

Novel Amidinating Cross-Linker for Facilitating Analyses of Protein Structures and Interactions

Matthew A. Lauber and James P. Reilly*

Department of Chemistry, Indiana University, Bloomington, Indiana 47405

A novel bifunctional thioimide cross-linking reagent (diethyl suberthioimide) that modifies amines without sacrificing their native basicity is developed. Intermolecular cross-linking of neurotensin and intramolecular cross-linking of cytochrome c under physiological conditions is investigated with this reagent. Because it does not perturb the electrostatic properties of a protein, it is unlikely to lead to artifactual conclusions about native protein structure. The interpeptide cross-links formed with this reagent are easily separated from other tryptic fragments using strong cation exchange chromatography, and they have a readily identified mass spectrometric signature. The use of this novel amidinating protein cross-linking reagent holds great promise for efficient, large-scale structural analysis of complex systems.

Chemical cross-linking has long been used as a tool for elucidating protein structures and interactions.^{1–8} In most cross-linking analyses, protein functional groups are targeted for derivatization with a molecule that contains two reactive groups separated by a spacer arm of known length. Functional groups within the approximate distance of the spacer arm are therefore capable of reacting to form cross-links, and identification of the cross-linked species can accordingly lead to the mapping of protein structure or protein–protein interactions.

Advancing technologies in mass spectrometry, including improvements in resolution and mass accuracy, have led to a resurgent interest in this method of studying protein structures and interactions, because it is now possible to detect peptide cross-links, formed from proteolysis of derivatized protein samples, with sensitivity and high confidence.^{9,10} Recent mass spectrometry-based investigations of cross-linked proteins have improved our

understanding of systems not amenable to other methods of structural analysis.^{11–13}

Despite the potential of cross-linking studies, analytical challenges hinder what is usually accomplished. In particular, the detection of cross-linked peptides in the proteolytic digests of derivatized proteins is often impaired by the combination of their low stoichiometric yield and the presence of other peptide species. For example, in addition to the desired cross-linking products, unmodified peptides and peptides modified with partially hydrolyzed reagent (so-called “dead-ends”) are abundantly present. To address this issue, enrichment of reagent-modified peptides by affinity purification of biotin tagged reagents^{14–16} and reaction-based selective retrieval of a specially designed reagent have been performed.¹⁷ Although effective at removing unmodified peptides from samples, these techniques intrinsically enrich any reaction products containing the reagent, including dead-ends. Development of a simple and efficient technique for enriching cross-linked peptides from all components of proteolytic digests, including dead-end modified peptides, is therefore still very desirable.

Another cross-linking challenge involves ensuring that the reaction of a particular covalent probe does not have a deleterious effect on the stability of native protein structure. Amine-reactive succinimidyl esters, despite being popular for their reaction efficiency, modify basic primary amines to yield neutral amides. Modification with succinimidyl esters therefore eliminates the positive charge amines exhibit at physiological pH, perturbs the surface charge distribution of a protein, and can potentially result in disturbing protein tertiary and quaternary structure. Interestingly, an alternative functional group, known as an imide, has previously been utilized for protein modification.¹⁸ The imide functional group is highly selective for primary amines, and its reaction yields amidinated residues with pKa values about 2 units

* To whom correspondence should be addressed. E-mail: reilly@indiana.edu.

- (1) Das, M.; Fox, C. F. *Annu. Rev. Biophys. Bioeng.* **1979**, *8*, 165–93.
- (2) Singh, P.; Panchaud, A.; Goodlett, D. R. *Anal. Chem.* **2010**, *82* (7), 2636–2642.
- (3) Novak, P.; Giannakopoulos, A. E. *Eur. J. Mass Spectrom.* **2007**, *13* (2), 105–113.
- (4) Kluger, R.; Alagic, A. *Bioorg. Chem.* **2004**, *32* (6), 451–472.
- (5) Peters, K.; Richards, F. M. *Annu. Rev. Biochem.* **1977**, *46*, 523–51.
- (6) Vasilescu, J.; Figeys, D. *Curr. Opin. Biotechnol.* **2006**, *17* (4), 394–9.
- (7) Trakselis, M. A.; Alley, S. C.; Ishmael, F. T. *Bioconjugate Chem.* **2005**, *16* (4), 741–750.
- (8) Loster, K.; Josic, D. *J. Chromatogr. B Biomed. Sci. Appl.* **1997**, *699* (1–2), 439–61.
- (9) Sinz, A. *Mass Spectrom. Rev.* **2006**, *25* (4), 663–82.
- (10) Young, M. M.; Tang, N.; Hempel, J. C.; Oshiro, C. M.; Taylor, E. W.; Kuntz, I. D.; Gibson, B. W.; Dollinger, G. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97* (11), 5802–6.

- (11) Forwood, J. K.; Thakur, A. S.; Guncar, G.; Marfori, M.; Mouradov, D.; Meng, W.; Robinson, J.; Huber, T.; Kellie, S.; Martin, J. L.; Hume, D. A.; Kobe, B. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (25), 10382–7.
- (12) Lanman, J.; Lam TuKiet, T.; Barnes, S.; Sakalian, M.; Emmett, M. R.; Marshall, A. G.; Prevelige, P. E., Jr. *J. Mol. Biol.* **2003**, *325* (4), 759–72.
- (13) Back, J. W.; Sanz, M. A.; De Jong, L.; De Koning, L. J.; Nijtmans, L. G. J.; De Koster, C. G.; Grivell, L. A.; Van Der Spek, H.; Muijsers, A. O. *Protein Sci.* **2002**, *11* (10), 2471–2478.
- (14) Chowdhury, S. M.; Du, X.; Tolic, N.; Wu, S.; Moore, R. J.; Mayer, M. U.; Smith, R. D.; Adkins, J. N. *Anal. Chem.* **2009**, *81* (13), 5524–32.
- (15) Tang, X.; Munske, G. R.; Siems, W. F.; Bruce, J. E. *Anal. Chem.* **2005**, *77* (1), 311–318.
- (16) Trester-Zedlitz, M.; Kamada, K.; Burley, S. K.; Fenyoe, D.; Chait, B. T.; Muir, T. W. *J. Am. Chem. Soc.* **2003**, *125* (9), 2416–2425.
- (17) Yan, F.; Che, F.-Y.; Rykunov, D.; Nieves, E.; Fiser, A.; Weiss, L. M.; Hogue Angeletti, R. *Anal. Chem.* **2009**, *81* (17), 7149–7159.
- (18) Hunter, M. J.; Ludwig, M. L. *Methods Enzymol.* **1972**, *25* (Pt. B), 585–96.

higher than the already basic primary amines.¹⁹ In contrast to succinimidyl ester modification, amidination thus results in essentially no perturbation of a protein's electrostatic properties because, at physiological pH, an amidinated residue will be protonated and positively charged, just like a primary amine.¹⁸ Unfortunately, imidates are prone to extensive side reactions and consequently make poor cross-linking reagents when reaction conditions are below pH 10.^{20–22}

The problem with these imidates arises from the leaving group being an alkyl alcohol. When the pH of a reaction is below 9, protonation and subsequent elimination of ammonia predominates, leading to undesirable reaction products. These are outlined in Supporting Information Scheme 1. The use of imidate cross-linking reagents has consequently not been favored, even though they have desirable characteristics, because efficient cross-linking can only be achieved under the alkaline conditions in which many proteins denature.^{1,9} Interestingly, thioimidates have been employed as alternatives to imidates for protein modification. Pfeleiderer and co-workers introduced the thio-analog alternative of the chemical labeling reagent O-methylacetimidate, S-methylacetimidate (SMTA), and demonstrated that by replacing the alcohol leaving group of an imidate with a better one, namely a thiol, a more useful reagent could be obtained, specifically a reagent with no observed side reaction products across the entire physiological pH regime.²³ We have employed this molecule and the analogous S-methyl thiopropionimidate (SMTP) in a number of structural studies of protein and ribosome complexes with mass spectrometric analysis of reaction products.^{24–29} On the basis of this experience, it was apparent to us that a cross-linking reagent with thioimide reactive groups would be better suited to probing native protein structure than commonly used and commercially available reagents. We also suspected that such a reagent would be useful for overcoming the aforementioned detectability issue, as the cross-links it would form would contain a significant number of basic residues and thus be amenable to enrichment by strong cation exchange chromatography.

We have synthesized a novel homobifunctional thioimide cross-linking reagent, diethylsuberthioimide (DEST), that contains thiol leaving groups and effectively modifies proteins via amidination under physiological conditions. This makes DEST a compelling alternative to both imidate and succinimidyl ester reagents for structural studies, since it does not require alkaline pH for reactivity, does not perturb the electrostatic properties of

a protein, and is thus less likely to lead to artifactual conclusions about native protein structure. We also demonstrate that because this reagent preserves the basicity of targeted amines upon modification, the cross-links it forms can be easily separated from other components of tryptic digests using strong cation exchange (SCX) chromatography.

EXPERIMENTAL SECTION

Materials. Acetonitrile (ACN), hydrochloric acid, sodium hydroxide, trifluoroacetic acid (TFA), and water were purchased from EMD Chemicals (Gibbstown, NJ). Anhydrous diethyl ether was obtained from Fisher (Fair Lawn, NJ). Bovine pancreas trypsinogen (T-1143), horse heart cytochrome c (C-7752), neurotensin (N-6383), proteomics grade trypsin (T-6567), and Trizma base were purchased from Sigma (St. Louis, MO). Formic acid (FA) and suberionitrile were obtained from Aldrich (Milwaukee, WI). Calcium chloride dihydrate, dichloromethane, sodium chloride, sodium phosphate monobasic, and type 3A molecular sieves were purchased from Mallinckrodt Baker (Phillipsburg, NJ). Dimethyl suberimidate (DMS) was purchased from Pierce (Rockford, IL). Ethanethiol was obtained from Acros Organics (Pittsburgh, PA). Hydrogen chloride (technical grade) was purchased from Matheson (Cucamonga, CA).

Synthesis of Diethyl Suberthioimide (DEST). Diethyl suberthioimide (DEST) was prepared from ethanethiol and suberionitrile via the Pinner synthesis (Supporting Information Scheme 2).³⁰ Suberionitrile (9 mmol) in anhydrous dichloromethane (1:3 v/v) was added to ice-cold ethanethiol (90 mmol). The reaction mixture, under constant stirring, was sparged with hydrogen chloride gas for one hour and subsequently kept at 4 °C for an additional 16 h. Anhydrous diethyl ether was then added to aid precipitation, and the mixture was stored at –20 °C until a solid had formed. This solid was washed several times with anhydrous diethyl ether and stored in a vacuum desiccator at room temperature. The structure of the synthesis product was confirmed to be diethyl suberthioimide by ¹H NMR using a 400 MHz Varian Inova NMR spectrometer (Varian NMR, Palo Alto, CA). ¹H NMR (CDCl₃): δ 1.44 (t, 6H), 1.50 (t, 4H), 1.86 (m, 4H), 2.90 (t, 4H), 3.51 (q, 4H).

Protein Cross-Linking. Horse heart cytochrome c was reacted with either DEST or DMS. Reactions were carried out in 20 mM sodium phosphate/150 mM sodium chloride (pH 7) with the protein at a final concentration of 5 μM and reagent to protein molar ratios of 20:1, 100:1, 200:1, 500:1, or 1000:1. After proceeding at room temperature for 12 h, reactions were quenched by adding 0.5 M Tris to a final concentration of 50 mM. After another 12 h, the resulting reaction mixtures were desalted and cleared of hydrolyzed reagent using Microcon YM-10 centrifugal filter devices (Millipore, Eschborn, Germany). The concentrate was then dried and stored at –20 °C until later analysis or sample preparation. Cytochrome c modified by this method was checked for interprotein cross-linking using size exclusion chromatography (See Supporting Information Figure 1).

LC-MS of Cross-Linked Cytochrome C. Samples of intramolecularly cross-linked cytochrome c were analyzed by LC-MS using a 254 μm i.d. PEEK capillary column packed with C4 silica beads (5 μm, 300 Å, Phenomenex, Torrance, CA) and a Waters 2795

(19) Hermanson, G. T.; Ed., *Bioconjugate Techniques*; Academic Press: San Diego, 1995.

(20) Browne, D. T.; Kent, S. B. H. *Biochem. Biophys. Res. Commun.* **1975**, *67* (1), 133–8.

(21) Browne, D. T.; Kent, S. B. H. *Biochem. Biophys. Res. Commun.* **1975**, *67* (1), 126–32.

(22) Dihazi, G. H.; Sinz, A. *Rapid Commun. Mass Spectrom.* **2003**, *17* (17), 2005–2014.

(23) Thumm, M.; Hoeses, J.; Pfeleiderer, G. *Biochim. Biophys. Acta* **1987**, *923* (2), 263–7.

(24) Running, W. E.; Reilly, J. P. *J. Proteome Res.* **2009**, *8* (3), 1228–1246.

(25) Liu, X.; Reilly, J. P. *J. Proteome Res.* **2009**, *8* (10), 4466–4478.

(26) Lauber, M. A.; Running, W. E.; Reilly, J. P. *J. Proteome Res.* **2009**, *8* (9), 4193–4206.

(27) Liu, X.; Broshears, W. C.; Reilly, J. P. *Anal. Biochem.* **2007**, *367* (1), 13–19.

(28) Beardsley, R. L.; Running, W. E.; Reilly, J. P. *J. Proteome Res.* **2006**, *5* (11), 2935–2946.

(29) Janecki, D. J.; Beardsley, R. L.; Reilly, J. P. *Anal. Chem.* **2005**, *77* (22), 7274–7281.

(30) Roger, R.; Neilson, D. G. *Chem. Rev.* **1961**, *61*, 179–211.

liquid chromatograph (Waters, Milford, MA). Reversed-phase effluent resulting from separation of 300 pmol of protein sample was infused into the ESI source of a quadrupole time-of-flight mass spectrometer (Q-TOF, Waters, Manchester, UK) at a flow rate of 5 $\mu\text{L}/\text{min}$. In all experiments, the voltage applied to the ESI needle was +3.0 kV. Mass spectra (600–2500 m/z) were acquired over the time interval corresponding to a 25 min gradient between 0.1% TFA/0.2% FA in 80:20 water/ACN and 0.1% TFA/0.2% FA in 30:70 water/ACN. Whole protein masses were obtained by summing together raw spectra corresponding to chromatographic peaks and deconvolution using MassLynx and MaxEnt 1.

Proteolytic Digestion of Cross-Linked Cytochrome c.

Cross-linked cytochrome c (25 μg) and proteomics grade trypsin (1.25 μg) were reconstituted in solution, such that the digestion was carried out in 100 mM Tris and 10 mM calcium chloride (pH 8) at a 20:1 (w/w) ratio of substrate to protease. Each digest reaction was allowed to proceed at 37 °C for 24 h and subsequently quenched by adding 2% TFA to a final concentration of 0.5%.

SCX Enrichment of DEST Interpeptide Cross-Links. Interpeptide cross-links were enriched by strong cation exchange (SCX) chromatography by exploiting the fact that they contain more basic residues and thus more positive charges at low pH than non-cross-linked species. Tryptic digest of cross-linked cytochrome c (5 μg) was loaded on an SCX column (TSKgel SP-NPR, 4.6 \times 35 mm, Tosoh Bioscience, Montgomeryville, PA) using 0.1% TFA in water as mobile phase and a flow rate of 0.3 mL/min. Non-cross-linked species were removed from the adsorbed sample in ten column volumes of mobile phase containing 300 mM NaCl. Interpeptide cross-links were then eluted from the SCX column onto a C18 trapping column (Thermo Hypersil-Keystone Javelin, 1.0 \times 20 mm, Bellefonte, PA) in ten column volumes of mobile phase containing 1000 mM NaCl. After it was desalted with mobile phase A at 0.3 mL/min for 10 min, the contents of the C18 trapping column were eluted in a 5 min gradient (flow rate = 50 $\mu\text{L}/\text{min}$) from 0.1% FA in water to 0.1% FA in ACN and 10 min isocratic hold. This fraction enriched for interpeptide DEST cross-links was dried under vacuum and resuspended in 25 μL of mobile phase A for subsequent LC-MS/MS analysis.

Capillary LC-ESI-MS/MS of Proteolyzed Cross-Linked Cytochrome c. Capillary LC-ESI-MS/MS of peptides was performed using an IntegraFrit capillary trapping column packed with 1.5 cm of C18 (150 μm \times 11 cm, New Objective, Woburn, MA; Magic C18, 5 μm , 200 Å, Michrom BioResources, Auburn, CA), a capillary analytical column packed with 15 cm of C18 (75 μm \times 15 cm, Magic C18, 5 μm , 100 Å, Michrom BioResources, Auburn, CA), an LTQ Orbitrap mass spectrometer (Thermo Electron, Bremen, Germany), and a Dionex chromatography system (Ultimate 3000, Dionex, Sunnyvale, CA). In each experiment, approximately 1 μg of protein digest was injected onto a trapping column to remove salts and contaminants by flushing for 10 min with mobile phase A (0.1% FA in 97:3 water/ACN) at a flow rate of 10 $\mu\text{L}/\text{min}$. The flow rate was then reduced to 0.3 $\mu\text{L}/\text{min}$, effluent from the trapping column was directed to the capillary LC column, and a 75-min gradient between 0 and 40% mobile phase B (0.1% FA in ACN) was implemented. Eluting peptides were electrosprayed into a Thermo LTQ-Orbitrap mass spectrometer operating in data-dependent mode to acquire a full MS scan (300–2000 m/z) and subsequent CID MS/MS scans of the five

most intense precursor ions in the LTQ at 35% normalized collision energy. The resolution for the MS scan was set to 30 000. Dynamic exclusion was employed with the following settings: a 45 s exclusion duration time, maximum exclusion list of 500 with one repeat count, and repeat duration of 15 s. In addition, charge state rejection was enabled for 1+ charge states. MS and MS/MS spectra were subjected to data reduction using Mascot Distiller, such that precursor masses were redetermined via interpretation of isotopic distributions and MS/MS spectra were deisotoped. Processed data were subsequently searched against the sequence of horse heart cytochrome c for interpeptide cross-links using the web-interface of xQuest.³¹ Search parameters and details on the determination of a false discovery rate for our analysis can be found in Supporting Information.

Cytochrome c Crystal Structure. Intramolecular distances between alpha carbons of lysine residues in crystalline horse heart cytochrome c were determined using the measurement feature of PyMOL, version 0.99 (DeLano Scientific, www.pymol.org) and the crystal structure of oxidized horse heart cytochrome c (PDB 1HRC).^{32,33}

RESULTS AND DISCUSSION

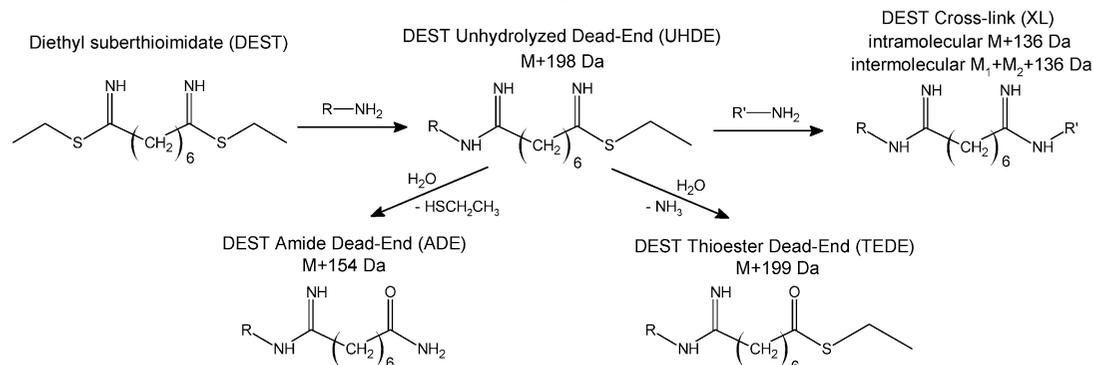
Development of DEST. A novel homobifunctional thioimide cross-linking reagent has been readily synthesized from a simple one-pot reaction, in which suberonitrile is added to a solution of ethanethiol saturated with anhydrous hydrogen chloride. Suberonitrile was chosen as a starting material, because it provides a bifunctional thioimide with an 11 Å, or 8-atom, spacer arm. This length is intermediate relative to other cross-linkers, meaning that the reagent should be useful for studying both protein tertiary and quaternary structures and that it should provide a sufficient number of cross-links for analysis. In developing the linker, we also sought to have a small leaving group to minimize steric bulk. Unfortunately, the smallest thiol, methanethiol, is gaseous at room temperature, making it more difficult to handle. Ethanethiol, a liquid at room temperature, was therefore chosen as the nucleophile for the synthesis, establishing the leaving group of the synthesized reagent. The water-soluble compound synthesized from these starting materials is diethyl suberthioimide (DEST). It is the analog of the commercially available cross-linking reagent, dimethyl suberimidate (DMS), in targeting primary amines and preserving their positive charge through formation of amidine linkages. Based on similar chemistry, DEST, like DMS, was expected to form two types of products, cross-links and dead-ends (Scheme 1). Cross-links (XL) between two amines introduce 136 Da mass shifts. If it occurs intramolecularly, the product has a mass equal to the mass of the targeted molecule plus the mass of the cross-link ($M + 136$). If it forms intermolecularly, the product's mass is the sum of the masses of the two linked molecules plus the mass of the cross-link ($M_1 + M_2 + 136$). Dead-ends, which form when one end of the reagent reacts with an amine while the other end is hydrolyzed, lead to mass shifts of either 154 or 199 Da, depending on whether hydrolysis

(31) Rinner, O.; Seebacher, J.; Walzthoeni, T.; Mueller Lukas, N.; Beck, M.; Schmidt, A.; Mueller, M.; Aebersold, R. *Nat. Methods* **2008**, *5* (4), 315–8.

(32) DeLano, W. L. *The PyMOL Molecular Graphics System*; DeLano Scientific: Palo Alto, CA, 2002.

(33) Bushnell, G. W.; Louie, G. V.; Brayer, G. D. *J. Mol. Biol.* **1990**, *214* (2), 585–95.

Scheme 1. Major Reaction Products of DEST Cross-Linking



involved elimination of the thiol and formation of an amide (ADE) or elimination of ammonia and formation of a thioester (TEDE). These possibilities have been confirmed by model peptide experiments that are fully described in Supporting Information.

Intramolecular Protein Cross-Linking with DEST. Cytochrome c is an N-terminally acetylated small protein consisting of 104 amino acid residues. Crystal structure data suggest that every one of its 19 primary amines (Lys residues), in addition, is solvent accessible and thus modifiable. As such, it is an ideal substrate for probing the potentially complex array of cross-linking reaction products caused by modification with DEST. These data also indicate that cytochrome c contains 83 lysine pairs with inter α -carbon distances less than 20 Å.³⁴ Such pairs are well within the maximum cross-linking distance of an 11 Å spacer arm reagent, given the inherent flexibility of lysine side chains. As the lysine side chain is approximately 6.5 Å long, the maximum α -carbon to α -carbon cross-linking distance that DEST is capable of bridging is 24 Å. This suggests that DEST is too large to gain very useful structural information from intramolecular cross-linking of relatively small proteins, such as cytochrome c; however, it is the chemistry of the cross-linking reaction and the analysis of the products that is of interest in the work that follows.

As expected, LC-ESI-MS revealed DEST-modified cytochrome c to be a complex mixture of intramolecular cross-links (M + 136 Da), ADE (M + 154 Da), and TEDE (M + 199 Da) dead-ends. Figure 1 includes the deconvoluted ESI mass spectra for cytochrome c modified by DEST at 20:1, 100:1, and 200:1 reagent to protein molar ratios for 12 h. Upon modification at a 20:1 molar ratio, up to 2 cross-linking products were incorporated into cytochrome c, although unmodified cytochrome c remained the most abundant species. Modification at a 100:1 molar ratio resulted in a shift to 2–5 incorporated cross-linking products, with products ranging from 4 intramolecular cross-links to 2 amide dead-ends. Modification at a 200:1 molar ratio resulted in a shift to 5–7 incorporated cross-linking products. Considering that the protein has 19 modifiable primary amines, the largest number of intramolecular cross-links that could be observed is 9. This number would only be seen if all sites were accessible and not blocked by dead-ends. Thus, without any attempt to optimize the extent of cross-linking, we are observing more than half of the number possible. Not surprisingly, modification of cytochrome c with DEST was strikingly different than modification of cytochrome c with DMS.

Modification with the imidate analog at a 100:1 molar ratio yielded no incorporation of cross-linking products. Even at an increased molar ratio of 1000:1, there was no indication of intramolecular cross-linking and unmodified cytochrome c remained the most abundant species (Supporting Information Figure 2).

As evident from the neurotensin experiments in Supporting Information and the cytochrome c work described above, DEST is clearly more effective than DMS. Furthermore, DEST, to our knowledge, represents the first reagent capable of cross-linking amines in a single stage reaction at physiological pH without eliminating their native basicity.

Peptide Analysis of Cross-Linked Cytochrome c. The main objective of a mass spectrometry based cross-linking experiment is to extract the structural information encoded in peptides produced by proteolysis of a cross-linked protein or cross-linked protein complex. Consequently, we studied DEST-modified cytochrome c through peptide MS/MS analysis, to test the true utility of our reagent. Cytochrome c modified by DEST at a 100:1 ratio was chosen as the appropriate sample for this analysis because of its intermediate extent of modification, an important consideration because under-modification can lead to needlessly low abundances of cross-linked peptides while overmodification can minimize proteolysis, undesirably complicate the peptides generated, and potentially disturb protein structure.

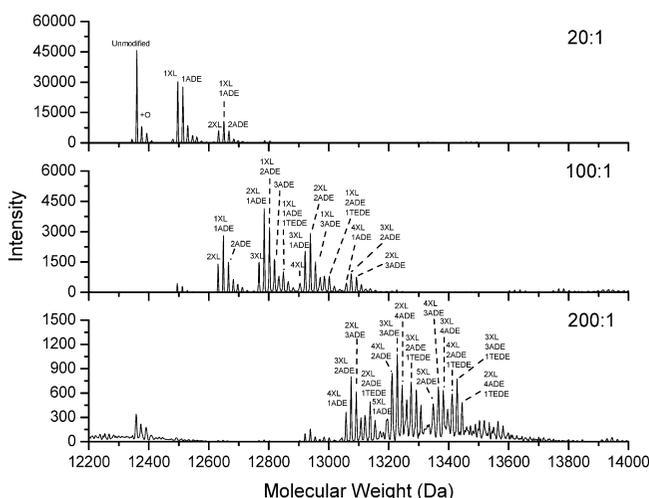


Figure 1. Deconvoluted ESI-QTOF mass spectra of DEST-modified cytochrome c. The products of three reactions involving differing molar ratios of reagent to protein were analyzed (20:1, 100:1, and 200:1). Reaction products are labeled as follows: amide dead-end (ADE), thioester dead-end (TEDE), and intramolecular cross-link (XL).

(34) Lee, Y. J.; Lackner, L. L.; Nunnari, J. M.; Phinney, B. S. *J. Proteome Res.* 2007, 6 (10), 3908–17.

Table 1. LC-MS/MS Analysis of Digested, DEST-Modified Cytochrome c

cross-link	distance (C_{α} – C_{α} , Å)	xQuest matches
K5–K13	13	4
K7–K39	23	1
K22–K100	12	1
K27–K73	23	1
K27–K79 ^a	13	4
K27–K86	24	1
K39–K55	8	3
K39–K60 ^a	10	17
K39–K72	17	2
K53–K73	15	5
K53–K79	12	1
K55–K72 ^a	12	3
K55–K73	12	6
K60–K86	19	2
K72–K86 ^a	14	9
K73–K86 ^a	13	5
K87–K99	19	2
K88–K99 ^a	17	10

^a Cross-links observed in all analyses.

Tryptic peptides generated from three replicate samples of cross-linked cytochrome c were analyzed by means of nanoLC-ESI-MS/MS with an LTQ-Orbitrap high resolution hybrid mass spectrometer. These experiments produced accurate precursor masses and fragmentation spectra that were analyzed by the recently developed search algorithm, xQuest.³¹ For our analysis, xQuest was used to search for theoretical cross-linked peptides with masses that matched measured precursor masses to within 5 ppm. In addition, searching was ensured to be of high confidence (false discovery rate less than 0.5%) by requiring that at least 15% of the fragment ion signal in MS/MS spectra match predicted b and y-type ions.

Our analysis of DEST-modified cytochrome c resulted in the assignment of 77 MS/MS spectra as corresponding to interpeptide cross-links and the identification of 18 unique cross-linked residue pairs. A summary of the identified cross-linked residues along with their corresponding number of xQuest matches is shown in Table 1, and detailed information of the observed cross-links can be found in Supporting Information Tables 1–3. It is noteworthy that certain pairs of residues appear much more often in the data than others. This can be the result of several effects, including the varying solvent accessibility and nucleophilicity of different amine sites and the varying ionization and detectability of different peptide sequences. The identified cross-linked residue pairs were completely consistent with the known crystal structure of cytochrome c,³³ the maximum expected cross-linking distance for an 8-atom, 11 Å spacer arm, and the potential flexibility of the targeted lysine side chain. In particular, the distances between all alpha carbons of unambiguously identified cross-linked residues were less than or equal to 24 Å, or the approximate distance of two lysine side chains and the cross-linker spacer arm. Interestingly, pairs detected via more than one xQuest match involved residues within 19 Å, and cross-links detected in all three LC-MS/MS runs involved residues within 17 Å (Supporting Information Figure 3.) This indicates, as would be expected, that the rate of cross-linking tends to increase with decreasing inter-residue distance.

The fact that DEST modification does not neutralize the basicity of targeted amines could in some cases be very advanta-

geous with respect to ionization efficiency. Unfortunately, we have not yet observed a significant effect in this respect, as the number of cross-links identified in our work is comparable to previous observations made by others while studying the cross-linking of cytochrome c with a succinimidyl ester analog of DEST (BS3).³⁴ The main reason for this is, likely, that trypsin has been employed as the protease in these studies. That is, cross-linked tryptic peptides inherently contain 4 strongly basic residues no matter the type of amine reactive cross-linker used, so the difference in ionization efficiency of DEST and BS3 cross-links analyzed in these studies is likely modest. This is supported by the fact that the average charge state of cross-links detected in this study was approximately 3.5, or about the same as the average charge state observed for cross-links identified in the previously mentioned BS3 cross-linking study.³⁴ It is likely that the higher charge states of DEST cross-links, $\geq 5+$, are not prominently observed because of a dramatic decrease in the gas phase basicity of multiply charged molecules³⁵ in this size regime (1500–4000 Da) as a result of charge repulsion. Nevertheless, compared with acylating reagents, such as succinimidyl esters, DEST may have a particular advantage in nontryptic digestion schemes in which some peptides may contain no other strongly basic residues than their N-termini. As discussed below, the higher basicity of DEST-linked peptides can be readily exploited in the liquid phase.

Fragmentation of DEST interpeptide cross-links was of interest, as there is potential for unique dissociation because of the localization of charge at the sites of cross-linking. Figure 2 shows two MS/MS spectra representative of the fragmentation observed for interpeptide DEST cross-links. Remarkably, there appears to be little difference between these spectra and those previously observed for analogous succinimidyl ester derived cross-links.^{17,34,36–38} The majority of ions, in particular, can be assigned as single cleavage fragments of the peptide backbones. There were nevertheless a few noteworthy observations. Ions involving C-terminal rearrangements ($b_x + 18$ ions), for example, were noticeably abundant in some spectra, such as Figure 2A. This phenomenon has been documented in MALDI MS/MS of succinimidyl ester cross-linked tryptic peptides,³⁹ indicating that formation of these product ions is favored when protons are localized away from the basic C-termini of tryptic peptides. As the 3+ charge state of the noted DEST cross-link cannot support simultaneous protonation of the highly basic amidine moieties and the basic C-terminal residues, competition for the ionizing protons would expectedly lead to an uncharged C-terminal residue and consequently favored C-terminal rearrangement. The fact that C-terminal rearrangement ions were significantly less intense in the MS/MS spectrum of the 4+ charge state of this DEST cross-link is consistent with C-terminal rearrangement being a charge deficient phenomenon (Figure 2B). Another type of noteworthy feature apparent in the spectra of some interpeptide DEST cross-links was the presence of low intensity ions related to fragmentation of the amidine cross-

(35) Cole, R. B. *J. Mass Spectrom.* **2000**, *35* (7), 763–72.

(36) Gardner, M. W.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (3), 344–357.

(37) Schilling, B.; Row Richard, H.; Gibson Bradford, W.; Guo, X.; Young, M. M. *J. Am. Soc. Mass Spectrom.* **2003**, *14* (8), 834–50.

(38) Gaucher, S. P.; Hadi, M. Z.; Young, M. M. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (3), 395–405.

(39) King, G.; Jones, A.; Kobe, B.; Huber, T.; Mouradov, D.; Hume, D.; Ross, I. *Anal. Chem.* **2008**, *80*, 5036–5043.

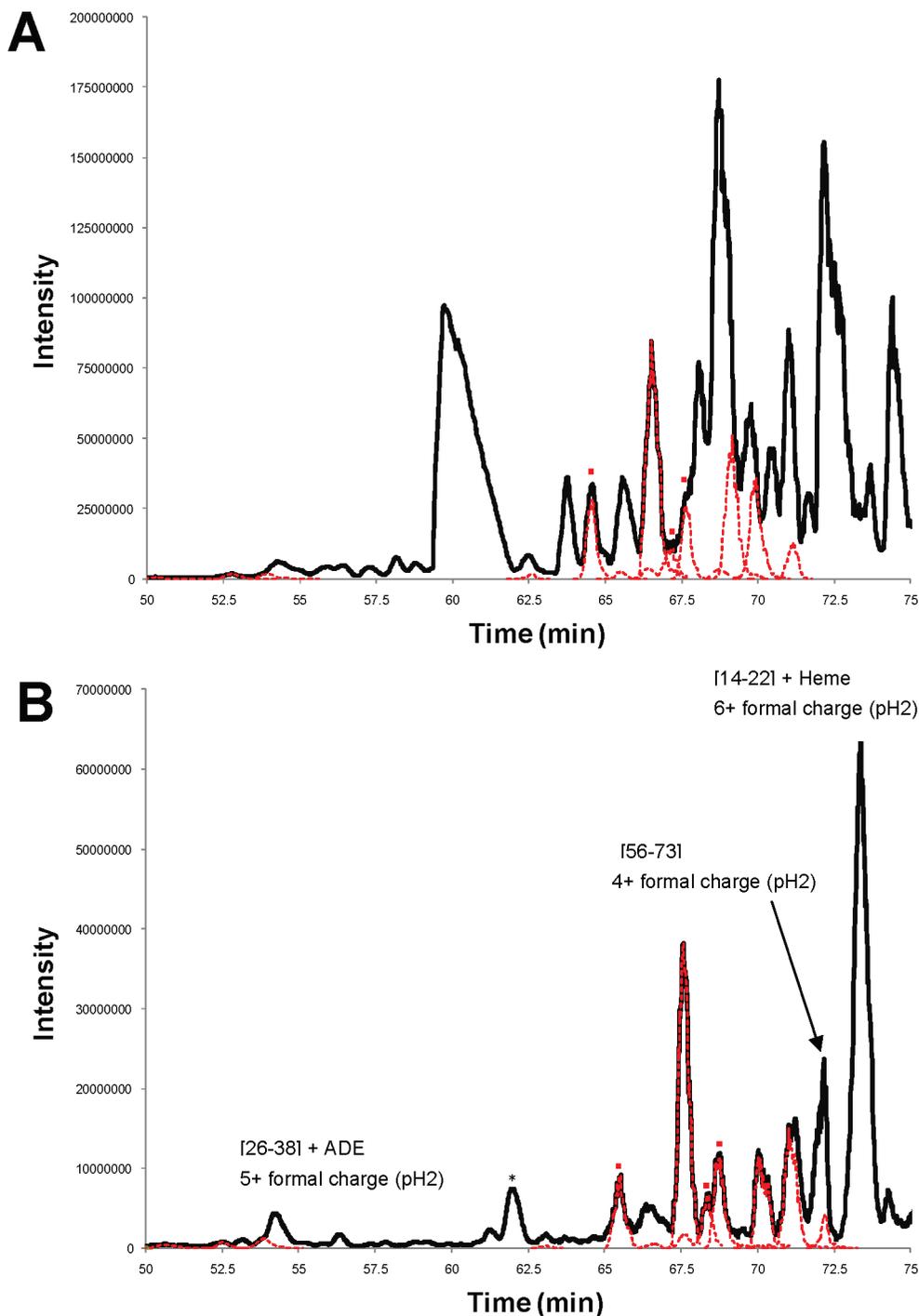


Figure 3. LC-MS of trypsinized, DEST-modified cytochrome c (A) before and (B) after SCX enrichment. Base peak chromatograms (black) are shown along with reconstructed selected ion chromatograms of interpeptide cross-links (red). Three interpeptide cross-links, marked with (■), were identified manually based on similarity to xQuest matches.⁴³ Two of these involve truncated versions of tryptic peptides,⁴³ and one is an interpeptide cross-link also modified with a dead-end. The asterisk marks a spurious peak due to neurotensin from previous experiments.

exchange chromatography of DEST-modified samples, unlike previous cross-link enrichment techniques that involve biotin or functional group tagged reagents, is thus capable of separating interpeptide cross-links from both unmodified and dead-end modified peptides.

This enrichment technique is quite simple. A tryptic digest of DEST-modified cytochrome c was loaded onto an SCX column and washed with a concentration of sodium chloride capable of displacing unmodified and dead-end modified peptides. We found that 300 mM sodium chloride was most effective at clearing the

sample of low charge species without causing loss of more highly charged interpeptide cross-links. The latter were then displaced from the SCX column onto a reversed phase C18 trapping column using 1000 mM sodium chloride. The enriched sample was subsequently rinsed free of salt, collected by elution with organic mobile phase, and subjected to peptide analysis as previously done for the whole digest.

Base peak chromatograms of the tryptic digest before and after strong cation exchange enrichment are shown in black in Figure 3. The red curve represents a reconstructed interpeptide cross-

link chromatogram. It was generated by identifying all cross-linked peptides from multiple xQuest matches and manually summing their selected ion chromatograms. Each panel displays data acquired from a single LC run. Therefore, the black chromatograms show the total composition of the sample while the red chromatograms highlight the cross-linked portion of the sample. Before enrichment, most cross-links were buried by low charge components of the digest, primarily individual, unlinked peptides, and only two cross-links were detected as resolved chromatographic peaks. With SCX enrichment, however, the complexity of the digest was significantly reduced, and most cross-links were well resolved.

The presence of a few non-cross-linked species in the enriched sample is due to the anomalous primary structure of cytochrome c. For instance, the small peak at 54 min is a dead-end modified peptide that although fully tryptic contains 5 basic residues, due to the presence of two histidines. This is unusual, since histidine has a low rate of occurrence in proteins.⁴⁰ A second notable noncross-linked species that eluted at 72 min is an unmodified tryptic peptide with two missed cleavages and having the sequence GITWKEETLMEYLENPKK. For many proteins, missed cleavages can be minimized by optimizing digestion conditions. Cytochrome c, however, exhibits a primary sequence that is particularly prone to missed cleavages. This peptide, in particular, has two consecutive sites known to impede cleavage; one site is surrounded by acidic residues (that can form salt bridges) and the other falls immediately before another cleavage site.^{41,42} The most dominant non-cross-linked species present in the enriched sample eluted at 73 min. The most prominent feature in the chromatogram, it is a fully tryptic unmodified peptide that coincidentally bears the exact number of positive charges (6) at low pH as interpeptide DEST cross-links. This peptide contains a histidine residue as well as a covalently bound heme group, which under the pH 2 conditions used here has a 3+ formal charge (Supporting Information Figure 5). Obviously, this is related to the relatively unique structure of cytochrome c. Hence, SCX enrichment of interpeptide DEST cross-links appears to be very effective. The reduction of sample complexity demonstrated in this proof of concept will undoubtedly improve the detectability of cross-links in complex mixtures. Furthermore, this pragmatic method does not require tagged reagents or any additional reactions and is capable of separating cross-links from dead-ends, making it easier to implement and potentially more effective than previous enrichment strategies.^{14–17}

(40) Voet, D.; Voet, J. G. In *Biochemistry*, 3rd ed.; J. Wiley & Sons: Hoboken, NJ, 2004; p 67.

(41) Monigatti, F.; Berndt, P. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (1), 13–21.

(42) Yen, C.-Y.; Russell, S.; Mendoza, A. M.; Meyer-Arendt, K.; Sun, S.; Cios, K. J.; Ahn, N. G.; Resing, K. A. *Anal. Chem.* **2006**, *78* (4), 1071–1084.

(43) Alves, P.; Arnold, R. J.; Clemmer, D. E.; Li, Y.; Reilly, J. P.; Sheng, Q.; Tang, H.; Xun, Z.; Zeng, R.; Radivojac, P. *Bioinformatics* **2008**, *24* (1), 102–109.

CONCLUSIONS

We have introduced a novel bifunctional thioimide cross-linking reagent (diethyl suberthioimide, DEST) that effectively modifies proteins under physiological conditions and does so without sacrificing the native basicity of targeted amines. DEST represents the only reagent currently capable of doing so in a single stage reaction at physiological pH. This makes it a compelling alternative to both imidate and succinimidyl ester reagents for structural studies, because it does not require alkaline pH for reactivity and does not perturb the electrostatic properties of a protein. Similar to amidinating labeling reagents, DEST is thus less likely to lead to artifactual conclusions about native protein structure. The additional steric bulk of a cross-linker could introduce small perturbations, but considering the short alkyl chain that connects the reactive groups, this is not expected. The fact that this reagent preserves the basicity of targeted amines upon modification also means that the cross-links it forms can be easily separated from other components of tryptic digests, including dead-end modified peptides, using strong cation exchange chromatography. The use of this novel amidinating cross-linking reagent holds great promise for efficient, large-scale structural analysis of complex systems.

ACKNOWLEDGMENT

We would like to thank Oliver Rinner for his assistance with xQuest, as well as Zaneer Segu and Yehia Mechref for providing access to and assistance with the LTQ-Orbitrap. This research was supported in part by the Indiana METACyt Initiative of Indiana University, funded in part through a major grant from the Lilly Endowment, Inc., and by National Science Foundation Grants CHE-0518234 and CHE-0832651.

SUPPORTING INFORMATION AVAILABLE

Experimental information about peptide cross-linking experiments, size exclusion chromatography of DEST-modified cytochrome c, and xQuest database searching, supplemental results and discussion about peptide cross-linking experiments, scheme portraying the side reaction of imidates below pH 9, scheme for the synthesis of DEST, size exclusion chromatogram of DEST-modified cytochrome c, whole protein mass spectra of DMS-modified cytochrome c, tables of xQuest MS/MS matches for interpeptide cross-links from cytochrome c, figure of the crystal structure of cytochrome c, figure indicating the fragmentation of the amidine bond of a DEST cross-link, and figure demonstrating the protonation of ferric heme as a function of pH. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Received for review June 17, 2010. Accepted August 10, 2010.

AC101586Z