


REVIEW

Do calmodulin binding IQ motifs have built-in capping domains?

Arantza Muguruza-Montero^{1,6}  | Rafael Ramis^{2,4} | Eider Nuñez^{1,6} | Oscar R. Ballesteros^{3,4} | Markel G. Ibarluzea^{2,4} | Ariane Araujo^{1,6} | Sara M-Alicante¹ | Janire Urrutia^{1,5} | Aritz Leonardo^{2,4} | Aitor Bergara^{2,3,4} | Alvaro Villarroel^{1,6}

¹LaboKCNQ, Barrio Sarriena, Leioa, Spain

²Donostia International Physics Center, Donostia, Spain

³Centro de Física de Materiales CFM, CSIC-UPV/EHU, Donostia, Spain

⁴Departamento de Física, Universidad del País Vasco, UPV/EHU, Leioa, Spain

⁵Department of Physiology, Faculty of Medicine and Nursery, UPV/EHU, Leioa, Spain

⁶Instituto Biofisika, CSIC-UPV/EHU, Leioa, Spain

Correspondence

Arantza Muguruza-Montero and Alvaro Villarroel, LaboKCNQ, Barrio Sarriena, Leioa 48940, Spain.

Email: arantza.muguruza.montero@gmail.com (A. M.-M.) and alvaro.villarroel@csic.es (A. V.)

Funding information

Government of the Autonomous Community of the Basque Country, Grant/Award Numbers: IT1165-19, KK-2020/00110; Spanish Ministry of Science and Innovation, Grant/Award Numbers: PID2019-105488GB-I00, RTI2018-097839-B-100; Fondos Europeos Desarrollo Regional (FEDER)

Abstract

Most calmodulin (CaM) targets are α -helices. It is not clear if CaM induces the adoption of an α -helix configuration to its targets or if those targets are selected as they spontaneously adopt an α -helical conformation. Other than an α -helix propensity, there is a great variety of CaM targets with little more in common. One exception to this rule is the IQ site that can be recognized in a number of targets, such as those ion channels belonging to the KCNQ family. Although there is negligible sequence similarity between the IQ motif and the docking site on SK2 channels, both adopt a similar three-dimensional disposition. The isolated SK2 target presents a pre-folded core region that becomes fully α -helical upon binding to CaM. The existence of this pre-folded state suggests the occurrence of capping within CaM targets. In this review, we examine the capping properties within the residues flanking this core domain, and relate known IQ motifs and capping.

KEYWORDS

capping, calmodulin, IQ motif, SK channels

Abbreviations: CaM, calmodulin; CRD, calcium responsive domain; Θ , hydrophobic-bulky residue; H-bonds, hydrogen bonds; Hp, hydrophobic core; BB-BB, backbone-to-backbone; SC-SC, side-chain-to-side-chain; SK, small conductance Ca^{2+} -activated potassium channels.

1 | INTRODUCTION

Calmodulin (CaM) is a highly conserved eukaryotic protein formed by two similar globular domains, the N- and C-lobes, linked by a very flexible sequence.^{1,2} CaM can

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Authors. *Protein Science* published by Wiley Periodicals LLC on behalf of The Protein Society.

bind up to four Ca^{2+} ions and, depending on site occupancy, jumps through open, semi-open, and close conformations adjusting to the target α -helix,³ much like two hands holding onto a rope, where the thumbs represent one EF-hand, and the other four fingers constitute the other EF hand. CaM is a very flexible protein that can recognize a large number of diverse target sequences that almost invariably adopt an α -helical configuration.² CaM targets often have two anchoring bulky hydrophobic residues separated by 6, 10/11, 14, and 16 residues, that is, with an α -helical periodicity (3.6 residues per turn)⁴⁻⁶ (Figure 1). CaM versatility is so that it can even recognize the same α -helix in two orientations.^{5,8,9} It is not clear if CaM recognizes already formed α -helices (conformational selection) or if CaM induces the formation of the helix as a part of the target recognition process.

2 | CaM AS A FOLDING INDUCTOR FOR THE SK2 CHANNEL CALCIUM RESPONSIVE DOMAIN

Small conductance Ca^{2+} -activated potassium (SK) channels contribute to after-hyperpolarization following an action

potential. SK channels mediate the intrinsic excitability of many cells,¹⁰ modulate the activation of immune responses,¹¹ and contribute to the regulation of vascular tone.¹²

These Ca^{2+} operated channels have a six transmembrane domain (6TM) architecture (S1–S6) and both N- and C-termini are located in the intracellular region. The calcium responsive domain (CRD) is formed by CaM docked into two antiparallel α -helices (hA and hB) that run parallel to the plasma membrane. CaM remains constitutively bound through the C-lobe anchored at the distal part of hA, which harbors the sequence “LRxxWL,” conserved among the SK channel family. These residues within the helix A adopt an orientation that is structurally equivalent to that of IQ motifs of other targets anchored to the CaM C-lobe⁵ (Figure 2a). The N-lobe is freely moving and is responsible for gating. Upon Ca^{2+} loading, the N-lobe pulls down the S4-S5 α -helix linker and this mechanical force is transmitted to the pore domain, leading to opening of the gate.¹³

The IQ motif (IQxx Θ R, where a hydrophobic bulky residue often occupies the Θ position) is the best-recognized site for CaM anchoring and upon binding to CaM adopts an α -helical disposition within the C-lobe. The name of this well-known motif refers to the first two

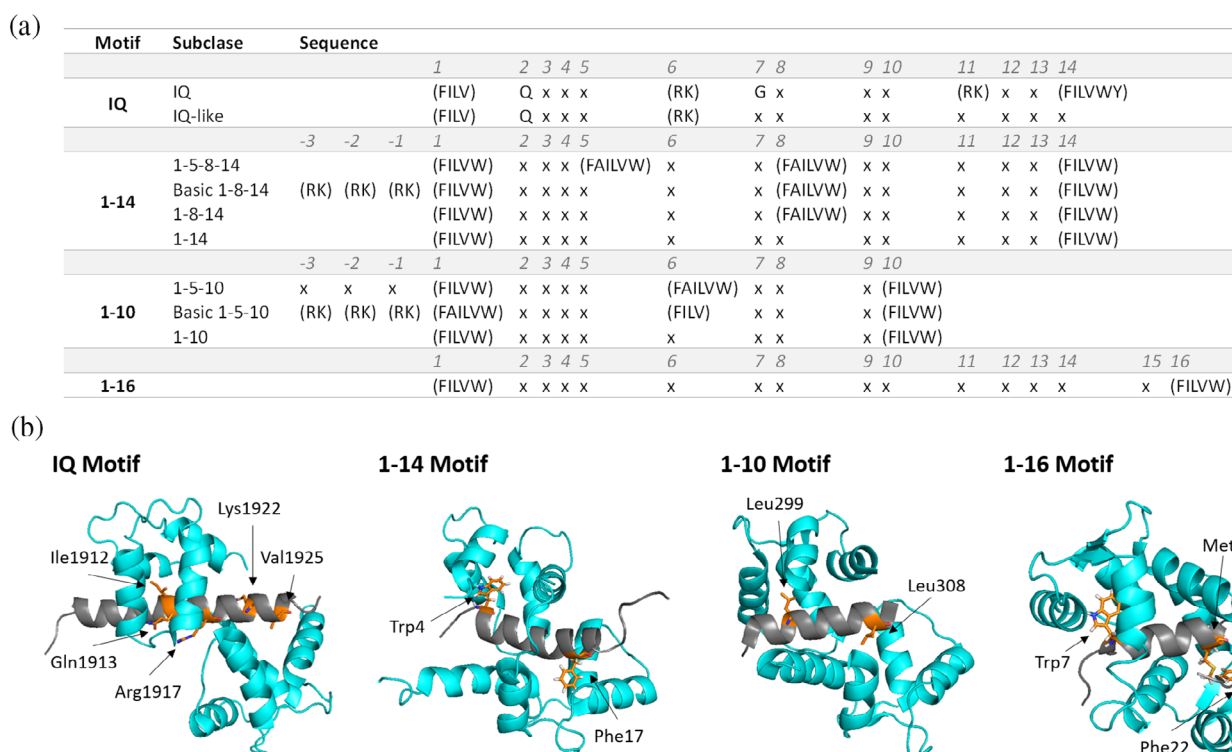


FIGURE 1 Two bulky hydrophobic residues define most CaM binding motifs. (a) Classification of CaM binding domains and motif alignments using one-letter amino acid codes. The residues within brackets often appear at the indicated position. “x” refers to any amino acid. Residue positions are indicated relative to the IQ motif (adapted from Calmodulin Target Database⁷). (b) Selected examples of helices (in grey) engaged with CaM (in cyan), indicating the anchoring residues as orange sticks (PDBs: 4JPZ for IQ motif, 2BBN for 1–14, 1CDM for 1–10 and 1CCK for 1–16)

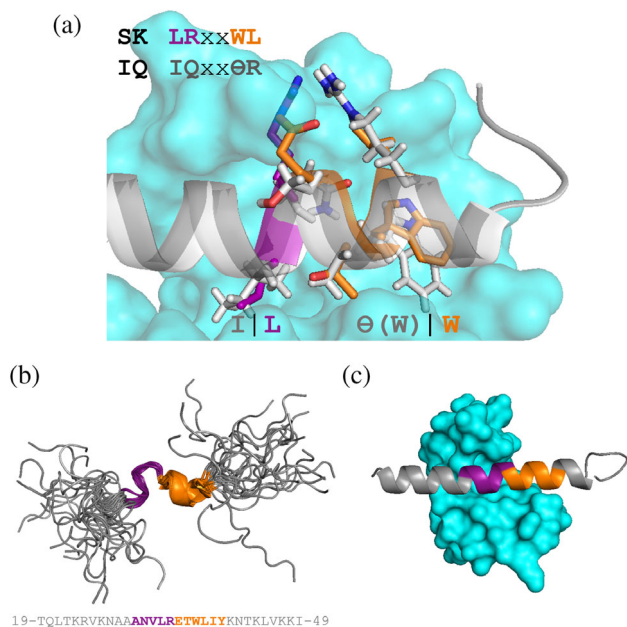


FIGURE 2 Structure of the hA region of SK2 in the absence and presence of CaM. (a) Superimposition of SK2 CRD (dark grey) and an IQ motif (from PDB: 6FEG) (light grey) in complex with C-lobe. The lateral chain of the CaM anchoring residues are shown as stick in grey for the IQ and in purple and orange for SK2 CRD, using the same color patterns as in (b). (b) Solution SK2 CRD structure resolved by NMR (PDB: 1KKD). The structural conformations of the 5-residue-long turn (purple) and 6-residue-long α -helix (orange) are similar in all 23 models. The sequence is shown at the bottom of the panel. (c) Structure of SK2 CRD in complex with the CaM C-lobe (PDB: 3SJQ). The surface of the C-lobe is shown in cyan and the color patterns of the CRD are as in (b). Images were rendered using PyMol 1.3

amino acids: isoleucine (I) and glutamine (Q). It was first described in 1988 as a CaM binding motif in neuromodulin,¹⁴ and thereafter, it was characterized in the myosin family of proteins.¹⁵ This amphipathic motif is present in several cell signaling, transport, and cytoskeletal proteins.⁴ Nevertheless, the IQ motif is just one of many sequences recognized by CaM. CaM can structurally recognize more than 300 target proteins with little sequence similarity, triggering diverse regulatory mechanisms.^{16,17}

Within the CRD of SK channels, the Ile position of the IQ motif, which is often replaced by Val or Leu, is occupied by Leu. The Θ position, which holds a hydrophobic residue in IQ peptides, is occupied by Trp432 in SK2.⁵ The equivalent region to the IQ motif is in a pre-folded state in the absence of CaM. The NMR structure shows an incipient helix region in solution.¹⁸ This structure (PDB: 1KKD) consists of a well-ordered core region, forming an unusually folded helical domain between Ala28 and Asn 42 (note that there are Ala423 and Asn437 in the full channel

enumeration) flanked by highly flexible N- and C-termini (Figure 2b). The CRD of SK2, however, forms an α -helix in complex with CaM (Figure 2c). Constitutive association with CaM does not occur after the introduction of mutations within hA (V32G/L33G), which abolish Ca^{2+} -dependent channel activation. Superposition of the backbone of the 15 best structures for the V32G/L33G mutant show a clear difference in comparison with WT. Whereas the WT peptide maintains a well-organized core region, the mutant peptide does not, indicating that these two residues are involved in the stabilization of the core helical region. These observations led to the proposal of the existence of a pre-folded state for some CaM targets that mediates the initial steps during target recognition. How is this pre-folded state maintained? A requirement for the formation of an α -helix is a phenomenon named “capping” that we address below.

3 | WHAT IS CAPPING?

The 1950s was the golden age for structural biology. Pauling et al. proposed the secondary structural motifs of proteins:¹⁹ α -helices and β -sheets. The first one was confirmed when Kendrew et al. reported the three-dimensional X-ray structure of myoglobin, which was the first resolved protein.²⁰ Since then, this research area has grown significantly, revealing all kind of complex domains, combining α -helices, β -sheets, turns, and other components. As its name indicates, α -helices have spiral geometry, forming a right-handed helix. β -sheets, instead, are an extended segment of the polypeptide chains, linked by turns kept together by a network of interactions between parallel or antiparallel strands.

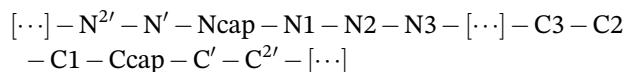
How those complex forms are maintained is of utmost importance. For example, an α -helix is stable due to the main-chain hydrogen bonds (H-bond) formed by an amino acid and the fourth residue located upstream and downstream in the helix. However, these characteristic H-bonds are missing at the helix termini, and thus, some other kinds of interactions are required. Stabilization of the terminal regions of α -helices occurs by a process named capping. Although capping was first recognized 70 years ago, there are still many open questions.

As mentioned above, α -helices are characterized by consecutive main-chain $i \rightarrow i + 4$ H-bonds between each amide hydrogen and a carbonyl oxygen from the adjacent helical turn.²¹ The repetitive turns have canonical values for their backbone dihedrals ($\varphi = -64^\circ \pm 7^\circ$; $\psi = -41^\circ \pm 7^\circ$), which allow these interactions. However, the termini of the helix cannot be stabilized because no helix turn follows, and the required H-bonds are missing. Presta and Rose observed different H-bond patterns capable

of satisfying backbone amide hydrogen and carbonyl oxygen groups in these terminal turns,²² that were named “helix capping.”²³ Interestingly, some amino acids have particular positional preference at the helix boundary, forming clusters of residues, known as capping domains, which initiate helix formation during protein folding.²⁴ Indeed, several folding disorders, such as spongiform encephalopathies²⁵ or diabetes mellitus produced by a misfolded transcription factor,²⁶ have been described as capping defects.

Because the first and last turns of a helix cannot form the characteristic main-chain $i \rightarrow i + 4$ H-bonds, capping stabilizes the boundaries of the α -helices, providing alternative ways to satisfy the required interactions. Probably, as a direct consequence, isolated α -helices are marginally stable structures and unfold readily.^{27,28} In addition, capping can be recognized as an early event during folding of barnase²⁹ and lysozyme.³⁰

The nomenclature of capping residues is as follows:



where N1–N2–N3–[...]-C3–C2–C1 participate in the helix backbone hydrogen-bonding network and have helical backbone dihedral angles. Ncap and Ccap residues are the first (amino termini) and the last residue (carboxyl termini), respectively, which participate in the hydrogen-bonding network. Ncap and Ccap do not adopt dihedral angles compatible with α -helix fold. Residues labeled with one or more apostrophes neither participate in helix backbone H-bonding nor have helical dihedral angles. Thus, they are turns out of the helical region.²²

Multiple strategies are possible for helix capping. Whereas most of the capping interactions involve backbone-to-backbone (BB-BB) H-bonds, approximately one-third of the interactions setting caps are side-chain-to-backbone H-bonds. In fact, side-chain capping is more common at Ncap than at Ccap, since the helix geometry predisposes side chains towards upstream backbone positions, promoting the H-bonds with the amide hydrogen of the first helical turn.³¹

H-bonds are only one of many types of stabilizing interactions. Hydrophobic interactions also stabilize caps.³² Predominantly, residues nearby in sequence provide hydrophobic caps and they often form hydrophobic clusters. Aurora and Rose redefined, in 1998, the phenomenon of helix capping to include both H-bonding and hydrophobic interactions. Their classification was based on the structural observation of 1,316 helices, which supports that at least there was one short- or mid-range interaction (interactions between residues separated by 3–4 or 4–7 amino acids, respectively, in the

Ncap/Ccap sequence) at the Ncap in 81% of cases and at the Ccap in 87% of cases.³¹

Salt bridges between side chains of opposite charge could replace hydrophobic capping due to a positional preference for charged polar residues (e.g., a high-normalized positional preference for Asp at N' and Arg or Lys at N4). Still, some positions occupied by Lys or Arg may function as hydrophobic residues because their long and flexible alkyl side-chain moieties can interact with buried apolar surfaces.³¹

The solvent plays a major role in folding events, interacting with the helices and affecting their stability. The dynamics of folding and unfolding might involve many water molecules, which have different anchoring sites with different helix intermediates.³³ Water molecules bind to helices, either externally to the backbone carbonyl O atom or internally by tilting open the helix H-bond and lodging between the backbone carbonyl O atom and the amide group. When the distance between donor and acceptor atoms is long, a water molecule, acting as a bridge, can stabilize the structure enabling the formation of two H-bonds within the two residues.^{34,35}

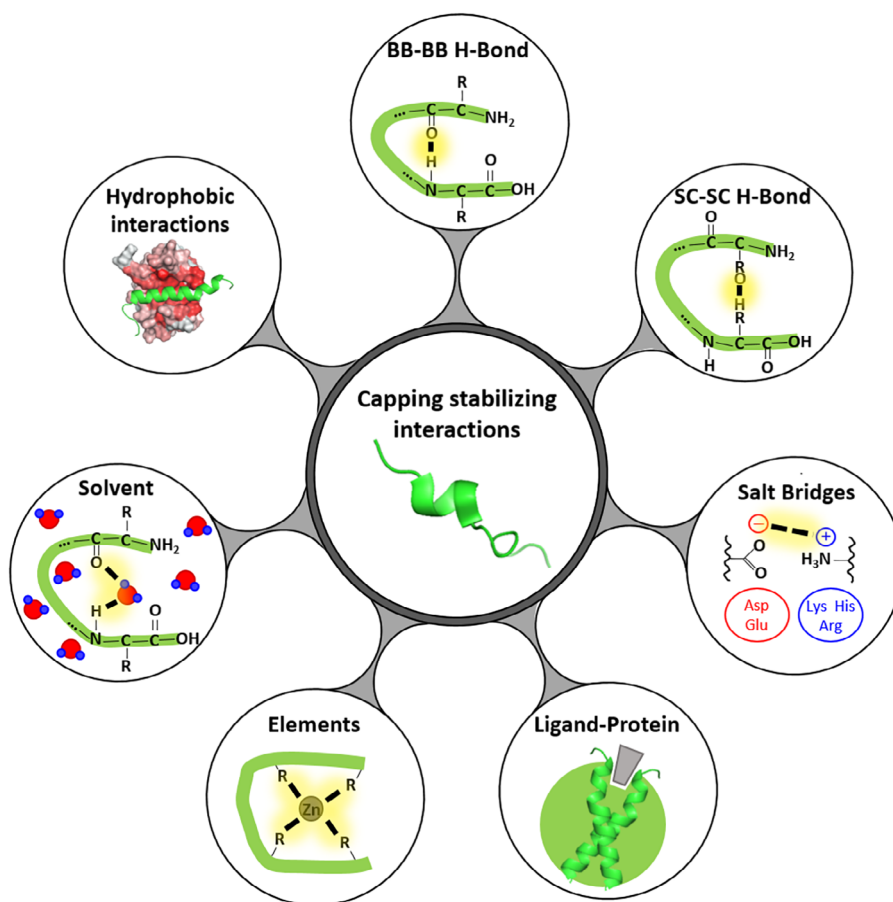
Intermolecular interactions might be also relevant for helix capping and stability. Ligand–protein interactions are of paramount importance in stabilizing the buried isolated charges and in electrostatic interactions of diverse nature for L-Arabinose-, sulfate-, and leucine/isoleucine/valine binding proteins.³⁶ Ligand–protein interaction might be exploited for pharmacological applications. For example, Imidazolium IL 1-butyl-3-methylimidazolium chloride induces the inactivation of the endocellulase 1 from *Acidothermus cellulolyticus*, destabilizing capping domains and propagating the unfolding process in an irreversible way.³⁷

Another reported intermolecular capping case is the Zn^{2+} -protein capping in a coiled-coil model domain.³⁸ This model protein is rich in Ala residues. Because the –CH₃ side chain of the alanine core has relatively low hydrophobicity, the peptide has low inherent thermodynamic stability, and is unfolded in aqueous solution. In the presence of Zn^{2+} , the protein adopts a coiled-coil trimer configuration, because Zn^{2+} forms three N-capping domains (Figure 3).

4 | CAPPING MOTIFS

A great effort has been made to organize capping motifs depending on either sequence or structural patterns. Seven patterns were identified within a large data set,³¹ which included some already known as “big box” or “Schellman motif”. The systematic nomenclature of the patterns is based on the shortest hydrophobic interaction at the helix terminus. Hydrophobic interactions between residues at positions A and B are written as $A \rightarrow B$, where

FIGURE 3 Diagram of stabilizing interactions involved in capping. Caps can be stabilized by different interactions shown in this figure. BB–BB, backbone–backbone interactions; SC–SC, side-chain–side-chain interactions



the arrow points from the hydrophobic non-helical residue to the hydrophobic residue within the helix. Several motifs are further differentiated by the presence of a particular residue located at a preferred position. Such cases are annotated by appending a slash followed by the position and its one-letter code. For example: $A \rightarrow B/C$ X means that the shortest hydrophobic interaction will be between residues at position A and B, and at position C, residue X will be the preferred one.

There are three N-capping patterns:

- | | | |
|-------|-------------------------|-------------------------|
| (I) | $N^i \rightarrow N3$ | $N^i \rightarrow N4$ |
| (II) | $N^{2i} \rightarrow N3$ | $N^{2i} \rightarrow N4$ |
| (III) | $N^{3i} \rightarrow N3$ | $N^{3i} \rightarrow N4$ |

There are four C-capping patterns:

- | | | |
|-------|-----------------------------|-----------------------------|
| (IV) | $C^{2i} \rightarrow C3/C'G$ | |
| (V) | $C^{3i} \rightarrow C3/C'n$ | $C^{4i} \rightarrow C3/C'n$ |
| (VI) | $C^{3i} \rightarrow C3/C'G$ | $C^{4i} \rightarrow C3/C'G$ |
| (VII) | $C^{4i} \rightarrow C3/C'P$ | $C^{5i} \rightarrow C3/C'P$ |

Where “n” refers to a non- β -branched residue (any residue except Val, Ile, Thr, or Pro). These patterns are

summarized in Table 1. Both the systematic and classical nomenclatures are included. Schematic representations, theoretical structures, and dihedral angles of each motif are shown in Figure 4.

Almost all N-capping domains present a preference for a proline at N1, which plays a key role (Table 1). The N1 position is well suited to the steric and chemical characteristics of proline⁴⁰ because it lacks the flexibility of other residues. Furthermore, its backbone dihedral angle is similar to that of an ideal helix.⁴¹ However, when proline adopts a helical conformation, the bulky pyrrolidine ring affects the conformation of the preceding residue, that is, forces the Ncap residue into a non-helical conformation.

Schellman and α L motifs are classic names for C-capping motifs,³² which are characterized by a glycine at C'. The other two motifs are also named Non-Glycine-Schellman and Proline motifs. These are characterized by a non- β -branched residue or a proline at C', respectively. Proline and glycine are commonly known as “helix breakers”: while proline is structurally too rigid, glycine is very flexible. The entropy effect drives a structural rearrangement from a defined secondary structure to a loop.⁴² In addition to preferences, all the C-capping motifs are able to form at least one H-bond interaction apart from having hydrophobic interactions (Figure 4).

TABLE 1 Residue preferences in capping sequences

Classic name	Nomenclature	Residue preference in N-capping positions								
		$N^{3'}$	$N^{2'}$	N'	N_{cap}	$N1$	$N2$	$N3$	$N4$	
<i>N capping motifs</i>										
I Box	$N' \rightarrow N3$			A (3.0)	D (2.8)	P (2.1)	E (2.2)	M (3.1)		
				F (2.9)	P (2.3)	K (1.8)	D (2.1)	H (2.9)		
				I (2.4)	S (1.9)	R (1.5)	A (1.4)	L (2.1)		
				L (2.1)						
	$N' \rightarrow N4$				M (4.9)	T (4.6)	E (3.8)	E (3.8)	E (7.3)	M (4.1)
					L (3.5)	S (2.9)	P (2.6)	D (2.5)	Q (5.0)	F (2.7)
					I (3.0)	D (2.4)	R (1.4)	A (1.5)	D (2.5)	V (2.4)
										I (2.4)
II Big box	$N^{2'} \rightarrow N3$		A (4.1)	D (3.1)	D (4.4)	P (3.3)	E (2.7)	M (4.4)		
			F (3.3)	G (2.9)	G (1.8)	E (2.7)	D (1.7)	F (3.6)		
			Y (2.7)	S (2.3)	T (1.8)	R (2.7)	A (1.7)	A (3.0)		
	$N^{2'} \rightarrow N4$		F (5.5)	D (2.9)	T (3.5)	P (4.6)	E (3.6)	Q (6.3)	R (4.0)	
			I (3.9)	T (2.8)	S (2.7)	A (1.6)	D (2.9)	E (5.8)	M (3.7)	
				S (2.1)	N (2.3)		S (1.5)	D (2.9)	L (2.8)	
					D (2.2)					
III β -box	$N^{3'} \rightarrow N3$	A (3.9)	T (3.1)	D (3.6)	D (4.8)	P (3.0)	E (6.5)	Y (7.6)		
		I (3.8)	N (2.7)	S (2.7)	E (2.6)	A (2.3)	K (2.3)	F (3.6)		
		F (3.6)	E (2.6)	N (2.7)	N (2.7)	E (2.6)		R (3.3)		
	$N^{3'} \rightarrow N4$	F (3.7)	E (4.5)	G (3.8)	C (4.4)	K (3.1)	Q (5.0)	E (6.2)	R (4.6)	
		Y (3.9)	R (3.4)	N (2.7)	D (2.4)	E (1.8)	D (4.9)	Q (3.8)	F (3.7)	
		I (2.6)	S (3.0)	E (2.7)	T (2.8)	I (1.8)	A (2.1)	N (3.6)	V (2.5)	
<i>C capping motifs</i>										
IV Schellman	$C^{2'} \rightarrow C3/C'G$	L (3.1)	Q (2.4)	E (3.0)	H (2.4)	G (12.0)	I (2.2)			
		A (2.9)	R (2.2)	A (3.0)	K (1.9)		V (1.8)			
		M (2.2)	K (2.2)	K (2.4)	L (1.6)		K (1.5)			
V Non-glycine Schellman	$C^{3'} \rightarrow C3/C'n$	L (2.8)	R (2.8)	R (2.8)	H (3.6)	N (4.6)	L (2.5)	L (2.7)		
		A (3.4)	Q (2.3)	E (2.8)	N (2.1)	K (3.1)	Y (1.6)	I (2.6)		
		M (2.2)	I (2.3)	K (2.4)	L (1.9)	D (2.5)		M (2.5)		
	$C^{4'} \rightarrow C3/C'n$	L (2.5)	R (3.0)	E (4.7)	F (3.2)	N (8.0)	V (2.5)	D (2.9)	I (5.3)	
		A (3.4)	Q (2.2)	K (2.1)	L (2.5)	D (3.6)	T (2.0)	T (2.5)	V (3.3)	
		R (2.0)	I (2.0)	A (1.9)	H (1.5)	K (2.0)		P (2.5)	Y (2.3)	
VI α_L	$C^{3'} \rightarrow C3/C'G$	L (2.9)	K (2.8)	R (2.8)	L (2.6)	G (12.0)	G (2.5)	L (2.7)		
		A (2.8)	L (2.3)	E (2.8)	R (2.1)		N (2.5)	I (2.4)		
		K (2.2)	N (2.1)	K (2.4)	T (1.9)		T (2.5)	M (2.2)		
	$C^{4'} \rightarrow C3/C'G$	L (3.0)	Q (3.1)	R (3.7)	T (3.2)	G (12.0)	E (4.0)	E (3.9)	V (5.3)	
		F (3.0)	L (2.8)	E (3.0)	H (2.5)		S (3.0)	S (3.5)	A (4.0)	
		K (2.9)	R (2.8)	V (2.9)	I (1.5)		D (3.0)	G (2.7)	L (2.4)	
VII Proline	$C^{4'} \rightarrow C3/C'P$	L (3.5)	K (3.5)	K (2.7)	H (4.8)	P (22.0)	E (5.3)	D (2.2)	H (2.8)	
		Y (3.0)	R (2.8)	E (2.5)	D (2.0)		Q (2.2)	T (2.1)	F (2.6)	
		Y (3.0)	E (2.2)	Q (2.2)	L (1.8)		N (2.0)	R (1.8)	L (2.0)	
	$C^{5'} \rightarrow C3/C'P$	K (3.0)	K (3.9)	E (3.0)	Y (4.2)	P (22.0)	E (4.0)	R (4.0)	K (5.3)	V (4.2)
		L (2.0)	L (1.9)	H (3.0)	C (2.5)		S (2.0)	Q (2.1)	Q (4.0)	I (3.0)
		F (2.0)	R (1.8)	F (2.7)			Q (2.0)	E (2.0)		F (2.1)

Note: Motifs are classified as N-capping motifs (top) and C-capping motifs (bottom). The classic and systematic names are indicated at the first and second columns, respectively. Residue positions in caps are indicated within the line with a grey background. The relative preference for each position is indicated in brackets. This number is computed for each i residue: $f = (\text{fraction of } i \text{ in motif}) / (\text{fraction of } i \text{ in data set})$. Adapted from Aurora and Rose.³¹

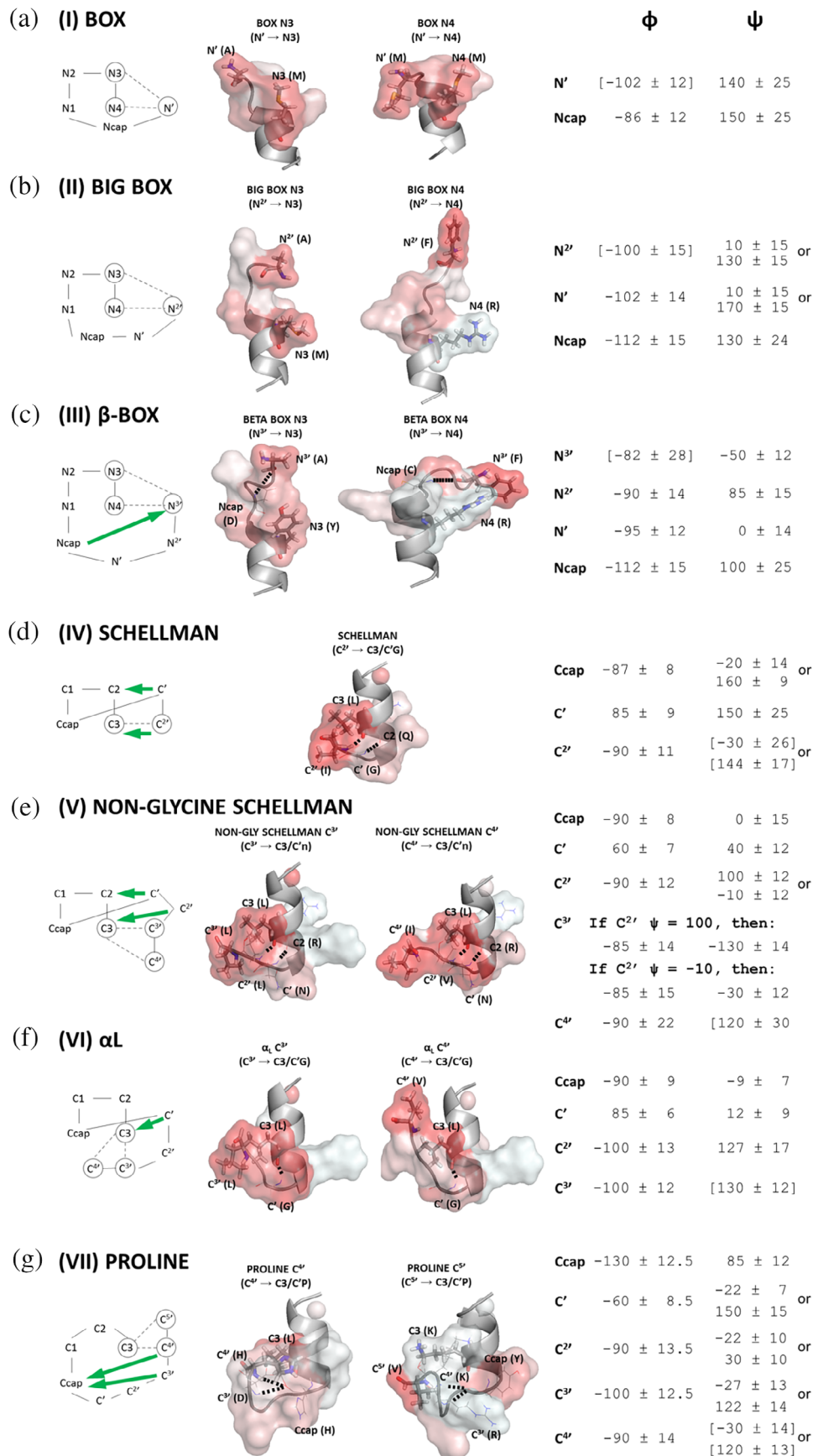


FIGURE 4 Legend on next page.

Looking at theoretical structures generated using VMD⁴³ and PyMol⁴⁴ (Figure 4, middle panel), the hydrophobicity coloring does not always match with theory. Note that some of the surfaces of the hydrophobic interacting residues (shown as sticks) are white instead of red (hydrophobic residues). This occurs for two arginine residues (for Big and Beta Box N4 motifs) and for a lysine (in C3 of Proline C^{5'} motif). However, these two amino acids need to be accurately analyzed in terms of interactions because their long alkyl chain can act as hydrophobic donors/acceptors.

The number of capping domains identified is steadily increasing. Bioinformatics has grown leading to the deployment in 2000 of an interactive and user-friendly database, CAPS-DB, that creates a large structural classification of capping domains based on geometry.⁴⁵

5 | CAPPING INTERACTIONS OF THE SOLUTION STRUCTURE OF SK2 CRD

Does the pre-folded core of the SK2 CRD or the IQ sites have properties of capping motifs? To address this question, the sequence was compared to those preferred sequences shown in Table 1. The solution NMR structure of SK2 CRD presents a structurally conserved N-capping domain, which stabilizes the core region (shown in purple in Figure 2b). In contrast, a C-capping domain could not be detected in any of the 23 models in PDB 1KKD. Although helix stability should be maintained by both N- and C-capping domains, we will focus on N-capping interactions, as they can be analyzed based on structural conservation within NMR SK2 models. The N-terminus of the core domain (Figure 4a, shown in purple) stabilizes the little helix (shown in orange), being E35 at the Ncap position. For the N-capping motifs, the only glutamic acid that appears at the Ncap position is in the β -Box, yet the rest of the sequence does not match with this CDR of

SK2. Sequence searching in CAPS-DB returns caps from already resolved structures but does not find the solution structure of SK2 (PDB: 1KKD), thus, pre-folded states of proteins are not taken into account.

Within the hydrophobic interactions, V32 interacts with W37 and L38. In other words, N^{3'} interacts with N2 and N3. There are also hydrophobic interactions between A30 (N^{5'}), L33 (N^{2'}), and V32 (N^{3'}) (Figure 5a). Therefore, V32 has the interactions required to be considered a stabilizing residue, in agreement with mutagenesis.¹⁸ H-bonds may also play a

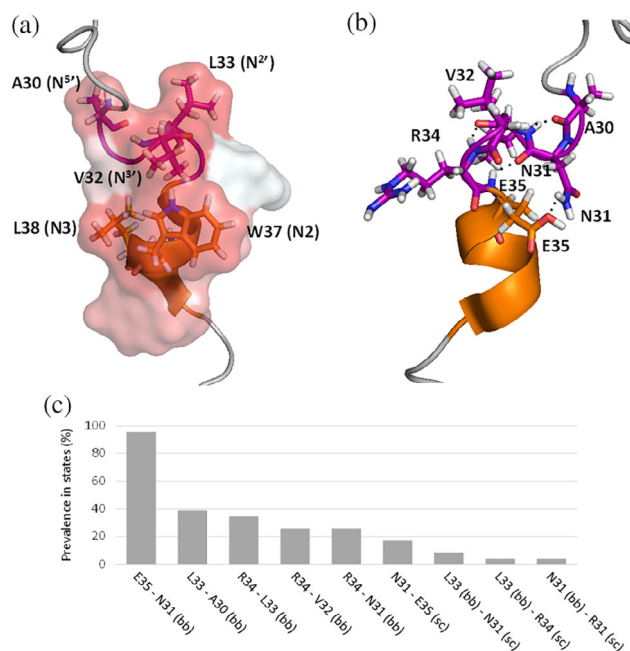


FIGURE 5 Structural analysis of the pre-folded state of SK2 CRD (PDB: 1KKD). The backbone of the protein follows the color code of Figure 4. (a) Hydrophobic residues (red surface) are labeled and the theoretical N-capping positions are indicated between brackets. (b) H-bonds are indicated with black, dashed lines (State 22). (c) Prevalence (%) of H-bonds within the 23 models of the 1KKD structure. Images were rendered using PyMol 1.3

FIGURE 4 Structural characteristics of capping motifs. N-capping motifs are (a) Box, (b) Big Box, and (c) β -Box; C-capping motifs are (d) Schellman, (e) Non-glycine Schellman, (f) α_L , and (g) Proline. The capping patterns are indicated below the name. (Left) Schematic representation of the capping domains using positional nomenclature. Black lines indicate peptide bond. Grey-dashed lines connecting two residues surrounded by circles represent hydrophobic interactions. Three circles represent the two types of hydrophobic interactions detected for that motif, for example, for the Box motif the hydrophobic interaction can occur between N' \rightarrow N3 or N' \rightarrow N4. Green arrows indicate H-bonds. (Middle) Theoretical structure created using the Molefacture Protein Builder plugin of VMD for peptide generation. Red intensity corresponds to Eisenberg hydrophobicity scale. Dihedral angles selected for each amino acid are indicated on the right panel. The preferred sequence for each motif was selected from Table 1 (first line of each motif) and a 5-residue-long peptide was added to ensure visualization of a helix in a cartoon representation. The residues contributing to capping with just an H-bond interaction are shown as lines. The interacting residues are labeled, naming the position and the one-letter-code of the residue in brackets. H-bonds are represented with black, dashed lines. (Right) Dihedral angles (Φ ; Ψ) for each non-helical position in the motif. For completeness, dihedral angles at terminal positions that are irrelevant for the conformation of the motif, are included within square brackets. This figure is adapted from Aurora and Rose.³¹ PyMol has been used for image generation. The *Color h* script³⁹ was used for hydrophobic-surface-coloring in red. Stick representation was employed for residues involved in hydrophobic interactions

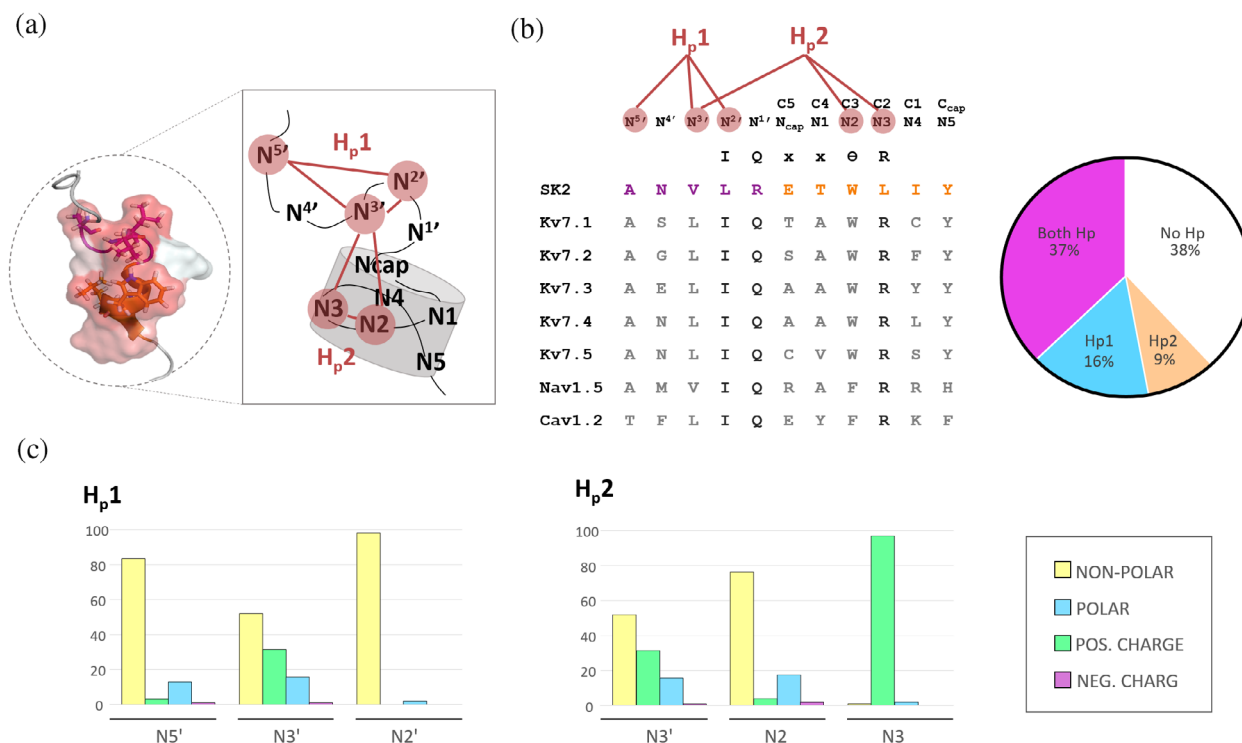


FIGURE 6 Capping within IQ motifs. (a) Schematic representation of Ncap stabilizing interactions. Inset, the structure presented in Figure 5a is shown to ease the schematic interpretation. (b) Left, example of IQ sequence alignment of some ion channels. The SK2 CRD sequence is colored as in Figure 2b. The complete alignment is shown in Table S1. Right, percentage of sequences that can form both Hp, just one of them or no one from 102 sequences (Table S1). (c) Percentage of non-polar, polar and positively and negatively charged (pos. and neg. respectively) residues for each Ncap position in the first hydrophobic core and in the second hydrophobic core (left and right, respectively). The percentage of each residue for all the Ncap positions is collected in Table S2

role in helix stabilization. Apart from those canonical α -helical H-bonds in the core zone (shown in orange in Figure 5; E35 \rightarrow I39; T36 \rightarrow Y40), there is just one H-bond, a backbone-to-backbone one, present in 22 out of the 23 models of the dynamic structure (Figure 5c). The rest of the H-bonds are present in less than 40% of the models. This BB-BB H-bond between E35 (Ncap) and N31 (N^{4'}) underscores the importance of the Ncap residue.

In short, the SK2 helix seen in solution deviates from the classic capping motifs, as its interaction pattern is different. The SK2 helix has predominantly Ncap hydrophobic interactions or cores (Hp1: N^{3'}–N2–N3 and Hp2: N^{5'}–N^{3'}–N^{2'}) (Figure 6a), and some transient H-bonds may contribute to cap stabilization.

However, the described Ncap interactions confer a static vision of the core formation. State-of-the-art molecular dynamics simulations could envision the dynamism of the pre-folded SK2 helix. To illustrate the dynamic interactions, we performed a replica exchange molecular dynamics simulation using the solution structure of the CRD of SK2 as the starting configuration (PDB 1KKD; see Supporting Information). The simulation confirms that the SK2 CRD can form a stable α -helix between E35 (Ncap) and N42, whereas the upstream residues

(N27–R34) remain mostly in a turn, coil, or bend configuration (Figure 7a), as described in the static structure. As for the intermediate- and long-range interactions stabilizing this α -helix (Figure S1b,d, respectively), most of them are of a hydrophobic nature. The residues with the greatest short-range hydrophobic contribution are mainly V32 (N^{3'}) and W37 (N2) (Figure 7b). These two residues interact with each other during the simulation stabilizing the helix, and they also interact with most of the residues of the N-terminus of the core domain, suggesting that V32 and W37 are key residues for helix stability and capping. Y40 may also play a role in helix stability since it partially interacts with some N-terminal residues such as N27, A28, A30, and N31 (Figures 7b and S1b,d). When focusing on the static hydrophobic interactions or cores described above (Hp1: N^{5'}–N^{3'}–N^{2'} and Hp2: N^{3'}–N2–N3) (Figures 5 and 6a), at a long-range interaction (9.5 Å) all the residues interact with the rest except for A30 (N^{5'}) and L38 (N3) (Figure S1e), suggesting that those two hydrophobic cores are responsible for maintaining the pre-folded α -helix. When the attention is focus on these residues forming the hydrophobic cores at a short-range distance (cutoff: 4.5 Å; Figure 7c) it can be observed that the most relevant interaction is between

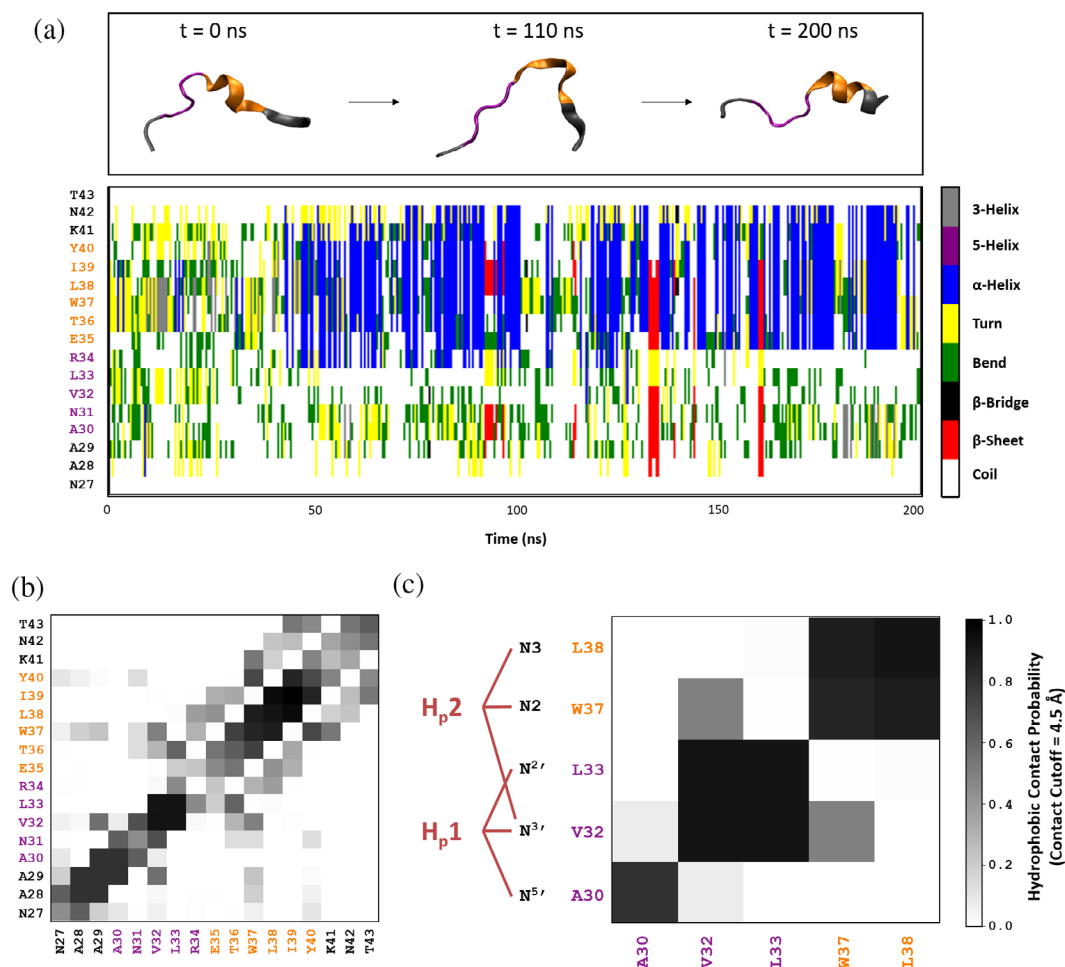


FIGURE 7 Replica exchange molecular dynamics simulation reveal stability of the pre-folded helix. (a) Secondary structure of the SK2 CRD, according to the DSSP definition, along a 200-ns Hamiltonian replica exchange molecular dynamics simulation. (b) Hydrophobic contact probabilities in the SK2 CRD α -helical conformations sampled along the simulation. (c) A more detailed view of the hydrophobic contact probabilities of the simulation among the residues involved in the hydrophobic cores (H_p) of the structure. For (b) and (c), a contact cutoff of 4.5 Å has been used (see Supporting Information, Molecular Dynamics section). In all panels, same colors as in Figure 2b are used (related to Figure S1)

the V32 and W37, which maintains together the pre-folded helix with the cap (colored in orange and in purple in Figure 2b, respectively). However, the rest of the residues involved in these hydrophobic cores may play an important role in stabilizing the helix as they do interact with residues belonging to the same hydrophobic core.

Overall, the molecular dynamics simulation illustrates the stability observed in the structure of the SK2 CRD that is maintained by two hydrophobic cores and the interaction between V32 and W37 may particularly contribute to this stability.

6 | CAN THE IQ MOTIF CONFER A PRE-FOLDED STATE?

The question that we address now is whether the IQ motifs could also confer a pre-folded state as seen in the

SK2 structure (Figure 5). Since there are no available structures of isolated peptides bearing an IQ motif in solution, sequence-based criteria have to be applied to address this question. A total of 102 IQ and IQ-like motifs were aligned to the SK2 core domain sequence, taking E35 as the Ncap residue (Table S1). The prevalence of residue type was then quantified (Figure 6). For the first hydrophobic core (H_p1 ; $N^{5'}-N^{3'}-N^{2'}$) the residue at position $N^{2'}$ for the IQ motifs is the first amino acid of this conserved motif (Ile). Ninety-eight percent of the residues at this position are hydrophobic (Figure 6c). At the $N^{5'}$ position, three positions upstream of Ile, about 83% of the residues are hydrophobic. Interestingly, the $N^{3'}$ position, involved in both H_p , is the less conserved amino acid in terms of hydrophobicity (52%). Lys and Arg are quite conserved in this position (18 and 11%, respectively; see Table S2). As the side chains of Arg and Gln have a long alkyl tail, they can form local hydrophobic

interactions. Taking these residues as partially hydrophobic, a 53% of the analyzed IQ motifs can form Hp1 (Table S2).

In the second hydrophobic core (Hp2; N^{3'}—N2—N3), 76% of the sequences have a hydrophobic residue at the N2 position, the θ residue of IQ motif (Figure 6). The N3 position is the highly conserved Arg of IQ motifs (93%) (Table S2) which could form the hydrophobic interaction with the alkyl chain.

Considering both hydrophobic cores, about 37% of the IQ domains could adopt similar capping geometry (Figure 6b, right; Table S1). About 25% of the sequences can form just one Hp, and 38% sequences cannot form any Hp. Interestingly, a hydrophobic residue is not present in the N^{3'} position, the common residue for both Hp, in 80% of the sequences that cannot form any Hp, yet they do have hydrophobic residues in the other two Hp positions.

The interaction stabilizing the pre-folded helix with the cap, between residues V32 (N^{3'}) and W37 (N2) for the SK CRD, could play a role in the stability of the IQ motifs, as 47% of them have apolar residues at those positions. The N2 position is the θ residue of the IQ motif.

Interestingly, some residues may participate within both N- and C-capping domains for the SK2 CRD. Note that the residue at position N5 is the last residue of the SK2 CRD α -helix in solution, and, therefore, it can be regarded as Ccap. For instance, the N2 and N3 positions, which are involved in N-capping, also correspond to C3 and C2, respectively (Figure 6b). As mentioned before, the different NMR models display structural diversity, opening a question mark for the role of these C3 and C2 residues in C-capping stability. Regarding the IQ sequences, there is an intriguing preponderance of arginine at position C^{2'} (66% of the cases analyzed), suggesting that this amino acid may play a particular role in stabilizing the C-capping domain.

A unique hydrophobic pattern was not evident in many IQ motifs, implying the existence of varied recognition processes, with some IQ motifs presenting pre-folded states and others not doing so. The still intriguing CaM recognition might be a crucial process for α -helix formation. The rapidly evolving structural techniques together with the increasing computational power will be crucial for resolving all the unknowns about the ability of CaM to recognize a huge diversity of seemingly unrelated sequences.

7 | CONCLUSION

It is not clear if CaM recognizes already formed α -helices (conformational selection) or if CaM induces the

formation of the helix as a part of the target recognition process. The solution structure of SK2 CRD reveals a pre-folded helix stabilized with a structurally conserved N-capping domain, suggesting that CaM probably induces the formation of the remaining α -helix when recognizing this pre-folded domain. Choosing the correct folding pathway may depend critically of the N-capping stabilizing properties imposed by the core domain, as they are the first residues that start folding as translation progressed within the ribosome. Many proteins start folding as the nascent chain move through the exit tunnel of the ribosome (reviewed in Liutkute et al.⁴⁶). CaM target recognition may occur during this process, and it may contribute to the correct folding of many targets during the co-translational synthesis.⁴⁷

In this review, we have mainly focused on the N-capping domains of the IQ motifs. The cap stabilizing this pre-folded helix is different from classical capping motifs. The pre-folded helix is stabilized mainly by two very stable hydrophobic cores (Hp1: N^{5'}—N^{3'}—N^{2'}/Hp2: N^{3'}—N2—N3), as revealed by molecular dynamics. Besides, the interaction between V32 (N^{3'}) and W37 (N2) in the SK2 CRD is important for maintaining the pre-folded helix and the cap together. The SK2 CRD is analogous to the IQ domains, raising the question if a similar recognition process is used when this CaM binding motif is present. Analysis of the several IQ motifs reveals heterogeneity regarding capping properties, and about 37% could form two hydrophobic interactions adequate for a pre-folded SK2-like recognition mechanism. On the other hand, the majority of IQ sites are likely to be unfolded, and some sort of mutual structural induction between the target and CaM is likely involved.

We are far from understanding the ability of CaM to recognize such a diverse array of targets, and current methods fail to identify all existing targets. Solving the structure of IQ domains in solution, extended molecular dynamics, such as replica exchange, and the use of machine learning techniques encompassing the properties of different CaM binding domains are three possible avenues to reveal the secrets of the mysterious CaM-target recognition process.

ACKNOWLEDGMENTS

The Government of the Autonomous Community of the Basque Country (IT1165-19 and KK-2020/00110) and the Spanish Ministry of Science and Innovation (RTI2018-097839-B-100 to A.V. and PID2019-105488GB-I00 to A.B., A.L., and O.R.B.) and FEDER funds provided financial support for this work. A.M.-M. is supported by predoctoral contracts from the Basque Government administered by University of the Basque Country.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Arantza Muguruza-Montero: Project administration (equal); writing – original draft (lead); writing – review and editing (lead). **Rafael Ramis:** Investigation (equal); resources (supporting); writing – review and editing (supporting). **Eider Nuñez:** Writing – review and editing (equal). **Oscar R. Ballesteros:** Writing – review and editing (supporting). **Markel G. Ibarluzea:** Writing – review and editing (supporting). **Ariane Araujo:** Writing – review and editing (supporting). **Sara M-Alicante:** Writing – review and editing (supporting). **Janire Urrutia:** Writing – review and editing (equal). **Aritz Leonardo:** Writing – review and editing (supporting). **Aitor Bergara:** Writing – review and editing (supporting). **Alvaro Villarroel:** Conceptualization (equal); project administration (equal); supervision (lead); writing – original draft (supporting); writing – review and editing (equal).

ORCID

Arantza Muguruza-Montero  <https://orcid.org/0000-0001-8713-4949>

REFERENCES

1. Finn BE, Forsén S. The evolving model of calmodulin structure, function and activation. *Structure*. 1995;3:7–11.
2. Ikura M, Ames JB. Genetic polymorphism and protein conformational plasticity in the calmodulin superfamily: Two ways to promote multifunctionality. *Proc Natl Acad Sci*. 2006;103:1159–1164.
3. Nelson MR, Chazin WJ. An interaction-based analysis of calcium-induced conformational changes in Ca^{2+} sensor proteins. *Protein Sci*. 1998;7:270–282.
4. Rhoads AR, Friedberg F. Sequence motifs for calmodulin recognition. *FASEB J*. 1997;11:331–340.
5. Villarroel A, Tagliatalata M, Bernardo-Seisdedos G, et al. (2014) The ever changing moods of calmodulin: How structural plasticity entails transductional adaptability. *J Mol Biol* 426:2717–2735.
6. Osawa M, Tokumitsu H, Swindells MB, et al. A novel target recognition revealed by calmodulin in complex with Ca^{2+} -calmodulin-dependent kinase kinase. *Nat Struct Biol*. 1999;6:819–824.
7. Yap KL, Kim J, Truong K, Sherman M, Yuan T, Ikura M. Calmodulin target database. *J Struct Funct Genomics*. 2000;1:8–14.
8. Hovey L, Fowler CA, Mahling R, et al. Calcium triggers reversal of calmodulin on nested anti-parallel sites in the IQ motif of the neuronal voltage-dependent sodium channel NaV1.2. *Biophys Chem*. 2017;224:1–19.
9. Ataman ZA, Gakhar L, Sorensen BR, Hell JW, Shea MA. The NMDA receptor NR1 C1 region bound to calmodulin: Structural insights into functional differences between homologous domains. *Structure*. 2007;15:1603–1617.
10. Tokimasa T. Intracellular Ca^{2+} -ions inactivate K^{+} -current in bullfrog sympathetic neurons. *Brain Res*. 1985;337:386–391.
11. Kirkwood A, Simmons MA, Mather RJ, Lisman J. Muscarinic suppression of the M-current is mediated by a rise in internal Ca^{2+} concentration. *Neuron*. 1991;6:1009–1014.
12. Ikeda SR, Kammermeier PJ. M current mystery messenger revealed? *Neuron*. 2002;35:411–412.
13. Lee C-H, MacKinnon R. Activation mechanism of a human SK-calmodulin channel complex elucidated by cryo-EM structures. *Science*. 2018;360:508–513.
14. Alexander KA, Wakim BT, Doyle GS, Walsh KA, Storm DR. Identification and characterization of the calmodulin-binding domain of neuromodulin, a neurospecific calmodulin-binding protein. *J Biol Chem*. 1988;263:7544–7549.
15. Cheney RE, Mooseker MS. Unconventional myosins. *Curr Opin Cell Biol*. 1992;4:27–35.
16. Tidow H, Nissen P. Structural diversity of calmodulin binding to its target sites. *FEBS J*. 2013;280:5551–5565.
17. Urrutia J, Aguado A, Muguruza-Montero A, et al. The crossroad of ion channels and calmodulin in disease. *Int J Mol Sci*. 2019;20:400.
18. Wissmann R, Bildl W, Neumann H, et al. A helical region in the C terminus of small-conductance Ca^{2+} -activated K^{+} channels controls assembly with Apo-calmodulin. *J Biol Chem*. 2002;277:4558–4564.
19. Pauling L, Corey RB, Branson HR. The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. *Proc Natl Acad Sci U S A*. 1951;37:205–211.
20. Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature*. 1958;181:662–666.
21. Pauling L, Corey RB. Configurations of polypeptide chains with favored orientations around single bonds: Two new pleated sheets. *Proc Natl Acad Sci U S A*. 1951;37:729–740.
22. Presta L, Rose G. Helix signals in proteins. *Science*. 1988;240:1632–1641.
23. Richardson J, Richardson D. Amino acid preferences for specific locations at the ends of alpha helices. *Science*. 1988;240:1648–1652.
24. Koscielska-Kasprzak K, Cierpicki T, Otlewski J. Importance of α -helix N-capping motif in stabilization of $\beta\beta\alpha$ fold. *Protein Sci*. 2003;12:1283–1289.
25. Iovino M, Falconi M, Petruzzelli R, Desideri A. Role of the helix capping in the stability of the mouse prion (180-213) segment: Investigation through molecular dynamics simulations. *J Biomol Struct Dyn*. 2001;19:237–246.
26. Narayana N, Phillips NB, Xin HQ, Jia W, Weiss MA. Diabetes mellitus due to misfolding of a β -cell transcription factor: Stereospecific frustration of a Schellman motif in HNF-1 α . *J Mol Biol*. 2006;362:414–429.
27. Marqusee S, Baldwin RL. Helix stabilization by Glu...Lys⁺ salt bridges in short peptides of de novo design. *Proc Natl Acad Sci U S A*. 1987;84:8898–8902.
28. Lyu P, Liff M, Marky L, Kallenbach N. Side chain contributions to the stability of alpha-helical structure in peptides. *Science*. 1990;250:669–673.
29. Serrano L, Matouschek A, Fersht AR. The folding of an enzyme. *J Mol Biol*. 1992;224:847–859.

30. Radford SE, Dobson CM, Evans PA. The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature*. 1992;358:302–307.
31. Aurora R, Rose GD. Helix capping. *Protein Sci*. 1998;7:21–38.
32. Aurora R, Srinivasan R, Rose GD. Rules for α -helix termination by glycine. *Science*. 1994;264:1126–1130.
33. Sundaralingam M, Sekharudu YC. Water-inserted α -helical segments implicate reverse turns as folding intermediates. *Science*. 1989;244:1333–1337.
34. Karle IL, Flippen-Anderson J, Uma K, Balaram P. Aqueous channels within apolar peptide aggregates: Solvated helix of the alpha-aminoisobutyric acid (Aib)-containing peptide Boc-(Aib-ala-Leu)₃-Aib-OMe.2H₂O.CH₃OH in crystals. *Proc Natl Acad Sci U S A*. 1988;85:299–303.
35. Fisher BF, Guo L, Dolinar BS, Guzei IA, Gellman SH. Heterogeneous H-bonding in a foldamer helix. *J Am Chem Soc*. 2015;137:6484–6487.
36. Quioco FA, Sack JS, Vyas NK. Stabilization of charges on isolated ionic groups sequestered in proteins by polarized peptide units. *Nature*. 1987;329:561–564.
37. Summers SR, Sprenger KG, Pfaendtner J, Marchant J, Summers MF, Kaar JL. Mechanism of competitive inhibition and destabilization of *Acidothermus cellulolyticus* endoglucanase 1 by ionic liquids. *J Phys Chem B*. 2017;121:10793–10803.
38. Liu J, Dai J, Lu M. Zinc-mediated helix capping in a triple-helical protein. *Biochemistry*. 2003;42:5657–5664.
39. Eisenberg D, Schwarz E, Komaromy M, Wall R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J Mol Biol*. 1984;179:125–142.
40. Yun RH, Anderson A, Hermans J. Proline in α -helix: Stability and conformation studied by dynamics simulation. *Proteins Struct Funct Genet*. 1991;10:219–228.
41. Schulz GE, Schirmer RH principles of protein structure. New York: Springer Verlag, 1979.
42. Imai K, Mitaku S. Mechanisms of secondary structure breakers in soluble proteins. *Biophysics (Oxford)*. 2005;1:55–65.
43. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. *J Mol Graph*. 1996;14(27–28):33–38.
44. Anon. The PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC.
45. Segura J, Oliva B, Fernandez-Fuentes N. CAPS-DB: A structural classification of helix-capping motifs. *Nucleic Acids Res*. 2012;40:479–485.
46. Liutkute M, Samatova E, Rodnina MV. Cotranslational folding of proteins on the ribosome. *Biomolecules*. 2020;10:97.
47. Urrutia J, Aguado A, Gomis-Perez C, et al. An epilepsy-causing mutation leads to co-translational misfolding of the Kv7.2 channel. *BMC Biol*. 2021;19:109.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Muguruza-Montero A, Ramis R, Nuñez E, et al. Do calmodulin binding IQ motifs have built-in capping domains? *Protein Science*. 2021;30:2029–41. <https://doi.org/10.1002/pro.4170>