Studies on secondary structure of caldesmon and its C-terminal fragments

Edward A. CZURYŁO,* Sergei Yu. VENYAMINOV† and Renata DĄBROWSKA*‡

*Nencki Institute of Experimental Biology, Department of Muscle Biochemistry, 3 Pasteur Str., 02-093 Warszawa, Poland, and †Institute of Protein Research, Pushchino, Russia

Evaluation of the secondary structure of caldesmon from c.d. spectra revealed that it contains 51 % helix, 9 % β -strand and 40 % of remainder structures. These values agree well with the predicted ones from amino acid sequence, assuming an extended chain structure for caldesmon. The estimates of the secondary-structure elements in C-terminal 34 kDa and 19 kDa fragments

are: 11 and 12 % helix, 22 and 20 % β -strand, 13 and 17 % β -turns and loops, and 54 and 50 % of remainder structure respectively. The best fit of experimental data was obtained assuming the globular state of the fragments. On the basis of structural analysis and fragmentation by proteolytic and chemical cleavages the six-domain model of caldesmon is proposed.

INTRODUCTION

Caldesmon is a protein actively involved in the regulation of smooth-muscle contraction (for reviews, see [1,2]). It binds to the protein components of contractile apparatus (actin, myosin and tropomyosin) and affects cycling of cross-bridges between actin and myosin [3,4]. Functioning of caldesmon can be regulated by Ca^{2+} and calmodulin.

Smooth-muscle caldesmon is a thin elongated molecule approx. 76 nm in length and 2 nm in diameter [5]. The central portion is 30–40 nm long and serves as a 'spacer' between two active domains each about 20 nm long with more flexible conformations [6]. Since the latter domains are involved in the binding of caldesmon to actin, tropomyosin, calmodulin and myosin, it is believed that they are responsible for the functional properties of caldesmon.

Previous studies on the secondary structure of caldesmon have yielded a range of results with estimates of helicity from c.d. spectroscopy of 10 % [7] or 40 % [8], whereas the value predicted from the amino acid sequence was 80–85 % [9].

There are several methods commonly used for evaluation of the secondary structure of protein from c.d. spectra (for a review, see [10]). However, almost all of them are based on the X-ray structure of globular proteins. In order to account for the rod-like shaped caldesmon, we have used a modification of the method of Provencher and Glöckner [11] in which 16 reference c.d. spectra of globular proteins were extended to a total of 20, including four with non-globular structure [12]. The original method [11] was used for evaluation of the secondary structure of C-terminal 34 kDa and 19 kDa fragments of caldesmon that, like the intact protein, bind to actin and inhibit actin-activated ATPase activity of myosin [13,14]. The estimations of the secondary structure were compared with the data predicted from primary structure of caldesmon and its C-terminal fragments on the basis of the ALB algorithm [15].

MATERIALS AND METHODS

Preparation of proteins

Chicken gizzard caldesmon was prepared by the procedure of Bretscher [16]. The C-terminal 34 kDa fragment was obtained by chymotryptic cleavage of caldesmon as described by Szpacenko and Dąbrowska [13] and purified as described by Czuryło et al. [17]. The C-terminal 19 kDa (25 kDa from SDS/PAGE) frag-

ment of caldesmon was prepared by chemical cleavage by the method of Nefsky and Bretscher [14]. The concentration of the caldesmon and its fragments was calculated from nitrogen content determined by a micromethod of Jaenicke [18].

U.v. spectroscopy

Absorption spectra of caldesmon or its fragments were recorded on a Cary 219 (Varian) spectrophotometer over wavelength range 240–400 nm at 20 °C at a protein concentration range of 0.6–1.4 mg/ml in 5 mm rectangular cell; an analogous cell containing solvent alone was placed in the reference beam. U.v. spectra of caldesmon and its C-terminal fragments were recorded in 50 mM cacodylate/NaOH buffer, pH 7.0, containing 0.5 mM β -mercaptoethanol, 0.5 mM EGTA and 0.02 mM NaN_3 . Calculations of the absorption coefficients were corrected for solution turbidity as described in [19] by plotting the dependence of log absorbance of the solution versus log wavelength and extrapolating the linear dependence between these values over the absorption range of 320–400 nm to 240–300 nm. Absorption coefficients obtained from experimental data were compared with those calculated using the values of the molar absorption coefficients of tyrosine and tryptophan (1405 and 5579 $\text{M}^{-1} \cdot \text{cm}^{-1}$ respectively [20]; see Table 1).

C.d. spectroscopy

C.d. spectra of caldesmon and its C-terminal fragments were recorded, in the same buffer as u.v. spectra, on a J-600 (Jasco) spectropolarimeter in 0.093-mm and 0.185-mm-path-length cylindrical cells in the far-u.v. region (184–250 nm), and 10-mm-path-length cylindrical cell in the near-u.v. region (240–310 nm) over a protein concentration range of 0.6–1.4 mg/ml. Before c.d. measurements were made, the protein solutions were clarified by centrifugation (20000 *g* for 15 min). Recording of each spectrum, including the baseline, was repeated twice. The c.d. spectra were calculated by averaging the values obtained in cells of different pathlengths at different concentrations of several independent batches of caldesmon or its fragments.

Data are presented in terms of molar ellipticity per one amino acid residue calculated according to the equation:

$$[\theta]_{\text{m.r.w.}}(\lambda) = \frac{\theta(\lambda) \cdot 10^3 \cdot \text{m.r.w.}}{l \cdot c} (\text{degrees} \cdot \text{cm}^2 \cdot \text{dmol}^{-1})$$

‡ To whom correspondence should be addressed.

Table 1 Absorption coefficients (and parameters used for their calculation) of caldesmon and its C-terminal fragments

Protein/fragment	N content (%)	Molecular mass (kDa)	$A_{279}^{0.1\%}$	$A_{\text{theor.}}$	$10^{-3} \times \epsilon_{\text{exptl.}}$	$10^{-3} \times \epsilon_{\text{theor.}}$
Caldesmon (5 Trp, 3 Tyr)	18.53	86.8	0.403 ± 0.017	0.370	35.1 ± 1.4	32.1
34 kDa fragment (3 Trp, 1 Tyr)	18.43	33.9	0.632 ± 0.022	0.535	21.4 ± 0.7	18.1
19 kDa fragment (3 Trp, 1 Tyr)	17.70	18.7	1.130 ± 0.025	0.969	21.2 ± 0.5	18.1

where $\theta(\lambda)$ denotes degrees of ellipticity and recorded by the instrument at wavelength λ , m.r.w. is the mean residue weight (determined from amino acid sequence [9,21]), l is the cell path length in mm and c is the concentration of protein in mg/ml. Measurements of the c.d. spectra were carried out over the range of 184–320 nm at a spectral slit-width of 1 nm. The cells were fixed in a home-built thermostatically controlled holder. Temperature was maintained at 10.0 ± 0.1 °C with a water thermostat. Calibration of wavelength and scale of spectropolarimeter were done according to the procedure described elsewhere [22].

Evaluation of the secondary structure from c.d. spectra

The secondary structure of the C-terminal fragments of caldesmon was calculated by the method of Provencher and Glöckner [11] based on linear combination of c.d. spectra of 16 referenced globular proteins of known structure. This method includes the constraints $\sum f_i = 1$ and $1 \geq f_i \geq 0$, where f_i denotes the fraction of the i th conformation, i.e.: helix, β -strand, β -turns and remainder (unordered) structure. To determine the secondary structure of caldesmon we have used the modified method of Provencher and Glöckner [11], in which the number of reference spectra was extended to 20 [12]. These additional spectra originated from: oxidized ribonuclease A (pH 2.0 at 20 °C), staphylococcal nuclease (pH 3.0 at 6 and 70 °C) and the peptide Glu-Lys-Lys-Leu-Glu-Glu-Ala (pH 2.0 at 10 °C) which, under these conditions, do not form tertiary structure [23].

Prediction of the secondary structure

The algorithm ALB [15] was used for prediction of the secondary structure of caldesmon and its C-terminal fragments from the amino acid sequence. This non-statistical method uses short- and middle-range interactions to predict the secondary structure and long-range interactions (in the simplest version of the 'floating logs' model) to predict tertiary structure. Therefore, the algorithm ALB offers possibility of predicting the structure of polypeptide chains with and without globular structure, and was successfully applied for many proteins (for details, see [24]).

RESULTS

C.d. spectroscopy

C.d. spectra of intact caldesmon and its C-terminal 34 kDa and 19 kDa fragments are presented in Figure 1(a). The c.d. spectrum of caldesmon showing double minimum ($[\theta]_{221} = -19\,500$ and $[\theta]_{206} = -23\,100$ degrees $\cdot \text{cm}^2 \cdot \text{dmol}^{-1}$) and a maximum ($[\theta]_{191} = 26\,300$ degrees $\cdot \text{cm}^2 \cdot \text{dmol}^{-1}$) suggests high content of helical structure in this molecule. The c.d. spectra of C-terminal 34 kDa and 19 kDa fragments have very weak first minima at 222 nm and blue-shifted second minima $[\theta]_{200} = -13\,100$ and $[\theta]_{199} = -21\,400$ degrees $\cdot \text{cm}^2 \cdot \text{dmol}^{-1}$ respectively, whereas the maxima are beyond the instrument range.

Figure 1(b) illustrates the dependence of mean residue ellipticity at 222 nm on temperature. Neither the intact caldesmon nor its fragments show co-operative transition over the range of 0–90 °C. However, a linear increase in the mean residual ellipticity for intact caldesmon and its 34 kDa fragment indicate continuous unfolding of their secondary structures. The increase of mean residual ellipticity for the 34 kDa fragment is much less steep than for whole caldesmon. In the case of the 19 kDa fragment above room temperature, a linear slow decrease without change in slope of the ellipticity curve takes place. Both the intact caldesmon and its fragments undergo complete refolding after cooling to room temperature (results not shown).

Secondary-structure estimation and prediction

The estimates of the secondary-structure elements of intact caldesmon and its fragments from the experimental data shown in Figure 1(a) are presented in Table 2, together with the values predicted from the amino acid sequence. In the case of intact caldesmon the best agreement of the experimental data with the predicted values is obtained by assuming an extended state. In this state caldesmon has 51 % of amino acid residues involved in helical structure, 11 % in β -strands, 9 % in β -turns and loops, while the remaining 30 % of the protein is unstructured. If caldesmon is assumed to be in a globular state, then the percentage of helical and β -structural amino acid residues would decrease to about one-half of that in a rod-shaped model, and unstructured amino acid residues would increase to about 70 %. The probable locations of the helical and β -strands regions of caldesmon are presented in Figure 2.

On the basis of the secondary-structure prediction of caldesmon it can be supposed that it is composed of six structurally different domains (Figure 3 and Table 3). Domain 1 has about 53 % helical structure and contains five helices of different length and one short β -strand region, while domains 3 and 5 form strong helices, composed of 172 and 67 residues respectively. Domains 2 and 6 have α/β character. In the former there is more β -strand than helix, while in the latter the converse is true. Both domains may also be folded into a globule-like structure. Domain 4 is less structured and contains two short helical regions.

The C-terminal 34 kDa fragment of caldesmon is comprised of the second half of domain 4 and domains 5 and 6, whereas the 19 kDa fragment is solely domain 6, missing only the 14 N-terminal amino acid residues (Figure 3). The results of the secondary-structure predictions for these fragments are the closest to those calculated from the c.d. spectra, assuming their globular conformation (Table 2). Predictions made assuming an extended structure for these fragments gave values significantly different from the experimental data and also from the values predicted for the fragments in the intact protein (Table 2). The probability of distribution of the predicted structured regions of both fragments are presented in Figure 4.

Table 2 also shows predictions of the secondary structure for

Figure 1 C.d. spectroscopy of caldesmon and its C-terminal fragments

(a) C.d. spectra and (b) thermal unfolding of caldesmon (—), the 34 kDa fragment (---) and the 19 kDa fragment (.....). All measurements were performed in 50 mM cacodylate buffer, pH 7.0, containing 0.5 mM EGTA, 0.5 mM β -mercaptoethanol and 0.02 mM NaN_3 in the range of the protein concentrations of 0.6–1.4 mg/ml.

Table 2 Comparison of the secondary structure of caldesmon and its fragments calculated from c.d. spectra and predicted from amino acid sequence

^aEstimations of the secondary structure from c.d. spectra were performed by the method of Provencher and Glöckner, unmodified [11] in the case of 34 kDa and 19 kDa fragments and modified (spectrum database was extended with the data for extended polypeptide chain) in the case of caldesmon. Predictions were based on the Ptitsyn and Finkelstein [15] ALB algorithm assuming: ^bextended structure of caldesmon and its fragments in the native molecule, ^cextended structure of isolated fragments and ^dglobular structure of isolated fragments.

Protein/fragment	Secondary structure (%)				Method of estimation
	Helix	β -Strand	β -Turns	Unordered	
Caldesmon	51	9	0	40	Experimental ^a
	51	11	9	30	Predicted ^b
34 kDa fragment	16	20	13	51	Experimental ^a
	42	10	10	34	Predicted ^b
	38	28	17	17	Predicted ^c
	21	16	15	48	Predicted ^d
19 kDa fragment	12	20	17	51	Experimental ^a
	28	14	18	40	Predicted ^b
	28	34	20	18	Predicted ^c
	15	22	21	42	Predicted ^d
33 kDa (CT54) fragment	55–58	ND	ND	ND	Experimental [25]
	66	6	7	11	Predicted ^b
	71	11	3	15	Predicted ^c

central 33 kDa (CT54) fragment of caldesmon. This fragment covers approximately the C-terminal half of domain 2, domain 3 and the N-terminal half of domain 4 (residues 166–450) of

caldesmon. Figure 5 shows probability of distribution of secondary-structure elements of the 33 kDa fragment isolated and in intact caldesmon.

DISCUSSION

Our estimates, based on c.d. spectra, of the types of secondary structures found in caldesmon and the degree to which they are formed are internally consistent if two criteria are met: a rod-like protein shape must be assumed, and the basis set of c.d. spectra used to fit our data must be extended to include non-globular proteins. The extended Provencher and Glöckner method [11] based on linear combination of c.d. spectra of proteins of known secondary structure seemed to be the most suitable, since the estimates of all secondary-structure elements of caldesmon using this method appeared to be in good agreement with the values predicted from amino acid sequence of caldesmon assuming a non-globular state. Consequently, long-distance intramolecular interactions should not play any important role in the process of formation of the secondary structure of caldesmon molecule.

The absence of long-distance interactions in caldesmon polypeptide chain is confirmed by the results of temperature unfolding experiments showing that structural transitions occur in a gradual linear fashion without a clearly identifiable transition region. The melting process occurs in a step-by-step fashion starting at less stable structural regions and finishing at the most stable ones without any co-operativity. Similar temperature melting features of caldesmon were reported also by Wang et al. [25].

Our results prompted us to conclude that caldesmon molecule contains six distinct structural domains. The model of domain structure of caldesmon proposed recently by Marston and Redwood [1] on the basis of structural analysis and proteolytic cleavage of this protein takes into account only four structural domains devoid of unstructured regions susceptible for proteolytic attack. The main difference between these two models is in the location of domain 1 and in the presence of additional

Figure 2 Prediction of the secondary-structure elements of caldesmon assuming an extended state

The probability of helix (—) and β -strand (---). H, S and T show the positions of helical (H), β -strand (S) and turns or loop (T) regions respectively (these symbols are also used in Figures 4 and 5). Lines denote low probability, other symbols high probability of respective structures. Amino acid residues are numbered as described by Bryan et al. [9].

Figure 3 Structural domains of caldesmon

The six domains model is based on analysis of predicted structure of caldesmon (see Figure 2) and its proteolytic and chemical cleavages shown below the model. Calculated molecular masses of fragments with apparent ones derived from SDS/polyacrylamide gels (in parentheses) are used to name the peptides. ▨, helical domain, ▩, α/β domain, □, mostly unstructured domain.

Table 3 Characteristics of caldesmon domains

Domain no.	Amino acid residue no.	No. of residues	Helix			β -Strand		
			Region	No. of residues	Percentage in domain	Region	No. of residues	Percentage in domain
1	1–110	110	5	58	53	1	5	5
2	111–235	125	2	14	11	6	34	27
3	236–410	175	1	172	98	0	0	0
4	411–490	80	2	20	25	0	0	0
5	491–565	75	1	67	90	0	0	0
6	566–756	191	6	48	25	6	29	15

Figure 4 Prediction of the secondary structure elements of the C-terminal fragments of caldesmon

(a) Isolated 34 kDa fragment assuming a globular state. (b) Isolated 19 kDa fragment assuming a globular state. (c) The 34 kDa and 19 kDa fragments in intact caldesmon (compare with Figure 2). Amino acid residues are numbered as in intact caldesmon. Descriptions are as in legend to the Figure 2.

domain 2 in our model. Domain 3 in our model and domain 2 in four-domains model containing 'spacer' are basically the same. Similarly, domain 5 in the six-domains model corresponds to domain 3 in the four-domains model. Unstructured region

Figure 5 Prediction of the secondary-structure elements of the central 33 kDa (CT54) fragment of caldesmon

(a) Assuming an extended state. (b) Structure in intact caldesmon (compare with Figure 2). Amino acid residues are numbered as in intact caldesmon. Descriptions are as in the legend to the Figure 2.

between domains 3 and 4 in four-domains model are included into domain 6 in six-domains model. The location of helical domains (1, 3 and 5) in our model is in good agreement with the predictions of Wang et al. [25], calculated according to Gibrat et al. [27].

Lack of unstructured peptides connecting domains in our model enables a description of the structure of non-muscle isoforms of caldesmon which are devoid of a central, helical domain. In the case of four-domains model [1] removal of domain 2 evokes complication in extension of unstructured domain 1–3-linked region to around 70–100 amino acid residues (depending on the source of non-muscle caldesmon). The existence of such a long, unstructured region in the centre of the protein is very improbable. In our model the unstructured region spans only 30 amino acid residues. Moreover, detailed analysis of these parts of polypeptide chain of caldesmon that connect domains (in four-domains model) suggests that they are not always unstructured (see Figure 2). However, existence of loops and turns in these regions facilitates proteolytic cleavage. For

example, as is shown on Figure 3, the first split of caldesmon by chymotrypsin takes place at the Trp⁴⁵⁰ of unstructured domain 4 [13]. Further chymotryptic digestion, which produces the central 33 kDa and 15 kDa fragments, occurs within domains 2 and 6, which are rich in turns and loops [25,28]. Thrombic cleavage of caldesmon producing 15 kDa fragment [29] takes place in the vicinity of the border of domain 5, and the cleavage sites at Lys⁴⁹⁷ and Arg⁵⁹³ are located between structured regions.

It should be noted that, for the C-terminal 150 amino acid residues of caldesmon, the probability of both helix and β -strand formation, although higher than in previous predictions [1,25,26], is still very low (Figure 4). Moreover, the predicted secondary structure of the isolated C-terminal 34 kDa and 19 kDa fragments (assuming their globular state) differs from the structure of these fragments in intact caldesmon (Table 2; Figure 4). These complications originate from the imperfection of the program for this particular sequence. However, it cannot be excluded that differences in structure mentioned are due to the isolation procedure.

Comparison of the predicted helical content of the central 33 kDa fragment with that estimated from c.d. measurements by Wang et al. [25] (Table 2; Figure 5) indicates that its secondary structure is the same regardless of whether it exists in the intact protein or the isolated form.

We thank Dr A. V. Finkelstein for providing the ALB computer program. This work was supported by a statutable grant from the State Committee for Scientific Research to the Nencki Institute.

REFERENCES

- Marston, S. B. and Redwood, C. S. (1991) *Biochem J.* **279**, 1–16
- Walsh, M. P. (1991) *Biochem. Cell Biol.* **69**, 771–800
- Brenner, B., Yu, L. C. and Chalovich, J. M. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 5739–5743
- Marston, S. B. and Redwood, C. S. (1992) *J. Biol. Chem.* **267**, 16796–16800
- Graceffa, P., Wang, C.-L. A. and Stafford, W. F. (1988) *J. Biol. Chem.* **263**, 14195–14202
- Mabuchi, K. and Wang, C.-L. A. (1991) *J. Muscle Res. Cell. Motil.* **12**, 145–151
- Lynch, W. P., Riseman, V. M. and Bretscher, A. (1987) *J. Biol. Chem.* **262**, 7429–7437
- Graceffa, P. and Jancsó, A. (1990) *J. Muscle Res. Cell Motil.* **11**, 440 (abstr.)
- Bryan, J., Imai, M., Lee, R., Moore, F., Cook, R. G. and Lin, W. G. (1989) *J. Biol. Chem.* **264**, 13873–13879
- Yang, J. Y., Wu, C.-S. C. and Martinez, H. M. (1986) *Methods Enzymol.* **130**, 208–269
- Provencher, S. W. and Glöckner, J. (1981) *Biochemistry* **20**, 33–37
- Venyaminov, S. Yu., Baikalov, I. A., Shen, Z. M., Wu, C.-S. C. and Yang, J. T. (1993) *Anal. Biochem.*, in the press
- Szpacenko, A. and Dąbrowska, R. (1986) *FEBS Lett.* **202**, 182–186
- Nefsky, B. and Bretscher, A. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 3549–3553
- Ptitsyn, O. B. and Finkelstein, A. V. (1983) *Biopolymers* **22**, 15–25
- Bretscher, A. (1984) *J. Biol. Chem.* **259**, 12873–12880
- Czuryło, E. A., Emelyanenko, V. I., Permyakov, E. A., Dąbrowska, R. (1991) *Biophys. Chem.* **40**, 181–188
- Jaenicke, L. (1974) *Anal. Biochem.* **61**, 623–627
- Winder, A. F. and Gent, W. L. G. (1971) *Biopolymers* **10**, 1243–1251
- Mihaly, E. (1968) *J. Chem. Eng. Data* **13**, 179–182
- Hayashi, K., Kanda, K., Kimizuka, F., Kato, L. and Sobue, K. (1989) *Biochem. Biophys. Res. Commun.* **164**, 503–511
- Venyaminov, S. Yu. and Gogia, Z. V. (1982) *Eur. J. Biochem.* **126**, 299–309
- Privalov, P. I., Tiktopulo, E. I., Venyaminov, S. Yu., Griko, Yu. V., Makhatadze, G. I. and Khechinashvili, N. N. (1989) *J. Mol. Biol.* **205**, 737–750
- Ptitsyn, O. B. and Finkelstein, A. V. (1989) *Protein Eng.* **2**, 443–447
- Wang, C.-L. A., Chalovich, J. M., Graceffa, P., Lu, R. G., Mabuchi, K. and Stafford, W. F. (1991) *J. Biol. Chem.* **266**, 13958–13963
- Bryan, J. (1990) *Ann. N.Y. Acad. Sci.* **599**, 100–110
- Gibrat, J. F., Garnier, J. and Robson, B. (1987) *J. Mol. Biol.* **198**, 425–443
- Katayama, E. (1989) *J. Biochem. (Tokyo)* **106**, 993–998
- Leszyk, J., Mornet, D., Audemard, E. and Collins, J. H. (1989) *Biochem. Biophys. Res. Commun.* **160**, 210–216