The NMDA Receptor NR1 C1 Region Bound to Calmodulin: Structural Insights into Functional Differences between Homologous Domains

Zeynep Akyol Ataman,1 Lokesh Gakhar,1 Brenda R. Sorensen,1 Johannes W. Helt,2 and Madeline A. Shea1,*

1Department of Biochemistry
2Department of Pharmacology
Roy J. and Lucille A. Carver College of Medicine, University of Iowa, Iowa City, IA 52242-1109, USA
*Correspondence: madeline-shea@uiowa.edu
DOI 10.1016/j.str.2007.10.012

SUMMARY

Calmodulin (CaM) regulates tetrameric N-methyl-D-aspartate receptors (NMDARs) by binding tightly to the C0 and C1 regions of its NR1 subunit. A crystal structure (2HQW; 1.96 Å) of calcium-saturated CaM bound to NR1C1 (peptide spanning 875–898) showed that NR1 S890, whose phosphorylation regulates membrane localization, was solvent protected, whereas the endoplasmic reticulum retention motif was solvent exposed. NR1 F880 filled the CaM C-domain pocket, whereas T886 was closest to the N-domain pocket. This 1-7 pattern was most similar to that in the CaM-MARCKS complex. Comparison of CaM-ligand wraparound conformations identified a core tetrad of CaM C-domain residues (FLMMC) that contacted all ligands consistently. An identical tetrad of N-domain residues (FLMMN) made variable sets of contacts with ligands. This CaM-NR1C1 structure provides a foundation for designing mutants to test the role of CaM in NR1 trafficking as well as insights into how the homologous CaM domains have different roles in molecular recognition.

INTRODUCTION

The ionotropic N-methyl-D-aspartate receptor (NMDAR) is a major source of calcium flux into neurons in the brain and has a critical role in learning, memory, neural development, and synaptic plasticity (Mori and Mishina, 1995). Mammalian NMDARs have two families of subunits designated NMDAR1 (NR1) and NMDAR2 (NR2). Understanding the roles of NR1 and NR2 subunits in brain function is complicated by variable developmental and spatial expression of their mRNA, and by the presence of eight variants of NR1 arising from N-terminal sequence variations and alternative splicing of four exons encoding the C0 (membrane-proximal), C1, C2, and C2’ regions (Figure 1A). The C-terminal tail of NR1 binds several proteins, including calmodulin (CaM). CaM has two highly homologous domains (N and C); each domain has two EF-hand sites that bind calcium cooperatively. CaM regulates a wide array of target proteins, including kinases, phosphatases, and ion channels (cf. Bhattacharya et al., 2004; Chin and Means, 2000; Saimi and Kung, 2002; Vetter and Leclerc, 2003).

Upon calcium influx, CaM induces inactivation of NMDAR (reducing its open rate and mean open time) by binding to NR1 C0 and C1 (Ehlers et al., 1996b). Inactivation releases NMDAR from the neuronal cytoskeleton by disrupting interactions between NR1 and z-actinin2 (Krupp et al., 1999; Zhang et al., 1998). CaM binding to NR1 C1 significantly enhances NMDAR inactivation (Ehlers et al., 1996b), which absolutely requires C0 (Zhang et al., 1998). Dissociation constants for (Ca2+)4-CaM binding to peptides representing the CaM-binding domains (CaMBDs) of NR1 showed that its affinity for C1 (NR1C1p; aa 875–898) was 20-fold more favorable than its affinity for C0 (NR1C0p; aa 838–863) (Ehlers et al., 1996b). The Kd values (4 and 87 nM, respectively) suggest that (Ca2+)4-CaM binds both C0 and C1 intracellularly, where [CaM]free is 50–75 nM (Wu and Bers, 2007). C1 has an endoplasmic reticulum (ER) retention motif (R893–R895; Figure 1A) and a Protein Kinase C (PKC) phosphorylation site (S896) that are needed for proper maturation and ER release (Scott et al., 2001; Standley et al., 2000). NR1 S890 phosphorylation by PKC disrupts surface-associated clusters of NR1, causing an even distribution throughout fibroblasts (Tingley et al., 1997), and affects receptor potentiation (Zheng et al., 1999). The position of CaM relative to these regulatory motifs in NR1 is not known.

To explore the role of CaM in NR1 trafficking, it is not possible to make a viable CaM knockout organism because CaM is essential and has multiple targets. Our strategy was to determine the binding interface to provide a platform for targeted mutagenesis of NR1. Here, we report a crystal structure of (Ca2+)4-CaM bound to NR1C1p (1.96 Å resolution) in which the CaM domains wrapped around helical NR1C1p (Figures 2A and 2B). The ER retention motif (R893–R895) and S896 were solvent exposed, whereas S890 was buried in the N domain of CaM. NR1C1p was predicted to be a 1-12 CaMBD (Yap et al., 2000), but it was found to be a 1-7 CaMBD. NR1 F880
was anchored in the C domain of CaM, whereas T886 (rather than F891) contacted the highest number of residues in the CaM N domain. The same 1-7 motif and nearly identical primary contact residues were identified for CaM bound to a MARCKS (myristoylated, alanine-rich, PKC substrate) peptide (1IWQ). In that case, CaM binding interrupts attachment of the actin cytoskeleton to the plasma membrane (Aderem, 1992). Parallels between proteins regulating MARCKS attachment to the cytoskeleton and proteins binding C1 suggest a similar mechanism of regulation of C1 in the formation of NR1-rich clusters (Ehlers et al., 1996b; Tingley et al., 1997).

The 1-7 motif found here is unusual among CaM-target interfaces in 17 complexes in which both the N and C domains of CaM contacts the target. The only other known case is CaM bound to MARCKS. To determine whether the nature as well as the spacing of CaM residues critical to molecular recognition were different among these structures, we analyzed the CaM-target contacts in all 17 complexes and found that 4 residues in the C domain (F92, L105, M124, and M144: FLMMC) were used consistently by CaM to contact targets. Although a structurally equivalent tetrad (F19, L32, M51, and M71: FLMMN) was observed in the CaM N domain, these CaM residues were not used identically by all targets.

RESULTS

Binding of CaM to NR1C1p and NR1C0p

Titrations of NR1C1p with (Ca2+)4-CaM and apo (Ca2+-depleted) CaM (Figure 1B) yielded a K_d of 2.0 ± 0.1 nM for...
Calmodulin Bound to NMDAR NR1 C1 Region

Table 1. Crystallographic Data Collection and Refinement Statistics

<table>
<thead>
<tr>
<th>Crystallographic Data Collection and Refinement Statistics</th>
<th>(Ca2+)4-CaM-NR1C1p Crystal</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Space group</strong></td>
<td>P321</td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
<td>a = 40.361 Å; b = 40.361 Å; c = 175.765 Å; α = 90°; β = 90°; γ = 120°</td>
</tr>
<tr>
<td><strong>Resolution (Å)</strong></td>
<td>19.69–1.90 (2.00–1.90)</td>
</tr>
<tr>
<td><strong>Rsym or Rmerge (%)</strong></td>
<td>0.034 (0.255)</td>
</tr>
<tr>
<td><strong>I/σI</strong></td>
<td>17.36 (3.92)</td>
</tr>
<tr>
<td><strong>Completeness (%)</strong></td>
<td>98.4 (92.6)</td>
</tr>
<tr>
<td><strong>Redundancy</strong></td>
<td>3.01 (2.56)</td>
</tr>
<tr>
<td><strong>Resolution (Å)</strong></td>
<td>8.57–1.90</td>
</tr>
<tr>
<td><strong>Number of reflections</strong></td>
<td>13,379</td>
</tr>
<tr>
<td><strong>Rwork/Rfree (%)</strong></td>
<td>20.6/24.8</td>
</tr>
<tr>
<td>**B factor for protein (Å) **</td>
<td>33.8</td>
</tr>
<tr>
<td><strong>B factor for ligand</strong></td>
<td>43.5</td>
</tr>
<tr>
<td><strong>B factor for ions</strong></td>
<td>31.4</td>
</tr>
<tr>
<td><strong>B factor for water</strong></td>
<td>41.8</td>
</tr>
<tr>
<td><strong>Rmsd bond lengths (Å)</strong></td>
<td>0.019</td>
</tr>
<tr>
<td><strong>Rmsd bond angles (°)</strong></td>
<td>1.585</td>
</tr>
<tr>
<td><strong>Number of protein atoms</strong></td>
<td>1,089</td>
</tr>
<tr>
<td><strong>Number of ligand atoms</strong></td>
<td>178</td>
</tr>
<tr>
<td><strong>Number of ions</strong></td>
<td>4</td>
</tr>
<tr>
<td><strong>Number of water atoms</strong></td>
<td>66</td>
</tr>
<tr>
<td><strong>Ramachandran plot (% residues)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Most favored</strong></td>
<td>94.2</td>
</tr>
<tr>
<td><strong>Additionally allowed</strong></td>
<td>5.8</td>
</tr>
<tr>
<td><strong>Disallowed</strong></td>
<td>0.0</td>
</tr>
</tbody>
</table>

Values in parentheses refer to the highest-resolution shell.

Structure

Calmodulin, which binds with an affinity of 4 nM (Ehlers et al., 1996b). A dissociation constant of 0.3 ± 0.025 μM (Kd of 6.33E-3 M⁻¹) was resolved for apo CaM binding to NR1C1p. Compared to NR1 C0, (Ca2+)4-CaM binds CO with a Kd of 87 nM (Ehlers et al., 1996b) and apo CaM binds CO with a Kd of 2.25 μM (Akyol et al., 2004). Comparison of simulated equilibrium titrations of CaM binding to NR1C1p and NR1C0p (Figure 1C) showed that the midpoints for apo and (Ca2+)4-CaM binding to NR1C1p differed by four orders of magnitude.

**Structure of the (Ca2+)4-CaM-NR1C1p Complex**

The crystal structure of (Ca2+)4-CaM bound to NR1C1p was determined to 1.96 Å resolution (Figures 2A–2C). It adopted the canonical CaM-target conformation in which both the N and C domains of CaM contacted peptide to form a compact, ellipsoidal complex. CaM residues 1, 2, 75–80, and 148 were disordered and were not included in the model. Figure 2A shows the electron density map for NR1C1p (residues 897 and 898 were disordered); the peptide was built into the final model manually. Refinement statistics are given in Table 1.

This structure was compared to 16 other similar and nonredundant (Ca2+)4-CaM-target structures (listed in Experimental Procedures). To evaluate the overall structural variability in these 17 compact CaM-target complexes, we compared each one to an average structure as described in Experimental Procedures. Among 17 structures, the average rmsd of the CaM N domain was 0.75 Å, whereas that of the C domain was 0.59 Å (Figures 2D and 2E). The complexes that showed the highest deviation from the average backbone conformation in the C domain were those determined by NMR (CaM with CNG channel [1SY9], CaMKK [1CKK], and skMLCK [2BBM]); skMLCK also had the highest rmsd bond angles (1.585 °). Two structures of the drug TFP bound to CaM (1A29 for the N domain; 1LIN for the C domain) had the smallest rmsd values.

**Accessibility of NR1C1p Motifs**

Processing and localization of C1-containing NR1 subunits is regulated by an ER retention motif (R893–R895) and by phosphorylation of S896. These residues have a high fractional solvent-accessible surface area (SASA): R893, 89.4%; R894, 71.6%; R895, 92.4%; S896, 87.1%; average SASA, 85% (Figure 3A). In contrast, S890 (implicated in subunit clustering and receptor potentiation by PKC) was protected by the N domain of CaM, having only 41% SASA. Two CaM N-domain residues (M36, M51) were within 4.5 Å of NR1 S890 as determined by using Contacts of Structural Units (CSU) (Sobolev et al., 1999). The hydroxyl of S890 was 3.22 Å from the sulfur of M36 and 4.01 Å from that of M51, suggesting that these residues interact in the complex (Figure 3B). No contacts were observed between NR1 S890 and any residues in the CaM C domain.

**Interface Contacts**

To explore the CaM-NR1C1p interface, CSU was used to determine CaM residues within 4.5 Å of NR1C1p residues (Figure 4). Among the ordered side chains, this analysis identified 36 residues in the N domain (residues 3–74) and 34 in the C domain (residues 81–147) that met this criterion. As shown in Figures 4A and 4B, contacts with the CaM N domain were well distributed across the length of NR1C1p: 17 with the N-terminal half (residues 875–885; gray) and 19 with the C-terminal half (residues 886–896; black). In contrast, contacts with the CaM C domain were skewed: 27 with the N-terminal half of NR1C1p and only 7 with the C-terminal half (Figures 4B and 4C). While CSU analysis showed that most NR1C1p residues contacted a single CaM domain, side chains of NR1C1p K875, K876, T879, and L887 contacted 2 or more residues in each domain. The 7 contacts of K875 are shown in purple and are underlined in Figure 5A. NR1C1p F880 contacted the highest number of residues.
within a single domain of CaM (Figure 4B). F880 was within 4.5 Å of F92, I100, L105, M124, A128, F141, and M144 (red letters in Figure 5A). All have hydrophobic side chains located in the C domain of CaM. The peptide residue making the second highest number of contacts within a single domain of CaM was T886. Its partners in CaM were F19, L32, M36, M51, and M72, all hydrophobic side chains in the N domain (blue letters in Figure 5A). In the common parlance of CaM-peptide interactions, F880 and T886 qualify as peptide anchors in the hydrophobic clefts of the C and N domains, respectively; however, the anchors are usually both hydrophobic. Alignment of the CaM sequence by its calcium-binding sites (Figure 5A) illustrated that a tetrad of the CaM residues contacting F880 and T886 residues were a set of identical side chains (FLMM) in corresponding positions: F19/F92, L32/L105, M51/M124, and M71/M144 (boxed, Figure 5A).

The domains of CaM in complex with NR1C1p were aligned by minimizing the distance between the Cα atoms of the FLMM tetrad residues in each domain; their backbone structures were closely aligned (Figure 5B). The rmsd for the Cα atoms of the FLMM tetrad in each domain was 0.208 Å, and this value was 0.573 Å for a comparison of the whole domain. Thus, the FLMM residues in each domain (i.e., FLMMN and FLMMC) appeared to adopt nearly identical spatial conformations (Figure 5C). The electron density overlap of side chains of FLMMN residues and NR1C1p F880 is shown in Figure 5D. The perpendicular orientation of NR1C1p F880 relative to CaM F92 allows for a favorable π–π interaction between the two aromatic rings (Singh and Thornton, 1992).

Identifying CaM Residues Commonly Used for Target Interactions

There is only one other structure of CaM bound to a 1–7 CaMBD motif, but there are numerous compact, ellipsoidal CaM-target structures. To explore how the CaM-NR1C1p interface related to those complexes, we used CSU to conduct a statistical analysis of CaM residues contacting targets in 16 other compact (Ca²⁺)₄-CaM-target structures (12 CaM-peptide, 4 CaM-drug complexes; listed in Experimental Procedures). In the set of 17 structures analyzed, 3 CaM residues (F92, M124, and M144) contacted every target; in all but one structure, L105 also contacted the target (Figure 6A). Thus, FLMMN consistently serves as the vertices of the C-domain hydrophobic pocket in these structures. An overlay of domains aligned according to the Cα atoms of the FLMMN tetrad shows that the positions of these FLMM residues in all structures is in Figure 6B.

A corresponding analysis of the N domain showed that although residues in the FLMMN tetrad were contacted in at least 12 of the analyzed structures, these residues were not the 4 residues contacted most frequently (Figure 6C). Instead, E11 was the only residue found to be within 4.5 Å of the target peptide or drug in all 17 of the structures that were examined. However, 2 of those 17 structures (1CTR.pdb and 1A29.pdb) were CaM-drug complexes in which the pocket of the N domain was vacant. Thus, E11 interacted with target molecules bound exclusively in the hydrophobic cleft of the C domain. The second most commonly used N-domain residue, A15, also contacts the target associated with the C domain of CaM. In the 15 structures that had the hydrophobic cleft of the N domain occupied, F19 contacted the target in all of them, as did E14. However, the frequency of use of other FLMMN residues (L32, M51, and M71) was lower and was dispersed among other hydrophobic N-domain residues (L18, M72, M36, F68, and L39) that contacted the target in as many or more structures. An overlay of 17 N domains aligned according to Cα atoms of the FLMMN tetrad (Figure 6D) shows that the pocket formed is very similar in all structures.

A comparison of side chain orientations of each FLMM residue in these 17 structures is shown in Figure 6E. Each FLMM residue was aligned with the corresponding residue in 2HQW; rmsds ranged from 0.2 to >1.6 Å. All FLMM residues, except F92, had side chain orientations that deviated by <1.0 Å in most structures; the smallest deviations were observed for M residues. For residues F19, L32, and F92, deviations ranged from 1.2 to 1.4 Å. In these, F19 and F92 were rotated by ~90° relative to the orientation observed in 2HQW (Figures 6F and 6H),
Figure 4. Distribution of CaM N- and C-Domain Contacts in the CaM-NR1C1p Complex

(A) N-domain residues ≤ 4.5 Å of NR1C1p shown as sticks; 17 contacts were made with NR1 residues 875–885 (gray), and 19 contacts were made with residues 885–896 (black).

(B) Sequence map of CaM residues ≤ 4.5 Å of NR1C1p. Residues in NR1C1p that make the highest number of contacts exclusively with the C domain (F880) and the N domain (T886) are boxed; the ER retention signal is underlined.

(C) C-domain residues ≤ 4.5 Å of NR1C1p shown as sticks; 27 contacts were made with residues 875–885, and 7 contacts were made with residues 885–896. Ca²⁺ ions and binding sites (yellow in [A] and [C]) are designated I, II, III, and IV. The figure was made with MacPymol.
and L varied most at the C_y and C_a atoms (see Figure 6G; Table S1, see the Supplemental Data available with this article online).

Identifying Target Residues that Contact the FLMM Tetrad of CaM
An analysis of the chemical characteristics of the target residues that contact FLMM_N and FLMM_C revealed that these tetrads were not used identically by the targets. Figure 7A shows sequences of the peptide in 13 compact CaM-target complexes. These were aligned according to the peptide residue (red box) that contacted the highest number of FLMM_C residues. These were aligned according to the peptide residue (red box) that contacted the highest number of FLMM_C residues. In 11 of 13 complexes, this residue also made the highest number of contacts with all C-domain residues of CaM (Figure 8, red bars). In two cases, there were 2 residues (Y1627 and F1628 of the Cav1.2 channel in 2BE6, and W3620 and L3623 of hRyR1 in 2BCX) that each contacted 3 FLMM_C residues; F1628 in 2BE6 and W3620 in 2BCX had the highest number of contacts with all C-domain residues of CaM. In 12 of these 13 structures, the residue that contacted the highest number of FLMM_C residues had a large aromatic moiety (7 F, 5 W, 1 Y) and two (CaMKII [1CDM] and hRyR1 [2BCX]) had a leucine residue in the cavity defined by the FLMM_C tetrad. A structural alignment of these complexes according to the C_a atoms of the FLMM_C (Figure 7B) revealed that the orientations of the residue contacting the majority of these FLMM_C residues in all 13 structures were well conserved.

The residue in each peptide that contacted the highest number of the FLMM_N residues is boxed in blue in Figures 7A and 8. In the case of CaM bound to CaMKII (1CDM), smMLCK (2BBM), and hRyR1 (2BCX), more than one residue contacted an equal number of FLMM_N residues. In each structure, the residue that contacted the majority of FLMM_N residues was also the one that had the highest number of contacts with all N-domain residues of CaM, with the exception of the CaM-CaMKI structure (1MXE), in which R317 contacted one more residue than M316. However, unlike contacts in the C domain, there were other residues in the target that had the same number of contacts as these. For example, in the CaM-CaMKII structure (1CDM), R297, G301, L304, T305, and A309 all contacted the same number of N-domain residues (4) in CaM; however, of these, only T305 and A309 contacted the majority of the FLMM_N residues.

A large variation was observed in the size, chemical characteristics, and spacing of the residues that
contacted the majority of the FLMM residues relative to the primary anchor residue at the reference position of “1” (Yap et al., 2000). In 11 of the sequences, the residue contacting FLMM was a hydrophobic amino acid, but its position varied from 10, 11, 14, 16, or 17. The structures of CaM-NR1C1p (2HQW), CaM-MARCKS (1IWQ), and CaM-CaMKIIa (1CDM) were unusual in that the residue contacting the majority of the FLMM residues was polar (Ser or Thr) and at position 7. A structural alignment of these complexes according to the Ca atoms of the FLMM residues (Figure 7C) revealed a much larger variation in the position of the target residue in the FLMM cavity than was observed for the FLMMC cavity. In some structures (i.e., complexes with peptides from MARCKS, CaMKIIz, NR1 C1, eNOS, and Myosin VI), the FLMM cavity was empty or only partially occupied, as illustrated in Figure 7D for the CaM-NR1C1p structure. The C-domain primary contact residue of the peptide in this structure (F880) was observed to fill the FLMM cavity of CaM. However, T886, which contacts the majority (3) of the FLMM residues and makes the highest number of contacts with the N domain of CaM, appears to contact only the rim of the cavity defined by the FLMM tetrad (Figure 7D).

To explore the general availability of the FLMM tetras to binding of a hydrophobic moiety that is not restricted by the orientation and chemical linkage of residues in a target peptide, this analysis of the FLMM tetras was focused to include those in four CaM-drug compact complexes. Three of these have TFP (Trifluoperazine; 10-[3-(4-methyl-piperazin-1-yl)-propyl]-2-trifluoromethyl-10-phenothiazine) bound in 3 CaM:drug ratios (1:1, 1:2, 1:4), and the fourth has DPD (N-[3,3,-diphenylpropyl]-N0-[1-R-(3,4-BIS-butoxylphenyl)-ethyl]-propylenediamine) bound in a 1:2 ratio of CaM:drug (1QIV.pdb). In the structures with a 1:1 ratio of CaM:TFP (1CTR.pdb) and with a 1:2 ratio (1A29.pdb), only the FLMMC tetrad was occupied, and in the structure with a 1:4 ratio of CaM:TFP (1LIN.pdb) both tetras were occupied by TFP. A structural alignment of the two domains of 1LIN.pdb according to the positions of the Ca atoms of the FLMM tetras illustrated that TFP...
was capable of binding both FLMM cavities in the same orientation, and that the two domains of CaM bound to TFP have similar structures (rmsd of 0.492 Å). This similarity between domains was also observed in the CaM-DPD structure, in which both FLMM tetrads were occupied with the same moiety of a DPD molecule in the same orientation (Figure 7E). The N and the C domains of CaM in this structure align nearly as well as the domains of CaM when bound to NR1C1p (Figure 5B) (rmsd of 0.652 Å in the DPD complex versus 0.573 Å in the NR1C1p complex).

**DISCUSSION**

The C1 region of the NR1 subunit of the NMDA receptor has been shown to regulate receptor trafficking and decrease PKC-induced receptor potentiation. The structure of (Ca$^{2+}$)$_4$-CaM bound to NR1C1p (2HQW) presented...
here demonstrates that C1 is a 1-7 motif, and it indicates which residues contribute to the interface between CaM and NR1C1. It provides a foundation for the interpretation of NR1C1 mutants that would disrupt association with CaM and serve to test the role of CaM in NR1 trafficking.

Figure 8. Interface Analysis of 13 CaM-Peptide Complexes
Residues in the N domain (gray) and C domain (black) of CaM within 4.5 Å of a peptide residue determined with CSU. Red indicates the peptide residue contacting the highest number of C-domain residues; blue indicates the peptide residue contacting the highest number of N-domain residues.

Structure Calmodulin Bound to NMDAR NR1 C1 Region

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Number of Residue Contacts (≤ 4.5 Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1CDL</td>
<td>R RK W Q K T G A V R A I G R L S S</td>
</tr>
<tr>
<td>1SY9</td>
<td>G G F R R I A R L V G Y L R E W A Y R</td>
</tr>
<tr>
<td>1CDM</td>
<td>F N A R K R L K G A I L T Y M L A T</td>
</tr>
<tr>
<td>2BBM</td>
<td>R H K W K K N F I A V S A A N R F K R I S S S G A L</td>
</tr>
<tr>
<td>1CKK</td>
<td>F P N G F S R K R H G M A K V L I V T W S P I L K V</td>
</tr>
<tr>
<td>2BCX</td>
<td>K S K K A V W H K L S K Q R R A V V A C F R M T P</td>
</tr>
<tr>
<td>1IQ5</td>
<td>R F P N G F R K K R H G M A K V L I V T W S P I L K V</td>
</tr>
<tr>
<td>2BE6</td>
<td>V L G Q E K R K K K R F Y E Q I L F T A Y F K Q V T Y E</td>
</tr>
<tr>
<td>1IWQ</td>
<td>K K R F S F K K S F K L S Q F S F K</td>
</tr>
<tr>
<td>1MXE</td>
<td>I K K N F A S K K W K Q A F N A T A V V H R M K</td>
</tr>
<tr>
<td>2HQW</td>
<td>K K K A T F R A I T S T L A S S F K K R R R S</td>
</tr>
<tr>
<td>1NIW</td>
<td>R K K F K E V A N A V K I S A S L M</td>
</tr>
</tbody>
</table>
Physiological Significance of the CaM-NR1C1p Complex

NR1 subunits are expressed in an ~10-fold excess over NR2 subunits in the cell; however, only 40%–50% of these NR1 subunits reach the cell surface in cultured hippocampal neurons (Okabe et al., 1999). There are eight splice variants of NR1, and the major variant in the brain contains the C0-C1-C2 regions (Mori and Mishina, 1995); however, this variant has the lowest fraction of cell-surface expression (Okabe et al., 1999). ER retention of the NR1 subunit is mediated by a sequence of three contiguous R residues on the C1 region and is controlled by the phosphorylation of S896 after the ER retention signal (Scott et al., 2001). The structure of (Ca\(^{2+}\))\(_4\)-CaM bound to NR1C1p showed that the ER retention signal and S896 were not occluded by CaM. While this structure alone cannot rule out the possibility of a direct role of CaM in ER retention, it is more likely that CaM might serve as an indirect modulator by interacting with a kinase or other protein that has a role in ER retention.

Studies of NR1 splice variants containing the C1 region showed that this region was necessary and sufficient for the formation of discrete subcellular receptor clusters that are associated with the plasma membrane when expressed in fibroblast cells (Ehlers et al., 1995). Phosphorylation of S890 within the C1 region by PKC disrupts the receptor-enriched clusters, resulting in an even distribution of the NR1 subunit (Tingley et al., 1997). This residue had only 41% solvent accessibility when complexed with (Ca\(^{2+}\))\(_4\)-CaM, and its hydroxyl group was within 4 Å from the sulfur groups of M36 and M51 of CaM. We have shown that (Ca\(^{2+}\))\(_4\)-CaM binds NR1C1p with a high affinity (K\(_D\) = 2 nM). Ehlers and coworkers (Ehlers et al., 1996b) reported comparable affinities for NR1C1p and a larger fusion protein of NR1C1p (~150 kDa; K\(_D\) = 4 nM). Together, these data suggest that CaM can protect this region from other proteins in the presence of calcium.

Common Features of Hydrophobic Cavities in CaM Domains

Early studies of CaM-peptide interactions demonstrated the importance of hydrophobic residues found on both the target and CaM, particularly the role of tryptophan in a peptide and the methionine “puddles” of the CaM hydrophobic clefts (O’Neil and DeGrado, 1990). Subsequent structural comparisons have concluded that large hydrophobic residues of the target bind the hydrophobic pockets of CaM, and that orientation of binding is determined by the electrostatic characteristics of the domains of CaM (cf. Bhattacharya et al., 2004; Hoeflisch and Ikura, 2002; Ishida and Vogel, 2006; Vetter and Leclerc, 2003). The roles of individual Met residues on CaM in enzyme activation have been investigated mutagenesis (Chin et al., 1997). Dynamics simulations further support the thesis that a set of conserved methionines in each domain of CaM are flexible and allow for accommodation of variable peptide residues (Fiorin et al., 2006). A structural overlay of the domains of CaM in seven wrap-around CaM-target structures available in the PDB (Ishida and Vogel, 2006) highlighted a group of 7 hydrophobic residues (4 M, 2 F, 1 L) in each domain of CaM that surround the hydrophobic pockets in these structures. On the basis of a comprehensive statistical analysis of the contact distances in 17 wrap-around CaM-target structures, we have identified the FLMM\(_2 \) tetrad as prongs that hold a hydrophobic residue in all canonical CaM-target complexes studied here, whereas the corresponding FLMM\(_3 \) tetrad is not as well contacted.

Conservation of FLMM Tetrads in CaM Sequences

Given the prevalence of the FLMM tetrads in the target-binding pockets of CaM, it was expected that these residues would be conserved in the sequence of CaM across species. Comparison of 102 CaM sequences (Figure 9) revealed that 4 of the 8 FLMM residues (F19, L32, F92, and M124) were completely conserved, and 2 (M51, L105) were 99.98% conserved: position 51 was M in all but two sequences (Calm_Yeast and Calm_KLULA have L), and position 105 was L in all but two sequences (Calm2_Pethy has V and Calm_Mousc has W; see Table S2 for complete analysis and accession numbers). In the 102 CaM sequences analyzed, this was the only occurrence of W at any position. It is possible that a sequencing ambiguity could account for the two substitutions observed at this position. Codons for W (UGG) and V (UGU) differ by a single base from the sequence for L (UUG), raising the possibility that L105 is also completely conserved.

The hydrophobicity of the remaining 2 FLMM tetrad residues (M71, M144) corresponding to the fourth position of the tetrad in both domains of CaM (Figure 5A) was highly conserved among the 102 sequences compared; however, these 2 residues showed higher sequence variation than the other 6 FLMM residues. Position 71 was M in 62 sequences and L in the other 40. Position 144 was M in 67 sequences, V in 23, L in 10, and I in 2. The smaller size of the variant side chains may allow the pocket to accommodate larger anchor residues.

Conservation of FLMM Tetrads

A search of the SWISS- PROT knowledgebase and the Protein Data Bank (PDB) for all proteins with identical spacing of the primary sequence for FLMM (e.g., F-(x12)-L-(x18)-M-(x19)-M) with PROSITE (Sigrist et al., 2002) on the ExPASy Proteomics Server (http://ca.expasy.org) (Gasteiger et al., 2003) identified 361 nonredundant sequences in these databases (i.e., for sequences found both in the PDB and the SWISS- PROT knowledgebase, the PDB sequence was omitted; Table S3). Of these, 132 were sequences of CaM with 2 listings per accession number, 1 for each domain. There were six structures available in the PDB of non-CaM sequences that contained an identical FLMM primary sequence: CaM-like protein 3 (1GGZ-A), the C domain of CaM-like protein 5 (2B1U-A), the N domain of centrin (caltractin; 2AMI-A), cytochrome P450 (2IJ5-A), aspartyl-tRNA synthase (1C0A-A), and the M chain of the photosynthetic reaction center (1AIJ-M). Alignment of Ca\(_X\) atoms of the
In these structures with Ca\(_{\alpha}\) atoms of the FLMMC residues in 2HQW indicated that only the two CaM-like proteins and centrin form an FLMM pocket similar to that observed in CaM (see Figure S1). The rmsds for the remaining three structures ranged from 9.5 to 13.0 Å, compared to 0.36–2.0 Å observed for the two CaM-like proteins and centrin. Notably, other EF-hand proteins that bind calcium, but cannot substitute for CaM in cellular function (e.g., troponin C, S100B, calbindin, and parvalbumin), lack the FLMM tetrad. An alignment of these proteins with rat CaM (3CLN) by using ClustalW (Chenna et al., 2003) indicated that Troponin C (4TNC) has I51 and L124; S100B (1PSB) has Y19, V51, and F71; calbindin has Y19, L51, and L71; and parvalbumin (1B8R) has C19, V32, A51, and F71.

**Common Features of Residues that Occupy the Hydrophobic Pockets of CaM**

The diversity among sequences that represent the CaMBDs of target proteins makes it difficult to classify them and predict precisely how they will interact with domains of CaM. Common classification schemes rely on identifying hydrophobic side chain residues presumed to anchor the peptide to the N and C domains of CaM and matching the number of residues between the presumed anchors to a known CaMBD. However, the prediction that NR1C1p was a 1-12 motif (Yap et al., 2000), with F880 as the bulky hydrophobic residue at position 1 and F891 at position 12, did not correspond to the observed structure of CaM bound to NR1C1p. Although F880 was found in the C-domain hydrophobic pocket, F891 was not in or near the corresponding pocket of the N domain. F891 made contacts with only 4 CaM residues (Q41, E84, E87, A88), 3 of which were in the C domain rather than the N domain, and only one of which was aliphatic. This emphasizes the difficulty in predicting which residues in a CaMBD will make close contacts with CaM and occupy the pocket of each domain. On the basis of this structure, we propose that mutation of NR1 F880 to any residue other than Trp would decrease the affinity of NR1 for CaM. Mutation of NR1 K875 (which also contacts 7 CaM residues) is anticipated to affect binding orientation and/or affinity.
Variable Use of FLMM₂ and FLMMᵡ by Targets
An alignment of 17 CaM-target complexes according to the C2 atoms of FLMM₂ and FLMMᵡ illustrated that the backbone structures of the N and C domains of CaM when bound to these targets differed by less than 1.0 Å (see Figures 2D and 2E). Our analysis showed that despite the high degree of homology between the structures and sequences of the N and C domains of CaM, they are not used in the same way by the 13 canonical CaM-peptide structures studied here. The cavity formed by the FLMMᵡ residues was consistently occupied by a large aromatic moiety, whereas that formed by the FLMM₂ residues was occupied by a variety of smaller, hydrophobic or polar residues. However, this does not seem to be due to a structural restraint of the N domain in binding a large aromatic residue because structural alignments of the N domain and C domain of CaM that have the same drug molecules in their hydrophobic pockets show that the drugs can bind both domains in a similar manner (Figure 7E). These findings suggest that a reason for the different usages of the FLMM₂ and the FLMMᵡ pockets could be because of the tertiary constraints that the peptides have, such as the distances between the two potential anchors or their orientation on the helix (i.e., NR1 F880 and F891 both face toward the C domain).

Summary
We present a crystal structure of (Ca²⁺)₄-CaM in complex with a peptide corresponding to the NR1 C1 subunit of the NMDA receptor. The complex displays the canonical wrap-around CaM-target-binding conformation and illustrates the position of CaM with respect to ER retention and phosphorylation sites. This study also sheds light on the structural basis for functional differences between the highly homologous domains of CaM. The C domain of CaM provides the major anchoring site for NR1C1p, whereas that formed by the FLMM₂ residues was occupied by a variety of smaller, hydrophobic or polar residues. However, this does not seem to be due to a structural restraint of the N domain in binding a large aromatic residue because structural alignments of the N domain and C domain of CaM that have the same drug molecules in their hydrophobic pockets show that the drugs can bind both domains in a similar manner (Figure 7E). These findings suggest that a reason for the different usages of the FLMM₂ and the FLMMᵡ pockets could be because of the tertiary constraints that the peptides have, such as the distances between the two potential anchors or their orientation on the helix (i.e., NR1 F880 and F891 both face toward the C domain).

EXPERIMENTAL PROCEDURES

Materials
NR1C1p (KKKATFRAITSTLASSFKRRRSSK) and fluorescein-labeled NR1C1p (F1-NR1C1p) were purchased from Keck Biotechnology Resource Center (New Haven, CT). Recombinant mammalian CaM was prepared as described previously (Pedigo and Shea, 1995; Sorensen and Shea, 1998). CaM was depleted of Ca²⁺ by successive dialyses at 4°C against 50 mM HEPES, 100 mM KCl, 1 mM MgCl₂, and 0.5 mM NTA with 5 mM EGTA and then with 0.05 mM EGTA.

Optimization of Crystallization Conditions
Crystals of the CaM-NR1C1p complex grew at 4°C with a reservoir solution of 7% PEG6000, 5 mM CaCl₂, 100 mM Na-acetate (pH 4.0) by using the hanging-drop vapor-diffusion technique. Two microliters of a solution containing 0.9 mM NR1C1p and 0.7 mM (Ca²⁺)₂-CaM was mixed with an equal volume of the reservoir solution. Crystals were cryoprotected in mineral oil and flash cooled in liquid nitrogen.

Diffraction Data Collection and Processing
Crystals diffracted to 1.96 Å resolution. Data were collected at the Advanced Photon Source (APS; Argonne National Laboratory) at beamline 17-ID, Industrial Macromolecular Crystallography Association Collaborative Access Team (IMCA-CAT). Data were processed by using o’TREK (Pflugrath, 1999) from Rigaku/MSC and REFMAC5 (Murshudov et al., 1997); the starting model was 1WQ (Yamauchi et al., 2003). NR1C1p was modeled with O (Jones et al., 1991). Water was
Calmodulin Bound to NMDAR NR1 C1 Region

added by using Arp/Warp (part of CCP4 suite [CCP4, 1994]) and was verified in O. TLS refinement used groups defined as N domain (chain A, residues 3–74 and Ca²⁺ I and II), C domain (chain A, residues 81–146 and Ca²⁺ III and IV), and peptide (chain B, residues 875–896). Electron densities for side chains of residues 893–896 of NR1C1p were not well defined; rotamers that minimized electrostatic clashes for 893–896 were selected for the final model, which was verified with PROCHECK and accepted for deposit by the PDB. Table 1 summarizes crystallographic statistics.

Solvent-Accessible Surface Area Calculations
A 1.4 Å probe in WHAT IF (Vriend, 1990) and the VACCx command were used to calculate % solvent-accessible surface area (SASA) for each residue relative to its accessibility in an unfolded peptide (see Write-up and ACCESS links at http://swift.cmbi.kun.nl/whatif). The %SASA reported was an average value for the side chain atoms. The model was colored by using PyMOL script color_b.py (Queens University, http://adelie.biochem.queensu.ca/~rlc/work/pymol).

Structure Alignment and Analysis
MacPymol 0.99rc6 (DeLano Scientific, CA) was used to align Cx atoms of the FLMM residues of domains in a single structure (e.g., 2HQW [NR1C1p; Figures 5B–5D], 1LIN [TFP, 1:4], and 1QIV [DPD; Figure 7E] or between multiple structures (see Figures 6 and 7) in which the four Cx atoms of the FLMM within another structure were aligned with the corresponding atoms of 2HQW. Rmsd values for domain alignments within a single structure were calculated for residues 8–73 (66 atoms) and 81–146 (66 atoms) by using MacPymol. These were chosen to minimize contributions from highly flexible regions (e.g., termini and linker) of the complexes.

For CaM-peptide complexes compared to 2HQW, the peptide, resolution, and PDB ID were CaM-dependent protein kinase II (CaMK II, 2 Å; 1CDM), smooth muscle myosin light-chain kinase (smMLCK, 2.2 Å; 1CDL), nematode CaM-dependent protein kinase kinase (CaMKK, 1.8 Å; 1QOS), myristoylated alanine-rich C-kinase substrate (MARCKS, 2 Å; 1WQ), endothelial nitric oxide synthase (NOS, 2.05 Å; 1NIW), myosin VI motor protein (Myosin VI, 2.4 Å; 2BK8), skeletal muscle myosin light-chain kinase (MLCK, NMR; 2BBM), olfactory CNG channel (Cng-2, NMR; 1SY9), CaM-dependent protein kinase I (CaMKI, 1.7 Å; 1MXE), CaM-dependent kinase kinase (CaMKK, NMR; 1CKK), voltage-dependent L-type calcium channel α-1C subunit (Ca.1.2, 2 Å; 2BE6) and the ryanodine receptor Type 1 (RYR1, 2.0 Å; 2BCX). Four other structures were CaM-drug complexes: CaM:TFP 1:1, 2.45 Å (1CTR); CaM:TFP 1:2, 2.74 Å (1A29); CaM:TFP 1:4, 2 Å (1LIN); and CaM:DPD 1:2, 2.64 Å (1QIV). Distances between residues in CaM and a target were calculated by using CSU (Sobolev et al., 1999). Protons were added to crystal structures by using WHAT IF (Vriend, 1990) to optimize the hydrogen-bonding network. A separation ≤ 4.5 Å qualified as a contact.

Calculation of a Reference Average Structure and Root-Mean-Square Deviations
Cα atoms for N-domain residues 5–72 (68 aa) or C-domain residues 81–145 (62 aa) were superimposed with corresponding atoms of 2HQW by using MacPymol. Resulting Cα positions were averaged with AVEPDB for O (Kleywegt et al., 2001) available from http://xray.bmc.uu.se/ufst. The rmsd of the N and C domains of each structure from the average structure was normalized for the number of residues (62 for N, 68 for C). The average rmsd value of all 17 structures was the average of the individual rmsd values.

CaM Titration of NR1C1p
The fluorescence anisotropy of Fl-NR1C1p (Figure 1B) was monitored with a Fluorolog 3 (Jobin Yvon, Horiba) spectrofluorimeter with a λex of 496 nm, a λem of 520 nm, and bandpasses of 2 nm excitation and 10 nm emission (Akyol et al., 2004). Averages of three readings with a 1 s integration time were recorded. Samples of 25 nM FI-NR1C1p in the absence of calcium or 3.33–6.22 nM FI-NR1C1p in the presence of 10 mM CaCl₂ in 50 mM HEPES, 100 mM KCl, 50 μM EGTA, 5 mM NTA, 1 mM MgCl₂ (pH 7.4) at 22°C were titrated with CaM stocks ranging from 1.33 to 732 μM. Each titration was replicated at least three times.

The FI-NR1C1p-CaM complex had a 1:1 stoichiometry. To estimate the affinity of CaM for FI-NR1C1p, data were fit to a one-site binding isotherm in Equation 1,

\[ \frac{Y_1}{K_p} = \frac{[\text{CaMtotal}]}{1 + [\text{CaMtotal}]} \]  

where \( K_p \) represents the association constant, and [CaMtotal] is bound CaM, calculated from the independent variable (total concentration of CaM, [CaMtotal]) (Figure 1B) according to the quadratic equation in Equation 2,

\[ [\text{CaMtotal}] = \frac{-b \pm \sqrt{b^2 - 4K_p([\text{Fl-NR1C1p}] - [\text{CaMbound}]})}{2K_p} \]  

where \( b = (1 + K_0[\text{Fl-NR1C1p} - \text{NR1C1p}_{\text{total}}] - K_0[\text{CaMbound}] \) Under apo conditions, the affinity for FI-NR1C1p was weak and [CaMtotal] = [CaMbound]. However, calcium increased the affinity such that [Fl-NR1C1p][CaMbound] = [CaMtotal]. Under these conditions, [CaMtotal] was limited and estimated iteratively in the nonlinear least-squares analysis as the difference between [CaMtotal] and [CaMbound] ([Fl-NR1C1p][CaMtotal] \( \gamma \)), a constant binding estimated this way is highly correlated with the numerical value of [Fl-NR1C1p][CaMtotal]; therefore, limiting values of \( \gamma \) were reported. Experimental variations in endpoints of individual titrations were accounted for by Equation 3,

\[ \text{Signal} = f(X) = \frac{Y_{\text{Fl}} + \gamma}{Y_{\text{Fl}} + \gamma + \text{Span}} \]  

where \( \gamma \) is the average fractional saturation of FI-NR1C1p (Equation 2), \( Y_{\text{Fl}} \) is the anisotropy of FI-NR1C1p alone, and Span (\( Y_{\text{Hi}} - Y_{\text{Lo}} \)) accounts for the magnitude and direction of signal change upon titration. The upper and lower endpoints of titrations in saturating calcium were well defined; but, in the absence of calcium, the upper endpoint was not well defined at the final [CaMtotal] tested (52 μM). In the nonlinear least-squares analysis, \( Y_{\text{Hi}} \) was fixed to be equivalent to the anisotropy after addition of calcium to the apo CaM:FI-NR1C1p sample.

Supplemental Data
Supplemental Data include a figure illustrating the conservation of the FLMM tetrads in six non-CaM proteins in the SWISS-PROT knowledgebase and PDB (Figure S1) and three tables showing the sequence homology of 101 CaM species (Table S1), the occurrences of the FLMM tetrad in identified protein sequences (Table S2), and rmsd values for variations in the side chain orientations of individual FLMM residues relative to those in 2HQW (Table S3) and are available at http://www.structure.org/cgi/content/full/15/12/1603/DC1/.

ACKNOWLEDGMENTS
This work was supported by AHA 0410078Z (Z.A.-A.), National Institutes of Health (NIH) R01 GM57001 (M.A.S.), and NIH R01 NS046450 (J.W.H.). We acknowledge use of the University of Iowa Protein Crystallography Facility; S. Ramaswamy for resources and training; and D. Ferraro and E. Brown for advice and data collection at the Advanced Photon Source (APS). Use of the IMCA-CAT beamline 17-ID at the APS was supported by the Industrial Macromolecular Crystallography Association through a contract with the Center for Advanced Radiation Sources at the University of Chicago. Use of the APS was supported by the U.S. Department of Energy, W-31-109-Eng-38.

References


**Accession Numbers**

Atomic coordinates were deposited in the Protein Data Bank under accession code 2HQW.