Sample Manipulation

In addition to reporter-based labeling of proteins to enable measurements ("additive methods"), subtractive methods of sequential or targeted protein degradation can unintutively impart additional information. Here, the investigator is able to take measurements before and after degradation, and by knowing something about how the degradation occurs (e.g. which amino acids could have fragmented), can infer information that may not be encoded by the reporters themselves.

Chemical Degradation

The best-characterized method of chemical protein degradation is the so-called Edman degradation, in which an isothiocyanate is selectively reacted with the N-terminus of the target protein to form a reactive phenylthiocarbamyl (PTC) intermediate. In the presence of anhydrous acid the amide bond closest to the intermediate is cleaved, releasing the intermediate and revealing the N-terminus of the next amino acid in the sequence. Historically, the reaction was carried out en-masse at the e.g. picomole scale and the cleaved PTC derivatives were collected and analyzed in bulk [1]. This yields the squences of proteins in bulk, but is unsuitable for the kind of single molecule measurements we are interested in here. One approach that *is* consistent with our requirements, proposed in [2], involves taking a single molecule measurement (e.g. via TIRF microscopy, see section **??**) of a sequence of length N (i.e. integrating information from all reporters), followed by Edman degradation and a second single molecule measurement of a sequence of length N-1 and inferring the nature of the cleaved reporter via ratiometic comparision to the original measurement.

The effect of cyanogen bromide on proteins should also be mentioned here, being a means of selective cleavage at methionine residues[3].

Enzymatic Degradation

Many enzymes - such as proteases or aminopeptidases - have evolved to cleave proteins in exquisitely selective fashion, and the additional information imparted via fragmentation can again be harnessed to ease protein sequence analysis. An exhaustive inventory falls outside the scope of this report. However, we note that trypsin, which cleaves c-terminal to lysine and arginine residues [4], and GluC, which cleaves c-terminal to glutamic acid residues [5], are both well characterized tools. Aminopeptidases may also be of interest of a chemical Edman degradation cannot be carried out [6].

References

[1] Laursen, R. A. Solid-phase edman degradation. *European Journal of Biochemistry* **20**, 89–102 (1971).

[2] Swaminathan, J., Boulgakov, A. A. & Marcotte, E. M. A theoretical justification for single molecule peptide sequencing. *PLoS computational biology* **11**, e1004080–e1004080 (2015).

[3] Gross, E. The cyanogen bromide reaction. *Methods in enzymology* **11**, 238–255 (1967).

[4] Chen, E. I., Hewel, J., Felding-Habermann, B. & Yates, J. R. Large scale protein profiling by combination of protein fractionation and multidimensional protein identification technology (mudpit). *Molecular & Cellular Proteomics* **5**, 53–56 (2006).

[5] Drapeau, G. R., Boily, Y. & Houmard, J. Purification and properties of an extracellular protease of staphylococcus aureus. *Journal of Biological Chemistry* **247**, 6720–6726 (1972).

[6] Sträter, N., Sun, L., Kantrowitz, E. & Lipscomb, W. N. A bicarbonate ion as a general base in the mechanism of peptide hydrolysis by dizinc leucine aminopeptidase. *Proceedings of the National Academy of Sciences* **96**, 11151–11155 (1999).