

# Mapping the Topology and Determination of a Low-Resolution Three-Dimensional Structure of the Calmodulin–Melittin Complex by Chemical Cross-Linking and High-Resolution FTICRMS: Direct Demonstration of Multiple Binding Modes<sup>†</sup>

Daniela M. Schulz,<sup>‡</sup> Christian Ihling,<sup>‡</sup> G. Marius Clore,<sup>§</sup> and Andrea Sinz<sup>\*,‡</sup>

Biotechnological-Biomedical Center, Faculty of Chemistry and Mineralogy, University of Leipzig, D-04103 Leipzig, Germany, and Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520

Received December 1, 2003; Revised Manuscript Received February 13, 2004

**ABSTRACT:** Calmodulin serves as a calcium-dependent regulator in many metabolic pathways and is known to bind with high affinity to various target proteins and peptides. One such target is the small peptide melittin, the principal component of honeybee venom. The calmodulin–melittin system was used as a model system to gain further insight into target recognition of calmodulin. Using chemical cross-linking in combination with high-resolution Fourier transform ion cyclotron resonance mass spectrometry (FTICRMS), we have determined the interacting regions within the calcium-dependent calmodulin–melittin complex and thus the orientation of bound melittin. Using ambiguous distance restraints derived from the chemical cross-linking data in combination with recently developed computational methods of conjoined rigid body/torsion angle simulated annealing, we were able to generate low-resolution three-dimensional structure models of the calmodulin–melittin complex, for which no high-resolution structure exists to date. Our data provide evidence for the first time that calmodulin can recognize target peptides in two opposing orientations simultaneously. The general procedure for mapping interacting regions within the complex involves conjugation of calmodulin and melittin with several cross-linking reagents possessing different specificities and spacer lengths, followed by enzymatic proteolysis of the cross-linked complex. The highly complex peptide mixtures were subsequently analyzed by nano-HPLC, which was online coupled to a FTICR mass spectrometer equipped with a nano-electrospray ionization source. The mass spectra obtained in this manner were screened for possible cross-linking products using customized software programs. This integrated approach, exemplified for mapping the topology of the calmodulin–melittin complex, is likely to have wide-ranging implications for structural studies on protein–protein interactions.

Calmodulin (CaM)<sup>1</sup> is a small (148 amino acids) acidic protein belonging to the class of EF-hand proteins, which is found ubiquitously in animals, plants, fungi, and protozoa (1). CaM serves as a calcium-dependent regulator in many metabolic pathways (2). Upon calcium binding, CaM adopts a dumbbell structure (3, 4) consisting of two lobes connected

by a flexible central helix (5, 6). CaM is known to bind with high affinity to various target proteins and peptides, with a dissociation constant in the low nanomolar range (7). One such target is the small, 26-residue peptide, melittin, the principal component of honeybee (*Apis mellifera*) venom. Concomitant with complex formation, melittin adopts an  $\alpha$ -helical conformation (8) consisting of two  $\alpha$ -helical segments interrupted by proline, with a bent rod-shaped appearance (9, 10). Our aim was to gain further insight into the determinants of target recognition by CaM, as exemplified by the CaM–melittin system. Melittin has been widely used as a tool for studying protein–protein interactions and is known to interact with a number of proteins, including CaM (11), myosin light chains (12), and calsequestrin (13). Melittin activates protein kinase C, Ca<sup>2+</sup>/CaM-dependent protein kinase II (14), and phospholipase A<sub>2</sub> (15) and inhibits P-type ion-motive ATPases (16–18). It has been demonstrated that melittin is one of the most potent inhibitors of CaM activity and as such is also a potent inhibitor of cell growth (19, 20). It has also been shown that melittin inhibits the melanotropin receptor in M2R melanoma cell membranes (21), and a recent report has revealed that bee venom inhibits carcinoma cell proliferation in vitro and tumor growth in vivo (22).

<sup>†</sup> Funding information: Saxon State Ministry of Higher Education, Research and Culture.

\* To whom correspondence should be addressed: Biotechnological-Biomedical Center, Faculty of Chemistry and Mineralogy, University of Leipzig, Linnéstrasse 3, D-04103 Leipzig, Germany. Phone: +49-341-9736078. Fax: +49-341-9736115. E-mail: sinz@chemie.uni-leipzig.de.

<sup>‡</sup> University of Leipzig.

<sup>§</sup> National Institutes of Health.

<sup>1</sup> Abbreviations: BS<sup>3</sup>, bis(sulfosuccinimidyl)suberate; CaM, calmodulin; CaMKK, CaM-dependent kinase kinase; CAMK, CaM-dependent protein kinase; CO-1, channel outlet 1; CO-2, channel outlet 2; sulfo-DST, disulfosuccinimidyl tartarate; ESI-FTICRMS, electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry; EDC, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride; HPLC, high-performance liquid chromatography; LHRH, luteinizing hormone releasing hormone; MALDI-TOFMS, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; MARCKS, myristoylated alanine-rich C kinase substrate; sulfo-NHS, *N*-hydroxy-sulfosuccinimide; C20W, plasma membrane calcium pump, synthetic peptide; skMLCK, skeletal-muscle myosin light-chain kinase; sm-MLCK, smooth-muscle myosin light-chain kinase.

A common feature of most high-affinity CaM-binding targets is the presence of an amino acid sequence capable of adopting the structure of an amphiphilic  $\alpha$ -helix (23). With respect to CaM, target binding induces a conformational transition from the dumbbell structure to an overall globular-shaped complex with the two lobes of CaM coming close to one another (24, 25). Thus, the target peptide is engulfed in a hydrophobic channel (26, 27). Hydrophobic interactions involve bulky hydrophobic anchor residues of the CaM-binding target and hydrophobic patches located within the two lobes of CaM (26, 28, 29). A cluster of basic amino acid residues in the sequences of CaM-binding peptides can also interact with the highly abundant negatively charged residues of CaM (30, 31). To date, several structures of CaM-CaM-binding target complexes have been solved by NMR spectroscopy and X-ray crystallography; these include CaM complexed with skMLCK (skeletal-muscle myosin light-chain kinase) (26), smMLCK (smooth-muscle myosin light-chain kinase) (31), CaMKII $\alpha$  (CaM-dependent protein kinase II $\alpha$ ) (32), C20W (plasma membrane calcium pump synthetic peptide) (33), CaMKK (CaM-dependent kinase kinase) (30, 34), CaMKI (CaM-dependent protein kinase I) (35), MARCKS (myristoylated alanine-rich C kinase substrate) peptide (36), plant glutamate decarboxylase (37), calcium-activated K<sup>+</sup> channel (38), and anthrax adenylate cyclase exotoxin (39). Unexpected aspects of CaM-target interaction have been observed in the latter two structures, demonstrating that the mechanism of activation may be less general than previously thought.

Chemical cross-linking in combination with high-resolution Fourier transform ion cyclotron resonance (FTICR) mass spectrometry provides an analytical approach that can serve as a low-resolution counterpart to NMR spectroscopy and X-ray crystallography for the structural characterization of protein complexes. While this approach is certainly not able to compete with the high-resolution structural data offered by either NMR spectroscopy or X-ray crystallography, it has the potential to overcome some of the inherent limitations of NMR and X-ray crystallography. Both NMR spectroscopy and X-ray crystallography require large amounts of material and are intrinsically time-consuming techniques; NMR spectroscopy is restricted to relatively small ( $\leq 50$  kDa) proteins and protein complexes, and crystals of macromolecular complexes that diffract to high resolution are often difficult to obtain. In recent studies, mass spectrometric analysis of peptide mixtures derived from chemical cross-linking and subsequent proteolytic digestion has proven to represent an efficient tool for the elucidation of spatial organizations in proteins (40, 41) and protein complexes (42, 43). The mass range of the protein complex under scrutiny is theoretically unlimited, because it is the proteolytic peptides that are analyzed. Analysis is generally fast and requires only minimal amounts of analyte (44). Further, the broad range of specificities available for cross-linking reagents toward certain functional groups, such as primary amines or carboxylic acids, and the wide range of distances different cross-linking reagents can bridge offer various options for experimental design (45). Despite these advantages, however, identification and correct assignment of the cross-linking products can be hampered by the complexity of the peptide mixtures derived from intermolecular cross-linking of two proteins. Tagging cross-linking reagents or

proteins with isotope (46, 47) or fluorogenic (48) labels has been used in attempts to facilitate identification of the cross-linking products.

In the present paper of the CaM-melittin complex, we employed chemical cross-linking in combination with nano-HPLC/nano-ESI-FTICRMS and computational structure determination methodologies derived from NMR spectroscopy (49–51). FTICRMS offers distinct advantages for the analysis of biological samples, including excellent mass resolution, ultrahigh mass measurement accuracy, high sensitivity, and a wide mass range (52, 53). Overall, the most decisive advantage of FTICRMS is that it permits identification of cross-linking products based solely on exact mass measurements without the need for additional labeling of either cross-linking reagents or proteins (40, 54). By making use of cross-linking reagents with different spacer lengths, we were able to map the relative (i.e., upper bound) distances of cross-linked atoms. Thus, we were able to determine interacting regions between CaM and melittin and therefore the orientation of the peptide within the complex. Previous limited proteolysis and cross-linking experiments had suggested that melittin is inversely oriented in the CaM complex relative to the peptides derived from MLCK, CaMKI, or CaMKII (55). From the ambiguous distance restraints derived from chemical cross-linking, we were able to generate low-resolution structure models of the CaM-melittin complex, for which no high-resolution structure exists to date. This is the first demonstration that it is possible to truly determine a low-resolution three-dimensional structure of a protein-peptide complex from mass spectrometric data when high-resolution structures of the free proteins are available. The integrated approach described here proves for the first time that CaM is able to recognize peptides in two opposing orientations simultaneously, which reveals that target recognition by CaM is more complex and varied than was previously supposed.

## EXPERIMENTAL PROCEDURES

**Materials.** Bovine brain CaM and melittin were obtained from Calbiochem (Schwalbach am Taunus, Germany) and were used without further purification. Purity was checked by MALDI-TOFMS and SDS-PAGE. The cross-linking reagents EDC [1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride], sulfo-NHS (*N*-hydroxysulfosuccinimide), BS<sup>3</sup> [bis(sulfosuccinimidyl)suberate], and sulfo-DST (disulfosuccinimidyl tartarate) were purchased from Pierce (Rockford, IL). Trypsin, endoproteinase AspN, and LysC (all sequencing grade) were obtained from Roche Diagnostics (Mannheim, Germany). LHRH (luteinizing hormone releasing hormone), sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid), and proteins for MALDI-TOFMS calibration were purchased from Sigma (Taufkirchen, Germany). Nano-HPLC solvents were spectroscopic-grade (Uvasol, VWR, Darmstadt, Germany). Water was purified with a Direct-Q5 water-purification system (Millipore, Eschborn, Germany).

**Cross-Linking Reactions.** An equimolar mixture of CaM and melittin (10  $\mu$ M each, final concentration) containing 1 mM CaCl<sub>2</sub> (100 mM MES buffer at pH 6.5, final volume 1 mL) was incubated at room temperature for 20 min prior to addition of the zero-length cross-linker EDC in combination with sulfo-NHS. EDC was dissolved in water (shortly prior

to addition) and added in 500-, 1000-, and 2000-fold molar excess over the protein/peptide mixture. Sulfo-NHS, dissolved in water, was added immediately in 500-fold molar excess to all three EDC-containing reaction mixtures, thus resulting in EDC/sulfo-NHS ratios of 1:1, 2:1, and 4:1, respectively. For quenching the reactions, 200- $\mu$ L aliquots were taken from the reaction mixtures after 5, 15, 30, 60, and 120 min and DTT (40 mM final concentration) was added.

For cross-linking experiments with the homobifunctional cross-linking reagent BS<sup>3</sup> or sulfo-DST, an equimolar mixture of CaM and melittin (10  $\mu$ M each, final concentration) containing 1 mM CaCl<sub>2</sub> (20 mM HEPES buffer at pH 7.4) was incubated for 20 min at ambient temperature. BS<sup>3</sup> or sulfo-DST was dissolved in DMSO (stock solutions: 10, 50, and 100 mM), and 10  $\mu$ L was added to result in a 10-, 50-, and 100-fold molar excess over the proteins. To quench the reactions, 200- $\mu$ L aliquots were taken from the reaction mixtures after 5, 15, 30, 60, and 120 min and NH<sub>4</sub>HCO<sub>3</sub> (20 mM final concentration) was added.

**SDS–PAGE and Enzymatic Proteolysis.** After separation of the reaction mixtures by one-dimensional SDS–PAGE [15%, Coomassie staining (56)], the bands of interest were excised and in-gel digested as described previously (57). Depending on the volume of the gel pieces, between 5 and 10  $\mu$ L of enzyme solution was added, and the digest was incubated at 37 °C for 16 h. Trypsin alone (50 ng/ $\mu$ L) was used for all the cross-linking reaction mixtures, whereas endoproteinase AspN (40 ng/ $\mu$ L), trypsin/AspN (both 100 ng/ $\mu$ L), and LysC/AspN (both 100 ng/ $\mu$ L) were used as additional digestion enzymes of the CaM–melittin complex cross-linked with EDC/sulfo-NHS. When employing a combination of two proteases, only half of the volume of each was added to result in a 2-fold dilution (i.e., 50 ng/ $\mu$ L) of the initial concentration.

**MALDI-TOFMS.** Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOFMS) was performed on a Voyager DE RP Biospectrometry Workstation (Applied Biosystems, Foster City, CA) equipped with a nitrogen laser (337 nm). The instrument was run in positive ionization mode, and measurements were performed in linear mode using sinapinic acid as the matrix. The matrix was prepared in 30% acetonitrile, 70% water, and 0.1% trifluoroacetic acid. Samples were desalted employing Microcon YM-10 filters (Millipore, Eschborn, Germany) and prepared using the dried droplet method. A total of 100 shots were added to one spectrum in the mass range of  $m/z$  2000–42 000. The instrument was calibrated using cytochrome *c* ( $[M + H]^+_{\text{average}} = 12\,361.6$ ) and myoglobin ( $[M + H]^+_{\text{average}} = 16\,952.6$ ). Data acquisition and processing were performed using Voyager version 5.1 and Data Explorer version 4.0 (Applied Biosystems).

**Nano-HPLC/Nano-ESI-FTICRMS.** The peptide mixtures from the enzymatic digests were separated by nano-HPLC. Nano-HPLC was carried out on an Ultimate Nano-LC system (LC Packings, Amsterdam, The Netherlands) equipped with a Switchos II column switching module and a Famos Micro Autosampler with a 5- $\mu$ L sample loop. Samples were injected by the autosampler and concentrated on a trapping column (PepMap, C18, 300  $\mu$ m  $\times$  1 mm, 5  $\mu$ m, 100 Å, LC Packings) with water containing 0.1% formic acid at flow rates of 20  $\mu$ L/min. After 2 min, the peptides were eluted onto the

separation column (PepMap, C18, 75  $\mu$ m  $\times$  150 mm, 3  $\mu$ m, 100 Å, LC Packings), which had been equilibrated with 95% A (A being water and 0.1% formic acid). Peptides were separated using the following gradient: 0–30 min: 5–50% B, 30–31 min: 50–95% B, 31–35 min: 95% B (B being acetonitrile and 0.1% formic acid) at flow rates of 200 nL/min and detected by their UV absorptions at 214 and 280 nm.

The nano-HPLC system was coupled online to an Apex II FTICR mass spectrometer equipped with a 7 T superconducting magnet (Bruker Daltonics, Billerica, MA) and a nano-electrospray ionization source (Agilent Technologies, Waldbronn, Germany). For nano-ESIMS, coated fused-silica PicoTips (tip IDs = 8 and 15  $\mu$ m, New Objective, Woburn, MA) were applied. The capillary voltage was set to –1400 V. Mass spectral data were acquired over an  $m/z$  range of 400–2000; four scans were accumulated per spectrum, and 400 spectra were recorded for each LC/MS run. MS data acquisition was initialized with a trigger signal from the HPLC system at 5 min after initiation of the LC gradient. Data were acquired over 34.5 min. Calibration of the instrument was performed with CID fragments (capillary exit voltage = 200 V) of the LHRH peptide [y5 ( $m/z$  499.2987), b4 ( $m/z$  522.2096), LHRH ( $[M + H]^{2+}$   $m/z$  592.2358), y6 ( $m/z$  662.3620), y7 ( $m/z$  749.3941), b7 ( $m/z$  855.3784), y8 ( $m/z$  935.4734), b8 ( $m/z$  1011.4795), LHRH ( $[M + H]^+$   $m/z$  1183.5643)]. Data acquisition and processing were performed using XMASS versions 5.0.10 and 6.0 (Bruker Daltonics). Processing of the raw data was performed using the “Projection” tool in XMASS (58).

**Identification of Cross-Linking Products.** Cross-linking products were identified using GPMaw (General Protein Mass Analysis for Windows) versions 5.12 $\beta$ 3 and 6.00 (Lighthouse Data, Odense, Denmark) (available at <http://welcome.to/gpmaw>), the ExpASY Proteomics tools in the Swiss-Prot Database (available at [www.expasy.ch](http://www.expasy.ch)), and ASAP (Automatic Spectrum Assignment Program) version 1.09 (available at <http://roswell.ca.sandia.gov/~mmyoung/asap.html>). In case of doubt, cross-links assigned by the software were not considered as such when discrimination between a peptide and a cross-linking product or between two different cross-linking products, coincidentally having the same mass, was ambiguous. Cleavages by proteases at modified amino acids, such as the trimethylated K115 in CaM as well as amino acids modified by cross-linking reagents, were excluded. Both the N-terminus of CaM and the C-terminus of melittin were excluded from possible cross-linking because the former is acetylated and the latter amidated.

**Molecular Modeling.** Structures of the CaM–melittin complex were calculated by conjoined rigid body/torsion angle simulated annealing (49–51) using the molecular structure determination package Xplor-NIH (59) (available online at <http://nmr.cit.nih.gov/xplor-nih>), and structures were viewed using the VMD-XPLOR visualization package (60) (available at <http://vmd-xplor.cit.nih.gov>). Distance restraints derived from cross-linking data were represented by empirical  $\langle r^{-6} \rangle^{-1/6}$  averages. This ensures that at least one of the distances in the restraint lies within the specified target range, but no penalty ensues if the remaining distances within the restraint are much longer than the target distance. The distances were classified into three ranges,  $\leq 5$ ,  $\leq 8.5$ , and



<b>Calmodulin</b>		
<sup>1</sup> Ac-ADQLTEEQIA	<sup>11</sup> EFKEAFSLFD	<sup>21</sup> KDGDGITITTK
<sup>31</sup> ELGTVMRSLG	<sup>41</sup> QNPTEAELQD	<sup>51</sup> MINEVDADGN
<sup>61</sup> GTIDFPEFLT	<sup>71</sup> MMARKMKDTD	<sup>81</sup> SEEEIREAFR
<sup>91</sup> VFDKDGNGYI	<sup>101</sup> SAAELRHVMT	<sup>111</sup> NLGE <u>KL</u> TDEE
<sup>121</sup> VDEMIREADI	<sup>131</sup> DGDGQVNYEE	<sup>141</sup> FVQMMTAK
<b>Melittin</b>		
<sup>1</sup> GIGAVLKVLT	<sup>11</sup> TGLPALISWI	<sup>21</sup> KRKRQQ-NH <sub>2</sub>

FIGURE 1: Amino acid sequences of CaM and melittin. CaM was found to be *N*-acetylated (Ac-Ala) and to contain a trimethylated lysine in position 115 (underlined). The C-terminal amino acid of melittin is amidated.

≤ 13.4 Å, corresponding to cross-links obtained using EDC, sulfo-DST, and BS<sup>3</sup>, respectively. The starting coordinates employed for CaM and melittin were derived from the 2.2 Å resolution crystal structure of the CaM–smMLCK peptide complex [PDB accession code 1CDL (31)] and the 2 Å resolution crystal structure of helical melittin [PDB accession code 2MLT (9)], respectively. The N- (residues 5–70) and C-terminal (residues 81–146) halves of CaM and the N- (residues 1–9) and C-terminal (residues 13–26) halves of melittin were treated as rigid bodies. Residues 71–80 of CaM, residues 10–12 of melittin (i.e., the location of the kink between the two helical segments), and all interfacial side chains were given full torsional degrees of freedom. The nonbonded term in the target function comprised a quartic van der Waals repulsion term (61) and a torsion angle database potential of the mean force derived from high-resolution crystal structures (62). The latter ensures that the side chain torsion angles lie within energetically allowed rotamer ranges. A qualitative interpretation of the cross-linking data indicated that melittin could bind CaM in two opposing orientations. Consequently, both orientations were refined simultaneously by including two complete sets of nonoverlapping coordinates (i.e., no interactions were allowed between the two sets).

## RESULTS

**Characterization of CaM and Melittin.** The basis underlying the investigation of intermolecular interaction sites between CaM and melittin in the CaM–melittin complex is the assessment of detailed descriptions of their respective primary structures. Peptide mass fingerprinting of both CaM and melittin yielded full-sequence coverage. CaM was found to be *N*-terminally acetylated and to contain a trimethylated lysine at position 115 (63), whereas the C terminus of melittin was confirmed to be amidated. Accordingly, exact mass measurements, after spectral deconvolution, of the intact protein and peptide performed by ESI-FTICRMS yielded a base peak at  $m/z$  16 790.921 for CaM (base peak from simulation at  $m/z$  16 790.884,  $\Delta m = 0.037$  mass units, 2.1 ppm) and a monoisotopic signal at  $m/z$  2845.756 for melittin (calculated mass:  $[M + H]^+_{\text{monoisotopic}} = 2845.762$ ,  $\Delta m = 0.006$  mass units, 2.1 ppm). The amino acid sequences of CaM and melittin are shown in Figure 1.

**Cross-Linking Reactions.** Cross-linking experiments with CaM and melittin were performed as described under the Experimental Procedures. Figure 2 provides a schematic overview of the analytical strategy employed. The zero-length cross-linker EDC in combination with sulfo-NHS, as well as the homobifunctional, amine-reactive cross-linking re-

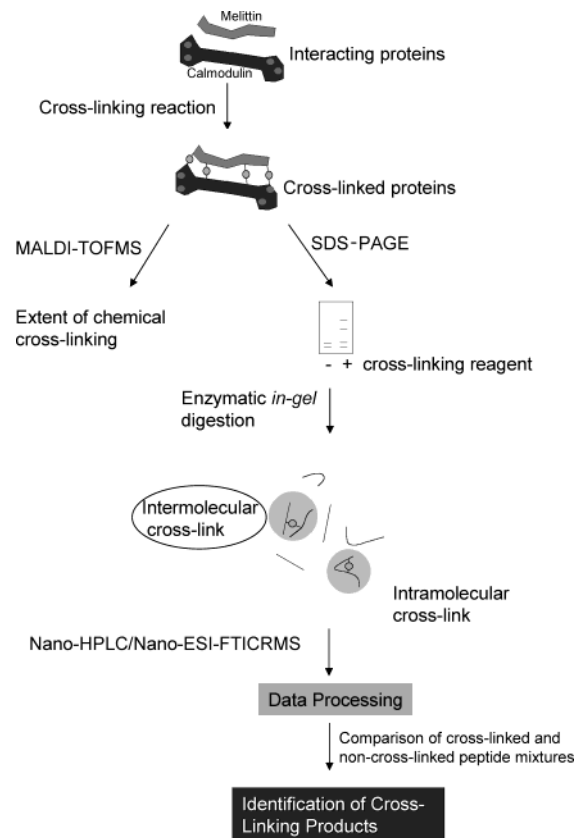


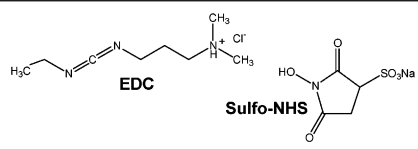
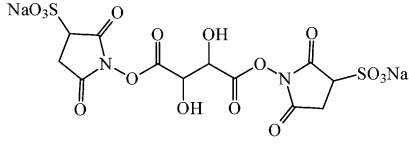
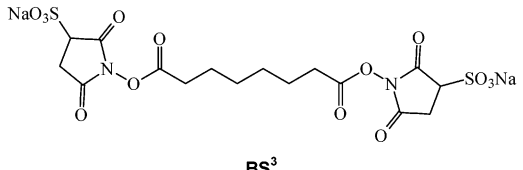
FIGURE 2: General analytical strategy for mapping the interacting sequences between CaM and melittin.

agents sulfo-DST and BS<sup>3</sup> (Table 1), was used. EDC reacts with the carboxylic acid group of an aspartic or glutamic acid or the C terminus of a protein to form a short-lived active *O*-acylurea intermediate. Sulfo-NHS reacts with the EDC active-ester complex and extends the half-life of the activated carboxylate. In the presence of an  $\epsilon$ -amino group of lysine or the free N-terminus of a protein, an amide bond is formed, leading to the loss of an H<sub>2</sub>O molecule and thus a mass decrease of 18.011 mass units per cross-link.

Homobifunctional cross-linking reagents contain two identical functional groups on either side of the molecule that are separated by a spacer bridging a defined distance. Homobifunctional sulfo-NHS esters, such as sulfo-DST and BS<sup>3</sup>, are highly reactive toward  $\alpha$ - and  $\epsilon$ -amino groups, and their sulfonate groups provide sufficient water solubility. However, they are susceptible to hydrolysis. Upon cross-linking, sulfo-DST and BS<sup>3</sup> produce amide bond cross-linked molecules causing a mass shift of 113.995 and 138.068 mass units, respectively, whereas partially hydrolyzed cross-linkers exhibit a mass increase of 132.006 and 156.079 mass units, respectively.

**SDS-PAGE.** The cross-linking reaction mixtures were separated by one-dimensional SDS-PAGE, and the gels were stained with Coomassie Brilliant Blue. Figure 3 shows a gel of melittin, CaM, and cross-linking reaction mixtures of an equimolar mixture of both proteins employing EDC/sulfo-NHS at a 2:1 ratio after incubation times of 5, 15, 30, 60, and 120 min. CaM and melittin possess molecular masses of around 16.8 and 2.8 kDa, respectively. Thus, the cross-linked complex between CaM and melittin is located at around 19.6 kDa depending on the extent of chemical cross-linking. After an incubation time of 5 min, three distinct

Table 1: Cross-Linking Reagents Used for CaM–Melittin Cross-Linking, with Their Respective Spacer Lengths and Specificities

Cross-Linking Reagent	Spacer length	Specificity
 <p>EDC Sulfo-NHS</p>	zero-length	amide bond formation between primary amine and carboxylic acid
 <p>Sulfo-DST</p>	6.4 Å	homobifunctional, reactive towards primary amines
 <p>BS<sup>3</sup></p>	11.4 Å	homobifunctional, reactive towards primary amines

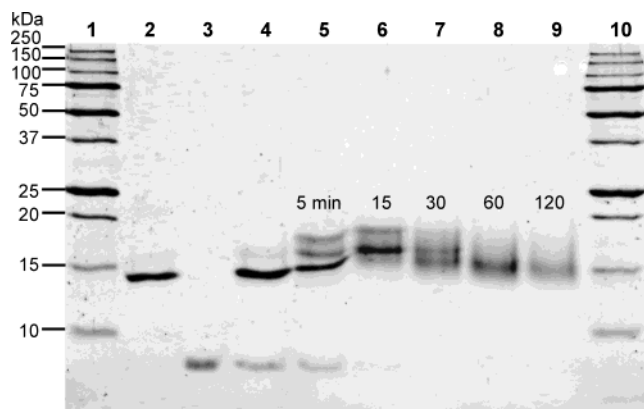


FIGURE 3: SDS–PAGE of cross-linking reaction mixtures (10  $\mu$ M CaM/10  $\mu$ M melittin) with EDC (10 mM)/sulfo-NHS (5 mM). Lanes 1 and 10: protein markers. Lane 2: CaM (10  $\mu$ M). Lane 3: melittin (10  $\mu$ M). Lane 4: CaM and melittin (10  $\mu$ M each) without EDC/sulfo-NHS. Lane 5: 5 min incubation time. Lane 6: 15 min incubation time. Lane 7: 30 min incubation time. Lane 8: 60 min incubation time. Lane 9: 120 min incubation time.

bands between 17 and 20 kDa are observed, as well as a faint band of melittin, indicating that cross-linking was not complete. After 15 min, only two bands can be clearly distinguished. When the incubation times are increased to 30, 60, and 120 min, the bands fuse and the mobility on the gel changes, indicating that alongside intermolecular cross-linking intramolecular cross-linking had also occurred. The bands on the gels were excised separately whenever it was possible to tell them apart; otherwise the whole region was cut out, and the gel pieces were subsequently used for enzymatic in-gel digestion. Aggregation of proteins caused by excessive cross-linking was not observed for any of the cross-linking reagents as evidenced by the absence of any gel bands in the higher mass range.

**MALDI-TOFMS.** MALDI-TOFMS was employed to estimate the extent of chemical cross-linking over the course of the cross-linking reaction. Figure 4 shows MALDI-TOF mass spectra of the nondigested reaction mixtures after

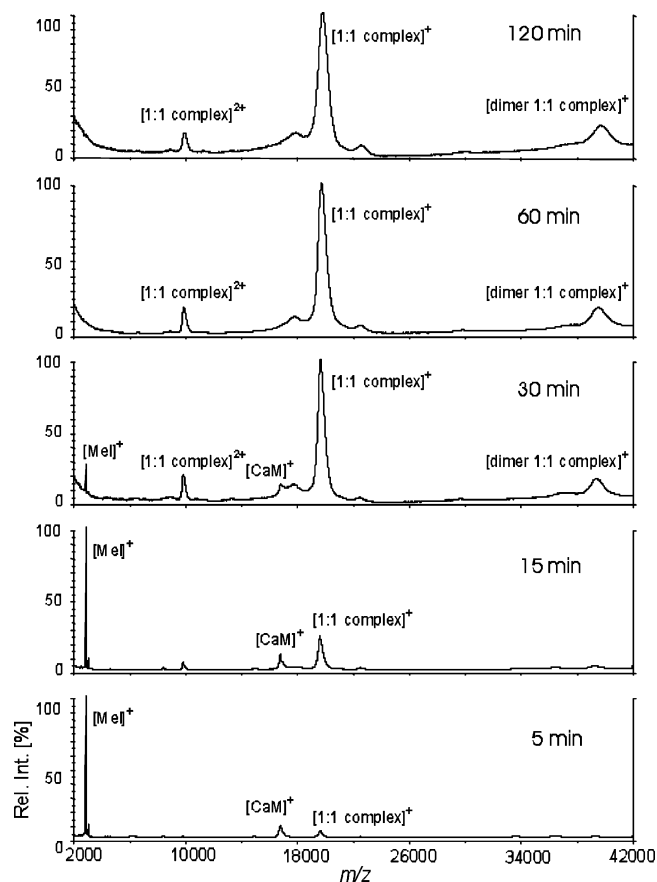


FIGURE 4: MALDI-TOF mass spectra of the nondigested reaction mixtures after incubation times of 5, 15, 30, 60, and 120 min (bottom to top), in 1000-fold excess of EDC and EDC/sulfo-NHS (2:1).

incubation times of 5, 15, 30, 60, and 120 min (bottom to top) with an EDC/sulfo-NHS ratio of 2:1. At an incubation time of 5 min, three distinct peaks appear at  $m/z$  2848.4, 16 792.3, and 19 622.9 corresponding to the singly charged ions of melittin, CaM, and the cross-linked CaM–melittin

(1:1) complex, respectively. The observed average masses are in good agreement with the molecular mass of protonated melittin ( $[M + H]^+_{\text{average}} = 2847.5$ ,  $\Delta m = 0.9$  Da), CaM with the posttranslational modifications stated above ( $[M + H]^+_{\text{average}} = 16\,791.6$ ), and the complex ( $[M + H]^+_{\text{average}} = 19\,638.1$ ) lacking 18.01 mass units because of one cross-link in the molecule. With increasing incubation time, the relative intensity of the singly charged signal of the CaM–melittin complex increases, whereas the relative intensity of the melittin signal steadily decreases and is no longer observed at incubation times beyond 60 min. Furthermore, the peak of the CaM–melittin (1:1) complex broadens indicating the formation of intramolecular cross-links in addition to intermolecular cross-linking products, when incubation times are increased. Average masses for the CaM–melittin (1:1) complexes range from  $m/z$  19 602.4 to 19 833.2, steadily increasing during the time course. An explanation for a mass increase in the cross-linked complexes with increasing incubation time could be that the sulfo-NHS group remains attached to the protein if a reactive amino group is not present for the cross-linking reaction. Although MALDI preferentially generates singly charged ions, signals of the doubly charged ions of the CaM–melittin (1:1) complex are also observed between  $m/z$  9779.7 and 9881.1. Signals of the singly charged ions of the complex dimer are visible between  $m/z$  39 239.2 and 39 479.3.

**In-Gel Enzymatic Digestion.** After separation by SDS–PAGE, the excised gel pieces of the complex were subjected to enzymatic in-gel digestion with various proteases (Figure 2). For comparative purposes, gel bands of CaM and melittin were excised and treated in the same fashion as the cross-linked complex. For enzymatic in-gel digestion, the proteases trypsin and AspN were applied in addition to combinations of trypsin/AspN and LysC/AspN. The latter two protease mixtures were exclusively applied to isolated CaM and melittin as well as to cross-linking products created using EDC in 1000-fold molar excess over the protein/peptide (2:1 EDC/sulfo-NHS ratio). Cross-linked complexes derived from cross-linking with sulfo-DST and BS<sup>3</sup>, both applied in a 50-fold molar excess over the protein/peptide, were digested by trypsin alone.

**Analysis of Cross-Linking Products.** The peptide mixtures generated from enzymatic in-gel digestion were subsequently separated by nano-HPLC and analyzed by nano-ESI-FT-ICRMS (Figure 2). These mixtures were rather complex containing peptides from CaM and melittin, as well as inter- and intramolecular cross-linking products. The experimentally obtained monoisotopic masses were compared with calculated masses of peptides and cross-linking products employing the GPMW and ASAP software packages. In the Discussion that follows, the cross-linking products obtained will be grouped into two different categories, namely, according to the two different binding orientations melittin was found to adopt in the CaM–melittin complex.

**Peptide Mass Mapping.** To obtain additional information on the regions of CaM involved in the interaction with melittin, we compared peptide maps from cross-linked CaM–melittin (1:1) complexes with those from non-cross-linked CaM and melittin. Sequence coverage maps were generated for samples cross-linked with different amounts of EDC/sulfo-NHS, which had been in-gel digested with trypsin and endoprotease AspN. With trypsin, full-sequence

coverage (100%) was obtained for CaM as well as for CaM–melittin complexes cross-linked with EDC/sulfo-NHS (1:1 and 2:1). However, when EDC/sulfo-NHS was applied at a ratio of 4:1, the sequence coverage decreased to 89.19%. Only 132 of the 148 amino acids of CaM were detected, whereas residues 75–90 were missing, which belong to the central  $\alpha$ -helix of CaM.

Digestion with endoprotease AspN led to different results. Peptide mass mapping with AspN yielded a sequence coverage of 98.65% for CaM, lacking amino acids D56 and A57, because the mass of this peptide is too small to be detected. For EDC/sulfo-NHS (1:1), sequence coverage of 87.84% was obtained with amino acids D56 and A57, as well as the sequence from residues 88 to 103 not being detected. With increasing amounts of EDC, a decline in the sequence coverage was observed. Digestion of the complex cross-linked with EDC/sulfo-NHS (2:1) yielded a sequence coverage of 45.27%, lacking the sequences comprising residues 20–49, 56–60, and 93–148. For the complex cross-linked with EDC/sulfo-NHS (4:1), only 16.89% of the CaM sequence was covered, lacking residues 20–86 and 93–148. These findings gave additional hints that the C- and N-terminal domains, as well as the central  $\alpha$ -helix, of CaM were involved in complex formation with melittin.

**Cross-Linking Products Obtained with EDC/Sulfo-NHS.** The results of the cross-linking experiments with EDC/sulfo-NHS are summarized in Table 2. From the reaction mixtures created with EDC/sulfo-NHS (1:1), no cross-linking products were detected after digestion with trypsin; however, when employing an EDC/sulfo-NHS ratio of 2:1 (Table 2, samples III), a number of cross-linking products were observed. In the cross-linking reaction mixtures with EDC/sulfo-NHS (4:1) (Table 2, samples IV), three cross-linking products were found, one of which had not been observed when working with EDC/sulfo-NHS (2:1). In addition to tryptic digestion of the cross-linked complexes, endoprotease AspN, trypsin/AspN, and LysC/AspN mixtures were used for digestion of the EDC/sulfo-NHS (2:1) cross-linking reaction mixtures. However, when using AspN alone as the digestion enzyme, we did not detect any cross-linking products. Extending HPLC elution of the digestion products with 95% acetonitrile to 60 min did not overcome this problem. Overall, most of the cross-linking products were observed in more than one sample (Table 2), and often the analogues containing an oxidized methionine were detected as well, which further increased the certainty of a correct assignment.

In total, twelve cross-linking products were observed for the CaM–melittin (1:1) complex cross-linked with EDC/sulfo-NHS after digestion with trypsin, trypsin/AspN, and LysC/AspN (Table 2). Because it is not possible in most cases to exactly identify the actual cross-linked amino acid, Table 2 lists a number of amino acids that are potentially involved in cross-linking.

Eight cross-linking products point conclusively to an orientation of melittin in the complex, which is inverted relative to that of the majority of the other CaM binding peptides studied to date, and we designate this binding mode as orientation A (Figure 5). However, we also found four additional cross-linking products between melittin and CaM, which did not match the orientation of melittin in the complex as suggested by orientation A (Figure 5). For example, residues 127–148 in the C-terminal domain of

Table 2: Cross-Linking Products of CaM and Melittin with the Cross-Linking Reagent EDC/Sulfo-NHS<sup>a</sup>

CaM	modified residue	melittin	modified residue	[M + H] <sup>+</sup> <sub>calcd</sub>	[M + H] <sup>+</sup> <sub>exp</sub>	ppm	sample	protease(s)
<b>Orientation A</b>								
1–13	D2, E6, E7, E11	8–22	K21	3212.757	3212.742	5	III15	T
					3212.750	2	IV30	
					3212.757	0	IV60	
					3212.759	1	IV120	
14–30	E14, D20, D22, D24	8–22	K21	3493.894	3493.884	3	IV120	T
78–86	D78, D80, E82, E83, E84	1–7	G1	1731.876	1731.872	2	III5	T
					1731.872	2	III15	
					1731.877	1	III30	
104–121	E114, D118, E119, E120	22–24	K23	2595.404	2595.393	4	III120	T/A
107–126		23–24	K23	2701.365	2701.363	1	III5	T
	E114, D118, D122, E119, E120, E123			2717.360	2701.362	1	III15	
					2717.356	1	III60	
118–130	E119, E120, D122, E123, E127, D129	22–24	K23	2003.981	2003.967	7	III30	T/A
					2003.970	6	III120	
120–126	D122, E123	23–24	K23	1191.615	1191.615	0	III120	T/A
127–148	E127, D129, D131, D133, E139, E140	23–24	K23	2774.276	2774.276	0	III5	T
				2790.271	2774.268	3	III15	
					2774.267	3	III30	
					2790.271	0	III5	
					2790.271	0	III15	
<b>Orientation B</b>								
1–13	D2, E6, E7, E11	1–7	G1	2202.165	2202.159	3	III5	T
					2202.157	4	III15	
					2202.159	3	III30	
					2202.166	0	IV60	
					2202.167	1	IV120	
1–13	D2, E6, E7, E11	1–7	G1	2202.165	2202.176	5	III30	L/A
14–30	E14, D20, D22, D24	1–7	G1	2483.303	2483.298	2	III5	T
					2483.301	1	III15	
					2483.307	2	III30	
					2483.302	0	III60	
127–148	E127, D129, D131, D133, E139, E140	1–7	G1	3144.486	3144.478	3	III15	T
				3160.481	3160.487	2	III15	

<sup>a</sup> EDC/Sulfo-NHS Were Applied in 1000-Fold/500-Fold Excess over the Protein/Peptide, (Ratio 2:1, Samples III) and 2000-Fold/500-Fold Excess over Protein (Ratio 4:1, Samples IV). Digestions Were Performed with the Proteases Trypsin (T), Trypsin and AspN (T/A), and LysC and AspN (L/A). Reaction Times (in Minutes) Are Given in Numbers (5–120).

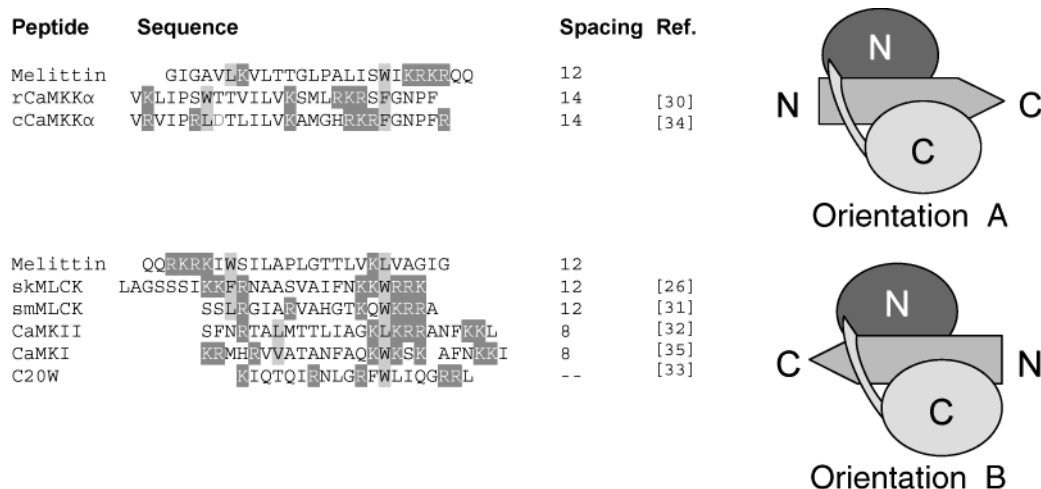


FIGURE 5: Sequence alignment of different CaM target sequences and schematic models for CaM binding. In the upper panel, the peptides are oriented from N- to C-terminus; in the lower panel, peptides are oriented from C- to N-terminus. Hydrophobic anchor residues as well as basic amino acids are marked. Figures of models A and B are adapted from ref 30.

CaM (containing acidic residues E127, D129, D131, D133, E139, and E140) were found to be cross-linked to residues 1–7 of melittin via the free N terminus of the peptide corresponding to signals at *m/z* 3144.486 (M144 or M145 oxidized) and 3160.481 (both M144 and M145 oxidized) when using trypsin as the digestion enzyme (Figure 6). Conclusively, we did not detect cross-linking products in the

N-terminal domain and central linker region of CaM when employing the proteases trypsin and AspN simultaneously.

When using a LysC/AspN combination for digestion of the cross-linked complex, we detected only one cross-linking product from samples of the CaM–melittin complex cross-linked with an EDC/sulfo-NHS ratio of 2:1 (Table 2). The identified cross-linking product comprised CaM residues



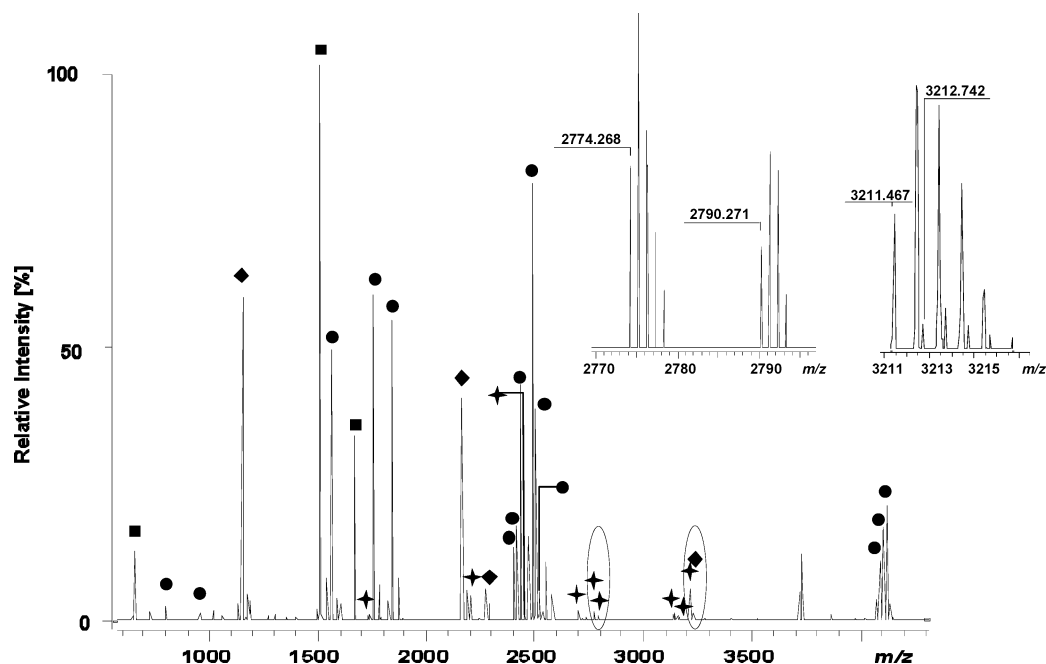


FIGURE 6: Nano-ESI-FTICR mass spectrum of the tryptic peptide mixture from the CaM-melittin complex cross-linked with EDC/sulfo-NHS (2:1) at an incubation time of 15 min: (star) cross-linking product, (■) melittin peptide, (●) CaM peptide, and (◆) autolytic tryptic peptide. The magnified inserts show a cross-linking product between CaM residues 1-13 and melittin residues 8-22 ( $[M + H]^+_{\text{calcd}} m/z$  3212.757, 4.7 ppm) and an autolytic peptide of trypsin ( $[M + H]^+_{\text{calcd}} m/z$  3211.474, 2.2 ppm, residues 160-190), as well as a cross-linking product between CaM residues 127-148 and melittin residues 23-24 ( $[M + H]^+_{\text{calcd}} m/z$  2774.276, 2.9 ppm), which was also detected with one methionine residue oxidized ( $[M + H]^+_{\text{calcd}} m/z$  2790.271, 0 ppm). The monoisotopic masses of the cross-linking products (left to right) in this spectrum are  $[M + H]^+_{\text{calcd}} m/z$  1731.876, 2202.165, 2483.303, 2701.365, 2774.276, 2790.271, 3144.486, 3160.481, and 3212.757. For detailed information on the cross-linking products, see Table 3. Peptides derived from melittin: residues 1-7 ( $[M + H]^+_{\text{calcd}} m/z$  675.429), residues 8-22 ( $[M + H]^+_{\text{calcd}} m/z$  1511.920), and residues 8-23 ( $[M + H]^+_{\text{calcd}} m/z$  1668.021). Peptides from CaM: residues 31-37 ( $[M + H]^+_{\text{calcd}} m/z$  805.424), residues 14-21 ( $[M + H]^+_{\text{calcd}} m/z$  956.472), residues 1-13 with acetylated N-terminus ( $[M + H]^+_{\text{calcd}} m/z$  1563.754), residues 91-106 ( $[M + H]^+_{\text{calcd}} m/z$  1754.871), residues 14-30 ( $[M + H]^+_{\text{calcd}} m/z$  1844.891), residues 107-126 with trimethyllysine K115 ( $[M + H]^+_{\text{calcd}} m/z$  2401.174, 2417.169 (1 $\times$  MSO), and 2433.163 (2 $\times$  MSO)), residues 127-148 ( $[M + H]^+_{\text{calcd}} m/z$  2490.080, 2506.075 (1 $\times$  MSO), and 2522.070 (2 $\times$  MSO)), residues 38-74 ( $[M + H]^+_{\text{calcd}} m/z$  4085.841 (1 $\times$  MSO), 4101.836 (2 $\times$  MSO), and 4117.831 (3 $\times$  MSO)). Peptides derived from trypsin autolysis:  $[M + H]^+_{\text{calcd}} m/z$  1153.574, 2163.056, 2273.159, 2289.154, and 3211.474.

1-13 and melittin residues 1-7 ( $m/z$  2202.165), pointing toward any of the four acidic residues in CaM residues 1-13 (D2, E6, E7, or E11) being conjugated with the  $\alpha$ -amino group of melittin's free N-terminus. The identical cross-linking product was observed in the tryptic digest. We believe that the poor detection of cross-linking products when using a LysC/AspN combination is due to AspN not cleaving melittin, because melittin does not contain any acidic amino acids. Because the lysine residues (K7, K21, and K23) of melittin are at least partly modified by the cross-linking reagent, LysC is also unable to cleave. Thus, the cross-linking products created in this manner are not amenable to observation in the optimum ESI-FTICRMS detection range between  $m/z$  400 and 2000.

The combined results provide evidence that melittin is also oriented in a way that corresponds to the peptide orientation found in the majority of the CaM-peptide complexes studied to date (26, 31-33, 35) and that we denote as orientation B (Figure 5).

An example of the quality of the data is provided by Figure 6, which shows the deconvoluted nano-ESI-FTICR mass spectrum obtained from the tryptic digest of the CaM-melittin complex cross-linked with EDC/sulfo-NHS (2:1) after an incubation time of 15 min. Signals of CaM and melittin, as well as signals of trypsin autolytic peptides, are visible, with signals of the cross-linking products displaying only low intensities. The magnified inserts demonstrate the

high resolution and mass accuracy provided by FTICRMS. For example, the cross-linking product at  $m/z$  3212.742 is clearly resolved from the signal at  $m/z$  3211.467 belonging to an autolytic fragment (sequence 160-190, M180 oxidized) of trypsin. The mass accuracies for both signals were 4.7 and 2.2 ppm, respectively; the average mass accuracy in the cross-linking experiment was 1.7 ppm using EDC/sulfo-NHS as the cross-linking reagents and trypsin as the protease.

**Cross-Linking Products Obtained with Sulfo-DST.** In addition to EDC/sulfo-NHS, CaM and melittin were cross-linked with the homobifunctional, amine-reactive cross-linker sulfo-DST (Table 3). After tryptic digestion of the CaM-melittin (1:1) complex, three intermolecular cross-linking products were observed, one of them containing an additional hydrolyzed cross-linker, and three peptides were found to be modified by a hydrolyzed cross-linker alone (Table 3). Peptides modified by a hydrolyzed cross-linker do not provide any structural information on protein interfaces but can yield important information on surface accessibility of certain amino acids.

Two cross-linking products are consistent with orientation A of melittin in the complex (Figure 5). Cross-linking between K13 of CaM (amino acid sequence 1-21) and K23 of melittin within amino acid sequence 22-24 (signal at  $m/z$  3073.511) was observed for all five incubation times.

Only one cross-linking product consistent with orientation B was observed; the product between residues 76-90 of



Table 3: Cross-Linking Products of CaM and Melittin Using the Cross-Linking Reagent Sulfo-DST (Tryptic Cleavage), 50-Fold Excess of Cross-Linker over Proteins<sup>a</sup>

CaM	modified residue	melittin	modified residue	[M + H] <sup>+</sup> <sub>calcd</sub>	[M + H] <sup>+</sup> <sub>exp</sub>	ppm	sample
<b>Orientation A</b>							
1–21	K13	22–24	K23	3073.511	3073.529	6	50–5
					3073.491	7	50–15
					3073.503	3	50–30
					3073.485	9	50–60
					3073.532	7	50–120
1–37 +OH-DST	K13, K21, K30	22–24	K23	4896.337	4896.327	2	50–5
					4896.385	10	50–15
<b>Orientation B</b>							
76–90	K77	8–22	K21	3636.858	3636.864	2	50–15
					3636.878	6	50–60
					3636.869	3	50–120
<b>Hydrolyzed Linker</b>							
14–30 +OH-DST	K21	1–7 +OH-DST	G1	789.435	789.433	3	50–15
75–86 +OH-DST	K75, K77			1976.897	1976.894	2	50–120
				1628.696	1628.701	3	50–15

<sup>a</sup> Reaction Times (in Minutes) Are Given in Numbers (5–120).

Table 4: Cross-Linking Products of CaM and Melittin Using the Cross-Linking Reagent BS<sup>3</sup> (Tryptic Cleavage) 50:50-Fold Excess of Cross-Linker over Proteins<sup>a</sup>

CaM	modified residue	melittin	modified residue	[M + H] <sup>+</sup> <sub>calcd</sub>	[M + H] <sup>+</sup> <sub>exp</sub>	ppm	sample
<b>Orientation A</b>							
14–30	K21	23–24	K23	2285.166	2285.168	1	50–60
76–90	K77	1–7	G1	2666.334	2666.334	0	50–15
<b>Orientation B</b>							
14–30	K21	1–7	G1	2639.381	2639.383	1	50–60
					2639.363	7	50–120
<b>Hydrolyzed Linker</b>							
14–30 +OH-BS <sup>3</sup>	K21			2000.970	2000.968	1	50–30
					2000.968	1	50–60
					2000.960	5	50–120
75–90 +OH-BS <sup>3</sup>	K75, K77			2785.360	2785.339	8	50–60

<sup>a</sup> Reaction Times (in Minutes) Are Given in Numbers (5–120).

CaM and residues 8–22 of melittin was observed at  $m/z$  3636.858 in three samples, pointing to K77 of CaM being cross-linked to K21.

In summary, for sulfo-DST, we detected only cross-linking products created between melittin and the N-terminal domain of CaM and the central helix of CaM, but no cross-linking between melittin and the C-terminal domain of CaM was observed.

*Cross-Linking Products Obtained with BS<sup>3</sup>.* Table 4 summarizes the cross-linking products obtained with BS<sup>3</sup> after tryptic digestion of the CaM–melittin (1:1) complex. As in the case of the experiments with sulfo-DST, two cross-linking products were observed consistent with orientation A, and one was observed consistent with orientation B.

The cross-linking product between residues 14–30 of CaM and residues 23–24 of melittin (signal at  $m/z$  2285.166) was observed once (in sample 50–60), indicative of a cross-link between K21 of CaM and K23 of melittin. In addition, residues 76–90 of CaM and residues 1–7 of melittin were found to be cross-linked (signal at  $m/z$  2666.334) via K77 of CaM and the N terminus of melittin.

One cross-linking product indicative of the orientation B was observed with a signal at  $m/z$  2639.381, corresponding to a cross-linking product between residues 14–30 of CaM and 1–7 of melittin. In this instance, K21 of CaM was cross-linked with the free N terminus of melittin. As in the case

of sulfo-DST, we only observed cross-linking products from BS<sup>3</sup> between melittin and the N-terminal domain of CaM and the central helix of CaM, but none were observed between melittin and the C-terminal domain of CaM. The reason we did not observe any cross-linking products between melittin and the C-terminal lobe of CaM with neither sulfo-DST nor BS<sup>3</sup> is probably due to the fact that there are only two lysine residues in the C-terminal domain of CaM. One is K115, which is trimethylated, and the other one is K148, which represents the C-terminal amino acid of CaM.

*Structural Model of the CaM–Melittin Complex.* Figure 7 provides a schematic overview of the combined results from cross-linking experiments between melittin and CaM using the three cross-linking reagents EDC/sulfo-NHS, sulfo-DST, and BS<sup>3</sup> and applying different proteases. Low-resolution structural models were calculated from ambiguous distance restraints derived from the cross-linking data using conjoined rigid body/torsion angle simulated annealing, as described in the Experimental Procedures. The backbone and noninterfacial side chains of the N- and C-terminal domains of CaM were treated as rigid bodies connected by a linker with full-torsional degrees of freedom. Likewise, the two helical halves of melittin were treated as rigid bodies connected by a linker (residues 10–13) with torsional degrees of freedom. All interfacial side chains were given torsional degrees of freedom. The distance restraints are represented

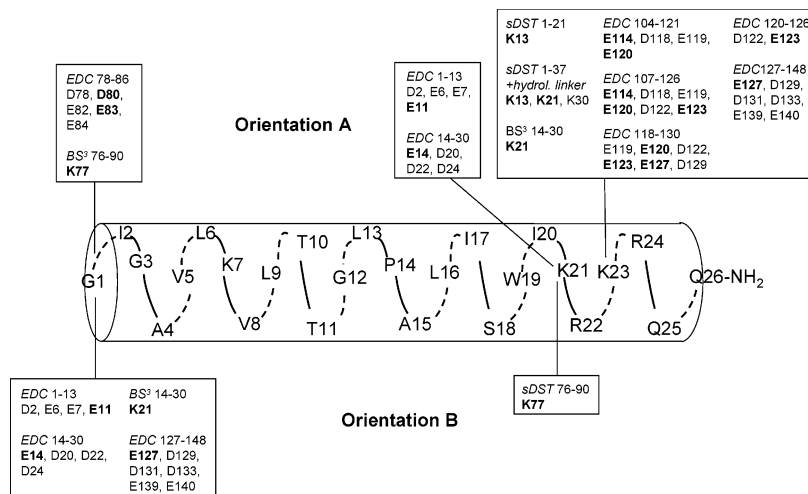


FIGURE 7: Schematic model of cross-linked residues between melittin and CaM. Cross-linked residues of CaM are located in the C- or N-terminal domain or in the central linker and can be assigned to orientation A or B (Figure 5). Amino acids, which are most likely to be involved in cross-linking, are printed in bold.

as an empirical  $\langle r^{-6} \rangle^{-1/6}$  average and allow for both ambiguity in terms of cross-linking assignments and melittin orientation. The two orientations of bound melittin were calculated simultaneously. A ribbon diagram of a representative pair of structures with opposite orientations of melittin in the CaM–melittin complex is displayed in Figure 8. The principal limitation in generating higher resolution structures from the cross-linking data is due to the presence of only three hinge residues for melittin (K21, K23, and the N terminus).

## DISCUSSION

*Chemical Cross-Linking and Mass Spectrometry for Studying Mechanisms of Protein–Protein Interactions.* In the present paper, we demonstrate that chemical cross-linking techniques combined with mass spectrometry and appropriate computational methods have the potential of providing the basis for an efficient low-resolution three-dimensional structural characterization of protein complexes. When an integrated approach was utilized comprising chemical cross-linking, high-resolution FTICRMS, and conjoined rigid body/torsion angle simulated annealing, we were able to gain novel structural information on the interaction mechanism underlying target peptide recognition by CaM, using the CaM–melittin complex as a model system.

The promise of mass spectrometry for this kind of protein–protein interaction study is apparent because of its high sensitivity enabling rapid analysis of the complex mixtures obtained from enzymatic digests of cross-linking reaction mixtures. If one can obtain a sufficient number of cross-links, these can be used as distance restraints, albeit ambiguous, to determine a low-resolution structure of a protein–protein complex, provided that high-resolution structures of the free proteins are available. In this regard, this approach presents a low-resolution alternative to NMR spectroscopy and X-ray-crystallography-based methods for low-resolution structural studies of macromolecular complexes.

*Structures of CaM–Peptide Complexes.* The structures of the CaM–peptide complexes solved to date (26, 30–32) identified several important features for target peptide binding to CaM. The  $\alpha$ -helical peptide is clamped by the N- and C-terminal domains of CaM, each of which contains a

hydrophobic cleft that can easily accommodate bulky hydrophobic anchor residues. Two channel outlets have been defined for CaM, with channel outlet 2 (CO-2) being more negatively charged than channel outlet 1 (CO-1). Thus, CO-2 is thought to contribute to peptide binding orientation through electrostatic interactions with the peptide's basic cluster at its one side (30). The two domains of CaM are shifted relative to each other with the N-terminal domain being more oriented toward CO-1 and the central linker, whereas the C-terminal domain of CaM is shifted toward CO-2. Therefore, CO-1 is partially covered by CaM's central linker, causing CO-1 to be slightly smaller than CO-2.

Figure 5 depicts several target peptide sequences, as well as their respective orientations, within CaM–peptide complexes. In most previously known CaM–target peptide complexes, such as CaM complexed with skMLCK peptide (26), smMLCK (31), and CaMKII (32), the peptide  $\alpha$ -helix is positioned in a way that the N-terminal portion of the peptide interacts with the C-terminal domain of CaM, whereas the C-terminal portion binds to the N-terminal domain of CaM (orientation B of Figure 5). However, there are also complexes in which the peptide's orientation is reversed, which is the case for CaM–CaMKK $\alpha$  complexes (30, 34). For these peptides, the N-terminal portion interacts with the N-terminal domain of CaM, whereas the C-terminal portion of the peptide binds to the C-terminal domain of CaM as shown in orientation A (Figure 5).

In addition to electrostatic interactions, there appear to be steric determinants that underlie the orientation of the peptide, which are defined by the distance that separates the two hydrophobic anchor residues in the target peptide. In melittin, as well as in the skeletal- and smooth-muscle MLCK peptides, there is a spacing of 12 residues between the two anchoring amino acids W and L (for melittin and smMLCK peptide) or W and F (for skMLCK peptide) (Figure 5). In rCaMKK $\alpha$ , however, two hydrophobic anchor residues are separated by a stretch of 14 amino acids, which causes the peptide to contain an additional well-defined  $\beta$  hairpin-like loop that folds back into the complex (30). The synthetic peptide C20W binds to only the C-terminal domain of CaM owing to the absence of a second hydrophobic anchor residue.

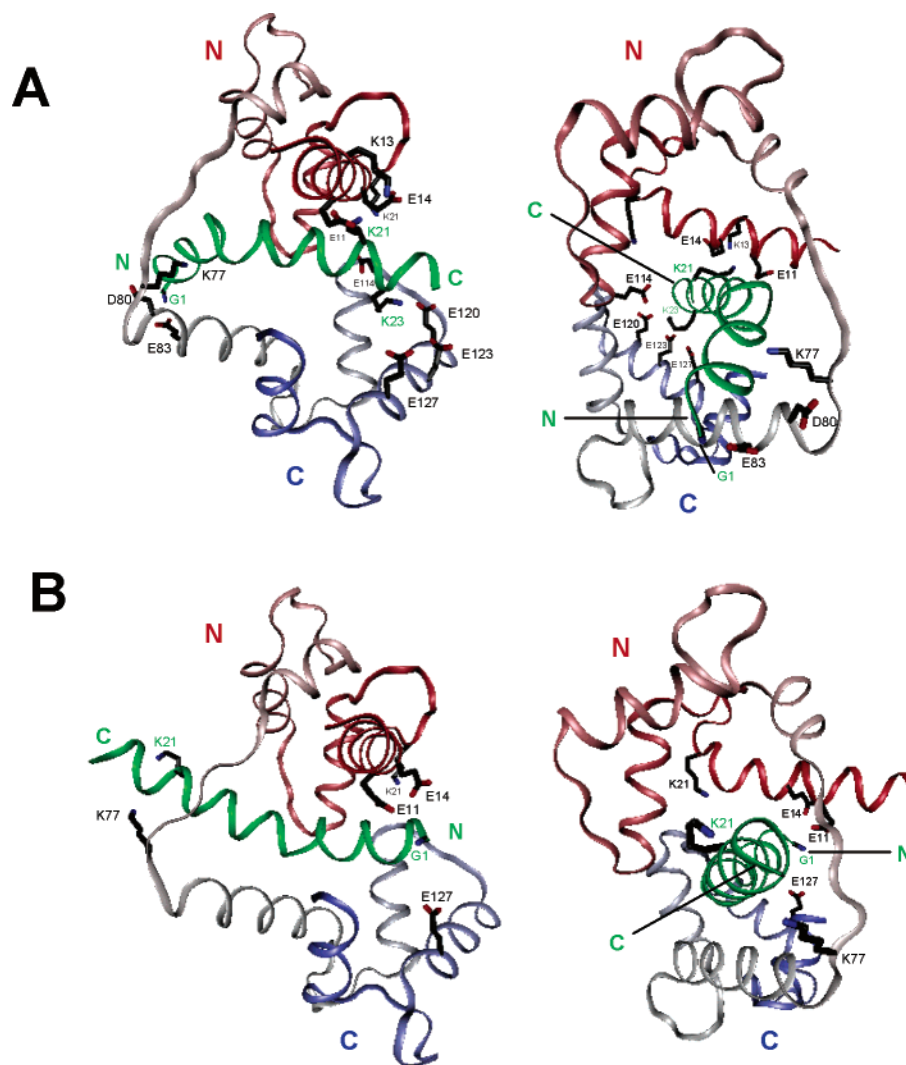


FIGURE 8: Parallel (A) and antiparallel (B) modes of binding of melittin in the CaM–melittin complex calculated from ambiguous distance restraints derived from the cross-linking data using conjoined rigid body/torsion angle simulated annealing. The amino acid side chains of CaM and melittin that are involved in cross-linking are shown for both orientations A and B. Each orientation is represented by two views: one with the helix axis of melittin parallel (left) and one with the helix axis perpendicular (right) to the plane of the page.

Recently, two novel structures of CaM complexed with large target fragments, namely, the calcium-activated  $K^+$  channel (38) and the anthrax exotoxin (39), have been solved and revealed unexpected binding modes. This included dimer formation with the fragments of the  $K^+$  channel and insertion of CaM between two domains of the anthrax toxin. This suggests that structural data on CaM complexed with large target proteins yield additional information on the mechanism of CaM target recognition and demonstrate that CaM possesses a variety of modes of interaction with its targets.

**Structure of the CaM–Melittin Complex.** To date, there has only been one structural study, on the basis of limited proteolysis and chemical cross-linking experiments with EDC, of the CaM–melittin complex (55). In this previous study, an exclusive orientation of melittin in the complex, corresponding to the parallel orientation of orientation A, was suggested based on the identification of only two cross-linking products connecting the N-terminus and K23 of melittin with residues 77–84 and E11 of CaM, respectively (55). Our present paper confirms and extends these findings by detecting cross-linking products between the N-terminus of melittin and residues 78–86, as well as between K21 of melittin and residues 1–13 of CaM (Table 2). Another report

on the CaM–melittin complex had suggested that 80% of the melittin population binds to CaM in orientation A; however, no data were shown supporting this statement (64).

To generate low-resolution models of the CaM–melittin complex based on the experimental cross-linking data, we made use of conjoined rigid body/torsion angle simulated annealing (49–51), a technique that has proven to be very powerful for solving structures of macromolecular complexes from NMR data (65–67). The limitation in generating structures from our cross-linking data is due to the presence of only three hinge residues, K21, K23, and the N-terminus, of melittin. K7, which could potentially form cross-linking products with CaM, was never found to be cross-linked in our experiments. The most likely reason is that cross-linking products involving K7 are not amenable to detection under the mass spectrometric conditions employed, because the proteases used are unable to cleave after K7 under conditions where it is modified.

From a qualitative examination of the cross-linking data (Table 2), it became immediately apparent that the data were not compatible with a unique orientation of melittin in the CaM–melittin complex. Structures were generated from the cross-linking data ascribing melittin to two opposite orienta-



tions within the CaM-peptide complex (Figure 8), one parallel (orientation A) and the other antiparallel (orientation B). The majority of cross-links originate from orientation A (Table 2). This could have been anticipated from the basic amino acid cluster (K21–K24) located close to the C terminus of melittin. However, there are a number of cross-linking-derived distance restraints that are not satisfied in the structures within the distance limits given by spacer lengths of the respective cross-linkers. One possible explanation is that, even within the context of two opposing orientations of melittin, multiple modes of binding exist and that cross-linking may pick up minor conformational states because of its fast reaction kinetics. We are aware that it is problematic to quantify the distribution of orientations of a protein/peptide in the complex from the number of cross-links obtained. Certain cross-linking products might be created in the course of the reaction but are not amenable to mass spectrometric detection. During the cross-linking reaction, a variety of different complexes are created but only those that can be picked up during mass spectrometric detection are considered for building a structural model.

The finding that melittin can bind to CaM in two opposite directions suggests that the presence of the basic amino acid cluster in the target peptide is not the only crucial determinant of peptide orientation within the CaM-melittin complex. Our studies demonstrate that CaM is able to recognize its target peptides in multiple modes, which is also evidenced by the recently published structures of complexes between CaM and the larger target proteins mentioned above (38, 39). It seems that the two hydrophobic anchor residues in melittin, L6 and W19, can be accommodated in each of the hydrophobic patches and that the bulky amino acid W19 does not necessarily have to be accommodated in the C-terminal domain of CaM. Apparently, the basic amino acid cluster in melittin, comprising K21–R24, can interact with negatively charged amino acids (glutamic or aspartic acids) in either CO-2 or CO-1 of CaM. From the number of cross-linking products obtained for each orientation, it is likely that melittin binds predominantly in the parallel orientation A (Figures 5 and 8).

Using chemical cross-linking in combination with high-resolution mass spectrometry and recently developed computational methods on the CaM-melittin model system, we demonstrate that this technique is likely to have wide-ranging implications for structural studies on protein-protein interactions.

## REFERENCES

- Crivici, A., and Ikura, M. (1995) Molecular and structural basis of target recognition by calmodulin, *Annu. Rev. Biophys. Biomol. Struct.* 24, 85–116.
- Carafoli, E. (2002) Calcium signaling: A tale for all seasons, *Proc. Natl. Acad. Sci. U.S.A.* 99, 1115–1122.
- Babu, Y. S., Sack, J. S., Greenhough, T. J., Bugg, C. E., Means, A. R., and Cook, W. J. (1985) 3-Dimensional structure of calmodulin, *Nature* 315, 37–40.
- Babu, Y. S., Bugg, C. E., and Cook, W. J. (1988) Structure of calmodulin refined at 2.2 Å resolution, *J. Mol. Biol.* 204, 191–204.
- Persechini, A., and Kretsinger, R. H. (1988) The central helix of calmodulin functions as a flexible tether, *J. Biol. Chem.* 263, 12175–12178.
- Barbato, G., Ikura, M., Kay, L. E., Pastor, R. W., and Bax, A. (1992) Backbone dynamics of calmodulin studied by <sup>15</sup>N relaxation using inverse detected two-dimensional NMR spectroscopy: The central helix is flexible, *Biochemistry* 31, 5269–5278.
- Comte, M., Maulet, Y., and Cox, A. J. (1983) Ca<sup>2+</sup>-dependent high-affinity complex formation between calmodulin and melittin, *Biochem. J.* 209, 269–272.
- Maulet, Y., and Cox, J. A. (1983) Structural changes in melittin and calmodulin upon complex formation and their modulation by calcium, *Biochemistry* 22, 5680–5686.
- Eisenberg, D., Terwilliger, T. C., and Tsui, F. (1980) Structural studies of bee melittin, *Biophys. J.* 32, 252–254.
- Terwilliger, T. C., and Eisenberg, D. (1982) The structure of melittin. II. Interpretation of the structure, *J. Biol. Chem.* 257, 6016–6022.
- Kataoka, M., Head, J. F., Seaton, B. A., and Engelmann, D. M. (1989) Melittin binding causes a large calcium-dependent conformational change in calmodulin, *Proc. Natl. Acad. Sci. U.S.A.* 86, 6944–6948.
- Malencik, D. A., and Anderson, S. R. (1988) Association of melittin with the isolated myosin light chains, *Biochemistry* 27, 1941–1949.
- He, Z., Dunker, A. K., Wesson, C. R., and Trumble, W. R. (1993) Ca(2+)-induced folding and aggregation of skeletal muscle sarcoplasmic reticulum calsequestrin. The involvement of trifluoperazine-binding site, *J. Biol. Chem.* 268, 24635–24641.
- Raynor, R. J., Zheng, B., and Kuo, J. F. (1991) Membrane interactions of amphiphilic polypeptides mastoparan, melittin, polymyxin B, and cardiotoxin. Differential inhibition of protein kinase C, Ca<sup>2+</sup>/calmodulin-dependent protein kinase II and synaptosomal membrane Na, K-ATPase, and Na<sup>+</sup> pump and differentiation of HL60 cells, *J. Biol. Chem.* 266, 2753–2758.
- Fletcher, J. F., and Jiang, M. S. (1993) Possible mechanisms of action of cobra snake venom cardiotoxins and bee venom melittin, *Toxicon* 31, 669–695.
- Cuppoletti, J., Blumenthal, K. E., and Malinowska, D. H. (1989) Melittin inhibition of the gastric (H<sup>+</sup>/K<sup>+</sup>) ATPase and photoaffinity-labeling with [I-125] azidosalicyl melittin, *Arch. Biochem. Biophys.* 275, 263–270.
- Cuppoletti, J. (1990) [I-125] azidosalicyl melittin binding domains—evidence for a polypeptide receptor on the gastric (H<sup>+</sup>/K<sup>+</sup>) ATPase, *Arch. Biochem. Biophys.* 278, 409–415.
- James, P., Maeda, M., Fischer, R., Verma, A. K., Krebs, J., Penniston, J. T., and Carafoli, E. (1988) Identification and primary structure of a calmodulin binding domain of the Ca<sup>2+</sup> pump of human erythrocytes, *J. Biol. Chem.* 263, 2905–2910.
- Hait, W. N., Cadman, E., Benz, C., Cole, J., Weiss, B. (1983) Inhibition of growth of L1210-leukemic cells by inhibitors of cyclic-nucleotide phosphodiesterase (PDE) and calmodulin (CAL), *Proc. Am. Assoc. Cancer Res.* 24, 50.
- Lee, G. L., and Hait, W. N. (1985) Inhibition of growth of C6 astrocytoma cells by inhibitors of calmodulin, *Life Sci.* 36, 347–354.
- Gest, J. E., and Salomon, Y. (1987) Inhibition by melittin and fluphenazine of melanotropin receptor function and adenylate cyclase in M2R melanoma cell membranes, *Endocrinology* 121, 1766–1772.
- Orsolic, N., Sver, L., Verstovsek, S., Terzic, S., and Basic, I. (2003) Inhibition of mammary carcinoma cell proliferation in vitro and tumor growth in vivo by bee venom, *Toxicon* 41, 861–870.
- Cox, J. A., Comte, M., Fitton, J. E., and DeGrado, W. F. (1985) The interaction of calmodulin with amphiphilic peptides, *J. Biol. Chem.* 260, 2527–2534.
- Clore, G. M., Bax, A., Ikura, M., and Gronenborn, A. M. (1993) Structure of calmodulin-target peptide complexes, *Curr. Opin. Struct. Biol.* 3, 838–845.
- O'Neil, K. T., and DeGrado, W. F. (1990) How calmodulin binds its targets: sequence independent recognition of amphiphilic α-helices, *Trends Biochem. Sci.* 15, 59–64.
- Ikura, M., Clore, G. M., Gronenborn, A. M., Zhu, G., Klee, C. B., and Bax, A. (1992) Solution structure of a calmodulin-target peptide complex by multidimensional NMR, *Science* 256, 632–638.
- Yuan, T., Weljie, A. M., and Vogel, H. J. (1998) Tryptophan fluorescence quenching by methionine and selenomethionine residues of calmodulin: Orientation of peptide and protein binding, *Biochemistry* 37, 3187–3195.
- Vetter, S. W., and Leclerc, E. (2003) Novel aspects of calmodulin target recognition and activation, *Eur. J. Biochem.* 270, 404–414.



29. Itakura, M., and Iio, T. (1992) Static and kinetic studies of calmodulin and melittin complex, *J. Biochem.* **112**, 183–191.
30. Osawa, M., Tokumitsu, H., Swindells, M. B., Kurihara, H., Orita, M., Shibamura, T., Furuya, T., and Ikura, M. (1999) A novel target recognition revealed by calmodulin in complex with Ca<sup>2+</sup>-calmodulin-dependent kinase kinase, *Nat. Struct. Biol.* **6**, 819–824.
31. Meador, W. E., Means, A. R., and Quirocho, F. A. (1992) Target enzyme recognition by calmodulin: 2.4 Å structure of a calmodulin-peptide complex, *Science* **257**, 1251–1255.
32. Meador, W. E., Means, A. R., and Quirocho, F. A. (1993) Modulation of calmodulin plasticity in molecular recognition on the basis of X-ray structures, *Science* **262**, 1718–1721.
33. Elshorst, B., Hennig, M., Försterling, H., Diener, A., Maurer, M., Schulte, P., Schwabe, H., Griesinger, C., Krebs, J., Schmid, H., Vorherr, T., and Carafoli, E. (1999) NMR solution structure of a complex of calmodulin with a binding peptide of the Ca<sup>2+</sup> pump, *Biochemistry* **38**, 12320–12332.
34. Kurokawa, H., Osawa, M., Kurihara, H., Katayama, N., Tokumitsu, H., Swindells, M. B., Kainosho, M., and Ikura, M. (2001) Target-induced conformational adaptation of calmodulin revealed by the crystal structure of a complex with nematode Ca<sup>2+</sup>/calmodulin-dependent kinase kinase peptide, *J. Mol. Biol.* **312**, 59–68.
35. Clapperton, J. A., Martin, S. R., Smerdon, S. J., Gamblin, S. J., and Bayley, P. M. (2002) Structure of the complex of calmodulin with the target sequence of calmodulin-dependent protein kinase I: Studies of the kinase activation mechanism, *Biochemistry* **41**, 14669–14679.
36. Yamauchi, E., Nakatsu, T., Matsubara, M., Kato, H., and Taniguchi, H. (2003) Crystal structure of a MARCKS peptide containing the calmodulin-binding domain in complex with Ca<sup>2+</sup>-calmodulin, *Nat. Struct. Biol.* **10**, 226–231.
37. Yap, K. L., Yuan, T., Mal, T. K., Vogel, H. J., and Ikura, M. (2003) Structural basis for simultaneous binding of two carboxy-terminal peptides of plant glutamate decarboxylase to calmodulin, *J. Mol. Biol.* **328**, 193–204.
38. Schumacher, M. A., Rivard, A. F., Bächinger, H. P., and Adelman, J. P. (2001) Structure of the gating domain of a Ca<sup>2+</sup>-activated K<sup>+</sup> channel complexed with calmodulin, *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) Structural basis for the activation of anthrax adenyl cyclase exotoxin by calmodulin, *Nature* **415**, 396–402.
40. Dihazi, G. H., and Sinz, A. (2003) Mapping low-resolution three-dimensional protein structures using chemical cross-linking and Fourier transform ion-cyclotron resonance mass spectrometry, *Rapid Commun. Mass Spectrom.* **17**, 2005–2014.
41. Young, M. M., Tang, N., Hempel, J. C., Oshiro, C. M., Taylor, E. W., Kuntz, I. D., Gibson, B. W., and Dollinger, G. (2000) High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5802–5806.
42. Bennett, K. L., Kussmann, M., Björk, P., Godzwon, M., Mikkelsen, M., Sørensen, P., and Roepstorff, P. (2000) Chemical cross-linking with thiol-cleavable reagents combined with differential mass spectrometric peptide mapping—A novel approach to assess intermolecular protein contacts, *Protein Sci.* **9**, 1503–1518.
43. Rappsilber, J., Siniossoglou, S., Hurt, E. C., and Mann, M. (2000) A generic strategy to analyze the spatial organization of multi-protein complexes by cross-linking and mass spectrometry, *Anal. Chem.* **72**, 267–275.
44. McLafferty, F. W., Fridriksson, E. K., Horn, D. M., Lewis, M. A., and Zubarev, R. A. (1999) Biomolecule mass spectrometry, *Science* **284**, 1289–1290.
45. Hermanson, G. T. (1996) *Bioconjugate Techniques*, 1st ed., Academic Press, San Diego, CA.
46. Pearson, K. M., Pannell, L. K., and Fales, H. M. (2002) Intramolecular cross-linking experiments on cytochrome *c* and ribonuclease A using an isotope multiplet method, *Rapid Commun. Mass Spectrom.* **16**, 149–159.
47. Taverner, T., Hall, N. E., O'Hair, R. A. J., and Simpson, R. J. (2002) Characterization of an antagonist interleukin-6 dimer by stable isotope labeling, cross-linking, and mass spectrometry, *J. Biol. Chem.* **277**, 46487–46492.
48. Sinz, A., and Wang, K. (2001) Mapping protein interfaces with a fluorogenic cross-linker and mass spectrometry: Application to nebulin-calmodulin complexes, *Biochemistry* **40**, 7903–7913.
49. Clore, G. M. (2000) Accurate and rapid docking of protein-protein complexes on the basis of intermolecular nuclear Overhauser enhancement data and dipolar couplings by rigid body minimization, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9021–9025.
50. Schwieters, C. D., and Clore, G. M. (2001) Internal coordinates for molecular dynamics and minimization in structure determination and refinement, *J. Magn. Reson.* **152**, 288–302.
51. Clore, G. M., and Bewley, C. A. (2002) Using conjoined rigid body/torsion angle simulated annealing to determine relative orientation of covalently linked protein domains from dipolar couplings, *J. Magn. Reson.* **154**, 329–335.
52. Comisarov, M. B., and Marshall, A. G. (1974) Fourier transform ion cyclotron resonance spectroscopy, *Chem. Phys. Lett.* **25**, 282–283.
53. Marshall, A. G. (2000) Milestones in Fourier transform ion cyclotron resonance spectrometry technique development, *Int. J. Mass Spectrom.* **200**, 331–356.
54. Kruppa, G. H., Schoeniger, J., and Young, M. M. (2003) A top down approach to protein structural studies using chemical cross-linking and Fourier transform mass spectrometry, *Rapid Commun. Mass Spectrom.* **17**, 155–162.
55. Scaloni, A., Miraglia, N., Orrù, S., Amodeo, P., Motta, A., Marino, G., and Pucci, P. (1998) Topology of the calmodulin-melittin complex, *J. Mol. Biol.* **277**, 945–958.
56. Laemmli, U. K. (1970) Cleavage of structural proteins during assembly of head of bacteriophage-T4, *Nature* **227**, 680–685.
57. Jensen, O. N., Shevchenko, A., and Mann, M. (1997) in *Protein Structure, A Practical Approach* (Creighton, T. E., Ed.) p 48, Oxford University Press, Oxford, U.K.
58. Bruker Daltonics (1996) *BioAPEX User's Manual, Version 1.1*, Billerica, MA.
59. Schwieters, C. D., Kuszewski, J., Tjandra, N., and Clore, G. M. (2003) The XPLOR–NIH molecular structure determination package, *J. Magn. Reson.* **160**, 65–73.
60. Schwieters, C. D., and Clore, G. M. (2001) The VMD-XPLOR visualization package for NMR structure refinement, *J. Magn. Reson.* **149**, 239–244.
61. Nilges, M., Gronenborn, A. M., Brünger, A. T., and Clore, G. M. (1988) Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints. Application to crambin, potato carboxypeptidase inhibitor and barley serine proteinase inhibitor 2, *Protein Eng.* **2**, 27–38.
62. Clore, G. M., and Kuszewski, J. (2002)  $\chi^1$  rotamer populations and angles of mobile surface side chains are accurately predicted by a torsion angle database potential of mean force, *J. Am. Chem. Soc.* **124**, 2866–2867.
63. Watterson, D. M., Sharief, F., and Vanaman, T. C. (1980) The complete amino acid sequence of the Ca<sup>2+</sup>-dependent modulator protein (calmodulin) of bovine brain, *J. Biol. Chem.* **255**, 962–975.
64. Brokx, R. D., Lopez, M. M., Vogel, H. J., and Makhatadze, G. I. (2001) Energetics of peptide binding by calmodulin reveals different modes of binding, *J. Biol. Chem.* **276**, 14083–14091.
65. Wang, G., Louis, J. M., Sondej, M., Seok, Y. J., Peterkofsky, A., and Clore, G. M. (2000) Solution structure of the phosphoryl transfer complex between the signal transducing proteins HPr and IIA<sup>Glucose</sup> and HPr of the *Escherichia coli* phosphoenolpyruvate: sugar phosphotransferase system, *EMBO J.* **19**, 5635–5649.
66. Cornilescu, G., Lee, B. R., Cornilescu, C. C., Wang, G., Peterkofsky, A., and Clore, G. M. (2002) Solution structure of the phosphoryl transfer complex between the cytoplasmic A domain of the mannitol transporter II<sup>Glucose</sup> and HPr of the *Escherichia coli* phosphotransferase system, *J. Biol. Chem.* **277**, 42289–42298.
67. Cai, M., Williams, D. C., Wang, G., Lee, B. R., Peterkofsky, A., and Clore, G. M. (2003) Solution structure of the phosphoryl transfer complex between the signal-transducing protein IIA<sup>Glucose</sup> and the cytoplasmic domain of the glucose transporter IICB<sup>Glucose</sup> of the *Escherichia coli* glucose phosphotransferase system, *J. Biol. Chem.* **278**, 25191–25206.