Journal of
Proteomics & Bioinformatics
ISSN: 0974-276X

Luo and Ranish, J Proteomics Bioinform 2011, 4:12
DOI: 10.4172/jpb.100000e8

Editorial Open Access

Crosslinkomics---A New Era of Mapping Protein-Protein Interactions

Jie Luo* **and Jeff Ranish**

Institute for Systems Biology, 401 Terry Ave N, Seattle, WA 98109, USA

A major goal of proteomics is to figure out how proteins interact with one another through networks and within complexes. The development of integrated chemical crosslinking, Mass spectrometric and computation approaches (CXMS) have recently emerged as powerful technologies to study protein interactions and have the potential to significantly advance the field of protein-interaction mapping. The identified crosslinked peptides (two peptides linked by specific crosslinkers), can be used to infer sites of protein-protein interactions and put distance constraints on interacting sites based on the properties of the crosslinkers. Recent reports on the on mapping of interactions within large assemblages such the 15 subunit RNA polymerase II-TFIIF complex [*EMBO J 29*, 717-726] , the 53 subunit Ribosome[*JPR*, 10(8):3604-3616] and a partially purified prepartion of the 12 subunit RNA polymerase II [*MCP*, mcp.M111.008318] suggest that the field of CXMS has matured to the point where its use for mapping protein interactions in large compelxes will soon become more routine. CROSSLINKOMICS, which uses mass spectrometry and computational approaches to identify chemically crosslinked peptides in a high throughput and unbiased manner, is emerging as a new and promising OMICS which has the potential to revolutionize the study of protein-protein interactions.

For the past decade, however, the CXMS approach has been limited primarily to the analysis of single proteins or small complexes. Three major hurdles have to be conquered for crosslinkomics to be useful for the large scale and routine study of protein-protein interactions, especially for the global mapping of protein-protein interactions. The first hurdle involves the detection of the typically low abundance crosslinked peptides in samples of high complexity. Several approaches have been devised to facilitate detection of peptide crosslinks during MS analysis. They include the use of crosslinking reagents that produce diagnostic fragmentation patterns during collision-induced dissociation (CID), isotope-coded crosslinkers or proteins, isotopic labeling of peptides derived from crosslinking reaction, SCX fractionation and enrichment of crosslinked products via affinity handles. Most of these approaches, however, require relatively large amounts of starting material which is due in part to generally poor crosslinking and enrichment efficiency. The requirement for relatively large amounts of starting material is the major hurdle at present that prevents the widespread usage of this technique for mapping protein-protein interactions.

The second hurdle is how to identify the crosslinked spectra and the corresponding crosslinked peptides. Identification of the crosslinked peptides is a formidable challenge because crosslinked peptides typically generate fragmentation spectra that are very complex and difficult to interpret. The explosion of database search space that occurs when combinations of all possible crosslinked peptide pairs is considered, severely limits the ability of search algorithms to distinguish true positives from false positives. Thus, most search algorithms require construction of sample-specific databases to limit the number of potential peptide pairs to consider during database searching. Current algorithms can handle on the order of 30-50 proteins. One strategy that facilitates confident identification of crosslinked peptides in complex samples is the use of "MS labile" crosslinkers that fragment either by in source decay or by CID (MS2) to transform crosslinked peptides into two modified peptides which can in turn be selected for CID (MS3) by data dependent routines, and identified by search algorithms such as Sequest or Mascot that are commonly used to identify linear peptides. The reduced complexity of the MS3 spectra facilitates confident peptide identification and alleviates the limitation of searching sample-specific databases. Various MS labile crosslinkers have been developed, including linkers that contain an Aspartyl-Prolyl-bond (D-P), various forms of a carbon-sulfur (C-S) bond, a urea moiety, and a Rink moiety. These methods have also encounters issues such as hydrophobicity, crosslinking and/or enrichment efficiency, multiple MS2 fragments and/or MS3 sensitivity. This issue is the main issue most labs work on at present and major progress has been made in identifying the crosslinking peptides.

The third hurdle for crosslinkomics is how to interpret the crosslinking data, which provides information about the vicinity of the crosslinking sites, not necessary the sites of interactions. Computational tools need to be built to integrate various sources of crosslinking data, with other structural approaches to permit generation of refined models of prtoein complexes structure. This hurdle has been least addressed so far, but it will become more and more important as CXMS becomes more routine. Until then, crosslinkomics can become an efficient and powerful approach to study protein-protein interactions. Thus, the furture direction to the realization of mapping protein-protein interactions in test tubes will rely on the progresses from all three areas: crosslinking efficiency and MS sensitivity increasement, the confidence and the efficiency of spectra identification and the structure modeling.

***Corresponding author:** Jie Luo, Institute for Systems Biology, 401 Terry Ave N, Seattle, WA 98109, USA, E-mail: jluo@systemsbiology.org

Received December 12, 2011; **Accepted** December 15, 2011; **Published** December 17, 2011

Citation: Luo J, Ranish J (2011) Crosslinkomics---A New Era of Mapping Protein-Protein Interactions. J Proteomics Bioinform 4: viii-viii. doi:10.4172/jpb.100000e8