Solution Structure of the Calponin Homology (CH)-Domain from the Smoothelin-Like 1 Protein: a Unique Apo-Calmodulin Binding Mode and the Possible Role of the C-terminal Type 2 CH-Domain in Smooth Muscle Relaxation

Hiroaki Ishida1, Meredith A. Borman2, Janina Ostrander2, Hans J. Vogel1, and Justin A. MacDonald2,3

From Departments of Biological Sciences and Biochemistry & Molecular Biology2, University of Calgary, Calgary, Alberta, Canada

Address correspondence to: Justin A. MacDonald, Libin Cardiovascular Institute of Alberta and Department of Biochemistry & Molecular Biology, University of Calgary, Faculty of Medicine, 3330 Hospital Drive N.W., Calgary, Alberta, T2N 4N1, Canada. Tel: 1-403-210-8433; Fax: 1-403-270-2211; E-mail: jmacdo@ucalgary.ca

Running Title: Complex structure of the CH-domain from SMTNL1

The smoothelin-like 1 protein (SMTNL1) is a recently discovered component of smooth muscle tissues [Borman M. A., MacDonald J. A., and Haystead T. A. J., (2004) FEBS Lett. 573, 207-213]. This 459-residue protein contains a single type-2 CH-domain at its C-terminus that shares sequence identity with the smoothelin family of smooth muscle specific proteins. In contrast to the smoothelins, SMTNL1 does not associate with F-actin in vitro, and although it is known to be phosphorylated during cGMP-mediated Ca2+-desensitization and relaxation, its specific role in smooth muscle remains unclear. In addition, the biological function of the C-terminal CH-domains found in the smoothelin proteins is also poorly understood. In this work, we have determined the solution structure of the CH-domain of mouse SMTNL1 (SMTNL1-CH; residues 346-459). The secondary structure and the overall fold for the C-terminal type-2 CH-domain is very similar to that of other CH-domains. However, two clusters of basic residues form a unique surface structure that is characteristic of SMTNL1-CH. Moreover, the protein has an extended C-terminal α-helix, which contains a calmodulin (CaM)-binding IQ-motif, that is also a distinct feature of the smoothelins. We have characterized the binding of apo-CaM to SMTNL1-CH through its IQ-motif by isothermal titration calorimetry and NMR chemical shift perturbation studies. In addition, we have used the HADDOCK protein-protein docking approach to construct a model for the complex of apo-CaM and SMTNL1-CH. The model revealed a close interaction of SMTNL1-CH with the two Ca2+-binding loop regions of the C-domain of apo-CaM; this mode of apo-CaM binding is distinct from previously reported interactions of apo-CaM with IQ-motifs. Finally, we comment on the putative role of the CH-domain in the biological function of the SMTNL1 protein.

Calponin is a key regulator of smooth muscle contraction (reviewed in 1,2), and the calponin homology (CH)-domain4 was identified in the N-terminal portion of this protein as a ~110 amino acid region that contributed to its actin binding properties (3). CH-domains have since been identified in a number of cytoskeletal and signaling proteins (4-6). The CH-domain has a highly conserved structure that is associated with diverse biological functions. Although the various CH-domains share relatively little amino acid sequence identity, a number of strictly conserved hydrophobic residues give rise to an almost invariant hydrophobic core (Fig. 1). Thus, all of the CH-domain structures that have been determined to date are very similar. In spite of a common overall fold, different CH-domains serve to interface with a wide variety of proteins involved in cytoskeletal dynamics and/or signal transduction. Therefore, the divergence in CH-domain function is thought to result from discrete sequence elements that are exposed on the protein surface. CH-domains have been classified into several families (summarized in 5). The type-1 and type-2 CH-domains are normally arranged in tandem and are found in many actin-binding proteins, including members of the spectrin, α-actinin, dystrophin, and fimbrin protein families. Single CH-domains are found in several proteins, such as calponin and IQGAP, and are usually classified as type-3 CH-domains. The type-2 CH-domain can also exist as an isolated CH-domain, and it is found in a few proteins, including smoothelins, MICALs, and RP/EBs (6).

Calponin and other CH-domain proteins may regulate smooth muscle contractility via the thin filament regulatory system. In a previous report, Borman and colleagues (7) identified a novel ~60 kDa protein that was phosphorylated by PKG during cGMP-induced Ca2+-desensitization in ideal smooth muscle. This protein was shown to contain a single type-2 CH-domain at its C-terminus, which shared sequence similarity with the smoothelin family of smooth muscle specific proteins (reviewed by 8). The 459-residue protein, initially called calponin-homology associated smooth muscle (CHASM) protein, is termed smoothelin-like 1 protein.
(SMTNL1). However, unlike the smoothelins, it did not associate with actin filaments in vitro, and hence the specific role of SMTNL1 in smooth muscle relaxation remains undefined. Also, the interaction of smoothelin proteins with actin is mediated by additional N-terminal actin binding domains such that their CH-domain was neither necessary nor sufficient for actin binding (9). Therefore, the biological role of the C-terminal type-2 CH-domain in the smoothelin protein family remains unclear.

Close inspection of their amino acid sequences reveals that the CH-domains of the smoothelins and SMTNL1 contain a putative calmodulin (CaM)-binding IQ-motif sequence (Fig. 1). It has been reported that CaM can bind to many other CH-domain proteins, even though these generally do not possess IQ-motifs. These proteins include calponin (10, 11), spectrin (12), dystrophin (13), and filamin A (14), where CaM modulates the Ca$^{2+}$-dependency of actin-binding (13-15). Some CH-domain proteins, such as IQGAP, also contain IQ-motifs, though these are located outside of the CH-domain region of the protein (16). In this study, we provide the first insight into the structure of the C-terminal type-2 CH-domain from SMTNL1 (SMTNL1-CH) and discuss its structural characteristics with respect to previously reported CH-domains. We also demonstrate the binding of apo-CaM to SMTNL1-CH via the IQ-motif sequence and calculate a docking model for the apo-CaM/SMTNL1-CH complex using the HADDOCK protein-protein docking program (17).

**EXPERIMENTAL PROCIDURES**

**Expression and purification of SMTNL1-CH.** The full-length mouse SMTNL1 cDNA was generated from I.M.A.G.E. clone 3593616 as previously described (7). A fragment of SMTNL1 encoding the CH-domain of SMTNL1 (SMTNL1-CH; residues 346-459) was then amplified by standard PCR techniques and subcloned into the pGEX-6P1 vector (GE Healthcare) using BamHI/NotI sites. The construct was verified by DNA sequencing.

The GST-SMTNL1-CH fusion protein was produced in *Escherichia coli*, strain BL21 (DE3) in LB media. Uniformly $^{15}$N- and $^{13}$C-labeled GST-SMTNL1-CH proteins were prepared in M9 media containing 0.5 g/L $^{15}$NH$_4$Cl and 1 g/L $^{13}$C$_6$-glucose (or unlabeled glucose). The fusion proteins were isolated using glutathione-Sepharose 4B resin and cleaved ‘on-column’ by treatment with PreScission Protease (GE Healthcare). The eluted protein contained the cloning artifact “GPLGS” at its N-terminus. The SMTNL1-CH construct was concentrated and exchanged into 1 mM sodium phosphate buffer with an Amicon centrifugal filter (Millipore).

**Expression and purification of Calmodulin (CaM).** Chicken CaM was expressed from the pET30b(+) vector in *Escherichia coli* strain BL21 (DE3) grown in LB media as described previously (18). Uniformly $^{15}$N-labeled CaM was prepared in M9 media containing 0.5 g/L $^{15}$NH$_4$Cl. CaM proteolytic fragments, the N-terminal domain of CaM (CaM-n: residues 1-77) and the C-terminal domain of CaM (CaM-ct: residues 78-148) were produced as previously described (19).

**NMR sample preparation.** All NMR samples contained 0.5-1.0 mM $^{15}$N- or $^{13}$C, $^{15}$N-labeled SMTNL1-CH, 1 mM sodium phosphate buffer (pH 7.0), 10 mM $^2$H$_2$-labeled dithiothreitol (DTT), 0.03% NaN$_3$ and 0.5 mM 2,2-dimethyl-2-silapentane-5-sulfate (DSS) in 90% H$_2$O/10% D$_2$O. Unlabeled SMTNL1-CH sample was also prepared in 99.99% D$_2$O. The samples used for residual dipolar coupling (RDC) measurements also contained 300 mM KCl, 20 mM Bis-Tris (pH 6.9) and 10 mg/ml filamentous phage P22 (Asla Labs). All samples to monitor the interaction between SMTNL1-CH and apo-CaM also contain an additional 1 mM EDTA.

**Isothermal titration calorimetry (ITC) measurements.** All ITC experiments were performed on a MicroCal ITC microcalorimeter. Solutions of ~0.5 mM CaM, CaM-n or CaM-ct with 20 mM HEPES (pH 7.0) and 1 mM EDTA or 5 mM CaCl$_2$ were sequentially injected into a sample cell containing 20 µM SMTNL1-CH in the same buffer. All buffers contained 1 mM 2-mercaptoethanol to prevent inter-molecular disulfide bonding in SMTNL1-CH. The concentration of each protein was determined using their predicted molar extinction coefficients (cm$^{-1}$·M$^{-1}$): CaM, $\varepsilon_{280} = 2560$, CaM-ct, $\varepsilon_{280} = 2980$, CaM-n, $\varepsilon_{250} = 742$, SMTNL1-CH, $\varepsilon_{280} = 18450$. All titrations were performed at 30 °C, and the data were fit to a one-site binding model (MicroCal Origin software) to obtain dissociation constants (Kd).

**NMR measurements.** All NMR experiments were carried out at 20 °C on Bruker Avance 500 or 700 MHz NMR spectrometers equipped with triple resonance inverse Cryoprobes with a single z-axis gradient. Sequential assignments of HN, N, CO, C$\alpha$ and C$\beta$ resonances of SMTNL1-CH were achieved using two dimensional (2D) $^{15}$N, $^2$H-HSQC and a series of three dimensional (3D) experiments including HNCA,CB, CBCA(CO)NH, HNCO and HN(CO)CA. Side-chain assignments were performed using 3D-C(CCO)NH-TOCSY, H(CCO)NH-TOCSY and HBHA(CBCACO)NH experiments. 2D-
(HB)CB(CGCD)HD and (HB)CB(CGCDCE)HE experiments were also employed for the aromatic resonance assignments (20). All NOESY experiments including the 3D-15N-NOESY-HSQC, 3D-13C-NOESY-HSQC and 2D-NOESY were measured with a mixing time of 100 ms. 1H-15N heteronuclear NOE experiments were acquired on a 700 MHz spectrometer with a recycle delay of 5 sec (21). The experiments were repeated three times and averaged. H-N RDC measurements were performed using the in-phase/anti-phase (IPAP)-1H, 15N-HSQC experiment (22). The chemical shift perturbation (CSP) studies were performed by monitoring the 1H, 15N-HSQC spectra of 15N-labeled SMTNL1-CH and 15N-labeled apo-CaM by adding unlabeled apo-CaM and unlabeled SMTNL1-CH, respectively. The CSP value was then evaluated as a weighted average chemical shift difference of 1H and 15N resonances, using the following equation.

\[ \text{CSP} = \sqrt{(\Delta HN)^2 + (\Delta N/5)^2} \]  

(23)

Chemical shifts in all spectra were referenced using DSS to obtain 1H, 15N and 13C chemical shifts (24). All spectra were processed using NMRPipe (25) and analyzed using the NMRView software (26).

Structure calculations. The initial SMTNL1-CH structure was calculated with CYANA v2.0 (27) using distance restraints obtained from the automatic NOE assignment protocol. Dihedral angle restraints (\( \phi, \psi \)) were predicted with TALOS (28), and hydrogen bond restraints were based on secondary structure from a chemical shift index (CSI) for the C\( \alpha \) and C\' atoms. Further structural refinement with the addition of RDC restraints were performed by XPLOR-NIH (29). Initial estimates for the axial component of the molecular alignment tensor (Da) and the rhombicity (R) were obtained on the lowest energy structure calculated by CYANA using PALES (30). Finally, the 30 lowest energy structures from a total of 200 were selected and analyzed.

Docking model for the SMTNL1-CH-apo-CaM complex. Based on the CSP data, a docking model for the SMTNL1-CH-apo-CaM complex was calculated with the HADDOCK2.0 program in conjunction with CNS (31). The HADDOCK (High Ambiguity Driven biomolecular DOCKing) program is originally designed for the generation of protein-protein complex structures based on available experimental data (17). According to the outcome of the CSP and ITC data, only the C-terminal domain of apo-CaM was employed in this calculation. The 30 structure ensemble for SMTNL1-CH (determined in this study) and the 10 structure ensemble for the C-domain of apo-CaM (residues 82-148; PDB code: 1F71; 32) were used as the starting structures. The average solvent accessibilities per residues in the ensemble structures were calculated by NACCESS (http://www.bioinf.manchester.ac.uk/naccess/). The high solvent accessible residues (>40%) that exhibited a CSP above 0.04 and 0.025 for SMTNL1-CH and the C-domain of apo-CaM, respectively, were designated as active residues (see below). The residues neighboring those active residues with a high solvent accessibility (>40%) were designated as passive residues. The active residues in SMTNL1-CH were G4, K6, N7, D93, K95, T99, E103, R106 and G112. The associated passive residues were S5, S94, C96, Q102, V109, Q110 and K111. For the C-domain of apo-CaM, the active residues were K94, D95, G96, Y99, D131, G134 and Q135, and the passive residues were N97, S101, G132, D133 and N137. The positions of the active residues for both proteins are indicated on their structure (see below). In the first stage of the calculation, an initial set of 1000 rigid-body docking models was generated. The 200 lowest energy complex models were then selected and submitted for a second stage of calculations with semi-flexible simulated annealing. The 50 lowest energy models in the second stage were refined in water solvent, and clustered using a 3.0 Å r.m.s.d. cut-off criterion. Finally, the 15 lowest energy complex models were selected from the most populated cluster with the lowest HADDOCK score, and used for the analysis. All molecular graphics were created with MOLMOL (33).

RESULTS

Structure determination of SMTNL1-CH- Using the HNCACB, HN(CO)CACB, HN(CA)CO, and HNCO experiments, all amide-resonances except for Gly1, Pro47, Pro52, and Pro92 were assigned in the 1H, 15N HSQC spectrum (Fig. 2). The side-chain resonance assignments were mainly obtained from the C(CO)NH-TOCSY, H(CO)NH-TOCSY and HBHA(CBCACO)NH spectra. Since SMTNL1-CH contains a total of 16 aromatic residues, the (HB)CB(CGCD)HD and (HB)CB(CGCDCE)HE experiments were also carried out to obtain unambiguous resonance assignments of the aromatic side-chains. Consequently, 94.6% of the total 1H resonances were assigned and used in the structure calculation. The initial structures were calculated using the torsion angle dynamics and the automatic NOE assignment protocol with CYANA. At the first stage, a total of 3521 NOE signals were manually identified on the 3D-15N- and 13C-NOESY-HSQC spectra as well as the 2D-NOESY spectrum acquired in D2O. 3261 NOE
signals were automatically assigned by CYANA and this generated 2136 distance restraints. The distribution of the distance restraints was determined as a function of residue number (Fig. 3a). The backbone r.m.s.d. for the well-folded region (residues 5-111) of the calculated 30 structures together with the hydrogen bond and TALOS dihedral restraints was 0.29 ± 0.05 Å (data not shown). To determine accurate inter-helical angles/positions, we employed H-N RDC restraints during the second stage of the structure calculation with the program XPLOR-NIH. The RDC values were available for 87.4% of the residues and ranged between -27.1 and 17.0 Hz with a digital resolution of 1.24 Hz. All restraints used for the final stage of the structure calculation are summarized in Table 1. The average number of restraints per residue was 18.7. The superimposition of the final 30 structures and the ribbon representation of the lowest energy structure of SMTNL1-CH are shown in Fig. 4a and b, respectively. The r.m.s.d. values for the backbone and the entire heavy atoms in the well-folded region (residues 5-111) were 0.27 ± 0.03 and 0.80 ± 0.04 Å, respectively. The experimentally determined RDC values correlate well with the best-fit RDC values calculated from the final structures of SMTNL1-CH with correlation (R) and quality (Q) factors of 1.00 and 0.015, respectively. The structures were validated with the PROCHECK program (34), and 84% of the residues were found in favored regions of the Ramachandran plot; the remaining residues were all found in the additionally allowed regions (Table 1). The SMTNL1-CH is a globular molecule containing five α-helices and one 3₁₀-helix with no β-sheet structure (Fig. 4b). The helices are located from residues 5-15, 30-32 (3₁₀-helix), 36-42, 60-75, 83-89, and 94-111, which agrees with the CSI values obtained from the Ca and C’ atoms (Fig. 3b). Several irregular 3₁₀-helices were also identified for residues 47-49, 52-54 and 57-59.  

**Structural comparison of SMTNL1-CH with other CH-domains**—A ribbon diagram of the lowest energy structure of SMTNL1-CH was superimposed with the CH-domain structures of smoothelin<sup>5</sup> (Fig. 5a, PDB code: 2D87), α-actinin 1 (Fig. 5b, PDB code: 2EYI) (35), spectrin (Fig. 5c, PDB code: 1BKR) (36), and MICAL-1 (Fig. 5d, PDB code: 2DK9) (37). The superimpositions were performed using the well-structured regions of SMTNL1-CH and the corresponding regions of other CH-domain structures that were defined by sequence alignments (Fig. 1). The backbone r.m.s.d. was 1.15, 3.47, 3.72, and 2.88 Å with smoothelin, α-actinin 1, spectrin, and MICAL-1, respectively.

**Characterization of the Apo-CaM interaction with SMTNL1-CH**—The binding of apo-CaM with SMTNL1-CH was first characterized by ITC experiments. apo-CaM exhibited an exothermic interaction with SMTNL1-CH (K<sub>d</sub>; 2.7 ± 10<sup>-6</sup> M) in the presence of EDTA, whereas no heat of binding was detected in the presence of Ca<sup>2+</sup> (Fig. 6a). We also performed ITC experiments with the proteolytic C- and N-terminal fragments of apo-CaM (CaM-nt and CaM-ct, respectively). CaM-ct exhibited an exothermic interaction with a similar K<sub>d</sub> (4.0 × 10<sup>-6</sup> M) to that observed with intact CaM. However, CaM-nt did not bind to SMTNL1-CH (Fig. 6b).

CSP values, determined from the <sup>1</sup>H, <sup>15</sup>N-HSQC spectrum of SMTNL1-CH following addition of unlabeled CaM, were plotted as a function of residue number (Fig. 7a). Amides that experience CSP greater than 0.04 could be mapped to residues located either on the IQ-motif sequence of SMTNL1-CH or to regions that were in close proximity (Fig. 7c). Similarly, the CSPs were monitored in the <sup>1</sup>H, <sup>15</sup>N-HSQC spectrum of apo-CaM by adding unlabeled SMTNL1-CH into the NMR sample (Fig. 7b). Residues with CSP values greater than 0.025 were only found in the C-domain of CaM and were located around its two Ca<sup>2+</sup>-binding loops (Fig. 7d). The backbone structure of 15 refined complex models and a ribbon representation of the lowest energy model for the interaction of SMTNL1-CH with the C-domain of apo-CaM was generated (Fig. 8). The average backbone r.m.s.d. for the 15 complex models was 0.95 ± 0.33 Å. In Fig. 8b, the side-chains which form inter-molecular hydrogen bonds at the interface of the docking model are also indicated. Those side chains were defined as belonging to K6, K95, E103, R106, K115 on SMTNL1-CH that form a hydrogen bond to the side chains of D95, N97, Y99, D133, D131 on apo-CaM, respectively.

**DISCUSSION**

We have undertaken a structural analysis of the C-terminal type-2 CH-domain of the smoothelin-like 1 protein to assist in our understanding of the biological function of this novel member of the smoothelin family of smooth muscle specific proteins. Herein, we illustrate that the CH-domain of SMTNL1 maintains a helix-rich, globular structure that is typical of that found for the CH-domains of many other proteins (Fig. 4b). A number of conserved hydrophobic residues, including the two Trp residues in SMTNL1-CH, contribute to a hydrophobic core that is conserved in the structure of other CH-domains (Fig. 1 and 4b). The
average backbone $\{^{1}H\}$-$^{15}$N NOE value in the well folded region (residues 5-111) is >0.8 without any considerable flexible regions, suggesting a very rigid protein conformation (Fig. 3a). This observation is similar to the reported $\{^{1}H\}$-$^{15}$N NOE data for the type-3 CH-domain of calponin (38). The SMTNL1-CH domain shares high sequence identity with the CH-domain of smoothelin, and not surprisingly, the backbone r.m.s.d. is 1.15 Å when the SMTNL1-CH structure was superimposed on the recently deposited CH-domain structure of smoothelin (Fig. 5a). In spite of the relatively low level of sequence identity (37.6%, 34.2% and 22.2% to the CH-domains of actinin-1, spectrin and MICAL-1, respectively), the 3D-structures were very similar to one another (Fig. 5b-e). The most remarkable difference in the structure of SMTNL1-CH as compared to the other CH-domains is the extended C-terminal α-helix (helix VI) (Fig. 5).

The unique basic cluster, “KTKKK” at the tail of helix VI bestows a highly basic surface to the SMTNL1-CH molecule (Fig. 1 and 4c). Another basic surface is formed by residues 59-61 and 88 (Fig. 1 and 4d) and is also unique to the CH-domains of the smoothelin family proteins. Unlike the majority of the CH-domains, which are generally located at the N-terminus of proteins, helix VI of SMTNL1-CH is located at the C-terminus of the SMTNL1 protein and is exposed to the solvent. The recently reported solution structure of the MICAL-1 type-2 CH-domain (37) also demonstrated a fairly long helix at its C-terminus. However, this CH-domain is located in the middle of the MICAL-1 protein, and the basic “KTKKK” cluster is absent (39).

A most interesting new finding from this work is the presence of a CaM binding IQ-motif on helix VI of SMTNL1-CH that is also conserved in the other smoothelin family members (Fig. 1b). The typical IQ-motif, with consensus sequence IQxxxxRGxxxR, was first characterized in the heavy chain of many myosin motor proteins as multiple tandem repeats (40) targeting either CaM or CaM-related molecules (41, 42). The IQ-motif sequence with minor substitutions has also been identified in many CaM binding proteins such as unconventional myosins, PEP-19, and IQGAP-like proteins (42-44). We tested whether CaM was capable of binding to SMTNL1-CH and the Kd value obtained by isothermal titration calorimetry was $2.7 \times 10^{-6}$ M for apo-CaM (Fig. 6a). This is in line with many previously examined apo-CaM target protein interactions. For example, the reported Kd values are $10^{-5}$ M for neurogranin (45) and PEP-19 (46), and $10^{-7}$ M for neuromodulin (47) and IQGAP1 (48), whereas Kd’s vary from $10^{-5}$ to $10^{-8}$ M for the IQ motifs in myosin V (IQ1-6) (49). These apo-CaM complexes are distinct from the typical Ca$^{2+}$-CaM complexes that represent a tight binding (Kd ~$10^{-8}$-$10^{-9}$ M) that lead to enzyme activation. The formation of apo-CaM complexes prevents CaM from diffusing in the cytoplasm. The localization of CaM at specific sites in the resting cell is believed to facilitate a direct response to an influx of Ca$^{2+}$ (50, 51). The $^{1}$H, $^{15}$N HSQC titration experiments showed the formation of a 1:1 complex with slow exchange on the NMR time scale (data not shown). The stoichiometry derived from ITC experiments was lower than one (Fig. 6a), and this discrepancy was probably caused by the difficulty of determining the binding stoichiometry of the relatively weak interaction by ITC experiments under the conditions of low protein concentrations (~20 µM). While most IQ-motifs are known to bind to apo-CaM, some proteins and peptides containing IQ-motifs are also capable of binding Ca$^{2+}$-CaM (42). Our ITC experiments suggested either extremely weak or no binding of Ca$^{2+}$-CaM with SMTNL1-CH (Fig. 6a).

Consistent with the outcome of these calorimetry studies, in $^{1}$H, $^{15}$N HSQC titration experiments with Ca$^{2+}$-CaM, we only observed a few signals that belong to the N-terminus of SMTNL1-CH that showed minor perturbations in the fast exchange regime on the NMR time scale, which might be caused by slight changes in the NMR sample conditions (data not shown). We therefore conclude that only apo-CaM binds efficiently to SMTNL1-CH. This behavior closely resembles that of apo-CaM binding to neuromodulin, neurogranin, and IQGAP1, where Ca$^{2+}$-CaM poorly interacts with these proteins (42, 43, 48). Hence, in the remainder of our discussion, we will focus on the novel interaction identified between apo-CaM and SMTNL1-CH.

The CSP study and the $^{1}$H, $^{15}$N-HSQC experiments clearly demonstrated that the interface of apo-CaM binding was located around the IQ-motif sequence of SMTNL1-CH (Fig. 7c). Previous models of the crystal structure of myosin V complexed with the CaM-like protein, Mlc1p, together with FRET distance measurements, indicated that CaM bound to IQ4 and IQ6 using only its C-domain while the N-domain remained free in solution (52). In SMTNL1-CH, the CaM-ct showed a similar binding affinity to that observed with intact CaM, whereas the CaM-nt did not show binding (Fig. 6b). The amount of heat (ΔH) generated by intact apo-CaM binding to SMTNL1-CH was however almost eight times higher than that by CaM-ct (Fig. 6b). On the other hand, the clear CSPs observed in $^{1}$H, $^{15}$N-HSQC spectrum of apo-CaM that were induced by binding to SMTNL1-CH only locate to the C-domain of apo-CaM (Fig 7b and d), which is
consistent with the ITC-data, and only small CSPs are observed in the N-domain. These results suggest that apo-CaM binds to SMTNL1-CH mostly through its C-domain, whereas the N-domain possibly contributes minor interactions. As the magnitude of the CSP observed in both spectra is relatively small, there is no significant alteration of the structures of both proteins that is induced by the binding. Based on these data, we can therefore construct a docking model for the SMTNL1-CH complex with the C-domain of apo-CaM using the HADDOCK program. The interaction appears to be electrostatic where several acidic residues (including D95, D131, and D133) of the Ca\(^{2+}\)-binding loops of the C-domain of apo-CaM form intermolecular hydrogen bonds to basic residues (including K6, K95, R106, and K115) of SMTNL1-CH. In this calculation, the potential active residues, K117 and K118 that are located on the disordered C-terminal region of SMTNL1-CH structure were not employed. However, considering their relatively large CSP values, as well as their close proximity to the CaM protein, K117 and/or K118 are also likely to form additional electrostatic interactions to CaM. The recent crystal structure of apo-CaM complexed with IQ-motifs from myosin V suggested that the C-domain of CaM adopted a semi-open conformation to grab the first half of the IQ-motif through a hydrophobic interaction (53). In this complex, the Ile residue at the first position and a hydrophobic residue at the fifth position (Ile or Val) of the IQ-motif sequence make important hydrophobic contacts to the semi-open hydrophobic patch of the C-domain of apo-CaM (Fig. 1). In SMTNL1-CH, the first Ile residue, I101, is however buried in the protein structure and is therefore not available to contribute such a hydrophobic interaction. In addition, the hydrophobic residue at position five of the myosin V IQ-motif (Ile or Val) is replaced with Tyr in SMTNL1-CH (Fig. 1). Many side chains of the C-domain of apo-CaM which form major electrostatic contacts to SMTNL1-CH including residues D95, N97, D131 and D133 also serve as the Ca\(^{2+}\)-binding ligands, suggesting that this interaction cannot be made once CaM binds Ca\(^{2+}\) ions. This is in agreement with our ITC and NMR results in which only apo-CaM seems capable of binding to SMTNL1-CH (e.g. Fig. 6a).

The addition of SMTNL1 can induce relaxation of permeabilized ileal smooth muscle strips (7), and our recent measurements of muscle force using SMTNL1 fragments suggest that the removal of SMTNL1-CH is critical for the relaxant potential of the protein (J.A. MacDonald; unpublished observations). Sufficient CaM is retained after skinning (demembranation) of rat caudal arterial smooth muscle in the presence of EGTA to support Ca\(^{2+}\)-evoked contraction (55), and the addition of exogenous SMTNL1-CH in the absence of Ca\(^{2+}\) (presence of 5 mM EGTA) could release CaM from the Triton-skinned tissue (J.A. MacDonald, unpublished observations). On the other hand, the mechanism whereby SMTNL1 can induce muscle relaxation is not understood. Although at this moment, the role of the SMTNL1-CaM interaction in muscle has not yet been defined, we speculate that the association of apo-CaM with SMTNL1 has physiological relevance. There are several mechanisms whereby a functional SMTNL1-apo-CaM interaction could modulate the contractility of smooth muscle. 1) The co-localization of CaM with the CH-domain of SMTNL1 could be required for the formation of ternary complexes with other target proteins of SMTNL1. 2) CaM binding might influence the phosphorylation efficiency of SMTNL1, or alternatively, phosphorylation may influence CaM binding efficiency under conditions of Ca\(^{2+}\)-desensitization. The PKG phosphorylation site (7) is located outside of the CH-domain, but in close proximity to the CaM binding site. 3) The C-terminal basic cluster of SMTNL1 that also exists in smoothelin could interact with other acidic molecules and because of the overlap with the CaM-binding site, CaM might compete with those targets. 4) As SMTNL1 interacts with common acidic residues on CaM for Ca\(^{2+}\)-binding (54), it could modulate the Ca\(^{2+}\)-association and/or dissociation of CaM, as has also been discussed in the case of PEP-19 (56, 57). 5) Finally, as we mentioned, binding of apo-CaM to proteins, followed by dissociation of the Ca\(^{2+}\)-CaM, has often been described as a method for intracellular localization of CaM (50). Such localization would influence the kinetics for Ca\(^{2+}\)-CaM regulated processes in vivo.

As CaM binding is one of the miscellaneous functions of the CH-domain of calponin, SMTNL1-CH might also be a multifunctional domain. As we have described in this manuscript, the C-terminal type-2 CH-domain has distinct structural properties from the other types of CH-domains. Therefore, identification of additional proteins that interact with the CH-domains of smoothelin and SMTNL1 in smooth muscle cells are necessary for a more complete understanding of the function of the C-terminal type-2 CH-domain.

REFERENCES
1. Winder S. J. and Walsh M. P. (1993) Cell Signal. 5, 677-686
2. Szymanski P. T. (2004) J. Muscle Res. Cell Motil. 25, 7-19
3. Castresana J. and Saraste M. (1995) FEBS Lett. 374, 149-151
4. Stradal T., Kranwitter W., Winder S. J., and Gimona M. (1998) FEBS Lett. 431, 134-137
5. Gimona M., Djinovic-Carugo K., Kranwitter W. J., and Winder S. J. (2002) FEBS Lett. 513, 98-106
6. Korenbaum E. and Rivero F. (2002) J. Cell Sci. 115, 3543-3545
7. Borman M. A., MacDonald J. A., and Haystead T. A. (2004) FEBS Lett. 573, 207-213
8. van Eys G. J., Niessen P. M., and Rensen S. S. (2007) Trends Cardiovasc. Med. 17, 26-30
9. Quences C., Kramer J., Cardoso M. C., and Leonhardt H. (2002) J. Cell Biochem. 85, 403-409
10. Takahashi K., Hiwada K., and Kokubu T. (1986) Biochem. Biophys. Res. Commun. 141, 20-26
11. Winder S. J., Walsh M. P., Vasulcak A., and Johnson J. D. (1993) Biochemistry. 32, 13327-13333
12. Sobel J. S., Goldstein E. G., Venuti J. M., and Welsh M. J. (1988) Dev. Biol. 126, 47-56
13. Jarrett H. W. and Foster J. L. (1995) J. Biol. Chem. 270, 5578-5586
14. Nakamura F., Hartwig J. H., Stossel T. P., and Szymanski P. T. (2005) J. Biol. Chem. 280, 32426-32434
15. Kołakowski J., Makuch R., Stepkowski D., and Dabrowska R. (1995) Biochem. J. 306 (Pt 1), 199-204
16. Briggs M. W. and Sacks D. B. (2003) EMBO Rep. 4, 571-574
17. Dominguez C., Boelens R., and Bonvin A. M. (2003) J. Am. Chem. Soc. 125, 1731-1737
18. Nakashima K., Maekawa H., and Yasawa M. (1996) Biochemistry 35, 5602-5610
19. Itakura, M. and Illo, T. (1992) J. Biochem. (Tokyo) 112, 183-191
20. Yamazaki T., Forman-Kay J. D. and Kay L. E. (1993) J. Am. Chem. Soc. 115, 11054-11055
21. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoeelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Biochemistry 33, 5984-6003
22. Ottiger M., Delaglio F., and Bax A. (1998) J. Magn. Reson. 131, 373-378
23. Grzesiek S., Bax A., Clore G. M., Gronenborn A. M., Hu J. S., Kaufman J., Palmer I., Stahl S. J., and Wingfield P. T. (1996) Nat. Struct. Biol. 3, 340-345
24. Wishart D. S., Bigam C. G., Yao J., Abildgaard F., Dyson H. J., Oldfield E., Markley J. L., and Sykes B. D. (1995) J. Biomol. NMR 6, 135-140
25. Delaglio F., Grzesiek S., Vuister G. W., Zhu G., Pfeifer J., and Bax A. (1995) J. Biomol. NMR 6, 277-293
26. Johnson B. A. and Blevins R. A. (1994) J. Biomol. NMR. 4, 603-614
27. Jee J., and Guntert P. (2003) J. Struct. Funct. Genomics. 4, 179-189
28. Cornilescu G., Delaglio F., and Bax A. (1999) J. Biomol. NMR. 13, 289-302
29. Schwieters C. D., Kuszewski J. J., Tjandra N., and Clore G. M. (2003) J. Magn. Reson. 160, 65-73
30. Zweckstetter M. and Bax A. (2000) J. Am. Chem. Soc. 122, 3791-3792
31. Brünger A. T., Adams P. D., Clore G. M., DeLano W. L., Gros P., Grosse-Kunstleve R. W., Jiang J. S., Kuszewski J., Nilges M., Pannu N. S., Read R. J., Rice L. M., Simonson T., and Warren G. L. (1998) Acta. Crystallogr. D Biol. Crystallogr. 54, 905-921
32. Chou J. J., Li S. P. and Bax A. (2000) J. Biomol. NMR. 18, 217-227
33. Koradi R., Billeter M., and Wuthrich K. (1996) J. Mol. Graph. 14, 29-32
34. Morris A. L., MacArthur M. W., Hutchinson E. G., and Thornton J. M. (1992) Proteins 12, 345-364
35. Borrego-Díaz E., Kerff F., Lee S. H., Ferron F., Li Y., and Dominguez R. (2006) J. Struct. Biol. 155, 230-238
36. Banuelos S., Saraste M., and Carugo K. D. (1998) Structure 6, 1419-1431
37. Sun H., Dai H., Zhang J., Jin X., Xiong S., Xu J., Wu J., and Shi Y. (2006) J. Biomol. NMR. 36, 295-300
38. Bramham J., Hodgkinson J. L., Smith B. O., Uhrin D., Barlow P. N., and Winder S. J. (2002) Structure 10, 249-258
39. Suzuki T., Nakamoto T., Ogawa S., Seo S., Matsumura T., Tachibana K., Morimoto C., and Hirai H. (2002) J. Biol. Chem. 277, 14933-14941
40. Cheney R. E. and Moosiker M. S. (1992) Curr. Opin. Cell Biol. 4, 27-35
41. Espindola F. S., Suter D. M., Partata L. B., Cao T., Wolenski J. S., Cheney R. E., King S. M., and Moosiker M. S. (2000) Cell Motil. Cytoskeleton. 47, 269-281
42. Bahler M. and Rhoads A. (2002) FEBS Lett. 513, 107-113
43. Jurado L. A., Chockalingam P. S., and Jarrett H. W. (1999) Physiol. Rev. 79, 661-682
44. Yamniuk A. P. and Vogel H. J. (2004) Mol. Biotechnol. 27, 33-57
45. Cui Y., Wen J., Hung Sze K., Man D., Lin D., Liu M., and Zhu G. (2003) Anal. Biochem. 315, 175-182
46. Dickerson J. B., Morgan M. A., Mishra A., Slaughter C. A., Morgan J. I., and Zheng J. (2006) Brain Res. 1092, 16-27
47. Alexander K. A., Wakim B. T., Doyle G. S., Walsh K. A., and Storm D. R. (1988) J. Biol. Chem. 263, 7544-7549
48. Li Q. and Stuenkel E. L. (2004) Biochem. Biophys. Res. Commun. 317, 787-795
49. Martin S. R. and Bayley P. M. (2004) FEBS Lett. 567, 166-170
50. Alexander K. A., Cimler B. M., Meier K. E., and Storm D. R. (1987) J. Biol. Chem. 262, 6108-6113
51. Tsvetkov P. O., Protasevich I. I., Gilli R., Lafitte D., Lobachov V. M., Haiech J., Briand C., and Makarov A. A. (1999) J. Biol. Chem. 274, 18161-18164
52. Terrak M., Rebowski G., Lu R. C., Grabarek Z., and Dominguez R. (2005) Proc. Natl. Acad. Sci. U. S. A. 102, 12718-12723
53. Houdusse A., Gaucher J. F., Krementsova E., Mui S., Trybus K. M., and Cohen C. (2006) Proc. Natl. Acad. Sci. U. S. A. 103, 19326-19331
54. Gifford J. L., Walsh M. P. and Vogel H. J. (2007) Biochem. J. 405, 199-221
55. Wilson D.P., Sutherland C., and Walsh M.P. (2002) J. Biol. Chem. 277, 2186-92
56. Putkey J. A., Kleerekoper Q., Gaertner T. R., and Waxham M. N. (2003) J. Biol. Chem. 278, 49667-49670
57. Putkey J. A., Waxham M. N., Gaertner T. R., Brewer K. J., Goldsmith M., Kubota Y., and Kleerekoper Q. K. (2008) J. Biol. Chem. 283, 1401-1410

FOOTNOTES

*This work was supported by research grants from the Heart and Stroke Foundation of Canada (HSFC) to J.A.M. and the Canadian Institutes of Health Research to H.J.V. H.I. and M.A.B. were the recipients of an Alberta Heritage Foundation for Medical Research (AHFMR) Fellowship and an HSFC Fellowship, respectively. J.A.M is the recipient of a Canada Research Chair (Tier II) in Smooth Muscle Pathophysiology and a Scholarship from the HSFC. H.J.V. holds a Scientist award from the AHFMR.

The Protein Data Bank entries for SMTNL1-CH and its complex with the C-terminal domain of apo-CaM are 2JV9 and XXX, respectively.

To whom correspondence should be addressed: Smooth Muscle Research Group and Department of Biochemistry & Molecular Biology, University of Calgary, Faculty of Medicine, 3330 Hospital Drive N.W., Calgary, Alberta, T2N 4N1, Canada. Tel: 1-403-210-8433; Fax: 1-403-270-2211; E-mail: jmacdo@ucalgary.ca

The abbreviations used are: CH-domain, calponin homology-domain; CaM, calmodulin; CaM-nt, the N-terminal domain of CaM; CaM-ct, the C-terminal domain of CaM; SMTNL1, smoothelin-like 1; SMTNL1-CH, CH-domain of smoothelin-like 1; PKG, cGMP-dependent protein kinase; HSQC, heteronuclear single quantum correlation; DSS, 2,2-dimethyl-2-silapentane-5-sulfate; RDC, residual dipolar coupling; NOE; nuclear Overhauser effect; ITC, isothermal titration calorimetry; CSP, chemical shift perturbation.

Tomizawa et al., This structure was determined as part of a structural genomics project. Only the atomic coordinates are currently available in the Protein Data Bank for this structure and no further information was provided.

FIGURE LEGENDS

Fig. 1. a, Sequence alignment of type-2 CH-domains. The invariant residues among all CH-domains used for the alignment are colored in magenta. The conserved hydrophobic residues are colored in green. A schematic drawing of the secondary structure elements of SMTNL1-CH is also included. The residues which form the unique basic surface on SMTNL1-CH are boxed. UniprotKB accession numbers are: human α-actinin 1, P12814; human α-actinin 3, Q08043; human spectrin, Q01082; human plectin, Q15149; human utrophin, Q5SZ57; human dystrophin, P11532; human MICAL-1, Q8TDZ2; mouse smoothelin isoform S/L, Q9JLU7/Q9R0D0; mouse SML-1, Q99LM3. b, Sequence alignment of the IQ motifs or IQ motif-like sequence from various CaM-binding proteins. The IQ motif is based on the consensus sequence IQxxxRGxxxR. The invariant residues among the IQ motifs displayed are colored in magenta. Other consensus residues are colored in green. UniprotKB accession numbers are: human
neuromodulin, P17677; human Cav1.2, Q13936; mouse myosin V, Q99104; Saccharomyces cerevisiae Myo2p, P19524.

Fig. 2. The assigned $^1$H, $^15$N-HSQC spectrum of SMTNL1-CH. The spectrum was acquired at 20 °C, pH 7.0. The folded signals are indicated by an asterisk.

Fig. 3. a, Structural data for SMTNL1-CH. The number of distance restraints for the structure calculation (top), r.m.s.d’s of the 30 calculated structures (middle), and the backbone $^{1}$H, $^{15}$N-NOE (bottom) are shown as a function of residue number. In the top panel, white, gray, dark gray, and black bars indicate the number of intra-residue, sequential, medium range (2-4), and long range (5+) NOEs, respectively. In the middle panel, the r.m.s.d’s for the backbone and all the heavy atoms are shown with a solid and dashed line, respectively. b, CSI values for Cα and C’ atoms. The secondary structures obtained from the calculated structures are also shown. The boxes with residue numbers indicate the position of the helices. Three irregular $3_{10}$-helices are also found in the positions indicated by the boxes with a dashed line.

Fig. 4. a, The backbones of 30 structures of SMTNL1-CH are superimposed for the well-folded region (residues 5-111). The helix numbers are labeled as in the text. b, The ribbon representation of the lowest energy structure. The conserved hydrophobic residues among the CH-domains are also shown. c and d, The surface electrostatic properties of SMTNL1-CH are depicted from different angles.

Fig. 5. Structural comparison of SMTNL1-CH with other type-2 CH-domains. The lowest energy structure of SMTNL1-CH is superimposed on the CH-domain structure from a, smoothelin, b, α-actinin 1, c, spectrin, and d, MICAL-1. SMTNL1-CH is colored navy in all panels.

Fig. 6. ITC characterization of the CaM-SMTNL1-CH interaction. a, CaM is titrated into SMTNL1-CH in the presence of EDTA or Ca$^{2+}$. The bottom panel shows the derived binding isotherms of the experiments with EDTA (■) and Ca$^{2+}$ (●). The heat of dilution control experiment with EDTA is shown in the inset. b, SMTNL1-CH was titrated with the half-fragments of CaM, CaM-nt and CaM-ct. The bottom panel includes the derived binding isotherms of SMTNL1-CH interactions with CaM-nt (▲) and CaM-ct (●). All experiments were carried out at 30 °C.

Fig. 7. Identification of the interfaces for SMTNL1-CH-apo-CaM interaction. a, CSP induced by apo-CaM binding to SMTNL1-CH is plotted as a function of the residue number. b, CSP induced by SMTNL1-CH binding to apo-CaM is plotted as a function of the residue number. The locations of the N- and C-domains of CaM are indicated. c, Amide proton atoms of residues with CSP > 0.04 are mapped as spheres on the SMTNL1-CH structure. The region which contains the IQ-motif like sequence is circled. d, Amide proton atoms with CSP > 0.025 are mapped as spheres on the structure of the C-terminal domain of apo-CaM. The spheres are colored in the same manner as in panel c and d. The residues that are employed as active residues in the HADDOCK calculation are shown in blue while the others are shown in green.

Fig. 8. HADDOCK-derived structure of SMTNL1-CH complexed with the C-domain of apo-CaM. a, Superposition of 15 models of the complex of SMTNL1-CH (green) with the C-domain of apo-CaM (blue). b, Ribbon representation of the lowest energy model. The side-chains that contribute to the formation of intermolecular hydrogen bonds are also shown. The acidic and basic side-chains are colored in pink and blue, respectively, while a Tyr is shown in green.
| Number of experimental restraints |  |
|----------------------------------|--|
| Distance restraints from NOEs     | 1821 |
| Intra-residue                    | 444  |
| Sequential                       | 475  |
| Medium range                     | 431  |
| Long range                       | 471  |
| Hydrogen bond distance restraints| 74   |
| Dihedral angle restraints (TALOS) | 228  |
| H-N RDC restraints               | 104  |

| Average r.m.s.d.s from experimental restraints |  |
|-----------------------------------------------|--|
| Distance restraint violation (Å)              | 0.044 ± 0.002 |
| Dihedral angle restraint violation (deg)      | 0.122 ± 0.109 |
| RDC restraint violation (Hz)                  | 0.111 ± 0.007 |

| Average r.m.s.d.s from idealized covalent geometry |  |
|---------------------------------------------------|--|
| Bonds (Å)                                         | 0.004 ± 0.000 |
| Angles (deg)                                      | 0.632 ± 0.007 |
| Impropers (deg)                                   | 0.503 ± 0.010 |

| PROCHECK Ramachandran analysis                   |  |
|--------------------------------------------------|--|
| Residues in favored regions (%)                  | 86.4 |
| Residues in additional allowed regions (%)       | 13.6 |
| Residues in generously allowed regions (%)       | 0.0  |
| Residues in disallowed regions (%)               | 0.0  |

| Coordinate precision of folded regions (residues 5-111) |  |
|---------------------------------------------------------|--|
| Backbone (Å)                                            | 0.27 ± 0.03 |
| All heavy atoms (Å)                                     | 0.80 ± 0.04 |
Figure 3

a

b

Figure 3
Figure 4

Part (a) and (b) show molecular structures, with parts labeled with Roman numerals and colors for different components. Part (c) and (d) display molecular surfaces with specific residues and labels indicated.
Figure 5
Figure 6

Fig. 6

\[ K_d = (2.7 \pm 0.3) \times 10^{-6} \]

\[ K_d = (4.0 \pm 0.7) \times 10^{-6} \]
Figure 8
