

Molecular Cloning of a Partial cDNA Clone Encoding the C Terminal Region of Chicken Breast Muscle Connectin

Authors: Yajima, Hirohiko, Ohtsuka, Hiroshi, Kume, Hideaki, Endo, Takeshi, Maruyama, Kei, et al.

Source: Zoological Science, 13(1) : 119-123

Published By: Zoological Society of Japan

URL: <https://doi.org/10.2108/zsj.13.119>

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Molecular Cloning of a Partial cDNA Clone Encoding the C Terminal Region of Chicken Breast Muscle Connectin

Hirohiko Yajima¹, Hiroshi Ohtsuka¹, Hideaki Kume²,
Takeshi Endo¹, Kei Maruyama², Sumiko Kimura¹
and Koscak Maruyama¹

¹Department of Biology, Faculty of Science, Chiba University,
Chiba 263 and ²National Institute of Physiological Sciences,
Okazaki 444, Japan

ABSTRACT—The cDNA sequence encoding the C terminal region of chicken skeletal muscle connectin was described. Its predicted amino acid sequence had 1,021 amino acids comprising six motif IIs (immunoglobulin C2 domain) and five interdomains. The sequence showed 70–75% homology with that of human cardiac connectin, but 168 amino acids including one motif II were missing in chicken skeletal muscle connectin. The C terminal sequence of chicken skeletal muscle connectin reported by the previous work (Maruyama *et al.*, 1994) was erroneous due to the accidental ligation of the cDNA clone encoding a N terminal region of connectin with a partial porin cDNA clone.

INTRODUCTION

Connectin/titin is the known largest peptide with a molecular mass of about 3,000 kDa. The complete sequence of the human cardiac connectin has recently been elucidated (Labeit and Kolmerer, 1995). The string-like single molecule spans from the Z line to the M line on the myosin filament in vertebrate striated muscle sarcomeres (cf. Maruyama, 1994).

Previously we reported the sequence of a cDNA clone, which we predicted to code the C terminal region of chicken breast muscle connectin (Maruyama *et al.*, 1994). However, it has turned out that this clone was a part of the N terminal region of connectin. The partial connectin clone was accidentally ligated with another cDNA fragment with a stop codon at its 3' terminus. In this report, we present the sequence of the 3' region of the chicken breast muscle connectin cDNA, which is similar to that of human cardiac connectin but 168 amino acids are missing.

MATERIALS AND METHODS

cDNA cloning and sequencing

A cDNA library was constructed from adult chicken breast muscle poly(A)⁺ RNA. A human skeletal muscle connectin cDNA clone (clone 5) was kindly donated by Dr. H. Sorimachi and Professor K. Suzuki of the University of Tokyo (cf. Fig. 1A of Sorimachi *et al.*, 1995). The clone 5 was digested with *Eco*RI to produce 2.7 kb cDNA encoding the very C terminal region of human skeletal muscle connectin. The probe for library screening was obtained by labeling this 2.7 kb cDNA with DIG-dUTP using DIG

DNA Labeling Kit (Boehringer Mannheim, Germany). The screening of the library and the analyses of the clone were performed according to the manufacturer's instructions. Sequencing was carried out by dideoxy nucleotide chain-termination method using BcaBEST Sequencing Kit (TAKARA, Japan).

Northern blotting

Northern blotting was carried out as described before (Endo and Nadal-Ginard, 1987). RNAs were prepared from various tissues of 12 day-old embryos and adults of chicken as described previously (Endo and Nadal-Ginard, 1987). The probes used were the whole length of Cn1 cDNA (3.33 kb) (Maruyama *et al.*, 1994), the 5' *Hinc*II fragment (1.87 kb), and the 3' *Hinc*II fragment (0.82 kb).

RESULTS

The predicted amino acid sequence of C terminal region of connectin

A cDNA clone of 3660 bp, CC4, obtained using the human clone (Sorimachi *et al.*, 1995) as probe, was predicted to have a coding region of 3065 bp, and to encode 1021 amino acids ending with the stop codon (Fig. 1). There are six motif II (immunoglobulin C2) domains with five interdomains (Fig. 2). They have 70–75% homology with the C terminal region of human cardiac connectin (Gautel *et al.*, 1993). In the interdomain between the third and fourth motif IIs there are 13 amino acids in chicken connectin, whereas there are 181 amino acids including one motif II in human connectin (Fig. 2). Thus 168 amino acids are missing in chicken connectin.

In the interdomain between the second and third motif II, there are four KSP regions that are very similar to those of human connectin (Fig. 3a). Only one out of 22 amino acids is different between the two species (proline in chicken and alanine in human). In the interdomain between the fifth

Accepted December 11, 1995

Received November 28, 1995

```

1 AA
3 AAGAGCAGACTTAGGAGAAGGAGGGAAGAGAAATAACTGAGATCACATCAGAAGAAGAAGAAATAGAAATG
1 K S R L R R R R E R E I T E I T S E E E E I E M
78 GTGCAACATGTTACAGGGAATTTTCACTCCTTCCGATTACTAAGGAGAAGAAGATCCCTTTCTCCACTTAC
26 V Q H V H R E F S P P S R L L R R R R S L S P T Y
153 ATTGAGTTGATGCGCCCTGTATCTGAATTAATTAGACCCGATCAGACCTCCTGAGGAAAGTGAAGAAGATCC
51 I E L M R P V S E L I R P R S R P P E E S E R R S
228 CCAACACCAGAGAAGTTCGACCAGTTTACCCAGCCCGTTTCTACTGAAAGATCACTCTCTCGGTTTGAAGG
76 P T P E R T R P R S P S P V S T E R S L S R F E R
303 ATGGCAAGATTTGACATTTTCTCGATACGAGTCCATGAAATCTGCTTTGAAAACACAGAAAATATGGAAGG
101 M A R F D I F S R Y E S M K S A L K T Q K T M E R
378 AAATACGAAGTCTAACTCAACAGCCTTTACCCCTTGACCATGCTCCTCGCATCACCTGAGAATGCGCTCACAC
126 K Y E V L T Q Q P F T L D H A P R I T L R M R S H
453 CGTGTGCCGTGTGCCATAACACCAGGTTCTTCTGAATGTCCAGTCAAACCAACTGCTGACGTGAAGTGGTAC
151 R V P C G H N T R F I L N V Q S K P T A D V K W Y
528 CACAATGGTATCGAGCTCCAAGAGAGCAGTAAGATACATTTCTCGAATACCAGCGGGGTATTAACACTGGAAT
176 H N G I E L S S K I H F S N T S G V L T L E I
603 CTCGACTGTATATTGATGACAGTGAACGTACAGGGCGGTATGCACAACTACAAAGGAGAATGTTCTGATTAC
201 L D C H I D D S G T Y R A V C T N Y K G E C S D Y
678 GCAACGTTAGATTACTGGAGGATTACACTACACTCTTCCAGCGGAGAGATGAAGAAGTACCAAGGTCA
226 A T L D V T G G D Y T T Y S S Q R R D E E V P R S
753 ATTTTACTGACCTGACAAGAACAGGCGTATGACAGTCTCGTCAATTTAGAAAGCAACTGCTGCAGAACGAGT
251 I L P D L T R T E A Y A V S S F K K A T A A G S H
828 TCCTCAGTAAGAGAGGTCAAATCAGAAGTGTACGACGAGAGAGTCACTTTTATCGTATGAACACCAGCTTCT
276 S S V R E V K S E V S A T R E S L L S Y E H H A S
903 TCCGAGGAAAAATCACTGCCTCTGAAGAAAAATCATTGGAAGAAAGGACTGTCCATAAGGCATTTAAGGCACT
301 S E E K I T A S E E K S L E E R T V H K A F K S T
978 CTGCCAGCAACGATCCTTACGAAACCACGGTCCATCACTGTTTCTGAAGGCGAGACTGCAGATTTTCTGTGAT
326 L P A T I L T K P R S I T V S E G E T A R F S C D
1053 GTTGACGGTGAGCCAGCACCAACAATAACATGGGTCCGTGCGGGTCAACCTATTGTTTCTCCAGGCGCTTCCAA
351 V D G E P A P T I T W V R A G Q P I V S S R R F Q
1128 ATAACACGGACACAGTACAAATCTACTTTCGAGATTTCTTGGTCCAGATTGCAGATGAAGGAAGTACACTGTT
376 I T R T Q Y K S T F E I S L V Q I A D E G S Y T V
1203 GTGGTGGAAAATCTGAAGGAAGACAGGAAGCGCACTTCACTTTGACCGTTCAAAGAAAAAGAAATCCCTGAAAA
401 V V E N S E G R Q E A H F T L T V Q R K R I P E K
1278 GCAATTACATCGCCTCCGAGAATCAAATCTCCTGAGCCTCGTGTAAAGTCTCCAGAGCCAGTTAAGTCTCCTAAA
426 A I T S P P R I K S P E P R V K S P E P V K S P K
1353 CGAGTGAATCTCCTGAACCCATCAGTACCCCTCAAAGCCAAGTACCCGCTGGAGACAAAACAGCAGTGA
451 R V K S P E P I S T P S K A K S P P G D K T A P V
1428 GAGAAAGTACAATTACCAACTGCCTCTCCTCCAAAGATAAAAGAGCAACTAAAAGCAGAAACCCTTGGAGATAAG
476 E K V Q L P T A S P P K I K E Q L K A E T L G D K
1503 GTGAAGTTATCCTGCGCAGTTGAAAGTAGTGTTTTAAGCATCAGGGAAGTGGCTTGGTATAGGACGGCAAAAG
501 V K L S C A V E S S V L S I R E V A W Y K D G K K
1578 CTGAAAGAAGACCATCATTCAAGTTTTCATTATGACGAGATGGCAGATATGAGTCAAAATCCACAACCTCACAC
526 L K E D H H F K F H Y A A D G T Y E L K I H N L T
1653 GAATCTGACAAAGGAGAATATACTTGTGAAATTTGGGAGAAGGAGGTATTTCAAAAATACTTCCAGTTTACC
551 E S D K G E Y T C E I M G E G G I S K T N F Q F T
1728 GGCCAAGTATTTAAAACATCCATTCCAGGTACAAAGTGTATCTGAAACCCCTAAATCAGTTGAGAAGGAGAC
576 G Q V F K N I H S Q V Q S V S E T P K S V E K G D
1803 AAAGTACTTGCTGTTTCTACCCAGAAGAAATCATCAGCGCAACTGAAGAAAAAGCTCAATTGAAGAAGTCATT
601 K V L A V S T Q K K S S A A T E E K A A I E E V I
1878 AAAAAGTCAATTGTTACTGAAGATGTCAAACAATTGCAAGCAGAAATTAGAGCTTCTTCGACTCAAATGACTGTG
626 K K S I V T E D V K Q L Q A E I R A S S T Q M T V
1953 TCTGAAGGCGAGAAAGTGAAGTCTGAAAGCAACATTCTGGTGCCTCTGAAGTAAATGGGTGCTAAATGGAATG
651 S E G Q K V T L K A N I P G A S E V K W V L N G M
2028 GAGCTAAGAAATTCAGATGATTACAGATATGGCATTTCGGAAGTAATCACACGCTAACTATCAAAAAGGCTAGC
676 E L R N S D D Y R Y G I S G S N H T L T I K K A S
2103 AATAAAGATGAAGGTATACTCACCTGTGAAGGTAAAACCTGACGAAGGCACTATTAATGCCAGTACGTGCTTACC
701 N K D E G I L T C E G K T D E G T I K C Q Y V L T
2178 TTCTCTAAGAGCCCTCAAACGAACCGGCTTTATCAGCAGCCGAAGTCTCAAATGTCAATGAAGGACAAGAC
726 F S K E P S N E P A F I T Q P K S Q N V N E G Q D

```

(continued)

and sixth motif IIs there is serine rich region: 30 serine residues are found in 95 amino acids (Fig. 3b). There are two repeats of 4 serine residues, five repeats of 2 serine residues and twelve serine residues. This serine rich region is also similar to that of human cardiac connectin. There are only two serines more in human connectin.

Cn1 is not the 3' terminal region of connectin cDNA

In the previous paper (Maruyama *et al.*, 1994), we described that Cn1, a clone screened with an antibody to connectin (Pc 1200), hybridized to an mRNA of chicken muscle with a mobility beyond the upper limit of resolution

(~25 kb). However, it also cross-hybridized to 2.0 and 1.5 kb mRNA, suggesting the possibility that Cn1 contained another gene. The amino acid sequence at the C terminal region, 775 to 805, had a homology to that of porin, voltage dependent channel protein (Kayser *et al.*, 1989).

Reexamination of the sequence of Cn1 showed that N terminal 774 amino acids are derived from connectin at the N terminal region not at the C terminal region (Yajima *et al.*, unpublished data). The C terminal 37 amino acids come from porin. In the present study, the 175 amino acid sequence of porin starting the 108th amino acid up to its C terminus was determined as shown in Figure 4. This new sequence had

```

2253 GTGTGTTTACCTGTGAAGTTTCTGGCGATCCTTCTCTGAAGTTGAATGGCTTAGGAATAATCAACCTATTGCC
751 V L F T C E V S G D P S P E V E W L R N N Q P I A
2328 GTTTCATCACATGAGAGCGACTCGCTCTAAAAACACTATTCTCTTGAGATTGCGAAATGCTGAGTGAGTGAC
776 V S S H M R A T R S K N T Y S L E I R N A A V S D
2403 ACTGGAAAATACACAGTTAAGGCTAAAAATTACCACGGCAATGCTCTGCTACAGCATCTTTGACTGTATTCCCT
801 T G K Y T V K A K N Y H G Q C S A T A S L T V F P
2478 CTAATTGAAGAACCTCCAAAAGAGGTAGTATTGAAGACAAGCGGCGACGCAAGCATGCACGAAAGTTTCTTCT
826 L I E E P P K E V V L K T S G D A S M H E S F S S
2553 CAATCCTTTCAAATGGCTGCTTCTAAACAAGAGGCTTCATTACAGCAGTTTCAGCAGTAGCAGCATGACTGAAACG
851 Q S F Q M A A S K Q E A S F S S F S S S S M T E T
2628 AAATTTGCAAGCATGTCTGCCAAAAGCATGTCCTCCATGAAAGAATCCTTTGTAGAGATGAGCTCCAGCAGTATT
876 K F A S M S A K S M S M K E S F V E M S S S I
2703 ATGGGAAATCTAGTATGGCTCAACTGGAAGTTCAACTAGTAAGATGCTTAAATCAGGTGTGAGAGGAGTACCA
901 M G K S S M A Q L E S S T S K M L K S G V R G V P
2778 CCAAAAATTGAAGCTCTGCCATCTGACATCAGTATTGATGAAGGAAAAGTTCTGACATTATCTTGCCTTTTTTCG
926 P K I E A L P S D I S I D D E G K V L T L S C A F S
2853 GGTGAACCTGCTCTGAAATAACATGGTACTGACGAGGGAGAAAATTACCAGTCAGGACCAGCAGCAGATTC
951 G E P A P E I T W Y C R G R K I T S Q D Q Q G R F
2928 CACATTGAAACCAGCGAAGATCTAACCACCTGATCATCATGGATGTCAGAAGAATGACGGTGAATTTTACACT
976 H I E T S E D L T T L I I M D V Q K N D G G I Y T
3003 CTGAATTTAGGAAATGAATTCGGTACCGATTCTGCCACTGTGAATATTAATATTTCGATCAATTTAGAAAGCTTGA
1001 L N L G N E F G T D S A T V N I N I R S I *
3078 TCTGTATTATTTACACTTGTCTTTTTTACAGTGATTCATATATGGACTGATCTGTAAATATTAATAAAAAAAAAA
3153 ATATTTTGTACCTACCTGACTGAGACAAAAGTCCAGAGATATACGTTATGACTTATAGTCATTATTGACCTAAGA
3228 ATTTTAAACACTTTTTCCTGACATGTACATACTGTATATAGCCGAAGTTAACGGTTATGAAGTTTGTACCTT
3303 TATTTTATGACATTTTGAATGTAACCTTTTGAATTAACAGTTGGTAGGAGAAAGTTTCTAGCAACAGCATATC
3378 CTGCTCAACATTTAACTAATTTTGTGCCTCAACTATTGTTGATGTCTAAGAATGCCTCAACAGGTAGAGAGGC
3453 TCCCTGTTGAAGATTACCTACAACCAAGAGAGGATACTGTGCATAGTATTTTCATACATGCACAAATCATTATGTT
3528 GAATCAGAAATGTAAGGCACTGGTGATATTACAATTACCTCCTGTAAGTATTACTTTAATGTTTACATTGGA
3603 TCTGATTACGTCGTAATTTCCATGCCTGACTGACAATTTTGAACAAAGTGCAAGCGG

```

Fig. 1. Nucleotide and predicted amino acid sequences of the 3' terminal region of chicken skeletal muscle cDNA (CC4).

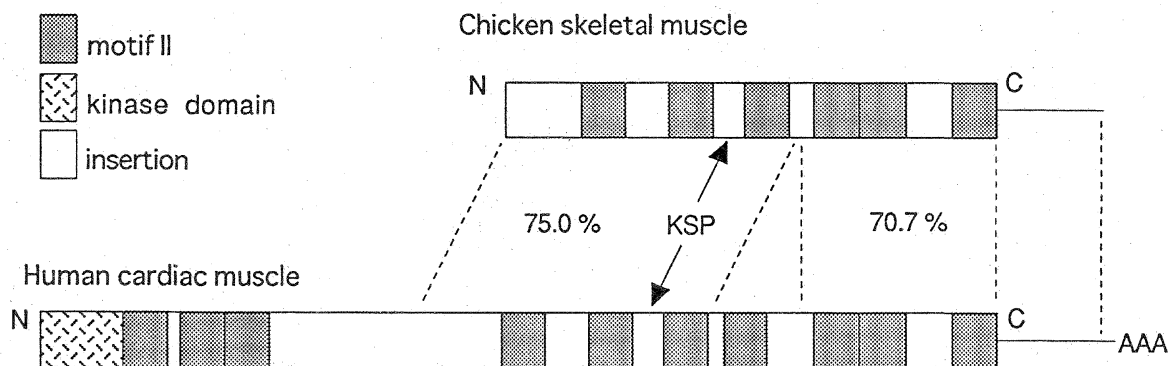


Fig. 2. Comparison of the domains of chicken skeletal muscle and human cardiac connectins at the C terminal region. Chicken skeletal muscle (this paper); human cardiac muscle (Labeit and Kolmerer, 1995). KSP, KSP repeat region.

74% homology with that of human porin.

Evidence by Northern blotting

When Northern blotting was carried out by using a probe of the whole length of Cn1 cDNA, two distinctive sizes of mRNAs were detected in embryonic and adult tissues (Fig. 5A). One of the mRNAs had a mobility beyond the upper limit of resolution (~25 kb) and was specific for skeletal and cardiac muscles as reported previously (Maruyama *et al.*, 1994). According to its extremely large size and tissue specificity, the hybridization band represents connectin mRNA.

In contrast, another mRNA was ~1.5 kb and present in all the tissues examined. To address whether or not the band was generated by the hybridization to the porin-like sequence, Northern blotting was carried out by use of the 5'

HincII fragment (1.87 kb) and the 3' *HincII* fragment (0.82 kb) of the cDNA. The former fragment did not contain the porin-like sequence but the latter contained exclusively the sequence. The 5' fragment hybridized only to the mRNA with an apparent size of ~25 kb in striated muscles (Fig. 5B), whereas the 3' fragment did exclusively to the ubiquitous ~1.5 kb mRNA. These results strongly suggest that Cn1 cDNA is a ligated form of connectin and porin-like sequences. These two cDNAs were likely to be ligated into one λ gt11 vector during the course of cDNA library construction (Fig. 4).

DISCUSSION

Connectin/titin is the largest peptide ever known. Human cardiac connectin consists of 244 repeated motifs

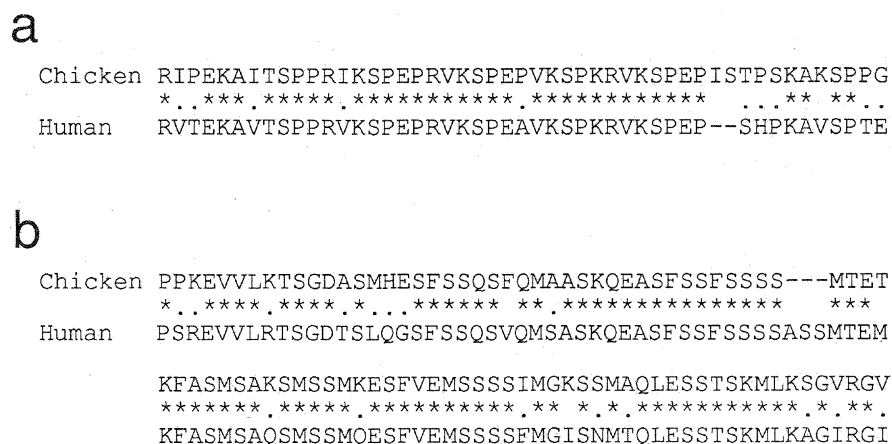


Fig. 3. Comparison of the amino acid sequences of chicken skeletal muscle and human cardiac connectins at the KSP and serine rich regions. Chicken skeletal muscle (this paper); human cardiac muscle (Labeit and Kolmerer, 1995). a, KSP region; 72.4% homology for 58 amino acids. b, serine rich region; 77.6% homology for 98 amino acids.

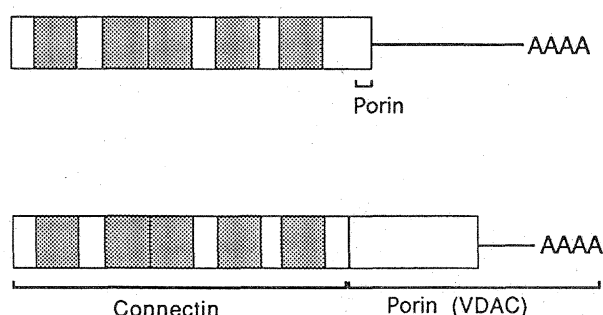


Fig. 4. The structure of Cn1, a connectin-porin ligated cDNA. Upper : Maruyama *et al.*, 1994 ; Lower : this paper. VDAC, voltage dependent channel protein. 5' Terminal region of connectin cDNA was ligated with a portion of porin cDNA.

(motif I, fibronectin type III and motif II, immunoglobulin C2 domain) with a molecular mass of 3,000 kDa (Labeit and Kolmerer, 1995). This paper describes the cDNA sequence of the C terminal region (1020 amino acids) of chicken skeletal muscle. Comparison of the sequence with that of human cardiac connectin showed some 70–75% homology, but 168 amino acids including one motif II are lacking in chicken skeletal muscle connectin. Another cDNA clone, CC2, also lacked these 168 amino acids region to support the sequence in the clone CC4 described above. This deletion may have derived from alternative splicing as shown by Labeit and Kolmerer (1995).

This paper is also to correct our previous paper (Maruyama *et al.*, 1994): a cDNA encoding a N terminal region of connectin accidentally ligated with cDNA encoding porin (Kayser *et al.*, 1989). This possibility was independently pointed out by Sebestyen *et al.* (1995).

ACKNOWLEDGMENTS

We thank Dr. H. Sorimachi and Prof. K. Suzuki of the University of Tokyo for their kind supply of human skeletal muscle connectin cDNA.

REFERENCES

- Endo T, Nadal-Ginard B (1987) Three types of muscle-specific gene expression in fusion-blocked rat skeletal muscle cells: translational control in EGTA-treated cells. *Cell* 49: 515–526
- Gautel M, Leonard K, Labeit S (1993) Phosphorylation of KSP motifs in the C-terminal region of titin in differentiating myoblasts. *EMBO J* 12: 3827–3834
- Kayser H, Kratzin HD, Thinnies FP, Gotz H, Schmidt WE, Eckart K, Hilschmann N (1989) Identification of human porins. II. Characterization and primary structure of a 31-kDa porin from human B lymphocytes (porin 31 HL). *Biol Chem Hoppe-Seyler* 370: 1265–1278

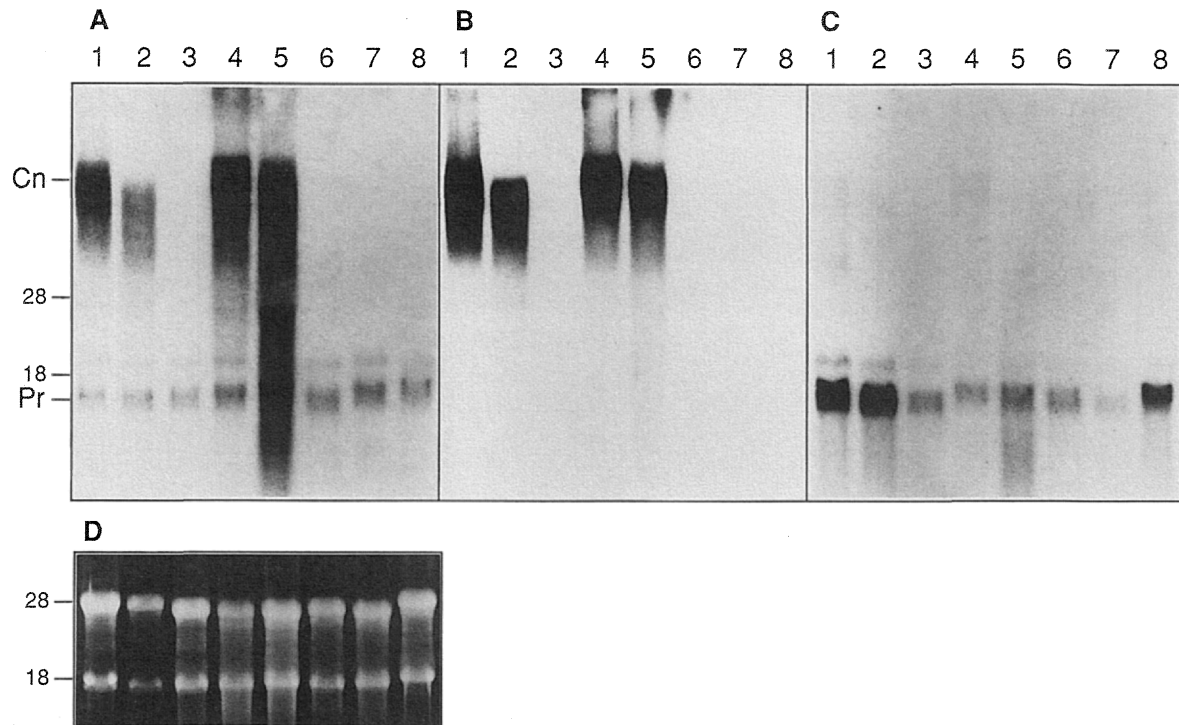


Fig. 5. Northern blotting of RNAs from various tissues probed with Cn1 cDNA. (A) Northern blotting probed with the whole length of Cn1 cDNA. (B) Blotting probed with the 5' *HincII* fragment of the cDNA. (C) Blotting probed with the 3' *HincII* fragment. (D) Ethidium bromide staining pattern showing 28S and 18S rRNAs. Ten micrograms of the following total RNAs were loaded to each lane: lane 1, embryonic skeletal muscle; 2, embryonic cardiac muscle; 3, embryonic gizzard smooth muscle; 4, adult skeletal muscle; 5, adult cardiac muscle; 6, adult gizzard smooth muscle; 7, adult brain; 8, adult liver. The positions of connectin and porin-like mRNAs are indicated by Cn and Pr, respectively. The positions of 28S and 18S rRNAs are denoted by 28 and 18, respectively.

- Labeit S, Kolmerer B (1995) Titins: Giant proteins in charge of muscle ultrastructure and elasticity. *Science* 270: 293–296
- Maruyama K (1994) Connectin, an elastic protein of striated muscle. *Biophys Chem* 50: 73–86
- Maruyama K, Endo T, Kume H, Kawamura Y, Kanzawa N, Kimura S, Kawashima S, Maruyama K (1994) A partial connectin cDNA encoding a novel type of RSP motifs isolated from chicken embryonic skeletal muscle. *J Biochem* 115: 147–149
- Sebestyen MG, Wolff JA, Greaser ML (1995) Characterization of a 5.4 kb cDNA fragment from the Z-line region of rabbit cardiac titin reveals phosphorylation sites for proline-directed kinases. *J Cell Sci* 108: 3029–3037

- Sorimachi H, Kinbara K, Kimura S, Takahashi M, Ishiura S, Sasagawa N, Sorimachi N, Shimada H, Tagawa K, Maruyama K, Suzuki K (1995) Muscle-specific calpain, p94, responsible for limb girdle muscular dystrophy type 2A, associates with connectin through IS2, a p94-specific sequence. *J Biol Chem* 270: 31158–31162

Addendum

The sequence data are available from DDBJ/EMBL/GenBank database under the accession number D83008.