

UNIVERSIDAD AUTÓNOMA DE MADRID
DEPARTAMENTO DE BIOLOGÍA MOLECULAR



**Bases Moleculares de la Presentación de Ligandos
Fosforilados por Moléculas HLA-B**

TESIS DOCTORAL

ADÁN ALPÍZAR MORÚA

MADRID, 2016

DEPARTAMENTO DE BIOLOGÍA MOLECULAR
FACULTAD DE CIENCIAS
UNIVERSIDAD AUTÓNOMA DE MADRID



**Bases Moleculares de la Presentación de Ligandos
Fosforilados por Moléculas HLA-B**



Adán Alpizar Morúa

Licenciado en Química

Director: Dr. Miguel Marcilla Goldaracena

Centro Nacional de Biotecnología, CNB-CSIC

Madrid, 2016

Summary

Major Histocompatibility Complex (MHC, HLA in humans) class I molecules play a key role in the eradication of viral infections and tumor transformation processes. Its primary function is the presentation to cytotoxic T lymphocytes (CTLs) of a broad peptide repertoire generated from the degradation of endogenous proteins. Recognition of an antigenic peptide in the context of a class I molecule leads to the activation of the CTL which exerts its cytotoxic activity on the antigen presenting cell. The peptidome displayed by HLA-I molecules includes phosphorylated ligands. These have been proposed as optimal targets for the treatment of cancer by T cell-based immunotherapy.

Using HLA-B*40 as a model, we set out to study the peptidome and phosphopeptidome presented by HLA-I molecules and to determine the molecular basis of the presentation of phosphorylated ligands. LC-MS/MS analysis of the peptide repertoire associated with B*40 allowed us to characterize more than 9500 unique ligands with a common anchor motif: acidic residues –mainly Glu– at P2 and, aliphatic residues Met or Phe at PΩ. The structural determinants of this motif were analyzed by solving the crystallographic structure of B*40 in complex with an endogenous ligand.

By using phosphopeptide enrichment methods, we were able to identify around 150 phosphorylated ligands displayed by this allotype. Analysis of this set of sequences revealed two striking facts: 1) some ligands had phosphoserine (pSer) instead of Glu at P2 and 2) phosphorylation frequently affected the residue at P4 and was accompanied by the presence of a basic residue at P1. By means of binding assays to B*40 and X-ray crystallography we could determine that the presentation of ligands phosphorylated P2 is explained by molecular mimicry between the residues Glu and pSer. Furthermore, we demonstrated that the preference for phosphorylated peptides at P4 is related to the presence of the residue Arg62 in the heavy chain of this allotype and that the selection of basic residues at P1 is a consequence of the structure of the A pocket.

Finally, the analysis of the phosphopeptidome displayed by HLA-B*39 showed that this is also enriched in ligands phosphorylated at P4 but with a lower frequency of positively charged residues at P1. This allows to anticipate that the preference for peptides phosphorylated at P4 is a common feature for most HLA-B molecules while the bias towards basic residues at P1 is allotype-dependent. This understanding of the molecular basis for the presentation of phosphopeptides by HLA-B molecules may contribute to the identification and prediction of phosphorylated epitopes for the treatment of cancer by T cell-based immunotherapy.

Resumen

Las moléculas de clase I del Complejo Mayor de Histocompatibilidad (MHC, HLA en humanos) desempeñan un papel fundamental en la erradicación de infecciones virales y procesos de transformación tumoral. Su función primordial es la presentación a los linfocitos T citotóxicos (CTLs) de un amplio repertorio peptídico generado a partir de la degradación de proteínas endógenas. El reconocimiento de un péptido antigénico en el contexto de una molécula de clase I conduce a la activación del CTL que ejerce su acción citolítica sobre la célula presentadora. El peptidoma presentado por las moléculas HLA-I incluye ligandos fosforilados. Estos han sido propuestos como dianas óptimas para el tratamiento del cáncer por inmunoterapia basada en células T.

Usando HLA-B*40 como modelo, nos propusimos estudiar el peptidoma y fosfopeptidoma presentado por moléculas HLA-I y determinar las bases moleculares de la presentación de ligandos fosforilados. El análisis por LC-MS/MS del repertorio peptídico asociado a B*40 nos permitió caracterizar más de 9500 ligandos únicos con un motivo de anclaje común: residuos ácidos –fundamentalmente Glu– en P2 y residuos alifáticos, Met o Phe en P Ω . Los determinantes estructurales de este motivo fueron analizados mediante la resolución de la estructura cristalográfica de B*40 en complejo con un ligando endógeno.

Gracias al uso de métodos de enriquecimiento de fosfopéptidos, pudimos identificar en torno a 150 ligandos fosforilados presentados por este alotipo. El análisis de este conjunto de secuencias reveló dos hechos llamativos: 1) algunos ligandos presentaban fosfoserina (pSer) en lugar de Glu en P2 y 2) la fosforilación afectaba frecuentemente al residuo en P4 e iba acompañada por la presencia de un residuo básico en P1. Mediante ensayos de unión a B*40 y cristalografía de rayos X pudimos determinar que la presentación de ligandos fosforilados en P2 se explica por el mimetismo molecular entre los residuos de pSer y Glu. Por otra parte, demostramos que la preferencia por péptidos fosforilados en P4 está relacionada con la presencia del residuo Arg62 en la cadena pesada de este alotipo y que la selección de residuos básicos en P1 es consecuencia de la estructura de la subcavidad A.

Por último, el análisis del fosfopeptidoma presentado por HLA-B*39 demostró que este también está enriquecido en ligandos fosforilados en P4 aunque con una menor frecuencia de residuos cargados positivamente en P1. Este hecho, permite anticipar que la preferencia por péptidos fosforilados en P4 es una característica común a la mayoría de moléculas HLA-B mientras que el sesgo hacia residuos básicos en P1 es dependiente de alotipo. El conocimiento de las bases moleculares de la presentación de fosfopéptidos por moléculas HLA-B puede contribuir a la identificación y predicción de epítomos fosforilados para el tratamiento del cáncer mediante inmunoterapia basada en células T.

Índice

| | |
|--|----|
| Abreviaturas | 5 |
| I-Introducción | 7 |
| I.1-Estructura de las moléculas del MHC de clase I | 7 |
| I.2-Características generales de los alotipos HLA-B*39 y HLA-B*40 | 9 |
| I.2.1-HLA-B*39..... | 9 |
| I.2.2-HLA-B*40..... | 10 |
| I.2.3-Diferencias en la estructura primaria de HLA-B*39 y -B*40..... | 10 |
| I.3-Ruta de procesamiento y presentación de antígeno mediada por moléculas del MHC de clase I ... | 11 |
| I.4-Presentación de péptidos fosforilados por moléculas del MHC de clase I..... | 15 |
| I.5-Identificación de fosfopéptidos mediante espectrometría de masas | 16 |
| I.5.1-Técnicas de enriquecimiento de fosfopéptidos..... | 18 |
| I.5.2-Esquemas de fragmentación alternativos al CID | 19 |
| II-Objetivos | 23 |
| III-Materiales y Métodos | 25 |
| III.1-Líneas celulares y anticuerpos monoclonales | 25 |
| III.2-Aislamiento del peptidoma unido a HLA-B*40:02 y -B*39:01 | 25 |
| III.3-Lisis celular y digestión trípica..... | 26 |
| III.4-Enriquecimiento de fosfopéptidos | 26 |
| III.5-Análisis mediante espectrometría de masas en tándem acoplada a HPLC (LC-MS/MS) | 28 |
| III.5.1-Análisis del peptidoma y fosfopeptidoma de HLA-B*40 y HLA-B*39 mediante espectrometría de masas ESI-Q-TOF | 28 |
| III.5.2-Análisis del peptidoma y fosfopeptidoma de HLA-B*40 mediante espectrometría de masas ESI-LTQ-Orbitrap | 29 |
| III.5.3-Análisis del proteoma y fosfoproteoma de la línea celular C1R-B*40 mediante espectrometría de masas ESI-LTQ-Orbitrap | 30 |
| III.5.4-Análisis de péptidos sintéticos por espectrometría de masas..... | 31 |
| III.6-Búsquedas MS/MS e identificación de péptidos | 31 |
| III.7-Análisis estadístico..... | 32 |
| III.8-Síntesis de péptidos..... | 33 |
| III.9-Ensayos de unión de péptido a HLA-B*40 | 33 |
| III.10-Cristalografía..... | 34 |

| | |
|---|----|
| III.10.1-Clonaje, expresión y purificación de proteínas recombinantes | 34 |
| III.10.2-Replegamiento <i>in vitro</i> y cristalización | 35 |
| III.10.3-Difracción de rayos-X y resolución de estructuras | 37 |
| III.11- Predicción <i>in silico</i> de ligandos de HLA-B*40 | 38 |
| IV-Resultados | 39 |
| IV.1-Análisis del peptidoma presentado por HLA-B*40 y bases estructurales de la presentación de péptidos por este alotipo..... | 39 |
| IV.1.1-Distribución por tamaño del peptidoma asociado a HLA-B*40 | 39 |
| IV.1.2-Characterización de los motivos de anclaje de los péptidos presentados por HLA-B*40..... | 41 |
| IV.1.3-Estructura de HLA-B*40 en complejo con un ligando endógeno..... | 44 |
| IV.2-Análisis del fosfopeptidoma presentado por HLA-B*40 y bases estructurales de la presentación de fosfopéptidos por este alotipo..... | 47 |
| IV.2.1-Identificación de ligandos fosforilados unidos a HLA-B*40..... | 47 |
| IV.2.2-El fosfopeptidoma asociado a HLA-B*40 está enriquecido en ligandos derivados de proteínas implicadas en mitosis y regulación del ciclo celular | 51 |
| IV.2.3-HLA-B*40 puede unir eficientemente péptidos fosforilados en P2 | 51 |
| IV.2.4-La interacción de los residuos de fosfoserina o ácido glutámico en P2 con la subcavidad B es muy similar estructuralmente..... | 54 |
| IV.2.5-Los fosfopéptidos unidos a HLA-B*40 se encuentran fosforilados mayoritariamente en P4 y tienden a presentar residuos básicos en P1 | 55 |
| IV.2.5.1-El motivo de fosforilación de los péptidos asociados a HLA-B*40 no está determinado por la especificidad del quinoma celular | 56 |
| IV.2.5.2-La fosforilación en P4 disminuye ligeramente la afinidad de unión a HLA-B*40 | 57 |
| IV.2.5.3-Efecto de la fosforilación en diferentes posiciones del péptido en los ensayos de unión a HLA-B*40..... | 58 |
| IV.2.5.4-Estructuras de HLA-B*40 en complejo con un ligando endógeno fosforilado en P4 y su homólogo no fosforilado | 59 |
| IV.3-El fosfopeptidoma asociado a HLA-B*39 también se encuentra enriquecido en ligandos fosforilados en P4 | 62 |
| V-Discusión | 65 |
| V.1-Characterización y bases estructurales del motivo de unión de péptidos de HLA-B*40..... | 66 |
| V.2-Characterización del fosfopeptidoma unido a HLA-B*40..... | 68 |
| V.2.1-HLA-B*40 puede presentar ligandos con residuos fosforilados en P2..... | 70 |

| | |
|--|-----|
| V.2.2-Los fosfopéptidos unidos a HLA-B*40 se encuentran mayoritariamente fosforilados en P4 y tienden a presentar residuos básicos en P1 | 71 |
| V.3-La preferencia por ligandos fosforilados en P4 podría ser común a la mayoría de moléculas HLA-B | 73 |
| VI-Conclusiones | 75 |
| VII-Referencias | 77 |
| VIII-Anexos | 89 |
| VIII.1-Anexo..... | 89 |
| VIII.1.1-Anexo | 91 |
| VIII.1.2-Anexo | 177 |
| VIII.1.3-Anexo | 203 |
| VIII.2-Anexo..... | 235 |

Abreviaturas

| | |
|------------|--|
| Å | Amstrong |
| ATP | Adenosín trifosfato |
| β_2m | Beta-2-microglobulina |
| CID | Disociación inducida por colisión |
| CTL | Linfocito T citotóxico |
| DHB | ácido 2,5 dihidroxibenzóico |
| DMEM | Medio de Eagle modificado por Dulbecco |
| DRiPs | Productos defectivos de la síntesis ribosomal |
| DTT | Ditiotreitol |
| EDTA | Ácido Etilendiaminotetraacético |
| ERAP | Aminopeptidasa asociada al retículo endoplásmico |
| ETD | Disociación mediante transferencia electrónica |
| ETHcD | Disociación por transferencia electrónica/colisión de alta energía |
| FBS | Suero fetal bovino |
| FDR | Tasa de falsos positivos |
| FWHM | Full Width at Half Maximum |
| HC | Cadena pesada |
| HLA | Antígeno leucocitario humano |
| HTLV-1 | Virus linfotrópico de células T Humanas 1 |
| IMAC | Cromatografía de afinidad con metal inmovilizado |
| IPTG | Isopropil- β -D-tiogalactósido |
| kDa | Kilodalton |
| LC-MS/MS | Cromatografía líquida acoplada a espectrometría de masas en tándem |
| MALDI-TOF | Desorción-ionización mediante láser asistida por matriz – Tiempo de vuelo |
| MHC | Complejo mayor de histocompatibilidad |
| MOAC | Cromatografía de afinidad con óxido metálico |
| MWCO | Corte de peso molecular |
| P | Posición peptídica |
| PBS | Tampón salino de fosfato |
| RE | Retículo Endoplásmico |
| SD | Desviación estándar |
| SDS-PAGE | Electroforesis en geles de poliacrilamida en presencia de dodecil sulfato sódico |
| TAP | Transportador asociado a procesamiento antigénico |
| TCR | Receptor de las células T |
| TFA | Ácido Trifluoroacético |

Introducción

I-Introducción

Gracias al sistema inmunitario, estamos provistos de mecanismos altamente eficaces que protegen nuestro cuerpo frente a agresiones externas o internas, reconociendo lo propio y eliminando lo extraño al organismo. Dentro de los muchos elementos implicados en esta tarea, los linfocitos T citotóxicos (CTLs) son los encargados de reconocer la presencia de patógenos intracelulares o procesos de transformación tumoral a través del receptor de las células T (TCR). Dicho reconocimiento se realiza en el contexto de las denominadas moléculas del Complejo Mayor de Histocompatibilidad (MHC).

En humanos, los genes del MHC conforman el denominado sistema HLA (*Human Leukocyte Antigens*), nombre debido a su descubrimiento como antígenos en los leucocitos que podían ser detectados mediante anticuerpos. Este engloba una familia de genes ubicados en el extremo corto del cromosoma 6 que juegan un papel fundamental en la defensa inmunológica del organismo frente a patógenos y, debido a su alto grado de polimorfismo, constituyen la principal barrera al trasplante de órganos y de células madre [1]. Se han identificado tres regiones en el MHC: de clase I, II y III. Dentro de la región de clase I, se encuentran los loci HLA-A, HLA-B y HLA-C, que codifican las denominadas moléculas del MHC de clase I clásicas cuya función es presentar a los CTLs fragmentos peptídicos derivados de proteínas endógenas [2]. La presencia de un péptido extraño en el contexto de una molécula de clase I conduce al reconocimiento y a la activación de CTLs que ejercen su acción citolítica sobre la célula diana.

El repertorio peptídico presentado por estas moléculas procede fundamentalmente de proteínas degradadas en el citosol o el núcleo celular y su diversidad se ve incrementada por la presentación de ligandos modificados postraduccionalmente [3]. En este contexto, la presentación de péptidos fosforilados por moléculas de clase I ha despertado un interés en los últimos años ya que se ha propuesto que los ligandos fosforilados específicos de tumor podrían ser de utilidad como dianas específicas para la inmunoterapia del cáncer mediante células T [4, 5]. El análisis de las bases moleculares de la presentación de este tipo de ligandos es el objetivo troncal de la presente tesis doctoral.

I.1-Estructura de las moléculas del MHC de clase I

Los genes del MHC de clase I codifican glicoproteínas de membrana formadas por la asociación no covalente de una cadena pesada polimórfica de 44 kDa, una cadena ligera de 12 kDa, denominada beta

2-microglobulina (β_2m), y un péptido cuyo tamaño oscila entre, típicamente, 8 y 13 aminoácidos. Se ha logrado determinar de una manera muy precisa la estructura atómica y tridimensional de estas proteínas gracias al gran número de estudios cristalográficos realizados sobre moléculas del MHC de clase I (Figura 1A) [6-11]. La cadena pesada consta de una región extracelular de 274 residuos formada por tres dominios proteicos globulares $-\alpha_1, \alpha_2$ y α_3- de aproximadamente 90 aminoácidos cada uno [12], una porción transmembrana de unos 25 residuos y un tallo citoplasmático de alrededor de 30 residuos.

El péptido se aloja en una cavidad o sitio de unión que se forma entre los dominios α_1 y α_2 . Esta estructura consta de una lámina β de 8 bandas antiparalelas constreñida lateralmente por dos hélices α . El dominio α_3 , situado en la proximidad de la membrana plasmática, presenta una estructura similar a los dominios constantes de las inmunoglobulinas: dos láminas β unidas a través de un puente disulfuro, a modo de sándwich. La β_2m se encuentra asociada a la cadena pesada, contribuyendo a la estabilidad del complejo [8].

El sitio de unión presenta unas dimensiones de $25 \times 12 \text{ \AA}$ [8, 13-16] y aloja el péptido con los extremos amino y carboxilo anclados a la cadena pesada mediante una red de puentes de hidrógeno (Figura 1B). Los ligandos de 8 o 9 residuos de aminoácido adoptan una conformación extendida a lo largo de la cavidad entre α_1 y α_2 , mientras que la región central de los péptidos más largos se curva hacia el exterior del sitio de unión, conservándose los anclajes N y C-terminales. La asociación entre la molécula de clase I y el péptido no se encuentra únicamente restringida por la longitud de éste. Un análisis detallado del sitio de unión permite definir 6 subcavidades denominadas desde la A a la F, cuya topología, tamaño y polaridad vienen determinadas por las cadenas laterales de los residuos que las conforman [14, 17, 18]. Las cadenas laterales de algunos residuos del péptido interactúan con el interior de las distintas subcavidades. Las subcavidades B y F, que son habitualmente las que alojan los residuos de anclaje principales, interactúan con los aminoácidos en posición (P) 2 y C-terminal, respectivamente [14, 18]. Los residuos ubicados en la región central del péptido suelen estar orientados hacia el solvente y participan en la interacción con el TCR [13, 16].

Obviamente, las propiedades químicas y estructurales de las subcavidades imponen restricciones al tipo de ligandos que pueden interactuar con el surco de unión, seleccionando un subconjunto de péptidos para su presentación al sistema inmunitario. Un rasgo característico de las moléculas del MHC de clase I es el grado extremo de polimorfismo que presentan los dominios α_1 y α_2 [19, 20] y que determina la especificidad de unión de péptidos de cada alotipo [21]. Según la base de datos IMGT-HLA (www.ebi.ac.uk/ipd/imgt/hla), en marzo de 2016 se encontraban reportados 10 437 alelos HLA-A, -B o -

C que codificaban 7534 proteínas distintas. Asimismo, el repertorio peptídico asociado a un alelo en particular puede incluir más de 10.000 especies diferentes con motivos estructurales comunes y bien definidos que permiten su unión a la molécula de clase I [22-24].

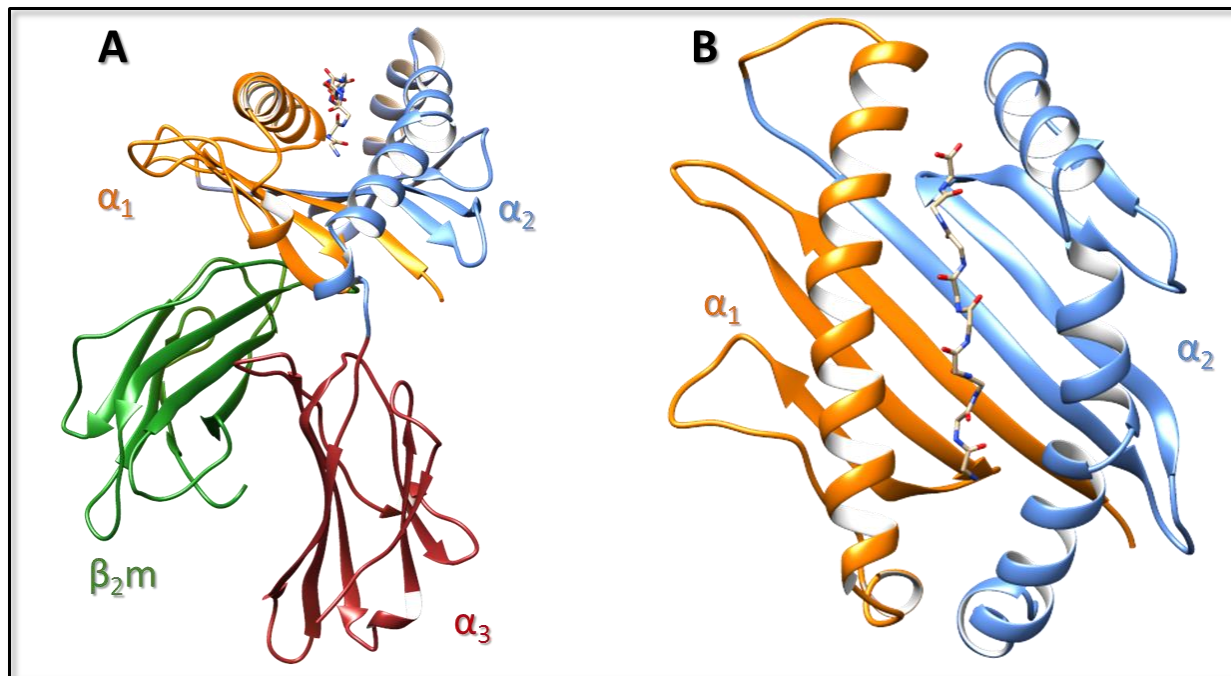


Figura 1. Estructura de una molécula del MHC clase I. (A) Se representa la región extracelular de la cadena pesada compuesta por α_1 , α_2 y α_3 , de color naranja, azul y rojo respectivamente. La β_2m se muestra en color verde y los carbonos alfa del ligando, ubicado entre α_1 y α_2 , usando el código de color CPK (carbono en gris, nitrógeno en azul y oxígeno en rojo). (B) Vista del surco de unión en el que se aloja el péptido. Nuevamente, α_1 y α_2 se representan de color naranja y azul, respectivamente. El péptido se representa mediante el código de color antes mencionado.

I.2- Características generales de los alotipos HLA-B*39 y HLA-B*40

En la presente tesis, se han usado como modelo biológico las moléculas del MHC de clase I HLA-B*39:01 y HLA-B*40:02. A continuación se describen sus características básicas respecto a su distribución geográfica, asociación a patologías y sus motivos de unión de péptido.

I.2.1- HLA-B*39

HLA-B*39 incluye 151 alelos que codifican 68 proteínas diferentes. La prevalencia de este antígeno es especialmente alta entre la población amerindia –con una frecuencia superior al 9%– aunque también está presente entre el 2 y el 3% de los individuos caucosoides, africanos, aborígenes australianos y orientales [25].

HLA-B*39 se asocia con espondilitis anquilosante en pacientes HLA-B*27 negativos y con artritis reumatoide pauciarticular juvenil [26]. Asimismo, ha sido sugerido como factor de riesgo en la arteritis de Takayasu [27], artritis asimétrica [28], artritis psoriásica [29-31], cálculos biliares [32] y enfermedad de Still [30]. Por último, se ha observado que determinadas complicaciones osteoarticulares derivadas de episodios de brucelosis parecen estar asociadas con HLA-B*39 [33].

A partir de estudios de secuenciación de Edman en *pool* [34] y de análisis mediante LC-MS/MS [35] se ha podido determinar que el alotipo HLA-B*39:01 une preferentemente ligandos con residuos de His o Arg en P2 y residuos hidrofóbicos en el C-terminal.

I.2.2-HLA-B*40

HLA-B*40 forma parte del supertipo B44 [36] y, a fecha de marzo de 2016, incluye 420 alelos que codifican 120 proteínas distintas. Es especialmente frecuente (9-12%) entre los aborígenes australianos y las poblaciones amerindias y orientales. Está también presente en el 3% de individuos caucasoides y casi ausente en personas de origen africano [25].

Se ha demostrado que este alelo es un factor de riesgo de leucemia T del adulto, un tipo de linfoma causado por el virus linfotrópico de células T Humanas 1 (HTLV-1) [37]. Al parecer, esta asociación se explica por la capacidad limitada de HLA-B*40 para presentar epítomos derivados de HTLV-1 y desencadenar una respuesta citotóxica adecuada. Asimismo, se encuentra asociado a espondiloartropatías B*27 positivas y negativas [38-41] y se ha sugerido como marcador de susceptibilidad de daño pulmonar en pacientes de artritis reumatoide [42].

El primer estudio sobre el motivo de unión de HLA-B*40 se basó en la identificación de 22 ligandos de esta molécula [43] y demostró la preferencia de este alotipo por residuos de Glu en P2 e hidrofóbicos en el C-terminal. Aunque los autores no llegaron a caracterizar el subtipo de B*40 que analizaron, el motivo de anclaje propuesto (X-E-I/F-X-X-X-X-I) es muy similar al que se describió para el alotipo B*40:01 en base a la identificación de 60 ligandos endógenos: X-E-X-X-X-I/V/L-X-L [44].

I.2.3-Diferencias en la estructura primaria de HLA-B*39 y -B*40

HLA-B*40:02 y HLA-B*39:01 difieren en 13 residuos en la secuencia de aminoácidos de los dominios α_1 o α_2 , que forman el surco de unión de ligandos (Figura 2). Como se ha comentado, el motivo de anclaje de ambas moléculas es bastante diferente. HLA-B*40 une residuos ácidos como Glu

en P2 [44], mientras que HLA-B*39 prefiere residuos como His o Arg en esta posición [34, 35]. Sin embargo, ambos alotipos presentan un motivo estructural parcialmente similar en el C-terminal.

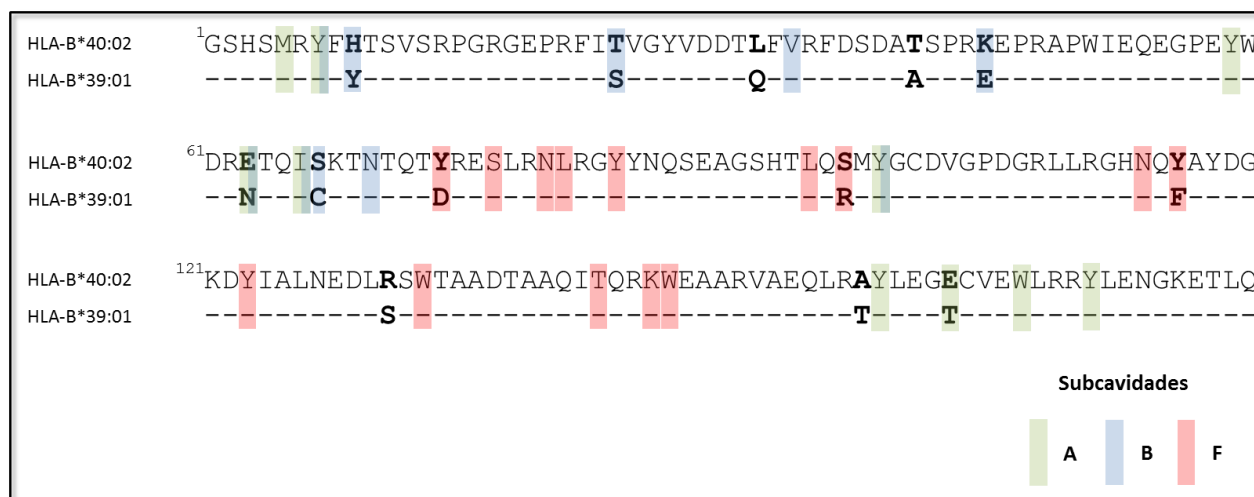


Figura 2. Comparación de la estructura primaria de los dominios α_1 y α_2 de HLA-B*40:02 y HLA-B*39:01. Los residuos conservados en ambas secuencias se presentan en líneas discontinuas y las posiciones polimórficas en negrita. Los residuos de las subcavidades A, B y F de ambas moléculas son representados en color verde, azul y rojo, respectivamente.

I.3-Ruta de procesamiento y presentación de antígeno mediada por moléculas del MHC de clase I

La presentación de péptidos antigénicos a los CTLs por moléculas HLA de clase I es el resultado de una serie de mecanismos celulares y procesos bioquímicos que permiten que péptidos generados por proteólisis sean transportados al lumen del retículo endoplásmico donde se unen a la cadena pesada y a la β_2m (Figura 3). Finalmente, este complejo formado es exportado a la superficie celular donde puede ser reconocido por el TCR de las células T CD8+.

Como se ha comentado anteriormente, la mayor parte de los péptidos unidos a moléculas de clase I proceden de la degradación de proteínas endógenas en el núcleo o el citosol [45]. Este proceso de degradación es llevado a cabo, fundamentalmente, por el proteasoma [46, 47]. En los primeros estudios sobre presentación antigénica mediada por este tipo de moléculas, se consideraba que los ligandos de clase I derivan de proteínas maduras que son ubiquitinadas y degradadas al final de su vida útil. No obstante, cada vez ha ido cobrando más fuerza la hipótesis de que los principales sustratos de la vía son productos defectivos de la síntesis proteica generados por el ribosoma (DRiPs) [48-52].

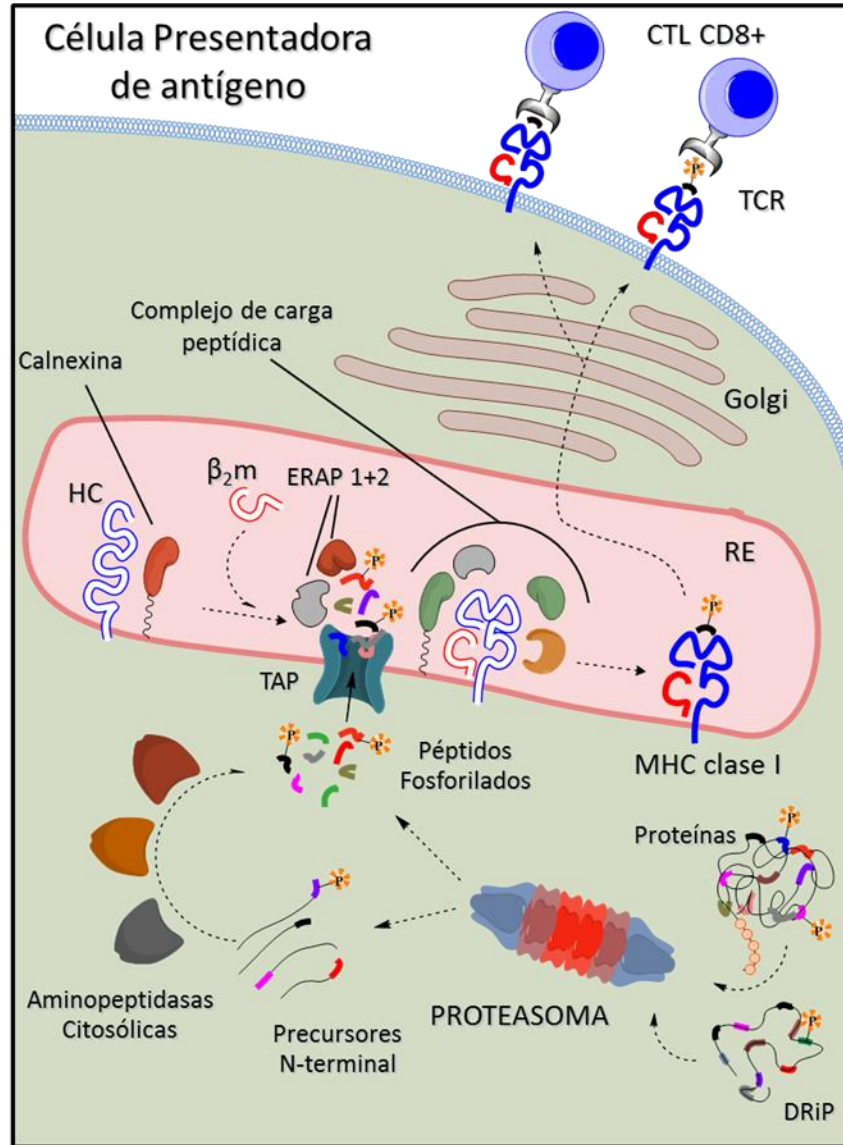


Figura 3. Ruta de procesamiento y presentación antigénica medida por moléculas del MHC de clase I. El proteasoma es el encargado de degradar DRiPs y proteínas que se encuentran al final de su vida útil en el núcleo o el citosol celulares. Los ligandos definitivos o sus precursores con extensiones amino terminales son transportados a través de TAP al retículo endoplásmico (RE). Una vez dentro del RE, la longitud de los péptidos es optimizada por ERAP1 y ERAP2 para su unión a la molécula MHC-I en un proceso asistido por chaperonas. Finalmente, tras la unión de un péptido de alta afinidad, se produce la disociación del complejo de carga y la exportación del complejo trimérico a la superficie celular para ser presentado a los linfocitos T.

El proteasoma es un complejo multiproteico grande y muy conservado cuya función principal es la degradación enzimática de proteínas defectuosas o que se encuentran al final de su vida útil [53-56]. El núcleo central, o proteasoma 20S, es un cilindro formado por 28 subunidades [56] distribuidas en dos

anillos exteriores constituidos por siete subunidades α y dos anillos interiores formados por siete subunidades β . Tres de las subunidades β – β_1 , β_2 y β_5 – son catalíticamente activas [57-59]. Los extremos N-terminales de las subunidades α forman una “compuerta” que puede abrirse para permitir el paso de las proteínas diana al interior del proteasoma. A través de esta región el proteasoma interacciona con determinados complejos reguladores como PA700, o complejo 19S [60, 61], que está implicado en el reconocimiento y la degradación de proteínas ubiquitinadas de manera dependiente de ATP [62, 63]. En células eucariotas las proteínas se marcan para su degradación mediante la adición de varias moléculas de ubiquitina. Esta “etiqueta” de poli-ubiquitina es reconocida por algunas subunidades del complejo 19S facilitando el despliegue de la proteína y su ingreso al canal central del proteasoma donde será accesible a los centros activos de las subunidades catalíticas. Globalmente, este proceso genera péptidos de entre 3 y 22 residuos de aminoácido que salen del proteasoma a través de uno o ambos poros abiertos [64].

Sólo un pequeño porcentaje de los péptidos generados por el proteasoma tiene el tamaño adecuado para su unión a moléculas del MHC de clase I [64]. El consenso general es que el proteasoma tiende a generar precursores con el C-terminal definitivo, mientras que el corte amino terminal no siempre es tan preciso, dando lugar a precursores extendidos que requieren de un procesamiento proteolítico posterior [65]. Se ha sugerido la participación de diversas aminopeptidasas citosólicas en el recorte de estos precursores, entre ellas la oligopeptidasa TOP (*thimet oligopeptidase*) [66], la bleomicín hidrolasa (BH) [67], la aminopeptidasa sensible a puromicina (PSA) [67, 68], la tripeptidil peptidasa II (TPPII) [68], o la nardilisina [69]. No obstante, la contribución global a la generación de ligandos de clase I de estas aminopeptidasas es, probablemente, de escasa magnitud.

Los péptidos generados en el citosol por el proteasoma pasan al retículo endoplásmico por medio del transportador asociado con procesamiento antigénico (TAP), un miembro de la familia de transportadores ABC (*ATP Binding Cassette*). TAP es un heterodímero formado por dos subunidades, TAP1 y TAP2, ambas necesarias para formar un canal que permite el tránsito de los péptidos de forma unidireccional [70]. TAP transporta ligandos con longitudes comprendidas entre 8 y 16 residuos de forma dependiente de ATP y su especificidad de sustrato está restringida por el reconocimiento del residuo C-terminal y los 3 primeros residuos N-terminales [71-73]. En este sentido, TAP favorece el transporte de péptidos con residuos C-terminales básicos o, fundamentalmente, hidrofóbicos siendo esta preferencia compatible con los motivos de unión de las moléculas HLA de clase I [74]. Se sabe, no obstante, que algunos alotipos de clase I son capaces de unir y presentar ligandos con motivos N-

terminales desfavorables para su transporte. Por ejemplo, los péptidos con Pro en posición 2 presentan una muy baja afinidad por TAP [75]. Estos son translocados en forma de precursores N-terminales para ser procesados mediante una actividad aminopeptidasa en el lumen del retículo endoplásmico que genera la versión definitiva del ligando que será presentado por la molécula de clase I [55, 76].

El recorte de las extensiones N-terminales de los precursores de ligando depende fundamentalmente de ERAP1 y ERAP2 (*Endoplasmic Reticulum Aminopeptidase 1/2*). Estas son dos metalopeptidasas dependientes de zinc residentes en el retículo endoplásmico. Estudios basados en digestiones *in vitro* de péptidos [77] revelan que ERAP1 recorta de forma óptima péptidos con una longitud entre 9 y 16 aminoácidos y los productos raramente tienen menos de 8 ó 9 aminoácidos, lo que sugiere que ha evolucionado para generar péptidos del tamaño adecuado para la unión a moléculas de MHC de clase I [78-80]. En cuanto a su especificidad, ERAP1 prefiere residuos grandes e hidrofóbicos tanto en el extremo amino como en el carboxilo del precursor, mientras que aminoácidos cargados o pequeños e hidrofóbicos se recortan con muy baja eficiencia [77, 81]. Esto es consistente con el tamaño y la especificidad de TAP para unir y translocar péptidos al retículo endoplásmico así como con las preferencias de unión de las moléculas de clase I. ERAP1 degrada péptidos mediante un mecanismo de “regla molecular” [77] de modo que cuando el péptido alcanza una longitud de 8 o 9 residuos deja de ser un buen sustrato para la peptidasa. Esta longitud mínima viene determinada por la distancia entre el sitio de unión del extremo C-terminal y el sitio catalítico de la enzima [82, 83].

ERAP1 actúa de forma concertada con ERAP2 [84-87], esta última muestra una fuerte preferencia por sustratos con residuos básicos tanto en el extremo amino como en el carboxilo [88, 89]. Sin embargo, los experimentos realizados por Saveanu y cols demostraron que, por sí sola, ERAP2 no es capaz de eliminar este tipo de residuos si no que requeriría de la acción conjunta de ERAP1 [86]. Evnouchidou y cols han propuesto que la formación de heterodímeros de ERAP1 y ERAP2 juega un papel crítico en la producción de los péptidos que serán presentados por moléculas del MHC de clase I, aumentando la eficiencia y facilitando el recorte de péptidos precursores [85]. No obstante, los efectos de la dimerización sobre la función ERAP aún tienen que ser analizados en profundidad siendo necesarios más estudios funcionales y estructurales para determinar la naturaleza exacta de estos heterodímeros y las modificaciones alostéricas provocadas por la interacción de ERAP1 y ERAP2.

La unión del péptido a la molécula de clase I en el interior del retículo endoplásmico se ve asistida por el denominado complejo de carga. Éste es una estructura multiproteica formada por TAP, la cadena

pesada, la β_2m , la oxidoreductasa ERp57 [90], la calreticulina –una chaperona que estabiliza la molécula de clase I en ausencia de ligando [91]– y la tapasina [92, 93] que actúa de puente con TAP. El papel del complejo de carga es favorecer el plegamiento de la molécula de clase I y la incorporación de un ligando de alta afinidad al complejo trimérico. Cuando esto sucede, el complejo de carga se disocia y la molécula de clase I es transportada a través del aparato de Golgi hasta la superficie celular donde puede ser reconocida por los linfocitos T [45].

I.4-Presentación de péptidos fosforilados por moléculas del MHC de clase I

Las moléculas de clase I presentan péptidos con modificaciones postraduccionales en la superficie celular [3, 94, 95]. Entre éstas, se encuentran la acetilación N-terminal [96], la fosforilación [4, 5, 97], la metilación [98], la glicosilación [99-101] o la oxidación de cisteína a cistina [102].

Los péptidos fosforilados asociados a moléculas del MHC de clase I presentan un especial interés ya que han sido propuestos como dianas potenciales para la inmunoterapia del cáncer basada en células T [4, 5, 103]. La idoneidad de este tipo de ligandos se fundamenta, por un lado, en que las células tumorales muestran frecuentemente patrones de fosforilación aberrantes [104] y, por otro, en que los péptidos fosforilados pueden ser transportados desde el citosol al retículo endoplásmico mediante TAP y formar complejos estables con la cadena pesada y la β_2m que pueden ser reconocidos por CTLs de manera específica (Figura 3) [97, 105]. La caracterización del fosfopeptidoma asociado a moléculas de clase I y la identificación de fosfopéptidos derivados de tumores son necesarios para poder implementar estos enfoques inmunoterapéuticos. No obstante, y a pesar de la relevancia fisiológica de este tipo de ligandos, el número de estudios al respecto es limitado y, hasta la fecha, solo se ha caracterizado el fosfopeptidoma presentado por unos pocos alotipos.

Por ejemplo, Zarlino y cols identificaron, a partir de células de un carcinoma pulmonar y varias líneas linfoblastoides, 15 fosfopéptidos asociados a diversas moléculas de clase I, entre ellos 6 ligandos de HLA-B*27 y 6 de HLA-B*07 [97]. En otro estudio del mismo grupo, se secuenciaron 36 fosfopéptidos presentados por HLA-A*02 a partir de células de melanoma, carcinoma de ovario y células linfoblastoides [4]. Meyer y cols caracterizaron 11 ligandos fosforilados, 8 de ellos asociados a HLA-B*07, a partir de un carcinoma renal y la línea linfocítica JY [5]. Por último, Cobbold y cols describieron, a partir de células primarias de linfoma y tejido sano, 10 y 85 fosfopéptidos restringidos por HLA-A*02 y HLA-B*07, respectivamente [106]. Curiosamente, este conjunto de trabajos sugiere la existencia de un motivo estructural común en el fosfopeptidoma asociado a estos diferentes alotipos. En primer lugar, en

promedio, un 73% de los ligandos fosforilados presentan esta modificación en P4 de la secuencia peptídica. Además, en torno al 77% de los péptidos fosforilados en P4 incluye un residuo básico en P1 (Tabla 1).

Las bases estructurales de estas preferencias solo han sido investigadas con cierta profundidad en el caso de HLA-A*02. Mohammed y *cols* resolvieron la estructura cristalográfica de cuatro complejos de esta molécula con péptidos que estaban fosforilados en P4 y presentaban residuos cargados positivamente en P1 [103]. En todos los casos, la estructura del complejo se encontraba estabilizada por dos tipos de interacciones que involucraban al grupo fosfato: 1) interacción de la fosfoserina con los residuos Arg65 y/o Lys66 de la cadena pesada y 2) formación de un puente salino entre el grupo fosfato en P4 y el residuo básico en P1. Un estudio posterior por Petersen y *cols* determinó la estructura del mismo alotipo en complejo con dos ligandos fosforilados en P4, uno de ellos con Arg en P1 [95]. En este caso, no se observaron interacciones electrostáticas entre los residuos P1 y P4, sugiriendo que estas son dependientes de epítipo. Por el contrario, se confirmó la importancia de la interacción de la fosfoserina con los residuos Arg65 y Lys66 para la estabilización del complejo. El hecho de que estos dos residuos, en especial Arg65, se encuentren ampliamente conservados permite suponer que la presentación de ligandos fosforilados preferentemente en P4 puede ser una constante en las moléculas HLA-A [103].

| Alotipo | Número de Fosfopéptidos | XXX(pS/pT) | (K/R)XX(pS/pT) | Referencia |
|----------|-------------------------|------------|----------------|--------------------------|
| HLA-B*07 | 6 | 5 (83%) | 5 (100%)* | Zarling et al 2000 [97] |
| HLA-B*07 | 8 | 5 (63%) | 5 (100%)* | Meyer et al 2009 [5] |
| HLA-B*07 | 85 | 62 (73%) | 43 (69%)* | Cobbold et al 2013 [106] |
| HLA-A*02 | 36 | 24 (67%) | 18 (75%)* | Zarling et al 2006 [4] |
| HLA-A*02 | 10 | 7 (70%) | 7 (100%)* | Cobbold et al 2013 [106] |
| HLA-B*27 | 6 | 6 (100%) | 5 (83%)* | Zarling et al 2000 [97] |

Tabla 1. Características moleculares de los fosfopéptidos asociados a diferentes moléculas HLA de clase I. *Porcentaje referido al número absoluto de péptidos fosforilados en P4.

I.5-Identificación de fosfopéptidos mediante espectrometría de masas

La fosforilación es una modificación post-traducciona reversible clave en la mayoría de los procesos biológicos de procariontas y eucariontas. Respuestas fisiológicas como la adaptación a los cambios del entorno celular están reguladas por la fosforilación y defosforilación de múltiples proteínas llevadas a cabo por kinasas y fosfatasa, respectivamente. Debido a su papel fundamental en la biología celular, no es sorprendente que, tal y como se mencionó anteriormente, los procesos de transformación

tumoral, que cursan con una fuerte desregulación de la célula, se encuentren asociados con patrones de fosforilación aberrantes [104]. En eucariotas la fosforilación afecta, fundamentalmente, a residuos de Ser, Thr y Tyr en una relación del orden de 1800:200:1 [107] aunque también puede ocurrir en residuos de histidina, arginina o lisina [108].

Los abordajes proteómicos actuales permiten estudiar los patrones de fosforilación de un buen número de proteínas y constituyen una herramienta muy poderosa para analizar respuestas celulares a nivel global [109]. Gracias a ellos, ha sido posible efectuar los estudios antes mencionados relativos al fosfopeptidoma presentado por moléculas del MHC de clase I [4, 5, 110]. No obstante la identificación de fosfopéptidos por espectrometría de masas no está exenta de problemas. Existen tres obstáculos fundamentales que dificultan la identificación de especies fosforiladas mediante esta aproximación. En primer lugar, su baja estequiometría: normalmente, las proteínas fosforiladas se encuentran en menor proporción respecto a sus equivalentes no fosforiladas [111]. En segundo lugar, la presencia de un grupo fosfato en la secuencia del péptido afecta negativamente a la eficiencia de ionización lo que hace que incluso a concentraciones equivalentes las especies fosforiladas sean más difíciles de detectar que las no fosforiladas [112, 113]. Por último, la fragmentación de fosfopéptidos mediante CID (*Collision-Induced Dissociation*) –la técnica más empleada hoy en día– tiende a generar espectros MS/MS con escasa información estructural [114].

Respecto a este último punto, es sabido que el grupo fosfato de los residuos de fosfoserina y fosfotreonina es particularmente lábil y sufre reacciones de beta eliminación durante la fragmentación por CID, liberándose una molécula de ácido fosfórico (H_3PO_4) y dando lugar a dehidroalanina y ácido dehidroaminobutírico, respectivamente [114, 115]. Como consecuencia, en el correspondiente espectro de fragmentación, se observa una pérdida de 98 (carga 1+), 49 (carga 2+) o 32.7 (carga 3+) unidades de m/z respecto a la masa del ion precursor. Aunque es menos frecuente, la eliminación del grupo fosfato puede producirse también a través de la liberación de ácido metafosfórico (HPO_3) que se refleja en el espectro MS2 una diferencia respecto al ion precursor de 80 (carga 1+), 40 (carga 2+) o 26.7 (carga 3+) unidades de m/z. En ambos casos, la señal correspondiente a la pérdida del grupo fosfato suele ser muy prominente en el espectro MS/MS apantallando el resto de señales a partir de las que se deduce la secuencia del péptido precursor y dificultando la identificación no ambigua del mismo.

La fosfotirosina, por su parte, no es susceptible de perder el grupo fosfato mediante beta eliminación debido a la naturaleza aromática de su cadena lateral [114, 115]. En este caso, tras la fragmentación por CID, únicamente se observa la pérdida de ácido metafosfórico y con una frecuencia

muy inferior a la observada en especies con fosfoserina o fosfotreonina. Esto facilita la identificación de péptidos fosforilados en tirosina aunque, como se mencionó más arriba, la frecuencia de estos es mínima dentro del proteoma.

Para solventar las dificultades relacionadas con la identificación de fosfopéptidos por espectrometría de masas, las aproximaciones fosfoproteómicas actuales se basan en la combinación de dos estrategias: 1) el enriquecimiento de las especies fosforiladas antes del análisis por LC-MS/MS y 2) la utilización de esquemas de fragmentación alternativos al CID que eviten la pérdida del grupo fosfato durante el proceso de fragmentación.

I.5.1-Técnicas de enriquecimiento de fosfopéptidos

El enriquecimiento de fosfopéptidos a partir de mezclas complejas puede llevarse a cabo mediante diversas técnicas como: la inmunoafinidad mediante anticuerpos anti-fosfoserina [116, 117], anti-fosfotreonina [118] o anti-fosfotirosina [119, 120]; la cromatografía de intercambio iónico [121] o de interacción hidrofílica [122]; y mediante el uso de cromatografía de afinidad de óxido metálico (MOAC, *Metal Oxide Affinity Chromatography*) o de metal inmovilizado (IMAC, *Immobilized Metal Affinity Chromatography*). De lejos, la cromatografía de afinidad es el abordaje más utilizado actualmente debido a su alta eficiencia, relativo bajo coste y facilidad de implementación.

La cromatografía de tipo MOAC usando TiO_2 como ligando (Figura 4A) fue propuesta por Pinkse y cols quienes demostraron que la capacidad de este óxido metálico para interaccionar específicamente con fosfopéptidos aumentaba bajo condiciones ácidas [123]. Los fosfopéptidos eran subsecuentemente eluidos de la columna usando un tampón alcalino y, tras concentrarlos en una columna de fase reversa, eran analizados por LC-MS/MS. La especificidad de la interacción del TiO_2 con los fosfopéptidos aumenta cuando el proceso de enriquecimiento se lleva a cabo en presencia de determinados ácidos. En este sentido, el uso de ácido 2,5 dihidroxibenzóico (DHB) disminuye el número de especies no fosforiladas en la fracción enriquecida [111, 124]. Aparentemente, este hecho se basa que las moléculas de DHB compiten eficientemente con los péptidos no fosforilados por la interacción con el TiO_2 , mientras que la unión de los fosforilados no se ve afectada. Como alternativa, la selectividad de la matriz de TiO_2 por fosfopéptidos puede incrementarse aún más mediante el uso de ácido glicólico durante la separación cromatográfica [125].

Por su parte, el IMAC (Figura 4B) se basa en la unión de fosfopéptidos, cargados negativamente, a iones metálicos (típicamente Fe^{3+} , Ga^{3+} o Al^{3+} o Zr^{4+}), y ha sido usada extensivamente para el análisis de fosfoproteomas [110, 126]. De manera general, la muestra es disuelta en una solución ácida antes de su aplicación a la columna cromatográfica. El pH de elección suele ser de 1,0-2,9 para promover interacciones entre los fosfopéptidos y la fase estacionaria, y para minimizar la unión de los péptidos no fosforilados cargados negativamente [127]. Esto se debe a que el pKa del grupo fosfato es 1, y el del grupo carboxilo es 4 de tal forma que, en condiciones muy ácidas, el primero presentará carga negativa mientras que el segundo estará protonado y será eléctricamente neutro.

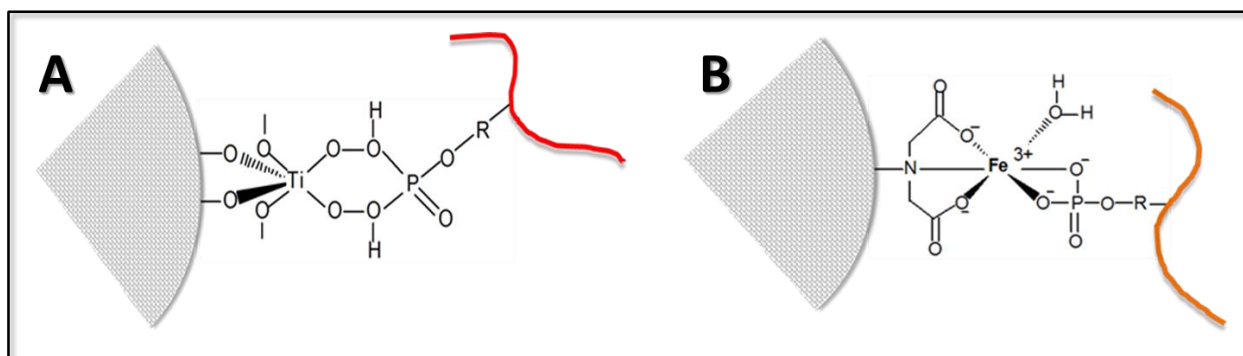


Figura 4. Representación esquemática de (A) matriz MOAC utilizando TiO_2 como ligando unida a un fosfopéptido (en color rojo) y (B) matriz IMAC utilizando Fe^{3+} como ligando unida a un fosfopéptido (en color naranja).

Además de los metales usados habitualmente para el enriquecimiento de fosfopéptidos mediante IMAC, algunos autores han demostrado las ventajas de otros ligandos para este cometido. En este sentido, Zhou y cols emplearon microesferas recubiertas por un espaciador flexible con un grupo fosfato en su extremo libre [128-130]. Los grupos fosfatos de la superficie actúan como quelantes de átomos de Ti^{4+} creándose una matriz capaz de interaccionar específicamente con fosfopéptidos. Algunos autores consideran que este abordaje ofrece una mayor selectividad y eficiencia en comparación con otros métodos como el IMAC- Fe^{3+} , IMAC- Zr^{4+} , MOAC- TiO_2 o MOAC- ZrO_2 [129].

I.5.2-Esquemas de fragmentación alternativos al CID

Los primeros métodos de fragmentación alternativos al CID, que surgieron a finales del pasado siglo y principios de este, son el ECD (*Electron Capture Dissociation*) [131] y el ETD (*Electron Transfer Dissociation*) [132]. En ambas técnicas se transfiere un electrón al péptido precursor para crear un radical catiónico inestable. En el caso del ETD, la transferencia de dicho electrón no se produce directamente sino que requiere un compuesto intermediario que suele ser un radical aniónico del

fluoranteno o el antraceno. La fragmentación del péptido tras la captura del electrón se produce por la rotura de los enlaces N-C α generando fundamentalmente iones de las series c y z. A diferencia del CID, la fragmentación del péptido no se ve influida por la secuencia del mismo, lo que permite identificaciones con mayor cobertura. Además, estos métodos preservan el grupo fosfato unido al péptido lo que facilita enormemente la identificación de especies fosforiladas [114].

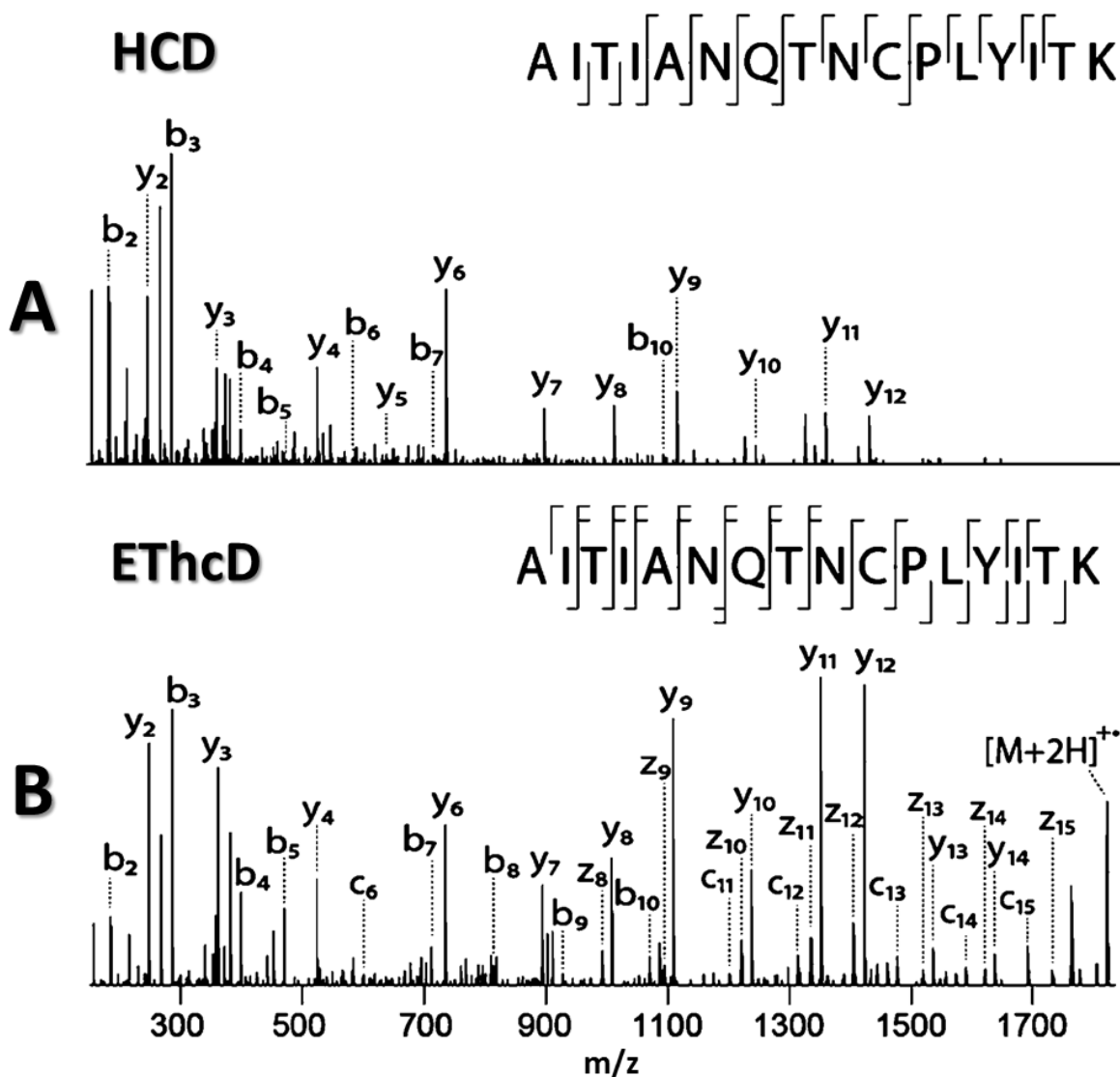


Figura 5. Fragmentación de un péptido tríptico doblemente cargado mediante HCD o ETHcD. (A) Espectro de fragmentación por HCD (CID) en el que se observan 18 iones pertenecientes a las series b/y. (B) Espectro de fragmentación del mismo péptido mediante ETHcD. Nótese que en este caso se identifican 35 iones correspondientes a las series b/y y c/z. Imagen cortesía de Frese *et al.* Anal Chem. 2012 [133].

Uno de los problemas asociados con los esquemas ECD y ETD es su baja eficiencia de fragmentación que hace que la señal correspondiente al ion que ha capturado el electrón sea muy abundante en el espectro de fragmentación [134]. Para solventar este inconveniente se han diseñado varias aproximaciones que aumentan la energía interna del péptido a fragmentar, típicamente mediante activación por colisión. Por ejemplo, Swaney y cols introdujeron la fragmentación ETcaD, en la que el radical catiónico no disociado se fragmenta mediante CID de baja energía [135]. Por su parte, Liu y Lai idearon el esquema ET/CID (*Electron-Transfer coupled with Collision-Induced Dissociation*) donde el precursor de carga reducida que no ha sido fragmentado es aislado para una nueva ronda de fragmentación (MS3) mediante CID [136]. Por último, Frese y cols desarrollaron la fragmentación ETHcD (*Electron-Transfer/Higher-energy Collision Dissociation*) en la que tras la reacción de ETD se induce la fragmentación del precursor y de los iones producto en una celda de colisión HCD [133], esta última es un tipo de fragmentación CID característica de los espectrómetros de masas con analizador Orbitrap que se lleva a cabo en una celda de colisión independiente de la trampa de iones.

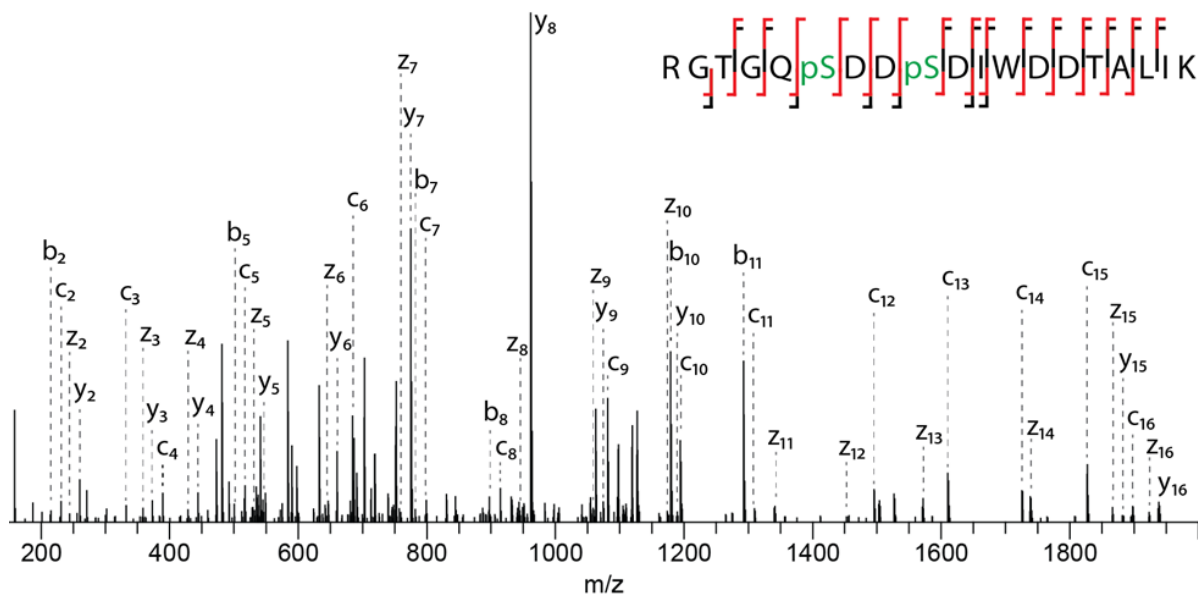


Figura 6. Espectro de fragmentación por ETHcD de un péptido triptico doblemente fosforilado. Se generan abundantes iones de las series c/z y b/y que permiten la identificación precisa de los sitio de fosforilación. Imagen cortesía de Frese *et al.* J Proteome Res. 2013 [137].

A diferencia de las dos aproximaciones anteriores, el ETHcD genera espectros de fragmentación ricos en iones de tipo c/z y b/y lo que aumenta sensiblemente la cobertura de secuencia en la identificación de péptidos (Figura 5). También se ha demostrado su alta eficacia a la hora de caracterizar especies fosforiladas y de identificar con precisión su sitio de fosforilación (Figura 6) [137]. Por último,

cuando se aplicó al estudio de ligandos de moléculas del MHC de clase I, la fragmentación EThcD arrojó mejores resultados que el CID o el ETD tanto en número absoluto de identificaciones como en porcentaje de espectros MS2 con secuencia asignada [24]. Cabe destacar que en este último estudio se identificaron 59 ligandos fosforilados a pesar de que no se llevó a cabo ningún paso de enriquecimiento de fosfopéptidos, demostrando que la fragmentación EThcD es una herramienta muy útil para el análisis de fosfopeptidomas asociados a moléculas de clase I.

Objetivos

II-Objetivos

Como propósito central de esta tesis, nos propusimos aislar y caracterizar en profundidad el peptidoma y el fosfopeptidoma presentado por moléculas HLA de clase I. Para ello, planteamos los siguientes objetivos concretos:

- 1- **Caracterización del peptidoma asociado a HLA-B*40:** determinación del motivo de unión de péptidos de este alotipo y análisis de las bases moleculares de dicho motivo.
- 2- **Caracterización del fosfopeptidoma asociado HLA-B*40:** puesta a punto de una estrategia para la identificación masiva de ligandos fosforilados y estudio de los determinantes estructurales de la presentación de fosfopéptidos por esta molécula.
- 3- **Caracterización del fosfopeptidoma asociado a otras moléculas de clase I:** aislamiento e identificación de ligandos fosforilados presentados por HLA-B*39.

Materialles y Métodos

III-Materiales y Métodos

III.1-Líneas celulares y anticuerpos monoclonales

Hmy2.C1R (C1R) es una línea linfoide con baja expresión de sus antígenos de clase I endógenos HLA-B*35:03 y HLA-C*04:01 [138]. El transfectante estable de HLA-B*39:01 en esta línea (C1R-B*39) ha sido descrito previamente [35]. Para generar un transfectante estable que expresara HLA-B*40:02 (C1R-B*40), el cDNA correspondiente a este alelo se obtuvo mediante transcripción inversa a partir de la línea celular LCL 143.2 [139] para clonarlo posteriormente en el vector RSV5neo. En torno a 10^7 células C1R fueron transfectadas con 10 μ g de DNA por electroporación a 250 mV y 960 μ F. Para seleccionar transfectantes estables, se cultivaron las células en presencia de 1 mg/ml de geneticina (Invitrogen). La expresión en superficie de HLA-B*40:02 se confirmó mediante citometría de flujo. Las células fueron cultivadas en DMEM suplementado con un 7,5% de FBS (ambos de Sigma) a 37°C en botellas de vidrio, con la ayuda de un agitador de rodillos, hasta una densidad aproximada de $1,5 \cdot 10^6$ células/ml. A continuación, fueron lavadas 2 veces con PBS, centrifugadas y guardadas a -80°C para su posterior uso.

El anticuerpo monoclonal W6/32 (IgG2a específico para un determinante monomórfico de moléculas HLA de clase I) ha sido descrito anteriormente [140].

III.2-Aislamiento del peptidoma unido a HLA-B*40:02 y -B*39:01

Los ligandos unidos a HLA-B*40, y -B*39 fueron aislados siguiendo un protocolo descrito anteriormente [141] con pequeñas variaciones. Se lisaron unas 10^{10} células C1R-B*40 o C1R-B*39 durante 1h a 4°C en Tris 20mM, NaCl 150mM, Igepal 1% CA-630 (Sigma-Aldrich), NaN_3 0,02% a pH 7,5 en presencia de una mezcla de inhibidores de proteasas (Complete-Midi, Roche) y de fosfatasas¹ (PhosSTOP, Roche). El lisado se sometió a centrifugación diferencial durante 10 min a 2000-g, 30 min a 10.000-g y 1h a 100.000-g. Tras la ultracentrifugación, la fracción soluble se aplicó sobre una precolumna de sefarosa bloqueada con TRIS. El eluido de la precolumna se hizo pasar dos veces sobre una columna de sefarosa (CNBr-activated Sepharose 4B, GE Healthcare) acoplada al anticuerpo monoclonal W6/32. Se realizaron tres lavados exhaustivos a la columna con 1) NET con un 10% de una solución saturada de NaCl e Igepal CA-630 0,5%; 2) NET con un 5% de una solución saturada de NaCl e Igepal CA-630 0,5% y

¹ La muestra obtenida a partir de la línea C1R-B*40 para su análisis por espectrometría de masas ESI-Q-TOF fue aislada en ausencia de inhibidores de fosfatasas.

3) NET. El tampón NET, en todos los casos, fue TRIS 50 mM, NaCl 150 mM, EDTA 5 mM, pH 7,4. La elución de la columna se llevó a cabo a temperatura ambiente con ácido trifluoroacético (TFA) al 0,1%. Posteriormente las fracciones con mayor absorbancia a 280 nm se filtraron a través de un Vivacon 2 (Sartorius Stedim) para separar los péptidos de la cadena pesada y la β_2m . La mezcla peptídica resultante fue concentrada en un *speedvac*, hasta un volumen aproximado de 100 μ l y fue posteriormente desalada con puntas OMIX C18 (Varian). Finalmente, la muestra se secó completamente y se guardó a -20°C para su posterior análisis.

III.3-Lisis celular y digestión triptica

Para el análisis del proteoma y fosfoproteoma de la línea C1R-B*40 se lisaron 10^6 células en urea 8M, bicarbonato amónico 50 mM, pH 8, en presencia de inhibidores de proteasas (Complete mini EDTA-free, Roche) e inhibidores de fosfatasa (PhosSTOP, Roche). El lisado celular se sonicó 3 veces en hielo y se centrifugó a 20.000-g durante 20 min a 4°C. La concentración de proteína en la fracción soluble se cuantificó mediante un ensayo de Bradford. Las proteínas fueron reducidas con DTT 2 mM a 56°C durante 25 min y alquiladas con iodoacetamida 4 mM a temperatura ambiente y en oscuridad. La digestión se llevó a cabo con Lys-C durante 4 horas a 37°C usando una relación enzima-sustrato de 1:75. A continuación, la muestra fue diluida 4 veces con bicarbonato amónico 25 mM. Para completar la digestión, se usó tripsina en una relación enzima-sustrato de 1:100 durante toda la noche a 37°C. La reacción se detuvo añadiendo ácido fórmico hasta una concentración final del 1%. Se hicieron alícuotas de 200 μ g que fueron desaladas utilizando cartuchos Sep-Pack C18 (Waters), secadas mediante *speedvac* y guardadas a -20°C. Este procedimiento se llevó a cabo en colaboración con el laboratorio del Dr. Albert Heck (Netherlands Proteomics Centre, Utrecht).

III.4-Enriquecimiento de fosfopéptidos

El enriquecimiento de los fosfopéptidos asociados a HLA-B*40 y HLA-B*39 se llevó a cabo mediante la aplicación en tándem de dos cromatografías de afinidad de tipo IMAC (*Immobilized-Metal Affinity Chromatography*) y MOAC (*Metal Oxide Affinity Chromatography*) utilizando como ligandos Fe^{3+} y TiO_2 , respectivamente, siguiendo protocolos previamente descritos [142, 143]. La mezcla de péptidos fue reconstituida en 200 μ l de solución de carga (acetonitrilo 50%, TFA 0,3%, pH 1,5) y se incubó a temperatura ambiente con 15 μ l de matriz Phos-Select (Sigma). Posteriormente se empaquetó dicha matriz en una microcolumna que se conectó en tándem con otra empaquetada con matriz *Oligo R3* (Applied Biosystems). La mezcla de péptidos se hizo pasar con ayuda de una jeringa y se colectó la

fracción no retenida para aplicarla posteriormente sobre la columna de TiO₂. Ambas columnas se sometieron a lavados exhaustivos; primero con solución de carga, seguido de solución de transferencia (ácido fosfórico 1%), y finalmente con solución de lavado (acetonitrilo 2%, TFA 0,1%). La elución se llevó a cabo empleando dos soluciones a distinto pH: 1) acetonitrilo 50%, TFA 0,1% y 2) acetonitrilo 30%, hidróxido amónico 0,06 mM. Ambos eluidos se mezclaron, se secaron, se resuspendieron en ácido fórmico al 0,1% y se almacenaron a -20°C hasta su posterior análisis (nos referiremos a esta muestra como fracción IMAC a lo largo de esta tesis).

El *flow-trough* de la cromatografía previamente descrita se llevó a sequedad y se resuspendió en buffer de carga para la cromatografía basada en TiO₂ (ácido glicólico 1 M, acetonitrilo 80%, TFA 5%). Esta solución se aplicó sobre una microcolumna empaquetada con resina Titansphere (GL Sciences). Tras hacer pasar la muestra por la columna de afinidad, la fracción no retenida se colectó y acidificó añadiendo 10 µl de ácido fórmico 0,1%. A continuación, se secó y resuspendió en TFA 0,1% para su posterior desalado con ayuda de una ZipTip C18 (Eppendorf). La muestra desalada se secó y resuspendió en ácido fórmico 0,1% y se guardó a -20°C hasta su posterior análisis mediante LC-MS/MS. (nos referiremos a esta muestra como fracción FT a lo largo de esta tesis). Por su parte la columna de TiO₂ fue lavada con acetonitrilo 80%, TFA 5% y los péptidos se eluyeron con dos soluciones: 1) hidróxido amónico 0,3 M y 2) hidróxido amónico 0,3 M, acetonitrilo 30%. Ambos eluidos fueron mezclados, acidificados mediante la adición de 10 µl de ácido fórmico al 100% y llevados a sequedad. Posteriormente, la muestra fue desalada utilizando una ZipTip C18 (Eppendorf) y secada. Se resuspendió en ácido fórmico 0,1% y almacenó a -20°C hasta su posterior análisis (nos referiremos a esta muestra como fracción TiO₂ a lo largo de esta tesis).

El enriquecimiento de fosfopéptidos provenientes de la digestión trípica del proteoma total de la línea C1R-B*40 se llevó a cabo mediante un protocolo ya descrito [129] en colaboración con el grupo del Dr. Albert Heck (Netherlands Proteomics Centre, Utrecht). El digerido trípico se resuspendió en acetonitrilo 80%, TFA 6% y se aplicó sobre una microcolumna empaquetada con microesferas de Ti⁴⁺-IMAC. El *flow-trough* de esta cromatografía se secó en un *speedvac*, se resuspendió en ácido fórmico 0,1% y se almacenó a -20°C hasta su análisis. Se realizó un primer lavado con acetonitrilo 50%, TFA 0,5%, NaCl 200 mM, seguido de un segundo lavado con acetonitrilo 50%, TFA 0,1%. Los péptidos retenidos se eluyeron con dos soluciones: 1) NH₃ 10% y 2) acetonitrilo 80%, ácido fórmico 2%. Los dos eluidos se combinaron, se secaron en un *speedvac* y se resuspendieron en ácido fórmico 0,1% para su posterior análisis mediante LC-MS/MS.

III.5-Análisis mediante espectrometría de masas en tándem acoplada a HPLC (LC-MS/MS)

Todos los análisis mediante espectrometría de masas realizados en este documento se resumen en la Tabla 2. Los experimentos que implicaron el uso de espectrómetros de masas equipados con analizadores Orbitrap, fueron llevados a cabo en colaboración con el laboratorio del Dr. Albert Heck (Netherlands Proteomics Centre, Utrecht).

III.5.1-Análisis del peptidoma y fosopeptidoma de HLA-B*40 y HLA-B*39 mediante espectrometría de masas ESI-Q-TOF

Se analizaron dos réplicas técnicas de las fracciones FT, IMAC y TiO₂ obtenidas tras el enriquecimiento de fosfopéptidos descrito anteriormente. Las muestras fueron analizadas utilizando un HPLC nano-LC Ultra (Eksigent) acoplado en línea con un espectrómetro de masas 5600 triple TOF (AB Sciex) mediante una fuente de ionización Nanospray III (AB Sciex). La fuente de ionización se equipó con una punta de sílice fundida PicoTip (10 µm X 12 cm, New Objective). Se empleó una columna de atrapamiento de fase reversa C18 ChromXP (350 µm x 0.5 mm, diámetro de partícula: 3 µm, de tamaño de poro: 120 Å; Eksigent) y una columna C18 ChromXP (75 µm x 150 mm, diámetro de partícula: 3 µm, tamaño de poro: 120 Å; Eksigent). Se utilizó ácido fórmico 0,1% como solvente A y ácido fórmico 0,1% en acetonitrilo como solvente B. La carga en la columna de atrapamiento se llevó a cabo en condiciones isocráticas de solvente A con flujo de 2 µl/min durante 10 min. Para el análisis del peptidoma unido a HLA-B*40:02, la cromatografía se desarrolló a un flujo de 300 nl/min a 40°C con un tiempo total de análisis de 150 min bajo condiciones de elución en gradiente, como sigue: 2% de solución B durante 1 min, aumento lineal a 30% de B en 109 min, aumento lineal a 40% de B en 10 min, aumento lineal a 90% de B en 5 min y 90% de B durante 5 min. Por su parte el 50% del peptidoma asociado a HLA-B*39 se analizó como se ha descrito para el caso de HLA-B*40. Para el restante 50%, se empleó un tiempo total de análisis de 250 min y el siguiente gradiente: 2% de solución B durante 1 min, aumento lineal a 30% de B en 181 min, aumento lineal a 40% de B en 24 min, aumento lineal a 90% de B en 15 min y 90% de B durante 10 min. El HPLC fue controlado mediante el programa Eksigent Control (v3.12, Eksigent).

La adquisición se llevó a cabo en modo de ionización positiva y con un voltaje en la fuente de iones (ISVF) de 2800 V, presión de gas cortina (CUR) de 20 psi, temperatura de la interfaz (IHT) de 150°C, presión de gas 1 (GS1) de 20 psi y *declustering potential* (DP) de 85 V. En los análisis de 150 min, los ciclos de adquisición consistieron en un barrido MS (350-1250 m/z) de 250 ms seguido por un máximo

de 50 espectros MS₂ (100-1500 m/z) con un tiempo de acumulación de 50 ms. El ciclo de adquisición en los análisis de 250 min de duración consistió en un espectro MS (350-1250 m/z) de 250 ms seguido de un máximo de 25 espectros MS₂ (100-1500 m/z) con un tiempo de acumulación de 100 ms. Los iones más intensos de acuerdo al espectro de masas fueron seleccionados para su posterior fragmentación. Se excluyeron los iones monocargados para evitar la fragmentación de contaminantes no peptídicos. A cada ion fragmentado se le aplicó una ventana de exclusión de 20 s. El tiempo total de cada ciclo fue de 2,8 s. El espectrómetro de masas fue controlado mediante el programa Analyst TF (v1.5, AB Sciex).

III.5.2-Análisis del peptidoma y fosopeptidoma de HLA-B*40 mediante espectrometría de masas ESI-LTQ-Orbitrap

Tres réplicas biológicas de las fracciones FT, IMAC y TiO₂ obtenidas tras el enriquecimiento de fosfopéptidos a partir del peptidoma asociado a HLA-B*40 fueron analizadas mediante nanoLC-MS/MS. Para ello, se usó un HPLC EASY-nLC 1000 (Thermo Fisher Scientific) acoplado en línea a un espectrómetro de masas LTQ Orbitrap Elite equipado con un módulo ETD (Thermo Fisher Scientific). Se utilizó una columna de atrapamiento casera de 100 µm x 20 mm con matriz Reprosil C18 (Dr. Maisch, diámetro de partícula: 3 µm) seguida de una columna analítica de 50 µm x 500 mm empaquetada con matriz Poroshell C18 (Agilent, diámetro de partícula: 2.7 µm). Se usó ácido fórmico 0,1% como solvente A y ácido fórmico 0,1% en acetonitrilo como solvente B. Las fracciones IMAC y TiO₂ fueron analizadas en un gradiente de 2 a 27% de solución B en 91 min y un tiempo total de análisis de 120 min. Luego del gradiente la columna fue lavada durante 3 min incrementando al 100% de B y finalmente fue llevada a condiciones iniciales de 98% solución A durante 25 min. Para el análisis de las fracciones FT se utilizó un gradiente de 2 a 27% de solución B en 160 min con un tiempo total de análisis de 180 min.

Los espectros MS se adquirieron mediante el analizador Orbitrap barriendo el rango entre 300 y 1200 unidades de m/z a una resolución de 60.000 (FWHM). Los 10 iones con mayor intensidad en el espectro de masas fueron seleccionados para su posterior fragmentación por EThcD. Fueron excluidos los iones monocargados para evitar la fragmentación de contaminantes no peptídicos. Los tiempos máximos de acumulación fueron de 200 ms y de 1500 ms para los barridos MS y MS/MS, respectivamente. Los iones fragmento fueron detectados en el Orbitrap a una resolución de 15.000 (FWHM). Se aplicó una ventana de exclusión de 60 s sobre cada ion fragmentado. El espectrómetro de masas fue controlado mediante el software Xcalibur Qual Browser v3.0.63 (Thermo Fisher Scientific).

III.5.3-Análisis del proteoma y fosfoproteoma de la línea celular C1R-B*40 mediante espectrometría de masas ESI-LTQ-Orbitrap

El análisis de los péptidos tripticos derivados del proteoma y fosfoproteoma de la línea C1R-B*40 se llevó a cabo utilizando un HPLC EASY-nLC 1000 (Thermo Fisher Scientific) acoplado en línea a un espectrómetro de masas Orbitrap Fusion (Thermo Fisher Scientific) equipado con un módulo ETD. Se utilizó ácido fórmico 0,1% como solvente A y ácido fórmico 0,1% en acetonitrilo como solvente B. Se empleó un gradiente no lineal para lograr la máxima separación de los péptidos en ambas muestras, comenzando por una variación de 2 a 3% de solución B en 8 min, de 3 a 8% de B en 50 min, de 8 a 12% de B en 20 min, de 12 a 15% de B en 10 min, de 15 a 19% de B en 10 min, de 19 a 24% de B en 10 min y de 24 a 30% de B en 10 min. El tiempo total de análisis fue de 140 min.

La adquisición de espectros MS se realizó en el analizador Orbitrap, barriendo el rango entre 350 y 1400 unidades de m/z a una resolución de 60.000 (FWHM). Se utilizó un método de máxima velocidad, donde los iones precursores más abundantes fueron seleccionados en ciclos de trabajo de 5 segundos para las muestras enriquecidas en fosfopéptidos y de 3 segundos para las no enriquecidas. Los precursores fueron aislados en el cuadrupolo con una ventana de 1,6 Da. Para las muestras fosforiladas, se utilizaron métodos diferentes de análisis: uno mediante ETHcD suplementado con una energía de colisión HCD extra del 40% y adquisición de los espectros MS2 en el analizador Orbitrap a una resolución de 15.000 (FWHM) y otro alternando entre fragmentación HCD y ETcaD y adquiriendo los espectros MS2 en la trampa de iones en modo de barrido rápido mediante un procedimiento similar al descrito por Frese y cols [134]. El tiempo de reacción del ETD fue calibrado basándose en la carga del precursor para ambas fragmentaciones (ETHcD y ETcaD). El tiempo máximo de acumulación de iones fue de 50 ms y 25 ms para los barridos MS y MS2, respectivamente. El control de ganancia automático (AGC) fue fijado en 400.000 para el barrido MS y en 10.000 para el MS2. El tiempo máximo de inyección fue de 100 ms en el método ETHcD y de 70 ms en el HCD/ETcaD.

Para analizar la muestra no enriquecida en fosfopéptidos, se mantuvieron los mismos parámetros anteriormente descritos excepto que la fragmentación se llevó a cabo exclusivamente mediante HCD con un 35% de energía de colisión y la lectura MS2 se efectuó en la trampa de iones. En este caso, el AGC fue fijado en 30.000 para el barrido MS2, mientras que el tiempo de inyección fue de 35 ms como máximo. El espectrómetro de masas fue controlado mediante el software Xcalibur Qual Browser v3.0.63 (Thermo Fisher Scientific).

| Línea celular | Tipo de análisis | Espectrómetro de masas | Fragmentación | Muestras | Réplicas biológicas |
|-----------------|----------------------------|------------------------------------|-------------------|---------------------------------|---------------------|
| C1R-B*40 | Peptidoma y fosfopeptidoma | ESI-Q-TOF (TTOF 5600) | CID | 3 (IMAC, TiO ₂ , FT) | 1 |
| C1R-B*39 | Peptidoma y fosfopeptidoma | ESI-Q-TOF (TTOF 5600) | CID | 3 (IMAC, TiO ₂ , FT) | 1 |
| C1R-B*40 | Peptidoma y fosfopeptidoma | ESI-LTQ-Orbitrap (Orbitrap Velos) | ETHcD | 3 (IMAC, TiO ₂ , FT) | 3 |
| C1R-B*40 | Proteoma y fosfoproteoma | ESI-LTQ-Orbitrap (Orbitrap Fusion) | ETHcD, HCD, ETcaD | 2 (Ti ⁴⁺ -IMAC, FT) | 1 |

Tabla 2. Resumen de las muestras analizadas por LC-MS/MS.

III.5.4-Análisis de péptidos sintéticos por espectrometría de masas

Los péptidos sintéticos fueron resuspendidos en una mezcla de ácido fórmico 0,5% y acetonitrilo 20% a una concentración estimada de 250 fmol/ μ l antes de ser infundidos directamente en el espectrómetro de masas 5600 triple TOF a un flujo de 1-3 μ l/min para la obtención de los espectros MS2. Adicionalmente, para determinar los tiempos de retención de los fosfopéptidos unidos a HLA-B*40, 100 fmol se analizaron de cada uno de ellos por LC-MS/MS en las mismas condiciones que el correspondiente peptidoma endógeno.

Los péptidos sintéticos analizados mediante ETHcD en el espectrómetro de masas LTQ Orbitrap Elite fueron resuspendidos en ácido fórmico 10% a una concentración estimada de 10 ng/ μ l. Aproximadamente, 50 ng de péptido fueron analizados mediante nanoLC-MS/MS en un gradiente cromatográfico de 45 min de duración.

III.6-Búsquedas MS/MS e identificación de péptidos

Tanto para el peptidoma como el fosfopeptidoma presentado por HLA-B*40 y HLA-B*39, los datos crudos generados por los espectrómetros de masas fueron procesados con el software Peakview (v1.1, AB Sciex) o Proteome Discoverer (Thermo Fisher Scientific) para generar un archivo con extensión *.mgf* que se utilizó como entrada para las búsquedas de iones MS2. La base de datos que se empleó en las búsquedas, contenía 88.266 entradas y se creó concatenando la sección *Homo sapiens* de Uniprot (a fecha de noviembre de 2013) con una base de datos de entradas señuelo compuesta por secuencias invertidas generadas a partir de la anterior con la herramienta DBtoolkit (v4.1.4).

Las búsquedas se llevaron a cabo mediante la combinación de varios motores y con distintos parámetros de búsqueda. Los motores utilizados fueron: MASCOT (Matrix Science, v2.4.0), OMSSA (National Center for Biotechnology Information, v2.1.9), X!Tandem (The Global Proteome Machine Organization, v2013-02-01-1) y X!Tandem con k-score (LabKey Software, v2.3-7806). Se consideraron las

siguientes modificaciones variables: oxidación de la metionina, formación de ácido piroglutámico a partir de glutamina o ácido glutámico N-terminales, acetilación en el extremo amino de la proteína y fosforilación de serina, treonina, y tirosina. Los *scores* de los distintos motores de búsqueda se convirtieron a una escala de probabilidad común tal y como se ha descrito previamente [144] y solo se consideraron aquellas secuencias identificadas con una $FDR \leq 1\%$ a nivel de péptido.

Para la identificación de fosfopéptidos mediante CID, debido a que esta técnica no es la idónea para caracterizar especies fosforiladas, se aplicaron criterios más estrictos. La estrategia utilizada se basó en no considerar positiva una identificación hasta que la secuencia implicada no hubiera sido confirmada por comparación del espectro de fragmentación experimental con el del correspondiente péptido sintético. Para evitar desechar secuencias de verdaderos ligandos fosforilados que no superaran el umbral de $FDR \leq 1\%$, se consideraron todas las secuencias de fosfopéptidos con los motivos de unión a B*40 o B*39 asignadas a espectros MS2 con un score de MASCOT ≥ 20 (B*40) o una $FDR \leq 65\%$ a nivel de péptido (B*39). Todos estos espectros fueron inspeccionados manualmente en busca de iones diagnóstico de péptidos fosforilados, fundamentalmente pérdidas neutras del grupo fosfato (-98, -49 y -32,7 unidades de m/z para iones simple, doble y triplemente cargados, respectivamente). Aquellas asignaciones que parecían correctas se confirmaron o rechazaron mediante fragmentación del péptido sintético correspondiente.

Para las muestras obtenidas mediante digestión triptica las búsquedas se realizaron combinando los siguientes motores de búsqueda: MASCOT v2.4.0, OMSSA v2.1.9, X!Tandem v2013-02-01-1 y Myrimatch (Mass Spectrometry research Center, Vanderbilt University, v2.2.140) utilizando tripsina como enzima de corte permitiendo hasta dos *misscleavages* y considerando la carbamidometilación de cisteínas como modificación fija. Se consideraron como modificaciones variables: oxidación de metionina, formación de ácido piroglutámico a partir de glutamina o ácido glutámico N-terminales, acetilación en el extremo amino de la proteína y fosforilación de serina, treonina, y tirosina. La tolerancia en el MS fue de 10 ppm y en MS/MS de 0.05 Da (ET_hcD) o 0.6 Da (HCD y ET_caD). La combinación de resultados y la estimación de la FDR se realizó igual que en caso de los peptidomas de B*40 y B*39. Solo se consideraron aquellas secuencias identificadas con una $FDR \leq 1\%$ a nivel de péptido.

III.7-Análisis estadístico

Para determinar el motivo de unión de HLA-B*40:02, los péptidos identificados fueron agrupados de acuerdo a su longitud. La frecuencia de cada residuo en cada posición del péptido (f_{obs}) se comparó

con la frecuencia del mismo aminoácido en la base de datos (f_{exp}), bajo la hipótesis nula de que $f_{obs} \leq f_{exp}$. Los p-valores preliminares se calcularon asumiendo una distribución binomial con $p = f_{exp}$. Los P-valores definitivos se obtuvieron aplicando la corrección de Bonferroni según la siguiente ecuación:

$$P = 1 - (1 - p)^{20k}$$

donde k es el número de residuos de cada grupo de péptidos (por ejemplo, $k = 9$ para los nonámeros). Se consideró que una diferencia era estadísticamente significativa cuando $p < 0,05$. Los residuos que mostraban una relación $f_{obs}/f_{exp} > 5$ fueron considerados motivos fuertes de unión. Los residuos que mostraban un incremento estadísticamente significativo pero con $f_{obs}/f_{exp} < 5$ veces fueron considerados motivos débiles.

III.8-Síntesis de péptidos

La síntesis de péptidos y fosfopéptidos en fase sólida se realizó en un sintetizador automatizado MultiPep (Intavis) utilizando química estándar Fmoc. Los péptidos fueron purificados mediante cromatografía de fase reversa en un HPLC *Smartline* (Knauer) equipado con una columna C18 218TP52 (Vydac) o utilizando una microcolumna empaquetada con matriz Oligo R3 (Applied Biosystems). Los candidatos destinados a ensayos de unión, se cuantificaron mediante análisis de aminoácidos en un analizador Biochrom 30 (Biochrom). El péptido GEFGGCGSV fue marcado después de la síntesis con 5-iodoacetamidafluoresceína (Thermo) siguiendo las instrucciones del fabricante. Seguidamente fue purificado mediante HPLC de fase reversa y cuantificado por absorbancia a 491 nm como se ha descrito previamente [145].

III.9-Ensayos de unión de péptido a HLA-B*40

Los ensayos de unión se realizaron con pequeñas variaciones sobre un protocolo ya descrito [145, 146]. En resumen, las células C1R-B*40 se lavaron dos veces con PBS e incubaron durante 5 min a 4°C. A continuación, se añadieron 500 μ l de ácido cítrico 0,263 M, Na_2HPO_4 0,123 M, BSA 1%, pH 3 durante 90 s. Después de dos lavados en DMEM con 2% de FBS a 4°C, las células fueron incubadas en DMEM suplementado con FBS al 5%, β_2m 2 μ g/ml (Calbiochem), el péptido fluorescente GEFGGXGSV 400 mM (donde X representa cisteína marcada con fluoresceína) y diferentes concentraciones de los péptidos candidatos entre 50 μ M y 23 nM en diluciones seriadas de 1/3. Tras incubar durante toda la noche a 4°C, se midió la fluorescencia en un citómetro de flujo Cytomics FC500 (Beckman Coulter). La inhibición de la unión del péptido fluorescente se representó frente a la concentración del péptido candidato, y la IC_{50}

(concentración del péptido candidato que da una inhibición del 50%) fue estimada ajustando los valores a una curva sigmoide tal y como ha sido descrito por Kessler y cols [145].

III.10-Cristalografía

Los experimentos de cristalografía fueron llevados a cabo en colaboración con el Dr. Cesar Santiago (Centro Nacional de Biotecnología, Madrid).

III.10.1-Clonaje, expresión y purificación de proteínas recombinantes

El DNA codificante de los residuos 1-274 de HLA-B*40:02 se amplificó por PCR a partir del cDNA de longitud completa, se clonó en el vector pET-22b (Novagen) utilizando el sitio NdeI y se usó para transformar la cepa Rosetta-GAMI (DE3) pLysS de *E. coli*. Por su parte, la β_2m se expresó a partir de un cDNA clonado en el vector pET-30a (Novagen) proporcionado por el profesor James McCluskey (Universidad de Melbourne, Parkville, Australia). La expresión de esta proteína se llevó a cabo en la cepa BL21 (DE3) pLysS de *E. coli*. Las bacterias fueron cultivadas en medio LB suplementado con ampicilina 100 mg/ml en el caso de la cadena pesada o kanamicina 50 mg/ml para la β_2m .

La expresión de ambas proteínas fue inducida mediante la adición de IPTG 0,5 mM, cuando los cultivos alcanzaron una densidad óptica a 600nm igual a 0,6. Se continuó con la incubación a 37°C y en agitación durante 4 horas (β_2m) y 5 horas (cadena pesada). Las bacterias fueron lavadas con Tris 20 mM, NaCl 150 mM, pH 7,5 y centrifugadas a 4000 rpm durante 10 min a 4°C. El sedimento fue guardado a -80°C hasta su posterior uso. La expresión de ambas proteínas se confirmó mediante SDS-PAGE (Figura 7).

La purificación y solubilización de los cuerpos de inclusión se llevó a cabo siguiendo un protocolo ya publicado [147] con ligeras modificaciones. Las bacterias se resuspendieron en TRIS 50 mM, NaCl 0,2 M, EDTA 5 mM, DTT 5 mM, pH 8,0 y se añadió lisozima a una concentración final de 1 mg/ml. Tras una incubación de 15 min en hielo, se añadió Triton X-100 y ácido deoxicólico a una concentración final de 0,5%. La lisis se llevó a cabo mediante 2 ciclos de congelación y descongelación tras los cuales se sonicó el lisado dos veces en ciclos de 1 min. El lisado celular se centrifugó y el precipitado insoluble se lavó primero con TRIS 50 mM, NaCl 0,1 M, EDTA 1 mM, DTT 1 mM, Triton X-100 0,5%, pH 8,0; luego con TRIS 50 mM, NaCl 2 M, urea 1 M, EDTA 1 mM, DTT 1 mM, pH 8,0 y, finalmente, con TRIS 50 mM, EDTA 1 mM, DTT 1 mM, pH 8,0. Se realizaron cuatro lavados con cada tampón y –después de cada centrifugación– el sedimento se resuspendió y se sonicó dos veces en ciclos de 1 min. Después de la última etapa de

centrifugación, las proteínas insolubles se resuspendieron en MES 25 mM, urea 8 M, EDTA 10 mM, DTT 1 mM, pH 6,5. Ambas proteínas se cuantificaron midiendo su absorbancia a 280 nm.

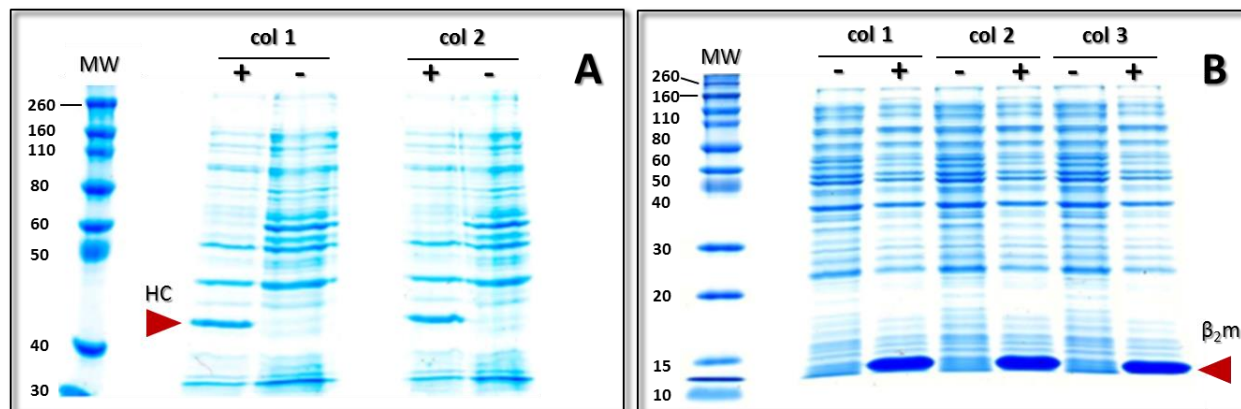


Figura 7. Confirmación mediante geles SDS-PAGE de la expresión en bacterias de la cadena pesada de HLA-B*40 y la β_2m . (A) Se señala la banda correspondiente a la cadena pesada (HC) en dos colonias independientes. (B) Se indica la banda correspondiente a la β_2m en tres cultivos independientes. Los signos + y - indican la presencia o ausencia de IPTG 0.5 mM en el medio de cultivo.

III.10.2-Replegamiento *in vitro* y cristalización

El replegamiento de HLA-B*40 se llevó a cabo de manera similar a la descrita por Reid y cols [148] en 1 l de Tris-HCl 0,1 M, L-arginina 400 mM, EDTA 2 mM, Cisteamina 6,3 mM, Cistamina 3,7 mM, PMSF 0,1 mM, pH 8,3. Se emplearon 15 mg de cadena pesada, 20 mg de β_2m y 10 mg de péptido (relación molar 1:4:20). En un primer paso se añadieron gota a gota 5 mg de β_2m a la solución de replegamiento y se incubó a 4°C con agitación moderada durante 8 h. El resto de la β_2m se combinó con la cadena pesada y el péptido y, siguiendo el mismo procedimiento, se incorporó a la solución de replegamiento en cuatro pulsos en intervalos de 12 h.

Tras la última incubación, la mezcla fue concentrada mediante ultrafiltración a través de una membrana de 10 kDa omega 62 mm (Pall) hasta un volumen aproximado de 50 ml. A continuación, la mezcla se concentró mediante filtros VIVASPIN TURBO 15 (10.000 MWCO, Sartorius) hasta unos 3 ml. Finalmente, el complejo trimérico replegado se purificó mediante cromatografía de exclusión molecular. Para ello se aplicó el concentrado proteico sobre una columna HiLoad 16/60 Superdex 200 prep grade empacada con sephadex G75 o G200 (GE Healthcare) pre-equilibrada con TRIS 25 mM, NaCl 100 mM, pH 8 a 10°C y a un flujo de 0,5 ml/min. El eluido de la columna se colectó en fracciones de 2,5 ml que posteriormente fueron analizadas mediante SDS-PAGE para determinar su pureza y composición

(Figura 8A). La presencia del péptido se corroboró mediante análisis por espectrometría de masas MALDI-TOF (Figura 8B). Las fracciones que contenían el complejo trimérico se concentraron usando un VIVASPIN TURBO 4 (5000 MWCO, Sartorius) hasta una concentración final de 15-20 mg/ml. Esta mezcla se conservó a 4°C hasta su utilización en ensayos de cristalización.

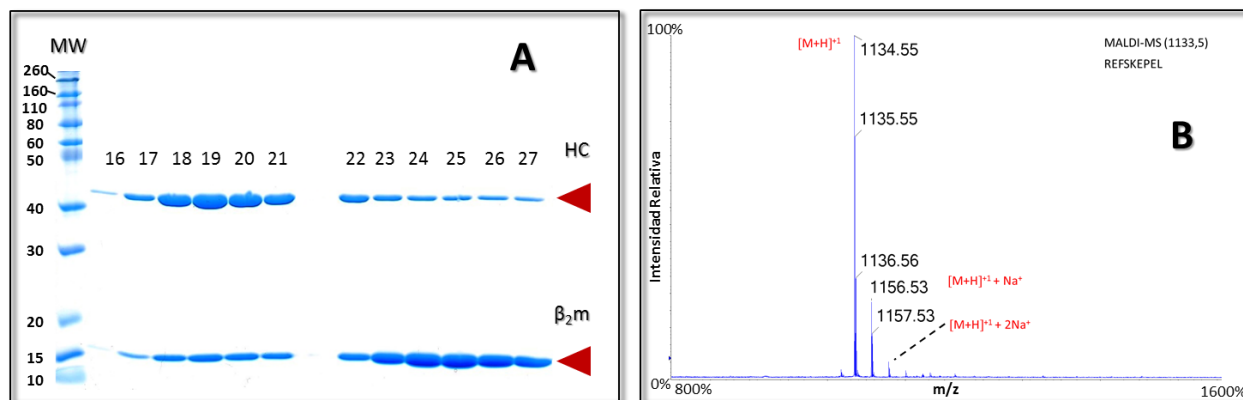


Figura 8. Caracterización molecular de los complejos triméricos replegados. (A) Fracciones colectadas (16-27) mediante cromatografía de exclusión molecular del complejo trimérico y verificadas mediante SDS-PAGE. La cadena pesada (HC) y la β_2m muestran pesos moleculares aparentes en torno a 44 y 12 kDa, respectivamente. En este caso, se utilizaron las fracciones 18-20 para los experimentos de cristalización. (B) Espectro de masas para corroborar la presencia del ligando peptídico, en este caso para el complejo B*40-REFSKEPEL.

La cristalización del complejo se llevó a cabo por el método de difusión de vapor con gota sentada. Para ello se mezclaron gotas de 150 nl del complejo trimérico con 150 nl de las distintas soluciones de cristalización disponibles en el kit Pi-Minimal (Jena Bioscience) utilizando un robot TECAN (Genesis). Se obtuvieron cristales para todos los complejos HLA-B*40-péptido al cabo de tres días (Figura 9) en las siguientes condiciones: condición A (G(pS)FSRFYSL) PEG 8000 26%, HEPES 150 mM pH 7,5, tartrato potásico 60 mM; condición B (GEFSRFYSL) PEG 20 000 18.6%, citrato tri-sodio 150 mM pH 5,5, tartrato de sodio 20 mM; condición C (REF(pS)KEPEL) PEG MME 5000 20%, TRIS 150 mM pH 8, etilenglicol 8% y condición D (REFSKEPEL) PEG 8000 30%, TRIS 150 mM pH 8,5, isotiocianato 70 mM. Posteriormente se recogieron los cristales con la ayuda de un capilar de vidrio y se trasladaron a una placa con tampón de cristalización correspondiente y crioprotector -20% glicerol para los complejos A, B y D; y un 20% etilenglicol para el complejo C- durante unos segundos. Se recogieron cristales individuales con la ayuda de lazos y rápidamente fueron congelados en nitrógeno líquido para su transporte y colección de datos.

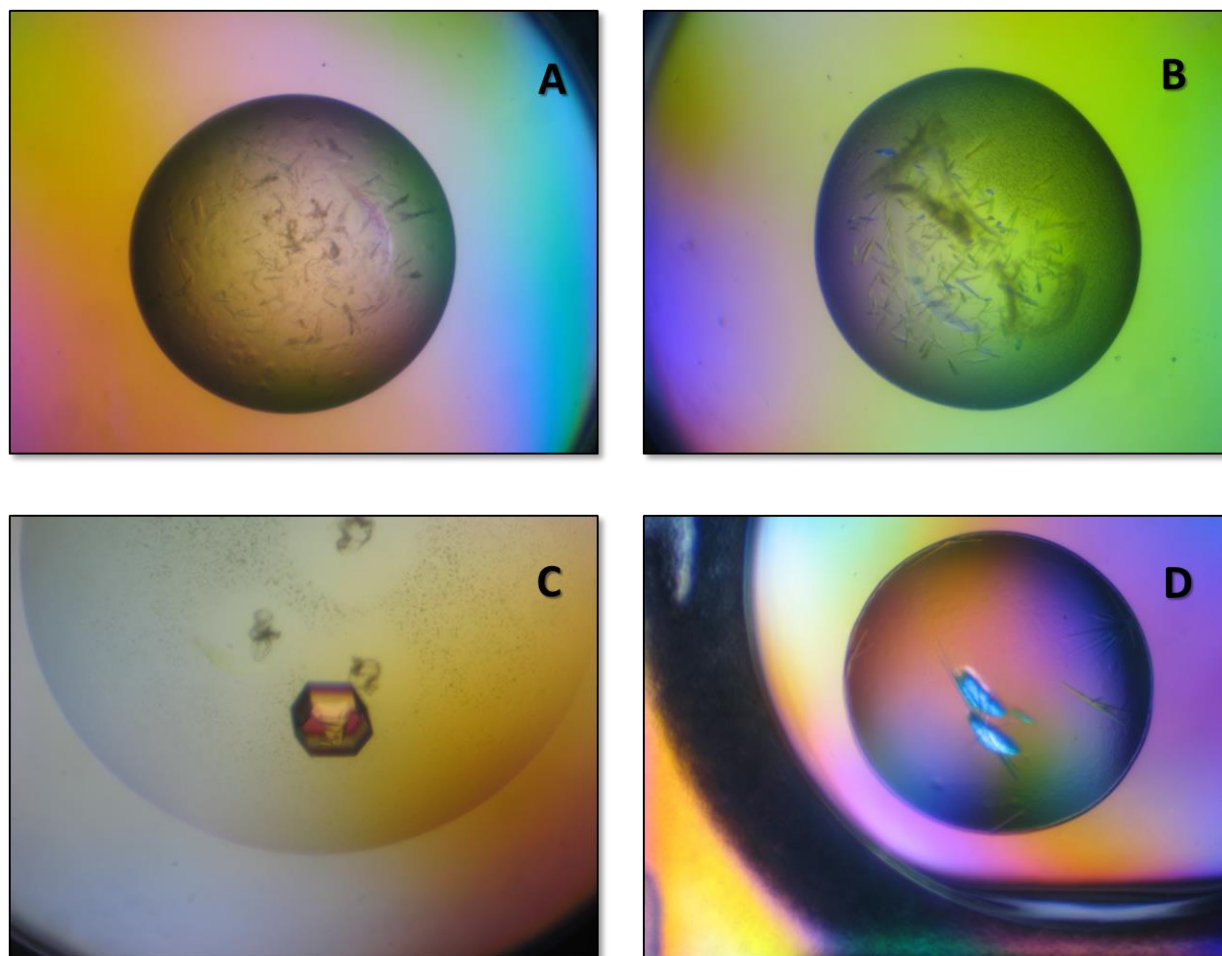


Figura 9. Cristales de los cuatro complejos HLA-B*40-péptido obtenidos mediante el método de difusión de vapor con gota sentada. (A) B*40-G(pS)FSRFYSL, (B) B*40-GEFSRFYSL, (C) B*40-REF(pS)KEPEL y (D) B*40-REFSKEPEL.

III.10.3-Difracción de rayos-X y resolución de estructuras

Los experimentos de difracción se llevaron a cabo en la línea BL13-XALOC [149] del sincrotrón ALBA (Cerdanyola del Vallès, Barcelona), utilizando una longitud de onda de 1 Å con tiempos de exposición de 0,25 s e incrementos en el ángulo de rotación de 0,25°. Las imágenes obtenidas fueron integradas con XDS [150], posteriormente escaladas con AIMLESS (CCP4 suite) [151] y los datos resultantes introducidos en PHASER (CCP4 suite) para la resolución de la estructura mediante reemplazo molecular. Para ello se usó como molde la estructura de HLA-B*41:04 (PDB: 3LN5) que comparte un 97,1% de identidad con HLA-B*40:02 en los dominios α_1 , α_2 y α_3 . Una vez obtenida una solución fiable se cambió el registro de aminoácidos manualmente (COOT) [152] por el de HLA-B*40:02 y se procedió a su refinamiento mediante Phenix v1.10-2148 [153].

III.11- Predicción *in silico* de ligandos de HLA-B*40

Los péptidos trípticos fosforilados identificados a partir del proteoma celular de C1R-B*40 fueron asignados a las secuencias de sus proteínas parentales (extraídas de Uniprot) y se consideraron todos los posibles nonámeros que contenían fosfoserina (pSer) o fosfotreonina (pThr). A partir de este listado, se seleccionaron todos los péptidos que se ajustaban al patrón X-E-X-X-X-X-X-[FLIVM] y se registró la posición del residuo fosforilado. Se aplicó el mismo procedimiento para las secuencia de 10 y 11 aminoácidos (patrones X-E-X-X-X-X-X-X-[FLIVM] y X-E-X-X-X-X-X-X-X-[FLIVM]) respectivamente. Estos análisis fueron llevados a cabo en colaboración con el Dr. Florencio Pazos (Centro Nacional de Biotecnología, Madrid).

Resultados

IV-Resultados

IV.1-Análisis del peptidoma presentado por HLA-B*40 y bases estructurales de la presentación de péptidos por este alotipo

IV.1.1-Distribución por tamaño del peptidoma asociado a HLA-B*40

Para determinar la distribución por tamaño de los ligandos asociados a HLA-B*40 se inmunoprecipitó dicha molécula y se analizó su repertorio peptídico mediante cromatografía líquida acoplada a espectrometría de masas. Se llevaron a cabo cuatro purificaciones independientes del peptidoma, con sus correspondientes pasos de enriquecimiento de fosfopéptidos, a partir de 10^{10} células C1R-B*40. En uno de los casos, la mezcla peptídica se analizó por LC-MS/MS usando CID como esquema de fragmentación mientras que en los tres restantes se empleó EThcD. La identificación de péptidos se llevó a cabo combinando varios motores de búsqueda y los resultados se filtraron a una FDR $\leq 1\%$ a nivel de péptido según se describe en materiales y métodos.

Se identificaron un total de 5909 y 7340 péptidos únicos mediante fragmentación CID y EThcD, respectivamente (Tabla 3) de los que 5801 (98%) y 7223 (98%) correspondían a secuencias de entre 8 y 13 residuos, la longitud típica de los ligandos de moléculas del MHC de clase I. Dado que la línea C1R expresa niveles bajos de HLA-B*35 y HLA-C*04 [138], los péptidos que presentaban motivos de unión a estos alotipos [154-156] fueron descartados del análisis. El número final de péptidos que presumiblemente correspondían a ligandos de HLA-B*40 fue de 5532 en la aproximación basada en CID y de 6952 en el abordaje mediante EThcD.

| Fragmentación | FDR $\leq 1\%$ | 8-13 | HLA-B*35:03 | HLA-C*04:01 | HLA-B*40:02 |
|---------------|----------------|------|-------------|--------------|----------------|
| CID | 5909 | 5801 | 47 (1%)* | 222 (4%)* | 5532 (95%)* |
| EThcD | 7340 | 7223 | 59 (1%)* | 212 (3%)* | 6952 (96%)* |

Tabla 3. Caracterización del peptidoma aislado a partir de la línea celular C1R-B*40. Se muestra el total de péptidos identificados mediante las fragmentaciones CID y EThcD a una FDR $\leq 1\%$, el número de péptidos con una longitud de 8 a 13 residuos de aminoácidos y el número de ligandos asignados a cada alotipo de clase I. * Este porcentaje es relativo al número total de péptidos de 8 a 13 residuos de longitud.

En torno al 30% de los ligandos descritos fueron secuenciados mediante ambas técnicas, mientras que 2554 y 3974 se identificaron solo por CID o EThcD, respectivamente (Figura 10). Todas las

secuencias peptídicas así como sus proteínas de origen se encuentran reflejadas en el material suplementario S1 y S2 incluido en el CD adjunto.

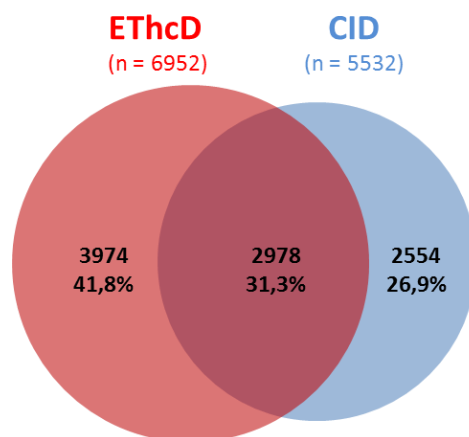


Figura 10. Representación del número de ligandos asociados a HLA-B*40 identificados por CID y/o ETHcD.

La distribución de tamaños de ambos grupos de ligandos siguió un patrón muy similar con una masa molecular promedio (\pm SD) de $1090,9 \pm 131,6$ y $1099,9 \pm 128,9$ Da para los péptidos identificados por CID y ETHcD respectivamente. No se observaron grandes diferencias en cuanto a la longitud de los ligandos identificados (Figura 11).

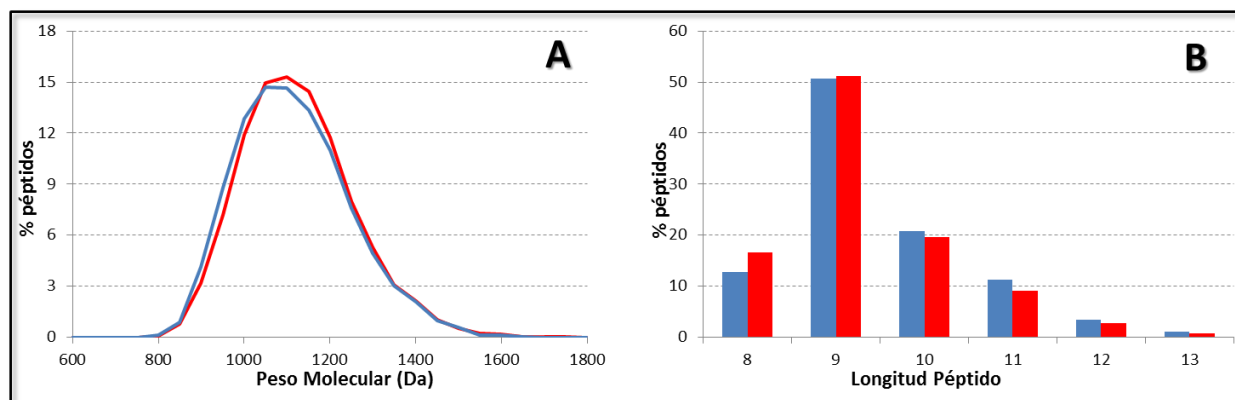


Figura 11. Distribución por tamaño del ligandoma presentado por HLA-B*40 e identificado mediante CID (color azul) o ETHcD (color rojo). (A) Se representan el % de todos los ligandos respecto a su peso molecular. (B) Distribución por tamaños del total de ligandos identificados.

Como era esperable, los nonámeros fueron la especie más representada constituyendo el 51% del repertorio peptídico unido a HLA-B*40 seguidos de decámeros (20-21%), octámeros (13-16%), undecámeros (9-11%), dodecámeros (3%) y tridecámeros (1%) (Tabla 4).

| Fragmentación | HLA-B*40:02 | Distribución por tamaño de ligandos en HLA-B*40:02 | | | | | |
|---------------|-------------|--|------------|------------|-----------|----------|---------|
| | | 8 (%) | 9 (%) | 10 (%) | 11 (%) | 12 (%) | 13 (%) |
| CID | 5532 | 705 (13%) | 2806 (51%) | 1146 (21%) | 620 (11%) | 193 (3%) | 62 (1%) |
| ETHcD | 6952 | 1153 (16%) | 3552 (51%) | 1369 (20%) | 628 (9%) | 194 (3%) | 56 (1%) |

Tabla 4. Distribución por tamaño del número de ligandos de HLA-B*40 identificados mediante CID o ETHcD.

IV.1.2- Caracterización de los motivos de anclaje de los péptidos presentados por HLA-B*40

Para determinar el motivo de anclaje de los péptidos unidos a HLA-B*40 se comparó la frecuencia de cada residuo de aminoácido en cada posición del péptido con la frecuencia de dicho residuo en la base de datos contra la que se efectuó la búsqueda. Se consideró que un residuo estaba siendo seleccionado en una posición determinada si la frecuencia observada (f_{obs}) era significativamente superior ($p < 0,05$) a la esperable a partir de la frecuencia de dicho residuo en la base de datos (f_{exp}). Para este análisis sólo se consideraron los ligandos con una longitud de entre 8 y 12 residuos ya que el número de tridecámeros era demasiado bajo para permitir un análisis estadístico robusto. Arbitrariamente, se clasificaron los motivos de anclaje como fuertes ($f_{obs}/f_{exp} > 5$) o débiles ($f_{obs}/f_{exp} < 5$) (Figura 12).

El análisis de las secuencias caracterizadas permitió identificar la posición 2 del péptido (P2) como el motivo fundamental de anclaje a HLA-B*40. La mayoría de los ligandos mostraron una preferencia casi absoluta por residuos ácidos (Glu y Asp) en esta posición (Figura 13). No obstante, la frecuencia de P2-Asp era apreciablemente menor que la de P2-Glu y disminuía según aumentaba la longitud de los péptidos. Por ejemplo, entre los octámeros identificado por ETHcD, se observaron 795 (69%, $f_{obs}/f_{exp} = 9,5$) y 268 (23%, $f_{obs}/f_{exp} = 4,8$) ligandos con los motivos P2-Glu y P2-Asp, respectivamente. En el caso de los nonámeros, P2-Glu estaba presente en 3263 ligandos (92%, $f_{obs}/f_{exp} = 12,7$) mientras que P2-Asp solo en 172 (5%, $f_{obs}/f_{exp} = 1,0$). Muy pocos decámeros (3), undecámeros (2) y dodecámeros (3) presentaban Asp en P2. En oposición, las frecuencias de P2-Glu alcanzaron el 97% ($f_{obs}/f_{exp} = 13,4$), 94% ($f_{obs}/f_{exp} = 13,0$) y 85% ($f_{obs}/f_{exp} = 11,8$), respectivamente, en estos mismos grupos de ligandos. Los resultados obtenidos en la aproximación basada en CID arrojaron resultados casi idénticos (Tabla 5).

La segunda posición principal de anclaje a B*40 resultó ser el extremo C-terminal del péptido (P Ω). Aquí, la especificidad fue ligeramente menor que en el caso de P2 (Figura 14). Se observó, una fuerte selección de Leu que afectaba a 696 octámeros (60%, $f_{obs}/f_{exp} = 6,1$), 2039 nonámeros (57%, $f_{obs}/f_{exp} = 5,8$), 863 decámeros (63%, $f_{obs}/f_{exp} = 6,4$), 417 undecámeros (66%, $f_{obs}/f_{exp} = 6,7$) y 122 dodecámeros (62%, $f_{obs}/f_{exp} = 6,4$) entre los péptidos identificados mediante ETHcD. De nuevo, el patrón

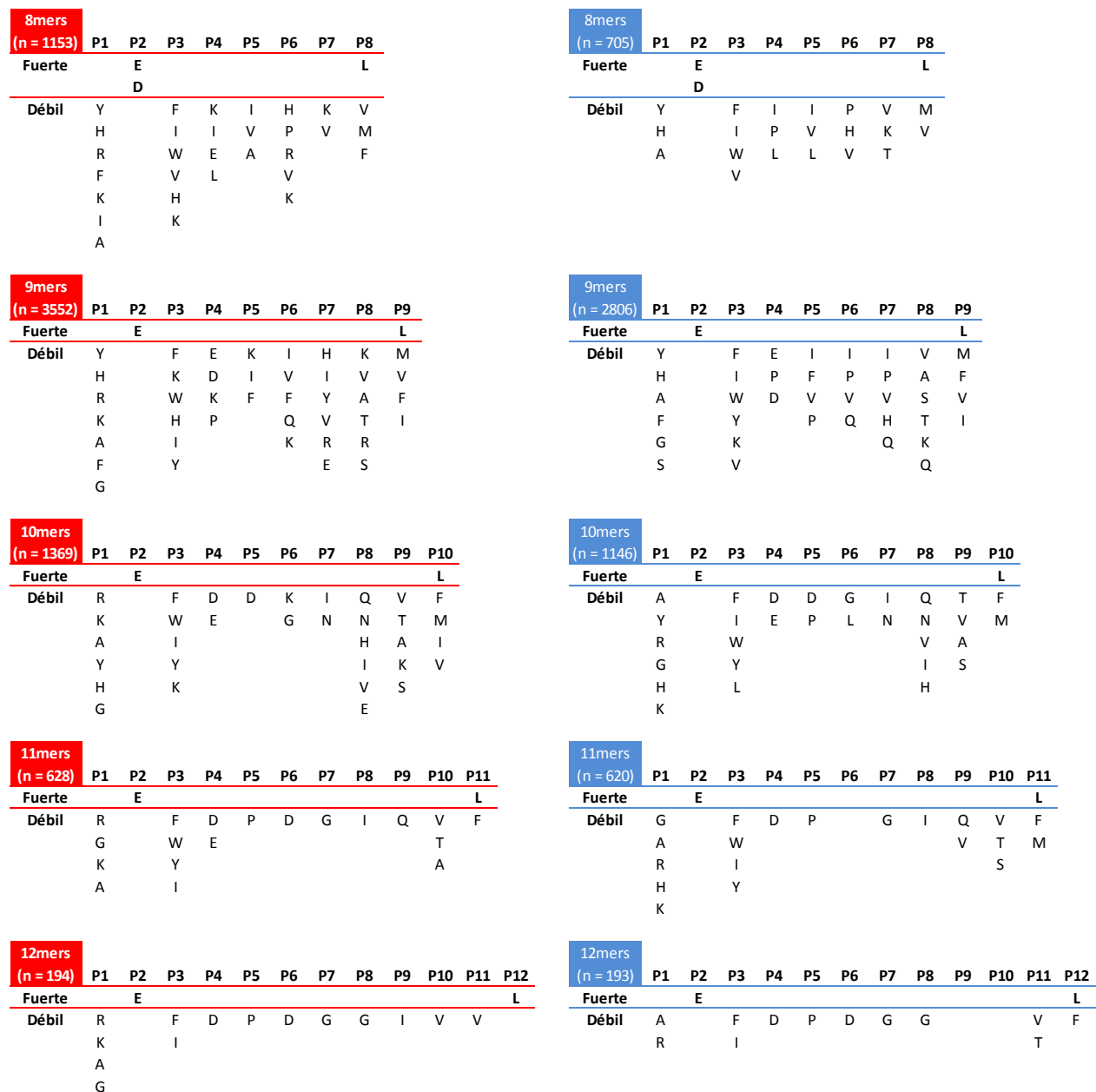


Figura 12. Motivo de unión a HLA-B*40. Se muestran los residuos estadísticamente sobrerrepresentados ($p < 0,05$) para los distintos tamaños de ligando. Las preferencias de unión se clasificaron como fuertes ($f_{obs}/f_{exp} > 5$) o débiles ($f_{obs}/f_{exp} < 5$).

observado resultó ser independiente del esquema de fragmentación (Tabla 5). Otros residuos como Met, Phe, Val o Ile también se encontraban sobrerrepresentados de manera significativa. Finalmente, Ala también estaba presente en P Ω , aunque con una frecuencia más baja de lo esperable.

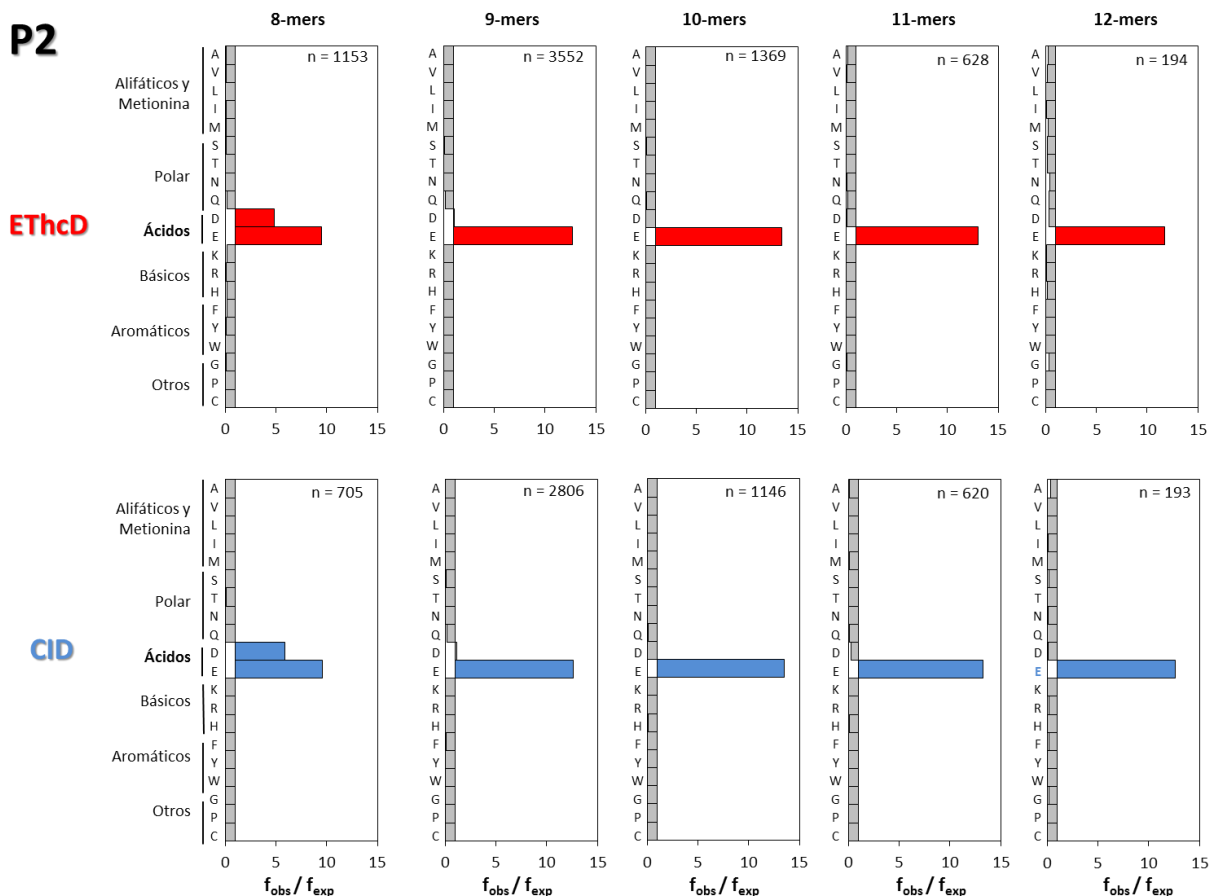


Figura 13. Uso de residuos en P2 entre los ligandos de HLA-B*40 identificados por ETHcD (rojo) o CID (azul). La frecuencia de cada residuo observado (f_{obs}) se comparó con la frecuencia del mismo residuo en la base de datos (f_{exp}). Las diferencias estadísticamente significativas ($p < 0,05$) se representan en color rojo (ETHcD) o azul (CID).

Además de los dos anclajes principales antes comentados, se observaron preferencias en las restantes posiciones de los ligandos, aunque mucho más sutiles que las que afectaban a P2 y P Ω . En este sentido, los patrones detectados mediante los dos esquemas de fragmentación empleados fueron comparables. No obstante, la mayor potencia estadística en el análisis por ETHcD –debida al mayor número de identificaciones– evidenció un número ligeramente mayor de residuos sobrerrepresentados en distintas posiciones (Figura 12).

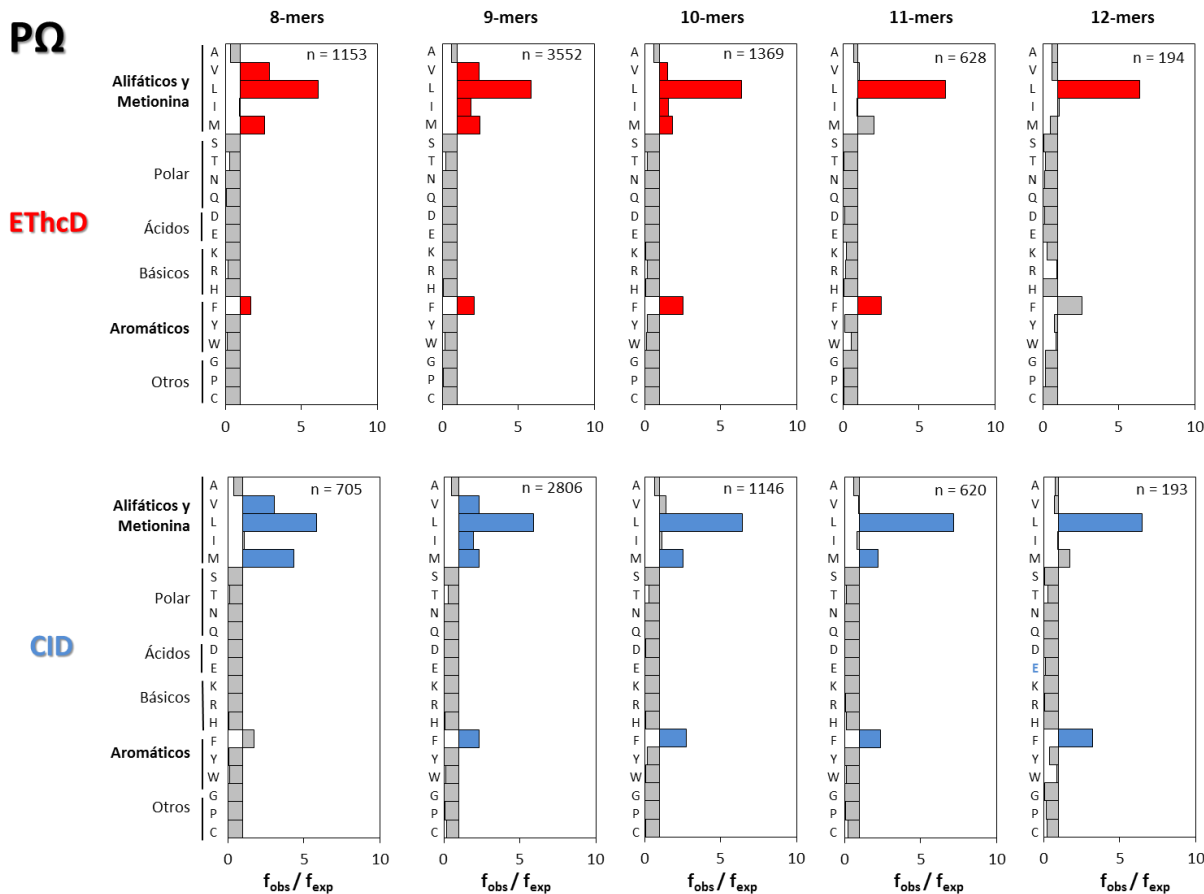


Figura 14. Uso de residuos en PΩ entre los ligandos de HLA-B*40 identificados por ETHcD (rojo) o CID (azul). La frecuencia de cada residuo observado (f_{obs}) se comparó con la frecuencia del mismo residuo en la base de datos (f_{exp}). Las convenciones son las mismas que en la Figura 13.

| | | P2 | | | | | | | | | | PΩ | | | | |
|-------|-------------------|---------|-----|---------|-----|----------|------|----------|-----|----------|-----|-----|-----|-----|-----|-----|
| | | 8 meros | | 9 meros | | 10 meros | | 11 meros | | 12 meros | | Leu | Met | Phe | Val | Ile |
| CID | Péptidos (%) | 69 | 28 | 91 | 6 | 98 | 0,2 | 96 | 1 | 91 | 0,5 | 61 | 6 | 8 | 12 | 6 |
| | f_{obs}/f_{exp} | 9,5 | 5,9 | 12,6 | 1,2 | 13,5 | 0,04 | 13,3 | 0,2 | 12,6 | 0,1 | 6,2 | 2,6 | 2,4 | 2 | 1,5 |
| ETHcD | Péptidos (%) | 69 | 23 | 92 | 5 | 97 | 0,2 | 94 | 0,3 | 85 | 1,5 | 60 | 5 | 8 | 13 | 7 |
| | f_{obs}/f_{exp} | 9,5 | 4,8 | 12,7 | 1 | 13,4 | 0,05 | 13 | 0,1 | 11,8 | 0,3 | 6,0 | 2,3 | 2,2 | 2,1 | 1,5 |

Tabla 5. Motivos principales de anclaje de HLA-B*40. Se indica el porcentaje de péptidos y los ratios f_{obs}/f_{exp} de los principales residuos seleccionados en P2 y PΩ.

IV.1.3-Estructura de HLA-B*40 en complejo con un ligando endógeno

Para analizar las bases estructurales que establecen las preferencias de unión de HLA-B*40, se determinó la estructura cristalográfica de este alotipo en complejo con el péptido REFSKEPEL (Figura 15,

PDB: 5IEK). La estructura se resolvió a una resolución de 1,8 Å mediante reemplazo molecular. Los parámetros cristalográficos y estadísticas del refinamiento se resumen en la Tabla 6.

El análisis de la estructura reveló que el péptido REFSKEPEL adoptaba una conformación extendida a lo largo del surco de unión de HLA-B*40 (Figura 15A, B y C). Los residuos P2-Glu y P9-Leu se hallaban anclados en las subcavidades B y F, respectivamente. La cadena lateral de P6-Glu se encontraba orientada hacia abajo mientras que P1-Arg, P4-Ser, P5-Lys y P8-Glu se disponían hacia el solvente siendo potencialmente accesibles a la interacción con el TCR. Por último, la cadena lateral de P7-Pro se disponía de manera aproximadamente paralela al suelo del surco de unión y P3-Phe se ubicaba en el interior de la subcavidad D.

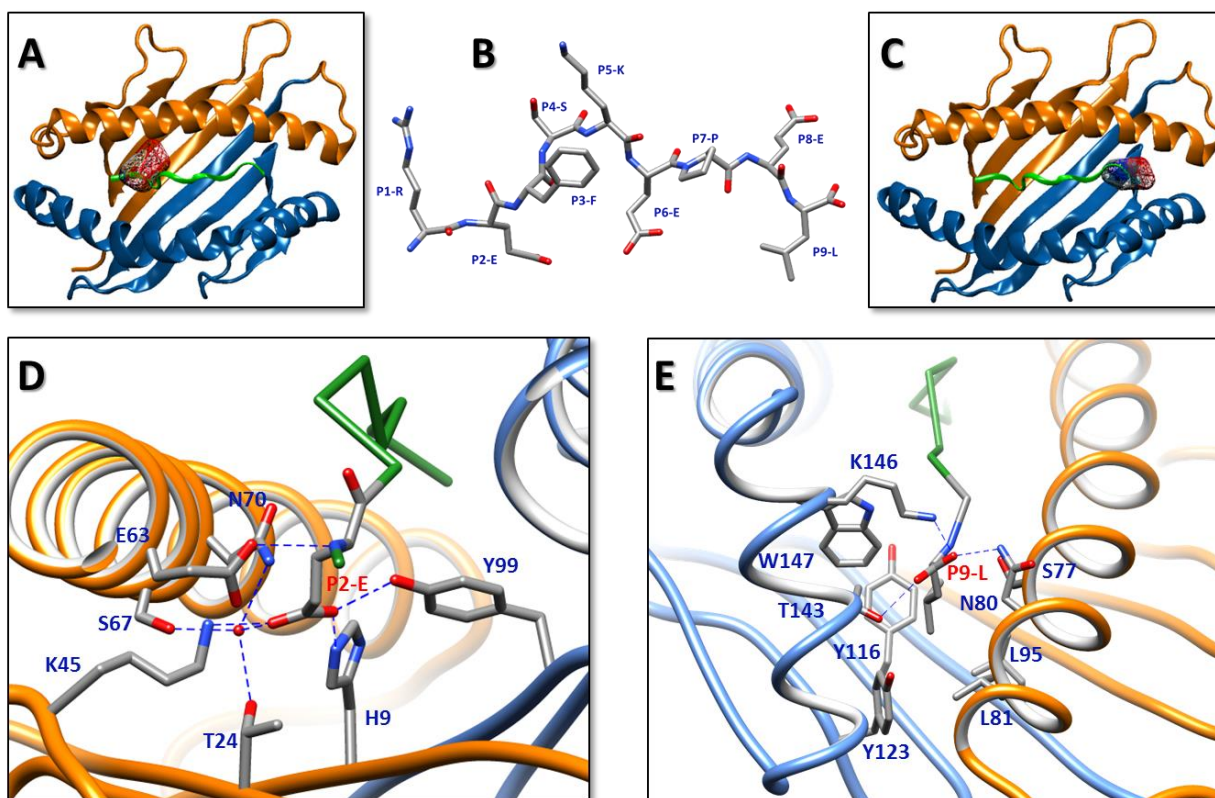


Figura 15. Estructura cristalográfica de HLA-B*40 en complejo con el ligando REFSKEPEL. α_1 y α_2 se representan en naranja y azul, respectivamente. (A y C) Vista superior surco de unión mostrando el ligando en conformación extendida. Se muestra la envoltura de los residuos P2-Glu (A) y P Ω -Leu (C). (B) Estructura del péptido en el surco de unión de HLA-B*40. Se muestran de color azul y en código de una letra los residuos que componen dicho ligando. (D) Interacción del residuo P2-Glu del ligando con algunos de los residuos ubicados en la subcavidad B de la molécula de clase I. (E) Interacción del residuo P Ω -Leu del ligando con algunos de los residuos ubicados en la subcavidad F de la molécula de clase I. Las interacciones electrostáticas y los puentes de hidrógeno se representan con líneas discontinuas de color azul. Los residuos de aminoácidos se indican con el código de una letra y de color azul (cadena pesada) o rojo (péptido). La cadena principal del péptido se muestra de color verde.

Respecto al residuo P2-Glu, su grupo α -amino interactuaba mediante un puente de hidrógeno con el grupo γ -carboxilo del residuo Glu63. Por otro lado, su cadena lateral se ubicaba en el interior de la subcavidad B estableciendo un gran número de interacciones con diferentes residuos de la cadena pesada incluyendo His9, Lys45 y Tyr99 –que contactaban directamente con el grupo carboxilo de P2-Glu– y Thr24, Ser67 y Asn70 –que lo hacían a través de una molécula de agua–. En el extremo C-terminal, por su parte, el grupo carboxilo de P9-Leu, se estabilizaba mediante puentes de hidrógeno con los residuos Asn80 y Thr143 mientras que su cadena lateral se posicionaba en el interior de la subcavidad F estabilizándose mediante interacciones hidrofóbicas con los residuos Leu81, Leu95, Tyr116, Tyr123 y Trp147.

| Colección de Datos | HLA-B*40 + G(pS)FSRFYSL | HLA-B*40 + REF(pS)KEPEL | HLA-B*40 + REFSKEPEL |
|--|--|---|---|
| PDB ID | Pendiente | 5IEH | 5IEK |
| Wavelength (Å) | 1,00555 | 1,00556 | 1,00556 |
| Space group | P ₄ ₃ 2 ₁ 2 | P2 ₁ 2 ₁ 2 ₁ | P2 ₁ 2 ₁ 2 ₁ |
| Cell dimensions (a,b,c) (Å) | 50,115 50,115 338,271 90 90 90 | 51,044 82,249 110,084 90 90 90 | 51,26 82,201 111,145 90 90 90 |
| Resolution range (Å) | 25-2,5 (2,589 – 2,5) | 25-1,5 (1,554 – 1,5) | 25-1,8 (1,864 – 1,8) |
| Total reflections | 32076 (3068) | 149752 (14732) | 88506 (8693) |
| Unique reflections | 16050 (1535) | 74898 (7370) | 44288 (4349) |
| Multiplicity | 2,0 (2,0) | 2,0 (2,0) | 2,0 (2,0) |
| Data completeness (%) | 1,00 (1,00) | 1,00 (1,00) | 1,00 (1,00) |
| I/σ | 11,77 (3,06) | 32,48 (8,51) | 6,19 (1,63) |
| Wilson B-factor | 34,79 | 13,75 | 19,74 |
| R_{merge} | 0,04578 (0,2563) | 0,01038 (0,07625) | 0,04711 (0,2432) |
| R_{meas} | 0,06474 (0,3624) | 0,01468 (0,1078) | 0,06662 (0,344) |
| CC^{1/2} | 0,997 (0,857) | 1 (0,982) | 0,996 (0,873) |
| CC* | 0,999 (0,961) | 1 (0,995) | 0,999 (0,965) |
| Refinamiento Estadístico | | | |
| Reflections used in refinement | 16048 (1535) | 74897 (7370) | 44243 (4341) |
| Reflections used for R_{free} | 837 (84) | 3697 (363) | 2211 (221) |
| R_{work} | 0,2029 (0,2793) | 0,1611 (0,1791) | 0,1613 (0,2208) |
| R_{free} | 0,2413 (0,3723) | 0,1837 (0,2100) | 0,1870 (0,2638) |
| Non-hydrogen atoms | 3287 | 3816 | 3716 |
| Macromolecules | 3174 | 3181 | 3190 |
| Ligands | - | 4 | - |
| Water | 113 | 621 | 526 |
| Protein residues | 384 | 383 | 385 |
| Rms deviations from ideality | | | |
| Bond lengths (Å) | 0,005 | 0,008 | 0,009 |
| Bond angles (°) | 0,76 | 0,99 | 1,04 |
| Ramachandran plot (%) | | | |
| Favored region | 98 | 99 | 99 |
| Allowed region | 2,4 | 1,3 | 1 |
| Outliers region | 0 | 0 | 0 |
| Rotamer outliers | 2,1 | 0,88 | 3,2 |
| B-factors (Å²) | 37,29 | 18,09 | 24,56 |
| Average macromolecules | 37,38 | 15,77 | 22,88 |
| Average ligands | - | 23 | - |
| Average solvent | 34,49 | 29,84 | 34,72 |

Tabla 6. Colección de datos cristalográficos y refinamiento estadístico de la estructura.

IV.2-Análisis del fosfopeptidoma presentado por HLA-B*40 y bases estructurales de la presentación de fosfopéptidos por este alotipo

IV.2.1-Identificación de ligandos fosforilados unidos a HLA-B*40

Como se comentó en la introducción a esta tesis, la caracterización de fosfopéptidos basada en fragmentación CID entraña una considerable dificultad, debido a que los espectros MS2 de las especies fosforiladas son poco informativos. De hecho, usando esta técnica, no pudimos identificar ningún péptido fosforilado a una $FDR \leq 1\%$ a partir de la búsqueda con espectros MS/MS. Para solventar esta dificultad, los espectros MS2 asignados a secuencias fosforiladas con un *score* de MASCOT superior a 20 fueron inspeccionados a mano. Aquellos cuya identificación tenía visos de ser correcta fueron sintetizados y fragmentados mediante CID para validarlos o descartarlos definitivamente (Figura 16 y Anexo 1.1). Mediante este procedimiento se confirmaron 85 secuencias fosforiladas (N = 1-85 en la Tabla 7).

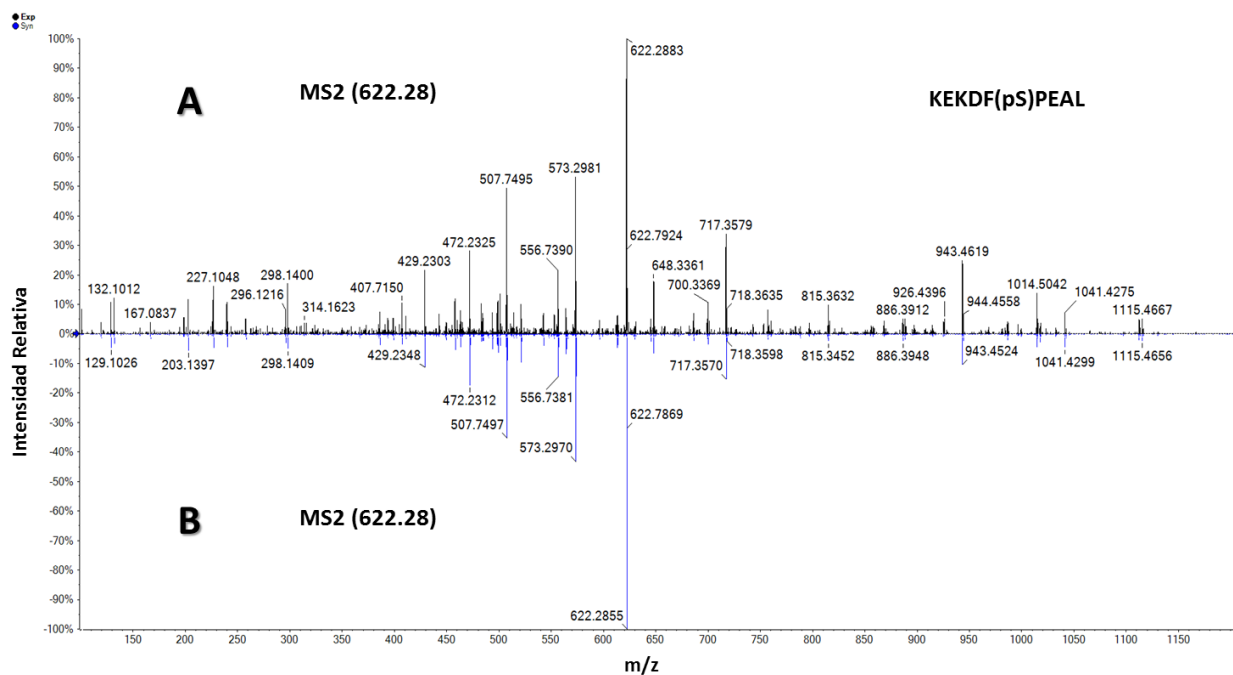


Figura 16. Espectro de fragmentación por CID del péptido fosforilado KEDF(pS)PEAL. (A) Ligando endógeno. (B) Péptido sintético.

Adicionalmente, comparamos los tiempos de retención de los 85 fosfopéptidos sintéticos con los determinados experimentalmente en nuestra muestra, obteniéndose una altísima correlación ($R^2 = 0,9991$) entre ambas medidas (Figura 17). Entre los fosfopéptidos identificados, nos llamó la atención la

presencia de tres secuencias (N = 57, 65 y 81 en la Tabla 7) que contenían fosfoserina (pSer) en P2, en lugar de los residuos de Glu o Asp que eran esperables. Esto nos llevó a hipotetizar que la similitud estructural existente entre los residuos de Glu y pSer podría explicar el uso de este último como anclaje principal a HLA-B*40.

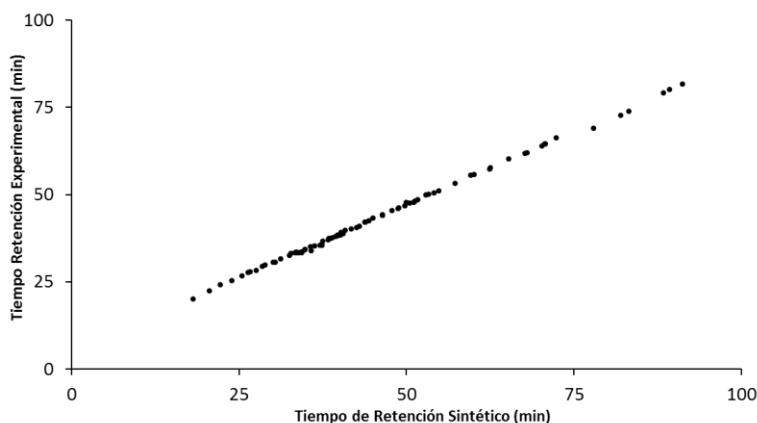


Figura 17. Confirmación de la identificación por CID de péptidos fosforilados presentados por HLA-B*40. El tiempo de retención de los candidatos endógenos se representa frente al tiempo de retención de los correspondientes péptidos sintéticos. Los datos se ajustan a una función lineal ($R^2 = 0,9991$).

A diferencia del abordaje basado en CID, el uso de espectrometría de masas con fragmentación EThcD supuso una gran ventaja para la identificación de fosfopéptidos. La aplicación de esta técnica permitió la identificación con una $FDR \leq 1\%$ de 113 secuencias fosforiladas con Glu, Asp o pSer en P2. Para evaluar la fiabilidad de estas identificaciones, se seleccionaron 26 secuencias, se sintetizaron los correspondientes fosfopéptidos y se fragmentaron mediante EThcD. Este procedimiento permitió la confirmación de 25 (96%) secuencias y la identificación de 1 falso positivo que fue descartado para los análisis posteriores (Anexo 1.2).

Como resumen, la Tabla 7 muestra los 148 fosfopéptidos –derivados de 135 proteínas– identificados en el peptidoma eluído de HLA-B*40 que son el resultado de combinar las identificaciones mediante fragmentación CID y EThcD. En este grupo de ligandos, la fosforilación afectaba a residuos de Ser (93%) y de Thr (7%) exclusivamente. Las 148 secuencias fosforiladas incluían 137 sitios únicos de fosforilación, de los que 68 (50%) estaban ya anotados en las bases de datos Uniprot (<http://www.uniprot.org>) y/o HPRD (*Human Protein Reference Database*, <http://www.hprd.org>). Los 69 restantes (50%), por lo tanto, corresponden a nuevos sitios de fosforilación descritos en este trabajo por primera vez.

| N | Secuencia | P | NA | Nombre de la proteína | CID | ETHcD | Syn | no-P | Ref |
|----|-----------------|---|--------|--|-----|-------|-----|------|-------|
| 1 | AEL(pS)PVEQKL | 4 | Q9Y5P8 | Protein phosphatase 2A 48 kDa regulatory subunit | + | + | + | - | - |
| 2 | REL(pS)PQRNL | 4 | Q96F44 | E3 ubiquitin-protein ligase TRIM11 | + | - | + | - | - |
| 3 | TER(pS)PGPRL | 4 | O14978 | Zinc finger protein 263 | + | - | + | - | - |
| 4 | IES(pS)PQYRL | 4 | Q8WWM7 | Ataxin-2-like protein | + | - | + | + | [157] |
| 5 | KEKDF(pS)PEAL | 6 | O15372 | Eukaryotic translation initiation factor 3 subunit H | + | + | + | + | [158] |
| 6 | GEW(pS)ASLPHRF | 4 | O43164 | E3 ubiquitin-protein ligase Praja-2 | + | + | + | + | - |
| 7 | SEA(pS)PSREAL | 4 | Q13111 | Chromatin assembly factor 1 subunit A | + | + | + | - | [159] |
| 8 | YEG(pS)PIKVTL | 4 | P06748 | Nucleophosmin | + | + | + | + | [160] |
| 9 | HERSI(pS)PLL | 6 | Q9UBT6 | DNA polymerase kappa | + | - | + | - | - |
| 10 | VEKKIR(pS)L | 8 | O75394 | 39S ribosomal protein L33, mitochondrial | + | - | + | + | - |
| 11 | AEN(pS)PTRQQF | 4 | Q86XP3 | ATP-dependent RNA helicase DDX42 | + | + | + | - | [160] |
| 12 | QEV(pT)PSRSTL | 4 | P49454 | Centromere protein F | + | - | + | - | [161] |
| 13 | RE(pS)PPPSAL | 3 | Q86VE0 | Myb-related transcription factor, partner of profilin | + | - | + | - | - |
| 14 | REK(pT)PSEDKL | 4 | Q8NFC6 | Biorientation of chromosomes in cell division protein 1-like 1 | + | + | + | - | - |
| 15 | GEDR(pS)PPTAL | 5 | Q9UBU5 | MAP kinase-interacting serine/threonine-protein kinase 1 | + | - | + | + | [157] |
| 16 | REV(pS)PAPAVA | 4 | O60292 | Signal-induced proliferation-associated 1-like protein 3 | + | - | + | - | [162] |
| 17 | KEV(pT)PKPVL | 4 | P17544 | Cyclic AMP-dependent transcription factor ATF-7 | + | - | + | - | - |
| 18 | HEL(pS)PPQKRM | 4 | Q9BXP5 | Serrate RNA effector molecule homolog | + | - | + | - | [159] |
| 19 | RE(pT)PSRPM | 3 | Q9UQ35 | Serine/arginine repetitive matrix protein 2 | + | - | + | - | [163] |
| 20 | REYG(pS)PLKA | 5 | Q9UQ88 | Cyclin-dependent kinase 11A | + | + | + | - | [157] |
| 21 | KEM(pS)PTRQL | 4 | Q4G0N7 | Protein FAM229B | + | + | + | - | - |
| 22 | GENK(pS)PLLL | 5 | Q8IZD2 | Histone-lysine N-methyltransferase MLL5 | + | + | + | - | - |
| 23 | VEFPH(pS)PEI | 6 | Q9BR61 | Acyl-CoA-binding domain-containing protein 6 | + | + | + | + | - |
| 24 | YEKAN(pT)PEL | 6 | Q9H8V3 | Protein ECT2 | + | + | + | - | [159] |
| 25 | VEK(pS)PRPPL | 4 | Q7L590 | Protein MCM10 homolog | + | + | + | - | - |
| 26 | QEK(pS)PKQAL | 4 | O43683 | Mitotic checkpoint serine/threonine-protein kinase BUB1 | + | + | + | - | [157] |
| 27 | GEI(pS)PQREV | 4 | Q8WWM1 | LIM domain only protein 7 | + | + | + | - | [160] |
| 28 | HEFH(pS)QENL | 5 | Q9Y3M8 | StAR-related lipid transfer protein 13 | + | + | + | - | [164] |
| 29 | REG(pS)GRFSL | 4 | P57764 | Gasdermin-D | + | - | + | - | - |
| 30 | GEN(pS)GIGKL | 4 | Q6P0N0 | Mis18-binding protein 1 | + | + | + | - | [157] |
| 31 | SEI(pS)PIKGSV | 4 | Q7Z5K2 | Wings apart-like protein homolog | + | + | + | - | [160] |
| 32 | SEK(pS)LENETL | 4 | P23588 | Eukaryotic translation initiation factor 4B | + | - | + | + | [160] |
| 33 | RED(pS)QQRGAHL | 4 | Q32P51 | Heterogeneous nuclear ribonucleoprotein A1-like 2 | + | - | + | + | [165] |
| 34 | RED(pS)JGFSL | 4 | O75569 | Protein activator of the interferon-induced protein kinase | + | + | + | - | [157] |
| 35 | REA(pS)PSRLSV | 4 | O75122 | CLIP-associating protein 2 | + | + | + | - | - |
| 36 | SEQ(pS)LGQKL | 4 | P33991 | DNA replication licensing factor MCM4 | + | - | + | + | [159] |
| 37 | REV(pS)PAPAV | 4 | O60292 | Signal-induced proliferation-associated 1-like protein 3 | + | - | + | - | [162] |
| 38 | SERM(pS)LPPL | 5 | Q15398 | Disks large-associated protein 5 | + | - | + | - | [166] |
| 39 | SES(pS)PPKAM | 4 | Q9UHB6 | LIM domain and actin-binding protein 1 | + | + | + | - | [157] |
| 40 | SEV(pS)PSGVGF | 4 | Q6ULP2 | Aftiphilin | + | + | + | - | - |
| 41 | REK(pS)PGRML | 4 | O14578 | Citron Rho-interacting kinase | + | + | + | - | [162] |
| 42 | SEF(pS)DVDKL | 4 | O75665 | Oral-facial-digital syndrome 1 protein | + | + | + | + | - |
| 43 | REA(pS)PLSSNKL | 4 | Q9Y462 | Zinc finger protein 711 | + | + | + | - | - |
| 44 | REIES(pS)PQYRL | 6 | Q8WWM7 | Ataxin-2-like protein | + | - | + | + | [157] |
| 45 | REY(pS)PPYAP | 4 | Q9UPT8 | Zinc finger CCCH domain-containing protein 4 | + | - | + | - | [157] |
| 46 | KEKEV(pS)QEL | 6 | A8MX76 | Calpain-14 | + | - | + | - | - |
| 47 | KEL(pS)PPGGL | 4 | Q7Z456 | Kinesin-like protein KIF21A | + | - | + | - | [157] |
| 48 | RER(pS)PSPSF | 4 | P49585 | Choline-phosphate cytidylyltransferase A | + | + | + | - | [160] |
| 49 | GET(pS)LMRTL | 4 | O96017 | Serine/threonine-protein kinase Chk2 | + | - | + | - | [167] |
| 50 | AEK(pS)PGPIV | 4 | Q86WB0 | Nuclear-interacting partner of ALK | + | + | + | - | [157] |
| 51 | REN(pS)PVQSI | 4 | Q7Z221 | Treslin | + | - | + | - | - |
| 52 | RET(pS)PNRIGL | 4 | Q7Z460 | CLIP-associating protein 1 | + | + | + | - | - |
| 53 | REG(pS)FRVTTA | 4 | P49757 | Protein numb homolog | + | + | + | - | - |
| 54 | PEV(pS)PRPAL | 4 | Q96L73 | Lysine N-methyltransferase 3B | + | + | + | - | - |
| 55 | KEL(pS)PAGSI | 4 | O95453 | Poly(A)-specific ribonuclease PARN | + | + | + | - | [159] |
| 56 | REG(pS)YDRYL | 4 | Q9HCD5 | Nuclear receptor coactivator 5 | + | + | + | - | [168] |
| 57 | S(pS)YGNIRAV | 2 | Q8N122 | Regulatory-associated protein of mTOR | + | + | + | - | [169] |
| 58 | GEPHS(pS)PEL | 6 | Q9NUA8 | Zinc finger and BTB domain-containing protein 40 | + | - | + | - | - |
| 59 | REK(pS)PGRMLST | 4 | O14578 | Citron Rho-interacting kinase | + | - | + | - | - |
| 60 | SEL(pS)PGRSV | 4 | Q8NFF5 | FAD synthase | + | + | + | + | [162] |
| 61 | HEFS(pS)PSHL | 5 | Q9H3D4 | Tumor protein 63 | + | - | + | + | - |
| 62 | QEK(pS)PEERL | 4 | Q7RTP6 | Protein-methionine sulfoxide oxidase MICAL3 | + | + | + | - | [157] |
| 63 | KEQ(pS)PEPHL | 4 | Q5VV52 | Zinc finger protein 691 | + | + | + | - | [170] |
| 64 | NEA(pS)PVKRPRL | 4 | Q9H8U3 | AN1-type zinc finger protein 3 | + | - | + | - | [171] |
| 65 | G(pS)FSRFYSL | 2 | P22695 | Cytochrome b-c1 complex subunit 2, mitochondrial | + | - | + | - | - |
| 66 | REF(pS)KEPEL | 4 | Q9NQS7 | Inner centromere protein | + | + | + | - | [161] |
| 67 | SEKPV(pS)PKSGTL | 6 | Q14155 | Rho guanine nucleotide exchange factor 7 | + | + | + | - | [168] |
| 68 | REL(pS)PLISL | 4 | P51825 | AF4/FMR2 family member 1 | + | + | + | - | [162] |
| 69 | RE(pT)PVPPAM | 3 | Q9Y2F5 | Uncharacterized protein KIAA0947 | + | - | + | - | [160] |
| 70 | REP(pS)PSSSSL | 4 | O60336 | Mitogen-activated protein kinase-binding protein 1 | + | - | + | - | - |

Resultados

| | | | | | | | | | |
|-----|--------------------|----|----------|---|---|---|---|---|-------|
| 71 | LEKRT(pS)PQVL | 6 | Q7Z417 | Nuclear fragile X mental retardation-interacting protein 2 | + | + | + | - | [157] |
| 72 | REL(pS)GTIKEI | 4 | P30050 | 60S ribosomal protein L12 | + | + | + | - | - |
| 73 | GEA(pS)PSHII | 4 | Q9ULL5-3 | Proline-rich protein 12 | + | - | + | - | - |
| 74 | GEQ(pS)PNVSL | 4 | Q58WW2 | DDB1- and CUL4-associated factor 6 | + | + | + | - | [157] |
| 75 | KE(pT)PDKVEL | 3 | Q13610 | Periodic tryptophan protein 1 homolog | + | - | + | - | [161] |
| 76 | NEFH(pS)PIGL | 5 | Q06413 | Myocyte-specific enhancer factor 2C | + | + | + | + | - |
| 77 | REV(pS)PEPIV | 4 | Q8N103 | T-cell activation Rho GTPase-activating protein | + | + | + | - | - |
| 78 | NEK(pS)PYASL | 4 | Q6IQ21 | Zinc finger protein 770 | + | + | + | - | - |
| 79 | REL(pS)GTIKEIL | 4 | P30050 | 60S ribosomal protein L12 | + | - | + | - | - |
| 80 | SERIMQL(pS)L | 8 | Q8IY63 | Angiomotin-like protein 1 | + | - | + | - | - |
| 81 | R(pS)FPTLPTL | 2 | Q13761 | Runt-related transcription factor 3 | + | - | + | - | - |
| 82 | KES(pS)PIPSP | 4 | Q01082 | Spectrin beta chain, non-erythrocytic 1 | + | - | + | - | [157] |
| 83 | SE(pS)GSNLHAL | 3 | Q6P4F7 | Rho GTPase-activating protein 11A | + | - | + | + | - |
| 84 | KELARQ(pS)F | 8 | Q9Y385 | Ubiquitin-conjugating enzyme E2 J1 | + | - | + | - | - |
| 85 | KEN(pT)PPAL | 4 | P31350 | Ribonucleoside-diphosphate reductase subunit M2 | + | - | + | - | [162] |
| 86 | FEH(pS)PLSEDRF | 4 | Q02078-2 | Myocyte-specific enhancer factor 2A | - | + | + | - | - |
| 87 | REA(pS)PSRLSVA | 4 | O75122 | CLIP-associating protein 2 | - | + | + | - | - |
| 88 | TELNL(pS)KGSLS | 6 | Q5W0B1 | RING finger protein 219 | - | + | - | - | [172] |
| 89 | REH(pS)NMAIDL | 4 | Q9UKS7 | Zinc finger protein Helios | - | + | - | - | - |
| 90 | REG(pS)FRVTT | 4 | P49757 | Protein numb homolog | - | + | - | - | - |
| 91 | VDRHMESLSQSK(pS)FR | 13 | Q96F86 | Enhancer of mRNA-decapping protein 3 | - | + | - | - | - |
| 92 | REN(pS)FGSPLEF | 4 | Q9UQ35 | Serine/arginine repetitive matrix protein 2 | - | + | - | - | [162] |
| 93 | RER(pS)KDFQV | 4 | Q5VUA4 | Zinc finger protein 318 | - | + | + | - | - |
| 94 | EEVAKRN(pS)L | 8 | Q43169 | Cytochrome b5 type B | - | + | + | - | [157] |
| 95 | AEV(pS)PDAKTSL | 4 | Q68DQ2 | Very large A-kinase anchor protein | - | + | + | - | - |
| 96 | AFT(pS)PPSLFK | 5 | Q6UN15 | Pre-mRNA 3'-end-processing factor FIP1 | - | + | - | - | [157] |
| 97 | SEK(pS)EQAAL | 4 | Q9UPV0 | Centrosomal protein of 164 kDa | - | + | - | - | - |
| 98 | KEFDGK(pS)LVSV | 7 | P08238 | Heat shock protein HSP 90-beta | - | + | + | - | [173] |
| 99 | GEF(pS)PNVPKSL | 4 | Q9Y2R2 | Tyrosine-protein phosphatase non-receptor type 22 | - | + | - | + | - |
| 100 | KEQ(pS)PQTLEL | 4 | Q96T49 | Protein phosphatase 1 regulatory inhibitor subunit 16B | - | + | - | + | - |
| 101 | RELH(pS)PSSEL | 5 | P41161 | ETS translocation variant 5 | - | + | - | - | - |
| 102 | AEINK(pS)LLAL | 6 | O00139 | Kinesin-like protein KIF2A | - | + | - | + | - |
| 103 | KEN(pS)PEKFAL | 4 | Q12824 | SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily B member 1 | - | + | - | + | - |
| 104 | QEK(pS)PVKTV | 4 | Q5VWN6 | Protein FAM208B | - | + | - | - | [160] |
| 105 | RES(pS)RSHSVF | 4 | Q02241 | Kinesin-like protein KIF23 | - | + | + | - | - |
| 106 | R(pS)IGETISL | 2 | Q9Y2G0 | Protein EFR3 homolog B | - | + | - | - | - |
| 107 | RER(pS)MSENAV | 4 | Q9GZY8 | Mitochondrial fission factor | - | + | - | - | [160] |
| 108 | AEK(pS)LLEKL | 4 | Q6UB99 | Ankyrin repeat domain-containing protein 11 | - | + | - | - | - |
| 109 | SEL(pS)PIREEL | 4 | Q8N573 | Oxidation resistance protein 1 | - | + | - | - | - |
| 110 | KEI(pS)PYENKM | 4 | Q8NG31 | Protein CASC5 | - | + | - | - | [157] |
| 111 | S(pS)PVNVKCLK | 2 | Q00839 | Heterogeneous nuclear ribonucleoprotein U | - | + | - | - | [159] |
| 112 | KEA(pS)PLRVI | 4 | Q5QP82 | DDB1- and CUL4-associated factor 10 | - | + | - | - | - |
| 113 | REA(pS)PSRLS | 4 | O75122 | CLIP-associating protein 2 | - | + | - | - | - |
| 114 | RET(pS)FQRKEL | 4 | Q2M3G4 | Protein Shroom1 | - | + | + | - | - |
| 115 | RES(pS)LGFQL | 4 | Q5TCZ1 | SH3 and PX domain-containing protein 2A | - | + | - | - | - |
| 116 | REK(pS)PVREP | 4 | Q86U06 | Probable RNA-binding protein 23 | - | + | - | - | [160] |
| 117 | KER(pS)PQQNIV | 4 | Q8NEZ4 | Histone-lysine N-methyltransferase 2C | - | + | - | - | [174] |
| 118 | REDSG(pT)FSL | 6 | O75569 | Interferon-inducible double-stranded RNA-dependent protein kinase activator A | - | + | - | - | [157] |
| 119 | VEH(pS)PFSSF | 4 | Q7LBC6 | Lysine-specific demethylase 3B | - | + | - | - | [157] |
| 120 | GELK(pS)PKVL | 5 | Q969H0 | F-box/WD repeat-containing protein 7 | - | + | - | - | - |
| 121 | GEK(pS)PEPLGL | 4 | Q9UPN7 | Serine/threonine-protein phosphatase 6 regulatory subunit 1 | - | + | - | + | - |
| 122 | AEK(pS)RTYVA | 4 | P40222 | Alpha-taxilin | - | + | - | - | - |
| 123 | KEF(pS)PFGTITSA | 4 | P11940 | Polyadenylate-binding protein 1 | - | + | - | + | [157] |
| 124 | SEH(pS)MDTLL | 4 | O00244 | Copper transport protein ATOX1 | - | + | - | + | - |
| 125 | KEI(pS)ITHHV | 4 | Q15418 | Ribosomal protein S6 kinase alpha-1 | - | + | - | - | - |
| 126 | GEN(pS)RVARL | 4 | O43847 | Nardilysin | - | + | + | - | - |
| 127 | GER(pS)GLKPPL | 4 | Q9Y6A5 | Transforming acidic coiled-coil-containing protein 3 | - | + | - | + | - |
| 128 | REN(pS)FHSSI | 4 | Q96Q89 | Kinesin-like protein KIF20B | - | + | - | - | - |
| 129 | REL(pS)PRAPGP | 4 | Q99558 | Mitogen-activated protein kinase kinase kinase 14 | - | + | - | - | - |
| 130 | R(pS)YSRLETL | 2 | Q96FF9 | Sororin | - | + | - | - | - |
| 131 | KEH(pS)LQVSP | 4 | O95235 | Kinesin-like protein KIF20A | - | + | - | - | [157] |
| 132 | YEG(pS)PIKVT | 4 | P06748 | Nucleophosmin | - | + | + | + | [160] |
| 133 | GERAG(pS)PDVL | 6 | Q8NDX6 | Zinc finger protein 740 | - | + | - | - | [159] |
| 134 | KEN(pS)LHRAL | 4 | Q2TAY7 | WD40 repeat-containing protein SMU1 | - | + | - | - | - |
| 135 | AER(pT)PELVEL | 4 | Q9NS56 | E3 ubiquitin-protein ligase Topors | - | + | + | - | - |
| 136 | REKEP(pS)RERDIA | 6 | Q8NI27 | THO complex subunit 2 | - | + | + | - | [160] |
| 137 | KER(pS)GVSLA | 4 | P10412 | Histone H1.4 | - | + | - | - | [159] |
| 138 | FEKPE(pS)PRA | 6 | Q92538 | Golgi-specific brefeldin A-resistance guanine nucleotide exchange factor 1 | - | + | - | - | [159] |

| | | | | | | | | | |
|-----|----------------|---|--------|---|---|---|---|---|-------|
| 139 | REE(pS)PLRIKM | 4 | Q9NYF8 | Bcl-2-associated transcription factor 1 | - | + | - | - | [160] |
| 140 | SEV(pS)KIVTV | 4 | Q562F6 | Shugoshin-like 2 | - | + | - | + | [175] |
| 141 | REG(pS)FRVT | 4 | P49757 | Protein numb homolog | - | + | + | - | - |
| 142 | AEGPPKRV(pS)L | 9 | Q14160 | Protein scribble homolog | - | + | - | - | [160] |
| 143 | REIE(pS)SPQYRL | 5 | Q8WWM7 | Ataxin-2-like protein | - | + | - | + | - |
| 144 | REG(pS)FHSNDL | 4 | Q96F15 | GTPase IMAP family member 5 | - | + | + | - | - |
| 145 | AEK(pS)PQLGL | 4 | Q9P2S5 | WD repeat-containing protein WRAP73 | - | + | + | + | - |
| 146 | KE(pS)TLHLVL | 3 | P0CG48 | Polyubiquitin-C | - | + | - | + | - |
| 147 | G(pS)PIKVTL | 2 | P06748 | Nucleophosmin | - | + | - | - | [160] |
| 148 | VEMP(pS)PKSL | 5 | Q8TBP0 | TBC1 domain family member 16 | - | + | - | - | - |

Tabla 7. Ligandos fosforilados asociados a HLA-B*40. N: número de ligando, P: posición fosforilada, NA: número de acceso en Uniprot, CID: péptido identificado por CID, EThcD: péptido identificado por EThcD, Syn: secuencia confirmada por fragmentación del péptido sintético correspondiente, no-P: identificación del péptido no fosforilado correspondiente en el repertorio presentado por B*40, Ref: sitio de fosforilación reportado previamente en Uniprot y/o HPRD

IV.2.2-El fosfopéptidoma asociado a HLA-B*40 está enriquecido en ligandos derivados de proteínas implicadas en mitosis y regulación del ciclo celular

Con el fin de identificar posibles asociaciones funcionales entre las proteínas parentales de los fosfopéptidos presentados por HLA-B*40, se efectuó un análisis de enriquecimiento usando STRING [176]. Se observó que un número significativo de proteínas –en torno a 30– estaban asociadas a términos GO relacionados con mitosis y ciclo celular (Tabla 8). Dado que las células C1R derivan de una línea linfoblastoide transformada con el virus de Epstein-Barr y sometida a mutagénesis para eliminar la expresión de moléculas de clase I endógenas, nos preguntamos si el patrón de fosforilación de este modelo biológico, presumiblemente aberrante, podría explicar esta observación. Para evaluar esta posibilidad, llevamos a cabo un análisis exhaustivo del fosfoproteoma de la línea celular C1R-B*40. Se extrajo el proteoma total de estas células y tras una digestión con tripsina se procedió al aislamiento de los péptidos fosforilados mediante IMAC-Ti⁴⁺ y a su identificación mediante LC-MS/MS. Este análisis permitió la identificación de 10.153 fosfopéptidos trípticos únicos (FDR ≤ 1% a nivel de péptido) derivados de 2369 proteínas (material suplementario S3 incluido en el CD adjunto).

Cuando este grupo de proteínas se usó como *background* en el análisis mediante STRING, no se observó ningún enriquecimiento de términos relacionados con mitosis o ciclo celular (Tabla 8), indicando que la sobrerrepresentación de ligandos fosforilados derivados de este tipo de proteínas es una consecuencia de la abundancia de estas en el fosfoproteoma de la línea celular C1R-B*40.

IV.2.3-HLA-B*40 puede unir eficientemente péptidos fosforilados en P2

Entre los 148 fosfopéptidos compilados, se encontraban 7 con un residuo de pSer en P2 (N = 57, 65, 81, 106, 111, 130, 147 en la Tabla 7). Como se menciona más arriba, esta excepción al motivo de

anclaje que habíamos descrito para este alotipo (preferencia por Glu o Asp en P2) podría ser explicada por un mecanismo de mimetismo molecular entre los residuos de Glu y pSer (Figura 18), donde el grupo fosfato aportaría el carácter ácido para permitir la interacción con los residuos de la subcavidad B.

| GO id | Término GO | Nº de genes | FDR (Genoma) ^a | FDR (Fosfoproteoma) ^b |
|------------|---|-------------|---------------------------|----------------------------------|
| GO:1903047 | mitotic cell cycle process | 26 | 1,04·10 ⁻⁰⁹ | 1,00 |
| GO:0000278 | mitotic cell cycle | 27 | 1,12·10 ⁻⁰⁹ | 1,00 |
| GO:0007049 | cell cycle | 32 | 9,22·10 ⁻⁰⁹ | 1,00 |
| GO:0022402 | cell cycle process | 28 | 9,22·10 ⁻⁰⁹ | 1,00 |
| GO:0051301 | cell division | 20 | 4,50·10 ⁻⁰⁸ | 1,00 |
| GO:0007067 | mitotic nuclear division | 17 | 4,60·10 ⁻⁰⁸ | 1,00 |
| GO:0000280 | nuclear division | 18 | 1,61·10 ⁻⁰⁷ | 1,00 |
| GO:0048285 | organelle fission | 18 | 3,32·10 ⁻⁰⁷ | 1,00 |
| GO:0000236 | mitotic prometaphase | 10 | 1,47·10 ⁻⁰⁶ | 1,00 |
| GO:0000087 | mitotic M phase | 13 | 2,48·10 ⁻⁰⁶ | 1,00 |
| GO:0000279 | M phase | 13 | 2,48·10 ⁻⁰⁶ | 1,00 |
| GO:0022403 | cell cycle phase | 14 | 6,92·10 ⁻⁰⁶ | 1,00 |
| GO:0098763 | mitotic cell cycle phase | 14 | 6,92·10 ⁻⁰⁶ | 1,00 |
| GO:0000090 | mitotic anaphase | 11 | 6,92·10 ⁻⁰⁶ | 1,00 |
| GO:0044848 | biological phase | 14 | 6,92·10 ⁻⁰⁶ | 1,00 |
| GO:0044772 | mitotic cell cycle phase transition | 13 | 5,90·10 ⁻⁰⁵ | 1,00 |
| GO:0044770 | cell cycle phase transition | 13 | 6,52·10 ⁻⁰⁵ | 1,00 |
| GO:1902589 | single-organism organelle organization | 29 | 2,71·10 ⁻⁰⁴ | 1,00 |
| GO:0051726 | regulation of cell cycle | 19 | 5,97·10 ⁻⁰⁴ | 1,00 |
| GO:0051983 | regulation of chromosome segregation | 6 | 1,38·10 ⁻⁰³ | 1,00 |
| GO:0007059 | chromosome segregation | 9 | 1,88·10 ⁻⁰³ | 1,00 |
| GO:0045786 | negative regulation of cell cycle | 13 | 1,90·10 ⁻⁰³ | 1,00 |
| GO:0007264 | small GTPase mediated signal transduction | 14 | 5,40·10 ⁻⁰³ | 1,00 |

Tabla 8. Análisis de enriquecimiento funcional sobre las proteínas que dan origen a los fosfopéptidos unidos a HLA-B*40. ^aFDR usando el genoma humano como *background*. ^bFDR usando el fosfoproteoma de la línea C1R-B*40 como *background*).

Para determinar si los péptidos identificados con pSer en P2 eran realmente ligandos de HLA-B*40, decidimos llevar a cabo una serie de ensayos de unión a esta molécula. Para ello, se sintetizaron 2 de los 7 péptidos fosforilados en P2, S(pS)YGNIRAV y G(pS)FSRFYSL (N = 57 y 65 en la Tabla 7), junto con dos variantes en las que el residuo de fosfoserina se sustituyó por ácido glutámico: SEYGNIRAV y GEFSRFYSL. Dichos péptidos se utilizaron en un ensayo competitivo de unión a HLA-B*40 basado en la disociación por tratamiento ácido de esta molécula en la superficie celular y su posterior replegamiento en presencia de los péptidos a ensayar y un péptido de referencia fluorescente (ver materiales y métodos).

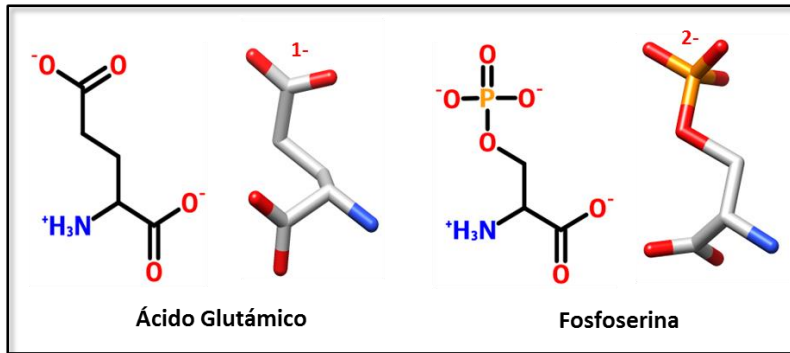


Figura 18. Similitud estructural entre el ácido glutámico y la fosfoserina. Se usa el código de color CPK

Como se muestra en la Figura 19, los dos ligandos fosforilados se unieron a HLA-B*40 con una alta afinidad. El valor de IC_{50} –la concentración de péptido que produce una inhibición del 50 % en la unión del péptido de referencia– fue de 1,8 y 0,3 μ M, respectivamente (Figura 19). Del mismo modo, los péptidos mutantes SEYGNIRAV y GEFSRFYSL mostraron afinidades muy ligeramente superiores a la de sus homólogos fosforilados (IC_{50} de 1,3 y 0,2 μ M, respectivamente). Estos resultados confirman los péptidos S(pS)YGNIRAV y G(pS)FSRFYSL como ligandos endógenos de HLA-B*40 y demuestran que la sustitución de Glu por pSer tiene un efecto mínimo o nulo sobre afinidad de unión a B*40.

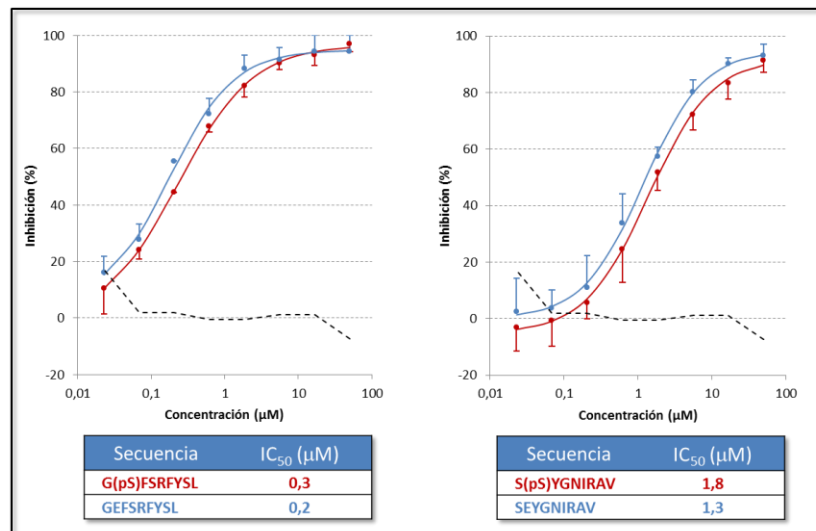


Figura 19. Ensayos de unión competitivos de ligandos con pSer (rojo) o Glu (azul) en P2 frente a un ligando fluorescente de referencia. Se muestra el porcentaje de máxima fluorescencia frente a la concentración de los péptidos. Se representa la media (puntos) y la desviación estándar de tres ensayos independientes. Estos valores se ajustaron a una curva sigmoideal (línea continua) para permitir la estimación de los valores de IC_{50} (concentración de péptido que produce una inhibición del 50% en la unión del péptido de referencia). La línea de color negra discontinua corresponde al péptido GRIDKPIIK (ligando de HLA-B*27) utilizado como control negativo de unión a B*40.

Como control negativo se incluyó el péptido GRIDKPILK, un ligando endógeno de HLA-B*27 que carece de los motivos de unión a HLA-B*40. Como se muestra en la Figura 19, este ligando no fue capaz de inhibir la unión del péptido de referencia demostrando que el resto de curvas de inhibición son consecuencia de la interacción específica con B*40 de los péptidos ensayados.

IV.2.4-La interacción de los residuos de fosfoserina o ácido glutámico en P2 con la subcavidad B es muy similar estructuralmente

Para probar la hipótesis de que la unión a B*40 de péptidos fosforilados en P2 se basa en el mimetismo molecular entre los residuos de Glu y pSer, decidimos cristalizar el complejo entre esta molécula de clase I y los péptidos G(pS)FSRFYSL y GEFSRFYSL. Los datos relativos al procesamiento estadístico y el refinamiento de la estructura del primer complejo se resumen en la Tabla 6. La estructura de este complejo se resolvió a una resolución de 2,5 Å mediante reemplazo molecular utilizando como molde la estructura de HLA-B*41:04 (PDB: 3LN5) que comparte con HLA-B*40:02 un 97,1% de identidad en los dominios α_1 , α_2 y α_3 . Respecto al complejo HLA-B*40-GEFSRFYSL, aún estamos trabajando en la resolución de la estructura. No obstante, los datos de que disponemos actualmente permiten determinar con precisión el plegamiento de los dominios α_1 y α_2 y la conformación del ligando peptídico.

En ambos casos, los péptidos se unieron en conformación extendida con los residuos principales de anclaje P2-pSer/P2-Glu y P9-Leu ubicados respectivamente en el interior de las subcavidades B y F y con P1-Gly interaccionando con residuos de la subcavidad A (Figura 20). P4-Ser, P5-Arg, P7-Tyr y P8-Ser quedaron expuestos hacia el solvente y accesibles a la interacción con el TCR. Por su parte el residuo P6-Phe se orientó hacia el interior del surco de unión. Finalmente, la cadena lateral de P3-Phe se situó cerca de la hélice α_2 siendo parcialmente accesible. Solo se detectaron diferencias en la posición de la cadena lateral de P5-Arg, probablemente debidas a la elevada movilidad de esta. Confirmando nuestra hipótesis, la red de interacciones que estabilizaba el residuo de pSer del ligando G(pS)FSRFYSL en P2 era prácticamente idéntica a la que se establecía entre el Glu y la subcavidad B en el caso del péptido GEFSRFYSL (Figura 20). Esta incluía contactos directos del grupo fosfato (G(pS)FSRFYSL) o del grupo carboxílico (GEFSRFYSL) en P2 con las cadenas laterales de los residuos His9, Lys45 y Tyr99 (Figura 20). Por último, en ambos casos, el grupo α -amino del residuo en P2 interaccionaba mediante un puente de hidrógeno con el grupo carboxilo de la cadena lateral de Glu63. En general, la interacción del residuo en P2 en estos dos complejos era bastante similar a la que se observaba en la descrita anteriormente para el ligando REFSKEPEL (Figura 15)

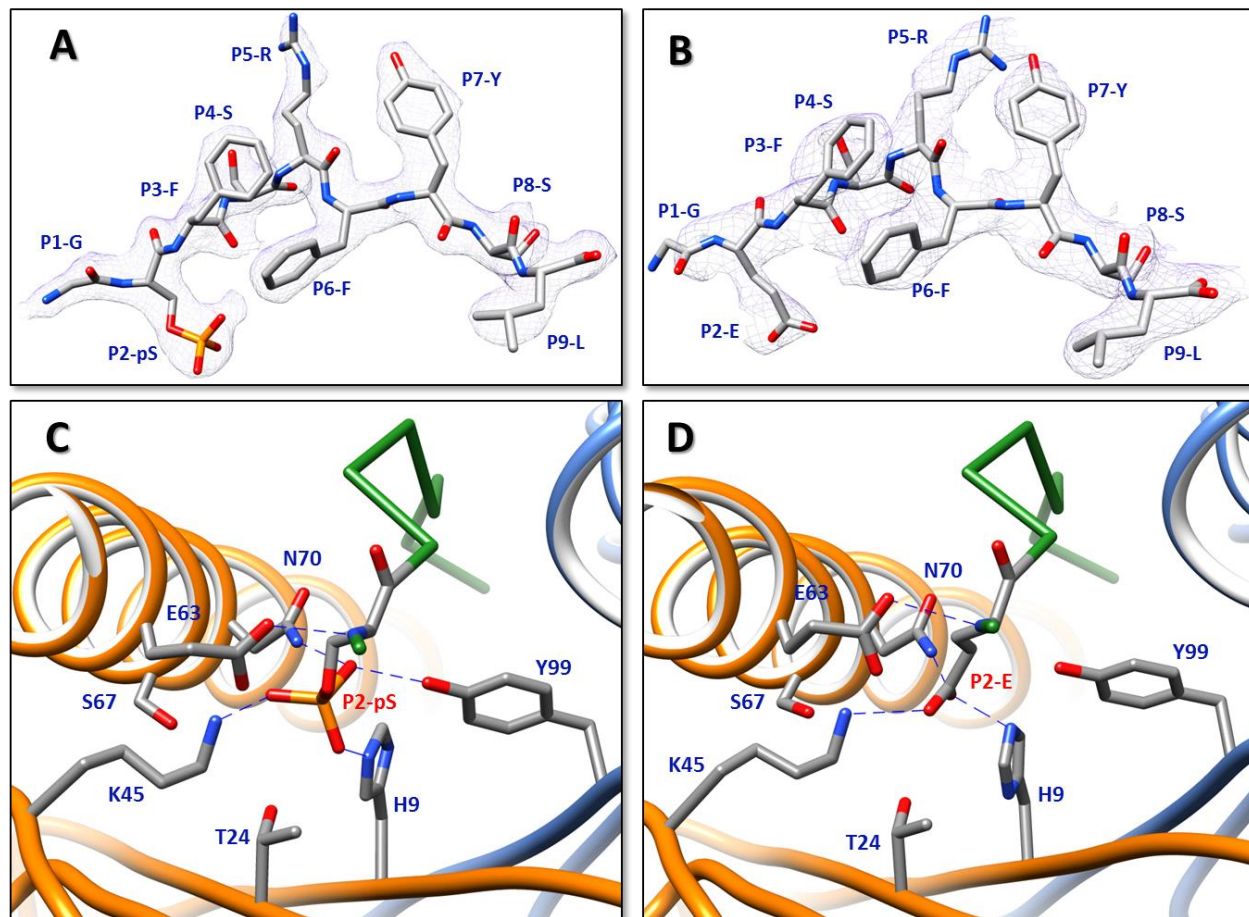


Figura 20. Estructura cristalográfica de HLA-B*40 en complejo con el péptido G(pS)FSRFYSL (A y C) o GEFSRFYSL (B y D). (A) Estructura adoptada por el péptido G(pS)FSRFYSL en el surco de unión de B*40. (B) Estructura adoptada por el péptido GEFSRFYSL en el surco de unión de B*40. (C) Interacción de P2-pSer con la subcavidad B de HLA-B*40. (D) Interacción de P2-E con la subcavidad B de HLA-B*40. Las interacciones se muestran con líneas discontinuas de color azul. Los residuos de aminoácidos correspondientes a la subcavidad B de la molécula de clase I se indican en el código de una letra y de color azul. El residuo en P2 del cada ligando se indica en color rojo. La cadena principal del péptido se muestra de color verde.

IV.2.5-Los fosfopéptidos unidos a HLA-B*40 se encuentran fosforilados mayoritariamente en P4 y tienden a presentar residuos básicos en P1

Un análisis detallado de las 148 secuencias indicadas en la Tabla 7, reveló varias características reseñable de los ligandos fosforilados asociados a HLA-B*40. En primer lugar, la fosforilación se encontraba en P4 en 101 péptidos (68%, Figura 21A). De estos, 65 (64%) presentaban a su vez un residuo de Pro en P5 mientras que en las restantes 36 secuencias (36%), P5 se encontraba enriquecida en Leu, Phe, Val, Ile y Met (Figura 21B). Es más, 56 de los ligandos fosforilados en P4 (55%) tenían también un residuo básico –fundamentalmente Arg– en el extremo N-terminal (Figura 21B). Un análisis

similar sobre las secuencias no fosforiladas del repertorio peptídico unido a HLA-B*40 evidenció que estas características eran específicas de los ligandos fosforilados. Las características generales del peptidoma no fosforilado venían definidas por el motivo de unión a este alotipo: Glu, Asp o pSer en P2 y Met, Phe o residuos alifáticos en el extremo C-terminal (Figura 21C). Curiosamente, la preferencia de HLA-B*40 por péptidos fosforilados en P4 y con residuos cargados positivamente en P1 se observa también en el contexto de otras moléculas de clase I como HLA-A*02 [4, 106], -B*27 [97] y -B*07 [5, 97, 106] (Tabla 1).

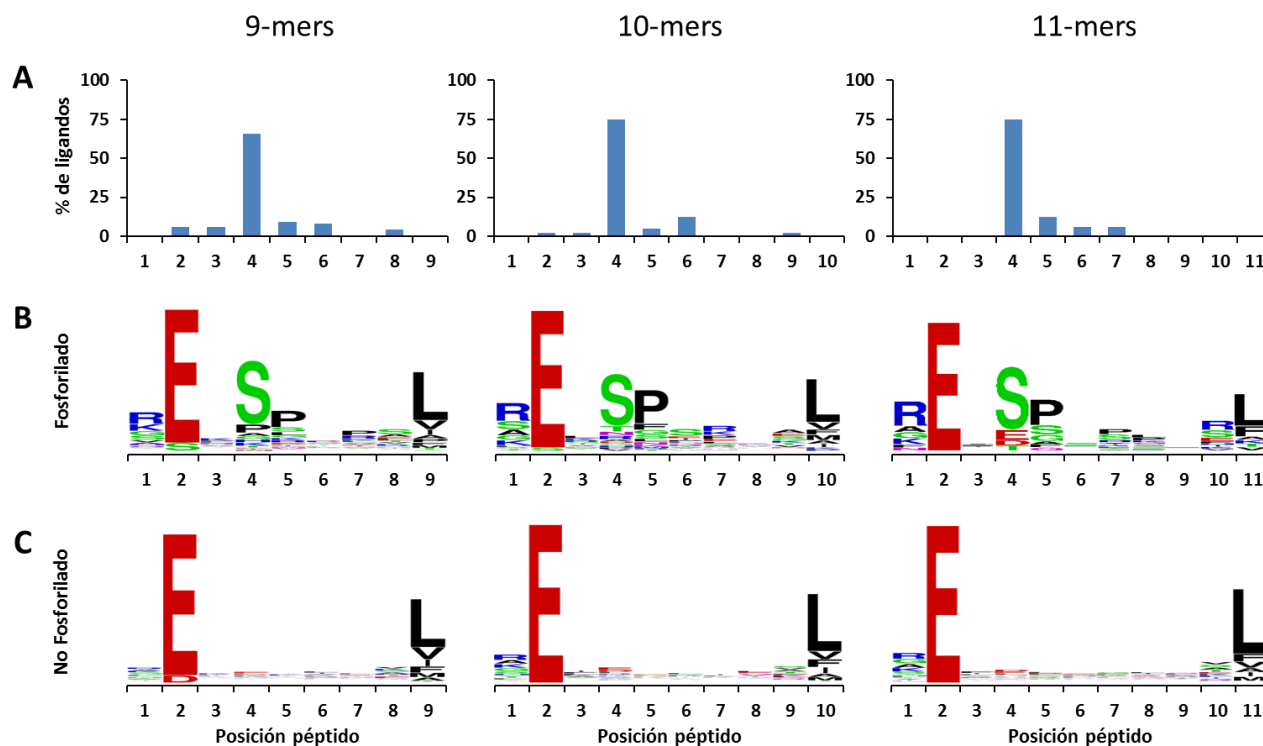


Figura 21. Análisis de las secuencias fosforiladas y no fosforiladas del peptidoma de HLA-B*40. (A) Frecuencia de fosforilación en cada posición peptídica de los ligandos fosforilados de 9, 10 y 11 residuos de longitud. (B y C) Uso de residuos en cada posición peptídica de los ligandos fosforilados (B) o no fosforilados (C).

IV.2.5.1-El motivo de fosforilación de los péptidos asociados a HLA-B*40 no está determinado por la especificidad del quinoma celular

A continuación, nos preguntamos si el motivo de fosforilación común observado en los fosfopeptidomas asociados a HLA-B*40 y a otras moléculas de clase I podría estar determinando la especificidad de una o varias quinasas particularmente abundantes o eficientes. En este sentido, algunas protein-quinasas como Plk1 reconocen secuencias peptídicas que solapan parcialmente con el motivo de

fosforilación observado en los fosfopéptidos presentados por HLA-B*40 [177] lo que podría explicar el motivo de fosforilación observado.

Para ello, a partir de los 10.153 fosfopéptidos trípticos identificados en el análisis del fosfoproteoma de la línea C1R-B*40 descrito más arriba, se llevó a cabo una predicción *in silico* de secuencias fosforiladas de entre 9 y 11 residuos que tuvieran los motivos de unión a HLA-B*40 (Glu en P2 y Leu, Ile, Val, Met o Phe en el extremo C-terminal). Como resultado, se compilaron un total de 2017 secuencias teóricas que cumplían con estas restricciones, no observándose ningún enriquecimiento de pSer/pThr en P4. Por el contrario, la fosforilación se encontraba distribuida equitativamente entre todas las posiciones peptídicas (Figura 22). Es más, entre las secuencias fosforiladas en P4 (284 ligandos), solo 50 (18%) presentaban residuos básicos en P1, muy lejos de la frecuencia observada en el fosfopeptidoma endógeno. Por su parte, observamos cierta sobrerrepresentación de residuos de Pro en P5 (67 ligandos, 24%). Estos resultados sugieren que: 1) ni la fosforilación preferente de los residuos en P4 ni el motivo Arg1/Lys1 vienen impuestos por la especificidad del quinoma celular, si no que probablemente reflejan una restricción estructural impuesta por el surco de unión de HLA-B*40 y 2) la elevada frecuencia de Pro en P5 está parcialmente determinada por quinasas que reconocen sitios TP o SP como se ha propuesto para HLA-B*07 [106].

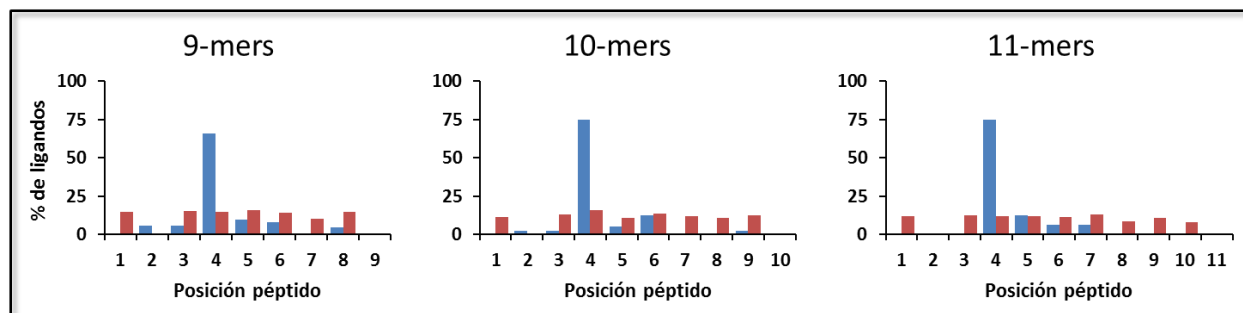


Figura 22. Frecuencia de fosforilación en cada posición peptídica en los fosfopéptidos presentados por HLA-B*40 (azul) y las secuencias predichas a partir del análisis bioinformático (rojo).

IV.2.5.2-La fosforilación en P4 disminuye ligeramente la afinidad de unión a HLA-B*40

La alta prevalencia de ligandos fosforilados en P4 podría ser debida a un incremento en la estabilidad del complejo MHC-péptido causada por la presencia del grupo fosfato. Para evaluar esta hipótesis decidimos llevar a cabo una serie de ensayos de unión competitiva a HLA-B*40 con 7 ligandos endógenos fosforilados y sus homólogos no fosforilados (Figura 23A).

La mayoría de los fosfopéptidos y de los ligandos no fosforilados mostraron valores de IC_{50} en el rango nanomolar o bajo micromolar (0,3-3,8 μ M), correspondiente a interacciones de afinidad media o alta (Figura 23B). Por su parte, el ligando REASPSRLS arrojó unos valores de IC_{50} de 9,9 μ M, 14,6 μ M para las especies no fosforilada y fosforilada, respectivamente, correspondientes a una afinidad baja por HLA-B*40. Llamativamente, en todos los ensayos de unión, la IC_{50} de los fosfopéptidos fue un tanto mayor – unas 2 veces, en promedio– que la de los ligandos no fosforilados. Este hecho indica que, lejos de favorecer la unión, la fosforilación en P4 tiene un efecto ligeramente negativo sobre la estabilidad de la interacción entre el péptido y HLA-B*40.

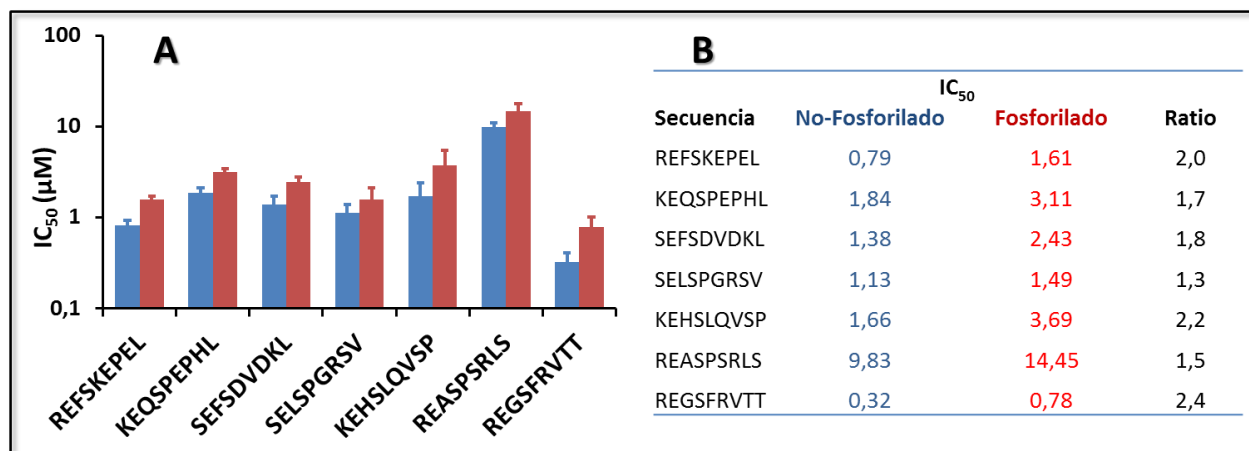


Figura 23. Ensayos de unión competitiva a B*40 de siete ligandos endógenos fosforilados o no en P4. (A) Se representan los valores de IC_{50} de los ligandos fosforilados (rojo) y no fosforilados (azul). Se muestra la media y la desviación estándar de tres experimentos independientes. (B) Valores de IC_{50} de los péptidos ensayados y ratio de los IC_{50} de los péptidos fosforilados frente a los no fosforilados.

IV.2.5.3-Efecto de la fosforilación en diferentes posiciones del péptido en los ensayos de unión a HLA-B*40

Teniendo en cuenta que la fosforilación en P4 aparentemente no mejora la unión a HLA-B*40, nos preguntamos si la alta frecuencia de este motivo era el resultado de una selección negativa que operaba en el resto de posiciones peptídicas. En consecuencia, se evaluó la unión de un conjunto de análogos basado en secuencias de poliglicina a HLA-B*40. Los mismos fueron sintetizados preservando los motivos fundamentales de unión a este alelo en P2 y P Ω y variando la posición del residuo de pSer a lo largo del péptido (Figura 24A). Como se muestra en la Figura 24B, las mayores eficiencias de unión a HLA-B*40 se observaron cuando la pSer se ubicó en P4 (0,5 μ M) y P8 (0,6 μ M), mientras que la fosforilación en P5, P6 y P7 se asociaba a una disminución de afinidad moderada (1,9 μ M, 1,8 μ M y 1,1

μM , respectivamente). Por último, la fosforilación en P3 producía un incremento de la IC_{50} (25,1 μM) que refleja un enorme efecto negativo sobre la unión a HLA-B*40.

Para evaluar el efecto de la presencia de un residuo cargado positivamente en el extremo N-terminal, se realizó un ensayo muy similar al anterior con un conjunto de análogos de poliglicina que además de las restricciones antes mencionadas, presentaban un residuo de Arg en P1. Independientemente de la posición fosforilada, la presencia de P1-Arg mejoró la afinidad de unión para todos los candidatos estudiados (Figura 24). Dicha mejora fue especialmente evidente cuando la fosforilación afectaba a P3. En este caso, la presencia de un aminoácido básico en P1 provocó una disminución de 3 veces en el valor de IC_{50} (de 25,1 a 7,9 μM). Este hecho explicaría la alta frecuencia de residuos básicos en el extremo N-terminal observada entre los fosfopéptidos presentados por HLA-B*40.

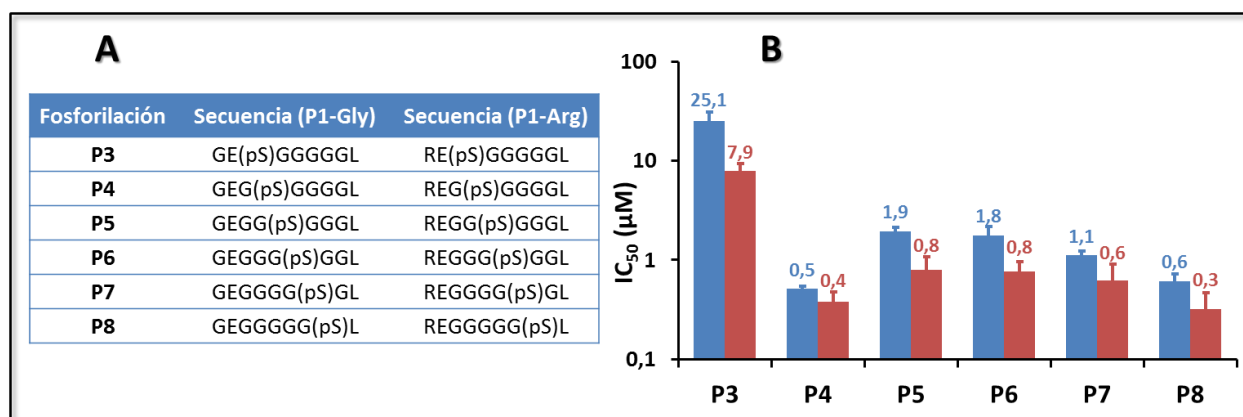


Figura 24. Ensayos de unión competitiva de análogos de poliglicina a HLA-B*40. (A) Secuencias de los péptidos ensayados. (B) Se representan los valores de IC_{50} de los péptidos con Arg1 (rojo) o Gly1 (azul) en función de la posición de la fosforilación. Se muestra la media y la desviación estándar de tres experimentos independientes.

IV.2.5.4-Estructuras de HLA-B*40 en complejo con un ligando endógeno fosforilado en P4 y su homólogo no fosforilado

Para determinar las bases estructurales de la presentación de péptidos fosforilados en P4 por HLA-B*40 decidimos resolver la estructura cristalina del complejo entre esta molécula y el fosfopéptido REF(pS)KEPEL (PDB: 5IEH) y compararla con la descrita más arriba para el complejo B*40-REFSKEPEL (PDB: 5IEK). La estructura se resolvió a una resolución de 1,5 Å mediante reemplazo molecular utilizando como molde HLA-B*40:02 asociado con el ligando G(pS)FSRFYSL. Los parámetros cristalográficos y estadísticas del refinamiento se resumen en la Tabla 6.

Ambas estructuras presentaban un plegamiento muy similar. Los dos ligandos adoptaron la típica conformación extendida con los extremos N y C-terminales anclados en el surco de unión (Figura 25A y

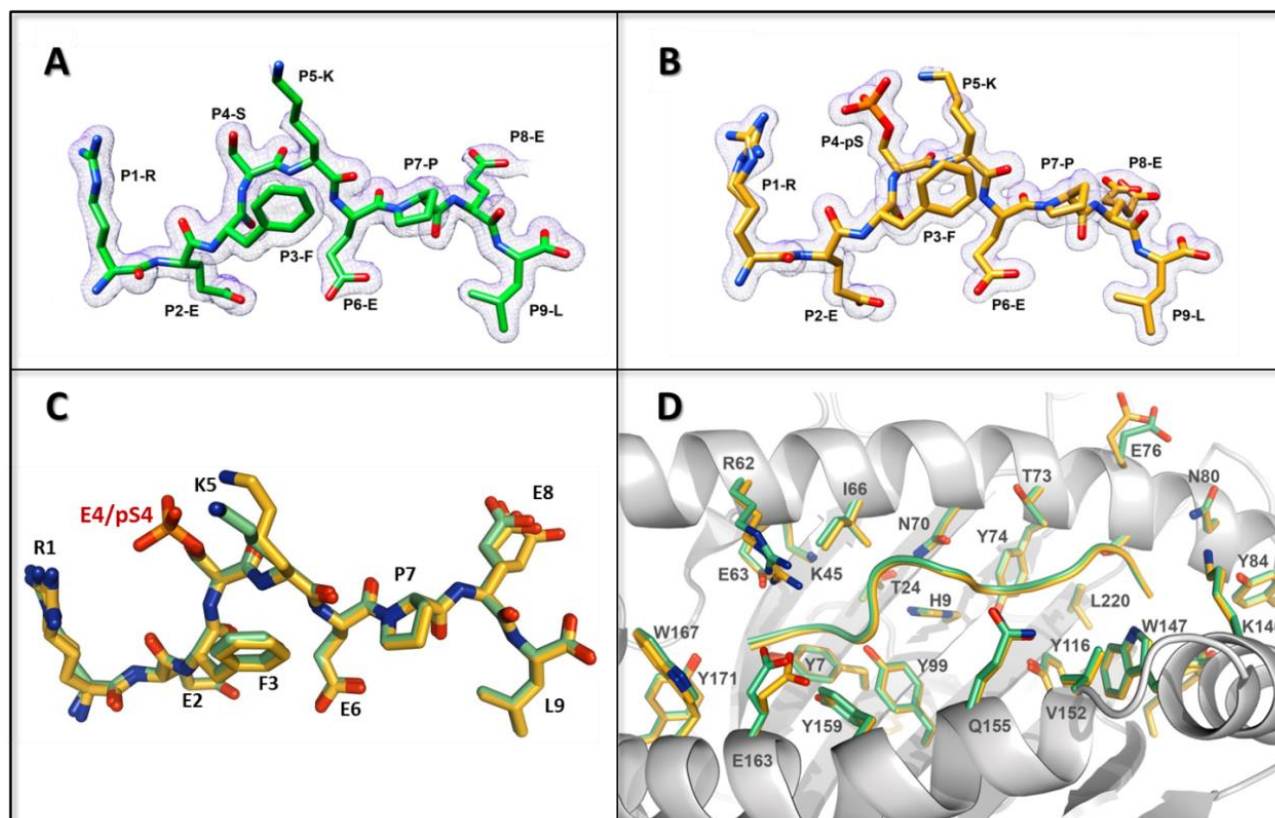


Figura 25. Estructura cristalográfica de HLA-B*40 en complejo con los ligandos REFSKEPEL (verde) o REF(pS)KEPEL (amarillo). (A y B) Estructura de los péptidos en el surco de unión de HLA-B*40. Se muestra el mapa de densidad electrónica (2Fo-Fc) a 1σ como una malla azul. (C) Superposición de ambos ligandos. (D) Superposición de ambas estructuras mostrando los residuos de la cadena pesada implicados en la interacción con el péptido.

B). Como era esperable, los residuos P2-Glu y P9-Leu se encontraban ubicados en las subcavidades B y F, respectivamente. P6-Glu se dispuso apuntando hacia el suelo del surco de unión mientras que P1-Arg, P4-pSer/P4-Ser, P5-Lys y P8-Glu se orientaban hacia el solvente, siendo potencialmente accesibles al TCR. A pesar de esta estructura análoga, se detectaron pequeñas variaciones en las posiciones de los residuos en P1, P5 y P8 que a su vez inducían la reorientación de las cadenas laterales de Arg62, Glu76 y Glu163. Tanto P1-Arg como P8-Glu mostraban dos conformaciones alternativas en el péptido fosforilado (Figura 25C y D) contribuyendo a las mencionadas diferencias en la posición de los residuos Arg62, Glu76 y Glu163 entre los dos complejos. En lo relativo a P5-Lys, la baja densidad electrónica detectada en torno a su cadena lateral es probablemente indicativa de una alta flexibilidad. No obstante, en el péptido REF(pS)KEPEL, el grupo ϵ -amino de este residuo tendía a orientarse hacia el grupo fosfato del

residuo P4-pSer, probablemente debido a la influencia de la carga negativa de este último (Figura 25B y C). En el ligando fosforilado, P4-pSer contactaba directamente con Arg62, P5-Lys y a través de 2 moléculas de agua con Glu163 e interaccionaba mediante una molécula de agua con P1-Arg (Figura 26B). Por el contrario, en el caso del péptido no fosforilado, P4-Ser no participaba en ninguna interacción significativa y la falta de la carga negativa del grupo fosfato causaba un ligero cambio en la

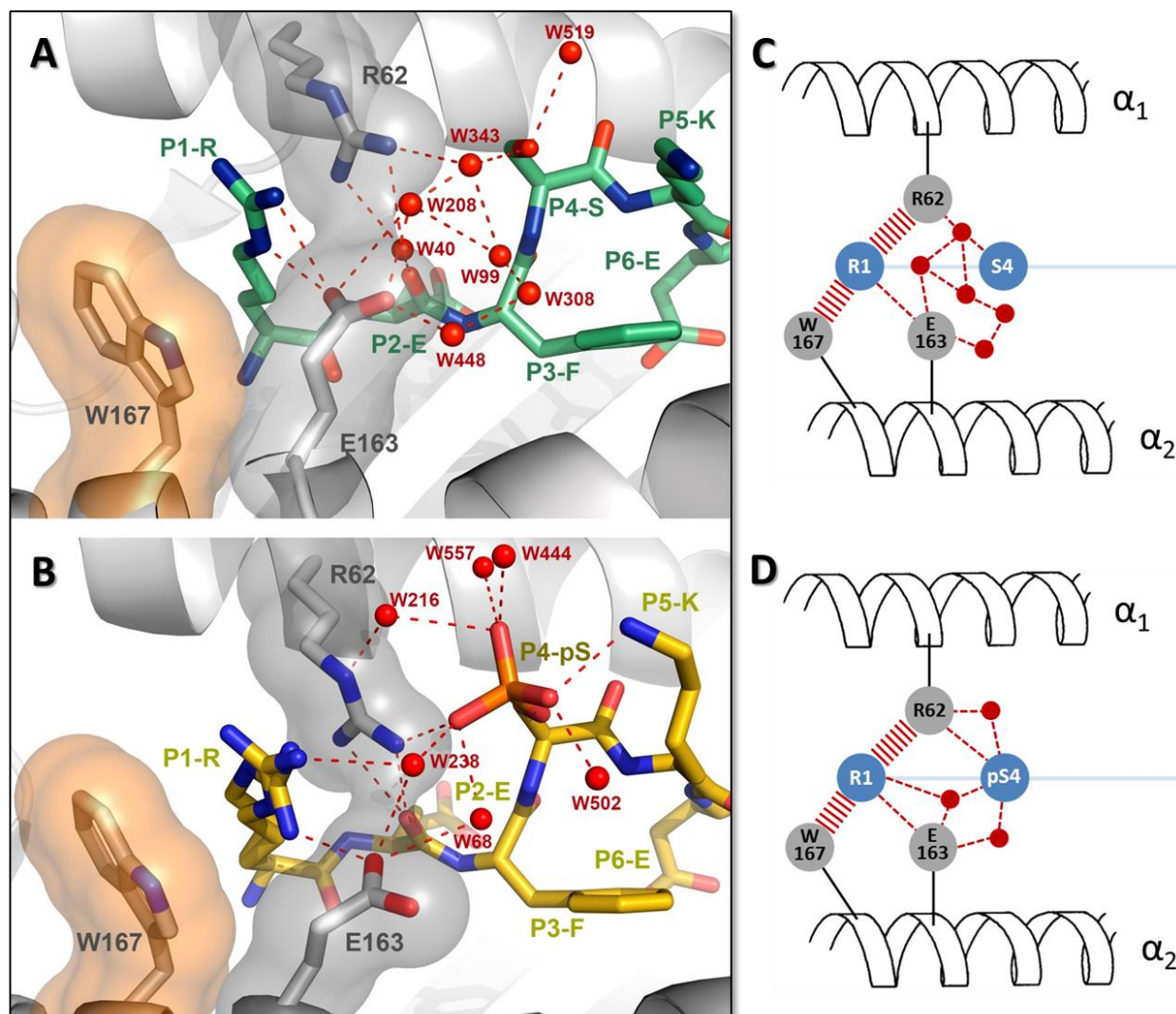


Figura 26. Estructura cristalográfica de los péptidos REFSKEPEL (A y C) y REF(pS)KEPEL (B y D) en complejo con HLA-B*40. (A) Interacción de los residuos P1 y P4 del péptidos REFSKEPEL (verde) con el surco de unión de B*40. Las moléculas de agua se representan como esferas de color rojo. Se muestra la superficie de los residuos pertenecientes a la cadena pesada que interaccionan con las posiciones P1 y P4 en ambos ligandos. Las líneas discontinuas de color rojo representan puentes de hidrógeno e interacciones electrostáticas. (B) Misma representación para el péptido REF(pS)KEPEL (amarillo). (C y D) Representación esquemática de las interacciones mostradas en A y B, respectivamente.

orientación de Arg62 y Glu163 (Figura 25D y 26A). En ambos casos, P1-Arg se estabilizó en el interior de la subcavidad A mediante 3 interacciones principales: 1) apilamiento π - π de su grupo guanidinio con el de la Arg62, 2) interacciones hidrofóbicas entre la región alifática de su cadena lateral y el anillo aromático del Trp167 y 3) un puente salino establecido entre su grupo guanidinio y el grupo carboxilo del residuo Glu163 (Figura 26A y B). Esta forma de unión proporciona una explicación para el aumento de la afinidad por HLA-B*40 observada para fosfopéptidos con residuos básicos en P1.

IV.3-El fosfopeptidoma asociado a HLA-B*39 también se encuentra enriquecido en ligandos fosforilados en P4

Una posible explicación para la conservación del motivo de fosforilación en los fosfopeptidomas asociados a HLA-B*40 y a otras moléculas como HLA-B*07 o -B*27 es la conservación del residuo Arg62 –clave para la estabilización del grupo fosfato en P4– en la mayoría de las proteínas codificadas en el locus HLA-B. Para averiguar si esta correlación podría extenderse a otras moléculas de clase I, el peptidoma presentado por HLA-B*39:01 en células C1R se purificó y se sometió a un enriquecimiento de fosfopéptidos como se describe en materiales y métodos. El peptidoma y el fosfopeptidoma de este alotipo fueron entonces analizados por LC-MS/MS mediante fragmentación CID. Se identificaron un total de 1157 péptidos únicos a una FDR \leq 1% a nivel de péptido (material suplementario S4 incluido en el CD adjunto). De ellos, 961 (83%) presentaron His o Arg en P2, siendo este el motivo de unión canónico de este alelo [34], mientras que 16 (1%) y 54 (5%) mostraban los motivos de unión a HLA-B*35 y HLA-C*04 [154-156]. Las 126 secuencias restantes (11%) no fueron asignadas a ningún alotipo. No obstante, observamos en este último grupo un total de 55 péptidos (44%) que presentaban Gln en P2, lo que sugiere que HLA-B*39 podría unir péptidos con este motivo de anclaje subóptimo.

El análisis por LC-MS/MS de las fracciones enriquecidas en fosfopéptidos permitió la caracterización de 31 ligandos fosforilados con His o Arg en P2 derivados de 30 proteínas (Tabla 9). Debido a que la adquisición de los espectros MS2 se llevó a cabo usando un esquema de fragmentación CID (que no es el ideal para la identificación de fosfopéptidos) todas estas secuencias se confirmaron mediante fragmentación del péptido sintético correspondiente (Anexo 1.3). Al igual que ocurría en el caso de HLA-B*40, las proteínas parentales de estos fosfopéptidos tendían a estar implicadas en mitosis y regulación del ciclo celular, reflejando las particularidades del fosfoproteoma de la línea C1R (Tabla 10).

| N | Secuencia | P | NA | Nombre de la Proteína | Syn | Ref |
|----|----------------|---|--------|---|-----|-------|
| 1 | NHV(pT)PPNVSL | 4 | P17275 | Transcription factor jun-B | + | - |
| 2 | RRH(pS)ASNLHAL | 4 | P47974 | Zinc finger protein 36, C3H1 type-like 2 | + | [157] |
| 3 | HH(pS)PGQTPTL | 3 | Q55W79 | Centrosomal protein of 170 kDa | + | [157] |
| 4 | HRA(pS)PNPVAM | 4 | Q9H0E3 | Histone deacetylase complex subunit SAP130 | + | [159] |
| 5 | GHG(pS)PPFSL | 4 | P19484 | Transcription factor EB | + | [159] |
| 6 | NRH(pS)LPFSL | 4 | P22681 | E3 ubiquitin-protein ligase CBL | + | [178] |
| 7 | RHK(pS)DSISL | 4 | Q00987 | E3 ubiquitin-protein ligase Mdm2 | + | - |
| 8 | TRM(pS)TVSEL | 4 | Q99661 | Kinesin-like protein KIF2C | + | [161] |
| 9 | TRL(pS)PAKIVL | 4 | Q8TEK3 | Histone-lysine N-methyltransferase, H3 lysine-79 specific | + | [157] |
| 10 | LRS(pS)LPISL | 4 | P19532 | Transcription factor E3 | + | - |
| 11 | TRL(pS)GGQVL | 4 | Q8NCN4 | E3 ubiquitin-protein ligase RNF169 | + | [162] |
| 12 | HHLG(pS)PSRL | 5 | Q9H8M2 | Bromodomain-containing protein 9 | + | [160] |
| 13 | VRQ(pS)PGPAL | 4 | P85037 | Forkhead box protein K1 | + | [172] |
| 14 | SRG(pS)PLPVL | 4 | P24864 | G1/S-specific cyclin-E1 | + | [162] |
| 15 | SRS(pS)PLKL | 4 | O75410 | Transforming acidic coiled-coil-containing protein 1 | + | - |
| 16 | THL(pS)QSSNL | 4 | Q06413 | Myocyte-specific enhancer factor 2C | + | - |
| 17 | RRY(pS)DLTTL | 4 | P25205 | DNA replication licensing factor MCM3 | + | [157] |
| 18 | THIQT(pS)QQAL | 6 | Q6PJP8 | DNA cross-link repair 1A protein | + | - |
| 19 | MRL(pS)RELQL | 4 | Q15051 | IQ calmodulin-binding motif-containing protein 1 | + | - |
| 20 | HHGG(pS)PIHWVL | 5 | Q96MU7 | YTH domain-containing protein 1 | + | [157] |
| 21 | ARA(pS)PEPEL | 4 | Q9BYX4 | Interferon-induced helicase C domain-containing protein 1 | + | [179] |
| 22 | HRAG(pT)LTSL | 5 | Q9UL54 | Serine/threonine-protein kinase TAO2 | + | - |
| 23 | THRPP(pT)PSL | 6 | J3KNZ9 | Probable fibrosin-1 | + | - |
| 24 | NRL(pS)VSSLL | 4 | Q9ULT8 | E3 ubiquitin-protein ligase HECTD1 | + | - |
| 25 | RR(pS)IPIITV | 3 | | | | |
| | RRS(pS)IPIITV | 4 | P33992 | DNA replication licensing factor MCM5 | + | - |
| 26 | FH(pT)TPNTSL | 3 | | | | |
| | FHT(pT)TPNTSL | 4 | O43683 | Mitotic checkpoint serine/threonine-protein kinase BUB1 | + | [159] |
| 27 | RR(pS)SDIISL | 3 | | | | |
| | RRS(pS)SDIISL | 4 | Q9UJX6 | Anaphase-promoting complex subunit 2 | + | - |
| 28 | LR(pS)TGVAL | 3 | | | | |
| | LRS(pT)GVAL | 4 | Q15437 | Protein transport protein Sec23B | + | - |
| 29 | SH(pS)SIHRL | 3 | | | | |
| | SHS(pS)SIHRL | 4 | Q9UL54 | Serine/threonine-protein kinase TAO2 | + | [174] |
| 30 | RR(pS)SIGLRV | 3 | | | | |
| | RRS(pS)IGLRV | 4 | Q96GN5 | Cell division cycle-associated 7-like protein | + | [168] |
| 31 | NR(pS)STRSLL | 3 | | | | |
| | NRS(pS)TRSLL | 4 | Q6ZUJ8 | Phosphoinositide 3-kinase adapter protein 1 | + | - |

Tabla 9. Ligandos fosforilados asociados a HLA-B*40. N: número de ligando, P: posición fosforilada, NA: número de acceso en Uniprot, Syn: secuencia confirmada por fragmentación del péptido sintético correspondiente, Ref: sitio de fosforilación reportado previamente en Uniprot y/o HPRD

La posición del residuo fosforilado se asignó de forma inequívoca en 24 de los 31 ligandos. Estos 24 fosfopéptidos incluían 24 sitios únicos de fosforilación, de los cuales 14 (58%) se encontraban ya

| GO id | Término GO | Nº de genes | FDR (Genoma) ^a | FDR (Fosfoproteoma) ^b |
|------------|----------------------------|-------------|---------------------------|----------------------------------|
| GO:0051301 | cell division | 7 | 2,13·10 ⁻² | 1,00 |
| GO:1903047 | mitotic cell cycle process | 8 | 2,13·10 ⁻² | 1,00 |
| GO:0007049 | cell cycle | 10 | 2,13·10 ⁻² | 1,00 |
| GO:0022402 | cell cycle process | 9 | 2,13·10 ⁻² | 1,00 |
| GO:0006270 | DNA replication initiation | 3 | 2,24·10 ⁻² | 1,00 |
| GO:0000278 | mitotic cell cycle | 8 | 2,48·10 ⁻² | 1,00 |

Tabla 10. Análisis de enriquecimiento funcional sobre las proteínas que dan origen a los fosfopéptidos unidos a HLA-B*39. ^aFDR usando el genoma humano como *background*. ^bFDR usando el fosfoproteoma de la línea C1R-B*40 como *background*.

anotados en las bases de datos Uniprot y/o HPRD, mientras que las restantes 10 (42%) correspondían a nuevos sitios de fosforilación descritos en este trabajo. En estos 24 ligandos fosforilados, la fosforilación en P4 afectó a 18 péptidos (75%) de los cuales 8 (44%) presentaron un residuo Pro en P5. En cuanto a P1, y a diferencia de lo observado en el caso de HLA-B*40, solo 3 ligandos (17%) presentaron residuos cargados positivamente en el extremo N-terminal. De cualquier modo, esta última cifra es, significativamente mayor que la frecuencia de residuos básicos en P1 dentro del peptidoma no fosforilado de este alotipo (1%, Figura 27).

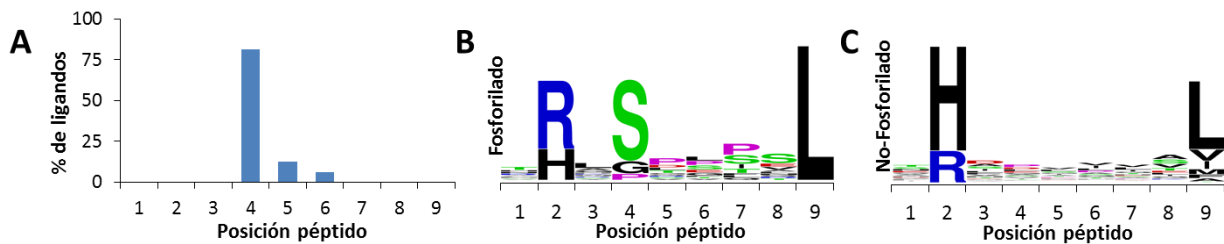


Figura 27. Análisis de las secuencias fosforiladas y no fosforiladas del peptidoma de HLA-B*39. (A) Frecuencia de fosforilación en cada posición peptídica de los ligandos fosforilados de 9 residuos de longitud. (B y C) Uso de residuos en cada posición peptídica de los ligandos fosforilados (B) o no fosforilados (C).

Discusión

V-Discusión

Hoy en día, la combinación de cromatografía líquida de alto rendimiento, análisis por espectrometría de masas y búsquedas en bases de datos, permite la identificación masiva de péptidos y proteínas con un elevado nivel de confianza. Estas mismas estrategias están siendo muy utilizadas en la identificación de los repertorios peptídicos presentados por moléculas del MHC de clase I [22-24, 180]. Sin embargo, esta tarea es significativamente más compleja debido a la menor eficiencia de los motores de búsqueda en ausencia de especificidad enzimática y a la baja cantidad de muestra típicamente disponible para el análisis (en nuestras manos son necesarios 10^{10} transfectantes del alelo de interés en C1R para obtener de 2 a 4 μg de péptidos). En cualquier caso, gracias al gran avance tecnológico que han experimentado los espectrómetros de masas, la identificación a gran escala de ligandos de HLA de clase I es ahora factible y algunos autores han propuesto la puesta en marcha del Proyecto Inmunopeptidoma Humano, análogo al Proyecto Proteoma Humano [181], para caracterizar sistemáticamente los peptidomas unidos a moléculas HLA de clase I y II [182].

En esta tesis, nos centramos fundamentalmente en la caracterización del repertorio peptídico presentado por HLA-B*40 expresado de manera estable en la línea celular C1R. Como comentamos en la introducción, HLA-B*40 ha sido poco estudiado hasta la fecha. Además de su asociación a espondiloartropatías B*27 positivas y negativas [38-41] y su posible papel como marcador de susceptibilidad de daño pulmonar en pacientes de artritis reumatoide [42], B*40 está estrechamente asociado al linfoma T del adulto causado por el virus HTLV-1 [37].

HLA-B*40 fue expresado en la línea celular C1R y purificado por afinidad. Seguidamente, se extrajo su repertorio peptídico mediante elución ácida. Los ligandos fosforilados fueron enriquecidos secuencialmente por medio de IMAC y cromatografía de afinidad con TiO_2 obteniéndose tres fracciones diferentes: IMAC, TiO_2 , y FT, que a continuación, se analizaron mediante LC-MS/MS usando CID o EThcD como mecanismo de fragmentación. Para soslayar la dificultad asociada con la ausencia de especificidad enzimática en las búsquedas con espectros MS/MS, empleamos un abordaje basado en la combinación de varios motores. Este tipo de análisis, ha sido usado previamente con éxito en la caracterización de los repertorios peptídicos presentados por moléculas del MHC de clase I [180]. Además, para minimizar el impacto de las identificaciones incorrectas, adoptamos un criterio de identificación muy estricto ($\text{FDR} \leq 1\%$ a nivel de péptido).

Tal y como se ha señalado [154], debido a que la línea C1R presenta niveles bajos de expresión de HLA-C*04 y -B*35, la identificación de ligandos presentados por estos alotipos podría introducir sesgos en el análisis del repertorio peptídico de HLA-B*40. Para evitarlo, las secuencias que presentaban el motivo de unión a C*04 [154, 155] o B*35 [156] fueron descartadas. Tras este filtrado, pudimos secuenciar más de 9500 ligandos de HLA-B*40. El número de péptidos identificados fue algo superior en el análisis por EThcD si bien esto podría achacarse al hecho de que se analizaran tres réplicas biológicas con esta técnica y solo una mediante CID.

V.1- Caracterización y bases estructurales del motivo de unión de péptidos de HLA-B*40

La masa promedio y la distribución de tamaños de los ligandos de HLA-B*40 identificados fueron muy similares, independientemente del mecanismo de fragmentación empleado. Los nonámeros fueron, de lejos, la especie más representada, como en la mayoría de moléculas HLA de clase I [183], suponiendo el 50% del repertorio. Adicionalmente, se detectó una alta frecuencia de octámeros (15%). Aunque no es excesivamente común, esta característica es compartida por otros miembros del supertipo B44 como HLA-B*37 o -B*18 [44].

Para determinar el motivo de unión a HLA-B*40, agrupamos los péptidos de acuerdo a su longitud. Este agrupamiento es necesario ya que los ligandos de clase I están anclados a través de sus extremos N y C-terminales al surco de unión [8]. En consecuencia, los péptidos cortos están unidos en una conformación extendida, mientras que los ligandos más largos sobresalen en la región central del surco de unión [6, 15, 184, 185] lo que dificulta el alineamiento preciso entre secuencias de distinta longitud, en particular entre las posiciones P4 y PΩ.

De nuevo, no se observaron grandes diferencias entre las secuencias caracterizadas por CID o EThcD, lo que indica que este último mecanismo de fragmentación no introduce sesgos significativos en los péptidos identificados. La restricción más fuerte para la unión a B*40 se observó en P2, donde la preferencia por residuos ácidos (Glu, Asp) fue prácticamente total. La base estructural de esta preferencia es evidente a partir de la estructura cristalográfica del complejo B*40-REFSKEPEL. En este caso, el grupo carboxilo interacciona directamente con los residuos His9, Lys45 y Tyr99 de la cadena pesada y, a través de una molécula de agua con Thr24, Ser67 y Asn70. Este tipo de interacción es muy similar a la observada en otras moléculas del supertipo B44 como HLA-B*41 que presenta una subcavidad B idéntica a la de B*40 [6]. Como norma general, la posición 45 es crítica para la selectividad

de la subcavidad B. En este sentido, alotipos con Lys45 como B*41 [6] o B*44 [10] unen ligandos con residuos ácidos en P2, mientras que, por ejemplo, HLA-B*27, que presenta Glu45, muestra una preferencia casi absoluta por Arg en P2 [11, 180, 186, 187].

Curiosamente, se observó una modulación del motivo en P2 en función de la longitud del péptido. Mientras que entre los octámeros asociados a B*40 las frecuencias de Glu y Asp en esta posición eran elevadas, sólo en torno a un 5% de los nonámeros contenían P2-Asp. Es más, en los ligandos de mayor tamaño, los péptidos con P2-Asp fueron prácticamente inexistentes. Por el momento, no tenemos una explicación satisfactoria para este hecho y probablemente habría que resolver más estructuras cristalográficas de B*40, con péptidos de distintas longitudes, para dilucidarlo.

Al igual que en otras moléculas de clase I, la segunda posición más influyente para la unión a B*40 se encontró en el extremo C-terminal del péptido, donde, en la mayoría de los casos, aparecían aminoácidos con carácter hidrofóbico. En este sentido, Leu estaba presente en aproximadamente el 60% de los nonámeros identificados. Además de Leu, HLA-B*40 era capaz de acomodar con menor frecuencia ligandos con otros residuos alifáticos (Met, Val, Ile) o aromáticos (Phe) en esta posición. De acuerdo a las estructuras cristalográficas presentadas en esta tesis, esta preferencia por residuos apolares en el extremo C-terminal del péptido viene dictada por la hidrofobicidad de la subcavidad F debida, en gran medida, a la contribución de los residuos Leu81, Leu95, Tyr116, Tyr123 y Trp147.

Respecto a las dos principales posiciones de anclaje, un motivo de unión similar ha sido descrito para HLA-B*40:01 [44]. No obstante, este alotipo es más restrictivo que HLA-B*40:02, presentando exclusivamente péptidos con Glu en P2 y no permitiendo la unión de Phe en P Ω . La base estructural para esta discrepancia no está clara, ya que ambas moléculas presentan subcavidades B y F muy similares. Es probable que el limitado número de ligandos identificados a partir B*40:01 –56 en el estudio mencionado– o efectos indirectos que impliquen residuos de anclaje secundarios puedan explicar esta divergencia.

Por último, además de estos dos motivos fuertes de anclaje en P2 y P Ω , se observaron ciertas preferencias más sutiles en casi todas las posiciones del ligando. Estas eran de menor magnitud cuanto mayor era la longitud de los péptidos, lo que refleja, probablemente, la conformación abombada que adoptan los péptidos más largos para ajustarse al surco de unión [6, 15, 184, 185]. En consecuencia, la región central del péptido no establece ningún contacto con la cadena pesada, y por lo tanto no hay limitaciones estructurales que establezcan la selección de motivos particulares.

V.2- Caracterización del fosfopeptidoma unido a HLA-B*40

La inmunoterapia basada en células T es una estrategia prometedora para el tratamiento del cáncer, especialmente en la fase tardía de la enfermedad o en aquellos casos que no responden a los tratamientos convencionales [188]. Este enfoque depende críticamente de la identificación de antígenos específicos de tumor presentados por moléculas del MHC de clase I, capaces de desencadenar una respuesta antitumoral efectiva [189]. En células tumorales, a menudo, la actividad de quinasas y fosfatasa se ve desregulada, provocando fosforilaciones aberrantes en proteínas [104] que posteriormente pueden ser procesadas y presentadas por moléculas de clase I [97, 105]. En este sentido, varios autores han propuesto que este tipo de ligandos podrían ser dianas óptimas para la inmunoterapia de los procesos cancerígenos [4, 5, 103, 190]. De hecho, se ha demostrado recientemente que es posible reducir el crecimiento tumoral *in vivo* mediante células T CD8+ dirigidas contra dos epítomos fosforilados derivados del sustrato 2 del receptor de insulina y del regulador del ciclo celular CDC25b [190]. Gran parte del trabajo expuesto en esta memoria consistió en el aislamiento e identificación de los fosfopéptidos presentes en el repertorio restringido por HLA-B*40 con un doble objetivo: 1) la puesta a punto de un flujo de trabajo que permitiera caracterizar el fosfopeptidoma presentado por moléculas del MHC de clase I y 2) la determinación de las bases moleculares de la presentación de este tipo de ligandos.

Sin ninguna duda, la espectrometría de masas se ha revelado como una herramienta idónea para la identificación de fosfopéptidos. No obstante, este tipo de moléculas presentan una serie de peculiaridades que dificultan su caracterización mediante abordajes proteómicos, tales como su baja estequiometría, su ineficiente ionización y la poca información contenida en sus espectros de fragmentación por CID [109, 114, 115]. En este trabajo adoptamos dos estrategias para soslayar estas dificultades: 1) la aplicación de técnicas de enriquecimiento de fosfopéptidos antes del análisis por espectrometría de masas y 2) el uso de fragmentación mediante EThcD en el análisis por LC-MS/MS. En relación con este último punto, se ha demostrado la eficiencia del EThcD en el mapeo de sitios de fosforilación de proteínas [137] y en la caracterización de los peptidomas asociados a moléculas de clase I [24] y II [191]. Adicionalmente, en los casos en que se empleó espectrometría de masas en tándem basada en CID, todas las secuencias de fosfopéptidos fueron inspeccionadas manualmente y confirmadas por fragmentación del correspondiente péptido sintético.

Como estrategias de enriquecimiento de péptidos fosforilados, aplicamos cromatografía de afinidad usando Fe^{3+} y TiO_2 como ligandos antes del análisis por LC-MS/MS. Por su parte, el uso del esquema de fragmentación ETHcD supuso una gran ventaja para la identificación de fosfopéptidos. Esta técnica permitió la identificación de más de 100 secuencias fosforiladas presentadas por HLA-B*40 (FDR $\leq 1\%$). En comparación, ni un solo fosfopéptido superó este umbral de FDR en la aproximación basada en CID, lo que demuestra la enorme ventaja de la fragmentación ETHcD en este tipo de análisis en línea con otros trabajos recientes [24, 192].

La combinación de los flujos de trabajo basados en CID y ETHcD, permitió la identificación de casi 150 fosfopéptidos asociados a HLA-B*40. Hasta donde sabemos, este es el mayor registro de ligandos fosforilados de moléculas de clase I descrito hasta la fecha. Como reflejo de las peculiaridades del fosfoproteoma de las células C1R, un número significativo de péptidos fosforilados de nuestro listado derivaban de proteínas implicadas en mitosis y control del ciclo celular, lo que confirma que las alteraciones en el patrón de fosforilación de proteínas relevantes en procesos de transformación tumoral son potencialmente detectables por las células T. Tras analizar las características moleculares de los fosfopéptidos descritos, observamos que la fosforilación se producía en Ser (93%) y Thr (7%) pero no en residuos de Tyr. Esta distribución en la frecuencia de fosforilación es muy parecida a la observada en el proteoma: 90%, 10% y $<0,05\%$ para fosfoserina, fosfotreonina, y fosfotirosina, respectivamente [109]. En cerca de un 60% de los casos, la fosforilación se produjo antes de un residuo de Pro, lo que probablemente refleja la especificidad de algunas quinasas como las pertenecientes a las familias MAP o CDK, que reconocen y fosforilan sitios SP o TP [193]. Por último, la mitad de los sitios de fosforilación descritos en este estudio habían sido previamente reportados, aunque no en el contexto del repertorio peptídico de moléculas de clase I. El 50% restante, ha sido descrito por primera vez en esta tesis, demostrando que la peptidómica del MHC puede contribuir a la caracterización, validación y anotación de modificaciones postraduccionales en proteínas.

Además de estas características, el análisis detallado de las secuencias fosforiladas reveló dos particularidades que nos llamaron poderosamente la atención: 1) la presencia de varios ligandos con residuos fosforilados en P2 y 2) el hecho de que la fosforilación afectara fundamentalmente a la posición P4 y fuera frecuentemente acompañada de residuos básicos en P1. Estas dos cuestiones se discuten en detalle a continuación.

V.2.1-HLA-B*40 puede presentar ligandos con residuos fosforilados en P2

Uno de los principales hallazgos de esta tesis fue la identificación de 7 ligandos que carecían del motivo de unión canónico a HLA-B*40 en P2. Demostramos que estas asignaciones no eran falsos positivos, ya que los espectros de fragmentación de los péptidos sintéticos correspondientes eran idénticos a los experimentales. Las 7 secuencias presentaban un residuo pSer en P2 en lugar del motivo de anclaje canónico de este alotipo: Asp2 o Glu2. Los ensayos de competición realizados con las secuencias S(pS)YGNIRAV y G(pS)FSRFYSL, demostraron la alta afinidad de estos péptidos, lo que confirma que son realmente ligandos endógenos de HLA-B*40. Dadas las similitudes estructurales entre pSer y Glu en términos de tamaño de su cadena lateral y distribución de carga, planteamos la hipótesis de que ambos podrían interactuar con la subcavidad B de manera análoga, lo que llevaría a la formación de complejos triméricos estables. De hecho, estos dos ligandos se comportaron de modo similar, en términos de afinidades de unión, que sus secuencias mutantes respectivas: SEYGNIRAV y GEFSRFYSL.

Corroborando nuestra hipótesis, la determinación estructural mediante difracción de rayos X de los péptidos G(pS)FSRFYSL y GEFSRFYSL en complejo con HLA-B*40 demostró que la interacción del residuo fosforilado con la subcavidad B era prácticamente idéntica que la se establecía en el caso del ligando con P2-Glu. Ambos péptidos presentaban una alta similitud estructural, con algunas diferencias menores en la conformación rotacional de P5-Arg en ambos complejos. Esta alta flexibilidad es probablemente debida a que este residuo está muy expuesto hacia el solvente y no interacciona con ninguna de las subcavidades de HLA-B*40.

La existencia de ligandos de moléculas de clase I que presentan residuos fosforilados en su posición principal de anclaje podría ser relevante de cara al diseño de estrategias de inmunoterapia del cáncer basadas en células T. Para que este tipo de abordajes sean efectivos, además de un epítipo diana –fosforilado en este caso– es necesaria una respuesta específica de CTLs con el fin de erradicar las células transformadas. Cuando la fosforilación se produce en residuos que no son fundamentales para el anclaje, es muy probable que se presenten tanto la especie fosforilada como la no fosforilada. De hecho, en torno al 20% de los fosfopéptidos unidos a HLA-B*40 descritos en esta tesis se identificaban también en su versión sin fosforilar. En esta situación, una célula T podría reconocer ambas especies por reactividad cruzada. Al respecto, aunque es posible generar CTLs específicos frente a fosfopéptidos [97], en algunos casos se detecta cierta actividad citotóxica contra su contrapartida no fosforilada [105]. Por lo tanto, incluso si un CTL que reconoce un ligando fosforilado pudiera escapar de la selección negativa

en el timo, la reactividad cruzada con el epítipo no fosforilado podría impedir su uso como agente terapéutico. Por el contrario, los péptidos fosforilados en P2 no presentarían esta limitación, debido a que la modificación postraduccional del residuo es esencial para la unión a la molécula de clase I, no existiendo base para una reacción cruzada contra el epítipo no modificado. Por ello, un posible objetivo para un trabajo futuro sería la identificación de ligandos de HLA de clase I fosforilados en P2 en muestras reales de tumores.

V.2.2-Los fosfopéptidos unidos a HLA-B*40 se encuentran mayoritariamente fosforilados en P4 y tienden a presentar residuos básicos en P1

En nuestro registro de fosfopéptidos presentados por HLA-B*40, era muy aparente que la fosforilación aparecía preferentemente en P4. A su vez, la mayoría de ligandos fosforilados en P4 presentaban un residuo básico en P1. Curiosamente, estas características moleculares se observan también en todos los alotipos cuyo fosfopeptidoma ha sido estudiado como HLA-A*02 [4, 106] o B*-07 [5, 97, 106].

El mecanismo molecular que subyace a este motivo de fosforilación común ha sido estudiado con cierto detalle para el caso de HLA-A*02. Como se comentó en la introducción de esta memoria, Mohammed y *cols* demostraron que la fosforilación de los ligandos endógenos en P4 a menudo producía un aumento en la afinidad de unión a la molécula de clase I [103]. Además, la determinación estructural de HLA-A*02 en complejo con cuatro péptidos fosforilados en P4, ayudó a los autores a identificar una serie de interacciones que vinculaban el grupo fosfato con los residuos Arg65 y/o Lys66 en la cadena pesada y un puente salino con el residuo cargado positivamente en P1. En un siguiente estudio [95], Petersen y *cols* confirmaron el papel clave de las interacciones que implican el grupo fosfato y los residuos Arg65 y/o Lys66 y demostraron que la interacción electrostática entre las cadenas laterales de los residuos en P1 y P4 era dependiente de epítipo. Por ejemplo, la presencia de P5-Pro y P6-Thr en uno de los ligandos estudiados no permitía una orientación adecuada de P4-pSer para interactuar con el residuo básico en P1.

Los ensayos de unión que hemos realizado para los péptidos fosforilados en P4 demuestran que, a diferencia de HLA-A*02, la fosforilación en esta posición tiene un pequeño efecto negativo sobre la afinidad de unión a HLA-B*40. Aunque este resultado aparentemente está en conflicto con la elevada frecuencia de fosforilación en este residuo, los experimentos realizados con análogos de poliglicina demuestran que la fosforilación en P4 (y también en P8) se asocia con una mayor afinidad de unión a

B*40 en comparación con otras posiciones del péptido. Estructuralmente, este efecto puede ser explicado por la interacción mediante un puente salino del grupo fosfato de la P4-Ser con la cadena lateral de Arg62 y su interacción mediada por agua con el grupo carboxilo de Glu163.

Los ensayos de unión realizados también dan cuenta de la alta frecuencia de residuos básicos en el extremo N-terminal en el fosfopeptidoma que presenta HLA-B*40 ya que la presencia de un residuo básico en esta posición incrementa la estabilidad del complejo, compensando el efecto negativo sobre la unión causado por la fosforilación en P4. Esto es probablemente debido al conjunto de interacciones que se establecen entre el residuo básico ubicado en P1 con algunos residuos de la subcavidad A de HLA-B*40. En las estructuras de B*40 en complejo con los péptidos REF(pS)KEPEL y REFSKEPEL, P1-Arg se estabiliza mediante contactos con Arg62, Glu163 y Trp167. Finalmente, el puente salino entre el grupo fosfato en P4 y el residuo básico en P1 descrito en algunos [103] pero no todos [95] los complejos HLA-A*02-fosfopéptido está ausente en la estructura de HLA-B*40 con el péptido REF(pS)KEPEL presentada en esta tesis. No obstante, los residuos P1-Arg y P4-pSer de alguna manera están vinculados por una interacción de larga distancia mediada por los residuos Arg62 y Glu163 y por una molécula de agua.

Las estructuras de los complejos B*40-REFSKEPEL y B*40-REF(pS)KEPEL presentan una alta similitud estructural, difiriendo solo en la existencia de dos orientaciones alternativas para los residuos P1-Arg y P8-Glu del ligando fosforilado y en las distintas conformaciones rotacionales que adopta P5-Lys en ambas estructuras. Esta flexibilidad del residuo en P5 se debe probablemente a la presencia de un grupo cargado negativamente como es el grupo fosfato en P4. En este sentido, nótese que P5-Lys se orienta hacia el fosfato cuando este se encuentra presente en el complejo.

Un hecho llamativo a partir de nuestros datos es la baja frecuencia de fosforilación en P8 entre los ligandos fosforilados de B*40. La afinidad de unión a este alotipo de los análogos de poliglicina fosforilados en P8 o en P4 es muy similar y sin embargo la tasa de fosforilación en esta última posición es aproximadamente 25 veces mayor que en P8. Esto sugiere la existencia de mecanismos adicionales que contribuyen a la presentación de fosfoligandos por B*40. Una posible razón es que, dado que los sitios TP y SP son motivos frecuentemente reconocidos por quinasas, un número significativo de péptidos con pSer o pThr en P8 presentarían Pro en el extremo C-terminal. Aunque B*40 es capaz de unir ligandos con Pro en esta posición, el análisis del motivo de unión de este alelo demuestra que no es un anclaje particularmente favorable. Una segunda explicación no excluyente, podría estar relacionada con eventos en la ruta de procesamiento antigénico tales como la translocación de péptidos a través de

TAP o el recorte de precursores N-terminales de ligandos por aminopeptidasas. Como se ha explicado en la introducción de esta tesis, la especificidad de TAP viene impuesta fundamentalmente por la naturaleza del residuo ubicado en el C-terminal y los tres residuos N-terminales [71-73], siendo improbable que la fosforilación en P4 o P8 pueda afectar al transporte de fosfopéptidos al retículo endoplásmico. Por el contrario, el recorte por aminopeptidasas como ERAP1 o ERAP2 teóricamente podría ser modulado por la secuencia interna de sus sustratos. Por ejemplo, aunque la eficiencia de corte de ERAP1 depende en gran medida del residuo N-terminal [81], se ha demostrado que la actividad de esta enzima se ve incrementada por la presencia de cadenas laterales cargadas positivamente cerca del C-terminal y reducida cuando existen residuos ácidos en las posiciones internas del péptido [194]. En cualquier caso, el papel, si lo hubiera, de ERAP1 y ERAP2 en la configuración del fosfopeptidoma presentado por moléculas del MHC clase I está aún por determinar.

V.3-La preferencia por ligandos fosforilados en P4 podría ser común a la mayoría de moléculas HLA-B

El hecho de que el residuo Arg62 –clave para la estabilización del grupo fosfato en P4– se encuentre conservado en la gran mayoría de los alotipos HLA-B podría ser la base para el motivo de fosforilación común observado en HLA-B*07 [5, 97, 106], -B*27 [97] y -B*40 (Tabla 11).

| Alotipo | Número de Fosfopéptidos | XXX(pS/pT) | (K/R)XX(pS/pT) | Referencia |
|----------|-------------------------|------------|----------------|--------------------------|
| HLA-B*07 | 6 | 5 (83%) | 5 (100%)* | Zarling et al 2000 [97] |
| HLA-B*07 | 8 | 5 (63%) | 5 (100%)* | Meyer et al 2009 [5] |
| HLA-B*07 | 85 | 62 (73%) | 43 (69%)* | Cobbold et al 2013 [106] |
| HLA-B*27 | 6 | 6 (100%) | 5 (83%)* | Zarling et al 2000 [97] |
| HLA-B*40 | 149 | 102 (68%) | 57 (56%)* | Esta Tesis |
| HLA-B*39 | 24 | 18 (75%) | 3 (17%)* | Esta Tesis |

Tabla 11. Características moleculares de los fosfopéptidos asociados a distintas moléculas HLA-B. *Porcentaje referido al número absoluto de péptidos fosforilados en P4.

El análisis del fosfopeptidoma asociado a HLA-B*39 –que mantiene el residuo Arg62– parece apoyar esta hipótesis, ya que también estaba enriquecido en péptidos fosforilados en P4. Sin embargo, a diferencia de los alelos mencionados anteriormente, la frecuencia de residuos básicos en P1 no resultó ser particularmente elevada, si bien era significativamente mayor en comparación con el resto de ligandos no modificados. Este hecho podría estar relacionado con la presencia de Thr163 en lugar de Glu163 en esta molécula, que haría a la subcavidad A menos propensa a acomodar cadenas laterales

cargadas positivamente. En este sentido, los péptidos endógenos de algunas moléculas HLA-B con Thr163 tales como HLA-B*14 [195], -B*18 [44] o -B*44 [44] incluyen muy pocos residuos básicos en el N-terminal. Esta correlación se rompe, sin embargo, por alotipos como HLA-B*37, donde un 33% de los péptidos que se presentan tienen Arg o Lys en P1 [44]. Del mismo modo, la sustitución de Glu por Thr en la posición 163 en HLA-B*27 no impide la presentación de este tipo de péptidos, aunque disminuye la estabilidad térmica de la molécula [196]. Por lo tanto, Glu163 no puede ser el único determinante de la preferencia dependiente de alotipo por fosfopéptidos con residuos básicos en P1.

En conjunto, los resultados expuestos en esta memoria demuestran que la serina fosforilada puede funcionar como anclaje principal a moléculas de clase I, al menos en el caso de HLA-B*40 y, probablemente, en el de otras moléculas con preferencia por residuos ácidos en P2, como los distintos miembros del supertipo B44. La presentación de esta clase de ligandos contribuye al incremento de la diversidad del repertorio peptídico presentado y podría ser relevante como fuente de neoantígenos en el contexto de un proceso de transformación tumoral. Nuestros datos también indican que la fosforilación preferente en P4 puede ser una característica común a los fosfopeptidomas presentados por la mayoría de moléculas HLA-B. Esta preferencia va acompañada de una selección de residuos básicos en P1 cuya magnitud es claramente dependiente de alotipo. El conocimiento de las bases moleculares de la presentación de fosfopéptidos por moléculas HLA-B puede contribuir a la identificación y predicción de epítopos fosforilados para el tratamiento del cáncer mediante inmunoterapia basada en células T.

Conclusiones

VI-Conclusiones

- 1- El flujo de trabajo descrito en esta tesis, basado en enriquecimiento de fosfopéptidos, análisis mediante espectrometría de masas y combinación de motores de búsqueda, es una herramienta eficaz para la caracterización del peptidoma y el fosfopeptidoma presentado por moléculas del MHC de clase I.
- 2- En comparación con el esquema de fragmentación CID, el ETHcD es mucho más eficiente en la identificación de ligandos fosforilados asociados a moléculas del MHC de clase I.
- 3- HLA-B*40:02 une preferentemente ligandos con residuos ácidos –en especial Glu– en P2, siendo esta su posición de anclaje principal. Esta restricción viene impuesta por la naturaleza de los residuos que conforman la subcavidad B y, en particular, por la presencia de Lys45, His9 y Tyr99 en este alotipo.
- 4- HLA-B*40:02 une fundamentalmente péptidos con residuos alifáticos, Met o Phe en su extremo C-terminal debido a la elevada hidrofobicidad de la subcavidad F de esta molécula que, a su vez, viene determinada por los residuos Leu81, Leu95, Tyr116, Tyr123 y Trp147 de la cadena pesada.
- 5- Gracias a la elevada similitud estructural entre los residuos de pSer y Glu, HLA-B*40:02 puede presentar péptidos fosforilados en P2. Estos ligandos podrían constituir dianas óptimas para el tratamiento del cáncer por inmunoterapia basada en células T.
- 6- El fosfopeptidoma presentado por HLA-B*40:02 está enriquecido en péptidos fosforilados en posición 4. Este hecho se explica por la interacción del grupo fosfato con el residuo Arg62 de la cadena pesada. El alto grado de conservación de este residuo sugiere que esta puede ser una característica compartida por la mayoría de moléculas HLA-B, incluidas HLA-B*07, -B*27 y B*39.
- 7- La frecuencia de residuos básicos en P1 entre los ligandos fosforilados de distintas moléculas HLA-B se encuentra incrementada respecto a la observada en el peptidoma no modificado correspondiente. La magnitud de este incremento es dependiente de alotipo y podría estar relacionada, al menos parcialmente, con la estructura de la subcavidad A.

Referencias

VII-Referencias

1. Auffray, C. and J.L. Strominger, *Molecular genetics of the human major histocompatibility complex*. Adv Hum Genet, 1986. **15**: p. 197-247.
2. Klein, J. and A. Sato, *The HLA system. First of two parts*. N Engl J Med, 2000. **343**(10): p. 702-9.
3. Engelhard, V.H., et al., *Post-translational modifications of naturally processed MHC-binding epitopes*. Curr Opin Immunol, 2006. **18**(1): p. 92-7.
4. Zarling, A.L., et al., *Identification of class I MHC-associated phosphopeptides as targets for cancer immunotherapy*. Proc Natl Acad Sci U S A, 2006. **103**(40): p. 14889-94.
5. Meyer, V.S., et al., *Identification of natural MHC class II presented phosphopeptides and tumor-derived MHC class I phospholigands*. J Proteome Res, 2009. **8**(7): p. 3666-74.
6. Bade-Doding, C., et al., *The impact of human leukocyte antigen (HLA) micropolymorphism on ligand specificity within the HLA-B*41 allotypic family*. Haematologica, 2011. **96**(1): p. 110-8.
7. Strominger, J.L., *Structure of class I and class II HLA antigens*. Br Med Bull, 1987. **43**(1): p. 81-93.
8. Madden, D.R., *The three-dimensional structure of peptide-MHC complexes*. Annu Rev Immunol, 1995. **13**: p. 587-622.
9. Archbold, J.K., et al., *Natural micropolymorphism in human leukocyte antigens provides a basis for genetic control of antigen recognition*. J Exp Med, 2009. **206**(1): p. 209-19.
10. Macdonald, W.A., et al., *A naturally selected dimorphism within the HLA-B44 supertype alters class I structure, peptide repertoire, and T cell recognition*. J Exp Med, 2003. **198**(5): p. 679-91.
11. Madden, D.R., et al., *The three-dimensional structure of HLA-B27 at 2.1 Å resolution suggests a general mechanism for tight peptide binding to MHC*. Cell, 1992. **70**(6): p. 1035-48.
12. Orr, H.T., et al., *Complete amino acid sequence of a papain-solubilized human histocompatibility antigen, HLA-B7. 2. Sequence determination and search for homologies*. Biochemistry, 1979. **18**(25): p. 5711-20.
13. Bjorkman, P.J., et al., *The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens*. Nature, 1987. **329**(6139): p. 512-8.
14. Saper, M.A., P.J. Bjorkman, and D.C. Wiley, *Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Å resolution*. J Mol Biol, 1991. **219**(2): p. 277-319.
15. Guo, H.C., et al., *Different length peptides bind to HLA-Aw68 similarly at their ends but bulge out in the middle*. Nature, 1992. **360**(6402): p. 364-6.
16. Madden, D.R., D.N. Garboczi, and D.C. Wiley, *The antigenic identity of peptide-MHC complexes: a comparison of the conformations of five viral peptides presented by HLA-A2*. Cell, 1993. **75**(4): p. 693-708.
17. Garrett, T.P., et al., *Specificity pockets for the side chains of peptide antigens in HLA-Aw68*. Nature, 1989. **342**(6250): p. 692-6.
18. Bouvier, M. and D.C. Wiley, *Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules*. Science, 1994. **265**(5170): p. 398-402.

19. Lopez de Castro, J., et al., *Primary structure of papain-solubilized human histocompatibility antigen HLA-B40 (-Bw60). An outline of alloantigenic determinants*. *Biochemistry*, 1983. **22**(16): p. 3961-9.
20. Young, A.C., S.G. Nathenson, and J.C. Sacchettini, *Structural studies of class I major histocompatibility complex proteins: insights into antigen presentation*. *FASEB J*, 1995. **9**(1): p. 26-36.
21. Falk, K., et al., *Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules*. *Nature*, 1991. **351**(6324): p. 290-6.
22. Hassan, C., et al., *The human leukocyte antigen-presented ligandome of B lymphocytes*. *Mol Cell Proteomics*, 2013. **12**(7): p. 1829-43.
23. Bassani-Sternberg, M., et al., *Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation*. *Mol Cell Proteomics*, 2015. **14**(3): p. 658-73.
24. Mommen, G.P., et al., *Expanding the detectable HLA peptide repertoire using electron-transfer/higher-energy collision dissociation (EThcD)*. *Proc Natl Acad Sci U S A*, 2014. **111**(12): p. 4507-12.
25. Marsh, S.G.E., P. Parham, and L.D. Barber, *Part 2 - HLA-B*, in *The HLA FactsBook*. 2000, Academic Press: London. p. 151-241.
26. Yamaguchi, A., et al., *Association of HLA-B39 with HLA-B27-negative ankylosing spondylitis and pauciarticular juvenile rheumatoid arthritis in Japanese patients. Evidence for a role of the peptide-anchoring B pocket*. *Arthritis Rheum*, 1995. **38**(11): p. 1672-7.
27. Kitamura, H., et al., *Association of clinical manifestations with HLA-B alleles in Takayasu arteritis*. *Int J Cardiol*, 1998. **66 Suppl 1**: p. S121-6.
28. Shiozawa, S., *HLA-B39 and asymmetric arthritis*. *Intern Med*, 2000. **39**(1): p. 8-9.
29. Salvarani, C., et al., *The cervical spine in patients with psoriatic arthritis: a clinical, radiological and immunogenetic study*. *Ann Rheum Dis*, 1992. **51**(1): p. 73-7.
30. Maeda, H., et al., *A family with cases of adult onset Still's disease and psoriatic arthritis*. *Intern Med*, 2000. **39**(1): p. 77-9.
31. Gladman, D.D. and V.T. Farewell, *The role of HLA antigens as indicators of disease progression in psoriatic arthritis. Multivariate relative risk model*. *Arthritis Rheum*, 1995. **38**(6): p. 845-50.
32. Mendez-Sanchez, N., et al., *The Amerindian's genes in the Mexican population are associated with development of gallstone disease*. *Am J Gastroenterol*, 2004. **99**(11): p. 2166-70.
33. Bravo, M.J., et al., *HLA-B*39 allele confers susceptibility to osteoarticular complications in human brucellosis*. *J Rheumatol*, 2003. **30**(5): p. 1051-3.
34. Falk, K., et al., *Peptide motifs of HLA-B38 and B39 molecules*. *Immunogenetics*, 1995. **41**(2-3): p. 162-4.
35. Yague, J., et al., *The South Amerindian allotype HLA-B*3909 has the largest known similarity in peptide specificity and common natural ligands with HLA-B27*. *Tissue Antigens*, 1999. **53**(3): p. 227-36.

36. Sidney, J., et al., *HLA class I supertypes: a revised and updated classification*. BMC Immunol, 2008. **9**: p. 1.
37. Yashiki, S., et al., *HLA-A*26, HLA-B*4002, HLA-B*4006, and HLA-B*4801 alleles predispose to adult T cell leukemia: the limited recognition of HTLV type 1 tax peptide anchor motifs and epitopes to generate anti-HTLV type 1 tax CD8(+) cytotoxic T lymphocytes*. AIDS Res Hum Retroviruses, 2001. **17**(11): p. 1047-61.
38. Wei, J.C., et al., *HLA-B60 and B61 are strongly associated with ankylosing spondylitis in HLA-B27-negative Taiwan Chinese patients*. Rheumatology (Oxford), 2004. **43**(7): p. 839-42.
39. Cortes, A., et al., *Major histocompatibility complex associations of ankylosing spondylitis are complex and involve further epistasis with ERAP1*. Nat Commun, 2015. **6**: p. 7146.
40. Robinson, W.P., et al., *HLA-Bw60 increases susceptibility to ankylosing spondylitis in HLA-B27+ patients*. Arthritis Rheum, 1989. **32**(9): p. 1135-41.
41. Rubin, L.A., et al., *Investigating the genetic basis for ankylosing spondylitis. Linkage studies with the major histocompatibility complex region*. Arthritis Rheum, 1994. **37**(8): p. 1212-20.
42. Charles, P.J., et al., *HLA-B40: a marker for susceptibility to lung disease in rheumatoid arthritis*. Dis Markers, 1991. **9**(2): p. 97-101.
43. Harris, P.E., et al., *Naturally processed HLA class I bound peptides from c-myc-transfected cells reveal allele-specific motifs*. J Immunol, 1993. **151**(11): p. 5966-74.
44. Hillen, N., et al., *Essential differences in ligand presentation and T cell epitope recognition among HLA molecules of the HLA-B44 supertype*. Eur J Immunol, 2008. **38**(11): p. 2993-3003.
45. Rock, K.L. and A.L. Goldberg, *Degradation of cell proteins and the generation of MHC class I-presented peptides*. Annu Rev Immunol, 1999. **17**: p. 739-79.
46. Rock, K.L., et al., *Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules*. Cell, 1994. **78**(5): p. 761-71.
47. Rock, K.L., et al., *Re-examining class-I presentation and the DRiP hypothesis*. Trends Immunol, 2014. **35**(4): p. 144-52.
48. Palmer, A.L. and B.P. Dolan, *MHC class I antigen presentation of DRiP-derived peptides from a model antigen is not dependent on the AAA ATPase p97*. PLoS One, 2013. **8**(7): p. e67796.
49. Yewdell, J.W., L.C. Anton, and J.R. Bennink, *Defective ribosomal products (DRiPs): a major source of antigenic peptides for MHC class I molecules?* J Immunol, 1996. **157**(5): p. 1823-6.
50. Yewdell, J.W., U. Schubert, and J.R. Bennink, *At the crossroads of cell biology and immunology: DRiPs and other sources of peptide ligands for MHC class I molecules*. J Cell Sci, 2001. **114**(Pt 5): p. 845-51.
51. Yewdell, J.W. and C.V. Nicchitta, *The DRiP hypothesis decennial: support, controversy, refinement and extension*. Trends Immunol, 2006. **27**(8): p. 368-73.
52. Anton, L.C. and J.W. Yewdell, *Translating DRiPs: MHC class I immunosurveillance of pathogens and tumors*. J Leukoc Biol, 2014. **95**(4): p. 551-62.
53. Dick, L.R., et al., *Degradation of oxidized insulin B chain by the multiproteinase complex macropain (proteasome)*. Biochemistry, 1991. **30**(10): p. 2725-34.

54. Boes, B., et al., *Interferon gamma stimulation modulates the proteolytic activity and cleavage site preference of 20S mouse proteasomes*. J Exp Med, 1994. **179**(3): p. 901-9.
55. Craiu, A., et al., *Two distinct proteolytic processes in the generation of a major histocompatibility complex class I-presented peptide*. Proc Natl Acad Sci U S A, 1997. **94**(20): p. 10850-5.
56. Coux, O., K. Tanaka, and A.L. Goldberg, *Structure and functions of the 20S and 26S proteasomes*. Annu Rev Biochem, 1996. **65**: p. 801-47.
57. Arendt, C.S. and M. Hochstrasser, *Identification of the yeast 20S proteasome catalytic centers and subunit interactions required for active-site formation*. Proc Natl Acad Sci U S A, 1997. **94**(14): p. 7156-61.
58. Heinemeyer, W., et al., *The active sites of the eukaryotic 20 S proteasome and their involvement in subunit precursor processing*. J Biol Chem, 1997. **272**(40): p. 25200-9.
59. Groll, M., et al., *Structure of 20S proteasome from yeast at 2.4 Å resolution*. Nature, 1997. **386**(6624): p. 463-71.
60. Peters, J.M., W.W. Franke, and J.A. Kleinschmidt, *Distinct 19 S and 20 S subcomplexes of the 26 S proteasome and their distribution in the nucleus and the cytoplasm*. J Biol Chem, 1994. **269**(10): p. 7709-18.
61. Voges, D., P. Zwickl, and W. Baumeister, *The 26S proteasome: a molecular machine designed for controlled proteolysis*. Annu Rev Biochem, 1999. **68**: p. 1015-68.
62. DeMartino, G.N., et al., *PA700, an ATP-dependent activator of the 20 S proteasome, is an ATPase containing multiple members of a nucleotide-binding protein family*. J Biol Chem, 1994. **269**(33): p. 20878-84.
63. Braun, B.C., et al., *The base of the proteasome regulatory particle exhibits chaperone-like activity*. Nat Cell Biol, 1999. **1**(4): p. 221-6.
64. Kisselev, A.F., et al., *The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes. Implications for understanding the degradative mechanism and antigen presentation*. J Biol Chem, 1999. **274**(6): p. 3363-71.
65. Shastri, N., S. Schwab, and T. Serwold, *Producing nature's gene-chips: the generation of peptides for display by MHC class I molecules*. Annu Rev Immunol, 2002. **20**: p. 463-93.
66. Saric, T., et al., *Major histocompatibility complex class I-presented antigenic peptides are degraded in cytosolic extracts primarily by thimet oligopeptidase*. J Biol Chem, 2001. **276**(39): p. 36474-81.
67. Stoltze, L., et al., *Two new proteases in the MHC class I processing pathway*. Nat Immunol, 2000. **1**(5): p. 413-8.
68. Levy, F., et al., *The final N-terminal trimming of a subaminoterminal proline-containing HLA class I-restricted antigenic peptide in the cytosol is mediated by two peptidases*. J Immunol, 2002. **169**(8): p. 4161-71.
69. Kessler, J.H., et al., *Antigen processing by nardilysin and thimet oligopeptidase generates cytotoxic T cell epitopes*. Nat Immunol, 2011. **12**(1): p. 45-53.
70. Vos, J.C., et al., *Membrane topology and dimerization of the two subunits of the transporter associated with antigen processing reveal a three-domain structure*. J Immunol, 1999. **163**(12): p. 6679-85.

71. Peters, B., et al., *Identifying MHC class I epitopes by predicting the TAP transport efficiency of epitope precursors*. J Immunol, 2003. **171**(4): p. 1741-9.
72. van Endert, P.M., et al., *The peptide-binding motif for the human transporter associated with antigen processing*. J Exp Med, 1995. **182**(6): p. 1883-95.
73. Momburg, F., et al., *Selectivity of MHC-encoded peptide transporters from human, mouse and rat*. Nature, 1994. **367**(6464): p. 648-51.
74. Uebel, S. and R. Tampe, *Specificity of the proteasome and the TAP transporter*. Curr Opin Immunol, 1999. **11**(2): p. 203-8.
75. Neefjes, J., et al., *Analysis of the fine specificity of rat, mouse and human TAP peptide transporters*. Eur J Immunol, 1995. **25**(4): p. 1133-6.
76. Elliott, T., et al., *Processing of major histocompatibility class I-restricted antigens in the endoplasmic reticulum*. J Exp Med, 1995. **181**(4): p. 1481-91.
77. Chang, S.C., et al., *The ER aminopeptidase, ERAP1, trims precursors to lengths of MHC class I peptides by a "molecular ruler" mechanism*. Proc Natl Acad Sci U S A, 2005. **102**(47): p. 17107-12.
78. Saric, T., et al., *An IFN-gamma-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides*. Nat Immunol, 2002. **3**(12): p. 1169-76.
79. Serwold, T., S. Gaw, and N. Shastri, *ER aminopeptidases generate a unique pool of peptides for MHC class I molecules*. Nat Immunol, 2001. **2**(7): p. 644-51.
80. Serwold, T., et al., *ERAAP customizes peptides for MHC class I molecules in the endoplasmic reticulum*. Nature, 2002. **419**(6906): p. 480-3.
81. Hearn, A., I.A. York, and K.L. Rock, *The specificity of trimming of MHC class I-presented peptides in the endoplasmic reticulum*. J Immunol, 2009. **183**(9): p. 5526-36.
82. Kochan, G., et al., *Crystal structures of the endoplasmic reticulum aminopeptidase-1 (ERAP1) reveal the molecular basis for N-terminal peptide trimming*. Proc Natl Acad Sci U S A, 2011. **108**(19): p. 7745-50.
83. Nguyen, T.T., et al., *Structural basis for antigenic peptide precursor processing by the endoplasmic reticulum aminopeptidase ERAP1*. Nat Struct Mol Biol, 2011. **18**(5): p. 604-13.
84. Lorente, E., et al., *Concerted in vitro trimming of viral HLA-B27-restricted ligands by human ERAP1 and ERAP2 aminopeptidases*. PLoS One, 2013. **8**(11): p. e79596.
85. Evnouchidou, I., et al., *ERAP1-ERAP2 dimerization increases peptide-trimming efficiency*. J Immunol, 2014. **193**(2): p. 901-8.
86. Saveanu, L., et al., *Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum*. Nat Immunol, 2005. **6**(7): p. 689-97.
87. Birtley, J.R., et al., *The crystal structure of human endoplasmic reticulum aminopeptidase 2 reveals the atomic basis for distinct roles in antigen processing*. Biochemistry, 2012. **51**(1): p. 286-95.
88. Hattori, A., et al., *Molecular cloning of adipocyte-derived leucine aminopeptidase highly related to placental leucine aminopeptidase/oxytocinase*. J Biochem, 1999. **125**(5): p. 931-8.

89. Tanioka, T., et al., *Human leukocyte-derived arginine aminopeptidase. The third member of the oxytocinase subfamily of aminopeptidases*. J Biol Chem, 2003. **278**(34): p. 32275-83.
90. Dick, T.P., et al., *Disulfide bond isomerization and the assembly of MHC class I-peptide complexes*. Immunity, 2002. **16**(1): p. 87-98.
91. Sadasivan, B., et al., *Roles for calreticulin and a novel glycoprotein, tapasin, in the interaction of MHC class I molecules with TAP*. Immunity, 1996. **5**(2): p. 103-14.
92. Wearsch, P.A. and P. Cresswell, *The quality control of MHC class I peptide loading*. Curr Opin Cell Biol, 2008. **20**(6): p. 624-31.
93. Vigneron, N., et al., *Functional significance of tapasin membrane association and disulfide linkage to ERp57 in MHC class I presentation*. Eur J Immunol, 2009. **39**(9): p. 2371-6.
94. Petersen, J., A.W. Purcell, and J. Rossjohn, *Post-translationally modified T cell epitopes: immune recognition and immunotherapy*. J Mol Med (Berl), 2009. **87**(11): p. 1045-51.
95. Petersen, J., et al., *Phosphorylated self-peptides alter human leukocyte antigen class I-restricted antigen presentation and generate tumor-specific epitopes*. Proc Natl Acad Sci U S A, 2009. **106**(8): p. 2776-81.
96. Yague, J., et al., *An N-acetylated natural ligand of human histocompatibility leukocyte antigen (HLA)-B39. Classical major histocompatibility complex class I proteins bind peptides with a blocked NH(2) terminus in vivo*. J Exp Med, 2000. **191**(12): p. 2083-92.
97. Zarling, A.L., et al., *Phosphorylated peptides are naturally processed and presented by major histocompatibility complex class I molecules in vivo*. J Exp Med, 2000. **192**(12): p. 1755-62.
98. Yague, J., J. Vazquez, and J.A. Lopez de Castro, *A post-translational modification of nuclear proteins, N(G),N(G)-dimethyl-Arg, found in a natural HLA class I peptide ligand*. Protein Sci, 2000. **9**(11): p. 2210-7.
99. Kastrup, I.B., et al., *Lectin purified human class I MHC-derived peptides: evidence for presentation of glycopeptides in vivo*. Tissue Antigens, 2000. **56**(2): p. 129-35.
100. Dengjel, J., H.G. Rammensee, and S. Stevanovic, *Glycan side chains on naturally presented MHC class II ligands*. J Mass Spectrom, 2005. **40**(1): p. 100-4.
101. Haurum, J.S., et al., *Presentation of cytosolic glycosylated peptides by human class I major histocompatibility complex molecules in vivo*. J Exp Med, 1999. **190**(1): p. 145-50.
102. Mannering, S.I., et al., *The insulin A-chain epitope recognized by human T cells is posttranslationally modified*. J Exp Med, 2005. **202**(9): p. 1191-7.
103. Mohammed, F., et al., *Phosphorylation-dependent interaction between antigenic peptides and MHC class I: a molecular basis for the presentation of transformed self*. Nat Immunol, 2008. **9**(11): p. 1236-43.
104. Blume-Jensen, P. and T. Hunter, *Oncogenic kinase signalling*. Nature, 2001. **411**(6835): p. 355-65.
105. Andersen, M.H., et al., *Phosphorylated peptides can be transported by TAP molecules, presented by class I MHC molecules, and recognized by phosphopeptide-specific CTL*. J Immunol, 1999. **163**(7): p. 3812-8.

106. Cobbold, M., et al., *MHC class I-associated phosphopeptides are the targets of memory-like immunity in leukemia*. *Sci Transl Med*, 2013. **5**(203): p. 203ra125.
107. Mann, M., et al., *Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome*. *Trends Biotechnol*, 2002. **20**(6): p. 261-8.
108. Ciesla, J., T. Fraczyk, and W. Rode, *Phosphorylation of basic amino acid residues in proteins: important but easily missed*. *Acta Biochim Pol*, 2011. **58**(2): p. 137-48.
109. Paradela, A. and J.P. Albar, *Advances in the analysis of protein phosphorylation*. *J Proteome Res*, 2008. **7**(5): p. 1809-18.
110. Abelin, J.G., et al., *Complementary IMAC enrichment methods for HLA-associated phosphopeptide identification by mass spectrometry*. *Nat Protoc*, 2015. **10**(9): p. 1308-18.
111. Larsen, M.R., et al., *Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns*. *Mol Cell Proteomics*, 2005. **4**(7): p. 873-86.
112. Larsen, M.R., et al., *Analysis of posttranslational modifications of proteins by tandem mass spectrometry*. *Biotechniques*, 2006. **40**(6): p. 790-8.
113. Zhao, Y. and O.N. Jensen, *Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques*. *Proteomics*, 2009. **9**(20): p. 4632-41.
114. Boersema, P.J., S. Mohammed, and A.J. Heck, *Phosphopeptide fragmentation and analysis by mass spectrometry*. *J Mass Spectrom*, 2009. **44**(6): p. 861-78.
115. DeGnore, J.P. and J. Qin, *Fragmentation of phosphopeptides in an ion trap mass spectrometer*. *J Am Soc Mass Spectrom*, 1998. **9**(11): p. 1175-88.
116. Gronborg, M., et al., *A mass spectrometry-based proteomic approach for identification of serine/threonine-phosphorylated proteins by enrichment with phospho-specific antibodies: identification of a novel protein, Frigg, as a protein kinase A substrate*. *Mol Cell Proteomics*, 2002. **1**(7): p. 517-27.
117. Goedert, M., R. Jakes, and E. Vanmechelen, *Monoclonal antibody AT8 recognises tau protein phosphorylated at both serine 202 and threonine 205*. *Neurosci Lett*, 1995. **189**(3): p. 167-9.
118. Scimeca, J.C., et al., *Tyrosine and threonine phosphorylation of an immunoaffinity-purified 44-kDa MAP kinase*. *Biochemistry*, 1991. **30**(38): p. 9313-9.
119. Zhang, Y., et al., *Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules*. *Mol Cell Proteomics*, 2005. **4**(9): p. 1240-50.
120. Rush, J., et al., *Immunoaffinity profiling of tyrosine phosphorylation in cancer cells*. *Nat Biotechnol*, 2005. **23**(1): p. 94-101.
121. Motoyama, A., et al., *Anion and cation mixed-bed ion exchange for enhanced multidimensional separations of peptides and phosphopeptides*. *Anal Chem*, 2007. **79**(10): p. 3623-34.
122. Matsumoto, H., E.S. Kahn, and N. Komori, *Separation of phosphopeptides from their nonphosphorylated forms by reversed-phase POROS perfusion chromatography at alkaline pH*. *Anal Biochem*, 1997. **251**(1): p. 116-9.

123. Pinkse, M.W., et al., *Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns*. Anal Chem, 2004. **76**(14): p. 3935-43.
124. Fila, J. and D. Honys, *Enrichment techniques employed in phosphoproteomics*. Amino Acids, 2012. **43**(3): p. 1025-47.
125. Jensen, S.S. and M.R. Larsen, *Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques*. Rapid Commun Mass Spectrom, 2007. **21**(22): p. 3635-45.
126. Thingholm, T.E. and O.N. Jensen, *Enrichment and characterization of phosphopeptides by immobilized metal affinity chromatography (IMAC) and mass spectrometry*. Methods Mol Biol, 2009. **527**: p. 47-56, xi.
127. Posewitz, M.C. and P. Tempst, *Immobilized gallium(III) affinity chromatography of phosphopeptides*. Anal Chem, 1999. **71**(14): p. 2883-92.
128. Zhou, H., et al., *Specific phosphopeptide enrichment with immobilized titanium ion affinity chromatography adsorbent for phosphoproteome analysis*. J Proteome Res, 2008. **7**(9): p. 3957-67.
129. Zhou, H., et al., *Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography*. Nat Protoc, 2013. **8**(3): p. 461-80.
130. de Graaf, E.L., et al., *Single-step enrichment by Ti⁴⁺-IMAC and label-free quantitation enables in-depth monitoring of phosphorylation dynamics with high reproducibility and temporal resolution*. Mol Cell Proteomics, 2014. **13**(9): p. 2426-34.
131. Zubarev, R.A.K., N.L.; McLafferty, F.W., *Electron Capture Dissociation of Multiply Charged Protein Cations*. ACS Publications, 1998. **120**: p. 2.
132. Syka, J.E., et al., *Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry*. Proc Natl Acad Sci U S A, 2004. **101**(26): p. 9528-33.
133. Frese, C.K., et al., *Toward full peptide sequence coverage by dual fragmentation combining electron-transfer and higher-energy collision dissociation tandem mass spectrometry*. Anal Chem, 2012. **84**(22): p. 9668-73.
134. Frese, C.K., et al., *Improved peptide identification by targeted fragmentation using CID, HCD and ETD on an LTQ-Orbitrap Velos*. J Proteome Res, 2011. **10**(5): p. 2377-88.
135. Swaney, D.L., et al., *Supplemental activation method for high-efficiency electron-transfer dissociation of doubly protonated peptide precursors*. Anal Chem, 2007. **79**(2): p. 477-85.
136. Liu, C.W. and C.C. Lai, *Effects of electron-transfer coupled with collision-induced dissociation (ET/CID) on doubly charged peptides and phosphopeptides*. J Am Soc Mass Spectrom, 2011. **22**(1): p. 57-66.
137. Frese, C.K., et al., *Unambiguous phosphosite localization using electron-transfer/higher-energy collision dissociation (ETHcD)*. J Proteome Res, 2013. **12**(3): p. 1520-5.
138. Zemmour, J., et al., *The HLA-A,B "negative" mutant cell line C1R expresses a novel HLA-B35 allele, which also has a point mutation in the translation initiation codon*. J Immunol, 1992. **148**(6): p. 1941-8.

139. Ramos, M., et al., *Primary structure of a novel HLA-B39 allele (B*3909) from the Warao Indians of Venezuela. Further evidence for local HLA-B diversification in South America*. Tissue Antigens, 1995. **46**(5): p. 401-4.
140. Barnstable, C.J., et al., *Production of monoclonal antibodies to group A erythrocytes, HLA and other human cell surface antigens-new tools for genetic analysis*. Cell, 1978. **14**(1): p. 9-20.
141. Marcilla, M., J.J. Cragnolini, and J.A. Lopez de Castro, *Proteasome-independent HLA-B27 ligands arise mainly from small basic proteins*. Mol Cell Proteomics, 2007. **6**(5): p. 923-38.
142. Navajas, R., A. Paradela, and J.P. Albar, *Immobilized metal affinity chromatography/reversed-phase enrichment of phosphopeptides and analysis by CID/ETD tandem mass spectrometry*. Methods Mol Biol, 2011. **681**: p. 337-48.
143. Thingholm, T.E., et al., *SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides*. Mol Cell Proteomics, 2008. **7**(4): p. 661-71.
144. Ramos-Fernandez, A., et al., *Generalized method for probability-based peptide and protein identification from tandem mass spectrometry data and sequence database searching*. Mol Cell Proteomics, 2008. **7**(9): p. 1748-54.
145. Kessler, J.H., et al., *Competition-based cellular peptide binding assay for HLA class I*. Curr Protoc Immunol, 2004. **Chapter 18**: p. Unit 18 12.
146. van der Burg, S.H., et al., *An HLA class I peptide-binding assay based on competition for binding to class I molecules on intact human B cells. Identification of conserved HIV-1 polymerase peptides binding to HLA-A*0301*. Hum Immunol, 1995. **44**(4): p. 189-98.
147. Jimenez, D., et al., *Contribution of N-linked glycans to the conformation and function of intercellular adhesion molecules (ICAMs)*. J Biol Chem, 2005. **280**(7): p. 5854-61.
148. Reid, S.W., et al., *Production and crystallization of MHC class I B allele single peptide complexes*. FEBS Lett, 1996. **383**(1-2): p. 119-23.
149. Juanhuix, J., et al., *Developments in optics and performance at BL13-XALOC, the macromolecular crystallography beamline at the ALBA synchrotron*. J Synchrotron Radiat, 2014. **21**(Pt 4): p. 679-89.
150. Kabsch, W., *Xds*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 125-32.
151. Vagin, A.A., et al., *REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use*. Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2184-95.
152. Emsley, P., et al., *Features and development of Coot*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.
153. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.
154. Schittenhelm, R.B., et al., *A comprehensive analysis of constitutive naturally processed and presented HLA-C*04:01 (Cw4)-specific peptides*. Tissue Antigens, 2014. **83**(3): p. 174-9.
155. Buchsbaum, S., et al., *Large-scale analysis of HLA peptides presented by HLA-Cw4*. Immunogenetics, 2003. **55**(3): p. 172-6.
156. Steinle, A., et al., *Motif of HLA-B*3503 peptide ligands*. Immunogenetics, 1996. **43**(1-2): p. 105-7.

157. Dephoure, N., et al., *A quantitative atlas of mitotic phosphorylation*. Proc Natl Acad Sci U S A, 2008. **105**(31): p. 10762-7.
158. Damoc, E., et al., *Structural characterization of the human eukaryotic initiation factor 3 protein complex by mass spectrometry*. Mol Cell Proteomics, 2007. **6**(7): p. 1135-46.
159. Olsen, J.V., et al., *Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis*. Sci Signal, 2010. **3**(104): p. ra3.
160. Olsen, J.V., et al., *Global, in vivo, and site-specific phosphorylation dynamics in signaling networks*. Cell, 2006. **127**(3): p. 635-48.
161. Wang, B., et al., *Evaluation of the low-specificity protease elastase for large-scale phosphoproteome analysis*. Anal Chem, 2008. **80**(24): p. 9526-33.
162. Daub, H., et al., *Kinase-selective enrichment enables quantitative phosphoproteomics of the kinome across the cell cycle*. Mol Cell, 2008. **31**(3): p. 438-48.
163. Kyono, Y., et al., *Successive and selective release of phosphorylated peptides captured by hydroxy acid-modified metal oxide chromatography*. J Proteome Res, 2008. **7**(10): p. 4585-93.
164. Stokes, M.P., et al., *Profiling of UV-induced ATM/ATR signaling pathways*. Proc Natl Acad Sci U S A, 2007. **104**(50): p. 19855-60.
165. Molina, H., et al., *Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry*. Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2199-204.
166. Yu, C.T., et al., *Phosphorylation and stabilization of HURP by Aurora-A: implication of HURP as a transforming target of Aurora-A*. Mol Cell Biol, 2005. **25**(14): p. 5789-800.
167. Lovly, C.M., et al., *Regulation of Chk2 ubiquitination and signaling through autophosphorylation of serine 379*. Mol Cell Biol, 2008. **28**(19): p. 5874-85.
168. Rigbolt, K.T., et al., *System-wide temporal characterization of the proteome and phosphoproteome of human embryonic stem cell differentiation*. Sci Signal, 2011. **4**(164): p. rs3.
169. Carriere, A., et al., *Oncogenic MAPK signaling stimulates mTORC1 activity by promoting RSK-mediated raptor phosphorylation*. Curr Biol, 2008. **18**(17): p. 1269-77.
170. Ruse, C.I., et al., *Motif-specific sampling of phosphoproteomes*. J Proteome Res, 2008. **7**(5): p. 2140-50.
171. Brill, L.M., et al., *Phosphoproteomic analysis of human embryonic stem cells*. Cell Stem Cell, 2009. **5**(2): p. 204-13.
172. Gauci, S., et al., *Lys-N and trypsin cover complementary parts of the phosphoproteome in a refined SCX-based approach*. Anal Chem, 2009. **81**(11): p. 4493-501.
173. Lees-Miller, S.P. and C.W. Anderson, *Two human 90-kDa heat shock proteins are phosphorylated in vivo at conserved serines that are phosphorylated in vitro by casein kinase II*. J Biol Chem, 1989. **264**(5): p. 2431-7.
174. Bian, Y., et al., *An enzyme assisted RP-RPLC approach for in-depth analysis of human liver phosphoproteome*. J Proteomics, 2014. **96**: p. 253-62.
175. Van Hoof, D., et al., *Phosphorylation dynamics during early differentiation of human embryonic stem cells*. Cell Stem Cell, 2009. **5**(2): p. 214-26.

176. Jensen, L.J., et al., *STRING 8--a global view on proteins and their functional interactions in 630 organisms*. Nucleic Acids Res, 2009. **37**(Database issue): p. D412-6.
177. Nakajima, H., et al., *Identification of a consensus motif for Plk (Polo-like kinase) phosphorylation reveals Myt1 as a Plk1 substrate*. J Biol Chem, 2003. **278**(28): p. 25277-80.
178. Pedraza-Alva, G., et al., *Regulation of Cbl molecular interactions by the co-receptor molecule CD43 in human T cells*. J Biol Chem, 2001. **276**(1): p. 729-37.
179. Huttlin, E.L., et al., *A tissue-specific atlas of mouse protein phosphorylation and expression*. Cell, 2010. **143**(7): p. 1174-89.
180. Marcilla, M., et al., *Comparative Analysis of the Endogenous Peptidomes Displayed by HLA-B*27 and Mamu-B*08: Two MHC Class I Alleles Associated with Elite Control of HIV/SIV Infection*. J Proteome Res, 2016. **15**(3): p. 1059-69.
181. Legrain, P., et al., *The human proteome project: Current state and future direction*. Mol Cell Proteomics, 2011.
182. Admon, A. and M. Bassani-Sternberg, *The Human Immunopeptidome Project, a suggestion for yet another postgenome next big thing*. Mol Cell Proteomics, 2011. **10**(10): p. O111 011833.
183. Trolle, T., et al., *The Length Distribution of Class I-Restricted T Cell Epitopes Is Determined by Both Peptide Supply and MHC Allele-Specific Binding Preference*. J Immunol, 2016. **196**(4): p. 1480-7.
184. Probst-Kepper, M., et al., *Conformational restraints and flexibility of 14-meric peptides in complex with HLA-B*3501*. J Immunol, 2004. **173**(9): p. 5610-6.
185. Tynan, F.E., et al., *A T cell receptor flattens a bulged antigenic peptide presented by a major histocompatibility complex class I molecule*. Nat Immunol, 2007. **8**(3): p. 268-76.
186. Ben Dror, L., et al., *The HLA-B*2705 peptidome*. Arthritis Rheum, 2010. **62**(2): p. 420-9.
187. Lopez de Castro, J.A., et al., *HLA-B27: a registry of constitutive peptide ligands*. Tissue Antigens, 2004. **63**(5): p. 424-45.
188. Wang, M., et al., *Current advances in T-cell-based cancer immunotherapy*. Immunotherapy, 2014. **6**(12): p. 1265-78.
189. Kessler, J.H. and C.J. Melief, *Identification of T-cell epitopes for cancer immunotherapy*. Leukemia, 2007. **21**(9): p. 1859-74.
190. Zarling, A.L., et al., *MHC-restricted phosphopeptides from insulin receptor substrate-2 and CDC25b offer broad-based immunotherapeutic agents for cancer*. Cancer Res, 2014. **74**(23): p. 6784-95.
191. Mommen, G.P., et al., *Sampling from the proteome to the HLA-DR ligandome proceeds via high specificity*. Mol Cell Proteomics, 2016.
192. Marino, F., et al., *Extended O-GlcNAc on HLA Class-I-Bound Peptides*. J Am Chem Soc, 2015. **137**(34): p. 10922-5.
193. Weigel, N.L. and N.L. Moore, *Kinases and protein phosphorylation as regulators of steroid hormone action*. Nucl Recept Signal, 2007. **5**: p. e005.
194. Evnouchidou, I., et al., *The internal sequence of the peptide-substrate determines its N-terminus trimming by ERAP1*. PLoS One, 2008. **3**(11): p. e3658.

195. Merino, E., et al., *Two HLA-B14 subtypes (B*1402 and B*1403) differentially associated with ankylosing spondylitis differ substantially in peptide specificity but have limited peptide and T-cell epitope sharing with HLA-B27*. J Biol Chem, 2005. **280**(43): p. 35868-80.
196. Gomez, P., et al., *Presentation of cytosolically stable peptides by HLA-B27 is not dependent on the canonic interactions of N-terminal basic residues in the A pocket*. J Immunol, 2009. **182**(1): p. 446-55.

Anexos

Anexo 1

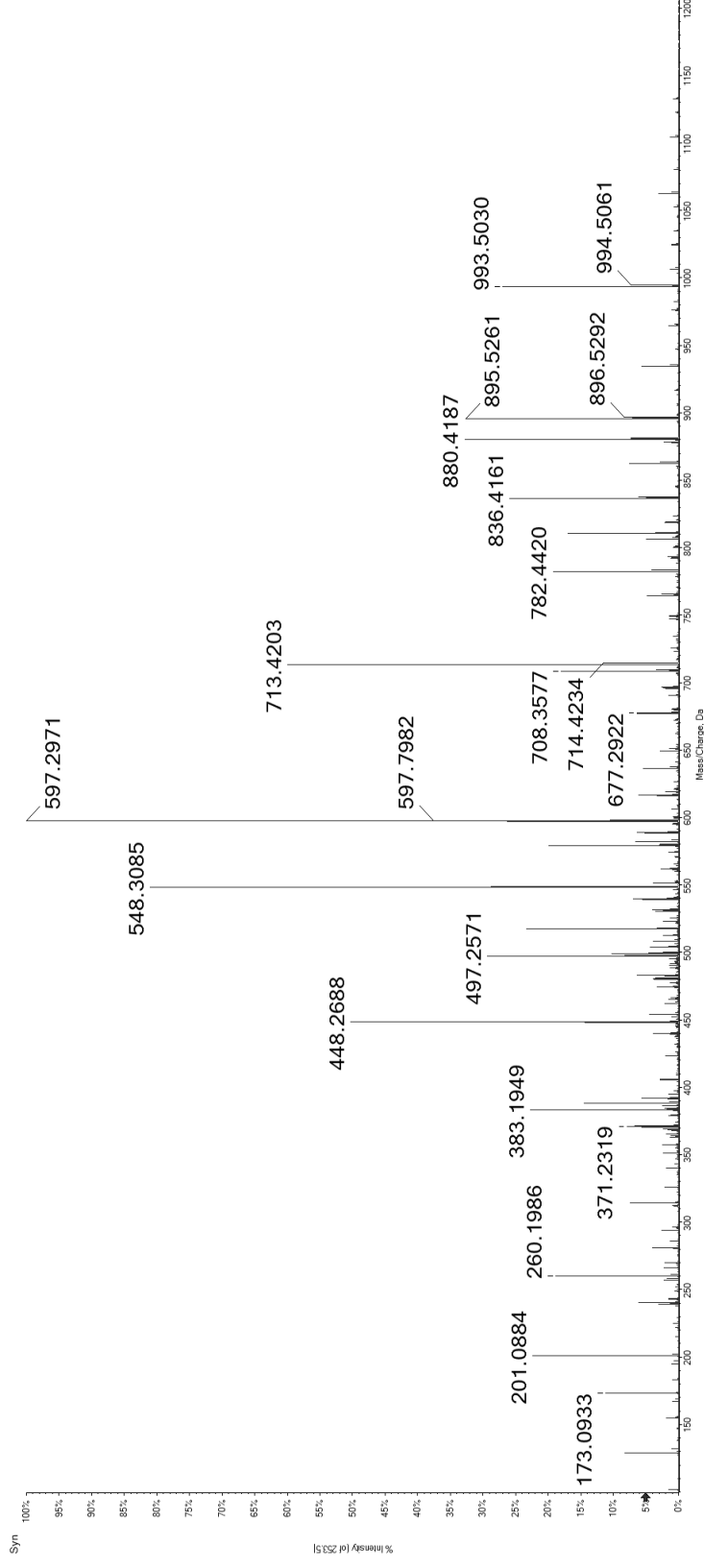
Anexo 1.1: Espectros MS2 de los fosfopéptidos identificados mediante CID.

Cada espectro MS2 candidato se compara con el espectro de fragmentación del péptido sintético correspondiente. Las comparaciones están ordenadas como se muestran en la Tabla 7. Los puntos rojos representan picos de iones que no están presentes en alguno de los dos espectros. Para confirmar la identidad del péptido en estos casos se comparó el tiempo de retención de ambas especies.

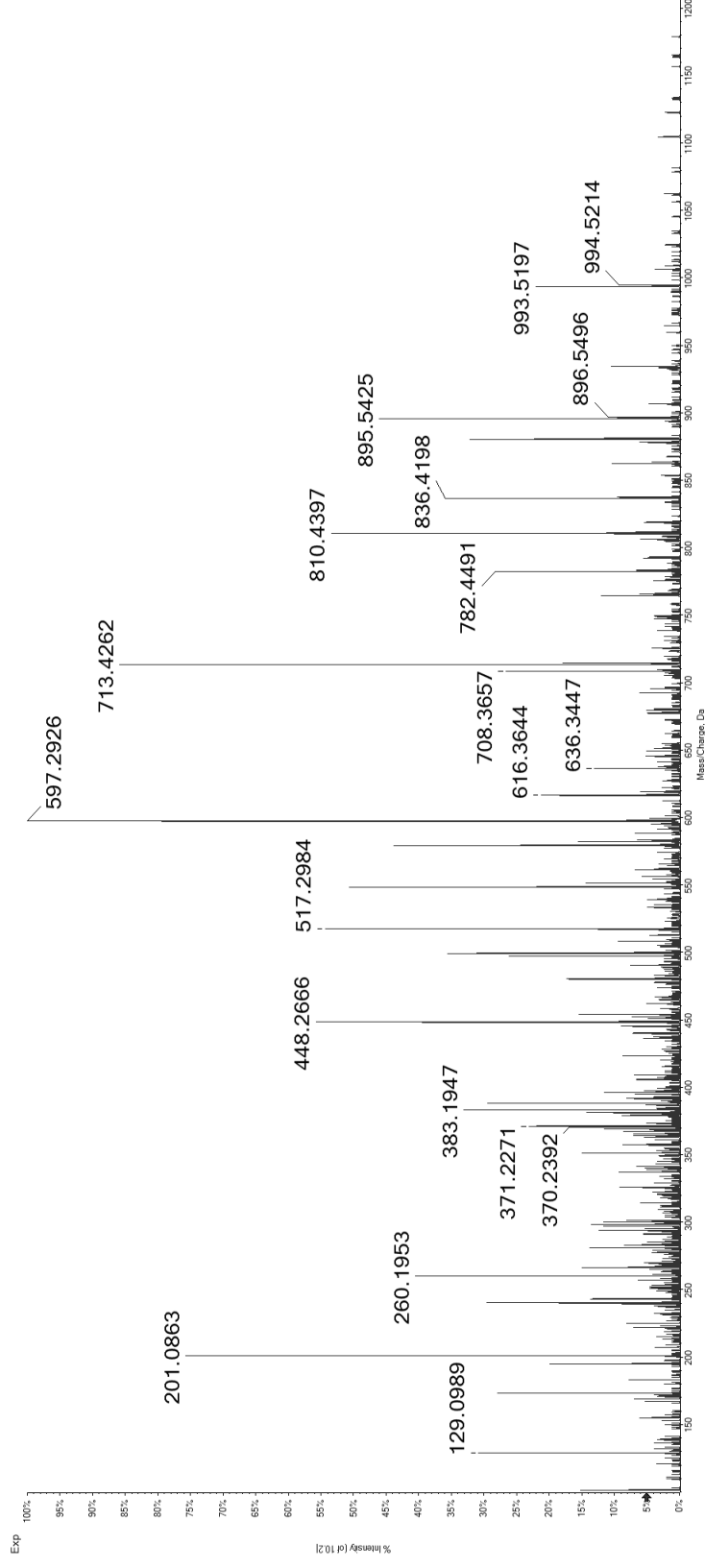
1

AEL(pS)PVEQKL

Synthetic Peptide



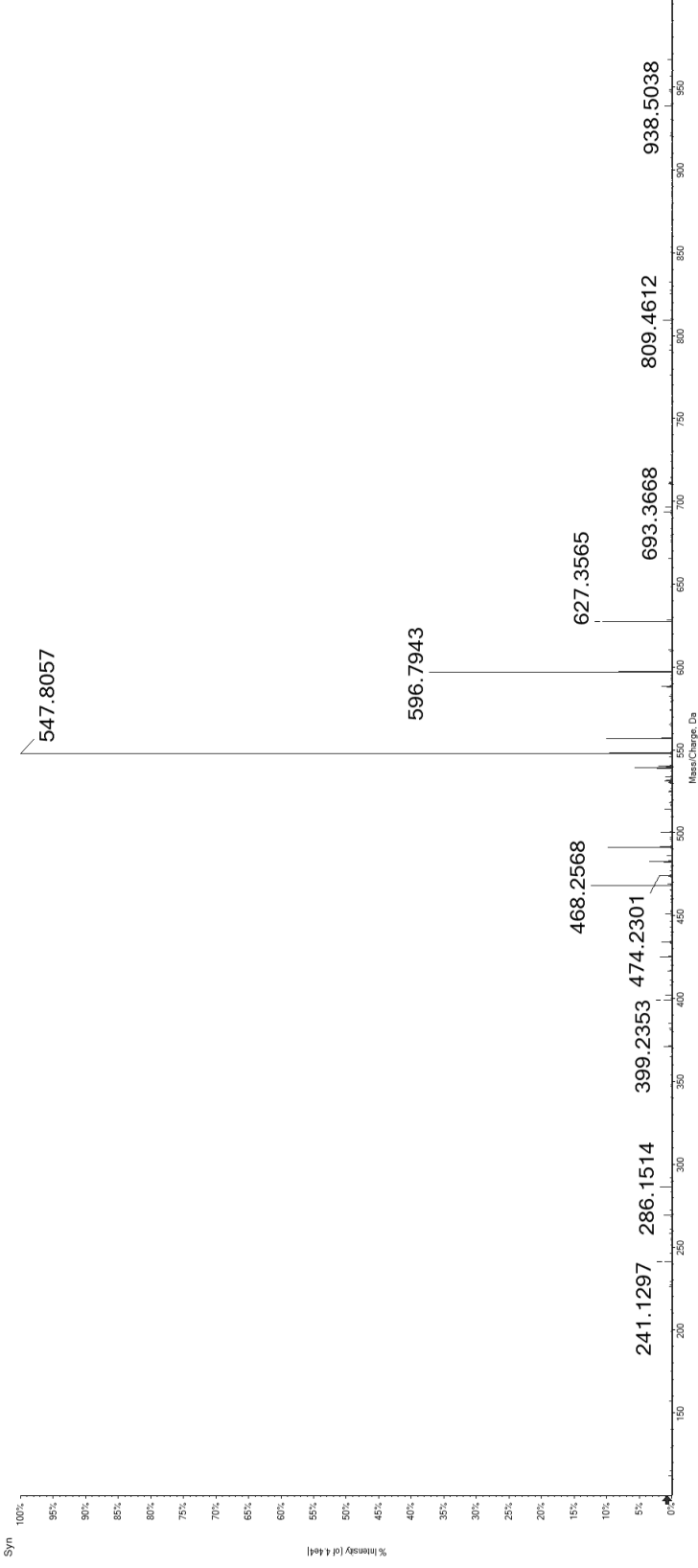
Candidate Spectrum



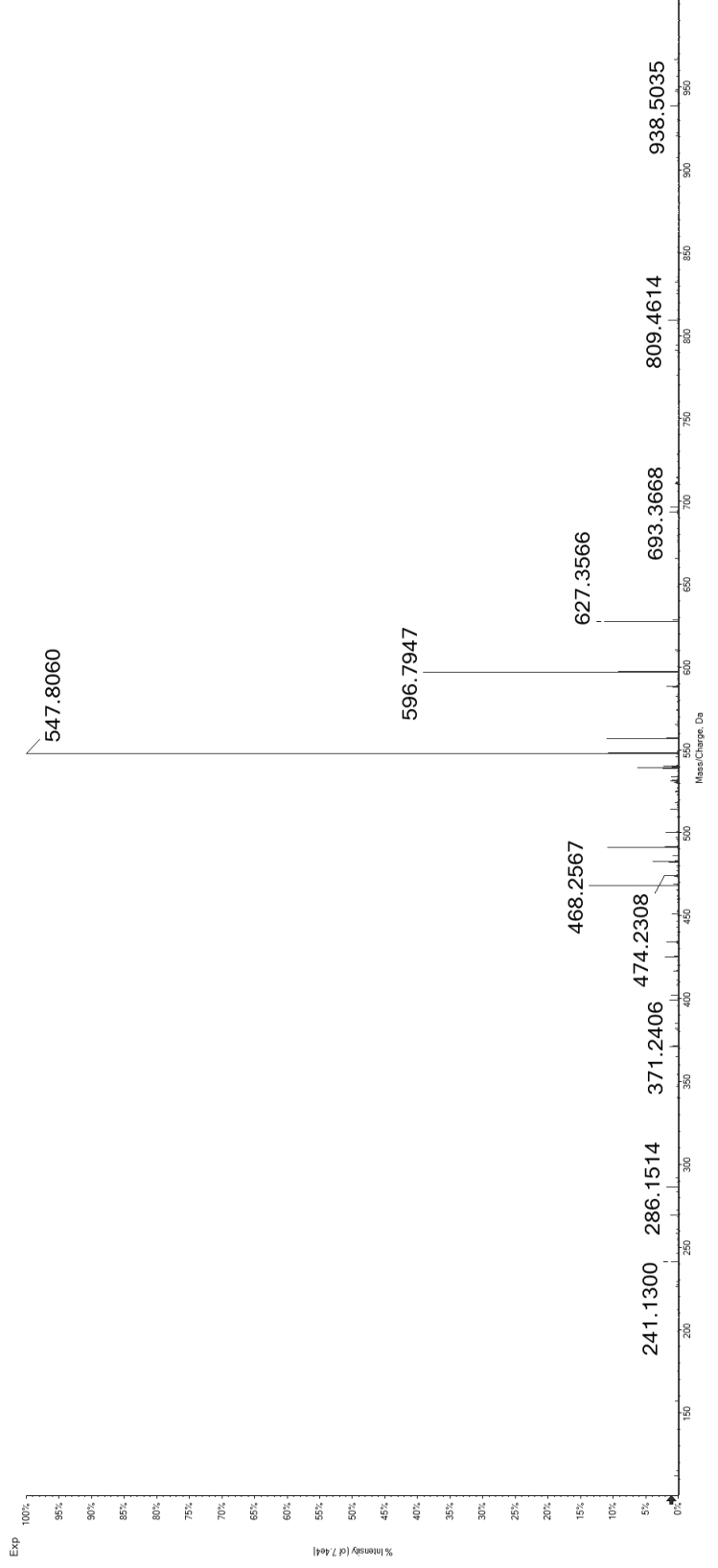
2

REL(pS)PQRNL

Synthetic Peptide



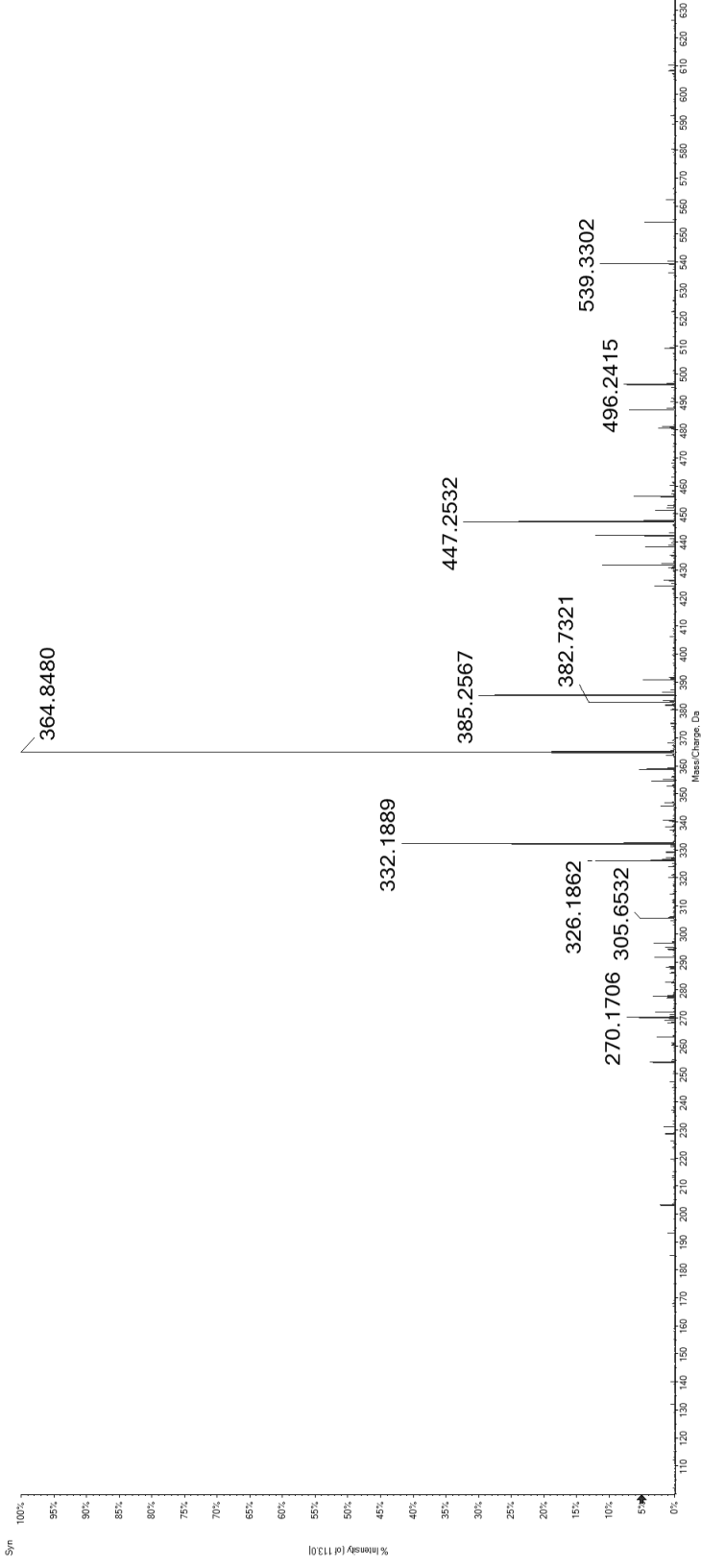
Candidate Spectrum



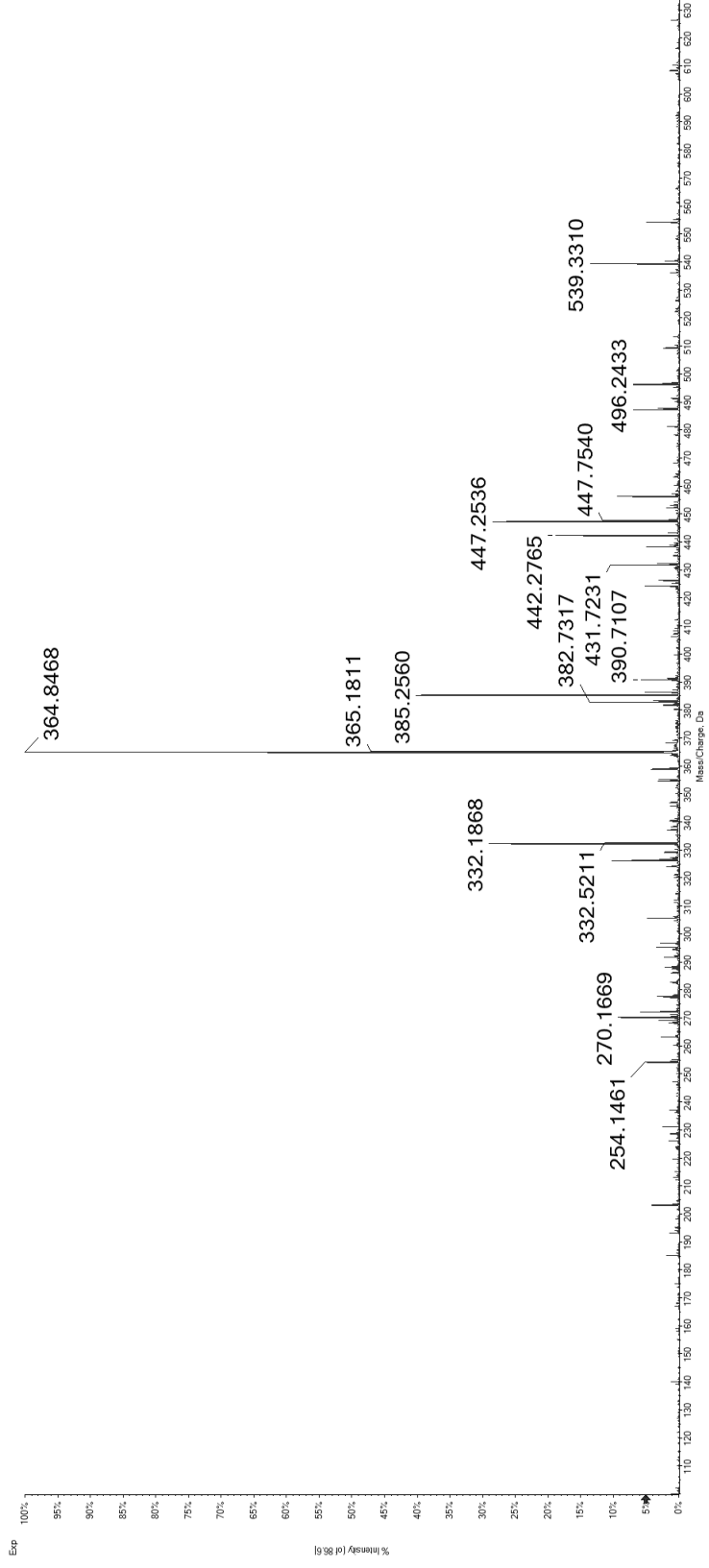
3

TER(pS)PGPRL

Synthetic
Peptide

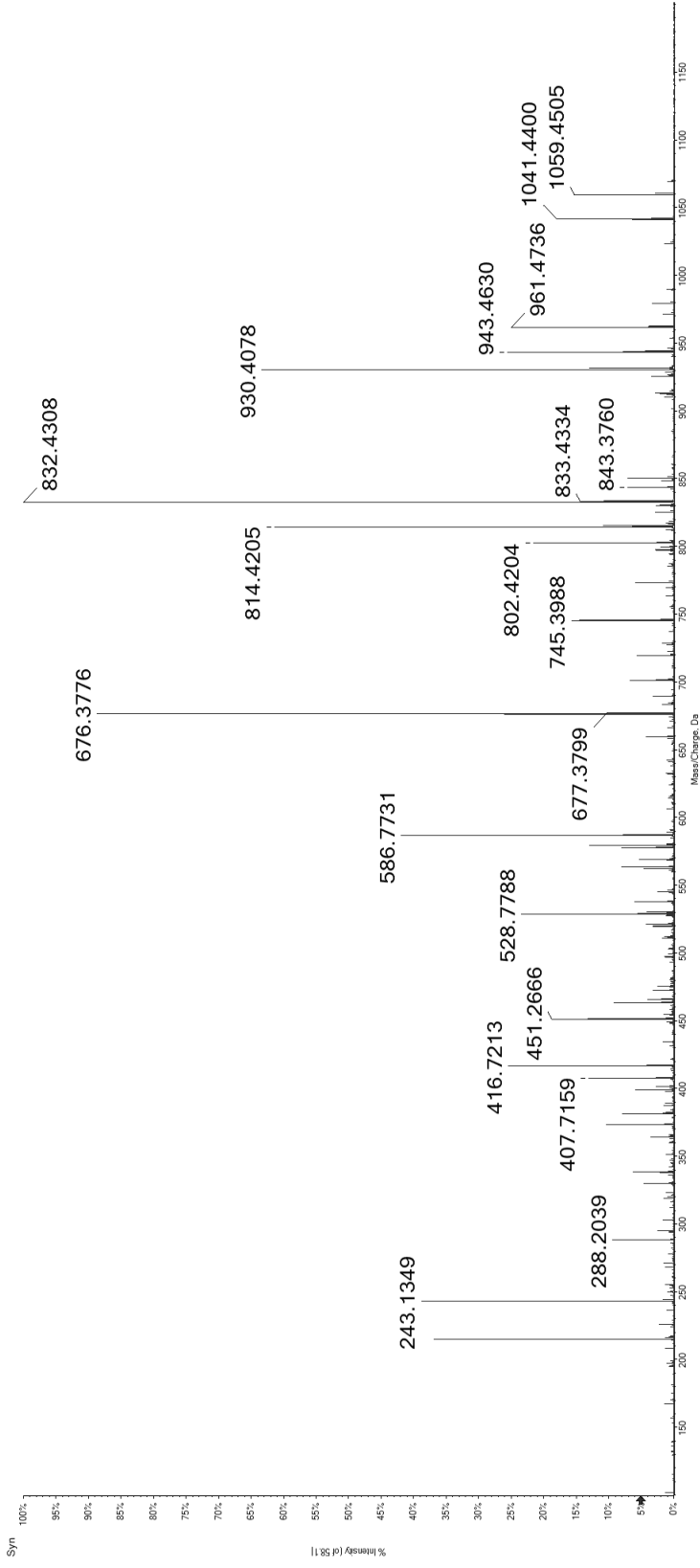


Candidate
Spectrum

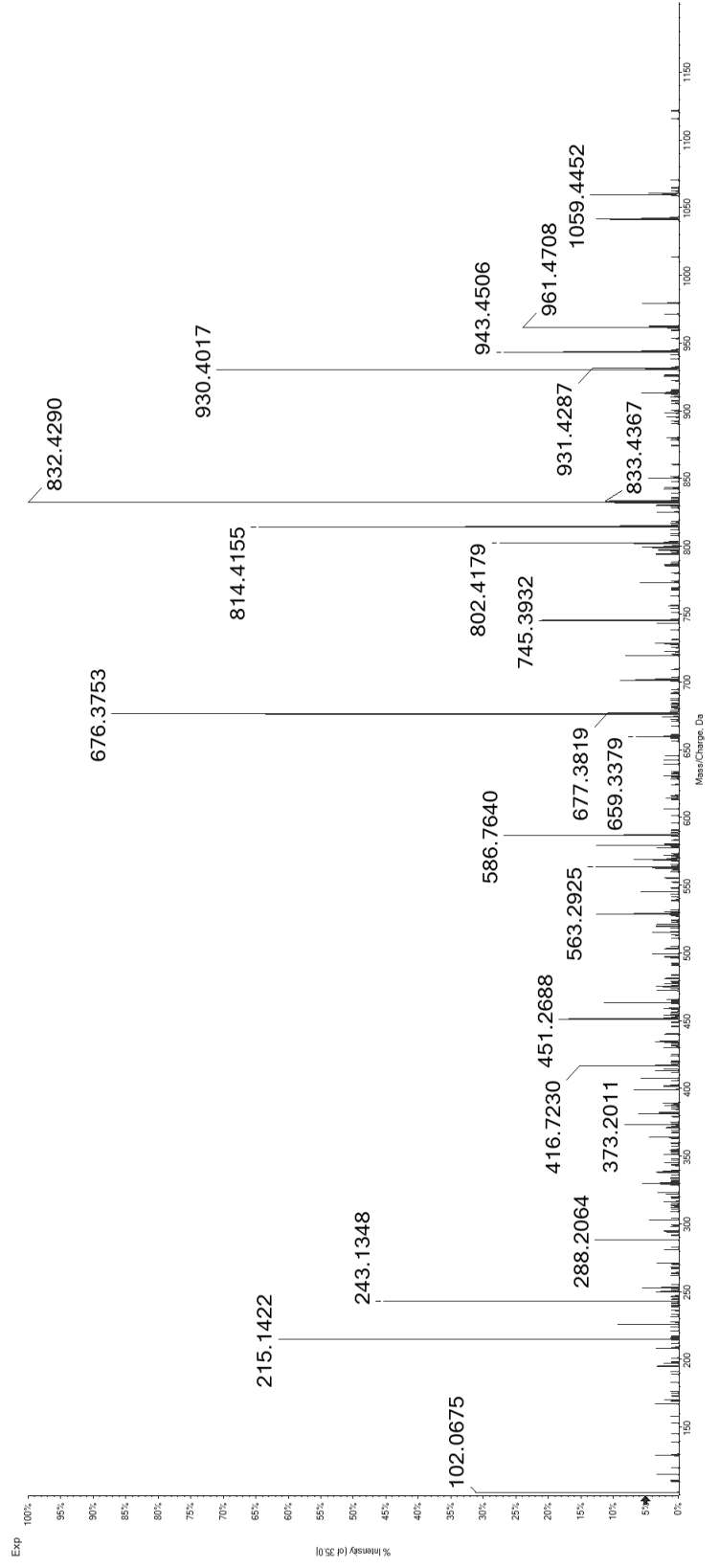


IES(ps)PQYRL

Synthetic Peptide



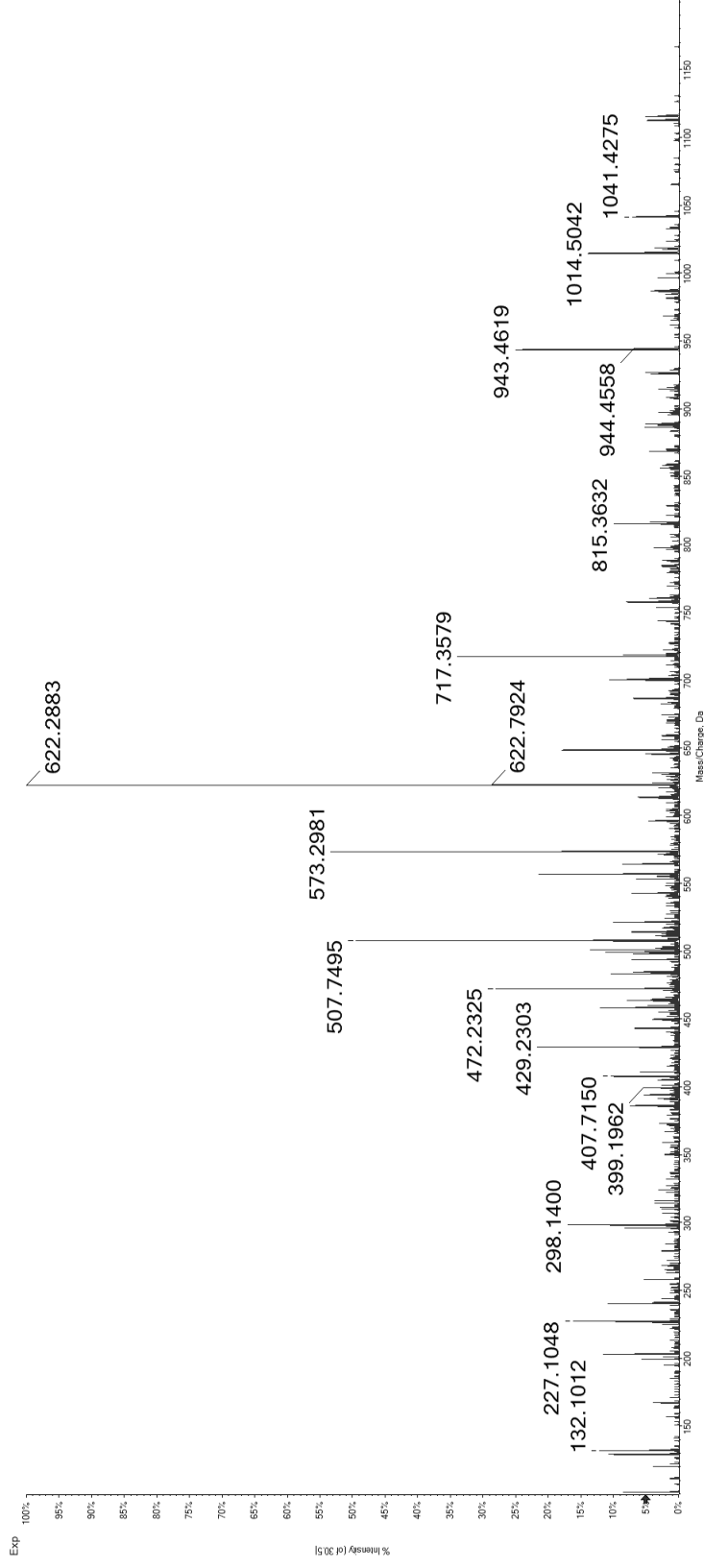
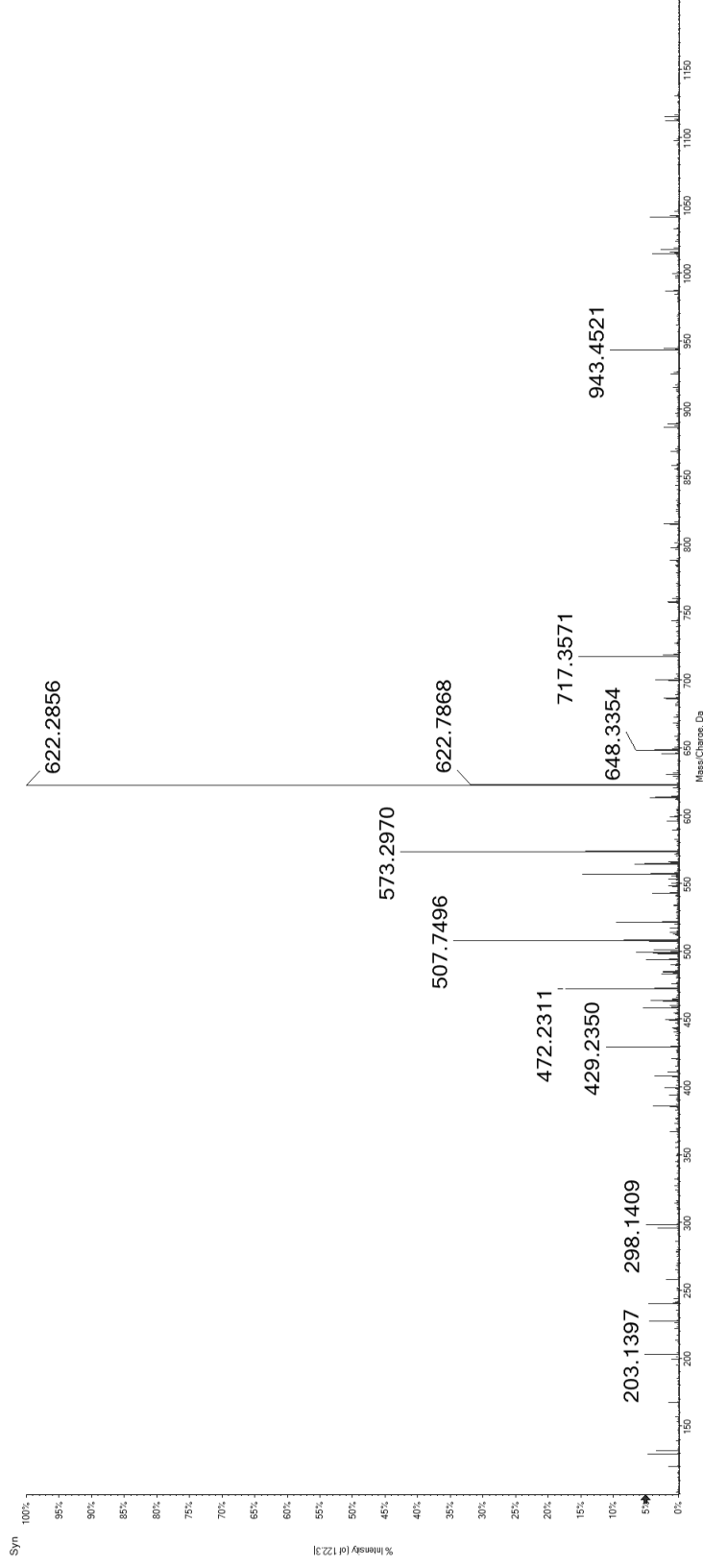
Candidate Spectrum



5

KEKDF(pS)PEAL

Synthetic
Peptide

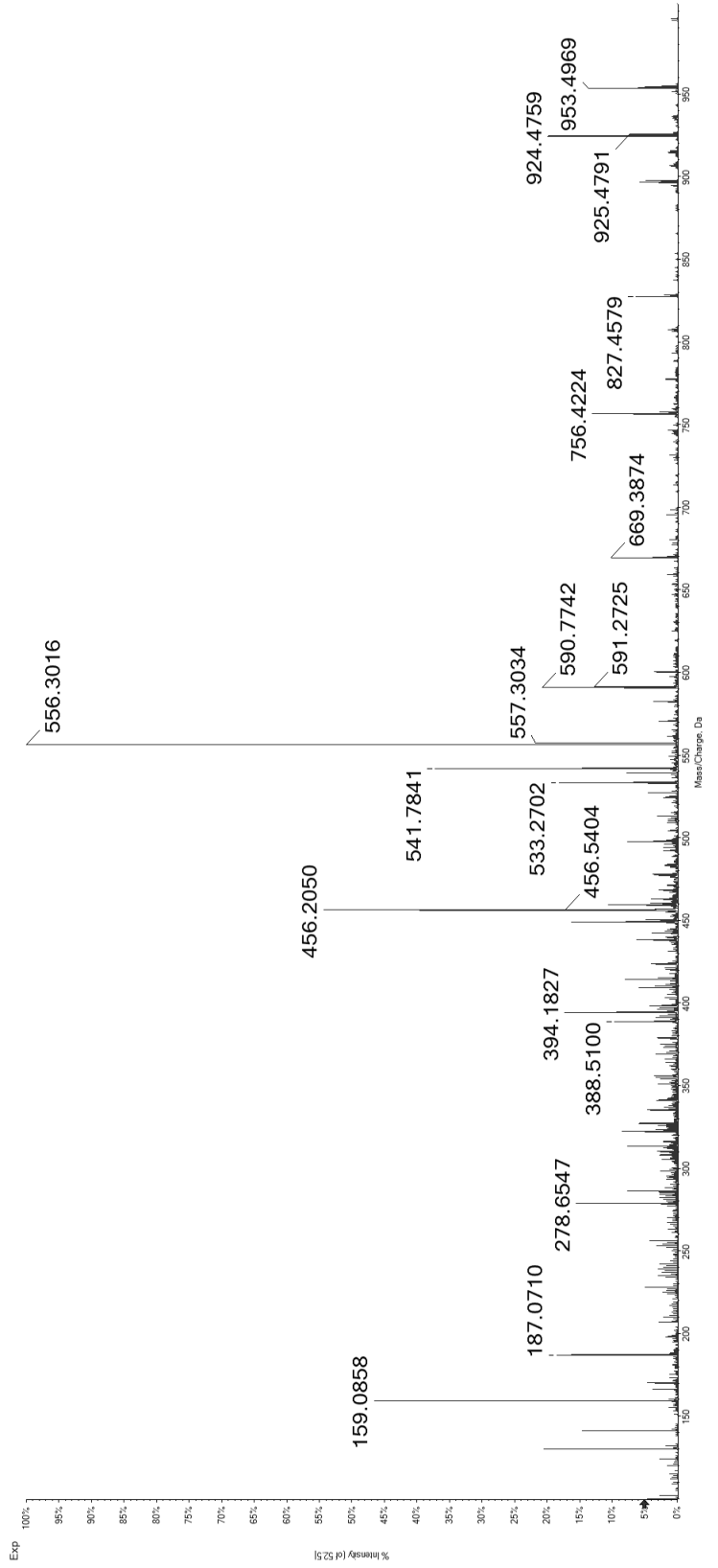
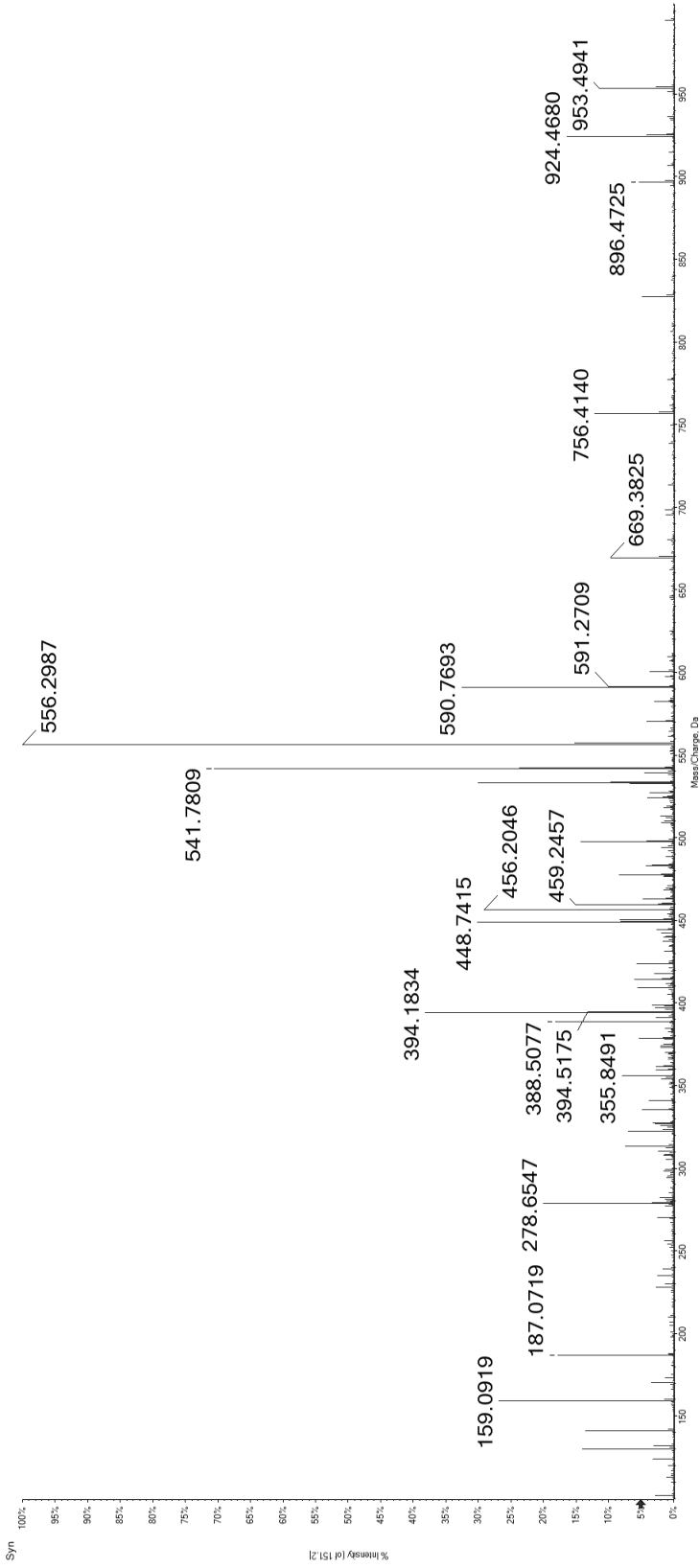


Candidate
Spectrum

6

GEW(ps)ASLPHRF

Synthetic Peptide

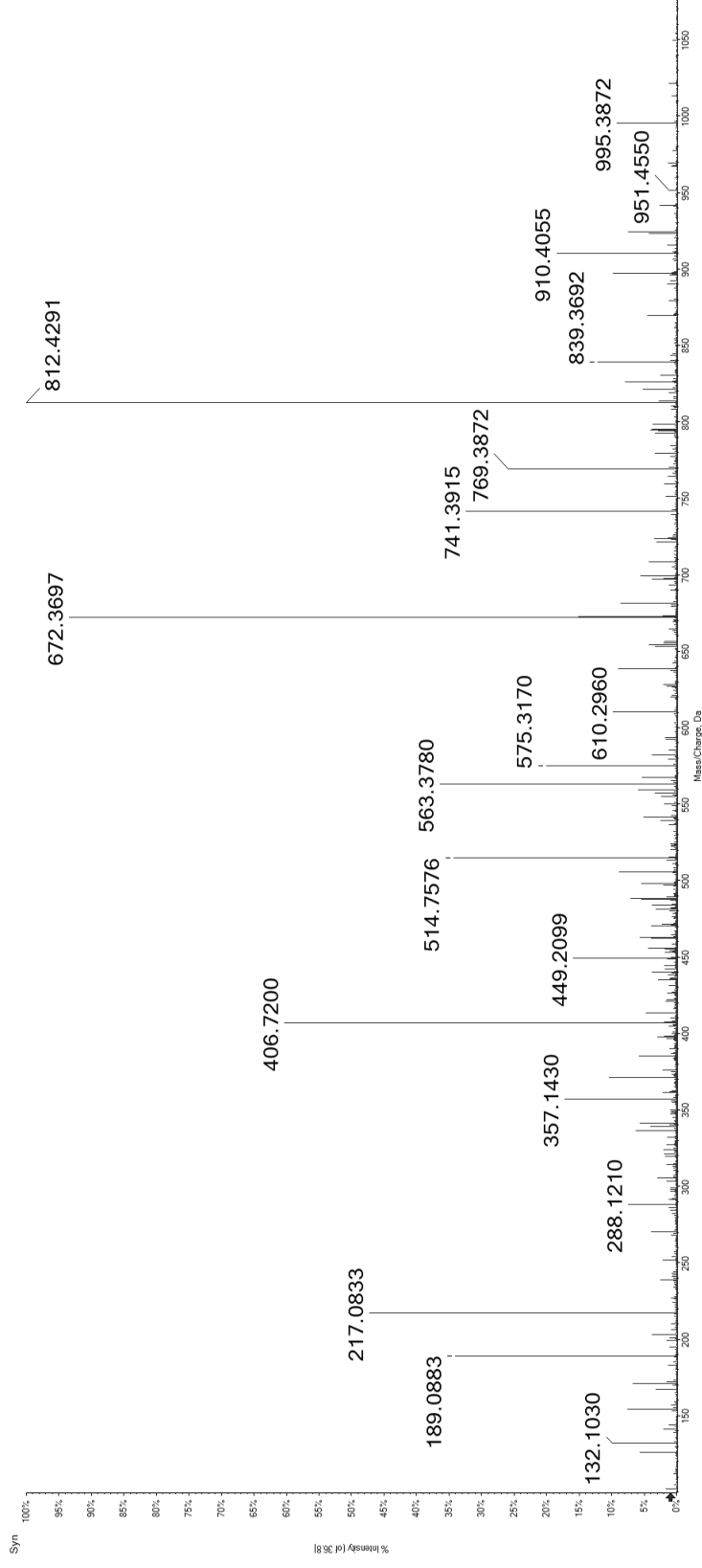


Candidate Spectrum

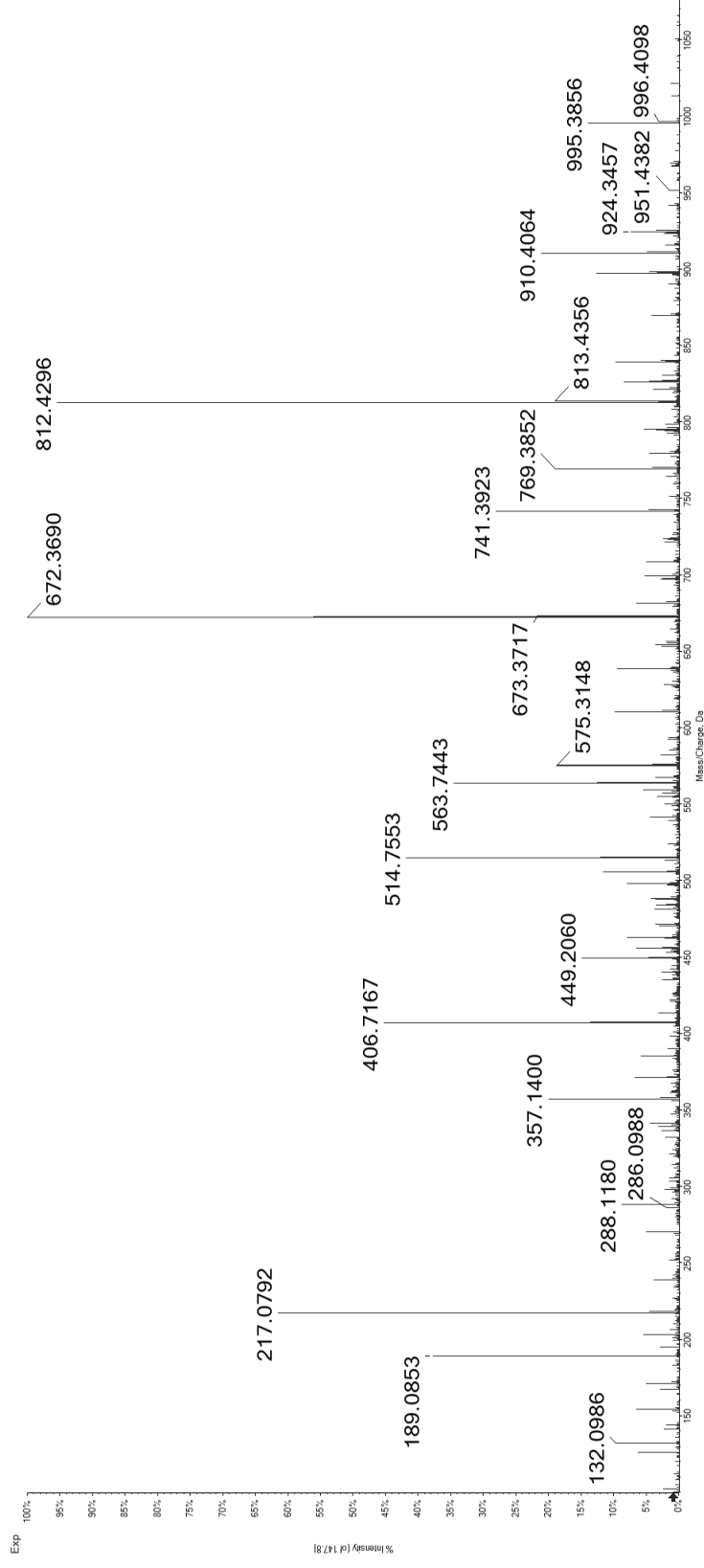
7

SEA(pS)PSREALI

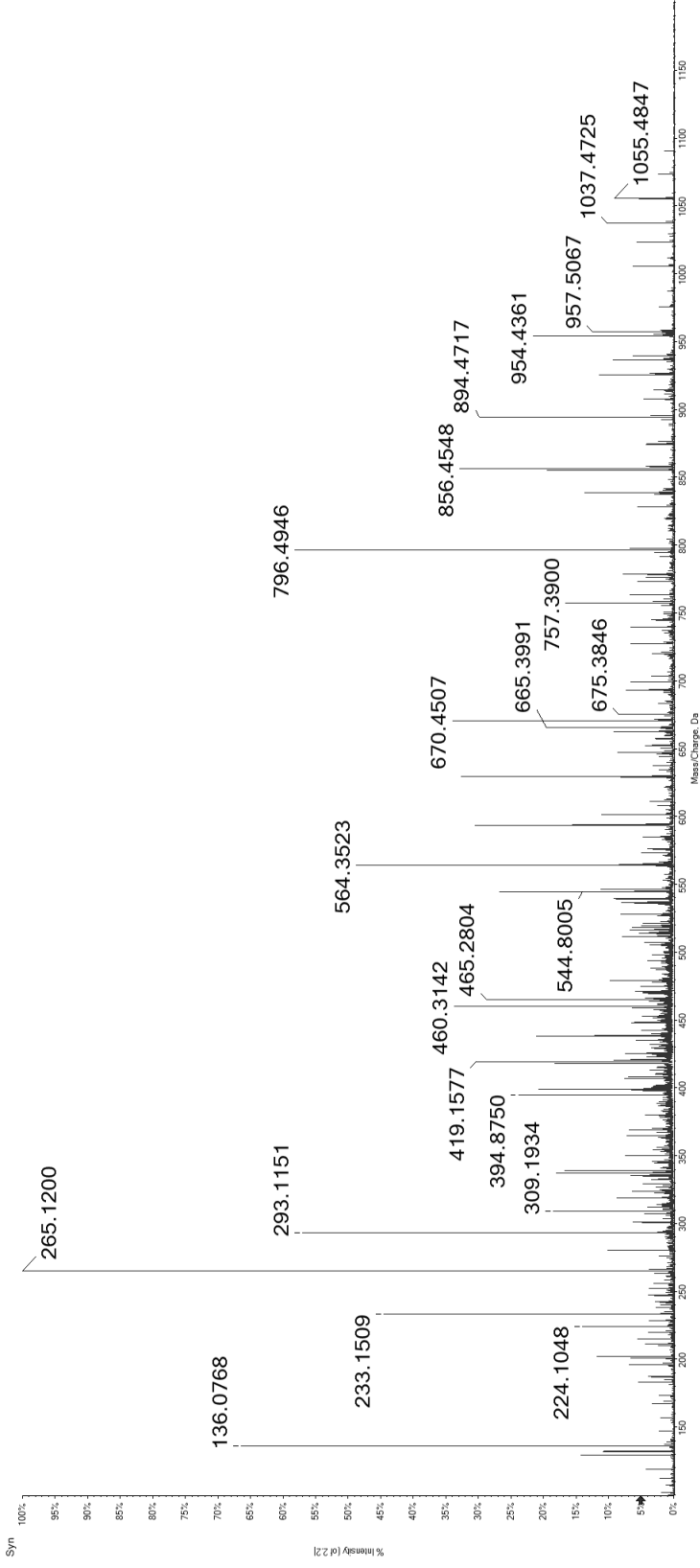
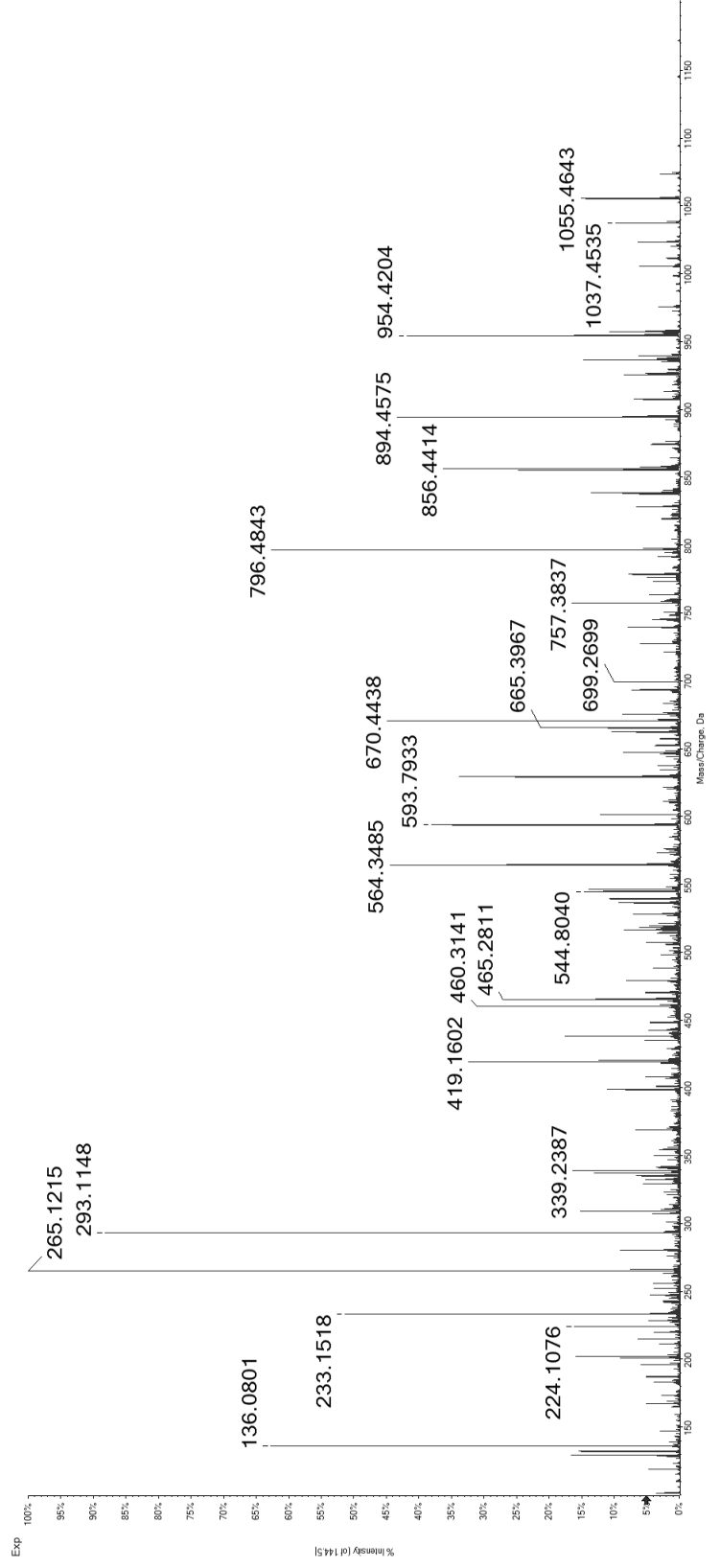
Synthetic Peptide



Candidate Spectrum



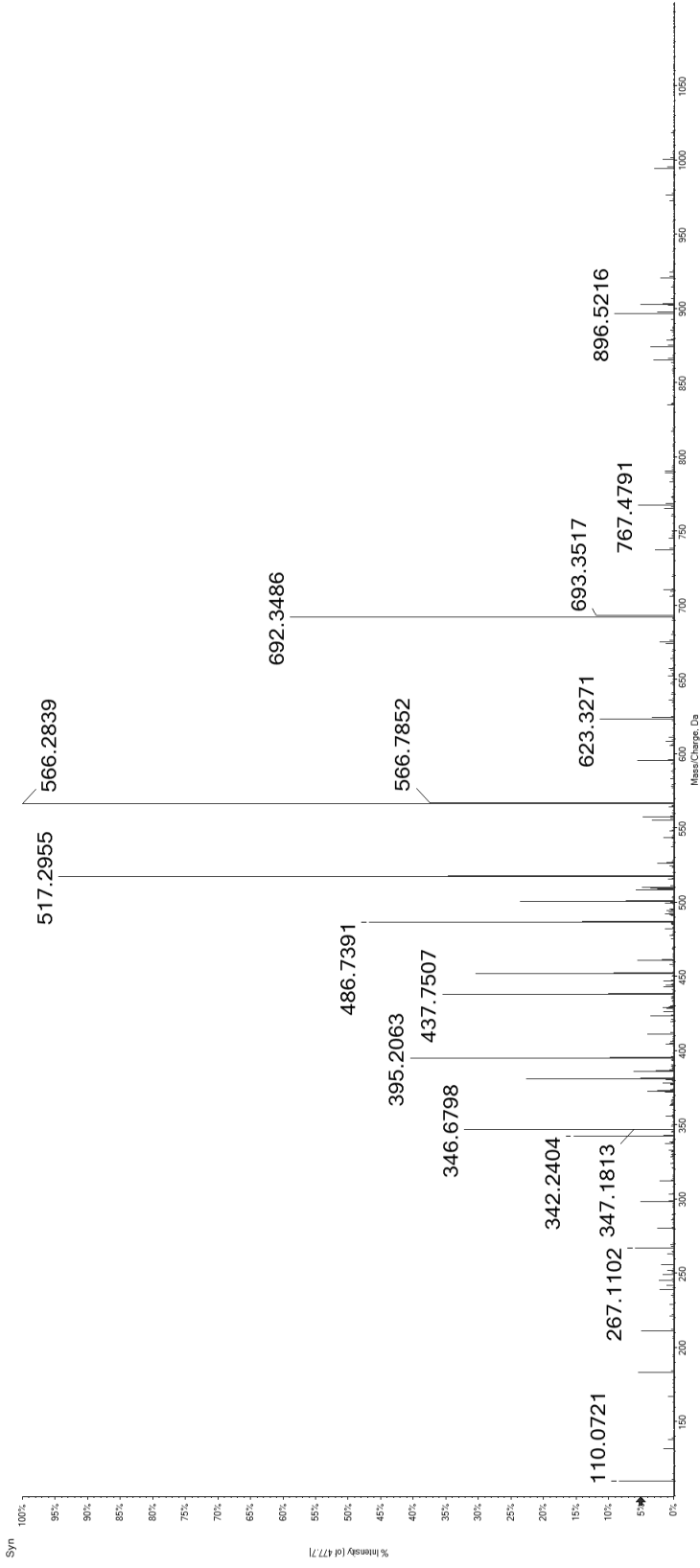
YEG(pS)PIKVTL

Synthetic
PeptideCandidate
Spectrum

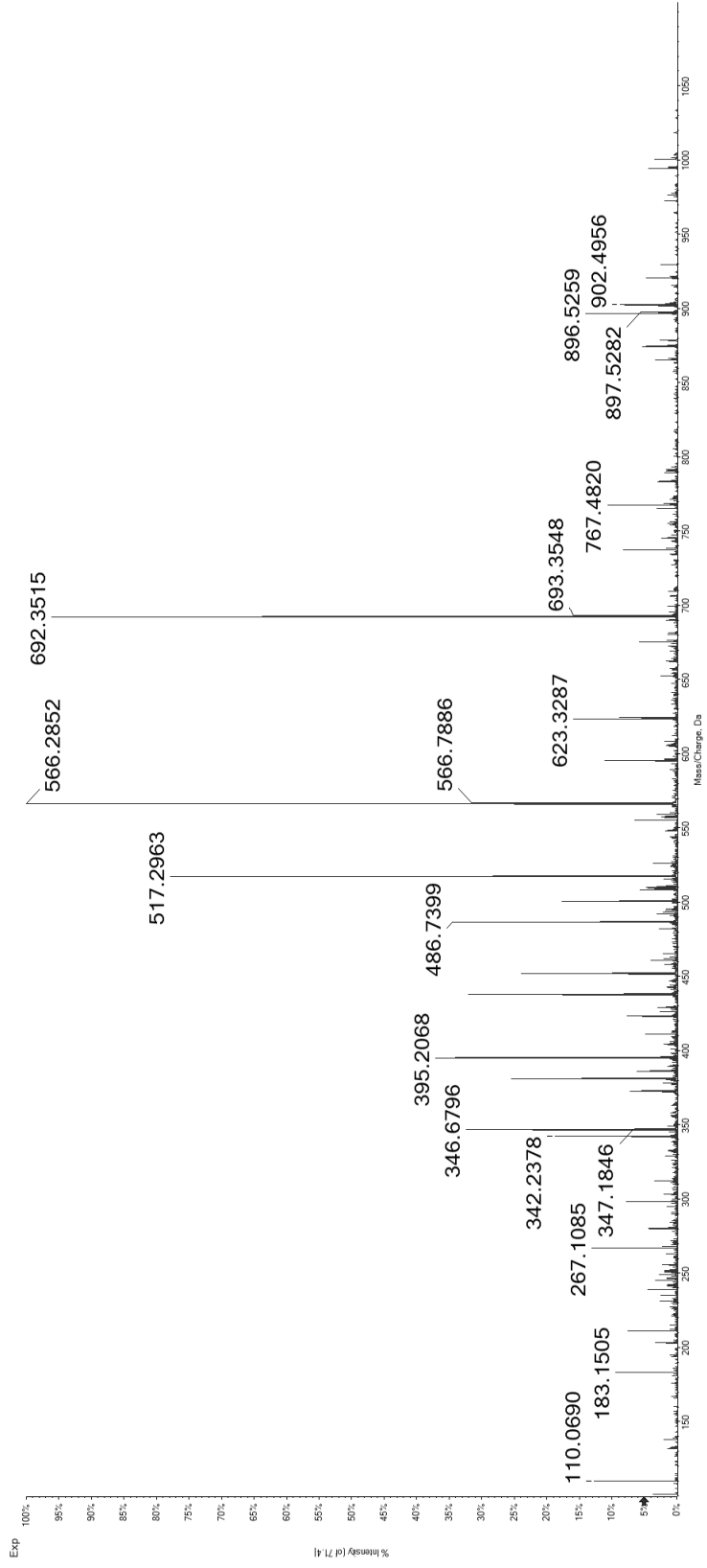
9

HERSI(ps)PLL

Synthetic Peptide

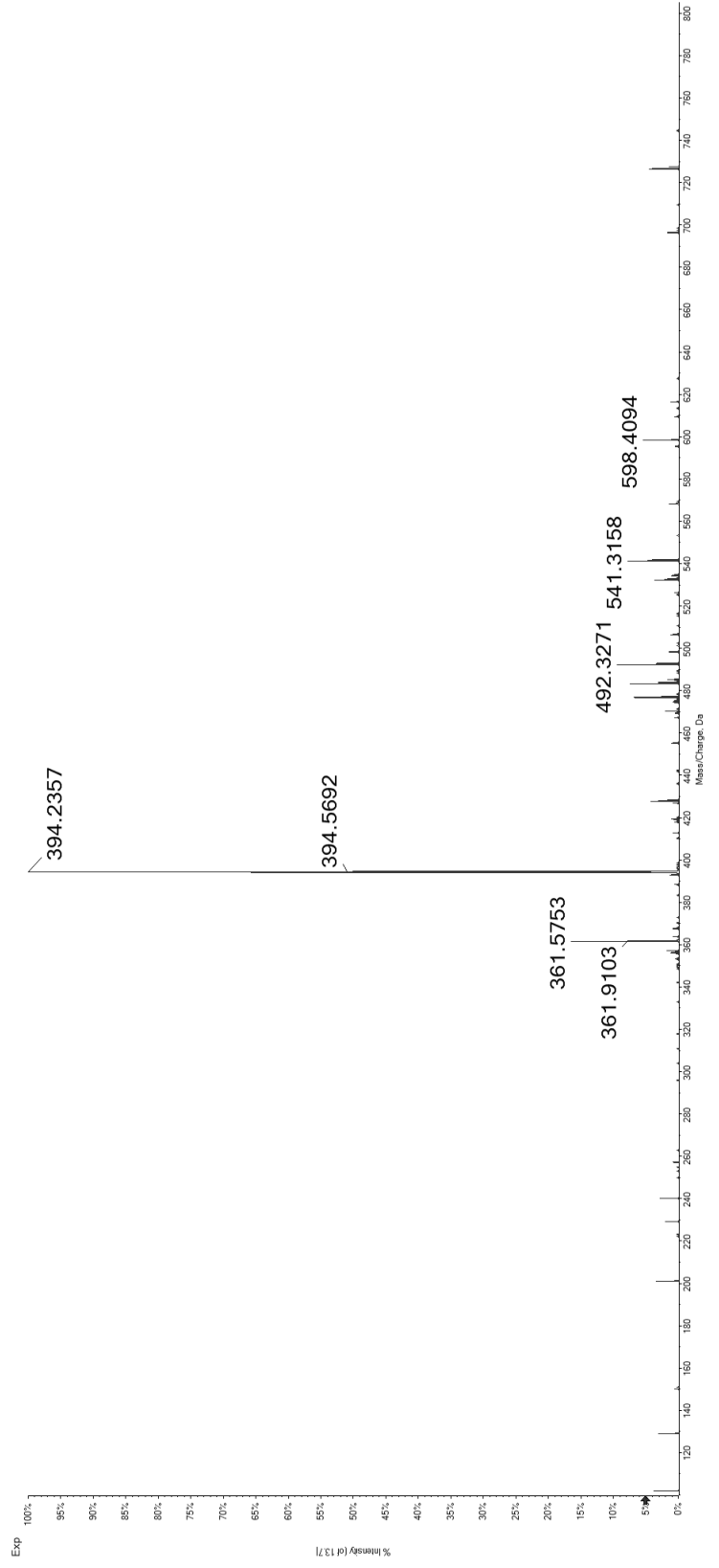
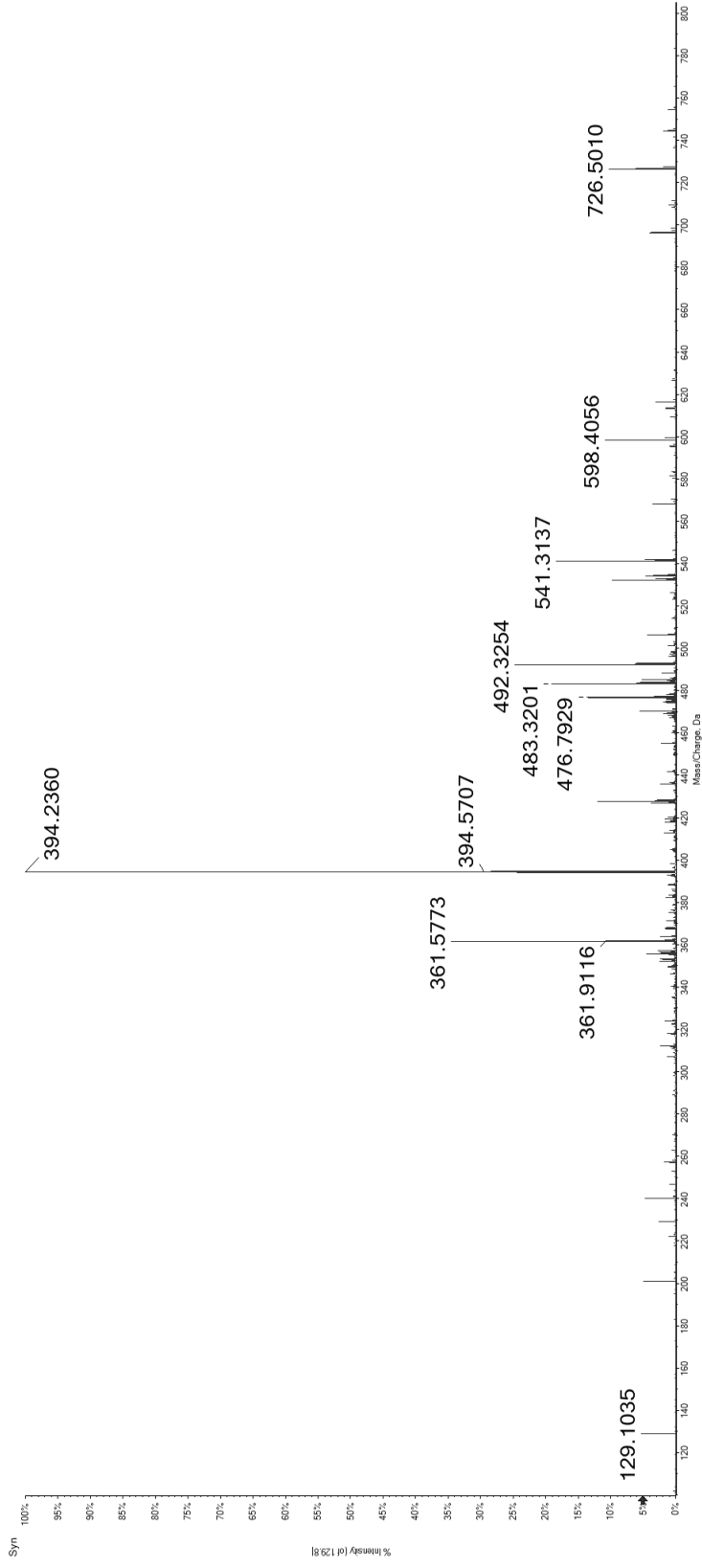


Candidate Spectrum



10 VEKKKIR(ps)L

Synthetic
Peptide

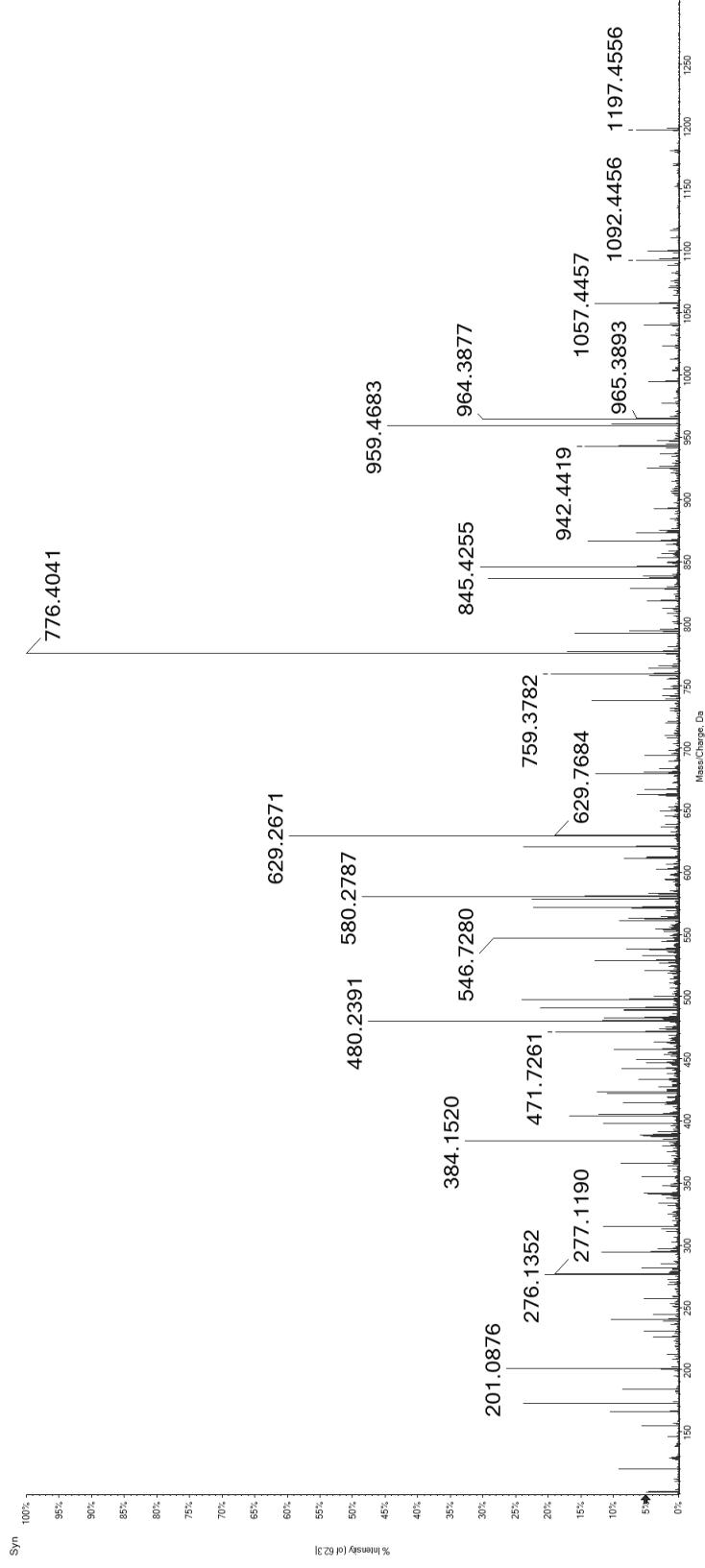


Candidate
Spectrum

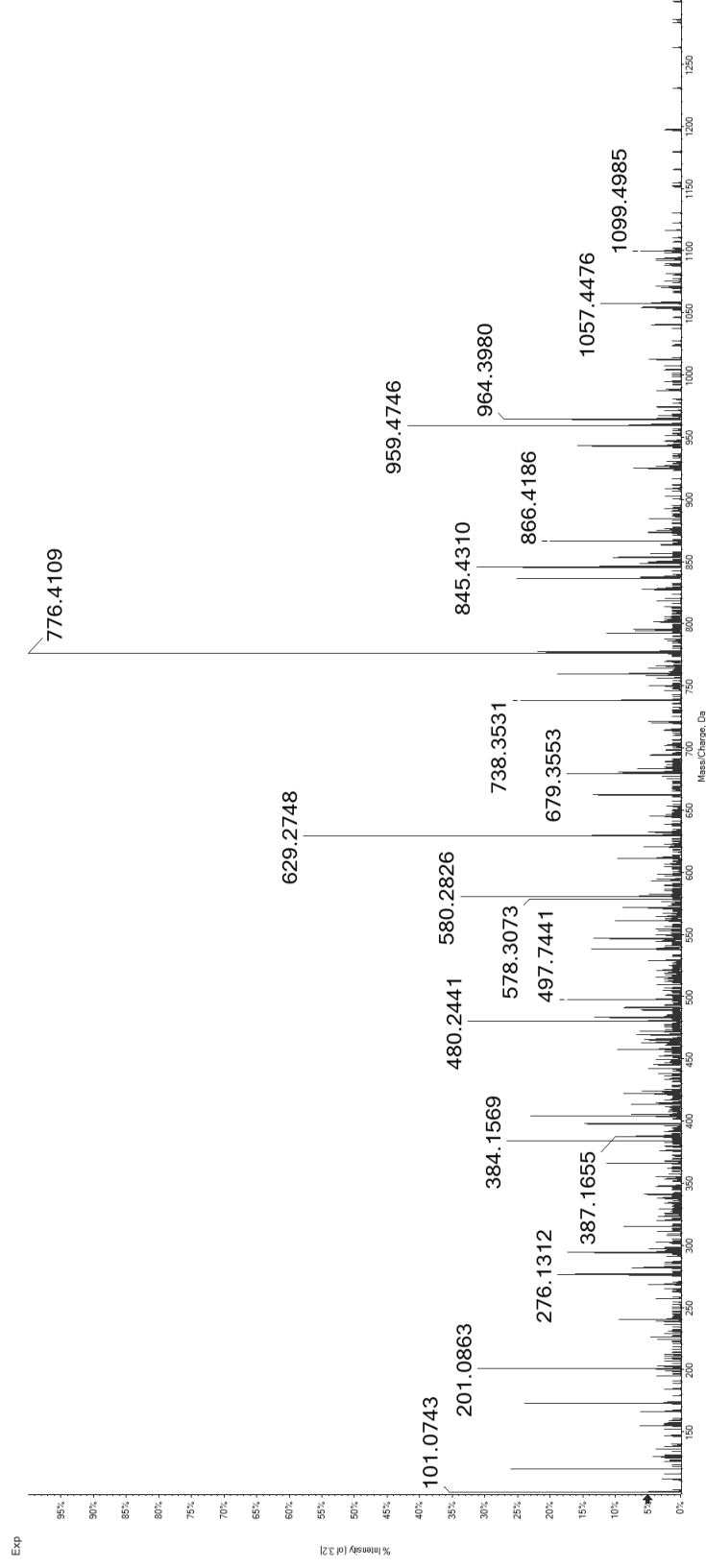
11

AEN(ps)PTRQQF

Synthetic Peptide



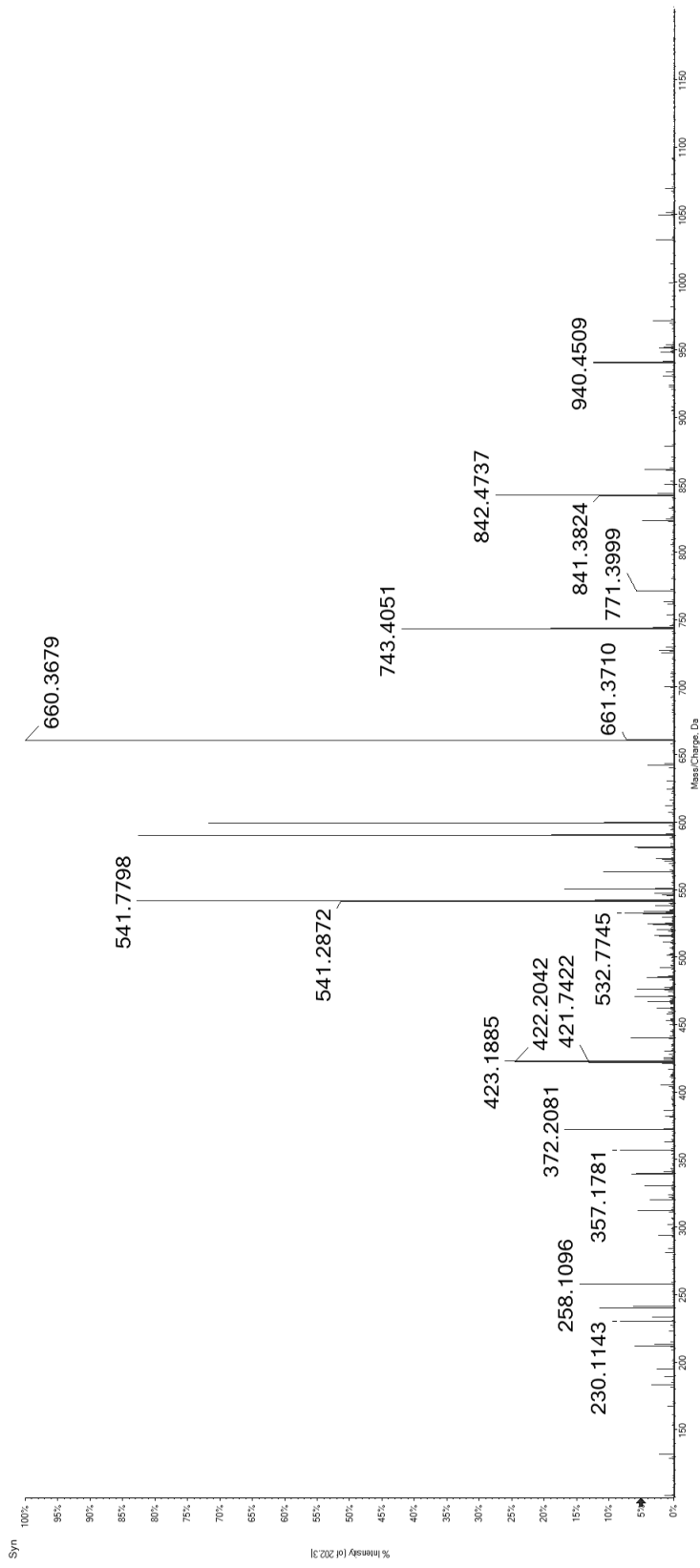
Candidate Spectrum



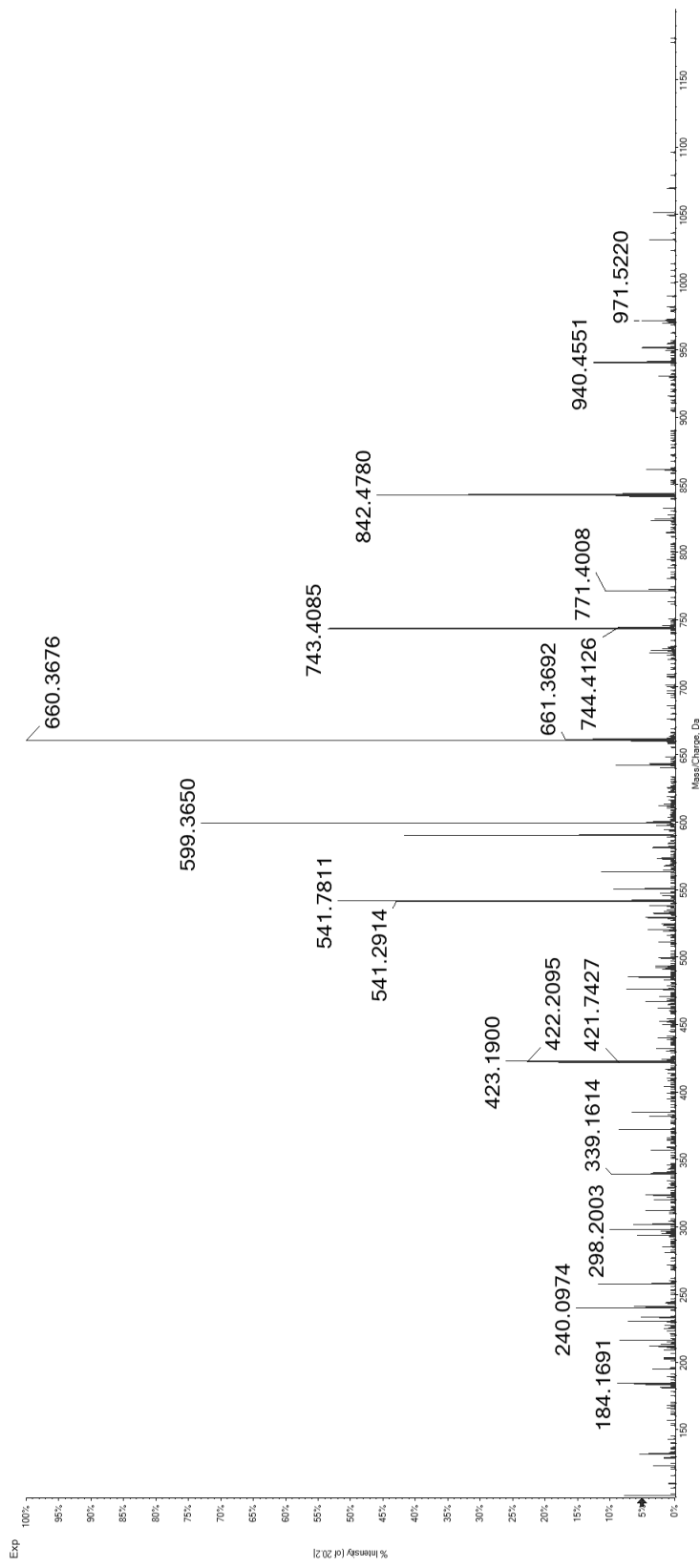
12

QEV(pT)PSRSTL

Synthetic Peptide

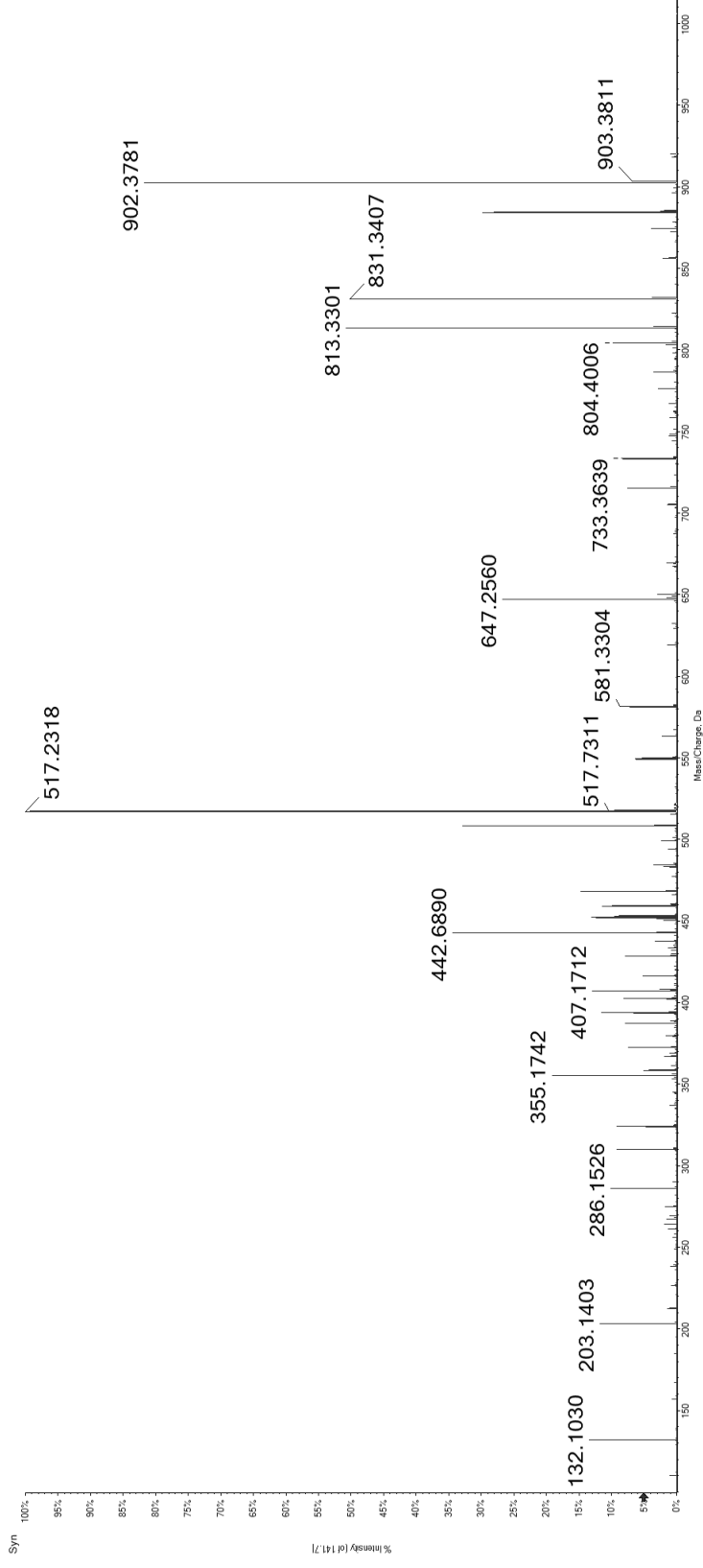


Candidate Spectrum

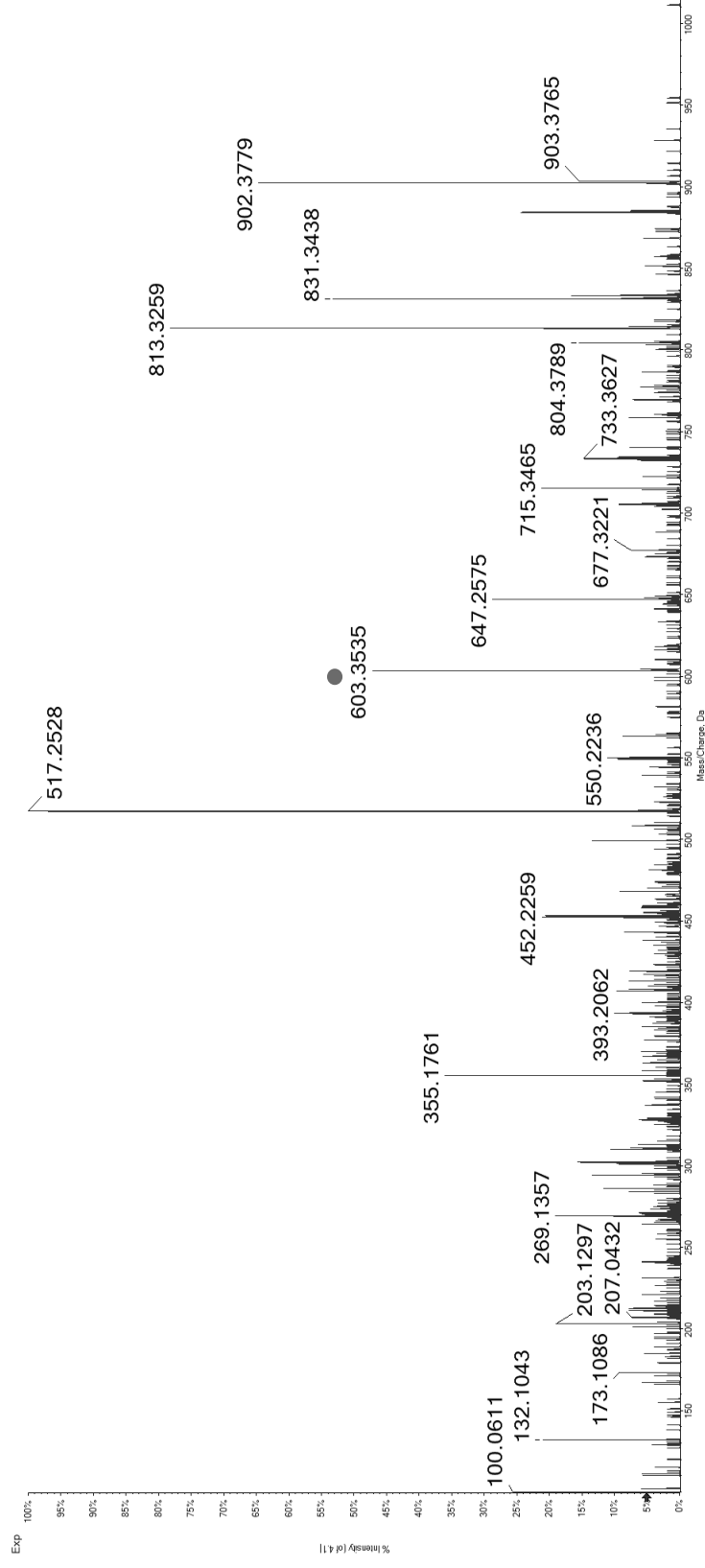


13 RE(pS)PPPSAL

Synthetic
Peptide
(Relative RT – 1,13)



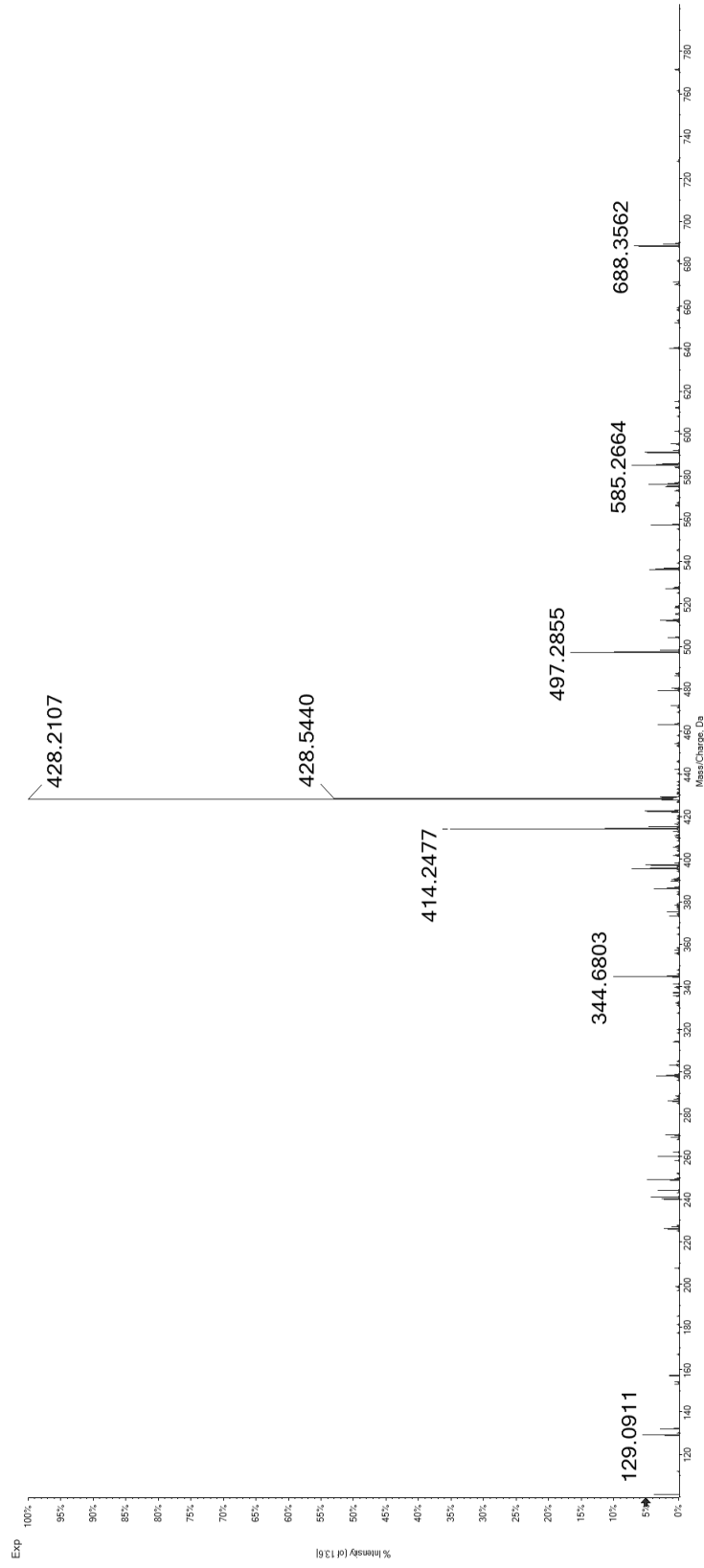
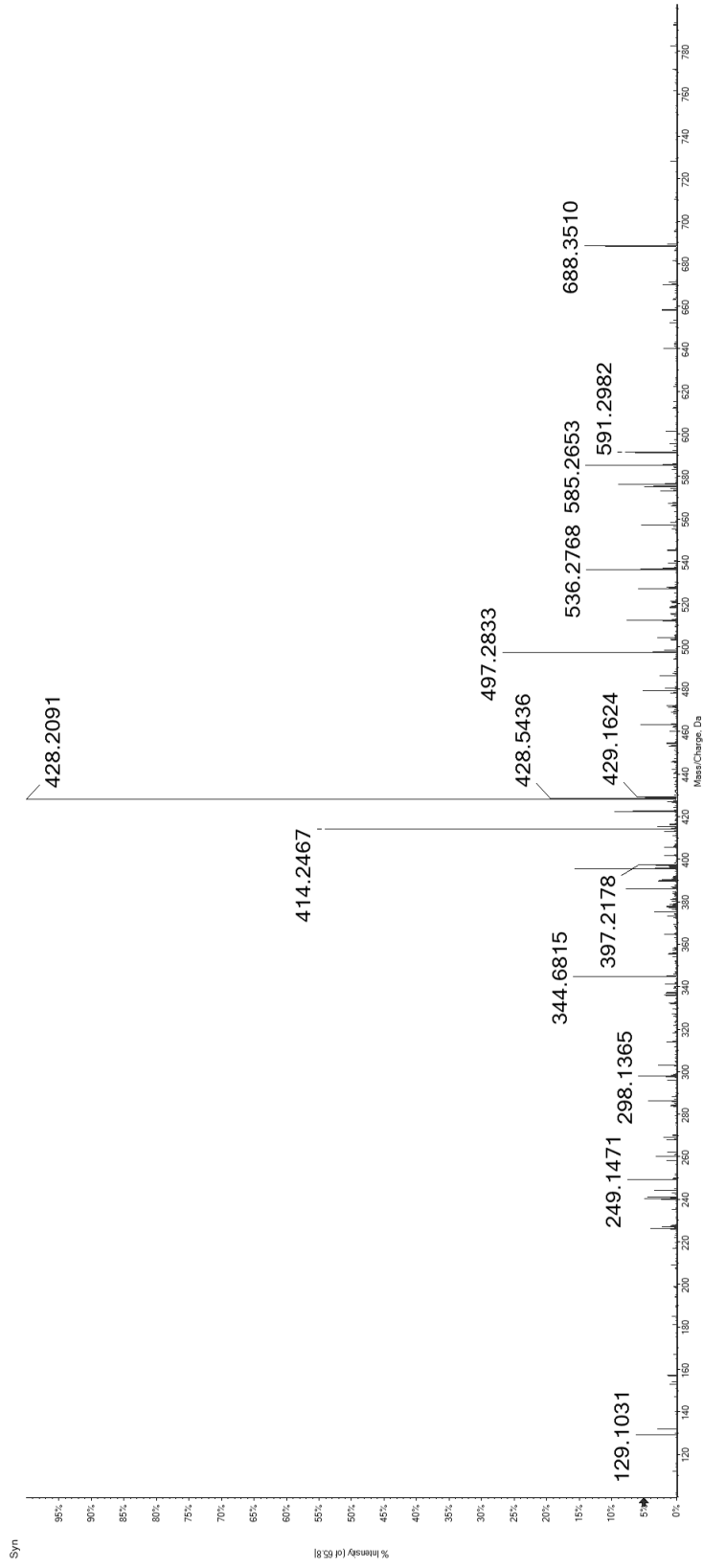
Candidate
Spectrum
(Relative RT – 1,15)



14

REK(pT)PSEDKL

Synthetic Peptide

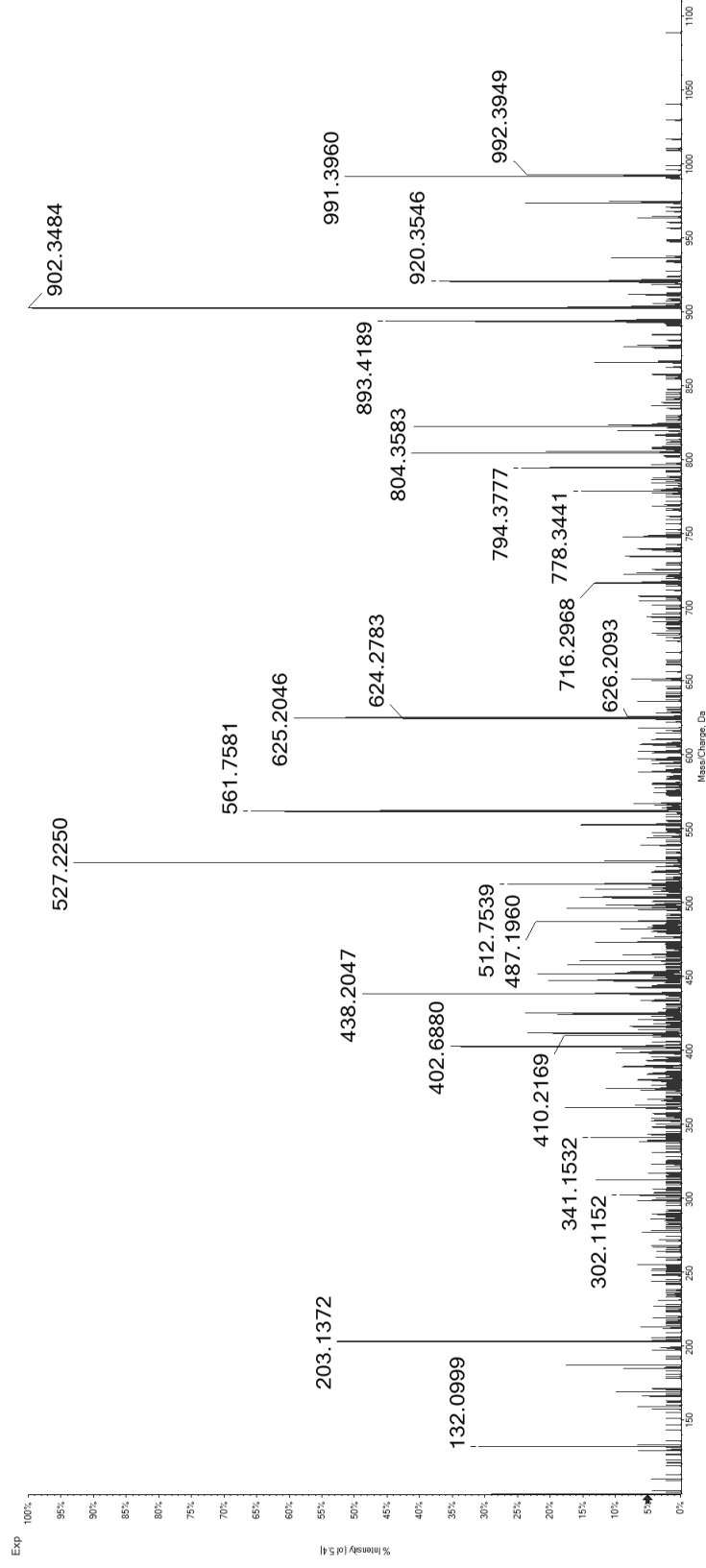
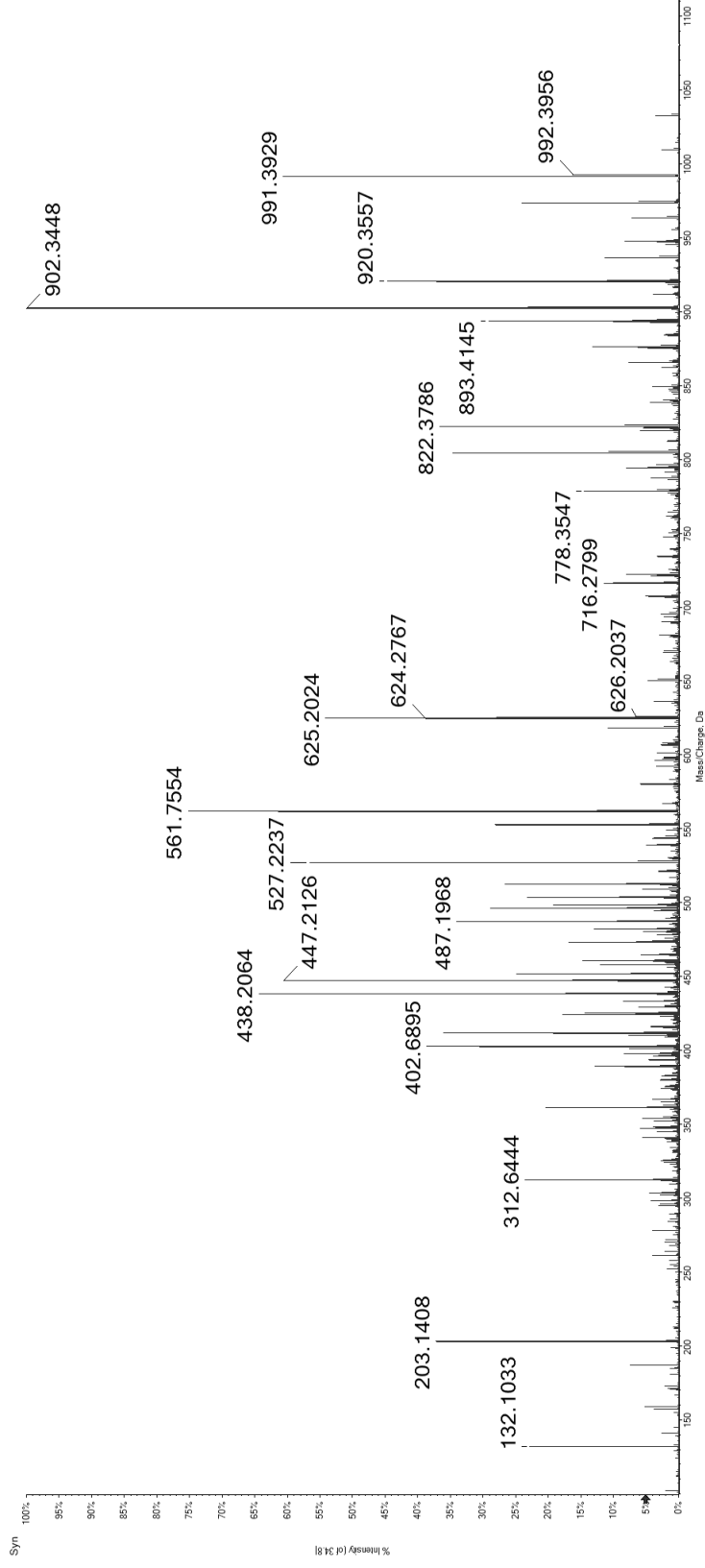


Candidate Spectrum

15

GEDR(ps)PPTAL

Synthetic Peptide

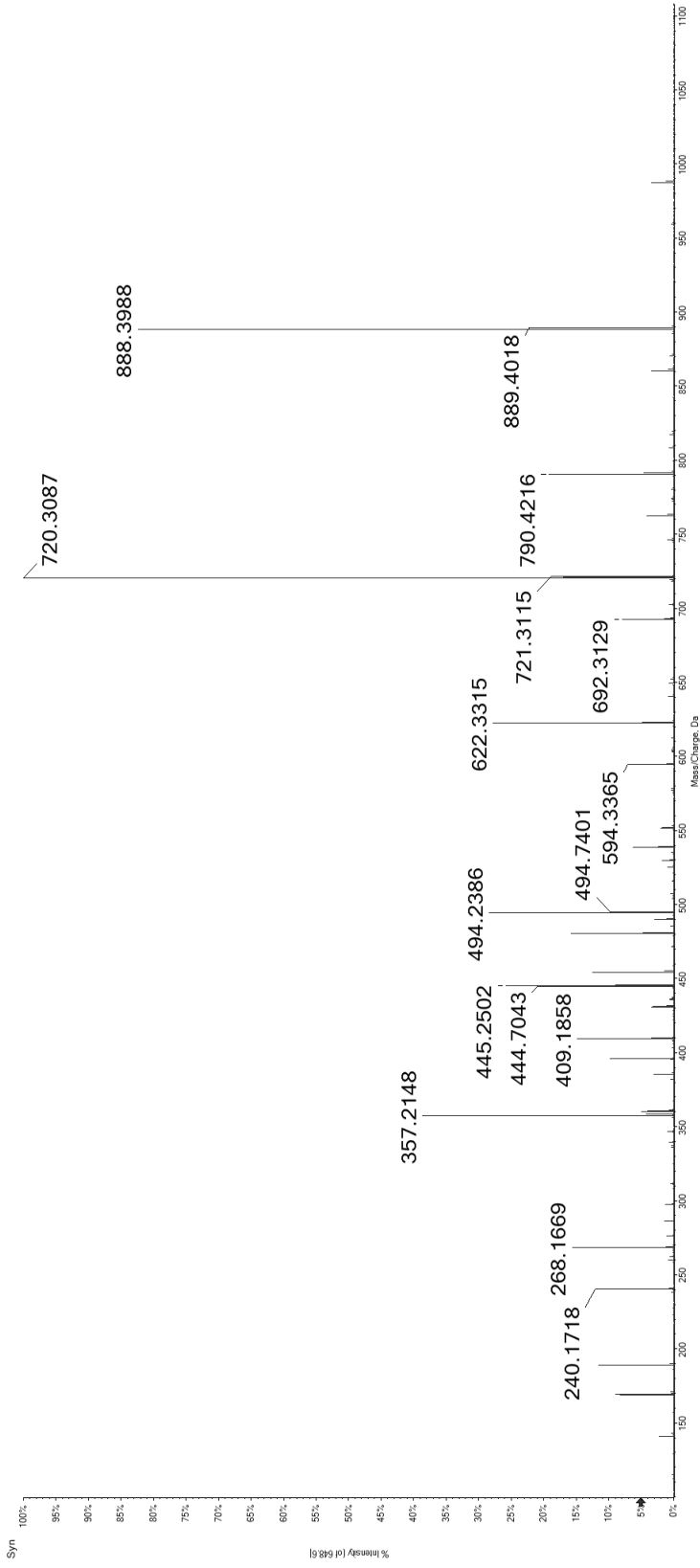


Candidate Spectrum

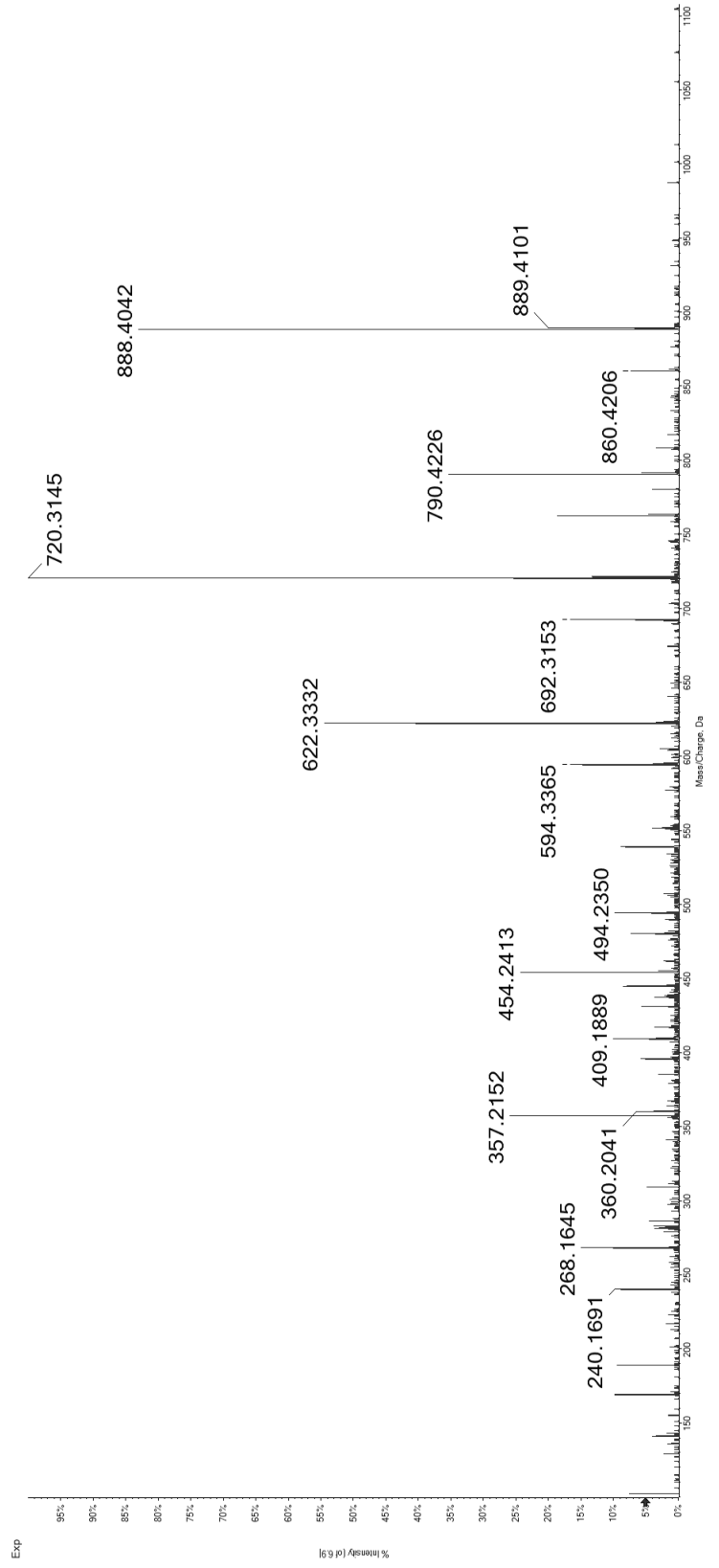
16

REV(pS)PAPAVA

Synthetic Peptide



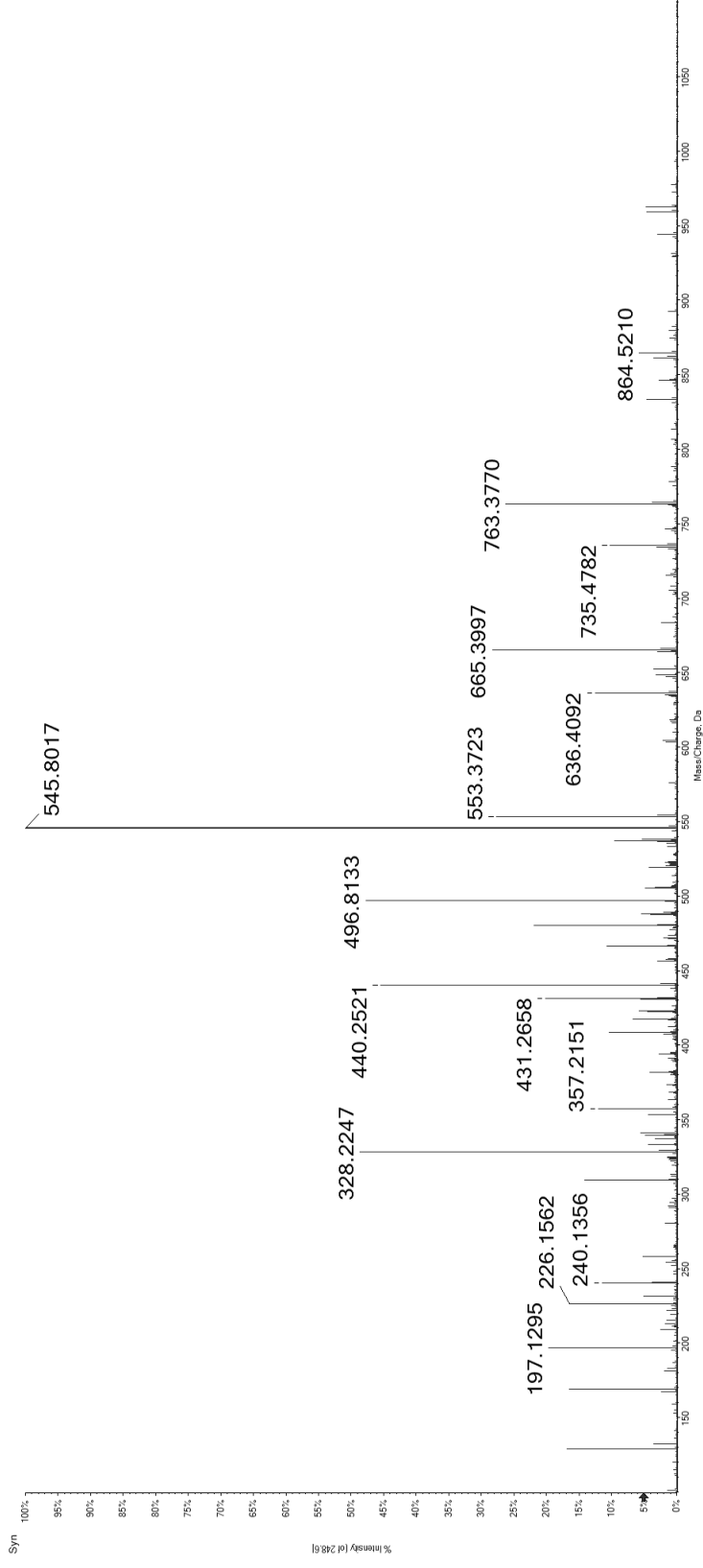
Candidate Spectrum



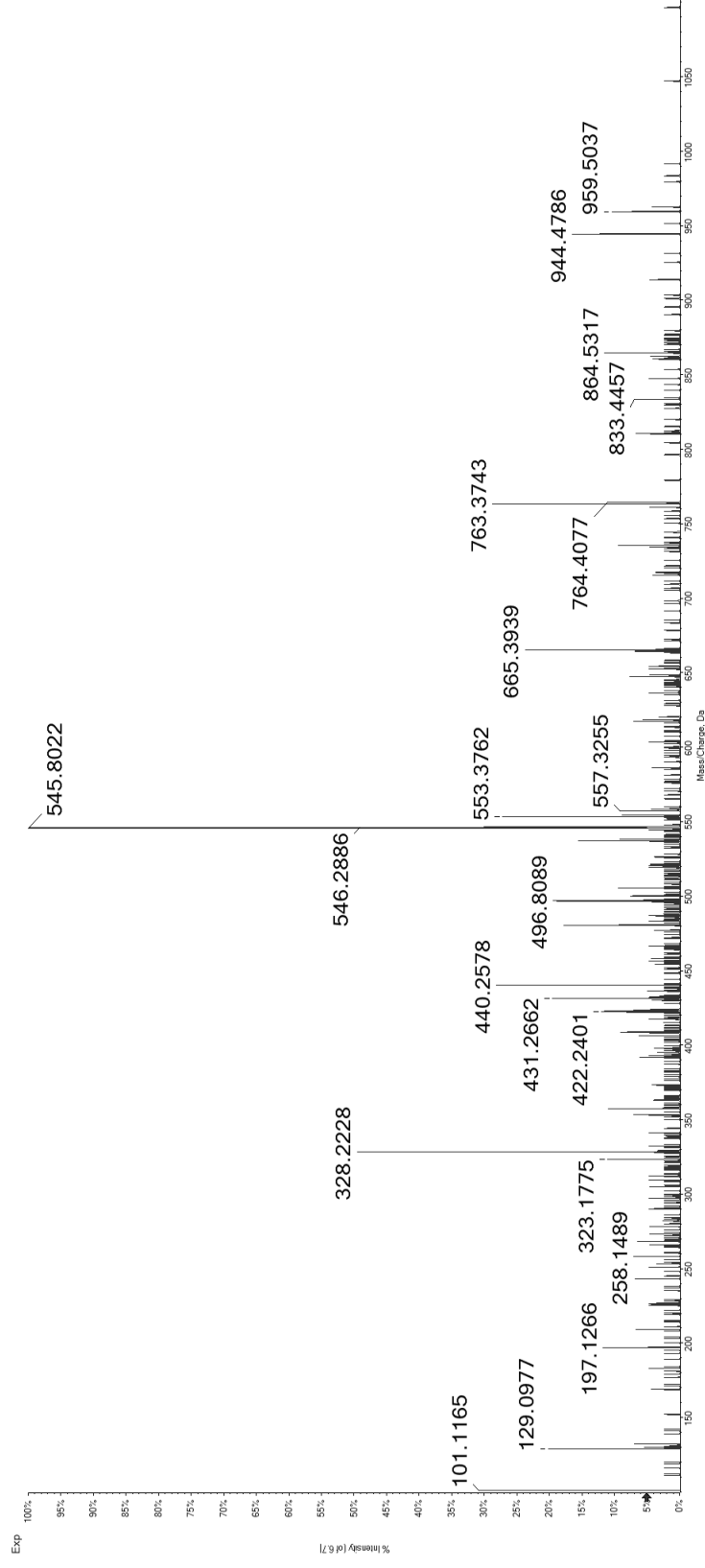
17

KEV(pT)PKPVL

Synthetic Peptide



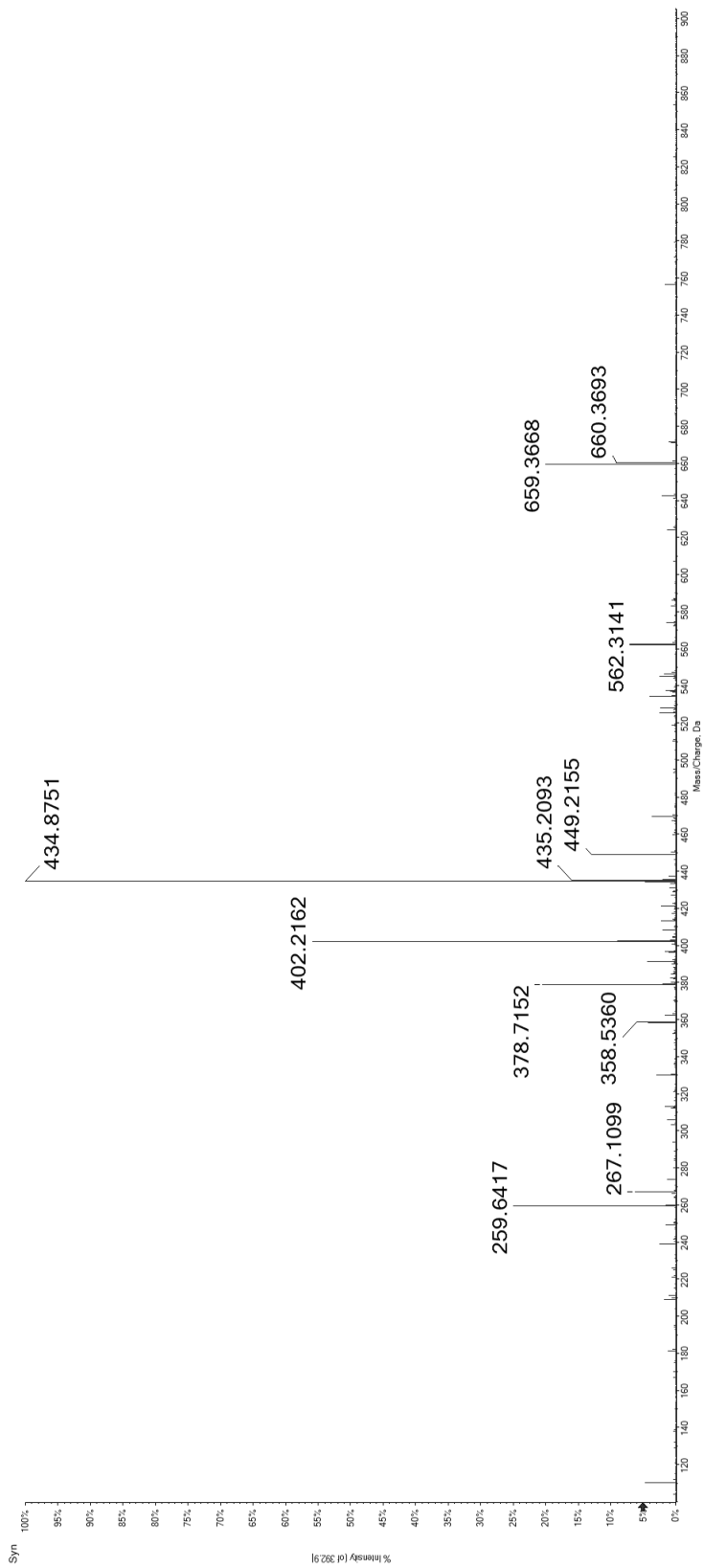
Candidate Spectrum



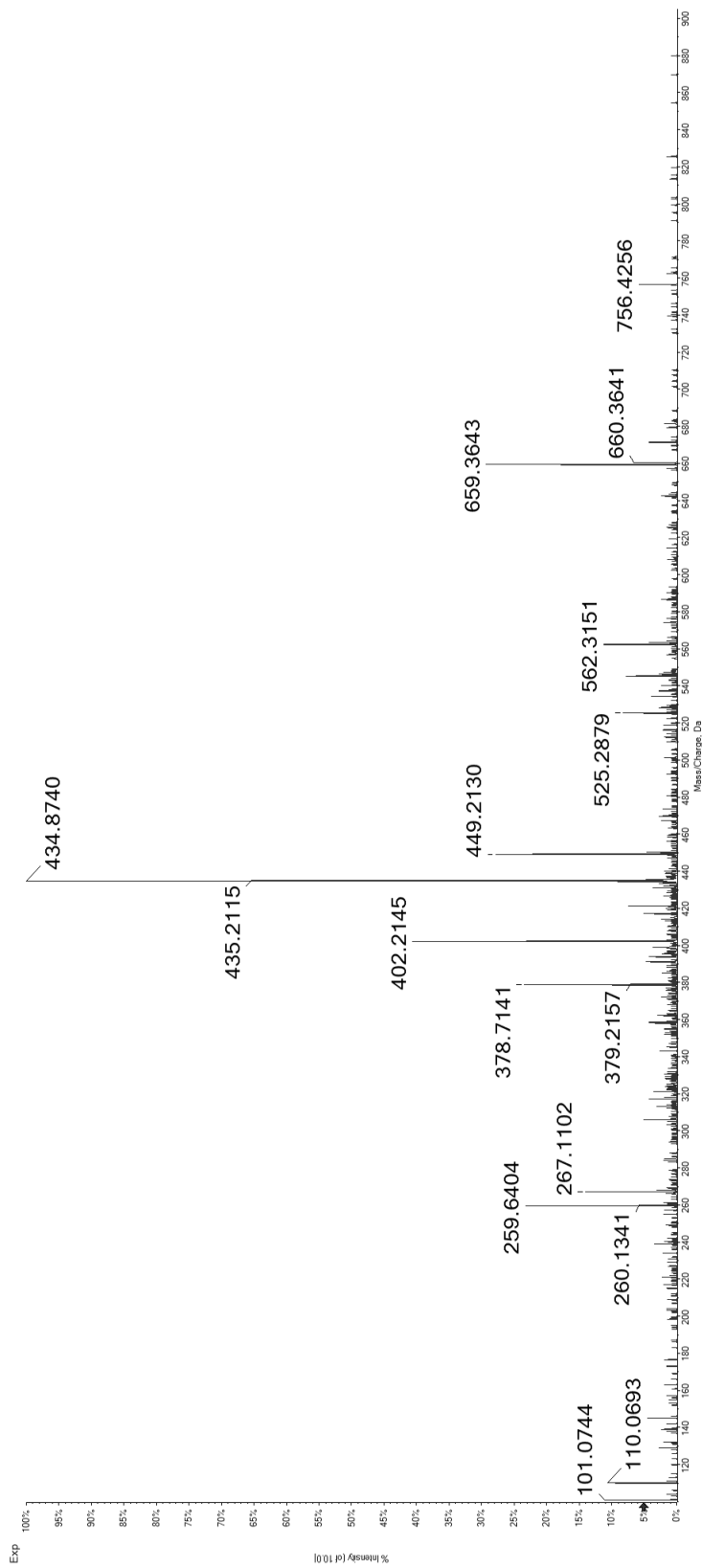
18

HEL(pS)PPQKRM

Synthetic Peptide



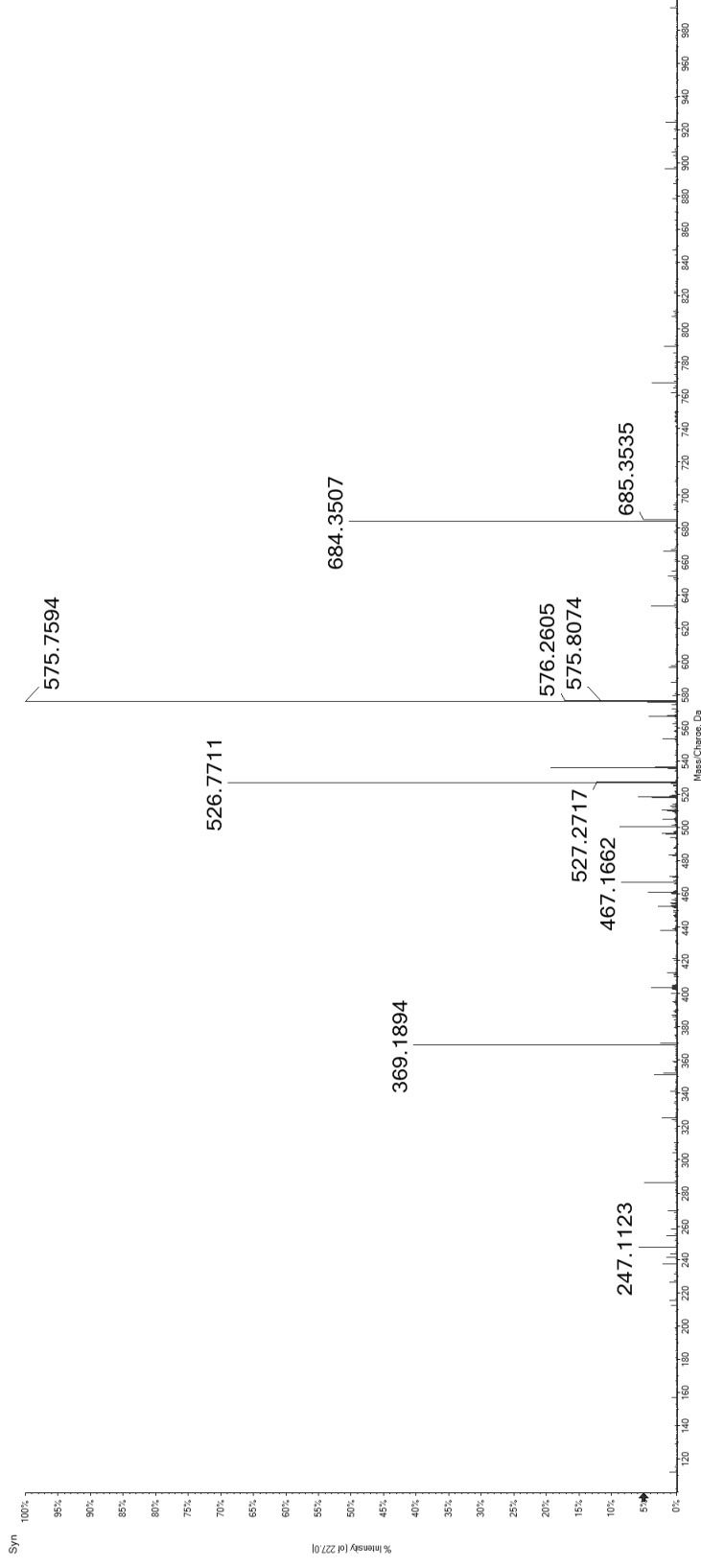
Candidate Spectrum



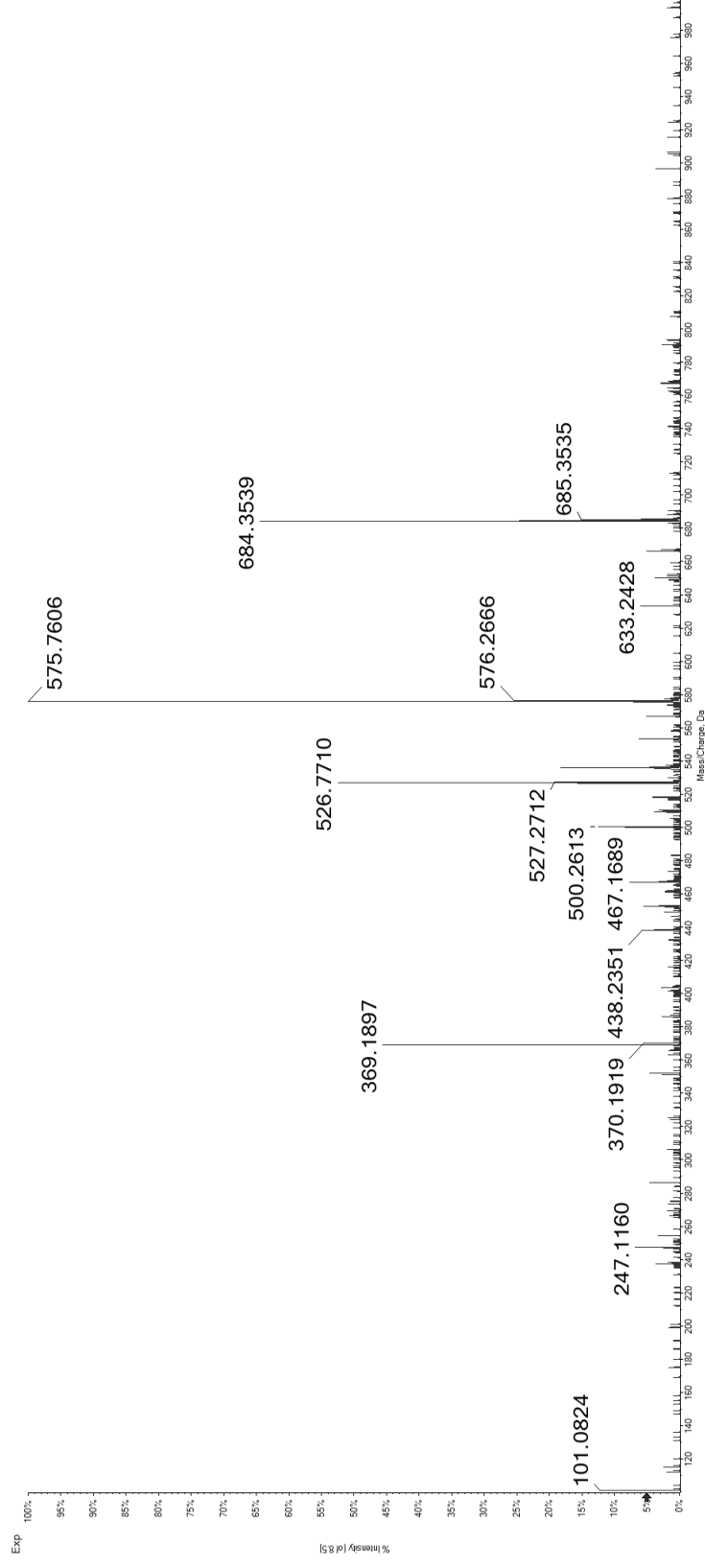
19

RE(pT)PSRPPM

Synthetic Peptide

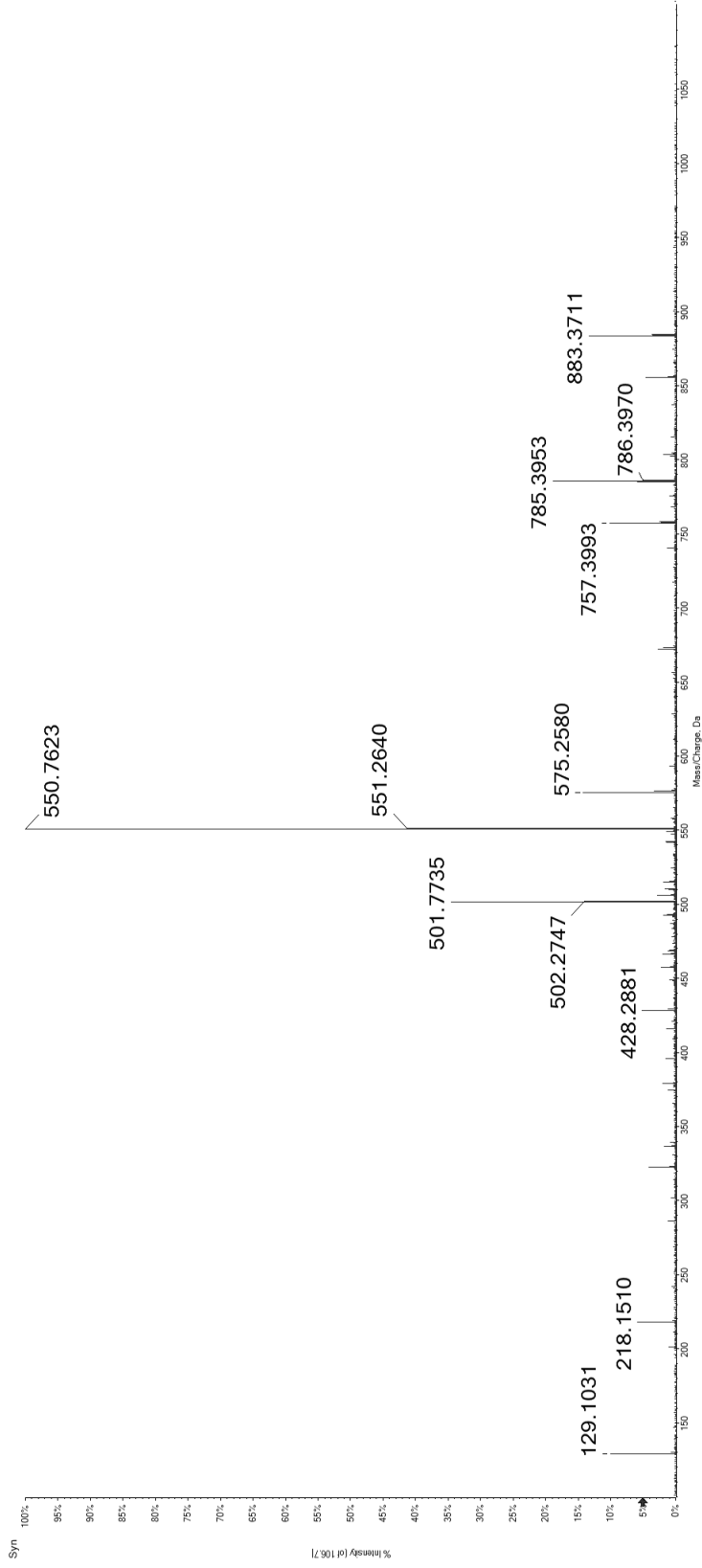


Candidate Spectrum

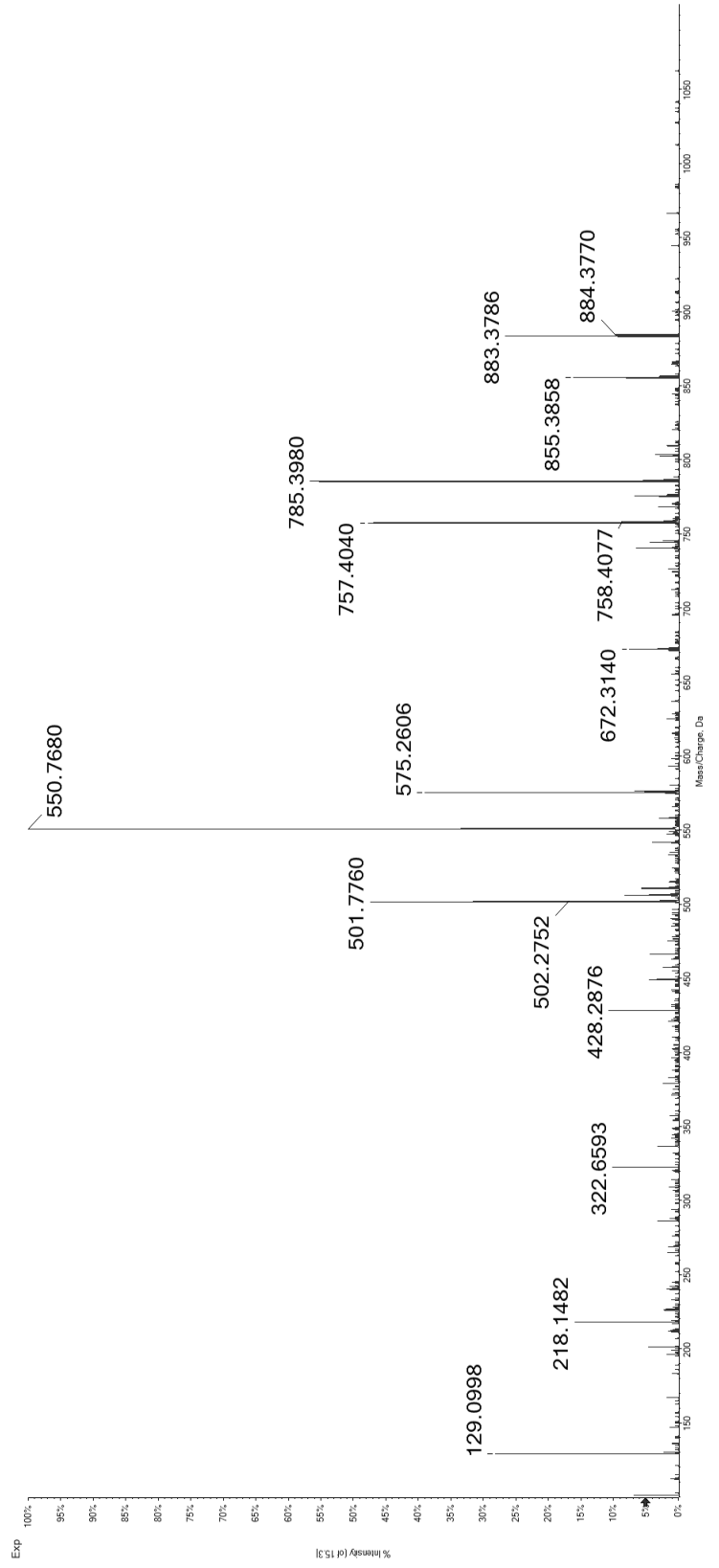


20 REYG(pS)PLKA

Synthetic
Peptide



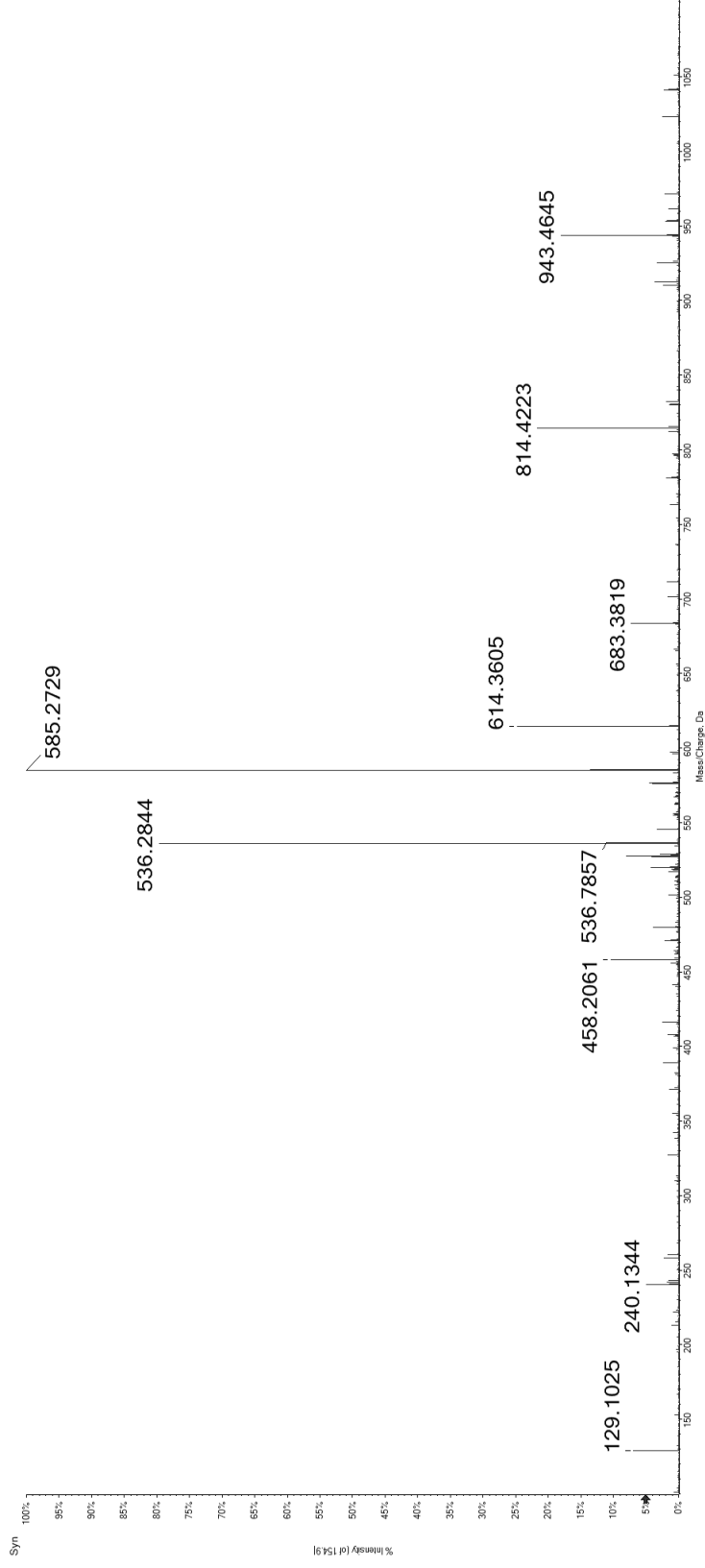
Candidate
Spectrum



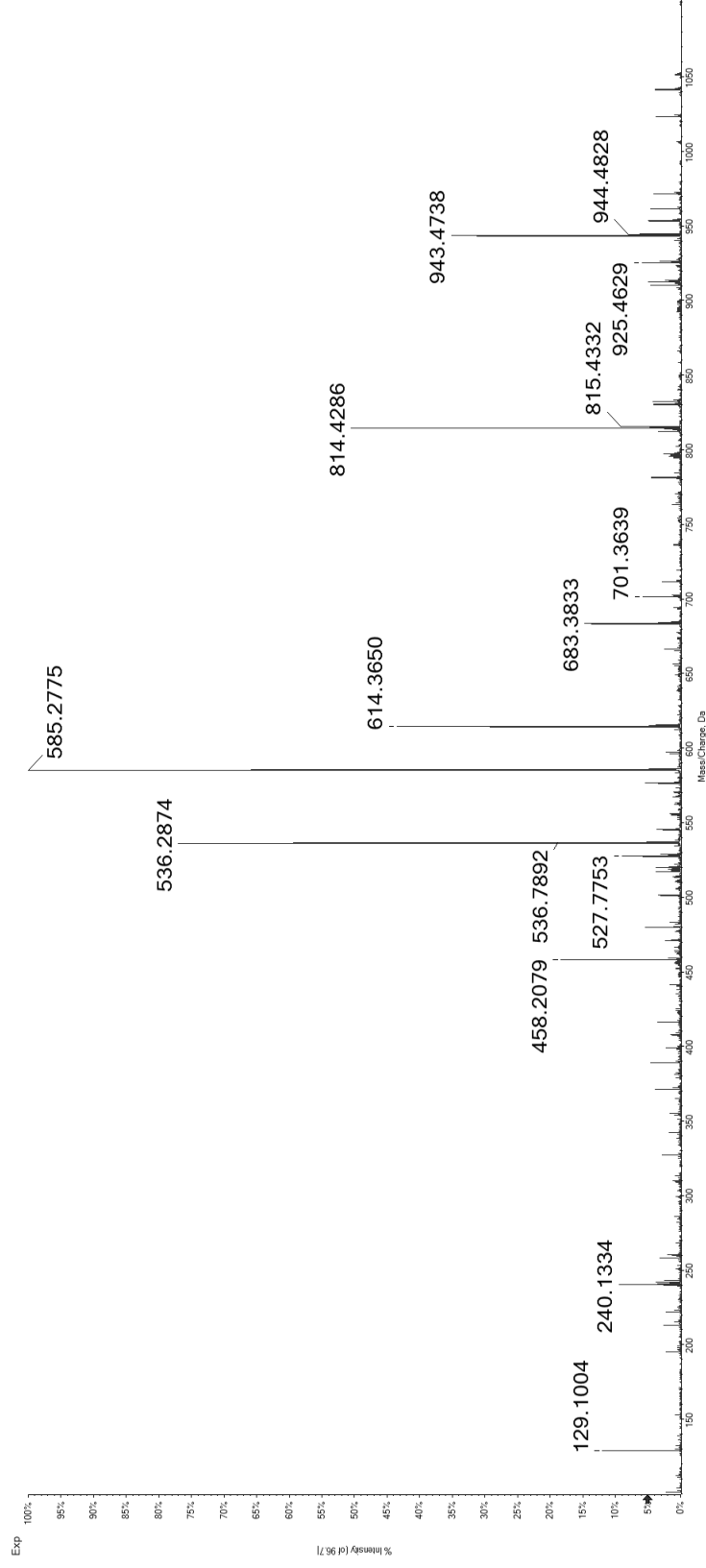
21

KEM(pS)PTRQL

Synthetic Peptide



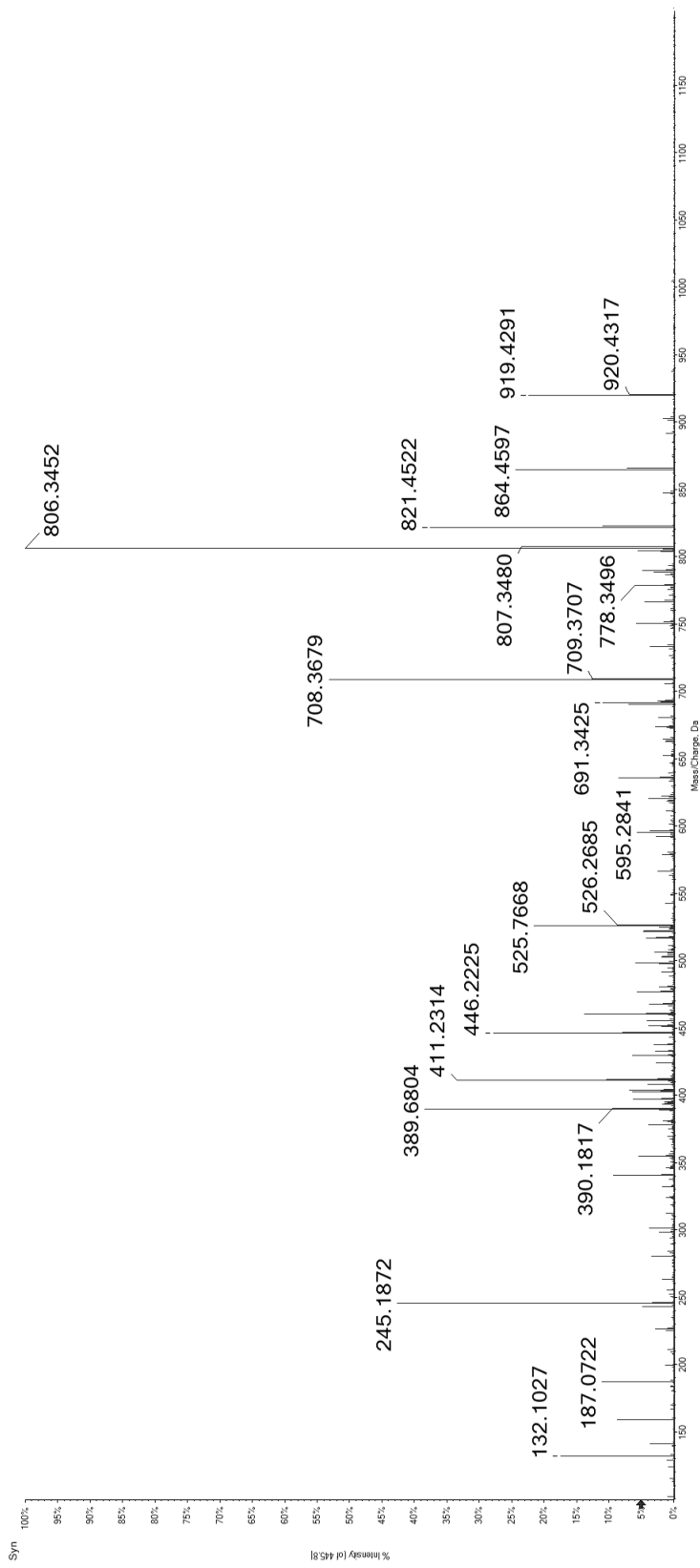
Candidate Spectrum



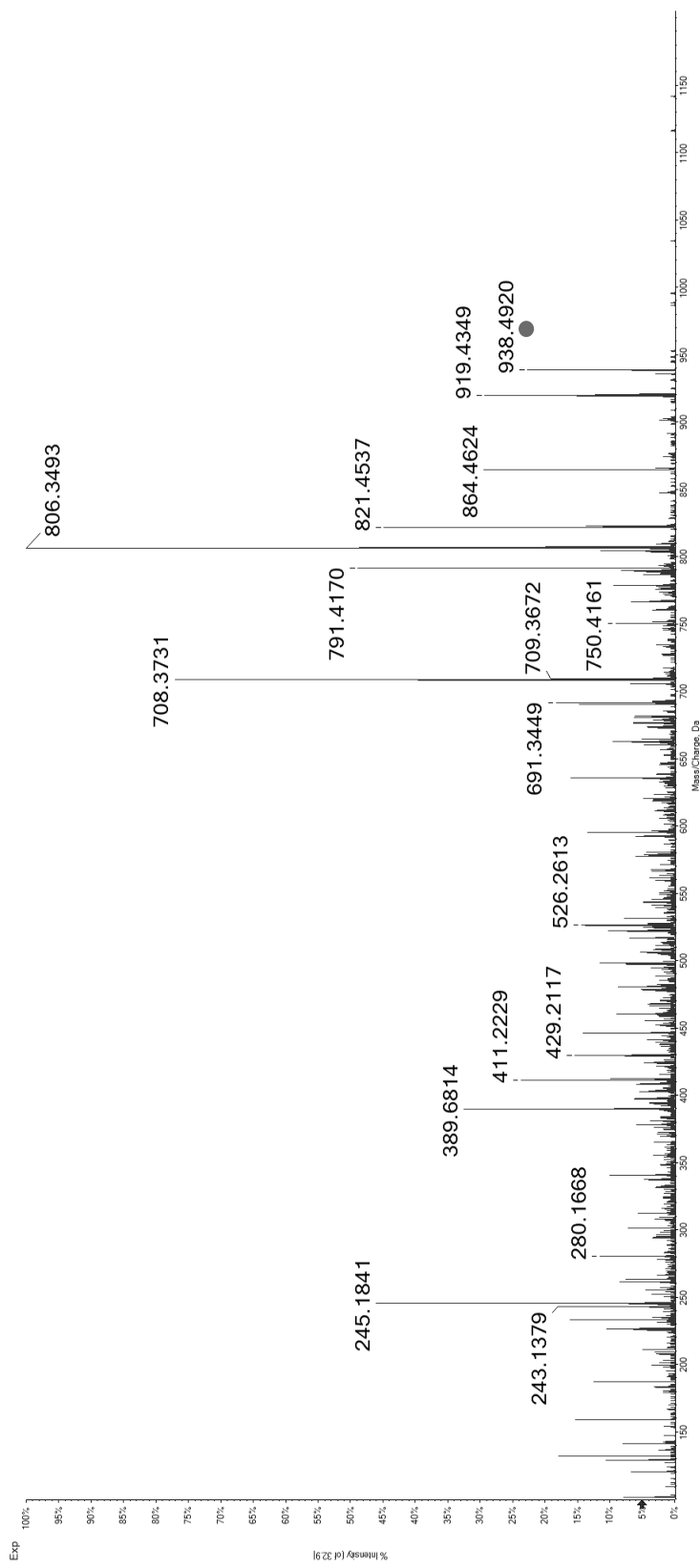
22

GENK(ps)PLLL

Synthetic
Peptide
(Relative RT – 2,01)

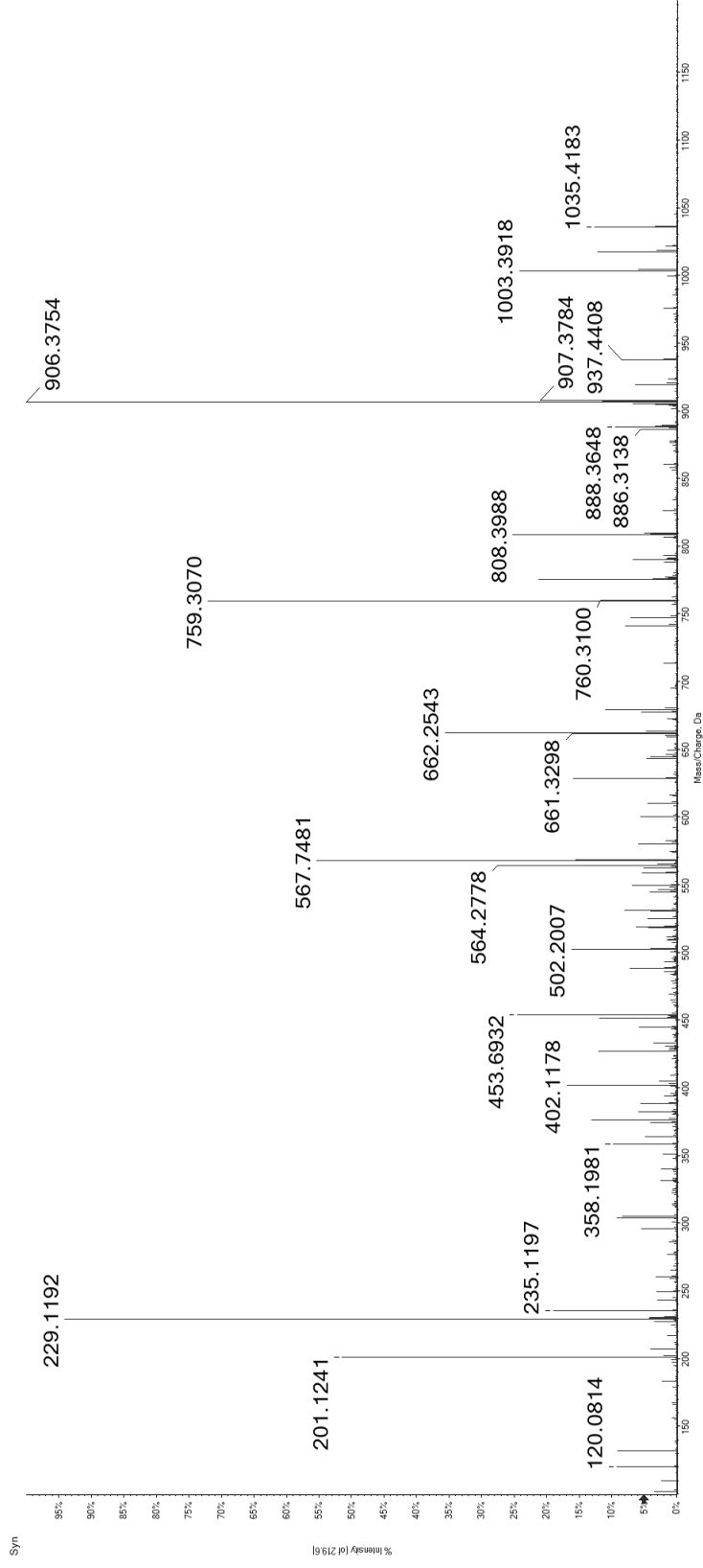


Candidate
Spectrum
(Relative RT – 1,97)

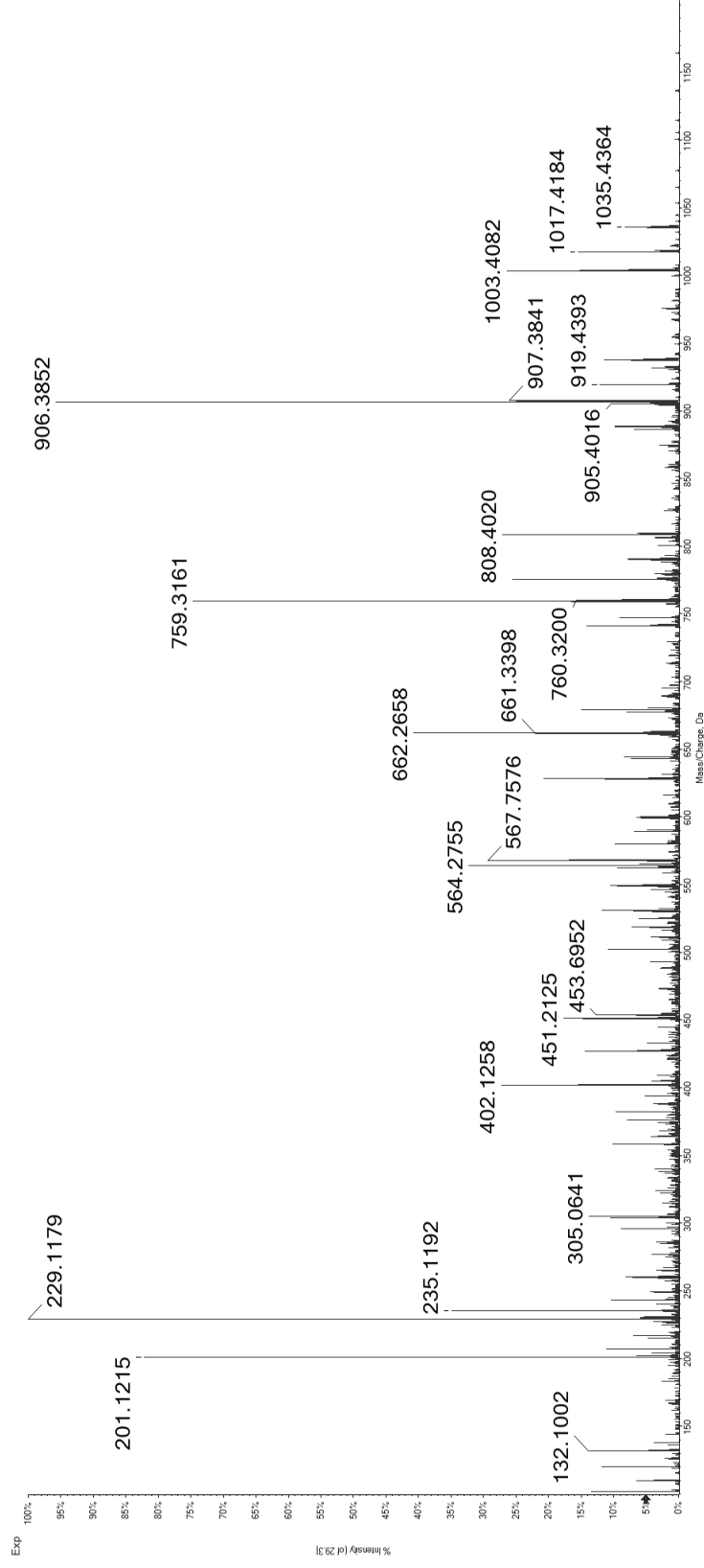


23 VEFPH(pS)PEI

Synthetic Peptide

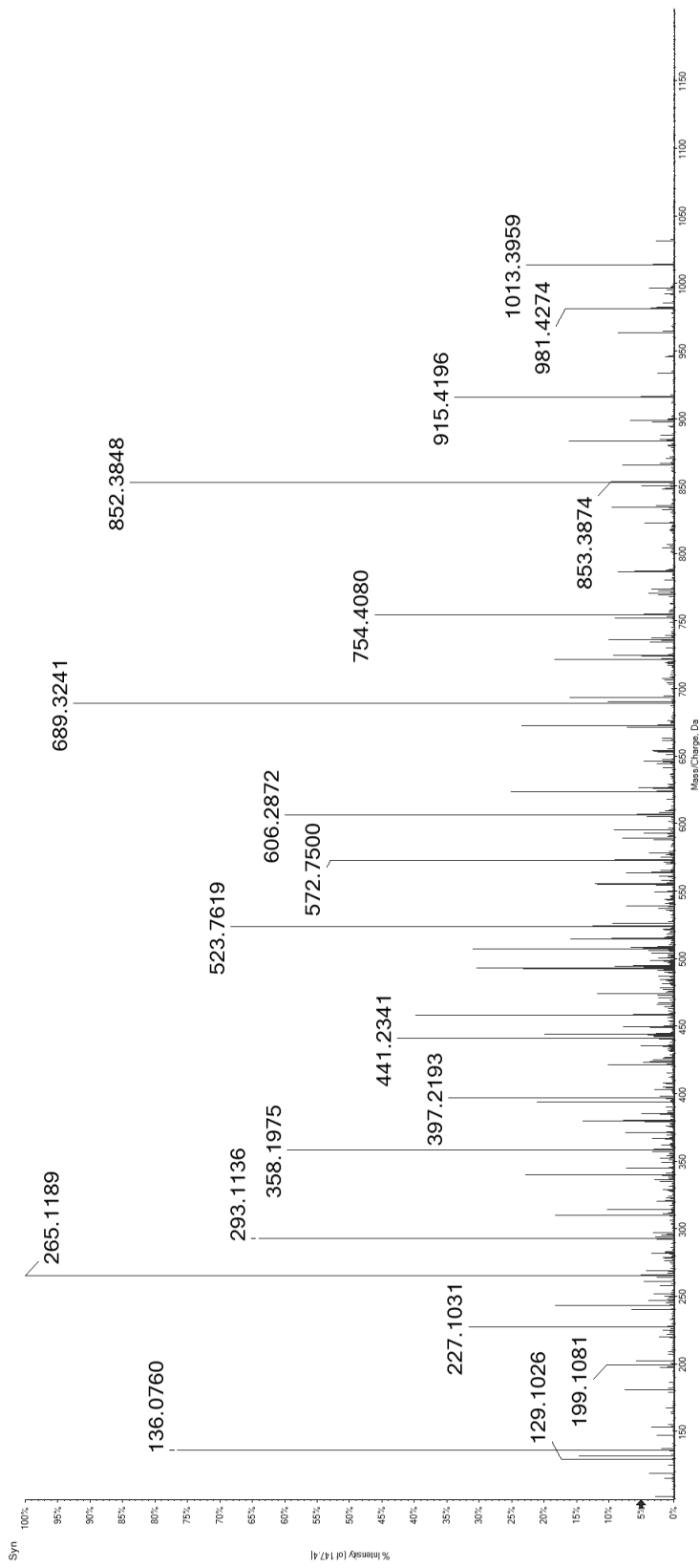


Candidate Spectrum

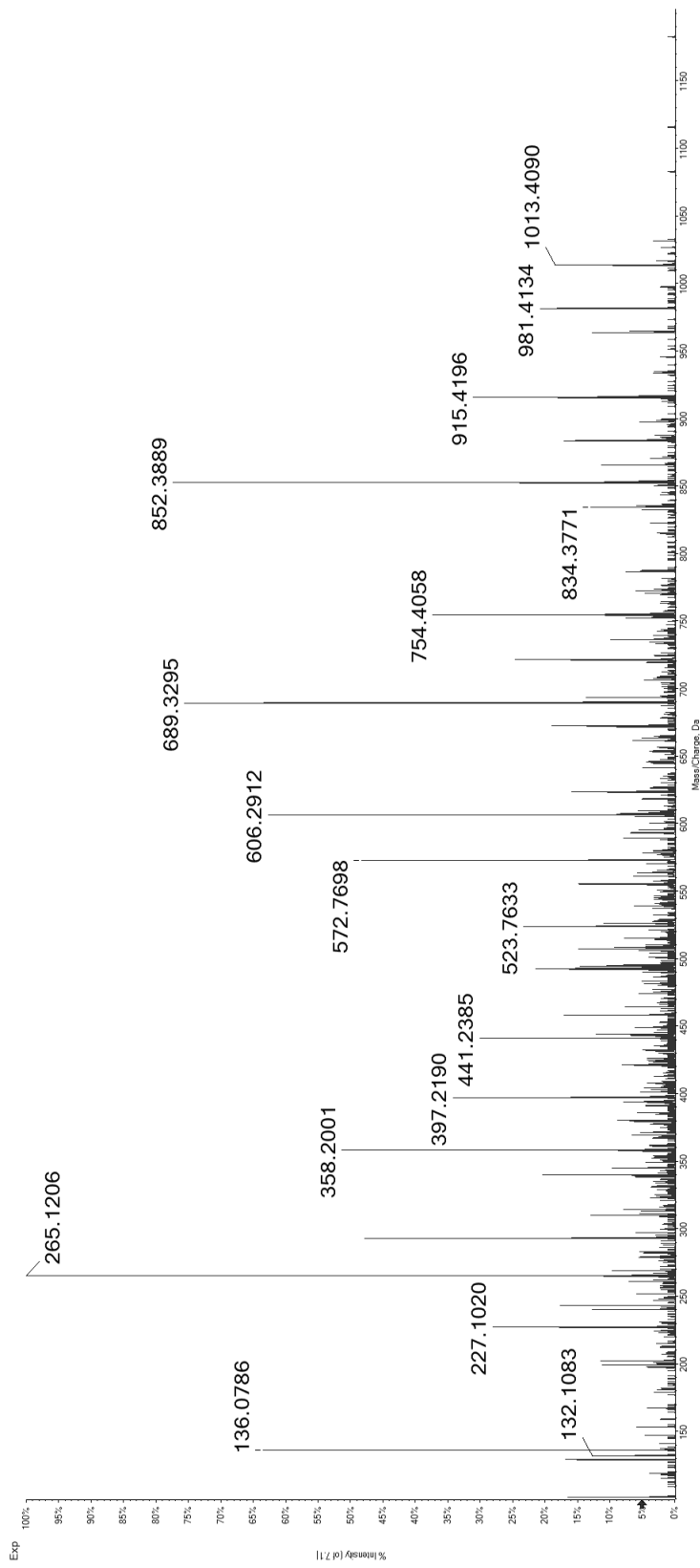


24 YEKAN(pT)PEL

Synthetic Peptide



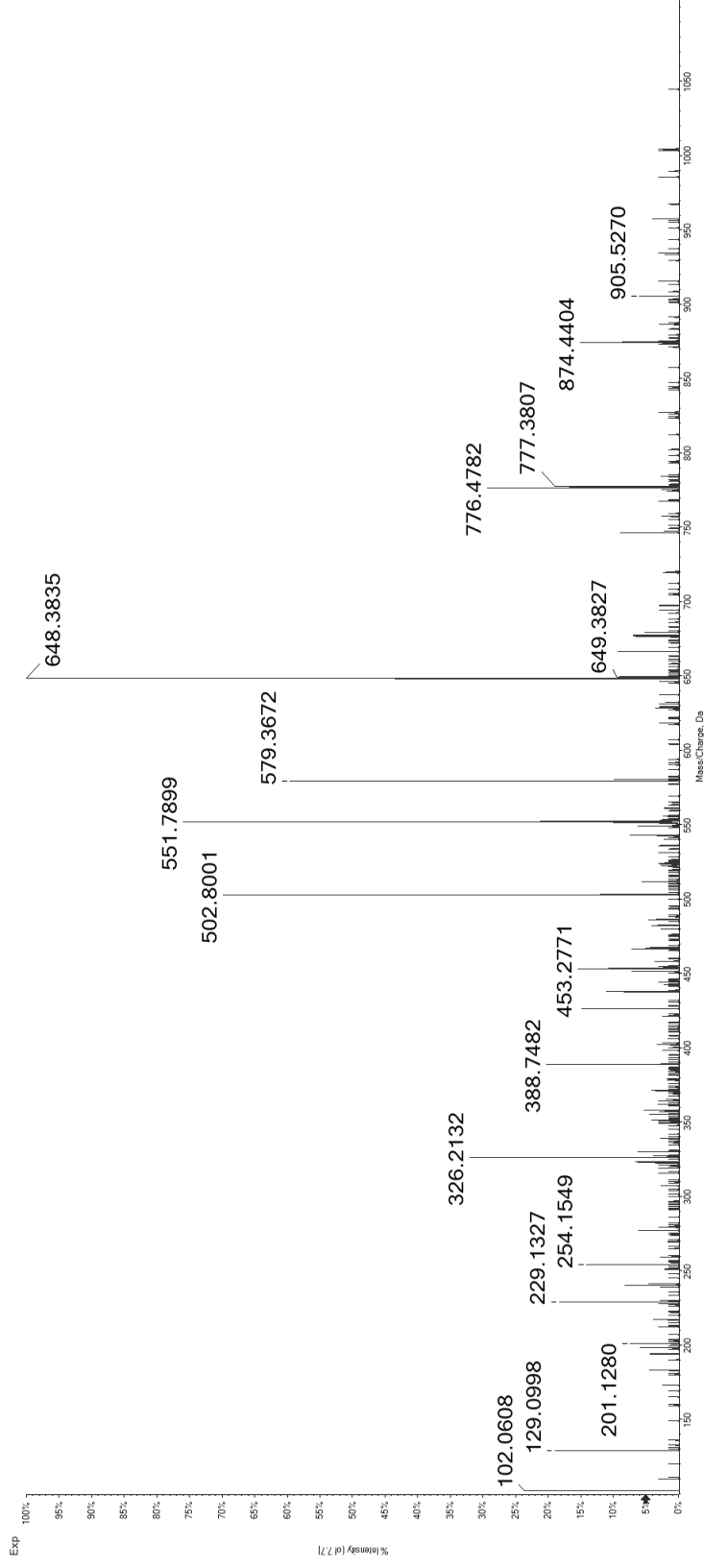
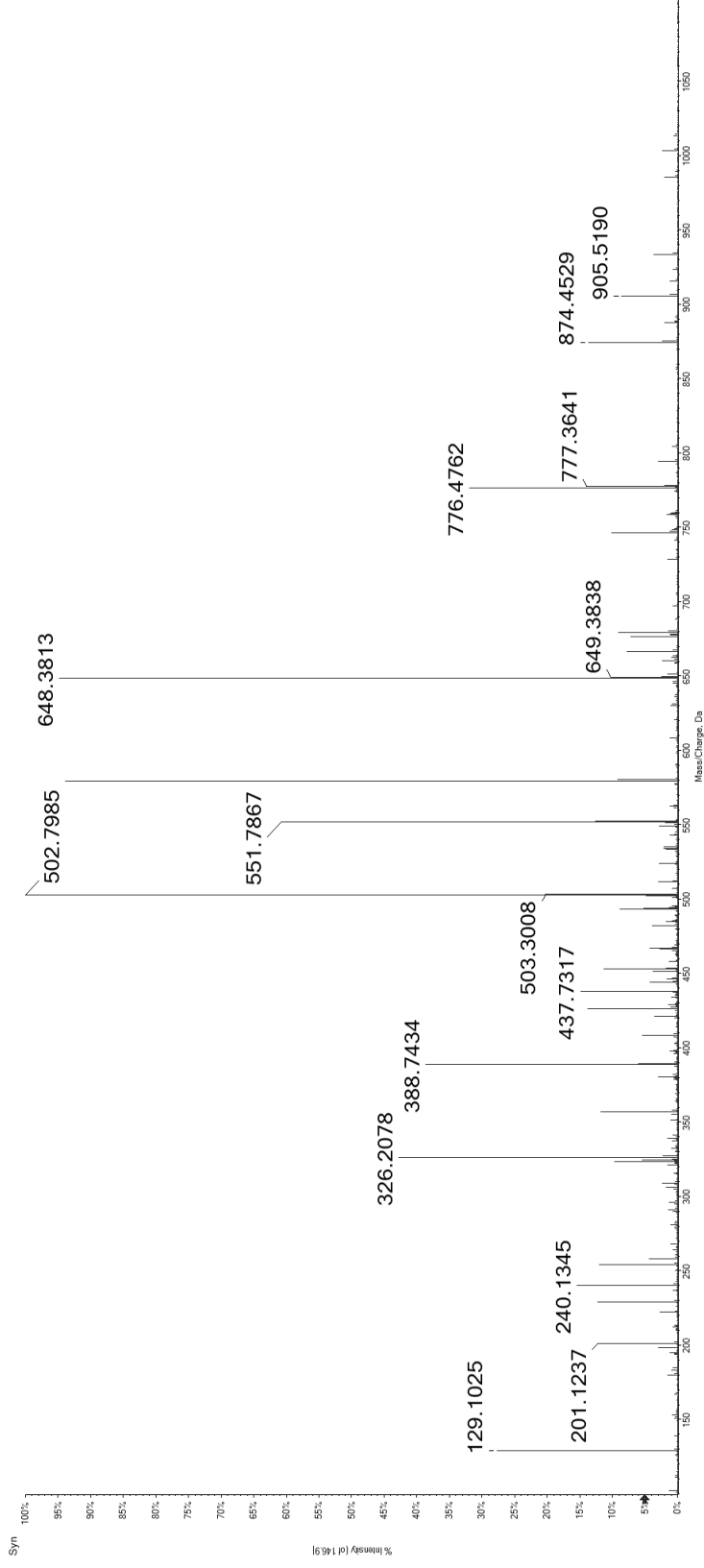
Candidate Spectrum



25

VEK(pS)PRPPL

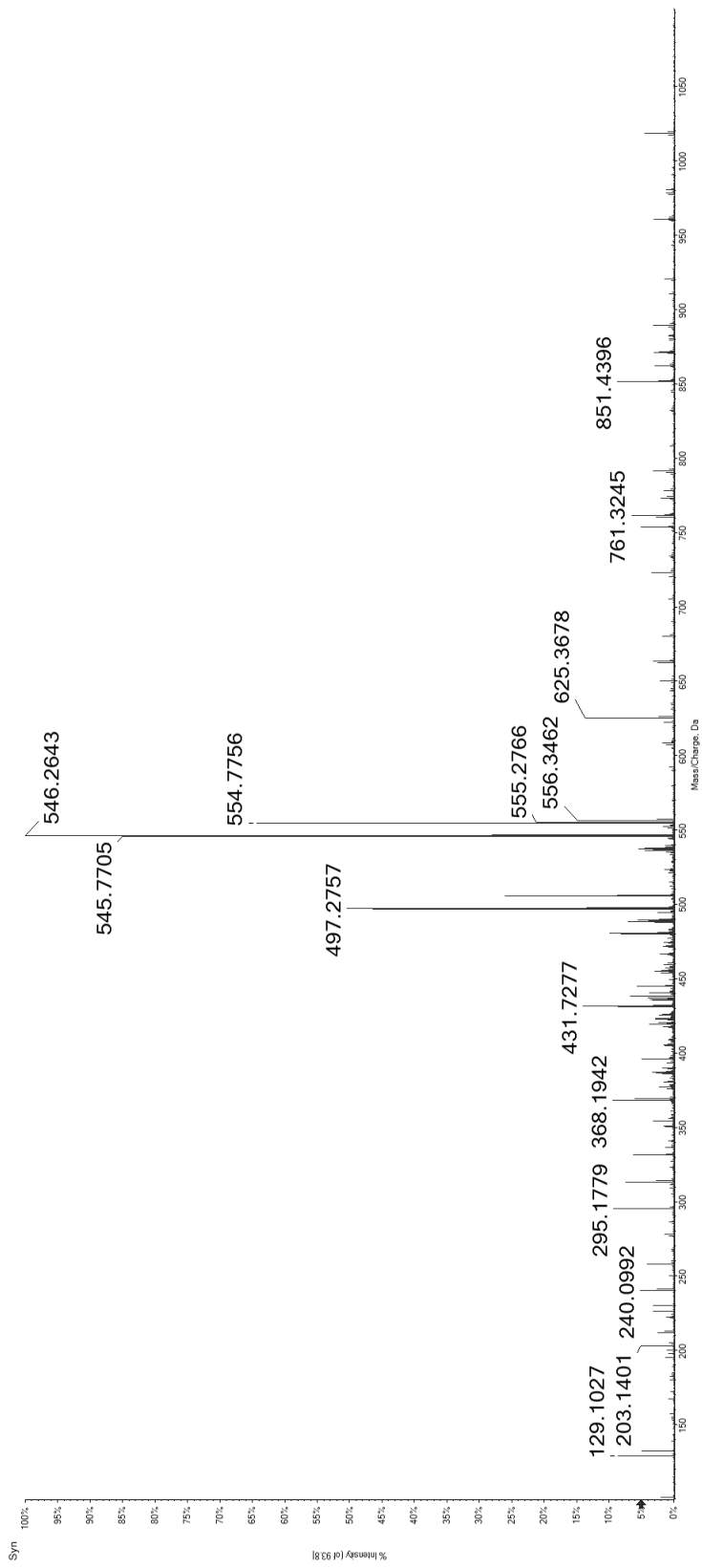
Synthetic Peptide



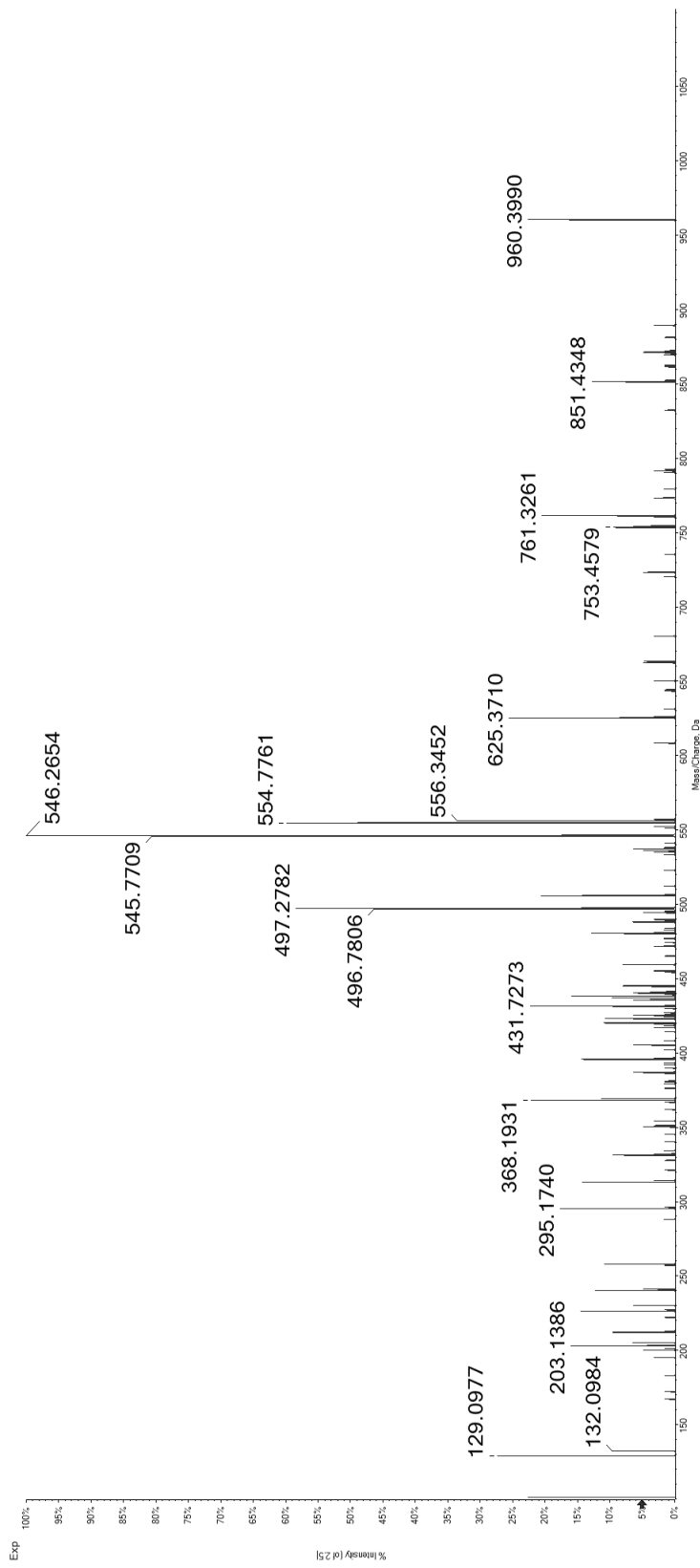
Candidate Spectrum

26 QEK(pS)PKQAL

Synthetic
Peptide



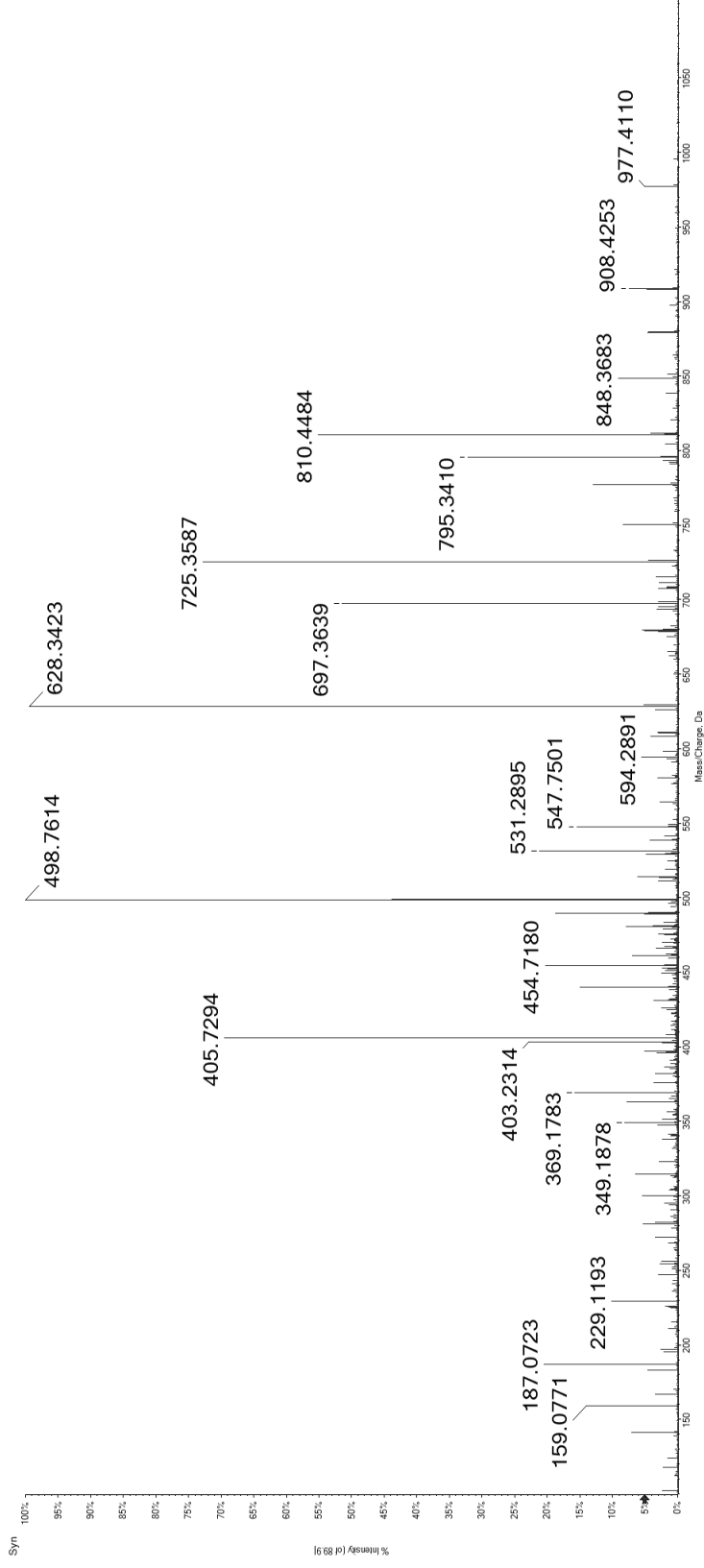
Candidate
Spectrum



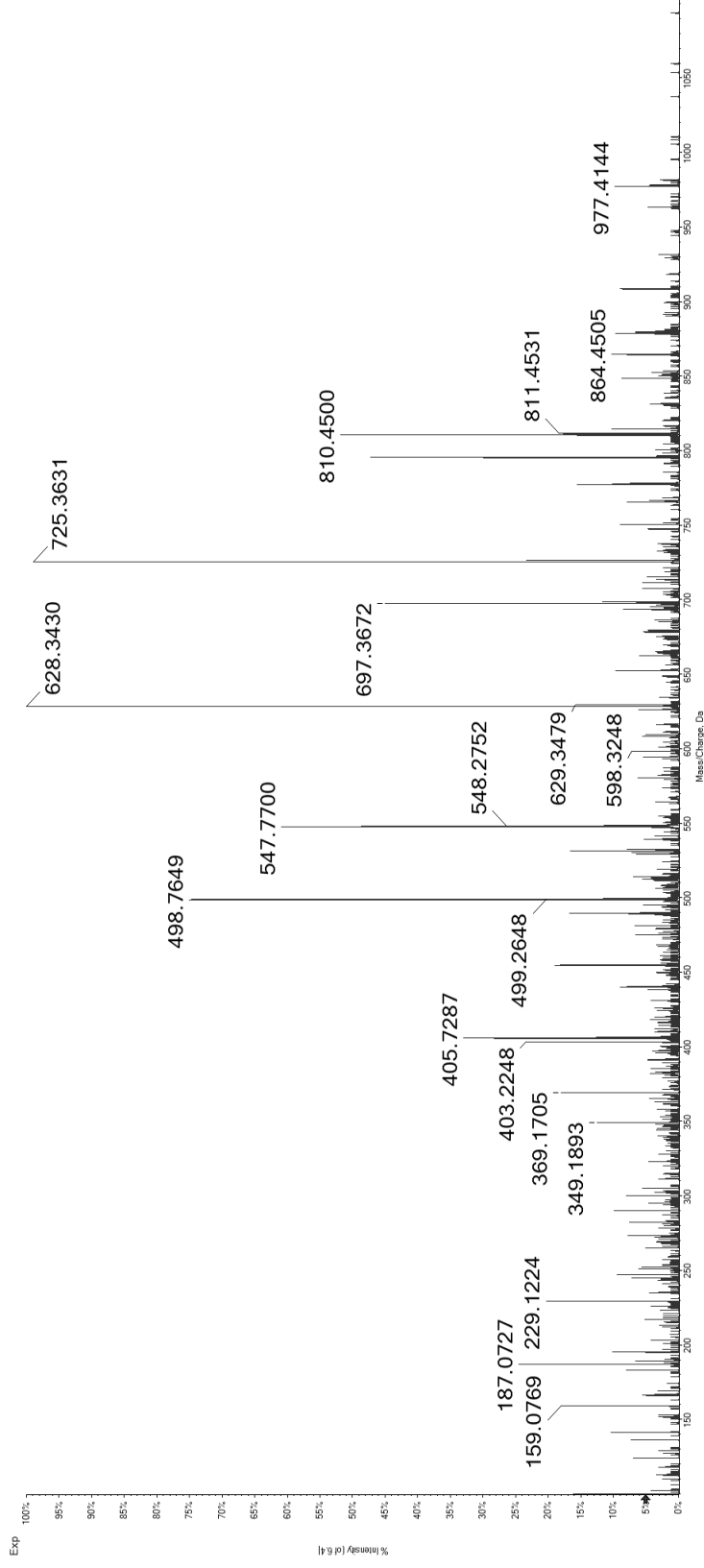
27

GEI(ps)PQREV

Synthetic Peptide



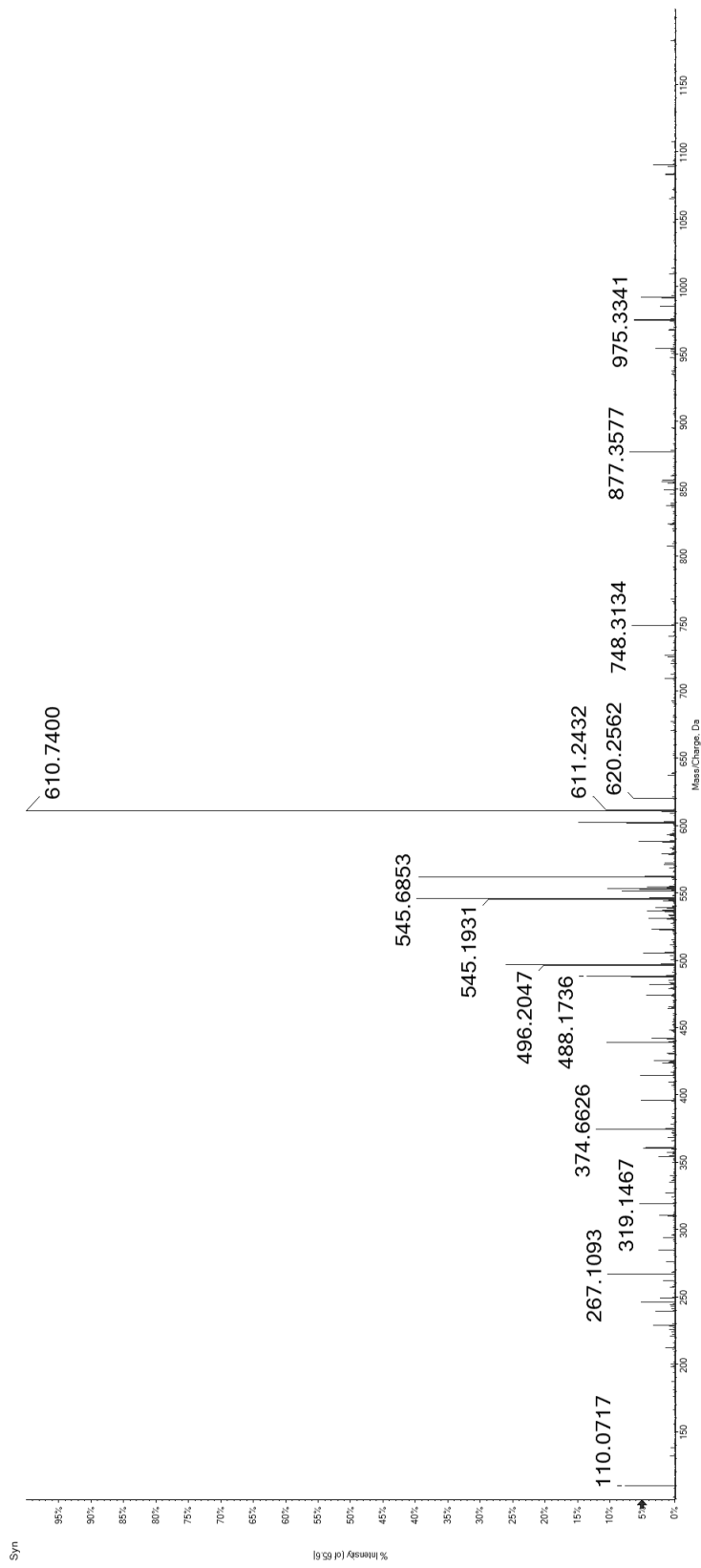
Candidate Spectrum



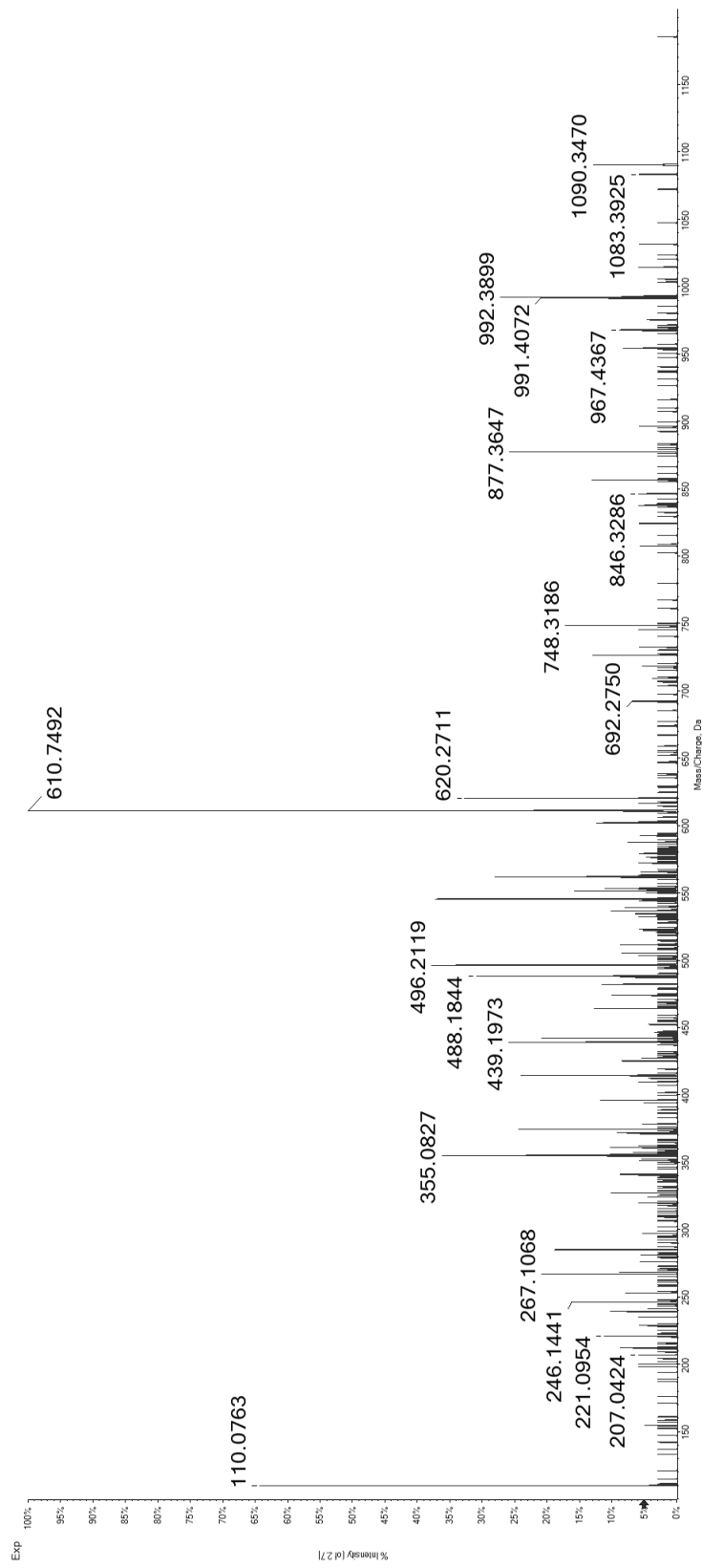
28

HEFH(pS)QENL

Synthetic Peptide

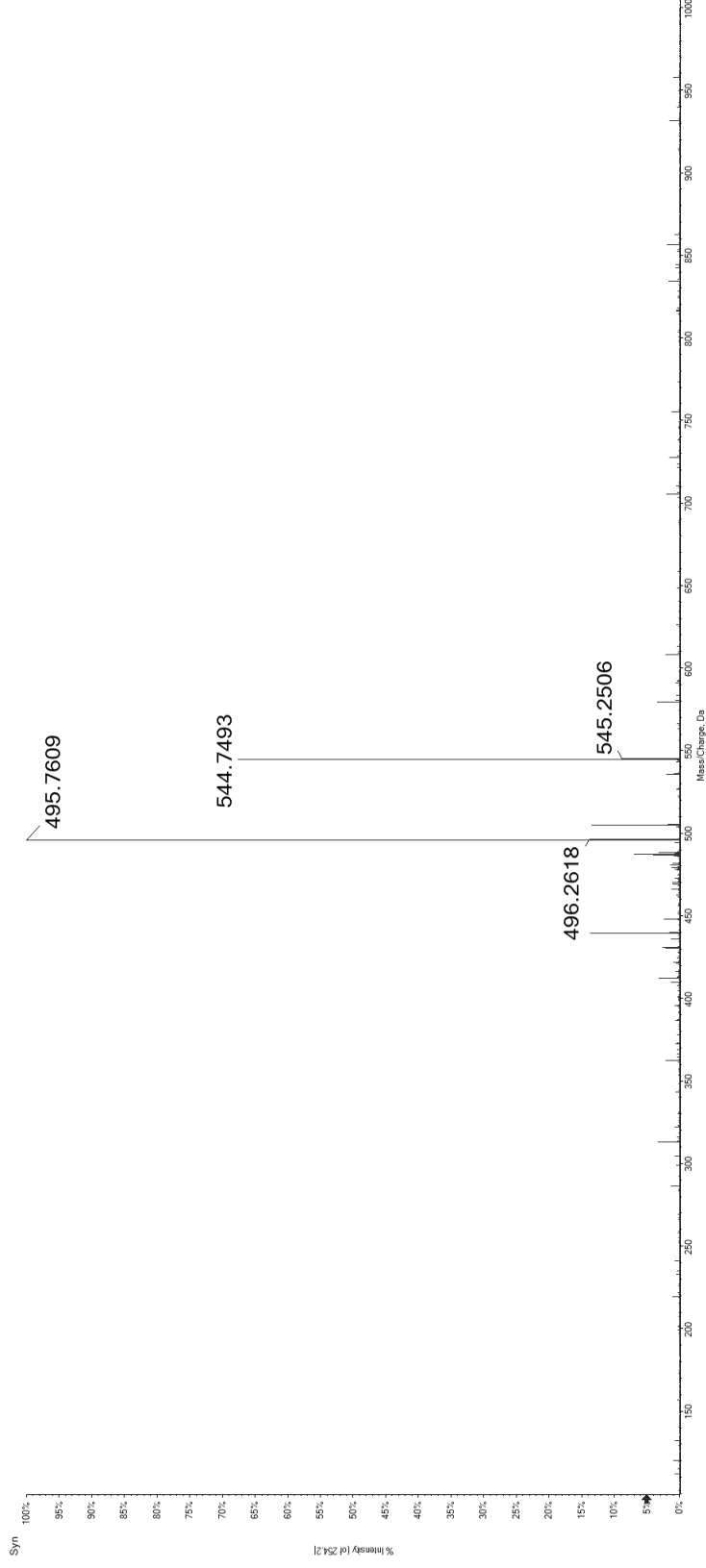


Candidate Spectrum

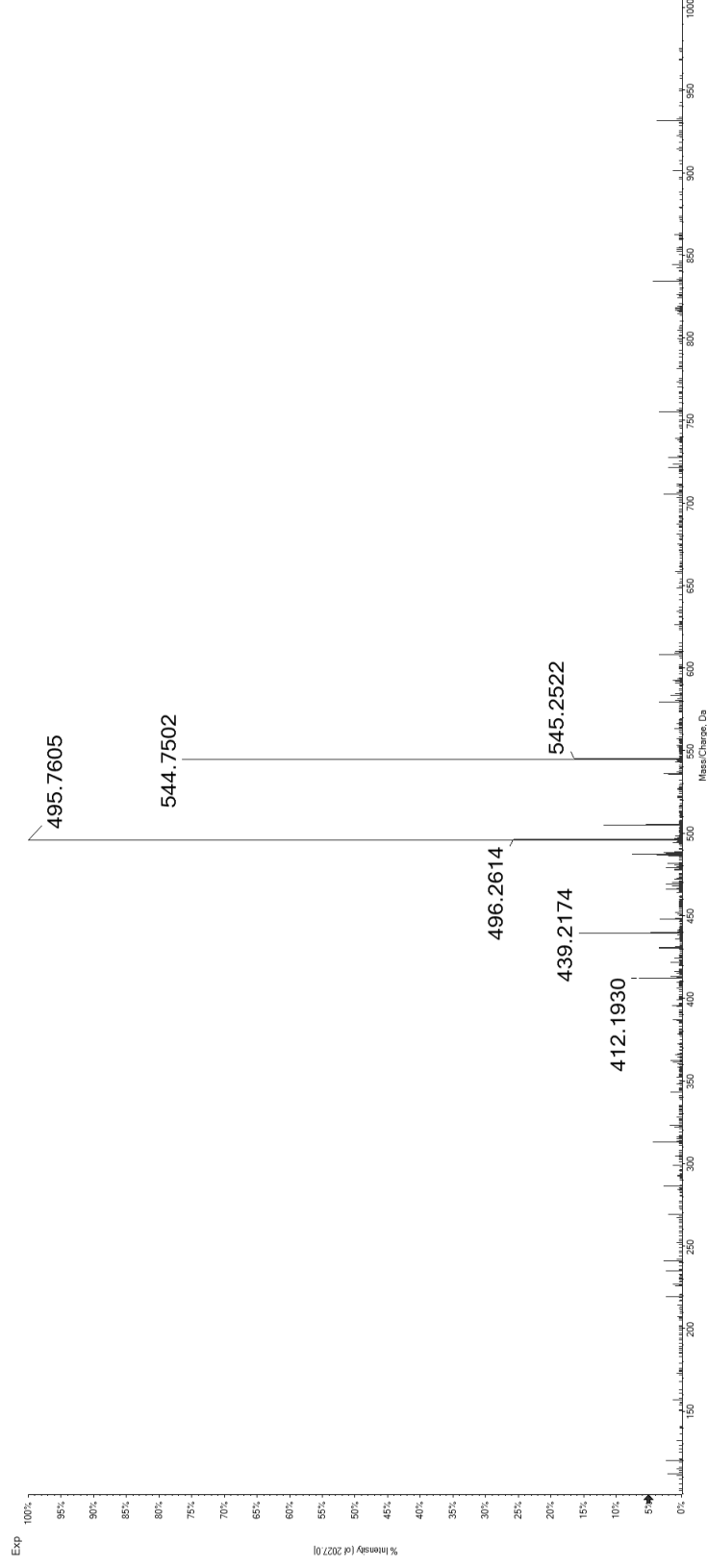


29 REG(ps)GRFSL

Synthetic
Peptide



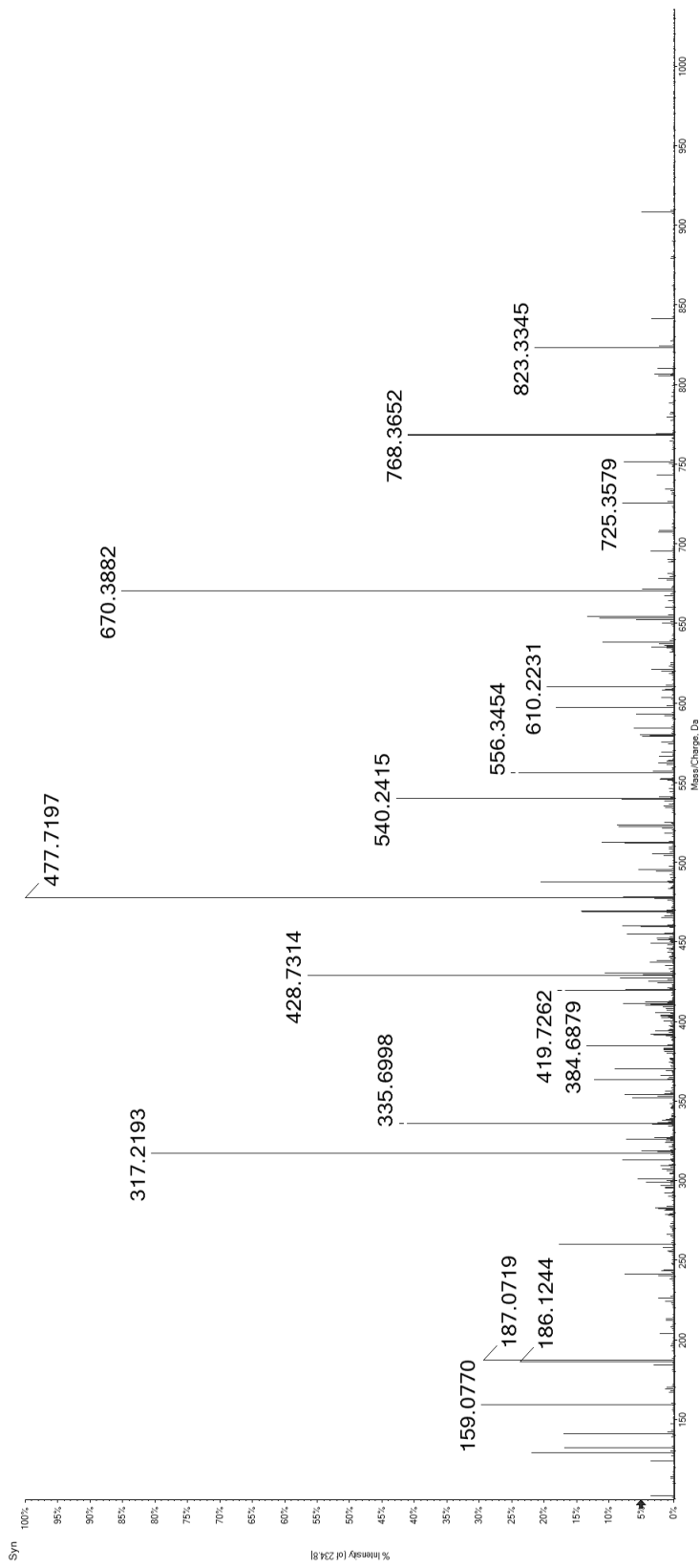
Candidate
Spectrum



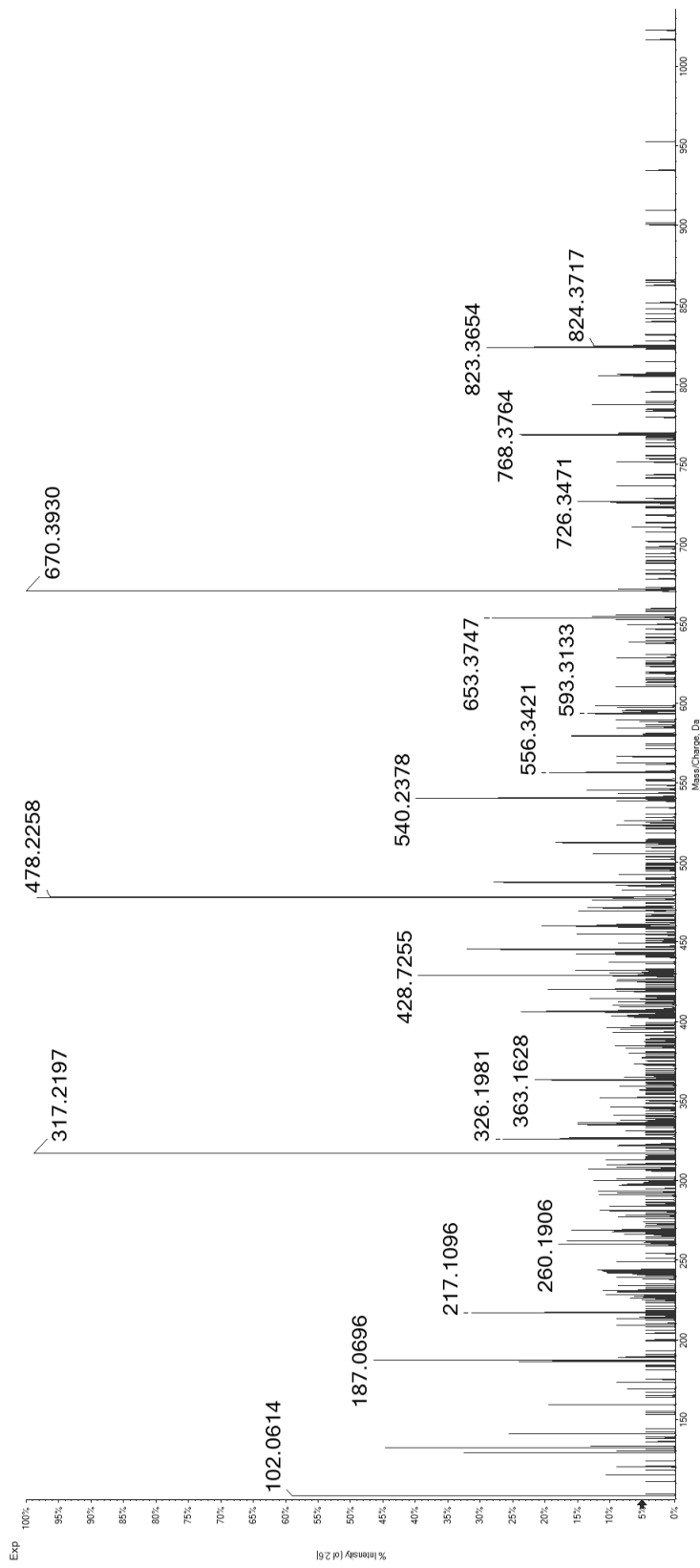
30

GEN(ps)GIGKL

Synthetic Peptide

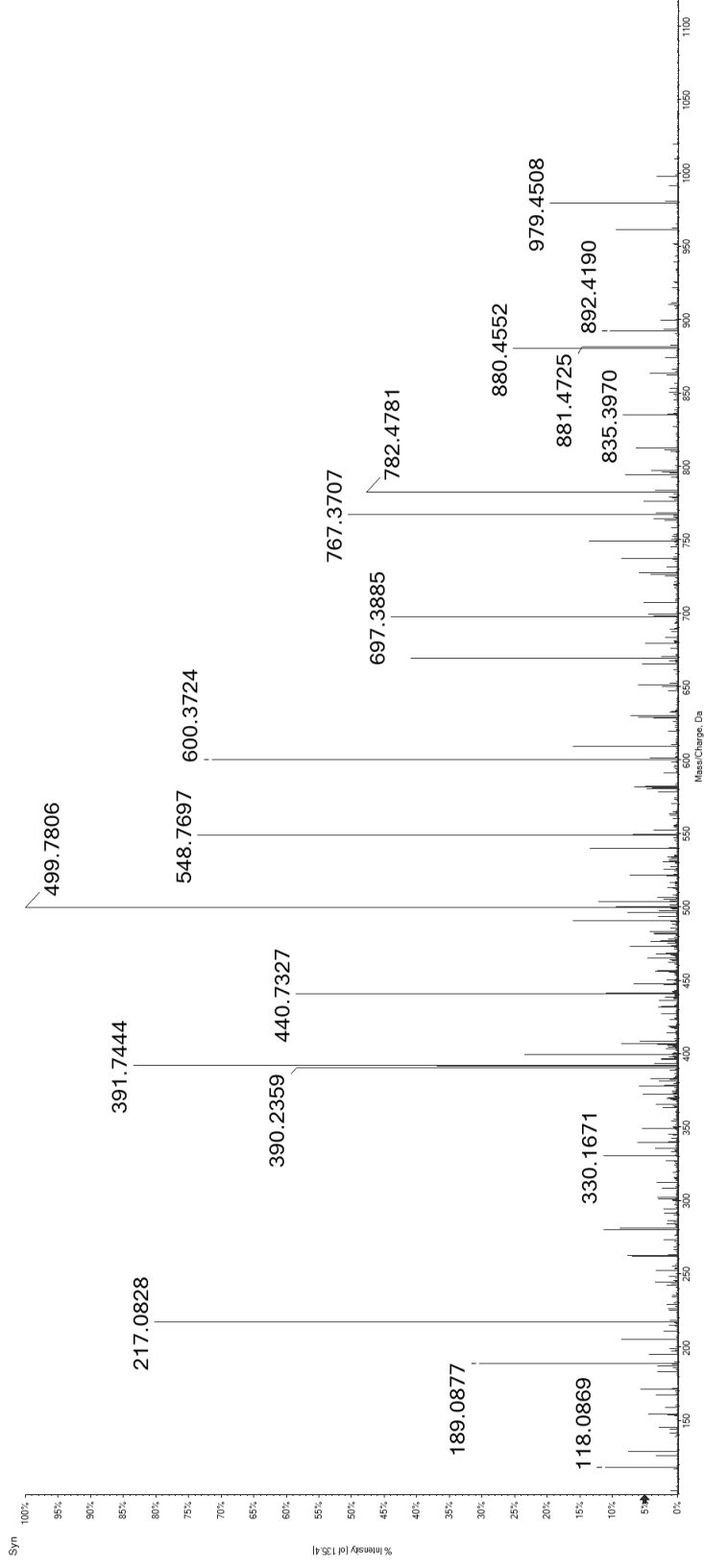


Candidate Spectrum

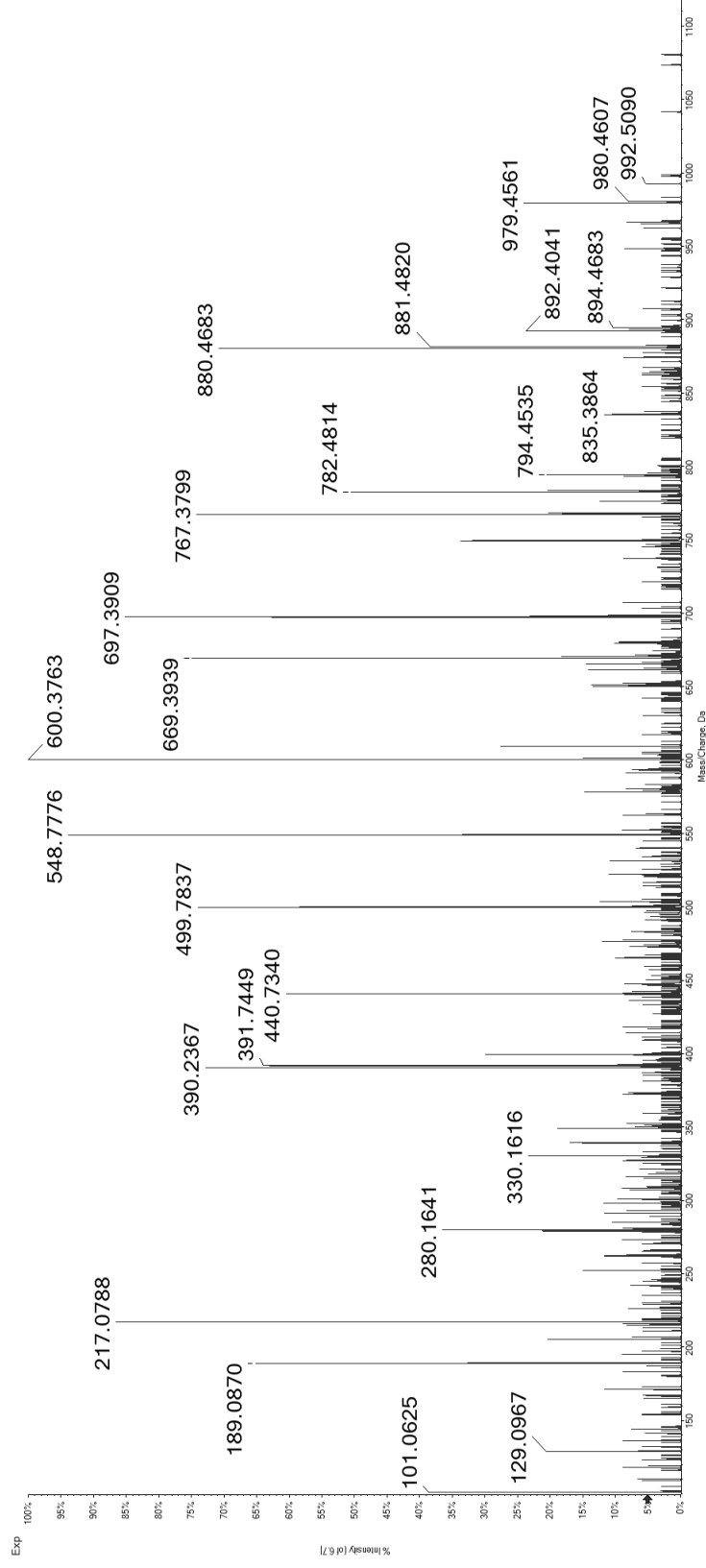


31 SEI(ps)PIKGSV

Synthetic
Peptide

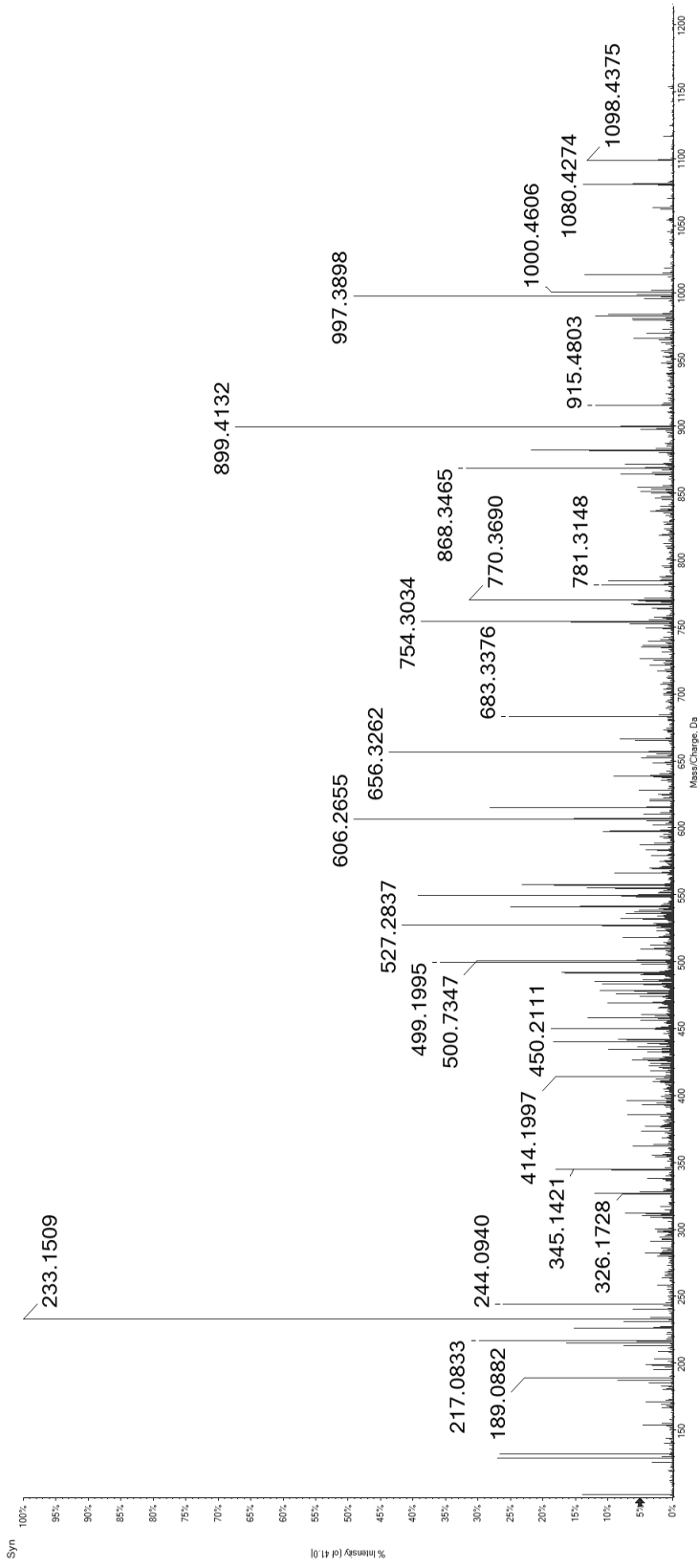


Candidate
Spectrum

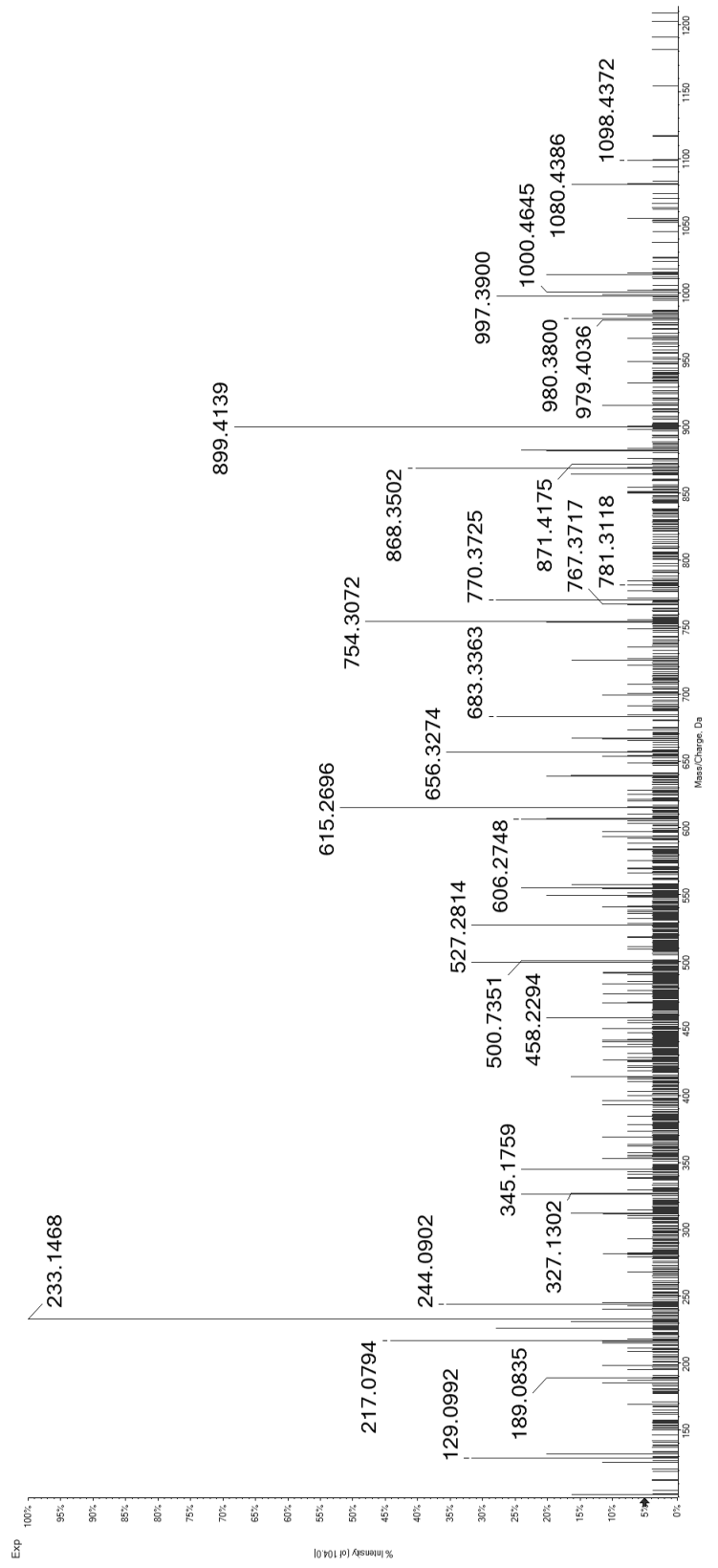


32 SEK(ps)LENETL

Synthetic
Peptide



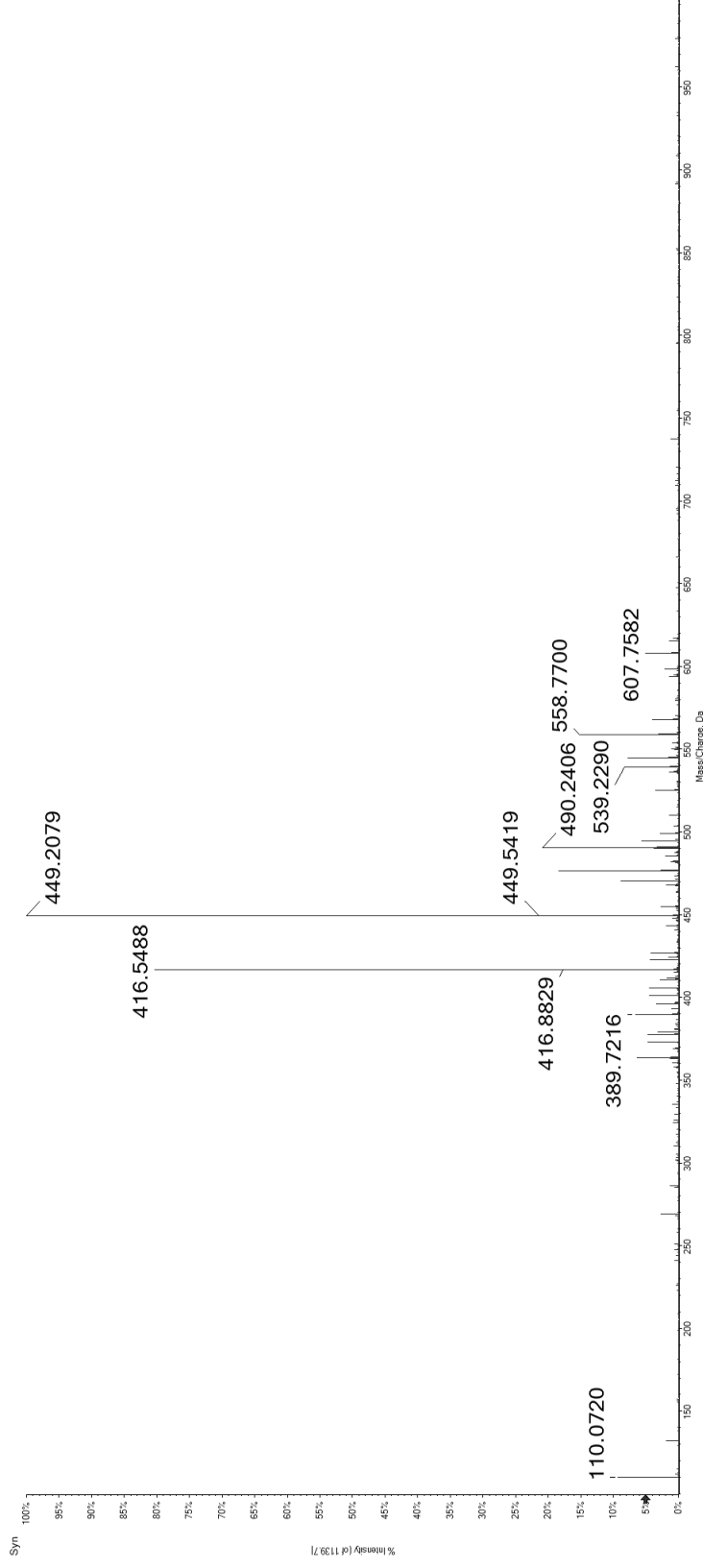
Candidate
Spectrum



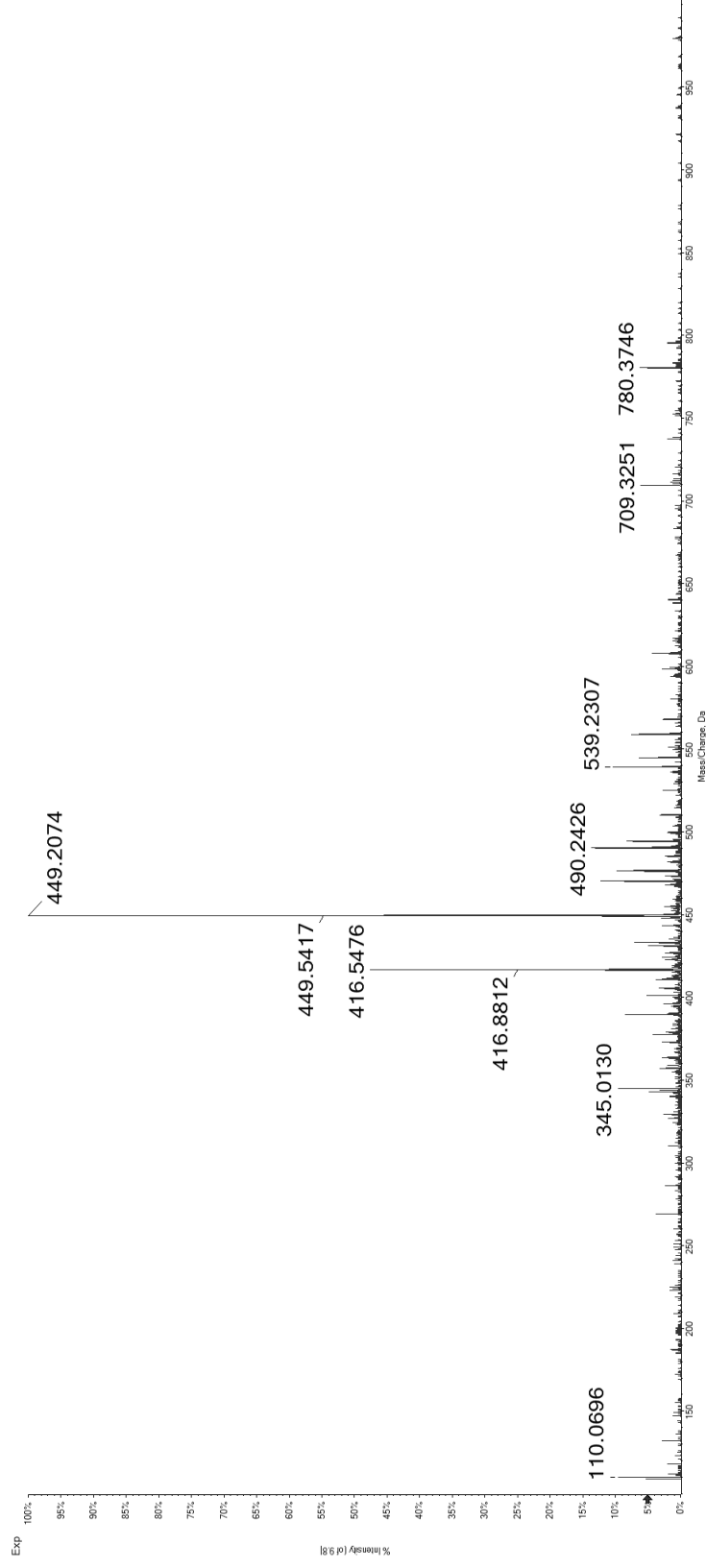
33

RED(pS)QRPGAHL

Synthetic Peptide



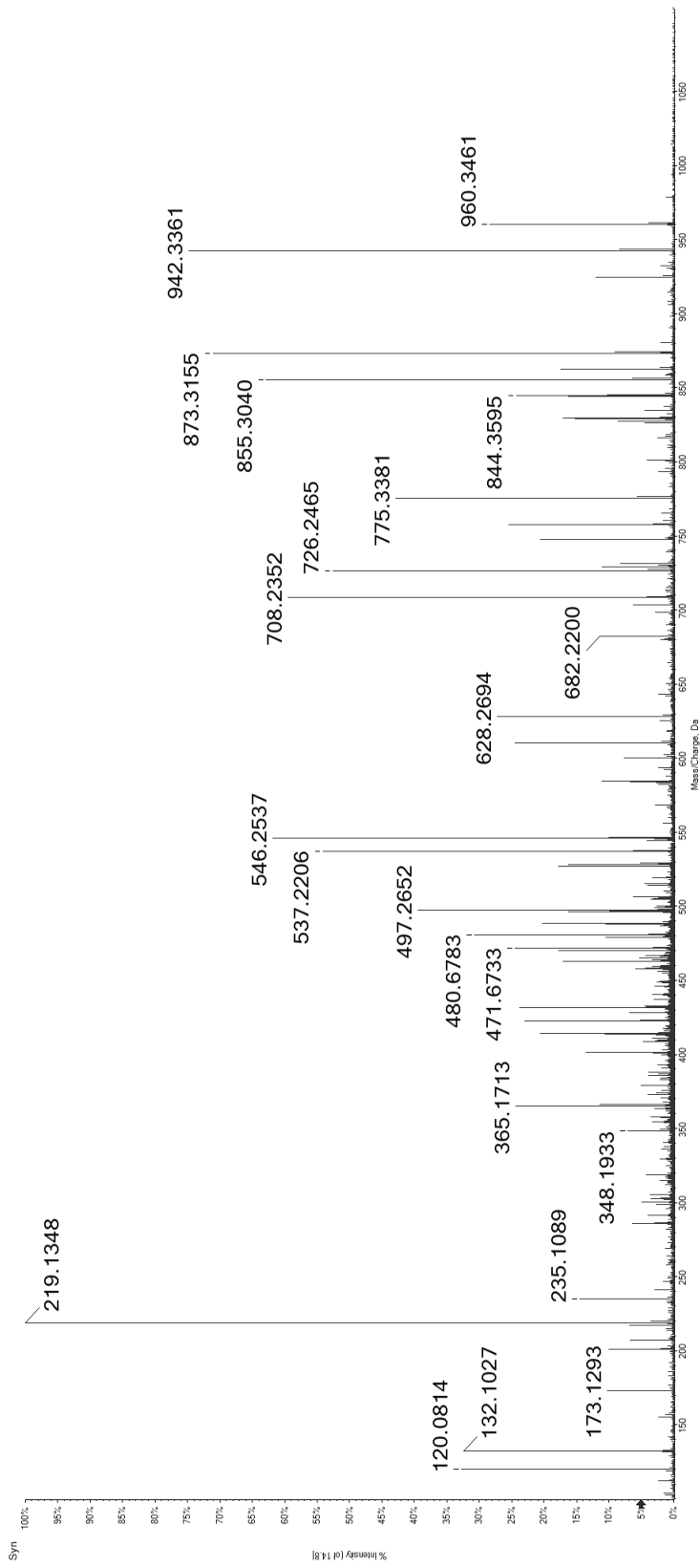
Candidate Spectrum



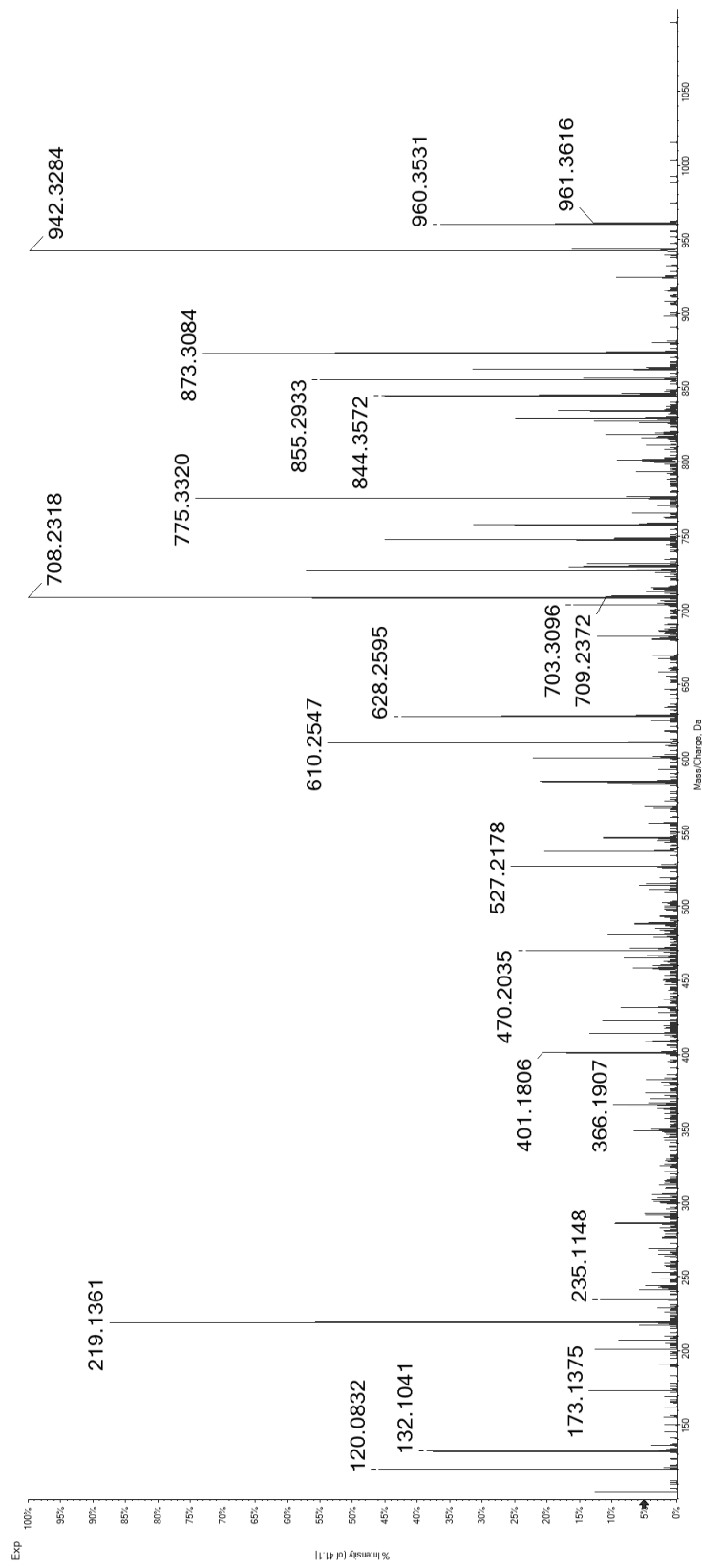
34

RED(ps)GTFSL

Synthetic Peptide



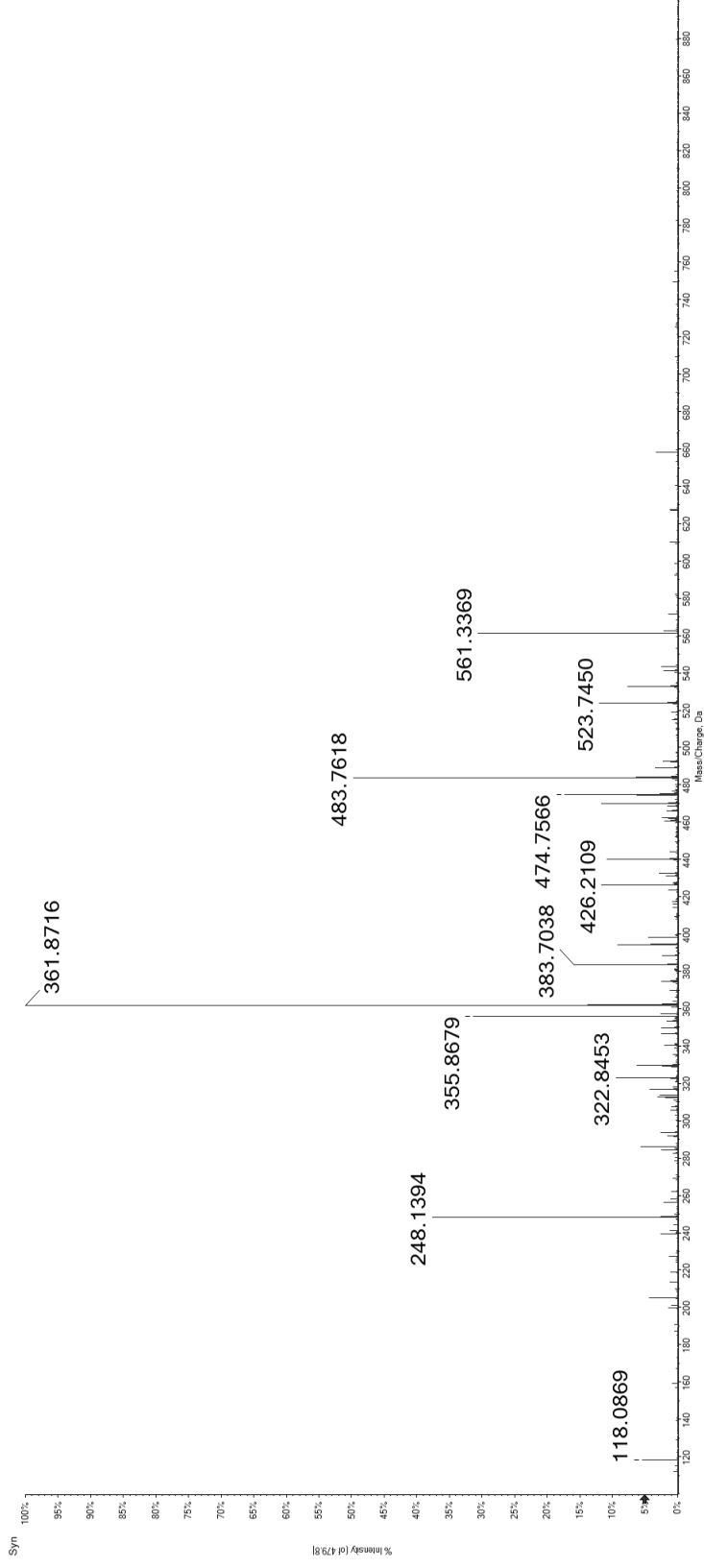
Candidate Spectrum



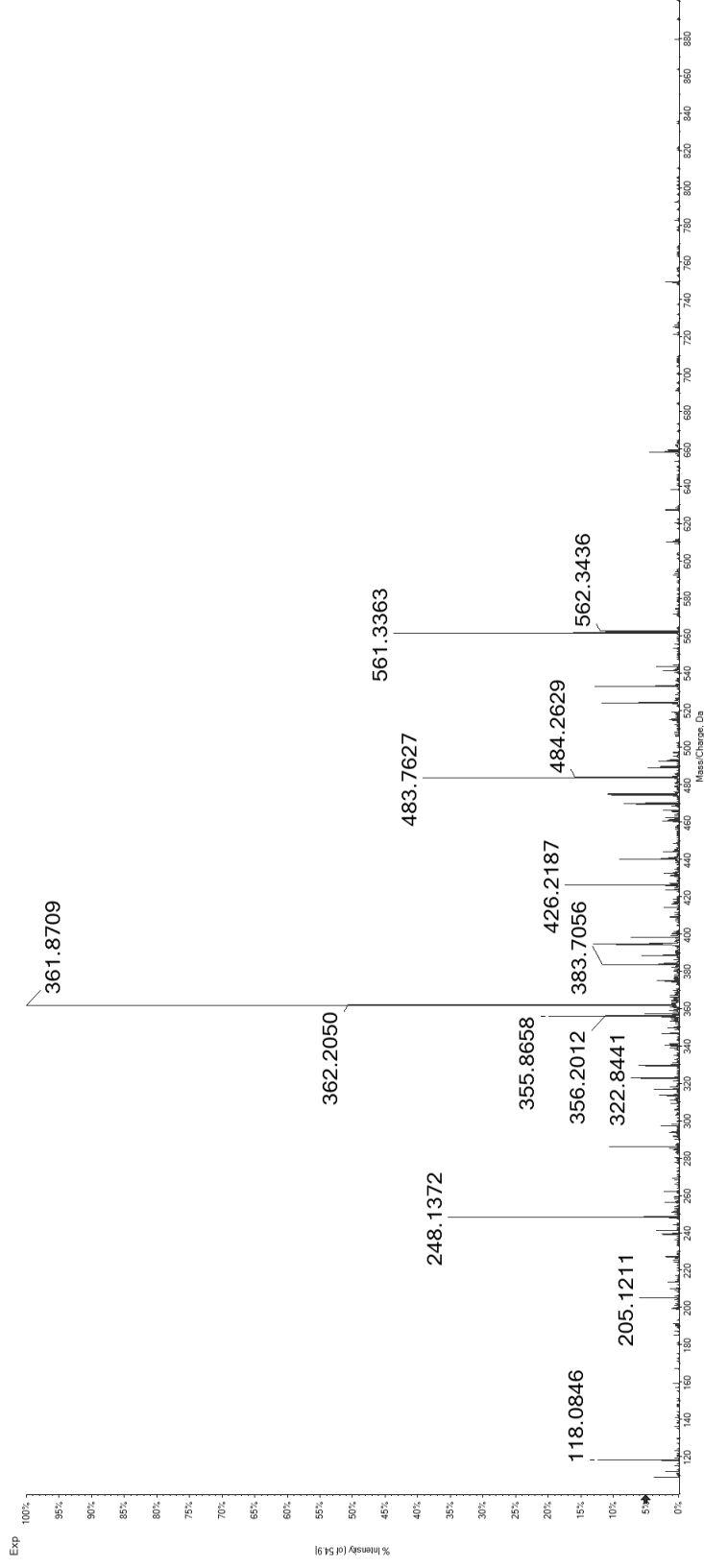
35

REA(pS)PSRLSV

Synthetic Peptide



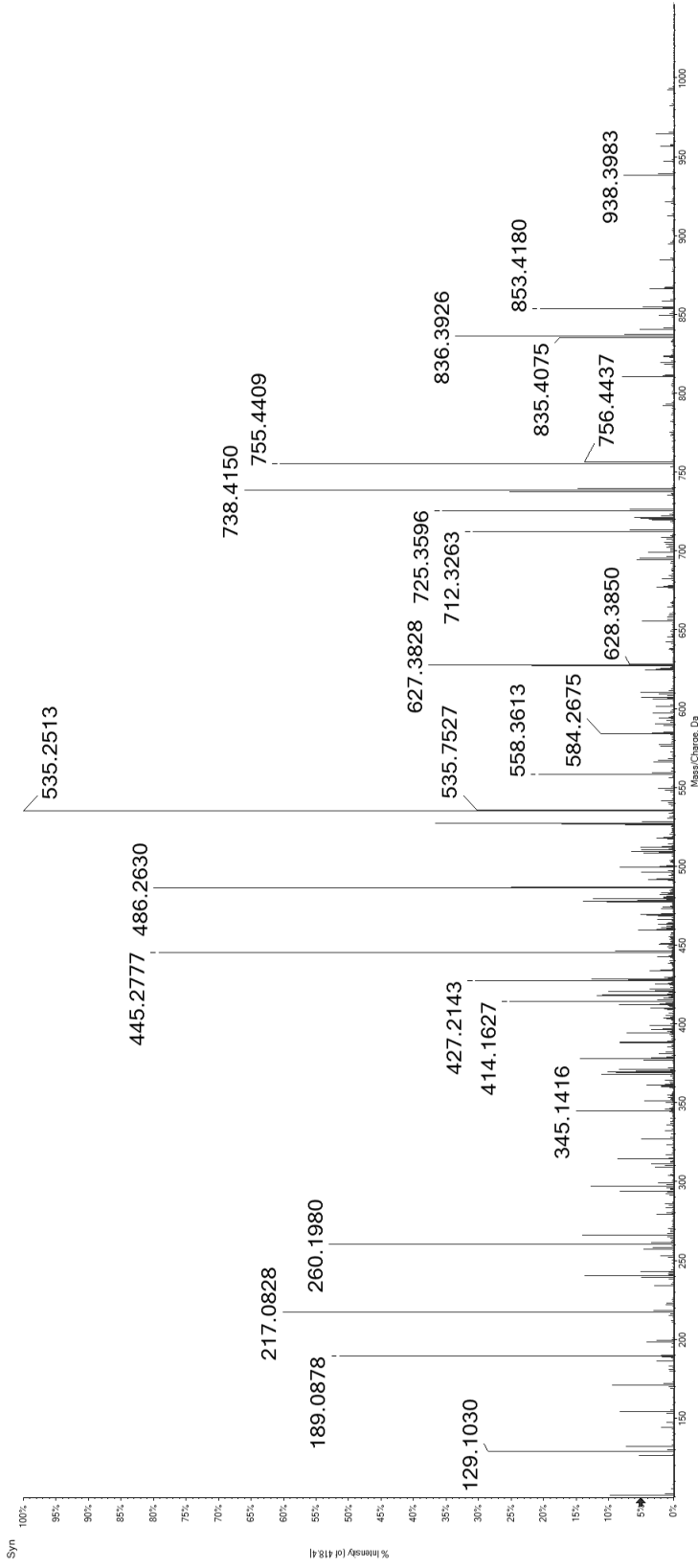
Candidate Spectrum



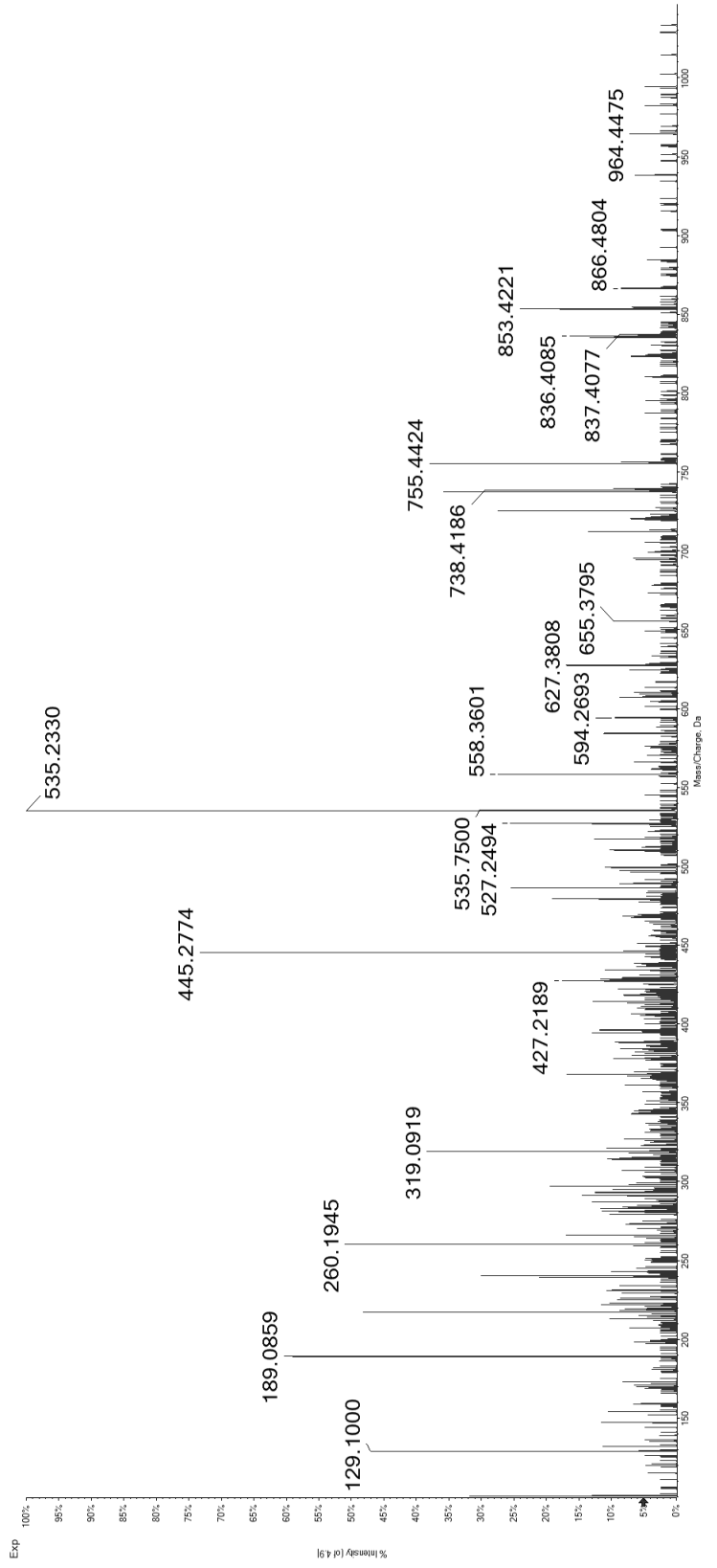
36

SEQ(pS)LGQKL

Synthetic Peptide



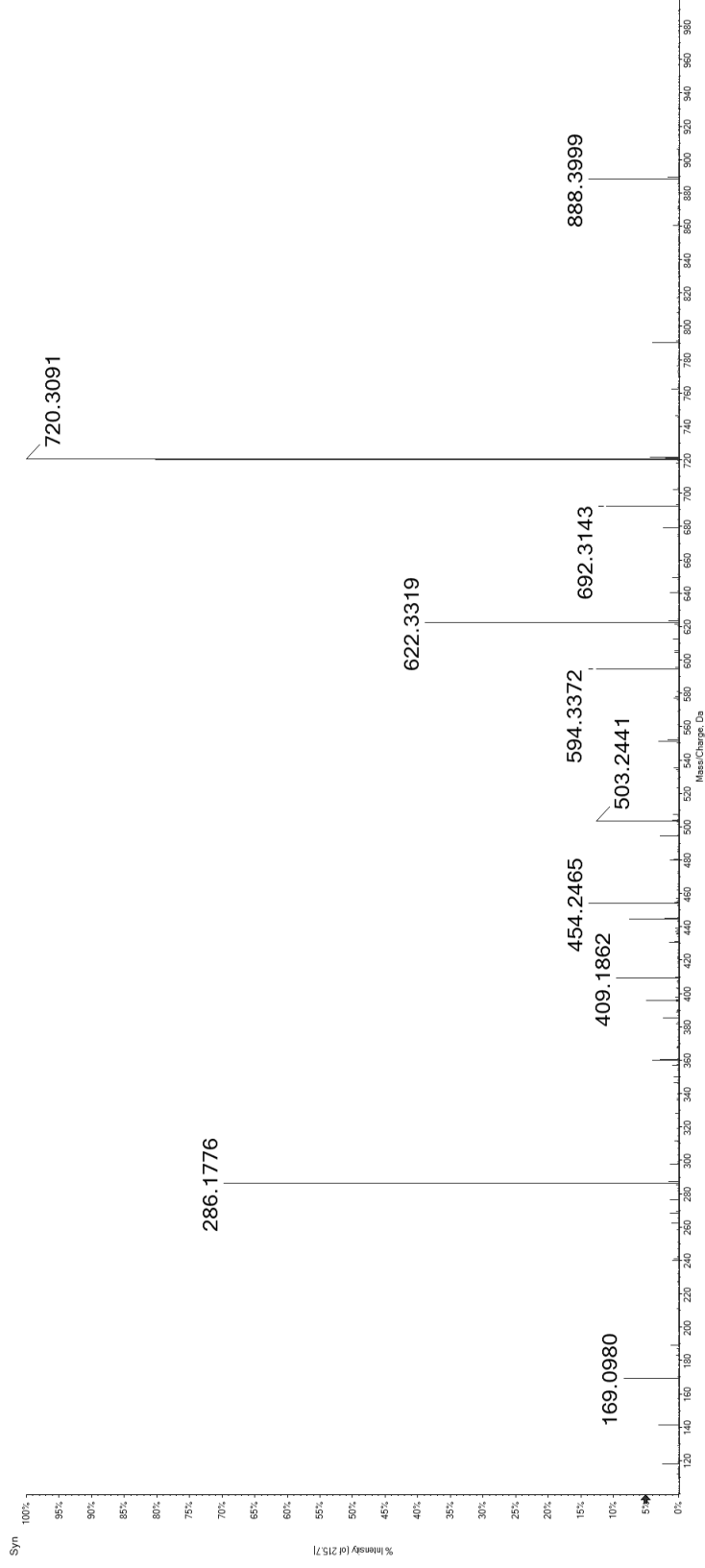
Candidate Spectrum



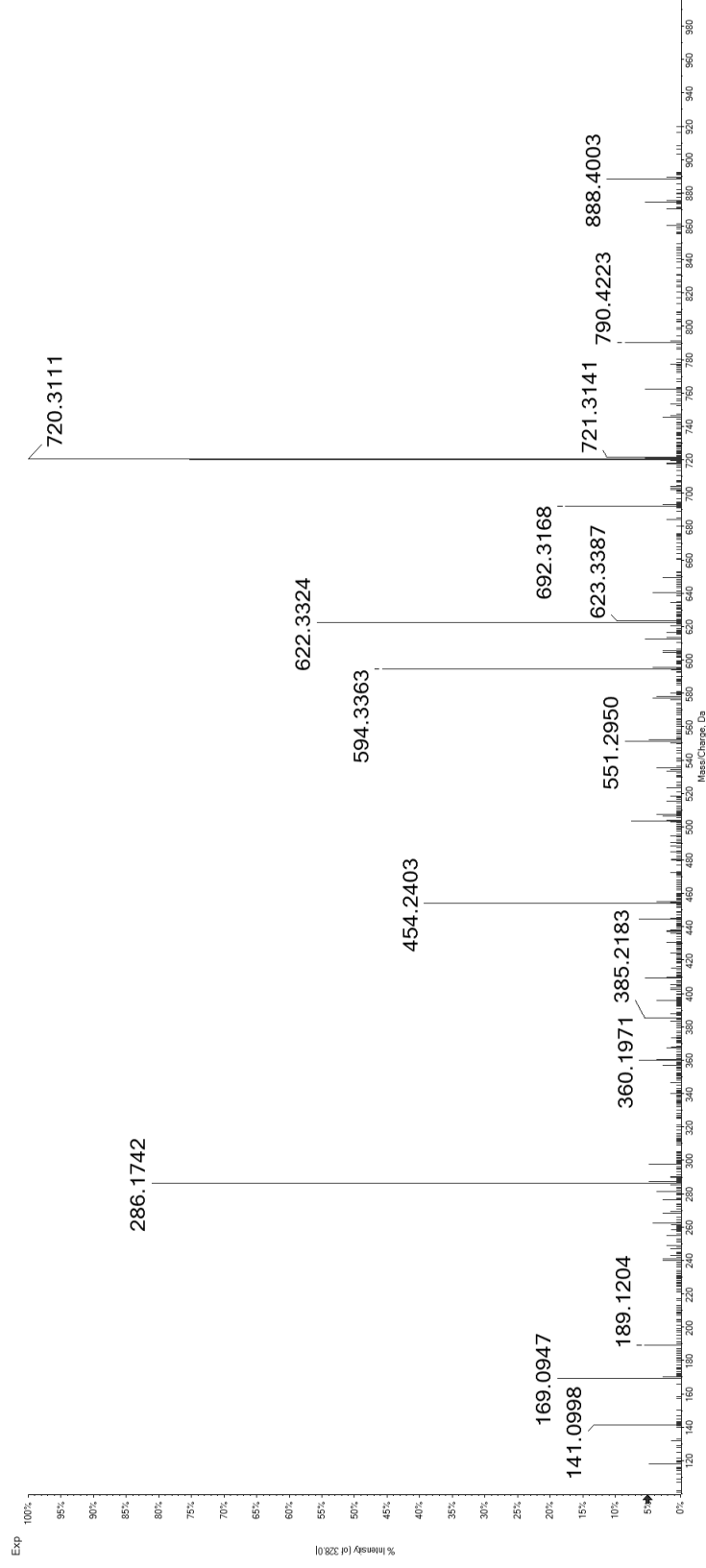
37

REV(pS)PAPAV

Synthetic Peptide



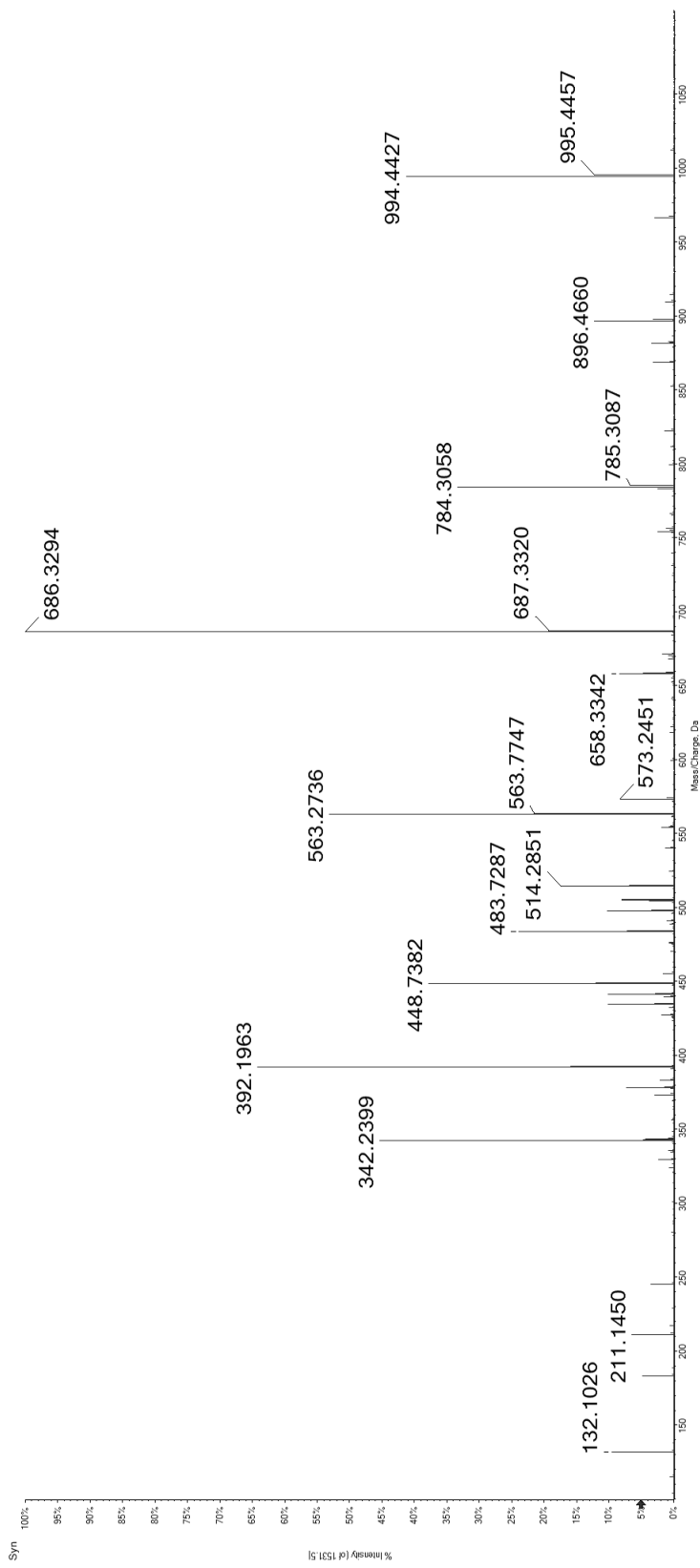
Candidate Spectrum



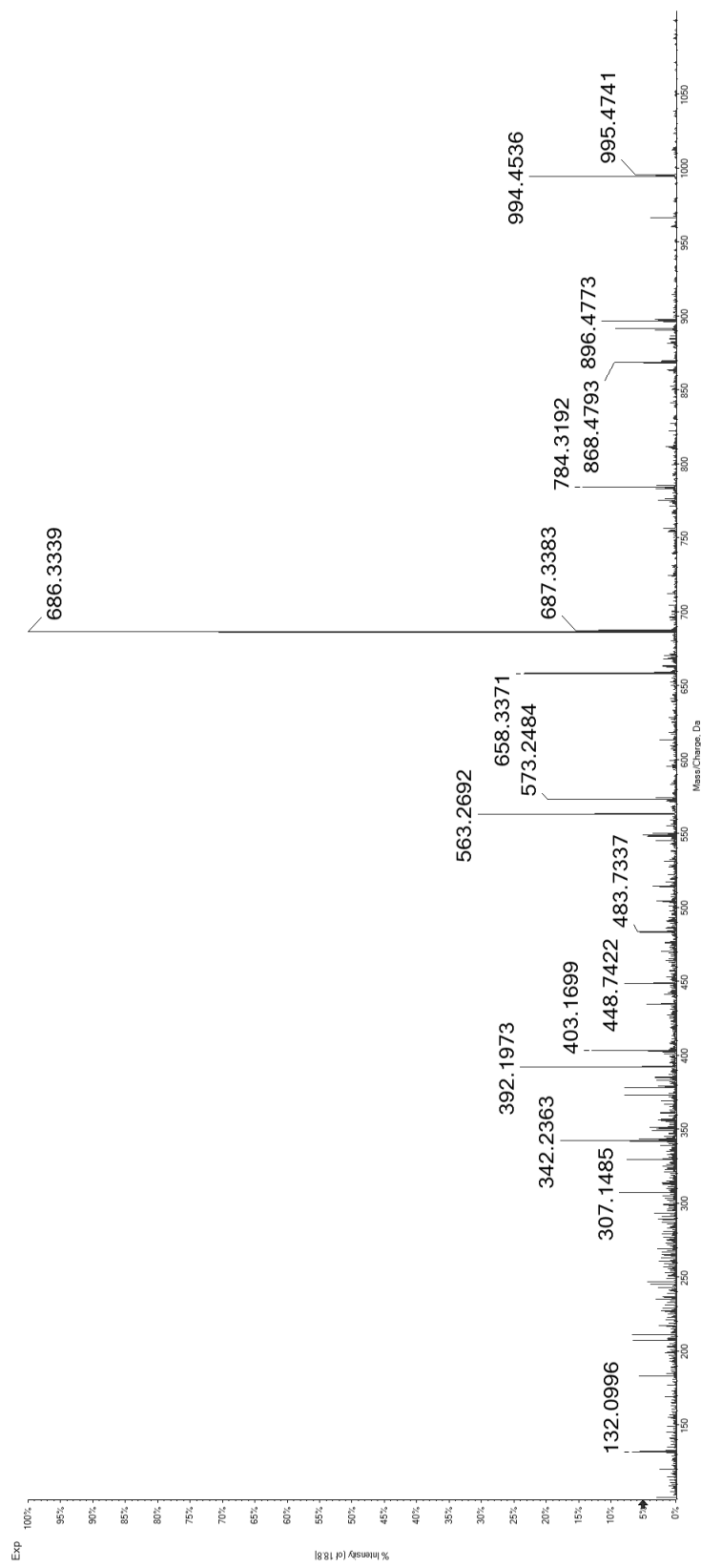
38

SERM(pS)LPLL

Synthetic Peptide



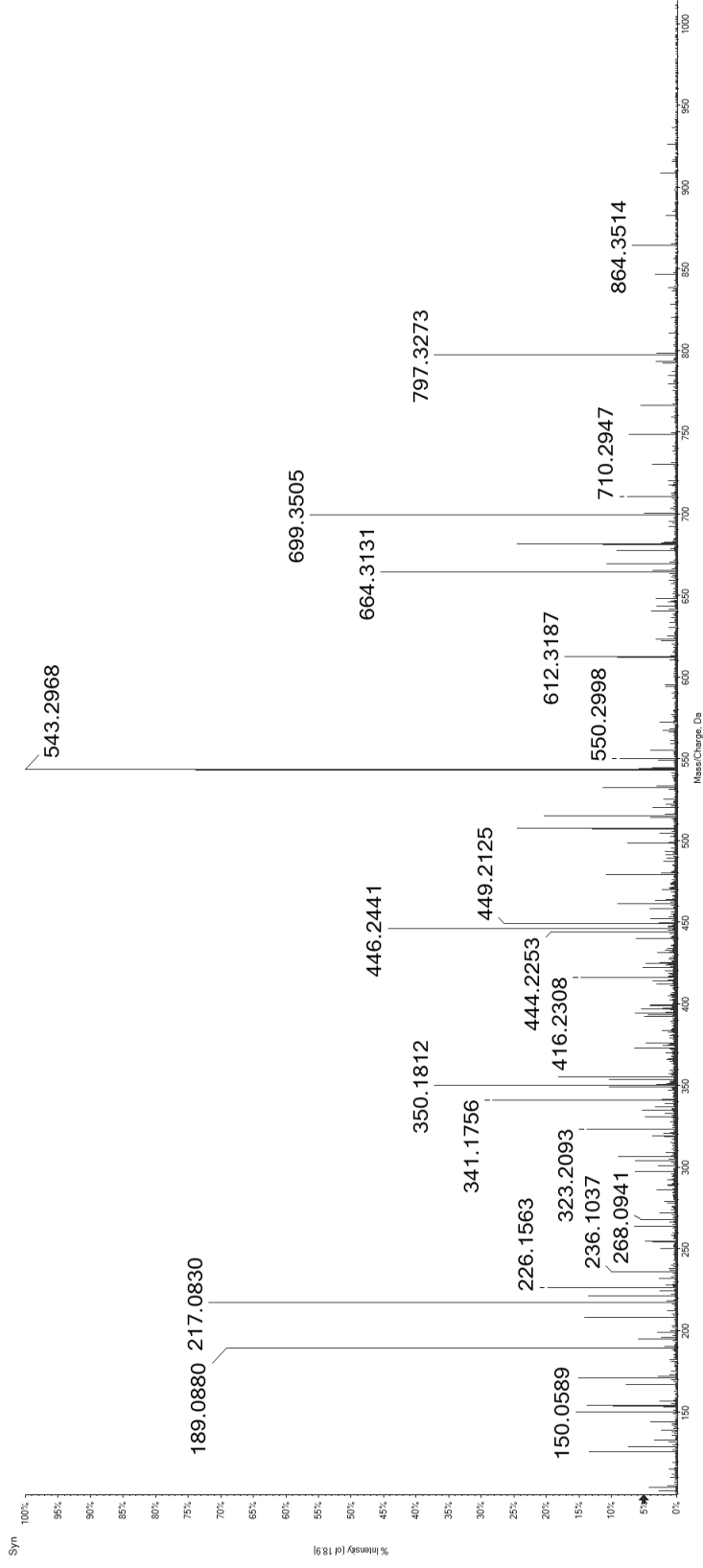
Candidate Spectrum



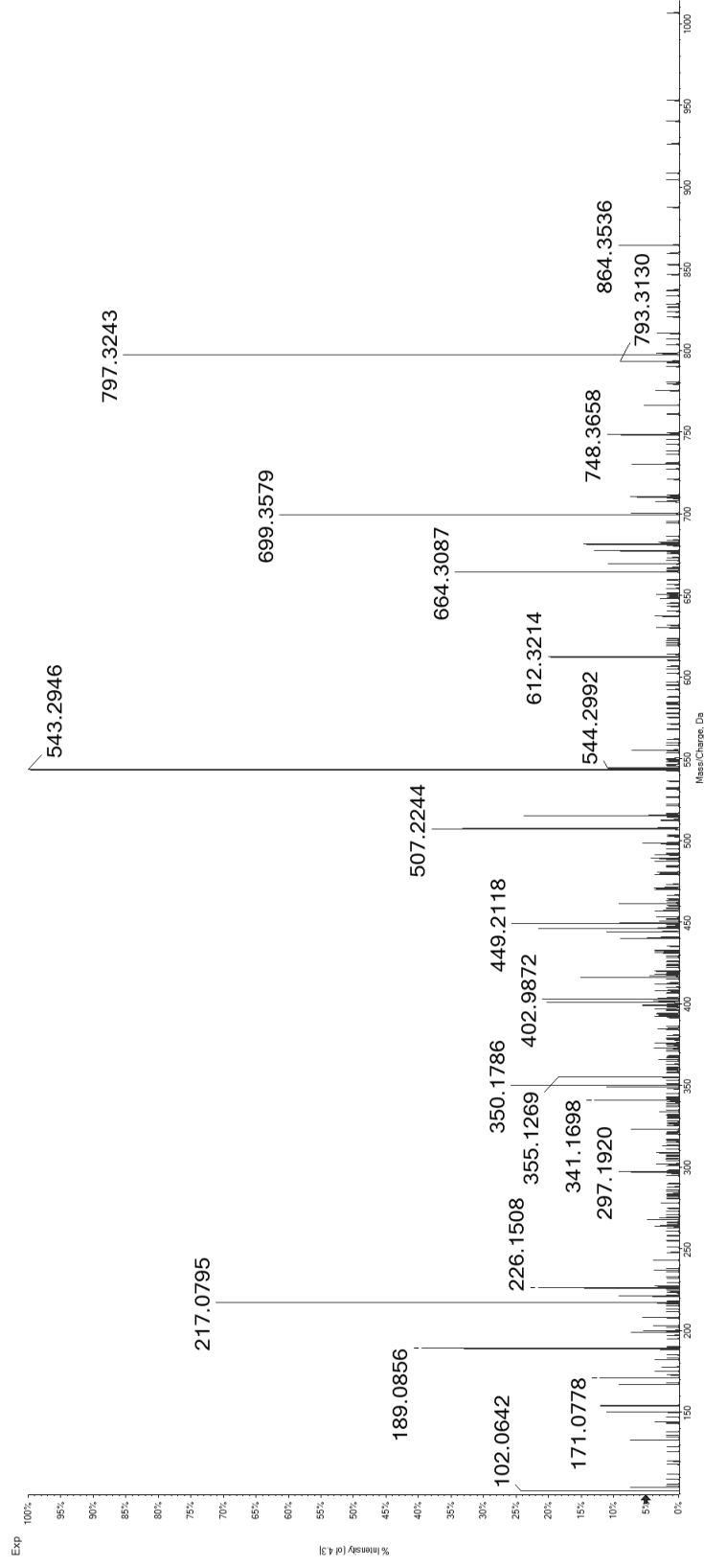
39

SES(pS)PPKAM

Synthetic Peptide



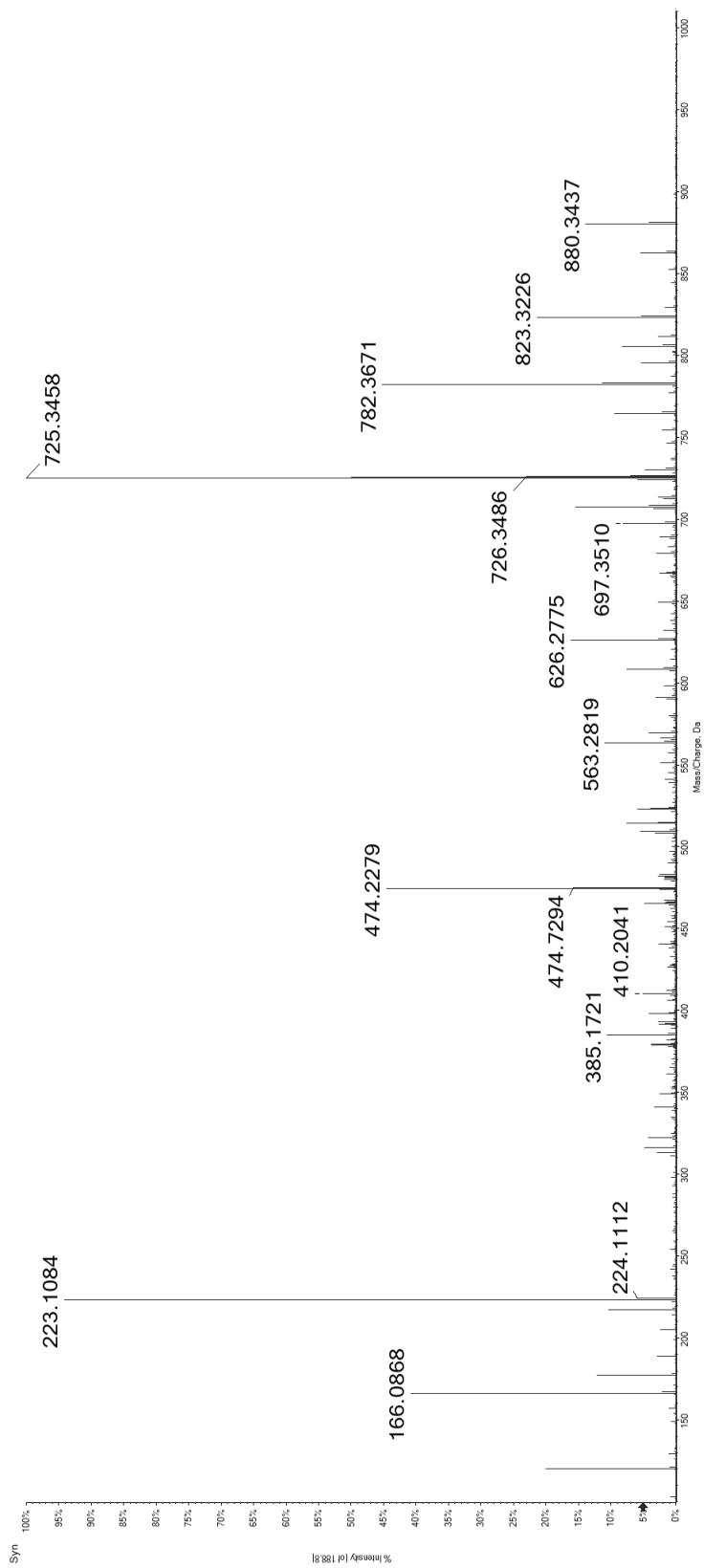
Candidate Spectrum



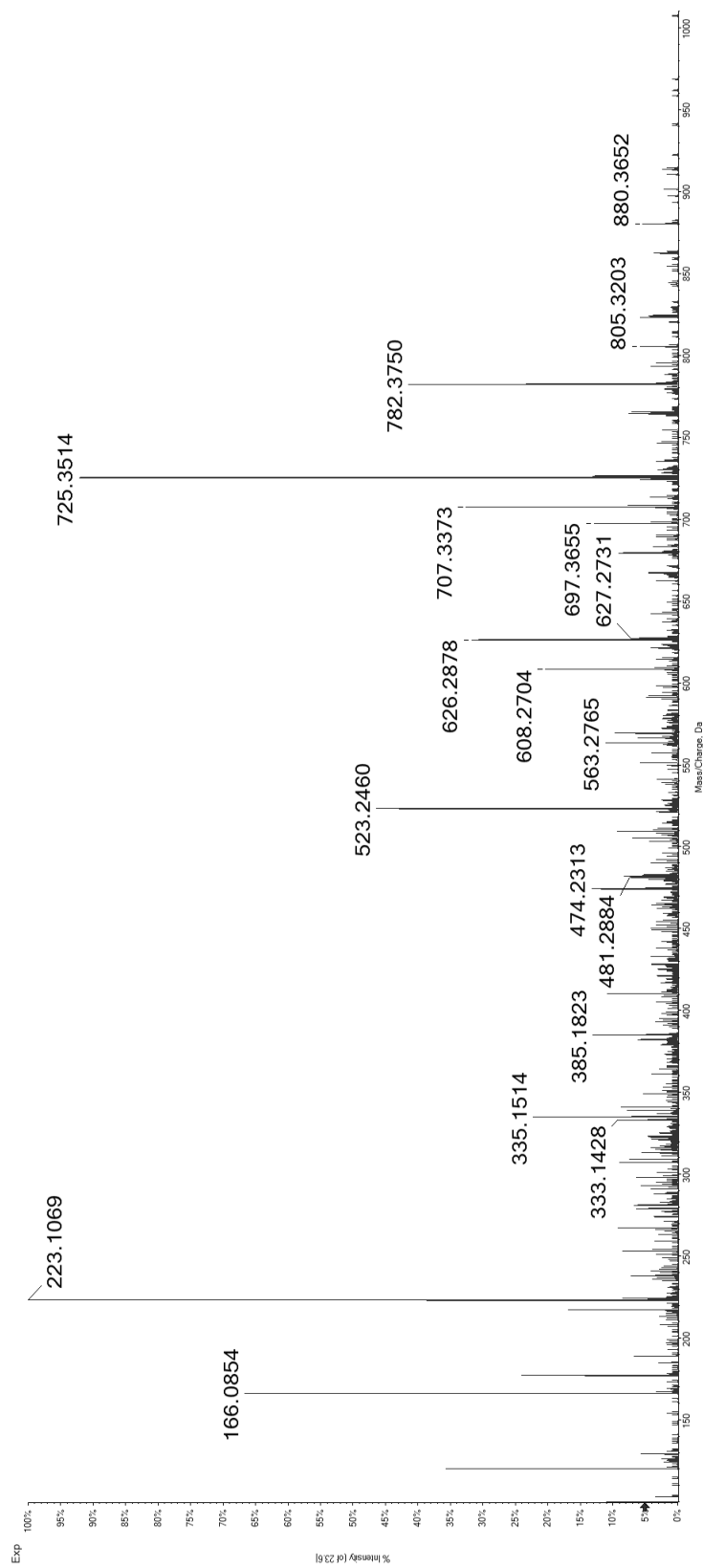
40

SEV(pS)PSGVGF

Synthetic Peptide



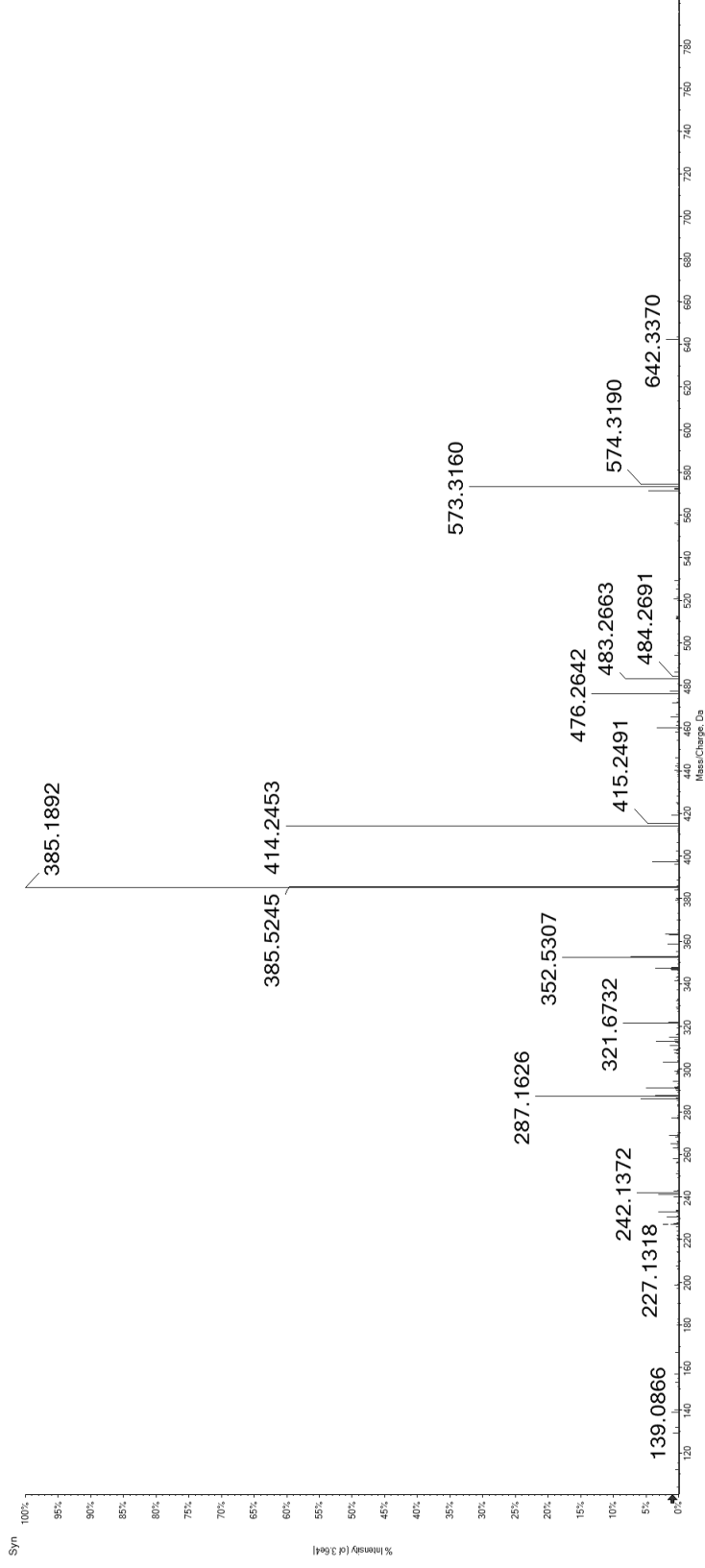
Candidate Spectrum



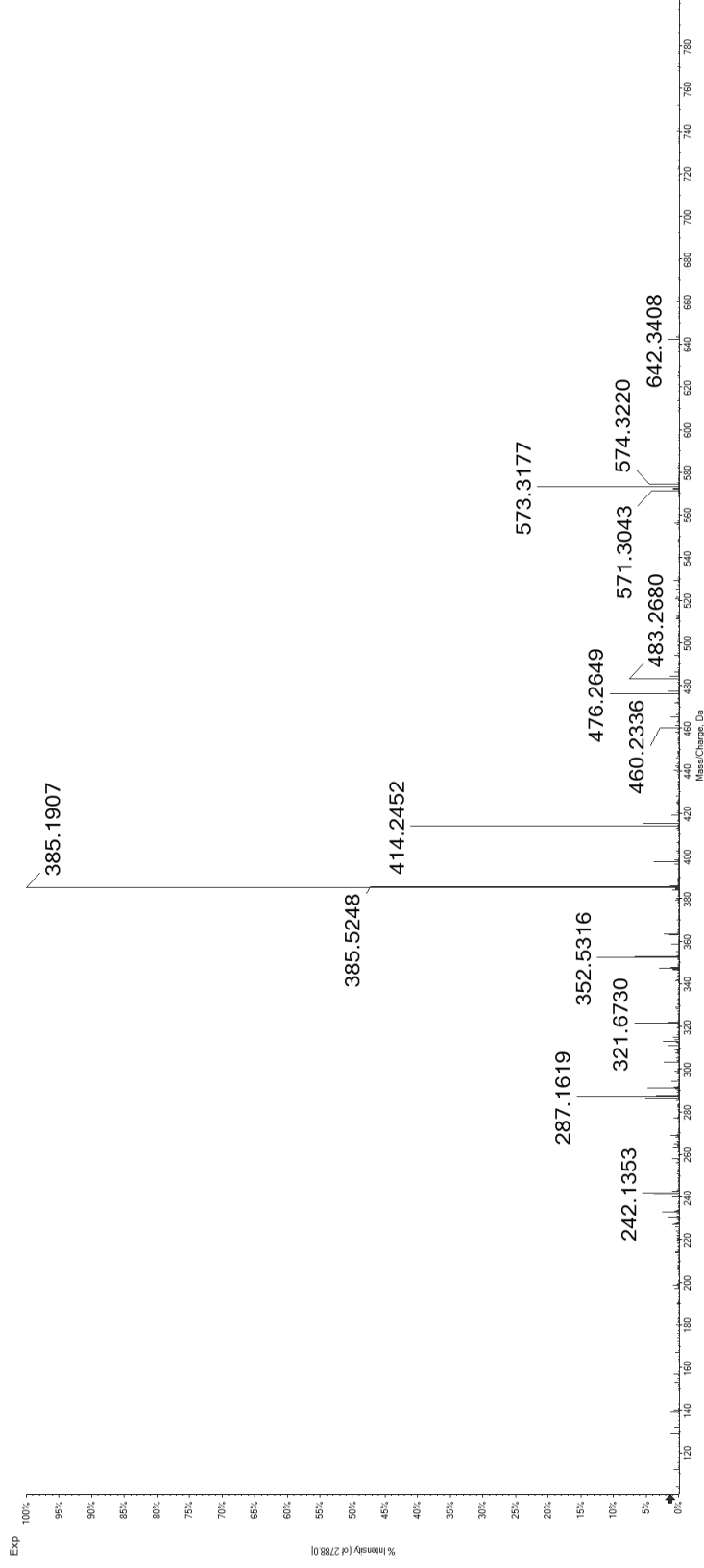
41

REK(pS)PGRML

Synthetic Peptide

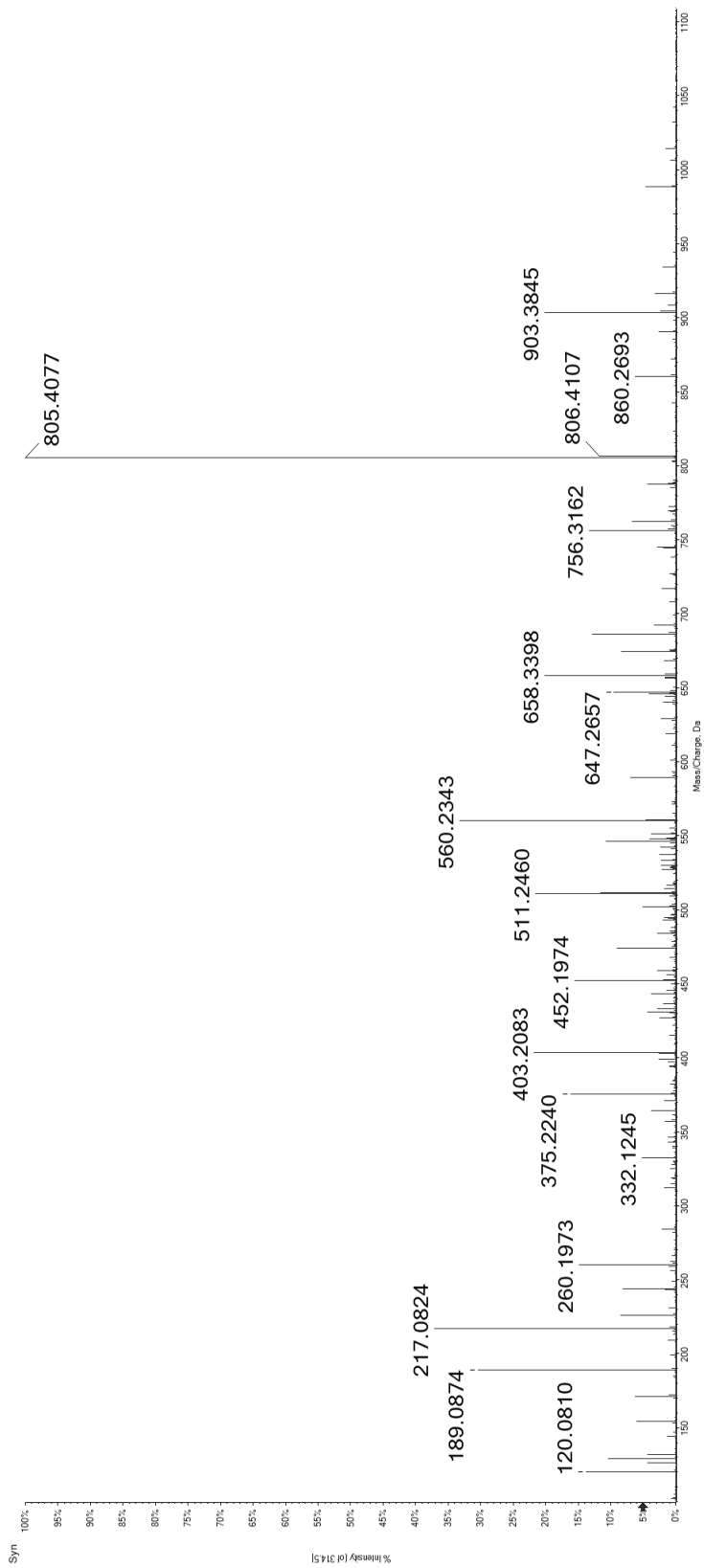


Candidate Spectrum

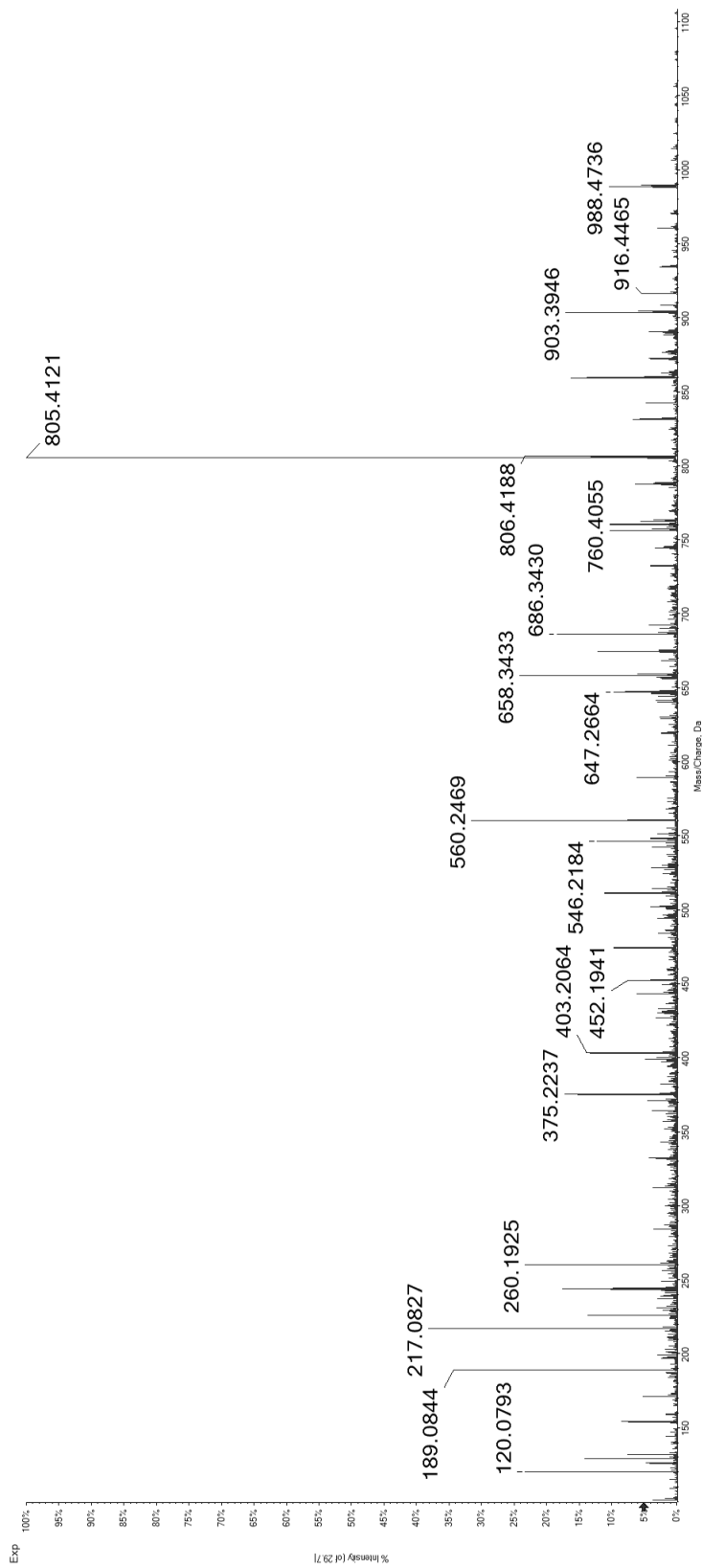


42 SEF(pS)DVDKL

Synthetic
Peptide



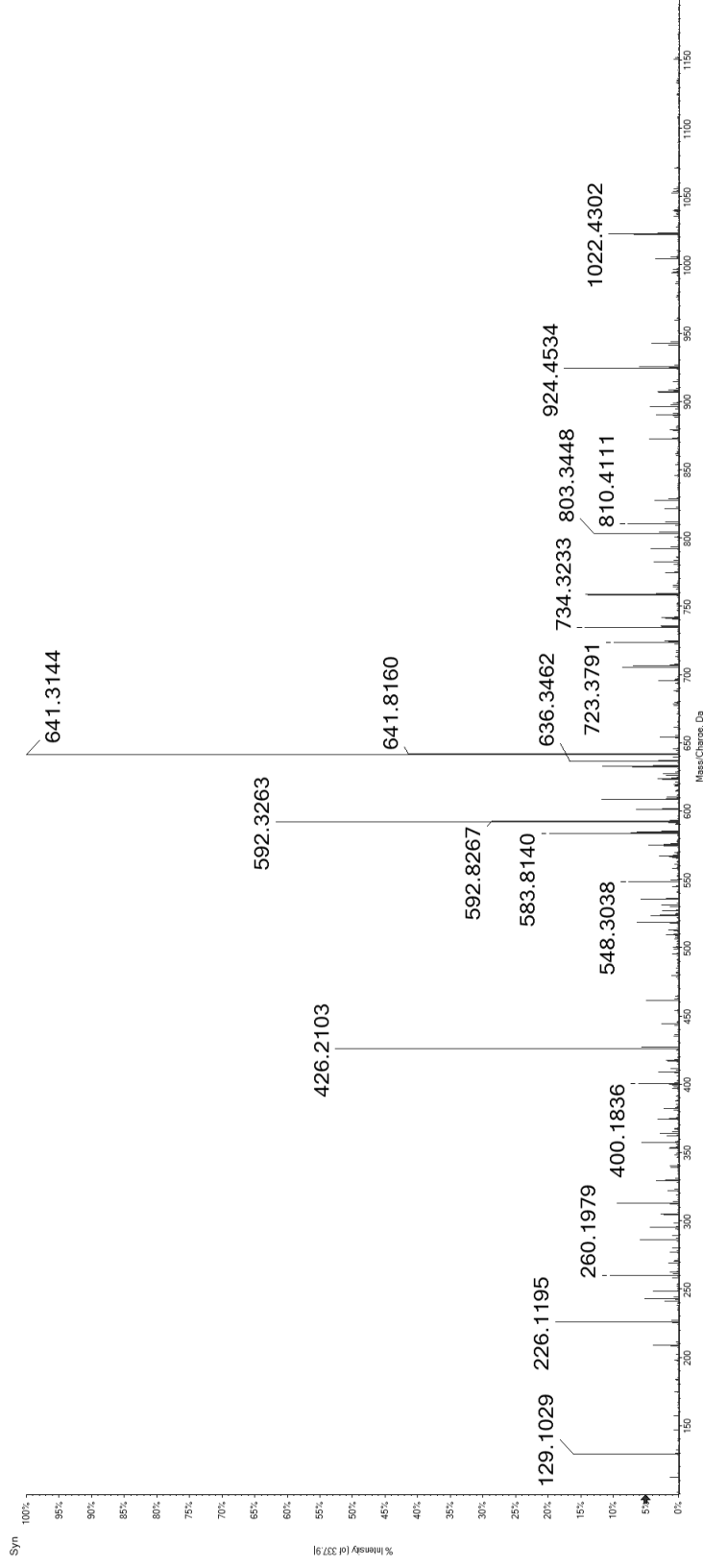
Candidate
Spectrum



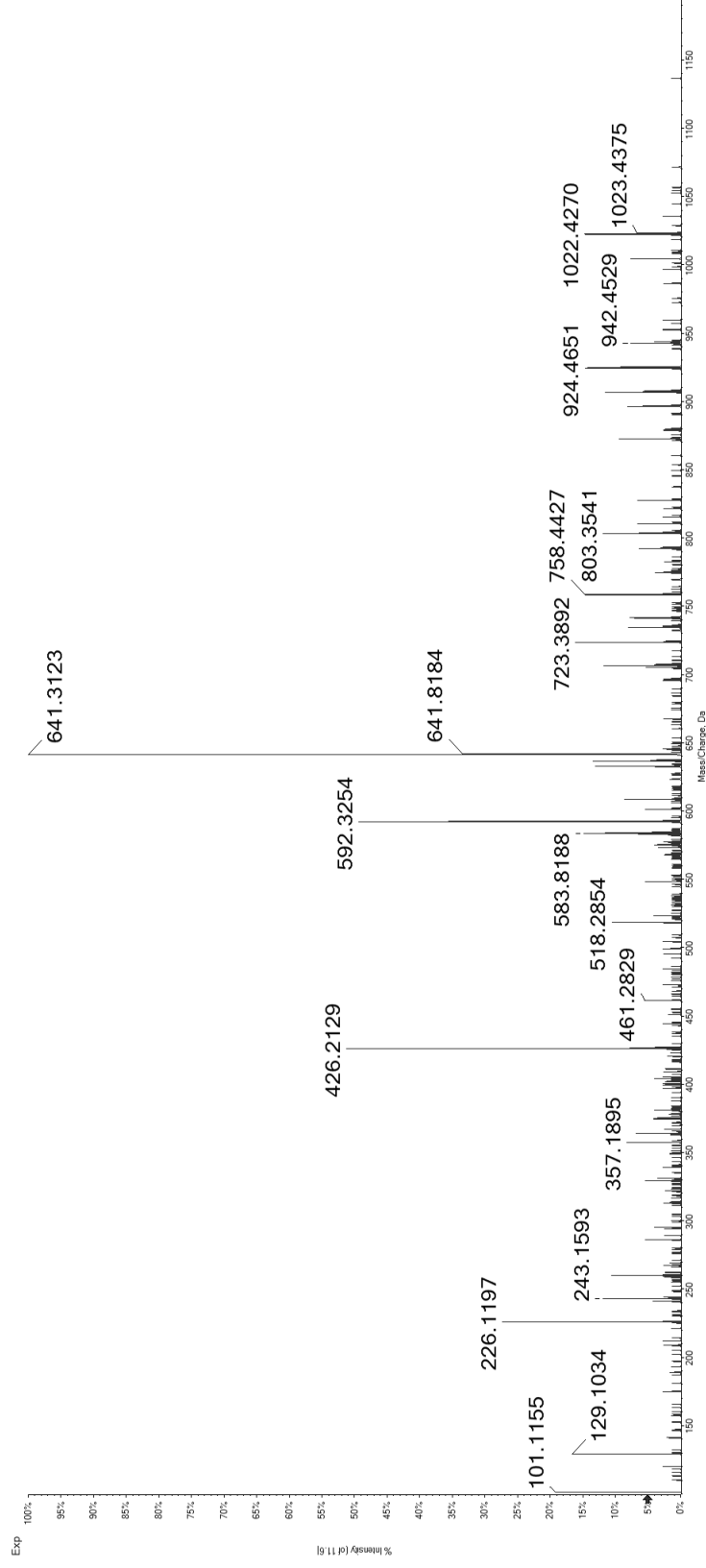
43

REA(pS)PLSSNKL

Synthetic Peptide



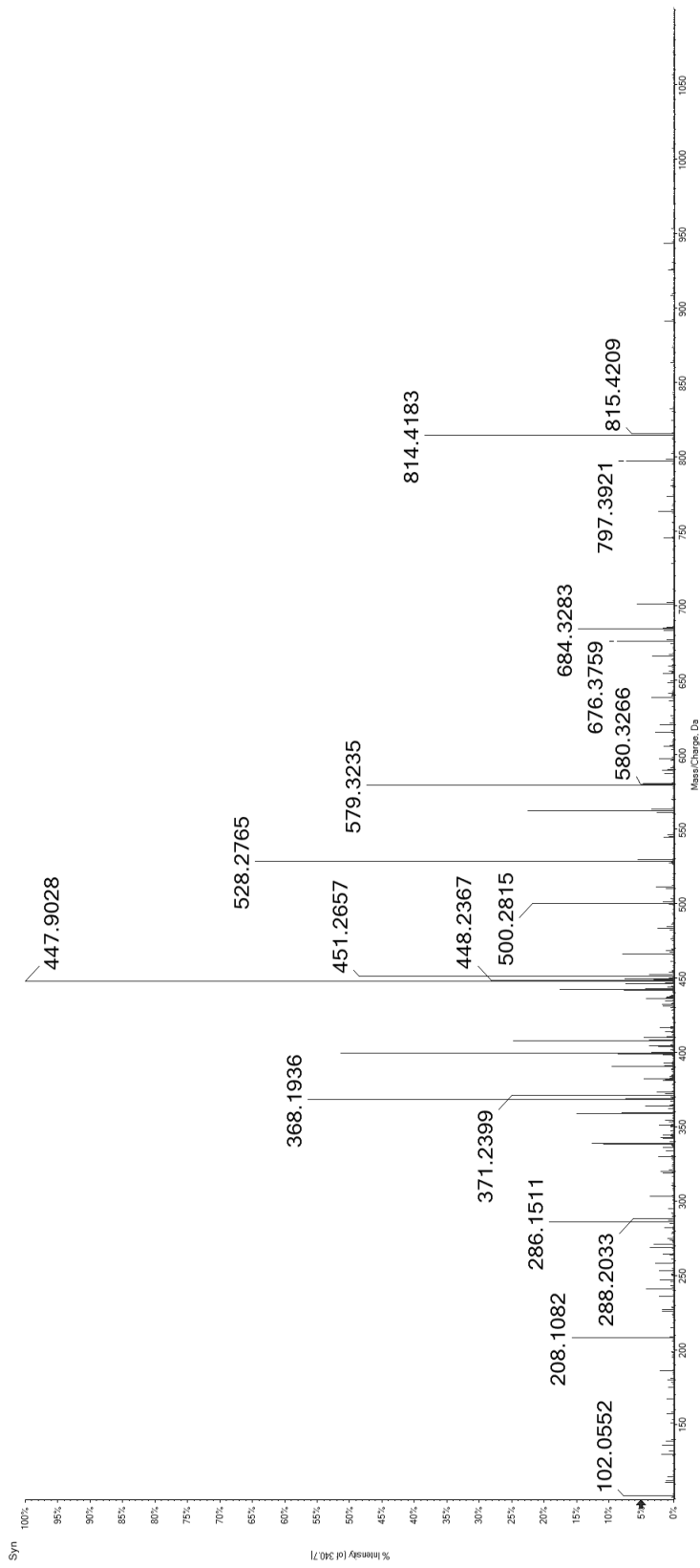
Candidate Spectrum



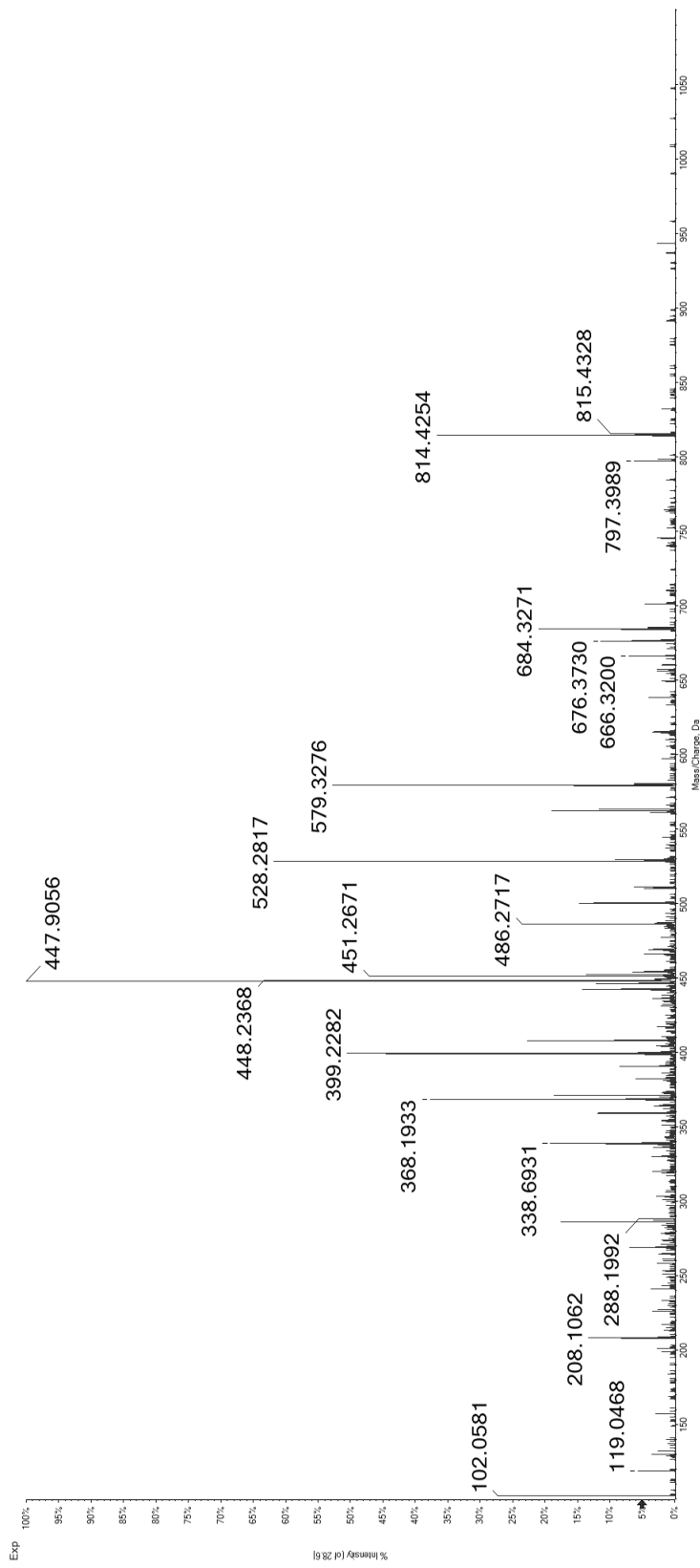
44

REIES(pS)PQYRL

Synthetic Peptide



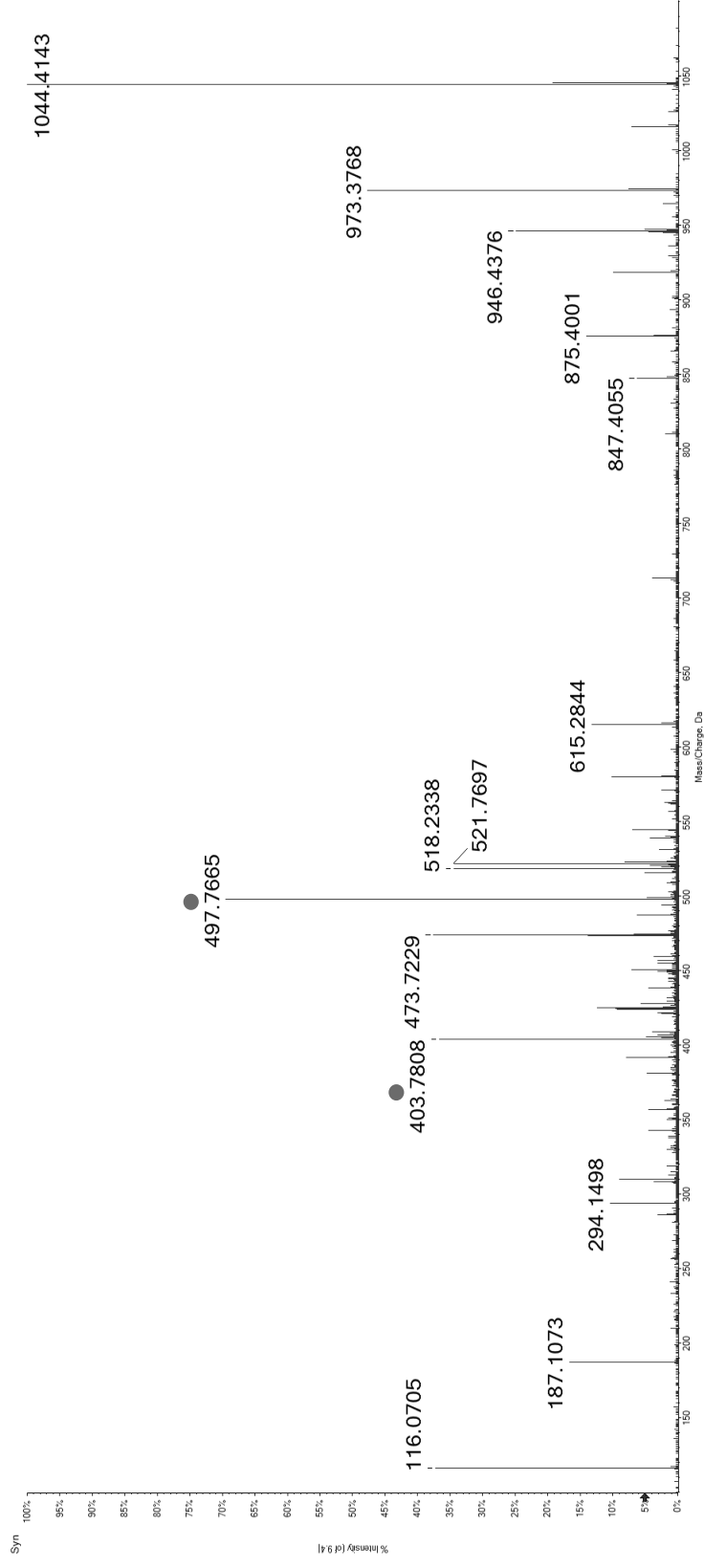
Candidate Spectrum



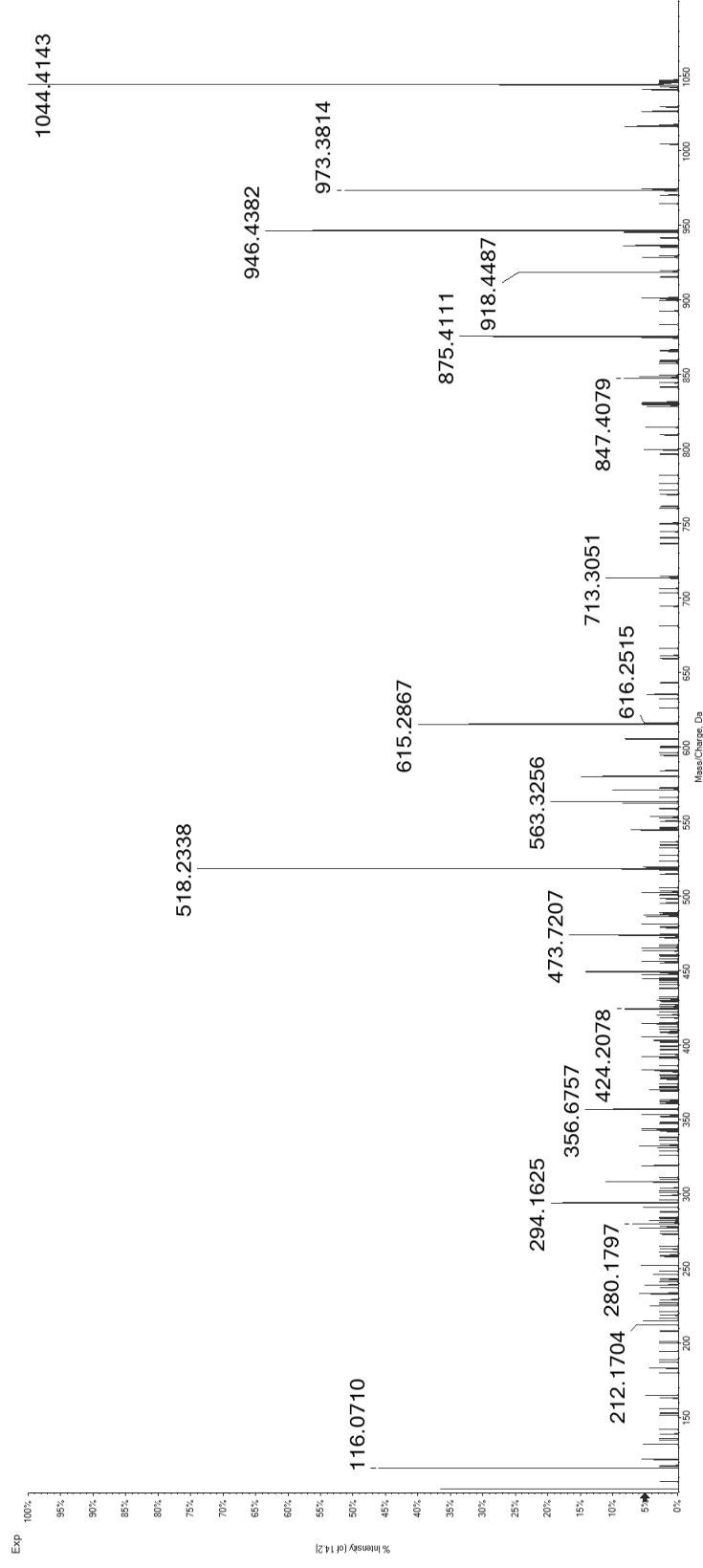
45

REY(pS)PPYAP

Synthetic
Peptide
(Relative RT – 1,61)



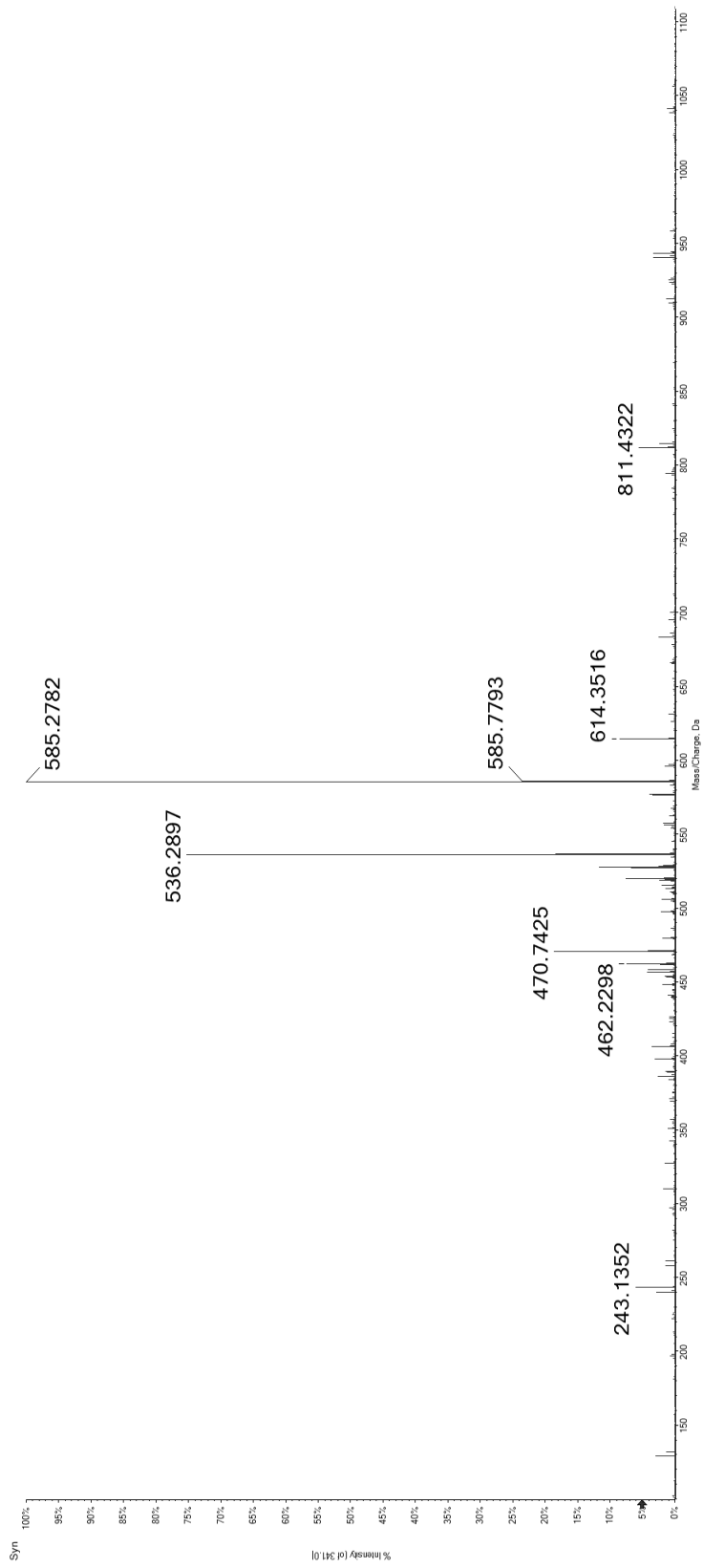
Candidate
Spectrum
(Relative RT – 1,60)



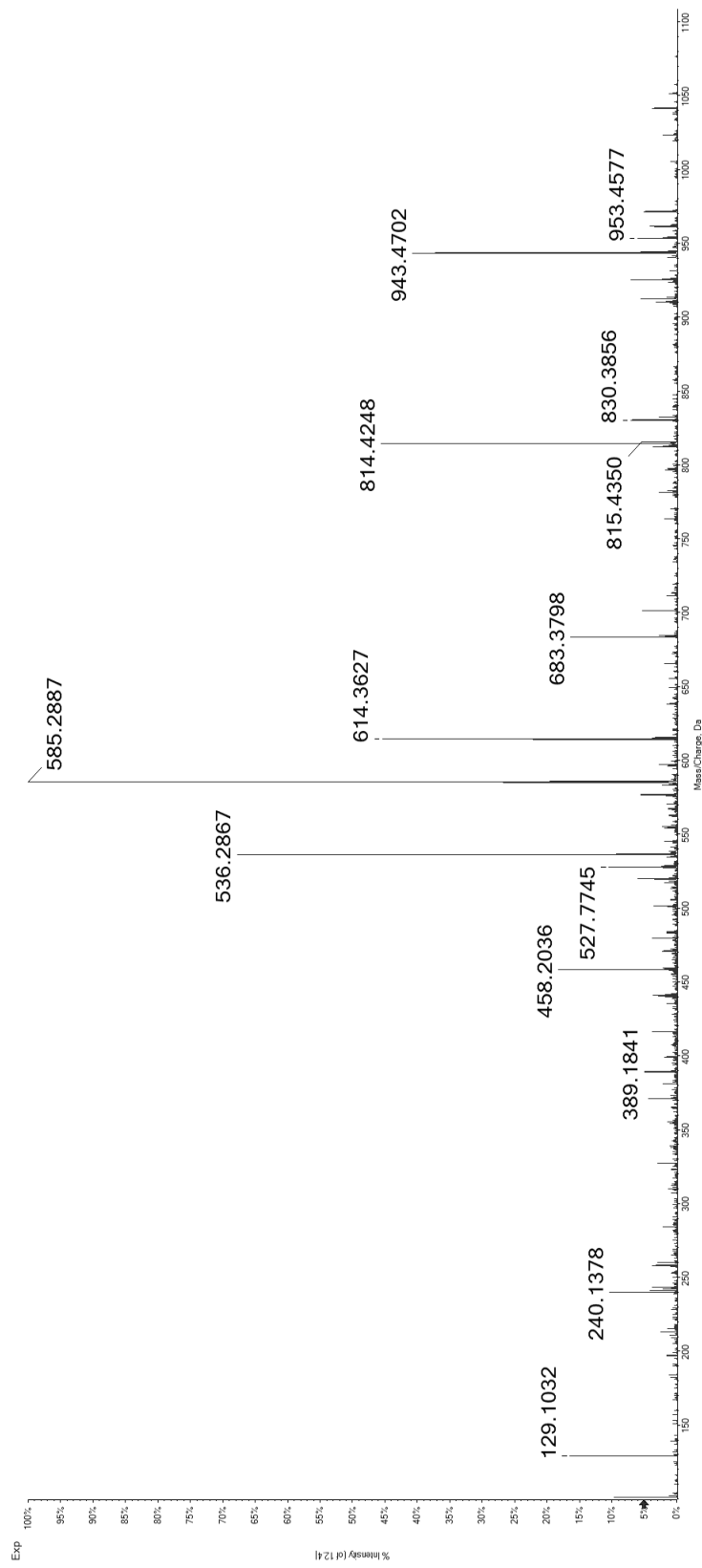
46

KEKEV(ps)QEL

Synthetic Peptide



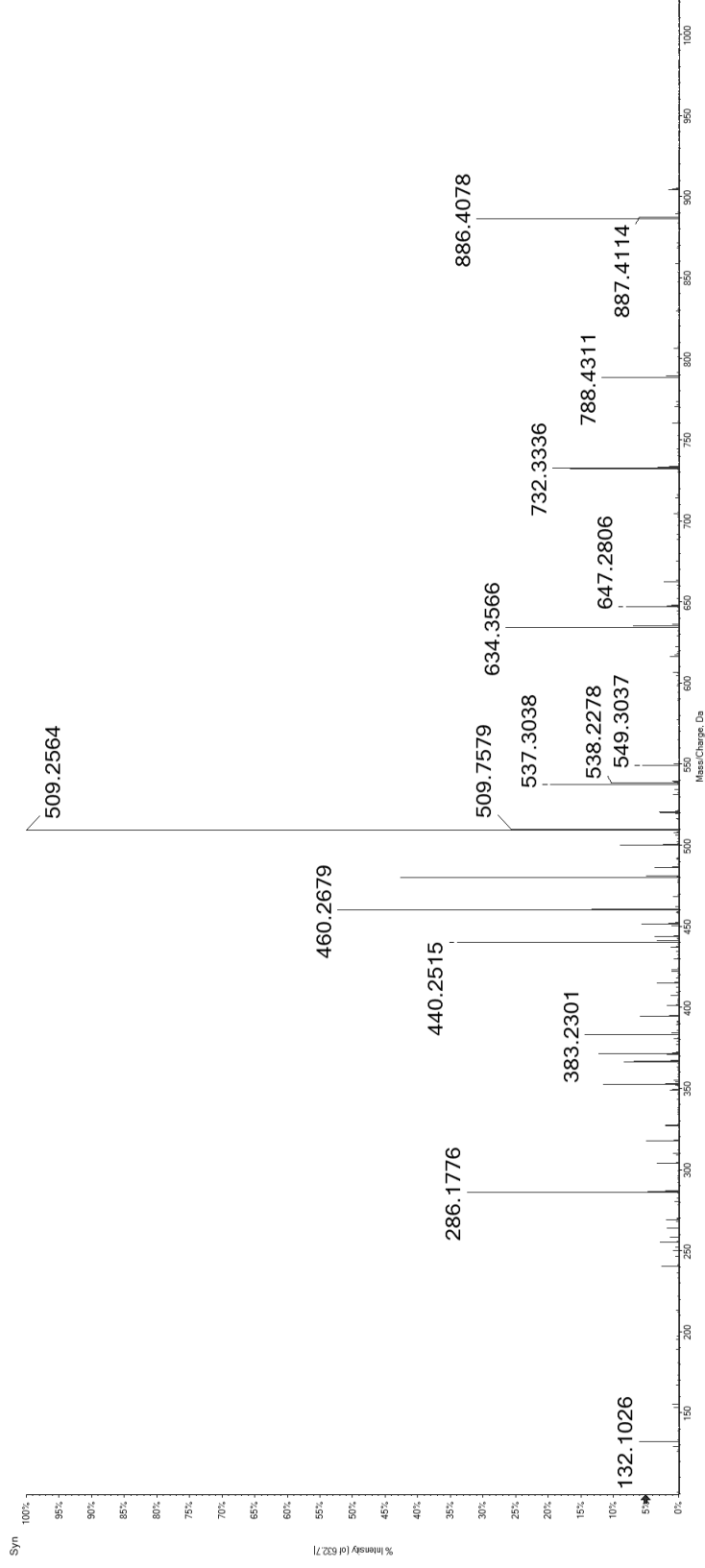
Candidate Spectrum



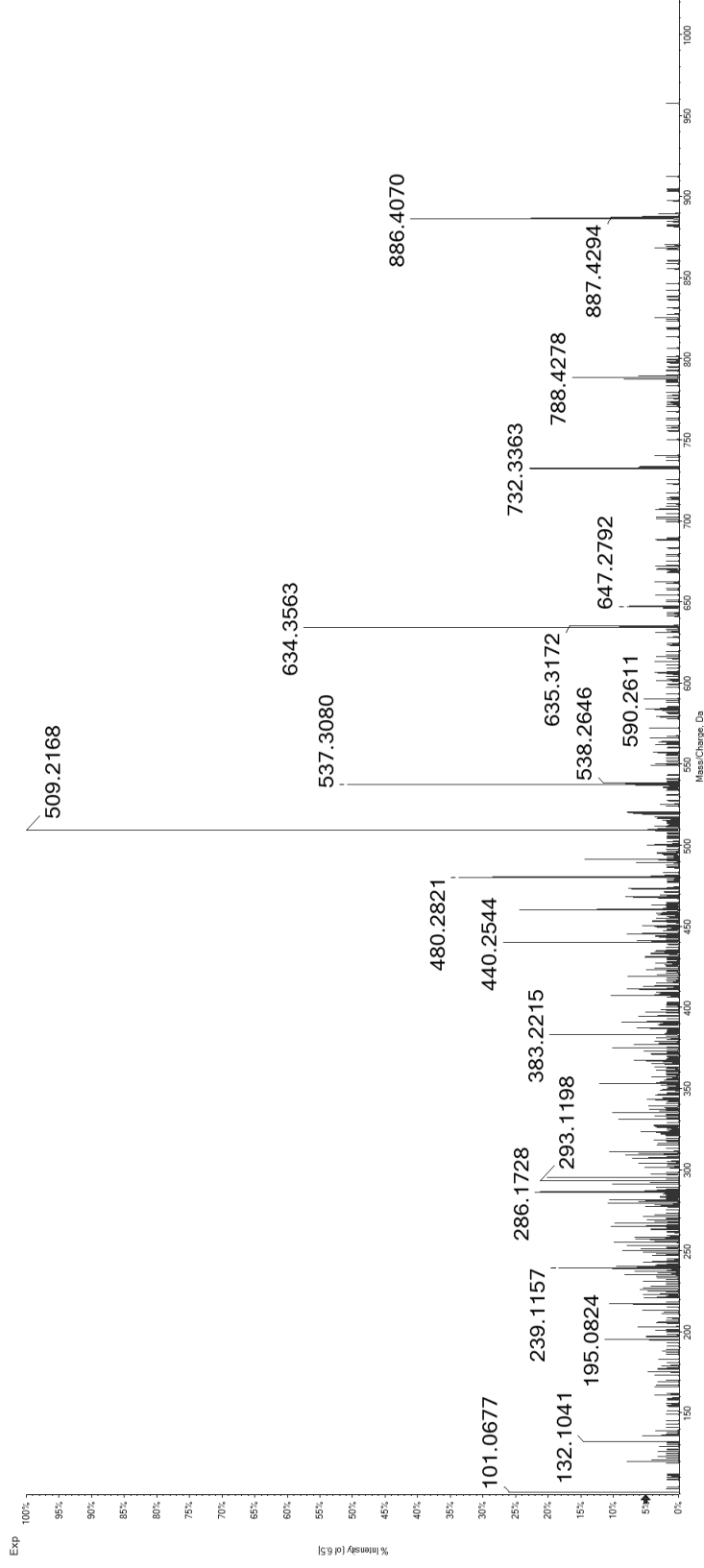
47

KEL(pS)PPPGL

Synthetic Peptide



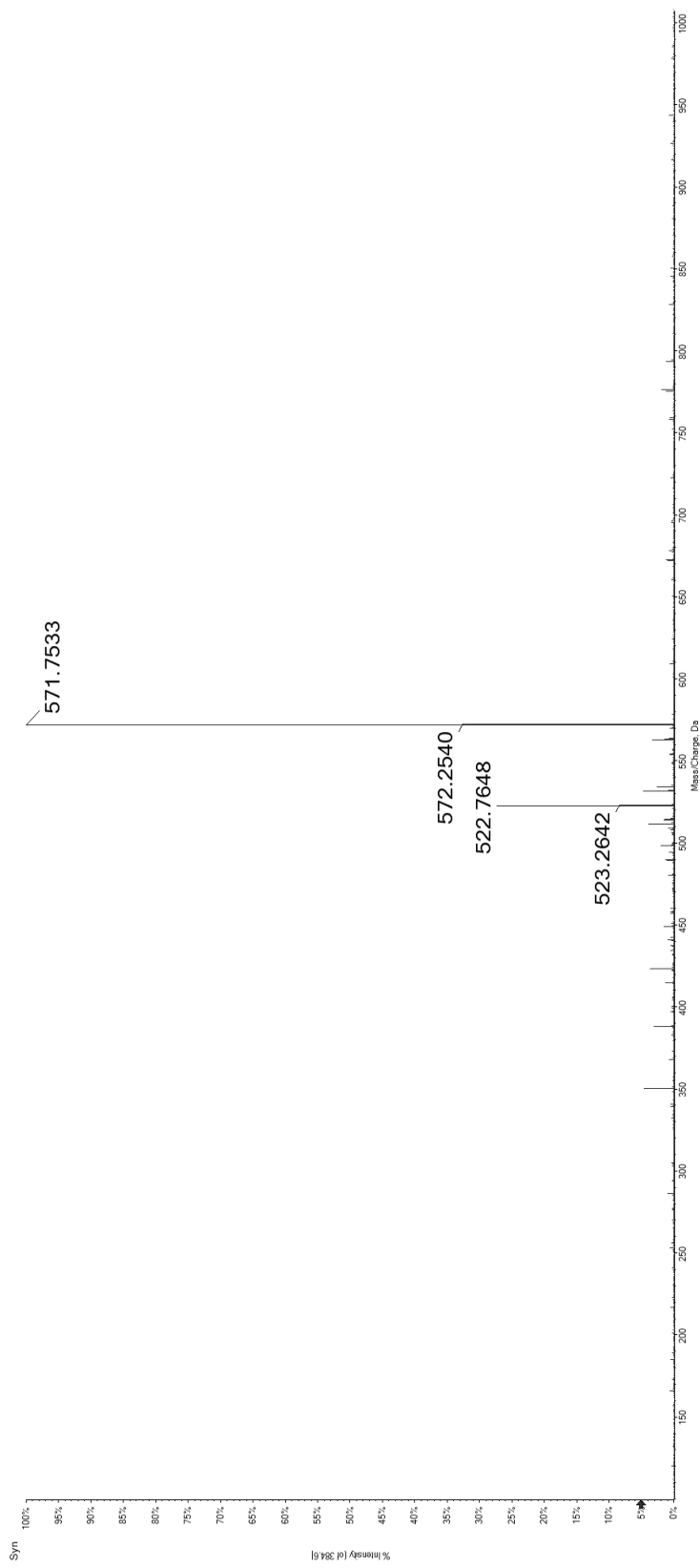
Candidate Spectrum



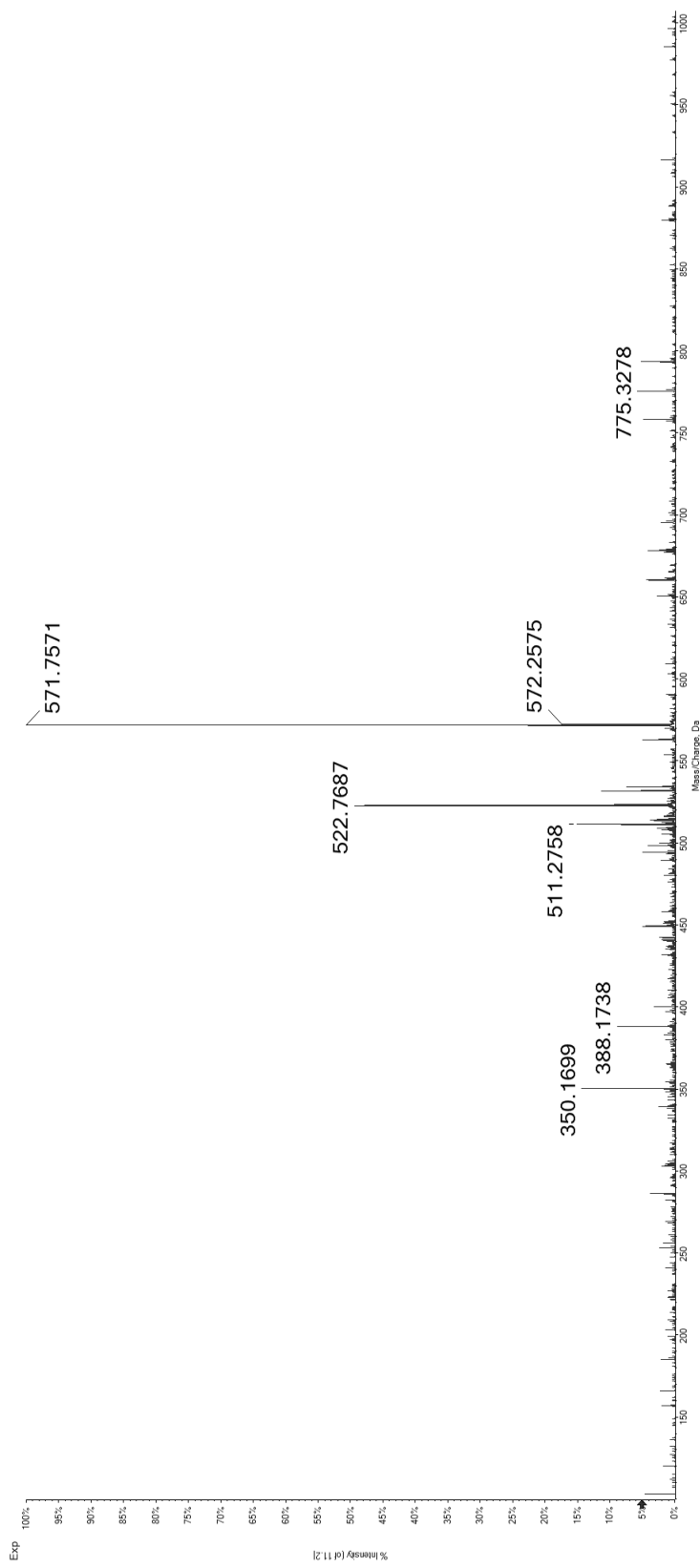
48

RER(pS)PSPSF

Synthetic Peptide



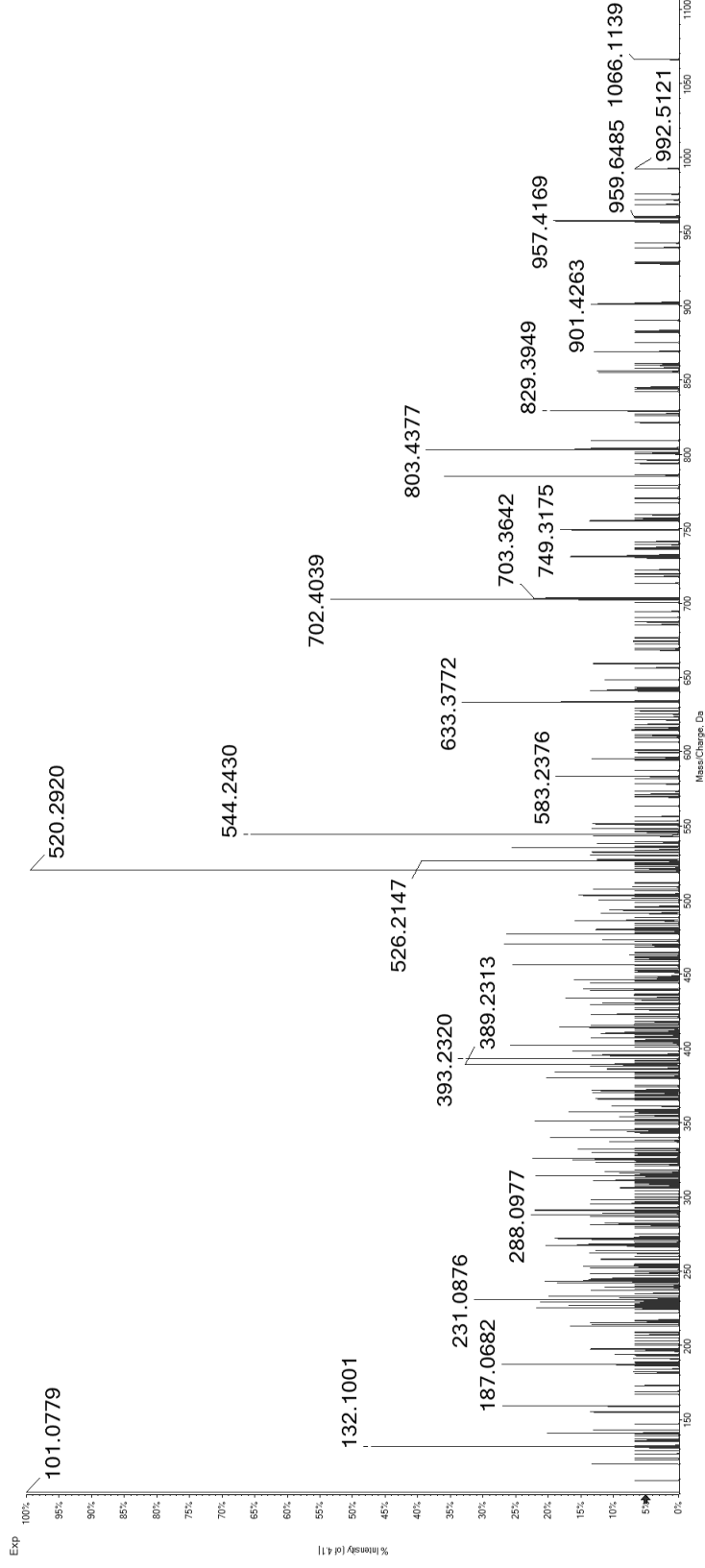
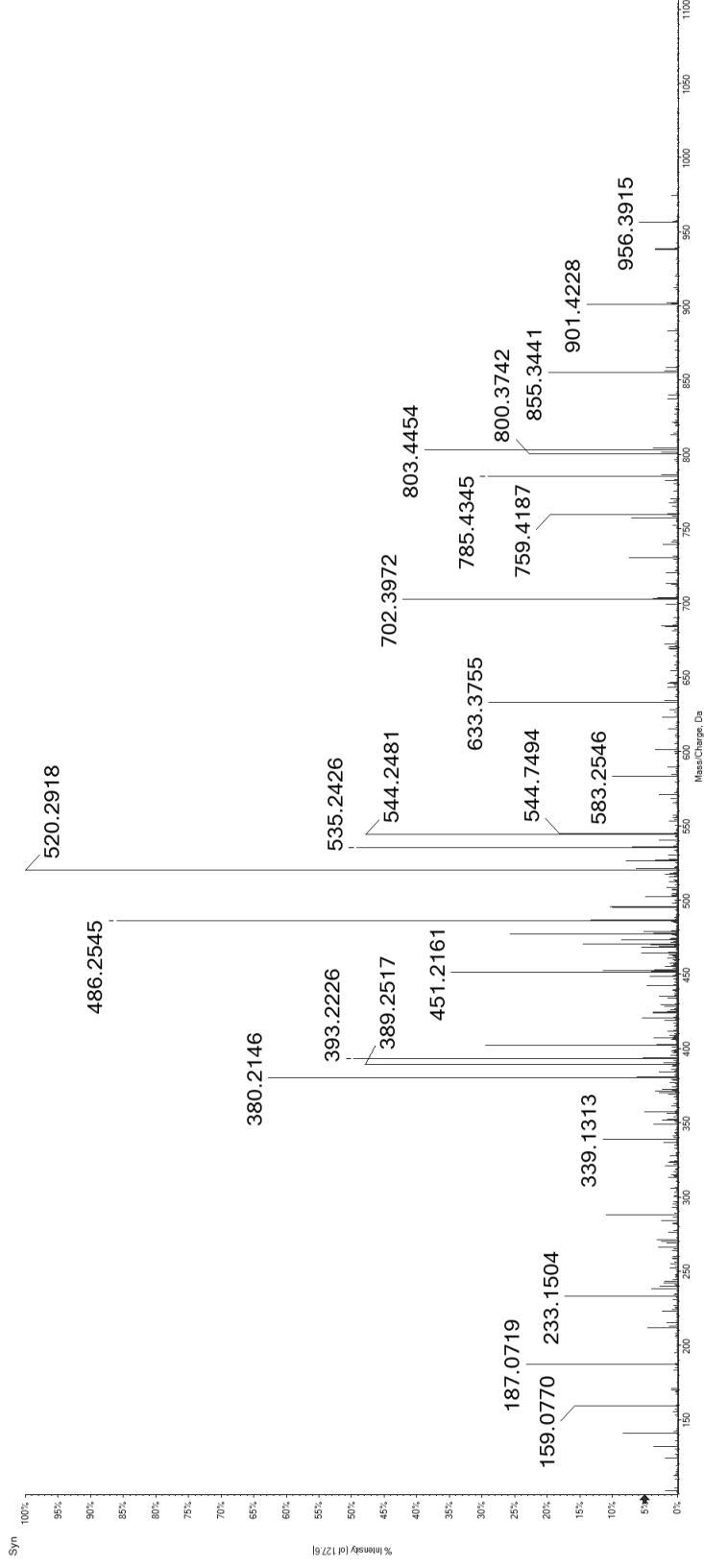
Candidate Spectrum



49

GET(pS)LMRTL

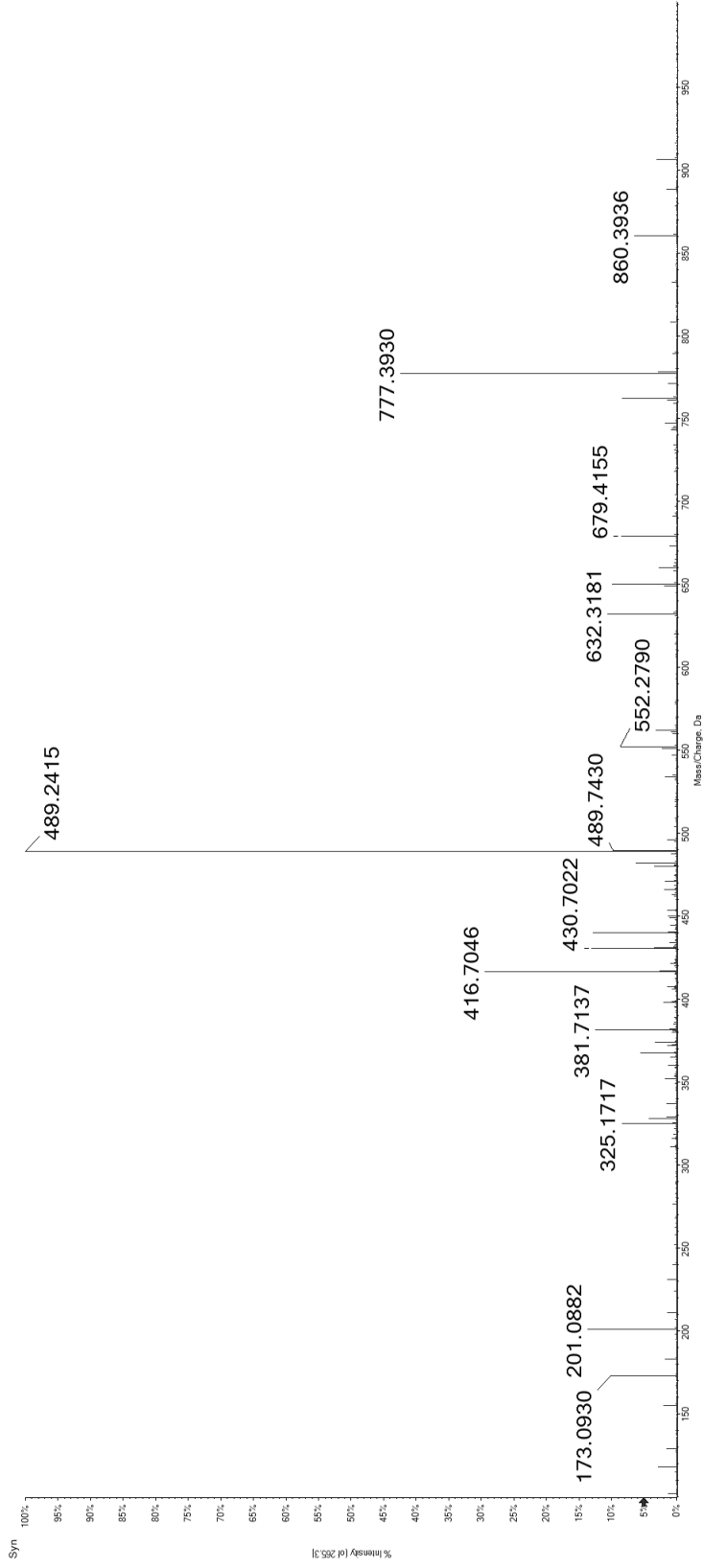
Synthetic Peptide



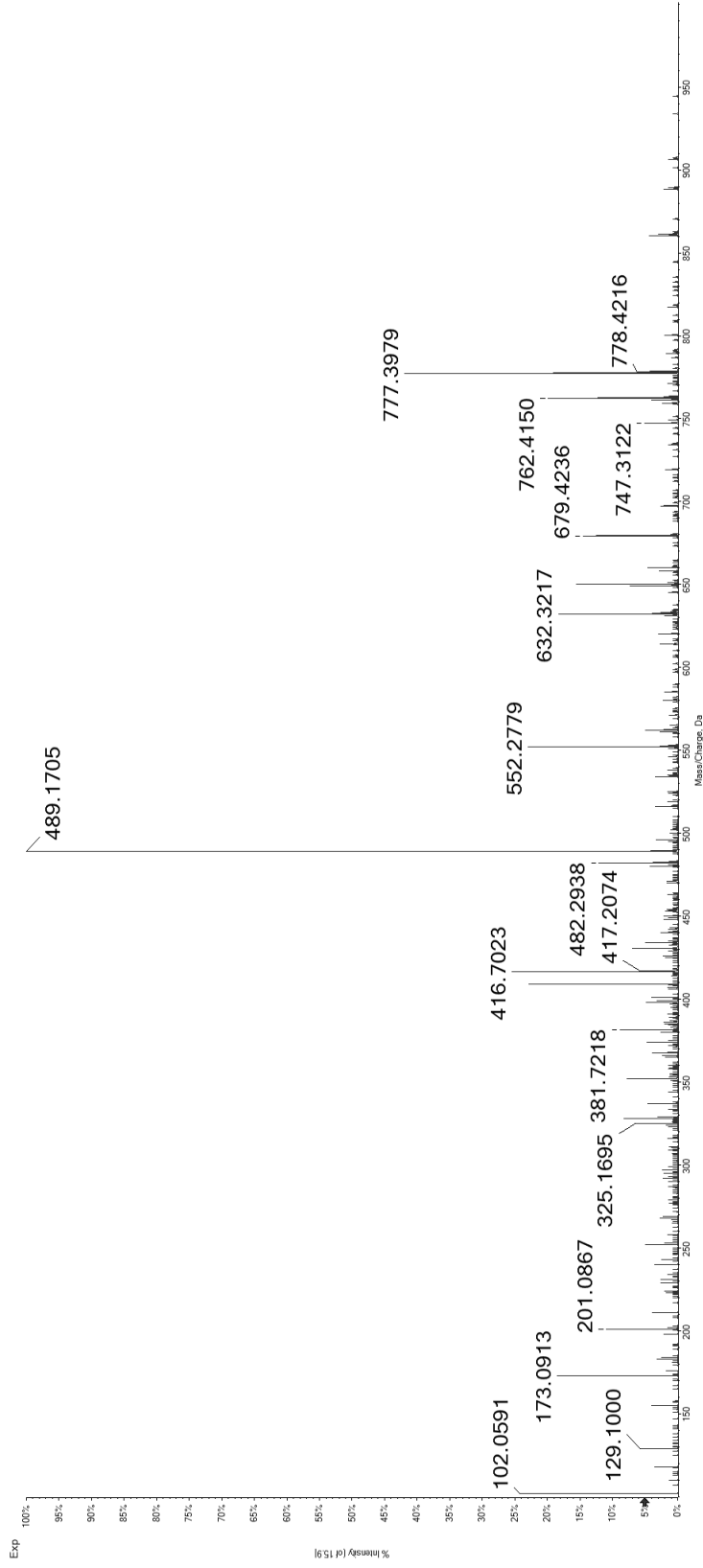
Candidate Spectrum

50 AEK(pS)PGPIV

Synthetic
Peptide



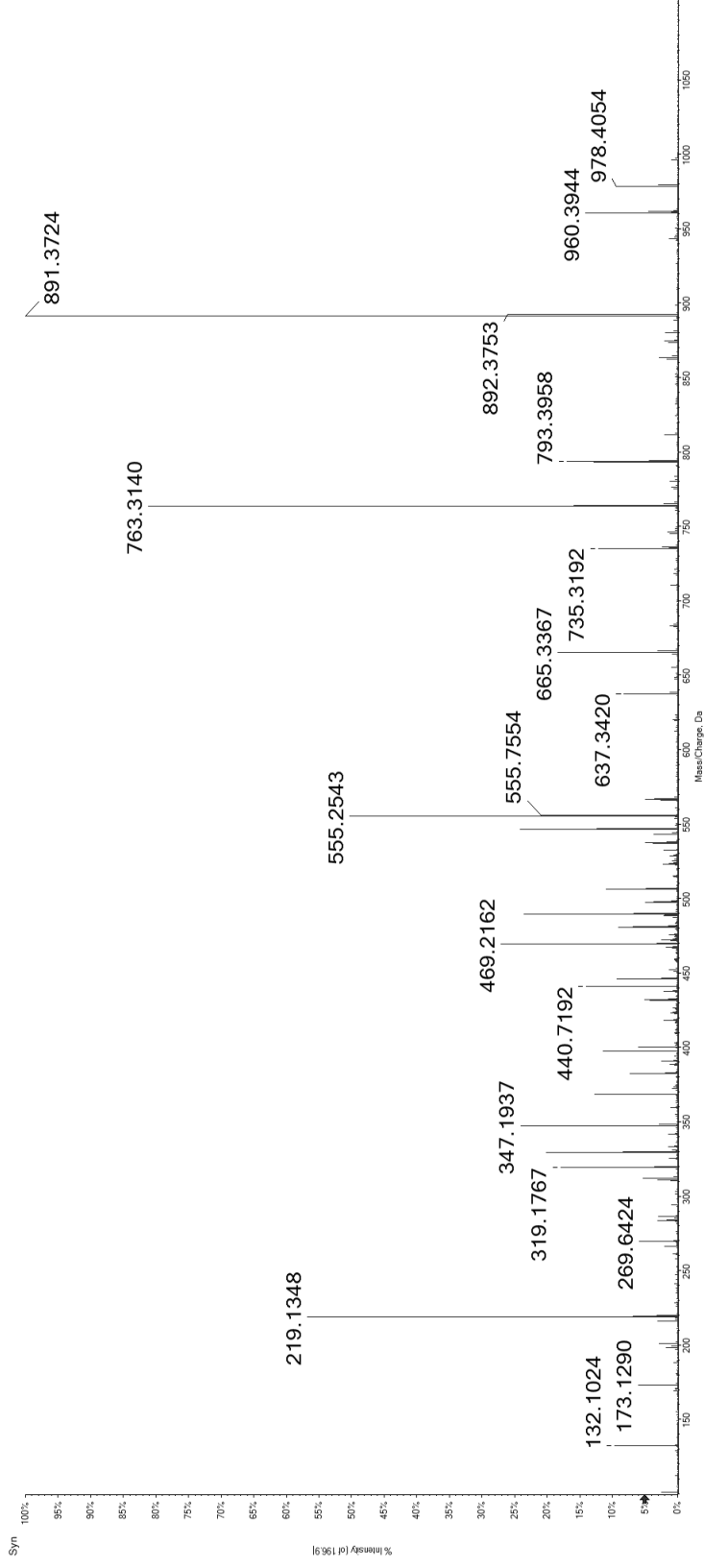
Candidate
Spectrum



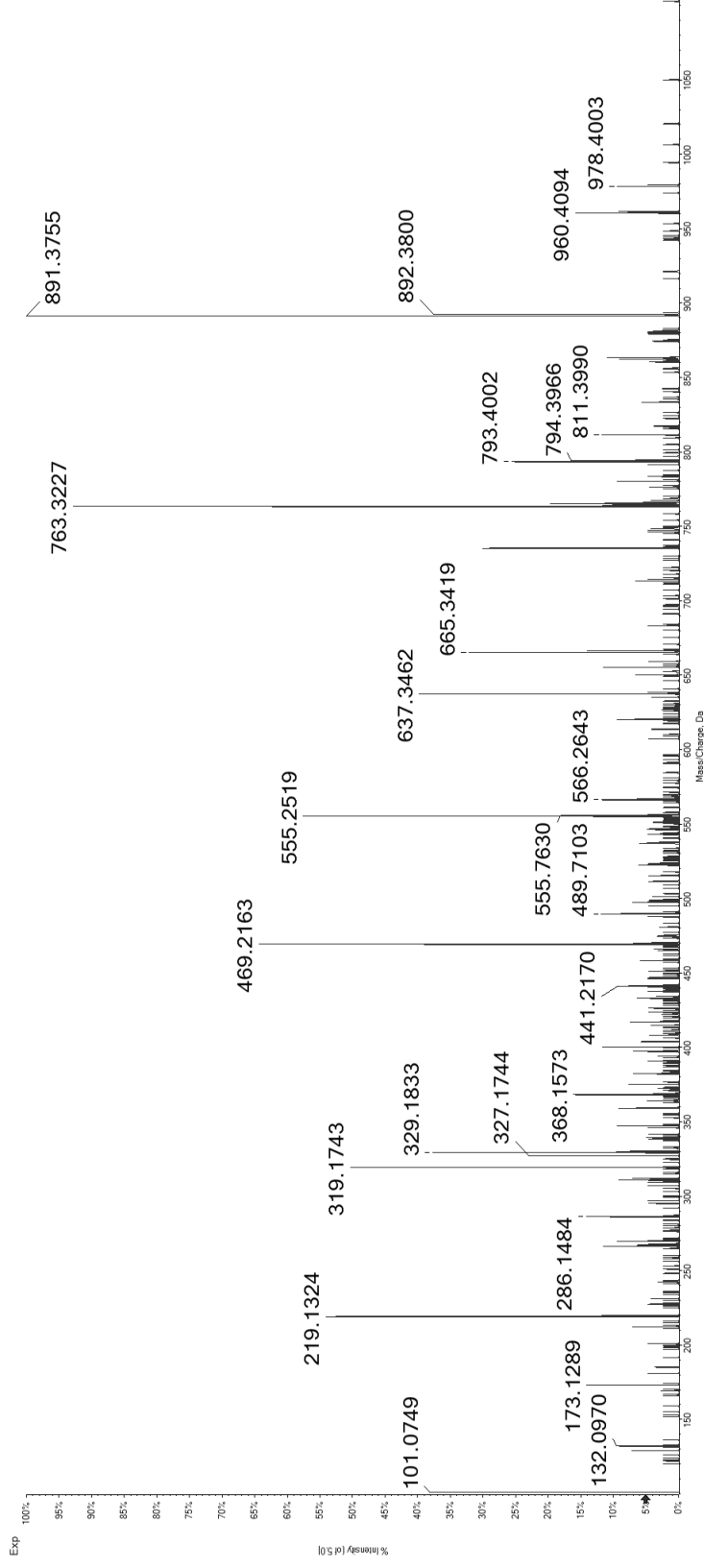
51

REN(pS)PVOQSI

Synthetic Peptide

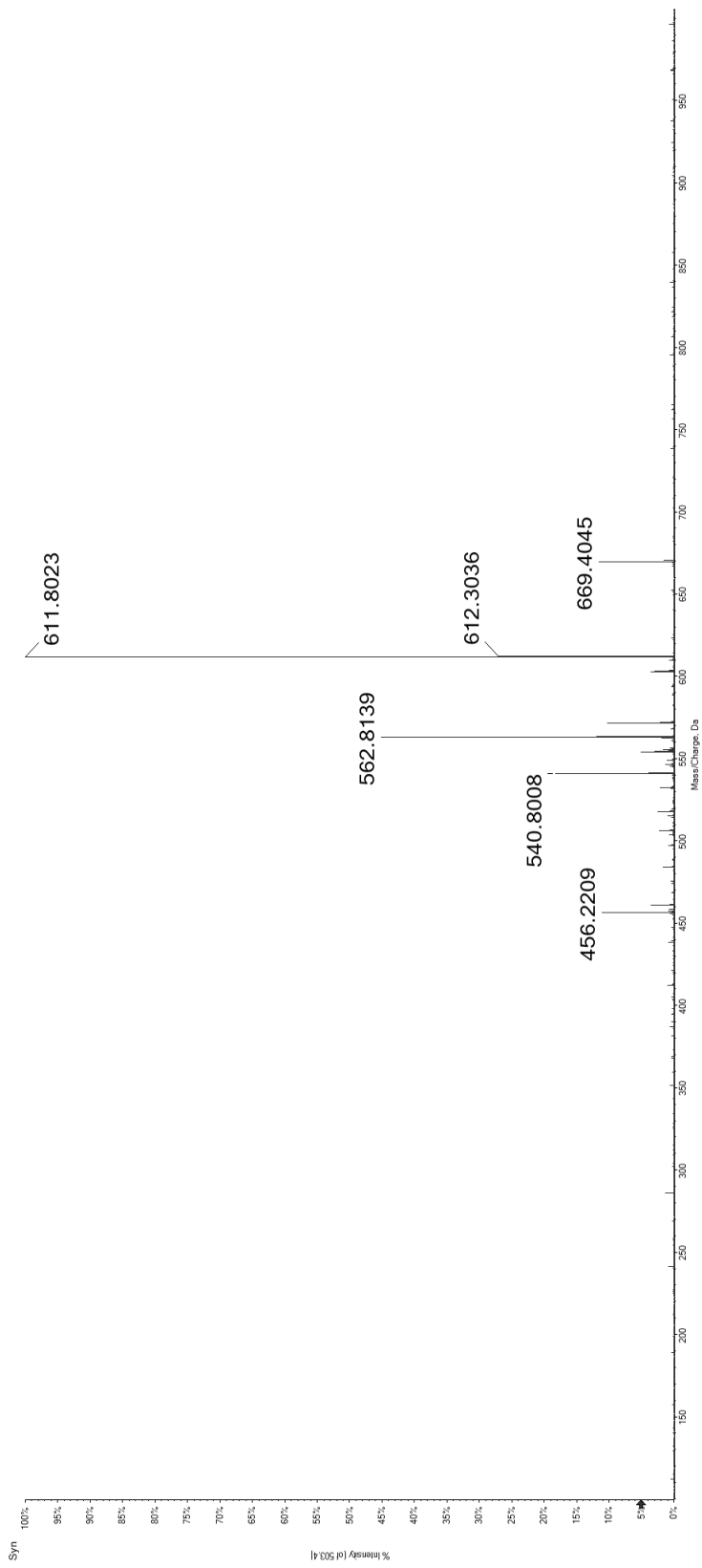


Candidate Spectrum

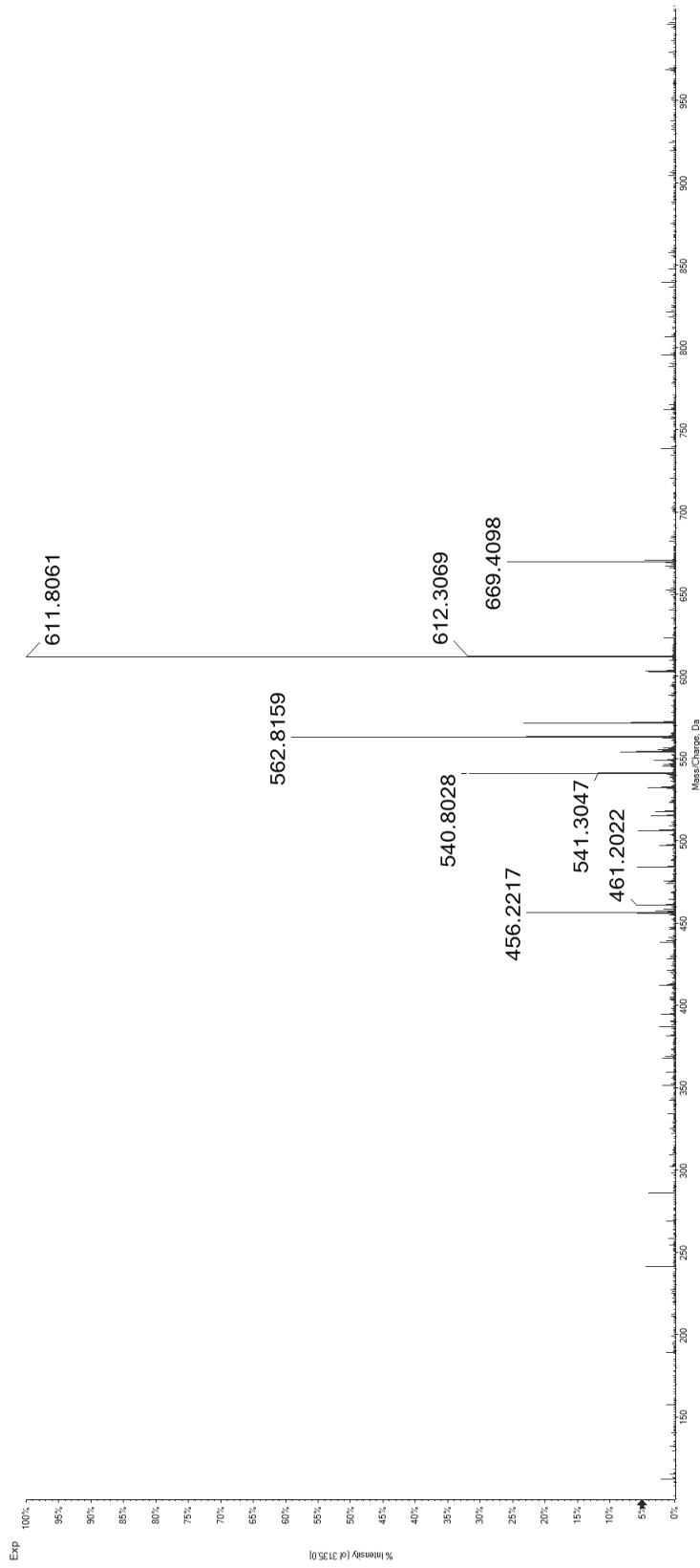


52 RET(pS)PNRIGL

Synthetic
Peptide



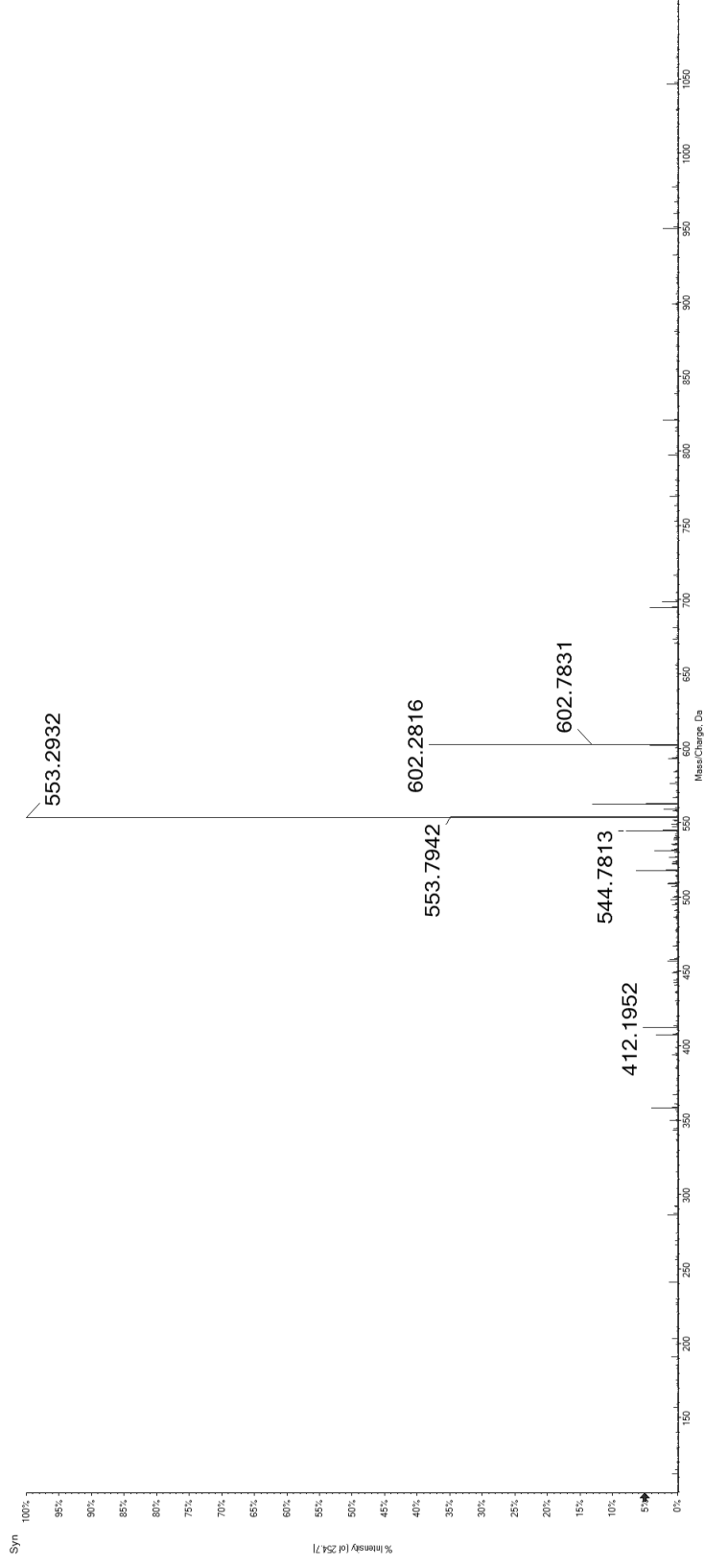
Candidate
Spectrum



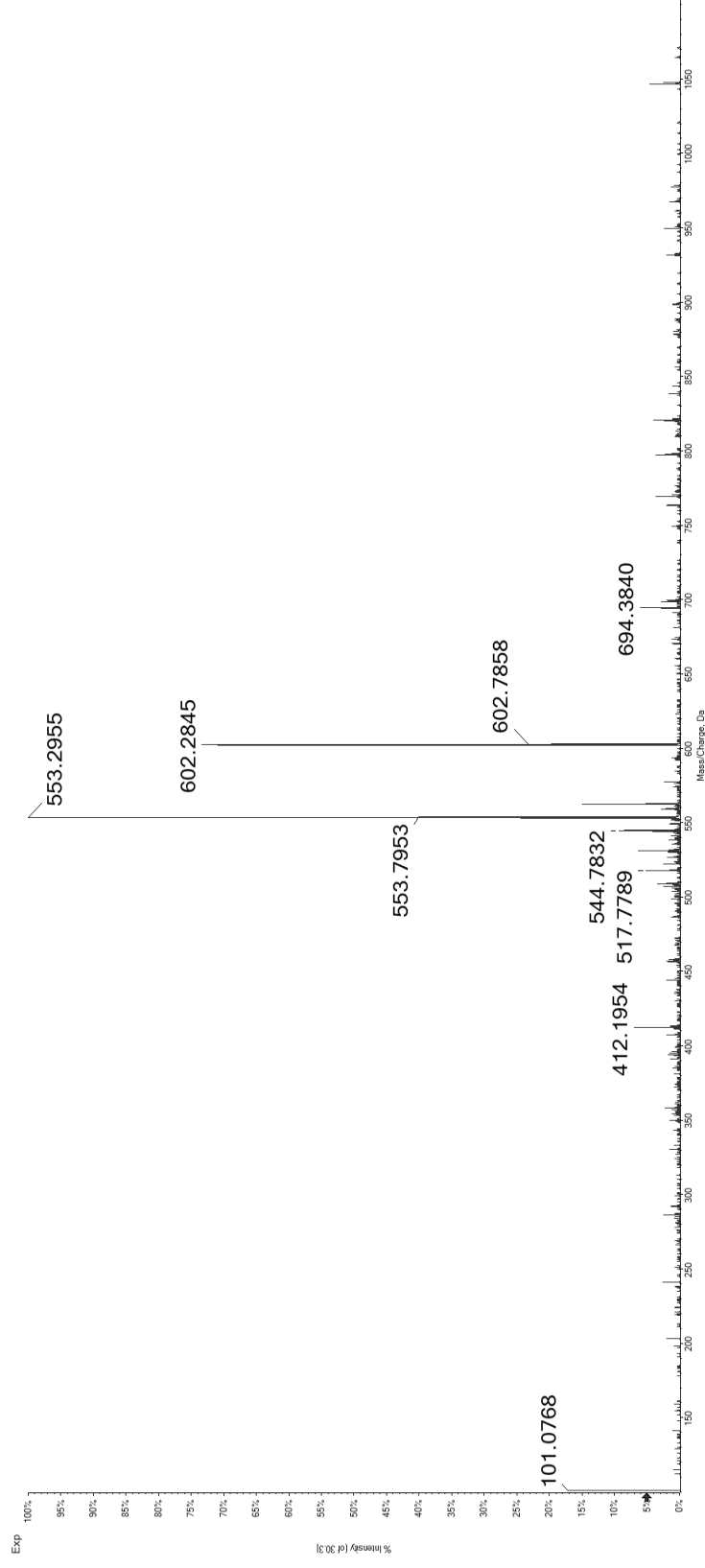
53

REG(pS)FRVTTA

Synthetic Peptide

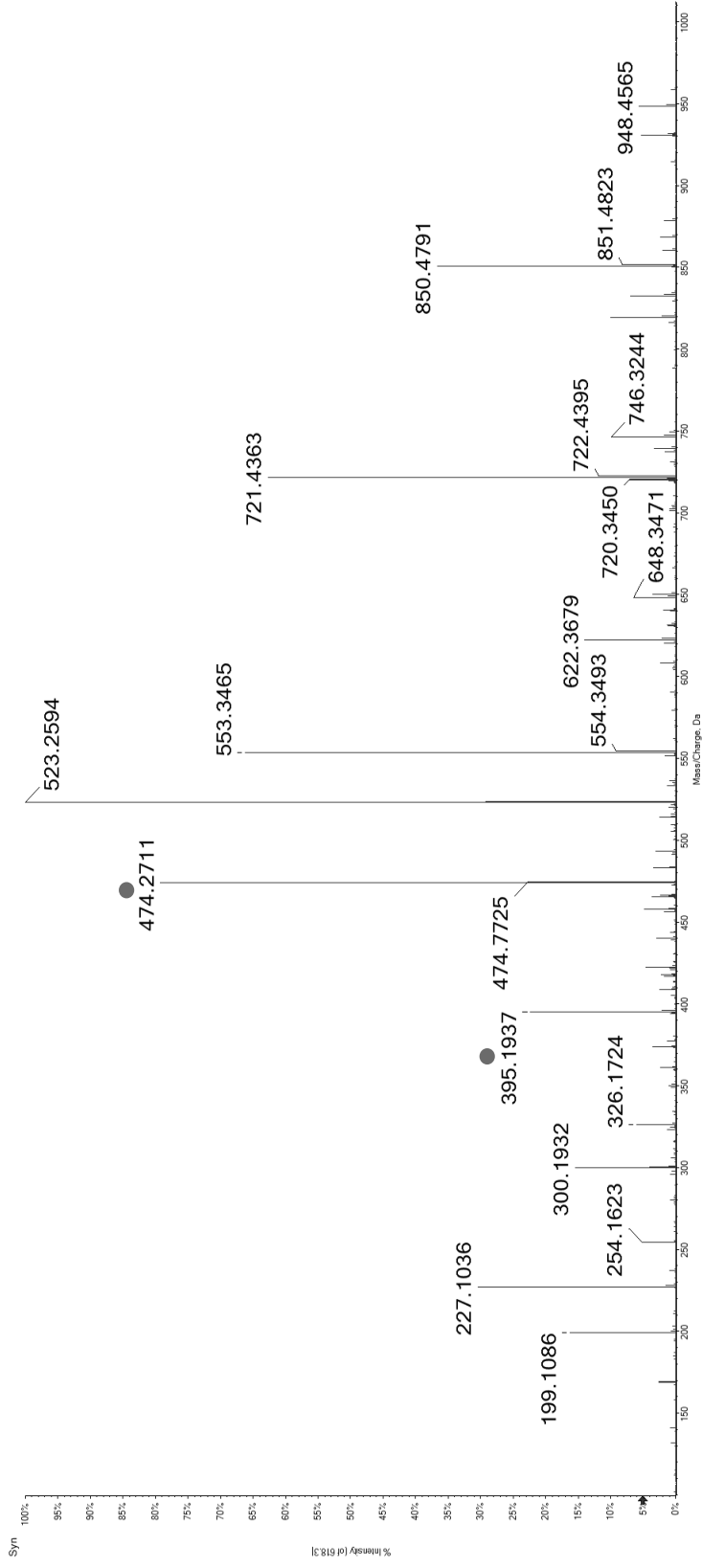


Candidate Spectrum

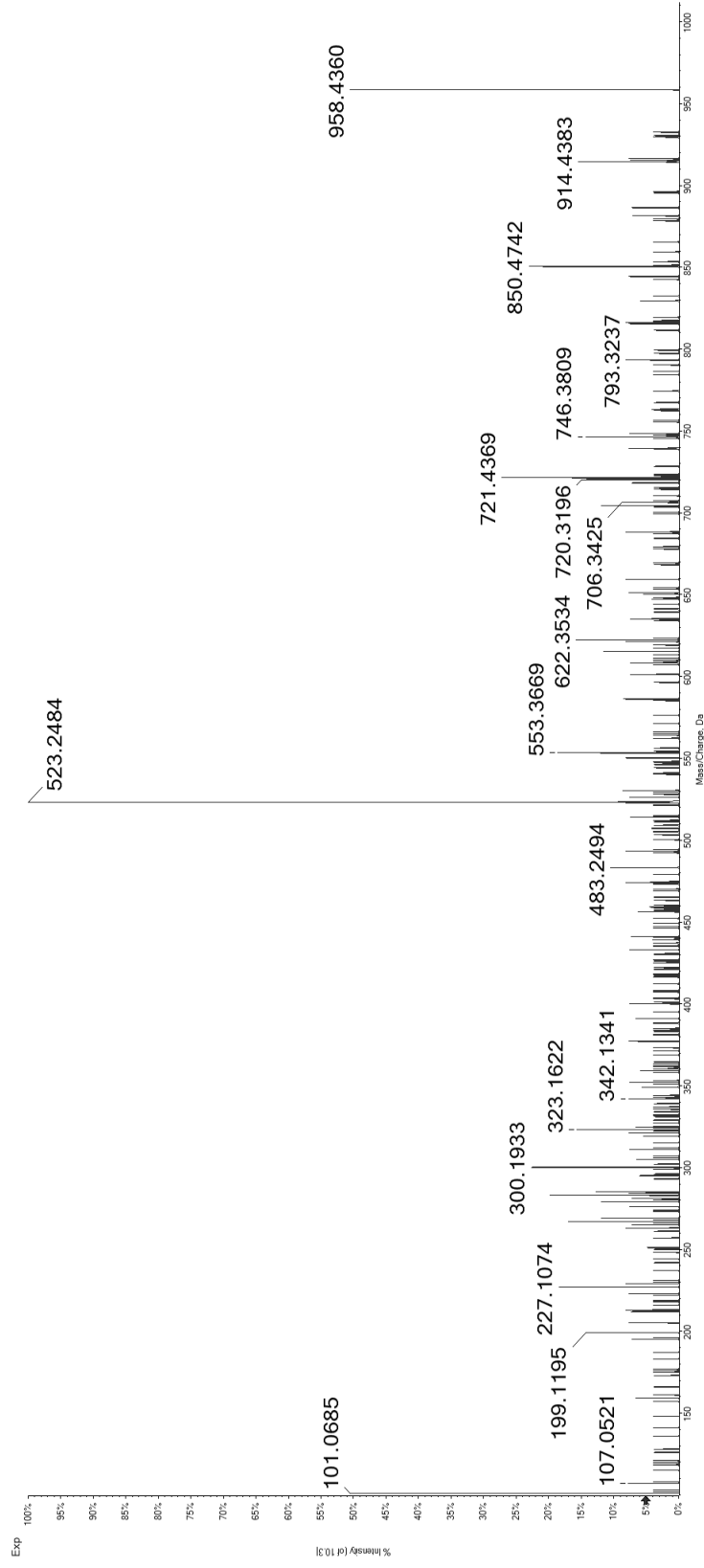


54 PEV(pS)PRPAL

Synthetic
Peptide
(Relative RT – 1,59)



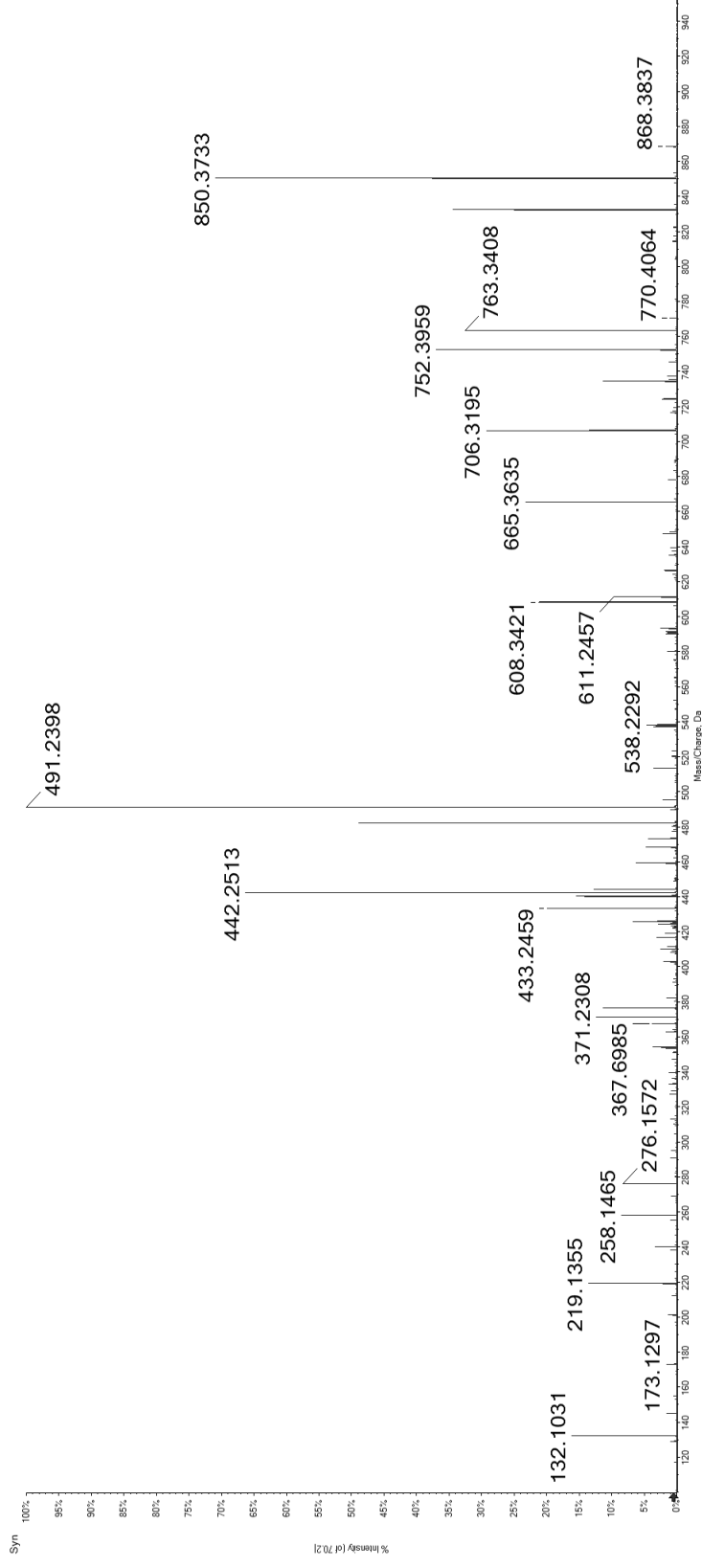
Candidate
Spectrum
(Relative RT – 1,57)



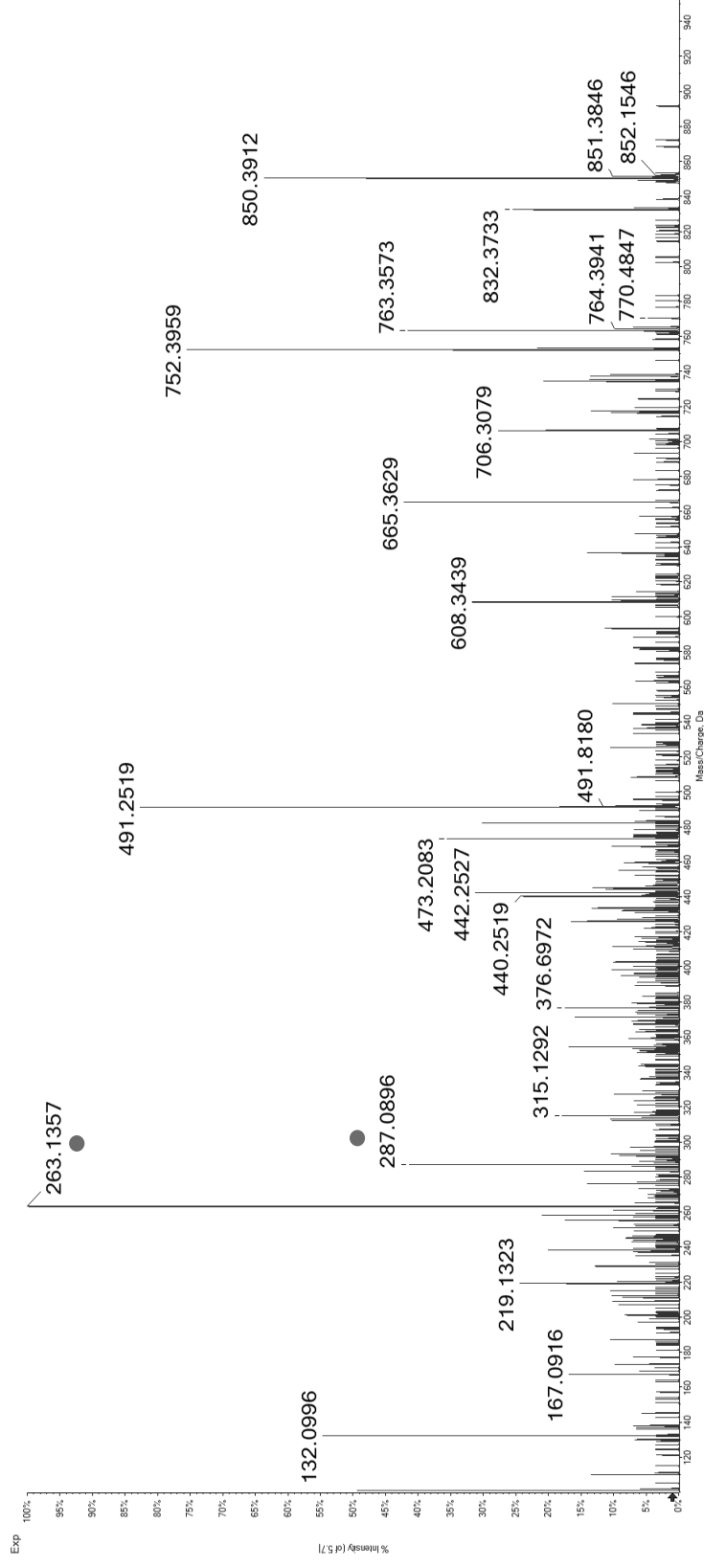
55

KEL(ps)PAGSI

Synthetic
Peptide
(Relative RT – 1,44)

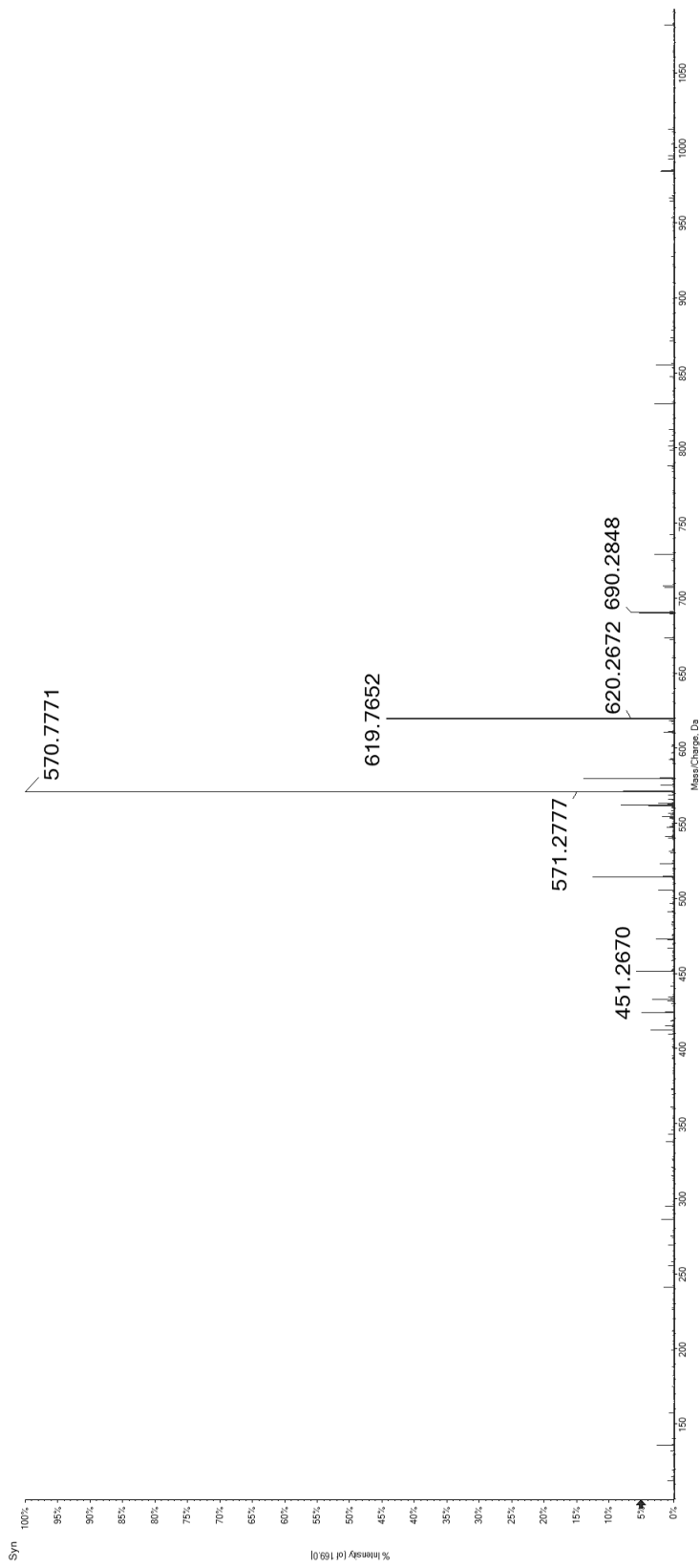


Candidate
Spectrum
(Relative RT – 1,44)

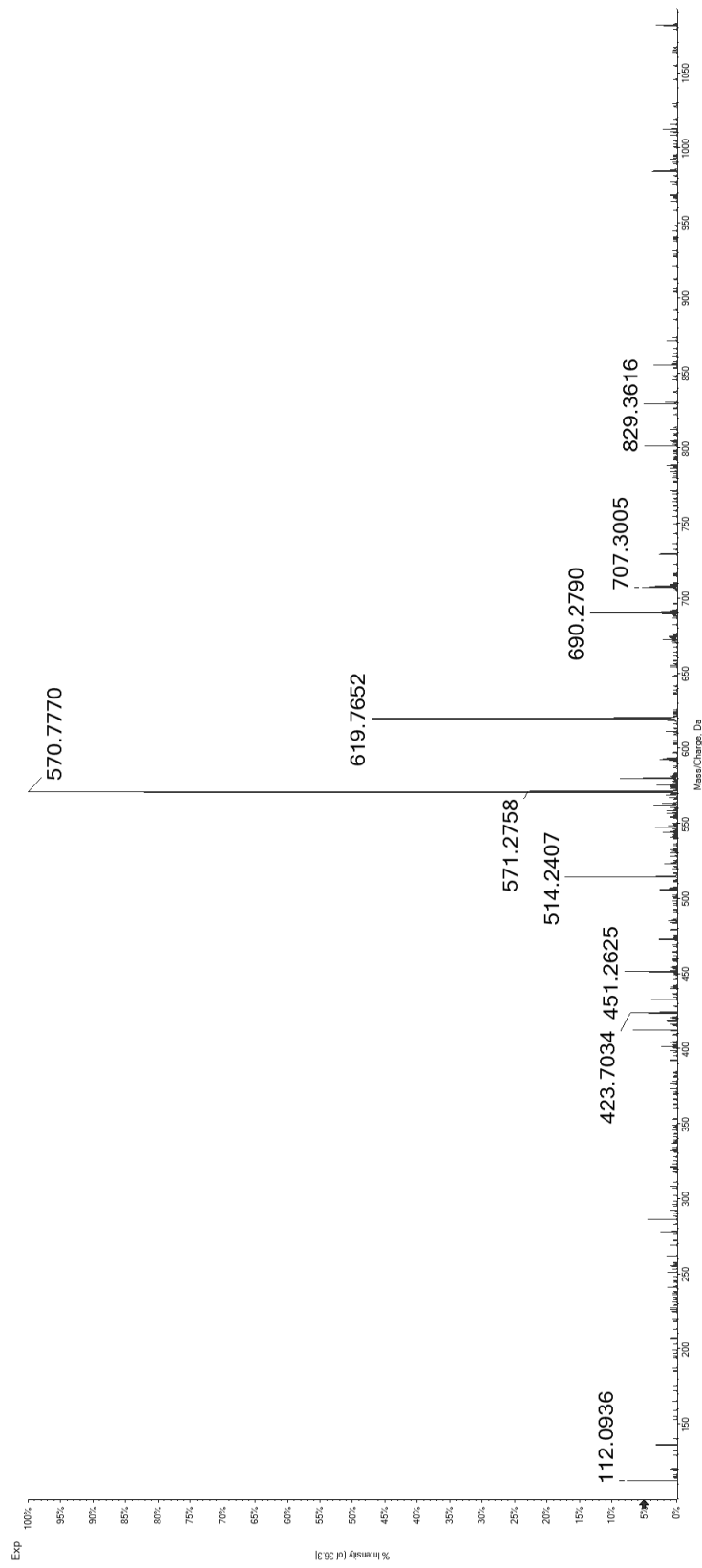


56 REG(pS)YDRYL

Synthetic
Peptide



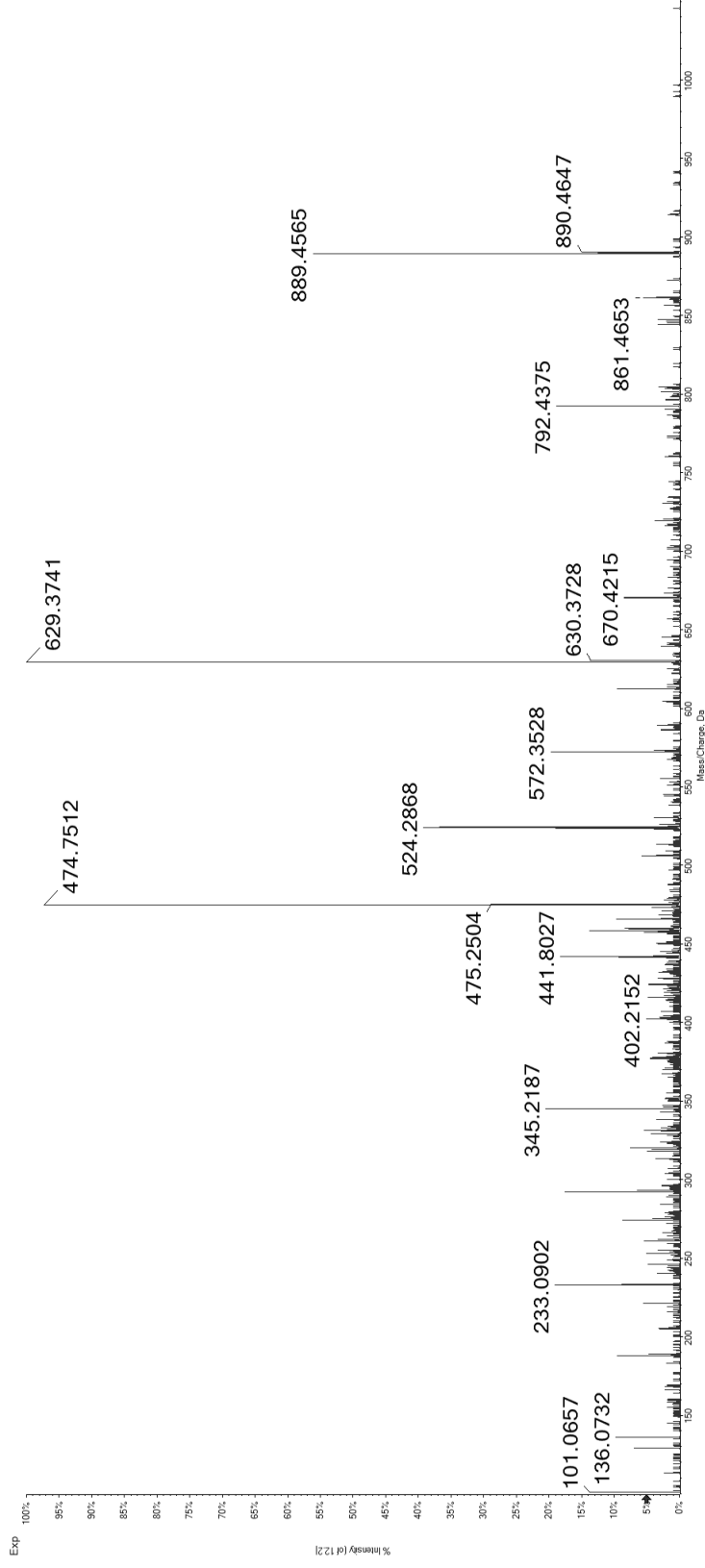
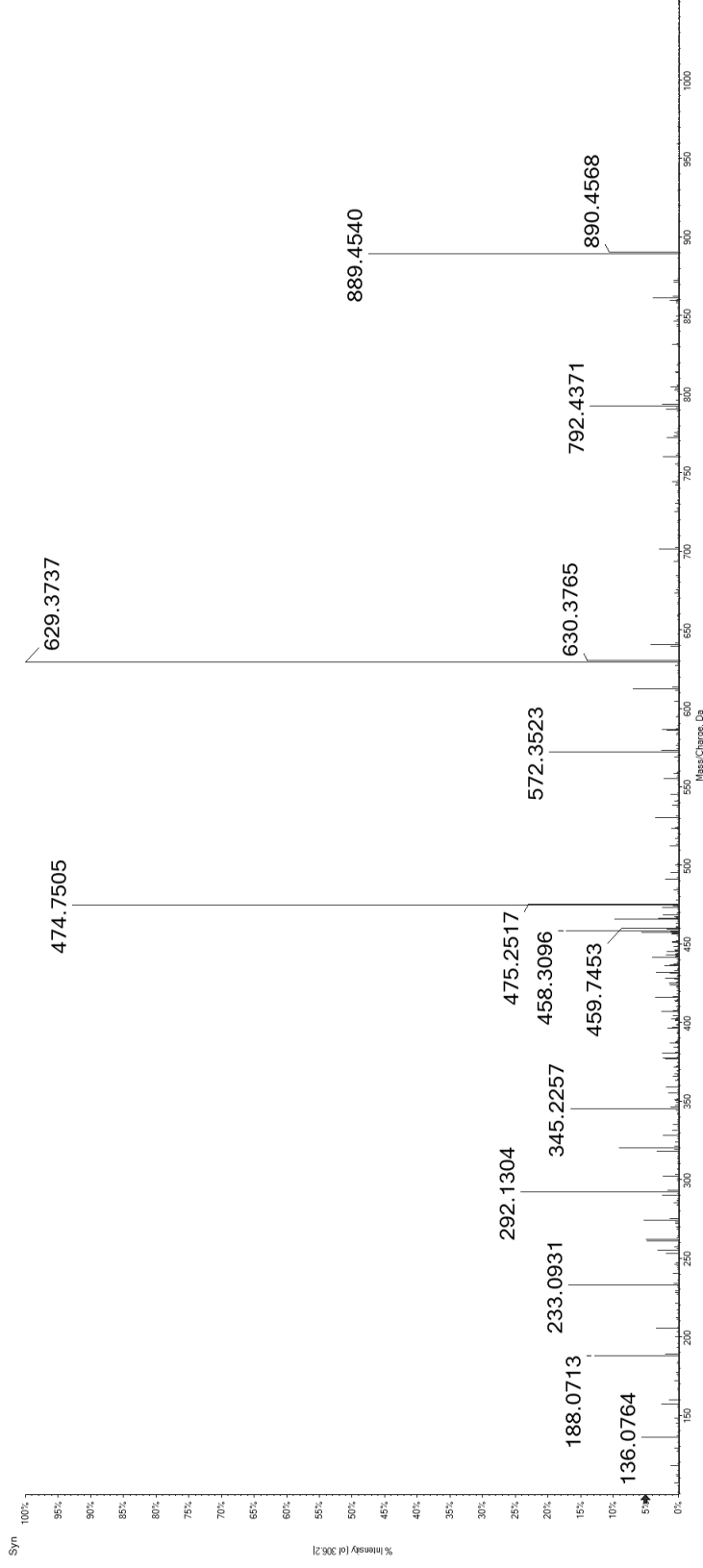
Candidate
Spectrum



57

S(pS)YGNIRAV

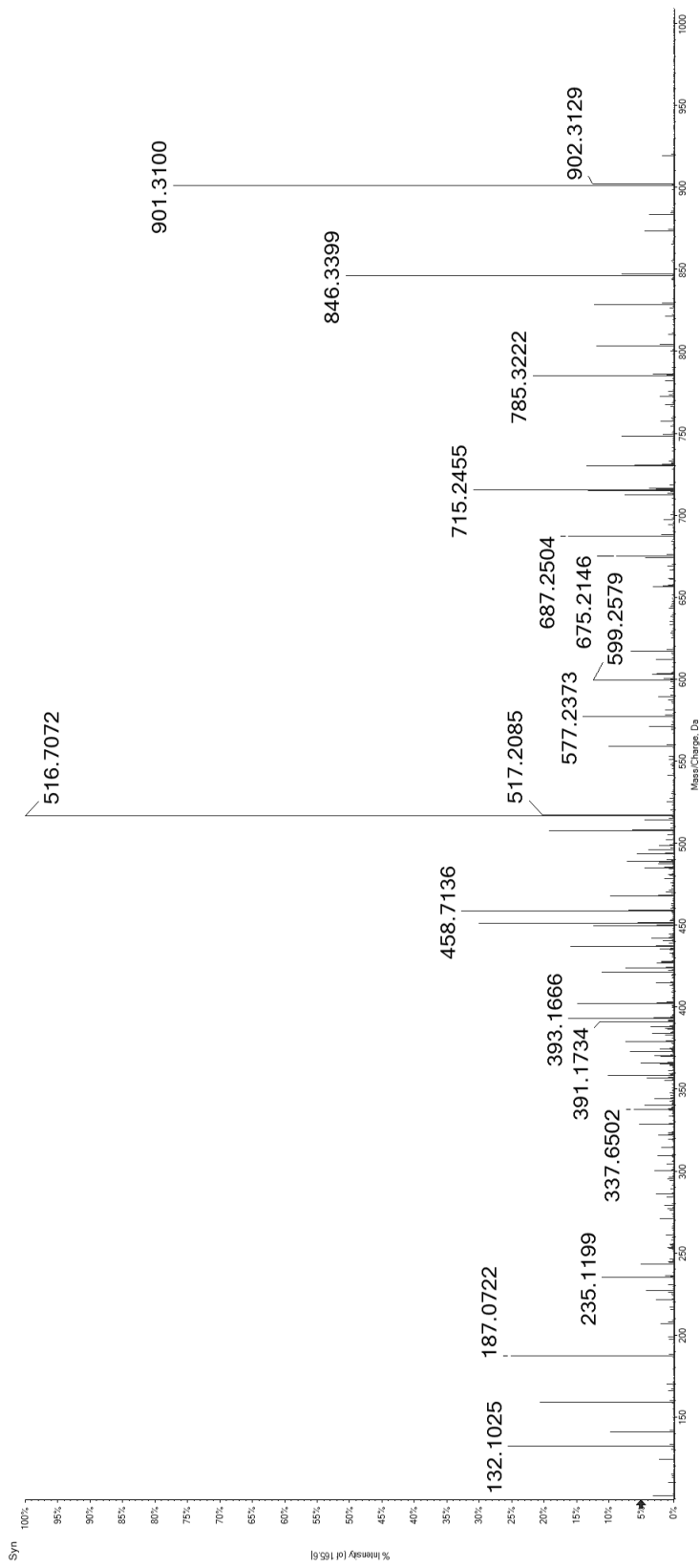
Synthetic
Peptide



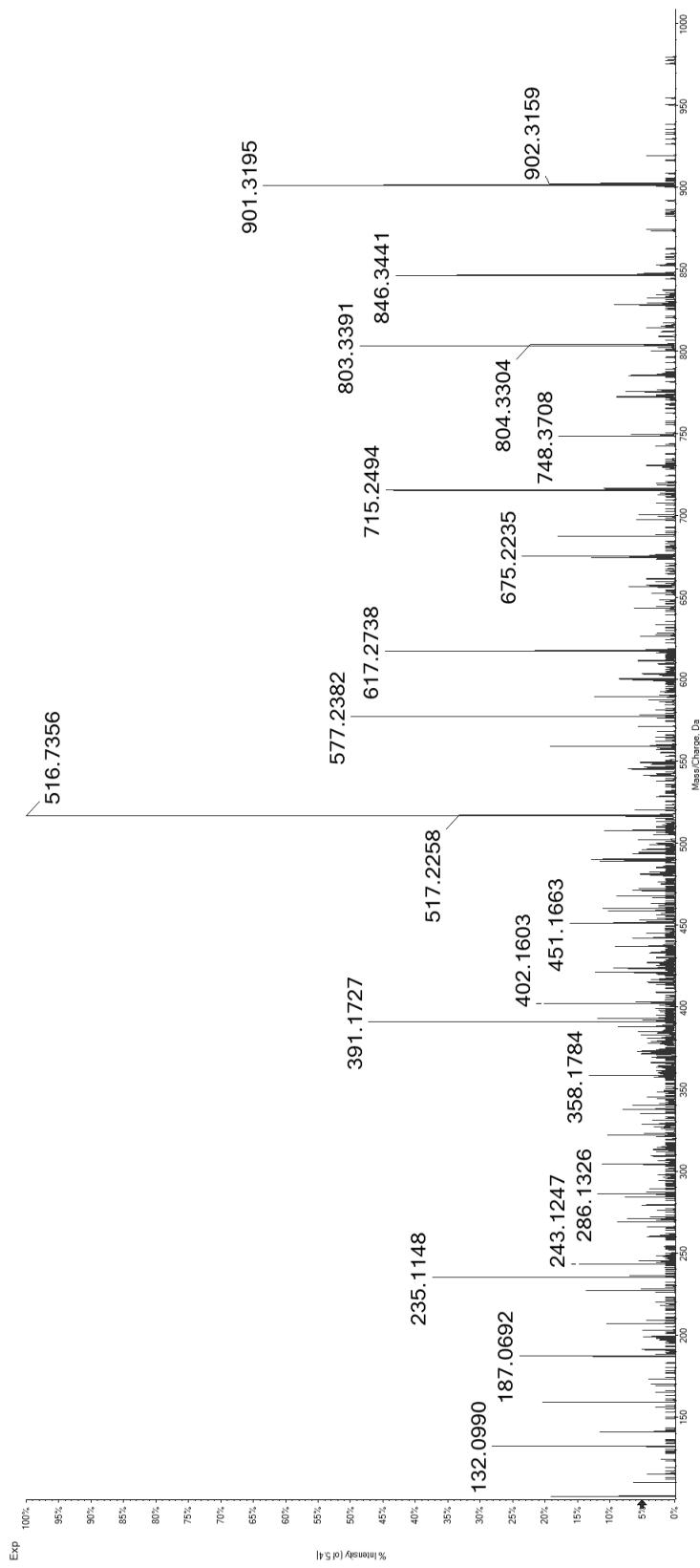
Candidate
Spectrum

58 GEPHS(pS)PEL

Synthetic
Peptide



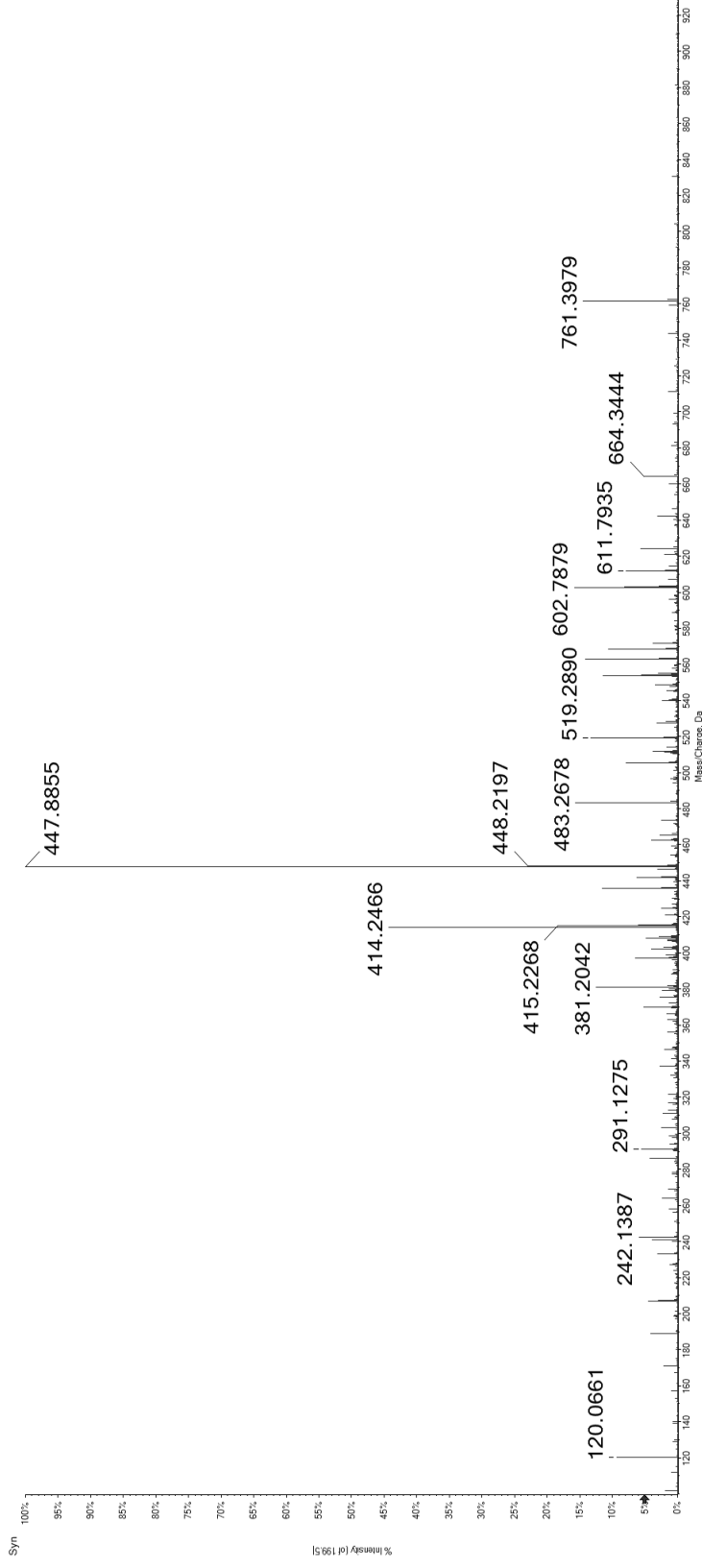
Candidate
Spectrum



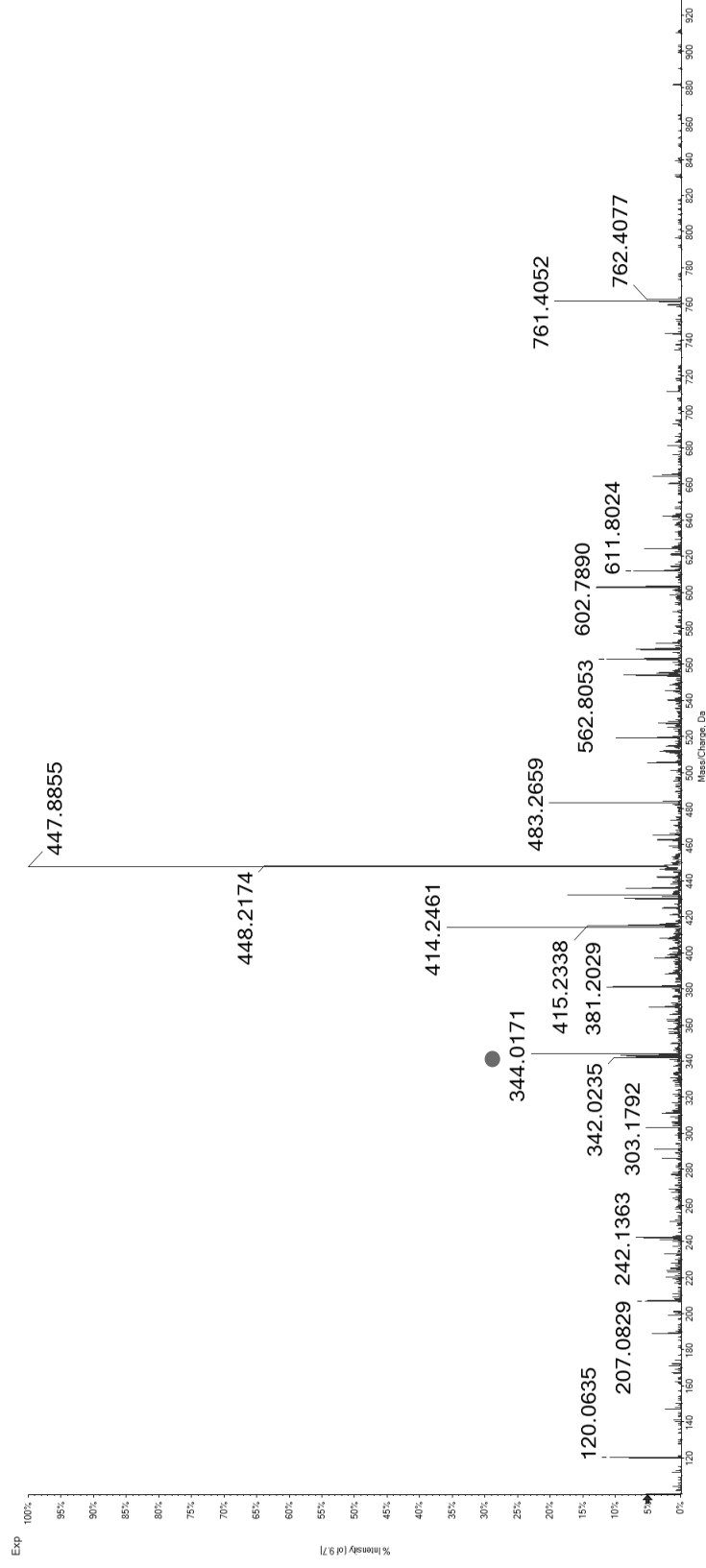
59

REK(pS)PGRMLST

Synthetic
Peptide
(Relative RT – 0,78)

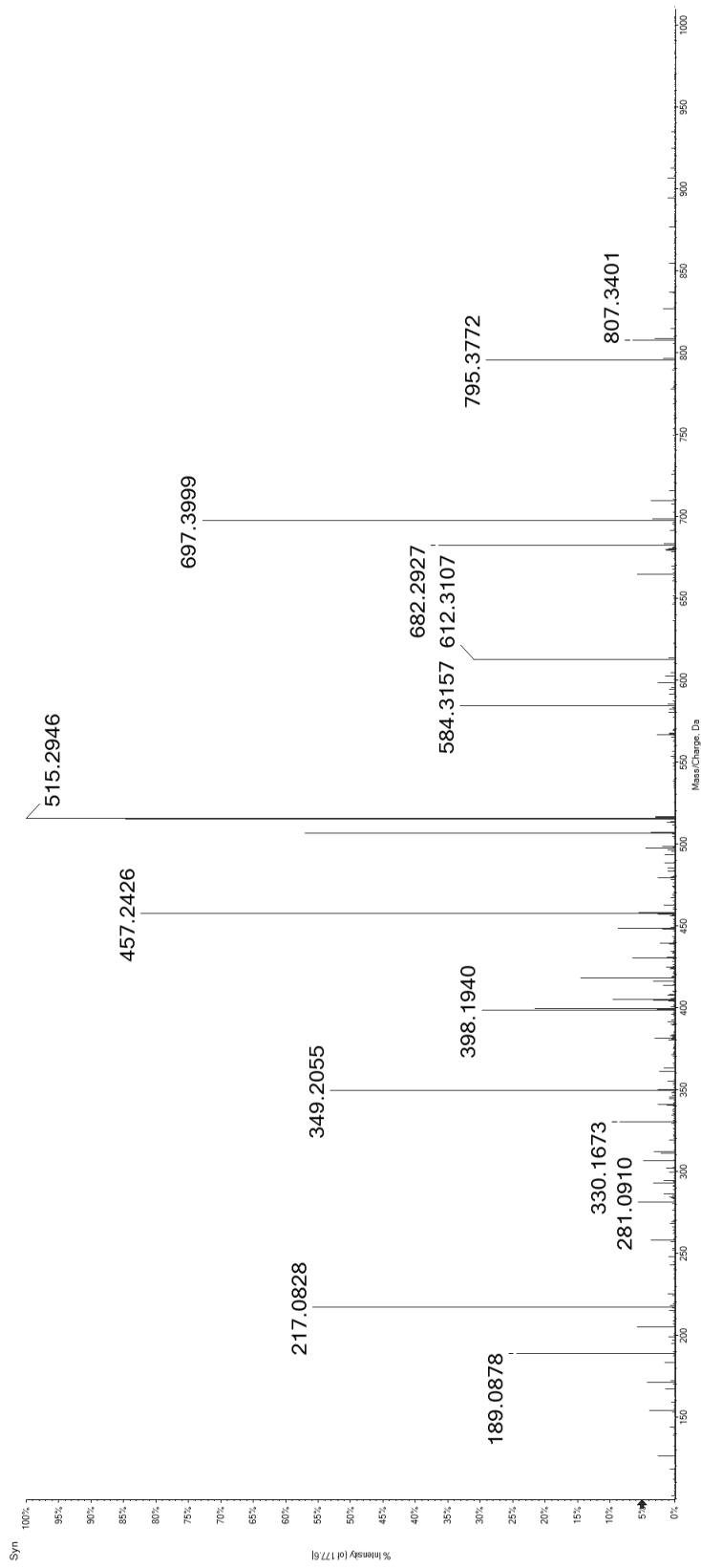


Candidate
Spectrum
(Relative RT – 0,82)

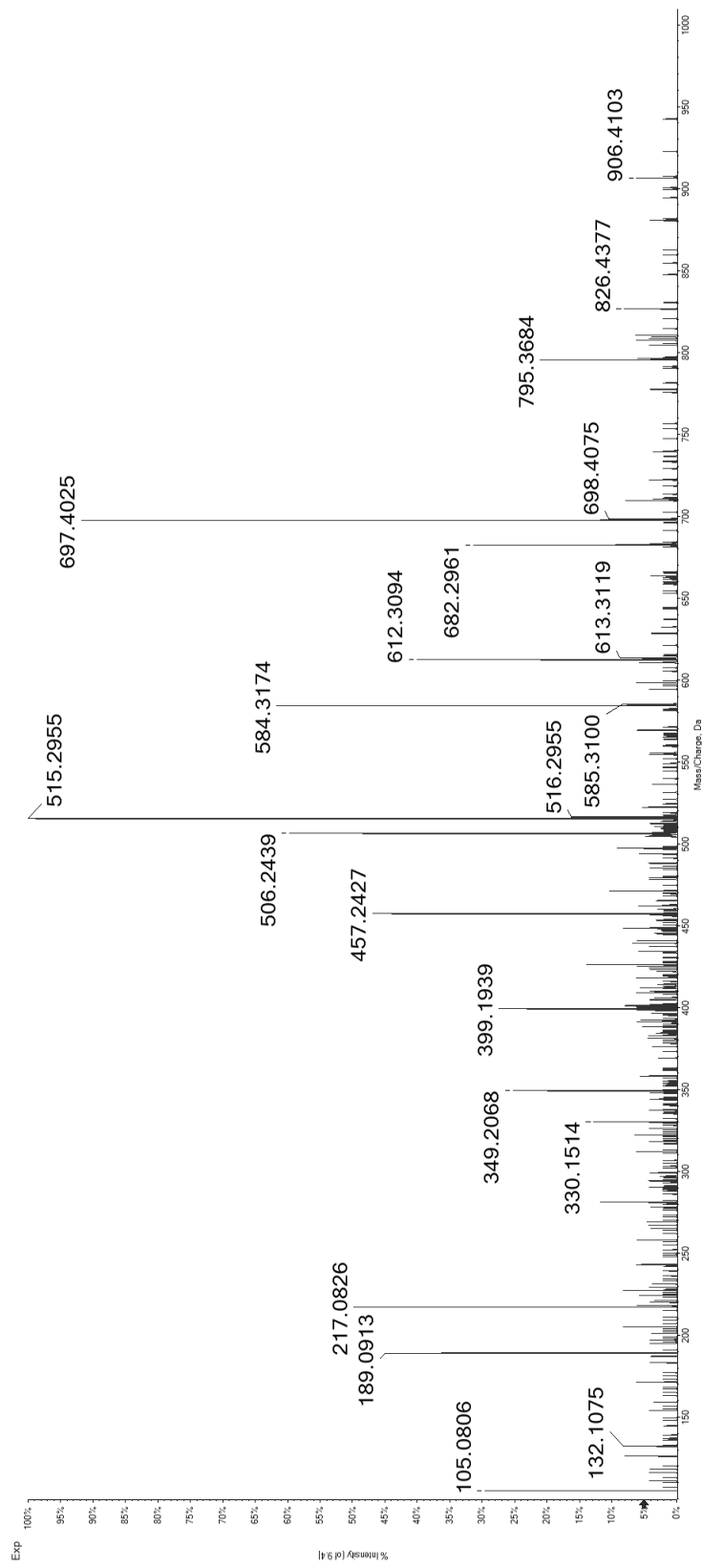


60 SEL(ps)PGRSV

Synthetic
Peptide



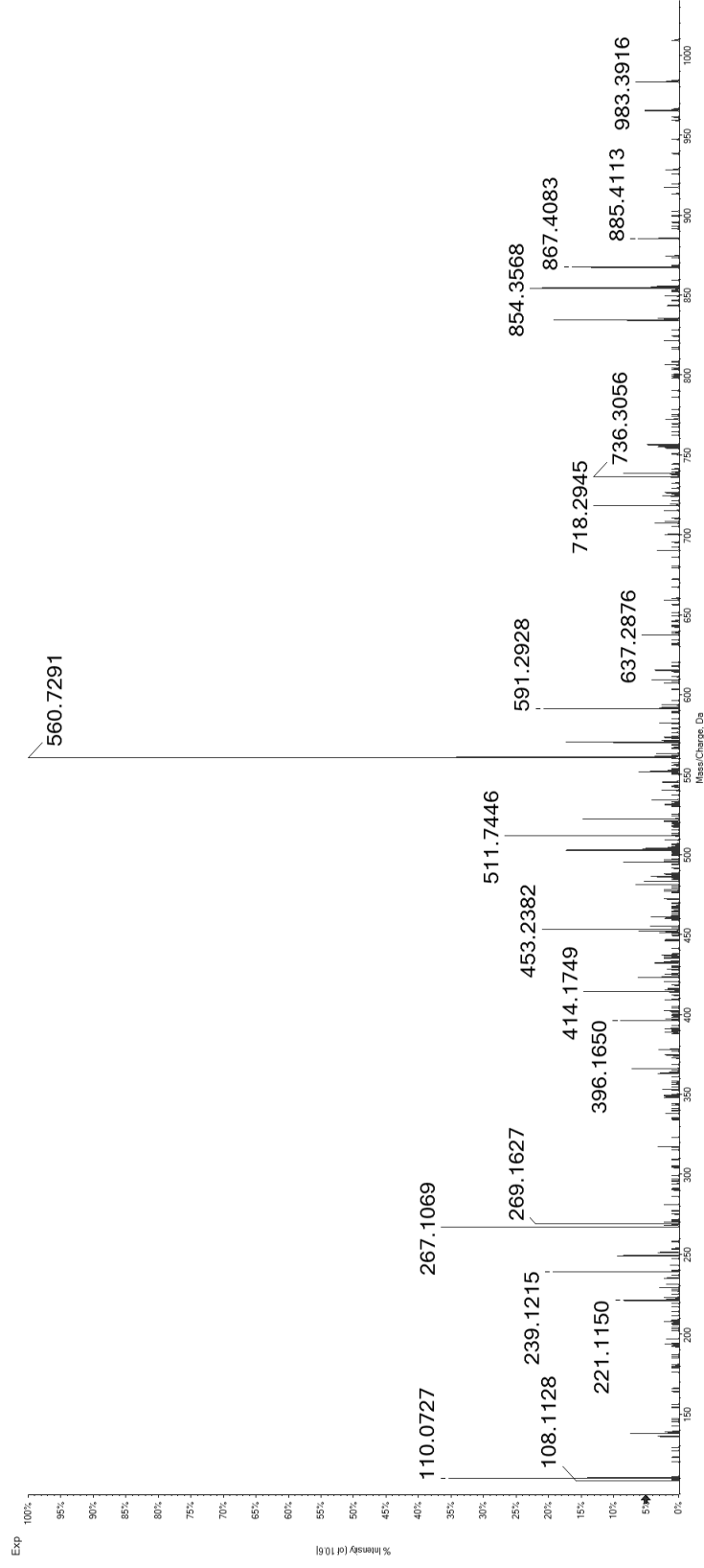
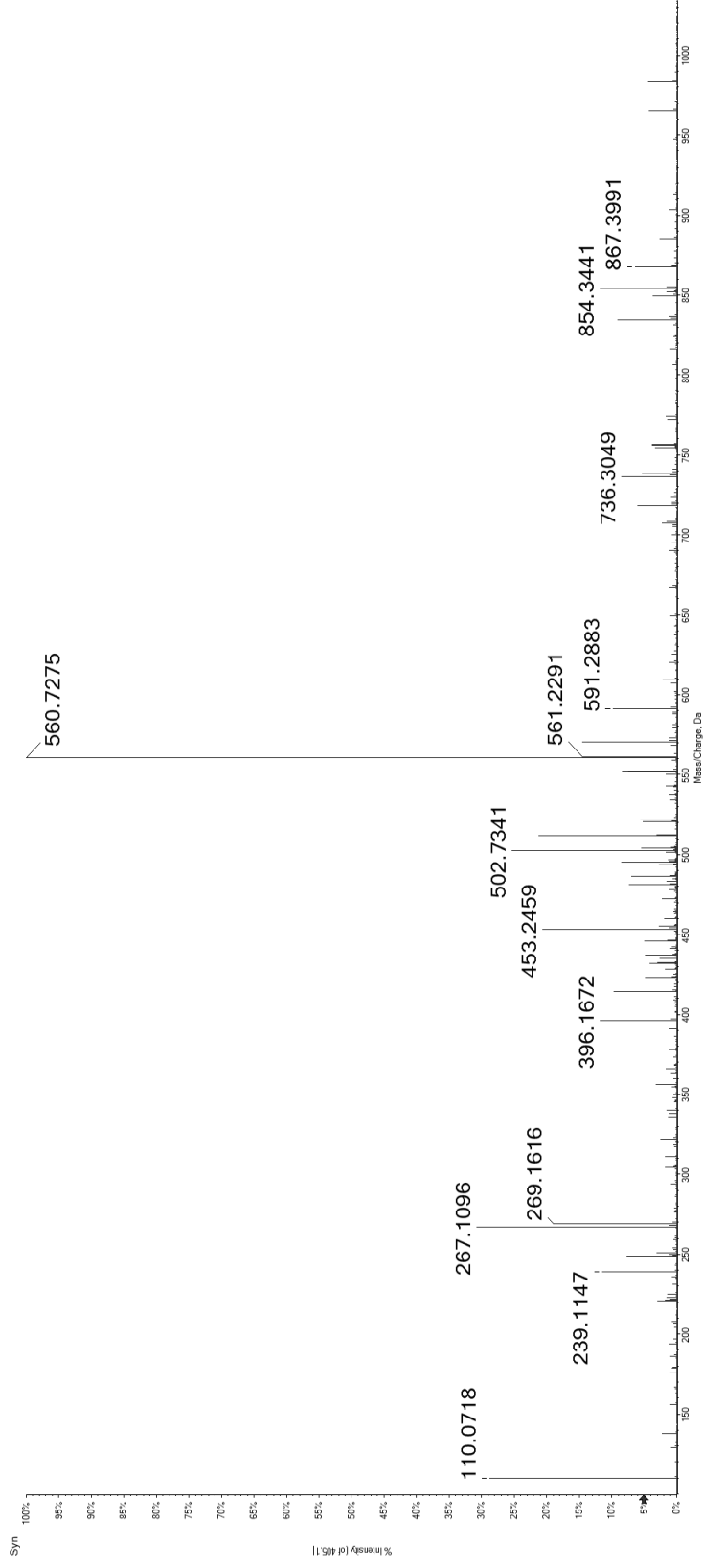
Candidate
Spectrum



61

HEFS(pS)PSHL

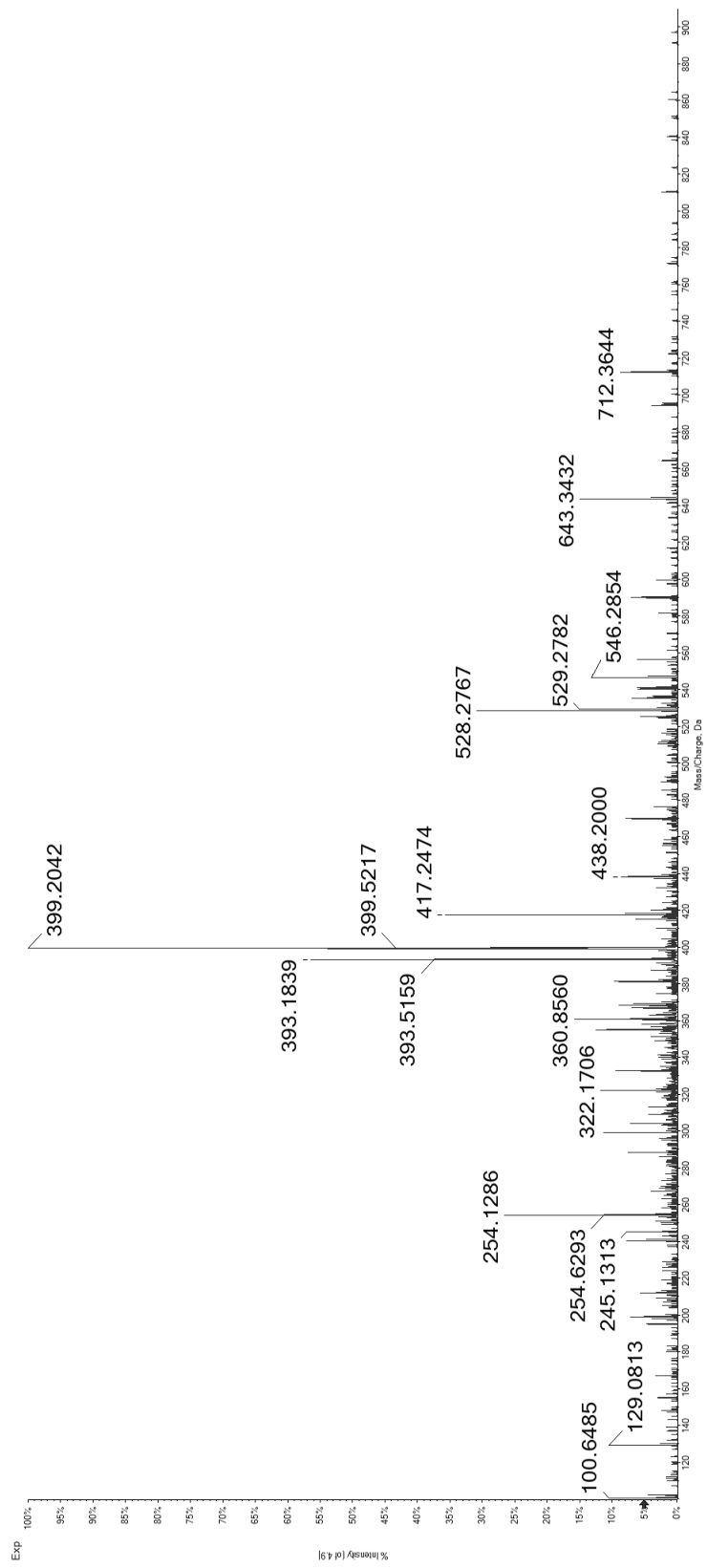
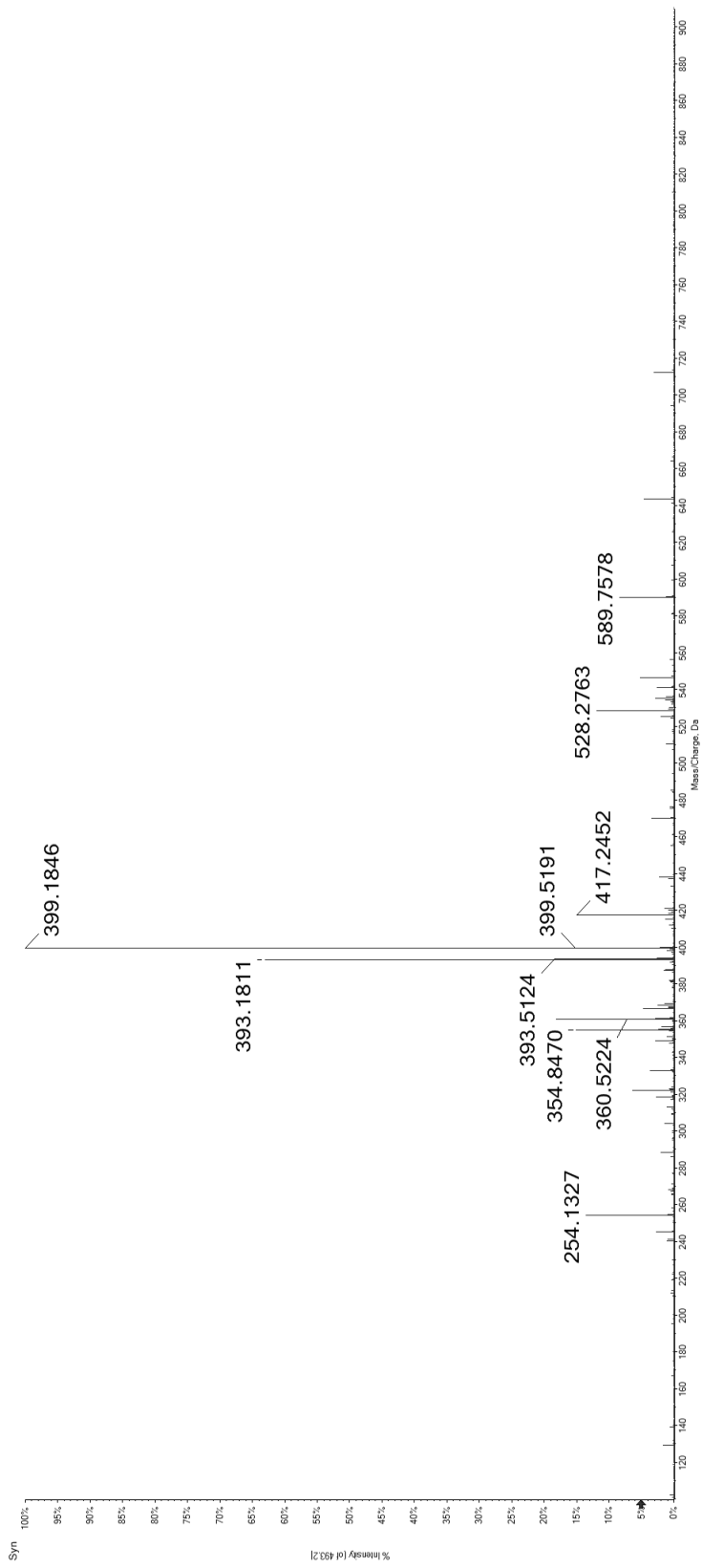
Synthetic Peptide



Candidate Spectrum

62 QEK(pS)PEERL

Synthetic
Peptide

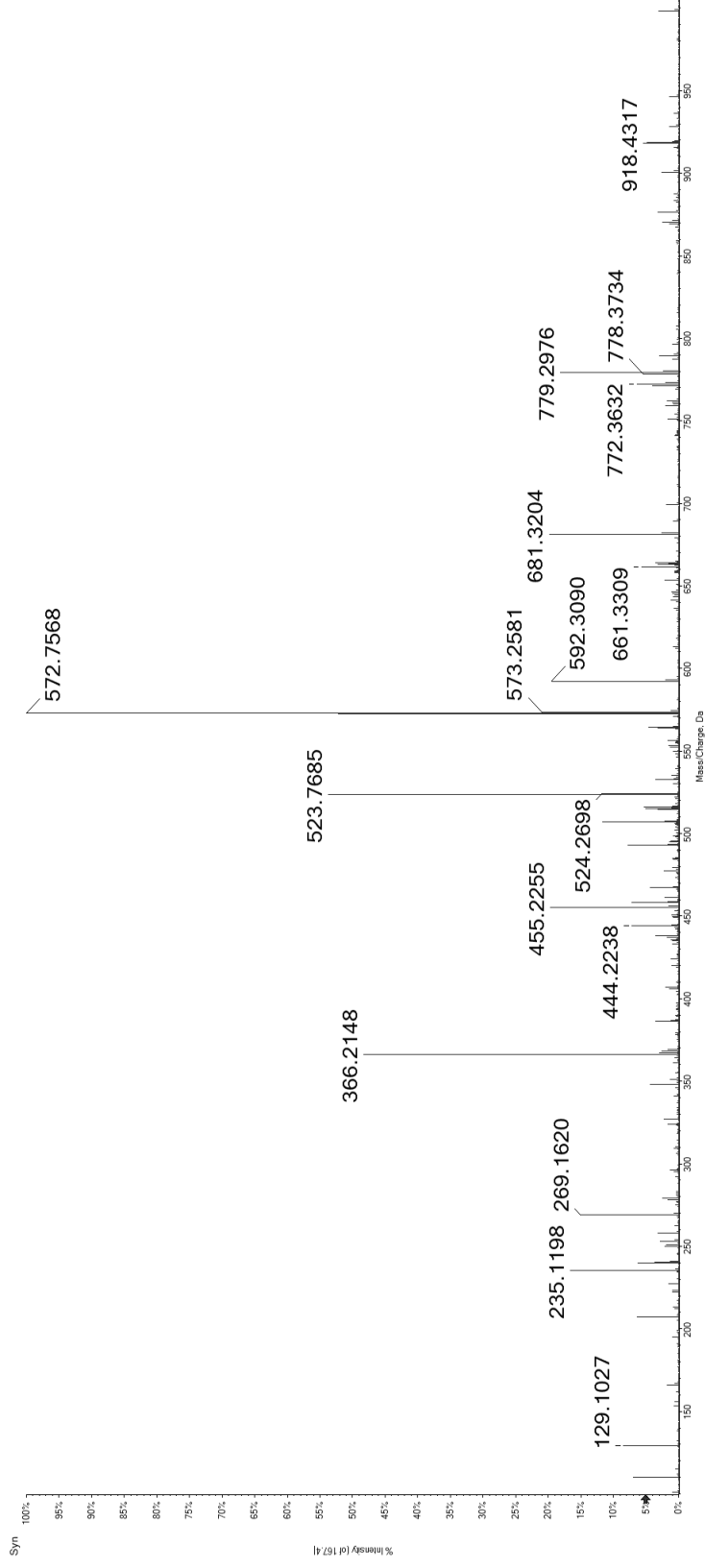


Candidate
Spectrum

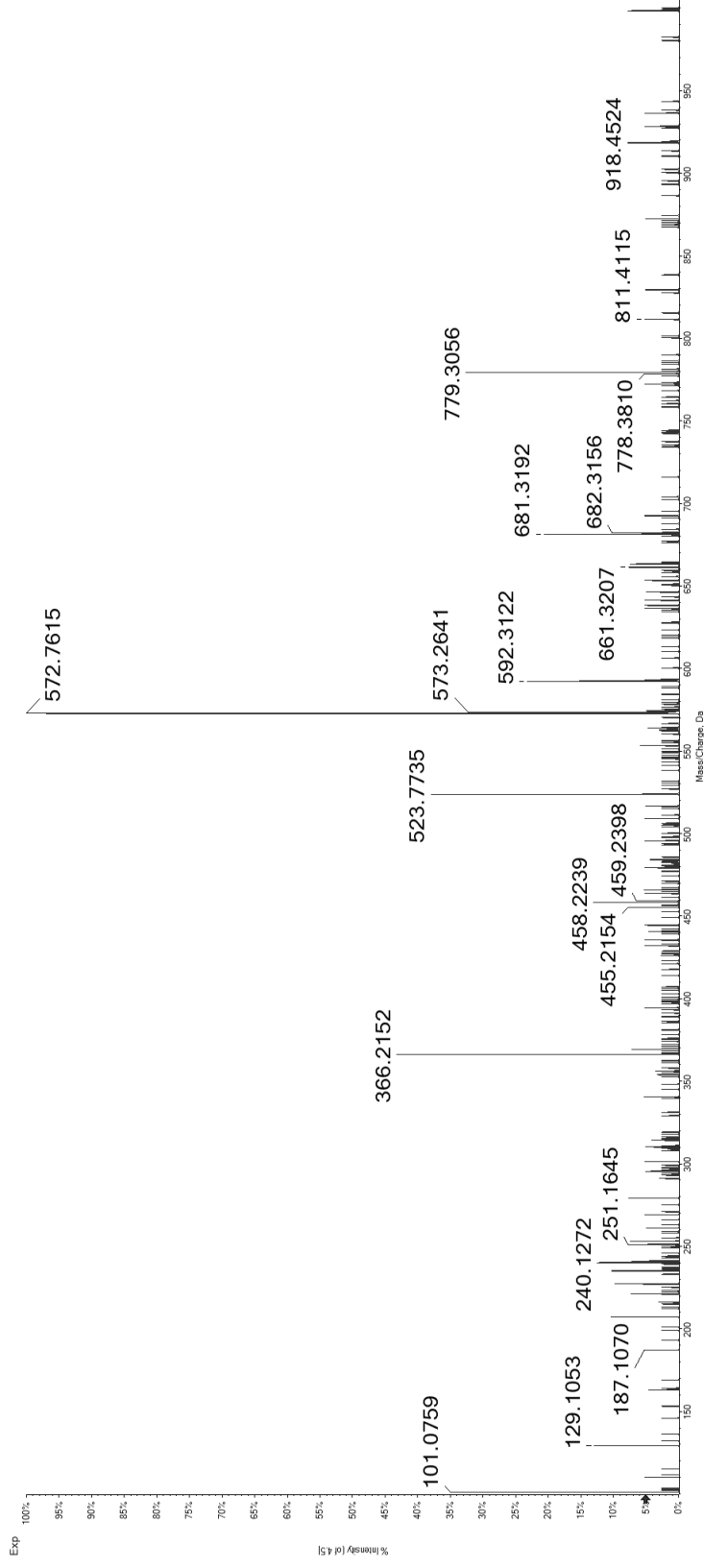
63

KEQ(pS)PEPHL

Synthetic Peptide



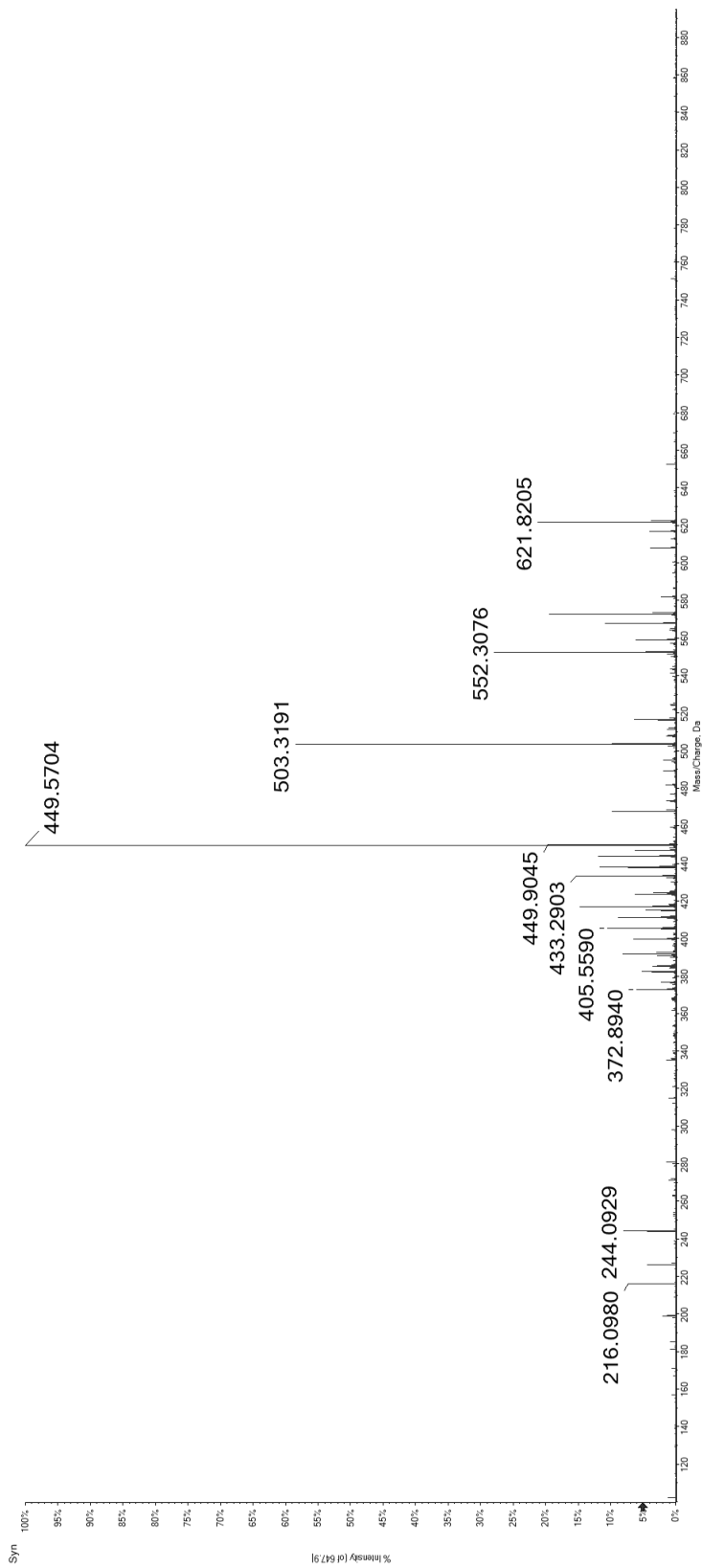
Candidate Spectrum



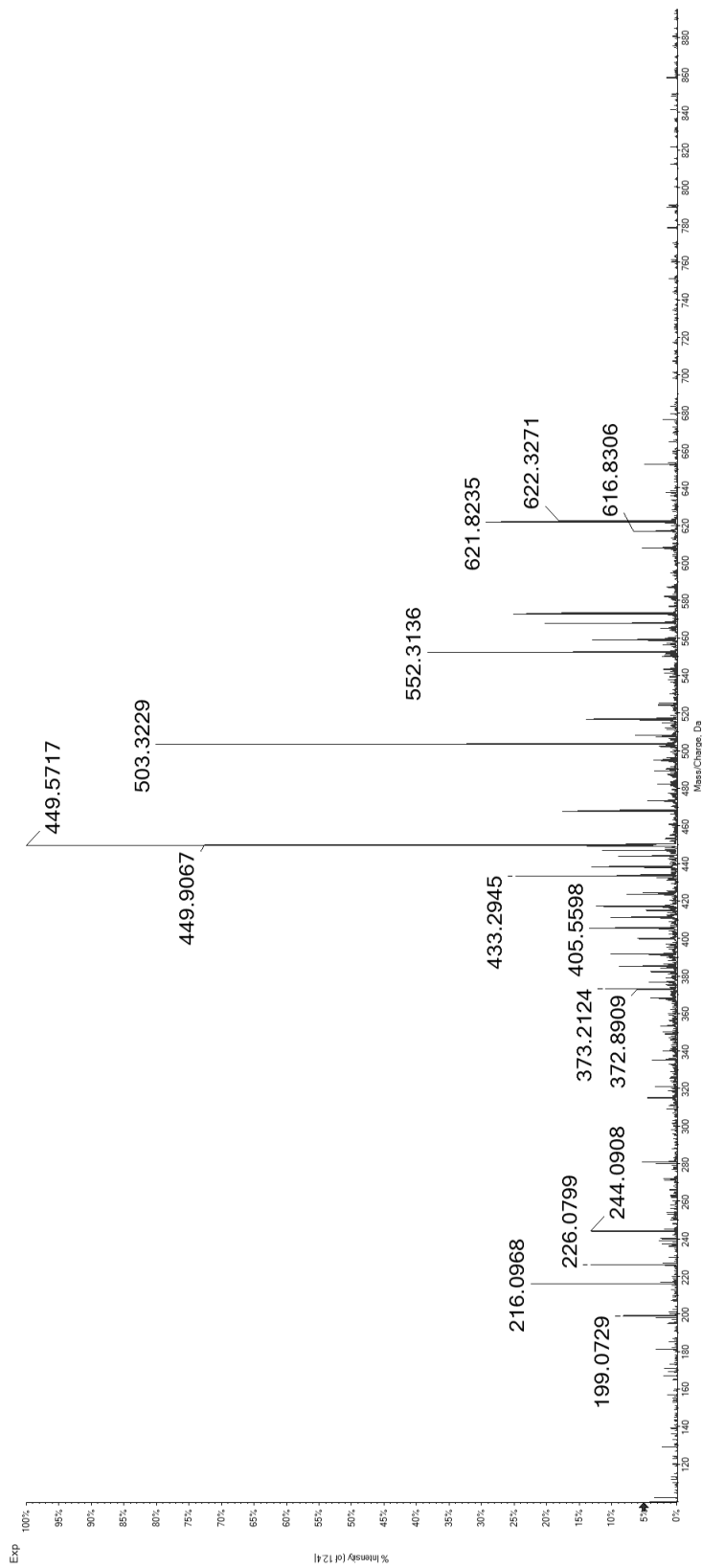
64

NEA(pS)PVKRPRL

Synthetic Peptide



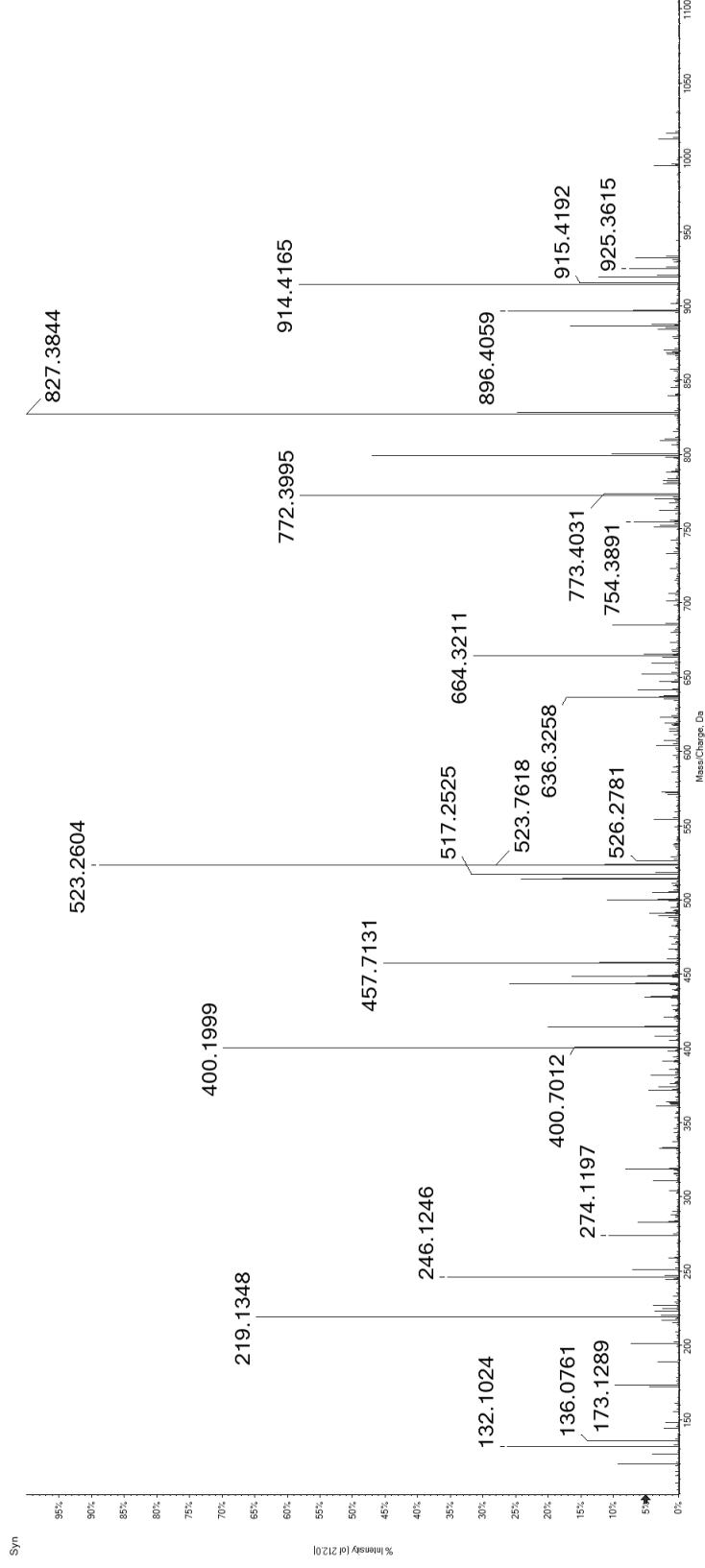
Candidate Spectrum



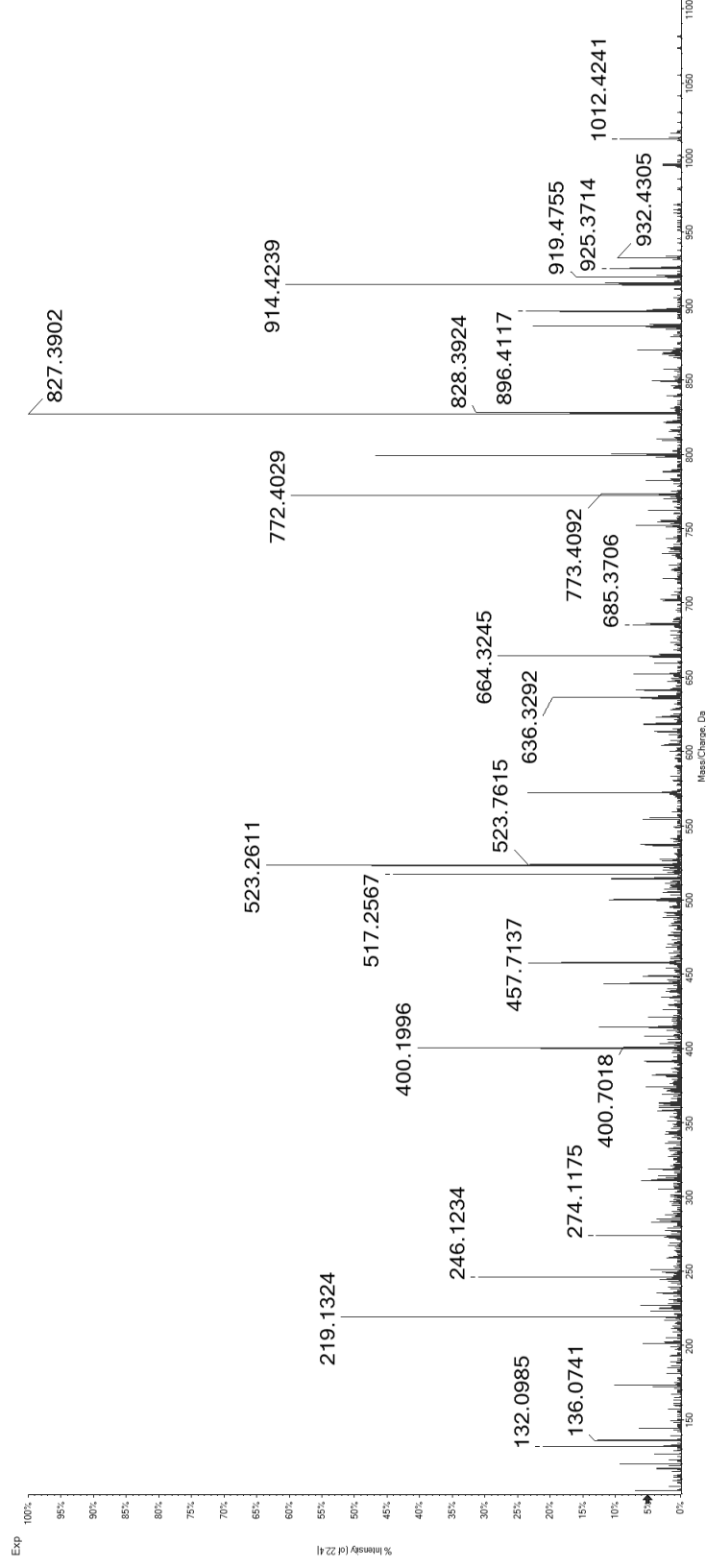
65

G(p)sFSRFYSL

Synthetic Peptide



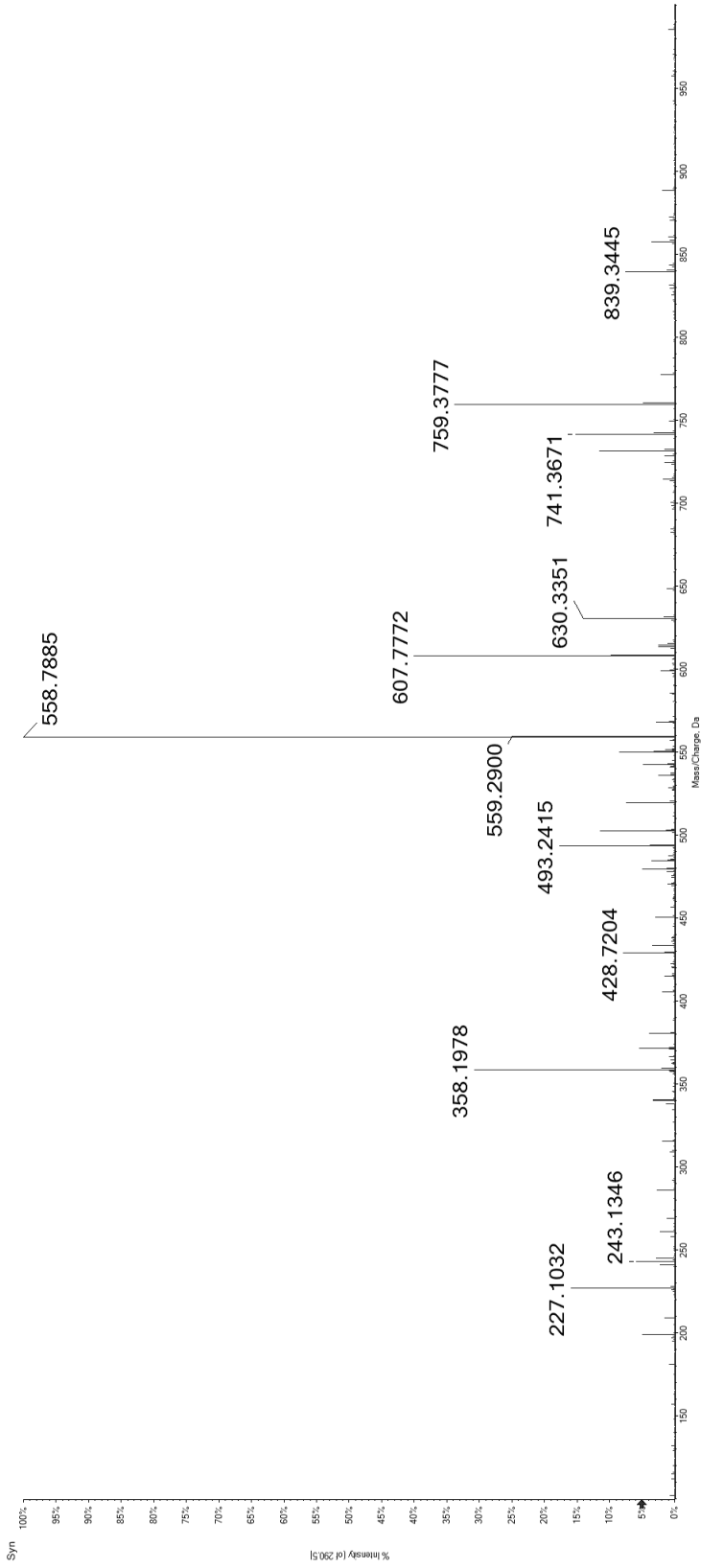
Candidate Spectrum



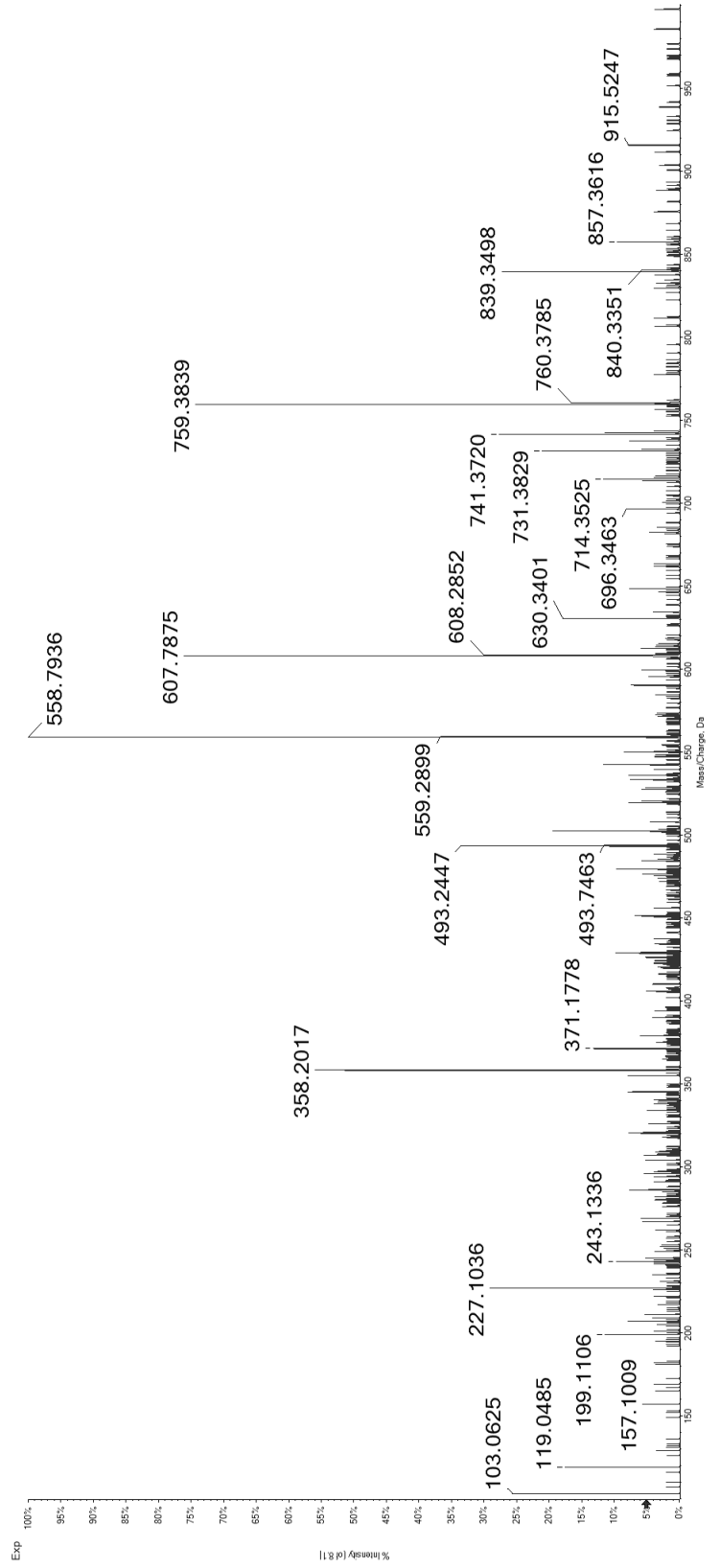
66

REF(pS)KEPEL

Synthetic Peptide



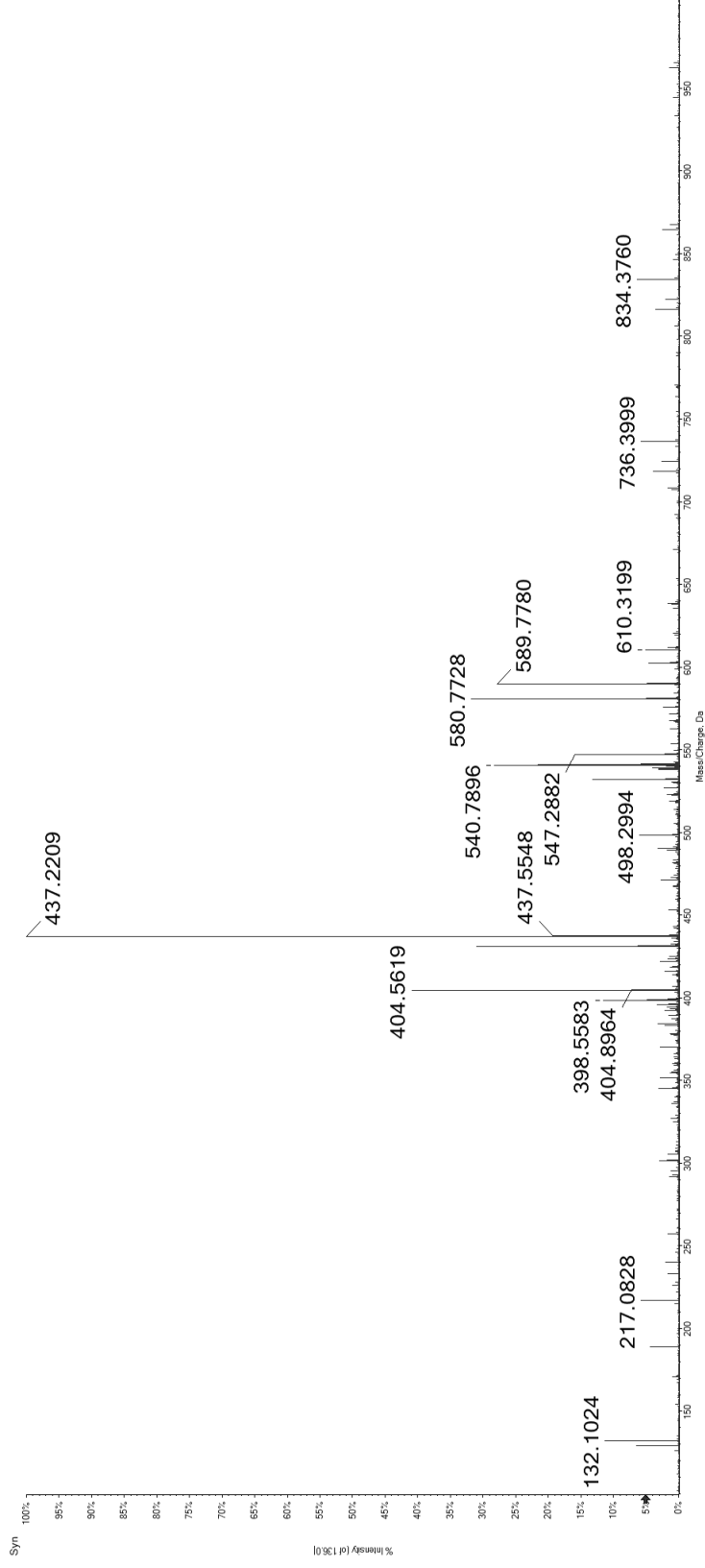
Candidate Spectrum



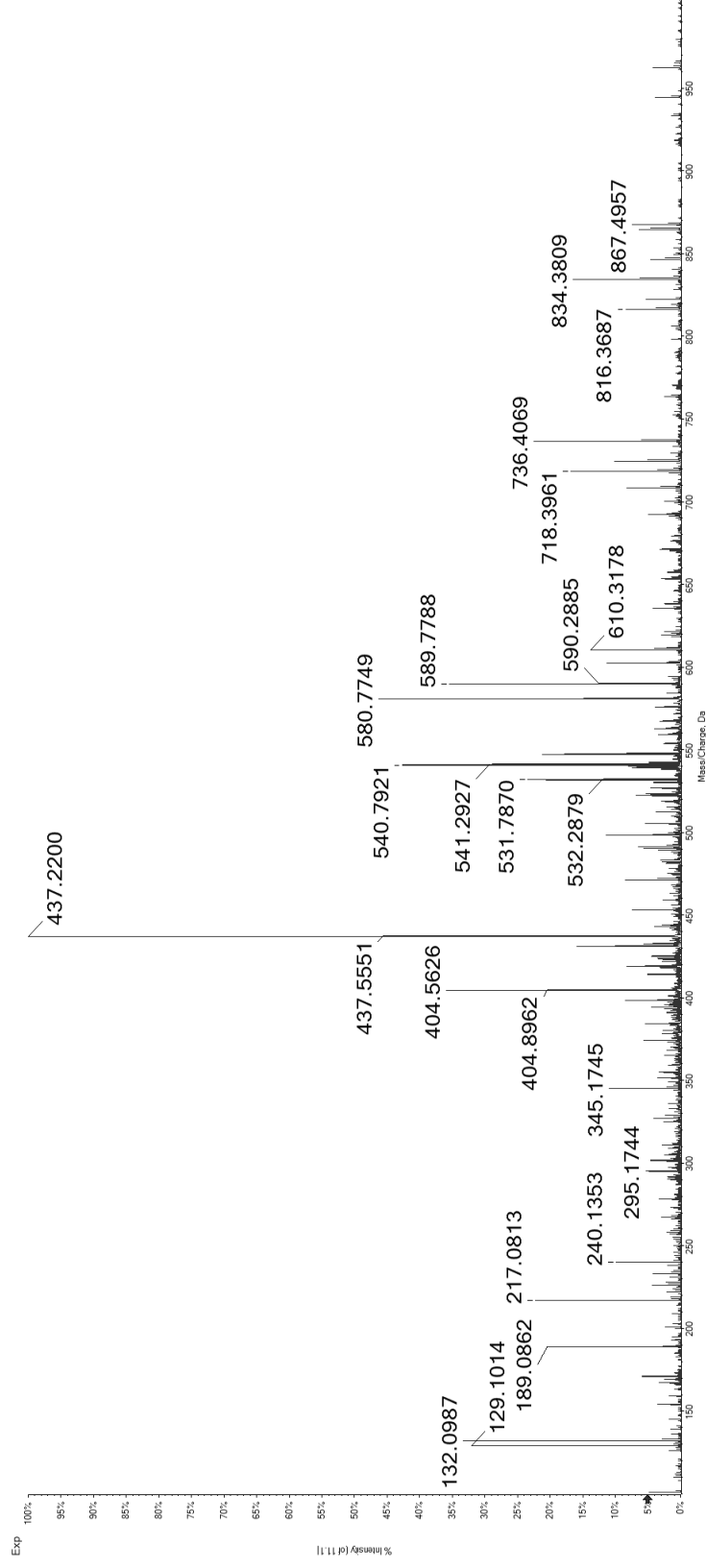
67

SEKPV(ps)PKSGTL

Synthetic Peptide



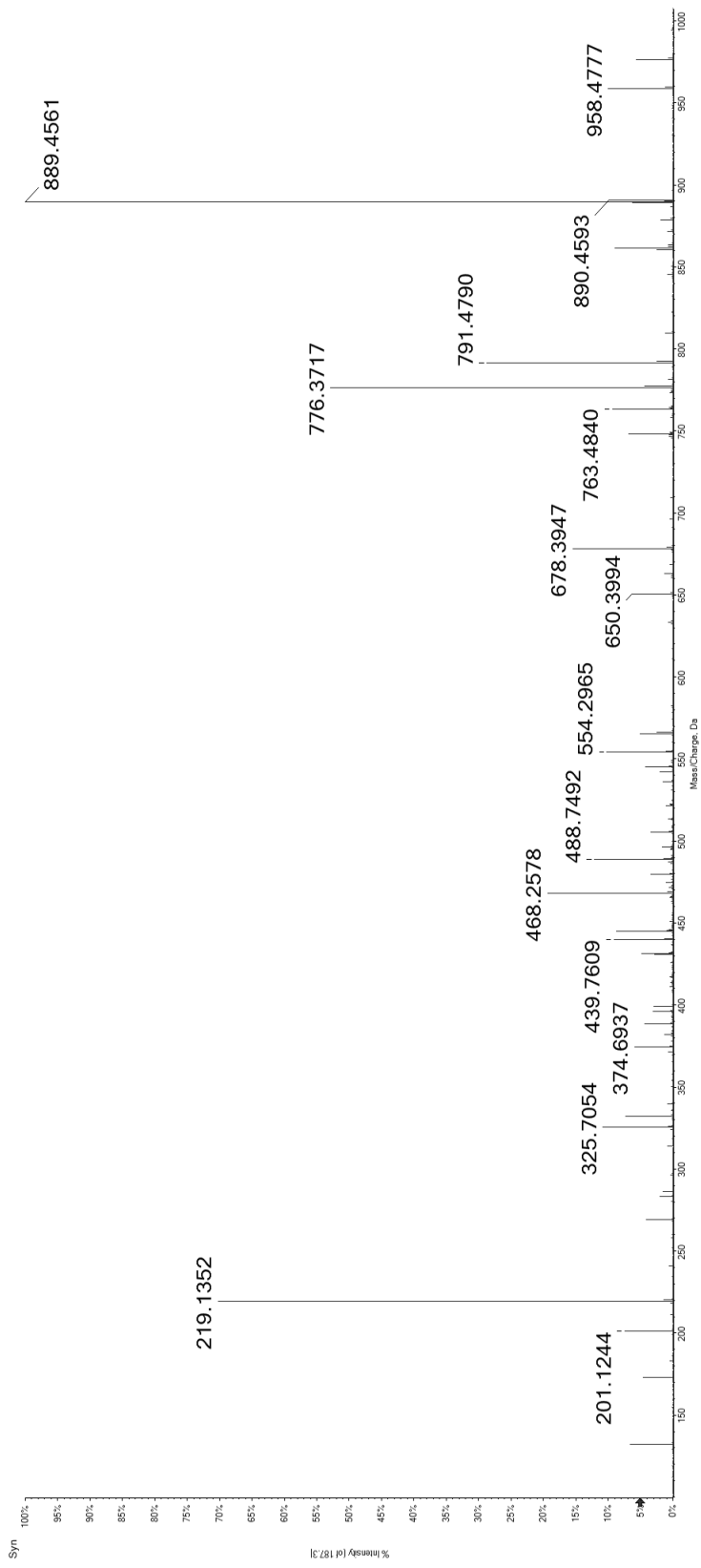
Candidate Spectrum



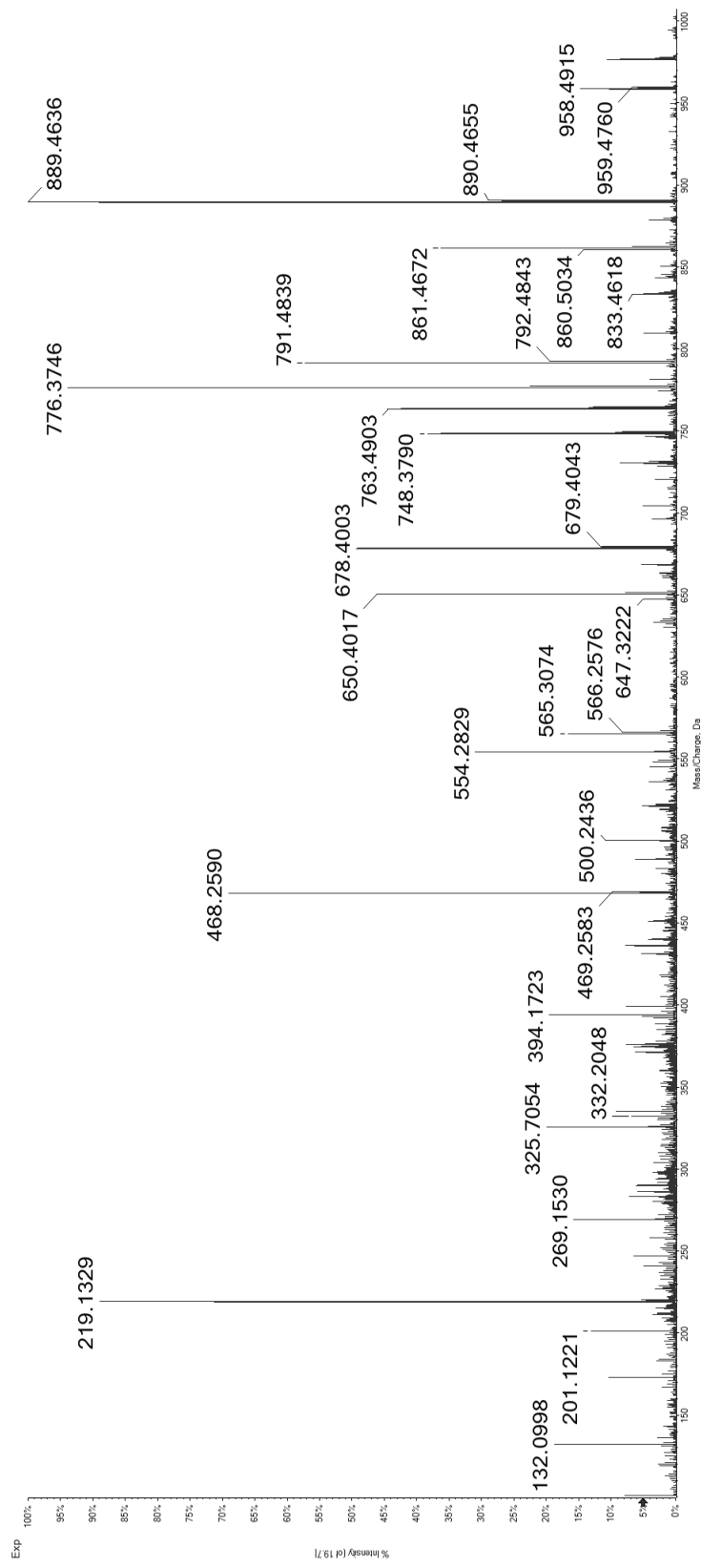
68

REL(ps)PLISL

Synthetic Peptide



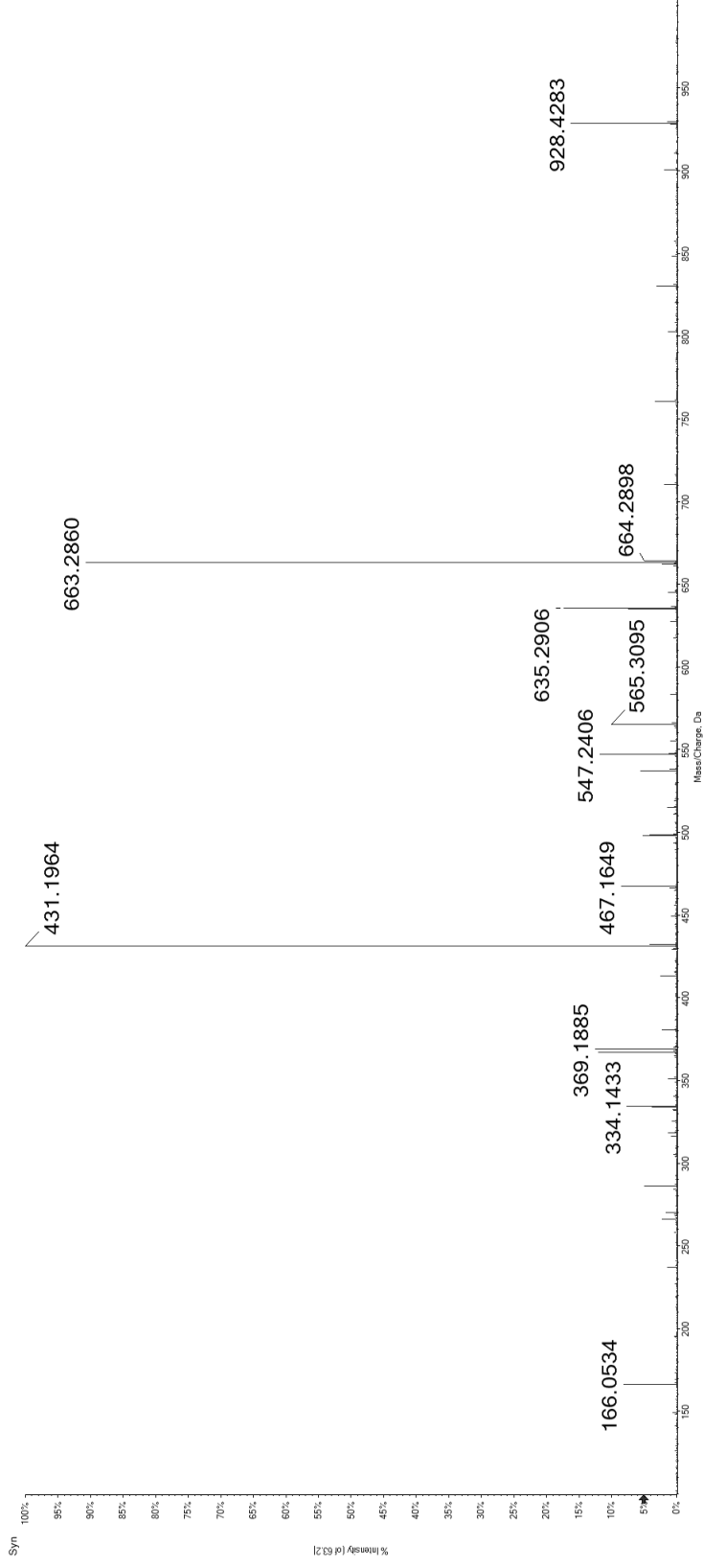
Candidate Spectrum



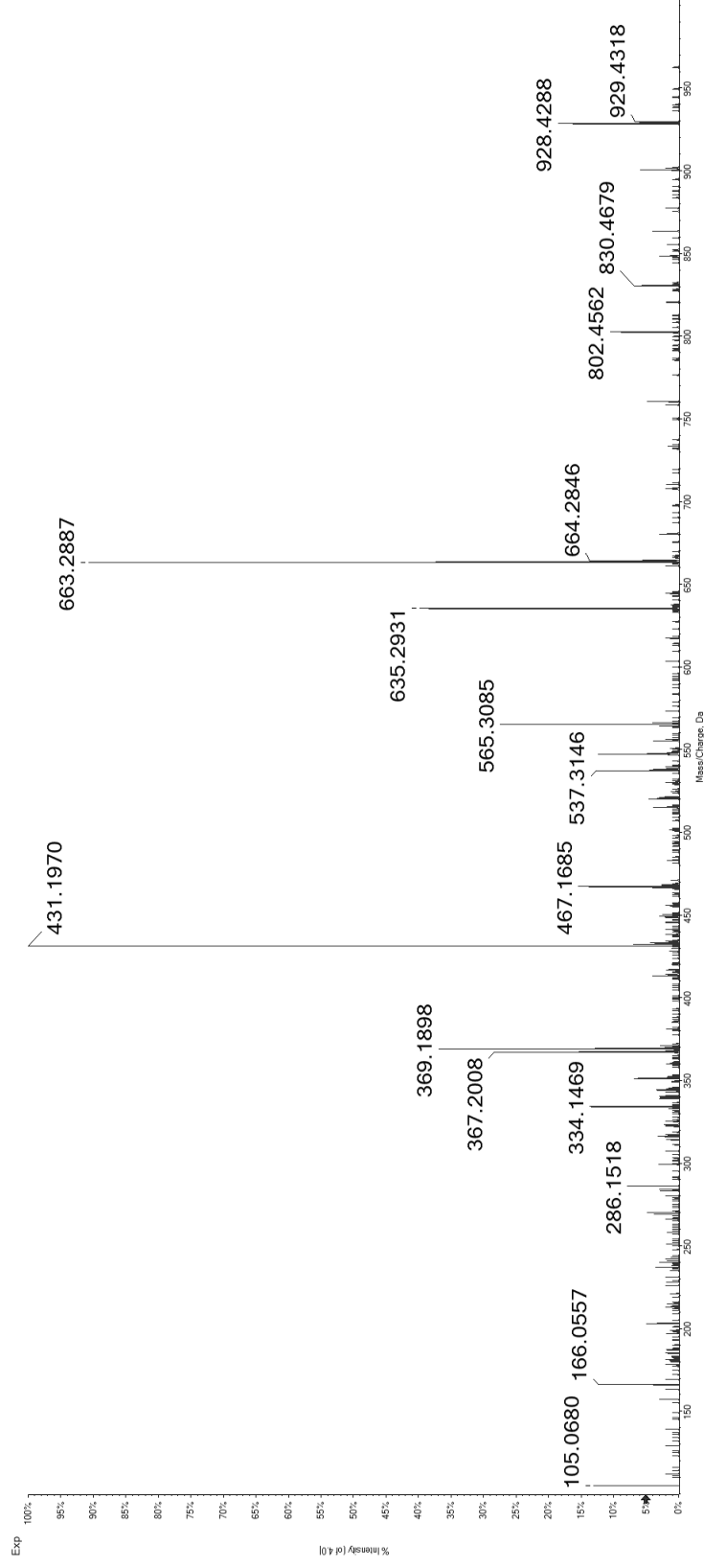
69

RE(pT)PVPPAM

Synthetic Peptide



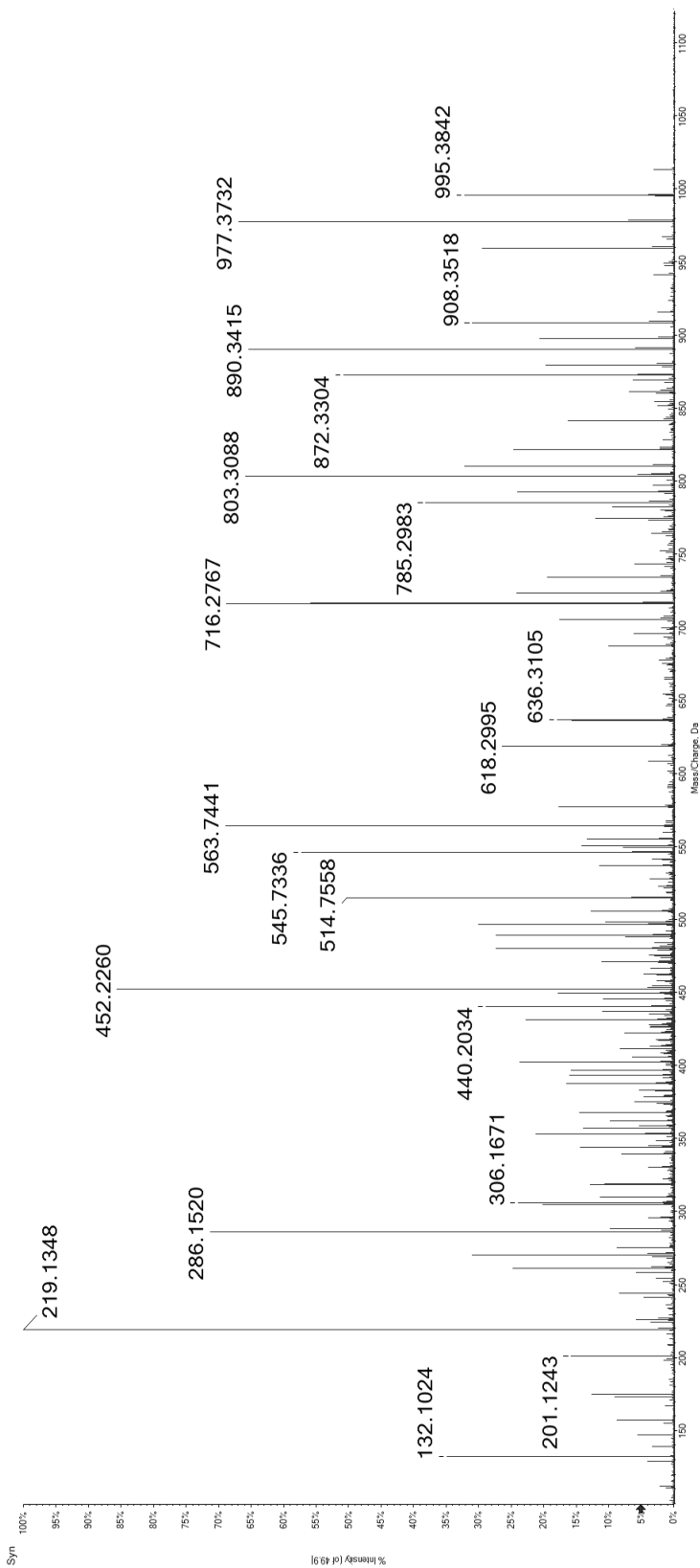
Candidate Spectrum



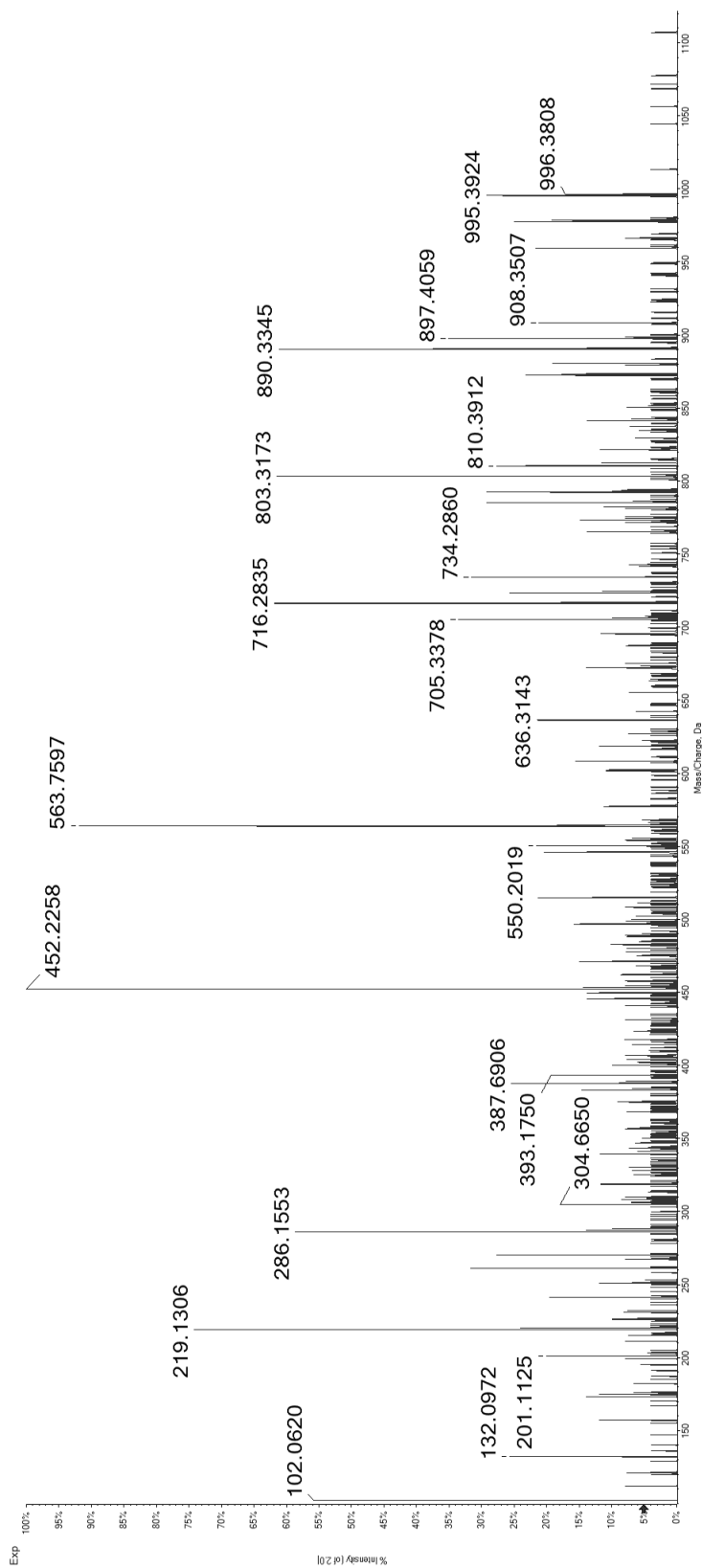
70

REP(pS)PSSSSL

Synthetic Peptide



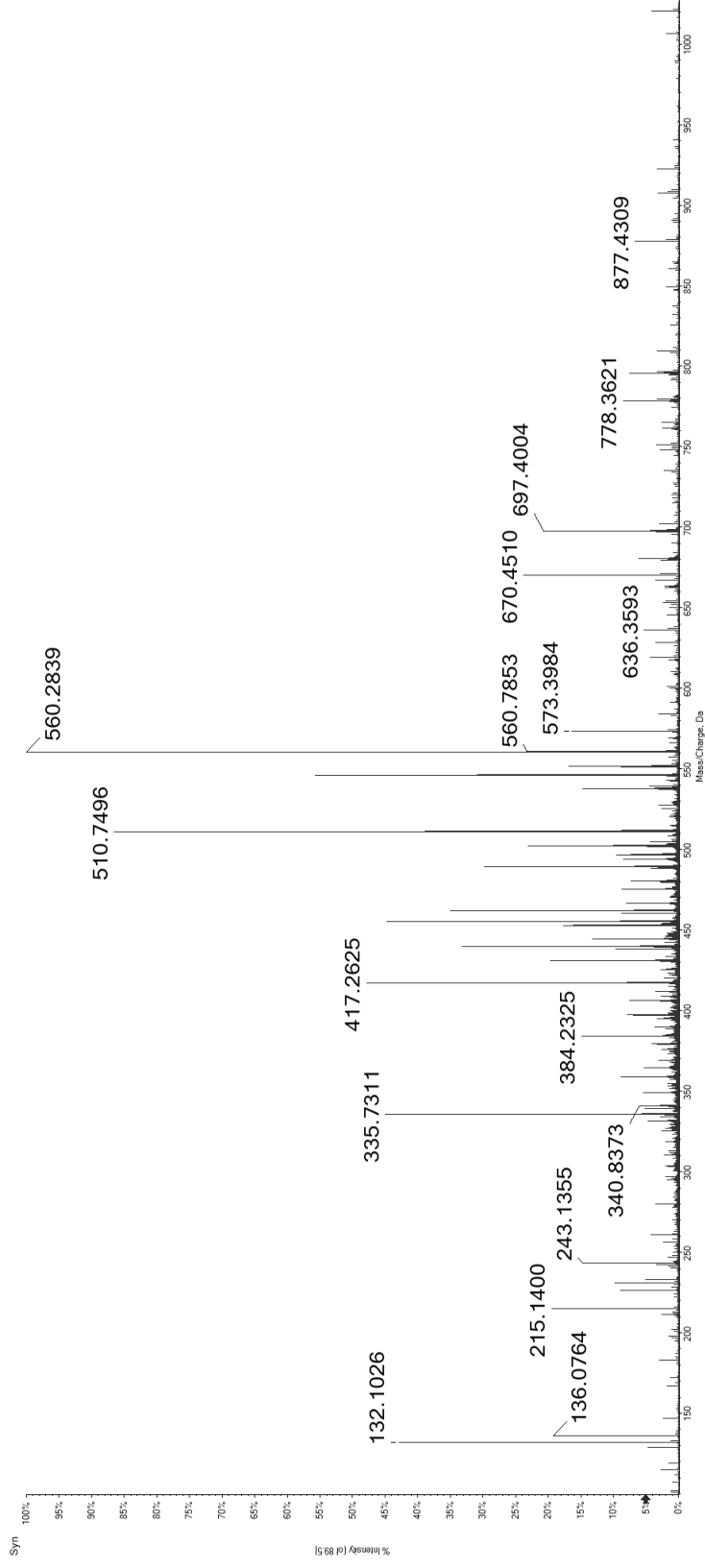
Candidate Spectrum



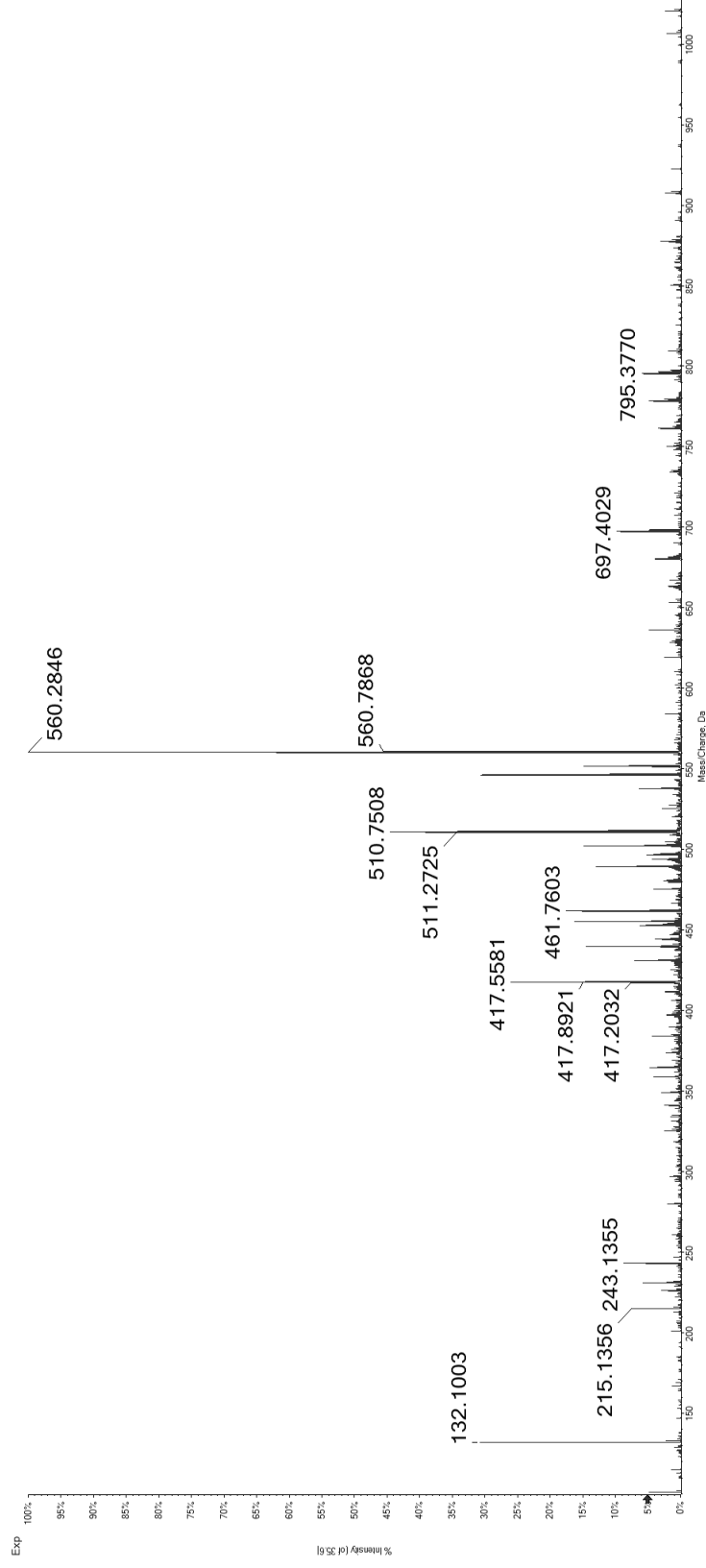
71

LEKRT(pS)PQVL

Synthetic Peptide



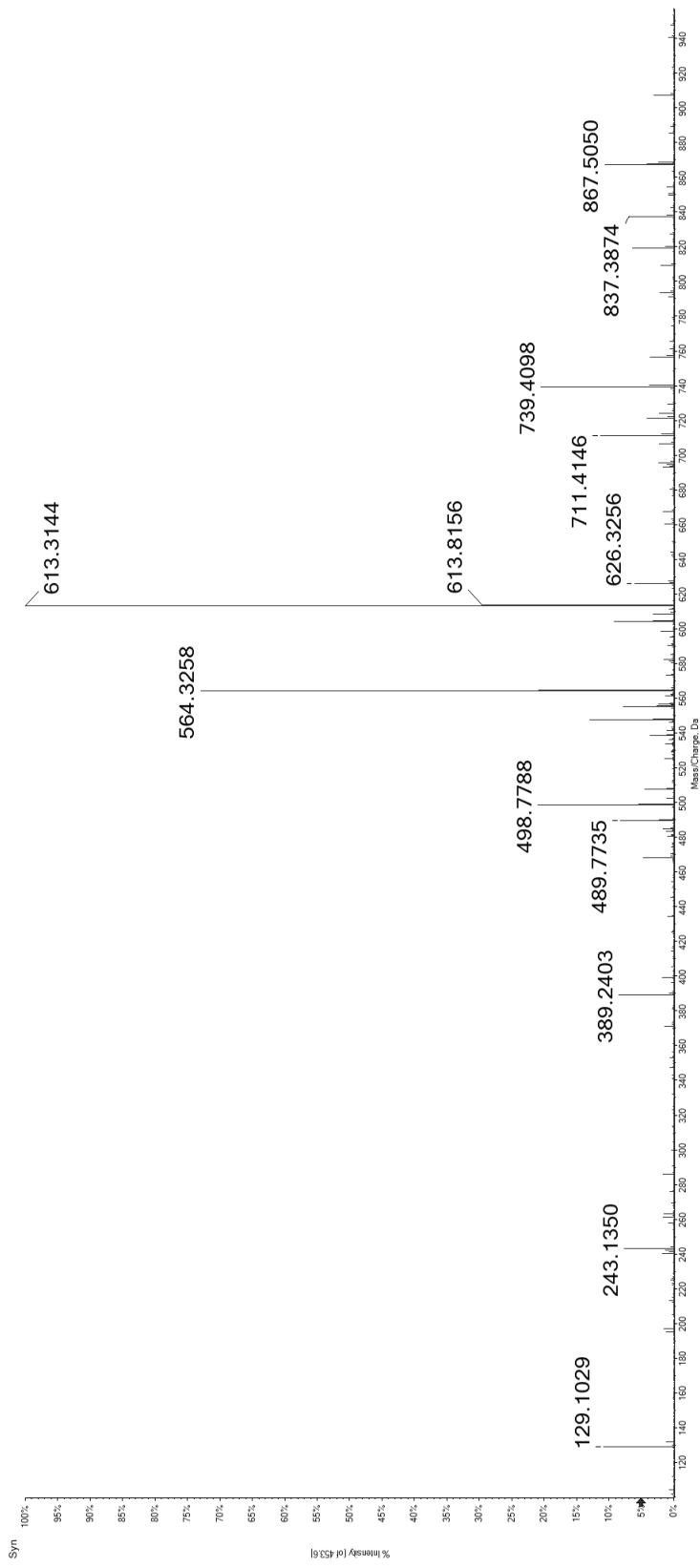
Candidate Spectrum



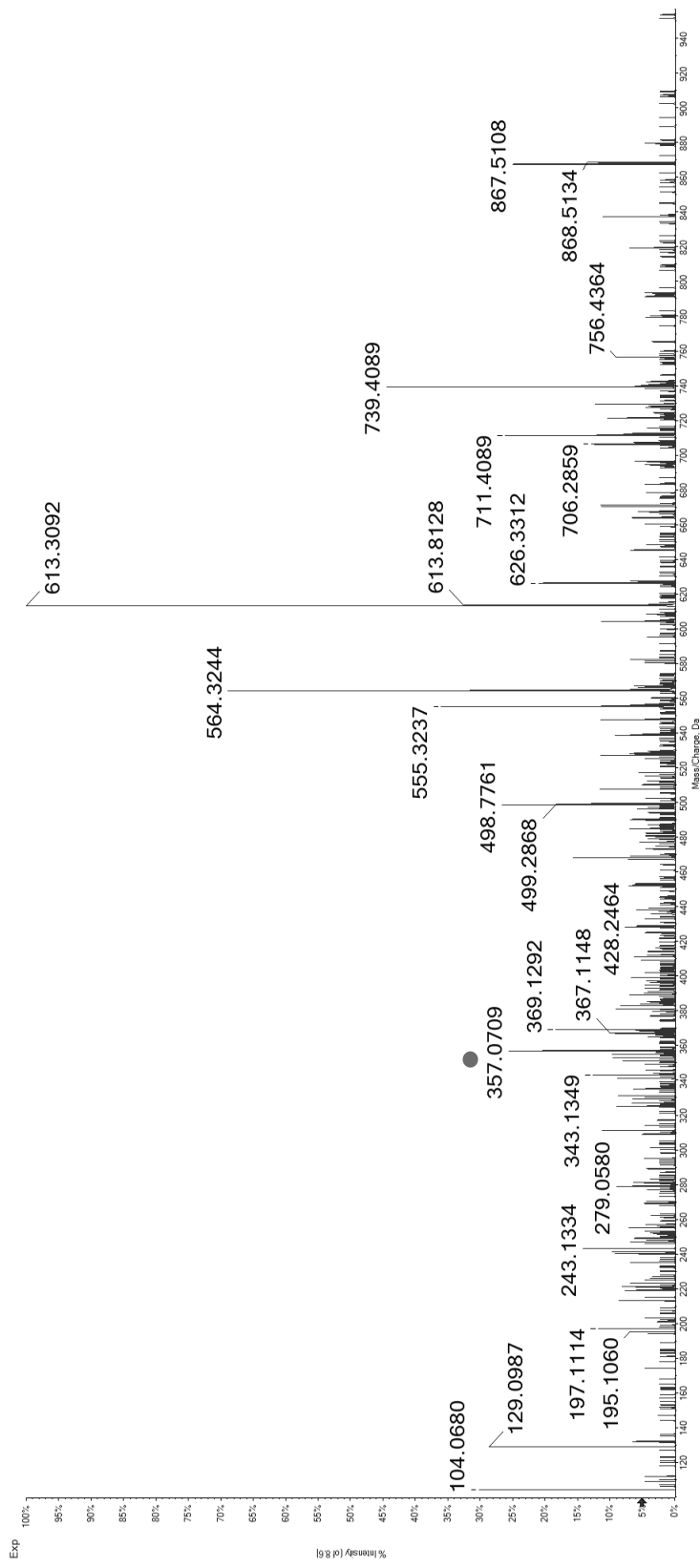
72

REL(ps)GTIKEI

Synthetic
Peptide
(Relative RT – 1,74)



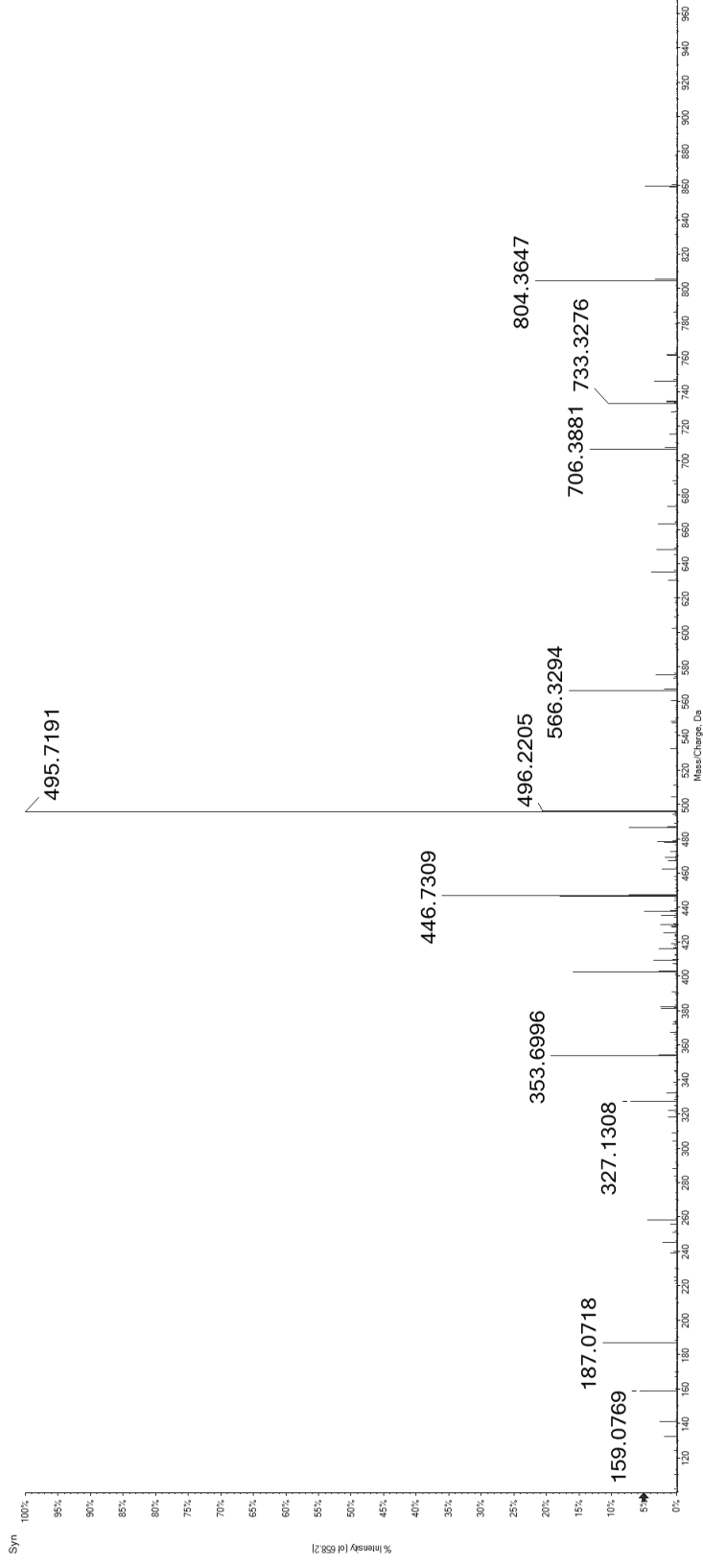
Candidate
Spectrum
(Relative RT – 1,63)



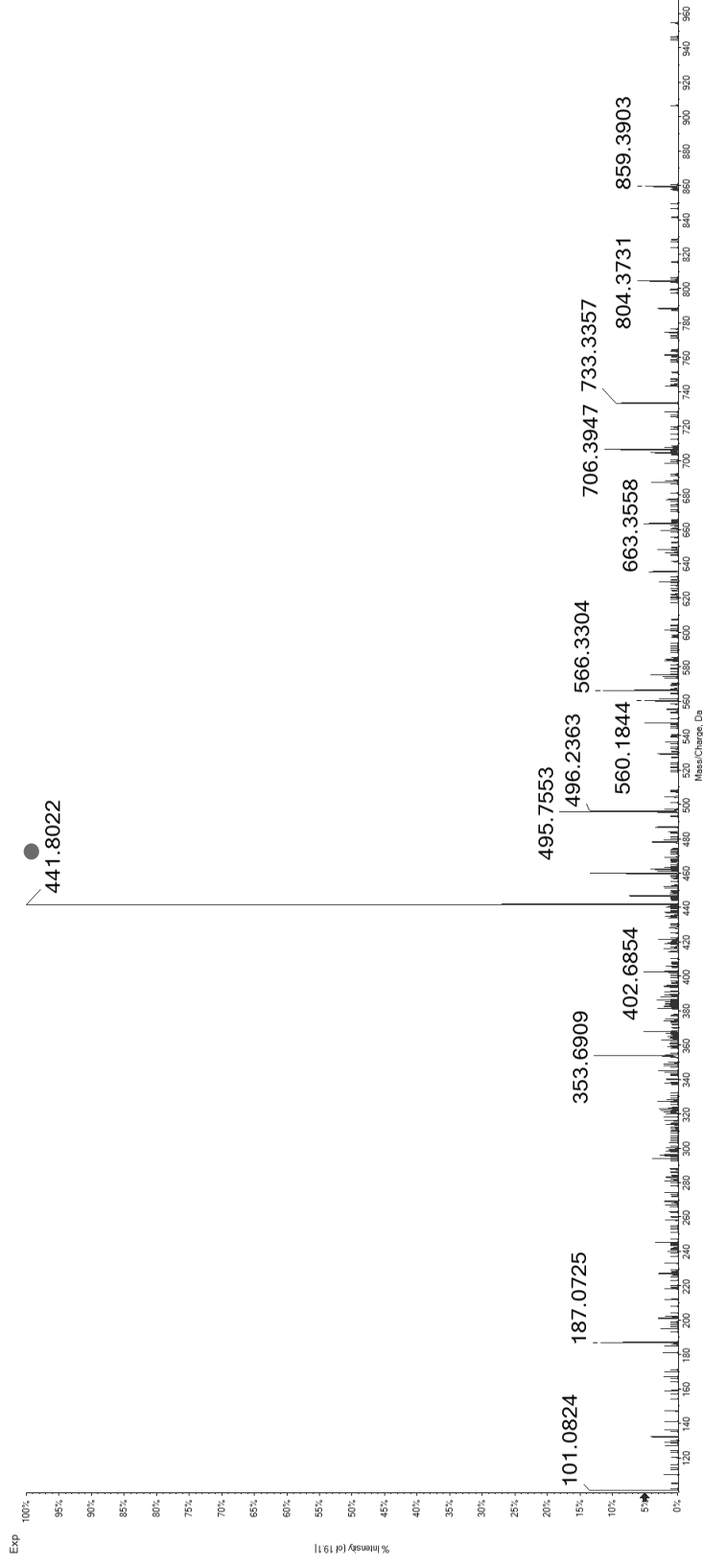
73

GEA(ps)PSHII

Synthetic
Peptide
(Relative RT – 1,60)



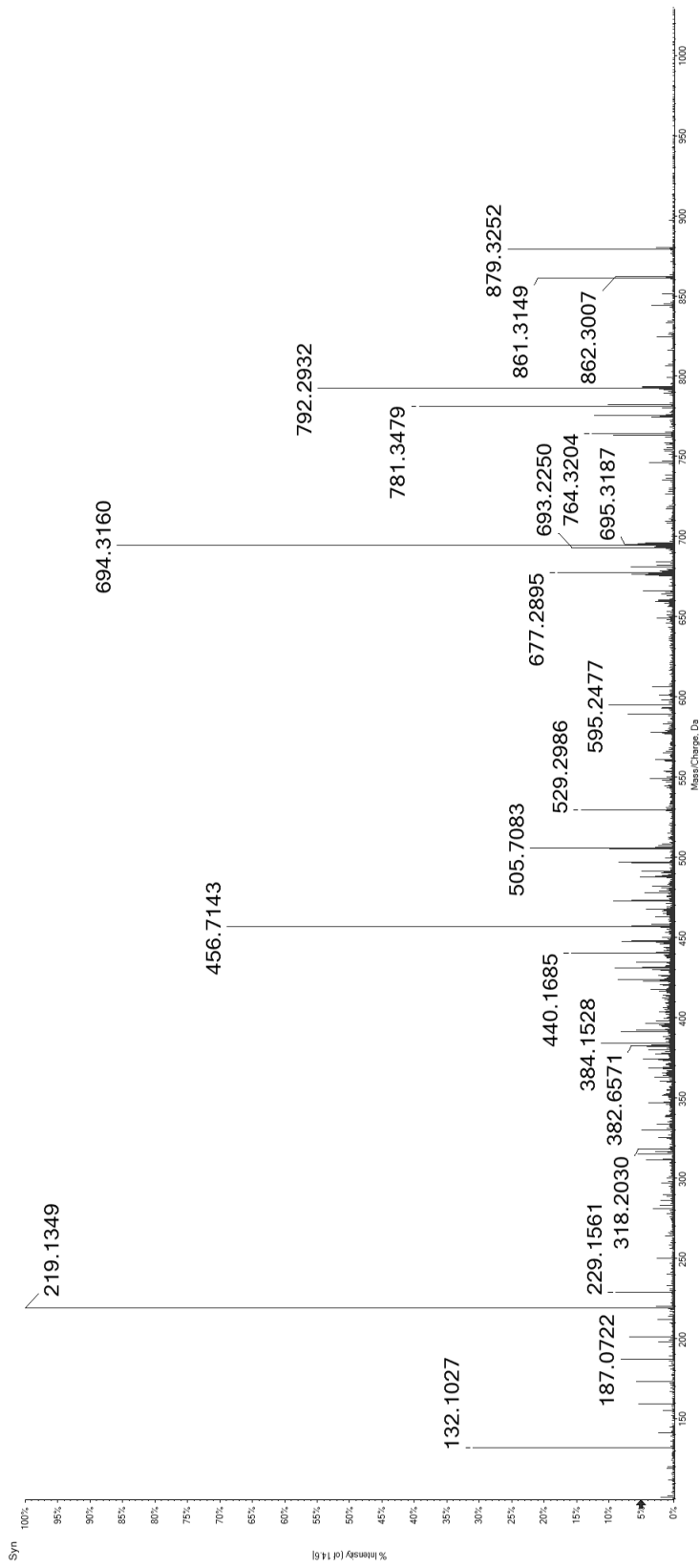
Candidate
Spectrum
(Relative RT – 1,52)



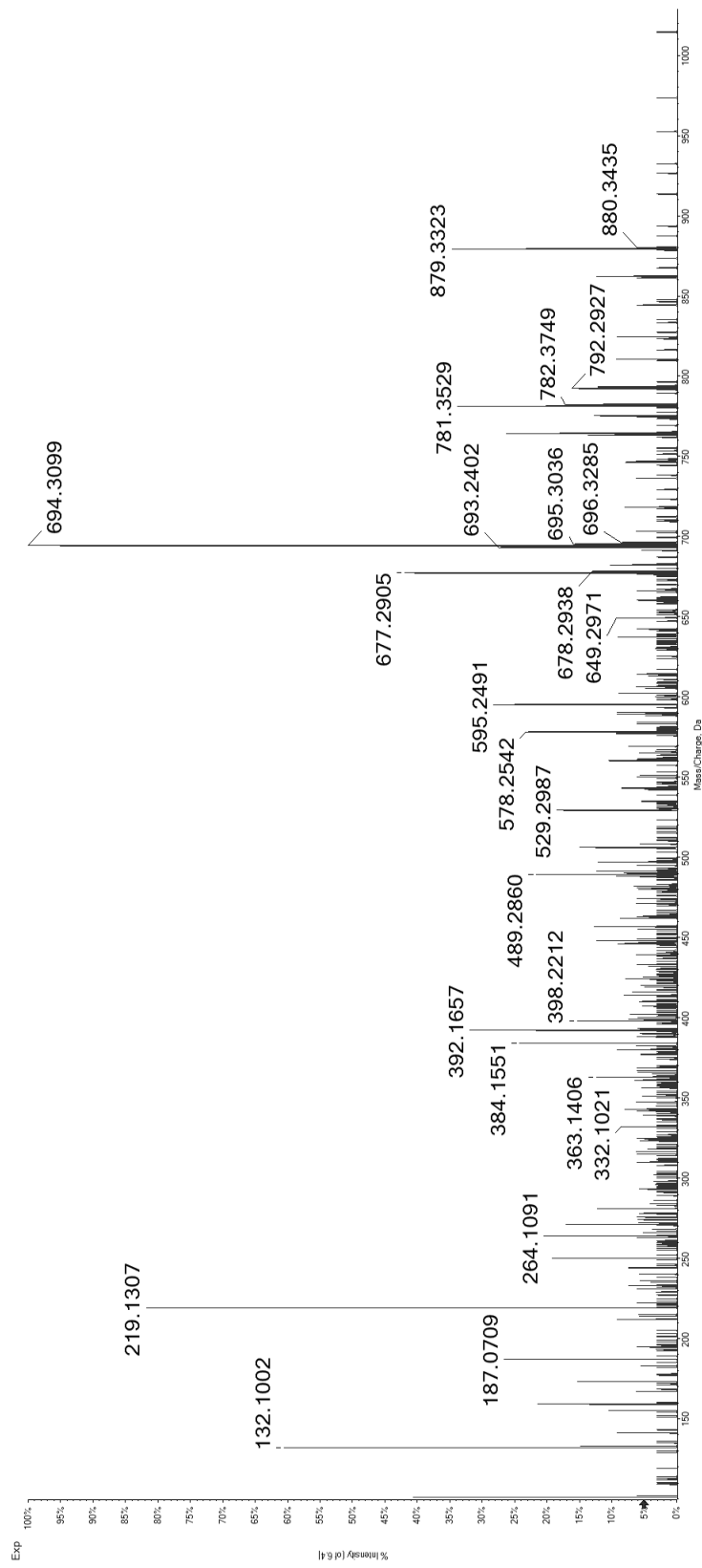
74

GEQ(ps)PNVSL

Synthetic Peptide



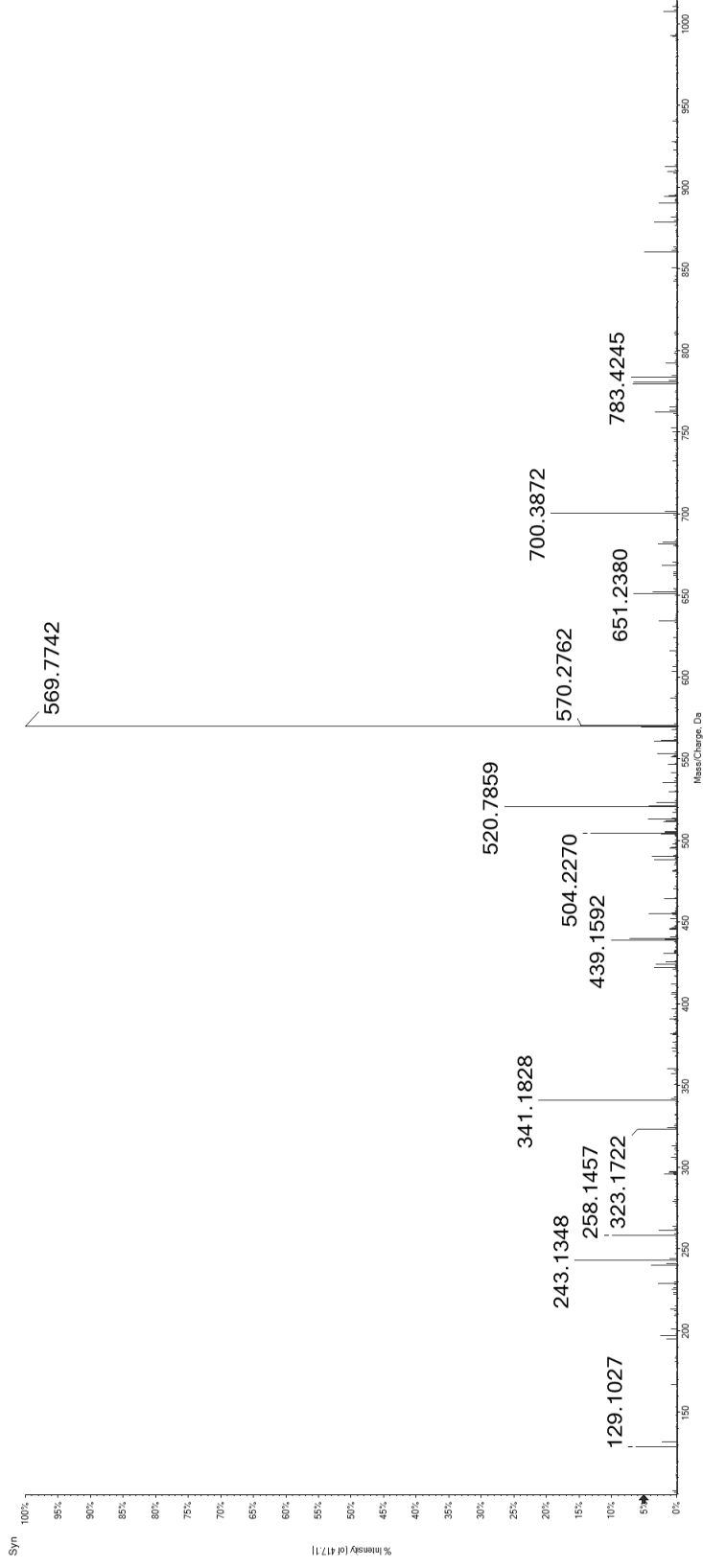
Candidate Spectrum



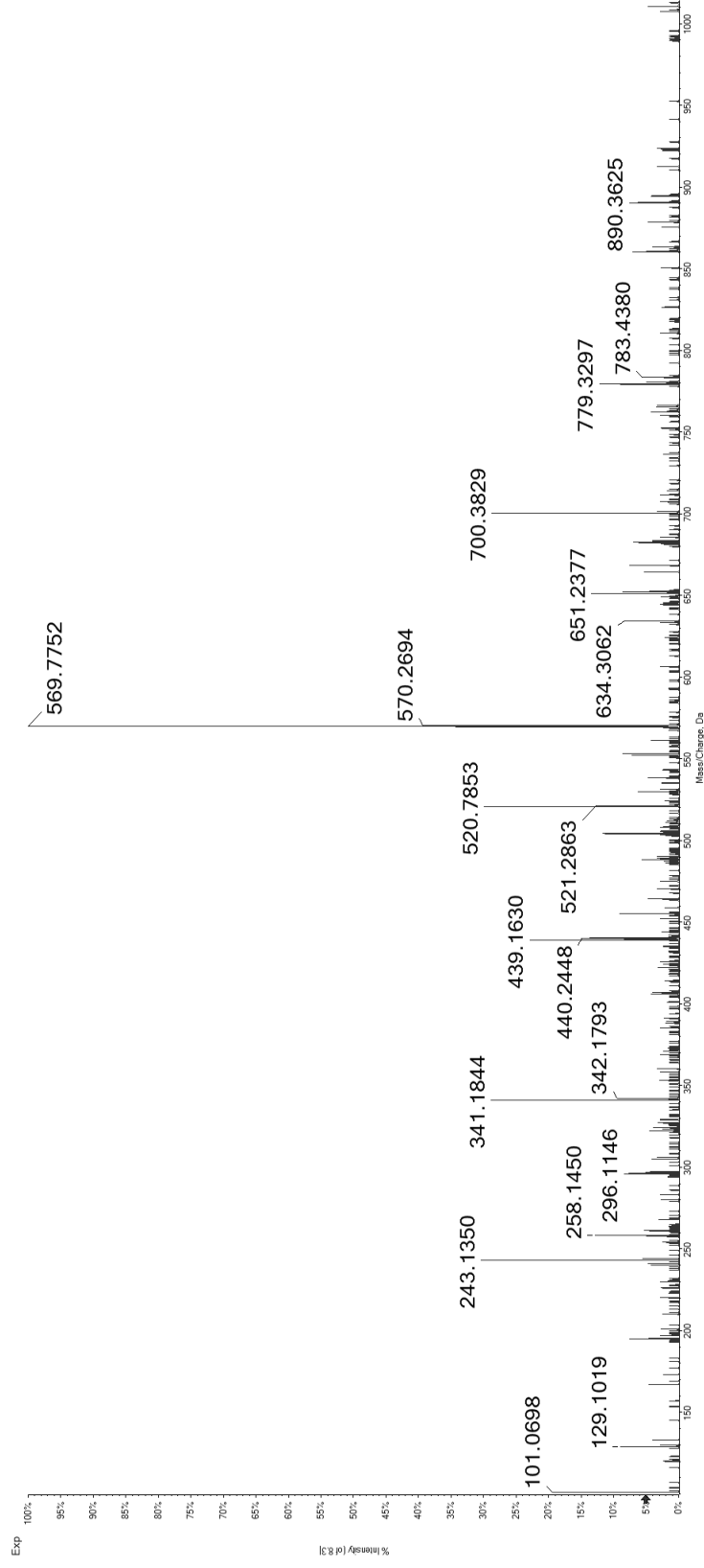
75

KE(pT)PDKVEL

Synthetic Peptide



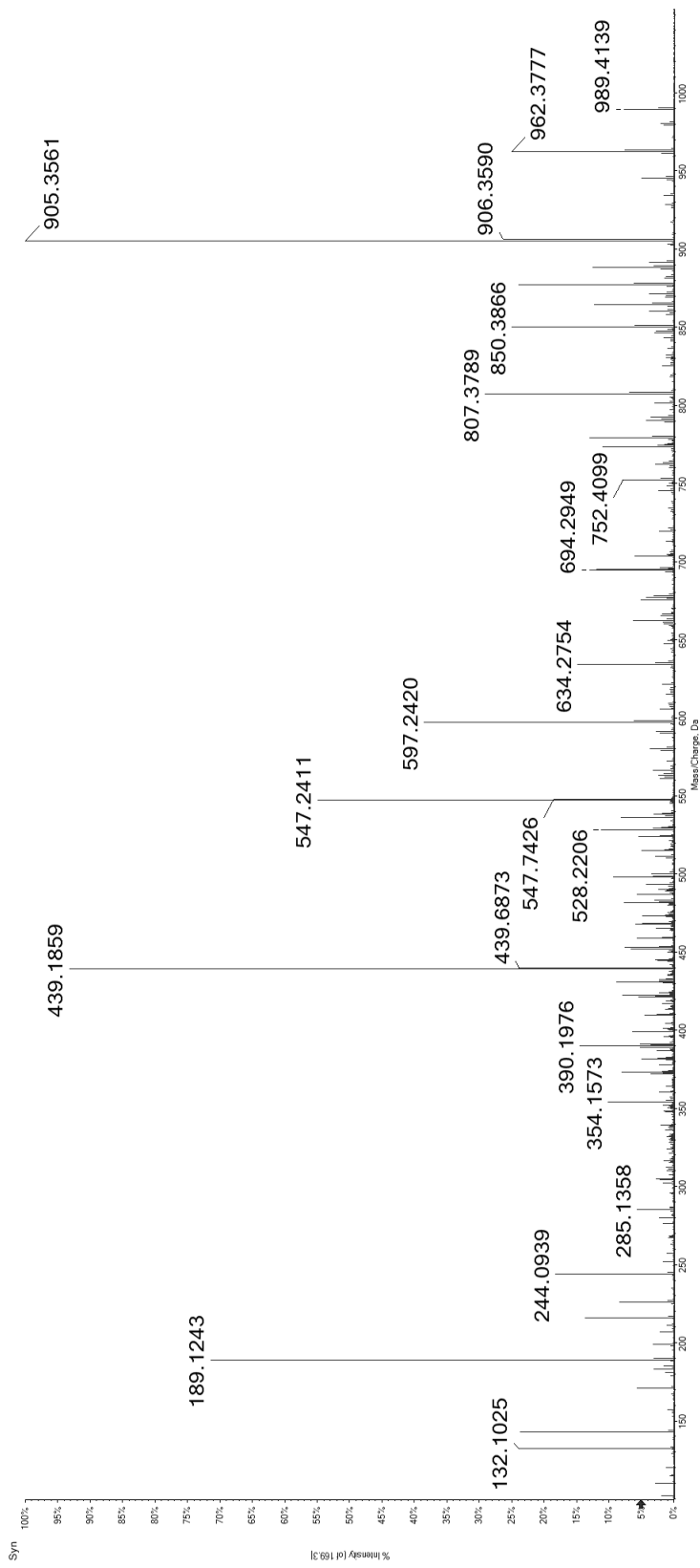
Candidate Spectrum



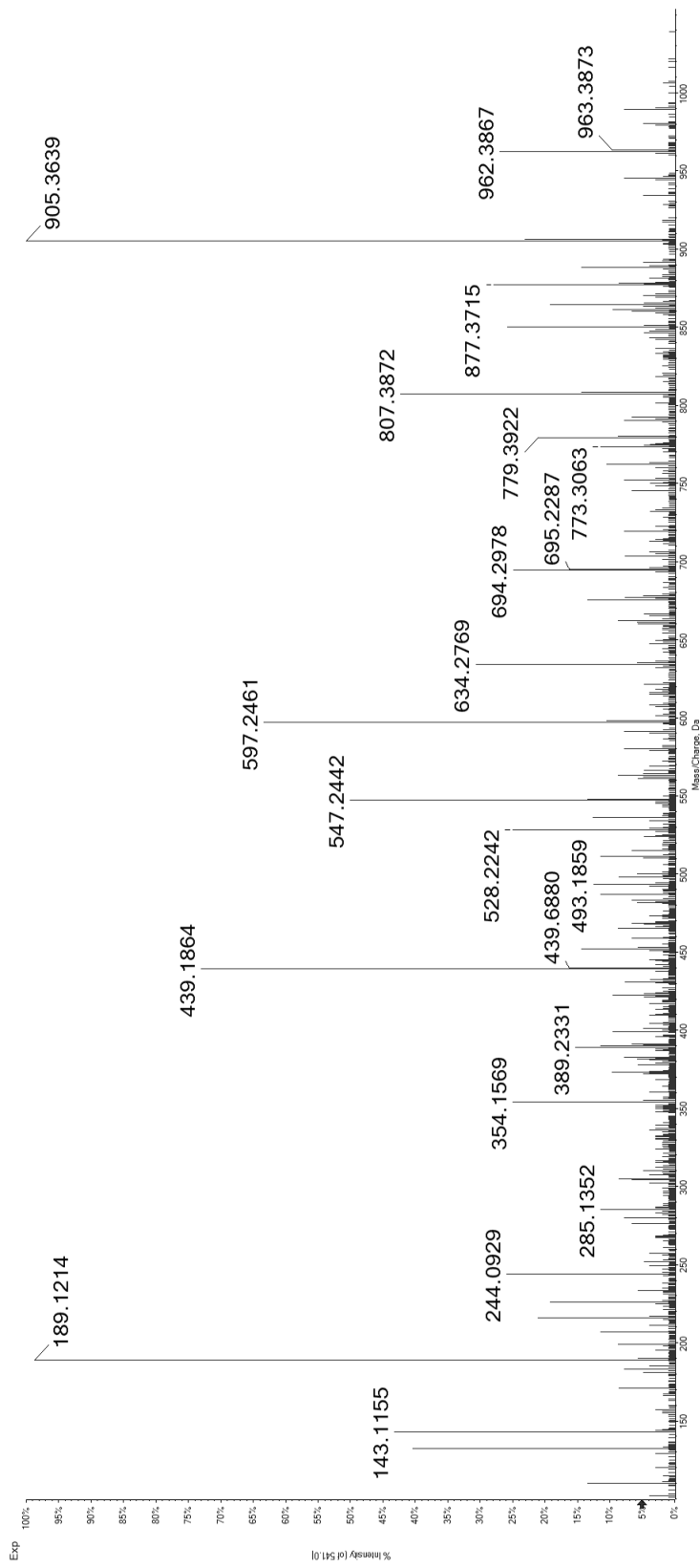
76

NEFH(pS)PIGL

Synthetic Peptide



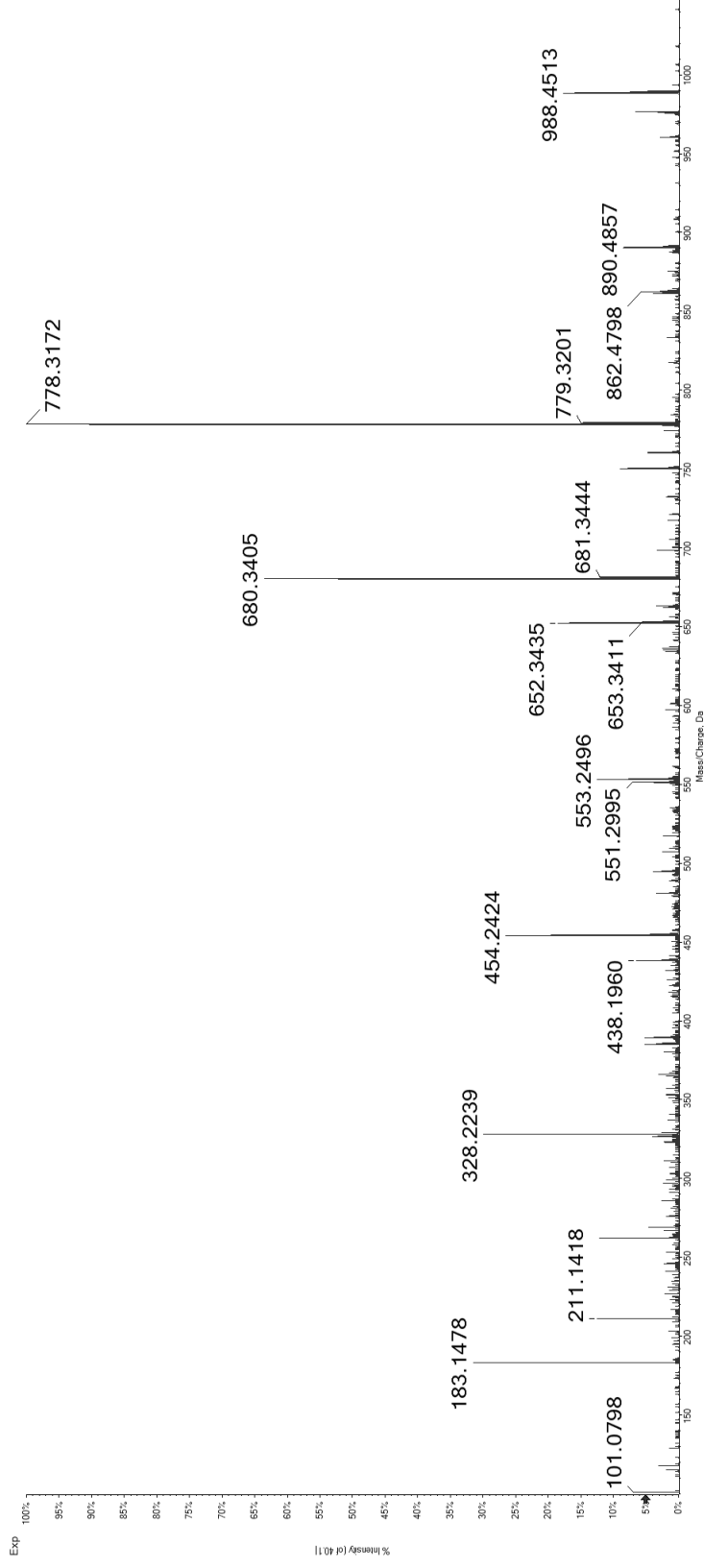
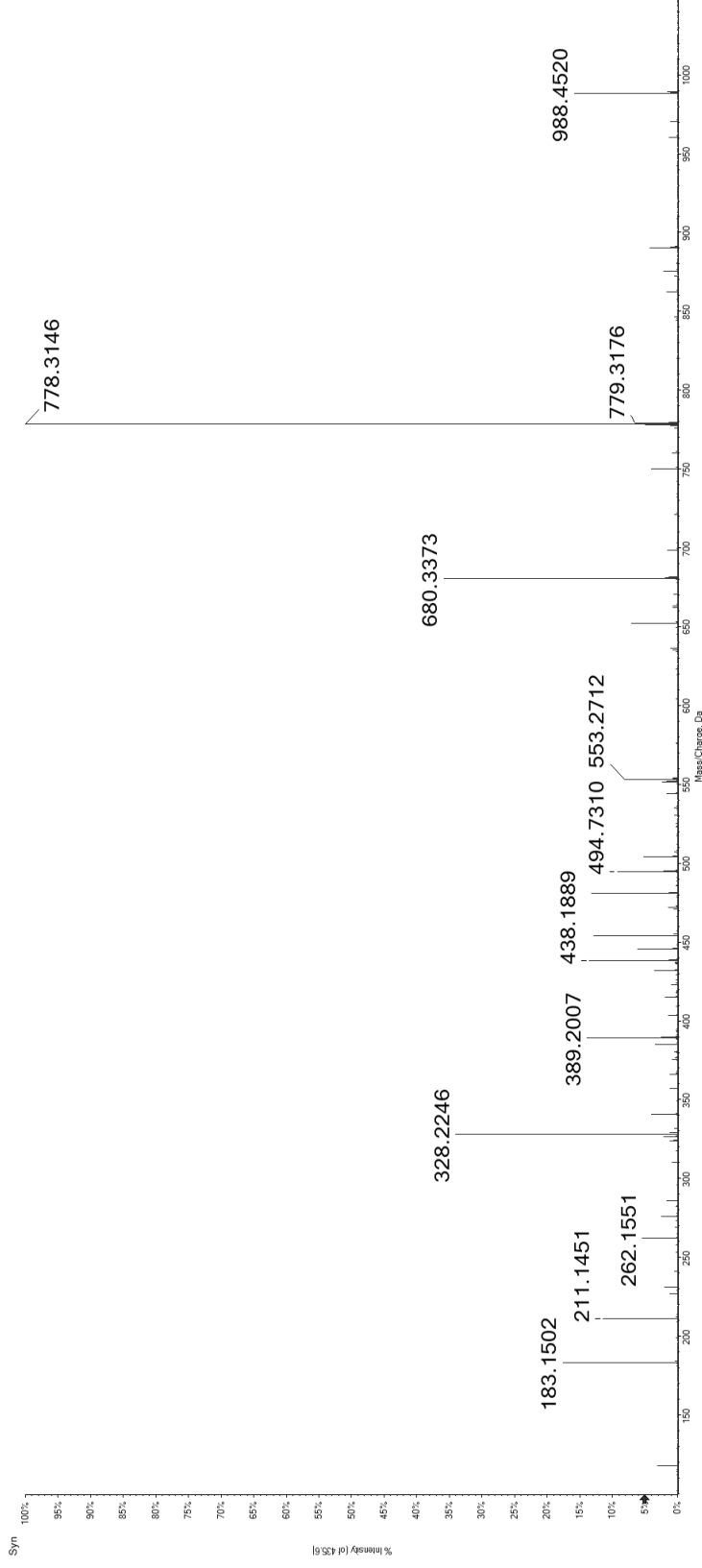
Candidate Spectrum



77

REV(pS)PEPIV

Synthetic Peptide

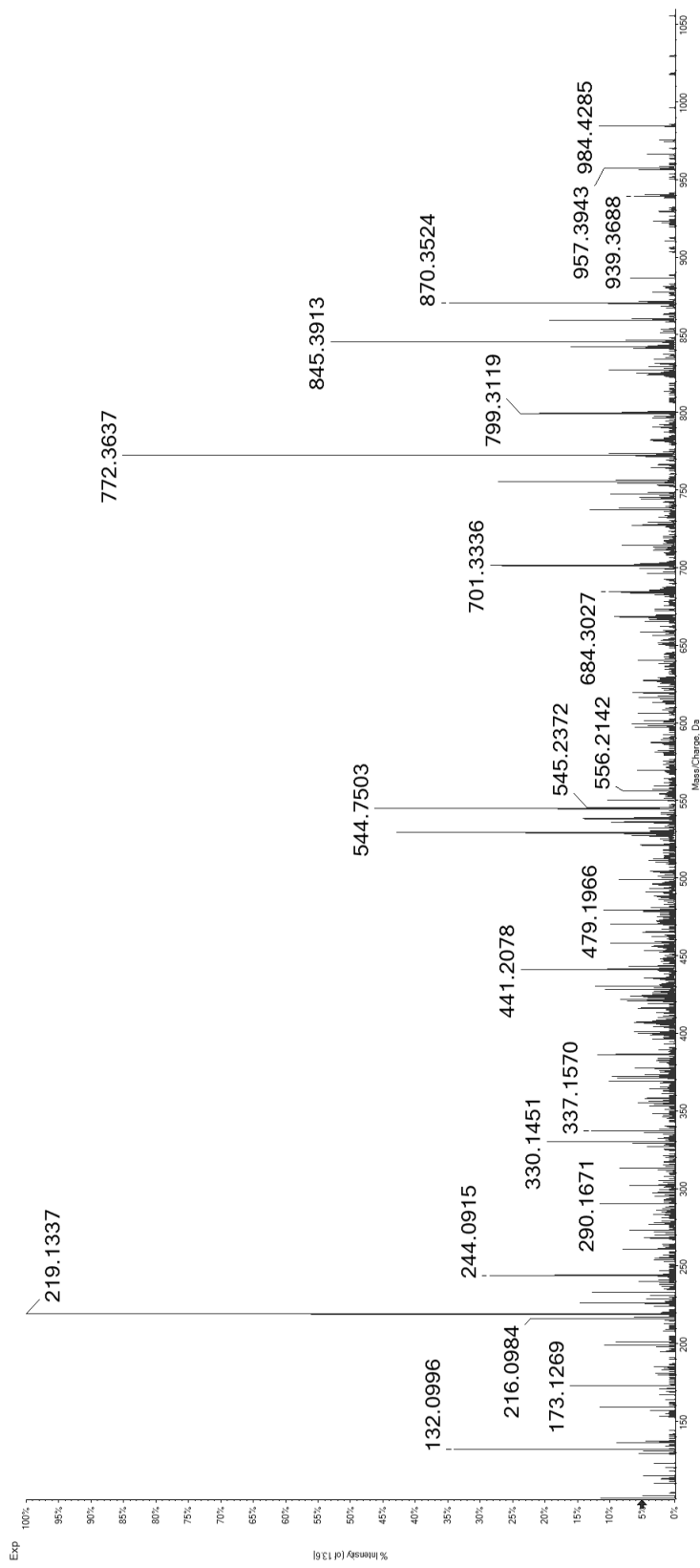
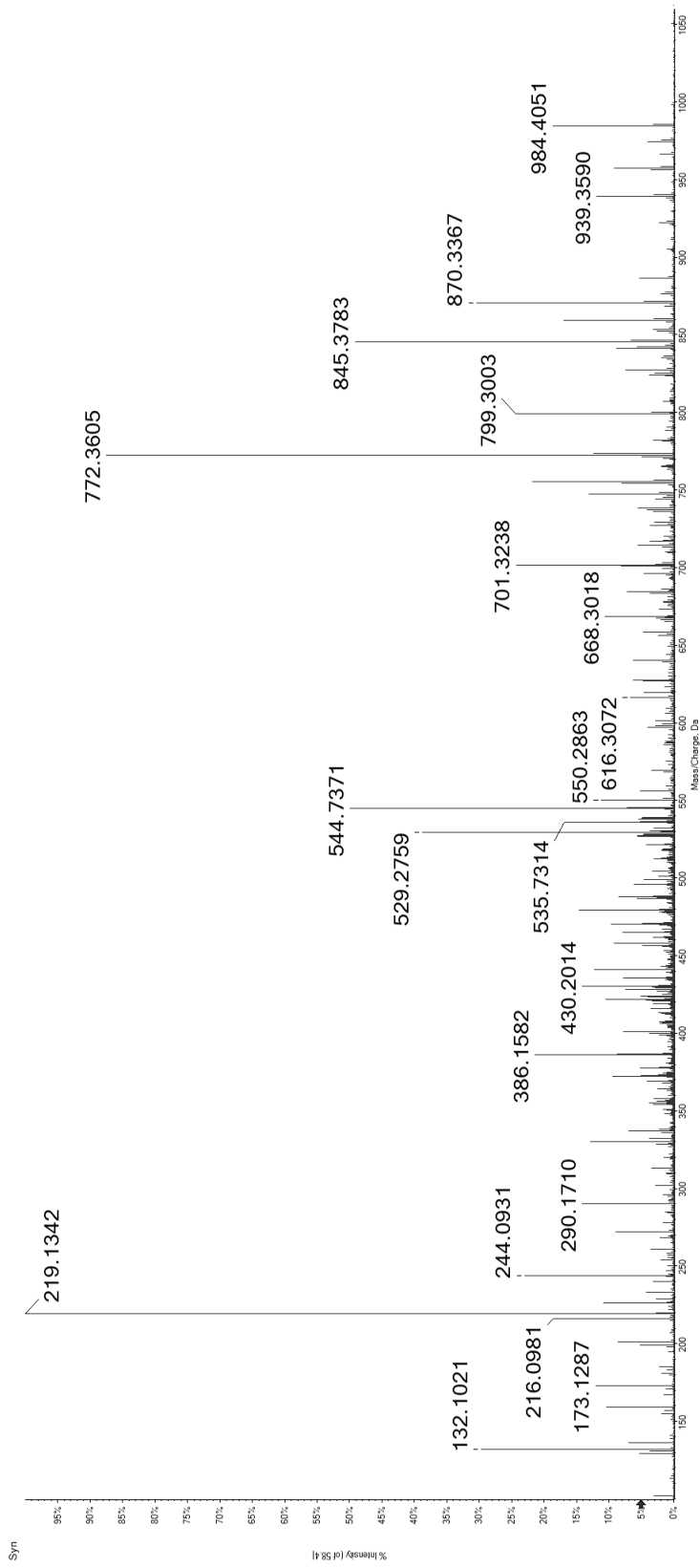


Candidate Spectrum

78

NEK(pS)PYASL

Synthetic Peptide

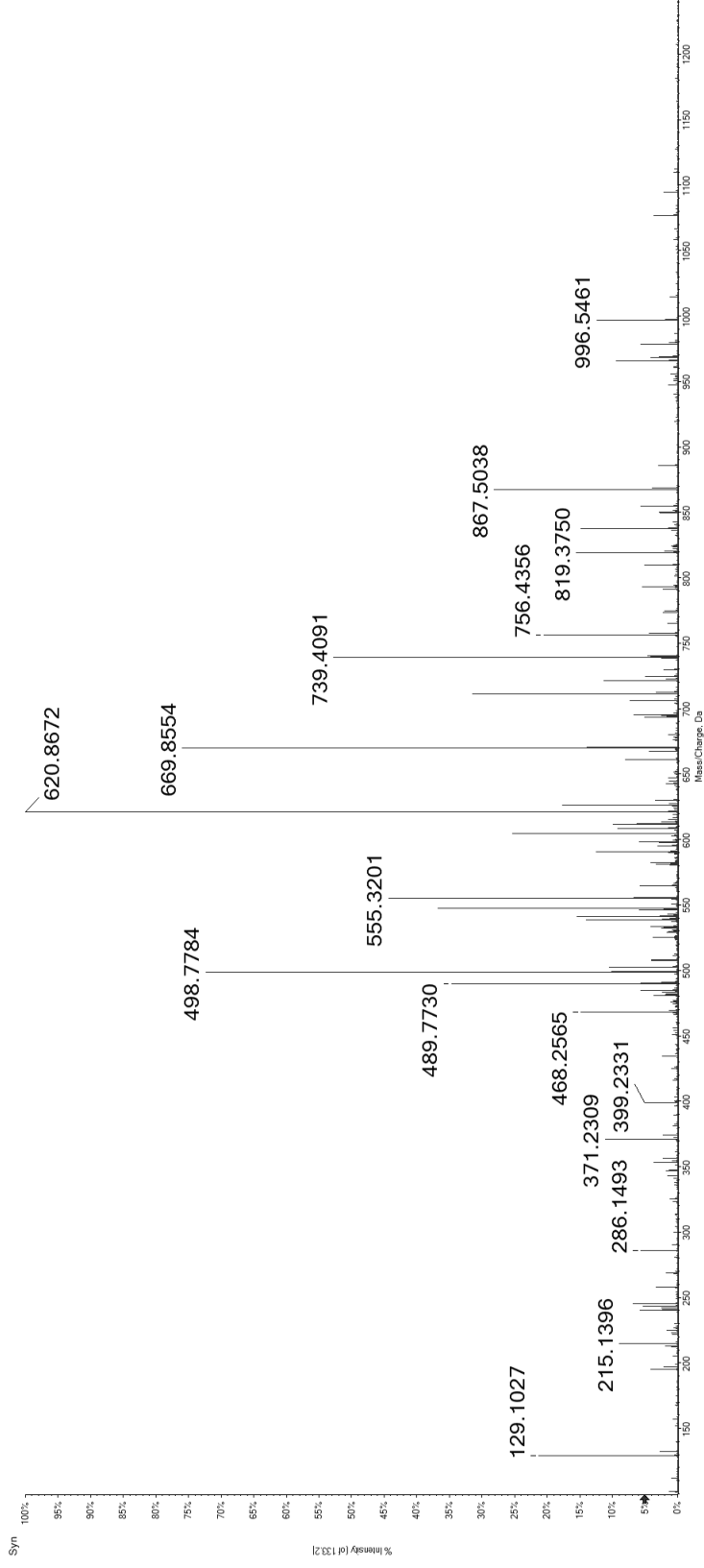


Candidate Spectrum

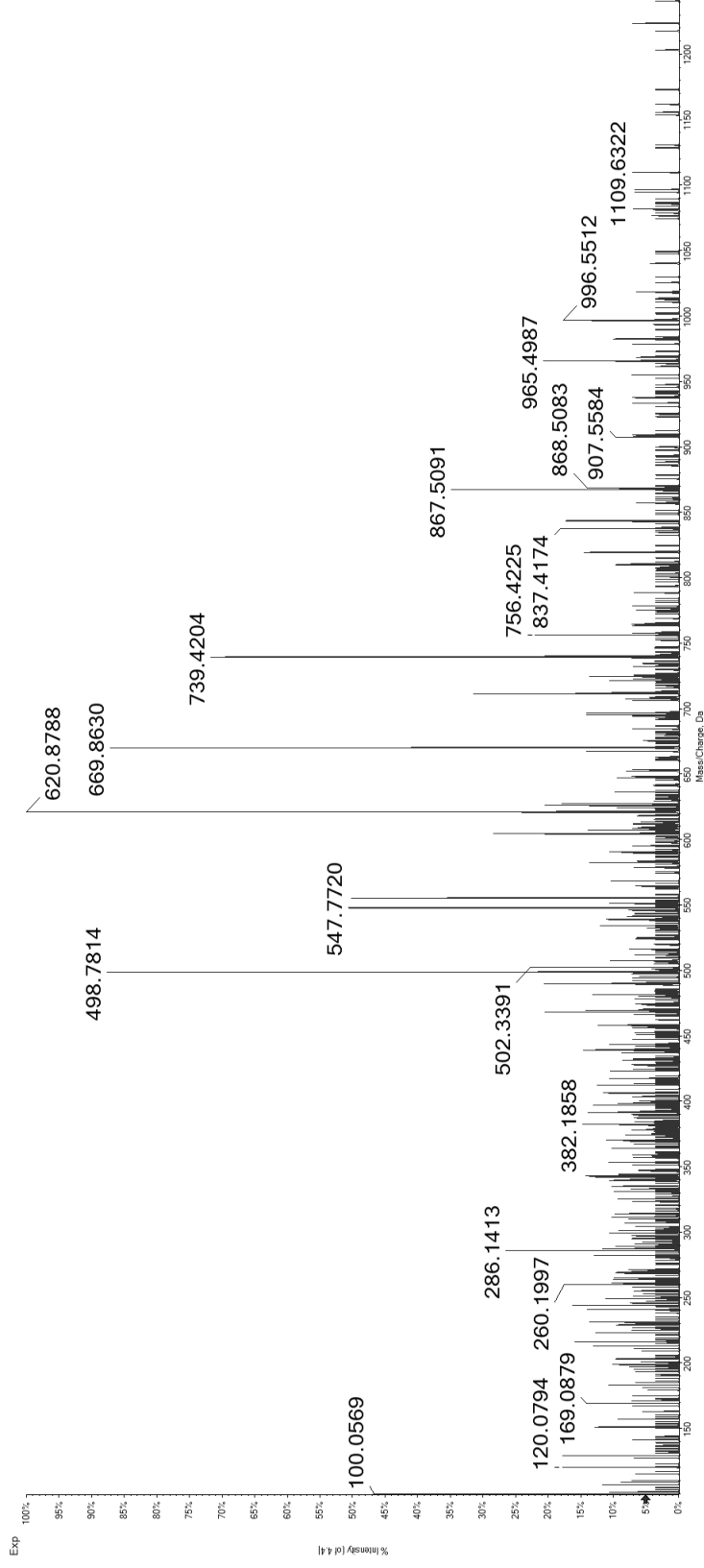
79

REL(ps)GTIKEIL

Synthetic Peptide

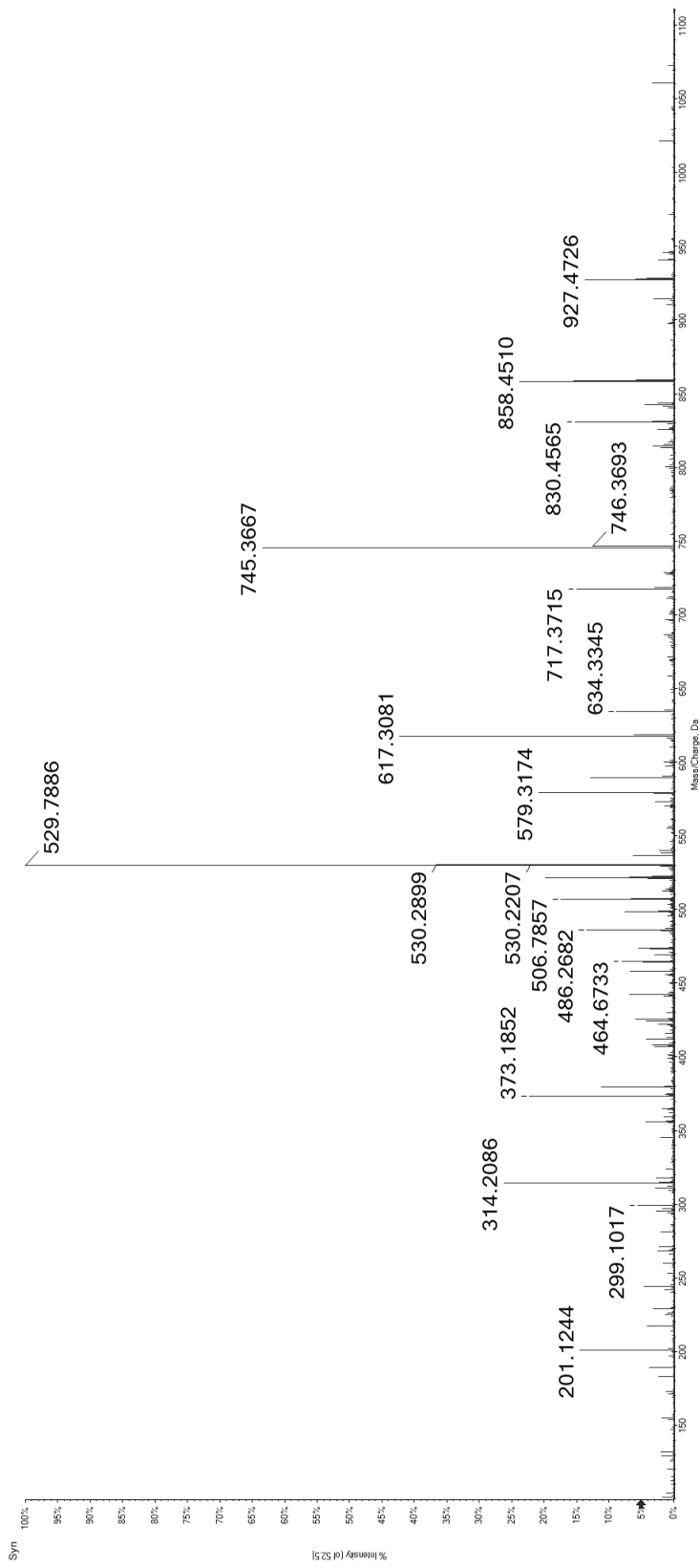


Candidate Spectrum

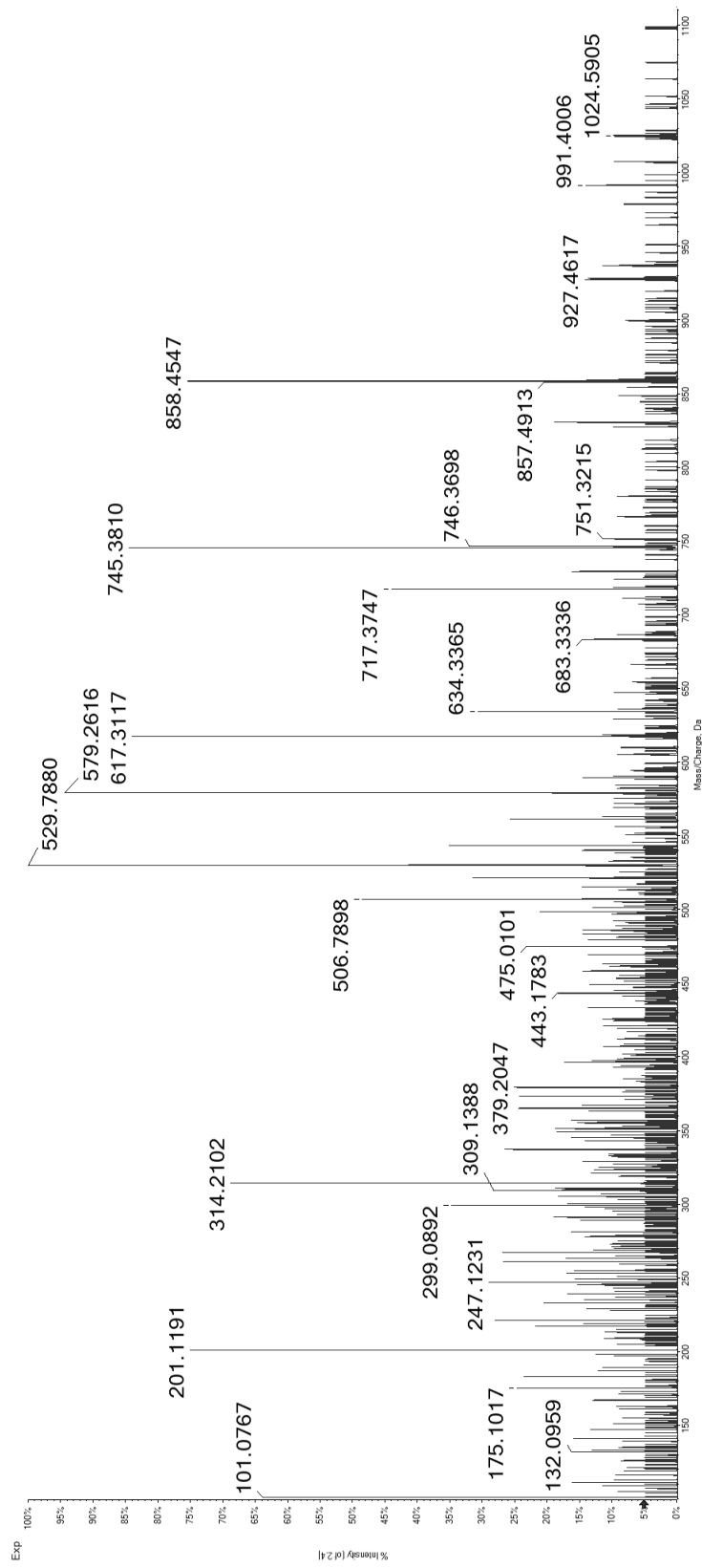


80 SERIMQL(ps)L

Synthetic
Peptide



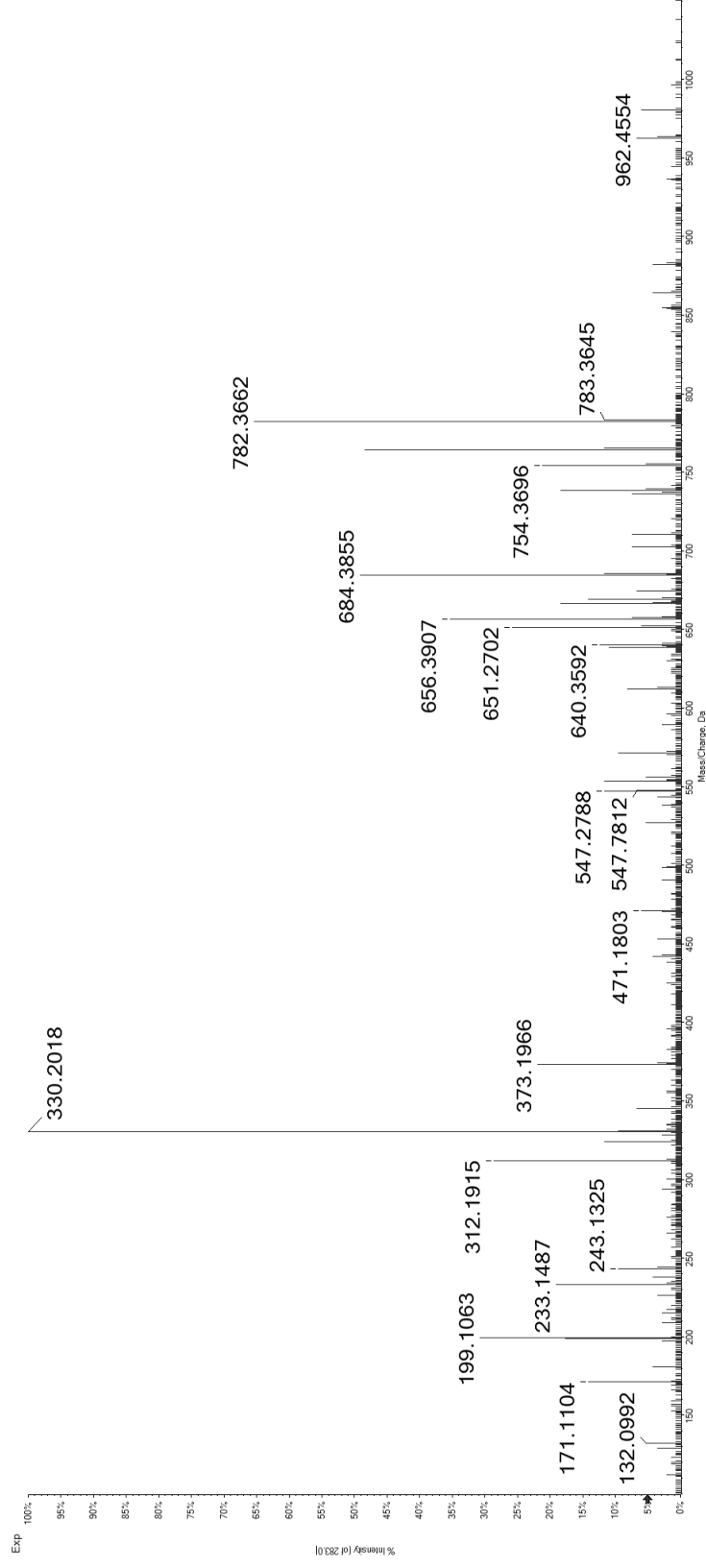
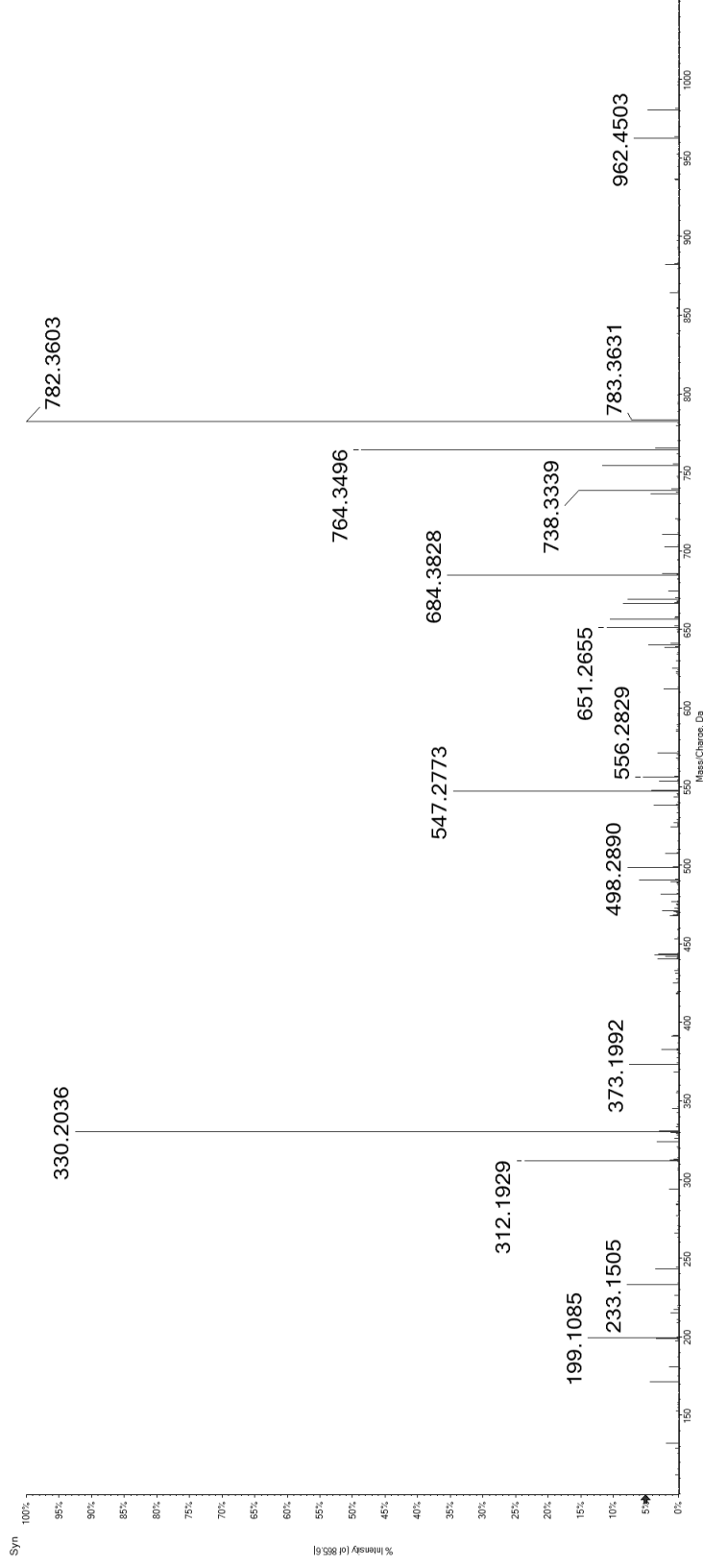
Candidate
Spectrum



81

R(pS)FPTLPTL

Synthetic Peptide

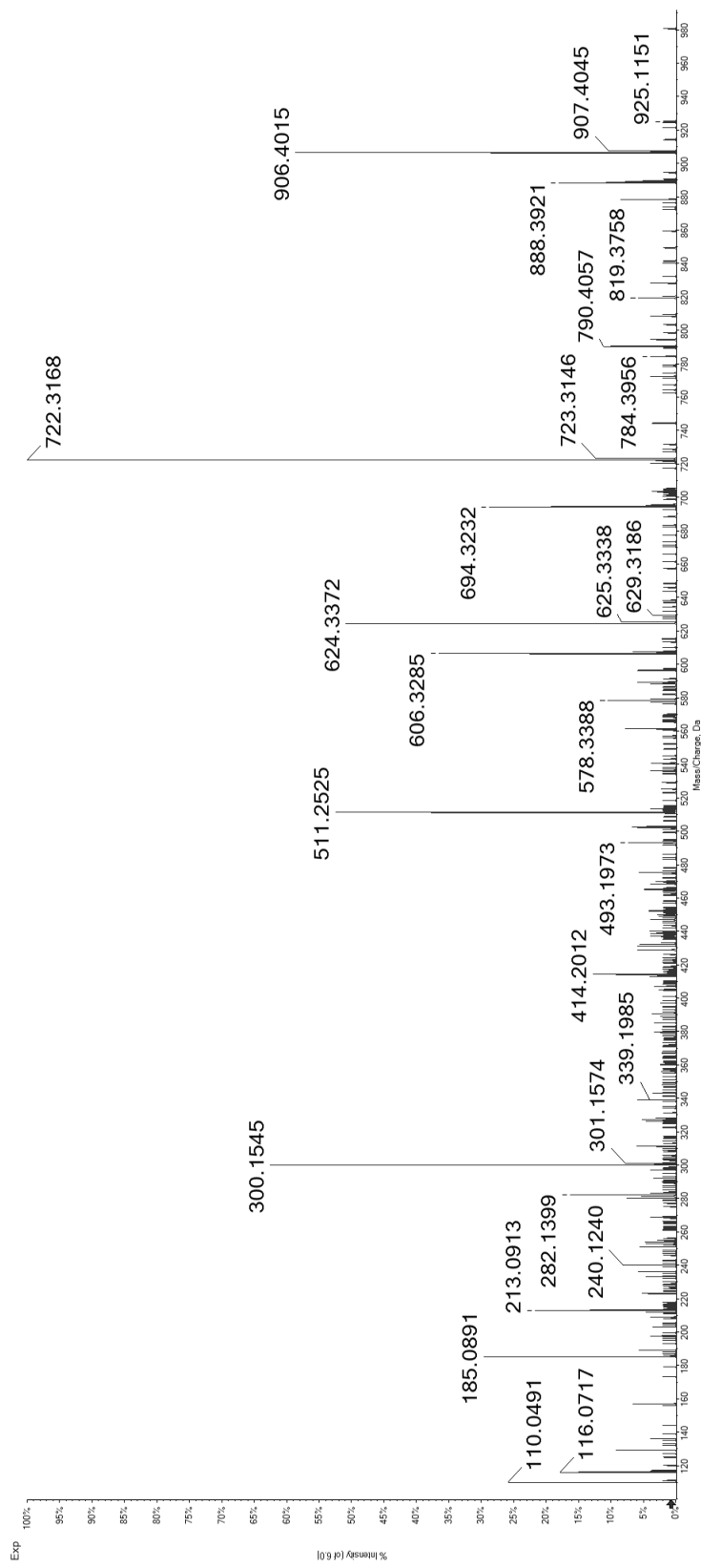
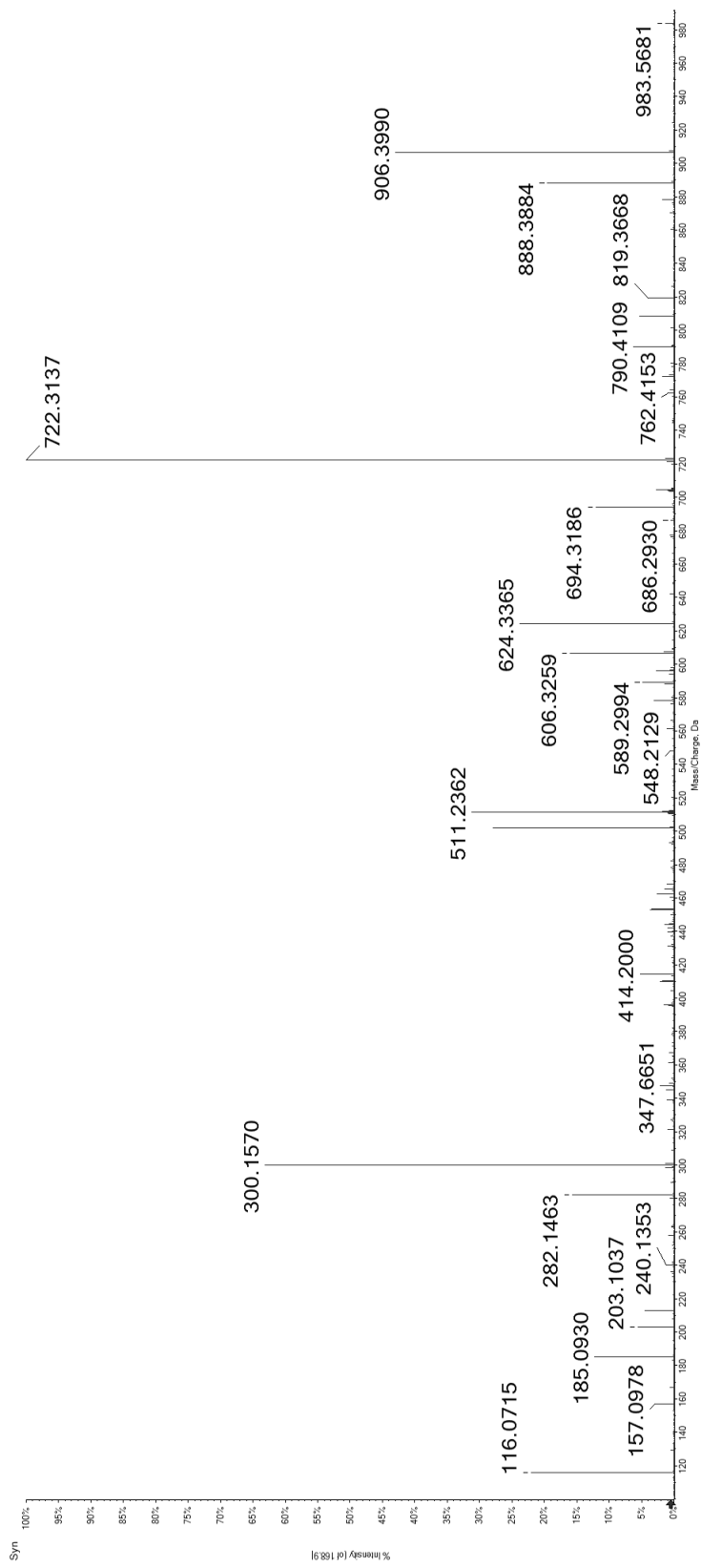


Candidate Spectrum

82

KES(ps)PIPSP

Synthetic Peptide

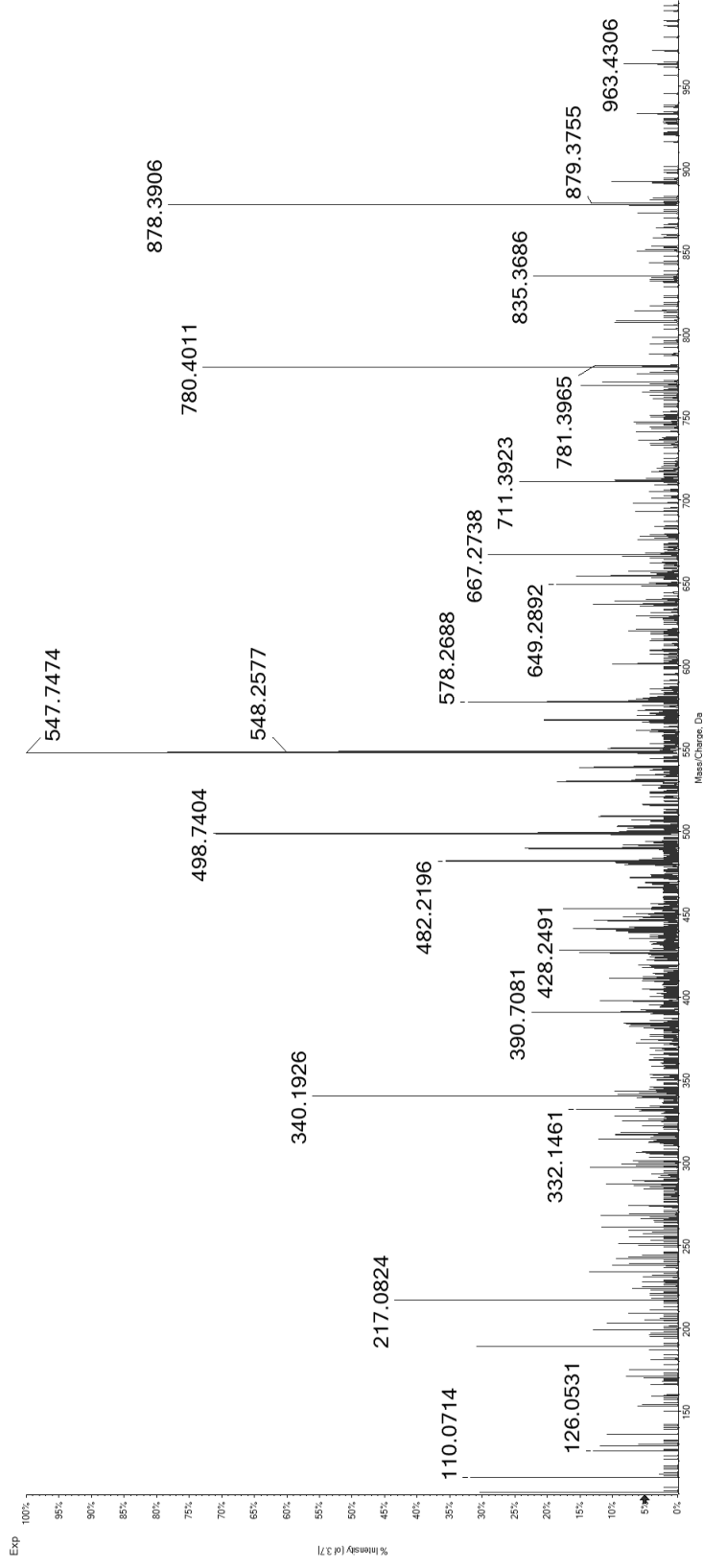
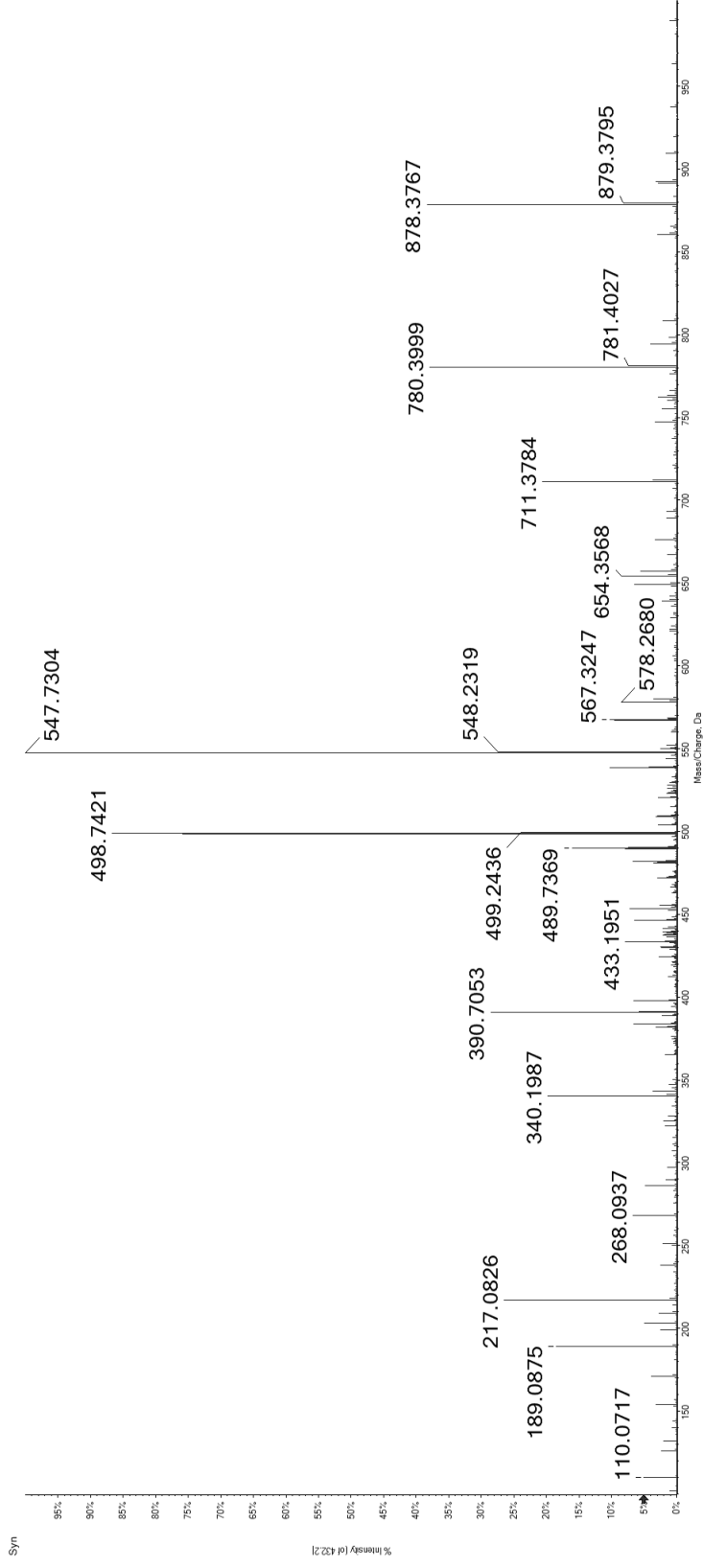


Candidate Spectrum

83

SE(ps)GSLHAL

Synthetic Peptide

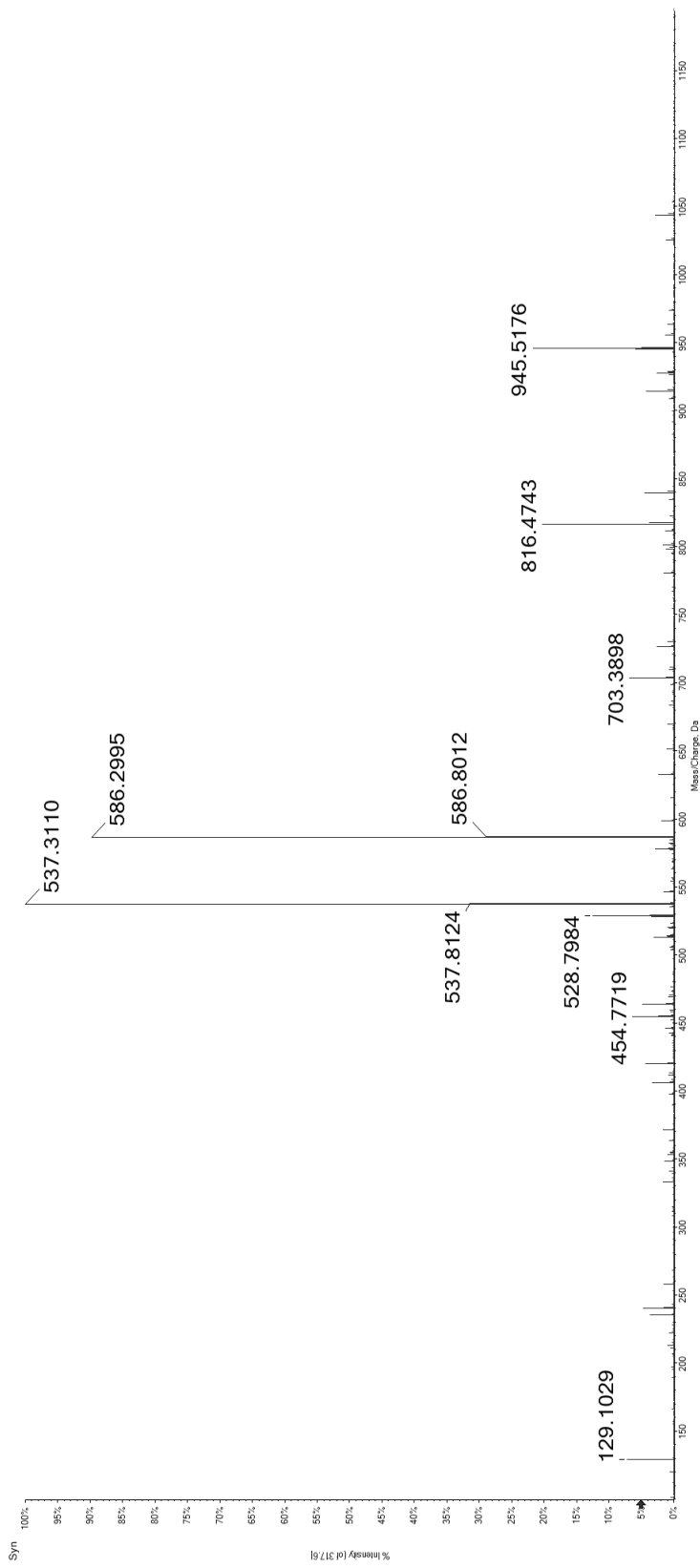


Candidate Spectrum

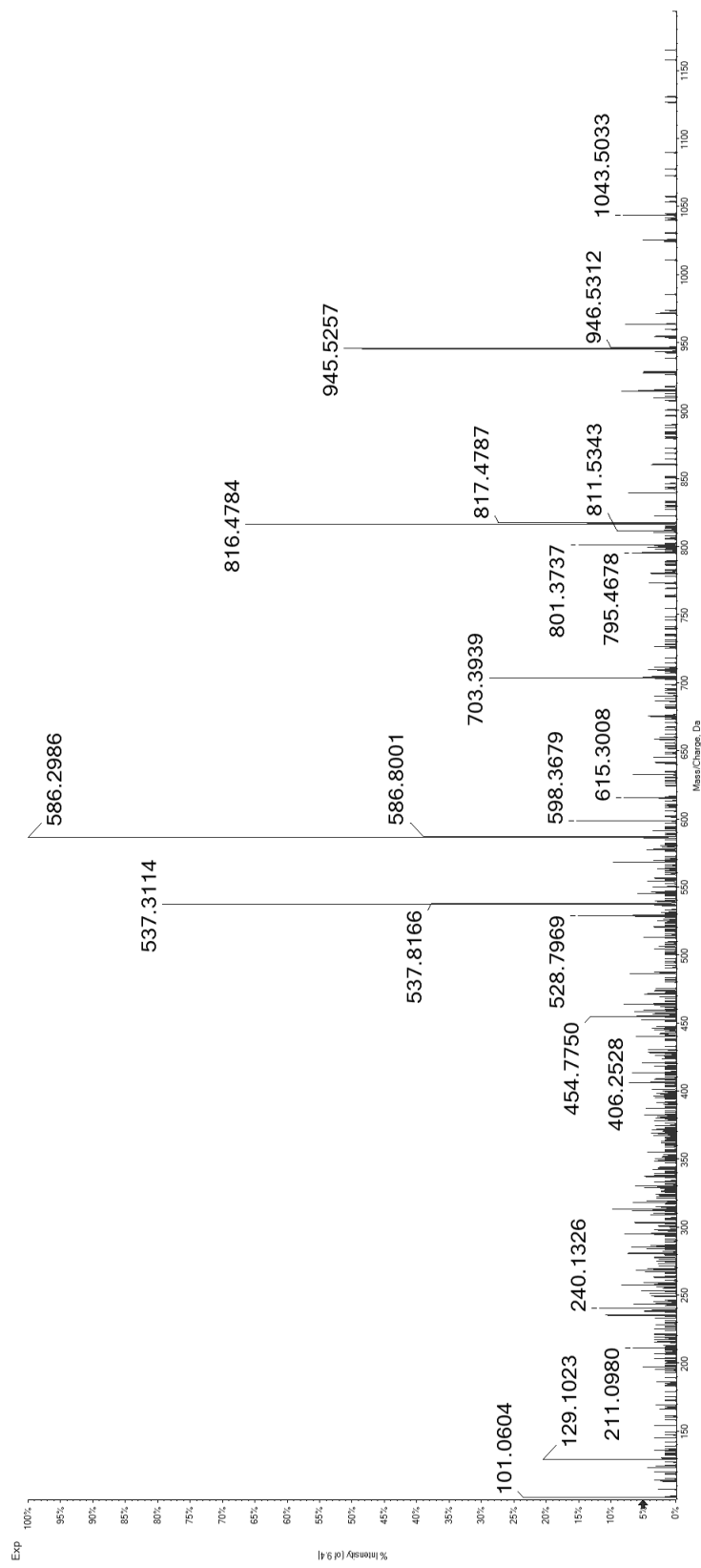
84

KELARQI(ps)F

Synthetic Peptide



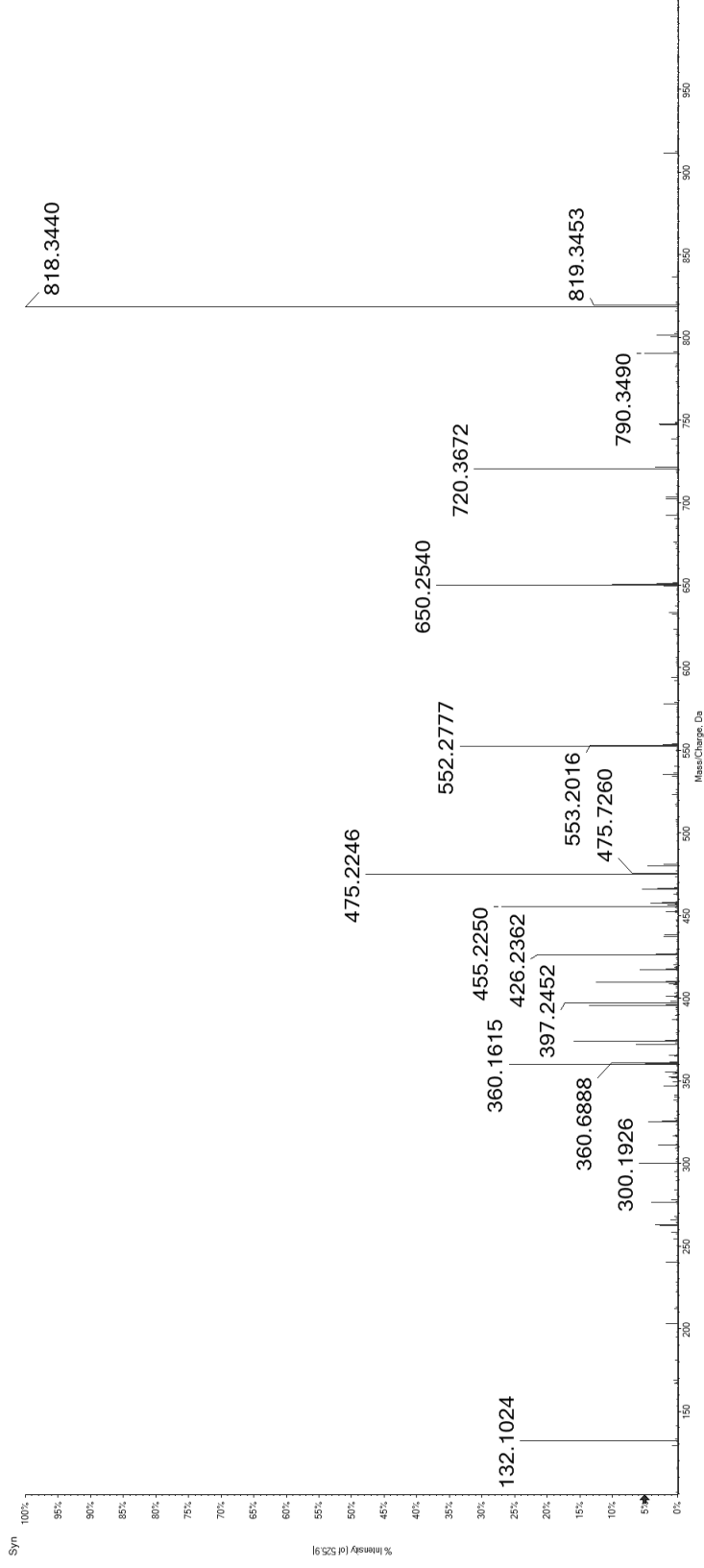
Candidate Spectrum



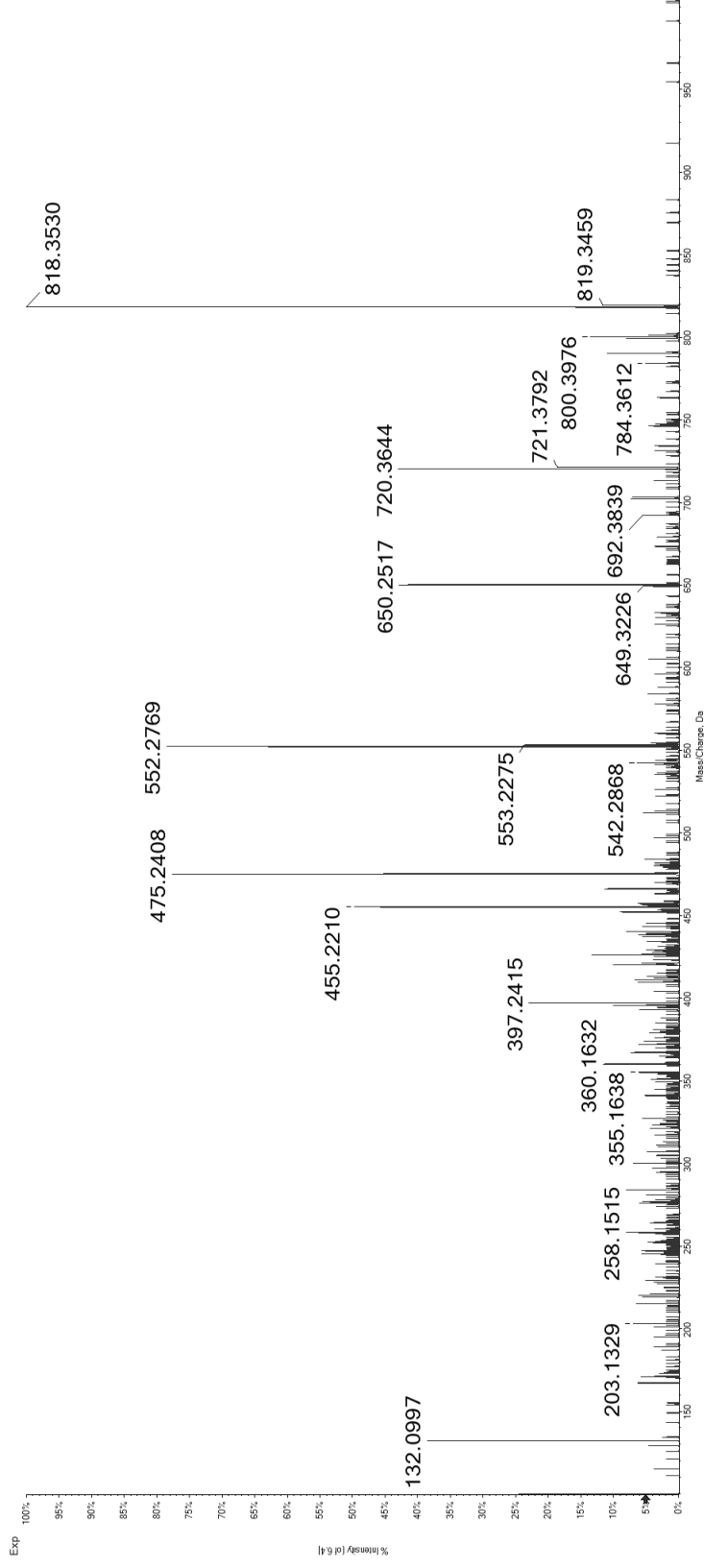
85

KEN(pT)PPAL

Synthetic Peptide



Candidate Spectrum



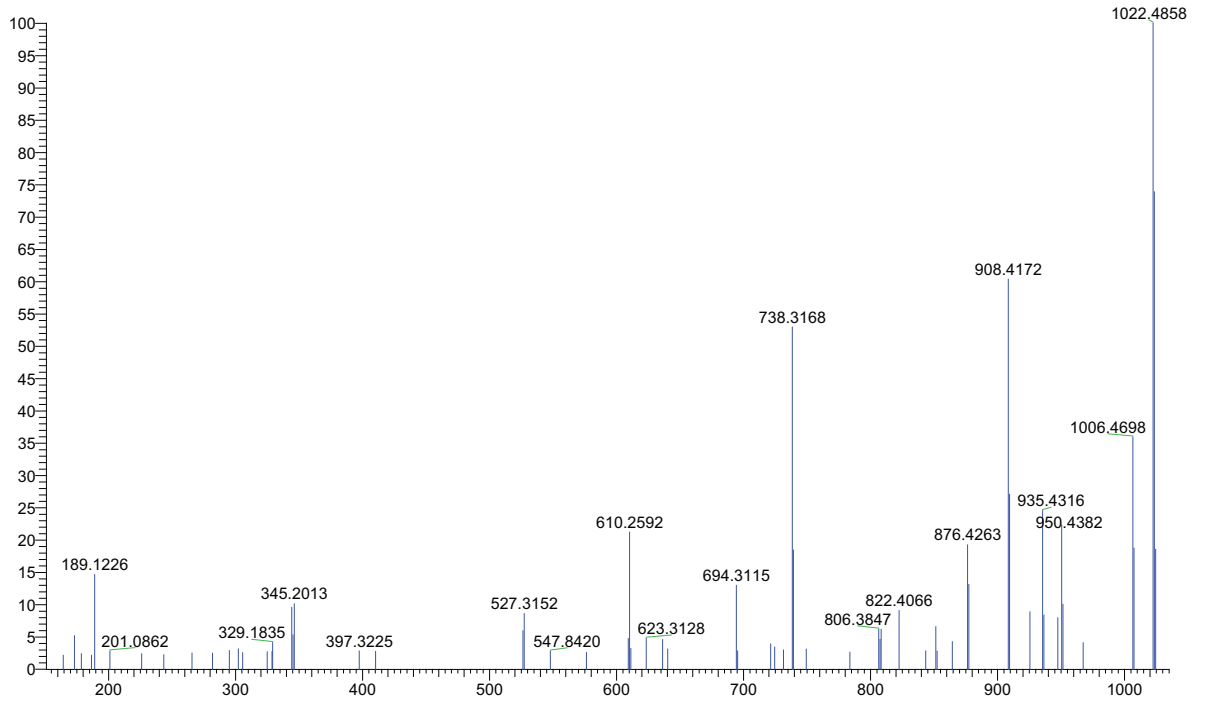
Anexo 1.2: Espectros MS2 de los fosfopéptidos identificados mediante EThcD.

Cada espectro MS2 candidato se compara con el espectro de fragmentación del péptido sintético correspondiente.

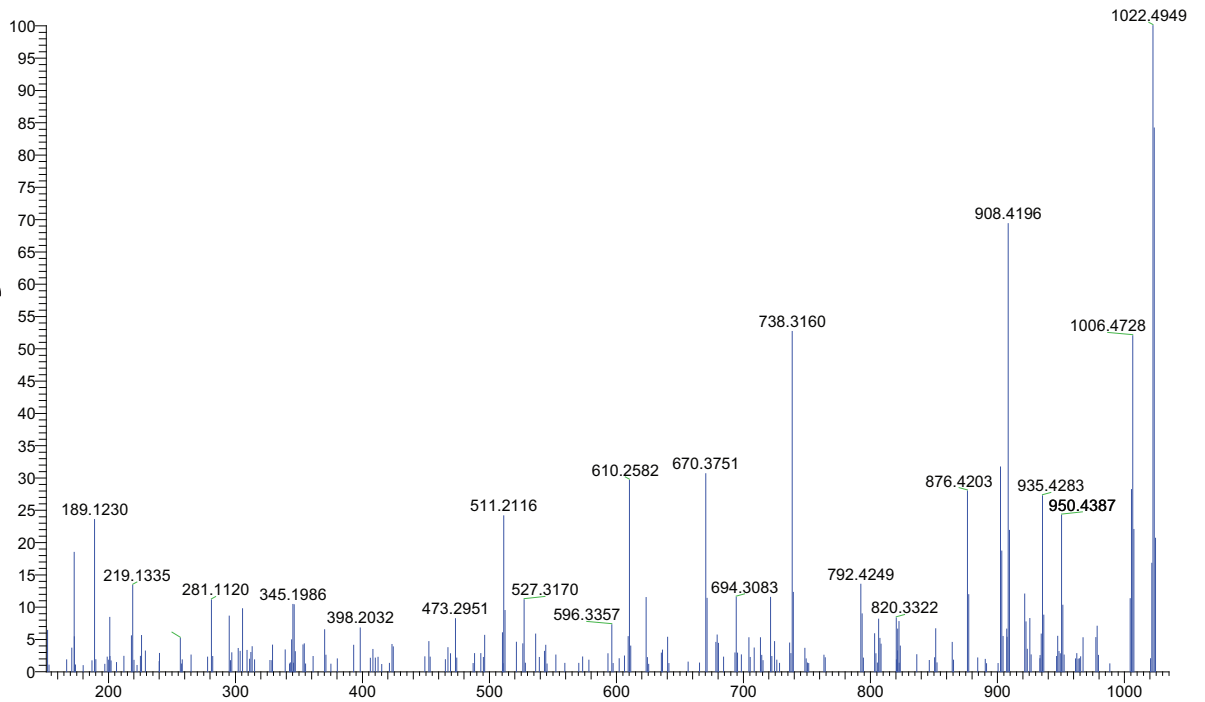
1

AEK(pS)PQLGL

Synthetic Peptide



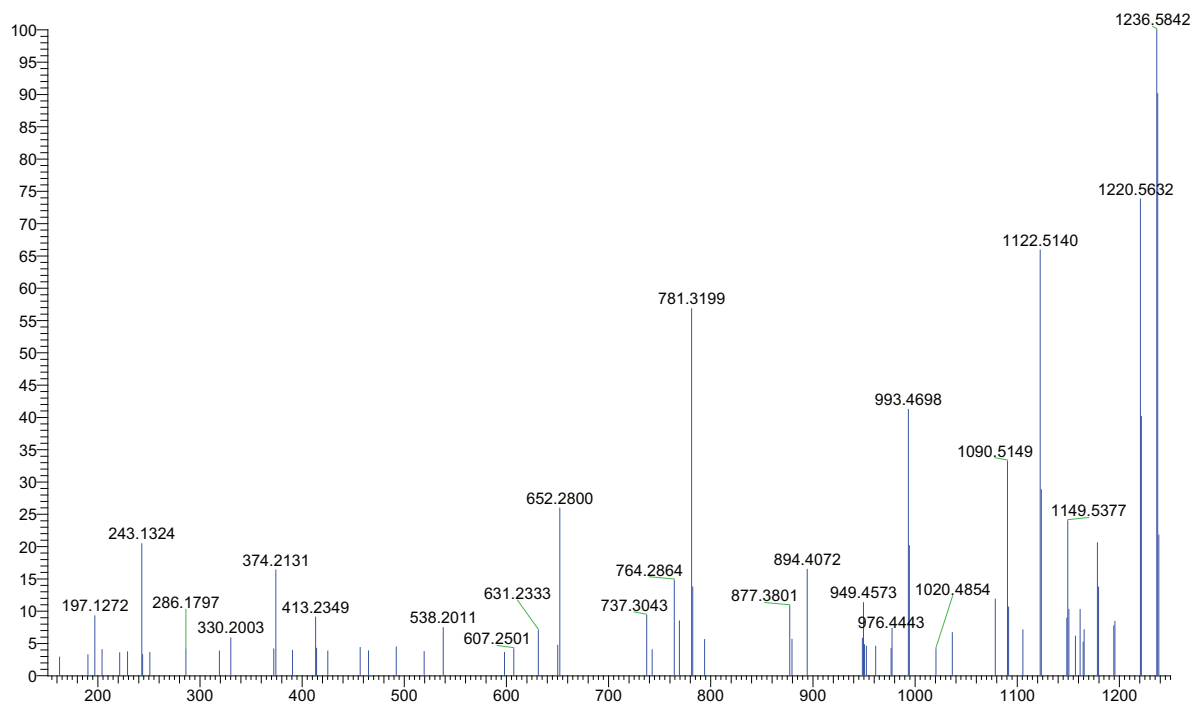
Candidate spectrum



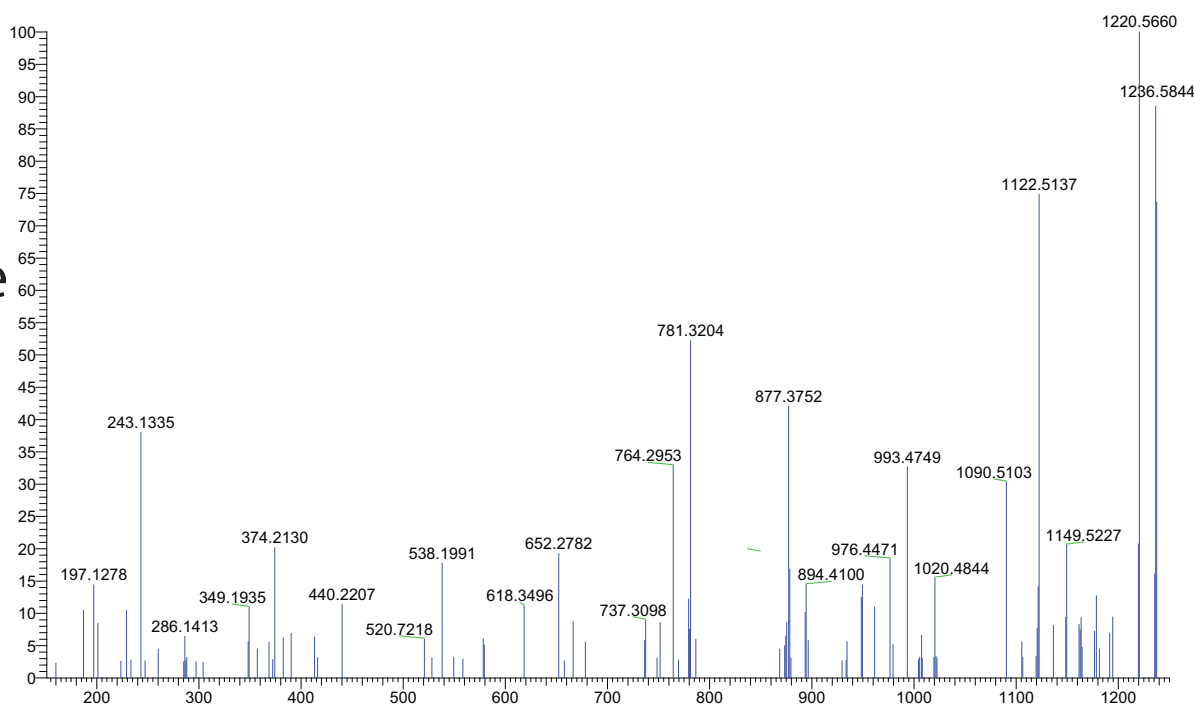
2

AER(pT)PELVEL

Synthetic Peptide



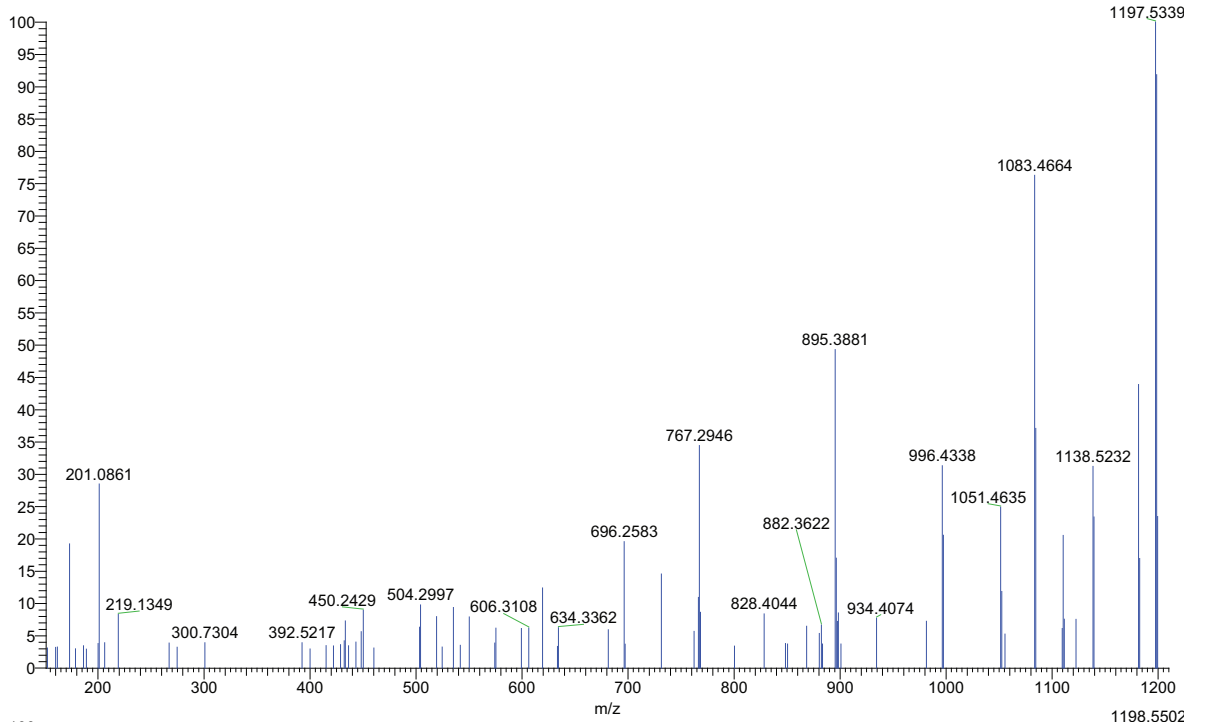
Candidate spectrum



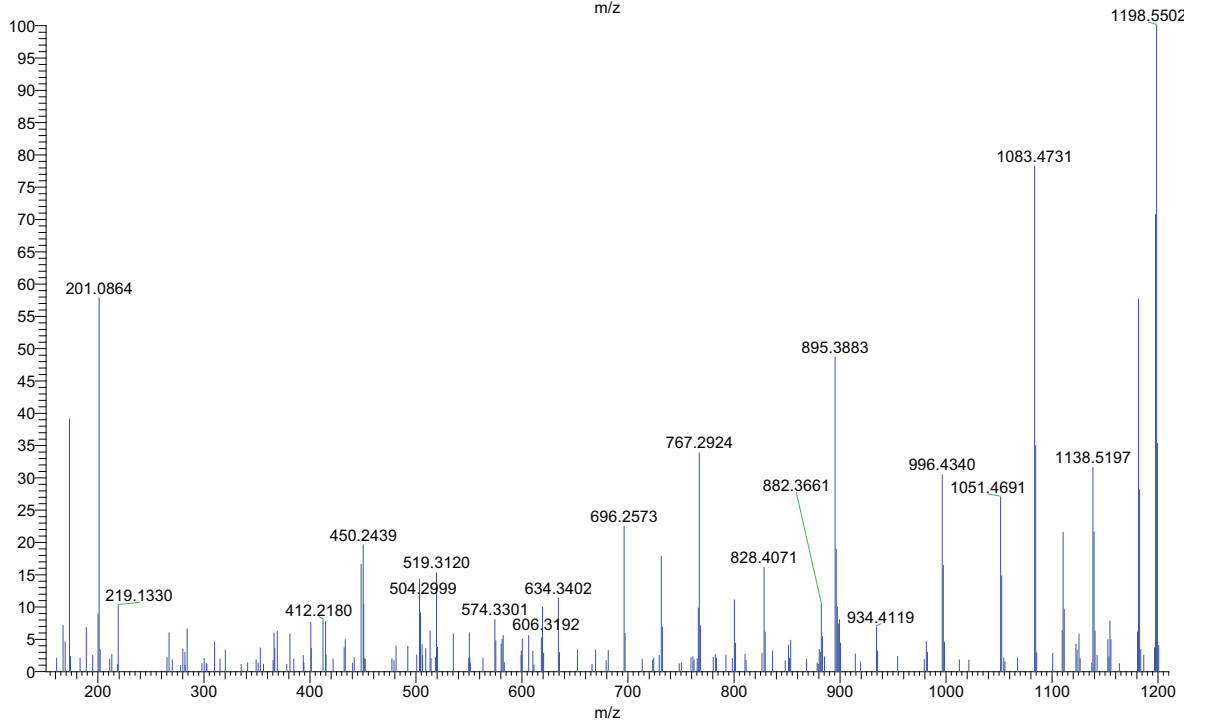
3

AEV(pS)PDAKTSL

Synthetic Peptide



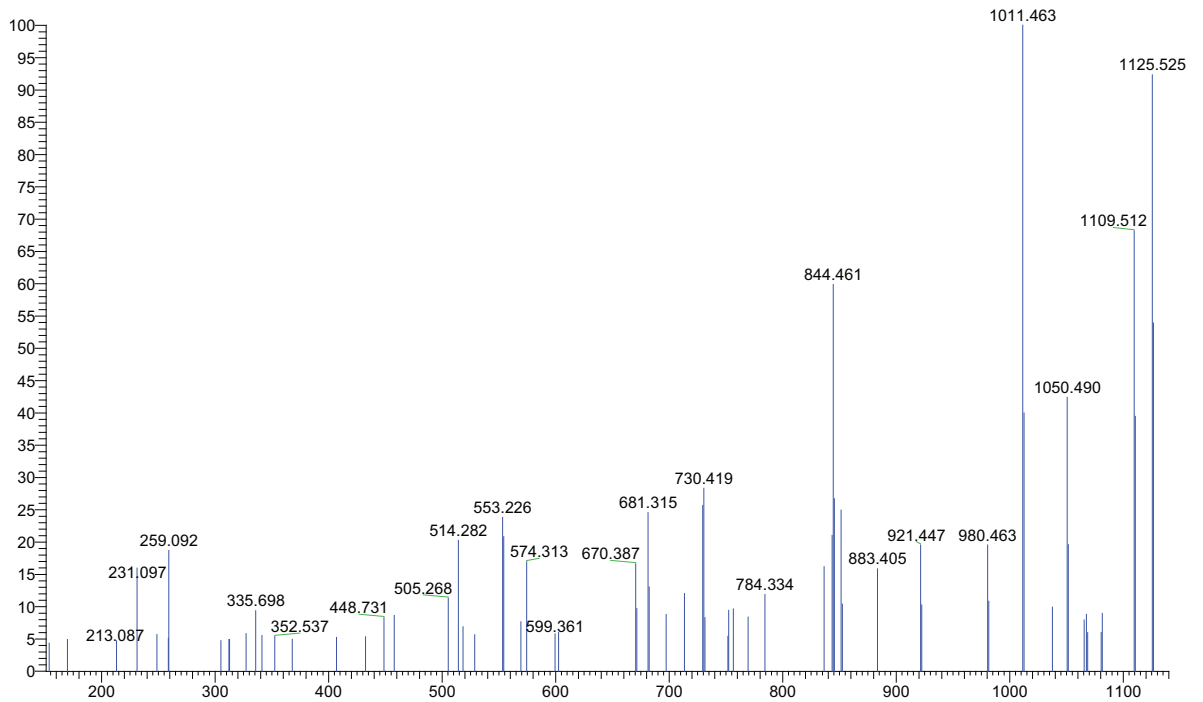
Candidate spectrum



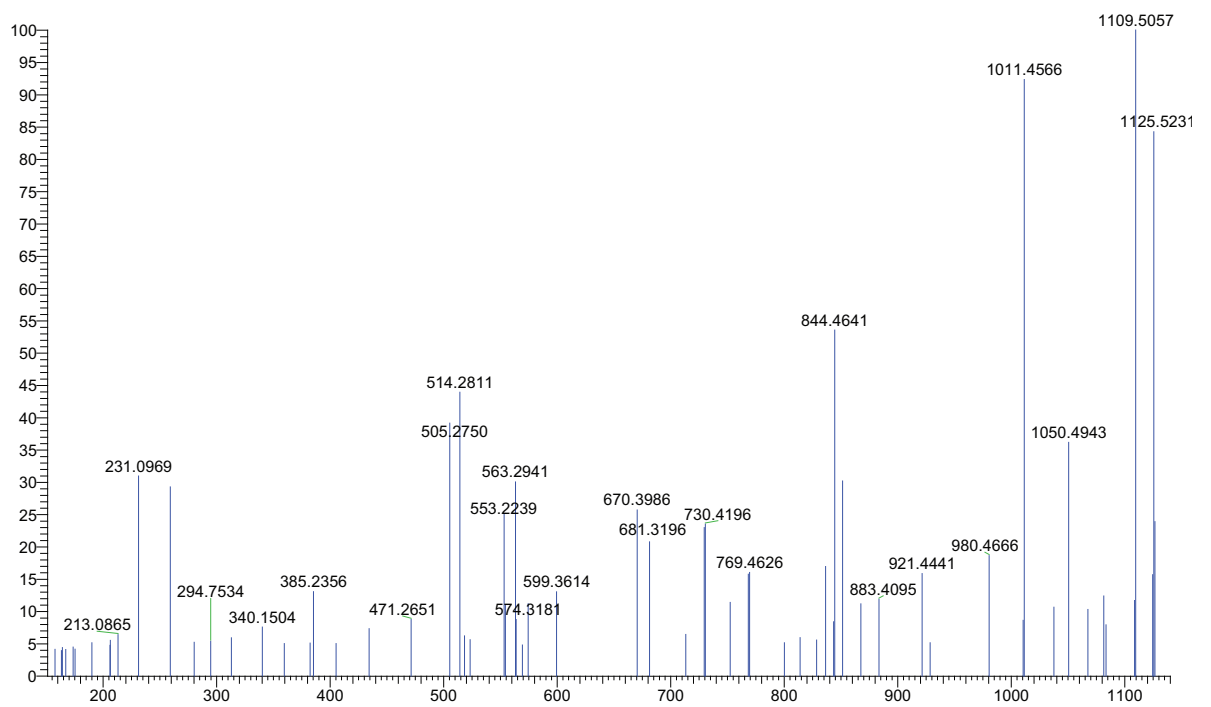
4

EEVAKRN(pS)L

Synthetic Peptide



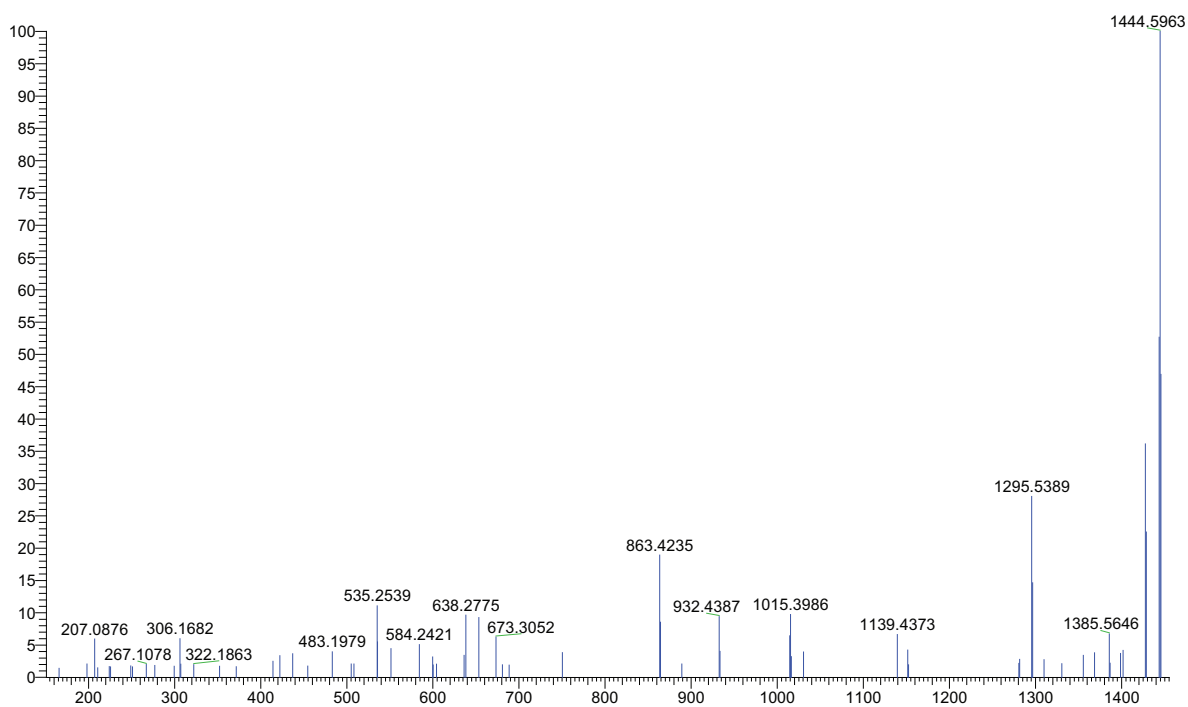
Candidate spectrum



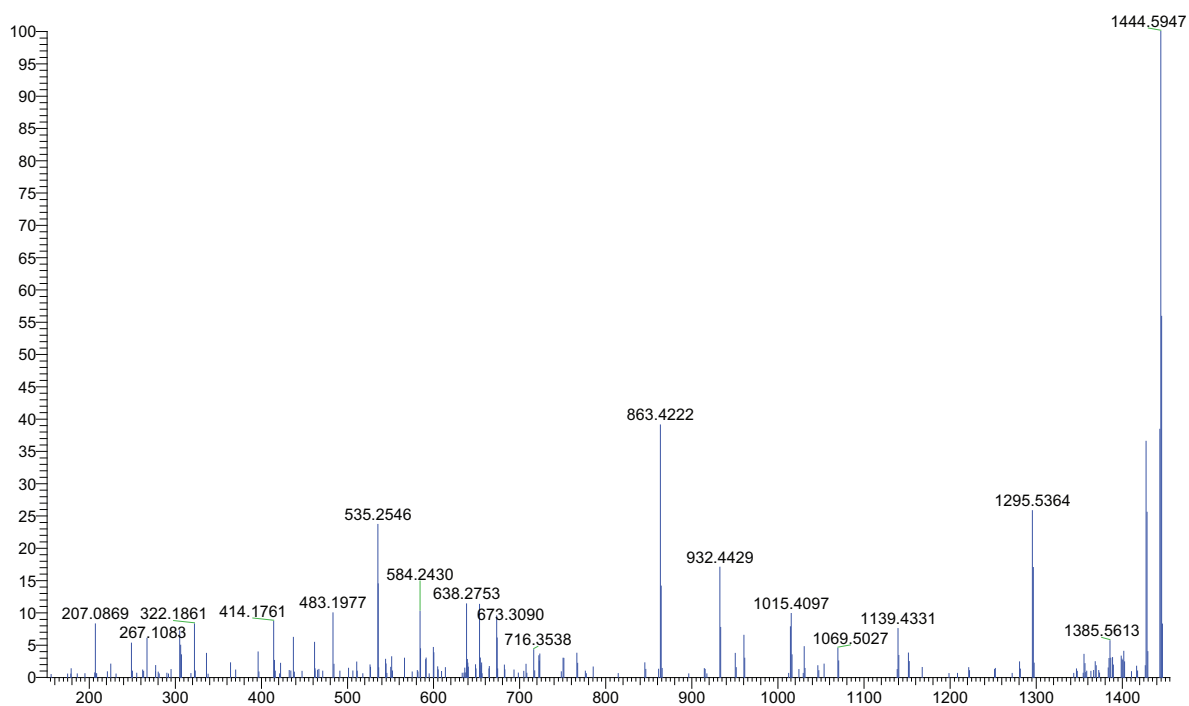
5

FEH(pS)PLSEDRF

Synthetic Peptide



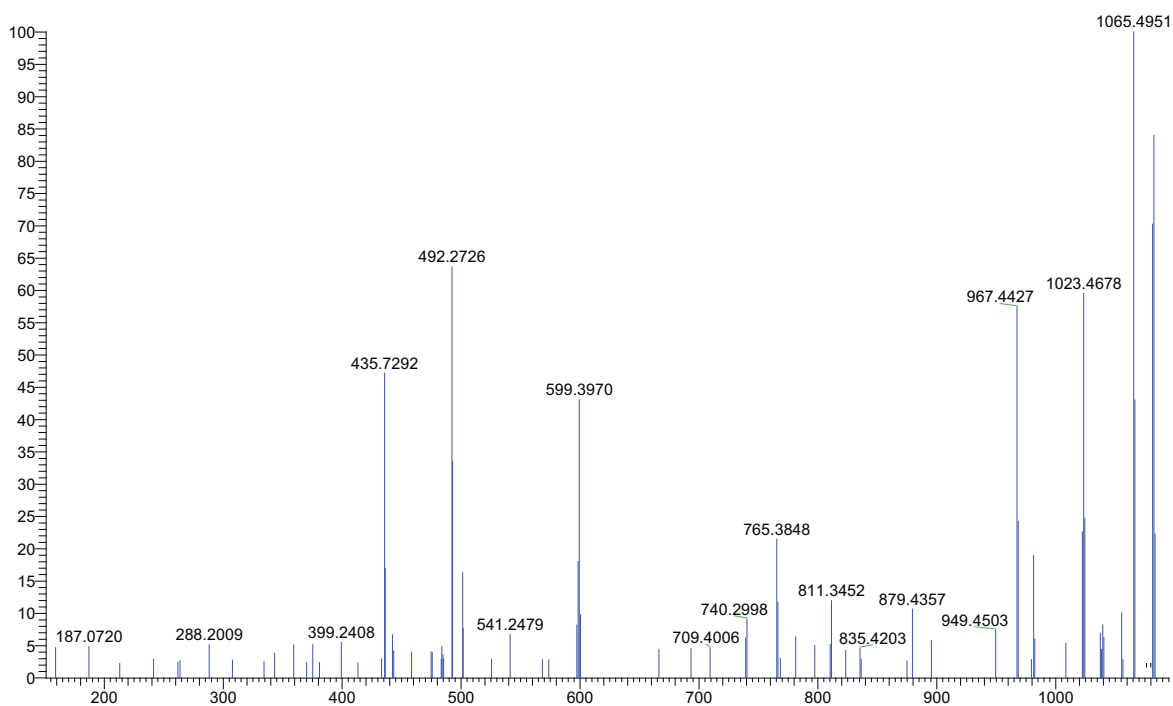
Candidate spectrum



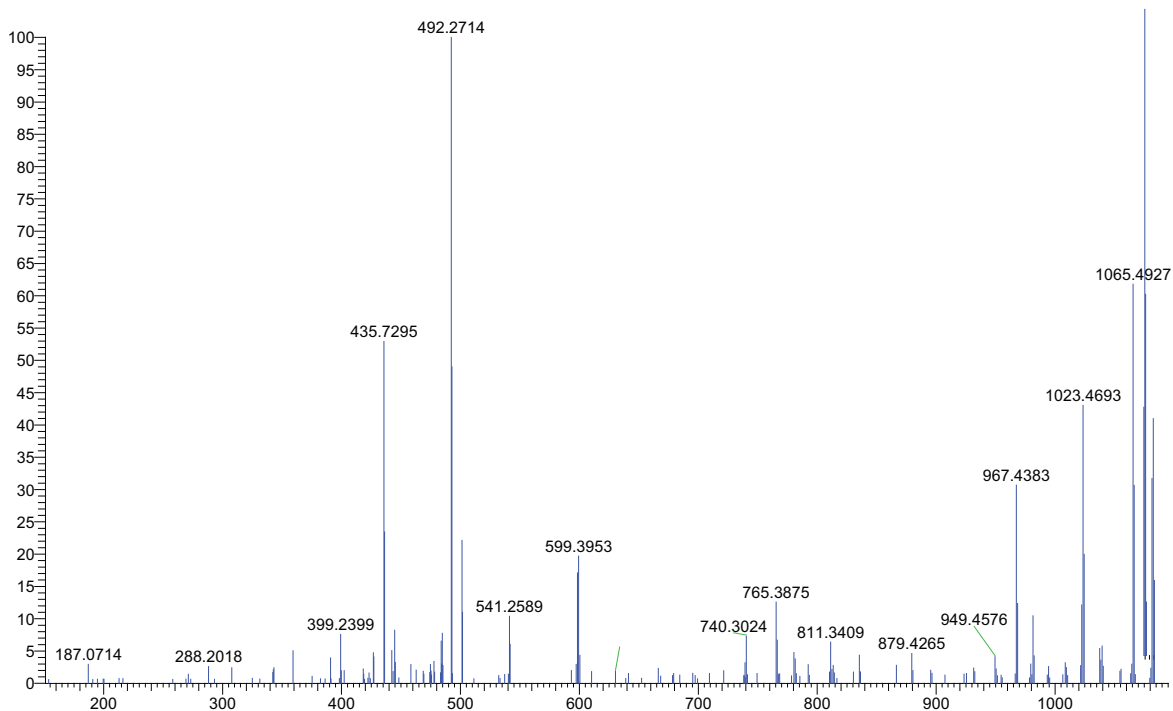
6

GEN(pS)RVARL

Synthetic Peptide



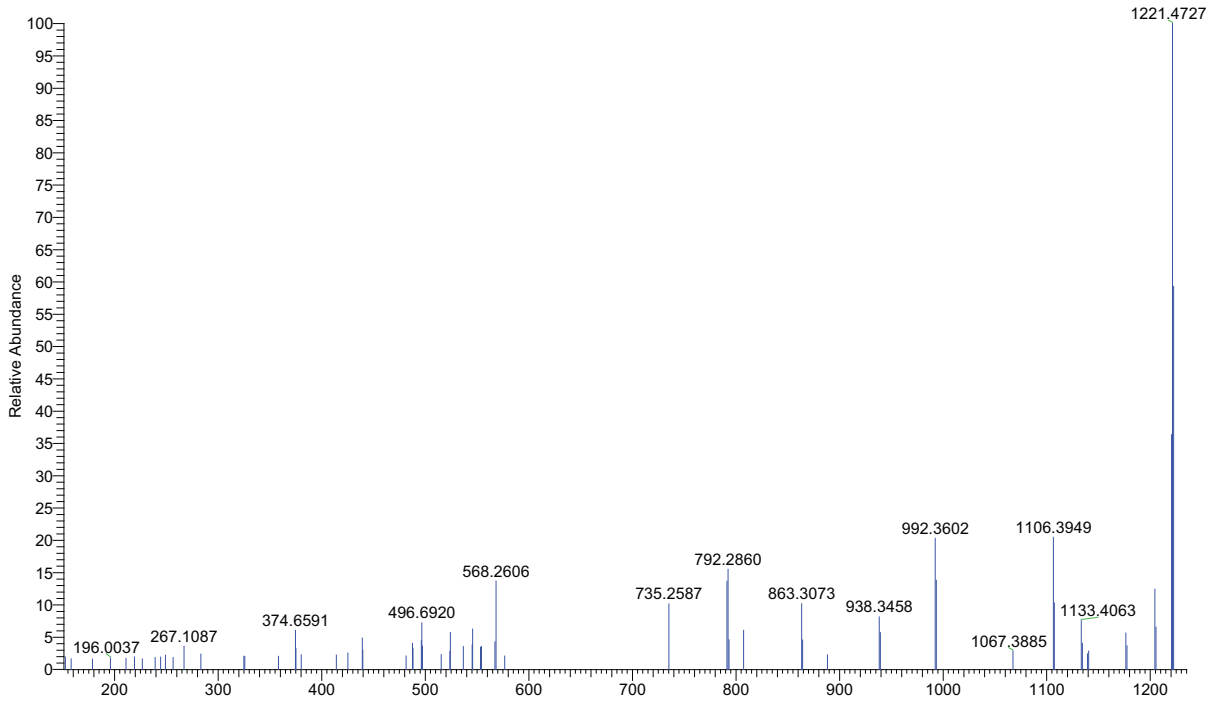
Candidate spectrum



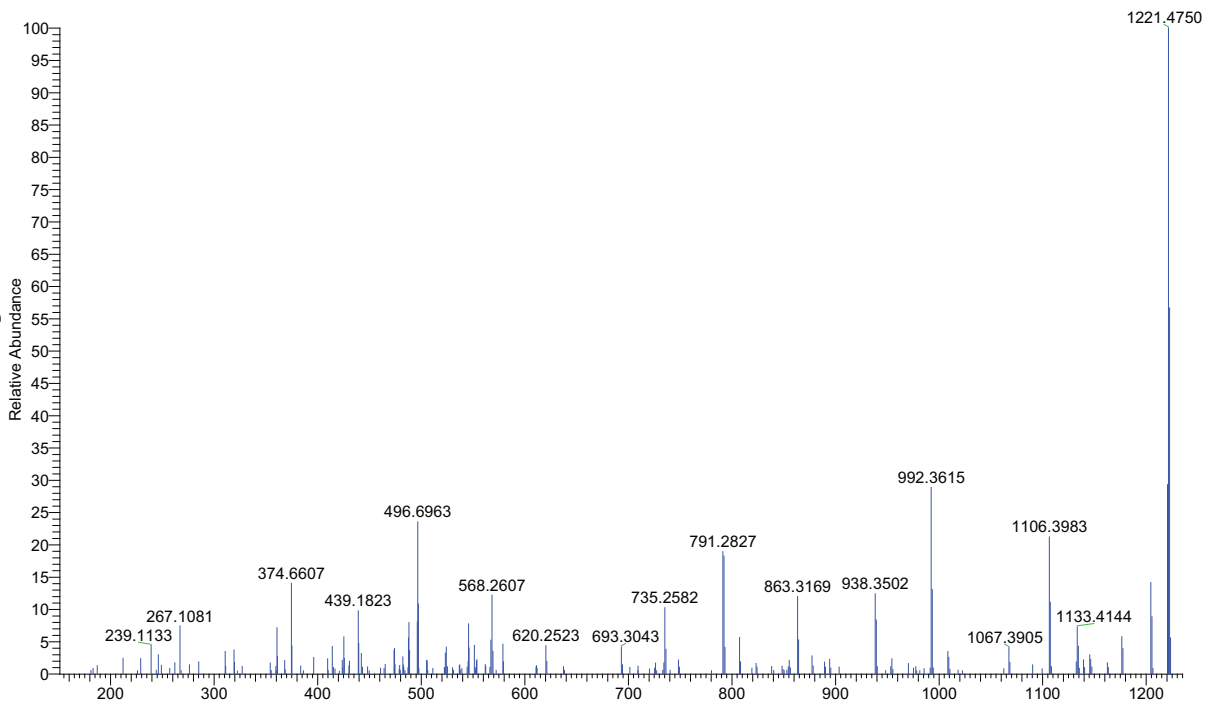
7

HEFH(pS)QENL

Synthetic Peptide



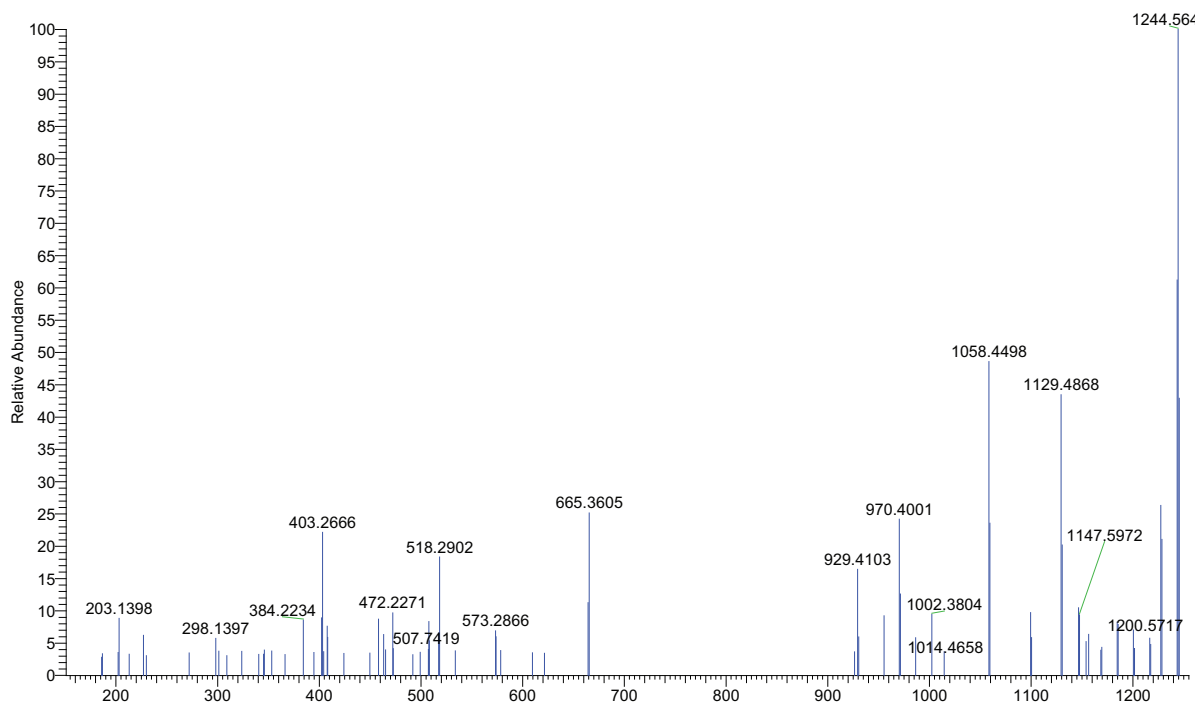
Candidate spectrum



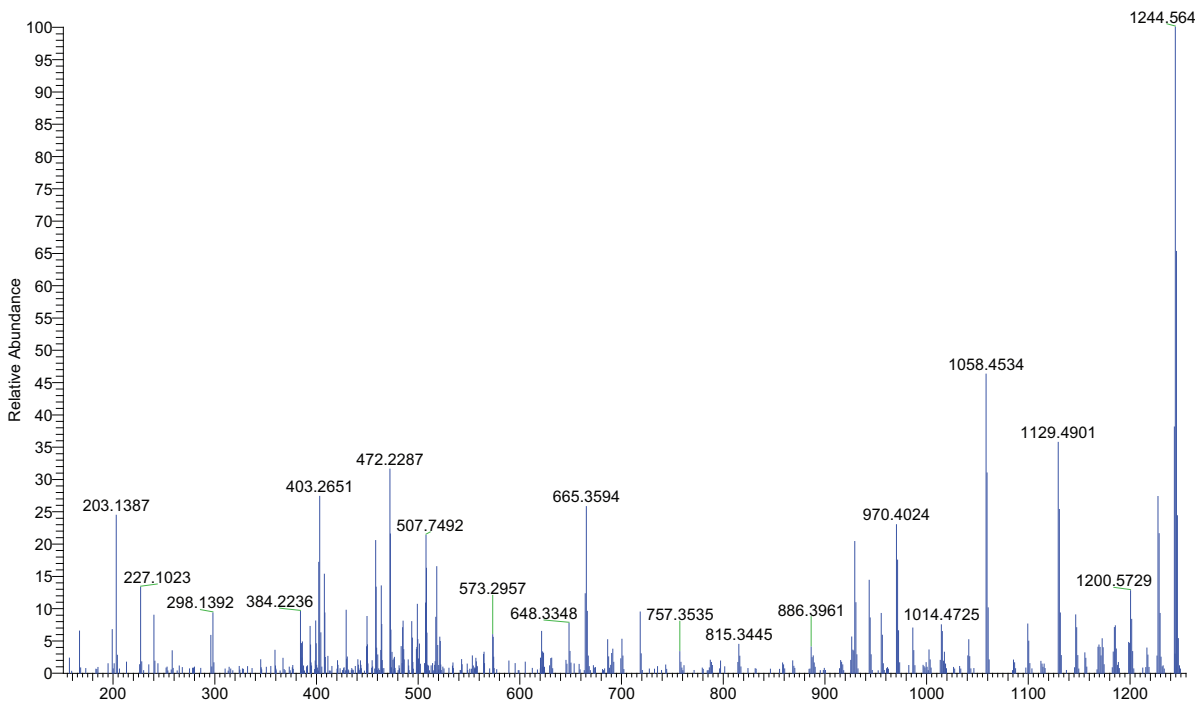
8

KEKDF(pS)PEAL

Synthetic Peptide



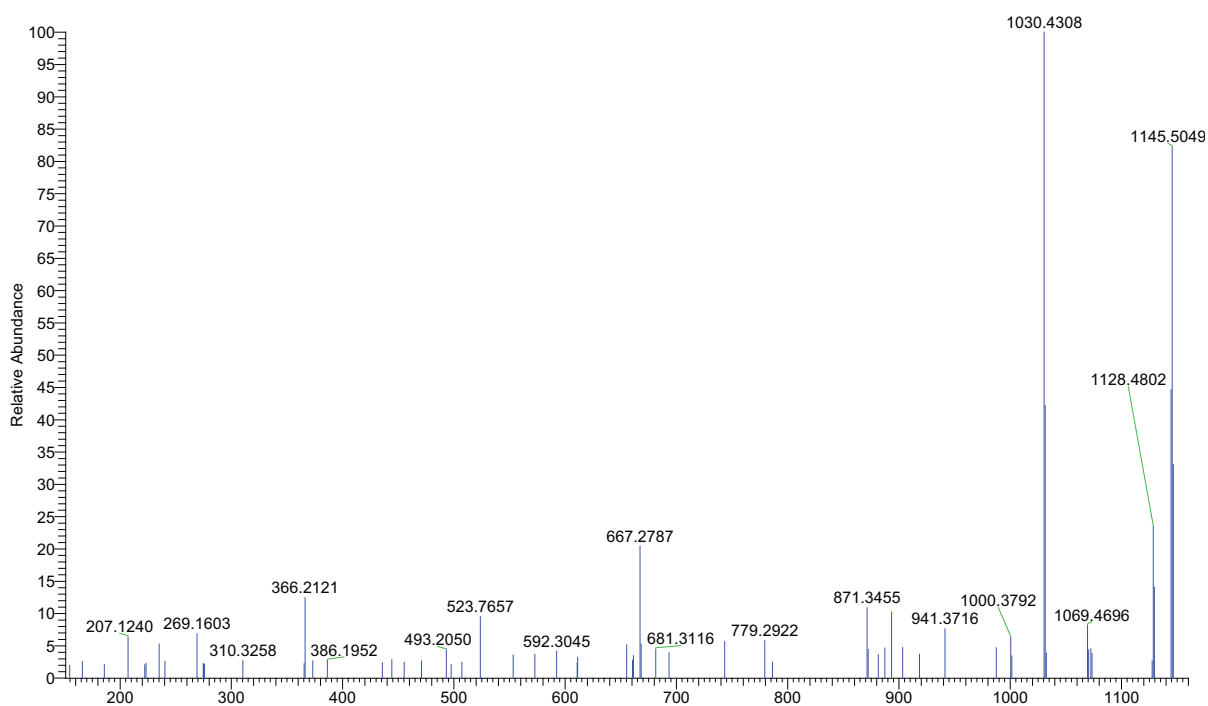
Candidate spectrum



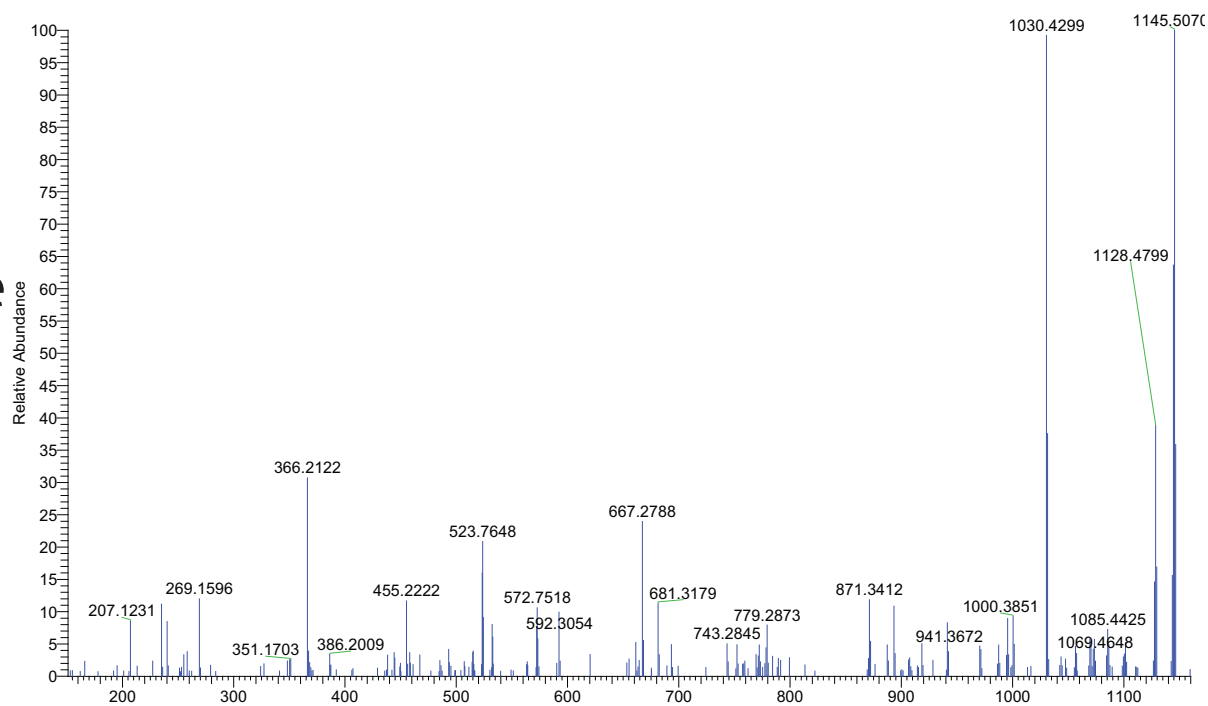
9

KEQ(pS)PEPHL

Synthetic Peptide



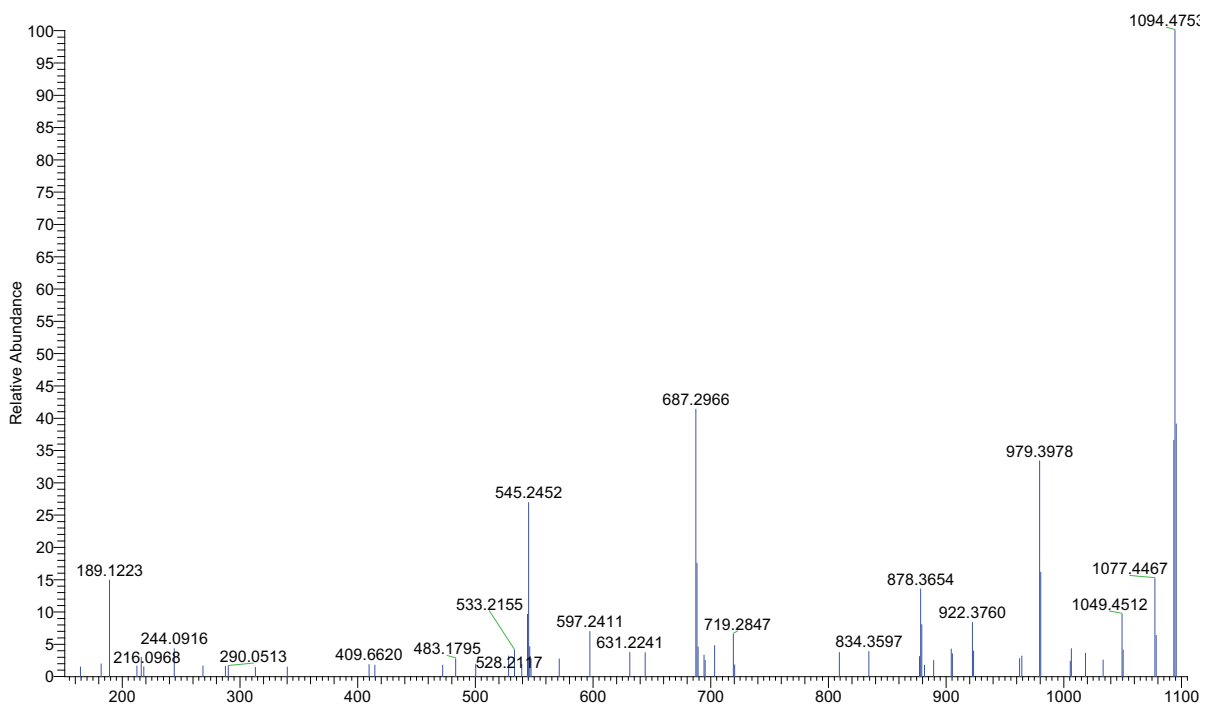
Candidate spectrum



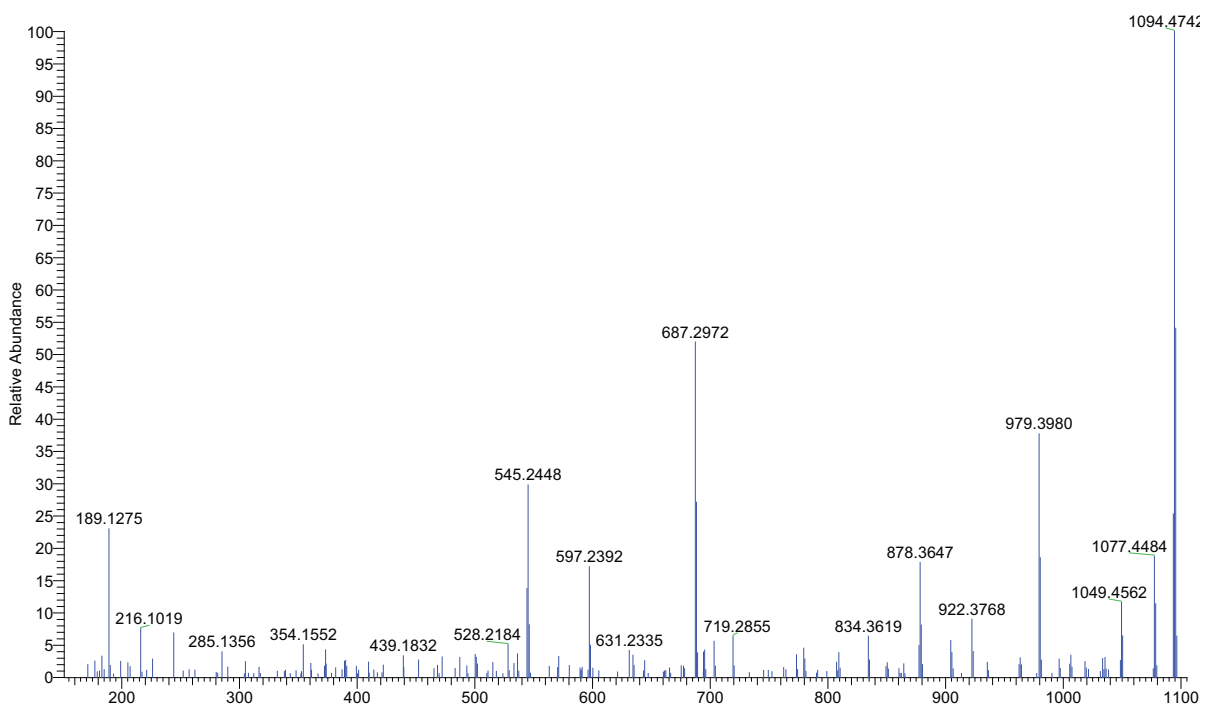
10

NEFH(pS)PIGL

Synthetic Peptide



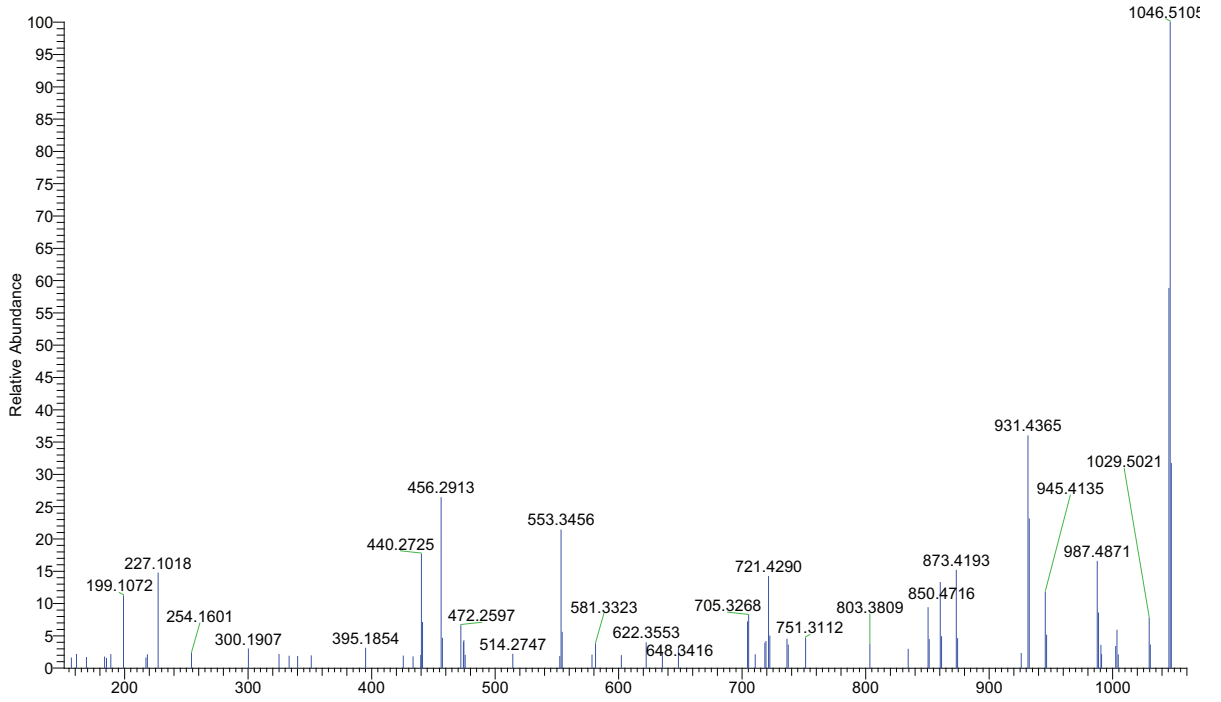
Candidate spectrum



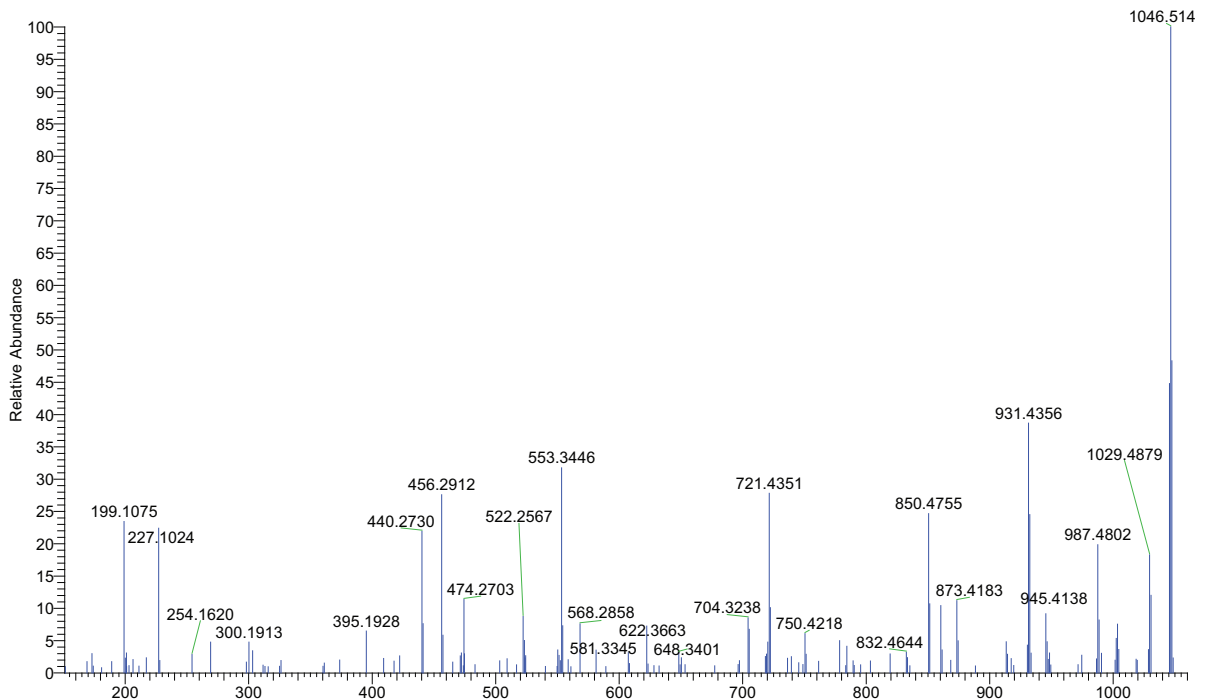
11

PEV(pS)PRPAL

Synthetic Peptide



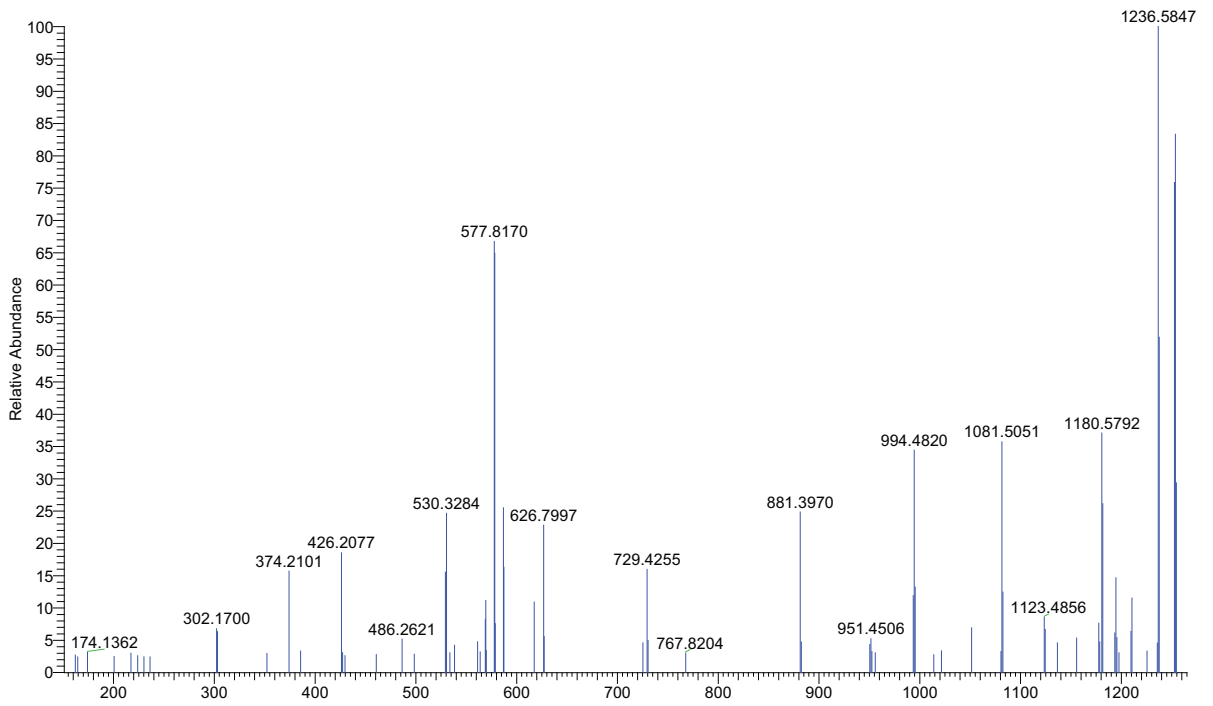
Candidate spectrum



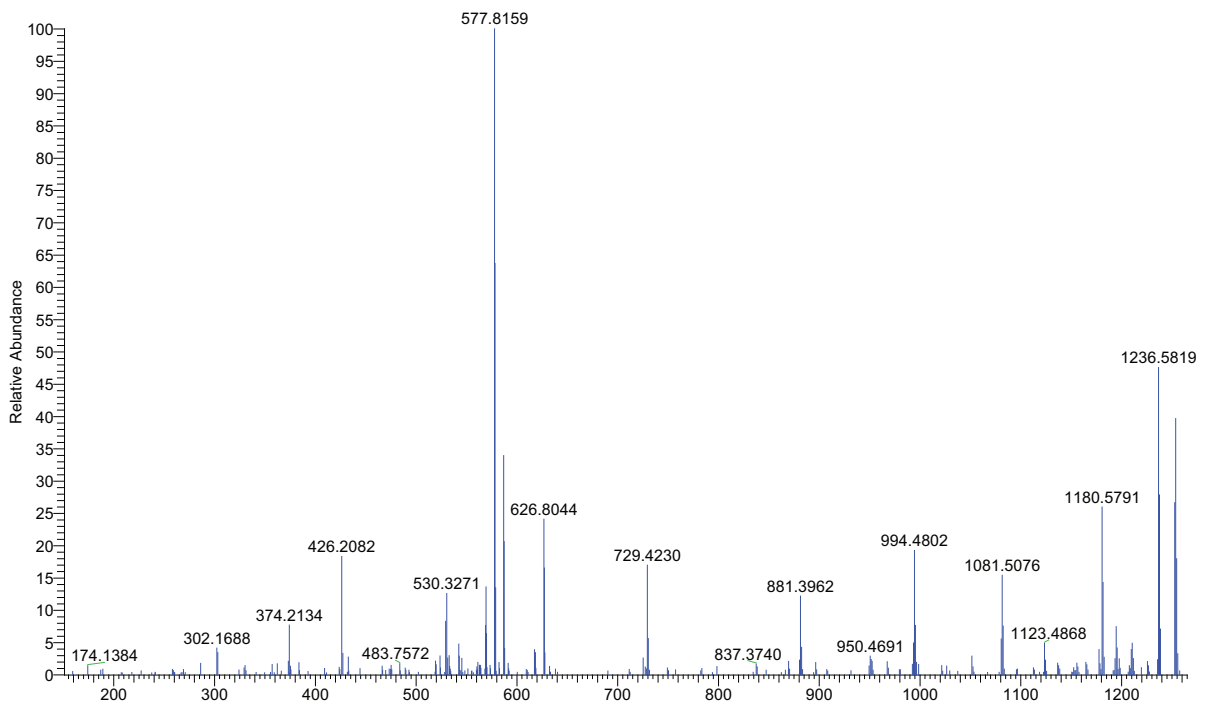
12

REA(pS)PSRLSVA

Synthetic Peptide



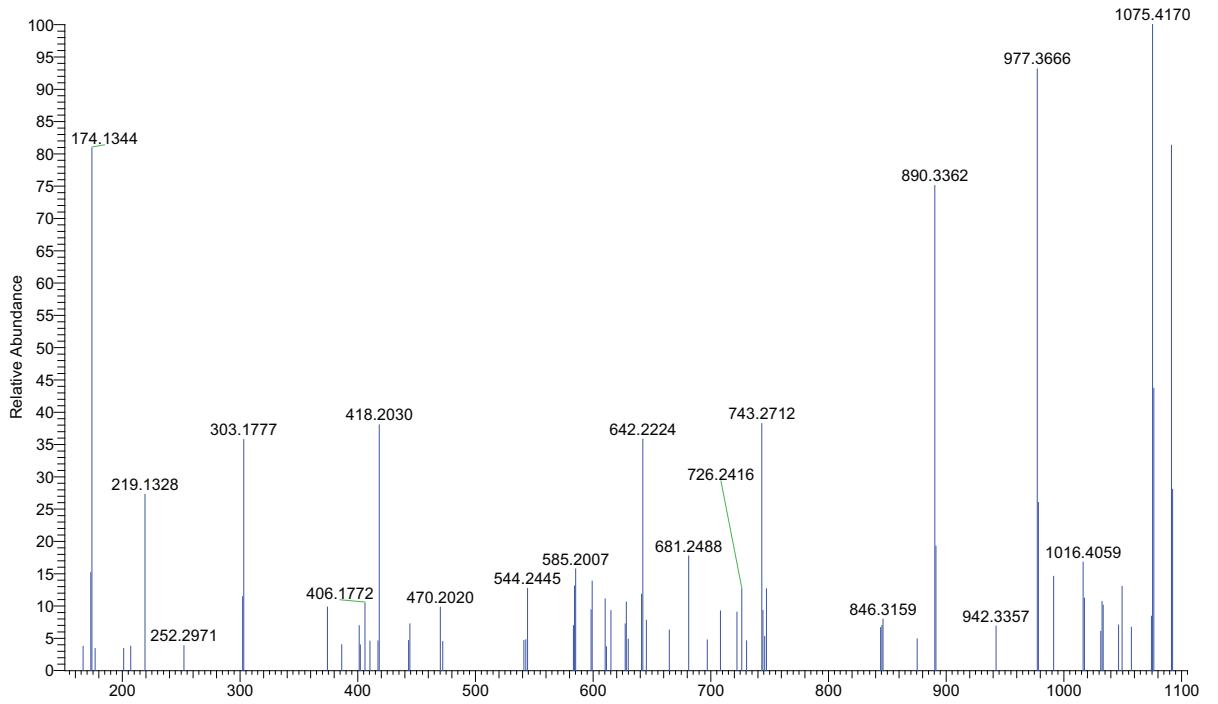
Candidate spectrum



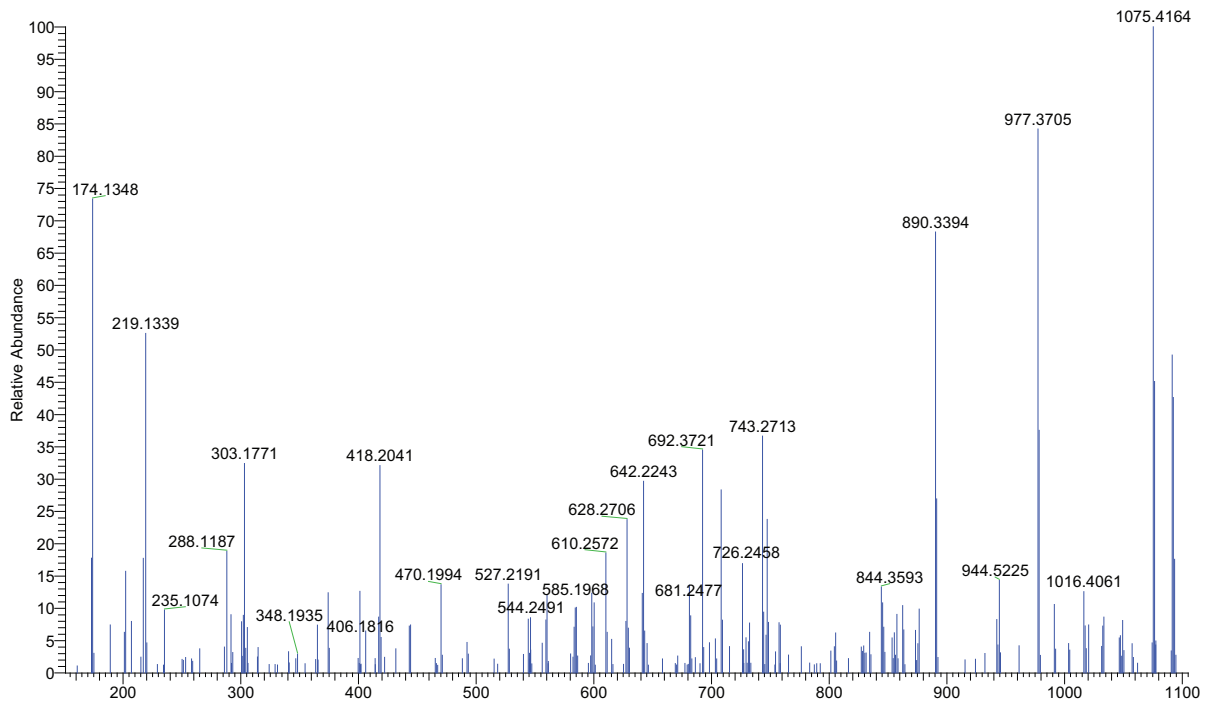
13

RED(pS)GTFSL

Synthetic Peptide



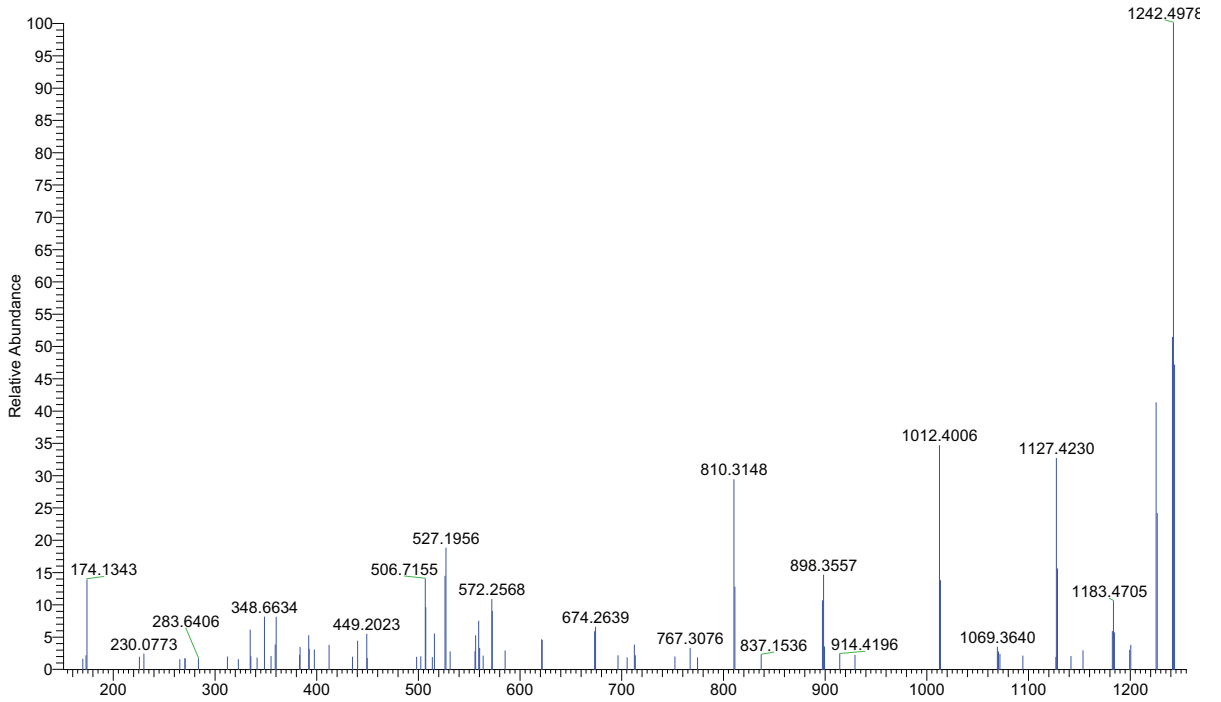
Candidate spectrum



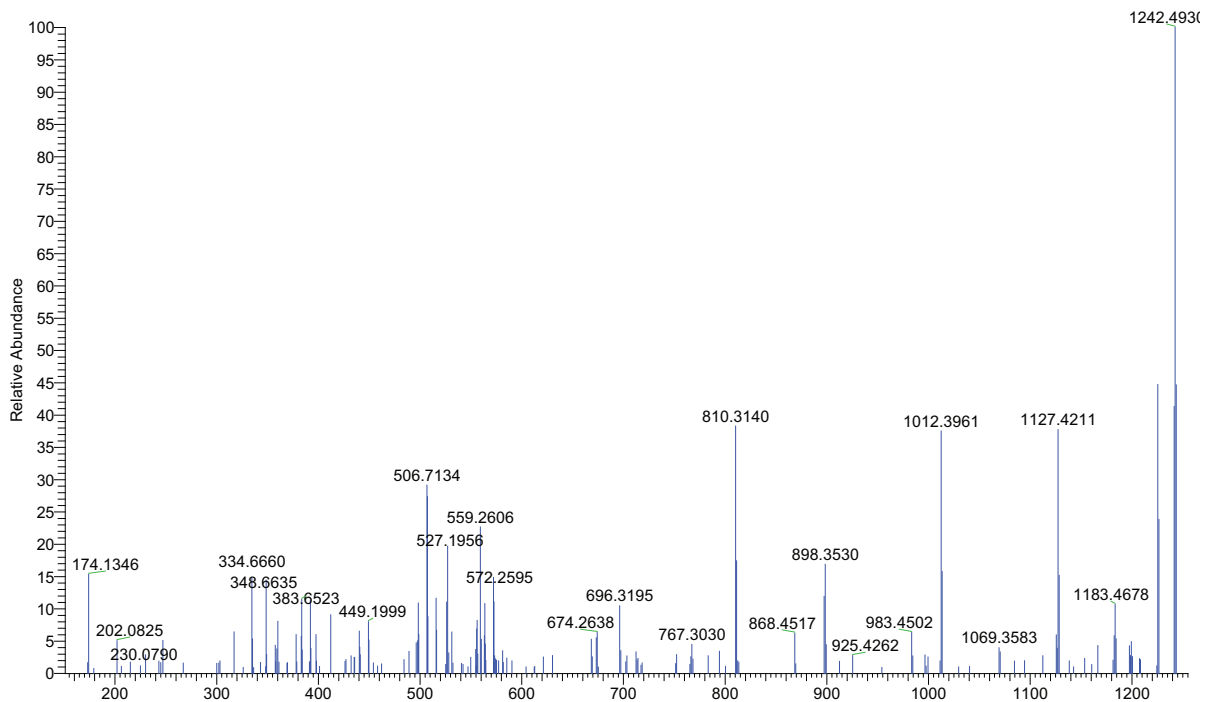
14

REG(pS)FHSNDL

Synthetic Peptide



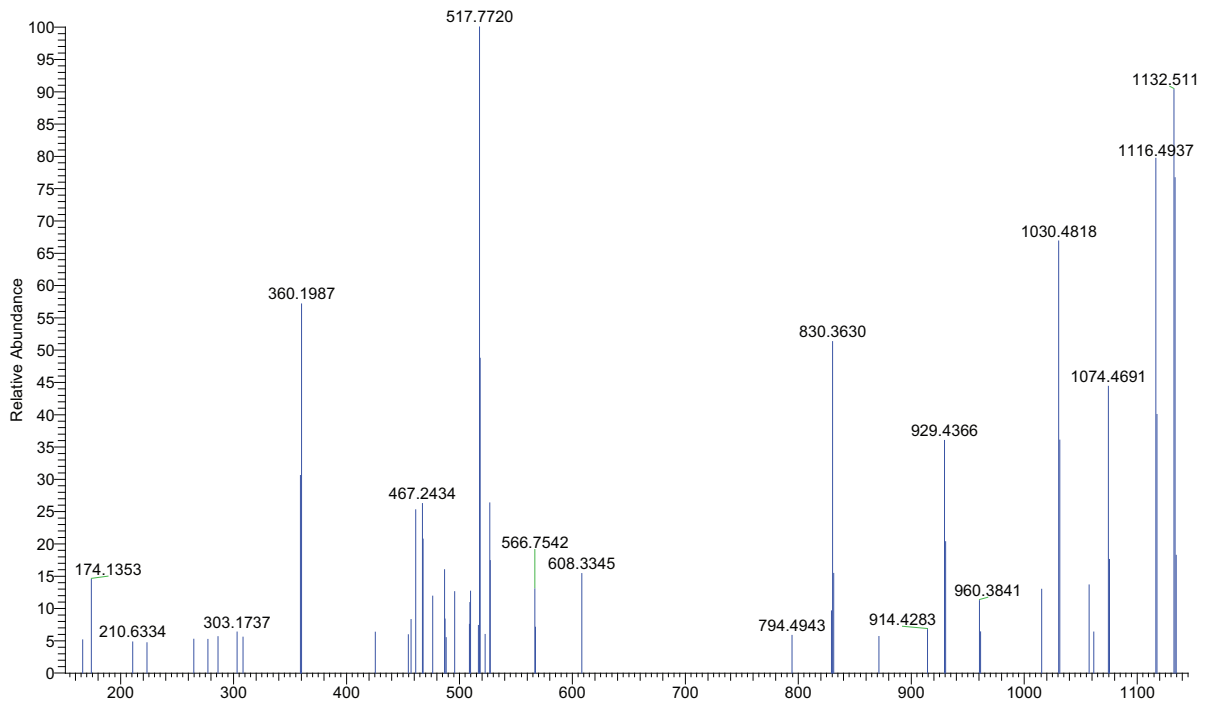
Candidate spectrum



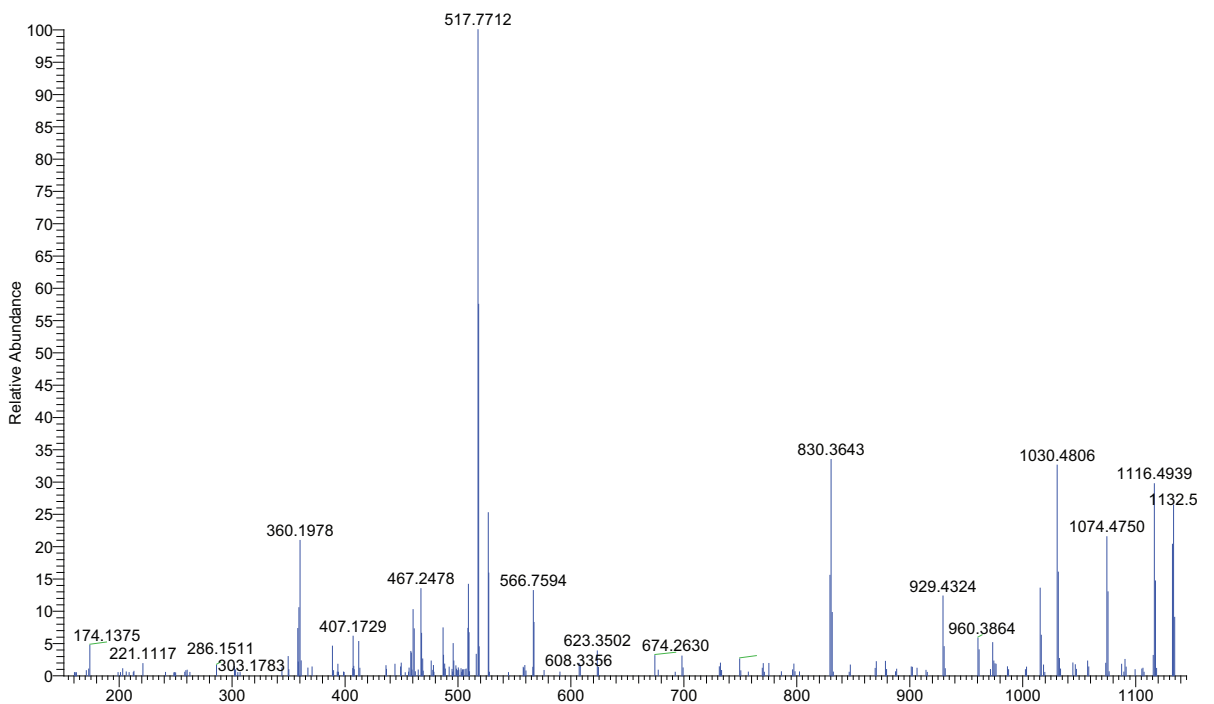
15

REG(pS)FRVTT

Synthetic Peptide



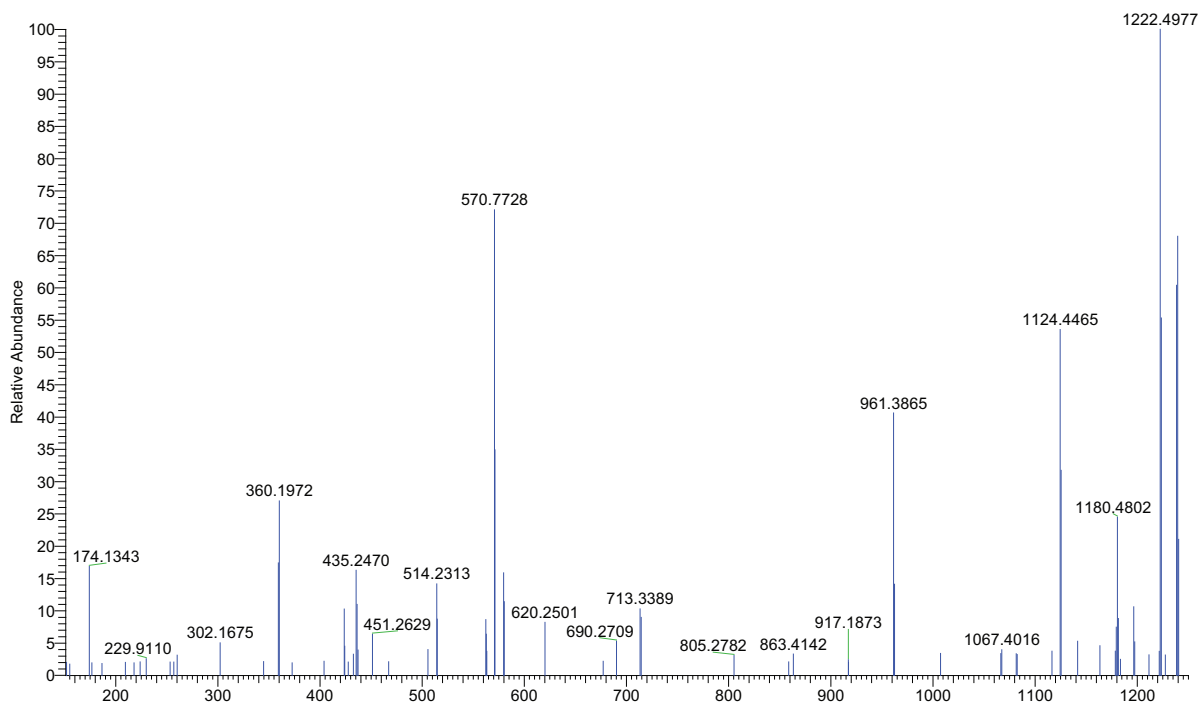
Candidate spectrum



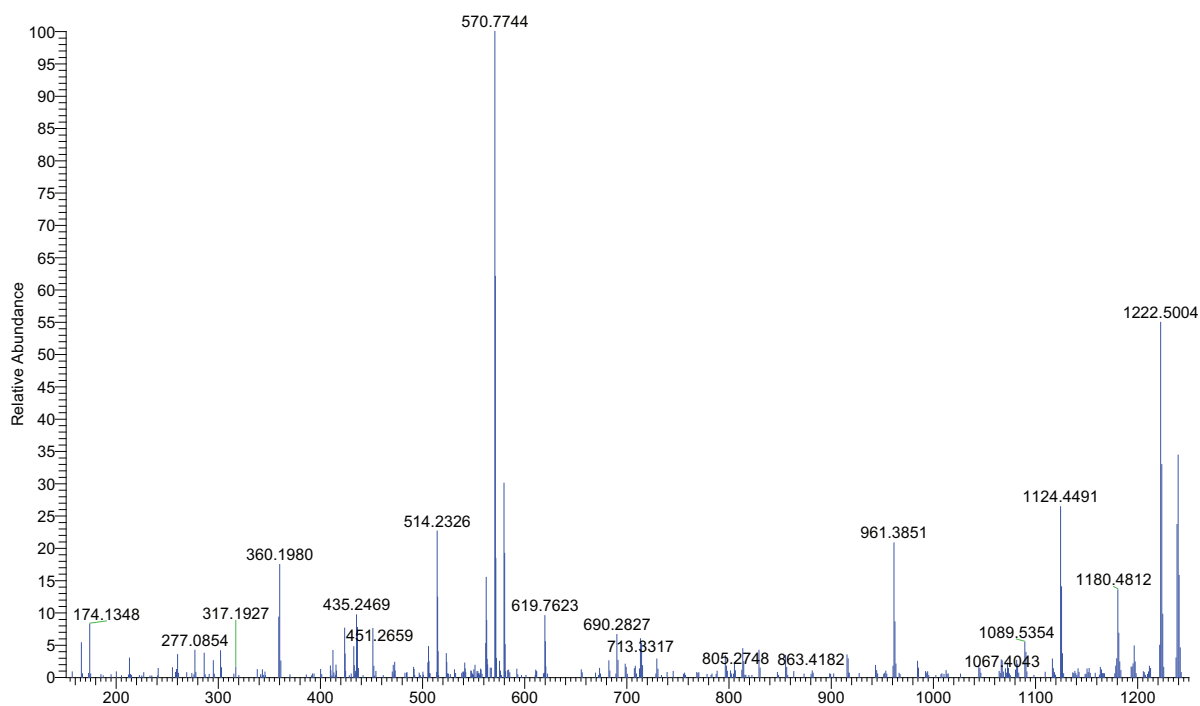
16

REG(pS)YDRYL

Synthetic Peptide



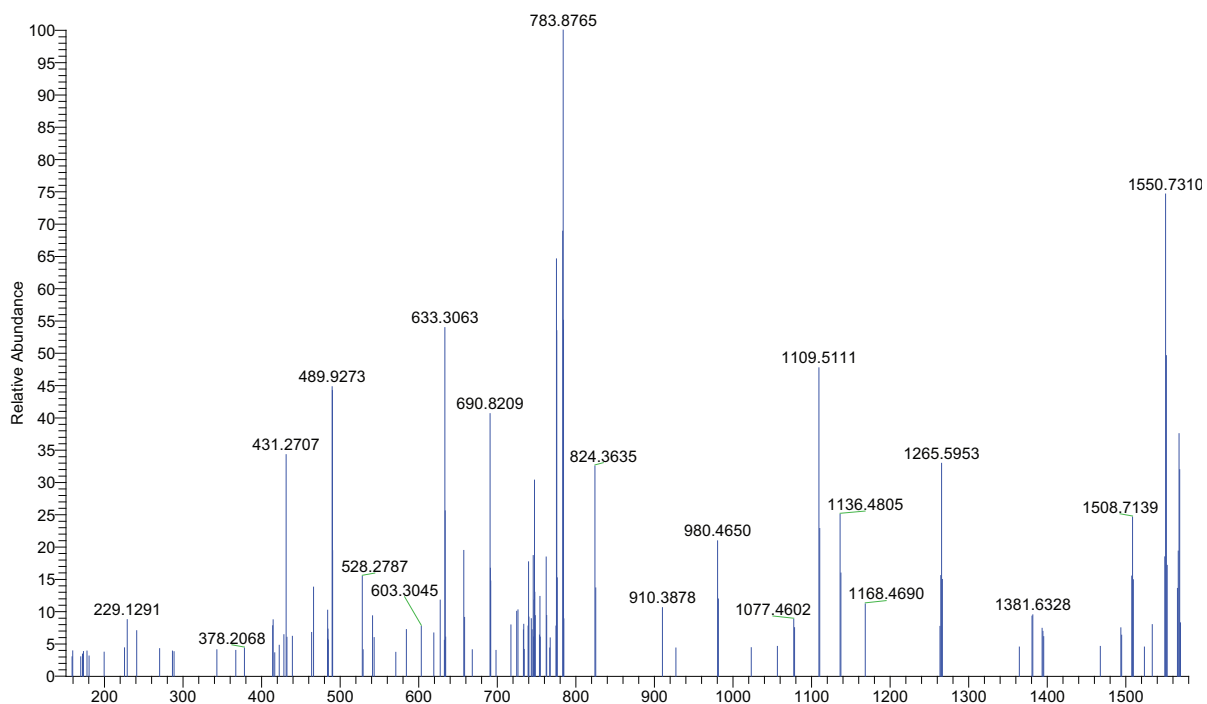
Candidate spectrum



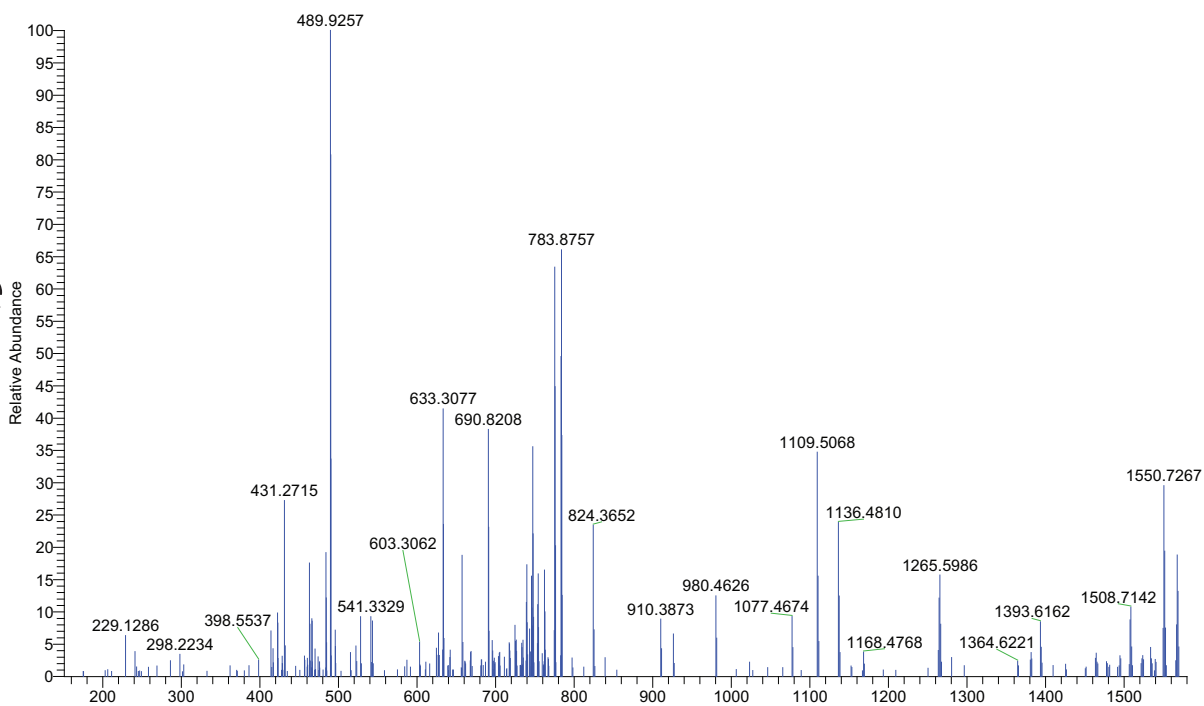
17

REKEP(pS)RERDIA

Synthetic Peptide



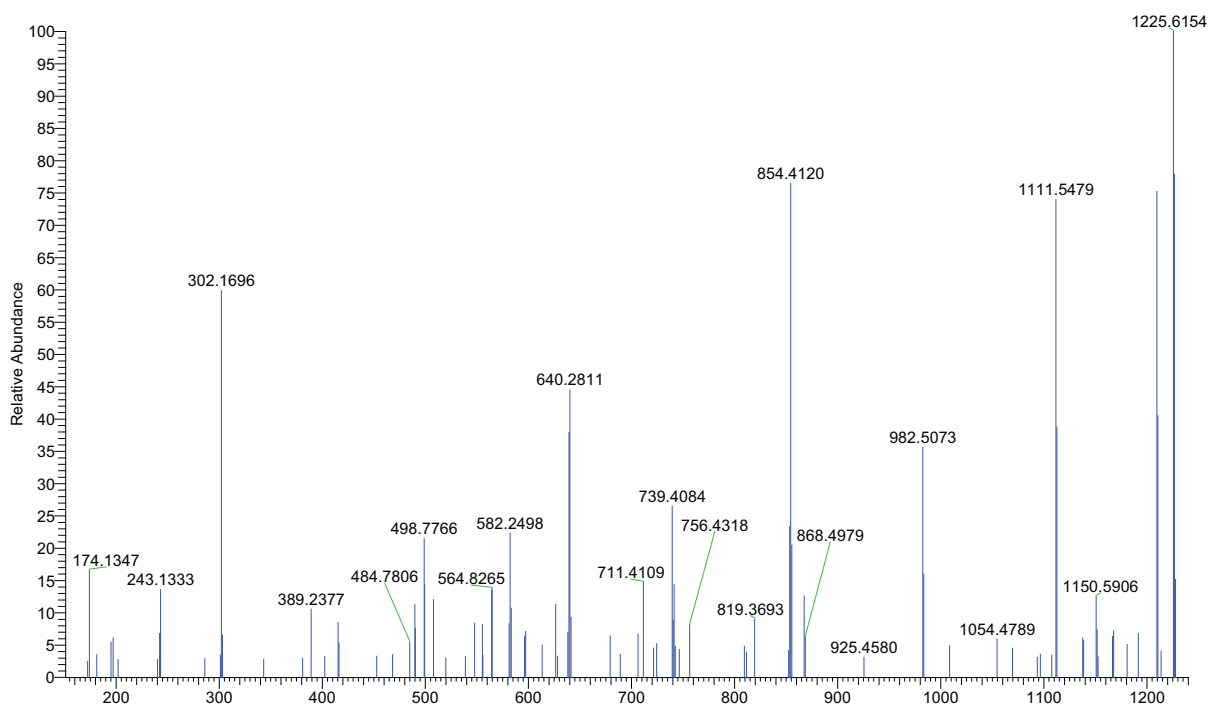
Candidate spectrum



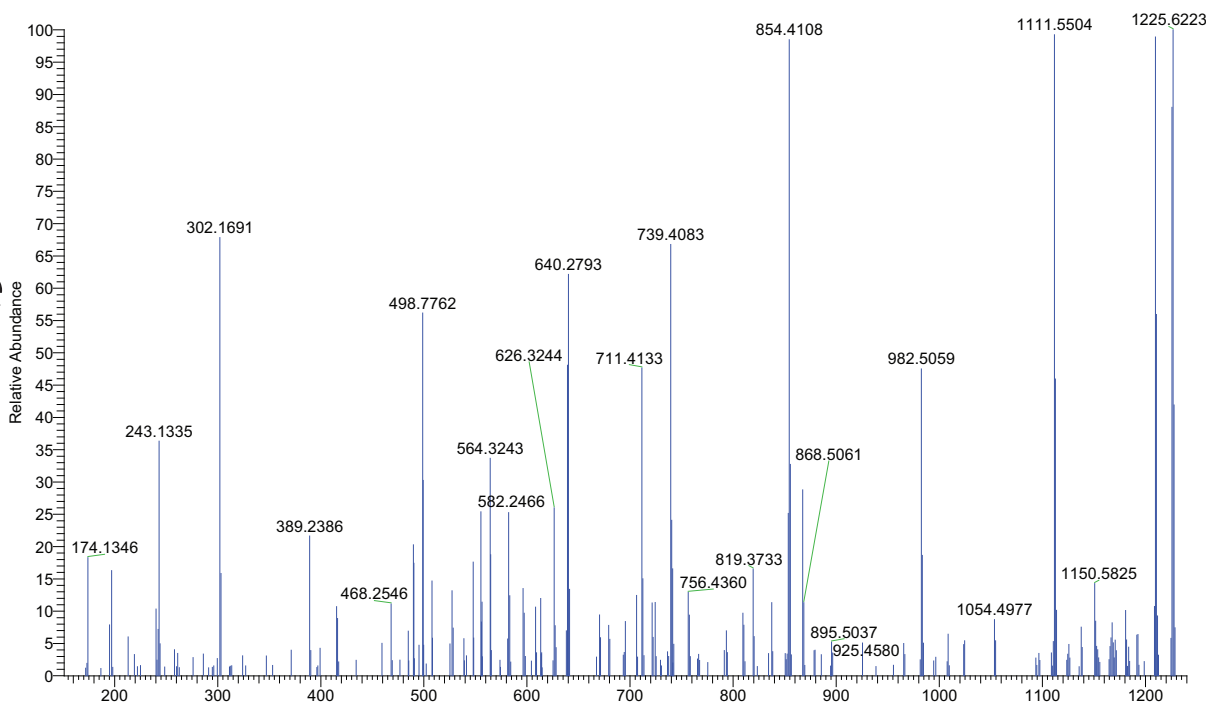
18

REL(pS)GTIKEI

Synthetic Peptide



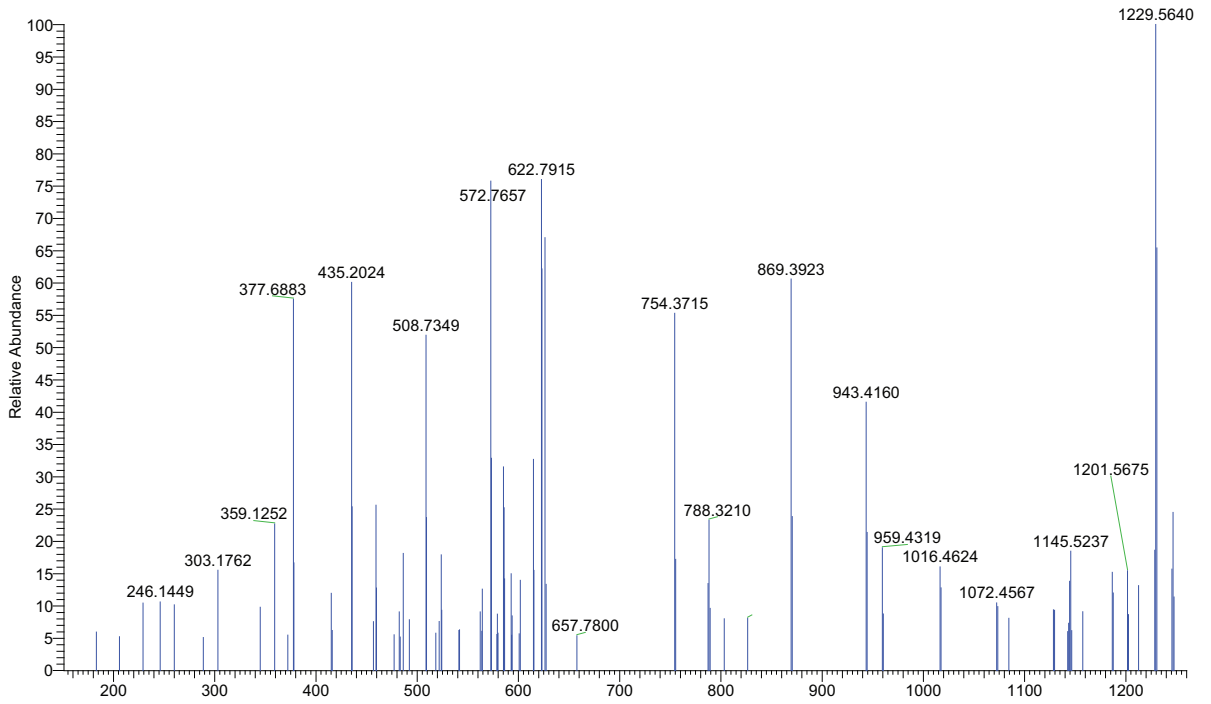
Candidate spectrum



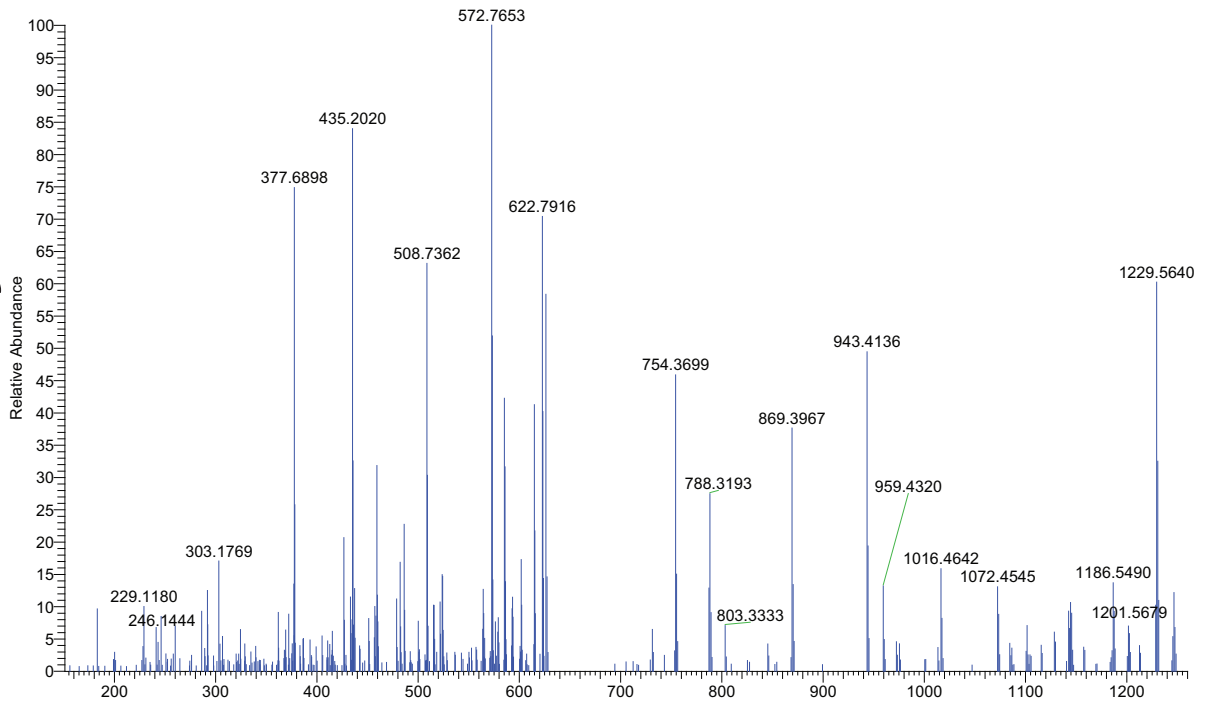
19

RER(pS)KDFQV

Synthetic Peptide



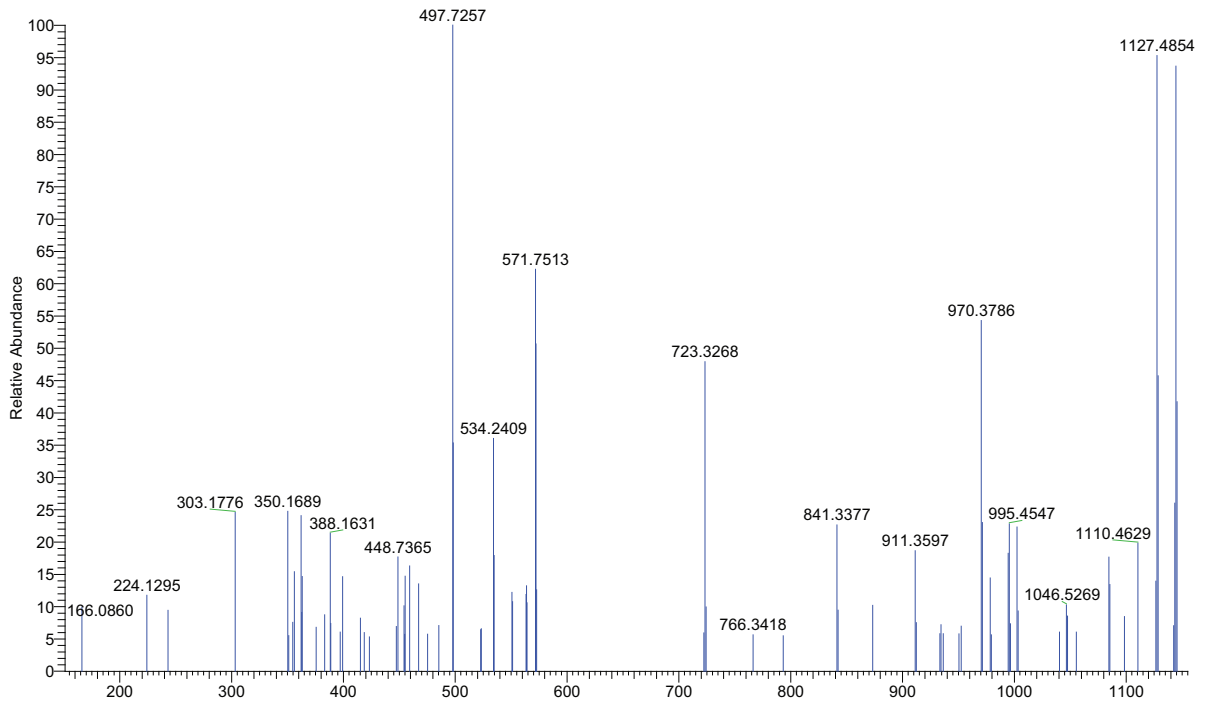
Candidate spectrum



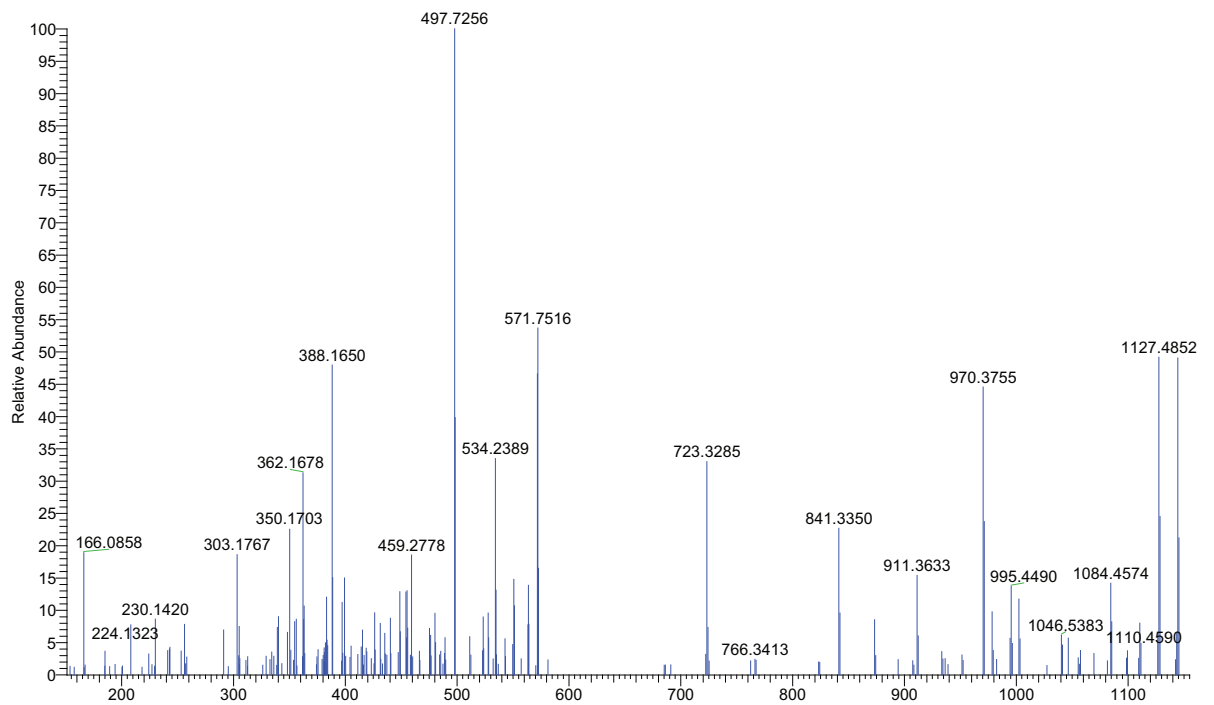
20

RER(pS)PSPSF

Synthetic Peptide



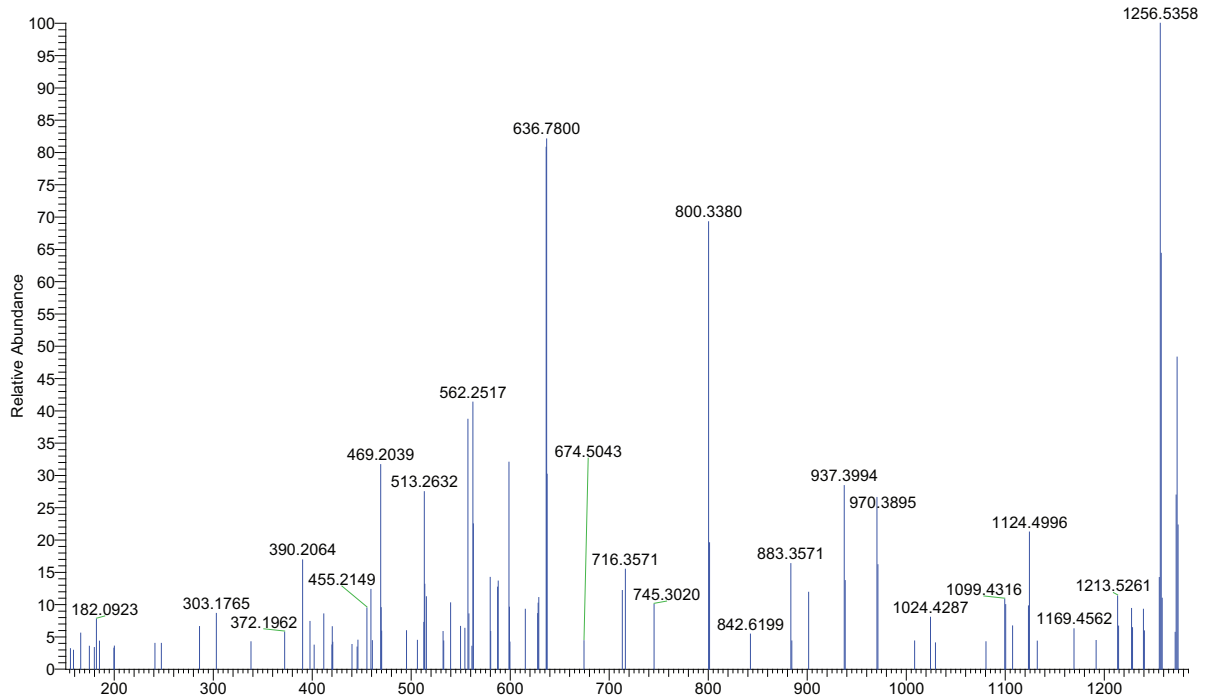
Candidate spectrum



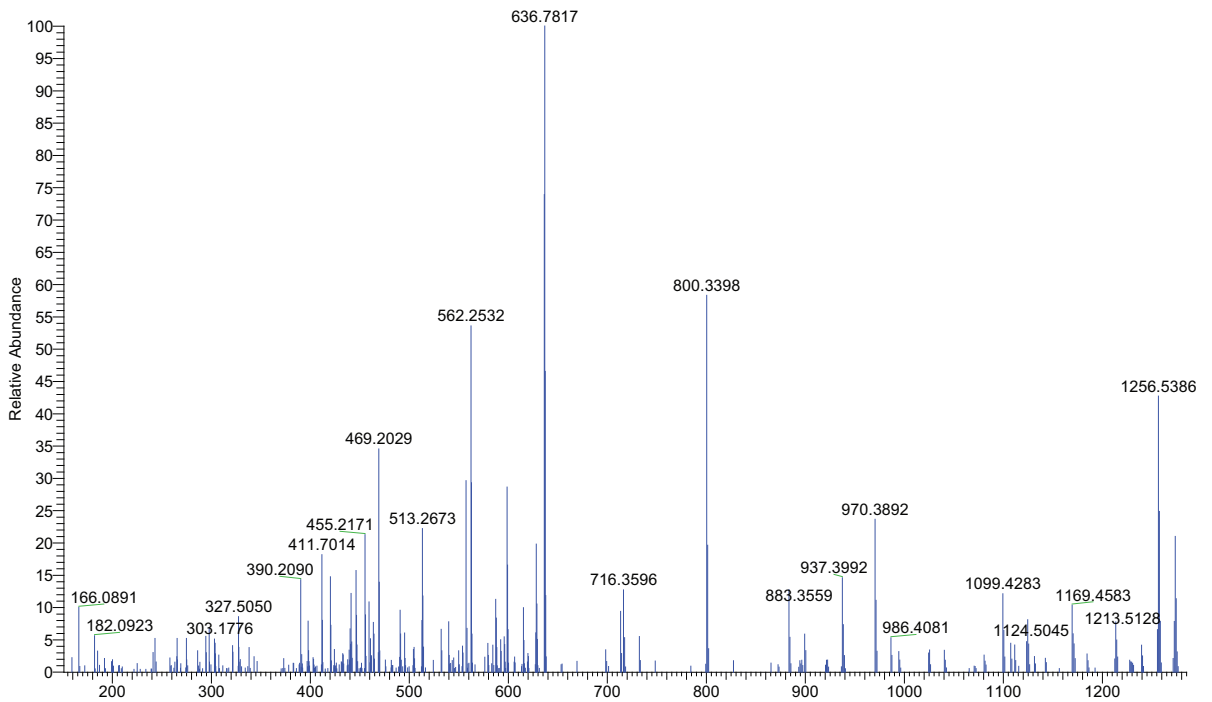
21

RES(pS)RSHSVF

Synthetic Peptide



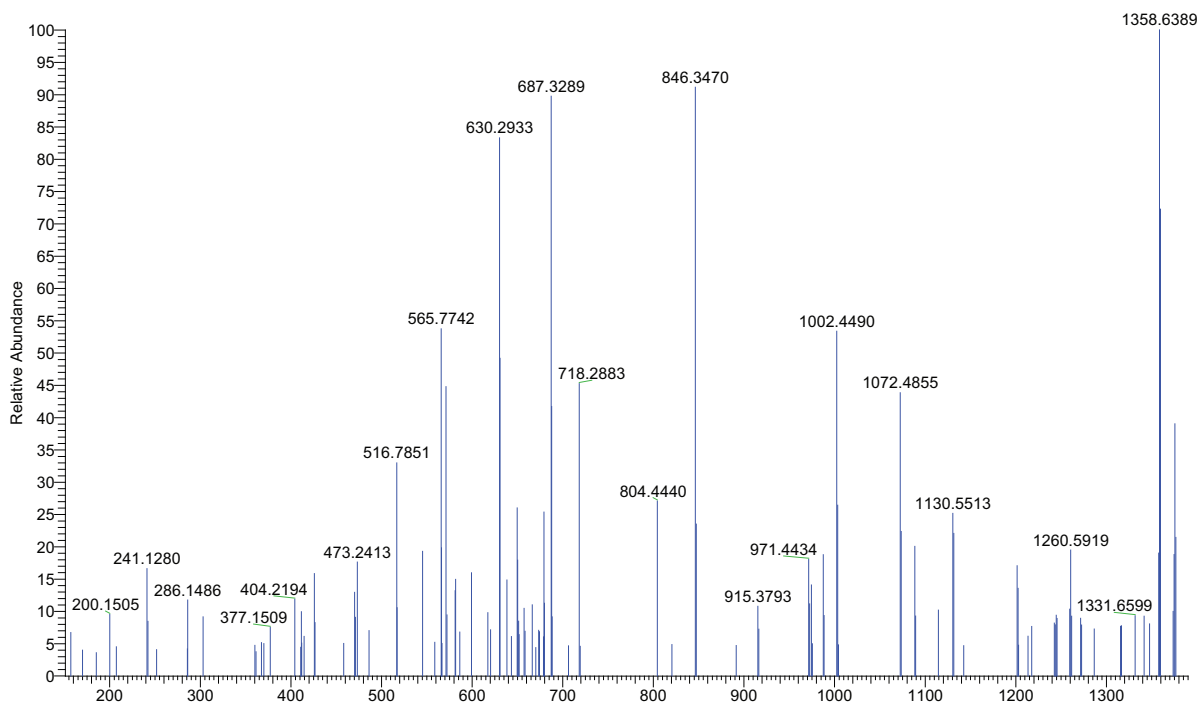
Candidate spectrum



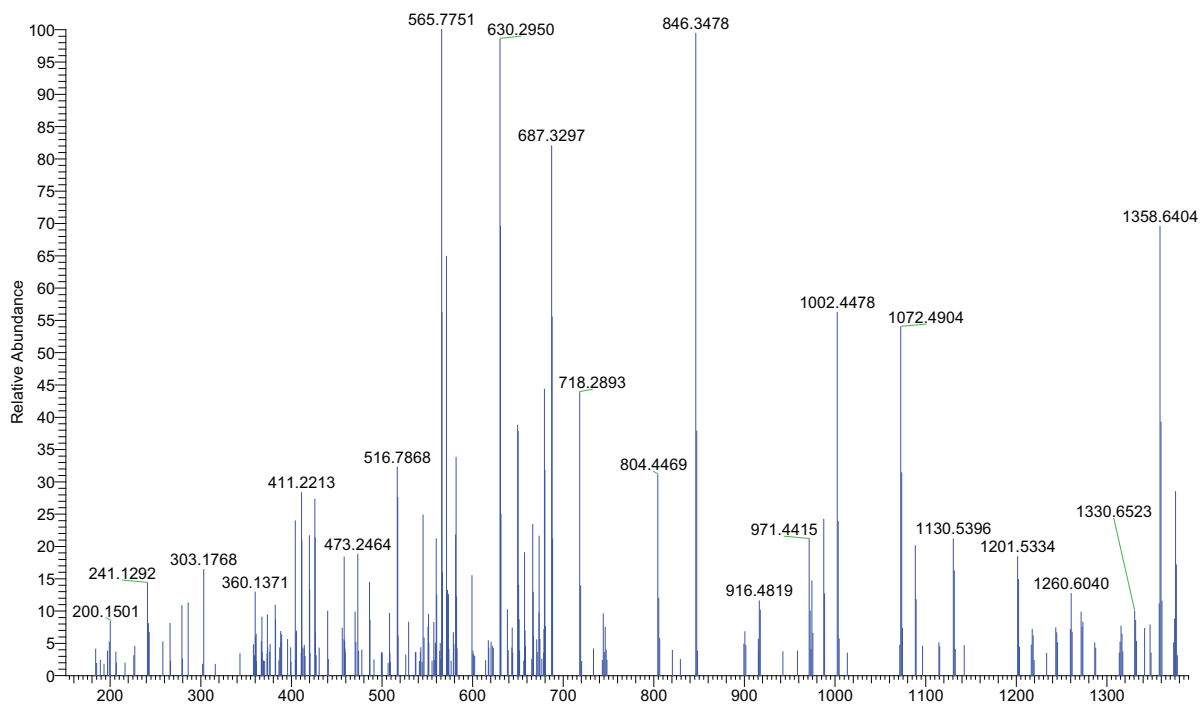
22

RET(pS)FQRKEL

Synthetic Peptide



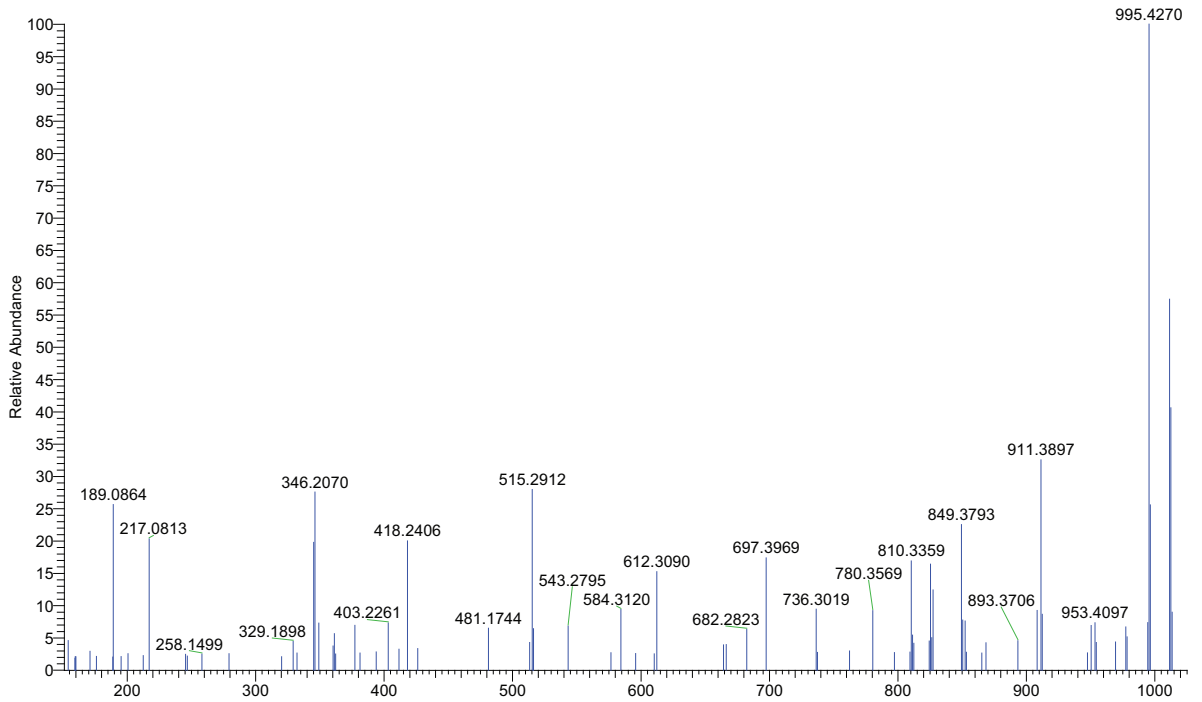
Candidate spectrum



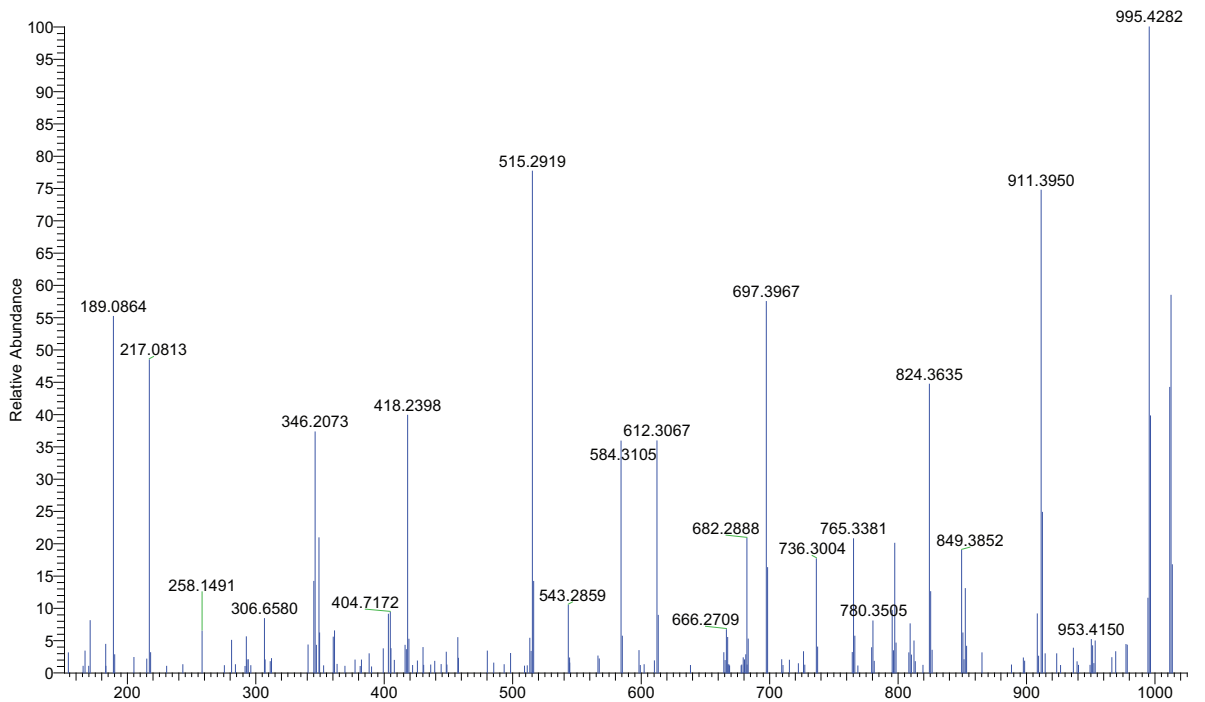
23

SEL(pS)PGRSV

Synthetic Peptide



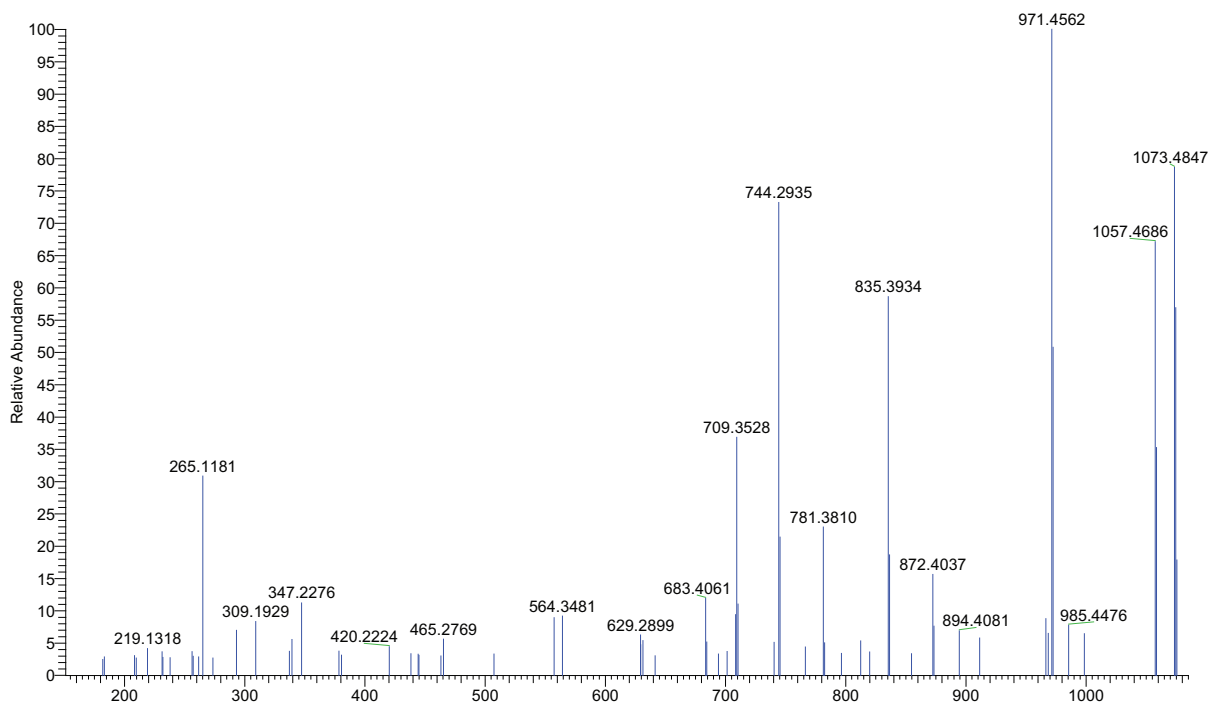
Candidate spectrum



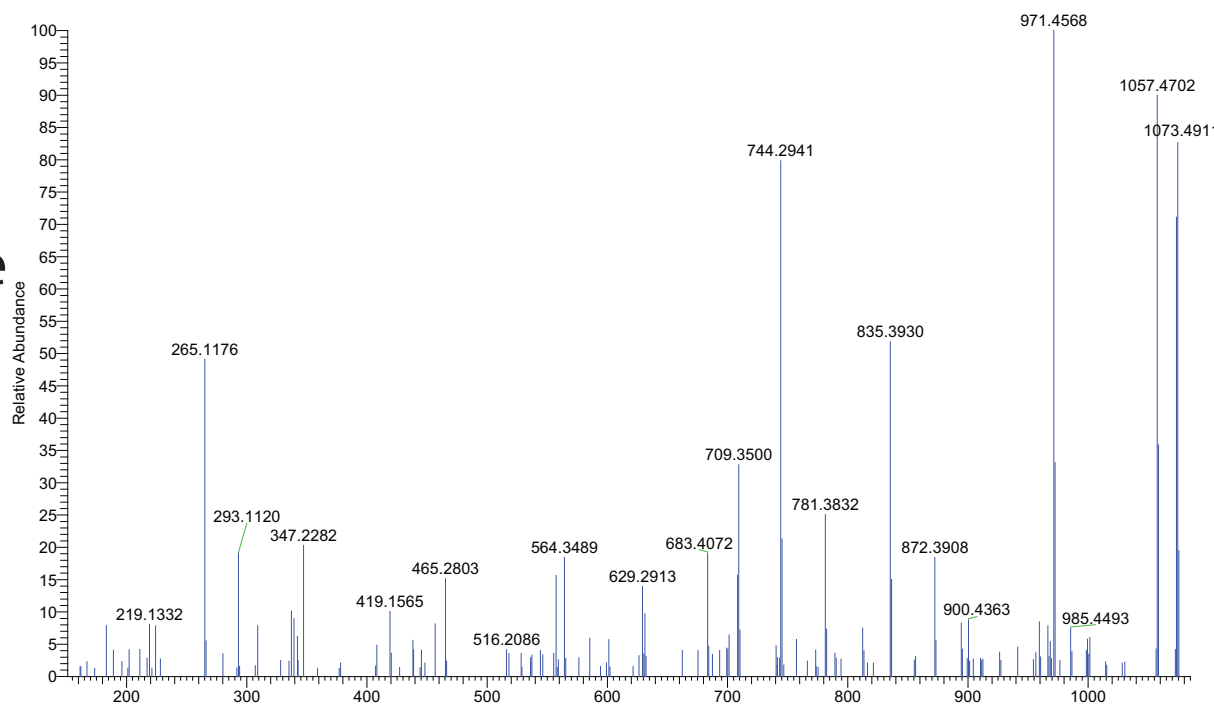
24

YEG(pS)PIKVT

Synthetic Peptide



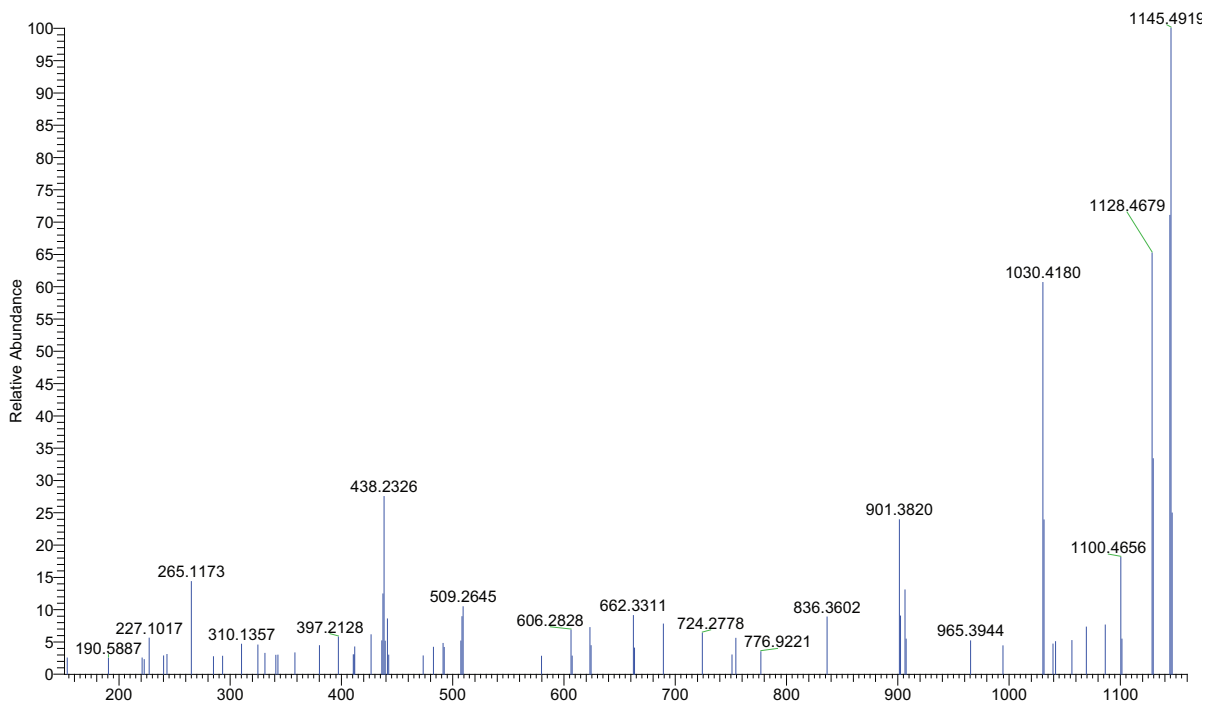
Candidate spectrum



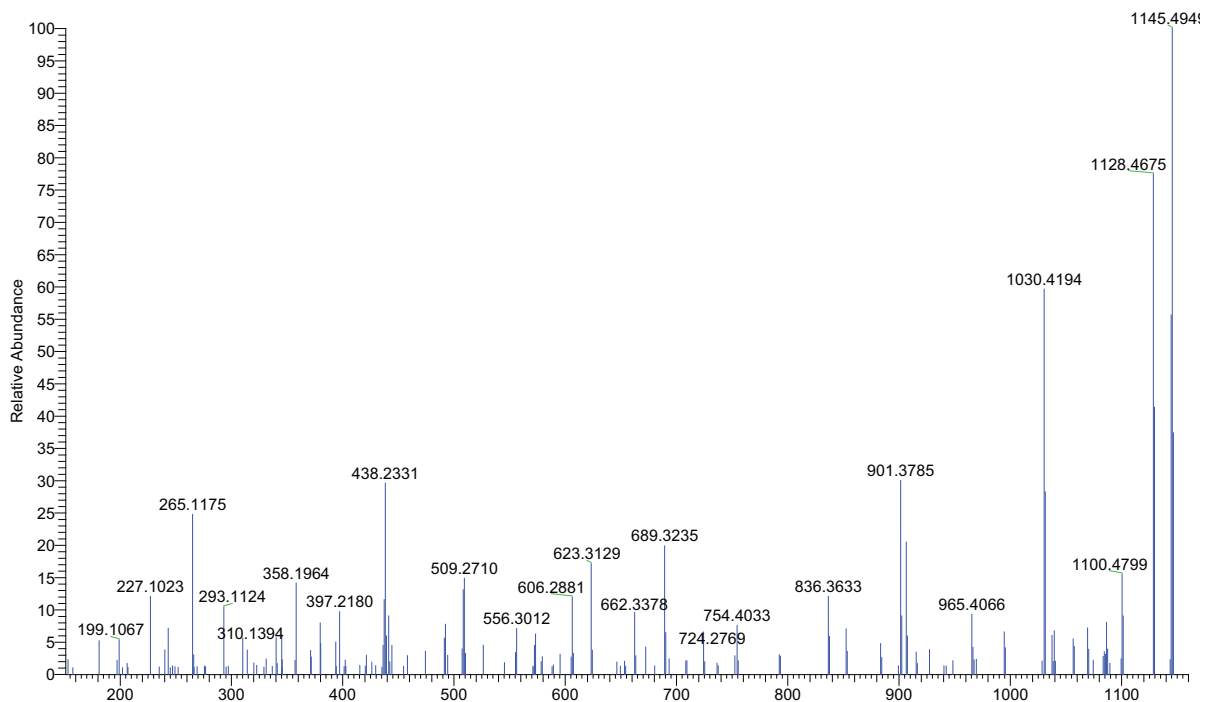
25

YEKAN(pT)PEL

Synthetic Peptide



Candidate spectrum



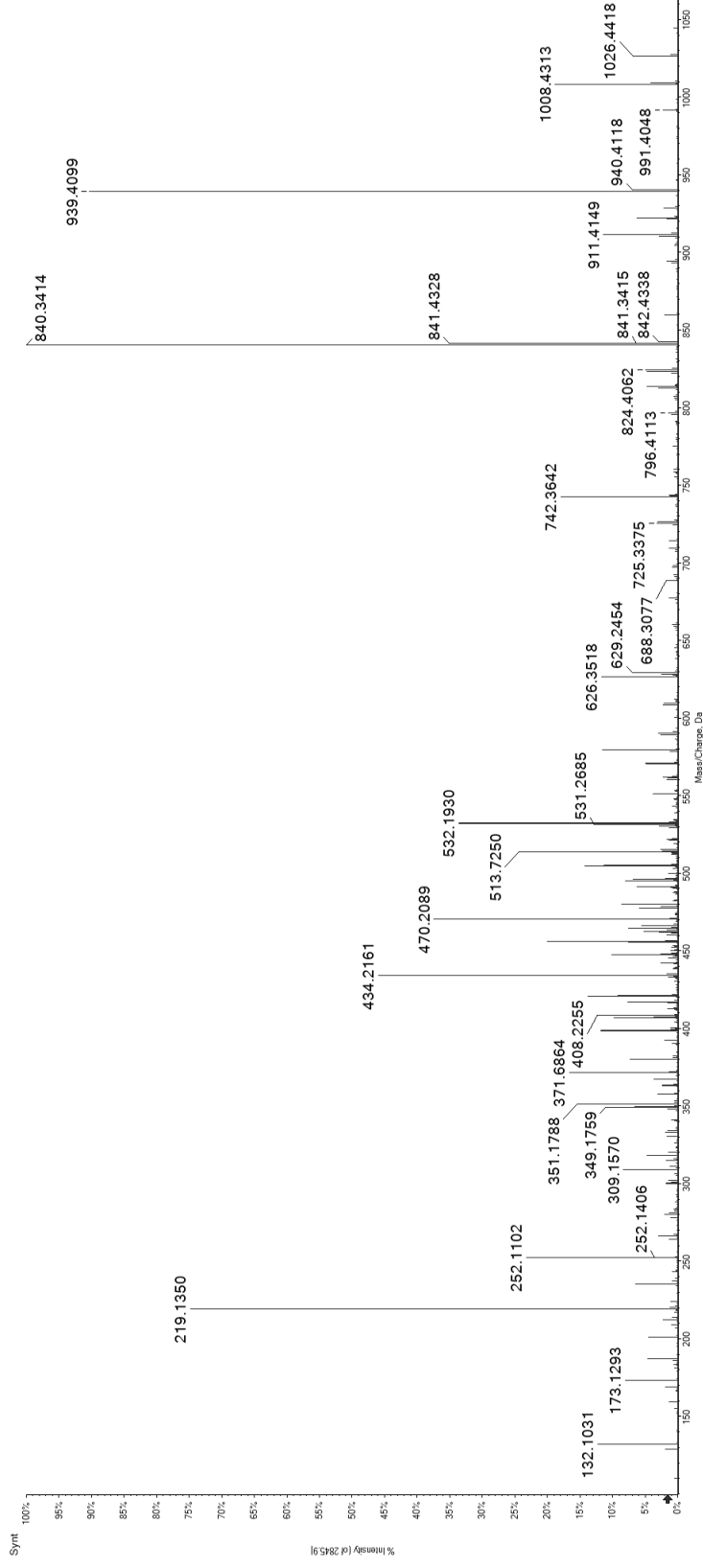
Anexo 1.3: Espectros MS2 de los fosfopéptidos identificados mediante CID.

Cada espectro MS2 candidato se compara con el espectro de fragmentación del péptido sintético correspondiente. Las comparaciones están ordenadas como se muestran en la Tabla 9.

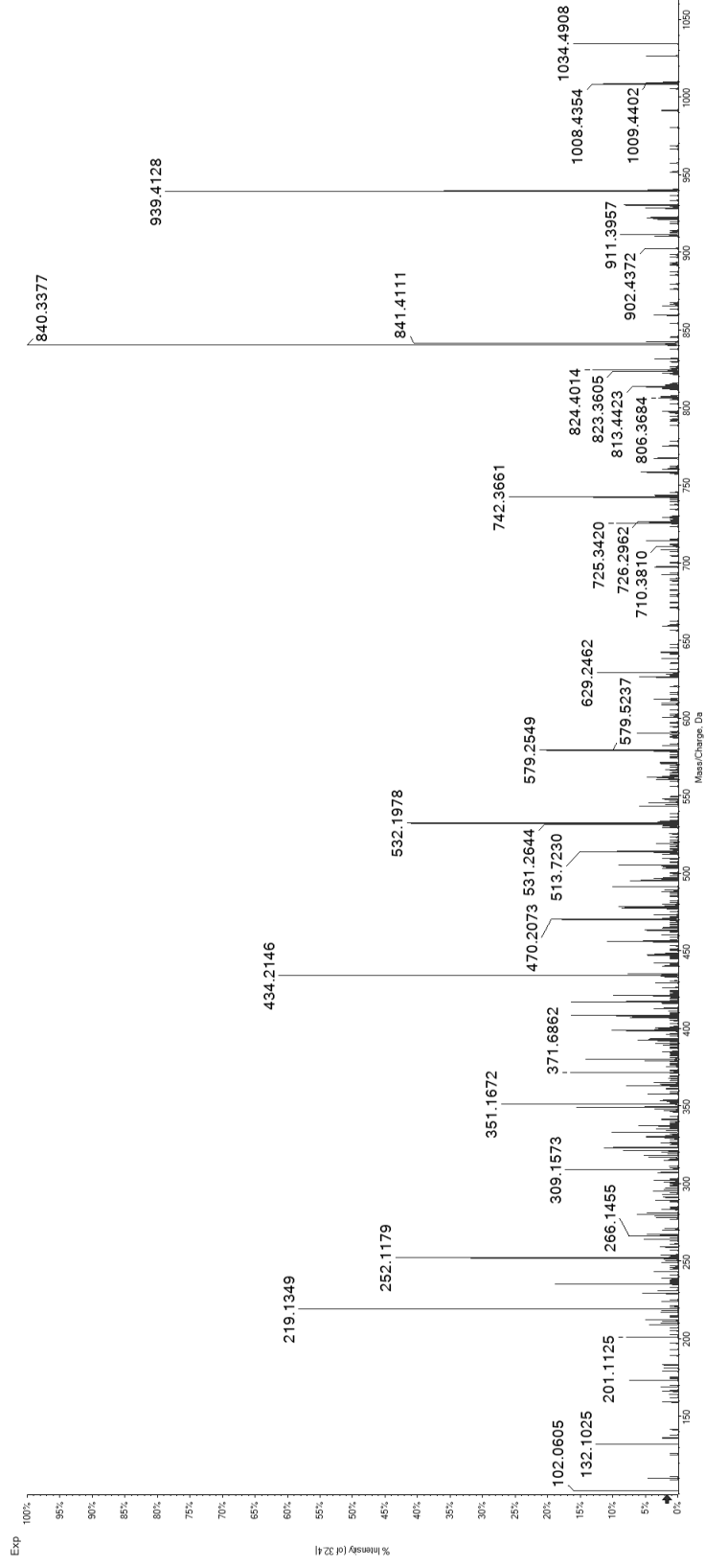
1

NHV(pT)PPNVSL

Synthetic Peptide



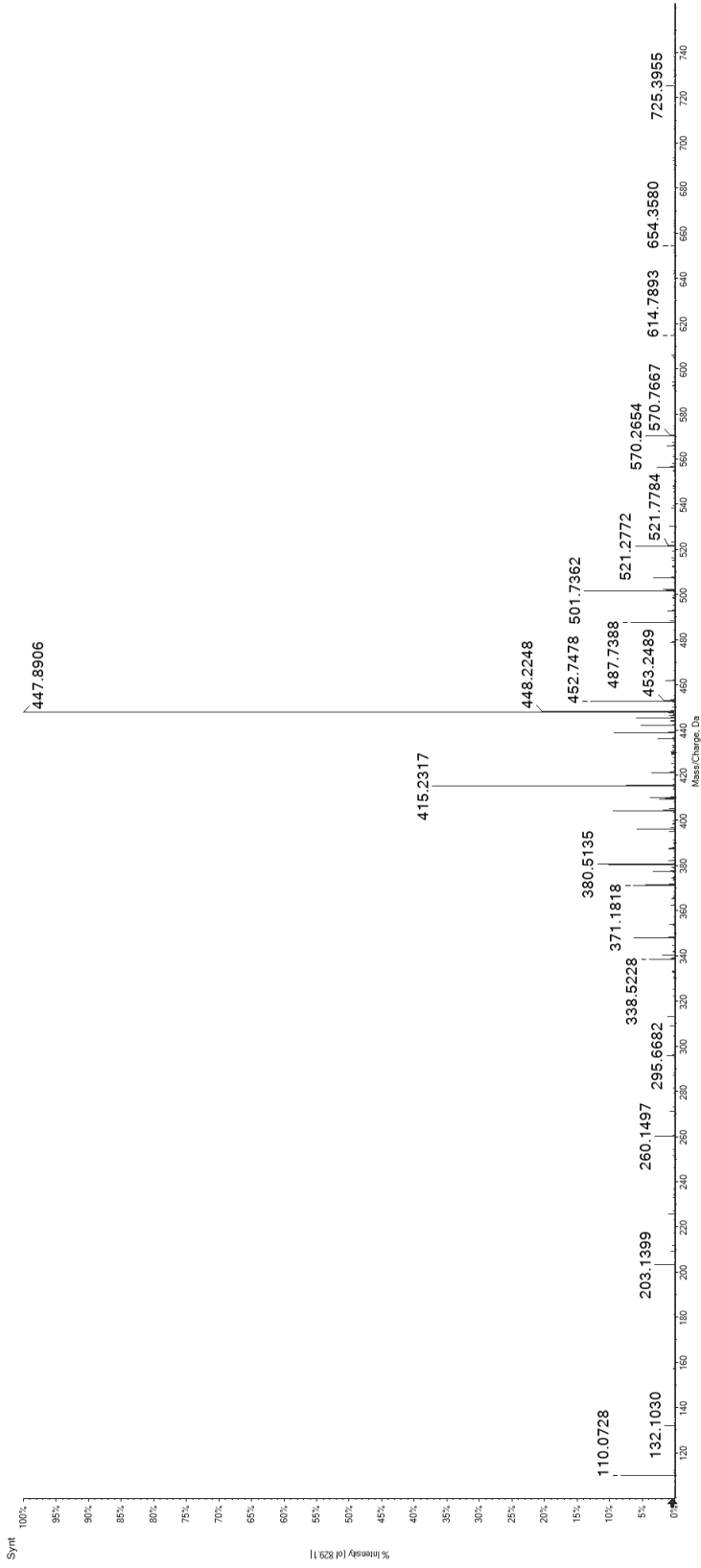
Candidate Spectrum



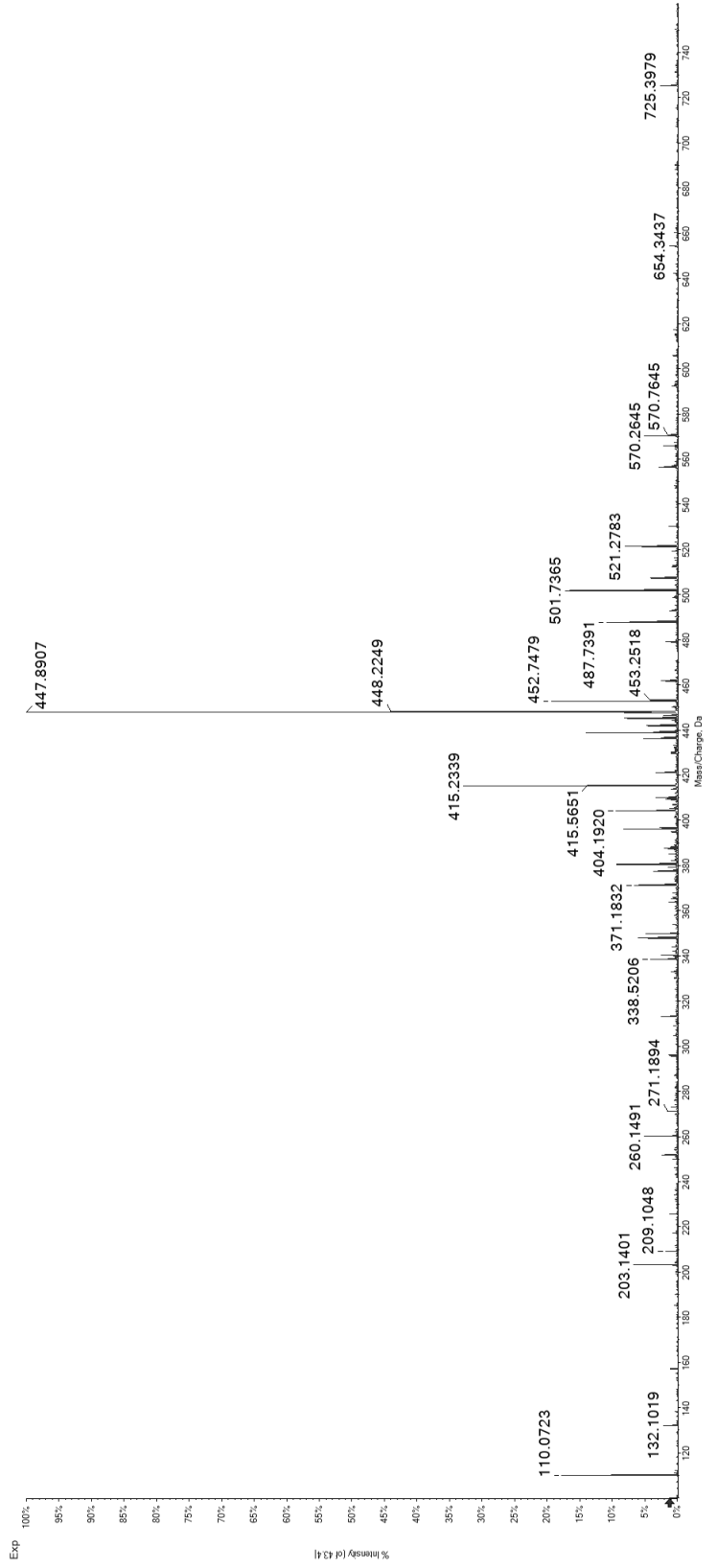
2

RRH(pS)ASNLHAL

Synthetic Peptide



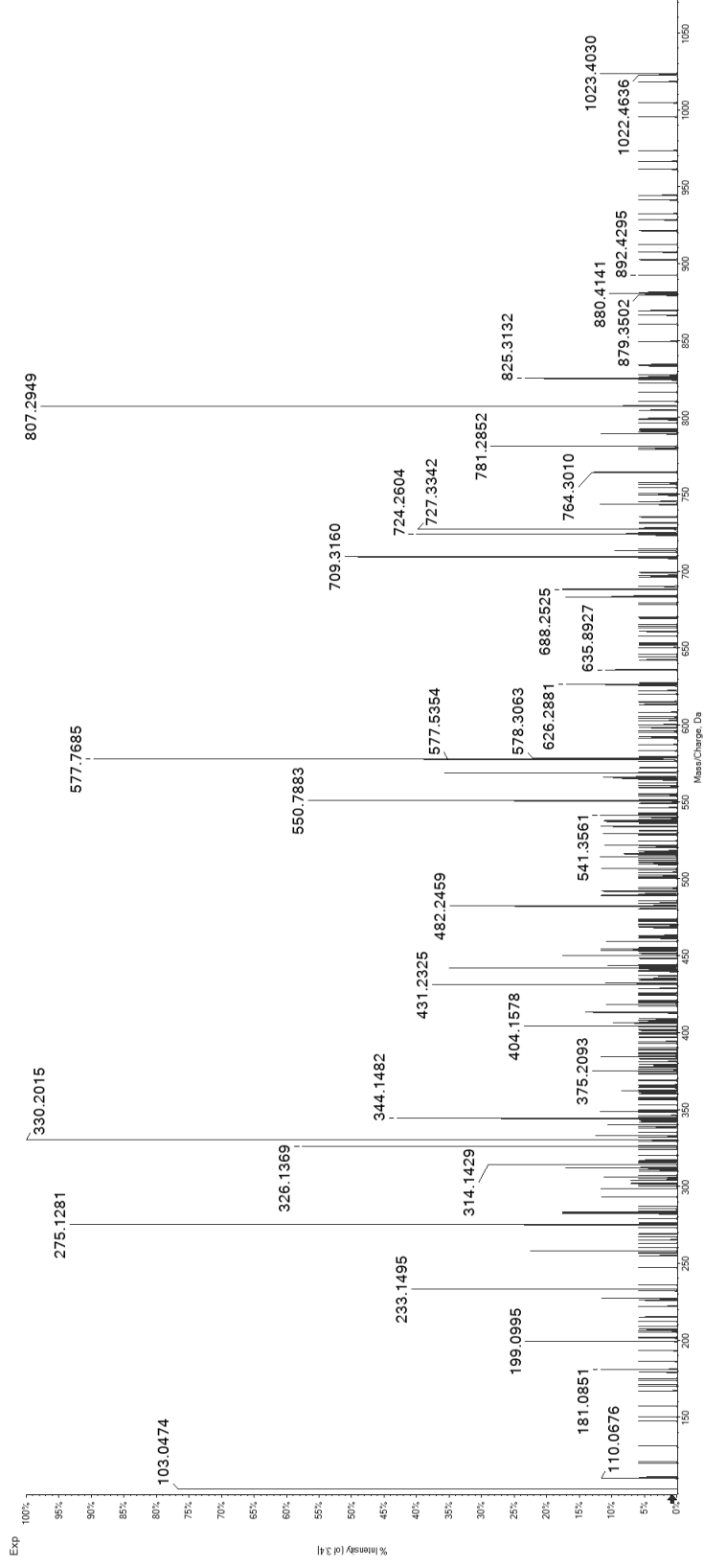
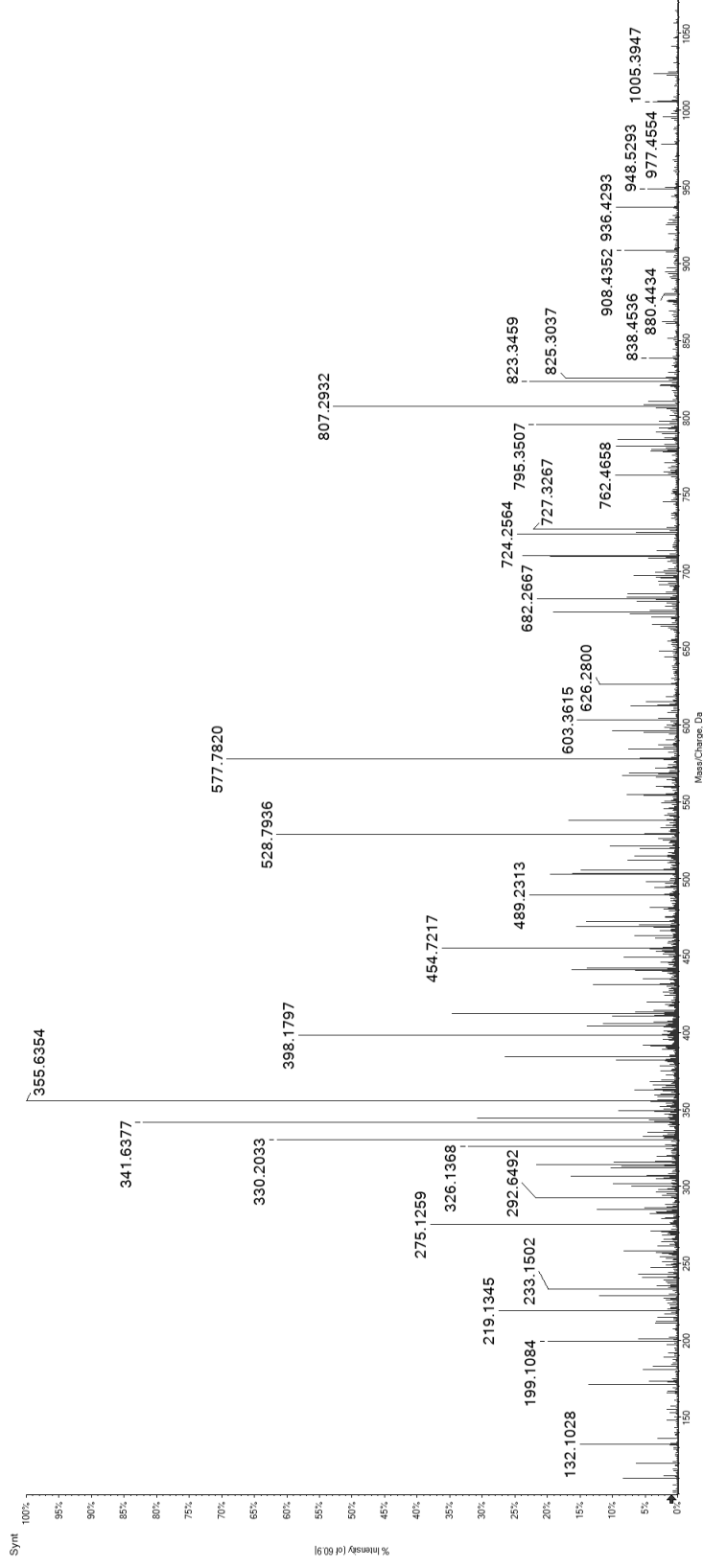
Candidate Spectrum



3

HH(pS)PGQTPTL

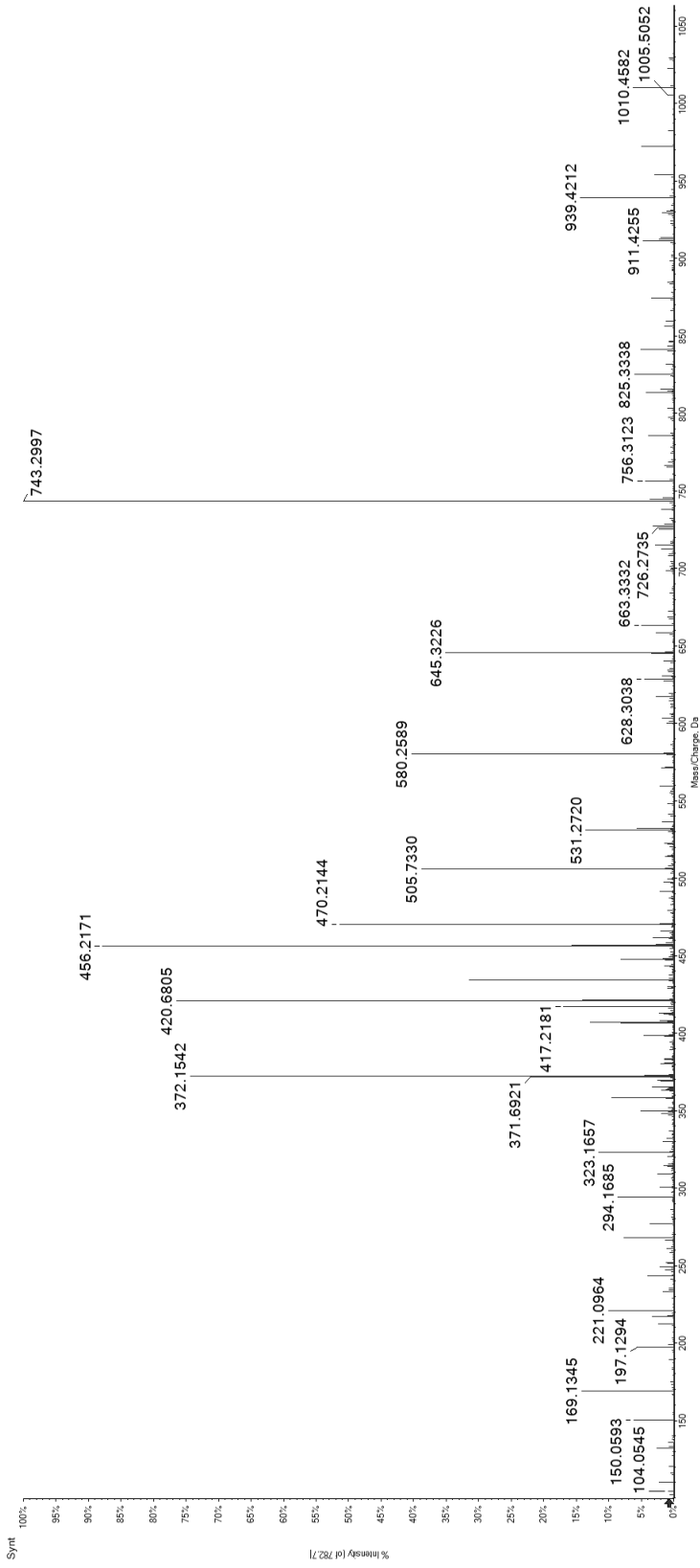
Synthetic Peptide



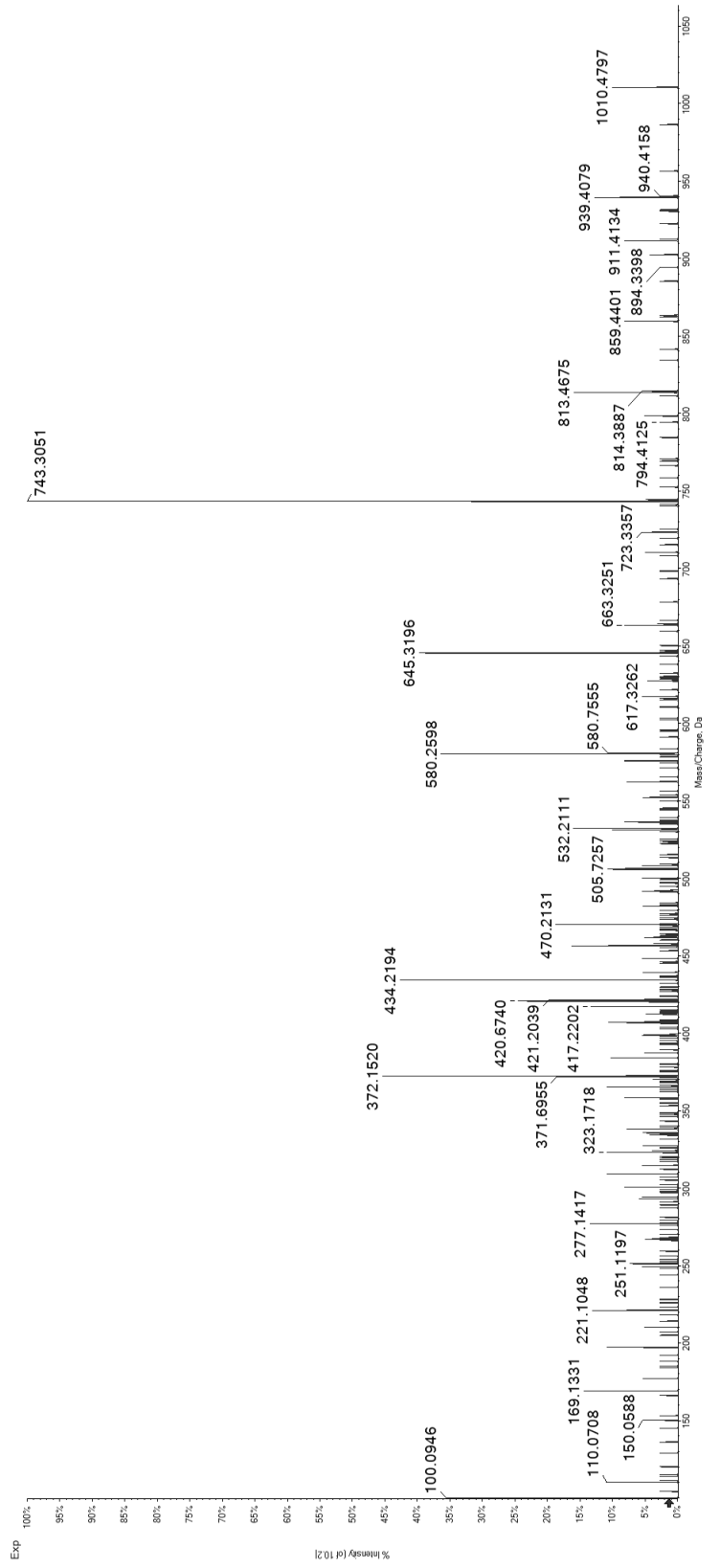
Candidate Spectrum

HRA(pS)PNPVAM

Synthetic
Peptide



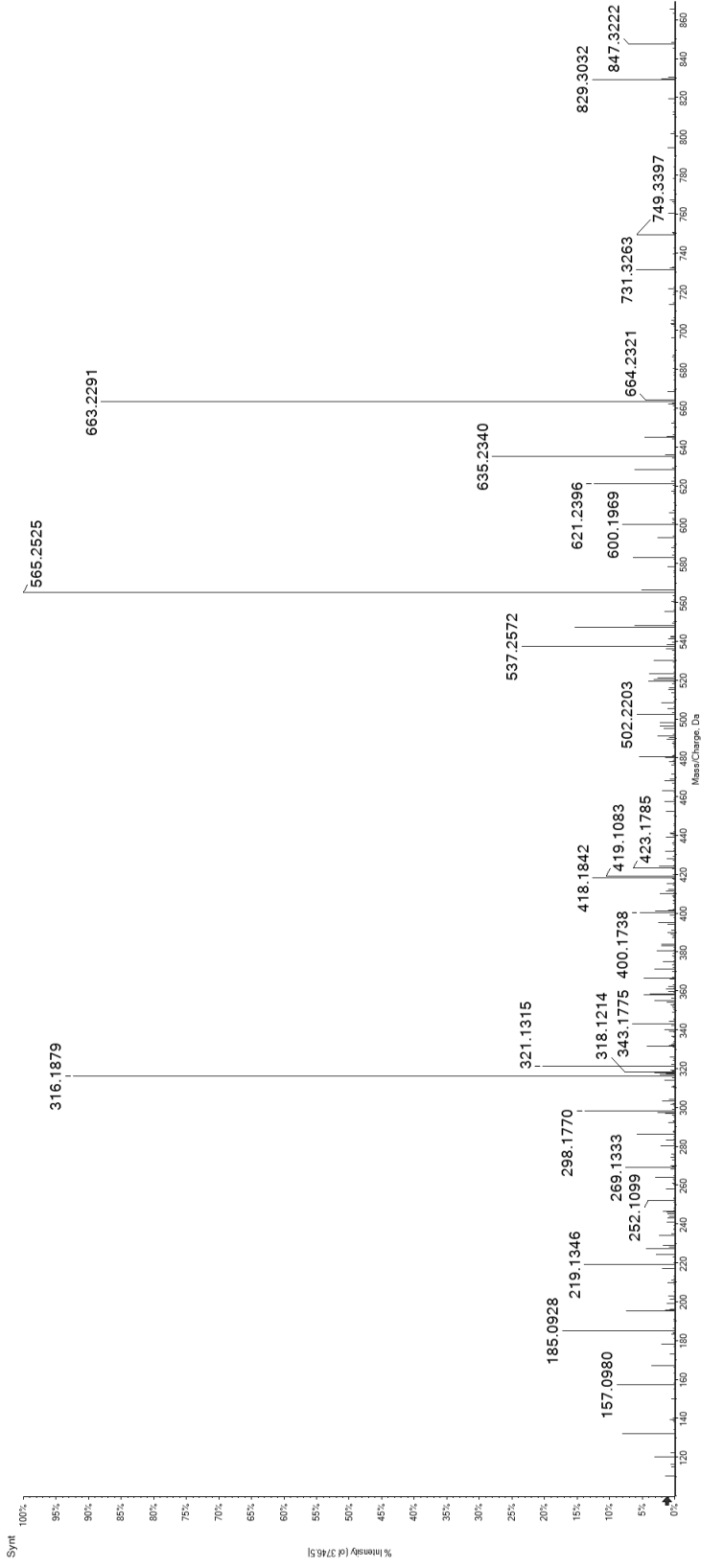
Candidate
Spectrum



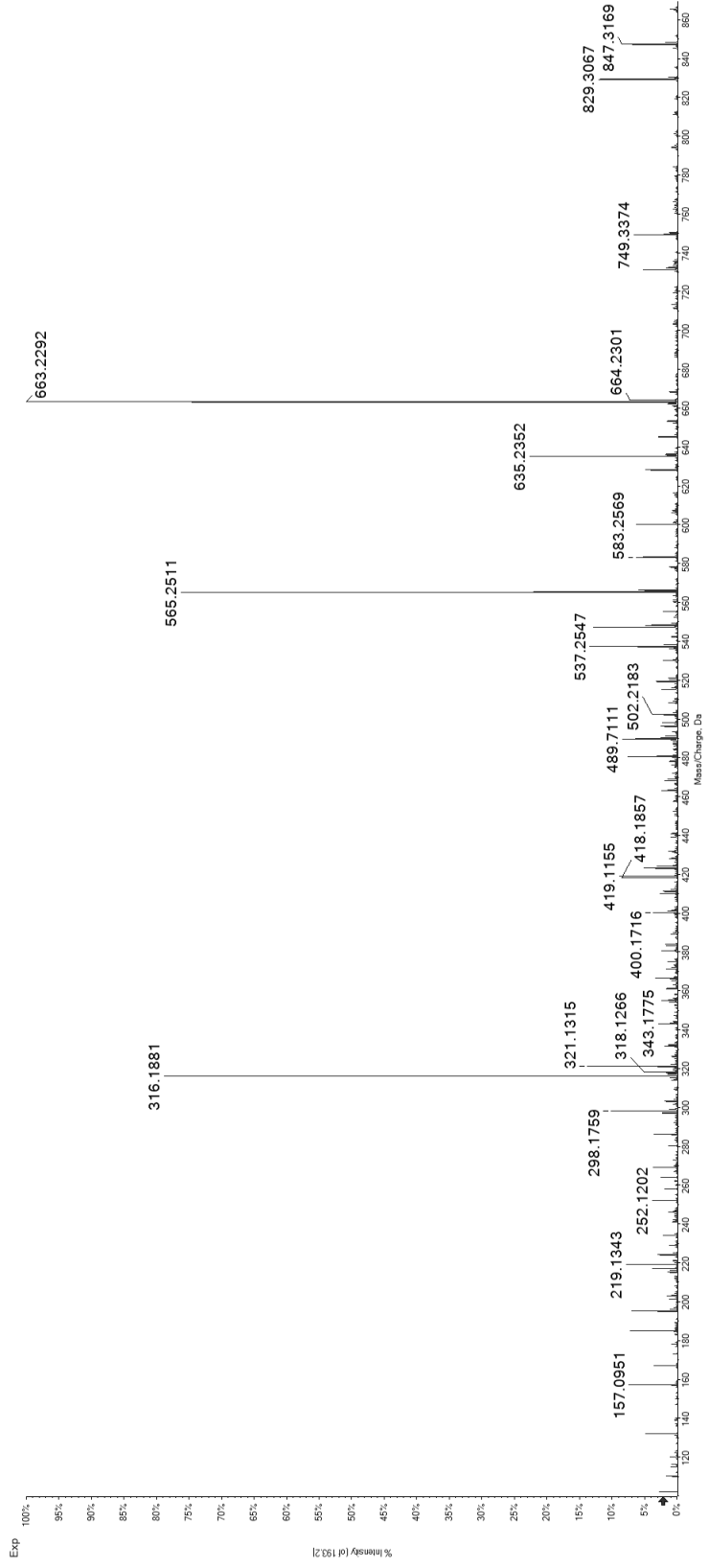
5

GHG(ps)PFPSL

Synthetic Peptide



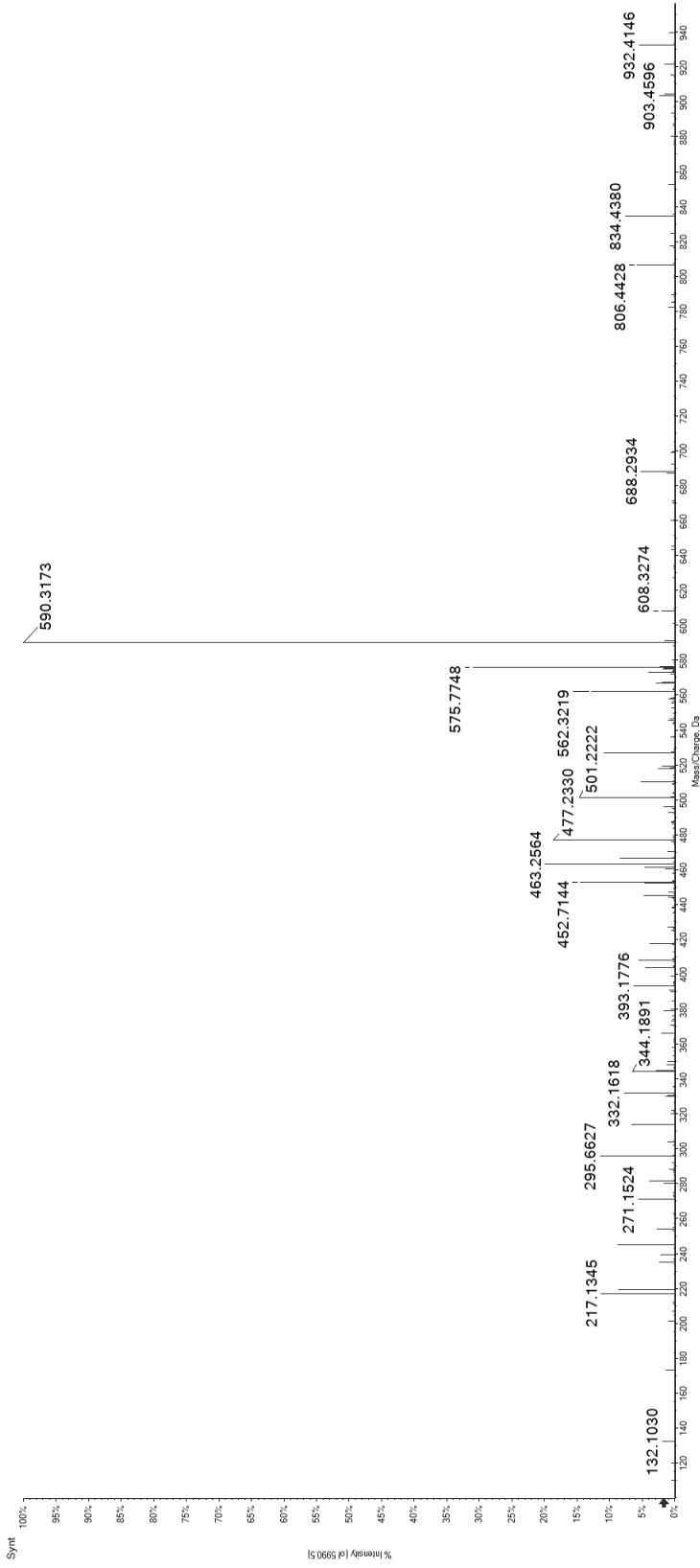
Candidate Spectrum



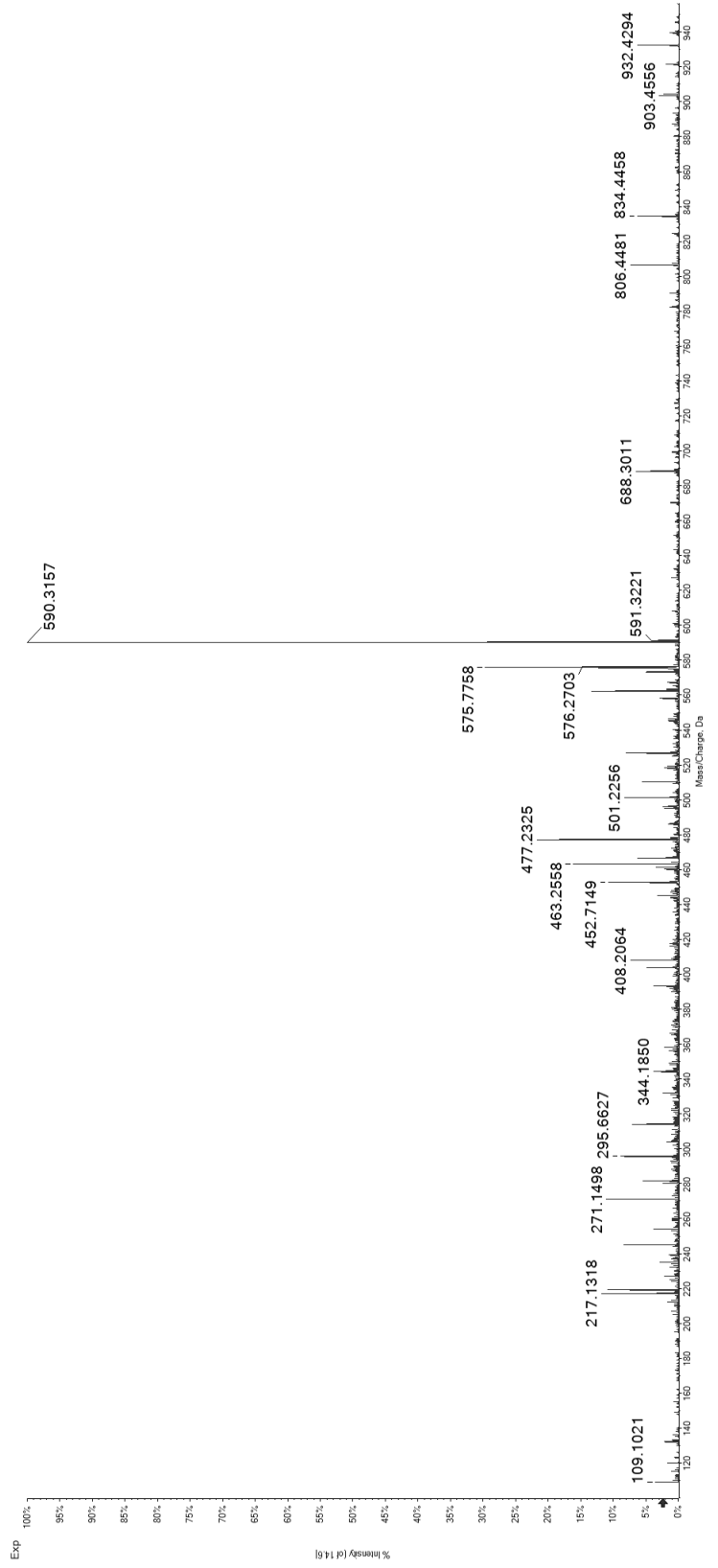
6

NRH(pS)LPFSL

Synthetic Peptide



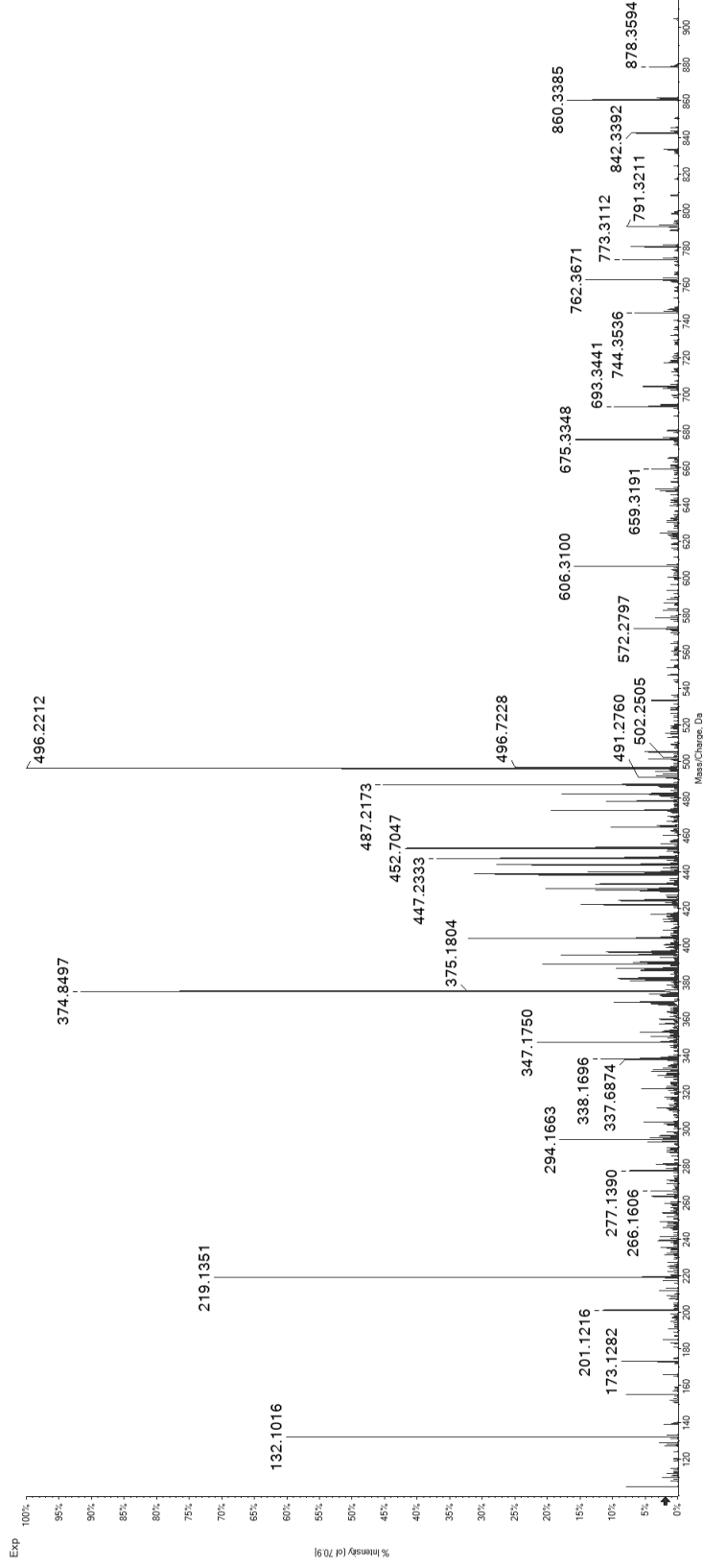
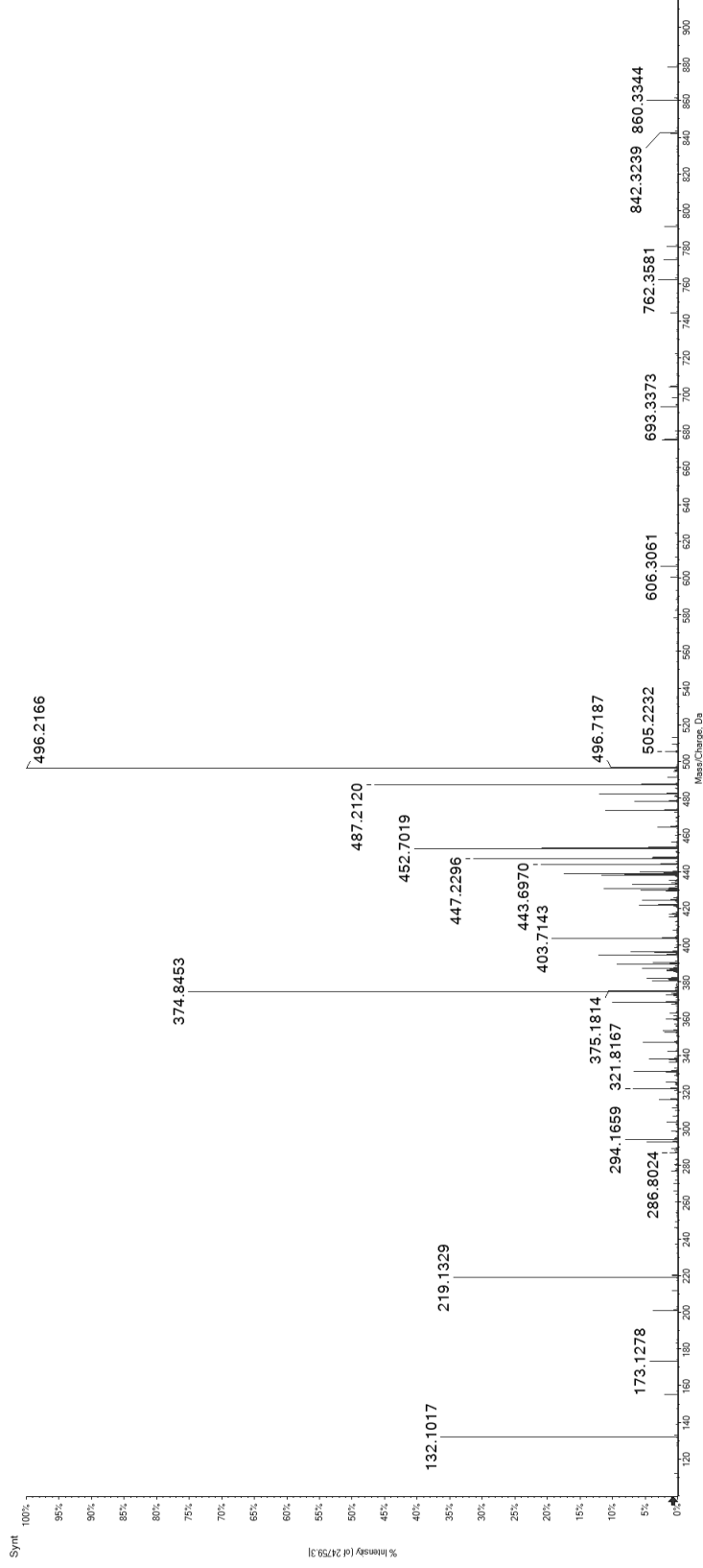
Candidate Spectrum



7

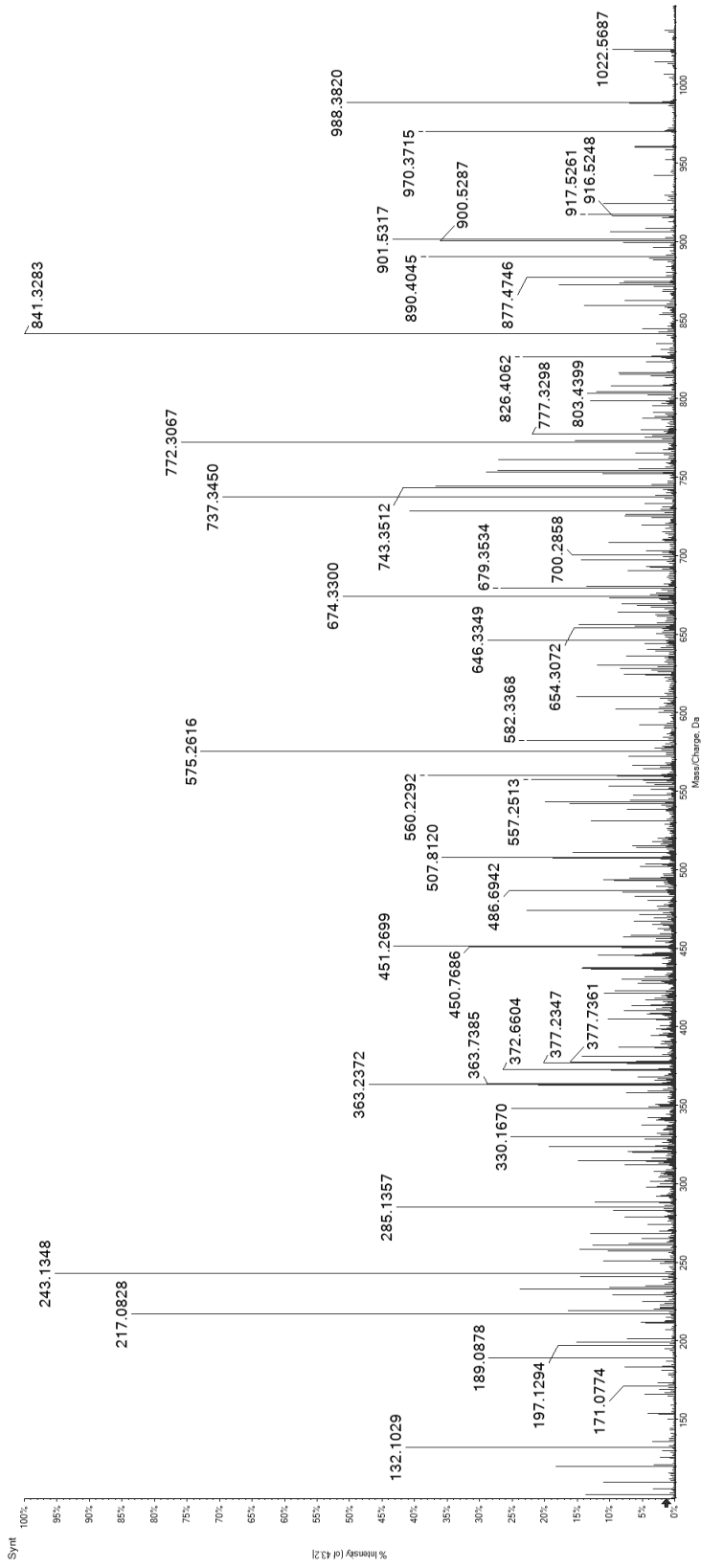
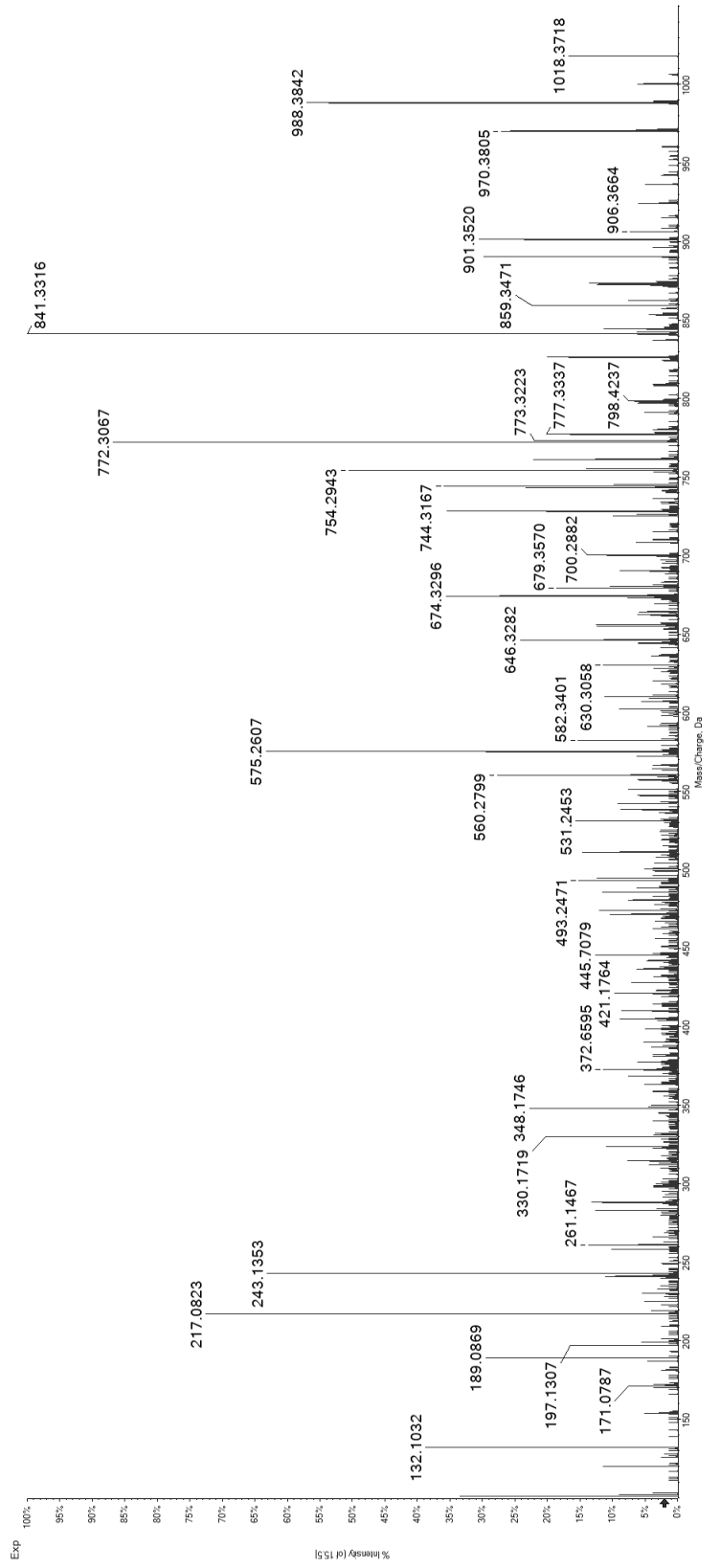
RHK(pS)DSISL

Synthetic Peptide



Candidate Spectrum

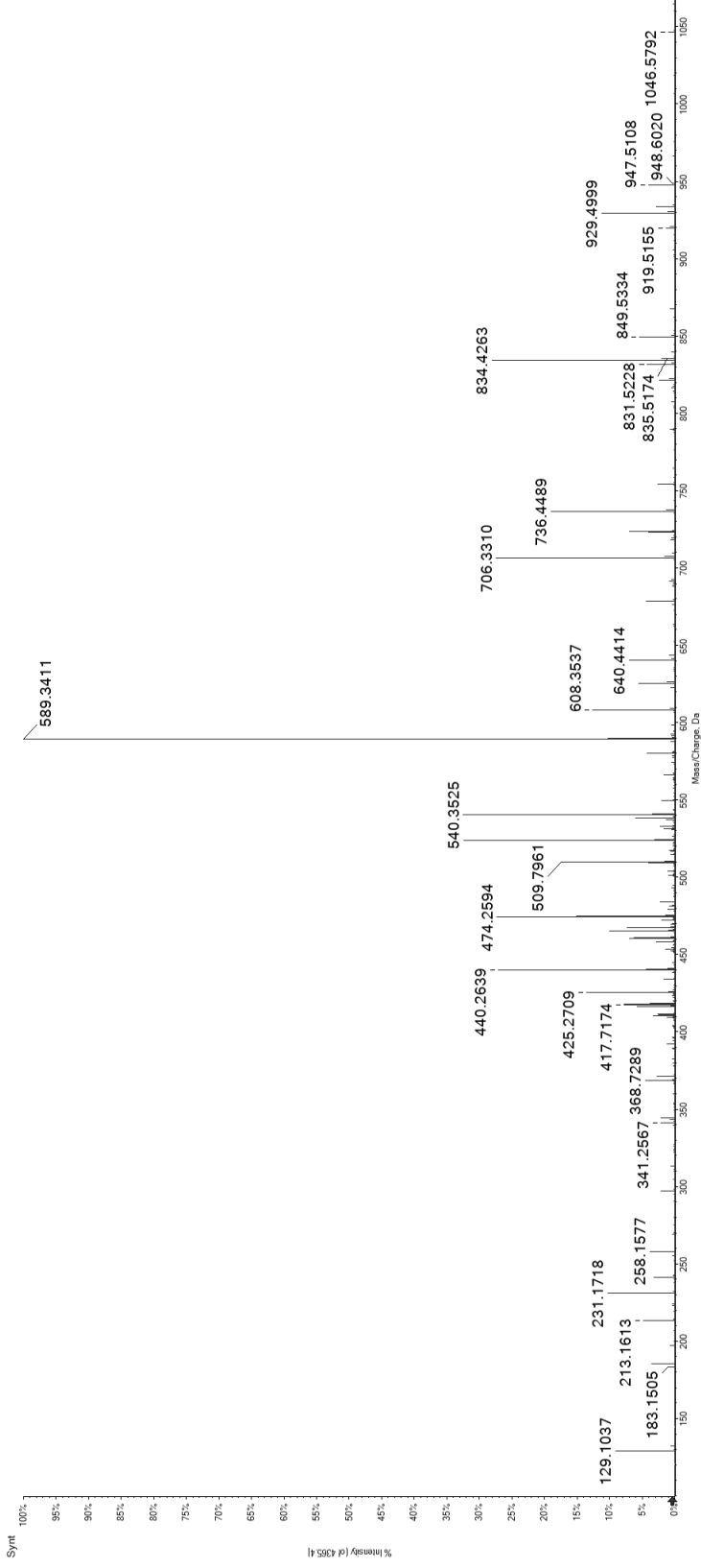
TRM(pS)TVSEL

Synthetic
PeptideCandidate
Spectrum

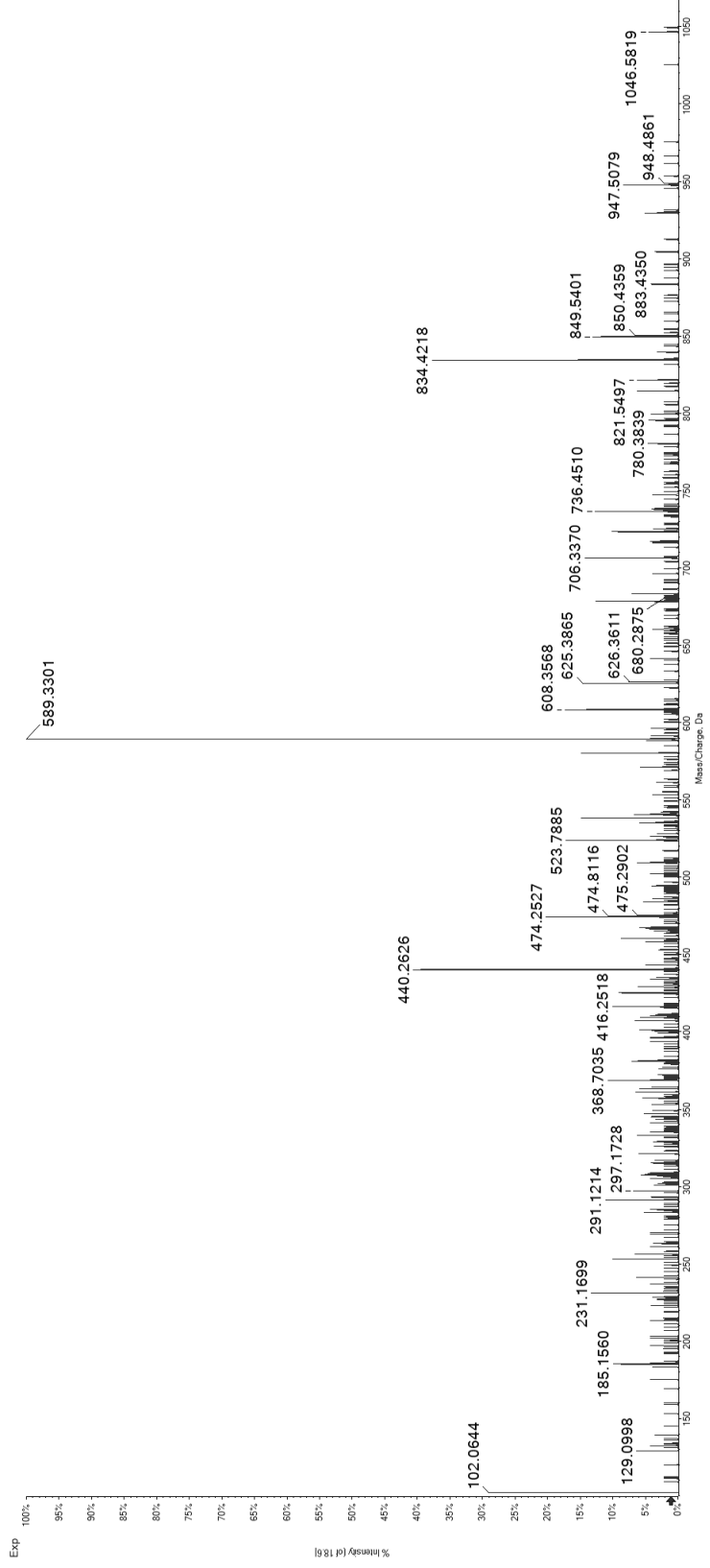
9

TRL(pS)PAKIVL

Synthetic Peptide



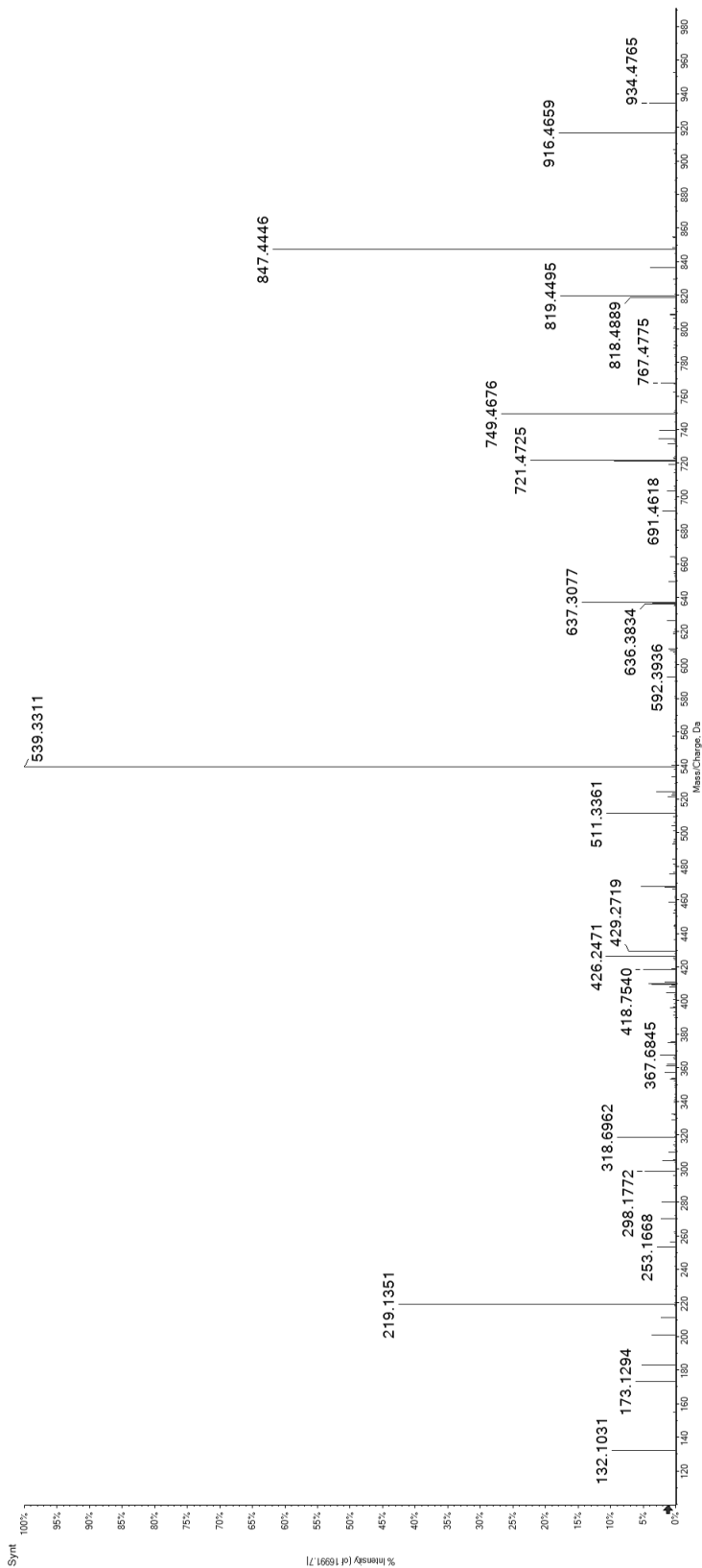
Candidate Spectrum



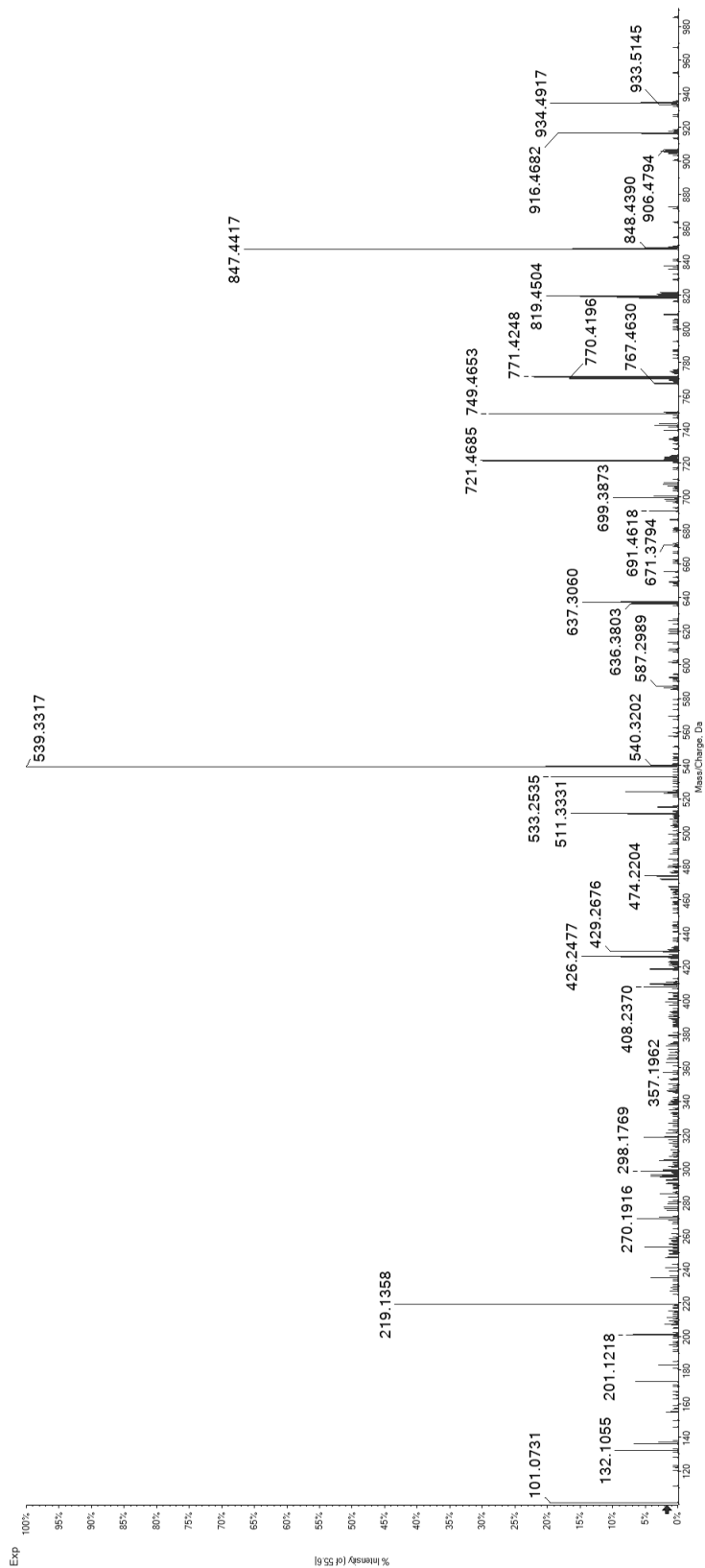
10

LRS(ps)LPISL

Synthetic Peptide



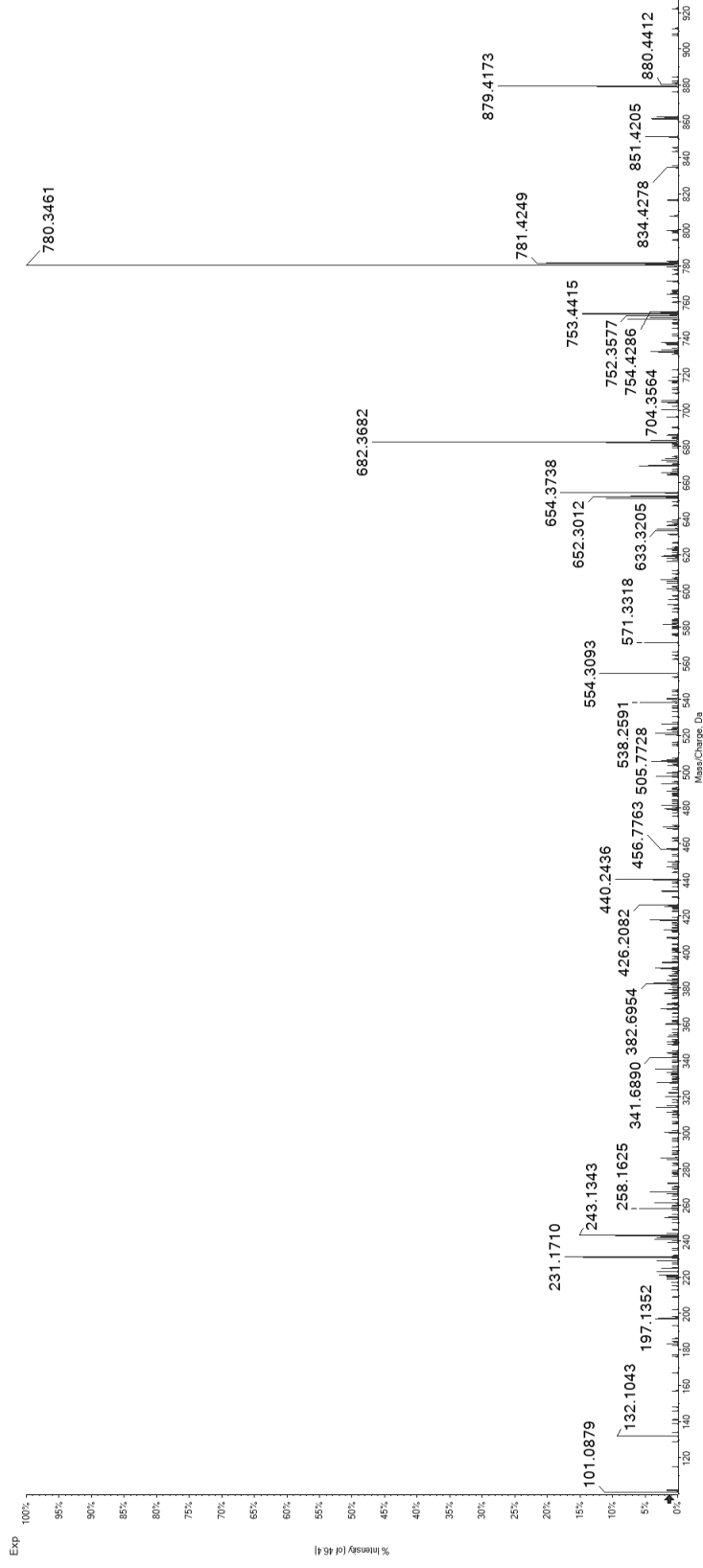
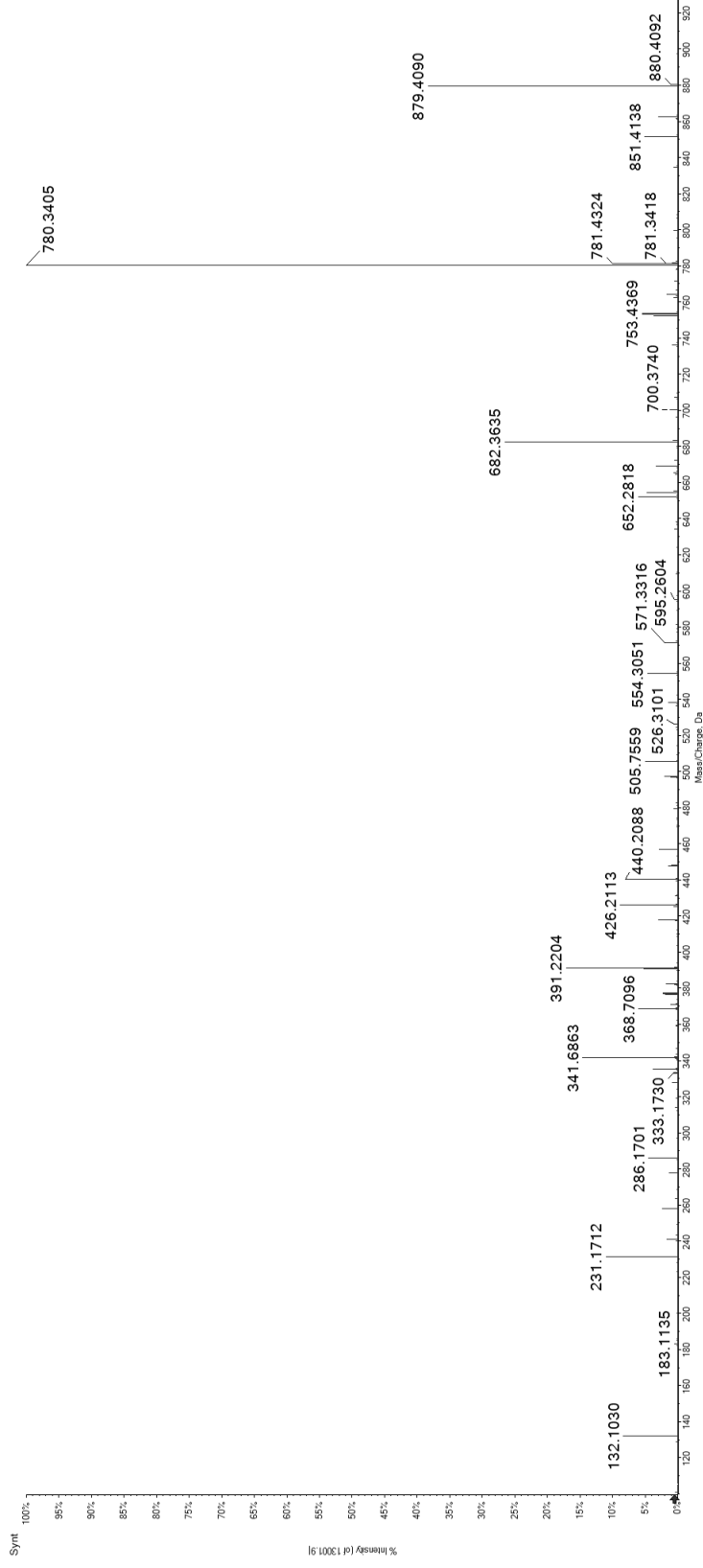
Candidate Spectrum



11

TRL(ps)GGQVL

Synthetic Peptide

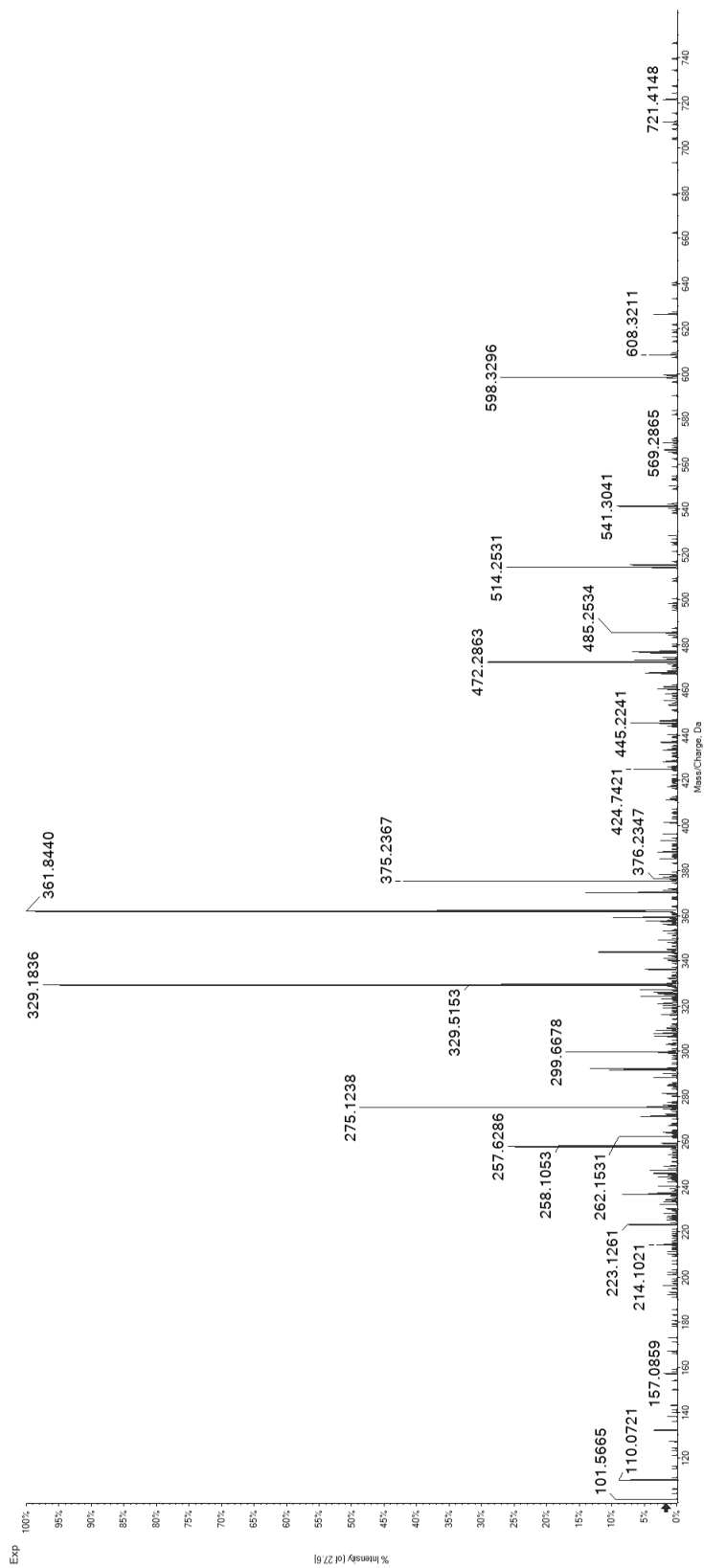
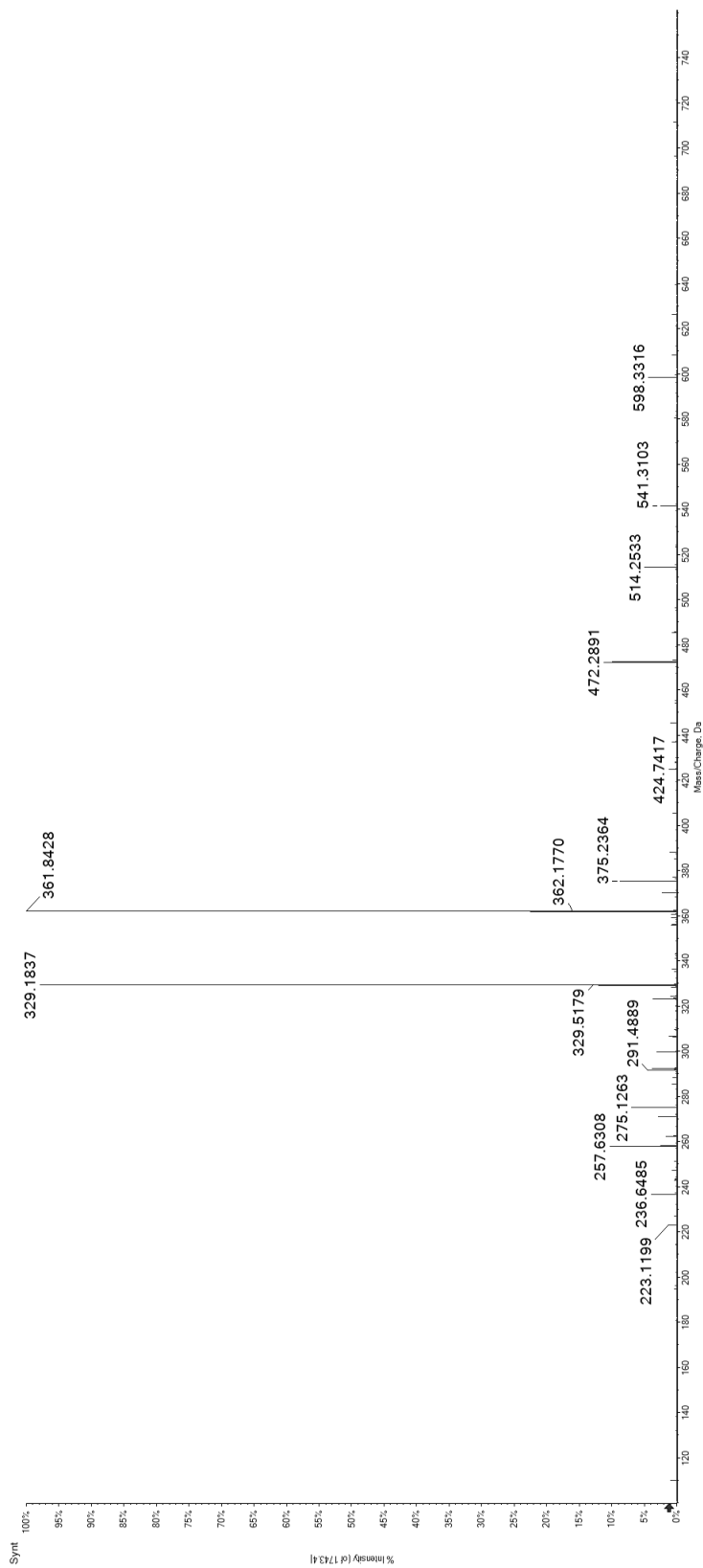


Candidate Spectrum

12

HHLG(ps)PSRL

Synthetic Peptide

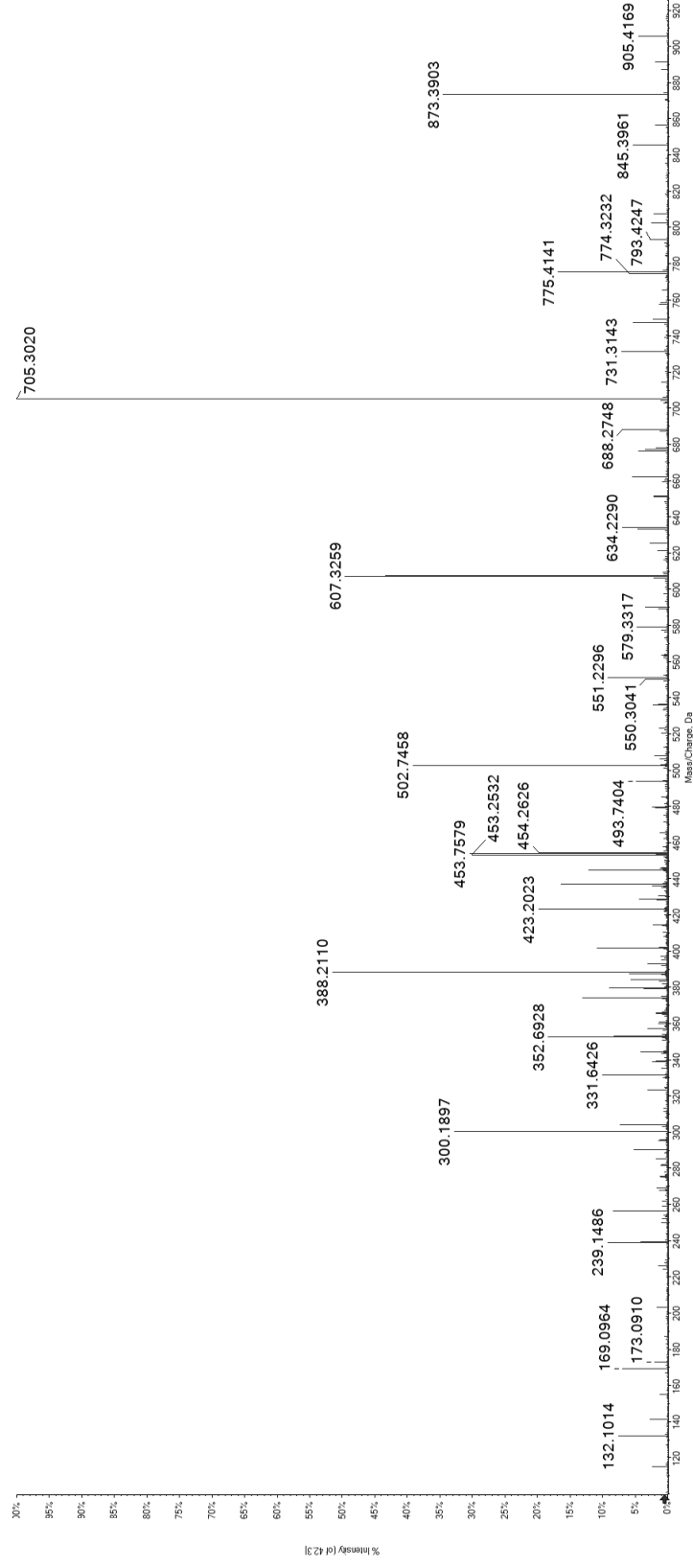


Candidate Spectrum

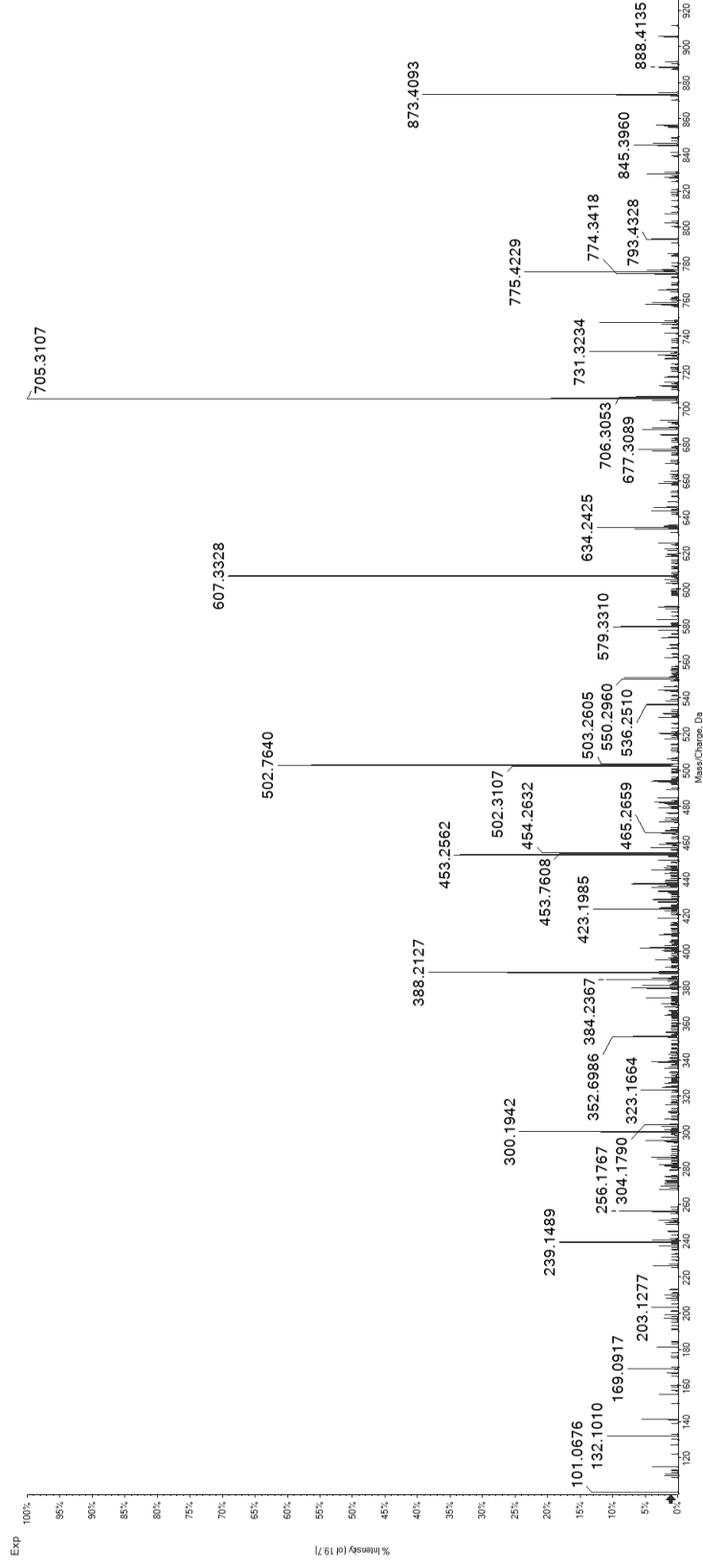
13

VRQ(ps)PGPAL

Synthetic Peptide



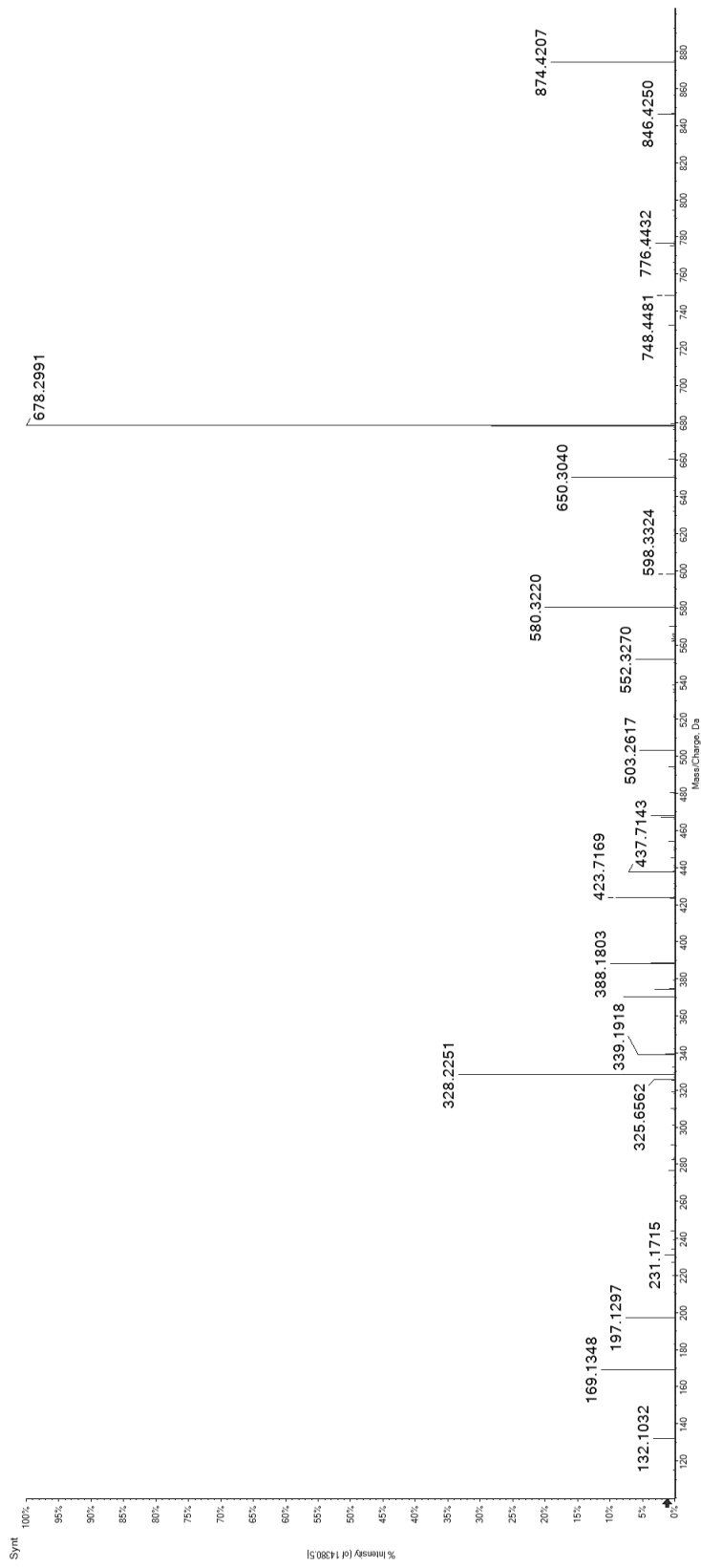
Candidate Spectrum



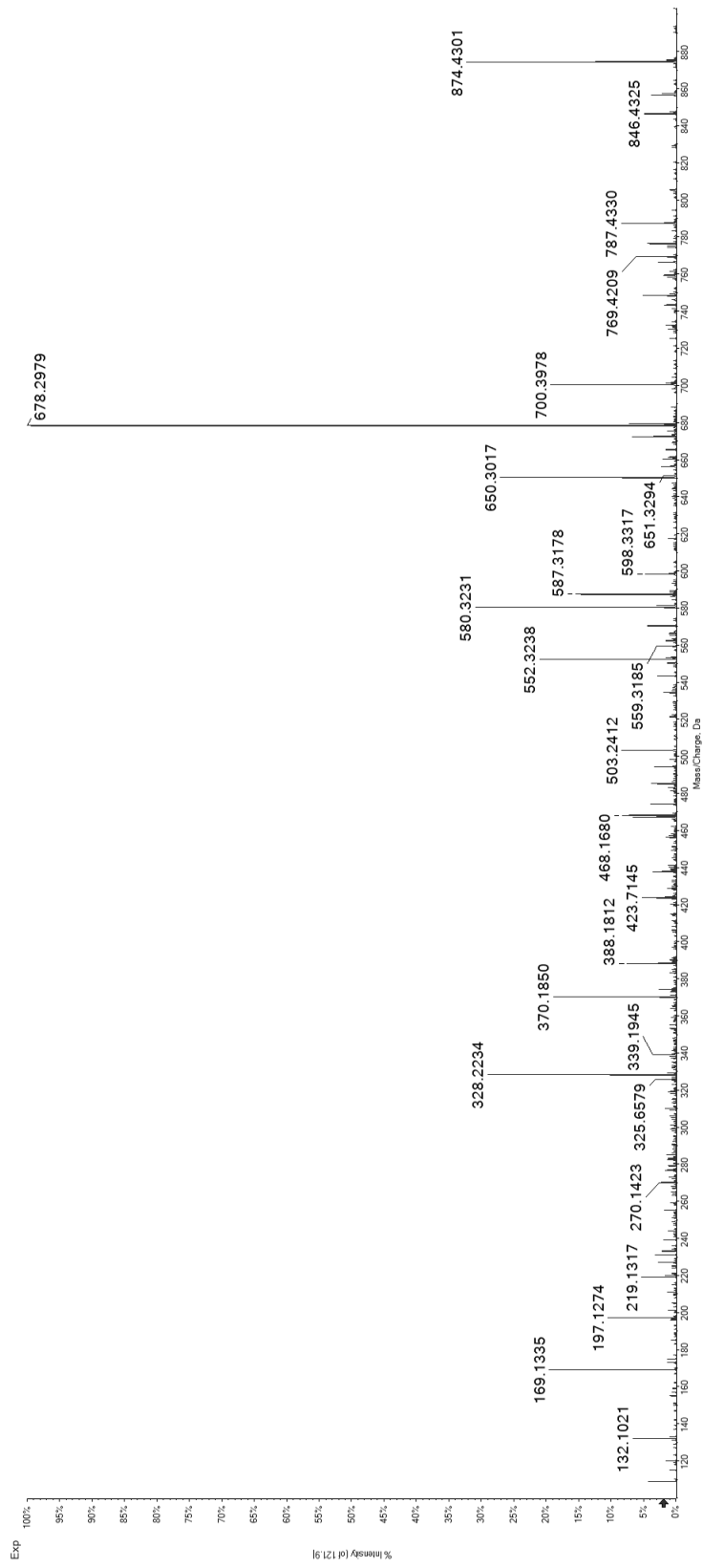
14

SRG(pS)PLPVL

Synthetic Peptide

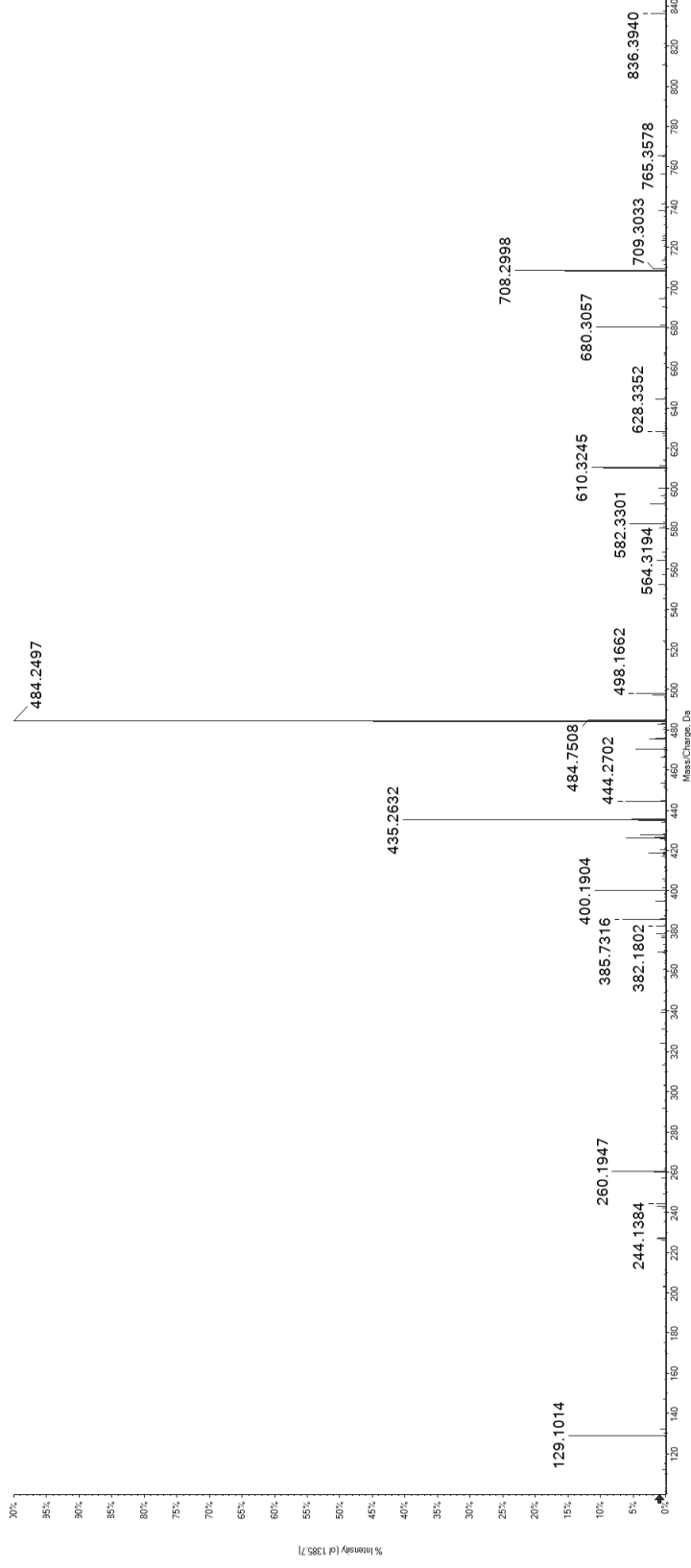


Candidate Spectrum

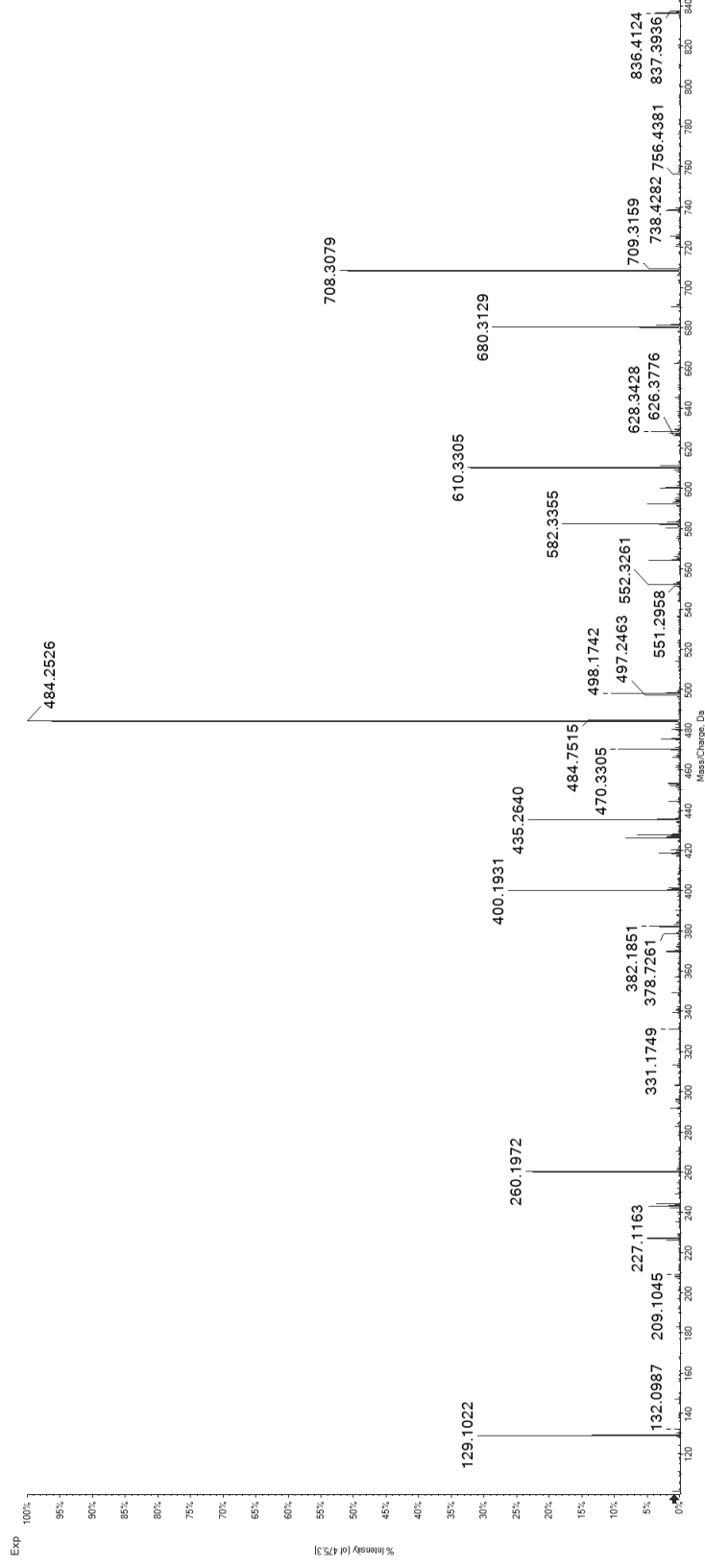


15 SRS(ps)PLKL

Synthetic
Peptide



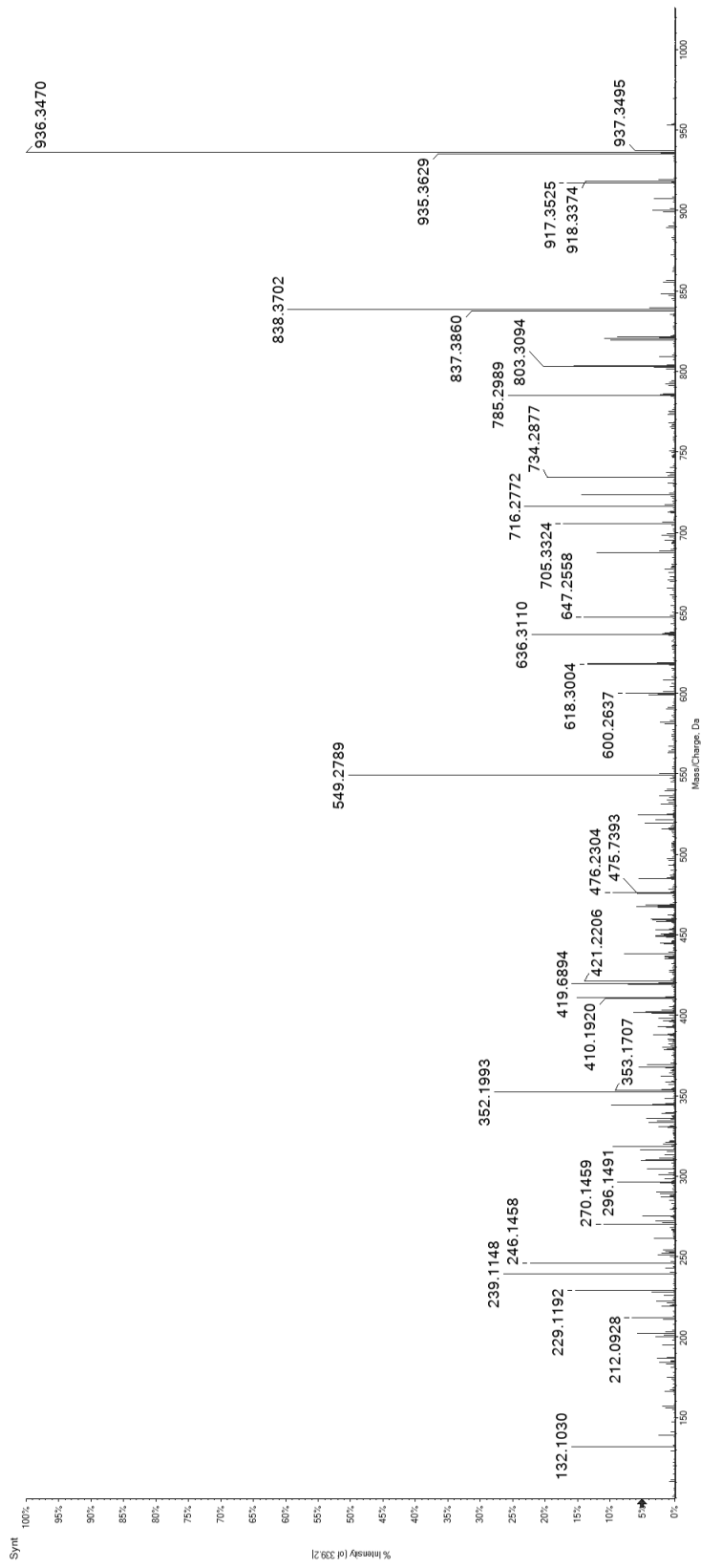
Candidate
Spectrum



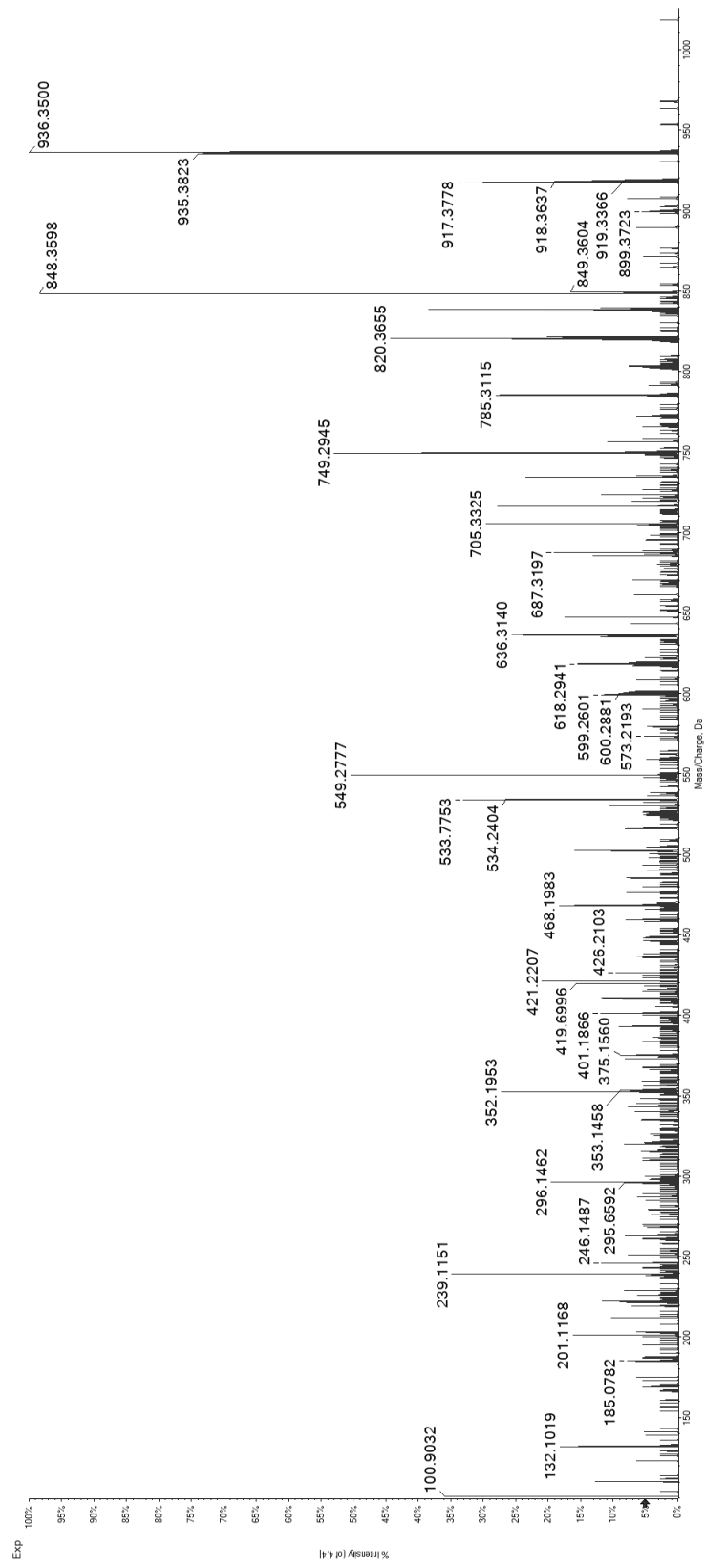
16

THL(pS)QSSNL

Synthetic Peptide



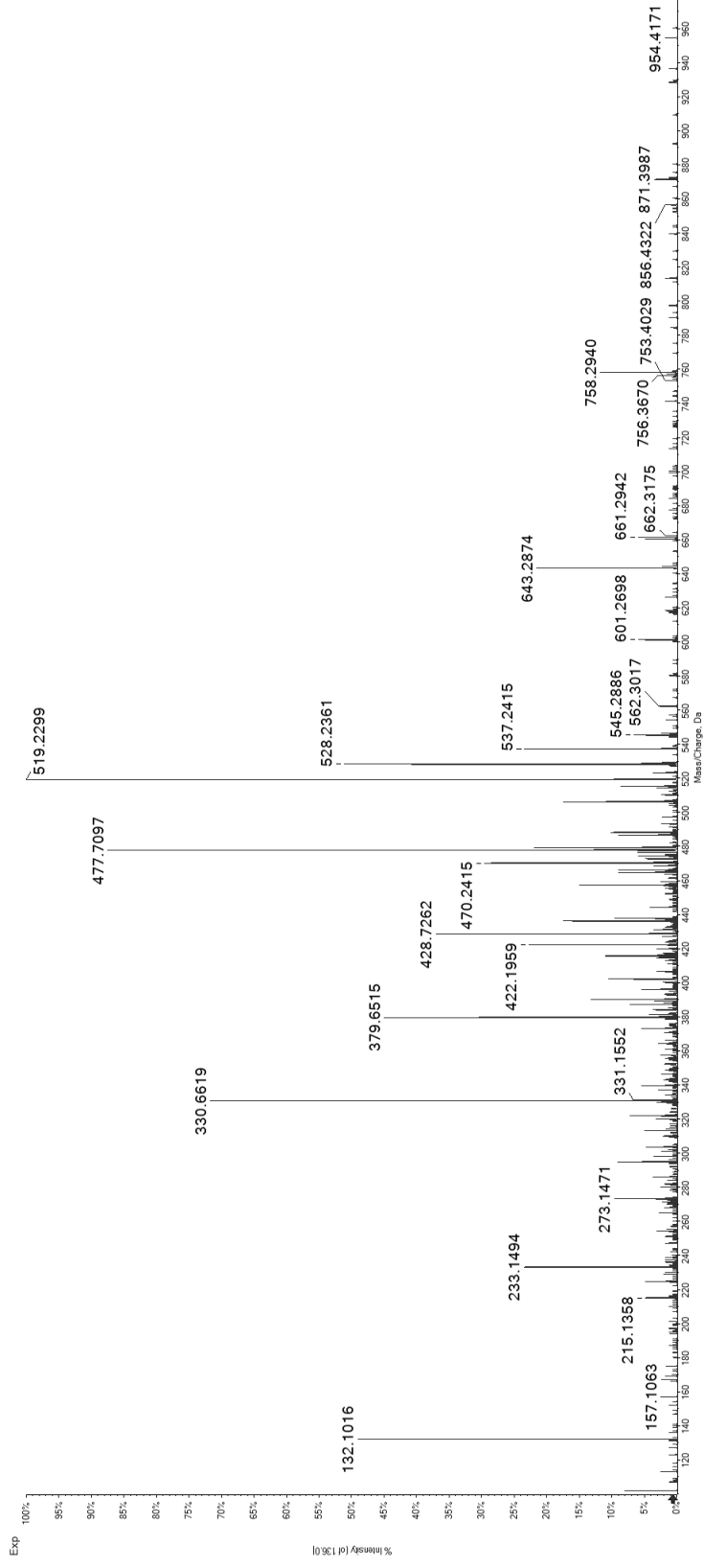
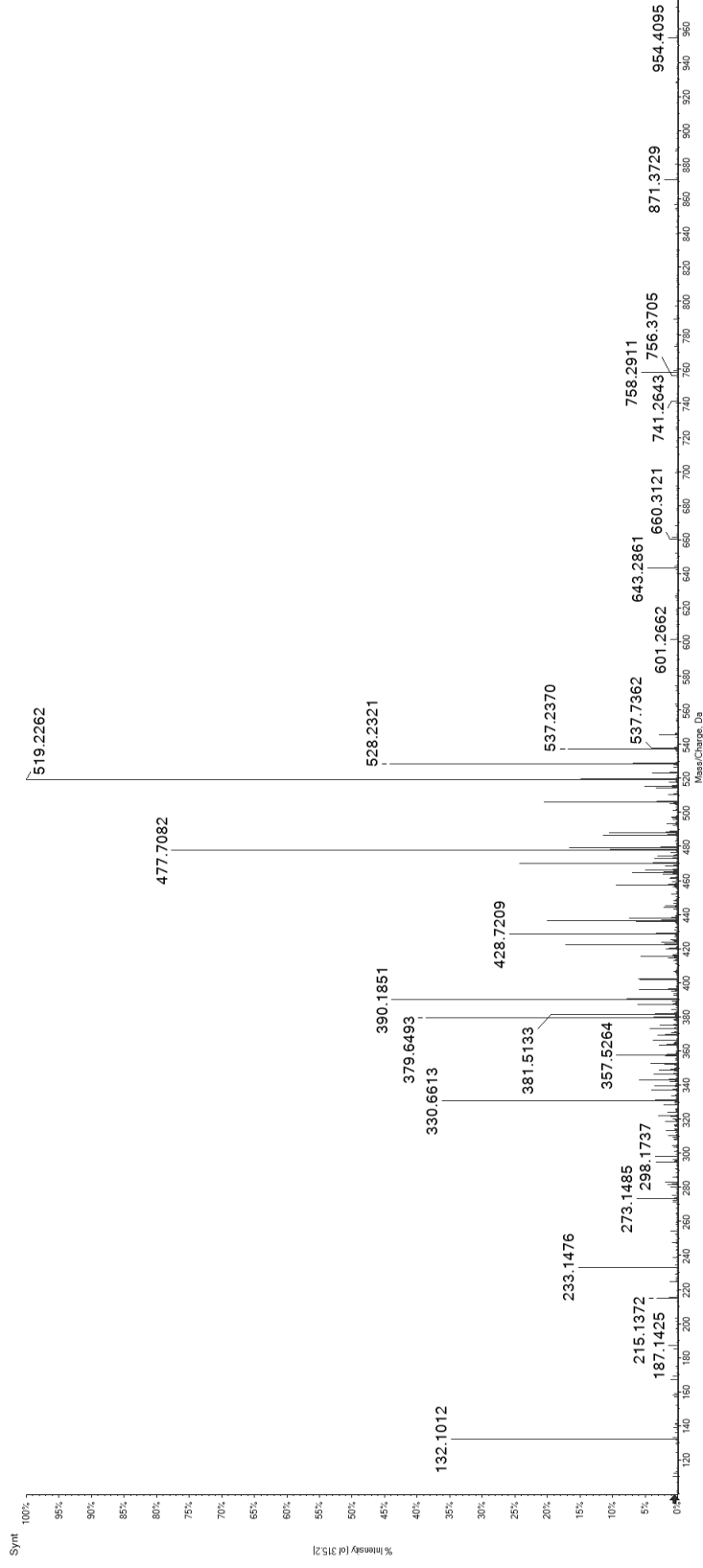
Candidate Spectrum



17

RRY(pS)DLTTL

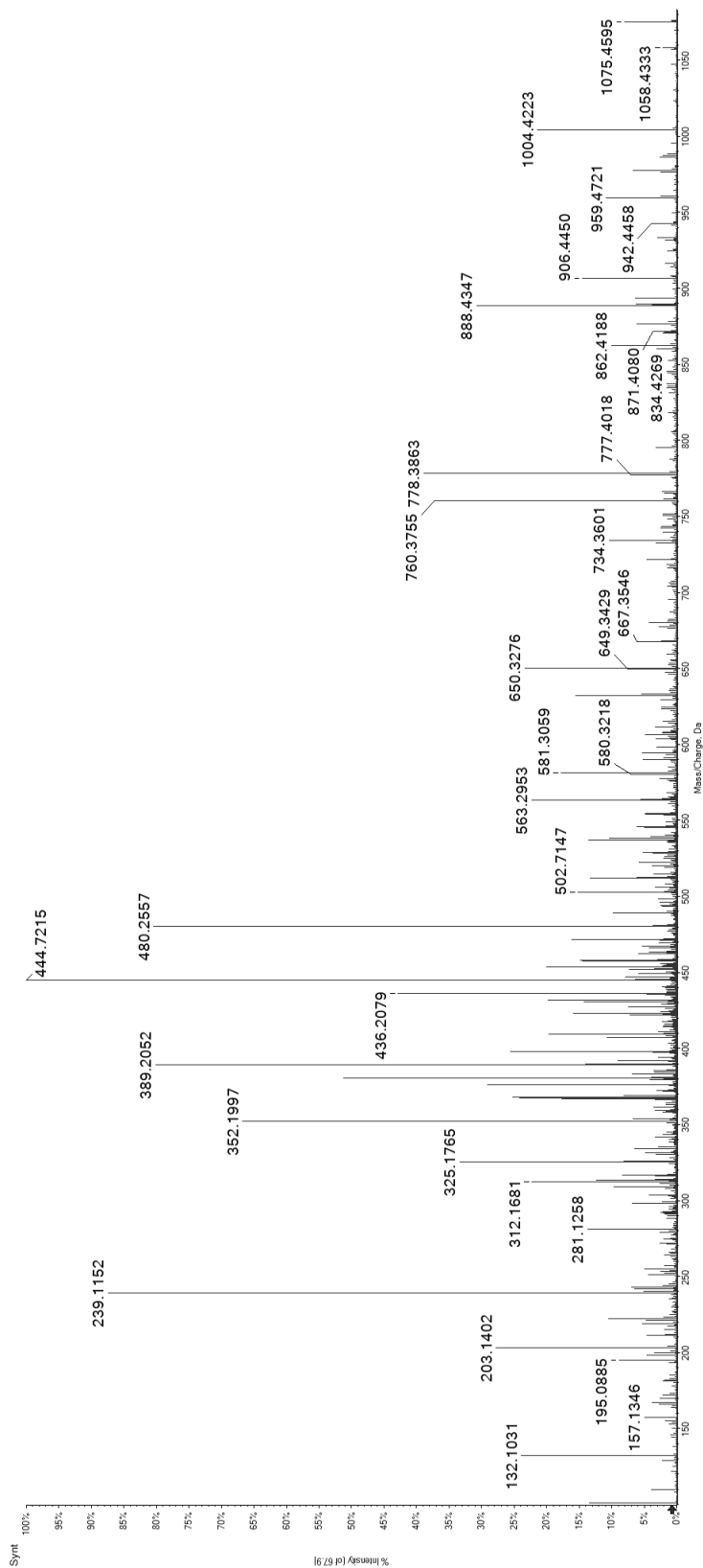
Synthetic Peptide



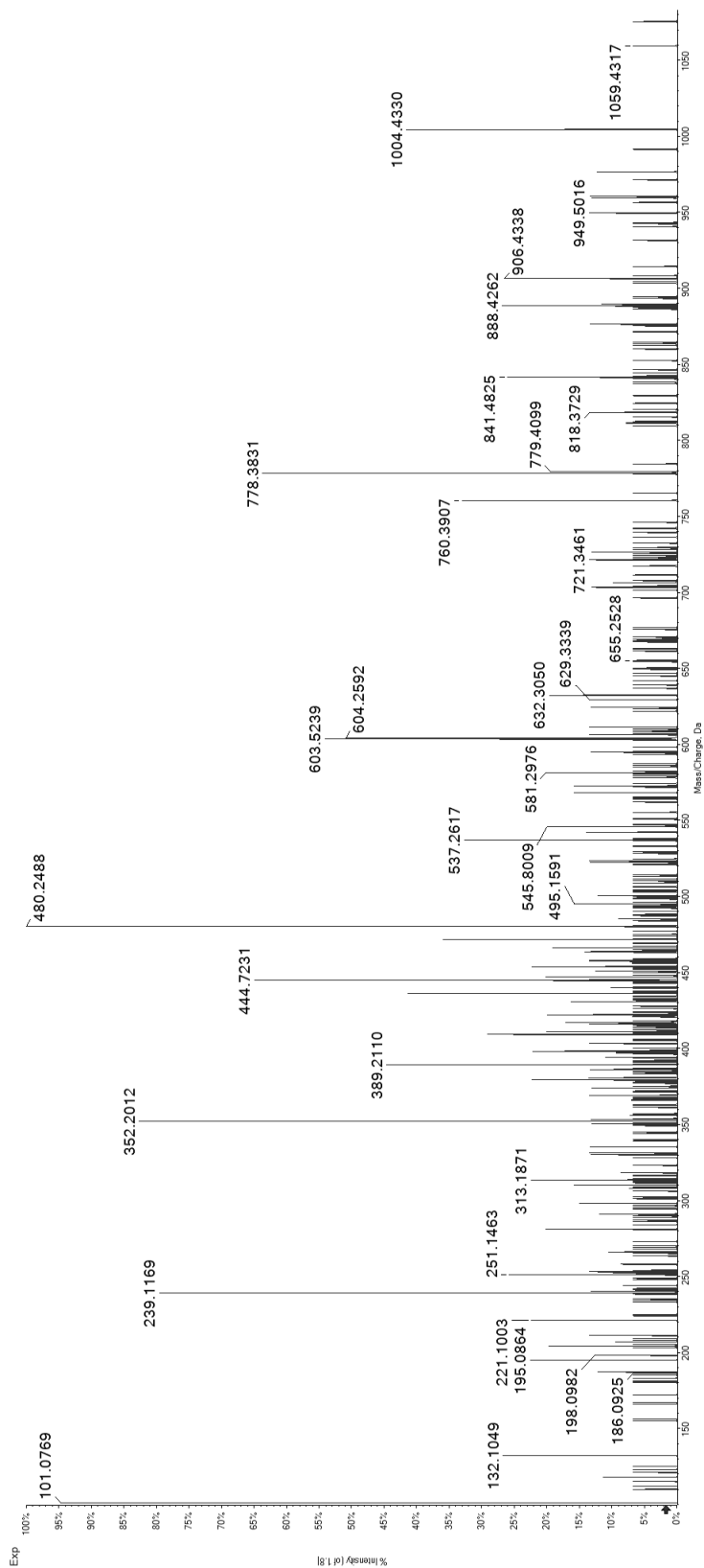
Candidate Spectrum

THIQT(ps)QQAL

Synthetic
Peptide



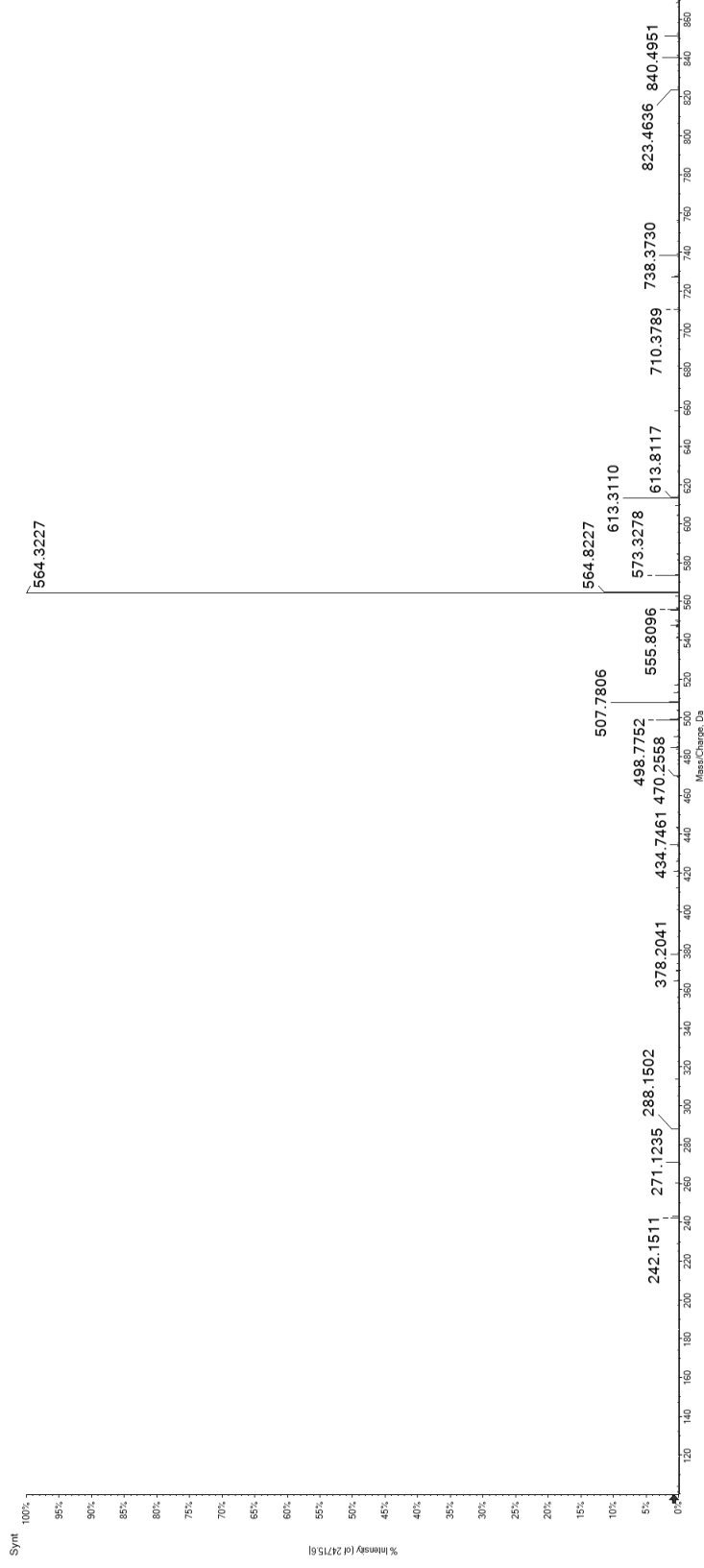
Candidate
Spectrum



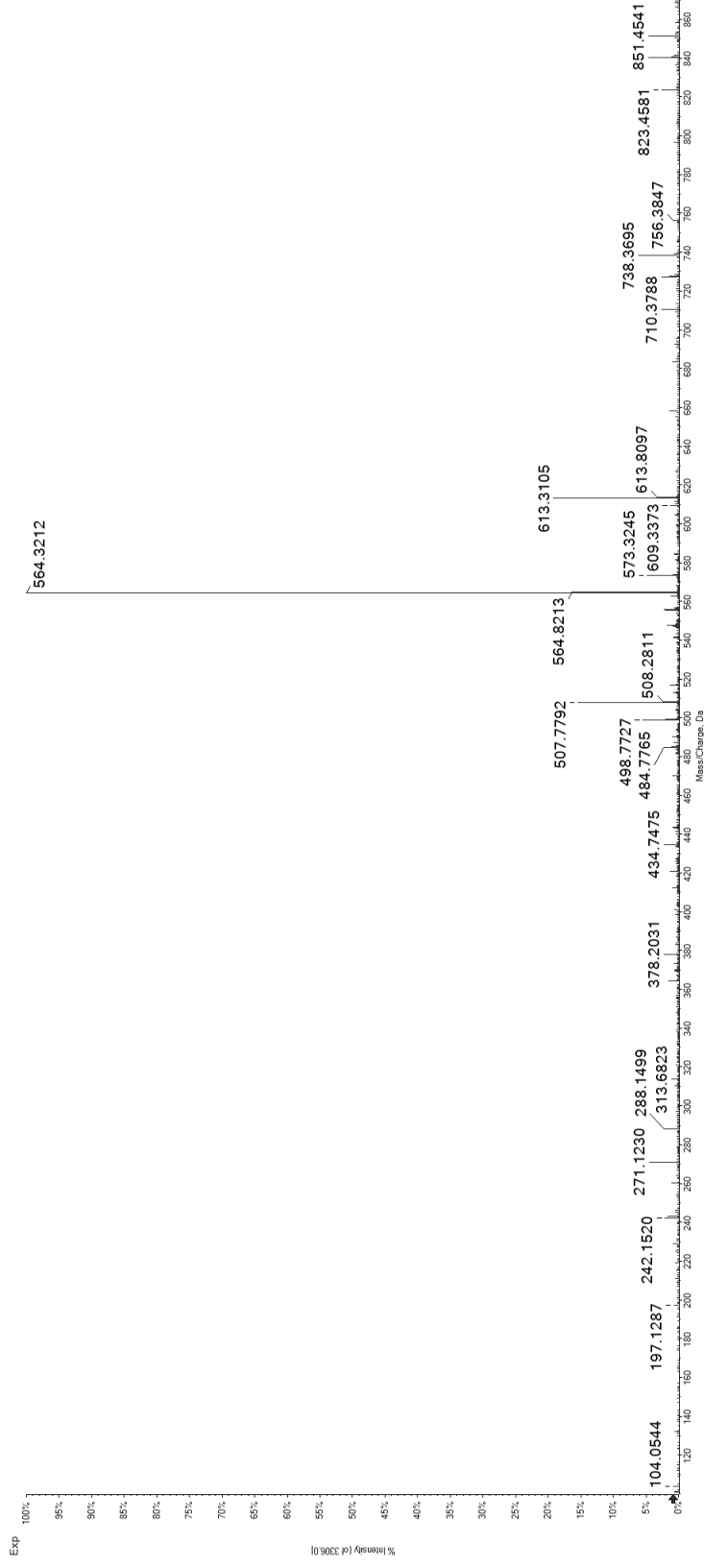
19

MRL(pS)RELQL

Synthetic Peptide



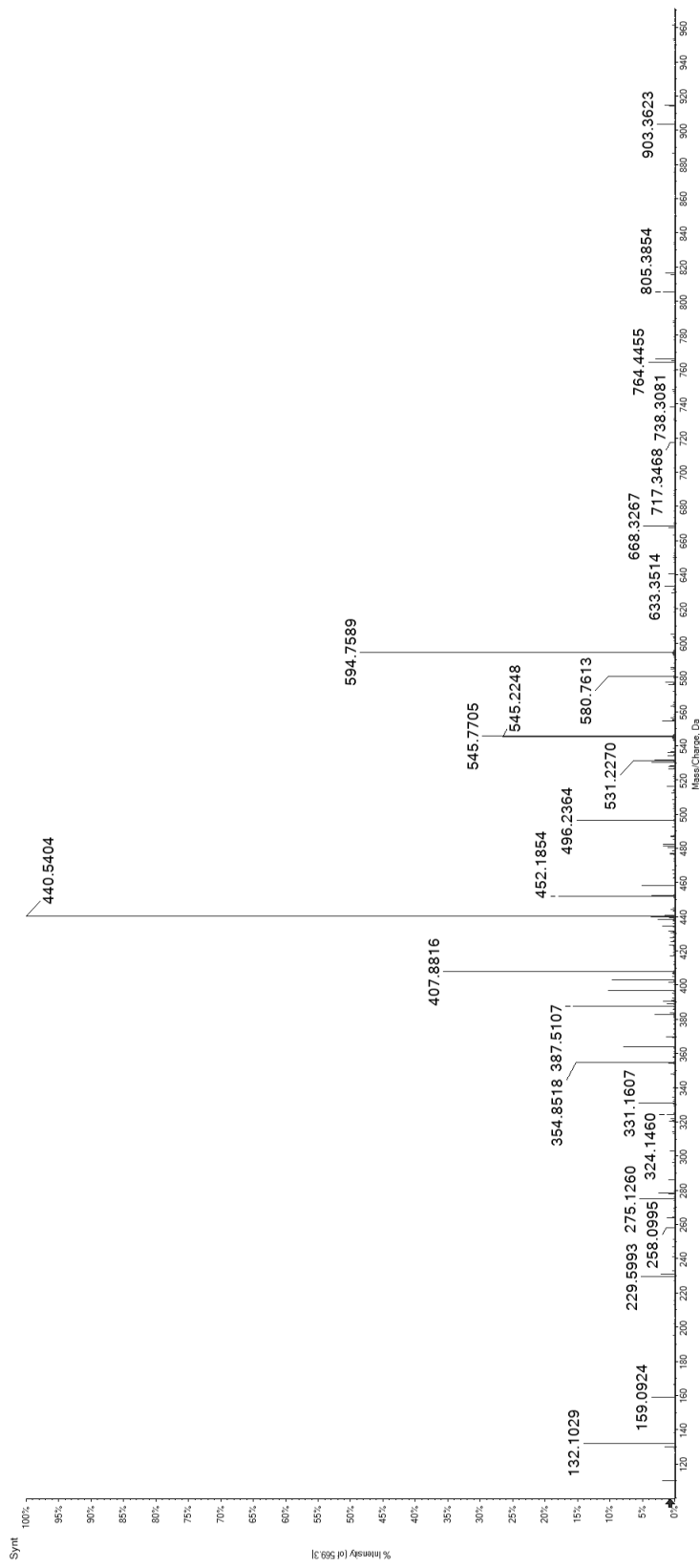
Candidate Spectrum



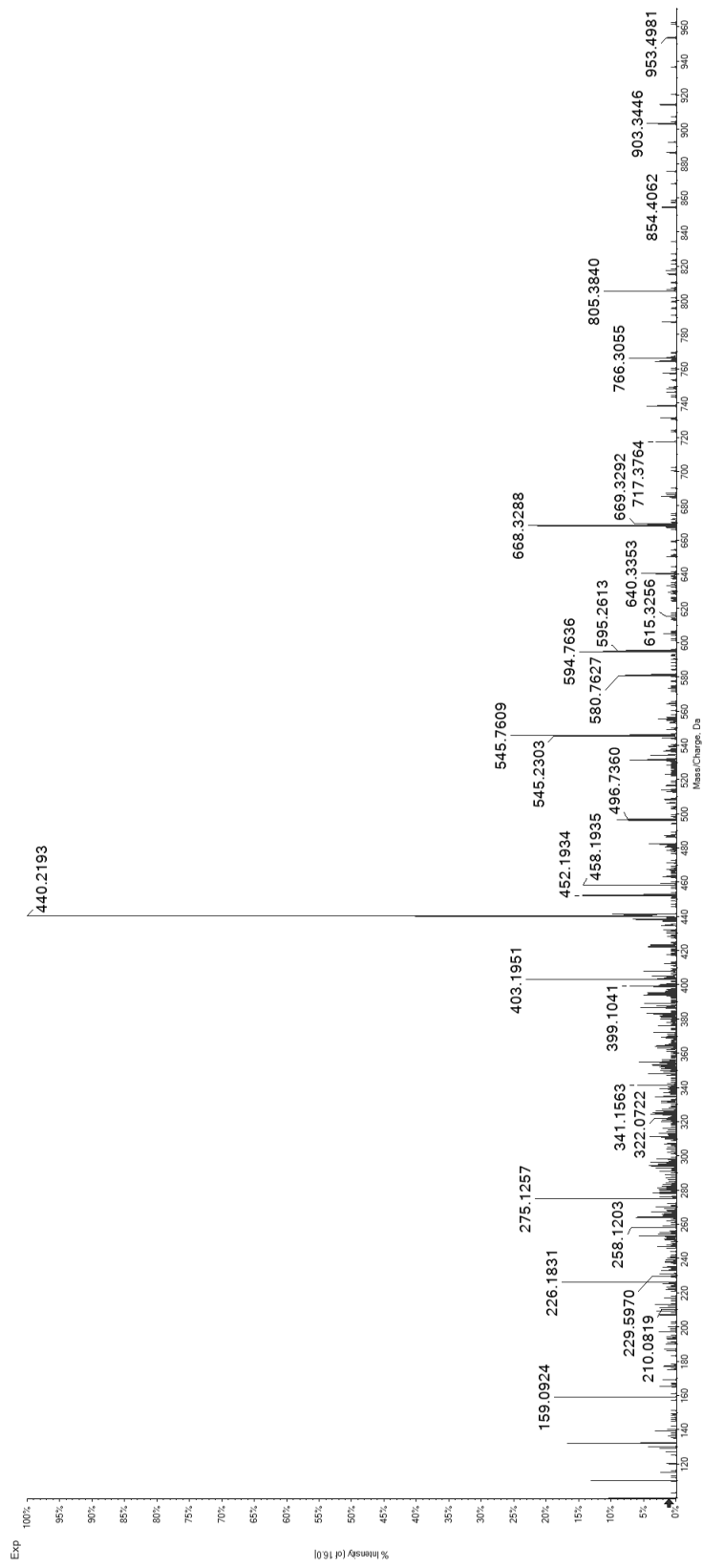
20

HHGG(ps)PIHWVL

Synthetic Peptide



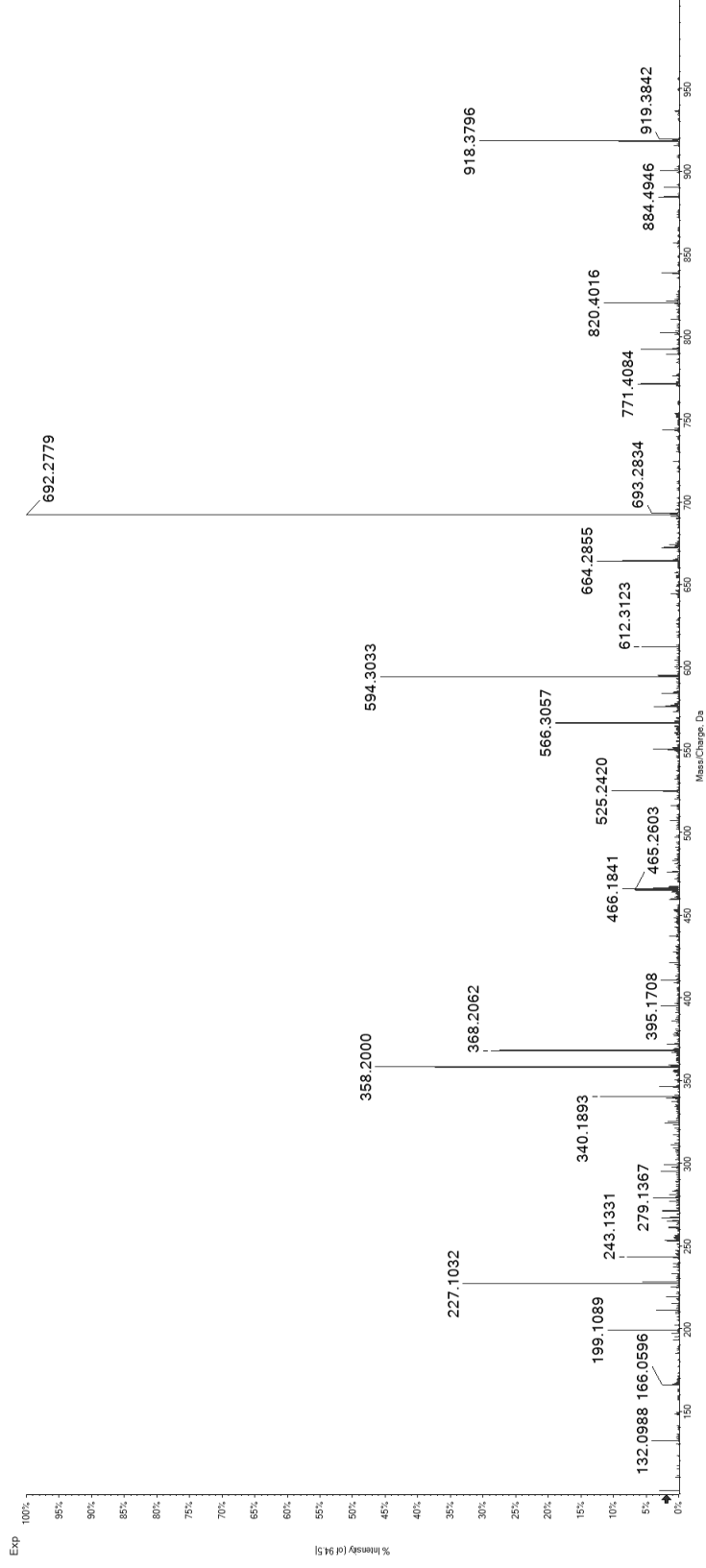
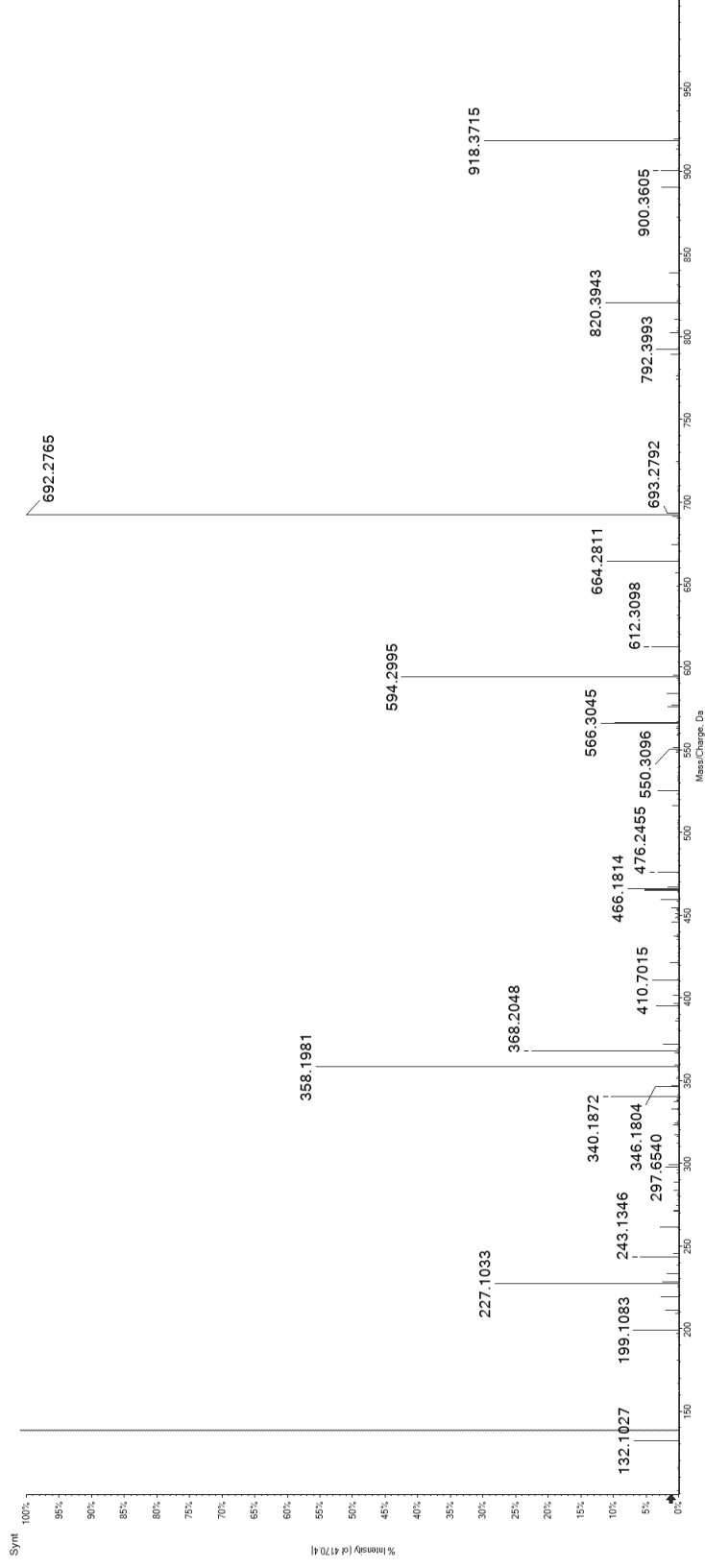
Candidate Spectrum



21

ARA(pS)PEPEL

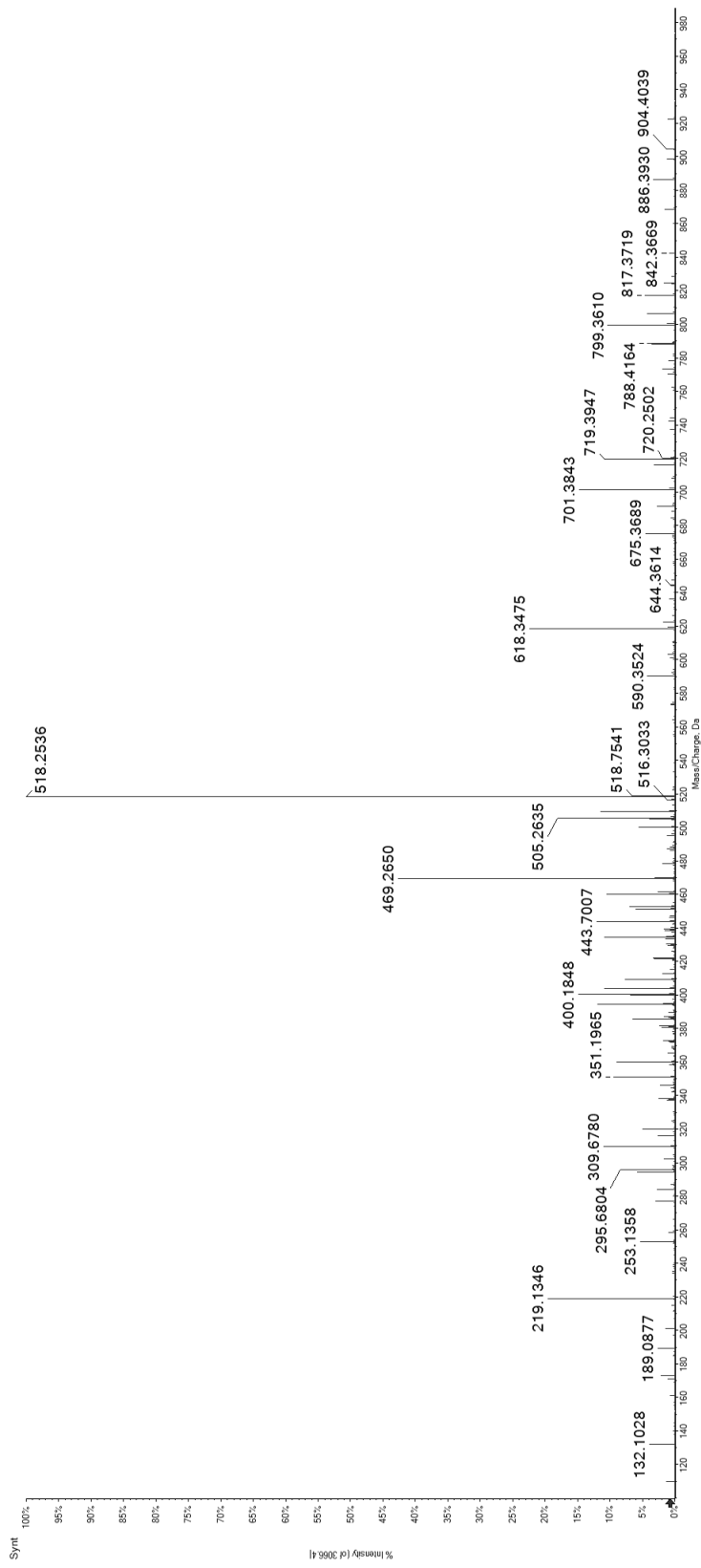
Synthetic Peptide



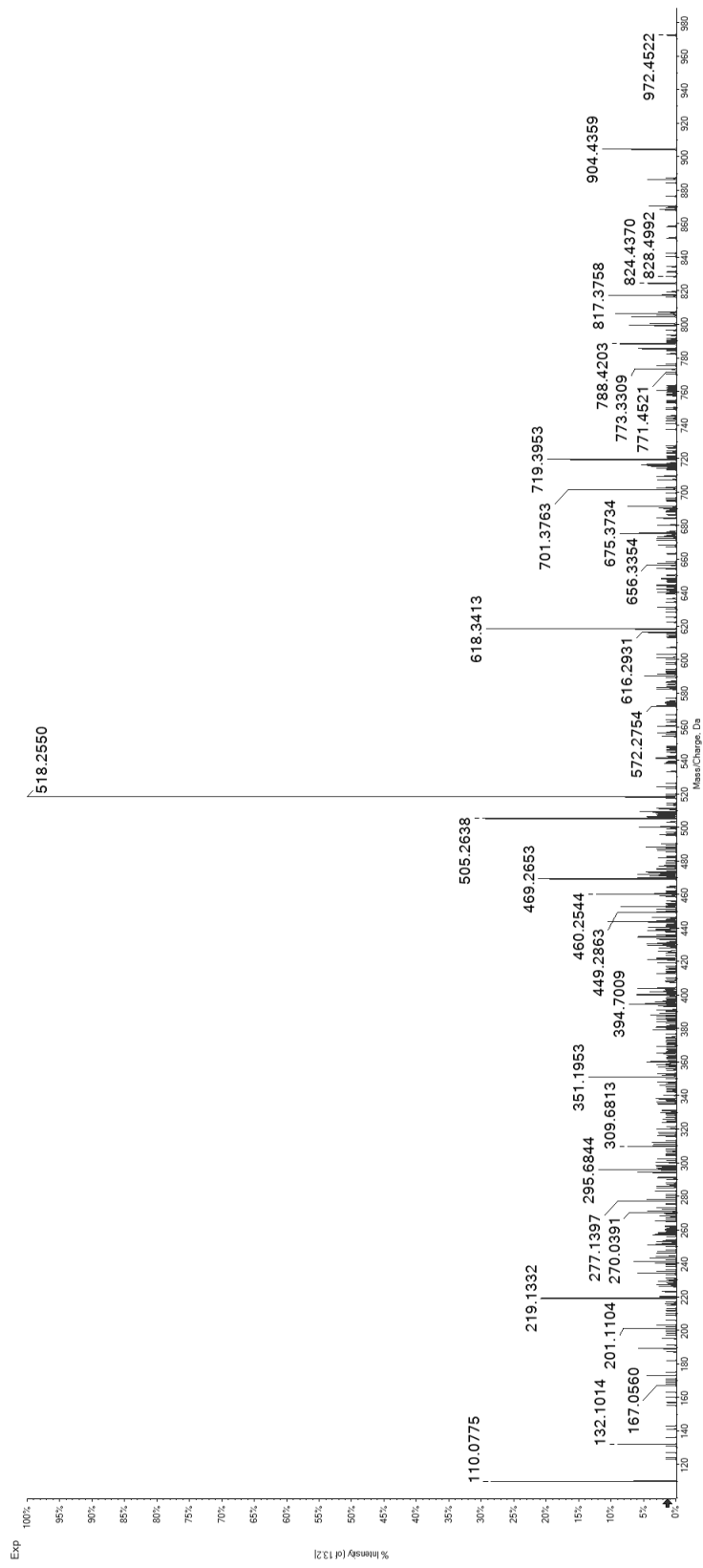
Candidate Spectrum

HRAG(pT)LTSL

Synthetic
Peptide



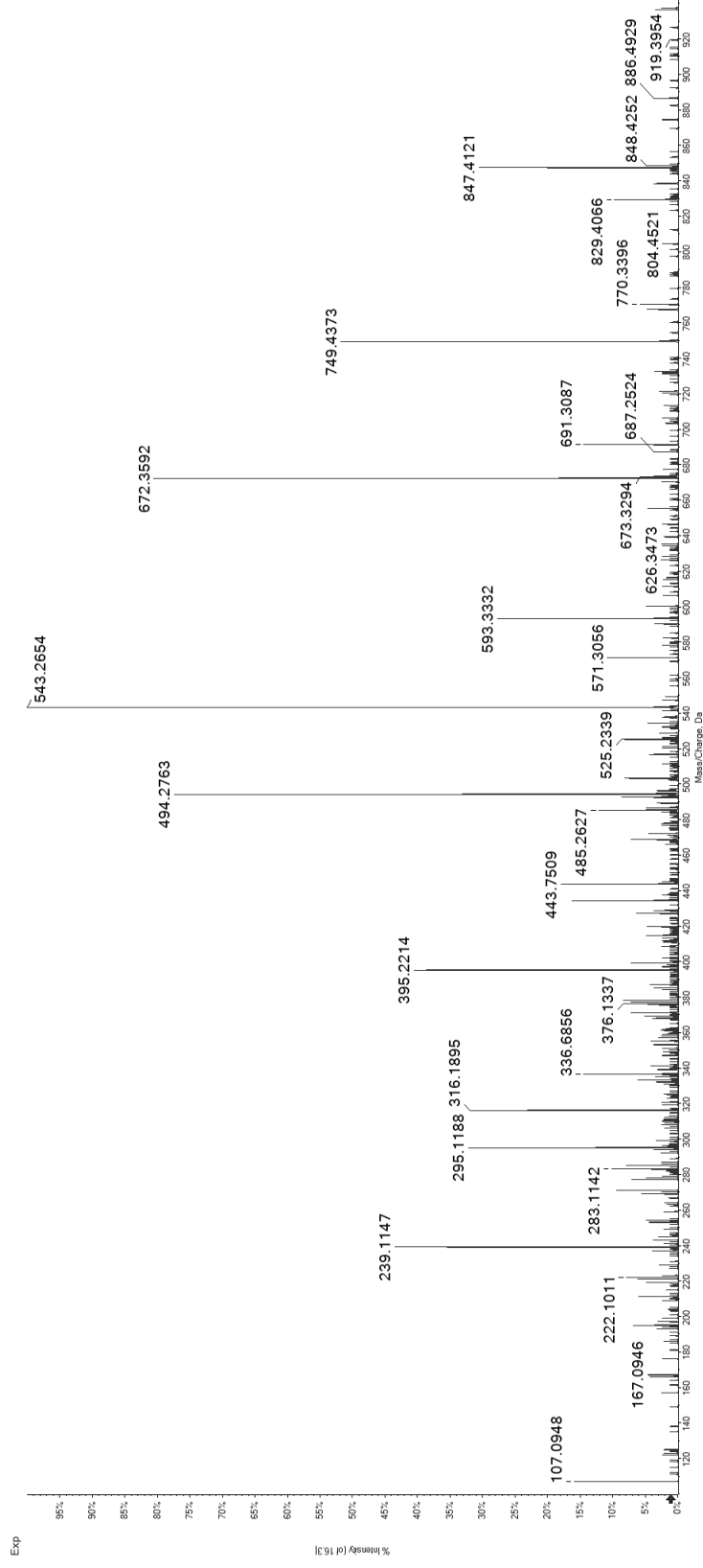
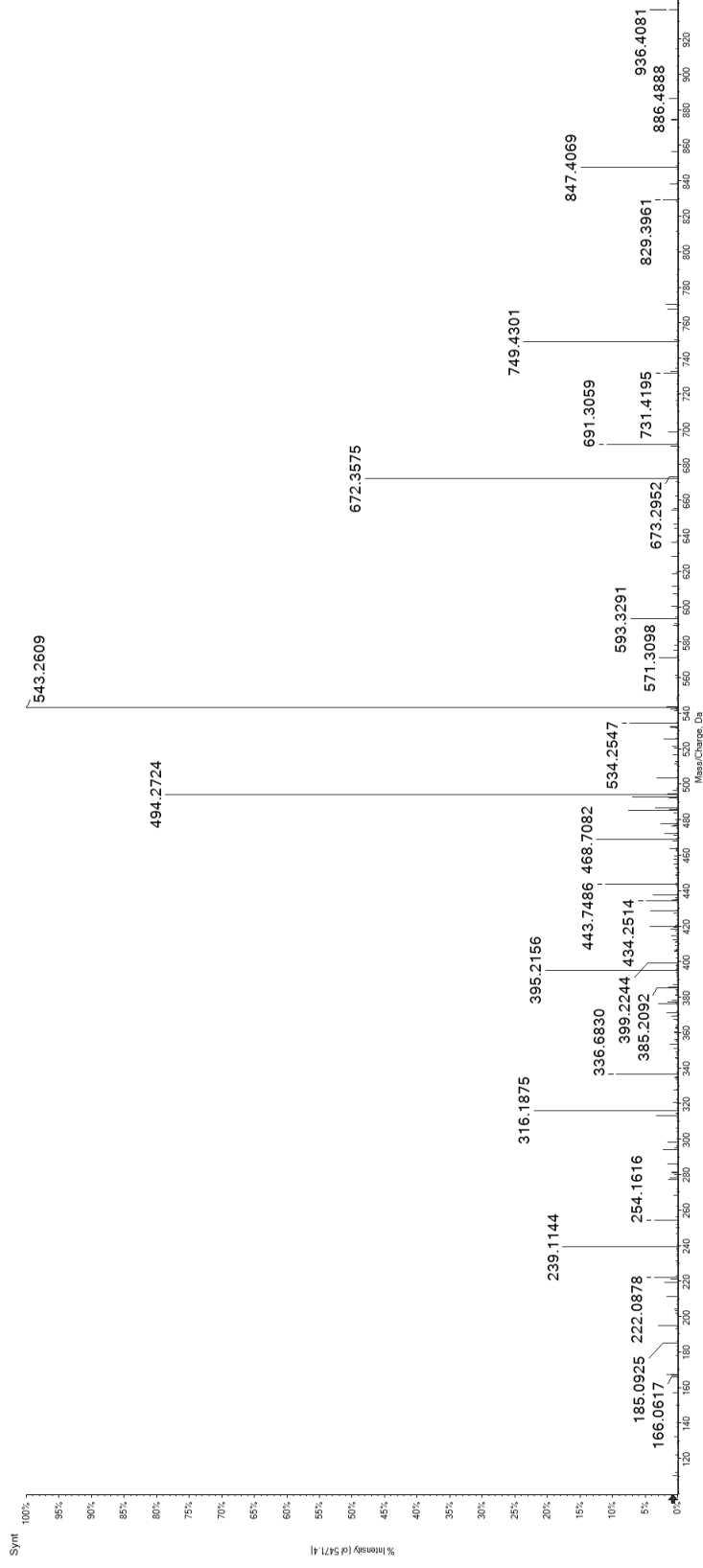
Candidate
Spectrum



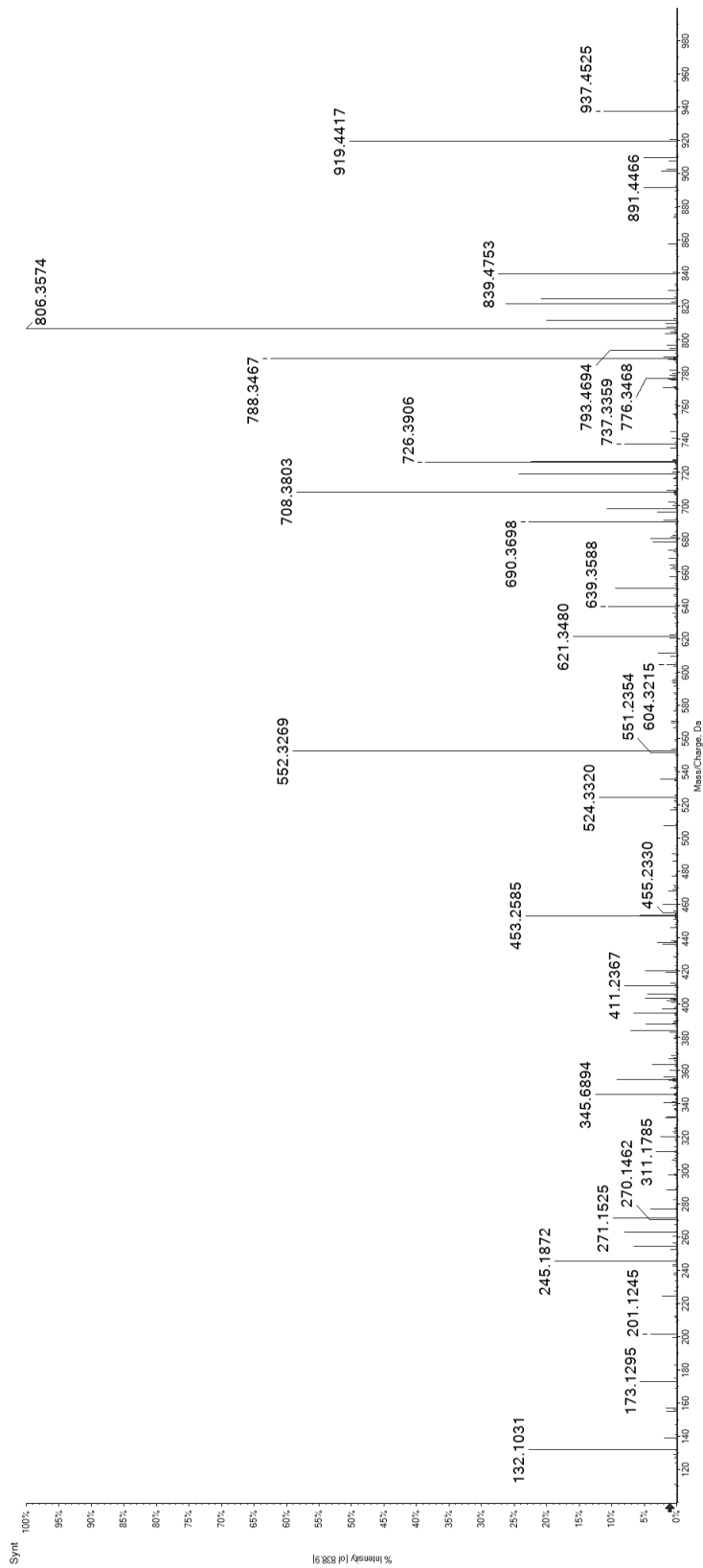
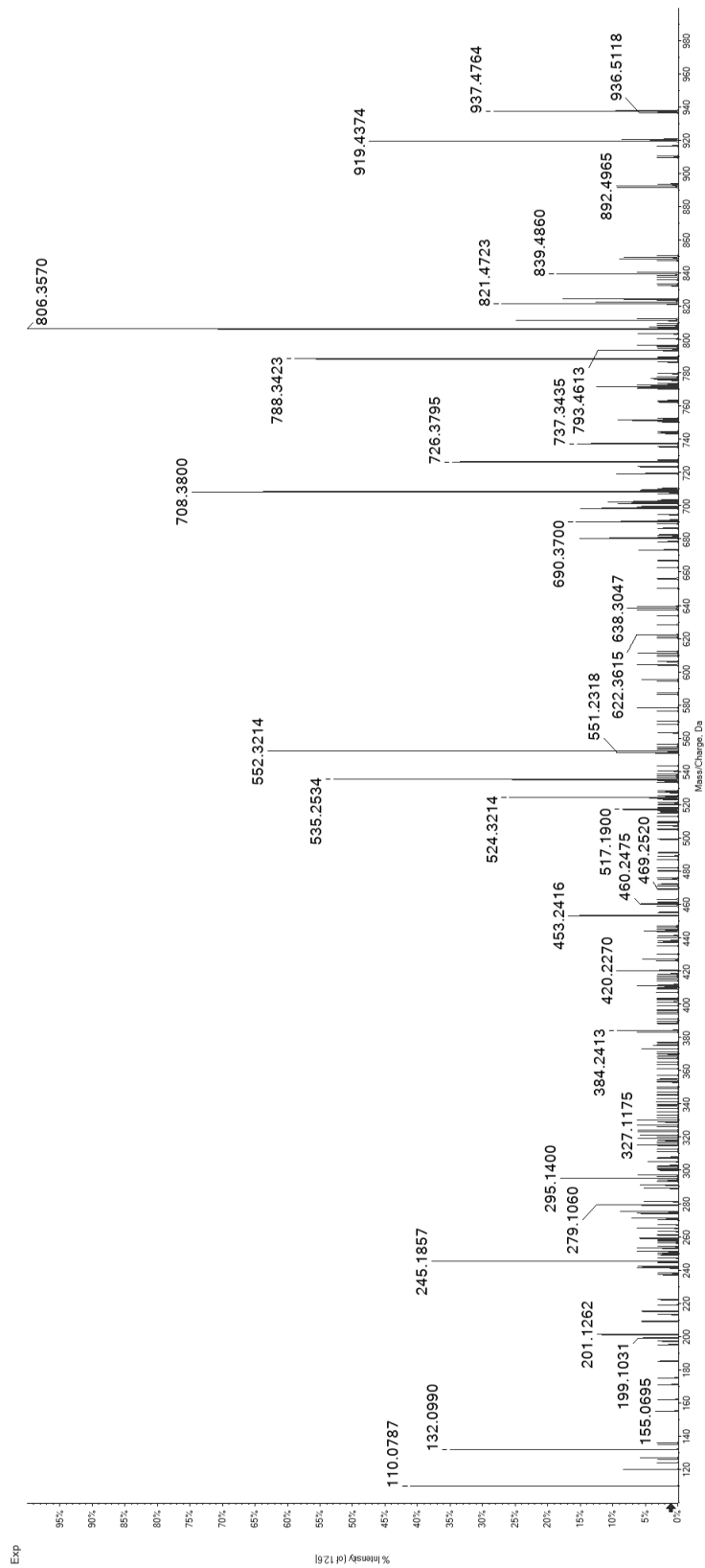
23

THRPP(pT)PSL

Synthetic Peptide



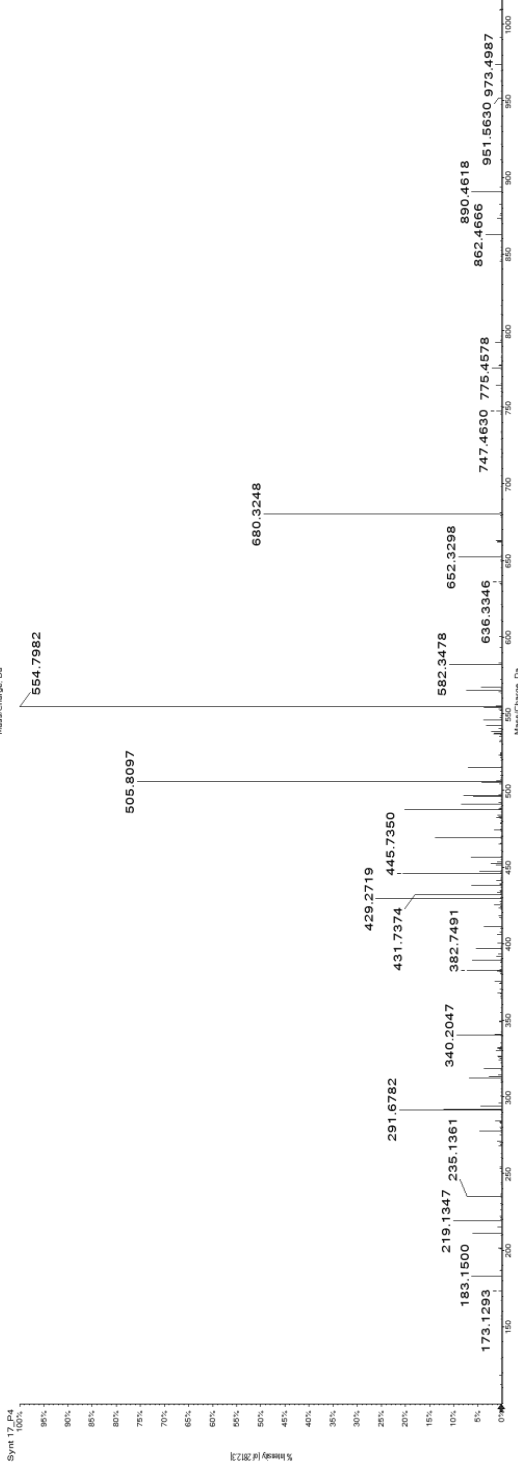
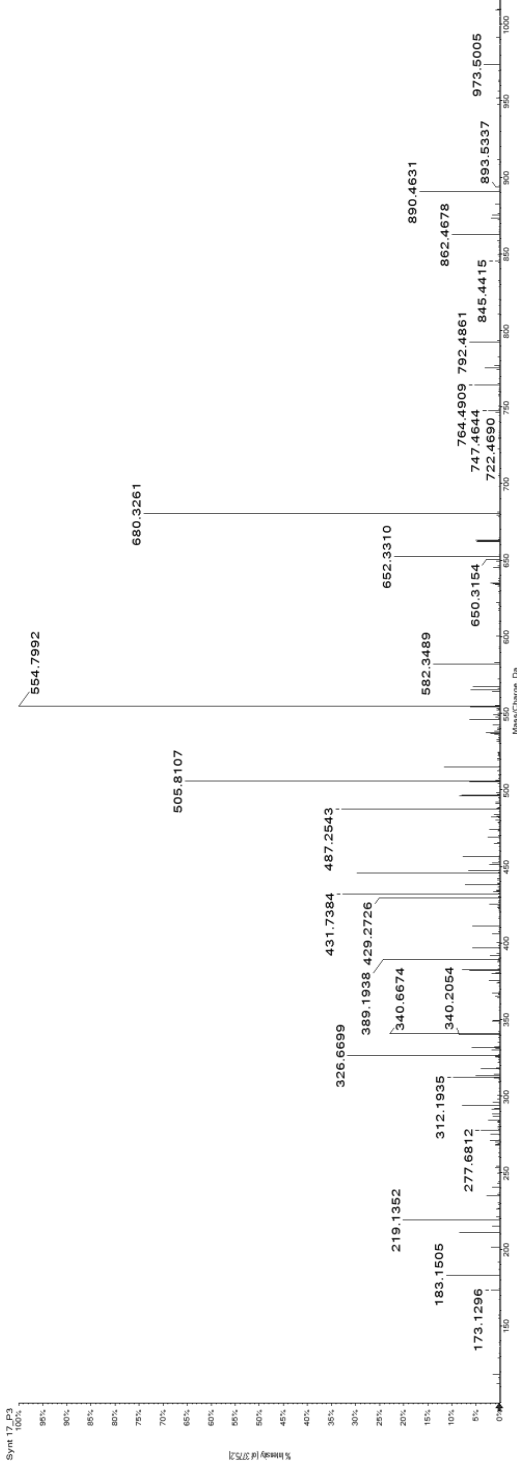
Candidate Spectrum

Synthetic
PeptideCandidate
Spectrum

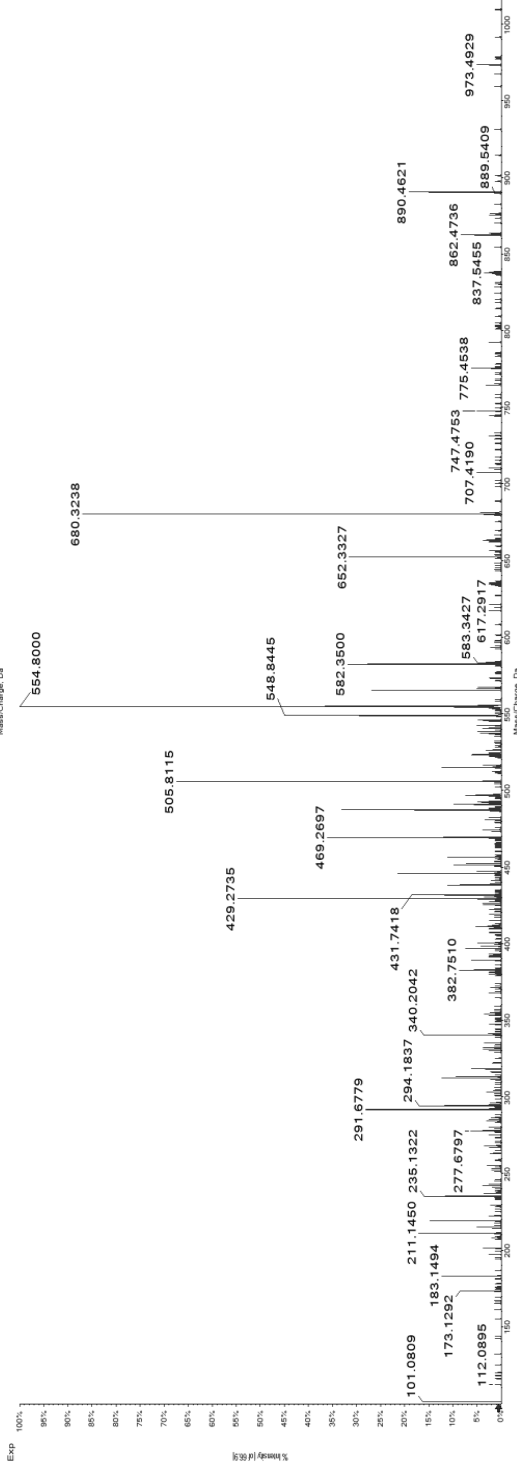
25

RR(pS)SIPITV
RRS(pS)IPITV

Synthetic
Peptide
RR(pS)SIPITV



Synthetic
Peptide
RRS(pS)IPITV

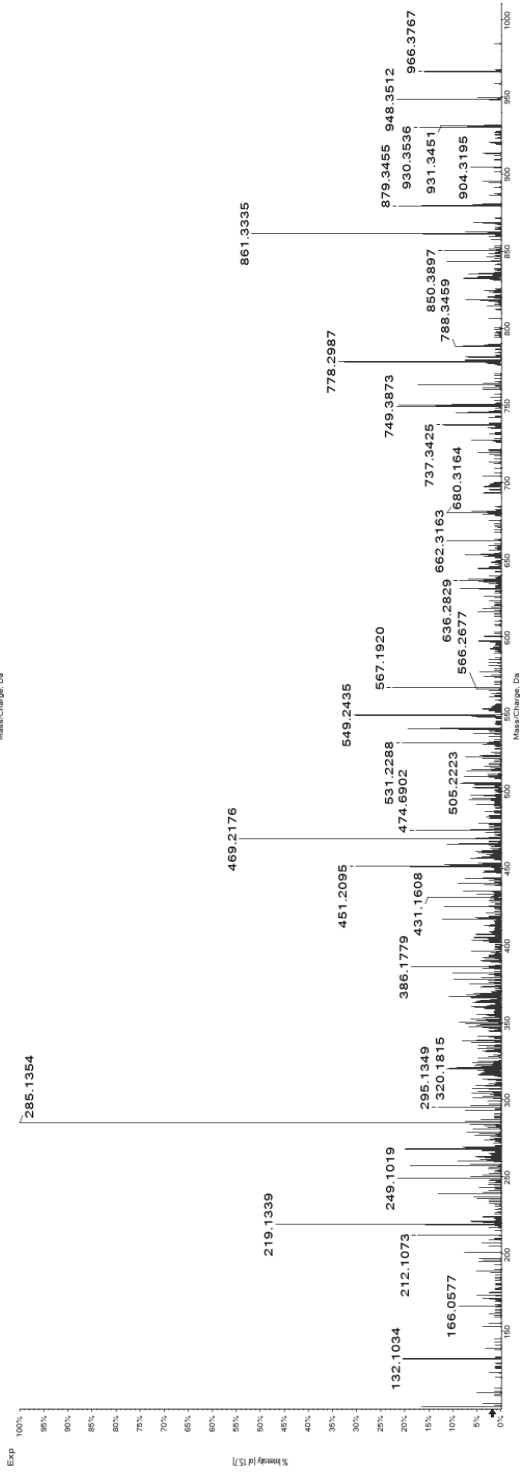
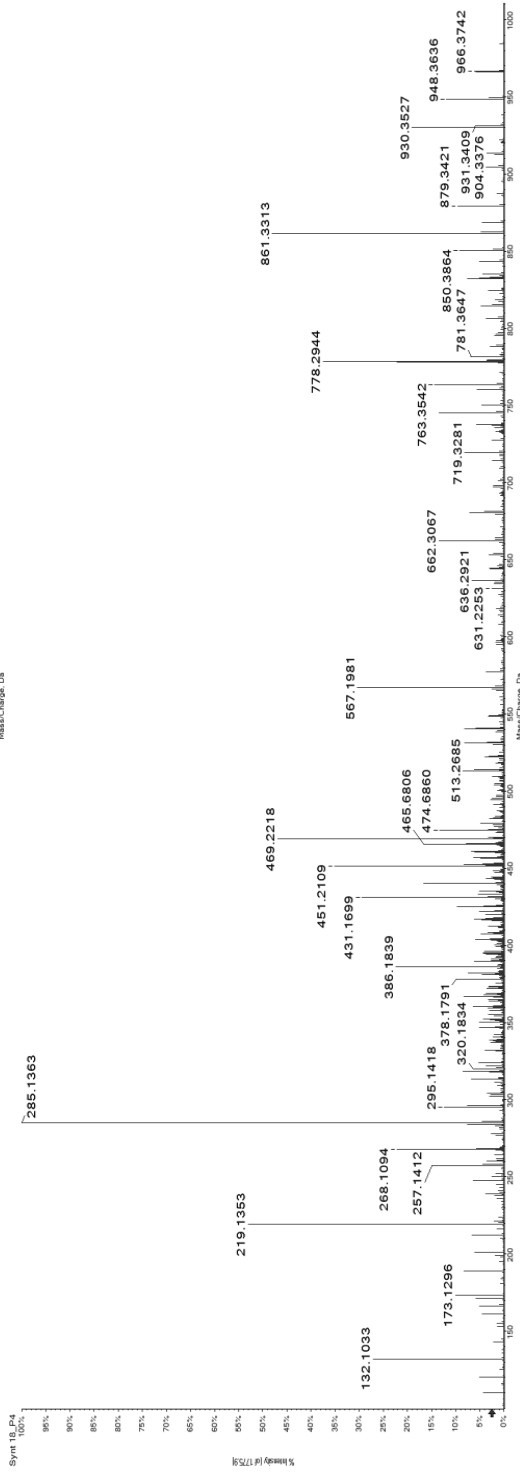
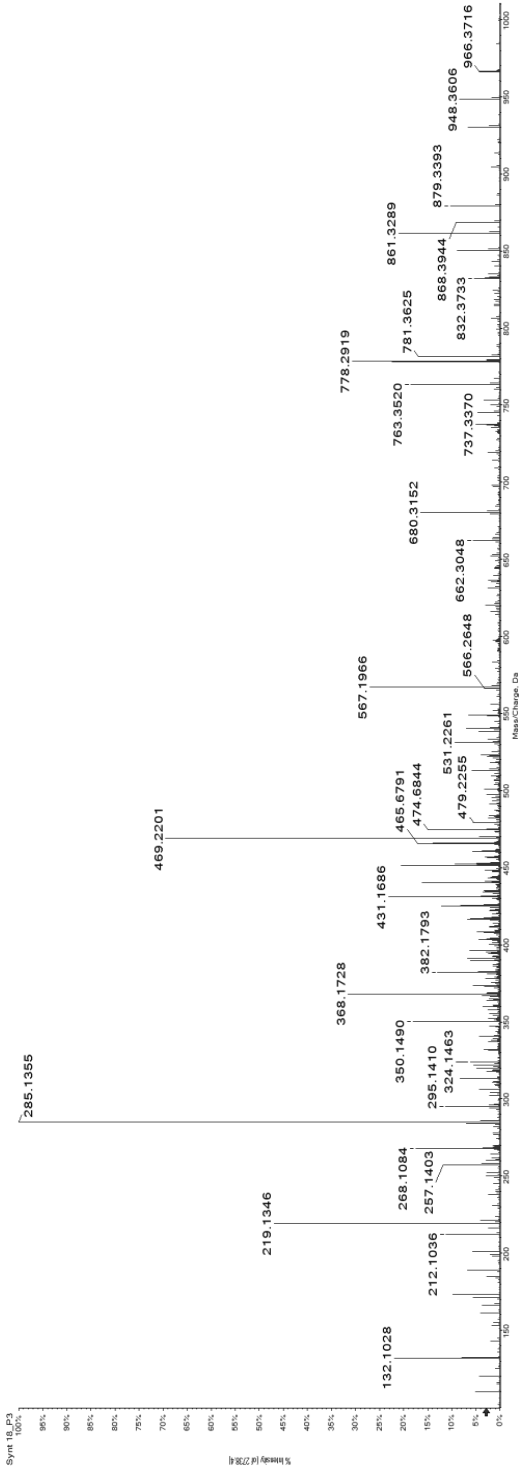


Candidate
Spectrum

26

FH(pT)TPNTSL
FHT(pT)PNTSL

Synthetic
Peptide
FH(pT)TPNTSL



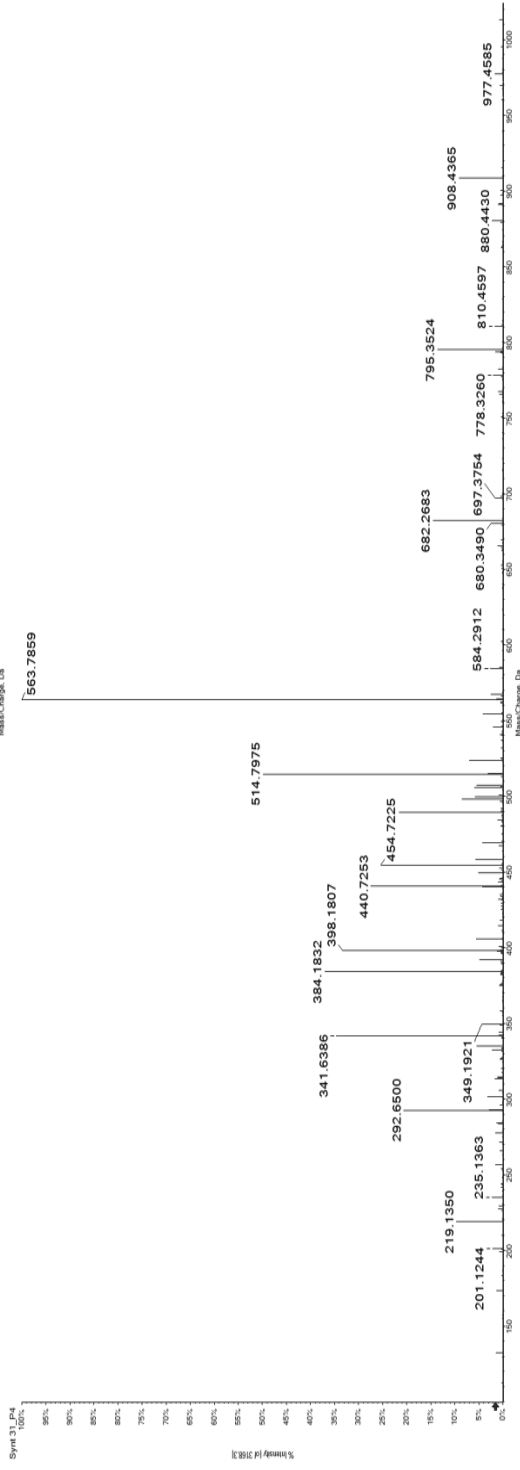
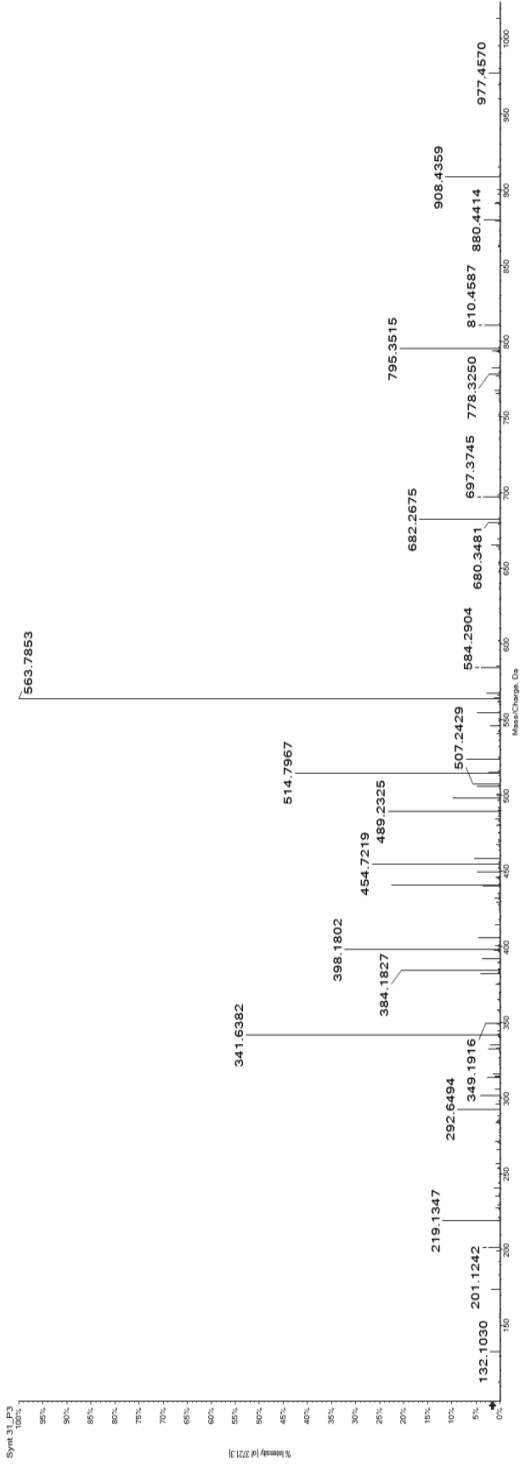
Candidate
Spectrum

229

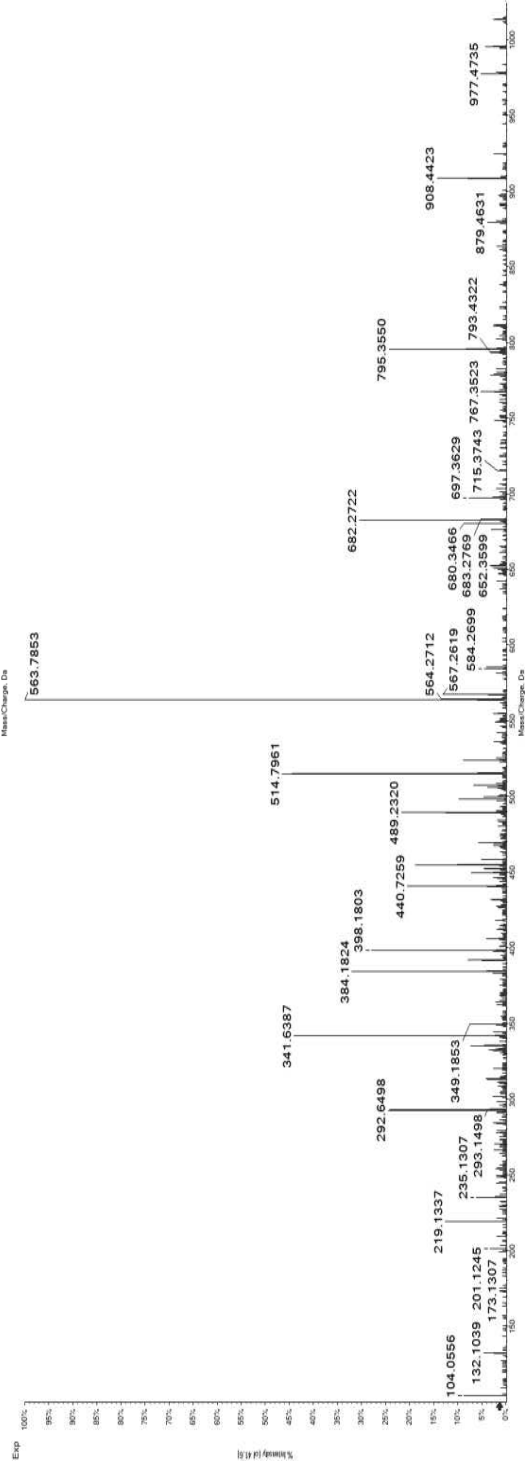
27

RR(pS)SDIISL
RRS(pS)DIISL

Synthetic
Peptide
RR(pS)SDIISL



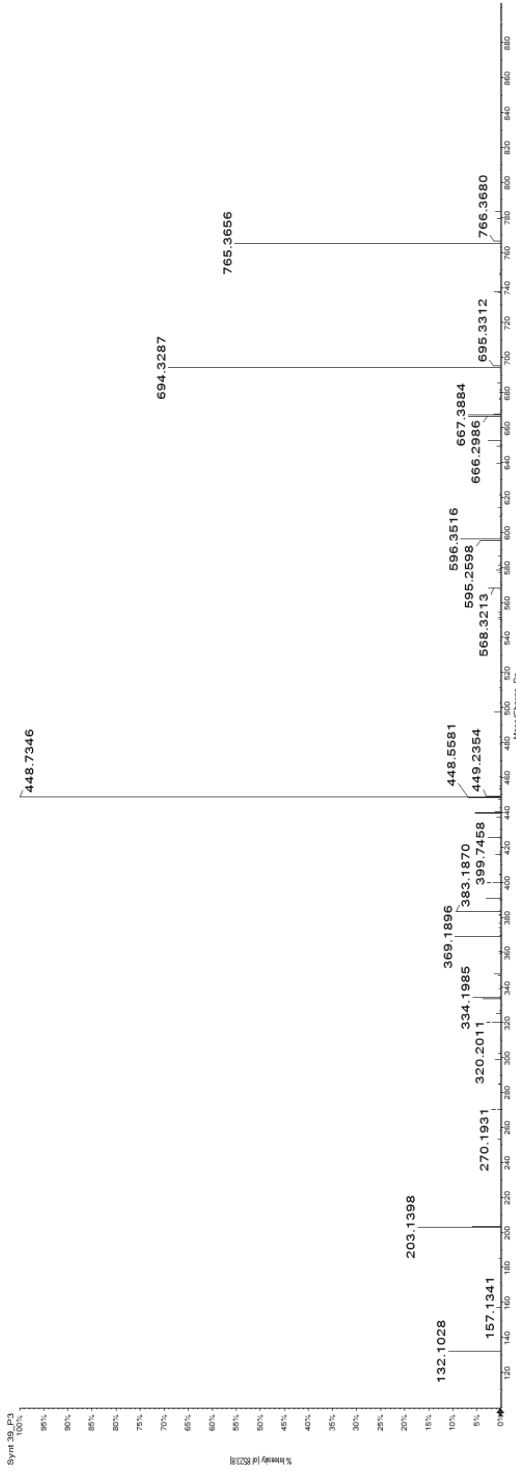
Synthetic
Peptide
RRS(pS)DIISL



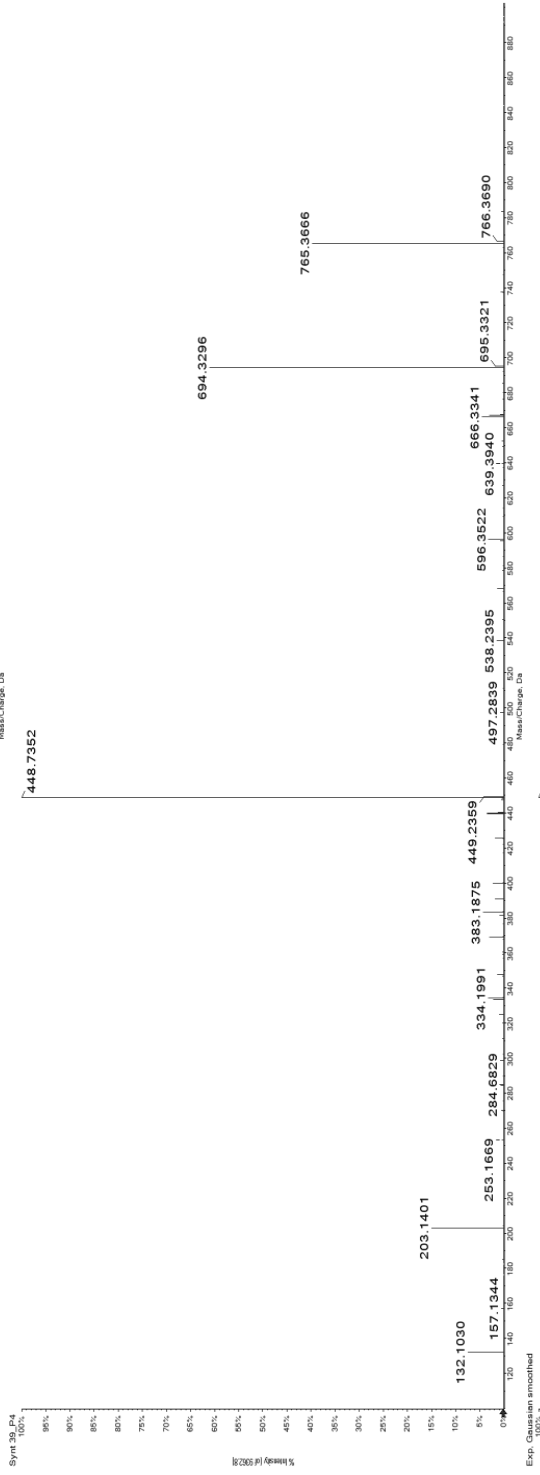
Candidate
Spectrum

LR(pS)TGVAL LRS(pT)GVAL

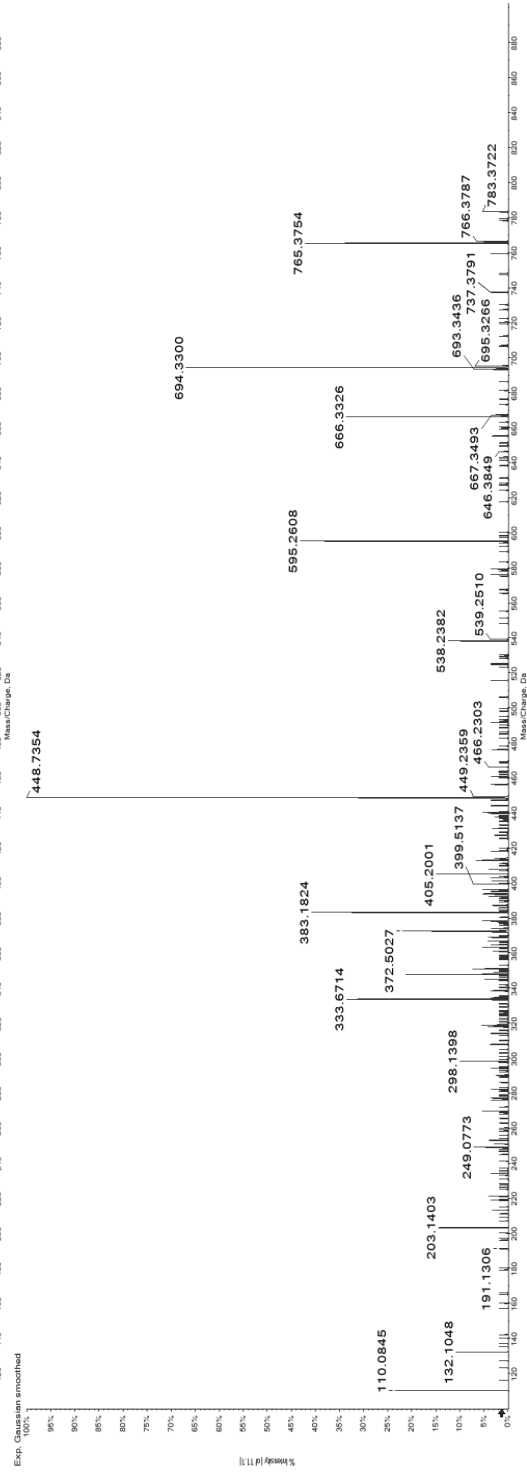
Synthetic Peptide LR(pS)TGVAL



Synthetic Peptide LRS(pT)GVAL



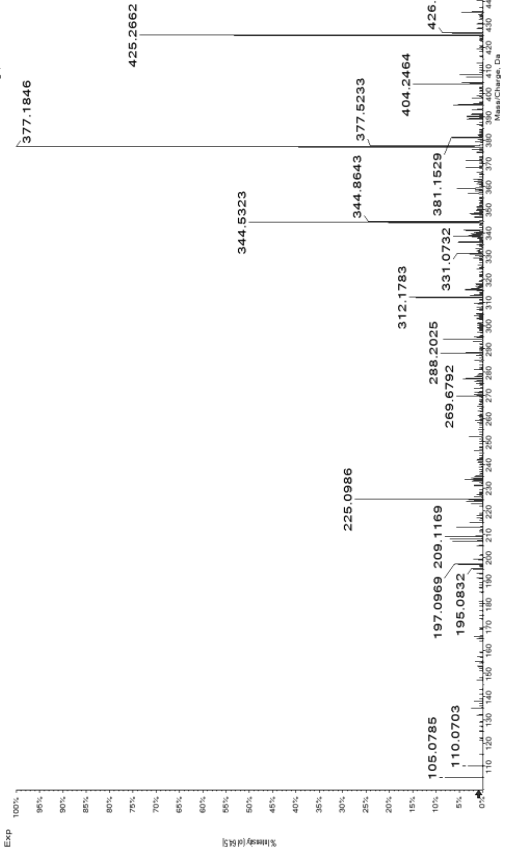
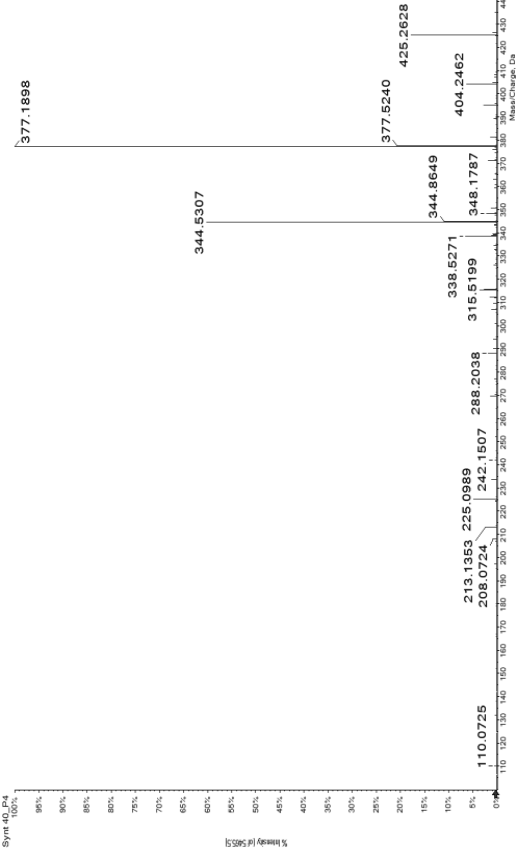
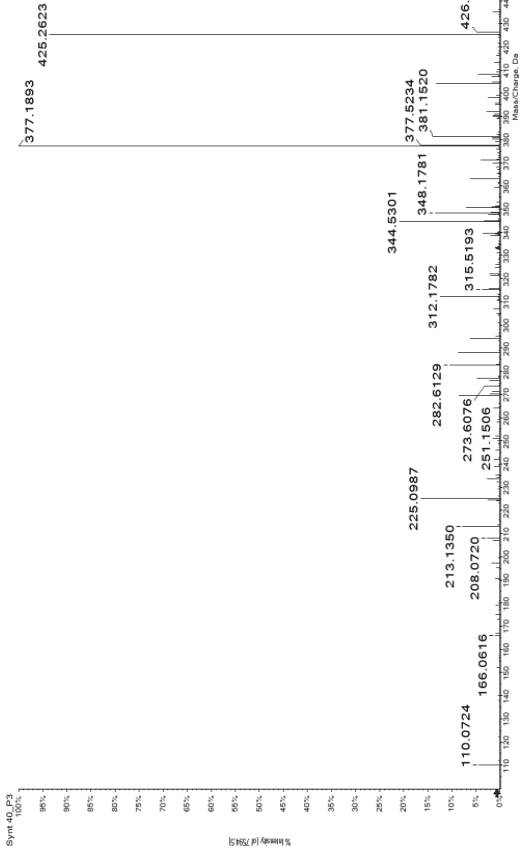
Candidate Spectrum



29

SH(pS)SIHRL
SHS(pS)IIHRL

Synthetic
Peptide
SH(pS)SIHRL



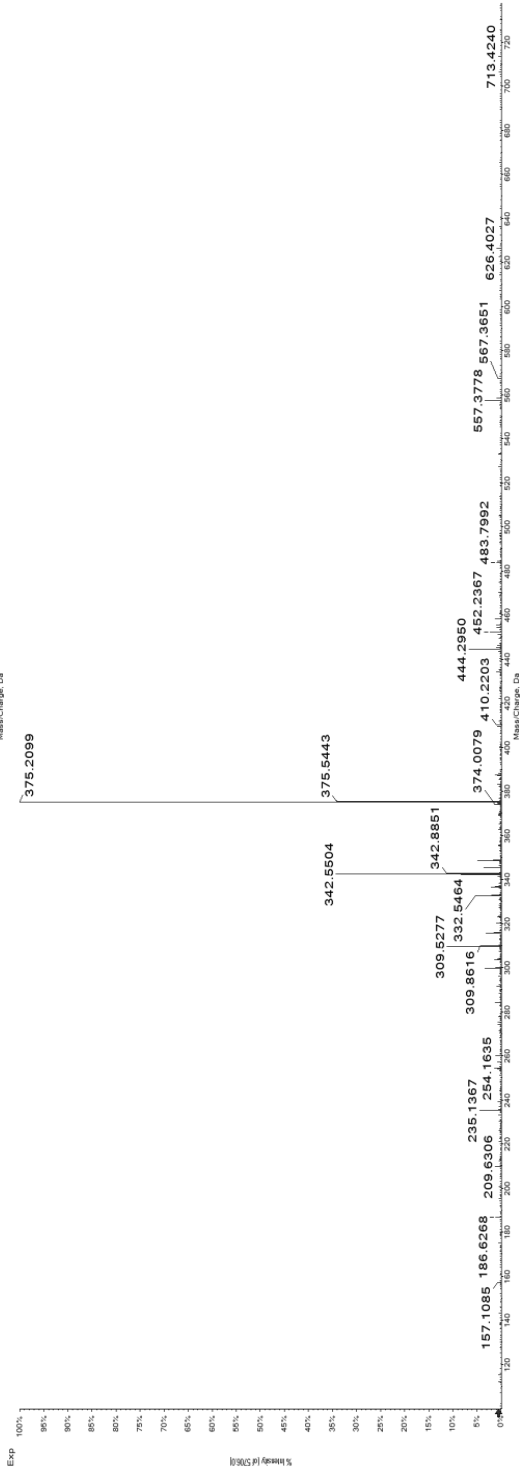
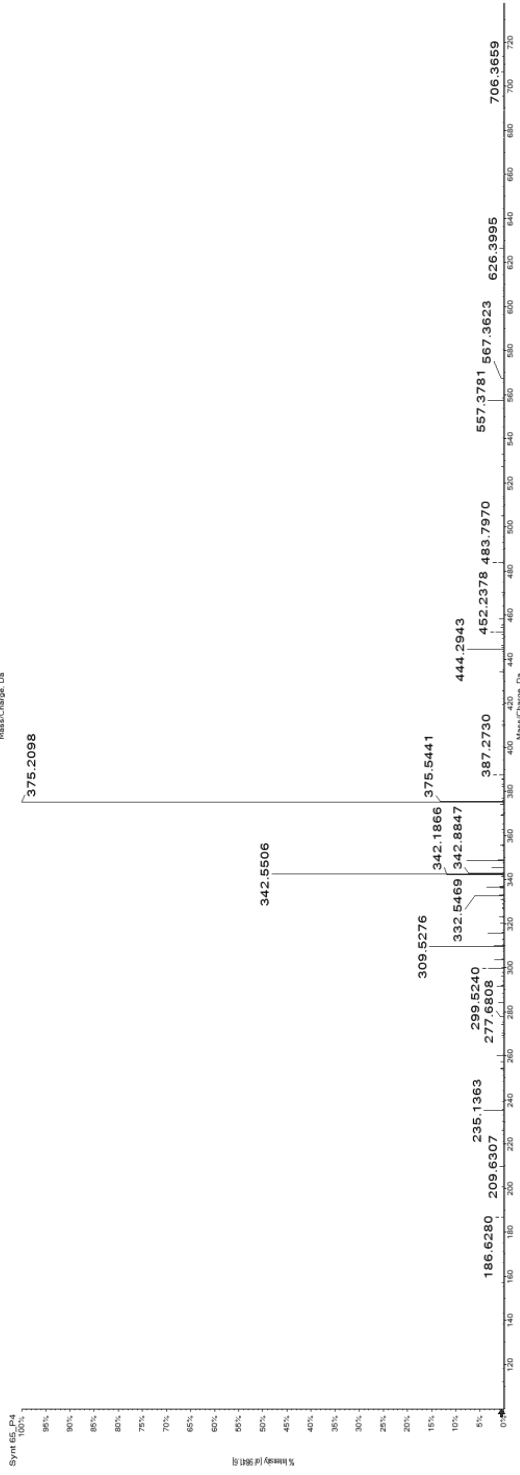
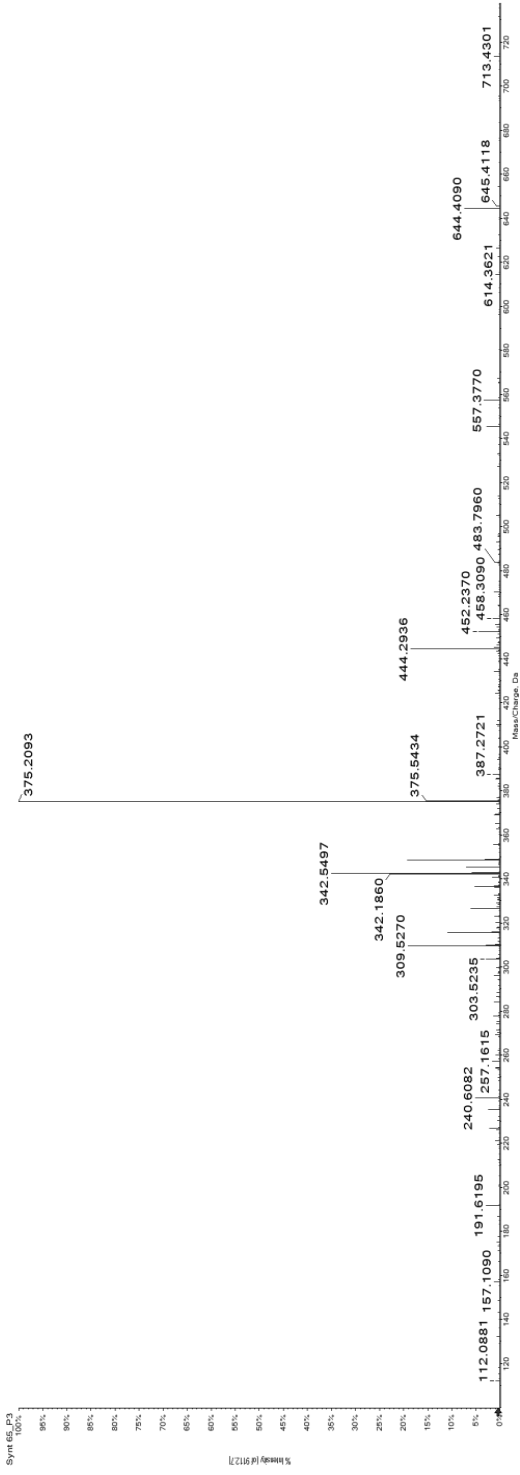
Synthetic
Peptide
SHS(pS)IIHRL

Candidate
Spectrum

30

RR(pS)SIGLRV
RR(pS)SIGLRV

Synthetic
Peptide
RR(pS)SIGLRV



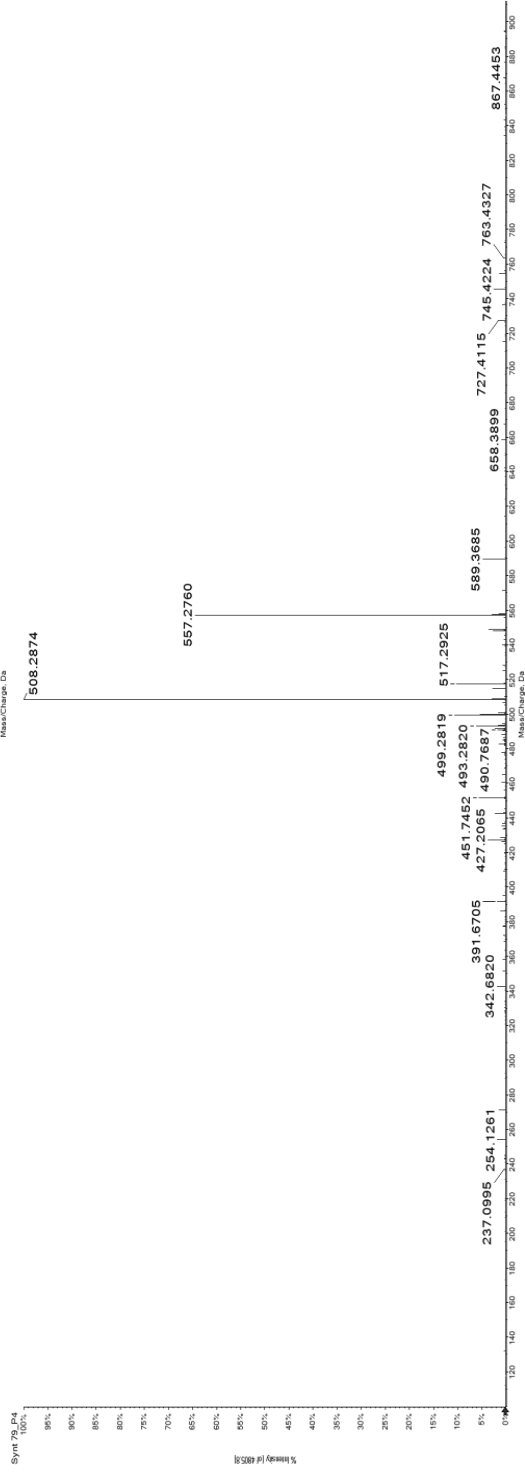
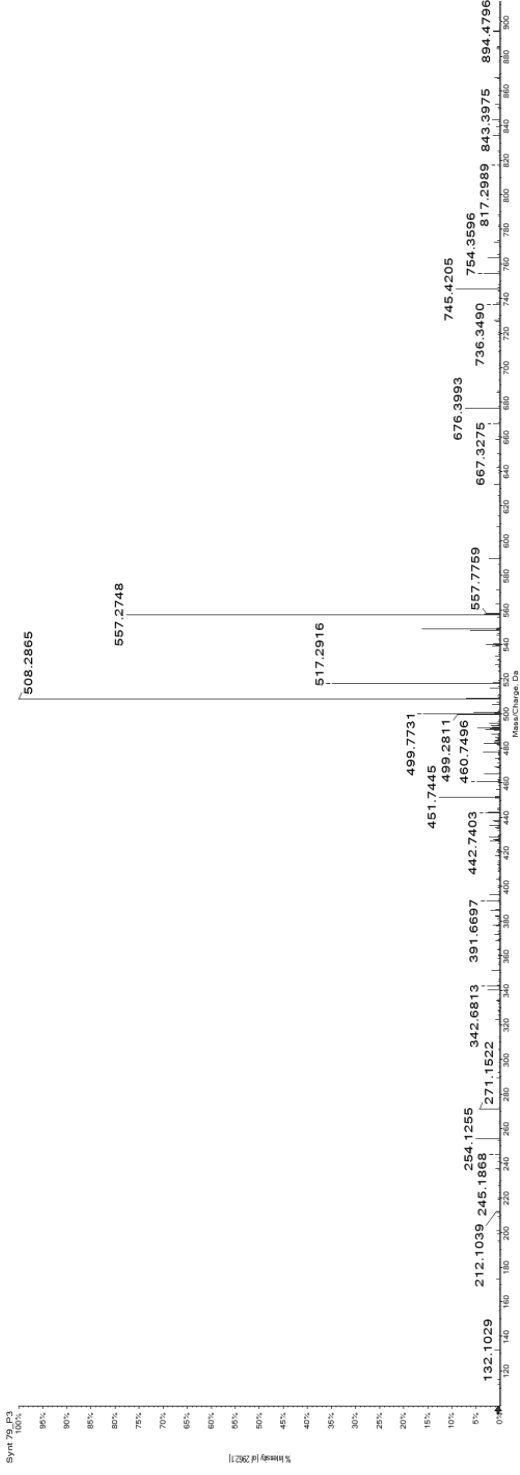
233

Candidate
Spectrum

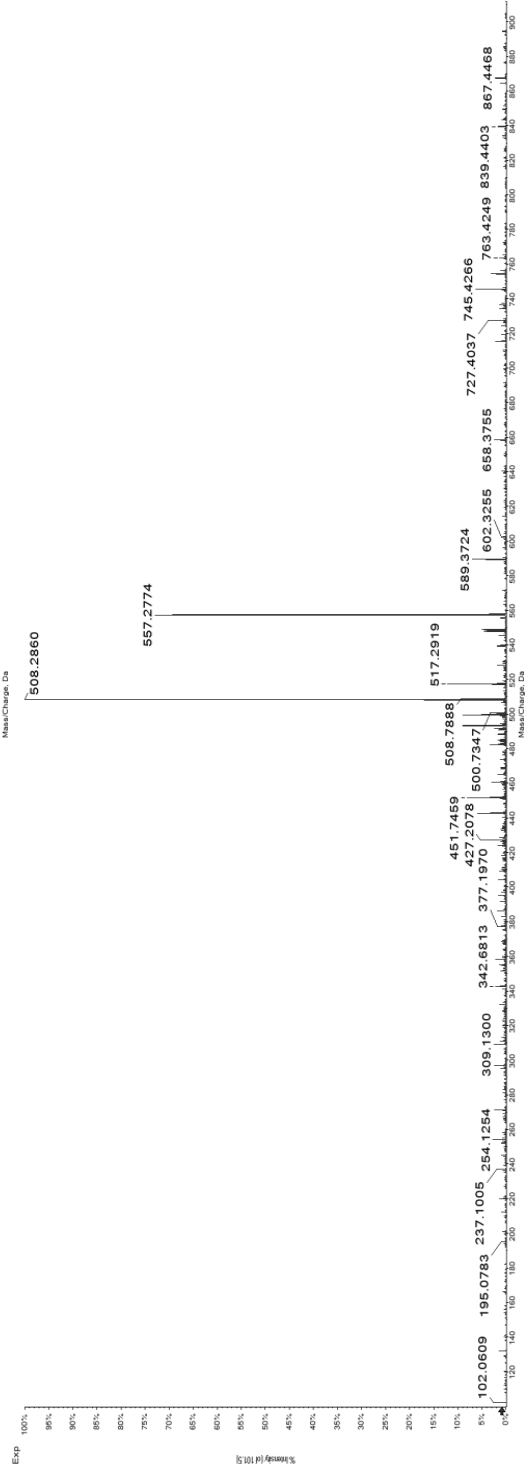
31

NR(pS)STRSLL
NRS(pS)TRSLL

Synthetic
Peptide
NR(pS)STRSLL



Synthetic
Peptide
NRS(pS)TRSLL



Candidate
Spectrum

Anexo 2

Publicaciones generadas durante el periodo de elaboración de esta tesis:

- 1- Marcilla M, **Alpizar A**, Paradela A, Albar JP. (2011) A systematic approach to assess amino acid conversions in SILAC experiments. *Talanta*. Apr 15; 84(2):430-6.
- 2- Marcilla M, **Alpizar A**, Lombardía M, Ramos-Fernandez A, Ramos M, Albar JP. (2014) Increased Diversity of the HLA-B40 Ligandome by the Presentation of Peptides Phosphorylated at Their Main Anchor Residue. *Mol. Cell Proteomics*. Feb; 13(2):462-74.
- 3- Caron E, Espona L, Kowalewski DJ, Schuster H, Ternette N, **Alpizar A**, Schittenhelm RB, Ramarathinam SH, Lindestam Arlehamn CS, Koh CC, Gillet LC, Rabsteyn A, Navarro P, Kim S, Lam H, Sturm T, Marcilla M, Sette A, Campbell D, Deutsch EW, Moritz RL, Purcell AW, Rammensee HG, Stevanovic S, Aebersold R. (2015) An open-source computational and data resource to analyze digital maps of immunopeptidomes. *Elife*. Jul 8; 4.
- 4- **Alpizar A**, Marino F, Ramos-Fernandez A, Lombardía M, Jeko A, Pazos F, Paradela A, Santiago C, Heck AJ, Marcilla M. (2016) A molecular basis for the presentation of phosphorylated peptides by HLA-B antigens. Enviado.