Characterization of the Secondary Structure of Calmodulin in Complex with a Calmodulin-Binding Domain Peptide†

Sharon M. Roth,‡,§ Diane M. Schneider,‡,§,∥ Laura A. Strobel,‡ Mark F. A. Van Berkum,⊥,# Anthony R. Means,⊥,▽ and A. Joshua Wand*,‡,§,○

Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, Pennsylvania 19111, Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, and Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030

Received August 5, 1991; Revised Manuscript Received October 31, 1991

ABSTRACT: The interaction between calcium-saturated chicken calmodulin and a peptide corresponding to the calmodulin-binding domain of the chicken smooth muscle myosin light chain kinase has been studied by multinuclear and multidimensional nuclear magnetic resonance methods. Extensive ^1H and ^{15}N resonance assignments of calmodulin in the complex have been obtained from the analysis of two- and three-dimensional nuclear magnetic resonance spectra. The assignment of calmodulin in the complex was facilitated by the use of selective labeling of the protein with α-^{15}N-labeled valine, alanine, lysine, leucine, and glycine. These provided reference points during the main-chain-directed analysis of three-dimensional spectra of complexes prepared with uniformly ^{15}N-labeled calmodulin. The pattern of nuclear Overhauser effects (NOE) seen among main-chain amide NH, $C_\alpha H$, and $C_\beta H$ hydrogens indicates that the secondary structure of the globular domains of calmodulin in the complex closely corresponds to that observed in the calcium-saturated state of the protein in the absence of bound peptide. However, the backbone conformation of residues 76–84 adopts an extended chain conformation upon binding of the peptide in contrast to its helical conformation in the absence of peptide. A sufficient number of NOEs between the globular domains of calmodulin and the bound peptide have been found to indicate that the N- and C-terminal regions of the peptide interact with the C- and N-terminal domains of calmodulin, respectively. The significance of these results are discussed in terms of recently proposed models for the structure of calmodulin–peptide complexes.

Calmodulin is a small acidic protein which binds four calcium ions with dissociation constants in the micromolar range. Its primary function is the modulation of the activity of many enzymes in response to changes in calcium concentrations [for a review, see Means (1988)]. Calcium-dependent regulation occurs through a tight binding interaction of CaM[1] with specific domains of the regulated enzymes. CaM-binding domains have been identified and characterized for a number of enzymes, including the smooth (Kemp et al., 1987) and skeletal (Blumenthal et al., 1985) muscle myosin light chain kinases. These domains are generally small (20 amino acid residues), contain predominantly hydrophobic and basic amino acids, have a propensity for helix formation, and bind to calmodulin in a calcium-dependent manner. Over the past few years a general model for the structure of calmodulin-binding domain peptides bound to CaM has been developed [e.g., McDowell et al. (1985), O'Neil et al. (1987), and O'Neil and DeGrado (1990)]. Its central feature is the formation of an α-helical structure with a hydrophobic face in contact with CaM and a basic face in contact with solvent. The structure of the smooth muscle myosin light kinase calmodulin-binding domain peptide (smMLCKp) bound to CaM has recently been determined (Roth et al., 1991). The general features of the backbone conformation of the bound peptide are consistent with the amphiphilic helix model although distorted helical turns were found in the center and near the C-terminal end of the bound peptide (Roth et al., 1991).

The effects of the binding of small peptides to CaM have been studied by a variety of physical methods. A central issue in these studies has been the degree to which the formation of a complex disturbs the secondary and tertiary structure of calmodulin. Studies with peptides corresponding to the skeletal muscle myosin light chain kinase calmodulin-binding domain have indicated that the secondary structure of the two globular domains of CaM in the absence of peptide (Babu et al., 1988; Ikura et al., 1990) is retained in the complex (Ikura et al., 1991; Seeholzer & Wand, 1989) with the calcium-binding domains being stabilized by the binding of peptide (Seeholzer & Wand, 1989). On a more global scale, low-angle X-ray scattering studies have indicated that the binding of peptide drastically alters the molecular dimensions of calmodulin in a manner consistent with a significant distortion of the long

† This work was supported by NIH Research Grants DK-39806 (A. J.W.) and GM-33926 (A.R.M.), by NIH Grants CA-06927 and RR-05539, by an appropriation from the Commonwealth of Pennsylvania, and by a grant from the F. Ripple Foundation awarded to the Institute for Cancer Research. D.M.S. is the recipient of an NIH postdoctoral fellowship (GM-12594). S.M.R. is the recipient of an NIH predoctoral fellowship administered by the University of Pennsylvania (GM-07229).

* Address correspondence to this author at the University of Illinois.
‡ Institute for Cancer Research.
§ University of Illinois at Urbana–Champaign.
∥ Present address: Sterling-Winthrop Pharmaceutical Research Group, Malvern, PA 19355.
⊥ Baylor College of Medicine.
Present address: Department of Molecular and Cellular Biology, University of California, Berkeley, CA 94720.
▽ Present address: Department of Pharmacology, Duke University Medical Center, Durham, NC 27710.
○ Present address: Department of Biochemistry, University of Illinois at Urbana–Champaign, IL 61801.

[1] Abbreviations: CaM, $(Ca^{2+})_4$-calmodulin; COSY, *J*-correlated spectroscopy; DQF, double-quantum filter; HMQC, heteronuclear multiple-quantum correlation; MCD, main chain directed; MLCK, myosin light chain kinase; NAB, amide NH–$C_\alpha H$–$C_\beta H$ subspin system; NOE, nuclear Overhauser effect; NOESY, NOE-correlated spectroscopy; skMLCKp and smMLCKp refer to the skeletal and smooth muscle MLCK calmodulin-binding domain peptides defined by Seeholzer and Wand (1989) and Roth et al. (1991), respectively; TOCSY, total correlation spectroscopy.

0006-2960/92/0431-1443$03.00/0

central helix seen to join the two globular domains in both the crystalline (Babu et al., 1988) and solution states (Ikura et al., 1991) of CaM (Heidorn et al., 1989).

Here we report a multinuclear and multidimensional NMR study of the solution structure of calcium-saturated calmodulin in complex with a peptide based on the smooth muscle myosin light chain kinase calmodulin-binding domain. The vast majority of the main-chain ^1H resonances of calmodulin have been assigned and their short distance relationships characterized by the pattern of NOEs occurring between them. These results indicate that the secondary structure of the globular domains of calmodulin does not change significantly upon association with the smMLCKp peptide while the central residues of the putative central helix appear to adopt an extended chain conformation. These results are consistent with previous results obtained from studies of the complex of CaM with homologous skeletal muscle MLCK CaM-binding domain peptides (Seeholzer & Wand, 1989; Ikura et al., 1991). Finally, analysis of two-dimensional NOESY spectra provides a sufficient number of unequivocal NOEs to determine that the orientation of the peptide is the opposite of that proposed by Persechini and Krestinger (1988) for their type III complex and is consistent with the orientation recently proposed by O'Neil and DeGrado (1990).

Materials and Methods

Preparation of Calmodulin and Peptides. The plasmid harboring the chicken calmodulin gene (Putkey et al., 1985) was used to transform *Escherichia coli* EMG-2 cells. Transformed cells, selected on the basis of ampicillin resistance, were further transformed with the cI plasmid carrying λ repressor genes required to make the induction of protein host independent. Doubly transformed cells, selected on the basis of ampicillin and tetracyclin resistance, were adapted to minimal media containing M9 salts, without $(NH_4)_2SO_4$ and essential vitamins (Sambrook et al., 1989), 1 g/L NH_4Cl, and 200 mg/L of each amino acid. Selectively ^{15}N-labeled calmodulin was prepared using this media with the desired amino replaced with its α-^{15}N-labeled counterpart. [α-^{15}N]Amino acids were from ISOTEC and were used without further purification. Uniformly ^{15}N-labeled calmodulin was prepared using M9 salts and essential nutrients as above, 2 g/L glucose, and 1 g/L $^{15}NH_4Cl$. Cells were grown at 30 °C to an OD_{600} of 2.5 and induced at 42 °C for 1.5 h. *E. coli* expressed calmodulin was purified from cell lysates by phenyl-Sepharose affinity chromatography (Gopalarishna & Anderson, 1982) as described previously (Seeholzer & Wand, 1989) except that gel filtration was unnecessary to obtain pure (>95%, as judged by electrophoresis) protein. If required, a second purification by affinity chromatography was done. The peptide (smMLCKp) used in these studies is based upon the primary sequence of the chicken smooth muscle myosin light chain kinase calmodulin-binding domain (Lukas et al., 1986; Kemp et al., 1987). The sequence of the peptide used is

smMLCKp

Ac-A-R-R-K-W-Q-K-T-G-H-A-V-R-A-I-G-R-L-S-NH_2

The unlabeled smMLCKp peptide was prepared by solid-phase peptide synthesis on an Applied Biosystems automatic peptide synthesizer as described previously for the analagous skeletal muscle peptide (Seeholzer & Wand, 1989). ^{15}N-Labeled peptide was synthesized with [α-^{15}N]tBOC amino acids (lysine, leucine, valine, alanine, and glycine) as described previously (Roth et al., 1991). Both labeled and unlabeled peptides were purified as described previously for the skeletal MLCK calmodulin-binding domain peptide (Seeholzer & Wand, 1989). NMR samples of calmodulin–peptide complexes were prepared from stock solutions of CaM and smMLCKp as described previously (Seeholzer & Wand, 1989) except that ultrafiltration rather than lyophilization was employed. All samples of the complex were between 0.5 and 3 mM in 10 mM deuterated imidazole and 40 mM Ca^{2+} at pH 6.5 (uncorrected for the isotope effect).

NMR Spectroscopy. All NMR spectra shown were obtained on a Bruker AM600 NMR spectrometer at 25 °C. Standard pulse sequences were used to obtain phase-sensitive DQF COSY (Rance et al., 1983), spin-locked TOCSY (Bax & Davis, 1985), and NOESY spectra (Macura & Ernst, 1981). ^1H–^1H correlation spectra were generally derived from data sets composed of 700–800 free induction decays (FIDs) of 1024 complex points. ^{15}N–^1H heteronuclear multiple-quantum correlation (HMQC) spectra were obtained using the pulse sequences described by Bax et al. (1983). A delay of 4.5 ms was used to create ^{15}N–^1H coherence in all experiments. ^{15}N–^1H correlation spectra were generally derived from data sets composed of 400–600 FIDs of 1024 complex points. Three-dimensional NOESY-HMQC (50-ms mixing time) and TOCSY-HMQC (35-ms mixing time) spectra were obtained essentially as described by Kay et al. (1989), except that a simple inversion pulse was used to remove the effects of ^{15}N precession and *J*-coupling during the incremented ^1H time domain. All experiments employed direct on-resonance presaturation of the H_2O solvent line. A 128 (^1H) × 120 (^{15}N) × 512 (^1H acquisition) complex point data set was acquired. Each free induction decay was the average of 24 scans. Spectral widths of 7352 Hz for ^1H and 2294 Hz for ^{15}N were used. The spectra were not folded. Time-proportional phase incrementation was used in all experiments to provide quadrature detection during the incremented time domains (Marion & Wüthrich, 1983). Two-dimensional spectra were processed to 2K × 2K real points using the FTNMR and FELIX software from Hare Research (Bothell, WA). Three-dimensional spectra were processed to 512 × 128 × 512 real points with the upfield half of the acquisition frequency domain being discarded. Linear prediction of the first time point was used as required. All spectra shown are referenced to external 3-(trimethylsilyl)propionate-*2,2,3,3-d*$_4$ at 0.0 ppm for ^1H chemical shifts and to external $(^{15}NH_4)_2SO_4$ at 24.93 ppm for ^{15}N chemical shifts.

Results

Resonance Assignment Strategy. The titration of calcium-saturated calmodulin (CaM) with substoichiometric amounts of the peptide corresponding to the calmodulin-binding domain of the smooth muscle myosin light chain kinase (smMLCKp) results in a ^1H NMR spectrum composed of two distinct sets of resonances. One set corresponds to uncomplexed calmodulin while the other corresponds to the 1:1 complex (Roth et al., 1991). The complex is, on the NMR time scale, in slow exchange with its dissociated components. The ^1H NMR spectrum of 1:1 complex is significantly different than that of CaM alone and, because of the slow exchange properties, requires that the resonance assignment of the complex be done de novo. The basic strategy used follows closely the main-chain-directed assignment strategy outlined for the analysis of two-dimensional ^1H NMR spectra of proteins (Englander & Wand, 1987; Wand & Nelson, 1991). To facilitate the analysis, three-dimensional ^{15}N-edited ^1H–^1H TOCSY and NOESY spectra were relied upon to resolve chemical shift degeneracies associated with the definition of amino acid main-chain subspin systems and amide correlated NOEs, re-

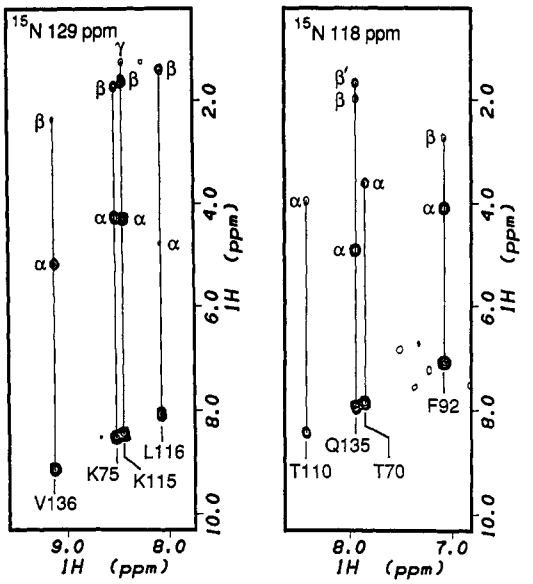

FIGURE 1: Expansions of two-dimensional contour plots at selected ^{15}N frequencies of a three-dimensional TOCSY-HMQC spectrum of the 1:1 complex of uniformly ^{15}N-enriched calmodulin and selectively ^{15}N-labeled smMLCKp peptide in H_2O. A 35-ms spin-locking period was used. These expansions illustrate the use of this spectrum in defining amino acid residue amide NH–C_αH–C_βH subspin systems.

spectively. The basic strategy, nonetheless, followed the main-chain-directed approach where the main-chain amide NH–C_αH–C_βH (NAB) sets of each residue were defined at the outset of the analysis. This was followed by a search for the closed loop MCD patterns corresponding to the various elements of secondary structure recognized by this approach (Englander & Wand, 1987; Wand & Nelson, 1991). To provide sequence alignment of MCD-defined units of secondary structure both standard spin system analysis and analysis of selectively ^{15}N-labeled complexes were used. In what follows, we briefly describe the various arguments and approaches that were used to overcome a number of difficulties in obtaining the resonance assignments listed in Table I.

Definition of NAB Sets. The first step in the MCD approach to the assignment of ^1H resonances of proteins is the definition of the main-chain resonances associated with each amino acid residue (Englander & Wand, 1987; Nelson et al., 1991). The three-dimensional TOCSY-HMQC spectrum of uniformly ^{15}N-labeled calmodulin in complex with the smMLCKp peptide provided the bulk of the NAB set assignments. Using an MLEV-17 mixing sequence of 35 ms, the TOCSY-HMQC provided sufficient resolution and correlations to unambiguously identify amide NH–C_αH resonance pairs for 90% of the residues of calmodulin in the complex. Only about 40% of the companion C_βH resonances could be unambiguously defined in the first pass through this spectrum. Several examples of the correlations obtained between NAB set protons are shown in Figure 1. Comparison of the three-dimensional TOCSY-HMQC and NOESY-HMQC spectra with ^{15}N-decoupled two-dimensional NOESY, TOCSY, and DQF COSY spectra of the complex allowed the number of complete NAB sets to be raised to over 70% of those expected. Many of the remaining ambiguities in the beta hydrogen(s) of incomplete NAB sets were resolved later in the analysis.

Definition of MCD Antiparallel Sheet Patterns. Titration of CaM with peptides corresponding to the skeletal muscle myosin light chain kinase calmodulin-binding domain results in the appearance of downfield-shifted C_αH resonances (Klevit et al., 1985; Seeholzer & Wand, 1989). These have been assigned to the antiparallel β-sheet regions formed by association of the loops of the two EF-hands of each globular domain (Seeholzer & Wand, 1989). Similar spectral changes are seen upon binding of the smooth muscle myosin light chain kinase calmodulin-binding domain peptide (smMLCKp). Given the likelihood that these resonances also arise from residues in the anticipated loop–loop interaction, the definition of antiparallel β-sheet MCD patterns was undertaken before delineation of helical MCD patterns, in contrast to the recommended procedure (Englander & Wand, 1987; Nelson et al., 1991). As was the case for the skMLCKp–CaM complex (Seeholzer & Wand, 1989), hybrid antiparallel sheet MCD patterns were found in both globular domains. An example of the amide NH $\leftrightarrow$ C_αH and C_αH $\leftrightarrow$ C_αH NOEs involved in this NOE pattern are shown in Figures 2 and 3. The residues grouped by this search were placed in the primary sequence on essentially the same basis as before (Seeholzer & Wand, 1989) with additional confirmation provided by comparison to HMQC spectra of complexes prepared with selectively ^{15}N-labeled CaM (see below).

Definition of MCD Helical Patterns. Helical MCD patterns are composed of closed loops of NOE connectivities involving backbone amide NH, C_αH, and C_βH protons (Englander & Wand, 1987; Wand & Nelson, 1991). The most direct manner to search for these patterns while taking full advantage of the resolution of the three-dimensional NOESY-HMQC experiment is to examine each vector along the ^1H chemical shift axis at each ^1H–^{15}N correlation coordinate. This entire set of vectors was examined for amide–amide NOEs which are potential participants in helical MCD NOE patterns. For each vector containing such an NOE, a search was then begun for a second vector which reciprocated the amide–amide NOE. Once found, the two vectors were checked for the remaining connectivities required by the simplest helical MCD pattern (i.e., amide NH $\leftrightarrow$ C_βH). The discovery of the closed loop [$NH_i \leftrightarrow NH_{i+1} \leftrightarrow C_\beta H_i \leftrightarrow NH_i$] connectivity also provides the chain orientation of the involved residues. Examples are shown in Figure 5. This process was continued until exhausted. The assembled blocks of NAB sets were then checked for other inter-NAB set NOEs (Wand & Nelson, 1991) which provided further confirmation of the grouping of NAB sets. At this point, the incompleteness of the definition of the NAB sets, due to degeneracy and/or insufficient long-range connectivities in *J*-correlated spectra and occasional degeneracies of amide NH resonaces of residues adjacent in the primary sequence, results in overlapping helical MCD patterns composed, on average, of four to five residues.

Sequence Alignment of MCD-Defined Helical Regions. Alignment of groups of NAB sets, generated by the first pass through the NOESY-HMQC data set, was undertaken by two approaches. The first involved comparison of HMQC spectra of complexes prepared with selectively ^{15}N-labeled calmodulins to the HMQC spectrum of the uniformly labeled CaM–smMLCKp complex. Calmodulin was selectively labeled with [α-^{15}N]valine (7 residues), alanine (11 residues), leucine (9 residues), glycine (11 residues), and lysine (8 residues). Though not auxotrophic for these amino acids, the *E. coli* strain used did not "scramble" the alanine, leucine, and lysine labels. A small percentage of the labeled nitrogen of valine was converted to alanine (Figure 4). The glycine label was extensively converted to serine (not shown). The HMQC spectra of each selectively labeled sample were compared to the (^{15}N–^1H) coordinates of each NAB set that had been aligned within an MCD-defined helical segment. With respect to the sequence alignment of helical regions, the valine and,

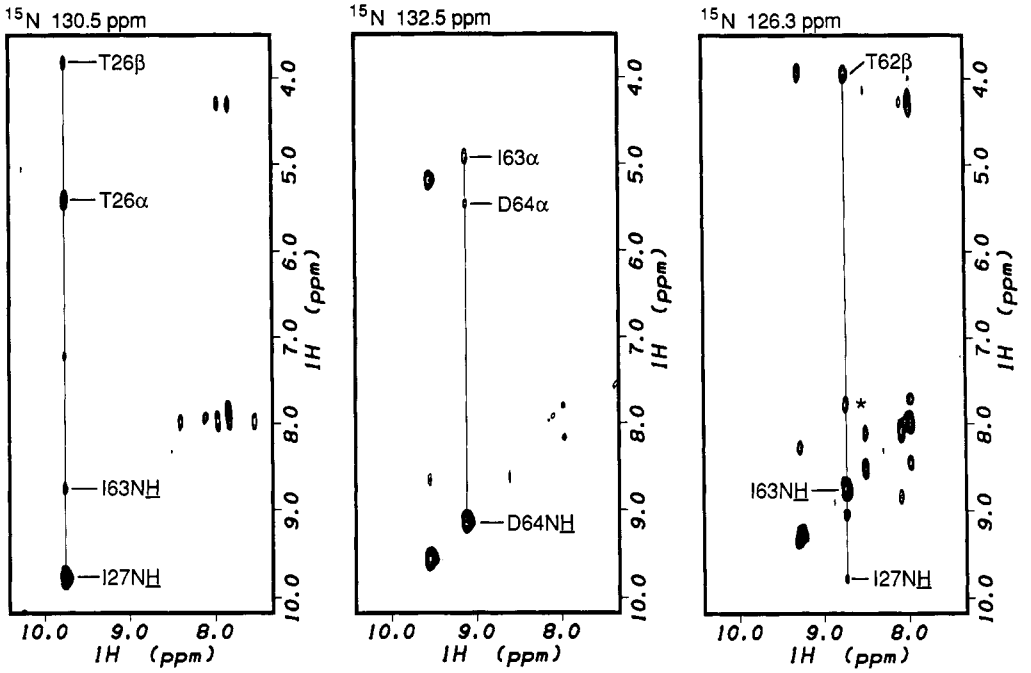

FIGURE 2: Expansions of two-dimensional slices at selected ^{15}N frequencies of the three-dimensional NOESY-HMQC spectrum obtained in H_2O of the 1:1 complex of uniformly ^{15}N-enriched calmodulin and selectively ^{15}N-labeled smMLCKp peptide. The spectrum was obtained with a 50-ms mixing time. Indicated are the intra- and interstrand NOEs between residues I27, T26, I63, and D64 that compose an MCD hybrid NOE pattern. The interstrand $C_\alpha H \leftrightarrow C_\alpha H$ NOE completing the hybrid MCD NOE pattern is shown in Figure 3. Definition of the hybrid pattern may be found in Wand and Nelson (1991).

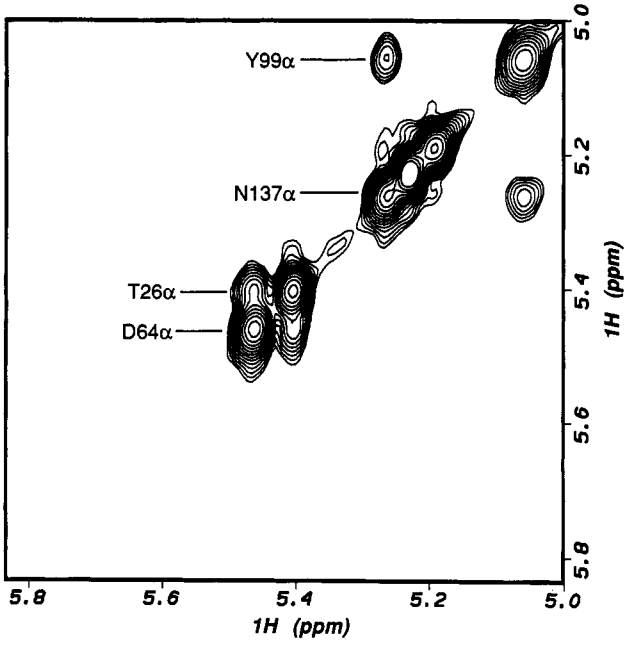

FIGURE 3: Expansion of the two-dimensional NOESY spectrum of the 1:1 complex of uniformly ^{15}N-enriched calmodulin and selectively ^{15}N-labeled smMLCKp peptide. Indicated are the $C_\alpha H \leftrightarrow C_\alpha H$ NOEs of the antiparallel sheets of the globular domains of calmodulin.

in some cases, the alanine labeled complexes proved to be the most immediately useful and directly allowed the alignment of NAB sets to regions of the complex corresponding to helices II, III, V, VI, VII, and VIII of the crystal structure of CaM. Reference to the complex prepared with calmodulin selectively labeled wit glycine (serine) allowed helix II to be completed and helix V to be completed and extended into its C-terminal turn. The [^{15}N]leucine-labeled complex was less useful at this point owing to significant (near) degeneracies of the associated amide ^{1}H resonances. The ^{15}N-labeled lysine complex was also unhelpful at this stage as most labeled lysines did not appear in MCD-defined helical regions.

At this point in the assignment procedure the aligned NAB sets were associated with their side-chain spin systems in two-dimensional ^{1}H–^{1}H *J*-correlated spectra. Many of the correlations observed in the standard two-dimensional ^{1}H–^{1}H TOCSY spectrum of the complex could then be unraveled for a number of the side chains whose identities were now known. In particular, most of the more simple spin systems (e.g., Thr, AMX_2) were classified and associated with their parent NAB sets. These, together with the selectively labeled complexes, then allowed NAB sets to be assigned to all regions of calmodulin except for the C-terminal half of helix I, the N-terminal portion of helix III, and the center residues of helix IV.

Resolution of Degeneracies and Assignment of Regions of Extended Chain Conformation. The placement of the vast majority of NAB sets to specific locations in the primary sequence provides a sufficient set of constraints to resolve the ambiguities arising in remaining helical regions of CaM. Consideration of the amino acid classifications provided by complexes prepared with calmodulin selectively labeled with [^{15}N]alanine and [^{15}N]leucine allowed the tentatively aligned NAB sets in the N-terminal helix to be confirmed. Careful consideration of all available information (MCD patterns, selective labeling, spin system analysis) allowed the nonhelical regions joining MCD-defined helices to be completed. Spectra of complexes prepared with calmodulin labeled with [^{15}N]-lysine were particularly useful at this stage. At this point, NAB sets arising from all residues except for Phe 141 and the seven middle residues of the central helix (residues 78–84) had been assigned. The helical MCD patterns entering (residues 68–76) and exiting (residues 85–97) this region of calmodulin are seen to end abruptly (Figure 6). Interresidue NOEs consistent with an extended chain conformation are seen between the central residues (residues 77–84) of the putative central helix (Figure 6).

Discussion

Through the application of recently developed three-dimensional ^{15}N-separated TOCSY-HMQC and NOESY-

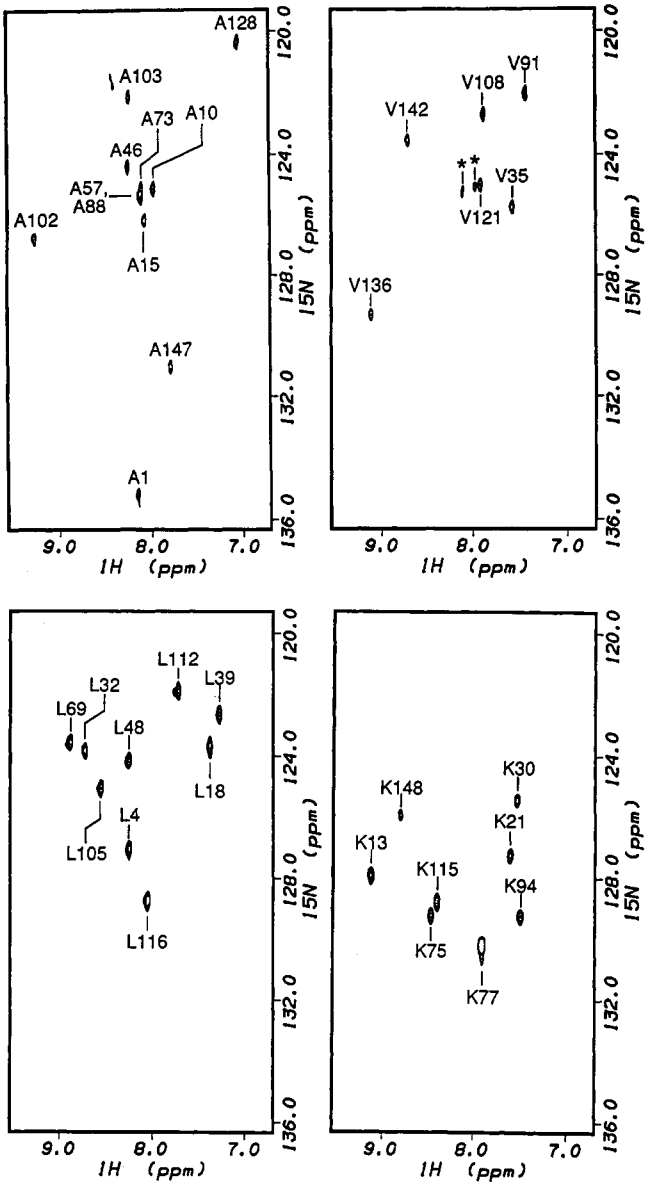

FIGURE 4: Expansions of HMQC spectra of 1:1 complexes of selectively ^{15}N-labeled calmodulins and unlabeled smMLCKp peptide obtained in H_2O. Indicated are the assignments of ^{15}N–^{1}H correlations in spectra of complexes prepared with calmodulin labeled with [^{15}N]alanine, [^{15}N]valine, [^{15}N]leucine, and [^{15}N]lysine. The asterisks indicate [^{15}N]valine that has been metabolized to alanine.

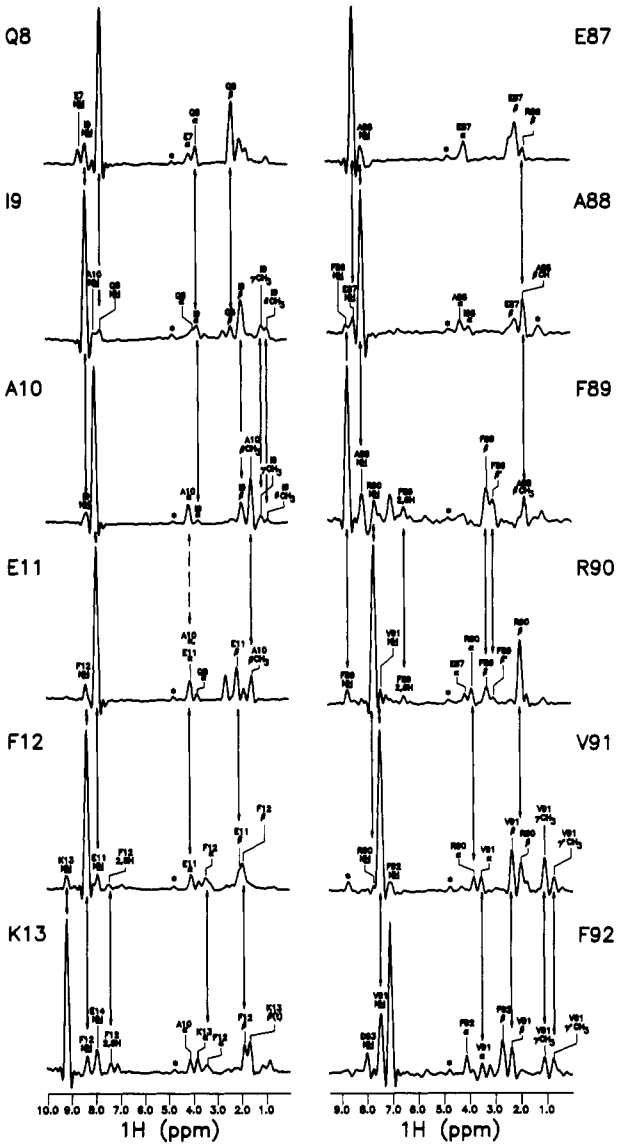

FIGURE 5: Examples of MCD helical connectivities observed in the three-dimensional NOESY-HMQC spectrum of uniformly ^{15}N-enriched calmodulin in 1:1 complex with the smMLCKp peptide in H_2O. A 50-ms mixing period was used. Each panel contains a single vector along the ^{1}H chemical shift axis at the ^{15}N–^{1}H chemical shift coordinate of the indicated amino acid residue. The panels on the right illustrate the helical MCD NOE pattern observed for residues Glu 87–Phe 92 and is a particularly simple case because of the absence of degeneracy amongst the participating NAB sets. A slightly more complicated case is illustrated by the panels on the left where the NOE connectivities involving the main-chain resonances of residues Gln 8–Lys 13 are shown. In this case, a degeneracy of two amide proton resonances (Ala 10 and Glu 11) is overcome by appealing to the long-range (i,i+3) NOEs (i.e., Gln 8 $C_\alpha H$ ↔ Glu 11 NH; Ala 10 $C_\alpha H$ ↔ Lys 13 NH). These NOEs are components of the higher order MCD patterns defined for helices (Wand & Nelson, 1991). The asterisks indicates the residual solvent peak.

HMQC experiments, in conjunction with the main-chain-directed assignment strategy, we have obtained essentially complete assignments for the backbone ^{1}H resonances of chicken calmodulin in complex with a peptide corresponding to the calmodulin-binding domain of the chicken smooth muscle myosin light chain kinase. This approach was successful owing to the resolution provided by the correlation of ^{1}H–^{1}H NOEs with the ^{15}N chemical shift of the bonded amide nitrogen(s) and the robustness of the MCD patterns. Through a large system (ca. 19 kDa), a sufficient number of long-range correlations were provided by the TOCSY-HMQC experiment to undertake the initial definition of NAB sets with a high degree of success. The use of a relatively short mixing time (50 ms) for the NOESY-HMQC experiment resulted in a spectrum which provided the fundamental short distance interactions required by the applied MCD patterns but not so long as to render the spectrum hopelessly confused by extensive spin diffusion and longer range correlations. This strategy takes advantage of the high degree of correlation found between distances defining helical MCD patterns (Wand & Nelson, 1991). The procedure was made difficult only when extended chain conformations were encountered (for which there are no MCD patterns) or when significant amide ^{1}H degeneracies were present in the same unit of secondary structure. However, the use of selective labeling and classical spin system identification served to overcome many of the problems presented and raises the confidence of the resonance assignments presented here to a high level. In particular, the use of selective labeling was particularly important in this respect.

The studies presented here represent the second charac-

Table I: Chemical Shifts of Assigned Resonances of Calmodulin in Complex with the smMLCKp Peptide[a]

residue	chemical shift (ppm)					residue	chemical shift (ppm)				
	^{15}N	N^1H	$C_\alpha{}^1H$	$C_\beta{}^1H$	other ^1H		^{15}N	N^1H	$C_\alpha{}^1H$	$C_\beta{}^1H$	other ^1H
A1	134.96	8.15	4.18	1.51		K75	129.18	8.53	4.28	1.75	
D2	120.22	7.47	4.53	2.29		M76	127.70	8.55	4.74	2.03	
Q3	117.05	7.17	4.12	2.70		K77	130.17	7.98	4.29t		
L4	126.99	8.29	4.67	1.70		D78	123.68	8.06	4.27		
T5	116.84	8.67	4.45	4.75	1.32 γm	T79	122.48	8.79	5.01		1.93 γm
E6	124.10	9.00	3.98	2.05		D80	126.22	7.26	4.66	2.80	
E7	123.40	8.59	4.06	2.03		S81	127.63	8.58	3.46	3.33	
Q8	123.04	7.74	3.79	2.33		E82	119.80	8.66	5.24	2.78	
I9	123.40	8.36	3.72	1.92	1.09 γm	E83	124.81	8.93	4.25	2.60	
A10	125.16	7.98	4.10	1.54		E84					
E11	122.13	7.93	4.07	2.16		I85	125.94	8.10	3.95	2.30	1.18 γm
F12	121.85	8.36	3.45	1.93		R86	126.92	8.50	4.14	1.84	
K13	127.84	9.19	3.82	1.91		E87	122.62	8.48	4.12	2.14	
E14	122.76	7.98	4.48	1.62		A88	125.30	8.15	4.29	1.24	
A15	125.94	8.10	4.29	2.07		F89	122.69	8.72	4.48	3.31, 3.05	
F16	122.48	8.81	3.82	3.35		R90	119.80	7.69	3.87	2.00	
S17	117.69	8.07	4.09	2.76		V91	121.99	7.47	3.55	2.35	1.07, 0.71
L18	123.61	7.38	4.01	1.78		F92	118.60	7.09	4.10	2.73	
F19	121.14	7.94	4.07	2.14		D93	120.30	7.98	4.50		
D20	127.06	8.45	3.88	2.69t		K94	129.18	7.55	3.91	1.84	
K21	127.20	7.66	3.98	1.89		D95	117.69	8.26	4.56	2.65	
D22	117.69	8.06	4.56			G96	113.10	7.82	3.85		
G23	112.82	7.66	3.87			N97	123.54	8.39	4.37	1.97	
D24	124.67	8.45	4.48	2.46t		G98	116.42	10.63	4.07	3.44	
G25	116.56	10.58	4.23	3.44		Y99	119.80	7.66	5.04	2.51	
T26	116.35	8.15	5.35	3.76	1.02 γm	I100	130.94	10.27	4.77	2.00	
I27	130.45	9.75	4.80	1.77t		S101	127.56	8.99	5.04	4.03	
T28		8.26	4.91			A102	126.57	9.29	3.91	1.50	
T29	115.78	9.27	3.79	4.25	1.31 γm	A103	122.13	8.26	4.03	1.45	
K30	125.58	7.60	4.15	1.84		E104	124.03	7.98	4.44	2.21	
E31	125.79	7.76	3.98	2.35		L105	125.02	8.58	4.15	1.97	
L32	123.68	8.75	4.18	1.86		R106	122.27	8.75	3.90	1.91	
G33	109.37	8.81	3.96	3.58		H107	122.83	8.02	4.28	2.30	
T34	121.63	8.01	3.96	4.42	1.26 γm	V108	122.62	7.88	3.41	1.97	0.83, 0.36
V35	125.58	7.60	3.71	2.05	0.90, 0.58	M109	119.24	8.15	4.01	2.52	
M36	121.00	8.39	3.84	2.00t		T110	118.25	8.44	3.99	4.21	1.20 γm
R37	121.56	8.72	3.68	1.97		N111	126.50	7.98	4.36	2.73	
S38	122.34	8.06	4.20	4.11		L112	121.85	7.74	4.04	1.81	
L39	122.48	7.27	4.36	1.73		G113	107.96	7.52	4.29	3.60	
G40	110.42	7.79	4.34	3.77		E114	125.37	7.96	4.29	1.65	
Q41	122.48	7.83	4.07	2.16		K115	128.68	8.45	4.33	1.67	
N42	123.63	8.59	3.49	3.35		L116	128.76	8.07	4.75	1.50	
P43						T117	117.47	8.96	4.50		1.29 γm
T44	116.63	8.81	4.48	4.04t	1.34 γm	D118	124.03	7.74	4.02	2.37	1.86
E45	122.48	8.80	4.03	2.03		E119	123.19	8.70	4.13	2.05	
A46	124.24	8.29	4.09	1.39		E120	121.92	8.40	4.04		
E47	122.62	7.77	4.04	2.37		V121	125.16	7.96	3.42	2.24	0.99, 0.91
L48	124.24	8.29	4.37	2.05t		D122	123.68	8.02	4.31	2.73	
Q49	121.85	8.25	3.79	2.18		E123	123.19	8.17	3.98	2.16	
D50	124.03	7.98	4.04	2.78		M124	122.62	7.87	3.96	1.99	
M51	122.97	8.06	4.09	2.30		I125	124.67	8.21	3.87	2.40	
I52	121.00	7.68	3.42	1.89		R126	121.63	8.55	4.04	1.89	
N53	120.86	8.50	4.44	2.98		E127	120.51	8.01	3.99	2.22	
E54	119.80	7.36	4.06	2.13		A128	120.30	7.08	4.29	1.05	
V55	112.47	7.11	4.40	2.26	0.88	D129	121.63	8.01	4.42	2.84	
D56	125.65	7.55	4.53	2.71		I130	131.72	8.25	3.85	1.92	
A57	125.30	8.15	4.29	1.83		D131	120.30	8.36	4.52		
D58	117.69	8.27	4.75	3.91		G132	112.40	7.60	3.96	3.82	
G59	112.40	7.60	3.98	3.82		D133	124.67	8.36	4.45	2.97	
N60	122.62	8.17	4.59	2.65		G134	116.42	10.30	4.06	3.42	
G61	116.56	10.58	4.37	3.74		Q135	118.60	7.96	4.88	1.99, 1.72	
T62	112.61	7.69	4.77		1.11 γm (t)	V136	129.25	9.13	5.19	2.40	1.09, 0.77
I63	126.29	8.72	4.96	1.99		N137	132.92	9.56	5.19		
D64	132.56	9.13	5.46	2.87		Y138	121.99	8.34	3.24	2.33, 2.05	
F65	123.61	9.13	3.95	2.86		E139	122.34	8.07	3.57	2.00	
P66			3.72	2.26	1.99 γ, 3.81 δ	E140	123.61	8.70	3.84		
E67	126.57	7.98	4.17	2.73t		F141					
F68	126.99	8.45	3.87	3.46, 3.06		V142	123.68	8.75	2.95	1.81	0.67, 0.41
L69	123.54	8.91	3.28	1.24, 0.97		Q143	121.99	7.50	3.80	2.00	
T70	118.74	7.87	3.61	4.12	1.23 γm	M144	120.22	7.23	4.25	1.58	
M71	123.96	7.19	3.88	1.84t		M145	117.69	7.60	4.26	1.62	
M72	119.80	7.66	4.15	2.00		T146	114.23	7.52	4.78	4.29	112 γm
A73	125.16	8.15	4.23	1.29		A147	131.01	7.83	4.31	1.39	
R74	122.20	6.71	3.98	1.72		K148	125.94	8.89	4.34	1.91	

[a] ^1H Chemical shifts referenced to external 3-(trimethylsilyl)propionate-*2,2,3,3-d*$_4$ (coaxial capillary) at 0.0 ppm. ^{15}N Chemical shifts are referenced to external $^{15}NH_4Cl$ at 24.93 ppm. Chemical shifts are for *E. coli* expressed chicken calmodulin in the 1:1 complex with the smMLCKp peptide in 10 mM deuterated imidazole and 40 mM Ca^{2+}, at pH 6.5 and 25 °C. Tentative assignments are noted with a "t". When ambiguity exists, methyl resonances are indicated with an "m".

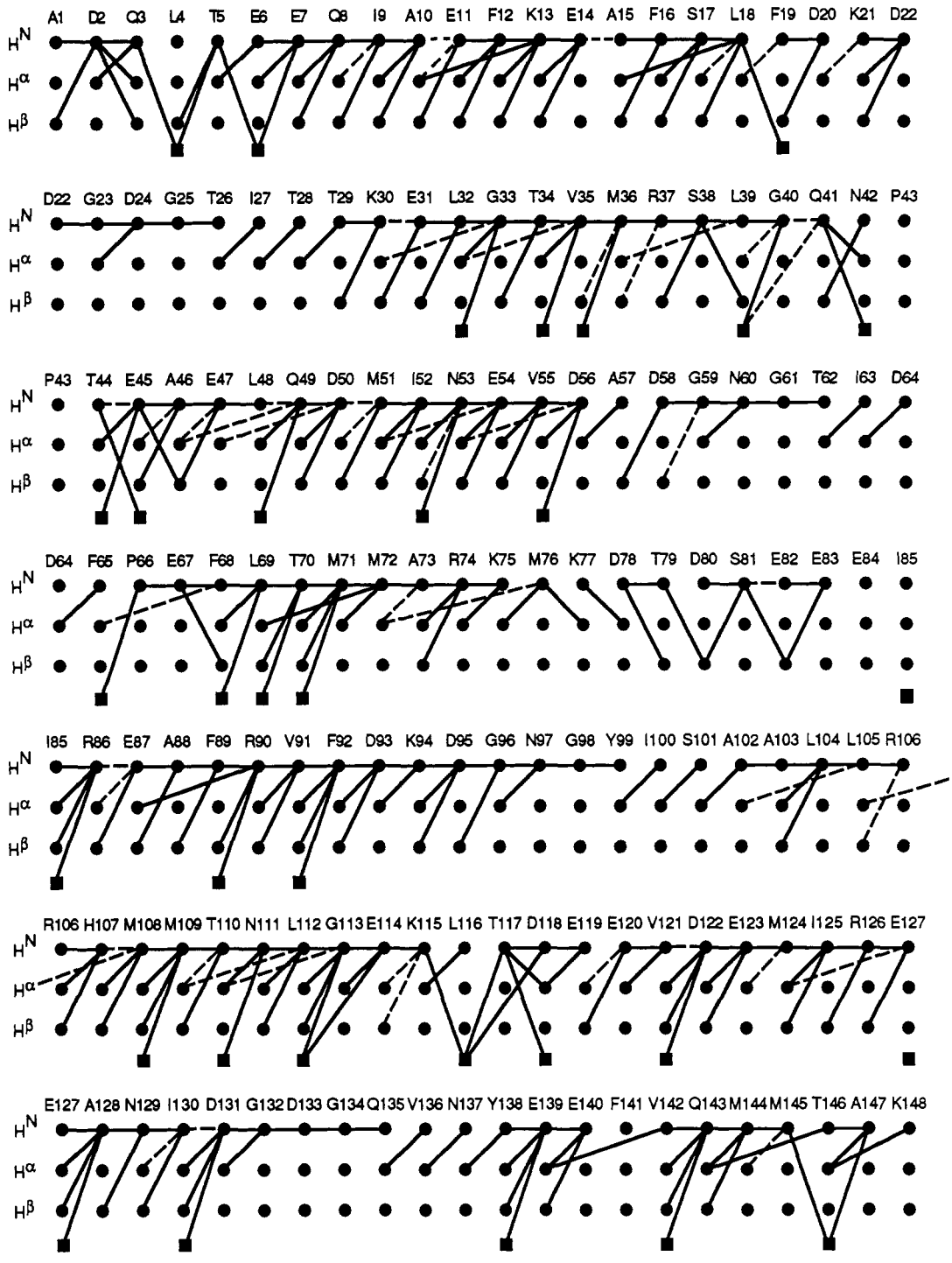

FIGURE 6: Schematic summary of the NOEs used in the assignment of calmodulin in complex with the smMLCKp peptide. Solid lines indicate unambiguous NOE connectivities. Dashed lines indicate the presence of degeneracy of one or both of the participating resonances. Individual NAB set members are represented by sold circles and are labeled as amide, α, and β. Squares indicate NOEs to other resonances of the amino acid residue that were paricularly helpful in resolving ambiguities in the sequence specific assignment. When used, the δ-hydrogen of proline is placed at the amide position in the diagram. For clarity, intraresidue NOEs have been omitted.

terization of the secondary structure of calmodulin in complex with a peptide. A parallel study of the homologous rabbit skeletal muscle myosin light chain kinase calmodulin-binding domain peptide in complex with *Drosophila melanogaster* calmodulin has recently been reported by Ikura et al. (1991). Using a distinctly different assignment methodology, these workers reported main-chain amide ^{15}N and ^{1}H, α-^{13}C and ^{1}H, and carbonyl ^{13}C resonance assignments. After adjusting for a small systematic difference, the corresponding chemical shift assignments reported here are quite similar throughout the globular regions of calmodulin. Significant differences are seen only in the region of residues 76–85. This is somewhat surprising as calmodulin is a highly conserved protein and calmodulin from chicken and *D. melanogaster* differ at only three residues [Y(F)99, Q(T)143, A(S)147], all of which are remote from the central helix. These chemical shift differences may therefore reflect local effects of the differing sequences of the bound peptide used, in particular, the presence of a histidine in the center of the smMLCKp peptide. It should also be mentioned that, to our knowledge, the CaM-binding domain of the *D. melanogaster* MLCK has not been defined. This also raises the possibility that the complex between the rabbit skeletal muscle MLCK CaM-binding domain peptide and calmodulin from *D. melanogaster* may be somewhat

different than the complex studied here.

The resonance assignments presented in Table I were obtained, in part, by analysis of the interresidue NOEs observed among main-chain amide NH–C_αH–C_βH protons. Keeping in mind the relatively short mixing time used to generate the NOE, one can define the helical regions of CaM in the complex by the presence of consecutive and overlapping fundamental helical MCD patterns. Using this criterion, helices are found to span residues 6–20, 29–38, 44–55, 68–76, 85–95, 106–113, 121–131, and, perhaps, 138–144. Though the boundaries differ slightly, these regions of helicity correspond, with one exception, quite well to the MCD-defined helical regions predicted by the crystal structure of CaM (Babu et al., 1988). In addition, as was determined previously for complexes of CaM with peptides corresponding to the skeletal MLCK calmodulin-binding domain (Seeholzer & Wand, 1989; Ikura et al., 1991), the antiparallel sheet structures associated with the calcium-binding domains of CaM remain intact in the CaM–smMLCKp complex. One may conclude, therefore, that the fundamental assumptions regarding the correspondence of the structure of the globular domains of CaM in isolation and in complex with the skeletal muscle MLCK CaM-binding domain peptide made in recent modeling studies (Persechini & Krestinger, 1988; O'Neil & DeGrado, 1990) are applicable to the CaM–smMLCKp complex. There is only one significant difference between the secondary structure of CaM and that of the CaM–smMLCKp complex, and it involves residues 76–84 of the "central helix". In this region of the CaM–smMLCKp complex, the absence of the [$NH_i \leftrightarrow C_\beta H_i \leftrightarrow NH_{i+1} \leftrightarrow NH_i$] fundamental helical MCD NOE patterns is inconsistent with even a distorted helical conformation [see Wand and Nelson (1991) and Nelson et al. (1991)]. Furthermore, the presence of several $NH_i \leftrightarrow C_\alpha H_{i+1}$ NOEs in this region clearly indicates an extended chain conformation in contrast to the generally α-helical conformation seen in the crystal (Babu et al., 1988) and solution (Ikura et al., 1990) structures of CaM. This behavior is also seen in the complex between the rabbit skeletal muscle MLCK CaM-binding domain peptide and calmodulin from *D. melanogaster*, though there are subtle differences in the pattern of main-chain NOEs involving residues 78–80.

The resonance assignments presented here, in conjunction with those obtained for the bound smMLCKp peptide (Roth et al., 1991), provide a basis for searching for contacts between the peptide and CaM. At this time, NOEs between the ring protons of Trp 5, the methyls of Ala 14 and Leu 18 of the peptide and the methyls of Val 108, and the ring of Phe 68 and the α- and β-hydrogens of Val 35 of CaM, respectively, have been identified. The short distances indicated by these NOEs are consistent with the symmetric variant of the type III complex described by Persechini and Krestinger (1988) (i.e., the bound peptide has an opposite orientation to that of the type III complex described) and are essentially the same as the model recently proposed by O'Neil and DeGrado (1991). However, our preliminary modeling studies indicate that, using the structure determined for the bound smMLCKp peptide (Roth et al., 1991), a more significant distortion of the central helix than used by Persechini and Krestinger (1988) is required to meet the requirements of these NOE contacts. Greater distortion of the central helix is required by the fact that, although generally helical, the backbone of the peptide does display local regions of nonhelical conformation. This, in turn, results in the need for larger deformations of the central helix in order to meet the constraints of the CaM–smMLCKp NOEs listed above and the structure determined for the peptide. This larger distortion of the central helix, relative to previous modeling studies, is consistent with the pattern of NOEs seen in the central region of CaM where the backbone confirmation is unequivocally nonhelical in nature. Finally, it is also interesting to note that the orientation of smMLCKp peptide bound to calmodulin determined here is opposite to that found for melittin bound to CaM (Seeholzer et al., 1987).

ACKNOWLEDGMENTS

We are indebted to Dr. Yiqing Feng for undertaking the preliminary modeling studies of the smMLCKp–CaM complex. We are grateful to Dr. D. Hansburg and S. Nakajima for preparation of the smMLCKp peptides and to Dr. R. Katz for providing the cI plasmid.

REFERENCES

Babu, Y. S., Bugg, C. E., & Cook, W. J. (1988) *J. Mol. Biol. 204*, 191–204.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Bax, A., Griffey, R. H., & Hawkins, B. L. (1983) *J. Am. Chem. Soc. 105*, 7188–7190.

Blumenthal, D. K., Takio, K., Edelman, A. M., Charbonneau, H., Titani, K., Walsh, K. A., & Krebs, E. G. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3187–3191.

Englander, S. W., & Wand, A. J. (1987) *Biochemistry 26*, 5953–5958.

Gopalakrishna, R., & Anderson, W. B. (1982) *Biochem. Biophys. Res. Commun. 104*, 830–836.

Heidorn, D. B., Seeger, P. A., Rokop, S. E., Blumenthal, D. K., Means, A. R., Crespi, H. L., & Trewhella, J. (1989) *Biochemistry 28*, 6757–6764.

Ikura, M., Kay, L. E., & Bax, A. (1990) *Biochemistry 29*, 4659–4667.

Ikura, M., Kay, L. E., Krinks, M., & Bax, A. (1991) *Biochemistry 30*, 5498–5504.

Kemp, B. E., Pearson, R. B., Guerriero, V., Jr., Bagchi, I. C., & Means, A. R. (1987) *J. Biol. Chem. 262*, 2542–2548.

Klevit, R. E., Blumenthal, D. K., Wemmer, D. E., & Krebs, E. G. (1985) *Biochemistry 24*, 8152–8157.

Lukas, T. J., Burgess, W. H., Prendergast, F. G., Lau, W. & Watterson, D. M. (1986) *Biochemistry 25*, 1458–1464.

Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95–117.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

McDowell, L., Sanya, G., & Prendergast, F. G. (1985) *Biochemistry 24*, 2979–2984.

Means, A. R. (1988) *Recent Prog. Horm. Res. 44*, 223–261.

Nelson, S. J., Schneider, D. M., & Wand, A. J. (1991) *Biophys. J. 59*, 1113–1122.

O'Neil, K. T., & DeGrado, W. F. (1990) *Trends Biochem. Sci. 15*, 59–64.

O'Neil, K. T., Wolfe, H. R., Jr., Erickson-Viitanen, S., & DeGrado, W. F. (1987) *Science 236*, 1454–1456.

Persechini, A., & Kretsinger, R. H. (1988) *J. Cardiovasc. Pharmacol. 12* (Suppl. 5), S1–S12.

Putkey, J. A., Slaughter, G. R., & Means, A. R. (1985) *J. Biol. Chem. 260*, 4704–4712.

Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.

Roth, S. M., Schneider, D. M., Strobel, L. A., van Berkum, M. B., Means, A. R., & Wand, A. J. (1991) *Biochemistry 30*, 10078–10084.

Sambrook, J., Fritsch, E. F., & Maniatas, T. (1989) *Molecular Cloning*, 2nd ed., p A.3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Seeholzer, S. H., & Wand, A. J. (1989) *Biochemistry 25*, 4011–4020.
Seeholzer, S. H., Cohn, M., Wand, A. J., Crespi, H. L., Putkey, J. A., & Means, A. R. (1987) in *Calcium-Binding Proteins in Health and Disease*, pp 360–371, Academic Press, New York.
Wand, A. J., & Nelson, S. J. (1991) *Biophys. J. 59*, 1101–1112.

Localization of the Binding Site for Streptococcal Protein G on Human Serum Albumin. Identification of a 5.5-Kilodalton Protein G Binding Albumin Fragment†

Cecilia Falkenberg, Lars Björck, and Bo Åkerström*
Department of Medical and Physiological Chemistry, P.O. Box 94, University of Lund, S-221 00 Lund, Sweden
Received May 21, 1991; Revised Manuscript Received September 23, 1991

ABSTRACT: Protein G is a streptococcal cell wall protein with separate and repetitively arranged binding domains for immunoglobulin G (IgG) and human serum albumin (HSA). In this work, the binding of protein G to HSA was studied. The results suggest that a single binding site is present on HSA: the apparent size of the HSA–protein G complex (230 kDa) corresponded to two or three HSA molecules bound to one protein G molecule, and Ouchterlony immunodiffusion did not yield any precipitate between protein G and HSA. HSA was cleaved by pepsin and CNBr into several fragments which were identified by SDS–PAGE and N-terminal amino acid sequencing, and the binding of protein G to the fragments was studied in Western blot experiments. The results indicated that the binding area was located in disulfide loops 6–8, involving both the second (loop 6) and the third (loops 7 and 8) domain of HSA. One of the protein G binding pepsin fragments, with an apparent molecular mass of 5.5 kDa, located in loops 7 and 8, was isolated and found to completely inhibit the binding between protein G and the intact HSA, again suggesting a single protein G binding site on serum albumin. Reducing the disulfide bonds of HSA, and subsequent alkylation of the half-cystine residues, significantly decreased the affinity for protein G. Protein G bound to albumin from baboon, cat, guinea pig, hamster, hen, horse, man, mouse, and rat, but not to albumin from cow, dog, goat, pig, rabbit, sheep, snake, or turkey.

Protein G is an immunoglobulin G (IgG)[1] binding protein expressed by group C and G streptococci (Björck & Åkerström, 1990). The protein was originally enzymatically solubilized from the streptococcal cell wall (Reis et al., 1984; Björck & Kronvall, 1984), but the protein G gene has also been expressed in *Escherichia coli* (Guss et al., 1986; Fahnestock et al., 1986; Björck et al., 1967). Comparative studies (Åkerström et al., 1985; Guss et al., 1986; Åkerström & Björck, 1986) showed that the IgG Fc binding properties of protein G are similar to those of protein A, the IgG-binding protein of *Staphylococcus aureus* (Forsgren & Sjöquist, 1966; Langone, 1982), although protein G has a wider range of IgG-binding activity among mammalian species. The physicochemical properties of proteins A (Langone, 1982) and G (Åkerström & Björck, 1986) indicate that they are both fibrous proteins and the binding site on IgG Fc is identical or very similar for the two proteins (Stone et al., 1989). However, the IgG-binding domains of proteins A and G showed no sequence homology (Uhlén et al., 1984; Guss et al., 1986; Fahnestock et al., 1986). Apart from IgG, protein G also shows affinity for human serum albumin (HSA) (Björck et al., 1987). On the protein G molecule, HSA binding was found to be separate from IgG binding and located to repeated domains in the N-terminal half of the streptococcal protein, whereas the IgG-binding domains reside in the C-terminal half (Åkerström et al., 1987; Sjöbring et al., 1988). This unique organization of protein-binding domains will allow the protein G expressing *Streptococcus* to cover itself with an outer layer of albumin and an inner layer of IgG, which should influence the host–parasite relationship during infections with these bacteria.

In the present work, we have investigated the binding between protein G and HSA in order to localize the binding site on the HSA molecule. Albumin has been described as a cigar-shaped protein, consisting of three spherical domains. The domains are approximately equal in size, and each contains six disulfide bonds except domain 1 which contains five. Together these form disulfide loops 1–9 [for a review, see Peters (1985)]. On the basis of the exon arrangement of the gene, each domain is also divided into two subdomains (Minghetti et al., 1986). The three-dimensional structure of HSA has been determined, and confirms the model with three domains and six subdomains (Carter et al., 1989). By the position and size of the loops, the three-dimensional structure and, by amino acid sequence comparison, the three domains are homologous. In this work, HSA was fragmented by pepsin and CNBr treatment, and the binding of protein G to each fragment was analyzed. The results suggest that protein G binds to a single site located in the second and third domains. The mapping of the binding to this region is described, and

† This work was supported by grants from the Swedish Medical Research Council (Projects 7144 and 7480), the Gustav V:s 80-year Foundation, the Foundations of A. Österlund, Crafoord, G. and J. Koch, and O. E. and Edla Johansson, the Swedish Society for Medical Research, the Royal Physiographic Society, Läkaresällskapet, the Medical Faculty, University of Lund, and Excorim KB, Lund.
* To whom correspondence should be addressed.

[1] Abbreviations: HSA, human serum albumin; SPRIA, solid-phase radioimmunoassay; Ig, immunoglobulin; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; PBS, phosphate-buffered saline; kDa, kilodalton(s).

0006-2960/92/0431-1451$03.00/0