Homodimerization of calpain 3 penta-EF-hand domain

Ravikiran RAVULAPALLI, Beatriz GARCIA DIAZ, Robert L. CAMPBELL and Peter L. DAVIES¹

Department of Biochemistry, Queen's University, Kingston, ON, Canada K7L 3N6

Calpains 1 and 2 are heterodimeric proteases in which large (relative molecular mass M_r 80 000) and small (M_r 28 000) subunits are linked through their respective PEF (penta-EF-hand) domains. The skeletal muscle-specific calpain 3 is believed not to form a heterodimer with the small subunit but might homodimerize through its PEF domain. Size-exclusion chromatography and analytical ultracentrifugation of the recombinant PEF domain of calpain 3 show that it forms a stable homodimer that does not

dissociate on dilution. Molecular modelling suggests that there would be no barriers to the dimerization of the whole enzyme through the PEF domains. This orientation would place the catalytic centres at opposite ends of the dimer.

Key words: calpain 3, homodimerization, molecular modelling, muscular dystrophy, p94, penta-EF-hand domain (PEF domain).

INTRODUCTION

Calpains form a widely distributed superfamily of Ca^{2+} -dependent cytosolic cysteine proteases. Numerous calpain isoforms have been identified in vertebrates [1–3], invertebrates [1,3–6], plants [7,8] and microorganisms [1,3,6]. These enzymes are supposed to play a role in many intracellular processes linked to calcium signalling, including cell motility, apoptosis, cell differentiation and cell-cycle regulation [9–11]. They function by modulating the biological activities of their substrates through limited proteolysis [12]. Abnormal changes in calpain activity, including those due to defects in calcium homeostasis or mutations in calpain genes, contribute to pathologies such as ischaemic injury, Alzheimer's disease and other neurological disorders, muscular dystrophy, cancer and Type II diabetes [3,13].

The two abundant ubiquitous mammalian calpains 1 and 2 (μ - and m- respectively) are the best-characterized members of the family. They are heterodimers of a large (relative molecular mass M_r 80 000) and a small (M_r 28 000) subunit. The M_r 80 000 catalytic subunit has four structural domains (I–IV) [14,15]. The first two domains (I and II) make up the papain-like catalytic core common to all calpains, whereas domains III and IV are the C2-like and PEF (penta-EF-hand) domains respectively [1]. The small subunit contains two domains (V and VI). Its PEF domain (VI) makes extensive contact with the homologous PEF DIV (domain IV) of the large subunit through pairing of the fifth EF hand [16]. Both isoforms are catalytically inactive in the absence of Ca^{2+} . They undergo a series of conformational changes on Ca^{2+} binding to several domains [17–19]. In addition to these widely distributed μ - and m-isoforms, a dozen more distinctive isoforms have been identified in humans. Although the proteolytic core domains (I and II) are conserved in all isoforms of the calpain family, the other domains can be quite varied. Some homologues seem to lack the small subunit (calpains 3, 5, 6, 7 and 10), suggesting that they may not form heterodimers.

Calpain 3 (p94), the first identified 'tissue-specific' isoform, is produced predominantly in skeletal muscles [2]. p94 is implicated in myofibrillogenesis and sarcomere remodelling. Mutations in p94 are genetically linked to limb girdle muscular dystrophy type 2A, signifying a crucial role for p94 in muscle physiology

[20,21]. It has a domain organization similar to that of the μ - and m-calpain large subunits, and its amino acid sequence is 54 and 51 % identical with their respective large subunits [2]. p94 is M_r 14 000 larger than the μ - and m-calpain large subunits. This is due to the presence of three additional unique sequences; an N-terminal extension sequence of 47 amino acids (referred to as NS) that replaces the anchor helix of m-calpain, and two insertion sequences of 48 and 77 amino acids (referred to as IS1 and IS2 respectively). IS1 is located in domain II as an insert that flanks the active site cleft and acts as an internal propeptide [22]; IS2 is located between domains III and IV and binds titin [23].

Attempts to isolate p94 from muscles or to produce it using recombinant systems have been hampered by its rapid autolysis [24–26]. Owing to this instability, the physiological role of p94 and the structure–function aspects of its different domains and additional unique sequences remain obscure. When inactive [C129S (Cys¹²⁹ → Ser)] rat p94 was expressed in COS cells, it was stable, and during size-exclusion chromatography, it eluted with an M_r ~180 000, as though it were a dimer [27]. However, there have been no definitive studies performed on the issue of the oligomeric state of p94 *in vitro* or *in vivo*. Also, we note that PEF proteins do not naturally occur as monomers [28], suggesting that the PEF domain DIV could be the logical site for dimerization. In the present study, we describe experiments that establish the intrinsic capability of the p94 PEF domain to homodimerize, and we show by molecular modelling that there are no obvious barriers to dimerization of the whole enzyme through this interaction.

MATERIALS AND METHODS

Cloning of p94DIV

PCR primers (5'-TATACATATGGGCAGCTCTGATCAGGAAAGTG-3' and 5'-CACCTCGAGGGCATACTGGTGAGCTGCAGCCACTC-3') were designed to generate the cDNA fragment encoding the DIV region of p94 by PCR amplification of reverse transcripts (reverse transcriptase–PCR) of total RNA from human skeletal muscles (Stratagene, La Jolla, CA, U.S.A.) using an RT-Thermoscript kit (Invitrogen, Carlsbad, CA, U.S.A.) and Expand high-fidelity DNA polymerase (Roche, Indianapolis, IN, U.S.A.).

Abbreviations used: DIV, domain IV; DTT, dithiothreitol; IS, insertion sequence; Ni-NTA, Ni²⁺-nitrilotriacetate–agarose; NS, N-terminal sequence; PEF, penta-EF-hand.

¹ To whom correspondence should be addressed (email daviesp@post.queensu.ca).

Co-ordinates of the reported protein structure have been deposited in the Protein Data Bank (PDB ID 1Y9V).

Figure 1 Human calpain p94 domain IV and its purification

(A) The cDNA sequence of domain IV of human calpain p94 is shown above the encoded amino acid sequence. (B) SDS/PAGE of fractions obtained during purification. Lane 1 shows the molecular-mass standards indicated at the side of the gel. Lanes 2 and 3 were loaded with an aliquot of *E. coli* supernatant and pellet respectively. Lanes 4–6 represent pooled fractions from purifications steps on DEAE-Sepharose, Ni-NTA and Sephadex G-75 columns respectively. The amount of protein in lane 5 is twice that loaded in lane 6.

The amplified product was cloned into the pCR 2.1-TOPO vector (Invitrogen). DNA sequences of positive clones were verified. The verified cDNA sequence was ligated into the *NdeI* and *XhoI* sites of the pET24a expression vector (Novagen, Madison, WI, U.S.A.) to incorporate an N-terminal Met codon and a C-terminal histidine tag into the p94DIV construct (Figure 1A).

Protein expression and purification

The plasmid encoding p94DIV was transfected into *Escherichia coli* BL21(DE3) (Novagen) grown in LB (Luria–Bertani) medium under kanamycin selection to an absorbance A_{600} of 0.8–1.0 at 37°C. Protein expression was induced with 0.4 mM isopropyl β -D-thiogalactoside for 3 h. Cells harvested from the culture by centrifugation were resuspended in lysis buffer (25 mM Tris/HCl, pH 7.6, 5 mM EDTA, 5%, v/v, glycerol, 10 mM 2-mercaptoethanol and 0.1 mM PMSF) and lysed by sonication. The lysate was centrifuged at 27 000 *g* for 45 min and the supernatant was applied to a DEAE-Sepharose column. Protein was eluted by a 0–0.75 M NaCl salt gradient in lysis buffer, and fractions containing p94DIV were identified by the presence of a prominent *M_r* 22 000 band on SDS/PAGE. These fractions were pooled and applied

to a 10 ml NTA (Ni^{2+} -nitrilotriacetate–agarose) column (Qiagen, Chatsworth, CA, U.S.A.). The column was washed with N-buffer (50 mM Tris/HCl, pH 7.6, 100 mM NaCl, 5 mM imidazole and 0.01% sodium azide) and eluted with N-buffer and 250 mM imidazole. The fractions containing p94DIV were further purified by gel-permeation chromatography.

Size-exclusion chromatography

A Sephadex G-75 column (2.6 cm $\times$ 100 cm; Amersham Biosciences) was equilibrated overnight with 25 mM Tris/HCl (pH 7.6), 150 mM NaCl, 5 mM EDTA, 2% glycerol and 0.05% sodium azide at a flow rate of 0.5 ml/min. The column was then calibrated with the following protein standards: Blue Dextran (*M_r* 2 000 000), BSA (*M_r* 66 000), ovalbumin (*M_r* 44 000), carbonic anhydrase (*M_r* 29 000), myoglobin (*M_r* 17 000) and cytochrome *c* (*M_r* 12 400), with inorganic phosphate to mark the total volume. The protein sample containing p94DIV was loaded on to the column and eluted at 0.5 ml/min. All chromatographic runs were performed at 4°C. Peak volumes are standardized to a *K_D* (diffusion constant) value [30] given by $K_D = (V_e - V_0)/(V_t - V_0)$, where V_e is the elution volume of the centre of a protein peak, V_0 the void volume (determined by elution of Blue Dextran) and V_t the total elution volume (determined by elution of inorganic phosphate). The absorbance of each fraction was measured at 280 nm, and those in which Blue Dextran and myoglobin eluted were measured at 256 and 415 nm respectively. The presence of phosphate in the eluate was detected by a colorimetric assay [31].

Purified p94DIV was concentrated using a Centricon *M_r* 10 000 concentrator (Millipore, Bedford, MA, U.S.A.) to a final protein concentration of 25 mg/ml in 10 mM Hepes (pH 7.5) and 10 mM DTT (dithiothreitol), then divided into 400 μ l aliquots and flash-frozen in liquid nitrogen for storage.

CD spectroscopy

CD spectra were recorded using an Olis RSM 1000 CD spectrophotometer (Bogart, GA, U.S.A.). The CD spectra spanning wavelengths 260–180 nm were collected using a 0.1 mm path length quartz cuvette at 4°C. p94DIV was dialysed into 5 mM Tris/HCl (pH 7.6), 0.1 mM DTT and 0.2 mM EDTA. The protein concentration was 2.2 mg/ml. Five scans were recorded, averaged and corrected for buffer background signal. Deconvolution of the data was performed using CDNN software [32]. Results are expressed in terms of mean residue ellipticity $[\theta]$ (deg $\cdot$ cm² $\cdot$ dmol^{−1}).

Intrinsic tryptophan fluorescence studies

Intrinsic tryptophan fluorescence measurements were performed using a PerkinElmer LS50B fluorescence spectrophotometer equipped with a stirrer-adapted 4 ml cuvette (Hellma, Concord, Ontario, Canada) at a constant temperature of 24°C. Excitation and emission wavelengths were set at 280 and 340 nm respectively. The reaction buffer was 50 mM Tris/HCl (pH 7.6) containing 100 mM NaCl, 1 mM EDTA and 1 mM DTT. p94DIV (2 μ M) was added to the assay buffer and intrinsic fluorescence was recorded for 10 min. At this point, 50 mM CaCl_2 dissolved in the assay buffer was pumped continuously into the cuvette at 4 μ l/min using an injector pump (Harvard Apparatus pump 22). Similar experiments were performed under the same conditions but using MgCl_2 instead of CaCl_2 . In a control experiment, intrinsic tryptophan fluorescence changes were monitored during CaCl_2 addition to purified rat μ I-II (1 μ M) [18] by the method described above for p94DIV.

Analytical ultracentrifugation

Ultracentrifugation experiments were performed on p94DIV at 20°C in a Beckman (Palo Alto, CA, U.S.A.) Optima XL-I analytical ultracentrifuge equipped with automated scanning absorption optics (at 280 nm) by the method of Laue and Stafford [29]. Protein solutions were exhaustively dialysed against 20 mM Pipes (pH 6.5), containing 150 mM NaCl, 0.1 mM DTT and either 1 mM EDTA or 2 mM CaCl_2 . The rotor used was a Beckman An60-Ti. Sedimentation velocity runs were performed at two different speeds, 30 000 and 42 000 rev./min, for 7 and 4 h duration respectively and a final protein concentration of 0.25 mg/ml both in the absence and presence of Ca^{2+} (2 mM). The results were analysed by SEDFIT software [33]. Sedimentation equilibrium experiments were performed at three different speeds, 12 000, 16 000 and 20 000 rev./min, and at six different protein concentrations (0.05, 0.1, 0.2, 0.3, 0.4 and 0.5 mg/ml) both in the presence and absence of Ca^{2+} (2 mM). Each speed was maintained until there was no significant difference in $r^2/2$ versus absorbance. After 24 h of equilibration at each speed, a scan was taken at each hour for the next 5 h (total five scans). The results were analysed using Origin4.1 software (Beckman).

Homology modelling of the p94DIV dimer

Modelling of the p94DIV dimer was performed using PyMOL, MODELLER and GROMACS [34–37]. A total of 100 models of p94DIV was created with the program MODELLER by mapping DIV (Glu⁶⁴⁹–Ala⁸²¹) on to the crystal structure of the rat calpain small subunit DVI homodimer (Glu¹–Ser¹⁷³) [38] to verify that there was sufficient sequence identity to form a dimer. The structures were examined with PyMOL to detect clashes at the interface, and the complementarity of the models was measured with the program sc from the CCP4 program suite [39]. The model with the highest complementarity was subjected to energy minimization for 10 000 steps or until convergence (tolerance, 1000 kJ · mol⁻¹ · nm⁻¹) and molecular dynamics (3 ns or 30 000 steps, at 300 K, with a constant temperature bath) using GROMACS. Coordinates for the p94DIV dimer have been deposited in the Protein Data Bank (PDB ID 1Y9V). A second model was generated to examine the possibility of heterodimerization of p94DIV with DVI. In this case, the sequence of p94DIV was mapped on to one subunit of the rat calpain DIV homodimer, whereas the sequence of the human DVI was mapped on to the other. As a control, the process of energy minimization and molecular dynamics was repeated for the crystal structure of the rat DVI homodimer (PDB ID 1DVI) and for the DIV–DVI heterodimer (PDB ID 1KFU) from the Ca^{2+} -free human calpain crystal structure. The trajectories from the molecular dynamics simulations were analysed for energetic stability and complementarity between the subunits.

Graphic model of the m-calpain large-subunit dimer

A homodimer of two m-calpain large subunits was generated by superimposing DIV of one m-calpain heterodimer on to the small subunit (DVI not shown) of a second m-calpain heterodimer (Figure 6). This computer-graphics model of a full-length m-calpain large-subunit dimer associated through their PEF domains was constructed using PyMOL. The structure was visually examined for structural clashes but no other manipulations were performed.

RESULTS

Expression and purification of recombinant human p94 domain IV

The DNA sequence encoding p94DIV is shown in Figure 1(A) above the amino acid sequence. The predicted protein has 189

Figure 2 CD spectra and intrinsic tryptophan studies on p94DIV

(A) Plot of the molar ellipticity θ as a function of wavelength. Spectra of p94DIV were collected over the far- and near-UV range (180–260 nm). (B) Ca^{2+} -induced conformational change in p94DIV (black) and μ I-II (light grey). Intrinsic fluorescence intensity was monitored at 340 nm (I_{340}) by exciting p94DIV and μ I-II at 280 nm while continuously adding CaCl_2 . Under similar conditions, no change in protein conformation was observed when MgCl_2 was added to p94DIV (dark grey).

residues, including the His₆ C-terminal tag, with a theoretical pI of 5.92 and a calculated M_r of 22 181. The amino acid sequence is approx. 52 % identical with DIV of μ -calpain and approx. 47 % identical with DIV of m-calpain. p94DIV was produced as the major soluble product in *E. coli* supernatant (Figure 1B, lane 2) and was purified using a combination of three column separations, DEAE-Sepharose, Ni-NTA and Sephadex G-75. The pooled fractions from the DEAE-Sepharose column contained several higher M_r contaminating *E. coli* proteins, which were largely removed in the subsequent Ni-NTA affinity chromatography step. The final purification product from Sephadex G-75 was estimated to be more than 95 % pure, as indicated by the SDS/PAGE (Figure 1B, lane 6). Soluble p94DIV was stable throughout the purification and concentration steps. A final yield of approx. 250 mg of protein was obtained from a 4 litre culture.

CD spectra and intrinsic tryptophan fluorescence studies

To assess that recombinant p94DIV is indeed properly folded, we determined its secondary structure by CD (Figure 2A). The CD spectra of p94DIV are those of a predominantly α -helical protein. Deconvolution of the CD spectra showed a high α -helical content (~ 59 %) and a very low percentage of random coil. A comparison of the secondary structure profile of p94DIV with that of the homologous rat small subunit determined from its crystal structure (PDB ID 1DVI) shows that the two proteins share high structural

Table 1 Comparison of secondary structure content of p94DIV, based on CD spectra or modelling, with that of the rat M_r 21 000 small-subunit dimer obtained from its crystallographic structure (PDB ID 1DVI)

	Rat small subunit (%)	p94DIV (%)	p94DIV model (%)
Helix	63	59	64
Random coil	17	16	17
β -Strand	5	5	5
β -Turn	14	13	14
Total	99	93	100

Figure 3 Size-exclusion (Sephadex G-75) chromatography standard curve

K_D values obtained from peak volumes of different protein standards are plotted against $\log M_r$. The points a–e correspond to the protein standards used: (a) BSA (M_r 66 000), (b) ovalbumin (M_r 44 000), (c) carbonic anhydrase (M_r 29 000), (d) myoglobin (M_r 17 000) and (e) cytochrome *c* (M_r 12 400). Point x represents the intersection of the K_D value of p94IV obtained from its elution volume with the standard curve giving a $\log M_r$ value. The M_r was then calculated to be approx. 40 000 with an error value of ± 5000 .

similarity (Table 1). There are only a few percentage points difference in the abundance of each secondary structure type. The same is true for the model of p94DIV. These results strongly suggest that p94DIV was produced as a well-folded protein in *E. coli*.

To study the effect of Ca^{2+} on the conformation of p94DIV, we monitored its intrinsic tryptophan fluorescence in the presence and absence of this ion. The p94DIV monomer has three tryptophan residues and five putative EF-hand Ca^{2+} -binding sites. p94DIV displayed a significant shift in fluorescence intensity on the addition of Ca^{2+} (Figure 2B), which is indicative of a conformational change in the protein caused by binding Ca^{2+} . Under the same conditions, Mg^{2+} did not produce a change in intrinsic tryptophan fluorescence. For a positive control, we used the proteolytic core of μ -calpain (μ I–II), which undergoes a large conformational change when it binds Ca^{2+} (Figure 2B, light grey trace) [18].

P94DIV is a stable dimer in solution

The fraction obtained after Ni-NTA purification was chromatographed on Sephadex G-75, both as a purification step and to determine its M_r . A standard curve was generated by plotting the K_D values for proteins of known molecular masses against the logarithm of their molecular masses (Figure 3). Recombinant p94DIV eluted immediately after ovalbumin as a single macromolecular species with an $M_r \sim 40\,000 \pm 5000$, estimated by interpolation of its K_D value on to the standard curve. This is twice the expected molecular mass (M_r 22 181) of p94DIV. These results suggest that p94DIV forms a homodimer in solution. To confirm this and obtain quantitative information on the strength of dimerization, ultracentrifugation analysis was performed.

Figure 4 Sedimentation equilibrium distribution for p94IV

The M_r is plotted against the cell loading concentration c (mg/ml). The graph shows that the M_r of p94DIV is constant as a function of concentration (c), indicating an ideal single component in sample. Graphical data are consistent with the expected molecular mass range for the p94DIV domain dimer ($\sim 43\,500$) throughout the concentration range in the absence of Ca^{2+} . Similar values were obtained in the presence of 2 mM Ca^{2+} .

Sedimentation velocity analysis of recombinant p94DIV showed that the sample solution was homogeneous both in the presence and absence of 2 mM CaCl_2 . A value of $s_{20,w}^0 \sim 3.41$ was obtained, which in this case corresponds to an M_r of approx. 44 000, assuming the protein to be globular in shape. From sedimentation equilibrium studies, an M_r of 43 580 was obtained, which is consistent with the expected theoretical value for the p94DIV dimer (44 360). The dependence of M_r on the ultracentrifuge cell-loading concentration c is shown in Figure 4 for the entire range of protein concentrations used (approx. 0.05–0.6 mg/ml). M_r is constant as a function of c , indicating that the sample behaves as an ideal single component. No evidence for dissociation of the homodimer (at concentrations of 0.05, 0.1 and 0.2 mg/ml) was observed either in the presence or absence of Ca^{2+} , even at the lowest concentrations tested, suggesting strong dimer interaction.

Homology modelling of p94DIV

Mapping of the sequence of p94DIV on to the rat calpain DVI homodimer showed that the dimer interface is well conserved (Figure 5). Energy minimization and molecular dynamics did not show any evidence that the p94DVI homodimer is unstable. Calculation of the average complementarity from every 100th structure in the dynamics trajectories for the human m-calpain DIV–DVI heterodimer, the rat DVI homodimer and the p94DIV homodimer models gave values of 0.76, 0.74 and 0.73 respectively. Average complementarity of the p94DIV homodimer lies within the minimum and maximum values calculated for the trajectory for the human m-calpain DIV–DVI heterodimer (0.70–0.80 respectively). This suggests that the p94DIV is capable of dimerization. When a model of the p94DIV–DVI heterodimer was generated, there were no obvious steric clashes or charge repulsions that would prevent heterodimerization and the average complementarity value 0.72 was comparable with the others.

DISCUSSION

The premise for our study of p94DIV dimerization stems from the circumstantial evidence that p94 might associate with itself rather than with the small subunit (M_r 28 000) found in calpains 1 and 2. This small subunit has not been seen in preparations of p94 from skeletal muscles [27], even though muscle

Figure 5 Homology model of p94DIV

(A) Molecular surface of the 'A' chain at the dimer interface between the two monomers of the rat DVI dimer. Residues that are identical between rat DVI and p94DIV are coloured white, while those that differ are coloured black. (B) Alpha-carbon trace of the superimposed structures of rat DVI dimer (black) and the p94DIV dimer (light grey) created with the modeller program. The view in this Figure is rotated 90° about the vertical axis from that in (A). The 'A' chain of each dimer is drawn in thick tubes, while the 'B' chain is drawn in thin tubes. The spheres represent Ca^{2+} ions.

cells do produce the small-subunit-containing calpains [40]. Also, yeast two-hybrid assays on p94 (both inactive and active) failed to show binding of the small subunit to p94 [23,41]. Evidence for homodimerization comes mainly from reports that an active-site mutant of p94(C129S) elutes as an M_r 180 000 protein during size-exclusion chromatography [27]. This is consistent with homodimer formation, although an asymmetrical p94 monomer could also produce the same result. In a comparable system, the homologous large subunits of chicken calpain (M_r 80 000) have been reported to form homodimers when expressed alone in SF9 cells, but form heterodimers when co-expressed with the M_r 28 000 small subunit [27,42]. Most PEF proteins are known to form either heterodimers or homodimers. This has been established by biochemical analyses, co-immunoprecipitation and yeast two-hybrid assays on PEF proteins such as calpain, ALG-2, grancalcin, sorcin and peflin [28]. This observation that PEF domains are typically paired as heterodimers [3,14,15,43] or homodimers [44–46] argues for p94 homodimerization and points to the PEF domain as being the probable homodimerization interface.

p94DIV was massively produced in *E. coli* and gave yields that were comparable with those obtained with the homodimer of the rat calpain small subunit [46]. It accumulated in the soluble fraction and behaved during column chromatography as a single entity with a sharp elution profile, consistent with it being a stable, well-folded protein. This was confirmed by analytical ultracentrifugation and CD spectroscopy studies. On Ca^{2+} binding, it underwent a conformational change that was detected by an increase in intrinsic tryptophan fluorescence. Evidence for p94DIV

being a homodimer comes from both size-exclusion chromatography and analytical ultracentrifugation studies. The M_r values of the p94DIV domain deduced from size-exclusion chromatography ($40\,000 \pm 5000$) and analytical ultracentrifugation studies (43 500) are consistent with a calculated theoretical dimer molecular mass of 44 362. The value determined by size-exclusion chromatography is also deemed to be reliable because crystallography of calpain PEF domain homo- and heterodimers shows them to be roughly globular in shape ($47 \text{ \AA} \times 48 \text{ \AA} \times 54 \text{ \AA}$; $1 \text{ \AA} = 10^{-10} \text{ m}$) both with and without Ca^{2+} present [38,47]. The dimer molecular mass calculated from analytical ultracentrifugation (M_r 43 500) is constant over a 10-fold concentration range, suggesting that the dimer is very stable.

To investigate further the implications of our biophysical studies described above, we performed some modelling studies. Initially, we mapped the p94DIV sequence on to the rat small subunit (DIV) and observed that the dimer interface is well conserved. We generated a homology model of p94DIV and performed energy minimization and molecular dynamics studies. These showed that p94DIV is capable of forming a homodimer through the same interface as the rat DVI homodimer and the human m-calpain heterodimer (Figure 5).

Given the known propensity of PEF domain proteins to form dimers, we explored the possibility that full-length p94 could dimerize through the pairing of its PEF domain. In the absence of a p94 structure, we made a computer-graphics model of an m-calpain large-subunit homodimer (Figure 6). The m-calpain large subunit has a domain composition similar to that of p94 but lacks the three unique additional sequences (NS1, IS1 and IS2).

Figure 6 Graphic model of the large subunit dimer

(**A**) Schematic representation of the large subunit domain structures of m-calpain and p94. The domain organization of the two proteins is similar, but p94 has three additional unique sequences; an N-terminal (NS) extension sequence and two insertion sequences (IS1 and IS2). (**B**) Homodimer of two m-calpain large subunits generated by superimposing domain VI of one m-calpain heterodimer on to domain IV of the other (domain VI not shown). They associate with each other through the PEF domain IV. The locations of IS1 and IS2 and the NS sequence in p94 (coloured purple and cyan respectively) are indicated by thick arrows. The inset to (**B**) shows a side view (rotated 90° about the vertical axis).

The regions of the human m-calpain that are replaced by these additional sequences (NS, IS1 and IS2) in p94 are highlighted in Figure 6. Since these additional sequences have no known structural homologues, we were not able to model their structures with confidence. Nevertheless, we can draw some conclusions from this mock-up. The IS1 sequence is remote from the dimerization contact and is unlikely to be affected by dimer formation. The NS sequence of p94 replaces the anchor helix of m-calpain. Although it appears to be in close proximity to the DIV homodimer, its sequence is rich in proline and glycine residues and shows no homology to the anchor helix of m-calpain. Therefore we predict that it will be intrinsically disordered and will not occupy the same space as the m-calpain anchor helix. Thus it is unlikely to interfere with dimer association. The IS2 sequence is present in the m-calpain 'transducer' region, which lies away from the dimer interface. Although we have no structural information about IS2, there is nothing in our model that suggests that it

should interfere with homodimerization of p94. Also, since IS2 is the binding site for titin, we assume that it would project outwards to favour titin association. In our model, the catalytic cores are orientated in opposite directions (at the two extreme ends), and they are unlikely to autoproteolyse each other. Dimerization removes the fifth EF-hand of DIV as a potential binding site for the interactions of p94 with other proteins. However, dimerization means that all other possible contact sites are duplicated, which could result in stronger binding or simultaneous binding to two ligands (e.g. binding sites on titin) [23].

Homodimerization in this manner might apply to all of the single-chain calpains that have a PEF domain, such as the *Drosophila* calpains A–C. Furthermore, it should not be assumed that other PEF-containing mammalian calpains (e.g. 8, 11, 12 and 13) will necessarily form a heterodimer with the small subunit in preference to a homodimer, until this is proven by experimentation as with calpains 1, 2 and 9 [48].

This work was supported by the Heart and Stroke Foundation of Ontario. We are grateful to the Protein Function Discovery (PFD) Facility of Queen's University and K. Munro for assistance with analytical ultracentrifugation and CD analysis and to S. Gauthier for technical assistance. R. R. is supported by a tuition bursary and a QGA award from Queen's University. P. L. D. holds a Canada Research Chair in Protein Engineering.

REFERENCES

- Sorimachi, H. and Suzuki, K. (2001) The structure of calpain. *J. Biochem. (Tokyo)* **129**, 653–664
- Sorimachi, H., Imajoh-Ohmi, S., Emori, Y., Kawasaki, H., Ohno, S., Minami, Y. and Suzuki, K. (1989) Molecular cloning of a novel mammalian calcium-dependent protease distinct from both m- μ -types. Specific expression of the mRNA in skeletal muscle. *J. Biol. Chem.* **264**, 20106–20111
- Goll, D., Thompson, V., Li, H., Wei, W. and Cong, J. (2003) The calpain system. *Physiol. Rev.* **83**, 731–801
- Pinter, M., Stierandova, A. and Friedrich, P. (1992) Purification and characterization of a Ca^{2+} -activated thiol protease from *Drosophila melanogaster*. *Biochemistry* **31**, 8201–8206
- Barnes, T. M. and Hodgkin, J. (1996) The tra-3 sex determination gene of *Caenorhabditis elegans* encodes a member of the calpain regulatory protease family. *EMBO J.* **15**, 4477–4484
- Huang, X., Czerwinski, E. and Mellgren, R. (2003) Purification and properties of the *Dictyostelium* calpain-like protein, Cpl. *Biochemistry* **42**, 1789–1795
- Lid, S. E., Gruis, D., Jung, R., Lorentzen, J. A., Ananiev, E., Chamberlin, M., Niu, X., Meeley, R., Nichols, S. and Olsen, O. A. (2002) The defective kernel 1 (*dek1*) gene required for aleurone cell development in the endosperm of maize grains encodes a membrane protein of the calpain gene superfamily. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5460–5465
- Margis, R. and Margis-Pinheiro, M. (2003) Phytocalpains: orthologous calcium-dependent cysteine proteinases. *Trends Plant Sci.* **8**, 58–62
- Carafoli, E. and Molinari, M. (1998) Calpain: a protease in search of a function? *Biochem. Biophys. Res. Commun.* **247**, 193–203
- Santella, L. (1998) The role of calcium in the cell cycle: facts and hypotheses. *Biochem. Biophys. Res. Commun.* **244**, 317–324
- Wang, K. K. (2000) Calpain and caspase: can you tell the difference? *Trends Neurosci.* **23**, 20–26
- Glading, A., Lauffenburger, D. A. and Wells, A. (2002) Cutting to the chase: calpain proteases in cell motility. *Trends Cell Biol.* **12**, 46–54
- Huang, Y. and Wang, K. K. (2001) The calpain family and human disease. *Trends Mol. Med.* **7**, 355–362
- Hosfield, C. M., Elce, J. S., Davies, P. L. and Jia, Z. (1999) Crystal structure of calpain reveals the structural basis for Ca^{2+} -dependent protease activity and a novel mode of enzyme activation. *EMBO J.* **18**, 6880–6889
- Strobl, S., Fernandez-Catalan, C., Braun, M., Huber, R., Masumoto, H., Nakagawa, K., Irie, A., Sorimachi, H., Bourenkow, G., Bartunik, H. et al. (2000) The crystal structure of calcium-free human m-calpain suggests an electrostatic switch mechanism for activation by calcium. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 588–592
- Kretsinger, R. H. (1997) EF-hands embrace. *Nat. Struct. Biol.* **4**, 514–516
- Tomba, P., Emori, Y., Sorimachi, H., Suzuki, K. and Friedrich, P. (2001) Domain III of calpain is a Ca^{2+} -regulated phospholipid-binding domain. *Biochem. Biophys. Res. Commun.* **280**, 1333–1339
- Moldoveanu, T., Hosfield, C. M., Lim, D., Elce, J. S., Jia, Z. and Davies, P. L. (2002) A Ca^{2+} switch aligns the active site of calpain. *Cell (Cambridge, Mass.)* **108**, 649–660
- Reverter, D., Strobl, S., Fernandez-Catalan, C., Sorimachi, H., Suzuki, K. and Bode, W. (2001) Structural basis for possible calcium-induced activation mechanisms of calpains. *Biol. Chem.* **382**, 753–766
- Kramerova, I., Kudryashova, E., Tidball, J. G. and Spencer, M. J. (2004) Null mutation of calpain 3 (p94) in mice causes abnormal sarcomere formation *in vivo* and *in vitro*. *Hum. Mol. Genet.* **13**, 1373–1388
- Richard, I., Broux, O., Allamand, V., Fougerousse, F., Chiannikulchai, N., Bourg, N., Brenguier, L., Devaud, C., Pasturaud, P., Roudaut, C. et al. (1995) Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. *Cell (Cambridge, Mass.)* **81**, 27–40
- Diaz, B. G., Moldoveanu, T., Kuiper, M. J., Campbell, R. L. and Davies, P. L. (2004) Insertion sequence 1 of muscle-specific calpain, p94, acts as an internal propeptide. *J. Biol. Chem.* **279**, 27656–27666
- Sorimachi, H., Kinbara, K., Kimura, S., Takahashi, M., Ishiura, S., Sasagawa, N., Sorimachi, N., Shimada, H., Tagawa, K., Maruyama, K. et al. (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
- Sorimachi, H., Ishiura, S. and Suzuki, K. (1993) Muscle-specific calpain, p94, is degraded by autolysis immediately after translation, resulting in disappearance from muscle. *J. Biol. Chem.* **268**, 10593–10605
- Federici, C., Eshdat, Y., Richard, I., Bertin, B., Guillaume, J. L., Hattab, M., Beckmann, J. S., Strosberg, A. D. and Camoin, L. (1999) Purification and identification of two putative autolytic sites in human calpain 3 (p94) expressed in heterologous systems. *Arch. Biochem. Biophys.* **363**, 237–245
- Branca, D., Gugliucci, A., Bano, D., Brini, M. and Carafoli, E. (1999) Expression, partial purification and functional properties of the muscle-specific calpain isoform p94. *Eur. J. Biochem.* **265**, 839–846
- Kinbara, K., Ishiura, S., Tomioka, S., Sorimachi, H., Jeong, S. Y., Amano, S., Kawasaki, H., Kolmerer, B., Kimura, S., Labeit, S. et al. (1998) Purification of native p94, a muscle-specific calpain, and characterization of its autolysis. *Biochem. J.* **335**, 589–596
- Maki, M., Kitaura, Y., Satoh, H., Ohkouchi, S. and Shibata, H. (2002) Structures, functions and molecular evolution of the penta-EF-hand Ca^{2+} -binding proteins. *Biochim. Biophys. Acta* **1600**, 51–60
- Laue, T. M. and Stafford, III, W. F. (1999) Modern applications of analytical ultracentrifugation. *Annu. Rev. Biophys. Biomol. Struct.* **28**, 75–100
- Andrews, P. (1964) Estimation of the molecular weights of proteins by Sephadex gel-filtration. *Biochem. J.* **91**, 222–233
- Mackereth, F. J. H. (1978) Water Analysis: Some Revised Methods for Limnologists, Freshwater Biological Association, Ambleside
- Bohm, G., Muhr, R. and Jaenicke, R. (1992) Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng.* **5**, 191–195
- Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606–1619
- DeLano, W. L. (2003) The PyMOL molecular graphics system. DeLano Scientific, San Carlos, CA, <http://www.pymol.org>
- Sali, A. and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815
- Lindahl, E., Hess, B. and van der Spoel, D. (2001) A package for molecular simulation and trajectory analysis. *J. Mol. Model.* **7**, 306–317
- Berendsen, H. J. C., van der Spoel, D. and van Drunen, R. (1995) GROMACS: a message-passing parallel molecular dynamics implementation. *Comp. Phys. Commun.* **91**, 43–56
- Blanchard, H., Grochulski, P., Li, Y., Arthur, J. S., Davies, P. L., Elce, J. S. and Cygler, M. (1997) Structure of a calpain Ca^{2+} -binding domain reveals a novel EF-hand and Ca^{2+} -induced conformational changes. *Nat. Struct. Biol.* **4**, 532–538
- Collaborative Computational Project Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763
- Farkas, A., Tompa, P. and Friedrich, P. (2003) Revisiting ubiquity and tissue specificity of human calpains. *Biol. Chem.* **384**, 945–949
- Herasse, M., Ono, Y., Fougerousse, F., Kimura, E., Stockholm, D., Beley, C., Montarras, D., Pinset, C., Sorimachi, H., Suzuki, K. et al. (1999) Expression and functional characteristics of calpain 3 isoforms generated through tissue-specific transcriptional and posttranscriptional events. *Mol. Cell. Biol.* **19**, 4047–4055
- Kinbara, K., Sorimachi, H., Ishiura, S. and Suzuki, K. (1998) Skeletal muscle-specific calpain, p49: structure and physiological function. *Biochem. Pharmacol.* **56**, 415–420
- Kitaura, Y., Matsumoto, S., Satoh, H., Hitomi, K. and Maki, M. (2001) Peflin and ALG-2, members of the penta-EF-hand protein family, form a heterodimer that dissociates in a Ca^{2+} -dependent manner. *J. Biol. Chem.* **276**, 14053–14058
- Blanchard, H., Li, Y., Cygler, M., Kay, C. M., Simon, J., Arthur, C., Davies, P. L. and Elce, J. S. (1996) Ca^{2+} -binding domain VI of rat calpain is a homodimer in solution: hydrodynamic, crystallization and preliminary X-ray diffraction studies. *Protein Sci.* **5**, 535–537
- Hamada, H., Okochi, E., Oh-hara, T. and Tsuruo, T. (1988) Purification of the Mr 22000 calcium-binding protein (sorcini) associated with multidrug resistance and its detection with monoclonal antibodies. *Cancer Res.* **48**, 3173–3178
- Graham-Siegenthaler, K., Gauthier, S., Davies, P. L. and Elce, J. S. (1994) Active recombinant rat calpain II. Bacterially produced large and small subunits associate both *in vivo* and *in vitro*. *J. Biol. Chem.* **269**, 30457–30460
- Lin, G. D., Chattopadhyay, D., Maki, M., Wang, K. K., Carson, M., Jin, L., Yuen, P. W., Takano, E., Hatanaka, M., DeLucas, L. J. et al. (1997) Crystal structure of calcium bound domain VI of calpain at 1.9 Å resolution and its role in enzyme assembly, regulation, and inhibitor binding. *Nat. Struct. Biol.* **4**, 539–547
- Lee, H. J., Tomioka, S., Kinbara, K., Masumoto, H., Jeong, S. Y., Sorimachi, H., Ishiura, S. and Suzuki, K. (1999) Characterization of a human digestive tract-specific calpain, nCL-4, expressed in the baculovirus system. *Arch. Biochem. Biophys.* **362**, 22–31