stress. *Nutrients* **2019**, *11*, 2161-2173.

[18] Mizushige, T.; Kanegawa, N.; Yamada, A.; Ota, A.; Kanamoto, R.; Ohinata, K. Aromatic amino acid-leucine dipeptides exhibit anxiolytic-like activity in young mice. *Neurosci. Lett.* **2013**, *543*, 126-129.

[19] Kanegawa, N.; Suzuki, C.; Ohinata, K. Dipeptide Tyr-Leu (YL) Exhibits Anxiolytic-Like Activity after Oral Administration via Activating Serotonin 5-HT1A, Dopamine D1 and GABA Receptors in Mice. *FEBS Lett.* **2010**, *584*, 599-604.

[20] Kawasaki, T.; Seki, E.; Osajima, K.; Yoshida, M.; Asada, K.; Matsui, T.; Osajima, Y. Antihypertensive Effect of Valyl-Tyrosine, a Short Chain Peptide Derived from Sardine Muscle Hydrolyzate, on Mild Hypertensive Subjects. *J. Hum. Hypertens.* **2000**, *14*, 519-523.

[21] Takagi, H.; Shiomi, H.; Ueda, H.; Amano, H. A Novel Analgesic Dipeptide from Bovine Brain is a Possible Met-Enkephalin Releaser. *Nature* **1979**, *282*, 410-412.

[22] Wu, M.; Xu, Y.; Fitch, W. L.; Zheng, M.; Merritt, R. E.; Shrager, J. B.; Zhang, W.; Dill, D. L.; Peltz, G.; Hoang, C. D. Liquid Chromatography/Mass Spectrometry

Methods for Measuring Dipeptide Abundance in Non-Small-Cell Lung Cancer. *Rapid Commun. Mass Spectrom.* **2013**, *27*, 2091-2098.

[23] Gudasheva, T. A.; Deeva, O. A.; Mokrov, G. V.; Yarkov, S. A.; Yarkova, M. A.; Seredenin, S. B. The first dipeptide ligand of translocator protein: Design and anxiolytic activity. *Dokl. Biochem. Biophys.* **2015**, *464*, 290-293.

[24] Vasconcelos, S. N.; Drewes, C. C.; de Vinci Kanda Kupa, L.; Farsky, S. H.; Stefani, H. A. Evaluation of toxicity on epithelial and tumor cells of biaryl dipeptide tyrosines. *Eur. J. Med. Chem.* **2016**, *114*, 1-7.

[25] Cimlova, J.; Kruzberska, P.; Svagera, Z.; Husek, P.; Simek, P. In situ derivatization-liquid liquid extraction as a sample preparation strategy for the determination of urinary biomarker prolyl-4-hydroxyproline by liquid chromatography-tandem mass spectrometry. *J. Mass Spectrom.* **2012**, *47*, 294-302.

[26] Mazzi, G.; Fioravanzo, F.; Burti, E. New marker of bone resorption: hydroxyproline-containing peptide. High-performance liquid chromatographic assay without hydrolysis as an alternative to hydroxyproline determination: a preliminary report. *J. Chromatogr. B Biomed. Appl.* **1996**, *678*, 165-172.

[27] Hirayama, A.; Igarashi, K.; Tomita, M.; Soga, T. Development of quantitative method for determination of gamma-glutamyl peptides by capillary electrophoresis

tandem mass spectrometry: an efficient approach avoiding matrix effect. *J. Chromatogr. A* **2014**, *1369*, 161-169.

[28] Soga, T.; Sugimoto, M.; Honma, M.; Mori, M.; Igarashi, K.; Kashikura, K.; Ikeda, S.; Hirayama, A.; Yamamoto, T.; Yoshida, H., et al. Serum metabolomics reveals gamma-glutamyl dipeptides as biomarkers for discrimination among different forms of liver disease. *J. Hepatol.* **2011**, *55*, 896-905.

[29] Yamamoto, S.; Shiga, K.; Kodama, Y.; Imamura, M.; Uchida, R.; Obata, A.; Bamba, T.; Fukusaki, E. Analysis of the Correlation between Dipeptides and Taste Differences among Soy Sauces by Using Metabolomics-Based Component Profiling. *J. Biosci. Bioeng.* **2014**, *118*, 56-63.

[30] de Person, M.; Sevestre, A.; Chaimbault, P.; Perrot, L.; Duchiron, F.; Elfakir C. Characterization of Low-Molecular Weight Peptides in Champagne Wine by Liquid Chromatography/Tandem Mass Spectrometry. *Anal. Chim. Acta* **2004**, *520*, 149-158.

[31] Desportes, C.; Charpentier, M.; Duteurtre, B.; Maujean, A.; Duchiron, F. Liquid Chromatographic Fractionation of Small Peptides from Wine. *J. Chromatogr. A* **2000**, *893*, 281-291.

[32] Takahashi, K.; Tokuoka, M.; Kohno, H.; Sawamura, N.; Myoken, Y.; Mizuno, A. Comprehensive Analysis of Dipeptides in Alcoholic Beverages by Tag-Based

Separation and Determination Using Liquid Chromatography/Electrospray Ionization Tandem Mass Spectrometry and Quadrupole-Time-of-Flight Mass Spectrometry. *J. Chromatogr. A* **2012**, *1242*, 17-25.

[33] Nakashima, E. M. N.; Qing, H. -Q.; Tanaka, M.; Matsui, T. Improved Detection of Di-Peptides by Liquid Chromatography-Tandem Mass Spectrometry with 2,4,6-Trinitrobenzene Sulfonate Conversion. *Biosci. Biotechnol. Biochem.* **2013**, *77*, 2094-2099.

[34] Zhou, G.; Wang, M.; Li, Y.; Peng, Y.; Li, X. Rapid and Sensitive Analysis of 27 Underivatized Free Amino Acids, Dipeptides, and Tripeptides in Fruits of *Siraitia grosvenorii* Swingle Using HILIC-UHPLC-QTRAP((R))/MS (2) Combined with Chemometrics methods. *Amino Acids* **2015**, *47*, 1589-1603.

[35] Uwaje, N. C.; Mueller, N. S.; Maccarrone, G.; Turck, C. W. Interrogation of MS/MS search data with an pI Filter algorithm to increase protein identification success. *Electrophoresis* **2007**, *28*, 1867-1874.

[36] Kofuji, S.; Hirayama, A.; Eberhardt, A. O., et al. IMP Dehydrogenase-2 Drives Aberrant Nucleolar Activity and Promotes Tumorigenesis in Glioblastoma. *Nat. Cell. Biol.* **2019**, *21*, 1003-1014.

[37] Nagata, Y.; Hirayama, A.; Ikeda, S.; Shirahata, A.; Shoji, F.; Maruyama, M.;

Kayano, M.; Bundo, M.; Hattori, K.; Yoshida, S.; Goto, Y.; Urakami, K.; Soga, T.; Ozaki, K.; Niida, S. Comparative Analysis of Cerebrospinal Fluid Metabolites in Alzheimer's Disease and Idiopathic Normal Pressure Hydrocephalus in a Japanese Cohort. *Biomarker Res.* **2018**, *6*, 5.

[38] Saito, Y.; Morine, Y.; Iwahashi, S.; Ikemoto, T.; Imura, S.; Yamanaka-Okumura, H.; Hirayama, A.; Soga, T.; Tomita, M.; Shimada, M. Changes of Liver Metabolites Following Hepatectomy with Ischemia Reperfusion towards Liver Regeneration. *Ann. Gastroenterol. Surg.* **2018**, *2*, 204-211.

[39] Satoh, K.; Yachida, S.; Sugimoto, M., et al. Global Metabolic Reprogramming of Colorectal Cancer Occurs at Adenoma Stage and is Induced by MYC. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, E7697-E7706.

[40] Ozawa, H.; Hirayama, A.; Shoji, F.; Maruyama, M.; Suzuki, K.; Yamanaka-Okumura, H.; Tatano, H.; Morine, Y.; Soga, T.; Shimada, M.; Tomita, M. Comprehensive Dipeptide Analysis Revealed Cancer-Specific Profile in the Liver of Patients with Hepatocellular Carcinoma and Hepatitis. *Metabolites* **2020**, 10, 442.

[41] Yang, J. D.; Roberts, L. R. Hepatocellular carcinoma: A global view. *Nat. Rev. Gastroenterol. Hepatol.* **2010**, *7*, 448-458.

[42] Sakamoto, M. Early HCC: diagnosis and molecular markers. *J. Gastroenterol.*

**2009**, *44 Suppl. 19*, 108-111.

[43] Ryder, S. D. British Society of G. Guidelines for the diagnosis and treatment of hepatocellular carcinoma (HCC) in adults. *Gut* **2003**, *52 Suppl. 3*, iii1-8.

[44] Adams, E.; De Maesschalck, R.; De Spiegeller, B.; Vander Heyden, Y.; Smeyers-Verbeke, J.; Massart, D. L. Evaluation of dissolution profiles using principal component analysis. *Int. J. Pharm.* **2001**, *212*, 41-53.

[45] Greg Wells, Harry Prest, Charles William Russ IV. Signal, Noise, and Detection Limits in Mass Spectrometry. Agilent Technologies.

[46] Soga, T.; Heiger, D. N. Amino Acid Analysis by Capillary Electrophoresis Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2000**, *72*, 1236-1241.

[47] Handbook of Chemistry: Pure Chemistry, 5th ed.

[48] Gertsman, I.; Gangoiti, J. A.; Barshop, B. A. Validation of a Dual LC-HRMS Platform for Clinical Metabolic Diagnosis in Serum, Bridging Quantitative Analysis and Untargeted Metabolomics. *Metabolomics* **2014**, *10*, 312-323.

[49] Roux, A.; Xu, Y.; Heilier, J. -F.; Olivier, M. -F.; Ezan, E.; Tabet, J. -C.; Junot, C. Annotation of the Human Adult Urinary Metabolome and Metabolite Identification Using Ultra High Performance Liquid Chromatography Coupled to a Linear Quadrupole Ion Trap-Orbitrap Mass Spectrometer. *Anal. Chem.* **2012**, *84*, 6429-6437.

[50] Toyo'oka, T. Recent Advances in Separation and Detection Methods for Thiol Compounds in Biological Samples. *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.* **2009**, *877*, 3318-3330.

[51] Kumar, A.; Bachhawat, A. K. Throwing Light on a Lightly Studied Metabolite. *Curr. Sci.* **2012**, *102*, 288-297.

[52] Kim, H.-J.; Kim, J. H.; Noh, S.; Hur, H. J.; Sung, M. J.; Hwang, J. -T.; Park, J. H.; Yang, H. J.; Kim, M. -S.; Kwon, D. Y.; Yoon, S. H. Metabolomic Analysis of Livers and Serum from High-Fat Diet Induced Obese Mice. *J. Proteome Res.* **2011**, *10*, 722-731.

[53] Mori, A.; Hatate, H.; Tanaka, R. Ability of Three Kind of Imidazole Dipeptides, Carnosine, Anserine, and Balenine, to Interact with Unsaturated Fatty Acid-Derived Aldehydes and Carbohydrate Derived Aldehydes. *Int. J. Pept. Res. Ther.* **2020**, *26*, 1651-1660.

[54] Mong, M.; Chao, C.; Yin, M. Histidine and Carnosine Alleviated Hepatic Steatosis in Mice Consumed High Saturated Fat Diet. *Eur. J. Pharmacol.* **2011**, *653*, 82-88.

[55] Stegen, S.; Everaert, I.; Deldicque, L.; Vallova, S.; de Courten, B.; Ukropcova, B.; Ukropec, J.; Derave, W. Muscle Histidine-Containing Dipeptides are Elevated by Glucose Intolerance in Both Rodents and Men. *PLoS One* **2015**, *10*, No. e0121062.

[56] Zhou, L.; Ding, L.; Yin, P.; Lu, X.; Wang, X.; Niu, J.; Gao, P.; Xu, G. Serum

metabolic profiling study of hepatocellular carcinoma infected with hepatitis B or hepatitis C virus by using liquid chromatography-mass spectrometry. *J. Proteome Res.* **2012**, *11*, 5433-5442.

[57] Hirayama, A.; Kami, K.; Sugimoto, M.; Sugawara, M.; Toki, N.; Onozuka, H.; Kinoshita, T.; Saito, N.; Ochiai, A.; Tomita, M., et al. Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry. *Cancer Res.* **2009**, *69*, 4918-4925.

[58] Dieckmann, R.; Neuhof, T.; Pavela-Vrancic, M.; Döhren H. Dipeptide synthesis by an isolated adenylate-forming domain of non-ribosomal peptide synthetase (NRPS). *FEBS Lett.* **2001**, *498*, 42-45.

[59] Petruzziello, A. Epidemiology of Hepatitis B Virus (HBV) and Hepatitis C Virus (HCV) Related Hepatocellular Carcinoma. *Open Virol. J.* **2018**, *12*, 26-32.

[60] Lu, X.; Li, Y.; Xia, B.; Bai, Y.; Zhang, K.; Zhang, X.; Xie, H.; Sun, F.; Hou, Y.; Li, K. Selection of small plasma peptides for the auxiliary diagnosis and prognosis of epithelial overian cancer by using UPLC/MS-based nontargeted and targeted analyses. *Int. J. Cancer* **2019**, *144*, 2033-2042.

# Abbreviates

## Amino Acids

### Non polarity

| | | |
|---|---|---|
| G | Gly | Glycine |
| A | Ala | Alanine |
| V | Val | Valine |
| L | Leu | Leucine |
| I | Ile | Isoleucine |
| M | Met | Methionine |
| F | Phe | Phenylalanine |
| Y | Tyr | Tyrosine |
| W | Trp | Tryptophan |
| P | Pro | Proline |

### Non-electric charge polarity

| | | |
|---|---|---|
| S | Ser | Serine |
| T | Thr | Threonine |
| C | Cys | Cysteine |

| | | |
|---|---|---|
| N | Asn | Asparagine |
| Q | Gln | Glutamine |

### Acidity

| | | |
|---|---|---|
| D | Asp | Aspartic Acid |
| E | Glu | Glutamic Acid |

### Basicity

| | | |
|---|---|---|
| K | Lys | Lysine |
| R | Arg | Arginine |
| H | His | Histidine |

## Scientific terms

| | |
|---|---|
| HCC | Hepatocellular carcinoma |
| HBV | Hepatitis B virus |
| HCV | Hepatitis C virus |
| Non B/C | Non B and non C hepatitis |
| MLC | Metastatic liver cancer |
| BMI | Body mass index |
| HFD | High-fat diet |

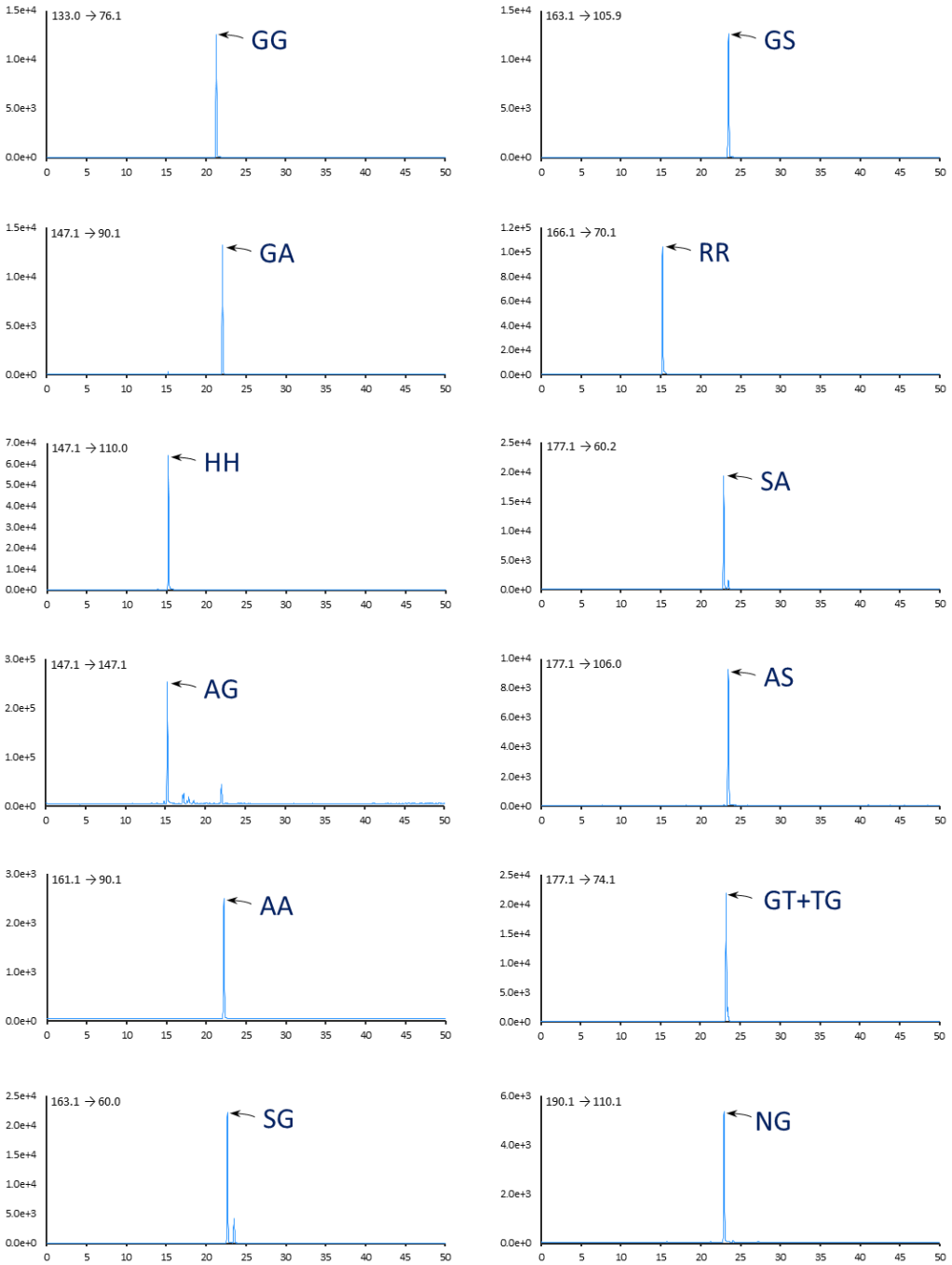
## *Abbreviates*

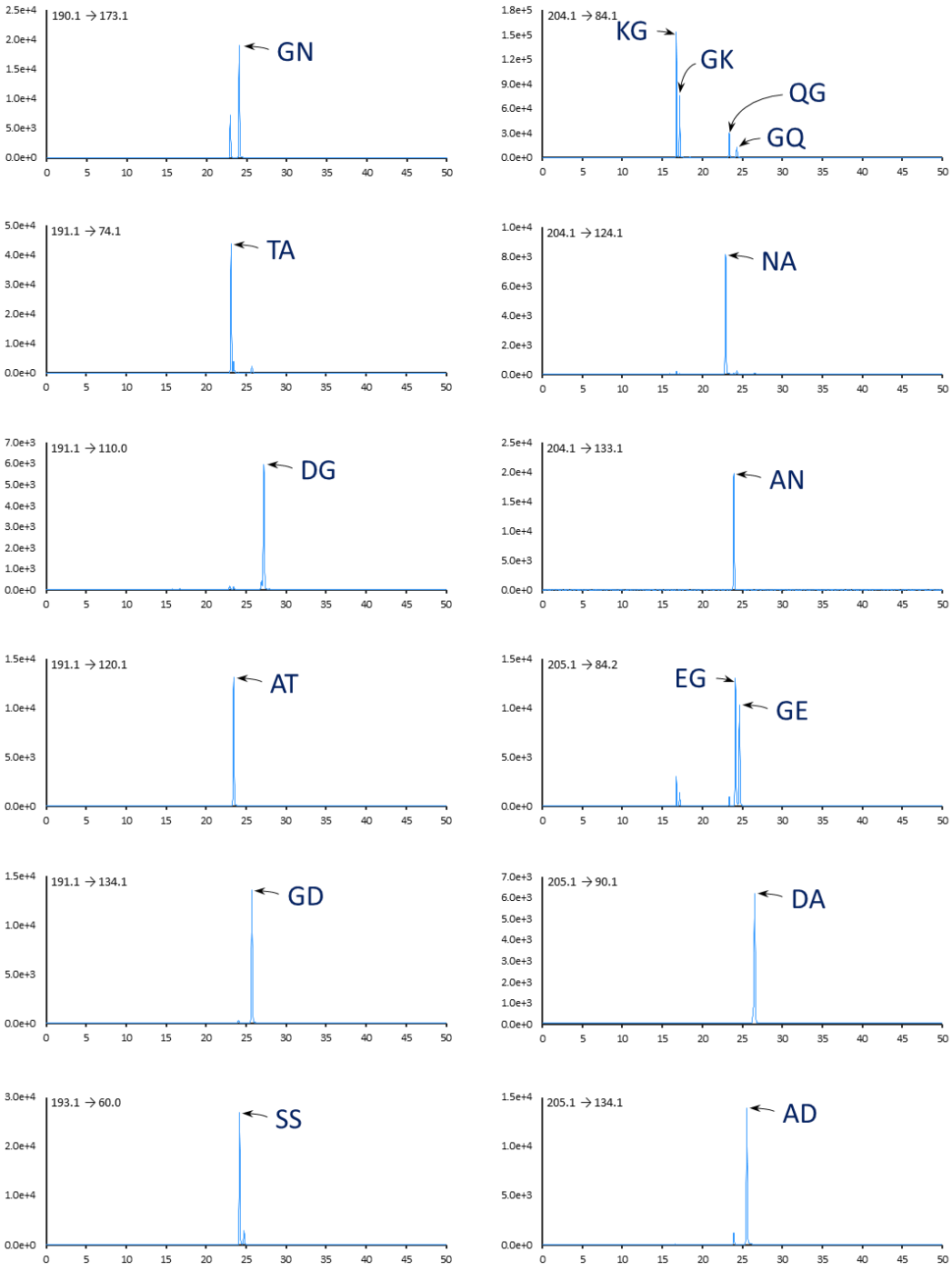
| | |
|---|---|
| GC-MS | Gas chromatography mass spectrometry |
| LC-MS | Liquid chromatography mass spectrometry |
| HPLC | High performance liquid chromatography |
| RPLC-MS | Reverse-phase liquid chromatography mass spectrometry |
| LC-MS/MS | Liquid chromatography tandem mass spectrometry |
| CE-MS | Capillary electrophoresis mass spectrometry |
| CE-MS/MS | Capillary electrophoresis tandem mass spectrometry |
| MRM | Multiple reaction monitoring |
| BGE | Background electrolyte |
| ESI | Electrospray ionization |
| AJS | Agilent jet stream |
| HSS | High strength silica |
| PFP | Penta fluoro phenyl |
| C18 | Octadecyl silyl |
| PCA | Principal component analysis |
| PC1 | First principal component |
| PC2 | Second principal component |
| CL | Control limit |

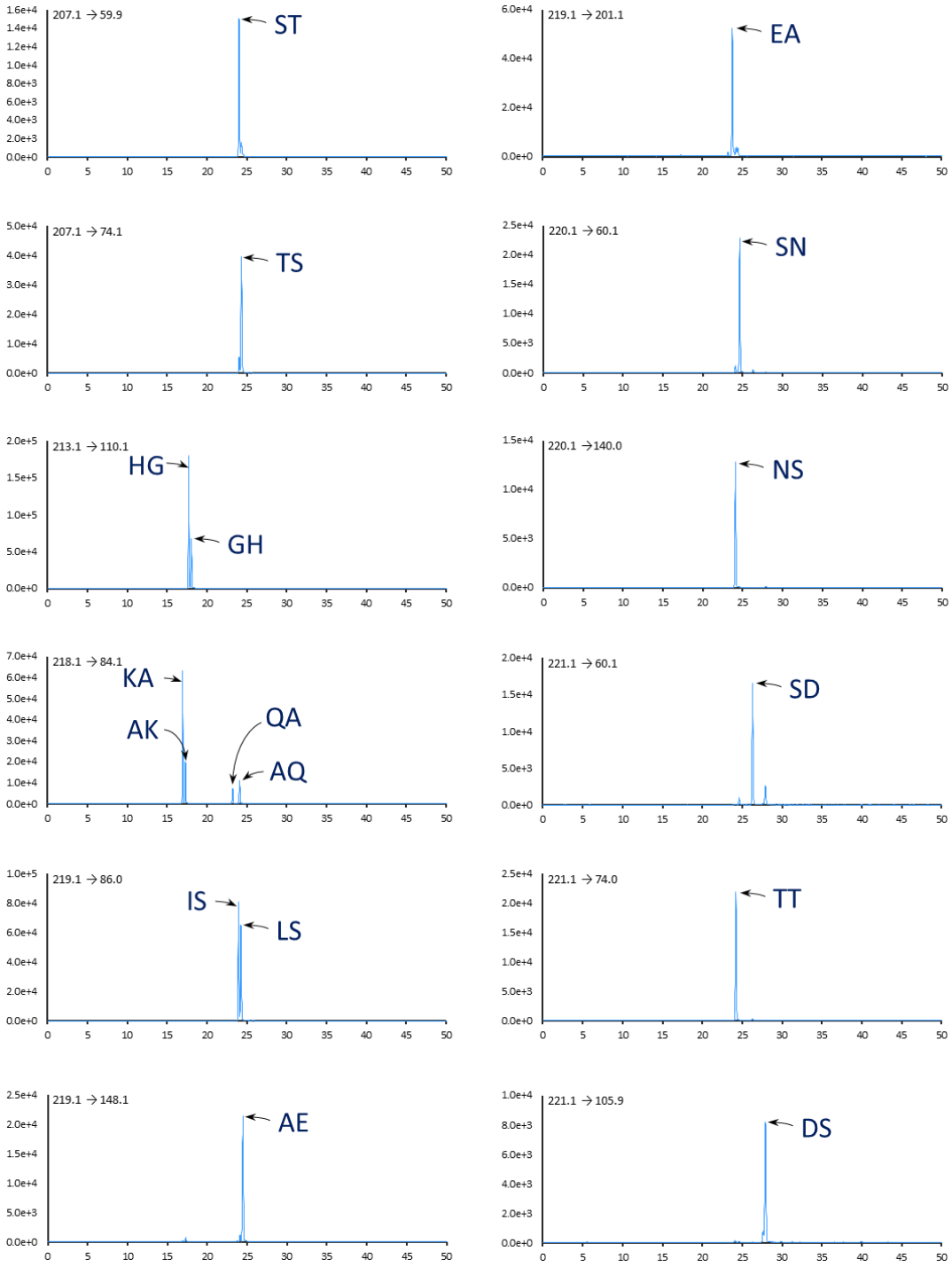
| | |
|---|---|
| UCL | Upper control limit |
| LOD | Limit of detection |
| IDL | Instrumental detection limit |
| STD | Standard |
| IS | Internal standard |

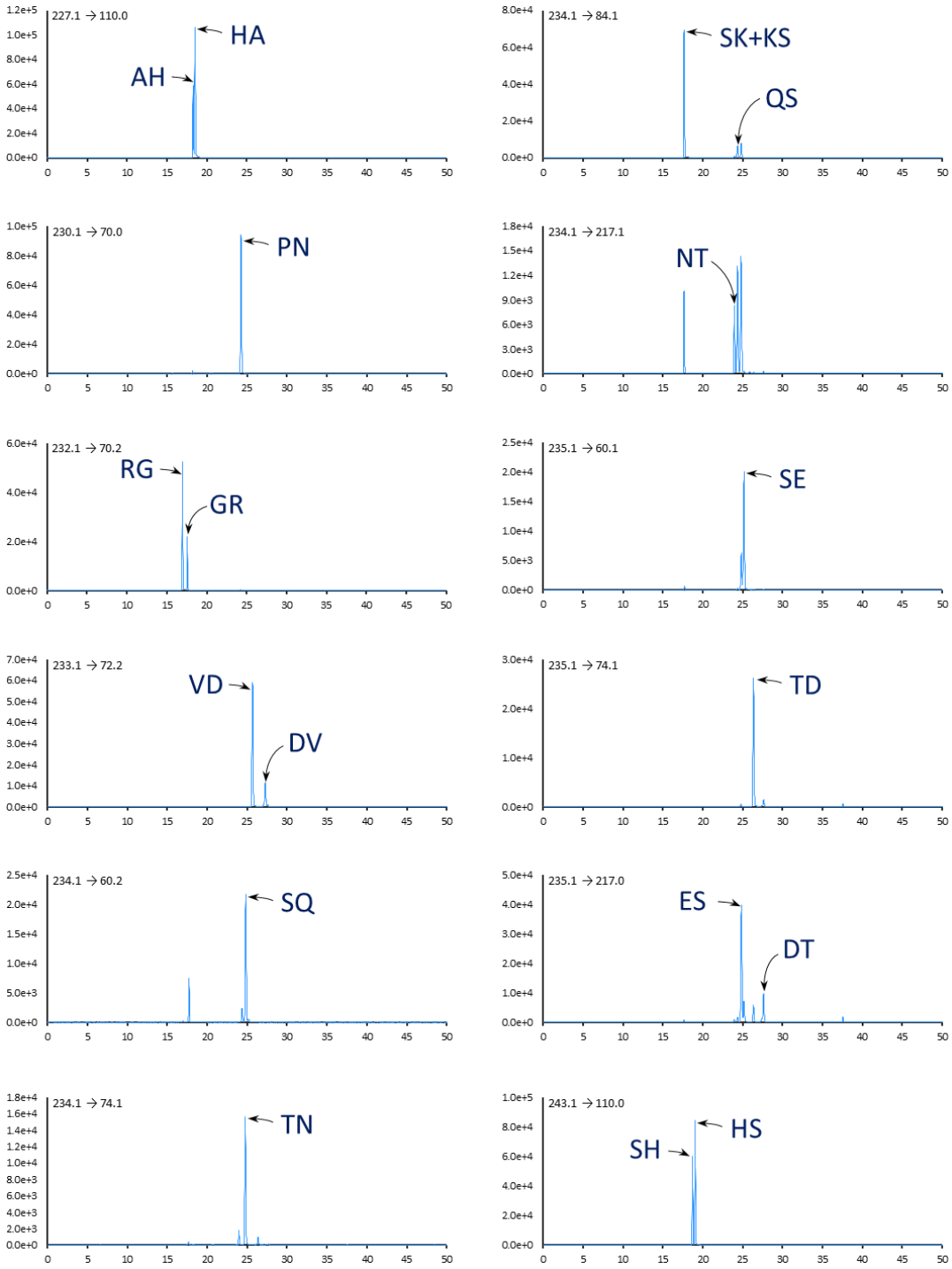
# Appendixes

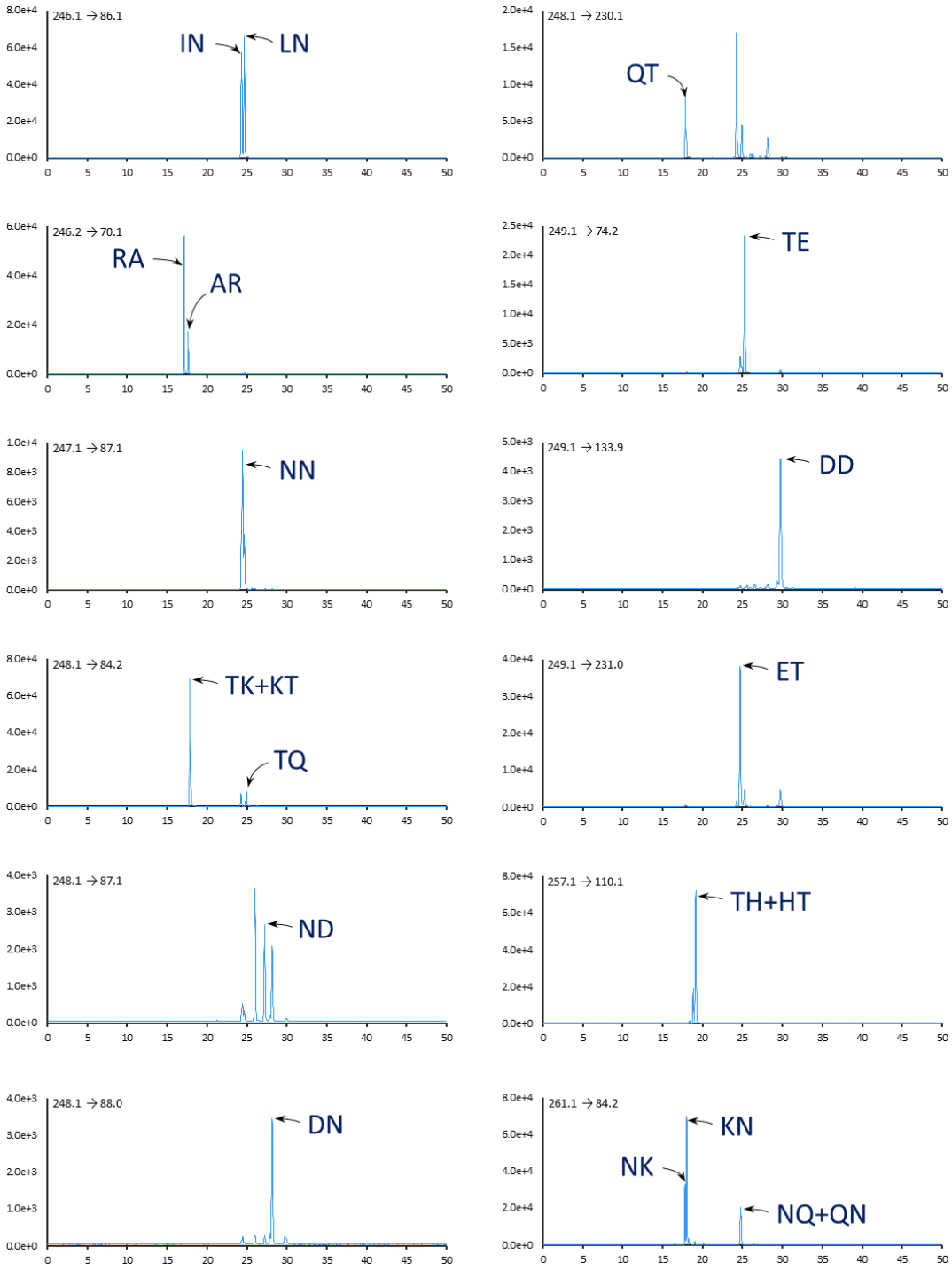
Figure A1. MRM electropherogram (blue) or chromatogram (orange) for a standard mixture of 361 dipeptides.

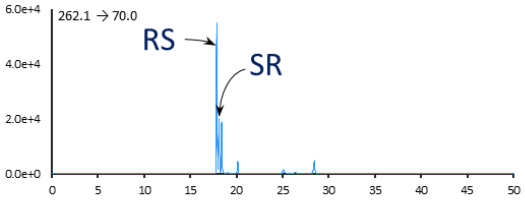


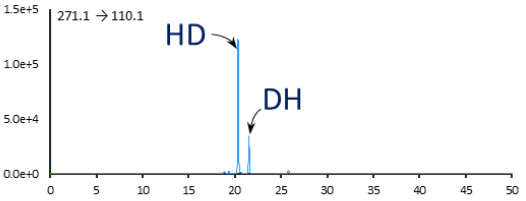

190.1 → 173.1
GN
204.1 → 84.1
KG
GK
QG
GQ
191.1 → 74.1
TA
204.1 → 124.1
NA
191.1 → 110.0
DG
204.1 → 133.1
AN
191.1 → 120.1
AT
205.1 → 84.2
EG
GE
191.1 → 134.1
GD
205.1 → 90.1
DA
193.1 → 60.0
SS
205.1 → 134.1
AD

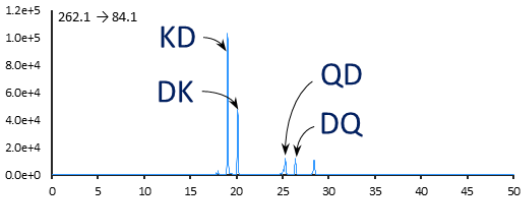

207.1 → 59.9
ST
219.1 → 201.1
EA
207.1 → 74.1
TS
220.1 → 60.1
SN
213.1 → 110.1
HG
GH
220.1 →140.0
NS
218.1 → 84.1
KA
AK
QA
AQ
221.1 → 60.1
SD
219.1 → 86.0
IS
LS
221.1 → 74.0
TT
219.1 → 148.1
AE
221.1 → 105.9
DS

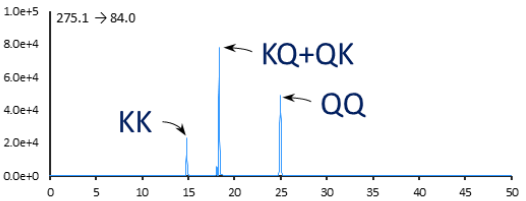

227.1 → 110.0
HA
AH
234.1 → 84.1
SK+KS
QS
230.1 → 70.0
PN
234.1 → 217.1
NT
232.1 → 70.2
RG
GR
235.1 → 60.1
SE
233.1 → 72.2
VD
DV
235.1 → 74.1
TD
234.1 → 60.2
SQ
235.1 → 217.0
ES
DT
234.1 → 74.1
TN
243.1 → 110.0
SH
HS

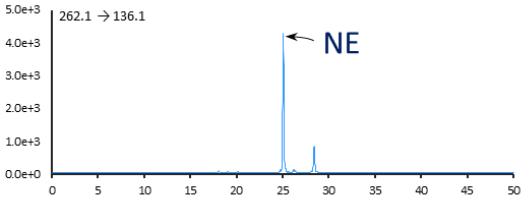

246.1 → 86.1
IN
LN
248.1 → 230.1
QT
246.2 → 70.1
RA
AR
249.1 → 74.2
TE
247.1 → 87.1
NN
249.1 → 133.9
DD
248.1 → 84.2
TK+KT
TQ
249.1 → 231.0
ET
248.1 → 87.1
ND
257.1 → 110.1
TH+HT
248.1 → 88.0
DN
261.1 → 84.2
KN
NK
NQ+QN

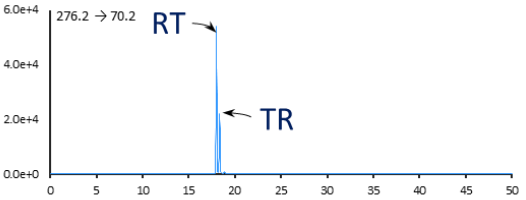

262.1 → 70.0
RS
SR

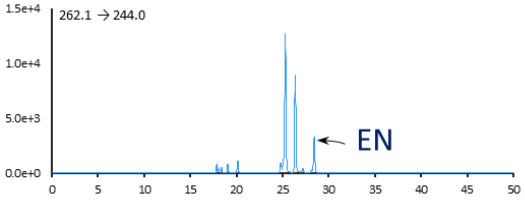

271.1 → 110.1
HD
DH

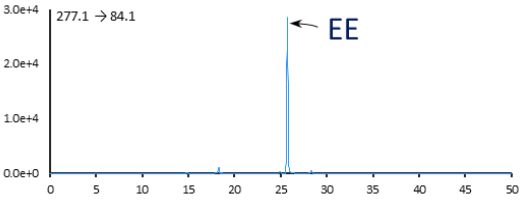

262.1 → 84.1
KD
DK
QD
DQ

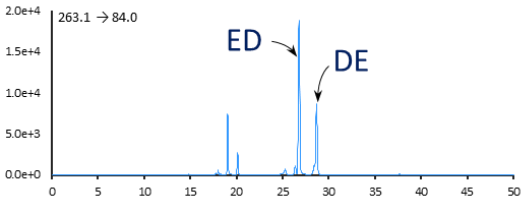

275.1 → 84.0
KQ+QK
QQ
KK

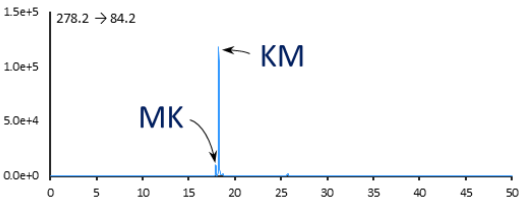

262.1 → 136.1
NE

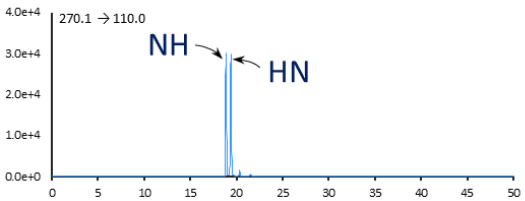

276.2 → 70.2
RT
TR

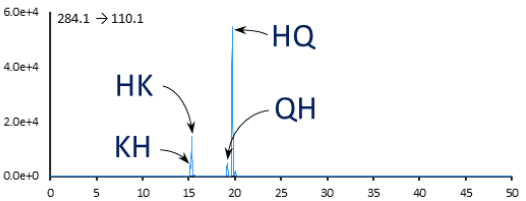

262.1 → 244.0
EN

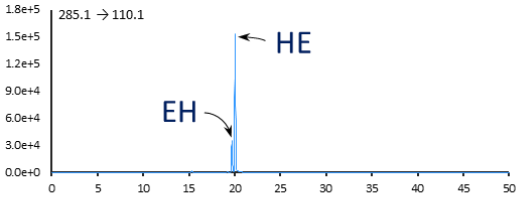

277.1 → 84.1
EE

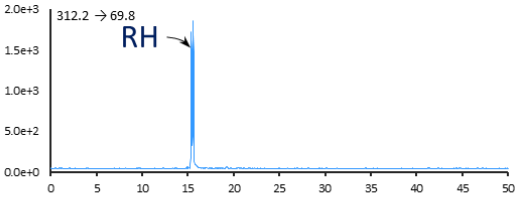

263.1 → 84.0
ED
DE

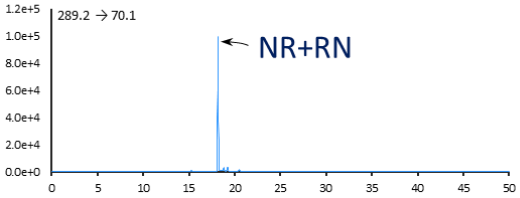

278.2 → 84.2
KM
MK

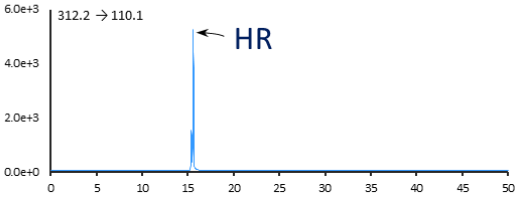

270.1 → 110.0
NH
HN

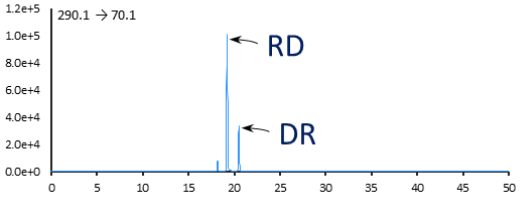

284.1 → 110.1
HQ
HK
QH
KH

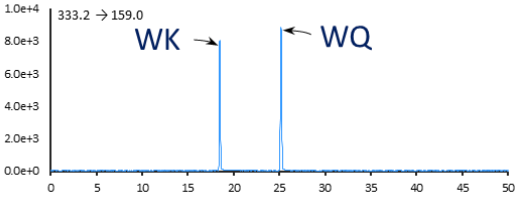

285.1 → 110.1
HE
EH

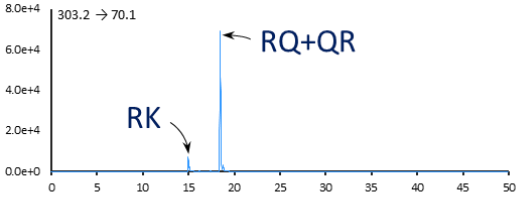

312.2 → 69.8
RH

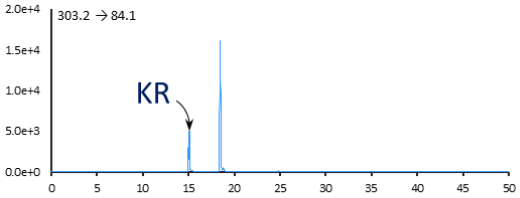

289.2 → 70.1
NR+RN

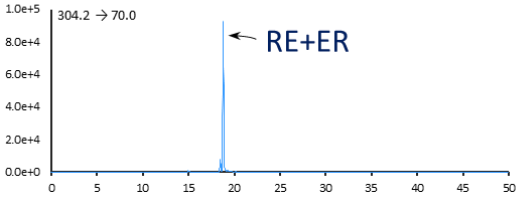

312.2 → 110.1
HR

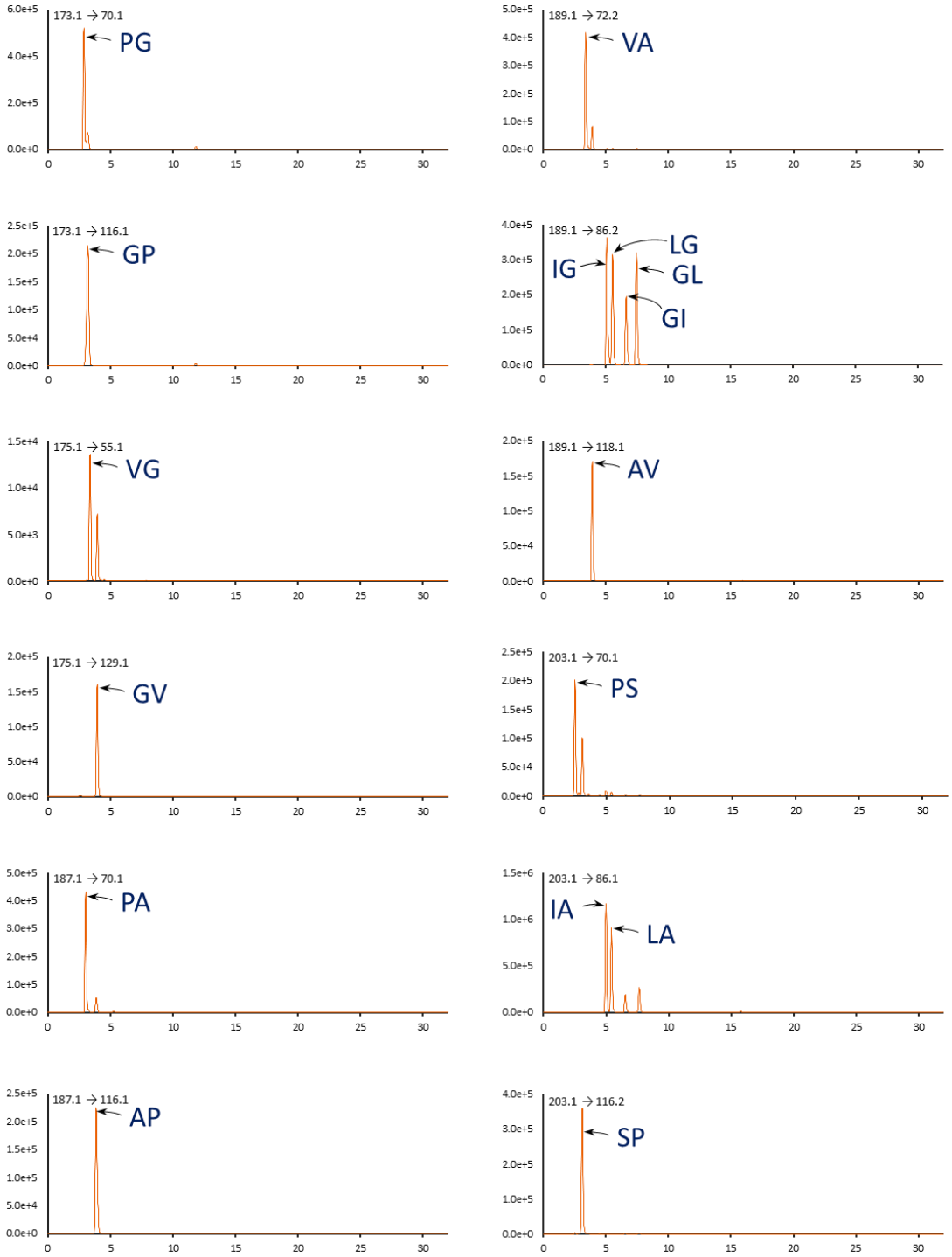

290.1 → 70.1
RD
DR

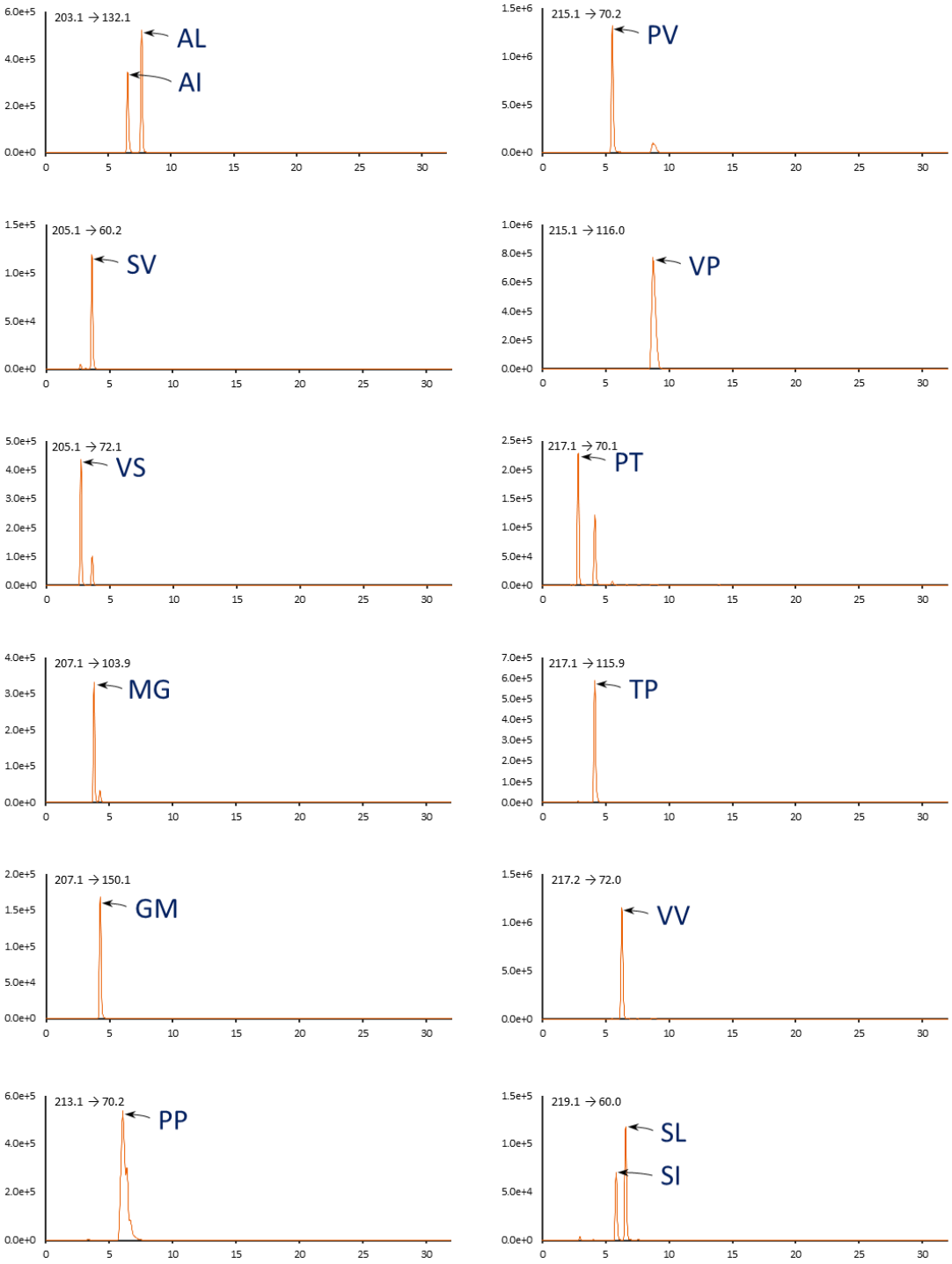

333.2 → 159.0
WK
WQ

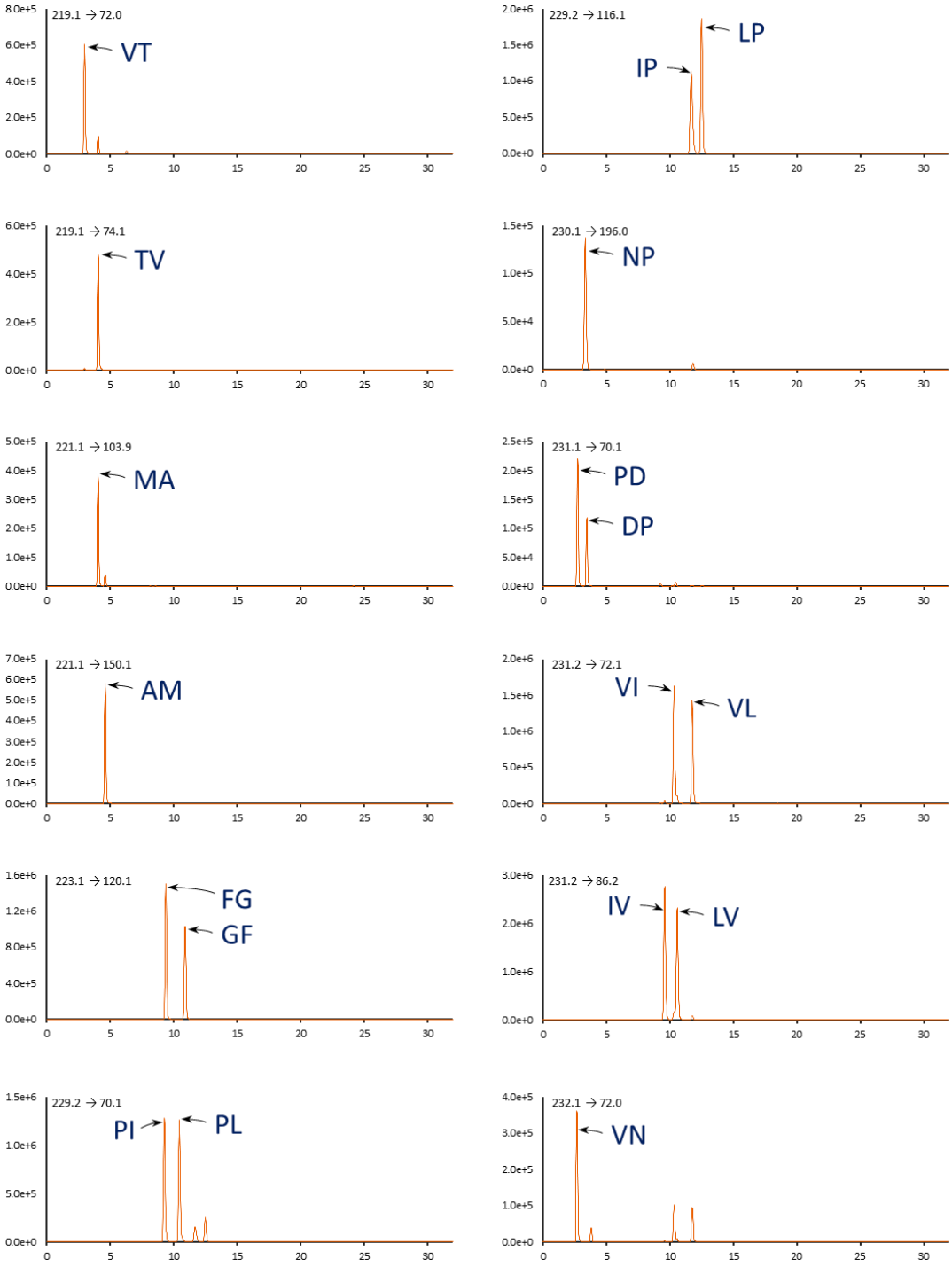

303.2 → 70.1
RQ+QR
RK

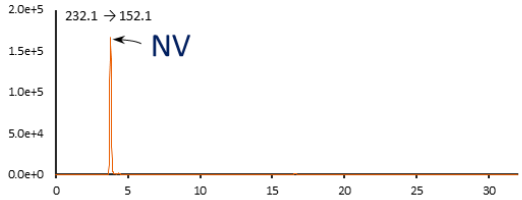

303.2 → 84.1
KR

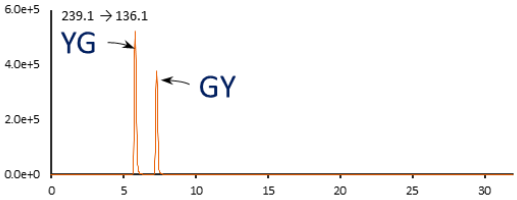

304.2 → 70.0
RE+ER

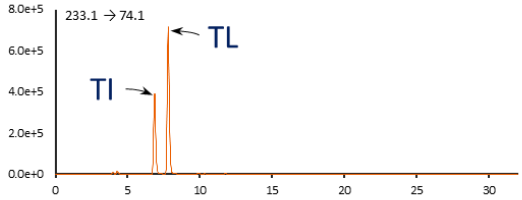

173.1 → 70.1
PG
189.1 → 72.2
VA
173.1 → 116.1
GP
189.1 → 86.2
LG
IG
GL
GI
175.1 → 55.1
VG
189.1 → 118.1
AV
175.1 → 129.1
GV
203.1 → 70.1
PS
187.1 → 70.1
PA
203.1 → 86.1
IA
LA
187.1 → 116.1
AP
203.1 → 116.2
SP

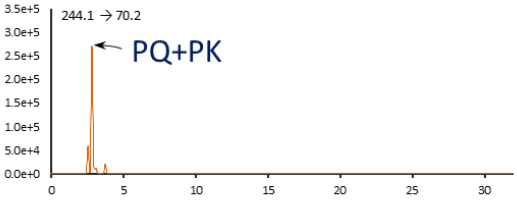

203.1 → 132.1
AL
AI
215.1 → 70.2
PV
205.1 → 60.2
SV
215.1 → 116.0
VP
205.1 → 72.1
VS
217.1 → 70.1
PT
207.1 → 103.9
MG
217.1 → 115.9
TP
207.1 → 150.1
GM
217.2 → 72.0
VV
213.1 → 70.2
PP
219.1 → 60.0
SL
SI

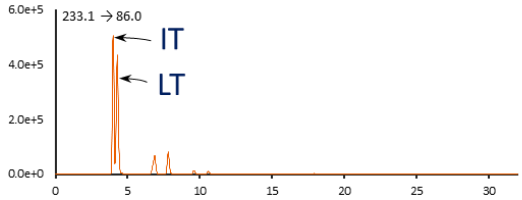

219.1 → 72.0
VT
229.2 → 116.1
LP
IP
219.1 → 74.1
TV
230.1 → 196.0
NP
221.1 → 103.9
MA
231.1 → 70.1
PD
DP
221.1 → 150.1
AM
231.2 → 72.1
VI
VL
223.1 → 120.1
FG
GF
231.2 → 86.2
IV
LV
229.2 → 70.1
PI
PL
232.1 → 72.0
VN

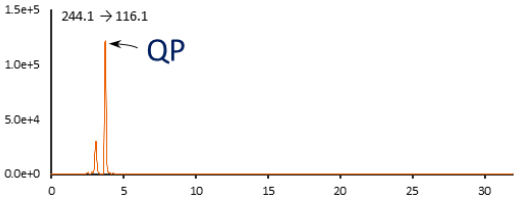

232.1 → 152.1
NV

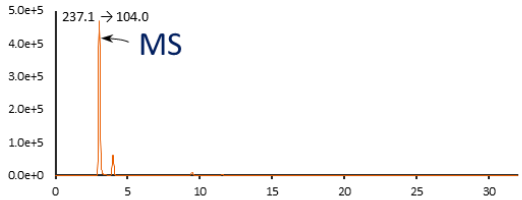

239.1 → 136.1
YG
GY

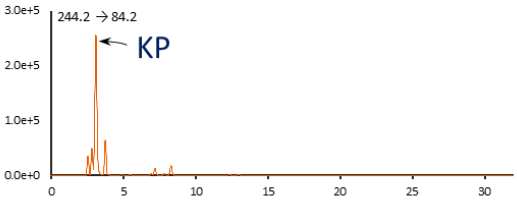

233.1 → 74.1
TL
TI

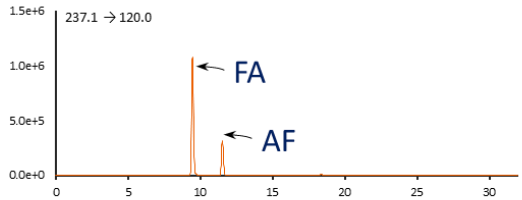

244.1 → 70.2
PQ+PK

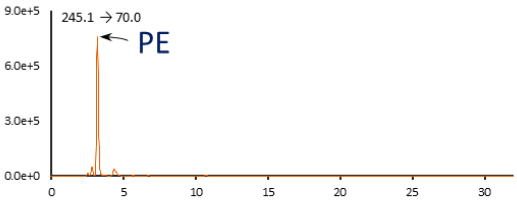

233.1 → 86.0
IT
LT

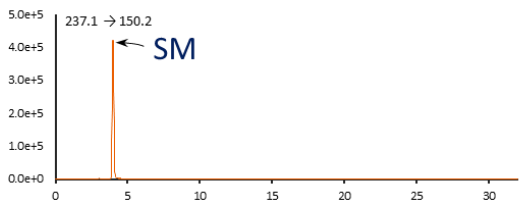

244.1 → 116.1
QP

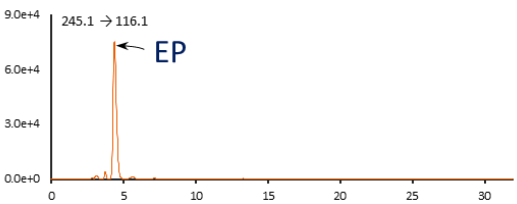

237.1 → 104.0
MS

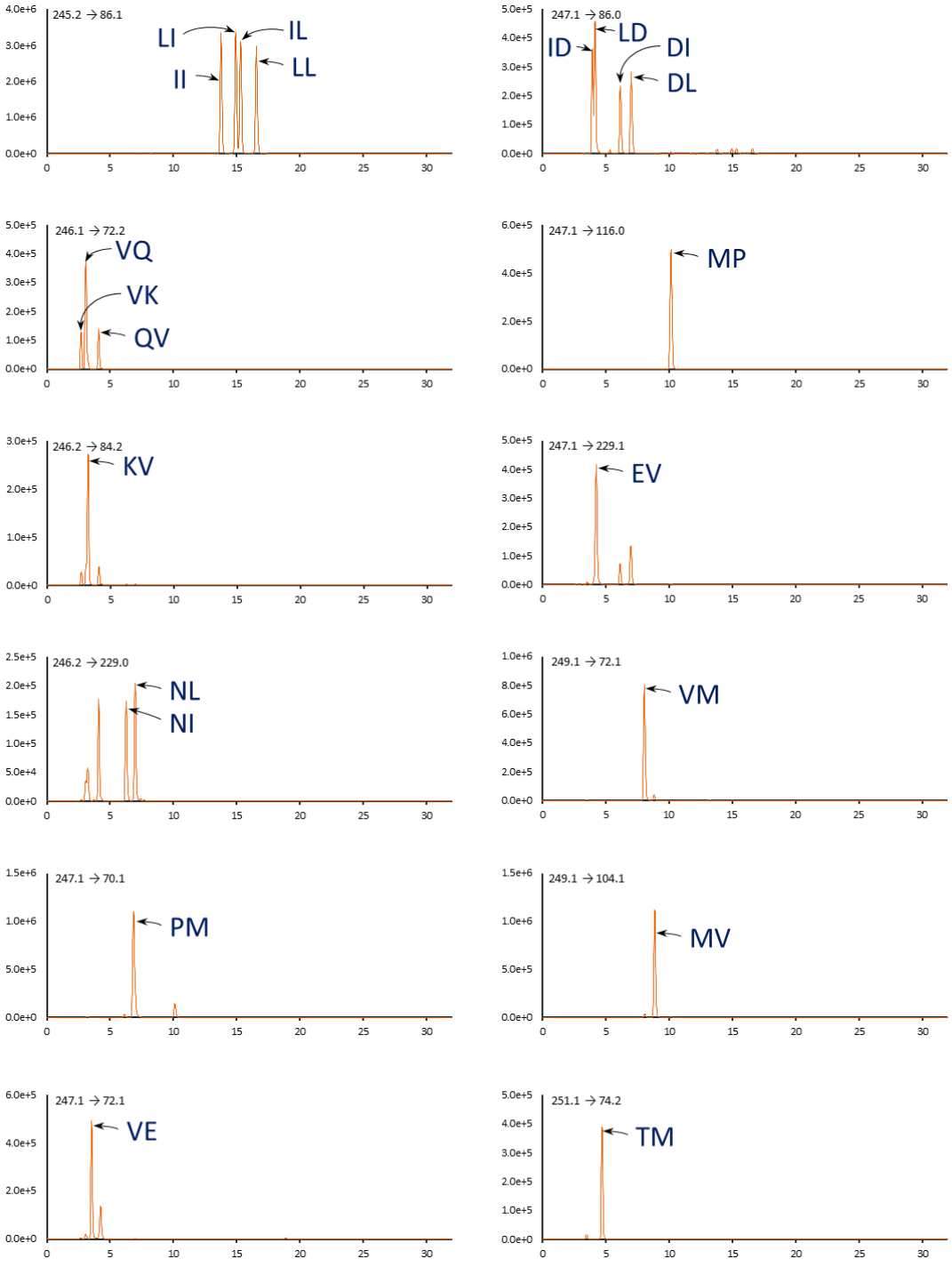

244.2 → 84.2
KP

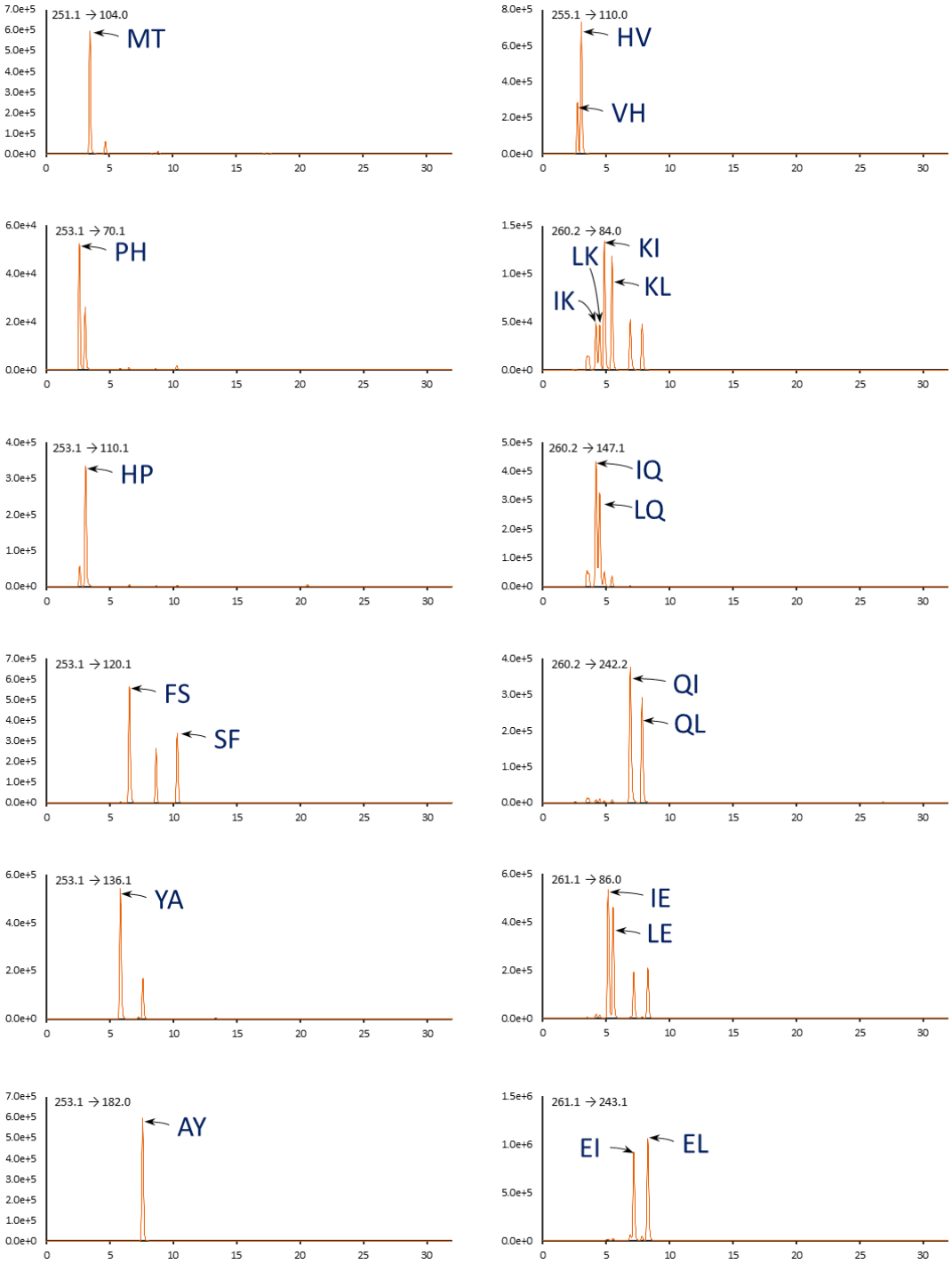

237.1 → 120.0
FA
AF

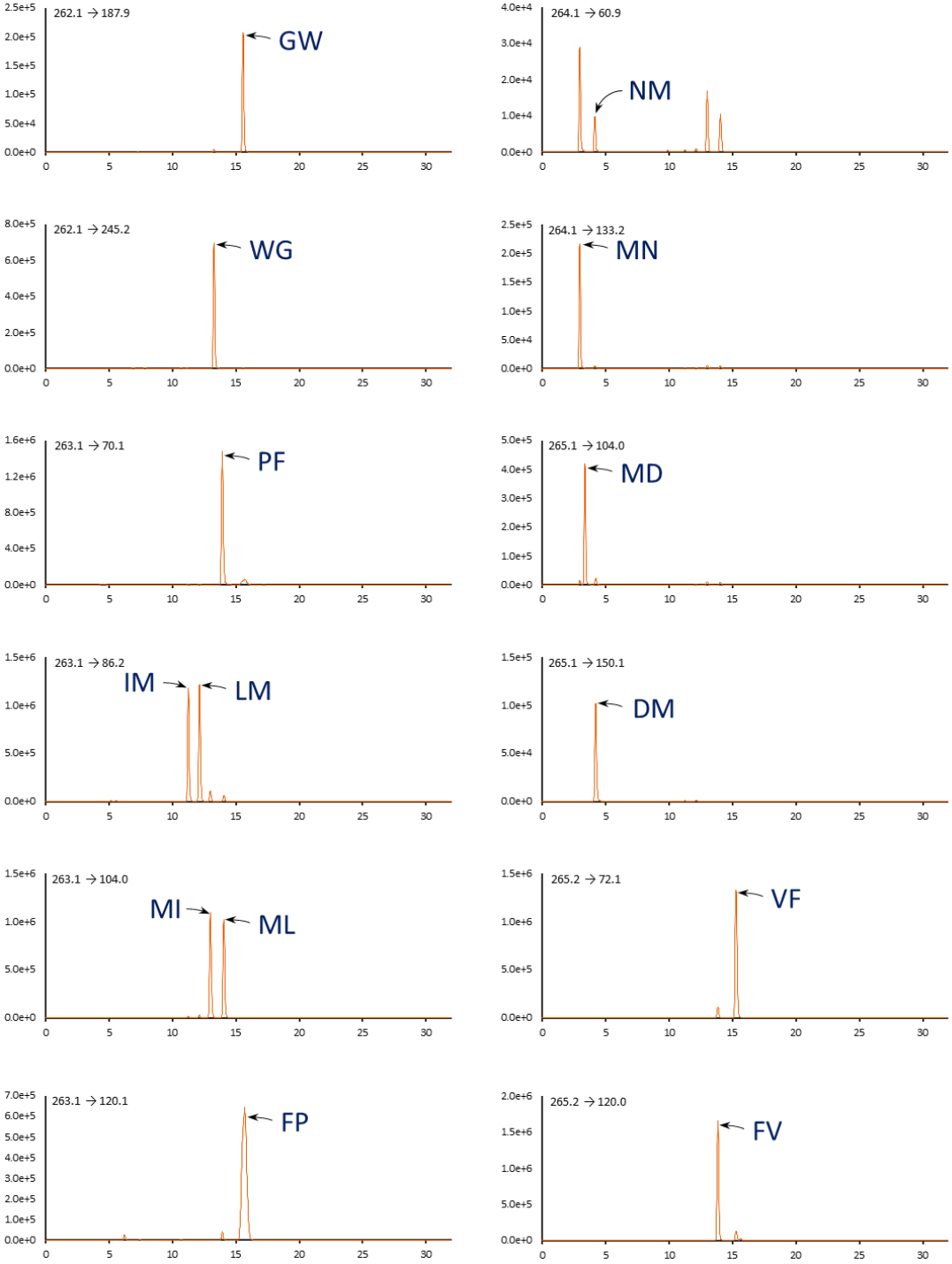

245.1 → 70.0
PE

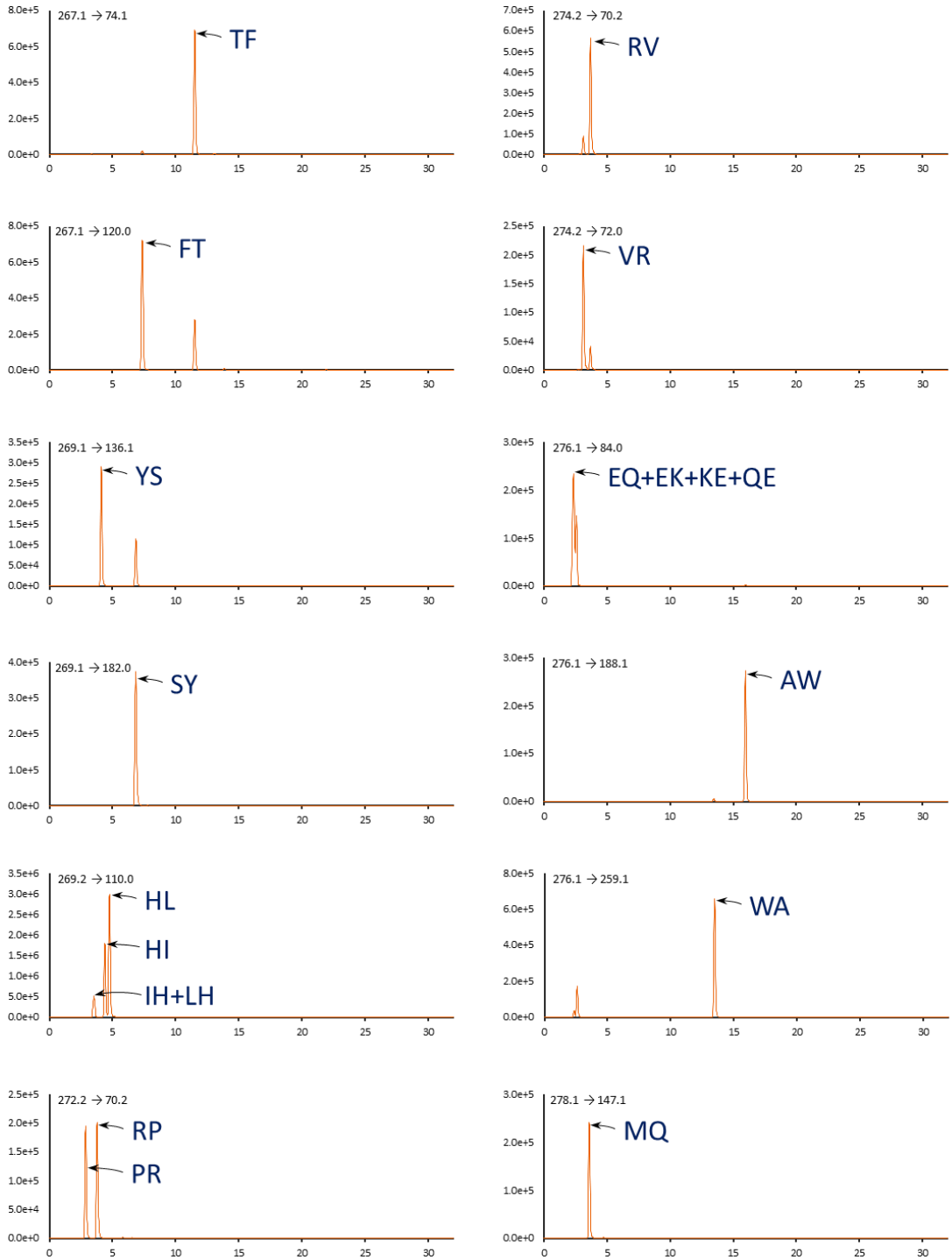

237.1 → 150.2
SM

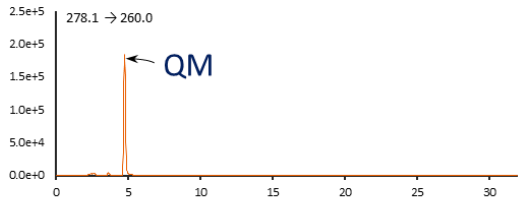

245.1 → 116.1
EP

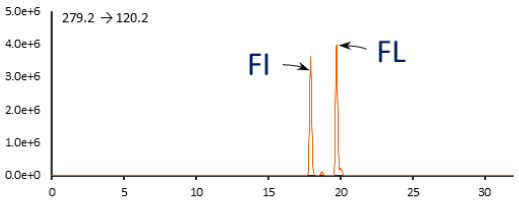

245.2 → 86.1
LI
IL
II
LL
247.1 → 86.0
ID
LD
DI
DL
246.1 → 72.2
VQ
VK
QV
247.1 → 116.0
MP
246.2 → 84.2
KV
247.1 → 229.1
EV
246.2 → 229.0
NL
NI
249.1 → 72.1
VM
247.1 → 70.1
PM
249.1 → 104.1
MV
247.1 → 72.1
VE
251.1 → 74.2
TM

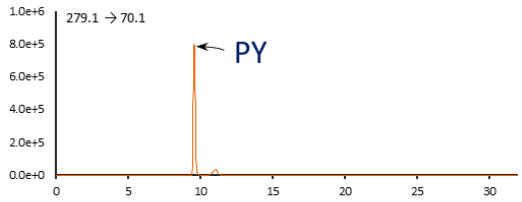

251.1 → 104.0
MT
255.1 → 110.0
HV
VH
253.1 → 70.1
PH
260.2 → 84.0
LK
KI
KL
IK
253.1 → 110.1
HP
260.2 → 147.1
IQ
LQ
253.1 → 120.1
FS
SF
260.2 → 242.2
QI
QL
253.1 → 136.1
YA
261.1 → 86.0
IE
LE
253.1 → 182.0
AY
261.1 → 243.1
EI
EL

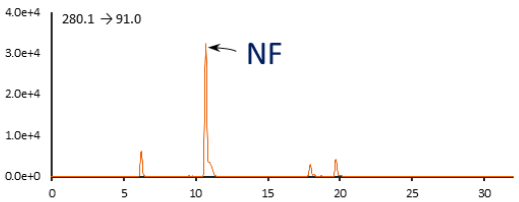

262.1 → 187.9
GW
264.1 → 60.9
NM
262.1 → 245.2
WG
264.1 → 133.2
MN
263.1 → 70.1
PF
265.1 → 104.0
MD
263.1 → 86.2
IM
LM
265.1 → 150.1
DM
263.1 → 104.0
MI
ML
265.2 → 72.1
VF
263.1 → 120.1
FP
265.2 → 120.0
FV

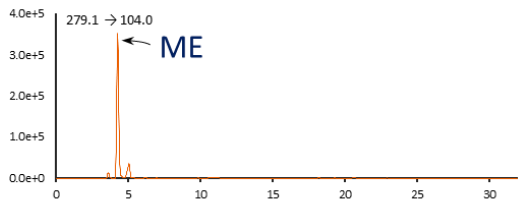

267.1 → 74.1
TF
274.2 → 70.2
RV
267.1 → 120.0
FT
274.2 → 72.0
VR
269.1 → 136.1
YS
276.1 → 84.0
EQ+EK+KE+QE
269.1 → 182.0
SY
276.1 → 188.1
AW
269.2 → 110.0
HL
HI
IH+LH
276.1 → 259.1
WA
272.2 → 70.2
RP
PR
278.1 → 147.1
MQ

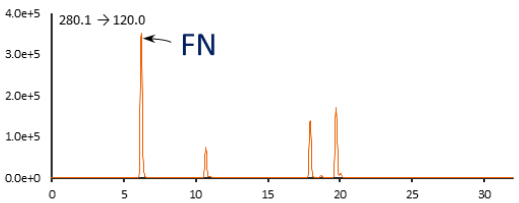

278.1 → 260.0
QM

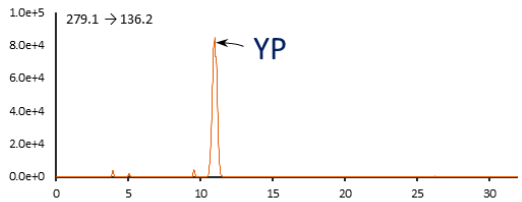

279.2 → 120.2
FI
FL

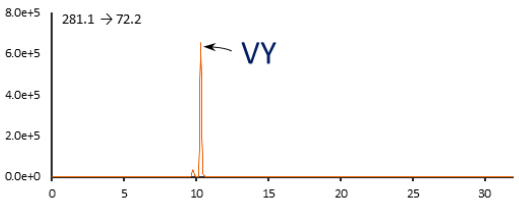

279.1 → 70.1
PY

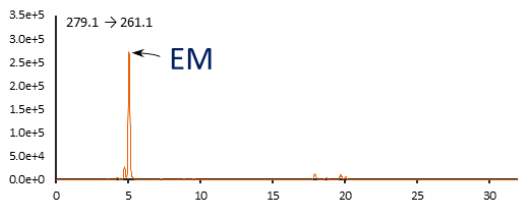

280.1 → 91.0
NF

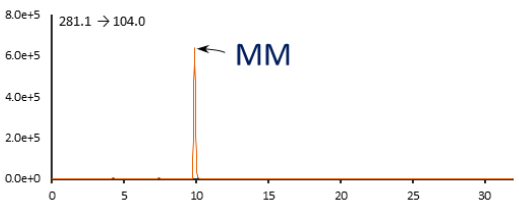

279.1 → 104.0
ME

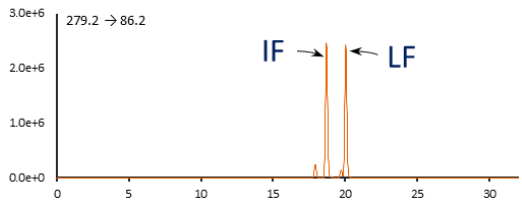

280.1 → 120.0
FN

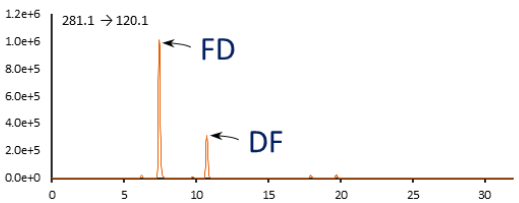

279.1 → 136.2
YP

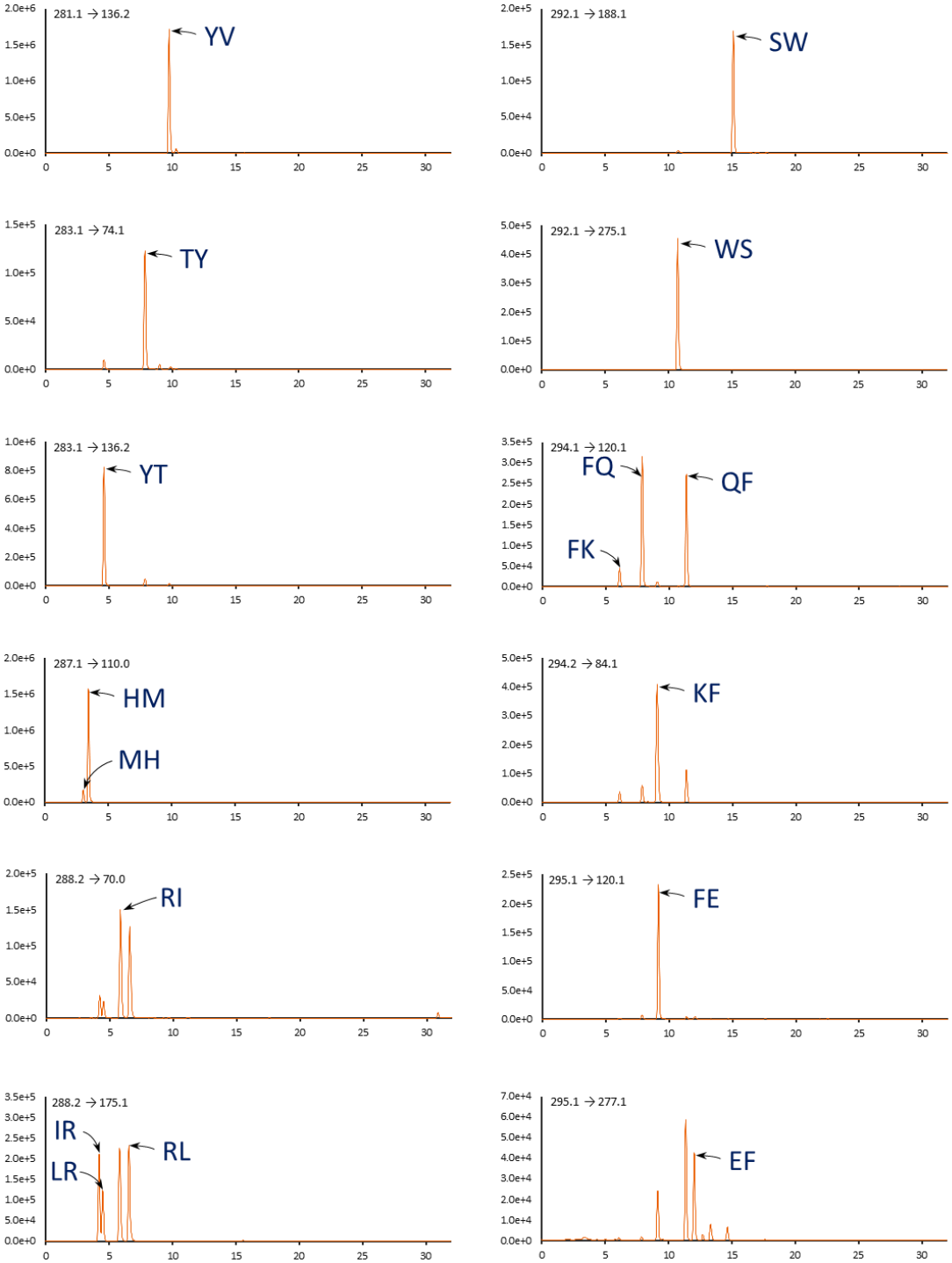

281.1 → 72.2
VY

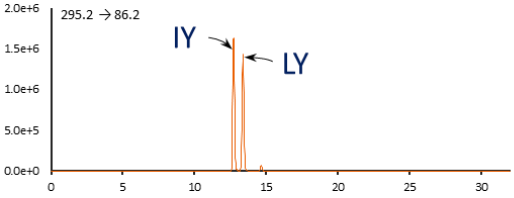

279.1 → 261.1
EM

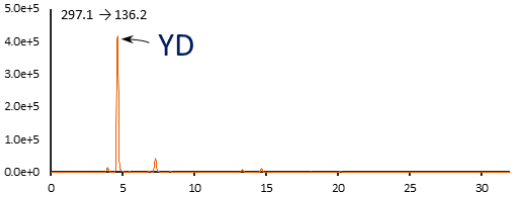

281.1 → 104.0
MM

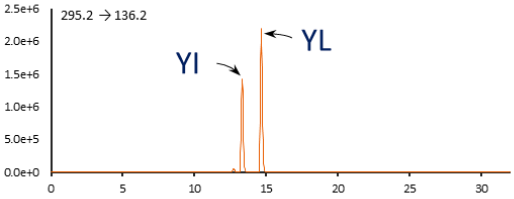

279.2 → 86.2
IF
LF

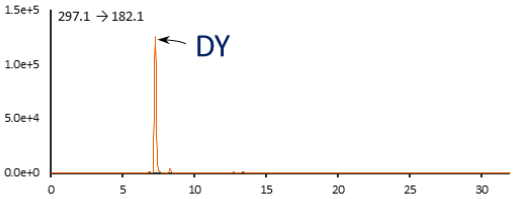

281.1 → 120.1
FD
DF

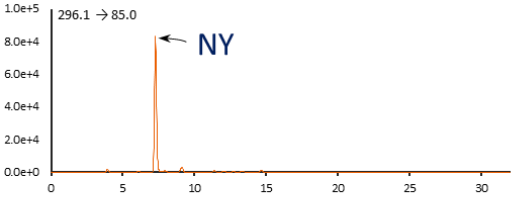

281.1 → 136.2
YV
292.1 → 188.1
SW
283.1 → 74.1
TY
292.1 → 275.1
WS
283.1 → 136.2
YT
294.1 → 120.1
FQ
QF
FK
287.1 → 110.0
HM
MH
294.2 → 84.1
KF
288.2 → 70.0
RI
295.1 → 120.1
FE
288.2 → 175.1
IR
RL
LR
295.1 → 277.1
EF

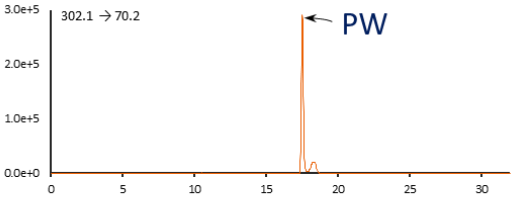

295.2 → 86.2
IY
LY

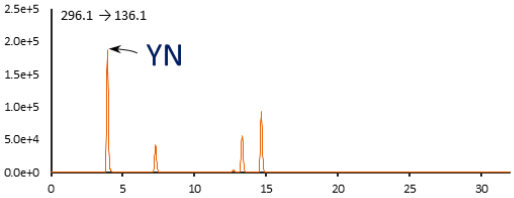

297.1 → 136.2
YD

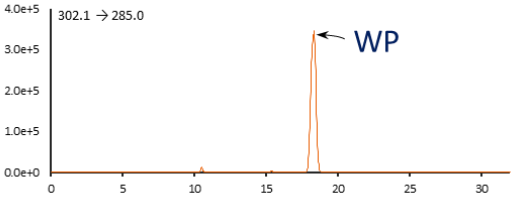

295.2 → 136.2
YI
YL

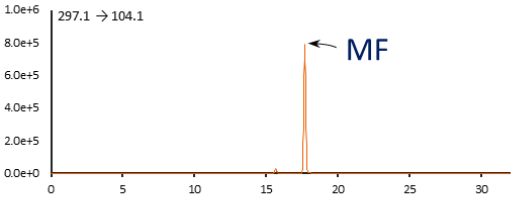

297.1 → 182.1
DY

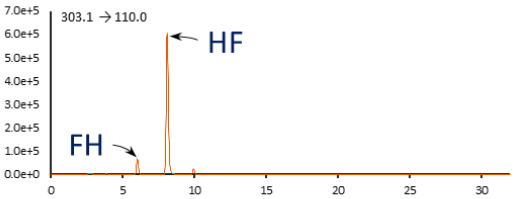

296.1 → 85.0
NY

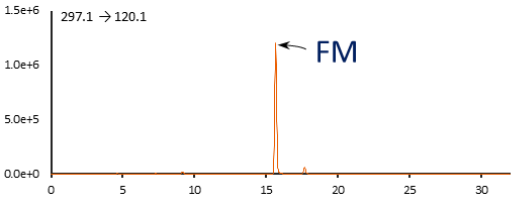

302.1 → 70.2
PW

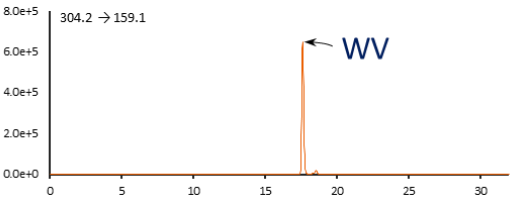

296.1 → 136.1
YN

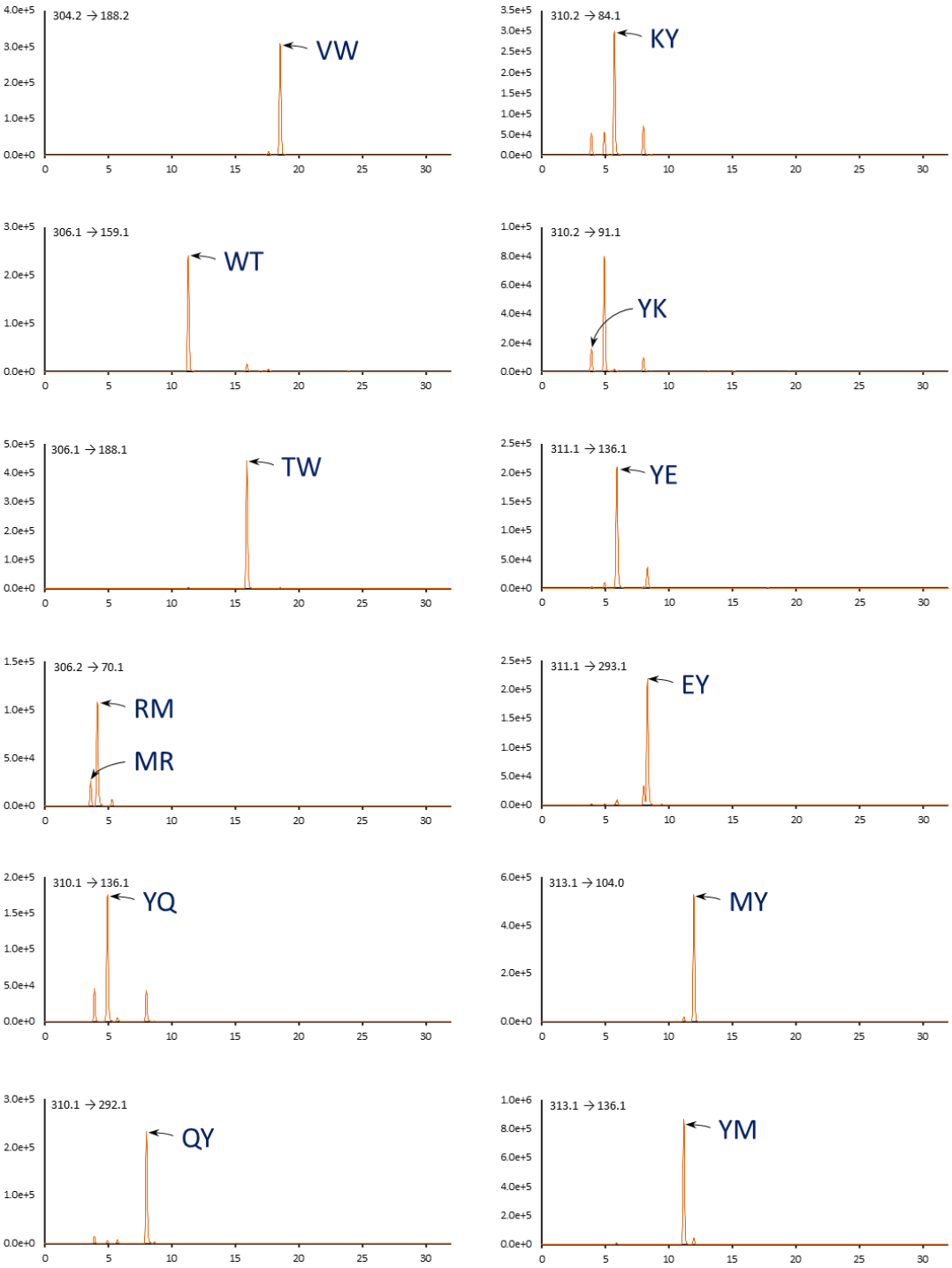

302.1 → 285.0
WP

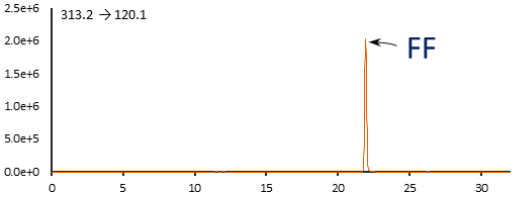

297.1 → 104.1
MF

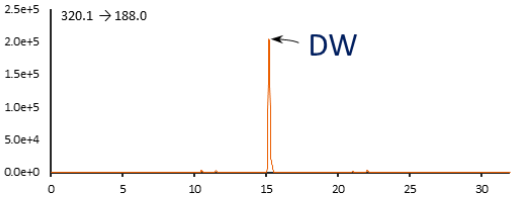

303.1 → 110.0
HF
FH

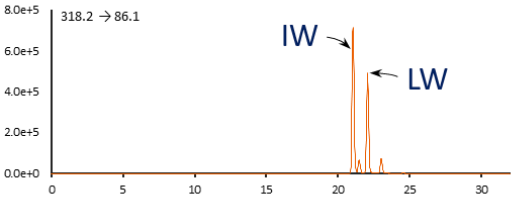

297.1 → 120.1
FM

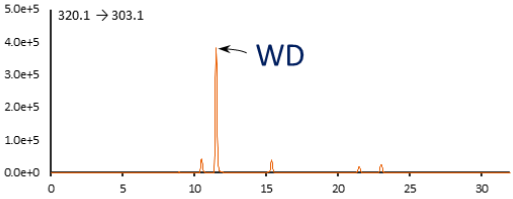

304.2 → 159.1
WV

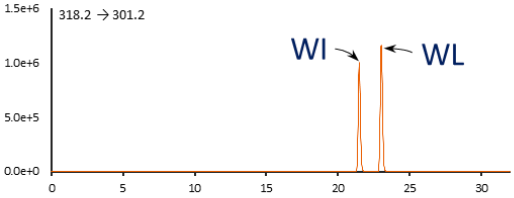

304.2 → 188.2
VW
310.2 → 84.1
KY
306.1 → 159.1
WT
310.2 → 91.1
YK
306.1 → 188.1
TW
311.1 → 136.1
YE
306.2 → 70.1
RM
MR
311.1 → 293.1
EY
310.1 → 136.1
YQ
313.1 → 104.0
MY
310.1 → 292.1
QY
313.1 → 136.1
YM

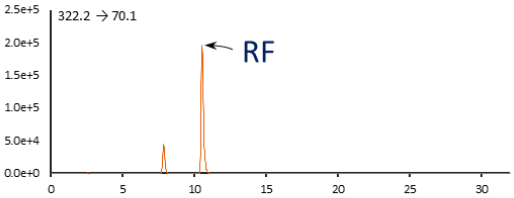

313.2 → 120.1
FF

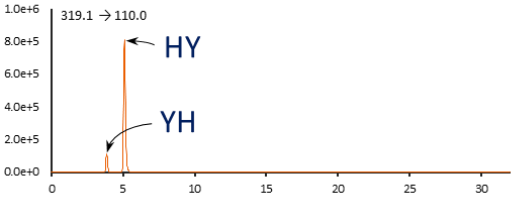

320.1 → 188.0
DW

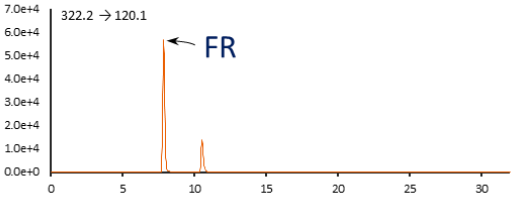

318.2 → 86.1
IW
LW

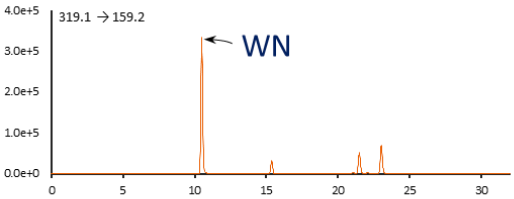

320.1 → 303.1
WD

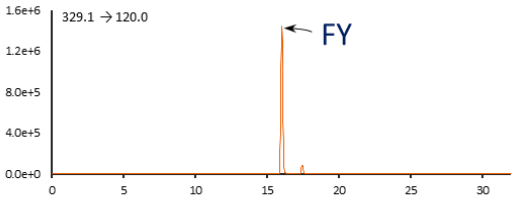

318.2 → 301.2
WI
WL

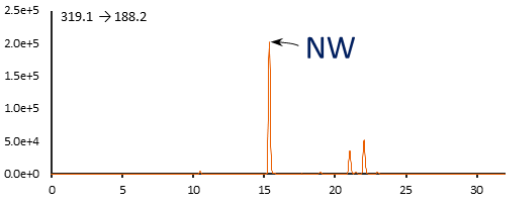

322.2 → 70.1
RF

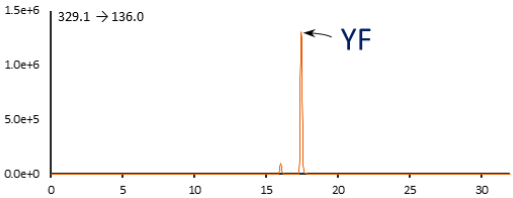

319.1 → 110.0
HY
YH

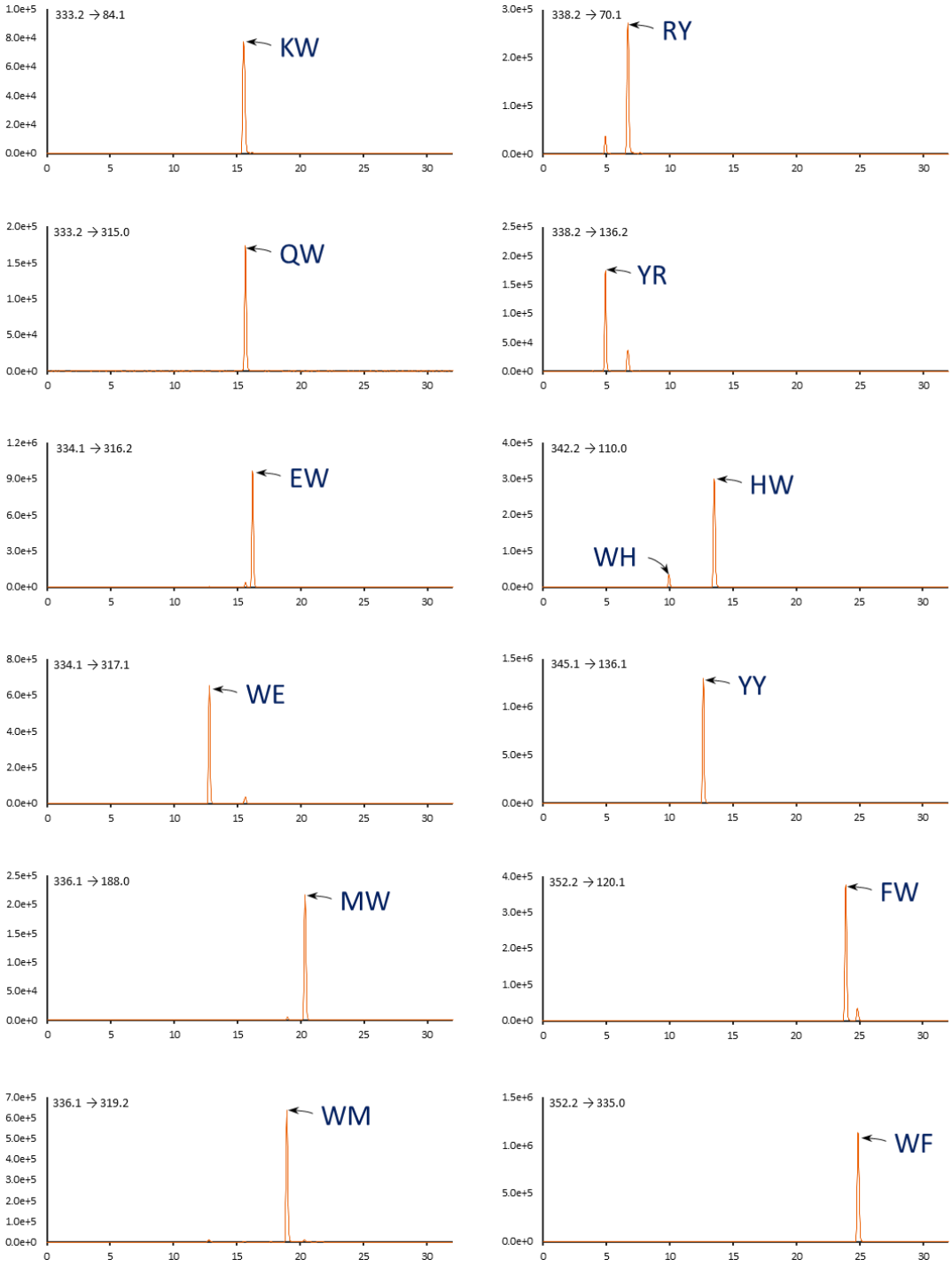

322.2 → 120.1
FR


319.1 → 159.2
WN


329.1 → 120.0
FY


319.1 → 188.2
NW


329.1 → 136.0
YF


333.2 → 84.1
KW
338.2 → 70.1
RY
333.2 → 315.0
QW
338.2 → 136.2
YR
334.1 → 316.2
EW
342.2 → 110.0
HW
WH
334.1 → 317.1
WE
345.1 → 136.1
YY
336.1 → 188.0
MW
352.2 → 120.1
FW
336.1 → 319.2
WM
352.2 → 335.0
WF

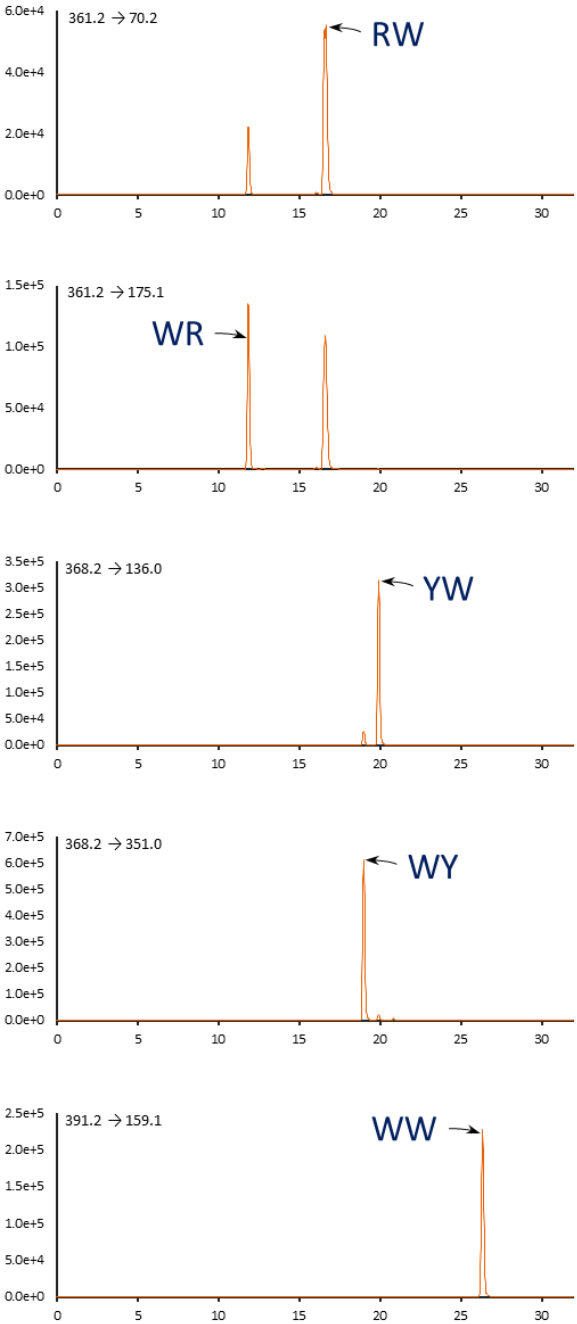

361.2 → 70.2
RW
361.2 → 175.1
WR
368.2 → 136.0
YW
368.2 → 351.0
WY
391.2 → 159.1
WW

Table A1. Optimized MRM settings of 361 dipeptides

| Dipeptide | Precursor Ion | Product Ion | Collision Energy | Mode |
|---|---|---|---|---|
| GG | 133 | 76.1 | 4 | CE |
| GA | 147.1 | 90.1 | 8 | CE |
| HH | 147.1 | 110 | 8 | CE |
| AG | 147.1 | 147.1 | 0 | CE |
| AA | 161.1 | 90.1 | 8 | CE |
| SG | 163.1 | 60 | 12 | CE |
| GS | 163.1 | 105.9 | 8 | CE |
| RR | 166.1 | 70.1 | 16 | CE |
| PG | 173.1 | 70.1 | 12 | LC |
| GP | 173.1 | 116.1 | 8 | LC |
| VG | 175.1 | 55.1 | 28 | LC |
| GV | 175.1 | 129.1 | 4 | LC |
| SA | 177.1 | 60.2 | 16 | CE |
| GT | 177.1 | 74.1 | 8 | CE |
| TG | 177.1 | 74.1 | 8 | CE |
| AS | 177.1 | 106 | 4 | CE |
| PA | 187.1 | 70.1 | 12 | LC |
| AP | 187.1 | 116.1 | 8 | LC |
| VA | 189.1 | 72.2 | 8 | LC |
| GI | 189.1 | 86.2 | 12 | LC |
| GL | 189.1 | 86.2 | 12 | LC |
| IG | 189.1 | 86.2 | 12 | LC |
| LG | 189.1 | 86.2 | 12 | LC |
| AV | 189.1 | 118.1 | 8 | LC |
| NG | 190.1 | 110.1 | 16 | CE |
| GN | 190.1 | 173.1 | 4 | CE |
| TA | 191.1 | 74.1 | 12 | CE |
| DG | 191.1 | 110 | 16 | CE |
| AT | 191.1 | 120.1 | 8 | CE |
| GD | 191.1 | 134.1 | 8 | CE |
| SS | 193.1 | 60 | 12 | CE |
| PS | 203.1 | 70.1 | 20 | LC |
| IA | 203.1 | 86.1 | 8 | LC |

| | | | | |
|---|---|---|---|---|
| LA | 203.1 | 86.1 | 8 | LC |
| SP | 203.1 | 116.2 | 8 | LC |
| AI | 203.1 | 132.1 | 4 | LC |
| AL | 203.1 | 132.1 | 4 | LC |
| GK | 204.1 | 84.1 | 20 | CE |
| GQ | 204.1 | 84.1 | 20 | CE |
| KG | 204.1 | 84.1 | 20 | CE |
| QG | 204.1 | 84.1 | 20 | CE |
| NA | 204.1 | 124.1 | 16 | CE |
| AN | 204.1 | 133.1 | 4 | CE |
| SV | 205.1 | 60.2 | 20 | LC |
| VS | 205.1 | 72.1 | 8 | LC |
| EG | 205.1 | 84.2 | 20 | CE |
| GE | 205.1 | 84.2 | 20 | CE |
| DA | 205.1 | 90.1 | 4 | CE |
| AD | 205.1 | 134.1 | 4 | CE |
| ST | 207.1 | 59.9 | 16 | CE |
| TS | 207.1 | 74.1 | 12 | CE |
| MG | 207.1 | 103.9 | 4 | LC |
| GM | 207.1 | 150.1 | 4 | LC |
| PP | 213.1 | 70.2 | 16 | LC |
| GH | 213.1 | 110.1 | 14 | CE |
| HG | 213.1 | 110.1 | 14 | CE |
| PV | 215.1 | 70.2 | 12 | LC |
| VP | 215.1 | 116 | 8 | LC |
| PT | 217.1 | 70.1 | 20 | LC |
| TP | 217.1 | 115.9 | 8 | LC |
| VV | 217.2 | 72 | 12 | LC |
| AK | 218.1 | 84.1 | 24 | CE |
| AQ | 218.1 | 84.1 | 24 | CE |
| KA | 218.1 | 84.1 | 24 | CE |
| QA | 218.1 | 84.1 | 24 | CE |
| SI | 219.1 | 60 | 22 | LC |
| SL | 219.1 | 60 | 22 | LC |
| VT | 219.1 | 72 | 8 | LC |
| TV | 219.1 | 74.1 | 12 | LC |

| | | | | |
|---|---|---|---|---|
| IS | 219.1 | 86 | 8 | CE |
| LS | 219.1 | 86 | 8 | CE |
| AE | 219.1 | 148.1 | 4 | CE |
| EA | 219.1 | 201.1 | 0 | CE |
| SN | 220.1 | 60.1 | 20 | CE |
| NS | 220.1 | 140 | 12 | CE |
| SD | 221.1 | 60.1 | 20 | CE |
| TT | 221.1 | 74 | 8 | CE |
| MA | 221.1 | 103.9 | 4 | LC |
| DS | 221.1 | 105.9 | 8 | CE |
| AM | 221.1 | 150.1 | 4 | LC |
| FG | 223.1 | 120.1 | 12 | LC |
| GF | 223.1 | 120.1 | 12 | LC |
| AH | 227.1 | 110 | 18 | CE |
| HA | 227.1 | 110 | 18 | CE |
| PI | 229.2 | 70.1 | 20 | LC |
| PL | 229.2 | 70.1 | 20 | LC |
| IP | 229.2 | 116.1 | 8 | LC |
| LP | 229.2 | 116.1 | 8 | LC |
| PN | 230.1 | 70 | 16 | CE |
| NP | 230.1 | 196 | 12 | LC |
| DP | 231.1 | 70.1 | 28 | LC |
| PD | 231.1 | 70.1 | 28 | LC |
| VI | 231.2 | 72.1 | 10 | LC |
| VL | 231.2 | 72.1 | 10 | LC |
| IV | 231.2 | 86.2 | 8 | LC |
| LV | 231.2 | 86.2 | 8 | LC |
| GR | 232.1 | 70.2 | 32 | CE |
| RG | 232.1 | 70.2 | 32 | CE |
| VN | 232.1 | 72 | 16 | LC |
| NV | 232.1 | 152.1 | 12 | LC |
| DV | 233.1 | 72.2 | 12 | CE |
| VD | 233.1 | 72.2 | 12 | CE |
| TI | 233.1 | 74.1 | 12 | LC |
| TL | 233.1 | 74.1 | 12 | LC |
| IT | 233.1 | 86 | 8 | LC |

| | | | | |
|---|---|---|---|---|
| LT | 233.1 | 86 | 8 | LC |
| SQ | 234.1 | 60.2 | 16 | CE |
| TN | 234.1 | 74.1 | 16 | CE |
| KS | 234.1 | 84.1 | 24 | CE |
| QS | 234.1 | 84.1 | 24 | CE |
| SK | 234.1 | 84.1 | 24 | CE |
| NT | 234.1 | 217.1 | 0 | CE |
| SE | 235.1 | 60.1 | 20 | CE |
| TD | 235.1 | 74.1 | 12 | CE |
| DT | 235.1 | 217 | 0 | CE |
| ES | 235.1 | 217 | 0 | CE |
| MS | 237.1 | 104 | 8 | LC |
| AF | 237.1 | 120 | 14 | LC |
| FA | 237.1 | 120 | 14 | LC |
| SM | 237.1 | 150.2 | 4 | LC |
| GY | 239.1 | 136.1 | 16 | LC |
| YG | 239.1 | 136.1 | 16 | LC |
| HS | 243.1 | 110 | 20 | CE |
| SH | 243.1 | 110 | 20 | CE |
| PK | 244.1 | 70.2 | 30 | LC |
| PQ | 244.1 | 70.2 | 30 | LC |
| QP | 244.1 | 116.1 | 16 | LC |
| KP | 244.2 | 84.2 | 20 | LC |
| PE | 245.1 | 70 | 16 | LC |
| EP | 245.1 | 116.1 | 12 | LC |
| II | 245.2 | 86.1 | 8 | LC |
| IL | 245.2 | 86.1 | 8 | LC |
| LI | 245.2 | 86.1 | 8 | LC |
| LL | 245.2 | 86.1 | 8 | LC |
| QV | 246.1 | 72.2 | 18 | LC |
| VK | 246.1 | 72.2 | 18 | LC |
| VQ | 246.1 | 72.2 | 18 | LC |
| IN | 246.1 | 86.1 | 10 | CE |
| LN | 246.1 | 86.1 | 10 | CE |
| AR | 246.2 | 70.1 | 32 | CE |
| RA | 246.2 | 70.1 | 32 | CE |

| | | | | |
|---|---|---|---|---|
| KV | 246.2 | 84.2 | 20 | LC |
| NI | 246.2 | 229 | 0 | LC |
| NL | 246.2 | 229 | 0 | LC |
| PM | 247.1 | 70.1 | 16 | LC |
| VE | 247.1 | 72.1 | 16 | LC |
| DI | 247.1 | 86 | 14 | LC |
| DL | 247.1 | 86 | 14 | LC |
| ID | 247.1 | 86 | 14 | LC |
| LD | 247.1 | 86 | 14 | LC |
| NN | 247.1 | 87.1 | 16 | CE |
| MP | 247.1 | 116 | 8 | LC |
| EV | 247.1 | 229.1 | 0 | LC |
| KT | 248.1 | 84.2 | 26 | CE |
| TK | 248.1 | 84.2 | 26 | CE |
| TQ | 248.1 | 84.2 | 26 | CE |
| ND | 248.1 | 87.1 | 12 | CE |
| DN | 248.1 | 88 | 16 | CE |
| QT | 248.1 | 230.1 | 4 | CE |
| VM | 249.1 | 72.1 | 12 | LC |
| TE | 249.1 | 74.2 | 12 | CE |
| MV | 249.1 | 104.1 | 8 | LC |
| DD | 249.1 | 133.9 | 4 | CE |
| ET | 249.1 | 231 | 0 | CE |
| TM | 251.1 | 74.2 | 16 | LC |
| MT | 251.1 | 104 | 8 | LC |
| PH | 253.1 | 70.1 | 28 | LC |
| HP | 253.1 | 110.1 | 20 | LC |
| FS | 253.1 | 120.1 | 14 | LC |
| SF | 253.1 | 120.1 | 14 | LC |
| YA | 253.1 | 136.1 | 12 | LC |
| AY | 253.1 | 182 | 4 | LC |
| HV | 255.1 | 110 | 22 | LC |
| VH | 255.1 | 110 | 22 | LC |
| HT | 257.1 | 110.1 | 22 | CE |
| TH | 257.1 | 110.1 | 22 | CE |
| IK | 260.2 | 84 | 26 | LC |

| | | | | |
|---|---|---|---|---|
| KI | 260.2 | 84 | 26 | LC |
| KL | 260.2 | 84 | 26 | LC |
| LK | 260.2 | 84 | 26 | LC |
| IQ | 260.2 | 147.1 | 8 | LC |
| LQ | 260.2 | 147.1 | 8 | LC |
| QI | 260.2 | 242.2 | 4 | LC |
| QL | 260.2 | 242.2 | 4 | LC |
| KN | 261.1 | 84.2 | 28 | CE |
| NK | 261.1 | 84.2 | 28 | CE |
| NQ | 261.1 | 84.2 | 28 | CE |
| QN | 261.1 | 84.2 | 28 | CE |
| IE | 261.1 | 86 | 14 | LC |
| LE | 261.1 | 86 | 14 | LC |
| EI | 261.1 | 243.1 | 4 | LC |
| EL | 261.1 | 243.1 | 4 | LC |
| RS | 262.1 | 70 | 32 | CE |
| SR | 262.1 | 70 | 32 | CE |
| DK | 262.1 | 84.1 | 28 | CE |
| DQ | 262.1 | 84.1 | 28 | CE |
| KD | 262.1 | 84.1 | 28 | CE |
| QD | 262.1 | 84.1 | 28 | CE |
| NE | 262.1 | 136.1 | 24 | CE |
| GW | 262.1 | 187.9 | 16 | LC |
| EN | 262.1 | 244 | 0 | CE |
| WG | 262.1 | 245.2 | 4 | LC |
| PF | 263.1 | 70.1 | 12 | LC |
| DE | 263.1 | 84 | 26 | CE |
| ED | 263.1 | 84 | 26 | CE |
| IM | 263.1 | 86.2 | 12 | LC |
| LM | 263.1 | 86.2 | 12 | LC |
| MI | 263.1 | 104 | 8 | LC |
| ML | 263.1 | 104 | 8 | LC |
| FP | 263.1 | 120.1 | 20 | LC |
| NM | 264.1 | 60.9 | 32 | LC |
| MN | 264.1 | 133.2 | 8 | LC |
| MD | 265.1 | 104 | 8 | LC |

| | | | | |
|---|---|---|---|---|
| DM | 265.1 | 150.1 | 8 | LC |
| VF | 265.2 | 72.1 | 12 | LC |
| FV | 265.2 | 120 | 16 | LC |
| TF | 267.1 | 74.1 | 12 | LC |
| FT | 267.1 | 120 | 20 | LC |
| YS | 269.1 | 136.1 | 12 | LC |
| SY | 269.1 | 182 | 8 | LC |
| HI | 269.2 | 110 | 22 | LC |
| HL | 269.2 | 110 | 22 | LC |
| IH | 269.2 | 110 | 22 | LC |
| LH | 269.2 | 110 | 22 | LC |
| HN | 270.1 | 110 | 26 | CE |
| NH | 270.1 | 110 | 26 | CE |
| DH | 271.1 | 110.1 | 26 | CE |
| HD | 271.1 | 110.1 | 26 | CE |
| PR | 272.2 | 70.2 | 32 | LC |
| RP | 272.2 | 70.2 | 32 | LC |
| RV | 274.2 | 70.2 | 28 | LC |
| VR | 274.2 | 72 | 24 | LC |
| KK | 275.1 | 84 | 28 | CE |
| KQ | 275.1 | 84 | 28 | CE |
| QK | 275.1 | 84 | 28 | CE |
| QQ | 275.1 | 84 | 28 | CE |
| EK | 276.1 | 84 | 28 | LC |
| EQ | 276.1 | 84 | 28 | LC |
| KE | 276.1 | 84 | 28 | LC |
| QE | 276.1 | 84 | 28 | LC |
| AW | 276.1 | 188.1 | 16 | LC |
| WA | 276.1 | 259.1 | 4 | LC |
| RT | 276.2 | 70.2 | 36 | CE |
| TR | 276.2 | 70.2 | 36 | CE |
| EE | 277.1 | 84.1 | 32 | CE |
| MQ | 278.1 | 147.1 | 8 | LC |
| QM | 278.1 | 260 | 4 | LC |
| KM | 278.2 | 84.2 | 26 | CE |
| MK | 278.2 | 84.2 | 26 | CE |

| | | | | |
|---|---|---|---|---|
| PY | 279.1 | 70.1 | 16 | LC |
| ME | 279.1 | 104 | 12 | LC |
| YP | 279.1 | 136.2 | 16 | LC |
| EM | 279.1 | 261.1 | 4 | LC |
| IF | 279.2 | 86.2 | 10 | LC |
| LF | 279.2 | 86.2 | 10 | LC |
| FI | 279.2 | 120.2 | 12 | LC |
| FL | 279.2 | 120.2 | 12 | LC |
| NF | 280.1 | 91 | 40 | LC |
| FN | 280.1 | 120 | 20 | LC |
| VY | 281.1 | 72.2 | 12 | LC |
| MM | 281.1 | 104 | 8 | LC |
| DF | 281.1 | 120.1 | 16 | LC |
| FD | 281.1 | 120.1 | 16 | LC |
| YV | 281.1 | 136.2 | 12 | LC |
| TY | 283.1 | 74.1 | 16 | LC |
| YT | 283.1 | 136.2 | 12 | LC |
| HK | 284.1 | 110.1 | 26 | CE |
| HQ | 284.1 | 110.1 | 26 | CE |
| KH | 284.1 | 110.1 | 26 | CE |
| QH | 284.1 | 110.1 | 26 | CE |
| EH | 285.1 | 110.1 | 26 | CE |
| HE | 285.1 | 110.1 | 26 | CE |
| HM | 287.1 | 110 | 24 | LC |
| MH | 287.1 | 110 | 24 | LC |
| RI | 288.2 | 70 | 36 | LC |
| IR | 288.2 | 175.1 | 16 | LC |
| LR | 288.2 | 175.1 | 16 | LC |
| RL | 288.2 | 175.1 | 16 | LC |
| NR | 289.2 | 70.1 | 34 | CE |
| RN | 289.2 | 70.1 | 34 | CE |
| DR | 290.1 | 70.1 | 36 | CE |
| RD | 290.1 | 70.1 | 36 | CE |
| SW | 292.1 | 188.1 | 16 | LC |
| WS | 292.1 | 275.1 | 4 | LC |
| FK | 294.1 | 120.1 | 26 | LC |

| | | | | |
|---|---|---|---|---|
| FQ | 294.1 | 120.1 | 26 | LC |
| QF | 294.1 | 120.1 | 26 | LC |
| KF | 294.2 | 84.1 | 20 | LC |
| FE | 295.1 | 120.1 | 20 | LC |
| EF | 295.1 | 277.1 | 4 | LC |
| IY | 295.2 | 86.2 | 10 | LC |
| LY | 295.2 | 86.2 | 10 | LC |
| YI | 295.2 | 136.2 | 14 | LC |
| YL | 295.2 | 136.2 | 14 | LC |
| NY | 296.1 | 85 | 20 | LC |
| YN | 296.1 | 136.1 | 16 | LC |
| MF | 297.1 | 104.1 | 8 | LC |
| FM | 297.1 | 120.1 | 16 | LC |
| YD | 297.1 | 136.2 | 12 | LC |
| DY | 297.1 | 182.1 | 8 | LC |
| PW | 302.1 | 70.2 | 24 | LC |
| WP | 302.1 | 285 | 8 | LC |
| FH | 303.1 | 110 | 24 | LC |
| HF | 303.1 | 110 | 24 | LC |
| QR | 303.2 | 70.1 | 42 | CE |
| RK | 303.2 | 70.1 | 42 | CE |
| RQ | 303.2 | 70.1 | 42 | CE |
| KR | 303.2 | 84.1 | 36 | CE |
| ER | 304.2 | 70 | 40 | CE |
| RE | 304.2 | 70 | 40 | CE |
| WV | 304.2 | 159.1 | 16 | LC |
| VW | 304.2 | 188.2 | 12 | LC |
| WT | 306.1 | 159.1 | 12 | LC |
| TW | 306.1 | 188.1 | 20 | LC |
| MR | 306.2 | 70.1 | 42 | LC |
| RM | 306.2 | 70.1 | 42 | LC |
| YQ | 310.1 | 136.1 | 16 | LC |
| QY | 310.1 | 292.1 | 8 | LC |
| KY | 310.2 | 84.1 | 24 | LC |
| YK | 310.2 | 91.1 | 48 | LC |
| YE | 311.1 | 136.1 | 16 | LC |

| | | | | |
|---|---|---|---|---|
| EY | 311.1 | 293.1 | 4 | LC |
| RH | 312.2 | 69.8 | 36 | CE |
| HR | 312.2 | 110.1 | 32 | CE |
| MY | 313.1 | 104 | 12 | LC |
| YM | 313.1 | 136.1 | 12 | LC |
| FF | 313.2 | 120.1 | 16 | LC |
| IW | 318.2 | 86.1 | 16 | LC |
| LW | 318.2 | 86.1 | 16 | LC |
| WI | 318.2 | 301.2 | 8 | LC |
| WL | 318.2 | 301.2 | 8 | LC |
| HY | 319.1 | 110 | 24 | LC |
| YH | 319.1 | 110 | 24 | LC |
| WN | 319.1 | 159.2 | 20 | LC |
| NW | 319.1 | 188.2 | 16 | LC |
| DW | 320.1 | 188 | 16 | LC |
| WD | 320.1 | 303.1 | 8 | LC |
| RF | 322.2 | 70.1 | 44 | LC |
| FR | 322.2 | 120.1 | 32 | LC |
| FY | 329.1 | 120 | 16 | LC |
| YF | 329.1 | 136 | 20 | LC |
| KW | 333.2 | 84.1 | 28 | LC |
| WK | 333.2 | 159 | 26 | CE |
| WQ | 333.2 | 159 | 26 | CE |
| QW | 333.2 | 315 | 8 | LC |
| EW | 334.1 | 316.2 | 4 | LC |
| WE | 334.1 | 317.1 | 8 | LC |
| MW | 336.1 | 188 | 20 | LC |
| WM | 336.1 | 319.2 | 8 | LC |
| RY | 338.2 | 70.1 | 36 | LC |
| YR | 338.2 | 136.2 | 28 | LC |
| HW | 342.2 | 110 | 30 | LC |
| WH | 342.2 | 110 | 30 | LC |
| YY | 345.1 | 136.1 | 16 | LC |
| FW | 352.2 | 120.1 | 20 | LC |
| WF | 352.2 | 335 | 8 | LC |
| RW | 361.2 | 70.2 | 48 | LC |

| | | | | |
|---|---|---|---|---|
| WR | 361.2 | 175.1 | 24 | LC |
| YW | 368.2 | 136 | 20 | LC |
| WY | 368.2 | 351 | 8 | LC |
| WW | 391.2 | 159.1 | 20 | LC |