Structures of human calpain-3 protease core with and without bound inhibitor reveal mechanisms of calpain activation

Qili Ye, Robert L. Campbell, and Peter L. Davies

From the Department of Biomedical and Molecular Sciences, Queen’s University

Kingston, Ontario K7L 3N6, Canada

*Running title: Internal propeptide of calpain-3 is unstructured

1To whom correspondence should be addressed: Peter L. Davies, Department of Biomedical and Molecular Sciences, Queen’s University, Kingston, ON, Canada, K7L 3N6; Phone: (613) 533-2983; Fax: (613) 533-2497; E-mail: daviesp@queensu.ca

Key words: calpain-3, calcium-binding protein, autoproteolysis, complex, protease inhibitor, propeptide, crystal structure, drug design, small-angle X-ray scattering

ABSTRACT

Limb-girdle muscular dystrophy type 2a arises from mutations in the Ca\(^{2+}\)-activated intracellular cysteine protease calpain-3. This calpain isoform is abundant in skeletal muscle and differs from the main isoforms, calpain-1 and -2, in being a homodimer and having two short insertion sequences. The first of these, IS1, interrupts the protease core and must be cleaved for activation and substrate binding. Here, to learn how calpain-3 can be regulated and inhibited, we determined the structures of the calpain-3 protease core with IS1 present or proteolytically excised. To prevent intramolecular IS1 autoproteolysis, we converted the active-site Cys to Ala. Small-angle X-ray scattering (SAXS) analysis of the C129A mutant suggested that IS1 is disordered and mobile enough to occupy several locations. Surprisingly, this was also true for the apo version of this mutant. We therefore concluded that IS1 might have a binding partner in the sarcomere and is unstructured in its absence. After autoproteolytic IS1 removal from the active Cys-129 calpain-3 protease core, we could solve its crystal structures with and without the cysteine protease inhibitors E-64 and leupeptin covalently bound to the active-site cysteine. In each structure, the active state of the protease core was assembled by the cooperative binding of two Ca\(^{2+}\) ions to the equivalent sites used in calpain-1 and -2. These structures of the calpain-3 active site with residual IS1 and with bound E-64 and leupeptin may help guide the design of calpain-3-specific inhibitors.

Calpains are intracellular cysteine proteases transiently activated by calcium signaling (1–3). Triggering of these complex multi-domain proteases causes them to make transformative cuts in their protein targets to accomplish cellular events like reorganization of the cytoskeleton needed for cell movement (4), membrane repair (5), and vesicle transport (6). In mammals, calpains exist as a family of over a dozen different isoforms, all but one of which share a two-domain protease core (7). Many of the isoforms have a C-terminal penta-EF hand (PEF) domain, which serves as a dimerization module (8, 9). The structural basis for this dimerization is that calcium-binding EF-hands are typically paired. Thus, the fifth EF-hand that is not paired within the PEF domain is available to hetero- or homodimerize with another unpaired fifth EF-hand from a second PEF domain.

In the ubiquitously expressed main calpains -1 and -2, the PEF domain is heterodimerized to a two-domain small subunit through its C-terminal PEF domain (10, 11). In contrast, calpain-3 uses the fifth EF-hand for
Internal propeptide of calpain-3 is unstructured

homodimerization (Fig. 1). Evidence for this association came initially from the absence of small subunit in calpain-3 preparations (12) and from the apparent molecular weight of the enzyme (180,000 Da) being twice that expected for the monomer (94,000 Da) when autolysis was suppressed by mutation of the catalytic Cys to Ser (13). Subsequently, we showed direct evidence of homodimerization of the calpain-3 PEF domain (14, 15) and have subsequently solved the crystal structure of the Ca$^{2+}$-bound PEF homodimer (16). Other structural differences from the large subunit of calpain-1 and -2 are the presence of two insertion sequences (IS) that are not seen in any other calpain isoforms. IS1 is 48 residues long and interrupts the second domain (PC2) of the protease core after Asp 267. The other, IS2, is a similar-sized insertion present in the linker region between the calpain-type β-sandwich domain (CBSW) and the C-terminal PEF(L) domain (17), which is reported to bind titin in muscle (18). The N-terminal leader sequence (NS) is also different from all others in the calpains and has no homologs in the database.

Aside from these structural differences, calpain-3 is unique in being especially abundant in skeletal muscle (19, 20). It is encoded by a single gene, and loss-of-function mutations cause limb girdle muscular dystrophy (LGMD) type 2a (21, 22). This progressive wasting disease affects a characteristic set of muscles with onset in early adulthood. Almost 500 distinct LGMD type 2a mutations have been mapped in different families across the world according to the Leiden muscular dystrophy database (http://www.dmd.nl/). Some of these introduce stop codons that lead to truncated, non-functional calpain-3; others change an obviously essential amino acid, like one of the catalytic triad residues. However, many mutations affect residues somewhat removed from key sites, and it is not clear from models of apocalpain-3 (23) or the Ca$^{2+}$-bound active form (24) how these changes affect the enzyme. In an attempt to answer this question, we previously transferred several unexplained mutations into rat calpain-2 and found that they typically predisposed the enzyme to rapid autoproteolytic turnover, which could translate into accelerated loss of function of calpain-3 in the sarcomere (24).

Autoproteolysis of isolated or recombinant calpain-3 is a characteristic of this isoform, and has proven difficult to prevent in vitro (13). Cleavage initially occurs in the IS1 region. Modeling studies suggest the N-terminal part of IS1 lies in the catalytic cleft within range of the catalytic Cys (C129) (25). Thus, the instant the enzyme is activated, it can cut the N-terminal region of IS1 in an intramolecular reaction (26). When this autoproteolysis was studied on a recombinant version of the calpain-3 protease core the rate of the first IS1 cleavage was independent of enzyme concentration (26). This zero order reaction is consistent with intramolecular cleavage. Furthermore, when an attenuated version (C129S) of the calpain-3 core was incubated with active wild-type core (C129), the latter was unable to cut the C129S version, although it was able to cut itself. This confirmed that the activating first cut in IS1 is the result of an intramolecular reaction, whereas the subsequent cut at the other end of the insertion sequence could occur by intermolecular proteolysis. Although these cleavages have been regarded as autoproteolytic inactivation of the enzyme (13), we deem at least the first cleavage as an essential step to activation through removal of an internal propeptide that otherwise blocks substrate access to the catalytic cleft (25).

Here, as part of our ongoing structural analysis of the whole enzyme and its binding partners, we have determined structures of the calpain-3 protease core with IS1 present and with it proteolytically excised. The insertion sequence appears to be largely unstructured, both in the absence and presence of calcium. The predicted explanation for the intramolecular cleavage of IS1, that the N-terminal region lies in the catalytic cleft in range of the active site Cys, is here confirmed by X-ray crystallography. In addition, we have solved the first inhibitor-bound structures of the calpain-3 core, which is a step towards making inhibitors more specific for this calpain isoform. Selective inhibition of calpain-3 activity may have a sparing effect on those LGMD2A mutants that show accelerated autoproteolytic loss (24).

RESULTS

Most of IS1 was not seen in the crystal structure of human calpain-3 C129S protease core—The calpain-3 protease core immediately follows a unique N-terminal sequence of 47 amino
Internal propeptide of calpain-3 is unstructured

acids (NS), which is predicted by PSIPRED (27, 28) to be highly disordered (Fig. 2). Since this region is rapidly autolyzed (26), the first 45 residues of the NS sequence were deleted from the expression construct used here. The 48-residue insertion sequence 1 (IS1), which interrupts the protease core domain PC2 and acts as an internal autoinhibitory peptide (25), is also rapidly autolyzed. To slow down this internal autolysis, and thereby provide an opportunity to determine its structure and function within the core, we replaced the catalytically active residue Cys129 with Ser (26). This mutation produced a high yield of soluble protein that was amenable to crystallization.

The crystal of calpain-3 protease core ΔNS C129S formed in the presence of Ca$^{2+}$ diffracted to 2.3-Å resolution and contained four molecules in the asymmetric unit (A-D). No electron density was seen for the N-terminal residues Ile46 to Ile54 of any molecule, or even up to Glu58 in some molecules (Fig. 3). Whereas this N-terminal end of the core is flexible, the C terminus is more rigid and is only missing the last residue, Asp419, in one of the four molecules. Molecule B is described in the text below unless otherwise noted.

The overall structure of the Ca$^{2+}$-bound calpain-3 protease core is similar to the structures of other active protease cores of the calpain family. Superimposing the protease core structure of calpain-3 on the human calpain-1 protease core (PDB ID 2ARY) and rat calpain-2 protease core (PDB ID 1MDW) gave root mean square deviation (RMSD) values of 1.37 Å and 1.86 Å for alignment of 318 Ca atoms, respectively. The catalytic triad residues of PC1 and 2 are suitably spaced for proteolysis. Specifically, the distance of Ser129 Oγ to His334 N8 of imidazole ring is 3.6 Å, which is a functional distance for nucleophilic attack. Two Ca$^{2+}$ ions reside in the core that at the conserved Ca$^{2+}$ binding sites of PC1 and 2. Ca$^{2+}$ coordination at both sites matches that seen in other calpain protease cores (29–31). In PC1 the Ca$^{2+}$ is coordinated by the side chains of Glu199 and Asp120, the main-chain carbonyl groups of Ile113 and Gly115, and forms electron pairs with two water molecules. In PC2, the Ca$^{2+}$ is coordinated by the side chains of Glu364, Asp371 and Asp394, the main-chain carbonyl group of Thr392 and Glu396, and forms an electron pair with one water molecule. The PC1 and 2 domains are connected by a well-defined double salt-bridge (Arg118 to Glu396) adjacent to the Ca$^{2+}$ binding sites.

Unexpectedly, most of the unique IS1 region of the calpain-3 protease core was not seen in the structure, either because it was disordered or missing. Following the clear continuing electron density from the N terminus of IS1 (Asp268), eight residues were manually built in molecule B, and seven residues in molecules C and D, but only three residues could be traced in molecule A due to insignificant electron density thereafter. The C-terminal end of IS1 is missing, and electron density is not traceable until part way into PC2 at Pro319 in molecule B, Val320 in molecule C, Gln321 in molecule D and Glu323 in molecule A. The missing portion of IS1 lies between the 1b and 2b autolysis sites (25, 26).

Replacement of the catalytic Cys by Ala prevents autolysis—Because the region of IS1 lying between the two autolysis sites in calpain-3 protease core was not seen in the C129S crystal structure, we hypothesized that the C129S mutant might have had enough protease activity to excise IS1 during crystallization over many weeks in high CaCl$_2$ concentration (0.1M). To examine this possibility, we washed some C129S crystals with buffer, dissolved those in water, and analyzed the protein by SDS-PAGE (Fig. 4A) and MALDI-TOF mass spectrometry (data not shown). Both methods showed only 26- and 13-kDa species present. The former matched in size the ΔNS core up to the 1b IS1 cleavage site in PC2, and the latter to the remainder of PC2 from the 2b IS1 cleavage site onwards. These results from the C129S crystals are in agreement with previous experiment using the C129 protease core (25) over a much shorter time frame. It suggests that IS1 of the C129S crystals was indeed auto-proteolyzed during crystallization.

To investigate the source of proteolysis we did a 4-week incubation of the calpain-3 core with a variety of reagents that would distinguish between autoproteolysis and cleavage by exogenous proteases. Calpain-3 core C129S at 48μM was mixed with crystallization solution except for PEG8000, which affected visualization of protein bands on SDS-PAGE. Individual assays were done with the following additives: 0.2% NaN$_3$, 10 mM EDTA, 20 μM E-64, 0.5 mM PMSF, 100 mM NaCl and 10% Glycerol. To stop protease digestion on aliquots removed during the time course, the
samples were mixed with 3x SDS-PAGE loading buffer and boiled for 5 min. A trace amount of autoproteolysis of the protease core of calpain-3 C129S was detectable after 24 h (Fig. 4B). SDS-PAGE analysis exhibited three bands; a prominent one at 43-kDa, and faint ones at 26.6-kDa and 17-kDa. After 2 weeks, a 13-kDa band was noticeable and gradually increased in intensity in parallel with the 26.6-kDa band at the expense of the 43-kDa starting material. None of the protease inhibitors (EDTA, E-64 and PMSF), the bactericidal agent (NaNO3), or the protein stabilizers (glycerol and NaCl) were able to prevent enzyme autolysis (Fig. 4D-I). Therefore, it is confirmed that the protease core of calpain-3 C129S retains weak protease activity; presumably that of a serine protease because Ser129 is equally well aligned with His334 for catalysis, as is Cys 129 in the native protein.

In order to completely abolish protease activity, we mutated Cys129 to Ala129 and assayed its activity as we did for C129S. No autolysis bands were generated during the 4-week incubation with this mutant (Fig. 4C). Crystallization of this enzymatically inactive calpain-3 construct was undertaken, and C129A crystals were obtained in half the time (two weeks) it took to form C129S crystals.

The C129A crystal diffracted to 2.75 Å and, once again, four protease cores were present in the asymmetric unit. Even though there was no autolysis of C129A, most of IS1 still could not be built due to poor quality or missing electron density. Only the N-terminal seven residues of IS1 could be traced based on clear electron density. Therefore, we contend that IS1 is highly flexible in the Ca2+-bound core. Despite the difference in autolysis between C129S and C129A, the structures are superimposable, with two Ca2+ bound and the PC1 and 2 domains rotated to form the catalytic cleft. IS1 is exposed at the surface of the protein and, therefore, its absence from the crystal structure due to autolysis or flexibility has no effect on the core structure.

**IS1 of the calpain-3 core is highly mobile in solution as revealed by SAXS**—To better understand how IS1 relates to the enzyme, we looked at the conformation of the insertion sequence in solution using SAXS. To avoid any autolysis, even during the shorter time frame of the SAXS experiments, we used only the C129A version of the calpain-3 core. SAXS data were collected on calpain-3 core C129A in the Ca2+-free state, with 5mM EDTA present, and in the Ca2+-bound state in 2mM CaCl2 (Table 1). The scattering patterns of both samples were similar, and based on the Guinier plot at low q values there were no signs of aggregation or intermolecular repulsion for the Ca2+-free core and the Ca2+-bound core (Fig. 5A). The Kratky plots showed bell-shaped curves for both samples (Figs. S1&S2 D, E, F). They indicated that the calpain-3 protease cores are globular molecules and were well folded. Molecular masses of 42.02 and 42.8 kDa were calculated from the Porod volume*0.625 for the Ca2+-free and -bound samples, respectively. These values are close to the 43.3 kDa value calculated from the amino-acid sequence of the monomer.

Structural predictions based on sequence analysis (32) indicate IS1 lacks well-defined tertiary structure (Fig. 2). Therefore, the missing IS1 region was built in using the Ensemble Optimization Method (EOM) (33). A large conformational pool (10,000 possibilities) was generated using sequence information and the crystal structure of Ca2+-bound calpain-3 C129S core as a template, but with PC1 and PC2 unfixed and, therefore, free to rotate. However, only those conformers that were consistent with the SAXS experimental scattering data were retained. The final ensemble of conformers of the calpain-3 core with the IS1 structural model inserted was in excellent agreement with both experimental scattering data, chi^2 values of 1.304 and 0.497 corresponding to the Ca2+-free calpain-3 core and the Ca2+-bound core, respectively. The calculated flexibility of the ensemble conformers by EOM showed that the Ca2+-free core value was relatively high in solution (Rflex 63.2/Rσ 1.2), while the Ca2+-bound core was more compact (Rflex 58.0/Rσ 0.34). Both structural models from EOM were then used to calculate a CRYSOL scattering curve (34) to compare against experimental scattering. The curves of theoretic and experiment are well matched with chi^2 values of 1.94 and 0.62 for the Ca2+-free core and Ca2+-bound cores, respectively. Also, these ensemble models aligned quite closely with the ab initio envelope profiles developed from DAMAVER calculations (Fig. 6). However, the Ca2+-free core model shows a higher chi^2 value than that of the Ca2+-bound form, which indicates
that domains PC1 and 2 are more dynamic in the apo state than with Ca\(^{2+}\) bound.

In the structural model of the Ca\(^{2+}\)-free core, IS1 lies in the vicinity of the active site groove between PC1 and 2 (Fig. 7A&B). The N-terminal end of IS1 up to residue Thr270 enters the active site and is close to PC1. A little beyond the C terminus of IS1, structure is regained at Gln321 of PC2. In between these points the IS1 region forms an elongated random structure that protrudes out of the enzyme cleft, and is exposed to the solvent. The profile of the modeled IS1 is different in the Ca\(^{2+}\)-bound core, being more extended and flexible in the solvent space around the front of active site than in the Ca\(^{2+}\)-free sample (Fig. 7C&D). A change in IS1 disposition is to be expected because when Ca\(^{2+}\) binds to the core there is substantial rotation of the PC1 and 2 domains relative to each other and the ends of IS1 are shifted in their locations between the Ca\(^{2+}\)-bound and -free forms. Using the SAXS data, the distances between the IS1 ends in the Ca\(^{2+}\)-free and -bound forms are 33.4 Å and 16.5 Å, respectively. The BILBOMD program (35) was applied to monitor IS1 dynamic motions at a high temperature (1500K). PC1 and 2 were treated as rigid bodies while allowing IS1 to remain flexible. Rg was limited to the range between 20–26 Å and 400 conformations per Rg value were sampled. The Ca\(^{2+}\)-free and Ca\(^{2+}\)-bound structures resulted in large average r.m.s.d. values of 9.0 Å and 12.3 Å, respectively, when compared to the initial IS1 position, indicating that IS1 is flexible and can adopt diverse conformations.

Comparison of the P(r) functions (Pairwise distance distribution function) for both samples (Fig. 5B) shows that the calpain-3 cores have different distance distributions in solution even though the Ca\(^{2+}\)-free and Ca\(^{2+}\)-bound calpain-3 cores are globular molecules. The Ca\(^{2+}\)-free protein has a more extended shape (Dmax =84.7 Å and Rg = 25.2±0.04 Å) than the Ca\(^{2+}\)-bound version (Dmax=83.7 Å and Rg =24.9±0.04 Å). This suggests that there is some flexibility between the PC 1 and 2 domains in both Ca\(^{2+}\)-free and Ca\(^{2+}\)-bound states in solution. In the absence of supporting domains in the whole enzyme structure it seems that the PC1 and 2 can pivot around a hinge region (29), such that the two halves of the core are not always correctly aligned as they would be in the native state. Similar conclusions were reached with the isolated calpain-1 core. Flexibility was also observed by superposition of the crystal structure of Ca\(^{2+}\)-bound calpain-3 core on the solution structures with Ca\(^{2+}\) absent/present. These comparisons yielded r.m.s.d. values of 13.7 Å for the Ca\(^{2+}\)-free solution structure, and 7.6 Å for the Ca\(^{2+}\)-bound solution structure (data not show).

**Role of IS1 of calpain-3 protease core in auto-inhibition**—The basis for the inactivity of apo-calpains was immediately clear when the first crystal structures of calpain-2 were solved (10, 11). The catalytic cleft was not properly formed and remains this way until two Ca\(^{2+}\) bind cooperatively to sites in the core that causes a rotation of the PC1 and 2 domains such that the catalytic triad residues are properly aligned for catalysis (29). Based on modeling, the same calcium activation process would seem to be necessary to form the catalytic cleft of calpain-3. But in addition to the metal ion requirement it is known that calpain-3 has an additional inhibitory feature, which is IS1 acting as an internal autoinhibitory propeptide to block the enzyme’s active site (25). Although in the Ca\(^{2+}\)-bound crystal structure most of IS1 is not visible, what can be seen of IS1 in the active site cleft supports the auto-inhibition role. Firstly, the N-terminal segment of IS1 (Met272-Thr273-Tyr274-Gly275) protrudes into the active-site cleft through the S3, S2, S1 and S1 ′ sub-sites and makes extensive contacts with residues on either side of the catalytic cleft (Fig. 8). (See detailed description below). Secondly, residues Val320-Gln321-Tyr322 located just beyond the C terminus of IS1 are part of PC2 that lies on the top of the active-site, in which Val320 forms an interaction with residue Glu125 OE2 of PC1 (2.9 Å). Furthermore, an extended hydrophobic patch is formed by Val320, Tyr322 on top of the active site together with cleft residues Trp360, Pro359 and His334. Of these residues Trp360 NE1 interacts with carbonyl oxygen of Gly275 (3.3 Å) of IS1 and van de Waals contacts are formed between the catalytic His334 and Gly275 of IS1 (Fig. 8A). These contacts restrict access to the catalytic cleft and are not seen in the calpain-1 and calpain-2 cores. This region of IS1 that occupies the active site cleft is perfectly conserved in the comparison of the IS1 sequences from 22 different mammals, which reinforces its functional importance (Fig. S3).
Internal propeptide of calpain-3 is unstructured

Autolytic removal of IS1 paves the way for solving inhibitor-bound calpain-3 core structures—Previous studies by our group proposed that the loss of IS1 was required for calpain-3 activation. Pre-incubation of the calpain-3 core with Ca²⁺ present increased its proteolytic activity over a 24-h period as IS1 was autolysed away (25). At this time point calpain activity could be inhibited by adding cysteine protease inhibitors E-64 and leupeptin. To obtain inhibitor-bound structures for the calpain-3 core (C129) the autolysis reaction was allowed to proceed for even longer (5 days) before reacting the Ca²⁺-bound core with inhibitors E-64 and leupeptin prior to crystallization.

Cysts of the core bound to E-64 and leupeptin diffracted to 2.8 Å- and 3.2 Å-resolution, respectively. Both crystal types had four molecules in the asymmetric unit and had the same space group (P2₁2₁2₁) as the inhibitor-free structures (C129S-core-Ca²⁺ and C129A core-Ca²⁺). Thus the molecular replacement method was performed using the C129S-core-Ca²⁺ structure as a search model. The initial 2Fo-Fc electron density maps for both co-crystals showed continuous density in the active site; subsequently the E-64 molecule and leupeptin molecule were unambiguously placed there and were well defined (Table 2). Electron density was missing for the N-terminal 8-13 residues of the construct in all four molecules of both complex structures. Almost the entire IS1 (residues 268-324) was missing except in molecule A of the calpain-3 core-Ca²⁺-E-64 complex structure (missing residues 270-321) and in molecule D (missing residues 269-321) from the structure of calpain-3 core-Ca²⁺-leupeptin complex. Molecules A and D are described in the text below unless otherwise noted. Both complex structures display similar overall architecture to the inhibitor-free structure of C129S-core-Ca²⁺ with r.m.s.d. values of 0.65 Å (E-64) and 0.41 Å (leupeptin), using 314 and 306 Ca atoms for alignment, respectively.

The E-64 molecule is accommodated in the active site cleft of the enzyme core by occupying the S1, S2 and S3 sub-sites in place of IS1, after the latter has been autolysed away (Fig. 9). Note that the N-terminal region of IS1 is not observed because it has been displaced from the active site by the inhibitor and is too disordered to have stable electron density. The N-terminal epoxy and central leucyl moieties of E-64 fit well into the electron density, while the C-terminal 4-guanidinobutane moiety of E-64 has weak electron density that suggests this region is flexible (Fig. 10A). There is strong continuous electron density between the C2 of E-64’s epoxy group and the catalytic Cys129 S7, which directly confirms covalent bond formation (1.64 Å) between E-64 and calpain-3. One of the oxygen atoms from the epoxysuccinic motif of E-64 is pointing to the oxyanion hole and forms hydrogen bonds with Gln123 NE2 (2.51 Å) and the main-chain amide of Cys129 (3.0 Å) (Fig. 9A). In addition, it makes an electrostatic contact with the main-chain amide of Asp128 (3.4 Å). The second oxygen atom of the epoxysuccinic motif produces a hydrogen bond with catalytic His334 ND1 (2.9 Å) at the P1-S1 site. Also, the main-chain amide of the leucyl group of E-64 makes electrostatic contacts with the main-chain carbonyl group of the conserved Gly222 (3.3 Å) from PC1 and with the conserved Gly333 (3.5 Å) from PC2 at the P2-S2 sub-site. The 4-guanidinobutane moiety of E-64 is pointing to the solvent and no P-S contact is observed in the S3 sub-site.

The leupeptin molecule is also positioned in the active site pocket occupying the S1, S2 and S3 sub-sites with a good fit to the electron density map (Fig. 10B). Leupeptin adopts the same orientation as natural substrates, whereas E-64 is in the reversed orientation. The sequence Leu-Leu-Arginal binds to the S3-S2-S1 sites, respectively. Continuous electron density is seen between the aldehyde carbonyl group of leupeptin and the catalytic Cys129 S7 (Fig. 10B). The density indicates a 2.0-Å hemithioacetal bond is formed between leupeptin and calpain-3. The interactions between leupeptin and the calpain-3 protease core are weaker than those observed in the E-64-bound calpain-3 protease core (Fig. 9B). The hemiacetal oxygen atom of leupeptin faces the oxyanion hole and forms three electrostatic contacts with the main-chain amide of catalytic Cys129 (3.6 Å), Gln123 NE2 (3.5 Å) and catalytic His334 (3.9 Å), respectively. Elsewhere in the active-site pocket, P1 Arg3 interacts with the carbonyl oxygen of Gly333 (3.2 Å) of PC2. Additionally, the backbone of the P2 Leu2 residue makes a parallel electrostatic contact with the main-chain amide and carbonyl oxygen of Gly222 (3.0 Å and 3.4 Å) of PC1 at the P2-S2 sub-site. However, the N-terminal acetyl-leucine moiety of leupeptin makes no contact with
neighboring residues in the S3 sub-site, much the same way the C-terminal 4-guanidinobutane moiety of E-64 fails to contact the S3 sub-site in its complex (Fig. 9B).

Solving the inhibitor-bound structures has also given us a chance to see if, and how, the residual N-terminal IS1 section might interfere with substrate or inhibitor binding to the cleft of the calpain-3 core. Comparing the structures of the residual IS1 in the Ca^{2+}-bound C129S core with the E-64/leupeptin-bound C129 core structures, the backbone of Thr273-Tyr274 from IS1 overlaps well on the P1 to P2 moieties of E-64 and leupeptin, but their side chains are in different orientations and make different contacts with S1’ to S2 sub-site residues (Fig. 8). The carbonyl oxygen of Gly275 of IS1 projects into the S1’ site and forms polar contacts with Gln123 NE2 (3.2 Å) and Trp360 NE1 (3.4 Å), while the Gly275 α-carbon also makes van der Waals contacts with catalytic His334 CE1 (3.2 Å). The latter two interactions were not seen in the inhibitor complex structures (Fig. 9). Interestingly, Tyr274 of IS1 takes over the role of epoxysuccinic motif of E-64 and Arg3 of leupeptin in the complex structures. Although there is no covalent bond or intermediate hemiacetal bond formed between the Ser129 side-chain (vs. active Cys129 in the complexes) and IS1, the amide of Tyr274 of IS1 makes an electrostatic contact with the Ser129 Oγ group (3.5 Å). Furthermore, the main-chain carbonyl oxygen of Tyr274 from IS1 form hydrogen bonds with Gln123 NE2 (3.1 Å), the amide of Ser129 (2.8 Å) and the amide of Asp128 (3.3 Å) from PC1, respectively (Fig. 8B). The amide of Tyr274 from IS1 also makes a hydrogen bond with the carbonyl oxygen of Gly333 (3.2 Å) from PC2 at the S1 sub-site. These bonds between IS1 and residues of the S1 sub-site clearly show the residual IS1 fragment blocking the catalytic center of the enzyme. In further keeping with this stability, Thr273 of IS1 imitates the role of the P2 Leucine residue of the inhibitors in making a double electrostatic contact with conserved Gly222 (3.3 Å and 3.6 Å) from PC1 and van der Waals contacts with the conserved Gly333 from PC2 at the S2 site. Moreover, additional contacts are seen between IS1 and residues of the S3 sub-site. These are Asn271 OD1 of IS1 with Lys410 NZ (3.1 Å) and with Asn223 ND2 (4.0 Å), respectively. By comparing these results, it seems that IS1 is potentially more stable in the active site as an internal auto-inhibitory propeptide than are the exogenous inhibitors E-64 and leupeptin. What gives these inhibitors the edge in occupying the catalytic cleft is that they can form covalent bonds with Cys129. Although the residual portion of IS1 makes a network of interactions in the active site, these connections can be broken when the PC1 and 2 domains rotate around each other. However, this N-terminal IS1 peptide can obviously compete with substrates and inhibitors for access to the catalytic cleft whenever it forms.

**DISCUSSION**

Activating cleavage of the internal IS1 propeptide in calpain-3 is regulated by calcium signaling—In the crystal structure of Ca^{2+}-bound calpain-3 C129S protease core, the IS1 region, with the exception of the N-terminal eight residues, cannot be observed. According to SDS-PAGE analysis of Ca^{2+}-bound calpain-3 C129S protease core crystals sampled, and the time course of autoproteolysis of the enzyme in solution, the calpain-3 C129S core with Ca^{2+} present shows weak protease activity. The explanation for this activity is that the serine replacement for cysteine in the catalytic triad can be deprotonated and act as a nucleophile to hydrolyze IS1. This activity came to light during the weeks required for crystallization, and could potentially occur within the crystal since it is a zero-order reaction. C129 catalyzed cleavage of IS1 is observable in a matter of minutes, and is complete within hours (25, 26). This explanation is supported by the C129A mutation completely abolishing the protease activity of the calpain-3 core. In all four active structures presented here Ca^{2+} ions are located in the expected sites of PC1 and 2. The core structures have clearly undergone the same cooperative realignment of the two core domains seen in the calpain-1 core when Ca^{2+} ions bind that brings the catalytic triad residues into alignment for catalysis. At least in this in vitro artificial system, where the calpain-3 core is examined in isolation, the driving force for activation is Ca^{2+} signaling and the IS1 peptide can only delay access to the catalytic cleft but not prevent it.

SAXS solution structural models of C129A, with and without Ca^{2+} bound, show that most of IS1 is a long extended flexible loop exposed to the solvent. Only a short N-terminal portion of IS1 is
visible and is located in the active site cleft as shown by the crystal structures. We had earlier anticipated that IS1 would be structured in at least the calcium-free form, since there is a central region of IS1 with high alpha-helix propensity (25). Also, the protein sequence alignment of IS1 from humans and 21 other mammals shows most of the insertion sequence has been well conserved during the evolution with 87.7% identity and 95.2% similarity (Fig. S3). Such conservation suggests that IS1 performs an important biological function rather than simply be a disordered protein loop. In general, molecular recognition can induce unstructured proteins to fold in the cell. The folding transition can occur when the unstructured protein region binds to its target, such as other proteins or a nucleic acid (36). Indeed, this is how calpastatin recognizes and inhibits calpain-2. Based on SAXS solution structural models with/without Ca$^{2+}$ bound and the Ca$^{2+}$-bound crystal structures, we propose that IS1 is intrinsically unstructured as predicted by PSIPRED (27, 28) (Fig. 2) and has affinity for an unknown protein. In a calcium-free environment, IS1 of the protease core could associate with the unknown partner and form a stable complex that is enzymatically inactive. When there is an influx of calcium ions to the sarcomere and two Ca$^{2+}$ bind to the calpain-3 core they could activate the enzyme through reorientation of PC1 and 2. In this model the IS1-binding protein could potentially be released from the Ca$^{2+}$-activated enzyme, allowing IS1 to be cut to help open up the cleft to calpain-3 substrates. Another possibility is that the binding partner serves to hold the IS1 loop away from the cleft making it fully accessible to substrates. In this model there would be no need for proteolysis. Activation would be fully reversible without a physical change to the enzyme. Another advantage of this model is that the N-terminal end of IS1 would not be lying in the catalytic cleft as it does after cleavage, needing to be displaced by substrates. It used to be thought that autoproteolysis was needed for activation of calpains-1 and -2, but this was disproven by the structure of the activated calpain-2 in complex with calpastatin showing that the anchor helix is released on calcium binding, rather than being proteolysed away (37, 38). Again, the advantage of a non-proteolytic activation is a return to the inactive state after calcium signaling without damage or loss to the calpain, so that the system is ready to respond to the next signaling event.

**IS1 as a guide for designing inhibitors of calpain-3**—Crystal structures show the N-terminal portion of IS1 is located in the active site cleft after cleavage, with the residues closest to the cleavage site making extensive hydrophilic interactions with residues of PC1 and 2 from the S3 to S1´ sub-sites, as detailed in Results. Although E-64 and leupeptin bind in the active site of calpain-3 core and form a covalent bond to the catalytic Cys129, both exogenous inhibitors have fewer contacts with residues of the active site compared with N terminus of IS1. Thus, this section of IS1 can serve as a model for the design of a peptidomimetic inhibitor of calpain-3. For exogenous calcium-dependent cysteine-protease inhibitors, the irreversible/reversible covalent bond between inhibitor and catalytic cysteine thiol is crucial for effectiveness (39). IS1 forms a weak electrostatic contact with Oγ group of Ser129 (catalytic Cys129), which suggests the possible location of a warhead to turn the IS1 peptide from a competitor peptide into a covalent inhibitor.

E-64 is an epoxy succinic inhibitor with broad effectiveness against cysteine proteases. It forms an irreversible covalent bond with the active-site cysteine thiol. Leupeptin is a peptide aldehyde inhibitor against serine, cysteine and threonine proteases and makes a transient reversible covalent bond with the cysteine thiol. Both inhibitors had no effect on calpain-3 autolysis, presumably because they could not displace the uncleaved IS1 from the catalytic cleft (12). The calpain-3 splice variant Lp82, which lacks IS1, could be inhibited by E-64 (40, 41), and the calpain-3 isoform from human peripheral blood mononuclear cells (PBMC), which lacks IS1 and 2, could also be inhibited by leupeptin (42). Although E-64 and leupeptin bind in the active site of calpain-3 core after IS1 cleavage, and form a covalent bond to the catalytic Cys129, both exogenous inhibiators have many fewer contacts with residues of the active site than does the N terminus of IS1. The electrostatic interactions made by E-64 and leupeptin in the calpain-3 core are also fewer than those made in the calpain-1 core crystal structures (43). This further suggests that E-64 and leupeptin are not optimal active-site inhibitors of the calpain-3 protease core. Thus, based on IS1 there are options for designing
superior inhibitors of calpain-3 that might compete with IS1 and block activity even before the initial autolytic event.

**EXPERIMENTAL PROCEDURES**

*Site-directed mutagenesis, expression, and purification of human calpain-3 protease core—*

The protease core (PC1 and PC2) of human calpain-3 active form Cys129 with its 45-residue N-terminal extension sequence deleted (ΔNS), and the same construct with its catalytic Cys129 mutated to Ser (C129S) have been previously described (25, 26). These two constructs include residues Ile46 to Asp419 with a C-terminal 6xHis tag. In a third construct the catalytic Cys129 was mutated to Ala using the QuikChange site-directed mutagenesis method (Stratagene) according to the manufacturer’s protocol. The three calpain-3 protease core constructs ΔNS C129, C129S and C129A were expressed and purified as previously described (26).

*Derivatization of calpain-3 protease core ΔNS C129 with inhibitors E-64 and leupeptin—*

Calpain-3 protease core ΔNS C129 was purified through the DEAE anion-exchange, Ni-NTA, and Sephacryl S-200 size-exclusion chromatography steps. Fractions containing purified calpain-3 protease core were combined and analyzed by 12% SDS–PAGE, and then incubated in 10 mM CaCl$_2$, 25 mM Tris-HCl pH 7.6 and 150 mM NaCl for 5 days at room temperature until IS1 autolysis reached its maximum extent as judged by SDS–PAGE. Subsequently a ten-fold molar excess of inhibitor E-64 was mixed with autolyzed protein solution and incubated for 3 h at room temperature. The derivatized protein was further purified by Q-Sepharose chromatography (Pharmacia). All purified proteins were buffer exchanged into 10 mM Na HEPES (pH 7.6), 10 mM DTT. For making the complex with leupeptin, after 5 days of IS1 autolysis, the protein was mixed with a ten-fold molar excess of leupeptin and incubated for 1 h. Crystallization trials were done without further purification of the leupeptin-modified core.

Protein solutions of C129A for SAXS experiments were dialyzed exhaustively against 10 mM Na HEPES (pH 7.6), 3% glycerol, 5 mM DTT and 2 mM CaCl$_2$, (or without CaCl$_2$ but with 5 mM EDTA instead). Protein molarities were determined by absorbance at 280 nm using a Nano-drop spectrometer with a molecular mass 43318.08 Da and a theoretical molar extinction coefficient of 80370 M$^{-1}$ cm$^{-1}$.

*Molecular weight determination of calpain-3 core from protein crystals—*

Calpain-3 protease core C129S crystals were transferred to a new drop containing crystallization well solution devoid of calpain-3 core C129S solution. The crystals were then washed three more times with well solution. Subsequently, these crystals were dissolved in Millipore water and subjected to both SDS-PAGE and MALDI-TOF Mass spectrometry analysis.

*Autoproteolysis assays—*

The degree of autoproteolysis of calpain-3 protease core constructs was assessed by SDS-PAGE. Calpain-3 protease core C129S and C129A (48 µM final concentration) were tested for autolysis in 0.1M MES (pH 6.5), 0.1M CaCl$_2$. This buffer mimicked crystallization conditions except for the absence of PEG 8000, which affected protein band visualization on SDS-PAGE. The various additives assessed for their effect on protease core C129S autolysis included: the cysteine protease inhibitor E-64 (0.02 mM), serine protease inhibitor PMSF (0.2mM), chelating agent EDTA (10mM) and chemicals that could potentially protect the protein from degradation, namely 100 mM NaCl, 0.2% Na$_2$N$_2$O$_4$ and 10% glycerol. Samples were incubated at 22 °C in a final volume of 60 µl. Aliquots were removed at various times (0, 1, 3, 6, 14, 21 and 28 days) and the reaction was stopped by the addition of 3x SDS-PAGE loading buffer, which contained 187.5 mM Tris-HCl (pH 6.8), 6% SDS, 30% glycerol, and 0.03% bromophenol blue. Protein bands were visualized after 12% SDS-PAGE by Coomassie Blue staining. The band-intensity was measured by densitometry. The control assay consisted of calpain-3 core C129S or C129A incubated with 5 mM EDTA in place of CaCl$_2$ as described above.

*Crystallization and data collection—*

Purified proteins were concentrated to 9-13 mg/ml for crystallization trials. The hanging drop vapor diffusion method was used at 22°C. Precipitant conditions were 8% - 11% (w/v) PEG 8000, 0.1 M MES (pH 6.5) and 0.1 M CaCl$_2$. Irregular, soft
crystals of protease core ΔNS C129S were initially obtained after four weeks. Streak-seeding was carried out immediately after new drops were set up. Through several rounds of seeding, rectangular shaped crystals grew over a 4-week period to reach diffraction size. For protease core ΔNS C129A crystallization, a cross-seeding procedure was performed. Initial seeds were from the C129S crystals, afterwards the seeds were all from C129A crystals. Clusters of long needle-shaped crystals of C129 – E-64 derivative and crystals of co-crystallized C129 – Leupeptin appeared within one week and reached diffraction size in three weeks. Because all of the crystals were fragile and diffracted poorly, dehydration was used. Crystals were transferred into stabilizing solutions that contained 18% (w/v) PEG 8000, 0.1 M MES (pH 6.5), 0.1 M CaCl2 and sorbitol, with the latter increasing in 5% increments from 5% to 30% over a period of 3 days, before crystals were flash-frozen in liquid nitrogen.

**Structural determination and refinement**—For crystals of calpain-3 protease core ΔNS C129S and C129A, X-ray diffraction data were collected on the X6A beam line at the Brookhaven National Synchrotron Light Source (NSLS). For crystals of calpain-3 protease core ΔNS C129 – E-64 derivative and – Leupeptin derivative, X-ray diffraction data were collected on 23ID-B beam line at the Argonne National Laboratory (APS). All crystals belonged to the P2_1_2_1_2 space group with four molecules per asymmetric unit. The datasets were indexed, integrated and scaled using XDS (44). Molecular replacement calculations were carried out with PHASER (45) through the CCP4 GUI (46). Calpain-2 protease core structure (PDB ID code 1MDW), but with Ca^{2+} omitted, was used as a search model. The structure was then refined with REFMAC5 (47) and PHENIX (48). All inspections and manual manipulations were completed with COOT (49, 50). Figures were generated in PyMOL (www.pymol.org). Crystallographic data collection and refinement statistics are summarized in Table 2.
Acknowledgements: We thank Dr. John Allingham for access to his home source X-ray diffractometer, and acknowledge the synchrotron facilities at Brookhaven (NSLS), Chicago (APS), Cornell (CHESS), USA and Grenoble, France (ESRF) for assistance with data collection. We are grateful to Sherry Gauthier for technical assistance with site-directed mutagenesis and to Dr. Shuaiqi Guo for SAXS data collection at ESRF. This research was funded by a grant from the Canadian Institutes of Health Research to P.L.D., who holds the Canada Research Chair in Protein Engineering.

Conflict of interests: The authors declare that they have no conflict of interest with the contents of this article.

Author contributions: QLY designed and performed the experiments, solved the structures, analyzed data and drafted the manuscript. RLC helped with the design of the experiments and structure analysis, and revised the manuscript. PLD conceived the experiments, helped analyze and interpret the results, and edited the manuscript. All authors reviewed the results and approved the final version of the manuscript.

References

1. Ono, Y., Saito, T. C., and Sorimachi, H. (2016) Calpain research for drug discovery: challenges and potential. Nat. Rev. Drug Discov. 15, 854–876
2. Campbell, R. L., and Davies, P. L. (2012) Structure-function relationships in calpains. Biochem. J. 447, 335–351
3. Croall, D. E., and Ersfeld, K. (2007) The calpains: modular designs and functional diversity. Genome Biol. 8, 218
4. Glading, A., Lauffenburger, D. A., and Wells, A. (2002) Cutting to the chase: calpain proteases in cell motility. Trends Cell Biol. 12, 46–54
5. Mellgren, R. L., Miyake, K., Kramerova, I., Spencer, M. J., Bourg, N., Bartoli, M., Richard, I., Greer, P. A., and McNeil, P. L. (2009) Calcium-dependent plasma membrane repair requires m- or mu-calpain, but not calpain-3, the proteasome, or caspases. Biochim. Biophys. Acta. 1793, 1886–1893
6. Khoutorsky, A., and Spira, M. E. (2009) Activity-dependent calpain activation plays a critical role in synaptic facilitation and post-tetanic potentiation. Learn. Mem. Cold Spring Harb. N. 16, 129–141
7. Ono, Y., and Sorimachi, H. (2012) Calpains: an elaborate proteolytic system. Biochim. Biophys. Acta. 1824, 224–236
8. Maki, M., Kitaura, Y., Satoh, H., Ohkouchi, S., and Shibata, H. (2002) Structures, functions and molecular evolution of the penta-EF-hand Ca2+-binding proteins. Biochim. Biophys. Acta. 1600, 51–60
9. Maki, M., Maemoto, Y., Osako, Y., and Shibata, H. (2012) Evolutionary and physical linkage between calpains and penta-EF-hand Ca2+-binding proteins. FEBS J. 279, 1414–1421
10. 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
11. Strobl, S., Fernandez-Catalan, C., Braun, M., Huber, R., Masumoto, H., Nakagawa, K., Irie, A., Sorimachi, H., Bourenkow, G., Bartunik, H., Suzuki, K., and Bode, W. (2000) The
Internal propeptide of calpain-3 is unstructured

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

12. Kinbara, K., Sorimachi, H., Ishiura, S., and Suzuki, K. (1998) Skeletal muscle-specific calpain, p49: structure and physiological function. Biochem. Pharmacol. 56, 415–420

13. Kinbara, K., Ishiura, S., Tomioka, S., Sorimachi, H., Jeong, S. Y., Amano, S., Kawasaki, H., Kolmerer, B., Kimura, S., Labeit, S., and Suzuki, K. (1998) Purification of native p94, a muscle-specific calpain, and characterization of its autolysis. Biochem. J. 335 (Pt 3), 589–596

14. Ravulapalli, R., Diaz, B. G., Campbell, R. L., and Davies, P. L. (2005) Homodimerization of calpain 3 penta-EF-hand domain. Biochem. J. 388, 585–591

15. Ravulapalli, R., Campbell, R. L., Gauthier, S. Y., Dhe-Paganon, S., and Davies, P. L. (2009) Distinguishing between calpain heterodimerization and homodimerization. FEBS J. 276, 973–982

16. Partha, S. K., Ravulapalli, R., Allingham, J. S., Campbell, R. L., and Davies, P. L. (2014) Crystal structure of calpain-3 penta-EF-hand (PEF) domain - a homodimerized PEF family member with calcium bound at the fifth EF-hand. FEBS J. 281, 3138–3149

17. Ono, Y., and Sorimachi, H. (2015) Amino acid sequence alignment of vertebrate CAPN3/calpain-3/p94. Data Brief. 5, 366–367

18. Sorimachi, H., Kinbara, K., Kimura, S., Takahashi, M., Ishiura, S., Sasagawa, N., Sorimachi, N., Shimada, H., Tagawa, K., and Maruyama, K. (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

19. 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- and mu-types. Specific expression of the mRNA in skeletal muscle. J. Biol. Chem. 264, 20106–20111

20. Ono, Y., Ojima, K., Shinkai-Ouchi, F., Hata, S., and Sorimachi, H. (2016) An eccentric calpain, CAPN3/p94/calpain-3. Biochimie. 122, 169–187

21. Richard, I., Broux, O., Allamand, V., Fougerousse, F., Chiannilkulchai, N., Bourg, N., Brenguier, L., Devaud, C., Pasturaud, P., and Roudaut, C. (1995) Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. Cell. 81, 27–40

22. Kramerova, I., Beckmann, J. S., and Spencer, M. J. (2007) Molecular and cellular basis of calpainopathy (limb girdle muscular dystrophy type 2A). Biochim. Biophys. Acta. 1772, 128–144

23. Jia, Z., Petrounevitch, V., Wong, A., Moldoveanu, T., Davies, P. L., Elce, J. S., and Beckmann, J. S. (2001) Mutations in calpain 3 associated with limb girdle muscular dystrophy: analysis by molecular modeling and by mutation in m-calpain. Biophys. J. 80, 2590–2596

24. Garnham, C. P., Hanna, R. A., Chou, J. S., Low, K. E., Gourlay, K., Campbell, R. L., Beckmann, J. S., and Davies, P. L. (2009) Limb-girdle muscular dystrophy type 2A can result from accelerated autoproteolytic inactivation of calpain 3. Biochemistry (Mosc.). 48, 3457–3467

25. 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

26. Rey, M. A., and Davies, P. L. (2002) The protease core of the muscle-specific calpain, p94, undergoes Ca2+-dependent intramolecular autolysis. FEBS Lett. 532, 401–406

27. Jones, D. T., and Cozzetto, D. (2015) DISOPRED3: precise disordered region predictions with annotated protein-binding activity. Bioinforma. Oxf. Engl. 31, 857–863

28. McGuffin, L. J., Bryson, K., and Jones, D. T. (2000) The PSIPRED protein structure prediction server. Bioinforma. Oxf. Engl. 16, 404–405
Internal propeptide of calpain-3 is unstructured

29. 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*. **108**, 649–660
30. Moldoveanu, T., Hosfield, C. M., Lim, D., Jia, Z., and Davies, P. L. (2003) Calpain silencing by a reversible intrinsic mechanism. *Nat. Struct. Biol.* **10**, 371–378
31. Davis, T. L., Walker, J. R., Finerty, P. J., Mackenzie, F., Newman, E. M., and Dhe-Paganon, S. (2007) The crystal structures of human calpains 1 and 9 imply diverse mechanisms of action and auto-inhibition. *J. Mol. Biol.* **366**, 216–229
32. Buchan, D. W. A., Minneci, F., Nugent, T. C. O., Bryson, K., and Jones, D. T. (2013) Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res.* **41**, W349-357
33. Tria, G., Mertens, H. D. T., Kachala, M., and Svergun, D. I. (2015) Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ*. **2**, 207–217
34. Svergun, D., Barberato, C., and Koch, M. H. J. (1995) CRYSOI – a Program to Evaluate X-Ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Crystallogr.* **28**, 768–773
35. Pelikan, M., Hura, G. L., and Hammel, M. (2009) Structure and flexibility within proteins as identified through small angle X-ray scattering. *Gen. Physiol. Biophys.* **28**, 174–189
36. Wright, P. E., and Dyson, H. J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293**, 321–331
37. Hanna, R. A., Campbell, R. L., and Davies, P. L. (2008) Calcium-bound structure of calpain and its mechanism of inhibition by calpastatin. *Nature*. **456**, 409–412
38. Moldoveanu, T., Gehring, K., and Green, D. R. (2008) Concerted multi-pronged attack by calpastatin to occlude the catalytic cleft of heterodimeric calpains. *Nature*. **456**, 404–408
39. Matsumoto, K., Mizoue, K., Kitamura, K., Tse, W. C., Huber, C. P., and Ishida, T. (1999) Structural basis of inhibition of cysteine proteases by E-64 and its derivatives. *Biopolymers*. **51**, 99–107
40. Nakamura, Y., Fukiage, C., Ma, H., Shih, M., Azuma, M., and Shearer, T. R. (1999) Decreased sensitivity of lens-specific calpain Lp82 to calpastatin inhibitor. *Exp. Eye Res.* **69**, 155–162
41. Fukiage, C., Nakajima, E., Ma, H., Azuma, M., and Shearer, T. R. (2002) Characterization and regulation of lens-specific calpain Lp82. *J. Biol. Chem.* **277**, 20678–20685
42. De Tullio, R., Stifanese, R., Salamino, F., Pontremoli, S., and Melloni, E. (2003) Characterization of a new p94-like calpain form in human lymphocytes. *Biochem. J.* **375**, 689–696
43. Moldoveanu, T., Campbell, R. L., Cuerrier, D., and Davies, P. L. (2004) Crystal structures of calpain-E64 and -leupeptin inhibitor complexes reveal mobile loops gating the active site. *J. Mol. Biol.* **343**, 1313–1326
44. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr*. **66**, 125–132
45. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr*. **40**, 658–674
46. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr*. **67**, 235–242
47. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr*. **67**, 355–367
48. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart,
Internal propeptide of calpain-3 is unstructured

P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221

49. Emsley, P. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132

50. Incardona, M. -F., Bourenkov, G. P., Levik, K., Pieritz, R. A., Popov, A. N., and Svensson, O. (2009) EDNA: a framework for plugin-based applications applied to X-ray experiment online data analysis. J. Synchrotron Radiat. 16, 872–879

51. Nielsen, S. S., Toft, K. N., Snakenborg, D., Jeppesen, M. G., Jacobsen, J. K., Vestergaard, B., Kutter, J. P., and Arleth, L. (2009) BioXTAS RAW, a software program for high-throughput automated small-angle X-ray scattering data reduction and preliminary analysis. J. Appl. Crystallogr. 42, 959–964

52. Svergun, D. I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys. J. 76, 2879–2886

53. Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J. Appl. Crystallogr. 25, 495–503

54. Svergun, D. I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophys. J. 76, 2879–2886

55. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. J. Appl. Crystallogr. 36, 860–864

56. Karplus, P. A., and Diederichs, K. (2012) Linking crystallographic model and data quality. Science. 336, 1031-1033

FOOTNOTES

Atomic coordinates and structure factors have been deposited in the Protein Data Bank (http://www.rcsb.org) under the codes 6BDT (C129S mutant of the calpain-3 protease core), 6BGP (C129A mutant of the calpain-3 protease core), 6BJD (calpain-3 protease core in complex with E-64) and 6BKJ (calpain-3 protease core in complex with leupeptin).

FIGURE LEGENDS

Figure 1. Schematic structures of conventional calpains-1 and -2, and calpain-3. (A) Conventional calpains-1 and -2 form heterodimers through the penta-EF-hand domains of the large PEF(L) and small PEF(S) subunits. (B) Calpain-3 forms a homodimer through PEF(L) domain connections, and has unique sequences of NS, IS1 and IS2 shown in grey. The protease core domains (PC1 and 2) are colored orange and yellow, respectively. C, H and N are the catalytic triad residues, Cys, His and Asn. The red section preceding PC1 is the anchor helix that contacts the small subunit. Other domains are the calpain beta-sandwich domain (CBSW) in green; the PEF(L) domain in light blue; the glycine-rich domain (GR) in pink; and PEF(S) in purple. Numbers indicate residues at the domain boundaries. Arrows indicate autolysis sites (1b and 2b) in IS1.

Figure 2. Prediction of disorder in the calpain-3 protease core sequence. The calpain-3 protease core sequence, including NS, was analysed for the probability of disorder using a web-based program (http://www.bioinf.cs.ucl.ac.uk). The likelihood of a region being disordered is plotted as a confidence score against the amino acid sequence. The blue tracing represents sequence in a disordered state and the orange tracing corresponds to disordered residues likely to be involved in protein – protein interactions.
Internal propeptide of calpain-3 is unstructured

Figure 3. Crystal structures of calpain-3 protease core C129S (A), and C129 in complex with E-64 (B). Ribbon tracings colored orange and yellow represent domains PC1 and 2 of the calpain-3 protease core, respectively. Residual IS1 in (A) and (B) is coloured blue. Green stick represents E-64 in (B). Catalytic residues, and the visible N- and C-termini of IS1 and the protease core are labeled with their residue numbers. Magenta spheres represent Ca^{2+}.

Figure 4. Time course of autoproteolysis of calpain-3 protease cores analyzed by SDS-PAGE. Lanes kDa, M and C represent molecular masses, molecular mass markers and a control protease core without proteolysis (day 0), respectively. All protein samples were incubated in 0.1M MES (pH 6.5) and 0.1 M CaCl_2. (A) Analysis of C129S core crystals. (B) C129S core in solution. (C) C129A core in solution. (D) – (I) C129S core in solution incubated with the protease inhibitors or chemicals indicated below the gel images. Band intensity as measured by densitometry is shown below the gel images of (B) – (I). Blue, orange, grey and yellow lines represent 43-kDa, 26.6-kDa, 17-kDa and 13-kDa bands, respectively. The time course of the incubations was measured in days (1-28) indicated above the gel images.

Figure 5. SAXS analyses of Ca^{2+}-free and Ca^{2+}-bound calpain-3 C129A protease core. (A) Guinier plot scattering patterns of the C129A calpain-3 core. Ca^{2+}-free and Ca^{2+}-bound samples are in green and red, respectively. (B) Pair distribution functions P(r) of Ca^{2+}-free (green) and Ca^{2+}-bound (red) samples. Both samples give similar maximum dimensions (Dmax) for the particles (84.7 Å for Ca^{2+}-free and 83.7 Å for Ca^{2+}-bound, but the P(r)-distribution profiles are different.

Figure 6. Agreement between structural models and SAXS experimental scattering patterns for the calpain-3 core. A) and B) are the theoretical scattering patterns calculated by CRYSOL from structural models based on the Ensemble Optimization Method for the Ca^{2+}-free and Ca^{2+}-bound calpain-3 cores, respectively. C) and D) show solution structural models fit to envelops, which correspond to Ca^{2+}-free and Ca^{2+}-bound sample, respectively. Green ribbon represents the PC1 and 2 domains of the calpain-3 core. Red spheres represent IS1 residues. N- and C-termini of the core are indicated by N and C, respectively.

Figure 7. Calpain-3 core structures derived from SAXS analysis. The C129A calpain-3 protease core solution structure is shown in surface representation with PC1 colored orange, PC2 colored yellow, and IS1 in red. A) and B) are Ca^{2+}-free structure shown with a 40° difference in orientation. C) and D) are the Ca^{2+}-bound structure in the same orientations as A) and B), respectively.

Figure 8. Contacts between IS1 and residues of calpain-3 C129S protease core active-site cleft. Ribbon and stick in orange represent PC1, and yellow represents PC2. The ribbon and stick in green are IS1. The black dashed lines indicate hydrophilic interactions between residues. A) displays the active-site cleft from S5’ to S3 subsites and B) displays the active-site cleft from S1’ to S3 subsites.

Figure 9. Interactions of inhibitors in the active-site of the calpain-3 protease core. Orange and yellow ribbon and stick represent domains PC1 and PC2 of the calpain-3 core, respectively. The cyan stick in A) is E-64, and the grey stick in B) is leupeptin. Black dashes indicate the electrostatic interactions between E-64/leupeptin and residues in active-site of the calpain-3 core. Magenta spheres represent Ca^{2+}. 15
Internal propeptide of calpain-3 is unstructured

Figure 10. Covalent attachment of inhibitors to the catalytic C129 of the calpain-3 core. A) E-64 and B) leupeptin are in green; Cys129 is in gray. Bonds between inhibitors and Cys129 are in yellow. O atoms are red and N atoms are blue. The light blue mesh shows the 2Fo – Fc electron density map contoured at the 1σ level.
Table 1. Small-angle X-ray scattering parameters

| Data-collection                  | Calpain-3 protease core Ca²⁺-bound | Calpain-3 protease core Ca²⁺-free |
|--------------------------------|-------------------------------------|----------------------------------|
| Instrument                      | ESRF BM29                           | CHESS G1                         |
| Detector distance (m)           | 2.867                               | 1.5027                           |
| Wavelength (Å)                  | 0.9918                              | 1.055                            |
| q range (Å⁻¹)                   | 0.025-5.0                           | 0.007-0.7                        |
| Exposure time per frame (s)     | 1                                   | 1                                |
| Frames per concentration        | 10                                  | 20                               |
| Measured concentrations (mg/ml) | 1, 3, 5                             | 1, 3, 5                          |
| Temperature (k)                 | 287.15                              | 277.15                           |
| Structural parameters for averaged data |                                |                                  |
| I(0) (from Guinier analysis)    | 0.43±0.03                           | 1.83±0.13                        |
| Rg (Å) (from Guinier analysis)  | 24.7±0.01                           | 25.3±0.06                        |
| I(0) (from P(r))                | 4.07±0.05                           | 1.84±0.18                        |
| Rg (Å) (from P(r))              | 24.9±0.04                           | 25.2±0.04                        |
| Dmax (Å)                        | 83.7 (error 2.5%)                   | 84.7 (error 1.5%)                |
| Porod volume (Å³)               | 68481.2                             | 67229.2                          |
| Molecular mass [from (Vp*0.625)] (kDa) | 42.8                                | 42.02                            |
| Monomer molecular mass from sequence (kDa) | 43.3                                | 43.3                             |

Software employed

Primary data reduction: Automated SAXS pipeline
Data processing: ATSAS package
Ab initio analysis: DAMMIF
Validation and averaging: DAMAVER
Modeling an ensemble of conformations: EOM/BILBOMD
Computation of model intensities: CRYSOL
Internal propeptide of calpain-3 is unstructured

Table 2. Crystallographic data collection and refinement statistics

| Protein sample          | C129S     | C129A     | C129-E64  | C129-Leupeptin |
|-------------------------|-----------|-----------|-----------|----------------|
| **Data collection statistics** |           |           |           |                |
| Space group             | P2₁2₁2₁   | P2₁2₁2₁   | P2₁2₁2₁   | P2₁2₁2₁        |
| Unit Cell dimensions (Å)| 59.81, 105.49, 225.26 | 60.1, 106.18, 225.62 | 60.28, 105.36, 225.36 | 54.93, 106.69, 234.63 |
| Resolution range (Å)*   | 18-2.30 (2.32) | 18-2.75 (2.77) | 29.3-2.8 (2.82) | 20-3.2 (3.22) |
| R<sub>meas</sub>%*      | 0.20 (1.37) | 0.17 (1.44) | 0.17 (0.58) | 0.65 (1.75) |
| CC 1/2*                 | 99.7 (79.5) | 99.8 (75.8) | 99.0 (91.4) | 98.3 (84.1) |
| <I/σ (I) >*            | 11.39 (2.15) | 16.42 (2.38) | 9.54 (3.29) | 4.48 (1.9) |
| Completeness (%)*       | 97.9 (96.6) | 99.7 (100) | 99.3 (99.7) | 98.8 (99.1) |
| Redundancy*             | 14.1(13.7) | 14.67 (14.9) | 4.29 (4.3) | 9.55 (9.5) |
| **Refinement statistics** |           |           |           |                |
| Resolution (Å)          | 2.3       | 2.75      | 2.8       | 3.2            |
| No. of reflections      | 59691     | 36402     | 34197     | 21642          |
| R<sub>work</sub>/R<sub>free</sub> | 0.21/0.24 | 0.20/0.24 | 0.21/0.27 | 0.27/0.34 |
| No. of atoms            |           |           |           |                |
| Protein                 | 10380     | 10266     | 10159     | 10182          |
| CA ion                  | 8         | 8         | 8         | 8              |
| CL ion                  | 4         | 4         | 4         | 4              |
| Ligand                  | -         | -         | E-64      | Leupeptin      |
| Water                   | 113       | 50        | 86        | 47             |
| B-factors (Å<sup>2</sup>) | 60       | 68.18     | 62.86     | 132            |
| Protein                 | 48.11     | 57.41     | 62.55     | 131.84         |
| Ligand/Ion              | 44.75     | 40.72     | 89.4/44.59 | 137.59/91.45  |
| Water                   | 48.22     | 37.7      | 43.37     | 62.39          |
| R.m.s. deviations#      |           |           |           |                |
| Bond length (Å)         | 0.011     | 0.011     | 0.008     | 0.01           |
| Bond angle (°)          | 1.46      | 1.34      | 1.17      | 1.53           |
| Ramachandran plot statistics (%) |        |           |           |                |
| Most favoured region    | 97.26     | 95        | 95        | 95.75          |
| Additionally allowed regions | 2.74   | 4         | 4         | 3.75           |
| PDB accession IDs       | 6BDT      | 6BGP      | 6BJD      | 6BKJ           |

* The numbers in parentheses are for the highest resolution shell.

# Root mean square deviation to ideal values.
Internal propeptide of calpain-3 is unstructured

Figure 1
Internal propeptide of calpain-3 is unstructured

Figure 2
Internal propeptide of calpain-3 is unstructured

Figure 3
Internal propeptide of calpain-3 is unstructured

Figure 4
Internal propeptide of calpain-3 is unstructured

Figure 5
Internal propeptide of calpain-3 is unstructured

Figure 6
Internal propeptide of calpain-3 is unstructured

![Diagram showing structural representations of calpain-3]

**Figure 7**
Internal propeptide of calpain-3 is unstructured
Internal propeptide of calpain-3 is unstructured

Figure 9

Figure 10
Structures of human calpain-3 protease core with and without bound inhibitor reveal mechanisms of calpain activation
Qilu Ye, Robert L. Campbell and Peter L. Davies

*J. Biol. Chem.* published online January 30, 2018

Access the most updated version of this article at doi: 10.1074/jbc.RA117.001097

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts