The application of mass spectrometry to Ginkgo biloba analysis and identification of phosphorylated proteins in response to elevated level of cCMP.

Ding, Shujing

The application of mass spectrometry to *Ginkgo biloba* analysis and identification of phosphorylated proteins in response to elevated level of cCMP

By

Shujing Ding

A thesis submitted for the degree of Doctor of Philosophy in the University of Wales Swansea

September 2006

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

DECLARATION

This work has not previously been accepted in substance for any degree, and is not being concurrently submitted in candidature for any degree.

Signed .... (Candidate)

Date ..... 23 Nov. 2006 .....

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed .... (Candidate)

Date ..... 23 Nov. 2006 .....

Signed .. .. (Supervisor)

Date ..... 23 Nov. 2006 .....

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organizations.

Signed ..... (Candidate)

Date ..... 23 Nov. 2006 .....

ACKNOWLEDGEMENTS

I would firstly like to thank my supervisors, Prof. A.G. Brenton and Prof. R.P. Newton, for providing me with the opportunity of entering the field of mass spectrometry and conducting the challenging research that I am interested in, for their encouragement, full support and advice.

I would like to thank Dr Ed Dudley for his training, assistance, advice, support and friendship during my time in BAMS facility, for correcting part of my thesis.

I would like to thank Dr Sue Plummer and Dr James Tang at the Obsidian Research Limited, Port Talbot, for providing standard materials and urine samples.

I would like to thank Dr Liz Bond, Dr Sarah Forbes-Roberston, Dr Chris Smith, Dr Ping Hu, Dr Qingbao Song and Dr Lijuan Chen for their informative discussions and suggestions. I would like to thank Penny Diffley, Alun Jones and Brian Cooper for their technical support.

I would also like to thank my friends and fellow colleagues in BAMS facility and MSRU, too numerous to mention fully here, for their kind help and friendship.

I would like to thank Zhejiang University of Technology for providing me the opportunity to study in UK and Chinese government for the financial support.

Finally, I would like to thank all my family, especially my husband Xiaochuan Yang and my son Ting Yang, for their understanding and constant love, otherwise this would not be possible.

Summary

Mass spectrometry is widely used nowadays especially in the fields of pharmaceutical and proteomics research. *Ginkgo biloba* is one of the top selling phytopharmaceuticals in the US and Europe. The two major active components of *Ginkgo* leaf extract are the flavonoids and terpene lactones. Identification, determination, as well as the physiological effects of these two sets of compounds have been of increasing interest over the last 20 years. In this thesis, systematic qualitative and quantitative studies of the flavonoids and terpene lactones in *Ginkgo biloba* by liquid chromatography / mass spectrometry have been undertaken. Also in this thesis, mass spectrometric methodology was developed and applied to the identification of the proteins specifically phosphorylated in response to cCMP.

Structural information of *Ginkgo biloba* flavonoids and terpene lactones, the fragment of compounds were obtained on both a LCQ ion trap and Q-TOF mass spectrometer. The tentative fragment pathways were proposed and used for structural elucidation of some unknown components in *Ginkgo biloba* commercial products. Capillary column separation of *Ginkgo biloba* commercial product was evaluated and fingerprint profiles of five *Ginkgo biloba* commercial products were compared.

A reverse-phase high-performance liquid chromatography electrospray ionisation (RP-HPLC/ESI) mass spectrometry method was developed and validated for the simultaneous determination of ten major active components in *Ginkgo biloba* extract (bilobalide, ginkgolides A, B, C, quercetin, kaempferol, isorhamnetin, rutin hydrate, quercetin-3-β-D-glucoside and quercitrin hydrate).

The quantitative determination of flavonoids and terpene lactones by LC/MS in human urine after consumption of *Ginkgo biloba* product was developed. The online solid-phase extraction and capillary column with column-switch technique require minimum sample pre-treatment and both flavonoids and terpene lactones can be detected simultaneously.

The mass accuracy at high molecular weight by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry was investigated to resolve a question on mass accuracy which had been observed to be relatively low for high mss proteins. Bovine serum albumin (BSA) was employed as a model compound and strategies to improve mass measurement at high mass were examined.

LC/MS was applied in part of the cyclic nucleotide project in the School of Biological Science. Since cAMP and cGMP are recognized second messengers and play important roles in signal transduction, to elucidate the function of cCMP in signal transduction, efforts were made to identify the cCMP-responsive protein kinase substrates. Methodology of specific enrichment of phosphopeptides using immobilized metal affinity chromatography (IMAC) was developed, phosphorylated proteins responding specifically to cCMP were proposed, and this supports the relationship of cCMP with cell hyperproliferation.

Publications

Journal papers

1. Ding, S., Dudley, E., Plummer, S., Tang, J., Newton, R.P. and Brenton, A.G., Quantitative determination of major active components in *Ginkgo biloba* dietary supplements by liquid chromatography mass spectrometry, Rapid Commun. Mass Spectrom., **20**, 2753 (2006)

2. Ding, S., Dudley, E., Chen, L., Plummer, S., Tang, J., Newton, R.P. and Brenton, A.G., Determination of active components of *Ginkgo biloba* in human urine by capillary HPLC/MS with column switch on-line purification, Rapid Commun. Mass Spectrom., **24**, 3619 (2006)

Poster / Presentations

1. Ding, S., Dudley, E., Brenton, A.G. and Newton, R.P., A study of cAMP dependent protein kinase in brain tissue, 27th BMSS annual meeting, Derby 2004

2. Brenton, A.G., Ding, S., Dudley, E. and Newton, R.P., A new approach to high-mass calibration in linear time-of-flight, 27th BMSS annual meeting, Derby 2004

3. Ding, S., Dudley, E., Bond, L., Newton, R.P. and Brenton, A.G., Identification by mass spectrometry of selective proteomic responses in the mouse brain to challenge with cCMP, 17th International mass spectrometry conference, Prague, 2006

Abbreviations

α	selectivity factor
ac	alternating current
Abs	absorbance
AC	adenylyl cyclase
ACN	acetonitrile
AMP	adenosine monophosphate
amu	arbitrary mass unit
APCI	atmosphere pressure chemical ionisation
APS	ammonium persulfate
ATP	adenosine 5'-triphosphate
BL	bilobalide
BSA	bovine serum albumin
cAMP	adenosine 3',5'-cyclic monophosphate
cCMP	cytidine 3',5'-cyclic monophosphate
cGMP	guanosine 3',5'-cyclic monophosphate
CI	chemical ionisation
CID	collision induced dissociation
Da	Dalton
DC	direct current
DNA	deoxyribonucleic acid
DTT	dethiothreitol
EDTA	ethylenediaminetetra-acetic acid disodium salt
ESI	electrospray ionisation

FAB/MIKES	fast- atom bombardment/mass analysed ion kinetic energy scanning
fmol	femtomole
FTICR	Fourier transform ion cyclotron resonance
GA	ginkgolide A
GB	ginkgolide B
GC	ginkgolide C
GC (Chapter 7)	guanylyl cyclase
GC/MS	gas chromatography/mass spectrometry
GDP	guanosine diphosphate
GTP	guanosine 5'-triphosphate
h, hr	hour
HCl	hydrochloric acid
HPLC	high performance liquid chromatography
HPLC/MS	high performance liquid chromatography/mass spectrometry
i.d.	internal diameter
IEF	isoelectric focusing
IMAC	immobilized metal affinity chromatography
IPG	immobilized pH gradient
IR	isorhamnetin
K'	capacity factor
K	partition coefficient
kDa	kilo Dalton
KF	kaempferol

LC	liquid chromatography
LC/MS	liquid chromatography/mass spectrometry
LLE	liquid-liquid extraction
LOD	limit of detection
LOQ	limit of quantitation
MALDI	matrix-assisted laser desorption/ionisation
mg	milligram
$[M+H]^+$	molecular mass + mass of proton
$[M-H]^+$	molecular mass - mass of proton
min	minutes
mL	millilitre
mm	millimeter
mM	millimolar
mRNA	messenger RNA
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MS^n	tandem mass spectrometry to the power n
m/z	mass to charge ratio
Nano ESI	nanoflow electrospray ionisation
nL	nanolitre
nm	nanometer
ns	nanosecond
ODS	octadecylsilane
PAGE	polyacrylamide gel electrophoresis

pI	isoelectric point
PKG	cGMP-dependent protein kinase
PMF	peptide mass fingerprinting
PMSF	phenylmethylsulfonyl fluoride
pmol	picomole
ppm	parts per million
PTM	post-translational modification
QC	quality control
QD	quercetin
QG	quercetin-3-β-D-glucoside
QH	quercetin-3-rhamnoside
Q-TOF	quadrupole time-of-flight
RF	radiofrequency
RH	rutin
r.p.m.	revolution per minute
RNA	ribonucleic acid
Rs	resolution
s, sec	second
SDS	sodium dodecyl sulphate
SPE	solid phase extraction
TEMED	N,N,N',N'-tetramethylethylenediamine
TFA	trifluoroacetic acid
TIC	total ion chromatogram
TLC	thin layer chromatography

TMS	trimethylsilyl
TOF	time-of-flight
Tris	tris-HCl
UV	ultra violet light
V	voltage
v/v	volume per volume
Vh	volt hours
Vo	void volume
Vt	total volume of the column
W	Watt
W	peak width
$W_{1/2}$	peak width at half height
w/v	weight per volume
XIC	extracted ion chromatogram
μg	microgram
μL	microlitre
μmol	micromole
1-D	one-dimensional
2-D	two-dimensional

Contents

Chapter 1

Introduction to HPLC and mass spectrometry

1.1 Introduction to chromatography

Chromatography is a technique which is widely used for the separation of analyte mixtures and the identification of their individual components. Chromatography has been defined as

"*A method used primarily for the separation of components of a sample, in which the components may be distributed between two phases, one of which is stationary while the other moves. The stationary phase may be a solid, liquid supported on a solid, or a gel. The stationary phase may be packed in a column, spread as a layer, or distributed as a film and the mobile phase may be gaseous or liquid*"

by the International Union of Pure and Applied Chemistry (IUPAC)[1]. The chromatography process occurs as a result of repeated absorption/desorption steps during the movement of analyte along the stationary phase. The separation is based on the differences in the equilibrium coefficients of the individual analytes in the sample. Based upon the physical nature of the mobile phase chromatography is classified accordingly: gas chromatography (GC), liquid chromatography (LC), capillary electrophoresis chromatography (CEC) and supercritical fluid chromatography (SFC).

1.1.1 History of chromatography

Chromatography is considered as one of the most dynamic and versatile analytical methods, it has advanced greatly since its beginnings in the twentieth century. A Russian botanist Tswett[2] first introduced the concept of chromatography in 1903, he reported the novel separation and isolation of various plant pigments by using a glass

column packed with calcium carbonate. In his experiment, coloured bands relating to the differing plant components passed through the column sequentially and for this reason the process was termed chromatography which means colour writing. Unfortunately, his remarkable work was regarded as being of little merit and was ignored by his contemporaries for nearly twenty years. It was not until 1931 that Kuhn and his co-workers[3] applied Tswett's method to the separation of isomers of polyene pigments which promoted the acceptance of the method. This research generated renewed interest and recognition of chromatography as a valuable separation method. In 1941, liquid-liquid partition chromatography was developed by Martin and Synge[4], of the Wool Industries Research Association in England. A solid support made of water retained on a solid support of silica gel was used as the stationary phase. In 1944, Martin with his colleagues Consden and Gordon introduced paper chromatography[5] in which paper acts as a flat "column". It was one of the key enabling-techniques that lead to Sanger's 1958 Nobel Prize for determining the first amino acid sequence of the protein insulin, although it was later replaced by thin-layer chromatography, Edman degradation and various forms of electrophoresis. In 1952, James and Martin developed the technique of gas-liquid chromatography and demonstrated the integrated theory of this separation technique[6]. Because of his remarkable contribution to chromatography Martin was award the Nobel Prize for Chemistry in 1952.

HPLC was invented in 1966, being named high-performance liquid chromatography by Horvath who reported a packing material comprising of silanized silica and demonstrated that high pressure and small particle size provide improved separation. The first commercial HPLC system was not available until 1969 when the high pressure and small particle size were applied to the commercial column. Theoretical

understanding of gas chromatography and later of liquid chromatography was developed when the concept of the "height equivalent to a theoretical plate" was put forward[7-9].

HPLC was popularized throughout the 1970s since it provided more precise and rapid separations required for many areas of biochemical research. After nearly 40 years development, chromatography plays an even more important role in a wide range of scientific studies. It is an extremely versatile and robust separation technique which is widely used in fields such as the analysis of compounds in pharmaceuticals, natural products and biology.

1.1.2 Theory of chromatography

Chromatography is a physical method of separation in which the components being separated and distributed between two immiscible phases, one phase is held static which is termed as stationary phase, the other phase, termed as the mobile phase, travels past the stationary phase under either gravity or pressure or electrophoretic pressure. The stationary phase is a dispersed medium with a large surface area and it can be either solid or liquid spread over an inert support or spread as a thin film on the walls of the column. The mobile phase can be gas (GC), liquid (LC), electrolyte (CE) or a supercritical fluid (SFC). Chromatographic theory was first proposed in 1941 by Nobel Prize winners Martin and Synge. In the 1950s, Craig[10] and Glueckauf[11] developed so-called "plate theory"; van Deemter[12] and Giddings[13] established the rate theory. The following sections discuss the important parameters underlying the theory of chromatography.

1.1.2.1 Distribution coefficient (K)

Chromatographic separation occurs due to the differential migration of analytes along the column. The average rate of migration of an analyte through the column depends on the fraction of time spent in the stationary phase and on the affinity of the analyte to the stationary phase. Components that tend to reside in the mobile phase will move more quickly than those that prefer the stationary phase. The different tendencies of analytes to preferentially exist in the stationary or mobile phase allow their separation due to the continual equilibration between the two phases[14]. Each analyte is distributed between the two phases with its characteristic distribution coefficient, K, which is defined as

$$K=C_s/C_m \qquad \text{(Eq. 1.1)}$$

where C_s is the molar concentration of the analyte in the stationary phase and C_m is its molar concentration in the mobile phase. Analyte with a high distribution coefficient will move slowly through the column. Separation cannot be achieved between analytes with the same distribution coefficient.

1.1.2.2 Retention time (t_R)

The retention time (t_R) is the time taken for an analyte to move from the point of injection until it reach the detector. A typical chromatogram (Figure 1.1) shows the retention times and peak widths (W_A and W_B) of the analyte compounds A and B relative to the non-retained components. In this figure, t_M is the dead time, which is the time for unretained components to reach the detector. It is identical for all analytes in a given chromatographic system since they migrate with the same velocity (that of the

mobile phase). The adjusted retention time t_R' is the time for which an analyte is retained compared with that of an unretained compound and is therefore directly associated with the interaction of the analyte with the stationary phase. Their relationship is given by

$$t_R' = t_R - t_M \qquad \text{(Eq. 1.2)}$$

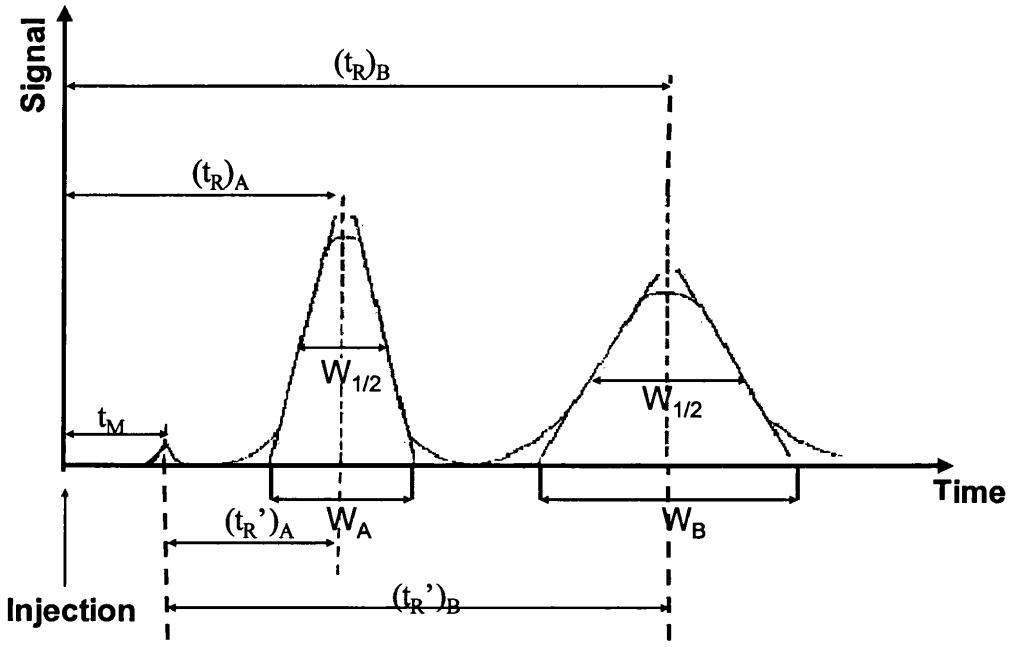


Figure 1.1 Chromatographic separation of two analytes A and B. t_m-dead time; t_R-retention time of an analyte; $W_{1/2}$-peak width at half maximum.

1.1.2.3 Capacity factor (k')

The capacity factor is used to described the migration rates of analytes through the chromatographic system, this can be calculated from the chromatogram by

$$k' = (t_R - t_M)/ t_M \qquad \text{(Eq. 1.3)}$$

It can be used to compare retentions on different instruments since it is independent of the mobile phase flow rate and the physical dimensions of the column. Ideally, the capacity factor of the components in the separation should be between 1 and 5. Values much lower than unity correspond to analytes eluting very rapidly which makes accurate determination of the retention times difficult. Values higher than 20 indicate long elution time and broad peak width.

1.1.2.4 Selectivity factor (α)

The selectivity factor (α) is defined as the ratio of the distribution coefficients of components A and B or the ratio of capacity factors of components A and B.

$$\alpha = K_B/K_A = k'_B/k'_A \qquad \text{(Eq. 1.4)}$$

The selectivity factor indicates how well a chromatographic system can separate two analytes A and B. B is more strongly retained component, where the selectivity factor is always greater than 1. If α=1 the retention times are identical and no separation can be obtained.

Although the selectivity factor describes the separation of peaks it cannot give accurate information, especially for LC systems, since it does not take into account peak widths. A better measure of separation of two neighbouring peaks is provided by column resolution, R_s, from which both retention difference and column efficiency are evaluated.

1.1.2.5 Resolution (R_s)

The resolution (R_s) between two components in a chromatogram is determined from the differences in the corresponding retention times and baseline peak widths (W_b). For symmetrical peaks of Gaussian shape it can be defined as

$$R_s = \frac{2[(t_R)_B - (t_R)_A]}{W_A + W_B} = \frac{1.18[(t_R)_B - (t_R)_A]}{W_{1/2A} + W_{1/2B}} \qquad \text{(Eq. 1.5)}$$

where $W_{1/2}$ is the peak width at half-height. Generally, baseline resolution can be achieved when $R_s > 1.5$ for peaks of similar size. An increase in the resolution value shows improved separation whereas smaller values illustrate that two analytes co-elute to some degree.

1.1.2.6 Separation efficiency

The efficiency of a given separation method is described in terms of the plate height (H) and the number of theoretical plates (N). Plate number originates from distillation theory and was first applied to chromatography by Martin and Synge[4] in 1941. Theoretical plates are considered as a series of narrow discrete sections in a chromatographic column or layer and equilibrium of the solute between the stationary and mobile phase is taken at each plate. Movement of analyte and mobile phase is viewed as a series of transfers from one plate to the next. The efficiency of a chromatographic system improves as the number of equilibrations and thus the number of theoretical plates increases. The plate number N of a chromatographic system is defined as,

$$N=16\,(t_R/w)^2 \qquad \text{(Eq. 1.6)}$$

$$\text{or } N=5.54\,(t_R/w_{1/2})^2 \qquad \text{(Eq. 1.7)}$$

The height of a theoretical plate H, is readily calculated providing the length of the column (normally in centimetres) is known,

$$H = L/N \qquad \text{(Eq. 1.8)}$$

where H = distance over which chromatographic equilibrium is achieved and is referred to as the height equivalent to a theoretical plate (HETP). Each theoretical plate is regarded as an equilibrium step, therefore the column efficiency increases as the number of theoretical plates increase, and the smaller the plate height, the more efficient the column. HETP is smaller for smaller column-packing materials, lower mobile phase flow rate, less viscous mobile phase and higher column temperatures.

1.1.2.7 van Deemter equation (band broadening theory)

In 1945, van Deemter[12] developed an equation that related the height of the theoretical plate to the linear mobile phase velocity and the various physical chemical properties of the solid phase and mobile phase,

$$\text{HETP} = A + \frac{B}{\upsilon} + (C_m + C_s)\upsilon = a(d_p) + \frac{b}{\upsilon} + c(d_p)^2\upsilon \qquad \text{(Eq. 1.9)}$$

The overall effect derived by the van Deemter equation plotting the HETP against mobile phase velocity is shown in Figure 1.2.

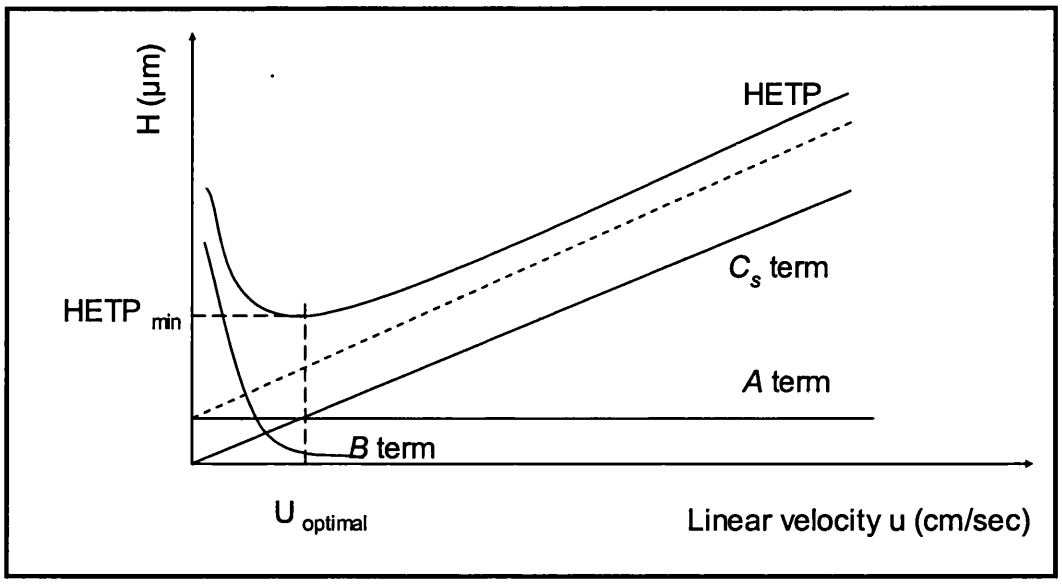


Figure 1.2 van Deemter plot, the height-equivalent of the theoretical plate versus linear mobile phase velocity

The terms used in the van Deemter equation are described below.

υ: the mobile phase velocity.

d_p: diameter of packing particles

A(a): the Eddy diffusion which represents the various of pathways by which a component finds its way through the column. The various possible pathways result in different retention times as the mobile phase carries sample molecules through the packed stationary phase and this effect is directly proportional to the diameter of the particles packing the column. To minimise this effect small and uniformity packed columns should be employed.

B(b): the longitudinal diffusion. It describe the band broadening process in which solutes diffuse from the concentrated centre of a zone to the more dilute regions ahead of and behind the zone centre. The contribution of the longitudinal diffusion is

inversely proportional to the mobile phase velocity. The B term in the van Deemter equation is negligible due to the small solute diffusion coefficient at practical flow rates in HPLC relative to gases (as in GC)[15].

C_s and C_m (c): the mass transfer effect which describes the time available for equilibrium of an analyte to be established between the mobile and stationary phases. The slow equilibrium will cause the chromatographic column to operate under non-equilibrium conditions. Analyte molecules at the front of a band are swept ahead before they have time to equilibrate with the stationary phase and thus be retained. Similarly, equilibrium is not reached at the trailing edge of the band, and molecules are left behind in the stationary phase by the fast moving mobile phase. The mass-transfer broadening is related to both the size of packing particles and the column flow rate. The faster the mobile phase moves, the bigger the packing size, the less time there is for equilibrium to be approached[7].

The van Deemter equation was first applied in gas chromatography but it has been found to be equally applicable to liquid chromatography. Although there are many other similar dispersion equations, the van Deemter equation remains one of the most accurate form and is used extensively in column design. HPLC has been developed extensively since it was realized that the separation efficiency can be largely improved by reducing particle-packing size and increased pressure.

1.1.3 An overview of types of liquid chromatography

The basic liquid chromatograph consists of a solvent system, a sample injection device, a column, a detector and a data system (Figure 1.3). The sample is introduced into the HPLC system by sample loop injection or by an autosampler. Mobile phase is delivered by the high pressure pump and carries the sample through the column where separation takes place with the analytes being detected, usually with a UV detector.

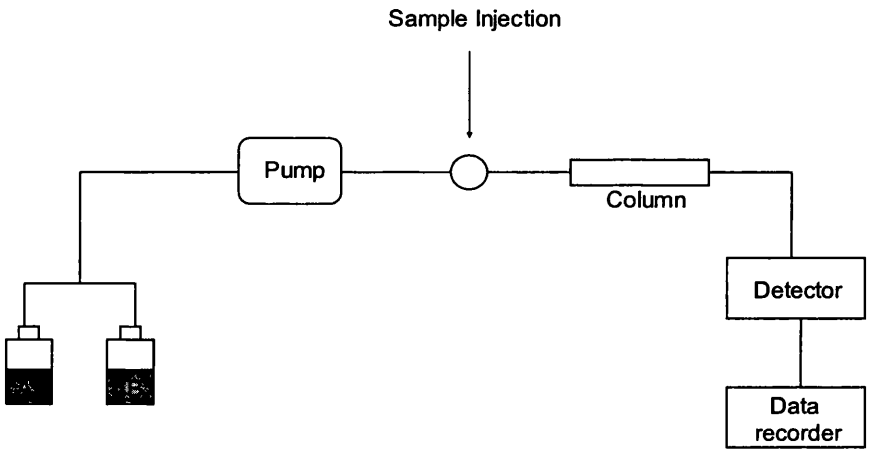


Figure 1.3 Schematic diagram of the components of a basic HPLC system

A number of column chromatography types are available for the analysis and separation of mixtures. They can be classified according to stationary or mobile phase, separation process and mechanism. Size-exclusion[16, 17], ion exchange[18, 19], affinity and reverse phase chromatography are the most widely used nowadays, only the types of chromatography employed in the work in this thesis are discussed.

1.1.3.1 Affinity chromatography

Affinity chromatography requires that an immobilized ligand, covalently coupled to the column's stationary phase, interacts specifically and reversibly with the solute of

interest. Figure 1.4 is a schematic representation that illustrates the processes involved in an affinity column based separation. As the sample passes through the column, the analyte of interest should bind in complementary fashion to the ligand (A) covalently attached to the solid support, while the rest of the solutes in the sample, incapable of specific binding, flow through without direct interaction. After the contaminants wash completely through the column, the analyte of interest (B) is then eluted via a variety of approaches.

Release of the bound analyte include ligand : analyte complex disruption by inclusion of competitive ligands in the mobile phase or by changes in mobile phase composition such as ionic strength or pH. Finally, the column is re-equilibrated for its next use.

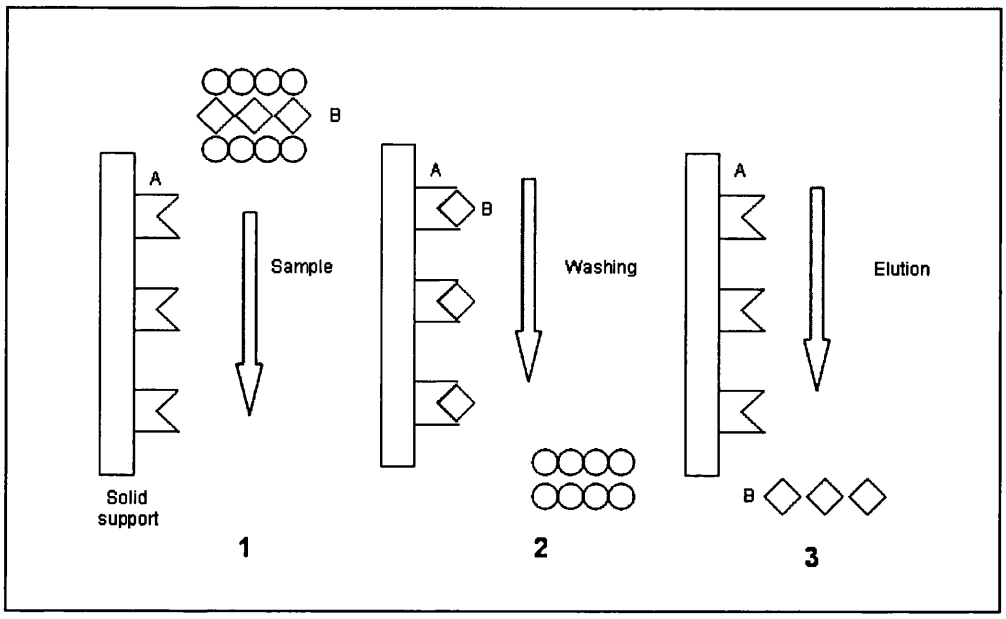


Figure 1.4 Illustration of affinity chromatography, in which sample B bind specifically to the column, whereas contaminants are washed off straight away, and sample B can be eluted later.

In order to be an effective affinity chromatographic system the column matrix must have certain properties: a) the ligand must contain functional groups that will allow for covalent coupling to the solid support without adversely altering its binding activity towards the analyte; b) the ligand must be capable of somewhat specific but reversible binding to the analyte of interest, and the affinity of the ligand:solute complex must be sufficient to provide for good binding; c) non-specific interaction between the column and unwanted sample components should be minimal; and d) the covalent linkages used to immobilize the ligand must be stable to all conditions employed during chromatography and column clean-up.

Affinity chromatography has been widely applied, for example, to the isolation of some naturally occurring products including enzymes such as chymotrypsin[20], alkaline phosphatase[21], lectins[22], specific peptides[23] and was utilised in the protein phosphorylation study in this thesis.

1.1.3.2 Reverse phase high-performance liquid chromatography (RP-HPLC)

Reverse phase HPLC is currently the most popular method of separation as it offers efficient separation with good resolutions over relatively short analysis times. About 80% of all chromatography applications use reverse phase separation.

1.1.3.2.1 Stationary phases in reverse phase chromatograph

In reverse phase HPLC, the packing materials are mainly silica-based and contain covalently bonded alkyl chains of different lengths. The most commonly used alkyl chains are n-octadecyl (C_{18}), octyl (C_8) and n-butyl (C_4), as well as phenyl (C_6H_5) and

amino (NH_2), each are responsible for imparting specific chromatographic character to the column. For example, the chemically bonded octadecylsilane (ODS) packing is the most commonly used. ODS packing silica has OH groups on its surface which react with an octadecyl group to give a strong non-polar stationary phase. The reaction is shown in Figure 1.5. Due to the steric hindrance of these functional groups, there are still some free silanol groups present and they will interact with the analytes causing peak tailing. To minimize this effect end capping[24] is introduced to cover the silanol group.

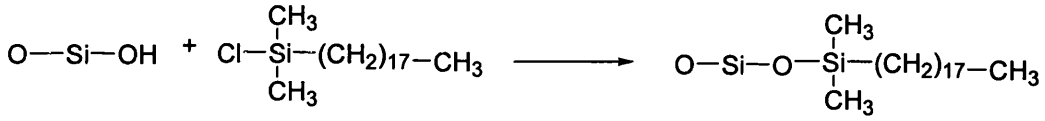


Figure 1.5 Formation of ODS reverse phase stationary phase

The silica-based supports are unstable in aqueous buffers at alkaline conditions since the hydrolysis of the siloxane bonds can occur. The use of polymer-based reversed phase resins such as polystyrene -divinylbenzene copolymer support offer the advantage of increased stability over a wide range of pH values from 1 to 13, as opposed to 2 to 8 for bonded silicas. Rapid developments in areas like proteomics led to the introduction of the monolith column, also known as a rod column which is made by filling capillaries with a monolith, a continuous separation bed is created inside the capillary using a polymerization mixture that undergoes chemical and physical changes inside the column. Generally monolithic columns can be considered as a single large "particle" of porous structure which provides higher performance than conventional particle packed columns[25]. The main advantage of the monolithic column is the elimination of the need

to use end frits to retain the stationary phase. The elimination allows a homogeneous structure of the entire column, rather than exhibiting different properties because of the packing particles and retaining frits which are claimed to be responsible for bubble formation during analysis[26].

Columns of i.d. 2-5 mm are generally used for analytical purposes. Wider columns of i.d. between 10 mm and 25.4 mm can be used for preparative work. Column lengths are normally 5, 10, 15, 25 cm long, if micro particulate stationary phases of 10 μm or less is used. Miniaturization in analytical techniques has become very attractive in recent years from an economic and ecologic point of view. The benefits include reduced chemical consumption (particularly solvent), improved separation, excellent selectivity and small amounts of sample required which is very important in biological analysis. The trends in column format address these new requirements by using packings with smaller particles to obtain more resolution, and sensitivity. Nomenclature for cylindrical column formats is given in Table 1.1, for example columns with internal diameters (i.d.) of 1-2.1 mm are classified as microbore columns. Such columns should, according to van Deemter theory, provide a gain in sensitivity over a conventional 4.6 mm HPLC column that is inversely proportional to the square of the internal diameter[27]. Experiments performed with columns of different dimensions have demonstrated an increase in sensitivity when using 1.0mm i.d. columns compared to 4.6 mm i.d. columns of 17-fold[28] and 16 to 18-fold[29]. For more enhanced sensitivity and ideal MS electrospray ionisation compatibility, the column i.d. was driven down to 0.3 mm, as well as using zero dead volume connections. The use of microbore columns and nanobore columns is well demonstrated nowadays and are widely used for high-efficiency separation of complex mixtures, for example, of peptides from the trypsin digestion of protein(s).

The physical format of columns is also changing. Columns with rectangular, square or other perimeter-shape conduits, i.e. on chips, have appeared and will soon become more widely used.

Table 1.1 Nomenclature of HPLC columns

Column description	Dimension	Typical flow rate
Preparative column	i.d.>10 mm	>20 mL/min
Semi-preparative column	5 mm < i.d. ≤10 mm	5.0-40 mL/min
Normal-bore column	4 mm < i.d. ≤ 5 mm	1.0-10.0 mL/min
Narrow-bore column	2.1 mm < i.d. ≤ 4 mm	0.3-3 mL/min
Microbore column	1 mm ≤ i.d. ≤ 2.1 mm	50-1000 μL/min
Capillary column	100 μm < i.d.< 1 mm	0.4-200 μL/min
Nanobore column	25 μm ≤ i.d. ≤ 100 μm	25-4000 nL/min

1.1.3.2.2 Mobile phase used in reverse phase HPLC

The mobile phase in the reservoir is first degassed to prevent air bubble formation in the pump heads and in the column, which can reduce the efficiency and produce serious noise. Degassing can be performed by purging the solvent with helium or by ultrasonic treatment. A pump is used to deliver solvent systems into the column and there are a number of different types of LC pumps available, the most common of which is the dual head-reciprocating pump. The mobile phases used in reverse phase HPLC generally are a mixture of water or an aqueous buffer and a water miscible solvent (e.g. methanol or acetonitrile). The elution of the analyte can be isocratic or gradient. In isocratic elution the composition of the mobile phase is constant for the whole analysis time, which is

suitable for some detectors that are sensitive to mobile phase changes, and once all analytes have been removed no regeneration of the starting conditions is required prior to the next run. Gradient elution needs to be considered if the chromatographic resolution is not adequate within a reasonable analysis time and separation is unsatisfactory. In a gradient elution, the composition of the mobile phase in the column can be continually changing using two distinct mobile phases which are combined in differing amounts in a mixing chamber before application to the stationary phase.

For RP-HPLC, the proportion of the organic solvent is increased over time, hence increasing the mobile phase strength therefore the hydrophilic analyte is eluted first and followed by the more hydrophobic analytes This technique is likely to provide optimum retention time and selectivity for the mixtures of a broad polarity range, however re-equilibration of the column to the starting conditions, after each run, is required prior to the next separation.

1.1.4 Detectors used in chromatography[30, 31]

Several detectors are available for detection of analytes after chromatographic separation. The most commonly used is the ultraviolet detector which relies upon the presence of an ultraviolet light absorbing chromaphore in the analyte. It is applicable to a large number of compounds, as it can be rather sensitive, has a wide linear range, is relatively unaffected by temperature fluctuations and is also compatible with gradient elution. The diode array detector allows the detection of the light absorbance at a number of wavelengths simultaneously and also allows the production of the absorbance spectra of the eluted components[32]. Other detectors include refraction index

detectors, fluorescence detectors[33], electrochemical detectors[34], radioactivity detectors[35] and more recently on-line LC/mass spectrometry[36] and LC/NMR[37, 38].

1.2 Introduction to mass spectrometry

Mass spectrometry is now a highly sophisticated technique, with computerized instrumentions playing a vital role and has steadily evolved over the last century. A mass spectrometric system is capable of forming, separating and detecting ions according to their mass-to-charge ratio (*m/z*) and the resulting mass spectrum is a plot of the relative abundance of the ions generated as the function of the *m/z* ratio. All mass spectrometers are composed of a sample inlet, ion source, mass analysis, ion detection and data processing components, as shown in Figure 1.6.

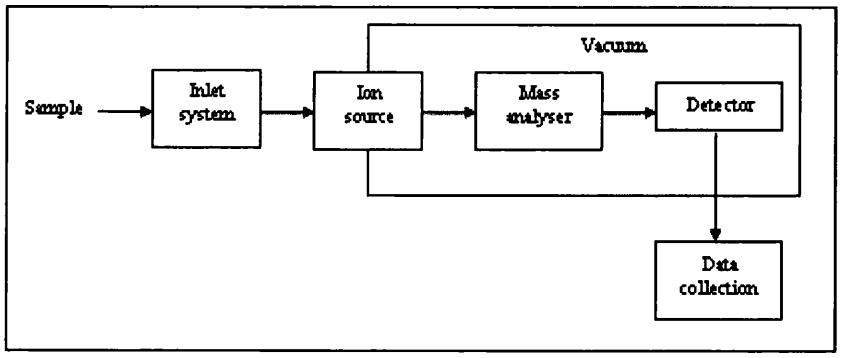


Figure 1.6 Schematic diagram of the components of a mass spectrometer

1.2.1 A brief history of mass spectrometry[39]

The history of MS began in 1897 with the discovery of the electron by Sir J. J. Thomson of the Cavendish Laboratory of the University of Cambridge, who studied electrical discharges in gases. In the first decade of the 20th century, Thomson constructed the

first mass spectrometer (then called a parabola spectrograph) for the determination of mass-to-charge ratios of ions[40]. In this instrument, ions generated in discharge tubes were passed into electric and magnetic fields, which for his design made the ions move through parabolic trajectories. The rays were then detected on a fluorescent screen or photographic plate. Thomson was awarded the 1906 Nobel Prize in Physics "in recognition of the great merits of his theoretical and experimental investigations on the conduction of electricity by gases". In 1918, the British scientist F.W. Aston, designed a mass spectrometer in which ions were dispersed by mass and focused by velocity, it improved the MS resolving power by an order of magnitude over the resolution Thomson had been able to achieve[41]. Aston won the 1922 Nobel Prize in Chemistry for isotope studies carried out with this type of instrument[42]. Around 1920, the American scientist Dempster developed a magnetic deflection instrument[43] with direction focusing and the first electron impact source which ionises volatilized molecules with a beam of electrons from a hot wire filament. This design is still widely used in modern mass spectrometers.

In the 1930s Mattauch and Herzog invented the high-mass resolution double focussing instruments and another design was developed by Nier and Johnson[44]. The magnetic sector was developed and became a commercial instrument in the U.S. by the 1940s and was manufactured by Consolidated Engineering Corporation (Pasadena California). In the early 1940s, the magnetic sector was developed by Neir to conduct isotopic analysis, which resulted in separation of uranium-235 from uranium-238 for the first atomic bomb during World War II[45]. In 1946, the time-of-flight mass spectrometer (TOF-MS) concept was proposed by William E. Stephens[46] and practical TOF instruments were first designed and constructed in the late 1940s and mid -1950s by Cameron and

Eggers[47]. The mass resolution of TOF was greatly improved in 1974 by Mamyrin[48]. Ion cyclotron resonance MS (ICR-MS) was brought to the attention of chemists in 1960s[39] by Baldesschweiler *et al.* In 1974, Marshall *et al.* invented Fourier transform ICR MS (FT-ICR MS) which gives the highest mass resolution of any instrument[49]. At the same time, APCI was discovered by Horning and his group[50] and LC-MS coupling was reported by McLafferty *et al.*[51]

The ESI-MS technique was first conceived in the 1960s by Dole[52], but it was put into practice in the early 1980s by Fenn[53]. MALDI-MS, a form of laser desorption MS, was reported by German scientists Hillenkamp and Karas[54], and independently by the research scientist Tanaka and coworkers at Shimadzu Corp., Kyoto, Japan[55]. Until the early 1980s, biomolecules were usually ionised by electron impact (EI) or chemical ionisation (CI) methods. An essential requirement for these two ionisations is to vaporize the sample. This is of no great concern in the analysis of small organic molecules or those amenable to gas chromatography. However, polar thermally labile samples cannot be analyzed by EI or CI without prior derivatization, for example, by trimethylsilylation (TMS)[56]. Over the last 20 years, biological MS has changed out of all recognition. This is primarily due to the development of "soft ionisation" methods, such as fast-atom bombardment (FAB), which permits the ionisation and vaporization of large, polar, and thermally labile biomolecules[57]. However, FAB whilst being the pioneering technique of the early 1980's is only used for some specialist chemical applications and is not generally used in biological mass spectrometry now. MALDI and ESI are the two main soft ionisation techniques which widely used in the study of large biomolecules nowadays, and considered as complimentary to each other. A brief description of the mass spectrometers used in this thesis is given below.

1.2.2 Introduction to LCQ electrospray ion trap

The LCQ ion trap is a product of the Thermo Electron Company. It is an analytical instrument which has both electrospray and APCI ionisation sources and has multiple MS/MS capacity, and is able to perform both quantitative and qualitative analyses.

1.2.2.1 The electrospray ionisation ion source

1.2.2.1.1 LC/MS Interfacing

Many separation techniques such as gas chromatography (GC), liquid chromatography (LC) or capillary electrophoresis (CE) can be coupled with mass spectrometry and these coupling techniques have been widely used in pharmaceutical, food and biological fields. A liquid chromatography/mass spectrometry (LC/MS) is an instrumental setup in which a high-performance liquid chromatograph (HPLC) and a mass spectrometer (MS) are combined. The LC is the input device for the mass spectrometer, effluent from the LC is introduced into mass spectrometer, and the mass spectrometer is the detector for the LC. An LC/MS interface is required to translate from the high-pressure environment of the LC to the very low-pressure environment in the mass spectrometer. The coupling of liquid chromatography/mass spectrometry provides sensitive, selective, rapid and information-rich analytical methodology in many fields and has become a robust and routine analytical tool.

Direct on-line combination of liquid chromatography to a mass spectrometer was considered for many years as incompatible[58], because the mass spectrometer requires a vacuum and the analytes of interest to be present in the gas phase whereas HPLC provides separation of involatile compounds in a liquid mobile phase. Different

methods were used to tackle these problems[59], e.g. moving belt coupling[60] and the particle beam[61] interface are based on the selective vaporization of the elution solvent before it enters the spectrometer source; continuous-flow fast-atom bombardment[62], relied on reducing the flow of liquid that is introduced into the interface in order to obtain a flow that can be pumped directly by the source. Later a series of HPLC/MS coupling methods such as, electrospray and atmospheric pressure chemical ionisation (APCI) were introduced. Amongst which electrospray is the most widely employed one.

1.2.2.1.2 The electrospray ionisation mechanism

Electrospray ionisation (ESI) is a soft ionisation technique. ESI sources ionise the sample at atmospheric pressure and then transfer the ions into the mass spectrometer. Electrospray is conceptually a more simple method for the ionisation of polar molecules compared to MALDI. Figure 1.7 shows the ESI interface and schematic representation of the electrospray process.

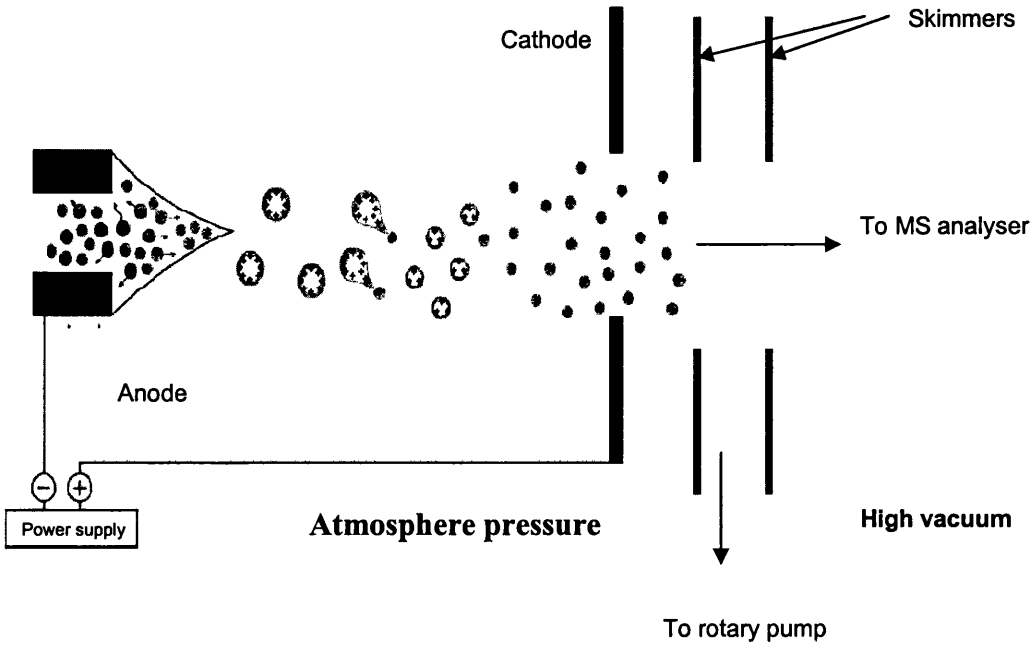


Figure 1.7 ESI interface and schematic representation of the ESI process[57]

In Electrospray, the sample of interest is dissolved in a solvent solution and introduced into the ESI source, either through a HPLC column or by syringe infusion, as a fine mist of droplets, which has an accompanying flow of nitrogen gas surrounding it. Electrospray is produced by applying a strong electric field to the liquid passing through a capillary tube. The electric field is obtained by applying a potential difference of 3-6 kV between this capillary and the counter-electrode, producing electric fields of 10^6 V m^{-1}. The small charged droplets are sprayed from the ES capillary into a bath gas of hot nitrogen (~100 °C) at atmospheric pressure and travel down a pressure and potential gradient towards an orifice in the mass spectrometer high-vacuum system. The electrospray field induces a charge accumulation at the liquid surface located at the end of the capillary, which breaks up to form highly charged droplets when the surface-Coulombic forces overcome surface-tension forces. The process continues until

gas-phase ions are produced from charged droplets in a series of solvent evaporation-Coulomb fission cycles. The gas, most often nitrogen, is injected coaxially at a low flow rate which allows dispersion of the spray to be limited in space so that sample droplets can pass either through a curtain of heated inert gas, or through a heated capillary to remove the solvent and enter the mass analyser. The exact mechanism of ion formation whether it is by ion evaporation or by complete solvent removal from the charged droplet is under debate[63]. The ionised analytes are protonated, e.g. $[M+H]^+$ or in negative mode de-protonated $[M-H]^-$,with the help of some volatile acid or base working as an ion-pairing agent. The ESI sources can tolerate flows up to 0.2 mL/min, which allow convenient coupling to a classical HPLC system. The flow can scale down to about 100 nL/min by using nano-electrospray techniques, which is compatible with HPLC using small-bore or capillary columns.

ESI is a soft ionisation technique and has made significant contributions to modern mass spectrometry. Any polar or ionic compound can be analyzed by ESI. Some heat-labile or high molecular weight compounds which were previous unsuitable for mass analysis can be analysed by ESI. ESI has become particularly useful in the mass analysis of polar analytes, such as, biological polymers (proteins, peptides, glycoproteins and nucleotides), pharmaceuticals and their metabolites[57]. The greatest advantage of ESI compared to MALDI is that it can be easily interfaced to hyphenated technique, such as, HPLC or CE which make the automated analysis of proteins and peptides faster and more reliable.

1.2.2.2 Components of the LCQ ion trap mass spectrometer

In the LCQ system, the ion optics, mass analyser, ion detection system and the atmospheric pressure ionisation stack, are all enclosed in a vacuum manifold, while the ion source is the link between atmospheric pressure and the vacuum system.

1.2.2.2.1 Ion source — Electrospray probe

The electrospray probe is the interface that couples the mass spectrometer to the HPLC and produce ions from the liquid phase. The ESI probe assembly of the LCQ ion trap is shown in Figure 1.8. The liquid sample from a syringe or a HPLC column is introduced into the ESI probe through the sample inlet. The sheath gas (nitrogen) nebulises the sample solution into a fine mist when the sample solution leaves the ESI nozzle. A high voltage is applied to the needle, the electric field near at the capillary tip is very high, and the solution leaves the electrospray needle (so called Taylor cone) as a fine mist of charged droplets which are drawn into the vacuum region of the mass spectrometer. The outer coaxial nitrogen auxiliary gas introduced through the ESI probe inlet assists in the nebulisation and evaporation of the sample solution.

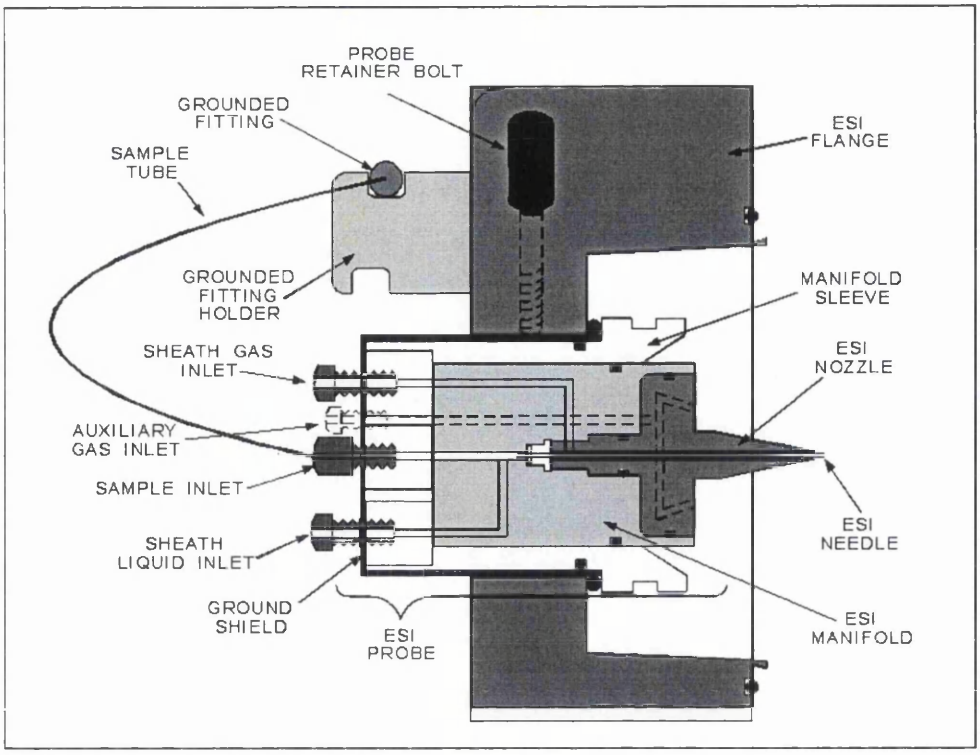


Figure 1.8 The ESI probe assembly of the LCQ ion trap[64]

1.2.2.2.2 Mass analysis — ion trap mass analyser

The ion trap mass analyser consists of three electrodes: a circular electrode, with two hyperbolic endcaps on the top and the bottom, the ring electrode is positioned symmetrically between the two-endcap electrodes as shown in Figure 1.9. These electrodes form a cavity and have functions including ion storage, ion isolation, collision-induced dissociation (CID for MS^n) and sequential mass ejection processes.

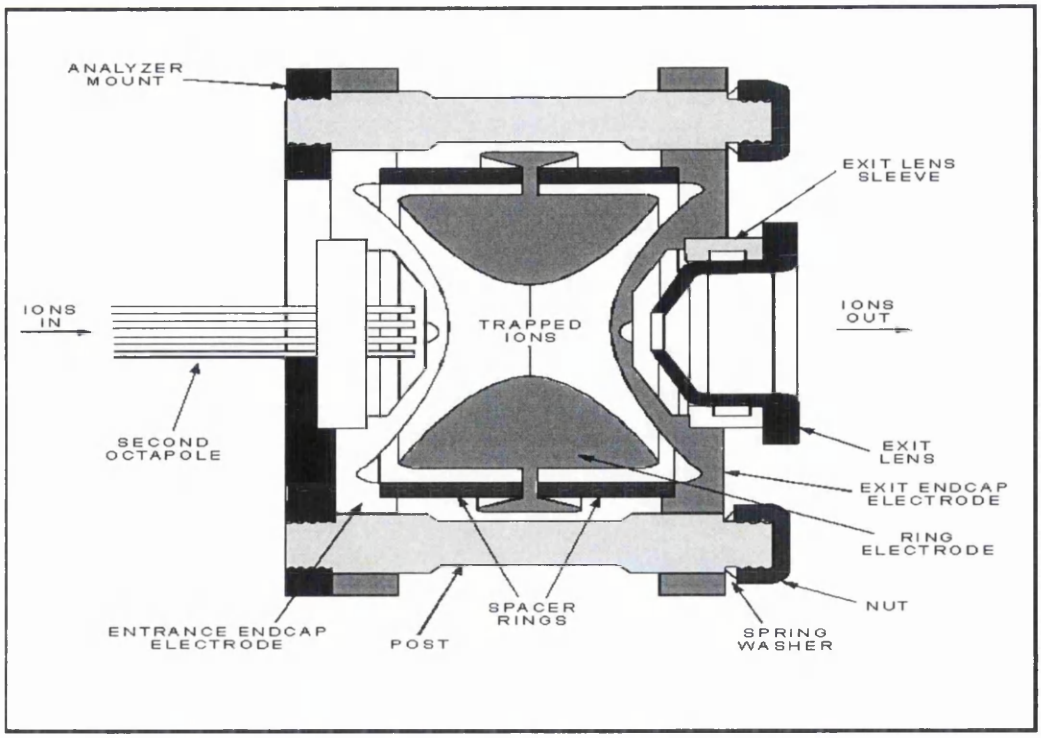


Figure 1.9 Cross sectional view of the LCQ ion trap mass analyser[64]

First, a DC offset voltage is applied to the mass analyser electrode to draw ions in from the inlet tube. An RF voltage is applied to the central ring electrode to create a three-dimensional rotationally-symmetrical quadrupole electric field. At low amplitude ions above a minimum of *m/z* can be stored in this field in stable orbits and as the RF voltage is increased, the trajectories of the ions become unstable in order of increasing *m/z* ratios. Ions with unstable orbits are ejected from the mass analyser towards the ion detection system. A huge advance in ion trap technology was the introduction of a buffer gas, such as, helium gas inside the mass analyser cavity which slowed ions down so that they become trapped by the potential well created by the RF field being applied. The helium gas also serves to dampen the amplitude of the oscillations of the trapped ions, thereby focusing them into the centre of the cavity. Sensitivity and resolution are

both enhanced due to these interactions with helium gas in the trap at pressures ~10^{-3} mbar.

An ion trap has a unique capability, which differentiates from all other mass spectrometers, in that it stores ions. Once an ion is stored, it can be manipulated in many different ways to perform, for example, multistage MS/MS experiments (MS^n). For other mass spectrometric techniques this would require additional and large analyser hardware (in tandem). An ion trap is especially sensitive and is frequently used for structural elucidation through MS/MS and MS^n experiments. For this purpose, the He bath gas in the trap is used for collision induced dissociation (CID). First, the ions of interest are isolated by eliminating all other *m/z* species from the trap. These mass selected ions are resonantly excited to larger orbits by applying a high amplitude voltage to the alternating dipolar field at the end caps and subsequently inducing collisions with the surrounding helium atoms. Compared with a triple quadrupoles, the ion trap is a physically small and a less expensive instrument, which nevertheless is highly sensitive and possesses MS^n capacities. However, it is not best suited for quantitative analysis due to its small trapping volume, and limited capacity for ion storage. Overfilling of the ion trap will cause the deterioration of the mass spectrum and loss of dynamic response range due to space charging. To avoid these effects, the number of ions introduced into the trap should be controlled carefully. Trapping of ions can also be performed in a linear ion trap device. Recently, the linear ion trap has been combined with quadrupole, time-of-flight and Fourier-transform ion cyclotron resonance mass spectrometry. A linear ion trap has two major advantages over the quadrupole ion trap, a larger ion storage capacity and a higher trapping efficiency[65], both greatly improve its performance.

1.2.2.2.3 Ion detection

The ion beam passes through the mass analyser and is transformed into a usable signal by a detector. In the LCQ system an electron multiplier detector which is located behind the mass analyser, ions emerging from the ion trap hit a conversion dynode, as shown in Figure 1.10, which improves sensitivity and performance in negative ion mode. The electron multiplier is the most commonly used detector in mass spectrometers because its low cost and good linear dynamic response range (~10^7). It is capable of producing a high signal-to-noise ratio, and it also permits voltage polarity interchange between the positive ion and negative ion mode of manipulation.

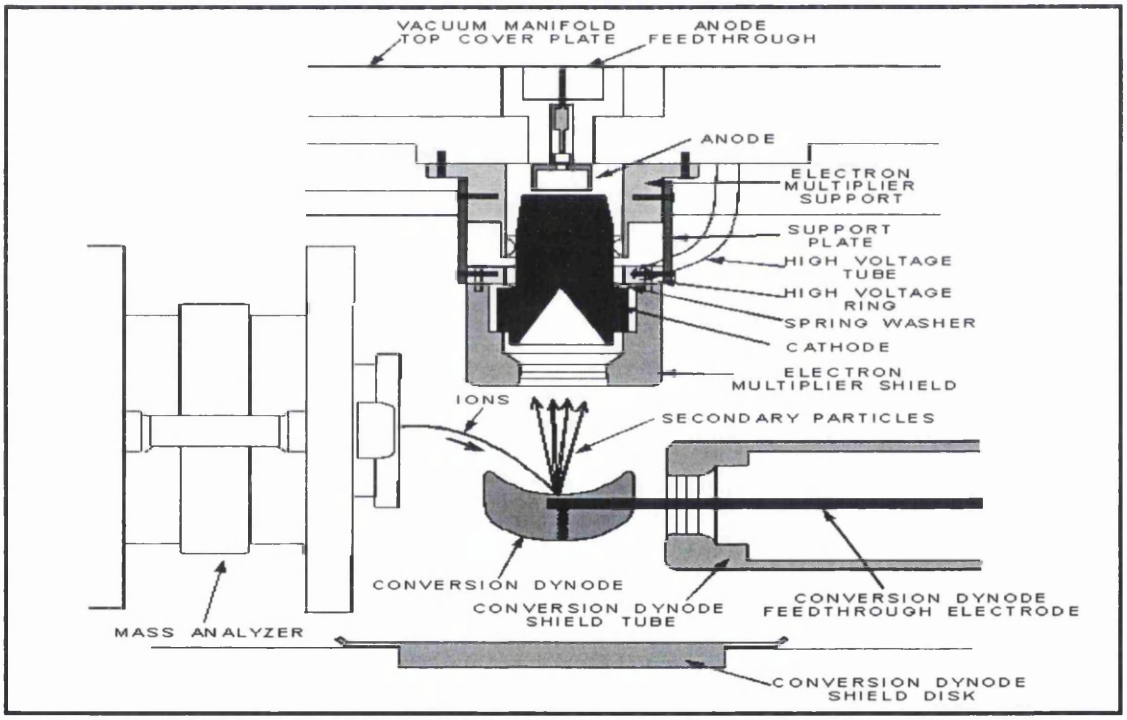


Figure 1.10 Cross sectional view of the LCQ ion detection system[64]

When a positive or negative ion reaches the conversion dynode, it causes the emission of several secondary particles include photons, positive ions, negative ions, electrons

and neutrals. When positive ions strike the negative high-voltage conversion dynode, the conversion dynode converts ions from the mass analyser into secondary particles (negative ions and electrons). When negative ions strike the positive high-voltage conversion dynode, the secondary particles of interest are positive ions or photons. The number of secondary particles is greater than that of ions striking the conversion dynode and the signal is amplified. These secondary particles are accelerated into the electron multiplier. They strike the cathode with sufficient energy to dislodge electrons as they collide with its curving inner wall. These electrons pass further into the electron multiplier, again striking the walls, causing the emission of more and more electrons as they travel towards the ground potential. Thus a cascade of electrons is created that finally results in a measurable current at the end of the electron multiplier. The current is typically amplified to the order of $10^5 - 10^6$ for a singly charged ion hitting the entrance of the electrode.

1.2.3 Introduction to matrix-assisted laser desorption ionisation time-of-flight (MALDI-TOF) mass spectrometry

MALDI-TOF is a technique offering fast and accurate determination of a number of polymer characteristics. It is now widely used in proteomics research.

1.2.3.1 Description of the MALDI source and the MALDI mechanism

MALDI is the acronym for matrix-assisted laser desorption/ionisation. MALDI's development came out of earlier successes in applying laser desorption ionisation (LDI) to biomolecules as well as plasma desorption and fast-atom bombardment (FAB). During the 11^{th} International MS Conference in Bordeaux in 1988, the MALDI

technique was described as suitable for proteins with molecular masses exceeding 10 kDa and attracted the attention of the MS community[54]. Simultaneously Japanese scientists demonstrated the desorption of protein ions with masses in excess of 60,000 Da[55]. Since then, MALDI has become an important part in the field of protein chemistry. It is mostly used in peptide mass mapping, where proteins separated by polyacrylamide gel electrophoresis (PAGE) are digested by a suitable enzyme and the resulting peptides are mass analyzed[66].

Compared with electrospray ionisation, MALDI has the following merits of its own:

a. It is more compatible with buffers, normally used in biological assays, reducing the need for sample cleanup.
b. It can analyze mixtures and different classes of biopolymers, including peptides, oligonucleotides, glycoconjugates and synthetic polymers. Whereas ESI generates multiply charged ions, especially as the molecular weight increases.
c. It tends to give abundant $[M+H]^+$ or $[M-H]^-$ ions which gives an immediate picture of the molecular species present. It is suitable for biopolymers with very high mass, the *m/z* ratio can be as high as 200,000.
d. It provides very high sensitivity and often requires only femtomoles of sample.
e. It can be automated with high throughput and simplicity of operation.

Despite of its widespread application, the fundamental processes of ion generation and desorption are still not well understood and still the subject of active research[67]. Some of the basic conditions which rule desorption are understood and molecular modelling is beginning to reveal some important aspects of the processes even though dynamic modelling of realistic volumes of μm^3 in size and with time scales of tens of

nanoseconds are still beyond the capacity of even very large computers[68]. Ionisation is even more of a challenge; Knochnemuss and Zenobi suggested a mechanism that the matrix primary ionisation is followed by matrix ion-analyte molecule reactions which lead to analyte ions[69]. Karas *et al*, on the basis of thermodynamics, proposed a new model for the underlying ionisation processes in MALDI which explains the observation that UV/MALDI spectra are dominated by singly charged ions[70]. There is strong experimental evidence that the matrix is essential to dilute and isolate analyte molecules from each other by the formation of a solid solution upon evaporation of the solvent and to function as a mediator for energy absorption[71]. The matrix is present in vast excess of the analyte and when mixed thoroughly with analyte isolates individual analyte molecules and promotes adequate matrix-analyte collisions or interactions. The matrix serves to absorb energy from the pulsed laser beam, and provides photoexcited acid or basic sites for ionisation of analyte molecules in ion/molecule collision[72]. MALDI is a relatively simple ionisation method where a sample (and suitable solvent) and matrix (and suitable solvent) are mixed thoroughly and dried on a sample plate. A UV laser is fired on the plate, the matrix absorbs energy and transfers it to sample so as to produce gas phase sample neutrals and, to a lesser extent ions, as illustrated in Figure 1.11[73]. Sample ions are generally formed by ion/molecule reactions in the gas-phase MALDI plume, by protonation and also processes, such as, sodiation are commonly observed. Nitrogen lasers are a common choice to illuminate the sample and have a principal wavelength in the UV range (337 nm) on commercial MALDI mass spectrometers. UV-absorbing aromatic compounds are most commonly used as matrices[74], and various aromatic acids provide excellent sensitivity for forming protonated ions.

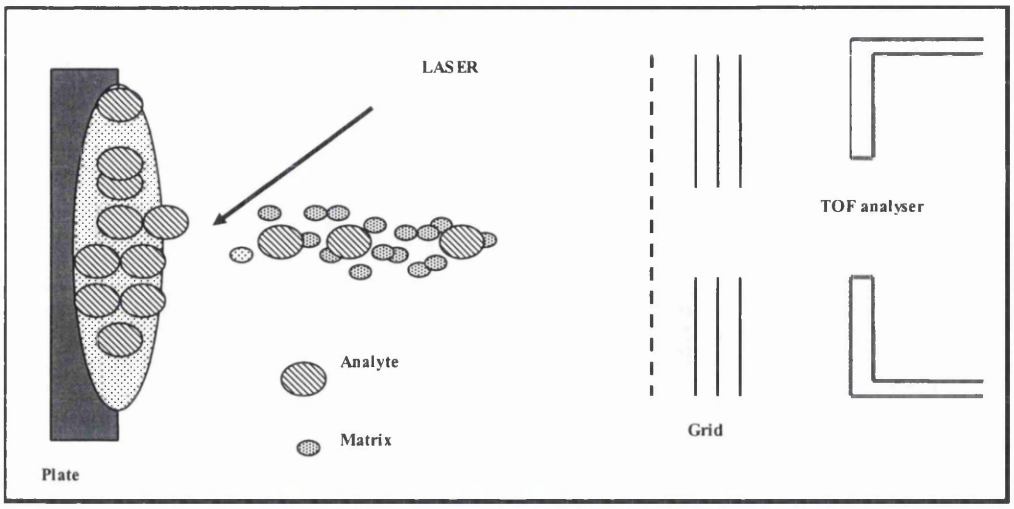


Figure 1.11 Illustration of a MALDI ion source

1.2.3.2 Time-of-flight (TOF) analyser

The concept of separating ions of different charge-to-mass ratios via time-of-flight (TOF-MS) were originally proposed in 1948[47]. However, in its first realization, all the key elements, such as pulsed ion source, electronics and detector, were so imperfect that the original instruments were of little practical use. It could hardly separate ions of very low masses. The first TOF mass spectrometers of any practical interest were proposed in the 1950s, and were linear time-of-flight mass analysers. The mass resolution ($m/\Delta m$) was 100-300 at a partial pressure of about 10^{-7} Torr in the ion source. At that time time-of-flight mass spectrometers were considered as low resolution instruments. The resolution of simple TOF instruments is restricted due to several practical and technical reasons, such as uncertainties in the time of ion formation, its initial location in the extraction field, the kinetic energy of the analyte ions and metastable fragmentation. A serious effect being due to the energy spread of ions created in the pulsed source which

considerably limits the mass resolution. Because of these limitations, a pause occurred in TOF development and application from 1960s to 1980s. In the 1970s, important progress was achieved when the electrostatic ion reflectron was introduced by Mamyrin *et al*[75]. The introduction of the reflectron enables the compensation for differences of times-of-flight in the field-free drift regions affected by the energy spread in the source. Mass resolutions up to 6000 (FWHM) have been reported for peptides up to about 3000 Da with reflectron TOF-MS[76] and resolution of up to 30,000(FWHM) can be achieved on lower mass ions.

Because time-of-flight requires the ions to be produced in bunches, it is especially suited for pulsed laser sources. Since the introduction of matrix-assisted laser desorption ionisation (MALDI) by Karas and Hillenkamp[54], time-of-flight has evolved rapidly as a routine analytical mass spectrometer for the analysis of biomolecules. The simplest time-of-flight mass analyser consists of a short source-extraction region, usually of a few centimeters, a drift region typically 0.5 to 4 m in length, and a detector. A high voltage is applied on the source (typically 6-30 kV) to accelerate ions to constant energy. The drift region is field free, ions move across this region with velocities that are inversely proportional to the square root of their mass. Ion flight times generally fall between 10 to 200 μs, and can be recorded by a digital oscilloscope to produce a mass spectrum.

Figure 1.12 shows the schematic diagram of MALDI with a reflectron-TOF tube. A sample, co-crystallized with the matrix, is irradiated by a laser beam, leading to sublimation and ionisation of analytes. About 100–500 ns after the laser pulse, a strong acceleration field is switched on (delayed extraction), which imparts a fixed kinetic

energy to the ions produced by the MALDI process. These ions travel down a flight tube and are turned around in an ion mirror, or reflector, to correct for initial energy differences. The ions can be detected sequentially according to their masses; lighter ions have higher velocities and arrive at the detector earlier than heavier ions.

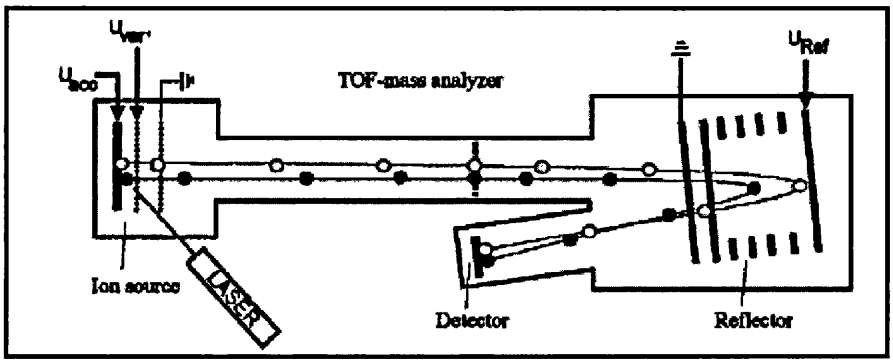


Figure 1.12 Schematic diagram of MALDI-TOF with a reflectron tube[77]

The analytical advantages of TOF include its relatively low cost, high sensitivity, large mass range and the ability to record a complete spectrum in a single acquisition.

1.2.4 Quadrupole time-of-flight (Q-TOF) mass spectrometer

1.2.4.1 Description of a quadrupole mass analyser

The quadrupole mass analyser is a device that uses the stability of the ion trajectories in oscillating electric fields to separate ions according to their *m/z* ratio. The analysers were first described in 1953 by Paul and Steinwegen, after continuing the work of Chrisophilos on strong focussing of ions. It is made up of four parallel rods of hyperbolic section. Ions travelling between the rods are subjected to the influence of a radio frequency (RF) electric field, the trajectory of an ion will be stable if it does not

strike the rods. The fundamental theory of a quadrupole mass filter is very complex and will not be dealt with here. However, there are two types of voltages applied to the rods: 1) a radio frequency voltage (V) of large amplitude (several kV) which entraps ions allowing their transmission and 2) a DC voltage, U, (~10 V) applied to the rods to enable mass selection. Whilst quadrupole theory is complex the ratio of U to V is an important parameter which is set by the computer so as to ensure unit mass resolution whilst scanning an appropriate mass range.

Due to the relative simplicity of the quadrupole design this device has been combined, in tandem, with like quadrupoles to achieve MS/MS. A triple quadrupole is composed of two quadrupole mass filters and a RF-only quadrupole, it is usually symbolized as Q-q-Q. The middle RF-only quadrupole has no DC voltage (i.e. U=0) and the mass resolution capacity falls away allowing all ions, of any mass to have a stable trajectory. Normally a collision gas is introduced at a pressure so that a mass selected ion from the previous quadrupole can undergo many collisions inducing the ion to fragment. The first and the third quadrupoles are mass analysers which can be either set to scan a mass range or set to transmit any single mass. This huge flexibility allows a range of "scanning modes" to be undertaken with this versatile device. Figure 1.13 displays the main scanning capabilities of a triple quadrupole, the most important of the MS/MS scans are: product ion scan, precursor ion scan and neutral loss scan. The triple quadrupole has an excellent reputation for quantitative analysis and is widely used in the pharmaceutical industry.

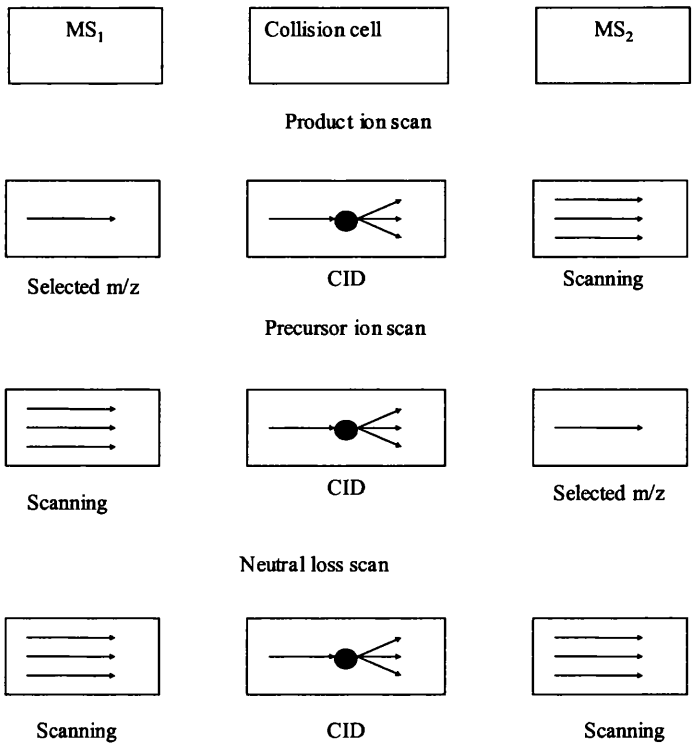


Figure 1.13 Illustration of MS/MS scanning modes by a triple quadrupole mass spectrometer.

1.2.4.2 Quadrupole time-of-flight (Q-TOF) mass spectrometry

Quadrupole time-of-flight (Q-TOF) mass spectrometers have only been commercially available in the last decade. They are now widely accepted and utilised by research groups for their ability to provide high sensitivity and accuracy in molecular weight determination, molecular structure information and quantitative analysis. For example, the Micromass Q-TOF Ultima used in this study is a hybrid quadrupole time-of-flight mass spectrometer and was the instrument used for the mass accuracy studies. The ion source of Q-TOF is an electrospray ion source, sample introduction is through an infusion pump, loop injection or from an HPLC column. The ESI interface, by using the

orthogonal design, which generates ions in a Z-spray source, has proved very successful in handling involatile buffers.

As introduced previously, quadrupole analysers are generally made up of four parallel hyperbolic rods. If a positive ion enters the space between the rods, it will be drawn toward a negative rod. If the potential changes polarity before the ion neutralises itself on the rod, the ion will change direction. Therefore the ion trajectories are controlled by a set of time dependent forces that can be generated by applying direct current and radiofrequency potentials to a set of electrodes. In a Q-TOF, the ions leaving the quadrupole analyser will flow into the orthogonal time-of-flight analyser. A collision cell and orthogonal acceleration (OA) pusher is located between the quadrupole and the TOF analyser. The collision cell serves to induce fragmentation in MS/MS experiments with the orthogonal acceleration pusher serving to either push the ions downward into the TOF analyser or to drift into the post acceleration photomultiplier detector which can record mass spectra. The OA pusher enables all ions to enter the flight tube at precisely the same time in the Q-TOF. The ions pushed out of the OA can travel either in a "V" or "W" path down the TOF tube to the reflectron lens at the end of the instrument, the reflectron then reflects ions back to the detector.

In Q-TOF, the quadrupole is operated as an ion guide in MS mode and as a mass selection device in MS/MS mode. A reflectron time-of-flight (TOF) analyser is placed orthogonally to the quadrupole and serves as a mass resolving device for both MS and MS/MS modes. In the TOF analyser ions are separated on the basis of their velocity differences. The lighter ions travel faster and reach the detector earlier than the heavier

ones. The final detector is a microchannel plate (MCP) in which an ion counting system is employed. The schematic diagram of Q-TOF is shown in Figure 1.14.

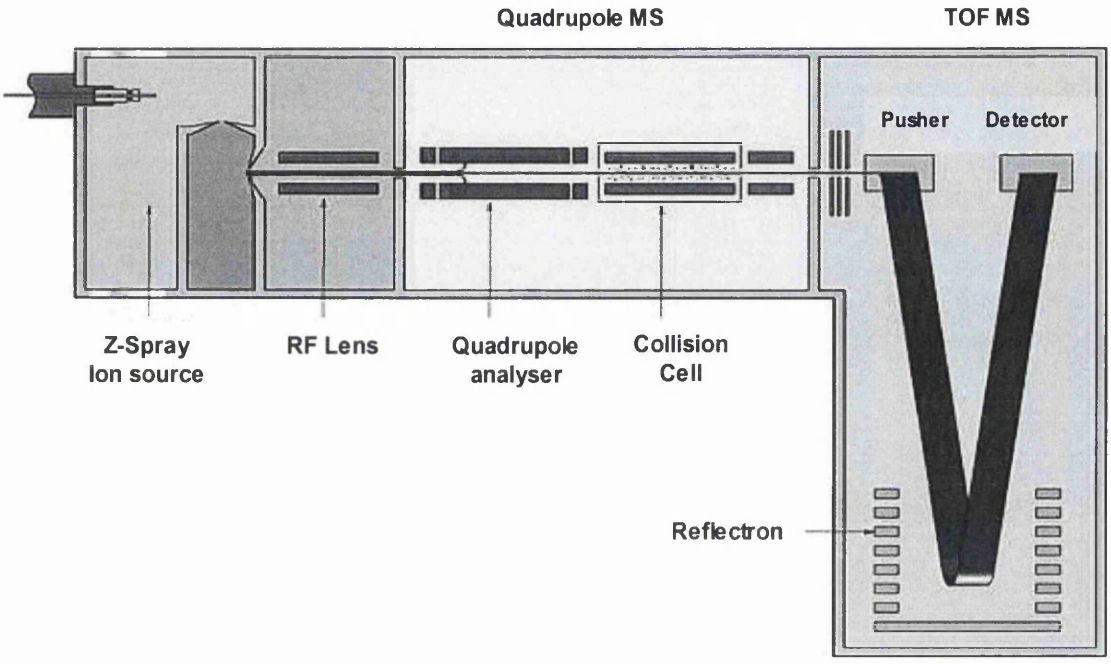


Figure 1.14 Schematic diagram of the Q-TOF mass spectrometer[78]

A Q-TOF can achieve high mass accuracy and is often used to obtain accurate mass of compounds. The use of "lockspray" allows the acquisition of exact mass data by providing the capability to introduce the reference compound (known lock mass), which is then used to correct for drifts in mass scale, through a different sample line from that of the analyte. Figure 1.15 shows the design of lockspray in Q-TOF. An oscillating baffle will switch between two sample lines at certain time intervals, thus allowing analysis of each spray independently, and avoids problems such as suppression of the analyte by the reference or vice versa, mass interference between the analyte and reference and clustering of the analyte and reference.

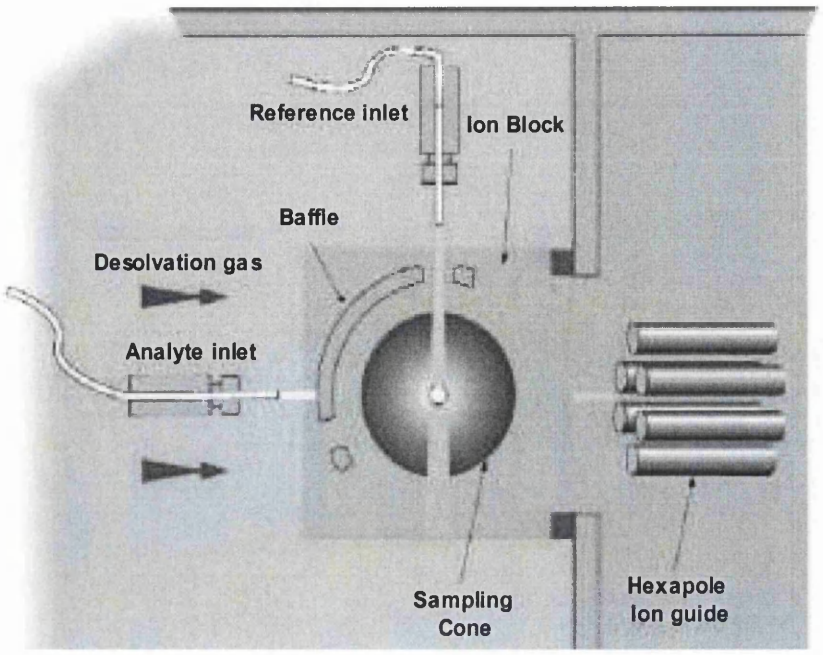


Figure 1.15 Lockspray schematic and the dual electrospray source including an infused internal lockmass from a second sprayer [78]

1.2.5 A brief review of other analyser techniques

There are several mass analysers available, such as time-of-flight, quadrupole, ion trap, Fourier transform ion cyclotron resonance. They work on different theoretical principles and are often suited to different applications. The comparison of different mass analyser is shown in Table 1.2, an appropriate mass analyser needs to be chosen to achieve good mass spectrometric result.

Fourier transform ion cyclotron resonance (FTICR) mass analyser was first used with internal ion sources and developed rapidly when external ion sources are combined. Today the most popular external sources are ESI and MALDI. FTICR is unsurpassed in performance amongst all other mass spectrometers in terms of resolution, mass accuracy and possibly MS/MS capabilities. This performance relies on a super conducting

magnet, which is inherently very stable to $1:10^8$. However, this has also limited its uptake due to its expense and operation.

Table 1.2 Comparison of different mass analysers

Mass analyser	*m/z* limit	Resolution	Mass accuracy (ppm)
Time-of-flight	Unlimited[a]	Low	5-10
Ion trap	< 3000[b]	Low	~100
Quadrupole	< 4000	Low	~100
Magnetic sector	$< 10^4$	High	2-5
FTICR	$< 10^5$	Very high	Typically 1-2, can be as low as 0.1

[a] Depends on detector

[b] Depends on the size of trap and the voltage which can be applied

ICR refers to ion cyclotron resonance. Its underlying characteristic is that ions can be trapped in a circular trajectory around a magnetic field. The analyser operates at high vacuum, 10^{-9} Torr, where little or no collisions occur when ions undergo cyclotron motion. The cyclotron frequency of the ion is proportional to the mass of ion which can then be measured to very high accuracy.

Magnetic sector is another mass analyser widely used. In a magnetic field, ions of different mass follow different circular trajectories. By placing slits to select and collect ions the magnetic analyser can select ions according to their *m/z* ratio. An electronic detector, such as, an electronic multiplier is placed behind the collector slit to directly measure the ion beam. However, the ion beam has an energy dispersion and angular dispersion, which causes the loss of resolution. Double-focusing mass spectrometers

address these problems by using a combination of a magnetic and electrical field to focus and sort ions. Therefore an ion beam of a given *m/z* is brought to a focus even when the ion beam is initially diverging and containing ions of different energy. Double focusing provides accurate mass measurement and can achieve excellent measurement of ion abundances over a wide dynamic range 10^7 (7 orders). It is used to determine elemental compositions of ions in mass spectra which lead to compound identifications. But the relatively high price and complexity limits its popularity even though its performance for many applications is superb.

The orbitrap mass analyser is a wholly new mass analyser invented by Makarov[79]. It can be considered as a modified form of quadrupole ion trap, although the orbitrap uses electrostatic field while the quadrupole ion trap uses a dynamic electric field, typically oscillating at ~1 MHz. The orbitrap can provide high mass resolution, high mass accuracy, and good dynamic range and has been applied to a wide range of analytes.

1.2.6 Scan modes in mass spectrometry

Mass spectrometers can be operated in different scan modes, they are chosen according to the analysis to be undertaken. Several different scan mode can be chosen for one analysis and computer control MS can automate such procedures. Full scan mode is a scan mode in which the mass analyser records all the ions within a certain mass range set in advance. Generally, full scan experiments are used to determine the molecular weights of unknown compounds or of each component in a mixture of unknown compounds. Full scan mode provides more information about an analyte than does selected ion monitoring (SIM, see below). Full scan mode cannot provide high

sensitivity due to the long scan time over a range of masses and limited time sampling each mass. Selected ion monitoring (SIM) scan is a common method to monitor ions of selected *m/z* ratio. This mode provides much higher sensitivity than full scan mode and is widely used for quantitation of known compounds.

Tandem mass spectrometry is any general method involving at least two stages of mass analysis. The first stage is to isolate a precursor ion which further undergoes a second stage or multistage to yield product ions and neutral fragments. Tandem mass spectrometry can provide more structure information and is very useful for structural elucidation of unknown components. Selected reaction monitoring (SRM) scan mode is a two-stage tandem mass spectrometry experiment in which a particular reaction or set of reactions, such as the fragmentation of an ion or the loss of a neutral moiety, is monitored. In the SRM scan mode, the precursor ions are excited so that they collide with collision gas. The collisions of the parent ions cause them to fragment to produce one or more product ions (by mass analyser CID for the ion trap or collision cell CID for the Q-TOF). Then, ions of interest for one or more mass-to-charge ratios are selected to produce SRM signal for each SRM process. Selected reaction monitoring allows for the very rapid analysis of trace components in complex mixtures. SRM mode offers high specificity since a limited number of product ions is monitored and thus the background noise (and chemical noise) is greatly reduced.

Pharmaceutical and biological sciences are currently the fields that LC/MS is most widely employed as an analytical technique. In the following chapters, LC/MS was applied in both these fields from an analytical point of view. One study is the quantitation of active components in *Ginkgo biloba* nutritional supplements and urine

samples after ingestion of *Ginkgo biloba*. In another study, immobilized metal ion affinity chromatography LC/MS was applied to the identification of phosphoproteins in mouse brain samples, with the purpose of identifying potential substrate proteins of cyclic CMP. For LC, normal-bore column (4.6 mm i.d.), capillary column (300 μm i.d.) and nano column (75 μm i.d.) were exploited. For mass spectrometry, an ion trap and QTOF were used in the *Ginkgo biloba* study, while an ion trap and MALDI-TOF were utilized in the proteomics study. The mass accuracy at high molecular weight in MALDI-TOF was also investigated in this thesis. The phenomenon of peak broadening observed for high mass protein peaks was explained, by a theoretical study, on the basis of ion fragmentation.

1.3 References

1. IUPAC, Pure Appl. Chem., **37,** 147 (1974)
2. Tswett, M.S., Proc. Warsaw Soc. Nat. Sci.(Biol.), **14**, No. 6 (1903)
3. Kuhn, R. and Lederer, E., Physiol. Chem., **141,** 197 (1931)
4. Martin, A.J.P. and Synge, R.L.M., J. Biochem., **35,** 1358 (1941)
5. Gonsden, R., Gordon, A.H. and Martin, A.J.P., J. Biochem., **38,** 224 (1944)
6. James, A.T. and Martin, J.P., J. Biochem., **50,** 679 (1952)
7. Giddings, J.C., *Dynamics of Chromatography Part I. Principle and Theory*, New York (1965)
8. Kirkland, J.J., J. Chromatogr. Sci., **7,** 7 (1969)
9. Huber, J.F.K., J. Chromatogr. Sci., **7,** 85 (1969)
10. Craig, L.C., Anal. Chem., **22,** 1346 (1950)
11. Gluekauf, E., Trans. Faraday Soc., **22,** 1346 (1955)
12. van Deemter, J., Zuiderweg, F.J. and Klinkenberg, A, Chem. Eng. Sci., **5**, 271 (1956)
13. Giddings, J.C., J. Chem. Phys., **31,** 1462 (1959)
14. Dudley, E., PhD Thesis, (2001)
15. Niessen, W.M.A., *Liquid Chromatography-Mass Spectrometry*, Marcel Dekker, Inc., (1999)
16. Porath, J. and Flodin, P., Nat., **4676**, 1657 (1959)
17. *Gel Filtration Handbook: Principles and Methods*, Amersham Biosciences
18. Mikes, O., in *Separation Methods*, (Ed. Dely, Z., Published by Elsevier) (1984)
19. *Ion Exchange Handbook: Principles and Methods*, Amersham Biosciences
20. Turkova, J. and Seifertova, A., J. Chromtogr., **148**, 293 (1978)

21. Bouriotis, V. and Dean, D.G., J. Chromatogr., **206**, 521 (1981)

22. Filka, K., Coupek, J. and Kocourek, J., Biochem. Biophys. Acta, **539**, 518 (1978)

23. Keski-Oja, J. and Yamada, M., Biochem., **193,** 615 (1981)

24. Meyer, V.R., *Practical High-Performance Liquid Chromatography*, 2nd Edition, John Wiley & Sons, (1998)

25. Klodzinska, E., Moravcova, D., Jandera, P. and Buszewski, J. Chromatogr. A, **1109**, 51 (2006)

26. Hayes, J.D. and Malik, A., Anal. Chem., **72,** 4090 (2000)

27. Oudhoff, K.A., Sangster, T., Thomas, E. and Wilson, I.D., J. Chromatogr. B, **832,** 191 (2006)

28. Mallett, D.N. and Law, B., J. Pharm. Biomed. Anal., **9**, 53 (1991)

29. Liu, J., Volk, K.J., Mata, M.J., Kerns, E.H. and Lee, M.S., J. Pharm. Biomed. Anal., **15,** 1729 (1997)

30. Scott, R.P.W., *Liquid Chromatographic Detectors*, Elsevier, (1986)

31. Yeung, S. and Synovec, R.E., Anal. Chem., **58,** 1237A (1986)

32. Huber, L., *Diode Array Detection in HPLC*, Marcel Dekker, (1993)

33. Weinberger, R. and Sapp, E., Int. Lab, **14**, 80 (1984)

34. Horvai, G. and Pungor, E., Crit. Rev. Anal. Chem., **21,** 1 (1989)

35. Veltkamp, A.C., Das, H.A., Frei, R.W. and Th Brinkman, U.A., Eur. Chromatogr. News, **1,** 16 (1987)

36. Slobodnik, J., van Baar, B.L.M. and Th Brinkman, U.A., J. Chromatogr. A, **703,** 81 (1995)

37. Laude, D.A. and Wilkins, C.L., Anal. Chem., **56,** 2471 (1984)

38. Albert, K., J. Chromatogr. A, **703,** 123 (1995)

39. Borman, S., Dagani, R. and Rawls, R. L., Chem. Eng. News, **76**, 39 (1998)

40. Thomson, J. J., *Rays of Positive Electricity and Their Application to Chemical Analysis*, Longmans, Green and Co. Ltd., (1913)

41. Aston, F.W., Philoss. Mag., **38**, 707 (1919)

42. Aston, F.W., *Mass Spectra and Isotopes*, Edward and Arnold and Co., (1942)

43. Dampster, A. J., Phys. Rev., **11**, 316 (1918)

44. Johnson, E.G. and Nier, A.O., Phys. Rev., **91**, 12 (1953)

45. Nier, A.O., Rev. Sci. Instrum., **11**, 252 (1940)

46. Stephens, W.E., Phys. Rev., **69**, 691 (1946)

47. Cameron A.E. and Eggers, D.F., Rev. Sci. Instrum., **19**, 605 (1948)

48. Karataev, V.I., Mamyrin, B.A. and Smikk, D.V., Sov. Phys.-tech. Phys., **16**, 1177 (1972)

49. Comisarow, M.B. and Marshall, A.G., Chem. Phys. Lett., **25**, 282 (1974)

50. Horning, E.C., Carroll, D.I., Dzidic, I., Haegele, K.D., Horning, M.D. and Stillwell, R.N., J. Chromatogr. Sci., **412**, 725 (1974)

51. Apino, P.J., Baldwin,M.A. and McLafferty, F.W., Biomed. Mass Spectrom., **1**, 80 (1974)

52. Dole, M., March, L.L., Hines, R.L., Mobley, R.C., Ferguson, L.D. and Alice, M.B., J. Chem. Phys., **49**, 2240 (1968)

53. Yamashita, M. and Fenn, J.B., J. Phys. Chem., **88**, 4451 (1984)

54. Karas,M. and Hillenkamp, F., Anal. Chem., **60**, 2299 (1988)

55. Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y. and Yoshida T., Rapid Commun. Mass Spectrom., **8**, 151 (1988)

56. Cremese, M., Wu, A.H., Cassella, G., O'Connor, E., Rymut, K. and Hill, D.W., J. Forensic Sci., **43**, 1220 (1998)

57. Griffiths, W.J., Jonsson, A.P., Liu, S., Rai, D.K. and Wang, Y, Biochem. J., **355**, 545 (2001)

58. Mermet, J.M., Otto, M. and Widmer, H.M. in *Analytical Chemistry*, (Ed. Kellner, R.,) Wiley VCH, (1998)

59. Ardrey, B., *Liquid Chromatography/Mass Spectrometry*, VCH, New York, (1998)

60. McFadden, W.H., Schwartz, H.L. and Evans, S., J. Chromatogr. A, **122**, 389 (1976)

61. Willoughby, R.C. and Browner, R.F., Anal. Chem., **56**, 2625 (1984)

62. Barber, M., Bordoli, R.S., Sedgwick, R.D. and Tyler, A.N., J. Chem. Soc. Chem. Commun., **15**, 325 (1981)

63. Kebarle, J., J. Mass Spectrom., **35**, 804 (2000)

64. LCQ manual, Getting Started, Finnigan

65. Hopfgartner, G., Varesio, E., Tschappat, V., Grivet, C., Bourgogne, E. and Leuthold, L.A., J. Mass Spectrom., **39**, 845 (2004)

66. Pappin, D.J.C. and Bleasby, A.J., Curr. Biol., **3**, 327 (1993)

67. Krüger, R. and Karas, M., Chem. Rev., **103**, 427 (2003)

68. Zhigilei L.V. and Garrison, B.J., J. Appl. Phys. A, **88**, 1281 (2000)

69. Knochenmuss, R., Stortelder, A., Breuker, K. and Zenobi, R., J. Mass Spectrom., **35**, 1237 (2000)

70. Karas, M., Gluckmann, M. and Schafer, J., J. Mass Spectrom., **35**, 1 (2000)

71. Dreisewerd, K., Schurenberg, M., Karas, M. and Hillenkamp, F., Int. J. Mass Spectrom. Ion Processes, **154**, 171 (1996)

72. Hillenkamp, F., Karas, M., Beavis, R.C. and Chait, B.T., Anal. Chem., **63**, 1193A (1991)

73. Cotter, R.J., *Time-of-flight Mass Spectrometry: Instrumentation and Applications to Biological Research*, American Chemical Society, Washington, DC, (1997)

74. Fitzgerald, M.C., Parr, G.R. and Smith, L., Anal. Chem., **65**, 3204 (1993)

75. Mamyrin, B.A., Int. J. Mass Spectrom. Ion Processes, **131**, 1 (1994)

76. Ingendoh, A., Karas, M., Hillenkamp, F. and Giessmann U., Int. J. Mass Spectrom. Ion Processes, **131**, 345 (1994)

77. Mann, M. and Hendrickson, R.C., Annu. Rev. Biochem., **70**, 437 (2001)

78. Micromass Q-TOF Users' Manual

79. Makarov, A., Anal. Chem., **72**, 1156 (2000)

Chapter 2

Tandem mass spectrometry of *Ginkgo biloba* flavonoids and terpene lactones

2.1 Introduction to *Ginkgo biloba*

2.1.1 The history of *Ginkgo biloba*

Ginkgo biloba is a dietary supplement derived from the leaves of *Ginkgo Biloba Linne*, an ornamental, deciduous tree. The *Ginkgo biloba* tree is considered to be one of the world's oldest living tree species with some specimens living over 1000 years old. *Ginkgo biloba* is also known as ginkgo, hill apricot, maiden hair tree, kew tree, oriental plum tree, silver apricot, silver fruit, and silver plume and is one of the traditional herbal remedies in both European and Chinese traditional medicine. In China, preparations of *Ginkgo biloba* have been used for 5000 years in the treatment of lung ailments such as asthma and bronchitis and as a remedy for cardiovascular disease[1]. The first publication concerning the internal use of the leaves of the *Ginkgo* tree for medical purpose dates back to 1505 A.D[2]. A standardized extract of *Ginkgo biloba* leaves (GBb or EGb 761) has been developed by Dr. Willmar Schwabe GmbH & Co. in Germany[3] and is the common intravenous formulation available in Europe and used in clinical trials. It is extracted by a 27 step process, and the sample '761', out of many thousands assayed samples gave the greatest number of purified components and is thus referred to as EGb761. At present, standardized extracts of *Ginkgo biloba* leaves are used in Europe in the treatment of peripheral and cerebral circulation disorders and are among the top selling dietary supplements in the botanical and herbal remedy market, with estimates of worldwide annual sales varying from 450 million US$ to over 1 billion US$ in 1998[4].

2.1.2 The clinical applications of *Ginkgo biloba* (GBE)

Ginkgo is one of the best researched herbs; the scientific and clinical evidence of the efficacy of *Ginkgo biloba* is centred on two major physiological systems.

2.1.2.1 Improvement in peripheral blood flow

Clinical studies[5, 6, 7] have shown that GBE improves blood circulation by dilating blood vessels and reducing the stickiness of blood platelets. In these studies significant improvements in clinical symptoms resulting from improved blood circulation were demonstrated. The evidence from these studies suggests that *Ginkgo biloba* extract may be particularly effective in treating ailments associated with decreased blood flow to the brain, particularly in elderly individuals. The effects are largely due to the fact that *Ginkgo* strongly inhibits the platelet aggregating factor (PAF) which has the effect of thickening the blood and contributes to a strong inflammatory response. *Ginkgo biloba* treatment is therefore appropriate for conditions, such as, varicose veins, phlebitis, haemorrhoids, intermittent claudication (pain resulting from circulatory impairment), Raynaud Syndrome, male impotence and general poor circulation.

2.1.2.2 Reduction of cerebral insufficiency

GBE effectiveness in reduction of cerebral insufficiency has been demonstrated in human trials[8, 9], especially with the elderly and with measurements including short term memory, period of concentration, task performance and speed of cognitive response in *Ginkgo biloba* has consistently shown significant clinical benefits. Based on studies conducted in laboratories, animals and humans, professional herbalists may

recommend *Ginkgo* for various health problems, such as, Dementia and Alzheimer's disease; eye problems; intermittent claudication; memory impairment; tinnitus as well as a range of other ailments (altitude sickness, asthma, depression, disorientation, headaches, high blood pressure, erectile dysfunction and vertigo). The positive effects of *Ginkgo biloba* are its excellent record of safe, long term use, making this a herbal nutritional supplement that could be used on a consistent basis for anyone with circulatory problems and also as a general "life improvement" supplement during the ageing process.

2.1.3 Main active components of *Ginkgo biloba* extract

2.1.3.1 Flavonoids

Flavonoids are among the most ubiquitous groups of plant secondary metabolites distributed in various foods and medicinal plants. Figure 2.1 shows basic structures of flavonoids commonly found in dietary sources. They are structurally similar to steroid hormones, particularly estrogens, and therefore have been studied extensively over the past several years. Their potential roles in prevention of hormone-dependent cancers including those of breast, prostate and colon, which are leading causes of morbidity and mortality in western countries, have been investigated. In recent years, there has been a significant increase in the number of papers published on the health beneficial effects of flavonoids[10]. Many studies have shown that flavonoids exhibit biological activities, including anti-allergenic, anti-viral, anti-inflammtory and vasodilating actions. Flavonoids are largely planar molecules and their structural variation comes in part from the pattern of substituents, such as, hydroxylation, methoxylation, prenylation, or glycosylation. Flavonoid aglycones are subdivided

into flavone, flavonol, flavonone, and flavanol types. Figure 2.1 shows different type of flavonoids.

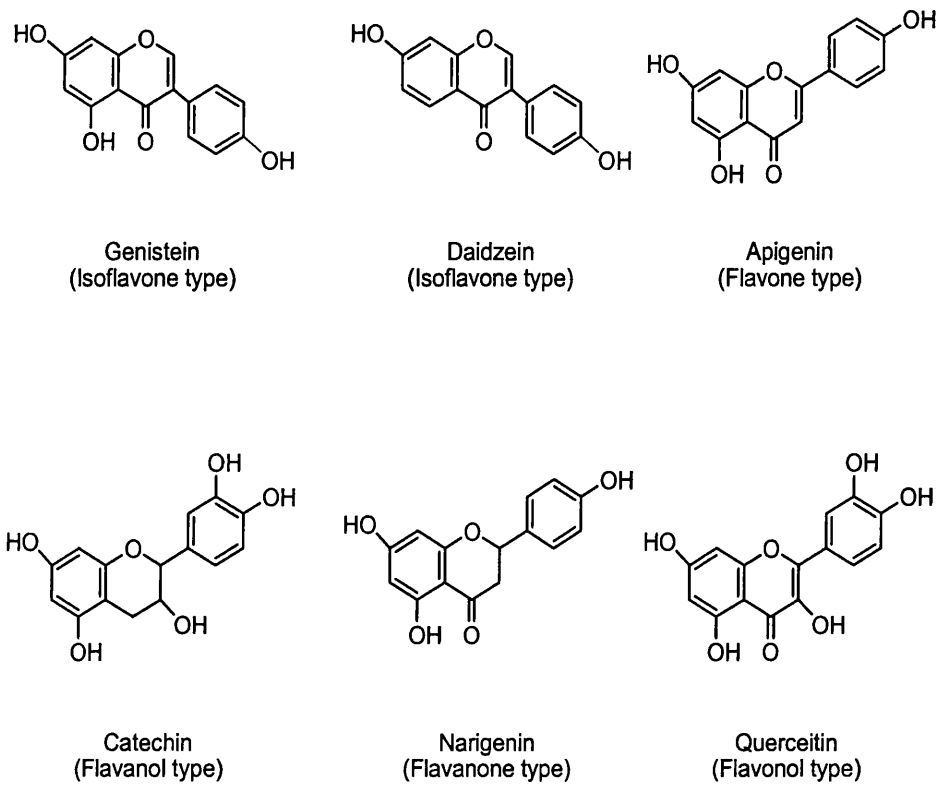


Figure 2.1 Basic structures of flavonoids commonly found in dietary sources

The flavonoid constituents of *Ginkgo* extract are essentially flavonol in nature. Although they are sometimes found as their aglycones, flavonoids most commonly occur in plant materials as flavonoid O-glycosides, in which one or more hydroxyl groups of the aglycones are bound to a carbohydrate moiety, e.g. glucose sugar, forming an acid-labile glycoside O-C bond. In *Ginkgo*, the glycosidic linkage is normally located in position 3 or 7 of a phenolic aglycone (quercetin, kaempferol or isorhamnetin) and the carbohydrate moiety usually being D-glucose, L-rhamnose or glucorhamnose[11,12]. Certain flavonol glycosides, and/or their metabolites, may play key

roles in the therapeutic actions of *Ginkgo* extract. Their structures and the position numbers are shown in Figure 2.2.

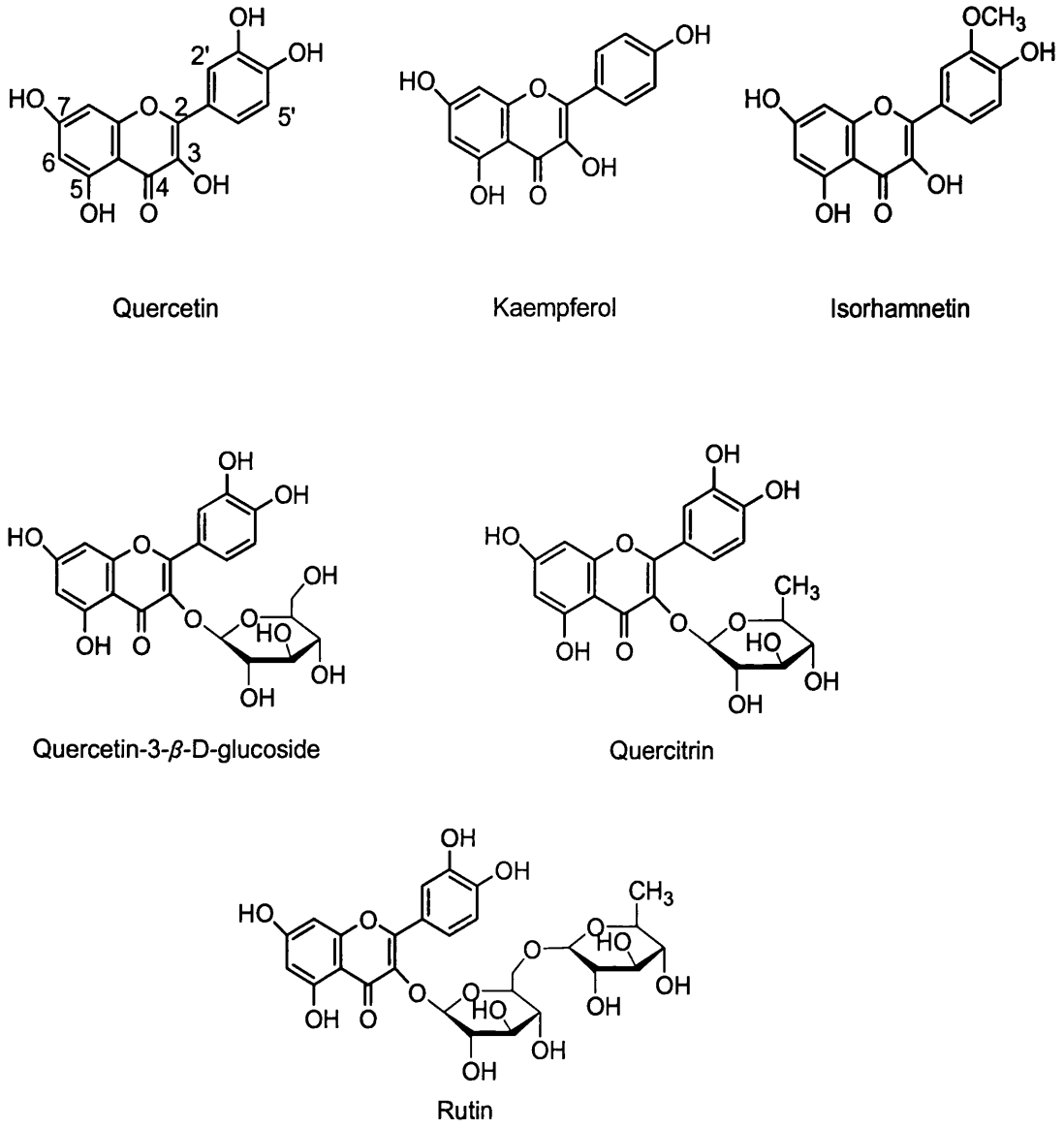
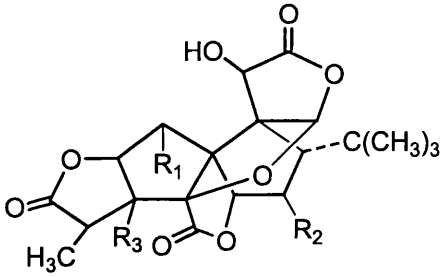


Figure 2.2 Structures of some flavonoids found in *Ginkgo* extract

2.1.3.2 Terpene lactones

Of all the compound classes present in *Ginkgo biloba*, the terpene trilactones have received by far the most attention. This is due to their chemical uniqueness, their

importance in quality control and their analytical challenge[4]. Terpene lactones in *Ginkgo biloba* include various 20-carbon diterpene lactone derivatives (ginkgolides A, B, C, J, and M) and a 15-carbon sesquiterpene (bilobalide). The structures of these highly oxidized terpenes are shown in Figure 2.3. The structures of ginkgolides were originally elucidated by two Japanese groups in the 1960s[13, 14]. In the 1980s, the interest in the ginkgolides suddenly soared when they were found to competitively inhibit the platelet-activating factor (PAF), thus preventing thrombus formation, bronchoconstriction and suppressing allergic reactions.

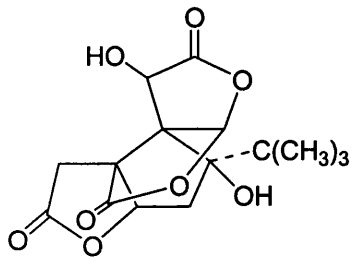


Ginkgolide A: R_1=R_2=H, R_3=OH

Ginkgolide B: R_1=R_3=OH, R_2=H

Ginkgolide C:R_1=R_2=R_3=OH

Ginkgolide J: R_1=H, R_2=R_3=OH



Bilobalide

Figure 2.3 Structures of terpene lactones in *Ginkgo biloba*

2.1.3.3 Other components

Other components in *Ginkgo biloba* include 6-hydroxykynurenic acid, organic acids (vanillic, ascorbic, shikimic, p-coumaric), iron-based superoxide dismutase, benzoic acid derivatives, carbohydrates, sterols and polyprenols[4].

Since *Ginkgo biloba*'s pharmacological activity has been linked to the first two groups of compounds, the flavonoids and terpene trilactones, most of the commercial standardized extracts contain no less than 24% of *Ginkgo* flavonol glycosides and not less than 6% of *Ginkgo* terpenes. Table 2.1 shows different classes of compounds present in the standardized *Ginkgo* extract EGb 761.

Table 2.1 Different classes of compounds present in the EGb761[4]

Compound class	%	Compound class	%
Flavonol glycosides	24.0	High molecular mass compounds	4.0
Terpene trilactones	6.0	Inorganic constituents	5.0
Proanthocyanidins	7.0	Water, solvent	3.0
Carboxylic acids	13.0	Various	3.0
Catechins	2.0	Unknown	13.0
Non-flavonol glycosides	20.0	Alkylphenols	≤5 ppm

2.2 Aim of study

Mass spectrometers have been used to study flavonoids existing as their glycosylated conjugates and their aglycone flavonoids structure. Mass spectra obtained under electron impact (EI) conditions have been widely used for structural investigation of flavonoids[15]. However, complex fragmentations may occur during EI due to the broad spread of internal energy carried by the initially produced $M^{+\cdot}$ ions, which may suppress the peak for $M^{+\cdot}$ and some important primary fragment ions containing structural information. Furthermore, EI is not suitable for the analysis of polyphenolic flavonoids, such as flavonols and O-glycosides because of their high polarity and low volatility.

The introduction of fast atom bombardment (FAB) with collision induced dissociation (CID) was applied to the structural characterisation of flavones[16], flavonols[17] and their O- and C-glycosides[18]. An LCQ ion trap mass spectrometer was used to determine the structures of various flavonoids in natural nutrition supplements[19].

To help the determination and identification of active components in *Ginkgo biloba* extract and biological samples, the mass spectrometric properties of six flavonoids and four terpene lactones were investigated. MS/MS of the flavonoids have been well studied but with mass spectrometry of comparatively low mass accuracy such as ion traps. No reports have studied the fragmentation of terpene lactones by tandem mass spectrometry. In this study, ESI/MS^n analyses were conducted on both flavonoids and terpene lactones using an ion trap mass spectrometer to generate MS/MS and MS^n spectra and accurate mass analyses were performed on a Q-TOF instrument in order to assist in the elucidation of the fragment pathways of these compounds.

2.3 Experimental

2.3.1 Materials

Ginkgolide A (**GA**), ginkgolide B (**GB**), ginkgolide C (**GC**), bilobalide (**BL**), quercetin dehydrate (**QD**), quercetin-3-β-D-glucoside (**QG**), quercetin-3-rhamnoside (**QH**), kaempferol (**KF**), isorhamnetin (**IR**) and rutin (**RH**) were purchased from Sigma (St. Louis, MO, USA).

HPLC grade solvents: methanol, acetonitrile, formic acid and acetic acid were purchased from Fisher chemicals (Loughborough, UK) and were used without further

purification. Water was purified with a Milli-Q deionisation unit (Millipore, Bedford, MA, USA). Gases used included oxygen free nitrogen and helium which were purchased from BOC Ltd (Surrey, UK).

2.3.2 MS and MS^n of standard reference compounds

The ESI-MS and further MS^n of standard flavonoids and terpene lactones were conducted using an LCQ ion trap mass spectrometer (Finnigan, Hemel hempstead, UK) equipped with an ESI source. The full scan analysis of the compounds, utilised to determine which polarity was the most useful in this study was initially conducted. For the MS^n analysis the instrument was utilised in negative ion mode and the parameter settings are shown in Table 2.2. Standards were dissolved in methanol to a final concentration of about 100 μg/mL, and standard solutions were infused into the ESI source by syringe at a flow rate of 3 μL/min.

Table 2.2 ESI-MS conditions

Parameter	Negative ion mode
Sheath gas flow (arbitrary units)	60
Auxiliary gas flow (arbitrary units)	15
Spray voltage (kV)	4.5
Capillary temperature (°C)	230
Capillary voltage (V)	20
Tube lens offset (V)	10

2.3.3 Accurate mass analysis of *Ginkgo biloba* reference standards

The accurate mass measurement of the product ions formed by the fragmentation of the *Ginkgo biloba* reference standards was performed on the Q-TOF Ultima mass spectrometer (Micromass, Manchester, UK) in negative mode equipped with a lock spray source allowing both a reference compound and analyte to be studied simultaneously therefore allowing the acquisition of accurate mass data. The reference standard stock solutions were diluted with 50% methanol aqueous solution to a concentration of approximately 10 pmol/μL and infused into mass spectrometer with a syringe at flow rate of 1 μL/min. The spray voltage used was 2.5 kV, source temperature of 80 °C, desolvation temperature of 120 °C, TOF voltage of 9.1 kV, MCP voltage of 2300 kV, a cone gas flow of 100 L/hr, desolvation flow of 250 L/hr and the collision energy was set to 4 eV during the full scan MS experiments and 15-35 eV for MS/MS experiments. Cyclic CMP was used as lock mass reference and was infused at the same time with the *Ginkgo biloba* standards at a flow rate of 3 μL/min and was used to adjust the active components mass peaks produced in the ESI-MS and MS/MS experiments to improve their mass accuracy.

2.4 Result and discussions

2.4.1 Full scan MS under positive ion mode (ESI ion trap)

Both positive and negative ionisation modes were evaluated for the investigation of the *Ginkgo* flavonoids and terpene lactones. Tables 2.3 and 2.4 show the ions generated by the standards in positive and negative ionisation full scan mode, respectively. In positive ionisation mode, the protonated molecular ion signal is very weak and sodium

or potassium adducts are dominant (Figure 2.4). In the spectra of all three flavonoid glycosides, the aglycone counterpart quercetin (*m/z*=303) can be observed.

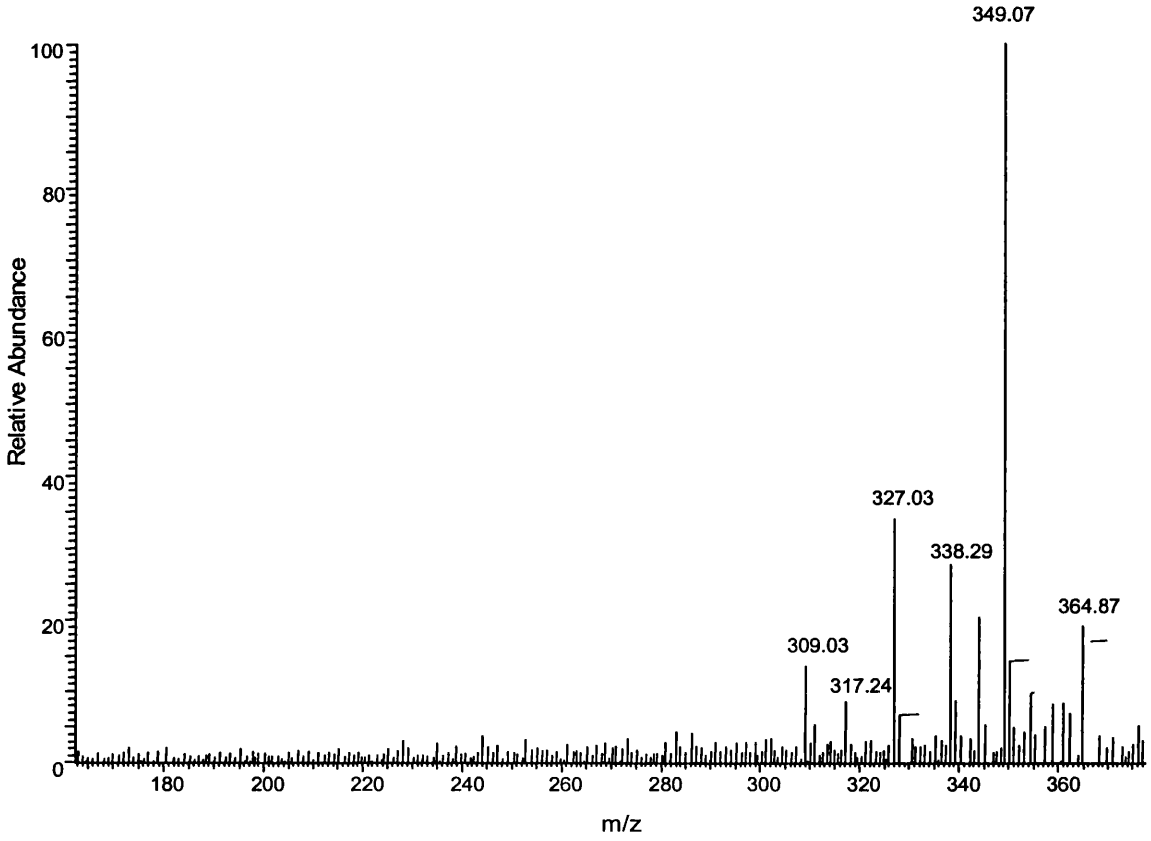


Figure 2.4 Mass spectrum of bilobalide in positive mode

$[M+H]^+$: 327, $[M+Na]^+$: 349

Table 2.3 ESI-MS with ion trap mass analyser of *Ginkgo biloba* standards *m/z* and relative abundance (%, in brackets) of major ions in positive ionisation mode

Ions	$[M+H]^+$	$[M+Na]^+$	$[M+K]^+$	$[2M+Na]^+$	$[M+H_2O]^+$	$[M+H]^+$-Glc[1]	$[M+H]^+$-Rha[2]	$[M+H]^+$-Glc-Rha[3]
BL	327(10)	349(30)	—	675(100)	—	—	—	—
GA	409(8)	431(100)	447(20)	—	—	—	—	—
GB	425(4)	367(28)	—	—	442(100)	—	—	—
GC		383(62)	—	—	—	—	—	—
IR	317(78)	339(30)	—	631(8)	—	—	—	—
KF	—	—	—	—	—	—	—	—
QD	303(11)	325(12)	—	627(100)	—	—	—	—
QG	465(5)	487(100)	—	—	—	303(5)	—	—
QH	449(20)	471(100)	—	—	—	—	303(21)	—
RH	611(8)	633(100)	—	—	—	—	—	303(4)

[1] loss of one glucoside
[2] loss of one rhamnoside
[3] loss of one glucoside and one rhamnoside

Table 2.4 ESI-MS with ion trap mass analyser of *Ginkgo biloba* standards *m/z* and relative abundance (%, in brackets) of major ions in negative ionisation mode

Ions	$[M-H]^-$	$[M-2CO]^-$	$[M-H]^-$-CO	$[2M-H]^-$	$[M-H]^-$-Glc	$[M-H]^-$-Rha	$[M-H]^-$-Glc-Rha
BL	325(100)			651(32)	—	—	—
GA	407(100)	351(84)	379(12)	—	—	—	—
GB	423(100)	367(28)	395(8)	—	—	—	—
GC	439(100)	383(62)	411(14)	—	—	—	—
IR	315(100)	—	—	631(8)	—	—	—
KF	285(100)	—	—	571(11)	—	—	—
QD	301(100)	—	—	603(37)	—	—	—
QG	463(100)	—	—	—	300(17)	—	—
QH	447(100)	—	—	—	—	300(22)	—
RH	609(1000	—	—	—	447(12)	—	300(19)

2.4.2 MS under negative ion mode (ESI ion trap)

In negative mode the deprotonated molecular ions of the reference standards give much stronger signals then those generated in positive mode. In this mode, the deprotonated molecular ions $[M-H]^-$ of all the standards are the base peaks, no adducts were observed (Figure 2.5). The aglycone counterpart quercetin (*m/z*=301) can be observed in all three flavonoid glycosides. For ginkgolide A, B and C, the ion corresponding to loss of one and two carboxyl group $[M-CO]^-$ and $[M-2CO]^-$ can also be observed (Table 2.4) and the signal intensity is much stronger in negative mode than positive mode as shown in Figure 2.6.

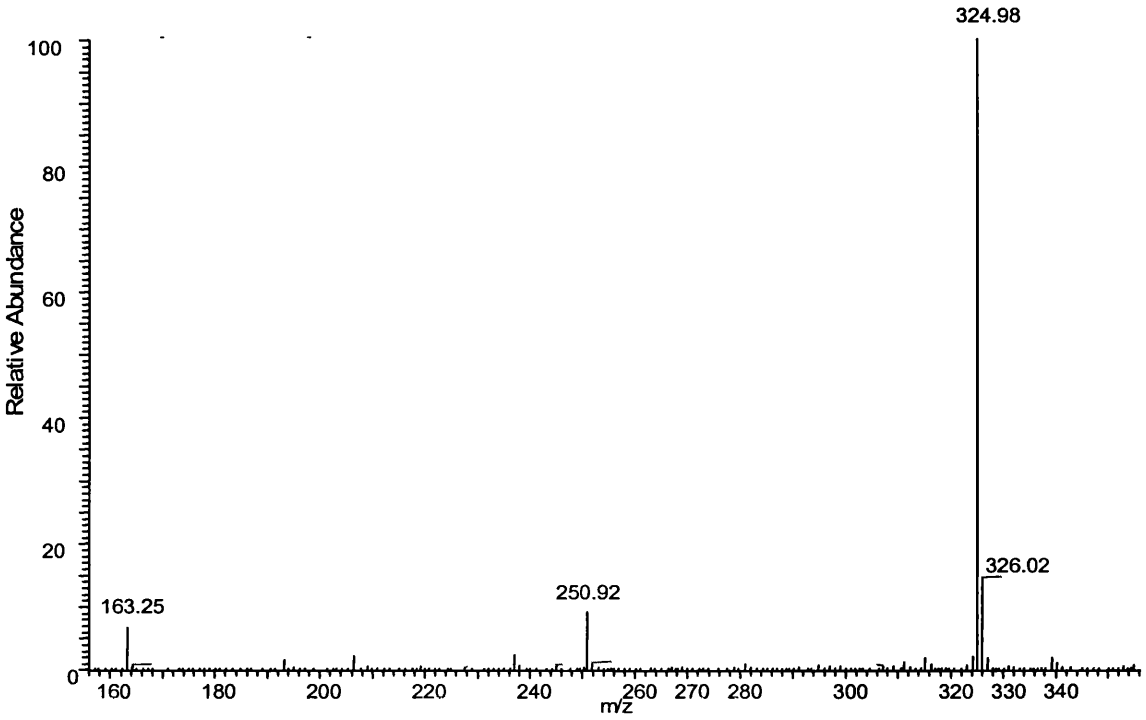


Figure 2.5 LC/ESI MS of bilobalide in negative mode $[M-H]^-$: 325

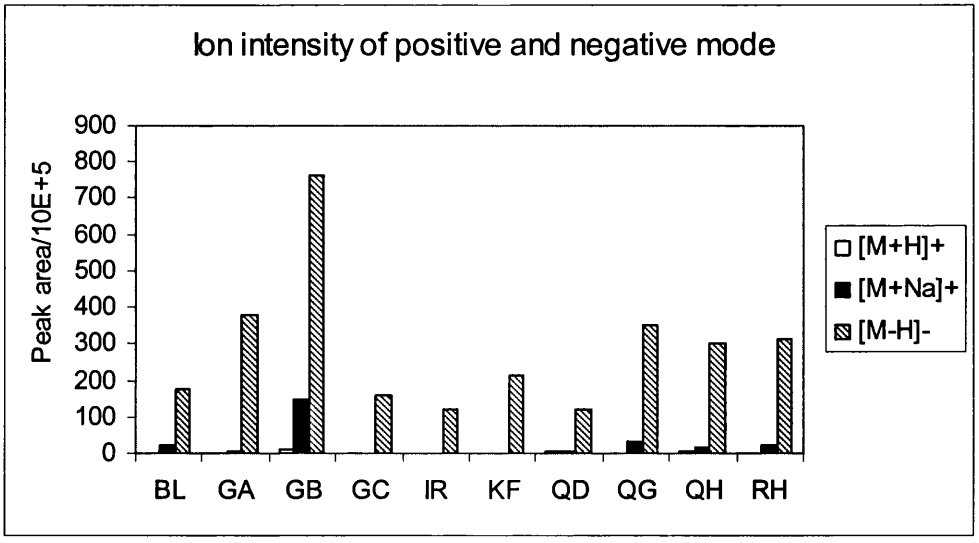


Figure 2.6 Comparison of ion intensity in positive and negative mode for the 10 components of *Ginkgo biloba*

2.4.3 Collision induced dissociation of *Ginkgo biloba* standards

2.4.3.1 *Ginkgo biloba* flavonoids glycosides

In all the tandem mass spectrometry, molecular ions $[M-H]^-$ were chosen as the precursor ion as they are the base peak ions in negative ionisation mode. Figures 2.7-2.9 show the ESI/MS/MS of three *Ginkgo* flavonoid glycosides when their deprotonated molecular ions were selected as precursor ions and the product ions recorded. As can be observed, the three flavonoid glycosides fragment to lose the glycoside and produce the corresponding aglycone, in this case quercetin, as the main fragment ion. This is very important characteristic which can be used to identify the flavonoid glycosides in *Ginkgo biloba* commercial samples as reported in Chapter 4. The proposed fragment pathway is shown in Figure 2.10.

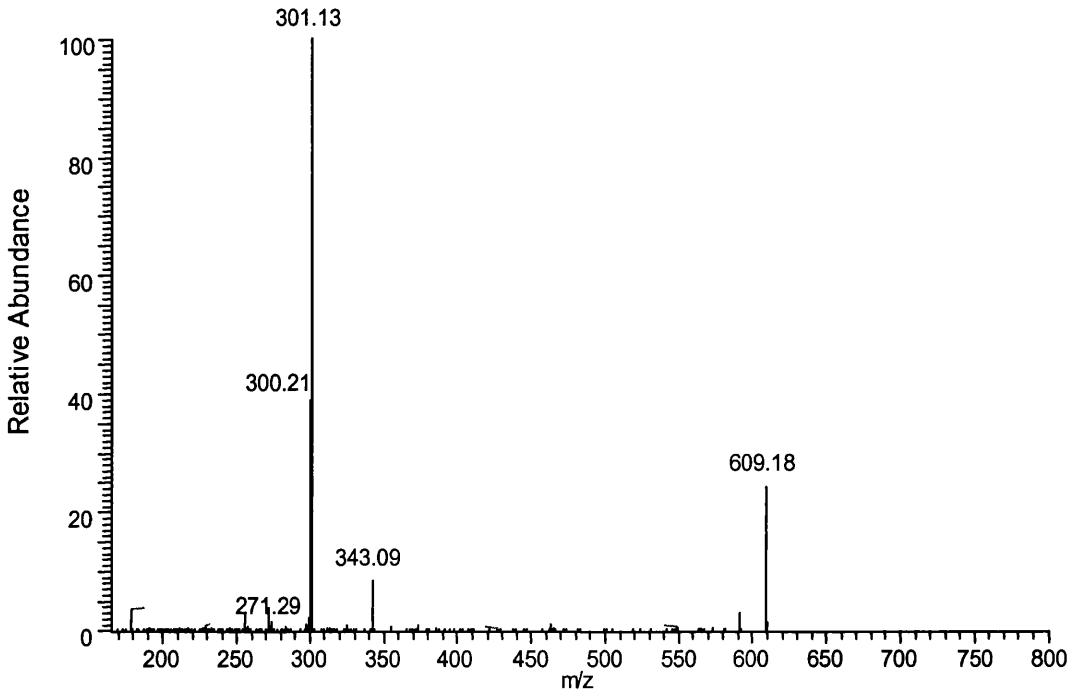


Figure 2.7 Negative ESI-MS/MS mass spectrum of rutin hydrate
Precursor ion $[M-H]^-$: 609; collision energy 28%.

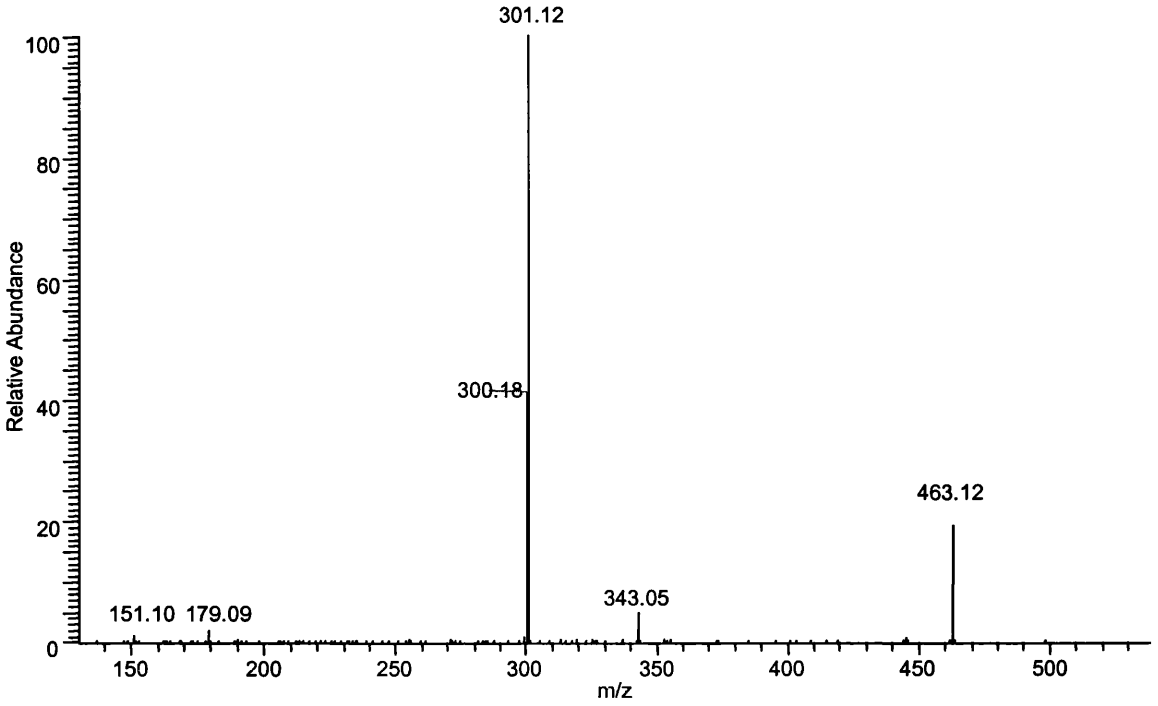


Figure 2.8 Negative ESI-MS/MS mass spectrum of quercetin-3-β-D-glucoside. Precursor ion $[M-H]^-$: 463; collision energy 27%.

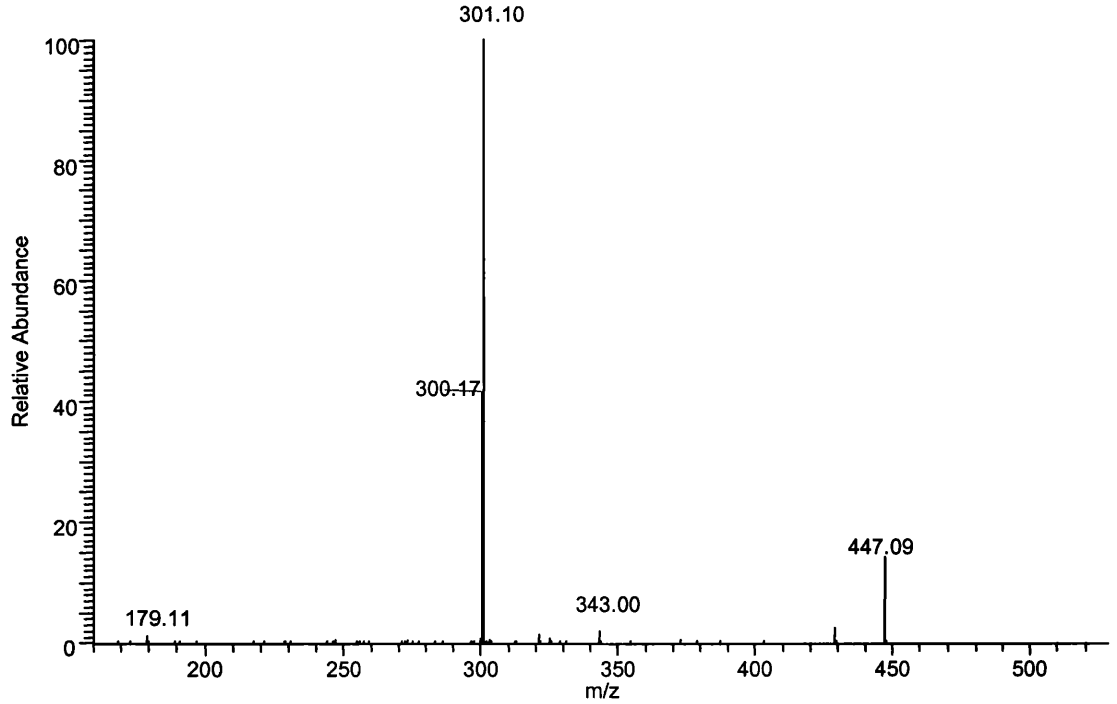
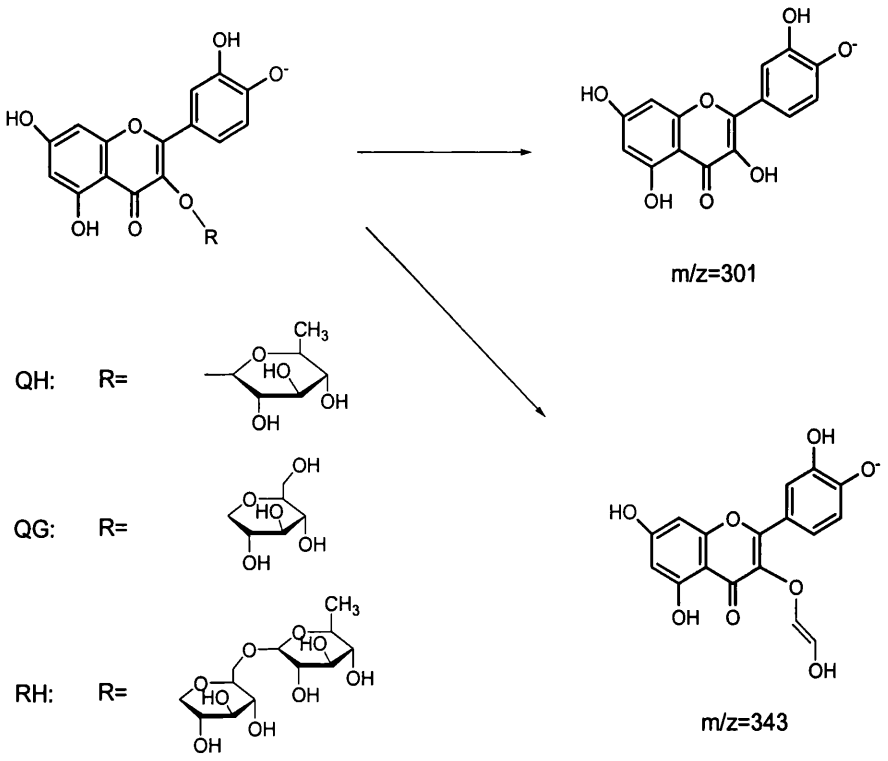


Figure 2.9 Negative ESI-MS/MS mass spectrum of quercetin hydrate. Precursor ion $[M-H]^-$: 447; collision energy 28%.

Figure 2.10 Proposed fragment pathways of *Ginkgo* flavonoid glycosides



It should be noted that both 300 and 301, corresponding to the quercetin fragment and a radical quercetin anion respectively, are the major ions in Figures 2.7-9. According to the literature[20], the relative abundance of the radical aglycone to the aglycone product ion is dependent on the collision energy with a relative increase in radical aglycone product ion formation detected with an increase in collision energy.

2.4.3.2 *Ginkgo biloba* flavonoid aglycones

A systematic ion nomenclature for flavonoid aglycones has been proposed by Claey and co-worker[21]. As shown in Figure 2.11, the symbols $^{i,j}A^+$ and $^{i,j}B^+$ are used to designate primary product ions containing intact A and B rings, respectively. The superscripts i and j refer to the bonds of the C-ring that have been broken.

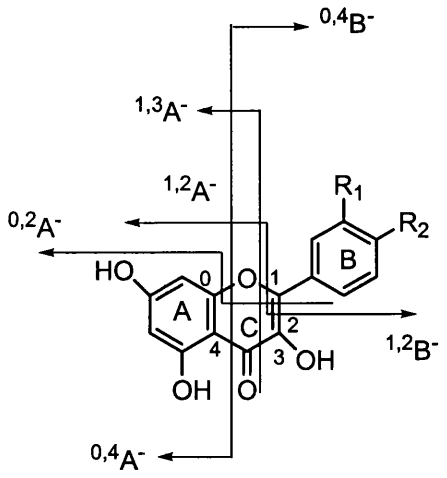


Figure 2.11 Nomenclature and diagnostic product ions of deprotonated flavonols formed by ESI-ion trap mass spectrometer

The MS/MS spectra of flavonoid aglycones quercetin, kaempferol and isorhamnetin are shown in Figures 2.12-14, respectively. Major MS/MS fragment ions of the three flavonoid aglycones are summarized in Table 2.5. In the case of isorhamnetin, the spectrum is dominated by the $[M\text{-}H\text{-}CH_3]^{\cdot-}$ with *m/z* 300, which is the loss of a CH_3 radical from the precursor ion. This fragment pathway was supported by accurate mass analysis. Further fragmentation of *m/z* 300 was performed by ion trap mass spectrometer and shown in Figure 2.15. Quercetin and kaempferol share similar fragment patterns while kaempferol gives more fragment ions. $^{1,3}A^-$, $^{0,4}A^-$ and $^{0,4}B^-$ were observed in the fragmentation spectra of both compounds. $^{1,2}A^-$ was observed in quercetin and $^{0,2}A^-$ in kaempferol. Ions corresponding to the loss of H_2O, CO, COO, ketene and their combined loss from precursor ions are observed in the spectra. The proposed fragment pathways determined by the combined ion trap MS^n and accurate mass measurement, are shown in Figure 2.16, 2.17, 2.18, respectively.

Table 2.5 Major product ions formed from ions using ESI MS/MS for the flavonoid aglycones quercetin and kaempferol in *Ginkgo biloba*

Ions	QD	KF	IR	*IR (M=300)
$[M-H]^-$	301	285	315	
$[M-H-CH_3]^-$			300	
$[M-OH]^-$				283
$[M-H-H_2O]^-$	283	267		
$[M-H-CO]^-$	273	257		271
$[M-H-COO]^-$	257	241		
$[M-H-CO-COO]^-$	229	213		227
$[M-H-C_2H_2O]^-$		243		
$[M-H-H_2O-CO]^-$		239		
$[M-H-2CO]^-$		229		243
$[M-H-H_2O-2CO]^-$		211		
$[M-H-3CO]^-$		201		215
$[M-H-2CO-COO]^-$		185		
$^{1,3}A^- + H_2O$	169	169		169
$^{1,2}A^-$	179			
$^{0,2}A^-$		163		163
$^{1,3}A^-$	151	151		151
$^{0,4}B^-$	193	187		
$^{0,4}A^-$	107	107		107

* This column is the MS^3 of the MS/MS product ion m/z=300

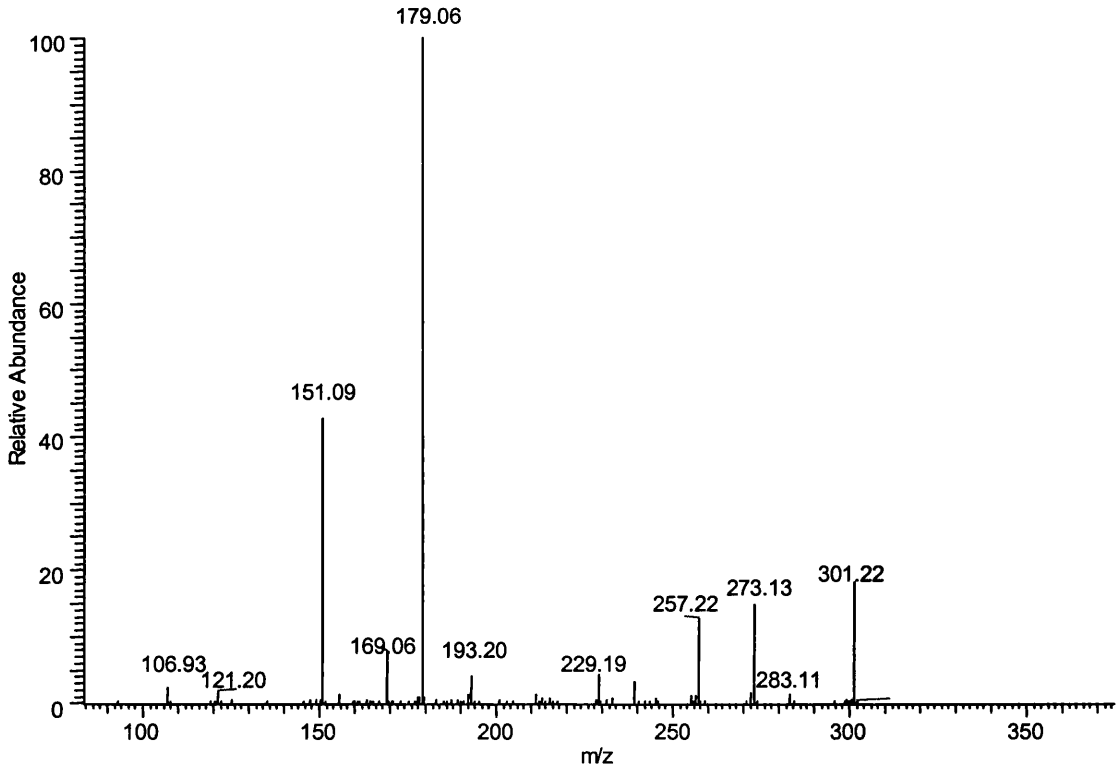


Figure 2.12 Negative LC/ESI-MS^2 mass spectrum of quercetin.

Precursor ion $[M-H]^-$: 301; collision energy 36%

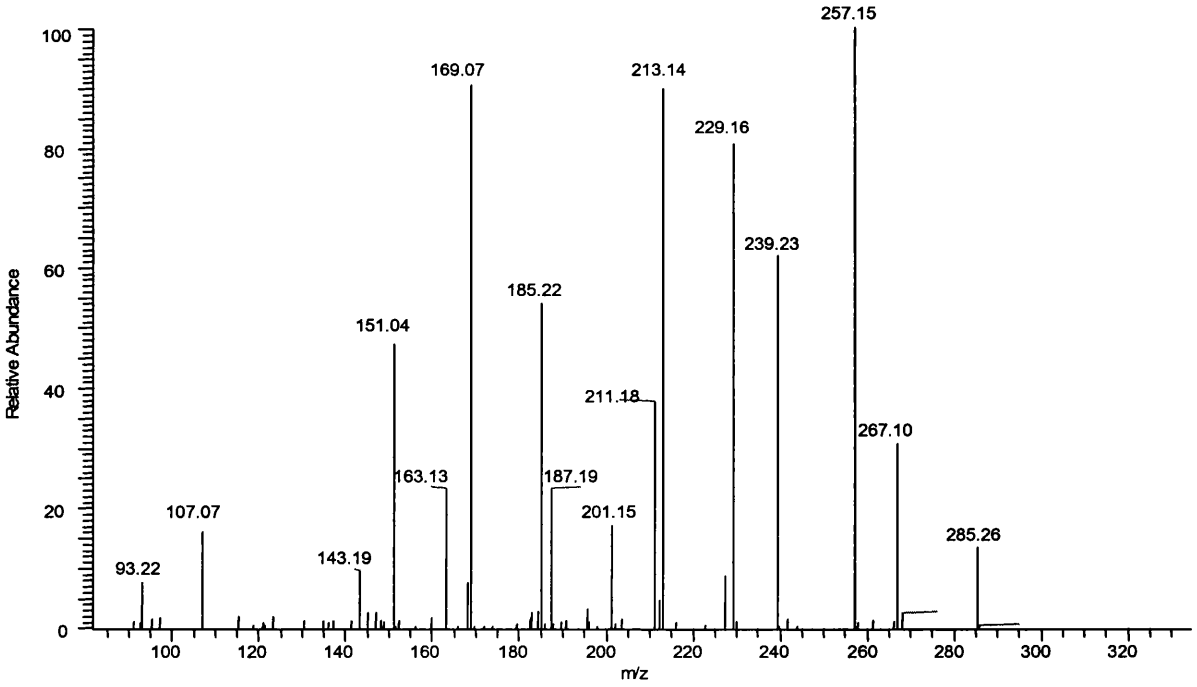


Figure 2.13 Negative LC/ESI-MS^2 mass spectrum of kaempferol.

Precursor ion $[M-H]^-$: 285; collision energy 46%

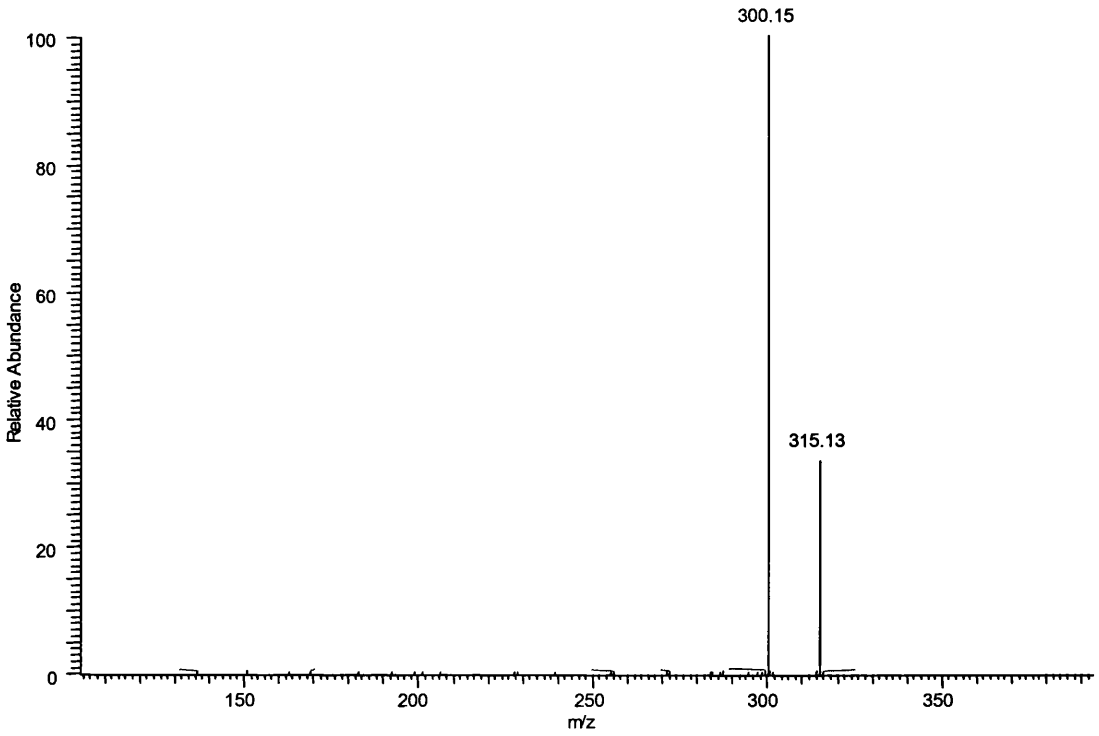


Figure 2.14 Negative LC/ESI-MS^2 mass spectrum of isorhamnetin

Precursor ion $[M-H]^-$: 315; collision energy 34%

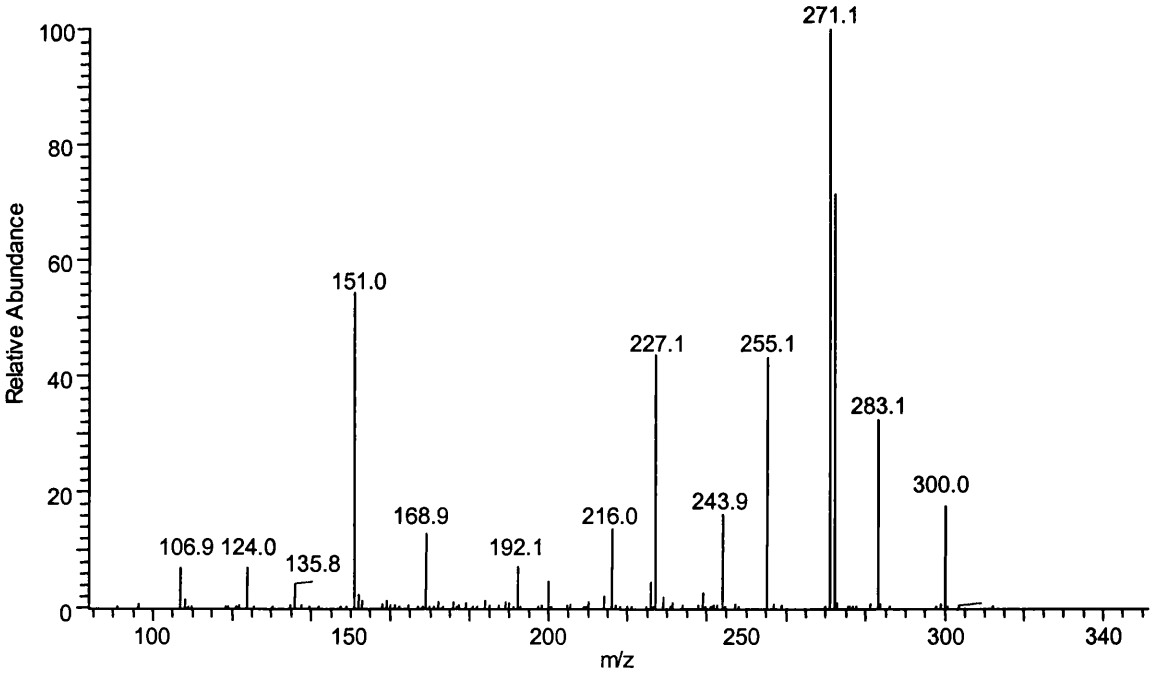


Figure 2.15 Negative LC/ESI-MS^3 mass spectrum of isorhamnetin

315 CID 34% 300 CID 35%

Figure 2.16 Proposed fragment pathway of quercetin MS² MS³ MS⁴ MS⁵

(mass accuracy in ppm shown in brackets)

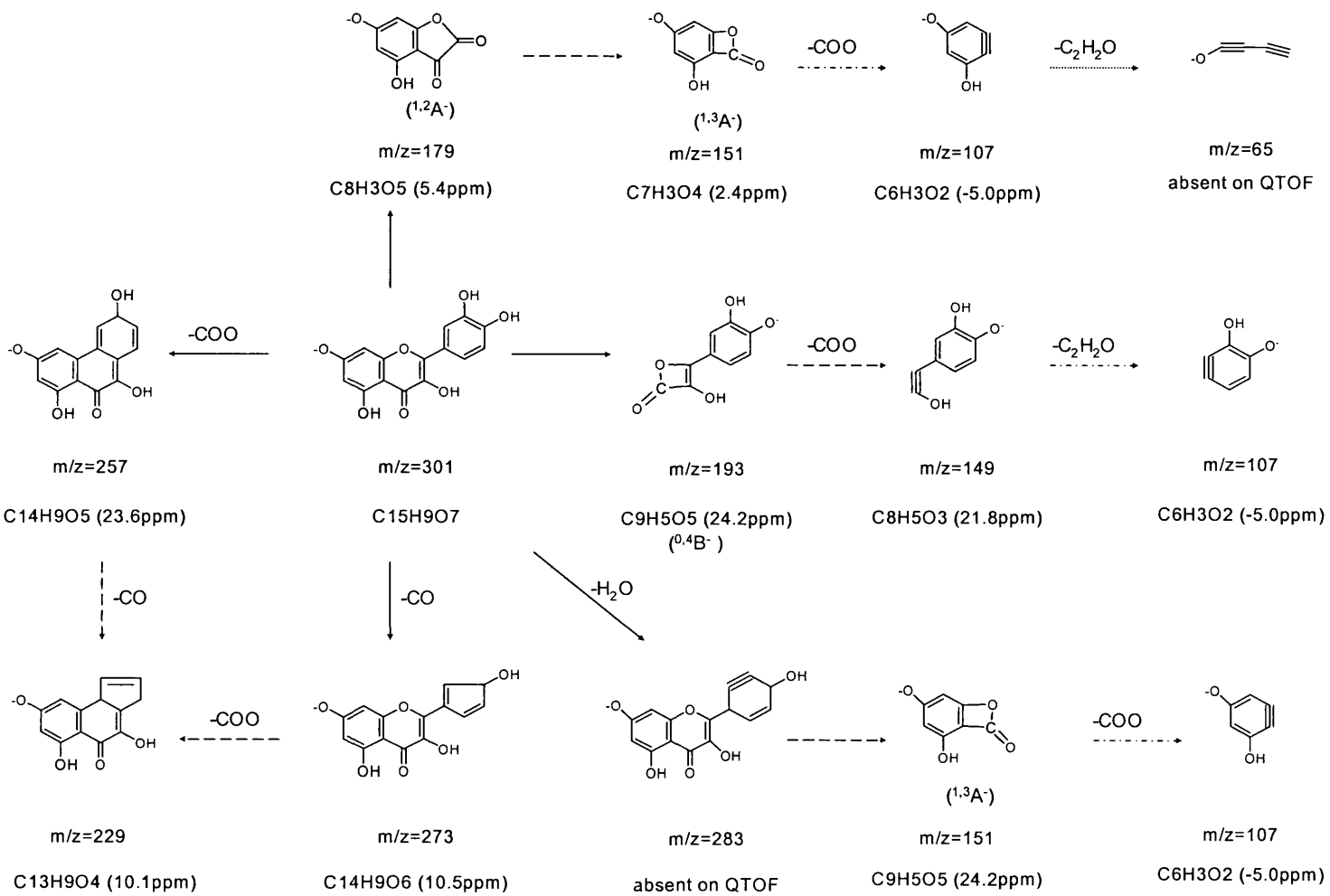


Figure 2.17 Proposed fragment pathway of kaempferol

(mass accuracy in ppm shown in brackets)

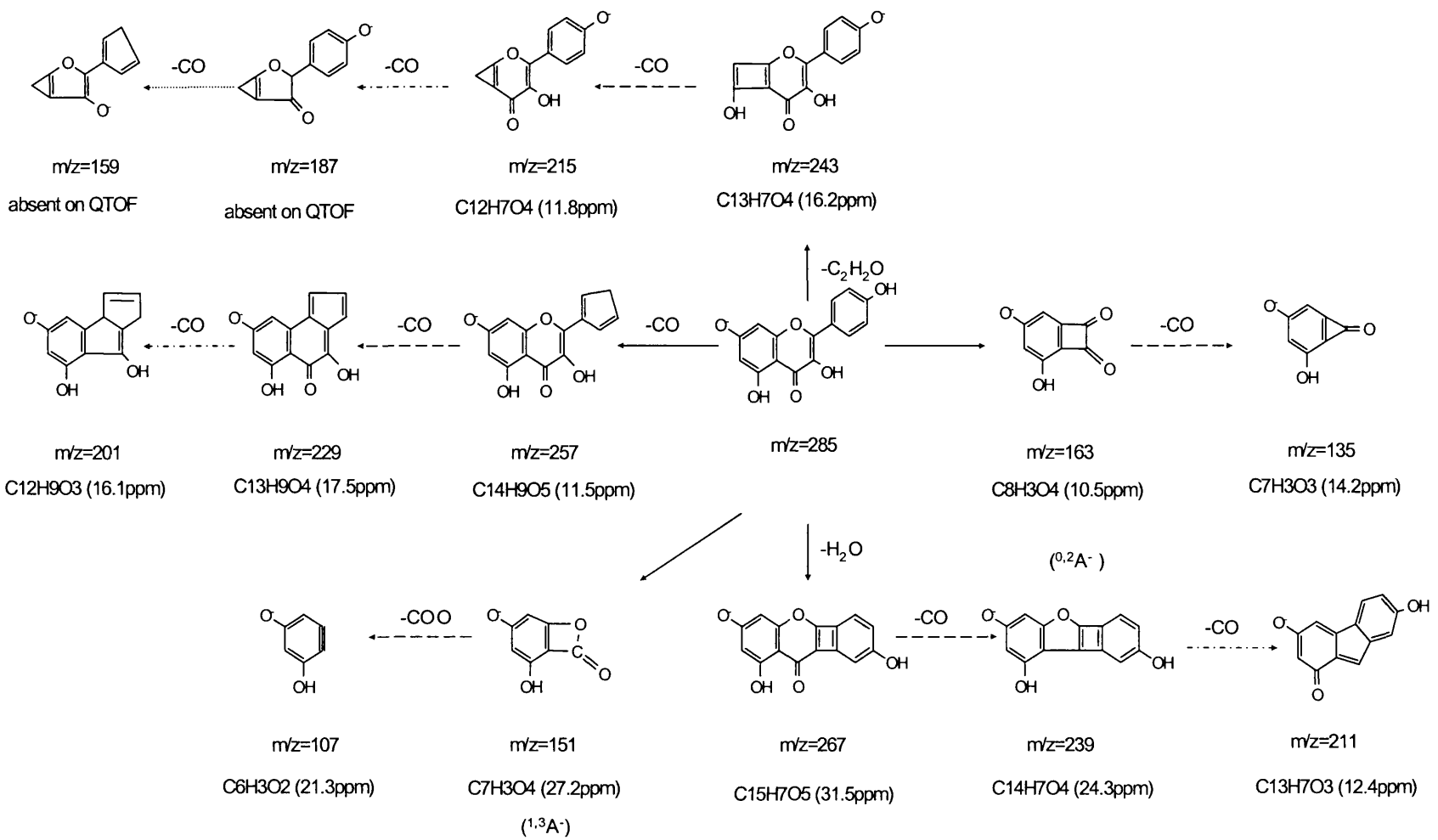
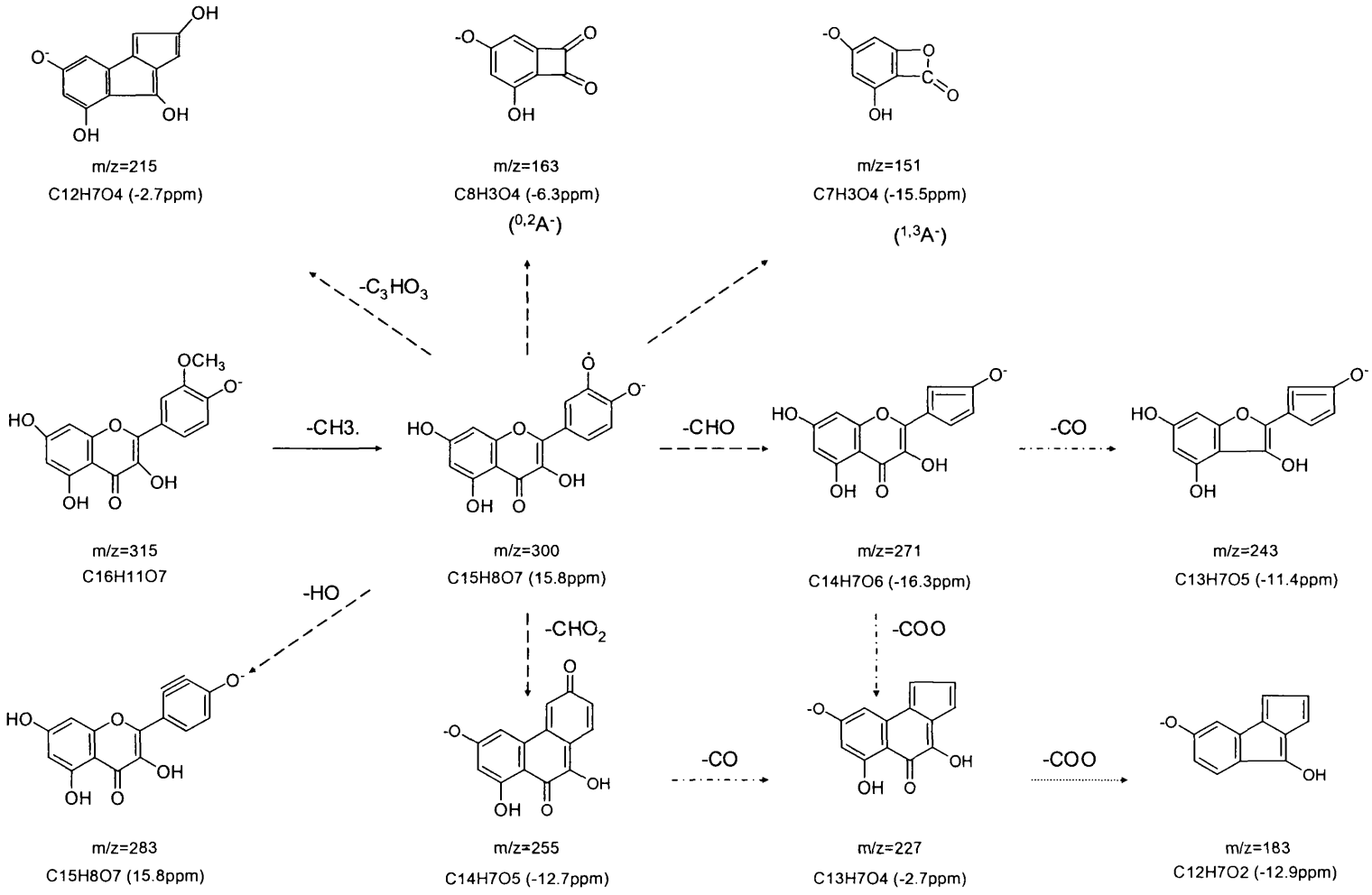


Figure 2.18 Proposed fragment pathway of isorhamnetin MS^2 MS^3 MS^4 MS^5

(mass accuracy in ppm shown in brackets)



2.4.3.3 *Ginkgo biloba* terpene lactones

The MS/MS spectra of $[M-H]^-$ ions of bilobalide and ginkgolide A, B, C are shown in Figures 2.19-2.22. All the *Ginkgo* terpene lactones exhibit similar fragmentation pathways. The summarized fragment ions are shown in Table 2.6. *Ginkgo* terpene lactones have unique chemical structures, they are highly oxidized terpenes with many carbonyl and hydroxyl function groups. Consecutive loss of H_2O, CO, COO from molecular ions was observed. Proposed fragment pathways for *Ginkgo* terpene lactones are shown in Figures 2.23-2.26 by combining ion trap MS to the end fragment information and accurate mass measurement results.

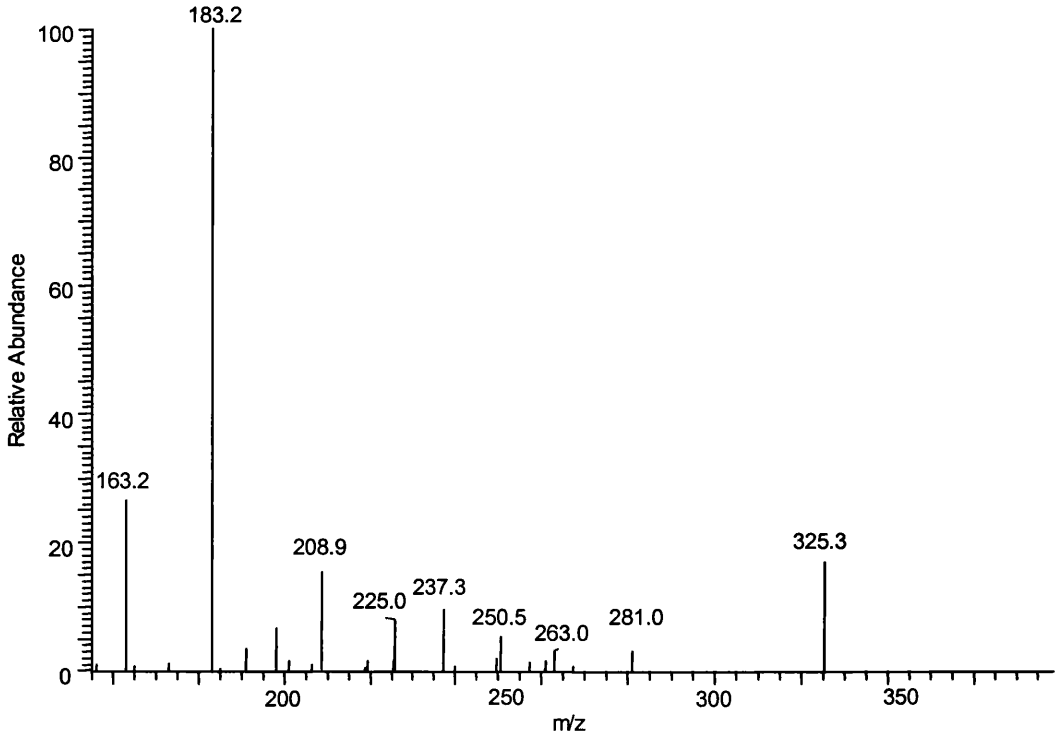


Figure 2.19 Negative LC/ESI-MS^2 mass spectrum of bilobalide
Precursor ion $[M-H]^-$: 325; collision energy 50%

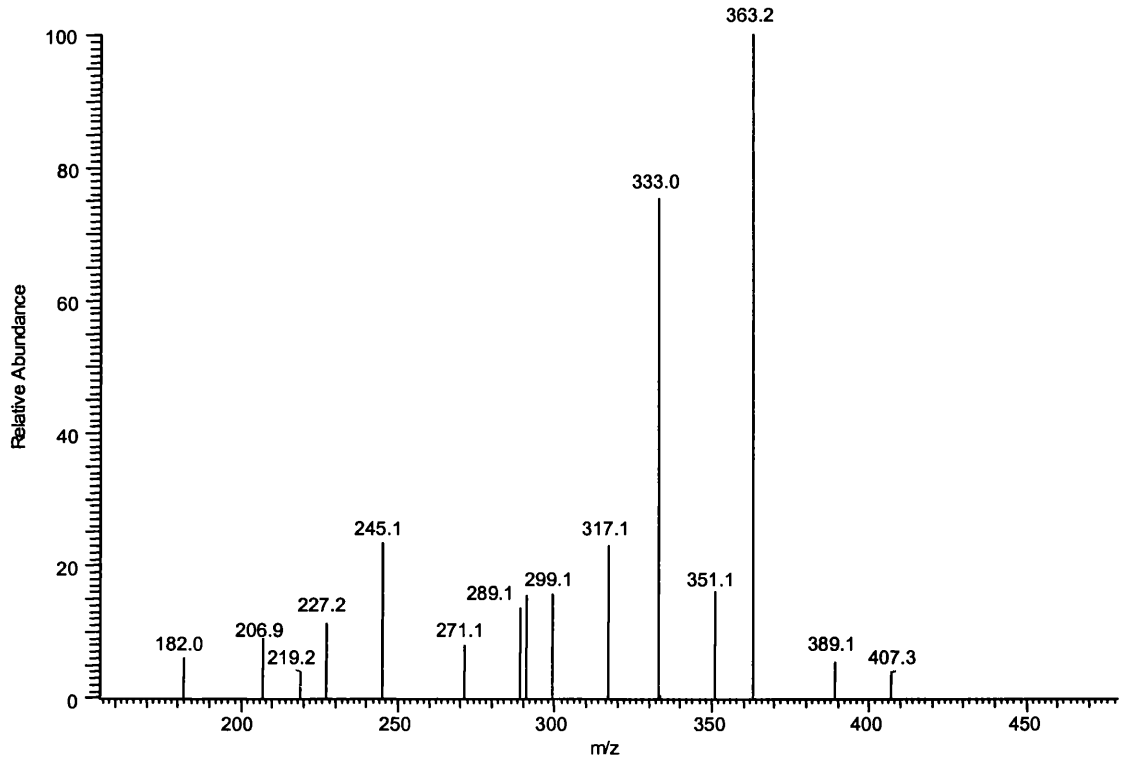


Figure 2.20 Negative LC/ESI-MS^2 mass spectrum of ginkgolide A

Precursor ion $[M-H]^-$:407; collision energy 28%

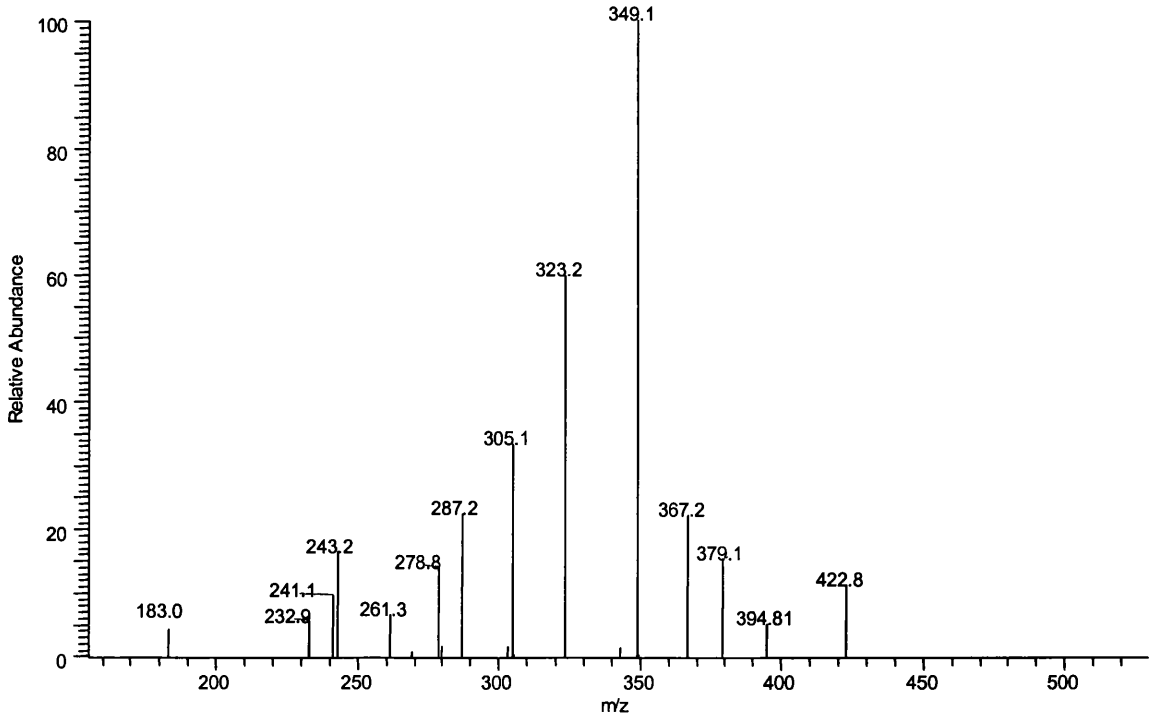


Figure 2.21 Negative LC/ESI-MS^2 mass spectrum of ginkgolide B

Precursor ion $[M-H]^-$:423; collision energy 28%

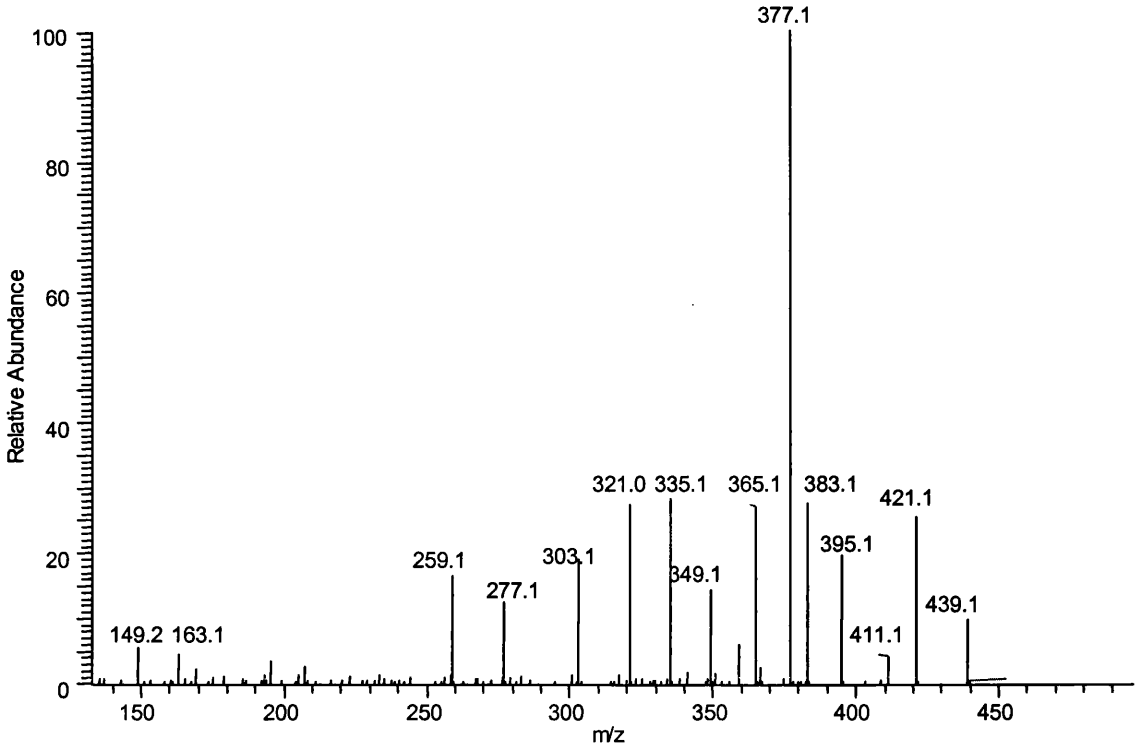


Figure 2.22 Negative LC/ESI-MS^2 mass spectrum of ginkgolide C

Precursor ion $[M-H]^-$:439 collision energy 26%

Table 2.6 Major product ions formed from ions using ESI MS/MS for the terpene lactones in *Ginkgo biloba*

Ion	GA	GB	GC	BL
$[M-H]^-$	407	423	439	325
$[M-H-H_2O]^-$	389	405	421	—
$[M-H-CO]^-$	—	395	411	
$[M-H-COO]^-$	363	379	395	281
$[M-H-2COO]^-$	—	—	—	237
$[M-H-2COO-CO]^-$	—	—	—	209
$[M-H-2CO]^-$	351	367	383	—
$[M-H-2CO-COO]^-$	—	323	—	225
$[M-H-2CO-2COO]^-$	—	279	—	—
$[M-H-H_2O-COO]^-$	—	361	377	
$[M-H-2CO-H_2O]^-$	333	349	365	251
$[M-H-2CO-H_2O-COO]^-$	289	305	321	—
$[M-H-2CO-2H_2O-COO]^-$	271	287	303	—
$[M-H-2CO-H_2O-2COO]^-$	245	261	277	163
$[M-H-2CO-2H_2O-2COO]^-$	227	243	259	—
$[M-H-4CO-CH_2O-]^-$				183

Figure 2.23 Proposed fragment pathway of bilobalide (MS^2 MS^3 mass accuracy in ppm shown in brackets)

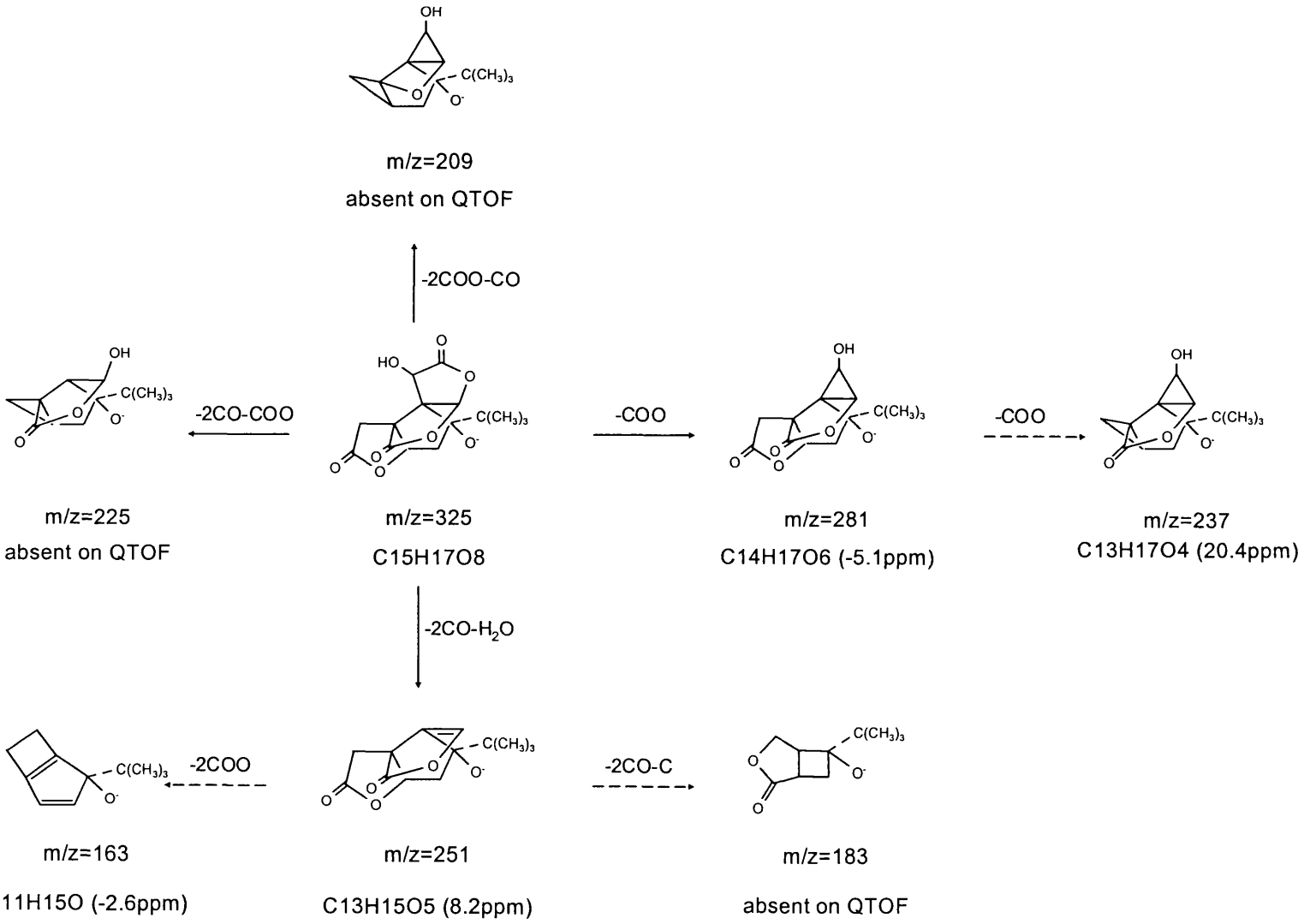


Figure 2.24 Proposed fragment pathway of ginkgolide A ($\xrightarrow{MS^2}$ $\xrightarrow{MS^3}$ $\xrightarrow{MS^4}$ mass accuracy in ppm shown in brackets)

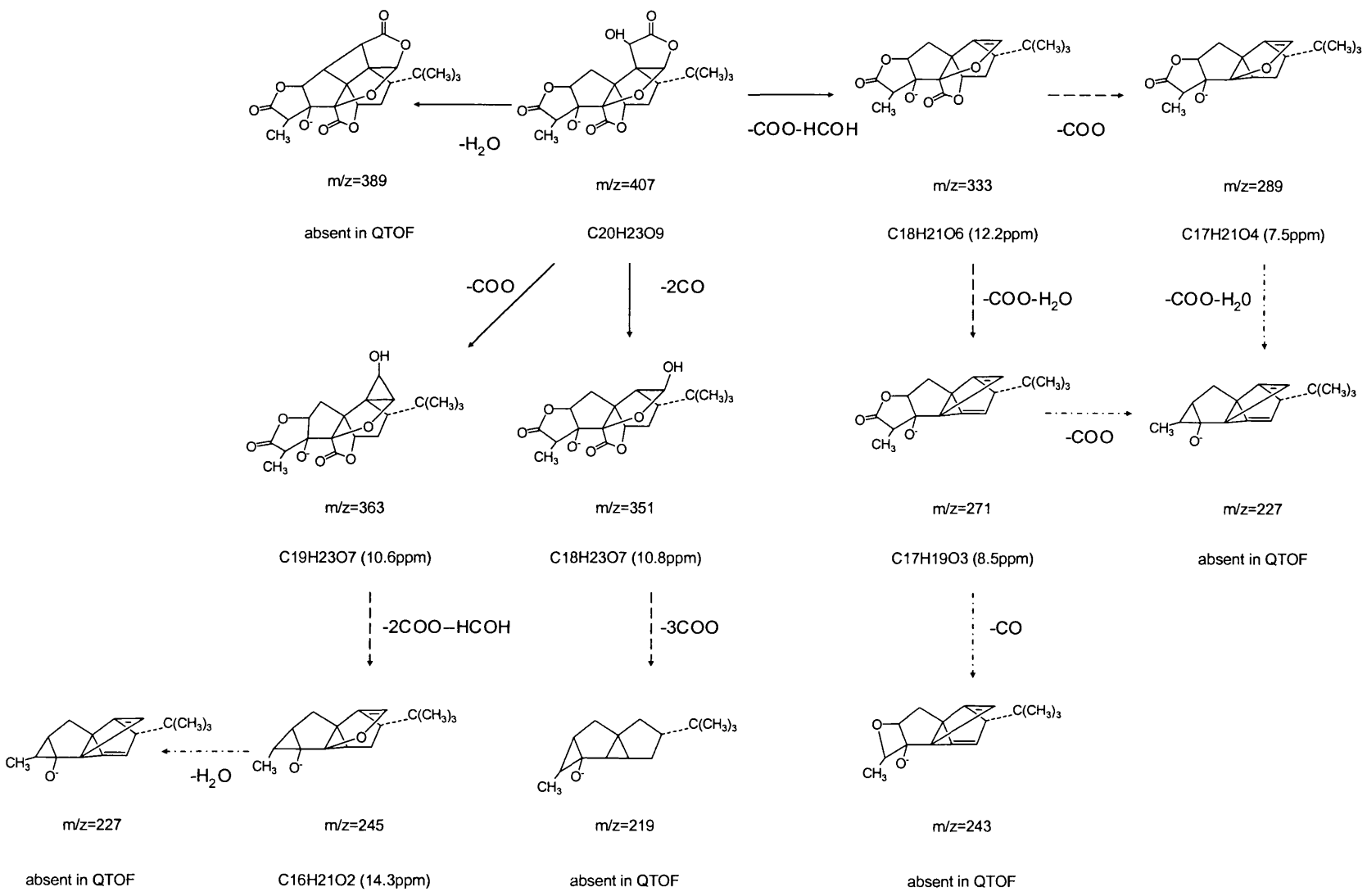


Figure 2.25 Proposed fragment pathway of ginkgolide B ($\xrightarrow{MS^2}$ $\dashrightarrow$ $\cdots\rightarrow$ mass accuracy in ppm shown in brackets)

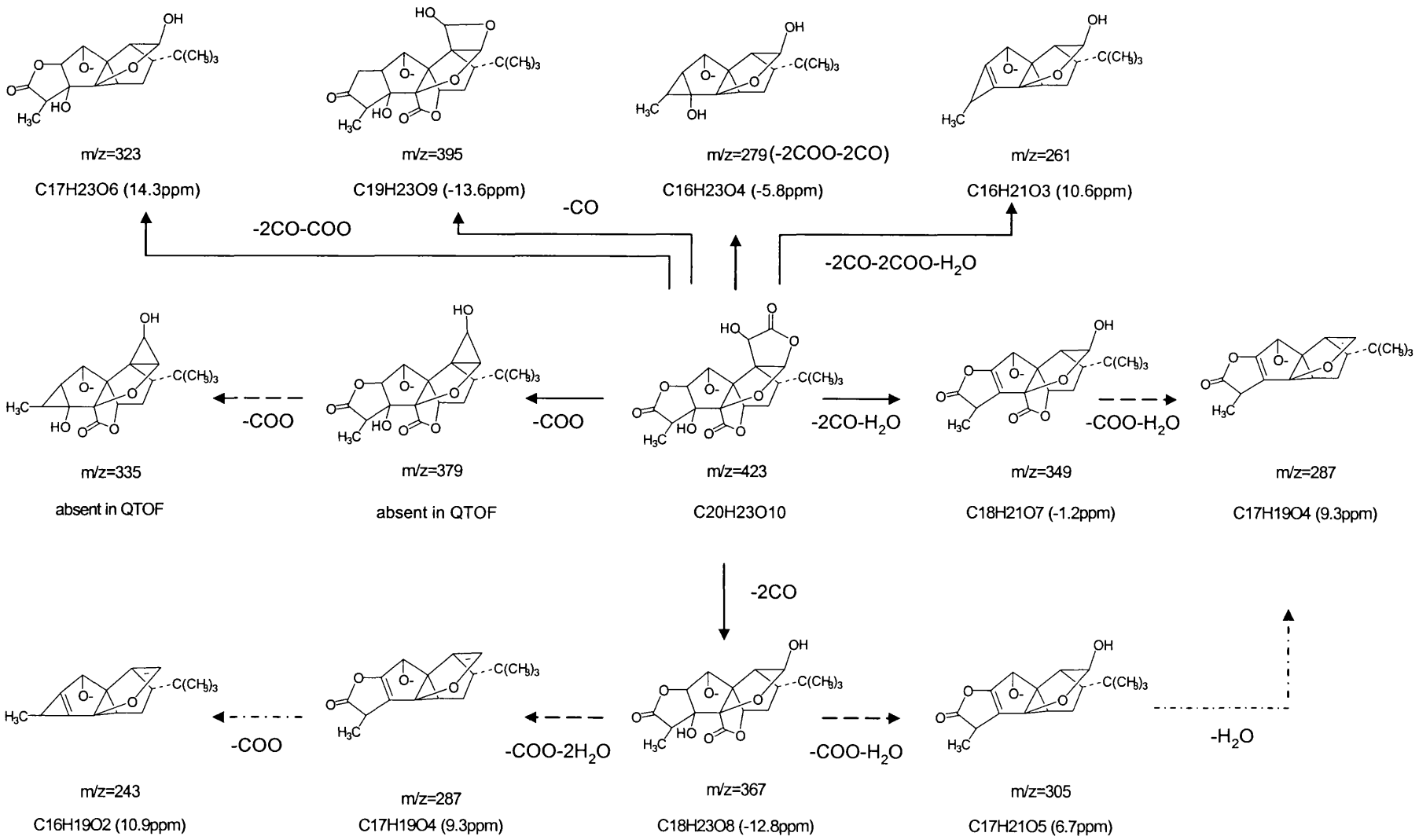
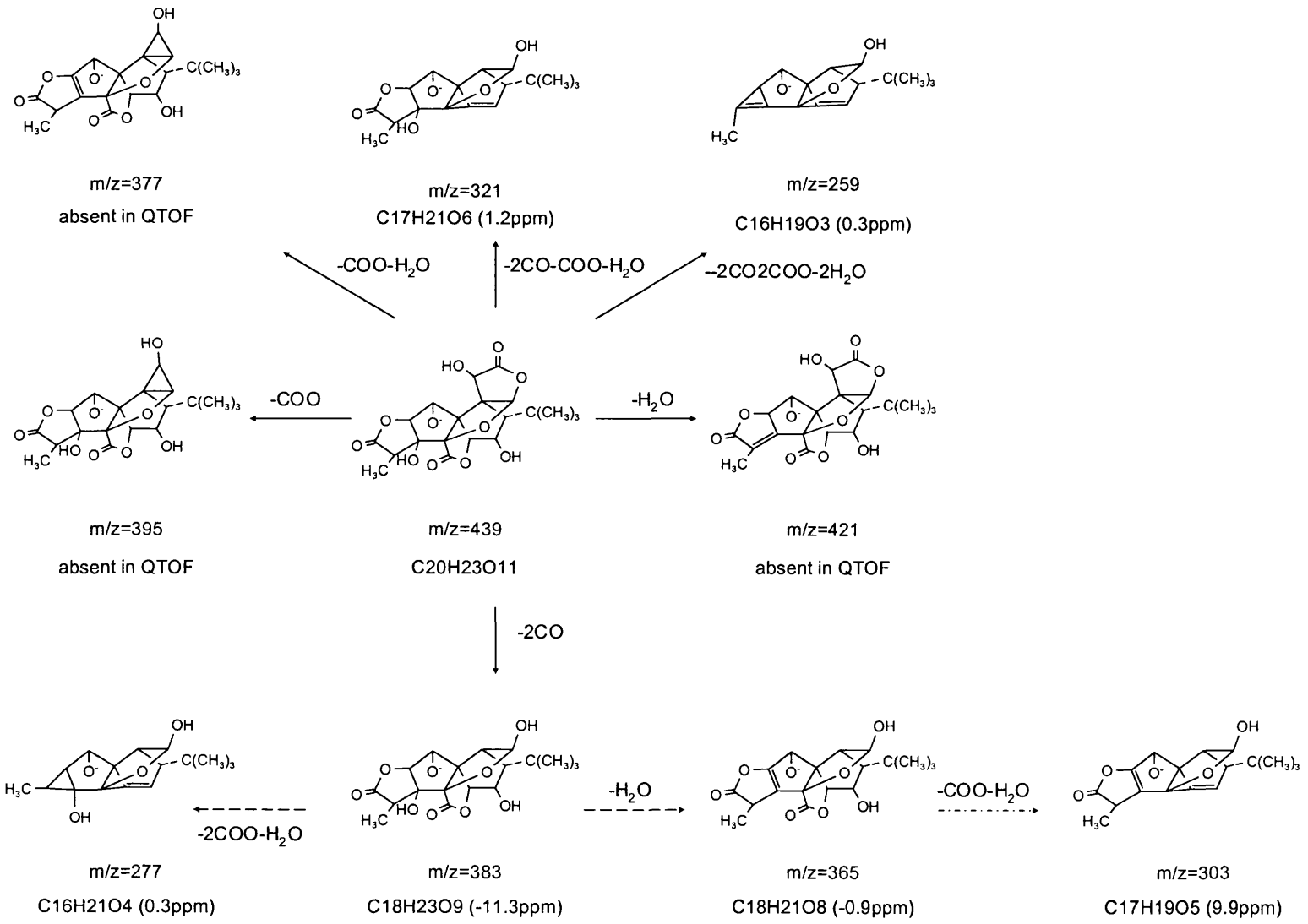


Figure 2.26 Proposed fragment pathway of ginkgolide C ($\xrightarrow{MS^2}$ $\dashrightarrow$ $\cdots\rightarrow$ mass accuracy in ppm shown in brackets)



2.4.4 Accurate mass in characterisation of fragmentation

A limitation of the ion trap is its mass accuracy which is not great enough to give the empirical formulae of the fragment ions, however Q-TOF instruments have a high mass accuracy capacity, and is able to list all the possible element composition of the fragment ions within a predefined ppm range. In this study, cCMP was chosen as the lock spray reference because all the *Ginkgo biloba* standards are in the same mass range. By using cCMP as the lock spray reference, the variations of all the mass of theoretical element compositions of the fragment ions and the mass recorded by mass spectrometer are generally between 0.3-20 ppm (as shown on the previous figures) and so the element composition provided by the Q-TOF gives greater confidence in the fragmentation pathways proposed.

The determination of the accurate mass of the fragment ions is of great benefit, especially when there is more than one possible pathway of fragmentation that could give an ion of a given *m/z* value but the resultant ions differ in empirical formulae and are not isomers. In our study, several such cases were encountered. One is observed in the fragmentation of quercetin, the product ion at *m/z* 179 could be attributed to either the $^{1,4}B^-$ ion Figure 2.27 (a) or the $^{1,2}A^-$ ion Figure 2.27 (b). Accurate mass analysis was therefore necessary (Table 2.7), the $^{1,2}A^-$ ion $C_8H_3O_5$ has a mass accuracy of 5.4 ppm while the next on the list -69.8 ppm, and it can therefore be concluded that the ion of *m/z* 179 is the $^{1,2}A^-$ ion of quercetin.

$C_9H_7O_4$ (a)

$C_8H_3O_5$ (b)

Figure 2.27 Possible fragment ions with *m/z* 179 in quercetin fragmentation

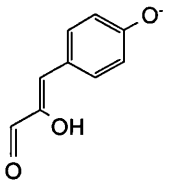
Table 2.7 Empirical formula of the *m/z* 179 peak by accurate mass analysis

Mass	Calc. Mass	mDa	PPM	DBE	Score	Formula
179.0008	178.9980	2.8	5.4	7.5	2	C8 H3 O5
	179.0133	-12.5	-69.8	11.5	4	C12H3O2
	179.0192	-18.4	-102.6	2.5	1	C5 H7O7

In kaempferol fragmentation, the fragment ion with *m/z* 163 can again be attributed to one of two possible formulae; the $^{1,4}B^-$ ion Figure2.28 (a) or the $^{0,2}A^-$ ion Figure2.28 (b). From Table 2.8, it can be concluded that ion with *m/z* 163 is $^{0,2}A^-$ ion of kaempferol, which proves that the characterisation of the *m/z* 163 ion as $^{1,2}A^-$ ion in the paper by Fabre *et al.* is wrong[22], actually, $^{1,2}A^-$ ions of both quercetin and kaempferol are found to be *m/z* 179.

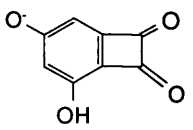
Table 2.8 Empirical formula of the *m/z* 163 peak by accurate mass analysis

Mass	Calc. Mass	mDa	PPM	DBE	Score	Formula
163.0104	163.0031	7.3	14.6	7.5	2	C8 H3 O4
	163.0184	-8.0	-49.0	11.5	4	C12 H3 O
	163.0395	-29.1	-178.6	6.5	1	C5 H7 O6



$C_9H_7O_3$

(a)

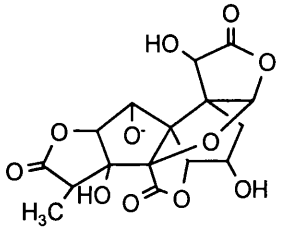


$C_8H_3O_4$

(b)

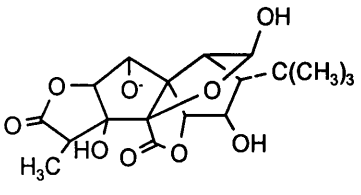
Figure 2.28 Possible fragment ion with *m/z* 163 in kaempferol fragmentation

In the study of terpene lactones similar examples of isobaric possible pathways were also found. In the fragmentation of GC, the fragment ion with *m/z* 383 could be attributed to the lost of side alkyl side chain from the precursor ion as in Figure 2.29(a), or the lost of two carboxyl group in the ring, as shown in Figure 2.29(b). The accurate mass of the ion is shown in Table 2.9, from which it can be concluded that the ion of *m/z* 383 is due to the loss of 2CO from the precursor ion, the same fragment pathway was observed in ginkgolide A and B.



$C_{16}H_{15}O_{11}$

(a)



$C_{18}H_{23}O_9$

(b)

Figure 2.29 Possible fragment ion with *m/z* 383 in ginkgolide C fragmentation

Table 2.9 Empirical formula of the *m/z* 383 peak by accurate mass analysis

Mass	Calc. Mass	mDa	PPM	DBE	Score	Formula		
383.1299	383.1342	-4.3	-11.3	7.5	2	C18	H23	O9
	383.0978	32.1	83.7	8.5	1	C17	H19	O10

Quadrupole time-of-flight (QTOF) mass spectrometry was reported to be able to produce valuable MS^3 and MS^4 data in addition to the usual MS^2 data[23], which is generally only obtained by ion trap and Fourier transform ion cyclotron resonance mass spectrometers. By making optimal use of in-source CID, product ions of the initial precursor ion can first be formed in the ESI source, hence MS^2 and even higher MS^n product ions become available for further fragmentation in the collision cell of the Q-TOF mass spectrometer. In this study, similar phenomena was observed and utilized in characterizing fragment pathway of reference compounds, for example, the precursor ion *m/z* 315 of flavonoid aglycone isorhamnetin, provided MS/MS data as shown in Figure 2.30, upon fragmentation under collision energy of 20 eV, which is similar to the MS/MS spectrum observed using an ion trap (Figure 2.14), the main fragment ion is *m/z* 300. When collision energy increased to 35 eV, the spectrum is shown in Figure 2.31, the further fragmentation of *m/z* 300 can be observed, and this spectrum is similar the MS^3 of isorhamnetin precursor ion obtained by ion trap mass spectrometry (Figure 2.15). Therefore it is important to optimizing collision energy when running MS/MS using Q-TOF to confirm the LCQ Data, By choosing suitable collision energy in the MS scan stage, Q-TOF mass spectrometry can obtain MS/MS spectrum similar to that of MS^3 or MS^4 as in ion trap.

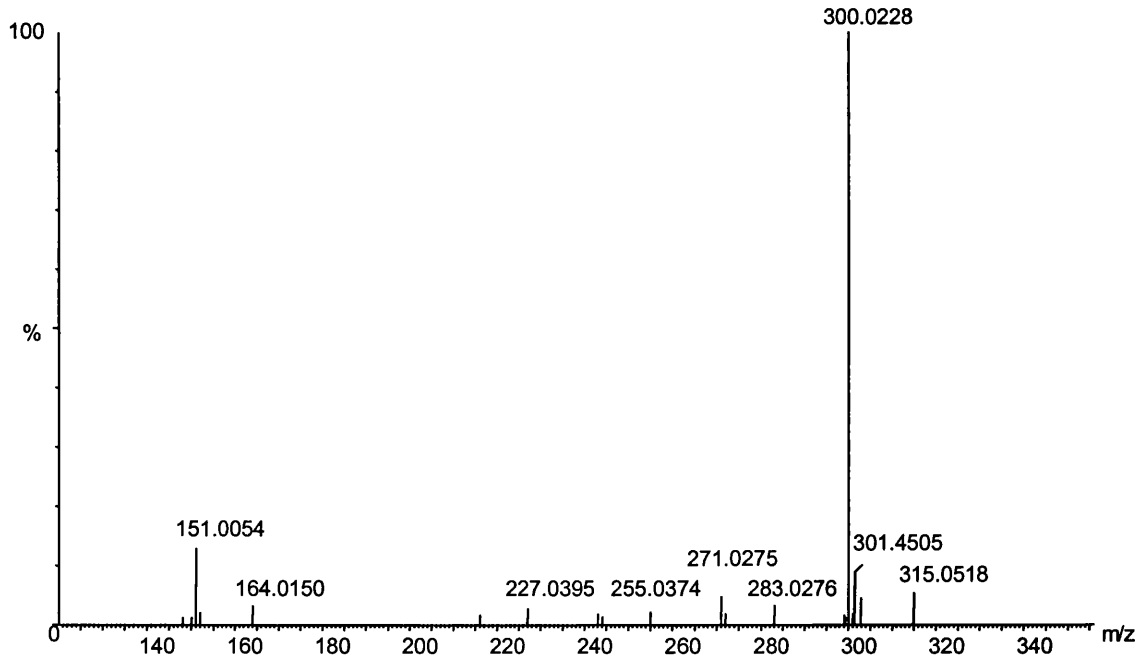


Figure 2.30 MS/MS of *m/z* 315 of isorhamnetin by Q-TOF mass spectrometry with collision energy 20 eV

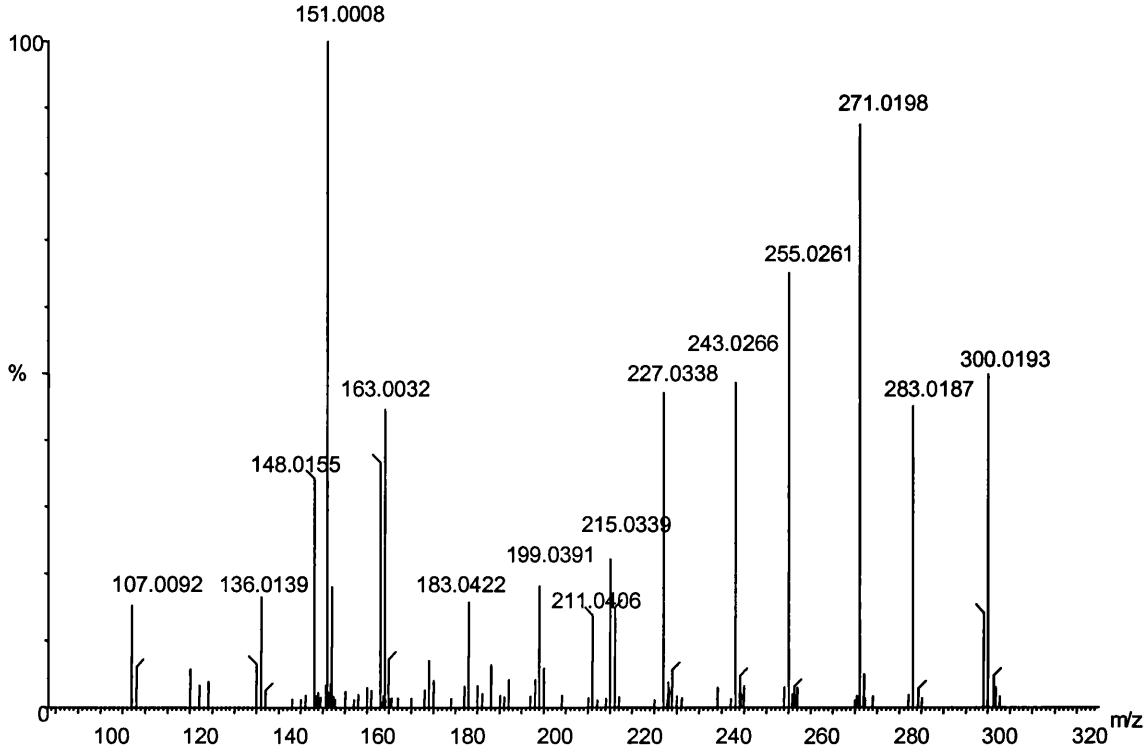


Figure 2.31 MS/MS of *m/z* 315 of isorhamnetin by Q-TOF mass spectrometry with collision energy 35 eV

2.5 Conclusion

Full scan mass spectrometry using ESI in negative mode shows very high sensitivity compared to analysis in positive ion mode. The LCQ ion trap has proved to be a powerful method for the study of the fragment pathway of flavonoids and terpene lactones since it can perform MS^n in a step-wise manner, where n can be as high as10 in theory. For flavonoids in this study MS^5 was the highest generation of fragmentation that could be accurately studied, which is very useful in structural analysis, however, the ion trap is not a high resolution mass spectrometer and it can only be used to obtain nominal mass. The Q-TOF compliments the ion trap because of its high mass accuracy. In this study, MS^n data obtained from ion trap and accurate mass measurement were combined to create a comprehensive fragment pathway of the flavonoids and terpene lactones in *Ginkgo biloba.* However, the fragmentation pathway of flavonoids is controversial, different pathways have been proposed by different authors[22, 24]. For the terpene lactones where multiple losses of CO_2 and CO occur, it is difficult to predict which section of the molecule is lost first. The combination of ion trap and accurate mass measurement gives the proposed fragment pathways a higher degree of confidence in comparison to other literature reports. To confirm the entire fragment pathways, more work such as isotope labels would need to be utilised in order to confirm these pathways however this was beyond the scope of the work detailed in this thesis. The MS/MS study performed here was required in order to understand the characteristic fragmentation pathways of the reference standards so that comparisons could be made with the MS/MS of unknown suspected flavonoids and terpene lactones in the later analysed *Ginkgo biloba* extract. In Chapter 4, a fingerprint profile of *Ginkgo biloba*

extract is obtained by LC/MS in data-dependent MS/MS scan mode and the knowledge of the fragmentation pathways of flavonoids and terpene lactones was applied to the determination of the structure of the active components in *Ginkgo biloba* presented in the fingerprint profile.

2.6 Reference

1. Perry LM, *Medicinal Plants of East and Southeast Asia: Attributed Properties and Uses*, MIT Press: Cambridge, MA ,(1984)

2. Foster, S. and Chongxi, Y., *Herbal Emissaries*, Healing Arts Press, Rochester, Vermont, USA, (1992)

3. O'Reilly, J. and Jaggy, H., European Patent 90123140.7 (03.12.90), (1990)

4. van Beek, T.A., J. chromatogr A, **967**, 21 (2002)

5. Chung, K.F., McCusker, M. and Page, C.P., Lancet, 248 (1987)

6. Jung, F., Mrowietz, C., Kieswetter, H. and Wenzel, E., Arzneimittelforschung **40**, 589 (1990)

7. Mouren, X., Caillard, P. and Schwartz, F., Angiology **45**, 413 (1994)

8. Le Bars, P.L., Katz, M.M., Berman, N., Itil, T.M., Freedman, A.M. and Schatzberg, F.A., J. Am. Med. Assoc., **278**,1327 (1997)

9. Allain, H., Raoul, P., Lieury, A. and LeCoz, F., Clinical Therapeutics **15**, 549 (1993)

10. Le March, L., Biomed. Pharmacother, **56**, 296 (2002)

11. Graefe, E.U., Derendorf, H. and Veit, M., Int. J. Clin. Pharmacol. Ther., **39**, 219 (1999)

12. Pietta, P., Mauri, P. and Rava, A., J. Pharm. Biomed. Anal, **10**, 1077 (1992)

13. Nakanishi, K., Pure Appl. Chem., **14**, 89 (1967)

14. Okabe, K., Yamada, K., Yamamura, S. and Takada, S., J. Chem. Soc., 2201 (1967)

15. Hedin, P.A. and Philips, V.A., J. Agric. Food Chem., **40**, 607 (1992)

16. Baracco, A., Bertin, G., Gnocco, E., Legorati, M., Sedocco, S., Catinella, S., Favretto, D. and Traldi, P., Rapid Commun. Mass Spectrom., **9**, 427 (1995)

17. de Koster, C.G., Heerma, W., Dijkstra, G. and Neimann, G.J., Biomed. Mass Spectrom., **12**, 596 (1985)

18. Li, Q.M. and Claeys, M., Biol. Mass Spectrom., **23**, 406 (1994)

19. Cunniff, J., Tiller, P., Harvey, M. and Land, A., Structural Determination of Flavonoids using MS^n, technical articles, Thermo Finnigan, Milford, MA.

20. Hvattum, E. and Ekeberg, D., J. Mass Spectrom., **38**, 43 (2003)

21. Ma, Y.L., Li, Q.M., Van den Heuvel, H. and Claeys, M., Rapid Commun. Mass Spectrom., **11**, 1357 (1997)

22. Fabre, N. and Rustan, I, J. Am. Mass Spectrum., **12**, 707 (2001)

23. Dudley, E., Tuytten, R., Bond, A., Lemiere, F., Brenton, A.G., Esman, E.L. and Newton, R.P., Rapid Commun. Mass Spectrom., **19**, 3075 (2005)

24. Chen, L., Zhao, X., Plummer, S., Tang. J., Games, D.E., J. Chromatogr. A, **1082**, 60 (2005)

Chapter 3

Quantitative determination of the active components in *Ginkgo biloba* extract nutritional supplements by LC/MS

3.1 Introduction

3.1.1 Quantitative Analysis: Precision, accuracy and sensitivity

Quantitative analysis is the determination of the concentration of components in the sample. HPLC and mass spectrometry can perform the quantitative determination of components. Method validation is very important in quantitative analysis since it is the process of proving that an analytical method is acceptable for its intended purpose. Method validation should be evaluated in terms of precision, accuracy and sensitivity.

Precision is the closeness of agreement or degree of scatter between a series of measurements obtained from multiple samplings of the same homogeneous sample[1]. A more comprehensive definition was proposed by the International Conference on Harmonization (ICH)[2]. It can mainly be considered from the following two aspects: Repeatability expresses the precision under the same operating conditions over a short interval of time. This is pure instrumental precision; it can be measured by the sequential, repetitive injection of the same sample 10 or more times, followed by the averaging of the measured values and determination of the relative standard deviation (RSD) of all measurement. It can be termed as intra-assay precision. %RSD (%relative standard deviation) can be calculated by the following equation:

$$\%RSD = 100 \times \frac{SD.}{Average} \qquad \text{(Eq. 3.1)}$$

Reproducibility evaluates the reproducibility of the whole analytical method, including sample extraction, sample clear-up and instrumental behavior as well. The

reproducibility of an LC/MS method is usually described by the coefficient of variation (CV)[3].

Accuracy of a method is the closeness of the measured value to the true (or expected) value for the sample. Accuracy is normally evaluated by recovery studies. There are different ways to determine the recovery. 1) comparison to a reference standard, 2) recovery of the standard spiked into a blank matrix, 3) standard addition of the analyte. For the quantitation of herbs, since it is not possible to prepare a blank sample matrix without the presence of the analyte, standard addition method is most often used. In this study, three samples of *Ginkgo biloba* nutritional supplements, each spiked with known quantities of reference standards in low, middle and high concentrations were extracted according to the same extraction method. The components in these samples were determined by LC/MS, the ratio of the amount of the standard obtained from the experiment to the spiked amount was the recovery of the extraction.

$$\%\,\mathrm{Recovery} = 100 \times \left(\frac{Calculated}{True\ value} \right) \qquad \text{(Eq. 3.2)}$$

Sensitivity is another important factor when evaluating a method, and is often described by the limit of detection (LOD) and limit of quantitation (LOQ). Limit of detection of an analyte is the concentration which gives an instrumental signal significantly different from the background signal (defined as the analyte concentration giving a signal equal to the blank matrix signal, y_B, plus three standard deviation of the blank matrix, s_B, i.e., (y_B+3s_B)[3]) and is often estimated by a signal-to-noise ratio of 3:1. Limit of quantitation is the minimum amount of analyte in a sample which give a response that

can be quantified with suitable accuracy and precision. It is estimated by a signal-to-noise ratio of 10:1[3].

3.1.2 Quantitation of the active components in *Ginkgo biloba* extract food supplement: literature review

Commercial *Ginkgo biloba* products are usually standardized mixtures, based on the content of flavonoids and terpene lactones. Many investigations have been carried out on commercial preparations. Normally, these two classes of compounds are measured separately. There are many analytical methods which can be used for the determination of *Ginkgo* flavonoids. Older methodologies for the assay of flavonoids include spectrophotometry, fluorometry, paper chromatography and thin layer chromatography[4]. Routine analysis of flavonoids is currently conducted by high-performance liquid chromatography with ultra-violet detection (HPLC-UV). It was reported that flavonoids occur in *Ginkgo* leaves and extracts as many different flavonoid glycosides, most of them are derivatives of quercetin, kaempferol and isorhamnetin. The aglycones themselves occur only in relatively low concentration. Hasler and coworkers published a reversed phase separation of 33 flavonoid glycosides, flavonoids and biflavones using a tertiary gradient system[5]. Unfortunately due to a lack of commercially available reference compounds no quantitation of all the individual glycosides is possible. The normal procedure for the quantitative analysis of flavonoid glycosides and aglycones in *Ginkgo* leaves and extracts is an acidic hydrolysis followed by a reverse phase liquid chromatography of the resulting aglycones[5]. Because only three flavonoid aglycones occur in significant concentrations, this procedure greatly simplifies the separation and facilitates analysis, the content of the flavonoid glycosides can be recalculated from

measurement of the flavonoid aglycones. However, there is some limitation in this procedure, for example, this method cannot differentiate between the fortified flavonoids (aglycones or glycosides) and the intact flavonoids that originate from the *Ginkgo* plant. The price of quercetin and rutin is very low and they are easily available; there is a possibility that some *Ginkgo* products are fortified using these low cost ingredients. Differentiation between the flavonoid aglycones and glycosides is important because bioavailability, pharmacodynamics and pharmacokinetics of intact flavonol glycosides and the hydrolysis products are different[6].

It has previously been difficult to develop a suitably fast and robust LC-UV method for the analysis of *Ginkgo* terpene lactones since they are poor chromophores with very weak absorption in the UV range, even trace impurities interfere with the detection of these compounds by UV[7]. Furthermore, tedious pre-purification procedures of the samples have been necessary in order to separate the terpene trilactones from compounds which would interfere in their detection[8, 9]. The quantification of the non-UV active *Ginkgo* constituents requires methods that are more selective. Gas chromatography coupled with a flame ionisation detector (GC/FID) and mass spectrometer equipped with an electron impact interface (GC-EI/MS) has been explored[10, 11], but these methods are rather time-consuming because derivatization of the sample is required. The application of HPLC with evaporative light scattering detection (ELSD) and refractive index (RI) detection provided a satisfactory determination of the ginkgolides and bilobalide[9, 12]. Advantages of RI over ELSD are its larger linear range, lower costs and its broader availability. Advantages of ELSD over RI are better baseline, compatibility with THF and gradients, small solvent peak and greater sensitivity. Both methods are suitable for the routine analysis of all terpene trilactones

after an RP-HPLC separation, but these methods also required a long process of sample preparation. HPLC-RI is currently the most widely used quality control method. Supercritical fluid chromatography (SFC) combined with ELSD detection has proved to be an interesting alternative to HPLC and efficient separation of the terpene lactones can be achieved in 10 minutes in the isocratic mode[13]. Quantitative NMR is another method proposed for the quantification[14].

Efforts were concentrated on simultaneous determination of the flavonoids and terpene lactones in *Ginkgo biloba* preparations. ELSD[15] and GC/MS[11] have been used to determine terpene lactones and flavonoid aglycones in *Ginkgo biloba* extract, these two methods include time-consuming extraction or derivatization procedure. Another problem is that in these two methods the flavonoid glycosides were hydrolysed into flavonoid aglycones before quantitation, so the information of flavonoid glycosides, which is important to evaluate the product quality and storage conditions, can not be obtained. Mass spectrometry is currently the most sensitive and selective analytical method for the rapid qualitative and quantitative analysis of known compounds as well as the identification of unknown compounds from crude and partially purified samples of natural supplements[16] and has been applied to the analysis of *Ginkgo biloba*, mostly for the determination of *Ginkgo* terpene lactones[17, 18]. Its unique ability to filter and isolate molecular ions with specific mass-to-charge (*m/z*) ratios from a complex mixture makes MS an invaluable tool for analytic chemistry. Recently there were reports on simultaneously detection of active components, terpene lactones and intact flavonoid glycosides in *Ginkgo* products by nanoelectrospray[19], but it is more like a fingerprint profile, no quantitative results of *Ginkgo* flavonoid glycosides and terpene lactones were presented. In this chapter, an integrated approach consisting of HPLC, LC/MS has

been used for the quantification of active components in *Ginkgo biloba* nutritional supplements. Simultaneous quantitative determination of terpene lactones, flavonoid aglycones and three intact flavonoid glycosides were obtained with minimum sample pre-treatment and Optimisation of the extraction and rapid separation method was investigated.

3.1.3 Aims of study

Ginkgo biloba's pharmacological activity has been linked to the two groups of compounds, flavonoids and terpene lactones, as introduced previously, most of the commercial extracts are standardized and claimed to contain no less than 24% of *Ginkgo* flavonoid glycosides and no less than 6% of *Ginkgo* terpene lactones. Due to the great difference in concentration and chromatographic property, flavonoids and terpene lactones in *Ginkgo biloba* are determined separately[15]. Most studies have quantified the flavonoid content of *Ginkgo biloba* in terms of the quercetin, kaempferol and isorhamnetin content, following an acid hydrolysis process[20]. More information is needed regarding the flavonoid glycoside conjugates and aglycones content. It was reported that the content of flavonoid glycosides and flavonoid aglycones will vary according to the season of harvest. Also, the different proportion of flavonoid glycosides and aglycones suggests different storage condition, or degradation of the product. More importantly, it may indicate fortified chemical preparation[19].

The principal objective of this chapter is to develop a routine quantitative method to determine quercetin, kaempferol or isorhamnetin, quercetin-3-β-D-glucoside, quercetin-3-rhamnoside, rutin, ginkgolide A, ginkgolide B, ginkgolide C and bilobalide

in *Ginkgo* nutrition supplements. A method for the quantitative determination of 10 active components in *Ginkgo biloba* was successfully developed by on-line LC/ESI-MS. The extraction and hydrolysis methods for the isolation of the flavonoids and terpene lactones from *Ginkgo biloba* were also studied.

3.2 Experimental

3.2.1 Chemicals and standards

Methanol, acetonitrile (HPLC grade solvents), formic acid and acetic acid were purchased from Fisher chemicals (Loughborough, UK) and were used without further purification. Ginkgolide A (**GA**), ginkgolide B (**GB**), ginkgolide C (**GC**), bilobalide (**BL**), quercetin dehydrate (**QD**), quercetin-3-β-D-glucoside (**QG**), quercitrin (**QH**), kaempferol (**KF**), isorhamnetin (**IR**), rutin (**RH**) and andrographolide were purchased from Sigma (St. Louis, MO, USA). Water was purified with a Milli-Q deionisation unit (Millipore, Bedford, MA, USA). Gases used included oxygen free nitrogen and helium which were purchased from BOC Ltd (Surrey, UK).

3.2.2 Standard stock solution and calibration solutions

Stock solutions were prepared (1 mg/5 mL methanol) to give a final concentration of 200 μg/mL and solutions were then used to prepare working standards for calibration curves and recovery experiments with *Ginkgo* extract samples. All solutions were placed in an ultrasonic bath for 10 minutes to ensure they were completely dissolved. The calibration solutions (working solutions) were accurately diluted with methanol just prior to use. All solutions were stored at -20 °C.

3.2.3 Sample preparation

Samples of commercial *Ginkgo* extract products were prepared by combining the contents of 10 capsules or 10 tablets pulverized into powder. 100 mg of *Ginkgo biloba* nutrition supplement were accurately weighed into a 20 mL vial. 20 mL of methanol were added to the sample and shaken briefly to mix. The sample vial was sonicated in an ultrasonic bath Bransonic 2510 (Branson, Danbury, CT) at a frequency of 42 kHz at 25 °C for 50 minutes, 1 mL of sample was then centrifuged for 10 minutes at 17,000 g to pellet insoluble material. 100 μL of the supernatant were removed into a sample vial and 10 μL of it were analyzed by HPLC/MS. The extraction procedure presented here minimized handling of the samples during preparation. The sample was extracted directly by methanol and transferred to the HPLC system without an evaporation procedure, solid-phase extraction (SPE) or liquid-liquid extraction (LLE) clean-up procedure.

3.2.4 Traditional QC (quality control) analysis

500 mg of *Ginkgo* product powder were accurately weighed and refluxed with acidified methanol for 5 hours, then 10 μL of the resulting solution were injected onto LC system, as below.

3.2.5 LC/MS methodology

The extracts were analysed using an LC/MS system that consisted of a Hewlett Packard 1100 HPLC system (Hewlett-Packard, Wilmington, DE, USA) with a HP autosampler, gradient pump, and a HP 1100 photodiode-array detection for HPLC separation

interfaced to a LCQ ion trap mass spectrometer (Finnigan, Hemel Hempstead, UK). The separation was achieved with a reverse phase analytical column, Luna C_{18} RP column (100×4.6 mm, 5 μm, Phenomenex, Torrance, CA, USA), using a gradient elution at a temperature of 35 °C. UV detection at 250 nm was applied. The mobile phase was composed of 0.1 %(v/v) acetic acid in water (A), 1:1(v/v) mixture of ACN and methanol (B). The flow rate was 0.8 mL/min, and a splitter was used to transfer only $\frac{1}{4}$ of the flow into the mass spectrometer, the gradient is shown in Table 3.1. During full scan MS mode analysis LC/ESI/MS was carried out in the negative ion mode from *m/z* 200-800 with electrospray ionisation for quantitation. The heated capillary temperature was heated to 190 °C and the electrospray voltage was 4.5 kV, sheath gas flow rate was 90 arbitrary units, and auxiliary gas flow was 15 arbitrary units. The system was calibrated according to manufacturer instructions and was tuned using the calibration solution. The ESI parameters were optimised by flow injection of standard solutions. The ions monitored in SIM and SRM mode is shown in Table 3.2. During SIM mode a mass window of ± 0.5 Da was used in order to specify the ion monitored and during SRM scanning mode the same window was utilized for both the precursor and product ion selection.

Table 3.1 Gradient elution program used from HPLC/MS analysis

Time (min)	A%	B%	Flow rate (mL/min)
0	80	20	0.8
50	55	45	0.8
52	0	100	0.8
57	0	100	0.8
60	80	20	0.8
70	80	20	0.8

Table 3.2 Ions monitored in SIM and SRM detection mode

Component	SIM (*m/z*)	SRM (*m/z*)	Relative collision energy (%)
BL	325	325→ 183	50
GC	439	439→ 377	26
RH	609	609→ 301	28
QG	463	463→ 301	28
QH	447	447→ 301	27
GA	407	407→ 363	28
GB	423	423→ 349	25
QD	301	301→ 179	36
KF	285	285→ 257	46
IR	315	315→ 300	34

3.2.6 Data analysis

Xcalibur ver. 1.2 software (Thermo Electron, San Jose, USA) was used for data acquisition and analysis. Microsoft Excel (ver. 5.0) was used for statistical calculations. Data are expressed as means ± SD (standard deviation of the mean). Linear regression analysis using the least squares method was used to evaluate the calibration curve of each analyte as a function of its concentration.

3.2.7 Reproducibility

Measurement of intra- and inter-day variability was utilized to determine the precision of the method. An extracted sample was analysed to determine the intra-day repeatability (examined in one day) and inter-day repeatability (determined over

3 consecutive days). The relative standard deviation (RSD) was calculated as a measurement of method reproducibility.

3.2.8 Recovery

The accuracy of the method was examined by using the standard addition method for recovery studies. *Ginkgo biloba* commercial products were spiked with three different amounts of standards. The spiked samples and unspiked sample were assayed using the same method and results expressed as mean recovery ±SD. The recoveries were determined by calculating the concentration difference between unspiked and spiked results and comparing the data to the spiked levels.

3.3 Method development

3.3.1 Optimisation of chromatographic conditions

HPLC separation was achieved with a C_{18} reverse phase analytical column (100×4.6 mm, 5 μm, Phenomenex), using a gradient elution at a temperature of 35 °C. UV detection at 250 nm was applied. Figure 3.1 shows the HPLC separation from a mixed gradient elution program (Table 3.1) of a mixture solution of ten standards, the six flavonoids can be separated, but no UV absorbance of terpene lactones can be observed. To optimise the LC separation, the mass spectrometer needed to be used in order to observe the elution of terpene lactones.

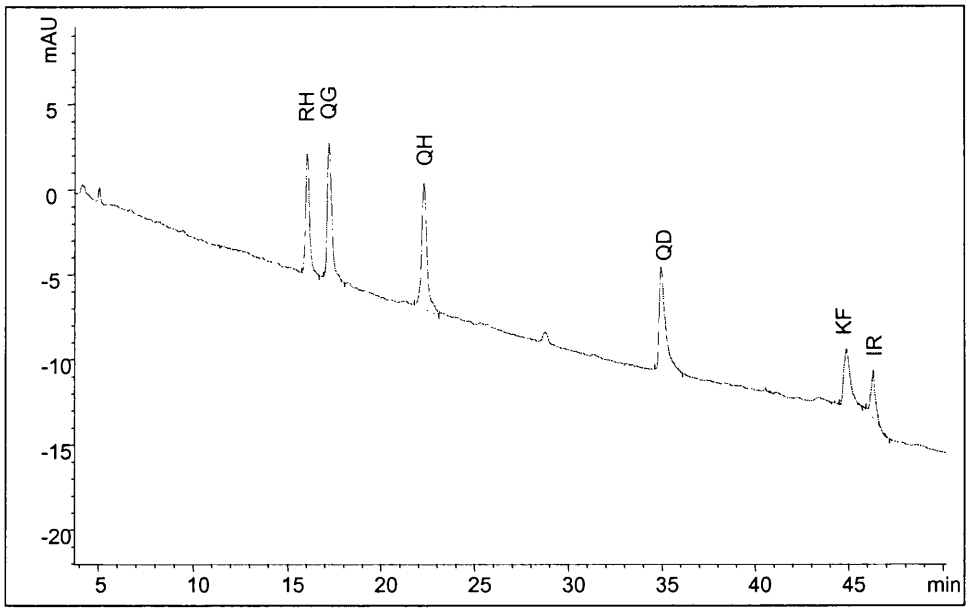


Figure 3.1 LC chromatogram of a mixture solution of 10 standard components, only 6 *Ginkgo* flavonoids can be observed, the 4 terpene lactones show no UV absorbance Experiment conditions: Luna C_{18} 100×4.6 mm, 5 μm; mobile phase A: 0.1 %HAC in H2O, B : 0.1% HAC in ACN; gradient see Table 3.1

Phosphate buffers are often used in the HPLC analysis of *Ginkgo* flavonoids to adjust the pH of mobile phase; however, these buffer solutions are not compatible with HPLC/MS on-line separation. In this case, the flavonoids and terpene lactones are basic compounds, acid needs to be added to adjust the pH of the mobile phase. For HPLC separation (column used in this study was a reverse C_{18} column), the compounds need to be neutralized so that they can equilibrate with the solid phase of the column and obtain good separation. However, for mass spectrometry, the compounds need to be ionised so that they can be detected by the mass spectrometer. Formic acid, acetic acid and TFA at different concentrations were compared for the sensitivity of ionisation, 0.1% acetic acid showed best separation on the HPLC column and good mass spectrometry sensitivity.

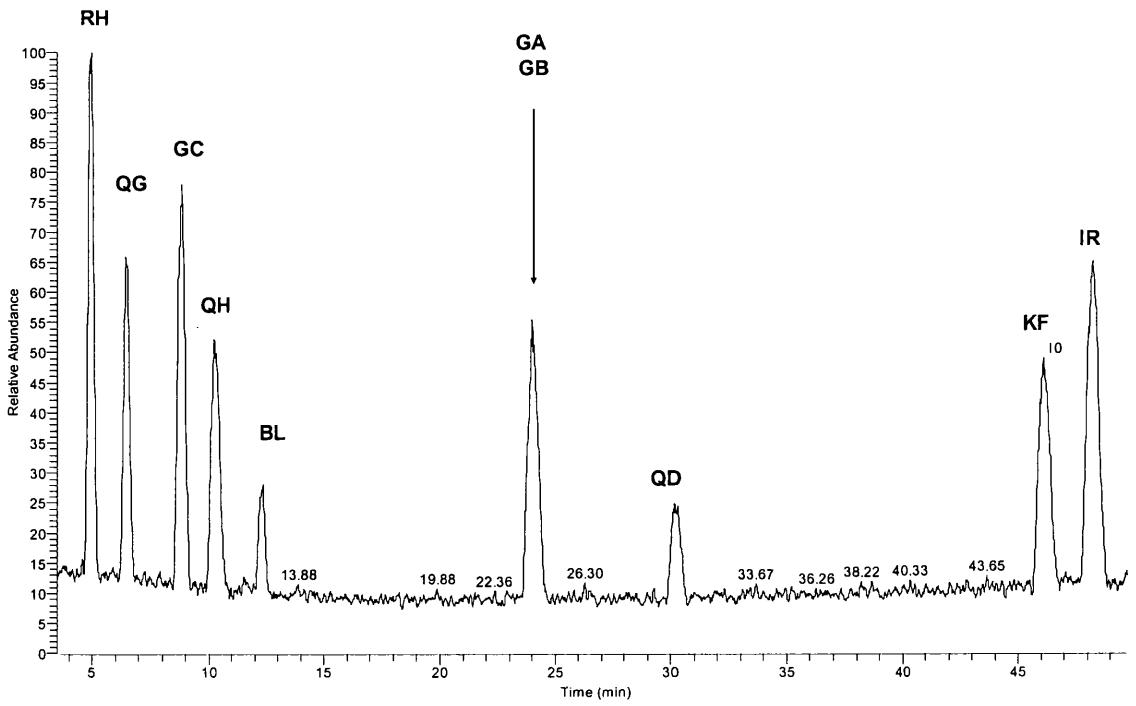


Figure 3.2 LC/MS chromatogram of a mixture solution of 10 standard components using ACN as mobile phase. Mobile phase: 0.1% acetic acid (A), ACN (B); gradient: 20% B for 5 min, 20-30% Bin for 40 min, 30-35% B for 10 min

A different mobile phase was evaluated for the separation of the standard mixture. Acetonitrile has a lower UV cut-off, lower viscosity and better mass transfer, while methanol is less expensive and more environmentally friendly. Both acetonitrile and methanol were compared for the separation. Figure 3.2 shows an LC/MS chromatogram of a mixture solution of 10 standard components, by using acetonitrile as mobile phase B. Flavonoids can be separated but ginkgolide A and ginkgolide B co-eluted at 24.02 minutes. The performance of methanol as a mobile phase is shown in Figure 3.3. By using methanol as the mobile phase, ginkgolide A and ginkgolide B were separated quite well at 26.93 and 29.14 minutes respectively, but rutin hydrate and

quercetin-3-β-D-glucoside co-eluted at 36.75 minutes and the separation took a long time whilst kaempferol and isorhamnetin were not eluted from the column after 1hr.

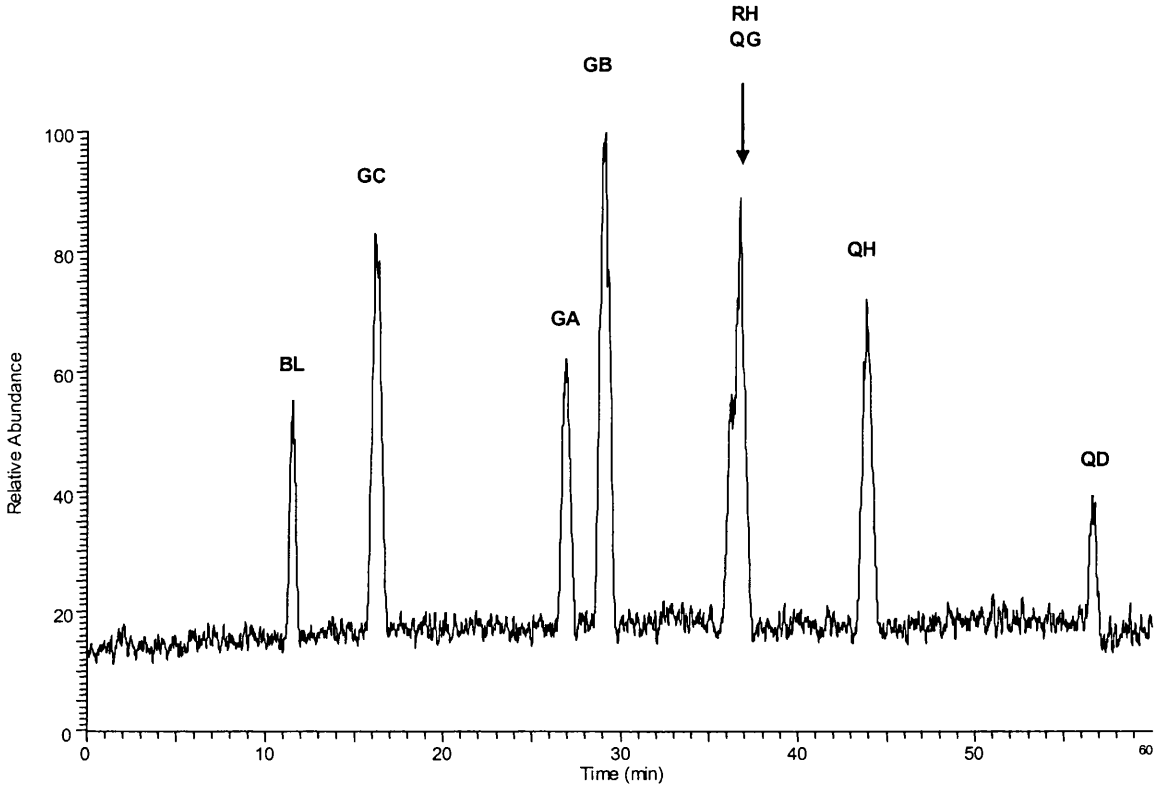


Figure 3.3 LC/MS chromatogram of a mixture solution of 10 standard components using methanol as mobile phase. Mobile phase 0.1% acetic acid (A), methanol (B); gradient: 25-50% B in 60 minutes

This clearly shows the difficulty in utilizing a single analysis for the determination of compounds of varied polarity such as these, however, good separation is very important to minimize the ion suppression during the electrospray process. Since the HPLC system has no tertiary pump, different ratios of ACN and methanol were used for the separation of the standards, it was found that 1:1 of ACN and methanol gives best separation. Figure 3.4 shows the total ion chromatogram (TIC) of the separation of a

mixture, in solution, of 10 standard components under the stated conditions (Section 3.2.5) and shows clearly that the 1:1 ratio provides good resolution for all ten standards. All components can obtain baseline separation in less than 50 minutes.

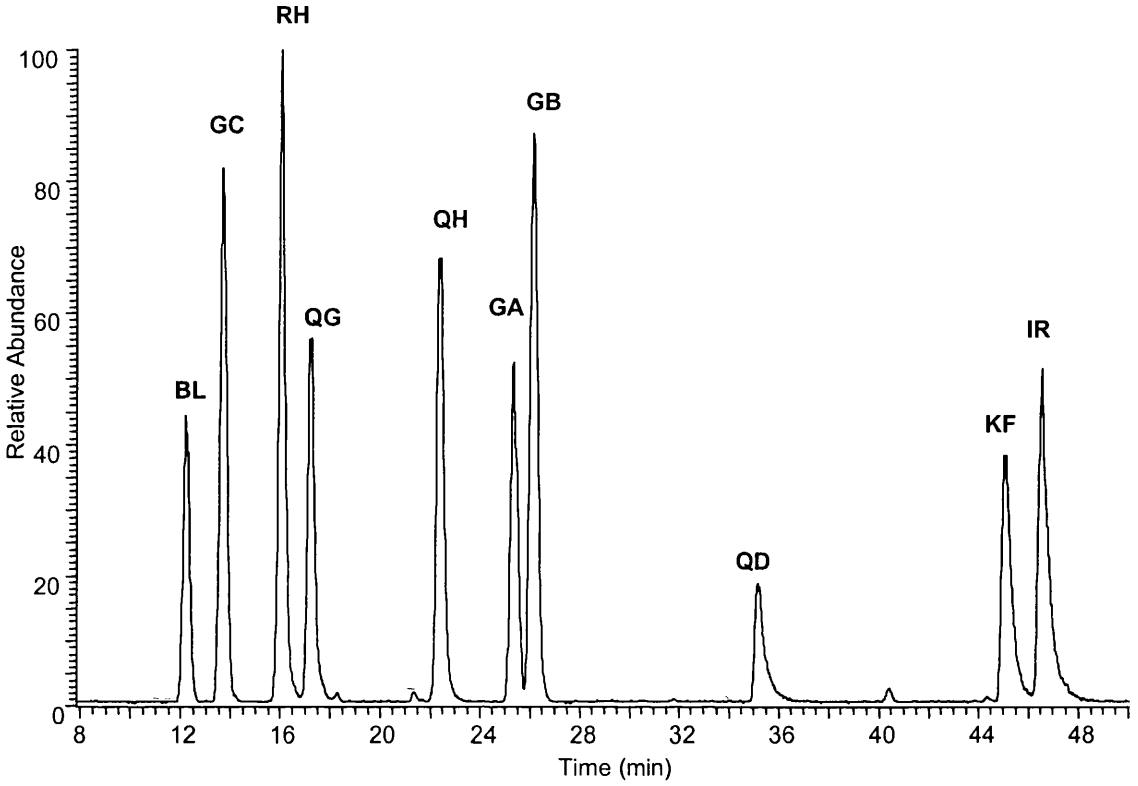


Figure 3.4 LC/MS chromatogram of a mixture solution of 10 standard components using a 1:1 mixture of acetonitrile and methanol as mobile phase B. Experiment conditions: reverse column, Luna C_{18} 100×4.6 mm, 5 μm; gradient see Table 3.1

3.3.2 Optimisation of mass spectrometric conditions

3.3.2.1 Calibration and tuning of the LCQ ion trap

Mass calibration of the instrument is very important and allows the MS detector to assign the correct mass values to the ion signals that it detects. The basic process of

mass calibration involves the acquisition of a data file (mass spectrum) using a standard mass calibration compounds. The data file is then compared with a mass calibration file of that compound, which has the correct mass assigned to each peak. Any difference between these two files is adjusted to bring the new data file into the line with the mass calibration file. This adjustment is then applied to all subsequent data files acquired. This process is called mass calibration. The LCQ ESI tuning and calibration solution contains caffeine, MRFA, Ultramark 1621 in 50:50 methanol: water containing 1% acetic acid. Caffeine provides an ESI singly charged peak at *m/z* 195.2. MRFA (L-methionyl-arginyl-phenylananyl-alanine acetate) provides an ESI singly charged peak at *m/z* 524.3. Ultramark provides ESI singly charged peaks at *m/z* 1022.1, 1122.1, 1222.1, 1322.1, 1422.1, 1522.1, 1622.1, 1722.1, 1822.1 and 1921.9, as shown in Figure 3.5.

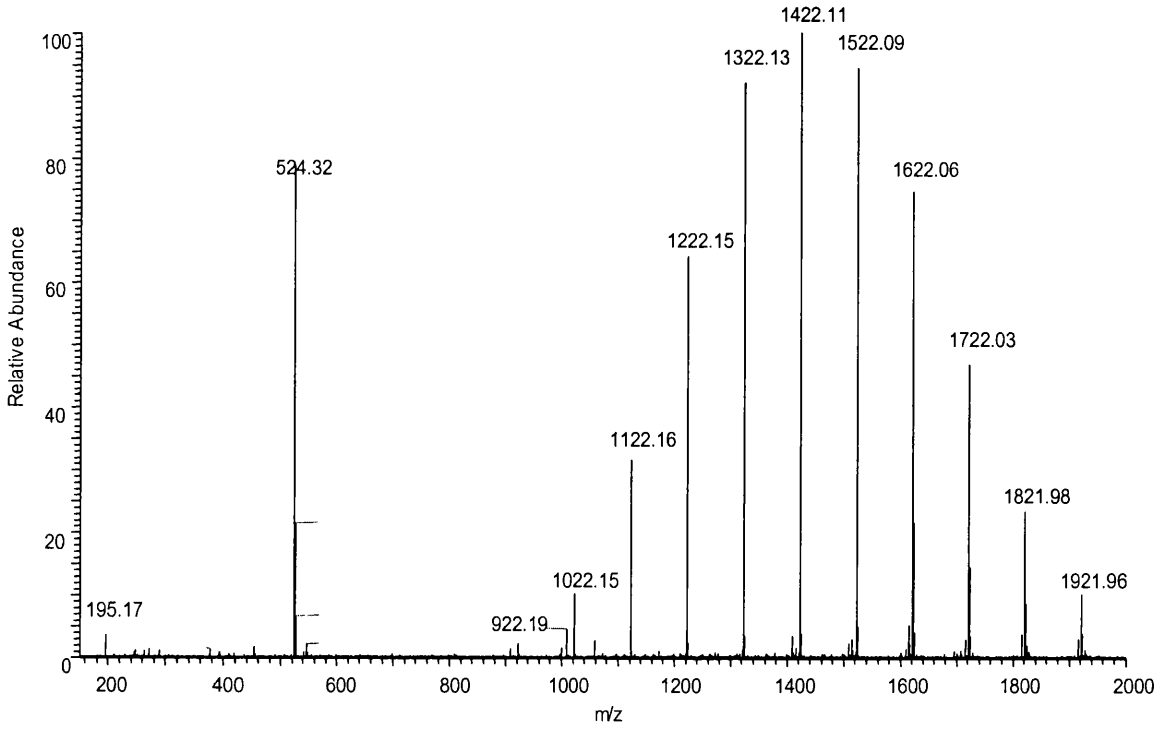


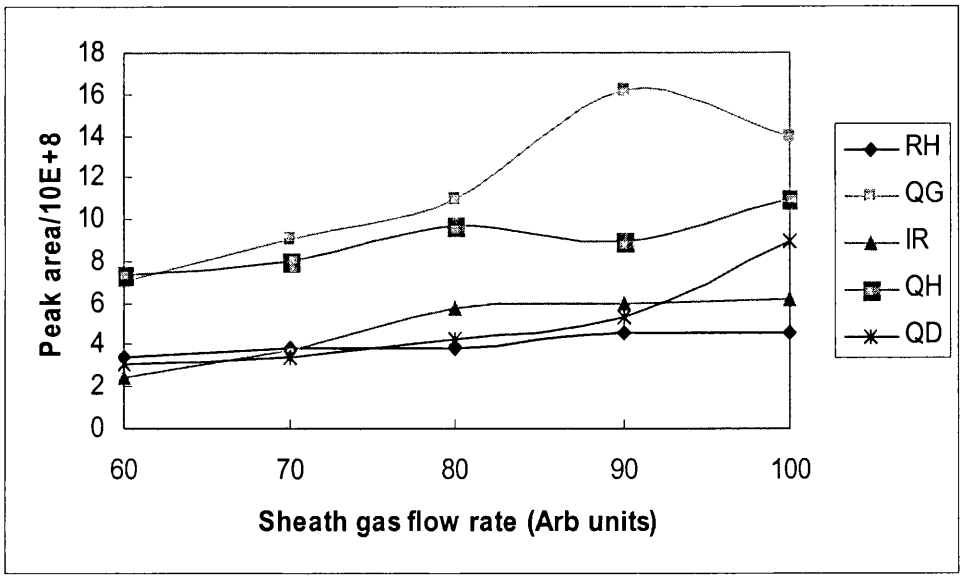
Figure 3.5 Mass spectrum of caffeine, MRFA and Ultramark 1621 tune solution

For better sensitivity of mass spectromic analysis, the "tune files" for the reference standards were set up, in which the variable parameters within the mass spectrometer are optimised for the analysis of the protonated or deprotonated molecule of the compound being tuned for. In order to get the best sensitivity for an LC/MS study, some important mass spectrometric parameters were optimised for an LC flow rate of 0.8 mL/min with a split flow of 0.2 mL/min to mass spectrometer. Since the study in Chapter 2 shows negative ionisation gives better sensitivity of molecular ion species $[M-H]^-$, the following Optimisation was carried in negative ionisation mode.

3.3.2.2 Optimisation of the sheath gas and auxiliary gas levels

The first stage of Optimisation was the adjustment of gas levels applied to the electrospray source in order to study the effect of different gas flows upon the ionisation of the compounds of interest. Sheath gas flow rate was altered between 60 and 100 arbitrary units, keeping all other parameters constant. The responses of *Ginkgo* analytes showed significant dependence upon the sheath gas flow rate, as shown in Figures 3.6 (a) and (b). All compounds showed an increased response in signal intensity as the sheath gas level was increased to 90 arbitrary units, except kaempferol which has maximum signal intensity at 80 arbitrary units. A further increase of sheath gas flow rate to 100 arbitrary units was observed to weaken the signal for 8 of 10 compounds, especially for QG and GA. Therefore for subsequent experimentation the sheath gas flow was set to 90 arbitrary units for LC/MS analysis.

(a)



(b)

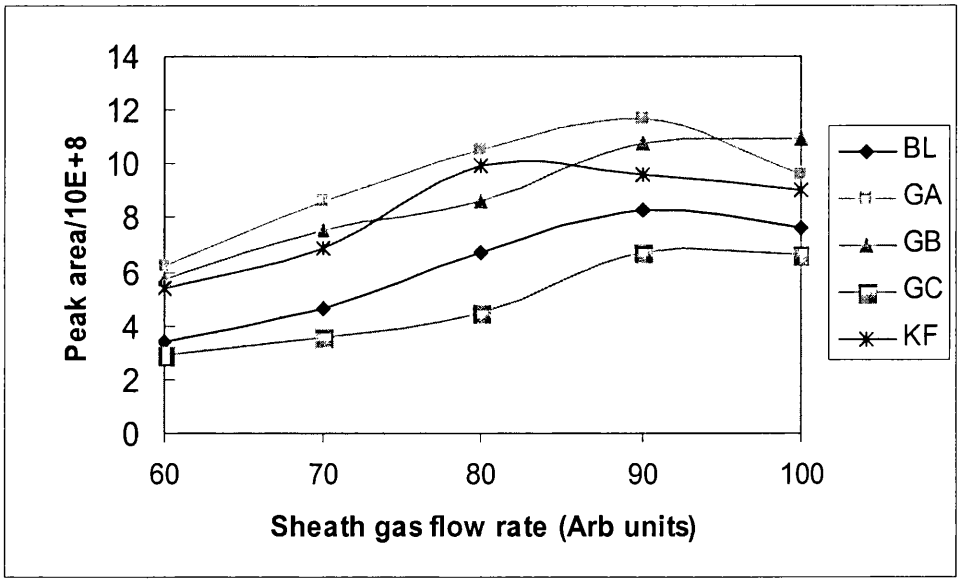
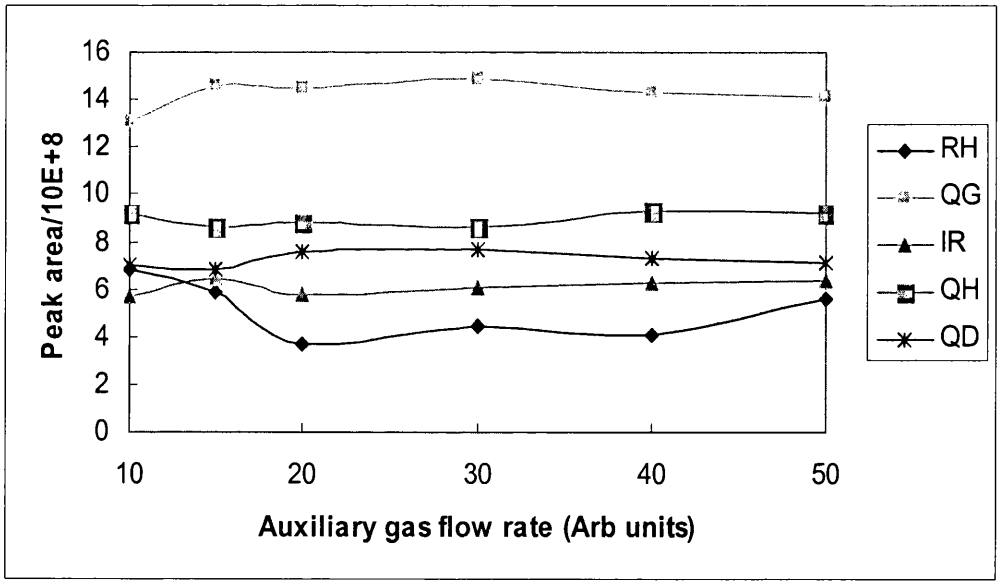


Figure 3.6 Effect of sheath gas flow rate (arbitrary units) upon ionisation of the flavonoids and terpene lactones

Next the auxiliary gas levels were altered and the response of the standard compounds monitored. These gas levels were varied between 10 and 50 arbitrary units (the maximum level allowed with a sheath gas flow of 90 arbitrary units), the relative intensities of the compounds affected by these changes in gas flow rate are shown in Figures 3.7 (a) and (b).

The variation of auxiliary gas levels showed less change than sheath gas flow on signal intensities. The majority of the compounds showed maximum signal intensity at flow rate of 15 arbitrary units, except RH showed maximum signal at lower flow rate, therefore auxiliary gas level of 15 arbitrary units was used in further studies.

(a)



(b)

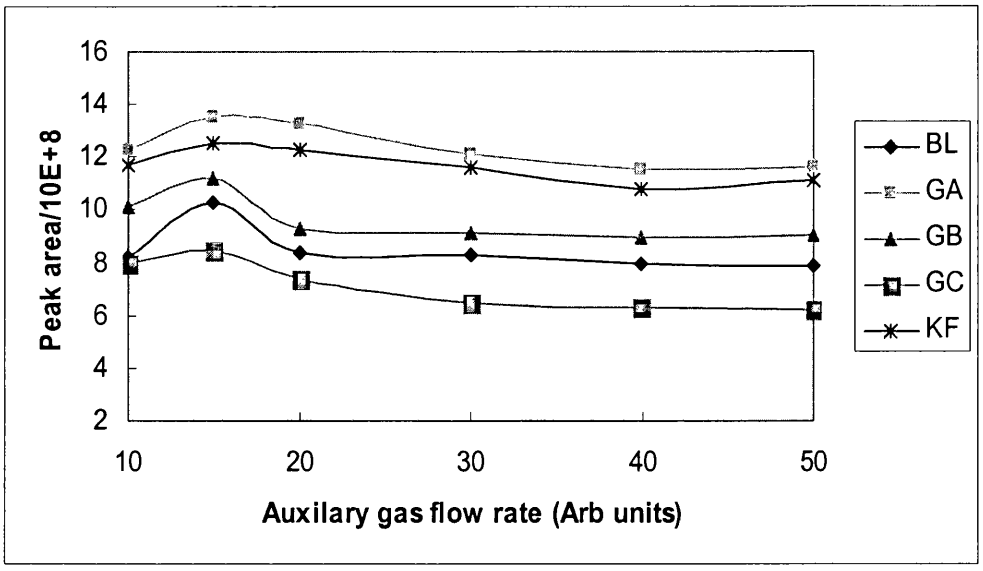
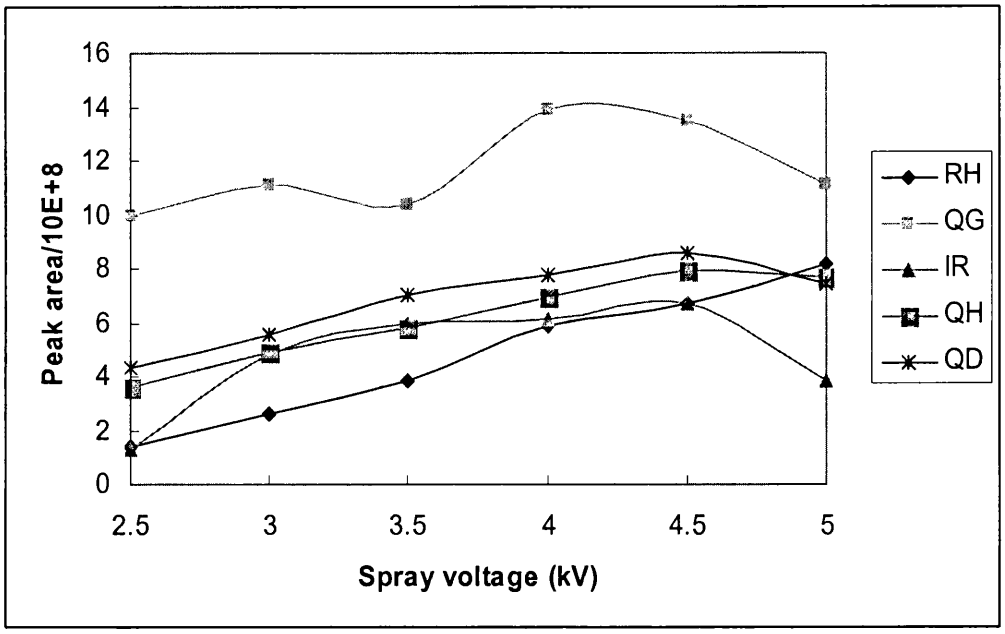


Figure 3.7 Effect of auxiliary gas flow rate (arbitrary units) upon ionisation of the flavonoids and terpene lactones

3.3.2.3 Optimisation of spray voltage

The spray voltage was altered between 2.5 and 5 kV and the signal of all the compounds tested were monitored. Of the compounds tested the results where a significant difference in signal was detected are shown in Figures 3.8 (a) and (b). All the compounds tested showed a gradual increase of signal intensity with increasing spray voltage, with maximum signals at 4.5 kV, although there are some fluctuation for QG and GA. Further increase of spray voltage caused signal reduction for most of the compounds (8 out of 10). Therefore a spray voltage of 4.5 kV was used for all further analysis.

(a)



(b)

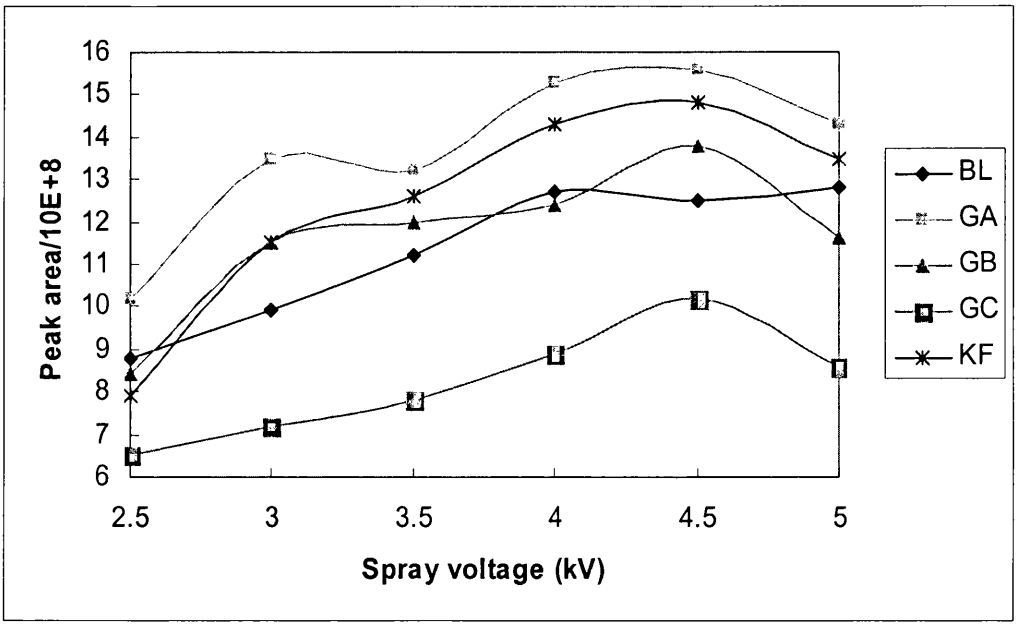
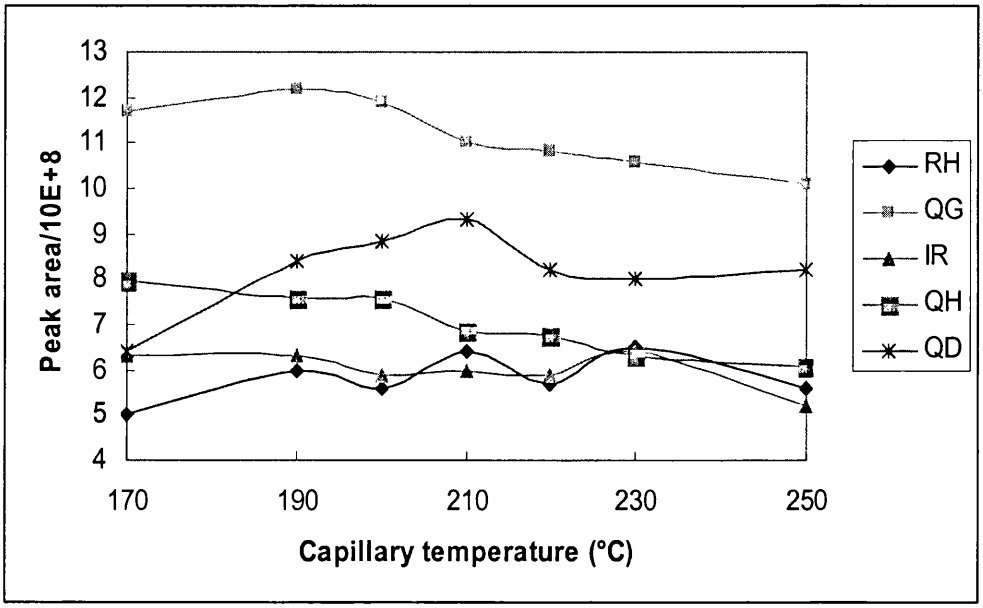


Figure 3.8 Effect of spray voltage upon ionisation of the flavonoids and terpene lactones

3.3.2.4 Optimisation of capillary temperature

The spray capillary temperature was altered between 170 and 250 °C and the intensity of the molecular ion species of the standard compounds monitored. As shown in Figures 3.9 (a) and (b), 8 out 10 of the compounds shows maximum signal intensity at 190 °C, therefore the capillary temperature of 190 °C was used for further analysis.

(a)



(b)

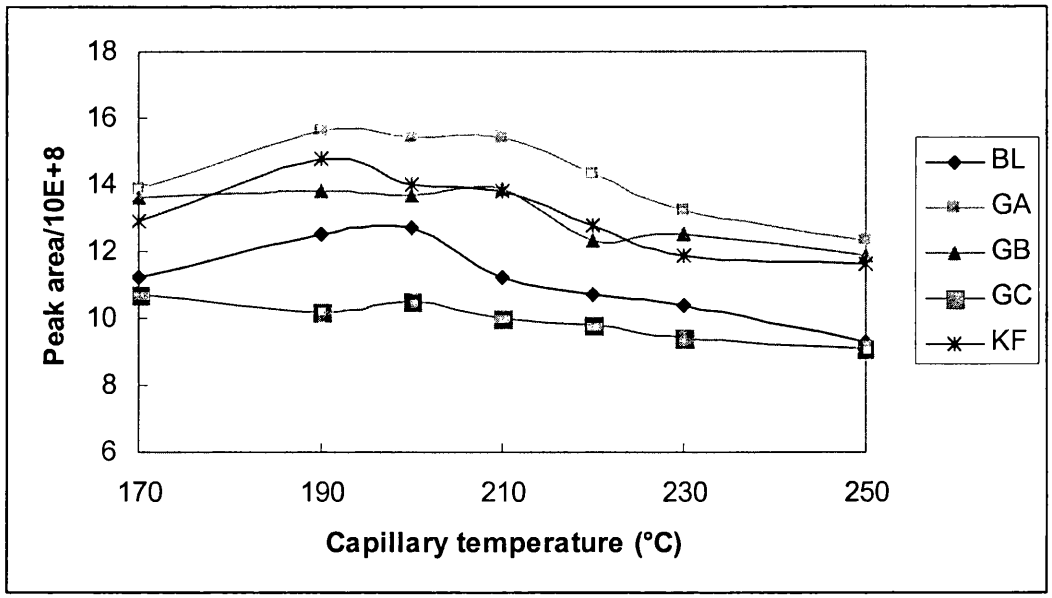


Figure 3.9 Effect of capillary temperature upon ionisation of the flavonoids and terpene lactones

3.3.3 Optimisation of extraction procedures for *Ginkgo biloba* nutritional supplement samples

Extraction of flavonoids and terpene lactones from *Ginkgo biloba* is a challenge. The traditional quality control method for determination of flavonoids in *Ginkgo* products is extraction by hydrolyzing in acidified methanol to convert all the flavonoid glycosides into their flavonoid aglycone counterparts. After that several sample clean up steps are needed including a solid phase extraction (SPE) procedure[5]. For extraction of terpene lactones from *Ginkgo* extract, there are considerable sample clean up and partition steps because there are many apolar compounds e.g. ginkgolic acid, biflavones and chlorophyll, as well as polar constituents e.g. flavonoid glycosides which can interfere with the UV detection if they are not removed. Many procedures have been published over the last two decades. Early procedures were very time-consuming and error-prone

which comprised up to 35 partitioning steps[21], since then there are significant developments in the sample clean-up of various *Ginkgo* samples for terpene lactones analysis[9, 22-24]. However, a SPE column or liquid-liquid extraction (LLE) step is still necessary to improve the rate of extraction of terpene lactones while filtering the intervening components. These laborious procedures make extraction and analysis of large numbers of samples difficult. The most simple sample clean-up is "no sample clean-up", just an extraction immediately followed by analysis. Ganzera *et al.*[25] investigated a *Ginkgo* extract with no sample clean-up using a HPLC system and ELSD detector. Although the method was validated in terms of recovery, peak purity, limit of detection, linearity, extraction efficiency and reproducibility, the sample amount was rather high and there were some criticisms on the peak purity evaluation[20]. However, the application of a mass spectrometry detector, due to its high sensitivity and selectivity, allows the sample size to be reduced and peak purity is much less of a problem. In this study, many different solvents were studied to extract the samples and an optimised extraction was developed for LC/MS quantitation.

3.3.3.1 Comparison of different solvent extraction systems

Methanol, acetonitrile, ethanol and 1:1(v/v) methanol: acetonitrile were evaluated for their efficiency in the extraction of active components for *Ginkgo biloba* nutritional supplement. 20 mL of the above solvent were added to 100 mg of *Ginkgo biloba* nutritional supplement powder respectively. The solutions were sonicated in a water bath at 25 °C for 1 hour, 1 mL of the solution was centrifuged, 10μL of the supernatant were injected on HPLC/MS system using the optimised LC/MS conditions previously described. The concentrations were calculated using the calibration curves obtained

from reference standards. Table 3.3 compares the extraction of *Ginkgo biloba* component by different solvents. As shown in this table, methanol is the solvent that can obtain the highest concentration of all the compounds tested, so methanol was chosen as the extraction solvent.

Table 3.3 Comparison of concentration of the different components in one *Ginkgo biloba* supplement extracted by different solvents. The concentrations were calculated using the calibration curves obtained from reference standards.

Solvent	Methanol	Acetonitrile	Ethanol	1:1 Methanol: ACN
BL(μg/mL)	6.73	5.32	6.02	5.54
GC(μg/mL)	9.32	6.49	8.75	6.95
RH(μg/mL)	32.8	26.3	29.5	23.6
QG(μg/mL)	1.45	0.89	1.26	0.97
QH(μg/mL)	0.80	0.46	0.75	0.50
GA(μg/mL)	10.9	5.94	9.79	6.85
GB(μg/mL)	9.56	4.21	9.21	5.38
QD(μg/mL)	6.59	4.16	6.43	5.17
KF(μg/mL)	1.18	0.71	1.07	0.99
IR(μg/mL)	1.12	0.79	1.10	0.95

3.3.3.2 Comparison of different methanol concentrations

As methanol showed the best extraction efficiency, different concentrations of methanol were compared for the extraction of the 10 active components in *Ginkgo biloba*. Table 3.4 shows 100% methanol gives the best result, considering that the standard solutions were prepared in 100% methanol, it was chosen as the extract solution of choice.

Table 3.4 Comparison of concentration of the different components in one *Ginkgo biloba* supplement extracted by different concentrations of aqueous methanol. The concentrations were calculated using the calibration curves obtained from reference standards.

Methanol (%, v/v)	50	60	70	80	90	100
BL(μg/mL)	5.82	4.91	5.84	6.60	6.13	6.75
GC(μg/mL)	8.14	7.38	8.47	8.99	9.07	9.36
RH(μg/mL)	27.6	26.1	32.5	35.1	34.4	33.2
QG(μg/mL)	1.04	1.16	1.46	1.37	1.48	1.56
QH(μg/mL)	0.66	0.61	0.69	0.68	0.74	0.82
GA(μg/mL)	8.31	7.61	7.45	8.64	8.68	11.1
GB(μg/mL)	8.06	73.3	8.17	8.34	8.85	9.58
QD(μg/mL)	3.90	5.14	5.98	6.09	6.37	6.64
KF(μg/mL)	0.78	1.33	1.56	1.57	1.39	1.21
IR(μg/mL)	0.60	1.00	1.13	1.15	1.17	1.18

3.3.3.3 Comparison of different sonication times on the efficiency of extraction

Sonication can greatly reduce the time of extraction; Table 3.5 shows the effect of sonication time on the extraction of flavonoids and terpene lactones in *Ginkgo biloba*. It shows the concentrations of all the standards are constant between 30-50 minutes in the sonicator. Too long a sonication will result in a decrease of the standards concentration, this might arise due to the heating caused by sonicating which may degradate the standards. Therefore, 50 minutes sonication time was chosen in this study.

Table 3.5 Comparison of concentration of the 10 components in one *Ginkgo biloba* supplement extracted by different sonication time. The concentrations were calculated using the calibration curves obtained from reference standards.

ɔnication time (min)	10	20	30	40	50	60	90	120
BL(μg/mL)	6.13	5.93	6.94	6.78	6.64	5.61	5.20	4.60
GC(μg/mL)	8.61	8.93	8.43	8.85	9.98	7.43	8.51	5.89
RH(μg/mL)	43.8	41.2	50.5	49.6	53.2	34.3	28.6	25.4
QG(μg/mL)	1.46	1.48	1.70	1.65	1.71	1.28	1.08	1.37
QH(μg/mL)	0.75	0.65	0.78	0.75	0.80	0.68	0.62	0.53
GA(μg/mL)	8.85	6.80	8.05	9.05	9.05	6.68	7.03	6.88
GB(μg/mL)	7.70	7.41	7.12	9.16	9.59	6.85	6.95	6.98
QD(μg/mL)	9.32	8.01	6.34	7.35	5.23	5.62	4.81	5.25
KF(μg/mL)	0.93	0.93	0.86	0.90	0.83	0.75	0.84	0.81
IR(μg/mL)	0.93	0.85	0.74	0.83	0.69	0.69	0.75	0.71

3.3.3.4 Optimisation of Soxhlet extraction procedure

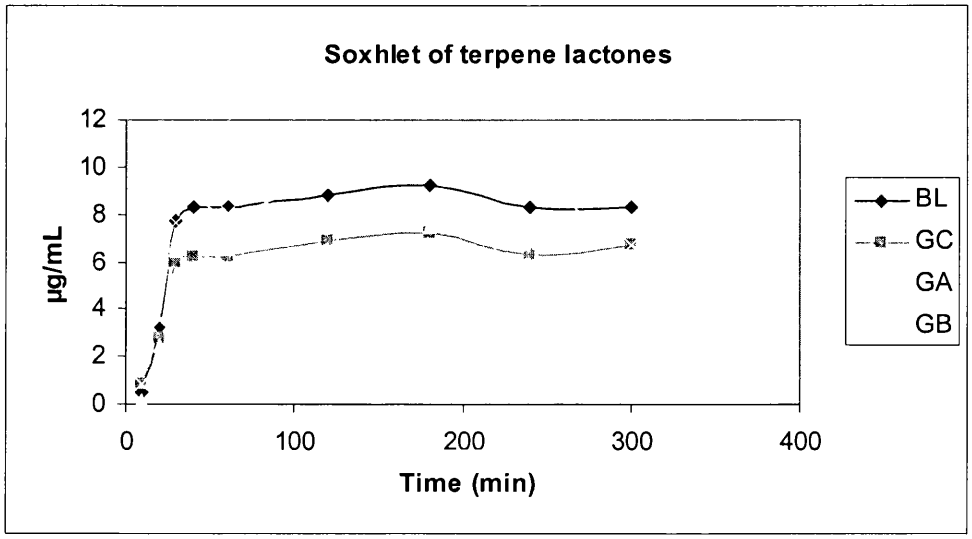
Traditional quality control of *Ginkgo biloba* extract is evaluated by total flavonoids, in this process, the flavonoid glycosides are converted to flavonoid aglycones under acidic reflux, then the total flavonoid content can be obtained by multiplying by 2.51 for the amount of aglycones[20]. In this study, the amount, of the 10 active components, upon different reflux times was studied.

About 500 mg of *Ginkgo biloba* nutritional supplement powder was accurately weighed and put into a Soxhlet thimble and refluxed with 100ml of methanol and 25%HCl (9:1). Figures 3.10 (a) to (d) show the change of concentration of each compound upon the

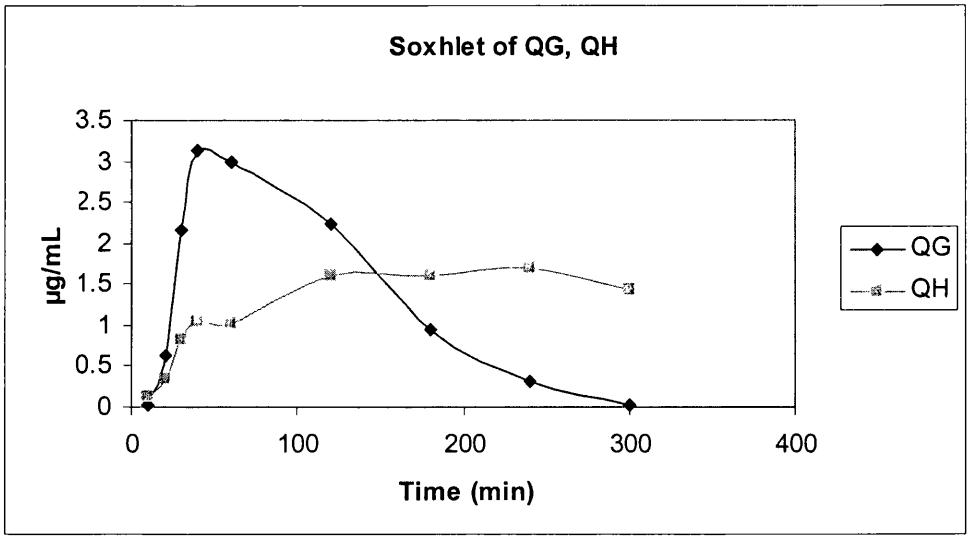
reflux time. The amount of terpene lactones remain unchanged up to 4 hours but longer reflux time causes ginkgolide B to degrade as shown in Figure 3.10 (a). Figures 3.10 (b) and (c) show that it takes about 40 minutes for flavonoids to reach their maximum concentration, then the concentration gradually reduces since they are hydrolysed to their aglycone counterparts. However, it was noted that the concentration of QH remained relatively constant even after 5 hours refluxing. Figure 3.10 (d) shows that it takes about 5 hours to convert all flavonoid glycosides into aglycones and the refluxing time for tradition QC (quality control) methods was set to 5 hours accordingly.

Figure 3.10 (a)-(d) The concentration of flavonoids and terpene lactones in *Ginkgo biloba* nutritional supplement upon refluxing time in Soxhlet extraction with 100ml of methanol and 25%HCl (9:1)

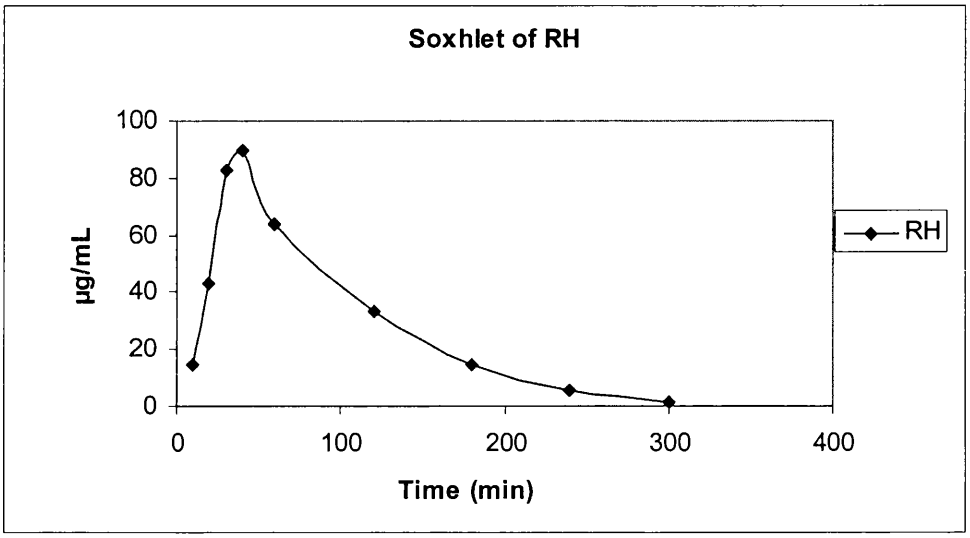
(a)



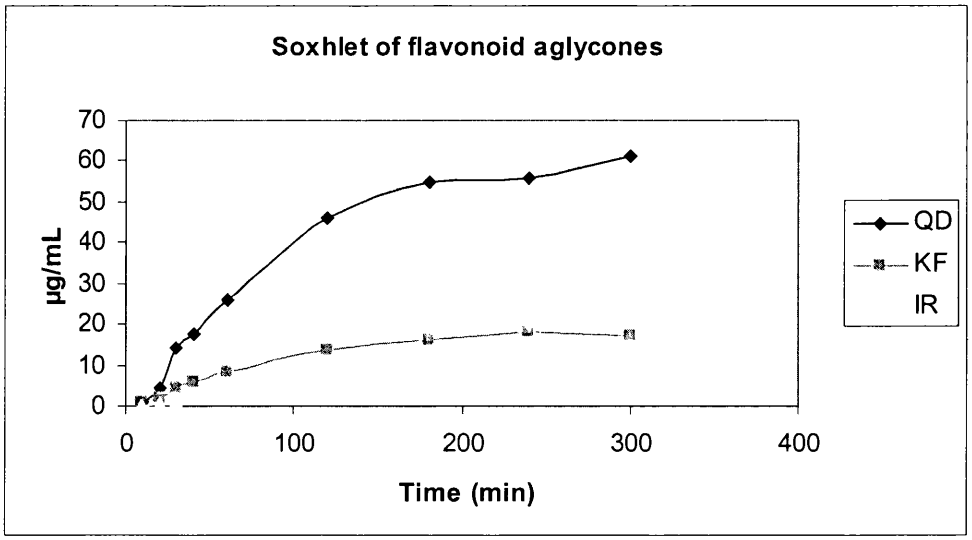
(b)


Soxhlet of QG, QH
3.5
3
2.5
2
1.5
1
0.5
0
µg/mL
0
100
200
300
400
Time (min)
QG
QH

(c)


Soxhlet of RH
100
80
60
40
20
0
µg/mL
0
100
200
300
400
Time (min)
RH

(d)


Soxhlet of flavonoid aglycones
70
60
50
40
30
20
10
0
µg/mL
0
100
200
300
400
Time (min)
QD
KF
IR

3.4 Results and discussions

3.4.1 Sensitivity

The limit of detection (LOD) was studied in order to investigate the sensitivity of the assay developed. In this study, the LOD was estimated by serial dilution of the analyte until the diluted sample gave a signal-to-noise ratio (S/N) of 3. This was conducted using both UV detection and detection by different forms of mass spectrometric analysis. Table 3.6 shows the comparison of LOD by UV detection and MS analysis using both single ion monitoring (SIM), of the $[M-H]^-$ ion, and single reaction monitoring (SRM), refer to Table 3.2 for the MS/MS transitions employed. It can be concluded from the data (Table 3.6) that mass spectrometry consistently gave better sensitivity compared to UV analysis and that selected ion monitoring gave the best sensitivity. Whilst the sensitivity of selected reaction monitoring mode was in most cases comparable with the signal obtained in full scan mode, the exceptions to the latter being those obtained for BL, QG and GA.

Theoretically, SRM analysis should be considered, over SIM analysis, when the background signal noise of the mass spectrometer is high. However, SRM was found to be significantly poorer than SIM mode. In this analysis there are 10 or more fragmentation processes for each compound, over which the ion signal is effectively diluted, thus reducing sensitivity in SRM mode. Therefore the SIM analysis was deemed to be the most beneficial for the further development of the assay.

Table 3.6 Comparison of limit of detection of the 10 components in *Ginkgo biloba* by UV and MS under different analytical modes of detection, namely UV, full scan mass spectra, selected ion monitoring (SIM) and selected reaction monitoring (SRM)

Component	UV (μg/mL)	Full scan (μg/mL)	SIM (μg/mL)	SRM (μg/mL)
BL	undetectable	0.73	0.044	0.185
GC	undetectable	0.10	0.085	0.10
RH	1.06	0.28	0.010	0.28
QG	0.86	0.35	0.042	0.089
QH	0.75	0.18	0.031	0.09
GA	undetectable	0.77	0.076	0.19
GB	undetectable	0.08	0.010	0.17
QD	2.05	1.81	0.40	1.80
KF	1.30	0.52	0.065	0.18
IR	1.36	0.56	0.068	0.28

* λ=250nm for UV analysis, ions monitored inSIM and SRM mode see Table 3.2

3.4.2 Linear range

Calibration curves were constructed by plotting peak areas of the SIM signal of $[M-H]^-$ ions against concentration, the calibration curves of the 10 reference standards are shown in Figures 3.11-3.16.

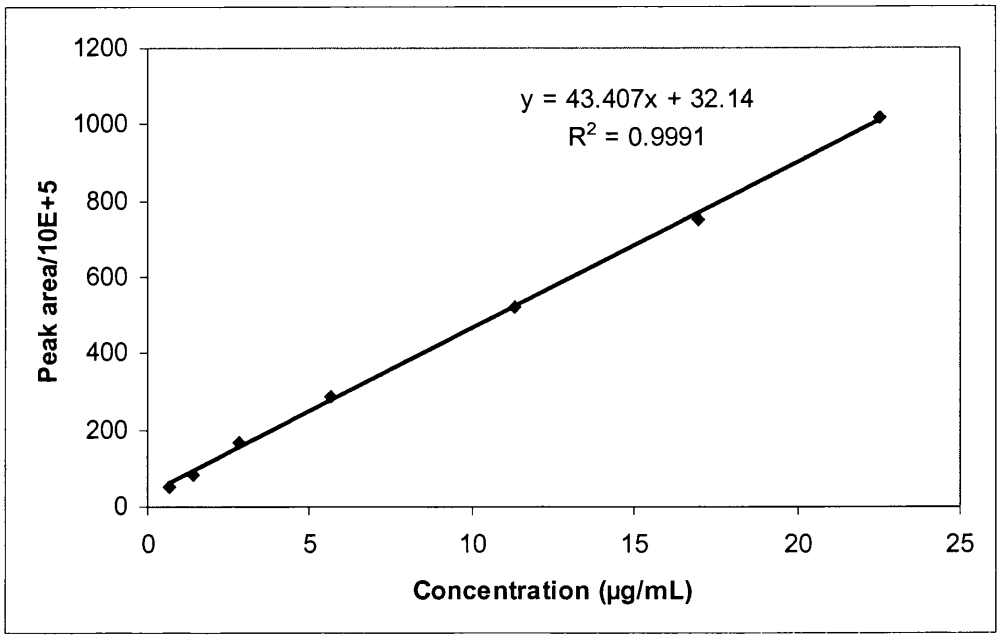


Figure 3.11 Calibration curve of bilobalide

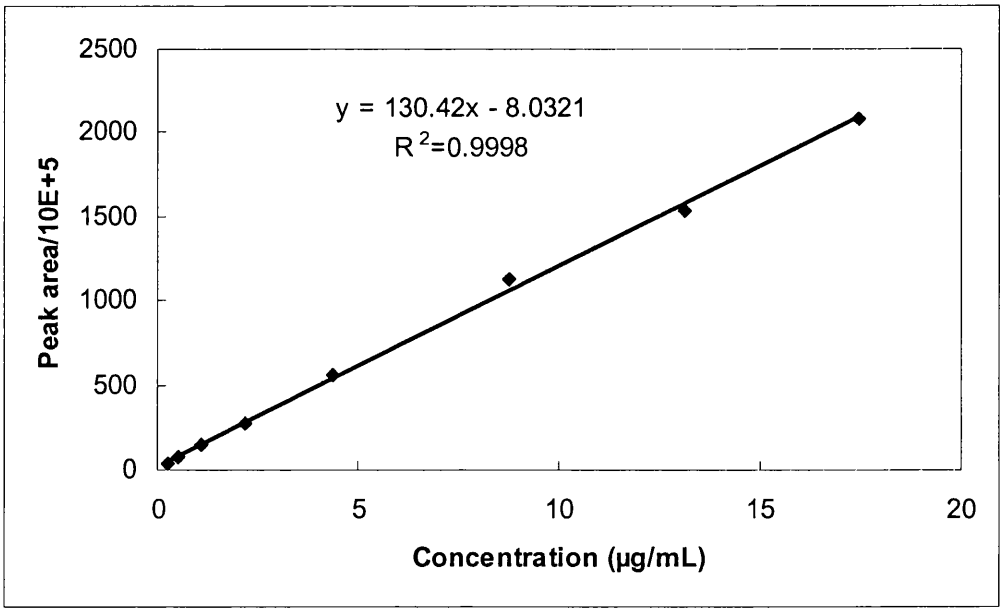


Figure 3.12 Calibration curve of ginkgolide C

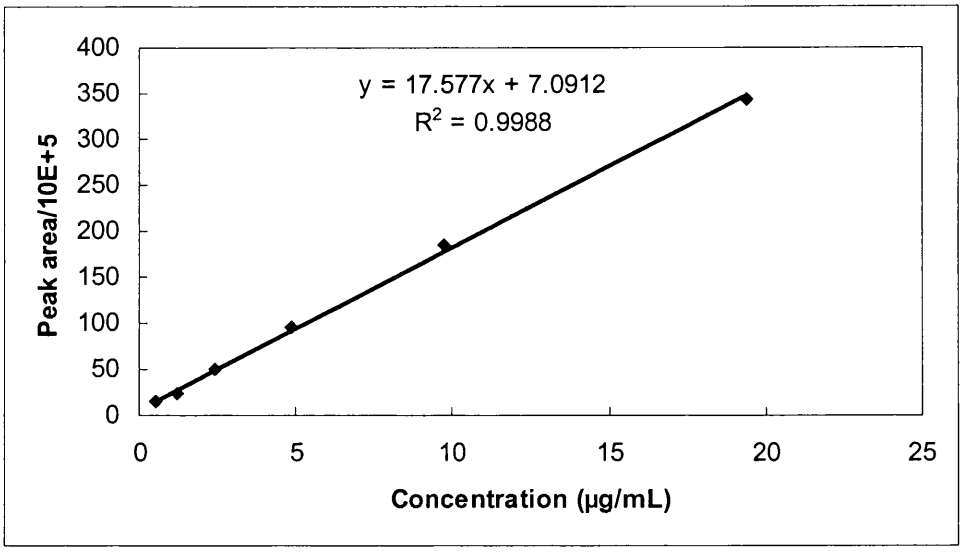


Figure 3.13 Calibration curve of ginkgolide A

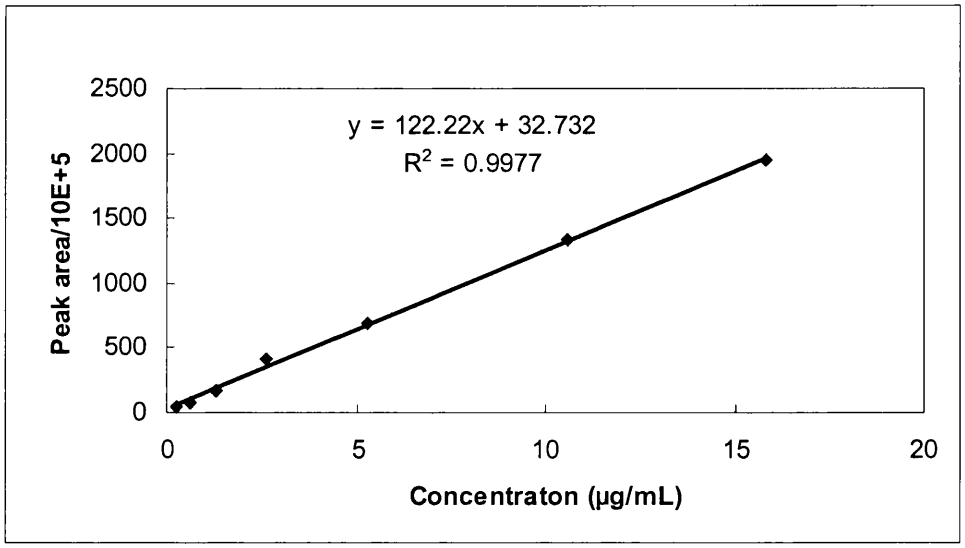


Figure 3.14 Calibration curve of ginkgolide B

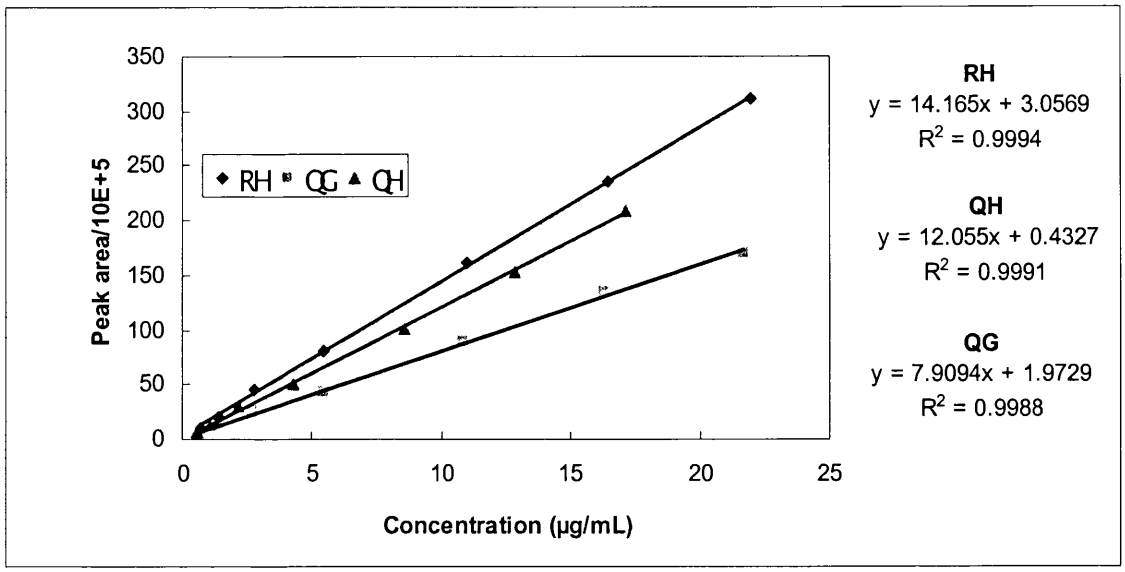


Figure 3.15 Calibration curve *of Ginkgo* flavonoids

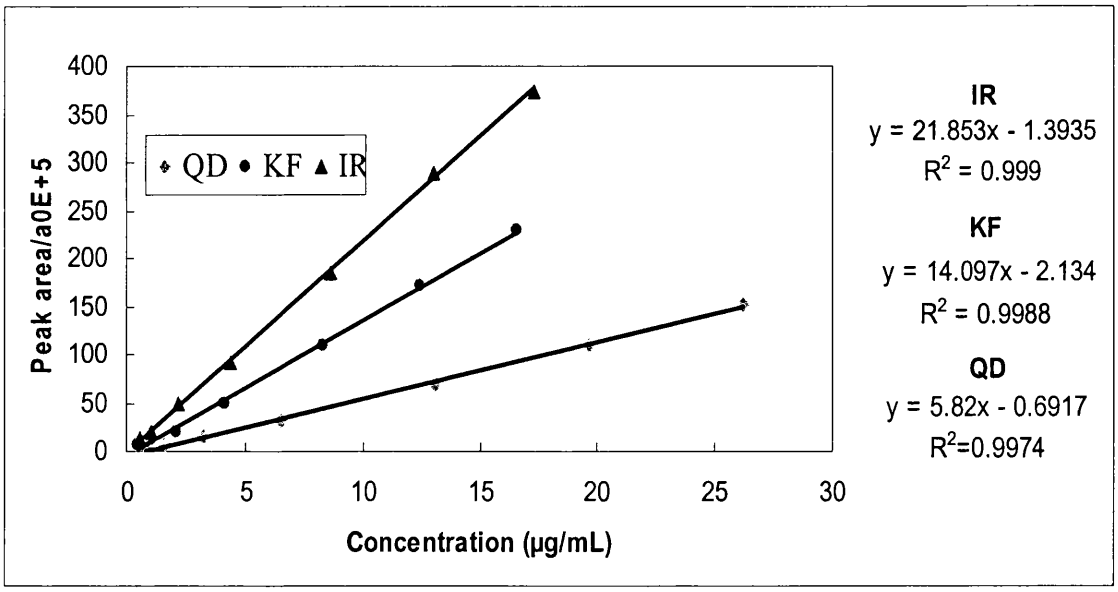


Figure 3.16 Calibration curve for *Ginkgo* flavonoid aglycones

Linear regression analysis was also performed for each reference standard. Table 3.7 compares the linear range and correlation coefficients obtained for the regression analysis based on a minimum of six data points with and without an internal standard added. The internal standard used was andrographolide. The use of an internal standard is generally considered a more accurate analytical method, however in this study, its use showed no improvement in the quantitative analytical result. The extraction process is very simple and has no drying or reconstitution steps and the LC/MS system used in this study is free of manual intervention steps, which provided several advantages over the conventional manual injection method. The high degree of automation provides consistent sampling, so the variation due to sample injection is thought not to be a major cause of error. As shown in Table 3.7, the method employed without an internal standard is better than that using an internal standard in terms of their comparative linearity. This suggests that manual sample preparation induces more errors into the entire extraction-analysis than automatic sample injection and mass spectrometric performance. The data also suggests that no loss of sensitivity is exhibited through continued analysis of the compounds and so the need for an internal standard, to compensate for loss of mass spectrometric performance, is not required.

Table 3.7 Linear range for the 10 active components in *Ginkgo biloba* by LC/ESI-MS in negative ion SIM mode

Component	Without internal standard		With internal standard	
	Linear range (μg/mL)	R^2	Linear range (μg/mL)	R^2
BL	0.7-23	0.9991	0.7-11.5	0.9727
GC	0.27-17.5	0.9998	0.27-17.5	0.9982
RH	0.68-22	0.9994	0.68-22	0.9982
QG	0.68-22	0.9988	0.68-22	0.9934
QH	0.53-18	0.9991	0.53-18	0.9972
GA	1.21-19.4	0.9986	1.21-19.4	0.9977
GB	0.66-15.8	0.9976	0.66-15.8	0.9953
QD	0.8-26	0.9974	0.8-26	0.9894
KF	0.52-17	0.9988	0.52-17	0.9982
IR	0.54-17	0.9990	0.54-17	0.9984

To test whether these two calibration methods differ in their precision, a two- tailed test was applied to the two sets of sample variance, the result of F-test on the normalised standard deviation of each calibration, as shown in Table 3.8. The critical F value for degree of freedom 6 is 5.82 (P=0.05)[3]. Eight out of ten *Ginkgo* calibrations show no significant difference between these two methods and only one sample BL shows significant difference between the two calibrations. Thus there is no significant difference using either an internal standard or not.

Table 3.8 F-test for the comparison of two calibration methods to examine if they are statistically different. The two methods being compared are I. Without internal standard and II. With internal standard.

Component	I. Without internal standard			II. With internal standard			$F = s_2^2/s_1^2$	Methods different
	slope	s_b	s_1 (s_b/slope)	slope	s_b	s_2 (s_b/slope)		
BL	43.41	0.5886	0.01356	0.9615	0.09303	0.09675	50.90	Yes
GC	118.5	2.049	0.01728	1.0393	0.02196	0.02112	1.49	No
RH	14.16	0.1547	0.01092	3.3884	0.07136	0.02105	3.71	No
QG	7.909	0.1207	0.01526	1.1353	0.04644	0.04089	7.18	Yes
QH	12.05	0.1594	0.013223	3.1733	0.08429	0.02656	4.03	No
GA	17.57	0.3082	0.01753	0.4083	0.00973	0.02382	1.84	No
GB	122.2	2.6105	0.02135	1.1606	0.03974	0.034237	2.56	No
QD	5.820	0.1331	0.02286	0.7702	0.02399	0.03114	1.85	No
KF	14.09	0.2163	0.01534	2.6837	0.05765	0.02147	1.96	No
IR	21.85	0.31	0.01418	4.0221	0.07975	0.01983	1.95	No

Notes:

1. S_b is the standard deviation of the slope[3]; s_1 and s_2 are the standard deviations of the normalised slopes of the calibration methods I and II.

2. The F-Test statistics were taken from the calculations in Ref. 3 and were generally the $F_{6,6}$ statistic.

As the internal standard method needs more sample preparation work, based on the comparison results in this study, a method that does not utilise an internal standard was chosen for quantitation. In order to further test the reliability of this method, it was thought necessary to test multiple extracts of real biological samples as well as pure standards in order to determine the robustness of the analyses and whether it can be relied upon without an internal standard.

3.4.3 Intra-day and inter-day reproducibility

The reproducibility of the proposed method was evaluated by carrying out five replicate quantitative determinations for each of the studied compounds present in a given sample on the same day and five on three consecutive days. Table 3.9 and 3.10 illustrate intra-day and inter-day RSDs of the quantitation of the 10 components in *Ginkgo biloba*, respectively. It was shown that good repeatability and reproducibility can be obtained by this method and so the system is applicable for quantitative determination. The %RSD values of intra- and inter-day precision ranged from 5.5 to 11.2% and 6.3-11.2%, respectively, indicating very little variation in the HPLC-MS performance or the extraction efficiency of the methanol solubilisation of the compounds.

Table 3.9 Intra-day precision of the 10 components in *Ginkgo biloba* extraction

Component	BL	GC	RH	QG	QH	GA	GB	QD	KF	IR
					External standard					
Mean	1.90	1.43	11.08	0.24	0.22	1.43	1.51	1.53	0.47	0.27
SD	0.16	0.08	0.73	0.018	0.012	0.16	0.13	0.13	0.036	0.017
CV	8.61	6.12	6.55	7.95	5.50	11.19	8.58	8.50	7.71	6.24
					Internal standard					
Mean	2.12	1.54	11.97	0.35	0.28	1.58	1.64	1.69	0.45	0.31
SD	0.21	0.12	0.95	0.029	0.025	0.18	0.15	0.16	0.042	0.025
CV	10.6	6.49	7.94	8.28	8.93	11.39	9.14	9.47	9.33	8.06

Table 3.10 Inter-day precision of the 10 components in *Ginkgo biloba* extraction

Component	BL	GC	RH	QG	QH	GA	GB	QD	KF	IR
					External standard					
Mean	1.96	1.52	11.65	0.35	0.41	1.52	1.87	1.48	0.53	0.38
SD	0.22	0.11	0.94	0.022	0.038	0.15	0.15	0.16	0.046	0.029
CV	11.2	7.24	8.07	6.29	9.26	9.87	8.02	10.81	8.68	7.63
					Internal standard					
Mean	2.23	1.46	11.38	0.32	0.35	1.48	1.72	1.66	0.49	0.33
SD	0.25	0.13	1.09	0.028	0.025	0.10	0.15	0.16	0.040	0.030
CV	12.3	8.90	9.58	8.75	7.14	6.76	8.72	9.64	8.16	9.09

3.4.4 Recoveries of the 10 compounds in *Ginkgo biloba* nutritional supplements

The recovery values of the methods were calculated as the percentage of analytes recovered by the assay. Recovery of the herbs' active components was obtained by standard addition because the components are already present in the herbal extract naturally and so this naturally occurring level must be determined and compensated. In this study, one of the *Ginkgo* samples was chosen for the recovery test and spiked at three levels, as described in the materials and methods. The results of the recovery experiments are shown in Table 3.11. The data shows that good recovery can be obtained by the extraction and analytical method described, the recovery values obtained were in all, but one case, greater than 90% (with the one outlier at 89.7% recovery) and the variation in recovery is generally better than 5% RSD with the greatest variation being only 11.2% at the lowest spiked level for GB. Among the compounds examined, kaempferol and isorhamnetin shows excellent recovery, the suspected reason for this observation is that the linearity of the calibration curve of these two compounds is better than others, although it has been previously reported that isorhamnetin exhibits a very low recovery by Deng *et al.*[26]

3.4.5 Active components in commercial samples

Five commercial samples were obtained in order to study the variations in the levels of the ten examined compounds in these products. Sample 1 and 2 were bought from a local market, sample 3 was bought from a nationwide health food shop and sample 4, 5 were obtained from Obsidian Research Ltd (Port Talbot, UK). Sample 5 was claimed to contain pure *Ginkgo biloba* extract.

Table 3.11 Recoveries of the 10 active components in different concentrations

Components / Spiked	BL	GC	RH	QG	QH	GA	GB	QD	KF	IR
Mean (10 μg/mL)*	89.7	93.4	98.0	92.5	93.9	92.1	94.2	91.9	99.6	98.8
SD	5.32	3.61	4.12	4.69	2.95	2.82	3.61	2.47	1.86	1.92
RSD (%)	5.93	3.87	4.20	5.07	3.14	3.06	3.83	2.69	1.87	1.94
Mean (5 μg/mL)*	90.2	95.3	95.5	96.2	102.8	98.5	100.6	94.5	99.6	97.7
SD	3.80	5.16	3.59	3.58	5.39	4.27	3.71	3.58	1.58	2.21
RSD (%)	4.22	5.41	3.76	3.72	5.24	4.34	3.30	3.79	1.59	2.26
Mean (1.25 μg/mL)*	109.6	100.3	100.6	104.2	102.9	105.1	99.9	98.6	99.5	99.5
SD	9.13	3.57	8.73	5.76	4.84	3.28	11.23	3.46	4.02	3.53
RSD (%)	9.77	3.56	8.71	5.53	4.70	3.12	11.24	3.51	4.03	3.54

* n=6

The traditional QC method of HPLC separation with UV detection of the hydrolysis formed aglycones was evaluated in this study. The quantitative results are shown in Table 3.12. From this table, all the five samples satisfy the quality control criteria with the total flavonoid levels higher than 24%, interestingly the proportion of the quercetin, kaempferol and isorhamnetin varies greatly.

Table 3.12 The content (%) of *Ginkgo* flavonoid aglycones after acidic hydrolysis

Sample number	**QD**(%)	**KF**(%)	**IR**(%)	Total aglycone (%)	Total flavonoids (%)
1	11.55	2.392	< LOQ	13.94	35.0
2	11.39	1.991	0.6741	14.06	35.3
3	5.806	4.349	1.275	11.43	28.7
4	6.542	3.613	1.023	11.18	28.1
5	8.741	1.855	0.5761	11.17	28.0

The five commercial samples were analyzed by the developed HPLC-SIM-MS method to obtain the levels of the active components in these samples. All the concentrations were converted to the percentage of each component in the *Ginkgo* sample as is commonly specified during quality control analyses. The results are shown in Table 3.13 where it can be observed that the content of the ten components varied greatly between the different commercial samples, although they all claimed the sample has >6% of *Ginkgo* terpene lactones and >24% *Ginkgo* flavonoids. From our study, most of the samples satisfied the stated criteria for the *Ginkgo* terpene lactones (i.e., their levels should combined be >6%), the content of each terpene lactone is generally similar with bilobalide and ginkgolide A the more abundant components. However, the individual

Table 3.13 Content of *Ginkgo* flavonoids and terpene lactones in *some Ginkgo biloba* commercial products *

Sample	BL%	GA%	GB%	GC%	TT%±SD[a]	RH%	QG%	QH%	QD%	KF%	IR%	TF%±SD[b]
1	1.665	1.902	1.383	1.178	6.127±0.04	12.11	0.3485	0.1576	0.9781	0.3351	0.2299	14.16±0.04
2	2.337	1.523	1.180	1.048	6.088±0.08	6.319	0.3864	0.2403	7.978	0.4134	0.3288	15.67±0.03
3	3.252	3.294	2.588	1.472	10.61±0.05	6.276	0.5004	0.3217	0.4606	0.3339	0.1990	8.091±0.02
4	2.908	2.597	1.675	1.738	8.918±0.06	10.29	0.3267	0.1676	0.6497	0.1088	0.1114	11.66±0.02
5(GBE[c])	1.592	1.723	1.084	1.171	5.570±0.03	10.99	0.1039	0.08830	1.618	0.2351	0.1718	13.21±0.03

*Sample was analyzed in triplicate.

a. TT, determined total terpene lactones ± SD; TT=BL+GA+GB+GC

b. TF, determined total flavonoids ±SD; TF=RH+QG+QH+QD+KF+IR

c. GBE, standardized *Ginkgo biloba* leaf extract

amount of specific compound within this class varied from sample to sample, the content of the same compound can differ as much as 200%. Taking ginkgolide A as an example, the content of ginkgolide A in sample 2 is 1.52%, but in sample 3 is 3.29% (see Table 3.13). As for *Ginkgo* flavonoids, rutin accounts for ¼ (6.2%) to ½ (12.1%) of total flavonoids, and a great difference, from 0.46% to 7.98%, in quercetin content was observed. The total flavonoid content determined is much lower than the quality control value of *Ginkgo biloba* extract product, which states that the levels should account for > 24%. This finding is not entirely surprising because the total flavonoids quantified by our analysis only accounts for the three flavonoid aglycones and three flavonoid glycosides which were commercially available. There are many other flavonoids glycosides in the sample which are undetermined because no standards were available and the total glycosides were not hydrolysed to their comparable aglycones prior to analysis in our method. Using the traditional QC method of HPLC separation with UV detection of the hydrolysis formed aglycones, the total flavonoid levels of all samples are higher than 24% suggesting that glycosides other than the ones tested for in our combined assay make up for the difference between our methods total flavanoid level and the level determined by QC methods.

Combining the result of HPLC and HPLC/MS, the content of rutin of sample 1 is too highly elevated (12.11%), and the content of quercetin in sample 2 is also too high (7.98%) in comparison to other samples. It can be concluded that samples 1 and 2 are fortified by low price quercetin or rutin additions, respectively. The *Ginkgo* sample 3 is the best quality product since it has the high content of both terpene lactones (10.6%) and flavonoids (28.7%) while the contents of possible fortified components, rutin and quercetin are low.

These results support previous reports of inconsistent content of commercial *Ginkgo biloba* products[23-26]. However, since the conventional quality control method involves hydrolysis followed by the analysis of flavonoid aglycones, it is not able to distinguish the fortified flavonoids from the natural flavonoid glycosides. The analysis of glycosides becomes important because product fortification is a common problem associated with natural products. The method described here avoids the hydrolysis process which makes the procedure more convenient experimentally and can differentiate some flavonoid aglycones and flavonoid glycosides. However, further availability of standards and their incorporation into the assay would improve the analysis. The method also has the advantage of being able to quantitate the flavonoids and terpene lactones in a single extraction and analysis. This method can be considered as a valuable tool for the quality evaluation of *Ginkgo biloba* dietary supplement products. The data presented here suggests that instead of setting a minimal amount for the total ginkgolides (> 6%) and flavonoids (> 24%), an alternative way of comparing the levels of the two bioactive compounds should be presented as quality control criteria.

3.5 Conclusion

A HPLC-ESI/MS method was developed with the aim of quantitative analysis of ten active components in commercial *Ginkgo biloba* nutritional supplements. Compared with HPLC with UV or ELSD detection, mass spectrometry has the advantage in terms of sensitivity and selectivity, with no tedious sample clean up procedures being required. The limit-of-detection for UV and the various mass spectrometric scan modes employed were determined and compared. Selected ion monitoring (SIM) in negative ionisation mode was chosen for the quantitation of 10 major active components in *Ginkgo biloba* as it showed itself to be the most sensitive from the above preliminary studies. The sample preparation and assay procedure involved is simple, rapid, with good accuracy and reproducibility being demonstrated. This method offers a simultaneous quantitation of ten major active components in a single run, terpene lactones, flavonoid glycosides and aglycones can be monitored at the same time, no hydrolysis of flavonoid glycosides to aglycones is required.

There is remarkable variation in the contents of the flavonoids and terpene lactones although all the *Ginkgo biloba* nutritional supplements satisfy the traditional quality control standards which set a minimum amount of the total flavonoids >24% and ginkgolide >6%. The analysis of flavonoid glycosides as well as aglycones can be useful indicators of sample stability, too high concentration of rutin or quercetin can be recognized as fortification, this is circumvented when samples are hydrolyzed to back calculate flavonoid glycoside content in traditional quality control methods. The

fortification of low cost components rutin or quercetin was observed in two out of five commercially available *Ginkgo biloba* products being analysed in this study. These results indicate that suitable quality control methods need to be implemented to ensure the quality of *Ginkgo biloba* nutritional supplements. This method may serve as a valuable tool for the quality evaluation of *Ginkgo biloba* dietary supplement products.

3.6 References

1. Joannou, G.E., Kelly, G.E. and Reeder, A.Y., J. Steroid Biochem. Mol. Biol., **54**, 167 (1995)

2. Bayer, T., Colnot, T. and Dekant, W., Toxicol. Sci., **62**, 205 (2001)

3. Miller, J.C. and Miller, J.N., *Statistics for Analytical Chemistry*, Ellis Horwood Ltd, Halsted Press, Chinchester (1988)

4. Jamshidi, A., Adivadi, M. and Husain, S.W., J. Planar Chromatogr., **13,** 57 (2000)

5. Hasler, A., Sticher, O. and Meier, B., J. Chromatogr., **605**, 41 (1992)

6. Graefe, E.U., Derendorf, H. and Veit, M., Int. J. Clin. Pharmacol. Ther., **39,** 219 (1999)

7. Pietta, P., Mauri, P. and Rava, A., J. Pharm. Biomed. Anal., **10**, 1077 (1992)

8. Camponovo, F.F., Wolfender, J.L., Maillard, M.P., Potterat, O. and Hostettmann, K., Phytochem. Anal., **6**, 141 (1995)

9. van Beek, T.A., Scheeren, H.A., Rantio, T., Melger, W.C. and Lelyveld, G,P., J. Chromatogr. **543**, 375 (1991)

10. Chauret, N., Carrier, J., Mancini, M., Neufeld, R., Weber, M. and Archambault, J., J. Chromatogr., **588**, 281 (1991)

11. Deng, F. and Zito, S.W., J. Chromatogr. A, **986**, 121 (2003)

12. Chen, P., Su, X.L., Nie, L.H. and Yao, S.Z., J. Chromatogr. Sci., **36**, 197 (1998)

13. Strobe, J.T.B., Taylor, L.T. and van Beek, T.A., J. Chromatogr. A, **738**, 115 (1996)

14. van Beek, T.A., van Veldhuizen, A., Lelyveld, G.P. and Piron, I., Phytochem. Anal., **4**, 261 (1993)

15. Li, W. and Fitzloff, J.F., J. Liquid Chromatogr. Related Tech., **25**, 2501 (2002)

16. Waridel, P., Wolfender, J.L. and Ndjoko, K., J. Chromatogr. A, **926**, 29 (2001)

17. Jensen, A.G., Ndjoko, K., Wolfender, J.L., Hostettmann, K., Camponovo, F. and Soldati, F., Phytochem. Anal., **13**, 31 (2002)

18. Sun, Y., Li, W., Fitzloff, J.F. and van Breemen, R.B., J. Mass Spectrom., **40**, 373 (2005)

19. Liu, C., Mandal, R. and Li, X.F., Analyst, **130**, 325 (2005)

20. van Beek, T.A., J. Chromatogr. A, **967**, 21 (2002)

21. Lobstein-Guth, A., Briancon-Scheid, F. and Anton, R., J. Chromatogr., **267**, 431 (1983)

22. Lang, Q., Yak, H.K. and Wai, C.M., Talanta, **54**, 673 (2001)

23. Lolla, E., Paletti, A. and Peterlongo, F., Fitoterapia, **69**, 513 (1998)

24. van Beek, T.A. and Taylor, L.T., Phytochem. Anal., **7**, 185 (1996)

25. Ganzera, M., Zhao, J. and Khan, I.A., Chem. Pharm. Bull., **49**, 1170 (2001)

26. Deng, F. and Zito, S.W., J. Chromatogr. A, **986**, 121 (2003)

Chapter 4

Fingerprint profile of *Ginkgo biloba* nutritional supplements by LC/ESI-MS/MS

4.1 Introduction

Quality control of traditional Chinese herbs by "fingerprint analysis" has been given much attention in past reports as there are many varying factors in herbal medicine, unlike synthetic drugs, raising concerns in quality control. Species variation, geographical source, cultivation, harvest, storage and process are all factors leading to a product of different quality and efficacy. Compared with synthetic drugs, plant extracts are very complex mixtures, and the therapeutic effect of the plant is often not caused by only one or two compounds but the combination of a group of components[1]. In traditional Chinese medicine, the combination of different herbs is crucial since the same plants, when combined in different ratios, will cause different effects, therefore it is important to have an overall view of all the components in the plant extract in order to evaluate the quality of plant products. This is sometimes achieved by a process of "fingerprint analysis" in which the experimental data from the chemical analysis of different extracts are compared without accurately quantitating every individual compound known to occur in the extract. The major task of the fingerprint is identification and evaluation the similarity and stability among the samples. Usually a reference is used in fingerprint study. For crude plant evaluation the reference fingerprint should be the chromatogram of the plant from the original place; for plant extract, the reference fingerprint is the chromatogram of the product that has been proved to be therapeutically effective. Although "pure" quantitation is not achieved, any deviation from the comparable untainted or unfortified extract is an indication of an impure or tampered with extract.

Each extract has its own chemical "fingerprint" and this has often been investigated using high pressure liquid chromatography (HPLC), thin layer chromatography (TLC) or nuclear magnetic resonance (NMR) analysis. Fingerprint chromatography was introduced in the 1990s' and subsequently accepted by World Health Organization for the quality control of herbal medicines[2]. The use of fingerprinting analysis for quality control and standardization of medicinal herbs has attracted interest in herb research in recent years[3]. Fingerprint analysis is used in the quality control of traditional Chinese medicines and their raw materials as stated by the Chinese Pharmacopoeia[4] and is also now a demand of the European Economic Council (EEC) guideline 75/318 "Quality of Herb Drugs"[5]. Chromatographic profiles of major components are used to evaluate herbal growers and suppliers, to standardize raw materials and to control formulation and tablet content uniformity and screen for adulterants. Thin layer chromatography (TLC) has been the most widely used classical method for fingerprint analysis in herbal extracts. More recently HPLC-UV or HPLC/MS have been used for the detailed profiling of active components of trace marker compounds[6]. Hasler and Sticher obtained a fingerprint chromatogram of *Ginkgo biloba* and characterized components using pure reference flavonoids using a HPLC three-pump system with a diode-array detector. Altogether 33 flavonoids were identified unambiguously[7], however, HPLC with UV detector is not suitable for terpene lactone analysis. Mauri *et al.* obtained a typical fingerprint-like spectrum by direct infusion in ESI-MS mode and assigned 9 flavonoid glycosides which are known to be present in *Ginkgo biloba*[8]. Recently nano-electrospray MS by direct infusion was used for the simultaneously detection of both active components[9], giving a fingerprint-like profile. Due to the possibility of ion suppression by more abundant compounds in biological mixtures, direct infusion is

thought not be able to provide a comprehensive profile of the plant extract and provide reliable quality control information. Among the various hyphenated instruments, liquid chromatography coupled with ion trap tandem mass spectrometry (LC/ITMS) is generally considered as one of the most powerful tools to perform on-line compositional and structural analysis of active constituents in plant extracts[10]. In this study, LC/ITMS was carried out in order to produce a fingerprint profile of *Ginkgo biloba* extract and the knowledge of the fragmentation pathways of active components was applied to determine the active components in *Ginkgo biloba* found in the fingerprint. Two LC/MS systems were used to compare the separation of *Ginkgo biloba* extracts in this study. One system is Hewlett Packard 1100 HPLC system (Hewlett-Packard, Wilmington, DE, USA) with LCQ ion trap mass spectrometer (Finnigan, Hemel Hempstead, UK) using normal size column Luna C_{18} RP column (100×4.6 mm, 5 μm, Phenomenex). The second system consisted of a LC Packings Ultimate Capillary LC system (Dionex, Amsterdam, Netherlands) with LCQ Deca XP ion trap (Thermo Finnigan, San Jose, CA), the column used in this system was an in-house made C_{18} capillary column. The use of characteristic fragmentation profiles was also investigated in order to determine whether signals unique to the flavonoids and terpene lactones could be obtained to give a clearer comparison of different extracts.

4.2 Experimental

4.2.1 Materials

HPLC grade solvents: methanol, acetonitrile, formic acid and acetic acid were purchased from Fisher chemicals (Loughborough, UK) and were used without further

purification. Water was purified with a Milli-Q deionisation unit (Millipore, Bedford, MA, USA). Gases used included oxygen free nitrogen and helium which were purchased from BOC Ltd (Surrey, UK).

4.2.2 Sample preparation

Five commercial *Ginkgo* extract products (as described in Section 3.4.5) were dissolved in methanol by sonicating at 25 °C for 50 minutes, and centrifuged for 10 minutes at 17,000 g (see 3.2.3 for detail) 10 μL of the supernatant were analyzed by HPLC/MS.

4.2.3 LC/MS of *Ginkgo biloba* nutrition supplement by normal-bore (4.6 mm i.d.) column

LC/MS was used to obtain a fingerprint chromatogram of methanol extracted *Ginkgo biloba* nutritional supplement. The LC system was a Hewlett Packard 1100 HPLC system (Hewlett-Packard, Wilmington, DE, USA) and the column used was a Luna C_{18} RP column (100×4.6 mm, 5 μm, Phenomenex). The mobile phase was composed of 0.1 %(v/v) acetic acid in water (A), 1:1(v/v) mixture of ACN and methanol (B). The gradient ran from 20% B to 45%B over 50 minutes, the column was then washed with 100% B for 5 minutes and re-equilibrated for 10 minutes. 10 μL of methanol extracted *Ginkgo biloba* nutritional supplement were injected onto the LC/MS system using an autosampler. The flow rate used was 0.8 mL/min, and a splitter was used to transfer only $\frac{1}{4}$ of the flow into the mass spectrometer. An LCQ ion trap mass spectrometer (Finnigan, Hemel Hempstead, UK) was used for MS analysis. LC/ESI/MS was carried out in the negative ion mode from *m/z* 200-1000 with electrospray ionisation. The heated capillary temperature was 190 °C and the electrospray voltage was 4.5 kV,

sheath gas flow rate was 90 arbitrary, and auxiliary gas flow was 15 arbitrary. The mass spectrometer was operated in a data-dependent mode composing of 3 scan events. The full scan mass spectrum was first obtained and followed by collision-induced dissociation of the highest abundant ion selected from the full scan. Next, in the third scan event, the highest abundant ion from the MS/MS was chosen for further fragmentation (MS^3).

4.2.4 LC/MS of *Ginkgo biloba* nutritional supplement by capillary column

The column used in this experiment was an in-house made capillary C_{18} column (300 μm i.d. x 15 cm). A slurry packing procedure was employed to prepare the column. Slurries were prepared by placing ~5 mg of 100 Å pore size C_{18} resin in a small vial, and adding 1 mL of isopropanol and a mini stirring magnet. A frit was prepared at one end of the 320 μm i.d. fused-silica capillary by dipping into a solution made of 375 μL of Kasil 1624 potassium silicate solution (PQ Corporation, Valley Forge, PA, USA) and 125 μL of formamide (Sigma, St. Louis, MO, USA) for 2 seconds and heating in oven at 100 °C for 4 hours. The column was then packed using a high pressure chamber at constant nitrogen gas pressure of 40 bar until the desired length was packed. The pressure was turned off and the column was further packed overnight as the pressure dissipated. During assembly a second frit (Valco, Houston, TX, USA) was placed at the other end of the column and the column encased in PEEK tubing in order to improve the stability of the column.

The LC/MS system was composed of a LC Packings Ultimate Capillary LC system (Dionex, Amsterdam, Netherlands) with LCQ Deca XP ion trap (Thermo Finnigan, San

Jose, CA). The mobile phase was composed of 0.1 %(v/v) acetic acid in water (A), 1:1(v/v) mixture of ACN and methanol (B), the flow rate was 4 μL/min. The gradient is shown in Table 4.1 and the electrospray settings are shown in Table 4.2. Again the mass spectrometer was performed in data-dependent mode as described for the 4.6 mm i.d. column LC/MS experiment.

Table 4.1 Gradient elution program used in capillary separation

Time (min)	A%	B%	Flow rate (mL/min)
0	90	10	4
120	40	60	4
125	2	98	4
135	2	98	4
138	90	10	4
140	90	10	4

Table 4.2 Mass spectrometry setting of capillary LC/ESI-MS

Parameter	Negative ion mode
Sheath gas flow (arbitrary units)	60
Auxiliary gas flow (arbitrary units)	10
Spray voltage (kV)	2.5
Capillary temperature (°C)	200
Capillary voltage (V)	15
Tube lens offset (V)	60

4.3 Result and discussions

4.3.1 Identification of active components in *Ginkgo biloba* nutritional supplement by LC/MS with normal-bore (4.6 mm i.d.) column

The LC/MS chromatogram of ten standards of major active components in *Ginkgo biloba* is shown in Figure 3.4 in Chapter 3. A Fingerprint chromatogram of a methanol extracted *Ginkgo biloba* nutritional supplement (sample 3) was obtained with the same LC/MS conditions as the standards but with the mass spectrometer running in a data-dependent mode as previously described (Section 4.2.3). The advantage of this mass spectrometry setting is that it can obtain the full scan spectrum of column-eluted components of the sample and the fragment spectrum of the eluted components can be recorded simultaneously without the mass of the compounds being known prior to the analysis.

Using normal-bore column separation, 11 peaks were observed in the extracted *Ginkgo biloba* sample, as shown in Figure 4.1. The information gained from data-dependent HPLC/MS/MS and the knowledge from previous MS/MS and MS^n fragmentation studies were next used to obtain the definite or tentative identifications of peaks in the resulting mass spectrum.

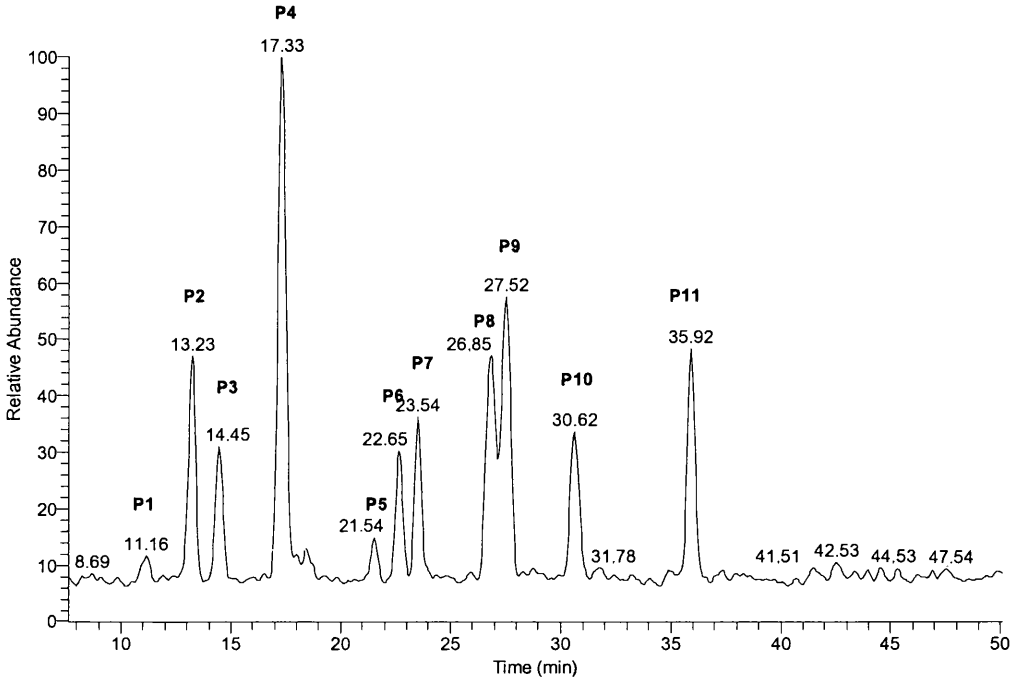


Figure 4.1 Full scan base peak chromatogram by using data-dependent LC/MS/MS of *Ginkgo biloba* sample 3

4.3.1.1 Peak 2, 3, 4, 8, 9

Peak 2 has a retention time of 13 minutes and mass spectrum shows a base peak at *m/z* 325 (Figure 4.2). The MS/MS spectrum of the peak is identical to that previously described for bilobalide (Figure 2.19). From further comparison of the retention time and MS/MS spectrum of the unknown with standard bilobalide, it can be concluded that peak 2 is bilobalide. By the same approach (utilisation of standards), peak 3 can be identified as ginkgolide C, peak 4 can be identified as rutin and peak 8, 9 as ginkgolide A and B, respectively.

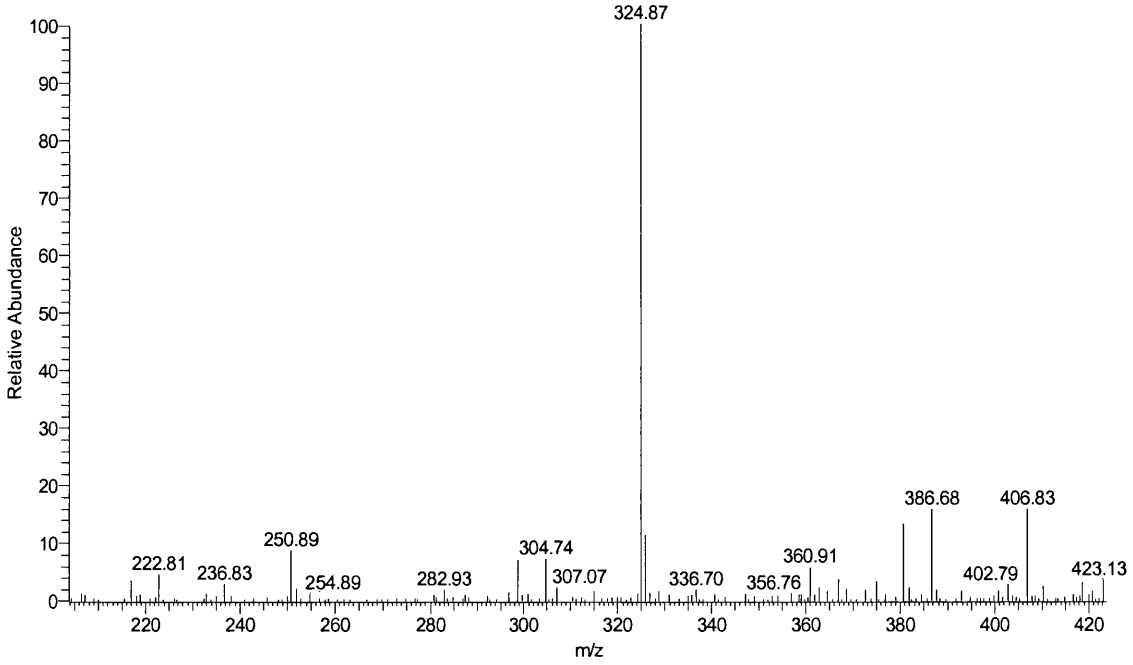


Figure 4.2 Full scan spectrum of peak 2 (Figure 2.33) of *Ginkgo biloba* sample 3

4.3.1.2 Peak 1

This peak occurs at a retention time of 11 minutes and shows a base peak at *m/z* 917 (Figure 4.3). Fragmentation of the *m/z* 917 ion yields product ion at *m/z* 755 with the loss of 162amu (as seen in Figure 4.4). Further fragmentation of *m/z* 755 caused a loss of 146amu and produced a spectrum with product ions *m/z* 609, 343 and 300 as shown in Figure 4.5. The *m/z* 609 and corresponding fragment ions match the ions of intact rutin and its product ions as described (Figure 2.7) and so the compound is thought to be a derivative of rutin. In this MS^3 spectrum the base peak product ion at *m/z* 300 is characteristic peak of quercetin aglycone. Since the mass of the glycoside functional groups glucosyl ($-C_6H_{10}O_5$) and rhamnosyl ($-C_6H_{10}O_4$) are 162 and 146, respectively,

peak 1 is thought to consist of the quercetin moiety with two each of the glucosyl and rhamnosyl groups as opposed to rutin which contains only one of each group. Peak 1 is therefore tentatively identified as 2rhamnosyl-2glucosyl quercetin, however it should be recognised that the order of the sugar moieties can not be distinguished by the data obtained.

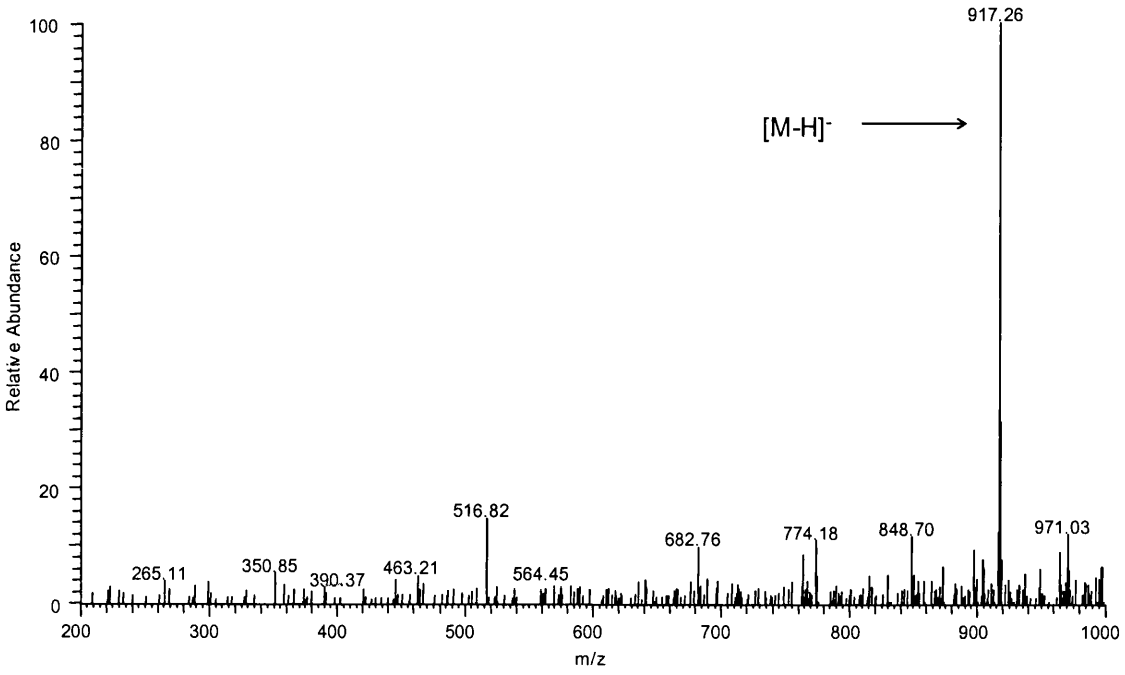


Figure 4.3 Full scan spectrum of peak 1 (Figure 4.1) of *Ginkgo biloba* sample 3

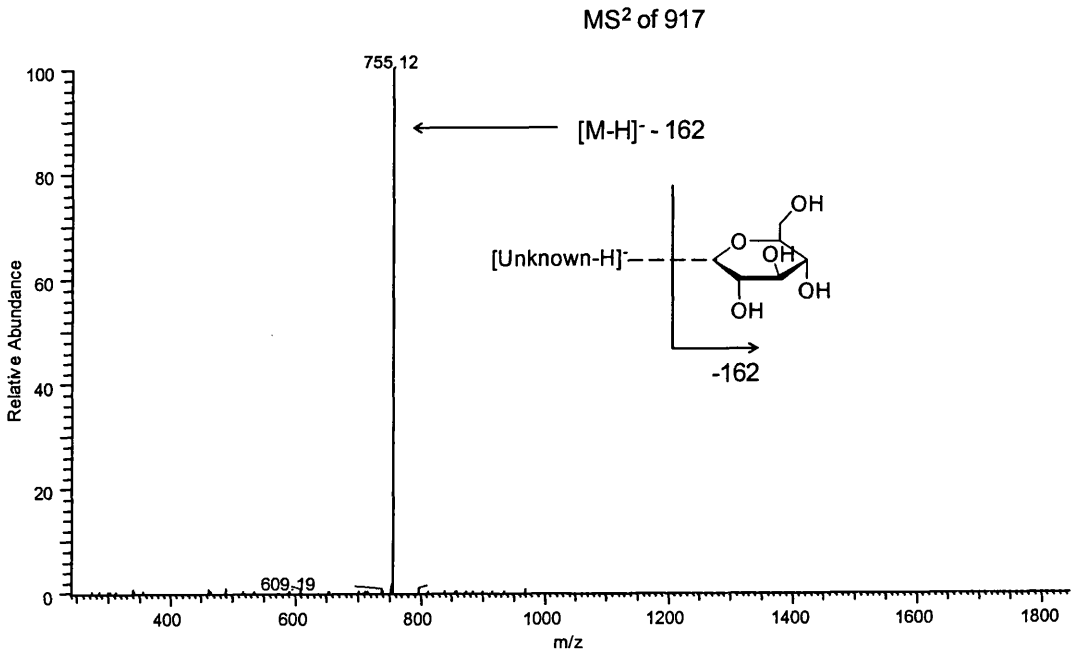


Figure 4.4 MS/MS scan of Peak 1 *m/z* 917

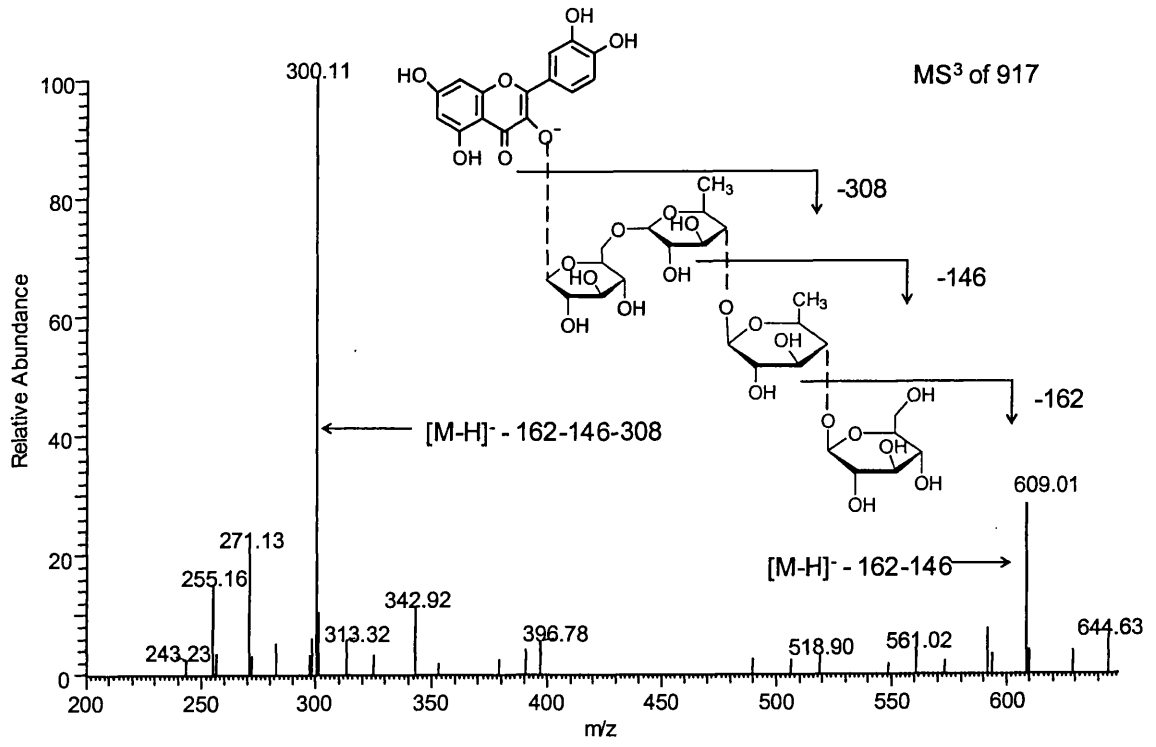


Figure 4.5 MS^3 scan of Peak 1 (Figure 2.33) *m/z* 917 → 755

4.3.1.3 Peak 5

Peak 5 shows a retention time of 21.5 minutes and a base peak of 609, the same base peak mass as Peak 4 (rutin), the MS/MS analysis produced a different fragmentation spectrum compared to the standard rutin, as shown in Figure 4.6, it shows a characteristic fragment ion of *m/z* 300 which indicates that the aglycone is quercetin. The fragment of *m/z* 300 (Figure 4.7) shows an identical spectrum to that obtained with the commercial standard quercetin (refer to Figure 2.12). The compound is therefore considered to be similar in structure to rutin however the sugar moieties must be bound to the quercetin structure in a different manner compared to rutin in order to explain the differences in MS/MS data obtained.

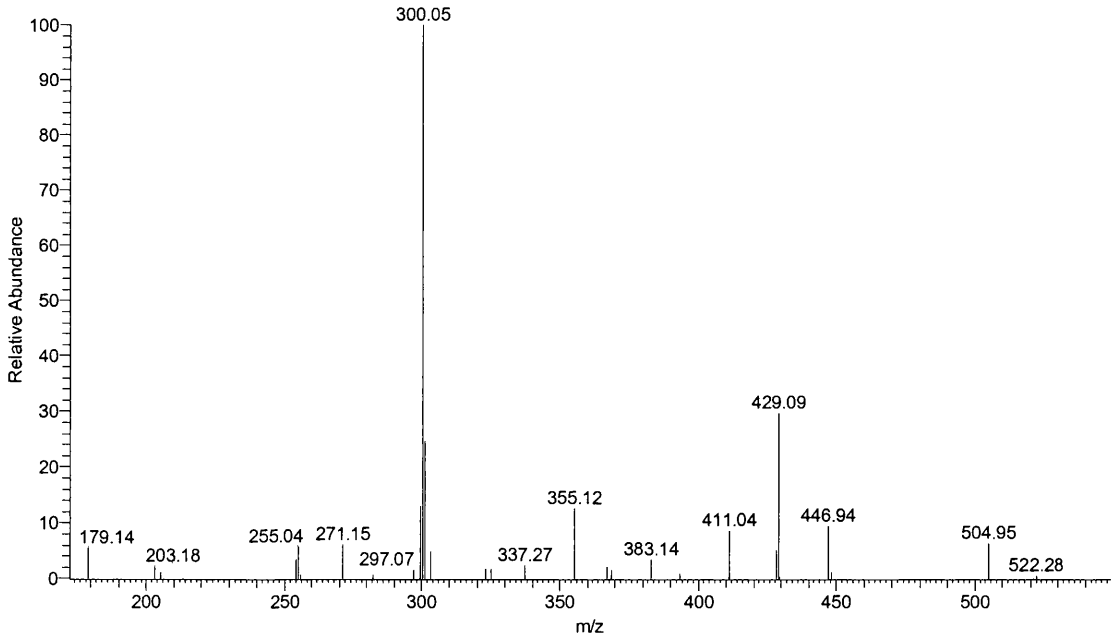


Figure 4.6 MS/MS spectrum of Peak 5

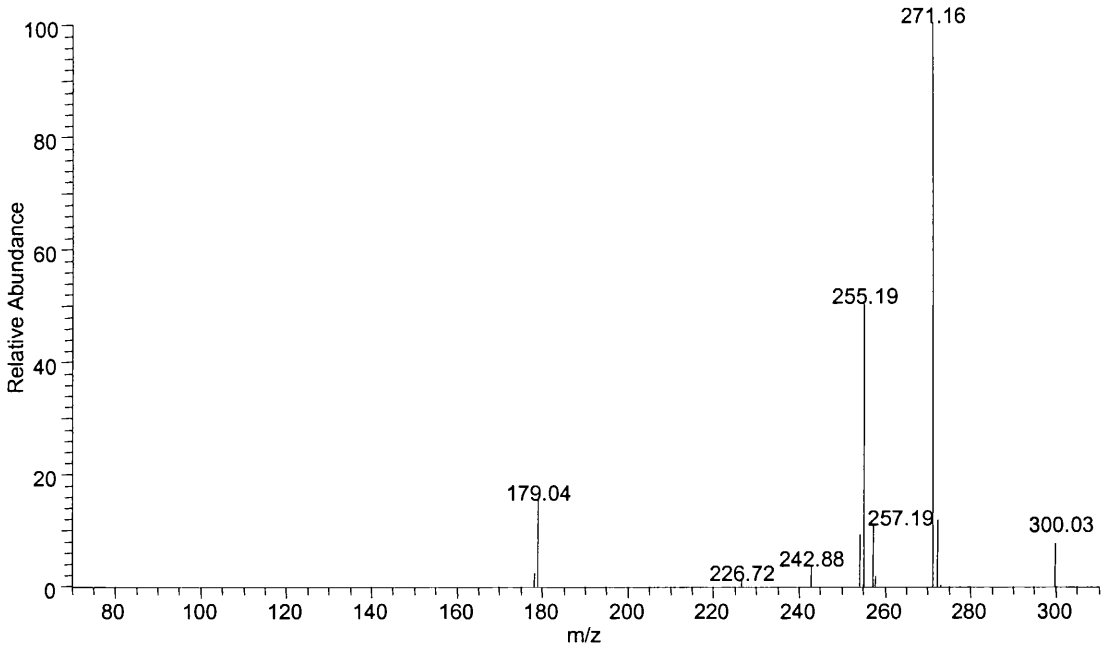


Figure 4.7 MS^3 scan of Peak 5 *m/z* 609→ 300

4.3.1.4 Peak 6

Peak 6 shows a retention time of 22.6 minutes and base peak at *m/z* 593, upon fragmentation, a major product ion *m/z* 285 is obtained representing a loss of 308 amu from previous compounds which is thought to occur via the loss of a rhamnosyl-glucosyl moiety (Figure 4.8). Further MS^3 spectrum of *m/z* 285 is shown in Figure 4.9, and this latter data shows similar major fragment ions to the standard kaempferol (as seen in Figure 2.13), peak 6 is hence tentatively determined to be rhamnosyl-glucosyl kaempferol.

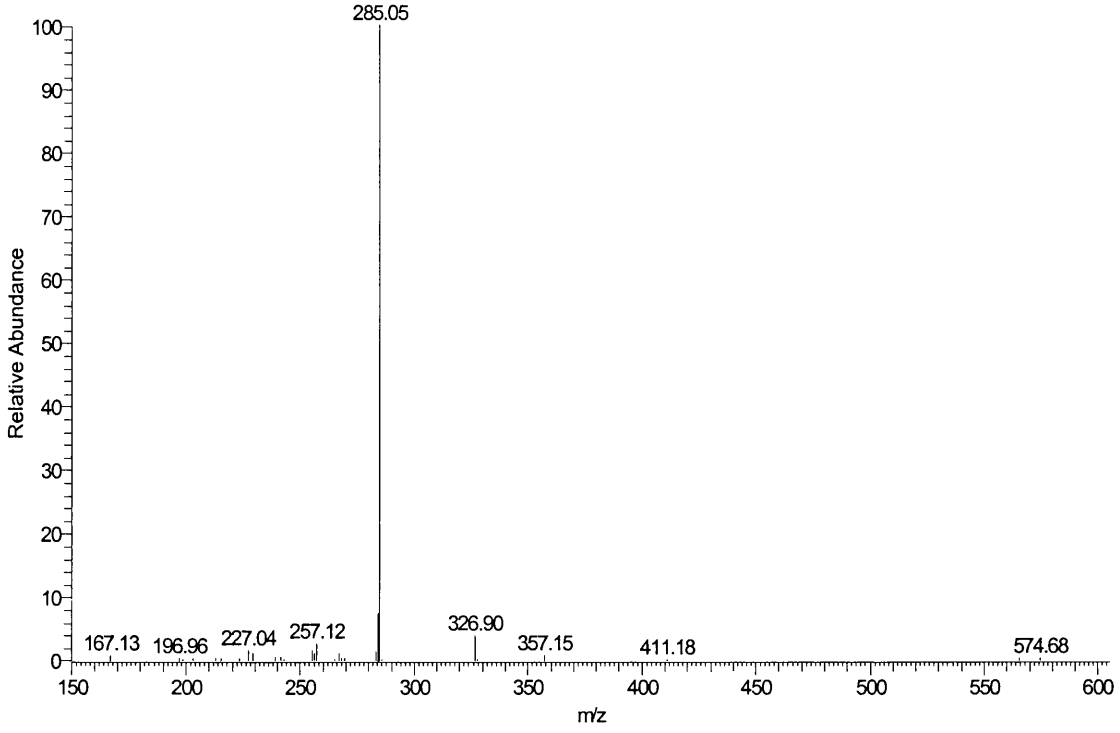


Figure 4.8 MS/MS spectrum of Peak 6 *m/z* 593

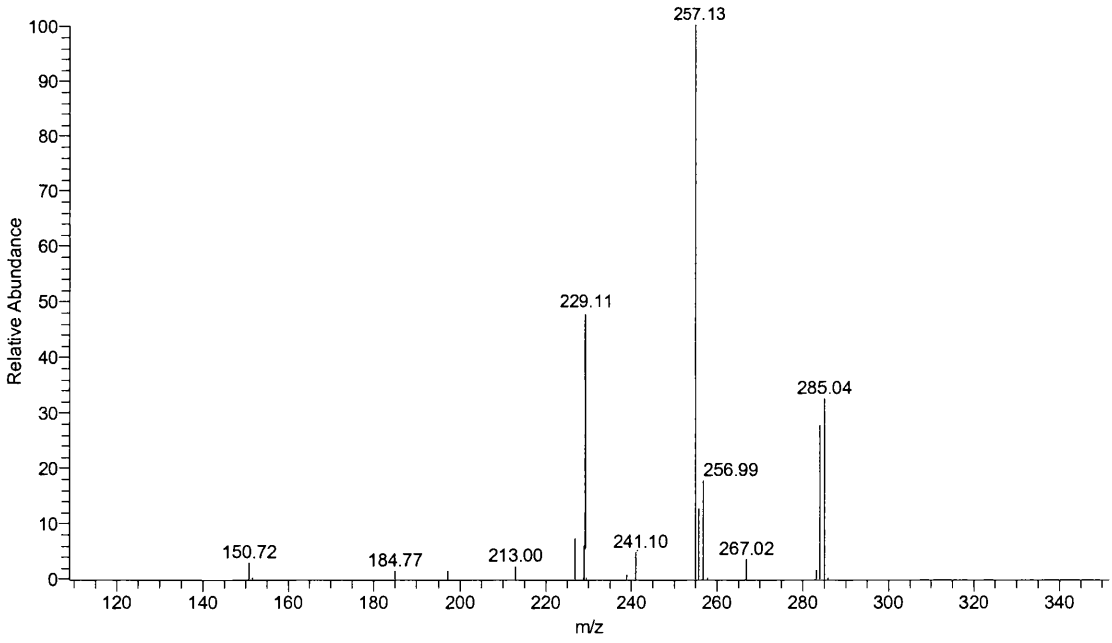


Figure 4.9 MS^3 spectrum of Peak 6 *m/z* 593 → *m/z* 285

4.3.1.5 Peak 7

Peak 7 shows a retention time of 23.5 minutes and base peak at *m/z* 623, upon fragmentation, a product ion at *m/z* 315 can be obtained. Again this is thought to represent a loss of a rhamnosyl-glucosyl moiety (Figure 4.10). Further MS^3 spectrum of *m/z* 315 is shown in Figure 4.11, and the experimental data obtained is identical to the MS/MS spectrum of standard isorhamnetin (Figure 2.14). Peak 7 is therefore thought to be rhamnosyl-glucosyl isorhamnetin.

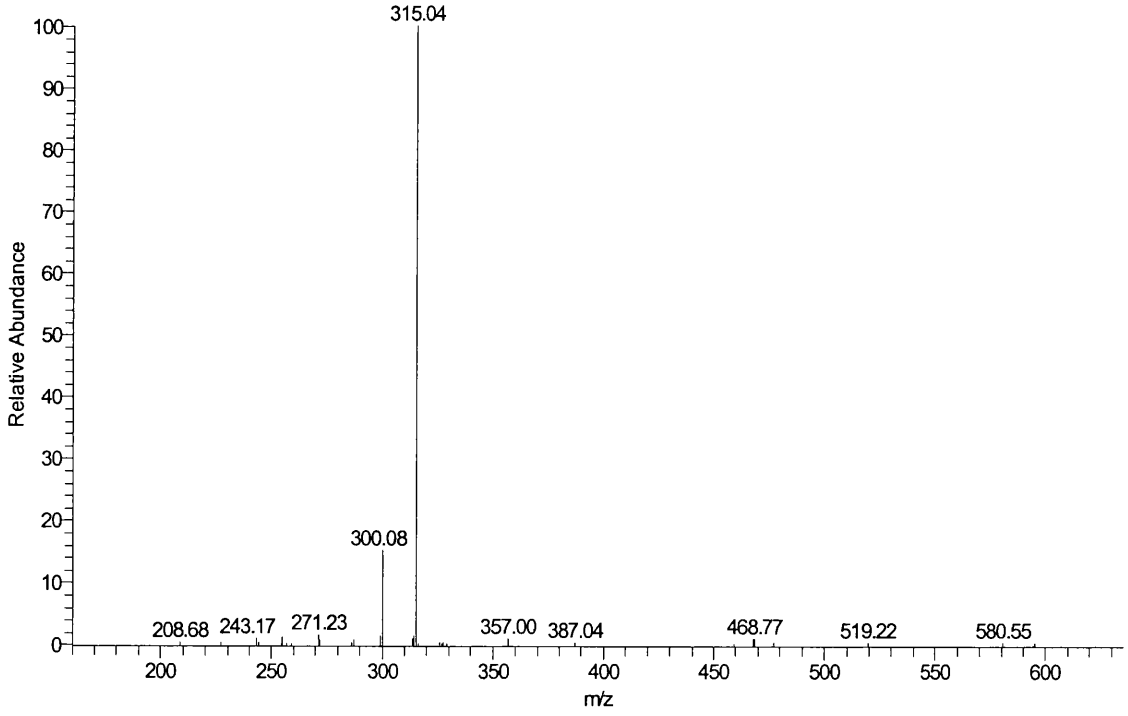


Figure 4.10 MS/MS spectrum of Peak 7, *m/z* 623

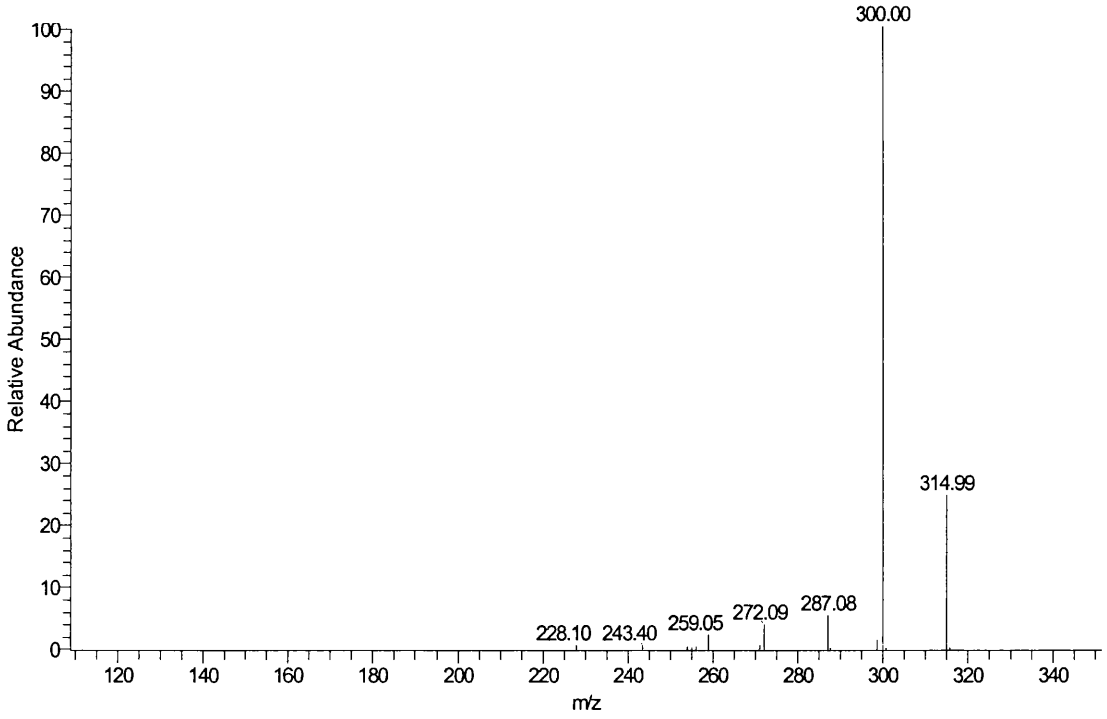


Figure 4.11 MS^3 spectrum of Peak 7, *m/z* 623 → *m/z* 315

4.3.1.6 Peak 10

Peak 10 shows a retention time of 30.5 minutes and a base peak of *m/z* 755. MS^2 of 755 shows a loss of a rhamnosyl moiety causing a mass loss of 146 amu with product ion of *m/z* 609 (Figure 4.12). Further fragmentation of *m/z* 609 shows a typical ion of quercetin with *m/z* 300, the spectrum is identical to the MS/MS data for rutin (rhamnosyl glucosyl-quercetin), so peak 10 can be assigned as 2rhamnosyl glucosyl-quercetin

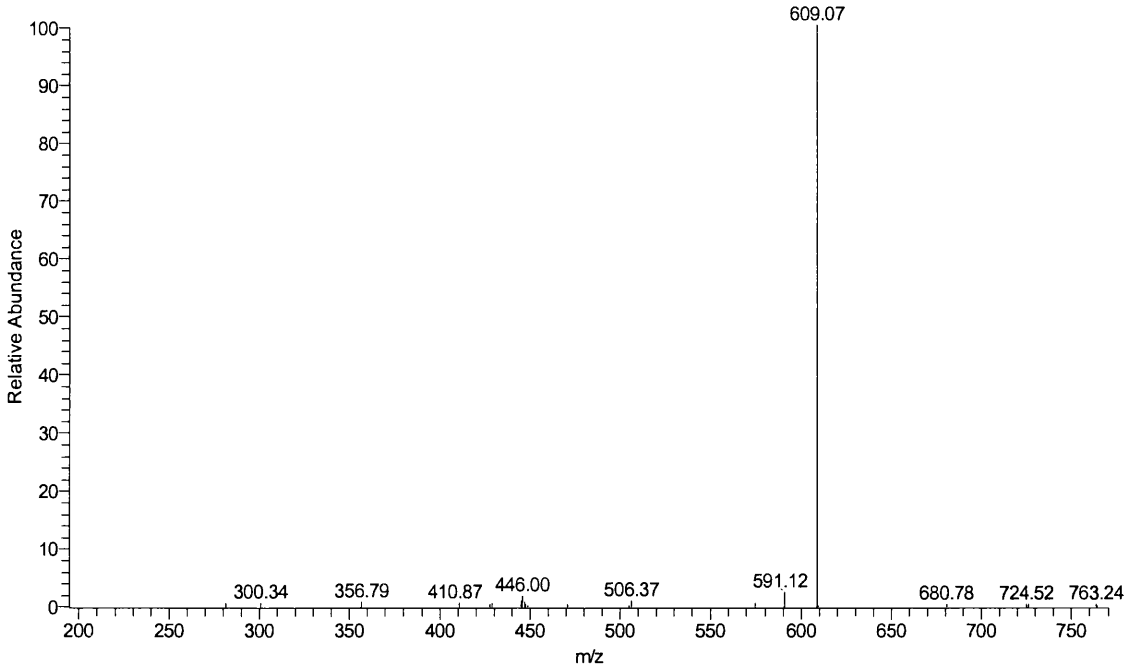


Figure 4.12 MS/MS spectrum of Peak10, *m/z* 755

4.3.1.7 Peak 11

Peak 11 shows a retention time of 36 minutes and base peak at *m/z* 739, upon fragmentation a base peak ion *m/z* 593, resulting from a loss of 146 amu, is obtained (Figure 4.13). Further MS^3 spectrum of *m/z* 593 is shown in Figure 4.14, in this spectrum, *m/z* 284, the negative kaempferol ion with the formation of a free radical species by further loss of one hydrogen without loss of the charge (as seen for quercetin[11]) is base peak ion, so Peak 11 can be assigned as 2rhamnosyl-glucosyl kaempferol.

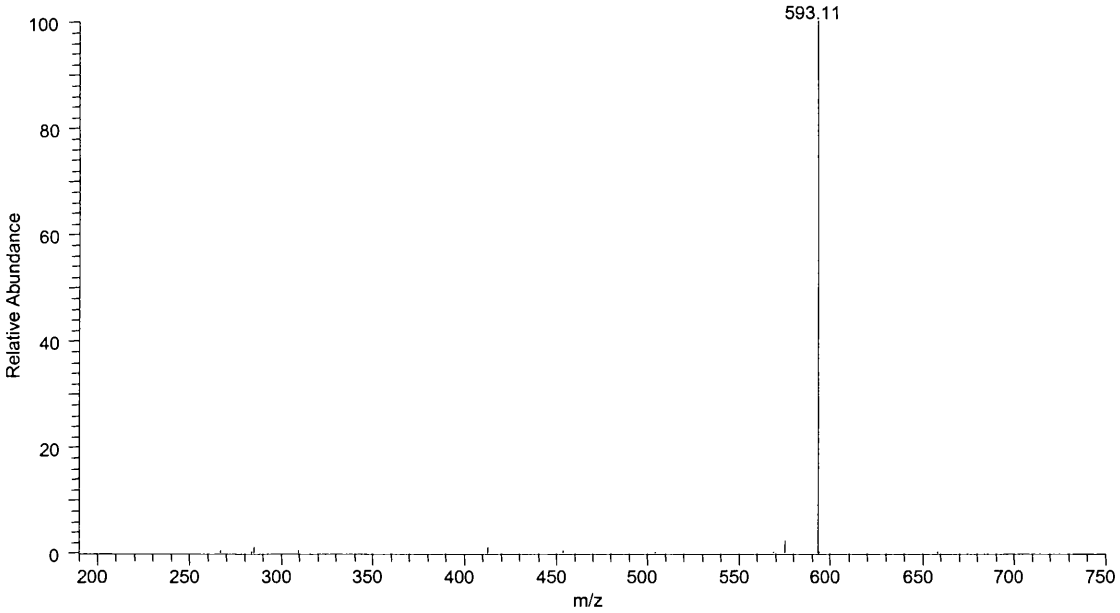


Figure 4.13 MS/MS spectrum of Peak11, *m/z* 739

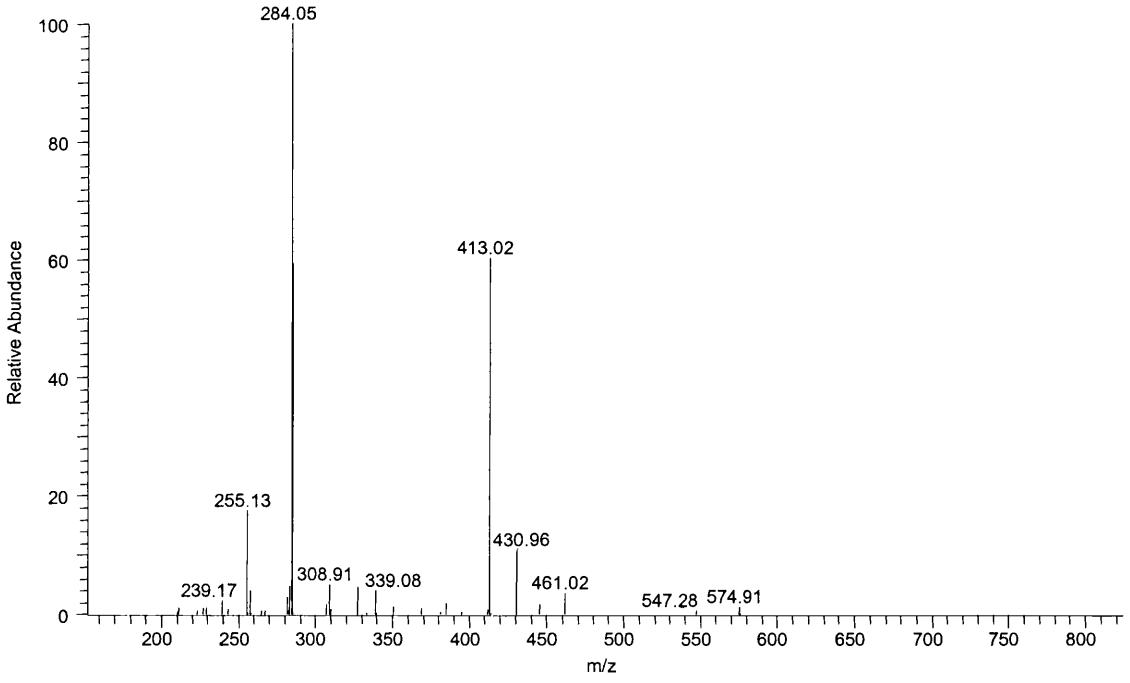


Figure 4.14 MS^3 spectrum of Peak 11, *m/z* 739 → *m/z* 593

4.3.2 LC/MS of *Ginkgo biloba* nutritional supplements by capillary column

An LC/MS system composed of a LC Packings Ultimate Capillary LC system (Dionex, Amsterdam, Netherlands) with LCQ Deca XP ion trap (Thermo Finnigan, San Jose, CA) was also evaluated for the analysis of *Ginkgo biloba* nutritional supplement. The LC and MS conditions are described in the method section, under these conditions, the standard reference materials exhibited satisfactory separation. The base peak MS full scan of the *Ginkgo biloba* extract itself shows more than 30 peaks (Figure 4.15) compared to the 11 peaks detected by the conventional bore column, this indicates that capillary column can obtain better sensitivity than the normal size column.

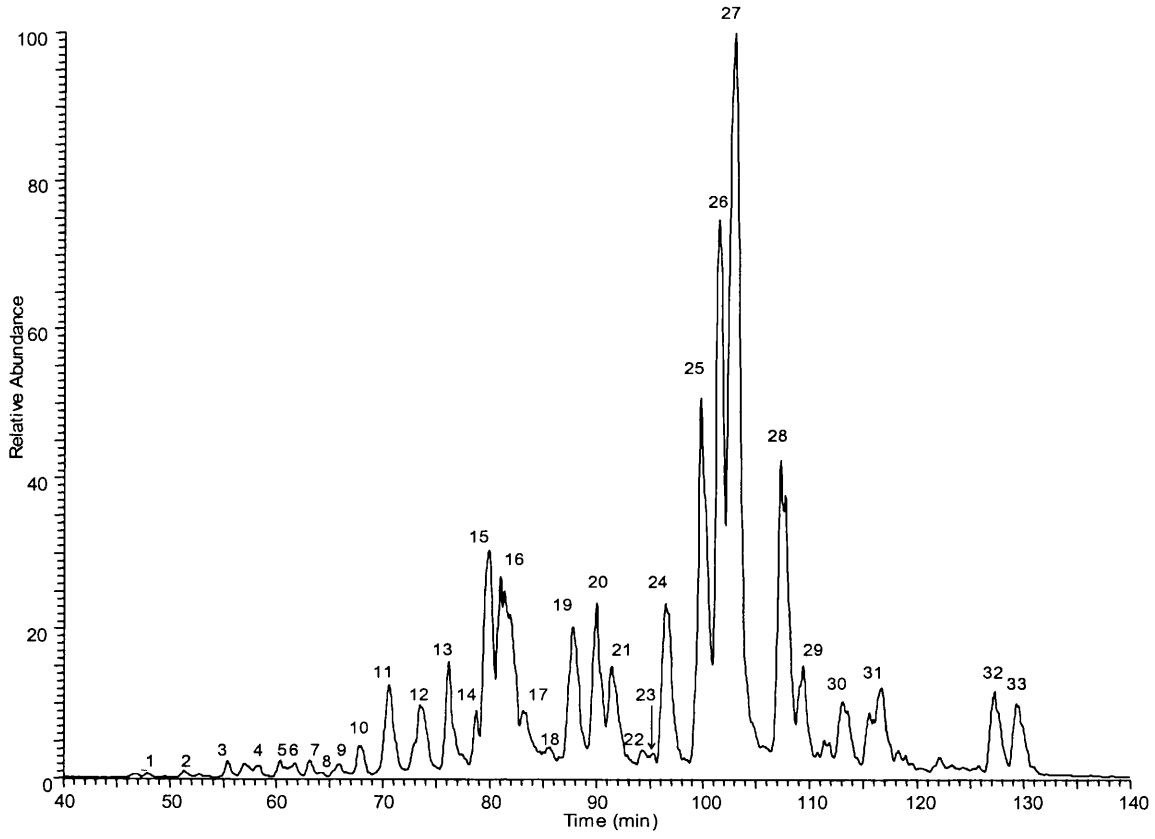


Figure 4.15 Full scan base peak of capillary LC/MS of *Ginkgo biloba* nutritional supplement sample 3

Upon further investigation, using the principle described for the identification in 4.6mm normal-bore column, the tentative identification can be obtained from information provided by the data-dependent analysis by capillary column. Table 4.3 shows the MS and MS^n of the peaks at different retention time and the proposed identification. The four terpene lactones (BL, GA, GB, GC) can be identified by comparing the full scan and MS^n data with standards, among these, GA, GB, and GC were detected with the dimer as the base peak. Another terpene lactone ginkgolide J (GJ), with *m/z* of 423 can be observed at retention time 77.7 minutes, GJ was detected from the mobile phase derived acetic acid adduct as the base peak. The MS/MS of *m/z* 423 is shown in Figure 4.16, the similar fragment ions of GB (379, 361, and 305) can be observed. The major components are identified as flavonoids, 35 flavonoids can be detected in this analysis, among these, 12 flavonoids are originated from quercetin, 10 flavonoids are originated from kaempferol, and 5 flavonoids are originated from isorhamnetin. 2 flavonoids are thought to originate from myrincetin (*m/z* 317) and 1 from methylmyrincetin (*m/z* 331). Peak 8, 10, 18, 23 are thought to be flavonoids originated from unknown aglycones with *m/z* of 345, 521, 417. Also Peak 22 (b) is recognized as aglycone with *m/z* of 285 from the MS/MS data however the identity of the sugar bound to it is unknown at present.

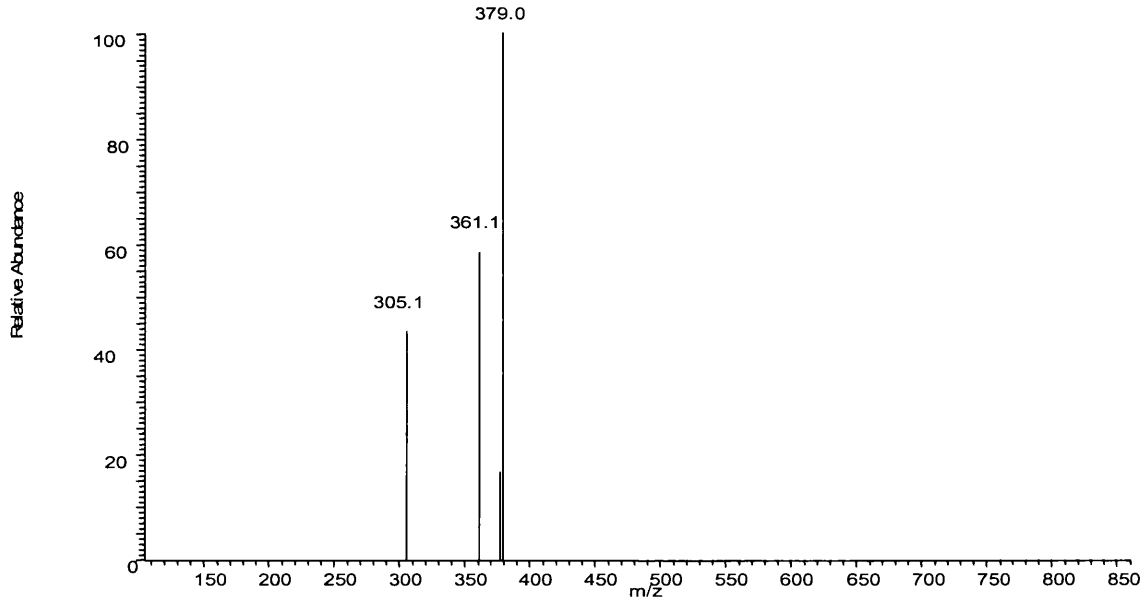


Figure 4.16 MS/MS spectrum of Peak 13 *m/z* 423 at retention time of 77.7 minutes

The flavonoids identified in this capillary LC/MS study differ from Hasler's HPLC fingerprint in which 33 flavonoids were identified[7]. This is thought to be because the detector utilised is different, UV is more sensitive to some flavonoids and mass spectrometry is sensitive to others and so these two should be deemed as complimentary. This LC/MS separation using capillary column obtains more peaks than Hasler's reports and by using data-dependent scan mode, the mass and fragment pathways can be recorded at the same time so more information is provided. An advantage of the LC/MS approach is the ability to also study the levels of the terpene lactones which are not readily detected by HPLC in combination with UV.

Table 4.3 MS, MS/MS and MS^n of the peaks in Figure 4.18 and the proposed identification

Peak	RT	MS	MS^2	MS^3	Identification [a]
1	47.8	917	755	301	Q+2G+2R
2	51.3	901	739	593, 285	K+2G+2R
3	55.6	771	609	301	Q+2G+R
4	58.0	801	639	331	Methylmyricetin+2G+R
5	60.1	771	609	301	Q+2G+R
6	61.6	755	593	285	K+2G+R
7	62.9	785	623	315	I+2G+R
8	64.3	815	653	345	Unknown flavonoid+2G+R
9	66.0	755	593	285	K+2G+R
10	67.7	683	521		Unknown flavonoid+G
11	70.4	755	300	271, 255, 179, 151	Q+G+2R
12 (a)	72.9	625	316	287, 271, 179, 151	Myricetin+R+G
12 (b)	73.3	917	755	609,300	Q+2G+2R
12 (c)	73.5	325	281, 251, 163		BL
13 (a)	76.1	739	593	285	K+G+2R
13 (b)	76.8	769	315	300	I+2R+G
13(c)	77.7	483 [b]	423	379,361,305	Ginkgolide J
14	78.8	901	739	593,285	K+2G+2R
15	79.7	879 [c]	439	395, 383, 365	GC
16	81.5	609	301	271, 179, 151	RH
17 (a)	83.2	639	331	316, 315	Methymyricetin+G+R
17 (b)	83.2	463 [d]			QG
18	85.5	579	417	402, 371, 205, 181, 166, 151	unknown flavonoid+G

Continues

Table 4.3 continued

Peak	RT	MS	MS^2	MS^3	Identification [a]
19	87.8	609	301	271, 255, 179, 151	Q+G+R
20	90.0	593	285	267, 257, 229, 151	K+G+R
21 (a)	91.3	623	315	300	I+G+R
21 (b)	92.8	447	259,241,215,151		QH
22 (a)	93.9	477	315	300, 285, 271, 243	I+G
22 (b)	94.2	913	739	593, 285	K+R+unknown sugar (*m/z* 174)
23	95.3	653	345	330, 315,	Unknown flavonoid+G+R
24	96.5	593	285	267,255, 229, 151	K+G+R
25	99.8	755	609	300, 271, 255	Q+G+2R
26	101.5	815 [c]	407	379, 363, 351,	GA
27	103.0	847 [c]	423	395, 367, 305	GB
28	107.2	739	593	285	K+G+2R
29	109.3	755	609	300, 271, 255	Q+G+2R
30	113.2	301	271, 255		QD
31	116.7	739	593	285	K+G+2R
32	127.3	285	257		KF
33	129.3	315	300	283, 255, 227,151	IR

[a] Q: quercetin, K: kaempferol, I: isorhamnetin, G: glucosyl R: rhamnosyl

[b] acetic acid adduct

[c] dimer

[d] no MS/MS obtained because mass spectrometer fragment the coeluted high abundant flavonoid

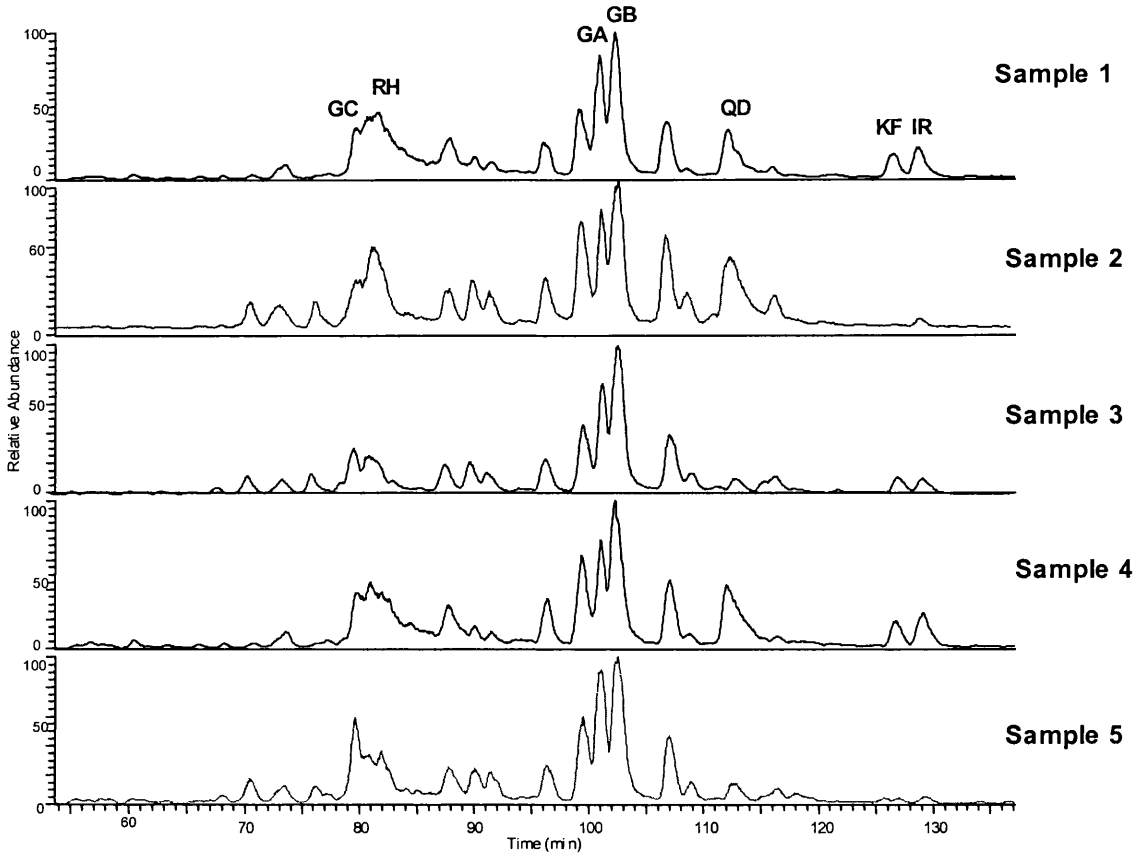


Figure 4.17 Bask peak full scan of the five *Ginkgo biloba* commercial samples using capillary column LC/MS/MS in data-dependent mode

Figure 4.17 shows the fingerprint profile of the five *Ginkgo biloba* commercial samples, the three peaks in the middle including GA and GB are similar in all the samples and can be used as reference peaks to evaluate the other components by comparing the areas of these peaks. The QD/GA ratio in sample 2 is much higher than other samples, also RH/GA ratio in sample 1 is higher than other samples, which is consistent with our previous quantitative results. This indicates that the LC/MS fingerprint is a useful technique in quality control of the plant extract, in our previous study[12] the quantitation of both flavonoid glycosides and flavonoid aglycones is able to distinguish the spiked fortification (QD in sample 2 and RH in sample 1) which can not be detected by

tradition quality control, but this fingerprint study can detect the fortification with no need of reference standards and this is especially useful when reference standards are unavailable.

It was considered that methanol extraction has the potential to extract many components from *Ginkgo biloba*, not just the active components flavonoids and terpene lactones that are the bioactive compounds of interest. To have a view of these compounds more specifically, different data processing techniques can be utilised. According to the fragmentation pathways of flavonoid glycosides previously described, all flavonoid glycosides will fragment to their aglycones upon fragmentation. So the extract ion chromatogram (XIC) of flavonoid aglycones *m/z* 301, 285, 315 can be used to analyse the flavonoid glycoside components in the sample. More than 10 peaks can be seen in each *m/z* range which indicates that *Ginkgo biloba* is a complex mixture of various flavonoid glycosides. Since no reference standard for the flavonoid glycosides are available, the method is not able to precisely identify each of the flavonoids, however, Figure 4.18 shows a profile of the flavonoids in *Ginkgo biloba* extract and how many of the flavonoids come from quercetin, kaempferol and isorhamnetin, respectively. This experiment was performed with an ion trap mass spectrometer in data-dependent mode. Using triple quadrupole mass spectrometers, it is possible to perform a precursor ion scan, in which the product ion can be specified and all the precursor ions that produce this specified ion under fragmentation will be recorded. Hence for a further development of the experimental analysis, a quadrupole mass spectrometer can be used to create a profile of all the flavonoids that give a product ion of quercetin, kaempferol and isorhamnetin, however, terpene lactones can not be detected by using this precursor ion scan.

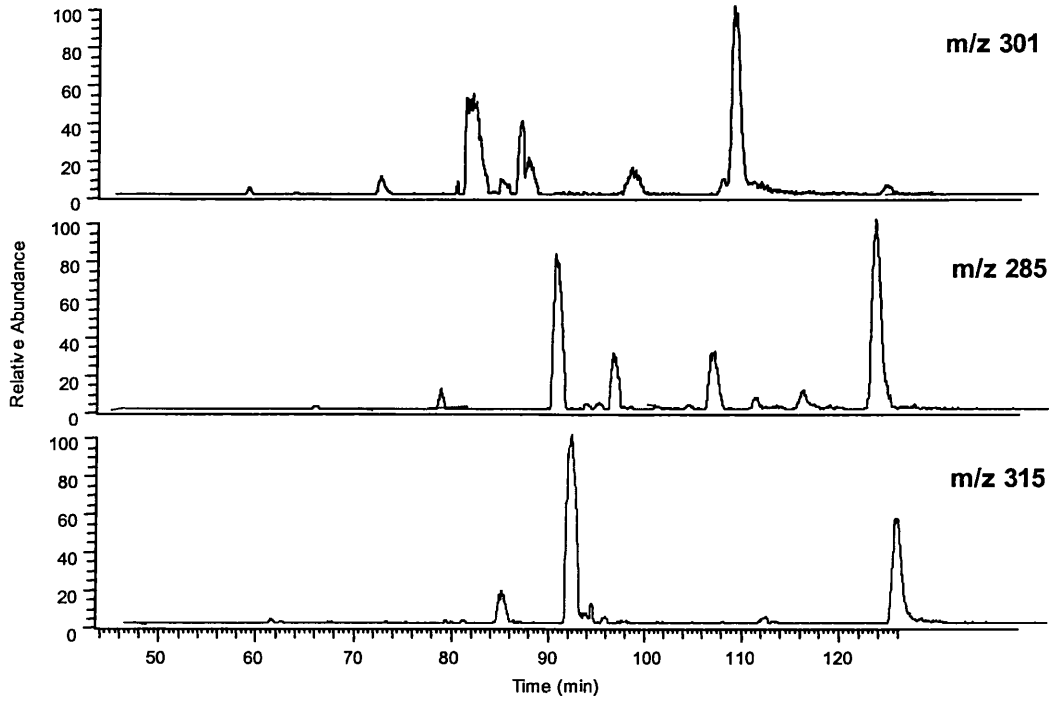


Figure 4.18 Capillary LC/MS extracted ion chromatogram (XIC) of *Ginkgo biloba* extract at *m/z* 301, 285, 315

Alternatively, since the detected flavonoids are mainly glucoside and rhamnoside conjugated (with additional masses of 162 and 146 Da respectively), constant neutral loss analysis of the characteristic sugar ring can be used to identify the flavonoids from the methanol extract and obtain a profile of which flavonoid aglycones are conjugated to which sugar moieties. As shown in Figure 4.19, by using neutral loss of m/z 162, 146, 308, 454, most of the flavonoids in Table 4.3 are present in this figure, and there are 3 peaks a, b and c, corresponding to 449→287 at 63.98 minutes, 449→287 at 69.44 minutes, 523→ 361 at 72.3 minutes, respectively, which were not detected by using full scan base peak but were observed by constant neutral loss data processing. The data also indicates that there are more glucoside flavonoids than rhamnoside flavonoids in *Ginkgo biloba* extract.

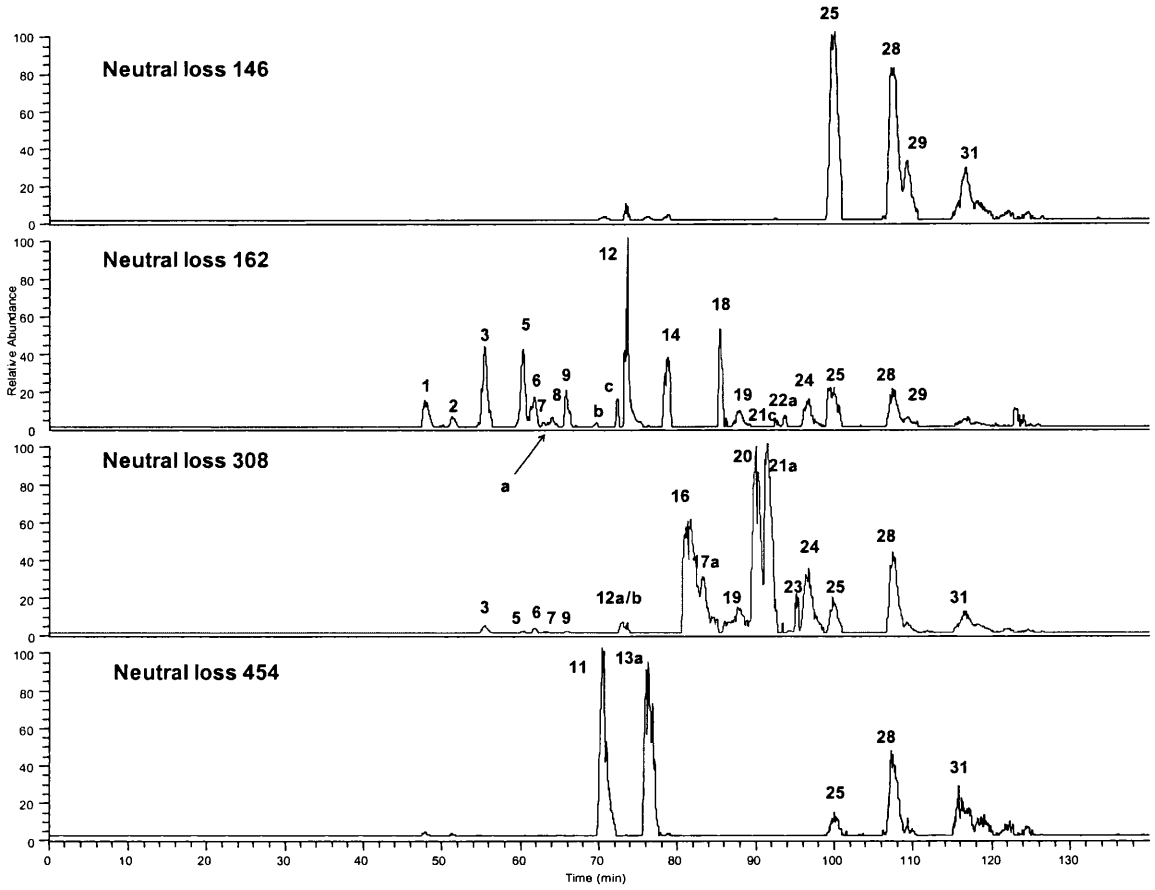


Figure 4.19 Capillary LC/MS extracted ion chromatogram (XIC) of *Ginkgo biloba* extract by neutral loss of *m/z* 162, 146, 308, 454 Da

The same principle can be applied to the terpene lactones. According to Table 2.6, the loss of 44 Da (COO), 74 Da ($2CO+H_2O$) and 162 Da ($2CO+H_2O+2COO$) can be used to profile the terpene lactones. However the neutral loss study of terpene lactones proved to be less successful because the dimer of the terpene lactones or the acetic acid adducted terpene lactones were more abundant in the mass spectra than the de-protonated molecule, and the characteristic fragmentation of the terpene lactones is only observed in the MS^3 scan and the quality and the signal to noise ratio of the observed losses is not as good as MS^2. Reducing the MS scan range could solve this problem of dimer fragmentation, however the increased mass range is required so as not to miss many of the flavonoids.

The neutral losses profile can also be utilized to compare different samples, for example, by comparing neutral loss of *m/z* 308, it can be seen that sample 1 has more rutin than other samples, which indicates fortification. Interestingly, by comparing neutral loss of *m/z* 162 as shown in Figure 4.20, the peaks 3, 5-7, 12, 14 and 18 (as labelled in the figure and listed in Table 4.3) show obvious variations between samples, this is thought to because of the different origins of the extract or storage conditions which are very important in the quality control of plant extracts.

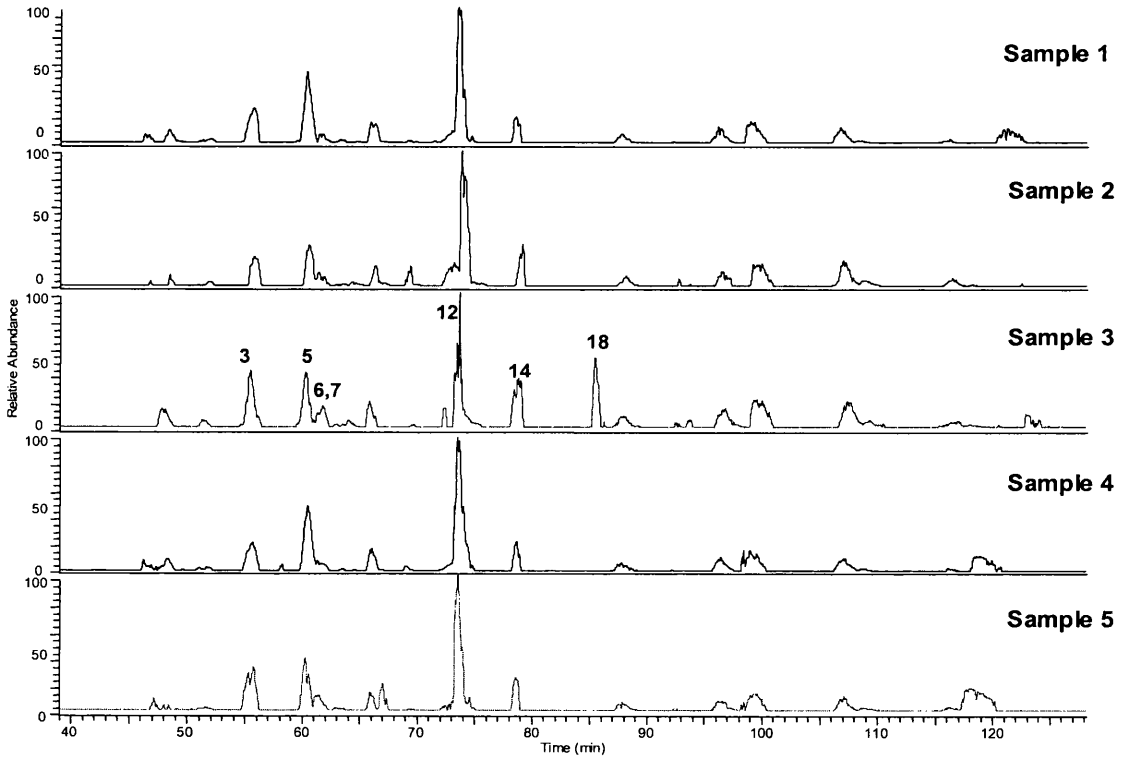


Figure 4.20 Capillary LC/MS extracted ion chromatogram (XIC) of neutral loss of *m/z* 162 Da of five *Ginkgo biloba* commercial samples

4.4 Conclusion

LC/MS is excellent for fingerprint quality control of plant extracts due to its high sensitivity and specificity. LC/MS with data-dependent scan mode not only gives a

chromatogram which shows different peaks, at the same time, the mass and fragmentation of the components can be provided. This is a major advantage over HPLC-UV and TLC methods which are currently the major instrumentations used for fingerprinting analysis. The fingerprint chromatograms of five commercial *Ginkgo biloba* samples show variations in rutin and quercetin concentration, which indicate fortification, and this can not be spotted by using traditional QC method.

Capillary column (i.d. 300 μm) and normal-bore column (i.d. 4.6 mm) were compared in the separation of *Ginkgo biloba* commercial samples. The capillary column shows better resolution and sensitivity than a normal size column. Many more peaks are present in the chromatogram of capillary column separation. According to the fragment pathway of *Ginkgo* flavonoids, the XIC of *m/z* 301, 285 and 315 were constructed and it shows that more than 30 flavonoids can be detected by capillary column, neutral loss also provides important information regarding identification and evaluation the similarity and stability among the samples.

Comparing precise quantitation and fingerprinting in quality control of *Ginkgo biloba* extract, precise quantitation is more accurate and the exact amount of components of interest can be obtained, but limited number of components are under investigation mainly due to the limited availability of commercial standards, essential for accurate quantitative analysis. Fingerprint analysis can generate an overview of all the components in the sample under investigation. By comparing the peak area with standard fingerprint, semi-quantitative data can be obtained. The fingerprint method is therefore especially useful when reference standards for all components can not be obtained and be quantitated. However, chromatographic fingerprints sometimes exhibit

variations in peak height and retention time of a given sample running through identical columns even under the same separation conditions. Therefore proper normalization of chromatographic fingerprints must be taken into consideration[13]. These two methods are thought to be complimentary in the quality control of plant extract.

It should be noted that in terms of the identification of bioactive components from *Ginkgo biloba* extracts by ion trap mass spectrometry no accurate mass was available. However by comparing the retention time and MS/MS fragments with the commercial standards, the identification of the 10 standards in the *Ginkgo* extracts can be confirmed. For those with no reference standards, the investigator is able to elucidate the aglycone from which the flavonoids originated, what sugar and how many sugars were added to the aglycone; but how the sugars are organized is unable to be decided from the MS/MS data. Further study may benefit from the collection of the LC fractions of the components of interest and characterization of these components with the help of other analysis methods such as NMR and X-ray. Although more than 30 peaks were resolved, the co-elution of the components was observed in this study, therefore slower gradient and longer elution time can be applied for better separation and more unknown identification.

Capillary column shows a better sensitivity and resolution than the normal-bore column, so in the next chapter it is used for urine analysis with the online SPE process by which time-consuming manual sample preparations can be avoided.

4.5 References

1. Xue, T.H. and Roy, R., Science, **300**, 740 (2003)
2. World Health Organization, *Guidelines for the Assessment of Herbal Medicines*, Munich, 28.6.1991, WHO, Geneva, (1991)
3. Kerns, E.H., Volk, K.J., Whitney, J.L., Rourick, R.A and Lee, M.S., J. Drug Inf., **32**, 471 (1998)
4. Committee of National Pharmacopoeia, *Pharmacopoeia of PR China*, Press of Chemical Industry, Beijing, (2000)
5. UNICEF/UNDP/World Bank/WHO, *Handbook of non-clinical safety testing*, (2004)
6. Cai, Z., Lee, F.S.C., Wang X.R. and Yu W.J., J. Mass Spectrom., **37**, 1013 (2002)
7. Hasler, A. and Sticher, O., J. Chromatogr., **605**, 41 (1992)
8. Mauri, P., Migliazza, B. and Pietta, P., J. Mass Spectrom., **34**, 1361 (1999)
9. Liu, C., Mandal, R., Li, X.F., Analyst, **130**, 325 (2005)
10. Ying, T., Zhao, L., Lin, Z., Wang, G. and Xiang, B., J. Chromatogr. Sci, **42**, 177 (2004)
11. Hvattum, E. and Ekeberg, D., J. Mass Spectrom., **38**, 43 (2003)
12. Ding, S., Dudley, E., Plummer, S., Tang, J., Newton, R.P. and Brenton, A.G., Rapid Commun. Mass Spectrom., **20**, 2753 (2006)
13. Draˇsar, P. and Moravcova, J., J. Chromatogr. B, **812**, 3 (2004)

Chapter 5

Determination of active components of *Ginkgo biloba* in human urine by capillary HPLC/MS with column switch on-line purification

5.1 Introductions

Terpene lactones and flavonoids are two active species in *Ginkgo biloba*[1]. Terpene lactones are associated with increased circulation to the brain and other parts of the body and may exert a protection action of the nerve cells[2]. The terpene lactones found in *Ginkgo biloba* are ginkgolides (ginkgolides A, B and C) and bilobalide. *Ginkgo biloba* also produces a large number of flavonoids in leaves, mainly derivatives of isorhamnetin, kaempferol and quercetin. Flavonoids in *Ginkgo biloba* are mainly flavonols which also occur naturally in many other plants. The flavonoids show biological properties through their free radical-scavenging antioxidant activities and metal-ion-chelating abilities[3]. Despite the benefits of these components, their bioavailability after oral administration is considered to be a limiting factor[4].

Flavonoids are well studied because of their ubiquitous existence. After ingestion, flavonoid glycosides are thought to be first hydrolyzed by microorganisms in the gastrointestinal tract to aglycones. The liberated aglycones can be absorbed from the intestinal wall and excreted in the urine and bile as glucuronides and sulphate conjugates[5]. A number of analytical techniques have been utilized to evaluate the metabolism and bioavailability of flavonoids *in vitro* and *in vivo*. The methods utilised previously for the study of urinary and other biologically important flavonoids levels include HPLC[6, 7, 8], and mass spectrometry (GC/MS[9, 10] and LC/MS[11]). Study of terpene lactones has been more limited to the study of plasma levels for which LC/ESI MS[12] and LC/APCI MS[13] have been reported. GC/MS was also report for terpene lactone analysis[14].

Mass spectrometry is one of the most effective techniques for the analysis of complex mixtures in biological samples because of its high sensitivity, specificity, and easy combination with chromatographic techniques. A common limitation of the previously developed techniques is that all of the methods reported require either a solid-phase extraction or a liquid-liquid extraction purification step prior to analysis, both very time-consuming. The techniques also examined only one of the two sets of compounds (flavonoids or terpene lactones); in contrast a previous HPLC/MS method developed by our group allowed the detection of both sets of compounds from the supplement itself by conventional bore HPLC/MS[15]. Although this method worked well for the analysis of the commercial herbal extract the analysis of the same components excreted in urine is complicated due to their lower levels and the large number of other high abundant species also excreted.

5.2 Aims of study

To evaluate the influence of flavonoids and terpene lactones of *Ginkgo biloba* and its beneficial health effects, it is important to monitor the concentration of the *Ginkgo* terpene lactones and flavonoids occurring in biological samples. Whilst pharmacological activities of *Ginkgo biloba* flavonoids and terpene lactones have been extensively studied, there has been little investigation into analytical methods for determination of terpene lactones and flavonoids simultaneously in such samples. In this study we developed a single quantitative assay for both flavonoids and terpene lactones in human urine by HPLC/MS with on-line cleanup of the urine sample. As a "proof of principle" study the assay was validated, shown to be sensitive, accurate and

also function as well if not better than HPLC/MS detection with off-line purification as previously reported[15].

The separation and quantitation of flavonoids and terpene lactones was achieved by using an in-house made capillary column, the high sensitivity of the capillary column and the online SPE with column switching technique require minimum sample pre-treatment. Different sample treatments, solid-phase extraction, liquid-liquid extraction and online solid phase extraction using a trap column were compared.

5.3 Experimental

5.3.1 Chemicals and standards

HPLC grade solvents: methanol, acetonitrile, formic acid and acetic acid were purchased from Fisher chemicals (Loughborough, UK) and were used without further purification. Ginkgolide A (**GA**), ginkgolide B (**GB**), ginkgolide C (**GC**), bilobalide (**BL**), quercetin dehydrate (**QD**), quercetin-3-β-D-glucoside (**QG**), quercetin-3-rhamnoside (**QH**), kaempferol (**KF**), isorhamnetin (**IR**), rutin (**RH**), andrographolide (**Internal standard**) and β-glucuronidase/sulfatase (G0876) were purchased from Sigma (St. Louis, MO, USA). Water was purified with a Milli-Q deionisation unit (Millipore, Bedford, MA, USA). Gases used included oxygen free nitrogen and helium which were purchased from BOC Ltd (Surrey, UK). Stock solutions were prepared from 1 mg of each of the ten standards on a 0.0001 mg balance (Sartorius, UK) and dissolved in 5 mL methanol to give a final concentration of 200 μg/mL. All solutions were placed in an ultrasonic bath for 10 minutes to ensure

they were completely dissolved. The quantitation standards (working solutions) were accurately diluted with methanol just prior to use. All solutions were stored at -20 °C.

5.3.2 Collection and preparation of urine samples

Urine samples were collected from five volunteers before and 4 hours after the ingestion of a single tablet of *Ginkgo biloba*, the tablet was stated to contain 28.8 mg flavonoids and 7.2 mg terpene lactones. The blank (before) samples were used for spiking experiments to generate validation data. Urine samples were kept in a freezer at -20 °C. Before analysis the urine aliquots were centrifuged at 14000 g for 10 minutes in order to remove particulate matter and the supernatant was taken for analysis. 100 μL of urine aliquot were transferred to a 1.5 mL Eppendorf tube, then 25 μL of 1M sodium acetate buffer and 1 μL of enzyme (containing 114 units of β-glucuronidase and 3.3 units of sulfatase) were added and incubated for 1 h at 37 °C. A 10 μL aliquot of internal standard was added to each sample and 5 μL of the sample were injected directly onto the HPLC/MS system.

5.3.3 Preparation of analytical column and online trap column

The analytical column and online trap column used in this experiment were prepared in-house and were both capillary columns of internal diameter of 320 μm, with lengths of the HPLC column and online trap column 15 cm and 3 cm, respectively. Two different solid phases C_{18} (PepMap stationary phase, Dionex, UK) and C_{30} (YMC, UK) were used to pack the columns tested in this study. Each stationary phase was prepared by placing ~5 mg of resin in a small vial, and adding 1 mL of isopropanol and a mini

stirring magnet to each. A slurry packing procedure described in Section 4.2.3 was employed to prepare the columns.

5.3.4 Evaluation of compound trapping during column switching

A column switch pattern was designed using an Ultimate / Switchos / Famos pump and valve switching system (Dionex, Amsterdam, Netherlands) in order to record both the passing of sample from the trap to the waste and the elution of the sample retained by the trap column. The configuration is shown in Figure 5.1. Initially, Valve A and B are in the position 1-2 and so sample injected by autosampler will enter Port 4 of Valve A and follow the solid line and enter Port 8 of Valve B, so that the components that cannot bind to the trap column and washed away to the waste can be monitored by the mass spectrometer. After 3 minutes, Valve A and B are switched to position 1-10, components retained on trap column are then eluted using high organic composition mobile phase from the Ultimate pump and all the components that bind on the trap column can again be monitored by mass spectrometry (following the dashed lines). By using this method, all the components before and after the switch of the valve can be monitored simultaneously.

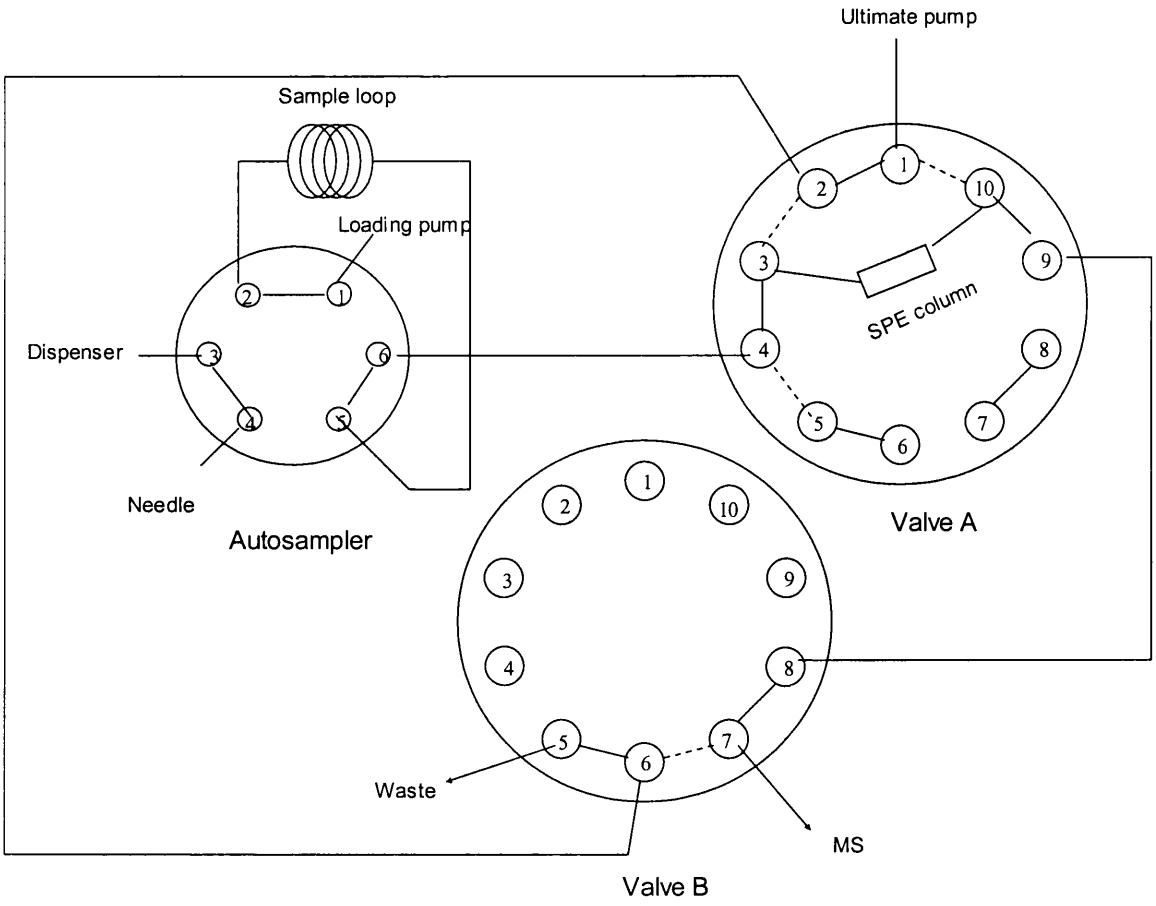


Figure 5.1 Schematic illustration of online solid-phase extraction:

(a) Valve A and B in position 1-2, the components wash though the precolumn were monitored by mass spectrometry

(b) Valve A and B in position 1-10, components trapped on precolumn was eluted and monitored by mass spectrometry

5.3.5 LC/MS conditions

The HPLC/MS system was composed of an LC Packings Ultimate Capillary HPLC system with a Switchos loading pump and FAMOS autosampler (Dionex, Amsterdam, Netherlands) and an LCQ Deca XP ion trap (Thermo Finnigan, San Jose, CA). The column used for the separation was a C_{18} capillary column. The mobile phase was

composed of 0.1% (v/v) acetic acid in water (A), 1:1 (v/v) mixture of ACN and methanol (B), the flow rate was 4 μL/min and the gradient was based on a previous separation developed by our group[15]. The gradient used for the analytical separation is shown in Table 5.1. The loading pump was run at 30 μL/min 0.1 %(v/v) acetic acid in water and 5μL were injected per sample. The electrospray settings is shown in Table 5.2. The mass spectrometer was used in negative mode and selected ion monitoring of the $[M-H]^-$ ion was used to monitor each component analysed. The setup of the column switching system is shown in Figure 5.2.

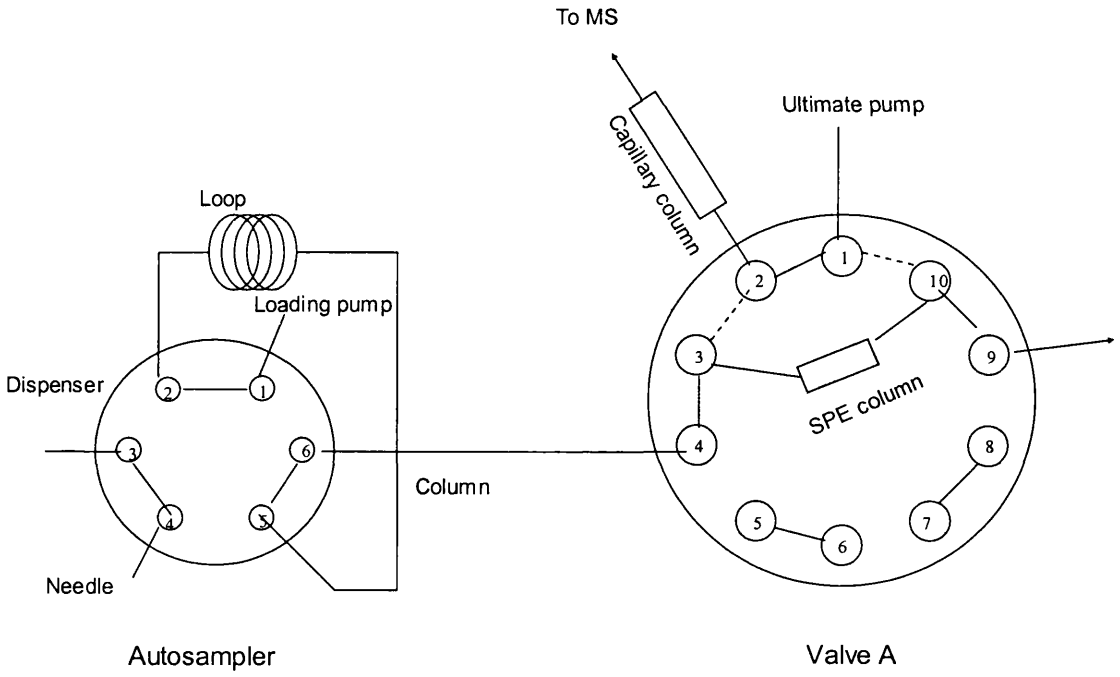


Figure 5.2 Illustration of column switch LC/MS. The components to be analysed were retained on the C_{18} precolumn while the salt in the sample went to waste. After 3 minutes, Valve A was switched and components trapped on the precolumn were eluted to C_{18} analytical column and detected by mass spectrometry

Table 5.1 Gradient elution program used in capillary column separation

Time (min)	A%	B%	Flow rate (mL/min)	Valve A position
0	90	10	4	1-2
3	90	10	4	1-10
120	40	60	4	1-10
125	2	98	4	1-10
135	2	98	4	1-10
138	90	10	4	1-2
140	90	10	4	1-2

Table 5.2 Mass spectrometry setting of capillary LC/ESI-MS

Parameter	Negative ion mode
Sheath gas flow (arbitrary units)	60
Auxiliary gas flow (arbitrary units)	10
Spray voltage (kV)	2.5
Capillary temperature (°C)	200
Capillary voltage (V)	15
Tube lens offset (V)	60

5.3.6 Offline solid-phase extraction (SPE) and liquid -liquid extraction (LLE)

For the off-line solid-phase extraction experiment Bond Elute C_{18} SPE cartridges (Varian, Oxford, UK) were conditioned by sequentially passing 1 mL of methanol and 1 mL of 0.1% HAC through under gravity. After applying the sample (1 mL blank urine spiked with standards) the cartridge was washed with 1 mL of 0.1% HAC and then eluted with 1mL of methanol. The eluent was blown to dryness with a gentle stream of

nitrogen before being re-suspended and analysed by HPLC/MS without online trapping as described previously[15]. Before LLE, 1 mL of blank urine was again spiked with standards, 3 x 2 mL of ethyl ester was next added to extract the reference standards from urine sample. A rotating mixer was used for approximately 30 minutes to facilitate the extraction. The sample was centrifuged at 2000 g for 5 minutes, the ethyl ester layer was transferred to a glass vial and concentrated to dryness with nitrogen stream. The residual was then dissolved in the proper volume of methanol and analysed by HPLC/MS as for the offline SPE experiment.

5.4 Results and discussions

5.4.1 Test of reference standard retention of different in-house made SPE columns

The contaminating components of the urine samples (such as inorganic salts etc.) have the possibility of interfering with the combined HPLC and mass spectrometry analysis. In this study, an online trap column was used to preconcentrate the components of interest whilst the contaminants were washed to waste. The trapped components of interest could then be eluted onto the analytical column and separated before mass spectrometric detection. As a first step towards this goal the trap column needed to be tested in order to evaluate its potential for the retention of the components of interest. Due to the varying polarity of the flavonoids and terpene lactones two types of trap column were tested in this study, C_{30} and C_{18} columns. C_{30} columns have been reported to show significantly greater shape selectivity and improved retention compared to C_{18} phases[16]. Whilst retention was achieved in our case the compounds were found to bind too tightly to the column and proved difficult to elute from the trap column in a sharp,

resolved peak as would be required for on-column focussing onto the analytical column, as per the benefit of column switching techniques (data not shown). The C_{18} trap column exhibited good retention with no TIC response before the column switch at 3 minutes as shown in Figure 5.3. The components of interest were eluted from the trap column at 3.4 minutes after the application of 90% acetonitrile mobile phase from the Ultimate pump and the full scan mass spectrum between retention time of 3.15 to 3.85 minutes is shown in Figure 5.4. All the ions of interest (10 standards and 1 internal standard) are present in the 3.15-3.85 minutes mass spectrum. Therefore both good retention and highly resolved elution are exhibited by this online trapping column and so it was chosen for the following experiments. The time point at which the valve switch occurs is critical in this experiment and so is a parameter that needed to be optimized before further experimentation. The time point depends on the injection volume, the dead volume prior to and including the trap itself and sample loading velocity. If the valve was switched too soon the sample was not desalted thoroughly and the HPLC/MS response suffered (data not shown). If switched too late sample loss of at least some of the components was detected. In our case, the valve was switched at 3 minutes, the dead volume (including the injection loop) was less than 6 μL, the loading pump flow rate 30 μL/min. The trap column was therefore flushed with a total volume of 90 μL 0.1% HAC which is 18 times the sample injection volume and the salt washed off thoroughly during this time. The recovery of each component needed investigation in order to determine whether any of the compounds themselves were being significantly lost to waste during the trapping period.

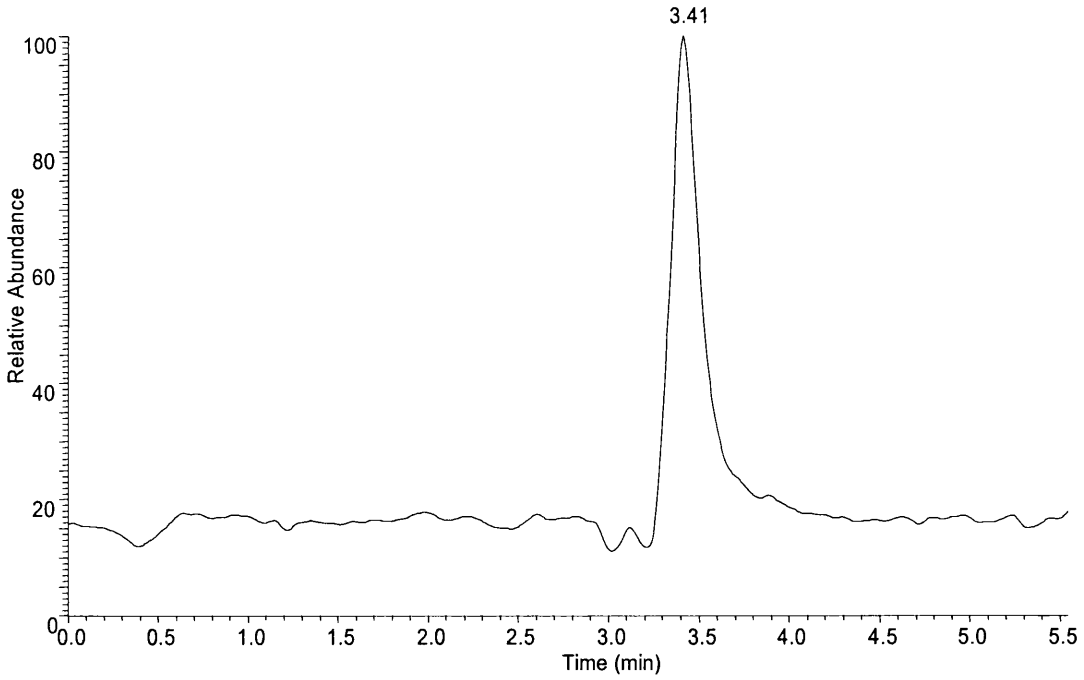


Figure 5.3 Test of C_{18} SPE trap column sample loading solution 0.1% HAC, Valve switch time 3 minutes, elute solution 90% ACN

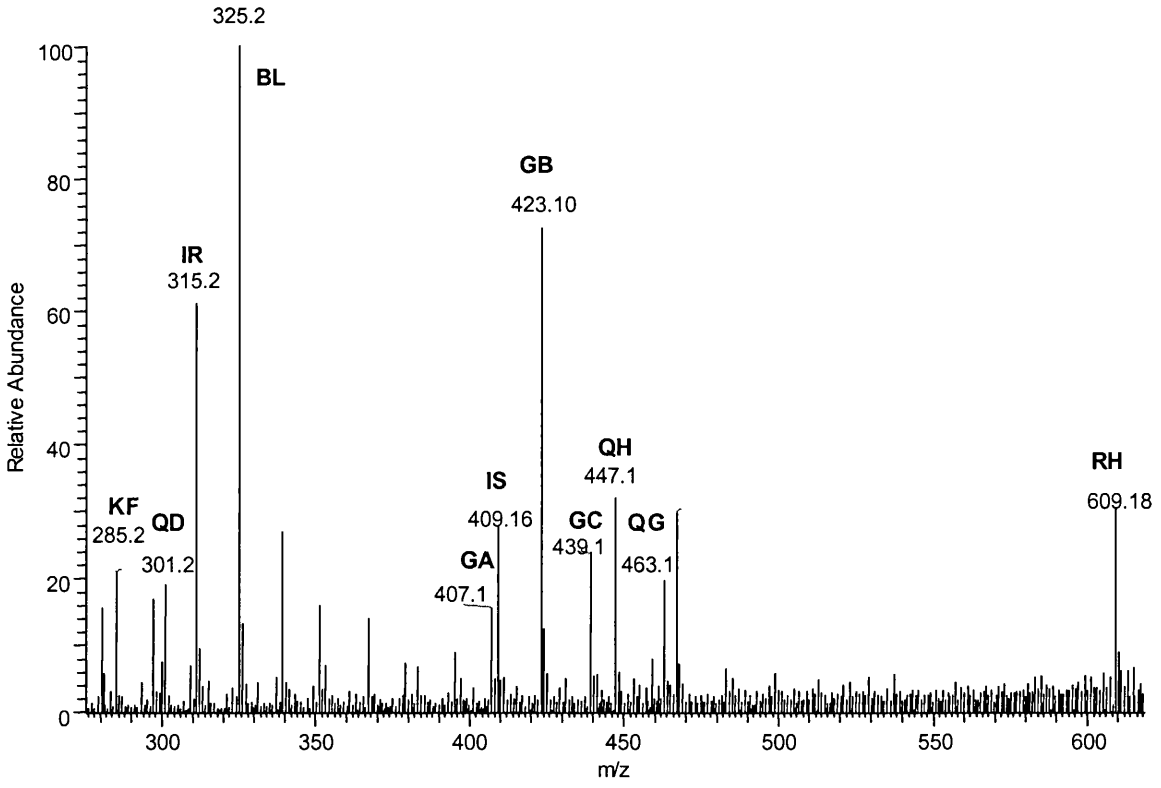


Figure 5.4 Full scan mass spectrum of the peak in Figure 5.3 retention time 3.15-3.85 minutes

5.4.2 Comparison of extraction recovery of online SPE, offline SPE, and LLE

The extraction recovery of the online purification method, conventional offline SPE and LLE methods was obtained by comparing the peak area of the recovered components SIM signal to that of the pure standards at the levels subjected to the purification protocols. The results are shown in Table 5.3 and show that the recovery of online purification is higher than 75% and exhibits higher recovery of all 10 standards except isorhamnetin compared to the commonly utilised LLE strategy of purification. Online trapping has similar recovery to offline SPE and therefore can be used in sample pre-treatment in order to avoid time consuming offline SPE and improve the automation of the whole analysis. The losses detected for the on-line trap HPLC/MS system were similar to those found in other purification protocols and so it is concluded that the switching time of the valve is adequate for the combined system.

Table 5.3 Comparison of recovery of solid phase extraction, liquid-liquid extract and online solid phase extraction of *Ginkgo* active components in blank urine matrix

Component	On-line SPE (%)	Off-line SPE (%)	LLE (%)
BL	75.6	79.9	74.5
GC	82.3	76.7	70.6
RH	98.5	98.3	83.3
QG	96.8	95.2	82.6
QH	92.4	93.6	76.8
GA	95.7	93.9	90.0
GB	97.6	92.6	92.3
QD	93.2	91.5	77.8
KF	90.4	78.8	88.6
IR	82.1	81.8	88.0

5.4.3 The method validation of quantitative analysis of active components of *Ginkgo biloba* in urine

In this study, the quantitative analysis of 10 active components of *Ginkgo biloba* in urine sample was undertaken using online-purification HPLC/MS with column switching technology. The total ion chromatogram (TIC) and extracted ion chromatogram (XIC) of the $[M-H]^-$ ions of blank urine spiked with mixture of standards of active components in *Ginkgo biloba* are shown in Figure 5.5 and 5.6.

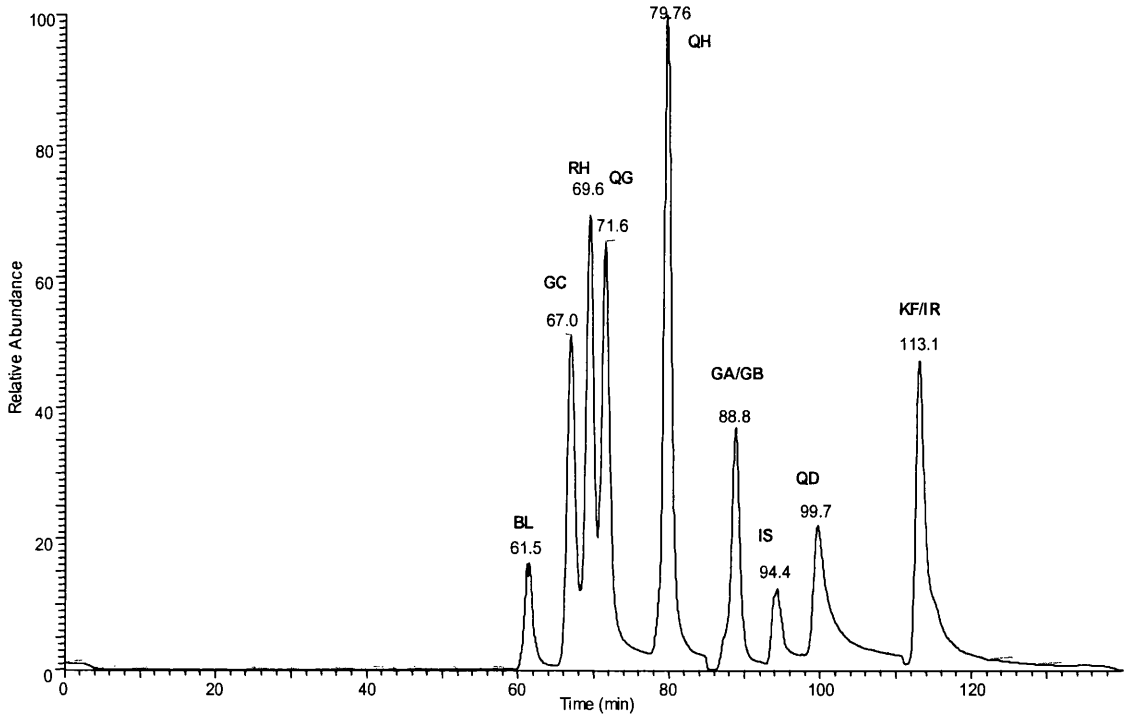


Figure 5.5 TIC of standard compounds spiked into urine prior to analysis

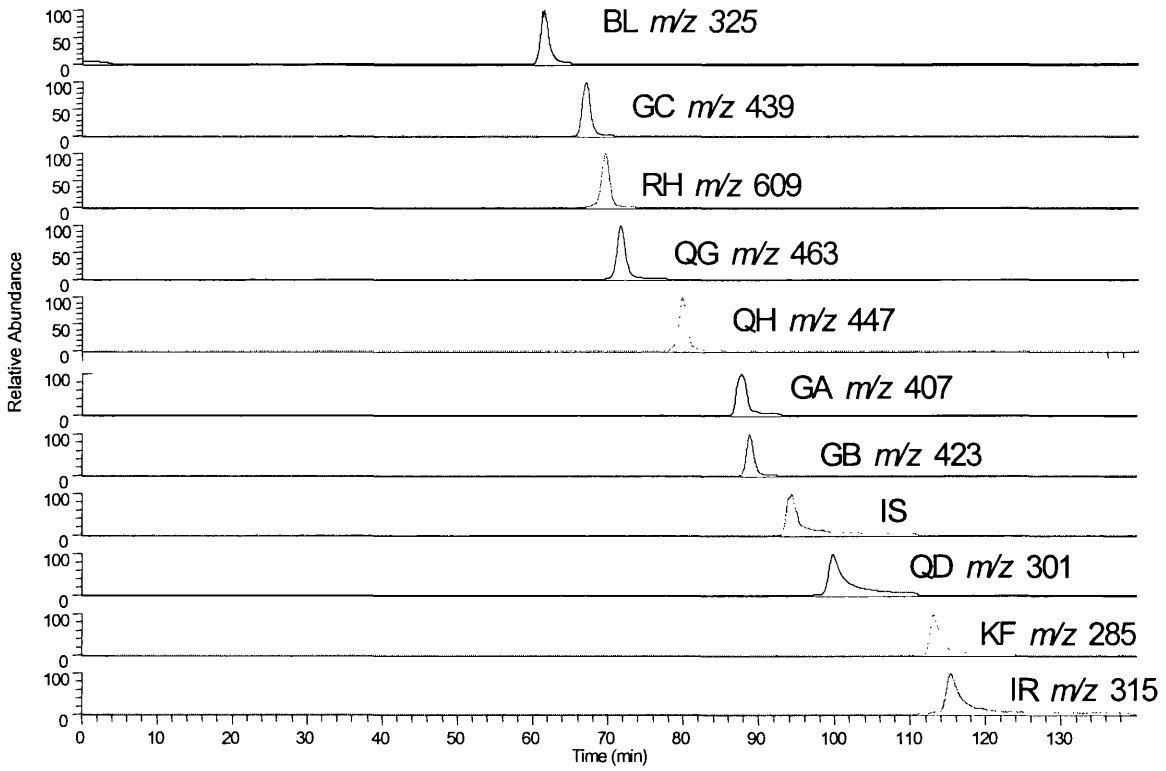


Figure 5.6 XIC of $[M-H]^-$ ions of standard compounds spiked into urine prior to analysis

It can be seen that the 10 components and internal standard obtain a reasonable separation by C_{18} capillary chromatography with a few compounds co-eluting from the column. Further experiments did not improve the separation of these compounds but their molecular ions are distinct enough to allow their individual study. Compared to a normal-bore column, the flow rate of the capillary column is as low as 4 μL/min, which can be connected directly to mass spectrometer giving an increase in the sensitivity of the combined system as it acts in a concentration dependent manner.

Calibration curves were obtained by analyzing a series of blank human urine samples (100 μL) spiked with reference standards of the concentration range 10-10000 ng/mL, with andrographolide used as internal standard. Calibration curves were constructed by

plotting peak areas ratio of the SIM signal of $[M-H]^-$ ions of reference standard and internal standard against concentration, the calibration curves are shown in Figures 5.7-5.13.

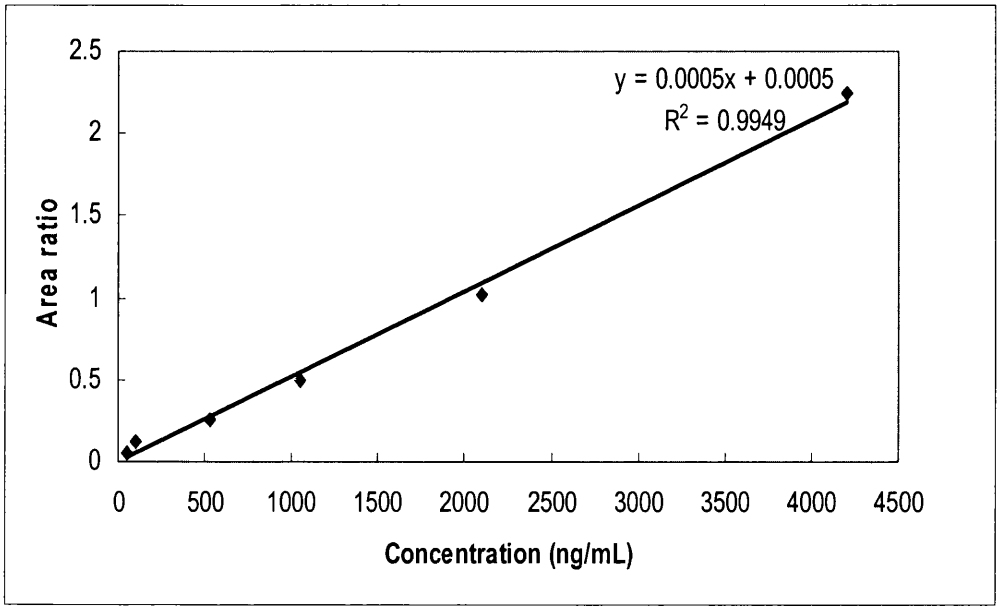


Figure 5.7 Calibration curve of bilobalide

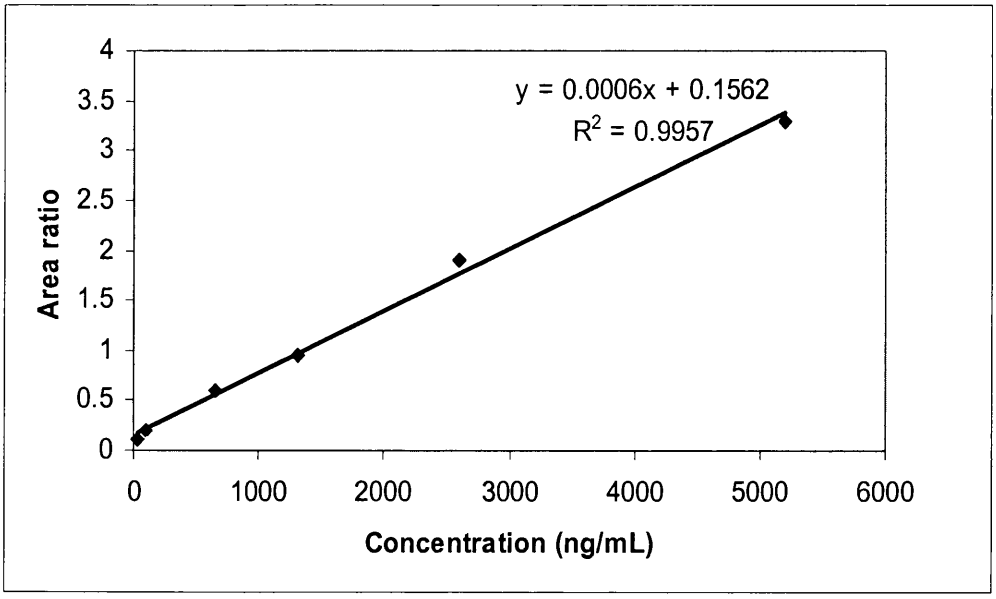


Figure 5.8 Calibration curve of ginkgolide C

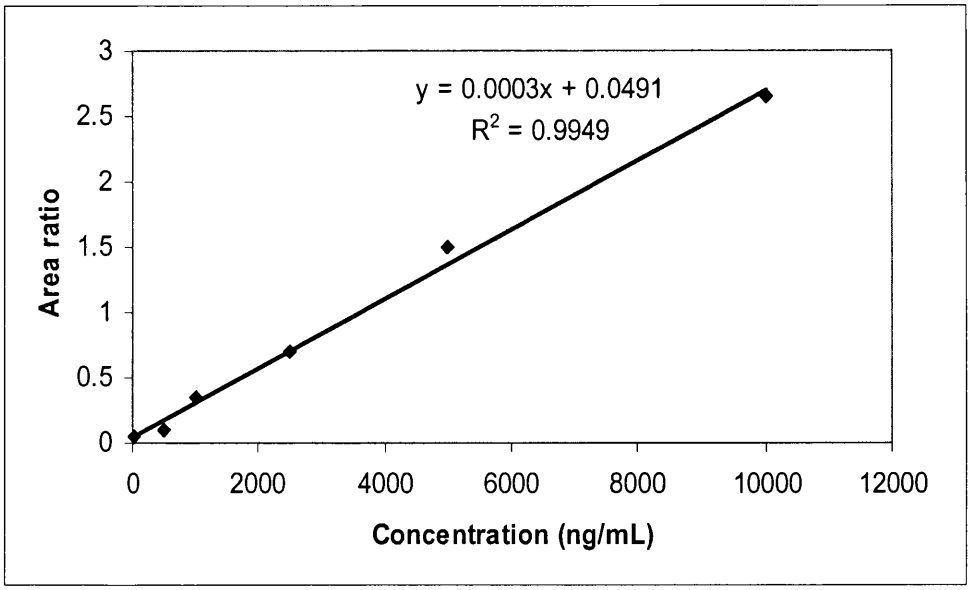


Figure 5.9 Calibration curve of ginkgolide A

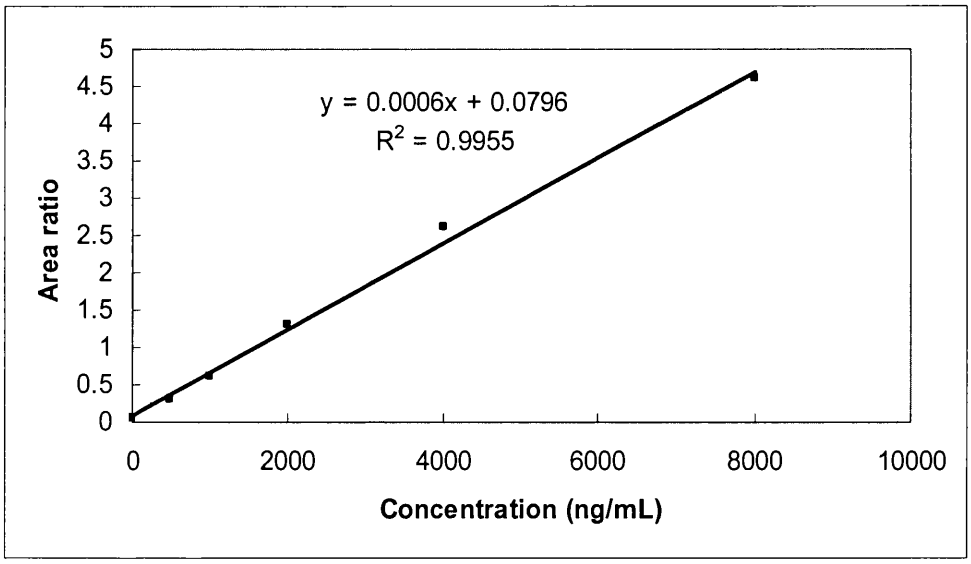


Figure 5.10 Calibration curve of ginkgolide B

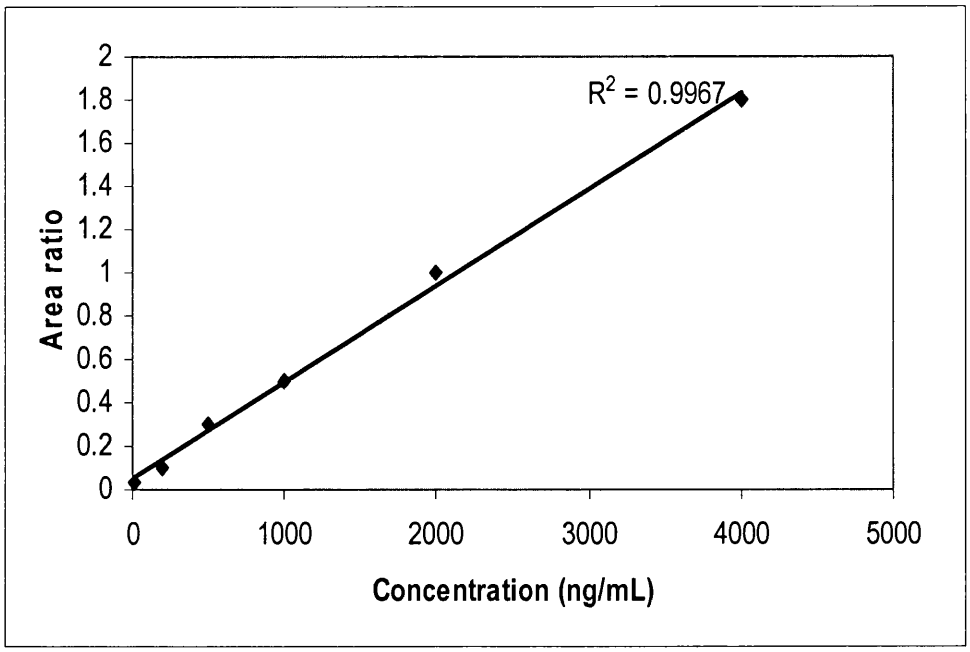


Figure 5.11 Calibration curve of quercetin

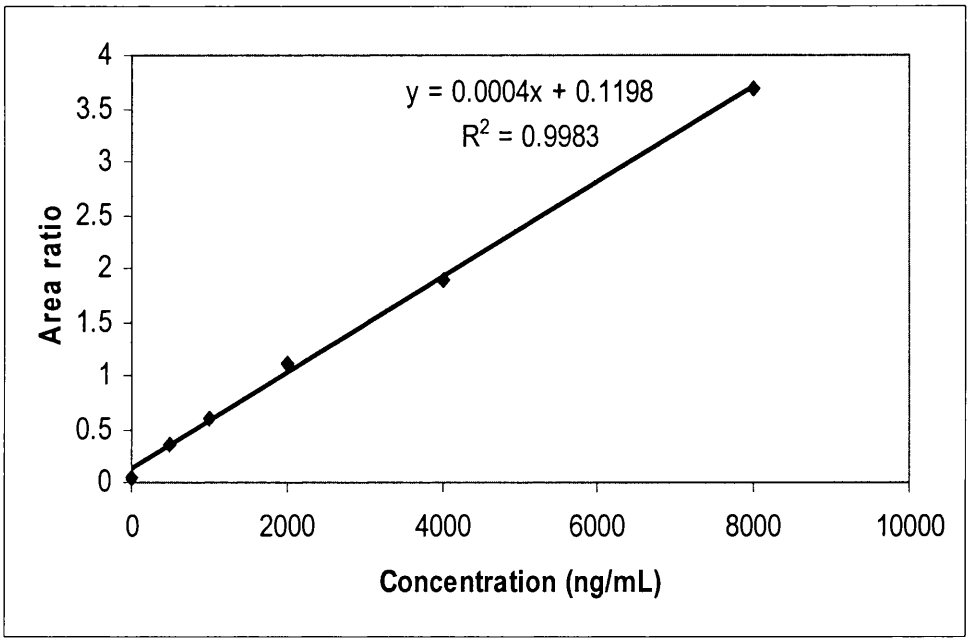


Figure 5.12 Calibration curve of kaempferol

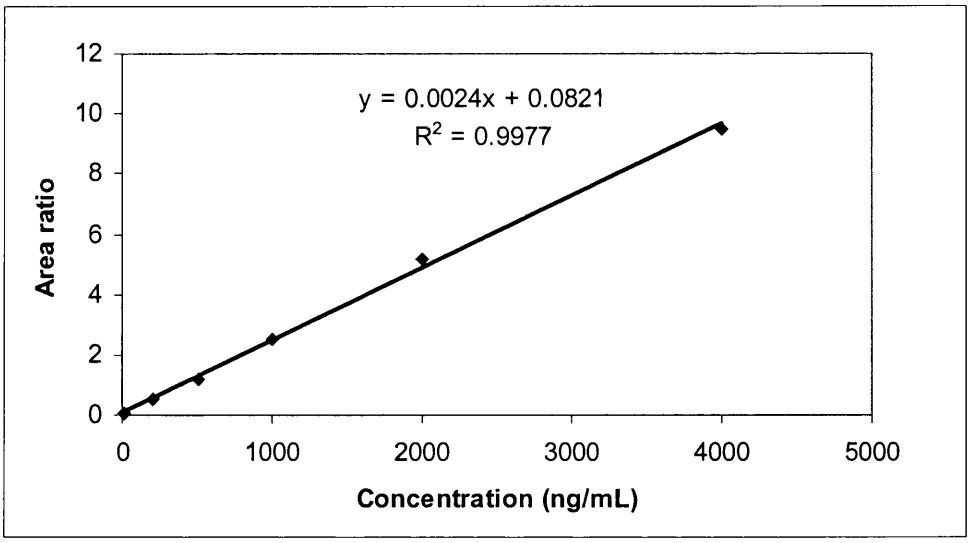


Figure 5.13 Calibration curve of isorhamnetin

Linear regression analysis was also performed for each reference standard and the linear range and correlation coefficients obtained for the regression analysis based on a minimum of six data points was shown in Table 5.4. The limit of detection (LOD) was studied in order to investigate the sensitivity of the assay developed. In this study, the LOD was estimated by spiking the blank urine with a serial dilution of the analytes until the diluted sample gave a signal-to-noise ratio (S/N) of 3. LOD data is shown in Table 5.4. The LOD of the reference standards is between 1-20 ng/mL, which is an adequate sensitivity for the trace analysis in biological sample. The reproducibility of the proposed method was evaluated by carrying out five replicate quantitative determinations for each of the studied compounds spiked in blank urine on the same day and five on three consecutive days. Intra-day and inter-day RSDs of the quantitation of the 10 components in *Ginkgo biloba* is shown in Table 5.4, it shows that good repeatability and reproducibility can be obtained by this method and so the system is applicable for quantitative determination. The %RSD values of intra- and inter-day

precision ranged from 8.5-13.6% and 11.6-17.2%, respectively, indicating little variation in the HPLC/MS performance.

Table 5.4 Limit of detection, linearity range, intra- and inter-day precision of *Ginkgo* active components in urine matrix

Component	LOD (ng/mL)	Linearity (ng/mL)	R^2	Intra-day (%CV)	Inter-day (%CV)
BL	18.3	50-4000	0.9942	11.9	14.7
GC	12.2	40-5000	0.9957	10.4	13.1
RH	1.9	10-2000	0.9985	11.4	17.2
QG	4.2	10-2000	0.9988	8.5	15.6
QH	1.1	10-1000	0.9990	10.2	17.5
GA	6.9	20-10000	0.9949	8.5	11.6
GB	3.6	10-8000	0.9955	9.6	12.2
QD	3.4	10-4000	0.9967	13.6	17.0
KF	3.1	10-8000	0.9983	8.4	14.4
IR	3.2	10-4000	0.9977	7.9	13.6

Table 5.5 Concentration (ng/mL) of *Ginkgo* terpene lactones and *Ginkgo* flavonoid aglyones in urine following ingestion of *Ginkgo biloba* tablets and enzyme hydrolysis

Component	BL	GC	GA	GB	QD	KF	IR
Sample 1	1795	127	2092	841	50.5	79.4	37.9
Sample 2	1659	100	1729	679	42.8	62.8	24.5
Sample 3	1219	67	1204	455	17.1	18.1	9.6
Sample 4	1510	89	1618	663	37.2	52.4	18.9
Sample 5	1597	103	1654	623	31.4	55.3	26.7
Average (ng/mL)	1556	97	1659	652	35.8	53.6	23.52
STD	215	22	317	138	12.6	22.4	10.4

5.4.4 Assay of active components in human urine

The developed method was applied to the analysis of the urine samples of 5 volunteers and the results are shown in Table 5.5. From urine analysis, *Ginkgo* terpene lactones are not affected by the inclusion of the β-glucuronidase/sulfatase (Figure 5.14). BL, GA, GB were excreted in their original form and exhibit high concentrations in the urine sample, GC has a relatively low concentration, which is consistent with report from Mauri *et al*[17] that GC is metabolised to its methylated form. Figure 5.14 shows terpene lactones and *Ginkgo* flavonoids before and after enzyme hydrolysis. *Ginkgo* flavonoid aglycones were lower than limit of quantitation before enzyme treatment, but can be readily identified and quantitated after enzyme hydrolysis, however they were detected at far lower concentrations than the ingested amount which is consistent with a previous report[9]. This suggests that after absorption through the intestinal wall the flavonoid aglycones are excreted in the urine not only as glucuronides and sulphate conjugates,

but also in hydro-, dihydro- or methylate form which may account for reduced excretion values. The content of terpene lactones and flavonoids differ greatly due in some part to differing metabolism and also to differing urinary volumes however since the established method has good resolution, sensitivity and wide linear range, the active components can be quantitated simultaneously from urinary samples allowing further development of high throughput techniques to monitor *Ginkgo* bioavailability.

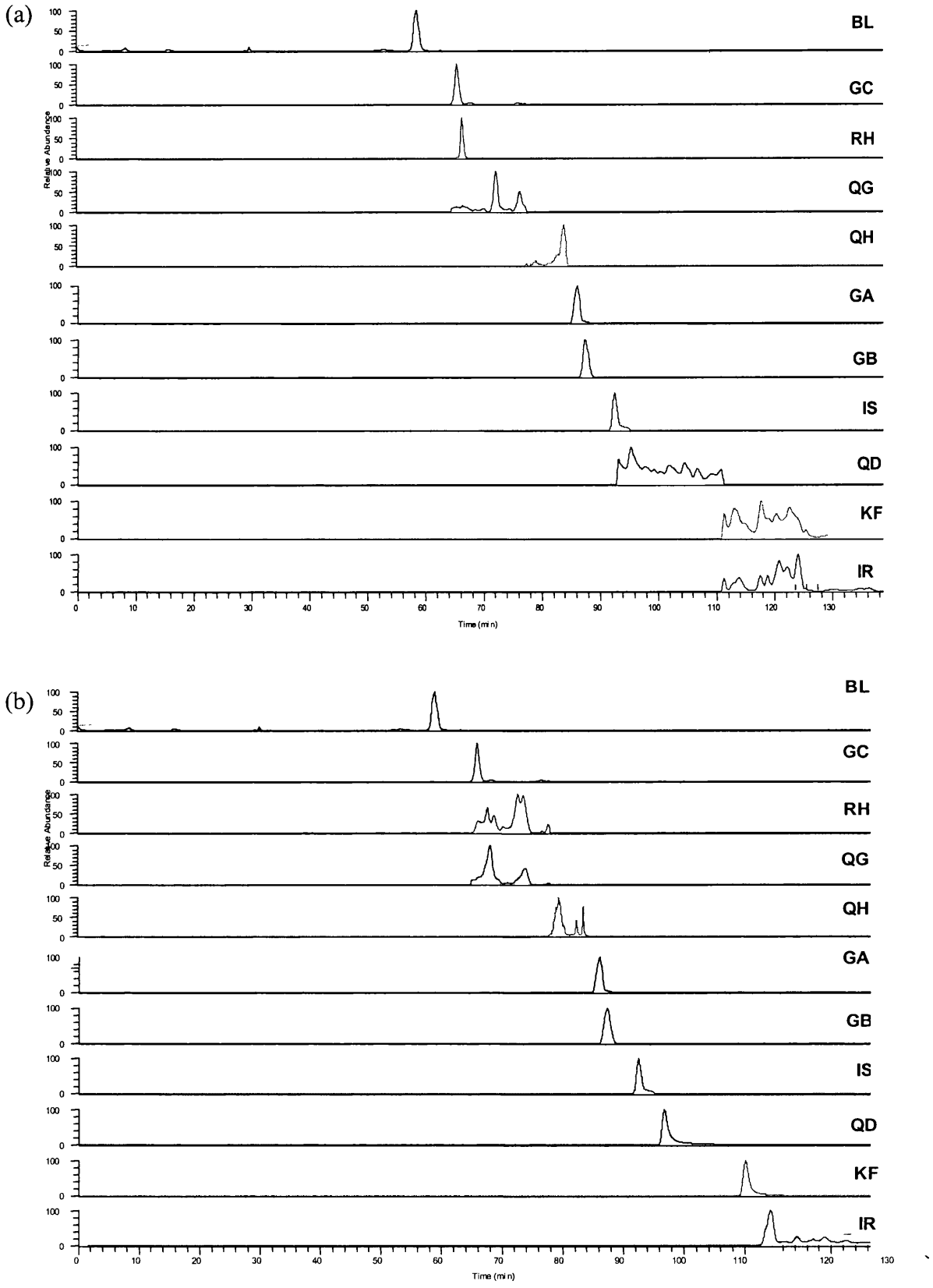


Figure 5.14 Analysis of urine sample 4 hours after ingestion of the *Ginkgo biloba* tablet before (a) and after (b) enzyme treatment of the sample

5.5 Conclusion

An on-line column switching HPLC/MS was successfully developed for the analysis of urinary excreted flavonoids and terpene lactones derived from *Ginkgo* extract in a single analysis in this study. The reverse phase trap column was shown to allow the quick and efficient clean up of the injected urine sample which required minimal prior sample preparation thereby facilitating higher throughput and greater automation. The experiments were performed as a "proof of principle" study and were validated in order to demonstrate adequate recovery of the analytes, linear range of the analysis and acceptable limits of detection. The comparison of the developed method with off-line purification methodologies show that the on-line purification does not suffer from any significant loss of analyte or interference from other matrix materials. The assay was applied to study the levels of these compounds in the urine of 5 subjects 4 hours after ingesting a *Ginkgo biloba* extract tablet. Although no attempt was made in this study to quantitate the full amount of each analyte excreted by measuring total urinary volume and or utilising a normalising factor such as urinary creatinine in this present study however the results nevertheless show that the excreted levels could be analysed from as little as 100 μL of urine.

5.6 References

1. van Beek, T.A., J. Chromatogr. A, **967**, 21 (2002)
2. McKenna, D.J., Jones, K. and Hughes, K., Altern. Ther. Health. Med., **7**, 70 (2001)
3. Rice-Evans, C.A., Miller, N.J. and Paganga, G., Trends Plant Sci., **2**, 152 (1997)
4. DeFeudis, F.V. and Drieu, K., Current Drug Targets, **1**, 25 (2000)
5. Prasain, J.K. Wang, C.C. and Barnes, S., Free Radical Biol. Med., **37**, 1324 (2004)
6. Nielsen, S.E. and Dragsted, L.O., J. Chromatogr. B, **707**, 81 (1998)
7. Erlund, I., Alfthan, G., Siren, H., Ariniemi, K. and Aro, A., J. Chromatogr. B, **727**, 179 (1999)
8. Wang, F.M., Yao, T.W. and Zeng, S., J. Pharm. Biomed. Anal., **33**, 317 (2003)
9. Watson, D.G. and Oliveira, E.J., J. Chromatogr. B, **723**, 203 (1999)
10. Watson, D.G. and Pitt, A.R., Rapid Commun. Mass Spectrom., **12**, 153 (1998)
11. Oliveira, E.J. and Watson, D.G., FEBS letters, **471**, 1 (2000)
12. Hua, L., Guangji, W., Hao, L., Minwen, H., Haitang, X., Chenrong, H., Jianguo, S. and Tian L., J. Pharm. Biomed. Anal., **40**, 88 (2006)
13. Mauri, P., Minoggio, M., Iemoli, L., Rossoni, G., Morazzoni, P., Bombardelli, E. and Pietta, P., J. Pharm. Biomed. Anal., **32**, 633 (2003)
14. Biber A and Koch E., Planta Med., **65**, 192 (1999)
15. Ding, S., Dudley, E., Plummer, S., Tang, J., Newton, R.P. and Brenton, A.G., Rapid Commun. Mass Spectrom., **20**, 2753 (2006)
16. Srinivasan, G., Meyer, C., Welsch, N., Albert, K. and Müller, K., J. Chromatogr. A, **1113**, 45 (2006)

17. Mauri, P., De Palma, A., Pozzi, F., Basilico, F., Riva, A., Morazzoni, P., Bombardelli, E. and Rossoni, G., J. Pharm. Biomed. Anal., **40**, 763 (2006)

Chapter 6

An investigation of mass accuracy at high molecular weight by matrix-assisted laser desorption/ionisation

6.1 A brief overview of mass measurement

The term accurate mass generally refers to the mass that allows the elemental formula to be deduced unequivocally. For low molecular weight ions an accurate mass tends to be close in value to the nominal mass. To obtain the nominal mass of a compound, the mass spectrometer should have a mass accuracy of at least ±0.4 u. As molecular size increases so does the fractional component of mass. This arises especially with large numbers of H (mass = 1.0078 u) and gives accurate masses with large fractional components. The mass of a molecule with about 128 H atoms will add one mass unit to the nominal mass, that means the mass scale rolls over about every 1000 u for hydrogen rich molecules[1]. It should be noted that the IUPAC unit for mass is defined 12 u (not amu) is equal to the mass of ^{12}C. However, the term Da (Dalton) is frequently and interchangeably used and is equivalent and equal to u.

The improvement of the mass accuracy of instruments has allowed scientists to determine the mass of the molecular ion accurately and from this information a molecular formula can be assigned although not necessarily uniquely. The traditional and continuing justification for high resolution mass spectrometry is for the identification and confirmation of the molecular formulae of new compounds. An early example of the utility of accurate mass is for the detailed characterization of higher boiling fractions of petroleum in terms of functional group composition[2].

Mass spectrometric exact mass measurements are widely used to determine or confirm the elemental composition of low molecular weight organic compounds (<850 u). Most of the elements have isotopes; these are responsible for the peaks in the mass spectrum appearing as isotopic clusters. They are characteristic of the elemental composition and provide important analytical data[3]. The relative abundances of isotopes in a molecular ion (or in a fragment ion) result from their statistical distribution. Mathematical methods for the calculation of theoretical relative abundances within the isotopic cluster, for comparison with experiment, usually rely on expansion of the polynomial expression based on an extension of the binominal probability distribution[4]. Because the number of isotope combinations becomes very large at higher mass, requiring huge numbers of computer calculations, approximate mathematical methods have been sought to improve calculation times[5, 6]. Rockwood's method uses Fourier transforms to conduct the multiple convolutions required to determine molecular isotope distributions. Discrete Fourier transforms are very efficient, and this method for calculating isotope distributions is fast, accurate and economical in its use of computer memory so it has significant practical implication.

Theoretically, the isotopic abundances for a given formula can be calculated and compared with experimental values, the consequence of which is that the elemental composition of any ion, or fragment, can be calculated from its measured mass if this is sufficiently accurately determined. The technique consists of measuring the mass of the monoisotopic peak characteristic of the intact molecule, with accuracy and precision

sufficient to distinguish it from the masses of all or most other plausible elemental compositons[7]. It is true, for low molecular weight compounds, that the monoisotopic peak is the base peak of the molecular ion cluster (refer to Figure 6.1). In general, accurate mass measurement is non-ambiguous up to a couple of hundred mass units, for empirical formulae containing a relatively simple set of elements C, H, N and O, due to the small number of potential molecular formulae[8]. According to Marshall *et al*[9] the highest mass for unique elemental composition determination is 895u for current instrumentation. For organic compounds, the carbon isotope is always a major consideration. When the carbon numbers start to get very high, the isotope distribution becomes dominated by the carbon content. For an increasing number of carbon atoms, i.e., higher mass compounds, the base peak is no longer the monoisotopic peak. In the case of, for example, high molecular proteins the isotopic distribution tends to be a Gaussian shape where the monoisotopic peak is hardly detectable (see Figure 6.1). Mass measurement accuracy is used to evaluate the performance of the instrument where the mass of an unknown species is measured against a known entity. Usually mass measurement accuracy is expressed in parts per million (ppm), the measurement indicates the deviation of the instrument mass measurement from the calculated monoisotopic mass. Different molecular weight ranges exhibit different mass accuracy, the highest mass accuracy of ±0.5 ppm can be achieved by FTICR over mass range from 90-300 Da and ±1 ppm from 250-1000 Da[9].

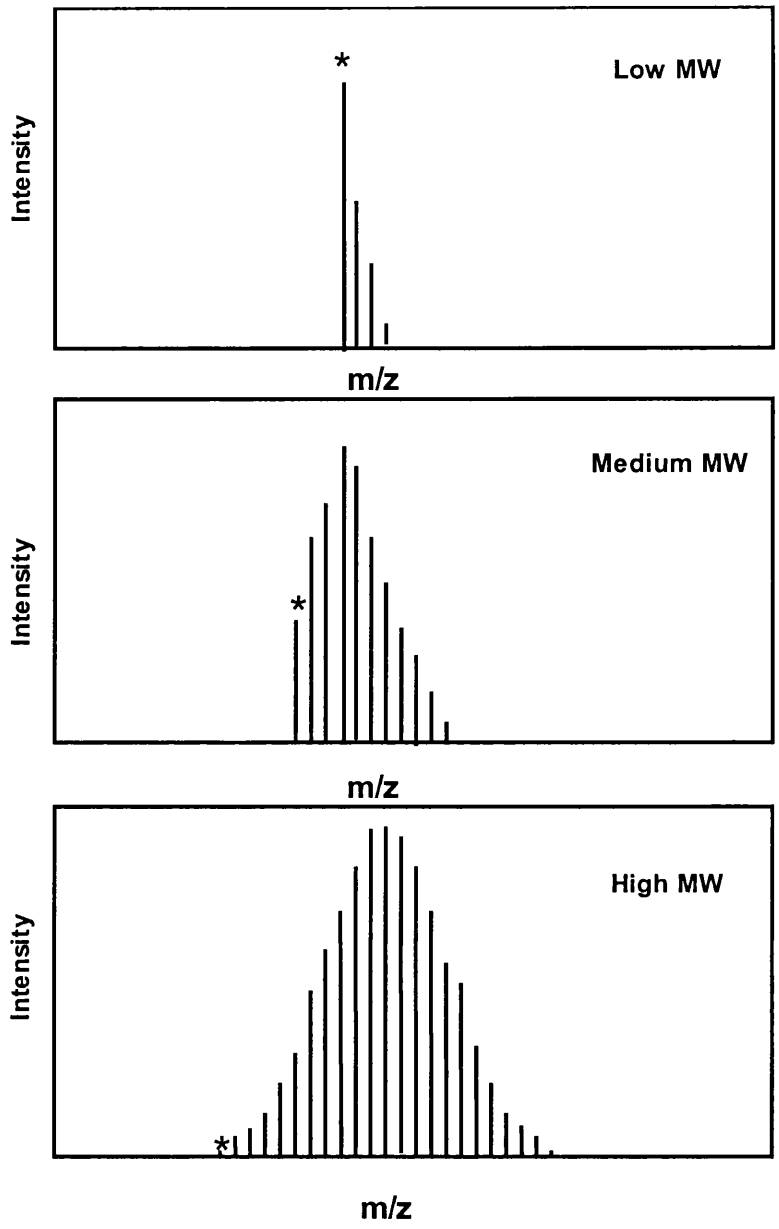


Figure 6.1 Isotopic distribution of low, medium and high molecular weight compounds, showing the reduction in the intensity of the monoisotopic ion with increasing molecular weight. The monoisotopic ion is indicated by *.

Currently, the main area which drives developments in mass spectrometry is biological and medical sciences, in particular the analysis of mixture of peptides, proteins and oligonucleotides[10]. The technical requirements for the mass spectrometer to meet the demands of biological science are high sensitivity, high scan speed, good mass accuracy, automation and the ability to obtain structural information[11]. In proteomics research, the normal procedures to identify a protein are first the digestion of proteins with enzymes, for example, trypsin, then mass spectrometry is employed to obtain information on the masses or amino acid sequence of the peptides. The experimental MS results are inputted to a protein database and compared with the masses or sequence of peptides which are digested theoretically from proteins by the database software[12]. It is probable to find many matched target proteins from the database. The more accurate the masses of the peptides, the more reliable the target protein identification[13]. For example, the amino acids Lysine and Glutamine have monoisotopic residual masses of 128.174 and 128.131 u, respectively. This shows that mass measurement accuracy could be very important because any inaccuracy will cause false matches, or no match. So the mass spectrometrist must dedicate themselves to the development of mass analysers for high mass accuracy, and now there are commercial mass spectrometers which have mass accuracy as high as three decimal places, such as, FTICR[9] and the very popular MS/MS instruments, such as, QTOF[14] and MALDI-TOF-TOF[15].

When MALDI was discovered and implemented on linear TOF mass spectrometers, the accuracy of mass measurement was extremely poor, i.e. ± 2 Da for peptides using

external calibration. This poor accuracy and poor mass resolution (<500 (M/ΔM), FWHM) were largely caused by formation of ions with both a broad initial kinetic energy distribution and mass-independent initial velocities[16]. Refocusing of ions using an electrostatic mirror (or reflectron) boosted the achievable resolution to >2000[17]. Such mass spectrometers are commonly referred to as reflectron-TOF MS. The implementation of delayed extraction (DE) also alleviated the ion energy spread problem[18, 19]. DE accompanied by the modern high-voltage electronic switches and fast digitization electronics enable MALDI-TOF instruments with reflectors and longer flight tube lengths to achieve mass resolution >10,000 with accompanying mass accuracy of about 10 ppm or better[20].

Reflectron-TOF MS is not suitable for high molecular weight proteins. According to the principles of time-of-flight, the time-of-flight of a protein is proportional to the square root of its molecular mass. The larger the protein, the longer the flight time. Studies by Chait[21] and Demirev *et al.*[22] have shown that the metastable fragmentation time frame for small proteins (5,000 to 20,000 Da) is of the order of tens of microseconds. Ions spend half of their time in the drift region for a reflectron TOF system and most of their time for a linear TOF system, so that fragmentation of these ions occurs predominantly in the flight tube. A discussion of the affects of fragmentation of ions in the acceleration region and field free drift tubes has been made by Cotter[23], for fragmentation during acceleration Cotter predicts "ions will contribute to the general baseline noise" between the precursor ion and the product ion. Whilst in the flight tube he predicts

"fragmentation in this region is largely inconsequential" as the "product ions and neutrals will all have the same time-of-flight as the precursor ion". So MALDI-TOF of proteins is often performed in linear mode where sensitivity will be considerably higher as the molecular species will be detected, at the correct mass, even if it fragments during flight as the fragment ion has the same mean velocity as the precursor ion[24]. However, the discussion by Cotter does not explain the high mass tailing we observed at high mass and prompts a more thorough approach.

6.2 Calibration of the mass scale of a mass spectrometer

6.2.1 Introduction

To obtain an accurate mass spectrum, it is essential to calibrate the mass scale of the spectrometer and establish a mass scale over a defined mass range. If the calibration is poor, even a high precision mass measurement will be poor in terms of accuracy. Physically a mass scale does not exist, it has to be derived from another physical quantity usually time. Now time can be measured to very high accuracy and precision because time may be measured by a frequency counter. To achieve nominal mass, the calibration accuracy needs to be less than 0.4 u, higher mass accuracy is required to obtain the elemental formula. Mass accuracy of approximately 35 ppm at 100 u and 0.018 ppm at 750 u are needed to achieve an unique formula[1]. So the mass accuracy required varies over the mass scale and has a practical limit defined by technology.

Calibration of the *m/z* scale of the mass spectrometer is achieved using a reference compound yielding ions of known *m/z*. Appropriate instrument calibration is vital for good mass measurement accuracy. Two calibration protocols, described below, are used with accurate mass measurement.

6.2.2 External calibration

In external calibration mode, calibration is carried out prior to analysis. The spectrum of a known reference compound is recorded and the peaks in the spectrum are assigned to their exact masses by the instrument data system to produce the correct calibration. To achieve the best results using external calibration, the calibration needs to be conducted as soon as possible to the unknown accurate mass measurement to minimize instrument drift due to any instability of the electronics and temperature variation.

6.2.3 Internal calibration

In internal calibration mode, the *m/z* scale is calibrated using a single ion or a series of ions of known *m/z* from a reference compound that is introduced into the mass spectrometer ion source at the same time as the analyte. The method is superior to external calibration since the effects of instrumental drift are counteracted. To achieve optimum mass accuracy the number of calibration points need to be sufficient and equally spaced over the intended mass range. Ideally the calibration ions should be as close in mass to the analyte ions under investigation, as this reduces mass measurement

errors. A single mass continuously to lock the mass calibration scale is commonly referred to as a "lock mass". This method is used when the calibration compound is not continuously present.

The type of calibration compound depends on mass range, whether a nominal or accurate mass is required, resolution of the instrument and ionisation mode. Some common reference compounds are shown in Table 6.1.

Table 6.1 Reference compounds for mass marking

Ionisation mode	**Reference compounds for mass marking**	**Mass range (probable ions)**
EI	PFK (Perfluorokerosine)	69 - ~1000
	Heptacosafluoro-tri-N-butylamine	up to 614
CI	Poly (dimethysiloxane)	90-1274
	FAME (Fatty acid methyl esters mixture)	≤1000[1]
ESI	PEG (polyethylene glycol)	80 - > 900
	PEI (Poly ethylenimine)	61-1997
	Alkyl ethoxy sulfate	265-1027
	RbI/ NaI	100-2000
MALDI	*Mixture of peptides (Mixture 1)*	
	Des-Arg1-Bradykinin	905.05
	Angiotensin I	1297.51
	Glu1-Fibrinopeptide B	1571.61
	Neurotensin	1673.96
	Mixture of proteins (Mixture 3)	
	Insulin bovine	5734.59
	Insulin bovine (2+)	2867.80
	Thioredoxin from E. coli	11674.48
	Thioredoxin from E. coli (2+)	5837.74
	Apomyoglobin from horse	16952.56
	Apomyoglobin from horse (2+)	8476.78

[1] Dependent on FAME used.

Methods of calibrating are instrument dependent, for time-of-flight, quadrupole and sector calibration time is used; for the ion trap, Orbitrap and FTICR calibration frequency is used. A comparison of precision for different mass analyser is listed in Table 6.2[25].

Table 6.2 Typical mass accuracy of different mass analyser[25]

Mass analyser type	***m/z* range**	**Mass measurement accuracy**
TOF	10^6	> 2 ppm[1]
FTICR	10^5	~0.1 ppm
Sector full scan	< 10^4	3 ppm
Sector peak matching	< 10^4	0.3 ppm
Sector narrow scanning	< 10^4	0.3 ppm
Quadrupole	< 10^4	> 0.1 Da
Special Quadrupole	10^4	> 10 ppm
Ion trap	10^3-10^4	0.1 Da
Ion trap very slow scan	10^3-10^4	50-100 ppm

[1] Reflectron TOF typically for ions < 5000 Da

From this table, it is clear that FTICR currently offers the highest mass resolution of any mass spectrometer and is routinely used for accurate mass measurement requiring mass resolution in excess of $1x10^6$. For MALDI the easiest mass analyser to employ is TOF because MALDI is a pulsed technology like TOF. In addition, TOF has very high sensitivity as all ions are detected. Mass accuracy of ~5 to 10 ppm and mass resolving powers around 20,000 to 30,000 (FWHM) can be achieved with reflectron time-of-flight[26]. When high-mass compounds have to be run they are often not amenable to reflectron TOF-MS as the sensitivity is low due to ion fragmentation (e.g.

$M^+ \rightarrow F^+ + N$) within the flight path. These fragment ions generally do not reach the detector because their kinetic energy and path changes suddenly upon fragmentation and neutrals (N) formed on fragmentation pass straight through the reflectron and are lost[23]. This phenomenon is observed to increase with increasing mass. Characterization of high mass proteins is thus usually conducted in linear-TOF mode, however accurate mass measurement is less well defined in linear TOF mode since the focusing action of the reflectron is absent, mass measurements are typically considered to be 100 ppm.

When using MALDI-TOF to obtain the mass spectra of proteins in linear mode, smoothing of the raw data affects the result. Part of this study was aimed to examine a number of strategies to improve mass measurement at high mass, one of which being how best to treat raw data.

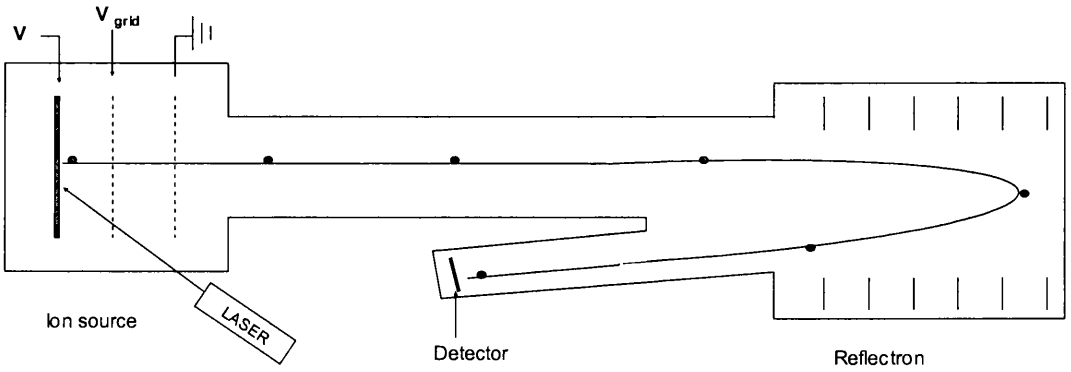


Figure 6.2 Schematic diagram of a MALDI-TOF with a reflectron

V=source voltage

V_{grid}=variable grid voltage used for focusing

In MALDI-TOF, ions are formed at the surface of the MALDI plate when the laser is fired (as seen in Figure 6.2), these are accelerated through the entire source-extraction region to the same final kinetic energy (zV)[23]

$$\frac{mv^2}{2} = zV \qquad \text{Eq. (6.1)}$$

where m=mass of the ion, v=velocity, z=ne, n=number of charges, e=electronic charge ($1.6\text{x}10^{-19}$ C), and V=source voltage.

The ions cross the drift region towards the detector with velocities

$$v = \left(\frac{2zV}{m}\right)^{1/2} \qquad \text{Eq. (6.2)}$$

and flight times

$$t = d\left(\frac{m}{2zV}\right)^{1/2} \qquad \text{Eq. (6.3)}$$

which depend upon the square root of their masses, where d=length of flight tube.

If d and V are considered constant, the above equation can be written as

$$t = c\left(\frac{m}{z}\right)^{1/2} \qquad \text{Eq. (6.4)}$$

where c is a constant.

However, for practical purposes a time offset t_0 is introduced to reflect uncertainty of the time origin, as follows

$$t = t_0 + c_0\left(\frac{m}{z}\right)^{1/2} \qquad \text{Eq. (6.5)}$$

In general, the time-of-flight for a real system, which has small deviations of linearity, the time-of-flight can be modeled using a polynomial equation of the general form

$$t = t_0 + c_0\left(\frac{m}{z}\right)^{1/2} + \sum_i c_i\left(\frac{m}{z}\right)^{n_i} \qquad \text{Eq. (6.6)}$$

When i=1, the following equation is obtained

$$t = t_0 + c_0\left(\frac{m}{z}\right)^{1/2} + c_1\left(\frac{m}{z}\right) \qquad \text{Eq. (6.7)}$$

In this equation, t_0 accounts for any internal delay in the acquisition system, $c_0(m/z)^{1/2}$ is the flight time of an ion with initial velocity, v, from the target surface to the detector, and $c_1(m/z)$ is a small flight time correction for the ion velocities at the onset of the extraction pulse. This equation is used for calibration of mass spectra with external or internal standards. The flight times of ions in TOF mass spectrometers are converted to ion masses using a calibration such as Eq. 6.7[27].

Moskovets *et al.*[28] reported a mass calibration equation with 4 terms which minimizes

the effect of systematic errors resulting from the fact that ion velocities are mass dependent due to the rise time of the extraction pulse. By this equation (Eq. 6.8), a mass accuracy of 1.6 ppm was obtained over the mass range of 1.0-4.0 kDa.

$$t = t_0 + c_{-1}\left(\frac{m}{z}\right)^{-1/2} + c_0\left(\frac{m}{z}\right)^{1/2} + c_1\left(\frac{m}{z}\right) \qquad \text{Eq. (6.8)}$$

MALDI-TOF mass spectrometers are equipped with fast digitizers and very stable power supplies, which allow time to be measured to within a precision of 0.05 ns (at best), corresponding to 1 ppm uncertainty in mass for ions of 100 μs flight times[29]. In practice TOF instruments use one of the above analytical methods and establish a difference table between the theoretical masses and observed masses and give a calibration graph with shows the mass "residuals" as a function of the mass scale. An example is shown in Figure 6.3.

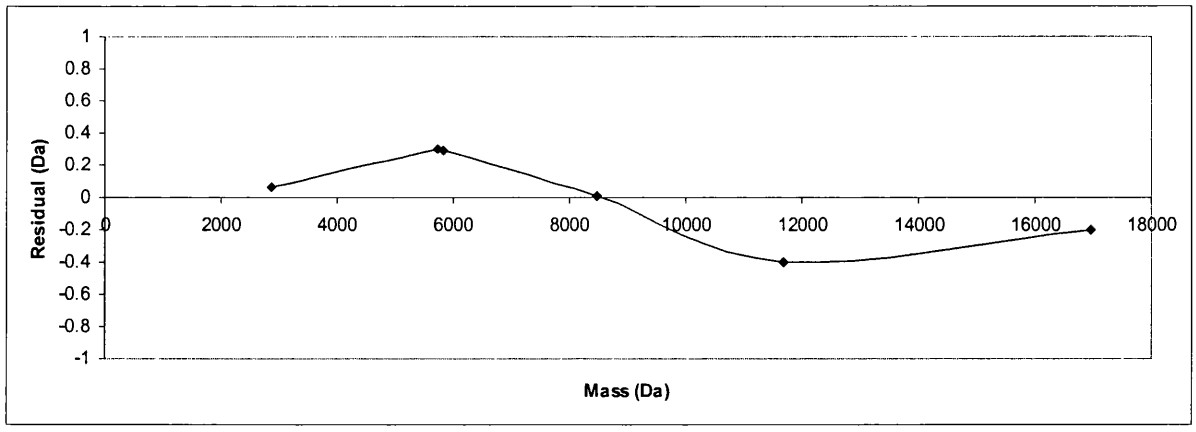


Figure 6.3 Calibration graph of the Voyager DE STR TOF mass analyser for protein mixture (Mixture **3**, listed in Table 6.1) as the calibration standard used in this study

6.3 Experimental methods and materials

6.3.1 Chemical and materials

Bovine serum albumin standard (Sequazyme BSA Test Standard Kit) was purchased from Applied Biosystems (Framingham, MA, USA). Sinapinic acid, acetonitrile, trifluoroacetic acid were obtained from Sigma Chemical Co. (St. Louis, MO, USA) and used without further purification. Water was purified with a Milli-Q deionisation unit (Millipore, Bedford, MA, USA).

6.3.2 Experiment method

Data acquisition and processing were performed on a Voyager DE STR mass spectrometer (Applied Biosystems, Warrington, UK) equipped with a delayed extraction ion source. The path length is approximately 2 metre in linear mode and 3 metre in reflectron mode[30]. A pulsed nitrogen laser is used and operates at 337 nm. For the experiments described in this chapter, the instrument was operated in linear mode. The accelerating voltage was 25000 V and the grid voltage (Grid 1 shown in Figure 6.2) was set to 94% of the accelerating voltage. The delay time before extraction is 800 nsec allowing ions of the same mass but different initial kinetic energies (energy or velocity focusing) to focus at the detector at the same time, thus achieving time focusing and high resolution[28].

The matrix used in this experiment was a saturated solution of 10 mg/mL sinapinic acid in 70/30 0.1%TFA/ACN. 1 μL of BSA standard was mixed with 25 μL of matrix solution, and 1 μL of the mixed solution was spotted on the MALDI plate. Normally 20 to 30 sample spots were placed on the MALDI plate at one time; the samples were sufficiently stable to accommodate this method of preparation. After the spot dried on the target plate, it was introduced into the Voyager DE STR MALDI-TOF. Fine control over the laser desorption process could be achieved using a joystick control to focus the laser at various areas of the sample. The laser power was set to 2500 (a software parameter usually set between 1500 and 3000 units) and remained unchanged to avoid mass measurement error related to power change. Signals from 100 shots were accumulated and averaged to generate one mass spectrum. The calibration of the mass spectrometer was carried out on one BSA spot, and then the laser was fired again at the same spot to obtain a mass spectrum of BSA. After one spectrum was acquired, it was calibrated and the whole procedure repeated to ensure the best possible mass measurements were obtained. Two methods for smoothing spectra were employed, "noise filter" and "Gaussian smooth", resulting in two sets of mass spectra each corresponding to a different signal filter.

6.4 Results and discussions

6.4.1 Mass accuracy and molecular weight

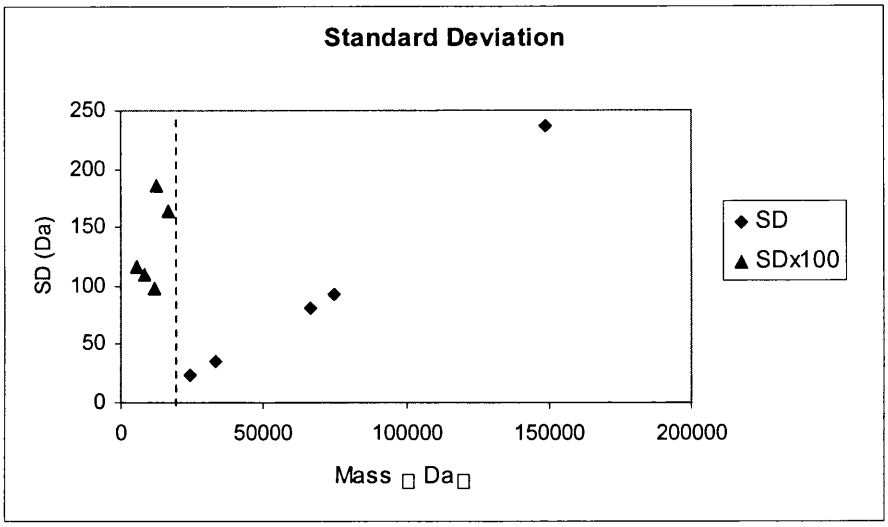
In the manual of the Voyager DE STR mass spectrometer, it is claimed that by external

calibration, 0.05% mass accuracy can be obtained in linear mode and 0.01% or better mass accuracy can be obtained in reflectron mode. By internal calibration 0.02% and 0.002% (20 ppm) mass accuracy can be obtained in linear and reflectron mode, respectively[27]. A variety of compounds (see Table 6.3) were used for mass measurement over the range 2 to 150 kDa, the standard deviation (SD) and relative standard deviation (RSD) is shown in Figure 6.4. It shows that mass error increases with molecular weight, especially at masses higher than 20,000 Da when the error increases linearly with increasing mass.

Table 6.3 Protein ion list for the study of mass measurement accuracy, reported in Figure 6.4

Protein ion	***m/z***
Insulin	5733
Apomyoglobin (+2)	8473
Thioredoxin	11668
Cytochrome C	12225
Apomyoglobin (+1)	16943
β-Casein	24078
BSA (+2)	33149
BSA (+1)	66415
IgG (+2)	74395
IgG (+1)	148804

(a)



(b)

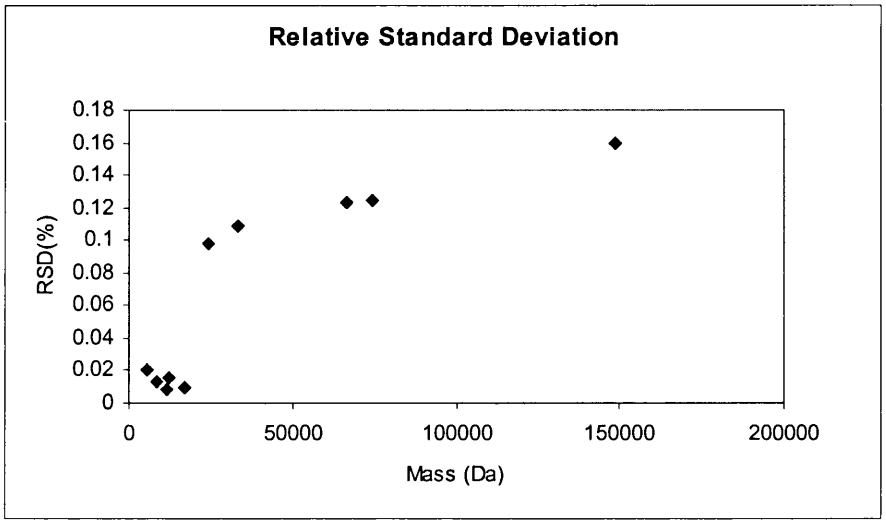


Figure 6.4 Standard deviation and relative standard deviation of mass measurement using calibation as a function of increasing molecular weight for the protein standards listed in Table 6.3

We aim to examine a number of strategies to improve mass measurement at high mass (>20,000 Da) and have employed the ubiquitous BSA (monoisotopic mass 66,384 Da, average mass 66431 Da) as a model compound. Albumin is an abundant blood plasma protein serving as an agent for osmotic regulation and fatty acid transportation. BSA has

been previously studied by electrospray analysis[31-33] and its elemental formula deduced from an ion spray experiment $C_{2935}H_{4583}N_{780}O_{899}S_{39}$[34].

6.4.2 Investigation of factors affecting the peak shape, mass resolution and assignment of the mass scale of high mass MALDI mass spectra

6.4.2.1 Peak shapes of high mass BSA

Normally the molecular ion in a mass spectrum can provide information on the molecular weight of a compound. By examining the relative abundance of isotopic peaks, or the isotopic profile, information concerning the elemental composition can be deduced. But the important prerequisite to using this approach is the identification of the peak corresponding to the monoisotopic ion. If the mass is high, in this case MW=66431 Da, the distribution of the isotopic envelope is Gaussian, the first visible mass is not the monoisotopic mass because it is at a low abundance and in the MALDI spectrum at the noise level[35]. Even the most abundant mass in the molecular ion envelope is not distinguishable because of a lack of mass resolution and the huge combinations of isotope peaks present at high mass.

Figure 6.5 shows the MALDI mass spectrum of BSA, a zoom of the molecular ion peak is shown in Figure 6.6 and shows the peak of BSA $[M+H]^+$ is very wide, the peak width at half maximum is about 850 Da. The peak is also observed to be asymmetric with a high mass tail extending several thousand mass unit.

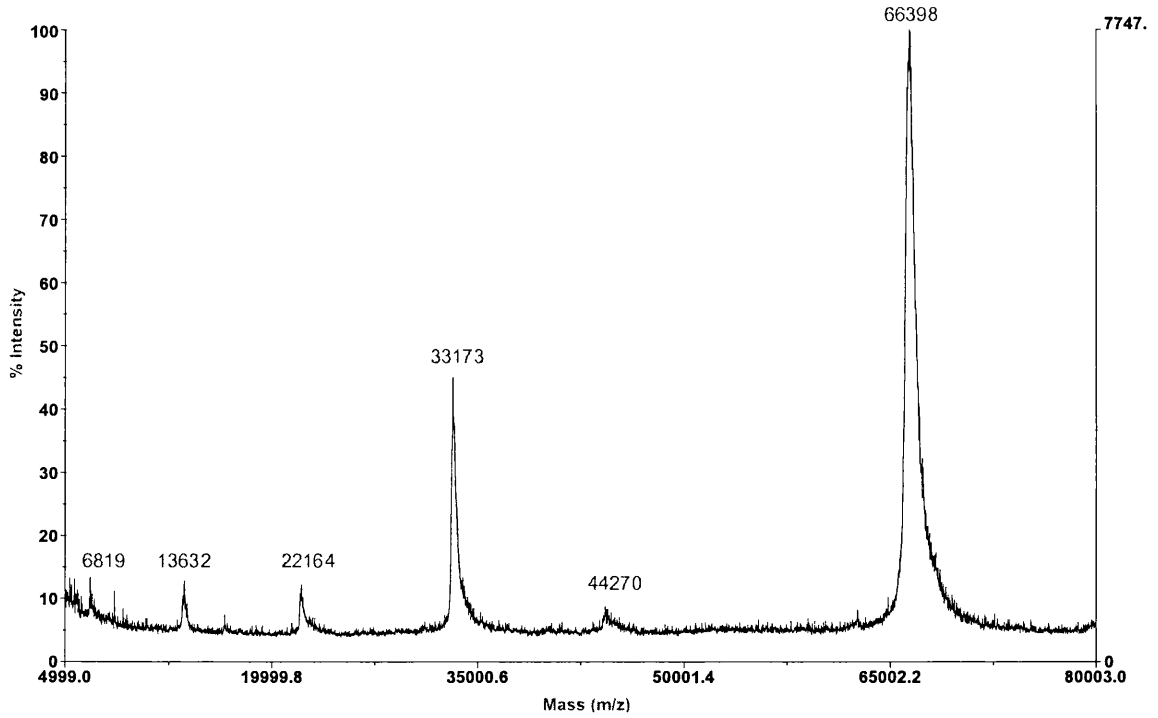


Figure 6.5 Raw MALDI-TOF spectrum of BSA in linear mode

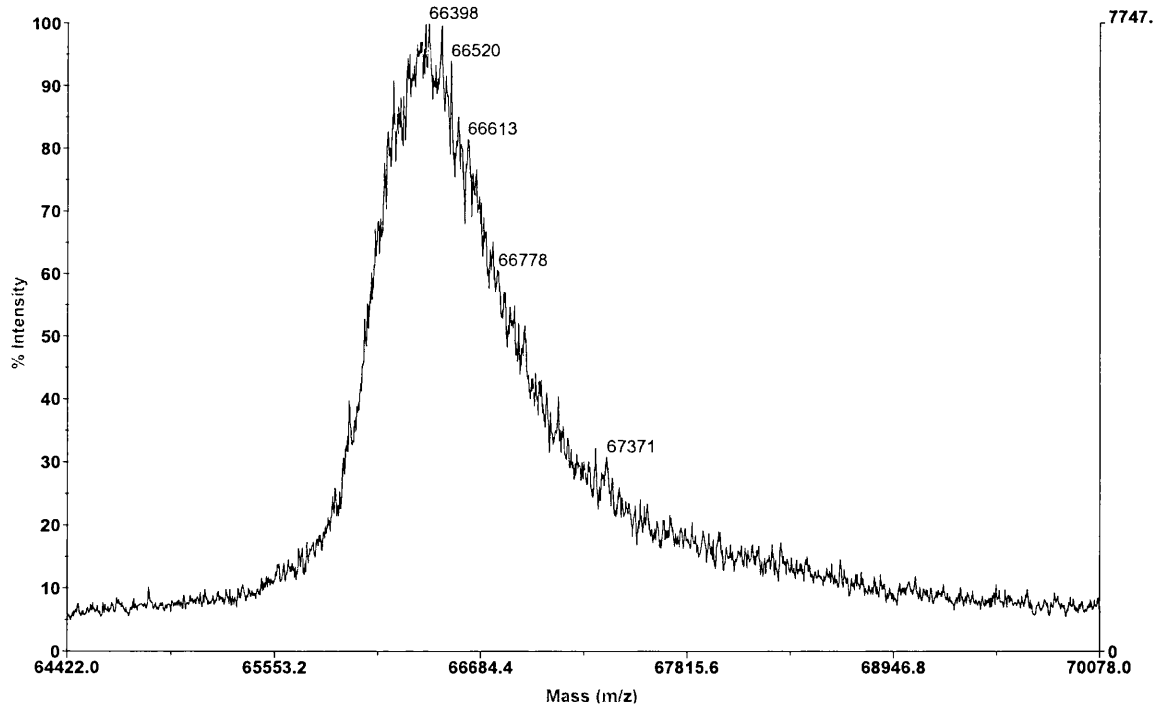


Figure 6.6 Expansion of the molecular ion region, of the singly charged BSA peak (from Figure 6.5)

For high mass compounds, such as proteins, unit mass resolution is not possible, but the presences of isotopes at their natural abundances should still display a Gaussian shape isotopic envelope. Different algorithms have been used to obtain the shape of the isotopic envelope. The Voyager DE STR mass spectrometer has a built in program, which is able to predict the theoretical isotopic envelope from a molecular formula of the compound. The theoretical isotope envelope of BSA was obtained using this program, as shown in Figure 6.7 and the FWHM is only about 20 Da.

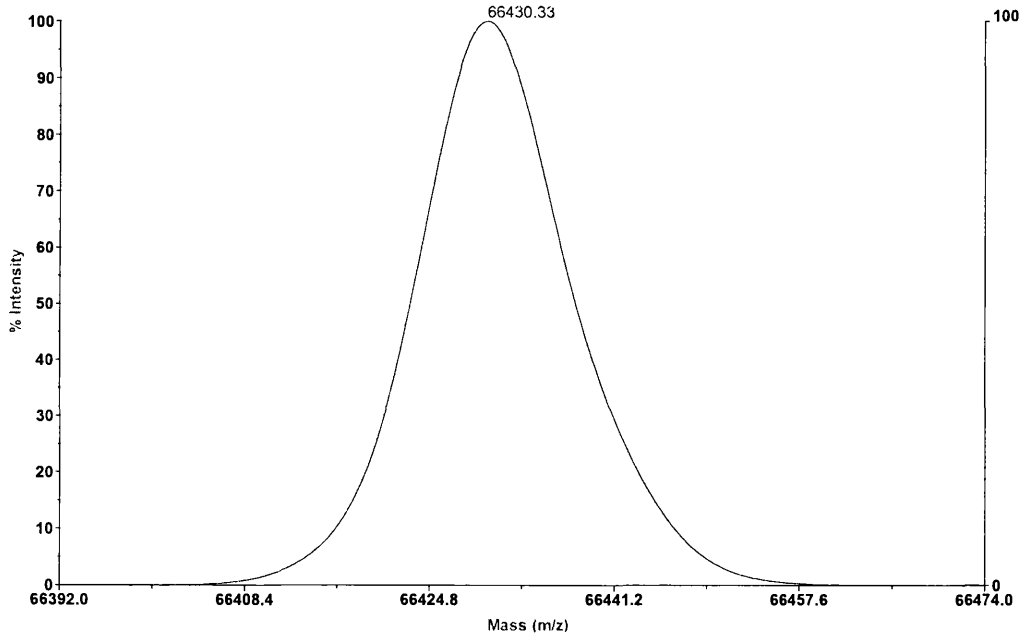


Figure 6.7 Theoretical isotope distribution for BSA predicted from Data Explorer software

When the theoretical isotope envelope is superimposed onto the MALDI spectrum of BSA, as shown in Figure 6.8, an obvious difference can be observed in the predicted envelope and the experimental spectrum.

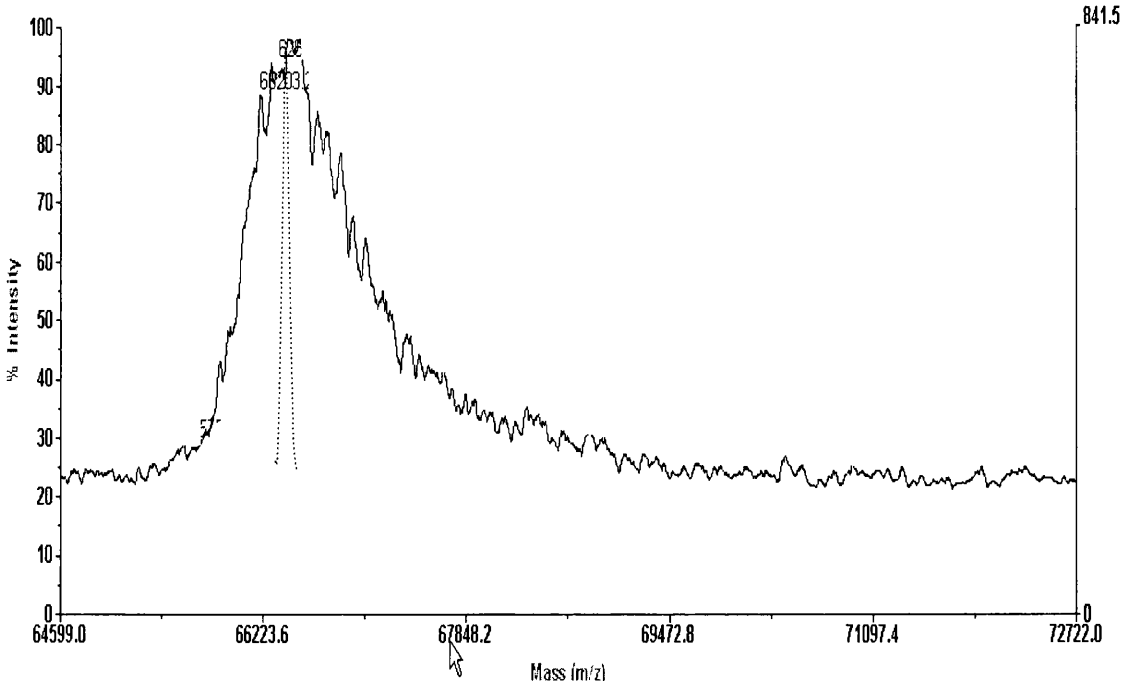


Figure 6.8 Comparison of theoretical isotope envelope and isotope envelope displayed by MALDI-TOF mass spectrometry

This difference could be due to

a. Insufficient or degradation of mass resolution at higher mass due to initial kinetic energy distribution of the ablated neutrals/ions.
b. Protein adduct formation, e.g. matrix adducts.
c. Some other physical effect in a linear TOF arrangement e.g. broadening of the peak due to ion fragmentation.

In section 6.4.4 the effect of the above was calculated and compared, where appropriate, to experiment. The effect of matrix adducts does not account for the lower-mass broadening and the literature indicates that typically one matrix unit is added. There is no evidence of adduction of many matrix molecules.

6.4.2.2 Establishing a mass scale in "linear TOF" mode

There are different choices to establish a mass scale for linear TOF when examining proteins at high mass. We can use the existing mass scale and extrapolate to high mass or use other high mass compounds to calibrate against. Also we can use a self calibration procedure, which can be applied to test the mass accuracy of the "linear TOF" mode for high mass proteins. The formula of BSA is $C_{2935}H_{4583}N_{780}O_{899}S_{39}$, giving an average MW of 66430.3 Da, in positive MALDI-TOF, the $[M+H]^+$ is 66431, and this was used for calibration. In our initial study using BSA, it was found that the method of smoothing affected the reported mass, for example, of the isotopic distribution maxima.

6.4.2.2.1 Smoothing

Post-processing of mass spectra is critical for optimal data interpretation and presentation. The Voyager 5.0 software includes routine data smoothing methods which can reduce background noise and improve the appearance of spectra. It has two methods for smoothing: a) Noise filter and b) Gaussian smoothing. The application of both of these is described below.

The "noise filter" programme is designed to identify correlated and uncorrelated features in data so that only the uncorrelated noise components are filtered. The degree of correlation is the parameter that determines how efficiently the noise is filtered with

respect to the signal. This value typically varies from 0 to 1 with a common default of 0.7. Figure 6.9 shows the BSA spectrum after the noise filter was applied using a coefficient of 0.7, no obvious improvement can be observed from the raw spectrum (see Figure 6.5), and the maxima is difficult to choose and has a large uncertainty.

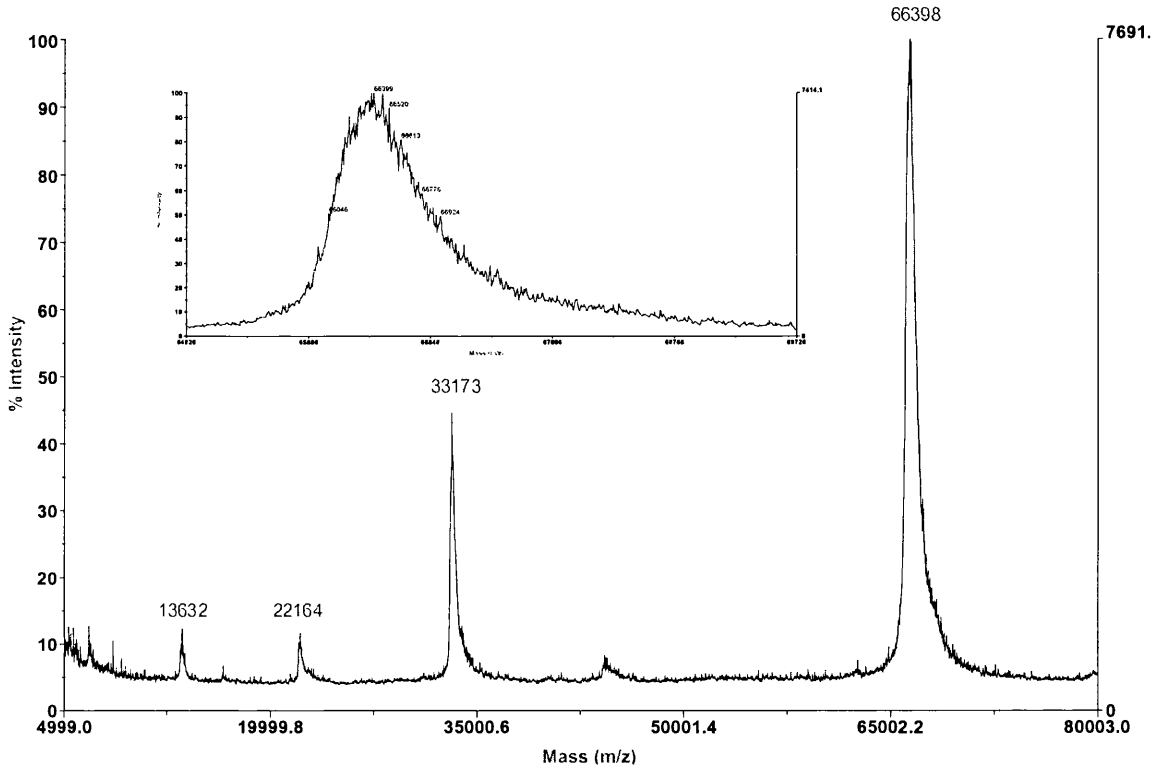


Figure 6.9 Mass spectrum of BSA after noise filter (correlation 0.7)
Insert: zoom of $[M+H]^+$ peak

The Gaussian smooth programme is a n-point filter, it is a widely used and well known data-smoothing process. The most general type of n-point smoothing was developed by Savitzky-Golay[36], it is sometimes called a 'moving window' average. In this process, the program produces a new point by averaging the neighboring n-points (n: number of

points to be averaged, always an odd number). The simplest form is 3-pt smoothing which takes the average of the "point" plus its two neighbours. n-pt smoothing generally applies a "weight", to the neighbouring points, which reduces as the neighbour becomes further from the central "point".

For high mass proteins where the peak is very wide, a wider averaging window needs to be selected for satisfactory results. For BSA, it was observed that smoothed spectra using 49-point smoothing did not cut the peak height by much, as shown in Figure 6.10, and was adopted. The resulting peak shape is symmetrical and the maximum is easier to identify with a small uncertainty.

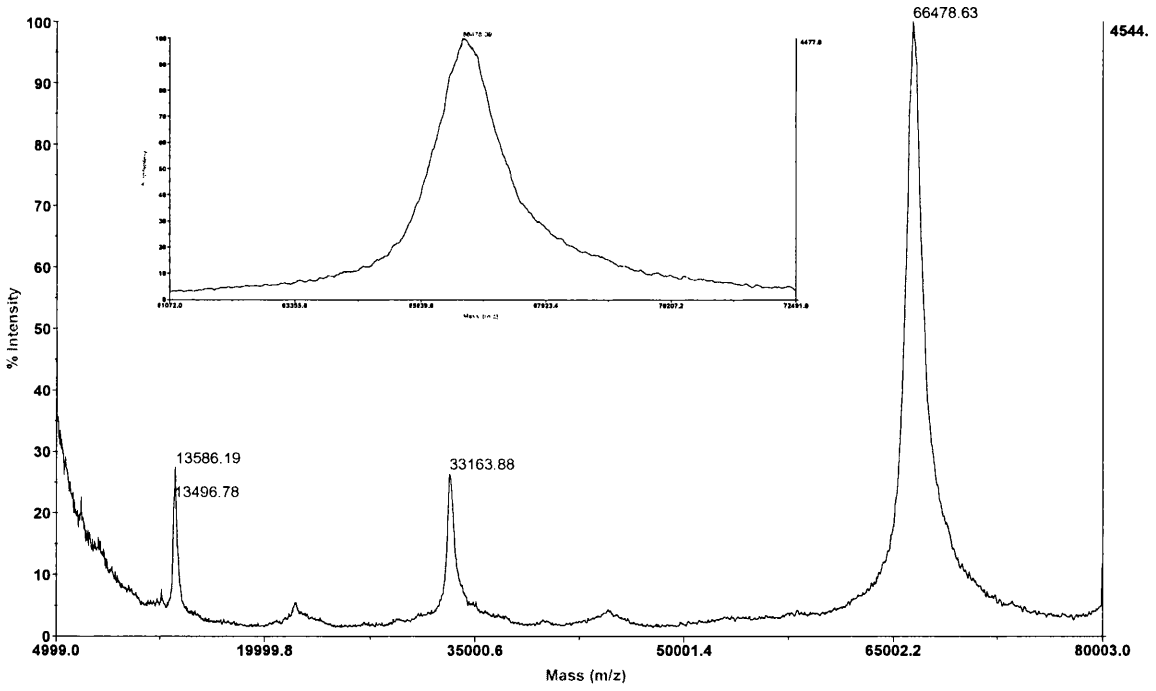


Figure 6.10 Mass spectrum of BSA after Gaussian smooth (49-point)
Insert: zoom of $[M+H]^+$ peak

To compare the two smoothing methods, we used BSA as a standard and used it to self-calibrate, i.e., to measure the mass error of BSA as a separate measurement. The measurement of BSA was repeated 30 times, for each of the methods and the results are shown in Figure 6.11. The percentage error for method 1 is within ±0.4%, and for method 2 within ±0.2%, respectively. It is difficult to assess the accuracy and precision of methods 1 and 2 from Figure 6.11, although method 2 appears to have better precision (less scatter of data points).

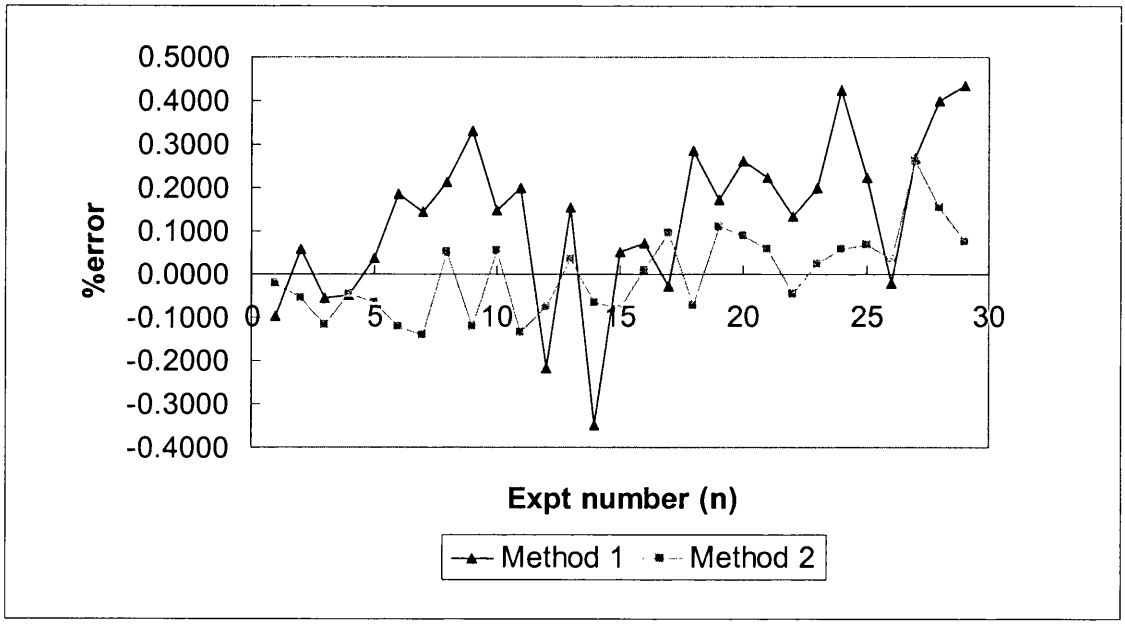


Figure 6.11 Linear TOF mass measurement accuracy of BSA, calibrating by Method 1 and Method 2.
Method 1. using the Noise filter algorithm (N=0.7)
Method 2. using the Gaussian smoothing (using n points-where n is odd and n~ MW/1000)

To assess this data further the accumulated average mass of BSA was calculated for each method and the results are shown in Figures 6.12 and 6.13. It can be clearly seen that the two methods bring the mass error under control and converging to a mass error of less than 1 Da after 20 to 30 measurements. Whereas the first series of measurements by method 1 appears to be a process "out of control", where the mass error continues to drift upward and exceeds +90 Da after 28 measurements.

The above observation left us with an intriguing problem. In practice 3-6 measurements are typically used. Figure 6.13 shows that higher than 28 measurements need to be averaged to give a result with a mass error of 1 Da, at high mass. Therefore other methods to process calibration data other than smoothing of the mass spectrum in the molecular ion region (66000-69000 Da) were studied to see if more effective methods can be developed. These are described in Section 6.4.3.

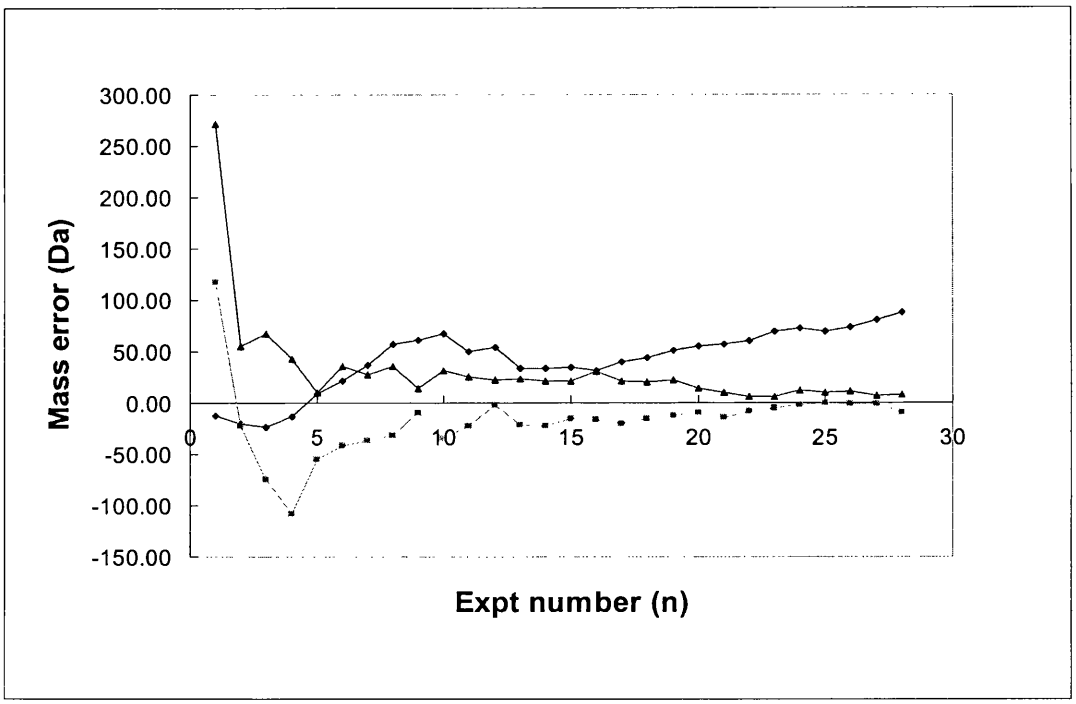


Figure 6.12 Cumulative mass error (Da) of the average molecular weight as a function of the number (n) of averaged mass scan using the "noise reduction" algorithm

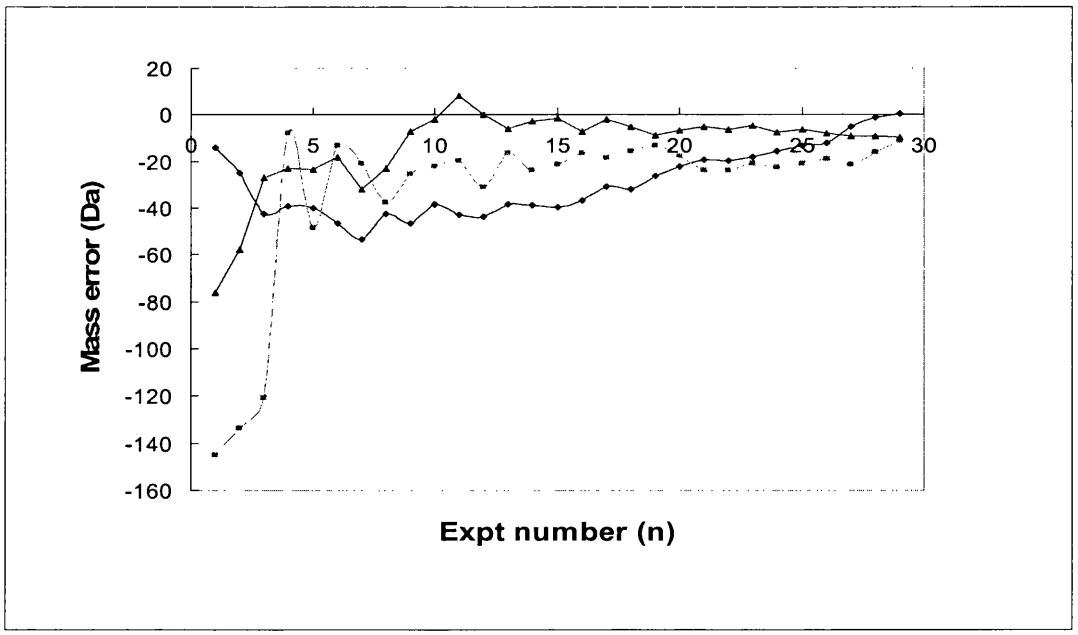


Figure 6.13 Cumulative mass error (Da) of the average molecular weight as a function of the number (n) of averaged mass scan using the "Gaussian smoothing" algorithm

6.4.2.2.2 Optimisation of peak shape

To obtain accurate mass values in the calibration spectra well-shaped and symmetrical peaks with best resolution are ideally needed. Smoothing of the spectra was carried out on raw data before calibration was made. In this study, it can be observed that the grid voltage and delay time are two parameters which affect the mass resolution of the mass spectrum. These two latter parameters were systematically investigated as a function of mass, the result is shown in Figures 6.14 and 6.15. Figure 6.14 shows that the change in peak width for masses of 33 kDa, 66 kDa, 74 kDa and 148 kDa. For the lower masses, the trend line is flat, but for higher masses distinct minima in the peak width, i.e., best resolution, can be observed at 70% grid voltage. So 70% is a good choice for all masses which offers mass independent resolution.

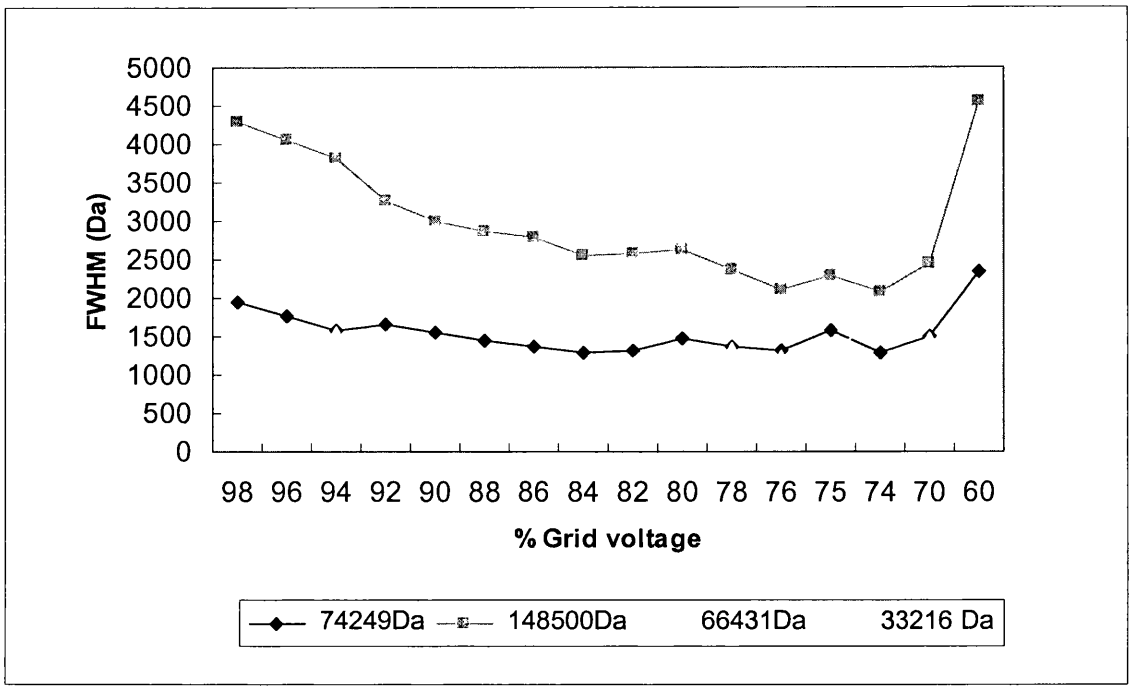


Figure 6.14 Peak width as a function of grid voltage. Optimisation of grid voltage for protein samples at four different *m/z* values with delay time 800 nsec. The peak width was measured at full width half-maximum (FWHM).

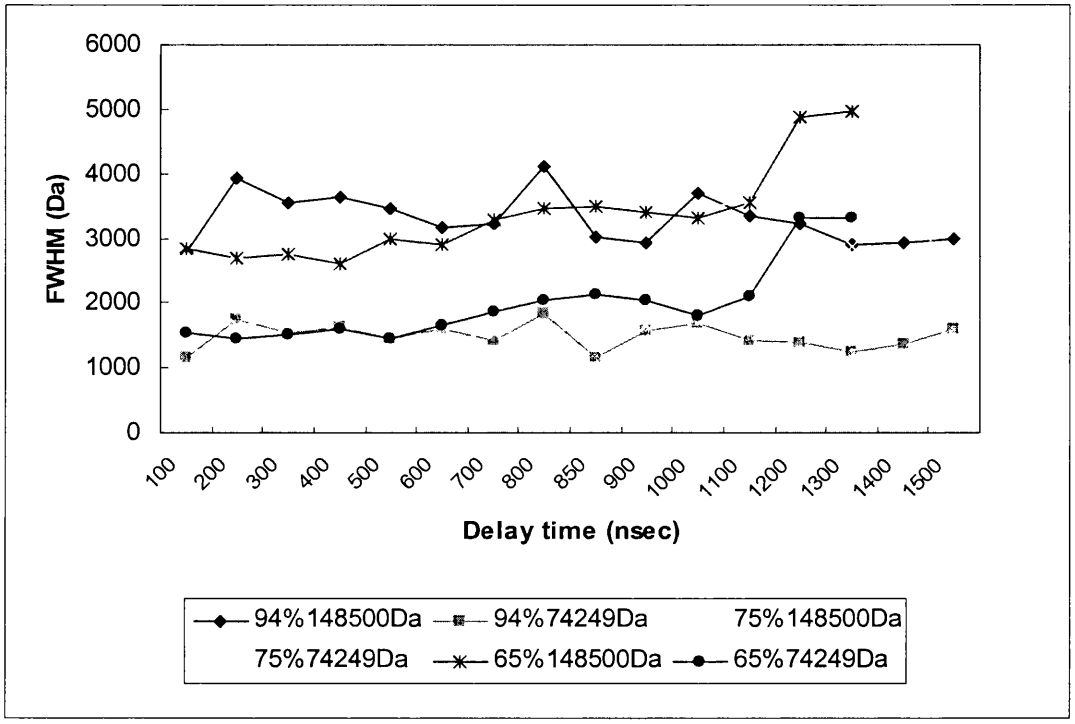


Figure 6.15 Peak width as a function of delay time for protein samples at two different *m/z* (74249 Da, 148500 Da) and three different grid voltages (94%, 75%, 65%) The peak width was measured at full width half-maximum (FWHM).

Figure 6.15 shows the effect of delay time on resolution, the curves are essentially flat, which indicate delay time is not as influential as grid voltage in resolution optimisation. No optimisation between grid voltage and delay time can be easily made as they are both appear mass dependent for these experiments. Ideally the grid voltage or delay time need to be changed with mass. This is impossible with current instrument design because the masses are not mass separated in the ion source, so switching either of the above two instrument parameters will be ineffective.

Even after optimisation and smoothing, the mass resolution is still low, and with n-point Gaussian smooth, approximately 30 acquisitions are needed to be carried out to obtain a satisfactory mass accuracy. This means 30x100, that is 3000 individual laser shots, are required and this is inefficient. In light of this finding, other methods were sought to determine the calibration mass and proposition for practical mass accuracy measurements are discussed below.

6.4.3 Investigation of fitting procedures to improve mass measurement at high mass

There are two alternative ways to detect the peak top instead of Gaussian smoothing. One is fitting to a function, curve fitting procedures are needed to fulfill this. The other one is by differentiation of the raw data, the peak top can be characterized when the first derivative is equal to 0, both of these approaches are investigated in this study.

6.4.3.1 Curve fitting procedures

Curving fitting is a mathematical area and there are readily accessible methods such as MS-Excel and a fitting program, used in this study, called EasyPlot[37].

In Excel, when a spectrum is drawn from raw data, a trend line with a formula can be created from the raw data. The trend line can be fitted to a number of predefined functions, i.e., linear, polynomial, exponential, logarithmic, power or exponential form which can be chosen by the user. The choices of formulae by different data trend processing are:

i. Linear, $y=mx+c$

ii. Logarithmic, $y=ln\ (x)$

iii. Polynomial, $y=c_0x_0+c_1x_1+c_2x_2+...+c_nx_n$

iv. Exponential, $y=x^n$

v. Power, $y=exp\ (x)$

Another method that has been used in this study is EasyPlot. It is a powerful data analysis software package which can be applied to many different kinds of scientific data. A feature of it is that it can create a curve from any arbitrary set function or data set. It also has algebraic methods built in, such as, differentiation which was applied to investigate where the peak top is located. The curve fitted function found in this study is of the form

$$y = \frac{a}{(x - m_1)^n}\left(\frac{1}{e^{b/(x-m_2)} - 1}\right) \qquad \text{(Eq. 6.9)}$$

where y is intensity, x is mass scale, m_1 and m_2 are mass offsets, *a* and *b* are constants. *a*, *b*, m_1 and m_2 are determined by the fitting procedure. This function can be parameterized to use in the calibration. The curve fitting result by EasyPlot is shown in Figure 6.16, it shows a similar trend curve (dotted) as "Gaussian smoothing" algorithm as Data explorer software and it takes only seconds to obtain the peak top. EasyPlot does not average the raw data in the same manner as the Gaussian smooth method, which makes the curve more similar to the actual peak shape.

Another function found for fitting is Maxwellian-like and of the form

$$y = x_0 + ax^n e^{-bx^m} \qquad \text{(Eq. 6.10)}$$

where x_0 is a mass offset, *a* and *b* are constants, *n* and *m* lie in the range 0 to 5 and would be equal to two for a Maxwellian peak shape.

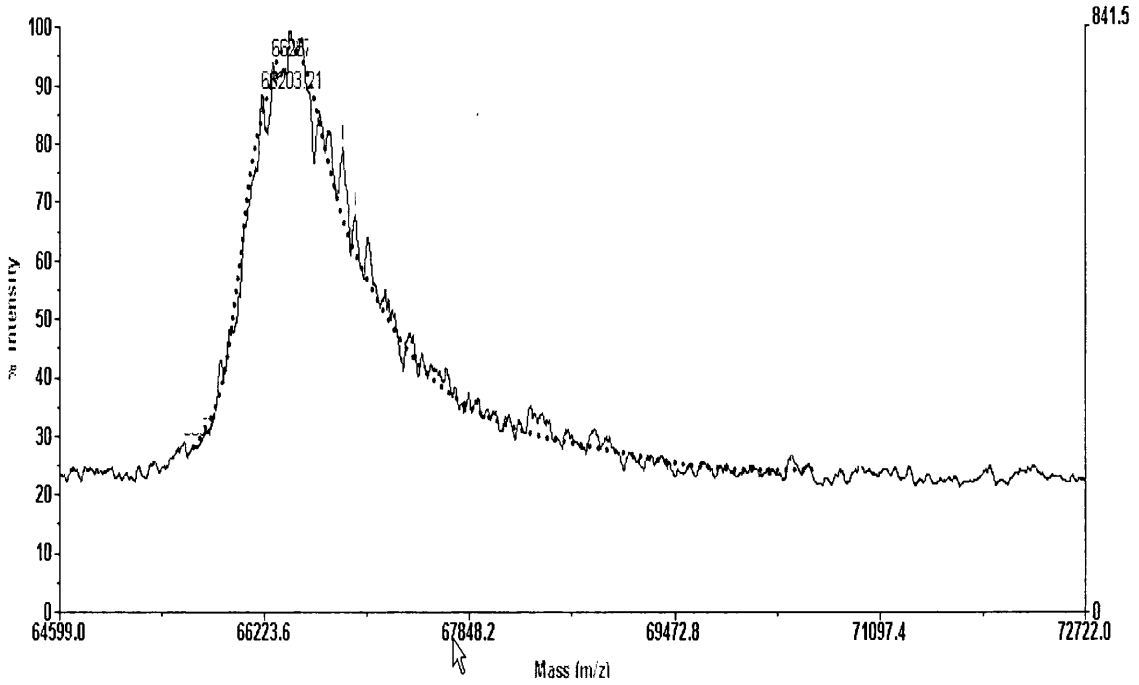


Figure 6.16 Curve fitting (dotted) using the EasyPlot software

6.4.3.2 First derivative method

The peak top is conceptually the zero point of a numerical first derivative. Differentiation is carried out by EasyPlot and the result is shown in Figure 6.17. The peak top can be exactly characterized by the point y=0. This is a potential alternative to a curve fitting procedure to find the peak top, i.e., a measure of the average mass of a calibration ion.

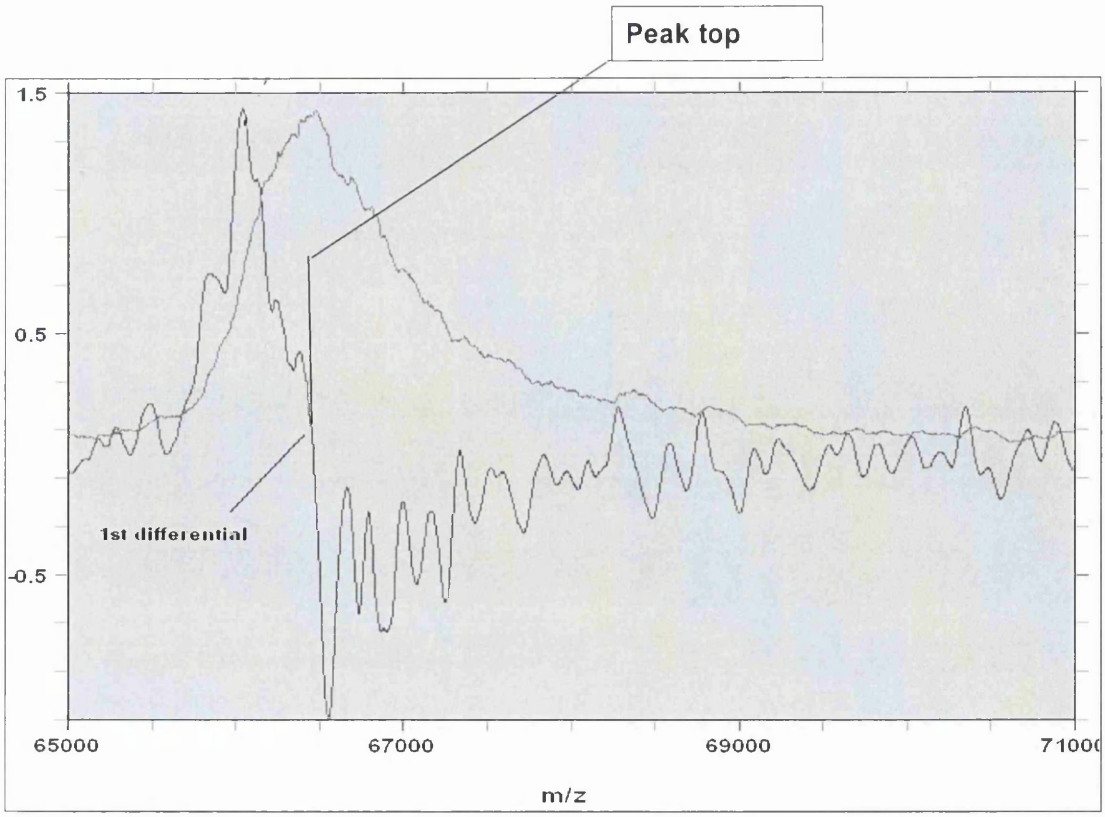


Figure 6.17 Calibrating to a fixed point using the first derivative of the mass spectrum

This is just a theoretical investigation and was not applied in practice since these procedures could not easily be integrated into Data Explorer. But it provides some ideas for programming the Applied Biosystems mass spectrometer software for future application.

6.4.4 Investigation of the peak broadening phenomenon observed in linear TOF at high mass

The current investigation is not exhaustive and only preliminary in nature. However, because of the significance of this observation and the impact on high mass MALDI, it

was considered in some detail. There are several possible causes of peak broadening considered in this study and they are discussed below.

6.4.4.1 Poor mass resolution performance of linear TOF: An ion source effect or initial kinetic energy distribution?

The design of ion sources is critical for good performance by any ionisation technique in mass spectrometry. This is fundamentally an ion optical problem. Cotter has reviewed the fundamentals of TOF ion source design extensively in his two books[23, 38]. However, he does not discuss the phenomenon observed here in terms of an ion optical problem. In 1955, Wiley and Mclaren[39] described two focusing techniques for time-of-flight mass spectrometry, that is, space and time-lag focusing. These two techniques solve the problem of space distribution of ions in the ion source. The reflectron time-of-flight solved the affects of kinetic energy distribution of ions[17]. It is not our intention to investigate ion optics here and mass dependent behavior at high mass seems improbable as it would represent a serious defect in the ion optics of this high performance design. Usually mass dependent behavior occurs when magnetic fields are present and are generally most evident at lower mass. It has been stated that the initial velocity spread (u_0) of an ion leads to an initial kinetic energy spread of $\frac{1}{2}mu_0^2$, Karas *et al*[40] have measured the initial velocity spreads of ions in the initial phase of the expansion of a MALDI plume. Their results show that for most matrices α-cyano-4-hydroxycinnamic acid, 2,5-dihydoxybenzoic acid, the velocity distribution ranging from approximately

300 to 600 m s^{-1}. Thus for increasing mass the kinetic energy distribution will increase. Table 6.4 gives the kinetic energy distribution calculated over the mass range up to 10^6 Da, additionally the effective time "spread" and mass "spread" corresponding to these kinetic energy distributions has been calculated and is given in Table 6.4. It can be seen from Table 6.4 that the expected mass spread due to the initial kinetic energy spread of proteins with a mass about 66000 Da could be as great as ±100 Da, at the base of the peak, accounting for almost 10% of the observed peak broadening. However, delayed extraction which is built into the instrument should help to reduce the effects of initial kinetic energy spread. This observation suggests that the effect of peak broadening on varying the initial delay time should be studied in detail. However, the data in Fig 6.15 shows that the effect of varying delay time is flat with little effect on the peak shape at half height.

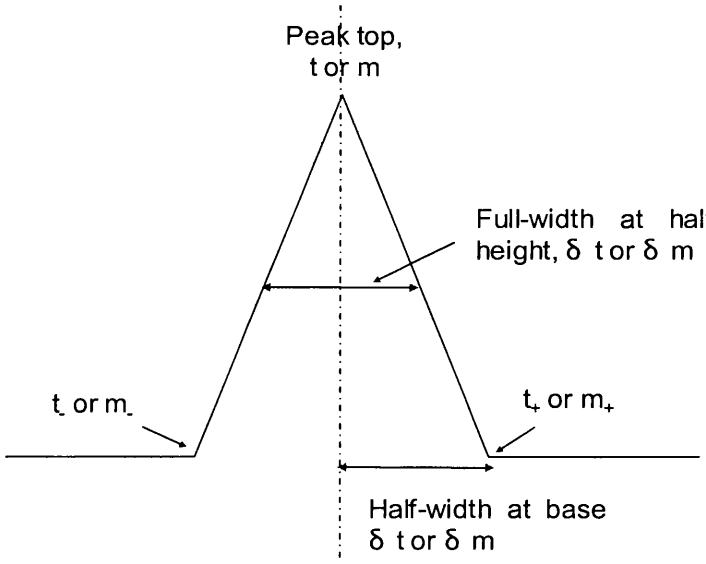


Figure 6.18 Nomenclature used to denote the broadening of a peak in terms of time and mass spread.

Table 6.4 Estimation of mass broadening of a MALDI peak as a function of the initial kinetic energy (KE), arising from an assumed velocity spread of 300m/s of the ablated molecule (or ion), over the mass range 2000-10^6 Da

Mass (Da)	Time-of-flight (s)	Velocity (m/s)	KE (eV)[1]	Time-of-flight[2] upper (s), t_+	Time spread (s)[3,5] δt	Lower mass limit (Da)[4] m-	Mass spread (±Da)[5] δm
2000	4.072E-05	300	0.93	4.072E-05	-7.594E-10	1999.93	0.07
5000	6.438E-05	300	2.33	6.438E-05	-3.00179E-09	4999.53	0.47
10000	9.105E-05	300	4.66	9.104E-05	-8.48975E-09	9998.14	1.86
20000	1.288E-04	300	9.33	1.287E-04	-2.40093E-08	19992.54	7.46
50000	2.036E-04	300	23.31	2.035E-04	-9.48653E-08	49953.41	46.59
100000	2.879E-04	300	46.63	2.876E-04	-2.68132E-07	99813.83	186.17
200000	4.072E-04	300	93.26	4.064E-04	-7.57335E-07	199256.71	743.29
500000	6.438E-04	300	233.15	6.408E-04	-2.98117E-06	495380.16	4619.84
1000000	9.105E-04	300	466.29	9.021E-04	-8.37398E-06	981689.83	18310.17

Notes:

1. KE is calculated via $\frac{1}{2}mv^2 = eV$
2. Time-of-flight of an ion receiving maximum kinetic energy.
3. Time-of-flight of spread of half the peak width, at peak base. As the kinetic energy has increased the time spread has a negative value.
4. Mass equivalent to the time-of-flight listed. Mass limit is referred to as lower as the mass will be reduced when the ions kinetic energy is increased.
5. Half width of the peak, at base (refer to Figure 6.18 for an explanation of terms).

Table 6.5 shows the mass broadening effect of different initial MALDI plume velocity over a mass range from 2000 to10^6 Da. Higher initial velocity spread will cause even greater peak broadening, whereas lower velocity spreads will cause less broadening. For high mass proteins such as BSA, a velocity spread 300-600 m s^{-1}should be considered although Karas *et al.* indicates the velocity spread is ~300 m s^{-1} for CHCA as matrix. The work of Karas *et al.* was obtained for lower mass species, in general, than those studied here. There are no experimental measurements of velocity spread at very high masses, especially for the higher mass range investigated here. The peak broadening for velocity spreads of 300 to 600 m s^{-1} is approximately 100 to 500 Da, as seen from the model calculations listed in Table 6.5. A full theoretical description of this effect would involve calculation for a distribution of initial velocities, e.g., a Maxwell-Boltzman like distribution. As yet there is little experimental data in this area, which would be needed for such calculations, however trends can be seen by the relatively simple model given here.

Table 6.5 Estimation of mass broadening of a MALDI peak as a function of the initial kinetic energy (KE), arising from assumed velocity spreads ranging from 200 to 1000 m s^{-1} of the ablated molecule (or ion), over the mass range 2000-10^6 Da

Mass	Mass width, δm, (Da)[1] at initial average velocity (m s^{-1})				
	200	400	600	800	1000
2000	0.03	0.13	0.30	0.53	0.83
5000	0.21	0.83	1.86	3.31	5.18
10000	0.83	3.31	7.46	13.25	20.68
20000	3.32	13.25	29.80	52.91	82.55
50000	20.72	82.76	185.82	329.40	512.79
100000	82.83	330.49	740.54	1308.98	2030.33
200000	331.04	1317.60	2940.40	5168.27	7959.72
500000	2063.86	8154.44	17980.94	31096.33	46945.70
1000000	8221.48	32094.35	69427.02	117102.42	171664.94

Notes:

1. Half width of the peak, at base.

Figure 6.19 shows the theoretical peak broadening as a function of mass, the curve is quadratic which is consistent with what was observed in the experiment, as shown in Figure 6.4 (a).

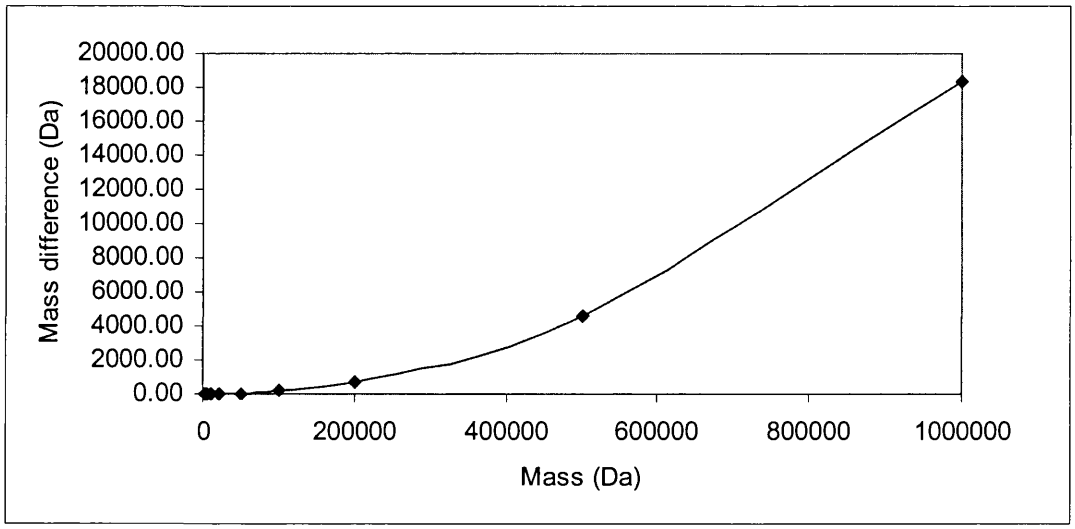


Figure 6.19 Mass differences (spread of peak at half base width) as a function of molecular mass for a velocity spread of 300 m s^{-1}

6.4.4.2 Fragmentation and metastable decay of the molecular species

Most ions have a given lifetime before they break (fragment) unless they are very stable, in which case the molecular species persists. Once ions are formed they tend to fragment, the fragmentation is a function of time, the usual experimental observation window ranges from 10^{-12} sec to seconds.

Prompt fragmentation is generally $\leq 10^{-8}$ sec and occurs in the source. If fragmentation occurs instantly in the MALDI source, the fragment ions will be accelerated to the same final kinetic energy as the molecular ion, both molecular ion and fragment ions will be recorded in the mass spectrum.

Fragmentation processes are driven by the amount of internal energy contained in the precursor ion. The more internal energy, the faster fragmentation is, the lower internal energy is, the slower the fragmentation will be[39]. The fragmentation will continue to the point where the product ion is stable. Molecular ions with very low internal energy, i.e., below the reaction barrier, will not fragment. Such ions can be created in special ion sources, for example, found in soft ionisation techniques such as chemical ionisation and electrospray.

A theory of ion fragmentation was developed by Rosenstock *et al.* in 1952 and is referred to as quasi equilibrium theory (QET). Slower fragmentation is referred to as

"metastable" decay and may contribute to ion loss in some mass spectrometer designs[41]. Initially, the term "metastable" decay was coined to describe processes which occur in the field free regions of a sector mass spectrometer[23]. However, it is now used to describe fragmentation of an ion between leaving the ion source and reaching the detector, so it covers a range, typically, 10^{-7} to 10^{-3} sec.

It has been observed by Demirev *et al.*[22], and other references cited therein, that the higher the mass of a protein, the greater the degree of metastable fragmentation. An important observation of ion fragmentation for high masses in TOF is the difference between reflectron-TOF and linear-TOF-MS. Reflectron-TOF does not show many ions in the high mass region whereas linear-TOF shows a greater proportion. In linear-TOF, if the following fragmentation occurs in the field free region,

$$M^+ \rightarrow F^+ + N$$

the parent ion M^+, fragment ion F^+ and neutral species N will have the same mean velocity[23]. In theory, they will arrive at the detector simultaneously and will not cause poor mass resolution. However, it is well documented[23] that when ions with high kinetic energy, of several thousand electron volts (keV) fragment in a field free region the fragments will have a very broad peak shape. This is discussed in more detail below.

The process of ion fragmentation is similar to radioactive decay and in the simplest case can be described by a single half life, τ. The number of ions, N, surviving after a time t

is given by

$$N = N_0 e^{-t/\tau} \qquad \text{(Eq. 6.11)}$$

where N_0= number of ions at t=0. Figure 6.20 shows ion intensity curves, calculated for the Voyager mass spectrometer in linear mode (path length 2 m). For simplicity, the plots assume the ion source region has zero length so that all ions are immediately accelerated to an energy eV. In practice the ion source region is divided into two ion acceleration regions. The curves in Figure 6.20 show the intensity of the survivor molecular ion of BSA for ion lifetimes (τ) varying between 10^{-3}-10^{-7} sec. It can be seen that the intensity falls off rapidly with shorter ion lifetime, although even for τ=10^{-4} s over 50% of ions decay before reaching the detector.

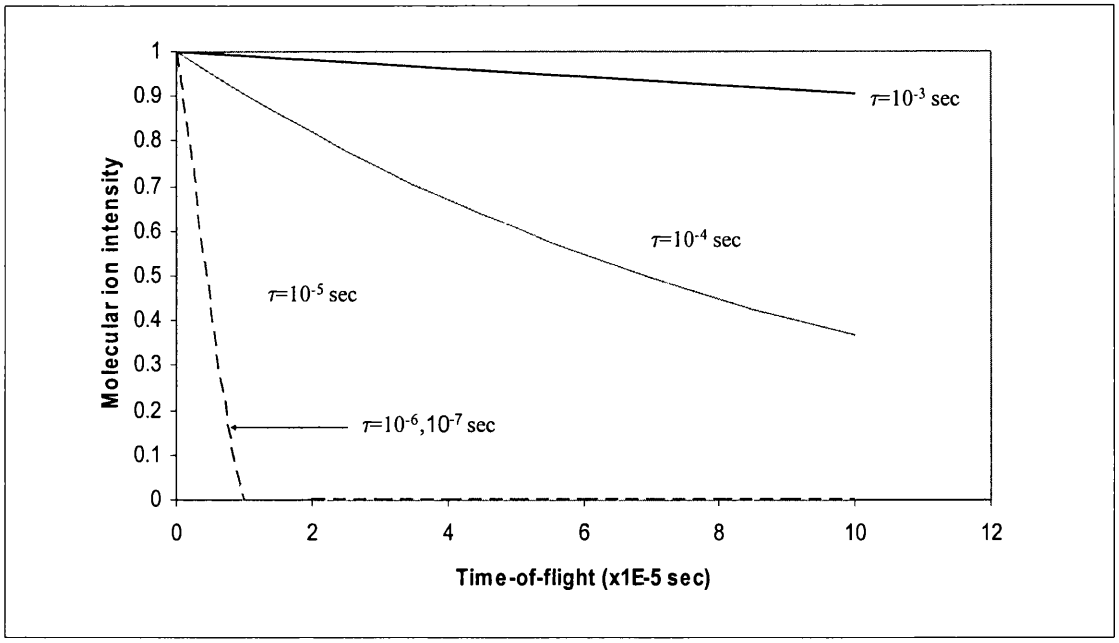


Figure 6.20(a) Illustration of the effect of ion lifetime, τ, on the intensity of the molecular ion species (survivor ions) for theoretical values, $\tau=10^{-3}$, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7} sec as a function of ion flight time. The curves have been calculated for BSA (MW=66431 Da), a path length of 2m and an acceleration voltage 25 kV. It is assumed the molecular ion is promptly accelerated to its final kinetic energy.

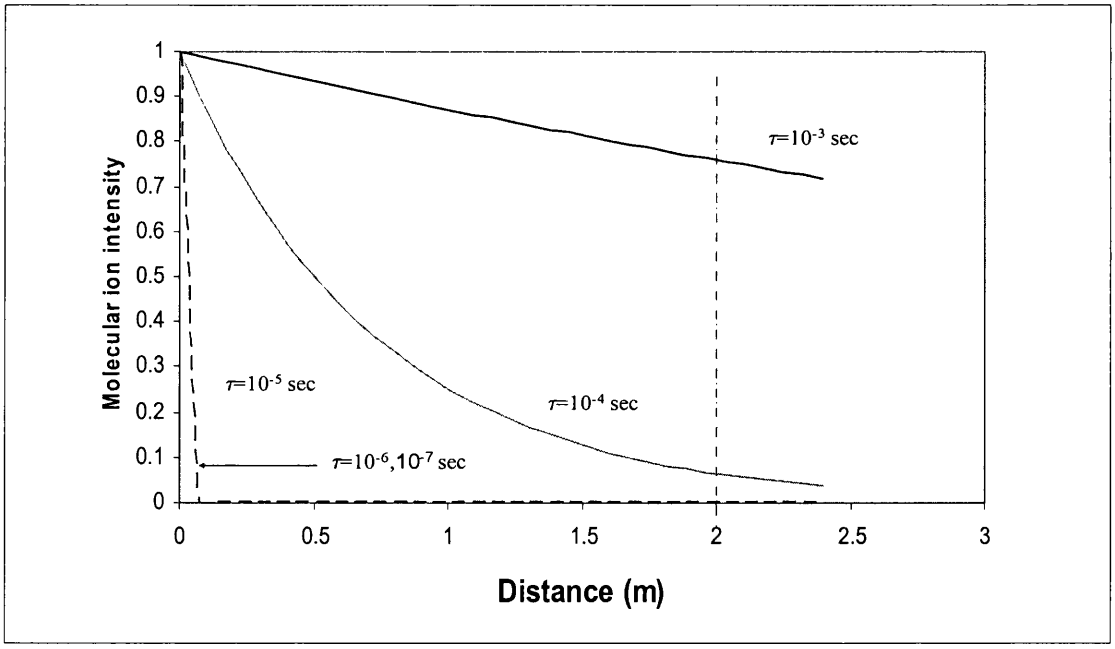


Figure 6.20(b) Illustration of the effect of ion lifetime, τ, on the intensity of the molecular ion species (survivor ions) for theoretical values, $\tau=10^{-3}$, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}sec as a function of distance travelled by the ion. The curves have been calculated for BSA (MW=66431 Da), and an acceleration voltage 25 kV. It is assumed the molecular ion is promptly accelerated to its final kinetic energy. Dashed line shows the detector position in the linear mode for the Voyager STR DE MALDI-TOF-MS

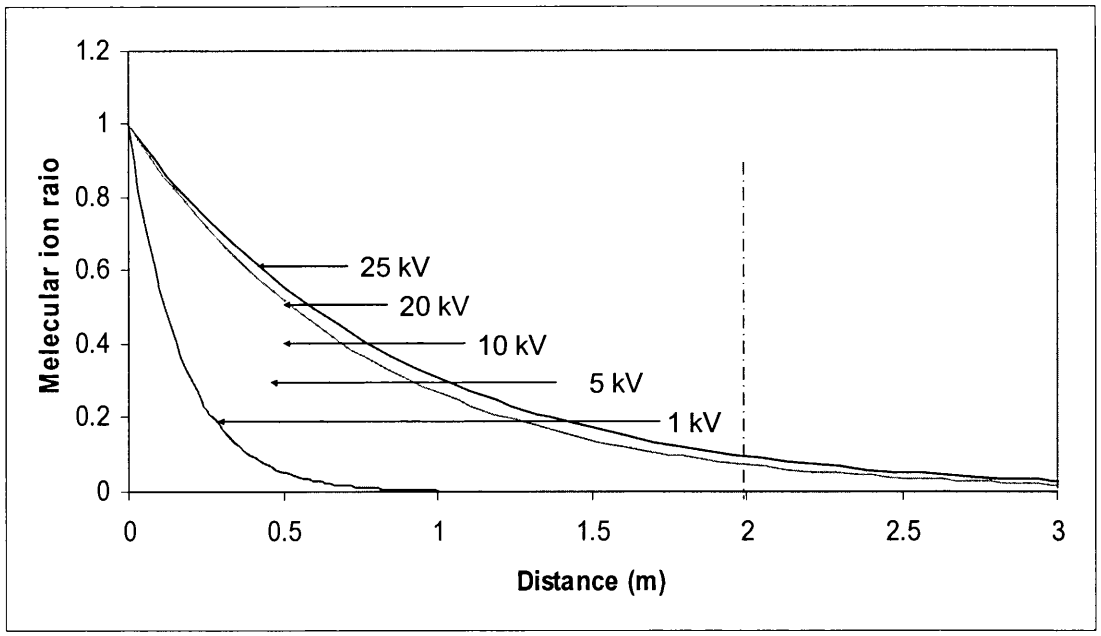


Figure 6.21 Effect of accelerating voltage, V on the intensity of the molecular ion (survivor ions) for V=1, 5, 10, 20 and 25 kV. The curves have been calculated for BSA (MW=66431 Da), a path length of 2 m and ion lifetime,τ=10^{-4} s. It is assumed the molecular ion is promptly accelerated to its final kinetic energy. Dashed line shows the detector position in the linear mode for the Voyager STR DE MALDI-TOF-MS

Figure 6.21 shows the effect on BSA survival as a function of accelerating voltage. From these figures it can be seen that fragmentation processes may extend all the way from the time of ionisation to the time that ions reach the detector, but for ions with lifetimes faster than10^{-5} sec, most of ions will fragment in the acceleration field.

If the fragmentation occurs in the acceleration region, it will be detrimental to the mass resolution since fragmentation at this time produces ions that arrive at the detector predicted by Cotter[23] to "contribute to the general baseline noise" between the precursor and product ion peaks. These ions contribute to the baseline noise in the spectrum and

tailing of molecular ion peak. This illustrates that for high mass ions, metastable fragmentation in the flight tube can be very extensive. This fragmentation will affect the mass spectra particularly in terms of achievable mass resolution.

6.4.4.3 Effect of internal energy release on peak width

In the previous section the effect of internal energy has been described. The greater the internal energy the faster the rate of ion fragmentation will be. Another important factor is the amount of internal energy partitioned to kinetic energy of the fragments F^+ and N which will lead to translational (kinetic) energy spread of the products in the laboratory frame of reference. In the book entitled 'Metastable Ions' by Cooks *et al.*[24] this is described in detail. The velocity spread of the products F^+ and N are given in below, a novel aspect of this calculation of the velocity spread of the neutral species, which is usually neglected but is a process which contributes to a linear-TOF mass spectrum.

The starting point is that the precursor ion is accelerated in the ion source through a potential V and will have a kinetic energy eV,

$$\frac{1}{2}mv^2 = eV \qquad \text{(Eq. 6.12)}$$

where m is the mass of ion in kg (1 Da =1.66×10^{-27} kg), v is the velocity in m s^{-1}, e is charge on the electron in Coulombs (1.602×10^{-19} C).

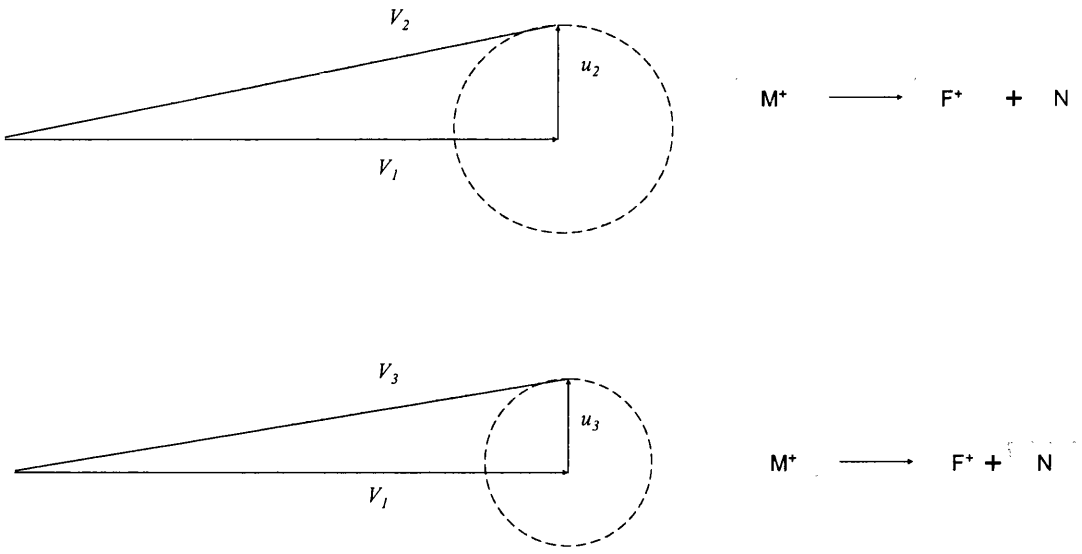


Figure 6.22 Velocity diagram describing the ion fragmentation process ($M^+ \rightarrow F^+ + N$)

Figure 6.22 shows the vector diagram describing ion fragmentation, in this figure

v_1 = velocity of the molecular ion (precursor)

v_2 = velocity of the product ion (F^+)

v_3= velocity of the product neutral (N)

u_2 = velocity of the product ion (F^+) in the centre of mass frame

u_3 = velocity of the product neutral (N) in the centre of mass frame

The maximum and minimum velocity spread for F^+ is given by[24]

$$v_{2x}(\max) = v_1 + \left(\frac{2m_3T}{m_1m_2}\right)^{1/2} \qquad \text{(Eq. 6.13)}$$

$$v_{2x}(\min) = v_1 - (\frac{2m_3T}{m_1m_2})^{1/2} \qquad \text{(Eq. 6.14)}$$

where T is the internal energy of M^+ that is converted into translational energy of the products in the fragmentation and

$$T + \frac{1}{2}m_1v_1^2 = \frac{1}{2}m_2(v_1+u_2)^2 + \frac{1}{2}m_3(v_1+u_3)^2 \qquad \text{(Eq. 6.15)}$$

The velocity of the neutral species is not shown by Cooks *et al*[24], and is derived here,

$$m_2u_2 + m_3u_3 = 0 \qquad \text{(Eq. 6.16)}$$

$$u_2 = \frac{-m_3u_3}{m_2} \qquad \text{(Eq. 6.17)}$$

$$T = \frac{m_1m_2}{2m_3}u_2^2 = \frac{m_1m_2}{2m_3}\left(\frac{-m_3u_3}{m_2}\right)^2 = \frac{m_1m_3}{2m_2}u_3^2 \qquad \text{(Eq. 6.18)}$$

$$v_{3x}(\max) = v_1 + (\frac{2m_2T}{m_1m_3})^{1/2} \qquad \text{(Eq. 6.19)}$$

$$v_{3x}(\min) = v_1 - (\frac{2m_2T}{m_1m_3})^{1/2} \qquad \text{(Eq. 6.20)}$$

From the above equations we can estimate the velocity spread of the F^+ (v_{2x}(max) to v_{2x}(min)) and N (v_{3x}(max) to v_{3x}(min)). In Table 6.6, the maximum velocity spreads for

BSA are calculated for energy releases (T) of 0.1, 1 and 5 eV. The internal energy releases above are only estimates based on previous measurements[24] based on lower mass species, this may be a deficiency when modeling high mass species, such as BSA, that can be corrected when the internal energy content of very high mass ions created by MALDI is known.

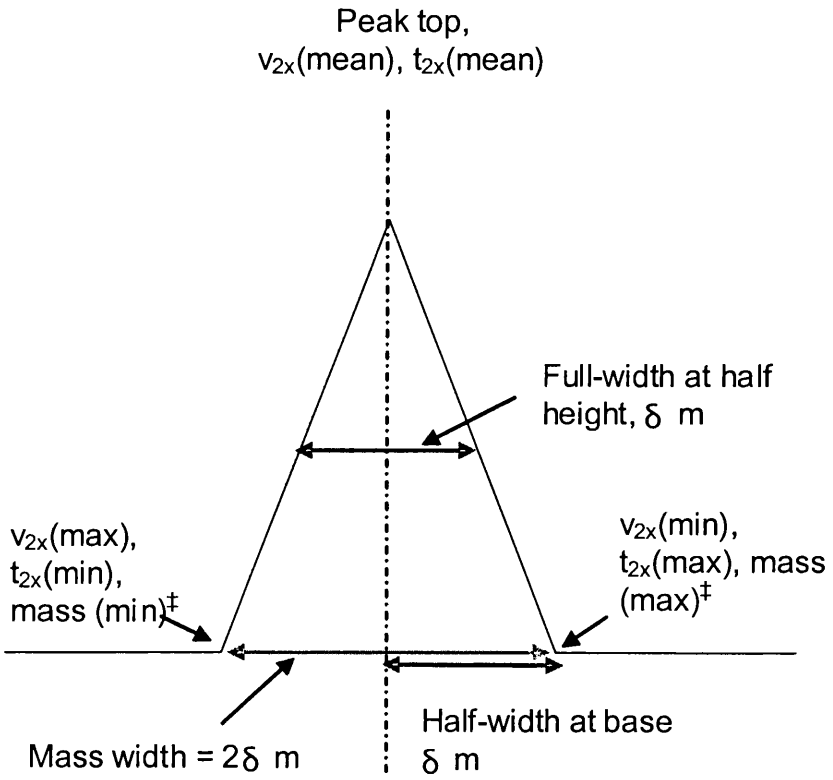


Figure 6.23 Nomenclature used to denote the broadening of a peak due to kinetic energy release, in terms of velocity, time and mass spread. ‡ refer to Table 6.6

Table 6.6 Calculated values of the maximum and minimum velocity, time-of-flight and mass spreads (refer to Figure 6.23 for an explanation of terms) for BSA fragment ion and neutral species F^+, N as a function of internal energy release to translational energy (T/eV) of the fragments. It is assumed in this model calculation, to illustrate the effect, that the fragmentation is $66431^+ \rightarrow 33215^+ + 33216$

Velocity and time-of-flight of the products		**Internal energy release** (*T*/eV)		
		0.1	***1***	***5***
Fragment ion, F^+	v_{2x}(mean)(m s^{-1})	8520	8520	8520
	v_{2x}(max) (m s^{-1})	8537	8574	8641
	v_{2x}(min) (m s^{-1})	8503	8466	8399
	$t_{2x(mean)}$ (sec)	2.347×10^{-4}	2.347×10^{-4}	2.347×10^{-4}
	$t_{2x(max)}$ (sec)	2.352×10^{-4}	2.362×10^{-4}	2.381×10^{-4}
	$t_{2x(min)}$ (sec)	2.343×10^{-4}	2.333×10^{-4}	2.315×10^{-4}
	$Mass_{(max)}$), δm, (u)	265‡	832‡	1840‡
	$Mass_{(min)}$), δm, (u)	-267‡	-848‡	-1920‡
	Mass width, 2δm, (u)	531†	1681†	3759†
Fragment neutral, N	v_{3x}(mean) (m s^{-1})	8520	8520	8520
	v_{3x}(max) (m s^{-1})	8537	8574	8641
	v_{3x}(min) (m s^{-1})	8503	8466	8399
	t_{3x}(mean) (sec)	2.347×10^{-4}	2.347×10^{-4}	2.347×10^{-4}
	t_{3x}(max) (sec)	2.352×10^{-4}	2.362×10^{-4}	2.381×10^{-4}
	t_{3x}(min) (sec)	2.343×10^{-4}	2.333×10^{-4}	2.315×10^{-4}
	$Mass_{(max)}$, δm, (u)	265‡	832‡	1840‡
	$Mass_{(min)}$), δm, (u)	-267‡	-848‡	-1920‡
	Mass width, 2δm, (u)	531†	1681†	3759†

‡ Mass difference from mean mass of 66431 u

† The mass difference is calculated = $Mass_{(max)} - Mass_{(min)} = 2\delta m$

Table 6.6 gives the velocity, time and mass spreads for the hypothetical fragmentation of BSA to one half its mass. In this case both the fragment ion and neutral have the same numerical values for their velocities (and velocity spread) and this is a result of choosing that the fragmentation leads to equal masses for the products. If the masses of F^+ and N were chosen unequal then the velocity spreads would be different, leading to different contributions to the velocity spread only. Note the mean velocities of the products, $v_{2x(\text{mean})}$ and $v_{3x(\text{mean})}$ are not different from their precursor, no matter what the fragment and neutral masses are. For energy releases of 0.1, 1 and 5 eV the time spreads around the mean time of 234.69 μs is ± 0.469, 1.485 and 3.32 μs, respectively. These correspond to mass spreads of ±265, 832 and 1840 u, respectively. The measured mass spread of the experimental spectrum is about 3200 u at the base (assuming a symmetrical peak and ignoring the high mass tail) and 850 u at half height. The above calculated values refer to the maximum velocity (and time spread) and therefore relates to the measurement taken at the base of the experimentally observed peak.

Thus if this process is due to kinetic energy release upon fragmentation then the above data suggest the kinetic energy release is in excess of 5 eV. This is a very large kinetic energy release, especially when values published in the literature[42, 43] are taken into account, as the largest energy releases are reported for ions, such as, doubly-charged ions where energy releases are of the order a few eV and largely arise from Coulombic repulsion. However, BSA is a very much larger species with greater ability to store internal energy in its (3n-3) degrees of freedom. Therefore it is proposed that a possible

mechanism for peak broadening is due to fragmentation of the ions as they fly through the analyser. The energy release needs to be of the order 1 to10 eV, which taking into account the high masses may not be an unreasonable amount of energy to store. The model for kinetic energy release may need some rethinking when comparing small molecule fragmentation considered by Physical Chemists in the 1970's and the situation for very large molecules.

The asymmetric shape of the peak is thought to be due to ion fragmentation in the acceleration region as the exponential decay plots shown in Figures 6.19 and 6.20 show there to be a considerable degree of fragmentation occurs early on, especially when the ion lifetime is less than 10^{-5}sec, i.e., fast metastable decay. Earlier work on ion fragmentation show this lifetime is quite probable and therefore we can expect this effect to be large and seems to agree with current observations. This study is only preliminary in nature but has uncovered a hitherto unreported effect in high mass MALDI- TOF-MS mass spectra.

6.5 Conclusion

In this study, the mass accuracy of MALDI-TOF of high mass proteins was investigated as a result of preliminary studies obtained for the cyclic nucleotide research work, reported in Chapter 7. For data processing to obtain peak tops (or centroids) for mass calibration Gaussian smoothing, using 49-point window, was found to be good for

proteins such as BSA when compared to the alternative noise filter method, which did not give good results when applied. However, the requirement for a large number of mass spectra to be averaged to obtain good mass accuracy is noted as a draw back of the method developed here. In a study to find alternatives two techniques involving curving fitting and differentiation of calibration peak shapes was investigated using readily available software, such as MS-Excel and a powerful scientific programme EasyPlot. These methods are proposed as alternatives to those available within the existing software of the Voyager MALDI-TOF-MS system. These data processing techniques could be applied by mass spectrometry manufacturers.

The phenomenon of peak broadening observed for high mass protein peaks was investigated and a theoretical study showed that these phenomena can be explained on the basis of the ion undergoing fragmentation whilst in flight and the product ion(s) and neutral(s) will have the same time-of-arrival and thus the same effective mass. However, the fragmentation process leads to broader peaks due to kinetic energy release. Model calculations show that ions as massive as BSA only need 0.1 to a few eV of internal energy to be released into kinetic energy and translational motion of the product to account for such broadening effects. The mass asymmetry on the higher mass side of peaks may be explained by in-source fragmentation process. Further work is needed to build up a more complete picture to this preliminary study.

6.6 References

1. Webb, K., Bristow, T., Sargent, M. and Stein, B., *Methodology for Accurate Mass Measurement of Small Molecules, Best Practice Guide*, LGC limited, (2004)
2. Budzikiewicz, H., Djerassi, C. and Williams, D.H., *Mass Spectrometry of Organic compounds*, Holden-Day, San Francisco (1967)
3. de Hoffmann, E. and Stroobant, V., *Mass Spectrometry Principles and Applications* 2nd John Wiley and Sons Ltd, (2001)
4. Yergey, J.A., Int. J. Mass Spectrom. Ion Processes, **52**, 337 (1982)
5. Rockwood, A.L., VanOrden, S.L. and Smith, R.D., Anal. Chem., **67**, 2699 (1995)
6. Rockwood, A.L. and VanOrden, S.L., Anal. Chem., **68**, 2027 (1996)
7. McLafferty, F.W., *Interpretation of Mass Spectra*, 3rd edition; University Scientific Books: Mill Valley, CA (1980)
8. Bristow, A.W.T. and Webb, K.S., J. Am. Soc. Mass Spectrom., **14**, 1086 (2003)
9. Marshall, A.G.., Int. J. of Mass Spectrom., **200**, 3231 (2000)
10. Chapman, J.R., *Protein and peptide analysis by mass spectrometry*, Humana Press, Totowa NJ (1996)
11. Hop, C.E.C.A., J. Mass Spectrom., **33**, 397 (1998)
12. Steen, H. and Mann, M., Nat. Rev., **5**, 699 (2004)
13. Clauser, K.R., Baker, P. and Burlingame, A.L., Anal. Chem., **71**, 2871 (1999)
14. de Jong, A., Mass Spectrom. Rev., **17**, 311 (1998)
15. Jardine, D.R., Mortgan, J., Alderdice, D.S. and Derrick, P.J., Org. Mass Spectrom., **27**,

1077 (1992)

16. Zhou, J., Ens, W., Standing, K.G. and Verentchikov, A., Rapid Commun. Mass Spectrom., 6, 671 (1992)
17. Mamyrin, B.A., Int. J. Mass Spectrom. Ion Processes, **131**, 1 (1994)
18. Colby, S.M., King, T.B. and Reilly, J.P., Rapid Commun. Mass Spectrom., **8**, 865 (1994)
19. Brown, R.S. and Lennon, J., Anal. Chem., **67**, 1998 (1995)
20. Juhasz, P. and Vestal, M.L., J. Am. Soc. Mass Spectrom., **7**, 995 (1998)
21. Chait, B.T., Int. J. Mass Spectrom. Ion Phys., **53**, 227 (1983)
22. Demirev, P., Olthoff, J.K., Fenselau, C. and Cotter R.J., Anal. Chem., **59**, 1951 (1987)
23. Cotter, R.J., *Time-of-Flight Mass Spectrometry*, American Chemical Society, Washington DC (1994)
24. Cooks, R.G., Beynon, J.H., Caprioli, R.M. and Lester, G.R., *Metastable Ions*, Elsevier Press, (1973)
25. Newton, R.P., Brenton, A.G., Smith, C.J. and Dudley, E., Phytochemistry, **65**, 1449 (2004)
26. Kovtoun, S.V., English, R.D. and Cotter, R.J., J. Am. Soc. Mass Spectrom., **13**, 135 (2002)
27. Users' manual Voyager DE STR mass spectrometer, Applied Biosystems, USA
28. Moskovets, E. and Karger, B.L., Rapid Commun. Mass Spectrom., **17**, 229 (2003)
29. Russell, D.H. and Edmondson, R.D., J. Mass Spectrom., **32**, 263 (1997)

30. Vestal, M. and Juhasz, P., J. Am. Soc. Mass Spectrom., **9**, 892 (1998)

31. Vorst, H.J., van Tilborg, M.W.E.M, van Veelen, P.A., Tjaden, U.R. and van der Greef, J., Rapid Commun. Mass Spectrom., **4**, 202 (1990)

32. Barber, M., Bordoli, R.S., Sedgwick, R.D. and Tyler, A.N., Biomed. Mass Spectrom., **8**, 337 (1981)

33. Poulter, L. and Taylor, L.C.E., Int. J. Mass Spectrom. Ion Proc., **91**, 183 (1989)

34. Feng, R., Konishi, Y. and Bell, A.W., J. Am. Soc. Mass Spectrom., **2**, 387 (1991)

35. Yergey, J., Heller, D., Hansen, G., Cotter, R.J. and Fenselau, C., Anal. Chem., **55**, 353 (1983)

36. Savitzky, A. and Golay M.J.E., J. Anal. Chem., **36**, 1627 (1984)

37. Marsh, P.L., Am. Statistician, **48**, 44 (1994)

38. Cotter, R.J., *Time-of-flight mass spectrometry: instrumentation and applications in biological research*, American Chemical Society, Washington DC (1997)

39. Wiley, W. C. and McLaren, I. H., Rev. Sci. Instrum., **26**, 1150 (1955)

40. Karas, M., Bahr, U., Fournier, I., Gluckmann, M. and Pfenninger, A., Int. J. Mass Spectrom., **226**, 239 (2003)

41. Rosenstock, H.M., Wallenstein, M.B., Wahrhaftig, A.L. and Eyring, H., Proc. Natl. Acad. Sci. USA, **38**, 667 (1952)

42. Brenton, A.G., Morgan, R.P. and Beynon, J.H., Ann. Rev. Phys. Chem., **30**, 51 (1979)

43. Beynon, J.H., Brenton, A.G. and Harris, F.M., Int J. Mass Spectrom. Ion Phys., **45**, 5 (1982)

Chapter 7

Mass spectrometric analysis of changes in the murine brain phosphoproteome elicited by cCMP

7.1 Introduction

7.1.1 Introduction to cyclic nucleotide biochemistry

Cyclic nucleotides have attracted the attention of scientists since the discovery of adenosine 3',5'-cyclic monophosphate (cAMP) by Rall and Sutherland in 1958[1]. The major significance of this pioneering work was that it provided the impetus for the exploration of cell regulation and signal transduction mechanisms. Signal transduction is any process by which a cell converts one kind of signal into another and is a vital aspect of inter-cell communication. The process often involves a sequence of biochemical reactions inside the cell, which are carried out by enzymes and linked through second messengers[2].

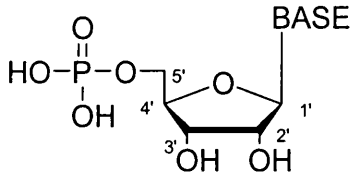


Figure 7.1 Structure of nucleotide

7.1.1.1 Nucleotides and cyclic nucleotides

Nucleotides (Figure 7.1) are molecules containing a heterocyclic nucleobase, a sugar ring and a phosphate moiety. The nitrogenous nucleobase includes purines such as adenine and guanine, and pyrimidines, such as cytosine, uridine and thymine (Figure 7.2). The sugar unit can be either D-ribose or 2'-deoxy-D-ribose. Nucleotides are the structural units of nucleic acids which contain genetic instructions and are involved in

various processes such as protein synthesis.

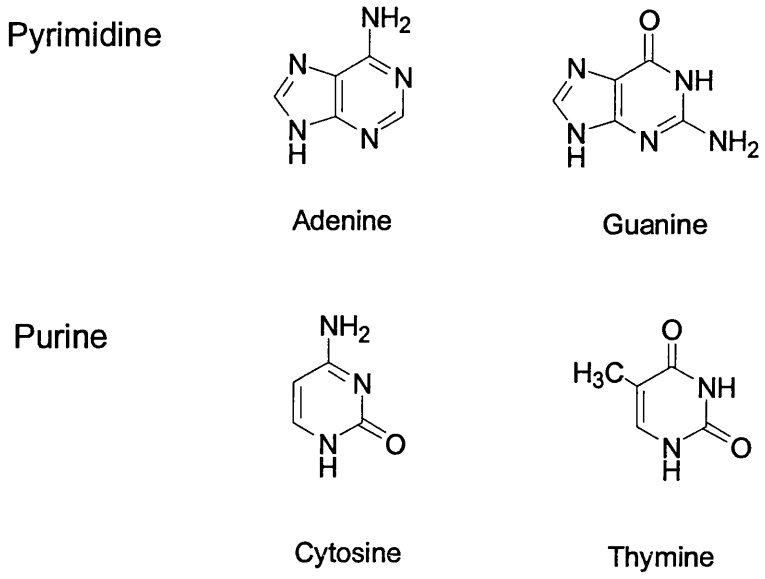


Figure 7.2 Structures of some naturally occurring purine and pyrimidine bases

A cyclic nucleotide is a nucleotide in which the phosphate group is bonded to two of the sugar's hydroxyl groups, either between 3', 5'-position or the 2', 3'- position, forming a cyclic structure (Figure 7.3). X-ray crystallography and proton-NMR have been used to examine the rigid conformations of cyclic nucleotides[3, 4], particularly the 3',5'-cyclic phosphodiester. This rigid conformation is the possible reason that chemical synthesis of cyclic nucleotides requires fairly drastic conditions and their biosyntheses are extremely endothermic[5]. Compared to their phosphomonoester counterparts, cyclic nucleotides are more resistant to weak acid and alkaline hydrolysis. The cyclic nucleotides are less polar than the analogous non-cyclic forms thus they are consequently less soluble in aqueous solutions[6]. The 2',3'-cyclic nucleotides have been proved to be mainly breakdown products of nucleic acid[7], 3',5'-cyclic nucleotides, on the other hand, have been found to play important roles in metabolic regulation in bacterial, plant and tissue

cells[8, 9].

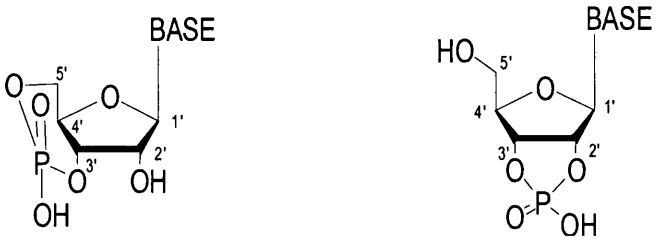


Figure 7.3 Structure of 3', 5'-cyclic nucleotide and 2', 3'-cyclic nucleotide

7.1.1.2 Cyclic nucleotides as second messengers

The concept of second messengers was first reported by Sutherland and Rall[1] in 1958 when they were investigating adrenalin, an important hormone which travels through the blood as a signal of an imminent danger. They studied the effects of both adrenalin and glucagon upon glycogen phosphorylase activity in dog liver and found that elevated levels of cAMP, brought about by glucagon and adrenalin, stimulated hepatic glycogenolysis, resulting in an increased level of glucose in the blood stream[10].

According to Sutherland's second messenger theory, adrenalin is a primary messenger secreted from the adrenal glands into the bloodstream that circulates until it binds to a specific receptor in the outer side of a membrane of a target cell. The binding activates adenylyl cyclase (AC), which converts adenosine 5'-triphosphate (ATP) to cAMP on the inner side of the membrane, through interaction with a G-protein. This triggers the production of the second messenger, cAMP, which stimulates the activity of cAMP-dependent protein kinase which in turn phosphorylates a number of physiologically active proteins, thereby altering their activity (Figure 7.4)[11]. The increase in concentration of the intracellular levels of cAMP also increases the activity

of cAMP phosphodiesterase which hydrolyzes the second messenger to its precursor AMP[12].

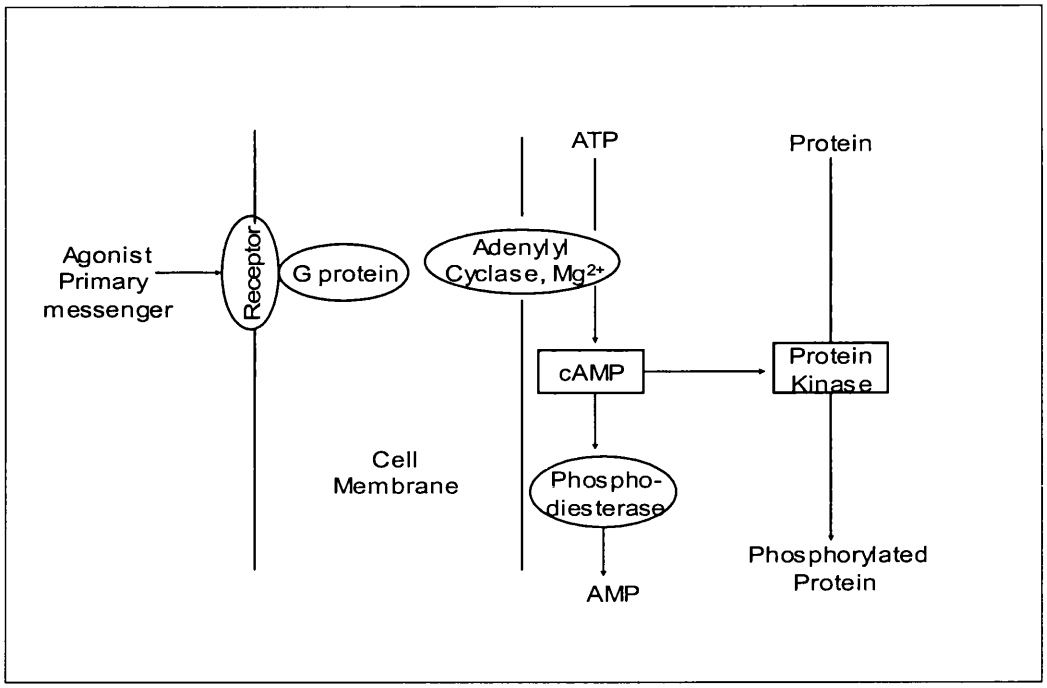


Figure 7.4 cAMP second messenger system

A consequence of this process and a major reason for studying cyclic nucleotide systems has been that they are very valuable drug targets. These drugs alter the activities of proteins by increasing or reducing the cAMP level or adenylyl cyclase or phosphodiesterase activity in the body. The simplest such agent is caffeine, present in coffee and tea, which inhibits phosphodiesterase activities; as a result, it inhibits the breakdown of cyclic AMP and thus mimics the effects of adrenaline[13].

7.1.1.3 Adenosine 3'5'-cyclic monophosphate

Adenosine 3'5'-cyclic monophosphate or cAMP (Figure 7.5) was the first cyclic nucleotide identified by Sutherland and Rall[1] and its discovery opened new horizons in biological research. The presence of cAMP has now been firmly established in higher plants[14], bacteria[15,16], and in nearly all animal organs and tissues[17]. cAMP is considered to be a vital component in all nucleated mammalian cells and proven to be involved as a regulatory agent in various short term, sudden effect hormones by recognizing and amplifying the signal inside the cell[18]. Studies have also shown that an abnormal level of cAMP relates to a variety of disorders of hormone function, such as diabetes mellitus[19], cystic fibrosis[20], allergic responses[21], asthma[22], hyperparathyroidism[23] and cancer[24, 25]. There are various cAMP related enzymes involved in signal transduction pathway with cAMP acting as second messenger.

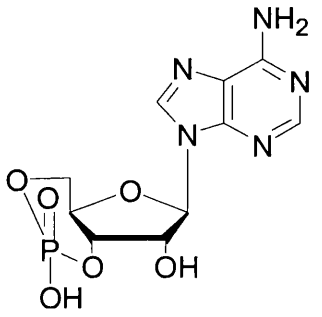


Figure 7.5 Structure of Adenosine 3',5'-cyclic monophosphate

7.1.1.3.1 Adenylyl cyclase

Adenylyl cyclase is responsible for the synthesis of cAMP from ATP. cAMP is released into the cytoplasm after G_s-protein binds to its receptor on the catalytic domains of adenylyl cyclase. After that, the cAMP binds to protein kinase to expose this enzyme's

active site, which continues the communication pathway[26]. Adenylyl cyclase plays a dual role in the regulation of cAMP. It has many stimulatory and inhibitory receptors that bind G proteins and moderate its catalytic subunit activity. When the G proteins bind to the receptors, the binding sites of ATP are exposed and the enzyme adenylyl cyclase is then able to catalyze the transformation of ATP to cAMP. Since there are both inhibitory and stimulatory receptor sites for the G-proteins, the binding of these proteins to adenylyl cyclases can increase the production of cAMP, but can also decrease the production of cAMP.

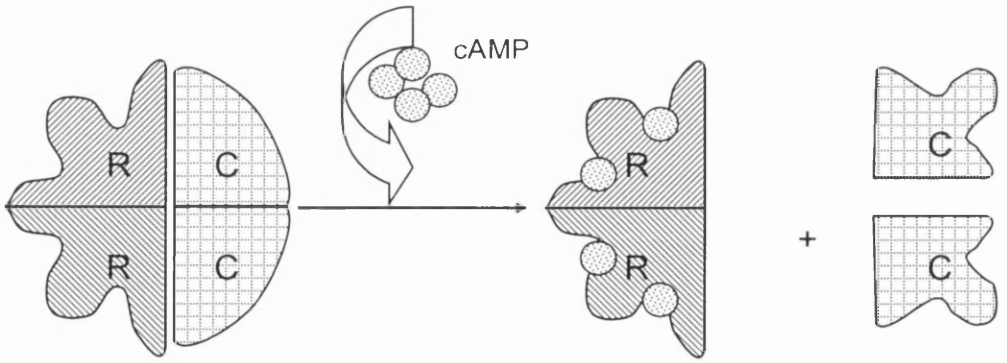
7.1.1.3.2 cAMP phosphodiesterase

Phosphodiesterase is the enzyme that catalyzes the hydrolysis of phosphodiester bonds of cyclic nucleotides and so acts as the 'off-switch' in the signal transduction pathway. Initially three isoforms were identified; CaM-PDE, cAMP-PDE and cGMP-PDE[27, 28]. The number of PDE isoforms identified increased with the years and in 1995, the nomenclature for the PDE family was standardized[29]. Today there are more than 11 families of phosphodiesterases, which differ significantly in distribution, specificity and kinetics[30]. Inhibitors of phosphodiesterases can prolong or enhance the effects of physiological processes mediated by these cyclic nucleotides so they too are pharmacological targets for selective inhibitors.

7.1.1.3.3 cAMP-dependent protein kinase

cAMP-dependent protein kinase, also known as protein kinase A (PKA), refers to a family of enzymes, which, in response to cAMP in the cell, initiate changes in the activity of extra-nuclear proteins by phosphorylating serine (S) or threonine (T)

residues[31]. Protein phosphorylation, first identified as a mechanism for regulating protein activity in the 1950s[32], is probably the most important mechanism for regulation in mammalian cells. Protein kinases are major intracellular targets of cyclic nucleotides signal transduction pathways and are known to regulate the majority of cellular pathways; up to 30% of all proteins may be modified by kinase activity[33]. Each cAMP-dependent protein kinase is a holoenzyme that consists of two regulatory and two catalytic subunits. Under low levels of cAMP, the regulatory subunits block the catalytic centre of the catalytic subunits, the holoenzyme remains intact and is catalytically inactive. When the concentration of cAMP rises, as a result of activation of adenylyl cyclase or inhibition of phosphodiesterases, cAMP binds to the two binding sites on the regulatory subunits, which then undergo a conformational change that releases the catalytic subunits (Figure 7.6)[34]. The free catalytic subunits can then catalyze the transfer of ATP terminal phosphates to protein substrates which usually results in a change in activity of the substrates, as a result of increased surface charge.



$$R_2C_2 + 4cAMP \rightleftharpoons R_2C_2(cAMP)_4 \rightleftharpoons 2(cAMP)_2R + 2C$$

Figure 7.6 Mechanism of protein kinase A activity

cAMP-dependent protein kinase has several functions in the cell, including phosphorylation of proteins and regulation of glycogen, sugar, and lipid metabolism. Proteins phosphorylated by cAMP-dependent protein kinase include glycogen phosphorylase kinase, protamine, glycogen synthetase, hormone sensitive lipase, casein, troponin, fructose-1,6-diphosphatase, pyruvate kinase and histone[35]. Despite its broad substrate specificity, PKA activity can be highly selective in a physiological setting[36]. Specific hormones, each capable of increasing intracellular cAMP levels, can result in the preferential phosphorylation of different target substrates[37]. The effects of cAMP-dependent protein kinase phosphorylation are generally short term and transient because protein phosphatases quickly dephosphorylate cAMP-dependent protein kinase substrates.

Cyclic AMP can also exert longer term effects. cAMP-dependent protein kinase is recognized as a means of regulation of gene expression since a protein called cAMP-response element binding protein (CREB) was found to be a target of PKA. On phosphorylation of the CREB protein dimerization occurs and in this state is able to interact with DNA at the cAMP response element (CRE), an 8-base pair palindrome, allowing initiation of transcription in the presence of two other proteins, CPB (CREB binding protein) and p300, both large proteins which only interact with CREB as its phosphorylated dimer[5]. Cyclic AMP effects independent of protein kinase include action on ion channels[38] and / or EPAC (Exchange protein directly activated by cyclic AMP)[39].

7.1.1.4 Guanosine 3',5'-cyclic monophosphate

Guanosine 3',5'-cyclic monophosphate is a cyclic nucleotide derived from guanosine triphosphate (GTP). It was first isolated from rat urine in 1963[40] followed by identification in various mammalian tissues[41], plant[42] and bacterial cells[43]. Although the intracellular concentration of GMP is typically 10 fold less than cAMP[44], cGMP acts as a second messenger much like cAMP, most notably by activating intracellular protein kinases in response to the binding of membrane-impermeable peptide hormones to the external cell surface[45] and regulating the movement of sodium ions and water across membranes although it has a more restricted function than cAMP[46]. Enzymes for cyclic GMP metabolisms analogous to those for cyclic AMP are present.

7.1.1.4.1 Guanylyl cyclase

The synthesis of cGMP is catalyzed by guanylyl cyclase (GC), the enzyme which converts GTP to cGMP. There are two main forms of guanylyl cyclase, membrane-bound GC and soluble GC[47]. Membrane-bound GC is activated by hormones such as the natriuretic factors[48], while nitric oxide and nitrosamines typically stimulate cGMP synthesis in soluble GC[49].

7.1.1.4.2 cGMP phosphodiesterase

cGMP phosphodiesterases hydrolyze cGMP and three classes are found in mammals; cGMP-stimulated phosphodiesterase, cGMP-specific phosphodiesterase and cGMP-inhibited phosphodiesterase. Among these, cGMP-stimulated phosphodiesterase and cGMP-inhibited phosphodiesterase also involve in the regulation of the action of

cAMP for example in adipose tissue lipolysis[50] and platelet aggregation[51].

7.1.1.4.3 cGMP-dependent protein kinases (PKG)

cGMP-dependent protein kinases catalyze the phosphorylation of several endogenous substrates in different tissues e.g. smooth muscle and cerebellar purkunje cells[52,53]. cGMP-dependent protein kinases are present in 5-10 fold less amounts compared to cAMP-dependent protein kinases in many tissues e.g. cerebellum, adrenal cortex, lung and heart[54, 55, 56]. Unlike PKA, PKG is a dimer consisting of one catalytic and one regulatory unit, with the regulatory unit blocking the active sites of the catalytic unit. Again cGMP binds to sites on the regulatory unit of PKG and activates the catalytic units, enabling them to phosphorylate their substrates however catalytic and regulatory units do not disassociate when cGMP-dependent protein kinases activated[57].

7.1.1.5 Cytidine 3',5'-cyclic monophosphate

After the demonstration of the natural occurrence of cAMP and cGMP and the identification of their roles as second messengers in a variety of cellular processes further investigations were carried out to determine if there are any other naturally occurring cyclic nucleotides and if so, their possible cellular functions. Evidence of the natural occurrence of cytidine 3',5'-cyclic monophosphate (cyclic CMP, see Figure 7.7) was initially found by Bloch and his colleagues in liver extracts and leukaemia L1210 cultures in 1974[58]. Wikberg and Wingren debated this issue by claiming that endogenous cyclic CMP-immunoreactive material could be separated chromatographically from authentic cCMP[59,60]. However, radioimmunoassay[61, 62] and

enzyme immunoassay[63, 64] methods provided new data that supported the claims of Bloch. Several factors were reported to stimulate increase in intracellular cyclic CMP concentrations, such as luteinizing hormone releasing hormone[65], long acting thyroid stimulator[66] and elevated cell proliferation rate[67]. Newton and his coworkers have now unequivocally demonstrated the natural occurrence of cyclic CMP by means of tandem spectrometric analysis of mammalian tissue extracts, higher and lower plants, as well as bacteria such as *E. coli* [68]. Radioimmunoassay showed that cyclic CMP concentrations were of similar levels in a variety of tissues, but the cCMP concentration is elevated in rapidly proliferating cells which suggests that cyclic CMP may have value as a clinical marker, as indicated by levels in leukaemic patients[69]. Further research confirmed the existence of enzymes capable of cyclic CMP synthesis (cytidylyl cyclase) and hydrolysis (cCMP phosphodiesterase).

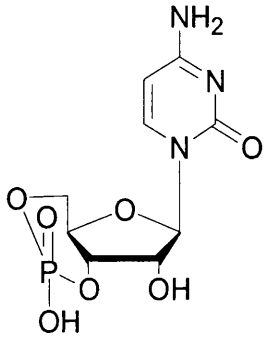


Figure 7.7 Structure of cytidine 3',5'-cyclic monophosphate

7.1.1.5.1 Cytidylyl cyclase

Cytidylyl cyclase is the enzyme that converts CTP to cCMP (see Figure 7.8) and was initially identified in mouse myeloid leukaemic tumours and normal mouse liver and

spleen[70, 71]. Although disputed by Gaion and Krishna[72], Newton proved that cyclic CMP was one of the products of the putative cytidylyl cyclase activity with the CTP substrate in 1988[73]. The controversy regarding the identities of both the cyclic nucleotide in tissue extract and the putative cyclic CMP product of the cytidylyl cyclase reaction was explained by the discovery of four novel cyclic CMP analogues, cytidine 3',5'-cyclic pyrophosphate, cytidine 2'-O-monophosphate-3',5'-cyclic monophosphate, cytidine 2'-O-glutamyl 3',5'-cyclic monophosphate and cytidine 2'-O-aspartyl-3',5' cyclic monophosphate (see Figure 7.9). The chromatographic systems for resolving cCMP were later used for the specific assay for cytidylyl cyclase[74] and a radioimmunoassay for extracted cCMP[75]. Cytidylyl cyclase has subsequently been established as existing in many rat and mammalian tissues and it was found that cytidylyl cyclase could be stimulated by Fe^{2+} and Mn^{2+}[76].

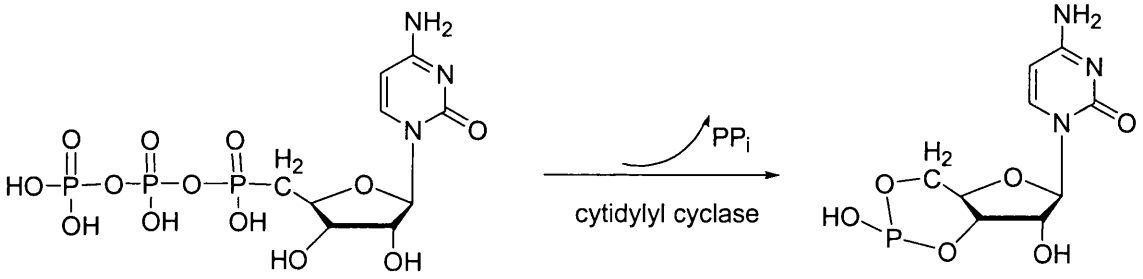
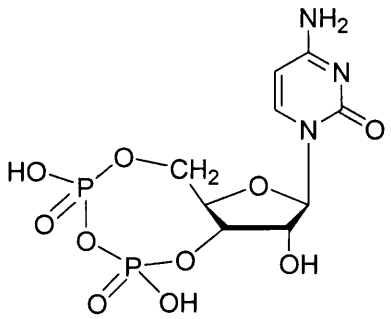
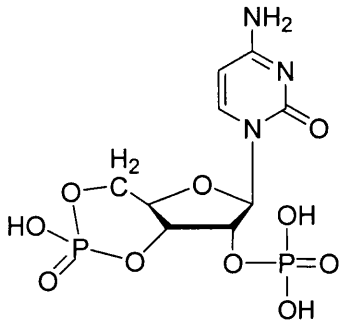


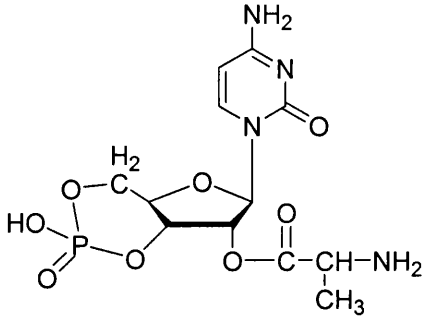
Figure 7.8 Reaction catalysed by cytidylyl cyclase



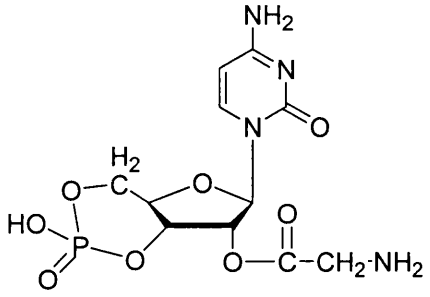
cytidine 3',5'-cyclic pyrophosphate



cytidine 2'-O-monophosphate
-3',5'-cyclic monophosphate



cytidine 2'-O-glutamyl
3',5'-cyclic monophosphate



cytidine 2'-O-aspartyl
-3',5' cyclic monophosphate

Figure 7.9 Analogues of cCMP also produced by cytidylyl cyclase

Application of the cytidylate cyclase activity assays has shown that cytidylyl cyclase is unaffected by effectors of adenylyl cyclase such as glucagon, cathcholamines, and muscarinic agents, or effectors of guanylyl cyclase such as ANP and nitroso-compounds, but it is activated by diethylstilboestrol, testosterone, calmodulin and to a lesser extent progesterone[17, 76].

7.1.1.5.2 cCMP phosphodiesterase

cCMP phosphodiesterases are enzymes that degrade cCMP to 5'-CMP (see Figure 7.10).

They can be classified into two groups, cCMP-specific phosphodiesterase and multifunctional phosphodiesterase. The former is substrate specific[68, 77] while the latter is capable of hydrolyzing both the 2',3'- and 3',5'- isomers of cyclic nucleotides and has activity with both purines and pyrimidines[78-80]. cCMP-specific phosphodiesterase was identified by Newton *et al* in a variety of rat organ tissues[81]. There are similarities as well as differences between the two classes of phosphodiesterase[82]. Both enzymes show similar optimal pH values, relative insensitivity to typical cAMP/cGMP phosphodiesterase inhibitors and require the involvement of cations as cofactors, whereas their differences lie on their specificity and effect of calmodulin on their activity[83]. cCMP-specific PDE is the only enzyme demonstrated to be inhibited by calmodulin.

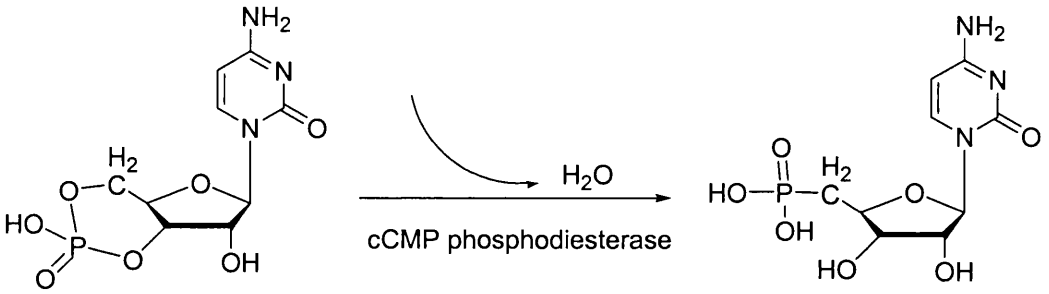


Figure 7.10 Reaction catalysed by cCMP phosphodiesterase

7.1.1.5.3 cCMP-responsive protein kinase

With evidence of the sensitivity of cytidylyl cyclase and cyclic CMP phosphodiesterase to hormone agonists, it is a reasonable hypothesis that a cyclic CMP-binding protein or cyclic CMP protein kinase must exist if cyclic CMP has a second messenger function. Proteins capable of binding cCMP were first reported in 1985[84], cCMP-responsive

protein kinase activity, with activation due to cCMP only, was demonstrated in rat tissues several years later. Newton *et al* have demonstrated several binding proteins, two of which possess protein kinase activity selectively sensitive to cyclic CMP[85], studies with [γ-^{32} P]-ATP incorporation have indicated at least 17 endogenous protein substrates of this kinase, but they have yet to be characterized[86]. In normal levels cCMP has no significant role in cell proliferation; Bond, by using IMAC ZipTip, identified Rab23 as a possible phosphorylation substrate of cCMP-responsive protein kinase and indicates this phosporylation could potentially lead to cell hyperproliferation and formation of a brain tumor[83]. The available data suggest that cCMP may be a third secondary messenger involved in cell proliferation growth and differentiation, but deduction of the function of cyclic CMP as a third cyclic nucleotide metabolic regulator will only be credible after the cyclic CMP binding proteins and the substrates of cyclic CMP-sensitive protein kinases have been characterized and demonstrated to be integral components of systems responsive to exogenously applied cyclic CMP, and to be shown sensitive to factors modifying cytidylyl cyclase and cyclic CMP phosphodiesterase activities[86].

7.1.1.5.4 Preliminary studies of biological effects of cCMP

Studies from radioimmunoassay have shown that levels of cCMP are elevated in rapidly dividing tissues and that there is an inverse relationship between brain cCMP levels and aggression in male mice, which may be linked to testosterone levels[87]. Many varied biological effects of cCMP have been reported by studying effects of dibutyryl cyclic CMP upon rat and mouse tissues including the stimulation of DNA synthesis[88], inhibition of RNA synthesis[89], an increase in concentration of free amino acids and lipid concentration, and decrease in total protein concentration[90]. Further research showed

that there is an increase in the contents of mRNA, rRNA and tRNA in brain cells 30 to 60 minutes following the separate injections of cCMP, cUMP, cytidine and uridine, showing a more specific effect of cytidine and cCMP on neurons and glial cells[62].

cCMP levels have been shown to be elevated in developing tissues such as kidney foetal tissue compared to adult tissue[91], corresponding to the lower levels in phosphodiesterase and higher levels of cytidylyl cyclase activity observed in tissues undergoing rapid regeneration and developing tissues[67]. Increased levels of cCMP have also been shown in cell hyperproliferation[63]. cCMP was found to be present at concentration 100 fold greater than normal cells in leukaemia L-1210 cells. In accordance with this, cytidylyl cyclase activities are increased in foetal and other rapidly dividing tissues, and decreased in 'older' tissues[17].

7.1.1.6 Other naturally occurring cyclic nucleotides

Inosine 3',5'-cyclic monophosphate (cIMP), uridine 3',5'-cyclic monophosphate (cUMP), and 2'-deoxythymidine 3',5'-cyclic monophosphate (cdTMP) have been identified as naturally occurring in mammalian tissues by Newton and colleagues using FAB/MIKES mass spectrometry[92]. The proteins capable of their biosynthesis, hydrolysis and binding have been identified in rat tissues[17], however the physiological significance of these cyclic nucleotides remains to be elucidated.

7.1.2 Introduction to proteomics

7.1.2.1 Proteome and proteomics

A rough draft of the human genome was completed by the Human Genome Project in early 2001[93, 94]. The completion of the Human genome sequence is one of the most important landmarks in biological research. A major challege in modern biology is to understand the expression, function, and regulation of the entire set of proteins encoded by an organism. It requires an investigation of genes, gene transcripts, proteins, and metabolites which have been termed the genome, transcriptome, proteome and metabolome respectively (Figure 7.11)[95]. Proteome is a term coined in 1994 by Marc Wilkins[96], a graduate student at Macquarie University in Australia, to provide an analogy to the term "genome". Proteomics is the study of the proteome which can be viewed as an experimental approach to explain the information contained in a genomic sequence in terms of the structure, function, and control of biological processes and pathways[97].

7.1.2.2 Challenges in proteomics, post-translational modification

While it is often viewed as the "next step" of genomics, proteomics is more complex than genomics. The genetic information in an organism is contained within the nucleus where it is arranged into genes encoded by DNA which constitutes the chromosomes. During gene expression the information encoded within the genome is converted by a process of transcription to a corresponding primary RNA transcript. This is subsquently porcessed within the nucleus to form mature messenger RNA (mRNA), mRNA is then transported to the cytoplasm where it is bound to ribosomes, to act as a template for

translation into the corresponding protein (Figrue 7.11). It has been estimated that one gene produces 6-8 proteins[98] because of splice variants and post transcriptional modifications[99, 100], which makes the study of the proteome more complex than genomic studies. Most importantly, while the genome is a constant entity, the proteome differs from cell to cell and in a single cell from time to time as it is constantly changing through its biochemical interactions with the genome and environment. One organism will have radically different protein expression in different parts of its body, at different stages of its life cycle and in different environmental conditions. Hence the identification of the proteins in their final state is required if the purpose of proteomics is to understanding of protein function and interaction. Once genes are transcribed they are edited and translated into proteins. Covalent modifications to the amino acid sequence can occur co-translationally or post-translationally and such modifications play a pivotal role in regulating protein activity. Identification of the type of modification and its location offer crucial information for understanding the function or regulation of a given protein. So far more than 200 different modifications have been reported, many of which are known to control signaling pathways and cellular processes[102, 103]. Post-translational modifications include alkylation, acetylation, glycylation, glycosylation and phosphorylation, usually to serine, tyrosine, threonine.

Much of the complexity of higher organisms is believed to reside in the specific post-translational modification of proteins. Post-translational modifications play critical roles in a proteins function and the reversible phosphorylation/dephosphorylation of proteins is one of the most common and important regulatory mechanism[104].

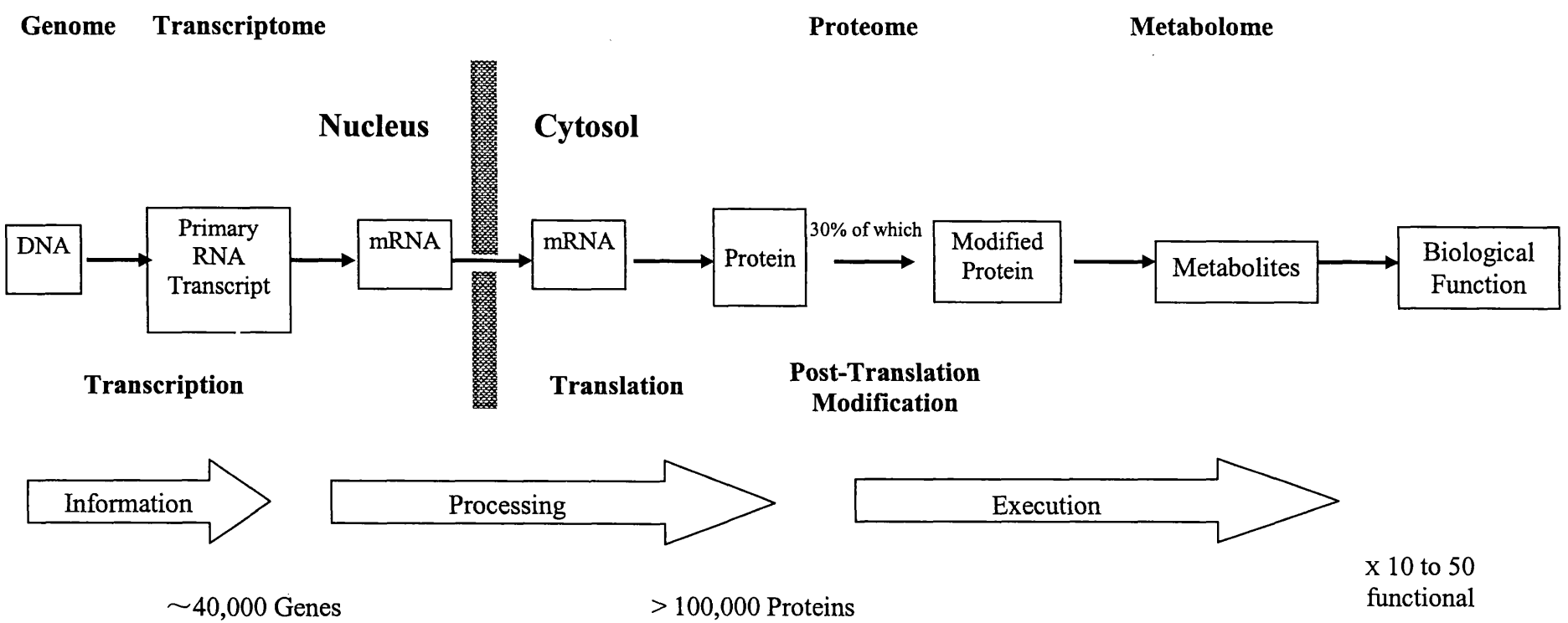


Figure 7.11 Scheme representing the transformation of genomic information into biological function

Phosphorylation plays a key role in the regulation of virtually all cellular events. Many crucial biological processes such as the cell cycle, cell growth, cell differentiation, and metabolism are tightly controlled by reversible protein phosphorylation, modulating protein activity, stability, interaction and localization[105, 106]. It has been estimated that more than one-third of all proteins can be modified by phosphorylation in mammalian cells, and that up to 5% of the genes in a vertebrate genome encode for either protein kinases or phosphatases[107]. The most common type of protein phosphorylation studied involves the formation of phosphate ester bonds with the hydroxyl side chains of serine, tyrosine, and threonine. Serine phosphorylation represents ca. 90% of all cellular phosphorylation, threonine phosphorylation ca. 10%, and tyrosine phosphorylation only 0.1%[108]. The collection of proteins which are phosphorylated is described as the phosphoproteome.

7.1.2.3 Key technologies used in proteomics research

Protein samples of biological origin are by nature highly complex and require sophisticated analytical tools to provide reliable analysis of the components. Proteomics especially challenges the need for robust, automated, wide dynamic range and sensitive high-throughput technologies. Several key technologies have emerged as important tools for the global profiling of the most functional compartment encoded for in the genome at the protein levels. The first technology is the separation of proteins which includes two-dimensional gel electrophoresis and/or chromatographic separation. The second technology, mass spectrometry combined with bioinformatics, has become an important tool in the analysis of protein sequence and structure, making the characterization of part of the proteome simultaneously possible[109]. The basic analytical

requirements in proteome analysis are high sensitivity, high resolution, high throughput and high-confidence protein identification. To reduce sample complexity prior to mass spectrometry for protein identification, one of the following two approaches is usually taken: a) proteins are first separated, and then digested, this is called 'top-down' proteomics[110], b) 'shotgun' proteomics, a complex protein mixture is first digested, peptides are then chromatographically resolved, this is referred to as 'bottom-up' proteomics[111]. In both cases, separation technologies play a critical role in protein identification and analysis. Considering the high complexity of mixtures of biological origin, pre-MS separation technology and dynamic range of pre-MS separation and mass spectrometry becomes the bottleneck for the successful study of biomolecules[112].

7.1.2.3.1 Two-dimensional gel electrophoresis

7.1.2.3.1.1 A brief history

Two-dimensional gel electrophoresis can also be referred to as 2-D gel electrophoresis or 2D-PAGE. Separation of protein mixtures by electrophoresis began early in the twentieth century. In 1956, Smithies and Poulik[113] described a combination of paper and starch gel electrophoresis for the separation of serum protein which is the first two dimensional co-ordinate data. In 1970, Laemmli UK produced the first gel image data by running 2-D gel electrophoresis for increased separation of structural proteins in a bacteriophage where they incorporated isoelectric focusing[114]. In 1975, Klose[115] and O'Farrell[116] published the protocol for high resolution 2-D gel electrophoresis for the analysis of *E. coli* proteins. Since then 2-D gel electrophoresis, which permit the simultaneous analysis of hundreds or even thousands of gene products, have been the

method of choice for separation of complex protein mixtures.

7.1.2.3.1.2 The theory of 2-D gel electrophoresis

7.1.2.3.1.2.1 Isoelectric focusing (IEF)

IEF is an electrophoretic method that separates proteins according to their isoelectric points (pI). Protein molecules carry charge depending on the pH of their surroundings, a protein molecule in solution at any pH other than its isoelectric point has a net average charge. This causes it to move in an applied electric field. Isoelectric focussing is a process that separates proteins on the basis of their mobility at a particular pH. Isoelectric focusing involves utilizing an IPG (immobilized pH gradient) strip and setting up a pH gradient and allowing proteins to migrate in the presence of an electric field to the point in the system where the pH equals their isoelectric point. To establish a pH gradient requires the use of polymeric buffer compounds which resemble proteins themselves, as they have large numbers of both positive and negative charges, and possess isoelectric points over a similar pH range (3-12). Such buffers are called ampholytes; they migrate to their isoelectric point, as there are hundreds or even thousands of individual ampholyte species, they spread across the whole gel slab between the cathode contact and the anode contact. Thus after a few hours of application of the electric field, the ampholytes have migrated and formed a pH gradient; the pH range depends on their composition. Any proteins present also move, since the proteins are mostly larger, they move more slowly, so the pH gradient becomes established before the proteins reach their isoelectric position. A protein applied to the gel will migrate until it reaches a pH equivalent to its own pI, provided that its pI falls within the range of the ampholytes used to generate the pH gradient. IEF concentrates

proteins into focused narrow zones because if they diffuse out of the zone, they will become charged and migrate back to the position of zero net charge under the influence of the electric field[117].

7.1.2.3.1.2.2 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE is currently the most commonly used electrophoretic technique for the analysis of proteins which separates proteins primarily by mass. Polyacrylamide gel is formed by the polymerization of monomers of acrylamide with monomers of a suitable bifunctional cross-linking agent, N, N'-methylene-bis-acrylamide, referred to as bis-acrylamide for short. A three dimensional network is formed by the cross-linking of randomly growing linear polyacrylamide chains by a mechanism of vinyl polymerization with the addition of the catalysts, N,N,N',N'-tetramethylethylethylenediamine (TEMED). The proportion of acrylamide and bis-acrylamide determines the extent of cross linking and are important in determining the physical properties of the gel, including pore size, elasticity, density and mechanical strength[118]. Once the polyacrylamide gel has set, proteins can be applied to the gel matrix which acts as a molecular sieve separating proteins according to their shapes. SDS gels involve running the electrophoresis after denaturing the proteins with the detergent sodium dodecyl sulphate. Dodecyl sulphate binds strongly to proteins and effectively masks the intrinsic charge of the polypeptide chain, consequently, separation can occur only as a result of the molecular sieving through the pores of the gel[119].

7.1.2.3.1.2.3 2-D gel electrophoresis

This technique separates proteins in two dimensions. The first dimension, isoelectric focusing (IEF), separates proteins according to their isoelectric points (pI) and the second dimension, SDS-PAGE, separates proteins according to their molecular weights. Each spot on the resulting gel corresponds to a single protein species in the sample. Thousands of different proteins can thus be separated and the pI and the molecular weight of a specific protein can be obtained.

7.1.2.3.1.3 Advantages and disadvantages of 2-D gel electrophoresis

2-D gel electrophoresis has been employed for more than 30 years and is still widely used. It is reported that 2-D gel electrophoresis can simultaneously separate approximately 2000 proteins, and by using optimised protocol, up to 10,000 proteins can be seen[120]. 2-D gel electrophoresis provides a very high resolution of proteins and high sensitivity based on different possible protein staining methods. The combination of two-dimensional gel electrophoresis with mass spectrometry (MS) is the most common analytical scheme in proteomics. The high resolving power of 2-D gel electrophoresis allows separation of a protein mixture into individual spots in a flat gel. Ideally, a single protein is isolated from a specific spot of interest and then digested to peptides. The identification of proteolytic fragments and hence of the precursor protein can be achieved by MS, usually by matching the MS results with an available database. Although 2-D gel electrophoresis and MS are not directly connected, fully integrated technology that combines the two techniques via robotic systems is now commercially available[121].

However, even with continuously improvements, 2-D gel electrophoresis has unavoidable drawbacks and intrinsic limitations[122]. Some disadvantages of 2-D gel electrophoresis are that it is very time-consuming methodology for reproducible protein separation; it has limited pI range and is not suitable for membrane proteins which count more than 30% of all proteins, and it can only detect high abundance proteins. Dynamic range and protein solubility issues complicate the detection and separation of low-abundance and hydrophobic proteins by 2-D gel electrophoresis. Proteins that occur in low copy number in the whole cell lysate e.g. signalling proteins and transcription proteins have not been successfully investigated using this approach. High mass proteins have great difficulty entering into the SDS gel so they cannot be analysed. Another limitation is that it is difficult to be directly coupled with MS since the necessity of off-line protein digestion.

These drawbacks have undermined the prospects for 2-D gel electrophoresis as a dominant separation technique in proteomics and have stimulated the development of alternative technologies. To achieve a wider dynamic range of analysis, obtain better reproducibility and increase analytical throughput, several multidimensional chromatographic separations have been proposed as alternatives to 2-D gel electrophoresis[123]. In particular, with the emergence of techniques such as nanoflow capillary high-performance chromatography in conjunction with electrospray ionisation tandem MS, attention has been focusing on developing comprehensive multidimensional liquid-based separation techniques.

7.1.2.3.2 Chromatographic separation of proteins and peptides

Over the past decade, the analysis of peptides and proteins by high performance liquid chromatography progressed rapidly due to its high resolving power, reproducibility and its compatibility with electrospray mass spectrometry. The chemical properties of proteins and peptides vary in charge, hydrophobicity, solubility, pI, these differences being the basis for their separation and purification by today's modern HPLC methods. One attractive feature of liquid chromatography is the broad selection of stationary and mobile phases which makes it a versatile and fundamentally important tool in proteomics. Proteins or peptides separated by RP-HPLC can be introduced directly into the mass spectrometer through an ESI source for identification and analysis. Increased resolution of proteins and peptides can be achieved by multiple steps of chromatography. Because of the high resolving power of LC, ion-suppression effects in MS caused by overlapping signals from high and low abundance ions can be reduced or even eliminated. It is also possible to enrich low abundance proteins or peptides presented in complex sample mixtures within a wide dynamic range of concentration using selected LC methods[124]. Although mass spectrometers can measure the mass of intact proteins, proteins can be difficult to handle and to resolve under the same conditions, and the sensitivity for proteins is much lower than for peptide[125].

The analysis of low levels of peptides originating from complex biological samples requires improved sensitivity of any given detection system. This is typically achieved by the use of columns with smaller i.d. such as capillary columns of 300 μm i.d. with a typical flow rate of about 4 μL/min or nanoscale LC-ESI/MS as employed in LCMS experiments. Columns with 50-200 μm internal diameter and HPLC systems with flow

splitters that can handle flow rate of a few 100 nL/min are commercially available. Currently, chromatographical separations are even performed at flow rates as low as 20 nL/min using columns with internal diameter down to 15 μm[126]. Such reduced flow rates offer greatly improved sensitivity in concentration-sensitive detectors such as mass spectrometers. With such a low flow rate, loading a sample onto the column is very time consuming, however, this can be avoided by making use of column switching techniques. The purpose of column switching is to segregate the loading of the sample and the actual separation to reduce the loading time as well as clean up the sample. This is achieved by first loading the sample at a high flow rate onto a short preconcentration-column (trap column) which has a larger diameter than the nano column[127]. Once the sample is loaded and desalted, the trap column is switched to be in line with the 75 μm i.d. nano analytical column to run the gradient at a lower flow rate. When coupled to a tandem mass spectrometer, peptides can be analyzed and selected for fragmentation in the collision cell of the MS instrument as they elute from the column. The sensitivity of nano HPLC compared to conventional HPLC increases approximately 3700 times according to the theoretical down-scale factor[128].

7.1.2.3.3 Mass spectrometry

Mass spectrometry is currently the technique of choice for identification and characterization proteins on a large scale. The power and efficiency of mass spectrometry was recently recognized through the Nobel price in Chemistry in 2002, where both electrospray and matrix-assisted laser desorption/ionisation principles (the two most widely applied ionisation techniques for proteomics work) were rewarded[129,130]. Traditionally, ESI has been associated with tandem MS and protein

sequencing while MALDI was associated with peptide mass fingerprinting (PMF) and these modes are complementary to each other. It has been demonstrated that proteins can be analyzed by mass spectrometry requiring very little or no manipulations of the samples, just by placing thin slice of frozen tissue section (for example a human breast needle biopsy or a mouse brain) directly on the MALDI target plate and applying the matrix solution[131-133]. However, as previously indicated (section 7.1.2.3), in many cases, different proteins are present in very different concentrations in tissues or cells which may contain many hundreds to thousands of different proteins. In these cases, the complexity of the protein sample requires a separation procedure prior to the mass spectrometric analysis of the sample.

Electrospray sources have been coupled to liquid chromatography because the electrospray process is capable of transferring ions from solution directly into the gas phase as they elute from the column. Electrospray mass spectrometry (ESI-MS) enables the detection of proteins due to its ability to form multiply charged ions and thus reducing the *m/z* values to levels that can be detected. The widely established protocol is that enzymatic digestions of the protein mixture are fractionated by C_{18} reverse phase liquid chromatography coupled directly to ESI-MS capable of data-dependent MS to produce fragmentation (MS/MS) spectra from as many components as possible. The resultant spectra are then subjected to database search algorithms that match the MS/MS spectra to amino acid sequence.

7.1.2.3.4 Protein identification by database searching

Proteins are long chain polymers of L-amino acids. All amino acids have in common a carboxyl group and an amino group bonded to a primary carbon. In the protein structure, amino acids are covalently coupled via peptide bonds to form the linear peptide chain. To identify a protein, the first step is to convert proteins to a set of peptides using a sequence-specific protease. Trypsin is an aggressive and stable protease, which cleaves proteins specifically on the carboxy-terminal side of arginine and lysine residues, thus creates information-rich peptides in the preferred mass range for sequencing[125].

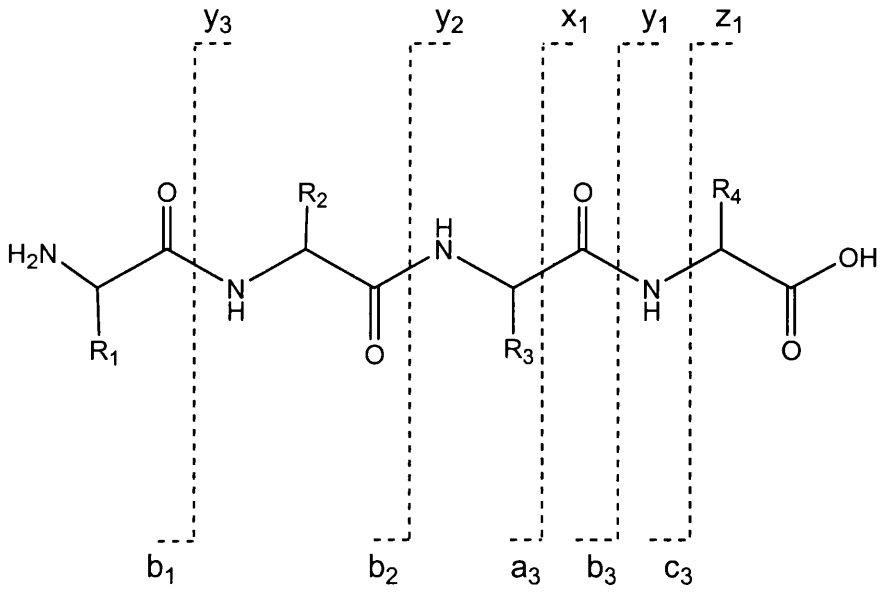


Figure 7.12 Designation for fragment ions when peptide backbone is fragmented

Figure 7.12 shows a peptide with 4 amino acids; in mass spectrometers, peptide fragmentation is induced by collision-induced dissociation, bond breakage mainly occurs through the lowest energy pathways, that is, the cleavage of the peptide bonds. This leads to the formation of b and y ions , b ions are formed when the charge is

retained on the amino-terminal side, y ions are formed when the charge is retained on the carboxy-terminal side, also a, c, x, z ions can be formed as shown in Figure 7.12. Ions are labelled consecutively from the original amino terminus a_m, b_m and c_m, in which m represents the number of amino acid R groups contained in these ions. They are also labelled consecutively from the original carboxyl terminus as z_n, y_n, x_n, in which n equals the number of R groups contained in these ions. Apart from these ions, fragment ions due to the further loss of NH_3 or H_2O can be produced, for example, b_m-NH_3 or y_n-H_2O.

As shown in Figure 7.12, each peptide fragment in a series differs from its neighbour by one amino acid. In theory, it is therefore possible to determine the amino acid sequence by considering the mass difference between neighbouring peaks in a series. This process is called *de novo* sequencing[134]. However, the information in MS/MS spectra is often not complete and there are intervening peaks which may confuse the analysis. Database matching is an alternative way to identify proteins. This is easier than *de novo* sequencing, because although a peptide fragmentation spectrum might not contain sufficient information to derive the complete amino acid sequence unambiguously, it might still have sufficient information to match it uniquely to a peptide sequence in the database on the basis of the observed and expected fragment ions. The disadvantage of database searching is that the experiments should be carried out using organisms whose genome have been sequenced, so that all the possible peptides are known. Peptides can be identified from databases by their fragmentation spectra using pattern searching with several different algorithms such as SEQUEST and Mascot, PeptideSearch, Sonar ms/ms and ProteinProspector. The SEQUEST algorithm developed in 1994 by Yates' research group[135] was utilised in this study. In SEQUEST algorithm, a signal processing

technique called autocorrelation is used to mathematically determine the overlap between the experimental spectrum and a theoretical spectrum derived from the sequence in the database. The overlap is given in the form of scores, including the Xcorrelation score and the Delta correlation score. These scores provide a quick assessment of each match, and they can be a filtering tool. The Xcorrelation score (XC) value describes how well the theoretical for each sequence cross correlates with the observed spectrum. XC scores above 2.0 are usually indicative of a good correlation. The Delta correlation (dCn) describes how different the first peptide match is from the second peptide match in the search. A general rule of thumb is that a dCn of 0.1 or greater is good[136]. The filter range of the scores set in this study is shown in Table 7.4 (Section 7.2.7).

In LC-MS/MS analysis of protein digests the mass spectrometer will cycle through a sequence that consists of obtaining a mass spectrum followed by obtaining tandem mass spectra of the most abundant peaks that were found in the first mass spectra. This process does not necessarily result in measurement of peptides covering the complete protein (100% sequence coverage), for example, very short peptides of only a few amino acids or very long peptides of >30 amino acids are often not detected[103]. However, at least some of the peptides can be sequenced by the mass spectrometer, which is sufficient for protein identification but not for complete protein characterization.

7.1.2.4 Phosphoproteome study by chromatographic enrichment of phosphopeptides

In signalling pathways, kinase cascades are often turned on and off by the reversible addition and removal of phosphate groups. Protein phosphorylation has been shown to play an essential role in a variety of fundamental cellular functions such as gene transcription, cell-cycle progression, energy storage, metabolic regulation and apoptosis[137,138]. Therefore, the identification of phosphorylated proteins and the exact localization of the modification level are essential in understanding a biological pathway at the molecular level. Analysis of the entire complement of phosphorylated proteins in cells has been investigated in recent years. This is the combined result of optimisation of enrichment protocols for phosphorylated proteins and phosphopeptides, better fractionation techniques, especially multidimensional chromatography, and the improvements in methods used to selectively visualize phosphorylated residues by mass spectrometry. It is estimated that there are approximately 100,000 potential phosphorylation sites in the human proteome of which fewer than 2000 are currently known[139]. Traditional procedures used to identify phosphorylated proteins include radioactive labelling with ^{32}P-labeled ATP followed by SDS polyacrylamide gel electrophoresis or thin layer chromatography then Edman sequencing, and the use of phosphospecific antibodies[140]. Most of the traditional methods are inadequate because it is impossible to obtain the large amount of proteins required for phosphorylation analysis.

Mass spectrometry is now the technique of choice for phosphorylation analysis because of its sensitivity, versatility and speed with MALDI-TOF and ESI-MS/MS commonly

used[141]. In the post-source decay (PSD) MALDI-TOF, phosphorylated peptides identification can be carried out by searching for peptides whose mass is shifted by 98 Da or 80 Da due to the loss of H_3PO_4 and HPO_3 from the predicted one. For serine and threonine phosphorylation, $(MH\text{-}H_3PO_4)^+$ is the more abundant ion observed. However, the loss of the phosphate moiety is not favoured in the case of phosphotyrosine, because the phosphate moiety is stabilized by the aromatic ring[142].

In the ESI-MS/MS mode, the side chains on phosphoserine, phosphothreonine and phosphotyrosine can fragment easily on both sides of the phosphoester bond upon collision-induced dissociation (CID). Peptides carrying a phosphate group can be identified by the neutral loss of H_3PO_4 from the precursor ions in the positive mode[143], that is 32.6, 49, 98 Da from the $(M+3H)^{3+}$, $(M+2H)^{2+}$, and $(M+H)^+$ ions, respectively.

Recent technological developments have made it increasingly feasible to directly analyze complex peptide mixtures by LC-MS/MS with the identification of hundreds of peptides and proteins being achieved within a single chromatographic run. However, the comprehensive analysis of phosphorylated proteins has encountered some biological and analytical limitations. From a biological point of view, phosphorylated proteins are often present at very low concentrations and are therefore difficult to be detected in complex mixtures containing high concentrations of other cellular proteins. From the analytical point of view, there are several factors that complicate the analysis of phosphorylated proteins. Firstly, negatively-charged modification can hinder proteolytic digestion by trypsin and secondly, the sequence near the phosphorylation site is usually very hydrophilic, leading to a selective loss of phosphopeptides during the standard methods used for sample preparation. Thirdly, mass spectrometers have a finite dynamic

range, peptides of very low abundance may not be picked up for sequencing in the presence of high abundant unphosphorylated peptides. Finally, phosphorylated peptides may undergo neutral loss of phosphate during MS, which can further limit their detection[142]. Major challenges in the analysis of phosphorylated proteins or phosphopeptides are the isolation of the phosphoprotein, and then the isolation or enrichment of the phosphopeptides from the overwhelming amount of nonphosphorylated peptides present in a complex protein total digest. Currently, immobilized metal affinity chromatography (IMAC) has proven particularly effective in the capture of phosphopeptides[144-146].

IMAC utilizes immobilized Fe^{3+}, Ga^{3+}, K^{+}, or Cu^{2+} ions to selectively retain phosphorylated peptides since phosphopeptides can be captured selectively through their negatively charged phospho-group on immobilized-metal affinity (IMAC) columns. This method relies on high affinity interaction between transition metal ions and phosphorylated side chains of serine, threonine or tyrosine[147]. In this technique, a stationary phase with an immobilized transition metal binding functionality, e.g. imidodiacetic acid (IDA) or nitrilotriacetic acid (NTA) is first charged with either Fe^{3+} or Ga^{3+}. The transition metals therefore form tight complexes with the stationary phase as shown in Figure 7.13. Protein digests are loaded onto the column at low pH, the adsorption of phosphopeptides is based on the coordination between immobilized metal ion and electron donor groups (phosphate group) in the peptide structure. After the washing step, the bound phosphopeptides can be released by changing the pH and disrupting phosphorylated proteins or peptides from their metal ligand complexes.

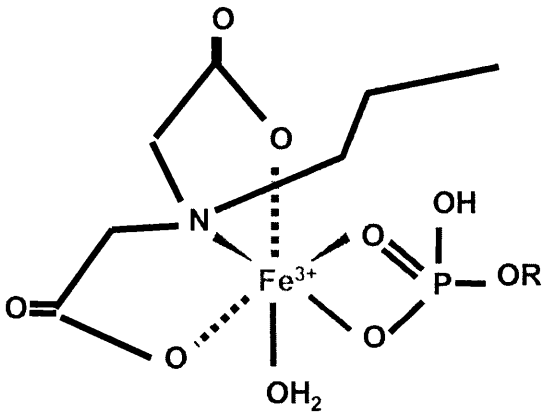


Figure 7.13 Illustration of IMAC chromatography

The advantages of IMAC for phosphorylated protein enrichment are ligand stability, high protein loading, mild elution conditions, simple regeneration, high applicability under denaturing conditions and low cost[148]. However, there are several limitations of IMAC, such as, possible losses of phosphopeptides during elution from the column and the presence of interfering peptides deriving from the acidic non-phosphorylated peptides that show affinity to the immobilized metal ions. To prevent non-phosphorylated peptides binding to the IMAC column through carboxyl groups, recently, Ficarro *et al.* reported a method based on the esterification and thereby neutralization of the negatively charged amino acid residues before the IMAC step, which allows identification of hundreds of phosphopeptides in yeast cell lysates[149]. Because binding of non-phosphorylated peptides is minimized, this approach is very sensitive, phosphopeptides at low femtomol levels were claimed to be identified. Recent research shows that the stainless steel surfaces available in LC/ESI MS hardware can play a principal part in adsorption and even trapping of phosphorylated analytes under acidic conditions[150] and this effect should be considered for future research.

7.1.3 Aims of study

As introduced previously, adenosine 3',5'-cyclic monophosphate (cyclic AMP) has well established functions as a biochemical second messenger, mediating the action of a wide range of mammalian hormones and neurotransmitters, while the second cyclic nucleotide, guanosine 3',5'-cyclic monophosphate (cyclic GMP) performs a similar but more restricted function. In view of the parallel functions of purine and pyrimidine nucleotides in nature, these roles of purine cyclic nucleotides as secondary messengers pose the question as to whether pyrimidine cyclic nucleotide second messengers also exist. The natural occurrence of cCMP has now been unequivocally demonstrated by means of tandem spectrometric analysis of sequentially purified tissue extracts. A number of factors which regulate the specific enzymes responsible for its synthesis and degradation have been identified and many varied biological effects of cCMP have been reported. However, the mechanism of its action and its precise cellular function are not yet fully elucidated. The identification of the phosphorylation substrate of cCMP-responsive protein kinase would be helpful in elucidating the function of cCMP in signal transduction. Bond[83] identified Rab23 as a possible phosphorylation substrate of cCMP-responsive protein kinase in 5min cCMP incubation using IMAC ZipTip. The aim of this study is to identify phosphorylated proteins which are either unique or increased in expression in response to elevated levels of cCMP in different incubation times, that is, 1, 5, 15 and 30 minutes.

Recent improvements in phosphopeptides enrichment by immobilized metal affinity chromatography (IMAC) and in mass spectrometry make it possible to identify phosphorylated proteins on a proteome-wide basis. IMAC coupled to LC-MS/MS has

been proved to be a powerful method to identify phosphorylated proteins present in complex mixtures of non-phosphorylated proteins.

To fulfil our aims in this study, mouse brain homogenates were incubated for a number of time periods with known amounts of ATP and three different cyclic nucleotides, cAMP, cGMP, cCMP or "blank" which contains only ATP without the addition of any cyclic nucleotide. Normal intracellular cyclic nucleotide (cAMP) level in brain is about 0.5 mM[151], hence 2 mM was chosen as elevated level without having adverse effect on functioning brain cells. ATP is phosphate donor, 5 mM was chosen so ATP is in excess and will not be a hindering factor in the study. An online IMAC-nano LC/MS platform for phosphoprotein profiling was set up and phosphorylated proteins in cCMP-incubated mouse brain homogenate were characterised. The identification of these phosphorylated proteins should provide new avenues for investigating the mechanism of cCMP signalling pathways and its cellular function.

7.1.4 Established methods

This study carried on from the work of Bond[83] in the Cyclic Nucleotide Biochemistry Research Laboratory and Biomolecular Analysis Mass Spectrometry (BAMS) Facility. In the work of Bond[83], mouse brain homogenates were incubated with ATP and different cyclic nucleotides and control (with only ATP) for different time periods. Then the homogenate was centrifuged and dialysed overnight. The dialysed brain homogenates were then enriched by IMAC ZipTip followed by trypsin digestion and HPLC/MS/MS for identification. Alternatively, the dialysed brain homogenates were first digested with trypsin and then enriched by IMAC ZipTip followed by HPLC/MS/MS for

identification. In this study, we carry on from the previous work with the aim of: a) the optimisation and the set up of online IMAC system. b) identification of the phosphorylated proteins unique expressed in response to cyclic CMP incubations of different incubation times.

7.2 Experimental

7.2.1 Materials

7.2.1.1 Chemicals

All chemicals were obtained from Sigma-Aldrich (Poole, Dorset, U.K.) with the exceptions: glycerol, acetic acid, hydrochloric acid (HCl), methanol (MeOH), acetonitrile (ACN), obtained from Fisher Scientific (Loughborough, Leic., U.K.). The MALDI matrices α-cyano-4-hydrocinnamic acid and sinapinic acid were purchased from Fluka (Sigma) and the MALDI calibration mixtures **1** and **3** were obtained from Applied Biosystems (Warrington, U.K.). Formic acid (HCOOH) was purchased from Analar (Dorset, U.K.), and the dialysis membrane for the dialysis of the whole brain homogenate purchased from Medicell International (London, U.K.). The oxygen-free nitrogen was supplied by BOC Ltd (Guildford, Surrey, U.K.). Milli-Q purified water was prepared 'in-house' in the BAMS Facility using the Elix® and Milli-Q® Ultrapure water purification system obtained from Millipore U.K. Ltd (Watford, U.K.). All the chemicals used in gel electrophoresis including IPG strips and 2-D Quanti Kit were obtained from Amersham Biosciences (Uppsala, Sweden) except Tris base and glycine, which were obtained from Melford laboratories Ltd (Ipswich, U.K.).

7.2.1.2 Animals

A total of 48 8-week old female mice were sacrificed during this study. The mice were obtained from Harlan UK Ltd (Shaw's Farm, Blackthorn Bicester, Oxon, UK), and were delivered as a batch of 24 mice a time, over two consecutive weeks. The sacrifice of the mice was conducted on the same day as delivery and performed as rapidly as possible by a blow to the base of the head followed by decapitation. The brains were then removed and pooled into an ice-cold buffer to avoid tissue decomposition due to post mortem changes.

7.2.1.3 Apparatus

Immobiline TM DryStrip Reswelling Tray for IPG strip rehydration, Ettan IPGphor platform for first dimension isoelectric focus, and second dimension gel electrophoresis apparatus were obtained from Amersham Biosciences (Uppsala, Sweden).

7.2.2 Sample preparation for the extraction of proteins from mouse brain tissue

The extraction buffer was 50 mM Tris-HCl (pH 7.4), the pH of which was adjusted with a 6 M HCl solution. This buffer also consisted of 78 mM of dithiothreitol (DTT), 1 mM of ethylenediamine tetra-acetic acid disodium salt (EDTA) solution, 1 mM PMSF, 1 μM pepstatin A and 20 μM leupeptin[83]. The buffer was freshly made immediately before use and was maintained at ice cold temperatures during the extraction of the mouse proteins. Mouse brains were placed in 1:9 (w/v) of the extraction buffer immediately after removal[83], then the tissue was disrupted by mechanical homogenization at 900 r.p.m. in a Potter S Elvehjem Homogenizer (Braun Melsunger, Germany) for four 10-second

bursts in an ice bath to ensure effective and full disruption of the cells and minimize protein denaturation due to heating. The resulting crude tissue homogenate was pipetted into 20 mL vials to make 16 aliquots of 5 mL homogenate and was utilised for the protein estimation experiment for the study of the effects of cyclic nucleotides upon murine brain protein phosphorylation.

A standard 2 mM solution of one cyclic nucleotide (cyclic AMP, cyclic GMP or cyclic CMP), and 5 mM ATP disodium salt was prepared in ice-cold 50 mM Tris-HCl buffer (pH 7.4). A control lacking cyclic nucleotides was also prepared. 4 x 5 mL of each of these solutions was pipetted into the previous prepared 5 mL mice brain homogenate solutions in 20 mL vials. These newly produced solutions were then vortexed and incubated at 37 °C for 1, 5, 15 and 30 minutes in a water bath to allow the incorporation of the phosphate groups into proteins catalysed by native protein kinases. At the allocated time the incubations were transferred to an ice bath and kept for 5 minutes to stop kinase activity. The samples were then stored at -80 °C.

Before further analysis, the samples were centrifuged at 16,000 g for 30 minutes to remove cell debris and insoluble material. The sample solution was then dialyzed against 50 mM Tris-HCl with 7 mM DTT and the protease inhibitors 1 mM EDTA and 1mM PMSF in the cold room overnight and separated into 200 μL aliquots and stored at -80 °C.

7.2.3 Protein assay

The Bradford method[152] was utilised for protein assay. The Coomassie brilliant blue G-250 dye, when dissolved in a strong acid, turns a red brown colour due to protonation, but when it binds to a positively charged protein, the blue colour is restored due to a shift in the pKa of the bound Coomassie blue. Bradford reagent: 100 mg of Coomassie Brilliant Blue G-250 was dissolved in 50 mL of 95% ethanol and then made up to a volume of 500 mL with distilled water. The solution was acidified using 100 mL of 85% (w/v) orthophosphoric acid and diluted further with distilled water to give a final concentration of 0.01% (w/v) Coomassie Brilliant Blue G-250.

A standard solution of 1 mg/mL Bovine Serum Albumin (BSA) was prepared and subsequently diluted with Tris-HCl (pH 7.4) to concentrations ranging from 0.2 mg/mL to 1 mg/mL. From these serial dilutions, 100 μL were taken and added to 5 mL of Bradford reagent. The tube were then vortex mixed and left for 2 minutes to allow the effective binding of proteins to the dye and their absorbance at 595 nm was measured using an Agilent 8453 UV-visible spectrophotometer. A calibration curve was produced by measuring the absorbance for each BSA dilution (Figure 7.14). 20 μL of the crude brain tissue homogenate sample and 80 μL Tris buffer were mixed with 5 mL of the Bradford reagent, the absorbance was then measured at 595 nm and concentration calculated from the calibration curve.

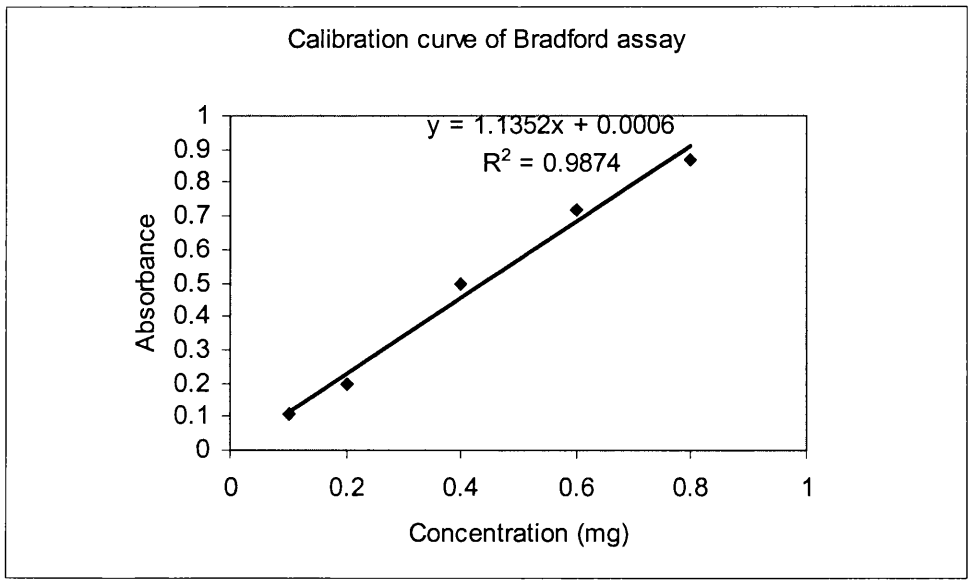


Figure 7.14 Calibration curve of Bradford method

7.2.4 2-D gel electrophoresis

The separation of proteins in the tested sample was primarily performed by two-dimensional gel electrophoresis. The original protocols that were used can be found in the Amersham manual for 2-D gel electrophoresis (code number 80-6429-60), however, several steps were altered to optimise the procedure.

7.2.4.1 TCA/acetone precipitation

Brain homogenates were suspended in 10% TCA in acetone, with 0.07% 2-mercaptoethanol overnight at -20 °C, proteins were pelleted by centrifugation and the pellets were washed with ice cold acetone containing 2-mercaptoethanol at least 3 times before the pellets were freeze dried.

7.2.4.2 Rehydration of the IPG strips

The pellets were redissolved in rehydration buffer and protein concentration was quantitated by 2-D Quanti Kit according to the users' manual[153]. A suitable volume of sample was pipetted into one lane in the Immobiline TM DryStrip Reswelling Tray. A 7 cm IPG gel strip was slipped into the lane with its gel side facing down and its pointed end towards the back of the tray to ensure maximum absorption of the rehydration solution. Then the strip was covered with about 4 mL of Dry Strip Cover Fluid to avoid any evaporation and crystallization of urea. The strip was then protected from dust and contamination by the protective tray cover and left to rehydrate overnight at room temperature.

7.2.4.3 Isoelectric focusing

After the rehydration of the IPG strips, the gel strip was rinsed with deionised water and transferred to the Ettan IPGphor platform and two electrodes were applied on the holder over the anodic and cathodic end of the strip. The gel was placed face up and was covered with 1-2 mL Dry Cover fluid to avoid the crystallization of urea and "burning" of the strip during the IEF. When the strip holder was properly assembled, it was covered with the appropriate lid and IEF was run by using the parameters (Table 7.1) programmed in the Ettan IPGphor.

Table 7.1 IEF conditions for a 7 cm mini gel

Voltage (V)	Voltage mode	Step duration (h:min)	Volt-hours (Vh)
100	Step and Hold	1:00	100
200	Step and Hold	1:00	200
500	Step and Hold	1:00	500
1000	Step and Hold	1:00	1000
5000	Gradient	0:20	830
5000	Step and Hold	2:00	10000
Total		6:20	12630

7.2.4.4 Casting of SDS-PAGE gels

A 7 cm resolving gels containing 10% acrylamide were prepared from 3.33 mL of 30% bis-acrylamide stock and 4.02 mL of distilled water, 2.5 mL of 1.5M Tris-HCl (pH 8.8) and 100 µL 10% (w/v) SDS stock. Just prior to casting, 5 µL of N, N, N^1, N^1-tetramethylenediamine (TEMED) and 50 µL of 10% ammonium persulphate were added to induce polymerization, giving a total volume of 10 mL. The mixture was immediately poured between two clean glass plates. Water saturated butanol was pipetted immediately onto each gel to form a flat upper surface. The homogenous gels were then allowed to polymerize for at least one hour before being ready for the electrophoresis.

7.2.4.5 Equilibration and loading of IPG strips on the gels

Each IPG strip was equilibrated in order to fix the pH range on the strip prior to placement onto the gel. This process was carried out in a suitable plastic tube (supplied by Amersham Biosciences) with the gel side of the strip facing up. The first

equilibration step involved the addition of DTT (10 mg/mL) to the equilibration buffer and introduction of each strip to individual tubes. The strip was left to equilibrate for 13 minutes. Then it was equilibrated with an iodoacetamide solution (25 mg/mL) for another 13 minutes and removed from the plastic tube, its gel side was rinsed with a small amount of Milli-Q water.

With the help of forceps the IPG strips were then positioned on the top part of the casted separation gel and along the gap of the plates, together with SDS electrophoresis buffer which help the strip slipping between the plates. A layer of warm agarose and bromophenol blue solution in 1 x SDS buffer was pipetted on top of the strip to ensure that it would not lose contact with the gel. The agarose was then allowed to set for approximately 5 minutes before the commencement of SDS-PAGE.

7.2.4.6 SDS-PAGE electrophoresis

After the gel was cast and IEF strip was loaded on top of the resolving gel, the gel was secured onto the gel electrophoresis apparatus and the electrical supply connected. Prior to running the buffer reservoir was filled with SDS buffer containing 0.5 g SDS, 1.5 g Tris and 7.2 g glycine made up to 500 mL with distilled water. The apparatus was then allowed to run for approximately 15 minutes at 50 volts and 45 minutes at 200 volts until the bromophenol blue dye reached the bottom of the gel. The schematic diagram of 2-D gel electrophoresis is shown in Figure 7.15.

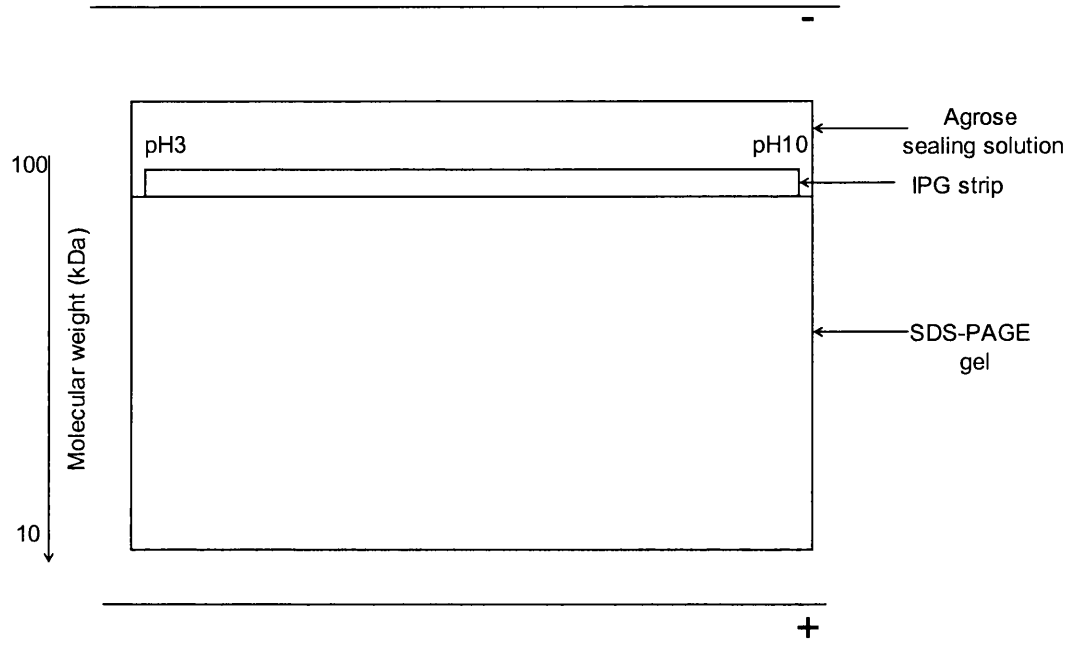


Figure 7.15 Schematic diagram of 2-D gel electrophoresis

7.2.4.7 Staining of proteins on gels

After the SDS gel electrophoresis, the IEF strip was discarded and the resolving gel was then carefully removed from the plates and stained in solution containing 0.12% Coomassie Blue G250, 10% ammonium sulphate, 20% methanol, 10% O-phosphoric acid. The gel was then destained in deionised water until spots of protein were clearly seen.

7.2.5 Protein identification by IMAC chromatography

7.2.5.1 Trypsin digestion

Sequencing grade modified trypsin (Promega, Madison, USA) was used for digestion of the freeze dried brain homogenates. The trypsin solution was made by first dissolving trypsin (20 μg) in 20 μL of 50 mM acetic acid, then diluting 50 times with 40 mM

NH_4HCO_3 with 10% acetonitrile. The trypsin solution was added to the brain homogenate to create a trypsin: protein ratio of 1:100 and incubated at 37 °C overnight. The reaction was terminated by freeze-drying the samples. The samples were resuspended in 0.1% acetic acid before application to further separation and LC/MS analysis.

7.2.5.2 Optimisation of IMAC techniques

7.2.5.2.1 IMAC Ziptip

IMAC ZipTips®$_{MC}$ were utilised following manufacturers instructions [154]. Each 10 µL ZipTip®$_{MC}$ was activated with the aspiration and expulsion to waste of 50% aqueous acetonitrile containing 0.1% acetic acid (Wash solution 1) three times. The metal chelating resins were charged by the aspiration and dispensation of 10 µL metal solution (200mM of ferric chloride in 10 mM HCl). The ZipTip®$_{MC}$ was then washed three times with Milli-Q grade water (Wash solution 2) and three times with 1% acetic acid/10% ACN (Wash solution 3), with all expulsions going to waste. The ZipTip®$_{MC}$ was then equilibrated by the aspiration of Binding solution: 50 mM 2-[N-morpholino] ethanosulfonic acid (MES) buffer containing 10% acetonitrile and adjusted to pH 5.5 with ammonium hydroxide and dispensing it to waste, then repeating this operation 4 times.

The binding of phosphorylated proteins was conducted by the aspiration on and dispension off the ZipTip of each sample in MES buffer (10 µL) to waste twenty times to ensure maximum binding was achieved. The ZipTip®$_{MC}$ was then washed three times with the binding solution followed by Wash solution 1 and Wash solution 2, each time

the resulting eluent was dispensed to waste.

The elution of bound phosphorylated proteins was achieved by ten times aspiration and dispersion in 10 μL of 0.3 M ammonium hydroxide solution. The eluted peptide solution was then concentrated and reconstituted in 0.1% formic acid and analysed by MALDI-TOF MS and LC/MS.

7.2.5.2.2 IMAC SPE

In this study, an in-house made IMAC SPE cartridge was evaluated for the enrichment of phosphorylated proteins and peptides.

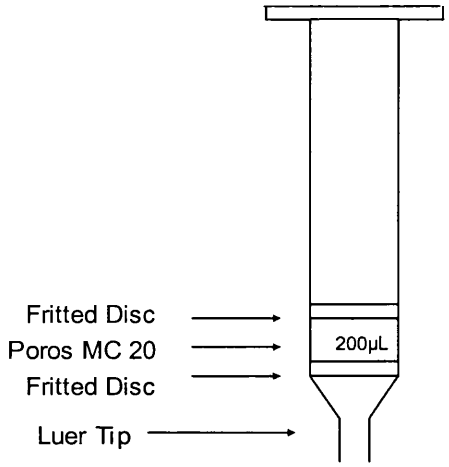


Figure 7.16 IMAC SPE cartridge (1mL)

IMAC SPE cartridges were made by packing 200 μL slurry of SelfPack Poros 20 MC Metal Chelate Affinity Packing (Applied Biosystems, Foster City, CA, USA) into an empty 1 mL C_{18} SPE cartridge under vacuum (Figure 7.16). The binding and elution procedures were optimised from Ficarro *et al*[150]. The digested peptide samples were

dissolved in 200 μL 0.1% HAC in 1:1:1 of methanol/ACN/H_2O. IMAC SPE cartridges were first washed off with 200 μL of 0.1%HAC, followed by 400 μL of 100mM $FeCl_3$ solution, 200 μL of 0.1%HAC, then peptide samples were loaded onto the cartridges, the unbound peptides washed with 200 μL of NaCl solution, followed by 200 μL of 0.1%HAC, the phosphorylated peptides eluted by 200 μL Na_2HPO_4. The cartridges are used under gravity and the steps of IMAC SPE are illustrated in Figure 7.17.

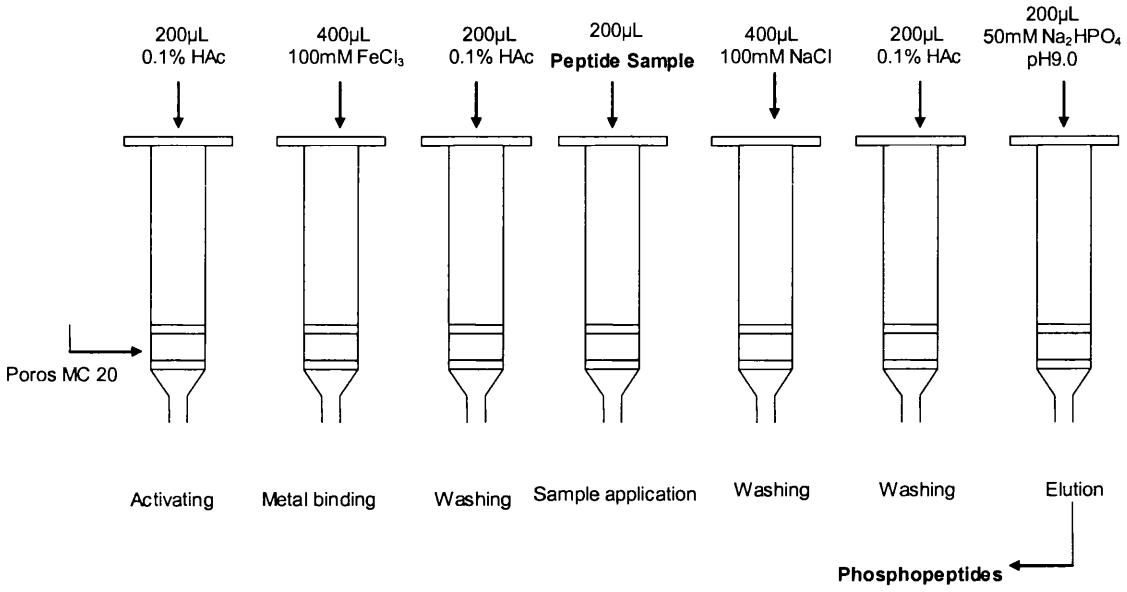


Figure 7.17 Steps of IMAC SPE

After the eluents from the IMAC SPE cartridge were collected, HAC was added to adjust the pH of the eluents before they were applied to C_{18} SPE cartridges to further clean up the sample. The 1 mL C_{18} SPE cartridges (Bond Elute, Varian, Harbor City, CA, USA) were activated by 800 μL of 100% acetonitrile, follow by 800 μL 0.1% TFA, the acidified peptides eluted from the IMAC SPE cartridges were then loaded onto the C_{18}

SPE cartridges. The unbound materials were washed off by 800 μL 0.1% TFA, then the desalted peptides eluted by 0.1% TFA in 75% ACN. The elutions were concentrated to dryness and analysed by MALDI-TOF and LC/MS.

7.2.5.3 On-line IMAC LC/MS for phosphopeptides enrichment and identification

The on-line IMAC experiments were performed using a LC Packings Ultimate Capillary LC system (Dionex, Amsterdam, Netherlands). A FAMOS microautosampler (Dionex) equipped with a 20 μL loop was used for injections. The on-line IMAC was designed with two pump switch systems. The IMAC trap column was installed in between port 6 of injection valve and port 2 of loading pump valve B (Figure 7.18), the loading pump flow rate was 5 μL/min. The analytical column was run at a flow rate of 200 nL/min using an Ultimate pump with a flowsplitter. The IMAC trap column was first connected to a waste line (port 3, see dashed line in Valve B). It was initially flushed with 20 μL of 50 mM EDTA for 8 minutes, followed by two washes with 20 μL of 100 mM $FeCl_3$ before 20 μL of sample were loaded onto the trap. Next 20 μL of 100 mM NaCl in 0.1% HAC were loaded in order to wash off any non-phosphorylated peptides, after that the loading pump valve B was switched so that port 2 was connected to port 1 (solid line of valve B) and port 4 of valve A, meantime the valve A was in position 1-2 (dashed line in valve A). 20 μL of elution solution 50 mM Na_2HPO_4 (pH 9.0) were then loaded onto the IMAC trap which eluted phosphorylated peptides onto a precolumn of 300 μm i.d. x 5 mm packed with the stationary phase 3 μm 100 Å PepMap C_{18} (Dionex) for desalting and preconcentration, via a SWITCHOS pump (Dionex).

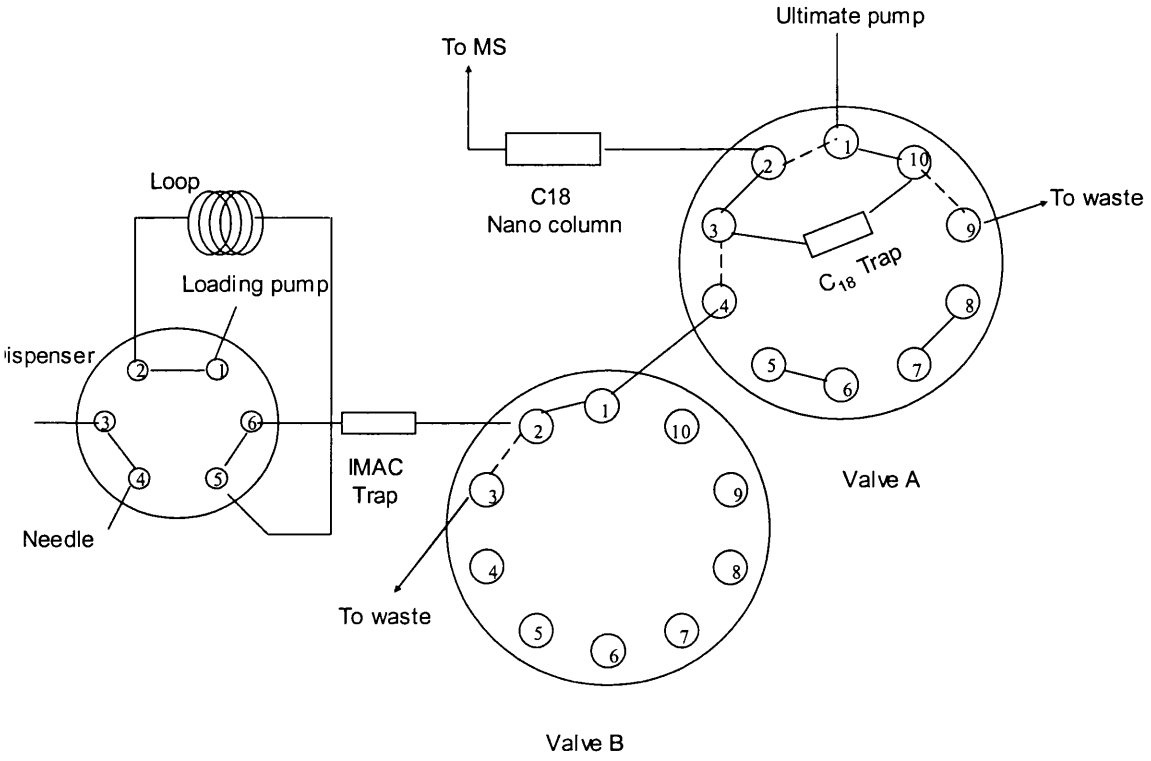


Figure 7.18 Schematic diagram of on-line IMAC system

Valve B: Peptides loaded onto IMAC trap, phophopeptides bound on the trap while unphosphorylated peptide washed off to waste at pH 3-4 (dotted line), bounded phosphopeptides eluted at pH 9.0 (solid line)

Valve A: enriched peptides were loaded onto the C_{18} precolumn and desalted (solid line), the trapped and cleaned peptides was back flushed to analytical column and mass spectrometry (dotted line)

The eluted phosphopeptides were trapped on the precolumn for 3 minutes to allow sufficient time to desalt the sample. After 3 minutes, valve A was switched and the peptides trapped on the C_{18} precolumn back flushed onto the analytical column. The Chromeleon software suite (Dionex) was used for instrument control, and also for triggering of MS data acquisition. The peptides were resolved on an in-house made analytical fused-silica nanocolumn of 15 cm x 75 μm i.d. packed with 3 μm particles,

100 Å pore C_{18} PepMap stationary phase (Dionex). Mobile phase A was 0.1% formic acid in 98% water and 2% acetonitrile while mobile phase B was 0.1% formic acid in 98% acetonitrile. Peptides were eluted from the column with a linear gradient of 0-45% B over 110 minutes. Finally, the column was re-equilibrated for 30 minutes with the initial mobile phase A between runs. Details of valve switches of on-line IMAC system are shown in Table 7.2.

Table 7.2 Details of valve switches of on-line IMAC system

Event	Valve B*	Valve A*	Time (min)
EDTA washing	1-10	1-2	8
Metal binding	1-10	1-2	8
Metal binding	1-10	1-2	12
Sample loading to IMAC trap	1-10	1-2	8
Washing off unbound peptides	1-10	1-2	12
Eluting from IMAC trap to C18 trap	1-2	1-2	15
Backflushing from C18 trap to MS	1-10	1-10	140

* For valve position 1-10 and 1-2 see Figure 7.17

The eluted peptides were analyzed by an LCQ Deca XP ion trap mass spectrometer equipped (Thermo Finnigan, San Jose, CA) with a nanoelectrospray ion source, the mass spectrometry conditions are shown in Table 7.3. The mass spectrometer was operated for 140 minutes in a data-dependent MS/MS mode which included four scan events. The full scan mass spectrum was first obtained and followed by collision-induced dissociation of the three highest abundant ions selected from the full scan. The mass scan range is 475-2000 Da, an isolation width of 2 Da and normalized collision energy of 35% (arbitrary units) was used in the MS/MS analysis.

Table 7.3 Parameter setting of nanoLC/MS and LC/MS/MS on LCQ ion trap

Sheath gas flow, aux gas flow rate	0
Spray voltage	1.7 kV
Capillary temperature	160 °C
Capillary voltage	10 V
Tube lens offset	10.00 V
Multipole 1 offset	-7.25 V
Lens voltage	-16.00 V
Multipole 2 offset	-10.00 V
Multipole RF Amplifier	400 V
Entrance lens	-48 V

7.2.6 Matrix-assisted laser desorption ionisation (MALDI-TOF) MS

The MALDI-TOF MS analysis was conducted on the Voyager-DE STR mass spectrometer (Applied Biosciences, Warrington, U.K.) in positive ion detection mode, the pulsed nitrogen laser operating at 337 nm and linear mode for protein analysis and reflector mode for peptide analysis. The peptide ion produced were extracted with a 400ns delay and accelerated to 25 kV. The accumulation of signals from 100 shots were averaged, the mass spectrum generated then the laser intensity altered accordingly and fired at various areas around the crystallised sample.

Before the analysis of homogenate incubates, the instrument was externally calibrated with calibration mix **1** (Applied Biosciences, Warrington, U.K.) for peptide analysis and calibration mix **3** (Applied Biosciences) for protein analysis. The calibration mixtures were prepared according to the protocols suggested by Perseptive Biosystems, where 1 in 24 dilution of each calibration mixture in the appropriate matrix is required.

0.5 μL of protein or peptide sample was applied to the MALDI plate, and then 0.5 μL of matrix was applied on top of it and allow it to dry at room temperature. The matrix used for the analysis of proteins was a saturated solution of 10 mg/mL sinapinic acid in 70/30 0.3%TFA/ACN. For peptide analysis, a saturated solution of 10 mg/mL α-cyano-4-hydroxy-cinnamic acid in 50% 0.1% TFA/ACN was used.

7.2.7 Protein sequence database searching

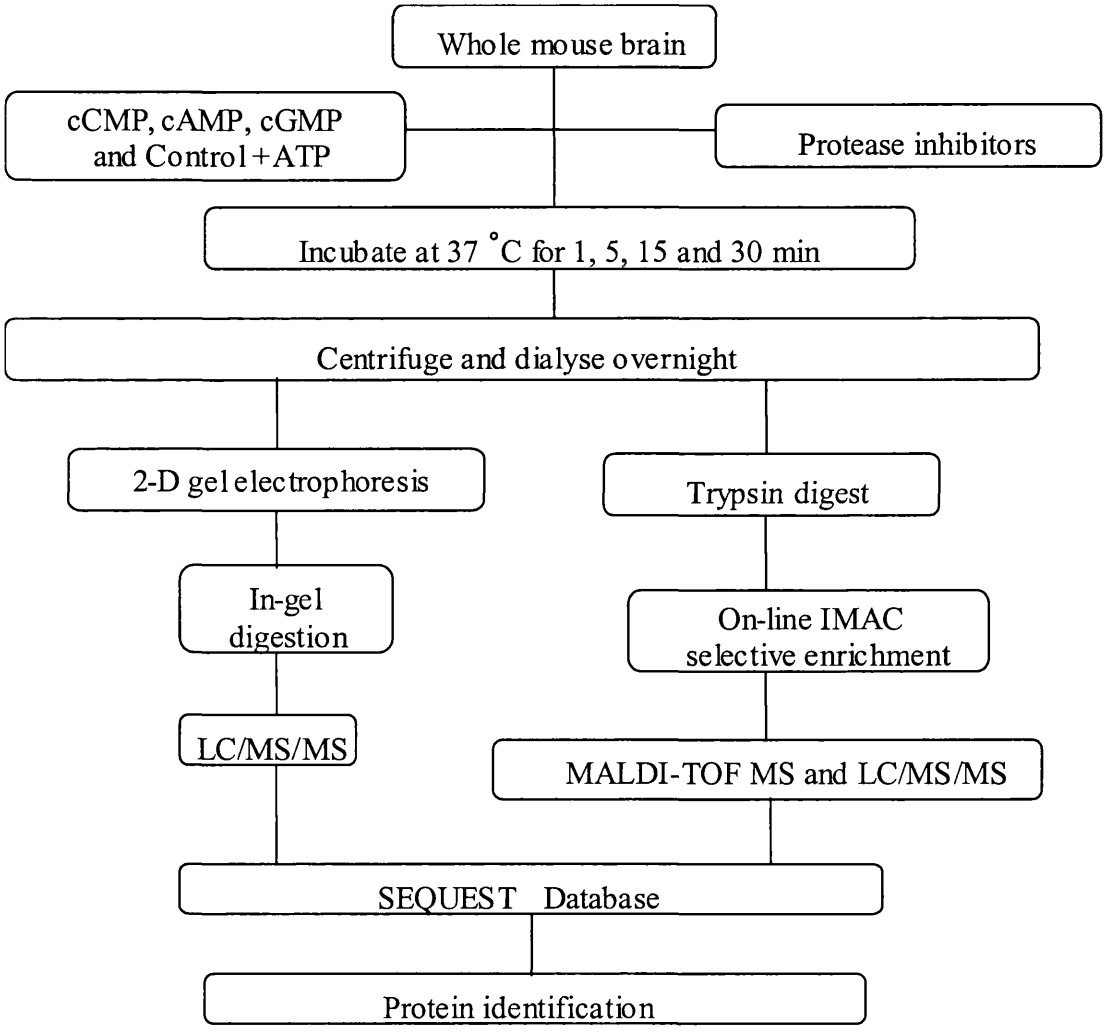
The SEQUEST database containing the sequence of mouse proteins from the Harvard microchemistry website (http://biowork.spc/intrachem.html) was used for the correlation of trypsin digested proteins. Software parameters were set to detect a differential modification of 80 Da on serine, threonine, and tyrosine. All data sets were searched using the following acknowledged constraints[136]: ±2.5 Da mass tolerance for peptide precursor mass searching; ±0.5 Da mass tolerance for the fragment ions. The output files from the correlation analysis were further studied using the XC scores, to produce a list of identified peptides and corresponding proteins as well as the comparison of the virtual spectra produced by SEQUEST and the actual MS spectra obtained from the nanoflow LC-ESI analysis. The positive identification of proteins follows the filter criteria as summarized in Table 7.4 and the assignments of phosphopeptide sequences were confirmed by comparing the acquired MS/MS spectra to the theoretical fragmentation pattern carefully.

Table 7.4 Criteria for SEQUEST database searching

Ion	XC	dCn	Sp	Different peptides
Charge +1	≥ 1.8	≥ 0.1	≥ 150	≥ 2
Charge +2	≥ 2.2	≥ 0.1	≥ 150	≥ 2
Charge +3	≥ 2.8	≥ 0.1	≥ 150	≥ 2

Where: XC-cross correlation value
dCn-delta correlation value
Sp-preliminary score
Different peptides-the number of times the peptide was scanned

7.2.8 Flow diagram of the protocols developed



7.3 Results and discussions

7.3.1 Protein quantitation

The Bradford method was applied to obtain an estimation of the amount of total protein present in the crude brain homogenate. The calibration curve produced was used to determine the amount of total protein present in the brain extract. The linear regression equations were derived from the linear trend line and were used to calculate the amount of protein present in the various protein preparations from the incubates after centrifugation and dialysis. The average protein concentration of the brain homogenate quantitated by the Bradford method was 4.2 mg/mL.

7.3.2 Protein separation by 2-D gel electrophoresis

2-D gel electrophoresis is the most widely used method in protein separation and was initially evaluated in this study. Figure 7.19 shows the 2-D PAGE images of the blank and cCMP incubations, the reproducibility was satisfactory, but the gel spots have vertical streaks. The reason of this is that the proteins have not fully dissolved and entered the gel properly, a common problem for gel analysis in this study. The protein homogenate was centrifuged at the speed of 16,000g to remove cell debris and insoluble material, so dialysis was then required to remove the lipid and salt and the membrane proteins were still in the supernatant. A disadvantage of 2-D gel electrophoresis is that it is not suitable for membrane proteins and brain proteins do have a high proportion of membrane proteins, furthermore many signal transduction-related proteins are membrane proteins[155]. If the brain homogenates were further ultracentrifuged at 50,000 g to obtain only cytosol of the brain proteins, improved 2-D PAGE images with

reduced streaking and tailing would be expected, but at the same time it may lose all the membrane proteins potentially important in our study. As there was no obvious difference between the two gels, no spot cutting and protein identification was carried out in this experiment.

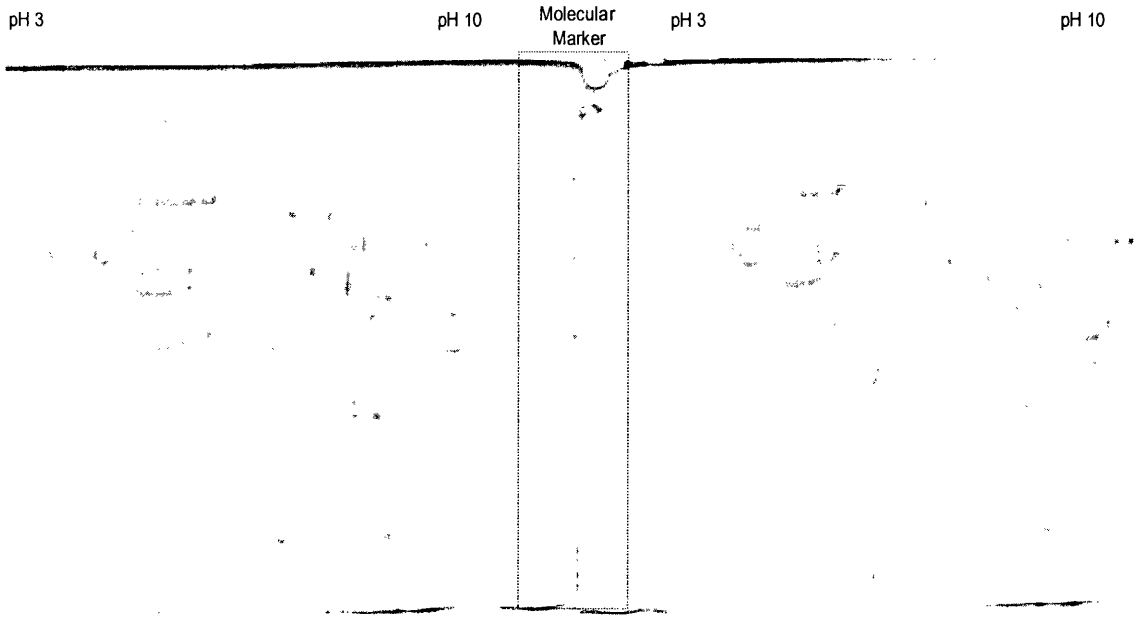


Figure 7.19 2D-PAGE comparison of Blank (Left) and cCMP charged incubates (Right)

7.3.3 Evaluation of different IMAC strategies

7.3.3.1 Evaluation of IMAC ZipTip

β-casein, molecular mass of 26.6 kDa, is a naturally occurring phosphorylated protein. According to literature[156], there are two β-casein trypsin digested phosphopeptides, one mono phosphorylated peptide, FQpSEEQQQTEDELQDK, and one tetraphosphorylated peptide, RELEELNVPGEIVE*p*SL*p*S*p*S*p*SEESITR. In this study, the trypsin digest of β-casein was used as a standard peptide mixture to evaluate the IMAC phosphopeptide

enrichment by using the protonated phosphopeptide molecule of *m/z* 2061.8 corresponding to the amino acid sequence FQ*p*SEEQQQTEDELQDK. The method development started with the existing method developed by Bond[83] in which IMAC ZipTips were used. MALDI-TOF spectra of standard peptides before and after passage through IMAC ZipTip are shown in Figure 7.20. In this figure, after the IMAC ZipTip, the peaks with *m/z* 1382.9, 2185.7, 2222.7, 2910.7 and 3111.7 are absent while the peak with *m/z* 2061.8 corresponding to the monophosphopeptide remains very strong, indicating that the ZipTip enriches the phosphopeptide by removing some unphosphorylated peptides. However, the tetraphosphorylated peptide can not be observed after the IMAC ZipTip and there is a peak at *m/z* 1967.3, corresponding to an unphosphorylated peptide, observed as being bound to the IMAC ZipTip.

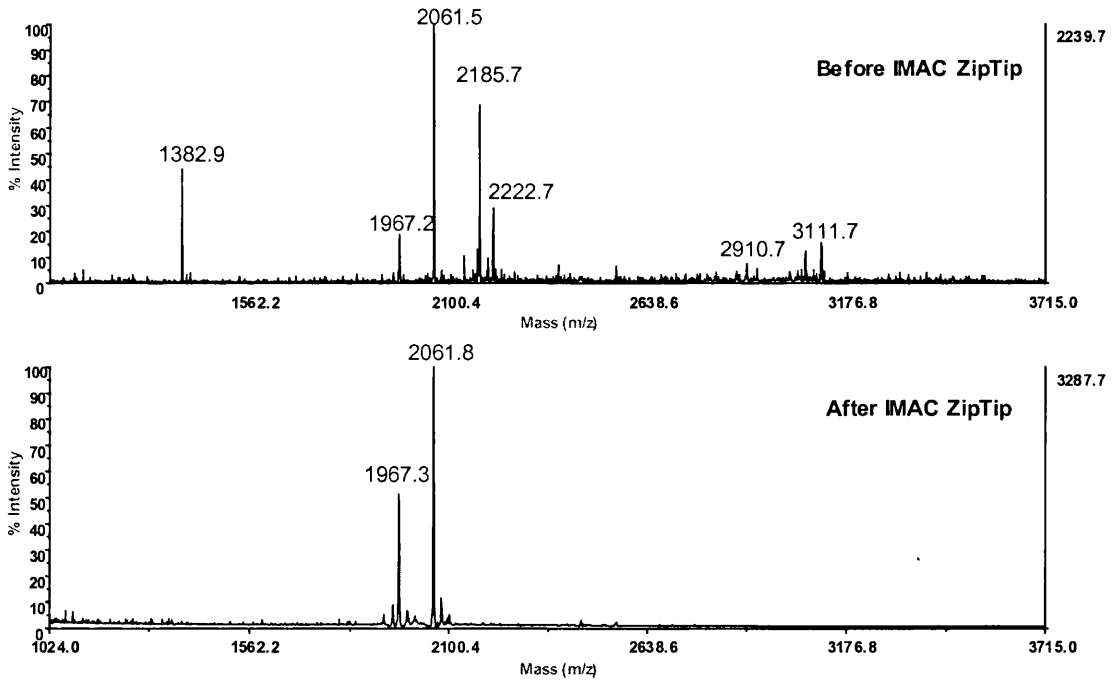


Figure 7.20 MALDI-TOF spectrum of β-casein trypsin digested peptide mixture before and after IMAC Ziptip

However, when the mouse brain homogenate trypsin-digested samples were applied to IMAC ZipTips, poor quality spectra were observed, as shown in Figure 7.21, in which hardly any peptides can be observed. Since protease inhibitor were added in the brain homogenate, larger peptides should be present in the digested sample. It indicates that IMAC ZipTips are not suitable for the study of phosphorylated peptides of complex protein mixtures such as brain homogenates. This finding is not surprising since the ZipTip is a 10 μL pipette tip with a micro-volume bed of chromatography media fixed at its end. It is intended for purifying and concentrating femtomoles to picomoles of protein or peptide samples before analysis and is not designed and consequently may not be suitable for complex protein digest mixtures such as these. Alternative methods to enrich phosphorylated peptides were therefore investigated.

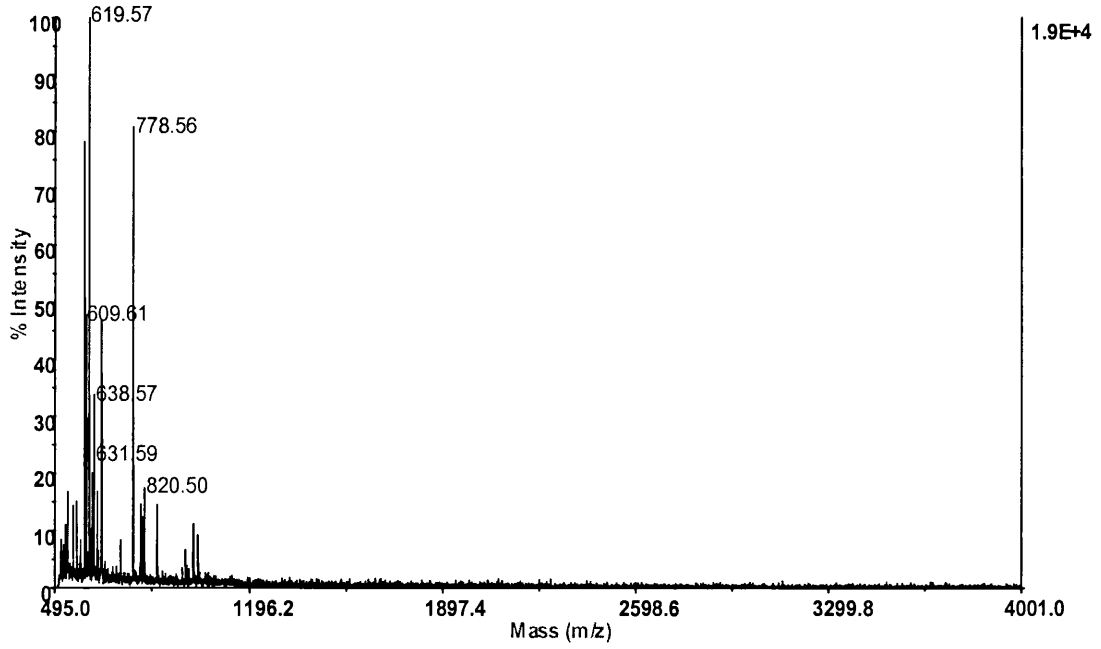


Figure 7.21 MALDI-TOF spectrum of elution of IMAC ZipTip of peptides digested from proteins in brain tissue homogenate

7.3.3.2 Evaluation of IMAC SPE

Standard peptides were tested on the IMAC SPE cartridge. Figure 7.22 shows the comparison of MALDI-TOF spectra of elution of IMAC ZipTip and SPE of trypsin digested peptides of standard protein β-casein; IMAC SPE and ZipTip show very similar spectra, both techniques enrich the phosphopeptides by removing some unphosphorylated peptides to some extent but do not perform perfectly. The tetraphosphorylated peptide was absent in both IMAC systems, and the peak with *m/z* 1967.3 corresponding to unphosphorylated peptide was also present in both IMAC; thus this is deduced to be an intrinsic shortcoming of IMAC itself, i.e., the tetraphosphopeptide might bind tightly on the metal chelating resin and can not be eluted while non-phosphorylated peptides can also bind to the IMAC column through carboxyl groups. However, it is still a method of choice to selectively enrich the phosphorylated peptides.

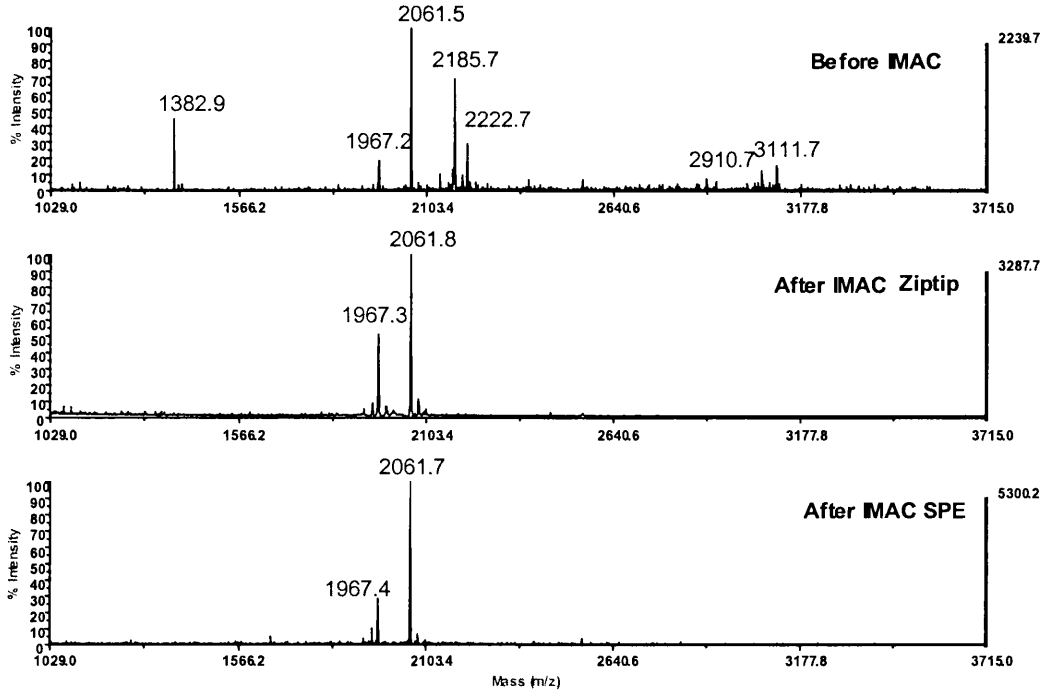


Figure 7.22 Comparison of IMAC ZipTip and SPE of trypsin digested peptides of standard protein β-casein

Since the bed volume of IMAC SPE is several times more than IMAC ZipTip, better loading capacity of IMAC SPE was to be expected and was proved by MALDI spectrum of the eluted peptides of brain homogenate, as shown in Figure 7.23, the MALDI-TOF spectrum of IMAC SPE of peptides digested from proteins in brain tissue homogenate exhibits many more peptides comparing to IMAC ZipTip enrichment (Figure 7.21).

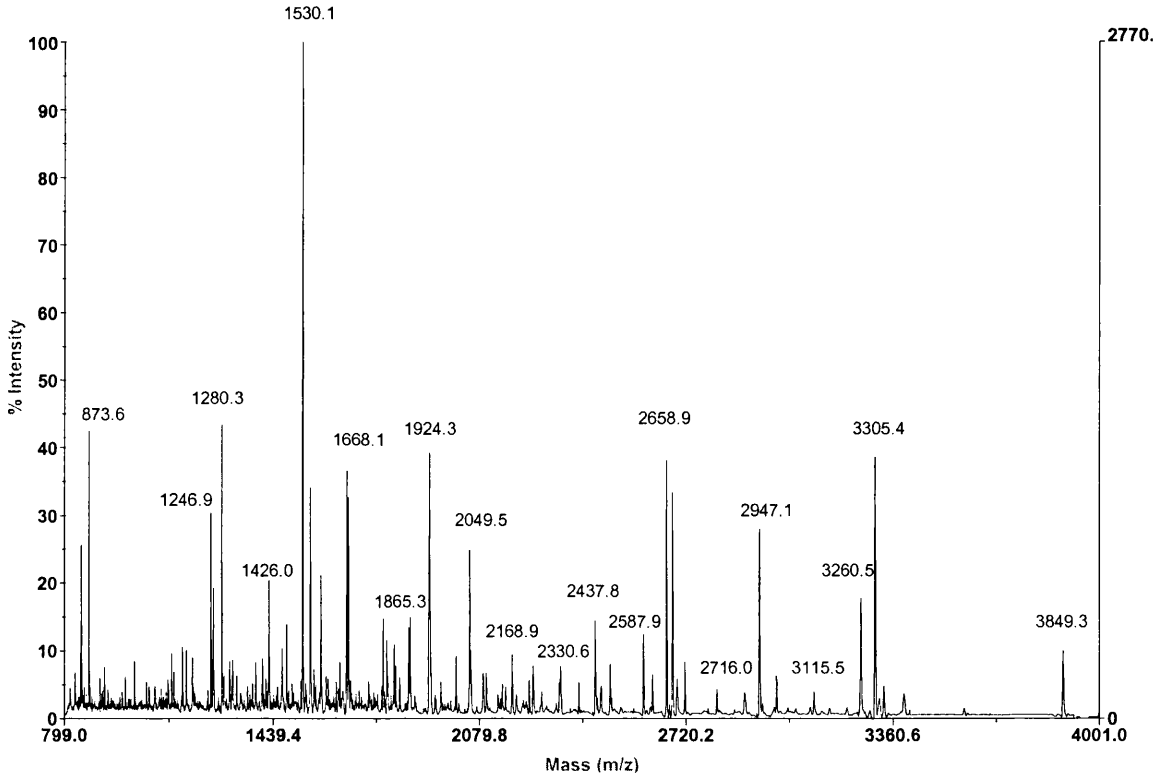


Figure 7.23 MALDI-TOF spectrum of IMAC SPE of peptides digested from proteins in brain tissue homogenate

7.3.3.3 Evaluation of online IMAC trap column for phosphopeptides enrichment

An online IMAC system was derived from IMAC SPE cartridge which had proved to be successful in the selective enrichment of phosphopeptides. The setup was described in

section 7.2.5.3. 10 μL of β-casein trypsin-digested peptides were loaded by autosampler onto the IMAC trap column, the eluents of sample loading, washing and Na_2HPO_4 elution were collected and analysed by MALDI-TOF. Figure 7.24 shows mixtures of peptides before loading onto the IMAC trap; the monophosphorylated peptide with *m/z* 2061 is relatively weak comparing to other unphosphorylated peptides. In Figure 7.25, the spectrum of the eluent of the sample loading step, no phosphorylated peptide was observed; similarly in Figure 7.26, the spectrum of the eluent of washing step, phosphorylated peptide was also not observed. Figure 7.27 is the spectrum of the eluent of pH 9.0 Na_2HPO_4 buffer; here the monophosphorylated peptide was highly enriched, tetraphosphorylated peptide was absent but there were also unphosphorylated peptides observed. These figures demonstrate that the on-line IMAC system set up in this study can enrich phosphorylated peptides efficiently as previously seen in IMAC ZipTip and SPE. The detection limit for known phosphopeptide from trypsin digested beta-casein was approximately 1 pmol. The spectra of brain homogenate trypsin digests elution from the IMAC trap column show a similar view of that of IMAC SPE (Figure7.23). Therefore the brain tissue homogenates incubated with cyclic nucleotides were first digested with trypsin and then applied to this on-line IMAC system before entering nano-LC/MS for further separation and identification.

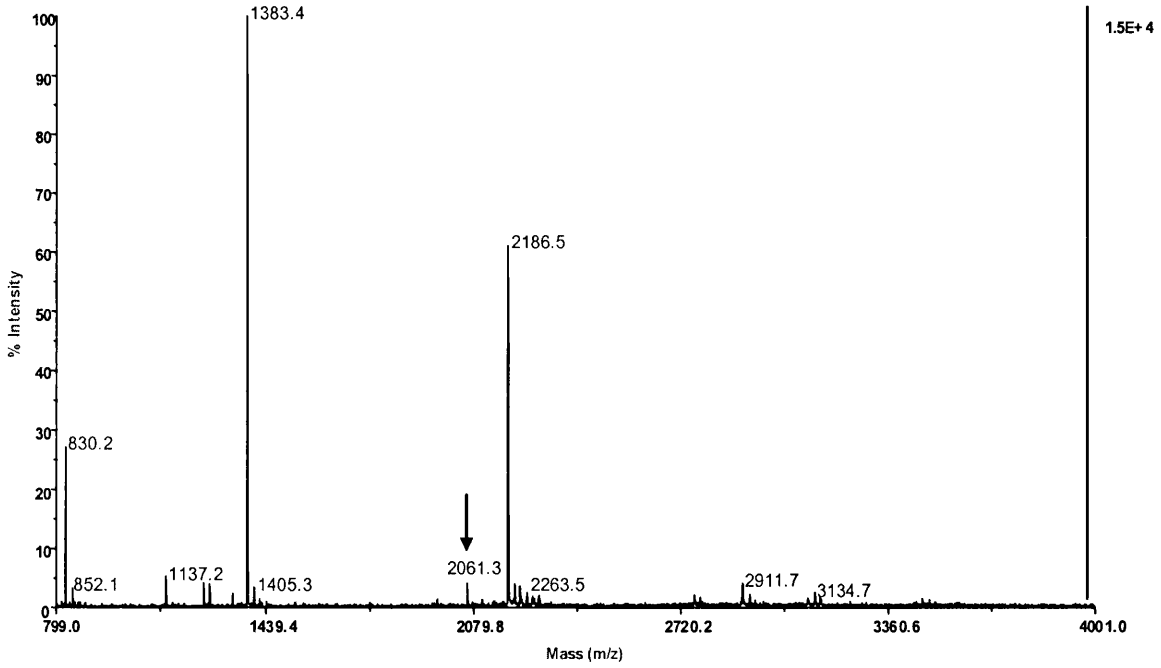


Figure 7.24 MALDI-TOF spectrum of β-casein trypsin digest (monophosphopeptide *m/z* 2061 is shown with arrow)

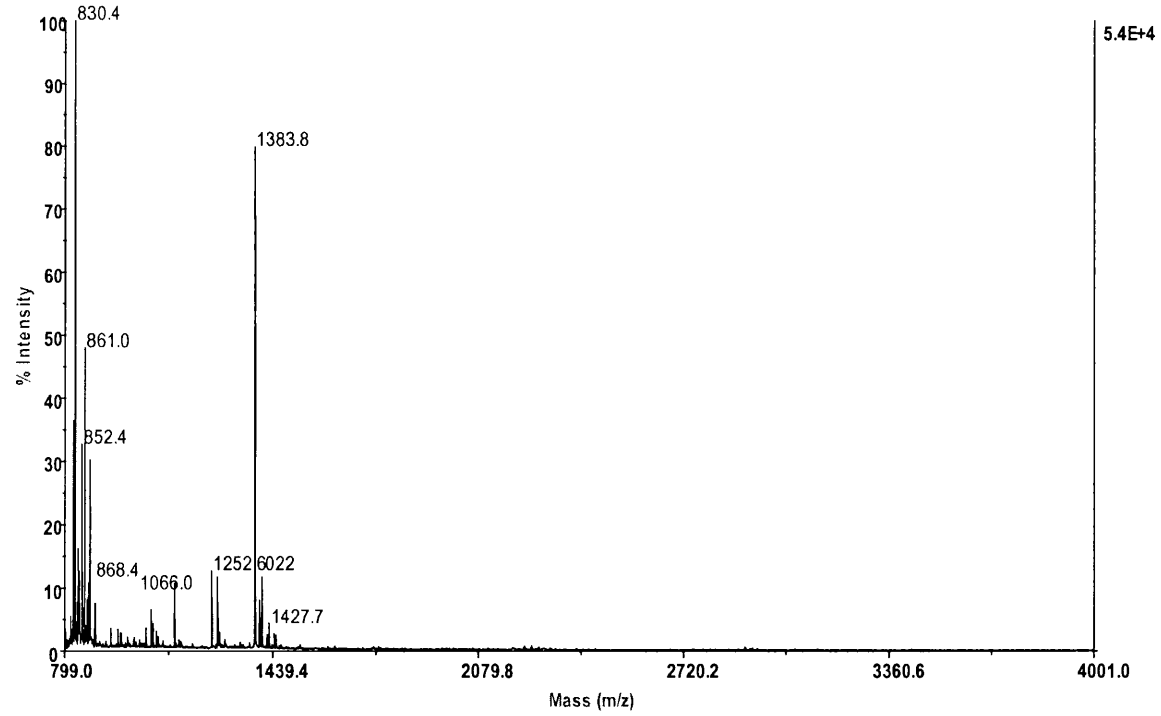


Figure 7.25 MALDI-TOF spectrum of eluent from IMAC trap column of β-casein trypsin digest as the peptides were loaded onto IMAC trap column

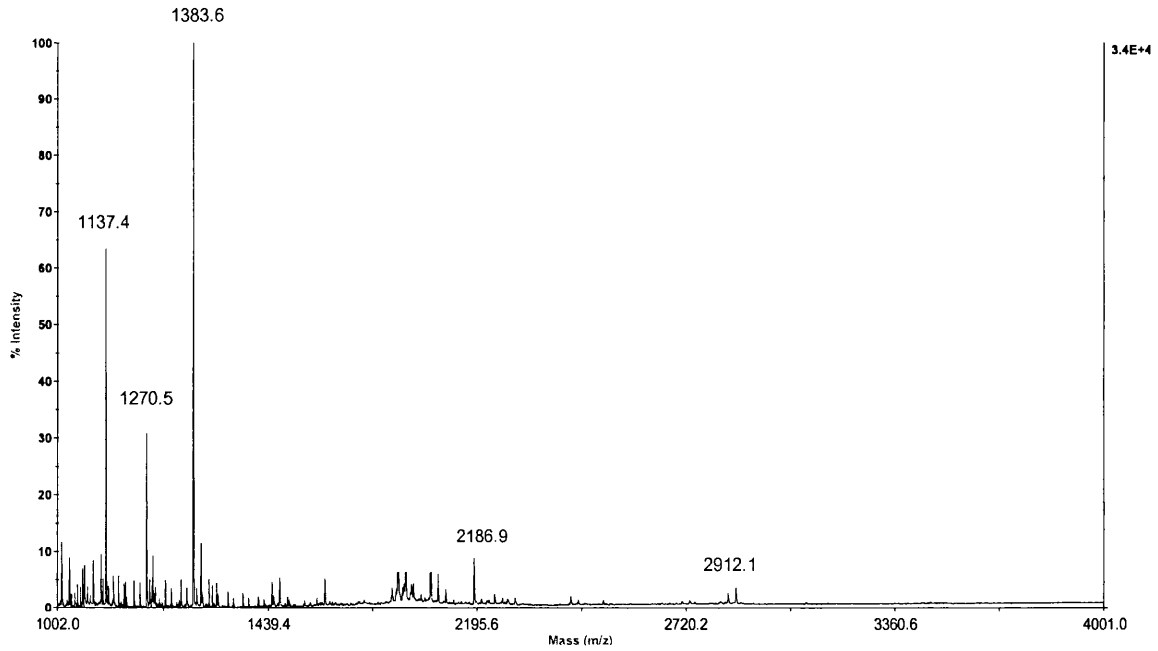


Figure 7.26 MALDI-TOF spectrum of eluent from IMAC trap column of β-casein trypsin digest as the IMAC trap column was washed by NaCl solution

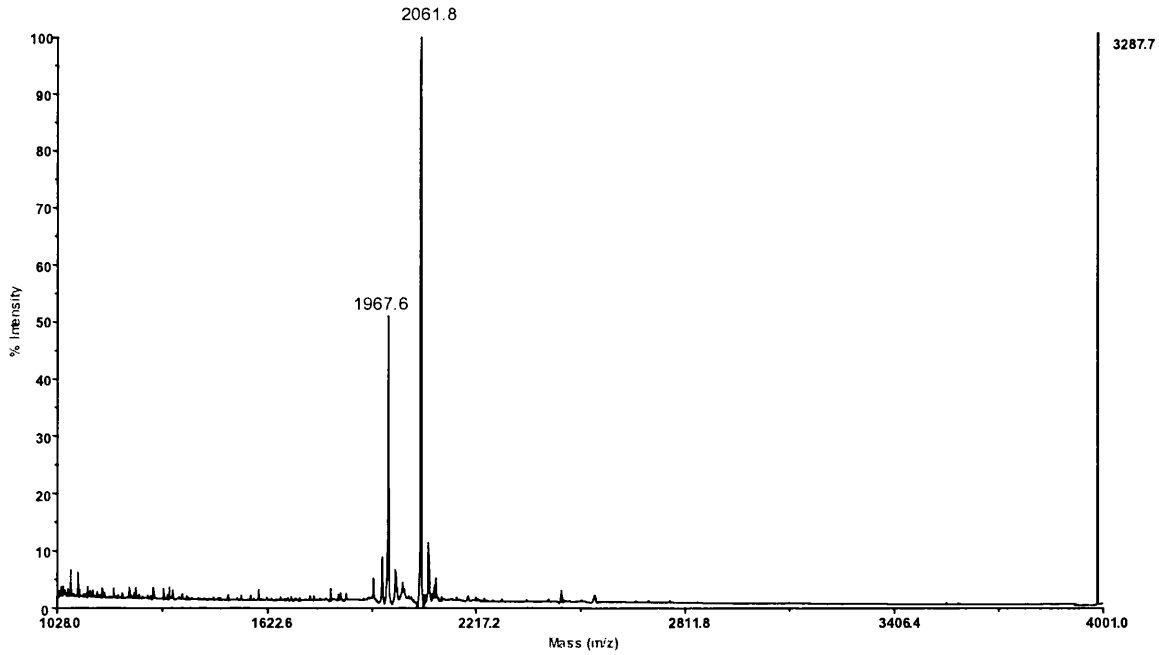


Figure 7.27 MALDI-TOF spectrum of eluent from IMAC trap column of β-casein trypsin digest as the bounded peptides were eluted with Na_2HPO_4 solution from the IMAC trap column

7.3.4 Comparison of peptide profile after IMAC enrichment by MALDI-TOF

Phosphorylation is a dynamic process, and this can be observed when comparing MALDI-TOF spectra of brain homogenate derived from different cyclic nucleotide incubation time points. Figures 7.28-7.31 show the MALDI-TOF spectra of the eluents from IMAC trap column of cCMP incubation at four different time points. The peptide with *m/z* of 1280 is most abundant in 5 minute incubation, while peptide with *m/z* 3304 is absent in 1 and 5 minutes, reaching an abundance peak at 15 minute, and is reduced in the 30 minute sample. More phosphorylated peptides can be observed with increasing incubation time.

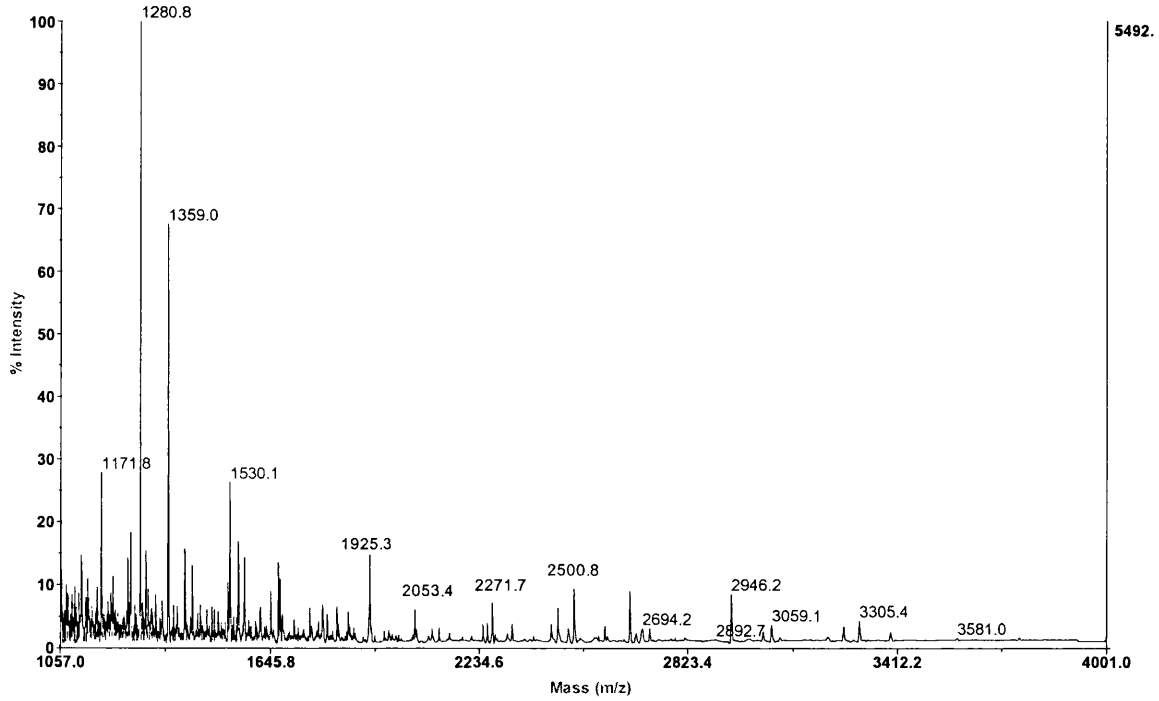


Figure 7.28 MALDI-TOF spectrum of eluent of 1 min cCMP incubation peptides from IMAC trap column

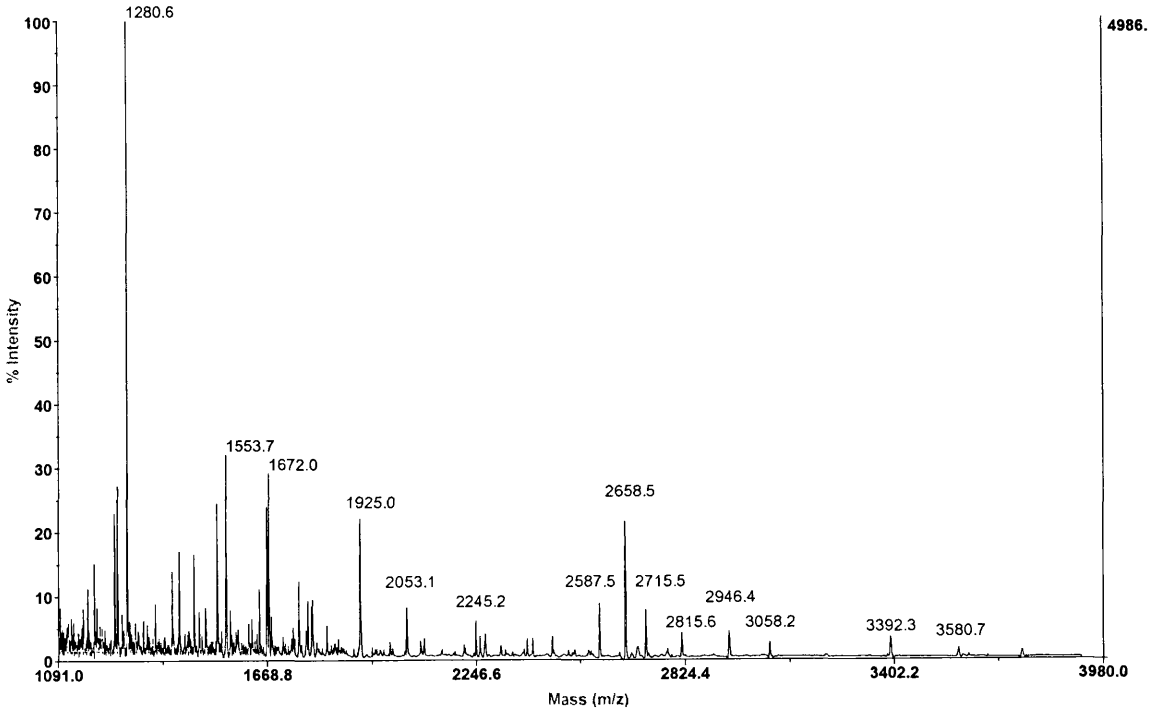


Figure 7.29 MALDI-TOF spectrum of eluent of 5 min cCMP incubation peptide after IMAC trap column

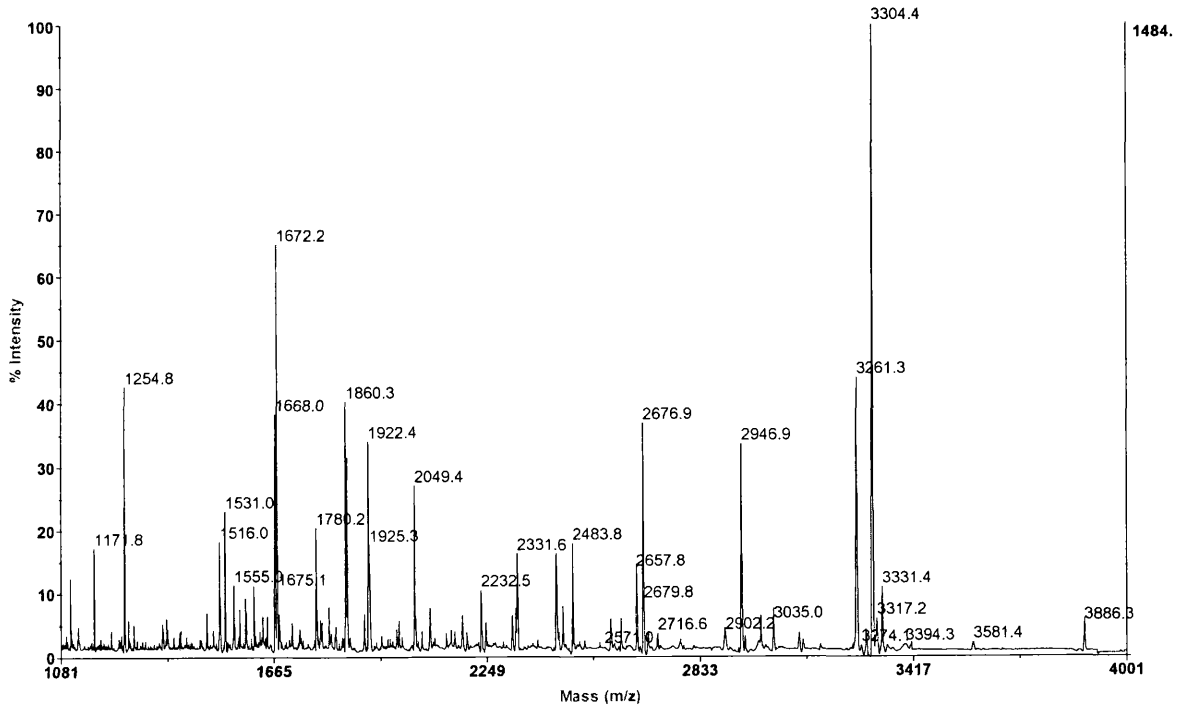


Figure 7.30 MALDI-TOF spectrum of eluent of 15 min cCMP incubation peptide after IMAC trap column

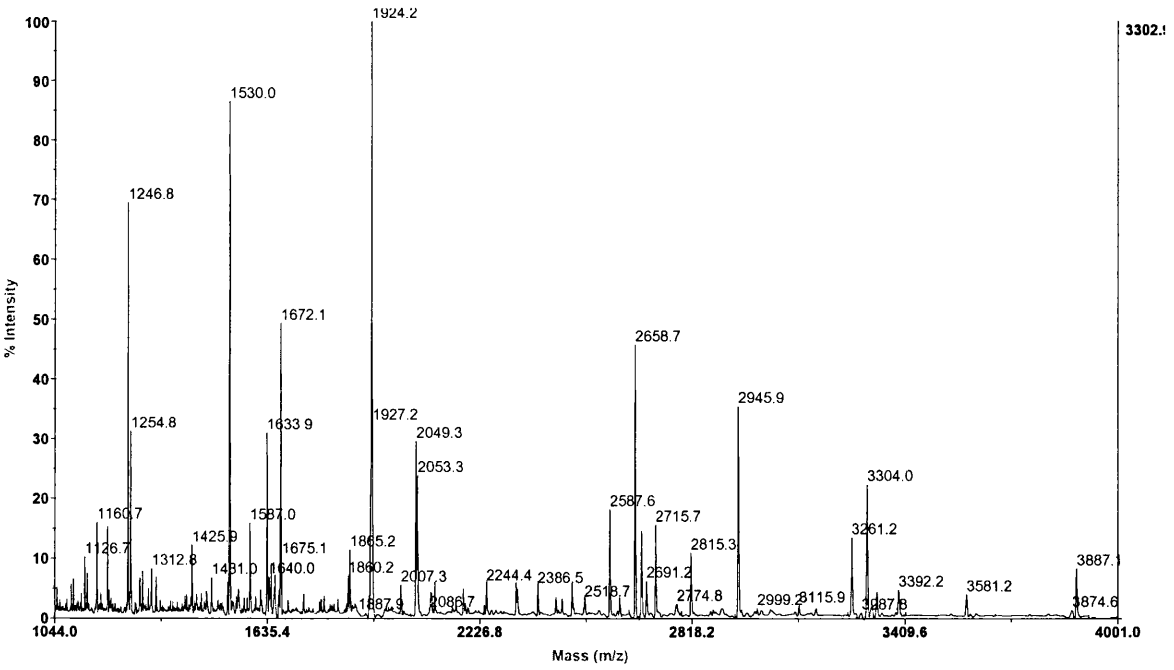


Figure 7.31 MALDI-TOF spectrum of eluent of 30 min cCMP incubation peptide after IMAC trap column

Figure 7.32 shows a typical comparison of brain homogenates, incubated with different cyclic nucleotides, by MALDI-TOF. The four incubations show similar spectra, but this does not mean phosphoproteome is closely similar for the four incubates, since the MALDI source ionises all the peptides at the same time, thus ion suppression cannot be avoided. However although in MALDI spectra only the relatively abundant peptides can be observed, there are still some differences which can be observed; for example, the peptide with *m/z* 1924 is observed at a greater relative abundance than *m/z* 1530 in cCMP than the other incubations. To identify the phosphopeptides LC/MS/MS was used, as the nano LC system allows the separation of the peptides before mass spectrometric analysis and the sensitivity is greatly increased.

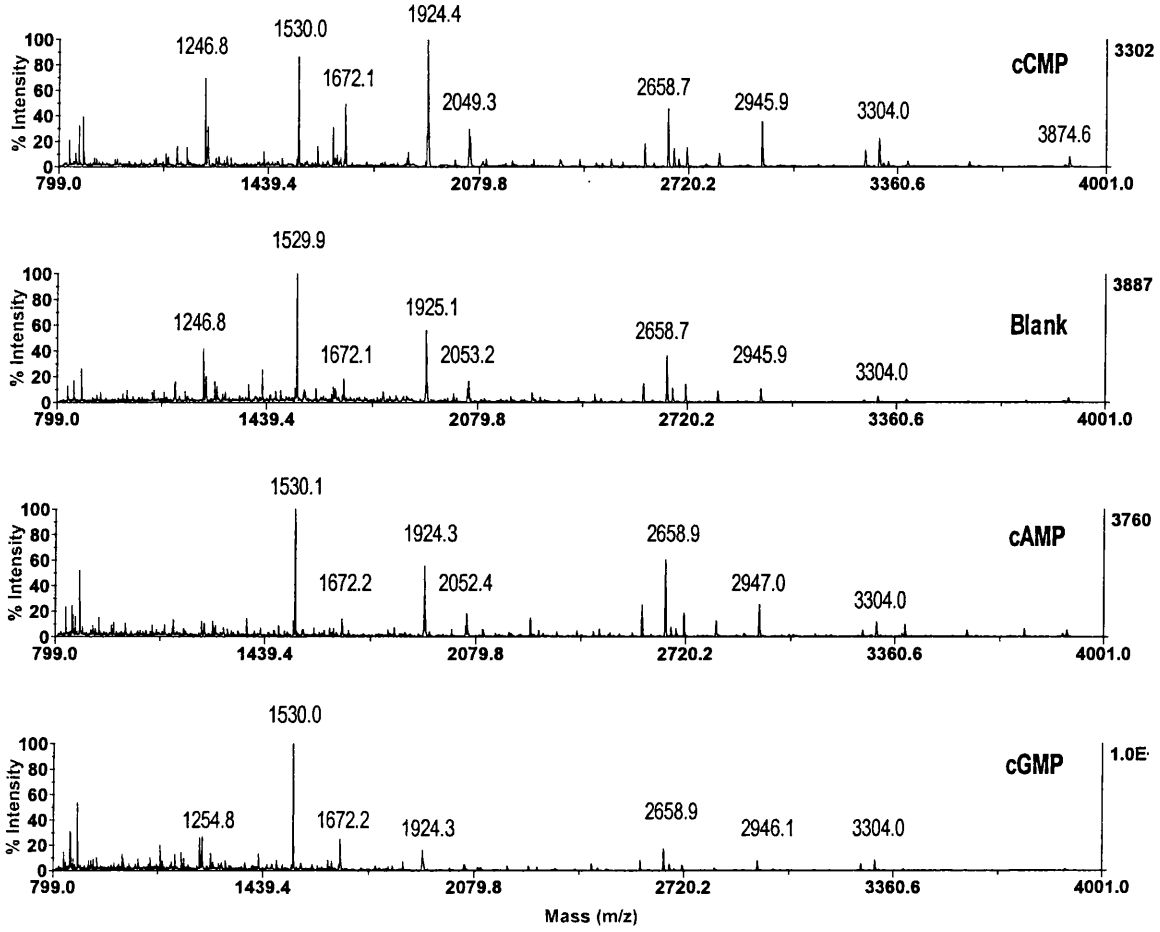


Figure 7.32 Comparison of protein phosphorylation of different cyclic nucleotide incubation (30min) by MALDI-TOF

7.3.5 Identification by online IMAC enrichment and LC/MS/MS of proteins phosphorylated in response to cCMP

Phosphorylated proteins were identified by inputting the MS and MS/MS data and searching the mouse database using SEQUEST software as described in Section 7.2.7. The software parameters were set to detect a differential modification of 80 Da on serine, threonine, and tyrosine. The filters (XC, dCn, Sp and number of peptides) shown

in Table 7.4 were applied to the database search and the assignments of phosphopeptide sequences were confirmed by comparing the acquired MS/MS spectra to the theoretical fragmentation pattern. On average, 40-50 phosphorylated proteins were identified in each time point of the cyclic nucleotide incubations by using the XC, dCn and Sp value filter criteria shown in Table 7.4, with some phosphorylated proteins present in more than one time incubation. Altogether about 120 phosphorylated proteins for each incubation were identified (data shown in Table A1-A16). Since the aim of this study is to identify phosphorylated proteins unique or increased in expression in response to cCMP incubation in the hope to elucidate the function of cCMP in signal transduction, the predicted and hypothetical proteins with unknown function are not considered and only phosphorylated proteins which exists in the cCMP incubation while apparently absent or of lower abundance from those with other cyclic nucleotides are studied.

The proteins identified by SEQUEST software were compared and the proteins unique to cCMP incubates listed. Phosphorylated proteins of the same incubation time range were compared initially. There are more phosphorylated proteins observed in the 30 mins incubation with some of these proteins being identified in cAMP, cGMP or blank incubation of incubation times other than 30 mins, these proteins were therefore not considered either. Since the samples are quite complex, there could be variations between analyses due to the information dependent acquisition of the MS/MS data which can vary from analysis to analysis, thence leading to variance in the specific peptides identified in any given analysis at any given time point. Therefore the identified phosphorylated proteins unique to cCMP were re-examined by studying the LC/MS data in order to ensure that the peptide from which the protein was identified was either unique in the cCMP incubation or the intensity of this peptide is significantly

higher in cCMP incubation compared to the other incubations. One example is shown below. The unique phosphoprotein, serologically defined colon cancer antigen 13, was identified in the 30 min cCMP incubation, the phosphopeptide matching the database is RHKGPGR*p*TGGLVISRP with the doubly charged precursor ion of *m/z* 758.24 at retention time of 89.2 min, the XIC of full scan base peak chromatogram with the *m/z* range of 758.24 of the four incubations is shown in Figure 7.33. There is strong peak at retention time 89.27 min, but for other incubations there is just background noise. Therefore, this peptide is characterized as unique to cCMP with high confidence, and serologically defined colon cancer antigen 13 was then designated as a putative cCMP-responsive protein kinase substrate.

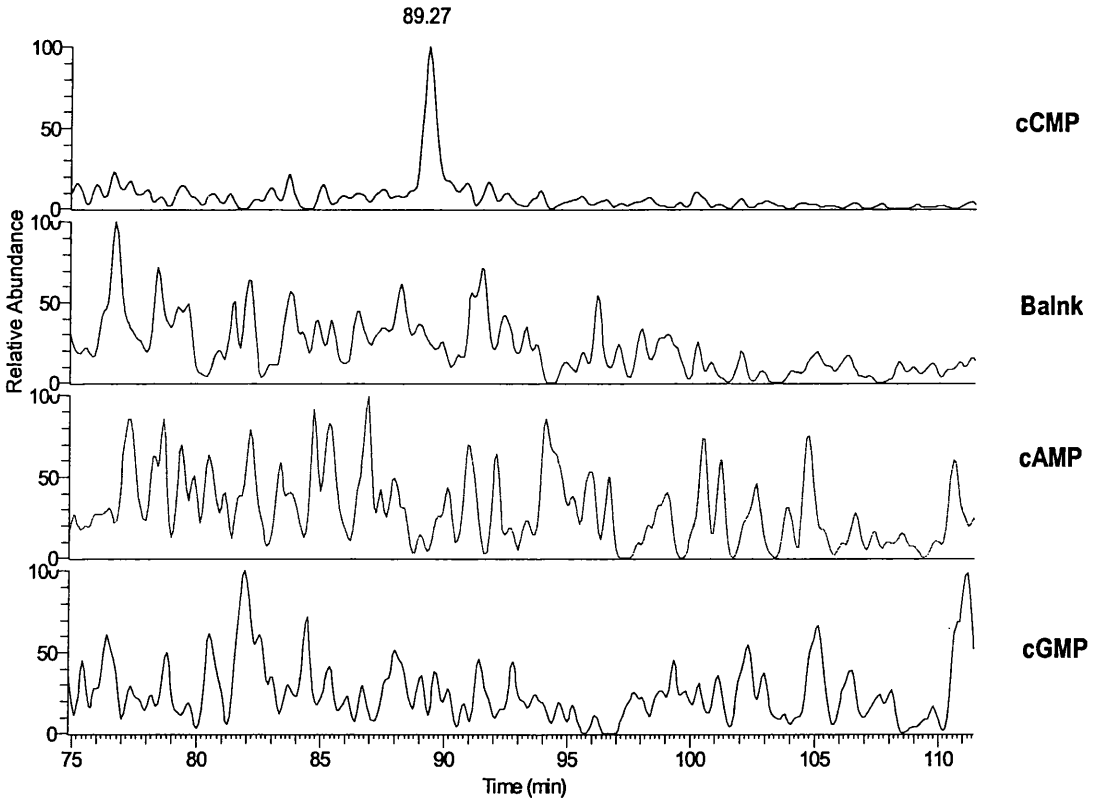


Figure 7.33 XIC of full scan base peak chromatogram at *m/z* range of 758.24 of the four cyclic nucleotide incubations using 75 μm i.d. C_{18} column with 0-45% B for 110 minutes（98%ACN with 0.1% formic acid）on LCQ Deca ion trap mass spectrometer

Another example shows the exclusion of a false positive identification. Cystathionase was identified only in the 30 min cCMP incubation, by the SEQUEST database searching. The fragmentation of doubly charged precursor ion is *m/z* 962.6 at retention time 62.6 minute matches the theoretical fragmentation phosphopeptide in the database with the sequence of KAGDEIICMDEVYGGT*p*NRY. The XIC of full scan base peak chromatogram with the *m/z* range of 962.6 of the four incubations is shown in Figure 7.34, all four cyclic nucleotide incubation have obvious peaks with similar intensity (highest in cAMP incubation) at that retention time. Therefore, this protein was proved as false identification although the peptides in cAMP, cGMP and Blank were not picked up by the mass spectrometer or the database.

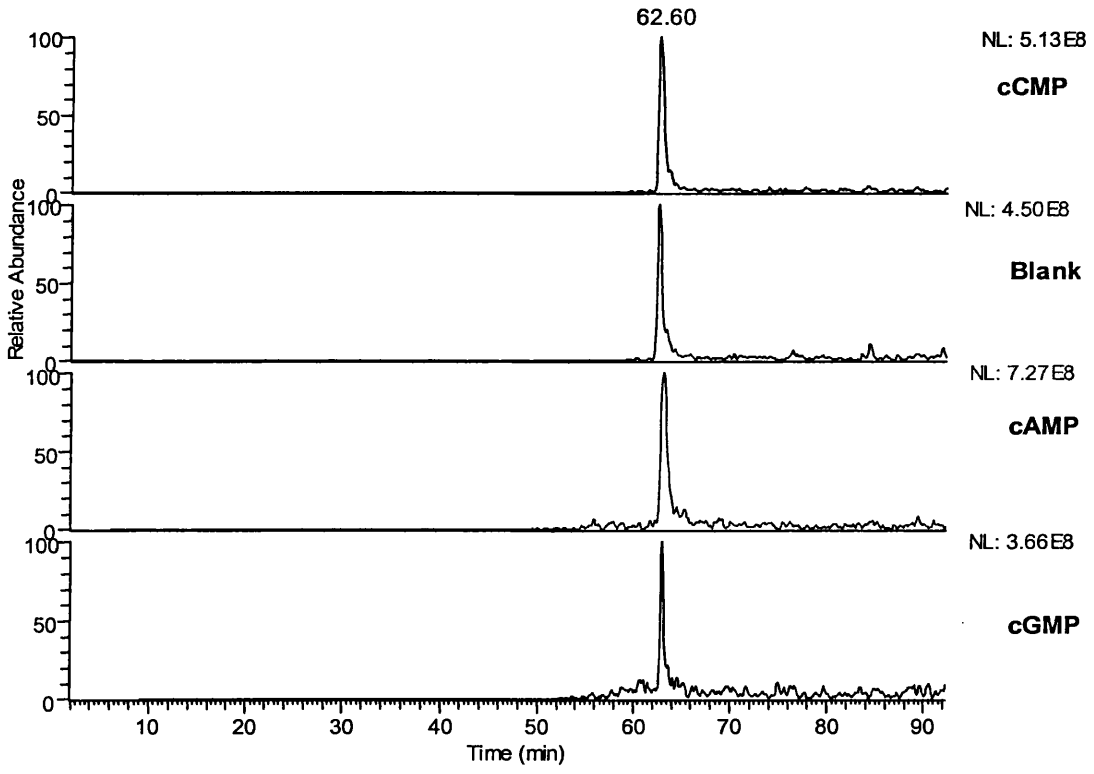


Figure 7.34 XIC of full scan base peak chromatogram at *m/z* range of 758.24 of the four cyclic nucleotide incubations using 75μm i.d. C_{18} column with 0-45% B for 110 minutes（98%ACN with 0.1% formic acid）on LCQ Deca ion trap mass spectrometer

By using the above screening methods, the proteins that are phosphorylated only because of the elevated level of cCMP can be identified. In this study, the proteins of interest are phosphorylated proteins. Peptides carrying a phosphate group can be identified by the neutral loss of H_3PO_4 from the precursor ions in the positive mode. So the screened proteins were examined again to determine whether the neutral loss of H_3PO_4 occurred, the peptides without neutral loss of H_3PO_4 were considered false identification. An example of one that fulfils this criteria is the peptide RHKGPGR*p*TGGLVISRP (serologically defined colon cancer antigen 13) described previously. Figure 7.35 shows the full scan base peak chromatogram of the IMAC-LC/MS/MS run of 30 min cCMP incubated brain homogenate trypsin-digested peptides. It displays a lot more peaks than the corresponding MALDI spectrum, and each peak is observed to consist of several peptides (see Figure 7.36).

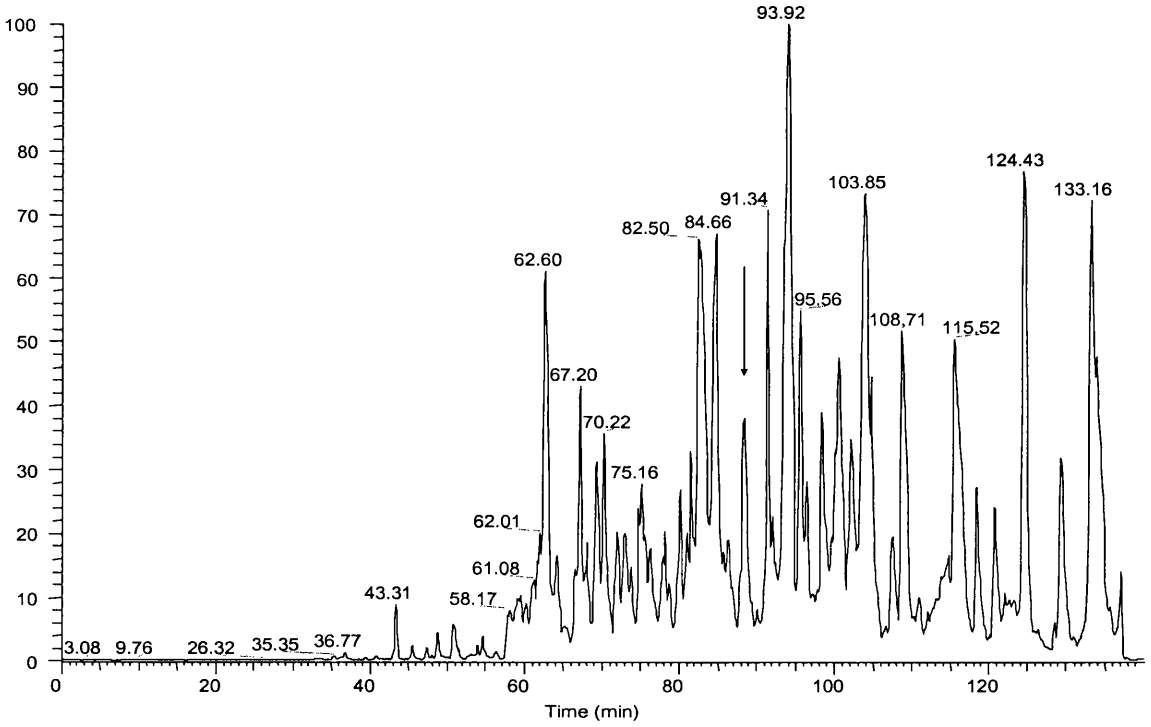


Figure 7.35 Full scan base peak chromatogram of trypsin digested peptides of brain tissue homogenate cCMP 30 min incubation with online IMAC enrichment and nano C_{18} column separation coupled to a LCQ Deca ion trap mass spectrometer

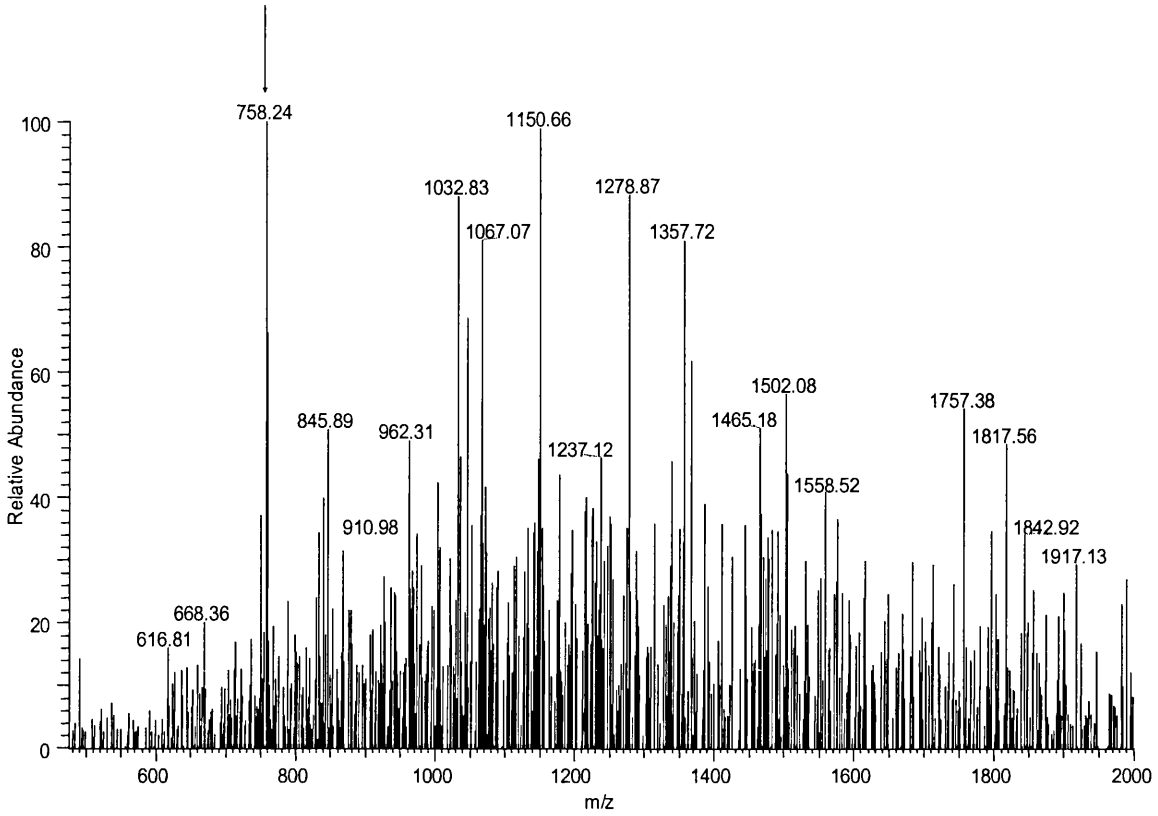


Figure 7.36 Full scan MS at retention time 89.27 min

Figure 7.36 shows full scan spectrum at retention time 89.27 min, it is still a mixture of peptides. MS/MS was then performed with the three highest abundant peptides with *m/z* of 758.24, 1150.66 and 1278.87. As shown in Figure 7.37, the peptide with *m/z* of 758.27 in Figure 7.36 was recognized as phosphopeptide with a sequence of RHKGPGR*p*TGGLVISRP. For the doubly charged peptide with a sequence RHKGPGR*p*TGGLVISRP, the major fragmentation should be the neutral loss of one phosphoric acid, and in this spectrum, the most abundant peak should be the *m/z* of precursor ion (758.27) minus 49, which is *m/z* 709.27 as shown in the spectrum (Figure 7.37).

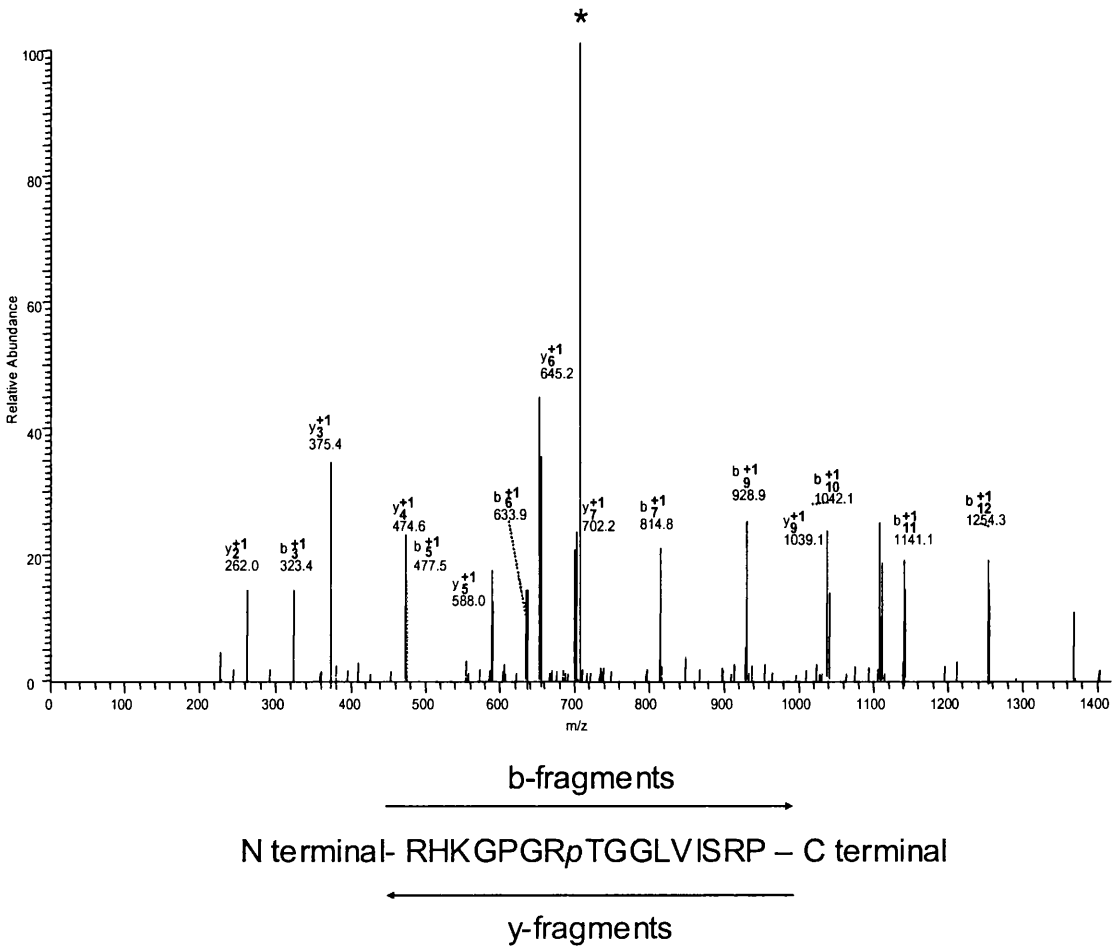


Figure 7.37 MS/MS spectrum of doubly charged peptide with sequence RHKGPGR*p*TGGLVISRP, precursor ion *m/z* 758.24, * neutral loss of 98 Da from precursor ions (MH_2^{2+}-49)

In this spectrum, 15 out of 26 ions of b and y ions of this peptide can be observed, the list of these ions is shown in Table 7.5. The phosphorylation site can be assigned to tyrosine as the mass difference between b_7 and b_6 ions is 101 amu + 80. The SEQUEST score is shown in Table 7.6, comparing with the criteria setting in Table 7.4 indicates a clear probability in identification.

Table 7.5 The *m/z* of all the b- and y- fragments generated from the MS/MS analysis of the peptide ion with *m/z* of 758.24, the *m/z* match the theoretical values of the identified sequence was shown in bold

	AA	B	Y	
1	H	138.15	-	14
2	K	266.32	1378.49	13
3	G	***323.37***	1250.31	12
4	P	420.49	1193.26	11
5	G	**477.54**	1096.15	10
6	R	**633.73**	**1039.09**	9
7	T@	**814.80**	882.91	8
8	G	871.85	**701.84**	7
9	G	**928.90**	**644.79**	6
10	L	**1042.06**	**587.73**	5
11	V	**1141.19**	**474.58**	4
12	I	**1254.35**	**375.44**	3
13	S	1341.42	**262.29**	2
14	R	-	175.21	1

Table 7.6 SEQUEST scores for *m/z* 758.24

MH+	Charge	XC	dCn	Sp	Ions
1515.63	+2	2.41	0.13	210.9	15/26

The phosphorylated proteins that are unique in cCMP incubation or higher in cCMP incubation at the different incubation times are listed in Table 7.7. The full scan spectrum comparison and MS/MS spectrum, b and y ions of the identified phosphopeptides of all these proteins in Table 7.7 except serologically defined colon cancer antigen 13 (which is discussed here as an example) are shown in Appendix 3. It can be observed that the number of phosphorylated proteins due to the inclusion of CMP increases with the incubation time. This could be explained due to the fact the blank sample is not actually blank, there are naturally occurring cCMP, cAMP and cGMP levels in the blank incubation, so at shorter incubation time, less difference among the samples may be expected as these levels are available. As time increases, the

Table 7.7 Phosphorylated proteins unique in elevated cCMP incubations from SEQUEST

Incubation	Protein	Mw	Peptide	MH^+	z	XC	dCn	Sp	Ions
1min	formin homology 2 domain containing 1	129541	RTLKSGLGDDLVQALGL*p*SKA	1896.06	2	2.36	0.18	466.0	20/34
5min	MAP-kinase activating death domain isoform 8	124489	REKT*p*TPFP*p*SLKGNRR	1635.61	2	2.23	0.32	493.2	10/24
15min	evolutionarily conserved G-patch domain containing	103007	KKED*p*SISEFL*p*SQARS	1670.56	2	2.65	0.14	651.9	16/24
	protein kinase, interferon-inducible double stranded RNA dependent	58279	KIGQ*p*TMYGTGSGVTKQEAKQ	1937.04	2	2.29	0.10	269.1	18/34
	lamin B2	67029	RMRVESL*p*S*p*YQLLGLQKQ	1926.03	2	2.35	0.49	896.2	16/27
30min	serologically defined colon cancer antigen 13	125055	RHKGPGR*p*TGGLVISRP	1515.63	2	2.41	0.13	210.9	15/26
	arrestin domain containing 1	46112	K*p*TGNVVL*p*TAS*p*TDLRG	1587.40	2	2.45	0.51	560.8	14/24
	myeloid/lymphoid or mixed lineage-leukemia translocation to 1 homolog	60785	RTLQTEDSNSEDEA*p*SFRS	1909.78	2	2.29	0.18	417.0	15/30
	low density lipoprotein receptor-related protein 1	50476	RILQEDFTCRAVNS*p*SCRA	1923.04	2	2.48	0.19	390.1	17/30
	centromere/kinetochore protein zw10 homolog	88062	KMLK*p*SL*p*SMELTVQKQ	1668.81	2	2.23	0.16	362.7	12/24
	deleted in azoospermia-like	33372	RNSLVTQDD*p*YFKDKR	1653.65	2	2.53	0.24	306.6	13/24

Mw- molecular weight (Da)
MH^+ -molecular ion of peptide identified
z-charge of the peptide
XC-cross correlation value
dCn-Delta correlation value
Sp-Preliminary score
Ions-number of experimental ions matching the theoretical ions for the peptide

cCMP naturally occurring in the brain is used up and the added cyclic nucleotide to saturation levels is responsible for the detected phosphorylation effects. This could effectively mean that the later timepoints represent the more cyclic nucleotide specific protein phosphorylations whilst in earlier timepoints these specific modifications are competing more with low level modifications caused by "naturally occurring" cyclic nucleotides.

7.3.6 Function as obtained from literature of putative cCMP-responsive protein kinase substrates

The function of the proteins apparently specifically phosphorylated in response to elevated level of cCMP can be grouped into several categories

Cell Proliferation:

- **Cell Cycle regulation**
  - lamin B2
  - myeloid/lymphoid or mixed lineage-leukaemia translocation to 1 homolog (MLL)
  - centromere/kinetochore protein zw10 homolog
  - deleted in azoospermia-like

- **Cell survival/anti-apoptosis and cell proliferation**
  - MAP-kinase activating death domain isoform 8
  - Pro-apoptotic: protein kinase, interferon-inducible double stranded RNA dependent
  - Low density lipoprotein receptor-related protein 1 (LRP1)

- **Tumour suppressors and oncogenes**
  - protein kinase, interferon-inducible double stranded RNA dependent
  - myeloid/lymphoid or mixed lineage-leukaemia translocation to 1 homolog (MLL) (leukaemia)

- **Cytoskeleton and cell organisation**
  - formin homology 2 domain containing 1
  - Low density lipoprotein receptor-related protein 1 (LRP1)

Transcription regulation:

- MAP-kinase activating death domain isoform 8
- protein kinase, interferon-inducible double stranded RNA dependent
- lamin B2 (possible via chromatin regulation)
- serologically defined colon cancer antigen 13

Cell patterning and development:

- **Homeobox and homeodomain-containing proteins and Regulators:**
  - serologically defined colon cancer antigen 13
  - myeloid/lymphoid or mixed lineage-leukaemia translocation to 1 homolog (MLL)

The above suggests that cCMP is involved in the phosphorylation of proteins that have various functions and many of these proteins are involved in cell proliferation, consistent with observations discussed earlier i.e. elevated levels of cCMP and cytidylyl cyclase activities in rapidly differentiating cells (Section 7.1.1.5.4).

7.3.6.1 Lamin B2 (Lmnb2)

Lamin B2 is a basic structural component of nuclear lamina, which are thought to provide a framework for the nuclear envelope and may also interact with chromatin[157]. The structural integrity of the lamina is strictly controlled by the cell cycle, as seen by the disintegration and formation of the nuclear envelope in prophase and telophase, respectively[158]. Increased phosphorylation of the lamins occurs before envelope disintegration and probably plays a role in regulating lamin associations.

In the cell cycle, progression through mitosis is thought to be governed by maturation promoting factor (MPF), a soluble molecule in the cytoplasm consisting of two proteins Cdk (cyclin-dependent kinase) and cyclin. When activated, MPF initiates a pathway leading to activation of a lamin-specific kinase, which will in turn phosphorylate lamin proteins of the nuclear lamina, and promote the nuclear envelope breakdown cycle[159]. More phosphorylated lamin was observed in cCMP incubation as shown in Appendix 3. The proposed cell cycle including lamin B2 with cCMP incursion is illustrated as Figure 7.38, with cCMP activating the serine / threonine kinase 7 (CDK7) which then activates MTF, the activated MPF then phosphorylating lamin B2 which promotes the nuclear membrane breakdown cycle.

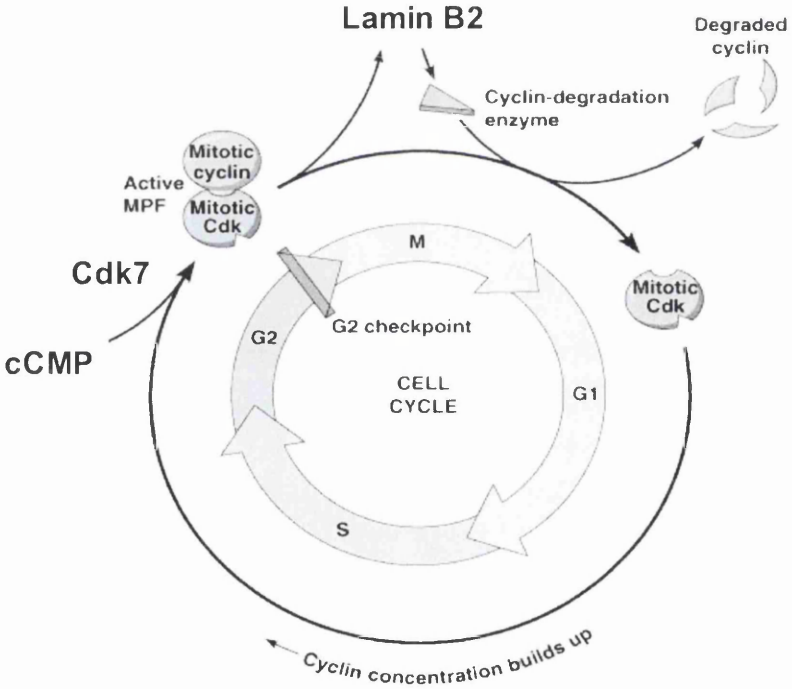


Figure 7.38 Proposed cell cycle with elevated level of cCMP

7.3.6.2 Myeloid/lymphoid or mixed lineage-leukaemia translocation to 1 homolog (MLL)

This is an oncogene protein involved in rearrangement in leukaemia[160]. The normal MLL gene plays a key role in developmental regulation of gene expression (including HOX genes). It regulates cell cycle through cyclin-dependent kinase inhibitors p27 and p18[161].

7.3.6.3 Centromere/kinetochore protein zw10 homolog (zw10)

This protein regulates the cell cycle through chromosome segregation during cell division. During the cell cycle, the zw10 protein moves from the centromere/kinetochore at prometaphase to kinetochore microtubules at metaphase, and then back to the centromere/kinetochore at anaphase which suggests that zw10 may act at the kinetochore as part of a tension-sensing checkpoint that renders anaphase onset dependent upon bipolar tension exerted across all centromeres[162].

7.3.6.4 Deleted in azoospermia-like (DAZL)

This protein is germ-cell-specific RNA-binding protein, essential for gametogenesis. It is able to stimulate translation and function at the level of translation initiation and thought to activate translationally silent mRNAs during germ cell development through the direct recruitment of PABPs (poly(A)-binding proteins)[163]. Also this protein is involved in progression into meiosis[164].

7.3.6.5 MAP-kinase activating death domain isoform 8

This protein is pro-apoptotic, via activation of MAPK (death domain-containing adaptor protein that interacts with the death domain of TNF-alpha receptor 1 to activate mitogen-activated protein kinase (MAPK) and propagate the apoptotic signal), but this isoform is pro-survival and lacks exon 16, which has pro-apoptotic domain[165].

7.3.6.6 Protein kinase, interferon-inducible double stranded RNA dependent (PKR)

Activated PKR is a tumour-suppressor and can induce apoptosis and inhibit translation. It is a regulator of diverse cellular response to stress, by working through TNF-a pathway, also it activates transcription[166].

7.3.6.7 Low density lipoprotein receptor-related protein 1 (LRP1)

This protein has a role in determining blood vessel structure and in angiogenesis, cell migration, proliferation and vascular permeability and is known to play roles in areas as diverse as lipoprotein metabolism, degradation of proteases, activation of lysosomal enzymes and cellular entry of bacterial toxins and viruses and is shown to interact with scaffolding and signalling proteins via its intracellular domain in a phosphorylation-dependent manner and to function as a co-receptor partnering with other cell surface or integral membrane proteins[167].

7.3 Conclusion

In this study, an online IMAC-nano LC/MS platform for phosphoprotein profiling was set up and phosphorylated proteins in mice brain homogenate incubated with cCMP were characterised. Phosphopeptide enrichment by immobilized metal affinity chromatography (IMAC), followed by nano liquid chromatography separation and mass spectrometry and protein database searching make it possible to identify abundant phosphorylated proteins on a proteome-wide basis. IMAC coupled to LC-MS/MS has proven to be a powerful method to identify the phosphorylated proteins present in complex mixtures containing non-phosphorylated proteins. On-line IMAC shows better loading capacity than IMAC ZipTip, other advantages of on-line IMAC system include (a) the samples can be processed automatically instead of the time-consuming ZipTip process, and (b) the avoidance of manual ZipTip process provides more reproducible result The further separation of phosphorylated peptides by nano-HPLC enables the complex mixture of phosphorylated peptides to be simplified prior to mass spectrometric analysis. The identification of the proteins via database searching (SEQUEST) was an integral part of this study and the identification of these phosphorylated proteins should provide new avenues for investigating the mechanism of cCMP signalling pathways and its precise cellular function.

It has been reported that cCMP has varied biological effects for example, the stimulation of leukaemia L-1210 cells and DNA synthesis, the inhibition of RNA synthesis as introduced in Section 7.1.1.5.4. In this study, several proteins were identified as unique (unique is described for this work as a peptide whose presence in the cCMP incubation is either unique only to cCMP or significantly stronger in cCMP compared to other

incubations) to elevated level of cCMP, and this protein phosphorylation can be linked to the biological function of cCMP. For example, strong evidence has been obtained for a causal link between phosphorylation of lamins by protein kinase C and disassembly of the nuclear lamina during mitosis[168] and it is reported that lamin B2 is hyperphosphorylated in leukemic cells than normal cells[169]. The function of each protein is grouped in Section 7.3.6. The data here support previous research implicating elevated level of cCMP a role in cell hyperproliferation in brain and other mammalian tissues. Also protein related to Leukaemia is found to be phosphorylated (myeloid/lymphoid or mixed lineage-leukaemia translocation to 1 homolog), this supports the view that cCMP is involved in the stimulation of leukaemia L – 1210 cells[58]. Proteins identified which involve in transcription regulation can be linked to the function as DNA and RNA synthesis.

Comparing four incubation timepoints, 30 min incubation has more specific phosphorylated proteins in response to cCMP, which indicates the phosphorylation was caused by the elevated level of cCMP rather than naturally occurring cCMP. For further research, the incubation time could be longer to investigate the long term effect of cCMP in cellular function.

It should be noted that there are also unique phosphorylated proteins identified in elevated level of cAMP and cGMP (as shown in Appendix 2. Table A1-16), but since the aim of this study is to identify phosphorylated proteins elicited by cCMP, no effort was made to investigate phosphoproteome of elevated level of cAMP and cGMP. Comparisons of LC/MS of the peptides identified shows the peptides are unique to cCMP incubation and are absent or lower in other incubations, which are evidences that

the phosphorylation was caused by elevated level of cCMP.

There are possibilities of effect on phosphatases by cCMP. Contrary to protein kinases, which phosphorylate the substrate, phosphatases reverse the phosphorylation, and this is a dynamic balance process. Further experiments could include phosphatase assays, for example with heavy isotope labelled phosphorylated proteins, to determine the activity of these enzymes and how significant their role is in dephosphorylating the proteins during the purification procedures. If the activity of these enzymes is determined to be significant, then the inclusion of phosphatase inhibitors such as calyculin A during the extraction procedure should be considered. However, if the activity of the phosphatases is low, it could be conceived that the protease inhibitors added during the extraction procedure could also be acting as inhibitors of esterases, thus inhibiting the phosphatase.

Although IMAC was utilised in this study in the purification of phosphorylated proteins from the sample mixture, it is not an entirely selective technique. A previous study[83] shows IMAC columns were also retaining proteins with GTP or GDP bound to them and therefore the response to cCMP could be GTP or GDP binding instead of a cCMP-responsive protein kinase effect, or the effect could be the phosphorylation of the GTP bond proteins.

There is also the possibility that cCMP inhibits the breakdown or even stimulates the synthesis of the identified proteins, and that their phosphorylation is via another non cCMP-responsive mechanism, or indeed that cCMP facilitates the release of these proteins from the membrane, or that the phosphorylation of proteins might be the indirect result of an accelerated cell cycle, the acceleration being the direct response to

cCMP. Many indirect effects are plausible, for example, the phosphorylation of MAP-kinase activating death domain isoform 8 might be explained as the result of cCMP blocking effects of cAMP as a competitive inhibitor, e.g. inhibiting PDE, weak agonist of PKA, PKG or regulative feedback to their cyclases.

In this study, by utilizing IMAC affinity chromatography, only the phosphorylated proteins are characterized. It cannot provide sufficient information to define the processes whereby the phosphorylation of target proteins is elevated in response to cCMP. The phosphorylation and dephosphorylation of the indicated proteins was searched in literature. However, for many of the proteins identified, there is no literature indicating whether phosphorylation activates the protein or inactivates it (except lamin B2, which is described in Section 7.3.6.1). To find out this, one possible choice is to extract the active protein using a phospho-specific antibody and to check phosphorylation state.

Further study should involve the comparison of whole protein identification (without IMAC enrichment) between incubations with different cyclic nucleotides. For such a complex brain homogenate, the sample separation remains a key factor of successful identification and comparison among different incubations. Future experiments could use simpler tissue or simpler organism and compare with the brain tissue. As a trial experiment, gel electrophoresis separation of whole brain homogenate proteins was evaluated in this study and some protocols and results are shown in the appendix which might be a good starting point for further research.

In order to determine whether the phosphorylations of proteins observed here are the

result of direct cCMP-responsive protein kinase activity or the end product of an indirect mechanism, definitive evidence could only be produced by using the purified cCMP-responsive protein kinase in incubation with the purified putative substrate proteins identified here.

In conclusion, in this study, the proteins in Table 7.7 were characterized as putative cCMP-responsive protein kinase substrates, because the phosphorylation could be also possibly due to one such response which leads to initiation or acceleration of chain reactions/cycles, which contain phosphorylation steps. Of the identified proteins, as can be observed from Figure 7.33 and Appendix 3, only two of the 11 proteins are shown to be unique to cCMP incubation while for the other 9 proteins, the corresponding peptide is higher in concentration due to elevated cCMP, and this might be because of the naturally occurring cCMP also present in all the other incubations. The Rab 23 reported by Bond[83] as putative cCMP-responsive protein kinase substrate was not found in this study, this could because different brain extracts (different mice, protease inhibitors when dialysis) and different IMAC processes were used (online IMAC), and it could be the phosphorylated protein is of very low level in the whole brain tissue extract.

7.5 References

1. Sutherland, E.W. and Rall, T.W., J. Biol. Chem., **232**, 1077 (1958)
2. Stryer, L., *Biochemistry*, Freeman, New York (1999)
3. Ts'O, P.O.P, *Basic Principles in Nucleic Acid Chemistry*, Academic Press, New york (1974)
4. Sundralingham, M., Ann. N.Y. Acad. Sci., **255**, 3 (1975)
5. Newton, R.P. and Smith, C.J., Phytochemistry, **65**, 2423 (2004)
6. Laukens, K., Roef, L., Witters, E., Slegers, H. and van Onckelen, H., FEBS Letters, 508, 75 (2001)
7. Newton, R.P., 'Cyclic Nucleotides' in *Cell Biology Labfax*, Bios Scientific Publishers, (1992)
8. Peterkofsky, A., Advances in Cyclic Nucleotide Research, **7**, 1 (1976)
9. Newton, R.P., Roef, L., Witters, E. and van Onckelen, H., New Phytologist, **143**, 427 (1999)
10. Sutherland, E.W., Rall, T.W. and Menon, T., J. Biol.Chem., **237**, 1220 (1971)
11. Hardie, D.G., *Biochemical Messengers, Hormones Neurotransmitters and Growth Factors*, Chapman and Hall, New York (1994)
12. Graves, J.D. and Krebs, E.G., Pharmacol. Ther., **82**, 111 (1999)
13. Sabie, F.T. and Gadd, G.M., Mycopathologia, **119**, 147 (1992)
14. Krebs, E.G. and Beavo, J.A., Annu Rev. Biochemistry, **48**, 923 (1979)
15. Makman, R.S. and Sutherland, E.W., J. Biol. Chem., **240**, 1309 (1965)
16. Rickenberg, H.V., Annu Rev. Microbiology, **28**, 353 (1974)
17. Hakeem, N.S., Thesis PhD, University of Wales Swansea, (1987)
18. Jost, J.P. and Rickenberg, H.V., Annu. Rev. Biochemistry, **40**, 741 (1971)

19. Exton, J.H., Adv. Cyclic Nucleotide Res., **12**, 319 (1980)

20. Buchwald, M. and Rioardan, J.R., Adv. Cyclic Nucleotide Res., **12**, 243 (1980)

21. Park, C.W., Adv. Cyclic Nucleotide Res., **12**, 181 (1980)

22. Conolly, M.E., Adv. Cyclic Nucleotide Res., **12**, 151 (1980)

23. Broadus, A.E., Adv. Cyclic Nucleotide Res., **12**, 509 (1977)

24. Ryan, W.L. and Heidrick, M.L., Adv. Cyclic Nucleotide Res., **4**, 81 (1974)

25. Boynton, A.L. and Whitfield, J.F., Adv. Cyclic Nucleotide Res., **15**, 193 (1983)

26. Purves, W.K., Sadava, D., Orian, G.H. and Heller, H.C., *Life: The Science of Biology*, 6^{th}, ed. Sinauer Associates, Inc. (2001)

27. Hidaka, H. and Endo, T., Adv. Cyclic Nucleotide Protein Phosphorylation Res., **16**, 245 (1984)

28. Nicholson, C.D., Challiss, R.A. and Shahid, M., Trend Pharmacol. Sci., **12**, 19 (1991)

29. Beavo, J.A., Physiol. Rev., **75**, 725 (1995)

30. Boswell-Smith, V., Spina, D. and Page, C.P., British J. Pharmacol., **147**, S252 (2006)

31. Hanks, S.K., Quinn, A.M. and Hunter, T., Science, **241**, 42 (1988)

32. Fisher, E.H. and Krebs, E.G., J. Biol. Chem., **216**, 121 (1955)

33. Mukherji, M., Expert Rev. Proteomics, **2**, 117 (2005)

34. Francis, S.H. and Corbin, J.D., Annu. Rev. Physiol., **56**, 237 (1994)

35. Nimmo, H.G. and Cohen, P., Adv. Cyclic Nucleotide Res., **8**, 145 (1977)

36. Murphy, B.J. and Scott, J.D., Trend in Cardiovascular Medicine, **8**, 89 (1998)

37. Livesey, S.A., Kemp, B.E., Re, C.A., Partridge, N.C. and Martin, T.J., J. Biol. Chem., **257**, 983 (1982)

38. Bradley, J., Li, J., Davidson, N., Lester, H.A. and Ziu, K., Proc. Natl. Acad. Sci. USA, **91**, 8890 (1994)

39. De Rooij, J., Zwartkruis, E.J.T. and Verheijen, M.H.G., Nature, 396, 474 (1998)

40. Ashman, D.F., Lipton, R., Melicow, M.M. and Price, T.D., Biochem. Biophy. Res. Commun., **11**, 330 (1963)

41. Goldberg, N.D., Dietz, S.B. and O'Toole, A.G., J. Biol. Chem., **244**, 4458 (1969)

42. Richard, H., Das, S., Smith, C.J., Pereira, L., Geisbrecht, A., Devitt, N.J., Game, D.E., van Geyschem, J., Brenton, A.G. and Newton, R.P., Photochemistry, **61**, 531 (2002)

43. Peterkofsky, A., Adv. Cyclic Nucleotide Res., **7**, 1 (1976)

44. Friedman, D.L., Johnson, R.A. and Zeiling, C.E., Adv. in Cyclic Nucleotide Res., **7**, 69 (1976)

45. Kuo, J.F., Kuo, W.N., Shoji, M., Davies, C.W., Seery, V.L. and Donelly, T.E., J. Biol. Chem., **251**, 1759 (1976)

46. Langridge, J.I., PhD Thesis, University of Wales, Swansea (1993)

47. Kimura, H. and Murad, F., J. Biol. Chem., **249**, 6910 (1974)

48. Waldman, S.A., Rapoport, R.M. and Murad, F., J. Biol. Chem., **259**, 14332 (1984)

49. Kimura, H., Mittal, C.K. and Murad, F., J. Biol. Chem., **250**, 8016 (1975)

50. Harrison, S.A., Reifsnyder, D.M., Gallis, B., Cadd, C.G. and Beavao, J.A., Molecular Pharmacology, **29**, 506 (1986)

51. Grant, P.G., Mannarino, A.F., Colman, R.W., Proc. Natl. Acad. Sci. USA, **85**, 9071 (1988)

52. de Bold, A.J., Borenstein, H.B., Veress, A.T. and Sonnenberg, H., Life Sci., **28**, 89 (1981)

53. de Bold, A.J., Science, **230**, 767 (1980)

54. Kuo, J.F. and Greengard, P., J. Biol. Chem., **245**, 2493 (1970)

55. Lincoln, T.M., Dills, W.L. and Corbin, J.D., J. Biol. Chem., **252**, 4269 (1977)

56. Gill, G.N., Holdy, K.E., Walton, G.M. and Kanstein, C.B., Proc. Natl. Acad. Sci. USA, **73**, 3918 (1976)

57. Francis, S.H. and Corbin, J.D., Crit. Rev. Clin. Lab Sci., **36**, 275 (1999)

58. Bloch, A. , Adv. Cyclic Nucleotide Res., **5**, 331 (1975)

59. Wikberg, J.E.S. and Wingren, G.B., Acta Pharmacol. Toxicol., **49**, 52 (1981)

60. Wikberg, J.E.S., Wingren, G.B. and Anderson, R.G., Acta Pharmacol. Toxicol., **49**, 452 (1981)

61. Cailla, H.L., Roux, D., Delaage, M. and Goridis, C., Biochem. Biophys. Res. Commun., **85**, 1503 (1978)

62. Murphy, B.E. and Stone, J.E., Biochem. Biophys. Res. Commun., **89**, 122 (1979)

63. Sato, T., Kuninaka, A. and Yoshino, H., Anal. Biochem., **123**, 208 (1982)

64. Yamamoto, I., Takai, T. and Tsuji, J., Immunopharmacol., **4**, 331 (1982)

65. Hierowski, M.T., Waring, A.J. and Schally, A.V., Biochim. Biophys. ACTA, **675**, 323 (1981)

66. Ochi, Y., Hosada, S., Hachiya, T., Yoshimura, M., Miyazaki, T. and Kajita, Y., Acta Endocrinol, **98**, 62 (1981)

67. Scavennec, J., Carcassorine, Y., Gastaut, J., Blanc, A. and Callia, H.L., Cancer Res., **41**, 3222 (1981)

68. Newton, R.P., Salvage, B.J. and Salih, S.G., Biochem. Soc. Trans., **11**, 354 (1983)

69. Newton, R.P., Salih, S.G., Salvage, B.J. and Kingston, E.E., Biochem. J., **221**, 665 (1984)

70. Cech, S.Y. and Ignarro, L.J., Biochem. Biophys. Res. Commun., **80**, 119 (1978)

71. Ignarro, L.J., Science, **203**, 672 (1979)

72. Gaion, R.M. and Krishna, G., Biochem. Biophys. Res. Commun., **86**, 101 (1979)

73. Newton, R.P., Hakeem, N.A., Salvage, B.J., Wassenaar, G. and Kingston, E.E., Rapid Commun. Mass Spectrom., **2**, 118 (1988)

74. Newton, R.P., Salvage, B.J. and Hakeem, N.A., Biochem. J., **265**, 581 (1990)

75. Newton, R.P., Evans, A.E., van Geyschem, J., Diffley, P.J., Hassan, H.G., Hakeem, N.A., Moyse, C.D., Cooke, R. and Salvage, B.J., J. Immunoassay, **15**, 317 (1994)

76. Salvage, B.J., PhD Thesis, University of Wales, Swansea (1985)

77. Newton, R.P. and Salih, S.G., Int. J. Biochem., **18**, 743(1986)

78. Kuo, J.F., Brackett, N.L., Shoji, M. and Tse, J., J. Biol. Chem., **253**, 2518 (1978)

79. Helfman, D.M. and Kuo, J.F., J. Biol. Chem., **257**, 1044 (1982)

80. Helfman, D.M., Shoji, M. and Kuo, J.F., J. Biol. Chem., **256**, 6327 (1981)

81. Newton, R.P. and Salih, S.G., Adv. Cyclic Nucleotide Protein Phosphorylation Res., **54**, 17a (1984)

82. Khan, J.A., PhD Thesis, University of Wales Swansea (1990)

83. Bond, A.E., PhD Thesis, University of Wales Swansea (2005)

84. Salih, S.G., PhD Thesis, University of Wales Swansea (1984)

85. Newton, R.P., Khan, J.A., Brenton, A.G., Langridge, J.I., Harris, F.M. and Walton, T.J., Rapid Commun. Mass Spectrom., **6**, 601 (1993)

86. Newton, R.P., Nucleosides Nucleotides, **14**, 743 (1995)

87. Walsh, D.A., Perkins, J.P. and Krebs, E.G., J. Biol. Chem., 243, 3763 (1968)

88. Chambers, D.A., J. Cell Biol., **67**, 60a (1975)

89. Pisarev, M.A. and Pisarev, D.L.K., Acta Endocrinol., **84**, 297 (1977)

90. Khachatryan, G.S., Galstyan, G.G., Antonyan, A.A., Alaverdyan, A.A., Khachatryan, V.G, Ninasyant, R.S., Vagradyan, A.G. and Admyan, M.K., J. Nerrokhimia, **6**, 552 (1987)

91. Cheng, Y.C. and Bloch, A., J. Biol. Chem., **253**, 2522 (1978)

92. Newton, R.P., Kingston, E.E., Hakeem, N.A., Salih, S.G., Beynon, J.H. and Moyse, C.D., Biochem. J., **236**, 431 (1986)

93. The International Human Genome Mapping Consortium, Nature, **409,** 934 (2001)

94. Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J. and Sutton, G.G., Science, **291,** 1304 (2001)

95. Nair, K.S., Jaleel, A., Asmann, Y.W., Short, K.R. and Raghavakaimal, S., Am. J. Physiol. Endocrinol. Metab., **286,** E863 (2004)

96. Wasinger, V.C., Cordwell, S.J., Cerpa-Poljak, A., Yan, J.X., Gooley, A.A., Wilkins, M.R., Duncan, M.W., Harris, R., Williams, K.L. and Humphery-Smith, I., Electrophoresis, **16,** 1090 (1995)

97. Aebersold, R. and Goodlett, D.R., Chem. Rev., **101,** 269 (2001)

98. Strohman, R., Biotechnology **12,** 156 (1994)

99. Wilkins, M.R., Sanchez, J.C., Willians, K.L. and Hochstrasser, D.F., Electrophoresis, **17,** 830 (1996)

100. Wilkins, M.R., Gasteiger, E., Goodey, A.A., Herbert, B.R., Molloy, M.P., Binz, P.A., Ou, K., Sanchez, J.C., Bairoch, A., Williams, K.L. and Hochstrasser, D.F., J. Mol. Biol., **289,** 645 (1999)

101. Yates, J.R., J. Mass Spectrom., **33,** 1 (1998)

102. Krishna, R.G. and Wold, F., Adv. Enzymol. Relat. Areas Mol. Biol., **67,** 265 (1993)

103. Mann, M. and Jenson, O.N., Nat. Biotech., **21,** 255 (2003)

104. Knight, Z.A., Schilling, B., Row, R.H., Kenski, D.M., Gibson, B.W. and Shokat, K.M., Nat. Biotech., **21,** 1047 (2003)

105. Krebs, E.G., Trends Biol. Sci., **19,** 439 (1994)

106. Faux, M.C. and Scott, J.D., Trends Biol. Sci., **21,** 312 (1996)

107. Kaufmann, B.J.E. and Fussenegger, M., Proteomics, **1,** 194 (2001)

108. Adamczyk, M., Gebler, J.C. and Wu, J., Rapid Commun. Mass Spectrom., **15,** 1481 (2001)

109. Shevchenko, A., Jenson, O.N., Podetelejnikov, A.V. and Sagliocoo, F., Proc. Natl. Acad. Sci. USA, **93,** 14440 (1996)

110. Kettman, J.R., Frey, J.R. and Lefkovits, I., Biomol. Eng., **18,** 207 (2001)

111. Wolters, D.A., Washburn, M.P. and Yates, J.R., Anal. Chem., **73**, 5683 (2001)

112. Walcher, W., Oberacher, H., Troiani, S., Holzl, G., Oefner, P., Zolla, L. and Huber, C.G., J. Chromatogr. B, **782,** 111 (2002)

113. Smithies, O., Poulik, M.D., Nature, **177**, 1033 (1956)

114. Laemmli, U.K., Nature, **227**, 680 (1970)

115. Klose, J., Humangenetik, **26,** 231 (1975)

116. O'Farrell, P.H., J. Biol. Chem., **250,** 4007 (1975)

117. Righetti, P.G., *Immobilized pH Gradients: Theory and Methodology*, Elsevier, Amsterdam, (1990)

118. Rothe, G.M., Maurer, W.D., *Gel Electrophoresis of Proteins*, Wright, Bristol, (1986)

119. Andrews, A.T., *Electrophoresis: Theory, Techniques and Biochemical and Clinical applications*, Clarendon Press, Oxford, (1986)

120. Klose, J. and Kobalz, U., Electrophoresis, **16**, 1034 (1995)

121. Lecchi, P., Gupte, A.R., Perez, R.E., Stockert, L.V. and Abramson, F.P., J. Biochem. Biophys. Methods, **56,** 141 (2003)

122. Gypi, S.P., Corthals, G.L., Zhang, Y., Rochon, Y. and Aebersold, R., Proc. Natl. Acad. Sci. USA, **97**, 9390 (2000)

123. Liu, H., Lin, D. and Yates, J.R., BioTechniques, **32**, 898 (2002)

124. Shi, Y., Xiang, R., Horvath, C. and Wilkins, J.A., J. Chromatogr. A, **1053**, 27 (2004)

125. Steen, H. and Mann, M., Nat. Rev., **5**, 699 (2004)

126. Shen, Y., Zhao, R., Berger, S.J., Anderson, G.A., Rodriguez, N. and Smith, R.D., Anal. Chem., **74**, 4235 (2002)

127. Oosterkamp, A.J., Gelpi, E. and Abian, J., J. Mass Spectrom., **33**, 976 (1998)

128. LC Packings, Application note 513, (2002)

129. Fenn, J., Mann, M., Meng, C.K., Wong, S.F. and Whitehouse, C.M., Science, **246**, 64 (1989)

130. Tanaka, K., Ido,Y., Akita, S., Yoshida, Y. and Yoshida T., Proc. Second Japan Symp. Mass Spectrometry, Osaka, Japan, 185 (1987)

131. Chaurand, P., Schwartz, S.A. and Caprioli, R.M., Curr. Opin. Chem. Biol., **6**,

676 (2002).

132. Xu, B. J., Caprioli, R. M., Sanders, M. E. and Jensen, R. A., J. Am. Soc. Mass Spectrom., **13**, 1292 (2002)

133. Chaurand, P., Schwartz, S.A. and Caprioli, R.M., J. Proteome Res., **3**, 245 (2004)

134. Taylor, J.A. and Johnson, R.S., Rapid Commun. Mass Spectrom., **11**, 1067 (1997)

135. Yates, J.R., Speicher, S., Griffin, P.R. and Hunkapiller, T., Anal. Biochem., **214**, 397 (1993)

136. Eng. J.K., McCormack, A.L., Yates, J.R., J. Am. Soc. Mass Spec., **5**, 976 (1994)

137. Johnson, G.L. and Lapadat, R., Science, **298**, 1911 (2002)

138. Gu, H. and Neel, B.G., Trends Cell Biol., **13**, 122 (2003)

139. Zhang, H., Zha, X., Tan, Y., Hornbeck, P.V., Mastrangelo, A.J., Alessi, D.R., Polakiewicz, R.D. and Comb, M.J., J. Biol. Chem., **277**, 39379 (2002)

140. Kalume, D.E., Molina, H. and Pandey, A., Current Opinion in Chem. Biology, **7**, 64 (2003)

141. Patterson, S.D. and Aebersold, R.H., Nat. Genet., **33**, 311 (2003)

142. Areces, L.B., Matafora, V. and Bachi, A., Eur. J. Mass Spectrom., **10**, 383 (2004)

143. Huddleston, M.J., Annan, R.S., Bean, M.F. and Carr, S.A., J. Am. Soc. Mass Spectrom., **4**, 710 (1993)

144. Nuwaysir, L.M. and Stults, J.T., J. Am. Soc. Mass Spectrom., **4**, 462 (1993)

145. Watts, J.D., Affolter, M., Krebs, D.L., Wange, R.L., Samelson, L.E. and Aebersold, R., J. Biol. Chem., **269**, 29520 (1994)

146. Vladka, G.P. and Menart, V., J. Biochem. Biophys. Methods, **49**, 335 (2001)

147. Olcott, M.C., Bradley, M.L. and Haley, B.E., Biochemistry, **33**, 11935 (1994)

148. Vladka, G.P. and Menart, V., J. Biochem. Biophys. Methods, **49**, 335 (2001)

149. Ficarro, S.B., McCleland, M.L., Stukenberg, P.T., Burke, D.J., Ross, M.M., Shabanowitz, J., Hunt, D.F. and White, F.M., Nat. Biotechnol., **20**, 301 (2002)

150. Tuytten, R., Lemiere, F., Witters, E., Van Dongen, W., Slegers, H., Newton, R.P., Van Onckelen, H. and Esmans, E.L., J. Chromatogr. A, **1104**, 209 (2006)

151. Ishii, T.M., Takano, M., Xie, L.H., Noma, A. and Ohmori, H, J. Biol. Chem., **274**, 12839 (1999)

152. Bradford, M.M., Anal. Chem., **72**, 248-254 (1976)

153. 2-D Quant Kit, Instructions 2D electrophoresis, Amersham Biosciences, (2003)

154. Millipore, User Guide for Reversed-Phase ZipTips®$_{MC}$ Pipette Tips

155. Wu, C.C. and Yates, J.R., Nat. Biotechnol., **21**, 262 (2003)

156. Yang, X., Wu, H., Kobayashi, T., Solaro, R.J. and van Breemen, R.B., Anal. Chem., **76**, 1532 (2004)

157. Gerace, L. and Burke, B., Annu. Rev. Cell Biol., **4**, 335 (1988)

158. Nigg, E.A., Curr. Opinion Cell Biol., **1**, 435 (1989)

159. Murray, A.W. and Kirschner, M.W., Nature, **339**, 275 (1989)

160. Greaves, M., Eguchi, M., Eguchi-Ishimae, M., Int. J. Hematol., **78**, 390 (2003)

161. Milne, T.A., Hughes, C.M., Lloyd, R., Yang, Z., Rozenblatt-Rosen, O., Schnepp, R.W., Krankel, C., Livolsi, V.A., Gibbs, D., Hua, X., Roeder,R.G., Meyerson, M. and Hess, J.L., Proc. Nat. Acad. Sci., **102**, 749 (2005)

162. Starr, D., Williams, B. C., Li, Z., Etemad-Moghadam,B., Dawe, R. K. and Goldberg, M. L., J. Cell Biol., **138**, 1289 (1997)

163. Collier, B., Gorgoni, B., Loveridge, C., Cooke, H.J., Gray, N.K., Eur. Mol. Biol. Org. J., **24**, 2656 (2005)

164. Vogel, T., Speed, R.M., Ross, A. and Cooke, H.J., Mol. Human Reproduction, **8**, 797 (2002)

165. Lim, K.M., Yeo, W.S. and Chow, V.T.K., Int. J. Cancer, **109**, 24 (2004)

166. Srivastava, S.P., Kumar, K.U. and Kaufman, R.J., J. Biol. Chem., **273**, 2416 (1998)

167. Lillis, A.P., Mikhailenko, I. and Strickland, D.K., J. Thrombosis and Haemostatic, **3**, 1884 (2005)

168. Kasahara, K., Chida, K., Tsunenaga, M., Kohno, Y., Ikuta, T. and Kuroki, T., J. Biol. Chem., **266**, 20018 (1991)

169. Meier, R., Mtiller, P.R., Hirtt, A., Leibundgutt, H., Ridolfi-Liithyt, A. and Wagner, H.P., Leukemia Research, **21**, 841 (1997)

Chapter 8 Conclusion

Summary

A HPLC-ESI/MS method was developed for simultaneous quantitation of 10 major active components in *Ginkgo biloba* nutritional supplements, by a single run, in this study. Mass spectrometry has the advantage in terms of sensitivity and selectivity, there is no tedious sample clean up procedures to follow. The sample preparation and assay procedure involved is simple, rapid and has demonstrated good accuracy and reproducibility. There is remarkable variation in the contents of the flavonoids and terpene lactones although all the *Ginkgo biloba* nutritional supplements satisfy the traditional quality control standards. Fortification by use of low cost components rutin and quercetin was observed in two out of five commercially available *Ginkgo biloba* products analyzed in this study. These results indicate that suitable quality control methods need to be implemented to ensure the quality of *Ginkgo biloba* nutritional supplements and this method may serve as a valuable tool for the quality evaluation of *Ginkgo biloba* dietary supplement products.

In this study, LC/MS was carried out to produce fingerprint profiles of *Ginkgo biloba* extracts, and the knowledge of the fragmentation pathways of active components was applied to determine the active components in *Ginkgo biloba* found in the fingerprint. Fingerprinting can generate an overview of all the components in the sample, under investigation, and is a complimentary quality control method to exact quantitation using

reference standards. Mass and fragmentation information of the resolved components can be simultaneously obtained by LC/MS employing data-dependent scan mode. Capillary column displays better sensitivity than normal-bore column in the identification of unknown active components in commercial samples of *Ginkgo biloba*.

In this thesis, an on-line purification, column switching HPLC/MS method was developed for the analysis of urinary excreted flavonoids and terpene lactones derived from *Ginkgo* extract in a single analysis. The reverse-phase trap column was shown to allow the quick and efficient clean up of the injected urine sample which required minimal previous sample preparation allowing higher throughput and greater automation to be achieved. Comparison of this method to off-line purification methodologies show that the on-line purification does not suffer from any significant loss of analyte or interference from other matrix related effects. The experiments were performed as a "proof of principle" study and were validated in order to demonstrate adequate recovery of the analytes, linear range of the analysis and acceptable limits of detection.

The mass accuracy at high molecular weight by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry was investigated to resolve an experimental observation of poor mass accuracy when studying high mss proteins. Bovine serum albumin (BSA) was employed as a model compound and strategies to

improve mass measurement at high mass were examined. The phenomenon of peak broadening observed for MALDI-TOF high mass protein peaks was investigated and a theoretical study showed that these phenomena can be largely explained on the basis of two factors, i.e., the initial velocity distribution and ion fragmentation in flight.

In this study, an online IMAC-nano LC/MS platform for phosphorylated protein profiling was set up and phosphorylated proteins in cCMP incubated mice brain homogenate were characterised. These proteins are thought to be substrates of cCMP-responsive protein kinase and many of them are thought to be involved in cell differentiation. The data obtained support previous research implicating cCMP has a role in cell hyperproliferation in brain and other mammalian tissues. Also proteins related to Leukaemia are found to be phosphorylated (FMS-like tyrosine kinase 1and myeloid/lymphoid or mixed lineage-leukaemia translocation to 1 homolog), this supports the view that cCMP is involved in the stimulation of leukaemia L – 1210 cells. The identification of these phosphorylated proteins should provide new avenues for investigating the mechanism of cCMP signaling pathways and its precise cellular function.

Appendices

Appendix 1 Q-TOF MS/MS spectra of some standards in *Ginkgo biloba*

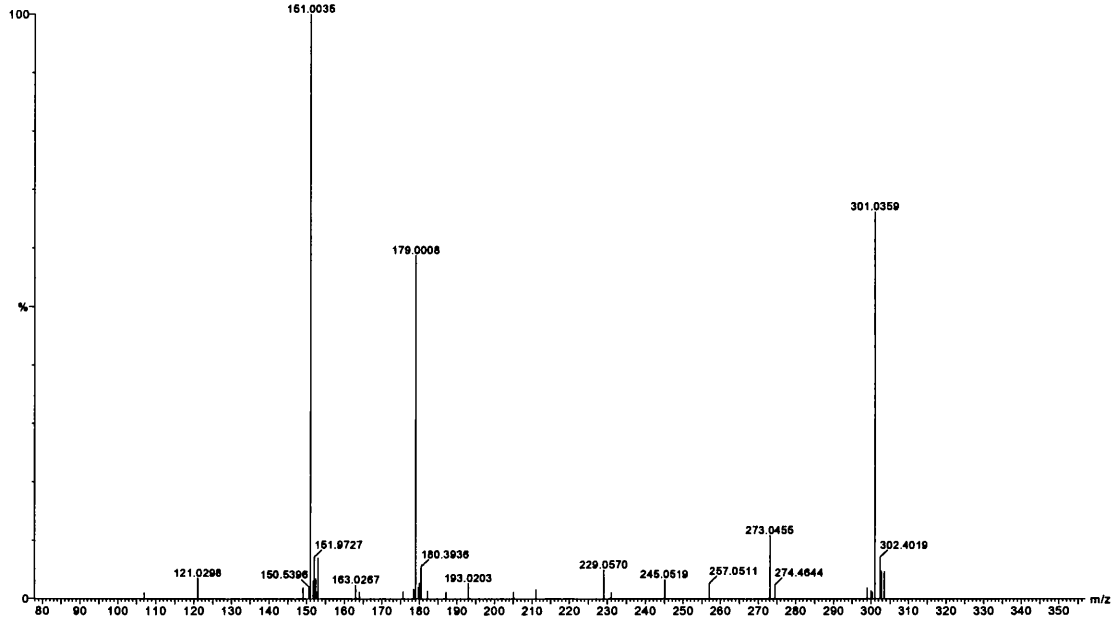


Figure A.1 MS/MS spectrum of QD by Q-TOF

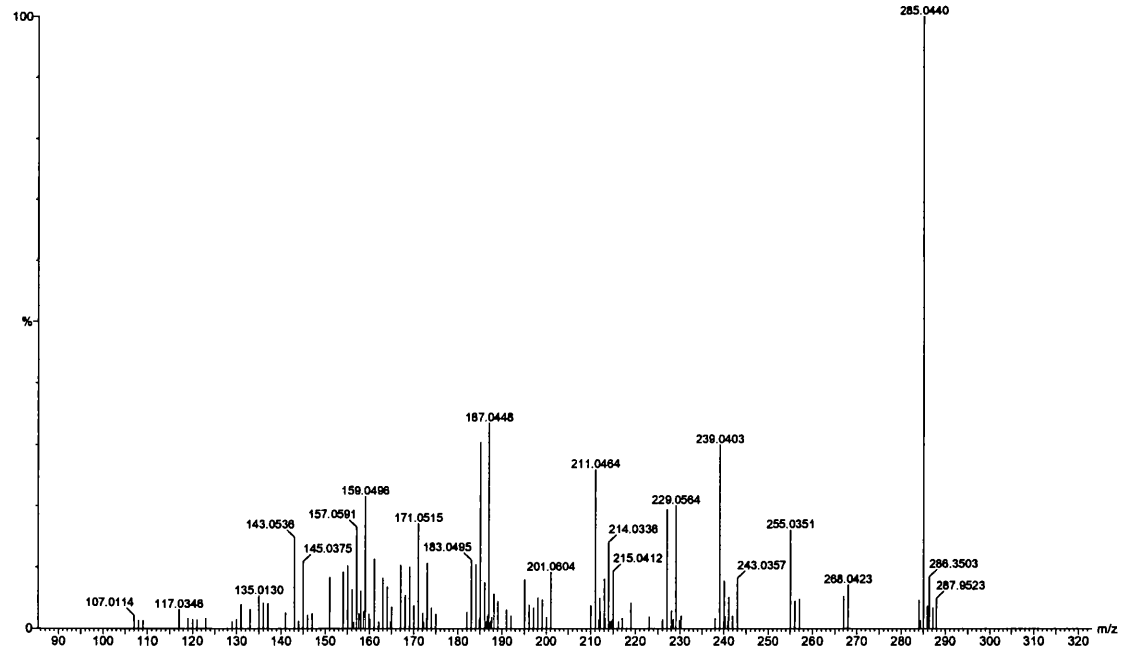


Figure A.2 MS/MS spectrum of KF by Q-TOF

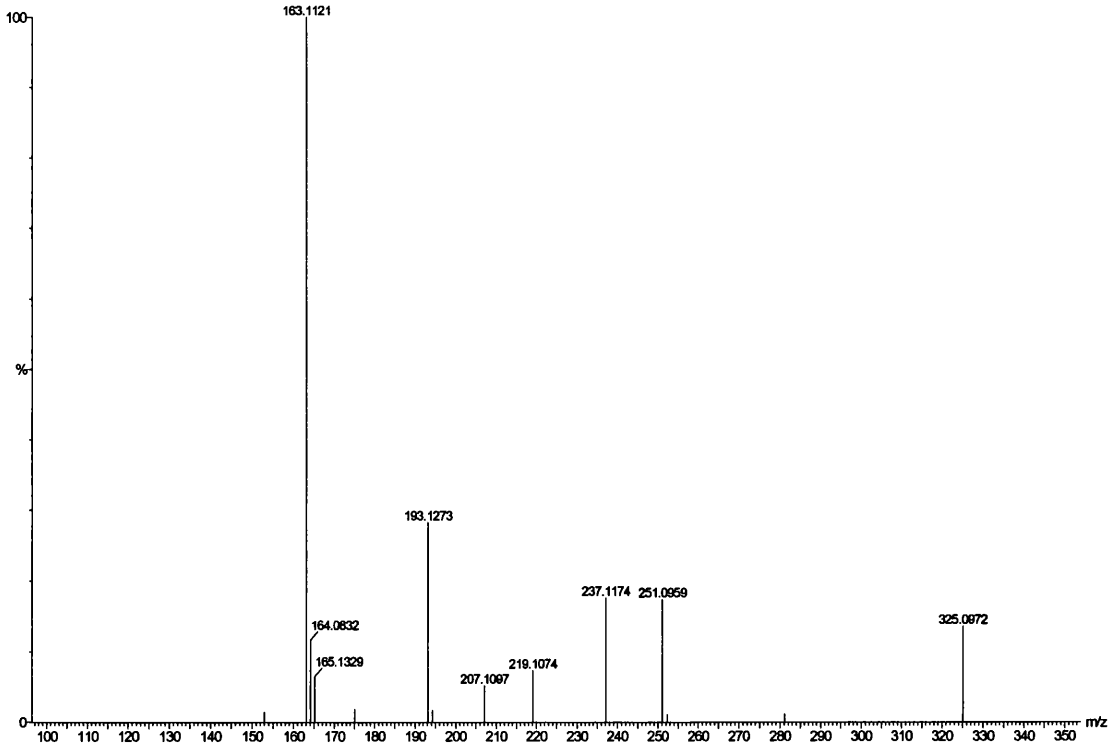


Figure A.3 MS/MS spectrum of BL by Q-TOF

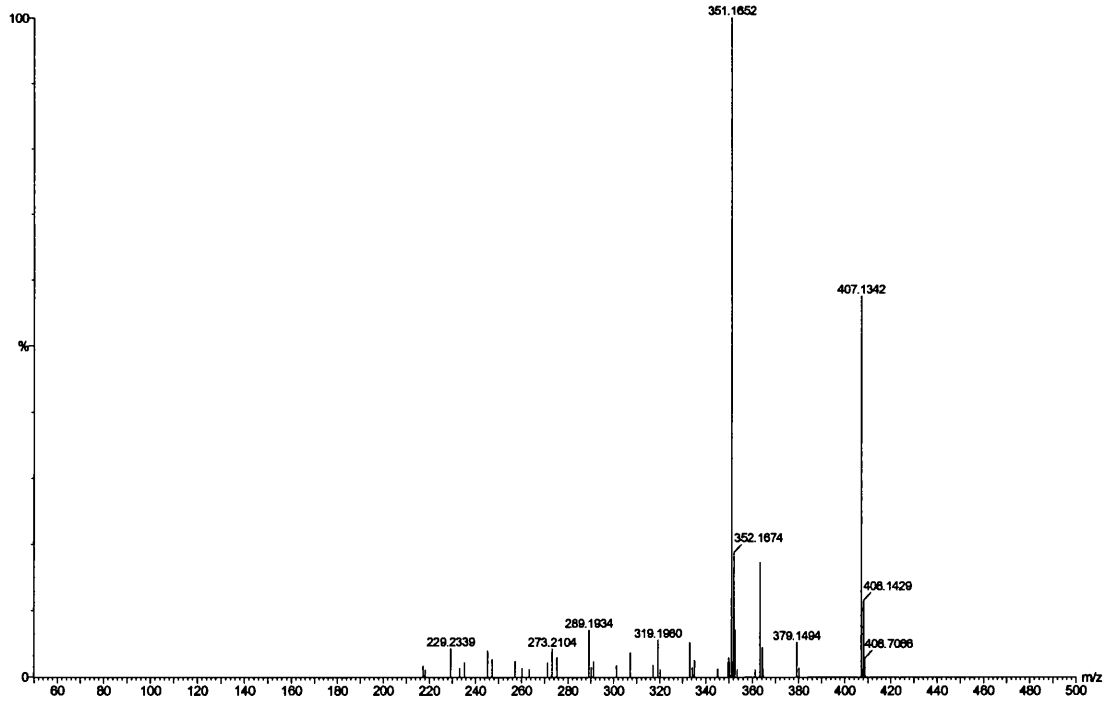


Figure A.4 MS/MS spectrum of GA by Q-TOF

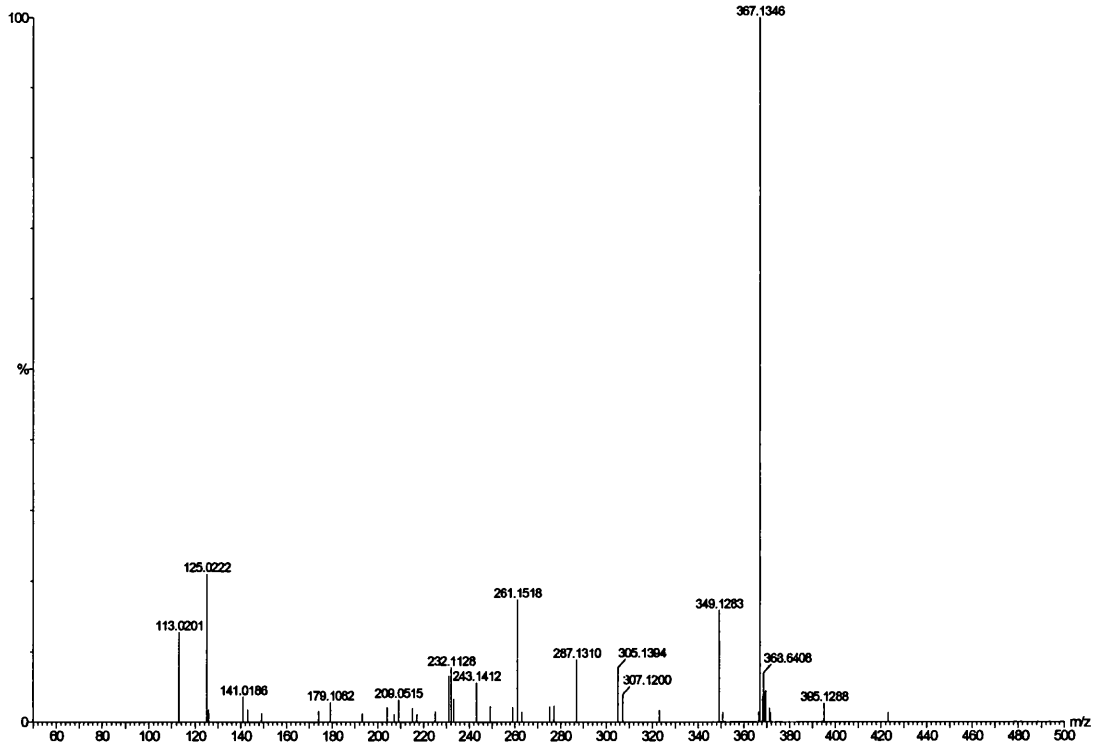


Figure A.5 MS/MS spectrum of GB by Q-TOF

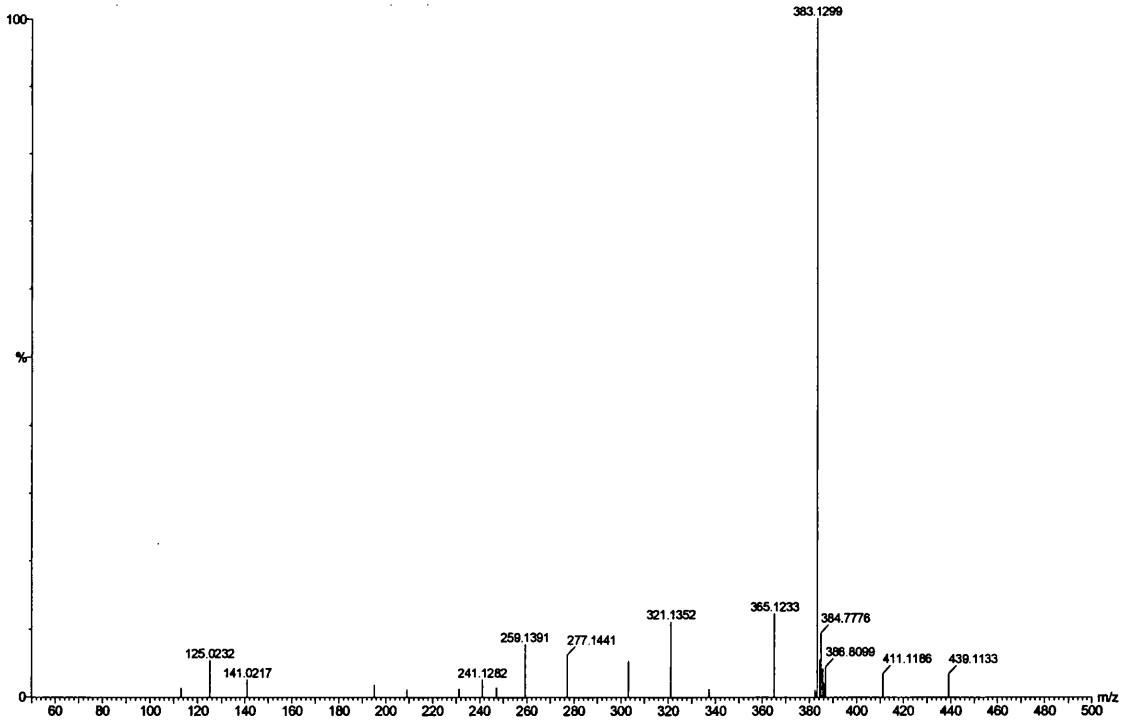


Figure A.6 MS/MS spectrum of GC by Q-TOF

Appendix 2 Phosphorylated proteins identified in cyclic nucleotide incubated brain homogenate at 4 time point

Table A.1 Phosphorylated proteins identified in cAMP incubation 1 min

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
odd Oz/ten-m homolog 3	RLSEILYD*p*STRV	1277.29	2	2.29	0.26	484.1	12/27
unc-13 homolog B	KI*p*TITVVCAQGLQAKD	1525.72	2	2.35	0.18	330.4	16/39
PREDICTED: hypothetical protein LOC69743	KQDVIRH*p*YNMHKKR	1649.78	2	2.47	0.39	475.1	16/33
dystonin isoform b	RSEAELGGGASEDGHL*p*SFTGSDRD	2260.15	2	2.44	0.17	155.1	12/63
jumonji domain containing 2A	MASE*p*SETLNPSARI	1473.47	2	2.71	0.13	900.0	17/36
MAP/microtubule affinity-regulating kinase 3	KIADFGF*p*SNEF*p*TVGSKL	1779.69	2	2.96	0.46	336.3	18/56
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54	KAR*p*SAQTGARALILSPTRE	1850.00	2	2.36	0.12	385.0	14/48
PREDICTED: similar to Gmps protein	KI*p*SQMPVILTPLHFDRD	1848.08	2	2.26	0.16	392.9	17/42
zinc finger, DHHC domain containing 5	KDEVQLKTT*p*Y*p*SKS	1472.38	2	2.23	0.17	851.4	18/40
PREDICTED: similar to MHC class I histocompatibility antigen H-2 Q4 al	KAHV*p*TCHHR*p*SDGDVTLRC	1964.90	2	2.42	0.34	393.8	23/60
Grp94 neighboring nucleotidase isoform 3	RTTLSVLECSLSEI*p*TAADKG	1962.09	2	2.35	0.56	327.7	16/51
gene rich cluster, C2f gene	KIG*p*T*p*SF*p*SVEDISDIRE	1779.57	2	2.49	0.53	510.6	20/65
transcriptional repressor NAC1	RAVLAA*p*S*p*SSYFRD	1332.25	2	2.30	0.44	306.0	15/40
PREDICTED: similar to L-NAME induced actin cytoskeletal protein	KTVEEKL*p*THRQ	1193.22	2	2.25	0.38	517.8	13/24
PREDICTED: similar to protein kinase, lysine deficient 3	KELQEL*p*YERL	1160.14	2	2.33	0.37	437.2	12/21
PREDICTED: similar to cajalin 2 isoform a	KNVSKSDSDLIAYP*p*SKD	1704.74	2	2.26	0.17	289.5	15/42
flotillin 1	KITLV*p*SSG*p*SGTMGAAKV	1540.53	2	2.33	0.16	275.1	23/56
NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 8	KAAAHHYGAQCDK*p*TNKE	1695.72	3	2.87	0.12	570.6	26/84
Rho GTPase activating protein 22	RTSSLDGPAAAVLSRT*p*SPPRL	1964.05	2	2.36	0.39	373.6	20/54
ELAV (embryonic lethal, abnormal vision, Drosophila)-like 4 (Hu antigen	RNCPSPMQTGAATDD*p*SKT	1703.71	2	2.49	0.16	260.3	16/45
hypothetical protein LOC399591	K*p*T*p*TEVKNTLPSNENKM	1849.74	2	2.43	0.21	407.9	21/56
dapper homolog 1, antagonist of beta-catenin (xenopus)	RAPSGVPPG*p*S*p*SVNFKNTKQ	1847.81	2	2.30	0.12	375.4	22/64
transforming, acidic coiled-coil containing protein 2 isoform c	KL*p*SCPEDSLPPSPKNRL	1720.81	2	2.24	0.11	425.2	16/42
serine/arginine-rich protein specific kinase 2	R*p*SR*p*TVSASSTGDLPKT	1566.46	2	2.24	0.17	169.4	19/52
pam, highwire, rpm 1	KPLP*p*TPR*p*SSPSGASSPRS	1754.69	3	2.82	0.42	534.5	32/120

Continues

Table A.1 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
solute carrier family 22 (organic cation transporter), member 18	KEASVQ*p*SAPQGG*p*TKA*p*SVFDLKA	2261.12	2	2.46	0.12	215.9	20/95
PREDICTED: similar to GREB1 protein isoform a	K*p*SVLQHIEATPKI	1303.37	2	2.24	0.33	428.3	13/30
PREDICTED: similar to hypothetical protein FLJ23834	KGTVVT*p*TNHLDVEKT	1493.52	2	2.22	0.20	520.2	15/36
general transcription factor II I	KFGEAIGMGFPVKVP*p*YRK	1849.07	2	2.46	0.47	618.3	20/45
PREDICTED: protein tyrosine phosphatase, receptor type	KPKNRYANVIA*p*YDH*p*SRV	1964.92	2	2.22	0.15	252.8	17/56

Table A.2 Phosphorylated proteins identified in cAMP incubation 5 min

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
PREDICTED: similar to Gm397 protein	KMFNSWLQPEKQ*p*TKE	1717.85	2	2.89	0.13	448.9	15/36
PREDICTED: similar to cajalin 2 isoform a	KNVSKSDSDLIAYP*p*SKD	1704.74	2	2.31	0.17	342.8	17/42
PREDICTED: RIKEN cDNA D630013G24 gene	KLL*p*STYGAEELYRAKR	1694.79	3	3.21	0.16	756.7	26/78
retinoblastoma-like 2	RSQVYR*p*SVLIKGKR	1458.61	2	2.25	0.43	153.4	13/33
heat shock protein 105	KLKVLG*p*TAFDPFLGGKN	1643.83	2	2.44	0.16	664.0	18/42
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54	KAR*p*SAQTGARALIL*p*SPTRE	1929.96	2	2.37	0.15	409.5	20/64
unc-13 homolog B	KITI*p*TVVCAQGLQAKD	1525.72	2	2.31	0.17	303.6	14/39
ATP-binding cassette transporter sub-family A member 16	KPNFP*p*YL*p*SEIA*p*TVKF	1719.60	2	2.43	0.21	292.5	20/60
TAF6 RNA polymerase II, TATA box binding protein (TBP)-associated factor	RQKLTT*p*SDID*p*YALKL	1656.62	2	2.45	0.38	233.5	15/48
ribosomal protein S6 kinase, polypeptide 5	R*p*TSTSTETRS*p*SSSESSRS	1930.56	2	2.25	0.16	416.6	24/75
shroom	K*p*SLSSELLTC*p*SPQHRK	1718.67	2	2.33	0.28	254.5	14/52
PREDICTED: myosin XVIIIb	KTRFG*p*SCESLLESGPCTRR	1924.03	2	2.28	0.57	670.1	19/48
lymphoblastomic leukemia	KLA*p*SPGPL*p*S*p*TEEVDHRN	1848.63	2	2.30	0.19	404.3	22/70
PREDICTED: similar to 40S ribosomal protein S26	RPST*p*SRIL*p*YL*p*SSLGSKM	1849.75	2	2.25	0.12	296.6	21/70
centromere autoantigen H	K*p*SNIEMK*p*SLAL*p*SRMKL	1848.85	2	2.25	0.14	438.8	20/65
ATP-binding cassette, sub-family C, member 10 isoform mrp7A	KLM*p*TELL*p*SGIRVIKF	1633.78	2	2.24	0.18	508.7	18/48
sarcoma antigen NY-SAR-41	KQRESS*p*TEKLRK	1314.31	2	2.21	0.16	364.1	13/27
PREDICTED: hypothetical protein LOC68964	K*p*TLEDLI*p*YKL	1155.08	1	1.91	0.12	653.4	17/28

Continues

Table A.2 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
PREDICTED: similar to Gamma-aminobutyric-acid receptor rho-3 subunit p	KNEQEA*p*TYARD	1162.07	2	2.56	0.17	404.8	12/24
PREDICTED: similar to protein kinase, lysine deficient 3	KELQEL*p*YERL	1160.14	2	2.29	0.16	349.3	10/21
glutamate receptor, ionotropic, kainate 2 (beta 2)	RTLLPN*p*T*p*TLT*p*YDTQKI	1849.70	2	2.22	0.16	245.7	19/65
synaptosomal-associated protein 91	RY*p*SRYLNEKA	1153.15	1	1.88	0.11	308.5	10/21
protocadherin beta 17	KYNITI*p*TV*p*SDLG*p*TPRL	1790.64	2	2.46	0.44	172.8	19/65
ATPase, H+ transporting, V1 subunit C, isoform 2	REF*p*Y*p*YDEKE	1153.97	2	2.20	0.12	326.4	12/24

Table A.3 Phosphorylated proteins identified in cAMP incubation 15 min

protein	Peptide	MH+	z	XC	dCn	Sp	Ions
HS1 binding protein	RIFEGVLE*p*SHAKPESPP	1848.96	2	2.26	0.11	345.2	14/45
PREDICTED: similar to Gmps protein	KI*p*SQMPVILTPLHFDRD	1848.08	2	2.35	0.19	524.5	16/42
integrin alpha 6	K*p*YTQELTLNRQKQ	1474.52	2	2.28	0.30	953.6	17/30
PREDICTED: hypothetical protein MGC31423	KVHVDPSGTFLAT*p*SCSDKS	1844.90	2	2.31	0.42	396.3	18/48
CBF1 interacting corepressor	KNML*p*YEELSSSHSDRGKA	1933.96	2	2.31	0.15	219.7	11/45
PREDICTED: similar to putative pheromone receptor	RENSDF*p*TKFIQ*p*TVNPKN	1928.88	2	2.45	0.13	286.4	17/56
protein phosphatase 1, regulatory (inhibitory) subunit 1C	K*p*SESLDEEEKLELQRR	1785.77	2	2.67	0.30	265.9	21/39
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54	KAR*p*SAQTGARALIL*p*SPTRE	1929.96	2	2.37	0.16	465.2	22/64
gene rich cluster, C2f gene	KIG*p*T*p*SF*p*SVEDISDIRE	1779.57	2	2.73	0.39	373.5	16/65
PREDICTED: hypothetical protein XP_357954	R*p*SNHVSLVDMIIKADRF	1778.94	2	2.43	0.18	441.5	15/42
adaptor-related protein complex 2, beta 1 subunit	RIQPGNPN*p*YTLSLKC	1525.61	2	2.33	0.18	431.8	17/36
hypothetical protein LOC70909	KS*p*FQQQLCDAIAIIKG	1658.82	2	2.36	0.40	512.1	15/39
PREDICTED: similar to protein kinase, lysine deficient 3	KELQEL*p*YERL	1160.14	2	2.22	0.39	448.3	12/21
PREDICTED: hypothetical protein LOC75173	RAA*p*TFTYSPLLYWINKR	1869.03	2	2.21	0.17	219.6	11/42
DVL-binding protein DAPLE	KENQQL*p*SKKIEKL	1425.49	2	2.21	0.12	726.6	18/30
PREDICTED: cholinergic receptor, nicotinic, alpha polypeptide 9	KV*p*YGKLPGSNLKA*p*SRN	1650.66	2	2.40	0.51	209.5	15/52

Continues

Table A.3 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
PREDICTED: similar to TATA element modulatory factor 1	R*p*SV*p*SEINSDDELPGKG	1650.48	2	2.23	0.55	384.4	20/52
unc-13 homolog B	KI*p*TITVVCAQGLQAKD	1525.72	2	2.24	0.21	234.4	12/39
jumonji domain containing 2A	MA*p*SESETLNPSARI	1473.47	2	3.04	0.17	773.1	17/36
phosphodiesterase 4B, cAMP specific	KN*p*SPKEFTASESEVCIKT	1849.92	2	2.35	0.10	308.3	13/45
PREDICTED: similar to Interferon-activatable protein 205 (IFI-205)	**KGLESMED*p*YQFR*p*TVKSLLRK**	**2333.41**	**3**	**2.85**	**0.13**	**460.6**	**30/136**
PREDICTED: RIKEN cDNA 1700065A05 gene	**KLDNIL*p*SGPKIDVLKS**	**1605.78**	**2**	**2.34**	**0.32**	**558.2**	**18/39**
stathmin 1	RSKESVPDFPL*p*SPPKK	1608.70	2	2.25	0.40	589.0	20/39
poly (ADP-ribose) polymerase family, member 2	KPE*p*SQLDLRVQELLKL	1748.88	2	2.23	0.28	299.1	13/39
hypothetical protein LOC194219	RKR*p*SAQTQMWE*p*SPVM*p*SQGKQ	2320.28	3	3.00	0.23	332.9	36/170
splicing factor 3a, subunit 2	KTGSGGVA*p*SS*p*SE*p*SNRDRRE	1963.64	2	2.22	0.30	470.6	22/80
PREDICTED: protein (peptidyl-prolyl cis/trans isomerase) NIMA-interact	K*p*SGSGKGGKGGAA*p*SG*p*SDSADKK	1921.60	2	2.38	0.29	331.8	22/95
hypothetical protein LOC70385	KYQ*p*SLKKQNAFTRD	1564.65	2	2.32	0.10	498.3	15/33
lymphoblastomic leukemia	KLA*p*SPGPL*p*S*p*TEEVDHRN	1848.63	2	2.45	0.23	500.3	22/70
PREDICTED: zinc finger protein 294	RL*p*SLN*p*YECFK*p*TDFKE	1848.74	2	2.44	0.11	286.6	19/60
AT motif binding factor 1	KLL*p*SNKNISAIIQGIGKD	1749.96	2	2.41	0.17	217.9	12/45
procollagen, type XXVII, alpha 1	RPDVSER*p*SCSQTPLSPAKQ*p*SARK	2405.39	2	2.30	0.14	154.8	18/80
phosphatidylinositol glycan, class W	KI*p*SLE*p*SEYNPAITCYRV	1919.89	2	2.29	0.18	225.0	14/56
G-protein coupled receptor 12	RYLSLYYAL*p*TYHSERT	1859.94	2	2.21	0.15	339.2	15/39
zinc finger protein 445	K*p*TSVSPETSVSEE*p*TGLKE	1811.68	2	2.20	0.15	427.2	18/60
macrophage activation 2 like	RL*p*SLPTE*p*TLQELLDVHVACEKE	2399.51	3	2.89	0.34	665.0	34/152
zinc finger, DHHC domain containing 5	KDEVQLK*p*TT*p*YSKS	1472.38	2	2.65	0.16	814.3	18/40
PREDICTED: protein tyrosine phosphatase, receptor type, D	KPKNRYANVIAYDH*p*SRV	1964.92	2	2.51	0.40	203.3	17/56
PREDICTED: vacuolar protein sorting 13C	K*p*SIGA*p*TL*p*TDVDDLIFKL	1848.71	2	2.41	0.18	294.4	20/70
annexin A6	REFIEK*p*YDKS	1152.16	2	2.30	0.22	447.7	14/21
PREDICTED: similar to hypothetical protein FLJ38281	REFGVAGPQAAPPW*p*SRV	1650.70	2	2.27	0.46	332.4	17/42
hypothetical protein LOC239591	KLKQGIPQDLS*p*S*p*SPKL	1658.64	2	2.25	0.29	213.5	21/52
piccolo (presynaptic cytomatrix protein)	RTIPK*p*SEVKVTEKC	1439.56	2	2.22	0.25	262.4	14/33
phospholipase A2, group X	**KYLFFP*p*SILCEKD**	**1440.61**	**2**	**2.20**	**0.44**	374.4	15/30
CLIP-associating protein CLASP2	**REASRE*p*SSRDTSPVRS**	**1657.61**	**2**	**2.26**	**0.49**	**221.6**	**11/39**
proprotein convertase subtilisin/kexin type 1	**KL*p*SIPYESF*p*YEALEKL**	**1849.82**	**2**	**2.24**	**0.10**	326.2	17/52
regulator of G-protein signaling 20	RASP*p*SA*p*SPADPGLPEG*p*SERT	1965.69	2	2.21	0.44	164.9	21/85

Table A.4 Phosphorylated proteins identified in cAMP incubation 30 min

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
growth associated protein 43	**KEGDGSATTDAAPAT*p*SPKA**	**1656.57**	**2**	**2.38**	**0.20**	**567.5**	**20/48**
nuclear, casein kinase and cyclin-dependent kinase substrate	**KTPSPQEEDEEAE*p*SPPEKKS**	**2108.04**	**3**	**3.50**	**0.36**	**834.6**	**35/102**
RNA binding motif protein 5	KCQGI*p*TAPIEAQVRL	1466.57	2	2.34	0.49	557.2	15/36
small glutamine-rich tetratricopeptide repeat (TPR) containing protein	RAPDRTPP*p*SEEDSAEAERL	1937.84	3	3.01	0.39	902.3	32/96
PREDICTED: similar to Gm397 protein	KMFNSWLQPEKQ*p*TKE	1717.85	2	2.23	0.16	332.2	14/36
heat shock protein 1, beta	KIEDVG*p*SDEEDDSGKDKKK	1946.84	3	3.12	0.25	1248.0	39/96
protein phosphatase 1, regulatory (inhibitor) subunit 1B	RVS*p*EHSSPEEEASPHQRT	1886.79	3	3.12	0.13	634.8	28/90
hypothetical protein LOC235472	KV*p*TLIEDDTALV*p*SWKP	1750.73	2	2.30	0.10	352.5	16/52
retinal degeneration, slow (retinitis pigmentosa 7)	K*p*SNQVEAEGADAGPAPEAG	1750.64	2	2.27	0.16	193.8	14/51
transcription elongation factor A (SII)-like 3	RGTDD*p*SPKNSQEDLQDRH	1885.76	2	2.77	0.19	546.8	26/45
ubiquitin specific protease 7	RAGFIQDTSLIL*p*YEEVKP	1907.03	2	2.31	0.17	239.2	12/45
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54 [Mus musculus]	KAR*p*SAQTGARALIL*p*SPTRE	1929.96	2	2.25	0.11	613.3	22/64
odd Oz/ten-m homolog 3	RLSEILYD*p*STRV	1277.29	2	2.28	0.34	499.4	13/27
coagulation factor V	KENRLDPPIVAR*p*YIRI	1792.94	2	2.33	0.24	296.0	14/39
hypothetical protein LOC244585	KNGDLSFLEKVD*p*SKI	1532.56	2	2.23	0.28	431.0	15/36
PREDICTED: similar to BC037112 protein	RIDAYVDEQM*p*TMK*p*TKQ	1833.87	2	2.22	0.43	383.5	17/52
ATP synthase, H+ transporting, mitochondrial F1 complex, O subunit	R*p*TVKVKSLNDI*p*TKR	1506.53	2	2.48	0.12	564.5	18/44
BCL2/adenovirus E1B 19kDa-interacting protein 1, NIP3 [Mus musculus]	KNSTLSEED*p*YIERRRE	1848.83	2	2.55	0.40	394.3	13/39
NACHT, leucine rich repeat and PYD containing 1	R*p*YIV*p*SS*p*SKLVEIIPKE	1816.80	2	2.33	0.38	248.1	23/65
phosphodiesterase 5A, cGMP-specific	RKATRDMVNAWF*p*SERV	1791.89	2	2.30	0.34	294.3	13/39
guanylate cyclase activator 2b	RTIATDECELCINVAC*p*TGC	1940.11	2	2.64	0.11	244.1	13/51
synaptotagmin 4	KSEVKGKAALPNL*p*SLHLDLEKR	2243.47	3	2.84	0.22	547.4	29/114
choroidermia	RFNIDLV*p*SKLLY*p*SRG	1728.77	2	2.55	0.14	377.7	16/48
regulator of G-protein signaling 13	**RFLK*p*SEMYQQLLKTVQSQ*p*SS**	**2406.50**	**3**	**2.85**	**0.18**	**605.7**	**35/144**
SET domain-containing protein 7	**RVYVADSLISSAGEGLF*p*SKV**	**1924.02**	**2**	**2.33**	**0.48**	**303.6**	**14/51**

Continues

Table A.4 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention recepto	RKTFDIEND*p*TFRL	1466.46	2	2.24	0.41	487.9	14/30
THUMP domain containing 3	RTASVDIIV*p*TDMPFGKRM	1831.01	2	2.26	0.12	277.2	16/45
PREDICTED: titin	RTDGGQY*p*TLKLSNVGGTKT	1819.87	2	2.31	0.12	333.2	20/48
hypothetical protein LOC74626	KLVNLNFQQ*p*SL*p*TEDQKLIDKG	2407.47	3	3.27	0.18	701.7	38/144
PREDICTED: SEC15-like 2	**RKNIGLTELVQIIINT*p*THLEKS**	**2438.61**	**3**	**2.90**	**0.22**	**531.9**	**34/152**
PREDICTED: down-regulated in metastasis	**KQKPGHPE*p*YDIL*p*TNVFAVLSAKN**	**2488.58**	**3**	**2.90**	**0.21**	**566.5**	**38/160**
KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention recepto	R*p*SCAGISGKSQLLFALVF*p*TTRY	2260.40	3	2.85	0.14	709.7	37/152
calpain 12	REV*p*TLLPDSLSQRW	1438.49	2	2.61	0.11	577.8	15/33
annexin A6	REFIEK*p*YDKS	1152.16	2	2.55	0.22	595.4	15/21
anion exchanger Slc26a11	KGFR*p*YF*p*T*p*TLEEAEKF	1831.64	2	2.26	0.10	245.6	19/60
bile acid beta-glucosidase	RQI*p*YGCPLGGIGGG*p*TI*p*TRG	1903.82	2	2.24	0.13	226.7	19/80
hypothetical protein LOC384806	KPKNL*p*S*p*S*p*TFEHLVHKM	1877.76	2	2.21	0.21	158.8	19/65
PREDICTED: hypothetical protein LOC78516 [Mus musculus]	RFSEEPGKV*p*SRASRA	1530.55	2	2.36	0.46	313.8	14/36
PREDICTED: DDHD domain containing 2	KEI*p*YQ*p*TQGVFLDQPLQ	1939.90	2	2.35	0.18	185.3	17/56
putative homeodomain transcription factor 1	R*p*SIL*p*SRHLN*p*SQVKKT	1750.67	2	2.30	0.13	214.6	15/60
hemolytic complement	KIEEQAAK*p*YKH	1160.18	2	2.26	0.34	399.7	13/24
component of oligomeric golgi complex 4	RLVQ*p*SNLMRN*p*SA*p*TEKI	1831.71	2	2.22	0.19	197.6	17/65
phosphatidylinositol-4-phosphate 5-kinase, type 1 gamma	RLLLHIGIIDILQS*p*YRF	1848.10	2	2.50	0.13	502.1	15/42

Table A.5 Phosphorylated proteins identified in Blank incubation 1 min

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
PREDICTED: similar to Gm397 protein	KMFNSWLQPEKQ*p*TKE	1717.85	2	2.51	0.20	248.4	12/36
EF hand domain containing 1	KAAAGELQED*p*SGLLALAKF	1737.81	3	3.03	0.45	533.8	25/96
zinc finger protein 638	KF*p*TQAGAQKMKRL	1346.46	2	2.27	0.21	423.3	12/30
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54	**KAR*p*SAQTGARALILSPTRE**	**1850.00**	**2**	**2.42**	**0.13**	**414.8**	**14/48**
insulin-like growth factor binding protein 6	**RGPSEETTKE*p*SKP**	**1273.21**	**2**	**2.32**	**0.19**	**356.0**	**12/30**

Continues

Table A.5 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
hypothetical protein LOC70713	RTTLFSAAFSLSGSLPLLRPP*p*SRL	2399.65	3	2.88	0.34	686.3	31/126
unc-13 homolog B	KITI*p*TVVCAQGLQAKD	1525.72	2	2.24	0.13	351.7	16/39
zinc finger, DHHC domain containing 5	KDEVQLK*p*T*p*TYSKS	1472.38	2	2.64	0.02	829.8	19/40
odd Oz/ten-m homolog 3	RLSEILYD*p*STRV	1277.29	2	2.28	0.41	511.4	13/27
PREDICTED: hypothetical protein LOC69743	**KQDVIRH*p*YNMHKKR**	**1649.78**	**2**	**2.47**	**0.45**	**399.5**	**15/33**
PREDICTED: protein tyrosine phosphatase, receptor type, D	**KPKNR*p*YANVIAYDH*p*SRV**	**1964.92**	**2**	**2.24**	**0.55**	**186.1**	**15/56**
PREDICTED: similar to Flj37300-A-prov protein	KGSNKYIN*p*YRD*p*SKL	**1605.49**	**2**	2.22	0.11	247.8	14/44
PREDICTED: similar to PI-3-kinase-related kinase SMG-1 isoform	KNSSST*p*SLNLKA	1131.10	2	2.44	0.08	440.0	15/27
jumonji domain containing 2A	MASE*p*SETLNPSARI	1473.47	2	3.03	0.26	1004.9	19/36
PREDICTED: similar to protein kinase, lysine deficient 3	KELQEL*p*YERL	1160.14	2	2.51	0.26	476.9	12/21
PREDICTED: similar to GREB1 protein isoform a	K*p*SVLQHIEATPKI	1303.37	2	2.36	0.31	388.5	14/30
growth factor receptor bound protein 2-associated protein 2	RA*p*S*p*SCE*p*TYEYPARG	1617.36	3	3.00	0.19	1163.5	37/110
multiple endocrine neoplasia 1	RSWLYLKG*p*SYMRCDRK	1859.05	2	2.25	0.46	338.7	15/39
myosin Ixb	RQNKHIQSCREES*p*S*p*THRE	2101.01	3	2.84	0.29	791.0	33/120
spir-2 protein	R*p*SRQTQ*p*SLYIPN*p*TRTLNFQ	2408.30	3	2.90	0.16	652.3	48/170
hypothetical protein LOC244585	KNGDLSFLEKVD*p*SKI	1532.56	2	2.26	0.41	288.0	11/36
PREDICTED: hypothetical protein LOC71841	RP*p*YLQEHSIQYERA	1643.66	2	2.21	0.13	324.8	13/33
sperm-associated cation channel 2	RVF*p*S*p*SIYVILWLLLGSIIFRN	2400.69	3	2.81	0.36	430.4	29/144
PREDICTED: hypothetical protein XP_357954	RSNHV*p*SLVDMIIKADRF	1778.94	2	2.37	0.37	579.3	17/42
phospholipase C-like 4	KIRDCEDPNDFSV*p*STL*p*SP*p*SGKL	2408.23	3	3.46	0.13	801.8	43/190
pre-mRNA processing factor 8	K*p*TI*p*S*p*SYTAFSRL	1373.13	2	2.37	0.11	337.5	16/45
crooked neck-like 1 protein	K*p*SYIDFEIEQEETERT	1868.81	2	2.27	0.14	413.0	16/39
hypothetical protein LOC226744	RV*p*SKE*p*TED*p*YLHSLLERC	2059.89	3	2.97	0.17	464.7	31/140
PREDICTED: similar to serologically defined breast cancer antigen NY-B	K*p*SVGL*p*SEELQLGHHQRR	1850.77	2	2.56	0.17	356.3	19/56
calpain 12	REV*p*TLLPDSLSQRW	1438.49	2	2.45	0.11	576.7	17/33
G patch domain containing 4	RKRQQCCEEDLDVS*p*SKD	1848.92	2	2.41	0.41	311.2	15/42
PREDICTED: RIKEN cDNA 4930485B16 gene	RY*p*TEQV*p*TPLLV*p*YAQRQ	1921.81	2	2.41	0.19	196.4	17/65
ring finger protein 20	**RSGSALLQ*p*SQSSTEDPKD**	**1795.64**	**2**	**2.28**	**0.14**	**300.1**	**20/60**
ring finger protein 12	**R*p*SR*p*SPLQP*p*TSEIPRRA**	**1864.73**	**2**	**2.21**	**0.10**	354.5	16/65

Continues

Table A.5 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
PREDICTED: similar to Discs, large homolog 5 (Placenta and prostate DL	RKVESLISQPMT*p*SGKN	1585.73	2	2.20	0.13	243.3	17/39
sex comb on midleg-like 2	RQ*p*SK*p*TPPTNEFKI	1437.34	2	2.41	0.17	1234.0	24/40
G protein-coupled receptor 126	KKGFNA*p*SYIRVAVSLRN	1761.93	2	2.38	0.56	279.1	15/42
four and a half LIM domains 5	**KCAAC*p*TKPI*p*TGLRGAKF**	**1650.75**	**2**	**2.29**	**0.46**	**435.2**	**21/56**
RuvB-like protein 1 [Mus musculus]	**KLDP*p*SIFE*p*SLQKE**	**1437.38**	**2**	**2.25**	**0.18**	**295.8**	**16/40**
ephrin B3	**RRCEAPPAPNLLL*p*TCDRP**	**1850.04**	**2**	**2.21**	**0.19**	**299.9**	**13/45**
PREDICTED: similar to Zinc finger protein 267 (Zinc finger protein HZF	RGVSGQCTQSPLYQQD*p*Y*p*SQEKL	2407.32	3	3.17	0.19	513.2	36/152
gene rich cluster, C2f gene	KMVSISN*p*YPLSAALTCAKV	1850.07	2	2.25	0.47	632.7	19/48
phospholipase A2, group IID	KIDGCKSL*p*TDNYKY	1437.48	2	2.24	0.16	288.9	15/33
secretogranin II	KY*p*PELLNT*p*NQLKRV	1649.63	2	2.22	0.46	211.4	13/44

Table A.6 Phosphorylated proteins identified in Blank incubation 5 min

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
EF hand domain containing 1	KAAAGELQED*p*SGLLALAKF	1737.81	3	3.19	0.44	720.5	28/96
growth associated protein 43	KEGDGSATTDAAPAT*p*SPKA	1656.57	2	2.68	0.18	562.2	20/48
hypothetical protein LOC70713 [Mus musculus]	RTTLFSAAFSLSGSLPLLRPP*p*SRL	2399.65	3	3.16	0.35	503.9	28/126
jumonji domain containing 2A	MASEvSETLNPSARI	1473.47	2	3.23	0.20	1084.6	20/36
PREDICTED: hypothetical protein LOC69743	KQDVIRH*p*YNMHKKR	1649.78	2	2.26	0.43	397.8	15/33
EGF-like module containing, mucin-like, hormone receptor-like sequence	KTHTLGVL*p*SEFKSKE	1527.63	2	2.45	0.17	484.8	17/36
rho/rac guanine nucleotide exchange factor (GEF) 18	RIIQNTEAG*p*TEDYKD	1562.54	1	2.03	0.13	514.2	18/36
PREDICTED: zinc finger protein 407	KCDYG*p*TNVPVEFRN	1480.51	3	2.82	0.30	726.0	24/66
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54	KAR*p*SAQTGARALILSPTRE	1850.00	2	2.55	0.20	319.8	13/48
dudulin 2	**KEAPKVGILG*p*SGDFARS**	**1597.68**	**2**	**2.22**	**0.34**	**331.3**	**17/42**

Continues

Table A.6 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
PREDICTED: MYST histone acetyltransferase (monocytic leukemia) 3	KEEPQGQERELE*p*TRV	1681.62	2	2.53	0.16	390.5	17/36
shroom	K*p*SLSSELLTC*p*SPQHRK	1718.67	2	2.29	0.15	466.4	19/52
Als2 C-terminal like [Mus musculus]	KGLYTQGVLD*p*TAALPPDP*p*S*p*STRK	2400.27	3	3.08	0.32	1088.5	54/200
zinc finger, DHHC domain containing 5	KDEVQLKTT*p*Y*p*SKS	1472.38	2	2.38	0.17	653.1	17/40
arylsulfatase B	KEYNNI*p*YSTNIF*p*TKRA	1923.86	2	2.22	0.50	364.6	16/52
hypothetical protein LOC320720	RAFC*p*SNIPHLKGKSAMKK	1813.06	2	2.42	0.25	253.1	15/45
syntaxin binding protein 1	RHKHIAEV*p*SQEVTRS	1694.63	3	2.89	0.10	642.1	33/96
RIKEN cDNA 2310044G17	RN*p*YLPILNS*p*Y*p*TKI	1566.42	2	2.24	0.13	158.9	15/50
predicted gene ICRFP703B1614Q5.5	RWLGRVGAAGG*p*SWVLARK	1736.88	2	2.27	0.43	256.4	16/45
PREDICTED: hypothetical protein LOC66498	KVAR*p*TD*p*SPDMPED*p*T	1674.41	2	2.44	0.27	415.0	21/60
DNA polymerase N	KSI*p*TVKP*p*SSTFREA*p*SRN	1906.76	1	2.02	0.12	211.8	17/70
SRY-box containing gene 30	RPRDDCRLG*p*TEAT*p*SNGLVH*p*SSKE	2484.33	3	2.91	0.05	532.8	40/200
claspin	KLDEDD*p*S*p*SSLL*p*TKE	1563.28	1	1.81	0.14	242.1	18/55
PREDICTED: similar to Tetratricopeptide repeat protein 6	RQFSQA*p*SDYF*p*S*p*TALKF	1833.62	2	2.28	0.48	315.9	20/65
cyclic nucleotide gated channel alpha 3	KLDIL*p*SLIPTDLA*p*YLKL	1848.96	2	2.33	0.12	338.3	20/56
transglutaminase 2, C polypeptide	R*p*SLVVGQKISTK*p*SVGRD	1719.77	2	2.25	0.10	323.0	19/56
transmembrane emp24 protein transport domain containing 5	KK*p*YI*p*TN*p*TDVLEMKL	1695.60	3	2.88	0.16	556.7	32/110
inhibitor of kappaB kinase epsilon	KKSGEVVAVKVFNSAS*p*YRR	1922.05	2	2.45	0.51	303.0	15/48
sex comb on midleg-like 2	RQ*p*SK*p*TPPTNEFKI	1437.34	2	2.36	0.12	1494.9	27/40
secretogranin II	K*p*YPELLN*p*TNQLKRV	1649.63	2	2.35	0.45	441.2	19/44
PREDICTED: DnaJ (Hsp40) homolog, subfamily C, member 13	R*p*SLQ*p*YGEQV*p*SEILSRS	1849.66	2	2.29	0.14	212.8	17/65
PREDICTED: similar to laminin, beta 4	REHPGALAG*p*TQAELRE	1530.55	2	2.26	0.43	866.9	19/39
1856 - 1858	REHPGALAG*p*TQAELRE	1530.55	2	2.24	0.45	392.7	14/39
coagulation factor II (thrombin) receptor-like 1	R*p*TVNRMQI*p*SL*p*SSNKF	1718.60	2	2.25	0.14	271.2	17/60
calcium channel, voltage-dependent, P/Q type, alpha 1A subunit	RKEMMAIWPNL*p*SQKT	1656.88	2	2.24	0.40	289.5	13/36
lipocalin 13	RD*p*SGENPEAMEEFKN	1563.51	1	1.83	0.12	204.1	13/36
ATPase, H+ transporting, V1 subunit C, isoform 2	REF*p*Y*p*YDEKE	1153.97	2	2.33	0.17	324.0	12/24
sperm associated antigen 1	KN*p*YKDAISK*p*YNECLKI	1849.85	2	2.24	0.17	333.1	19/52
PREDICTED: similar to NPAS3 (MOP6)	RFCCILRP*p*TLQLQIDGLSRNKV	2399.75	3	3.00	0.18	469.9	29/114

Table A.7 Phosphorylated proteins identified in Blank incubation 15 min

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
hypothetical protein LOC68423	RSEKMESV*p*SGYEAKV	1525.54	2	2.32	0.19	422.4	15/36
PREDICTED: similar to Gm397 protein	KMFNSWLQPEKQ*p*TKE	1717.85	2	2.53	0.19	241.3	12/36
jumonji domain containing 2A	MA*p*SESETLNPSARI	1473.47	2	2.93	0.16	915.5	19/36
G protein-coupled receptor 126	KKGFNA*p*SYIRVAVSLRN	1761.93	2	2.22	0.44	346.5	17/42
PREDICTED: similar to hypothetical protein 4932415M13	KG*p*SRDPTAG*p*SNVDNWRQ	1792.61	2	2.30	0.44	392.5	19/56
cell division cycle 25 homolog A	MELGP*p*SPPPRR	1161.24	2	2.32	0.09	610.9	15/27
dishevelled associated activator of morphogenesis 1	KQKEADAIDDTL*p*SSKLKV	1842.91	2	2.20	0.12	943.7	19/45
PREDICTED: hypothetical protein XP_357954	R*p*SNHVSLVDMIIKADRF	1778.94	2	2.22	0.25	325.7	13/42
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54	KAR*p*SAQTGARALILSPTRE	1850.00	2	2.45	0.14	413.9	15/48
unc-13 homolog B	KI*p*TITVVCAQGLQAKD	1525.72	2	2.31	0.19	238.8	13/39
Als2 C-terminal like	KGLYTQGVLD*p*TAALPPDP*p*SS*p*TRK	2400.27	3	3.24	0.34	928.0	49/200
zinc finger, DHHC domain containing 5	KDEVQLK*p*T*p*TYSKS	1472.38	2	2.50	0.18	573.8	16/40
PREDICTED: similar to mKIAA1754 protein	KAHV*p*SRIVGELVQAGRARG	1900.06	2	2.26	0.39	340.2	17/48
PREDICTED: RIKEN cDNA 1700065A05 gene	KLDNIL*p*SGPKIDVLKS	1605.78	2	2.39	0.38	541.7	18/39
interleukin 1 receptor, type I	RKY*p*TLITTLNI*p*SEVKS	1783.85	2	2.26	0.21	360.9	18/52
odd Oz/ten-m homolog 3	RLSEILYD*p*STRV	1277.29	2	2.34	0.37	376.2	12/27
DVL-binding protein DAPLE	KENQQL*p*SKKIEKL	1425.49	2	2.26	0.28	352.8	14/30
PREDICTED: similar to protein kinase, lysine deficient 3	KELQEL*p*YERL	1160.14	2	2.40	0.24	383.3	11/21
hypothetical protein LOC105351	KAIVE*p*SAENLEEVVSVVRV	1924.02	2	2.33	0.17	264.3	14/48
PREDICTED: similar to YTH domain containing 2	KT*p*SGRLNNGIPQVPVKR	1660.78	2	2.60	0.47	249.9	17/42
Werner syndrome protein	RAGL*p*TPETWKIIMDVIRN	1924.18	2	2.25	0.54	307.9	15/45
keratin complex 2, basic, gene 5	RNKL*p*TELEEALQKAKQ	1695.82	3	2.99	0.18	435.1	20/78
PREDICTED: hypothetical protein XP_622896	K*p*TYLTNTPNATAFKLFI*p*STGRN	2377.44	3	3.00	0.20	572.0	34/152
RAS protein-specific guanine nucleotide-releasing factor 1	KHTQDFDTDD*p*TLK*p*YRV	1915.76	2	2.20	0.37	306.3	15/52
fibroblast growth factor receptor 1	KDLV*p*SCA*p*YQVARGMEYLASKKC	2393.53	3	3.40	0.30	585.8	32/152
fibroblast growth factor receptor 3	KDLV*p*SCA*p*YQVARGMEYLASQKC	2393.49	3	3.37	0.39	585.8	32/152
glutathione S-transferase, mu 2	RLLLEY*p*TDT*p*S*p*YEDKKY	1958.78	2	2.38	0.16	493.2	22/65
PREDICTED: hypothetical protein LOC75173	RAATF*p*TYSPLLYWINKR	1869.03	2	2.22	0.16	252.8	11/42

Continues

Table A.7 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
conserved helix-loop-helix ubiquitous kinase	KVWAEAVH*p*YV*p*SGLKEDYSRL	2350.21	3	2.94	0.26	506.5	36/170
PREDICTED: hypothetical protein LOC73815	KLKH*p*SLLGQGRGAQEDSRS	1933.00	2	2.36	0.19	695.0	21/48
fatty acid synthase	RQL*p*TLRKLQEMSSKT	1642.83	2	2.31	0.18	377.5	13/36
BAP28 protein	RRLPILVQLV*p*T*p*TL*p*SAKK	1892.95	2	2.31	0.17	273.0	24/70
hexosaminidase B	KILEII*p*S*p*SLKKN	1304.36	2	2.28	0.13	242.1	16/36
plasma membrane associated protein, S3-12	K*p*YGLGHV*p*TEPRADTKT	1704.63	2	2.52	0.42	315.8	17/52
Grp94 neighboring nucleotidase isoform 3	RTTLSVLECSLSEI*p*TAADKG	1962.09	2	2.40	0.55	339.0	17/51
SET domain-containing protein 7	RVYVADSLISSAGEGLF*p*SKV	1924.02	2	2.35	0.47	376.6	15/51
PREDICTED: myosin XVIIIb	KTRFG*p*SCESLLESGPCTRR	1924.03	2	2.32	0.49	609.3	19/48
PREDICTED: similar to dedicator of cytokinesis 7	RNPI*p*SVLKLSAAEEGDRL	1779.85	2	2.25	0.37	753.4	21/45
PREDICTED: hypothetical protein XP_130476	KD*p*SLGMSDEDGAKD	1305.23	2	2.21	0.18	229.9	14/33
PREDICTED: similar to Ifi203 protein	RRQDELS*p*SSESLFINKE	1833.86	2	2.40	0.12	284.8	14/42
aldehyde dehydrogenase family 5, subfamily A1	RKW*p*YDLMIQNKDDLAKI	1962.14	2	2.28	0.56	152.9	13/42
PREDICTED: similar to hypothetical protein	RIYD*p*YKPFKT	1154.22	2	2.21	0.17	280.8	12/21
PREDICTED: hypothetical LOC218513	RGDL*p*SGSLQTSLVVQDAPSLSPSRS	2395.49	3	2.83	0.21	859.5	34/132

Table A.8 Phosphorylated proteins identified in Blank incubation 30 min

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
PREDICTED: MYST histone acetyltransferase (monocytic leukemia) 3	KEEPQGQERELE*p*TRV	1681.62	2	2.32	0.18	407.8	19/36
heterogeneous nuclear ribonucleoprotein A/B	RNRGNRGSGGGQG*p*STNYGKS	1847.77	2	2.33	0.48	242.6	16/51
eukaryotic translation initiation factor 2 alpha kinase 3	KALESVNGENAIIPLP*p*TIKW	1960.14	2	2.25	0.50	216.7	14/51
ligase IV, DNA, ATP-dependent	RVEKDMKQQSL*p*YIETKL	1921.09	2	2.23	0.15	304.5	14/42
odd Oz/ten-m homolog 3	RLSEILYD*p*STRV	1277.29	2	2.20	0.37	625.4	14/27
sine oculis-related homeobox 6 homolog	REW*p*YLQDPYPNP*p*SKKR	1925.88	2	2.71	0.47	568.3	18/52
PREDICTED: similar to Crocc protein	RH*p*SELVTKEAADLRA	1549.59	2	2.65	0.47	522.3	15/36

Continues

Table A.8 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
cDNA sequence AF366264	RK*p*YHDE*p*SGTEIE*p*SEALGQEPKR	2488.24	3	3.21	0.15	487.0	35/190
PREDICTED: similar to Ttc5 protein	KPL*p*STLQPGVNSGTVVLGKV	1848.01	2	2.71	0.37	680.9	20/51
PREDICTED: similar to putative pheromone receptor	RENSDF*p*TKFIQ*p*TVNPKN	1928.88	2	2.79	0.10	323.2	18/56
stathmin 1	KDL*p*SLEEIQKK	1155.16	2	2.21	0.13	754.7	19/24
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54	**KAR*p*SAQTGARALILSPTRE**	**1850.00**	**2**	**2.32**	**0.10**	367.3	**14/48**
PREDICTED: nebulin	**KLQSQYL*p*YVELATKERP**	**1922.05**	**2**	**2.73**	**0.17**	**230.0**	**12/42**
transcobalamin 2	RID*p*SQLVEKLGQRL	**1466.55**	2	**2.54**	**0.49**	364.4	13/33
ubiquitin specific protease 7	RAGFIQDTSLIL*p*YEEVKP	1907.03	2	**2.29**	0.11	258.9	13/45
soluble adenylyl cyclase	R*p*YMEGQVLHLQKQ	1426.55	2	2.22	0.50	488.7	13/30
PREDICTED: cell division cycle 27 homolog	KGGLTQP*p*SIND*p*SLEITKL	1833.78	2	2.38	0.17	322.1	20/60
A-kinase anchor protein 3	KSESY*p*SLISTKSRAGDPKL	1906.95	2	2.27	0.11	227.4	12/48
heterogenous nuclear ribonucleoprotein U	-.MS*p*SSPVNVKKL	1157.25	2	2.22	0.12	306.4	12/27
delta isoform of regulatory subunit B56, protein phosphatase 2A	KY*p*SGGPQIVKKE	1157.22	2	2.23	0.10	321.7	12/27
coagulation factor V	KENRLDPPIVAR*p*YIRI	1792.94	2	2.27	0.38	228.3	14/39
PREDICTED: similar to neurexin III	RQ*p*TGEGVGPQE*p*Y*p*TLVKL	1846.66	2	2.53	0.33	433.0	24/70
PREDICTED: similar to protein kinase, lysine deficient 3	KELQEL*p*YERL	1160.14	2	2.46	0.25	422.1	11/21
olfactory receptor 866	KIVFSILKIPS*p*S*p*SGKYKA	1928.07	2	2.44	0.14	620.0	18/60
shroom	K*p*SLSSELLTC*p*SPQHRK	1718.67	2	2.44	0.16	282.8	14/52
IIIG9 protein	RYLTTYNQG*p*YFENIPKG	1932.00	2	2.23	**0.19**	293.6	14/42
PDZ domain containing 1	KG*p*SNG*p*YGFYLRAGPEQKG	1904.82	2	2.23	0.14	473.1	19/60
SNF1-like kinase	RIPFFM*p*SQDCETLIRRM	1937.16	2	2.27	0.17	195.7	11/42
DEAD (Asp-Glu-Ala-Asp) box polypeptide 4	RKPAASDSGNGD*p*TYQ*p*SRS	1814.61	2	2.77	0.34	603.6	22/60
unc-13 homolog B	KITI*p*TVVCAQGLQAKD	1525.72	2	2.36	0.19	261.2	13/39
trinucleotide repeat containing 9	RASLVSKAAAE*p*SAEAQ*p*TIRS	1963.93	2	2.21	0.53	195.0	18/68
PREDICTED: similar to L-NAME induced actin cytoskeletal protein	KTVEEKL*p*THRQ	1193.22	2	2.26	0.18	479.5	13/24
glucosaminyl (N-acetyl) transferase 2 isoform B	K*p*YVLNTCGQDFPLKTNKE	1922.08	2	2.28	0.12	272.0	17/45
protein phosphatase 1K (PP2C domain containing)	KI*p*SLENVGCA*p*SLIGKRKE	1848.95	2	2.26	0.26	241.9	16/60
PREDICTED: RIKEN cDNA D630013G24 gene	**KLLST*p*YGAEELYRAKR**	**1694.79**	**3**	**2.93**	**0.16**	393.4	21/78
PREDICTED: similar to novel protein	**-.MALGAEL*p*SE*p*TLVEKP**	**1651.67**	**2**	**2.27**	**0.10**	403.3	19/52

Continues

Table A.8 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
PREDICTED: similar to HMG-1	KKQPDASVNFSEF*p*SKKC	1792.85	2	2.34	0.23	415.1	14/42
guanylate cyclase activator 2b	RTIATDECELCINVAC*p*TGC	1940.11	2	2.23	0.17	255.6	12/51
telomerase reverse transcriptase	RPIVNMSY*p*SMGTRALGRR	1834.04	2	2.46	0.32	374.4	17/45
leucine rich repeat containing 16	KPR*p*TA*p*SRPEDTPDSP*p*SGPSSPKV	2407.18	3	2.87	0.13	680.7	46/200
phospholipase A2, group IB, pancreas, receptor	**KNCAV*p*YKANKTLLPSNCA*p*SKH**	**2186.30**	**3**	**2.84**	**0.13**	**712.0**	**38/144**
natriuretic peptide receptor 3	**RFQMR*p*SNVKYPWGPLKL**	**1932.16**	**2**	**2.63**	**0.15**	**499.6**	**17/42**
serine/arginine-rich protein specific kinase 2	R*p*SR*p*TVSASSTGDLPKT	1566.46	2	2.32	0.10	157.9	19/52
hypothetical protein LOC234912	RLYDENIVRE*p*SGHIVKC	1932.92	2	2.28	0.11	300.8	17/56
PREDICTED: similar to OTTHUMP00000028561	R*p*SLKEAPQSD*p*SVGQQAGRS	1918.80	2	2.25	0.19	350.0	16/64
alpha thalassemia/mental retardation syndrome X-linked homolog	K*p*SA*p*S*p*SSSDAEGSSEDNKKQ	1926.53	2	2.23	0.13	186.5	16/80
phospholipase C, gamma 2	KPTEA*p*SADQLP*p*SP*p*SQLRE	1937.73	1	1.83	0.14	153.8	16/75
calpain 12	REV*p*TLLPDSLSQRW	1438.49	2	2.88	0.18	569.6	16/33
PREDICTED: myosin XVIIIb	KTRFG*p*SCESLLESGPCTRR	1924.03	2	2.42	0.45	511.9	18/48
mutS homolog 6	RLL*p*SKIHNVG*p*SPLKS	1566.63	2	2.31	0.12	205.7	17/48
1-phosphatidylinositol-4-phosphate 5-kinase	R*p*TAVQLR*p*SL*p*STVLKR	1656.59	2	2.29	0.48	333.2	18/60
PREDICTED: pleckstrin homology domain containing, family G (with RhoGe	KEN*p*SGQSPLYNSLGRK	1602.57	2	2.25	0.60	430.0	16/39
PL10 protein	KQYPISLVLAP*p*TRE	1438.58	2	2.69	0.23	573.7	15/33
suppressor of K+ transport defect 3	REGEVMKLLK*p*T*p*SE*p*TKY	1833.77	2	2.42	0.39	454.6	24/65
PREDICTED: similar to Casitas B-lineage lymphoma b [Mus musculus]	KI*p*YID*p*SLMKK*p*SKR	1566.53	2	2.27	0.15	183.1	16/50
PREDICTED: Na+/K+ -ATPase alpha 4 subunit [Mus musculus]	R*p*TSP*p*YAMLPLP*p*YQP	1718.64	2	2.21	0.19	403.4	16/60
kinesin family member 1B isoform a [Mus musculus]	KSLTTLGKVI*p*SALAEVSKKK	1926.17	2	2.45	0.44	682.3	19/51
PREDICTED: similar to BC037112 protein [Mus musculus]	RIDAYVDEQM*p*TMK*p*TKQ	1833.87	2	2.36	0.16	343.9	18/52

Table A.9 Phosphorylated proteins identified in cCMP incubation 1 min

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
growth associated protein 43	KEGDGSATTDAAPATSPKAEEP*p*SKA	2298.24	3	2.85	0.16	606.0	31/132
PREDICTED: similar to Gm397 protein	KMFNSWLQPEKQ*p*TKE	1717.85	2	2.68	0.10	372.8	14/36
hypothetical protein LOC70909	K*p*SFQQQLCDAIAIIKG	1658.82	2	2.38	0.44	305.5	11/39
zinc finger, DHHC domain containing 5	KDEVQLK*p*T*p*TYSKS	1472.38	2	2.26	0.15	800.6	18/40
unc-13 homolog B	KITI*p*TVVCAQGLQAKD	1525.72	2	2.34	0.17	359.9	15/39
PREDICTED: similar to Crocc protein	RH*p*SELVTKEAADLRA	1549.59	2	2.26	0.50	530.4	16/36
jumonji domain containing 2A	MASE*p*SETLNPSARI	1473.47	2	2.62	0.11	730.2	16/36
PREDICTED: similar to NALP-beta	RHLEI*p*SVNDLKDEGLKI	1790.88	2	2.51	0.40	156.7	11/42
SEC24 related gene family, member C	KFA*p*YRAVLN*p*SPVKT	1525.54	2	2.66	0.18	715.9	23/44
PREDICTED: RIKEN cDNA D630013G24 gene	KLLST*p*YGAEELYRAKR	1694.79	3	2.87	0.12	646.3	25/78
protein phosphatase 1, regulatory (inhibitory) subunit 1C	KSE*p*SLDEEEKLELQRR	1785.77	2	2.38	0.31	366.8	24/39
formin homology 2 domain containing 1	RTLKSGLGDDLVQALGL*p*SKA	1896.06	2	2.36	0.18	466.0	20/34
PREDICTED: similar to Gm444 protein	R*p*TPSLLGAD*p*TPVVKN	1458.44	2	2.40	0.19	428.8	19/48
multiple endocrine neoplasia 1	RSWLYLKG*p*SYMRCDRK	1859.05	2	2.31	0.36	381.4	16/39
PREDICTED: similar to hypothetical protein 4932415M13	KG*p*SRDP*p*TAGSNVDNWRQ	1792.61	2	2.21	0.43	288.1	18/56
hypothetical protein LOC235472	KV*p*TLIEDDTALV*p*SWKP	1750.73	2	2.20	0.19	444.4	17/52
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54	KAR*p*SAQTGARALIL*p*SPTRE	1929.96	2	2.34	0.17	386.0	20/64
PREDICTED: hypothetical protein XP_622534	RYK*p*YLYYEELSIRKS	1849.00	2	2.25	0.35	637.4	17/36
myosin Ixb	RQNKHIQSCREE*p*S*p*STHRE	2101.01	3	3.04	0.41	874.1	33/120
solute carrier family 22 (organic anion transporter), member 19	RKRICLL*p*SFVRF	1315.54	2	2.50	0.16	486.8	14/27
cerebral endothelial cell adhesion molecule 1	KLSWDLI*p*YLGRK	1316.41	2	2.34	0.16	336.8	13/27
stathmin 1	RSKESVPDFPL*p*SPPKK	1608.70	2	2.79	0.40	1020.5	25/39
shroom	K*p*SLSSELLTC*p*SPQHRK	1718.67	2	2.26	0.27	577.1	19/52
PREDICTED: hepatitis B virus x associated protein	RSVD*p*SGKCSAASALEEPAVKT	1930.01	2	2.33	0.16	252.8	16/54
spir-2 protein	R*p*SRQTQ*p*SLYIPN*p*TRTLNFQ	2408.30	3	3.26	0.17	560.4	42/170
PREDICTED: myeloid/lymphoid or mixed-lineage leukemia	KNIRQFIMPVVSAI*p*SSRI	1899.13	2	2.36	0.16	508.9	17/45
CEA-related cell adhesion molecule 14	K*p*SVTSKDSGYY*p*TLRILDANSRP	2407.38	3	2.95	0.10	446.7	30/152
PREDICTED: similar to cytochrome P450 monooxygenase CYP2T1	RLLNLF*p*SDNFCIM*p*SSRW	1920.99	2	2.39	0.42	253.3	15/56

Continues

Table A.9 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
supervillin	KE*p*SDR*p*YP*p*SGSEIPVVEDEEKV	2406.14	3	2.86	0.15	784.1	43/180
PREDICTED: similar to Ubn1 protein	KF*p*SKAGF*p*TALNA*p*SKEKK	1839.71	2	2.40	0.16	375.7	20/70
tenascin N	KKAN*p*TEGHTDID*p*SPKN	1673.52	2	2.25	0.16	315.6	16/52
hypothetical protein LOC72543	KV*p*TRYLCFTR*p*SF*p*SKE	1848.79	2	2.24	0.13	307.7	17/60
PREDICTED: similar to Ab2-404	**KLK*p*SY*p*SEPEKI**	**1321.10**	**2**	**2.22**	**0.15**	**353.2**	**20/40**
PREDICTED: similar to caterpillar 16.2	**KTPLEF*p*SN*p*TMACSDPKKE**	**1929.95**	**1**	**1.99**	**0.22**	**166.3**	**15/60**
argininosuccinate lyase	KGLP*p*S*p*TYNKDLQEDKE	1848.63	2	2.48	0.44	578.3	22/65

Table A.10 Phosphorylated proteins identified in cCMP incubation 5 min

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
small glutamine-rich tetratricopeptide repeat (TPR) containing protein	RAPDRTPP*p*SEEDSAEAERL	1937.84	3	3.10	0.33	887.5	32/96
heat shock protein 1, beta	KIEDVG*p*SDEEDDSGKDKK	1818.67	2	2.52	0.17	290.0	22/45
hypothetical protein LOC70385	K*p*YQSLKKQNAFTRD	1564.65	1	1.81	0.13	256.6	13/33
jumonji domain containing 2A	MA*p*SESETLNPSARI	1473.47	2	2.45	0.16	642.4	16/36
PREDICTED: similar to Crocc protein	RH*p*SELVTKEAADLRA	1549.59	2	2.33	0.52	325.7	12/36
transmembrane protein 24 [Mus musculus]	K*p*TPTKRSTLIISGV*p*SKV	1748.81	2	2.25	0.19	261.9	19/56
shroom	K*p*SLSSELLTC*p*SPQHRK	1718.67	2	2.35	0.17	254.0	14/52
PREDICTED: similar to protein kinase, lysine deficient 3	KELQEL*p*YERL	1160.14	2	2.32	0.40	432.3	11/21
rho/rac guanine nucleotide exchange factor (GEF) 18	RIIQNTEAG*p*TEDYKD	1562.54	1	2.31	0.13	613.2	18/36
unc-13 homolog B	KITI*p*TVVCAQGLQAKD	1525.72	2	2.68	0.10	365.8	14/39
PREDICTED: similar to enolase 1, alpha non-neuron	KYDLDFK*p*SPDDPSRY	1635.59	2	2.78	0.27	1392.0	24/36
PREDICTED: similar to Gamma-aminobutyric-acid receptor rho-3 subunit p	KNEQEA*p*TYARD	1162.07	2	2.49	0.18	607.9	15/24
protein phosphatase 1K (PP2C domain containing)	KI*p*SLENVGCA*p*SLIGKRKE	1848.95	2	2.29	0.17	425.5	19/60
zinc finger, DHHC domain containing 5	**KDEVQLKTT*p*Y*p*SKS**	**1472.38**	**2**	**2.38**	**0.18**	**743.8**	**18/40**
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54	**KAR*p*SAQTGARALIL*p*SPTRE**	**1929.96**	**2**	**2.39**	**0.19**	**420.4**	**19/64**

Continues

Table A.10 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
PREDICTED: RIKEN cDNA D630013G24 gene	KLLST*p*YGAEELYRAKR	1694.79	3	2.87	0.14	654.4	24/78
PREDICTED: marapsin 2	K*p*TMKNVCEGDSG*p*SPLVCKQ	1928.99	2	2.26	0.24	400.8	21/64
PREDICTED: G protein-regulated inducer of neurite outgrowth 1	RL*p*SSGQAERVSLVKTETLSSGKE	2258.40	3	2.85	0.22	473.3	32/120
PREDICTED: Na+/K+ -ATPase alpha 4 subunit	R*p*TSPYAMLPLP*p*YQP	1718.64	2	2.36	0.13	587.6	21/60
PREDICTED: gene model 431	KV*p*SGEANETQIAEALRR	1668.67	1	2.06	0.23	166.5	14/42
solute carrier family 22 (organic anion transporter), member 19	RKRICLL*p*SFVRF	1315.54	2	2.78	0.28	446.6	13/27
calcium channel, voltage-dependent, R type, alpha 1E subunit	K*p*STKVDGA*p*S*p*YFRHKE	1736.55	2	2.34	0.19	443.4	22/60
stathmin 1	KDL*p*SLEEIQKK	1155.16	2	2.31	0.12	680.5	17/24
PREDICTED: similar to calcium-sensing receptor related protein 3	R*p*YNNDFLG*p*TLAFKL	1563.50	2	2.29	0.11	411.5	17/44
PREDICTED: similar to mKIAA1741 protein [Mus musculus]	RLAAG*p*TLDWGGSGGRG*p*SPGKF	1904.82	1	1.81	0.13	297.0	18/72
cullin 4B	RLI*p*TYLDQ*p*T*p*TQKS	1564.40	2	2.28	0.16	269.0	16/50
MAP-kinase activating death domain isoform 8	REKT*p*TPFP*p*SLKGNRR	1635.61	2	2.23	0.32	493.2	10/24
hypothetical protein LOC73598	RPKG*p*YQLLLPSTTNRQ	1668.80	1	2.21	0.13	178.9	17/39
solute carrier organic anion transporter family, member 2a1	RE*p*Y*p*SLQENA*p*SGLI	1564.32	1	1.93	0.13	241.7	19/55
lymphoid nuclear protein related to AF4-like	RIPGKP*p*YKETEPPKG	1564.69	1	1.85	0.13	323.0	14/36
eukaryotic translation initiation factor 4A1	RSGS*p*SRVLITTDLLARG	1669.79	1	1.82	0.24	169.2	13/42

Table A.11 Phosphorylated proteins identified in cCMP incubation 15 min

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
PREDICTED: similar to NALP-beta	RHLEI*p*SVNDLKDEGLKI	1790.88	2	2.71	0.43	272.0	16/42
zinc finger, DHHC domain containing 5	KDEVQLK*p*TT*p*YSKS	1472.38	2	2.41	0.18	606.8	16/40
evolutionarily conserved G-patch domain containing	KKED*p*SISEFL*p*SQARS	1670.56	2	2.65	0.14	651.9	16/24
jumonji domain containing 2A	MASE*p*SETLNPSARI	1473.47	2	2.95	0.18	947.7	18/36
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54	KAR*p*SAQTGARALILSPTRE	1850.00	2	2.67	0.13	519.6	18/48
protein kinase, interferon-inducible double stranded RNA dependent	KIGQ*p*TMYGTGSGVTKQEAKQ	1937.04	2	2.29	0.10	269.1	18/34

Continues

Table A.11 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
PREDICTED: PR-domain zinc finger protein	KILYTL*p*TKVTELRK	1530.71	3	3.15	0.18	864.3	25/66
PREDICTED: similar to BC037112 protein	RIDAYVDEQM*p*TMK*p*TKQ	1833.87	2	2.35	0.38	550.3	20/52
hypothetical protein LOC68423	RSEKME*p*SVSGYEAKV	1525.54	2	2.33	0.14	348.1	15/36
unc-13 homolog B [Mus musculus]	KITI*p*TVVCAQGLQAKD	1525.72	2	2.22	0.19	266.1	14/39
gene rich cluster, C2f gene [Mus musculus]	**KIG*p*T*p*SF*p*SVEDISDIRE**	**1779.57**	**2**	**2.62**	**0.44**	**529.7**	**20/65**
angiotensinogen	**R*p*SLDL*p*STDPVLATEKI**	**1649.58**	**2**	**2.25**	**0.17**	**380.1**	**18/52**
PREDICTED: pleckstrin homology domain containing, family G (with RhoGe	KEN*p*SGQSPLYNSLGRK	1602.57	2	2.24	0.56	320.3	16/39
lamin B2	RMRVESL*p*S*p*YQLLGLQKQ	1926.03	2	2.35	0.49	896.2	16/27
WD repeat domain 51A	KDAV*p*TCVNF*p*SPSGHLLASGSRDKT	2422.42	3	2.85	0.41	480.9	42/168
hypothetical protein LOC107372	MLVS*p*SVSHALMNIPEA*p*SLTRQ	2317.48	3	3.52	0.15	471.1	35/152
PREDICTED: calmodulin binding transcription activator 1	R*p*SEPSNYY*p*STEGHKD	1659.45	2	2.74	0.23	175.3	17/48
argininosuccinate lyase	KGLP*p*S*p*TYNKDLQEDKE	1848.63	2	2.27	0.27	494.6	20/65
F-box protein 18	RYTA*p*TTKCSPSVDPERV	1735.78	2	2.20	0.18	196.7	19/42
UDP-glucuronate decarboxylase 1	K*p*TLK*p*TN*p*TIGTLNMLGLAKRV	2186.25	3	2.83	0.15	580.6	39/170
Cdc42 binding protein kinase beta	K*p*TS*p*SLLIL*p*TENENEKR	1831.64	2	2.45	0.40	239.8	16/65
hypothetical protein LOC269693	R*p*TILVE*p*SEIQKK	1320.28	2	2.32	0.16	282.6	16/36
type II transmembrane serine protease	KPGV*p*YTRVTAFRHWIASN*p*TGI	2407.48	2	2.22	0.18	155.6	16/76
PREDICTED: similar to hypothetical protein E330017N17	KCGIY*p*YV*p*YDLE*p*TQRCRF	2123.04	3	2.82	0.29	939.7	39/140
retinoic acid induced 3	RAQAPA*p*SP*p*YND*p*YEGRK	1779.49	2	2.71	0.40	535.3	21/65
hypothetical protein LOC70909	KKVD*p*SFAKES*p*S*p*SPKP	1650.45	2	2.29	0.35	358.6	21/60
microtubule associated serine/threonine kinase 1	RSSSGEAG*p*TPPVPIVVEPARP	1931.02	2	2.28	0.24	288.4	17/54
zinc finger protein 609	RFCDSP*p*TSDLEMRN	1481.51	2	2.20	0.42	336.2	17/33
PREDICTED: similar to PI-3-kinase-related kinase SMG-1 isoform 1 [Mus	RHDD*p*TRGHADIQNDEKG	1831.72	2	2.23	-	419.7	17/42
PREDICTED: similar to laminin, beta 4 [Mus musculus]	REHPGALAG*p*TQAELRE	1530.55	2	2.31	0.53	595.3	17/39
PREDICTED: similar to armadillo repeat containing, X-linked 4	R*p*TQ*p*SEARLDAAVDTKE	1665.55	2	2.30	0.26	303.4	15/52
Lutheran blood group (Auberger b antigen included)	RGWK*p*SSSLMVKVTSALSRE	1918.13	3	2.86	0.13	550.6	31/96
soluble adenylyl cyclase	**R*p*YMEGQVLHLQKQ**	**1426.55**	**2**	**2.20**	**0.50**	**473.3**	**12/30**

Continues

Table A.11 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
natriuretic peptide receptor 2	RYHKGAGSRLTL*p*SLRG	1639.76	2	2.21	-	284.8	13/39
PREDICTED: myosin XVIIIb	K*p*TRFGSCESLLESGPCTRR	1924.03	2	2.63	0.54	812.5	22/48
PREDICTED: filamin C, gamma	K*p*TPCEEVYVKH	1148.19	2	2.21	0.32	712.5	16/24
PREDICTED: similar to Wolf-Hirschhorn syndrome candidate 1							
protein iso	**KGIG*p*TPPN*p*T*p*TPIKN**	**1436.28**	**1**	**2.02**	**0.11**	**325.9**	**20/55**
supervillin	**KESDR*p*YP*p*SG*p*SEIPVVEDEEKV**	**2406.14**	**3**	**3.12**	**0.13**	**777.0**	**44/180**
RIKEN cDNA 5930412E23	**KLGMLSVLSTL*p*SGTIAIKH**	**1785.06**	**2**	**2.24**	**0.39**	237.2	14/48
PREDICTED: similar to protein kinase, lysine deficient 3	KELQEL*p*YERL	1160.14	2	2.22	0.10	360.1	10/21
PREDICTED: similar to ribosomal protein L34	R*p*TPGNRIV*p*YL*p*YTKK	1665.56	2	2.22	0.34	426.0	20/55

Table A.12 Phosphorylated proteins identified in cCMP incubation 30 min

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
growth associated protein 43	KEGDGSATTDAAPAT*p*SPKA	1656.57	2	2.37	0.11	687.1	21/48
potassium channel, subfamily K, member 3	RYNL*p*SEGG*p*YEELERVVLRL	2187.16	3	3.22	0.14	586.4	34/128
odd Oz/ten-m homolog 3	RLSEILYD*p*STRV	1277.29	2	2.20	0.38	489.1	14/27
coagulation factor V	KENRLDPPIVAR*p*YIRI	1792.94	2	2.76	0.42	493.2	17/39
heterogeneous nuclear ribonucleoprotein A/B	RNRGNRGSGGGQGS*p*TNYGKS	1847.77	2	2.24	0.42	273.7	17/51
PREDICTED: similar to fat3 [Mus musculus]	R*p*Y*p*SLVND*p*YNGRF	1441.17	2	2.45	0.19	433.8	20/45
acid phosphatase 5, tartrate resistant	RFQE*p*TFEDVFSDRALRN	1940.97	2	2.40	0.14	297.9	16/42
PREDICTED: MYST histone acetyltransferase (monocytic leukemia) 3	KEEPQGQERELE*p*TRV	1681.62	2	2.31	0.26	421.1	19/36
PREDICTED: similar to Crocc protein	RH*p*SELVTKEAADLRA	1549.59	2	2.21	0.55	532.2	16/36
HLA-B associated transcript 2	RPGTGGKTG*p*SGGGSSGGGGA GPGGRT	1910.82	2	2.34	0.16	516.9	22/69
heat shock protein 4	KLKVLATAFD*p*TTLGGRK	1643.79	2	2.38	0.13	462.6	16/42
PREDICTED: similar to BC037112 protein	**RIDAYVDEQM*p*TMK*p*TKQ**	**1833.87**	**2**	**2.30**	**0.39**	325.9	18/52
deleted in azoospermia-like	**RNSLVTQDD*p*YFKDKR**	**1653.65**	**2**	**2.53**	**0.24**	**306.6**	**13/24**

Continues

Table A.12 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
PREDICTED: hypothetical protein LOC68621	RLFCIAGDA*p*TVTIAHRYTPRE	2186.40	3	3.04	0.19	479.2	31/108
cathepsin 3 precursor	RKEMIEIPVP*p*TVKK	1464.68	2	2.45	0.39	574.0	14/33
serologically defined colon cancer antigen 13	RHKGPGR*p*TGGLVISRP	1515.63	2	2.41	0.13	210.9	15/26
PREDICTED: cell division cycle 27 homolog	KGGLTQPSIND*p*SLEI*p*TKL	1833.78	2	2.20	0.44	320.8	18/60
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54	**KAR*p*SAQTGARALILSPTRE**	**1850.00**	**2**	**2.49**	**0.11**	**486.0**	**17/48**
PREDICTED: similar to Gm397 protein	**KMFNSWLQPEKQ*p*TKE**	**1717.85**	**2**	**2.20**	**0.19**	**206.7**	**12/36**
adaptor-related protein complex 2, beta 1 subunit	**RIQPGNPN*p*YTLSLKC**	**1525.61**	**2**	**2.39**	**0.17**	328.2	13/36
evolutionarily conserved G-patch domain containing	KKED*p*SISEFL*p*SQARS	1670.56	2	2.65	0.14	651.9	21/48
hypothetical protein LOC320696	RIFPVEDQLETLK*p*SESQNKI	2186.29	3	3.04	0.16	498.6	30/102
PL10 protein	KQYPISLVLAP*p*TRE	1438.58	2	2.49	0.20	496.6	14/33
PREDICTED: empty spiracles homolog 1	RKQLAGSLSLSE*p*TQVKV	1669.78	2	2.28	0.15	537.1	18/42
heat shock protein 1, beta	KIEDVG*p*SDEEDDSGKDKK	1818.67	2	2.56	0.24	440.3	24/45
heat shock protein 105	KLKVLG*p*TAFDPFLGGKN	1643.83	2	2.66	0.12	680.8	19/42
PREDICTED: RIKEN cDNA D630013G24 gene	KLL*p*STYGAEELYRAKR	1694.79	3	2.96	0.10	411.2	23/78
DVL-binding protein DAPLE [Mus musculus]	KENQQL*p*SKKIEKL	1425.49	2	2.61	0.38	596.9	17/30
phosphodiesterase 4D, cAMP specific	RPM*p*SQISGVKKL	1155.27	2	2.43	0.21	392.6	15/27
myosin IXb	RQNKHIQSCREE*p*S*p*STHRE	2101.01	3	3.34	0.35	543.4	28/120
PREDICTED: similar to diacylglycerol kinase, eta isoform 2	RSRTKNLMW*p*YGVLGTRE	1863.06	2	2.23	0.18	237.8	11/42
olfactory receptor 866	KIVFSILKIPS*p*SSGK*p*YKA	1928.07	2	2.48	0.45	768.5	21/60
PREDICTED: hypothetical protein LOC71841	RP*p*YLQEHSIQYERA	1643.66	2	2.61	0.18	510.9	15/33
cystathionase	KAGDEIICMDEVYGG*p*TNRY	1923.98	2	2.28	0.51	228.2	16/48
argininosuccinate lyase	KGLP*p*S*p*T*p*YNKDLQEDKE	1848.63	2	2.27	0.27	494.6	20/65
PHD finger protein 1	RW*p*TDGLL*p*YLGTIKKV	1668.72	2	2.22	0.18	340.4	19/48
unc-13 homolog B	KI*p*TITVVCAQGLQAKD	1525.72	2	2.47	0.19	293.8	14/39
calpain 12	REV*p*TLLPDSLSQRW	1438.49	2	2.51	0.28	485.9	15/33
PREDICTED: similar to hypothetical protein FLJ25801	R*p*SLQ*p*TMITDLEKN	1439.42	2	2.42	0.26	396.0	16/40
bassoon	RPYV*p*SGGVTAVPLTSL*p*TRV	1878.91	2	2.38	0.24	336.0	21/64
ATPase, H+ transporting, V1 subunit C, isoform 2	REF*p*YYDEKE	**1153.97**	**2**	**2.33**	**0.15**	288.6	11/24
PREDICTED: hypothetical protein LOC68964	**R*p*SSTSVQVL*p*SFPPE*p*TKA**	**1847.69**	**2**	**2.48**	**0.44**	**510.2**	23/70

Continues

Table A.12 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
pleckstrin homology, Sec7 and coiled-coil domains, binding protein	R*p*SS*p*SLGDFSWSQRKV	1725.48	2	2.39	0.25	212.0	16/60
protein kinase, interferon-inducible double stranded RNA dependent	KIGQ*p*TMYGTGSGVTKQEAKQ	1937.04	2	2.29	0.10	269.1	16/51
FMS-like tyrosine kinase 1	KESCCSPPPDYNSVVL*p*S*p*SPPA	2373.36	3	3.48	0.12	827.3	37/160
PREDICTED: hypothetical protein XP_622534	RYKYL*p*YYEELSIRKS	1849.00	2	2.46	0.49	549.1	16/36
arrestin domain containing 1	K*p*TGNVVLTA*p*STDLRG	1587.40	2	2.45	0.51	560.8	14/24
death inducer-obliterator 1 isoform 2	KPTAL*p*SPTLLSKSMKDDRR	1941.12	2	2.38	0.16	210.9	14/32
inhibitor of kappaB kinase epsilon	KKSGEVVAVKVFN*p*SASYRR	1922.05	2	2.37	0.51	382.6	16/48
leukotriene A4 hydrolase	KS*p*SALQWLTPEQT*p*SGKQ	1793.72	2	2.32	0.30	595.0	18/56
transmembrane protein 24	RS*p*S*p*SCGDAELLGQA*p*TLPVGS PSRP	2373.23	3	3.51	0.11	882.8	47/210
PREDICTED: similar to mKIAA0136 protein	K*p*SEQ*p*SH*p*TEQSGIHVDLVSSPKP	2406.19	3	3.16	0.11	623.0	41/190
stathmin 1	KDL*p*SLEEIQKK	1155.16	2	2.73	0.11	684.4	18/24
serine/threonine kinase 7	RNLLK*p*TDDLSNDVCAVLKL	1942.10	2	2.51	0.15	302.8	13/32
PREDICTED: calpain 11	KI*p*SLPEVDDPEDD*p*SEKN	1848.66	2	2.39	0.10	283.7	16/56
phenylalkylamine Ca2+ antagonist (emopamil) binding protein	KHL*p*TSAQ*p*SVLD*p*SKVMKI	1884.82	2	2.36	0.18	377.1	20/70
PREDICTED: RIKEN cDNA 1810009A16	KNKDVPVEAL*p*T*p*TVKP	1574.56	2	2.30	0.11	756.1	21/48
myeloid/lymphoid or mixed lineage-leukemia translocation to 1 homolog	RTLQTEDSNSEDEA*p*SFRS	1909.78	2	2.29	0.18	417.0	15/30
nuclear receptor subfamily 1, group H, member 3	RYNPGSESI*p*TFLKD	1436.47	1	2.28	0.11	548.1	17/33
THUMP domain containing 3	RTASVDIIV*p*TDMPFGKRM	1831.01	2	2.25	0.12	432.8	17/45
annexin A6	REFIEK*p*YDKS	1152.16	2	2.25	0.29	339.9	14/21
translocase of outer mitochondrial membrane 40 homolog	R*p*SLCRRSAALGTG*p*SS*p*TGRG	1920.77	2	2.25	0.12	238.6	18/80
PREDICTED: hypothetical protein LOC66498	REYP*p*SEQIIVTEKT	1516.56	2	2.23	0.10	260.3	16/33
hypothetical protein LOC235472	KV*p*TLIEDDTALV*p*SWKP	1750.73	2	2.24	0.16	402.8	17/52
shroom	K*p*SLSSELL*p*TCSPQHRK	1718.67	2	2.23	0.10	388.8	17/52
PREDICTED: similar to ribosomal protein L38	KMNQI*p*SACVGGI*p*S*p*SFGPRV	1964.89	2	2.43	0.55	163.6	20/80
small conductance calcium-activated potassium channel protein 3	RSEDLEKQIGSLE*p*SKL	1643.66	2	2.42	0.46	485.9	20/39
PREDICTED: zinc finger protein 236	KCQYCMKSFS*p*T*p*SG*p*SLKV	1910.86	2	2.37	0.43	172.5	20/70
meteorin, glial cell differentiation regulator-like	KPFRD*p*S*p*SGANIYLEKT	1757.69	2	2.27	0.26	159.8	13/26

Continues

Table A.12 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
X-ray repair complementing defective repair in Chinese hamster cells 1	KLDLSLEDRKPP*p*SKP	1578.67	2	2.27	0.48	462.4	14/36
PREDICTED: ectonucleotide pyrophosphatase/phosphodiesterase 3	K*p*T*p*YLPTFE*p*TII	1438.29	1	2.01	0.13	340.0	18/45
nuclear autoantigenic sperm protein	KEAAIPGLNEDEVA*p*SGKT	1680.68	2	2.26	0.17	298.8	16/45
PREDICTED: hypothetical protein XP_619738	**KCSNLWLFLE*p*GQLPKD**	**1830.02**	**2**	**2.21**	**0.19**	**315.2**	**15/42**
flotillin 1	**KITLVS*p*SG*p*SGTMGAAKV**	**1540.53**	**2**	**2.20**	**0.18**	**222.9**	**17/56**
low density lipoprotein receptor-related protein 1	**RILQEDFTCRAVNS*p*SCRA**	**1923.04**	**2**	**2.48**	**0.19**	**390.1**	**17/30**
bagpipe homeobox gene 1 homolog	RADLAASLKLTE*p*TQVKI	1668.79	2	2.26	0.16	348.6	15/28
general transcription factor II I	RFVIKKHELLNS*p*TRE	1665.84	2	2.24	0.28	337.6	15/36
PREDICTED: glutamyl-prolyl-tRNA synthetase	KPVSA*p*TGAEDKDKKK	1426.44	2	2.23	0.37	371.9	17/36
teashirt 3	KLYG*p*SIFTGASKFRCKD	1758.95	2	2.21	0.30	186.3	12/24
vacuolar protein sorting 11	KL*p*TLL*p*TDPP*p*TARL	1438.29	1	1.95	0.12	548.7	22/50
ubiquitin specific protease 53	KAKSDSGTGYE*p*TDS*p*SQD*p*SRDKG	2375.00	3	3.40	0.13	1027.8	48/190
telomerase reverse transcriptase	RPIVNMSYSMG*p*TRALGRR	1834.04	2	2.30	0.49	299.3	16/45
centromere/kinetochore protein zw10 homolog	KMLK*p*SL*p*SMELTVQKQ	1668.81	2	2.23	0.16	362.7	12/24
lymphoid nuclear protein related to AF4-like	K*p*TEP*p*SKFPFPTKE	1439.40	2	2.23	0.18	371.2	17/40

Table A.13 Phosphorylated proteins identified in cGMP incubation 1 min

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
PREDICTED: similar to hypothetical protein 4932415M13	KG*p*SRDPTAG*p*SNVDNWRQ	1792.61	2	2.39	0.48	367.9	21/56
jumonji domain containing 2A	-.MA*p*SESETLNPSARI	1473.47	2	2.58	0.13	915.9	18/36
zinc finger, DHHC domain containing 5	KDEVQLKTT*p*Y*p*SKS	1472.38	2	2.29	0.18	916.1	19/40
centromere autoantigen H	K*p*SNIEMK*p*SLAL*p*SRMKL	1848.85	2	2.60	0.21	599.3	24/65
odd Oz/ten-m homolog 3	RLSEILYD*p*STRV	1277.29	2	2.30	0.38	580.6	14/27
hypothetical protein LOC209334	RIS*p*YLTQMNVKL	1277.40	2	2.21	0.39	250.9	10/27
heat shock protein 105	**KLKVLG*p*TAFDPFLGGKN**	**1643.83**	**2**	**2.41**	**0.19**	**622.5**	**19/42**
PREDICTED: hypothetical protein LOC71841	**RP*p*YLQEHSIQYERA**	**1643.66**	**2**	**2.45**	**0.12**	**530.1**	**14/33**

Continues

Table A.13 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
proteasome 26S non-ATPase subunit 8	KPED*p*STIPFTELAKQ	1528.57	2	2.33	0.14	237.0	12/36
EGF-like module containing, mucin-like, hormone receptor-like sequence	KTHTLGVL*p*SEFKSKE	1527.63	2	2.23	0.18	286.4	13/36
potassium voltage gated channel, Shab-related subfamily, member 1	K*p*TQ*p*SQPILN*p*TKE	1370.18	2	2.22	0.21	281.8	17/45
phospholipase A2, group X	**KYLFFP*p*SILCEKD**	**1440.61**	**2**	**2.26**	**0.38**	**445.5**	**16/30**
CAR-like membrane protein	**RARLVKPS*p*SS*p*SSG*p*SRS**	**1659.47**	**2**	**2.26**	**0.38**	**194.4**	**20/65**
RIKEN cDNA C330046G03	**KNE*p*YLLSLLAEERD**	**1530.58**	**3**	**2.91**	**0.11**	**848.1**	**28/66**
PREDICTED: hypothetical protein LOC69743	KQDVIRH*p*YNMHKKR	1649.78	2	2.29	0.31	298.5	13/33
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54	KAR*p*SAQTGARALIL*p*SPTRE	1929.96	2	2.41	0.17	582.4	24/64
baculoviral IAP repeat-containing 6	RL*p*SYLLPSARPELGVGPGRS	1963.15	2	2.31	0.19	293.1	15/51
splicing factor, arginine/serine-rich 16 isoform L	RSSS*p*SSASRTSSSR*p*SSSRS	1864.63	2	2.22	0.10	247.3	17/64
dynein, axonemal, heavy chain 5	KPVR*p*TDLN*p*YIAAVDLKT	1848.88	2	2.35	0.11	278.6	14/56
PREDICTED: RIKEN cDNA D630013G24 gene	KLL*p*STYGAEELYRAKR	1694.79	3	3.37	0.12	585.2	25/78
PREDICTED: similar to CDNA sequence BC016548	RPADSTNMDPWEL*p*SKTEPPTKE	2325.42	3	2.94	0.16	694.5	30/114
hypothetical protein LOC234069	RAIDSSNLKDDY*p*STAQRI	1864.83	2	2.31	0.18	509.4	15/45
PREDICTED: similar to protease inhibitor H; acrosin inhibitor	KVFCTRE*p*SDPLCGSDGQ*p*TYGNKC	2438.40	3	2.83	0.11	376.1	35/160
olfactory receptor 283	RYVAISSPLLYGQVM*p*SRQ	1865.07	2	2.44	0.15	498.5	17/45
Tal1 interrupting locus	KLLLLLQ*p*SDPKV	1220.36	2	2.40	0.19	501.4	17/27
parkin	KD*p*T*p*SVALNLIT*p*SNRR	1644.45	2	2.34	0.44	240.1	19/60
olfactory receptor 180	KTH*p*SIFKNA*p*SA*p*SMAR	1761.63	2	2.24	0.10	351.9	20/65
adaptor protein complex AP-1, gamma 1 subunit	KVPELMEMFLPA*p*TKN	1586.82	2	2.23	0.19	372.8	15/36
hypothetical protein LOC74107	R*p*SRYSSSSLFEQLEEKTKE	2100.15	3	3.02	0.43	567.0	26/96
fos-like antigen 2	RGS*p*SG*p*SPAHAESYSSGGGG QQKFRV	2400.22	3	2.82	0.28	524.1	38/176
kinesin family member 11	KQDLKLDIP*p*TGMTPERKK	1923.10	2	2.27	0.48	350.1	16/45
protein kinase C, alpha binding protein	RL*p*TIKK*p*YLDVKF	1381.45	2	2.23	0.39	386.1	19/36
arylsulfatase B	KE*p*YNNIYSTNIF*p*TKRA	1923.86	2	2.27	0.49	421.6	19/52
PREDICTED: similar to TATA element modulatory factor 1	R*p*SV*p*SEINSDDELPGKG	1650.48	2	2.21	0.45	304.2	19/52
PREDICTED: similar to Ifi203 protein	**RRQDELSS*p*SESLFINKE**	**1833.86**	**2**	**2.44**	**0.48**	**307.3**	**15/42**

Table A.14 Phosphorylated proteins identified in cGMP incubation 5 min

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
PREDICTED: similar to putative pheromone receptor	RENSDF*p*TKFIQ*p*TVNPKN	1928.88	2	2.63	0.10	232.8	15/56
small glutamine-rich tetratricopeptide repeat (TPR) containing protein	RAPDRTPP*p*SEEDSAEAERL	1937.84	3	3.13	0.37	740.2	29/96
growth associated protein 43	KEGDGSATTDAAPAT*p*SPKA	1656.57	2	2.35	0.03	839.2	23/48
PREDICTED: similar to Crocc protein	RH*p*SELVTKEAADLRA	1549.59	2	2.40	0.50	466.0	15/36
jumonji domain containing 2A	MASE*p*SETLNPSARI	1473.47	2	2.78	0.14	785.2	16/36
lymphoid nuclear protein related to AF4-like	RIPGKP*p*YKETEPPKG	1564.69	1	2.05	0.18	407.2	15/36
multiple endocrine neoplasia 1	RSWLYLKG*p*SYMRCDRK	1859.05	2	2.49	0.42	489.5	19/39
PREDICTED: hypothetical protein LOC69743	KQDVIRH*p*YNMHKKR	1649.78	2	2.59	0.45	358.0	14/33
EGF-like module containing, mucin-like, hormone receptor-like sequence	KTHTLGVL*p*SEFKSKE	1527.63	2	2.21	0.06	361.2	14/36
PREDICTED: similar to hypothetical protein 4932415M13	KG*p*SRDPTAGSNVDNWRQ	1792.61	2	2.24	0.43	561.0	22/56
hypothetical protein LOC244585	KNGDLSFLEKVD*p*SKI	1532.56	2	2.36	0.34	482.5	16/36
hepatoma-derived growth factor	RRAGDVLED*p*SPKR	1267.25	2	2.33	0.25	226.6	15/30
synaptojanin 2	KLMEFDQLQLQK*p*SSGKI	1832.98	2	2.26	0.17	423.9	17/42
hemolytic complement	KIEEQAAK*p*YKH	1160.18	2	2.31	0.26	371.5	12/24
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54	KAR*p*SAQTGARALIL*p*SPTRE	1929.96	2	2.56	0.18	495.3	21/64
centromere autoantigen H	K*p*SNIEMK*p*SLAL*p*SRMKL	1848.85	2	2.50	0.10	439.4	23/65
zinc finger, DHHC domain containing 5	KDEVQLKTT*p*Y*p*SKS	1472.38	2	2.69	0.13	835.8	18/40
odd Oz/ten-m homolog 3	RLSEILYD*p*STRV	1277.29	2	2.67	0.39	523.0	13/27
heat shock protein 1, beta	KIEDVG*p*SDEEDDSGKDKK	1818.67	2	3.17	0.30	566.7	26/45
protein phosphatase 1, regulatory (inhibitory) subunit 1C	K*p*SESLDEEEKLELQRR	1785.77	2	2.33	0.22	201.2	20/39
rho/rac guanine nucleotide exchange factor (GEF) 18	RIIQNTEAG*p*TEDYKD	1562.54	1	1.93	0.16	280.4	15/36
PREDICTED: similar to hypoxia-inducible HIG-1	RSLSQR*p*SSEME*p*YINKY	1832.77	2	2.39	0.47	181.4	14/52
myosin X	KKGGGSS*p*TL*p*SRR	1109.96	2	2.30	0.18	625.9	21/36
Tal1 interrupting locus	KLLLLLQ*p*SDPKV	1220.36	2	2.57	0.29	506.6	16/27
hypothetical protein LOC70385	KYQSLKKQNAFTRD	1564.65	1	1.86	0.13	219.2	12/33
semaphorin 3C [Mus musculus]	KKQQLYV*p*SSNEGV*p*SQV*p*SLHRC	2400.28	3	3.03	0.32	725.3	40/180
cytochrome P450, family 2, subfamily f, polypeptide 2	RSQDLLT*p*SL*p*TKL*p*SKE	1674.56	2	2.36	0.19	589.4	25/60

Continues

Table A.14 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
regulator of G-protein signaling 3	RNE*p*SPGAQPASKTDK*p*TTKS	1920.82	2	2.29	0.17	368.2	19/64
hypothetical protein LOC217705	R*p*SCDAATAGGGKK*p*SPQPTAT PLPRS	2452.37	3	2.88	0.15	711.9	49/220
PREDICTED: similar to enolase 1, alpha non-neuron	KYDLDFK*p*SPDDPSRY	1635.59	2	2.83	0.22	1266.8	23/36
potassium voltage gated channel, Shab-related subfamily, member 1	K*p*TQ*p*SQPILN*p*TKE	1370.18	2	2.47	0.11	337.3	17/45
centromere autoantigen C1	RNP*p*SKENNSPVV*p*YPNKK	1847.77	2	2.36	0.15	478.2	21/56
serine/arginine-rich protein specific kinase 2	R*p*SR*p*TVSASSTGDLPKT	1566.46	2	2.24	0.17	260.3	23/52
hypothetical protein LOC225898	RKAA*p*S*p*SANLLLR*p*SGSTESRG	2088.94	2	2.24	0.16	154.4	19/85
solute carrier family 22 (organic anion transporter), member 19	RKRICLL*p*SFVRF	1315.54	2	2.21	0.18	286.3	11/27
checkpoint kinase 1 homolog	KMCGTLP*p*YVAPELLKRKE	1900.22	1	1.81	0.18	275.6	14/45
voltage-gated potassium channel, subfamily H, member 2	KIPQI*p*TLNFVDLKG	1481.64	3	3.38	0.27	981.8	28/66
laforin	R*p*T*p*TDF*p*YFNIAGHQAMHYSRI	2400.26	3	3.27	0.28	572.5	36/170
PREDICTED: PR-domain zinc finger protein	KILYTL*p*TKVTELRK	1530.71	3	2.86	0.22	857.1	23/66
lymphoblastomic leukemia	KLA*p*SPGPL*p*S*p*TEEVDHRN	1848.63	2	2.58	0.13	314.2	20/70
PREDICTED: similar to ribosomal protein L38	KMNQI*p*SACVGGI*p*S*p*SFGPRV	1964.89	2	2.39	0.45	151.0	20/80
suppression of tumorigenicity 14 (colon carcinoma)	RNKPGV*p*Y*p*TRLPVVRD	1659.72	2	2.34	0.50	549.1	24/48
PREDICTED: peroxidasin	RPR*p*SPNDLLALFRY	1479.59	2	2.32	0.40	332.5	15/33
PREDICTED: similar to hypothetical protein	KPQMT*p*TE*p*TVEE*p*TKT	1634.43	2	2.25	0.23	638.1	24/55
NOK kinase	KGHL*p*SRVLLECSL*p*SDKL	1817.85	2	2.23	0.34	329.3	22/56
PREDICTED: hypothetical protein LOC68795	RY*p*TAFLKQ*p*SGKF	1303.25	2	2.20	0.36	219.4	13/36
PREDICTED: similar to PES1 protein	KYELGSANNYI*p*TRN	1481.47	3	3.00	0.29	472.1	22/66
T-cell acute lymphocytic leukemia 1	RPSP*p*YEMEISDGPH*p*TKV	1848.77	2	2.32	0.15	249.5	17/56
zinc finger protein 59 [Mus musculus]	KCKLCGS*p*SFPHKY	1287.42	2	2.29	0.50	454.0	17/30
vasoactive intestinal peptide receptor 1	KSQHPWGGSNGVSCS*p*TQVS MLTRV	2400.52	3	3.06	0.34	498.1	30/126
PREDICTED: zinc finger protein 407	KCD*p*YGTNVPVEFRN	1480.51	3	2.91	0.35	809.0	28/66
PREDICTED: marapsin 2	K*p*TMKNVCEGDSG*p*SPLVCKQ	1928.99	2	2.23	0.18	306.1	18/64

Table A.15 Phosphorylated proteins identified in cGMP incubation 15 min

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
growth associated protein 43	KEGDGSATTDAAPAT*p*SPKA	1656.57	2	2.28	0.15	495.7	18/48
PREDICTED: similar to Crocc protein	RH*p*SELVTKEAADLRA	1549.59	2	2.39	0.49	620.7	15/36
adaptor-related protein complex 2, beta 1 subunit	RIQPGNPN*p*YTLSLKC	1525.61	2	2.21	0.11	326.3	13/36
PREDICTED: cell division cycle 27 homolog	KGGLTQP*p*SINDSLEI*p*TKL	1833.78	2	2.48	0.44	350.2	20/60
hypothetical protein LOC69942	KQA*p*THNPMIQKTRGSAKQ	1848.99	2	3.02	0.45	818.6	21/45
hepatoma-derived growth factor	RRAGDVLED*p*SPKR	1267.25	2	2.98	0.19	503.7	22/30
odd Oz/ten-m homolog 3	RLSEILYD*p*STRV	1277.29	2	2.36	0.31	502.6	13/27
PREDICTED: RIKEN cDNA 1700065A05 gene	R*p*SHSVGGPLQNIDF*p*SQRP	1902.81	1	1.88	0.12	152.8	13/60
PREDICTED: similar to hypothetical protein	RIYD*p*YKPFKT	1154.22	2	2.29	0.17	294.8	11/21
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54	KAR*p*SAQTGARALIL*p*SPTRE	1929.96	2	2.33	0.10	454.6	22/64
PREDICTED: similar to Gamma-aminobutyric-acid receptor rho-3 subunit p	KNEQEA*p*TYARD	1162.07	2	2.87	0.16	481.2	13/24
PREDICTED: source of immunodominant MHC-associated peptides	RVDKAGSP*p*TLLNCLMYKM	1834.07	2	2.52	0.41	282.8	15/45
PREDICTED: similar to enolase 1, alpha non-neuron	KYDLDFK*p*SPDDPSRY	1635.59	2	2.27	0.21	899.8	22/36
PREDICTED: talin 2	KLDEG*p*TPPEPKG	1163.14	2	2.38	0.18	771.1	16/27
BCL2/adenovirus E1B 19kDa-interacting protein 1, NIP3	KNSTLSEED*p*YIERRRE	1848.83	2	3.13	0.47	477.8	15/39
heat shock protein 1, beta	KIEDVG*p*SDEEDDSGKD	1575.41	2	2.48	0.21	1092.2	25/39
unc-13 homolog B	KIT*p*TVVCAQGLQAKD	1525.72	2	2.36	0.16	309.3	14/39
PREDICTED: microtubule-associated protein 1 A	RVPSAPGQE*p*SPVPDTKS	1589.61	2	2.82	0.31	1036.4	24/42
mitochondrial DNA-directed RNA polymerase	RFC*p*SVS*p*SIKSLKS*p*SERA	1898.80	1	2.24	0.10	223.5	20/70
coagulation factor V	KENRLDPPIVAR*p*YIRI	1792.94	2	2.58	0.26	540.2	17/39
mucin 5, subtype B, tracheobronchial	KTTIL*p*T*p*TASSLATGTPRE	1751.72	2	2.23	0.13	593.0	22/60
PREDICTED: hypothetical protein LOC212153	RVEKA*p*SEFAI*p*SNTCSTRN	1903.85	2	2.27	0.48	270.8	15/60
SET domain-containing protein 7	RVYVADSLISSAGEGLF*p*SKV	1924.02	2	2.35	0.54	234.9	13/51
protein phosphatase 1, regulatory (inhibitory) subunit 1C	KSE*p*SLDEEEKLELQRR	1785.77	2	2.23	0.19	266.1	23/39
PREDICTED: hypothetical protein XP_622534	R*p*YKYLYYEELSIRKS	1849.00	2	2.84	0.43	608.3	18/36
a disintegrin and metalloprotease domain 32	KVTIVL*p*SSLELW*p*SDKN	1750.77	2	2.40	0.17	624.8	22/52
ubiquitin specific protease 9, Y chromosome	R*p*YPHQFEDKP*p*TL*p*SKV	1830.66	2	2.35	0.11	357.0	19/60

Continues

Table A.15 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
Tg737 protein	RRPPV*p*TAKIP*p*S*p*TAVSRP	1820.76	2	2.34	0.19	155.4	23/70
guanylate cyclase activator 2b	RTIATDECELCINVAC*p*TGC	1940.11	2	2.34	0.11	286.3	13/51
X-ray repair complementing defective repair in Chinese hamster cells 1	KPPSKP*p*SAGP*p*STLKR	1507.36	2	2.26	0.13	618.1	24/60
GLI-Kruppel family member GLI1	RRSS*p*SSSSMSSAY*p*TV*p*SRR	1920.68	2	2.21	0.11	417.6	21/75
rho/rac guanine nucleotide exchange factor (GEF) 18	RIIQNTEAG*p*TEDYKD	1562.54	1	2.03	0.12	320.9	16/36
PREDICTED: similar to novel protein similar to vertebrate otoferlin (O	RDFI*p*TQSRVFI*p*SEAEKK	1930.90	1	1.98	0.12	161.7	16/56
PREDICTED: RIKEN cDNA 4631422O05	KQEC*p*Y*p*SLKCNLEKE	1618.62	1	1.94	0.15	350.0	16/44
dentin sialophosphoprotein	KEKL*p*SPKDTRD	1154.18	1	1.83	0.11	468.7	13/24
REST corepressor 1	RQIQNIKQ*p*TN*p*SALKEKL	1903.92	2	2.64	0.44	778.7	22/56
calpain 12	REV*p*TLLPDSLSQRW	1438.49	2	2.46	0.12	541.2	16/33
annexin A6	REFIEK*p*YDKS	1152.16	2	2.45	0.13	469.5	15/21
PREDICTED: similar to mKIAA1109 protein [Mus musculus]	RRS*p*SMRAA*p*SLKD	1267.25	2	2.41	0.27	240.6	19/36
D-glucuronyl C5-epimerase [Mus musculus]	KLLTNGSV*p*SVVLE*p*T*p*TEKN	1930.82	1	1.93	0.13	199.7	17/75
PREDICTED: similar to Nuclear pore glycoprotein p62 (62 kDa nucleopori	KEQNGPLNMM*p*YVNKE	1618.74	1	1.85	0.16	417.6	16/36
carboxypeptidase N, polypeptide 1 [Mus musculus]	KDPA*p*TKRHRGPA	1286.31	2	2.21	0.17	239.1	14/30

Table A.16 Phosphorylated proteins identified in cGMP incubation 30 min

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
H2-K region expressed gene 2	KR*p*YESQLRDLERQ	1545.56	3	2.87	0.13	1056.7	23/60
protein phosphatase 1, regulatory (inhibitory) subunit 1C [Mus musculu	K*p*SESLDEEEKLELQRR	1785.77	2	2.38	0.19	289.4	20/39
growth associated protein 43	KEGDGSATTDAAPAT*p*SPKA	1656.57	2	2.38	0.20	567.5	20/48
SEC24 related gene family, member C	KFA*p*YRAVLN*p*SPVKT	1525.54	2	2.36	0.02	566.7	21/44
unc-13 homolog B	KITI*p*TVVCAQGLQAKD	1525.72	2	2.33	0.05	254.9	13/39
adaptor-related protein complex 2, beta 1 subunit	RIQPGNPN*p*YTLSLKC	1525.61	2	2.21	0.07	379.9	14/36
guanylate cyclase activator 2b	RTIATDECELCINVAC*p*TGC	1940.11	2	2.61	0.20	417.9	14/51

Continues

Table A.16 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
PL10 protein	KQYPISLVLAP*p*TRE	1438.58	1	1.98	0.04	345.1	14/33
vacuolar protein sorting 11	KL*p*TLL*p*TDPP*p*TARL	1438.29	1	2.26	0.23	589.1	24/50
DEAD (Asp-Glu-Ala-Asp) box polypeptide 54	KAR*p*SAQTGARALIL*p*SPTRE	1929.96	2	2.34	0.14	429.8	22/64
hepatoma-derived growth factor	RRAGDVLED*p*SPKR	1267.25	2	2.61	0.17	516.7	21/30
calcium/calmodulin-dependent protein kinase kinase 2, beta	RF*p*YFQDLIKG	1154.22	1	1.84	0.17	476.8	13/21
NACHT, leucine rich repeat and PYD containing 1	R*p*YIV*p*SS*p*SKLVEIIPKE	1816.80	2	2.30	0.02	393.2	25/65
BCL2/adenovirus E1B 19kDa-interacting protein 1, NIP3	KNSTLSEED*p*YIERRRE	1848.83	2	2.59	0.45	438.4	15/39
stathmin 1	KDL*p*SLEEIQKK	1155.16	2	2.21	0.11	607.6	17/24
polycystic kidney disease 1-like	RNLQ*p*SQS*p*SVNVIVKE	1576.54	2	2.23	0.25	548.9	20/48
heat shock protein 1, beta	KIEDVG*p*SDEEDDSGKD	1575.41	2	2.95	0.24	1023.5	25/39
hect (homologous to the E6-AP (UBE3A) carboxyl terminus) domain and RC	KI*p*SGANNCKPNRPSLAKI	1750.89	1	2.04	0.11	210.3	13/45
WAS protein family, member 2	RVQVKV*p*TQLDPKE	1335.46	2	2.28	0.03	849.1	19/30
supervillin	KASELA*p*TLIQ*p*TKR	1335.29	1	1.85	0.01	172.7	17/40
mitochondrial ribosomal protein S31	KNPYL*p*SVKQKV	1157.22	2	2.20	0.02	350.4	12/24
metallophosphoesterase 1	RNNPSFIMGSL*p*T*p*SRD *p*YALSKC	2342.26	3	3.05	0.12	480.7	34/180
putative MAPK activating protein PM20,PM21	RQPLGPDG*p*SQGFKQRR-	1751.81	2	2.40	0.03	547.7	17/42
spir-2 protein	RSRQ*p*TQ*p*SLYIPNTR*p*TL NFQ	2408.30	3	2.91	0.13	469.4	36/170
toll-like receptor 7	KKTVFVM*p*TQKYAKT	1524.73	2	2.31	0.12	258.7	15/33
calcium channel, voltage-dependent, beta 3 subunit	RD*p*SLMP*p*SDEA*p*SESSRQ	1751.45	2	2.34	0.09	384.2	22/65
mucin 5, subtype B, tracheobronchial	K*p*TTILT*p*TASSLATGTPRE	1751.72	2	2.38	0.07	344.4	19/60
sex comb on midleg-like 2	RQ*p*SK*p*TPPTNEFKI	1437.34	1	2.77	0.14	578.9	21/40
cathepsin 3 precursor	KYGK*p*TY*p*SLEEEGQKRA	1848.75	2	2.47	0.46	282.5	16/52
THUMP domain containing 3	RTASVDIIV*p*TDMPFGKRM	1831.01	2	2.41	0.17	436.8	18/45
annexin A6	REFIEK*p*YDKS	1152.16	2	2.39	0.17	474.8	15/21
angiopoietin-related protein 5	RSSEATQD*p*SELATLRM	1668.50	2	2.21	0.02	321.7	19/52
centromere autoantigen H	K*p*SNIEMK*p*SLAL*p*SRMKL	1848.85	2	2.21	0.00	376.8	18/65
RAN binding protein 2	RI*p*TPDM*p*TLQTMKG	1439.49	1	2.05	0.04	209.6	14/40

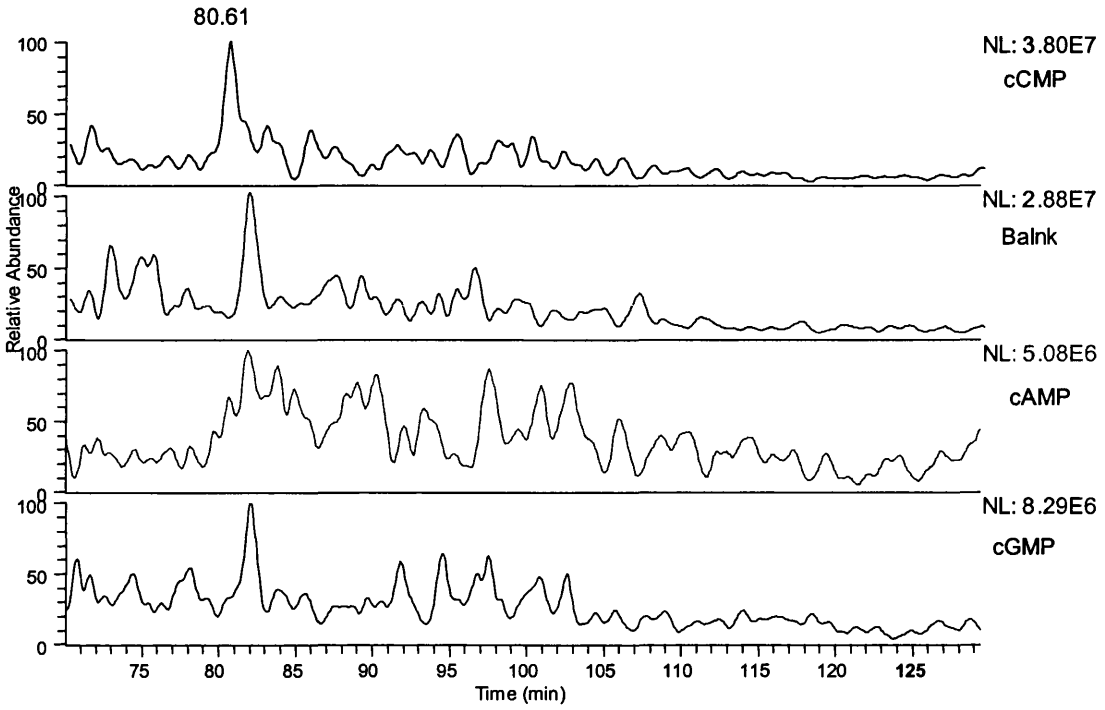
Continues

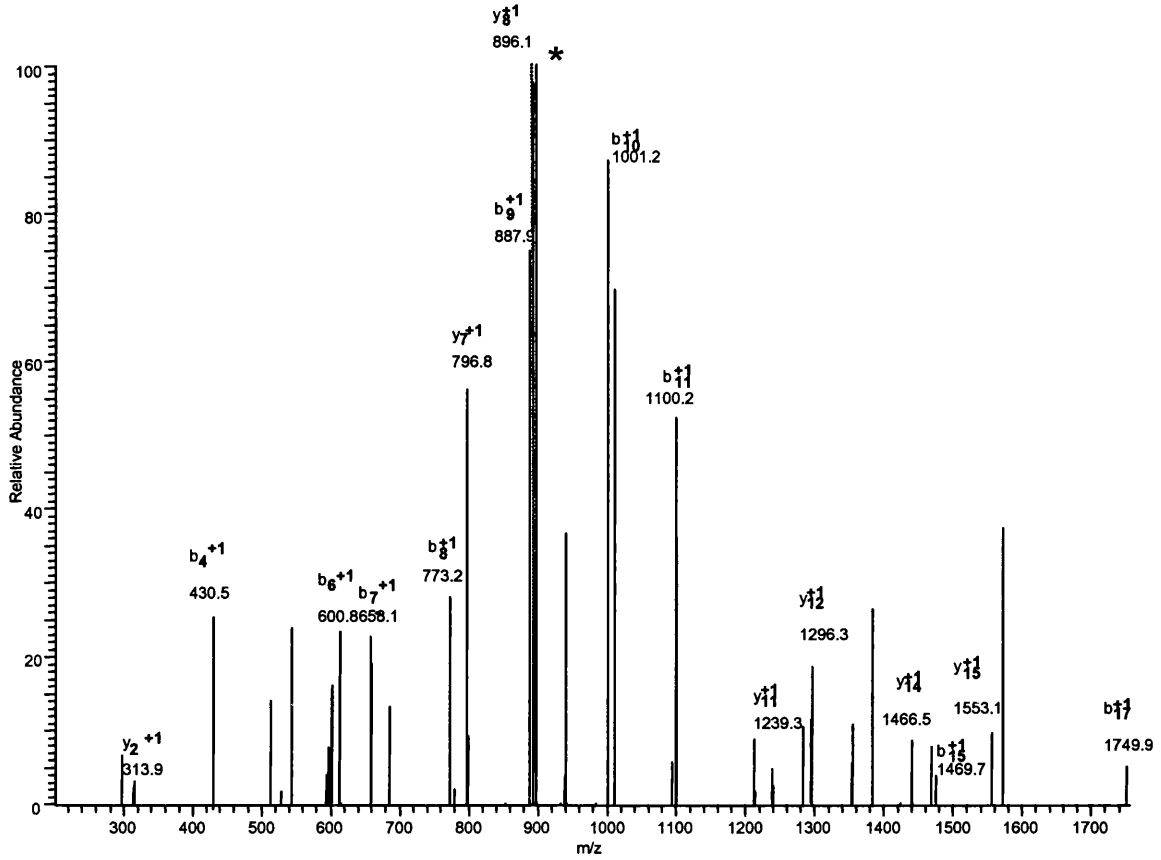
Table A.16 continued

Protein	Peptide	MH+	z	XC	dCn	Sp	Ions
SEC5 homolog	KK*p*SLEELLNKF	1154.22	1	2.04	0.00	449.8	14/24
inositol polyphosphate-5-phosphatase B	KLVFQLPFG*p*SH*p*TRT	1562.56	1	1.84	0.11	309.1	17/44
tryptophan hydroxylase 2	R*p*Y*p*TQ*p*SIEILKD	1335.17	1	1.87	0.06	143.3	20/40
cleavage and polyadenylation specific factor 4	RC*p*TKGHLAFL*p*SGQ	1422.40	2	2.55	0.43	306.3	18/44
purinergic receptor P2Y, G-protein coupled 2	KPPTEP*p*TP*p*SPQARR	1438.33	1	2.38	0.14	636.8	24/44
MUS81 endonuclease	R*p*TLYQLYC*p*SHSPL*p*S	1752.61	2	2.32	0.09	491.2	22/60
small conductance calcium-activated potassium channel protein 3	RSEDLEKQIGSLE*p*SKL	1643.66	2	2.23	0.39	283.0	16/39
PHD finger protein 1	RW*p*TDGLL*p*YLGTIKKV	1668.72	1	2.14	0.05	72.8	15/48
DEAH (Asp-Glu-Ala-His) box polypeptide 29	KNLPKSPNP*p*SFEKV	1438.49	1	1.98	0.04	448.1	17/33
PREDICTED: similar to proline-rich peptides 637K precursor, prostatic	KH*p*S*p*SP*p*TVQVLKM	1336.16	1	1.83	0.02	218.6	22/45
procollagen C-endopeptidase enhancer 2	KC*p*TWKI*p*TVPEGKV	1422.44	2	2.31	0.48	251.1	14/40
vacuolar protein sorting 41	KKD*p*SQNK*p*TLLKT	1335.29	2	2.26	0.03	900.2	23/36
eukaryotic translation initiation factor 4A1	RSG*p*SSRVLITTDLLARG	1669.79	1	2.06	-	98.5	15/42
PQ loop repeat containing 1	RSFAA*p*TDSKDEELRV	1549.50	2	2.39	0.51	251.8	13/36
gene rich cluster, C2f gene [Mus musculus]	KMVSISN*p*YPLSAALTCAKV	1850.07	2	2.34	0.52	519.7	16/48
microfibrillar-associated protein 3 [Mus musculus]	K*p*TLELAKV*p*TQFKT	1438.45	1	1.97	0.06	458.8	19/40
a disintegrin and metalloprotease domain 32 [Mus musculus]	KVTIVL*p*SSLELW*p*SDKN	1750.77	2	2.30	-	340.6	18/52

Appendix 3 Full scan spectrum comparison, MS/MS spectrum, b and y ions of the identified phosphopeptides of the proteins in Table 7.6

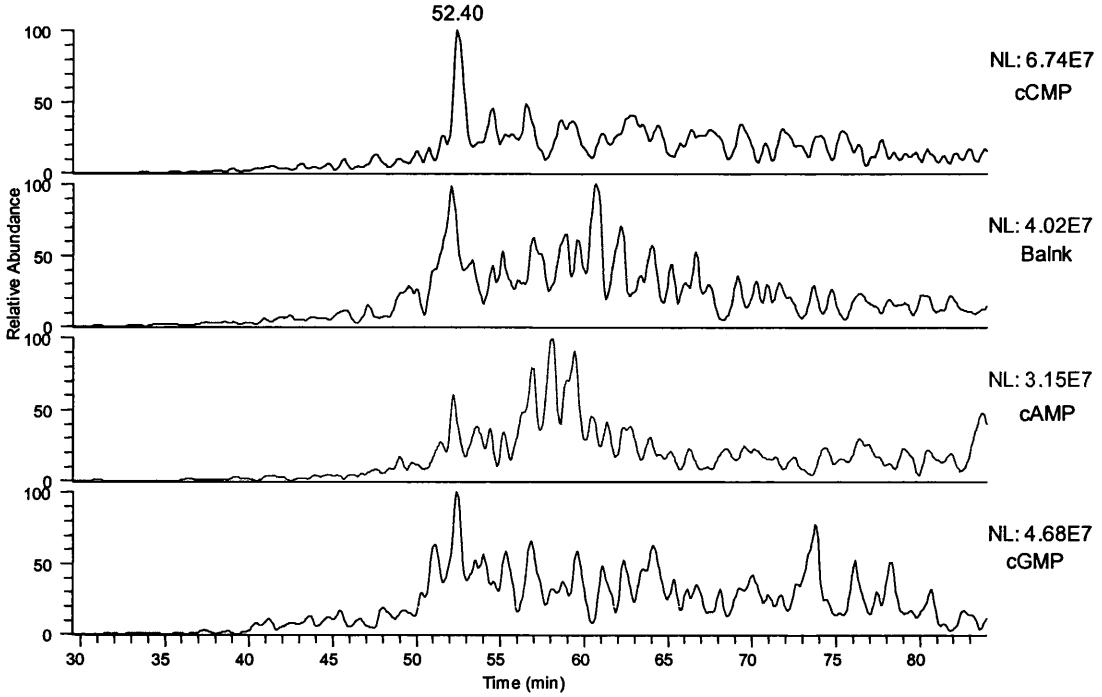
- **Formin homology 2 domain containing 1**

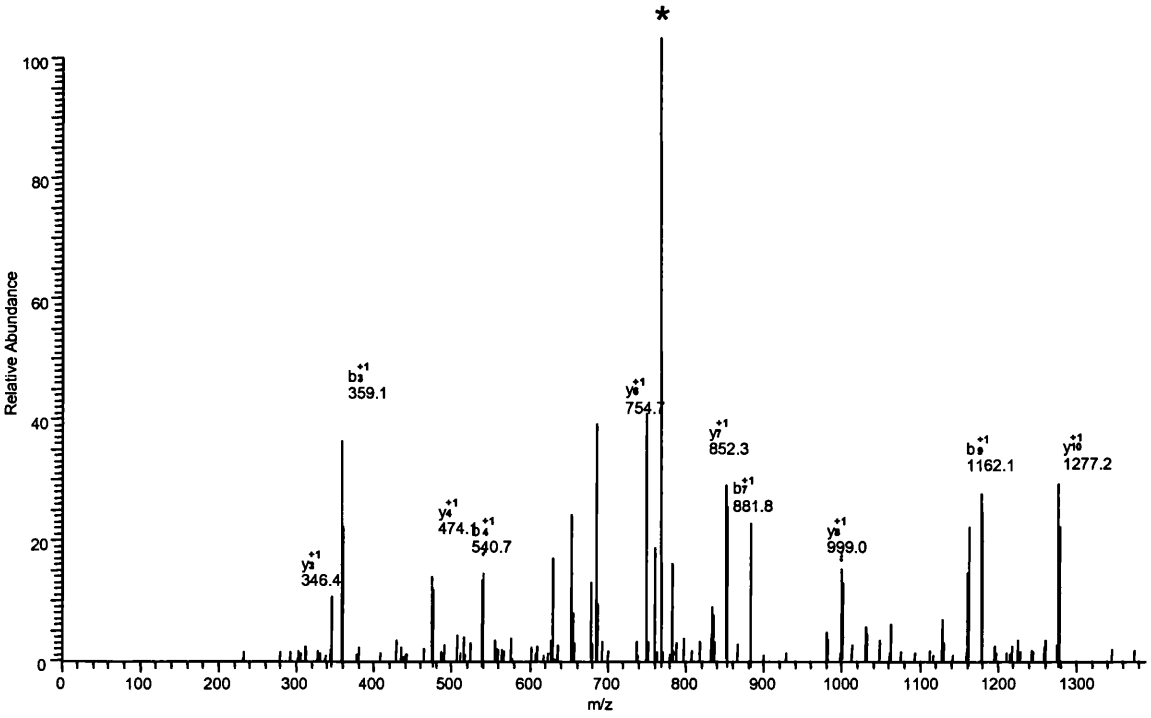




	AA	B	Y	
1	T	102.11	-	18
2	L	215.27	1794.95	17
3	K	343.44	1681.79	16
4	S	430.52	1553.62	15
5	G	487.57	1466.54	14
6	L	600.73	1409.49	13
7	G	657.78	1296.33	12
8	D	772.87	1239.28	11
9	D	887.96	1124.19	10
10	L	1001.11	1009.10	9
11	V	1100.25	895.95	8
12	Q	1228.38	796.81	7
13	A	1299.45	668.68	6
14	L	1412.61	597.61	5
15	G	1469.66	484.45	4
16	L	1582.82	427.40	3
17	S*	1749.86	314.24	2
18	K	-	147.19	1

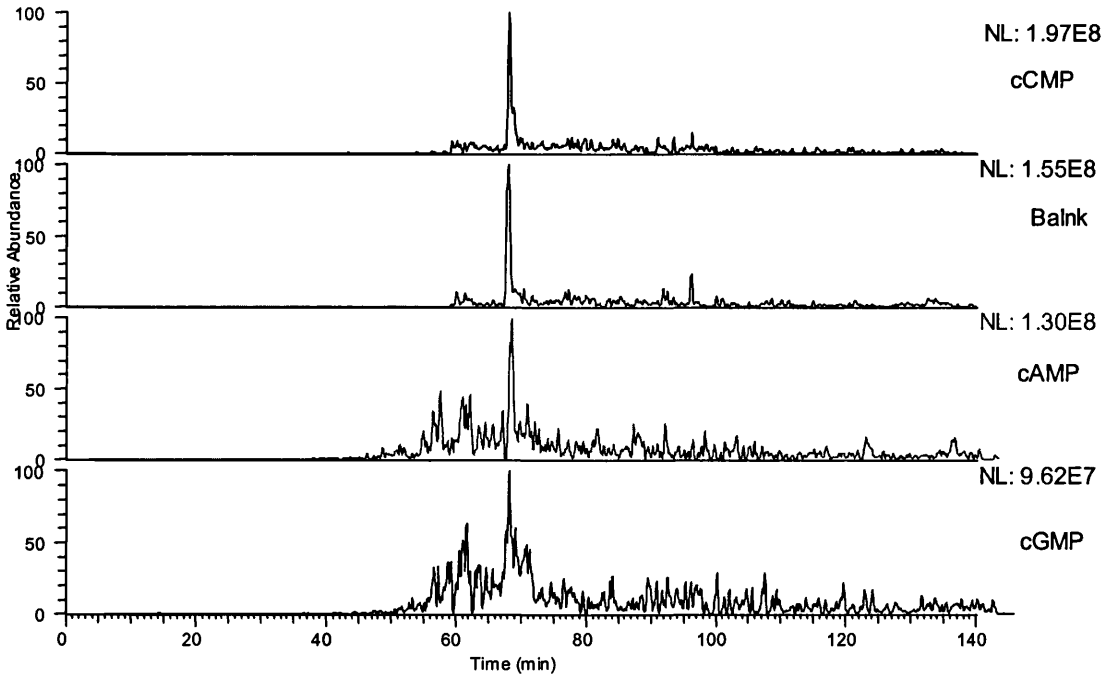
- **MAP-kinase activating death domain isoform 8**

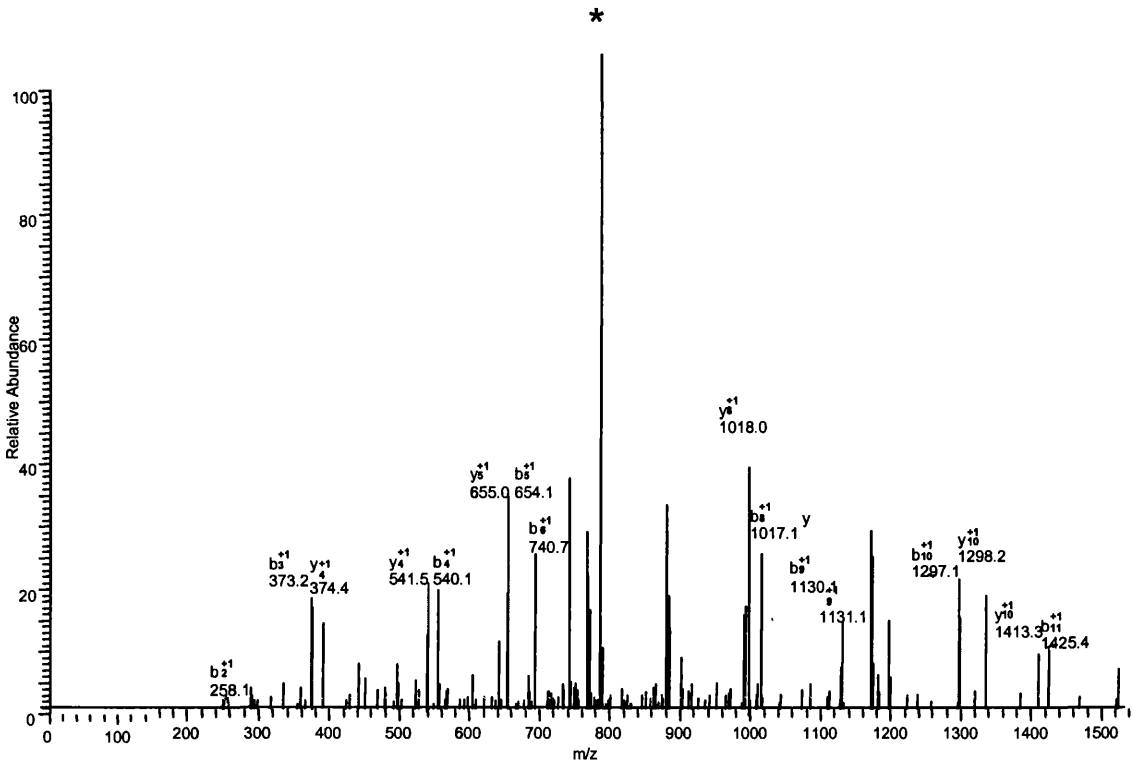




	AA	B	Y	
1	E	130.12	-	13
2	K	258.29	1506.49	12
3	T	**359.40**	1378.32	11
4	T@	**540.47**	**1277.22**	10
5	P	637.58	1096.14	9
6	F	784.76	**999.03**	8
7	P	**881.87**	**851.85**	7
8	S*	1048.92	**754.74**	6
9	L	**1162.08**	587.69	5
10	K	1290.25	**474.53**	4
11	G	1347.30	**346.36**	3
12	N	1461.40	289.312	2
13	R	-	175.209	1

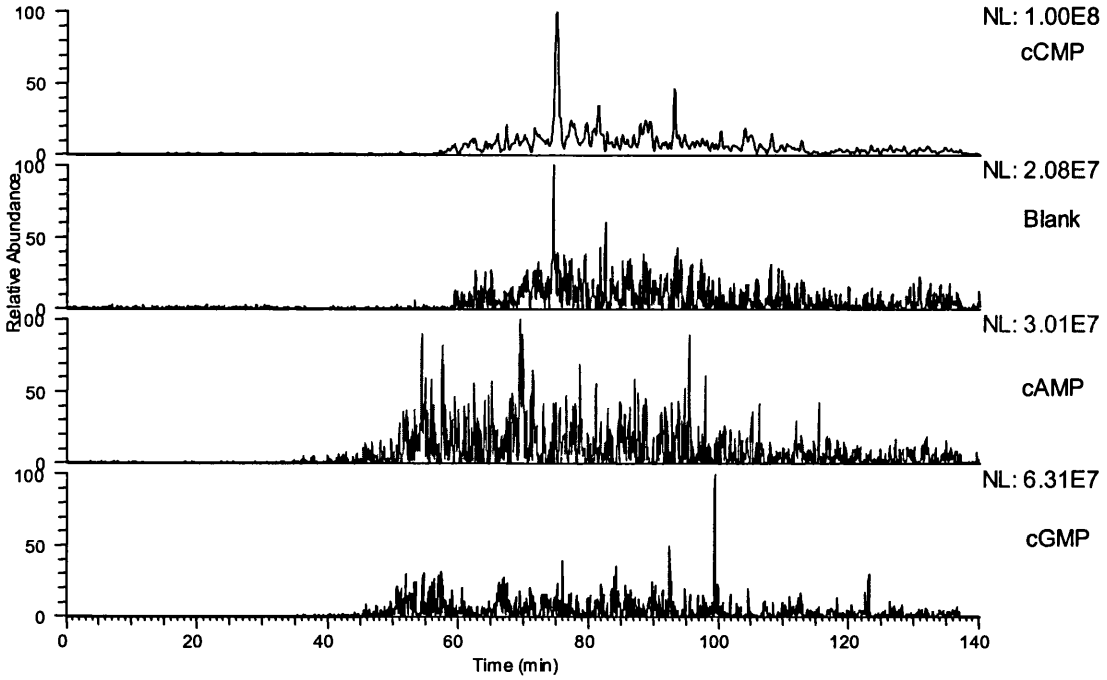
- **Evolutionarily conserved G-patch domain containing**

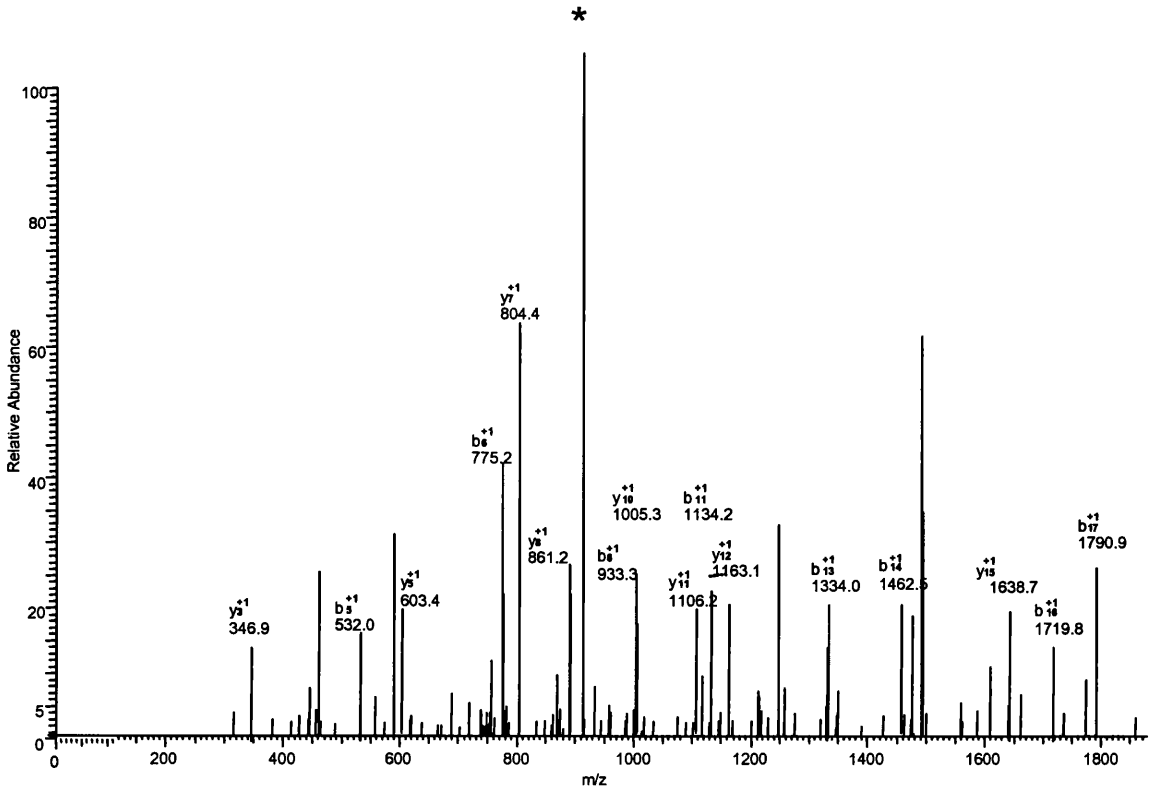




	AA	B	Y	
1	K	129.18	-	13
2	E	**258.29**	1542.39	12
3	D	**373.38**	**1413.27**	11
4	S*	**540.43**	**1298.19**	10
5	I	**653.58**	**1131.14**	9
6	S	**740.66**	**1017.98**	8
7	E	869.77	930.90	7
8	F	**1016.95**	801.79	6
9	L	**1130.11**	**654.61**	5
10	S*	**1297.15**	**541.46**	4
11	Q	**1425.28**	**374.42**	3
12	A	1496.36	246.28	2
13	R	-	175.21	1

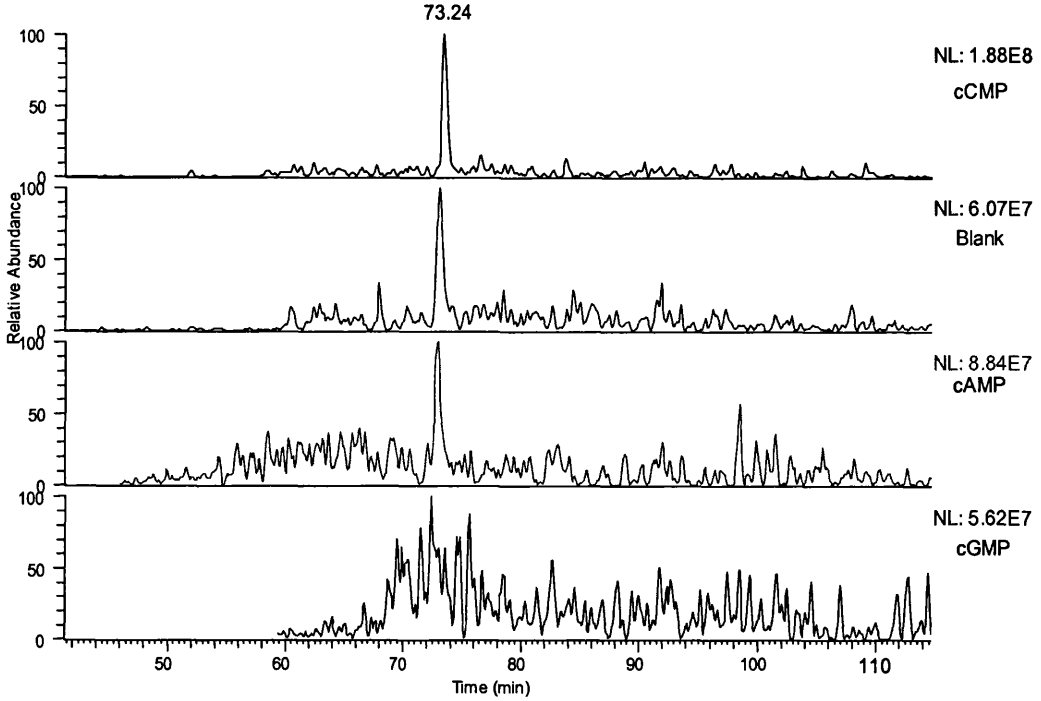
- **Protein kinase, interferon-inducible double stranded RNA dependent**

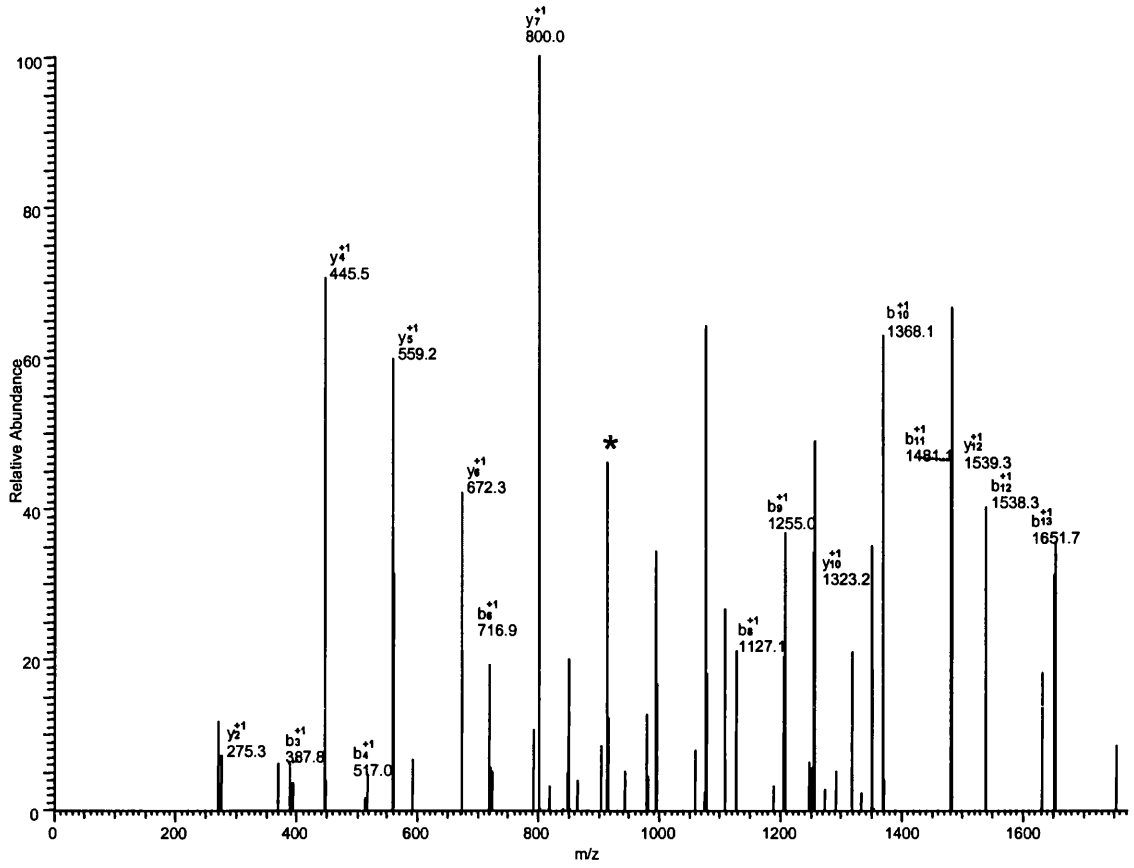




	AA	B	Y	
1	I	114.16	-	18
2	G	171.22	1823.88	17
3	Q	299.34	1766.83	16
4	T	400.45	**1638.70**	15
5	M	**531.65**	1537.60	14
6	Y#	**774.79**	1406.40	13
7	G	831.84	**1163.26**	12
8	T	**932.94**	**1106.21**	11
9	G	**989.99**	**1005.11**	10
10	S	**1077.07**	948.056	9
11	G	**1134.12**	**860.98**	8
12	V	1233.25	**803.93**	7
13	T	**1334.36**	704.79	6
14	K	**1462.53**	**603.69**	5
15	Q	1590.66	475.52	4
16	E	**1719.78**	**347.39**	3
17	A	**1790.85**	218.27	2
18	K	-	147.19	1

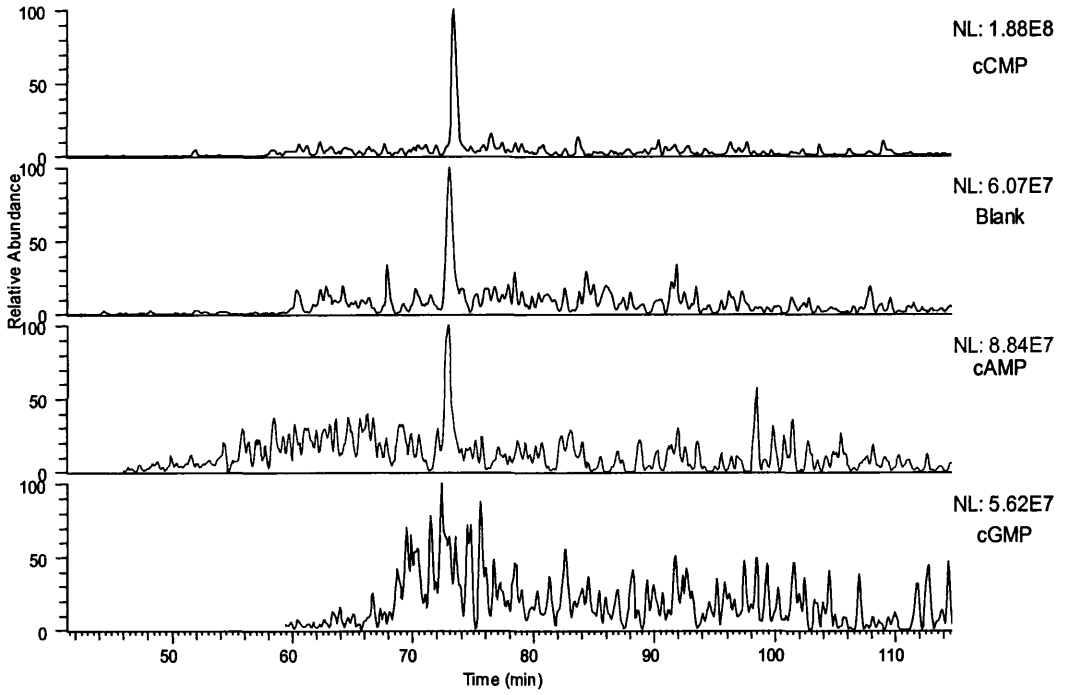
- **Lamin B2**

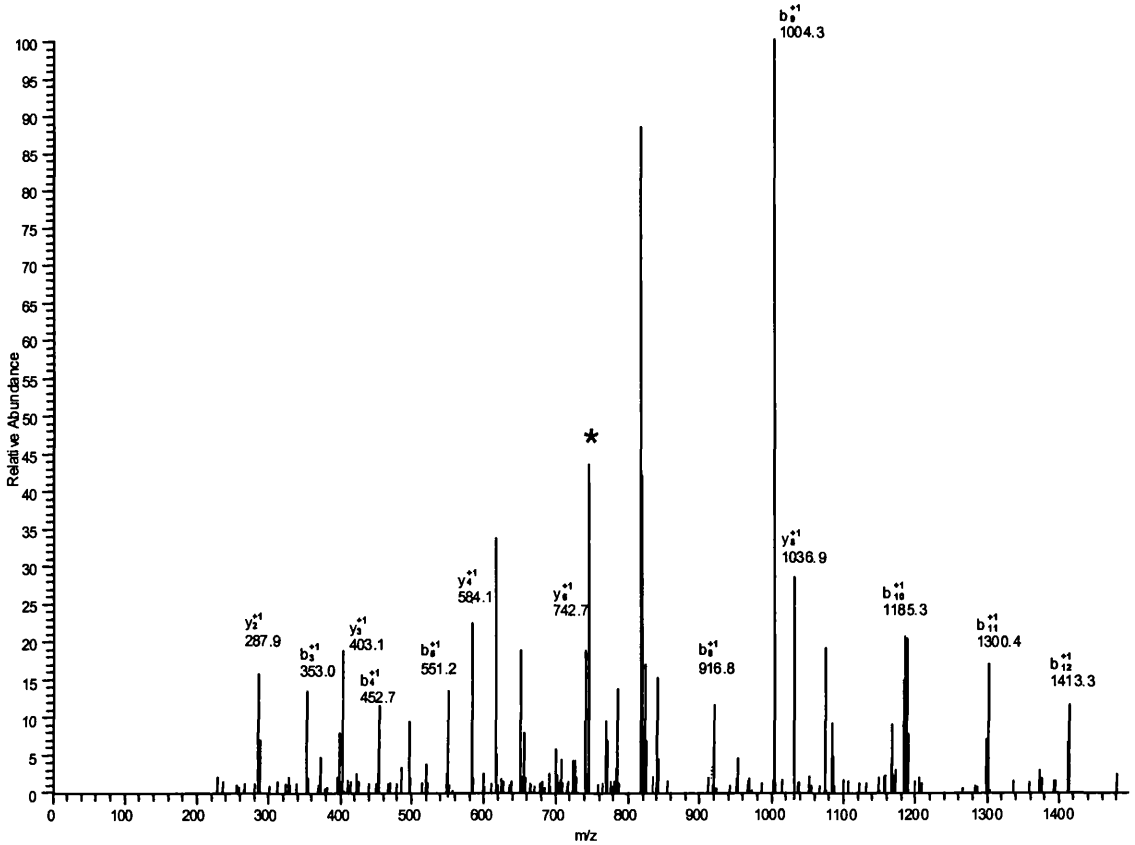




	AA	B	Y	
1	M	132.20	-	15
2	R	288.39	1794.83	14
3	V	**387.52**	1638.64	13
4	E	**516.64**	**1539.51**	12
5	S	603.71	1410.40	11
6	L	**716.87**	**1323.32**	10
7	S*	883.92	1210.16	9
8	Y#	**1127.06**	1043.12	8
9	Q	**1255.19**	**799.98**	7
10	L	**1368.34**	**671.85**	6
11	L	**1481.50**	**558.69**	5
12	G	**1538.55**	**445.54**	4
13	L	**1651.71**	**388.48**	3
14	Q	1779.84	**275.32**	2
15	K	-	147.195	1

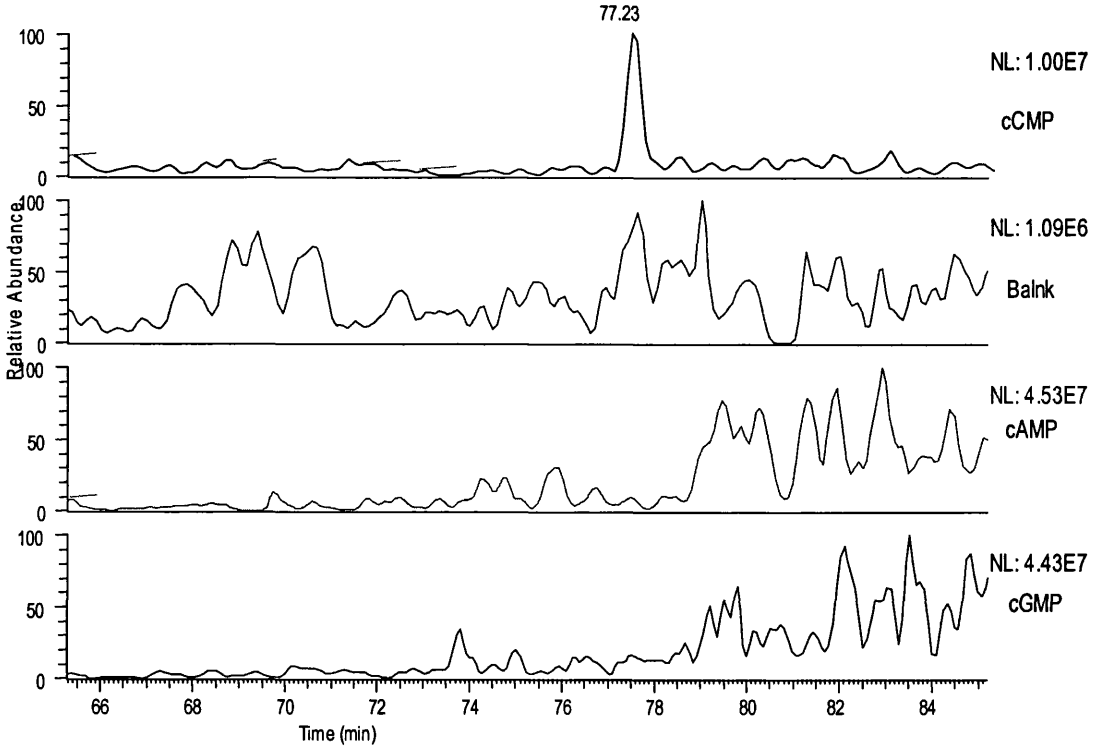
- **Arrestin domain containing 1**

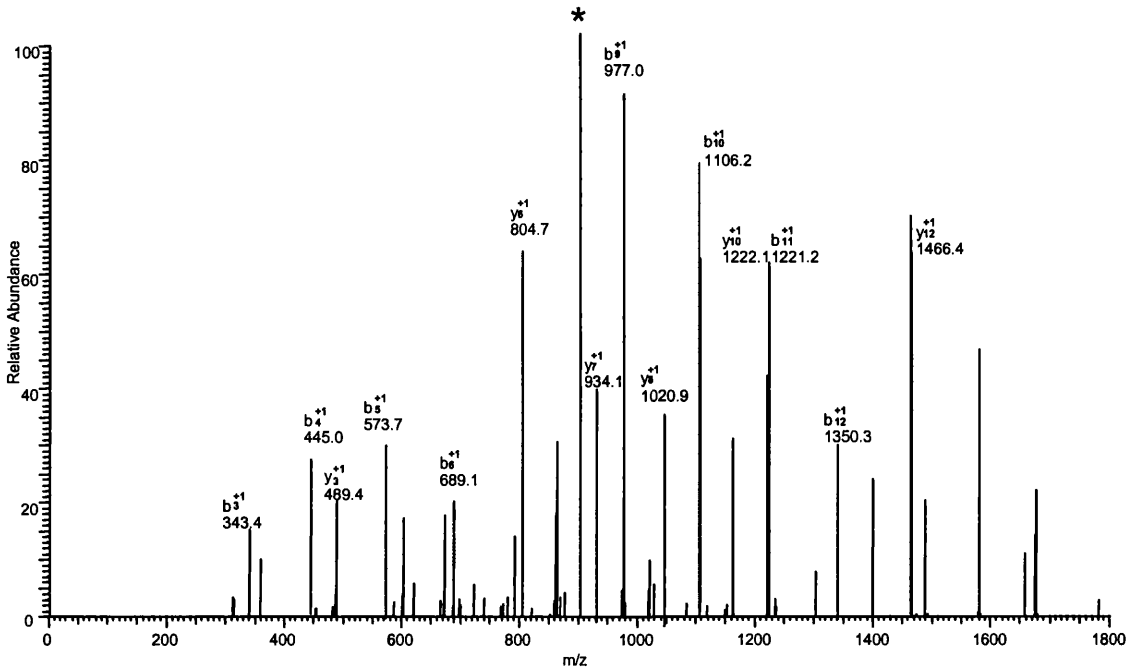




	AA	B	Y	
1	T@	182.08	-	13
2	G	239.13	1406.33	12
3	N	353.23	1349.27	11
4	V	452.36	1235.17	10
5	V	551.49	1136.04	9
6	L	664.65	1036.91	8
7	T@	845.72	923.75	7
8	A	916.80	742.68	6
9	S	1003.88	671.60	5
10	T@	1184.95	584.52	4
11	D	1300.04	403.45	3
12	L	1413.20	288.37	2
13	R	-	175.21	1

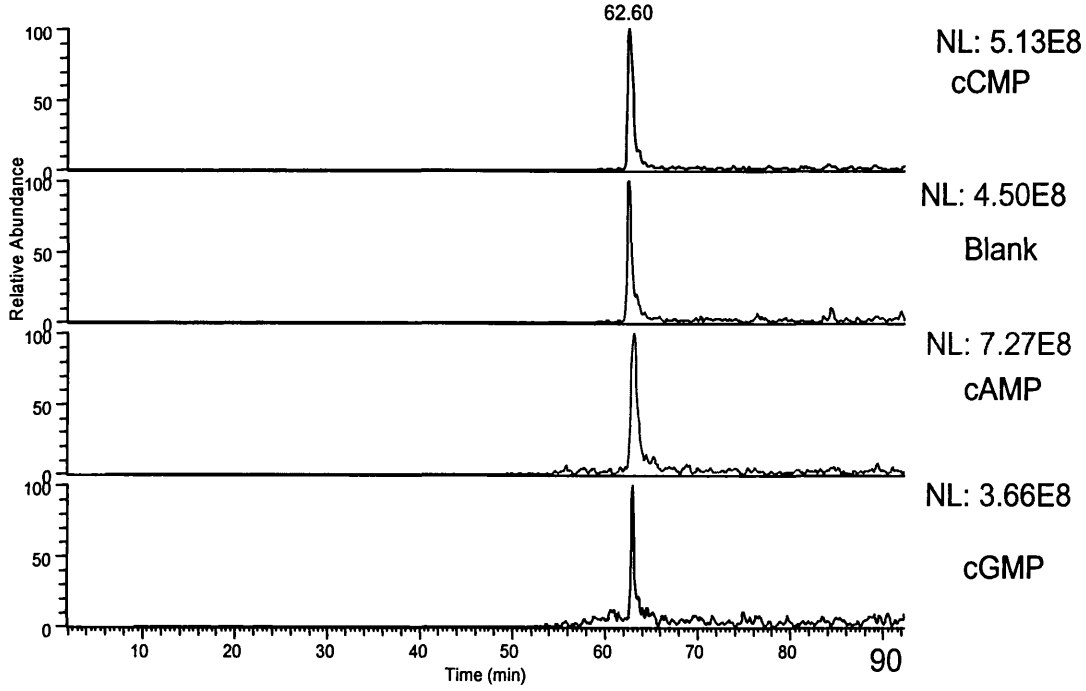
- **Myeloid/lymphoid or mixed lineage-leukemia translocation to 1 homolog**

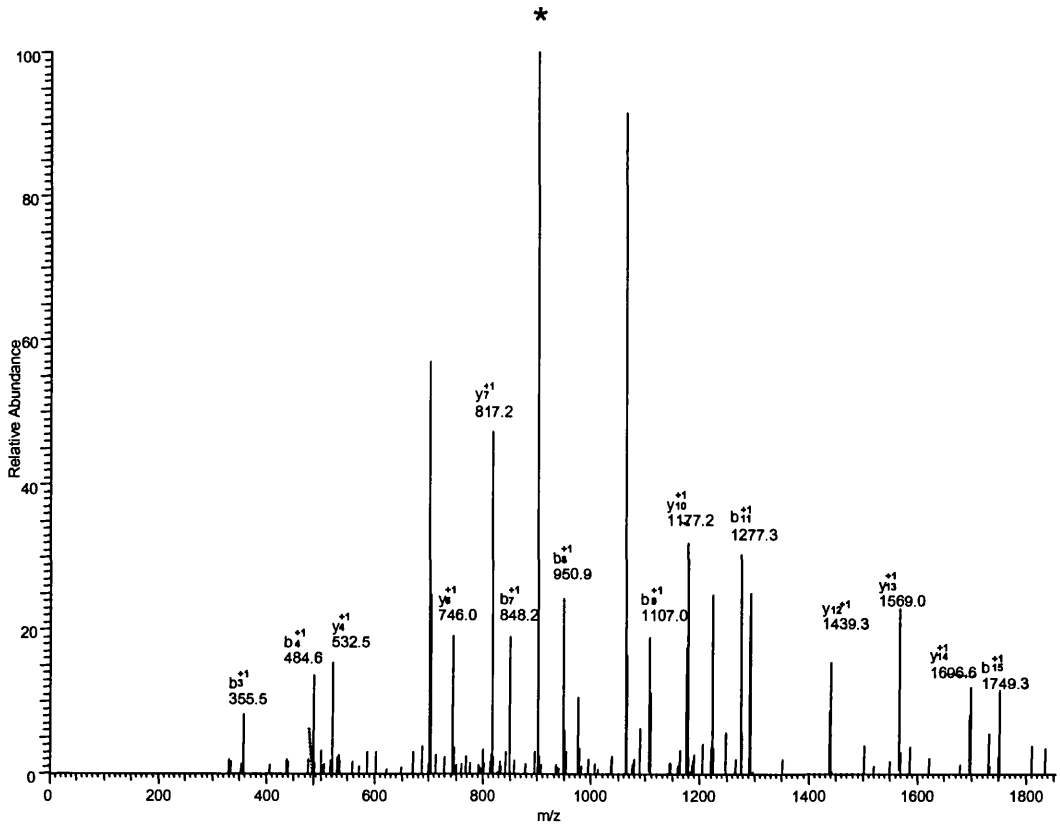




	AA	B	Y	
1	T	102.11	-	16
2	L	215.27	1808.67	15
3	Q	**343.40**	1695.52	14
4	T	**444.50**	1567.39	13
5	E	**573.62**	**1466.28**	12
6	D	**688.71**	1337.17	11
7	S	775.78	1222.08	10
8	N	889.89	1135.00	9
9	S	**976.96**	**1020.90**	8
10	E	**1106.08**	**933.82**	7
11	D	**1221.17**	**804.71**	6
12	E	**1350.28**	**689.62**	5
13	A	1421.36	**560.50**	4
14	S*	1588.40	**489.43**	3
15	F	1735.58	322.38	2
16	R	-	175.209	1

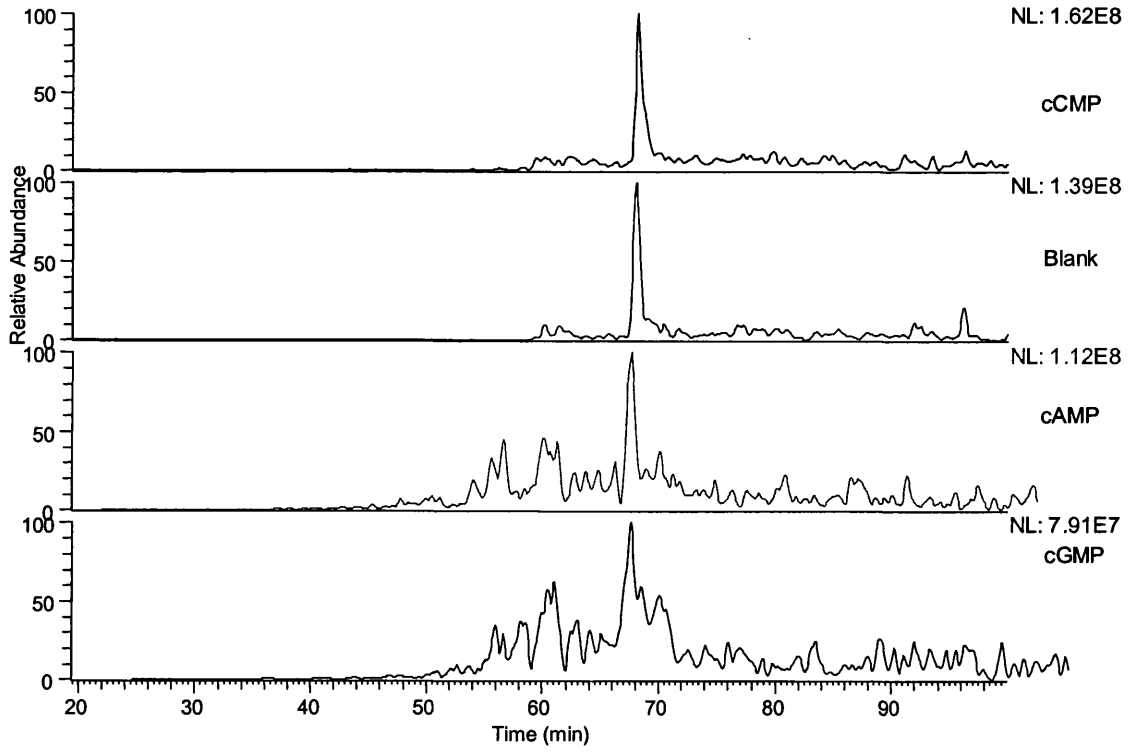
- **Low density lipoprotein receptor-related protein 1**

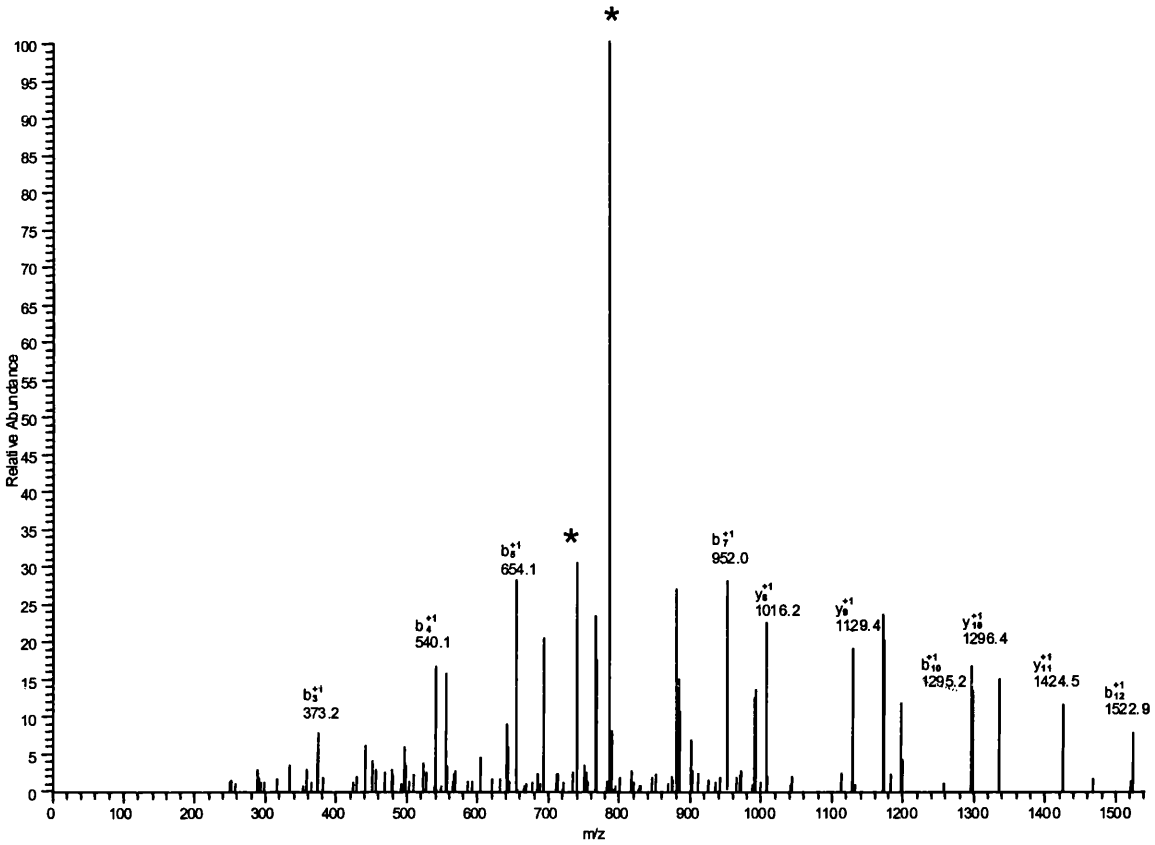




	AA	B	Y	
1	I	114.16	-	16
2	L	227.32	**1809.88**	15
3	Q	**355.45**	**1696.73**	14
4	E	**484.56**	**1568.60**	13
5	D	599.65	**1439.48**	12
6	F	**746.83**	1324.40	11
7	T	**847.93**	**1177.22**	10
8	C	**951.07**	1076.12	9
9	R	**1107.26**	972.97	8
10	A	**1178.34**	**816.79**	7
11	V	**1277.47**	**745.71**	6
12	N	1391.57	646.58	5
13	S	1478.65	**532.47**	4
14	S*	1645.70	445.40	3
15	C	**1748.84**	278.35	2
16	R	-	175.21	1

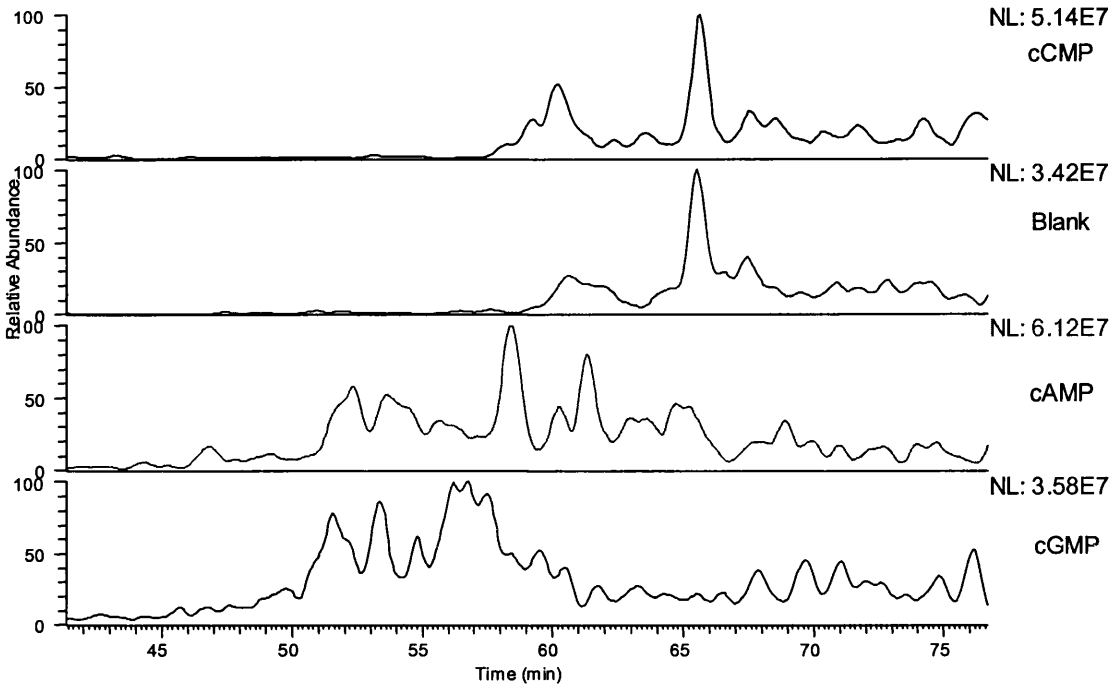
- **Centromere/kinetochore protein zw10 homolog**

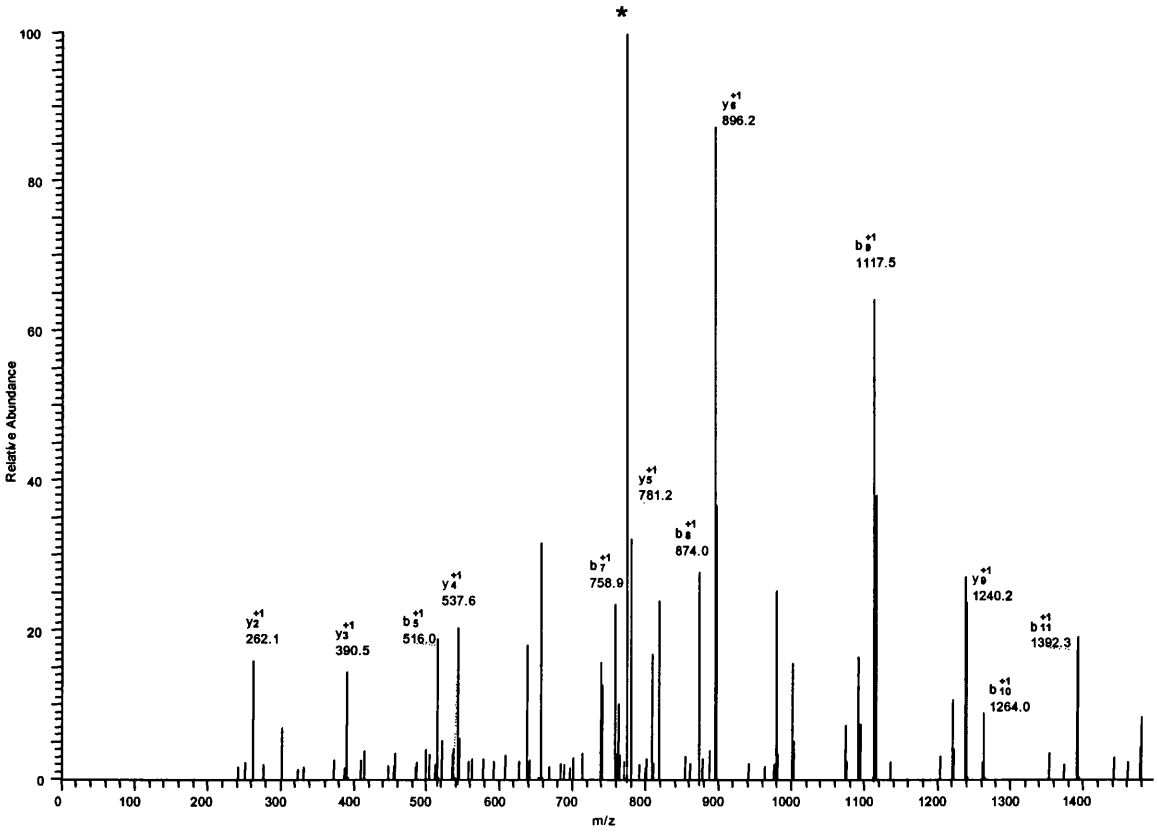




	AA	B	Y	
1	M	132.20	-	13
2	L	245.36	1537.61	12
3	K	**373.53**	1424.45	11
4	S*	**540.58**	1296.28	10
5	L	**653.74**	1129.23	9
6	S*	820.78	1016.07	8
7	M	**951.98**	849.03	7
8	E	1081.09	717.83	6
9	L	1194.25	588.72	5
10	T	**1295.35**	475.56	4
11	V	**1394.48**	374.46	3
12	Q	**1522.62**	275.33	2
13	K	-	147.20	1

- **Deleted in azoospermia-like**





	AA	B	Y	
1	N	115.11	-	13
2	S	202.19	1539.55	12
3	L	315.34	1452.47	11
4	V	414.48	1339.31	10
5	T	515.58	1240.18	9
6	Q	643.71	1139.08	8
7	D	758.80	1010.95	7
8	D	873.88	895.86	6
9	Y#	1117.02	780.77	5
10	F	1264.20	537.63	4
11	K	1392.37	390.46	3
12	D	1507.46	262.28	2
13	K	-	147.19	1

Appendix 4 Recipe for two 7 cm10% acrylamide gel

Gel type	Acrylamide	H2Od	1.5M Tris pH8.8	1 M Tris pH6.8	10%SDS	10%APS	TEMED
Resolving	5 mL	6.75 mL	3.75 mL	---	75 µL	75 µL	15 µL
Stacking	650 µl	3.25 mL	---	1.25 mL	25 µl	25 µl	5 µl

Appendix 5 Solutions to be prepared for 2D-Gel electrophoresis

Rehydration buffer:

7 M urea, 2 M thiourea, 4% CHAPS, 40 mM DTT, 0.5% IPG buffer, 0.002% Bromophenol blue, deionised water ($H2O^{d}$)

Lysis buffer 1

40 mM Tris-HCl (pH 7.4), 1 mM 2-mercaptoethanol, 1 M PMSF, 0.2 M Na_3VO_4, 1 mM NaF

Lysis buffer 2

7 M urea, 2 M thiourea, 4% CHAPS, 65 mM DTT, 1 mM EDTA, 0.5 mM EGTA, 1 M PMSF, 0.2 M Na_3VO_4, 1 mM NaF, 40 mM Tris-HCl

SDS Electrophoresis buffer

3 g Tris, 14.4 g glycine, 1 g SDS H_2O to make 1 L solution

SDS equilibration buffer

50 mM Tris-HCl (pH8.8), 6 M urea, 30% glycerol, 2% SDS, 400 µl of 1% bromophenol blue, H_2O^d to 200 mL

Agarose sealing solution

25 mM Tris, 192 mM glycine, 0.1 %(w/v) SDS, bromophenol blue, 0.5 % (w/v) agarose

Coomassie gel staining 1 liter

1.0 g coomassie Blue R-250, 450 mL methanol, 450 H_2O^d, 100 mL glacial acetic acid

Coomassie gel destain 1 L

100 mL methanol, 100 mL glacial acetic acid, 800 mL H_2O^d

Colloidal Coomassie Stain

0.12% Coomassie Blue G250, 10% ammonium sulphate, 20% methanol, 10% O-phosphoric acid

Destaining solution:

0.4 g potassium ferricyanide ($K_3Fe(CN)_6$) in 200 mL sodium thiosulphate (0.2 g/L $Na_2S_2O_3.5H_2O$)

Appendix 6 Parameters of active rehydration and IEF process of 24 cm IPG strip

Voltage(V)	Voltage mode	Step duration(h:min)	Volt-hours (Vh)
30	Step and Hold	12:00	36
100	Step and Hold	1:00	100
200	Step and Hold	1:00	200
500	Step and Hold	1:00	500
1000	Step and Hold	1:00	1000
8000	Gradient	0:30	2250
8000	Step and Hold	10:00	80000
Total		26:30	84086

Appendix 7 Homogeneous gel solutions for 2 x 24 cm SDS gel

Final %T	Volume required for (mL)		
	10%	12.5%	15%
Acrylamide stock	66.67	83.33	100
1.5 M Tris-HCl pH 8.8	50	50	50
Water	79	62.3	45.67
10% SDS	2	2	2
10% APS	2	2	2
TEMED	0.05	0.05	0.05

Appendix 8 Silver staining process

Step	Solution	Time on gentle shaker in solution
Fix	50% methanol 12% acetic acid 38% H_2O	>20 mins
Wash	50% ethanol 50% H_2O	2 x 20 mins
Sensitize	0.02% $Na_2S_2O_3.5H_2O$	1 mins
Rinse	Milli-Q H_2O	3 x 20 seconds
Silver	0.2% $AgNO_3$ in 0.026% (v/v) formaldehyde	20 mins
Rinse	Milli-Q H_2O	2 x 20 seconds
Develop	6% (w/v) Na_2CO_3, 4.0 x 10^{-4}%(w/v) $Na_2S_2O_3.5H_2O$ in 0.017% (v/v) formaldehyde	2-5mins
Wash	Milli-Q H_2O	2 x 2 mins
	H_2O:methanol:acetic acid	10 mins

Appendix 9 Sequential extraction protocol

375 mg mice brain was ground into powder in liquid nitrogen. This powder was placed into 1.7 mL lysis buffer1, sonicated in an ice-cold water bath for 6×10 s with 60 glass beads. Then the solution was centrifuged at 5000 g for 15 mins (4 °C), then centrifuged at 16000 g for 60 mins (4°C). This resulted in getting supernatant 1, and a pellet 1 of 329 mg. Next 1.6 mL of lysis buffer 2 was added to pellet 1, stirred for 1 hr in ice-cold water for 60 mins, and then centrifuged for 30 mins at 4 °C, supernatant 2 being produced.

Appendix 10 In-gel digestion protocol

Excised and destained gel band was cut into 1 mm^3 pieces, wash 3 times in 50% acetonitrile, 50 mM NH_4HCO_3 with vortexing for 30 mins at room temperature. After that the gel pieces were washed with 100% acetonitrile with vortexing for 10mins, before they were dried in a 'speedvac' vacuum condenser. Then the "swollen" gel pieces were placed on ice with 50 µL 20 µg/mL trypsin solution in 40 mM NH_4CO_3, 10% acetonitrile and incubated at 37 °C overnight. After incubation, extraction was started with 3.5% TFA in 50% acetonitrile with sonication for 15 mins, this was repeated three times and the combined solution was evaporated to dryness using a "speedvac" condenser. The digested peptides were then resuspended in 10 µL 0.1% TFA.

Appendix 11 C_{18} ZipTip of the digested peptides protocol

Millipore Reverse-phase ZipTips® C_{18} were used for both desalting and concentrating the peptides formed after trypsin digestion. The method used was based upon the documented protocol in the Millipore instruction guide and care was taken not to introduce air into the ZipTips at any point during any ZipTips procedure. Each 10 µL Ziptip was first activated by the aspiration and expulsion of 10 µL of 50% acetonitrile three times.

Equilibration of the C_{18} ZipTip then occurred with 0.1% TFA being aspirated and

expelled three times. The binding of peptides from the sample to the reverse phase packing was then conducted by aspirating the sample a minimum of ten times into ZipTip. The Ziptip then underwent a wash step with 0.1% TFA to remove any unbound components and the bound peptides were eluted with 10 μL of 75% ACN/0.1% TFA. The eluted sample was then concentrated to dryness and reconstituted in 4 μL of 50% ACN in 0.1% TFA for subsequent MALDI-ToF analysis or in 30 μL of 50% ACN in 0.1% TFA for ESI-MS/MS analysis. This was vortex-mixed to ensure equal distribution of sample components.

Appendix 12 Recovery of TCA/acetone precipitation

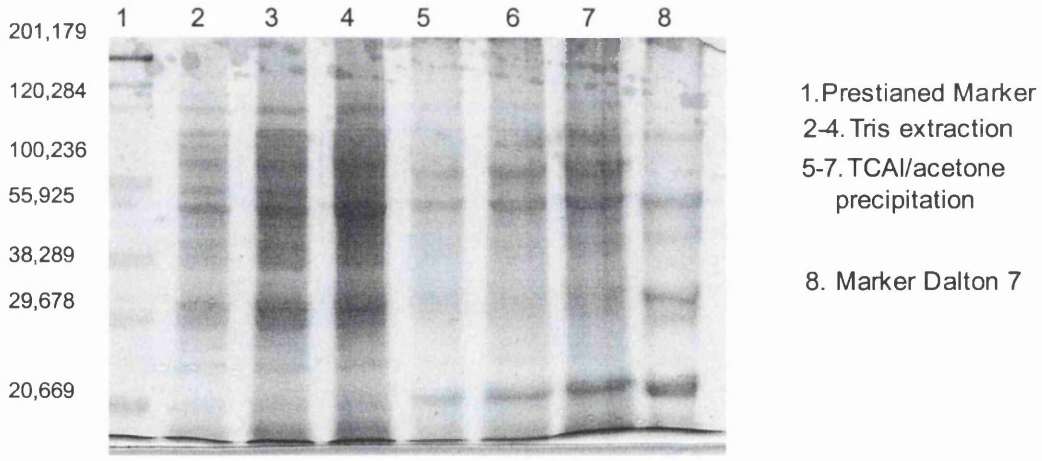


Figure A.7 Recovery of TCA/acetone precipitation by comparing the Tris extracted brain homogenate (Lanes 2-4) and TCA/acetone precipitation of the Tris extracted brain homogenate (Lane 5-7). The loss of some proteins comparing to the lanes without precipitation can be observed. But it is a still widely used to remove interfering substance and concentrate the proteins.

Appendix 13 7 cm 2D gel of mice brain tris extraction protein homogenate using IEF strip of different pH range

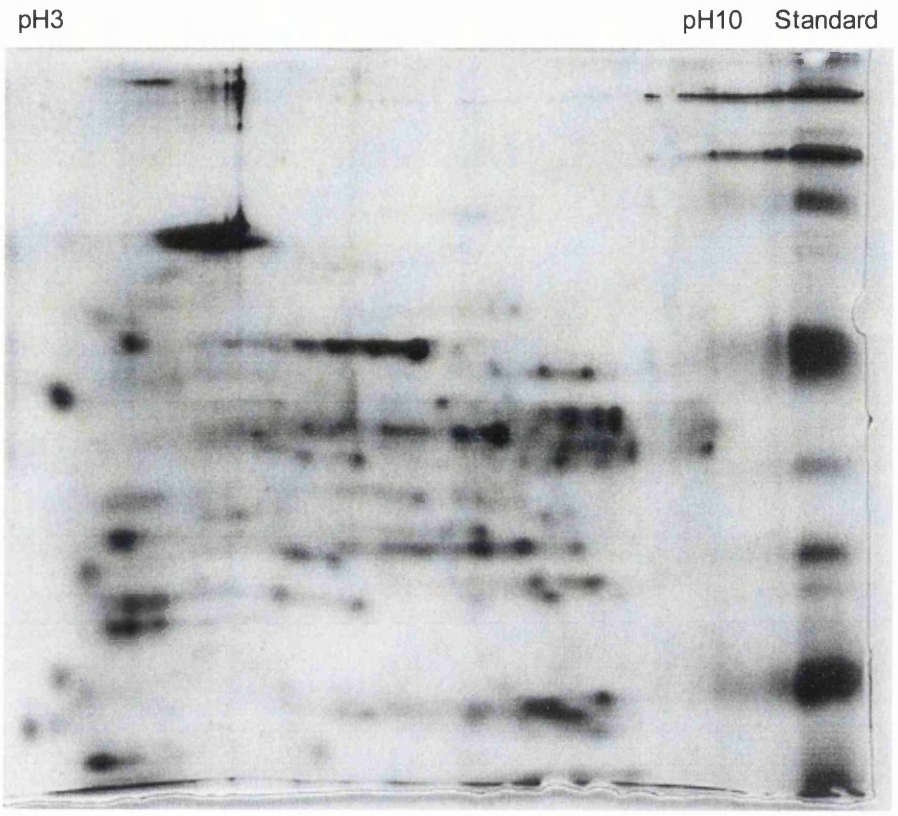


Figure A.8 10 μg of mice brain homogenate separated by pH 3-10 IPG strip and 9% SDS gel

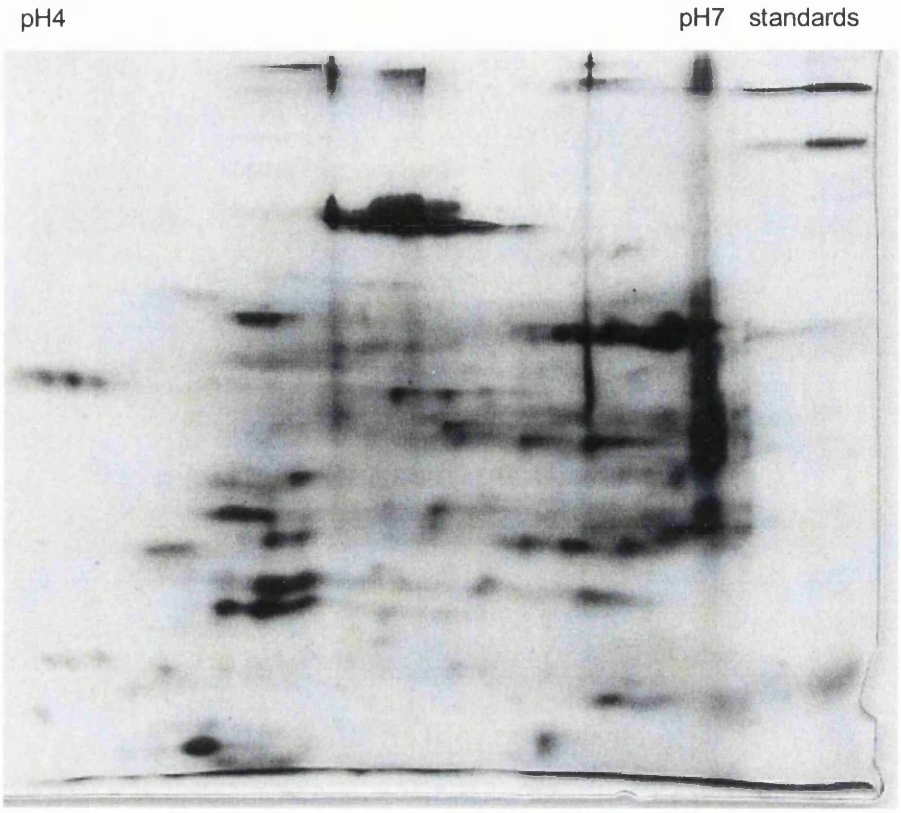


Figure A.9 20 μg of mice brain homogenate separated by pH 4-7 IPG strip and 9% SDS gel.

Normally, a narrow pH in the first dimension gives better resolution of the proteins, but better resolution was not observed in pH 4-7 gel from this experiment, the pH 4-7 gel seems not focused properly. The reason of this is because the two strips were run at the same time. It was found out that strips of different pH cannot run together, because the start resistance of two strips is different and therefore the electric current requirements conflict. When connected in parallel, it is difficult for them to attain equilibration at the same time under the same IEF condition, this is also suitable for sequential extraction, and the proteins obtained by two extraction methods are better run separately.

Appendix 14 7 cm 2-D gel of sequential extracted mice brain homogenate

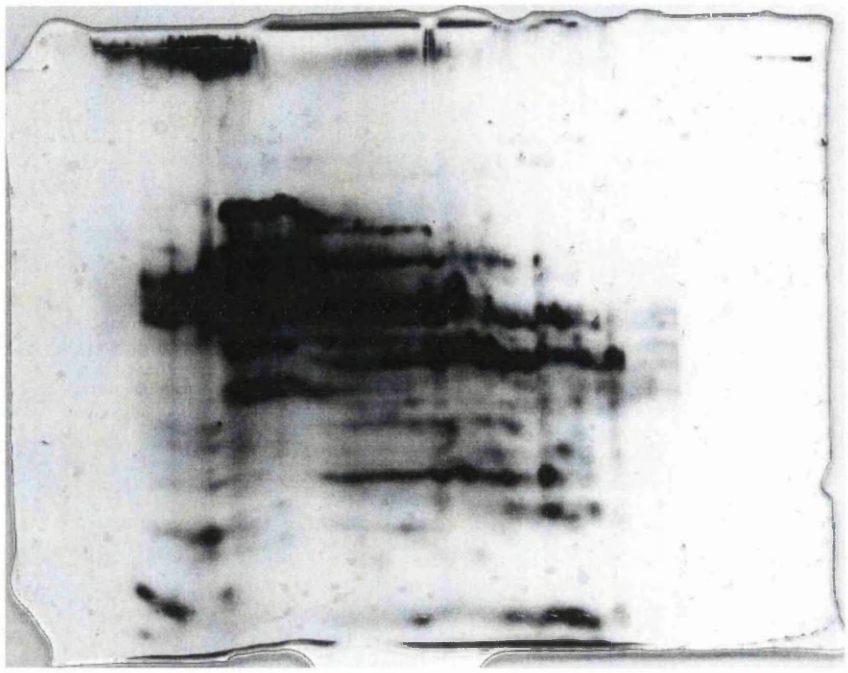


Figure A.10 7 cm 2-D gel electrophoresis of 40 µg of tris extraction of mice brain proteins followed by TCA/acetone precipitation and resuspension in lysis buffer. The gel shows proteins were overloaded for silver staining.

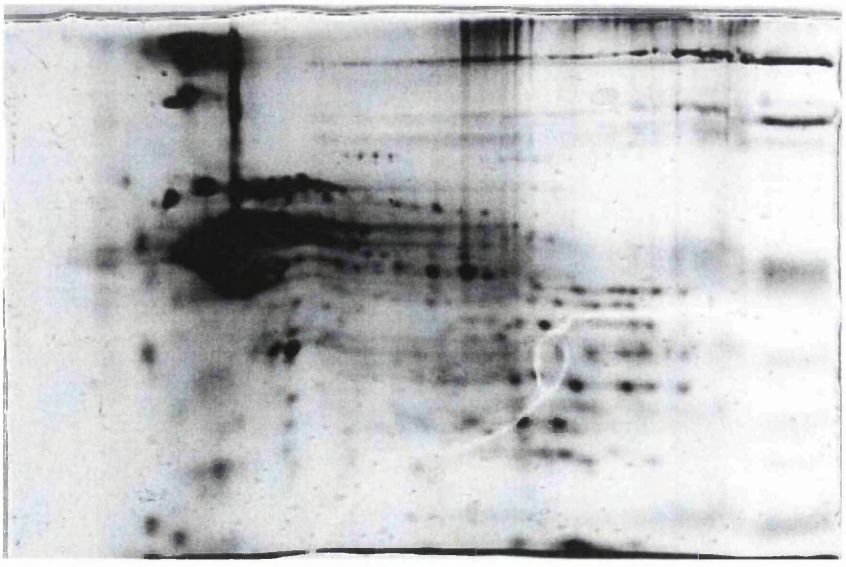


Figure A.11 7 cm 2-D gel electrophoresis of 30 µg of urea extraction of mice brain proteins after Tris extraction, followed by TCA/acetone precipitation and resuspension in lysis buffer the gel is better comparing to the Tris extraction one. It might be because the amount of proteins loaded is less.

Appendix 15 24 cm 2-D gel of tris extracted mice brain homogenate

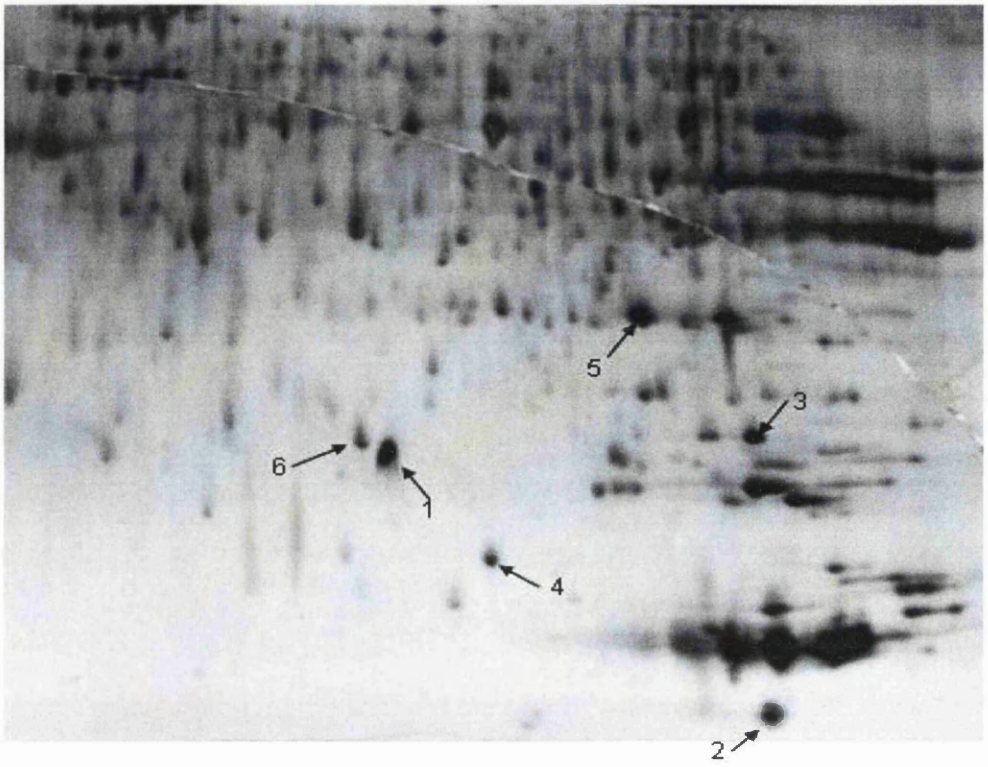


Figure A.12 24 cm 2-D gel electrophoresis of 100 μg of Tris extraction of mice brain proteins followed by TCA/acetone precipitation and resuspension in lysis buffer

The 24 cm gels displayed many more spots than the 7 cm smaller gels. But the larger gels are much more difficult to handle and break easily. The first dimension is very satisfactory because the protein spots are well focused, but there are vertical streaks in the gel. This is mainly due to the second dimension of electrophoresis. One reason of the streaks might be because the pH of the resolving gel, the other reason might be the iodoacetamide equilibration step (15 mins) is not long enough, another reason could be poor solubility of proteins in the gel.

Six spots (shown in arrows) were cut from the gel followed by destaining and in-gel digestion. After cleaning the sample by C_{18} ZipTip, they were analysed by MALDI ToF and LC/MS, no convincing identification was obtained from database search with LC/MS results. Although mass spectrometry compatible silver staining protocol was used in this experiment, the reason of this might be that the sensitivity of mass spectrometer is not as high as silver staining. The sensitivity of silver staining can be as low as 1-2ng (50fmol on average proteins), while the sensitivity of mass spectrometer is at pmol level.

Appendix 16 1-D gel for comparison of different incubations

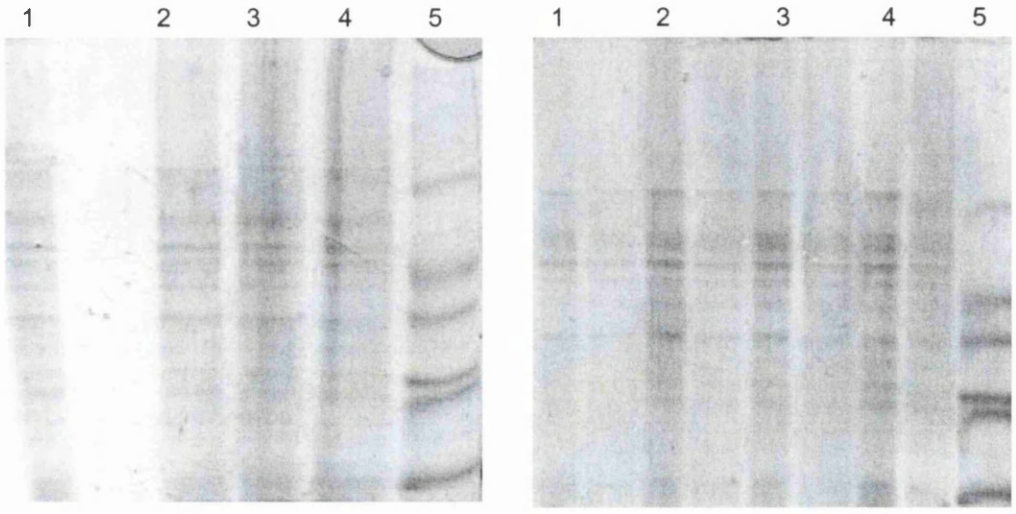


Figure A.13 1-D gel electrophoresis of cyclic nucleotides incubated mice brain homogenate. Lane 1: cAMP incubation; lane 2: Blank incubation; lane 3: cCMP incubation; lane 4: cGMP incubation; lane5: molecular marker. Left: colloid Coomassie staining. Right: Coomassie blue staining

No obvious differences were observed among different incubations from the 1-D gel image as shown in this Figure. The sensitivity of the colloid Coomassie staining was at the similar level to Coomassie blue, although it was claimed to be comparable to silver staining.