Prediction of the Tandem Mass Spectra of Cross-linked Peptides

Randy J. Arnold[1] ; Xin He[2] ; Emily Renzi[1] ; Predrag Radivojac[2] ; James P. Reilly[1] ; Haixu Tang[2]

[1]Department of Chemistry, [2]School of Informatics and Computing, Indiana University, Bloomington, IN 47405

Introduction

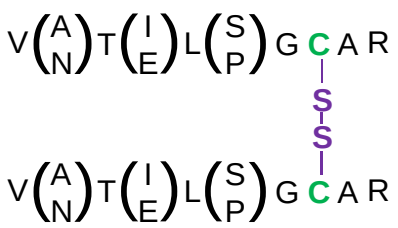
- Protein-protein interactions are involved in many biological processes, including both normal and abnormal growth.
- Mass spectrometry combined with covalent cross-likers has the potential to study many of these interactions as well as help to characterize protein structure.
- Both native (disulfide bond) and non-native (BS3, DEST, etc.) cross-linkers target specific amino acid side chains and have the potential to provide a wealth of biological information.
- Current understanding of how crosslinked peptides fragment in tandem mass spectrometry is limited.
- We propose to create libraries of crosslinked peptides that can be analyzed by standard proteomics methodology.
- Analysis of these samples can potentially improve our understanding of crosslinked peptide fragmentation and increase the accuracy of their identification.

Methods

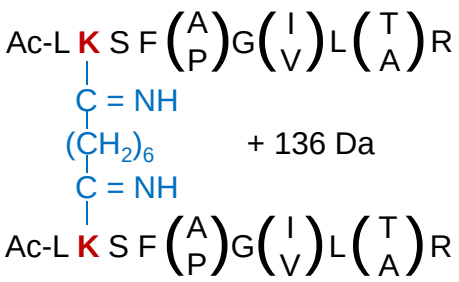
- Peptide libraries are synthesized using standard Boc-protected amino acids and resin-based split-and-mix combinatorial synthesis[2]
- Small 8-sequence libraries have three different positions that each contains one of two different amino acid residues
- Within the error of our ability to split the sample during synthesis, the two different residues at a given position are in equimolar abundance in the library
- When reacted to form cross-linked species, each 8-sequence library will produce 36 unique crosslinked peptides.
- Disulfide and DEST chemistries are illustrated here, but others are readily incorporated.

Crosslinked Peptides from Peptide Libraries

Disulfide crosslinker at C



DEST crosslinker at K[1]

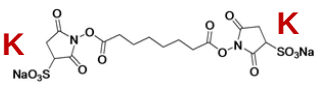


One library = 36 crosslinked peptides

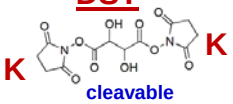
8 unique sequences, lettered A-H: AA, BB, CC, DD, EE, FF, GG, HH, AB, AC, AD, AE, AF, AG, AH, BC, BD, BE, BF, BG, BH, CD, CE, CF, CG, CH, DE, DF, DG, DH, EF, EG, EH, FG, FH, GH [CL = (US^2 + US)/2]

Other crosslinkers[2]

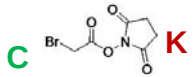
BS3



DST



SBA



Opportunities / Future Directions

- Synthetic peptide libraries allow for systematic study of various parameters of crosslinked peptide fragmentation, including sequence length, position of crosslinking, and crosslinking chemistry.
- Any crosslinking chemistry can be used, since peptides can be dissolved in any solvent and appropriate residues can be incorporated at any location.
- 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 advice and data. Funding was provided by NIH R01 RR024236-01A1 to PR.

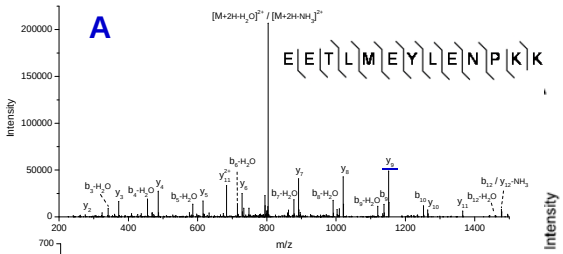
References

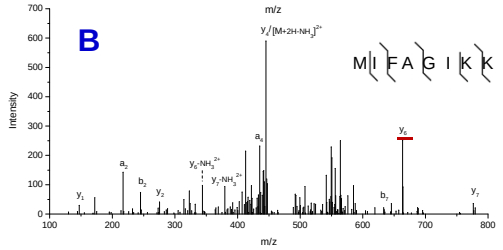
1. Lauber, M. A.; Reilly, J. P. *Anal. Chem.* **2010**, *82*, 7736-7743.
2. Bohrer, et al. *Anal. Chem.* **2010**, *82*, 6559-6568.
3. http://www.proteochem.com/proteincrosslinkers-c-1.html

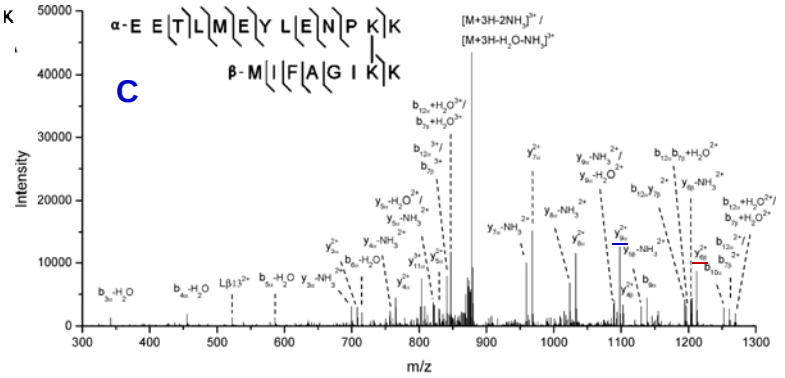
Contact : rarnold@indiana.edu

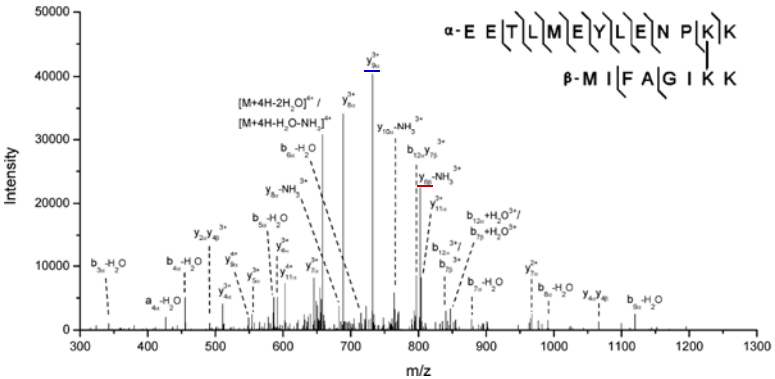
Example of DEST crosslinker [1]

MS/MS spectra for **A**) peptide EETLMEYLENPKK^{2+}, **B**) peptide MIFAGIKK^{2+}, **C**) DEST crosslinked EETLMEYLENPK*K/MIFAGIK*K^{3+}, and **D**) crosslinked EETLMEYLENPK*K/MIFAGIK*K^{4+}. Peptides were crosslinked in native cytochrome C and released upon digestion with trypsin. Note that prominent fragments y_9 from EETLMEYLENPKK and y_6 from MIFAGIKK also occur in the crosslinked peptide and that the charge states depend on the precursor charge state. K* indicates the location of the DEST crosslinking.









DEST = diethyl suberthioimidate