

# Molecular cloning and sequencing of the chicken smooth muscle myosin regulatory light chain

Neil G. Messer and John Kendrick-Jones

*MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England*

Received 7 April 1988

A cDNA probe was constructed from a chicken skeletal muscle regulatory light chain cDNA and was used to screen a chicken gizzard cDNA library. A clone containing the entire coding region of the chicken gizzard regulatory light chain was isolated and sequenced. The deduced protein sequence is identical to the most recently reported chemical sequence of the chicken smooth muscle regulatory light chain, and has homologies with other troponin C-like calcium-binding proteins.

Myosin light chain; cDNA cloning; Nucleotide sequence; (Chicken gizzard)

## 1. INTRODUCTION

The myosin regulatory light chains (RLCs) are small acidic polypeptides non-covalently bound to the neck region of the myosin head, which regulate the interaction of the myosin head with actin. On the basis of their regulatory properties, they can be divided into three functional classes: (i) molluscan RLCs regulate muscle contraction by inhibiting myosin-actin interaction in the absence of  $\text{Ca}^{2+}$ , and this inhibition is relieved by  $\text{Ca}^{2+}$  binding to the myosin head [1]; (ii) vertebrate smooth muscle RLCs exert a similar inhibitory effect, which is relieved by phosphorylation of the RLC by a specific  $\text{Ca}^{2+}$ -calmodulin activated kinase [2]; (iii) vertebrate sarcomeric myosin RLCs, though similarly phosphorylated, do not appear to play a primary role in the initial events in the regulation of contraction [3].

*Correspondence address:* J. Kendrick-Jones, MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England

*Abbreviations:* bp, base pair; Tris, tris(hydroxymethyl)aminomethane

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession number Y00983

The RLCs belong to the superfamily of calcium-binding proteins which includes calmodulin and troponin-C. These proteins are characterised by four EF hand  $\text{Ca}^{2+}$ -binding motifs [4]; in the case of the RLCs, mutations have eliminated  $\text{Ca}^{2+}$  binding in three of the EF hands, leaving only the first EF hand as a functional  $\text{Ca}^{2+}/\text{Mg}^{2+}$ -binding site [5,6].

A number of RLC cDNA clones have been isolated [7–11]. Using a chicken skeletal muscle RLC cDNA clone [9], we have expressed this RLC and explored structure-function questions by site-directed mutagenesis [12]. In order to extend these studies and compare the different functional classes of RLC, we have cloned and sequenced a chicken smooth muscle RLC.

## 2. MATERIALS AND METHODS

### 2.1. Materials

Enzymes were from New England Biolabs. [ $\alpha$ - $^{32}\text{P}$ ]dATP (3000 Ci/mmol, aqueous) and [ $\alpha$ - $^{35}\text{S}$ ]dATP (>400 Ci/mmol, aqueous) were from Amersham. Other reagents were from Sigma, BDH and BRL. A chicken gizzard cDNA library in the vector  $\lambda$ gt11, containing approx. 43000 independent recombinants [13], was kindly provided by Professor A.R. Means, Baylor College, Houston, TX.

### 2.2. Construction of probes and screening of library

The cDNA sequence chosen for a probe was a *Bam*HI restric-

tion fragment consisting of the first 331 bp of a chicken skeletal muscle RLC cDNA clone [9], together with the 17 bp of pLcII vector sequence [14] closest to the cloning site on the 5'-side. This region was chosen since there is around 75% amino acid homology between the N-terminal halves of the two RLCs, and it was assumed that it would therefore constitute a strongly hybridising and selective probe. The fragment was isolated from a 1% agarose gel [15] and labelled using a Pharmacia 'Oligolabelling' kit according to the manufacturer's instructions. The library was screened with this probe according to Benton and Davis [16]. Hybridisation positive plaques were picked, eluted into sterile  $\lambda$  diluent (10 mM Tris, pH 7.4, 15 mM MgSO<sub>4</sub>, 200 mM NaCl, 1 g/l gelatin), replated and rescreened until homogeneous.

### 2.3. Isolation and sequencing of positive clones

$\lambda$  DNA was isolated from a positive plate lysate [15]; the insert was digested out with *Eco*RI and subcloned into pUC8 [17] and M13mp18 [18]. Sequencing was according to Sanger et al. [19] using [ $\alpha$ -<sup>35</sup>S]dATP. The primers used were M13 universal primer and oligonucleotides synthesised by Mr T. Smith. Alignment of gel readings and analysis of the sequence were carried out using the programs DBAUTO [20], DBUTIL [20], ANALYSEQ [21] and IALIGN [22] on a VAX 8600 computer.

## 3. RESULTS

### 3.1. Isolation of a chicken smooth muscle RLC cDNA clone

Three hybridisation positive clones were

detected after screening approx.  $2 \times 10^5$  clones of the chicken gizzard muscle library (about four times the number of independent recombinants). The clone giving the strongest signal, denoted CGMRLC, was purified and the insert isolated on the assumption that it contained the longest sequence complementary to the probe, and therefore was likely to be the most complete RLC clone.

### 3.2. Nucleotide sequence of the CGMRLC clone

Using M13 universal primer, the sequences of the first 100 bp and the last 250 bp of the coding region were determined. Using this sequence information, oligonucleotide primers were synthesised which were used to determine the complete sequence of the clone on both strands (fig.1 summarises the sequencing strategy). Each base was sequenced on average approx. 6 times on each strand. The complete nucleotide sequence and the translation of the coding region are shown in fig.2; CGMRLC contains the complete coding region of the smooth muscle RLC, along with a 5'-untranslated region of 85 bp and a 3'-untranslated region of 59 bp.

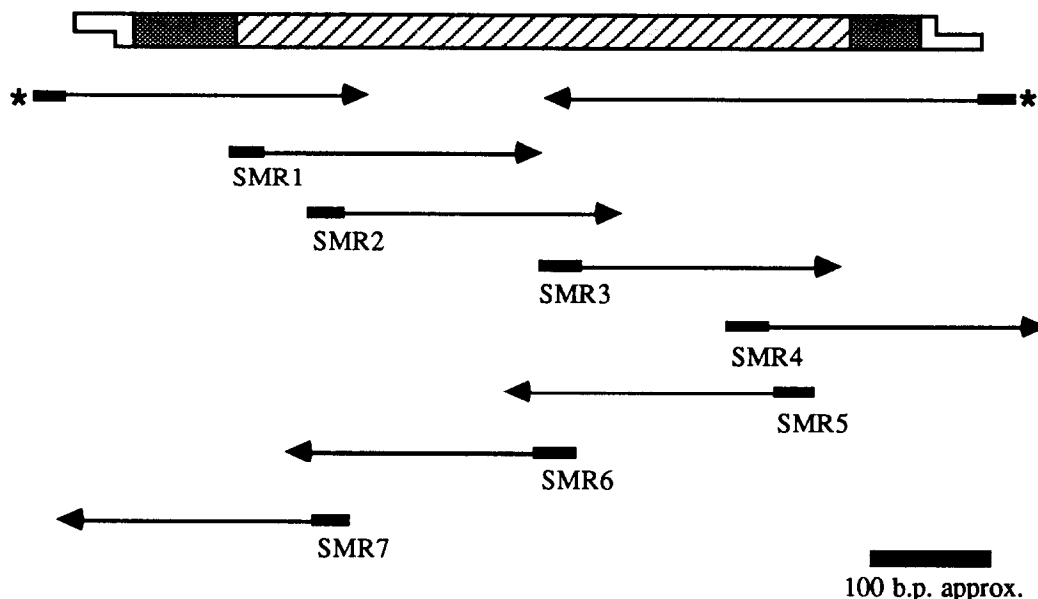


Fig.1. Summary of the sequencing strategy. The hatched area represents the coding sequence of CGMRLC, the stippled areas the 5'- and 3'-untranslated regions and the white areas the *Eco*RI linkers. \*— represents M13 universal primer, and — represents primers designed on the basis of the partial nucleotide sequence obtained using M13 primer (see text). Arrows represent the sequences read using each primer, and the bar gives the approximate scale.



of the tryptic peptides in the latter. In confirming our previous N-terminal sequence, our DNA sequence supports the observation [24] that the sequence of three basic residues (residues 11–13) upstream of the phosphorylatable serine (serine 19) is essential for phosphorylation by myosin light chain kinase (MLCK). The two phosphorylatable RLCs (skeletal and smooth muscle) shown in fig.3 both have this putative recognition sequence; the scallop muscle RLC, which contains a serine residue in the correct position for phosphorylation, but is not phosphorylated by MLCK, lacks this basic sequence.

The amino acid sequence shows the four putative  $\text{Ca}^{2+}$ -binding domains characteristic of all the members of the EF hand superfamily of  $\text{Ca}^{2+}$ -binding proteins. Fig.3 shows an alignment of the four domains of scallop striated [6], chicken smooth muscle [23,24] and chicken skeletal muscle [25] RLC sequences, that is, one representative of each functional class, together with the ancestral EF hand sequence of Baba et al. [26]. From this comparison, the deletions and non-conservative substitutions (often the presence of prolines) which have removed the  $\text{Ca}^{2+}$ -binding function of domains II, III and IV of the RLCs are apparent. Only domain I retains all the residues necessary for  $\text{Ca}^{2+}$  or  $\text{Mg}^{2+}$  binding; this is the divalent cation binding site observed in all RLCs. It is also clear that the sequences diverge much more in the C-terminal than the N-terminal halves of the proteins. Proteolytic cleavage studies indicate that the C-terminal halves, but not the N-terminal halves of the RLCs are able to bind to the myosin head [27]. It may be, therefore, that differences in this region lead to different interactions with the myosin head and thereby to differences in regulation. Expression of CGMRLC in the pLcII vector system [14], previously used to express the chicken skeletal muscle RLC [12], and protein engineering studies will allow us to test these ideas and compare the properties of the different RLCs.

*Note added in proof:* A chicken smooth muscle RLC has recently been cloned independently of this report (Zavodny, P.J. et al., Nucleic Acids Res. 16, 1214). The nucleotide sequences in the coding regions are identical except for four silent changes at the positions numbered 271, 274, 301 and 451 in our sequence.

*Acknowledgements:* We thank Professor Anthony Means for providing us with his chicken gizzard cDNA library, Terry Smith for oligonucleotide synthesis and Michael Way for much helpful advice and for critically reading the manuscript. N.G.M. thanks the Medical Research Council of Great Britain for a research studentship.

## REFERENCES

- [1] Kendrick-Jones, J., Szentkiralyi, E.M. and Szent-Gyorgyi, A.G. (1976) J. Mol. Biol. 104, 747–775.
- [2] Adelstein, R.S. and Eisenberg, E. (1980) Annu. Rev. Biochem. 49, 921–956.
- [3] Perschini, A., Stull, J.T. and Cooke, R. (1985) J. Biol. Chem. 260, 7951–7954.
- [4] Kretsinger, R.H. (1980) CRC Crit. Rev. Biochem. 8, 119–174.
- [5] Collins, J.H. (1976) Nature 259, 699–700.
- [6] Bagshaw, C.R. and Kendrick-Jones, J. (1979) J. Mol. Biol. 130, 317–336.
- [7] Hastings, K.E. and Emerson, C.R. jr (1982) Proc. Natl. Acad. Sci. USA 79, 1553–1557.
- [8] Arnold, H.H., Krauskopf, M. and Siddiqui, M.A.Q. (1983) Nucleic Acids Res. 11, 1123–1131.
- [9] Reinach, F.C. and Fischman, D.A. (1985) J. Mol. Biol. 181, 411–422.
- [10] Kumar, C.C., Cribbs, L., Delaney, P., Chien, K.R. and Siddiqui, M.A.Q. (1986) J. Biol. Chem. 261, 2866–2872.
- [11] Taubman, M.B., Grant, J.W. and Nadal-Ginard, B. (1987) J. Cell Biol. 104, 1505–1513.
- [12] Reinach, F.C., Nagai, K. and Kendrick-Jones, J. (1986) Nature 322, 80–83.
- [13] Guerrero, V., jr, Russo, M.A., Olson, N.J., Putkey, J.A. and Means, A.R. (1986) Biochemistry 25, 8372–8381.
- [14] Nagai, K. and Thøgersen, H.C. (1984) Nature 309, 810–812.
- [15] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- [16] Benton, W.D. and Davis, R.W. (1977) Science 196, 180–182.
- [17] Vieira, J. and Messing, J. (1982) Gene 19, 259–268.
- [18] Messing, J. (1983) Methods Enzymol. 101, 20–78.
- [19] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
- [20] Staden, R. (1982) Nucleic Acids Res. 10, 4731–4751.
- [21] Staden, R. (1985) Genet. Eng. Princ. Methods 7, 67–114.
- [22] Dayhoff, M.O. (1983) Methods Enzymol. 91, 524–545.
- [23] Maita, T., Chen, J.I. and Matsuda, G. (1981) Eur. J. Biochem. 117, 417–424.
- [24] Pearson, R.B., Jakes, R., John, M., Kendrick-Jones, J. and Kemp, B. (1984) FEBS Lett. 168, 108–112.
- [25] Matsuda, G., Suzuyama, Y., Maita, T. and Umegane, T. (1977) FEBS Lett. 84, 53–56.
- [26] Baba, M.L., Goodman, M., Berger-Cohn, J., Demaille, J.G. and Matsuda, G. (1984) Mol. Biol. Evol. 1, 442–455.
- [27] Scholey, J.M., Taylor, K.A. and Kendrick-Jones, J. (1981) Biochimie 63, 255–271.