Structural Basis of the Intrasteric Regulation of Myosin Light Chain Kinases

Daniel R. Knighton, Richard B. Pearson, Janusz M. Sowadski, Anthony R. Means, Lynn F. Ten Eyck, Susan S. Taylor, Bruce E. Kemp*

The smooth muscle myosin light chain kinase (smMLCK) catalytic core was modeled by using the crystallographic coordinates of the cyclic AMP–dependent protein kinase catalytic subunit (cAPK) and a bound pseudosubstrate inhibitor peptide, PKI(5-24). Despite only 30% identity in amino acid sequence, the MLCK sequence can be readily accommodated in this structure. With the exception of the short B-helix, all major elements of secondary structure in the core are very likely conserved. The active site of the modeled MLCK complements the known requirements for peptide substrate recognition. MLCK contains a pseudosubstrate sequence that overlaps the calmodulin binding domain and has been proposed to act as an intrasteric inhibitor and occupy the substrate binding site in the absence of Ca^{2+}-calmodulin. The pseudosubstrate sequence can be modeled easily into the entire backbone of PKI(5-24). The results demonstrate that the intrasteric model for regulation of MLCK by intramolecular competitive inhibition is structurally plausible.

Myosin light chain kinases (MLCKs), present in all eukaryotic cells, are members of the family of Ca^{2+}-calmodulin (Ca-CaM)–dependent protein kinases (*1*). In smooth muscle, MLCK (smMLCK) regulates the initiation of contraction. In the absence of Ca-CaM, MLCK is thought to be maintained in a latent form by an internal pseudosubstrate sequence bound within the active site (*2*). The binding of Ca-CaM to MLCK releases the autoinhibition and increases the enzyme's activity, predominantly by increasing the maximum rate (V_{max}) (*3*). The pseudosubstrate sequence of smMLCK overlaps the CaM binding region (Fig. 1) and contains a substrate-like sequence homologous to the recognition motif present in the myosin light chains (*4*). Synthetic peptides corresponding to the pseudosubstrate sequence act as potent inhibitors of MLCK in the presence of Ca-CaM (*2, 4*), and removal of the pseudosubstrate region by limited proteolysis rendered the enzyme constitutively active (*5, 6*). Proteolysis or truncation by mutagenesis within the CaM binding domain, leaving the pseudosubstrate region intact, generates an irreversibly inhibited enzyme (*6, 7*). These results, however, do not provide direct evidence that the pseudosubstrate sequence actually occupies the active site as originally proposed (*4*). Intramolecular competitive inhibition by the pseudosubstrate sequence is a central feature of the intrasteric model for the Ca-CaM–regulated protein kinases (*2*). Intrasteric regulation involves an autoinhibitory region that resembles the enzyme's substrate and is directed at the active site in contrast to an allosteric regulator that has a dissimilar structure and acts at a distinct site (*2*).

All of the eukaryotic protein kinases, including MLCK, share sequence similarities that extend throughout a conserved catalytic core (*8*). The crystal structure of the catalytic subunit of cAMP-dependent protein kinase (cAPK) with a bound 20-residue inhibitor peptide was recently solved at 2.7 Å resolution (*9, 10*). We have modeled both the smMLCK and skeletal muscle MLCK (skMLCK) using the crystal structure of cAPK as a template (*9, 10*). We evaluate whether the MLCK catalytic core residues can be accommodated in the coordinates of the catalytic subunit of cAPK and whether the proposed intrasteric model for regulation of smMLCK is structurally plausible.

The secondary structure of the catalytic subunit of cAPK (cat-cAPK) together with the modeled core of smMLCK are shown in Fig. 2 along with the color-coded sequence alignments between cAPK residues and the MLCKs (Fig. 3). In the aligned sequence of 261 cAPK residues, cAPK and smMLCK share 30% identity (conservative substitutions greatly increase the similarity). There is 46% identity between smMLCK and skMLCK. The sequence similarities extend over the full length of the catalytic core. Eleven highly conserved residues play essential roles in ATP (adenosine triphosphate) binding and catalysis and provide multiple, structurally conserved, fixed points of reference. The even distribution of invariant and essential residues and the sequence similarities throughout the core ensure that the overall architecture of the catalytic cores is similar. We assumed that all non–insert-deletion sequences were structurally conserved. Details of the modeling of the smMLCK core are given in (*11*), and the plausibility of the resulting smMLCK model was independently confirmed by three-dimensional (3-D) profile analysis (*12, 13*). Although the smMLCK model originated from the cAPK crystal structure, it is independently plausible for its sequence by this criterion. In aligning the sequences, there were only six regions with inserts or deletions in the smMLCK sequences, and the majority of these lie in loops between the conserved secondary structural features (Figs. 2 to 4). The B helix (residues 76 to 82) of the cAPK structure is probably the only secondary structural feature absent in the smMLCK. Details of the comparison between the smMLCK model and the cAPK structure are given in (*14*). For the structure comparison, both sets of residue numbers are used with one in parentheses.

Extensive peptide analog data were used to define the requirements for peptide recognition by smMLCK (*16, 17*). The 23-residue peptide substrate MLC(1-23) was

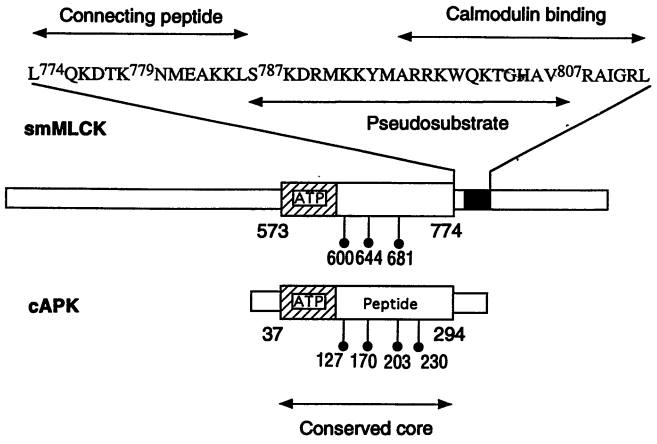

Fig. 1. Comparison of the domain structure of smMLCK and cat-cAPK. The conserved catalytic core is shown in the large box, with the ATP binding domain hatched. The acidic residues within the cAPK core implicated in substrate binding are indicated, along with the homologous residues in smMLCK. The regulatory domain of smMLCK, including the overlapping CaM-binding and pseudosubstrate sequences, is also shown.

D. R. Knighton, J. M. Sowadski, L. F. Ten Eyck, S. S. Taylor, Departments of Chemistry, Medicine, and Biology, University of California San Diego, La Jolla, CA 92093–0654.
R. B. Pearson and B. E. Kemp, St. Vincent's Institute of Medical Research, 41 Victoria Parade, Fitzroy, Victoria 3065, Australia.
A. R. Means, Department of Pharmacology, Duke University, Durham, NC 27710.
*To whom correspondence should be addressed.

phosphorylated with a Michaelis constant (K_m) of 2.7 μM (*16*). Within this sequence, the most influential four basic residue determinants were present in the 13-residue peptide MLC(11-23) [see Fig. 5 (*16, 17*)]. The peptide analog, KKRAARATS[19]NVFA (*18*) was phosphorylated at Ser[19] with a K_m = 8 μM and a V_{max} = 1.4 μmol min^{-1}, and the substitution of any of the four basic residues resulted in an increase in K_m of tenfold or more (*17*). Thus smMLCK shares a preference for basic residues at the P-3 and P-6 positions with the cAPK (Fig. 5) but has additional preferences at the P-7 and P-8 positions.

The inhibitor peptide, PKI(5-24) co-crystallized with cat-cAPK, binds primarily to residues in the large lobe (Fig. 5). The PKI sequence, R[18]RNAI[22], is referred to as the consensus site and is shared with many substrates and inhibitors of cAPK. Ala[21] replaces the phosphorylation site, P (see Fig. 5 and Table 1). The consensus portion of PKI(5-24) binds in an extended conformation in a cleft between the two lobes, whereas the NH_2-terminal portion of the peptide conveys high-affinity binding properties. This region forms an amphipathic α helix and binds to a hydrophobic groove on the surface of the large lobe (*10*). Four core acidic residues in particular are important for PKI(5-24) binding (Fig. 5). Glu[230] and Glu[170] interact with the P-2 Arg, and Glu[127] interacts with the P-3 Arg. Glu[203] contributes to recognition of the Arg at the P-6 position. In smMLCK, the residues corresponding to Glu[170] (Glu[644]), Glu[127] (Glu[600]), and Glu[203] (Glu[681]) in cAPK are conserved, explaining the importance of the P-3 and P-6 basic residues. However, Glu[230], which interacts with the P-2 Arg, is replaced by Ile[708] in smMLCK. The P-3 Arg[16] is required for efficient phosphorylation by smMLCK and is essential in directing the site of phosphorylation (*16*). Moving Arg[16] one position amino terminal to give KKRA**R**AATS[19]NVFA switched the site of phosphorylation completely to Thr[18] from Ser[19], whereas substituting it with Ala resulted in 60% phosphorylation on Thr[18]. In addition, the peptide RATSNVFS, containing the P-3 Arg alone, is an extremely poor substrate (K_m = 13 mM) and is phosphorylated largely at Thr[18] (71%), indicating that smMLCK may recognize a P-2 Arg. The interaction may be with Glu[644] (Glu[170]), but the kinetics are very unfavorable. skMLCK also exhibits a strong preference for Arg at the P-3 and P-6 positions in peptide substrates and shares corresponding acidic recognition sites in the core, but the skeletal MLC has Glu at the P-3 position; the corresponding peptide is a very poor substrate (*19*).

Unlike cAPK, additional basic residues at the P-7 and P-8 positions act as specificity determinants for smMLCK. Substitution of either Lys[11] or Lys[12] in the MLC substrate peptide resulted in 20-fold increases in K_m (*17*). These additional basic residue specificity determinants at the P-7 and P-8 positions probably interact with other acidic residues within smMLCK, such as Asp[719] and Glu[723]. However, skMLCK differs from smMLCK in having a greater preference for basic residues at the P-10 and P-11 positions than the P-7 and P-8 positions (*19*) but it is not clear from the model why this is so. The Arg residue present in the smMLC peptide

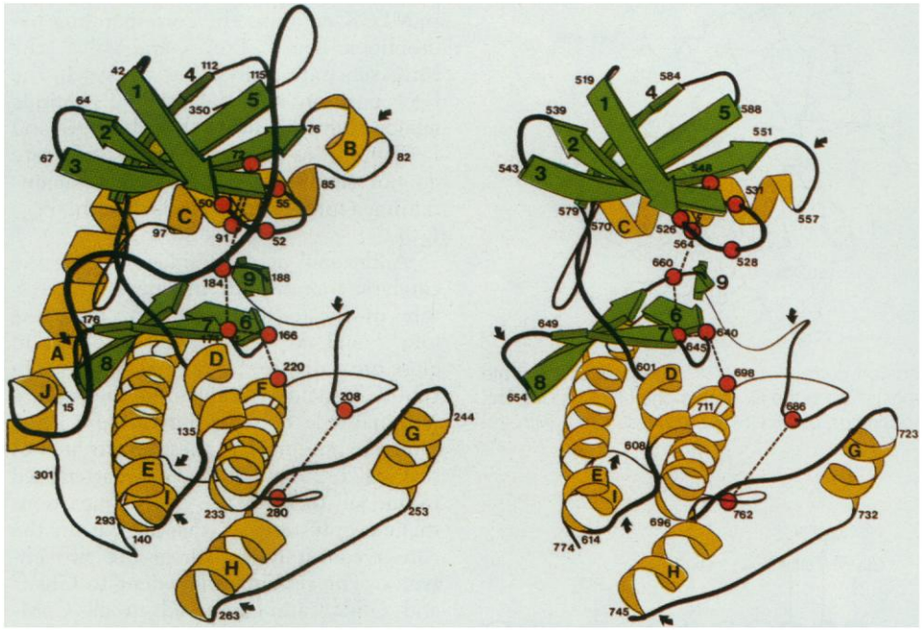

Fig. 2. Ribbon diagrams of cat-cAPK and the modeled core of smMLCK. The β strands are shown in green and the α helices in yellow. The residues that are invariant throughout the kinase family (Gly[50], Gly[52], Gly[55], Lys[72], Glu[91], Asp[166], Asn[171], Asp[184], Glu[208], Asp[220], and Arg[280] in cAPK) and the equivalent residues in smMLCK are shown in red. Regions where inserts or deletions occur in MLCK are indicated by the small arrows. The dotted lines link conserved residues whose side chains interact in the model.

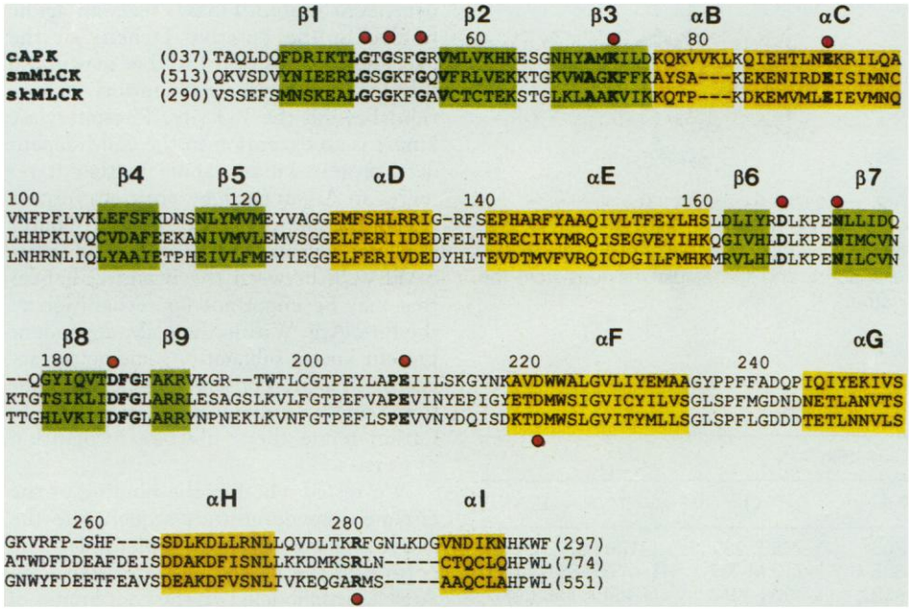

Fig. 3. Sequence alignment of cAPK and the MLCKs. The alignment has been made between residues 37 and 297 of the cAPK, 513 to 774 of the smMLCK, and 290 to 551 of skMLCK (β strands, green, and α helices, yellow). The invariant residues are indicated by closed circles.

substrate at P-15 may interact with Asp[734]. The substrate binding site for smMLCK thus extends along the cleft or major groove between the two lobes containing multiple sites for potential electrostatic interactions (Fig. 5). The smMLCK structure provides a remarkable series of electrostatic interactions matching the known basic residue determinants for the smMLCK peptide substrate but does not reveal the basis for some of the subtle differences in substrate specificity between smMLCK and skMLCK.

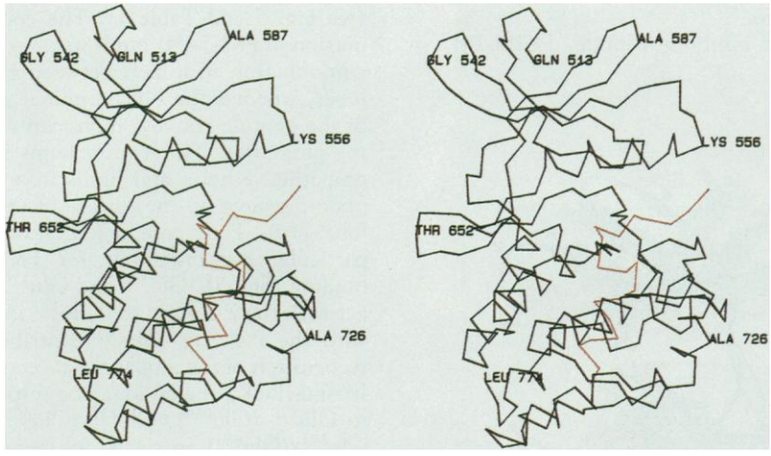

Fig. 4. Stereo view of the Cα backbone of the modeled core of smMLCK. The pseudosubstrate peptide MLCK(789-808) is shown in "red" superimposed upon the Cα backbone of smMLCK. The pseudosubstrate peptide was modeled on the coordinates of PK1(5-24), cocrystallized with the catalytic subunit of cAPK.

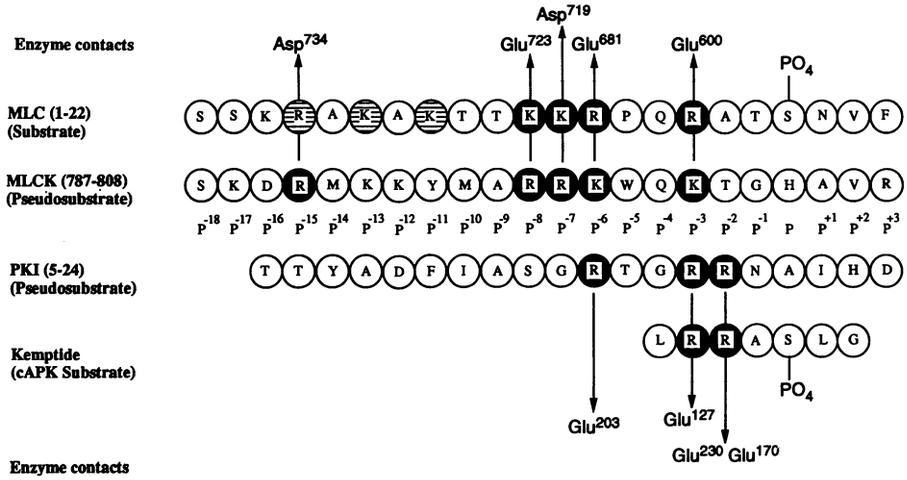

Fig. 5. Substrate and pseudosubstrate peptides for smMLCK and cAPK. The substrate and pseudosubstrate peptides for both MLCK and cAPK are aligned based on the substrate phosphorylation site (designated P). Each residue is assigned a position relative to the phosphorylation site. Basic residues involved in peptide recognition are shaded, and the acidic residues from the catalytic core thought to interact with these are indicated.

Table 1. Comparison of peptide recognition sites in the catalytic cores of cAPK and in the Ca-CaM–dependent protein kinases.

Protein kinase	Site				
	P-2	P-3	P-6	P-11	P + 1
cAPK	E-170 E-230	E-127	E-203	Y-235 F-239	L-198 P-202 L-205
smMLCK	E-644 I-708	E-600	E-682	L-713 M-717	L-677 P-681 V-684
skMLCK	E-421 M-485	E-377	E-459	L-490 L-494	N-454 P-458 L-461
Phosphorylase kinase γ	E-154 T-222	E-111	S-189	S-227 W-231	V-184 P-188 L-191
CaM II PKα	E-139 I-202	E-96	G-178	Y-210 W-214	F-173 P-177 L-180

The hydrophobic pocket in cAPK responsible for binding the P+1 residue is adjacent to the autophosphorylation site Thr[197] and consists of a pocket formed by Leu[198], Pro[202], and Leu[205]. Although the smMLCK contains the corresponding hydrophobic Leu[676], Pro[680], and Val[683], the MLC substrate peptide has an Asn in the P+1 position and Val at P+2. Peptide analog studies indicate that the P+1 and P+2 positions in the smMLCK substrate do not significantly contribute to maintaining a low K_m but do influence the V_{max} (*20*).

Additional acidic residues outside the catalytic core of both enzymes may play a role in substrate recognition, including Glu[331] and Asp[329] for cAPK that are in close proximity to the P-3 Arg position. In skMLCK, acidic residues on the amino-terminal side of the catalytic core affect substrate binding, but these effects are not strong (*21*, *22*) and we have not attempted to model them here. Our comparisons include only residues that lie within the conserved catalytic core of the two enzymes. The residues equivalent to Glu[170] and Glu[127] are conserved in all CaM-dependent protein kinases, but not a Glu at position 203 (*8*) (Table 1). None have a conserved acidic residue equivalent to Glu[230]. The absence of an acidic residue at position 230 is in keeping with the absence of a P-2 Arg requirement for these enzymes. Interestingly, all of the CaM-dependent protein kinases have an acidic residue in the putative D helix at the equivalent of position 130 that may play a role in basic residue recognition at positions beyond the P-3 site. Phosphorylase kinase is an exception in the CaM-dependent protein kinase family in that it requires an Arg in the P+2 position. Inspection of the phosphorylase kinase γ-subunit sequence reveals that it contains an insert SAEEVQE between the B and C helices that may be important for recognition of the P+2 Arg. Within the CaM-dependent protein kinase subfamily, some acidic acceptor sites are conserved, whereas additional ones in individual members serve to further refine their substrate recognition properties.

We tested whether the binding of the enzyme's pseudosubstrate sequence to the active site was structurally plausible. The portion of the pseudosubstrate, Asp[789] to Arg[808], was modeled on the coordinates of PKI(5-24) located in the extended cleft between the two lobes. Corresponding electrostatic interactions between the ba-

sic residues of the pseudosubstrate peptide and acidic residues located in the larger lobe were expected (Fig. 5), including Glu^{600} to Lys^{802} and Glu^{681} to Lys^{799}, that correspond to the cAPK recognition residues Glu^{203} and Glu^{127} for the P-3 and P-6 sites, respectively. Additional electrostatic interactions are possible for the basic residues extending beyond the P-6 position in the smMLCK pseudosubstrate. The P-7 Arg^{798} is located close to Asp^{719}, whereas the P-8 Arg^{797} is close to Glu^{723}. The closest acidic residue to the P-12 and P-13 residues, Lys^{793} and Lys^{792}, is Asp^{789} within the pseudosubstrate. Arg^{790} (P-15) is positioned near Asp^{734} in the highly acidic sequence $D^{734}FDDAEFDEISD^{745}$ (Fig. 5). Complementarity is also seen for hydrophobic interactions. For example, the model places Tyr^{794} in the hydrophobic pocket $L^{713}SPFM^{717}$. The corresponding hydrophobic pocket in cAPK is largely conserved and occupied by the P-11, Phe^{10} of PKI(5-24) (*10*). The surface adjacent to this hydrophobic pocket is provided by the D helix and is overall more hydrophobic in smMLCK and in cAPK. This extended hydrophobic surface complements Trp^{800} that occupies the P-5 site in the pseudosubstrate sequence. The α-helical section of the pseudosubstrate (Asp^{789} to Arg^{797}) is aligned antiparallel with the G helix so that helix dipoles may contribute to structural stability. The complementarity of the modeled autoinhibitor site with the smMLCK surface is seen in Fig. 6.

One of the most striking features of the modeled smMLCK structure is the proximity of the amino terminus of the pseudosubstrate sequence (Asp^{789}) to the carboxyl end of the catalytic core ($HPWL^{774}$), a distance of 28.8 Å from the Cα Trp^{773} to the Cα Asp^{789}. This distance is easily spanned by 14 residues (Figs. 4 and 6), although it is not possible to precisely predict the path of the connecting backbone.

The model indicates that there are a large number of possible electrostatic interactions between the catalytic core and the peptide connecting it to the pseudosubstrate sequence, suggesting that this peptide may strengthen the pseudosubstrate interaction. Also, the pseudosubstrate sequence likely assumes a conformation similar to PKI(5-24) with Tyr^{794} at the P-11 site analogous to Phe^{10} in the PKI. Although peptides derived from the connecting peptide region themselves do not act as pseudosubstrate inhibitors (*6*, *23*), their influence on the binding of the pseudosubstrate itself was not known. The peptide smMLCK(774-807) was synthesized, and it competitively inhibited phosphorylation by the constitutively active smMLCK truncated at Lys^{779} (*23*) with an extremely low K_i (0.33 nM) (Table 2). This inhibitor peptide exceeds the potency of the PKI(5-24) itself by a factor of 5 (*24*). Henderson analysis was used to assess the inhibition constant K_i. The smMLCK(774-807) is 30-fold more potent than the pseudosubstrate peptide smMLCK(787-807), indicating that the sequence L^{774}QKDTKNME-AKKL^{786} contributes significantly to the binding of the pseudosubstrate peptide as expected from the model. The extended peptide was also a potent noncompetitive inhibitor with respect to ATP, with a K_i of ~27 nM.

The potency of the cAPK inhibitor peptide PKI(5-24) is influenced by the presence of an aromatic residue (Phe^{10}) at the P-11

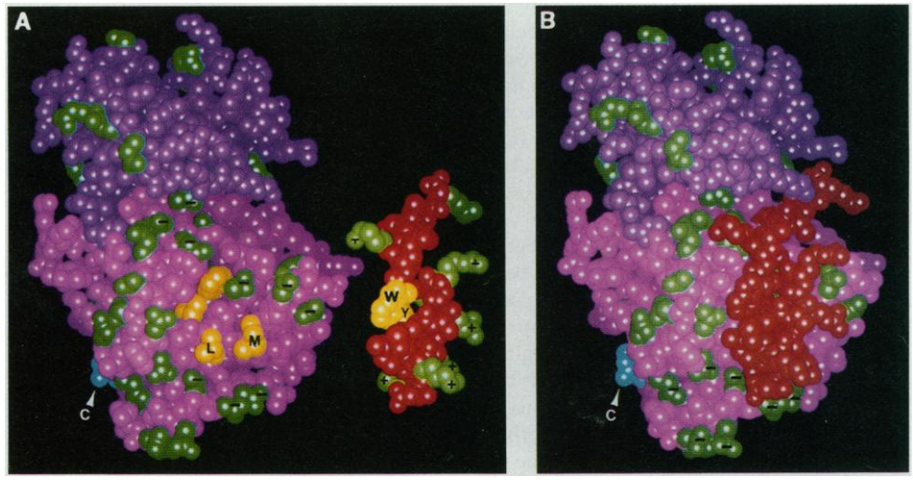

Fig. 6. Space-filling model of smMLCK showing the area of contact between the pseudosubstrate and catalytic core. The pseudosubstrate peptide is shown in red, and the acidic residues on the protein are indicated in green. The numbers of the residues between the D and E helices and from the loop between the G and H helices are indicated specifically. The basic residues in the pseudosubstrate peptide are shown in green. The two hydrophobic residues in the pseudosubstrate sequence Tyr^{794} and Trp^{800} are shown in yellow, as is the complementary hydrophobic surface on the smMLCK model.

Table 2. Inhibition of constitutively active smMLCK by pseudosubstrate peptide analogs. The CaM-independent 61-kD form of MLCK generated by trypsin digestion (*5*) was used because the pseudosubstrate peptides have CaM antagonist activity. The K_i (peptide substrate) values for the two shorter analogs were determined from secondary plots of K_m (determined by Lineweaver-Burk analysis) versus inhibitor concentration. The K_i (peptide substrate) for MLCK(774-807) was determined by Henderson analysis. Values shown are mean ± SEM for three independent experiments or average of duplicates that differ by less than 10%. The IC_{50} (ATP) values were determined from plots of MLCK activity versus inhibitor concentration at 13.5 μM ATP (K_m) and saturating peptide substrate (145.5 μM). Synthetic peptides were prepared by solid-phase peptide synthesis, purified by cation exchange chromatography and reversed-phase chromatography, and characterized by analytical high-performance liquid chromatography and quantitative amino acid analysis (*5*).

Peptide sequence	K_i (nM)		IC_{50} ATP (nM)
	Substrate	ATP	
L^{774}QKDTKMMEAKKLS^{787}KDRMKKYMARRKWQKTGHAV^{807}	0.33 ± 0.08	27.3 ± 6.8*	74
S^{787}KDRMKKYMARRKWQKTGHAV^{807}	11.7 ± 2.2	†	149
S^{787}KDRMKKAMARRKWQKTGHAV^{807}	21	2300*	1.1×10^4
L^{774}QKDTKNMEAKKLS^{787}	‡		$>1.4 \times 10^6$

*Although the secondary plot of K_m/V_{max} versus inhibitor concentration was linear, the Lineweaver-Burk plot showed mixed noncompetitive inhibition for [A794] MLCK 787-807 as did the Henderson plot for MLCK 774-807. †The secondary plot of K_m/V_{max} MLCK(787-807) concentration was nonlinear so a K_i could not be determined. ‡No inhibition of enzyme activity at limiting substrate (9 μM) and 1.4 mM MLCK (774-787).

position. This residue interacts with a complimentary hydrophobic pocket on the surface of cAPK containing $Y^{235}PPFF^{239}$. For smMLCK, Tyr^{794} is situated at the P-11 position in the pseudosubstrate peptide and the corresponding pocket on the surface of the enzyme contains $L^{713}SPFM^{717}$. We examined the effect of substituting the Tyr^{794} with Ala in the pseudosubstrate peptide, smMLCK(787-807) to determine its importance for peptide binding to smMLCK. The Ala^{794} analog peptide has a K_i of 21 nM, ~2-fold greater than the pseudosubstrate peptide, but much less than the 100-fold difference seen with the corresponding cAPK inhibitor peptide (*25*). Unexpectedly, the largest effect of the Ala^{794} substitution was on the inhibition of ATP binding with the IC_{50} value increasing ~75-fold from 149 to 11,030 nM (Table 2). The inhibition patterns for both peptides were complex, indicating a mixed noncompetitive mechanism. The secondary plots of K_m/V_{max} versus smMLCK(787-807) were nonlinear, so that a K_i value could not be estimated. Previous studies on the mechanism of MLC phosphorylation (*26*) have also revealed complex kinetics indicating that this is not restricted to the pseudosubstrate peptides.

Previously there have been a number of experimental observations that have been difficult to accommodate in the pseudosubstrate model of smMLCK regulation. If the pseudosubstrate acted as a competitive inhibitor at the active site, it was not obvious why very high concentrations of substrate peptide (300-fold greater than the K_m) failed to activate the enzyme by competition. The exceedingly high potency of the extended pseudosubstrate peptide, $K_i = 0.33 \pm 0.08$ nM, containing the pseudosubstrate sequence as well as the connecting peptide, explains why even high concentrations of substrate peptides do not overcome the intramolecular inhibition. The model also explains why a number of mutation experiments directed at the pseudosubstrate region failed to render the enzyme constitutively active. These included substitution of a number of basic residues in the pseudosubstrate region with either Ala or Glu (*27*) and even reversal of a segment of the pseudosubstrate sequence (*28*). The very tight binding of the pseudosubstrate sequence including the connecting peptide and the extensive number of electrostatic contacts over a large surface area make it apparent why classical proteolysis experiments (*5*, *6*) and truncation mutation experiments (*7*) were the most effective approaches, and why other mutation experiments (*27*, *28*) gave results that were interpreted as being inconsistent with the pseudosubstrate model. Further evidence in favor of the location of the pseudosubstrate sequence has been obtained by replacing part of it with the substrate sequence and showing that it becomes autophosphorylated in the absence of CaM (*29*).

REFERENCES AND NOTES

1. B. E. Kemp and J. T. Stull, in *Peptides and Protein Phosphorylation*, B. E. Kemp, Ed. (CRC Press, Boca Raton, FL, 1990), pp. 115–133.
2. B. E. Kemp and R. B. Pearson, *Biochim. Biophys. Acta* **1094**, 67 (1991).
3. R. B. Pearson *et al.*, unpublished results.
4. B. E. Kemp, R. B. Pearson, V. Guerriero, Jr., I. C. Bagchi, A. R. Means, *J. Biol. Chem.* **262**, 2542 (1987).
5. R. B. Pearson, R. E. H. Wettenhall, A. R. Means, D. J. Hartshorne, B. E. Kemp, *Science* **241**, 970 (1988).
6. R. B. Pearson *et al.*, *Eur. J. Biochem.* **200**, 723 (1991).
7. M. Ito, V. Guerriero, Jr., X. Chen, D. J. Hartshorne, *Biochemistry* **30**, 3498 (1991).
8. S. K. Hanks, A. M. Quinn, T. Hunter, *Science* **241**, 42 (1988); S. K. Hanks and A. M. Quinn, *Methods Enzymol.* **200**, 38 (1991).
9. D. R. Knighton *et al.*, *Science* **253**, 407 (1991).
10. D. R. Knighton *et al.*, *ibid.*, p. 414.
11. The model of the smMLCK core was prepared by replacing the side chains of the cAPK crystal structure model [Brookhaven Protein Data Bank (BPDB) entry 1CPK] with smMLCK side chains according to the alignment in Fig. 3. The relative inserts and deletions were modeled initially with plausible backbone fragments of the same length extracted from structures in the BPDB and given the smMLCK sequence. The conserved cAPK backbone atoms were not altered in the replacement procedure, and smMLCK side chains followed cAPK side chain conformations where possible. The core smMLCK model side chains were subjected to energy minimization with the program DISCOVER (version 2.70, Biosym Technologies, Inc., San Diego, CA, 1991) until the maximum gradient of any atom was less than 5 kcal mol^{-1} $Å^{-1}$. The preliminary fitted inserts-deletions of the catalytic core were subjected to simulated annealing with the program X-PLOR [A. T. Brünger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987)] from 1000 to 300 K against a fixed secondary structure, followed by additional loop energy minimization. Twenty residues from the smMLCK pseudosubstrate sequence were modeled onto the backbone of PKI(5-24), and the smMLCK side chains of both the core and the autoinhibitory sequence were adjusted manually to maximize hydrophobic and charge interactions. This assembly was subjected to 2400 conjugate gradient cycles of energy minimization with charges on Lys, Arg, Asp, and Glu side chains turned off and Cα atoms in the conserved regions fixed. Additional conjugate gradient cycles were run, with conserved Cα atoms harmonically restrained to initial positions, until the maximum gradient was less than 0.0005 kcal mol^{-1} $Å^{-1}$. The final smMLCK model with the pseudosubstrate sequence had deviations from ideal bond lengths of 0.009 Å and from ideal bond angles of 2.7°. The root-mean-square (rms) deviation in Cα positions in the conserved regions from those of the cAPK crystal structure was 0.19 Å.
12. In 3-D profile analysis, the empirical probabilities of finding a given amino acid in a specified residue environment in other protein structures are correlated with the residue environments evaluated for the modeled sequence, allowing an evaluation of the probability of an amino acid sequence occurring in a particular 3-D configuration [J. U. Bowie, R. Lüthy, D. Eisenberg, *Science* **253**, 164 (1991)]. Local and global sums of individual residue scores display different patterns for correctly and incorrectly folded sequences (*13*). The range of possible residue scores was −3.04 to 1.75. After mismatching the cAPK sequence to its 3-D profile by cyclically permuting by the 281-residue cAPK sequence with respect to the original 3-D profile, the total score was 0.51. In contrast, for the correct match the total score of the 281 residue from the cAPK crystal structure was 119 and for the corresponding 282 residues of the modeled smMLCK core with the pseudosubstrate sequence was 116. These are not greatly different, and both are consistent with the observed dependence of total score on sequence length. Examination of the window length–normalized scores of a 20-residue window moved along both cAPK and smMLCK sequences resembled the pattern for correct structures (*13*) in rarely falling below 0.2 and never below 0.0 (D. R. Knighton and L. F. Ten Eyck, unpublished results). The rare dips were attributable to the improbability of buried charged residues or to the occasional exposed hydrophobic residue. Several such buried charge groups occur in the cAPK crystal structure. For example, Lys^{72} and Glu^{91}, invariant protein kinase residues, are buried in the nucleotide binding site, and similarly invariant Glu^{208} and Arg^{280} form a buried ion pair.
13. R. Luthy, J. U. Bowie, D. Eisenberg, *Nature* **356**, 83 (1992).
14. In comparing the smMLCK model and the cAPK structure, the first three β strands that contain the glycine-rich loop and Lys^{72} (smMLCK, Lys^{548}) show no inserts or deletions in this alignment, and there is likely to be a high level of structural conservation in the residues 37 to 75 (smMLCK, 513 to 542). However, the B helix between residues 76 and 82 of the cAPK is probably absent in the smMLCK structure because of the absence of the sequence VKL^{82} (Fig. 3). skMLCK also contains a helix-breaking Pro in this region. This short B helix is unlikely to be conserved in many protein kinases (*8*). A shortened loop structure consisting of five residues was substituted here to connect β strand 3 with the C helix. The sequence extending from β strand 4 through the D helix contains no inserts or deletions. The sequence connecting the fifth β strand and the D helix links the small and the large lobes and is particularly conserved in these enzymes, with five of the seven residues identical. The loop between the D and E helices of smMLCK contains an additional Phe that is close to a hydrophobic pocket and a cluster of four acidic residues in the sequence $D^{607}EDFE^{611}$. This acidic sequence is not present in cAPK and may play a role in positioning the pseudosubstrate sequence within the active site. The loop that connects the sixth and seventh β strands shows strong homology with all of the Ser/Thr protein kinases. In cAPK, this loop, $D^{166}LKPEN$ (*18*), is referred to as the catalytic loop and is considered important for directing catalysis as well as positioning the substrate. The loop between the seventh and eighth β strands contains a two-residue insert (Lys^{651} and Thr^{652}) in the smMLCK sequence. The loop between β strands 8 and 9 contains the invariant $D^{184}FG$, where Asp^{184} (Asp^{660} in smMLCK) is essential for MgATP binding (J. Zheng *et al.*, in preparation). The region that connects β strand 9 with the F helix provides an extended surface in cAPK that contributes in part to recognition of the portion of the peptide that follows the site of phosphotransfer. In cAPK, this region contains a hydrophobic pocket comprised of the side chains of Leu^{198}, Pro^{202}, and Leu^{205} that serves to anchor the hydrophobic residue present in the P+1 position of the inhibitor peptide (Fig. 5). An additional two-residue insert (Gly^{671}-Ser^{672}) occurs here in the smMLCK; however, the side chains of Leu^{676}, Pro^{680}, and Val^{683} form an homologous hydrophobic pocket. A striking difference in smMLCK is the absence of a Thr^{197}, an autophosphorylation site in cAPK. In the cAPK structure, the Thr^{197} phosphate interacts directly with Arg^{165}, His^{87}, and Lys^{189} and is very resistant to phosphatases (*15*). Phosphorylation of Thr^{197} presumably contributes to the overall stabilization of this region of the enzyme (*9*). Arg^{165}, in particular, is highly conserved in nearly all protein kinases. In smMLCK,

Val[675] and Leu[639] substitute for Thr[197] and Arg[165], suggesting that hydrophobic interactions may substitute for this ion pair in smMLCK. The CaM-dependent protein kinases are conspicuous in their lack of a conserved Arg[165] compared to all other protein kinases. The region containing the F helix–loop–G helix is reasonably conserved. smMLCK contains two substitutions with acidic residues at positions Asp[721] and Glu[723] (Gln[243] and Ile[245] in cAPK). The loop between the G and H helices is extended in smMLCK by inserts totaling four residues. In this acidic region of smMLCK, 7 out of 15 residues are either Glu or Asp. These residues form part of a perimeter of acidic residues together with those derived from the loop between the D and E helices (see Fig. 6). A deletion of five residues occurs in smMLCK in the loop between the H and I helix. This loop contains the last conserved core residue, Arg[280] (smMLCK, Arg[762]) that forms an ion pair with Glu[208] (smMLCK Glu[686]) and is located under the E and F helices (Figs. 2 and 3). Similar deletions are found in almost all other protein kinases except members of the protein kinase C subfamily. The I helix in smMLCK may be deleted, but we argue that the loop region is more likely to be shortened than the I helix deleted. It seems likely that Trp[773] (cAPK Trp[297]) in the sequence HPW is the most carboxyl-terminal conserved structural feature between the cAPK and the catalytic core of the MLCKs. Although the mammalian cAPK has the sequence HKW preceding this residue, the yeast cAPK isoenzymes, TPK1, 2 and 3 all have the HPW equivalent to MLCK (*8*). The hydrophobic environment of the Trp[773] in smMLCK is contributed by residues from the E and F helices and appears to be conserved.

15. S. Shoji, K. Titani, J. G. Demaille, E. H. Fischer, *J. Biol. Chem.* **254**, 6211 (1979).
16. B. E. Kemp, R. B. Pearson, C. House, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 7471 (1983).
17. B. E. Kemp and R. B. Pearson, *J. Biol. Chem.* **260**, 3355 (1985).
18. 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.
19. C. H. Michnoff, B. E. Kemp, J. T. Stull, *J. Biol. Chem.* **261**, 8320 (1986).
20. R. B. Pearson, L. Y. Misconi, B. E. Kemp, *ibid.*, p. 25.
21. B. P. Herring, D. P. Fitzsimons, J. T. Stull, P. J. Gallagher, *ibid.* **265**, 16588 (1990).
22. S. A. Leachman, P. J. Gallagher, B. P. Herring, M. J. McPhaul, J. T. Stull, *ibid.* **267**, 4930 (1992).
23. N. J. Olson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 2284 (1990).
24. H. C. Cheng *et al.*, *J. Biol. Chem.* **261**, 989 (1986).
25. D. B. Glass, L. J. Lundquist, B. M. Katz, D. A. Walsh, *ibid.* **264**, 14579 (1989).
26. A. Sobieszek, *Eur. J. Biochem.* **199**, 735 (1991).
27. B. P. Herring, *J. Biol. Chem.* **266**, 11838 (1991).
28. M. O. Shoemaker *et al.*, *J. Cell Biol.* **111**, 1107 (1990).
29. I. R. Bagchi, B. E. Kemp, A. R. Means, *Mol. Endocrinol.* **6**, 621 (1992).

30. Supported by the NH&MRC and the Lucille P. Markey Foundation; by separate grants from the NIH (S.S.T., A.R.M., and J.M.S.), the American Cancer Society (J.M.S. and S.S.T.), NSF (S.S.T. and L.T.E.), NIH training grants T32CA09523 and T32DK07233 (D.R.K.), Anti-Cancer Council Victoria, and the NHF (B.E.K.). We thank the San Diego Supercomputing Center for use of the Advances Scientific Visualization Laboratory and the Cray Y-MP8/864, the USCD Department of Chemistry Modeling Facility, G. Hasegawa for help in preparation of the manuscript, and D. Appleby for help in preparation of the diagrams. Atomic coordinates have been submitted to the Brookhaven Protein Data Bank.

15 May 1992; accepted 29 June 1992

Lack of Protective Immunity Against Reinfection with Hepatitis C Virus

Patrizia Farci,* Harvey J. Alter, Sugantha Govindarajan, Doris C. Wong, Ronald Engle, Richard R. Lesniewski, Isa K. Mushahwar, Suresh M. Desai, Roger H. Miller, Norio Ogata, Robert H. Purcell

Some individuals infected with hepatitis C virus (HCV) experience multiple episodes of acute hepatitis. It is unclear whether these episodes are due to reinfection with HCV or to reactivation of the original virus infection. Markers of viral replication and host immunity were studied in five chimpanzees sequentially inoculated over a period of 3 years with different HCV strains of proven infectivity. Each rechallenge of a convalescent chimpanzee with the same or a different HCV strain resulted in the reappearance of viremia, which was due to infection with the subsequent challenge virus. The evidence indicates that HCV infection does not elicit protective immunity against reinfection with homologous or heterologous strains, which raises concerns for the development of effective vaccines against HCV.

HCV, the principal cause of non-A, non-B (NANB) hepatitis, accounts for more than 90% of posttransfusion hepatitis (PTH) in the United States (*1*) and results in chronic infection in over 50% of cases (*2*). Although NANB hepatitis was first recognized in the 1970s (*3*), HCV has only recently been identified and partially characterized (*4*). Before the discovery of HCV, the observation of multiple, distinct episodes of acute NANB hepatitis in the same individual suggested the existence of more than one NANB hepatitis agent (*5*). A similar clinical pattern was documented in cross-challenge studies in chimpanzees (*6*, *7*) although most of these studies did not show a second episode of acute NANB hepatitis (*8*, *9*). Interpretation of these data was difficult and was further complicated by the fact that recurrent episodes of hepatitis were also observed in chimpanzees that were not rechallenged (*10*) and in chimpanzees rechallenged with homologous inocula (*11*), which suggests that there might be explanations other than the existence of multiple NANB agents to account for these observations. With the discovery of HCV and the availability of serologic assays that detect antibodies to this virus (*12*), it has become evident that HCV is the major cause of NANB hepatitis and that if other NANB agents exist they account for only a minority of such cases (*13*).

Thus, the recurrence of NANB hepatitis could be explained by viral factors, such as the existence of multiple HCV serotypes or the emergence of mutant viruses that escape neutralization by the host's immune system. Alternatively, the inability of the host to mount a protective immune response could lead to reactivation or reinfection with the same virus. These questions can now be explored with sensitive and specific techniques such as the polymerase chain reaction (PCR) (*14*). To investigate whether primary HCV infection elicits protective immunity against reinfection with homologous or heterologous strains of virus, we have re-evaluated a series of cross-challenge experiments in chimpanzees, previously conducted at the National Institutes of Health.

The patterns of HCV viremia and humoral immune response were analyzed in five chimpanzees (*15*) sequentially inoculated with different HCV strains derived from five unrelated prospectively studied individuals (G, F, H, K, and R) with posttransfusion NANB hepatitis (*7*, *16*). Three chimpanzees were challenged twice and two were challenged four times. The animals were followed for a mean period of 32 months (range, 12 to 51 months). One chimpanzee (number 963) was rechallenged with the homologous inoculum and four (numbers 189, 196, 502, and 793) with heterologous inocula; the interval between the challenges ranged from 6 to 19 months (mean was 11.2 months). The challenge inocula were prepared from serum or plasma (Table 1) used for other successful chal-

P. Farci, D. C. Wong, R. H. Miller, N. Ogata, R. H. Purcell, Hepatitis Viruses Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892.
H. J. Alter, Department of Transfusion Medicine, Warren G. Margnuson Clinical Center, National Institutes of Health, Bethesda, MD 20892.
S. Govindarajan, Rancho Los Amigos Hospital, Downey, CA 90242.
R. Engle, Division of Molecular Virology and Immunology, Department of Microbiology, Georgetown University, Rockville, MD 20852.
R. R. Lesniewski, I. K. Mushahwar, S. M. Desai, Experimental Biology Research, Abbott Laboratories, North Chicago, IL 60064.
*To whom correspondence should be addressed.