Impact of crosslinker chemistry on peptide fragmentation spectra of crosslinked peptides

Randy J. Arnold\textsuperscript{1}; Suraj Saraswat\textsuperscript{1}; Chao Ji\textsuperscript{2}; Haixu Tang\textsuperscript{2}; Predrag Radivojac\textsuperscript{2}; James P. Reilly\textsuperscript{1}

\textsuperscript{1}Department of Chemistry, \textsuperscript{2}School of Informatics and Computing, Indiana University, Bloomington, IN 47405

**Introduction**

The study of protein-protein interactions is aided greatly by mass spectrometry-based identification of cross-linked peptides. However, cross-linked peptides present several challenges. First, these structures tend to form higher charge state precursors with as many as five or six charges common when modest-sized peptides are cross-linked, producing fragments that can have from one to five charges. Second, with their higher charges, the same normalized collision energy that can induce breakage of a single bond for doubly- or triply-charged peptides can produce multiple fragmentation events in the same cross-linked species, producing double-fragmentation product ions. Third, the structure of the cross-linker itself can have an impact on the fragmentation of the cross-linked peptide.

**Methods**

In our efforts to understand, and ultimately predict with reasonable accuracy, the fragmentation of cross-linked peptides, this work investigates the influence of two different lysine-to-lysine crosslinkers: DEST and BS3. DEST, or diethyl suberthimicidate, was developed in our lab\textsuperscript{1} and incorporates two easily ionized amidinated groups into the cross-linked peptide. BS3 (Bis[4-(N-succinimidyl)-butyl]ether) is a commercially- available crosslinker with two N-hydroxysulfosuccinimid esters (BS3) that produce two amide groups upon reaction.

**Figure 1** (right). MS/MS spectra for the peptides LIVVEKFSVEAPK / LLAQKLK crosslinked with A) DEST and B) BS3 and fragmented by CID of the 3\textsuperscript{+} charge state (precursor m/z 803.44 and 804.16 respectively). Strong peaks that can be assigned to the sequence are labeled with the assignment and weaker peaks that can be assigned are labeled with an asterisk (*). Fragmentation of peptide LIVVEKFSVEAPK is designated as alpha (\(\alpha\)) while fragmentation of peptide LLAQKLK is designated as beta (\(\beta\)). Notice that all labeled peaks correspond to fragmentation of peptide LIVVEKFSVEAPK.

**Figure 2** (below). Ratio of fragment ion intensity from the two peptides in a DEST-crosslinked peptide MS/MS spectrum when the smaller intensity peptide is divided by the larger intensity peptide for A) all 98 peptides identified from \textit{E. coli} ribosomal proteins, B) 35 peptides with difference in length less than 3, C) 35 peptides with difference in length between 3 and 7, and D) 28 peptides with difference in length greater than 7.

**Results and Observations**

DEST and BS3 produce chemically different crosslinks that fragment in unique ways, as shown in Figure 1. Also, both Figures 1 and 2 support the observation that cross-linked peptides tend to preferentially fragment on one, but not the other, peptide backbone. In addition to the lysine-to-lysine crosslinkers shown here, we have also utilized synthetic peptide libraries containing cysteine to create disulfide crosslinked peptide fragmentation. These spectra demonstrate similar features, including a preference for fragmentation of just one of the two peptides. This result is important in developing algorithms to identify MSMS fragmentation spectra of crosslinked peptides.

**Future Directions**

- MS/MS spectra of peptides can be predicted by machine learning models – a similar strategy that incorporates prior knowledge (fragmentation pattern of isolated peptides) has great potential for crosslinked peptide fragmentation spectra.
- Tendency of a peptide bond to fragment and amount of charge remaining on a fragment can be modeled using a non-linear regression model.
- Prediction of crosslinked peptide fragmentation spectra should facilitate more robust identification of these species in biological data.

**Acknowledgements**

The authors acknowledge Matt Lauber for sharing data. Funding was provided by NIH R01 RR024236-01A1 to PR.

**References**