J. Mol. Biol. (1991) 217, 621–623

Crystallization and Preliminary Analysis of Telokin, the C-Terminal Domain of Myosin Light Chain Kinase

Timothy A. Anderson[1], Masaaki Ito[2], David J. Hartshorne[2] and Ivan Rayment[1]†

[1]*Institute for Enzyme Research, Graduate School, and Department of Biochemistry, College of Agricultural and Life Sciences University of Wisconsin, Madison, WI 53705, U.S.A.*

[2]*Department of Animal Sciences, University of Arizona Tucson, AZ 85721, U.S.A.*

(Received 27 September 1990; accepted 16 October 1990)

Telokin, an acidic protein related to the C-terminal portion of smooth muscle myosin light chain kinase from turkey gizzard has been crystallized in a form suitable for a high-resolution diffraction analysis. The crystals were grown from solutions of polyethylene glycol 8000 using the hanging-drop vapor diffusion method. They belong to the trigonal space group $P3_121$ or $P3_221$ with cell parameters $a = 64{\cdot}0$ Å, $c = 59{\cdot}4$ Å and diffract to at least 2·7 Å resolution.

Myosin light chain kinase (MLCKase‡) is a key enzyme in the regulation of contractile activity in smooth muscle (Hartshorne, 1987). Phosphorylation of the 20,000 dalton light chains of myosin initiates contraction. This event is coupled to the intracellular Ca^{2+} transients *via* the formation of the Ca^{2+}–calmodulin complex and its subsequent binding to, and activation of, MLCKase. The enzyme is widely distributed and is found in striated muscle and also in many non-muscle cells (Hartshorne, 1987). Initially, studies on the structure of MLCKase used limited proteolysis as a means of defining functional areas (Foyt *et al.*, 1985; Walsh, 1985). A linear model was proposed (Foyt *et al.*, 1985) in which the active site is located approximately in the center of the molecule and is flanked on its C-terminal side by the calmodulin-binding site. This prototype model remains valid but has been refined using the sequence of gizzard MLCKase obtained from the cloned complete cDNA (Olson *et al.*, 1990). The gizzard MLCKase molecule contains 972 residues with a calculated M_r of 107,534. The central active site spans residues D517 to R762 and the calmodulin-binding site, residues A796 to S815. Between the active site and the calmodulin-binding site, possibly overlapping with the latter, is a sequence thought to inhibit the apoenzyme. One theory is that the active site recognizes this sequence because of its similarity to the light chain substrate (Kemp *et al.*, 1987). If one accepts this hypothesis, the pseudosubstrate sequence would be S787 to M805. To the C-terminal side of the calmodulin-binding site is an extension of the molecule that is found in the smooth muscle MLCKase but not in its skeletal muscle counterpart (Takio *et al.*, 1986; Roush *et al.*, 1988). The function of this part of the molecule is not known but it is intriguing that an independent protein probably corresponding to the MLCKase sequence I819 to E972 (the initiation site is not established but is between M816 and M822; the most likely site is M818) is found in abundance in smooth muscle. This is termed telokin (Ito *et al.*, 1989).

Telokin was isolated first as a byproduct of calmodulin purification from chicken gizzard (Dabrowska *et al.*, 1977) but unlike calmodulin did not activate MLCKase. More recent studies (Ito *et al.*, 1989) have shown it to be an acidic protein, $PI = 4{\cdot}5$, that stained metachromatically with Stains-all. Its apparent M_r, based on SDS/polyacrylamide gel electrophoresis is 24,000, but is calculated to be about 18,000 from its sequence. It contains one phosphorylation site for cAMP-dependent protein kinase, corresponding to S828 of the MLCKase molecule. Telokin is abundant in chicken and turkey gizzard and its concentration is at least 15 μM, compared to about 4 μM for

† Author to whom correspondence should be addressed.

‡ Abbreviations used: MLCKase, myosin light chain kinase; MLCK, gene for MLCKase; PEG, polyethylene glycol; kb, 10^3 bases or base-pairs.

0022–2836/91/040621–03 $03.00/0

MLCKase (Ngai & Walsh, 1985). To date it has been identified only in smooth muscle tissue (chicken oviduct, bovine stomach and aorta; M. Ito unpublished observations). The most convincing evidence for the relationship of telokin to MLCKase is based on sequence determination of tryptic peptides isolated from telokin. For four peptides the sequences covering 57 amino acid residues were identical with the sequence derived from gizzard cDNA. The N terminus of telokin is blocked and therefore it was suggested that this portion of the MLCK gene is expressed independently as a unique protein. This tentative conclusion was based also on earlier studies (Guerriero *et al.*, 1986; Russo *et al.*, 1987) that showed that the 2·1 kb DNA (to MLCKase) hybridized to two sizes of DNA of 5·5 kb and 2·7 kb. The larger of these corresponded to the mRNA for part of the MLCKase molecule (Guerriero *et al.*, 1986) and it was proposed that the translation product of the smaller message is probably telokin.

Telokin was isolated using the protocol established by Ito *et al.* (1989). Initial crystallization trials were carried out using the hanging-drop vapor diffusion method with 4 μl of protein solution mixed with 4 μl of precipitant solution equilibrated against 1 ml of precipitant solution. A wide variety of precipitants and conditions were tried; however, crystals were obtained only from polyethylene glycol (PEG) at low pH. Ionic precipitants tried included ammonium sulfate, sodium/potassium phosphate and magnesium sulfate over a pH range of 5 to 9. It was observed that the protein was less soluble at low pH but generally precipitated between salt concentrations of 1 to 2 M. 2-Methyl-3,4-pentanediol was also tried as a precipitant at low pH; however, this did not yield crystals. Solutions of PEG 8000 in the range 4 to 9% containing 50 mM-succinic acid, adjusted to pH 4 with potassium hydroxide and 0·5 mM-sodium azide, gave small crystals that showed a distorted hexagonal bipyramidal morphology. A variety of nucleation inhibitors including 1% (w/v) dimethyl sulfoxide, 2-methyl-3,4 pentanediol (v/v), 1,4-dioxane (w/v), glycerol (v/v), PEG 400 (v/v) and 5 to 50 mM-$CdSO_4$ were tried in an attempt to obtain larger crystals. The addition of NaCl in the range 100 to 200 mM did not improve the size of crystals or alter their morphology. It was found that ethylene glycol in the range 1 to 5% did suppress the number of nucleations and produced larger crystals. The best crystals were obtained by mixing a 4 μl drop containing 7 mg telokin/ml, 1 mM-dithiothreitol, 5 mM-NaH_2PO_4 at pH 7, with 4 μl of 4% PEG 8000, 0·5 mM-sodium azide, 1% (v/v) ethylene glycol buffered with 50 mM-potassium succinate (pH 4). These were equilibrated against 1 ml of 4 to 6% PEG 8000, buffered with 50 mM-potassium succinate at pH 4, 0·5 mM-NaN_3, 1% (v/v) ethylene glycol at room temperature. A precipitate occurs immediately on mixing the protein with the polyethylene glycol solution. Close examination of the precipitate indicated that micro-globules of a protein-rich phase had separated at the interface between the polyethylene glycol and protein solution. Over the period of several hours these microdroplets coalesced to form an oily film on the surface of the hanging drop that did not disperse with time. Removal of this precipitate by centrifugation completely abolished the growth of crystals. It is presumed that the crystals grow from this precipitate, since as the crystals grew the magnitude of the precipitate appeared to diminish. Small crystals of dimensions 50 μm × 50 μm × 50 μm were generally observed within 24 hours and grew to full size within a week. This strategy produced crystals of dimensions 0·5 mm × 0·5 mm × 0·5 mm. Crystallization occurs close to the isoelectric point for this protein, indeed dialysis of the protein against 50 mM-potassium succinate at pH 4 results in precipitation of the protein, although crystals do not appear to grow under these conditions.

The crystals were mounted in thin-walled quartz capillaries and examined by precession photography using CuKα radiation from a Rigaku RU200 rotating anode generator operated at 50 kV, 50 mA equipped with a 200 μm focal cup. Precession photographs (13°, crystal to film distance 100 mm, limit of resolution 3·5 Å, (1 Å = 0·1 nm)) were recorded for the *hk*0, *hhl* and *h*0*l* zones to obtain the cell dimensions and crystal symmetry. In addition unscreened 3° precession photographs were recorded for the three major zones. Examination of the zero level diffraction patterns and distribution of intensities in the upper levels on the unscreened precession photographs revealed the presence of a 3_1 or 3_2 screw axis and three 2-fold axes that lie perpendicular to a^* and b^*. The 3-fold symmetry was also clearly observable in the crystal morphology. This established that the crystals belong to the trigonal space group $P3_121$ or $P3_221$, with cell dimensions of $a = 64{\cdot}0$ Å and $c = 59{\cdot}4$ Å. On the assumption that there is one molecule of molecular weight 18,000 per asymmetric unit, the mass per unit volume for these crystals, V_m, is 1·95 Å^3/dalton, which is in the normal range for globular proteins (Matthews, 1968). Still photographs show that useful diffraction data can be recorded to at least 2·7 Å resolution. The crystals exhibit a low mosaic spread and are resistant to radiation damage. Even though the crystals were grown at room temperature, they can be cooled to 4 °C without any damage to their diffraction properties. With cooling, crystals continue to diffract to 2·8 Å after 24 hours of exposure to the X-ray beam. A search for isomorphous heavy-atom derivatives is in progress.

MLCKase is a large protein for which there is considerable evidence of conformational flexibility. In addition, the intact protein is highly susceptible to adventitious proteolysis, which for many years gave rise to variable estimates of its molecular weight. These attributes also make it difficult to crystallize the complete molecule. For this reason it is appropriate to study smaller fragments that contain the functional domains, with the goal of

reconstructing the intact molecule. The crystallization of telokin is the first step toward this goal.

This work was supported by grants from the National Institutes of Health to D.J.H. (HL23615, HL20984) and I.R. (AR35186). T.A.A. was supported by an NIH predoctoral training grant (GM08293).

References

Dabrowska, R., Aromatorio, D., Sherry, J. M. F. & Hartshorne, D. J. (1977). Composition of the Myosin Light Chain Kinase from Chicken Gizzard. *Biochem. Biophys. Res. Commun.* **78**, 1263–1272.

Foyt, H. L., Guerriero, V., Jr & Means, A. R. (1985). Functional Domains of Chicken Gizzard Myosin Light Chain Kinase. *J. Biol. Chem.* **260**, 7765–7774.

Guerriero, V., Jr, Russo, M. A., Olson, N. J., Putkey, J. A. & Means, A. R. (1986). Domain Organization of Chicken Gizzard Myosin Light Chain Kinase Deduced from a Cloned cDNA. *Biochemistry*, **25**, 8372–8381.

Hartshorne, D. J. (1987). Biochemistry of the Contractile Process in Smooth Muscle. In *Physiology of the Gastrointestinal Tract* (Johnson, L. R., ed.), 2nd edit., vol. 1, pp. 423–482, Raven Press, New York.

Ito, M., Dabrowska, R., Guerriero, V., Jr & Hartshorne, D. J. (1989). Identification in Turkey Gizzard of an Acidic Protein Related to the C-terminal Portion of Smooth Muscle Myosin Light Chain Kinase. *J. Biol. Chem.* **264**, 13971–13974.

Kemp, B. E., Pearson, R. B., Guerriero, V., Jr, Bagchi, I. C. & Means, A. R. (1987). The Calmodulin Binding Domain of Chicken Smooth Muscle Myosin Light Chain Kinase Contains a Pseudosubstrate Sequence. *J. Biol. Chem.* **262**, 2542–2548.

Matthews, B. M. (1968). Solvent Contents of Protein Crystals. *J. Mol. Biol.* **33**, 491–497.

Ngai, P. K. & Walsh, M. P. (1985). Properties of Caldesmon Isolated from Chicken Gizzard. *Biochem. J.* **230**, 695–707.

Olson, N. J., Pearson, R. B., Needleman, D. S., Hurwitz, M. Y., Kemp, B. E. & Means, A. R. (1990). Regulatory and Structural Motifs of Gizzard Myosin Light Chain Kinase. *Proc. Nat. Acad. Sci., U.S.A.* **87**, 2284–2288.

Pearson, B. R., Wettenhall, R. E. H., Means, A. R., Hartshorne, D. J. & Kemp, B. E. (1988). Autoregulation of Enzymes by Pseudosubstrate Prototypes: Myosin Light Chain Kinase. *Science*, **241**, 970–973.

Roush, C. L., Kennelly, P. J., Glaccum, M. B., Helfman, D. M., Scott, J. D. & Krebs, E. G. (1988). Isolation of the cDNA Encoding Rat Skeletal Muscle Myosin Light Chain Kinase, Sequence and Tissue Distribution. *J. Biol. Chem.* **263**, 10510–10516.

Russo, M. A., Guerriero, V., Jr & Means, A. R. (1987). Hormonal Regulation of a Chicken Oviduct Messenger Ribonucleic Acid that Shares a Common Domain with Gizzard Myosin Light Chain Kinase. *Mol. Endocrinol.* **1**, 60–67.

Takio, K., Blumenthal, D. K., Walsh, K. A., Titani, K. & Krebs, E. G. (1986). Amino Acid Sequence of Rabbit Skeletal Myosin Light Chain Kinase. *Biochemistry*, **25**, 8049–8057.

Walsh, M. P. (1985). Limited Proteolysis of Smooth Muscle Myosin Light Chain Kinase. *Biochemistry*, **24**, 3724–3730.

Edited by P. Wright