

Two Distinct *c-ski* cDNAs of Fish, Tilapia (*Oreochromis aurea*)

CHIU-JU HUANG, JER-YOUNG LIN, AND HUAI-JEN TSAI*

Institute of Fisheries Science, National Taiwan University, Taipei, Taiwan

ABSTRACT Two classes of tilapia *c-ski* cDNA (accession nos. AJ012011, AJ012012), designated as *tski1* and *tski2*, respectively encoded a 687 and a 714 AA protein and shared a 57% AA identity. Comparison with the Ski proteins of chickens, humans and *Xenopus*, tilapia TSki polypeptides shared a 60, 57, and 57% (TSki1) and 67, 63, and 61% (TSki2) AA identity, respectively. The most and the least abundant *c-ski* mRNAs are located in the brain and the skeletal muscle, respectively. Both *tski1* and *tski2* were widely expressed in the adult tissues examined, but *tski2* transcripts were at higher levels except in the ovary and oocytes: *tski1* transcripts were predominant in the ovary, whereas *tski2* transcripts were predominant in the testes. In the oocytes, the *tski1* mRNA was a maternally-inherited stockpile that subsequently was degraded, so that the expression ratio of *tski1* to *tski2* transcripts declined gradually as the fish developed from oocyte to 4-cm fry. *Mol. Reprod. Dev.* 54:223–231. © 1999 Wiley-Liss, Inc.

Key Words: cDNA; *c-ski*; fish; oncogene; Ski protein

INTRODUCTION

c-ski is proto-oncogene which encodes an 84 kDa nuclear protein (Nagase et al., 1990). The c-Ski protein stimulates proliferation, induces morphologic transformation, and promotes myogenic differentiation on cultured quail embryo fibroblasts (Stavnezer et al., 1981; Li et al., 1986; Stavnezer et al., 1986; Colmenares and Stavnezer, 1989; Colmenares et al., 1991). Transgenic mice show a distinctive muscular phenotype due to a hypertrophic effect on the formation of the fast muscle IIb fibers and concomitant decrease in body fat (Sutrave et al., 1990). c-Ski activates a muscle-specific enhancer resulting in the transcriptional regulation of muscle differentiation (Engert et al., 1995). *c-ski* is widely expressed in many tissues, although some studies that investigated skeletal muscle in *Xenopus* and chicken found no evidence of *c-ski* expression (Sutrave et al., 1990; Sleeman and Laskey, 1993). Namciu et al. (1994) reported *c-ski* expression at a high level in the skeletal muscle of mouse embryos at 12.5 days of gestation. Knockout mice die at birth and show a defective closure of the rostral neural tube and a decrease in skeletal muscle (Berk et al., 1997). Recently, Nicol and Stavnezer (1998) demonstrated that c-Ski acts as a repressor

rather than an activator, and depends on the cellular environment that allows c-Ski to bind to the GTC-TAGAC site. Ichikawa et al. (1997) also reported that c-Ski functions as suppressor of the myogenic gene in proliferating myoblasts and as a potentiator in post mitotic myotubes. The fact that c-Ski apparently plays many important roles suggests that its functionality as a transcriptional factor may be modified by interactions with various other transcriptional factors (Tarapore et al., 1997). It would therefore appear to be impossible to understand clearly and completely the biological properties of c-Ski without studying c-Ski in more diverse species. However, only the cDNA-derived *c-ski* gene of chicken (Sutrave and Hughes, 1989), human (Nomura et al., 1989), mouse (Namciu et al., 1995), *Xenopus* (Sleeman and Laskey, 1993) and axolotl (Ludolph et al., 1995) have been defined. To date, no such definition has been available for the cellular *c-ski* gene of the lowest vertebrates, fish.

As a step toward providing this information, we herein describe two classes of *c-ski* cDNAs of tilapia (*Oreochromis aurea*), one of the most popularly farmed fish species around the world. These data provide a wider base of knowledge on the primary structure and the gene expression of fish *c-ski*.

MATERIALS AND METHODS

RNA Isolation and RT-PCR (Reverse Transcription PCR)

RNAs were extracted from the whole bodies of tilapia fry (1.6 cm) using TRI reagent (Molecular Research Center). Polyadenylated RNA (0.2 µg) isolated with a poly(A) mRNA Isolation System (Promega) was reverse-transcribed using an oligo(dT)_{12–18} primer and SuperScript reverse transcriptase (BRL). The single strand cDNA products were subjected to PCR with primers 1'F (5'ATCTSTGCTTYGTGGTGGGMGGA3') and 1'R (5'ATCTSTGCTTYGTGGTGGGMGGA3'), which were designed from the predicted well-conserved sequences of exon 1 in chicken (Sutrave and Hughes,

Grant sponsor: National Science Council of the Republic of China; Grant number: NSC 87-2311-B002-028-824.

*Correspondence to: Huai-Jen Tsai, Institute of Fisheries Science, National Taiwan University, Taipei, Taiwan, 106.
E-mail: hjtsai@ccms.ntu.edu.tw

Received 11 March 1999; Accepted 18 June 1999

1989), human (Nomura et al., 1989) *Xenopus* (Sleeman and Laskey, 1993) *c-ski*. PCR was carried out for 34 cycles in a 50 μ l solution containing 500 ng of the oligo(dT)₁₂₋₁₈, 5 μ l of 10 \times PCR reaction buffer, 1 μ l 2.5 mM dNTP, 3 μ l 25 mM MgCl₂, 10 μ l of each amplifier and 2.5 U Taq polymerase (Promega). Each cycle consisted of 94°C, 1 min; 65°C, 1 min; and 72°C, 2 min. After the RT-PCR product was cloned and sequenced (Sanger et al., 1977), another RT-PCR was performed using primers E-1-5' (5'GAAACGGCTGTGTCTGCCG-CAGATTCT3') and P2R (5'AGGYTTCTCYTCYTGTT-TCAC3'). Primer E-1-5' was derived from the sequence of the first RT-PCR product using primers 1'F and 1'R. Primer P2R was the most well-conserved sequence of exon 6 of known *c-ski*.

Rapid Amplification of cDNA End (RACE)

For the 5' end, the first strand of cDNA was synthesized with 3 μ g of total RNA using SuperScript II (BRL) in the presence of 0.1 μ M of primer 5'RA1 (5'GGTCCGCTGCACCTGG3'). The cDNA was homotailed with poly(dA) at the 5' end using 18 U of terminal transferase Tdt (BM). The resulting tailed cDNA was then used to generate double stranded cDNA by amplification in the presence of 200 nM of each primer RAAP (5'GGC-CACGCGTGCAGTACTAC(T)₁₈3') and 5'RA2 (5'GTCCG-CACACCGAGTTGATCTG3'), 200 μ M dNTP, and 1 U Taq polymerase (Promega). The PCR reaction was run over 30 cycles of: 1 min at 94°C, 1 min at 64°C and 2 min at 72°C. The first PCR product were specifically amplified using a nested PCR performed with the RAUAP primer (5'GGCCACGCGTGCAGTACT3') and an internal 3' oligonucleotide 1-5'RA3 (5'GGGAGAAAT-CTCTGAGAACGCG3'; nucleotide 859-880) for *tski1* or 2-5'RA3 (5'GAAGTCCCGCAGCACCGTGT3'; nucleotide 709-687) for *tski2* under above PCR conditions. The procedure for 3' RACE was basically as same as that for 5' RACE except that (1) primer RAAP was used to synthesize the first strand of cDNA; (2) the RAUAP primer with either 5'CATGGCTCGGGCTCAGGAGAGG3' for *tski1* or 5'CGGTGTCAGCACCGCCCT3' for *tski2* were used to generate double strand cDNA; (3) the annealing temperature was 70°C; and (4) the RAUAP primer with either 5'CTTGACACCAAGGAG-GCGCG3' for *tski1* or 5'GAGGCAGGCGCTGGACAGC3' for *tski2* were used in a nested PCR amplification.

Deduced AA Sequence Analysis and Comparison

The deduced AA sequences of c-Ski from previously investigated species were retrieved from the GenBank (NCBI blast database search) for comparison. Alignment of the sequences was carried out by using the CLUSTAL 4 program (Higgins and Sharp, 1988) of DNASIS (Hitachi software). The weight matrix method (Bishop, 1994) based on the eukaryotic translation initiation context (Kozak, 1978) was used to analyze the initiation codon of TSKI1 polypeptide. Chou-Fasman (Chou and Fasman, 1978), Garnier-Osguthorpe-Robson (Garnier et al., 1978) and hydrophobic moment meth-

ods (Eisenberg et al., 1984) were used to analyze the secondary structure of TSKI polypeptides.

Genomic DNA Extraction and Southern Blot Analysis

Genomic DNA was extracted from red blood cells of tilapia (Sambrook et al., 1989). For the Southern blot analysis (Sambrook et al., 1989), 10 μ g of genomic DNA were digested respectively with *Bam*HI, *Eco*RI, *Pst*I, and *Sac*I, and hybridized to a [α -³²P]dCTP-labeled class-specific probe, which was obtained by digesting plasmid containing RT-PCR product of the *tski1* (1290 bp) or *tski2* (1452 bp) fragment with *Nco*I and *Sep*I, respectively. Hybridization was carried out before autoradiography on X-ray film (Kodak BioMax MS-1) for 16 hr at -70°C.

Specific Primers for Amplifying *tski1* and *tski2* From Genomic DNA

Class-specific primers were used to detect the *tski1* and *tski2* genes, respectively. For *tski1* detection, forward primer 5'GAGTGAGTGGGGAGACAG3' and reverse primer 1-5'RA3 were used, whereas *tski2* detection used forward primer 5'TCCGGGACTGCAAC-AAACT3' and reverse primer 2-5'RA3. PCR was carried out 2.5 U VioTaq DNA polymerase (Viogene), and consisted of 25 cycles at 94°C for 1 min, annealing at 58°C for *tski1* or 64°C for *tski2* for 1 min, and extension at 72°C for 1 min. The expected M_r of the generated PCR products were 688 and 305 bp for *tski1* and *tski2*, respectively.

RNase Protection Assay

The 1.3 and 1.4 kb PCR products containing cDNA sequences corresponding to nucleotide 831-2129 of *tski1* and nucleotide 659-2120 of *tski2* were inserted into the pGEM-T (Promega) and linearized at the *Xho*I and *Esp*I sites to generate two riboprobes of 517 and 340 bp, which were respectively designed to protect a 453 bp *tski1* mRNA fragment and a 244 bp *tski2* mRNA fragment. In vitro transcription for *tski1* and *tski2* were performed with 20 U of T7 and SP6 RNA polymerase (Promega), respectively, in the presence of 50 μ Ci [α -³²P]UTP. Total RNA samples (20 μ g) were extracted from oocytes, stage 15 embryos (freshly hatched embryos), stage 23 larvae (complete yolk-absorption), fry (1-, 2-, and 4-cm body length) and various tissues (brain, eye, gill, heart, intestine, skeletal muscle, stomach, ovary, and testis) using TriZol reagent, and then hybridized overnight at 42°C with 10 μ g of [α -³²P]UTP-labeled probe with a 5 \times 10⁸ cpm/ μ g specific activity. RNA-RNA hybrids were digested with RNase A (0.05 U) and T1 (20 U) (Ambion) and analyzed in a 5% (w/v) polyacrylamide gel. The dried gel was exposed to X-ray film for 2 hr at -70°C. For normalization purposes, tilapia β -actin mRNA hybridized with a β -actin specific riboprobe served as an internal control. RT-PCR was used to amplify a DNA fragment from total RNA of tilapia larva by using oligo(dA)₁₂₋₁₈ primer, forward primer 5'TGCGGTATCCATGAGACCAC3' (nucleotide

2883–2903) and reverse primer 5'GAAGCATTGCG-GTGGACGA3' (nucleotide 3295–3275). The last two primers were corresponding to exons 5 and 6 of the common carp β -actin genomic DNA (Liu et al., 1990). The 312 bp RT-PCR product was subcloned into the pGEM-T vector and linearized at the *DdeI* site to generate the riboprobe using T7 RNA polymerase. For quantifying the varying levels of *ski* expression in the tissues, the image of RNase protection assay data was taken from Kodak DC120 Zoom Digital Camera. Kodak Digital Science 1D Image Analysis Software was used to analyze the mean intensity. The total sum of intensity of two protection fragments resulting from β -actin probe were counted to normalize.

RESULTS

Two Classes of Products Produced by RT-PCR

A 350-bp DNA fragment was amplified after RT-PCR using primers 1'F and 1'R. After subcloning and sequencing, a forward primer (E-1–5') specific for exon 1 of tilapia *c-ski* was designed to perform another RT-PCR with a reverse primer P2R. Two DNA fragments around 1.3 and 1.5 kb were shown on the gel. The identities of the nucleotide sequence and deduced AA sequence between these two clones were 76.2 and 53.5%, respectively, and we therefore named the 1.3 kb clone *tski1* and the 1.5 kb clone *tski2*.

Primary Polynucleotide Sequences and Deduced AA of Tilapia c-Ski DNAs

Figure 1 depicts the full length polynucleotide sequences and deduced AA sequences of tilapia *tski1* and *tski2* cDNA. *tski1* cDNA contains a single open reading frame (ORF) of 2061 base (nucleotide 541–2601) encoding 687 AA residues. The stop codon TAA is followed by an untranslated region of 572 nucleotides at the 3' end, in which two potential instability sequences (ATTTA) (Sleeman and Laskey, 1993) and a potential polyadenylation signal are included. *tski2* cDNA contains a single ORF of 2057 (nucleotide 370–2511) bases which encode 714 AA residues. At the 3' end, one potential instability sequence and two potential polyadenylation signals are included.

Comparison of the Deduced AA Sequence of Tilapia TSki1 and TSki2 With c-Ski and SnoN of Some Other Known Species

In Figure 2, the deduced AA sequences of tilapia TSki1 and TSki2 are compared with those of the Ski proteins of other species. The shared identity between tilapia TSki1 and TSki2 is 57.2%. Tilapia TSki1 shares 59.7% identity with chicken *c-Ski*, 56.9% with human, and 57.1% with *Xenopus* Ski proteins. The corresponding shared identities for tilapia TSki2 are respectively 66.5, 62.7, and 61.3%. *ski*-related cDNA, sno, has been reported in humans (Nomura et al., 1989) and chickens (Boyer et al., 1993). However, the AA identity between human SnoN and tilapia TSki1 and TSki2 was as low as 38 and 39%, respectively. The comparison data also reveal that highly conserved domains are located at the

termini, exon 1 (AA 1 to 298 of TSki1) and exons 6 and 7 (AA 489 to 655 of TSki1). In contrast, the middle segment of tilapia TSki protein is a highly variable region. Several characteristics of Ski protein found in humans, chickens, and *Xenopus* are also conserved in tilapia TSki polypeptides. Both TSki1 and TSki2 contained heptad repeats (Nagase et al., 1993), a 25-mer repeat element and an alternating hydrophobic-basic motif at the C-terminus (Sleeman and Laskey, 1993).

Two *tski* Genes Detected by Southern Blot Analysis and PCR

Genomic DNA was isolated from tilapia blood, digested with four restriction enzymes, and analyzed using Southern blot hybridization with probes for *tski1* and *tski2*, respectively. The resulting restriction fragment analysis profiles were different (Fig. 3A), and only two common bands were found: an 8 kb *Bam*HI-band and a 27 kb *Eco*RI-band. Genomic DNA was extracted from individual fish and served as a template for PCR amplification using class-specific primers. PCR product around 700 bp was generated using the *tski1*-specific primers, while the *tski2*-specific primers generated a product of around 300 bp (Fig. 3B).

Expression of *tski1* and *tski2* Genes in Various Tissues and During Development

Tilapia *tski* genes were expressed in all tissues examined, with the brain and the skeletal muscle containing the most and the least abundant *c-ski* mRNAs of both classes, respectively (Fig. 4A). In general, tissues such as the heart, intestine, and brain had more *tski2* transcript than *tski1*, and in the testes, where *tski2* gene was also expressed dominantly, expression of the *tski1* gene was extremely rare. In contrast, in the ovary, *tski1* transcript was much more abundant in expression than the *tski2* transcript. Figure 4B shows that during development the expression of *tski1* appeared to undergo a stepwise decline as embryos developed, with a rapid decline over the hatching stage and relatively little change otherwise. However, the expression of *tski2* showed exclusively a high peak at the hatching stage. Consequently, the quantitative ratio of *tski1* to *tski2* decreased as the tilapia grew to fry of 4-cm body length.

DISCUSSION

Northern blot analysis of *c-ski* expression showed that there were two major mRNA species in chickens (Li et al., 1986) and humans (Nomura et al., 1989). This evidence in turn suggests that there might be more than one species of the *c-ski* gene, resulting either from an alternative splice and/or from two independent genes. Three classes of chicken *c-ski* cDNA have been found that the multiple cDNAs were derived from alternative splicing (Sutrave and Hughes, 1989). Sleeman and Laskey (1993) also detected two distinct fragments on Southern blot analysis of *Xenopus* genomic DNA at high stringency, but when the two respective genes were cloned out and sequenced, the

A ATG CTG CAG CTG CCC TGT AGC AGA GGG AGG AAC AGC TGG CTG AGG GAA GAG CCA GCC CGG GGA GTC AGT CAC ATT CCT GGC TGG GAT TCA ATG ACT GAA CCA GAC ACA 108

AAC AAA CAG AGA GGG CGA GAG AGA GAG AGA GGG AGG GAG GGA GAG AGA GAG AGA GAG AGA GAG CGA CTC ACA CAT TGA GTG AGT GGG GAG ACA GAC TGA ATG TAC GCC 216

AGA CTG AGC AGA CGA GGC ATA TTT CGG GCG CAC ATT AGG AGT GCG TAA AAC CGC AGG AGC ATA GAA ACA TAG TAC GGC CTG ATT GTA TGT TGG ATA ATA CGA CTT TGG 324

AGA CGC TTA GTT GTC GAC TTT AAT TGC GTG TGC TTA CAG GAA CTG CCC TAT GAT TTC AGC AGC TCT CAG TTT TCT GAA TGG ACG CGC TGG GAA GCT TCT CCG GTG ACA 4320

TGG AAG TGC GCG TTA TTG AGC GTG TAG CCG GAT GCT GAG GAC GCC CGA CCC GAT AAG TCG ATT TCA GCT CGC ATT TTG AGT TTT GAA CTC TCT TCT CTA ACC GAC AAC 540

ATG GAG GGC ACG AGC TTC CAG CCC CAT CCT GGA CTT CAG CAA ACT CTG CAG CAG TTT AAT CTG AGT TCC AGG CGC TCC CTC GGT GGA CCG GCG GCG TTC TCC GCT CGA 648
MET Glu Gly Thr Ser Phe Gln Pro His Pro Gly Leu Gln Gln Thr Leu Gln Gln Phe Asn Leu Ser Ser Arg Arg Ser Leu Gly Gly Pro Ala Ala Phe Ser Ala Arg 36

TGG CAC CAG GAC TCA CTC TTC GGG AAA GAC GGC AAA TCC GTC GAG ATG ATG CTC ACC CTG CCG CCT CAG ACA CCT CCA GTG ATG TCC GGT CCA CTT TTC ATC CCG TCC 756
Trp His Gln Asp Ser Leu Phe Gly Lys Asp Gly Lys Ser Val Glu Met Met Leu Thr Leu Pro Pro Gln Thr Pro Pro Val Met Ser Gly Pro Leu Phe Ile Pro Ser 72

GAC CGC TCC ACC GAG AGG TGC GAG ACG GTG CTG GAG CGG GAA CCT ATC TCC TGC TTC GTC GTC GGT GGT GAC AAA CGA CTG TGC CTA CCG CAG ATT CTC AAC AGC GTT 864
Asp Arg Ser Thr Glu Arg Cys Glu Thr Val Leu Glu Arg Glu Pro Ile Ser Cys Phe Val Val Gly Gly Asp Lys Arg Leu Cys Leu Pro Gln Ile Leu Asn Ser Val 108

CTC AGA GAT TTC TCC CTC CAG CAG ATC AAC TCG GTG TGC GAC GAT CTC CAC ATC TAT TGC TCC AGG TGC ACA GCG GAC CAG CTA GAG ATC CTC AAA GTG GTG GGT ATC 972
Leu Arg Asp Phe Ser Leu Gln Gln Ile Asn Ser Val Cys Asp Asp Leu His Ile Tyr Cys Ser Arg Cys Thr Ala Asp Gln Leu Glu Ile Leu Lys Val Val Gly Ile 144

CTG CCC TTC TCG GCA CCG TCC TGT GGG CTT ATC ACT CAG ACG GAT GCT GAA CGG CTT TGC AAC GCG CTC ATC TAT GGC GGT ACT TAC CCT CCT CAT TGC AAC AAG GAG 1080
Leu Pro Phe Ser Ala Pro Ser Cys Gly Leu Ile Thr Gln Thr Asp Ala Glu Arg Leu Cys Asn Ala Leu Ile Tyr Gly Gly Thr Tyr Pro Pro His Cys Asn Lys Glu 180

TCG GGC TCT CTG GAG TTG GAG CGA ACC GAG AAA AGC TTC AAA GTG TAT CAC GAA TGT TTT GGC CGG TGT AAA GGC CTG TTT GTC CCG GAA CTG TAC ACC TCT CCG GGT 1188
Ser Gly Ser Leu Glu Leu Glu Arg Thr Glu Lys Ser Phe Lys Val Tyr His Glu Cys Phe Gly Arg Cys Lys Gly Leu Phe Val Pro Glu Leu Tyr Thr Ser Pro Gly 216

GCC GCC TGC ATC CAG TGC ATG GAC TGC AGA CTC ATG TAC CCA CCT CAC AAG TTT GTT GTC CAC AGT CAC AAG AGA CTA GAG AAC CGG ACA GTC CAC TGG GGC TTT GAT 1296
Ala Ala Cys Ile Gln Cys Met Asp Cys Arg Leu Met Tyr Pro Pro His Lys Phe Val Val His Ser His Lys Arg Leu Glu Asn Arg Thr Val His Trp Gly Phe Asp 252

TCT GCC AAC TGG CGG GCT TAT GTG CTC TTA GAC CCG GAC TAC ACC GGG AAA GAG GAG AAG AGT CAC CTG GAG AAG CTG CTC AAA GAG TTA AAA GGA AAA TTT GAT CTG 1404
Ser Ala Asn Trp Arg Ala Tyr Val Leu Leu Asp Pro Asp Tyr Thr Gly Lys Glu Glu Lys Ser His Leu Glu Lys Leu Leu Lys Glu Leu Lys Gly Lys Phe Asp Leu 288

ACG GGA AAA CTG TCC AGT AAA TCT TGC AGA TCT CCC AGC CCC ATC CGA GCC AAG AGG TCC AAA TTC GAC AAA TTG CAG TCA GCT GAC AAA GAC AGG AAA CCC GAC TGG 1512
Thr Gly Lys Leu Ser Ser Lys Ser Cys Arg Ser Pro Ser Pro Ile Arg Ala Lys Arg Ser Lys Phe Asp Lys Leu Gln Ser Ala Asp Lys Asp Arg Lys Pro Asp Trp 324

CTC CAG TCA CTG TCA AAG TCT GCA CAC AAG GAT CTG AAA CAG GTC CAA CTA AAA CAG AGG CCC TCT GCT TTC CGC CCC TGG TCT CCT AAG CCA GCA GAA AAA GTG AAA 1620
Leu Gln Ser Leu Ser Lys Ser Ala His Lys Asp Leu Lys Gln Val Gln Leu Lys Gln Arg Pro Ser Ala Phe Arg Pro Trp Ser Pro Lys Pro Ala Glu Lys Val Lys 360

CCG GCC GCC AAG AAT GAG GTG GAG AGG TCG TGC TCG AGG AAT CAG GAG ACT CAC AAT TTG GCA TTT GCC CAC CTG GCC CCT GCG GTC CAT GCC AAG GAC AGC AAC ACT 1728
Pro Ala Ala Lys Asn Glu Val Glu Arg Ser Cys Ser Arg Asn Gln Glu Thr His Asn Leu Ala Phe Ala His Leu Ala Pro Ala Val His Ala Lys Asp Ser Asn Thr 396

CCT GAC AGG GGG ACA GCT GCC ATT TCC GTG CAG GAG CTG CAT AAT GGT GAC GCA CAG CCT ACA ACA AAG CCG GCC CAC TCC AGC AAC ACC AAC CGA GCC GAA GAC ATG 1836
Pro Asp Arg Gly Thr Ala Ala Ile Ser Val Gln Glu Leu His Asn Gly Asp Ala Gln Pro Thr Thr Lys Pro Ala His Ser Ser Asn Thr Asn Arg Ala Glu Asp Met 432

GAC ACA GAT GGA GAG ATT GAT GTG GAT GAC TGT GAT GAT GGT CCA GTG CAG TCT TCC TCC CTG GCT TCA CCT CCA TCT GCC TGC ACC AGT GTG TCT CAG ACT CTG ACT 1944
Asp Thr Asp Gly Glu Ile Asp Val Asp Asp Cys Asp Asp Gly Pro Val Gln Ser Ser Ser Leu Ala Ser Pro Pro Ser Ala Cys Thr Ser Val Ser Gln Thr Leu Thr 468

CCT CAG AGC ATG GCT CCG GCT CAG GAG AGG CCT TCC TGG CTG CCA GGG ACT GTT TGC CCA GAG ATG GGC ACC ATG AGA CAG ATG CTG TAT GCT GGT CTT GAC ACC AAG 2052
Pro Gln Ser Met Ala Arg Ala Gln Glu Arg Pro Ser Trp Leu Pro Gly Thr Val Cys Pro Glu Met Gly Thr Met Arg Gln Met Leu Tyr Ala Gly Leu Asp Thr Lys 504

GAG GCG CGG GAA AAA CTC CTG CAG GAG ATT GTC AGG ATG AGA GTG AAG CAG GAG GAG AAG CTG GCA GCT GCT CTT CAA GCT AAA CGC AGC CTT CAG CAG GAG CTG GAG 2160
Glu Ala Arg Glu Lys Leu Leu Gln Glu Ile Val Arg Met Arg Val Lys Gln Glu Glu Lys Leu Ala Ala Ala Leu Gln Ala Lys Arg Ser Leu Gln Gln Glu Leu Glu 540

TTT GTG AGA GTG GCT AAG AAA GGG GGT CTT GCG GAG GCC ATC GAG GCC AAG CGC AAC CTG CGA AAG GAG ATC GAA CGC CTT CCG GTG GAC TGG GAG AGG AAG ATG AGG 2268
Phe Val Arg Val Ala Lys Lys Gly Arg Leu Arg Glu Ala Ile Glu Ala Lys Arg Asn Leu Arg Lys Glu Ile Glu Arg Leu Arg Val Asp Trp Glu Arg Lys Met Arg 576

GAA GCG GAG GAG TCT TGT GGG CGG CTG AAG AGA GAG CTG GAG AGA GAG AGA CAG CTG CGA GTT TGC GAT AAA GGC TGC GAA GCC GAA CGC CTC CCG GTC AAG TAC TCT 2376
Glu Ala Glu Glu Ser Cys Gly Arg Leu Lys Arg Glu Leu Glu Arg Glu Arg Gln Leu Arg Val Cys Asp Lys Gly Cys Glu Ala Glu Arg Leu Arg Val Lys Tyr Ser 612

ACT CAG ATC GAA GAG CTT CAT GTG CAG CTG CAA CAG GCT GAA GCC GAT CGC GAG CAG CTG AGG CCG GAG CTG CAG CAG GAG AGA GAA GCT CGA CAG ACG CTG GAA AGC 2484
Thr Gln Ile Glu Glu Leu His Val Gln Leu Gln Gln Ala Glu Ala Asp Arg Glu Gln Leu Arg Arg Glu Leu Gln Gln Glu Arg Glu Ala Arg Gln Thr Leu Glu Ser 648

GTC GTC AAA GAC CTG CAA ACC CAG CTG GCC CTG CAG GCC GGC AGC ATC CTT CCT GGA GAA TGC AAG GAC ACG AGC ACA GAG GCG CAC AGA CAG ACC GCA CAA CCC ACT 2592
Val Val Lys Asp Leu Gln Thr Gln Leu Ala Leu Gln Ala Gly Ser Ile Leu Pro Gly Glu Cys Lys Asp Thr Ser Thr Glu Ala His Arg Gln Thr Ala Gln Pro Thr 684

AAT GGA TCC TAA AGT TGC ACA TTT AAA ACG AGA CAG AAA GAT TCA GTA AGA GGA GCA TGT TTA AGG GGA AAA AGA AAC AAA GAG TGT TTT TGG TAG GGG GAA AAA AAT 2700
Asn Gly Ser --- 687

CAG TCA TAT ACA ATA TGA TAC AGT ATA CGT CAT TAT TAT TAT GAC ATT ATG AAT TTT ATC AAA TGA ACT GGA AAG CAG CAC GAT GTC CGA GCA TGT AAT AGA GCG TGA 2808

AAC TGC GAG AGA CTT TCT TCA GCG ATG GAA AAA CAC ATC TTT CCG GTG ACG TCA GCG CCT GCG GTC AGA CAC CCG ACG CCG TGC AAG GAT TTC ATT TCA GTA TAA GCT 2916

ATT CTA AGA TGT GCA AAT ATA GCT ACA TAA CAC TAC TTG AAT ATG AAG GGA AGA AAC ACG TTG CCG CTG GTG CCG AGT TGA TGA AAC TCT TGT TGC CAC ATA TAC TGG 3024

TAC ACA CAA GGG TGT CAT TGC AGA CCA CTG GAA ACA AGA GAC ACG AGT TTA CCA CTA TTT ACT TTT AAT ATC TGC TTG CTT TTT ATA ACT GTG CAG AGG ATG TCC TTT 3132

TTA AGA AAT TCC CTT GCG CAG CAC TGT TAT TCC ACA AGC AC 3176

Fig. 1. Polynucleotide sequences of tilapia *tski1* (A) and *tski2* (B) cDNA and their deduced AA. The nucleotides are numbered beginning with the first nucleotide at the 5'-end, and where it appears, the second smaller number indicates the order of the AA position. The

potential initiation codons in the 5' region are boxed. The potential instability sequences (rapid decay sequences, Sleeman and Laskey, 1993) underscored with wavy lines and the possible polyadenylation signals are underlined.

B AGA CTT CTC TCC GCG CAG TCA GAG TGA ACT GCG GGG GCA GAC GGG GTG GAT GTT TCG TCC TGG ACA GCG GAA AGA ACC TCA AAT AAG GAT TTT TAC ATT TCG TTT TCA 108

CGC CTG ACA AAC GGA ACT TTT CGC CTC GGT GGA GTT TGG AAG CCG TTT TTT CAA ATA TGG ACA GTT GGC GCA CTG TTA CAG GGA GCG ACT GGT CGG TGT TGA CTC ACA 216

CAC GGG AAT AAA ACC GCA GCT GGG ACT TTT TTC CTG GAT GGT TGA TGA GGC GTT TGG GAC ACG GTG GAT TTA CCA GGC TGA AGG ACG CCT AOC TGG TTG TGA TTC CCG 324

GTC GAC CGT AGG AAG GAT TTA CGC ATC TAA AGT TGC AGG AAA AGC ATG GAG ACT GTA AGT CCG CCG AGC TTC CAG CCT CAT CCG GGA CTG CAA CAA ACT CTC AAG CAG 432
MET Glu Thr Val Ser Ara Pro Ser Phe Gln Pro His Pro Gly Leu Gln Gln Thr Leu Lys Gln 21

TTT CAC CTC AGC TCT ATG AGC TCG TTG GGT GGA CCC GCC GCT TTC TCT GCT CGC TGG CAA CAT GAG CTG CTC TTT AAG AAG GAC GGG AAG GAG CCC GAG CCG GTT CTG 540
Phe His Leu Ser Ser Met Ser Ser Leu Gly Cly Pro Ala Ala Phe Ser Ala Arg Trp Gln His Glu Leu Leu Phe Lys Lys Asp Gly Lys Glu Pro Glu Pro Val Leu 57

CAG CAT CTG CCG CCG CCC GTG ATG CCC GGC CCG CTC TTC GTC CCG TGG GAC CGC TCC ACG GAG AGG TGC GAG ACG GTC CTG GAG GGG GAG ACC ATC TCC TGC TTT GTG 648
Gln His Leu Pro Pro Pro Val Met Pro Gly Pro Leu Phe Val Pro Ser Asp Arg Ser Thr Glu Arg Cys Glu Thr Val Leu Glu Gly Glu Thr Ile Ser Cys Phe Val 93

GTC GGC GGG GAG AAA CCG CTG TGT CTG CCG CAG ATT CTC AAC ACG GTG CTG CCG GAC TTC ACC CTG CAG CAG ATC AAC TCC GTG TGC GAC GAG CTG CAC ATC TAC TGC 756
Val Gly Gly Clu Lys Arg Leu Cys Leu Pro Gln Ile Leu Asn Thr Val Leu Arg Asp Phe Thr Leu Gln Gln Ile Asn Ser Val Cys Asp Glu Leu His Ile Tyr Cys 129

TCC CGC TGC ACG GCC GAC CAG CTG GAA ATC CTT AAA GTC ATG GGG ATC CTG CCC TTC TCG GCG CCG TCC TGC GGT CTC ATC ACC AAG ACG GAC GCA GAG CCG CTC TGC 864
Ser Arg Cys Thr Ala Asp Gln Leu Glu Ile Leu Lys Val Met Gly Ile Leu Pro Phe Ser Ala Pro Ser Cys Cly Leu Ile Thr Lys Thr Asp Ala Glu Arg Leu Cys 165

AAC GCC CTG ATC TAC GGA GGC GCT TAC CCG CCG CGC TGC AAG AAG GAG ATG AAC GGA GGC TCG CTG GAG CTG CAG TTC ACC GAC AGG AGC TTC AAA GTC TAT CAC GAG 972
Asn Ala Leu Ile Tyr Gly Gly Ala Tyr Pro Pro Arg Cys Lys Lys Glu Met Asn Gly Gly Ser Leu Glu Leu Gln Phe Thr Asp Arg Ser Phe Lys Val Tyr His Gln 201

TGC TTC GGC AAG TGC AAA GGC TTG TTC GTG CCG GAG CTG TAC ACC AGC CCA AAC GCA GCG TGC ATC CAG TGC ATG GAC TGC AGA CTA ATG TAC CCG ACC CAC AAG TTT 1080
Cys Phe Gly Lys Cys Lys Gly Leu Phe Val Pro Glu Leu Tyr Thr Ser Pro Asn Ala Ala Cys Ile Gln Cys Met Asp Cys Arg Leu Met Tyr Pro Thr His Lys Phe 237

GTG GTG CAC GGC CAC AAG GCG CAG GAG AAC AGG ACT TGC CAC TGG GGC TTC GAC TCG GCC AAC TGG AGG GCG TAC ATC CTC CTG GGC CAG GAT TAC ACG GAG AAA GAG 1188
Val Val His Gly His Lys Ala Gln Glu Asn Arg Thr Cys His Trp Gly Phe Asp Ser Ala Asn Trp Arg Ala Tyr Ile Leu Leu Gly Gln Asp Tyr Thr Glu Lys Glu 309

GAG AAG GCC CGC CTG GAG CTG TTC CTG GAC GAG ATT AAG GAG AAA TTC GAC TTC GCC AAC AAG TAC AAG AGG AAA GCA TCA TCC AAG GTG TCC GAT CCC ATC CCG GTG 1296
Glu Lys Ala Arg Leu Glu Leu Phe Leu Asp Glu Ile Lys Glu Lys Phe Asp Phe Ala Asn Lys Tyr Lys Arg Lys Ala Ser Ser Lys Val Ser Asp Pro Ile Pro Val 309

AAA AAG TCT AAA CAT GAA GAC CTC TCA TCA CAA ACT CCG CTG GCC GAC AGA GAG AAG CAG CAC GAC TGG CTG CAG TCC CTG TCC ACT CCC AAT AAG GGG CTG AAC TGC 1404
Lys Lys Ser Lys His Glu Asp Leu Ser Ser Gln Thr Pro Leu Ala Asp Arg Glu Lys Gln His Asp Trp Leu Gln Ser Leu Ser Thr Pro Asn Lys Gly Leu Asn Cys 345

ATT CAG TCC AGG CAG AAG CCC TCA GCT TTC AGA CCC TGG TCC CCC CAC ATC TCT GCG GGT GAT AAG GAG CCC TCC AGT GAC CCG CTG GCC CTG CTG AGA GAC GGC TTC 1512
Ile Gln Ser Arg Gln Lys Pro Ser Ala Phe Arg Pro Trp Ser Pro His Ile Ser Ala Gly Asp Lys Glu Pro Ser Ser Asp Pro Leu Ala Leu Leu Arg Asp Gly Phe 381

TAC AAT TAC AAG AGC CTA GAG AAG TGG CCC CCA ACG TCG CCC TCA CTC CGC TGC CAC TAC GCA AGG TAC GGC CTC TGT CCC CAT CGA TAC CCA CCA CTT CCT ACC CAA 1620
Tyr Asn Tyr Lys Ser Leu Glu Lys Trp Pro Pro Thr Ser Pro Ser Leu Arg Cys His Tyr Ala Arg Tyr Gly Leu Cys Pro His Arg Tyr Pro Pro Leu Pro Thr Gln 417

GTG GAA CCG AAC CTC CAG GTG AGG GCG CAC ACC CCA ACA CCA GAC CTC GCA AGA GGA GAG CTA CCG AGG AGC TCC AGT CCC CAG CAC ATG AAA TCC CTG GGC CTG CTT 1728
Val Glu Ala Asn Leu Gln Val Arg Ala His Thr Pro Thr Pro Asp Leu Ala Arg Gly Glu Leu Pro Arg Ser Ser Ser Pro Gln His Met Lys Ser Len Gly Leu Leu 453

CCT CAG CAG CCA GCA AAC CTC CGC TGC CTC AGG ATG ACA AGG ACT CCG AGG TGG AGA TCG AGG TGG AGA GCC GCG ACG AAG CCG CTT CGT CCA TGT CCT CCC TCT GGT 1836
Pro Gln Gln Pro Ala Asn Leu Arg Cys Len Are Met Thr Arg Thr Pro Arg Trp Arg Ser Arg Trp Arg Ala Ala Thr Lys Pro Leu Arg Pro Cys Pro Pro Ser Gly 489

CGC CAT CTT TTA CCT CAT CCA GCA GCT CAG CCA AAG ACT TCA GCA TCG CCA GGT GTC CAA GGT CCC ATC TGC ACT ATC ATG TCA GCC GAA AGC ACG GCA CCC GCA GCC 1944
Arg His Leu Leu Pro His Pro Ala Ala Gln Pro Lys Thr Ser Ala Ser Pro Gly Val Gln Gly Pro Ile Cys Thr Ile Met Ser Ala Glu Ser Thr Ala Pro Ala Ala 525

ACA TCG GTG TCA GCA CCA GCC CCT GTC ACC TCT TCA GAG TCG GGC CTG GAG TCA GAG CTG GAG AGC CTG AGG CAG GCG CTG GAC AGC GGG CTG GAT TCC AAG GAG TCG 2052
Thr Ser Val Ser Ala Pro Ala Pro Val Thr Ser Ser Glu Ser Gly Leu Glu Ser Glu Leu Glu Ser Leu Arg Gln Ala Leu Asp Ser Gly Leu Asp Ser Lys Glu Ser 561

AAG GAG AAG TTT CTG CAC GAA ATA GTT AAG ATG AGG GTG AAG CAG GAG GAG AAG TTG GGA TCG GCC CTG CAG GCG AAA CGA AGC CTT CAG CAG GAG TTG GAG TTC TTG 2160
Lys Glu Lys Phe Leu His Glu Ile Val Lys Met Arg Val Lys Gln Glu Glu Lys Leu Gly Ser Ala Leu Gln Ala Lys Arg Ser Leu Gln Gln Glu Leu Glu Phe Leu 597

CGG GTG GCC AAG AAG GAG AAG CTG CCG GAG CCG ACC GAG GCA AAG CCG AAC CTC AGG AAG GAA ATT GAG CGT CTG CGA GCT GAG AGT GAG AAG AAG ATG AAG GAA GCC 2268
Arg Val Ala Lys Lys Glu Lys Leu Arg Glu Ala Thr Glu Ala Lys Arg Asn Leu Arg Lys Glu Ile Glu Arg Leu Arg Ala Glu Ser Glu Lys Lys Met Lys Glu Ala 633

AAC GAG TCA AGG ATC CGC CTG AAG CCG GAG CTA GAG CAG GCC CCG CAG CTG AGG GTC TGT GAT AAA GGC TGT GAG GCT GGG CGA CTC CGT GCA AAG TAT TCT GCA CAG 2376
Asn Glu Ser Arg Ile Arg Leu Lys Arg Glu Leu Glu Gln Ala Arg Gln Leu Arg Val Cys Asp Lys Gly Cys Glu Ala Gly Arg Leu Arg Ala Lys Tyr Ser Ala Gln 669

ATC GAG GAC CTC CAG ATG AAG CTG CAG CAT GCC GAG GCC GAC CGT GAG CAG CTT CGA GCC GAC TTG CTG CAT GAG CGA GAA GCT CCG GAG CAC CTG GAG AGG ATG GTG 2484
Ile Glu Asp Leu Gln Met Lys Leu Gln His Ala Glu Ala Asp Arg Glu Gln Leu Arg Ala Asp Leu Leu His Glu Arg Glu Ala Arg Glu His Leu Glu Arg Met Val 705

AAG GAG CTA CAG CAG CAA ATC AAG CAC TAA CCT AGA CAG GAC ACC CAA GGA CAT TCC CAT TGG TGA CAA CTA ACC AAG CAG CCA GCC TTA CAG CTG AAC CTT TTC CAT 2592
Lys Glu Leu Gln Gln Gln Ile Lys His --- 714

TTG CCA CCC TGA TCG CCA CCC CAC GAG TCC CAC AGG TCG CCC ACA TCA CAC CAC CGT TCC TGT GCT TGC CCC TCA CGC TAC TGA ACA CAA AAC GAC TTC AAA TGG AGA 2700

CCC AGA GAA GAA TGG ACA AGC GTT GTC CTT GTG CAT GAA AGG TCT AGT GGT ACC GTA GAA GAA TGA CTG GAA GCA TCC GTT TTC TGA AGC ATT GAC TTT TGC ACC TTC 2808

TGA ACT CGT AGA ACT GAG AAC TGG AGC ACA GTG AAC GGC TCT GTA GTG CTG AGT GCG GCT GTG CCT GGA TTT CCT CCA CAA GGA CAC AAG GAG AGG CTG TAG GAC TGT 2916

GAT ATA AAC TGA TAA CGA TAT TAC ATC AAC GTG TGC ACC TGC TGG GAG GTG TTA CGA TGA CAT GGC CAG GAG GGA AGG CCT CCA CCA GCG CCT CCT GCC ACG ATG AAC 3024

CTG TAG GAG TCT GAA GAC ATT ATA AGC TAT TTA AGA GTT ACA TAG CGA TAT GTA GAG TAC ACA AAG AGC GTT TTT CTG TGC TAT TGG AGT GAA AAA TGA ACA AAT GGA 3132

CTT CCC ACT GTG CTG CGA GGG TTT GTC GGT TTG TCA GGA AGC CAT CCC CGC CCA TCG GGC CGA CTC CCC AAA CCA CAA TTG GAT ATT AAA ACA TGG CAA ATG ATT TTA 3240

ACA TCA AAG ATT AGG TTT GGC TTT GTG TTC AAC AGG CAC ATT GTC CAT CTC GAA GCC TCA AAA CTG ATC CCT ACC AGT CAC CTG GTG ACG TGT TTT TAA ACA AGC TCT 3348

TTT ATT GGA TGC TCT GCA TAC AGC AGC CTC ACA GAC AAG TCG CCA TTT TGT TTG AGC ATT TGA GGT CTG TAC ACA CCC CAG ATG TTT TTG TGC CGT CGC AGT GCG CTG 3456

GGT GGT TGG TCT TGT CCG TAA CCA CAA ACT GTT CAT TTG AAT TAA ATT AAA GTA TGT ACA TAT AAA TAT 3525

Figure 1. (Continued.)

| | | |
|---------|---|-----|
| Tsk i1 | MEGTS-----FOHPHGLQQTLEQFHLSSRRSLGGPAAFSARWHQDSLFGKD-----GKSVEMMILTPP-QTTP-VMSGPLRTPSDRSTE | 77 |
| Tsk i2 | METVS--RPSFOHPHGLQQTLEQFHLSSMSLGGPAAFSARWHQLLFFKD-----GKEPEPVIQ--HLPPL-VMPGPLEVPSDRSTE | 78 |
| c-Ski | METVS--RSSFOHPHGLQQTLEQFHLSSMSLGGPAAFSARWQEM-YKRDNGKDPAE-----PVHLPPIQPPV-VMPGPFMPSPDRSTE | 82 |
| h-Ski | NEAAGGRGCFQHPHGLQQTLEQFHLSSMSLGGPAAFSARWQEA-YKESAKEAGAAVPAVPVPAATEPPPVHLPAIQPPVPLPGPFMPSPDRSTE | 99 |
| XSk i | METVS--RSSFOHPHGLQQTLEQFHLSSMSLGGPAAFSARWHTQDL-YKRECKGEP-----PEPTLHLPS-QPPVPIPGPLMPSPDRSTE | 82 |
| | | |
| Tsk i1 | RCETVLEGETISCFVVGGEKRLCLPQIINSVLRDESLQIINAVCDELHYCSRCTADOLEHLKVMGILPFSAPSCGLITKTDARLNCNALLYGGTYPPHC | 177 |
| Tsk i2 | RCETVLEGETISCFVVGGEKRLCLPQIINSVLRDESLQIINAVCDELHYCSRCTADOLEHLKVMGILPFSAPSCGLITKTDARLNCNALLYGGTYPPRC | 178 |
| c-Ski | RCETVLEGETISCFVVGGEKRLCLPQIINSVLRDESLQIINAVCDELHYCSRCTADOLEHLKVMGILPFSAPSCGLITKTDARLNCNALLYGGTYPPHC | 182 |
| h-Ski | RCETVLEGETISCFVVGGEKRLCLPQIINSVLRDESLQIINAVCDELHYCSRCTADOLEHLKVMGILPFSAPSCGLITKTDARLNCNALLYGGTYPPPC | 199 |
| XSk i | RCETVLEGETISCFVVGGEKRLCLPQIINSVLRDESLQIINAVCDELHYCSRCTADOLEHLKVMGILPFSAPSCGLITKTDARLNCNALLYGGTYPPRC | 182 |
| | | |
| Tsk i1 | NKESG----SELEERTEKSFKYHFECEGKCKGLFVPELYTSPGAACIQCMDCRLMYPPHKFVYHSHKALENRTCWGFDSANWRAYILLDPP-YTGKEE | 271 |
| Tsk i2 | --KKEMNG-SLELQFTDRSFKYHFECEGKCKGLFVPELYTSPNAACIQCMDCRLMYPTHNFVYHGHKAQENRTCWGFDSANWRAYILLQD-YTEKEE | 274 |
| c-Ski | --KKEFSSTIE--ELTEKSFKYHFECEGKCKGLFVPELYSNSAAGTQCLDCRLMYPPHKFVYHSHKALENRTCWGFDSANWRAYILLSD-YTGKEE | 277 |
| h-Ski | --KKELAASLAGLELSESRVRYHFECEGKCKGLFVPELYSSPSAAGTQCLDCRLMYPPHKFVYHSHKALENRTCWGFDSANWRAYILLSD-YTGKEE | 296 |
| XSk i | AKKSDFFPG-PLLELETEGSKFYHFECEGKCKGLFVPELYGHSAPGTCQCLDCRLMYPPHKFVYHSHKALENRTCWGFDSANWRAYILLARDVGGGDDE | 281 |
| | | |
| Tsk i1 | KSHLEKLEKELKGFDLTGKLSKSCR-----SPSIRAKRSKFDKL----QSADKDRKPDWQLS | 328 |
| Tsk i2 | KARLELFDLDEKELKGFDFANKYKRKASS-----KVSDPIPVKSKSHEDLSQITPL-ADREKQHDWQLS | 336 |
| c-Ski | KARLELQLEDEMEEKFDYNNKYKRKAPRNRESPRVQLRRTKMFKTMWDPAAGSAVLQRPDQNEVPSDPASKKTKIDDSASQSPASTEKEKQSSWRSL | 377 |
| h-Ski | QARLEGRCLDDVEKEKFDYGNKYKRVRPR-----VSSSEPPASIRPKTDDTSSQSPAPSEKDKPSSWRSL | 359 |
| XSk i | LARLEGRLEEEIKEKFDYSNRYKRKAAR-----LSSER-VAKKAKADDSI IHSPPSAEKDKSSWRSL | 343 |
| | | |
| Tsk i1 | SKSAHDLKQVQLKQRPFAFRPWSF--KPAEKVKPAKNEVERSCS---RNOETHNLAFAHAPAVHAKDSNTP-----DRGTAASVQELHN | 411 |
| Tsk i2 | ST-PNKLGNLQCSROKPSAFRPSFHSI SAGDKPESSDPLALLRDG-FYNYKSLKAWPPTSPSLRCHYARYGL---CPHRYPPLPQVEANLQVRAHTPT | 430 |
| c-Ski | SSSSNLSGCVHPQRLESAFRPWSFAVSAHEKELSTHLPALIRDSFSYKSFENAVAPNVALAPPAQKVVSNPPCATVVSRSSEPPSSAAQPRKRKHA | 477 |
| h-Ski | AGSSNLSLGCVHPQRLESAFRPWSFAVSAHEKELSPHLPALIRDS-FYSYKSFETAVAPNVALAPPAQKVVSSPPCAAASRAPELATCTQPRKRKLT | 458 |
| XSk i | SNI-NNAVGYHVPQRLESAFRPWSFAISANDKELSTHLPALIRDS-FYNYKSFENLAVPNAVALTPPVQKVVITSPPCVAVPRSTQSSGSPQSRKRRT | 441 |
| | | |
| Tsk i1 | GDAQPTTKPAHSSNRAEDMDTDGEIDVD---DC---DIDG---VQSSSLASPPSACTSVSQTLPQSMARAQERPSW-LPGTVC----- | 487 |
| Tsk i2 | ---PDLARGELPRSSSPQHMKSGLLPOQANLRLRMTRTPRWRSRWALSSPFTSSSSAKEL-----SSPGMLAPT-----VINTS | 513 |
| c-Ski | AETPAVPE-PVATVTAPEEDKSEAEIEVETREE-----FTSSLSSLSPPFTSSSSAKDM-----SSPGMQAPV-----PVNSS | 546 |
| h-Ski | VDTPGAPV-TLAPVAPEEDKSEAEVAVESREE-----FTSSLSSLSPPFTSSSSAKDL-----GSPGARA-----LPSA | 524 |
| XSk i | AELPIVPEAPAPVPIREEEKESETEIEVESREEC-----TFTSSLSSATKRLRCPPPSGRHLLPHPAQPKTASAPGVGQPICTIMSASTAPAATS | 527 |
| | | |
| | # ### # # # # | |
| Tsk i1 | -----PEMGTMRQMYAGDTKEAREKLIQELVIRRVKQEEKAAALQAKRSIQOELFLRVARKEKLEATEAKRNLKELERLRAENE | 572 |
| Tsk i2 | VSAPAPVTSSESGLSELESLRQALDSLEDSKEKEKFLHEIVKMRVQEEKIGSALQAKRSIQOELFLRVARKEKLEATEAKRNLKELERLRAENE | 627 |
| c-Ski | YEVAAHSDSHSSGLEAELEHLRQALDSLEDTKEAKEKFLHEIVKMRVQEEKINAALQAKRSIQOELFLRVARKEKLEATEAKRNLKELERLRAENE | 646 |
| h-Ski | VPDAAAPADAPSGLEAELEHLRQALDSLEDTKEAKEKFLHEIVKMRVQEEKISALQAKRSIQOELFLRVARKEKLEATEAKRNLKELERLRAENE | 624 |
| XSk i | YEVASHNDQCHSGLAELEHLRQALDSLEDSKEAKEKFLHEIVKMRVQEEKINAALQAKRSIQOELFLRVARKEKLEATEAKRNLKELERLRAENE | 613 |
| | | |
| | # ### # * * * * * | |
| Tsk i1 | RKREAEESCGRLKRELREERGLRVCKGCEAGRLRVKYSQIEELHVLQKAEADREQLRRELQOERARQTLQESWADLQTLALQAGSILPGELKDT | 671 |
| Tsk i2 | KKMKEANESRLRKRELEQARQIRVCKGCEAGRLRAKYSQIEELQMKLOHAEADREQLRADLHREAREHLEKVVRELQOELWPKSSSQSSSENTTS | 713 |
| c-Ski | KKMKEANESRLRKRELEQARQIRVCKGCEAGRLRAKYSQIEELQVKLOHAEADREQLRADLHREAREHLEKVVRELQOELWPKSSSQSSSENTTS | 746 |
| h-Ski | KKMKEANESRLRKRELEQARQIRVCKGCEAGRLRAKYSQIEELQVKLOHAEADREQLRADLHREAREHLEKVVRELQOELWPRARPEAAGSEGAA | 724 |
| XSk i | KKMKEANESRLRKRELEQARQIRVCKGCEAGRLRVKYSQIEELQSKLOHAEADREQLRTDLHREAREHLEKVVRELQOELWSKTSHLPSSEHTRK | 713 |
| | | |
| Tsk i 1 | STEAHROTAQPTNGS | 687 |
| Tsk i 2 | -----H | 714 |
| c-Ski | NME-----N | 750 |
| h-Ski | ELE-----P | 728 |
| XSk i | DIE-----N | 717 |

Fig. 2. Comparison of the deduced AA sequences of tilapia TSk proteins (Tsk i1 and Tsk i2) with those of chickens (c-Ski), humans (h-Ski) and *Xenopus* (XSk i) Ski proteins. The numbers start at the first AA residue. Identical AA residues are boxed in shadow. Dashed lines

represent gaps inserted to maximize the identity among all of the sequences compared. The well-conserved AA residues at the tandem repeat domain and at the leucine zipper domain are indicated by hashes (pound signs) and asterisks, respectively.

difference between them was only an extra C residue at position 848. Such a slight difference, however, cannot explain why two RNA bands of about 3.4 and 3.9 kb were detected when *Xsk i1* was used as a probe in a

Northern blot analysis (Sleman and Laskey, 1993); on the contrary, this latter evidence suggests that another class of *Xenopus c-ski* remains to be discovered. The present study is the first that clearly demonstrates that

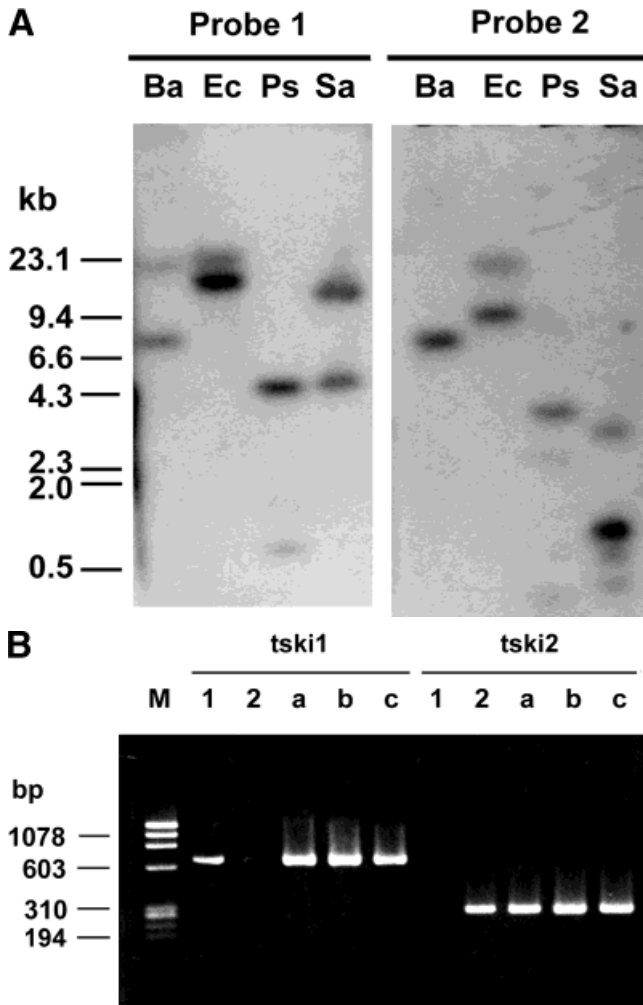


Fig. 3. (A) Southern blot analysis of enzyme-digested genomic DNA isolated from red blood cells of tilapia fry. *Bam*HI (Ba)-, *Eco*RI (Ec)-, *Pst*I (Ps)- and *Sal*I (Sa)-digested genomic DNA were electrophoresed on a gel, transferred to the nylon membrane and hybridized with [α - 32 P]dCTP-labeled probe 1 for *tski1* and probe 2 for *tski2*. Hybridization was carried out in a solution (Sambrook et al., 1989) of 50% formamide, 5 \times Denhardt's solution, 0.5% SDS, 6 \times SSPE and 100 mg/ml fragmented calf thymus DNA at 42 $^{\circ}$ C for 16 hr with a specific activity of 10 8 cpm/ μ g. Membranes were washed in 2 \times SSC and 0.5% (w/v) SDS for 25 min at room temperature, 2 \times SSC and 0.1% SDS for 20 min at room temperature, followed by another washing in 2 \times SSC and 0.5% SDS for 20 min at 37 $^{\circ}$ C. Molecular sizes in kilobase pairs (kb) were given by *Hind*III-digested lambda. (B) PCR analysis for the presence of *tski1* and *tski2* in tilapia genome. Genomic DNA was isolated from three individual tilapia fry (lanes a-c) and amplified using specific primers for *tski1* and *tski2*, respectively. Lane M, the molecular marker, *Hind*III-digested lambda genome; lane 1, plasmid containing *tski1* cDNA insert; and lane 2, plasmid containing *tski2* cDNA insert.

fish have two distinct *c-ski* genes. These two genes have their own expression patterns of both in tissues and during development.

Southern blot analysis of tilapia genomic DNA showed that the restriction patterns for the *tski1* and *tski2* probes were not identical. The *tski1* and *tski2* probes share common DNA sequences at both ends, corresponding to exon 1 and exons 6 and 7 of *c-ski*, but vary in the middle segment. The positive bands of the same M_r

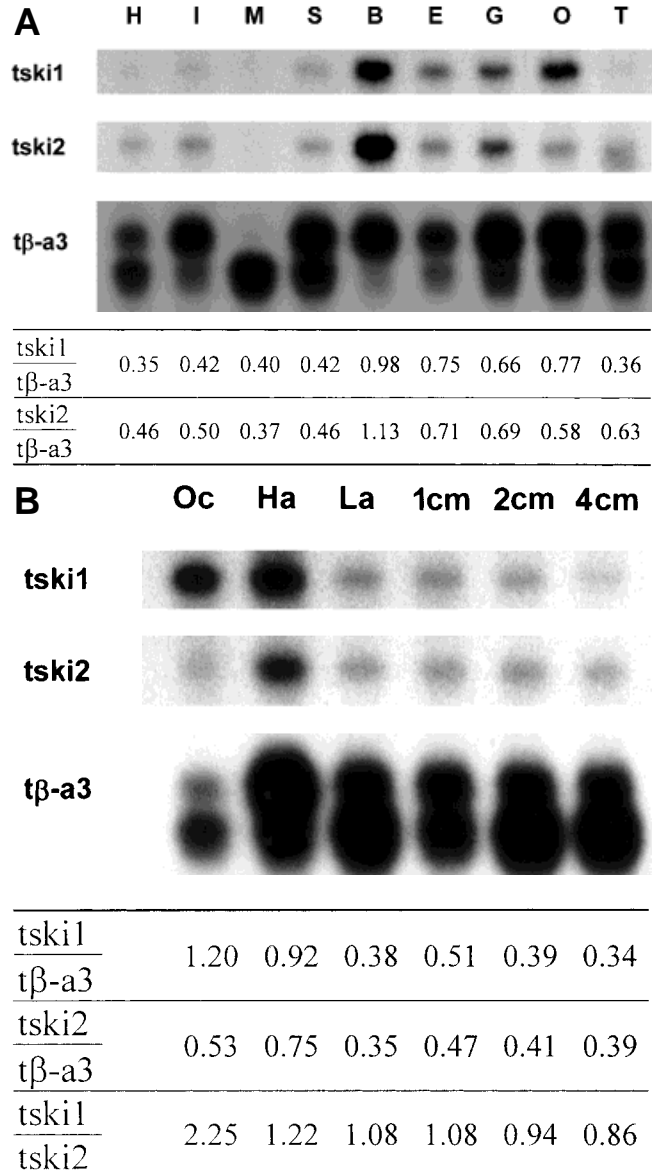


Fig. 4. RNase protection assay for tilapia *tski1* and *tski2* transcripts in (A) different tissues from adult fish, and (B) different developmental stages. The fragment sizes of *tski1* and *tski2* mRNAs protected by *tski1*- and *tski2*-specific antisense RNAs were 453 and 244 bp, respectively. The protection fragments' size of tilapia β -actin serving as internal control were 197 and 163 bp. H, heart; I, intestine; M, skeletal muscle; S, stomach; B, brain; E, eye; G, gill; O, ovary; T, testis; Oc, oocytes; Ha, freshly hatched embryos (stage 15); La, larva at stage 23 (yolk absorption); lanes 1, 2, and 4 cm, fry with 1, 2, and 4 cm body lengths, respectively. Quantified expression levels of *ski* gene (see Materials and Methods) are presented in tables under the figure as *tski1* over β -actin (*tski1*/ β -actin), *tski2* over β -actin (*tski2*/ β -actin) and *tski1* over *tski2* (*tski1*/*tski2*).

(Fig. 3A) may thus be accounted for by the fact that the two probes may hybridize with a common fragment, although each probe would also hybridize with its own specific bands. The different restriction fragment profiles suggest that there might be two independent *c-ski* genes. This hypothesis is supported by the PCR results: PCR products of 688 and 305 bp were generated from

each individual fish when *tski1*- and *tski2*-specific primers were respectively employed. All of this evidence strongly suggests that two independent but non-allelic *c-ski* genes exist in the tilapia genome. Although Northern blot analysis has commonly been used to study the expression and distribution of the *c-ski* gene in many previous studies, cross-hybridization may prevent this method from being sensitive enough to discriminate between two similar *c-ski* mRNAs. We therefore used an RNase protection assay with a class-specific RNA probe. The two classes of tilapia *tski* mRNAs are differentially expressed in some tissues and at certain developmental stages. Thus, like *Xenopus Xski1*, which is preferentially expressed in the female gonads of *Xenopus* (Sleeman and Laskey, 1993), tilapia *tski1* is expressed at a high level in the ovary, but at a low level in the testes. Surprisingly, however, we found that the tilapia *tski2* transcript was extremely rare in the ovary, but absolutely predominant in the testes, an observation that has not previously been reported. In fact, the tilapia *tski1* is maternally inherited, and in this respect it is similar to *Xski1* transcripts, which are maternally regulated during early development (Sleeman and Laskey, 1993). Thus as the tilapia embryos develop from oocytes to 4-cm fry and the *tski2* transcripts increase, the *tski1* transcripts decline, presumably because of the degradation of the maternally-inherited stockpile of *tski1* mRNA. It should therefore be very worthwhile to investigate what roles tilapia TSki1, an ovary-dominant class, and TSki2, a testis-dominant class, might play in gonad development.

Grimes et al. (1993) reported that in any cell a complete absence of *c-ski* expression is unlikely or extremely rare. Namciu et al. (1995) demonstrated that in mouse embryos at 12.5 days of gestation *c-ski* was expressed at a high level in the skeletal muscle. On the other hand, some reports stated that *c-ski* is widely expressed in many different tissues, but they were nevertheless unable to detect either chicken or *Xenopus c-ski* expression within skeletal muscle (Sleeman and Laskey, 1993; Engert et al., 1995). The present study demonstrated that either *tski1* or *tski2* gene expression is extremely rare in the skeletal muscle of adult tilapia. Tilapia *c-ski* transcript can also be detected in the heart, which is consistent with Claycomb and Lanson's (1987) report that *c-ski* was expressed in rodent cardiac tissue.

Nagase et al. (1993) reported a three tandem repeat of 25 AA at positions 571–645 in human c-Ski. Sleeman and Laskey (1993) stated that *Xenopus* c-Ski contained a novel sequence of four contiguous 25-mer repeat elements at positions 527–664 which form a pattern of five hydrophobic, one acidic and two basic residues. Meanwhile, Zheng et al. (1997) demonstrated that chicken c-Ski contains five imperfect tandem repeats of 25-mer and formed hydrophobic "bottoms." They proposed that the core consensus of the repeat motif was LXXELEXLR. In this study, however, we found that actually the last three 25-mer repeats (located at AA

535–593 in tilapia TSki1) are quite well-conserved in the c-Ski proteins of humans, chickens, *Xenopus*, and fish. This fact strongly suggests that LXXELEXRR would more accurately characterize the core consensus sequence of the tandem repeat element at the C-terminus of c-Ski. In this context it is also interesting to note that compared to TSki2, the AA composition of TSki1 is much less similar to the previously known c-Ski proteins. In addition, TSki1 is the only c-Ski protein that lacks the AA 509–529 segment that appears in TSki2. Thus, in TSki1, the first sequence of the five 25-mer tandem repeats is not perfectly conserved.

Based on a computer analysis of the predicted secondary structures of tilapia TSki, we found that the C-terminus (AA 522–659 for TSki1 and AA 539–714 for TSki2) of both classes of TSki encompasses an amphipathic α -helical domain, which has a tendency to form a coiled-coil structure. These are characteristics that are conserved in human, chicken, and *Xenopus* c-Ski sequences, presumably because they are structurally significant and functionally important in Ski proteins. At the C-terminal leucine zipper domain of the Ski protein, there is a periodically arranged AA element, LALELL, which is responsible for dimerization (Nagase et al., 1993) and is also essential for cooperation with the NF1 family of transcription factors (Tarapore et al., 1997). Tilapia TSki1 and TSki2 display a LALELL array at the C-terminus, located in precisely the same position reported for other c-Ski proteins. Because this heptad repeat is well-conserved among Ski proteins, it is highly likely that this sequence may also play an important role in forming the coiled-coil structure of fish TSki polypeptides.

Although *tski2* mRNA contains four separate AUGs within the 370-bp segment of the predicted mRNA 5'-leader, three of these, AUG-1, AUG-2, and AUG-3, located at nucleotide 50, 165, and 253, would terminate after translation of only 20, 19, and 40 residues, respectively. Thus, we propose that tilapia TSki2 starts at the AUG located at nucleotide 370. *tski1* mRNA, meanwhile, contains eight AUGs within a 540-bp segment of the predicted untranslated region, but again, the reading frame beginning with five of these eight AUGs would give ORF of only a few AA. This leaves AUG-2, AUG-3, and AUG-8, which might plausibly terminate *tski1* cDNA. These AUGs all keep the same ORF and share the same stop codon. Since authentic tilapia TSki1 has not yet been purified, the first AA cannot be conclusively determined. However, we propose that the initiation codon of *tski1* mRNA is more likely located at nucleotide 541 (AUG-8), as evidenced from the weight matrix analysis (Bishop, 1994). In this analysis, the score of the AUG-2 sequence (GATTCAATG) was -0.56 . The corresponding score for AUG-3 (GACTGAATG) was 0.72 , while the AUG-8 sequence (GACAACATG) scored highest with $+6.85$. Furthermore, a comparison of the known Ski proteins of humans (Nomura et al., 1989), *Xenopus* (Sleeman and Laskey, 1993) and chickens (Sutrave and Hughes,

1989), also suggested that the initiation codon most likely starts at AUG-8.

ACKNOWLEDGMENTS

The authors thank Dr. Huey-Jen Tsay, Department of Neuroscience, National Yang Ming University, for helping with the RPA experiment; Dr. Ueng-Cheng Yang, Department of Biochemistry, National Yang Ming University; and Mr. Yau-Hung Chen, for assisting in analyzing protein structures.

REFERENCES

- Berk M, Desai SY, Heyman HC, Colmenares C. 1997. Mice lacking the *ski* proto-oncogene have defects in neurulation, craniofacial patterning, and skeletal muscle development. *Genes Dev* 11:2029–2039.
- Bishop MJ. 1994. Guide to human genome computing. San Diego: Academic Press.
- Boyer PL, Colmenares C, Stavnezer E, Hughes SH. 1993. Sequence and biological activity of chicken *snoN* cDNA clones. *Oncogene* 8:457–466.
- Chou PY, Fasman GD. 1978. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv Enzymol* 47:45–148.
- Claycomb W, Lanson N. 1987. Proto-oncogene expression in proliferating and differentiating cardiac and skeletal muscle. *Biochem J* 247:701–706.
- Colmenares C, Stavnezer E. 1989. The *ski* oncogene induces muscle differentiation in quail embryo cells. *Cell* 59:293–303.
- Colmenares C, Suttrave P, Hughes SH, Stavnezer E. 1991. Activation of the *c-ski* oncogene by overexpression. *J Virol* 65:4929–4935.
- Eisenberg D, Sweet RH, Terwilliger TC. 1984. The hydrophobic moment detects periodicity in protein hydrophobicity. *Proc Natl Acad Sci USA* 81:140–144.
- Engert JC, Servaes S, Suttrave P, Hughes SH, Rosenthal N. 1995. Activation of muscle-specific enhancer by the *Ski* proto-oncogene. *Nucleic Acids Res* 23:2988–2994.
- Garnier J, Osguthorpe DJ, Robson B. 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J Mol Biol* 120:97–120.
- Grimes H.L., Ambrose MR, Goodenow MM. 1993. *c-ski* transcripts with and without exon 2 are expressed in skeletal muscle and throughout chick embryogenesis. *Oncogene* 8:2863–2868.
- Higgins DG, Sharp PM. 1988. CLUSTAL: a package for performing multiple sequence alignment on microcomputer. *Gene* 73:237–244.
- Ichikawa K, Nagase T, Ishii S, Asano A, Mimura N. 1997. Trans-regulation of myogenin promoter/enhancer activity by *c-ski* during skeletal-muscle differentiation: the c-terminal of the c-Ski protein is essential for transcriptional regulatory activity in myotubes. *Biochem J* 328:607–613.
- Kozak M. 1978. How do eucaryotic ribosomes select initiation regions in messenger RNA. *Cell* 15:1109–1123.
- Li Y, Turck CM, Teumer JK, Stavnezer E. 1986. Unique sequence, *ski*, in Sloan-Kettering avian retroviruses with properties of a new cell-derived oncogene. *J Virol* 57:1065–1072.
- Liu ZJ, Zhu ZY, Roberg K, Faras A, Guise K, Kapusinski AR, Hackett PB. 1990. Isolation and characterization of beta-actin gene of carp (*Cyprinus carpio*). *DNA Seq* 1:125–136.
- Ludolph DC, Neff AW, Parker MA, Mescher AL, Smith RC, Malacinski GM. 1995. Cloning and expression of the axolotl proto-oncogene *ski*. *Biochim Biophys Acta* 1260:102–104.
- Nagase T, Mizuguchi G, Nomura N, Ishizaki R, Ueno Y, Ishii S. 1990. Requirement of protein co-factor for the DNA-binding function of human *ski* proto-oncogene product. *Nucleic Acids Res* 18:337–343.
- Nagase T, Nomura N, Ishii S. 1993. Complex formation between proteins encoded by the *ski* gene family. *J Biol Chem* 268:13710–13716.
- Namciu S, Lieberman MA, Stavnezer E. 1994. Induction of the *c-ski* proto-oncogene by phorbol ester correlates with induction of megakaryocyte differentiation. *Oncogene* 9:1407–1416.
- Namciu S, Lyons GE, Micales BK, Heyman HC, Colmenares C, Stavnezer E. 1995. Enhanced expression of mouse *c-ski* accompanies terminal skeletal muscle differentiation in vivo and in vitro. *Dev Dynam* 204:291–300.
- Nicol R, Stavnezer E. 1998. Transcriptional repression by v-Ski and c-Ski mediated by a specific DNA binding site. *J Biol Chem* 273:3588–3597.
- Nomura N, Sasamoto S, Ishii S, Date T, Matsui M, Ishizaki R. 1989. Isolation of human cDNA clones of *ski* and the *ski*-related gene, *sno*. *Nucleic Acid Res* 17:5489–5500.
- Sambrook J, Fritsch EF, Maniatis T. 1989. In: Nolan C, editor. *Molecular cloning*. New York: Cold Spring Harbor Press. p 9.1–9.58.
- Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74:5463–5467.
- Sleeman JP, Laskey RA. 1993. *Xenopus c-ski* contains a novel coiled-coil protein domain, and is maternally expressed during development. *Oncogene* 8:67–77.
- Stavnezer E, Gerhard DS, Binari RC, Batzas I. 1981. Generation of transforming viruses in cultures of chicken fibroblasts infected with avian leukemia virus. *J Virol* 39:920–934.
- Stavnezer E, Barkas AE, Brennan LA, Brodeur D, Li Y. 1986. Transforming Sloan-Kettering viruses generated from the cloned *v-ski* oncogene by in vitro and in vivo recombinations. *J Virol* 57:1073–1083.
- Suttrave P, Hughes SH. 1989. Isolation and characterization of three distinct cDNAs for the chicken *c-ski* gene. *Mol Cell Biol* 9:4046–4051.
- Suttrave P, Kelly AM, Hughes SH. 1990. *ski* can cause selective growth of skeletal muscle in transgenic mice. *Genes Dev* 4:1462–1472.
- Tarapore P, Richmond C, Zheng G, Cohen SB, Kelder B, Kopchick J, Kruse U, Sippel AE, Colmenares C, Stavnezer E. 1997. DNA binding and transcriptional activation by the *Ski* oncoprotein mediated by interaction with NF1. *Nucleic Acids Res* 25:3895–3903.
- Zheng G, Blumenthal KM, Ji Y, Shardy DL, Cohen SB, Stavnezer E. 1997. High affinity dimerization by *Ski* involves parallel pairing of a novel bipartite α -helical domain. *J Biol Chem* 272:31855–31864.