Focus Issue on Peptide Fragmentation

The use of tandem mass spectrometry (MS/MS) to sequence peptides is now widespread, and central to the exploding fields of proteomics and systems biology. A plethora of sophisticated experimental strategies have been devised that combine various separation techniques with a wide variety of MS instrumentation to ionize, fragment and detect peptides. Automated MS/MS experiments can run 24/7 and produce enormous sets of raw data that must be processed by various sequencing software packages to derive primary peptide and protein sequence. In the last decade, numerous strategies were introduced that have turned these (originally) qualitative techniques quantitative and, therefore, established a convenient tool to characterize gene expression at the protein level.

The basic information unit in large-scale protein identification and quantification experiments is the peptide tandem mass spectrum (product-ion spectrum). Ion peaks must be assigned unambiguously for any robust biological information to be deciphered from MS/MS datasets. The sequencing software used to this end is based on bioinformatics modules and fragmentation models. The former generate and rank peptide sequences, create protein lists from peptide lists, and determine protein abundance ratios in quantitative proteomics experiments. The most widespread strategy to generate candidate peptide sequences uses searches of protein sequence databases. This technique can be combined with, or substituted by, *de novo* methods that attempt to sequence peptides without the use of sequence databases. Based on fragmentation models, *in silico* product-ion spectra (i.e., spectra generated by a computer) are created for candidate peptides and these theoretical spectra are compared to the experimentally measured fragmentation patterns. The similarity between the observed and theoretical data is used then to rank sequences, applying various scoring functions. This approach requires that the quality of the fragmentation model (and, thus, the theoretical spectrum used), directly effects the likelihood of a match being found. Consequently, the development of efficient fragmentation models is an active area of research at present. Achieving the aforementioned goal requires detailed understanding of the complex gas-phase dissociation chemistry of peptides.

Current peptide sequencing software works with a considerable error-rate, which leads to an undesirable ambiguity in MS-based protein identification and quantification. It is now widely accepted in the sequencing community that this is partly due to the oversimplified fragmentation models and related scoring functions. At present, most sequencing software routines consider only the most basic fragment ion series (for example *b* and *y* ions in collision-induced-dissociation (CID) and *c* and *z* ions in electron-capture/transfer-dissociation (EXD)) and poorly approximate or completely neglect fragment ion abundances. Conversely, numerous studies demonstrate that peptides fragment to produce complex dissociation patterns and fragment ion abundances (under the same/similar fragmentation conditions) can be reproduced to a significant extent. The application of superior fragmentation models that consider the diversity of peptide fragmentation pathways and (at least) relative fragment ion abundances should greatly improve the reliability of MS/MS-based peptide sequencing.

This *JASMS* Focus Issue on Peptide Fragmentation aims to advance our knowledge of the structure and reactivity of gas-phase peptide ions and their fragments. This focus issue contains twelve articles contributed by some of the leading laboratories in the field, providing representation of the major experimental, statistical and theoretical strategies currently used to study fragmentation of peptides. These papers report both experimental and theoretical approaches, in most of the cases combined to derive information in a synergistic manner. The applied experimental methodologies include site-specific labeling, H/D-exchange, multiple stage MS experiments, ion-mobility spectrometry and synthesis of model peptides with sequences designed specifically to probe particular structures and fragmentation chemistries. The applied computational strategies range from molecular dynamics to high level quantum chemical calculations. The chemistries of both even and odd electron peptide ions are probed, shedding light on structures and mechanisms involved in CID and EXD. The statistical studies analyze mass differences between modified and unmodified peptides, the fragmentation behavior of tryptic peptides and the information content of product-ion spectra in terms of *de novo* sequencing.

The fragmentation of peptide ions involves three major phases: pre-cleavage, amide bond cleavage, and post-cleavage processes. Pre-cleavage scrambling of amide hydrogens of anionic peptides is studied by Jørgensen and co-workers using an elegant experimental design based on region-selective deuteration of a model peptide and MS/MS experiments. Their results clearly indicate that, similar to protonated peptides, amide hydrogen scrambling is also prevalent for deprotonated peptides under CID. Pre-cleavage hydrogen atom migration plays an important role in the fragmentation of odd electron arginine-containing peptide ions as shown by Tureček and co-workers. These authors use collisions with Cs atoms to prepare radicals from singly and doubly charged dipeptides. The related fragmentation pathways and product ions are charac-

© 2008 Published by Elsevier Inc. on behalf of American Society for Mass Spectrometry.
1044-0305/08/$32.00
doi:10.1016/j.jasms.2008.10.009

terized by detailed computation, shedding light onto the chemistry underlying EXD.

The CID fragmentation pathways specific to β-Ala are used by O'Hair and co-workers to distinguish between isomeric dipeptides containing α- and/or β-alanine residues. Zubarev and co-workers report the fragmentation characteristics of doubly protonated tryptic peptides via statistical analysis of y ion intensities in a large database of validated peptide spectra. Principal component analysis indicates these spectra fall into two classes that are characterized by the relative abundance of the peptide bond cleavage after the first two N-terminal residues. The authors propose that peptides with this dominating amide bond cleavage form the diketopiperazine isomer of b_2 ion rather than the oxazolone. A further observation is that a_3 peaks are rarely recorded in the CID spectra of doubly protonated peptides. This phenomenon is also the subject of Glish, Paizs and co-workers in an independent study that investigates singly protonated peptides. These authors provide detailed theoretical calculations on the $b_4 \rightarrow a_4 \rightarrow b_3 \rightarrow a_3 \rightarrow b_2 \rightarrow a_2$ reaction cascade and show that the a_3 ion is indeed kinetically unstable.

The processes that consider reactions of b and a fragments belong to the post-cleavage phase of peptide fragmentation. It is worth noting that eight of the twelve focus papers discuss the post-cleavage phase to some extent, demonstrating the increased research effort being focused on the related chemistry. These studies will inevitably take advantage of the experimental design proposed by Árpád Somogyi that combines peptide fragmentation and H/D exchange experiments to derive structural information on fragment ions. Utilizing this approach, all fragment ions formed by CID or EXD simultaneously undergo HDX under the same experimental conditions. This method is often able to distinguish between isomeric species with linear and cyclic structures.

The chemistry related to these fragment ion structures is of critical importance for understanding the various scrambling pathways of N-terminal b and a CID fragments. Alex Harrison demonstrates that the CID spectra of the b_5 ions with the nominal $AAAAY_{oxa}$, $AAAYA_{oxa}$, $AAYAA_{oxa}$, $AYAAA_{oxa}$, and $YAAAA_{oxa}$ compositions are essentially identical. This observation strongly supports cyclization of the linear oxazolone-terminated species to form macro-cyclic structures that reopen at various amide bonds to give rise to scrambling of the primary peptide sequence. Gaskell and co-workers apply stable isotope labeling and ion mobility spectrometry with MS/MS to investigate the scrambling pathways of b and a ions. These experiments indicate the co-existence of multiple b and a ions isomers and suggest that cyclic structures play an important role in the underlying chemistry. These authors demonstrate that macro-cyclic structures exist for large b ions consisting of eight or nine amino acid residues. Van Stipdonk, Paizs and co-workers study the structure and reactivity of a and a^* fragments using stable isotope labeling, multiple-stage MS, and modeling. These investigations show that the $a \rightarrow a^*$ pathway involves post-cleavage proton-bound dimer intermediates which re-associate after reorganization. The isotope labeling data indicate that [a^*- CO] fragments also undergo scrambling, and the authors characterize the related chemistry by modeling. The important role of post-reaction proton-bound dimers in peptide fragmentation is further emphasized in the article by Siu and co-workers who report the dissociation chemistry of radical peptide ions and explain relative z and c ion abundance ratios based on the related fragment proton affinities.

The Focus Issue is completed by two papers that assess the information content of peptide MS and product-ion spectra. Spengler and co-workers show that a large fraction of singly protonated unmodified and singly phosphorylated peptides can be distinguished directly by their masses if measured by better than $\pm$ 0.1 ppm accuracy. This opens a new way of thinking in data-dependent analysis of peptide mixtures where the accurate precursor mass is used to determine whether further MS/MS sequencing is necessary. Mann and co-workers studied a large, high accuracy MS/MS spectral dataset to determine how much peptide sequence information is present in spectra generated in linear ion trap instruments. They conclude that at least 50% of their spectra contain enough sequence information to allow unambiguous *de novo* sequencing of the fragmenting peptides. Unfortunately no such *de novo* software that could unambiguously extract these peptide sequences currently exists.

The editors of this focus issue propose close collaboration between researchers carrying out experimental, statistical and theoretical peptide fragmentation studies and those developing sequencing software to overcome this computational bottleneck in proteomics (this sentiment echoes that of Mann and co-workers in their contribution). It is our sincere hope that this collection of articles will help stimulate subsequent discussion and cooperation between these fields leading to more synergistic approaches and exciting developments in the coming years within our field.

We thank the authors of this focus issue for their enthusiastic participation. We are also very grateful to Michael Gross for initializing the focus and co-editing the above papers. This work was kindly assisted by Joyce Neff and her efforts are also greatly appreciated.

Béla Paizs
Guest Editor
German Cancer Research Center
Germany

Michael Van Stipdonk
Guest Editor
Wichita State University
USA