Anal. Chem. **2005,** *77,* 495−503

Chemical Cross-Linking and High-Performance Fourier Transform Ion Cyclotron Resonance Mass Spectrometry for Protein Interaction Analysis: Application to a Calmodulin/Target Peptide Complex

Stefan Kalkhof,[†] Christian Ihling,[†] Karl Mechtler,[‡] and Andrea Sinz*,[†]

Biotechnological-Biomedical Center, Faculty of Chemistry and Mineralogy, University of Leipzig, D-04103 Leipzig, Germany, and Institute of Molecular Pathology, Dr. Bohrgasse 7, A-1030 Vienna, Austria

Chemical cross-linking has proved successful in combination with mass spectrometry as a tool for low-resolution structure determination of proteins. The integration of chemical cross-linking with Fourier transform ion cyclotron resonance (FTICR) mass spectrometry to determine protein interfaces was tested on the calcium-dependent complex between calmodulin (CaM) and a 26-amino acid peptide derived from the skeletal muscle myosin light chain kinase (M13). Different amine-reactive, homobifunctional cross-linkers and a "zero-length" cross-linker were employed. The covalently attached complexes were separated from nonreacted proteins by one-dimensional gel electrophoresis, and the bands of interest were excised and in-gel digested with trypsin. Digestion of the cross-linked complexes resulted in complicated peptide mixtures, which were analyzed by nano-HPLC/nano-ESI-FTICR mass spectrometry. The distance constraints obtained by chemical cross-linking were in agreement with the published NMR structure of the CaM/M13 complex, pointing to residues Lys-18 and Lys-19 of M13 being cross-linked with the central α-helix of CaM. Thus, the integrated approach described herein has proven to be an efficient tool for mapping the topology of the CaM/M13 complex. As such it is applicable as a general strategy for the investigation of the spatial organization of protein complexes and complements existing techniques, such as X-ray crystallography and NMR spectroscopy.

With recent thrusts in genome sequencing projects, the number of identified proteins has dramatically increased during the past few years. The physiological function of many of the newly discovered proteins, however, remains unclear. Closely related to studying the function of a protein is the analysis of its binding partners. After identification of a binding partner, the elucidation of the three-dimensional structure of a protein complex can yield important information about the function of a protein.

High-resolution structural analysis of proteins is currently accomplished by NMR spectroscopy, X-ray crystallography, and—to a certain extent—electron microscopy. These techniques, however, are time- and material-consuming methods, and moreover, they are not applicable to all proteins or protein complexes. A promising strategy with the potential to obtain low-resolution structural information on minute amounts of protein complexes within a few days is based on a combination of chemical cross-linking and mass spectrometry.[1] Using chemical cross-linking—a well-established technique in protein chemistry, where covalent bonds between different molecules (intermolecular) or between parts of a molecule (intramolecular) are formed—it is feasible to determine distance constraints between functional groups. The created cross-linking products are subsequently identified by mass spectrometric techniques.[2,3] The mass of the protein complex to be investigated is theoretically unlimited, since it is the proteolytic peptides that are analyzed. Analysis is generally fast and requires only minimal (down to femtomole or even attomole) amounts of analyte.[4] Furthermore, the broad range of specificities available for cross-linking reagents toward certain functional groups, such as primary amines, sulfhydryls, or carboxylic acids, and the wide range of distances different cross-linking reagents can bridge, offer various options for experimental design.[1] However, despite the straightforwardness of the cross-linking approach, identification of the cross-linking products can be hampered by the complexity of the created reaction mixtures. Several strategies have been employed to facilitate identification of the cross-linking products, such as using isotope-labeled cross-linkers or proteins, fluorogenic cross-linkers, or cross-linkers creating characteristic fragment ions during mass spectrometric detection.[1] An alternative is presented by using high-performance Fourier transform ion cyclotron resonance (FTICR) mass spectrometry, which offers unique analytical possibilities due to its excellent mass accuracy and high resolution.[5] Thus, a cross-linking product can be unambiguously

* Corresponding author. Phone: +49-341-9736078, Fax: +49-341-9736115, E-mail: sinz@chemie.uni-leipzig.de.

[†] University of Leipzig.

[‡] Institute of Molecular Pathology.

(1) Sinz, A. *J. Mass Spectrom.* **2003**, *38*, 1225−1237.
(2) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64−71.
(3) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299−2301.
(4) McLafferty, F. W.; Fridriksson, E. K.; Horn, D. M.; Lewis, M. A.; Zubarev, R. A. *Science* **1999**, *284*, 1289−1290.
(5) Comisarov M. B.; Marshall A. G. *Chem. Phys. Lett.* **1974**, *25*, 282−283.

10.1021/ac0487294 CCC: $30.25
Published on Web 12/07/2004

identified solely based on determining its exact mass. We recently presented the elucidation of the interacting regions between calmodulin (CaM) and its target peptide melittin by integrating chemical cross-linking and FTICR mass spectrometry.[6] In this study, we present the combination of chemical cross-linking and high-performance FTICR mass spectrometry as a generally applicable, rapid approach for identifying amino acids in protein complexes. We apply this technique to the complex between CaM and a target peptide.

CaM 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.[7] CaM serves as a calcium-dependent regulator in many metabolic pathways.[8] Upon calcium binding, CaM adopts a dumbbell structure[9,10] consisting of two lobes connected by a flexible central helix.[11,12] CaM is known to bind with high affinity to various target proteins and peptides, with dissociation constants ranging between 10^{-7} and 10^{-11} M.[7]

The myosin light chain kinase of the smooth muscle (sm-MLCK) is currently the functionally best-characterized representative of a class of CaM-dependent enzymes, which catalyze the phosphorylation of a serine residue at the N-terminus of the myosin light chain.[13] Another structurally similar member of this class of enzymes is the tissue-specific myosin light chain kinase of the skeletal muscle (skMLCK). The skMLCK consists of ~600 amino acids depending on the species (for example, the human isoform comprises 595, the rabbit isoform 607 amino acids). Binding studies used chymotryptic fragments of the rabbit skMLCK and synthetic peptides identified a C-terminal peptide of the skMLCK as the CaM-binding region.[14] This 26-amino acid peptide has been designated as M13, comprising amino acids 577−602 of the rabbit skMLCK (or amino acids 565−590 of the human skMLCK, respectively). Accordingly, M13 represents the complete CaM-interacting sequence of skMLCK and thus can be employed as a substitute for the total kinase in order to analyze the structures of CaM and the interacting MLCK sequence within the complex.[15] The three-dimensional structure of the CaM/M13 complex was solved several years ago by multidimensional NMR spectroscopy,[15] making this complex an ideal system to test our method.

In the present study, we aimed to demonstrate that a combination of chemical cross-linking and FTICR mass spectrometry can rapidly yield structural information on protein complexes. We show that the resulting data are in close agreement with data obtained from high-resolution methods, such as NMR spectroscopy.

(6) Schulz, D. M.; Ihling, C.; Clore, G. M.; Sinz, A. *Biochemistry* **2004**, *43*, 4703−4715.
(7) Crivici, A.; Ikura, M. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 85−116.
(8) Carafoli, E. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1115−1122.
(9) Babu, Y. S.; Sack, J. S.; Greenhough, T. J.; Bugg, C. E.; Means, A. R.; Cook, W. J. *Nature* **1985**, *315*, 37−40.
(10) Babu, Y. S.; Bugg, C. E.; Cook, W. J. *J. Mol. Biol.* **1988**, *204*, 191−204.
(11) Persechini, A.; Kretsinger, R. H. *J. Biol. Chem.* **1988**, *263*, 12175−12178.
(12) Barbato, G.; Ikura, M.; Kay, L. E.; Pastor, R. W.; Bax, A. *Biochemistry* **1992**, *31*, 5269−5278.
(13) Gallagher, P. J.; Herring, B. P.; Stull, J. T. *J. Muscle Res. Cell Motil.* **1997**, *18*, 1−16.
(14) Herring, B. P.; Stull, J. T.; Gallagher, P. J. *J. Biol. Chem.* **1990**, *265*, 1724−1730.
(15) Ikura, M.; Clore, G. M.; Gronenborn, A. M.; Zhu, G.; Klee, C. B.; Bax A. *Science* **1992**, *256*, 632−638.

EXPERIMENTAL SECTION

Materials. Bovine brain CaM was obtained from Calbiochem (Schwalbach am Taunus, Germany) and used without further purification. The skMLCK peptide (M13) was a generous gift from Dr. Ad Bax, National Institutes of Health. The purity was checked by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry and SDS−PAGE. The cross-linking reagents 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC), *N*-hydroxysulfosuccinimide (sulfo-NHS), bis-(sulfosuccinimidyl)suberate (BS^3), and disulfosuccinimidyl tartrate (sulfo-EGS) were purchased from Pierce (Rockford, IL). *N*-Hydroxysuccinimide, dicyclohexylcarbodiimide, 1,4-dioxane, and adipic acid were purchased from Sigma-Aldrich (Vienna, Austria). Trypsin (sequencing grade) was obtained from Roche Diagnostics (Mannheim, Germany). luteinizing hormone releasing hormone (LHRH), sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid), and proteins and peptides for MALDI-TOFMS calibration were purchased from Sigma (Taufkirchen, Germany). Buffer reagents were obtained from Sigma at the highest available purity. Nano-HPLC solvents were spectroscopic grade (Uvasol, VWR, Darmstadt, Germany). Water was purified with a Direct-Q5 water purification system (Millipore, Eschborn, Germany).

Synthesis of Disuccinimidyl Adipate (DSA). For synthesis of DSA-d_0/d_8, adipic acid-d_0/d_8 (0.727 mmol, 106 mg) and *N*-hydroxysuccinimide (1.46 mmol, 168 mg) were dissolved in 10 mL of anhydrous 1,4-dioxane. The flask was continuously flushed with dry argon. Dicyclohexylcarbodiimide (1.5 mL of a 1 M solution in CH_2Cl_2, 1.5 mmol) was added via a syringe under continuous stirring, and the reaction was allowed to proceed at room temperature for 20 h. The precipitated dicyclohexylurea was removed by vacuum filtration. Dioxane was removed by rotary evaporation, and the products were dried in high vacuum. The identity of the products (mp 165−167 °C) was confirmed by ESI mass spectrometry.

Cross-Linking Reactions. For cross-linking reactions with EDC/sulfo-NHS, an aqueous CaM stock solution (1 mg/mL, corresponding to 59.5 μM) was diluted with 20 mM MES buffer (pH 6.8) to give a CaM concentration of 10.6 μM (volume 946 μL). A 10-μL aliquot of a 100 mM $CaCl_2$ solution was added, and the mixture was incubated for 10 min at room temperature in order to ensure that CaM was fully loaded with calcium. A 34-μL aliquot of an aqueous M13 stock solution (1 mg/mL, corresponding to 337.6 μM) was added to give a final M13 concentration of 10 μM. After incubation for 10 min, 10 μL of a freshly prepared aqueous cross-linker solution, containing either 0.5, 1.0, or 2.0 M EDC in addition to 0.5 M sulfo-NHS, was added. Thus, molar excesses of 500, 1000, and 2000 of EDC over the protein/peptide concentration were obtained. For a control, 10 μL of water was added, instead of the cross-linker, to one CaM/M13 solution. The reaction mixtures were incubated at room temperature and 200-μL aliquots were taken after 5, 15, 30, 60, and 120 min. The reactions were quenched by adding 20 μL of an 220 mM aqueous DTT solution to each aliquot (final concentration 20 mM). Before SDS−PAGE and MALDI-TOFMS, the solutions were desalted using Microcon-YM-3 filters (Millipore, Eschborn, Germany).

For cross-linking experiments with the homobifunctional cross-linking reagents sulfo-EGS, BS^3 or DSA-d_0/d_8, an aqueous CaM stock solution (1 mg/mL) was diluted with 20 mM Hepes buffer

(pH 7.4) containing 1 mM $CaCl_2$ to give solutions containing CaM at a concentration of 10.44 μM (volume 957 μL). After an incubation time of 10 min, 33 μL of an M13 solution (final concentration 10 μM) was added. The cross-linking reactions were started by adding 10 μL of solutions containing either sulfo-EGS, BS^3 (each concentrations of 0.1, 0.05, or 0.01 M in DMSO), or DSA-d_0/d_8 (at concentrations of 0.1, 0.05, or 0.02 M in DMSO), thus yielding relative molar excesses of cross-linker over protein/peptide concentration of 100, 50, 10 (for sulfo-EGS and BS^3), and 20 (for DSA-d_0/d_8). One solution was prepared without adding cross-linker solution, but 10 μL of DMSO was added instead. The reaction mixtures were incubated at room temperature, and 200-μL aliquots were taken after 5, 15, 30, 60, and 120 min. The reactions were terminated by adding 20 μL of a 220 mM NH_4HCO_3 solution (final concentration 20 μM) to each aliquot. Before SDS−PAGE and MALDI-TOFMS, the solutions were desalted using Microcon-YM-3 filters (Millipore).

SDS−PAGE and Enzymatic Proteolysis. Following separation of the reaction mixtures by one-dimensional SDS−PAGE (15%, Coomassie staining[16]), the bands of interest were excised and in-gel digested as described previously.[17] Trypsin (50 ng/μL) was used as digestion enzyme for all samples. Depending on the volume of the gel pieces, between 2 and 5 μL of enzyme solution was added, and the digests were incubated at 37 °C for 16 h.

MALDI-TOF Mass Spectrometry. MALDI-TOF mass spectrometry 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 the linear mode (detection range m/z 16 000−24 000) as well as in the reflector mode (detection range m/z 2500−4000) using sinapinic acid and α-cyano-4-hydroxycinnamic acid as matrixes. A saturated matrix solution was prepared in 30% (v/v) acetonitrile, 69.9% (v/v) water, and 0.1% (v/v) TFA. Samples were prepared using the dried droplet method by spotting 0.5 μL of matrix solution and 0.5 μL of sample solution onto the target. Spectra from 100 laser shots were accumulated to one spectrum when operating the instrument in the reflector mode, and 300 laser shots were added when spectra were acquired in the linear mode. In the linear mode, the instrument was calibrated using cytochrome c ($[M + H]^+_{average}$ at m/z 12 361) and myoglobin ($[M + H]^+_{average}$ at m/z 16 952). For calibration in the reflector mode, signals of angiotensin I ($[M + H]^+_{mono}$ at m/z 1296.69), substance P ($[M + H]^+_{mono}$ at m/z 1347.74), and somatostatin (reduced form) ($[M + H]^+_{mono}$ at m/z 1637.72) were employed. Data acquisition and data processing were performed using the Voyager software version 5.1 and the Data Explorer software version 4.0 (Applied Biosystems, Foster City, CA).

Nano-HPLC/Nano-ESI FTICR Mass Spectrometry. Peptide mixtures from the enzymatic digests were separated by nano-HPLC. Nano-HPLC was carried out on an Ultimate Nano-LC system (Dionex, Idstein, Germany) equipped with a Switchos II column switching module and a Famos microautosampler with a 20-μL sample loop. Samples were injected by the autosampler and concentrated on a trapping column (PepMap, C18, 300 μm × 5 mm, 3 μm, 100 Å, Dionex) with water containing 0.1% formic acid at flow rates of 20 μL/min. After 3 min for desalting, the peptides were eluted onto the separation column (PepMap, C18, 75 μm × 150 mm, 3 μm, 100 Å, Dionex), which had been equilibrated with 95% A (solvent A: water containing 0.1% formic acid). Peptides were separated using the following gradient: 0−30 min, 5−50% B; 30−31 min, 50−95% B; 31−45 min, 95% B (solvent B: acetonitrile containing 0.1% formic acid) at flow rates of 200 nL/min and detected based on their UV absorptions at 214 and 280 nm.

The nano-HPLC system was coupled on-line to an APEX II FTICR mass spectrometer equipped with a 7-T superconducting magnet (Bruker Daltonics, Billerica, MA) and a nanoelectrospray ionization source (Agilent Technologies, Waldbronn, Germany). For nano-ESIMS, distal coated fused-silica PicoTips (tip i.d. 8 μm, New Objective, Woburn, MA) were applied. The capillary voltage was set to −1400 V. Mass spectral data were acquired in the broadband mode over an m/z range of 400−2000 with 256K data points, 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 7 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), LHRH ($[M + H]^{2+}$ m/z 592.2358), y7 (m/z 749.3941), and y8 (m/z 935.4734), LHRH ($[M + H]^+$ m/z 1183.5643)). Data acquisition was performed using the XMASS software (version 6.1.2); data processing was performed with the XMASS software (version 5.0.10) (Bruker Daltonics). Prior to Fourier transformation, the time-domain signals were doubly zero filled to enhance the digitization of the spectra and to improve their visualization followed by apodization with a sine function. A total of 400 single spectra were projected into one final mass spectrum using the Projection tool in the XMASS software.[18]

Identification of Cross-Linking Products. Cross-linking products were identified using the GPMAW (General Protein Mass Analysis for Windows) software, version 6.01 (Lighthouse Data, Odense, Denmark) (available at: http://welcome.to/gpmaw), the ExPASy Proteomics tools in the Swiss-Prot Database (available at: http://www.expasy.org), and the ASAP (Automatic Spectrum Assignment Program) software (available at: http://roswell.ca.sandia.gov/~mmyoung /asap.html). Proteolytic cleavages at modified amino acids, such as the trimethylated K115 in CaM as well as amino acids modified by cross-linking reagents were excluded. The N-terminus of CaM was excluded from possible cross-linking since it is acetylated.

Determination of Distances between Atoms in the CaM/M13 Complex. The NMR structure of the CaM/M13 complex is deposited in the RCSB Protein Data Bank (http://www.rcsb.org/pdb/) under the entry 2BBM.[15] Based on the atom coordinates of this structure, the distances between the nitrogen atoms of the ϵ-amine groups of lysine residues and the N-terminus of M13, respectively, were determined using the VMD-XPLOR visualization package[19] (available at: http://vmd-xplor.cit.nih.gov). The VMD-XPLOR software also calculated distances between

(16) Laemmli, U. K. *Nature* **1970**, *227*, 680−685.
(17) Jensen, O. N.; Shevchenko, A.; Mann, M. In *Protein Structure, A Practical Approach*; Creighton, T, E., Ed.; Oxford University Press: Oxford, U.K., 1997; pp 48.
(18) Bruker Daltonics *BioAPEX User's Manual* v 1.1; Billerica, MA, 1996.
(19) Schwieters, C. D.; Clore, G. M. *J. Magn. Reson.* **2001**, *149*, 239−244.

(A)

^{1}ADQLT ^{6}EEQIA ^{11}EFKEA ^{16}FSLFD ^{21}KDGDG ^{26}TITTK
^{31}ELGTV ^{36}MRSLG ^{41}QNPTE ^{46}AELQD ^{51}MINEV ^{56}DADGN
^{61}GTIDF ^{66}PEFLT ^{71}MMARK ^{76}MKDTD ^{81}SEEEI ^{86}REAFR
^{91}VFDKD ^{96}GNGTI ^{101}SAAEL ^{106}RHVMT ^{111}NLGEK ^{116}LTDEE
^{121}VDEMI ^{126}REAEI ^{131}DGDGQ ^{136}VNTEE ^{141}FVQMM ^{146}TAK

(B)

^{1}KRRWK ^{6}KNFIA ^{11}VSAAN ^{16}RFKKI ^{21}SSSGA ^{26}L

Figure 1. Amino acid sequences of (A) CaM and (B) M13. CaM was found to be N-acetylated (Ac−Ala) and to contain a trimethylated lysine in position 115 (the modified amino acids Ala-1 and Lys-115 are underlined).

C-atoms of carboxylic acids of glutamic or aspartic acid residues and the N-atoms of amine groups of lysine residues or the N-terminus of M13.

RESULTS AND DISCUSSION

Characterization of CaM and M13. The investigation of the interaction sites between CaM and M13 in the CaM/M13 complex depends on a detailed description of their respective primary structures. ESI-FTICRMS of CaM yielded the most abundant mass at m/z 16 790.921 (most abundant mass from simulation at m/z 16 790.884, $\Delta m = 0.037$ u, 2.1 ppm). Peptide mass fingerprinting of CaM using trypsin yielded full sequence coverage. CaM was confirmed to be N-terminally acetylated and to contain a trimethylated lysine at position 115.[6,20] The monoisotopic mass of M13 was determined by ESI-FTICR mass spectrometry as 2962.702 u (calculated 2962.717 u, $\Delta m = 0.015$ u, 5.0 ppm). The fragments created by tryptic cleavage did not cover the complete sequence of M13; however, when additional data from the cross-linked CaM/M13 complex were also considered, a total of 100% sequence coverage was achieved. The amino acid sequences of CaM and M13 are shown in Figure 1.

Cross-Linking Reactions. Cross-linking experiments between CaM and M13 were performed as described in the Experimental Section. 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 reagents sulfo-EGS, BS^3, and DSA-d_0/d_8 (Table 1), were used for conducting the cross-linking experiments.

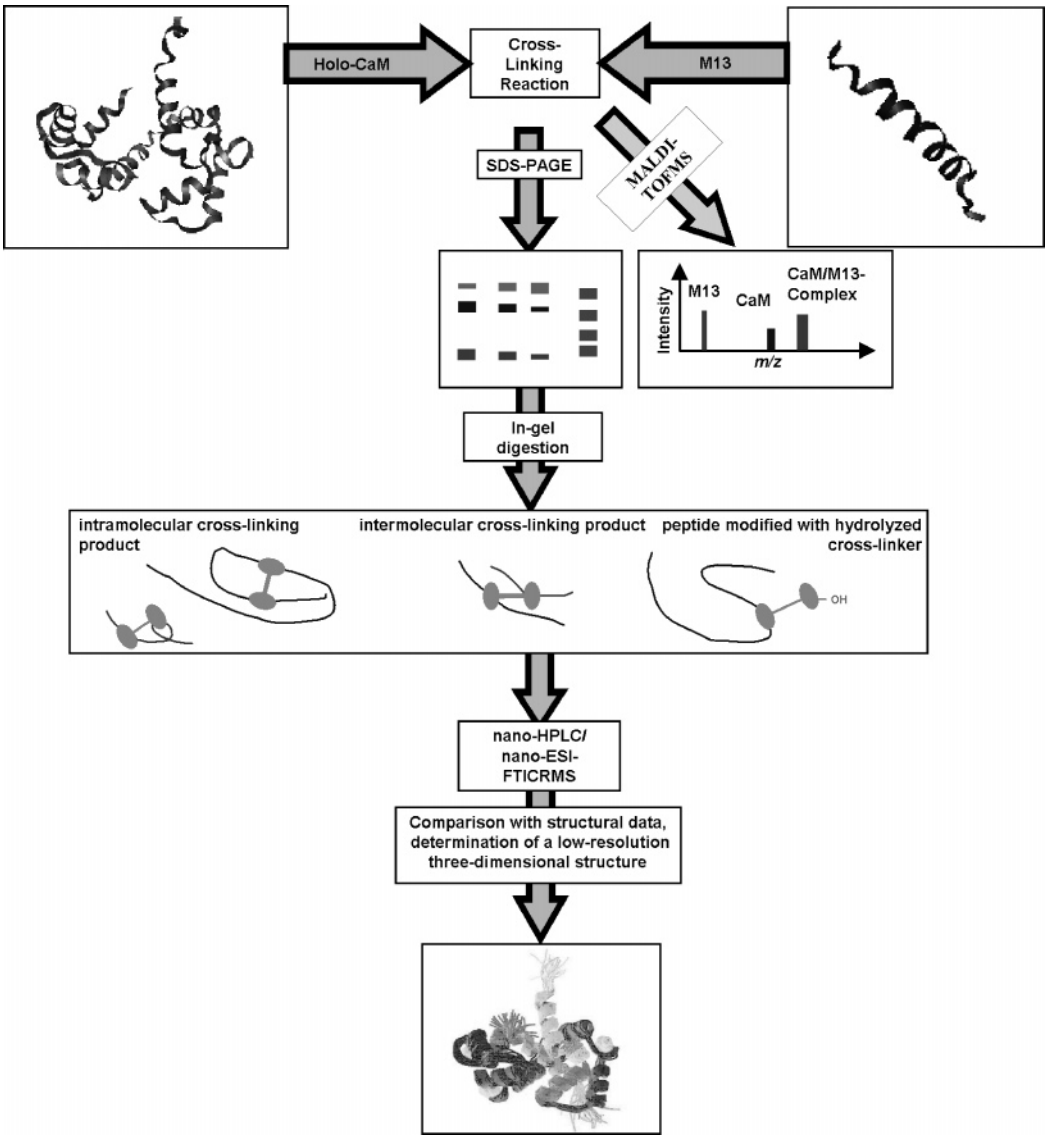
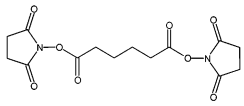
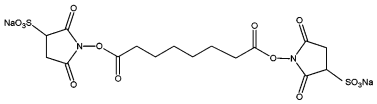
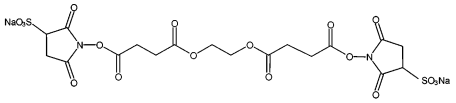
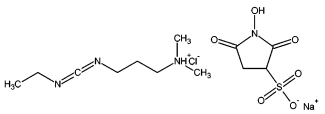


Figure 2. General analytical strategy for mapping the interacting sequences between CaM and M13.

Table 1. Structures and Spacer Lengths of the Cross-Linking Reagents Used for This Study

Name	Chemical Formula	Spacer Length
DSA		8.9 Å
BS^3		11.4 Å
Sulfo-EGS		16.1 Å
EDC / Sulfo-NHS		'zero-length'

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.[21] In the presence of an ϵ-amino group of lysine or the free N-terminus of a protein, an amide bond is formed, leading to the loss of an H_2O molecule and yielding a mass decrease of 18.011 u 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-EGS, BS^3, and DSA, are highly reactive toward primary amine groups; however, they are susceptible to hydrolysis. Upon cross-linking, sulfo-EGS, BS^3, and DSA produce amide bond cross-linked molecules causing a mass shift of 113.995, 138.068, and 110.0368 u (DSA-d_0), respectively, whereas partially hydrolyzed cross-linkers exhibit a mass increase of 132.006, 156.079, and 128.0472 u (DSA-d_0), respectively.

The optimum cross-linking reaction results in the high-yield formation of a cross-linked complex without creating oligomers and without distorting the tertiary structure of the complex by excessive cross-linking. Therefore, the right choice of both the cross-linker concentration and the reaction times is crucial for successfully conducting cross-linking experiments.

SDS−PAGE. After the cross-linking reaction, the reaction mixtures were separated by one-dimensional SDS−PAGE and the gels were stained with Coomassie Brilliant Blue. Figure 3 shows a one-dimensional gel of CaM, M13, and the cross-linking reaction product of an equimolar mixture of both proteins employing EDC/sulfo-NHS (20 mM/500 μM) after incubation times of 5, 15, 30, 60, and 120 min. CaM and M13 possess molecular weights of about 16 800 and 29 00, respectively. Thus, the cross-linked complex between CaM and M13 is expected to migrate on the gel at ~20 000 depending on the extent of chemical cross-linking. After incubation times of 5, 15, 30, and 60 min, two distinct bands

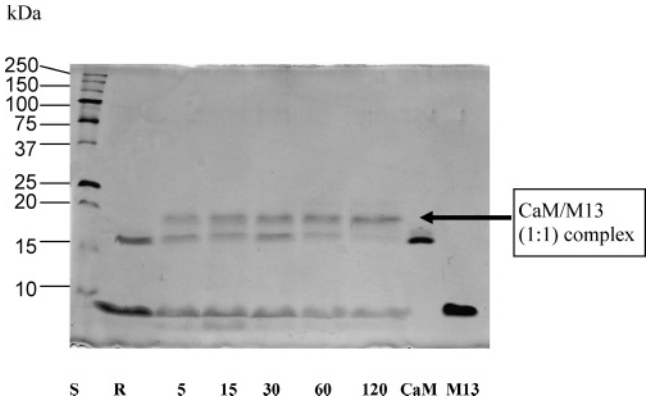


Figure 3. SDS−PAGE of cross-linking reaction mixtures of CaM (10 μM) and M13 (10 μM) with EDC (20 mM)/sulfo-NHS (500 μM). The samples are labeled as follows: Protein standard (S), CaM/M13 (1:1) mixture without cross-linker (R), reaction mixtures after 5 (5), 15 (15), 30 (30), 60 (60), and 120 min (120).

between ~17 000 and 20 000 were observed on the gel, corresponding to CaM and the CaM/M13 (1:1) complex (Figure 3). Additionally, a band in the low molecular weight region representing M13 indicated that cross-linking was not complete. The appearance of a band in the low molecular weight region in addition to the M13 band might be caused by intramolecular cross-linking of M13 causing a compact structure of the peptide with a higher mobility than M13 itself. Aggregation of proteins caused by excessive cross-linking was not observed for any of the cross-linking reagents as was evidenced by the absence of any gel bands in the higher mass range.

MALDI-TOF Mass Spectrometry. MALDI-TOFMS was employed to estimate the extent of chemical cross-linking over the course of the cross-linking reaction. MALDI-TOF mass spectrometry yields information on the number of intra- and intermolecular cross-linking products as well as on the number of modifications caused by hydrolyzed cross-linker. Figure 4 shows MALDI-TOF mass spectra of the nondigested reaction mixtures of CaM and M13 after incubation times of 5, 15, 30, 60, and 120 min with sulfo-EGS at 50 molar excess over the protein/peptide concentration. Based on the cross-linking experiments conducted with sulfo-EGS, complex formation was found to be complete after ~60 min when sulfo-EGS was applied at 50- or 100-fold molar excess. The number of incorporated cross-linker molecules increased from three to four during the second hour of the reaction (Figure 4). With increasing reaction time, the formation of a small amount of CaM/M13 (1:2) complex was also observed. Applying a low cross-linker concentration resulted in reduced complex formation as well as reduced modification with hydrolyzed cross-linker. For the reaction employing a 10-fold molar excess of cross-linking reagent over the protein/peptide concentration, we observed a highly reduced yield of cross-linking product after 120 min (data not shown). The cross-linker EDC/sulfo-NHS had to be used at a significantly higher molar excess in comparison to the homobifunctional amine-reactive cross-linkers employed in this study. Optimization of cross-linking reaction conditions resulted in the formation of a

(20) Watterson, D. M.; Sharief, F.; Vanaman, T. C. *J. Biol. Chem.* **1980**, *255*, 962−975.
(21) Hermanson, G. T. *Bioconjugate Techniques*, 1st ed.; Academic Press: San Diego, 1996.

Table 2. Intermolecular Cross-Linking Products between CaM and M13 with EDC/Sulfo-NHS[a]

$[M+H]^+_{calc}$	$[M+H]^+_{exp}$	CaM	M13	distance (Å)	Δm (ppm)	identified in sample
2224.113	2224.136	75−86	19−26		11	EDC_30min_2000x
	2224.129	D 78	K 19	13.76	7	EDC_120min_1000x
		D 80	K 19	12.47		
		E 82	K 19	11.62		
		E 83	K 19	3.66		
		E 84	K 19	6.90		
1836.882	1836.891	78−86	19−26		5	EDC_60min_2000x
		E 82	K 19	11.62		
		E 83	K 19	3.66		
		E 84	K 19	6.90		
		D 78	K 19	13.76		
		D 80	K 19	12.47		

[a] The reaction times and the excess of EDC in the samples, in which cross-links were identified, are given.

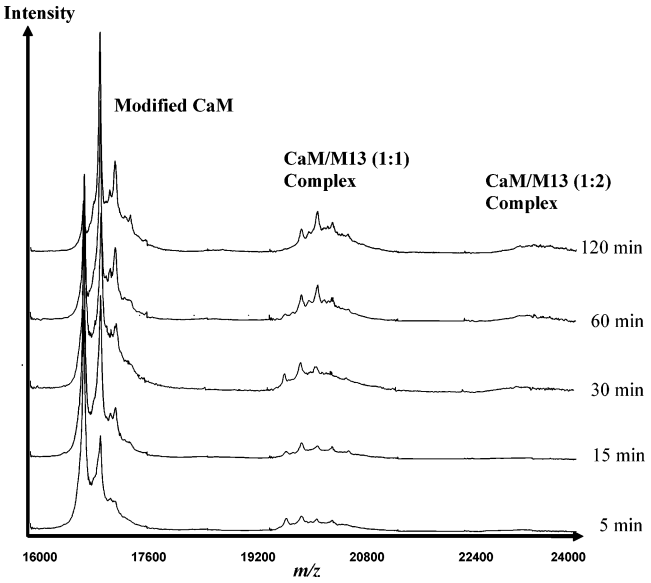


Figure 4. MALDI-TOF mass spectra (detection range *m*/*z* 16 t000−24t000) of the reaction mixture between CaM and M13 with 50-fold molar excess of sulfo-EGS at different reaction times (from bottom to top: 5−120 min).

sufficient amount of cross-linked complex after 15 min, if a 500−2000-fold molar excess of EDC in combination with a 500-fold molar excess of sulfo-NHS was employed (data not shown).

Analysis of Cross-Linking Products. Following SDS−PAGE separation of the cross-linking reaction mixtures, the cross-linked CaM/M13 (1:1) complexes were excised from the gel and subjected to enzymatic in-gel digestion with trypsin (Figure 2). The created peptide mixtures were separated by nano-HPLC and analyzed by nano-ESI-FTICR mass spectrometry (Figure 2). These mixtures were rather complex, as they contained peptides derived from CaM and M13 in addition to inter- and intramolecular cross-linking products and peptides modified by hydrolyzed cross-linker. The experimentally obtained monoisotopic masses were compared to calculated masses of peptides and cross-linking products employing the GPMAW and ASAP software packages.

Cross-Linking Products Obtained with EDC/Sulfo-NHS. The "zero-length" cross-linker EDC was employed in combination with sulfo-NHS at a molar excess of 500-, 1000-, and 2000-fold compared to protein/peptide concentration. Data analysis yielded two cross-linking products composed of the C-terminal sequence (amino acids 19−26) of M13 with amino acid sequences 75−86 and 78−86 of CaM (Table 2). The sequences of CaM contain as possible reactive sites two aspartic acid residues at positions 78 and 80 as well as three glutamic acid residues at positions 82, 83, and 84. Which amino acids were actually cross-linked by reaction with Lys-19 of M13, however, remains unclear without conducting MS/MS experiments.

Calculating the atomic distances between the nitrogen atoms of the amine groups of lysines and the C-atoms of carboxylic acid groups of aspartic acids and glutamic acids revealed that only a single distance is under 5 Å. This distance is between Glu-83 (in CaM) and Lys-19 (in M13). The possible cross-linking product between these two amino acids is in agreement with the fact that the zero-length cross-linker EDC is able to connect amine and carboxylic acid groups, which are within distances of under 5 Å.

Cross-Linking Products Obtained with Sulfo-EGS. In addition to the zero-length cross-linker EDC, three different amine-reactive homobifunctional cross-linking reagents with spacer lengths of 8.9 (DSA), 11.4 (BS^3), and 16.4 Å (sulfo-EGS) were employed. For all three reagents, the formation of a cross-linked CaM/M13 (1:1) complex was confirmed using MALDI-TOFMS and SDS−PAGE. Using sulfo-EGS, two cross-linking products were identified by nano-HPLC/nano-ESI-FTICRMS analysis of the proteolytic peptide mixtures (Table 3). A signal at *m*/*z* 1393.724 was assigned as cross-linking product between amino acids 75−77 of CaM and amino acids 19−26 of M13. As cross-linked amino acids, the residues Lys-19 (in M13) and Lys-75 (in CaM) were identified. As trypsin is expected not to cleave after a modified lysine residue, Lys-77 of CaM could be excluded as a cross-linked amino acid. Another cross-linking product composed of amino acids 17−19 (M13) and amino acids 76−86 (CaM) was identified based on the appearance of a signal at *m*/*z* 1999.916. In addition to the intermolecular cross-linking products between CaM and M13, an intramolecular cross-linking product within M13 consisting of the amino acid sequence 17−26 of M13 (Table 4) was identified in analyzing the CaM/M13 (1:1) complex. This cross-linking product could only have been created by cross-linking between M13 lysines at positions 18 and 19. Moreover, several modified peptides of M13 were identified as lysines at positions 5, 6, 18, and 19 in M13 derivatized with hydrolyzed sulfo-EGS (data not shown).

Table 3. Intermolecular Cross-Linking Products between CaM and M13 with the Homobifunctional, Amine-Reactive Cross-Linkers Sulfo-EGS, BS³, and DSA (d_0/d_8)[a]

cross-linker	$[M+H]^+_{calc}$	$[M+H]^+_{exp}$	CaM	M13	distance (Å)	Δm (ppm)	identified in sample
sulfo-EGS	1393.724	1393.734	75−77	19−26		7	EGS_120min_100x
		1393.735	K 75	K 19	18.12	8	EGS_120min_200x
	1999.916	1999.918	76−86	17−19		1	EGS_60min_200x
		1999.920	K 77	K 18	29.79	2	EGS_15min_200x
		1999.918				1	EGS_60min_100x
		1999.915				1	EGS_120min_100x
BS³	2380.191	2380.181	75−86	19−26		4	BS³_30min_100x
		2380.193	K 75	K 19	18.12	1	BS³_60min_100x
		2380.182	K 77	K 19	20.58	4	BS³_100x_120min
	2396.186 [M76 oxidized]	2396.185	75−86	19−26		0	BS³_60min_100x
			K 75	K 19	18.12		
			K 77	K 19	20.58		
	2699.360 (BS³+ linker)	2699.332	75−90	17−19		10	BS³_60min_100x
			K 75	K 18	26.24		
			K 77	K 18	29.79		
	2040.032	2040.032	75−86	17−19		6	BS³_60min_100x
			K 75	K 18	26.24		
			K 77	K 18	29.79		
DSA-d_0	1883.906	1883.904	76−86	17−19		1	DSA_120min_100x
		1883.902	K 77	K 18	29.79	2	DSA_120min_50x
DSA-d_8	1891.956	1891.948				4	DSA_120min_100x
		1891.950				3	DSA_120min_50x

[a] The reaction times and the excess of cross-linking reagent in the samples, in which cross-links were identified, are given.

Table 4. Intramolecular Cross-Linking Products in CaM and M13 with the Homobifunctional, Amine-Reactive Cross-Linkers Sulfo-EGS and BS³[a]

cross-linker	$[M+H]^+_{calc}$	$[M+H]^+_{ex}$	component	sequence	distance (Å)	Δm (ppm)	identified in sample
sulfo-EGS	1265.662	1265.650	M13	17−26 K18/K19	10.62	9	EGS_60min_50x
BS³	1175.667	1175.672	M 13	17−26	10.62	4	BS³_30min_100x
		1175.667		K18/K19		0	BS³_120min_100x
	2122.012	2122.007	CaM	75−90 K75/K77	5.38	2	BS³_120min_100x
	2138.007	2138.003	CaM [M76 oxidized]	75−90 K75/K77	5.38	2	BS³_120min_100x

[a] The reaction times and the excess of cross-linking reagent in the samples, in which cross-links were identified, are given.

Cross-Linking Products Obtained with BS³. Using BS³, four intermolecular cross-linking products identified between CaM and M13 pointed conclusively to a cross-linking between the sequence of amino acids 75−90 of CaM and the amino acid sequence 17−19 of M13 (Table 3). Figure 5 shows the deconvoluted ESI-FTICR mass spectrum of a tryptic peptide mixture of the CaM/M13 (1:1) complex, which had been cross-linked by BS³. The signal at m/z 2380.193 corresponds to the singly charged ion of a cross-linking product consisting of sequence 75−86 of CaM and sequence 19−26 of M13. The signal at m/z 2396.185 was identified as the Met-oxidized (Met-76 in CaM) derivative of the respective cross-linking product. Additionally, three intramolecular cross-linking products were detected (Table 4) and identified as amino acid Lys-18 of M13 cross-linked with Lys-19 and Lys-75 of CaM cross-linked with Lys-77. Moreover, we identified several peptides of both M13 and CaM that had been modified by hydrolyzed BS³. In M13, the N-terminal amino acid and in CaM, lysine residues at positions 21, 77, and 94 were found to be modified, indicating that the functional groups of these amino acids are highly solvent-exposed.

Cross-Linking Products Obtained with DSA. DSA was employed for the cross-linking reaction as a 1:1 mixture with its fully deuterated derivative (d_8). Therefore, an additional criterion

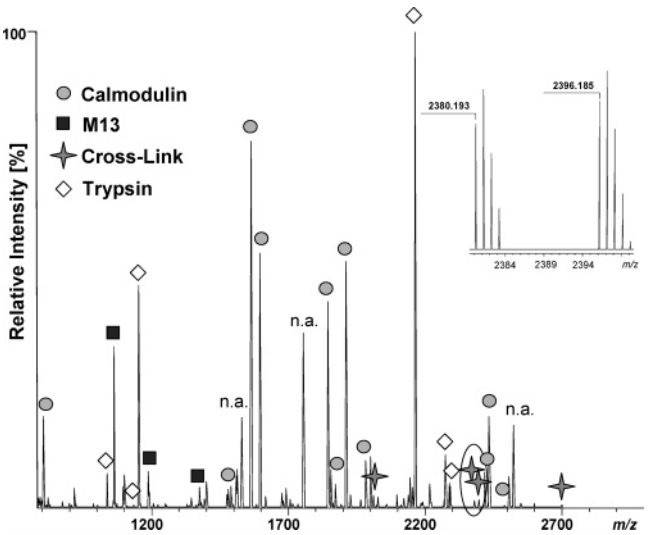


Figure 5. Deconvoluted ESI-FTICR mass spectrum of the tryptic peptide mixture from the CaM/M13 (1:1) complex cross-linked with BS³ (100-fold molar excess over protein/peptide concentration) at an incubation time of 60 min. The inset shows the magnified signal of a cross-linking product between CaM residues 75−86 and M13 residues 19−26 at m/z 2380.193, which was also detected with one methionine residue oxidized at m/z 2396.185; n.a., signal not assigned.

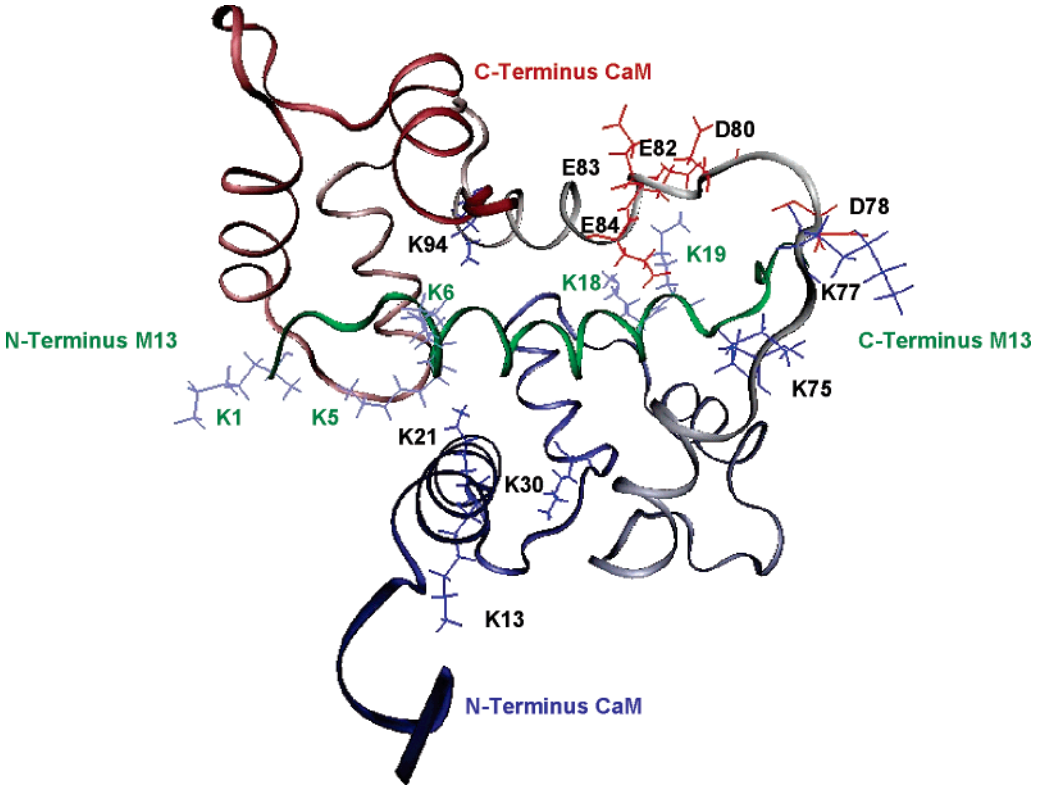


Figure 6. NMR structure of the CaM/M13 complex according to ref 15 (pdb entry 2BBM). The side chains of lysines and acidic amino acids that are potentially involved in cross-linking are indicated.

for the identification of cross-linking products was introduced, as every cross-linking product should exhibit two signals at a distance of 8 u with approximately the same intensity.[22] Two signals showing this mass difference were detected, which were identified as deuterated and nondeuterated cross-linking product between between the amino acid sequence 76−86 in CaM and the amino acid sequence 17−19 in M13 (Table 3). As trypsin was cleaving after Lys-19 in M13, lysine in position 18 could be unambiguously identified as the amino acid cross-linked with Lys-77 of CaM.

Intermolecular Cross-Linking Products between CaM and M13. Based on the published NMR structure (pdb entry 2BBM), the distances of N atoms of all lysine residues in CaM and M13 were calculated. The NMR structure[15] exhibits only one intermolecular distance lower than 10 Å, another distance under 15 Å, and nine intermolecular distances of under 20 Å. For the amine-reactive homobifunctional cross-linkers, we successfully identified seven cross-linking products containing sequences from the central α-helix of CaM (residues 75−90) and amino acids 18 and 19 of M13. Figure 6 shows the published NMR structure of the CaM/M13 complex, indicating the amino acid side chains that are potentially involved in cross-linking. It can be noted that all identified cross-linking products are in agreement with published structural data. However, we did not identify any cross-linking products containing lysines in positions 13, 21, 94, or 148 with any of the lysine residues of M13, although they are located less than 20 Å from each other.

Another interesting aspect is that Lys-18 of M13 was found to be cross-linked with Lys-77 of CaM with all amine-reactive cross-linkers used for this study: DSA, BS^3, and sulfo EGS (spacer lengths between ~9 and 16 Å). According to the published NMR structure,[15] the N-atoms of these two lysine residues are ~29 Å apart from each other (Figure 6). However, a closer look at the published structure reveals that the region comprising amino acids 74−82 in CaM is poorly defined and that the nitrogen atoms of Lys-77 in CaM and Lys-18 in M13 are within ~20 and 35 Å of each other. Apparently, the amine groups of Lys-77 in CaM and Lys-18 in M13 are coming close enough to be cross-linked by BS^3, DSA, and sulfo-EGS.

CONCLUSION

We successfully demonstrated the applicability of chemical cross-linking in combination with high-resolution FTICR mass spectrometry as a rapid method to determine interfaces between proteins or between a protein and a peptide. One advantage of the described strategy consists of the possibility of analyzing the topology of protein complexes within a few days—given that data analysis can be automated. (For the present study, however, analysis of the CaM/M13 complex required several weeks, which was mainly caused by the time-intense assignment of cross-linking products in the mass spectra.) Other advantages of the presented strategy include the low material consumption, allowing for analysis of femtomole amounts of proteins, as well as the possibility of obtaining solution structures of proteins. When lysines serve as targets for cross-linking reactions, valuable structural information of a protein can be obtained despite the flexibility of the lysine side chains. For a comprehensive structural analysis of a protein complex, however, one should apply a variety of cross-linking reagents with different reactivities, targeting both flexible and rigid amino acid side chains.

(22) Pearson, K. M.; Pannell, L. K.; Fales, H. M. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 149−159.

In general, the described strategy should not be considered as competing with high-resolution methods for three-dimensional structure analysis, such as NMR spectroscopy or X-ray crystallography, as only low-resolution structures are obtained. One major drawback of this approach is that not all predicted cross-linking products were found, although the functional groups with the right specificities were the correct distance for cross-linking. For example, no cross-linking products were identified between Lys-21 in CaM and Lys-6 in M13 although the two nitrogen atoms are within a distance of ~15 Å. One cannot draw any conclusion if a defined cross-linking product is not found, since the cross-linking product may not have been created at all or the created product may not have been amenable to subsequent mass spectrometric analysis. There is also a slight possibility that cross-linked peptides do not elute from the gel after enzymatic in-gel digestion and thus escape from analysis.

When studying relatively small complexes—as in the present case—for which the amino acid sequences of the binding partners are known, mass measurement accuracies of up to ~15 ppm are sufficient in order to unambiguously identify a cross-linking product. For larger complexes, additional information, e.g., by using isotope-labeled cross-linkers, is needed. Moreover, for an exact identification of cross-linked amino acids, MS/MS experiments need to be conducted. With our instrumentation, however, it is not feasible to perform on-line MS/MS experiments, for example, using electron capture dissociation or infrared multiphoton dissociation during an LC/MS analysis. For future studies, we are aiming to employ instruments such as the novel commercially available hybrid FTICR mass spectrometers,[23] which offer the possibility to conveniently acquire MS/MS data during an LC/MS run.

(23) Gershon, D. *Nature* **2003**, *424*, 581−587.

ACKNOWLEDGMENT

The junior research group of A.S. is funded by the Saxon State Ministry of Higher Education, Research and Culture. Financial support from the Thermo Electron Corp. (Mattauch-Herzog award of the German Society for Mass Spectrometry to A.S.) is also gratefully acknowledged. The authors are indebted to Dr. Ad Bax for the generous gift of M13, to Dr. Mike Lewis for critical reading of the manuscript, and to Ms. Daniela Schulz for valuable discussions.

Received for review August 26, 2004. Accepted October 22, 2004.

AC0487294