Pergamon

SCIENCE DIRECT®

Bioorganic & Medicinal Chemistry Letters 13 (2003) 4023–4026

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

Isotopically Labeled Crosslinking Reagents: Resolution of Mass Degeneracy in the Identification of Crosslinked Peptides

Christopher J. Collins,[a] Birgit Schilling,[a] Malin Young,[c] Gavin Dollinger[d] and R. Kiplin Guy[a,b,*]

[a]*Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, CA 94143, USA*
[b]*Department of Cellular and Molecular Pharmacology, University of California at San Francisco, San Francisco, CA 94143, USA*
[c]*Biosystems Research Department, Sandia National Laboratories, Livermore, CA 94551-0969, USA*
[d]*Small Molecule Drug Discovery, Chiron Corporation, Emeryville, CA 94608, USA*

Received 2 July 2003; revised 2 July 2003; accepted 26 August 2003

Abstract—Mass spectrometry in three dimensions (MS3D) is a newly developed method for the determination of protein structures involving intramolecular chemical crosslinking of proteins, proteolytic digestion of the resulting adducts, identification of crosslinks by mass spectrometry (MS), peak assignment using theoretical mass lists, and computational reduction of crosslinks to a structure by distance geometry methods. To facilitate the unambiguous identification of crosslinked peptides from proteolytic digestion mixtures of crosslinked proteins by MS, we introduced double ^{18}O isotopic labels into the crosslinking reagent to provide the crosslinked peptides with a characteristic isotope pattern. The presence of doublets separated by 4 Da in the mass spectra of these materials allowed ready discrimination between crosslinked and modified peptides, and uncrosslinked peptides using automated intelligent data acquisition (IDA) of MS/MS data. This should allow ready automation of the method for application to whole expressible proteomes.

We have recently reported a method called mass spectrometry in three dimensions (MS3D) to solve moderate resolution protein structures using a combination of chemical crosslinking, mass spectrometry, and computation.[1] The first step of MS3D is generation of crosslinks between amino acid side chains of the intact protein using a chemospecific crosslinking reagent. After purification of monomers the protein is proteolytically digested and the crude peptide mixture subjected to analysis by LC/MS. The resulting list of masses is compared computationally to a theoretical list of masses of all possible crosslinked peptides. Unambiguously identified crosslinks are used to generate a set of distance constraints that relate the reactive side chains to one another in a three dimensional sense. Finally, these distance constraints are substituted into distance geometry calculations to generate a protein structure of moderate resolution. With the exception of the identification of disulfide bridges,[2] crosslinking studies in conjunction with MS were exceedingly rare[3,4] before we published our initial report.

Frequently, a proteolytic digest of a crosslinked protein includes multiple species-crosslinked peptides, linear peptides, and singly modified peptides-all with masses undistinguishable the desired crosslinked peptides at normal instrument resolution. Although sophisticated MS experiments can unambiguously identify these peaks, our desire to automate MS3D for proteomics studies demanded an operationally simpler method to resolve such degeneracy—the incorporation of isotopic labels into the crosslinking reagent. A similar strategy has been employed for resolution of mass degeneracies in the analysis of protein interactions by crosslinking and MS,[5] and in the identification of intramolecular crosslinks.[6] We report herein that this strategy provides a simple method for discriminating between crosslinked peptides, peptides that have been modified only by addition of the crosslinking reagent (i.e., dead-end crosslinks), and linear, unmodified peptides.

The commercially available bis-carboxylic acid (**1a**, Fig. 1) was exchange-labeled using $H_2{}^{18}O$ at acidic pH to give the ^{18}O-labeled octanedioic acid **1b**–**1e**.[7–9] The resulting tetra-labeled bis-carboxylic acid then afforded a bis ^{18}O labeled bis(sulfosuccinimidyl) suberate (BS^3,

*Corresponding author. Tel.: +1-415-502-7051; fax: +1-415-514-0689; e-mail: rguy@cgl.ucsf.edu

0960-894X/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2003.08.053

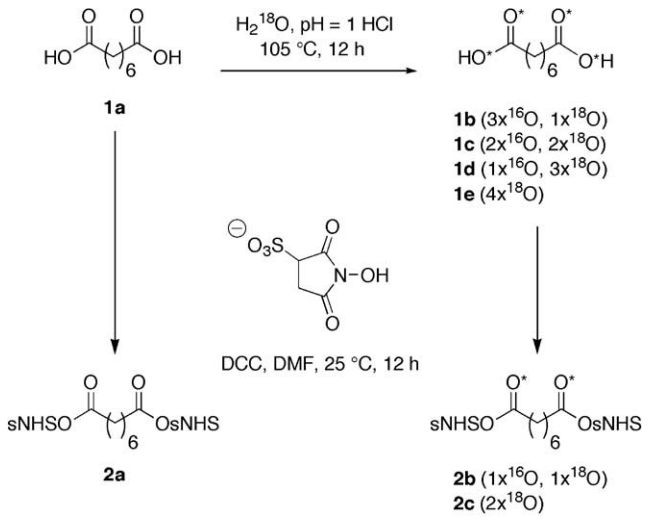

Figure 1. The synthesis of isotopically labeled crosslinking reagents **2a**, **2b**, and **2c**.

2c, Fig. 1) after treatment with sodium 1-hydroxy-2,5-dioxopyrrolidine-3-sulfonate (sNHS) and dicyclohexylcarbodiimide (DCC).[10] This procedure gave a favorable yield and provided, after a single precipitation, homogeneous crosslinking reagent.

Peptides crosslinked with a defined mixture of unlabeled and ^{18}O-labeled reagent (**2a** and **2c**, respectively, Fig. 1) now caused two sets of peaks, separated by 4 atomic mass units (amu). On the other hand, peaks caused by a noncrosslinked peptide exhibited only their normal isotopic distribution. Under the normal crosslinking conditions, dead-end peptides that were labeled but not crosslinked by the reagent also showed a 4-Da shift. These were easily resolved by executing a simultaneous second experiment using unlabeled **2a** in the presence of $H_2{}^{18}O$, which resulted in a 2-amu shift for singly modified peptide or by the base mediated exchange of the pendant carboxylate in water. Together, these methods allowed easy automation of the process of deciding which peaks arose from crosslinked peptides and which did not.

Commercially available pentapeptide thymopentin (RKDVY) was crosslinked with mixtures of varying compositions of **2a** and **2b**/**c** (Fig. 2). Because most peptides do not naturally form dimers, a peptide concentration of as high as 0.67 mM was used to force the model peptide into a crosslinked dimer structure (concentrations of 5–10 μM are usual for crosslinking of proteins). Crosslinking of this peptide could potentially afford three species: peptide dimer **3a** resulting from crosslinking with **2a**, which contains no ^{18}O-label; peptide dimer **3b** resulting from crosslinking with **2b**, which contains one ^{18}O-label; and peptide dimer **3c** resulting from crosslinking with **2c**, which contains two ^{18}O-labels. Reactions using different ratios of reagents **2a** and/or **2b**/**c** were monitored by matrix assisted laser desorption ionization (MALDI) MS as shown in Figure 2. Panel A shows the result when the test substrate was crosslinked with pure **2a**, revealing the normal isotopic

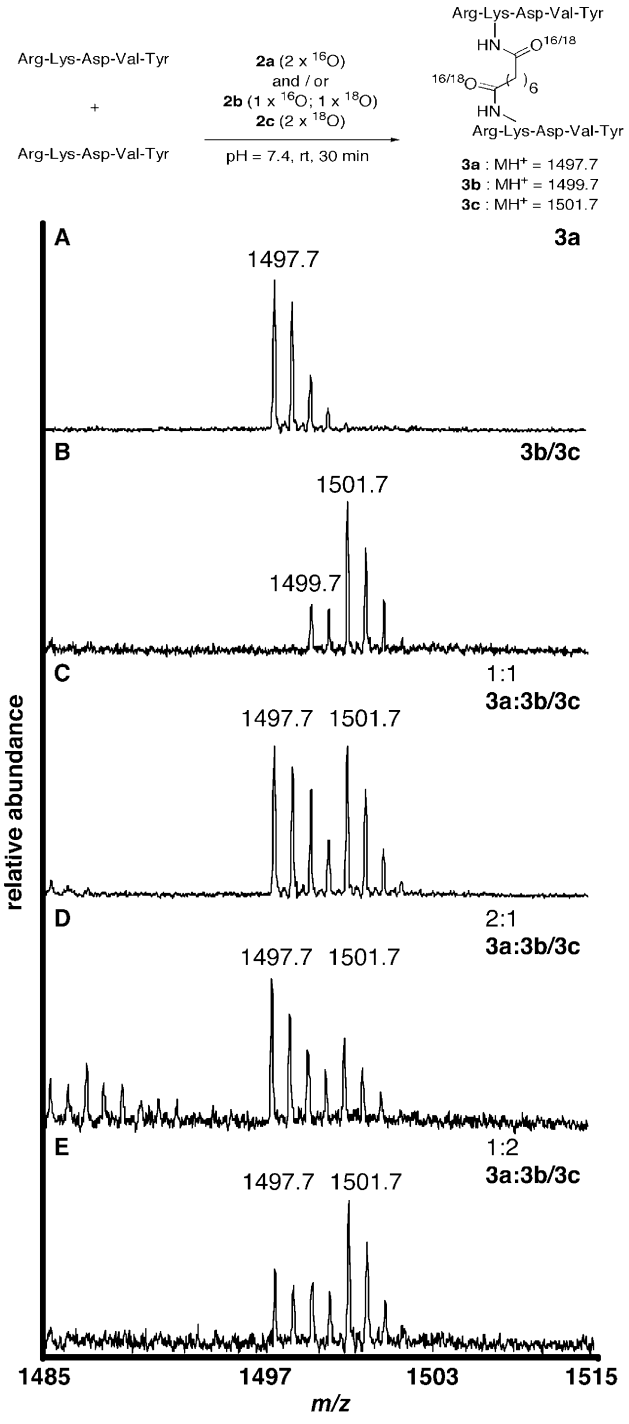

Figure 2. The crosslinking of a model peptide with isotopically labeled crosslinking reagents **2a**, **2b**, and/or **2c** to produce crosslinked peptides **3a**, **3b**, and **3c**. Lower panels show the positive-ion MALDI-MS spectrum of intermolecularly crosslinked RKDVY. Regular ^{16}O-containing crosslinking reagent **2a**, and ^{18}O-containing crosslinking reagent (defined mixture of **2b** and **2c**) were titrated and reacted with RKDVY yielding mixtures of crosslinked peptides **3a** and **3b**/**3c**. (A) reaction with crosslinker **2a**; (B) reaction with crosslinker **2b**/**2c**; (C) reaction with 1:1 mixture of **2a**:**2b**/**2c**; (D) reaction with 2:1 mixture of **2a**:**2b**/**2c**; and (E) reaction with 1:2 mixture of **2a**:**2b**/**2c**.

distribution caused by the natural abundance of carbon and nitrogen isotopes in the peptide. Panel B shows the results using unadulterated reagent **2b**/**c**, revealing the isotopic purity of this batch of **2b**/**c**. The isotopic dis-

tribution of the molecular ion of peptide **3b**/**c** was simulated with an isotope calculation program and an incorporation of 87.5% ^{18}O-label was determined. In addition, this experiment demonstrated that there is no back-exchange of ^{18}O-label to allow formation of **2a** under the reaction conditions. Panels C, D, and E show the results of crosslinking experiments using mixtures of **2a** and **2b**/**c** to give **2a** to **2c** ratios of 1:1, 1:2, and 2:1, respectively. The results of these experiments clearly show that the 4-Da shift provided by two ^{18}O-labels is sufficient to resolve doublets from peptides of this mass. Even for crosslinked peptides of higher masses, up to 2500 Da (data not shown), the isotope distribution is still distinctive and easily recognized. The readily obtained mixture of **2b**/**c** functions well for efficient labeling, and the 1:1 ratio of **2a** to **2c** provides a good relative signal without using too much of the ^{18}O-containing reagent.

Given the encouraging results obtained in the model study, the same reagent system was applied to the identification of crosslinked peptides of FGF-2. The crosslinking reaction conditions chosen for FGF-2 were similar to those described in our previous report,[1] with substitution of the mixture of crosslinking reagents **2a** and **2b**/**c** for BS3. Subsequent proteolytic digestion with trypsin was followed by HPLC fractionation and ESI-TOF-MS analysis. The observed crosslinked peptides were assigned to particular sequences using ASAP and the resulting assignments are listed in Table 1. In general, we observed good coverage of the expected peptides as observed in the previous study, regenerating 14 of 15 constraints produced in that study. We observed an effective signal to noise ratio enhancement of roughly tenfold due to the ability to use the isotope pattern to drive intelligent data acquisition of MS/MS data on the fly. This represents a significant technical advantage.

The constraints generated in this study were mapped to the known structure of hFGF-2 (Fig. 3). They show very good agreement with both the structure of the protein and prior work. Thus, all of the data needed to produce an accurate moderate resolution structure of the protein (as shown previously)[1] can be gathered in a single experiment with full automation of data acquisition and handling when using this reagent system.

Table 1. Crosslinked peptides identified from the digestion of FGF-2 crosslinked with **2a**, **2b**, and **2c**

Mass/Charge Ratio (m/z)	Intensity (counts)	Charge	Assignment	Error (ppm)
570.3253	94	5	121–146	14.99
583.8175[a]	128	6	1–22,73–81	37.21
593.337	101	6	111–120,126–146	10.57
597.8327[a]	116	4	111–129	11.44
600.3428	59	4	19–22,130–146	12.79
608.8216	177	5	121–129,45–60	152.81
619.8739	75	5	136–146,1–18	81.74
675.8824	94	4	40–52,73–81	44.25
695.4018[a]	113	4	121–129,130–145	18.95
695.9004[a]	136	5	78–97,126–135	67.88
765.9052[a]	107	4	47–60,110–119	8.79
844.4689	142	2	121–135	16.48
845.4672[a]	444	3	110–120,121–129 or 110–119,120–129	0.30
936.5293	125	3	19–22,34–52	28.75
937.531	359	m		

[a]First C13 isotope peak; m, multiple.

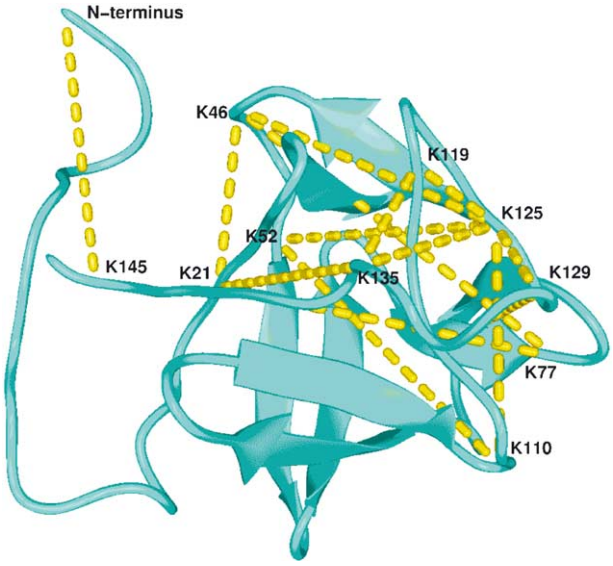

Figure 3. Crosslinks observed after treating recombinant hFGF-2 with a 1:1 mixture of crosslinking reagents **2a** and **2b**/**2c**. A single experiment run with intelligent data acquisition led to this constraint set, which previously required multiple experiments and secondary MS analysis.

The use of the reagent system described above represents a general strategy to introduce stable isotope labels into crosslinked peptides. Readily available ^{18}O-labeled crosslinking reagents provide a powerful tool for easy recognition of crosslinking events within a protein and its proteolytic digestion mixture. Characteristic isotopic distribution for molecular ion peaks of crosslinked peptides allows easy screening, which can be automated with appropriate software, a process that will be essential for high throughput experiments. Moreover, simple experimental modifications provide a means to differentiate between dead-end, or singly modified peptides from the desired crosslinked peptides. Another major advantage of introducing the stable isotope label is that the crosslinked peptides can easily be identified by MS without taking into account the specificities of the proteases used. Therefore, this method is compatible with the use of less specific proteases or combinations of proteases, which in turn would yield crosslinked peptides that are more likely to have masses <2500 Da, where the 4-Da mass shift in the molecular ion isotope pattern can be readily distinguished. Without the isotope tag, fewer constraints for the peptide search engines would allow too many possibilities and identification would be highly unlikely. In addition, ^{18}O-labeled peptides do not show any isotope effect during LC/MS, which provides an advantage over heavily deuterated stable isotope labels that can cause slightly different retention times. As we have shown, a mass shift of 4 Da is sufficient to easily distinguish the isotope distribution of isotopically coded peptides, but is also small enough to allow the selection of the entire

molecular ion isotope cluster for ESI-MS/MS studies. Therefore, this isotope tagging strategy not only facilitates the recognition of crosslinking events and offers an easy tool for automation of data analysis, it also helps to interpret fragmentation studies and confirm crosslinking assignments.

Acknowledgements

The Sandler Research Foundation, NSF (CHE-0118481), and UC BIOSTAR grant (S98-63, UCSF and Chiron Corporation) supported this project and NIH Postdoctoral Training Grant T35 CA 09270 provided stipend support for Dr. Collins.

References and Notes

1. 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*, 5802.
2. Bean, M. F.; Carr, S. A. *Anal. Biochem.* **1992**, *201*, 216.
3. Lacroix, M.; Rossi, V.; Gaboriaud, C.; Chevallier, S.; Jaquinod, M.; Thielens, N. M.; Gagnon, J.; Arlaud, G. J. *Biochemistry* **1997**, *36*, 6270.
4. Rossi, V.; Gaboriaud, C.; Lacroix, M.; Ulrich, J.; Fontecilla-Camps, J. C.; Gagnon, J.; Arlaud, G. J. *Biochemistry* **1995**, *34*, 7311.
5. Muller, D. R.; Schindler, P.; Towbin, H.; Wirth, U.; Voshol, H.; Hoving, S.; Steinmetz, M. O. *Anal. Chem.* **2001**, *73*, 1927.
6. Pearson, K. M.; Pannell, L. K.; Fales, H. M. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 149.
7. Steinschneider, A. B. M. I.; Buku, A.; Fiat, D. *Int. J. Peptide Res.* **1981**, *18*, 324.
8. Ponnusamy, E.; Jones, C. R.; Fiat, D. *J. Lab. Comp. Radiopharm.* **1986**, *26*, 773.
9. Gleispach, H.; Moser, R.; Mayer, B.; Esterbauer, H.; Skriletz, U.; Zierman, L.; Leis, H. J. *J. Chromatogr.* **1985**, *344*, 11.
10. Grabarek, Z.; Gergely, J. *Anal. Biochemistry* **1990**, *185*, 131.