

Determination of the Side Chain pK_a Values of the Lysine Residues in Calmodulin*

(Received for publication, May 28, 1993, and in revised form, July 6, 1993)

Mingjie Zhang and Hans J. Vogel‡

From the Department of Biological Sciences, The University of Calgary, 2500 University Dr. NW, Calgary, Alberta T2N 1N4, Canada

The 7 Lys residues in mammalian calmodulin (CaM) were reductively methylated with ^{13}C -enriched formaldehyde and studied by (^1H , ^{13}C)-heteronuclear multiple quantum coherence (HMQC) NMR. The apo- and Ca^{2+} -forms of CaM, as well as a complex with a 22-residue peptide which comprises the CaM binding region of myosin light chain kinase were studied. The complete assignment of the two-dimensional NMR spectra was obtained by site-directed mutagenesis (Lys→Gln) of all the Lys. The pK_a values for the individual Lys could be determined by pH titration experiments. In Ca^{2+} -CaM, the pK_a values range from 9.29 (Lys-75) to 10.23 (Lys-77). The Lys in apo-CaM have higher pK_a values than those in Ca^{2+} -CaM. The binding of the myosin light chain kinase peptide gives rise to an increase of the pK_a values of Lys-148 and Lys-75 by 0.5 and 0.8 pH units, respectively; this results from the relocation of their side chains to a completely solvent accessible state. The changes in the pK_a values upon binding Ca^{2+} or the myosin light chain kinase peptide show a remarkable correlation with earlier reported chemical reactivity changes. Thus, our results indicate that pK_a values, rather than structural and steric effects, play the dominant role in determining the reactivity of Lys side chains towards small electrophilic chemical modification reagents. The methodology used here could prove useful for the determination of individual pK_a values in other proteins.

Calmodulin (CaM)¹ is a ubiquitous acidic Ca^{2+} -binding protein of 148 amino acids that is found in all eukaryotic cells. The protein can bind and regulate various target enzymes in a Ca^{2+} -dependent manner (for reviews, see Klee and Vanaman (1982), Forsén *et al.* (1986), Hiraoki and Vogel (1987), and Means *et al.* (1991)). The binding of calcium induces major conformational changes which allow the protein to interact with its target enzymes. The three-dimensional structure of

the Ca^{2+} -form has been determined by high resolution x-ray methods (Babu *et al.* 1988; Chattopadhyaya *et al.*, 1992; Rao *et al.*, 1993). The protein was shown to be a dumbbell shaped molecule with the two domains of the protein linked by a long solvent exposed α -helix.

Mammalian CaM contains a total of 7 lysines and 1 trimethyllysine (Lys-115) residue; the side chains of these Lys residues are all located on the surface of the protein (Babu *et al.*, 1988). The ϵ -amino groups of the 7 Lys residues in CaM all show distinct reactivities towards chemical modification reagents (Giedroc *et al.*, 1985, 1987). Chemical modification of Lys residues can also influence the ability of CaM to activate enzymes. For example, acetylation of the Lys residues gives rise to a decreased affinity of CaM for calcineurin (Manalan and Klee, 1987), while carbamylation decreases the ability of CaM to activate phosphodiesterase, but has no effect on the activation of adenylate cyclase (Guerini *et al.*, 1987). Trace labeling techniques were also used to study the possible involvement of Lys residues when CaM interacts with target enzymes (Jackson *et al.*, 1986; Manalan and Klee, 1987; Winkler *et al.*, 1987). These chemical modification studies have provided useful information about changes in the reactivity of Lys residues in CaM upon the addition of calcium or target proteins. However, it is difficult to derive a detailed interpretation from these earlier results since it is not known whether the changes in reactivity were caused by adjustments of the pK_a or by structural changes. In an attempt to rectify this situation, we embarked on studies aimed at determining the individual pK_a values of all the Lys residues in CaM in its apo-, calcium-saturated and target protein bound states.

The pK_a value of the positively charged Lys side chain is a key property for a Lys residue in a protein, since it directly reflects the participation of the residue in salt linkages, hydrogen bonding, or other kinds of electrostatic interactions, etc. (Fersht, 1985; Burley and Petsko, 1988; Sancho *et al.*, 1992). Calculations of the pK_a value of Lys residues in proteins have been attempted for some time (Tanford and Roxby, 1972) and continues to be an active field of research (Bashford and Karplus, 1990; Yang *et al.*, 1993). However, accurate calculated values for the specific pK_a values of Lys residues in a protein have not been reported to date, which prompted us to pursue a new experimental approach.

In this work, reductive methylation, heteronuclear two-dimensional NMR spectroscopy, and site-directed mutagenesis were combined to determine the pK_a values of the individual Lys residues. The chemical modification served to introduce ^{13}C -labeled methyl groups into the Lys side chains. Heteronuclear two-dimensional NMR was used to give NMR spectra with good resolution and sensitivity. Following pH titration, the resonances and pK_a values were assigned to specific Lys residues with the help of site-directed mutagenesis. The methodology was critically assessed by studying the

* This work was financially supported by the Medical Research Council of Canada (MRC). The NMR spectrometer was funded by MRC and Alberta Heritage Foundation for Medical Research. The modeling computer was purchased with funds provided by the Erna and Victor Hasselblad Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Scholar of the Alberta Heritage Foundation for Medical Research (AHFMR). To whom correspondence should be addressed: Dept. of Biological Sciences, The University of Calgary, Calgary, Alberta, Canada, T2N 1N4. Tel.: 403-220-6006; Fax: 403-289-9311.

¹ The abbreviations used are: CaM, calmodulin; HMQC, heteronuclear multiple quantum coherence; K13Q-CaM, Lys-13 to Gln mutation of CaM etc.; MLCK, myosin light chain kinase; MOPS, 4-morpholinepropanesulfonic acid.

structural and dynamic effects of the chemical modification on CaM. The experimental approach described in this work provides a convenient and reliable technique to obtain the pK_a values for individual Lys residues in proteins. Here we report the determination of the pK_a values of all the Lys residues in apo- and Ca²⁺-CaM, and in a complex of CaM with a 22-residue peptide comprising the CaM binding region of skeletal muscle myosin light chain kinase (MLCK). We have found that the pK_a values play a critical role in the chemical reactivity of the Lys residues in CaM.

MATERIALS AND METHODS

Bovine testicles were purchased from Pel Freeze and stored at -80 °C before use. Carbon-13 enriched (99%) formaldehyde and ¹⁵N^α-Lys were purchased from MSD (Montreal, Canada). NaCNBH₃ is a product of Sigma. All the reagents used for site-directed mutagenesis were the products of Life Technologies Inc. or New England Biolabs. Bovine CaM was purified following standard procedures described in the literature (Andersson *et al.* 1983; Vogel *et al.*, 1983). CaM and its mutants were expressed in *Escherichia coli* and were purified according to the methods described by Putkey *et al.* (1986). The purity of the protein samples was greater than 95% as judged by SDS-polyacrylamide gel electrophoresis. The concentration of CaM was determined by UV absorption using $\epsilon_{276}^{1\%} = 1.8$.

The 22-residue peptide, KRRWKKNFIAVSAANRFKIKISS, which encompasses residues 577–598 of the amino acid sequence of skeletal muscle MLCK and is known to be its CaM binding domain (Blumenthal *et al.*, 1988), was synthesized by the Core Facility for Protein/DNA Chemistry, Queen's University, Canada. The purity of the peptide was >95% as judged by high pressure liquid chromatography and amino acid analysis. The concentration of the peptide was determined by UV absorption of the single Trp residue in the peptide ($\epsilon_{280}^{1\%} = 5.6 \times 10^3 \text{ cm}^2 \cdot \text{mol}^{-1}$).

Carbon-13 Methylation of CaM Samples—The reductive methylation of CaM samples with ¹³C-enriched (99%) formaldehyde (MSD) has been described by Jentoft and Dearborn (1979, 1983). Briefly, 10 mg of CaM was dissolved in 4 ml of 50 mM HEPES buffer, 10 mM CaCl₂, pH 7.5. Carbon-13-labeled formaldehyde and freshly prepared NaCNBH₃ (1 M stock solution) were added to this solution and the mixture was shaken gently and incubated overnight at 4 °C. NaCNBH₃ and [¹³C]formaldehyde were added at a 5–10-fold molar excess over the free amino groups of the protein. The reaction was stopped by extensive dialysis against 10 mM NH₄HCO₃, and the protein was subsequently freeze dried and stored at -20 °C prior to NMR studies.

Expression of CaM in *E. coli*—The synthetic gene encoding bovine CaM was a generous gift from Dr. T. Grundström (University of Umeå, Sweden). The gene was constructed with codons optimized for expression in *E. coli*. The plasmid carrying the CaM gene is a temperature sensitive "run-away" plasmid (Uhlen *et al.*, 1983; Brodin *et al.*, 1989). The expression of the gene was under the control of the *lacZ* promoter. The *E. coli* strain, MM294, was used as the host. To express CaM, the bacterial cells were grown at 30 °C in the presence of 50 µg/ml ampicillin in L-broth medium (or a suitable minimal medium, see below) until the A₆₀₀ reached ~1.5. Then an appropriate amount of prewarmed fresh L-broth medium was added into the above culture to bring the temperature to 37 °C. At the same time 160 mg/liter isopropyl-β-D-thiogalactopyranoside was added to the culture to induce the expression of protein. The culture was maintained at 37 °C for another 3–4 h, following which the cells were collected by centrifugation.

Selective Labeling of CaM with ¹⁵N^α-Lys—Selective labeling of CaM was achieved in a chemically defined MOPS minimal medium. The medium was made up as described in the literature except that the amino acid Lys was left out (Neidhardt *et al.*, 1974; Wanner *et al.*, 1977). Instead, ¹⁵N^α-Lys (50 mg/liter) was added to the growth medium together with isopropyl-β-D-galactopyranoside when the temperature of the culture was switched to 37 °C (3–4 h).

Site-directed Mutagenesis of CaM—The oligonucleotides were synthesized on a "Gene Assembler Plus" DNA synthesizer (Pharmacia LKB Biotechnology Inc.) and purified following the instructions provided by the manufacturer. All the mutations except Lys-13 were carried out in the pBluescript (pBSTM, Stratagene) plasmid in which the *KpnI*/*SacI* fragment of the CaM gene was subcloned. Since restriction sites are conveniently distributed throughout the CaM

gene (Fig. 1), the polymerase chain reaction could be used to generate all the mutations. The experimental procedures used were essentially those outlined by Kadowaki *et al.* (1989). The mutation of Lys-13 was carried out directly on the original run-away plasmid by utilizing the unique restriction sites of *XhoI* and *KpnI* on the plasmid. The mutations were identified by sequencing individual clones containing the polymerase chain reaction-generated DNA fragments which also provided for a check against any possible misincorporations that may have occurred during the polymerase chain reaction.

Sample Preparation for NMR Studies—Apo-CaM was prepared by passing a CaM solution through a Chelex-100 column equilibrated with 100 mM NH₄HCO₃ buffer, pH 8.0. About 10 mg of methylated CaM was dissolved in 0.4 ml of 99.9% D₂O containing 0.15 M KCl. The Ca²⁺-form of CaM was prepared by the addition of an appropriate amount of Ca²⁺ from a 0.25 M CaCl₂ stock solution to apo-CaM; this produced samples with 4.1 eq of Ca²⁺. The final volume of the methylated apo- and Ca²⁺-CaM samples for NMR analysis was approximately 0.45 ml. A complex of methylated CaM with an equimolar amount of the MLCK peptide was also prepared for NMR. A sample containing about 8 mg of methylated Ca²⁺-CaM was first prepared. Following adjustment of the pH to 7.5, 0.32 ml of a 1.5 mM MLCK peptide stock solution, pH 7.5, was added to the methylated CaM dropwise with gentle mixing. Subsequently, the volume of the CaM/MLCK peptide was reduced to 0.4 ml by spin vacuum drying without freezing. The complexes of methylated Lys CaM mutants (K13Q-, K21Q-, K30Q-, K75R-, K77Q-, K94Q-, and K148Q-CaM) with the MLCK peptide were prepared in a similar manner except that the final concentration of the protein and the peptide was about 0.4 mM.

The pH titrations were carried out by adding microliter amounts of 0.05–0.5 M KOD or DCl to the protein samples. The pH values were directly read from a pH meter using an Ingold electrode. No corrections were made for the isotope effect. The pH titrations typically covered the pH range of 2.0–12.0. Earlier studies have shown that CaM is stable over this pH range (Huque and Vogel, 1993). A small amount of ¹³CH₃OH (≈1 mM) was added to each sample to provide an internal chemical shift reference. The chemical shift values of the carbon and the proton of the methyl group were 49.5 and 3.36 ppm, respectively. A curve fitting program which was based on the Simplex algorithm (Caceci *et al.*, 1984) was used to calculate the pK_a values.

The NMR spectra were recorded at 25 °C on a Bruker AMX-500 spectrometer equipped with an X32 computer. All the spectra were processed using the Bruker UXNMR software package. The (¹H,¹³C)-HMQC spectra were recorded as described by Bax *et al.* (1983) in the phase-sensitive mode. Typically, each HMQC spectrum was recorded with 128 experiments in F1 and 1K complex data points in F2. The sweep width in the ¹H dimension covered 5 ppm with the carrier centered at ≈2.5 ppm and the ¹³C dimension sweep width covered 20 ppm with the carrier centered at ≈40 ppm. The data was zero filled once in the F2 dimension and twice in F1. A sine squared window

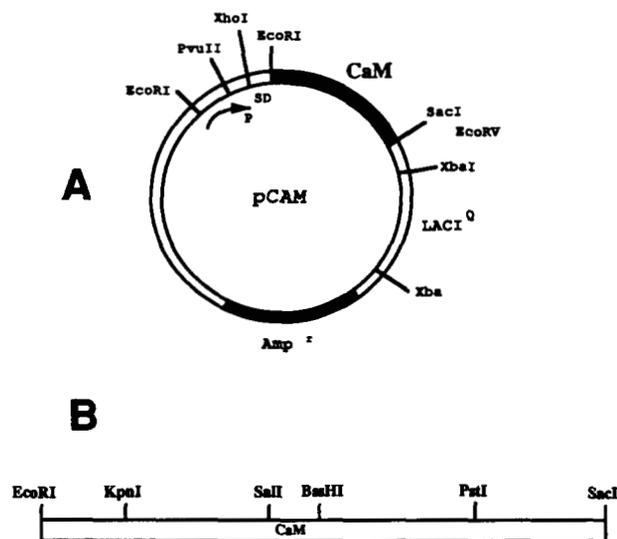


FIG. 1. Schematic diagrams of A, the map of the plasmid encoding the CaM gene; B, the restriction sites of the synthetic CaM gene.

function with 90° phase shifts was applied in both dimensions prior to the Fourier transformation.

The amide exchange rates of the Lys residues before and after methylation were measured in ^{15}N -Lys labeled CaM by a combination of saturation transfer and (^1H , ^{15}N)-HMQC as described by Spera *et al.* (1991). Two ^{15}N -Lys-labeled CaM samples (native protein and N^ε-Lys dimethylated protein) were prepared in 90% H_2O , 10% D_2O . The pH of both samples was adjusted to pH 6.0. To estimate the relative exchange rates, two experiments were performed with each sample, one with saturation of the water resonance during the preparation time (duration of 3 s and a width of ≈ 30 Hz) and the other with a 3-s relaxation delay. The ratio of the peak intensities determined with and without presaturation was used as a relative measure of the amide exchange rate. No attempts were made to measure the absolute value of the exchange rate.

An isotope filtered two-dimensional chemical exchange experiment was performed essentially as described by Montelione and Wagner (1989). The ^{13}C -methylated Ca^{2+} -CaM sample for this experiment was dissolved in 99.9% D_2O , 150 mM KCl, and the pH of the sample was adjusted to pH 5.98. 32 Scans were used for each experiment in F_1 with a total of 160 experiments. A 200-ms mixing time was used to observe cross-relaxation (as a result of chemical exchange or nOe) between the two slowly exchanging conformers.

RESULTS

Chemical Modification and Assignment of ^{13}C -Methylated Lys Residues—In this work, reductive methylation was employed to label the Lys residues in CaM by introducing isotopically ^{13}C -labeled methyl groups on their ϵ -amino groups. This labeling strategy has been used before, in combination with one-dimensional ^{13}C NMR, to determine the pK_a values of Lys residues in proteins (see for example, Jentoft and Dearborn (1979, 1983), Gerken (1984), and Huque and Vogel (1993)). Under the labeling conditions described under "Materials and Methods," over 90% of the ϵ - NH_2 group in each Lys residue becomes dimethylated, and only a small portion of ϵ - NH_2 is monomethylated. The ratio of dimethyllysine to monomethyllysine depends on the amounts of H^{13}CHO and NaCNBH_3 and the duration of the reaction; it also differs between the individual lysines (Jentoft and Dearborn, 1979, 1983; Huque and Vogel, 1993). Fig. 2 shows the two-dimensional (^1H , ^{13}C)-HMQC spectrum of ^{13}C -labeled Ca^{2+} -CaM at pH 10.1. Clearly, the monomethyl and dimethyl regions overlap in the ^1H dimension, but these two groups of resonance are well resolved in the ^{13}C dimension. Seven Lys peaks can be resolved in both the dimethyllysine and monomethyllysine regions. It is readily apparent that this approach provides a major advantage over ^{13}C NMR studies since the resolution in the ^1H dimension is better than in the ^{13}C dimension. The unique cross-peak positions indicate that all Lys residues are in distinct microenvironments. For apo-CaM, appreciable resonance overlap is observed in both regions (Fig. 3) suggesting that many Lys residues have similar microenvironments. The assignment of the ^1H - ^{13}C methyl groups to specific residues using the known side chain assignments of CaM (Ikura *et al.*, 1991) was unsuccessful because of the interruption of the magnetization transfer pathway from the ^{13}C of the methyl group to the ϵ - CH_2 by the ϵ -nitrogen atom in each Lys residue. Unfortunately various two- and three-dimensional NMR experiments proved to be not useful in obtaining the assignment (Zhang and Vogel, 1993). Therefore, in this study, the complete assignment of the monomethyllysine and dimethyllysine resonances was obtained by site-directed mutagenesis. All Lys residues in the protein were individually mutated to glutamine except for Lys-75, which was mutated to arginine, and the mutated proteins were ^{13}C -methylated and studied by (^1H , ^{13}C)-HMQC NMR. Fig. 4 shows how the resonances for Lys-13 and Lys-148 could be assigned. By comparison of the (^1H , ^{13}C)-HMQC spectrum of CaM (Fig. 2) with those for the K13Q- and

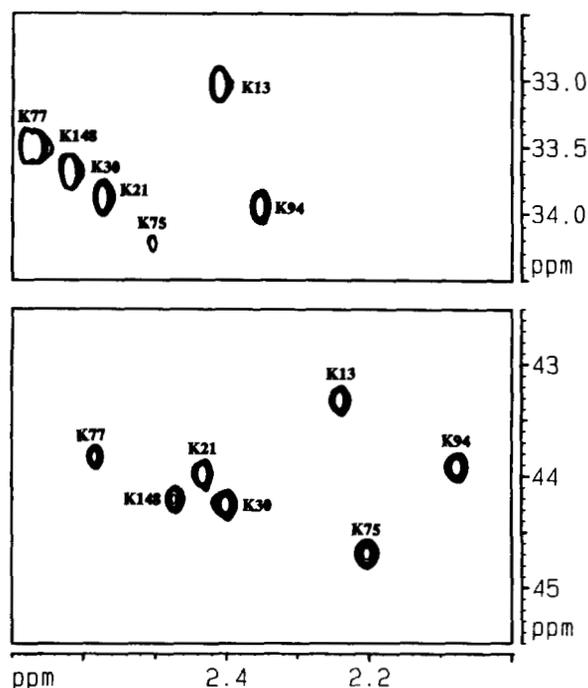


FIG. 2. (^1H , ^{13}C)-HMQC spectra of the N^ε-Lys ^{13}C -methylated Ca^{2+} -CaM at pH 10.0. The top panel and the bottom panel represent the monomethyllysine region and dimethyllysine region, respectively. The details of the experiment are described under "Materials and Methods."

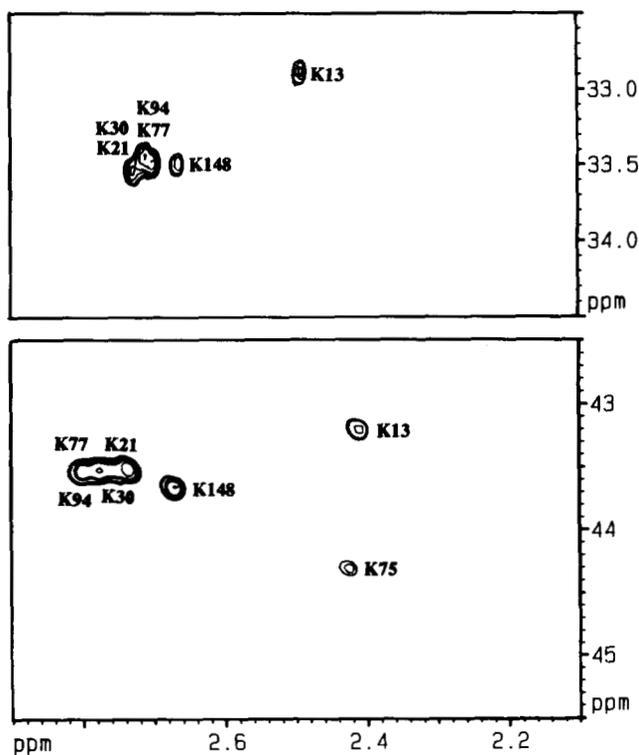


FIG. 3. (^1H , ^{13}C)-HMQC spectra of the N^ε-Lys ^{13}C -methylated apo-CaM at pH 10.0.

K148Q-CaM mutants (Fig. 4), it is readily apparent that the missing peaks in the Fig. 4, left and right, belong to Lys-13 and Lys-148, respectively. All the other Lys residues could be assigned in the same fashion because there were no noticeable chemical shift changes for the nonmutated Lys residues in the spectra of any of the mutants compared to that of bacte-

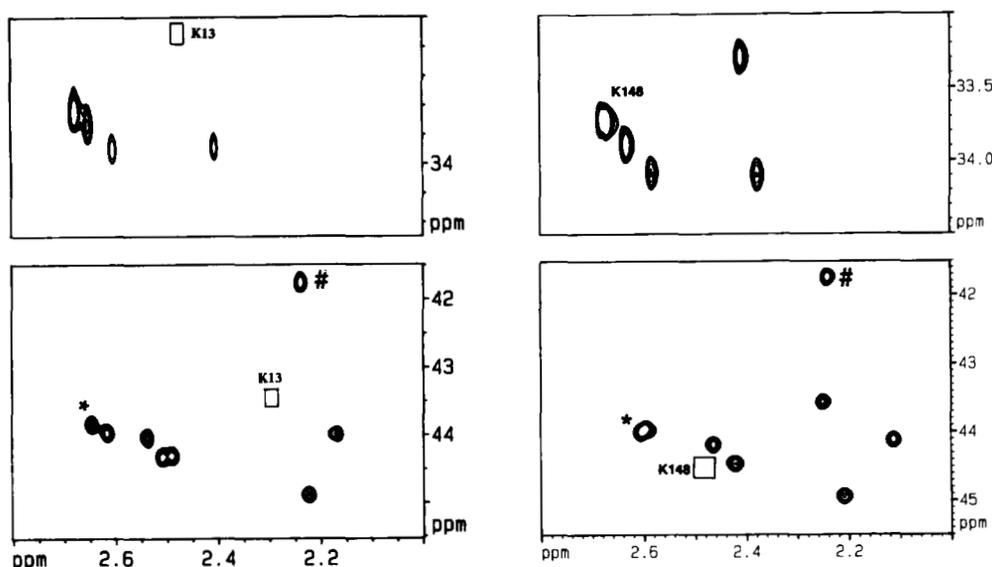


FIG. 4. (^1H , ^{13}C)-HMQC spectra of the N⁻Lys ^{13}C -methylated (left) K13Q-CaM mutant and (right) K148Q-CaM mutant recorded at pH 10.0. The experimental conditions are the same as described in the legend to Fig. 2. The resonance highlighted with an asterisk (*) results from Lys-115, and the resonance indicated with an # originates from the free amino group of the N-terminal residue. These two extra peaks arise because *E. coli* does not have the enzymes necessary to carry out these post-translational modifications. Since all the chemical reactivity studies with CaM were performed with the bovine protein, in which these two groups are unreactive, they are not considered further here.

rially expressed CaM. The lack of spectral changes upon mutation also suggests that the Lys residues do not directly influence each other; this notion is in agreement with their distribution over the protein surface (Babu *et al.*, 1988). The assignments are indicated in the (^1H , ^{13}C)-HMQC spectra of the methylated Ca^{2+} - and apo-form of CaM (Figs. 2 and 3, respectively). The assignments obtained by site-directed mutagenesis are consistent with the ones we reported earlier based on one-dimensional ^{13}C NMR studies of fully and partially modified Ca^{2+} -CaM and its proteolytic fragments (Huque and Vogel, 1993).

pH Titration and pK_a Values of the Methylated Lys Residues in CaM—pH titrations of methylated CaM were performed over the pH range from 2.0–12.0 for both the Ca^{2+} - and apo-forms. Fig. 5 shows part of the pH titration curves that were obtained for the dimethylated Lys residue in Ca^{2+} -CaM. From data such as those presented in Fig. 5, the pK_a values were calculated from the changes in the ^1H and ^{13}C chemical shifts. The pK_a values for both monomethyl Lys and dimethyllysine residues are listed in Tables I and II for Ca^{2+} -CaM and apo-CaM, respectively. The order of the pK_a values determined for monomethyl and dimethyl lysine residues are in agreement with each other. As expected from earlier studies with model compounds (Jentoft and Dearborn, 1983; Huque and Vogel, 1993), the pK_a values for monomethyllysine are about 0.8 pH units higher than those of the corresponding dimethyllysine residues. In addition, it is known that the pK_a values of dimethyllysines are very similar to those of unmodified lysines (Jentoft and Dearborn, 1983).

For most Lys residues, the ^1H and ^{13}C chemical shifts in the HMQC spectra of Ca^{2+} -CaM do not change once the pH of the sample is below ≈ 8 , with the exception of the resonances for Lys-13 and Lys-94. For Lys-13 two peaks are observed in the HMQC spectra when the pH of the sample is lower than 6.5. These collapse again into one peak once the pH of the sample was further decreased to pH ≈ 4.0 (Fig. 6). The resonance for Lys-94 undergoes a similar process except that it occurs in a different pH range (3.0–4.0). The presence

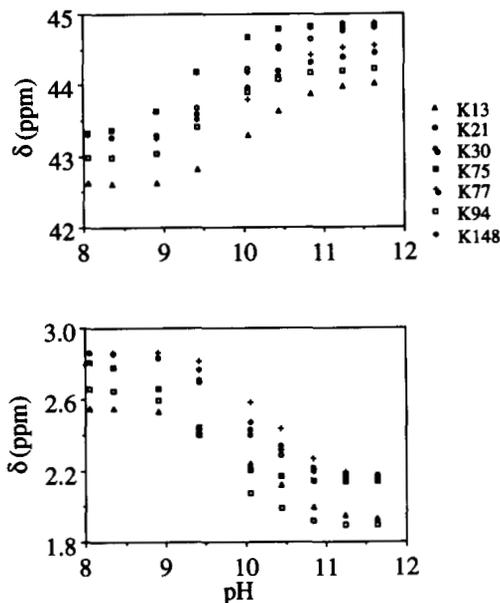


FIG. 5. pH titration curves of the dimethyllysine residues in Ca^{2+} -CaM as derived from (^1H , ^{13}C)-HMQC NMR spectra. The curves in the top panel and the bottom panel were derived from the ^{13}C and ^1H chemical shift changes, respectively.

of two peaks for Lys-13 between $4.0 < \text{pH} < 6.5$ and for Lys-94 between $3.0 < \text{pH} < 4.0$ can be the result of slow motions of part of their side chains in these pH ranges. Under such conditions the two methyl groups that are attached to the same ϵ -nitrogen can be in different microenvironments which could give rise to two different chemical shifts. A two-dimensional chemical exchange experiment was used to verify that the two peaks did indeed result from two methyl groups that are attached to the same Lys residue and that are in slow exchange; indeed the expected cross-peak between the two resonances was observed (Fig. 6). The nature of this exchange process is unclear at present. It could be related to a reduction

TABLE I
 pK_a values of the dimethyllysine and monomethyllysine residues in Ca^{2+} -calmodulin

Residue	pK_a values for dimethyllysine			pK_a values for monomethyllysine		
	From ^{13}C	From 1H	Mean	From ^{13}C	From 1H	Mean
13	10.09	10.09	10.09	10.74	10.76	10.75
21	9.88	9.88	9.88	10.58	10.54	10.56
30	9.83	9.84	9.84	10.65	10.64	10.65
75	9.28	9.29	9.29	10.23	10.28	10.26
77	10.20	10.25	10.23	10.92	11.03	10.98
94	9.65	9.64	9.65	10.32	10.31	10.32
148	9.88	10.02	10.00	10.92	11.03	10.98

TABLE II
 pK_a values of the dimethyllysine and monomethyllysine residues in apo-calmodulin

Residue	pK_a values for dimethyllysine			pK_a values for monomethyllysine		
	From ^{13}C	From 1H	Mean	From ^{13}C	From 1H	Mean
13	10.23	10.23	10.23	11.01	10.98	11.00
21	10.52	10.58	10.55	11.25	11.22	11.24
30	10.52	10.58	10.55	11.25	11.22	11.24
75	9.86	9.88	9.87	ND ^a	ND	ND
77	10.52	10.58	10.55	11.25	11.22	11.24
94	10.52	10.58	10.55	11.25	11.22	11.24
148	10.37	10.37	10.37	11.19	11.31	11.25

^a The intensity of the signal is so low that it was not possible to obtain pK_a values for this monomethyllysine.

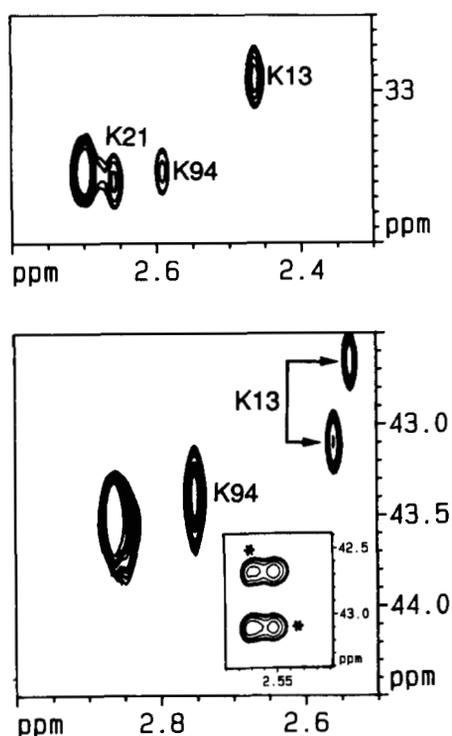


FIG. 6. $(^1H, ^{13}C)$ -HMQC spectra of N'-Lys methylated Ca^{2+} -CaM at the pH 5.98. The spectrum inserted in the bottom panel shows that the two resonances of Lys-13 display cross-peaks (highlighted with an asterisk) in the two-dimensional chemical exchange experiment recorded at this pH.

in the exchange rate of the protonated amine, alternatively it could be a reflection of the dissociation of calcium, or it could indicate the presence of a salt linkage as we suggested earlier (Huque and Vogel, 1993). However, salt linkages involving Lys residues were not observed in the crystal structure.

Effect of the Binding of the MLCK Peptide on the Lys pK_a Values—During the pH titration of the methylated CaM-MLCK peptide complex, it was found that the MLCK peptide binds tightly to CaM at pH values ranging from 7.5 to 11.3. Fig. 7 shows the monomethyllysine and dimethyllysine region of the $(^1H, ^{13}C)$ -HMQC spectrum of the CaM complex. The 7 Lys residues in the dimethyllysine region were well resolved. The assignment of the Lys residues in the complex was obtained by recording HMQC spectra of CaM-MLCK peptide complexes with the Lys mutant proteins. Fig. 8 shows an example for the assignment of Lys-13 and Lys-94 in the complex, the complete assignment is indicated in Fig. 7. By performing pH titrations we have again been able to obtain the pK_a values for each Lys residue of CaM in the complex, these are listed in Table III. Comparison with the pK_a values of the Lys residues in Ca^{2+} -CaM (see Table IV) shows that 4 Lys residues (Lys-13, Lys-30, Lys-77, and Lys-94) have virtually the same pK_a values when the MLCK peptide is bound to Ca^{2+} -CaM. The pK_a values of Lys-21 and Lys-148 increase by 0.25 and 0.5 units, respectively. Most interestingly, the pK_a of Lys-75 increases by 0.8 pH units, changing from the lowest pK_a value in Ca^{2+} -CaM to a very normal pK_a value in the CaM-peptide complex (Table IV).

Effect of the Chemical Modification on CaM—In order to ascertain that the dimethylation of CaM has no effect on the motions of the backbone of the protein, we studied the backbone amide exchange rates of the Lys residues since this parameter would likely be the most sensitive to any structural perturbations in the protein. To accomplish this, CaM was first selectively labeled with ^{15}N -Lys, and the amide exchange rates before and after chemical modification were studied for each Lys residue in the protein by the saturation transfer technique in combination with $(^1H, ^{15}N)$ -HMQC. Since a determination of the absolute values of the amide exchange rates is time consuming, we only compared the ratio of the peak volumes that were obtained with and without presaturation of the H_2O signal (I_{sat}/I_{unsat}). This ratio is known to be related to the amide exchange rate for fast exchanging amides (Pitner *et al.*, 1974). Since the amide exchange rates of the Lys residues in CaM fall in this category (Spera *et al.*, 1991),

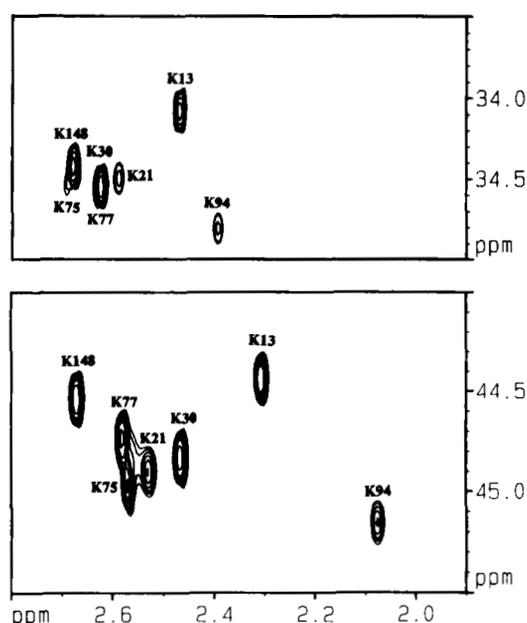


FIG. 7. $(^1H, ^{13}C)$ -HMQC spectrum of the complex of methylated bovine Ca^{2+} -CaM with the MLCK peptide at pH 10.0.

FIG. 8. (¹H,¹³C)-HMQC spectra of the N^ε-Lys methylated (left) K13Q-CaM/MLCK peptide complex and (right) K94Q-CaM/MLCK peptide complex at pH 10.0.

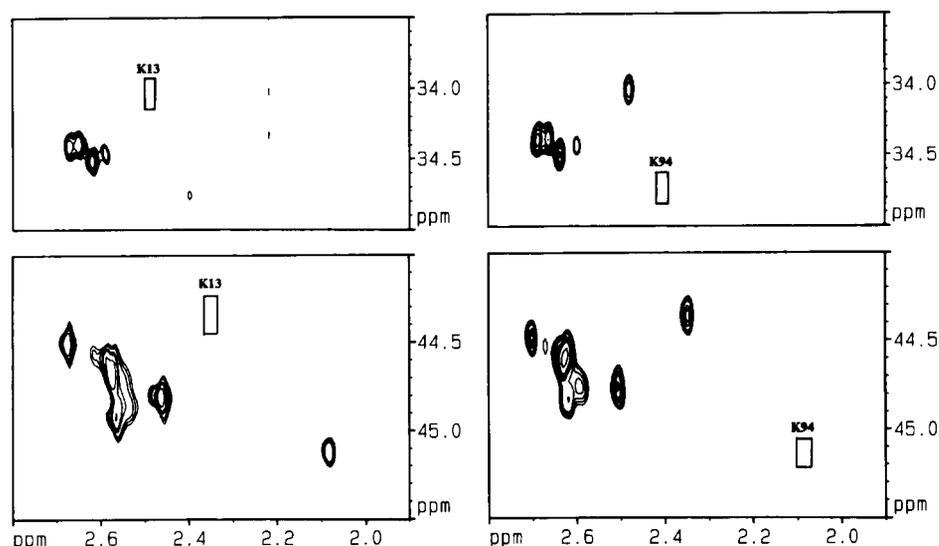


TABLE III
pK_a values of dimethyllysine and monomethyllysine in Ca²⁺-calmodulin·MLCK peptide complex

Residue	pK _a values for dimethyllysine			pK _a values for monomethyllysine		
	From ¹³ C	From ¹ H	Mean	From ¹³ C	From ¹ H	Mean
13	10.10	10.06	10.08	10.85	10.87	10.86
21	10.14	10.09	10.12	10.76	10.89	10.83
30	10.10	9.95	10.03	≈11.0 ^a	≈11.0 ^a	≈11.0 ^a
75	10.04	10.20	10.12	10.83	10.90	10.87
77	10.21	10.21	10.21	10.83	10.90	10.87
94	9.65	9.59	9.62	10.51	10.48	10.50
148	10.65	10.45	10.55	≈11.1 ^a	≈11.1 ^a	≈11.1 ^a

^a Since the complex was not stable at pH > 11.2, it was not possible to obtain accurate pK_a values for the monomethyllysines which have higher pK_a values than the dimethyllysine.

TABLE IV
pK_a value of the dimethyllysine residues in different forms of CaM

Residue	Apo-CaM	Ca ²⁺ -CaM	CaM/MLCK
13	10.23	10.09	10.08
21	10.55	9.88	10.12
30	10.55	9.84	10.03
75	9.87	9.29	10.12
77	10.55	10.23	10.21
94	10.55	9.65	9.62
148	10.37	10.00	10.55

this approach was followed here. Fig. 9 shows the (¹H,¹⁵N)-HMQC spectrum of ¹⁵N^ε-Lys selectively labeled CaM at pH 6.0. The assignment of the backbone amide resonances of the Lys residues was obtained from an earlier report (Ikura *et al.*, 1990). The (¹H,¹⁵N)-HMQC spectra are essentially the same before and after the chemical modification and this suggests that the modification does not perturb the protein structure (data not shown). Table V lists the values of $I_{\text{sat}}/I_{\text{unsat}}$ for each of the labeled amide protons in the native and modified protein. For most of the Lys residues the $I_{\text{sat}}/I_{\text{unsat}}$ ratio is identical (difference <5%) which means that they have the same amide exchange rates. A small difference (≈10%) was observed for Lys-75 and -77. This is probably related to the fact that the amide exchange rate of these 2 residues is extremely sensitive to the experimental conditions since they are in a very flexible region of the protein (Spera *et al.*, 1991; Barbato *et al.*, 1992), where a small difference in pH, ionic strength, or temperature would be sufficient to give rise to a

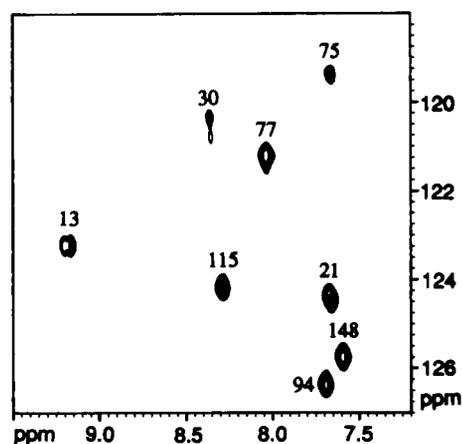


FIG. 9. (¹H,¹⁵N)-HMQC spectrum of ¹⁵N^ε-Lys-labeled Ca²⁺-CaM at pH 6.0. The assignment of the residues is indicated.

TABLE V
The relative amide exchange rate expressed as the $I_{\text{sat}}/I_{\text{unsat}}$ ratio of the lysine residues in calmodulin before and after chemical modification

Residue	Native	Modified
13	0.28	0.28
21	0.41	0.42
30	0.69	0.66
75	0.40	0.44
77	0.33	0.36
94	0.36	0.37
115	0.40	0.41
148	0.29	0.29

difference of 10% in the $I_{\text{sat}}/I_{\text{unsat}}$ ratio. Thus, these data confirm that the amide protons of the Lys residues in CaM have the same amide exchange rates before and after the chemical modification. We have also observed that the ¹H NMR spectra of the protein did not change upon methylation. Furthermore, the methylation of CaM does not change the ability of the protein to activate the CaM-dependent phosphodiesterase and smooth muscle myosin light chain kinase (data not shown), which indicates that the function of the protein is retained.

DISCUSSION

Many attempts have been made to determine the pK_a values of Lys residues in proteins by chemical modification studies

or via theoretical calculations. However, most of the chemical modification reagents perturb the Lys residues drastically (e.g. by attaching a large chemical group onto the Lys side chain) or they abolish the positive charge and/or pH titration behavior of the ϵ -NH₂ group of Lys residues (e.g. by acetylation or carbamylation) (for review, see Lundblad and Noyes (1984) and Fersht (1985)). Moreover, it can be time consuming to monitor the individual Lys residues in an intact protein after chemical modification. Theoretical calculations of the pK_a values of Lys residues in proteins have also been attempted. Unfortunately, to date the majority of these methods have only had modest success in deriving accurate pK_a values for the titratable groups of a protein (Warwicker and Watson, 1982; Bashford and Karplus, 1990; Yang *et al.*, 1993). Even for proteins for which a high resolution x-ray structure is available, calculated pK_a values for Lys residues can deviate significantly from experimental values. Errors in the theoretical calculation can arise from oversimplifications inherent in a model system, and more importantly by ignoring that the side chains of Lys residues in proteins are always in a dynamic motional state rather than in one fixed conformation. Theoretical calculations are simply not possible for proteins for which no three-dimensional structure is available.

In this paper, we have described an experimental strategy which combines established techniques, *viz.* chemical modification, two-dimensional HMQC NMR spectroscopy, and site-directed mutagenesis to determine the pK_a values of individual Lys residues in an intact protein using CaM as a model protein. This strategy is in principle applicable to any protein provided that the dimethylation does not alter the protein. The protein is first reductively methylated with ¹³C-labeled formaldehyde, and then the $-\text{CH}_3$ groups that are attached to the ϵ -NH₂ group of Lys residues can be monitored by (¹H, ¹³C)-HMQC spectra. This technique has a number of advantages for studying the Lys residues in a protein. 1) The chemical modification only adds one or two small methyl groups on to each Lys ϵ -nitrogen atom, and this introduces minimal perturbations in the structure of a protein (Jentoft and Dearborn, 1979, 1983; Gerken, 1984). Indeed, our amide exchange experiments show that the chemical modification has virtually no effect on the backbone dynamics of the Lys residues in CaM. It has also been shown by ¹H NMR that methylation does not affect the conformation of this protein (Huque and Vogel, 1993). Most importantly, the dimethylated Lys still retains its positive charge and pH titration behavior, thereby retaining the overall charge of the protein. Also, the chemical modification did not alter the function of CaM since the modified protein has the same ability to stimulate target enzymes as the native protein. 2) Inverse detected (¹H, ¹³C)-HMQC NMR spectroscopy provides excellent sensitivity for measuring the ¹³C-labeled methyl groups of Lys. A two-dimensional HMQC spectrum with very good signal to noise, that spans both the monomethyllysine and dimethyllysine region, can easily be acquired in 10 min of experimental time using a protein sample with a concentration of 1 mM. Recent technical advances such as gradient-NMR spectroscopy (Tyburn *et al.*, 1992) and HMQC without phase-cycling (Marion *et al.*, 1989) can even generate suitable spectra in less than 1 min.² This excellent sensitivity can facilitate NMR studies of other proteins that are not easy to obtain in large amounts or not very soluble in aqueous solution. 3) Another important advantage of this technique is that HMQC spectra provide excellent resolution for individual Lys residues as shown in Fig. 2. The 7 Lys residues in CaM are well resolved both in the monomethyl and in the dimethyl regions, and the movement of the resonances during the pH titration are easy to

follow. This is not always the case in one-dimensional ¹³C NMR (Jentoft and Dearborn, 1979, 1983; Gerken, 1984) or for one-dimensional ¹H NMR studies (Brown and Bradbury, 1973; Bradbury and Brown, 1975) of methylated proteins. Thus, the pK_a values of the Lys residues could be much more accurately determined by heteronuclear two-dimensional NMR than in one-dimensional ¹³C NMR experiments of the same protein (Huque and Vogel, 1993).

Correlation of the pK_a Values with Biochemical Data—The HMQC spectra of Lys ¹³C-labeled CaM in its apo-form and Ca²⁺-form are very different. In apo-CaM, the Lys resonances are less resolved, and 4 Lys residues (Lys-21, Lys-30, Lys-77, and Lys-94) have the same pH titration behavior and pK_a values, which means that they are in very similar chemical environments. In contrast, in Ca²⁺-CaM, each Lys residue has its own unique resonance and pK_a values, which indicates that each Lys residue is in a unique microenvironment. The results indicate that CaM undergoes a significant conformational change upon binding Ca²⁺. In general, the pK_a values of Lys residues in apo-CaM are higher than those in Ca²⁺-CaM. The higher pK_a values of the Lys residues in apo-CaM could mean that some of the Lys residues might be involved in salt linkages in order to stabilize the protein in the absence of Ca²⁺. However, a more likely explanation is that it arises from the increase of the total negative charge because of the dissociation of the four Ca²⁺ ions. Trace labeling experiments with acetic anhydride have shown that most of the Lys residues in Ca²⁺-CaM have higher reactivities than in apo-CaM (Giedroc *et al.*, 1985, 1987; Winkler *et al.*, 1987). The overall higher pK_a values of the Lys residues in apo-CaM suggest that this reactivity difference may be due to the differences in the pK_a values rather than to changes in the solvent accessibility of the Lys residues. In both apo- and Ca²⁺-CaM, Lys-75 has the lowest pK_a value, and this Lys has indeed the highest reactivity in both forms of the protein. The reactivity of Lys-75 increases significantly when calcium binds to apo-CaM, this is paralleled by a change in the pK_a values of this Lys from 9.87 to 9.29. The order of the reactivities of the Lys residues with acetic anhydride in Ca²⁺-CaM, as presented in the literature (Giedroc *et al.*, 1985, 1987; Winkler *et al.*, 1987), is in excellent agreement with the order of the pK_a values of the Lys residues presented in this study. Furthermore, no correlations are observed between the rate of the modification and the solvent accessibility of the side chains of the Lys residues as measured in the crystal structure of the protein (Babu *et al.*, 1988). Therefore, any suggestions that changes in structure or steric hindrance (Winkler *et al.*, 1987; Dwyer *et al.*, 1992) may lead to an increased reactivity of the Lys residues seems to be incorrect in the case of CaM. Whether the same applies to other proteins remains to be determined. However, the fact that the majority of the Lys side chains are exposed on the surface of proteins (Schulz and Schirmer, 1979) suggests that the correlation between pK_a and chemical reactivity will be widespread.

The binding of the MLCK peptide induces significant changes in the pK_a values of 2 Lys residues, notably Lys-148 and Lys-75. The pK_a values of Lys-75 increases from the lowest value, 9.29, in Ca²⁺-CaM to a very normal value, 10.12, in the CaM·MLCK peptide complex. As revealed in the x-ray structure, the side chain of Lys-75 is lying on top of the hydrophobic pocket of the N-terminal domain of the protein (Babu *et al.*, 1988). This position may be stabilized by partial polar interactions between the side chains of Lys-75 and Phe residues (Burley and Petsko, 1988). The low local dielectric constant around the side chain of Lys-75 may account for the low pK_a value of this Lys residue. The binding of the MLCK peptide to CaM induces a conformational change in the central helix of CaM (Ikura *et al.*, 1992; Meador

² M. Zhang and H. J. Vogel, unpublished results.

et al., 1992). In the complex, the MLCK peptide occupies the two hydrophobic pockets of CaM, thus forcing the side chain of Lys-75 away from the N-terminal hydrophobic pocket to a freely solvent accessible environment. Obviously, this change in the local environment of the side chain of Lys-75 is responsible for the change in its pK_a value from 9.29 to 10.12. The increase of the pK_a value of Lys-148 is also likely the result of similar changes in the local environment of this residue. Lys-148 is too flexible to be observed in any of the available crystal structures (Babu *et al.* 1988; Chattopadhyaya *et al.*, 1992; Rao *et al.*, 1993), however, it has been suggested that it lies on top of the hydrophobic surface in the C-terminal domain, and that it plays a similar role as Lys-75 (Strynadka and James, 1988). This suggestion is consistent with our pK_a data. Again, we noted that the pK_a differences accompanying the binding of the MLCK peptide to CaM are in excellent agreement with the changes reported for the reactivities of the residues towards acetic anhydride (Jackson *et al.*, 1986). Upon binding MLCK, Lys-75 experiences a dramatic decrease in reactivity which correlates with its large pK_a increase, while Lys-148 has a 2-fold decrease in its reactivity. The only other residue which also experiences a significant change in reactivity is Lys-21 (Jackson *et al.*, 1986). Again our data shows that it has a pK_a increase (0.25 pH unit) upon binding the MLCK peptide. At the same time, the pK_a values of other Lys residues (Lys-13, Lys-30, Lys-77, and Lys-94) remain nearly the same, and their reactivities towards acetic anhydride are also unchanged (Jackson *et al.*, 1986). Interestingly, the binding of calcineurin to CaM also results in a marked decrease in the reactivity of Lys-75, Lys-148, and Lys-21 (Manalan and Klee, 1987) suggesting that similar conformational changes occur in CaM upon binding this enzyme.

On the basis of the pK_a values determined here for the Lys residues of CaM, the pattern of the reactivities of the Lys residues towards a wide variety of chemical modification reagents can be grouped into three categories: 1) small electrophiles such as formaldehyde as used in this work, acetic anhydride (Giedroc *et al.*, 1985, 1987; Winkler *et al.* 1987; Jackson *et al.*, 1986) and nitrosourea (Mann and Vanaman, 1986), appear to be freely accessible to all the Lys residues in CaM, and the reactivity of the Lys residues is mainly dependent on their pK_a values; 2) for chemical modification reagents which contain hydrophobic groups that are not strong CaM binding antagonists (*i.e.* azidosalicylate *N*-hydroxysuccinimidylester (Dwyer *et al.*, 1992), *N*⁴-(9'-fluorenylmethyl)-oxycarbonyl-4-amino-1-oxyl-4-succinimidylloxycarbonyl-2,2,6,6-tetramethylpiperidine (Jackson and Puett, 1984)), the modification sites are determined by both the binding of the molecules onto the hydrophobic pockets of CaM and the pK_a values of the Lys residues; 3) for strong CaM binding antagonist derivatives such as reactive phenothiazines (Jarrett, 1984; Faust *et al.*, 1987; Newton and Klee, 1989), the modification sites are mainly determined by the binding of the molecules to CaM which results in the positioning of the reactive group in the vicinity of specific Lys residues in the protein.

In conclusion, we have described an approach which can be used to specifically determine the pK_a values of Lys residues in a protein. To our knowledge, this study makes CaM the first protein for which detailed information about the individual pK_a values of all the Lys residues in a number of physiological states is available. The high sensitivity and specificity of this technique makes the same approach a potentially general methodology for the determination of Lys pK_a values in other proteins provided that they are stable at the high pH values required for the titration experiments. The determination of the pK_a values of the Lys residues in three different forms of CaM, namely apo-, Ca²⁺-, and the MLCK-peptide

complex, provides the basis for understanding the many chemical modification studies of the Lys residues reported for this protein. By comparison with structural data we have found that the changes in the pK_a values of the Lys residues in different forms of CaM provide a sensitive measure of the local environment changes of the side chains of these residues. Our data indicate that the side chains of Lys-75 and Lys-148 are displaced from a location on their respective hydrophobic surfaces by the binding of the target proteins. In all likelihood, this displacement of Lys-75 plays a pivotal role in the unraveling of part of the central helix to form a flexible loop structure. This ultimately allows the two domains of CaM to bind in their proper orientation to the target proteins.

Acknowledgments—We thank Dr. T. Grundström at the University of Umeå for providing us with the plasmid encoding the synthetic CaM gene, and Dr. L. Gedamu at the University of Calgary for access to some of his equipment.

REFERENCES

- Andersson, A., Forsén, S., Thulin, E., and Vogel, H. J. (1983) *Biochemistry* **22**, 2039–2313
- Babu, Y. S., Bugg, C. E., and Cook, W. J. (1988) *J. Mol. Biol.* **204**, 191–204
- Barbato, G., Ikura, M., Kay, L. E., Pastor, R. W., and Bax, A. (1992) *Biochemistry* **31**, 5269–5278
- Bashford, D., and Karplus, M. (1990) *Biochemistry* **29**, 10219–10225
- Bax, A., Griffey, R. H., and Hawkins, B. L. (1983) *J. Magn. Reson.* **55**, 301–315
- Blumenthal, D. K., Takio, K., Edelman, A. M., Carbonneau, H., Titani, K., Walsh, K. A., and Krebs, E. G. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 3187–3191
- Bradbury, J. H., and Brown, L. R. (1973) *Eur. J. Biochem.* **40**, 565–576
- Brodin, P., Drakenberg, T., Thulin, E., Forsén, S., and Grundström, T. (1989) *Protein Eng.* **2**, 353–358
- Brown, L. R., and Bradbury, J. H. (1975) *Eur. J. Biochem.* **54**, 219–227
- Burley, S. K., and Petsko, G. A. (1988) *Adv. Protein Chem.* **39**, 125–189
- Caceci, M. S., and Cacheris, W. P. (1984) *Byte* **9**, 340–361
- Chattopadhyaya, R., Meador, W. E., Means, A. R., and Quijcho, F. A. (1992) *J. Mol. Biol.* **228**, 1177–1192
- Dwyer, L. D., Crocker, P. J., Watt, D. S., and Vanaman, T. C. (1992) *J. Biol. Chem.* **267**, 22606–22615
- Faust, F. M., Slisz, M., and Jarrett, H. W. (1987) *J. Biol. Chem.* **262**, 1938–1941
- Forsén, S., Vogel, H. J., and Drakenberg, T. (1986) *Calcium and Cell Function* (Cheung, W. Y., ed) Vol. 6, pp. 113–157, Academic Press, New York
- Fersht, A. (1985) *Enzyme Structure and Mechanism*, W. H. Freeman Press, San Francisco
- Gerken, T. A. (1984) *Biochemistry* **23**, 4688–4697
- Giedroc, D. P., Sinha, S. K., Brew, K., and Puett, D. (1985) *J. Biol. Chem.* **260**, 13406–13413
- Giedroc, D. P., Sinha, S. K., and Brew, K. (1987) *Arch. Biochem. Biophys.* **252**, 136–144
- Guerini, D., Krebs, J., and Carafoli, E. (1987) *Eur. J. Biochem.* **170**, 35–42
- Hiraoki, T., and Vogel, H. J. (1987) *J. Cardiovasc. Pharmacol.* **10**, 514–531
- Huque, M. E., and Vogel, H. J. (1993) *J. Protein Chem.*, in press
- Ikura, M., Kay, L. E., and Bax, A. (1990) *Biochemistry* **29**, 4459–4467
- Ikura, M., Spera, S., Barbato, G., Kay, L. E., Krinks, M., and Bax, A. (1991) *Biochemistry* **30**, 9216–9228
- Ikura, M., Clore, G. M., Gronenborn, A. M., Zhu, G., Klee, C. B., and Bax, A. (1992) *Science* **256**, 632–638
- Jackson, A. E., and Puett, D. (1984) *J. Biol. Chem.* **259**, 14985–14992
- Jackson, A. E., Carraway, K. L., III, Puett, D., and Brew, K. (1986) *J. Biol. Chem.* **261**, 12226–12232
- Jarrett, H. W. (1984) *J. Biol. Chem.* **259**, 10136–10144
- Jentoft, N., and Dearborn, D. G. (1979) *J. Biol. Chem.* **254**, 4359–4365
- Jentoft, N., and Dearborn, D. G. (1983) *Methods Enzymol.* **91**, 570–579
- Kadowaki, H., Kadowaki, T., Wondisford, F. E., and Taylor, S. I. (1989) *Gene (Amst.)* **76**, 161–166
- Klee, C. B., and Vanaman, T. C. (1982) *Adv. Protein Chem.* **35**, 213
- Lundblad, R. L., and Noyes, C. M. (1984) *Chemical Reagents for Protein Modification*, Vol. 1, CRC Press, Boca Raton, FL
- Manalan, A. S., and Klee, C. B. (1987) *Biochemistry* **26**, 1382–1390
- Mann, D. M., and Vanaman, T. C. (1985) *J. Biol. Chem.* **260**, 11284–11290
- Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) *J. Magn. Reson.* **85**, 393–399
- Meador, W. E., Means, A. R., and Quijcho, F. (1992) *Science* **257**, 1251–1254
- Means, A. R., VanBerkum, M. F. A., Bagchi, I., Lu, K. P., and Rasmussen, C. D. (1991) *Pharmacol. Ther.* **50**, 255–270
- Montelione, G. T., and Wagner, G. (1989) *J. Am. Chem. Soc.* **111**, 3096–3098
- Neidhardt, F. C., Bloch, P. L., and Smith, D. F. (1974) *J. Bacteriol.* **119**, 736–747
- Newton, D. L., and Klee, C. B. (1989) *Biochemistry* **28**, 3750–3757
- Pitner, T. P., Glickson, J. D., Dadok, J., and Marshall, G. R. (1974) *Nature* **250**, 582
- Putkey, J. A., Slaughter, G. R., and Means, A. R. (1985) *J. Biol. Chem.* **260**, 4704–4712
- Rao, S. T., Wu, S., Satyshur, K. A., Ling, K. Y., Kung, C., and Sundaralingam, M. (1993) *Protein Sci.* **2**, 436–447
- Sancho, J., Serrano, L., and Fersht, A. R. (1992) *Biochemistry* **31**, 2253–2258
- Schulz, G. E., and Schirmer (1979) *Principles of Protein Structure*, Springer-Verlag, New York

- Strynadka, N. C. J., and James, M. N. G. (1988) *Proteins Struct. Funct. Genet.* **3**, 1-17
- Spera, S., Ikura, M., and Bax, A. (1991) *J. Biomol. NMR* **1**, 155-165
- Tanford, C., and Roxby, R. (1972) *Biochemistry* **11**, 2192-2198
- Tyburn, J., Brereton, I. M., and Doddrell, D. M. (1992) *J. Magn. Reson.* **97**, 305-312
- Uhlen, B. E., Schweickart, V., and Clark, A. J. (1983) *Gene (Amst.)* **22**, 255-265
- Vogel, H. J., Lindahl, L., and Thulin, E. (1983) *FEBS Lett.* **157**, 241-246
- Wanner, B. L., Kodaira, R., and Neidhardt, F. C. (1977) *J. Bacteriol.* **130**, 211-222
- Warwicker, J., and Watson, H. C. (1982) *J. Mol. Biol.* **157**, 671-679
- Winkler, W. A., Fried, V. A., Merat, D. L., and Cheung, W. Y. (1987) *J. Biol. Chem.* **262**, 15466-15471
- Yang, A. S., Gunner, M. R., Sampogna, R., Sharp, K., and Honig, B. (1993) *Proteins Struct. Funct. Genet.* **15**, 252-265
- Zhang, M., and Vogel H. J. (1993) *Bull. Magn. Reson.* **15**, 95-97