Calmodulin and Myosin Light Chain Kinase: How Helices Are Bent

Robert H. Kretsinger

Protein phosphorylation is one of the most common posttranslational modifications and the final step in many cell signaling pathways. All eukaryotic kinases share a homologous catalytic core. Knighton *et al.* (*1*) present a model of the COOH-terminal catalytic domain of smooth muscle myosin light chain kinase (smMLCK) based on its inferred similarity with cyclic adenosine monophosphate–dependent protein kinase (cAPK), for which the crystal structure has been determined (*2*). Knighton *et al.* have used what is becoming an accepted procedure: Place homologous amino acids of the model, smMLCK, in the positions of their homologs in the known structure, cAPK. Confirm that insertions or deletions occur at loops or turns between fixed elements as α helices or strands of β sheets. Confirm that buried side chains in the model are hydrophobic. Proceed to energy minimization. A healthy skepticism should help refine this procedure, not reject it.

Enzymes that have polymers for substrates have clefts bounded by two (or three) domains. The first two domains of the kinase, residues 513 to 686 in smMLCK numbering, form the top part of the catalytic core; residues 513 to 774 form the bottom half. Then the fun begins.

Both smMLCK, as well as its close homolog, skeletal muscle (sk) MLCK, have a COOH-terminal peptide at residues 774 to 813 with three overlapping functional areas: connecting peptide 774 to 788, pseudosubstrate 787 to 807, and calmodulin (CaM)-binding region 796 to 813. The connecting peptide is long enough to allow the pseudosubstrate to lie in the catalytic cleft as does the inhibitor peptide, PKI(5-24), seen in the crystal structure of cAPK.

Binding of calcium-CaM (CaCaM) to the CaM binding region pries the pseudosubstrate out of the cleft and opens smMLCK to phosphorylate a serine of MLC (Fig. 1). Proteolytic cleavage of the connecting peptide achieves the same activation in vitro without resort to CaM.

Most schemes of intracellular signaling also involve Ca and CaM or one of its 30 homologs. CaM, which is thought to exist in every eukaryotic cell, consists of two lobes. Both contain a pair of helix-loop-helix Ca-binding domains known as EF-hands. The crystal structures of CaM and its close homolog troponin C (TnC) revealed that domains 1 and 2 are related by an approximate twofold axis, as are domains 3 and 4. The second (F) helix of domain 2, eight interdomain (or linker) residues, and the first (E) helix of domain 3 form a continuous 28-residue α helix (Fig. 2A). TnC (Fig. 2B) has 11 residues in its linker joining domains 2 and 3. It has the same dumbbell shape as does CaM but is 4.5 Å (3 × 1.5) longer. Two (CaM) or three (TnC) turns of α helix, the linker regions, are completely exposed to solvent. Most α helices, be they in coiled coils or in globular proteins, are stabilized by lateral contact with other parts of the protein. In its Ca-bound form, CaM activates a score of different enzymes or structural proteins.

This linker helix and the promiscuous interactions of CaM pose the fundamental question: How does CaM activate so many different targets? Persechini and Kretsinger (*3*) proposed that the ". . . linker region of the central helix of calmodulin functions as a flexible tether . . ." permitting the two lobes to enfold an α helix of the target, thereby removing the self-inhibition posed by this pseudosubstrate (Fig. 2C).

The solution structure of CaM complexed with the target peptide of skMLCK has recently been determined by multidimensional nuclear magnetic resonance (NMR) spectroscopy (*4*) (Fig. 2D). Meador *et al.* (*5*) described the crystal structure of CaM complexed with the homologous α-helical peptide of smMLCK (Fig. 2E). Four CaM-peptide complexes are present in the asymmetric unit. One infers that all four are nearly identical, although no comparison is offered. This inferred identity argues against this enfolding structure arising from constraints of crystal packing. The similarity of the model, the NMR structure, and the crystal structure is apparent.

The agreement between the two CaM structures attests to the development of multidimensional spectroscopy and the comple-

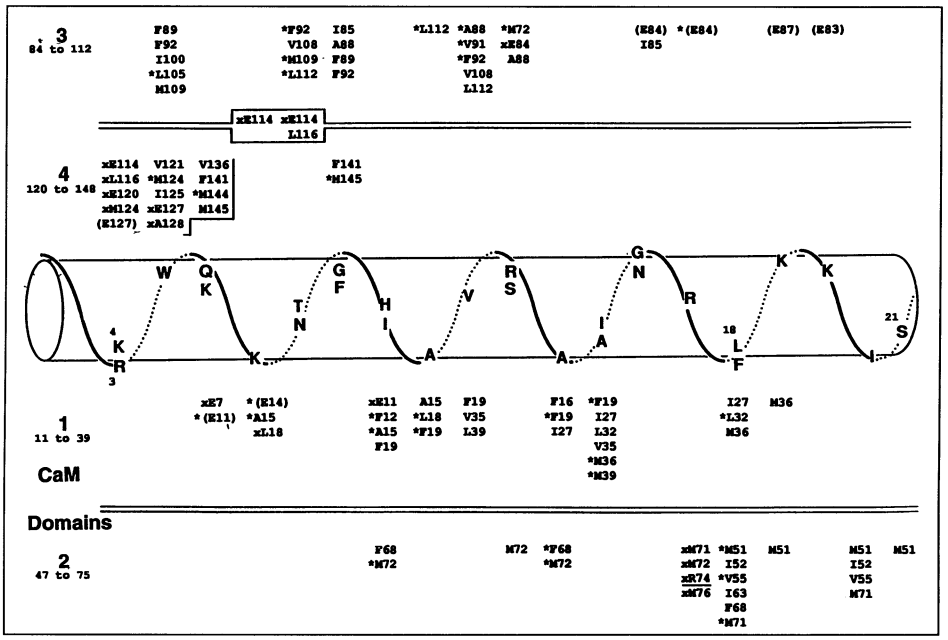

Fig. 1. Many interactions. Summary of CaM-target interactions redrawn from Ikura *et al.* (*4*). In the NMR structure, the peptide (skMLCK sequence 577 to 602) is bound by CaM; in the crystal structure the peptide (smMLCK sequence 796 to 815) is bound. The 26-residue skMLCK analog is helical from Arg^3 to Ser^{21}; the 20-residue smMLCK analog is helical from Lys^4 to Leu^{18}. Both analog sequences are shown, aligned by homology; where identical only one residue is shown. Interactions with CaM side chains are indicated: "*" indicates interactions reported for both solution and crystal complexes; "x" indicates interactions for only the crystal structure; no symbol indicates interactions observed for only the solution structure; residues in parentheses indicate interactions inferred but not observed by NMR. CaM side chains are indicated by domain; residues Glu^{114} and Leu^{116} lie between domains 3 and 4. Arg^{74} interacts with the Arg common to both peptides by van der Waals contact. As the pseudosubstrate, analog peptides are aligned with the natural substrate of MLCK, regulatory light chain (RLC) of myosin; Ser^{19} of RLC aligns with I9/H10 of the peptides.

The author is in the Department of Biology, University of Virginia, Charlottesville, VA 22901.

mentarity of the two disciplines as well as to the similarity of the two structures. The differences receive disproportionate emphasis simply because more space is required for their enumeration. In general, more NMR interactions are listed in Fig. 1, which may reflect a greater mobility and hence range of structures in solution or that some nuclear Overhauser effects (NOEs) may be observed at greater distance than the <4 Å cutoff for the crystal tabulation. More crystal contacts are cited for the first residue in helical conformation (smArg/skLys) and for the Arg common to both. This and the better definition of the linker region of CaM may reflect greater stability imposed by crystal packing. Part of the reason for the retention of function by CaM mutants having deletions within their linkers, Met^{76}-Glu^{83}, is that only two interactions of the linker with the target are cited and those involve residues 76 and 83. Ikura *et al.* acknowledged that the linker is a region of high mobility and that resonance and distance assignments here were tenuous. One anticipates that the temperature factors of the linker region in the crystal structure would be high. Of greater concern is that the relative positions of lobe 1,2 and lobe 3,4 differ slightly in the two structures. If the peptide helices in the two structures have (nearly) identical structures, then any relative displacement of the two lobes implies different interactions of CaM and peptide. The listed interactions reflect little difference in the positioning of the two lobes of CaM relative to the sm or the sk helices. Conversely, if the NMR structure has a small error in assignment of helical parameters of the peptide, the relative positions of the two lobes would be expected to twist with the target helix. The crystal structure should be especially reliable in showing the spatial relationships of components over long distances. The NMR structure will detect interactions in the range of 3 to 6 Å; long-distance relationships might reflect cumulative errors from adding many short-range interactions.

What of this linker helix? Prior to the crystal structure determinations of CaM and of TnC, the Kretsinger and Barry model (6) of a globular molecule with lobe 1,2 and lobe 3,4 resembling the CD domain and EF domain of parvalbumin was generally accepted. In this model, the linker was bent to allow the two hydrophobic faces of the two lobes to contact one another. After the crystal structure determinations, most experiments done in solution, including small-angle x-ray scattering and NMR, were interpreted in terms of a dumbbell-shaped structure. Many of the experiments that led Persechini and Kretsinger to propose the flexible tether model were based on mutant CaMs in which one to four residues were deleted from the linker region. These mutant CaMs are inferred to be dumbbell-shaped in solution as is the native. They activate several target enzymes despite the rotation of lobe 1,2 relative to lobe 3,4. It was assumed that there is a dynamic equilibrium between the straight and the bent forms of CaM and that binding to the target drives the distribution to the bent form by mass law. The crystal structure of CaM lacking Glu^{84} (Fig. 2F) (7), in the absence of a target, is bent in a similar manner to CaM in complex with the MLCK peptide.

We now have a plausible model of how the pseudosubstrate of smMLCK fits into the cleft of its own active site and the structure of CaM bound to a peptide representing the COOH-terminal half the pseudosubstrate. The value of a scientific discovery is often judged not by the questions it answers but by those that it poses.

What is the structure of the entire MLCK-CaM complex and of the other CaM-target complexes? We have seen one grip of four EF-hands. If the pseudosubstrate is pried out of its blocking position, what is the fulcrum? How do the surfaces of CaM and rest of MLCK interact? Is the flexible tether mechanism applicable to CaM in its interactions with a score of other enzymes and structural proteins?

There are 22 known subfamilies having four EF-hand domains, one each having five, six, or eight domains. Functions are known for only ten of the 22, none of the three. To what extent is the flexible tether paradigm applicable to these 25 and to the numerous others that will be discovered?

Returning to basics, how is stabilization and destabilization imparted from the two lobes of CaM to the linker region of its central helix? Wherein lies the flexibility?

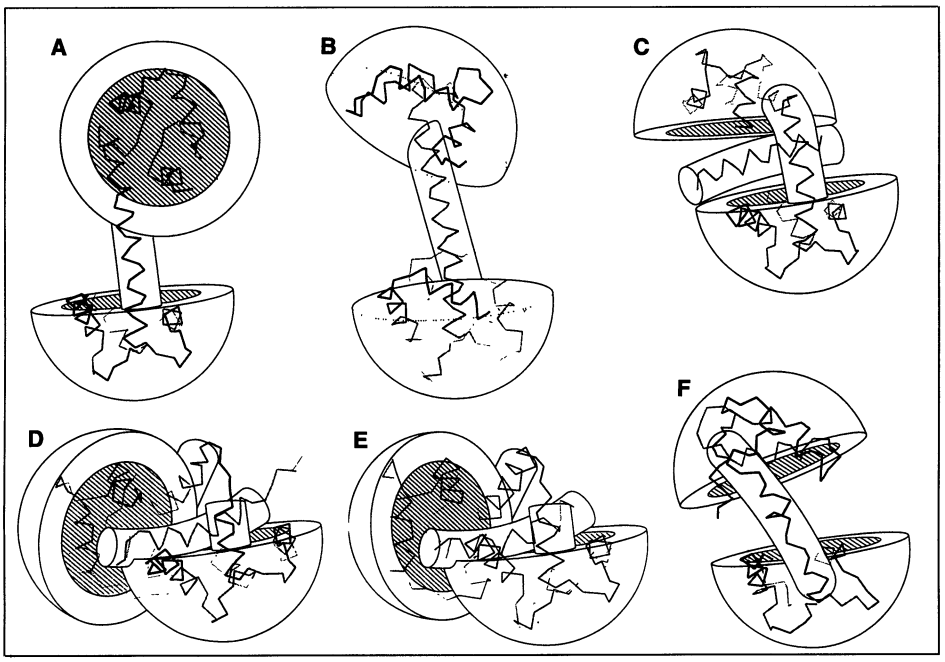

Fig. 2. Variations on a theme. The cartoons show the relative positions of lobe 1,2, of linker, and of lobe 3,4. The hemisphere of each cup, which symbolizes a lobe has an axis of rotational symmetry coincident with the approximate twofold rotation axis relating domains 1 and 2 or with the axis of 3 and 4. The hatched portion of the surface represents the hydrophobic patch, which is on the opposite side and 15 Å from the two calcium binding loops on the other surface. The handle of ladle 1,2 is almost perpendicular to the ladle face and consists of the COOH-terminus of helix F2 and the NH_2-terminal part of the linker. The handle of pot 3,4 is almost parallel to the pot face and consists of the COOH-part of the linker and the F-terminus of helix E3. In the crystal structures of TnC and CaM, the two handles are continuous and the linker region of the central helix is represented by a rod. The target helix of MLCKs to which CaM is bound is also a rod. Crystal structure of (**A**) CaM and (**B**) TnC. (**C**) Predicted structure of CaM bound to a target helix. (**D**) Solution NMR structure of CaM bound to skMLCK peptide. Crystal structures of (**E**) CaM bound to smMLCK peptide and (**F**) CaM lacking Glu^{84}

REFERENCES AND NOTES

1. D. R. Knighton *et al.*, *Science* **258**, 130 (1992).
2. D. R. Knighton *et al.*, *ibid.* **253**, 407 (1991); D. R. Knighton *et al.*, *ibid.*, p. 414.
3. A. Persechini and R. H. Kretsinger, *J. Biol. Chem.* **263**, 12175 (1988).
4. M. Ikura *et al.*, *Science* **256**, 632 (1992).
5. W. E. Meador, A. R. Means, F. A. Quiocho, *ibid.* **257**, 1251 (1992).
6. R. H. Kretsinger and C. D. Barry, *Biochim. Biophys. Acta* **405**, 40 (1975).
7. S. Raghunathan *et al.*, unpublished results.
8. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.