Australian lungfish neurohypophysial hormone genes encode vasotocin and [Phe^2]mesotocin precursors homologous to tetrapod-type precursors

(molecular cloning/precursor organization/molecular evolution/phylogenetic tree)

S<small>USUMU</small> H<small>YODO</small>*†, S<small>USUMU</small> I<small>SHII</small>‡, <small>AND</small> J<small>EAN</small> M. P. J<small>OSS</small>§

*Department of Biology, Graduate School of Arts and Sciences, University of Tokyo, Komaba, Meguro-ku, Tokyo 153, Japan; ‡Department of Biology, School of Education, Waseda University, Nishi-waseda, Shinjuku-ku, Tokyo 169-50, Japan; and §School of Biological Sciences, Macquarie University, New South Wales 2109, Australia

ABSTRACT   In view of the well-established role of neurohypophysial hormones in osmoregulation of terrestrial vertebrates, lungfishes are a key group for study of the molecular and functional evolution of the hypothalamo-neurohypophysial system. Here we report on the primary structure of the precursors encoding vasotocin (VT) and [Phe^2]mesotocin ([Phe^2]MT) of the Australian lungfish, *Neoceratodus forsteri*. Genomic sequence analysis and Northern blot analysis confirmed that [Phe^2]MT is a native oxytocin family peptide in the Australian lungfish, although it has been reported that the lungfish neurohypophysis contains MT. The VT precursor consists of a signal peptide, VT, that is connected to a neurophysin by a Gly-Lys-Arg sequence, and a copeptin moiety that includes a Leu-rich core segment and a glycosylation site. In contrast, the [Phe^2]MT precursor does not contain a copeptin moiety. These structural features of the lungfish precursors are consistent with those in tetrapods, but different from those in teleosts where both VT and isotocin precursors contain a copeptin-like moiety without a glycosylation site at the carboxyl terminals of their neurophysins. Comparison of the exon/intron organization also supports homology of the lungfish [Phe^2]MT gene with tetrapod oxytocin/MT genes, rather than with teleost isotocin genes. Moreover, molecular phylogenetic analysis shows that neurohypophysial hormone genes of the lungfish are closely related to those of the toad. The present results along with previous morphological findings indicate that the hypothalamo-neurohypophysial system of the lungfish has evolved along the tetrapod lineage, whereas the teleosts form a separate lineage, both within the class *Osteichthyes*.

Neurohypophysial hormones are nonapeptides regulating various physiological events related especially to water and salt metabolism and reproduction. Twelve distinct nonapeptide principles have been chemically characterized in a wide variety of vertebrates, and are classified into two groups: the vasopressin (VP) and the oxytocin (OT) families. They are believed to have developed from a common ancestral molecule by gene duplication (1). All vertebrate species, except for the cyclostomes, contain at least one VP family peptide and one OT family peptide. Several schemes have been proposed as pathways of hormonal molecular evolution based on amino acid sequences of nonapeptides and their phyletic distribution (1).

Complementary DNA and genomic analyses have shown that neurohypophysial nonapeptides are synthesized as large precursor molecules (2). Analysis of the precursor molecules, which are much larger than the nonapeptides, gives more precise and reliable information for estimating the molecular evolutionary relationships than does analysis based on substitutions only in the nine amino acid residues. Using statistical comparison of gene structures and the predicted amino acid sequences of precursors, Urano and colleagues (3–5) have proposed that teleost neurohypophysial hormone genes have their own evolutionary history separate from that of the tetrapod genes. Their hypothesis is further supported by the structural characteristics of neurohypophysial hormone precursors, including composition and the presence or absence of posttranslational modification sites (2–5).

Lungfishes are the closest surviving freshwater relatives of the Devonian fish group from which early amphibians are assumed to have evolved. Acher and his colleagues (6, 7) reported that lungfishes have mesotocin (MT), the non-mammalian tetrapod type hormone, as their OT family hormone, whereas all teleosts examined have isotocin (IT). They have proposed an evolutionary lineage of IT (teleosts)-MT (non-mammalian tetrapods)-OT (mammals) and suggest a close relationship between lungfishes and amphibians. However, the data on amino acid substitutions in nonapeptides are insufficient to estimate the molecular phylogeny of the whole neurohypophysial hormone genes. For their VP family peptide, lungfishes have vasotocin (VT), which is common to all non-mammalian vertebrate species. These previous reports led us to study lungfish neurohypophysial hormone precursors and their genes to further clarify the molecular evolution of neurohypophysial hormone genes.

As material, we used the Australian lungfish, *Neoceratodus forsteri*, which is considered to be more closely related to other freshwater fish than other lungfish species because of the well-developed gills on all gill arches. We isolated cDNAs encoding the neurohypophysial hormone precursors of this species. We report here characteristics of the nucleotide and deduced amino acid sequences of the isolated cDNAs and the structural organization of the precursors. These characteristics, together with the molecular phylogenetic analysis, clearly show that neurohypophysial hormone genes of the Australian lungfish have evolved along the tetrapod lineage, whereas the teleosts appear to form a separate lineage within the class *Osteichthyes*. Furthermore, we present evidence for a peptide, [Phe^2]MT, as a native neurohypophysial hormone in addition to VT in the Australian lungfish.

MATERIALS AND METHODS

Adult lungfish were collected from a tributary of the Mary River in southeastern Queensland in December 1993. Total RNA was extracted from two hypothalami, weighing 340 and 323 mg, with Isogen (Nippon Gene, Tokyo); poly(A)$^+$ RNA was then prepared by using Oligotex-dT30 (Japan Synthetic Rubber, Tokyo). A cDNA library was constructed by using the cDNA Synthesis System Plus and cDNA Cloning System λgt10 (Amersham).

Complementary DNAs encoding portions of VT and MT-like precursors were amplified by a PCR technique from a mixture of phage DNA prepared from the amplified cDNA library; their sequences were determined and were then used as screening probes to obtain full-length cDNAs. For PCR amplification, the following three primers were synthesized: forward and reverse primers that are complementary to regions close to *Eco*RI ends of the left and right arms of λgt10, respectively, and a degenerate VT/MT primer that consists of oligonucleotides predicted from the amino acid sequence of VT and MT (residues 1–7). The amplification was performed with the VT/MT primer and either the forward or the reverse primer in the presence of Ampli *Taq* DNA polymerase (Perkin–Elmer) using a Program Temp Control System, PC-700 (Astec, Fukuoka). After electrophoresis, major bands were excised and ligated into pBluescript II SK+ vector (Stratagene). The nucleotide sequences were determined by an automated DNA sequencer (Li-Cor, Lincoln, NE). The cDNA library was screened by plaque hybridization with ^{32}P-labeled partial VT and MT-like cDNAs. Inserts from positive clones were subcloned into pBluescript vector, and nucleotide sequences were determined as described above.

Sequence analysis indicated that the MT-like precursor contains a new nonapeptide: [Phe^2]MT. Because [Phe^2]MT can be changed to MT by a single nucleotide substitution, its genomic sequence was determined by PCR amplification. Genomic DNA of *Neoceratodus* was prepared from livers of two lungfishes separately with a RapidPrep Genomic DNA Isolation Kit (Pharmacia Biotech). PCR amplification was performed for individual genomic DNAs with sense (−37 to −18 from the translation initiation site) and antisense (119 to 102) primers. PCRs were conducted in duplicate by using different polymerases (Ampli *Taq* and TaKaRa Ex *Taq* polymerases). PCR products were subcloned into pCR-Script SK(+) (Stratagene), and sequences were then analyzed as described above.

Total RNAs (10 μg) extracted from the hypothalami and liver were electrophoresed in 1.2% agarose/formaldehyde gel, and transferred to a NYTRAN Plus membrane (Schleicher & Schüll). Nucleic acids retained by the filters were then hybridized with ^{32}P-labeled cDNA probes for VT and [Phe^2]MT precursors, and the filters were subjected to autoradiography.

Molecular phylogenetic trees of VP and OT family precursors were estimated by using a program package, MOLPHY (8). PROTML, a main program in MOLPHY, can provide evolutionary trees from amino acid sequences by a maximum likelihood method using the empirical amino acid transition matrix compiled by Jones *et al.* (9). Bootstrap probability was estimated by the RELL method developed by Kishino *et al.* (10). To estimate reliability of relationships among hormone precursors of the lungfish, tetrapods, and teleosts, we applied the maximum likelihood method to three hypothetical phylogenetic relationships (topologies) by using the lamprey VT precursor as the out group: A, monophyly of lungfish and toad; B, monophyly of lungfish and tetrapods; C, monophyly of lungfish and teleosts.

RESULTS

PCR amplification of the total lungfish cDNA in λgt10 with VT/MT and λgt10 primers gave two major bands of DNA which included nucleotide sequences for VT and MT. The cDNAs also included nucleotide sequences for Gly-Lys-Arg sequence and the conserved portion of neurophysin, supporting the notion that the cDNAs encode portions of lungfish neurohypophysial hormone precursors. By screening the library with these partial VT and MT cDNAs, 4 VT- and 7 MT-positive clones were obtained. One of four VT-positive clones contained a VT-specific sequence. Five of seven MT-positive clones encoded a MT-like sequence in which the second amino acid residue from the amino terminal was Phe, not Tyr. The remaining two MT-positive clones contained no sequence encoding neurohypophysial hormone. The MT-positive clones thus encoded a new peptide [Phe^2]MT as the mature OT family hormone instead of MT. Northern blot analysis revealed one band of 950 bases for the VT probe and one band of 800 bases for the [Phe^2]MT probe (data not shown).

VT Precursor. VT precursor is composed of four segments: a signal peptide, VT, a neurophysin and a copeptin (Figs. 1–3). VT is connected to its neurophysin by the Gly-Lys-Arg sequence which serves as a signal for proteolytic processing and carboxyl-terminal amidation (Fig. 1). All Cys residues in the neurophysin, which are considered to be important for its conformation, are completely conserved (Fig. 2). Lungfish copeptin includes a Leu-rich core segment and a glycosylation site, as in tetrapod copeptins (Figs. 2 and 3). Two sequential Arg residues are observed at the same site as the Arg residue in the mammalian VP and amphibian VT precursors (Fig. 2). These residues may serve as the processing signal between the neurophysin and the copeptin.

[Phe^2]MT Precursor. The [Phe^2]MT precursor is composed of three segments. Following a signal peptide, the hormone is connected to neurophysin by the Gly-Lys-Arg sequence (Figs. 1 and 2). The [Phe^2]MT cDNA has two possible translation start sites, which predict two signal sequences with 19 or 20 amino acid residues. Neither of the ATG codons matches completely with Kozak's rule (11). Amino acid sequence data for the [Phe^2]MT precursor is thus required to resolve the start codon. The central portion of the neurophysin is highly homologous with that of the lungfish VT neurophysin, and also with that of other vertebrate neurophysins (Fig. 2). However, the [Phe^2]MT precursor is shorter and, like the mammalian OT and toad MT precursors, does not contain a copeptin (Figs. 2 and 3).

Five positive clones had the same nucleotide sequence encoding [Phe^2]MT instead of MT. Genomic sequence was determined to confirm that the Phe residue in the second position in [Phe^2]MT precursor is actually encoded in the genome and not produced by an artifact in the course of cloning. PCR amplification was performed by using genomic DNAs separately prepared from the livers of two individuals as templates. PCR was performed in duplicate by using different polymerases. The analysis of the PCR products confirmed that the gene encodes the [Phe^2]MT precursor. Meanwhile, PCR amplification using the lungfish VT precursor-specific primers and the genomic DNAs confirmed the presence of VT sequence in the VT precursor. These PCR analyses further demonstrated that lungfish VT and [Phe^2]MT genes do not include an intron in their 5′-untranslated regions, signal peptide and hormone coding portions. In hagfish VT and salmonid IT genes, an additional intron is found in the 5′-untranslated region, just prior to the start codon, whereas tetrapod neurohypophysial hormone genes contain no intron in the 5′-untranslated region (12, 13).

Molecular Phylogenetic Tree. We first constructed a phylogenetic tree of VP and OT family precursors by the maximum likelihood method from the lamprey (5), chum salmon (3), masu salmon (14), white sucker (15), lungfish, toad (16), rat (17), and bovine (18) sequence data (Fig. 4). The lamprey VT precursor was used as the outgroup for both VP and OT

```
                                              ┌──────────────── Signal peptide ─────────────────┐
                                        -19  Met Pro Gly Thr Cys Leu Pro Leu Cys Phe Leu Cys Leu Leu Ala Phe Ser Ser Ala
Vasotocin cDNA        gacaagaaacttgagcaccctaaac ATG CCT GGA ACC TGT CTC CCA CTT TGC TTT CTG TGT TTA CTG GCA TTT TCT TCA GCA
[Phe²]MT cDNA  caatctttcaggctttctgaaggacgccagttttatcattcaga ATG ATG TCT TCA GCT TCT ATT CCC ATT TGC TTA ATC TGT CTG TTT GCT TTT ACT TCA GCC
                                        Met Met Ser Ser Ala Ser Ile Pro Ile Cys Leu Ile Cys Leu Phe Ala Phe Thr Ser Ala

 1  ┌──────── Hormone ────────┐   Processing and    ┌──────────── Neurophysin ──────────────▶
                                 amidation signal
    Cys Tyr Ile Gln Asn Cys Pro Arg Gly  Gly Lys Arg  Ser Phe Ile Asp Thr Glu Ile Arg Gln Cys Ile Pro Cys Gly Pro Gln Asn Arg Gly Arg Cys Phe
    TGC TAC ATT CAG AAT TGT CCC AGG GGA   GGA AAA AGA  TCA TTC ATA GAT ACA GAA ATC AGA CAG TGC ATC CCT TGT GGT CCC CAA AAC AGG GGC CGT TGT TTT
    TGC TTT ATT CAG AAT TGC CCC ATT GGA   GGT AAA AGA  TCT GTC TTG GAC ACA GAC CTC AGG AAG TGC ATT TCC TGT GGC CCT GGG AAC AGG GGC CGG TGC TTT
    Cys Phe Ile Gln Asn Cys Pro Ile Gly  Gly Lys Arg  Ser Val Leu Asp Thr Asp Leu Arg Lys Cys Ile Ser Cys Gly Pro Gly Asn Arg Gly Arg Cys Phe

35
    Gly Pro Tyr Ile Cys Cys Gly Glu Glu Leu Gly Cys Tyr Ile Gly Thr Ser Glu Thr Leu Arg Cys Leu Glu Glu Asn Tyr Leu Ser Ser Pro Cys Lys Ala Gly
    GGT CCT TAC ATT TGC TGT GGA GAG GAA CTT GGT TGT TAT ATT GGT ACC TCG GAG ACT CTG AGA TGT CTG GAG GAG AAT TAC TTA TCT TCC CCT TGT AAG GCC GGC
    GGT CCC AGC ATT TGT TGT GGA GAG GAG CTA GGC TGT TTC TTT GGT ACC ACA GAA GCT CTG AGA TGC CAG GAG GAG AAT TAT TGG CCT TCT CCT TGC CAG TCC GGT
    Gly Pro Ser Ile Cys Cys Gly Glu Glu Leu Gly Cys Phe Phe Gly Thr Thr Glu Ala Leu Arg Cys Gln Glu Glu Asn Tyr Trp Pro Ser Pro Cys Gln Ser Gly

70
    Gly Lys Leu Cys Ser Thr Asn Gly Gly Gln Cys Ala Ala Pro Gly Ile Cys Cys Thr Asp Glu Ser Cys Ala Met Asp Ser Ser Cys Leu Asp Gly Asp Ala Asp
    GGA AAA CTC TGC AGT ACA AAT GGC GGA CAA TGT GCC GCA CCT GGC ATT TGC TGT ACT GAT GAG TCC TGT GCC ATG GAC TCC AGC TGC CTT GAT GGA GAT GCT GAC
    GGA AAA CCC TGT GGT TCT GAG GGT GGA AGA TGT GCT GCC CCT GGG ATC TGC TGC AGT GAT GAA AGC TGC TCA GTG GAC TCA ACG TGC --- AAT CAA GAT GTC CTG
    Gly Lys Pro Cys Gly Ser Glu Gly Gly Arg Cys Ala Ala Pro Gly Ile Cys Cys Ser Asp Glu Ser Cys Ser Val Asp Ser Thr Cys --- Asn Gln Asp Val Leu

      ──(Copeptin) ──▶ N-glycosylation site                           Leucine-rich core segment
105
    Lys Arg Arg Met Phe Pro Glu Arg Asn Leu Thr Leu Leu Asp Gly Ala Thr Ser Asp Phe Leu Leu Lys Leu Ile His Leu Ala Asn Arg Gln Gln Gln Glu Glu
    AAA AGG CGG ATG TTT CCT GAA AGA AAC CTC ACC CTC CTG GAT GGA GCC ACG AGT GAC TTC CTC CTC AAA CTG ATT CAC TTG GCC AAC AGG CAA CAA CAG GAA GAA
    TTT CCT TGA accaactacactttcaaagataccaaggcgaacacttctgtatctagagtgtgccccaccttaaaacccaaagcctcaagatacagaacagtttacctatataatatatgccttttcaaaacatgctt
    Phe Pro Stop
    104

140
    Lys His Leu Leu Tyr Stop
    AAG CAT CTG CTG TAC TGA gaactgagaactaacaaccttacaacagcttcttttttatgtctaaacagtgaaacgcctttgagatgtttttactcaaaataaaacgctgacatttgaaaactct-(polyA)
    ttaggcaaatgtacaaatatgaatactgtcaataaaatatgtaaattacgtctacctaag-(polyA)
```
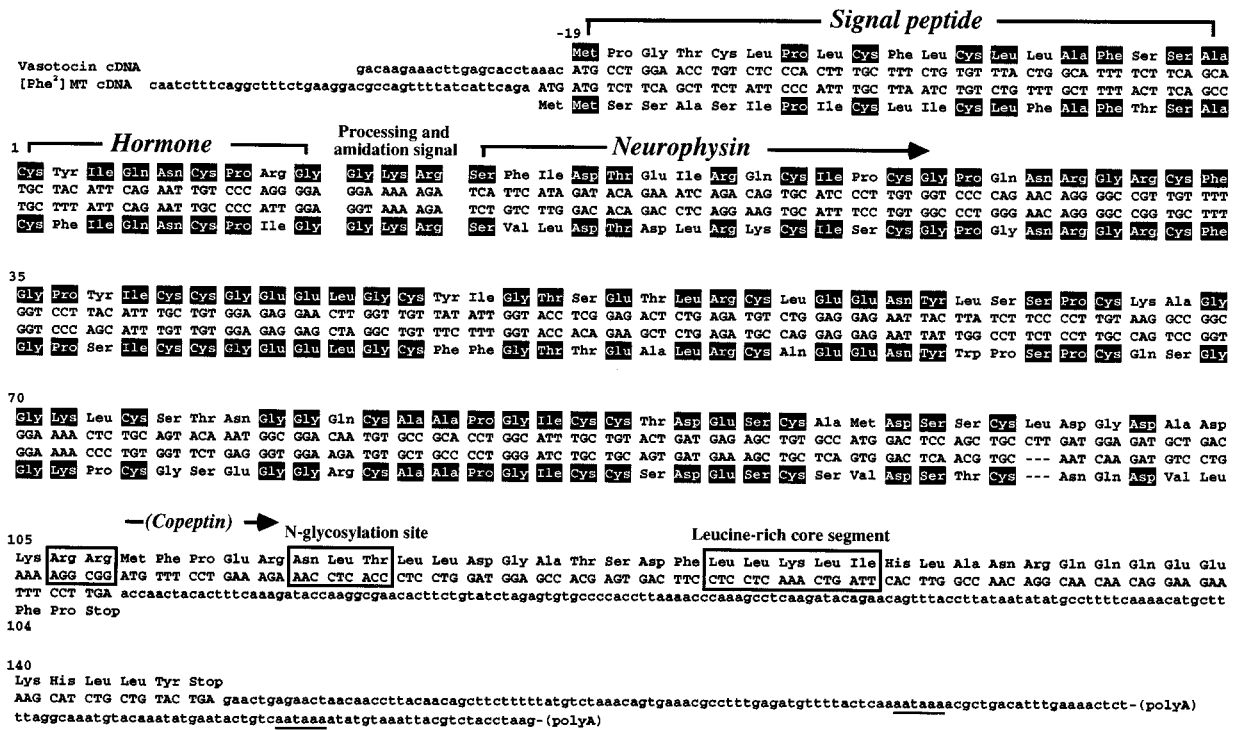
FIG. 1. Primary structures of the cDNAs encoding VT and [Phe²]MT precursors and the deduced amino acid sequences. The amino acid residues are numbered with the first residue (Cys) of hormones as 1. Identical amino acid residues are indicated in black. Sequential Arg residues, N-glycosylation site and Leu-rich core segment are boxed. The AATAAA sequence in the 3′-untranslated region is underlined.

families, because the cyclostome VT gene is considered to be a common ancestral gene for all neurohypophysial hormone genes in gnathostomes. In the VP family tree, the lungfish VT precursor formed a monophyletic group with the toad VT precursor and are equally closely related to mammalian group, whereas the lungfish [Phe²]MT precursor was linked with the tetrapod precursors in the OT family tree (Fig. 4).

Maximum likelihood analysis was then performed for the three topologies shown in Fig. 5 to estimate the degree of reliability for each of the relationships among lungfish, tetrapods, and teleosts. The most likely tree for VP family precursors was tree A (Table 1); tree C was completely ruled out. The bootstrap probability of tree A was higher than 90%. However, tree B could not be ruled out from the analysis, because the

log-likelihood difference of tree B from tree A did not exceed two times SE. Tree B was supported for the OT family precursors, although neither tree A nor tree C could be excluded. Tree C was the least probable, because the bootstrap probability of tree C was lower than 5%. This evidence supports monophyly of the Australian lungfish with the tetrapods and not with the teleosts for neurohypophysial hormone genes.

DISCUSSION

Structural characteristics of precursor molecules contain not only quantitatively but also qualitatively more useful information than amino acid sequences of nonapeptide hormones to

```
                       ┌── Signal peptide ─┐ Hormone    ┌────────────────── Neurophysin ──────────────────
Rat VP          : MMLNTTLSACFLSLLALTSA CYFQNCPRG GKR ATSD-MELRQCLPCGPGGKGRCFGPSICCADELGCFLGTAEALRCQEENYLPSP
Toad VT         : ---TAPVPACFLCLLALSSA CYIQNCPRG GKR SYPD-TAVRQCIPCGPGNRGNCFGPNICCGEDLGCYVGTPETLRCVEETYLPSP
Lungfish VT     : MP-GTCLPLCFLCLLAFSSA CYIQNCPRG GKR SFID-TEIRQCIPCGPQNRGRCFGPYICCGEELGCYIGTSETLRCLEENYLSSP
Chum salmon VT-I : MPYSTFQLLWVLGLLALSSA CYIQNCPRG GKR SFPD-LP-RQCMSCGPGDRGRCFGPNICCGEGMGCYMGSPEAAGCVEENYLPSP
Lamprey VT      : MARCAPLTLAVSVLSLVLISSA CYIQNCPRG GKR DLTD--SVRQCLPCGPGGQRCFGPRICCGEAMGCRLGGPDVAICRAERLMPSP
                                        *  *  ***  *  ***         *   ***  * **** ***   **  *      *   *     **

Chum salmon IT-I : MFGTSVSALCLLFLLSVCTA CYISNCPIG GKR SALAF-PSRKCMACGPGDRGRCFGPNICCGEGMGCYVGSPEAAGCVEENYLPSP
Lungfish [Phe²]MT : MMSSASIPICLICLFAFTSA CFIQNCPIG GKR SVLD-TDLRKCISCGPGNRGRCFGPSICCGEELGCFFGTTEALRCQEENYWPSP
Toad MT         : MSYTAL-AVTFFGWLALSSA CYIQNCPIG GKR SVIDFMDVRKCIPCGPRNKGHCFGPNICCGEELGCYFGTTETLRCQEENFLPSP
Rat OT          : MACPSL-ACCLLGLLALTSA CYIQNCPLG GKR AALD-LDMRKCLPCGPGGKGRCFGPSICCADELGCFVGTAEALRCQEENYLPSP
                                                                └─────── Conservative region of neurophysins

                  ┌──────────────────────────────── ─ ─ ─ ─ (Copeptin) ─ ─ ─ ─ ┐
Rat VP          : CQSGQKPCGS-GGRCGAAGICCSDESCVAEPEC-REGFFRLTRAREQSNATQLDGPARELLLRLVQLAGTQESVDSAKPRVY
Toad VT         : CEAGGKPCSS-GGRCAAPGVCCSDDTCVVDSSCLDEDS-ERRRVTPEQNMTQMDGSASDLLLRLMHMANRQQQ-SKHQFY
Lungfish VT     : CKAGGKLCSTNGGQCAAPGICCTDESCAMDSSCLDGDA-DKRRMFPERNLTLLDGATSDFLLKLIHLANRQQQEEKHLLY
Chum salmon VT-I : CEAGGRVCGSE-GSCAASGVCCDSESCVLDPDCL-EDS--KRQSPSEQNAALMGGLAGD-LLRILH-ATSRGRPQ
Lamprey VT      : CESRGEPCGH-GGKCGAPGLCCSSESCAEDASCGWEGGDSPGERPFPHSALRLQSPAAEAMLELIN-SNSLRD
                         *      *      *  *  **    *      *       *

Chum salmon IT-I : CEAGGRVCGSEEGRCAAPGICCDVEGCSIDQSC---TE-EDEAEYISQSVSSSHG--HDLLMKLINMISHTPPHRVHK
Lungfish [Phe²]MT : CQSGGKPCGSEGGRCAAPGICCSDESCSVDSTCNQDVLFP
Toad MT         : CESGRKPCGNNGGNCARSGICCNHESCTMDPACEQDSVFS
Rat OT          : CQSGQKPCGS-GGRCATAGICCSPDGCRTDPACDPESAFS-ER
                                                      ▲                  ▲
                                              glycosylation site    Leucine-rich core segment
```
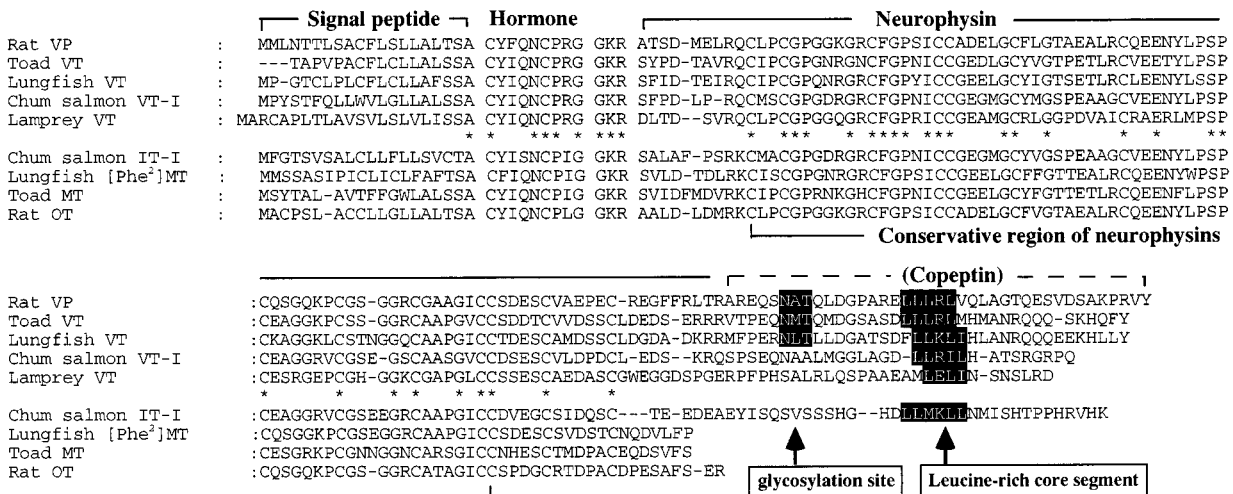
FIG. 2. Comparison of amino acid sequences of rat, toad, Australian lungfish, chum salmon, and lamprey neurohypophysial hormone precursors. Gaps marked by hyphens have been inserted to optimize homology. Asterisks denote identical amino acid residues among all precursors. The N-glycosylation site and the Leu-rich core segment are indicated in black.
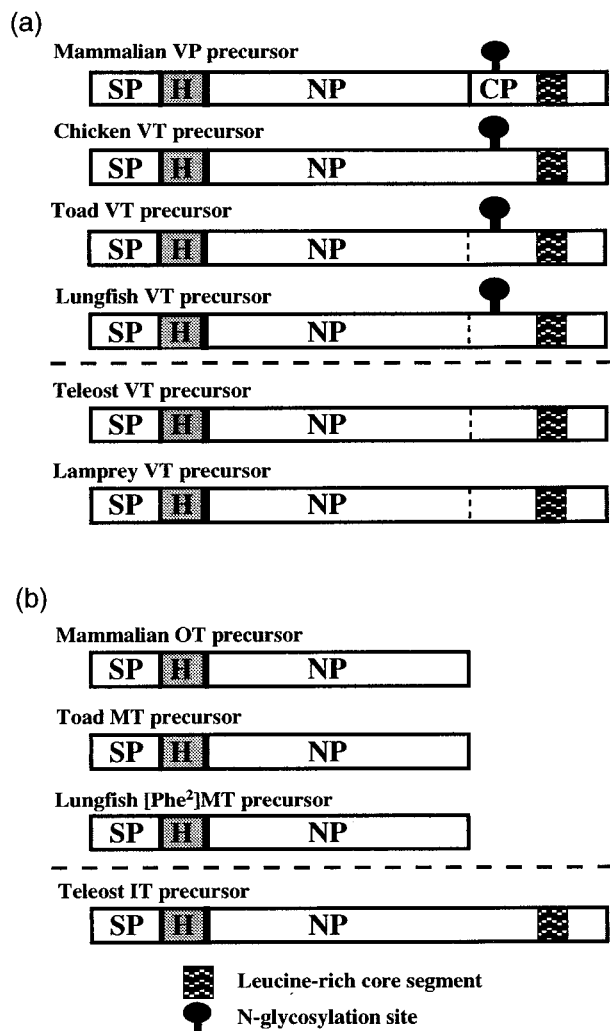
(a)

(b)

Leucine-rich core segment

N-glycosylation site

FIG. 3.   Schematic diagram showing the structural organization of the VP family (*a*) and OT family (*b*) precursors. SP, signal peptide; H, hormone; NP, neurophysin; CP, copeptin region. Vertical lines indicate possible cleavage site between NP and the CP-corresponding domain.

(a)

(b)

FIG. 4.   Phylogenetic trees of precursors of VP family (*a*) and OT family (*b*) hormones, from lamprey, chum salmon, masu salmon, white sucker, Australian lungfish, toad, rat, and bovine. The lamprey VT precursor was used as the outgroup for both VP and OT families. Numbers indicate bootstrap values for each branch.

study molecular evolution of neurohypophysial hormone genes. The present study clearly shows that lungfish VT and [Phe^2]MT precursors are closely related to tetrapod neurohypophysial hormone precursors. The lungfish VT precursor consists of a signal peptide, VT, neurophysin, and copeptin. The copeptin contains a Leu-rich core segment and a glycosylation site similar to those of tetrapod copeptins (Figs. 1–3). Cyclostome and teleost VT precursors also have copeptin-corresponding domains; however, they do not contain a glycosylation site (3, 5, 15). In the Australian lungfish VT precursor, two sequential Arg residues exist between the neurophysin and the copeptin. Arginine residue(s) between neurophysin and copeptin are also found in mammalian VP and amphibian VT precursors (Fig. 2) (16–19). In mammals, the single Arg residue is regarded as the specific recognition site for the processing enzyme that cleaves neurophysin and copeptin. In amphibian VT precursors, however, processing between neurophysin and copeptin does not occur, so that VT and "big neurophysin" are released together from the neurohypophysis (19). Further study is needed to clarify whether the lungfish neurophysin and copeptin domains are cleaved at the Arg residues *in vivo*.

The lungfish [Phe^2]MT precursor, like the tetrapod OT/MT precursors, is composed of a signal peptide, the hormone, and

neurophysin and lacks the portion corresponding to copeptin, so that the length of these precursors is less than that of the VP/VT precursors (Figs. 1–3). In contrast, the carboxyl terminals of teleost IT-neurophysins are about 30 amino acids longer than those of OT/MT neurophysins (Figs. 2 and 3) (3, 5, 14). The extended region in teleost neurophysins is homologous to the copeptin-corresponding domain and includes a Leu-rich core segment, so that the teleost IT precursor has the same structural organization as the teleost and cyclostome VT precursors. Furthermore, salmonid IT genes have a 4 exon/3 intron structure like the hagfish VT gene (13), suggesting that some teleost IT genes have conserved the features of the ancestral neurohypophysial hormone gene. In the lungfish, PCR analysis of genomic DNA showed no intron in the 5′-untranslated and signal peptide coding regions, suggesting a 3 exon/2 intron structure, as occurs in tetrapod neurohypophysial hormone genes. The overall structural organizations of lungfish VT and [Phe^2]MT precursors suggest that the lungfish neurohypophysial hormone genes are included in the tetrapod lineage (Fig. 3) and not the teleost lineage, suggesting an early separation between these two lineages in the *Osteichthyes*.

Molecular phylogenetic analysis also confirms the relationship between the lungfish neurohypophysial hormone genes and those of the tetrapods (Fig. 4). We verified the relationship among neurohypophysial hormone precursors of the Australian lungfish, tetrapods, and teleosts by the maximum likeli-
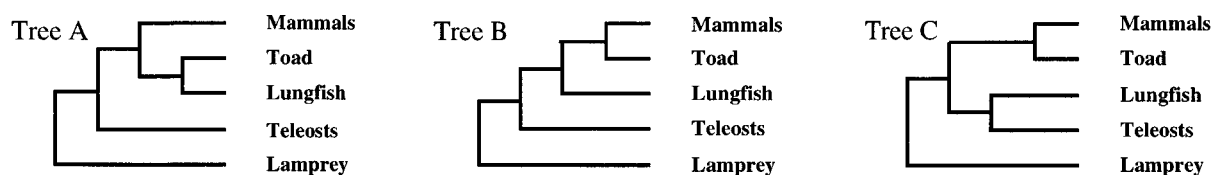
FIG. 5. Three possible topologies illustrating the relationships among neurohypophysial hormone precursors of Australian lungfish, toad, mammals, and teleosts. The lamprey VT precursor was used as outgroup. Tree A, monophyly of the lungfish and the toad; tree B, monophyly of the lungfish and tetrapods; tree C, monophyly of the lungfish and teleosts. The most likely tree among the topologies analyzed by the maximum likelihood method, as shown in Table 1, is tree A.

hood method by using three topologies (Fig. 5). For VP family genes, the monophyly of the lungfish and toad (tree A) was supported, whereas the monophyly of the lungfish and teleosts (tree C) was ruled out (Table 1). For the OT family, although no topology was completely excluded, the bootstrap probability of tree C (monophyly of the lungfish and teleosts) was lower than 5%. These results support the hypothesis that the lungfish neurohypophysial hormone genes belong to the same evolutionary lineage as the tetrapod genes.

The appearance of tetrapods is one of the most dramatic events in vertebrate evolution. Sarcopterygians (extinct rhipidistians, coelacanths, and lungfishes) are almost universally considered to be ancestral to tetrapods. Several hypotheses on the relationship among extant vertebrate groups have been proposed based on molecular and morphological data. Analyses of mitochondrial 12S rRNA and cytochrome *b* gene sequences suggested that the lungfish and tetrapods are included in the same clade (20). On the other hand, mitochondrial cytochrome oxidase I and 28S rRNA gene sequences support the hypothesis that lungfishes and coelacanths form a monophyletic group and are equally closely related to land vertebrates (20). Amino acid sequences of growth hormone (21) and the glycoprotein hormone α-subunit (22) of lungfishes have high homology with those of tetrapods. African lungfish prolactin, like tetrapod prolactins, contains three disulfide bonds and differs from teleost prolactins that lack the amino-terminal disulfide bond (23).

The lepidosirenid lungfishes (*Lepidosiren* and *Protopterus*) have a distinct neurohypophysis that is more similar to that of amphibians than to that of any other fish (24). In *Neoceratodus*, the thin neurohypophysis is located posterior to the adenohypophysis and not dorsal to it as is the case in the lepidosirenid lungfishes (25). Because *Neoceratodus* is considered to be closer to the Devonian ancestral dipnoans and *Lepidosiren* and *Protopterus* are considered to be of much more recent origin, parallel evolution between the lungfishes and modern amphibians has been proposed for the above characteristic observed in the lepidosirenid lungfishes. On the other hand, the pituitary of another extant sarcopterygian, the coelacanth *Latimeria*, is unique and different from the pituitaries of lungfishes and amphibians (26). Marshall and Schultze (27) have pointed out that the neurohypophysis is one character that appeared in the lineage prior to the splitting of lungfish and tetrapods. The morphological data, together with the present molecular re-

sults, suggest that the lungfishes have a hypothalamo-neurohypophysial system homologous to that of amphibians.

The molecular and morphological traits raise interest in the physiology of neurohypophysial hormones in lungfish. Intravenous injection of VT into free-swimming lungfish elicited a dose-dependent diuretic response, as occur in some freshwater teleosts (28). The onset of the diuretic response may depend on the rapid increase in dorsal aortic pressure induced by the administered VT. *Lepidosiren* and *Protopterus* live in shallow waters that are subject to seasonal evaporation. They are able to survive in the absence of environmental water by burrowing beneath the substratum and encysting in a water-impermeable cocoon. Lungfishes have a full complement of ornithine-urea cycle enzymes in the liver and become exclusively ureotelic during estivation (29). The cessation of water intake results in hemoconcentration and marked oliguria (30). Neurohypophysial hormones may play some roles in the estivation process.

In this study we used the degenerate VT/MT primer that encodes the consensus sequence for both VT and MT, as Acher and his colleagues (6, 7) have reported that VT and MT are native hormones in lungfishes (*Protopterus* and *Neoceratodus*). We cloned VT cDNA as the VP family member; however, [Phe^2]MT cDNA was the OT family cDNA. Although one nucleotide substitution can change MT to [Phe^2]MT (TAC/T to TTC/T), we are convinced that the [Phe^2]MT sequence is not an artifact for the following reasons: First, five positive clones from the cDNA library have the identical sequence. Second, genomic PCR amplification of the hormone-encoding sequence confirmed the coincidence of the genomic and cDNA sequences. We amplified from DNAs extracted from livers from two different fish separately: one of these was collected from the same animal as was used for cDNA cloning, and the other liver was obtained from a fish captured later in 1996. Furthermore, PCR amplifications were performed in duplicate using different *Taq* polymerases. Sequence analysis showed the same sequence encoding [Phe^2]MT in all amplified fragments. PCR amplification using the lungfish VT cDNA-specific primers showed the VT-specific sequence, but not [Phe^2]VT, supporting the contention that the codon for Phe residue in the [Phe^2]MT cDNA is not produced as an artifact in the course of cloning and PCR amplification. Third, the [Phe^2]MT gene is expressed in the lungfish hypothalamus, indicating that *Neoceratodus* used in the present study synthesizes [Phe^2]MT as a native hormone. We could find no evidence for MT gene expression in our lungfish hypothalami. At the moment, however, the following three possibilities can be considered: (*i*) an individual Australian lungfish has both [Phe^2]MT and MT, (*ii*) [Phe^2]MT and MT are carried by different individuals in the same wild population of the Australian lungfish, and (*iii*) the Australian lungfish has only [Phe^2]MT.

Table 1. Maximum likelihood analysis of relationships among lungfish, tetrapods, and teleosts based on neurohypophysial hormone precursors

	VP family		OT family	
Tree	1n L*	BP†	1n L*	BP†
A	ML	0.929	−1.2 ± 2.9	0.335
B	−7.5 ± 5.3	0.065	ML	0.618
C	−16.4 ± 7.6	0.006	−8.8 ± 5.8	0.047

*Numbers show the difference of log-likelihood (1n L) of each tree from that of the most likely (ML) tree. The ± indicates standard error.
†Bootstrap probability.

1. Acher, R. (1996) *Gen. Comp. Endocrinol.* **102,** 157–172.
2. Hyodo, S. & Urano, A. (1991) *Zool. Sci.* **8,** 1005–1022.
3. Hyodo, S., Kato, Y., Ono, M. & Urano, A. (1991) *J. Comp. Physiol. B* **160,** 601–608.
4. Urano, A., Kubokawa, K. & Hiraoka, S. (1994) *Fish Physiology*, eds. Sherwood, N. M. & Hew, C. L. (Academic, New York), Vol. 13, pp. 101–132.
5. Suzuki, M., Kubokawa, K., Nagasawa, H. & Urano, A. (1995) *J. Mol. Endocrinol.* **14,** 67–77.
6. Acher, R., Chauvet, J. & Chauvet, M. T. (1970) *Nature (London)* **227,** 186–187.
7. Michel, G., Chauvet, J., Joss, J. M. P. & Acher, R. (1993) *Gen. Comp. Endocrinol.* **91,** 330–336.
8. Adachi, J. & Hasegawa, M. (1992) MOLPHY, Programs for Molecular Phylogenetics I-PROTML (Inst. Stat. Math., Tokyo).
9. Jones, D. T., Taylor, W. R. & Thornton, J. M. (1992) *Comput. Appl. Biosci.* **8,** 275–282.
10. Kishino, H., Miyata, T. & Hasegawa, M. (1990) *J. Mol. Evol.* **30,** 151–160.
11. Kozak, M. (1981) *Nucleic Acids Res.* **9,** 5233–5252.
12. Heierhorst, J., Lederis, K. & Richter, D. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 6798–6802.
13. Kuno, S., Kubokawa, K., Nagasawa, H., Urano, A. & Ishii, S. (1996) in *Proceedings of the Third Congress of the Asia and Oceania Society for Comparative Endocrinology* (Sydney, Australia), pp. 211–212.
14. Suzuki, M., Hyodo, S. & Urano, A. (1992) *Zool. Sci.* **9,** 157–167.
15. Heierhorst, J., Morley, S. D., Figueroa, J., Krentler, C., Lederis, K. & Richter, D. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 5242–5246.
16. Nojiri, H., Ishida, I., Miyashita, E., Sato, M., Urano, A. & Deguchi, T. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 3043–3046.
17. Ivell, R. & Richter, D. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 2006–2010.
18. Ruppert, S., Scherer, G. & Schutz, G. (1984) *Nature (London)* **308,** 554–557.
19. Chauvet, J., Michel, G., Chauvet, M. T. & Acher, R. (1988) *FEBS Lett.* **230,** 77–80.
20. Yokobori, S., Hasegawa, M., Ueda, T., Okada, N., Nishikawa, K. & Watanabe, K. (1994) *J. Mol. Evol.* **38,** 602–609.
21. Rubin, D. A. & Dores, R. M. (1994) *Gen. Comp. Endocrinol.* **95,** 71–83.
22. Arai, Y., Kubokawa, K., Joss, J. M. P. & Ishii, S. (1997) in *Proceedings of the XIII International Congress of Comparative Endocrinology,* ed. Kawashima, S. (Monduzzi Editore, Bologna, Italy), in press.
23. Noso, T., Nicoll, C. S. & Kawauchi, H. (1993) *Biochim. Biophys. Acta* **1164,** 159–165.
24. Goossens, N., Dierickx, K. & Vandesande, F. (1978) *Cell Tissue Res.* **190,** 69–77.
25. Joss, J. M. P., Beshaw, M., Williamson, S., Trimble, J. & Dores, R. M. (1990) *Gen. Comp. Endocrinol.* **80,** 274–287.
26. Lagios, M. D. (1975) *Gen. Comp. Endocrinol.* **25,** 126–146.
27. Marshall, C. & Schultze, H.-P. (1992) *J. Mol. Evol.* **35,** 93–101.
28. Sawyer, W. H. (1970) *Am. J. Physiol.* **218,** 1789–1794.
29. Janssens, P. A. & Cohen, P. P. (1968) *Comp. Biochem. Physiol.* **24,** 887–898.
30. DeLaney, R. G., Lahiri, S., Hamilton, R. & Fishman, A. P. (1977) *Am. J. Physiol.* **232,** R10–R17.