

Extensive Characterization of *Tupaia belangeri* Neuropeptidome Using an Integrated Mass Spectrometric Approach

Filomena Petruzzello,[†] Laetitia Fouillen,[†] Henrik Wadensten,[‡] Robert Kretz,[†] Per E. Andren,[‡] Gregor Rainer,[†] and Xiaozhe Zhang^{*,†}

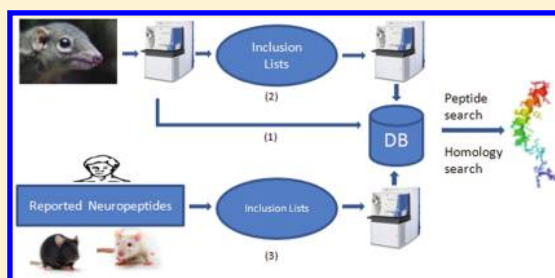
[†]Visual Cognition Laboratory, Department of Medicine, University of Fribourg, Chemin de Musee 5, Fribourg, CH-1700, Switzerland

[‡]Department of Pharmaceutical Biosciences, Medical Mass Spectrometry, Uppsala University, Biomedical Centre, Box 591, SE-75124 Uppsala, Sweden

S Supporting Information

ABSTRACT: Neuropeptidomics is used to characterize endogenous peptides in the brain of tree shrews (*Tupaia belangeri*). Tree shrews are small animals similar to rodents in size but close relatives of primates, and are excellent models for brain research. Currently, tree shrews have no complete proteome information available on which direct database search can be allowed for neuropeptide identification. To increase the capability in the identification of neuropeptides in tree shrews, we developed an integrated mass spectrometry (MS)-based approach that combines methods including data-dependent, directed, and targeted liquid chromatography (LC)–Fourier transform (FT)-tandem MS (MS/MS) analysis, database construction, de novo sequencing, precursor protein search, and homology analysis. Using this integrated approach, we identified 107 endogenous peptides that have sequences identical or similar to those from other mammalian species. High accuracy MS and tandem MS information, with BLAST analysis and chromatographic characteristics were used to confirm the sequences of all the identified peptides. Interestingly, further sequence homology analysis demonstrated that tree shrew peptides have a significantly higher degree of homology to equivalent sequences in humans than those in mice or rats, consistent with the close phylogenetic relationship between tree shrews and primates. Our results provide the first extensive characterization of the peptidome in tree shrews, which now permits characterization of their function in nervous and endocrine system. As the approach developed fully used the conservative properties of neuropeptides in evolution and the advantage of high accuracy MS, it can be portable for identification of neuropeptides in other species for which the fully sequenced genomes or proteomes are not available.

KEYWORDS: neuropeptide, neuropeptidomics, high accuracy tandem MS, directed and targeted LC–MS/MS analysis, *Tupaia belangeri*



INTRODUCTION

Neuropeptides are expressed as signaling molecules in all animal groups which possess a nervous system.^{1,2} The sequences of neuropeptides are determined by genes and highly conserved across species. Neuropeptides play signaling roles in both nervous and endocrine systems and thus are involved in diverse physiological processes such as mood, sleep, pain, reward, and social behaviors.^{1,3–6} Comprehensive characterization of neuropeptides in various species is essential to provide the fundamental understanding of neuropeptide function in basic neuroscience, pharmaceutical, and/or clinical research.

MS has become the method of choice for determining the exact sequence, including post-translational modifications (PTMs) of peptides and proteins from complex biological samples.^{7,8} Compared to traditional techniques such as immunocytochemistry,⁹ MS-based neuropeptidomic studies permit faster and more specific structural identification.^{6,10–16} The characterization of endogenous neuropeptides in animal models such as mice and rats has provided a wealth of information about signaling molecules in the nervous and endocrine systems.^{14,17}

Until now, the most commonly used neuropeptidomics strategies identify neuropeptides by using data-dependent LC–MS/MS analysis, where parent mass selection is usually determined solely according to the abundances of respective peptides, without considering their biological properties. This can result in low identification rates for neuropeptides, because many of them are present in the brain at low concentrations and are thus rarely selected for fragmentation. Neuropeptide identification also depends critically on the quality of MS and MS/MS spectra of a single sequence.^{18–21} Thus, it remains challenging in neuropeptidomics to confirm the sequences of neuropeptides that have low fragmentation information MS/MS spectra. This is particularly true when using the common collision-induced dissociation (CID) mode, where many neuropeptides with long sequences exhibit irregular fragmentation patterns and therefore cannot be identified with high confidence.²² Additional information is therefore required to overcome these

Received: July 28, 2011

Published: November 09, 2011

limitations, for example, knowledge about highly conserved structural motifs or chromatographic characteristics of known peptides.

Our present study targets neuropeptides from tree shrews and aims to provide an in-depth neuropeptidomics investigation. Tree shrews are small animals that are similar in size to rats but phylogenetically very close to primates.^{23,24} Neuropeptidomic research on an animal model such as the tree shrew is challenging because complete proteome information is not yet available. Therefore, a sequence search on its own protein databases, as commonly done for mice and rats, cannot be performed in this species. This complicates the comprehensive characterization of neuropeptides in the tree shrew. However, due to the highly conserved nature of neuropeptides,² we can predict that not all, but a large number of neuropeptides in tree shrews will have sequences identical to those in related mammalian species. Many of the other neuropeptides may have highly similar structures to their sequence equivalents, with only a small fraction of amino acid substitutions. Therefore, a MS-based approach that is capable of integrating known structural characteristics of neuropeptides will allow the identification of many endogenous tree shrew peptides in a fast and efficient manner.

Fully using the advantages of MS together with known properties of neuropeptides, we developed an integrated MS-based approach to characterize neuropeptides in tree shrews. The data acquisition was conducted by combining (i) data-dependent, (ii) directed, and (iii) targeted LC–FT-MS/MS analysis. We thus take advantage of recent developments in proteomics, which have demonstrated that the use of directed or targeted LC–MS/MS is able to improve detection capability and identification efficiency for peptides by rearrangement of parent ion selections.^{25,26} We used a directed LC–MS/MS analysis, conducted according to the results of a data-dependent analysis, while the targeted LC–FT-MS/MS analysis was undertaken using parent ion lists of neuropeptides or fragments that were reported in the literature for other mammalian species. All the MS information was then subjected to de novo sequencing-based hybrid spectral analysis.^{27,28} To identify the peptides that have low information content MS/MS spectra, we used truncated peptides and/or chromatographic characteristics to correct and confirm their sequences. Using our integrated approach, a large number of neuropeptides that have sequences identical or homologous to those in any other species were identified with high confidence from tree shrews.

EXPERIMENTAL SECTION

Materials

LC–MS grade acetonitrile and formic acid were purchased from Fisher Scientific (New Jersey, USA) and Fluka (Wisconsin, USA), respectively. Acetic acid was purchased from Fluka (Buchs, Switzerland). Pure water was prepared by GenPure system (TKA, Niederelbert, Germany). Siliconized microcentrifuge tubes (2 mL) were purchased from Eppendorf (Hamburg, Germany). Microcon centrifugal filter devices (Vivacon 500) were purchased from Sartorius AG (Goettingen, Germany).

Animals

Tree shrews (*Tupaia belangeri*) were used in experiments ($n = 3$). All animals were housed under constant temperature and humidity with free access to food and water. All procedures with

live animals were conducted with protocols approved by the veterinary office of Fribourg, Switzerland.

Sample Preparation

To reduce the interference peptides produced from fragmentation of proteins,¹⁴ restricted temperature control was used to minimize the degeneration of the endogenous neuropeptides. Two male tree shrews of six years old were sacrificed by decapitation after anesthetized with ketamine (100 mg/kg, Streuli Pharma AG, Uznach, Switzerland), then the brains were heat stabilized by using Denator irradiation (Denator AB, Gothenburg, Sweden) as described elsewhere.²⁹ The left and right striatum were dissected from the denaturated brain.

Neuropeptides were extracted three times from each striatum. The three-step extraction was conducted using a gradient of different methanol solutions: (i) 10% methanol containing 0.2% formic acid, (ii) 30% methanol containing 0.2% formic acid, and (iii) 50% methanol containing 0.2% formic acid, respectively. In each extraction step, 5 μ L of solution was used per 1 mg of tissue, and the sample was homogenized twice (each time 20 s) within 1 min by a Precellys 24 homogenizer (Bertin Technologies, Montigny-le-Bretonneux, France). After homogenization, the sample was centrifuged at 22000g for 60 min at 4 °C. All supernatants obtained from the three step extractions were mixed and filtered on a 10 kDa cutoff filter (Vivacon 500, Sartorius AG, Goettingen, Germany) by centrifuging for 90 min at 14000g at 4 °C. The usage of organic solvent in samples was demonstrated to be necessary to increase the recovery of large peptides in injection, separation, and storage.^{30,31} The calculated content of methanol in the final sample solution was 30%.

FT-MS Data Acquisition

The peptide extracts from the tree shrew brain were analyzed using a LTQ-Orbitrap Discovery (Thermo Fisher Scientific, Bremen, Germany) coupled to a 2D NanoLC (Eksigent Technologies, USA). In a LC–MS analysis, the extract was repeatedly injected 5 times (5 μ L/times) with a 1D pump (Chanel 1) to a trap column (100 μ m ID, 2 cm long), which was packed with a ReproSil-Pur C18 AQ particles (5 μ m, 100 Å; Dr. Maisch GmbH, Ammerbuch-Entringen, Germany) in a peek column holder (Upchurch, Oak Harbor, USA). The interval between each two injections was 3 min. The trap column was kept eluting with 2% acetonitrile and 98% water containing 0.2% formic acid. The elution direction of the trap column was reversed through a 10-port valve when it was switched to couple with the analytical column. The analytical column used C18 AQ (3 μ m, 100 Å; Dr. Maisch GmbH) as medium which was packed in a Picofrit capillary with an emitter tip of 10 μ m (NewObjective, Cambridge, USA). The mobile phase A and B in 2D pump (Chanel 2) were 0.2% formic acid and 95% acetonitrile containing 0.2% formic acid, respectively. The mobile phases were eluted on the analytical column at 300 nL/min with a gradient profile as 0–6 min, 2% B; 6–12 min, 2–20% B; 12–80 min, 20–50% B; 80–85 min, 50–60% B; 85–90 min, 60–95% B; and 90–100 min, 95% B.

Data acquisition on the LTQ-FTMS instrument consisted of a full FTMS scan event at a mass range of 350–2000 m/z . The lock mass (445.120025 from polydimethylcyclsiloxane) was used for real time internal recalibration.³² The mass resolution for each scan event was kept at 30 000. The first and fifth most intensive ions were selected and fragmented using CID. Dynamic exclusion was set as a repeat count of 1, exclusion duration of 180 s, and a repeat duration of 30 s. Other MS/MS parameters were set:

isolation width, m/z 2; normalized collision energy, 30%; activation Q, 0.25; and activation time, 50 ms.

The peptides in tree shrew brain extracts were analyzed using three methods which included data-dependent acquisition (DDA), directed and targeted LC–FT-MS/MS. The brain samples were first analyzed using DDA-LC–FT-MS/MS. Then the samples were analyzed using directed LC–FT-MS/MS as described elsewhere.²⁶ In brief, the m/z of peptides in the DDA-LC–FT-MS/MS data were examined with the software (Sieve, 1.2 version, Thermo, CA, USA) and filtered using charge states and intensities. The MS1 signals were determined with the following parameters: mass range of 0.02 Da, intensity threshold of minimal 100 000. The features with charge states ranging from 2 to 5 were collected in the inclusion mass list of tree shrew features for the directed LC–MS/MS analysis. In the targeted LC–MS/MS analysis, the inclusion lists of neuropeptides were collected from the neuropeptidomic studies conducted on mice,^{17,19} rats,^{11,33,34} and humans,^{35,36} respectively. The parent mass lists were loaded to the FT-MS/MS scan for selective fragmentation the target peptides. To acquire high quality MS/MS information, 500–750 ms maximum injection time was used in FT-MS/MS scan.

Spectral Interpretation

All the raw LC–FT-MS/MS data were subjected to Peaks Studio 5.2 (BSI, Canada) for spectral interpretation.²⁷ Peaks Studio 5.2 is composed of functions such as Data Refinement, Auto De Novo, Peaks Search, and Spider (homology search).³⁷ The Data Refinement program allows correcting of the parent mass and charge states to provide accurate monoisotopic mass of a peptide. The scans of quality value > 0.3 were kept for further sequence analysis. Data processing, including peak centroiding, charge deconvolution, and deisotope, was conducted for data refinement.

The refined data were subjected to Auto De Novo program for sequencing with the mass tolerance of parent ions and product ions set at 10 ppm and 0.05 Da, respectively. No enzyme was specified for cleavage. Variable PTMs, including, amidation (C-terminal), acetylation (N-terminal), pyroglutamalytion from glutamic acid, and glutamine (N-terminal) were selected in de novo sequencing.

De novo sequencing-based protein ID search was used to sequence tree shrew peptides that are identical to those from other species. The database search was conducted on a customized database constructed using predicted peptides from species in Euarchontoglires using a method described elsewhere.¹⁹ Briefly, the database was made with predicted neuropeptides of known precursors from Euarchontoglires species (50 377 entries). The cleavage is made according to the following template: (K/R)X_m(K/R)X_k(K/R)X_n(K/R) where X is any amino acid, K is lysine and R is arginine, and m and n correspond to 0, 2, 4, 6, and k corresponds to 3–50 amino acids. The residues X_k represent the predicted neuropeptide sequence.^{18,38} The database search was first performed for peptides with +1, +2, and +3 charge states using the mass tolerance of parent ions and product ions set at 10 ppm and 0.05 Da, respectively. Then the search was performed for peptides with $\geq +4$ charge states using the mass tolerance of parent ions and product ions set at 10 ppm and 2 Da, respectively. Estimation of false positives was conducted by searching all spectra against decoy databases. The cutoff p -value for peptide identification in Peaks search was $< 1 \times 10^{-3}$. Homology search (mass tolerance 0.5 Da) was used to find candidates

of peptides that have a high degree of homology to those from other species. The homology search was conducted on the de novo sequencing results using the same PTMs as used in Auto De Novo program.

To confirm sequences identified from the database search, all the search results were subjected to manual inspection. A sequence was considered correct only if they match all the following criteria: (1) the mass of a peptide must have been calculated from the monoisotopic ions of a peptide; (2) all the database search results were inspected with de novo sequencing results; (3) the peptide mass had to be within 10 ppm of the theoretical mass; (4) for peptides with +1, +2, and +3 charge states, the major fragments observed in MS/MS have to match within 0.05 Da to predicted monoisotopic fragmentation ions. For peptides with $\geq +4$ charge states, 1 or 2 Da shift is allowed to inspect the major fragments according to their predicted monoisotopic ions. (5) The fragmentation information must be enough to recognize the alignment of amino acids, in particular if they fall in the substitution positions across the adjacent species.

For comparison of neuropeptides between tree shrews and mice, the MS/MS data of mouse neuropeptide were searched using the same parameter as for tree shrew peptides except using the Swepep database (precursor, 2006-02-15 version).

The predicted fragmentation ions were calculated using the MS-Product tool in ProteinProsepector (v.5.7.2, Mass Spectrometry Facility, University of California, San Francisco). The sequence identity analysis was acquired by searching the peptide sequences in Uniprot. The statistic analysis was conducted with software Origin 7.5 (OriginLab Corporation, Massachusetts, US).

BLAST Analysis

The sequence identities of identified tree shrew peptides across different species were conducted using the BLAST analysis function in Uniprot. The statistic analysis of identities (paired t test) was performed with Origin 7.5 (OriginLab Corporation, Massachusetts, US). For truncated peptides, only the longest one was counted for acquiring the mean identity of tree shrew peptides. For example, when both the peptide PPEGVLGALLRV and its truncated peptide PPEGVLGALLR are identified, only the peptide PPEGVLGALLRV is counted to avoid repeated calculation of the peptide sequence derived from its neuropeptide precursor protein.

RESULTS

Combined LC–MS/MS Analyses of Tree Shrew Brain Samples

In the present study, we conducted a neuropeptidomic analysis of tree shrews by using an integrated approach that included data-dependent, directed LC–MS/MS analysis, and targeted LC–MS/MS analysis. The general workflow of the LC–MS/MS analysis methods described in this study is outlined in Figure 1. The data-dependent LC–MS/MS analysis was conducted with parent ions automatically selected for fragmentation. The subsequent directed LC–MS/MS analysis selectively used the features/peptides acquired in data-dependent LC–MS/MS analysis and removed undesired features/peptides. In the targeted LC–MS/MS analysis, inclusion lists were formed from collected m/z of reported neuropeptides from mice, rats, and humans, respectively. All the LC–MS/MS data were subsequently

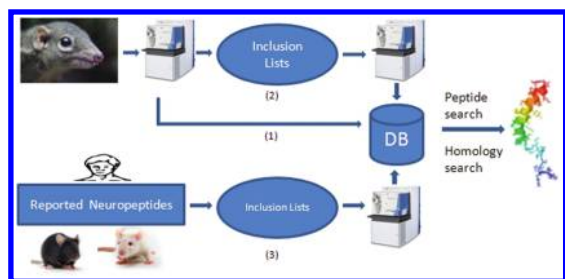


Figure 1. Workflow of tree shrew neuropeptide identification using combined LC–MS/MS analysis methods. Neuropeptide extracts from tree shrew brain were analyzed by data-dependent LC–MS/MS (1). Inclusion lists of tree shrew are formed by selectively extracting MS1 peaks from previously acquired LC–MS map and subject to the next round of directed LC–MS/MS analysis (2). Meanwhile, the inclusion lists of other species, including mice, rats, and human, are formed by collecting *m/z* of reported neuropeptides. These inclusion lists are loaded into instrumental methods separately for to further targeted LC–MS/MS analysis (3). De novo-based database (DB) search and homology search are conducted for LC–MS/MS data and finally generate a identification list of tree shrew peptides.

subjected to de novo sequencing, database search (Protein ID search) for characterization of peptides that have sequences identical to those in one or more other species. Meanwhile, the LC–MS/MS data acquired in the first and second steps were used for de novo sequencing-based homology search for characterization of neuropeptides with substituted amino acids in the sequence. The use of directed and targeted LC–MS/MS analysis allowed the number of identified neuropeptides to increase remarkably (from 70 to 107) with limited times of LC–MS/MS analysis.

Identification of Tree Shrew Peptides with Substituted Amino Acids in Their Sequences

To identify these tree shrew peptides, we used homology search based on de novo sequencing results. In total we identified 15 tree shrew peptides with sequences that have amino acid substitution(s) in comparison to their homologues in other species (Table 1). Figure 2 represents the spectral identification of a peptide LSDDDRVIWAEQQYQKERS with FT-MS/MS in high confidence. BLAST analysis showed this peptide has a high degree of identity to the peptide equivalent of endocrine convertase 1 from mice (89%), rats (89%), and humans (94%), respectively.

It is worth mentioning that clustering a group of truncated peptides will assist in prediction, identification, and confirmation of their correct sequences. Truncated peptides are commonly observed in neuropeptidomics studies, even if strict sample preparation procedures were used.^{11,29,33–35} Many truncated peptides detected in neuropeptide analysis were possibly derived from endogenous proteolytic processing and have their distinct physiological functions. Because of the overlap in the sequence of truncated peptides, identification of a truncated peptide will assist in sequencing the original peptide or other truncated peptides. For example, the peptide LGELFNYPYFDPLQWKSSR had an atypical fragmentation pattern (Supplementary Figure S1, Supporting Information). It was therefore difficult to identify the C-terminal sequence. In contrast, a truncated peptide, PYFDPLQWKSSRFE, had high-confident sequence based on the de novo sequencing results (Supplementary Figure S2, Supporting Information). With PYFDPLQWKSSRFE identified, the C-terminal

of the peptide LGELFNYPYFDPLQWKSSR was then confirmed. Further examination of homology search results allowed confirmation of the sequences of another two truncated peptides, LGELFNYPYFDPLQWKSS and LGELFNYPYFDPLQWKS. The same principle fully used the advantage of de novo sequencing and was therefore applied to other groups of truncated peptides in this study.

Characterization of Tree Shrew Neuropeptides Identical to Those in Other Species

Using the de novo sequencing-based protein ID search function, we characterized 92 peptides that have sequences identical to those from one or more other species (Table 1). These peptides included many classical neuropeptides and a large number of fragment peptides from neuropeptide precursors. These fragment peptides may potentially be biologically active neuropeptides. This hybrid spectral analysis used the advantage of the highly accurate FT-MS and FT-MS/MS data and generated high confidence in identification of these peptides, which allow the majority of the peptide fragment ions assigned, with a *p*-value lower than 1×10^{-3} , a parent mass error of <10 ppm, and a fragment ion mass error of <0.05 Da. Figure 3 depicts the examples of assignment of FT MS/MS spectra for neuropeptide K fragment (1–24). This peptide has information-rich tandem MS spectrum. The subsequent de novo sequencing plus Protein ID search against a customized database allows this peptide to be identified in a fast manner with high confidence.

In contrast to peptides that have information-rich tandem MS spectra, several identified long neuropeptides, such as neuropeptide Y and GAV, have low information content MS/MS spectra, and thus their sequences are still questionable despite high database search scores. Homology analysis indicated that most of these peptides have sequences identical to those in mice. To further verify the sequence of such peptides, we used mouse brain samples to examine their structures. Figure 4A outlines the workflow for the sequence verification on these peptides. The extracts from tree shrew brain, mouse brain, and their mixtures were analyzed using the same LC–MS/MS method. In this case, a tree shrew peptide can be considered as identical to the mice peptide only if they have the same chromatographic and mass spectrometric behaviors. Chromatographic and mass spectrometric characteristics are useful factors to confirm the sequence of peptides, especially with a high accuracy MS.³⁹ Figure 4B,C represent the process for confirming the sequence of neuropeptide Y. Neuropeptide Y has an uneven fragmentation pattern due to the existence of proline.³¹ Using high accuracy MS and MS/MS, neuropeptide Y was identified by database search based on the high information content in its N-terminal. The results showed that tree shrew neuropeptide Y has a fragmentation pattern highly similar to that of mouse neuropeptide Y (Figure 4B) with the same parent mass. However, its C-terminal fragmentation spectrum has much less information content, and importantly the amino acid substitutions of neuropeptide Y in different mammalian species occur in this region. It was difficult to determine by MS/MS alone whether tree shrew neuropeptide Y has sequence identical to mouse or human neuropeptide Y. Further comparative LC–MS/MS analysis demonstrated that neuropeptide Y in tree shrew, mouse samples, and their mixture had the same chromatographic behavior (retention time, peak shape) and mass spectrometric properties (parent mass and fragmentation patterns). On the basis of this

Table 1. Endogenous Peptides Identified from Tree Shrews^a

precursor	uniprot accession number	peptide name	sequence	P-value ^c or Spider score	M/Z	Z	M _r (calc)	mass error (ppm)
CART	P56388		IPYEKYGQVPM	1.4 × 10 ⁻⁴⁵	522.6190	3	1564.8323	1.87
CART	P56388		IPYE	5.9 × 10 ⁻¹²	634.3441	1	633.3373	0.87
cerebellin	Q9R171	cerebellin	SGSAKVAFAIRSTNH	1.4 × 10 ⁻⁴⁵	544.9521	3	1631.8381	2.09
cholecystokinin	P09240	cholecystokinin-33	NLQSLDPSHRIS ^b	8.9 × 10 ⁻¹⁶	456.2406	3	1365.7000	0.09
delta preprotachykinin I	Q9Z0K2		DAGHGQISHKMGYERSAMQNYE	1.6 × 10 ⁻¹³	628.0296	4	2508.0908	0.58
delta preprotachykinin I	Q9Z0K2		DAGHGQISHKMGYE ^b	2.8 × 10 ⁻¹³	510.5653	3	1528.6729	0.80
delta preprotachykinin I	Q9Z0K2		DAGHGQISHKMGYERSAMQNYE	4.3 × 10 ⁻⁸	628.0308	4	2508.0908	1.27
gamma preprotachykinin	Q9Z0K1		ALNSVGYERSAMQNYE ^b	1.4 × 10 ⁻⁴⁵	916.4210	2	1830.8208	3.60
gamma preprotachykinin	Q9Y494		LPEPFEHLQ ^b	4.7 × 10 ⁻¹²	611.8285	2	1221.6394	2.50
gonadoliberin	Q53HY8		Q(-17.03)HWSYGLRPG(-0.98)	1.4 × 10 ⁻⁴⁵	591.7975	2	1181.5730	6.41
neuroendocrine convertase 1	P63239		LSDDDRVITWAEQQYQKERS	41 ^e	789.3871	3	2365.1297	4.12
neuropeptide Y	Q6UA81		APAEVMARYSALRHHYNLITRQRY(-0.98)	1.4 × 10 ⁻⁴⁵	768.4037	4	3069.5715	4.61
neuropeptide Y	Q9XSW6	neuropeptide Y	YPSKPDNPGEDAPAEVMARYSALRHHYNLITRQRY(-0.98) ^d	1.6 × 10 ⁻⁸	854.8264	5	4269.0810	3.43
neuropeptide Y	P57774		APAEVMARYSALRHHYNLITRQRY(-0.98)	2.1 × 10 ⁻¹³	614.9217	5	3069.5715	0.24
neuropeptide Y	P57774		SKPDNPGEDAPAEVMARYSALRHHYNLITRQRY(-0.98)	3.4 × 10 ⁻⁹	802.8007	5	4008.9644	0.73
neuropeptide Y	Q6UA81		LITRQRY(-0.98)	1.2 × 10 ⁻¹²	474.7914	2	947.5665	2.00
neuropeptide Y	B6VRS8		SPETLISDLLMR	1.4 × 10 ⁻⁴⁵	687.8718	2	1373.7224	4.80
neurosecretory-proteinVGF	O15240		AQEEAEAEERLQEQELENYIEHVLLRRP	1.4 × 10 ⁻⁴⁵	742.1741	5	3705.8337	0.20
neurosecretory-proteinVGF	O15240		EQELENYIEHVLLRRP	1.4 × 10 ⁻⁴⁵	723.0429	3	2166.1070	0.00
neurosecretory-proteinVGF	O15240		Q(-17.03)EELENYIEHVLLRRP	1.4 × 10 ⁻⁴⁵	674.3549	3	2020.0376	2.54
neurosecretory-proteinVGF	O15240		RLQEQELENYIEHVLLRRP	1.4 × 10 ⁻⁴⁵	641.8451	4	2563.3508	0.19
neurotensin	Q9D3P9	neurotensin N	KIPYIL	3.6 × 10 ⁻³	373.7457	2	745.4738	4.18
neurotensin	Q9D3P9		IPYIL	2.6 × 10 ⁻⁵	618.3874	1	617.3788	2.18
neurotensin	Q9D3P9	neurotensin	Q(-17.03)LYENKPRRPYIL	1.4 × 10 ⁻⁴⁵	558.3120	3	1671.9097	2.63
neurotensin	Q9D3P9		QLYENKPRRPYIL	2.3 × 10 ⁻¹⁴	563.9854	3	1688.9363	1.16
pro MCH	P56952	neuropeptide E-1	EIGDEENSAPKFI(-0.98)	1.4 × 10 ⁻⁴⁵	724.3589	2	1446.6990	2.95
preprodynorphin	P01213		PKLKWQDNQ	1.4 × 10 ⁻⁴⁵	514.7805	2	1027.5449	1.54
preprodynorphin	P01213		YGGFLRKYP	1.4 × 10 ⁻⁴⁵	367.5362	3	1099.5813	5.00
preprodynorphin	P01213	α-neoendorphin	YGGFLRKYPK	1.3 × 10 ⁻¹⁵	410.2332	3	1227.6763	1.29
preprodynorphin	O35417	dynorphin A	YGGFLRRIRPKLKWQDNQ ^d	6.5 × 10 ⁻⁵	537.5599	4	2146.1912	8.99
preprodynorphin	O35417		YGGFLRRQF	4.0 × 10 ⁻¹²	381.8765	3	1142.5984	8.23
preprodynorphin	O35417	dynorphin B	YGGFLRRQFKVVT	5.7 × 10 ⁻¹²	393.4774	4	1569.8777	1.94
pro-enkephalin A	P01210		FAEALPSDEEGESYSKEVPEME	1.1 × 10 ⁻¹⁵	1237.0441	2	2472.0522	8.59
pro-enkephalin A	P47969		GGFMRLG	2.6 × 10 ⁻¹⁰	369.1925	2	736.3690	1.91
pro-enkephalin A	P01210		LPSDEEGESYSKEVPEM	1.4 × 10 ⁻⁴⁵	963.4210	2	1924.8247	1.40
pro-enkephalin A	P01210		LPSDEEGESYSKEVPEME	1.4 × 10 ⁻⁴⁵	1027.9431	2	2053.8674	2.02
pro-enkephalin A	P47969		RPEWWMQDYQ	1.4 × 10 ⁻⁴⁵	655.7842	2	1309.5549	0.75

Table 1. Continued

precursor	uniprot accession number	peptide name	sequence	P-value ^c or Spider score	M/Z	Z	M _r (calc)	mass error (ppm)
pro-enkephalin A	P47969		VGRPEWWMYD	1.4 × 10 ⁻⁴⁵	669.8011	2	1337.5864	0.91
pro-enkephalin A	P47969		VGRPEWWMYDQ	1.4 × 10 ⁻⁴⁵	733.8322	2	1465.6450	3.33
pro-enkephalin A	P47969	met-enkephalin-Arg-Phe	YGGFMR	1.4 × 10 ⁻⁴⁵	439.2056	2	876.3951	1.67
pro-enkephalin A	P47969	met-enkephalin-Arg-Gly-Leu	YGGFMRGL	1.4 × 10 ⁻⁴⁵	450.7238	2	899.4323	0.88
pro-enkephalin A	P47969		YGGFMRVGRPEWWMYDQ	1.4 × 10 ⁻⁴⁵	778.6969	3	2333.0623	2.83
pro-enkephalin A	P01210	peptide E	YGGFMRVGRPEWWMYDQKRYGGFL ^d	4.1 × 10 ⁻¹³	789.6378	4	3154.5168	1.63
pro-enkephalin A	P01210		AEALPSDEEGESYSKEVPEME ^b	1.4 × 10 ⁻⁴⁵	1164.0034	2	2324.9840	0.63
pro-enkephalin A	P01210		ALPSDEEGESYSKEVPEME ^b	1.4 × 10 ⁻⁴⁵	998.9421	2	1995.8618	3.98
pro-enkephalin A	P01210		ALPSDEEGESYSKEVPEME ^b	1.4 × 10 ⁻⁴⁵	1063.9615	2	2124.9043	0.46
pro-enkephalin A	P47969		GRPEWWMYDQ	1.4 × 10 ⁻⁴⁵	684.2968	2	1366.5767	1.79
pro-enkephalin A	P01210		PEWWMYDQ	1.4 × 10 ⁻⁴⁵	577.7350	2	1153.4539	1.48
pro-enkephalin A	P01210		RPEWWM	1.4 × 10 ⁻⁴⁵	452.7101	2	903.4062	0.54
pro-enkephalin A	P01210		VGRPEWWM ^b	1.4 × 10 ⁻⁴⁵	530.7548	2	1059.4961	0.92
pro-enkephalin A	P01210	leu-enkephalin	YGGFL	2.8 × 10 ⁻³	556.2761	1	555.2692	0.66
pro-enkephalin A	P47969	met-enkephalin	YGGFM	6.0 × 10 ⁻⁹	574.2348	1	573.2257	3.30
pro-enkephalin A	P01210		GGFMR	4.4 × 10 ⁻¹⁶	357.6759	2	713.3319	7.53
pro-enkephalin A	P47969		SPQLDDEAKEL ^b	28 ^c	622.8052	2	1243.5927	1.22
pro-enkephalin A	P47969		SPQLDDEAKELQ ^b	31 ^c	686.8340	2	1371.6517	1.26
pro-enkephalin A	P01210		EALPSDEEGESYSKEVPEME	1.6 × 10 ⁻¹⁵	1127.9802	2	2253.9468	0.43
prolactin	Q6A190		PVDLSKWSGLPSL ^b	14 ^c	699.8873	2	0.9927	3.09
pro-MCH	P56942		DFDMLRCMLGRVYRP	1.4 × 10 ⁻⁴⁵	624.6417	3	1870.9004	1.50
proSAAS	Q9UHG2		PPEGVIGALLR	1.4 × 10 ⁻⁴⁵	561.3392	2	1120.6604	3.05
proSAAS	Q9UHG2		PPEGVIGALLRV ^b	1.4 × 10 ⁻⁴⁵	610.8741	2	1219.7288	4.00
proSAAS	Q9QXU9		PRGEAAGAVQELARALAHLLLEAERQ ^d	1.4 × 10 ⁻⁴⁵	697.1256	4	2784.4630	3.77
proSAAS	Q9QXV0		VPPEGVIGALLRV ^b	20 ^c	660.4078	2	1318.7967	3.28
proSAAS	Q9QXU9		VPRGEAAGAVQELA	1.4 × 10 ⁻¹³	684.3682	2	1366.7205	0.98
proSAAS	Q9QXU9		AVPRGEAAGAVQELARALAHLLLEAERQ ^b	1.5 × 10 ⁻¹³	707.3912	4	2825.5256	3.49
proSAAS	Q9QXU9	GAV	AVPRGEAAGAVQELARALAHLLLEAERQ ^d	1.4 × 10 ⁻⁴⁵	739.6516	4	2954.5680	3.06
proSAAS	Q9QXU9		GEAAGAVQELARALAHLLLEAERQ ^e	1.4 × 10 ⁻⁴⁵	844.7787	3	2531.3090	2.12
proSAAS	Q60541		DADSSIEKQVALLKALYGHGQIS	1.4 × 10 ⁻⁴⁵	611.5777	4	2442.2754	2.60
proSAAS	Q60541	neuropeptideK (frag)	DADSSIEKQVALLKALYGHGQISH ^b	1.4 × 10 ⁻⁴⁵	645.8433	4	2579.3342	3.88
proSAAS	Q60541	neurokinin A	HKTDSFVGLM(-0.98)	1.4 × 10 ⁻⁴⁵	567.2936	2	1132.5698	2.59
proSAAS	Q60541		KPQQFFGLM(-0.98)	1.9 × 10 ⁻¹⁵	547.7940	2	1093.5742	0.67
proSAAS	Q60541		PKPQQFFGLM(-0.98)	5.4 × 10 ⁻¹⁵	596.3201	2	1190.6272	1.33
proSAAS	Q60541	substance P	RPKPQQFFGLM(-0.98)	1.4 × 10 ⁻⁴⁵	674.3741	2	1346.7283	3.99
proSAAS	Q60541		WSDWSDSQKEELPEPEHLLQ ^b	27 ^c	943.4431	3	2827.2977	3.45
secretogranin I	P16014		APQLDL	3.1 × 10 ⁻⁸	656.3609	1	655.3540	0.56
secretogranin I	P16014		DGVAELDQLLHY	1.4 × 10 ⁻⁴⁵	686.8423	2	1371.6670	2.31

Table 1. Continued

precursor	uniprot accession number	peptide name	sequence	P-value ^c or Spider score	M/Z	Z	M _r (calc)	mass error (ppm)
secretogranin I	A6MLG4		DQLLHY	1.4 × 10 ⁻⁴⁵	394.7017	2	787.3864	3.02
secretogranin I	P05060		FLGEGYHQVQ ^b	23 ^c	653.8080	2	1305.5987	2.08
secretogranin I	P05060		FLGEGYHQVQESRMDKA ^b	25 ^c	665.6519	3	1993.9317	1.08
secretogranin I	P16014		GVAELDQLLHY	1.4 × 10 ⁻⁴⁵	629.3293	2	1256.6401	3.21
secretogranin I	O35314		LGELEFNPYFDPLQWKSS ^b	27 ^c	977.4940	2	1952.9667	3.40
secretogranin I	O35314		LGELEFNPYFDPLQWKSS ^b	26 ^c	1021.0092	2	2039.9987	2.51
secretogranin I	O35314		LGELEFNPYFDPLQWKSSR ^b	47 ^c	733.0438	3	2196.1007	4.03
secretogranin I	O35314		PYFDPLQWKSSRFE ^b	37 ^c	900.4440	2	1798.8677	3.18
secretogranin I	P16014		Q(-17.03)YDGAELDQLLHY	1.4 × 10 ⁻⁴⁵	823.8908	2	1645.7625	2.82
secretogranin I	P16014		QYDGAELDQLLHY	1.4 × 10 ⁻⁴⁵	832.4036	2	1662.7891	2.20
secretogranin I	O35314		SGKELKDEEKENENENRKFVRL ^b	46 ^c	561.6972	5	2803.4457	1.39
secretogranin I	P16014		VAELELQLLHY	1.4 × 10 ⁻⁴⁵	600.8178	2	1199.6187	2.04
secretogranin II	Q8HZG8		ESVDFQELGKLTGPNNQ	1.4 × 10 ⁻⁴⁵	880.9470	2	1759.8743	2.91
secretogranin II	Q8HZG8		QELGKLTGPNNQ	1.4 × 10 ⁻⁴⁵	649.8410	2	1297.6626	3.76
secretogranin II	Q8HZG8		SVFQELGKLTGPNNQ	1.4 × 10 ⁻⁴⁵	816.4250	2	1630.8315	2.47
secretogranin II	P13521		FPVGLPKNDTTPNRQ ^b	37 ^c	566.6252	3	1696.8527	0.62
secretogranin II	Q8HZG8		ATLESVDFQELGKLTGPNNQ	1.4 × 10 ⁻⁴⁵	682.6909	3	2045.0432	3.82
secretogranin III	Q8WXD2		ELSAERPLNEQIA	1.4 × 10 ⁻⁴⁵	735.3858	2	1468.7522	3.32
secretogranin III	Q8WXD2		SPFDNKLNVEDVDST ^b	25 ^c	840.3933	2	1678.7687	1.98
secretogranin III	Q8WXD2		ELSAERPLNEQIAEAE	1.4 × 10 ⁻⁴⁵	899.9467	2	1797.8745	2.38
secretogranin V	Q6FHD0		SVNPLYQGQRLDNVVA	1.4 × 10 ⁻⁴⁵	886.9678	2	1771.9216	0.34
secretogranin V	Q6FHD0		SVNPLYQGQRLDN	1.4 × 10 ⁻⁴⁵	752.3831	2	1502.7478	2.60
somatostatin	Q7TSR3	somatostatin-14 (frag)	AGCKNFFWKT	1.4 × 10 ⁻⁴⁵	401.2003	3	1200.5750	3.56
somatostatin	Q7TSR3		AGCKNFFWKTFT	1.4 × 10 ⁻⁴⁵	483.9062	3	1448.6912	3.96
somatostatin	Q7TSR3	somatostatin-14	AGCKNFFWKTFTSC	1.4 × 10 ⁻⁴⁵	820.3738	2	1638.7322	0.52
somatostatin	P61279		AFSDPRLRQFLQKSLAAAAGKQLAKY	1.5 × 10 ⁻⁶	740.1661	4	2956.6243	3.72
somatostatin	Q7TSR3		FFWKTFTSC	1.4 × 10 ⁻⁴⁵	583.7717	2	1165.5266	1.89
somatostatin	Q7TSR3		NFFWKTFTSC	1.4 × 10 ⁻⁴⁵	640.7940	2	1279.5696	3.05
somatostatin	Q7TSR3		SANSNPAMAPRE	2.2 × 10 ⁻¹⁶	622.7880	2	1243.5615	0.10
thymosin β10	Q6ZWY8	thymosin β10	(Ac)ADKPDMEIASFDKAKLKKTTETQEKNTLPTKETTIEQEKRS ^d	<i>f</i>	823.2628	6	4933.5217	2.31
thymosin β4	P20065	thymosin β4	(Ac)SDKPDMAEIEKFKDKLKKTTETQEKNTLPSKETIEQKQAGES ^d	<i>f</i>	827.7576	6	4960.4847	3.47
VIP peptides	P32649	vasoactive intestinal peptide	Q(-17.03)MAVKYLSILN(-0.98)	1.4 × 10 ⁻⁴⁵	752.4232	2	1502.8279	2.60

^a PTMS: Delta mass -17.03, Pyro-Glu; Delta mass -42.01, amidation; Delta mass -0.98, acetylation. ^b Novel peptides not reported either in tree shrews or other species. ^c Peptides with P values less than 1 × 10⁻³ were selected for manual validation. ^d Further sequence validation using comparative LC-MS/MS analysis of mouse brain samples; ^e Homology search value. ^f Peptides identified by LC-MS/MS comparative analysis between tree shrew and mouse brain samples

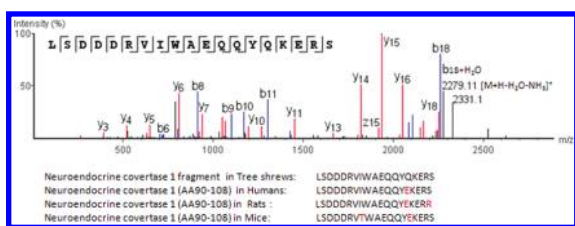


Figure 2. The neuroendocrine convertase 1 fragment LSDDDRVLWAEQQYQKERS has been sequenced via homology analysis of the MS/MS spectrum. The b- (blue) and y-ions (red) are assigned based on the delta mass difference between two fragmentation ions. De novo sequencing plus homology search allows this peptide to be unambiguously identified. The sequence of this peptide is highly conserved but not identical to its homologous peptide from mice, rats, and humans.

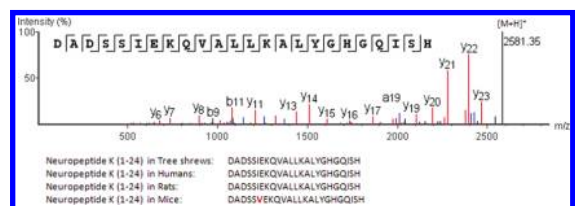


Figure 3. The neuropeptide K fragment (1–24) has been sequenced via analysis of the MS/MS spectrum. The b- (blue) and y-ions (red) are assigned based on the delta mass difference between two fragmentation ions. The sequence of this peptide is identical to that from human or rats but has one amino acid substitution in mice.

information, the sequence of neuropeptide Y was unambiguously determined. In addition to these classical neuropeptides characterized, thymosin β -4 and thymosin β -10, two peptide hormones that are not derived from prohormones, were also identified with this method (see Supplementary Figures S3 and S4, Supporting Information).

Post translational modifications (PTMs) were examined during peptide identification. PTMs have important biological influence because they can alter the functional property of neuropeptides, for example, increase the binding affinity to receptors or make peptides more resistant to enzymatic degradation.¹ In the present study, we characterized 22 peptides containing PTMs (see Table 1). Amidation is a PTM that was identified with the highest frequency in our spectral interpretation. C-terminal amidation is specific to endogenous peptides⁴⁰ and is required for the functional activation of many neuropeptides.⁴¹ In all cases where the peptide had a C-terminal amide group, the precursor contained Gly on the C-terminus; this is consistent with the known enzymatic reaction that produces C-terminal amides.⁴²

BLAST Analysis of Tree Shrew Neuropeptides

Peptides as well as neuropeptides in animals will present high identity when they have close relationship,⁴³ which can be used to understand their distance in evolution.⁹ As tree shrews are specific animals that are phylogenetically close to primates,^{23,24} it can be expected that neuropeptides in tree shrews will show overall higher homology to primates such as humans than the homologous neuropeptides in rodents such as mice and rats. To test this aspect, we determined the mean sequence identity of tree shrew peptides ($96.34 \pm 0.96\%$), rat peptides ($92.39 \pm 1.92\%$), and mice peptides ($91.15 \pm 1.98\%$) to human homologues. Figure 5 shows the difference between the sequence

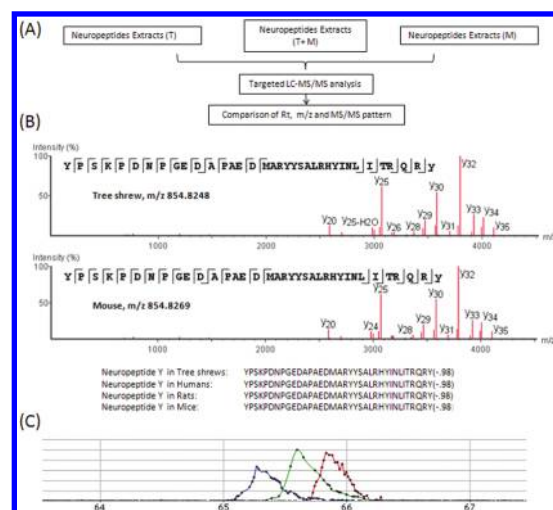


Figure 4. Applications of chromatographic and mass spectrometric information to verify the sequences of unevenly fragmented peptides. (A) Workflow to verify the sequence of neuropeptides that have low residue assignment. Targeted LC–MS/MS analysis are conducted on three types of neuropeptide samples: extracts of tree shrew brain, mice brain, and the mixtures of the two extracts (1:1, v/v). T: tree shrew, M: mouse. The retention time (Rt), parent ions, and fragmentation patterns are compared to confirm the sequence of unevenly fragmented tree shrew neuropeptides. (B) The MS/MS spectra and fragmentation ion assignments of neuropeptide Y from tree shrew and mouse samples. The C-terminal of tree shrew neuropeptide Y cannot be confirmed, although its fragmentation pattern in the N-terminal is highly similar to that of mouse neuropeptide Y. (C) The chromatographic behaviors of neuropeptide Y in extracts of tree shrew brain (blue), mice brain (red), and the mixed extracts of both (green). Neuropeptide Y has very close Rt in the three types of extracts. Importantly, the fact that no peak split of neuropeptide Y is observed in the mixed extracts indicate this peptide has an identical sequence with mouse neuropeptide Y.

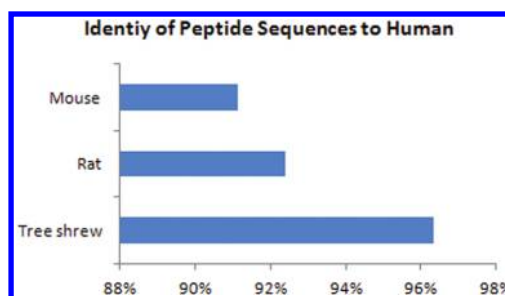


Figure 5. Sequence identity analysis of tree shrew, rat and mouse peptides against their equivalent peptides in human. The values are represented as mean \pm SEM. The comparison indicated that tree shrew peptides have significantly higher mean identify to their homologous peptides in humans than those in mice and rats (paired *t* test, $p < 0.01$).

identity of the peptides across the four species. Remarkably, the results showed that the peptides from tree shrew had significantly higher sequence identity to those from humans than from mice or rats.

DISCUSSION

Neuropeptides are important signaling molecules that regulate diverse physiological processes.⁷ They have been extensively characterized in standard laboratory animals such as rats and

mice, but to date there has been no systematic investigation in tree shrews, a small mammal that is a close relative of primates.^{23,24} We therefore focused on identifying as many neuropeptides as possible in the tree shrew, which we accomplished by developing an integrated neuropeptidomics approach. Currently, the tree shrew genome published online (see NCBI Web site) is still incomplete and the sequences are mostly unannotated. Therefore, a BLAST search on this genome can only yield a limited number of proteins. Thus, the currently available genome information for tree shrew cannot be used to predict a full list of neuropeptide precursors for database construction. This complicates data interpretation and would severely limit the neuropeptide identification rate.¹⁹ To circumvent these problems, we employed several complementary techniques to identify neuropeptides from tree shrews with high confidence. The current investigation represents a comprehensive survey of neuropeptides of the tree shrew using our integrated approach. Overall, we identified 107 peptides from 25 neuropeptide precursors. Of the identified peptides, some are classical neuropeptides, and the others may be potentially bioactive peptides. Despite the absence of a species-specific database, our results are nevertheless comparable, in terms of the total number of peptides identified, to previous studies using genetic or proteomic assay-assisted neuropeptidomics.^{13–15,44,45}

Our LC–MS/MS analysis strategy allowed highly efficient data acquisition. The analytical strategy combined data-dependent, directed and targeted LC–FT-MS/MS methods for analysis of tree shrew brain samples. In directed LC–MS/MS, the fragmentation of peptides was conducted using an inclusion list/parent mass list that was created based on analysis of the data-dependent LC–MS/MS data. The parent ions were selected according to charge states and intensities of peptides in full scans. This method allowed the removal of redundant signals and consequently increased the peptide identification efficiency.²⁶ Meanwhile, the targeted LC–MS/MS analysis allowed the screening of neuropeptides based on the prediction that a large number of tree shrew neuropeptides might be identical to their counterparts in one or more other mammals. To increase the probability of finding a peptide with identical sequence in the database, we created inclusion lists from four other species, in this study, from mice, rats, and humans. Current neuropeptidomic data were collected in these species using LC–ESI-MS/MS. The reported m/z and charge states of parent ions would therefore be the preferred states of peptides in ESI mode, which we used in our analysis. The creation of such inclusion lists is helpful to find neuropeptides, if they exist in the samples even at low concentrations. Additionally, when a low number of peptides were selected for fragmentation, some important CID parameters, such as maximum accumulation time, could be specially optimized to acquire high quality MS/MS spectra of all targeted peptides. Our targeted LC–MS/MS analysis is thus especially useful for detecting peptides with low abundance signals. Because of the high efficiency, directed/targeted LC–MS/MS are suited for very small quantities, such as those that can be obtained from brain tissue samples. Importantly, this method can also fully use the advantages of FT-MS/MS, because the application of FT-MS/MS was able to increase the confidence of sequencing,⁴⁶ but FT-MS/MS has lower sensitivity and slower scan speed than low resolution MS/MS conducted in other detectors such as ion traps.³² Using directed and targeted LC–MS/MS analysis, we were indeed able to identify a higher number of peptides, in total about 50% more, compared to our standard method of data-dependent LC–MS/MS analysis.

During evolution, the majority of neuropeptides have been highly conserved across species, with identical sequences occurring often across related species. On the basis of this principle, we conducted direct protein search and homology search to allow rapid identification of tree shrew peptides. Subsequently, we used combinational methods to verify the sequence of each neuropeptide. The validation of neuropeptidomic results remains challenging.²² Although high accuracy measurement of both parent mass and fragment mass significantly narrows the limits in database searching and increases the capability in the discovery of peptides,^{30,46} it sometimes fails to determine the full peptide sequence. This can happen especially for peptides with low information content MS/MS spectra. Prior knowledge of structural other peptide characteristics can support the validation of peptide sequence in such cases. Accordingly, we used the sequence and the retention time of mouse neuropeptides to validate the sequences of several tree shrew neuropeptides in question, for example, the long neuropeptides neuropeptide Y and GAV. Our results demonstrate that high accuracy mass spectrometric measurement plus a retention time criterion can be an efficient way of peptide identification, which is consistent with the previous observations in proteomic studies.³⁹

Tree shrews are interesting animal models used in brain research. Our study demonstrated that tree shrew peptides have a significantly higher degree of homology to equivalent sequences in humans than those in mice or rats, consistent with the close phylogenetic relationship between tree shrews and primates.^{23,24} As animal models, tree shrews thus provide a significantly closer approximation to humans in terms of biological signaling mechanisms than the standard laboratory animals such as rats and mice.

CONCLUSION

We present an integrated approach for the comprehensive characterization of neuropeptides in tree shrews. This approach exploits the high degree of conservation of neuropeptides across mammalian species and fully uses the advantages of directed and targeted LC–FT-MS/MS. Our study provides a complete overview of tree shrew neuropeptides, some of which are novel and most of which have not been functionally studied in tree shrews to date. Our finding now permits deep characterization of their function in the nervous and endocrine system of tree shrews in diverse biological studies.

ASSOCIATED CONTENT

Supporting Information

Figures S1–S4. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Phone: 0041-263008910. Fax: 0041-263009734. E-mail: xiaozhe.zhang@unifr.ch.

ACKNOWLEDGMENT

We thank Dr. Vaclav Ranc and Dr. Jennifer Lynn MILES for critical comments and revision on the manuscript preparation. This work was supported by the SNF R'Equip 316000-121308 and a EURYI award to G.R.

REFERENCES

- (1) Hokfelt, T.; Broberger, C.; Xu, Z. Q.; Sergeev, V.; Ubink, R.; Diez, M. Neuropeptides--an overview. *Neuropharmacology* **2000**, *39* (8), 1337–1356.
- (2) Holmgren, S.; Jensen, J. Evolution of vertebrate neuropeptides. *Brain Res. Bull.* **2001**, *55* (6), 723–735.
- (3) Brockmann, A.; Annangudi, S. P.; Richmond, T. A.; Ament, S. A.; Xie, F.; Southey, B. R.; Rodriguez-Zas, S. R.; Robinson, G. E.; Sweedler, J. V. Quantitative peptidomics reveal brain peptide signatures of behavior. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (7), 2383–2388.
- (4) Hokfelt, T.; Bartfai, T.; Bloom, F. Neuropeptides: opportunities for drug discovery. *Lancet Neurol.* **2003**, *2* (8), 463–472.
- (5) Insel, T. R.; Young, L. J. Neuropeptides and the evolution of social behavior. *Curr. Opin. Neurobiol.* **2000**, *10* (6), 784–789.
- (6) Hatcher, N. G.; Atkins, N., Jr.; Annangudi, S. P.; Forbes, A. J.; Kelleher, N. L.; Gillette, M. U.; Sweedler, J. V. Mass spectrometry-based discovery of circadian peptides. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (34), 12527–12532.
- (7) Aebersold, R.; Mann, M. Mass spectrometry-based proteomics. *Nature* **2003**, *422* (6928), 198–207.
- (8) Nilsson, T.; Mann, M.; Aebersold, R.; Yates, J. R., 3rd; Bairoch, A.; Bergeron, J. J. Mass spectrometry in high-throughput proteomics: ready for the big time. *Nat. Methods* **2010**, *7* (9), 681–685.
- (9) Hoyle, C. H. Neuropeptide families: evolutionary perspectives. *Regul. Pept.* **1998**, *73* (1), 1–33.
- (10) Che, F. Y.; Lim, J.; Pan, H.; Biswas, R.; Fricker, L. D. Quantitative neuropeptidomics of microwave-irradiated mouse brain and pituitary. *Mol. Cell Proteomics* **2005**, *4* (9), 1391–1405.
- (11) Dowell, J. A.; Heyden, W. V.; Li, L. Rat neuropeptidomics by LC-MS/MS and MALDI-FTMS: Enhanced dissection and extraction techniques coupled with 2D RP-RP HPLC. *J. Proteome Res.* **2006**, *5* (12), 3368–3375.
- (12) Hummon, A. B.; Richmond, T. A.; Verleyen, P.; Baggerman, G.; Huybrechts, J.; Ewing, M. A.; Vierstraete, E.; Rodriguez-Zas, S. L.; Schoofs, L.; Robinson, G. E.; Sweedler, J. V. From the genome to the proteome: uncovering peptides in the Apis brain. *Science* **2006**, *314* (5799), 647–649.
- (13) Nilsson, A.; Falth, M.; Zhang, X.; Kultima, K.; Skold, K.; Svenningsson, P.; Andren, P. E. Striatal alterations of secretogranin-1, somatostatin, prodynorphin, and cholecystokinin peptides in an experimental mouse model of Parkinson disease. *Mol. Cell Proteomics* **2009**, *8* (5), 1094–1104.
- (14) Svensson, M.; Skold, K.; Nilsson, A.; Falth, M.; Nydahl, K.; Svenningsson, P.; Andren, P. E. Neuropeptidomics: MS applied to the discovery of novel peptides from the brain. *Anal. Chem.* **2007**, *79* (1), 15–16.
- (15) Xie, F.; London, S. E.; Southey, B. R.; Annangudi, S. P.; Amare, A.; Rodriguez-Zas, S. L.; Clayton, D. F.; Sweedler, J. V. The zebra finch neuropeptidome: prediction, detection and expression. *BMC Biol.* **2010**, *8*, 28.
- (16) Zhang, X.; Pan, H.; Peng, B.; Steiner, D. F.; Pintar, J. E.; Fricker, L. D. Neuropeptidomic analysis establishes a major role for prohormone convertase-2 in neuropeptide biosynthesis. *J. Neurochem.* **2010**, *112* (5), 1168–1179.
- (17) Fricker, L. D. Analysis of mouse brain peptides using mass spectrometry-based peptidomics: implications for novel functions ranging from non-classical neuropeptides to microproteins. *Mol. Biosyst.* **2010**, *6* (8), 1355–1365.
- (18) Falth, M.; Skold, K.; Svensson, M.; Nilsson, A.; Fenyo, D.; Andren, P. E. Neuropeptidomics strategies for specific and sensitive identification of endogenous peptides. *Mol. Cell Proteomics* **2007**, *6* (7), 1188–1197.
- (19) Falth, M.; Skold, K.; Norrman, M.; Svensson, M.; Fenyo, D.; Andren, P. E. SwePep, a database designed for endogenous peptides and mass spectrometry. *Mol. Cell Proteomics* **2006**, *5* (6), 998–1005.
- (20) Falth, M.; Svensson, M.; Nilsson, A.; Skold, K.; Fenyo, D.; Andren, P. E. Validation of endogenous peptide identifications using a database of tandem mass spectra. *J. Proteome Res.* **2008**, *7* (7), 3049–3053.
- (21) Fricker, L. D.; Lim, J.; Pan, H.; Che, F. Y. Peptidomics: identification and quantification of endogenous peptides in neuroendocrine tissues. *Mass Spectrom. Rev.* **2006**, *25* (2), 327–344.
- (22) Menschaert, G.; Vandekerckhove, T. T.; Baggerman, G.; Schoofs, L.; Luyten, W.; Van Crieginge, W. Peptidomics coming of age: a review of contributions from a bioinformatics angle. *J. Proteome Res.* **2010**, *9* (5), 2051–2061.
- (23) Cao, J.; Yang, E. B.; Su, J. J.; Li, Y.; Chow, P. The tree shrews: adjuncts and alternatives to primates as models for biomedical research. *J. Med. Primatol.* **2003**, *32* (3), 123–130.
- (24) Martin, R. D. Are tree shrews primates? In *Primate Origins and Evolution*; Chapman & Hall: London, UK, 1990.
- (25) Domon, B.; Aebersold, R. Options and considerations when selecting a quantitative proteomics strategy. *Nat. Biotechnol.* **2010**, *28* (7), 710–721.
- (26) Schmidt, A.; Gehlenborg, N.; Bodenmiller, B.; Mueller, L. N.; Campbell, D.; Mueller, M.; Aebersold, R.; Domon, B. An integrated, directed mass spectrometric approach for in-depth characterization of complex peptide mixtures. *Mol. Cell Proteomics* **2008**, *7* (11), 2138–2150.
- (27) Ma, B.; Zhang, K.; Hendrie, C.; Liang, C.; Li, M.; Doherty-Kirby, A.; Lajoie, G. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **2003**, *17* (20), 2337–2342.
- (28) Nesvizhskii, A. I.; Vitek, O.; Aebersold, R. Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nat. Methods* **2007**, *4* (10), 787–797.
- (29) Svensson, M.; Boren, M.; Skold, K.; Falth, M.; Sjogren, B.; Andersson, M.; Svenningsson, P.; Andren, P. E. Heat stabilization of the tissue proteome: a new technology for improved proteomics. *J. Proteome Res.* **2009**, *8* (2), 974–981.
- (30) Taylor, S. W.; Andon, N. L.; Bilakovics, J. M.; Lowe, C.; Hanley, M. R.; Pittner, R.; Ghosh, S. S. Efficient high-throughput discovery of large peptidic hormones and biomarkers. *J. Proteome Res.* **2006**, *5* (7), 1776–1784.
- (31) Mohring, T.; Kellmann, M.; Jurgens, M.; Schrader, M. Top-down identification of endogenous peptides up to 9 kDa in cerebrospinal fluid and brain tissue by nano-electrospray quadrupole time-of-flight tandem mass spectrometry. *J. Mass Spectrom.* **2005**, *40* (2), 214–226.
- (32) Olsen, J. V.; de Godoy, L. M.; Li, G.; Macek, B.; Mortensen, P.; Pesch, R.; Makarov, A.; Lange, O.; Horning, S.; Mann, M. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell Proteomics* **2005**, *4* (12), 2010–2021.
- (33) Lee, J. E.; Atkins, N.; Hatcher, N. G.; Zamborg, L.; Gillette, M. U.; Sweedler, J. V.; Kelleher, N. L. Endogenous peptide discovery of the rat circadian clock: a focused study of the suprachiasmatic nucleus by ultrahigh performance tandem mass spectrometry. *Mol. Cell Proteomics* **2010**, *9* (2), 285–297.
- (34) Bora, A.; Annangudi, S. P.; Millet, L. J.; Rubakhin, S. S.; Forbes, A. J.; Kelleher, N. L.; Gillette, M. U.; Sweedler, J. V. Neuropeptidomics of the supraoptic rat nucleus. *J. Proteome Res.* **2008**, *7* (11), 4992–5003.
- (35) Zougman, A.; Pilch, B.; Podtelejnikov, A.; Kiehnopf, M.; Schnabel, C.; Kumar, C.; Mann, M. Integrated analysis of the cerebrospinal fluid peptidome and proteome. *J. Proteome Res.* **2008**, *7* (1), 386–399.
- (36) Sasaki, K.; Satomi, Y.; Takao, T.; Minamino, N. Snapshot peptidomics of the regulated secretory pathway. *Mol. Cell Proteomics* **2009**, *8* (7), 1638–1647.
- (37) Han, Y.; Ma, B.; Zhang, K. SPIDER: software for protein identification from sequence tags with de novo sequencing error. *J. Bioinform. Comput. Biol.* **2005**, *3* (3), 697–716.
- (38) Rholam, M.; Fahy, C. Processing of peptide and hormone precursors at the dibasic cleavage sites. *Cell. Mol. Life Sci.* **2009**, *66* (13), 2075–2091.
- (39) Strittmatter, E. F.; Ferguson, P. L.; Tang, K.; Smith, R. D. Proteome analyses using accurate mass and elution time peptide tags with capillary LC time-of-flight mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2003**, *14* (9), 980–991.

- (40) Cuttitta, F. Peptide amidation: signature of bioactivity. *Anat. Rec.* **1993**, *236* (1), 87–93, 172–173; discussion 93–95.
- (41) Eipper, B. A.; Stoffers, D. A.; Mains, R. E. The biosynthesis of neuropeptides: peptide alpha-amidation. *Annu. Rev. Neurosci.* **1992**, *15*, 57–85.
- (42) Prigge, S. T.; Mains, R. E.; Eipper, B. A.; Amzel, L. M. New insights into copper monooxygenases and peptide amidation: structure, mechanism and function. *Cell. Mol. Life Sci.* **2000**, *57* (8–9), 1236–1259.
- (43) Olivera, B. M. Conus peptides: biodiversity-based discovery and exogenomics. *J. Biol. Chem.* **2006**, *281* (42), 31173–31177.
- (44) Scholz, B.; Alm, H.; Mattsson, A.; Nilsson, A.; Kultima, K.; Savitski, M. M.; Falth, M.; Skold, K.; Brunstrom, B.; Andren, P. E.; Dencker, L. Neuropeptidomic analysis of the embryonic Japanese quail diencephalon. *BMC Dev. Biol.* **2010**, *10*, 30.
- (45) Svensson, M.; Skold, K.; Svenningsson, P.; Andren, P. E. Peptidomics-based discovery of novel neuropeptides. *J. Proteome Res.* **2003**, *2* (2), 213–219.
- (46) Lee, J. E.; Atkins, N., Jr.; Hatcher, N. G.; Zamdborg, L.; Gillette, M. U.; Sweedler, J. V.; Kelleher, N. L. Endogenous peptide discovery of the rat circadian clock: a focused study of the suprachiasmatic nucleus by ultrahigh performance tandem mass spectrometry. *Mol. Cell Proteomics* **2010**, *9* (2), 285–297.