Automatic Detection of Clipped Monoclonal Antibodies from High-Resolution Native Mass Spectrometry

Wilfred Tang[1] Marshall Bern[1] Andrew Nichols[1] K. Ilker Sen[1] Yong J. Kil[1] Eric Carlson[1] Tomislav Caval[2] Vojtech Franc[2] Albert J.R. Heck[2]

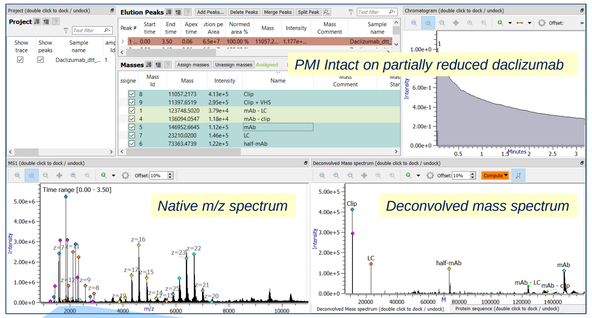
[1]Protein Metrics, San Carlos, CA [2]Biomolecular Mass Spectrometry and Proteomics, Utrecht University Contact: bern@proteinmetrics.com

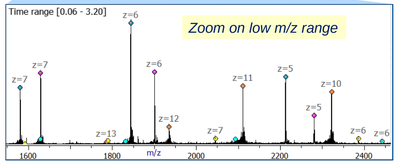
Introduction

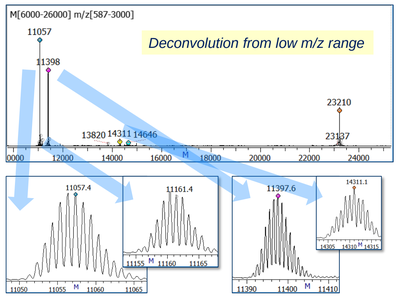
Protein therapeutics are subject to various *in vitro* and *in vivo* degradations, including "clipping", meaning in-solution fragmentation of the peptide backbone. Clipping can affect safety and efficacy and hence is considered a critical quality attribute in the biopharmaceutical industry.

Biopharma companies routinely assay monoclonal antibodies (mAbs) for clipping either by CE-SDS (capillary electrophoresis) or by mass spectrometry, most often intact mass analysis on reduced mAbs. MS offers the advantage that it can identify clipped species; however, manual assignment of mass peaks can be tedious, subjective, and error-prone.

Here we show the type of manual analysis that we aim to automate:







The peak at 11057.4 matches an initial sequence of the heavy chain (HC) up to GG.G

qVQLVQSGAEVKKPGSSVKVSC
KASGYTFTSYRMHWVRQAPGQG
LEWIGYINPSTGYTEYNQKFKD
KATITADESTNTAYMELSSLRS
EDTAVYYCAR**GG.G**VFDYWGQG
TTLTVSSGPSVFPLAPSSKST**S
G.G**…

which has computed average mass of 11057.3 and monoisotopic mass 11050.32. (Note that N-terminal clips weigh 18 Da more than b-ions.) The peaks at 11161.4 and 11397.6 match N-terminal extensions of S and VHS from signal peptide. The small peak at 14311 Da matches the clip at SG.G, but the mass is 4 Da off possibly due to interference at m/z 1790.

Methods – Experimental

We show data from three commercially available therapeutic mAbs, generous gifts from Genmab (Utrecht, The Netherlands):

- Daclizumab (trade name Zynbryta)
- Obinutuzumab (Gazyva)
- Rituximab (MabThera)

The daclizumab sample was already known to include an abundant clipped form as well as N-terminal extension by part of the signal peptide, as shown in the introduction and also in (Bern et al, 2018). We induced clipping in obinutuzumab and rituximab by exposing them to Cathepsin L, a host cell protease and a known cause of *in vitro* clips. The mAb was diluted to 1 mg / mL in 150 mM ammonium acetate at pH 5.2. Cathepsin L was added in 1:250 ratio, incubated for 2 days at 37°C, and buffer exchanged to 150 mM ammonium acetate pH 7 before native MS.

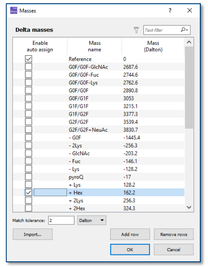
Samples were analyzed on an Exactive Plus Orbitrap EMR (Thermo Fisher Scientific) under native conditions as described previously (Rosati et al, 2014). Source fragmentation and collision energy were varied from 70 to 90 V and resolution was set to either 35,000 or 70,000.

Methods – Computational

We enhanced Protein Metrics Intact software by adding text commands to control automated clip matching. Clips are assumed to be either N- or C-terminal sequences of peptide chains. Proteins may consist of multiple chains, disulfide-bonded at cysteines, but the software does not consider disulfide-bonding pattern, so it does not distinguish clips that would divide the molecule in two and those that would not. User inputs include:

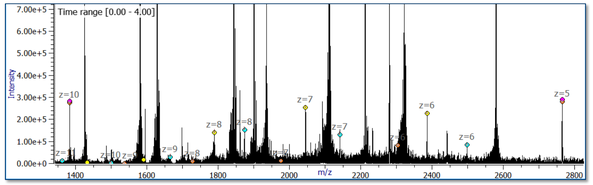
- Potential clips, either all possible cleavages, or only between specified amino acid types such as D.P or G.G, known clipping sites (Vlasek-Ionescu, 2011)
- Amino acid sequences of proteins
- Mass deltas for variable modifications (e.g., 22 for sodium, 1445.4 for G0F)
- Choice of pyro-Glu for N-terminal Glu
- Number of disulfide bonds (default is that all Cys are disulfide-bonded)
- Mass matching tolerance (in either Da or ppm) for calculated masses and observed peaks.

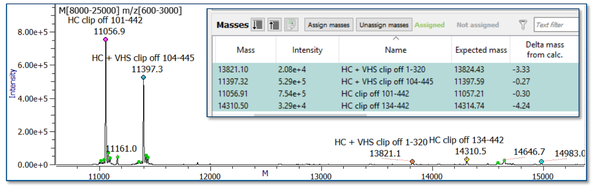
The matching algorithm is a simple greedy algorithm: each peak matches to the closest calculated mass within tolerance.



Screenshot of PMI Intact's mass delta table, showing 0 and 162 Da as selected deltas.

Daclizumab (Zynbrita)





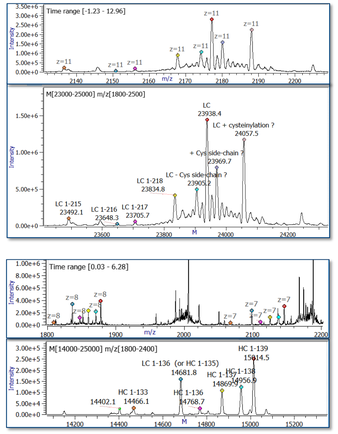
The FASTA file included both HC and HC starting VHS, so the peaks at 11057 and 11397 matched sequences. The peak at 14311 matches with poor mass accuracy (see the Introduction). The peak at 13821 is an artifact caused by slight charge uncertainty for the m/z peaks from 11057 (see co-located orange and magenta dots). The peak at 14983 is real but does not match any N- or C-terminal clip within 6 Da, a fact that would be tedious to determine manually.

Obinutuzumab (Gazyva)

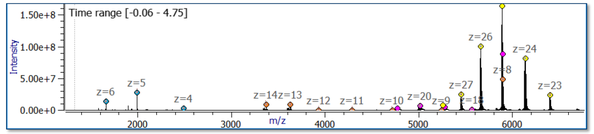
Obinutuzumab deconvolution showed many small peaks around 15, 24, and 48 kDa, which automatic clip matching localized to the start of the HC constant part, the end of the LC, and cleavages between the inter-chain disulfide bonds removing a mAb "arm" (also see Rituximab).

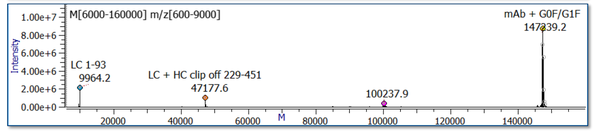
*The 24 kDa peaks show the full LC along with clips before the terminal Cys at sequence **N.R.G.E.C**. There are also peaks suggesting S-S bond fragments.*

*The 15 kDa peaks show clips at the HC constant sequence PLA**P.SS**K.S.T.S.G.G Interestingly, the prominent peak at 14681.8 Da is a better match (0.3 Da) to an LC clip than to the more likely HC clip (0.6 Da).*



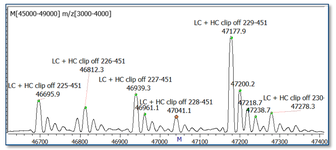
Rituximab (MabThera)





Deconvolution of the full m/z finds two clip regions: CDR L3 and the HC hinge. A focused deconvolution on 45 – 49 kDa shows each clip in the hinge (between the HC-LC Cys and the first HC-HC Cys).

*KS**C.D.K.T.H.T.C**PPCPAPELLG Complementary peaks at ~100 kDa help confirm clip assignments.*



Conclusions and References

We are continuing to improve and refine the automated detection of clipped mAbs. We draw the following conclusions from the work so far:

- Cathepsin L treatment induces a variety of low-abundance clipped mAbs.
- Native MS can assay in-solution clips, as it is not hampered by experimental artifacts from harsh conditions and in-source dissociation.
- Automated clip matching saves time and avoids human errors.
- At the current state of development, matched clips require human validation, due to false positives, which may result from (1) mass coincidences, (2) artifact peaks, (3) wide mass matching tolerance, and (4) large search space (e.g., inclusion of unlikely clips in sequence between S-S-bonded cysteines).

Planned future improvements include built-in knowledge of IgG disulfide bonding patterns, use of monoisotopic rather than average mass for isotope-resolved mass peaks, and statistical modeling of likely clip sites.

References

(Bern et al, 2018) Parsimonious charge deconvolution for native mass spectrometry, J Proteome Res, 2018, PMID: 29376659

(Rosati et al, 2014) Detailed mass analysis of structural heterogeneity in monoclonal antibodies using native mass spectrometry. Nat. Protoc, 2014, PMID: 24675736

(Vlasek-Ionescu, 2011) Fragmentation of monoclonal antibodies, mAbs, 2011, PMID: 21487244

(Kleemann et al, 2008) Characterization of IgG1 immunoglobulins and peptide-Fc fusion proteins by limited proteolysis in conjunction with LC-MS, Anal Chem 2008, PMID: 18293943