THE AMINO ACID SEQUENCE OF PROFILIN FROM CALF SPLEEN

L.-E. NYSTRÖM, Uno LINDBERG, John KENDRICK-JONES+ and Ross JAKES+
Cemo-group, Wallenberg Laboratory, Uppsala University, Box 562, 751 22 Uppsala, Sweden and
+*MRC Laboratory of Molecular Biology, Cambridge, England*

Received 27 February 1979

1. Introduction

Many kinds of cells contain a large pool of unpolymerised actin and there is evidence for conversion of unpolymerised to filamentous actin as a response to stimuli resulting in cell movements ([1–4], reviewed [5,6]). An unpolymerised form of actin has been purified from extracts of various kinds of cells as a 1:1 complex with a protein called profilin, which prevents the polymerisation of the actin. Profilin, isolated from calf spleen, thymus and brain and from human platelets can combine with rabbit skeletal muscle actin and prevent its polymerisation ([7,8], Blikstad, Sundquist and Eriksson, unpublished results). The profilin : actin complex (profilactin) from calf spleen has been crystallized [9] and crystals suitable for high resolution X-ray crystallography have been obtained. The amino acid sequence of actin and profilin will be useful in solving the crystal structure of profilactin and for the detailed characterisation of its biochemical properties. The amino acid sequence of actin from various sources has been reported [10,11]. This paper describes the amino acid sequence of profilin from calf spleen.

2. Materials and methods

Profilactin from calf spleen was purified as in [9]. Profilin was isolated from profilactin either by the 'urea method' [7] or by dissociation of the complex with 5 mM EGTA [12].

Profilin (1 μmol) was carboxymethylated with iodo-[^{14}C]acetic acid by the method in [13]. Modification of lysine residues with citraconic anhydride was performed using the conditions in [14].

Cleavage of carboxymethylated profilin (1 μmol) with cyanogen bromide was carried out at room temperature in 70% formic acid [15]. The peptide mixture was fractionated on Sephadex G-25 SF (Pharmacia, Sweden) in 25% formic acid. The peptides of pool II (fig.1a) were further purified by high-voltage paper electrophoresis at pH 6.5 and 2.1 according to [16] and by chromatography in the second dimension using the BAWP solvent system [17]. Fragments of profilin were also generated with [^{14}C]cyanide using the cyanocysteine cleavage method [18] and fractionated on Sephadex G-25 SF in 50 mM NH_4HCO_3. Two peaks containing the peptides Cya 2 and 1, respectively, were eluted. The N-terminal region of the Cya 1 peptide was sequenced by mass spectrometry. The Cya 2 peptide was further digested with chymotrypsin and the peptides obtained were sequenced using the Dansyl-Edman method (see below). The remaining fragments were retained by the column, but could be eluted with 50% acetic acid. However, this material was not analyzed further.

Carboxymethylated freeze dried profilin (1 μmol) was dissolved in 8 M urea, 0.4 M Tris–HCl (pH 8.0) and diluted to 1 M urea, 5 mM $CaCl_2$ before digestion with chymotrypsin. Chymotrypsin (3 X crystallized, Worthington) was added to a substrate : enzyme ratio (w/w) of 100:1. The mixture was incubated at 37°C with stirring. After 4 h digestion more chymotrypsin was added (final substrate : enzyme ratio of 50:1, w/w) and incubation continued over night. The insoluble residue left after digestion was removed by centrifugation and washed 3 times with 50 mM

NH_4HCO_3. The soluble peptides were fractionated on a Sephadex G-25 SF column in 50 mM NH_4HCO_3 (fig.1b) and further purified by high-voltage paper electrophoresis and chromatography as outlined above.

Trypsin (Worthington) was further purified by affinity chromatography on soybean trypsin inhibitor coupled to Sepharose 4B [19]. Carboxymethylated profilin (1 μmol) was suspended in 2 M urea, 50 mM NH_4HCO_3 (pH 8.5), 20 mM $CaCl_2$ and incubated with trypsin at a substrate : enzyme ratio of 50:1 (w/w) for 4 h at 37°C. Soluble tryptic peptides were separated and further fractionated by electrophoresis and chromatography.

The large peptide cT was prepared by tryptic digestion of citraconylated profilin (1 μmol) using a substrate : enzyme ratio of 100:1 (w/w) at 37°C. After 4 h the reaction was stopped and unblocking was performed by incubating with 2 vol. 5% formic acid/1% pyridine at room temperature overnight. The digest was fractionated on a Sephadex G-25 column (16 × 900 mm) in 10 mM HCl.

Protease from *Staphylococcus aureus* strain V8 (Miles Biochemical Co.) was used for further digestion of peptide Ch3 in 50 mM NH_4HCO_3 (pH 8.5) for 4 h at 37°C at a substrate : enzyme molar ratio of 50:1. The C-terminal sequence of peptide Ch3 was determined using DFP-treated carboxypeptidase A (Worthington) as described by Ambler using a substrate : enzyme molar ratio of 50:1 [20].

The sequences of most peptides were determined by the Dansyl-Edman method [21] using thin-layer chromatography to identify the dansylated amino acid [22]. The amides were assigned by the method in [23]. The large peptide cT (fig.2) was sequenced on a Beckman model 890-B automatic sequencer using the 0.1 M Quadrol program [24]. The PTH amino acids from each step were identified by thin-layer chromatography [25] gas–liquid chromatography [26], and after back hydrolysis [27] the resulting amino acids were determined on a Durrum D-500

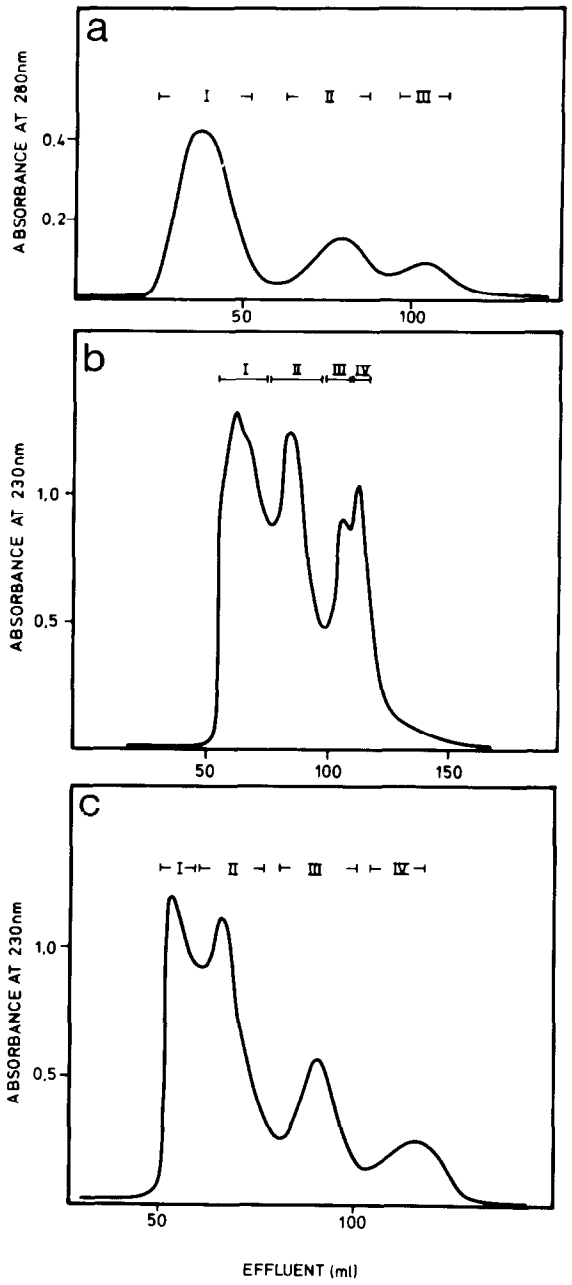

Fig.1. Fractionation of peptides from cleavage of profilin with cyanogen bromide, chymotrypsin and trypsin on Sephadex G-25 SF. (a) The cyanogen bromide peptides were separated on a 16 × 900 mm column in 25% formic acid. Fraction II contained peptides CNBr 1, 2 and 3 (see fig.2). (b) Chymotryptic peptides were fractionated on 25 × 500 mm column in 50 mM NH_4HCO_3. Fraction I contained peptides Ch3, Ch6 and Ch9: fraction II Ch5, Ch11; fraction III Ch4, Ch7, Ch8, Ch12; fraction IV Ch1, Ch2, Ch10. (c) Tryptic peptides were fractionated on a 25 × 500 mm column in 50 mM NH_4HCO_3. Fraction I contained peptides T1, T9, T10; fraction II T2, T5, T7; fraction III T6, T11; fraction IV T4, T8.

amino acid analyzer. The N-terminal region was determined for us by Drs Ann Dell and Howard Morris (Imperial College, London) using mass spectrometry [28].

3. Results and discussion

Profilin has mol. wt 15 000–16 000 as determined by SDS–polyacrylamide gel electrophoresis and it chromatographs as a monomer under non-denaturing conditions. It is a basic protein with a pI between pH 9.2 and 9.4. It has 2 tryptophanes and 4 tyrosines and its $E^{1\%}_{280}$ is 13. Profilin has 3 cysteine residues and its N-terminus is blocked. The amino acid composition of profilins from different sources are similar suggesting that profilin is a conserved protein [8].

The amino acid sequence of profilin was determined in the following way. Cleavage of carboxymethylated profilin with cyanogen bromide gave rise to peptides which were fractionated as shown in fig.1a. Three peptides (CNBr 1, 2, 3, see fig.2) were isolated from fraction II by paper electrophoresis and chromatography and sequenced. One of these peptides was the homoserine-lacking C terminal peptide. Fraction I contained peptides which proved difficult to separate and they were not analyzed further.

Carboxymethylated profilin was digested with chymotrypsin in 1 M urea. The addition of urea was essential as neither trypsin nor chymotrypsin would reproducibly degrade the profilin molecule in its absence. However, even after prolonged digestion with chymotrypsin, insoluble material remained. This material 'chymo-insoluble' was collected by centrifugation and analysed as described below. The soluble peptides were fractionated as described in section 2 and fig.1b. Peptides Ch2, Ch4, Ch5, Ch6, Ch7, Ch8, Ch10, Ch11 and Ch12 were sequenced by the manual Dansyl-Edman technique. The sequence of Ch9 was completed by sequencing the overlapping tryptic peptide T7 (see below). Only the first 6 amino acids of peptide Ch3 could be sequenced. This peptide was later shown to part of the N-terminal tryptic peptide T1. Its sequence and the sequence of the 'chymo-insoluble' remaining after digestion with chymotrypsin is described below.

Carboxymethylated profilin was digested with trypsin in the presence of 2 M urea. Insoluble material left after digestion contained one major peptide, T3. The supernatant was fractionated on Sephadex G-25 SF (fig.1c). The peptides T2, T4, T6, T8, T10 and T11 were sequenced. The sequence of T5 was established after comparison with the overlapping peptide Ch8. To obtain the sequences of T3 and T7 these peptides were further digested with chymotrypsin. The peptide T1 could not be sequenced using the Dansyl-Edman method. It had an amino acid composition similar to that of the chymotryptic peptide Ch3 and on chymotrypsin digestion T1 gave rise to Ch3, a peptide Ch2, free lysine and a tryptophane-containing peptide (Ch1). Sequencing of T1 by mass spectrometry gave Acetyl–Ala–Gly–Trp–Asn–Ala. This positioned Ch1 and Ch2. The overlap between Ch2 and Ch3 was obtained by the isolation of a peptide (res 4–10) arising from a partial cleavage of profilin with chymotrypsin. The sequence around position 20 was ascertained by the analysis of a peptide (res 19–23) obtained from peptide Ch3 by digestion using staphylococcal protease. Carboxypeptidase digestion gave the C-terminal sequence (res 20–24). The overlap between T1 and T2 was confirmed by tryptic cleavage of the cyanolation fragment Cya 2.

The 'chymo-insoluble' contained all the missing threonine residues of the molecule, and the overall composition suggested that it included the region which remained to be sequenced. It was heterogenous and digestion of citraconylated 'chymo-insoluble' with trypsin yielded small amounts of two peptides (res 66–88, res 84–88) together with a large peptide which became insoluble after removal of blocking groups. The N-terminal of the insoluble peptide indicated that it started at residue 89, but further sequence data was very difficult to obtain. The sequence of this region was obtained as follows. Peptides obtained by digesting the citraconylated profilin with trypsin were found to be soluble after unblocking the lysine residues and they could be fractionated on Sephadex G-25 SF in 10 mM HCl (data not shown). The material appearing with the void volume contained the expected peptide (cT, res 89–138) together with small amounts of a peptide representing T5 and T6. The solubility property of the peptide cT was radically different from that of 'chymo-insoluble' and its sequence from res 89–123 was determined using the spinning

cup sequencer. This gave an overlap into the known C-terminal sequence.

Profilin contains 142 amino acid residues giving it mol. wt 15 220. The N-terminal amino acid is acetylated. Using the computer program in [29,30], the primary structure of profilin was compared with the known sequence of proteins related to the actomyosin system: actin, light chains from rabbit skeletal muscle myosin and molluscan myosin, muscle tropomyosin, troponin subunits and pancreatic DNase I. No sequence homologies were found. The actin binding region of troponin I has been identified [31] and contains a high proportion of basic amino acids (a net positive charge at pH 7 or 8). The C-terminal region of profilin also contains a high net positive charge. This is of interest as the basic region of troponin I is suggested [31] to interact with the acidic N-terminus of actin.

A secondary structure of profilin was determined using a computer program based on the prediction rules in [32]. As shown in fig.2 the N-terminal half of the profilin molecule is suggested to contain mainly β-sheet structure and the C-terminal half to to contain α-helical structure. We eagerly await the tertiary structure of profilactin to test these predictions.

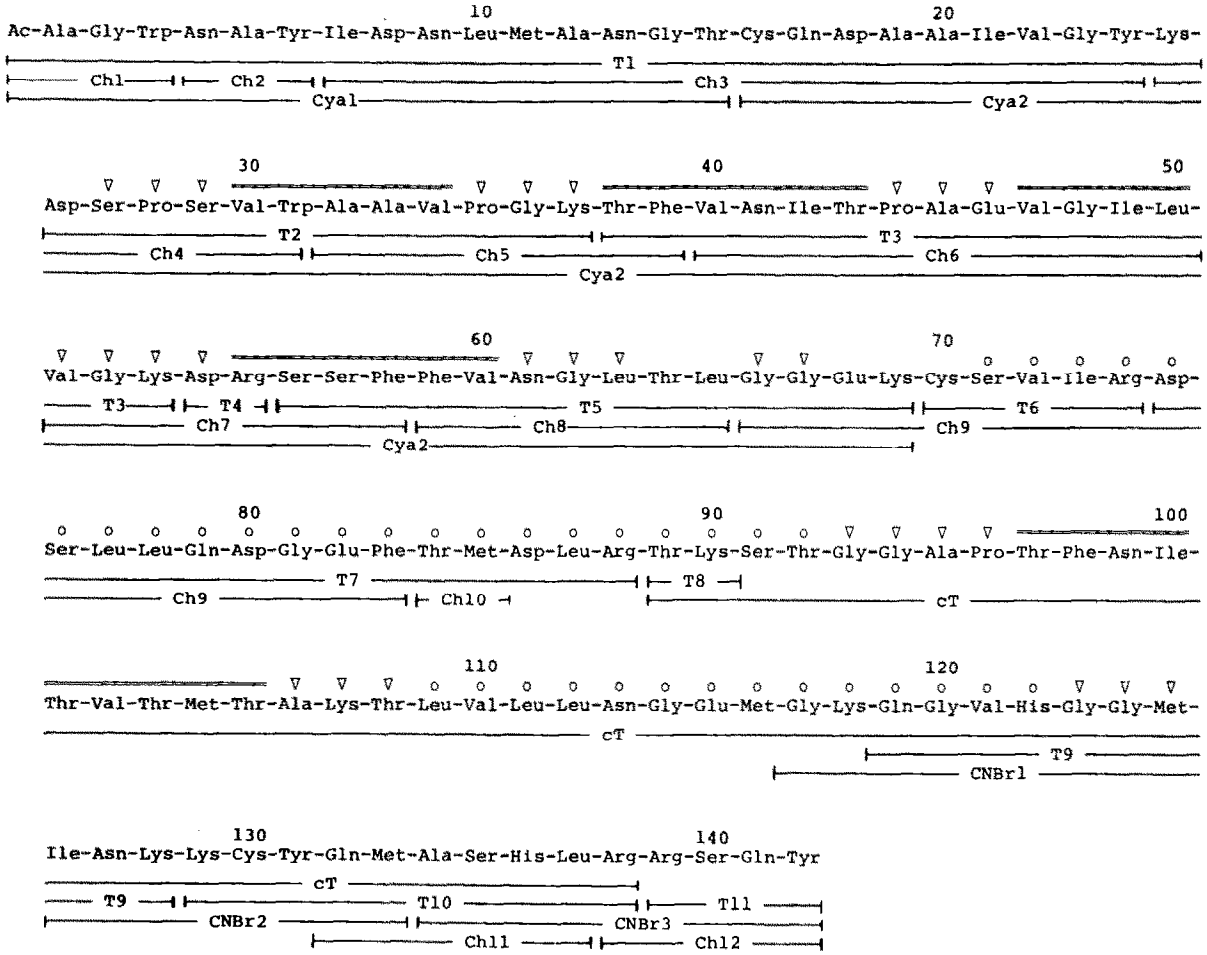

Fig.2. The amino acid sequence of calf spleen profilin. *Abbreviations:* chymotryptic peptides, Ch; tryptic peptides, T; cyanogen bromide fragments, CNBr; cyanolation fragments, Cya; and citraconylated tryptic fragment, cT.
The terminology used in the secondary structure predictions is (○) for α-helix, (=) for β-sheet, (▽) for turn.

Acknowledgements

The main part of the work was carried out at the MRC Laboratory of Molecular Biology, Cambridge and we thank Dr Hugh Huxley for continued interest and facilities. The profilactin used in this work was prepared for us by Håkan Larsson at the Wallenberg Laboratory, Uppsala. We gratefully acknowledge his efforts and the technical assistance of Christine Smalley. We are grateful to Drs Ann Dell and Howard Morris for determination of the N-terminal sequence, Dr Andrew McLachlan for the secondary structure prediction and Drs Lars Carlsson and John Walker for many constructive discussions. Finally we thank Elisabeth Aujalay for secretarial work. U.L. and L.E.N. acknowledge financial support from Uppsala University Reserve Fund and the Swedish Cancer Society.

References

[1] Tilney, L. G. (1976) J. Cell Biol. 69, 73–89.
[2] Behnke, O. (1976) in: Contractile Systems in Non-Muscle Tissues (Perry, S. V. et al. eds) pp. 105–115, Elsevier/North-Holland Biomedical Press, Amsterdam, New York.
[3] Wohlfart-Botterman, K. E. and Isenberg, G. (1976) in: Contractile Systems in Non-Muscle Tissues (Perry, S. V. et al. eds) pp. 297–308, Elsevier/North-Holland Biomedical Press, Amsterdam, New York.
[4] Markey, F. and Lindberg, U. (1979) in: 36th Colloq. Protides of the Biological Fluids (Peeters, ed) in press.
[5] Korn, E. D. (1978) Proc. Natl. Acad. Sci. USA 75, 588–599.
[6] Lindberg, U., Carlsson, L., Markey, F. and Nyström, L. (1979) in: The Cytoskeleton in Normal and Pathological Processes (Gabbiani, ed) S. Karger AG (in press).
[7] Carlsson, L., Nyström, L., Sundkvist, I., Markey, F. and Lindberg, U. (1977) J. Mol. Biol. 115, 465–483.
[8] Markey, F., Lindberg, U. and Erikson, L. (1978) FEBS Lett. 88, 75–79.
[9] Carlsson, L., Nyström, L., Lindberg, U., Kannan, K. K., Cid-Dresdner, H., Lövgren, S. and Jörnvall, H. (1976) J. Mol. Biol. 105, 353–366.
[10] Elzinga, M. and Lu, R. C. (1976) in: Contractile Systems in Non-Muscle Tissues (Perry, S. V. et al. eds) pp. 29–37, Elsevier/North-Holland Biomedical Press, Amsterdam, New York.
[11] Vanderkerckhove, J. and Weber, K. (1978) Eur. J. Biochem. 90, 451–462.
[12] Kendrick-Jones, J., Jakes, R., Nyström, L. E. and Lindberg, U. (1979) in: 36th Colloq. Protides of the Biological Fluids (Peeters, ed) in press.
[13] Szent-Györgyi, A. G., Szentkiralyi, E. M. and Kendrick-Jones, J. (1973) J. Mol. Biol. 74, 179–203.
[14] Butler, P. T. G. and Hartley, B. S. (1972) Methods Enzymol. 25, 191–199.
[15] Steers, E., Craven, G. R., Anfinsen, C. B. and Bethne, J. L. (1965) J. Biol. Chem. 240, 2478–2485.
[16] Michl, H. (1951) Mh. Chem. 82, 489–497.
[17] Waley, S. G. and Watson, J. (1953) Biochem. J. 55, 328–346.
[18] Jacobsen, G. R., Schaffer, M. H., Stark, G. R. and Vanaman, T. C. (1973) J. Biol. Chem. 248, 6583–6591.
[19] Affinity chromatography, principles and methods (1976) pp. 56–57, Pharmacia, Sweden.
[20] Ambler, R. P. (1972) Methods Enzymol. 11, 179–186.
[21] Gray, W. R. and Hartley, B. S. (1963) Biochem. J. 89, 59.
[22] Woods, K. R. and Wang, K. T. (1967) Biochim. Biophys. Acta 133, 369–370.
[23] Offord, R. E. (1966) Nature 211, 591–593.
[24] Brauer, A. W., Margolies, M. N. and Haber, E. (1975) Biochemistry 14, 3029–3035.
[25] Terhorst, C., Möller, W., Laursen, R. and Wittman-Liebold, B. (1973) Eur. J. Biochem. 34, 138–152.
[26] Pisano, J. J., Bronzert, T. J. and Brewer, H. B. (1972) Anal. Biochem. 45, 43–59.
[27] Mendez, E. and Lai, C. Y. (1975) Anal. Biochem. 68, 47–53.
[28] Morris, H. R., Williams, D. H. and Ambler, R. P. (1971) Biochem. J. 125, 189–201.
[29] Staden, R. (1977) Nucl. Acids Res. 4, 4037–4051.
[30] Staden, R. (1978) Nucl. Acids Res. 5, 1013–1015.
[31] Syska, H., Wilkinson, M. J., Grand, R. J. A. and Perry, S. V. (1976) Biochem. J. 153, 375–387.
[32] McLachlan, A. D. (1977) Int. J. Quant. Chem. 12 suppl. 1, 371–385.