Detection of aequorin-2 protein in dietary supplements using mass spectrometry protein sequence analysis

Lianji Jin, Ph.D.

U.S. Food and Drug Administration, Forensic Chemistry Center, Cincinnati, OH

Email: lianji.jin@fda.hhs.gov; Phone: 513-679-2713

ABSTRACT

Apoaequorin is the protein component of aequorin isolated from jellyfish that, in tandem with the green fluorescent protein, is responsible for its bioluminescence. Apoaequorin is a single chain polypeptide with 189 amino acid residues and three EF-hand Ca^{2+} binding sites, and belongs to the Ca^{2+} binding protein superfamily. Dietary supplements bearing label claims such as "healthier brains" and "sharper minds" were received at the Forensic Chemistry Center for analysis. Apoaequorin was listed as the active ingredient, reportedly produced by expression of the synthetic gene in a strain of *Escherichia coli*. Here, sequence analysis of a protein isolated from dietary supplements is described. Matrix assisted laser desorption ionization tandem time-of-flight mass spectrometry (MALDI-TOF/TOF-MS) and liquid chromatography high resolution accurate mass mass spectrometry (LC-HRAM MS) were used to perform intact protein mass analysis as well as peptide mass mapping and peptide sequencing. Capsules of three different dose strengths, chewable tablets and protein powders were found to contain a related compound ("Aequorin-2"; UniProtKB/Swiss-Prot accession# P02592). Aequorin-2 is comprised of 196 amino acid residues, which is equivalent to apoaequorin with a propeptide of 7 amino acid residues at the N-terminus.

Key words: apoaequorin, dietary supplement, MALDI-TOF/TOF-MS, LC-HRAM MS

INTRODUCTION

Aequorin is a photoprotein responsible for the bioluminescence of jellyfish *Aquorea Victoria* [1]. It was first isolated from the photogenic organs of jellyfish *Aquorea aquorea* by Shimomura et al. in the 1960s [2]. Aequorin is composed of two distinct units, the apoaequorin protein and the prosthetic group coelenterazine. When the protein binds calcium, aequorin decomposes via the action of molecular oxygen into apoaequorin, coelenteramide and CO_2, accompanied by the emission of blue light (the complex is often called the blue fluorescent protein, BFP) [3]. Luminescent jellyfish *Aequorea Victoria*, however, emits green light not blue light, which is due to the presence of the green fluorescent protein (GFP). In the presence of GFP, resonance energy transfer takes place from the excited state BFP to a chromophore in GFP to produce excited-state GFP, which emits green light as it relaxes to its ground state, identical to that emitted by jellyfish [4]. The apoprotein, apoaequorin (without the prosthetic group coelenteramide), can be produced and purified through expression of the apoaequorin cDNA in *Escherichia coli* [5]. This recombinant apoprotein has been used as a probe for measuring Ca^{2+} levels in intracellular organelles [6, 7]. Upon regeneration with coelenterazine, the recombinant aequorin was fully active with Ca^{2+}.

Apoaequorin is a single chain polypeptide comprised of 189 amino acid residues with a molecular mass of 21,459 Da, [1, 8], and contains a disulfide bond between Cys-145 and Cys-152 [9]. It contains three EF-hand Ca^{2+} binding sites, and is a member of the Ca^{2+} binding protein superfamily. These calcium binding properties were investigated recently for their effect in protecting neurons in an *in vitro* model of global ischemia in acute hippocampal brain slices [10]. The study showed that the hippocampal slices of rat brain infused with apoaequorin protected neurons against ischemic cell death, suggesting apoaequorin may be neurotherapeutic.

Dietary supplements, declared to contain apoaequorin and bearing label claims such as “healthier brains” and “sharper minds”, were received at the Forensic Chemistry Center for analysis. A related compound, known as aequorin-2, is produced through expression of the synthetic gene in a strain (BL21/DE3) of *Escherichia coli* [11]. The sequence is composed of 196 amino acid residues with a molecular mass of 22,285 Da (UniProtKB/Swiss-Prot Accession# P02592) (Figure 1). Compared to the apoaequorin sequence of 189 amino acid residues, the 196-residue sequence contains a propeptide of 7 amino acid residues at the N-terminus. Here, sequence analyses of a primary protein found in each of the dietary supplements tested are presented using matrix assisted laser desorption ionization tandem time-of-flight mass spectrometry (MALDI-TOF/TOF-MS) and liquid chromatography high resolution accurate mass mass spectrometry (LC-HRAM MS). The results indicate that aequorin-2, not apoaequorin as declared, was present in the capsules, chewable tablets and protein powders.

EXPERIMENTAL

Materials

Sequencing grade modified trypsin was purchased from Promega. Dithiothreitol (DTT), iodoacetamide (IAM), ammonium bicarbonate (NH_4HCO_3) and trifluoroacetic acid (TFA) were obtained from Sigma-Aldrich. RapiGest™ SF was bought from Waters. Acetonitrile (ACN), formic acid (FA), sinapic acid and Pierce LTQ Velos Calibration mix were obtained from Thermo Fisher. The mass standards kit for calibration of TOF/TOF instrument and alpha-cyano-

4-hydroxycinnamic acid (CHCA) were obtained from AB SCIEX. Deionized water (18.2 MΩ DI H_2O) was produced using an in-house System. All materials and reagents were used as purchased without further purification.

Dietary supplement description

Four different dietary supplements labeled to contain apoaequorin were analyzed using MALDI: 40 mg capsules (two different lots; "SC1"), 20 mg capsules (two different lots; "SC2"), 10 mg capsules (single lot) and 10 mg chewable tablets (single lot).

LC-HRAM MS analysis was conducted on SC1 and a supplement in powder form (single lot; "SPP"), labeled to contain 5 mg apoaequorin in 5 g of protein powder per packet.

Capsule and tablet preparation

MALDI analysis: Four individual capsules each from SC1 and SC2 were selected and the contents of each capsule were emptied into 15 mL conical tubes. Ten mL of DI H_2O were added to each tube, which was vortexed for 5 min and centrifuged for 2 min at 165×*g*. One mL of each supernatant was transferred into a 2 mL Eppendorf tube and centrifuged at 21000×*g* for 5 min. The new supernatant was collected and labeled as the sample stock solution. The sample stock solution was diluted 10-fold with DI H_2O for MALDI intact mass analysis and trypsin digestion. A 20-capsule composite for the 10 mg capsules, and a 20-tablet composite for the 10 mg chewables, were prepared similarly (data not shown).

LC-HRAM MS analysis: A five-capsule composite was prepared for SC1. Sample preparation was the same as MALDI analysis and was analyzed in duplicate.

Protein powder preparation

A composite was made of five protein powder packets. Duplicate portions of the composite, equivalent to ~ one-tenth of the packet content, were weighed out and reconstituted with 10 mL of DI H_2O in 15 mL conical tubes. The conical tubes were vortexed for 2 min and centrifuged for 20 min at 165×*g*. One mL of each supernatant was transferred into a 2 mL eppendorf tube and centrifuged for 10 min at 21000×*g*. The new supernatant was collected and labeled as the sample stock solution. The sample stock solution was aliquoted and stored at -20ºC until further analysis. The sample stock solution was diluted 50-fold with DI H_2O for trypsin digestion.

Trypsin digestion procedure

Individual capsules, composites and protein powders were prepared in the same manner except the digestion time and the amount of trypsin added depending on the mass spectrometry technique used and the amount of protein estimated. Briefly, 50 µL of the capsule/tablet sample (10X) or protein powder sample (50X) was mixed with 5 µL of 1% RapiGest™ SF in 400 or 500 mM NH_4HCO_3, followed by addition of 3 µL of 100 mM DTT. Following a brief vortex and spin down, the sample was incubated in a metal beads bath for 30 min at 60°C. After cooling to room temperature, 10 µL of 100 mM IAM was added to the sample which was left in the dark for 30 min. Sequencing grade trypsin (20 ng/µL), 34 µL for 10X or 50 µL for 50X, was then added to the sample, followed by incubation in a water bath at 37°C overnight for LC-HRAM MS sample or 2h for MALDI-TOF/TOF-MS sample. TFA was added to a final concentration of

~0.5%, followed by incubation for 30 min at 37°C. The sample was centrifuged for 10 min at 21000×*g*. The supernatant was collected for analysis.

MALDI-TOF/TOF MS Parameters

AB SCIEX™ TOF/TOF 5800 System was controlled via TOF/TOF™ Series Explorer™ software. Mass calibration was performed using AB SCIEX TOF/TOF mass calibration standard mix following manufacturer's recommendations. The intact mass measurement was recorded using Linear Mid High Mass method. The MS settings were as follows: mass range, *m/z* 4000-30,000; total shots/spectrum, 2000; stage velocity, 600 µm/sec; laser pulse rate, 400 Hz; laser intensity, 5500. The MS (full scan) spectra of the tryptic digests were acquired using the MS Reflectron Positive method. The MS settings were as follows: mass range, *m/z* 700-4000; total shots/spectrum, 400; stage velocity, 600 µm/sec; laser pulse rate, 400 Hz; laser intensity, 2800. The MSMS (product ion scan) spectra were acquired using the MSMS 2kV method. The MSMS settings were as follows: CID control, CID off; acquisition control, automatic; precursor mass window, relative, 200 resolution [FWHM]; metastable suppressor, on; total shots/spectrum, 4000; stage velocity, 1200 µm/sec; laser pulse rate, 1000 Hz; laser intensity, 3360. MALDI sample spotting was performed as follows: 1 µL of tryptic digest was mixed with 1 µL of CHCA matrix solution in 50:50:0.1% CH_3CN:H_2O:TFA (10 mg/mL), and 1 µL was spotted in each well on a OptiTOF 384 MALDI plate (SCIEX) which was allowed to air dry at ambient temperature. Sample spotting for protein samples was prepared in the same manner except that sinapic acid matrix solution was used.

LC-HRAM MS Parameters

An RSLCnano Ultimate 3000 UHPLC (Dionex) was interfaced with the Orbitrap Elite ETD mass spectrometer (Thermo) via a heated electrospray ionization (HESI) source. The MS system was controlled via XCalibur 2.1 software. Chromeleon Xpress (Dionex) was also used independently to control LC operation. An Acquity UPLC® Protein BEH C18 column, 1.0 mm X 150 mm, 1.7 µm, 300Å, (Waters, MA), was used for separations of tryptic peptides. Mobile phase A was 0.1% FA in H_2O and mobile phase B was 0.1% FA in CH_3CN. Tryptic peptides (2 µL injected) were eluted at a flow rate of 40 µL/mL with the following gradient: 2% B hold for 1 min, 2-40% B in 25 min, 40-90% B in 5 min and hold for 2 min, 90-2% B in 0.01 min and hold for 4.99 min. Total run time was 38 min with divert valve switched to waste during the first and the last minute. The Orbitrap was calibrated using Pierce LTQ Velos calibration solution per manufacturer's instructions. The MS spectra were recorded in FT mode with a scan resolution of 240,000 and a scan range of 300-2000 *m/z*. Automatic gain control (AGC) was $1e^6$. The MSMS spectra were recorded in data dependent mode using the ion trap in parallel. The ten most abundant ions were selected for generating product ion scans using the following parameters: Charge State: enable charge state screening, enable monoisotopic precursor selection, reject charge state +1; Activation Type: CID; Min Signal Required, 1000.0; Isolation Width: 2.00; Normalized Collision Energy, 35%; Default Charge State, 2; Activation Q, 0.250; Activation Time, 10.000. AGC target was $1e^4$ for CID. Dynamic exclusion was enabled as the following: Repeat Count, 2; Repeat Duration, 15.00; Exclusion List Size, 50; Exclusion Duration, 60.00; Exclusion Mass Width Low/High, 0.50000/1.50000.

Data processing and analysis

MALDI MS and MSMS spectra were analyzed and processed using Data Explorer® Software Version 4.10. Protein Prospector was used to perform *in silico* digestion and fragmentation to calculate the theoretical mass values. LC-HRAM MS data were searched against UniProtKB/ Swiss-Prot database via SEQUEST in Proteome Discoverer 1.2. The search parameters were as follows: mass range, 350-5000; protease, trypsin; missed cleavages, 2; static modification on cysteine residue, 57 Da; dynamic modification on methionine residue, 16 Da; precursor mass tolerance, ± 10 ppm; fragment ions mass tolerance, ± 0.8 Da. Search against Decoy database, FDR, 0.01.

RESULTS AND DISCUSSION

MALDI-TOF/TOF-MS of capsules and tablets

Intact mass analysis by MALDI. Capsules and tablets were extracted and analyzed to determine if any protein was present in the samples. Intact mass analysis of SC1 yielded a single mass peak at *m/z* ~22.7 kDa, as shown in Figure 2; similar results were obtained for SC2 and the capsule and tablet composites. The observed molecular mass is in line with the ~22.3 kDa of aequorin-2 with 196 amino acid residues. Note the mass measurement was performed under default calibration mode. Although mass accuracy could be improved by adding internal mass calibration standards to the sample to bracket the mass of interest and calibrate the mass range, highly pure protein standards of suitable molecular sizes can be difficult to obtain. Nonetheless, the intact mass analysis provided an estimated molecular weight, which is important in distinguishing the apoaequorin from aequorin-2. This is important, because once the protein undergoes enzymatic digestion for subsequent sequence analysis, the information on the propeptide portion might be lost. The rapid protein extraction protocol of the capsule contents proved to be effective. Potential matrix interference from the sample excipient materials was sufficiently diluted and tolerated by the MALDI ionization process.

Peptide mass mapping by MALDI. To determine the protein sequence, peptide mass mapping was first performed. Tryptic digestion of the capsule and tablet products was carried out with the aid of RapiGest SF surfactant, which allowed for an accelerated digestion in 2h. MS spectra of the tryptic peptides generated from SC1 and SC2 are shown in Figure 3. As can be seen, both SC1 and SC2 show consistent MS profiles, indicating both forms of capsules contain the same protein. Similar results were obtained with the 10 mg dosage forms (data not shown). The calculated versus observed tryptic peptides of aequorin-2 in SC1 and SC2 are provided in Table 1. Sequence coverage of 88% and 83% was obtained for SC1 and SC2, respectively. A peptide at *m/z* 2117 was observed in SC1, corresponding to the N-terminus peptide [1-18] of the 196-amino acid residue sequence, consistent with the intact mass measurement and the literature report [11]. This N-terminus peptide was also observed in SC2 but the S/N was below 3. Refining the sample extraction protocol for SC2 should enhance the signal intensity for this N-terminus peptide due to the lower dose strength. This peptide was also missing in the 10-mg capsule and 10-mg tablet preparations, presumably due to the lower dosage. Three short peptides, [WIGR, residues 19-22, *m/z* 531], [AYTK, residues 138-141, *m/z* 482], and [LYGGAVP, residues 190-196, *m/z* 676], have calculated mass values outside the mass detection window of 700-4000 *m/z* and thus were not detected. Peptide sequence [LATDELEK, residues 96-103, *m/z* 918] falls within the mass detection window but was not observed. Instead, it was observed in a longer peptide with one

missed cleavage [LATDELEKYAKNEPTLIR, residues 96-113, m/z 2104] in SC2 and in the 10 mg dosages, but not SC1. Additional peptides of unknown sources were also present in the spectra possibly resulting from the impurities after aequorin-2 protein purification from host cell proteins. Trypsin autolysis products at low level were also observed in addition to a notable peak at *m/z* 842.

Peptide sequencing by MALDI. Tryptic peptides of SC1 and SC2 were subjected to high energy collision induced dissociation to deduce the sequence at the amino acid level. For example, MSMS spectra at *m/z* 1813 of SC1 and SC2 are shown in Figure 4. The fragmentation profiles are consistent with each other, suggesting that they originated from the same peptide. The fragment ions are consistent with the sequence assigned as AAGIIQSSEDCEETFR [residues 142-157]. The remaining MSMS spectra of the tryptic peptides from SC1, corresponding to those also observed for SC2, are annotated and provided in Supplemental Figure S1. A complete list of the calculated versus observed fragment ion masses for all peptides of SC1 is given in Supplemental Table S1. Similar data were generated for the 10-mg dosage forms (data not shown). Note that the rapid digestion procedure produced long tryptic peptides resulting from missed cleavages. These long peptides are compatible with the MALDI process, and the high energy collision-induced dissociation produced information-rich fragmentation spectra for all peptides. As a result, peptide sequencing obtained sequence coverage of 54% for both SC1 and SC2. This sequence coverage reflects the fact that only abundant peptides were selected and fragmented for confirmation, which is a common practice in MALDI MSMS analyses.

LC-HRAM MS of capsules

SC1 capsules were also subjected to LC-HRAM MS analysis. Figure 5A shows the total ion chromatogram (TIC) of the tryptic digest of SC1. Figure 5B is the extracted ion chromatogram (EIC) of the tryptic peptide at *m/z* 906.9078 (+2 charge state), and Figures 5C and 5D show the MS and MSMS spectra for the peptide AAGIIQSSEDCEETFR. Note that this is the same peptide shown in Figure 4. Slight variations in the type and number of product ions are noticeable, resulting from using the two different types of mass spectrometers which operate under different experimental parameters and fragmentation mechanisms. For example, the short peptides that fell outside the detection window during MALDI analysis were observed in LC-HRAM MS analysis, in addition to the N-terminus short peptides MTSK [1-4, *m/z* 466] and QYSVK [5-9, *m/z* 624]. Accurate mass measurement of these singly charged short peptides was obtained with mass errors below 2 ppm. Also the two peptides at *m/z* 918 and *m/z* 676 (the C-terminus peptide), which were not observed using MALDI, were detected and sequenced. Taken together, sequence coverage of 100% and 90% was achieved for MS and MSMS using LC-HRAM MS. Table 2 summarizes all the tryptic peptides of SC1 observed in LC-HRAM MS analysis. Mass errors at or below 2 ppm were obtained for all peptides analyzed. The remaining MSMS spectra of the tryptic peptides of SC1 are annotated and provided in Figure S2. Note that multiple charge states may be present for a given peptide and only one charge state is selected and shown. Overall, comparing Figure S1 with Figure S2 shows that some peptides behave similarly in both fragmentation processes with mutual fragment ions and similar fragmentation profiles, while others provide complementary product ions resulting from the two different fragmentation processes, i.e., high energy versus low energy collision induced dissociation for MALDI and LC-MS, respectively.

During the proteomic database search, aequorin-1 was also a potential match with reasonable sequence coverage compared to that of aequorin-2 based on the sequence coverage and uniqueness of the peptide groups. Aequorin-1 (UniProtKB/Swiss-Prot accession# P07164) is homologous to aequorin-2 with 19 different amino acid residues throughout the entire sequence. The only unique peptide assigned to aequorin-1 was [LWGDALFDIIDKDQNGAISLDEWK, residues 114-137, calculated *m/z* 2762.3671, +1 charge state]. This peptide has the same theoretical mass as [IWGDALFDIVDKDQNGAITLDEWK, residues 114-137], which was assigned to aequorin-2. Fortunately, the fragment ions from b10 to b18 and y6 to y14 are unique for each peptide. Indeed, the y6 ion of aequorin-2 (calculated *m/z* 791.39) was observed at *m/z* 791.39 but the y6 ion of aequorin-1 (calculated *m/z* 777.38) was not present in the spectrum (Figure S2). This was also the case for the b10, y12, y13, and y14 ions. Thus, aequorin-1 was eliminated as a possibility.

LC-HRAM MS of protein powders

The SPP composite was analyzed only using LC-HRAM MS due to the complexity of the matrix and the relatively low abundance of aequorin-2 versus whey proteins. Figure 6A shows the TIC of the tryptic digest of SPP, and includes high abundance peptides from whey proteins including beta-lactoglobulin, alpha-lactalbumin, and serum albumin. The EIC of the peptide at *m/z* 906.9084 is shown in Figure 6B, with the corresponding MS and MSMS spectra in Figures 6C and 6D. Agreement with the accurate mass measurement in Figure 5C, and similarities to the MSMS spectra in Figure 5D, suggest these are the same peptide. Although aequorin-2 in SPP was present at low levels amid the whey protein background, several additional peptides were observed. The tryptic peptides of aequorin-2 in SPP observed in LC-HRAM MS are compared to those of aequorin-2 in SC1 in Table 2. The remaining MSMS spectra of the observed peptides of aequorin-2 in SPP (data not shown) resemble those of SC1 shown in Figure S2. Sequence coverage of 59% was obtained for both MS and MSMS. This sequence coverage is less than that obtained with SC1, because aequorin-2 in SPP was present at low levels, and suffered interference from the high levels of whey proteins in the matrix.

CONCLUSION

Three types of dietary supplements from the same manufacturer, i.e. capsules, chewable tablets and protein powders, that claimed to contain apoaequorin, were analyzed using MALDI-TOF/TOF MS and/or LC-HRAM MS. It was concluded that all of the products contained a related compound ("Aequorin-2"; UniProtKB/Swiss-Prot accession# P02592), a single chain polypeptide with 196 amino acid residues.

MALDI-TOF/TOF-MS allowed a fast, comprehensive evaluation of the capsules and chewable tablets that was facilitated by rapid protein extraction and trypsin digestion procedures. The analysis of two different dose strength capsules (SC1 and SC2) achieved a sequence coverage of 88% versus 83% for MS and 54% versus 54% for MSMS, respectively. The N-terminus peptide of aequorin-2 was detected and confirmed for SC1 whereas the same peptide signal was present with S/N < 3 for SC2.

LC-HRAM MS enabled detailed sequence analysis of aequorin-2 in both capsules and protein powders. Sequence coverage for SC1 was 100% and 90% for MS and MSMS, respectively.

Sequence coverage of 59% for both MS and MSMS was achieved for SPP due to its presence at low levels amid whey protein background. The N-terminus peptide was not detected and confirmed for SPP. Comparing LC-HRAM MS with MALDI-TOF/TOF MS found that some peptides behave similarly in both fragmentation processes (high energy versus low energy CID) with mutual fragment ions and similar fragmentation profiles, while others provide complementary product ions. For sequence analysis of aequorin-2 present in suspect protein powders, the LC-HRAM MS was the method of choice due to interference from complex matrix proteins.

REFERENCES

[1] H. Charbonneau, K.A. Walsh, R.O. McCann, F.G. Prendergast, M.J. Cormier, T.C. Vanaman, Amino acid sequence of the calcium-dependent photoprotein aequorin, Biochemistry, 24 (1985) 6762-6771.

[2] O. Shimomura, F.H. Johnson, Y. Saiga, Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea, J Cell Comp Physiol, 59 (1962) 223-239.

[3] J.F. Head, S. Inouye, K. Teranishi, O. Shimomura, The crystal structure of the photoprotein aequorin at 2.3 A resolution, Nature, 405 (2000) 372-376.

[4] H. Masuda, Y. Takenaka, Y. Shikamoto, M. Kagawa, H. Mizuno, F.I. Tsuji, Chromatography of isoforms of recombinant apoaequorin and method for the preparation of aequorin, Protein Expr Purif, 31 (2003) 181-187.

[5] S. Inouye, S. Aoyama, T. Miyata, F.I. Tsuji, Y. Sakaki, Overexpression and purification of the recombinant Ca2+-binding protein, apoaequorin, J Biochem, 105 (1989) 473-477.

[6] V. Granatiero, M. Patron, A. Tosatto, G. Merli, R. Rizzuto, The use of aequorin and its variants for Ca2+ measurements, Cold Spring Harb Protoc, 2014 (2014) 9-16.

[7] J.M. Kendall, M.N. Badminton, Aequorea victoria bioluminescence moves into an exciting new era, Trends Biotechnol, 16 (1998) 216-224.

[8] S. Inouye, M. Noguchi, Y. Sakaki, Y. Takagi, T. Miyata, S. Iwanaga, T. Miyata, F.I. Tsuji, Cloning and sequence analysis of cDNA for the luminescent protein aequorin, Proc Natl Acad Sci U S A, 82 (1985) 3154-3158.

[9] Y. Ohmiya, S. Kurono, M. Ohashi, T.F. Fagan, F.I. Tsuji, Mass spectrometric evidence for a disulfide bond in aequorin regeneration, FEBS Lett, 332 (1993) 226-228.

[10] J.A. Detert, E.L. Adams, J.D. Lescher, J.A. Lyons, J.R. Moyer, Jr., Pretreatment with apoaequorin protects hippocampal CA1 neurons from oxygen-glucose deprivation, PLoS One, 8 (2013) e79002.

[11] D.L. Moran, A.O. Tetteh, R.E. Goodman, M.Y. Underwood, Safety assessment of the calcium-binding protein, apoaequorin, expressed by Escherichia coli, Regul Toxicol Pharmacol, 69 (2014) 243-249.

Table 1. Calculated versus observed tryptic peptides of aequorin-2 in SC1 and SC2 by MALDI-TOF/TOF-MS.*

$[M+H]^+$ (observed) SC1	$[M+H]^+$ (observed) SC2	$[M+H]^+$ (calculated)	Start - End	Sequence
1064.6218	1064.6473	1064.5007	10 - 18	K.LTSDFDNPR.W
1204.7996	1204.8347	1204.6684	104 - 113	K.YAKNEPTLIR.I
1564.9091	1564.9392	1564.7577	65 - 79	R.HKDAVEAFFGGAGMK.Y
1812.9858	1813.0233	1812.8069	142 - 157	K.AAGIIQSSEDCEETFR.V
1838.0334	1838.0839	1837.8915	23 - 37	R.HKHMFNFLDVNHNGK.I
1883.0114	1883.0586	1882.8251	175 - 189	R.QHLGFWYTMDPACEK.L
1897.1625	1897.2101	1897.0138	47 - 64	K.ASDIVINNLGATPEQAKR.H
1942.1083	1942.1523	1941.9381	80 - 95	K.YGVETDWPAYIEGWKK.L
1982.0239	1982.0698	1981.8477	158 - 174	R.VCDIDESGQLDVDEMTR.Q
-	2104.3762	2104.1285	96 - 113	K.LATDELEKYAKNEPTLIR.I
2117.3606	-	2117.0332	1 - 18	-.MTSKQYSVKLTSDFDNPR.W
2762.6301	2762.7183	2762.3672	114 - 137	R.IWGDALFDIVDKDQNGAITLDEWK.A
2819.6011	2819.6902	2819.4495	38 - 63	K.ISLDEMVYKASDIVINNLGATPEQAK.R
3776.0337	3776.2866	3775.6368	142 - 174	K.AAGIIQSSEDCEETFRVCDIDESGQLDVDEMTR.Q

* All cysteine (C) residues are carbamidomethylated under the experimental conditions described.

Table 2. Calculated versus observed *m/z* of tryptic peptides from aequorin-2 in SC1 and SPP using LC-HRAM MS.

RT (min)	*m/z* (Obs.) SC1	RT (min)	*m/z* (Obs.) SPP	Charge state	*m/z* Calc.	Δ (ppm) capsules	Start- End	Sequence
3.51	381.2127	-	-	1+	381.2132	-1.3	104 - 106	K.YAK.N
3.20	466.2324	-	-	1+	466.2330	-1.3	1 - 4	-.MTSK.Q
4.23	482.2603	-	-	1+	482.2609	-1.2	138 - 141	K.AYTK.A
10.75	531.3031	-	-	1+	531.3038	-1.3	19 - 22	R.WIGR.H
7.51	624.3347	-	-	1+	624.3352	-0.8	5 - 9	K.QYSVK.L
14.54	338.6866	-	-	2+	338.6869	-0.9	190 - 196	K.LYGGAVP.-
	676.3661		-	1+	676.3665	-0.6		
13.01	421.7397	12.43	421.7405	2+	421.7401	-0.9	107 - 113	K.NEPTLIR.I
	842.4721		-	1+	842.4730	-1.1		
11.53	459.7420	-	-	2+	459.7411	2.0	96 - 103	K.LATDELEK.Y
	918.4765		-	1+	918.4749	1.7		
11.70	532.7536	11.09	532.7543	2+	532.7540	-0.8	10 - 18	K.LTSDFDNPR.W
	1064.4993		1064.5023	1+	1064.5007	-1.3		
18.01	549.2807	17.43	549.2812	2+	549.2809	-0.4	38 - 46	K.ISLDEMVYK.A
	1097.5535		1097.5554	1+	1097.5546	-1.0		
17.84	522.2577	17.34	522.2579	3+	522.2574	0.6	65 - 79	R.HKDAVEAFFGGAGMK.Y
	782.8822		782.8825	2+	782.8825	-0.4		
	1564.7574		-	1+	1564.7577	-0.2		
16.17	580.9764	15.58	580.9759	3+	580.9757	1.2	47 - 63	K.ASDIVINNLGATPEQAK.R
	870.9611		870.9608	2+	870.9600	1.3		
	1740.9113		-	1+	1740.9127	-0.8		
15.25	604.9408	14.59	604.9408	3+	604.9405	0.5	142 - 157	K.AAGIIQSSEDCEETFR.V
	906.9078		906.9084	2+	906.9071	0.7		
	1812.8041		-	1+	1812.8069	-1.5		
15.07	613.3018	-	-	3+	613.3020	-0.3	23 - 37	R.HKHMFNFLDVNHNGK.I
	919.4484		-	2+	919.4495	-1.2		
19.89	628.2800	-	-	3+	628.2799	0.2	175 - 189	R.QHLGFWYTMDPACEK.L
	941.9155		-	2+	941.9162	-0.7		
15.25	633.0097	14.77	633.0099	3+	633.0094	0.5	47 - 64	K.ASDIVINNLGATPEQAKR.H
	949.0095		-	2+	949.0105	-1.1		
20.92	647.9842	-	-	3+	647.9842	0	80 - 95	K.YGVETDWPAYIEGWKK.L
	971.4721		-	2+	971.4727	-0.6		
16.35	661.2879	15.77	661.2882	3+	661.2874	0.8	158 - 174	R.VCDIDESGQLDVDEMTR.Q
	991.4285		991.4285	2+	991.4275	1.0		
25.26	921.4619	24.90	921.4612	3+	921.4606	1.4	114 - 137	R.IWGDALFDIVDKDQNGAITLDEWK.A
	1381.6867		1381.6862	2+	1381.6872	-0.4		

Figure 1. The sequence of aequorin-2.

Aequorin-2 (*Aequorea victoria*) (UniProtKB/Swiss-Prot: P02592)

```
1   MTSKQYSVKL TSDFDNPRWI GRHKHMFNFL DVNHNGKISL DEMVYKASDI
51  VINNLGATPE QAKRHKDAVE AFFGGAGMKY GVETDWPAYI EGWKKLATDE
101 LEKYAKNEPT LIRIWGDALF DIVDKDQNGA ITLDEWKAYT KAAGIIQSSE
151 DCEETFRVCD IDESGQLDVD EMTRQHLGFW YTMDPACEKL YGGAVP
```

Figure 2. Intact mass analysis of SC1 by MALDI-TOF/TOF-MS.

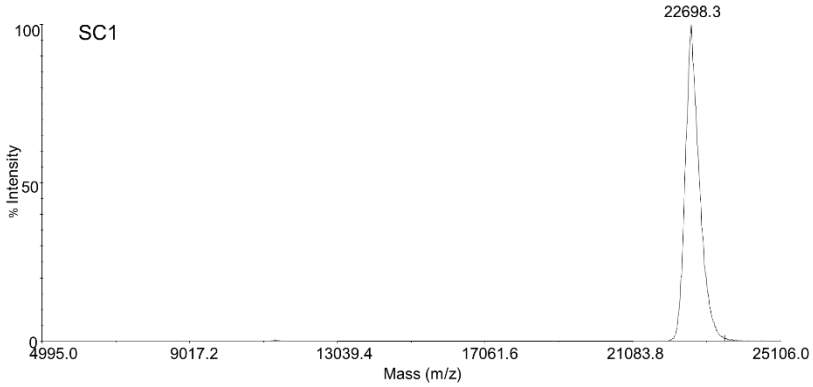


Figure 3. MS spectra of the tryptic digests of A) SC1 and B) SC2 using MALDI-TOF/TOF-MS.

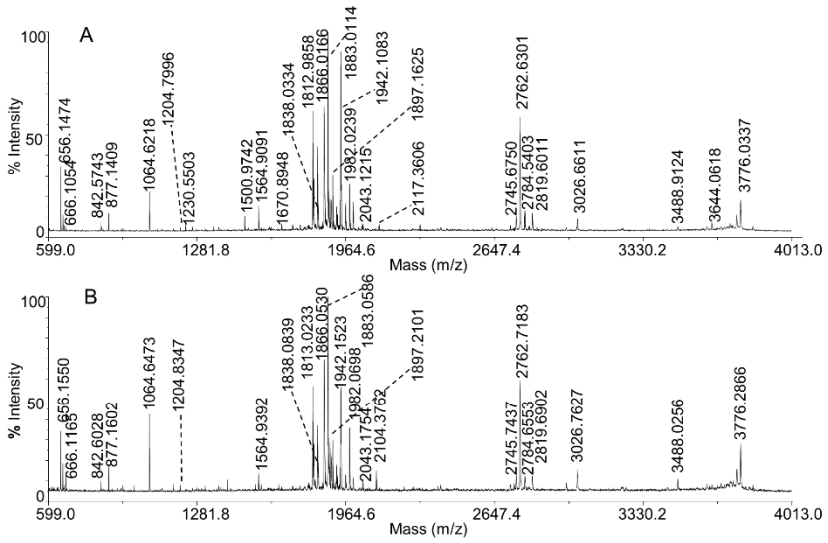


Figure 4. MSMS spectra of the tryptic peptide at *m/z* 1813 of A) SC1 and B) SC2 using MALDI-TOF/TOF-MS. Observed b and y ions are indicated by ⌉ and ⌊ on the sequence.

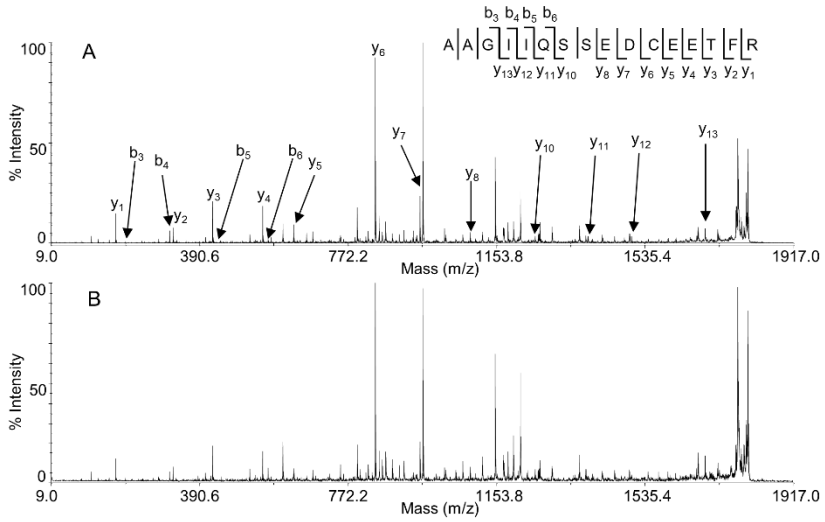


Figure 5. LC-HRAM MS of the tryptic digest of SC1. A) TIC. B) EIC at *m/z* 906.9078 (*z* = +2). C) MS spectrum. D) MSMS spectrum.

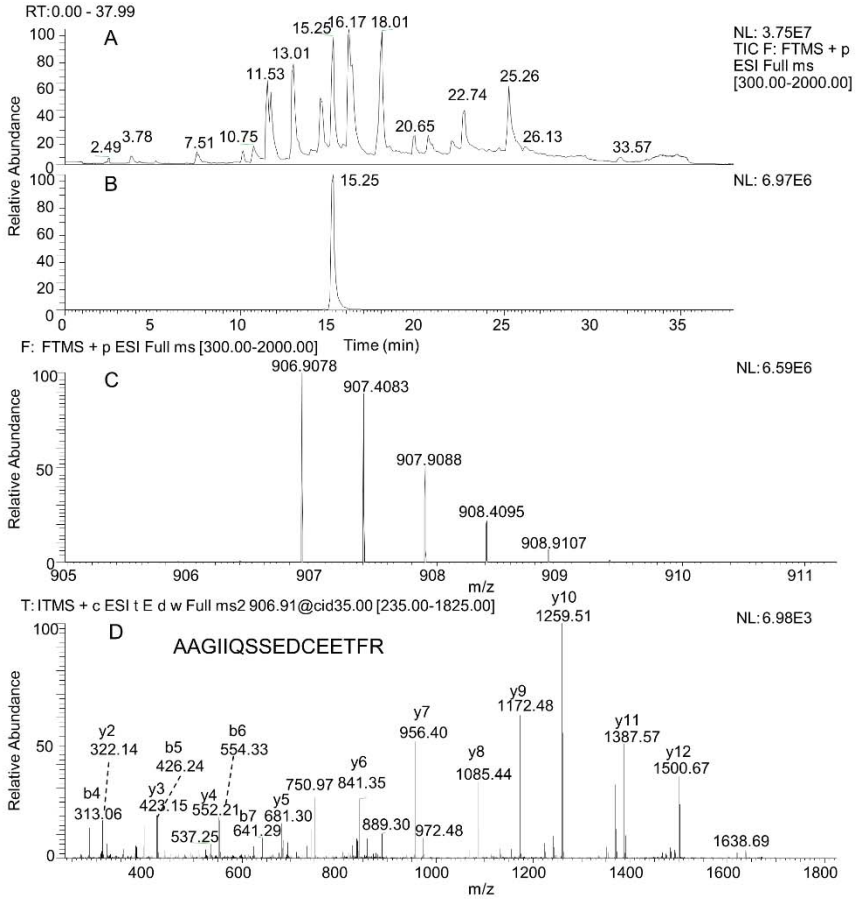


Figure 6. LC-HRAM MS of the tryptic digest of SPP. A) TIC. B) EIC at *m/z* 906.9084 (*z* = +2). C) MS spectrum. D) MSMS spectrum.

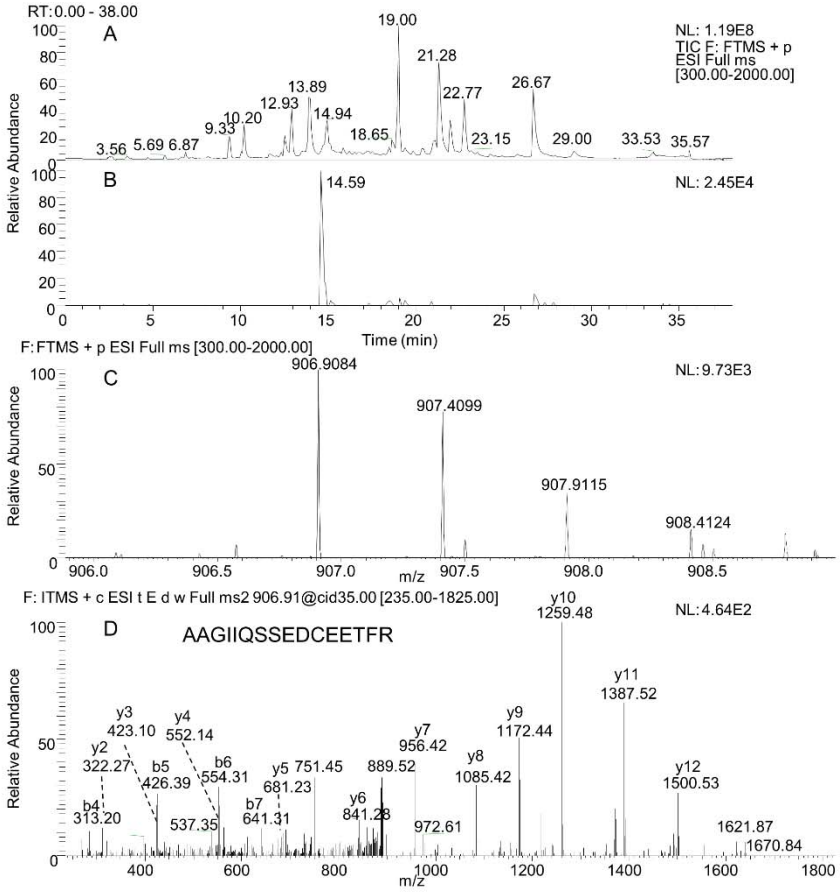


Table S1. Calculated versus observed fragment ions of tryptic peptides of aequorin-2 in SC1 using MALDI-TOF/TOF-MS.

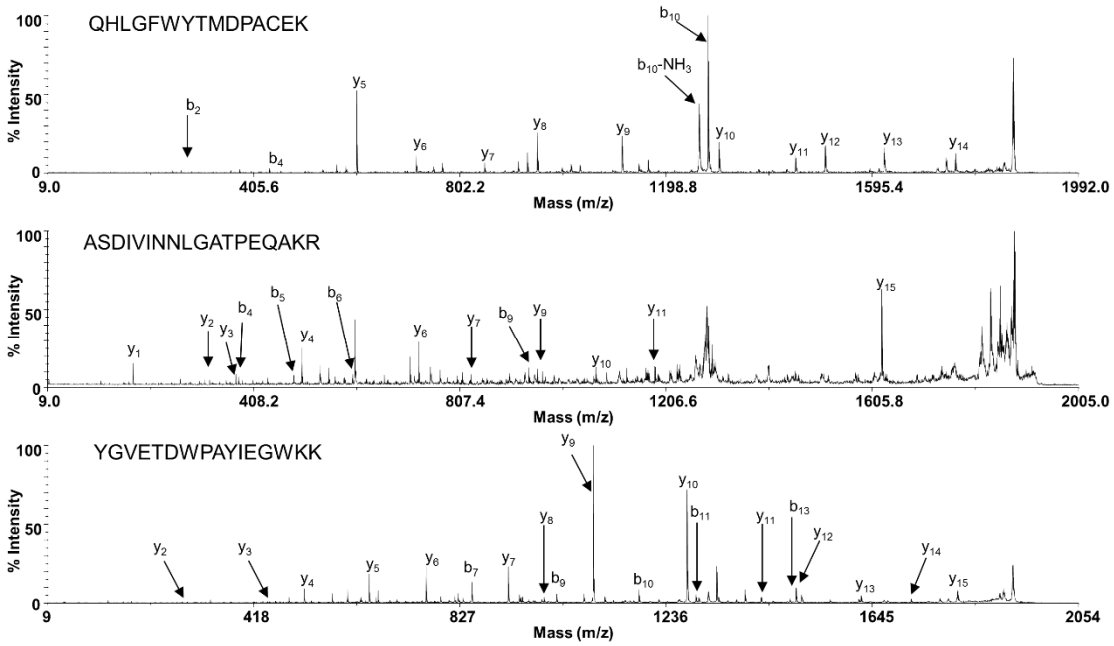
Peptide sequence	Fragment ion	Observed (*m/z*)	Calculated (*m/z*)	Δ Mass (*m/z*)
AAGIIQSSEDCEETFR	b_3	200.1649	200.1030	0.06
	b_4	313.2196	313.1870	0.03
	b_5	426.3181	426.2711	0.05
	b_6	554.3635	554.3297	0.03
	y_1	175.1535	175.1190	0.03
	y_2	322.2226	322.1874	0.04
	y_3	423.2835	423.2350	0.05
	y_4	552.3304	552.2776	0.05
	y_5	681.3722	681.3202	0.05
	y_6	841.4442	841.3509	0.09
	y_7	956.4809	956.3778	0.10
	y_8	1085.5197	1085.4204	0.10
	y_{10}	1259.5968	1259.4845	0.11
	y_{11}	1387.6233	1387.5430	0.08
	y_{12}	1500.7596	1500.6271	0.13
	y_{13}	1613.8204	1613.7112	0.11
QHLGFWYTMDPACEK	b_2	266.1267	266.1248	0.00
	b_4	436.2809	436.2303	0.05
	b_{10}-NH_3	1262.7784	1262.5299	0.25
	b_{10}	1279.7207	1279.5565	0.16
	y_5	604.3477	604.2759	0.07
	y_6	719.3979	719.3029	0.10
	y_7	850.4583	850.3434	0.11
	y_8	951.5138	951.3910	0.12
	y_9	1114.6016	1114.4544	0.15
	y_{10}	1300.6711	1300.5337	0.14
	y_{11}	1447.8174	1447.6021	0.22
	y_{12}	1504.7504	1504.6235	0.13
	y_{13}	1617.8738	1617.7076	0.17
	y_{14}	1755.8950	1754.7665	1.13
ASDIVINNLGATPEQAKR	b_4	387.1828	387.1874	-0.00
	b_5	486.3000	486.2558	0.04
	b_6	599.3854	599.3399	0.05
	b_9	940.5815	940.5098	0.07
	y_1	175.1546	175.1190	0.04
	y_2	303.2048	303.2139	-0.01
	y_3	374.2491	374.2510	-0.00
	y_4	502.3261	502.3096	0.02
	y_6	728.4595	728.4050	0.05
	y_7	829.5060	829.4526	0.05
	y_9	957.5851	957.5112	0.07
	y_{10}	1070.6461	1070.5953	0.05
	y_{11}	1184.7183	1184.6382	0.08
	y_{15}	1623.9971	1623.9177	0.08
YGVETDWPAYIEGWKK	b_7	851.4232	851.3570	0.07
	b_9	1019.5254	1019.4469	0.08
	b_{10}	1182.6179	1182.5102	0.11
	b_{11}	1295.6823	1295.5943	0.09
	b_{12}	1424.7264	1424.6369	0.09

Peptide sequence	Fragment ion	Observed (*m/z*)	Calculated (*m/z*)	Δ Mass (*m/z*)
	b_{13}	1481.7396	1481.6583	0.08
	y_2	275.1585	275.2078	-0.05
	y_3	461.2960	461.2871	0.01
	y_4	518.3222	518.3085	0.01
	y_5	647.3783	647.3511	0.03
	y_6	760.4924	760.4352	0.06
	y_7	923.5657	923.4985	0.07
	y_8	994.5901	994.5356	0.06
	y_9	1091.6667	1091.5884	0.08
	y_{10}	1277.7560	1277.6677	0.09
	y_{11}	1392.7421	1392.6947	0.05
	y_{12}	1493.8243	1493.7423	0.08
	y_{13}	1622.8455	1622.7849	0.06
	y_{14}	1721.9232	1721.8533	0.07
	y_{15}	1778.9144	1778.8748	0.04
VCDIDESGQLDVDEMTR	b_3	375.1095	375.1333	-0.02
	y_1	175.1460	175.1190	0.03
	y_2	276.1206	276.1666	-0.05
	y_3	407.2225	407.2071	0.02
	y_4	536.2847	536.2497	0.04
	y_5	651.3052	651.2767	0.03
	y_6	750.3993	750.3451	0.05
	y_7	865.4248	865.3720	0.05
	y_8	978.5182	978.4561	0.06
	y_9	1106.6250	1106.5146	0.11
	y_{10}	1163.6554	1163.5361	0.12
	y_{11}	1250.6636	1250.5861	0.08
	y_{12}	1379.6779	1379.6107	0.07
	y_{13}	1494.6758	1494.6377	0.04
	y_{14}	1607.7682	1607.7217	0.05
	y_{15}	1722.8605	1722.7487	0.11
IWGDALFDIVDKDQNGAITLDEWK	b_{13}	1488.7778	1488.7369	0.04
	b_{14}	1616.8353	1616.7955	0.04
	y_2	333.1030	333.1921	-0.09
	y_4	577.2369	577.2617	-0.02
	y_5	690.2910	690.3457	-0.05
	y_6	791.3696	791.3934	-0.02
	y_7	904.4226	904.4775	-0.05
	y_8	975.4813	975.5146	-0.03
	y_9	1032.5397	1032.5360	0.00
	y_{10}	1146.5892	1146.5790	0.01
	y_{11}	1274.6802	1274.6375	0.04
	y_{12}	1389.7260	1389.6645	0.06
	y_{13}	1517.8154	1517.7594	0.06
	y_{14}	1632.8435	1632.7864	0.06
	y_{15}	1731.9124	1731.8548	0.06
	y_{16}	1844.9968	1844.9389	0.06
	y_{17}	1960.0133	1959.9658	0.05
	y_{18}	2107.0852	2107.0342	0.05
	y_{19}	2220.1526	2220.1183	0.03
	y_{20}	2291.1987	2291.1554	0.04
	y_{21}	2406.3020	2406.1823	0.12
	y_{22}	2463.2496	2463.2038	0.05

Peptide sequence	Fragment ion	Observed (*m/z*)	Calculated (*m/z*)	Δ Mass (*m/z*)
AAGIIQSSEDCEETFRVCDIDESGQLDVDEMTR	y_{23}	2649.2441	2649.2831	-0.04
	b_{19}	2168.9412	2168.9233	0.02
	b_{21}	2397.0583	2397.0333	0.03
	b_{22}	2526.1187	2526.0759	0.04
	b_{27}	3026.2773	3026.2990	-0.02
	b_{29}	3241.3455	3240.3943	0.95
	y_{1}	175.2022	175.1190	0.08
	y_{3}	407.1771	407.2071	-0.03
	y_{4}	536.2051	536.2497	-0.04
	y_{5}	651.2446	651.2767	-0.03
	y_{6}	750.3159	750.3451	-0.03
	y_{7}	865.3205	865.3720	-0.05
	y_{8}	978.4174	978.4561	-0.04
	y_{9}	1106.4773	1106.5146	-0.04
	y_{10}	1163.5369	1163.5361	0.00
	y_{11}	1250.5017	1250.5681	-0.07
	y_{12}	1379.5813	1379.6107	-0.03
	y_{13}	1494.6445	1494.6377	0.01
	y_{14}	1607.7068	1607.7217	-0.01
	y_{15}	1722.8236	1722.7487	0.07
	y_{17}	1981.7988	1981.8477	-0.05
	y_{21}	2515.1226	2515.1075	0.02
	y_{23}	2804.2131	2804.1808	0.03
	y_{24}	2919.2427	2919.2077	0.04

Figure S1. The remaining MSMS spectra of the tryptic peptides of SC1 using MALDI-TOF/TOF-MS. A) *m/z* 1883, *m/z* 1897 and *m/z* 1942. B) *m/z* 1982, *m/z* 2762 and *m/z* 3775.

A



B

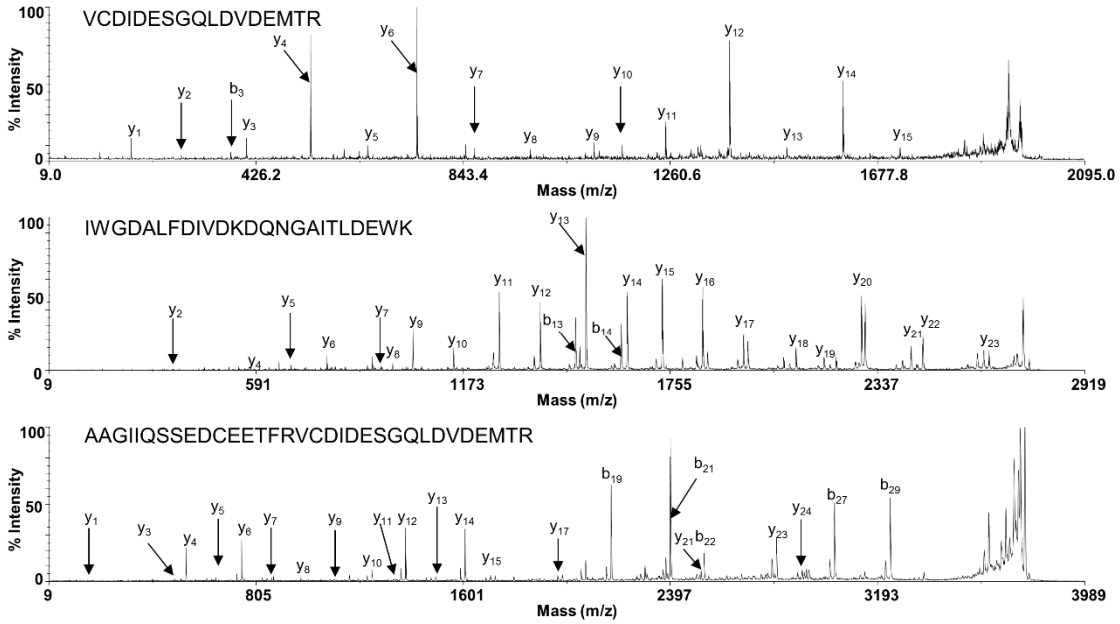
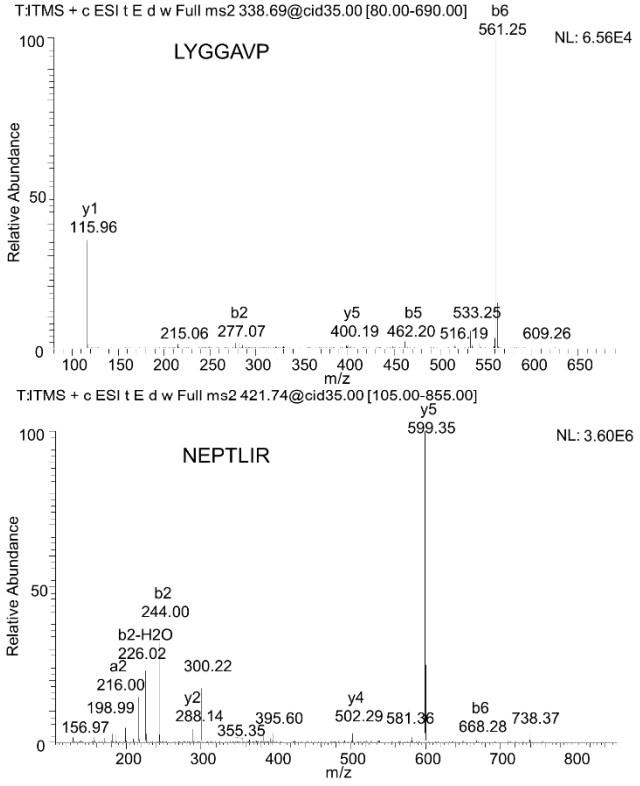
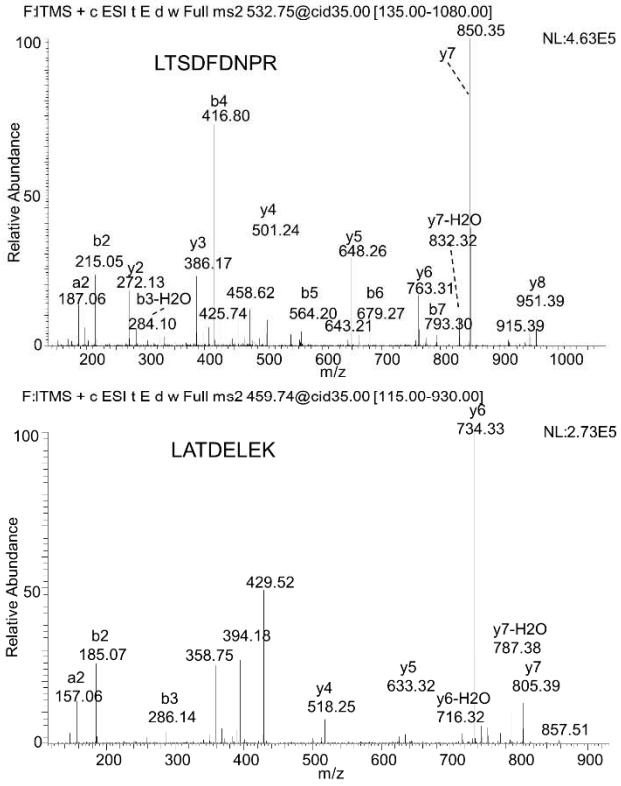
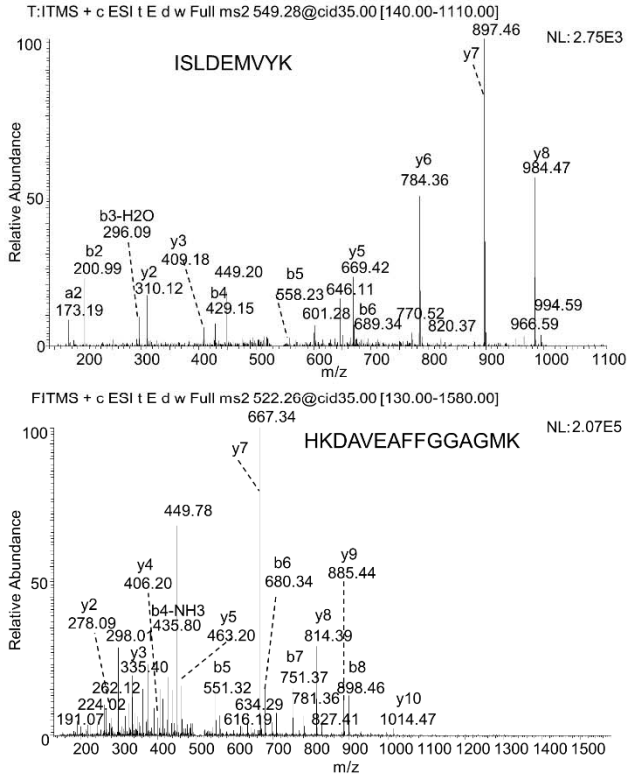
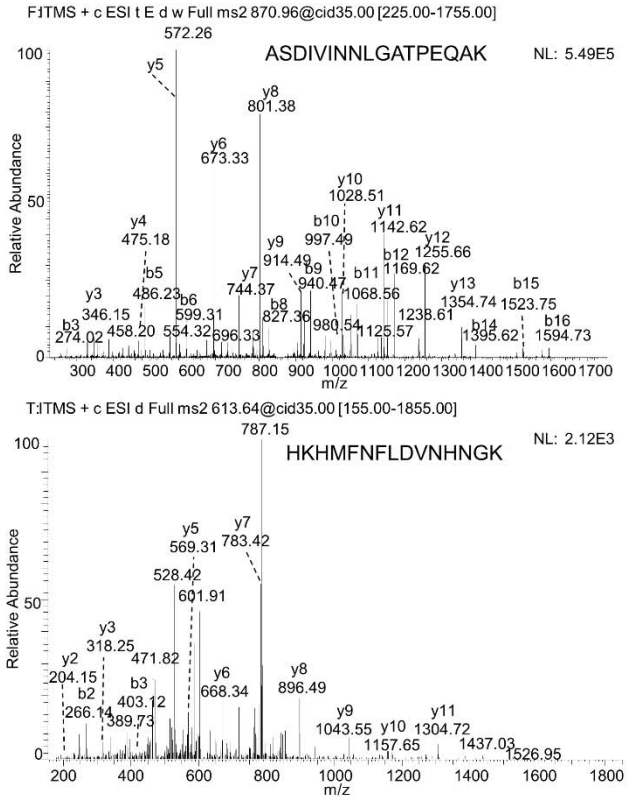

VCDIDESGQLDVDEMTR
% Intensity
Mass (m/z)
IWGDALFDIVDKDQNGAITLDEWK
% Intensity
Mass (m/z)
AAGIIQSSEDCEETFRVCDIDESGQLDVDEMTR
% Intensity
Mass (m/z)

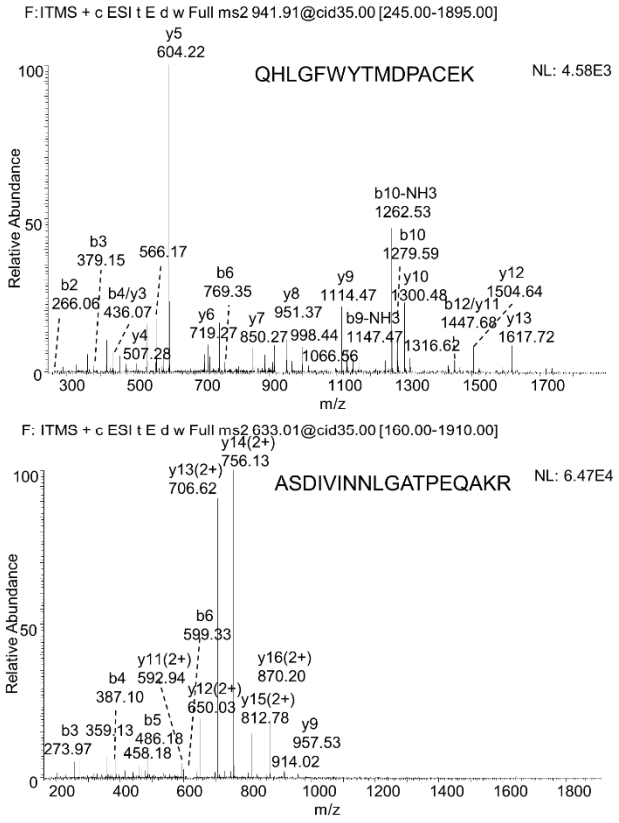
Figure S2. The remaining MSMS spectra of the tryptic peptides of SC1 by LC-HRAM MS.

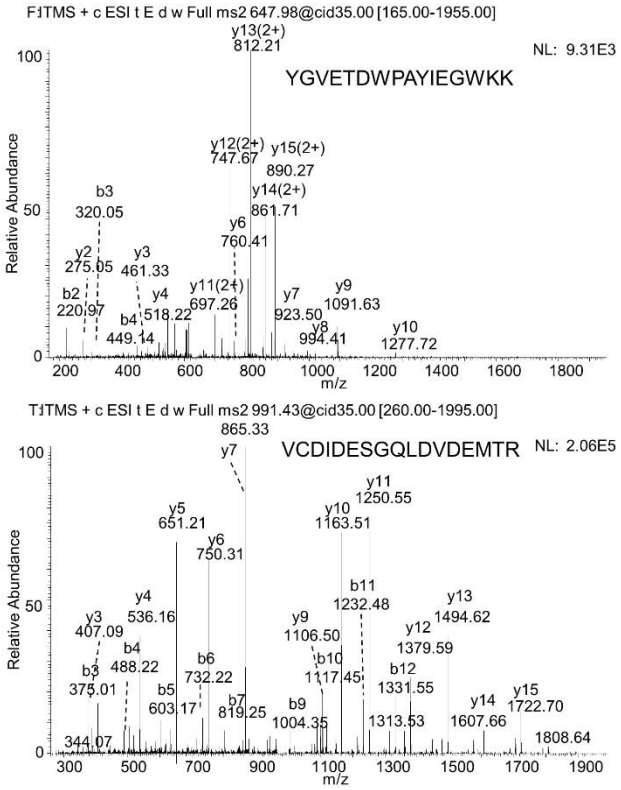


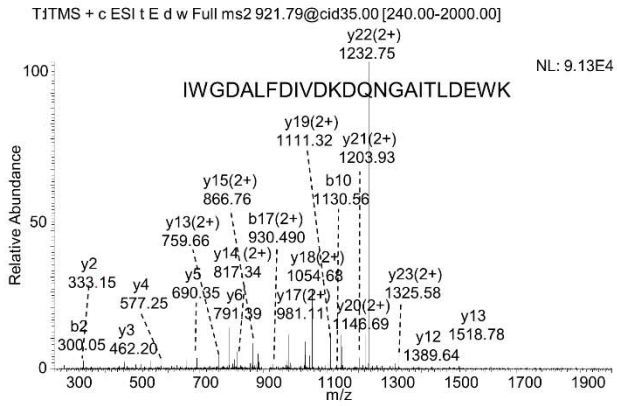

F:ITMS + c ESI t E d w Full ms2 532.75@cid35.00 [135.00-1080.00]
LTSDFDNPR
NL:4.63E5
850.35
y7
b4
416.80
y4
501.24
y7-H2O
832.32
y5
648.26
b2
215.05
y3
386.17
y2
272.13
a2
187.06
b3-H2O
284.10
425.74
458.62
b5
564.20
b6
679.27
y6
763.31
b7
793.30
643.21
y8
951.39
915.39
Relative Abundance
m/z
F:ITMS + c ESI t E d w Full ms2 459.74@cid35.00 [115.00-930.00]
LATDELEK
NL:2.73E5
y6
734.33
429.52
394.18
358.75
b2
185.07
a2
157.06
b3
286.14
y4
518.25
y5
633.32
y7-H2O
787.38
y7
805.39
y6-H2O
716.32
857.51
Relative Abundance
m/z


T:ITMS + c ESI t E d w Full ms2 549.28@cid35.00 [140.00-1110.00]
NL: 2.75E3
ISLDEMVYK
897.46
y7
y6
784.36
y8
984.47
b3-H2O
296.09
b2
200.99
y3
409.18
y2
310.12
449.20
b5
y5
669.42
a2
173.19
b4
429.15
558.23
646.11
601.28
b6
689.34
770.52
820.37
966.59
994.59
Relative Abundance
m/z
FITMS + c ESI t E d w Full ms2 522.26@cid35.00 [130.00-1580.00]
NL: 2.07E5
HKDAVEAFFGGAGMK
667.34
y7
449.78
y4
406.20
b6
680.34
y9
885.44
y2
278.09
b4-NH3
435.80
y5
463.20
298.01
y8
814.39
y3
335.40
b5
b7
751.37
b8
898.46
262.12
551.32
781.36
y10
1014.47
224.02
634.29
191.07
616.19
827.41
Relative Abundance
m/z


FTMS + c ESI t E d w Full ms2 870.96@cid35.00 [225.00-1755.00]
ASDIVINNLGATPEQAK
NL: 5.49E5
Relative Abundance
m/z
TTMS + c ESI d Full ms2 613.64@cid35.00 [155.00-1855.00]
HKHMFNFLDVNHNGK
NL: 2.12E3
Relative Abundance
m/z


F:ITMS + c ESI t E d w Full ms2 941.91@cid35.00 [245.00-1895.00]
QHLGFWYTMDPACEK
NL: 4.58E3
y5 604.22
b10-NH3 1262.53
b10 1279.59
b3 379.15
566.17
b6 769.35
y9 1114.47
y10 1300.48
y12 1504.64
b2 266.06
b4/y3 436.07
y8 951.37
b12/y11 1447.68
y13 1617.72
y6 719.27
y7 850.27
998.44
b9-NH3 1147.47
y4 507.28
1316.62
1066.56
Relative Abundance
m/z
F: ITMS + c ESI t E d w Full ms2 633.01@cid35.00 [160.00-1910.00]
ASDIVINNLGATPEQAKR
NL: 6.47E4
y14(2+) 756.13
y13(2+) 706.62
b6 599.33
y11(2+) 592.94
y16(2+) 870.20
b4 387.10
y12(2+) 650.03
y15(2+) 812.78
b5 486.18
b3 359.13
273.97
458.18
y9 957.53
914.02
Relative Abundance
m/z


FTMS + c ESI t E d w Full ms2 647.98@cid35.00 [165.00-1955.00]
y13(2+)
812.21
NL: 9.31E3
YGVETDWPAYIEGWKK
y12(2+)
747.67
y15(2+)
890.27
y14(2+)
861.71
b3
320.05
y6
760.41
y2
275.05
y3
461.33
b2
220.97
y4
518.22
y11(2+)
697.26
y7
923.50
y9
1091.63
b4
449.14
y8
994.41
y10
1277.72
Relative Abundance
m/z
TTMS + c ESI t E d w Full ms2 991.43@cid35.00 [260.00-1995.00]
865.33
y7
VCDIDESGQLDVDEMTR
NL: 2.06E5
y11
1250.55
y5
651.21
y10
1163.51
y6
750.31
b11
1232.48
y4
536.16
y13
1494.62
y3
407.09
y9
1106.50
y12
1379.59
b4
488.22
b6
732.22
b10
1117.45
b12
1331.55
b3
375.01
b5
603.17
b7
819.25
b9
1004.35
y15
1722.70
y14
1607.66
1313.53
1808.64
344.07
Relative Abundance
m/z


TITMS + c ESI t E d w Full ms2 921.79@cid35.00 [240.00-2000.00]
y22(2+)
1232.75
NL: 9.13E4
IWGDALFDIVDKDQNGAITLDEWK
y19(2+)
1111.32
y21(2+)
1203.93
y15(2+)
866.76
b10
1130.56
y13(2+)
759.66
b17(2+)
930.490
y14(2+)
817.34
y18(2+)
1054.68
y23(2+)
1325.58
y2
333.15
y4
577.25
y5
690.35
y6
791.39
y17(2+)
981.11
y20(2+)
1146.69
y13
1518.78
b2
300.05
y3
462.20
y12
1389.64
Relative Abundance
m/z