Received: 12 April 2010

(www.interscience.com) DOI 10.1002/jms.1775

Fragmentation behavior of a thiourea-based reagent for protein structure analysis by collision-induced dissociative chemical cross-linking

Mathias Q. Müller,^a Frank Dreiocker,^b Christian H. Ihling,^a Mathias Schäfer^{b*} and Andrea Sinz^{a*}

The fragmentation behavior of a novel thiourea-based cross-linker molecule specifically designed for collision-induced dissociation (CID) MS/MS experiments is described. The development of this cross-linker is part of our ongoing efforts to synthesize novel reagents, which create either characteristic fragment ions or indicative constant neutral losses (CNLs) during tandem mass spectrometry allowing a selective and sensitive analysis of cross-linked products. The new derivatizing reagent for chemical cross-linking solely contains a thiourea moiety that is flanked by two amine-reactive *N*-hydroxy succinimide (NHS) ester moieties for reaction with lysines or free *N*-termini in proteins. The new reagent offers simple synthetic access and easy structural variation of either length or functionalities at both ends. The thiourea moiety exhibits specifically tailored CID fragmentation capabilities – a characteristic CNL of 85 u – ensuring a reliable detection of derivatized peptides by both electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) tandem mass spectrometry and as such possesses a versatile applicability for chemical cross-linking studies. A detailed examination of the CID behavior of the presented thiourea-based reagent reveals that slight structural variations of the reagent will be necessary to ensure its comprehensive and efficient application for chemical cross-linking of proteins. Copyright © 2010 John Wiley & Sons, Ltd.

Supporting information may be found in the online version of this article.

Keywords: chemical cross-linking; protein structure analysis; CID; CNL

Introduction

Chemical cross-linking combined with mass spectrometry presents an emerging approach to study three-dimensional protein structures and to map protein interfaces.^[1-4] Yet, an unambiguous, sensitive, reliable and rapid identification of the amino acids involved in the covalent derivatization by chemical cross-linking still poses a challenge to the scientist. A mass spectrometric identification of the chemically modified amino acids in the respective protein is performed after enzymatic digestion of the cross-linking reaction mixture. This is often hampered by the complexity of the created peptide mixtures as often only a relatively small percentage of cross-linked products are present besides a large majority of unmodified peptides. One strategy for a facilitated mass spectrometric analysis of cross-linked products relies on the application of cross-linkers that fragment readily during collision-induced dissociation (CID) and give rise to characteristic fragment ions.[5-8]

In a previous publication,^[9] we reported the synthesis and application of a novel dissociative chemical cross-linker with a thiourea moiety, containing a highly nucleophilic sulfur for an attack at the adjacent glycylproline amide carbonyl. This structural feature allows for efficiently initiating cleavage of the cross-linker molecule upon CID. The characteristic fragmentation of the crosslinker molecule delivered highly indicative mass shifted product ions and constant neutral losses (CNLs), enabling a sensitive, selective and unambiguous detection of cross-linked species by MS/MS. Yet, the synthesis of this 'first-generation' cleavable crosslinker proved to be quite challenging. Therefore, we extended the thiourea-based concept of a dissociative cross-linker toward a simplified symmetrical thiourea analog, which is presented herein. The characteristic fragmentation of the thiourea-based reagent is directed by the influence of the thiourea moiety that is incorporated into the reagent. The applicability of our novel reagent as chemical cross-linker was examined with a set of model compounds, i.e. substance P (amino acid sequence RPKPQQFFGLM-NH₂), luteinizing hormone-releasing hormone (LHRH, amino acid sequence pEHWSYGLRPG) as well as the proteins hen egg lysozyme (14.3 kDa) and the ligand binding

Mathias Schäfer, Institute of Organic Chemistry, Department of Chemistry, Universität zu Köln, Greinstr. 4, D-50939 Köln, Germany. E-mail: mathias.schaefer@uni-koeln.de

- a Department of Pharmaceutical Chemistry and Bioanalytics, Institute of Pharmacy, Martin-Luther-Universität Halle-Wittenberg, Wolfgang-Langenbeck-Str.
 4, D-06120 Halle (Saale), Germany
- b Institute of Organic Chemistry, Department of Chemistry, Universität zu Köln, Greinstr. 4, D-50939 Köln, Germany

^{*} Correspondence to: Andrea Sinz, Department of Pharmaceutical Chemistry and Bioanalytics, Institute of Pharmacy, Martin-Luther-Universität Halle-Wittenberg, Wolfgang-Langenbeck-Str. 4, D-06120 Halle (Saale), Germany. E-mail: andrea.sinz@pharmazie.uni-halle.de



domain of the peroxisome proliferator-activated receptor alpha (PPAR α , ~32 kDa) with ESI (electrospray ionization)^[10] and matrixassisted laser desorption/ionization (MALDI)^[11] tandem mass spectrometry. The fragmentation behavior of the CID cleavable cross-linker was examined over a wide range of activation energies using both a linear quadrupole ion trap and a TOF/TOF instrument.

Experimental

Materials

All chemicals and solvents were used as purchased without further purification (Acros Organics, Geel, Belgium; ABCR, Karlsruhe, Germany). Toluene and diethyl ether were dried by distillation from benzophenone and sodium, dichloromethane and pyridine by distillation from calcium hydride. All other solvents were distilled over a column prior to use. Hen egg lysozyme, substance P, LHRH and buffer reagents were obtained from Sigma (Taufkirchen, Germany). The protease bovine trypsin was obtained from Roche Diagnostics (Mannheim, Germany). Nano-HPLC solvents were of spectrometric grade (Uvasol, VWR, Darmstadt, Germany). MALDI matrices and calibration standards were purchased from Bruker Daltonik (Bremen, Germany). Water was purified with a Direct-Q5 water purification system (Millipore, Eschborn, Germany). The ligand binding domain of PPAR α was expressed in *Escherichia coli* and purified according to a previously published protocol.^[12]

Synthesis of the thiourea cross-linker

Esterification of 4-amino-butyric acid (1, Scheme S1, Supporting Information) was achieved by reaction with thionyl chloride in methanol for 16 h. After the evaporation of the solvent from the reaction mixture, the crude product was stirred in ethyl acetate. The hydrochloride of 4-amino-butyric acid methyl ester (2) was isolated by filtration as a colorless solid at a yield of 98% (Scheme S1).^[13]

For the synthesis of 4-isothiocyanato-butyric acid methyl ester (**3**), **2** was suspended in a mixture of dichloromethane and carbon disulfide. After slowly adding triethylamine, the reaction mixture was stirred for 30 min before methyl chloroformate was added. After refluxing for 4 h, the mixture was cooled down to 20 °C and poured into water. The organic layer was washed with diluted hydrochloric acid and brine, dried over magnesium sulfate and concentrated under reduced pressure. The resulting yellowish oil was distilled in high vacuum to give **3** as a colorless oil at a yield of 80%.^[14]

For the reaction of the isothiocyanate **3** and the protected amine **2**, the amine was suspended in dichloromethane and treated with triethylamine. After 15 min, **3** was added at 0 °C and the reaction mixture was brought to 20 °C within 3 h. After washing with diluted hydrochloric acid and brine, the organic layer was dried over magnesium sulfate and the solvent was removed under reduced pressure. The resulting yellowish oil was purified by column chromatography with ethyl acetate/cyclohexane [1:2 (v/v)] to give 4-[3-(3-methoxycarbonyl-propyl)-thioureido]-butyric acid methyl ester (**4**) as colorless oil at a yield of 91%.^[15,16]

Hydrolysis of the two methyl ester functionalities of **4** was achieved with lithium hydroxide in a mixture of tetrahydrofurane, methanol and water for 15 h at 20 °C. After the removal of the solvent under reduced pressure, the residue was dissolved in water and acidified to pH 2 using 2 M hydrochloric acid. After seven extractions with ethyl acetate, the combined organic

layers were dried over magnesium sulfate and the solvent was removed under reduced pressure to give 4-[3-(3-carboxy-propyl)-thioureido]-butyric acid (**5**) as a colorless solid at a yield of 92%.^[17]

For activation of the two carboxylic acid functionalities to give the *N*-hydroxy succinimide (NHS) ester, **5** was dissolved in pyridine and treated with *N*-trifluoroacetoxy succinimide (**6**) – prepared from NHS and trifluoroacetic anhydride^[18] – at 20 °C. After 1 h, the solution was poured into diethyl ether and the precipitate was removed by filtration. The solid was washed with diethyl ether and dissolved in a mixture of dichloromethane and toluene. After the removal of the solvent under reduced pressure, the residue was dried under high vacuum to give 4-{3-[3-(2,5-dioxopyrrolidin-1-yloxycarbonyl]-propyl]-thioureido}-butyric acid 2,5dioxo-pyrrolidin-1-yl ester (**7**) as a colorless oil at a yield of 93%.^[19]

Cross-linking reactions

For cross-linking experiments, aqueous stock solutions of substance P (100 µg/ml), LHRH (100 µg/ml), lysozyme (10 mg/ml) or PPAR α ligand binding domain (2 mg/ml) were diluted with 20 mm 4-(2-hydroxyethyl)-piperazine-1-ethanesulfonic acid (HEPES) buffer (pH 7.4) to give 1 ml solution containing the protein/peptide at a concentration of 10 µм. The cross-linker (200 mм stock solution in dimethylsulfoxide) was added in 50-, 100- and 200-fold molar excess to the protein/peptide solution and the reactions were allowed to proceed for 5, 15, 30, 60 and 120 min. Reactions were quenched with ammonium bicarbonate (20 mm final concentration). One 200-µl aliquot was taken from each sample and stored at -20 °C before MS analysis. Cross-linked lysozyme and PPAR α were digested with trypsin (1:100 (w/w) enzyme to substrate ratio) overnight at 37 °C according to an existing protocol.^[20] The resulting digests were stored at $-20\,^\circ\text{C}$ before MS analysis.

MALDI-TOF mass spectrometry linear mode

MALDI-TOF mass spectrometry of intact cross-linked lysozyme and PPAR α was performed in linear and positive ionization mode on an Ultraflex III instrument (Bruker Daltonik, Bremen, Germany) equipped with a Smart beamTM laser using sinapinic acid as matrix. Five picomole of each sample was applied onto a ground steel target using the double layer method.

Nano-HPLC/MALDI-TOF/TOF mass spectrometry

Tryptic peptide mixtures of lysozyme and PPAR α were analyzed by offline coupling of a nano-HPLC system (Ultimate 3000, Dionex, Idstein, Germany) to a MALDI-TOF/TOF mass spectrometer (Ultraflex III), whereas the peptides substance P and LHRH were analyzed by MALDI-TOF/TOF-MS/MS without a preceding nano-HPLC separation. Samples were injected onto a precolumn (Acclaim PepMap, C18, 300 μ m \times 5 mm, 5 μ m, 100 Å, Dionex) and desalted by washing the precolumn for 15 min with 0.1% trifluoroacetic acid (TFA) before the peptides were eluted onto the separation column (Acclaim PepMap, C18, 75 μ m \times 150 mm, 3 μ m, 100 Å, Dionex), which had been equilibrated with 95% solvent A (A: 5% acetonitrile (ACN), 0.05% TFA). Peptides were separated with a 60-min gradient (0-60 min: 5-50% B, 60-62 min: 50-100% B, 62-67 min: 100% B, 67-72 min: 5% B, with solvent B: 80% ACN, 0.04% TFA) at a flow rate of 300 nl/min with UV detection at 214 and 280 nm. Eluates were fractionated into 18-s fractions using the fraction collector Proteineer fc (Bruker Daltonik), mixed with 1.1 µl

of matrix solution (0.8 mg/ml α -cyano-4-hydroxy cinnamic acid in 90% ACN, 0.1% TFA, 1 mM NH₄H₂PO₄) and directly prepared onto a 384 MTP 800 µm AnchorChip target (Bruker Daltonik). MALDI-TOF/TOF-MS analyses were conducted in the positive ionization and reflectron mode by accumulating 2000 laser shots in the range *m/z* 800–5000 to one mass spectrum. Mass spectra were externally calibrated using Peptide Calibration Standard II (Bruker Daltonik). Signals with a *signal-to-noise ratio* (S/N) >15 were isolated for laser-induced fragmentation. In-source decay (ISD) was performed manually on isolated cross-linked product candidates by increasing the laser energy by *ca* 10% relative to the usual laser energy in LID (laser-induced dissociation) MS/MS. Data acquisition was done automatically by the WarpLC 1.1 software (Bruker Daltonik) coordinating MS data acquisition (FlexControl 1.3) and data processing (FlexAnalysis 3.0) softwares.

Nano-HPLC/Nano-ESI-LTQ-Orbitrap mass spectrometry

Fractionation of tryptic peptide mixtures (lysozyme and PPAR α) was carried out on an Ultimate nano-HPLC system (Dionex Corporation, Idstein, Germany) using reversed phase C18 columns (precolumn: Acclaim PepMap, $300 \,\mu m \times 5 \,mm$, $5 \,\mu m$, $100 \,\text{\AA}$, separation column: Acclaim PepMap, 75 μ m \times 150 mm, 3 μ m, 100 Å, Dionex Corp., Idstein, Germany). After washing the precolumn for 15 min with water containing 0.1% TFA, the peptides were eluted and separated using a gradient from 0 to 50% B (90 min), 50 to 100% B (1 min) and 100% B (5 min), with solvent A: 5% ACN containing 0.1% formic acid (FA) and solvent B: 80% ACN containing 0.1% FA. The nano-HPLC system was directly coupled to the nano-ESI source (Proxeon, Odense, Denmark) of an LTQ (linear ion trap)-Orbitrap XL hybrid mass spectrometer (ThermoFisher Scientific, Bremen, Germany). MS data were acquired over 122 min in a data-dependent MS² mode: each high-resolution full scan (m/z 300–2000, resolution was set to 60 000) in the orbitrap was followed by three product ion scans in the orbitrap (resolution 15 000) for the three most intense signals in the full-scan mass spectrum (isolation window 1.5 u). MS³ in the orbitrap was performed on the two most intense signals in the MS² spectra with a resolution of 15 000. Dynamic exclusion (exclusion duration 180 s, exclusion window -1 to 2 Th (Thomson [m/z])) was enabled to allow detection of less abundant ions. Data acquisition was controlled via XCalibur 2.0.7 (ThermoFisher) in combination with DCMS link 2.0 (Dionex).

Offline nano-ESI-LTQ-Orbitrap-MS

Cross-linked substance P and LHRH were additionally analyzed by offline nano-ESI-LTQ-Orbitrap-MS, with the resolution in the orbitrap set to 100 000. MS/MS and MS³ experiments were conducted with a relative collision energy in the linear ion trap (LTQ) of 35% and selected product ions were analyzed in the orbitrap analyzer set to a resolution of 100 000.

Analysis of cross-linked products

Cross-linked products were identified by analyzing the MS data using General Protein Mass Analysis for Windows (GPMAW)^[21] version 8.10 (Lighthouse Data, Odense, Denmark, http://www.gpmaw.com) and the CoolToolBox software program, which is an upgrade of the VIRTUALMSLAB software program.^[22] The length of the cross-linker was calculated using the Viewer Light 5.0 software (Accelrys).^[23]

Results and Discussion

Cross-linker design

Our novel dissociative thiourea cross-linker was developed on the basis of a previously synthesized cross-linker containing a thiourea moiety that is connected to proline via a glycine residue.^[9] This previous analytical concept relied on the lability of the glycylproline bond, which is preferably cleaved by an intramolecular attack of the strongly nucleophilic thiocarbonyl sulfur. This attack is promoted by the preferably protonated proline amide nitrogen. Although we proved the general feasibility of the proposed concept, the synthesis of that 'first-generation' collisioninduced dissociative cross-linker proved to be quite challenging, which motivated us to develop a more easily accessible thioureacontaining cross-linker. Such a novel and simplified thiourea-based reagent should exhibit specifically tailored CID fragmentation capabilities, thus ensuring a sensitive and selective detection of derivatized peptides. Moreover, we aimed to examine whether the presence of a proline is essential for the preferred cleavage of the cross-linker. Hence, the novel compound solely contains a thiourea moiety flanked by two amine-reactive NHS ester moieties (Scheme 1), which react with lysines or free *N*-termini in proteins. The reactivity of our novel cross-linker is comparable to classical NHS ester-based reagents, such as bis(sulfosuccinimidyl) suberate. The abbreviation BuTuBu for the novel designed linker that is used throughout this article is deduced from thiourea (Tu) and amino butyric acid (Bu) (Scheme 1). The reaction scheme of the thiourea cross-linker is presented in Scheme 2. After one reactive site has reacted with a primary amine in a protein, the second reactive site might react with another amine group forming an intramolecular cross-link (1). Alternatively, the second NHS functionality either stays intact or is hydrolyzed or aminolyzed during reaction guenching (2). For guenching the cross-linking reaction, ammonium salts, i.e. NH₄HCO₃, are added. From initial experiments, it immediately turned out that intramolecularly crosslinked species $[M + H + 212 u]^+$ (1) are accompanied by an isobaric cyclic 2-alkylylimino-[1,3]thiazepan-7-one (3). The latter is formed upon collision activation by rearrangement of (1), but might also originate from a 'dead-end' cross-linked species (2), i.e. a peptide that has been modified by a partially hydrolyzed cross-linker (type 0 cross-link).^[24] Eventually, the rearrangement of (3) leads to the loss of a γ -lactame (85 u) to form the 7-alkylylimino-[1,3]oxazepane-2thione derivative (4). This reaction is perfectly suited for a general identification of BuTuBu-derivatized species and greatly facilitates the identification of cross-linked peptides.



Scheme 1. Structure of the novel homobifunctional chemical cross-linker NHS–BuTuBu–NHS. The abbreviations are indicated using parenthesis.



Scheme 2. Reactions of the amine-reactive thiourea cross-linker NHS-BuTuBu-NHS.

Cross-linking of substance P

For initial experiments with the thiourea cross-linker, the 11amino acid peptide substance P was employed as it contains merely two reactive sites for the amine-reactive thiourea crosslinker, the lysine residue at position three of the peptide and the free N-terminus. Therefore, substance P presents a straightforward model system to study the properties of the novel cross-linker. The fragmentation behavior was evaluated in both ESI and MALDI mass spectrometric experiments. In Fig. 1A, the MALDI-TOF mass spectrum of substance P after reaction with the thiourea cross-linker is presented. From the MALDI spectrum alone it is not clear whether the cross-linker molecule is attached to Lys-3 or to the N-terminus. Strikingly, the signals at m/z1559.834 for the protonated species $[M + H + 212 u]^+$ and at m/z 1581.827 for the sodium adduct $[M + Na + 212 u]^+$ were found to correspond to cross-linker-modified substance P. At this point, it is important to note that the identity of the derivatized peptide molecular ions observed in Fig. 1A is somewhat ambiguous (see Scheme 2): either an intramolecular cross-link between the N-terminus and the side chain amine group of lysine is formed (structure 1 in Scheme 2) or alternatively, the cross-linker has underwent an intramolecular cyclization to form a seven-membered 2-alkylylimino-[1,3]thiazepan-7-one (structure 3 in Scheme 2). This characteristic cyclization reaction depends on the pronounced nucleophilicity of the thiocarbonyl sulfur, which intramolecularly attacks the adjacent carbonyl, leading to the formation of the seven-membered heterocycle (structure 3 in Scheme 2). According to Scheme 2, a mass increase of 212 u compared to the peptide is observed for the cross-linked product. The signal at m/z 1559.834 was isolated and examined by CID in the MALDI-TOF/TOF mass spectrometer resulting in a product ion at m/z 1474.7 that arises from a CNL of 85 u (structure 4 in Scheme 2). To explain the characteristic and abundant loss of 85 u, we propose a rearrangement reaction forming both the neutral γ -lactame (85 u) and a [M + BuTu + H]⁺ (i.e. 7-alkylimino-[1,3]oxazepane-2-thione derivative (4) in Scheme 2). Yet, from MALDI-MS/MS data alone it was not possible to definitely assign the modification to Lys-3, the N-terminus or an intramolecular cross-link between both amine groups, nor is it possible to rule out the possibility of having mixed isomeric species of the named options. This is due to the absence of a characteristic b₂ product ion, while the observed product ions are indicated in the insets of Fig. 1B. However, the dominant abundance of the product ion corresponding to the CNL of 85 u (Fig. 1B) clearly demonstrates that the respective intramolecular nucleophilic attack of the thiocarbonyl sulfur definitely presents the preferred fragmentation channel, confirming the validity of the basic concept of our novel cross-linker. Further information on the created product ion is obtained by additionally conducted MALDI-ISD (in-source decay) MS/MS experiments (Fig. 1C).

In order to evaluate the potential of the thiourea cross-linker for ESI mass spectrometry, LTQ-Orbitrap mass spectrometric analyses (MS and MS/MS) were conducted (Fig. 2). In contrast to the MALDI mass spectrum, the signal of the doubly charged peptide that is modified with a cross-linker is visible at *m*/*z* 780.407 in the ESI mass spectrum (Fig. 2A). In addition, the signals of the triply charged ions $[M + BuTuBu + 2H + K]^{3+}$ at *m*/*z* 533.256 and [M + BuTuBu + 2H +Na]³⁺ at *m*/*z* 527.930 are detected. Yet, the signal with the highest abundance is the species at *m*/*z* 780.407 with a mass increase of 212 u compared to the unmodified peptide, indicating the derivatization with the cross-linker. In order to definitely elucidate the identity of this ion, CID experiments were performed in the linear ion trap (LTQ), whereas exact ion mass analysis of fragment ions was performed with high resolution ($R_{FWHM} = 15000$) in



Figure 1. (A) MALDI-TOF mass spectrum of substance P ($[M + H]^+$ at m/z 1347.746) after cross-linking reaction with the thiourea linker, resulting in a mass shift of 212 u (reaction according to Scheme 2). Please note that a number of signals originating from impurities in the substance P sample are visible in the mass spectrum. The ion at m/z 1575.883 is labeled as $[M_{ox} + BuTuBu + H]^+$ referring to oxidized substance P after cross-linking reaction with the thiourea linker (mass shift of 16 u). Asterisks indicate impurities in substance P. (B) CID fragment ion mass spectrum (MALDI-TOF/TOF-MS/MS) of the signal at m/z 1559.834. The respective isolated and fragmented species are presented in the insets. (C) ISD-MALDI fragment ion mass spectrum of $[M + BuTu + H]^+$ at m/z 1474.8. \$: signal is not assigned.





Figure 2. (A) ESI-LTQ-Orbitrap mass spectrum of substance P after cross-linking reaction with the thiourea cross-linker. (B) CID fragment ion mass spectrum (MS/MS) of the signal at m/z 780.407 showing substance P modified with BuTuBu (Scheme **2**, structure **1** or **3**). (C) CID fragment ion mass spectrum (MS³) of the signal at m/z 706.37. The signals assigned with \$ are frequently observed in the spectra obtained with our LTQ-Orbitrap mass spectrometer.



Figure 3. Linear MALDI-TOF mass spectrum of lysozyme with different excess of cross-linker and reaction times. The average molecular weight of protonated hen egg lysozyme (14 313 Da) matches perfectly with the observed signal at *m/z* 14 313.

the orbitrap. The ion at m/z 780.407 (either structure 1 or 3) loses 85 u to form the product ion (4) at m/z 737.88. Furthermore, sequence analysis on the basis of the ESI-LTQ-Orbitrap CID tandem mass spectrum allowed an unambiguous assignment of the crosslinker modification to the N-terminus of substance P (Fig. 2B). In particular, the presence of y_9 - and y_{10} -NH₃ ions indicate an unmodified lysine at position 3 and a cyclized seven-membered 7-alkylylimino-[1,3]oxazepane-2-thione cross-linker (4) (inset of Fig. 2B). However, the absence of any substance P ions showing two 'dead-end' (type 0) linker^[24] molecules attached to both the N-terminus and the Lys at position 3 could give a hint for the presence of an intramolecular cross-link of the respective amine functionalities. It can be hypothesized that during the cross-linking reaction, an intramolecular (type 1) cross-link has been created, but that during CID, a conversion of type 1 into a type 0 cross-linked species occurred. We are neither able to exclude nor to verify any of the potential cross-linker-modified species that are created by the thiourea cross-linker. This certainly presents a drawback for a comprehensive application of this reagent for cross-linking studies of proteins.

Cross-linking of LHRH

For monitoring the side reactivity with amino acids containing hydroxyl groups, cross-linking experiments were performed with LHRH as it contains two potentially reactive amino acids Ser-4 and Tyr-5. It has been shown that NHS esters also react with hydroxyl groups in serines, threonines and tyrosines at pH 7.4, albeit with a lower frequency compared to lysines.^[25,26] Due to the pyroglutamate the *N*-terminus cannot be modified. The base peak showed unmodified LHRH at *m/z* 1183.576 and a lower yield

of a modified species with the characteristic mass shift of 212 u at m/z 1395.682. MS/MS data revealed the indicative CNL of 85 u both in MALDI-TOF/TOF and ESI-LTQ-Orbitrap mass spectrometry (data not shown).

Cross-linking of lysozyme

For evaluating the applicability of the thiourea cross-linker for proteins, the 14.3 kDa protein lysozyme was derivatized with our novel cross-linker. Efficient cross-linking was proven by MALDI-TOF mass spectrometric measurements in the linear mode showing the attachment of up to six – possibly seven – cross-linker molecules to the protein depending on molar excess (50-, 100- and 200-fold excess) of cross-linker and reaction times ranging between 5 and 60 min (Fig. 3). For gaining insight into the reaction sites of the cross-linker at the protein, modified lysozyme was *insolution* digested with trypsin and the resulting peptide mixtures were analyzed by nano-HPLC/MALDI-TOF/TOF-MS(/MS) and nano-HPLC/nano-ESI-LTQ-Orbitrap-MS(/MS).

Exemplarily, a lysozyme peptide comprising amino acids 115-125 (CKGTDVQAWIR) was thoroughly analyzed as it contains a reaction site at lysine in position 116 and another – yet less likely – reaction site at the threonine residue in position 118. From the MALDI-TOF mass spectrum presented in Fig. 4A, a modification with the thiourea linker is apparent as a molecular ion of the peptide 115-125 with the characteristic mass shift of 212 u is observed at m/z 1545.740. This species is accompanied by an ISD fragment at m/z 1460.661, corresponding to the characteristic neutral loss of a γ -lactame with 85 u and the formation of the [1,3]oxazepane-2-thione derivative (see structure **4** in Scheme 2). The MALDI mass spectrum additionally exhibits two abundant





Figure 4. (A) MALDI-TOF mass spectrum of a nano-HPLC fraction containing a lysozyme peptide (amino acids 115-125, [M + BuTuBu + H]⁺ at m/z 1545.740) that is modified with the cross-linker. Please note that Cys-115 is carbamidomethylated. A signal at m/z 1460.661 corresponds to a species that is created after a CNL of 85 u from the species at m/z 1545.740. A signal at m/z 1471.762 could not be assigned so far. Also visible in the spectrum is a lysozyme peptide (amino acids 97-114) at m/z 2073.947. (B) CID fragment ion mass spectrum (MALDI-TOF/TOF-MS/MS) of the cross-linker-modified peptide (m/z 1545.740). (C) ISD-MALDI fragment ion mass spectrum of [M + BuTu + H]⁺ at m/z 1460.661. (D) Nano-ESI-LTQ-Orbitrap mass spectrum of a nano-HPLC fraction containing a cross-linker-modified lysozyme peptide (amino acids 115-125, [M + BuTuBu + 2H]²⁺ at m/z 773.369). (E) CID fragment ion mass spectrum (nano-ESI-LTQ-Orbitrap-MS/MS) of the signal at m/z 773.369, fragmentation in the LTQ, analysis of fragment ions in the orbitrap. The isolated and fragmented species are represented in the insets. (F) CID fragment ion mass spectrum (MS³) of the signal at m/z 764.36. The isolated and fragmented species are represented in the insets; \$: signal is not assigned.

ions at m/z 2073.947 and 1471.762. The signal at m/z 2073.947 was assigned as lysozyme peptide composed of amino acids 97–114 based on MS/MS experiments. The ion at m/z 1471.762 unambiguously results from an ISD fragmentation of the cross-linked precursor ion at m/z 1545.740 as the amino acid sequence of the ion at m/z 1471.762 was elucidated to be identical to that of the species at m/z 1545.7 and confirmed by additional MS/MS experiments (data not shown). The fragmentation mechanism and the mass difference of 74 u, formally corresponding to the loss of SOCN, are still a matter of debate.

In the following, different types of cross-linked products, namely modified molecular ions mass shifted by 212 u (either intramolecularly cross-linked (type 1) peptides (structure **1** in Scheme 2) or peptides modified by the intramolecularly cyclized sevenmembered 2-alkylylimino-[1,3]thiazepan-7-one (type 0 cross-link); structure **3** in Scheme 2) and type 2 (interpeptide) cross-links^[24] were analyzed both in MALDI and ESI-CID MS/MS experiments for gaining insight into the specific fragmentation patterns for each type of cross-linked product. In Fig. 4B, the product ion spectrum (MALDI-TOF/TOF-MS/MS) of the molecular ion with the character-



Figure 4. (continued).

istic mass shift of 212 u at *m/z* 1545.740 is presented (structure 3). Due to the presence of the C-terminal arginine, exclusively y-type ions are visible in the fragment ion mass spectrum, thus unambiguously demonstrating the modification at lysine-116. By far the most prominent signal in the MALDI-CID tandem mass spectrum is the signal at m/z 1460.6, i.e. the 7-alkylimino-[1,3]oxazepane-2-thione (4) that had been created by a CNL of 85 u from the precursor ion (3) (Fig. 4B). Additionally conducted MALDI-ISD MS/MS experiments of the latter species yielded a complete series of y-type ions and an almost complete b-type ion series (Fig. 4C). From the mass difference between the y_9 and y_{10} ions, the modification was unambiguously assigned to the lysine in position 116. As for substance P, ESI-LTQ-Orbitrap-MS, MS/MS and MS³ data were obtained for tryptic peptides derived from the digestion of lysozyme. The respective spectra are shown in Fig. 4D-F. The peptide comprising amino acids 115-125 that is modified with a cross-linker molecule is visible as doubly charged signal at m/z 773.369 in the nano-ESI-LTQ-Orbitrap mass spectrum (Fig. 4D). MS/MS and MS³ data confirm the modification with the cross-linker at lysine-116 (Fig. 4E and F).

Finally, an interpeptide (type 2) cross-linked product^[24] within lysozyme connecting the N-terminal lysine with lysine-97 was investigated. According to the crystal structure of hen egg lysozyme (pdb entry 193L) both N-atoms are at a distance of 23.2 Å and can readily be cross-linked by the thiourea linker that possesses a length of 12.5 Å due to the inherent flexibility of the lysine side chains. In previously conducted cross-linking studies, we found both lysines to be potential reaction sites of NHS ester-based cross-linkers.^[27] The MALDI-TOF mass spectrum shows the crosslinked product at m/z 2144.023 (Fig. 5A). MALDI-TOF/TOF-MS/MS analyses of this species revealed several y-type ions (Fig. 5B), whereas ESI-LTQ-Orbitrap-MS/MS also yielded a number of b-type ions (Fig. 5C). The interpeptide cross-link was thus unambiguously identified based on a number of fragment ions of the longer lysozyme peptide constituting the cross-linked product. The fact that the connected lysine represents only a single amino acid is readily explained as it represents the N-terminal amino acid of lysozyme. All other lysines in the amino acid sequence of lysozyme would not have been created as single amino acid after tryptic cleavage. In order for trypsin to cleave C-terminal to Lys-1, the





Figure 4. (continued).

cross-linking reaction must have occurred at the N-terminal amine group instead of the ε -amine group of the lysine side chain – a behavior that we frequently observe when cross-linking lysozyme.

Cross-linking of PPARα

As for lysozyme, MALDI- and ESI-MS and CID-MS/MS experiments were conducted for the 32-kDa ligand binding domain PPARa.^[11] From the b- and y-type ions of a PPAR α peptide (amino acids 213-222) in CID-MS/MS, a cross-linker modification was assigned to the lysine at position 216. The respective mass spectra are shown in Fig. S2 (Supporting Information).

Conclusions and Outlook

The novel thiourea cross-linker described herein is a simplified analog of our previous reagent incorporating a thiourea-Gly-Pro structure.^[9] This cross-linker exhibits a central thiourea moiety, which is flanked by amine-reactive NHS ester moieties for reaction with lysines or free N-termini in proteins. The activated ester moieties exhibit a high reactivity with peptides (substance P, LHRH) and proteins (hen egg lysozyme and PPAR α). The applicability of the thiourea compound as dissociative linker relies solely on the nucleophilicity of the thiocarbonyl sulfur, which is responsible for an efficient and preferred dissociation upon CID. The fragmentation behavior of the novel compound was probed under low-energy CID conditions (kinetic energy of the precursor ion <100 eV) in a linear quadrupole ion trap as well as under high-energy CID conditions (kinetic energy of the precursor ion in the range of keV) in a TOF/TOF system with respectively modified peptide ions. The results clearly confirm the feasibility of the proposed analytical concept as the amide bond adjacent to the thiocarbonyl sulfur is more easily cleaved than any other peptide backbone bond leading to the loss of a γ -lactame (85 u) and the formation of a 7-alkylimino-[1,3]oxazepane-2-thione (structure 4 in Scheme 2). On the basis of this characteristic 85 u-CNL, a rapid identification of peptide ions that are modified with the derivatizing agent is realized. However, our present investigations reveal that the fragmentation behavior of the thiourea compound is not unambiguous (Scheme 2), which hampers a general application of this particular reagent as a chemical cross-linker for protein structural studies. Current investigations with a number of thiourea compounds are underway to circumvent the apparent challenges that are inherently associated with the reagent described herein.



Figure 5. (A) MALDI-TOF mass spectrum of nano-HPLC fraction containing a lysozyme interpeptide (type 2) cross-link composed of the *N*-terminal amino acid connected to amino acids 97–112, in which Lys-1 is cross-linked via its *N*-terminus with Lys-97 ($[M + H]^+$ at *m/z* 2144.023). Other lysozyme peptides are also visible in the spectrum. (B) CID fragment ion mass spectrum (MALDI-TOF/TOF-MS/MS) of the cross-linked product (*m/z* 2144.023). (C) Nano-ESI-LTQ-Orbitrap-MS/MS data of the cross-linked product $[M + 3H]^{3+}$ at *m/z* 715.68. The isolated and fragmented species are represented in the insets.



Acknowledgements

MQM is supported by the DFG Graduiertenkolleg 1026 'Conformational Transitions in Macromolecular Interactions' at the Martin-Luther-Universität Halle-Wittenberg. AS gratefully acknowledges financial support by the DFG and the BMBF (ProNet-T³ project), MS and FD gratefully acknowledge financial support by the DFG.

Supporting information

Supporting information may be found in the online version of this article.

References

- A. Sinz. Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. J. Mass Spectrom. 2003, 38, 1225.
- [2] M. A. Trakselis, S. C. Alley, F. T. Ishmael. Identification and mapping of protein–protein interactions by a combination of cross-linking, cleavage and proteomics. *Bioconj. Chem.* 2005, *16*, 741.
- [3] A. Sinz. Chemical cross-linking and mass spectrometry to map threedimensional protein structures and protein-protein interactions. *Mass Spectrom. Rev.* 2006, 25, 660.
- [4] M. M. Young, N. Tang, J. C. Hempel, C. M. Oshiro, E. W. Taylor, I. D. Kuntz, B. W. Gibson, G. Dollinger. High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. *Proc. Nat. Acad. Sci. U.S.A.* 2000, *97*, 5802.
- [5] J. W. Back, A. F. Hartog, H. L. Dekker, A. O. Muijsers, L. J. de Koning, L. D. Jong. A new crosslinker for mass spectrometric analysis of the quaternary structure of protein complexes. J. Am. Soc. Mass Spectrom. 2001, 12, 222.
- [6] E. J. Soderblom, M. B. Goshe. Collision-induced dissociative chemical cross-linking reagents and methodology: applications to protein structural characterization using tandem mass spectrometry analysis. *Anal. Chem.* **2006**, *78*, 8059.
- [7] E. J. Soderblom, B. G. Bobay, J. Cavanagh, M. B. Goshe. Tandem mass spectrometry acquisition approaches to enhance identification of protein–protein interactions using low-energy collision-induced dissociative chemical crosslinking reagents. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 3395.
- [8] E. V. Petrotchenko, V. K. Olkhovik, C. H. Borchers. Isotopically coded cleavable cross-linker for studying protein – protein interaction and protein complexes. *Mol. Cell. Proteomics* 2005, *4*, 1167.
- [9] F. Dreiocker, M. Q. Müller, A. Sinz, M. Schäfer. Collision-induced dissociative chemical cross-linking reagent for protein structure characterization: applied Edman chemistry in the gas phase. J. Mass Spectrom. 2010, 45, 178.
- [10] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989**, *246*, 64.

- [11] M. Karas, F. Hillenkamp. Laser desorption ionization of proteins with molecular masses exceeding 10,000 Da. Anal. Chem. 1988, 60, 2299.
- [12] M. Q. Müller, C. Roth, N. Sträter, A. Sinz. Expression and purification of the ligand-binding domain of peroxisome proliferator-activated receptor alpha. *Prot. Exp. Purif.* **2008**, *62*, 185.
- [13] S. E. Holroyd, P. Groves, M. S. Searle, U. Gerhard, D. H. Williams. Rational design and binding of modified cell-wall peptides to vancomycin-group antibiotics: factorising free energy contributions to binding. *Tetrahedron* **1993**, *49*, 9171.
- [14] D. Hoppe, H. Follmann. 2-Thioxo-oxazolidine durch cycloaddition von α -metallierten alkylisothiocyanaten an carbonylverbindungen. *Chem. Ber.* **1976**, *109*, 3047.
- [15] P. Langer, A. Bodtke. Sequential cyclizations of 2isothiocyanatobenzonitrile with α -aminocarboxylic esters and acids. *Synlett* **2003**, *11*, 1743.
- [16] S. Reyes, K. Burgess. On formation of thiohydantoins from amino acids under acylation conditions. J. Org. Chem. 2006, 71, 2507.
- [17] A. B. Smith, S. S.-Y. Chen, F. C. Nelson, J. M. Reichert, B. A. Salvatore. Total syntheses of (+)-acutiphycin and (+)-trans-20,21didehydroacutiphycin. J. Am. Chem. Soc. 1997, 119, 10935.
- [18] T. S. Rao, S. Nampalli, P. Sekher, S. Kumar. TFA-NHS as bifunctional protecting agent: simultaneous protection and activation of amino carboxylic acids. *Tetrahedron Lett.* **2002**, *43*, 7793.
- [19] M. Adamczyk, Y. Y. Chen, J. C. Gebler, D. D. Johnson, P. G. Mattingly, J. A. Moore, R. E. Reddy, J. Wu, Z. Yu. Evaluation of chemiluminescent estradiol conjugates by using a surface plasmon resonance detector. *Steroids* **2000**, *65*, 295.
- [20] D. M. Schulz, S. Kalkhof, A. Schmidt, C. Ihling, C. Stingl, K. Mechtler, O. Zschörnig, A. Sinz. Annexin A2/p11 interaction: new insights into annexin A2 tetramer structure by chemical cross-linking, high-resolution mass spectrometry, and computational modeling. *Proteins* **2007**, *69*, 254.
- [21] S. Peri, H. Steen, A. Pandey. GPMAW a software tool for analyzing proteins and peptides. *Trends Biochem. Sci.* **2001**, *26*, 687.
- [22] L. J. de Koning, P. T. Kasper, J. W. Back, M. A. Nessen, F. Vanrobaeys, J. van Beeumen, E. Gherardi, C. G. de Koster, L. de Jong. Computerassisted mass spectrometric analysis of naturally occurring and artificially introduced cross-links in proteins and protein complexes. *FEBS J.* **2006**, 273, 281.
- [23] http://www2.lib.udel.edu/database/intermed/scifind/viewer.html; Accelrys ViewerLite 5.0 Software.
- [24] B. Schilling, R. H. Row, B. W. Gibson, X. Guo, M. M. Young. MS² assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides. J. Am. Soc. Mass Spectrom. 2003, 14, 834.
- [25] S. Kalkhof, A. Sinz. Chances and pitfalls of chemical cross-linking with amine-reactive *N*-hydroxy succinimide esters. *Anal. Bioanal. Chem.* 2008, 392, 305.
- [26] S. Mädler, C. Bich, D. Touboul, R. Zenobi. Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. J. Mass Spectrom. 2009, 44, 694.
- [27] G. H. Dihazi, A. Sinz. Mapping low-resolution three-dimensional protein structures using chemical cross-linking and Fourier transform ion-cyclotron resonance mass spectrometry. *Rapid Commun. Mass Spectrom.* 2003, 17, 2005.