Characterization of the shrimp eyestalk cDNA encoding a novel fushi tarazu-factor 1 (FTZ-F1)[1]

S.-M. Chan*, K.-M. Chan

Department of Zoology, The University of Hong Kong, Pokfulam Road, Hong Kong, China

Received 6 May 1999; received in revised form 31 May 1999

Abstract To study the role of ecdysone and the ecdysone inducible gene in the regulation of molting and development in crustaceans, we have cloned a cDNA encoding an orphan nuclear receptor family member from the eyestalk of the shrimp *Metapenaeus ensis*. The size of the cDNA is 4.3 kb with the longest open reading frame (ORF) encoding a protein of 545 amino acid residues. The deduced amino acid sequence of the shrimp cDNA consists of regions that are characteristic of those of the nuclear hormone receptors. It shows a high degree of amino acid sequence identity in the DNA binding domain, ligand binding domain and the FTZ box as compared to those of invertebrates and vertebrates. Unlike the insects *Drosophila melanogaster* and *Bombyx mori*, an AF2 transactivation domain was present in the shrimp FTZ-F1. Northern blot analysis using total RNA indicated that the FTZ-F1 mRNA could also be detected in the mature ovary. Northern blot analysis and RT-PCR analysis showed that the shrimp FTZ-F1 transcripts could be detected in the ovary, newly hatched nauplius, testis, eyestalk and epidermis of the adult shrimp. Although the cDNA clone was isolated from the eyestalk library, the shrimp FTZ-F1 appeared to express most abundantly in the mature oocytes. The presence of abundant FTZ-F1 specific maternal message in the late vitellogenic ovary and early nauplius indicates that it may be important for the early embryonic and larval development of the shrimp. Interestingly, shrimp FTZ-F1 can also be found in testis of the male shrimp. The presence of FTZ-F1 in other tissues such as epidermis suggests that it may also be involved in other physiological processes such as molting.

Key words: Shrimp; FTZ-F1; Nuclear receptor

1. Introduction

FTZ-F1 is a member of the orphan nuclear receptor family that interacts with the *ftz* transcription starting site and activates *ftz* transcription in *Drosophila* [1–3]. It plays a role in the blastoderm and nervous system development [3] and is one of the ecdysteroid inducible genes in molting and metamorphosis [4]. In addition to insects, several other FTZ-F1 family genes have also been cloned in different vertebrates. In vertebrates, the biosynthesis of steroid hormone is regulated by the tissue-specific expression of the cytochrome P450 steroid hydroxylases [5]. Analysis of the promoter of the steroid hydroxylase gene revealed the presence of elements specific for a transcription regulator, FTZ-F1. SF1 (or Ad4BP), the vertebrate homologue of the insect FTZ-F1, is a primary transcription regulator of several hormone and steroidogenic enzyme genes [6,7]. These genes are critical for normal physiological function of the hypothalamic-pituitary-gonadal axis in reproduction and sexual differentiation of gonads in vertebrates [8]. In crustaceans, molting is initiated by a decrease in the hemolymph level of the molt-inhibiting hormone (MIH). MIH inhibits the production of steroid ecdysone by the Y-organ. Therefore, the molting event is regulated by ecdysteroid and ecdysteroid inducible nuclear receptors. Much information has been accumulated on the molecular mechanism involving molting and reproduction in insects. However, only a few reports have described the cloning of steroid/nuclear receptor in crustaceans. For example, information on cDNA for th ecdysone receptor and retinoic X receptor has been reported for the fiddler crab *Uca pugilator* [9]. In the lobster *Homarus americanus*, the isolation of a potential ecdysone inducible gene that is homologous to the insect MHR3 has been reported [10]. We have also reported the cloning of a cDNA encoding the ecdysone inducible gene E75 homologue in the shrimp *Metapenaeus ensis* [11]. In summary, there is little information on the role of nuclear hormone receptor in the control of growth and development in crustaceans. Whether insect and crustacean share a similar mechanism in the control of growth, molting and reproduction is unknown. In our previous cloning of the E75 cDNA by RT-PCR, several partial cDNAs encoding the DNA binding domains of different steroid hormone receptors were obtained [11]. One of the partial cDNAs encoding a protein was highly homologous to the nuclear receptor FTZ-F1 of *Drosophila* and other vertebrates. To extend our study on the role of nuclear receptor in the regulation of gonad maturation and molt cycle in crustaceans, we report the cloning and characterization of a cDNA encoding a shrimp FTZ-F1.

2. Materials and methods

2.1. Experimental animals

Shrimp (14–18 g) were purchased from a local sea-food market and acclimated to laboratory conditions of 25°C, salinity at 28–30‰, in a 12 h light and 12 h dark photoperiod. Tissues (eyestalk, muscle, hepatopancreas and swimming leg) from animals of different molt stages were dissected for total RNA extraction.

2.2. RNA isolation/RT-PCR

Total RNAs from eyestalk, hepatopancreas, ovary, testis, muscles, epidermis, nerve cord and pleopod (swimming leg, containing cuticle, epidermis and muscle) were isolated by guanidinium thiocyanate extraction [12]. Poly(A)$^+$ RNA was prepared by a Poly(A)Tract mRNA isolation kit (Promega, USA) and used for cDNA library construction as well as reverse transcriptase-polymerase chain reaction (RT-PCR). Reverse transcription was performed in a final volume of 20 μl and 1×transcription buffer (50 mM Tris-HCl, 8 mM $MgCl_2$, 30 mM KCl, 2 mM each of dNTP, 10 mM DTT), 10 pmol of oligo(dT)$_{17}$ primer,

*Corresponding author. Fax: (859) 2857-4672.
E-mail: chansm@hkucc.hku.hk

[1] The nucleotide sequence for the shrimp MeFTZ-F1 has been submitted to the GenBank database.

PII: S0014-5793(99)00787-5

MeFTZ-F1

```
        5'  GAATTCGGCCGAGACAACAACTTGTTGACCCCGACTCTA ATG GAC TCT GGT TTA TTC CCC GGT GTA GCC ACC ACC
                                                    M  D  S  G  L  F  P  G  V  A  T  T    12
ACC CTG GAC TAC ACG TCT CTG GCC GAG CTC CCA GAC ACC AAG GAG GGC ATC GAG GAG CTG TGT CCT GTG TGC GGG
 T  L  D  Y  T  S  L  A  E  L  P  D  T  K  E  G  I  E  E  L  C  P  V  C  G    37
GAT AAG GTG TCA GGC TAC CAC TAT GGC CTG CTC ACC TGC GAG TCC TGC AAG GGA TTC TTC AAG CGG ACG GTG CAG
 D  K  V  S  G  Y  H  Y  G  L  L  T  C  E  S  C  K  G  F  F  K  R  T  V  Q    62
AAC AAG AAG GTG TAC ACC TGC GTC GCA GAT CGT TCC TGC CAG ATA GAC AAG ACG CAG AGG AAG CGA TGT CCG TAC
 N  K  K  V  Y  T  C  V  A  D  R  S  C  Q  I  D  K  T  Q  R  K  R  C  P  Y    87
TGT CGC TTC CAG AAG TGC CTC GAG GTT GGC ATG AAG CTG GAA GCG GTG CGT GCC GAC CGA ATG CGG GGT GGC AGA
 C  R  F  Q  K  C  L  E  V  G  M  K  L  E  A  V  R  A  D  R  M  R  G  G  R   112
AAC AAA TTT GGT CCC ATG TAC AAA CGT GAT CGA GCA CGA AAG CTG CAA TTG CTC CGG CAG AGA CAG CTG TCC CAG
 N  K  F  G  P  M  Y  K  R  D  R  A  R  K  L  Q  L  L  R  Q  R  Q  L  S  Q   137
CAG GGC ATC CTT TCG GGA GGT GCT CGA CAC ACC AGC TCT GGC GTG GCA TCA CGT ACT CCA CCC GGG GCT ACT CCT
 Q  G  I  L  S  G  G  A  R  H  T  S  S  G  V  A  S  R  T  P  P  G  A  T  P   162
CGG CGC CAT CCA CAC ACG TCC ACA TCA AGG AAG AGA TCC AGA GTC CCT TCC TCT CCT CGT CCA CAT CTT CGC CAG
 R  R  H  P  H  T  S  T  S  R  K  R  S  R  V  P  S  S  P  R  P  H  L  R  Q   187
ACT CCT CGC CGT CCC CCA TGG CAA TTG GGC GGC CTC GTC GCA GGC TCG GGC GGG CGT GGG CGG CCT AGT GCC TCC
 T  P  R  R  P  P  W  Q  L  G  G  L  V  A  G  S  G  G  R  G  R  P  S  A  S   212
GGC CCT GTG GCA CCC ATC CTC GCT GGG CCC GAT CCC GCC CTG TGG GTC ACC AAC GCC CAA TCC ACG GCG GGT GGC
 G  P  V  A  P  I  L  A  G  P  D  P  A  L  W  V  T  N  A  Q  S  T  A  G  G   237
GTA ACA ACG GGC ACA CCA CCT ACG GGA GGG CGT CGG CGG TGG AGG AGG AGG AGG AGG CGG GGG AGG AGG ACG
 V  T  T  G  T  P  P  T  G  G  R  R  R  W  R  R  R  R  R  R  R  G  R  R  T   262
CAC AGT ACG CCC TTG TAT CCC TTC ACC ATT ATC CGA GAA TTG GTG GAG ACC GTT GAC GAC CAG GAG TGG CAG GCG
 H  S  T  P  L  Y  P  F  T  I  I  R  E  L  V  E  T  V  D  D  Q  E  W  Q  A   287
TTG CTC TTC TCG CTT CTC CAG AAC CAG ACG TAT AAC CAA TGC GAA GTT GAC CTC TTC GAG CTC ATG TGC AAA GTC
 L  L  F  S  L  L  Q  N  Q  T  Y  N  Q  C  E  V  D  L  F  E  L  M  C  K  V   312
CTC CAA CAG AAC CTC TTC GCT CAG GTC GAC TGG CCG CGC AAC TCT TGC TTC TTC AAA GAC TTG AAG GTT GAC GAT
 L  Q  Q  N  L  F  A  Q  V  D  W  P  R  N  S  C  F  F  K  D  L  K  V  D  D   337
CAG ATG AAA CTT CTA CAA CAT TCA TGG TCA GAT TTA CTT ATC TTG GAT CAT CTA CAT CAA CGT ATA CAC AAC CGA
 Q  M  K  L  L  Q  H  S  W  S  D  L  L  I  L  D  H  L  H  Q  R  I  H  N  R   362
CTG CAA GAT GAA ACC ACA CTC CCT AAT GGG CAG AAG TTT GAT TTG CTT TCC CTG GCG CTG TTG GGA ACA ACT CAG
 L  Q  D  E  T  T  L  P  N  G  Q  K  F  D  L  L  S  L  A  L  L  G  T  T  Q   387
TTT GCA GAT CGT TTC CAT GCT ATT CTC AAC AAG CTG AGG GAC CTC AAC TTT GAT ATT TCA GAC TTT GTC TGT GTC
 F  A  D  R  F  H  A  I  L  N  K  L  R  D  L  N  F  D  I  S  D  F  V  C  V   412
AAA TTT ATT ATT CTA CTG AAT CCT GAC TCT ATT GCA GAT GTA CGG TTG CTT AGC GAT CGT AGA GCT GTG ATT GCT
 K  F  I  I  L  L  N  P  D  S  I  A  D  V  R  L  L  S  D  R  R  A  V  I  A   437
GCT CGG CCA GGT GCG ACA GGC ATT GAT GGA ATA TAC AGC CAG TGT TTG CCC AGA GTG AGT AGT GGA GAA ATC CAG
 A  R  P  G  A  T  G  I  D  G  I  Y  S  Q  C  L  P  R  V  S  S  G  E  I  Q   462
AAG CTC ATG GAC CTT TTG CCT GAA TTG CAC TTC CTT GCT GAA AAT GGG GAA AAG TAC TTG TAT TAC AAG CAC ATA
 K  L  M  D  L  L  P  E  L  H  F  L  A  E  N  G  E  K  Y  L  Y  Y  K  H  I   487
AAT GGA GCA GCA CCC ACA CAG ACC TTG CTA ATG GAG ATG TTG CAC AAC CAA AAG GAA ATA GAG GAA GTG GTA ACT
 N  G  A  A  P  T  Q  T  L  L  M  E  M  L  H  N  Q  K  E  I  E  E  V  V  T   512
CCC TGC TTT CAA CCG TCC AGT GCG CCA CAT GTG TAT ATA AAG GGT GGG TCT GTG AGA GTC TCG AGT CAC CAC GTC
 P  C  F  Q  P  S  S  A  P  H  V  Y  I  K  G  G  S  V  R  V  S  S  H  H  V   537
AGG CCA GAA GGA GGG TCT CCA CCA TGA CCAACACAACATCCAACATCCTGGAACGACCATCATGGGATTCTCAGCCTATGTTTCCACCAAA
 R  P  E  G  G  S  P  P  *                                                    545
CTATATAGTACATGGGAGCGGACGCCCGTTGACACTGCCCAGCCCACAGAAGCCAGCCTCGGGTCAGCTCATCAGTGCATCAGTCAGTAATTCAGAGTTA
ACAATGAAAGAAGAAAGGTAAAATCTGTCTTATAGAGAATAGAACTTTTGAACACTCATTTTTTACTTTAAGTTACTGAGTGTAAAACTAAGAGATCAAC
TCTGACCCTATACCATAACGGTCACAAGCATCCGTTTGGTGAAGTGCTGTGCTAGTCATACCCTTAATGGGCAGTATTGCATTCCTGAGCCTGAGACCTT
ACTGTGTAATTTATGGTGCTAAGCCTTCTTGGGGAAAGCCATGAGGTACAAACAGCCCAAAGAACTCTCTTGAGGCAAGTGGTTAGTGGTGAGTGCTAAA
TGGTCCCTGTCAGTTGTTACCAGAACGACCATGGTGCAATCAGGACTAATCTTAGGAAGTTGAGGCTCCAAGTGTTGATCAGTCAGTGGTGCGCTACTCT
CACTGAACGATATTATCCCTTACACATCCTCTTTCTTTGGCGCATCATACCTCATTCTCACTCTTTCTTATGGTGGCGGTCTCCTCACTCACTTTGGTAC
AAAGCTCTCACTTGTCAATATGTTTCACGACTTGTATAACTATTTATTTCTAAATATAATTGTAAAACCTGGGCACATGTCCAAAAAGCTGTTTGCCCAG
ATATAAATGATATCTTCATAATGCATAACACATAGTGTAAGTGTTGTGAAATAGTTTTTGATACTGTTGTTTTGTACAAGAGATTATTGTAGCAAAATAA
TGTATACTTGCATGAGCAAGCATCTTTACCCAGATATTGCTGCCACTTGTCCAGGATATGTGTTCATGCGAGGTAGGCAGGCTGGCTGCTCTACTTGTGT
ATGTCTGTCACTTAGAAGATAATGATTTGTAGTTTGTTAATTAAGAAGCAGAAGCATAAAATTCTAGTGCCTCAACTATGTTAACCTAGCATGTACAAAA
CACAGACATTACTGCAGTTGCGGCACATGAACCAGCACAGACAAGAGGCATATGTATGTTTATATCATTGGTTACATGCATTTCTCTTATTTTTAATAAC
AATGTACACTTTTATGGTTTTTATTTCATGAACTTTAGAAGGCATATTGTTTGGCTTTGTATTTCTTTTACAAAGAAATTTTTGCCTGCATTTTATGTCT
TTCAGTTAACAGAAAAATTTATTAGACAAAAGTAGTTCTGAATGTTATATTTAGAATTGAGAATTTGTAATGCAATGAAAATTTCAAATTAAGAATATAT
TGCCTGGCAATGAGAAGCAATACTTCTTTTCTTAGGTATTTTCTTTTTATAGTTTTTAATTGACTATTTAGTTAAGCATTGAATTAAGATTTGGAACATT
CCAATTCAAGGGATTCATGGGCATTCTTCACTTTGGTTATGCGTTCCCCTGGTGGTCTATTAGCGTAATCATTGTTGTCATCAATCATTAGTATCATCAC
TATCGCTGCTGCACATTCCGCATAGCTAATTATCATTTGGCTTTTTTTTGCTGCTAAAACACTGCATTTCTTCTCTCGTATTTCTTGGTTTTCTTTCATT
TCCTCTCTTGGCCACATATGTATAGTTTTCTTGTAGATATGTAAGTGACTTTTGTATTTTCTTCTCTTATTGCTCCTCATTGCAGGACTTCAAATTTTAA
CTACTTTAATAATTTTTTTATTTATTTAATGAAGGAAAGAACAAACATTGATAAGATGGTGTGCAACGGCAGCAGCTAACGGTACCCCTGCCACTTTTAA
CATTAATGCTGTATTTCCTATCTCTCACCTACTTACAATGTGCACATCTATTGTTTTTTATTGTTATTATAATTATTATTATAATTATTATCATTATAAT
TTTAGAAAATTTAAATATCTGTGCTTTTTCAAAAAGCATTACCAGTGAAAAATTATAAGGCATCTGGCCTTCACAAAAACAAATTATTTTTGTACCAATT
GAAAATTATTATCCCTGAAATCTACTATAGCTATTTACTATGTAATTGAAACACAGTCTGACTTAGATCTATCAGGATATGGTGTACAGCTGTTGAATTC
TAAGAGAGATTTATTTTTGGAACCAATGTTTTTCAGCACAAGAAAGGTGTTGGGATGAATCTAAGTCACACTGTGGTTATTCTTAATTCATTATTACCAT
GGAAATTGCTGAGTTTGAGCAATGGATATAAATGCAAGCACGAAGTGCGTATTGCATTCTGAATGGAATATGGACGCAATTGAGTTCTATGTTTATGGCT
CAGACTCACTAATTTCTTTTTACTTTTACTGGTCTACACAGTTCTTTCTTCACTCTAAAATGATTTGTTCAATGTCATTTTCCTTTGACATGTTTCTAGT
AAAAGAATGTGCTTGAAATCATCTATTTACTTTTTCATATGGCATTGAGCTACACTTCTAAACACTGAAAAAATTACATATTTCTCTCACGTCTGTAAAG
AAATAAACAAGTGTTCTTTGTAAAAATTGTATACTTGTTTATTGTAAAAAAAAAAACTGTACTCTAAGAAGAATATATTTTAATCTTCATCTAATGATTC
CCAGATTGTTTGTTCCTGTAAGGGAACATGCTAGACATGCTTTGAAACTTTTTCTTCATTATTTCACACCAGTGCCTATCGATAACTTTGGTCGGTATTG
GTGAATCATCACTTGAAGATTATGTTTTATATGAGACAACCCCAAACACGGAGTAGCCCTGGGATTGGTAGGTTCGAATCATCACTTGAAGATTATGTTT
TATATGAGACAACCCCAAACACGGAGTAGCCCTGGGATTGGTAGGTTCTCTTATTAGATTCTATTATTAATGCAAATAATGAGAATGGGATTAATAAAGT
GTTTTATTTAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
```

Fig. 1. Nucleotide and deduced amino acid sequence of the shrimp MeFTZ-F1. The putative DNA binding domain and the ligand binding domain are shown in light and dark gray boxes respectively. The FTZ-F1 box (italicized and bold) and the AF2 core (bold) characteristics of the FTZ-F1 are also indicated. Numbers on the right indicate amino acid positions. The putative poly-A signal sites (AATAAA) are in bold letters and the locations of the primers used in RT-PCR are underlined.
←

and 1 unit of reverse transcriptase (Promega, USA). The reaction mixture was incubated at 42°C for 2 h and used as a template in PCR. For RT-PCR, the final PCR mix (30 μl/reaction) consisted of 10 mM Tris-HCl, pH 8.0, 1.5 mM $MgCl_2$, 50 mM KCl, 10 pmol primers M_1 and M_2 and 2.0 μl reaction mix from the reverse transcription as described above. The PCR conditions included 25 cycles each of denaturation at 95°C for 1 min, annealing at 55°C for 1 min and extension at 72°C for 1 min. At the end of the last cycle, the PCR mix was further incubated at 72°C for 10 min for the completion of DNA synthesis.

2.3. Northern blot hybridization and RT-PCR Southern blot

For Northern blot analysis, 10 μg of total RNA from various tissues was run on a 1.2% agarose gel. RNAs were then transferred onto a nylon membrane (Boehringer, Germany). After pre-hybridization of the membrane at 42°C (0.5 M NaCl, 0.02 M PIPES, pH 6.5, 50% formamide, 1% SDS, and denatured salmon sperm DNA of 100 μg/ml) for 1 h, the membrane was hybridized in fresh hybridization buffer containing a probe spanning the ligand binding domain. For RT-PCR, cDNAs from various tissues were amplified with specific primer derived from the cDNA of the shrimp MeFTZ-F1. PCR products were transferred onto a nylon membrane for Southern blot analysis. The probe used in PCR contained DNA sequence exclusively derived from the internal sequence of the original PCR product. To confirm the absence of genomic DNA in RT-PCR that would contaminate the PCR product, genomic DNA was also amplified using the same pair of primers used in RT-PCR. Larger DNA fragments were amplified from genomic DNA using primers from RT-PCR (data not shown). This suggested that at least one intron was present in those two regions.

(a)

```
                                                                    -- 1ST ZINC FINGER --
MeFTZ-F1 -------------------------MDSGLFPG--VATTTLDYTSLAELPDTKEGIEELCPVCGDKVSGYHYGLLTCESCKGFFKRTVQN     63
DmFTZ-F1 NGNPNAGHMSSGSVGNGSGGAGNGGAGGNSGPGNPMGGTSATPGHGGEVIDFKHLFEELCPVCGDKVSGYHYGLLTCESCKGFFKRTVQN    540
ZFF1A    ------------MLPKVES-EYLGLARSHGEQGH-MPGNMQAPQFKMMDYSYDEDLDEMCPVCGDKVSGYHYGLLTCESCKGFFKRTVQN     76
mSF-1    HRRPIPARSRLVMLPKVET-EAPGLVRSHGEQGQ-MPENMQVSQFKMVNYSYDEDLEELCPVCGDKVSGYHYGLLTCESCKGFFKRTVQN    137

         -- 2ND ZINC FINGER          ------ FTZ-F1 BOX  ------
MeFTZ-F1 KKVYTCVADRSCQIDKTQRKRCPYCRFQKCLEVGMKLEAVRADRMRGGRNKFGPMYKRDRARKLQLLRQRQLSQQGILSGGARHTSSGVA    153
DmFTZ-F1 KKVYTCVAERSCHIDKTQRKRCPYCRFQKCLEVGMKLEAVRADRMRGGRNKFGPMYKRDRARKLQVMRQRQLALQALRNSMGPDIKPTPI    630
ZFF1A    NKRYTCIENQSCQIDKTQRKRCPYCRFQKCLTVGMKLEAVRADRMRGGRNKFGPMYKRDRALKQQKK-------ALIRANGLKLEAMTQV    159
mSF-1    QKRYTCIENQNCQIDKTQRKRCPYCRFKKCIDVGMKLEAVRADRMRGGRNKFGPMYKRDRALKQQKK-------ALIRANGLKLEAMSQV    220

MeFTZ-F1 SRTPPGATP----RRHPHTSTSRKRSRVP-SSPRPHLRQTPRRPPWQLGGLVA-----GSGGRGRPSASGPVAPILAGPDPALWVTNAQS    233
DmFTZ-F1 SPGYQQAYPNMNIKQEIQIPQVSSLTQSPDSSPSP-IAIALGQVNASTGGVIATPMNAGTGGSGGGGLNGPSSVGNGNSSNGSSNGNNNS    719
ZFF1A    MQTVPADLT--------ITSAIQNIHSASKGLPLSHHHHHHHHHHHHSSSSAGLP---PADFDRSPFVTSPVSMAMPPHAGG---LQGYQ    235
mSF-1    IQAMPSDL----------TSAIQNIHSASKGLPLSH---------------VALP---PTDYDRSPFVTSPISMTMPPHSS----LHGYQ    278

MeFTZ-F1 TAGGVTTG------TPPTGGRRRWRRRRRRRGRR--------------------------THST--------PLYPFTIIRELVETVDDQ    283
DmFTZ-F1 STGNGTSGGGGGNNAGGGGGGTNSNDGLHRNGGNDSSSCHEAGIGSLQNTADSKLCFDSGTHPSSTADALIEPLRVSPMIREFVQSIDDR    809
ZFF1A    AYGHFQSRTIKSEYPTPTQAR-QSPHGLPLRRSLR-------------------------SGSP--------PSFPHLVVELLKCEPDEP    291
mSF-1    PYGHFPSRAIKSEYPDPYSSSPESMMGYSYMDGYQ-------------------------TNSP--------ASIPHLILELLKCEPDEP    335

                                    ------ L   B   D    D O M A I N   I   -----
MeFTZ-F1 EWQALLFSLLQNQTYN---QCEVDLFEL-MCKVLQQNLFAQVDWPRNSCFFKDLKVDDQMKLLQHSWSDLLILDHLHQRIHNRLQDETTL    382
DmFTZ-F1 EWQTQLFALLQKQTYN---QVEVDLFELLMCKVLDQNLFSQVDWARNTVFFKDLKVDDQMKLLQHSWSDMLVLDHLHHRIHNGLPDETQL    896
ZFF1A    QVQAKILAYLQQEQASRGKHEKLNTFGL-MCKMADQTLFSIVEWARSSIFFRELKVDDQMKLLQKCWRELLILDHVFRQVMHAKEGSILL    380
mSF-1    QVQAKIMAYLQQEQSNRNRQEKLSAFGL-LCKMADQTLFSIVEWARSSIFFRELKVDDQMKLLQNCWSELLILDHIYRQVAHGKEGTIFL    424

                                    --  L B D   DOMAIN II --
MeFTZ-F1 PNGQKFD---LLSLALLGTTQFADRFHAILNKLRDLNFDISDFVCVKFIILLNPDSIADVRLLSDRRAVIAARPGATGIDGIYSQCLPRV    456
DmFTZ-F1 NNGQVFN---LMSLGLLGVPQPGDYFNELQNKLQDLKFDMGDYVCMKFLILLNP---S-VRGIVNRKTVSEGHDNVQAALLDYTLTCYPS    979
ZFF1A    VTGQQVDYALIASQAGATLNNLLSHAQELVSKLRSLQLDQREFVCLKFLVLFSLD----VKNLENFHLVESVQEQVNAALLDYVMCNYPQ    466
mSF-1    VTGEHVDYSTIISHTEVAFNNLLSLAQELVVRLRSLQFDQREFVCLKFLVLFSSD----VKNLENLQLVEGVQEQVNAALLDYTVCNYPQ    510

                                                AF2
MeFTZ-F1 SSGEIQKLMDLLPELHFLAENGEKYLYYKHINGAAPTQTLLMEMLHNQKEIEEVVTPCFQPSSAPHVYIKGGSVRVSSHHVRPEGGSPP    545
DmFTZ-F1 VNDKFRGLVNILPEIHAMAVRGE-----DHLITCTPSTVPAVRPPKR----------CSWRCCTPSARDRG-RENVTRNT---------   1043
ZFF1A    QTDKFGQLLLRLPEIRAISLQAEEYLYYKHLNGDVPCNNLLIEMLHAKRA-------------------------------------    516
mSF-1    QTEKFGQLLLRLPEIRAISKQAEDYLYYKHVNGDVPYNNLLIEMLHAKRA-------------------------------------    560
```

(b)

```
MeFTZ-F1 GFFKRTVQNKKVYTCVADRSCQIDKTQRKRCPYCRF 36
         :::::::::::..:::::: .:.: :.::.::  :::
AsFTZ-F1 GFFKRTVQNKRIYTCVADGKCEITKAQRNRCQPCRF 36
```

Fig. 2. a: Amino acid sequence alignment of the shrimp FTZ-F1 (MeFTZ-F1), the fly (DmFTZ-F1), zebrafish (ZFF1A) and mouse (mSF-1). Insertions (dashes) are added to allow for maximum degree of identity. The conserved regions are in gray boxes. A part of the N-terminal portion of DmFTZ-F1 and mSF-1 is deleted. b: Alignment of the amino acid sequence from the partial DBD of the shrimp MeFTZ-F1 and the brine shrimp AsFTZ-F1. Identical amino acids are shown by colons and similar amino acids are shown by dots.

Table 1
Comparison of the length and percentage identity of the amino acid sequences of the shrimp FTZ-F1 with other species

Region	Amino acid (length)							Identity (%)					
	Me	Bm	Dm	Dr	Xl	Mm	Hs	Bm	Dm	Dr	Xl	Mm	HS
A/B	32	88	509	40	45	12	12	25	6	9	12	9	9
C	66	66	66	66	66	66	66	92	98	89	88	86	86
D	183	158	233	181	166	160	159	30	31	14	13	13	15
F_1 box	28	28	28	28	28	28	28	100	100	96	96	93	93
E	112	112	112	115	115	115	115	69	71	41	43	37	36
F	124	103	95	81	81	81	81	12	15	20	20	18	18
AF2	6	–	–	6	6	6	6	–	–	83	83	83	83
Overall	545	555	1043	516	501	462	461	45	45	34	34	32	32

The species shown here and the GenBank accession numbers are *Metapenaeus ensis* (Me; submitted); *Bombyx mori* (Bm; P49867); *Drosophila melanogaster* (Dm; P33244); *Danio rerio* (Dr; AF014926), *Xenopus laevis* (Xl; A56543), *Mus musculus* (Mm; P33242), *Homo sapiens* (Hs; Q13285).

2.4. Library screening

An eyestalk library was screened with the shrimp specific FTZ-F1 cDNA probe (approx. 140 bp). The probe was synthesized by a random primer labeling kit using [α^{32}P]dCTP as a tracer. Hybridization of replica filters was performed at 65°C in hybridization buffer (0.5 M sodium phosphate, 1% SDS, 2 mM EDTA, 1×Denhardt's solution and 100 μg/ml denatured salmon sperm DNA) without formamide overnight. Filters were washed at a final concentration of 0.5×SSPE (20×SSPE: 3.6 M NaCl, 0.2 M $Na_2HPO_4{\cdot}7H_2O$ and 20 mM EDTA, pH 8.0), 0.1% SDS at 65°C twice for 30 min each time. After the third screening, potential positive phage plaques were purified and the recombinant pBK-CMV phagemids were rescued from the bacteriophage clones by in vivo excision according to the instructions of the manufacturer (Stratagene, USA). Potential positive cDNA clones were used for DNA sequence determination. DNA sequencing was performed using T_7 DNA polymerase and the dideoxynucleotide chain termination method [13].

2.5. DNA sequence analysis

DNA PCR sequencing was performed by a reaction kit purchased from Epicenter (USA) based on primer extension. Sequence translations and analysis were performed using DNAsis software package (International Biotechnologies, USA). Sequences in GenBank were searched using the BLAST sequence similarity algorithm available at the National Center for Biotechnology Information internet website.

3. Results

3.1. Cloning of the shrimp FTZ-F1

We used a pair of degenerate primers (forward: DBD1, CCGGAATTCACNTGYGARGGNTGYAAR; reverse: DBD2, TCCTTGNGCATGCCCACTTCGADGCACTT) to amplify cDNA obtained from the eyestalk of individual shrimp. Several partial cDNAs encoding the DNA binding domain of the nuclear receptors were isolated. The shrimp specific partial cDNA was used as a probe to screen the eyestalk cDNA library. The library was constructed from a mixed population of shrimp eyestalks. About 200 000 phage clones were screened with a radioactively labeled cDNA probe and two potential clones with inserts of similar sizes were isolated. After subcloning and DNA sequence determination, they were identified as identical clones with slightly different sizes in the 5′ end. Sequence analysis revealed that the largest size (4.78 kb) of these cDNA clones contained an open reading frame (ORF) of 1635 bp encoding a predicted protein with 545 amino acid residues and molecular mass of 62 kDa (Fig. 1). The cDNA consisted of a short 5′ uncoding region and a long 3′ untranslated region. Several putative polyadenylated signals were also identified (Fig. 1).

3.2. Characterization of the shrimp FTZ-F1

Comparison of the translated cDNA sequence with those of known nuclear receptors showed the highest similarity to the fushi tarazu-factor 1 (FTZ-F1). By alignment of the shrimp receptor sequence with those of the FTZ-F1 from insects and vertebrates, a five-domain structure [14,15] that was characteristic of all members of the steroid receptor superfamily could be identified (Fig. 4). Domain A/B represented the variable N-terminal end of the protein. Region C represented the DNA binding domain. Region E was the ligand binding domain, region D was the variable linker region joining the DBD and LBD and there was a region F. The deduced amino acid sequence of the cDNA revealed a high percentage of overall amino acid identity with the FTZ-F1 nuclear receptor superfamily members (Fig. 2a). Although no reports for other full-length cDNA sequence in crustacean can be found, the shrimp FTZ-F1 shows a high homology to the *Artemia* FTZ-F1 (Fig. 2b). In addition to the conservation of these domains, other regions with characteristics of the FTZ-F1 member were also present in the shrimp FTZ-F1. For example, both the FTZ-F1 box and the AF2 core were present in the shrimp MeFTZ-F1 cDNA (Table 1). Evolutionary study using UPMG analysis software indicates that the shrimp FTZ-

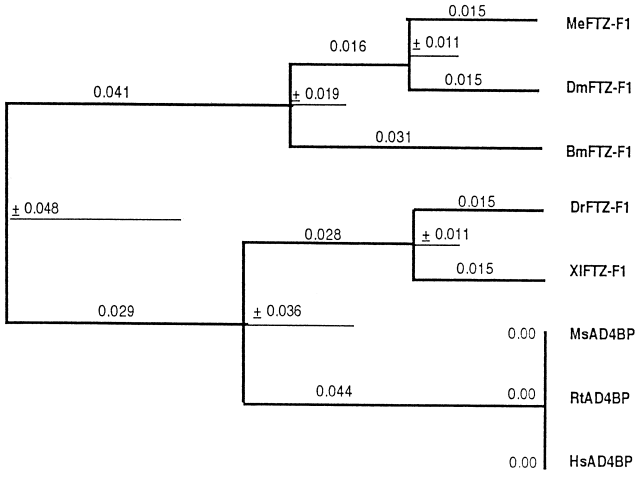

Fig. 3. Phylogenetic relationship of the shrimp FTZ-F1 with other members of the FTZ-F1 family. An UPGMA tree program was used to construct the tree and used to study the evolutionary relationship of different FTZ-F1 in the DNA binding domain (DBD) and the ligand binding domain (regions II and III combined).

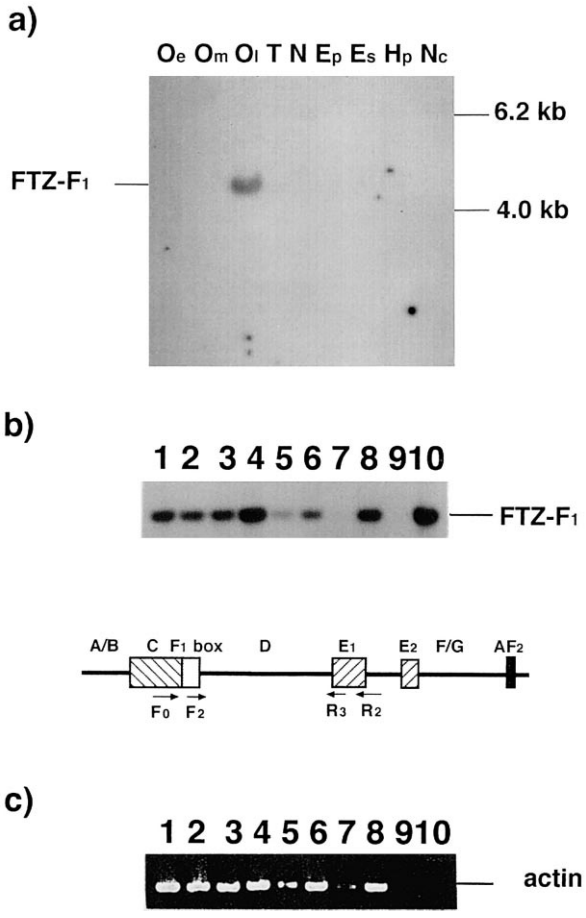

Fig. 4. (a) Tissue distribution of the shrimp FTZ-F1 by Northern blot analysis. Lanes contain total RNA extracted from ovary of early vitellogenic female (Oe), ovary from mid-vitellogenic stage (Om), ovary from late vitellogenic stage (Ol), testis (T), nerve cord (N), epidermis (Ep), eyestalks (Es), hepatopancreas (Hp) and nerve cord (Nc). (b) RT-PCR Southern blot analysis of RT-PCR products using an internal probe derived from F2/R3 PCR. Samples are: lane 1, immature ovary (oocyte diamter <20 μm); lane 2, early vitellogenic ovary (oocyte diameter <75 μm); lane 3, middle vitellogenic (oocyte diameter >150 μm); lane 4, mature ovary (oocyte diameter >250 μm); lane 5, pre-hatched nauplius; lane 6, 24 h nauplius; lane 7, 48 h nauplius; lane 8, testis; lane 9, negative control; lane 10, positive control. The arrows in the schematic sketch indicates primers located on the cDNA used in RT-PCR. Primers F0 and R2 were used in RT-PCR and primers F2 and R3 were used in PCR of FTZ-F1 cDNA to generate internal probe for Southern blot analysis. (c) RT-PCR expression of shrimp actin gene to demonstrate integrity of the RNAs and the relative amount of RNA used in PCR. Ethidium bromide staining of the PCR products was shown. Lanes 1–8 are same as in (b), lane 9 is control without template and lane 10 is the 1 kb DNA ladder.

F1 is most similar to the insect FTZ-F1 (Fig. 3). Compared to the vertebrate counterpart, the shrimp FTZ-F1 is most similar to the DrZF1 of the zebrafish.

3.3. Expression of MeFTZ-F1

By Northern blot analysis using total RNA, MeFTZ-F1 mRNA transcript was detected in the late stage of vitellogenic ovary (Fig. 4a). Other tissues such as the epidermis, eyestalks, hepatopancreas and nerve cord show no expression of MeFTZ-F1. To detect low levels of FTZ-F1 expression, RT-PCR was performed. The primers and conditions for RT-PCR were identical to the above description. With cDNA obtained from 10 μg of total RNA for RT-PCR, MeFTZ-F1 mRNA could be detected from the immature ovary and early larvae. However, the highest level of FTZ-F1 mRNA could be found in the late or mature ovary. Furthermore, FTZ-F1 mRNA could also be detected in the nauplius 24 and 48 h after hatching. Therefore, it appears that the shrimp FTZ-F1 may play a role in oocyte maturation and subsequent early larval development.

4. Discussion

A shrimp homologue for the FTZ-F1 has been cloned from the eyestalk cDNA library of the shrimp. The shrimp eyestalk consists of predominantly the optic nerve (including medulla terminalis of the X-organ-sinus gland complex), epidermis and muscles. However, further studies using techniques such as in situ hybridization could distinguish the exact location of FTZ-F1 expression. The shrimp FTZ-F1 shows the highest overall amino acid identity with that of the insects *Drosophila* and *Bombyx* (45%), followed by *Deniro* and *Xenopus* (34%), as well as the mouse and human (32%). Although molecular cloning of FTZ-F1 had also been reported in the brine shrimp, *Artemia salina*, only a partial nucleotide sequence of the DBD was reported [16]. Based on the published information, the partial DBDs of shrimp and brine shrimp shared only 72% amino acid identity. During the course of this study, we isolated another cDNA from the pleopod cDNA library. The DBD and LBD of this cDNA showed 92% and 70% amino acid sequence identity respectively as compared to the MeFTZ-F1 form (data not shown). Unlike the zebrafish which consists of 2–3 isoforms of ZFF-F1 with identical amino acid sequences in the DBD and LBD [17], the shrimp may consist of at least two distinctive forms of FTZ-F1. Although the phylogenetic analysis of the specific domains (DBD and LBD) of different FTZ-F1 indicates that the insect and crustacean are closely related to one another, the FTZ-F1 of *D. melanogaster* and the silkworm *B. mori* lack the AF2 core which is important for transactivation (Fig. 3). On the other hand, we speculate that both the FTZ-F1 of the shrimp and fruit fly may bind to the same response elements as the LBD and FTZ-F1 boxes of both groups are identical.

At present, the genes that are regulated by the shrimp FTZ-F1 have not yet been identified. However, in our study of the shrimp crustacean hyperglycemic hormone-like neuropeptide gene (CHH-like gene) [18], we have identified several potential binding sites for FTZ-F1 in the 5′ flanking region of the CHH-like neuropeptide gene (GenBank accession number AF109776). Removal of these sites reduced the promoter activity significantly (data unpublished). CHH-like neuropeptide has been implicated to play a role in the gonad maturation of the female lobster [19] and shrimp [18]. Since MeFTZ-F1 was isolated from the eyestalk cDNA library and the level of CHH-like expression is much higher in the reproductive females that undergo vitellogenesis, FTZ-F1 may play an important role in the regulation of CHH-like gene expression. Further study is required to establish the link between FTZ-F1 and CHH-like gene expression in shrimp. The presence of MeFTZ-F1 in the ovary and early larvae may represent maternal message deposited in the egg for embryonic development and early larval development. In *Drosophila*, FTZ-F1 is deposited in the egg during oogenesis [20] and it is an obliga-

tory co-factor for *ftz* gene expression [3]. Similarly in the silkworm, *B. mori*, BmFTZ-F1 mRNA could be detected in most of the tissues examined [21].

In summary, we have isolated and cloned a nuclear receptor from the eyestalk of the shrimp. Based on the amino acid sequence homology and the expression of the cDNA, it can be considered a potential nuclear receptor belonging to the FTZ-F1 subfamily. This study represents the first report on the cloning of the full-length cDNA for the FTZ-F1 homologue in crustacean. The high degree of amino acid identity in the conserved domains and the similarity in the distribution of mRNA in insects and shrimp suggest that the molting and reproductive events in these two groups of arthropods may be very similar.

Acknowledgements: This work was supported in part by a grant from the Hong Kong Government Research Grant Council (RGC 96M) and by an institutional grant from the HKU (CRCG) awarded to S.M.C.

References

[1] Ueda, H., Sun, G.C., Murata, T. and Hirose, S. (1992) Mol. Cell. Biol. 12, 5667–5672.
[2] Wong, M., Ramayya, M.S., Chrousos, G.P., Driggers, P.H. and Parker, K.L. (1996) J. Mol. Endocrinol. 17, 139–147.
[3] Yu, Y., Li, W., Su, K., Yussa, M., Han, W., Perrimon, N. and Pick, L. (1997) Nature 385, 552–555.
[4] Lavorgna, G., Karim, F.D., Thummel, C.S. and Wu, C. (1993) Proc. Natl. Acad. Sci. USA. 90, 3004–3008.
[5] Miller, W.L. (1988) Endocrin. Rev. 9, 295–318.
[6] Ohno, C.K., Ueda, H. and Petkovich, M. (1994) Mol. Cell. Biol. 14, 3166–3175.
[7] Lara, D.S., Rice, D.A. and Parker, K.L. (1992) Mol. Endocrinol. 6, 1249–1258.
[8] Ikeda, Y., Lala, D.S., Luo, X., Kim, E., Moisan, M.P. and Parker, K.L. (1993) Mol. Endocrinol. 7, 852–860.
[9] Durica, D.S. and Hopkins, P.M. (1996) Gene 171, 237–241.
[10] El Haj, A.J., Tamone, S.L., Peake, M., Reddy, P.S. and Chang, E.S. (1997) Gene 201, 127–135.
[11] Chan, S.-M. (1998) FEBS Lett. 436, 375–380.
[12] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[13] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[14] Green, S. and Chambon, P. (1988) Trends Genet. 4, 309–314.
[15] Evans, R.M. (1988) Science 240, 889–895.
[16] Escriva, H., Safi, R., Hanni, C., Langlois, M.C., Saumitou-Laprade, P., Stehelin, D., Capron, A.A., Pierce, R. and Laudet, V. (1997) Proc. Natl. Acad. Sci. USA 94, 6803–6808.
[17] Liu, D., Le Drean, Y., Ekker, M., Xiong, F. and Hew, C.L. (1997) Mol. Endocrinol. 11, 877–890.
[18] Gu, P.-L. and Chan, S.-M. (1998) FEBS Lett. 441, 397–403.
[19] de Kleijn, D.P., Janssen, K.P., Waddy, S.L., Hegeman, R., Lai, W.Y., Martens, G.J. and Van Herp, F. (1998) J. Endocrinol. 156, 291–298.
[20] Nomura, M., Bartsch, S., Nawata, H., Omura, T. and Morohashi, K. (1995) J. Biol. Chem. 270, 7453–7461.
[21] Sun, G.C., Hirose, S. and Ueda, H. (1994) Dev. Biol. 162, 426–437.