Identification of neuropeptides in gastropod mollusks.
- Classical and brand-new approaches –

F Morishita1,*, T Takahashi2, T Watanabe2, T Uto3, K Ukena4, M Furumitsu4 and T Horiguchi5

1Program of Basic Biology, Graduate School of Integrated Sciences for Life, Hiroshima University, Higashi-Hiroshima, Hiroshima, Japan
2Suntory Foundation for Natural Sciences, Seika, Kyoto, Japan
3Department of Biological Science, Graduate School of Science, Hiroshima University, Higashi-Hiroshima, Hiroshima, Japan
4Program of Life and Environmental Studies, Graduate School of Integrated Sciences for Life, Hiroshima University, Higashi-Hiroshima, Hiroshima, Japan
5Center for Health and Environmental Risk Research, National Institute for Environmental Studies, Tsukuba, Ibaraki, Japan

*E-mail: fumi425@hiroshima-u.ac.jp

Abstract. For the understanding of the neural regulation of homeostasis, reproduction and behaviors, elucidation of structure and function of neuropeptides is inevitable. A classical technique for peptide purification is the separation of peptidic extract from nervous tissue by the high-performance liquid chromatography, which was followed by screening of fractions with biological or immunological assays. Although this approach requires relatively large amount of tissues for extraction and the screening is time-consuming, we have identified more than 15 kinds of bioactive neuropeptides in the nervous tissue of a marine snail, Thais clavigera. The purified peptides include TEP (Thais excitatory peptide)-1/-2, FRFamide, WWamide and others. The molecular cloning of precursor for each neuropeptide demonstrated that structurally related peptides are aligned in tandem on most of the precursor proteins. To identify the peptides on the precursors, we conducted the de novo sequencing of peptides with the nanoLC-Orbitrap-MS/MS analysis. It identified most of the neuropeptides found on the precursor proteins in the extract from 50 of Thais ganglia. Thus, this technique is suitable for the comprehensive identification of peptides from relatively small amount of tissues. However, several peptides that had been identified by the classical technique, such as APGWamide and leucokinin, were not identified, suggesting the limitation of this brand-new technique.

1. Introduction
The nervous system and endocrine systems are essential to the regulation of physiological activities of various tissues, homeostasis such as osmoregulation, reproduction and behaviors of individuals such as locomotion and feeding. In these systems, various kinds of chemical substances are working as the signal
messenger. For example, acetylcholine, biogenic amines and amino acids are major neurotransmitters of the nervous system, while steroids are major hormones in the endocrine system. On the other hand, peptides work in both systems as signal molecules, namely, neuropeptides and peptide hormones.

The structure of neuropeptides/peptide hormones is diverse. In fact, enormous numbers of bioactive peptides with different structures have been identified from animals that evolved the nervous system. For instance, oxytocin/vasopressin super-family peptides have been found in various animals in both of vertebrate and invertebrates [1]. Amino acid sequences of those peptides are different from mammalian peptide, and their functions are not always osmoregulation. In this context, for the total understanding the functions of regulatory peptides in nervous and endocrine systems, it is important to identify neuropeptides or peptide hormone in the target animals, and then, clarify their functions.

Since the first discovery of a neuropeptide, FMRFamide, from a clam, *Macrocallista nimbosa*, in 1977 [2], identification of neuropeptides through the combination of HPLC-fractionation and screening of fractions by biological and immunological assays was very popular in mollusks, until early 2000’s. Considerable numbers of peptides have been identified in that period, and physiological functions of the peptides have been well investigated [3,4].

Some selected gastropods such as a sea hare, *Aplysia*, sea slug, *Tritonia*, land snails, *Helix* and *Achatina* and a freshwater snail, *Lymanea*, are suitable for investigation of the physiological functions of neuropeptides, because these animals show us stereo-typed behaviours regulated by simpler neural circuits consisted of identifiable large neurons. We, therefore, started peptide isolation project in Japanese sea hare, *Aplysia kurodai*, in late 90’s and reported several papers [5-8].

On the other hand, information of neuropeptides was very limited in neogastropods, a major class of gastropods. Most of the neogastropods adopt gonochorism for reproduction, in which male and female appear in different snails, while *Aplysia*, *Helix* and *Lymnaea* adopt hermaphrodite, in which male and female appear in the same animal. Neurons in the central nervous system (CNS) of neogastropods are relatively small, and the connection of the major ganglia still maintain the ancestral form, namely, the eight-figure torsion. These facts implied us that regulatory neuropeptide system in neogastropods is somewhat different from those in *Aplysia*, *Helix* and *Lymnaea*. Moreover, neogastropods are sensitive to endocrine-disrupting agent, such as organotin compounds. Therefore, we started peptide identification through the combination of HPLC-fractionation and screening in a marine snail, *T. clavigera*, in early 2000’s, and reported several papers [9-11]. One drawback of the peptide identification by this approach is that it requires large amount of animal-tissues as the starting material. This is because a part of extracted materials is lost during chromatography and screening, and the amino-acid sequencing of purified material by the Edman’s degradation is not so sensitive. Thus, collection of the animals may have impact to the ecological balance in nature.

Recent progress in techniques for mass spectrometry, together with the accumulated genetic information, enabled us to determine the amino acid sequence of purified peptides with high sensitivity. Especially, the nanoLC-Orbitrap MS/MS system is powerful tool for the determination of amino acid sequence of peptides. Such techniques reduce the impact to the ecology very much, because we can start peptide identification from tissues collected from smaller numbers of animals. Now, in the post-genome era, mass spectrometry supported by genetic information is the major approach to identification of peptides. Of course, none of the techniques are all-mighty. Even this brand-new approach has some drawbacks.

In this article, we will overview our work on peptide isolation from *T. clavigera* with a special reference to the technologies used in our studies. It includes the comparison of the classical and the brand-new approaches toward the identification of novel peptides, and show you the strong and weak points for respective approaches. Once neuropeptides are identified, their physiological functions are important issues to be clarified. We will give perspectives how neuropeptides are involved in regulation of reproduction in gastropod mollusks, because reproduction is essential event for the survival of species.
2. Materials and methods

2.1. Animals
The marine snail, *Thais clavigera*, was collected on the seashores around Ibaraki, Japan, and Mukaishima, Japan. Collected animals were maintained in a tank filled with artificial seawater (ASW), and fed with a clam, *Ruditapes sp*. The sea hare, *Aplysia kurodai*, was collected at Etajima and Mukaishima in Hiroshima, Ushimado in Okayama, Asamushi in Aomori and Oki in Shimane, respectively. Collected animals were maintained in a tank filled with ASW and fed with boiled sea weed. In both animals, temperature of ASW was maintained 15 °C.

2.2. Extraction and separation of peptides
Procedures for the extraction of peptides from *Thais* tissue was described previously [10]. In brief, five hundred of dissected tissues containing the CNS were powdered in liquid nitrogen (LN2), and boiled in 2% acetic acid solution for more than 10 min. Boiled tissue was homogenized and centrifuged to obtain the supernatant. The supernatant was condensed with a rotary evaporator, and it was applied on the Sep-Pak C18 column (Waters Corp., Milford, MA). Retained materials were eluted with 70% methanol (RM70).

2.3. HPLC-separation and screening
After removing methanol with a vacuum centrifuge, RM70 was separated by high-performance liquid chromatography (HPLC). When the sample was separated by reversed phase-HPLC, retained materials were eluted with a linear gradient of acetonitrile. When the sample was separated by cation-exchange HPLC, retained materials were eluted with a linear gradient of NaCl in the presence of 10 mM phosphate buffer (pH6.2-6.5). Eluted materials were fractionated in every 2-min, and an aliquot of each fraction (1/500-1/1000) were subjected to biological or immunological screening.

2.4. Screening
Protocol for biological screening was described previously [10]. In brief, dissected tissue (esophagus of *T. clavigera* or large hermaphroditic duct (LHD) of *A. kurodai*) was mounted in the test chamber. One end was tied to the bottom of the chamber, while the other end was tied to the force-transducer. When the tissue was contracted, strain gauge in the force-transducer bend downward, which modifies the electric current applied to the gauge. The change in the current was amplified and recorded on the chart recorder as the trace of motilities of the tissues.

When the fraction after the reversed phase HPLC were tested, an aliquot of each fraction (1/500-1/1000) was dried out, then dissolved in 10 μL of ASW. It was added to the external bathing media with a pipette. When the fraction after the cation-exchange HPLC were tested, an aliquot of each fraction was added to the external bathing media with a pipette, directly.

Immunological screening of HPLC-fractions was conducted by the dot-blot assay on the nitrocellulose paper [12]. In brief, bovine serum albumin (BSA) was immobilized on the nitrocellulose paper by baking (105 °C for 60 min.). An aliquot of each fraction obtained by reversed phase HPLC was dried out, then dissolved in 1 μL of DW. It was blotted on the BSA-coated paper. When the fraction of cation-exchange column was tested, 1 μL of each fraction was blotted on the on the BSA-coated paper, directly. Materials in the fraction was connected to BSA by formalin vapor (60 °C for 60 min.). After washing the paper with 1% sodium metabisulphate, it was incubated with primary antibody to previously known neuropeptides. We used anti-NdWFamide antibody [13], anti-GLWamide antibody [14] and anti-LEP (leech excitatory peptide) antibody [15] and anti-PRQFVamide antibody [8] for screening.

2.5. Analysis of structure of the purified materials
For the amino acid sequencing, a part of purified material (OD_{220} = 0.03 equivalent or more) was blotted on the polybrene-coated glass fiber disc, and it was analyzed with the automated gas-phase peptide
sequencer (PPSQ-10, Shimadzu Corp., Kyoto, Japan) based on the Edman’s degradation. Molecular mass of the purified material was determined by nanoflow electrospray ionization time-of-flight mass spectrometer, Q-Tof (Micromass UK Ltd., Manchester, UK).

2.6. Molecular cloning of precursor
Procedures for molecular cloning of precursor was described previously [16, 11]. In brief, mRNA was extraction from a hundred of *Thais ganglia* with oligo-dT cellulose column supplied with the Micro-FastTrack 2.0 (Invitrogen Co., Carlsbad, CA). The cDNA was synthesized by reverse transcription, and then, a unidirectional Uni-Zap λ-phage library was constructed with the ZAP-cDNA Synthesis Kit and Zap-cDNA Gigapack III Gold Cloning kit (Stratagene, La Jolla, CA). In brief, cDNA was incorporated into a λ-phage vector (Uni-Zap XR). After packaging in phage sheath, it was transfected and multiplied in *E. coli*, and cDNA library in Uni-Zap XR vector was extracted. For the amplification of precursor cDNA by PCR, degenerate reverse primer was made, according to the amino-acid sequence of the peptide. PCR amplicon was purified and cloned by the standard molecular biological techniques. Cloned cDNA was sequenced by 3130xl Genetic Analyzer (Applied Biosystems Inc., Waltham, MA).

2.7. Mass spectrometry with Orbitrap
For the mass spectrometry with LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific K.K., Tokyo, Japan), peptides were extracted from fifty ganglia of *T. clavigera* with the protocol described in the section 2.2. Final volume of the extract was around 100 μL. One-microliter of the extract was mixed with 14 μL of 0.1% formic acid, and then, 1 μL of the diluted sample was used for the MS/MS analysis.

For the nano-scale liquid chromatography (nanoLC), sample was condensed on a Trap column (Accliam PepMap100 C18, I.D 300 μm x 5 mm, Thermo Fisher Scientific K.K., Tokyo, Japan), then separated by a separation column (Nikkyo nano HPLC capillary column, 3 μm C18, I.D. 75 μm x 120 mm, Nikkyo Technos Co. Ltd., Tokyo, Japan) with a linear gradient of acetonitrile (4-35% over 100 min, 0.3 mL/min). Eluted materials were directly ionized by the nano-scale electrospray at source voltage between 1.5 and 2.5 kV.

Mass spectrum for the precursor ion was obtained by Fourier transform mass spectrometry mode with scan range from 200-1,500 (m/z) and resolution at 30,000. Then, the parent ion was fragmented by the collision to helium gas with collision energy at 35 and isolation width at 2.0 (m/z). The conditions for mass measurement of fragment ions was the same to that for the parent ion. Obtained mass data were analyzed by an application, PEAKS X (Bioinformatics Solutions Inc., Waterloo, Canada), for de novo sequencing.

2.8. Peptide synthesis
Peptides were synthesized by microwave-assisted solid-phase peptide synthesis using an automated peptide synthesizer (Syro Wave; Biotage Japan, Tokyo, Japan) based on the Fmoc-strategy, as described previously [10].

3. Results and discussion

3.1. Identification of peptides by classical approach
The isolation of bioactive peptides begins from the extraction. Most important point at this step is inactivation of protease and peptidase that digest peptides. Our standard procedures are quick freezing of dissected tissues in LN2, which was followed by powdering the frozen tissues in LN2. The powdered tissue in small portions was put into boiling 2% acetic acid, and it was kept boiling for more than 10 min. Boiled tissues were homogenized by the Waring blender or rotor-stator homogenizer, and supernatant was collected by centrifugation. Alternative procedures that we adopt are, freezing tissues by clamping two aluminums blocks that have been chilled in dry ice, and then, frozen tissues were kept in 60% acetone (final concentration) acidified with HCl for overnight at -20 °C. Then, it was homogenized and centrifuged as before.
Supernatant was condensed by the rotary evaporator or vacuum centrifuge, and it was applied to C18 column such as Sep-Pak C18 Vac column (Waters Corporation, Milford, MA). After the loading of tissue extract on the column, it was washed with 5% methanol to remove low-molecular weight neurotransmitters such as acetylcholine and biogenic amines, as well as salts. Then, retained materials were eluted with 70% methanol, and methanol was removed by vacuum centrifuge before applying it to HPLC-separation. We refer this extract as RM70 (retained materials 70).

Now, we came to the first check point. If RM70 doesn’t have any bioactivity (or immunoreactivity), no hope to identify bioactive peptides from the RM70. We used *Thais* esophagus for bioassay, because esophagus is sensitive to various neurotransmitters and neuropeptides. On the other hand, we aimed to isolate any neuropeptides that mediates reproduction of *T. clavigera*. For this purpose, bioassay using reproduction-associated organ is effective. However, reproduction-associated organs in this gastropod is small, and hard to make stable preparation for screening. So, we used the large hermaphroditic duct (LHD, complex of vas deference and oviduct) of *Aplysia* for the replacement. Fortunately, an aliquot of RM70 induced marked excitatory responses to *Thais* esophagus and *Aplysia* LHD, suggesting that RM70 contains any bioactive peptides. So, we started the HPLC-separation and screening cycle.

For the separation of RM70 by HPLC, first choice is reversed phase column that separates materials by their hydrophobicity, because this column retains various kinds of peptides. Cation-exchange column is also effective, because this column has bigger binding capacity than reversed-phase column at the same size. However, some peptides may have negative charge depending on pH of buffer for chromatography.

Figure 1 represents an example for separation of RM70 with reversed phase HPLC and a part of screening of bioactive fractions using Aplysia LHD. In this chromatography, RM70 was applied on the Inertsil ODS-80A column (GL Science, Tokyo, Japan), and retained materials were eluted by a linear gradient of acetonitrile (0-60% acetonitrile over 120 min, 0.5 ml/min), which were fractionated in every 2-min. An aliquot (1/1000) of each fraction was dried out once, and then, dissolved in ASW. It was applied to the isolated Aplysia LHD to see the bioactivity. For example, fractions 18-20 had little activity, while fractions 23, 26 and 27 had marked excitatory activity. The active fractions were grouped and subjected to the next separation-screening cycle. This cycle was repeated until bioactive material
was purified in a single absorbance peak on reversed phase column. Basically, reversed phase column and cation-exchange column are used alternatively, but final purification was conducted with reversed phase column, because salts in the buffer for cation-exchange HPLC may disturb the following analysis.

Another screening is immunological assay using antibody to previously known peptide. This screening depends on the similarity in the structure of peptides. We adopted dot-blot assay on the nitrocellulose paper, because this technique can handle many samples in parallel. However, low binding of oligo-peptides to nitrocellulose paper was anticipated. Therefore, nitrocellulose paper was coated with bovine serum albumin (BSA) by baking. Then, an aliquot of each fraction was blotted on the BSA-coated paper, and materials were immobilized on BSA by exposing the blotted paper to formalin vapor. Among the several poly-clonal antibodies tested, antibodies to GLWamide in *Hydra*, LEP in Leech and NdWFamide in *Aplysia* worked well to isolate novel peptides (Figure 2).

In the classical approach, purified peptides were analyzed by the peptide sequencer based on the Edman’s degradation. Edman’s degradation consisted of the two-step chemical reactions, namely, the labeling of N-terminal amide-moiety with phenyl isothiocyanate and cleavage of the labeled amino acid from the peptide with trifluoro acetic acid. Released amino acid was converted to stable phenyl thiohydantoin-amino acid, which is identified by chromatography. Since this labeling and cleavage cycle could be repeated on the same peptide, amino-acid sequence of the peptide is determined from N-terminus.

Peptide sequencing by Edman’s degradation has several drawbacks, if it was compared to the recent peptide sequencing by mass spectrometry. For instance, peptides with N-terminal modification, such as pyroglutamation and acetylation, can’t be analyzed, because phenyl isothiocyanate does not bind to such peptides. For the identification of cysteine residue, alkylation of peptide before analysis is necessary. Moreover, purified material at picomole-order is required for the analysis. Nevertheless, it was attractive that amino-acid residues in a peptide, including leucine and isoleucine, could be determined one by one with the peptide sequencer such as PPSQ-10 by Shimadzu Corp. or Model 447A by Applied Biosystems Inc.

Here, we exemplify the determination of structure of purified material numbered as TF37-2. Peptide sequencer determined the amino acid sequence of TF37-2 as Gly-Ser-Leu-Phe-Arg-Phe. The mass spectrometer, Q-Tof, determined the molecular mass of TF37-2 as 726.23 ([m+H]+), which is 1-mass smaller than the molecular mass predicted from the sequence. The result suggests that TF37-2 has C-terminal amide, because the molecular mass of -CONH₂ is 1-mass smaller than that of -COOH. Besides the C-terminal amide, mass spectrometry suggests us some modification of peptides, including disulfide-bond between a pair of cysteine residues and oxidation of methionine residue.

Once the structure of the purified peptide is deduced, the final step is confirmation of the analysis. Peptide was made by the analysis (GSLFRF-NH₂, this case), and elution time on HPLC was compared...
to purified substance. In fact, elution time of GSLFRF-NH$_2$ was exactly the same to that of TF37-2, and mixture of the two eluted in a single absorbance peak on both of reverse-phase and cation-exchange columns. Thus, we determined that structure of TF37-2 is GSLFRF-NH$_2$, with no ambiguity. Table 1 lists up *Thais* peptides identified by aforementioned approach.

Table 1. Peptides purified from *Thais clavigera* by immunological and biological assays.

| Name                  | Screening | ACN(a)(%) | Structure                        |
|-----------------------|-----------|-----------|---------------------------------|
| TEPb-1                | Dot-blot  | 29.2      | KCSDKWAIHACWGGNa                 |
| TEPb-2                |           | 31.4      | KACYGKWAMHCWGGNa                |
| FXXFamide             | Anti-GLWamide | 23.0 | PIAYGQFGKwa                    |
| FRFamide-1            | Anti-NdWFamide | 23.0 | SSLFRF$^a$                     |
| FRFamide-2            |           | 23.0      | GSLFEF$^a$                      |
| PentxFVamide          | Anti-PRQFVamide | 32.0 | LRDFV$^a$                      |
| TDP$^c$               | Bioassay  |           | Aplysia LHD                     |
| APGWamide             |           | 16.0      | GFRVNAAGRVAHGY$^a$              |
| FMRFamide             |           | 12.5      | APGW$^a$                        |
| FLRFamide             |           | 9.0       | APGW$^a$                        |
| Tachykinin-RP         |           | 22.0      | FHPMAGFGR$^a$                   |
| Leucokinin-RP         |           | 22.0      | APFSVWAa                        |
| Myomodulin-RP         |           | 20.0      | WMGLRL$^a$                      |
| Enterin-RP-1          | Inh$^b$   | 19.0      | VPSFGHRFV$^a$                   |
| Enterin-RP-2          | Inh$^b$   | 20.5      | RSFFTRFV$^a$                    |
| WWamide-RP            | Inh$^b$   | 18.5      | WKSMSVWA$^a$                    |

$^a$ concentration of acetonitrile (CAN) at the final purification by a reversed phase HPLC

$^b$ leech excitatory peptide

$^c$ *Thais* excitatory peptide

$^d$ This peptide has an intra-molecular disulfide bond between the two cysteine (C) residues

$^e$ Tetradecapeptide

$^f$ large hermaphroditic duct

$^g$ Excitatory action

$^h$ Inhibitory action

3.2. Identification of peptides by molecular cloning

Neuropeptides are made on the ribosome in a part of larger precursor protein. It is frequently observed that structurally related peptides are found on a single precursor. For instance, we isolated 10 kinds of enterins from *Aplysia* [6]. However, when we cloned the enterin precursor, we found that enterin is a family peptide consisted of 21 kinds of structurally related peptides. Another instance is precursor of ELH (egg-laying hormone). ELH is a peptide hormone that induce egg-laying behavior to matured *Aplysia*. ELH precursor contains other neuropeptides such as $\alpha$-BCP (bag cell peptide) that trigger discharge of the bag cell in abdominal ganglion [17]. Thus, a single precursor encodes different peptides with different function, which co-ordinate to induce proper behavior. These instances show us the importance for elucidation of neuropeptide precursor.

One of the practical ways to elucidate neuropeptide precursor is molecular biological approach. So, we prepared cDNA-library derived from *Thais* ganglia [11]. Using this library as template, a part of neuropeptide precursor cDNA was amplified by PCR (polymerase-chain reaction) with degenerate reverse primer. PCR amplicons were cloned and sequenced by the standard procedures for molecular biology. Once the cDNA encoding N-terminal region of precursor was cloned, gene specific forward
primer was made in the 5'-non-coding region to amplify the open-reading frame for the precursor protein.

**Figure 3.** Scale models showing the distribution of neuropeptides on respective precursor. Shaded column on each precursor represents location of neuropeptides. Note that alphabets and numbers on each model matches to those of each peptide indicated beneath each model.

S: signal peptide.

Of 17 kinds of neuropeptides purified from *T. clavigera*, we have identified precursors for 16 kinds of neuropeptides, currently. Figure 3 shows a scale model of two neuropeptide precursors, namely *Thais* leucokinin precursor and *Thais* leucokinin precursor. *Thais* leucokinin is a heptapeptide having structural similarity to leucokinin in insects [18]. However, cloned precursor consisted of as many as 931 amino acids. Besides the 14 copies of purified *Thais* leucokinin, 9 kinds of different leucokinins were found on the precursor. Moreover, two non-leucokinin type peptides having C-terminal amide were found on the precursor, although the biological function of them have not examined yet (Figure 3). Another example is *Thais* WWamide precursor. Purified WWamide is a heptapeptide having a pair of Trp residues on both of N- and C-termini. However, 17 kinds of different WWamide related peptides were found on the cloned precursor. These results demonstrated that neuropeptides on precursor proteins are diverse in this neogastropod.

### 3.3. Identification of peptides by mass spectrometry

Molecular cloning of neuropeptide precursors suggested the diversity of the neuropeptides in this neogastropod. Now, the question is, whether peptides found on the precursors are processed to be mature peptides? Because we can predict what kinds of peptides will be generated from each precursor, mass-spectrometry-based identification of peptides is effective. Therefore, we tried to identify peptides in the extract from Thais ganglia using the nanoLC-LTQ Orbitrap mass spectrometer. Because the analysis by Orbitrap mass spectrometer is much more sensitive than that by the peptide sequencer, we prepared RM70 from 50 of Thais ganglia, which is 1/10 of tissues used for the classical approach.

In nanoLC system, peptides in the RM70 was accumulated on the Trap column, and then, separated by C18 capillary column. Eluted materials were ionized by the nano-scale electrospray ionization, and the ions were introduced to the Orbitrap for the measurement of total mass of each ion (parent ion). Parent ion was fragmented by collision to helium gas, and molecular mass of each fragment ion was determined again by the Orbitrap. For the amino acid sequencing, an application, PEAKS X, aligned the mass data of each fragment ion generated from the same parent ion, and calculated the differences in molecular mass between the fragment ions to see which difference matches to the molecular mass of
amino acids. PEAKS X finally showed us the best match as the amino-acid sequence of the parent ion (Figure 4).

The analysis by nanoLC-Orbitrap MS/MS identified many peptides found on the precursors from the small amount of tissue extract. However, numbers of peptides identified were different among the precursors. For instance, of the 14 kinds of Thais enterins found on the precursor, 12 kinds of Thais enterins were identified by the analysis (Table 2). The analysis also identified most of the peptides on the precursors of ThFXXFamide, FRFamide, FMRFamide and pentaFVamide. By contrast, less than half of the peptides on the leucokinin (Table 2) and WWamide precursors were identified. Moreover, TEP-1, TEP-2 and APGWamide, which were purified by the classical approach was not identified by this analysis. Probably, difference in the efficiency of ionization among peptides is one of the reasons for this result. Moreover, several novel peptides such as Thais pedal peptide, FHWamide, FVRLamide were also identified by de novo sequencing. Our result demonstrated that nanoLC-LTQ Orbitrap mass spectrometer is a suitable equipment for the identification of peptides in small amount of tissues in marine snail.

In nanoLC system, peptides in the RM70 was accumulated on the Trap column, and then, separated by C18 capillary column. Eluted materials were ionized by the nano-scale electrospray ionization, and the ions were introduced to the Orbitrap for the measurement of total mass of each ion (parent ion). Parent ion was fragmented by collision to helium gas, and molecular mass of each fragment ion was determined again by the Orbitrap. For the amino acid sequencing, an application, PEAKS X, aligned the mass data of each fragment ion generated from the same parent ion, and calculated the differences in molecular mass between the fragment ions to see which difference matches to the molecular mass of amino acids. PEAKS X finally showed us the best match as the amino-acid sequence of the parent ion (Figure 4).

The analysis by nanoLC-Orbitrap MS/MS identified many peptides found on the precursors from the small amount of tissue extract. However, numbers of peptides identified were different among the precursors. For instance, of the 14 kinds of Thais enterins found on the precursor, 12 kinds of Thais enterins were identified by the analysis (Table 2). The analysis also identified most of the peptides on the precursors of ThFXXFamide, FRFamide, FMRFamide and pentaFVamide. By contrast, less than half of the peptides on the leucokinin (Table 2) and WWamide precursors were identified. Moreover, TEP-1, TEP-2 and APGWamide, which were purified by the classical approach was not identified by this analysis. Probably, difference in the efficiency of ionization among peptides is one of the reasons for this result. Moreover, several novel peptides such as Thais pedal peptide, FHWamide, FVRLamide were also identified by de novo sequencing. Our result demonstrated that nanoLC-LTQ Orbitrap mass spectrometer is a suitable equipment for the identification of peptides in small amount of tissues in marine snail.
Table 2. Peptides found on the cloned precursor.

| Thais enterin precursor | Thais leucokinin precursor |
|-------------------------|---------------------------|
| APTFGHLFVa *            | APSFVWAa *                |
| QPAFSYRFVa              | AKFSVWAa *                |
| VPSFGRHFVa *            | AAFSVWAa *                |
| APTDGHRFVa *            | GPFSUWAa *                |
| APSFGRHFVa *            | NLPARAFSVWAa              |
| pEPSFGRHFVa *           | AAFNVWAa                  |
| LPTFGHHFVa *            | AAFSSWAa                  |
| QPSFDHFSFVa             | TPFSAWAa                  |
| VPTFGHHFVa *            | EAFSAWAa                  |
| KPSFGHMFVa *            | RFSSWAa *                 |
| IPSFGRHFVa *            |                           |
| VPAFGRFVa *             | SANTRTAASARPa             |
| VPSFGRMFVa *            | SAMESPQFWPETa *           |
| VPQFGRHFVa *            |                           |

Note that pE represents N-terminal pyroglutamination, and Va, Aa, Pa and Ta represents C-terminal amidation of respective amino acids.

The peptides identified by Orbitrap MS/MS analysis are marked by asterisk, while those purified by HPLC and screening are boxed.

However, there are some drawbacks in mass-spectrometry-based peptide identification. For instance, this analysis may overlook some important neuropeptides, because of the difference in efficiency of ionization of the peptides. Any supportive information such as genetic information may necessary to finally determine amino-acid sequence of the peptide, because, for instance, it is hard to discriminate leucine residue and isoleucine residue by mass spectrometry alone. In the classical approach to the peptide isolation, we could obtain real peptides in our hand, which are used for confirmation of structure, or elucidation of bioactivity. By contrast, what we can obtain by mass spectrometry is just data, but not real peptide. So, confirmation of analysis is difficult, and peptide synthesis is necessary to find bioactivities of identified peptides.

As mentioned above, nanoLC-LTQ Orbitrap MS/MS analysis is a powerful tool for the comprehensive identification of peptides from relatively small amount of tissues. However, identified peptides are not always bioactive. To identify bioactive peptides from tissue extract, screening by bioassay is still attractive. Therefore, one of the practical approaches is purification of bioactive peptides by the HPLC-fractionation and screening of fractions by bioassay, then, determination of amino acid sequence by nanoLC-LTQ Orbitrap MS/MS. Because analysis by mass spectrometry is highly sensitive, it will be possible to determine structure of bioactive peptides from smaller amount of tissues.

For screening of bioactive peptides, digestive tract such as esophagus is an excellent organ, because it is easy to make preparation, durable for repeated application of HPLC-fractions, and sensitive to various kinds of peptides. In fact, using a part of digestive tract, so many kinds of peptides have been identified in mollusks and insects. On the other hand, innovation in the bioassay system is an important issue for peptide identification. For instance, although this is not a story on the gastropod, but on a jellyfish, novel peptides that promote oocyte maturation was identified by observing the induction of germinal vesicle breakdown in oocyte [19]. A unique assay system leads us to identification of unique peptides.

4. Perspectives
Reproduction is one of the major purposes of living organism. Disturbance of reproduction directly effects on the population of species, which may, in turn, effects on the ecological balance. In gastropod
mollusks, neuropeptides, rather than steroids, play important roles in regulation of reproduction. Here, we would like to give perspectives in relation between neuropeptide system and reproduction.

4.1. Reproduction and environmental pollution
One of the serious factors that threaten the ecological balance is environmental pollution. Since 1970’s, people were aware of worldwide decline in populations of marine snails (Prosobranch gastropods). In those animals, masculinization of females such as the secondary formation of penis and vas deference, or spermatogenesis in ovary were observed (imposex), which disturb ovulation and egg-release in female. Soon after, it was found that the trace amounts of organotin compounds such as triphenyltin (TPhT) and tributyltin (TBT) in seawater induced the imposex [7]. Organotins were released from antifouling paint on ships and fishing nets.

Several hypotheses on the molecular mechanism of imposex were proposed. Because organotins inhibit aromatese and acyl CoA-steroid acyltransferase, the enzymes catalyzing the steroid metabolism in vertebrates, one of the hypotheses is that organotins disturbs androgen-estrogen balance toward androgen predominant in female gastropods. However, hormonal action of steroid is still in controversial in gastropods [22].

Another hypothesis for imposex is malfunctioning of regulatory neuropeptide system. Considering the fact that neuropeptide plays crucial roles in both of nervous and endocrine systems in gastropods, this hypothesis is attractive. For instance, it was reported that the pedal ganglion releases peptides that promote penis formation (penis morphogenic factor, PMF) in male gastropods. Secretion of PMF is normally inhibited in females. However, neurotoxicity of organotins cancels the inhibition, which results in the penis formation in females. This story is simple, but the entity of PMF is still unclear.

It was also reported that injection of a molluscan neuropeptide, APGWamide, into female snails promotes penis elongation [19]. The result suggests that neuropeptide mediates penis formation by organotins. However, penis-elongation by APGWamide was very small. Thus, we have few evidences that support this hypothesis. More fundamental information on the structure and function of neuropeptides in marine snails is necessary to verify this hypothesis.

Now, it was demonstrated that target molecule of organotin compounds is one of nuclear receptors, the retinoid X receptor (RXR) in *T. clavigera* [8]. In fact, an endogenous RXR ligand in mammals, such as 9-cis-retinol, induces imposex in *T. clavigera*. In this hypothesis, binding of organotins to RXR promote the sequential gene expression that leads penis formation. Expression level of RXR gene is high in penis of normal male snails, whereas it is low in penis-forming area in normal female snails. However, expression of RXR gene is elevated in the penis forming area of imposex-exhibiting female snails, which were exposed to organotins, such as TPhT and TBT. Moreover, expression level of RXR gene correlates to the penis length of imposexed females.

As we already mentioned, imposex includes two different symptoms, namely, the secondary formation of penis and/or seminal duct in females and the spermatogenesis in ovary. In abalone, it was reported that spermatogenesis in ovary occurs without penis and/or seminal duct formation, suggesting that two symptoms are independent. Considering the facts that the ganglia in the central nervous system accumulate high concentrations of organotins, and expression of RXR gene is relatively high in the ganglion, possible scheme is that organotin promote the expression of neuropeptides through the activation of RXR in female ganglia. The neuropeptide, in turn, promotes the gonadal differentiation toward testis. This is an interesting hypothesis to be examined, because mechanism that regulate the expression of neuropeptide precursor gene has not been fully understood, yet.

Besides the organotin compounds, other pollutants may modify the expression of neuropeptide genes. Therefore, the attempt to use expression of particular neuropeptide precursor gene for the biomarker of the environmental pollution is an interesting challenge.

4.2. The master neuropeptide of reproduction in mollusks
Assuming that neuropeptides play essential roles in regulation of reproduction in mollusks, an interesting question is that, which one is the most important? If the master peptide is malfunctioned by
any environmental factors, it is anticipated that serious damage on reproduction of the animal leads the reduction of the population and disturbance of the ecological balance.

In mammal, steroid hormones such as testosterone and estradiol, play essential roles in sex-differentiation in embryo or gonadal maturation in adolescence. However, GnRH is the key peptide for reproduction, because it regulates the ovulation cycle through the regulation of FSH and LH-secretion. In this context, no GnRH, no reproduction.

In mollusk, the involvement of GnRH in reproduction is controversial. GnRH promotes proliferation of gonadal cells in the scallop, whereas the peptide regulates locomotion in *Aphysia*. In octopus, GnRH is multi-functional. Probably, GnRH is a multi-functional neuropeptide in invertebrates. In mollusk, physiological function of GnRH seems to be different among orders in mollusk. By contrast, GnRH is specialized for the regulation of reproduction in vertebrates. A question is raised here. Are there any master neuropeptide that trigger the cascades of processes that induce the reproduction in mollusk?

In *T. clavigera*, physiological function of GnRH is unclear, because GnRH has not been identified yet. For the purification of reproduction associated peptides from *T. clavigera*, we adopted *Aphysia* LHD for bioassay. Of the peptides identified by this assay system, *Thais* tetradecapeptide and APGWamide had inhibitory action on isolated penis of the marine snail. Some other peptides such as FLRFamide, myomodulin and WWamide, which were obtained by bioassay on esophagus or dot-blot assay, also modulated the motility of *Thais* penis. The result demonstrated that multiple neuropeptides are involved in the regulation of motility in a single organ, which suggests that regulation of reproduction by neuropeptide system is rather complex in gastropods.

In fact, recent studies with high throughput determination of neuropeptides and characterization of identified peptides demonstrated that multiple peptides work in concert or in competition to regulate the reproduction in mollusk [20-22]. For instance, 28 of bioactive neuropeptides were identified transcriptome and peptidome analysis in visceral ganglion, and bioassay using synthetic peptides revealed that multiple peptides such as APGWamide, buccalin and GnRH induced spawning and/or gonadal maturation. Again, mediation of multiple neuropeptides in the reproduction in mollusk was demonstrated. Unlike in mammals, the master peptide for reproduction has not been identified in mollusk, yet.

In this article, we have described techniques in neuropeptide identification. In the beginning of peptide identification, progress was made one by one, as the new peptide was purified by the classical approach, and the structure and function of identified peptide were elucidated. Now, advance in the mass spectrometer together with the accumulated genetic information such as genome and transcriptome data highly accelerated the identification of peptides. Efforts on the elucidation of physiological functions of identified peptides are making to catch up with accelerated identification of peptides.

Through these works, our knowledge on the neuropeptides expand widely in near future, which means that we obtain all the pieces of a big jigsaw puzzle. When we place the pieces at proper positions, we can see the whole picture of neuropeptide system that regulate reproduction or other activities for life. By looking at the picture, we may be able to save animals in danger of extinction by the artificial breeding. Or, such picture may save the aquaculture industries, because collecting eggs and larvae in the field is getting harder by the climate change. After all, research on neuropeptide is an essential work for our survival, living in the era of climate change by global warming.

**Acknowledgement**

We thank Ms Tomoko Amimoto, a technician in the Natural Science Center for Basic Research and Development, Hiroshima University, for her assistance of mass spectrometry of *Thais* peptides with nanoLC-Orbitrap mass spectrometer. We also thank to all the staffs in marine biological institutes in Hiroshima University, Okayama University, Tohoku University and Shimane University for collection and shipment of *A. kurodai*, to us.
References

[1] Beets I, Temmerman L, Janssen T and Schoofs L 2013 Worm 2 e24246
[2] Price D A and Greenberg M J 1977 Science 197 670-71
[3] Chase R 2002 Behavior and its neural control in gastropod molluscs (New York: Oxford University Press)
[4] Satake H 2006 Invertebrate neuropeptides and hormones: Basic knowledge and recent advances (Kerala, India: Transworld Research Network)
[5] Fujisawa Y, Furukawa Y, Ohta S, Ellis T A, Dembrow N C, Li L, Floyd P D, Sweedler J V, Minakata H, Nakamaru K, Morishita F, Matsushima O, Weiss K R and Vilim F S 1999 J. Neurosci 19 9618-34
[6] Furukawa Y, Nakamaru K, Wakayama H, Fujisawa Y, Minakata H, Ohta S, Morishita F, Matsushima O, Li L, Romanova E, Sweedler J V, Park J H, Romero A, Cropper E C, Dembrow NC, Jing J, Weiss K R and Vilim F S 2001 J. Neurosci 21 8247-61
[7] Morishita F, Sasaki K, Kanemaru K, Nakanishi Y, Matsushima O and Furukawa Y 2001 Peptides 22 183-89
[8] Furukawa Y, Nakamaru K, Sasaki K, Fujisawa Y, Minakata H, Ohta S, Morishita F, Matsushima O, Li L, Alexeeva V, Ellis T A, Dembrow N C, Jing J, Sweedler J V, Weiss K R and Vilim F S 2003 J. Neurophysiol 89 3114-27
[9] Morishita F, Minakata H, Takeshige K, Furukawa Y, Matsushima O and Horiguchi T 2006 Peptide Science 2005 43-46
[10] Morishita F, Minakata H, Takeshige K, Furukawa Y, Takata T, Matsushima O, Mukai S T, Saleuddin A S and Horiguchi T 2006 Peptides 27 483-492
[11] Morishita F, Furukawa Y, Kodani Y, Minakata H, Horiguchi T and Matsushima O 2015 Peptides 68 72-82
[12] Morishita F, Minakata H, Sasaki K, Tada K, Furukawa Y, Matsushima O, Mukai S T and Saleuddin A S 2003 Peptides 24 1533-44
[13] Morishita F, Nakanishi Y, Sasaki K, Kanemaru K, Furukawa Y and Matsushima O 2003 Cell Tissue Res 312 95-111
[14] Takahashi T, Kobayakawa Y, Muneoka Y, Fujisawa Y, Mohri S, Hatta M, Shimizu H, Fujisawa T, Sugiyama T, Takahara M, Yanagi K and Koizumi O 2003 Comp Biochem Physiol B Biochem Mol Biol 135 309-24
[15] Nagahama T, Ukena K, Oumi T, Morishita F, Furukawa Y, Matsushima O, Satake H, Takuwa K, Kawano T, Minakata H and Nomoto K 1999 Cell Tissue Res 297 155-62
[16] Morishita F, Furukawa Y and Matsushima O 2012 Peptides 38 291-301
[17] Newcomb R, Fisher J M and Scheller R H 1988 J Biol Chem 263 12514-521
[18] Holman G M, Cook B J and Nachman R J 1987 Comp Biochem Physiol C 88 27-30
[19] Takeda N, Kon Y, Quiroga A G, Lapebie P, Barreau C, Koizumi O, Kishimoto T, Tachibana K, Houliston E and Deguchi R 2018 Development 145
[20] York P S, Cummins S F, Degnan S M, Woodcroft B J and Degnan B M 2012 Front Zool 9 9
[21] Stewart M J, Favrel P, Rotgans B A, Wang T, Zhao M, Sohail M, O'Connor W A, Elizur A, Henry J and Cummins S F 2014 BMC Genomics 15 840
[22] In V V, Ntalamagka N, O'Connor W, Wang T, Powell D, Cummins S F and Elizur A 2016 Peptides 82 109-119