Biochemistry

Accelerated Publications

Detecting Protein Kinase Recognition Modes of Calmodulin by Residual Dipolar Couplings in Solution NMR[†]

Tapas K. Mal,[‡,§] Nikolai R. Skrynnikov,[§,||] Kyoko L. Yap,[‡,§] Lewis E. Kay,[§,||] and Mitsuhiko Ikura*,[‡,§]

Division of Molecular and Structural Biology, Ontario Cancer Institute, and Department of Medical Biophysics, University of Toronto, Toronto, Ontario M5G 2M9, Canada, and Protein Engineering Network Centres of Excellence and Departments of Molecular and Medical Genetics, Biochemistry, and Chemistry, University of Toronto, 1 King's College Circle, Toronto, Ontario M5S 1A8, Canada

Received July 8, 2002; Revised Manuscript Received September 10, 2002

ABSTRACT: Calmodulin-regulated serine/threonine kinases (CaM kinases) play crucial roles in Ca^{2+}-dependent signaling transduction pathways in eukaryotes. Despite having a similar overall molecular architecture of catalytic and regulatory domains, CaM kinases employ different binding modes for Ca^{2+}/CaM recruitment which is required for their activation. Here we present a residual dipolar coupling (RDC)-based NMR approach to characterizing the molecular recognition of CaM with five different CaM kinases. Our analyses indicate that CaM kinase I and likely IV use the same CaM binding mode as myosin light chain kinase (1−14 motif), distinct from those of CaM kinase II (1−10 motif) and CaM kinase kinase (1−16⁻ motif). This NMR approach provides an efficient experimental guide for homology modeling and structural characterization of CaM−target complexes.

Calmodulin (CaM)[1] is a ubiquitous intracellular calcium (Ca^{2+}) sensor protein present in all eukaryotes, from humans to yeast. CaM regulates the activity of a large array of different target proteins and enzymes, including protein kinases, protein phosphatases, ion channels and pumps, nitric oxide synthases, adenylyl cyclases, and phosphodiesterases. Among these diverse CaM targets, CaM-regulated serine/threonine kinases (CaM kinases) are the best characterized structurally and functionally, and have similar functional domain organization. The CaM kinase family includes CaM kinase I (CaMKI), CaM kinase II (CaMKII), CaM kinase IV (CaMKIV), CaM kinase kinase (CaMKK), myosin light chain kinase (MLCK), and several other CaM kinases whose functions and structures are not well characterized (1).

All known CaM kinases exhibit a similar mechanism of Ca^{2+}-dependent activation (2, 3). In the resting state of the cell, where the intracellular Ca^{2+} concentration is $<10^{-7}$ M, CaM kinases are autoinhibited by their regulatory domain which encompasses the pseudosubstrate and CaM binding regions. Upon elevation of the intracellular Ca^{2+} concentra-
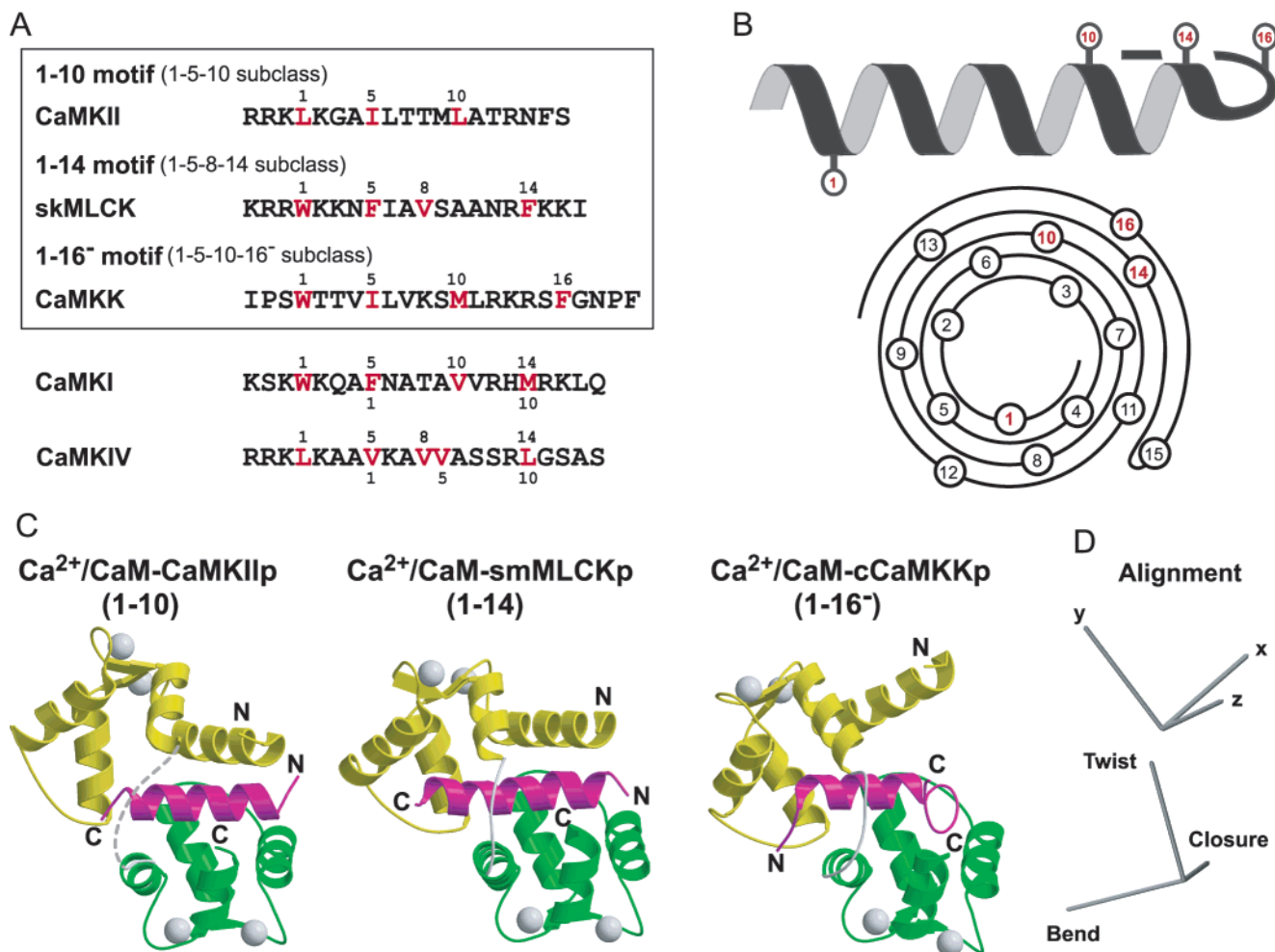
* To whom correspondence should be addressed: Division of Molecular and Structural Biology, Ontario Cancer Institute, and Department of Medical Biophysics, University of Toronto, Toronto, Ontario M5G 2M9, Canada. Fax: (416) 946-2055 or -6529. E-mail: mikura@uhnres.utoronto.ca.

[‡] Division of Molecular and Structural Biology, Ontario Cancer Institute, and Department of Medical Biophysics, University of Toronto.

[§] Protein Engineering Network Centres of Excellence, University of Toronto.

[||] Departments of Molecular and Medical Genetics, Biochemistry, and Chemistry, University of Toronto.

[1] Abbreviations: CaM, calmodulin; CaM kinase, calmodulin-regulated serine/threonine kinase; RDC, residual dipolar coupling; MLCK, myosin light chain kinase; CaMKI, CaM kinase I; CaMKII, CaM kinase II; CaMKIV, CaM kinase IV; CaMKK, CaM kinase kinase.

FIGURE 1: (A) Ca^{2+}/CaM binding sequences from CaMKII, skMLCK, rCaMKK, CaMKI, and CaMKIV. The possible anchoring residues binding to the hydrophobic pockets of Ca^{2+}/CaM are highlighted in red. (B) Schematic helical representation of the three distinct CaM binding motifs: $1-10$, $1-14$, and $1-16^{-}$. (C) Ribbon representation of three crystal structures of Ca^{2+}/CaM in complex with cCaMKKp (1.8 Å, PDB entry 1IQ5), smMLCKp (2.2 Å, PDB entry 1CDL), and CaMKIIp (2.0 Å, PDB entry 1CDM). These kinase-derived peptides are drawn in magenta; the N- and C-terminal domains of CaM are shown in yellow and green, respectively, and the linker region and Ca^{2+} are in gray. Panel C was generated using Molscript (*40*) and Raster3D (*41*). (D) Schematic representation of the principal axes of the alignment tensor for the structure of the CaM−cCaMKKp complex, and the axes of closure, bend, and twist used to relate different CaM conformations. Note that the orientation of the alignment axes in the CaM−peptide complex in solution with phage is determined mainly by electrostatic interactions, leading to a negative alignment (Aa = -1.56×10^{-3}) and a high rhombicity value ($R = 0.57$), in agreement with the results of Bax and co-workers (*30*).

tion to 10^{-6} M, CaM binds Ca^{2+} and becomes capable of binding with high affinity (typically, $K_d \sim 10^{-8}$–10^{-9} M) to a short peptide segment of a target kinase, thereby displacing the pseudosubstrate region. This results in conformational changes in both CaM and kinase catalytic domains, thus making the kinases functionally active. This mechanism has been supported by numerous biochemical studies on CaM kinases and several structural studies (*4*–*8*) on CaM in complex with kinase-derived peptides. Despite this common mechanism of activation, three different recognition modes of CaM kinases have been identified to date, which are termed the $1-10$, $1-14$, and $1-16^{-}$ motifs based on the position of two key anchoring hydrophobic residues in the target peptide (*2*, *9*) (Figure 1A,B). The $1-14$ motif was first identified in the NMR structure of CaM bound to the skeletal muscle MLCK peptide (skMLCKp) (*4*) and the crystal structure with smooth muscle MLCK peptide (smMLCKp) (*5*). The $1-10$ type was found in the structure of CaM complexed with a CaMKII peptide (CaMKIIp) (*6*). More recently, the NMR-derived structure of CaM bound

to the rat CaMKK peptide (rCaMKKp) (*8*) revealed the novel $1-16^{-}$ motif, which was also found in the high-resolution structure of CaM complexed with the *Caenorhabditis elegans* CaMKK peptide (cCaMKKp) (*7*). These CaMKK homologues from rat and *C. elegans* differ in their primary sequences at positions 1 (W versus L), 2 (T versus D), 4 (V versus L), 9 (S versus A), 11 (L versus G), 12 (R versus H), 13 (K versus R), 14 (R versus K), and 15 (S versus R). A striking feature of this $1-16^{-}$ motif is that the binding orientation of CaMKKp is the opposite of that of MLCKp and CaMKIIp with respect to the two CaM domains.

More than 350 CaM target proteins have been identified to date, the binding domains of which are classified in a database (http://calcium.uhnres.utoronto.ca/ctdb/) into several motif families based on sequence comparison (*10*). These sequence-based analyses, however, fail to predict CaM binding modes when no significant similarity is found between query and known sequences or when query sequences have several possible anchoring residues. Therefore, an efficient experimental procedure is needed to quickly

assess the CaM binding mode of a target protein. Here we demonstrate that residual dipolar couplings (RDCs) (*11*), which contain information about the orientation of atomic bond vectors in a protein, can be used to effectively characterize CaM−target interactions in solution. We find that partial backbone resonance assignments and corresponding amide ^{1}H−^{15}N RDC data are sufficient for probing differences in CaM recognition modes involving various CaM kinases.

MATERIALS AND METHODS

Sample Preparation. Uniformly ^{15}N- and ^{13}C-labeled recombinant *Xenopus laevis* CaM was overexpressed in *Escherichia coli* and purified to homogeneity as described previously (*12*). Peptides corresponding to the CaM-binding domain of rCaMKKp (α-isoform, residues 441−463), rat CaMKI (residues 300−320), rat CaMKIV (residues 319−339), and rabbit skMLCK (residues 578−597) were synthesized with a solid-phase peptide synthesizer (Peptide Synthesis Laboratory, Queen's University, Kingston, ON). A total of eight NMR samples were used for this study, each prepared in 500 μL of a 93% (v/v) H$_2$O/7% D$_2$O mixture (pH 6.8). The isotropic samples contained 20 mM Bis-Tris, 1 mM CaM, 5 mM CaCl$_2$, and 300 mM KCl. The aligned samples used for RDC measurements also contain 16 mg/mL filamentous phage Pf1 (Asla Labs). In each case, the molar ratio of CaM to peptide was 1:1.2 to ensure complete formation of the complex. The high salt concentrations used do not affect the stabilities or structures of CaM−target complexes, as the molecular interactions are very strong (K_d ∼ 10^{-8}−10^{-9} M) and hydrophobic in nature.

NMR Measurements. All NMR experiments were performed at 32 °C on a Varian Unity Inova 600 MHz spectrometer. The amide chemical shifts of all CaM−peptide complexes in the Pf1 liquid crystal are very similar to those measured in isotropic samples, indicating that perturbation of the protein structure by the alignment medium is minimal. Sequential assignments of the backbone resonances of CaM for the CaM−CaMKIp complex were determined using three-dimensional (3D) experiments [HNCACB, HBCBCA-(CO)NNH, and HNCO] (*13*). ^{1}H−^{15}N dipolar couplings were recorded using a two-dimensional (2D) IPAP-type sensitivity-enhanced (*14, 15*) ^{15}N−^{1}H HSQC correlation experiment. The IPAP strategy records $^1J_{HN}$-separated doublets that are either in phase (IP) or antiphase (AP), which are then combined to give spectra containing only upfield or downfield components (*16, 17*). All 2D IPAP experiments were carried out with 256 and 591 complex points in t_1 and t_2, respectively. Final data sets comprised 1024 and 4096 real points with digital resolutions of 1.8 and 2.2 Hz/point in F_1 and F_2, respectively. All NMR data were processed using NMRPipe and NMRDraw (*18*), and analyzed using XEASY (*19*).

Domain Orientation. Three crystal structures of CaM in complex with target peptides, including CaM−cCaMKKp (1.8 Å, PDB entry 1IQ5), CaM-smMLCKp (2.2 Å, PDB entry 1CDL), and CaM-CaMKIIp complexes (2.0 Å, PDB entry 1CDM), have been used in the analyses described above. All crystal structures are placed in the same coordinate frame by superimposing the C-terminal domain of each structure onto the C-terminal domain of 1CDM using MOLMOL (*20*). The rotations required to superimpose the N-terminal domains of the two structures with C-domains held fixed are described in terms of three consecutive rotations about orthogonal axes (closure, bend, and twist). The twist axis connects the centers of mass of the two domains in 1CDM. The closure axis is defined using the hinge axis about which a single rotation is performed to transform 1CDM into 1IQ5. This vector is then orthogonalized with respect to the twist axis to give the axis of closure. The cross product of unit vectors along the twist and closure axes defines the bend axis.

F Test Statistics. The *F* statistic was obtained from the relation $F = (\chi_i^2/\sigma_i^2)/(\chi_j^2/\sigma_j^2)$, where χ_k is the root-mean-square deviation between experimental and calculated DC values obtained from fitting RDC data using the complete structure of a CaM−peptide complex, σ_k is the average rmsd from fitting RDC data using only N- or C-terminal domains of CaM, and the index k denotes the specific crystal structure used for comparison. Probabilities (*P*) are evaluated using the function betainc (1/1+*F*, ν/2, ν/2) in Matlab (Matlab Inc.), where ν is the number of measured RDCs minus 5 (number of fitted alignment parameters). Note that specific values of *P* depend on σ_i and σ_j which are different for different crystal structures.

RESULTS AND DISCUSSION

Use of RDCs To Determine CaM Kinase Binding Modes. In recent years, RDCs have been successfully used to characterize molecular binding (*21−23*), assist in homology modeling (*24*), and determine relative domain orientations in multidomain proteins (*25, 26*). Here we first carried out NMR experiments on the structurally well-characterized CaM−skMLCKp and CaM−rCaMKKp complexes to assess the feasibility of using RDCs to determine the target recognition modes of CaM. We measured one-bond backbone ^{1}H−^{15}N RDCs of CaM in complex with peptides in a liquid crystalline medium containing 16 mg/mL filamentous phage Pf1 (*27*) (Figure 2). In the case of CaM−peptide complexes, the negatively charged phage Pf1 was an effective alignment medium, while other crystalline media such as bicelles (*11*) were unsuitable due to the presence of Ca^{2+} in the protein solution.

The experimentally measured ^{1}H−^{15}N RDCs of CaM in complex with rCaMKKp are analyzed in conjunction with the crystal structures of the CaM−cCaMKKp, CaM−smMLCKp, and CaM−CaMKIIp complexes (*5−7*) using a best-fit molecular alignment tensor obtained by the singular-value decomposition (SVD) method (*26, 28*). The comparison is performed for the structured regions comprising residues 6−72 from the N-terminal domain and residues 85−144 from the C-terminal domain, while residues 1−5, 73−84, and 145−148, which fall into unstructured and highly flexible regions, are excluded from the calculation. The best correlation is obtained when the experimental data of the CaM−rCaMKKp complex are compared with the RDC values calculated using the crystal structure of the CaM−cCaMKKp complex, as indicated by the correlation coefficient (R^2) of 0.98 and quality factor (*17*) (*Q*) of 13% (bottom panel in Figure 3A). The comparisons using the crystal structures of the CaM−smMLCKp and CaM−CaMKIIp complexes produce markedly poorer correlations,
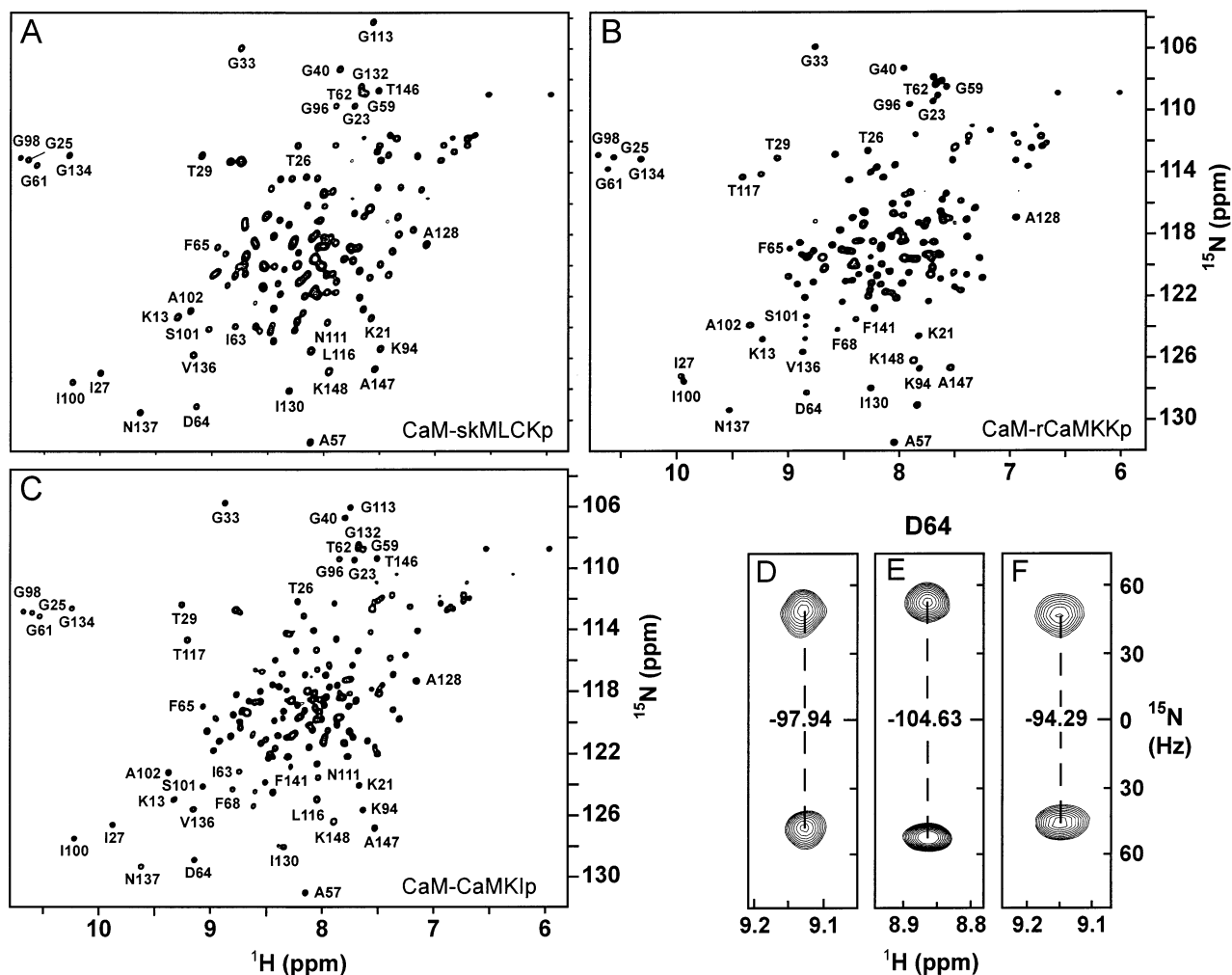
FIGURE 2: $^1H-^{15}N$ HSQC spectra of uniformly ^{15}N-labeled Ca^{2+}/CaM in complex with unlabeled (A) skMLCKp, (B) rCaMKKp, and (C) CaMKIp. A number of resonances easily assigned by visual inspection of a 3D HNCO data set are labeled in the plot. Resonances from residue D64 in $^{15}N-^1H$ IPAP-HSQC spectra of Ca^{2+}/CaM bound to (D) skMLCKp, (E) rCaMKKp, and (F) CaMKIp in Pf1 alignment media. The doublet is centered relative to the vertical axis, and the measured splitting is given in hertz. Dipolar couplings were determined by subtracting the splitting recorded from the samples with and without alignment medium.

with corresponding R^2 and Q values of 0.88 and 34% (middle panel) and 0.85 and 38% (top panel), respectively. An F test statistical analysis (*29*) shows a very minute probability (*P*) that the improvement noted for the cCaMKKp structure in relation to the other structures is the result of chance (7 × 10^{-4} and 2 × 10^{-8} for the comparisons with the CaM−smMLCKp and CaM−CaMKIIp complexes, respectively). Hence, the RDC data of CaM bound to rCaMKKp match far better to the CaM−cCaMKKp structure than to either the CaM−smMLCKp or the CaM−CaMKIIp structure. These results are also supported by separate experiments in which $^{15}N-^{13}C'$ RDC data have been measured where the following R^2 and Q values were obtained: 0.96 and 20% (CaM−cCaMKKp), 0.86 and 37% (CaM−smMLCKp), and 0.85 and 39% (CaM−CaMKIIp), respectively.

Similarly, the measured $^1H-^{15}N$ RDCs of the CaM−skMLCKp complex are compared with the best-fit values based on the crystal structures of three CaM−peptide complexes (Figure 3B). Not surprisingly, the experimental data fit remarkably well with the values calculated using the crystal structure of the CaM−smMLCKp complex (middle panel in Figure 3B), with R^2 and Q values of 0.97 and 18%, respectively, whereas significantly inferior correlations are

obtained when compared to either the CaM−CaMKIIp complex (0.90 and 31%, respectively, top panel) or the CaM−cCaMKKp complex (0.88 and 35%, respectively, bottom panel). The improvements obtained in fits with the CaM−smMLCKp complex relative to the CaM−CaMKIIp and CaM−CaMKKp complexes are significant with P values of 1 × 10^{-4} and 9 × 10^{-6}, respectively. Thus, the experimental RDC data of CaM in complex with skMLCKp are compatible only with the closely related CaM−smMLCKp structure that also displays the 1−14 type binding mode. Note that skMLCKp differs in the primary sequence from smMLCKp at positions 2 (Q versus K), 4 (T versus N), 5 (G versus F), 6 (H versus I) 9 (R versus S) 11 (I versus A) 12 (G versus N), and 14 (L versus F). The above two examples serve to illustrate that RDC data are highly sensitive to the structural parameters of CaM−target complexes and can be used to provide definitive information about the mode of binding of different peptides to CaM.

Different Conformations of CaM Bound to Kinases. The recognition modes of CaM with kinases (1−10, 1−14, and 1−16$^-$ motifs) differ from each other in the orientation of the two globular domains of CaM with respect to the helical binding region of the kinases (Figure 1C). The flexible
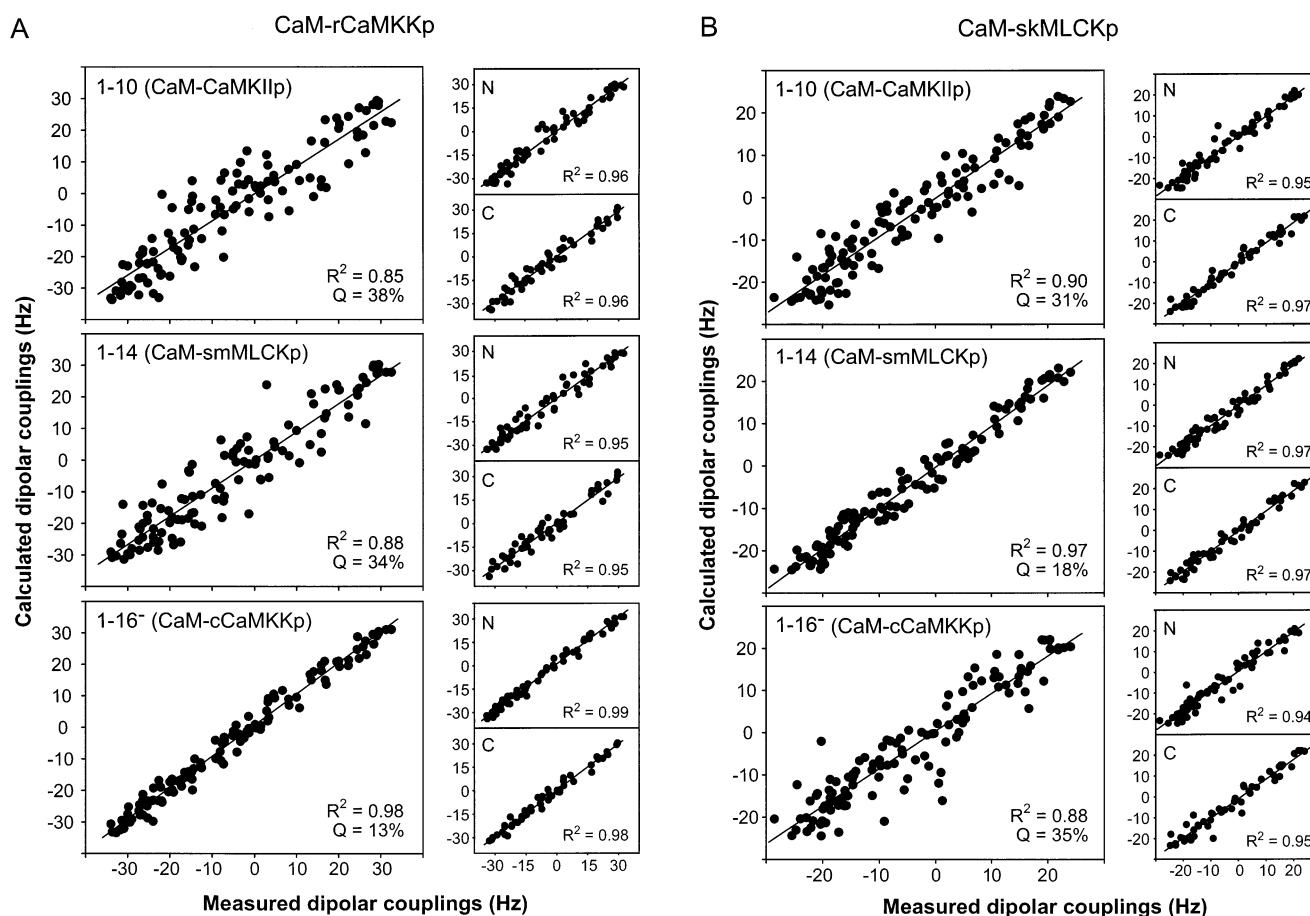
A

CaM-rCaMKKp

B

CaM-skMLCKp

FIGURE 3: Correlations between experimentally measured $^1H-^{15}N$ RDCs of (A) the CaM$-$rCaMKKp complex and (B) the CaM$-$skMLCKp complex and the best-fit RDC values calculated with the CaM$-$CaMKIIp (*6*) (1$-$10), CaM$-$smMLCKp (*5*) (1$-$14), and CaM$-$cCaMKKp (*7*) (1$-$16$^-$) crystal structures. The correlation coefficient R^2 and quality factor Q for each complex are given in the plots. The correlations derived for individual domains (N and C) of CaM are shown in the small panels to the right.

"central linker" between the two CaM domains is primarily responsible for rearrangement of the two domains in promoting an energetically favorable interaction with the target. This conformational variability of CaM can be described quantitatively in terms of rotations about what are termed closure, bend, and twist axes (Figure 1D) (for details, see Materials and Methods and ref *26*). The transformation of the CaM$-$smMLCKp structure to the CaM$-$CaMKIIp structure requires 17° closure followed by a -3° bend and an 8° twist. The transformation between the CaM$-$cCaMKKp and CaM$-$CaMKIIp structures requires more extensive rotations of 40°, 5°, and -13° for closure, bend, and twist, respectively. These analyses indicate that the orientations of the two CaM domains are very different for the 1$-$10 and 1$-$16$^-$ binding motifs, whereas the differences between the 1$-$10 and 1$-$14 motifs or between the 1$-$14 and 1$-$16$^-$ motifs are more subtle. Our RDC-based NMR approach can distinguish these subtle differences.

In this approach, the use of high-resolution crystal structures for the analysis of experimental RDC values is of critical importance. It is equally important to ensure that RDC analyses indeed reflect the differences in CaM domain orientation rather than variations in the local structure. This is verified by analyzing the experimental RDC data separately for N- and C-terminal domains of CaM. Very good correlations are obtained in this manner for all of the crystal structures considered (panels to the right in Figure 3),

indicating that structural changes within each of the two CaM domains upon binding to target peptides are minimal. Note that in the Ca^{2+}/CaM$-$peptide complexes studied here we do not observe the remarkable differences in the orientation of EF-hands in the N-terminal domain, which was described by Bax and co-workers (*30*) in a comparison of solution- and X-ray-derived structures of Ca^{2+}/CaM.

CaMKI Uses the Same Binding Mode as MLCK. To test this RDC-based approach on new target proteins, we studied the interaction between CaM and a 21-residue peptide encoding the CaM binding region of CaMKI (CaMKIp), for which no NMR or crystal structure is currently available. The CaMKIp sequence (*31*) suggests several possible anchoring residues for CaM (Figure 1), and experimental results are therefore necessary to clarify the precise mode of binding. The fitting of the $^1H-^{15}N$ RDC data from the CaM$-$CaMKIp complex using three different crystal structures is illustrated in Figure 4. The best correlation is observed for the crystal structure of the CaM$-$smMLCKp complex where $R^2 = 0.97$ and $Q = 17\%$ (middle panel) and much poorer correlations are obtained for the crystal structures of the CaM$-$cCaMKKp (0.89 and 33% with $P = 8 \times 10^{-4}$, bottom panel) and CaM$-$CaMKIIp (0.90 and 31% with $P = 8 \times 10^{-8}$, top panel) complexes. These results indicate that the recognition mode of CaM for CaMKI belongs to the 1$-$14 type, similar to that observed for MLCKs. Within the mapped CaM binding region of CaMKI
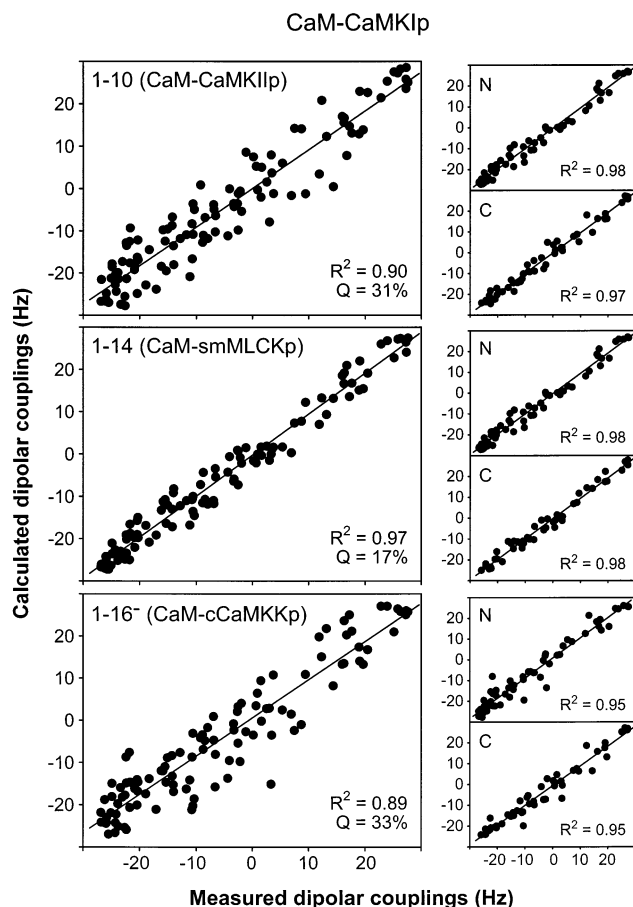
FIGURE 4: Correlation between experimentally measured $^1H-^{15}N$ dipolar couplings from the CaM−CaMKIp complex and the best-fit values obtained with the crystal structures of CaM−CaMKIIp (*6*) (1−10), CaM−smMLCKp (*5*) (1−14), and CaM−cCaMKKp (*7*) (1−16⁻) complexes. The correlations derived on the basis of individual domains (N and C) of CaM are shown in the small panels to the right.

(Figure 1), W303 and M316 are the only two residues that can anchor the 1−14 motif. In fact, Gomes et al. (*32*) have previously shown using tryptophan fluoresence and mutational studies that W303 is one of the anchoring residues for CaM binding. Taken together, we conclude that CaMKIp employs the 1−14 motif interaction with CaM using W303 and M316 as the key anchoring residues to the two CaM domains.

CaMKIV Likely Employs the 1−14 Type Binding Mode. Previous NMR studies of Ca²⁺-free and -bound CaM (*33−35*) and Ca²⁺/CaM−peptide complexes, including the CaM−skMLCKp (*4*), CaM−rCaMKKp (*8*), and CaM−C20W (*36*) (a peptide from the plasma membrane Ca²⁺ pump) complexes, have generated an extensive chemical shift (CS) database. Taking advantage of this wealth of information, we can easily assign approximately 35−45% of resonances in a $^1H-^{15}N$ HSQC spectrum of a CaM−peptide complex to specific amino acid residues by inspection of a 3D HNCO triple-resonance data set in which ^1H, ^{15}N, and $^{13}C'$ chemical shifts are recorded. In this way, many downfield-shifted resonances such as glycines G25, G61, G98, and G134 within the four Ca²⁺-binding loops and hydrophobic residues I27 and I100 within the β-sheet are easily assigned in $^1H-^{15}N$ HSQC spectra (Figure 2A−C). In the CaM−rCaMKKp complex, we were able to very rapidly assign 56 resonances
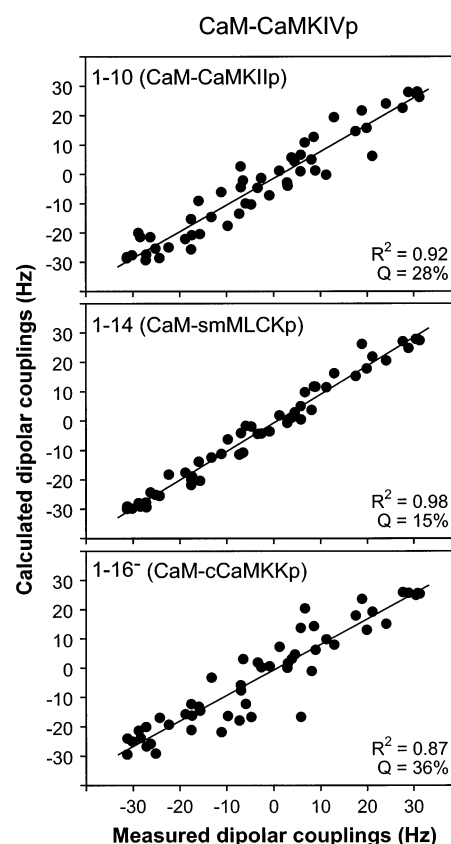
FIGURE 5: Correlation between experimentally measured $^1H-^{15}N$ dipolar couplings from the CaM−CaMKIVp complex and the best-fit values obtained with the crystal structures of CaM−CaMKIIp (*6*) (1−10), CaM−smMLCKp (*5*) (1−14), and CaM−cCaMKKp (*7*) (1−16⁻) complexes.

by simple visual inspection of $^1H-^{15}N$ HSQC and HNCO spectra and by using an in-house program that utilizes a CaM CS database. Of note, even with the reduced data set of RDC values, the R^2 and Q values remain high for the "correct" 3D structure (0.98 and 13%, respectively) compared to those for the "wrong" structures (0.90 and 32% for the CaM−smMLCKp complex and 0.85 and 39% for the CaM−CaMKIIp complex, respectively). This result suggests that even limited RDC data sets based on partial assignment can be sufficient to determine the CaM binding mode of a new target molecule.

We further explored the possibility of using reduced data sets for a complex of CaM with a CaMKIV peptide (CaMKIVp) for which structures and resonance assignments are not available. Fits of 52 RDCs from the CaM−CaMKIVp complex (assigned by visual inspection and with the use of an in-house program) using the CaM−CaMKIIp (1−10), CaM−smMLCKp (1−14), and CaM−cCaMKKp (1−16⁻) crystal structures produced the following R^2 and Q values: 0.92 and 28%, 0.98 and 15%, and 0.87 and 36%, respectively (Figure 5). Thus, the 1−14 mode of interaction appears to be the most likely for the CaMKIV peptide. The probability that the improvement in fit obtained with the CaM−smMLCKp model relative to the CaM−cCaMKKp structure results from chance is 0.016, effectively ruling out the 1−16⁻ mode of binding in this case. Statistical analyses cannot completely rule out the 1−10 binding mode, however. Moreover, an analysis of the sequence of CaMKIVp suggests that either the 1−14 or the 1−10 motif could be accom-

Table 1: Sequence Variation in the 1-5-8-14 Motif Identified in CaM Kinases[a]

	-2		1			5		8				14			
rabbit skMLCK (578 - 597)	K R R W K K N F I A V S A A N R F K K I
rat skMLCK 2 (580 - 599)	K R R W K K N F I A V S A R N R F K K I
human skMLCK (772 – 790)	Q R K W K K H F Y V V T A A N R L R K F
chicken smMLCK (797 – 816)	R R K W Q K T G H A V R A I G R L S S M
human MLCK 2 (1691 – 1710)	R R R W Q K T G N A V R A I G R L S S M
drosophila MLCK I (574 – 593)	R R K W Q K T G N A I R A L G R M A N L
rat CaMKI (300 – 319)	K S K W K Q A F N A T A V V R H M R K L
human CaMKIγ (301 – 320)	K S K W R Q A F N A A A V V H H M R K L
rat CaMKIβ (294 – 313)	R T H W K R A F N A T S F L R H I R K L

[a] The key hydrophobic and basic residues in CaM recognition are indicated in red and blue, respectively. Sequences: skMLCK (rabbit, A35021; rat, NP_476557; human, XP_040819), skMLCK (chicken, A35093; human, NP_444255; *Drosophila*, BAA19488), and CaMKI (rat, 1A06; human, NP_065172; rat isoform α, NP_058971).

modated (Figure 1). Thus, even in cases where complete assignments and structural data are unavailable, RDCs provide useful information about the mode of target binding.

CONCLUSIONS

The present DC-based NMR study has established that the CaM recognition mode of CaMKI is of the 1−14 motif type similar to that observed with MLCKs, and that it is likely that CaMKIV is also recognized in this manner. Inspection of the 1−14 motif sequences of all these CaM kinases reveals a variation of hydrophobic and basic residues at positions critical for CaM recognition (Table 1). In all known cases, the anchoring position 1 favors tryptophan over other hydrophobic amino acids for its binding to the CaM C-terminal binding domain, whereas position 14, which anchors to the CaM N-terminal domain, can be occupied by different hydrophobic residues varying from phenylalanine to leucine and even methionine. A number of different residues are found in positions 5 and 8, including phenylalanine, valine, threonine, isoleucine, alanine, and glycine, and these residues are therefore likely to be less critical for CaM binding. The N-terminal end (positions −2, −1, 0, 2, and 3) is comprised primarily of basic residues, consistent with its critical role in controlling the target helix orientation in the 1−14 type recognition motif (*8*).

The functional significance of the different binding modes of CaM kinases at a cellular level is still unclear at present. Previous biochemical and structural studies, however, indicate that the size of the binding interface and, accordingly, binding affinity depend on the type of the motif (total buried surface area varies from 2600 to 3000 Å^2 with K_d values ranging from 10^{-8} M for CaMKIIp to 10^{-9} M for skMLCKp and rCaMKKp) (*37*). It is possible that different binding modes may play a role in the kinetics or thermodynamics of phosphorylation and the subsequent activation of downstream target proteins. In this study, we have demonstrated that our RDC-based NMR approach can quickly determine the CaM binding mode of a new target protein. It is noteworthy that the entire process of NMR data collection and analysis takes only a couple of days. This approach can easily be combined with homology modeling methods to build a plausible structural model for a CaM−target complex in an efficient manner. As the number of determined structures for CaM−

target protein complexes increases, this approach will become more powerful in identifying other homologous CaM−target complexes. For instance, our method is expected to benefit from two recent structures of CaM in complex with Ca^{2+}-activated K$^+$ channels (*38*) and anthrax adenylyl cyclase (*39*) that reveal completely new CaM recognition modes. In conclusion, this RDC-based approach provides an efficient way of probing molecular recognition processes and will serve an important role in the postgenomic area, particularly for the characterization of protein−protein interactions in solution.

REFERENCES

1. Barnes, J. A., King, M. J., Kalra, J., and Sharma, R. K. (1992) *Biochem. Biophys. Res. Commun. 186*, 819−826.
2. Hoeflich, K. P., and Ikura, M. (2002) *Cell 108*, 739−742.
3. Chin, D., and Means, A. R. (2000) *Trends Cell Biol. 10*, 322−328.
4. Ikura, M., Clore, G. M., Gronenborn, A. M., Zhu, G., Klee, C. B., and Bax, A. (1992) *Science 256*, 632−638.
5. Meador, W. E., Means, A. R., and Quiocho, F. A. (1992) *Science 257*, 1251−1255.
6. Meador, W. E., Means, A. R., and Quiocho, F. A. (1993) *Science 262*, 1718−1721.
7. Kurokawa, H., Osawa, M., Kurihara, H., Katayama, N., Tokumitsu, H., Swindells, M. B., Kainosho, M., and Ikura, M. (2001) *J. Mol. Biol. 312*, 59−68.
8. Osawa, M., Tokumitsu, H., Swindells, M. B., Kurihara, H., Orita, M., Shibanuma, T., Furuya, T., and Ikura, M. (1999) *Nat. Struct. Biol. 6*, 819−824.
9. Rhoads, A. R., and Friedberg, F. (1997) *FASEB J. 11*, 331−340.
10. Yap, K. L., Kim, J., Truong, K., Sherman, M., Yuan, T., and Ikura M. (2000) *J. Struct. Funct. Genomics 1*, 8−14.
11. Tjandra, N., and Bax, A. (1997) *Science 278*, 1111−1114.
12. Ikura, M., Kay, L. E., and Bax, A. (1990) *Biochemistry 29*, 4659−4667.
13. Sattler, M., Schleucher, J., and Griesinger, C. (1999) *Prog. Nucl. Magn. Reson. Spectrosc. 34*, 93−158.
14. Schleucher, J., Sattler, M., and Griesinger, C. (1993) *Angew. Chem., Int. Ed. 32*, 1489−1491.
15. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663−10665.

16. Yang, D. W., and Nagayama, K. (1996) *J. Magn. Reson., Ser. A 118*, 117−121.
17. Ottiger, M., Delaglio, F., and Bax, A. (1998) *J. Magn. Reson. 131*, 373−378.
18. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
19. Bartels, C. H., Xia, T.-H., Billeter, M., Güntert, P., and Wüthrich, K. (1995) *J. Biomol. NMR 5*, 1−10.
20. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
21. Bolon, P. J., Al-Hashimi, H. M., and Prestegard, J. H. (1999) *J. Mol. Biol. 293*, 107−115.
22. Clore, G. M. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 9021−9025.
23. Koenig, B. W., Mitchell, D. C., Konig, S., Grzesiek, S., Litman, B. J., and Bax, A. (2000) *J. Biomol. NMR 16*, 121−125.
24. Chou, J. J., Li, S., and Bax, A. (2000) *J. Biomol. NMR 18*, 217−227.
25. Fischer, M. W., Losonczi, J. A., Weaver, J. L., and Prestegard, J. H. (1999) *Biochemistry 38*, 9013−9022.
26. Skrynnikov, N. R., Goto, N. K., Yang, D., Choy, W. Y., Tolman, J. R., Mueller, G. A., and Kay, L. E. (2000) *J. Mol. Biol. 295*, 1265−1273.
27. Hansen, M. R., Mueller, L., and Pardi, A. (1998) *Nat. Struct. Biol. 5*, 1065−1074.
28. Losonczi, J. A., Andrec, M., Fischer, M. W., and Prestegard, J. H. (1999) *J. Magn. Reson. 138*, 334−342.
29. Mosteller, F., and Tukey, J. W. (1977) *Data analysis and regression: A second course in statistics*, Addison-Wesley, Reading, MA.
30. Chou, J. J., Li, S., Klee, C. B., and Bax, A. (2001) *Nat. Struct. Biol. 8*, 990−997.
31. Goldberg, J., Nairn, A. C., and Kuriyan, J. (1996) *Cell 84*, 875−887.
32. Gomes, A. V., Barnes, J. A., and Vogel, H. J. (2000) *Arch. Biochem. Biophys. 379*, 28−36.
33. Finn, B. E., Evenas, J., Drakenberg, T., Waltho, J. P., Thulin, E., and Forsen, S. (1995) *Nat. Struct. Biol. 2*, 777−783.
34. Kuboniwa, H., Tjandra, N., Grzesiek, S., Ren, H., Klee, C. B., and Bax, A. (1995) *Nat. Struct. Biol. 2*, 768−776.
35. Zhang, M., Tanaka, T., and Ikura, M. (1995) *Nat. Struct. Biol. 2*, 758−767.
36. Elshorst, B., Hennig, M., Forsterling, H., Diener, A., Maurer, M., Schulte, P., Schwalbe, H., Griesinger, C., Krebs, J., Schmid, H., Vorherr, T., and Carafoli, E. (1999) *Biochemistry 38*, 12320−12332.
37. Crivici, A., and Ikura, M. (1995) *Annu. Rev. Biophys. Biomol. Struct. 24*, 85−116.
38. Schumacher, M. A., Rivard, A. F., Bachinger, H. P., and Adelman, J. P. (2001) *Nature 410*, 1120−1124.
39. Drum, C. L., Yan, S. Z., Bard, J., Shen, Y. Q., Lu, D., Soelaiman, S., Grabarek, Z., Bohm, A., and Tang, W. J. (2002) *Nature 415*, 396−402.
40. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
41. Merrit, E. A., and Bacon, D. J. (1997) *Methods Enzymol. 277*, 505−524.